

**PALACKÝ UNIVERSITY OLOMOUC**  
**FACULTY OF MEDICINE AND DENTISTRY**



Dissertation thesis

**Novel molecular approaches to ultrasensitive  
biomarker detection in the era of precision  
medicine**

**Anna Petráčková**

Study programme: Medical Immunology

Supervisor: Assoc. prof. Dr. Ing. Eva Kriegová

**Olomouc 2021**

## Statement

I hereby declare that this thesis entitled "Novel molecular approaches to ultrasensitive biomarker detection in the era of precision medicine" was written by me and all relevant resources are included in the reference part. The work was carried out at the Department of Immunology, Faculty of Medicine and Dentistry, Palacký University Olomouc under the supervision of assoc. prof. Dr. Ing. Eva Kriegová.

Olomouc, October 2021

.....

Mgr. Anna Petráčková



## Acknowledgement

I would like to thank my supervisor assoc. prof. Dr. Ing. Eva Kriegová for her leadership, continuous support, trust and great scientific discussions.

My thanks also belong to all my colleagues from the Department of Immunology, Department of Hemato–Oncology, Department of Internal Medicine III – Nephrology, Rheumatology and Endocrinology, Palacký University and University Hospital Olomouc and to bioinformaticians from the Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava.

Finally, I would like to thank my family for their support.

The reasearch was supported by Internal grant agency of Palacký University (IGA\_LF\_2021\_015).

## Bibliografická identifikace

Jméno a příjmení autora: Anna Petráčková  
Název práce: Nové molekulární přístupy k citlivé detekci biomarkerů  
v éře precizní medicíny  
Typ práce: Dizertační  
Pracoviště: Ústav imunologie, Lékařská fakulta Univerzity Pala-  
ckého v Olomouci  
Vedoucí práce: doc. Dr. Ing. Eva Kriegová  
Rok obhajoby práce: 2021  
Klíčová slova: biomarker, molekulární biologie, precizní medicína,  
sekvenování nové generace  
Počet stran: 89  
Jazyk: Anglický

## Bibliographical identification

Author's name and sur- Anna Petráčková  
name:  
Title: Novel molecular approaches to ultrasensitive  
biomarker detection in the era of precision medicine  
Type of thesis: Dissertation  
Department: Department of Immunology, Faculty of Medicine and  
Dentistry, Palacký University Olomouc  
Supervisor: Assoc. Prof. Dr. Ing. Eva Kriegová  
The year of defence: 2021  
Keywords: biomarker, molecular biology, precision medicine,  
next-generation sequencing  
Number of pages: 89  
Language: English

# Abstract

Recent technological advances in molecular and cellular biology have enabled unprecedented characterization of molecular changes underlying pathogenesis of human diseases, thus nominating novel biomarkers. This enables to enter into the era of precision medicine, an emerging approach for disease treatment and prevention that takes into account individual variability in genes, cellular and molecular profiles, environment, and lifestyle for each person. This thesis deals with the research of new biomarkers and the refinement of the analysis of existing biomarkers using novel molecular techniques in haemato-oncological and autoimmune diseases. In particular, it focuses on the analysis of germline and somatic variants that serve as genetic predictive biomarkers detected by next-generation sequencing technologies in chronic lymphocytic leukemia, multiple myeloma and mantle cell lymphoma, and refining their analysis at low variant allelic frequency. Moreover, it deals with serum protein candidate biomarkers of lupus nephritis and organ damage in the autoimmune disease systemic lupus erythematosus and gene expression biomarkers in systemic sclerosis and rheumatoid arthritis. Taken together, our studies have uncovered many candidate biomarkers in the studied diseases that may improve patient outcomes and tailor medical decisions and treatment as requested in the era of precision medicine. This thesis summarizes the results of nine original articles published in journals with impact factor.

## Abstract

Nedávné technologické pokroky v molekulární a buněčné biologii umožnily bezprecedentní charakterizaci molekulárních změn, které jsou základem patogeneze lidských onemocnění, a tím i stanovení nových biomarkerů. To také otevřelo éru precizní medicíny, nového přístupu k léčbě a prevenci nemocí, který zohledňuje u každého člověka individuální variabilitu v genech, buněčných a molekulárních profilech, prostředí a životním stylu. Tato práce se zabývá výzkumem nových biomarkerů a zpřesněním analýzy stávajících biomarkerů pomocí nových molekulárních technik u hematologických a autoimunitních onemocnění. Zejména se pak zabývá analýzou germinálních a somatických mutací, které slouží jako genetické prediktivní biomarkery detekované sekvenačními technologiemi nové generace u chronické lymfocytární leukemie, mnohočetného myelomu a lymfomu z plášťových buněk, a zpřesněním jejich analýzy při nízké alelické frekvenci, dále sérovými proteinovými kandidátními biomarkery lupusové nefritidy a orgánového poškození u autoimunitního onemocnění systémový lupus erythematosus a genovými expresními biomarkery u revmatoidní artritidy a systémové sklerodermie. Naše studie odhalily mnoho kandidátních biomarkerů u studovaných onemocnění, které mohou zlepšit péči o pacienty a přizpůsobit léčebné postupy dle požadavků precizní medicíny. Práce sumarizuje výsledky devíti původních prací publikovaných v časopisech s impakt faktorem.

# List of publications related to the thesis

## Original articles published in impact factor journals:

(Listed by first authorship and publication date.)

### APPENDIX A

**Petrackova A**, Savara J, Turcsanyi P, Gajdos P, Papajik T, Kriegova E. Rare germline pathogenic *ATM* variants are underestimated in cancer. – manuscript under review

### APPENDIX B

**Petrackova A**, Minarik J, Sedlarikova L, Libigerova T, Hamplova A, Krhovska P, Balcarkova J, Pika T, Papajik T, Kriegova E. Diagnostic deep-targeted next-generation sequencing assessment of *TP53* gene mutations in multiple myeloma from the whole bone marrow. *Br J Haematol.* 2020;189(4):e122-e125. (IF 2020: 6.998; Q1)

### APPENDIX C

**Petrackova A**, Horak P, Radvansky M, Fillerova R, Smotkova Kraiczova V, Kudelka M, Mrazek F, Skacelova M, Smrzova A, Kriegova E. Revealed heterogeneity in rheumatoid arthritis based on multivariate innate signature analysis. *Clin Exp Rheumatol.* 2020;38:289-298. (IF 2020: 4.473)

*Dean's award for the best student scientific publications in 2020*

### APPENDIX D

**Petrackova A**, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, Papajik T, Kriegova E. Standardization of sequencing coverage depth in NGS: recommendation for detection of clonal and subclonal mutations in cancer diagnostics. *Front Oncol.* 2019;9:851. (IF 2019: 4.848)

*Dean's award for the best student scientific publications in 2019*

### APPENDIX E

**Petrackova A**, Horak P, Radvansky M, Skacelova M, Fillerova R, Kudelka M, Smrzova A, Mrazek F, Kriegova E. Cross-disease innate gene signature: emerging diversity and abundance in RA comparing to SLE and SSc. *J Immunol Res.* 2019;3575803. (IF 2019: 3.327)

### APPENDIX F

**Petrackova A**, Smrzova A, Gajdos P, Schubertova M, Schneiderova P, Kromer P,

Snasel V, Skacelova M, Mrazek F, Zadrazil J, Horak P, Kriegova E. Serum protein pattern associated with organ damage and lupus nephritis in systemic lupus erythematosus revealed by PEA immunoassay. *Clinical Proteomics*. 2017;14:32. (IF 2017 3.516)

*Dean's award for the best student scientific publications in 2017*

#### APPENDIX G

Kriegova E, Fillerova R, Minarik J, Savara J, Manakova J, **Petrackova A**, Dihel M, Balcarkova J, Krhovska P, Pika T, Gajdos P, Behalek M, Vasinek M, Papajik T. Whole-genome optical mapping of bone-marrow myeloma cells reveals association of extramedullary multiple myeloma with chromosome 1 abnormalities. *Sci Rep*. 2021;11:14671. (IF 2020: 4.379; Q1)

#### APPENDIX H

Malarikova D, Berkova A, Obr A, Blahovcova P, Svaton M, Forsterova K, Kriegova E, Prihodova E, Pavlistova L, **Petrackova A**, Zemanova Z, Trneny M, Klener P. Concurrent TP53 and CDKN2A gene aberrations in newly diagnosed mantle cell lymphoma correlate with chemoresistance and call for innovative upfront therapy. *Cancers (Basel)*. 2020;12(8):2120. (IF 2020: 6.639; Q1)

#### APPENDIX I

Obr A, Klener P, Furst T, Kriegova E, Zemanova Z, Urbankova H, Jirkuvova A, **Petrackova A**, Malarikova D, Forsterova K, Cudova B, Sedlarikova L, Berkova A, Kasalova N, Papajik T, Trneny M. A high TP53 mutation burden is a strong predictor of primary refractory mantle cell lymphoma. *Br J Haematol*. 2020;191(5):e103-e106. (IF 2020: 6.998; Q1)

#### **Review articles published in impact factor journals:**

#### APPENDIX J

**Petrackova A**, Turcsanyi P, Papajik T, Kriegova E. Revisiting Richter transformation in the era of novel CLL agents. *Blood Reviews*. *Blood Rev*. 2021;49:100824. (IF 2020: 8.250; Q1)

#### APPENDIX K

Sedlarikova L\*, **Petrackova A**\*, Papajik T, Turcsanyi P, Kriegova E. Resistance-associated mutations in chronic lymphocytic leukaemia patients treated with novel agents. *Front Oncol*. 2020;10:894. (IF 2020: 6.244) \*Contributed equally

# Contents

1	INTRODUCTION . . . . .	1
1.1	Precision medicine . . . . .	3
1.2	Molecular diagnostics in haemato–oncology . . . . .	3
1.2.1	Chronic lymphocytic leukemia . . . . .	5
1.2.2	Multiple myeloma . . . . .	11
1.2.3	Mantle cell lymphoma . . . . .	12
1.2.4	NGS in molecular diagnostics of haemato–oncological diseases . . . . .	12
1.2.5	Whole–genome optical mapping for structural variant analysis in haemato–oncological diseases . . . . .	15
1.3	Molecular diagnostics of autoimmune diseases . . . . .	15
1.3.1	Serum biomarkers in systemic lupus erythematosus . . . . .	16
1.3.2	Gene expression biomarkers in autoimmune diseases . . . . .	18
2	AIMS . . . . .	20
3	MATERIALS AND METHODS . . . . .	21
3.1	Analysis of rare germline <i>ATM</i> variants in CLL . . . . .	21
3.1.1	Study subjects . . . . .	21
3.1.2	NGS and variant annotation . . . . .	21
3.1.3	Assessment of functional activity of <i>ATM</i> . . . . .	22
3.2	NGS analysis of <i>TP53</i> mutations in MM . . . . .	22
3.2.1	Study subjects . . . . .	22
3.2.2	The sample collection and plasma cell enrichment . . . . .	23
3.2.3	NGS mutation assessment . . . . .	23
3.3	NGS analysis of <i>TP53</i> clonal and subclonal mutations . . . . .	23
3.4	Whole genome optical mapping . . . . .	24
3.5	Analysis of serum biomarkers in SLE . . . . .	24
3.5.1	Study subjects . . . . .	24
3.5.2	Proximity extension immunoassay . . . . .	25
3.5.3	Statistics . . . . .	25
3.6	Gene expression biomarkers in autoimmune diseases . . . . .	25
3.6.1	Study subjects . . . . .	25

	3.6.2	Quantitative reverse transcription polymerase chain reaction . . . . .	26
	3.6.3	Statistics and data-mining methods . . . . .	27
4		RESULTS . . . . .	29
	4.1	Genetic biomarkers in haemato-oncology diseases . . . . .	29
	4.1.1	Rare germline <i>ATM</i> variants in CLL . . . . .	29
	4.1.2	Applicability of the whole bone marrow for analysis of <i>TP53</i> mutations in MM . . . . .	34
	4.1.3	Standardization of sequencing coverage depth in diagnostic NGS . . . . .	37
	4.1.4	<i>TP53</i> as biomarker in MCL . . . . .	40
	4.1.5	Whole-genome optical mapping for structural variant analysis in MM . . . . .	41
	4.2	Serum biomarkers in SLE . . . . .	41
	4.3	Gene expression biomarkers in autoimmune diseases . . . . .	46
5		DISCUSSION . . . . .	51
	5.1	Genetic biomarkers in haemato-oncology diseases . . . . .	51
	5.1.1	Rare germline <i>ATM</i> variants in CLL . . . . .	51
	5.1.2	Applicability of the whole bone marrow for analysis of <i>TP53</i> mutations in MM . . . . .	55
	5.1.3	Standardization of sequencing coverage depth in diagnostic NGS . . . . .	57
	5.2	Serum biomarkers in SLE . . . . .	60
	5.3	Gene expression biomarkers in autoimmune diseases . . . . .	64
6		CONCLUSION . . . . .	69
7		Abbreviations . . . . .	70
8		References . . . . .	72
9		Appendix - full text publications related to the thesis . . . . .	90



# 1 INTRODUCTION

Recent advances in high-throughput technologies have empowered unprecedented characterisation of molecular and cellular changes underlying the including development and progression of human diseases. Analysis of clinical samples at multiple molecular levels, including the genome, epigenome, transcriptome, proteome and metabolome, have opened the way to precision medicine. Precision medicine is, according to the Precision Medicine Initiative, "an emerging approach for disease prevention and treatment that takes into account people's individual variations in genes, environment, and lifestyle" [1, 2]. The hallmark of precision medicine is the use of the most appropriate treatment for each patient according to their individual variability, such as genes, environment, cellular and molecular profiles among others, which reflect differences in disease susceptibility, prognosis and response rates [3].

The stratification of patients is guided by biomarkers, which thus play a key role in precision medicine [4]. Biomarkers are quantifiable measurements of biologic homeostasis that define what is "normal", hereby providing a frame of reference for predicting or detecting what is "abnormal" [5]. The National Institutes of Health Biomarkers Definitions Working Group defined a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" [6]. Biomarkers can exist in many different forms including patient performance status, images (e.g. PET/CT scan), genetic alterations (e.g. *BRCA1/2* mutations), gene or protein expression profiles (e.g. serum protein electrophoresis for detection of monoclonal gammopathies), and cell-based markers (e.g. circulating tumour cells), among others (Figure 1) [5].

Biomarkers are used to determine disease and prognosis, guide treatment, and assess therapy response [4]. Six types of biomarkers are currently recognized: i) risk to identify the risk of developing an illness, ii) screening to screen for subclinical disease, iii) diagnostic to recognise overt disease, iv) staging to categorise disease severity [7], v) prognostic to identify patients with different outcome risks and vi)

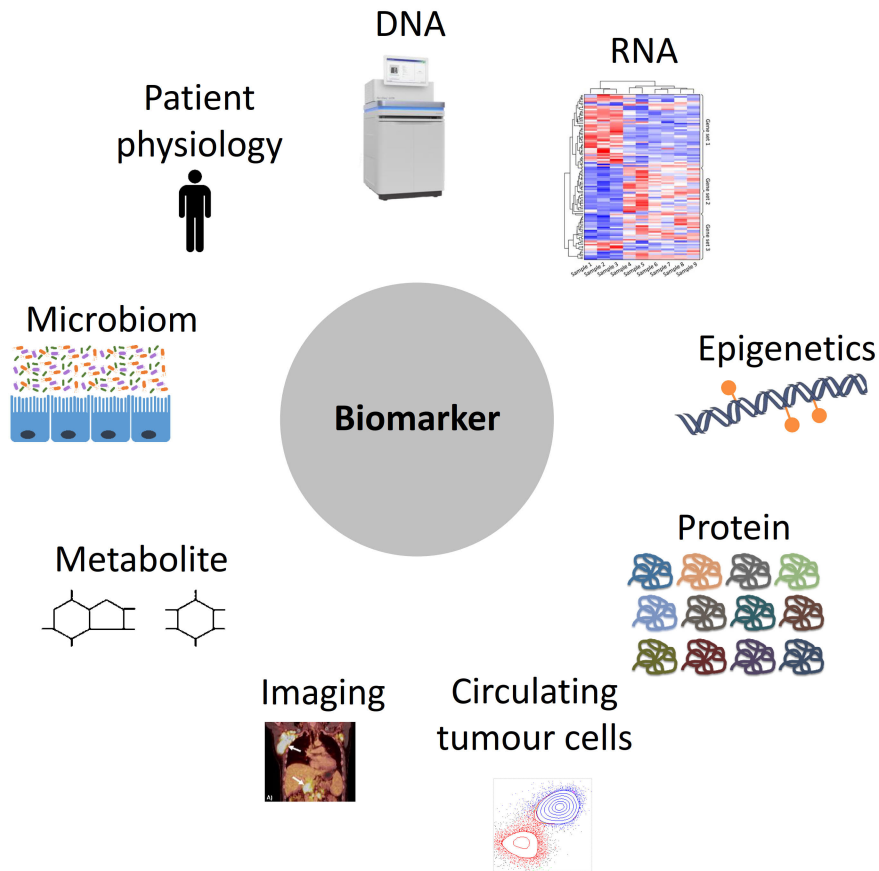


Figure 1: Biomarker types.

predictive to assist in predicting which patients will respond to a given therapy [8]. Regardless of their role, a reliable biomarker in a clinical setting must have sufficiently high sensitivity and specificity of the test or examination to provide a clear result either confirming the disease or other "abnormal" condition (high positive predictive value) or excluding it (negative predictive value) [7].

The main factors driving progress in precision medicine is biomarker discovery, validation and translation from bench to bedside. It involves two complementary strategies: the "knowledge-based" and the "unbiased"; the knowledge-based strategy relies on a direct understanding of the underlying biology of the disease process, while the unbiased approach relies on investigation of a high number of molecules using technological advances to characterize the biomolecular signature of a defined

disease phenotype. Evolution of biomarkers now represents a well-coordinated effort in a multidisciplinary environment allowing effective translation from the scientific research to clinical practice [7].

## 1.1 Precision medicine

The most prominent area of precision medicine and biomarker assessment is molecular diagnostics. It is a part of laboratory medicine, which relies on the detection of individual biologic molecules by molecular genetic techniques. The potential of molecular genetic tools was initially recognized by haemato-oncologists with the discovery of specific chromosomal translocations guiding the diagnosis of various leukemias and lymphomas [9, 10]. Nowadays, biomarkers are used in all fields of medicine.

## 1.2 Molecular diagnostics in haemato-oncology

Oncology is at the leading edge of precision medicine, moving beyond the current model of delivering anticancer treatment based on studies of largely unselected patients beyond a simple phenotypic marker and taking the lead in using the molecular profile of an individual's cancer genome to optimize treatment of their disease [11]. Thanks to the rapid improvement of sequencing methods, great progress has been made in the last 15 years in understanding the underlying mechanisms of malignancies, which would not have been possible with previous investigative techniques [12–14]. This knowledge has enabled the introduction of molecular tests that have become part of the standard treatment of patients; the identification of people with hereditary tumours in clinical oncology, as are a number of tests that help select the most effective treatment based on the molecular characteristics of the tumour tissue or some other biological parameters of the malignancy.

In haemato-oncology, molecular genetic tests are now routinely used in clinical laboratories. These mainly employ sequencing techniques (Sanger sequencing, next-generation sequencing (NGS)) and cytogenetics. Detection of genetic alter-

ations not only aids in diagnosis and treatment decisions, but also provides important prognostic information (Figure 2). In addition, genetic rearrangements associated with leukemia can be used as molecular markers allowing the detection of low-level residual disease [15].

Our research group focuses mainly on the search for genetic biomarkers and refinement of current sequencing biomarkers in chronic lymphocytic leukemia (CLL), multiple myeloma (MM) and mantle cell lymphoma (MCL) by novel approaches in molecular diagnostics, primarily using NGS technology. The addressed issues regarding molecular diagnostics of these diseases will be further elaborated.

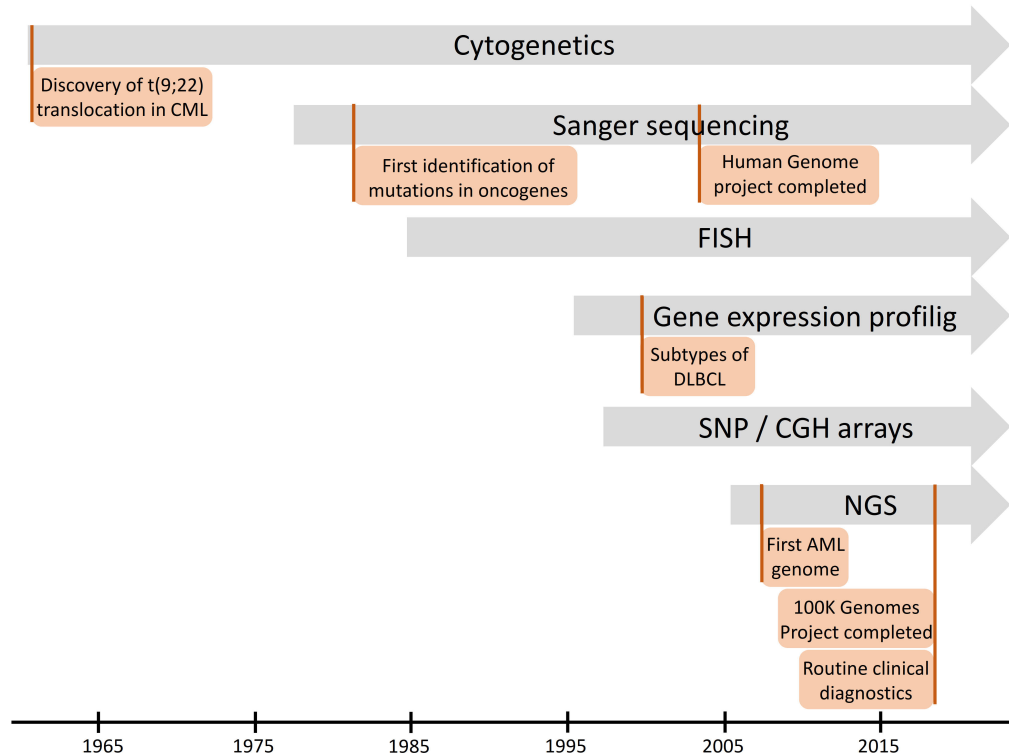


Figure 2: Timeline of genetic methods. Adapted from [14].

Legend: AML: acute myeloid leukemia; CML: chronic myeloid leukemia; DLBCL: diffuse large B-cell lymphoma.

### 1.2.1 Chronic lymphocytic leukemia

Chronic lymphocytic leukemia (CLL) is a malignancy of CD5<sup>+</sup> B cells and represent the most common leukemia in the Western world [16]. The most important molecular biomarkers in CLL include *TP53* disruption and *IGHV* (Ig heavy chain variable) gene mutation status. Based on international recommendations, molecular analysis of deletion 17p (del(17p)) which includes the locus of the tumour suppressor gene *TP53*, *TP53* mutations and (*IGHV*) mutation status should be assessed in each CLL patient before treatment as they are relevant for choice of therapy [16]. Patients with a del(17p) and/or *TP53* mutation have the poorest prognosis at least in the era of chemoimmunotherapy, with a median overall survival (OS) of 2–5 years [17, 18]. The prognosis of those patients has significantly improved with the introduction of B–cell receptor inhibitors and the BCL2 inhibitor venetoclax. Also CLL patients with unmutated *IGHV* status have a higher risk of adverse genetic aberrations and OS and time to treatment are significantly shorter in this patient group when compared to patients with mutated *IGHV* [16]. Since leukemic clones may evolve under pressure of therapy, del(17p) and *TP53* mutation analysis should be repeated before any line of therapy [19].

Moreover, additional molecular biomarkers are available to predict the prognosis of CLL patients [16]. These include somatic mutations in *ATM*, *NOTCH1*, *SF3B1*, *BIRC3*, *RPS15* genes, deletion 11q (del(11q)) and complex karyotype (defined by  $\geq 3$  or  $\geq 5$  abnormalities in chromosomal banding analysis) [16, 20–26].

### Richter transformation

In approximately 2–10% of CLL patients, aggressive lymphoma – most frequently diffuse large B–cell lymphoma (DLBCL) and rarely Hodgkin lymphoma (HL) – arise on the background of CLL, a phenomenon known as Richter transformation (RT). Despite recent advances in CLL treatment, RT also develops in patients on novel agents, usually occurring as an early event. RT incidence is lower in CLL patients treated with novel agents in the frontline compared to relapsed/refractory cases, but

higher in patients with *TP53* disruption. The genetic heterogeneity and complexity are higher in RT-DLBCL than CLL; the genetics of RT-HL are largely unknown mainly due to its rarity. In addition to *TP53*, aberrations in *CDKN2A*, *MYC*, and *NOTCH1* genes are common in RT-DLBCL (Figure 3); however, no distinct RT-specific genetic aberration is recognised yet [27]. Our research group recently addressed the topic of the RT in a review article [27] (Appendix J).

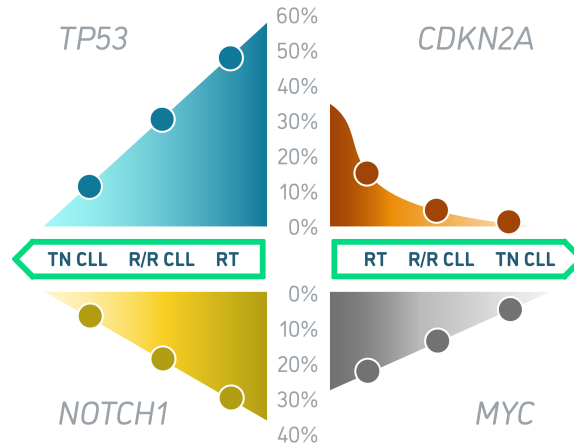


Figure 3: Frequency of genetic abnormalities in *TP53*, *NOTCH1*, *CDKN2A* and *MYC* genes in RT-DLBCL compared to treatment-naïve (TN) and relapsed or refractory (R/R) CLL patients.

### Therapy and resistance

Regarding the treatment of CLL, the approval of B-cell receptor (BCR) inhibitors (ibrutinib and idelalisib) and the BCL-2 antagonist venetoclax has led to the great shift in the CLL therapeutic management in the last decade (Figure 4). These novel agents have shown significant clinical efficacy in high-risk patients with relapsed/refractory disease with del(17p) and/or *TP53* mutation and complex karyotype as well as in previously untreated patients with/without poor-risk features [28–32]. Although these novel agents are superior to chemoimmunotherapy, the treatment is still failing in some patients and it is expected that the number of patients who progress or develop clinical resistance will increase with the growing number of patients indicated for this treatment and due to the long-term administration of these agents. Understanding the mechanisms that cause resistance and identifying the driving mutations and signalling pathways involved is therefore a

current need [28]. The topic of resistance-associated mutations in CLL patients treated with novel agents was reviewed by our research group [28] (Appendix K).

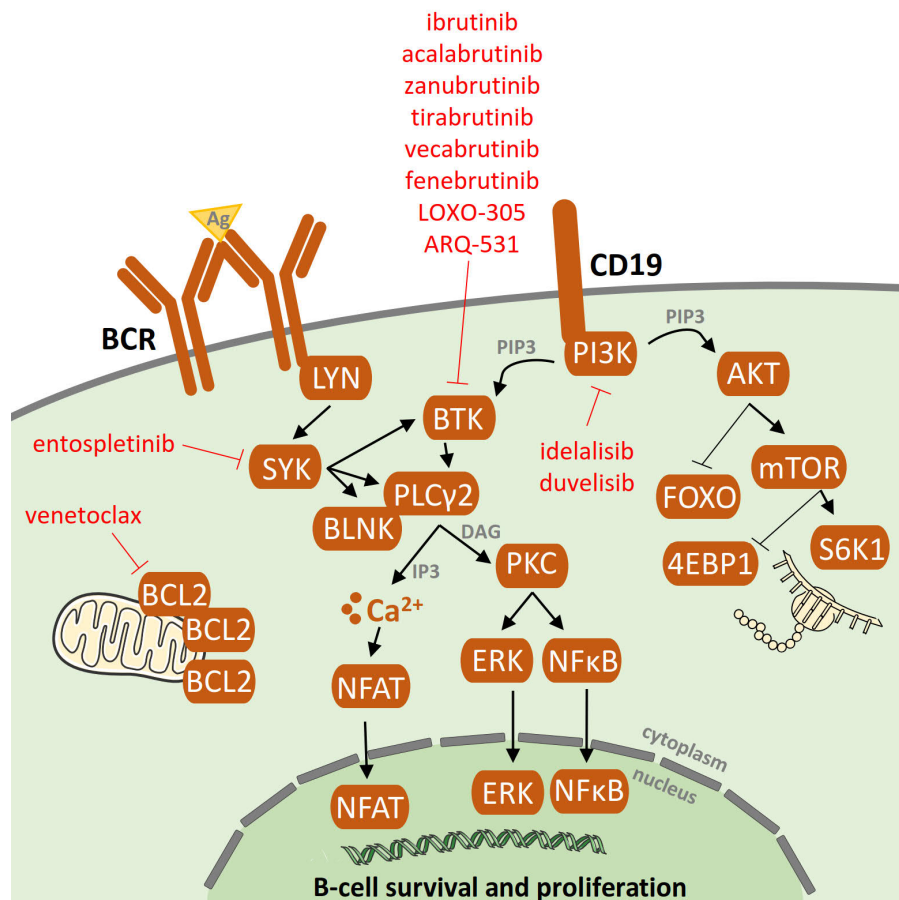


Figure 4: Mechanism of B-cell receptor (BCR) signaling pathway inhibitors [28]. Following antigen binding to the B-cell receptors, BCR signaling is chronically activated in CLL cells. As a consequence, the activated BCR pathway results in the production of second messengers and the activation of NF- $\kappa$ B and subsequent pro-proliferation and antiapoptotic pathways.

Primary resistance with no initial response to ibrutinib has been observed rarely and its mechanism is not yet understood [33–35]. Acquired secondary resistance to ibrutinib occurs in 8–13% of CLL cases who responded well to the treatment initiation [33,34] and are most commonly caused by the occurrence of resistance-associated mutations (Figure 5). A number of studies have confirmed the presence of *BTK* mutations in CLL patients relapsing on ibrutinib and have shown that those mutations were not present prior to drug administration [36–39]. The most common mutation (C481S) was found at the position of the binding site for ibrutinib thus reducing ibrutinib affinity for BTK [40,41]. *PLCG2* gain-of-function mutations are the second

most frequent mutations found in CLL patients who failed on the ibrutinib treatment. *PLCG2* is the gene encoding phospholipase C $\gamma$ 2, the protein immediately downstream of BTK. These mutations mostly have an activating effect resulting in continuous BCR signaling independently on BTK activation [36,42]. Although mutations in *BTK* and *PLCG2* genes are detected in  $\sim$ 80% of CLL patients who failed on ibrutinib [36–38], for 20% of patients, ibrutinib resistance-associated mutations remain unknown [38,39,43]. These data further support the presence of alternative mechanisms of drug resistance other than *BTK/PLCG2* mutations in a subset of patients who are still under investigation.

To date, several candidate loci/mutations that may contribute to resistance have been described. These include del(8p) and *SF3B1*, *PCLO*, *EP300*, *MLL2*, and *EIF2A* mutations [36,39,43]. More candidate genetic factors associated with resistance for ibrutinib-treated patients involve *BCL6* rearrangements, *MYC* gene abnormalities, del(17p), del(18p), 2p gain, *XPO1* overexpression, complex karyotype, epigenetic changes, and changes in the cell microenvironment [29,39,44,45]. However, it remains unclear whether these aberrations contribute causally to clinical resistance [28].

Venetoclax is an oral BH3 mimetic and highly selective inhibitor of the BCL-2 antiapoptotic protein, capable of restoring apoptosis tumour cells with high overall response rates as well as in heavily pretreated, high-risk CLL patients. Venetoclax is primarily available for CLL patients with *TP53* disruption, for patients who failed on ibrutinib or were not suitable for the treatment with BCR signaling inhibitors, as well as for patients refractory to chemoimmunotherapy [46–48]. There is already growing evidence about the role of acquired mutations in *BCL2* gene also within the BH3-binding site leading to the progression and failure of venetoclax [49,50].

As shown in the example of CLL, it is recommended to evaluate underlying genetic defects that serve as biomarkers at the time of diagnosis of haematological malignancies, as well as prior to and during therapy.

In addition to somatic aberrations with a clear impact on the pathogenesis and



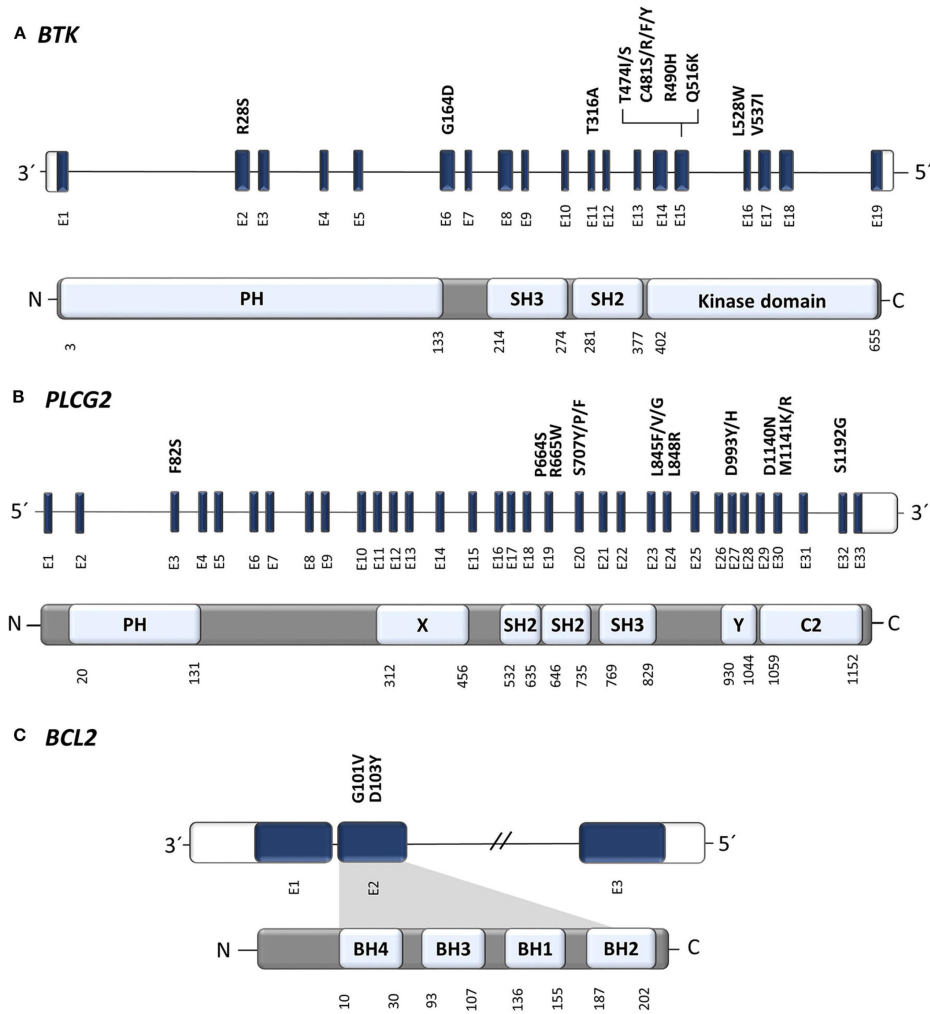


Figure 5: Position of known resistance-associated mutations in (A) *BTK*, (B) *PLCG2*, and (C) *BCL2* genes [28].

Legend: E, exon. Gene structure has been visualized according to a genome browser Ensembl (Ensembl release 99, January 2020, EMBL–EBI); protein domains and amino acid positions according UniProt Knowledgebase (2020-01 release, February 2020, UniProt Consortium).

therapy of haemato–oncological diseases, there are also emerging reports of the contribution of rare germline variants to the development and progression of cancer.

### Rare germline variations in *ATM*

Rare germline variations, occurring in the population with less than 0.5% frequency, have been recently revealed to have a crucial role in cancer etiology, especially when they occur in tumour suppressor genes [51–53]. Multiple lines of evidence showed that rare germline variants affect drug sensitivity, enhance the likelihood of addi-

tional somatic aberrations and accelerate cancer progression [54–56]. However, in cancer diagnostics, they are often i) referred to as variants of uncertain significance (VUS), as there is limited knowledge of their clinical and functional impact, or ii) overseen if paired tumour–normal variant analysis is used to filter tumour–only variants [57].

One of the genes, where rare, protein–altering germline variants occur, is *ATM*, a tumour suppressor essential for genome stability by regulating the DNA double–strand break response. Despite the known association of inherited rare variations in the *ATM* gene with the autosomal recessive disorder ataxia telangiectasia [58], there is growing evidence of their role in cancer pathogenesis [53, 59–61] including CLL [51, 62]. With the exception of truncating alleles causing ataxia telangiectasia [58], the clinical significance of most rare *ATM* germline variants is not fully elucidated as the majority of data is available only for *ATM* somatic disruption in cancer. In CLL, somatic *ATM* disruption (mutations/deletions) is recognised as a negative prognostic event, comparable to *TP53* abnormalities [26].

In our research study (Appendix A) we investigated the clinical and functional impact of rare germline *ATM* variants in CLL patients, particularly those on novel agents, and compared them with adverse somatic aberrations in CLL (del(11q), *ATM* mutation, del(17p), *TP53* mutation, *IGHV* status).

Our motivation for investigating rare germline, protein–altering variants in the *ATM* gene in CLL was the number of VUS variants in *ATM* that we have reported in our diagnostic tests of a panel of CLL–associated genes (*TP53*, *ATM*, *BIRC3*, *NOTCH1*, *SF3B1*, *POT1*, *MYD88*, *FBXW7*, *XPO1*, *EGR2*, *NFKBIE*, *RPS15*). Additionally, we analysed the annotation of rare germline *ATM* pathogenic/predicted pathogenic variants in diagnostic databases (ClinVar, VarSome) and their frequency in public whole exome sequencing (WES) tumour–normal datasets of CLL, MCL, metastatic breast and lung cancer patients. A comparison of published data on multiple cancers showed that the prevalence of rare *ATM* pathogenic/predicted pathogenic variants is underestimated across all cancers.

### 1.2.2 Multiple myeloma

Multiple myeloma (MM) is a clonal plasma cell proliferative disorder usually limited to a bone marrow (BM) microenvironment. Several recurrent translocations and copy number abnormalities can be found in MM patients [63], although at diagnosis, their analysis is usually limited to events that have an established prognostic role and may guide treatment: del(17p) and translocations t(4;14) and t(14;16) [64–66]. Somatic gene mutations are considered more likely to be secondary events associated with tumour progression [67, 68], and NGS studies have shown that MM is very heterogeneous in its spectrum of somatic mutations compared to other haematologic malignancies [66].

Del(17p) is among the important biomarkers of poor prognosis in MM [69] and its analysis is a part of the recommended risk assessment in newly diagnosed MM patients [70]. However, *TP53* mutation analysis in MM is not widely performed in diagnostics, particularly due to technical limitations regarding sample collection and plasma cell (PC) enrichment. The pitfalls are mainly connected with inter-individual variability in the sample amount, PC infiltration in BM, PC immunophenotypes and time-dependent losses of surface markers as well as haemodilution, patchy or site varied PC distribution, aspirate pull order and aggregation of PC in aspirated BM [71, 72], all together resulting in low PC recovery in some patients.

Therefore, there is a need to standardise the pre-processing of BM in clinical settings, as well as to have an alternative, when enriched tumour PCs are not available. In our research study [73] (Appendix B), we explored the applicability of the whole BM for analysis of somatic *TP53* mutations, which serve as predictive biomarkers in MM patients, together with *NRAS*, *KRAS* and *BRAF* hotspots, by deep targeted NGS in diagnostics.

### 1.2.3 Mantle cell lymphoma

Mantle cell lymphoma (MCL) is a subtype of B-cell lymphoma with a large number of recurrent cytogenetic/molecular aberrations. Approximately 5–10% of patients do not respond to frontline immunochemotherapy. Despite many useful prognostic indexes, new molecular prognostic predictors, which would reflect critical aspects of MCL biology and would help in therapy decisions, are still missing [74]. In our research studies, we analysed the prognostic impact of somatic mutations in *TP53* gene in two real-world cohorts of MCL patients using NGS [75, 76] (Appendices H and I).

### 1.2.4 NGS in molecular diagnostics of haemato–oncological diseases

In our research studies in the field of haemato–oncological diseases, we have primarily used NGS to investigate candidate genetic biomarkers or to refine their analysis.

NGS refers to sequencing methods that have reduced the time and cost compared to Sanger sequencing and thus significantly increased sequencing output. Sanger dideoxy synthesis sequencing method [77, 78] together with the Maxam–Gilbert chemical cleavage method [79] were the founding techniques of DNA sequencing.

The Maxam–Gilbert method is based on chemical modification of DNA and subsequent cleavage of the DNA backbone at sites adjacent to the modified nucleotides. Sanger sequencing uses specific chain-terminating nucleotides, dideoxy nucleotides, which lack the 3'-OH group, resulting in the termination of the growing DNA strand at this position. Nucleotide detection is enabled by radioactive or fluorescently labeled dideoxynucleotides on “sequencing” gels or automated sequencing machines. Sanger sequencing using the synthetic dideoxy method was developed in 1977 and has become the standard for sequencing [80].

The automation of Sanger sequencing allowed ever larger amounts of DNA to be read, culminating in the Human Genome Project [81]. The Human Genome Project was the largest collaborative biology project in the world to date, taking 13 years and costing nearly \$3 billion to complete. As a next step, large-scale sequencing

projects were undertaken to study human sequence variation. However, for these types of projects Sanger sequencing was too labor intensive, time consuming, and expensive [82]. In 2004, the National Human Genome Research Institute launched a program to reduce the cost of WGS (whole genome sequencing) to \$1000 over 10 years [83]. This accelerated the development of cheaper and faster methods, and in the following years NGS technologies generating thousands to millions of sequencing reactions per run emerged. For many years, different NGS technologies were developed and coexisted; however, today the market is largely dominated by Illumina platforms, especially in human medicine and clinical diagnostics.

Illumina uses sequencing by synthesis with optical base calling [84]. Following library preparation, which affixes adaptors to DNA fragments of approximately 150–500 bp, these fragments are hybridized to a glass slide covered with complementary adaptors. Once attached to the solid surface, the fragments are PCR amplified from one end only (single-end read) or from both ends (paired-end reads), producing millions to billions of clusters of clonal template DNA fragments that are to be sequenced simultaneously [84]. Sequencing itself is performed by synthesis using reversible terminator nucleotides, adaptation over Sanger sequencing. This permits one nucleotide to be incorporated at a time and the representative fluorescence to be recorded as a base call by high-resolution optical imaging, followed by cleavage of the terminal chemical modification, thereby allowing the next complementary nucleotide to be incorporated. This process is repeated for the length of the read to generate the sequence output, where read lengths are now typically between 75 to 250 bp [84–86].

Currently, deep targeted NGS of genes of interest is most commonly used for routine diagnostic testing in medicine. In addition to deep targeted NGS, other strategies include for example WES, which targets the entire set of human exons (approximately 2% of the human genome) [87], and WGS, which is theoretically capable of examining the entire human genome [88, 89].

NGS has rapidly expanded into the clinical setting in haemato-oncology and

oncology, as it may bring great benefits for diagnosis, selection of treatment, and/or prognostication for many patients [90]. Recently, several articles about the validation of deep targeted NGS in clinical oncology were published [91, 92], including a comprehensive recommendation by the Association for Molecular Pathology and the College of American Pathologists [90]. However, the lack of standardization of targeted NGS methods still limits their implementation in clinical practice [93].

One challenge in particular is the correct detection of mutations present at low variant allele frequencies (VAF) and standardization of sequencing coverage depth [90, 94, 95]. This is especially important for mutations that have clinical impacts at subclonal frequencies [90] such as the case of *TP53* gene mutations in CLL [96, 97]. *TP53* aberrations (*TP53* mutation and/or del(17p)) are among the strongest prognostic and predictive biomarkers guiding treatment decisions in CLL [98]. Nowadays, the European Research Initiative on Chronic Lymphocytic Leukemia (ERIC) recommends detecting *TP53* mutations with a limit of detection (LOD) of at least 10% VAF [99], and a growing body of evidence exists dedicated to the clinical impact of small *TP53* mutated subclones in CLL [96, 97].

Sanger sequencing and deep targeted NGS are currently the techniques most used for *TP53* mutation analysis [99] as well as for analysis of other genes with clinical impacts at low allele frequencies. Although Sanger sequencing provides a relatively accessible sequencing approach, it lacks the sensitivity needed to detect subclones due to its detection limit of 10–20% of mutated alleles [99]. NGS-based analysis has thus gained prominence in diagnostic laboratories for the detection of somatic variants and various technical developments of error correction strategies, both computational and experimental, are being developed for the accurate identification of low-level genetic variations [100]. In our research study [101] (Appendix D), we addressed the importance of the correct determination of sequencing depth in diagnostic NGS in order to obtain a confident and reproducible detection, not only of low VAF variants.

### 1.2.5 Whole-genome optical mapping for structural variant analysis in haemato-oncological diseases

Whole-genome optical mapping is a novel tool for detection of structural variations (SVs), including copy number aberrations (CNAs) in human genome. There is a growing body of evidence on the utility of optical mapping for resolving complex genomic architecture in haematology and solid tumours [102, 103]. Optical mapping has an advantage in detecting small and large structural rearrangements as well as complex rearrangements across the whole genome that are undetectable by traditional methods, such as sequencing and cytogenetics [103].

Our research group performed a pilot study on utility of whole-genome optical mapping to analyse the genomic architecture of extramedullary multiple myeloma (EMM) [104] (Appendix G). EMM represents a rare, aggressive and mostly resistant phenotype of MM and is frequently associated with high-risk cytogenetics, but their complex genomic architecture is largely unexplored.

## 1.3 Molecular diagnostics of autoimmune diseases

In the field of oncology, the era of precision medicine with intensive use of biomarkers is the most visible. However, molecular diagnostics is also advancing for autoimmune diseases that have a complicated molecular background and cannot be fully captured by traditional methods.

Autoimmune diseases can affect almost any organ and are caused by impaired immune tolerance to host-derived or “self” antigens. Their incidence has been steadily increasing in developed countries over the last four decades [105] and together they represent the third most common cause of morbidity and mortality after cardiovascular disease and cancer [106]. These diseases are characterised by a multifactorial etiology in which genetic factors interplay with environmental factors [107].

Our research group focuses in particular on the search for serum protein biomarkers and genetic expression (mRNA) biomarkers in selected chronic rheumatic autoimmune diseases: systemic lupus erythematosus (SLE), systemic sclerosis (SSc)

and rheumatoid arthritis (RA).

### 1.3.1 Serum biomarkers in systemic lupus erythematosus

SLE is a complex, multi-system autoimmune rheumatic disease with significant variability in the phenotypes and severity of the disease. The greatest challenges continue to be the prevention and management of irreversible organ damage and active lupus nephritis (LN), one of the most feared phenotypes in SLE.

Organ damage is a primary outcome in SLE, which is accrued not only during the disease course, but also by therapy itself [108]. Early damage is more likely to be linked to active inflammation, while late irreversible damage is often attributable to the side effects of drugs and especially to chronic and cumulative corticosteroid exposure [109]. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology SLICC/ACR Damage Index (SDI), divided into 38 items grouped in 12 organ systems, is a valid measure of irreversible organ damage in SLE [108]. Despite improvement in the survival of SLE patients in recent decades, significantly higher morbidity and mortality are reported in patients developing irreversible organ damage [108]. The patterns of organ damage vary among populations [110–112], but the musculoskeletal, cardiovascular, and renal systems are those most frequently affected [113]. Nowadays, prevention of irreversible damage is a major goal in the management of SLE patients and identification of the key molecules involved in the pathogenesis of organ damage is needed.

Lupus nephritis is a major manifestation associated with higher morbidity and mortality of SLE patients [114]. It has a considerable influence on treatment decisions, as well as long-term outcomes. The effective treatment of LN requires a correct diagnosis, timely intervention, and early treatment of any disease relapse. Renal biopsy is still the gold standard for diagnosis and deciding on therapy in LN but its invasive nature prevents it from being used repetitively in many cases [115]. Traditional clinical parameters such as proteinuria, glomerular filtration rate, urine sediments, anti-dsDNA antibodies, and complement levels are not sen-



sitive or specific enough to detect activity and early relapse of LN [116, 117]. Novel serum and urinary biomarkers such as cytokines and chemokines CCL2 [118], CCL3, CCL5 [119], IL17 [118], BlyS, APRIL [120], growth factor TGF $\beta$  [118] and others (TWEAK [121], IGFBP2 [122], OPG [123]) have recently been nominated for diagnosis and monitoring of LN. Although intensively investigated [124, 125], only a few biomarkers have been assessed for prediction of renal activity or prognosis. Identification of novel and reliable biomarkers or their combinations for LN reflecting also disease activity is, therefore, highly desirable.

In our research study we aimed to assess the serum protein pattern of SLE using a highly sensitive multiplex proximity extension immunoassay (PEA) on 92 inflammation-related proteins [126] (Appendix F). Special emphasis was given to serum patterns associated with irreversible organ damage and LN reflecting the renal disease activity and their usefulness in the prediction of these severe phenotypes.

PEA is immunoassay for high throughput detection of protein biomarkers in liquid samples. Each biomarker or analyte is recognized by a pair of oligonucleotide-labelled antibodies and when binding to their correct targets, they give rise to reporter amplicons which are amplified and quantified by microfluidic-based real-time PCR. The data obtained is normalized and used for the relative quantification of the concentration of each analyte [127]. The PEA offer the same level of performance as ELISA and comparable sensitivity to standard ELISA kits with much less sample and a higher dynamic range [126, 127].

PEA technique is based on immuno-PCR (iPCR), first described in 1992 [128], which amplifies biotinylated linear plasmid DNA associated with antigen/monoclonal antibody complexes immobilized on microtiter plate wells via streptavidin-protein A chimeras [129]. Additional modifications provided a more universal iPCR by replacing the fusion protein with commercially available biotinylated secondary antibodies and readout with fluorogenic PCR instead of gel electrophoresis [130].

The main drawback of iPCR is its non-homogeneous nature, which requires extensive washing steps to ensure minimal background signal. Proximity assays

address this problem and the first of which was the proximity ligation assay [129]. Their design can be variable, depending on the use of monoclonal or polyclonal antibodies, and also on a combination of both, but in principle they are based on two types of oligonucleotides bound to an antibody (one oligonucleotide per antibody), which are in close proximity and can therefore be joined by DNA ligation to form a template for PCR amplification.

Another variant of proximity assay is PEA where DNA ligation is not used but oligonucleotides are designed to be complementary. After hybridization, extension of the hybridizing oligonucleotide, bound to one of the probes, creates a DNA amplicon that can subsequently be detected and quantified by quantitative real-time PCR [129].

PEA, which uses two different types of antibodies and a specific DNA polymerase, allows to significantly minimize background noise and thus increase the sensitivity of the assay [127]. PEA is suitable for multiplexing [127] and has the advantages of very low sample consumption (up to 1  $\mu$ l) [131].

### 1.3.2 Gene expression biomarkers in autoimmune diseases

RA, SLE, and SSc are systemic autoimmune diseases characterized by overactivation of the innate immune system together with impaired downstream pathway of type I interferon- (IFN-) responding genes (IFN signature). Nevertheless, a certain heterogeneity in the IFN signature among those diseases has been recognized, and some patients even lack its presence [132–135].

Although the emerging role of the innate immunity in the pathogenesis of RA, SLE, and SSc has been demonstrated, there is no data on the cross-disease innate gene signature as well as its heterogeneity among those diseases yet. Numerous studies on individual innate immunity members in RA, SLE, and SSc showed the crucial role of Toll-like receptors (TLRs) and IL1 family [136, 137]. Notable examples of common innate pathways are (i) the involvement of the adapter protein MyD88 which is required for signal transduction by TLRs and receptors of the IL1

family, (ii) the activation of the type I IFN, and (iii) the presence of endogenous TLR ligands [138]. Besides shared innate pathways, disease-specific molecular and cellular mechanisms exist. In SLE, recent evidence has suggested a close relationship between the endosomal TLR activation and the disease onset [139, 140] with an essential role of endosomal TLRs in the generation of anti-nuclear antibodies and type I IFNs [141]. In RA, abundant activation of individual members of TLR and IL1 families was already evidenced with a proposed role for exogenous TLR ligands in the disease onset (i.e., *Proteus* infection of urinary tract, Epstein-Barr virus, and parvovirus B19) and for endogenous ligands in self-sustaining of the inflammatory loop [136, 142]. In SSc, signaling via TLR is increasingly recognized as a key player driving the persistent fibrotic response and is linked to the activity of TGF $\beta$ ; however, the pathological role of TLRs and their ligands in SSc still remains unclear [143].

In our research we aimed to elucidate the underlying differences in the innate immunity signature across three major autoimmune disorders using multivariate analysis [144, 145] (Appendices C and E). This first cross-disease analysis of the innate gene expression signature of 10 *TLRs*, 7 key members of the *IL1/IL1R* family, and interleukin 8 (*CXCL8*) in peripheral blood mononuclear cells (PBMC) from patients with active SLE, RA, and SSc revealed emerging diversity and abundance in RA compared to SLE and SSc. Our study contributes to further understanding of the innate signature underlying the immunopathology of major autoimmune diseases, with special emphases to discriminate shared and disease-specific expression patterns.

## 2 AIMS

To search for candidate molecular biomarkers using ultrasensitive molecular genetic methods in haemato-oncology and autoimmune diseases:

1. To search for candidate genetic biomarkers by next-generation sequencing (NGS) in haemato-oncology diseases.
2. To search for candidate protein biomarkers by immunoassay in autoimmune disease systemic lupus erythematosus.
3. To search for candidate mRNA biomarkers by high-throughput quantitative reverse transcription polymerase chain reaction (RT-qPCR) in autoimmune diseases.

## 3 MATERIALS AND METHODS

### 3.1 Analysis of rare germline *ATM* variants in CLL

#### 3.1.1 Study subjects

A real-world cohort of 336 patients with CLL was recruited between 2016–2020 in a single tertiary haematological centre at University Hospital Olomouc. All enrolled patients were diagnosed in accordance with the international criteria [146] for clinical and genetic characteristics see Appendix A, Table 1. The median follow-up was 25 months (min–max 1–193 months). Additionally, 198 healthy controls (113/83 female/male, median age 78) with no history of cancer or autoimmune diseases were recruited based on the patients' records.

#### 3.1.2 NGS and variant annotation

A complete *ATM* coding sequence was analysed by deep NGS of whole blood DNA as reported previously [76, 101]. A germline/somatic origin of detected *ATM* variants was confirmed by analysis of patients' germline saliva DNA (available for 93% of patients). Rare variants were defined as those having a minor allele frequency (MAF) <0.5%, according to the gnomAD database in any population [147]. Only protein-altering variants (missense, nonsense, frameshift and splice region) were investigated and annotated by VarSome and ClinVar databases; for missense variants, Sift and PolyPhen prediction tools were used. ATM activity was assessed in CLL and T cells from cryopreserved PBMC based on the phosphorylation of ATM-specific substrate KAP1 and ATM autophosphorylation of Ser-1981, after exposure to etoposide.

Moreover, following public whole exome/genome sequencing, datasets on tumour-normal samples in CLL (public datasets EGAD00001001464, EGAD00001001466), MCL (EGAD00001006159), metastatic breast (EGAD00001002747) and lung (EGAD00001004027, EGAD00001003960) were evaluated for rare germline and somatic variants, as well as CNA in the 11q region using the VarScan (v.2.4.4) tool.

Statistical analyses were performed using R software.

### 3.1.3 Assessment of functional activity of ATM

Patient's cryopreserved PBMC were thawed and rested overnight in a 1 ml suspension of  $2 \times 10^5$  of PBMC in complete RPMI-1640 medium in a 5% CO<sub>2</sub>, 37 °C, 100% relative humidity atmosphere.

A DNA double-strand-breaks-inducing agent, etoposide, was then added to give a final concentration of 25  $\mu$ M to activate ATM kinase. The cells were incubated at 37 °C, 5% CO<sub>2</sub>, 100% relative humidity for 1 h, then fixed by formaldehyde to a final concentration of 4% (vol/vol) and incubated for 20 min at room temperature. Samples were then centrifuged at 500 g/10 min and the supernatant discarded. The cells were permeabilized by three washes of 2 ml of phosphate-buffered saline (PBS) containing 0.1% Triton X-100 and washed three times with PBS containing 0.1% Tween-20. After 45 min incubation of cells with blocking solution (Reagent Diluent Concentrate 2, Bio-Techne) at room temperature, cell pellets were incubated overnight at 4 °C with pKAP1 (Rabbit polyclonal Anti-KAP1 primary antibody, Abcam) and pATM (Anti-phospho-ATM (Ser1981) Antibody, PE conjugated, Sigma-Aldrich) antibody according to manufacturers' recommendations. The next day, samples were washed three times with PBS containing 0.1% Tween-20 and subsequently stained with goat anti-rabbit secondary antibody (ab6564, Abcam) and other conjugated antibodies for CD3/5/19 markers (BioLegend). After 1 h incubation at room temperature in dark, samples were washed two times with PBS containing 0.1% Tween-20 and analysed by 8-color BD FACSCanto-II flow cytometer. All experiments were performed in duplicates.

## 3.2 NGS analysis of *TP53* mutations in MM

### 3.2.1 Study subjects

Diagnostic BM samples were obtained from 54 patients with MM (Appendix B, Table 1), diagnosed according to the International Myeloma Working Group criteria

[148].

### 3.2.2 The sample collection and plasma cell enrichment

Second-pull BM aspirates were collected into RPMI medium containing 25 U/ml heparin. BM cells were isolated either by density centrifugation on Histopaque-1077 (Sigma-Aldrich) at 400 g for 35 min at room temperature or by red blood cell lysis (RBC Lysis Solution, Qiagen). After washing with phosphate-buffered saline (300 g for 10 min at room temperature), PCs were enriched by MACSprep Multiple Myeloma CD138 MicroBeads kit (Miltenyi Biotec) or EasySep Human CD138 Positive Selection Kit II (STEMCELL) according to the manufacturer's recommendations. The percentage of PCs in BM and the purity of enriched PCs was measured by flow cytometry (BD FACSAria, Becton Dickinson) using a combination of CD19/CD38/CD45/CD56/CD138 antibodies (BioLegend).

Cytogenetic abnormalities were assessed by fluorescence in situ hybridisation (FISH) with immunophenotyping, as reported previously [149].

### 3.2.3 NGS mutation assessment

The full coding sequence of the *TP53* gene (exons 2–11, plus 5' and 3'-UTR; NM\_000546) together with hotspot regions in *NRAS* (exons 2–4; NM\_002524), *KRAS* (exons 2–4; NM\_004985) and *BRAF* (exons 11 and 15; NM\_004333) were analysed by NGS on MiSeq (2x151, Illumina) with a minimum target read depth of 5,000x as reported previously [101]. The detection limit was 1%, the variants within the range 1–3% were confirmed by replication.

## 3.3 NGS analysis of *TP53* clonal and subclonal mutations

*TP53* (exons 2–10 including 2 bp intronic overlap, 5' and 3'-UTR) was analyzed using 100 ng genomic DNA per reaction. Amplicon-based libraries were sequenced as paired-end on MiSeq (2x151, Illumina) with minimum target read depths of 5,000x. The LOD of *TP53* mutation analysis was set up to 1%, and the variants in

the range 1–3% were confirmed by replication.

### 3.4 Whole genome optical mapping

BM aspirates were obtained from an unselected cohort of 11 newly diagnosed MM patients with extramedullary disease ( $n = 4$ ) and without ( $n = 7$ ). PCs were enriched using an EasySep Human CD138 positive Selection Kit II (STEMCELL Technologies), according to the manufacturer’s instructions.

Isolation of PC high molecular weight genomic DNA from BM samples, labelling and analysis on Saphyr (Bionano) instrument was performed according to the manufacturer’s recommendations, targeting 100–300× human genome coverage.

### 3.5 Analysis of serum biomarkers in SLE

#### 3.5.1 Study subjects

Serum samples were obtained from 75 Czech SLE patients; all enrolled patients fulfilled the international ACR classification criteria [150]. The samples were aliquoted and stored at  $-80\text{ }^{\circ}\text{C}$  until further use. Organ damage was assessed by means of the SDI damage index (Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index) [108] and disease activity was evaluated by means of SLEDAI (Systemic Lupus Erythematosus Disease Activity Index) [151]. Subgroups were formed on the basis of (1) the SDI (SDI = 0,  $n = 33$ ; SDI = 1,  $n = 17$ ; SDI  $\geq 2$ ,  $n = 25$ ), (2) the biopsy-proven presence of LN (no LN,  $n = 48$ ; LN,  $n = 27$ ), and (3) the renal SLEDAI within LN subgroup, where renal SLEDAI score of  $\geq 4$  was taken as an indicator of active LN (inactive LN,  $n = 14$ ; active LN,  $n = 13$ ). The renal SLEDAI consists of the four renal parameters: hematuria, pyuria, proteinuria, and urinary casts [151]. The mean of LN duration in active LN patients was 7 years (range 0–19 years) and in inactive LN patients 8 years (range 1–18 years). The demographic and clinical features are described in Appendix F, Table 1. The age-matched control group of healthy subjects comprised 23 medical staff members (mean age 40, range 26–73, female/male 15/8), who gave statements about their



health status and excluded any medication used for SLE treatment (corticosteroids, antimalarials, immunosuppressant drugs).

### 3.5.2 Proximity extension immunoassay

The serum levels of 92 inflammation-related proteins were simultaneously measured by a PEA using the Proseek Multiplex Inflammation kit I (Olink Bioscience) according to the manufacturer's recommendation. Briefly, each analyte is recognized by a pair of oligonucleotide-labelled antibodies and when binding to their correct targets, they give rise to reporter amplicons which are amplified and quantified by microfluidic-based real-time PCR (BioMark HD System, Fluidigm Corporation). The data obtained is normalized and used for the relative quantification of the concentration of each analyte [127, 152].

### 3.5.3 Statistics

All statistical analyses were performed on linearized data (linear ddCq) for each analyte. Statistical tests (Mann-Whitney-Wilcoxon test, Benjamini-Hochberg correction, Spearman correlations, Receiver Operating Characteristic (ROC) curve analysis, and Bayesian probability model) were performed using the R statistical software with the Caret package. The p value for each protein was adjusted for multiple comparisons using the False Discovery Rate by the Benjamini-Hochberg procedure.  $P_{corr}$  value  $< 0.05$  was considered significant.

## 3.6 Gene expression biomarkers in autoimmune diseases

### 3.6.1 Study subjects

The study cohort consisted of 86 Caucasian patients with autoimmune diseases from a single rheumatology center in Olomouc, Czech Republic. All enrolled RA/SLE/SSc patients met the international 2010 ACR/EULAR classification criteria for RA [153], the ACR classification criteria for SLE [150], and the 2013 ACR/EULAR classification criteria for SSc, respectively [154]. To exclude heterogeneity due to the activity

and inactivity of the diseases, only cases with active phenotypes of the disease classified according to common activity scores (Disease Activity Score in 28 joints (DAS28), SLE Disease Activity Index (SLEDAI), and revised European Scleroderma Trials and Research group (EUSTAR) index) were included: RA ( $n = 36$ , DAS28  $\geq 3.2$ ), SLE ( $n = 28$ , SLEDAI  $> 6$ ), and SSc ( $n = 22$ , revised EUSTAR index  $> 2.25$ ).

The demographic and clinical features, used medication, duration of disease, and relative white blood count are described in Appendix E, Table 1. Distribution of lymphocyte, neutrophil, and monocyte counts did not differ between studied patient's groups ( $p > 0.05$ ). The healthy control cohort consisted of 77 subjects (mean age 51 years, min-max 24-90 years, female/male 58/19) out of which were formed three age-/gender-matched groups for each disease: 63 controls for RA (mean age 56 years, min-max 41-90 years, female/male 45/18), 33 controls for SLE (40, 24-50, 27/6, respectively), and 48 controls for SSc (58, 48-90, 34/14, respectively). In all healthy subjects, presence of inflammatory autoimmune diseases in first or second degree relatives, recent vaccination, infection, and usage of immunosuppressive drugs were excluded by questionnaire.

To further investigate the heterogeneity of the innate signature in only RA patients with a particular focus on active and inactive RA disease, subgroups were formed on the basis of the disease activity as assessed by means of the DAS28, with a DAS28 of  $\geq 3.2$  being taken as active RA (inactive RA,  $n=35$ ; active RA,  $n=32$ ).

### 3.6.2 Quantitative reverse transcription polymerase chain reaction

The PBMC were isolated from the peripheral blood by Ficoll density gradient centrifugation (Sigma-Aldrich) and stored in TRI Reagent (Sigma-Aldrich) at  $-80^{\circ}\text{C}$  until analysis. Total RNA was extracted using a Direct-zol RNA kit (Zymo Research) according to the manufacturer's recommendations. After reverse transcription with a Transcriptor First Strand cDNA Synthesis Kit (Roche), quantitative reverse transcription polymerase chain reaction (qPCR) was performed in a 100 nl

reaction volume containing a LightCycler 480 SYBR Green I Master mix (Roche) using a high-throughput SmartChip Real-Time-qPCR System (WaferGen) as reported previously [155]. Primer sequences are listed in [144] (Integrated DNA Technologies). The relative mRNA expression was calculated using phosphoglycerate kinase 1 as a reference gene [156].

In order to assess the innate immunity gene expression pattern, the expression of *TLR* (*TLR1–10*), *IL1/IL1R* family (21 members), and *CXCL8* was investigated in PBMC. Based on pilot evaluation of qPCR assays on a cohort of 20 RA, 20 SLE, and 20 SSc patients, 14 assays of *IL1/IL1R* family members (*IL1A*, *IL36RN*, *IL36A*, *IL36B*, *IL36G*, *IL37*, *IL38*, *IL33*, *IL1R2*, *IL18RAP*, *IL1RL1*, *IL1RL2*, *IL1RAPL1*, and *IL1RAPL2*) were below the limit of detection of the system and thus excluded from further analysis. The study continued therefore by expression profiling of 18 innate immunity genes: *TLR1–10*, 7 members of the *IL1/IL1R* family together with *CXCL8*.

### 3.6.3 Statistics and data-mining methods

Statistical analysis (Mann–Whitney U test, Benjamini–Hochberg correction) of relative gene expression values was performed using Genex (MultiD Analyses AB) and GraphPad Prism 5.01 (GraphPad Software). P value  $< 0.05$  was considered as significant.

In this study, a set of multivariate data-mining analyses to visualize and characterize the gene expression heterogeneity between and within the diseases was applied. For a flowchart of the analysis process used, see [144].

First, correlation networks using the LRNet algorithm [157] and Spearman's rank correlation coefficient were constructed and visualized to investigate the relationships between expressions of individual studied genes within the innate gene signature and to nominate the most representative molecules for the particular disease.

Second, Andrews curve analysis was applied for visualization of the structure

in multidimensional expression data [158–160]. The relative gene expression values of individual patients were transformed using Andrews’ formula; all calculations were performed by package “Andrews” from the R library. The Andrews curves were plotted to visualize the differences between particular diseases using a set of significantly deregulated genes and the whole set of studied genes. The difference is demonstrated by separation of the Andrews curve’s amplitudes and phase shift [158–160]. The curves of similar relative gene expression overlap between studied groups, while separation of curves demonstrates the differences in expression profiles [158–160].

Third, we applied association rule mining, a technique for finding frequent patterns, correlations, or associations among the given data set [161] to investigate the heterogeneity within the diseases themselves. Firstly, each gene data set was divided into low/high expression groups by arithmetic means of relative gene expressions within the whole data set. The applied package “arules” in the R system was used to extract rules (combinations of genes and its expression levels associated with the particular disease). Only a minimum number of top ranked rules describing the particular disease with a good confidence (threshold 0.75) and support were used.

All patients and controls included in all of the above studies provided written informed consent about the usage of biological materials for the purpose of this study, which was performed in accordance with the Helsinki Declaration and approved by the Ethics Committee of the University Hospital and Palacký University Olomouc.

## 4 RESULTS

### 4.1 Genetic biomarkers in haemato–oncology diseases

#### 4.1.1 Rare germline *ATM* variants in CLL

##### Prevalence and annotation of rare germline pathogenic/predicted pathogenic *ATM* variants in CLL

Four per cent of CLL patients carried rare, protein-altering germline *ATM* pathogenic/predicted pathogenic variants, as demonstrated in our patients (4.5%, 15/336, cohort A) and published datasets (3.4%, 15/445, cohort B, public datasets EGAD00001001464, EGAD00001001466). Of the detected rare *ATM* germline pathogenic/predicted pathogenic variants in CLL cohorts, 59% and 62% were classified as a VUS according to ClinVar and VarSome databases, respectively (Figure 6). In the ClinVar and VarSome databases, respectively, 14% and 10% of detected variants did not have rs numbers or were not listed. Besides *ATM* pathogenic/predicted pathogenic variants, rare variants with benign predictions were also detected in ~3% of CLL patients (A: 2.4%, 8/336; B: 3.6%, 16/445) and 1.5% of healthy controls (3/196).

*ATM* pathogenic/predicted pathogenic variants were associated with an increased risk of CLL (odds ratio 9.1, 95% CI: 1.2–69.5) when compared with healthy controls (1/196, 0.5%,  $p = 0.03$ ), but not with age at diagnosis and time to the first treatment when compared with CLL patients without these variants ( $p > 0.05$ ). All patients carried a single variant across the whole *ATM* gene with no apparent hotspot, except for one case with two variants. Variants were predominantly the missense type in both CLL cohorts (A, B) (69.0%, 20/29, Figure 6). Truncating *ATM* alleles (nonsense, frameshift) known to cause ataxia telangiectasia were less frequent (17.2%, 5/29). The MAF of *ATM* rare pathogenic/predicted pathogenic variants detected in CLL cohorts varied from extremely rare variants (singletons that were not found in population databases) to the highest MAF 0.005 observed in any population according to gnomAD. The majority (86%) of variants had  $MAF \leq 0.001$ ,

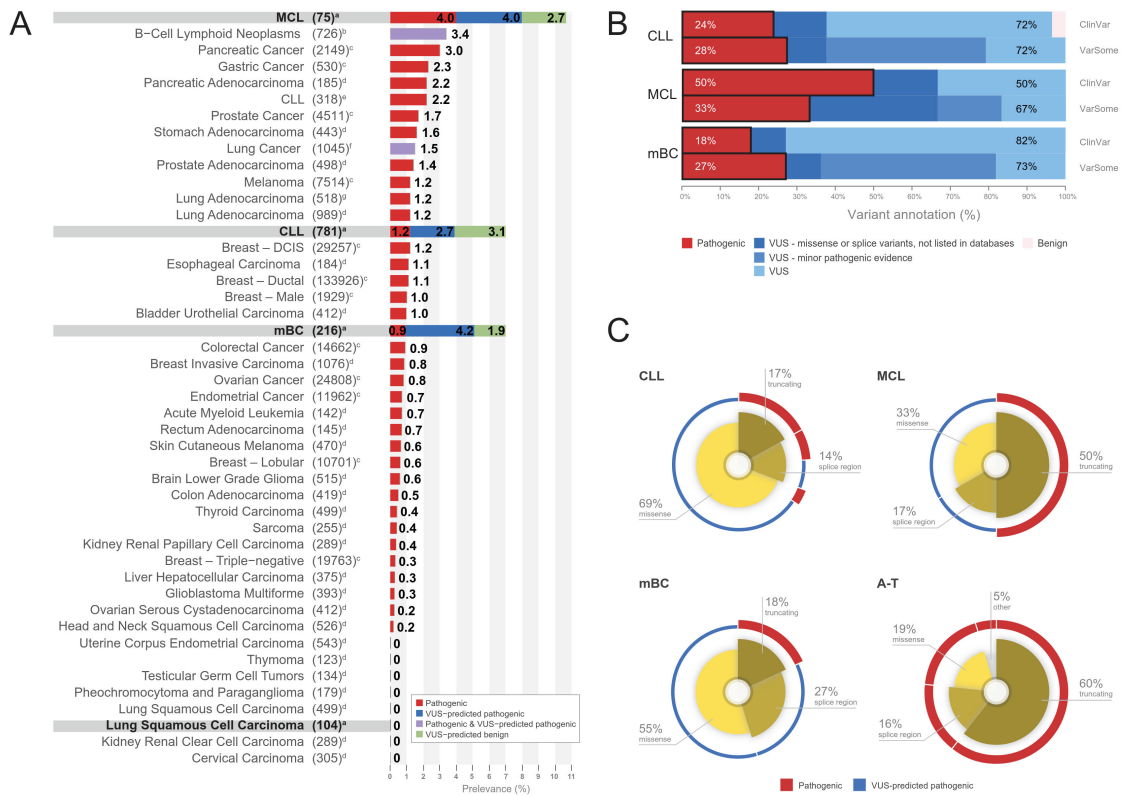


Figure 6: The prevalence and types of *ATM* rare germline variants in multiple cancers based on our data and published studies.

A) The prevalence of *ATM* rare germline variants across cancers. Patients with rare *ATM* germline pathogenic variants are red. Patients with *ATM* germline variants of uncertain significance (VUS)-predicted pathogenic variants are blue. Patients with either pathogenic or VUS-predicted pathogenic *ATM* rare variants are violet, as some studies combine these variants. Patients with *ATM* germline VUS-predicted benign variants are green. Reviewed cohorts and studies: (a) reanalysed public datasets, chronic lymphocytic leukemia (CLL) cohort A and cohort B are merged, (b) [262], (c) [61], (d) [261], (e) [62], (f) [260], (g) [53]. Four studies (c, d, e, g) excluded patients with VUS from analysis.

B) Annotation of rare germline *ATM* pathogenic/predicted pathogenic variants detected in CLL, metastatic breast cancer (mBC) and mantle cell lymphoma (MCL) by ClinVar and VarSome databases.

C) Spectrum of *ATM* rare germline pathogenic-predicted pathogenic variant types detected in CLL, MCL, mBC and ataxia telangiectasia (A-T). In CLL, cohort A and cohort B are merged.

of which 10 were singletons. A family history of cancer was higher in patients with rare *ATM* pathogenic/predicted pathogenic variants (87.5%) when compared with patients with wild type *ATM* (43.1%,  $p = 0.02$ ), based on the available self-reports from CLL cohort A.

### Rare germline pathogenic/predicted pathogenic *ATM* variants in multiple cancers

In other analysed cancers, rare *ATM* pathogenic/predicted pathogenic variants occurred in 8% of MCL (6/75, public dataset EGAD00001006159), 5% of metastatic breast cancer (11/216, EGAD00001002747) and 6% of lung cancer (2/36, EGAD00001004027, cancer subtype undefined) patients (Figure 6). In patients with squamous cell lung cancer, rare *ATM* pathogenic/predicted pathogenic variants were not detected (0/104, EGAD00001003960). All rare germline *ATM* pathogenic/predicted pathogenic variants in all cohorts evaluated were detected as heterozygous in the paired germline sample, and each individual had a different variant.

Additionally, we systematically reviewed the prevalence of rare germline *ATM* pathogenic/predicted pathogenic variants in multiple cancers and compared it with our data (Figure 6).

### Association of rare germline pathogenic/predicted pathogenic *ATM* variants with adverse somatic aberrations in CLL

Patients with rare germline *ATM* pathogenic/predicted pathogenic variants also had del(11q), 40% (6/15) in our CLL cohort A, and 27% (4/15) in cohort B. In both CLL cohorts, patients with rare germline *ATM* pathogenic/predicted pathogenic variants were  $\sim 3$  times more likely to have del(11q), always with a loss of the wild type allele, when compared with patients with wild type *ATM* and/or *ATM* somatic mutations (A: relative risk: 2.46, 95% CI: 1.25–4.84,  $p = 0.009$ ; B: 3.87, 1.55–9.66,  $p = 0.004$ ). In patients with del(11q), the median VAF of rare germline *ATM* variant in tumour samples reached 84% (A, min–max 50–93%) and 71% (B, 52–100%), respectively. Rare *ATM* variants with a benign prediction did not associate with the del(11q) ( $p > 0.05$ ). Similar to CLL, 83% (5/6) of MCL and 56% (6/11) of metastatic breast cancer patients with rare germline *ATM* pathogenic/predicted pathogenic variants had concurrent del(11q), as detected by CNA analysis in WES datasets.

Regarding somatic *ATM* variants, the vast majority of patients with rare *ATM*

pathogenic/predicted pathogenic variants did not have additional somatic *ATM* variants in both CLL cohorts (A: 93.3%, 14/15; B: 93.3%, 14/15) (Appendix A, Figure 2). Rare *ATM* pathogenic/predicted pathogenic variants were different from somatic *ATM* variants, except for one variant type (p.C2488Y) (Appendix A, Figure 3).

Regarding the association with *IGHV* status, rare germline *ATM* pathogenic/predicted pathogenic variants associated with unmutated *IGHV* that was detected in 93.3% of patients (14/15) in cohort A (A: relative risk: 1.50, 95% CI: 1.28–1.77,  $p < 0.0001$ , Appendix A, Figure 2). However, this association was not observed in cohort B when all variants, irrespective of position and variant type, were evaluated ( $p > 0.05$ ); 40% (6/15) of patients carrying rare germline *ATM* pathogenic/predicted pathogenic variants had unmutated *IGHV* status. Subanalysis in patients carrying only truncating or missense variants in the *ATM* kinase domain (aa2712-2962) (A:  $n = 5$ , B:  $n = 4$ ) revealed that all these patients developed CLL with unmutated *IGHV* (A: relative risk: 1.62, 95% CI: 1.49–1.77,  $p < 0.0001$ ; B: relative risk: 2.87, 95% CI: 2.52–3.27,  $p < 0.0001$ ).

Regarding *TP53* disruption (del(17p) and/or *TP53* mutation), the majority of patients carrying rare germline *ATM* pathogenic/predicted pathogenic variants did not acquire a *TP53* disruption (A: 80.0%, 12/15; B: 86.7%, 13/15; Appendix A, Figure 2). Of note, two patients harbouring rare germline *ATM* pathogenic/predicted pathogenic variants who, during the disease course, also acquired del(11q), del(17p) and *TP53* mutations were refractory to idelalisib–rituximab, ibrutinib and venetoclax, and one of those patients had a history of breast cancer and developed a RT eight months later.

*ATM* rare pathogenic/predicted pathogenic variants were not associated with other adverse somatic aberrations (trisomy 12, complex karyotype, mutated *SF3B1*, *BIRC3* and *NOTCH1*;  $p > 0.05$ ) in our CLL patients.



## Functional analysis of rare germline *ATM* pathogenic/predicted pathogenic variants in CLL

To evaluate the functional impact of rare germline *ATM* pathogenic/predicted pathogenic variants, we analysed a decrease in ATM activity in CLL and T cells obtained from 25 CLL patients, as assessed by KAP1 phosphorylation and ATM autophosphorylation of Ser-1981 (Figure 7).

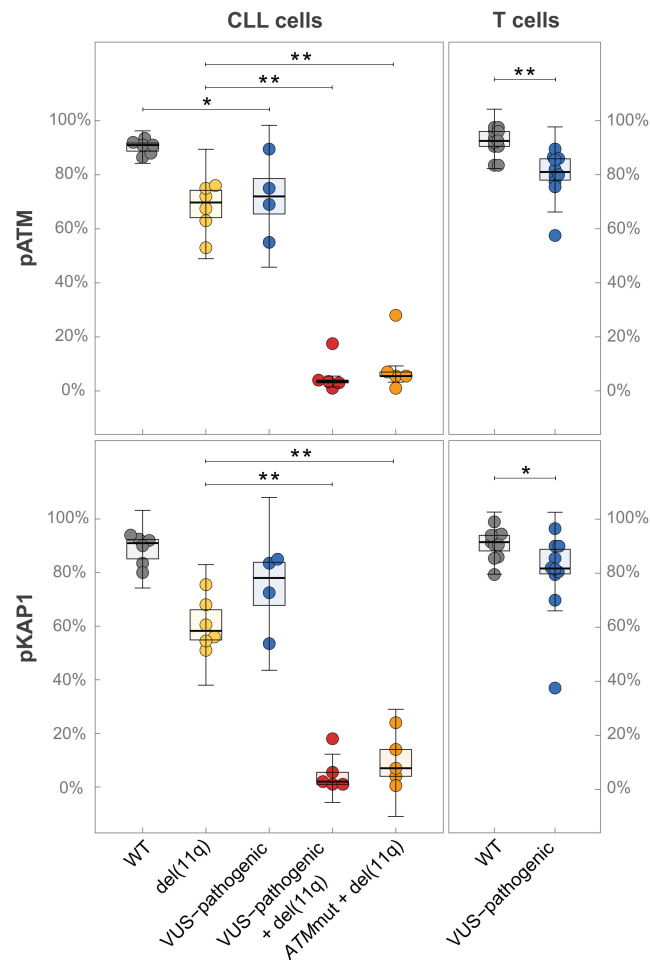


Figure 7: ATM activity in CLL and T cells obtained from chronic lymphocytic leukemia (CLL) patients with rare germline *ATM* pathogenic/predicted pathogenic variants, *ATM* wild type (WT) and somatic *ATM* disruption.

In a few patients with a high percentage of CLL cells in the peripheral blood, T cell analysis was not performed due to insufficient cell count.

All analysed rare germline *ATM* pathogenic/predicted pathogenic variants, except for one variant (R2443\*), were annotated by ClinVar and VarSome as VUS (missense variants: K224E, L480F, R717P, Y1442H, R2032K, Y2755S; splice region

variant: c.4236+4A>C). For comparison, the following somatic *ATM* variants (VAF min–max 8–94%) were analysed: L439H, L1283fs, C1674W, L2698F, Q2714del and T2773I. In samples with del(11q), this aberration was present in more than 50% of CLL cells as assessed by FISH analysis.

CLL cells from patients carrying rare *ATM* germline pathogenic/predicted pathogenic variants, together with del(11q), completely lost ATM activity, same as in patients with concurrent somatic *ATM* pathogenic variant and del(11q). Heterozygous rare germline *ATM* variants pathogenic/predicted pathogenic alone resulted in only a partial decrease in ATM activity (reduction of 10–20%) in both CLL and T cells, and the activity varied depending on the individual variant type (Figure 7).

### **Progression–free survival of treated CLL patients carrying rare germline *ATM* pathogenic/predicted pathogenic variants**

We evaluated the impact of rare germline *ATM* pathogenic/predicted pathogenic variants on progression–free survival (PFS) of patients treated with novel agents and chemoimmunotherapy. Patients on novel agents carrying rare germline *ATM* pathogenic/predicted pathogenic variants had shorter PFS (median 24 months) than patients with wild type *ATM* and/or somatic *ATM* and/or *TP53* disruption (49/40 months,  $p < 0.05$ ) and did not differ from patients with *ATM* and/or *TP53* somatic disruption (40 months,  $p > 0.05$ ) (Figure 8). Similar results were obtained for patients on chemoimmunotherapy. Subanalysis in only patients with unmutated *IGHV* revealed the same negative association of rare germline *ATM* pathogenic/predicted pathogenic variants on PFS in patients treated with novel agents (8).

#### **4.1.2 Applicability of the whole bone marrow for analysis of *TP53* mutations in MM**

Only in three–quarters of diagnostic BM samples was it possible to enrich with at least 100 000 cells needed to perform deep targeted NGS in our study. Of the 54 BM samples (median (range) 4.5 (1.0–12.0) ml) with PC infiltration (10.5%, range 2.0–67.0%), 72.9% samples were successfully enriched (PC purity: 90.9%, range

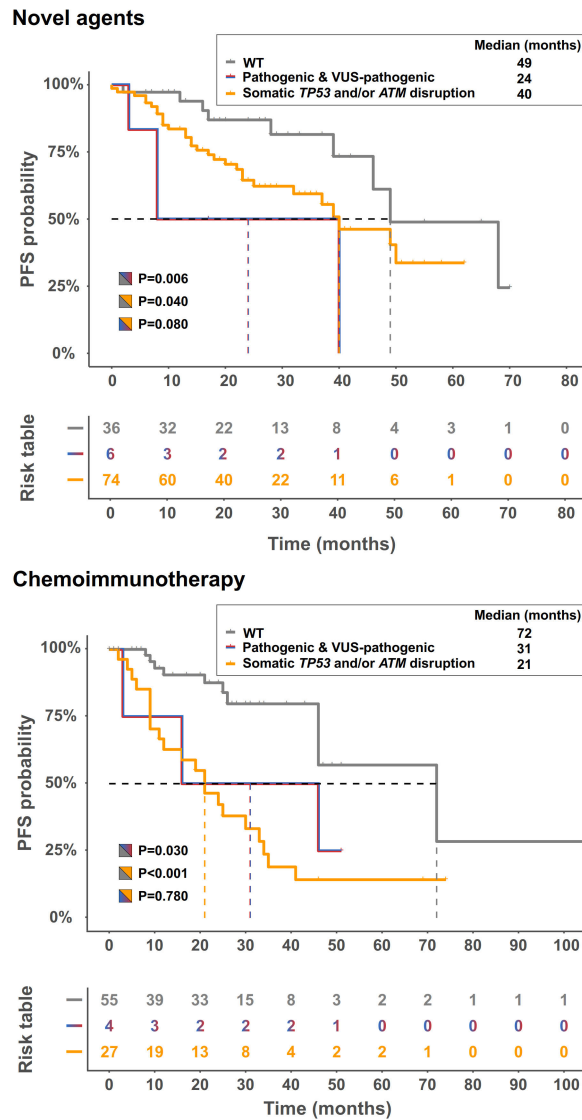


Figure 8: Progression-free survival of patients with chronic lymphocytic leukemia (CLL) treated with novel agents and chemoimmunotherapy stratified by the presence of rare germline *ATM* pathogenic/predicted pathogenic variants and somatic *ATM* and/or *TP53* disruption.

39.0–99.1%; PC absolute counts:  $1.0 \times 10^6$ , range  $0.1\text{--}35.30 \times 10^6$ ; yield of PC: 32.2%, range 7.0–100%). Lower PC recovery was observed using Ficoll gradient separation (24.1%, range 7.3–46.5%) compared to red blood cell lysis (43.0%, range 6.7–95.2%). Enrichment in 27.1% samples failed due to low recovery and/or sample amount.

To evaluate the utility of whole BM for mutation NGS analysis, we investigated the whole BM and matched enriched PCs ( $n = 27$ ). There was concordance between *TP53*, *NRAS*, *KRAS*, and *BRAF* mutations in 85.2% patients (Figure 9). Four

samples, in which mutations were found only in enriched samples, had PC infiltration in BM of 4.0%, 7.0%, 20.8%, and 26.5%. The VAF of detected mutations from the whole BM corresponded or were lower than the PC infiltration in BM. Overall, the enrichment of PCs increased the mutation VAF on average 23.3% (range 3.0–53.0%), which is not proportional to the increase of the tumour PCs fraction in the sample achieved by PC enrichment (on average 70.7%) (Figure 9).

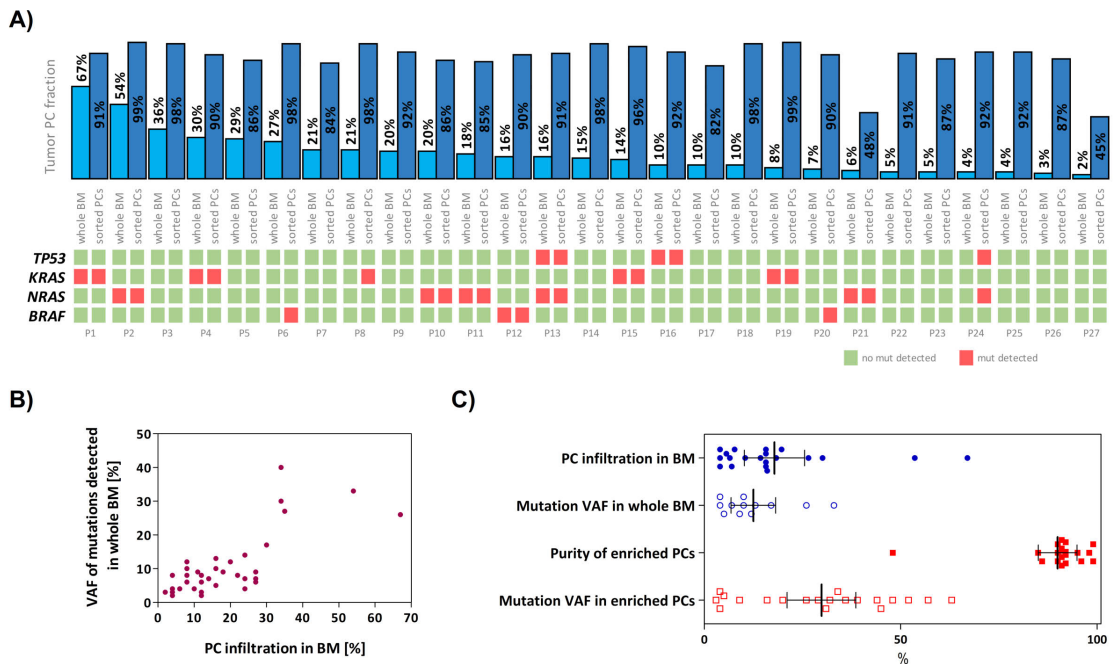


Figure 9: A) Concordance of mutations found in matched whole bone marrow (BM) and paired enriched tumour plasma cell (PC) samples. Heatmap represents detection of individual mutations in a series of paired whole BM and enriched tumour PC samples. Columns represent infiltration of tumour PCs in whole BM and purity of enriched PCs. B) Relationship between variant allele frequency (VAF) of detected mutations from the whole BM and the PC infiltration in the BM. C) Comparison of VAF of mutations found in matched whole BM and enriched tumour PC samples and their relationship to the PC fraction in the samples.

In our present cohort, 28% (15/54) of patients had *TP53* disruption and of them three patients at diagnosis (one only *TP53* mutation, two patients del(17p)). Only *TP53* mutations were detected in 47% (7/15), *TP53* mutations together with del(17p) in 33% (5/15) and only del(17p) in 20% (3/15) of patients with *TP53* disruption. Of 22% (12/54) of patients carrying *TP53* mutations, five had  $\geq 2$  (maximum 7) *TP53* mutations. Additionally, three patients at diagnosis had chro-

mosome 17 trisomy with FISH confirmation of three *TP53* gene copies, of them one patient also carried *TP53* mutation. Regarding other genes, *KRAS* mutations were detected in 18.5% (10/54) patients, *NRAS* mutations in 13.0% (7/54), and *BRAF* mutations in 9.3% (5/54). Mutations in *KRAS* and *NRAS* were mutually exclusive.

### 4.1.3 Standardization of sequencing coverage depth in diagnostic NGS

#### NGS sequencing depth and error rate

NGS sequencing depth directly affects the reproducibility of variant detection: the higher the number of aligned sequence reads, the higher the confidence to the base call at a particular position, regardless of whether the base call is the same as the reference base or is mutated [90]. In other words, individual sequencing error reads are statistically irrelevant when they are outnumbered by correct reads. Thus, the desired coverage depth should be determined based on the intended LOD, the tolerance for false positive or false negative results, and the error rate of sequencing [90, 100].

Using a binomial distribution, the probability of false positive and false negative results for a given error rate as well as the intended LOD can be calculated, and the threshold for a variant calling for a given depth can be estimated [90]. For example, given a sequencing error rate of 1%, a mutant allele burden of 10%, and a depth of coverage 250 reads, the probability of detecting 9 or fewer mutated reads is, according to the binomial distribution, 0.01%. Hence, the probability of detecting 10 or more mutated reads is 99.99% (100–0.01%), and the threshold for a variant calling can be defined. In other words, a coverage depth of 250 with a threshold of at least 10 mutated reads will have a 99.99% probability that 10% of the mutant allele load will not be missed by the variant calling (although it can be detected in a different proportion). In this way, the risk of a false negative result is greatly minimized. On the other hand, the probability of false positives heavily depends on the sequencing error rate (as the accuracy of all analytical measurements depends on the signal-to-noise ratio) [90, 100]. In our example, the probability of a false positive

result is 0.025%; however, the rate of false positives is not negligible when decreasing the LOD to the value close to the error rate. Conventional intrinsic NGS error rates range between 0.1 and 1% (Phred quality score of 20–30) [90,100] depending on the sequencing platform, the GC content of the target regions [162], and the fragment length, as shown in Illumina paired-end sequencing [163]. Therefore, the detection of variants at VAFs <2% is affected by a high risk of a false positive result, regardless of the coverage depth. It is also important to mention that the sequencing error rate applies only for errors produced by sequencing itself and does not include other errors introduced during DNA processing and library preparation, particularly during amplification steps, which further increase error rates [90,100].

### Minimum sequencing coverage in clinical settings

According to the binominal data distribution, a coverage depth of 250 should indeed be sufficient to detect 5% VAF with a threshold of variant supporting reads  $\geq 5$  (Figure 10). On the other hand, NGS analysis with a coverage depth of 100 along with a requirement of at least 10 variant supporting reads as recommended by the ERIC consortium [99] would result in a false negative of 45% for samples with a LOD of 10%. To confirm these theoretic calculations, we performed two independent dilution experiments to estimate the performance of *TP53* NGS analysis to detect 10% VAF at a depth of coverage of 100 reads. Indeed, we detected 30% of false negatives (5 positive samples of 7 true-positive samples and 9 positive samples of 13 true-positive samples) in two independent sequencing runs.

### Frequency of *TP53* subclonal mutations in CLL detected through diagnostic NGS

In order to evaluate the occurrence of low VAF in real-world settings, we reviewed our cohort of CLL patients examined for *TP53* mutations in our diagnostic laboratory. Of the diagnostic cohort of 859 CLL patients, 25% (215/859) were positive for *TP53* mutations, and of those, 53% (113/215) carried variants with VAF at 10% or lower.

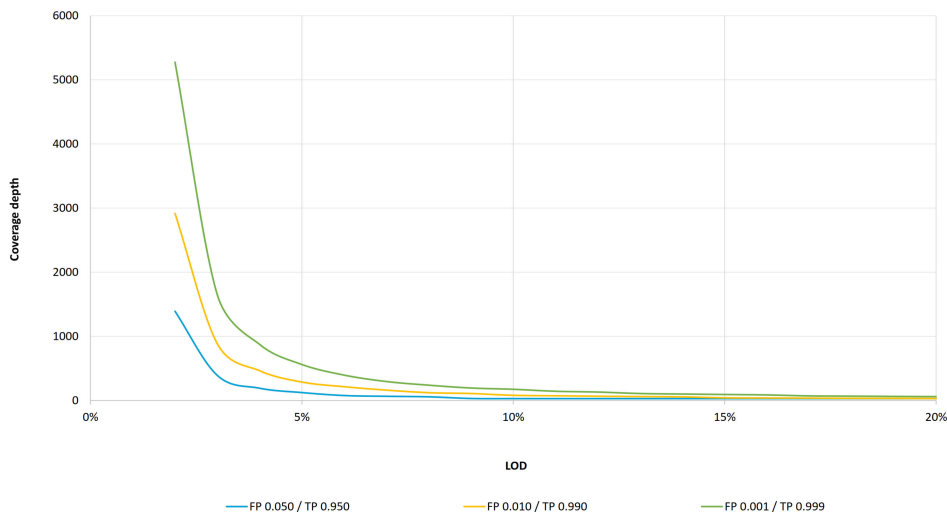


Figure 10: Limit of detection (LOD) as a function of coverage depth according to the binomial distribution.

Coverage depth needed to maintain an intended LOD (within 3–20% VAF range) for three cumulative probability settings: for false positive probability of 0.001 and true positive of 0.999, a LOD of 20% is achieved at 61 coverage depth, a LOD of 10% at 175, a LOD of 5% at 562, and a LOD of 3% at 1,650. For the false positive probability of 0.010 and true positive of 0.990, a LOD of 20% is achieved at 31, a LOD of 10% at 81, a LOD of 5% at 288, and a LOD of 3% at 886 coverage depth, respectively. For the false positive probability of 0.050 and true positive of 0.950, a LOD of 20% is achieved at 30, a LOD of 10% at 30, a LOD of 5% at 124, and a LOD of 3% at 392 coverage depth, respectively.

### Calculator for diagnostic NGS settings for detection of subclonal mutations

To assist laboratories with the determination of the minimum proper coverage parameters, we are providing a simple, user-friendly theoretical calculator (software) based on the binomial distribution (Figure 11). A web (or desktop) application and stand-alone source codes in R are accessible on Github: <https://github.com/mvasinek/olgen-coverage-limit>. Using this calculator, the correct parameters of sequencing depth and the corresponding minimum number of variant reads for a given sequencing error rate and intended LOD can easily be determined. Moreover, users can also take into account other errors by simply adding assay-specific errors to the sequencing error rate and using this overall error as an input to the calculator. For example, in our case of *TP53* mutational analysis we calculated with the overall

A OLGEN Coverage Limit Calculator	B OLGEN Coverage Limit Calculator
Variant allele frequency (%) 10.0 ?	Variant allele frequency (%) 3.0 ?
Sequencing error rate (%) 1.0 ?	Sequencing error rate (%) 1.0 ?
Probability of false positive result (%) 0.1 ?	Probability of false positive result (%) 0.1 ?
Probability of true positive result (%) 99.9 ?	Probability of true positive result (%) 99.9 ?
Minimum of variant reads (optional) ? ?	Minimum of variant reads (optional) ? ?
<a href="#">Calculate coverage</a>	<a href="#">Calculate coverage</a>
<b>Recommended coverage: 175</b>	<b>Recommended coverage: 1650</b>
<b>Minimum of variant reads: 7</b>	<b>Minimum of variant reads: 30</b>

Figure 11: OLGEN Coverage Limit calculator – a simple theoretical calculator suitable for determining the correct sequencing depth and corresponding minimum number of variant reads according to the binomial distribution for a given sequencing error rate and intended LOD recommended for diagnostic NGS.

Examples of calculated sequencing depths and the corresponding minimum number of variant reads recommended for variants with A) 10% VAF and 99.9% probability of detection and B) 3% VAF and 99.9% probability of detection.

error of  $\sim 1.16\%$ , thus we set up our minimum coverage depth requirements to 2,000 with threshold of minimum 40 reads for 3% VAF.

#### 4.1.4 *TP53* as biomarker in MCL

Our research group demonstrated the prognostic impact of somatic mutations in *TP53* gene in two real-world cohorts of MCL patients using NGS [75, 76] (Appendices H and I). We revealed that i) a high *TP53* mutation burden served as a predictive biomarker of chemoresistance in younger patients with newly diagnosed MCL regardless of the routinely used treatment strategy and ii) concurrent aberration of *TP53* (deletion and/or mutation) and deletion of *CDKN2A* gene represented the most significant predictive biomarkers of short event-free survival and OS in patients with MCL.



#### 4.1.5 Whole-genome optical mapping for structural variant analysis in MM

Our research group performed a pilot study on utility of whole-genome optical mapping to analyse the genomic architecture of extramedullary multiple myeloma (EMM) [104] (Appendix G). Our study revealed an association of chromosome 1 abnormalities in BM MM cells with extramedullary progression, nominating it as a candidate biomarker for EMM [104].

## 4.2 Serum biomarkers in SLE

In order to assess the serum protein pattern associated with SLE, we compared the serum protein levels obtained by PEA immunoassay in the SLE patients and healthy controls. Of 92 biomarkers that were analyzed, the levels of 14 analytes (IL1A, IL2, sIL2RB, IL4, IL5, IL13, IL20, sIL20RA, IL33, TSLP, ARTN, TNF, LIF, NRTN) were below the LOD in our sample set and therefore they were excluded from further analysis. Comparing SLE and the controls, 29 proteins were upregulated and sDNER downregulated in SLE ( $p_{corr} < 0.05$ ; Appendix F, Table 2). The distribution of the serum levels of top-upregulated proteins (sirtuin 2, IL18, caspase 8, sCD40/sTNFRSF5, sSLAMF1, sTNFRSF9, axin 1, sulfotransferase 1A1, STAMBP, CCL19/MIP-3 $\beta$ , IL10, and CCL4/MIP-1 $\beta$ ;  $p_{corr} < 0.003$ ) is shown in Appendix F, Figure 1. For the serum protein pattern associated with SLE and the changes in protein levels between SLE and the controls for top-deregulated analytes see Figure 12.

Because of the suggested central role of the IFN pathway in SLE pathogenesis by promoting feedback loops progressively disrupting peripheral immune tolerance and driving disease activity, we investigated the IFN protein signature of nine IFN-regulated cytokines. Because of the reported association of an increased IFN gene expression signature with disease activity [164, 165], we performed correlation analysis among the protein levels of IFN-regulated chemokines and disease activity as assessed by SLEDAI. The analysis revealed elevation of six IFN-regulated

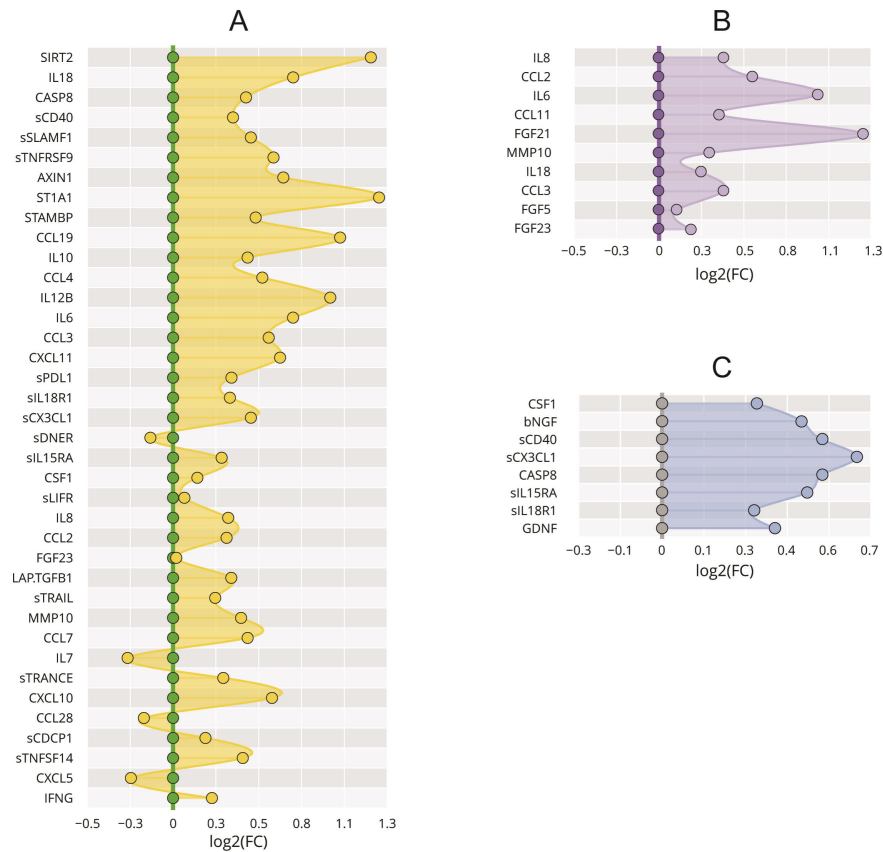


Figure 12: Protein serum fingerprints associated with A) SLE, B) organ damage, and C) active lupus nephritis (LN). Fingerprints are presented as FC (fold change of group medians) of serum levels of all deregulated serum proteins between particular groups.

cytokines (IL6, CCL2/MCP1, CCL3/MIP-1 $\alpha$ , sCD40, CXCL11, and CCL19;  $p_{corr} \leq 0.01$ ) in SLE and three (CCL8/MCP2, CXCL9, and CXCL10) did not reach significance ( $p_{corr} > 0.05$ ). Interestingly, only a mild positive correlation ( $r = 0.25$ ,  $p = 0.03$ ) was observed between the levels of IFN-regulated chemokines and disease activity as assessed by SLEDAI. Disease activity assessed by SLEDAI correlated better with the following analytes: IL8, GDNF, CX3CL1/fractalkine ( $r \geq 0.403$ ,  $p \leq 0.0003$ ), and CCL7/MCP3, IL15RA, VEGFA, and MMP10 ( $r \geq 0.355$ ,  $p \leq 0.002$ ).

### Protein pattern of organ damage

To obtain the protein pattern associated with organ damage, we compared the serum patterns from SLE patients with/without organ damage and subgroups ac-

cording to the SDI ( $\text{SDI} \geq 2/\text{SDI} = 1/\text{SDI} = 0$ ). In the patients with organ damage ( $\text{SDI} \geq 1$ ), elevated serum levels of IL8, CCL2, IL6, CCL11/eotaxin, FGF21, MMP10, IL18, CCL3, FGF5, and FGF23 ( $p_{\text{corr}} < 0.05$ ) were detected (Appendix F, Table 2, Figure 4). The serum protein pattern associated with organ damage and the changes in protein levels between SLE patients with/without organ damage are shown in Figure 12. Although the serum level of CCL11 did not differ between the controls and SLE patients as a whole, the patients with organ damage had higher levels of CCL11 in comparison to those with no organ damage, as well as to the control group. We did not observe differences in serum protein pattern between patients with  $\text{SDI} = 1$  and  $\text{SDI} \geq 2$  ( $p_{\text{corr}} > 0.05$ ).

Among organ damage associated analytes, the cumulative dose of glucocorticoids correlated positively with levels of IL8, CCL11 ( $r \geq 0.326$ ,  $p \leq 0.004$ ), CCL2 and MMP10 ( $r \geq 0.249$ ,  $p < 0.05$ ). Additionally, cumulative dose of glucocorticoids correlated with BDNF, CCL25, CXCL1, GDNF, IL17C, sADA, sCD137, sIL18R1, sSCF, and sTGF $\alpha$  ( $p < 0.05$ ). Moreover, IL8 ( $r = 0.416$ ,  $P = 0.0002$ ), MMP10 ( $r = 0.355$ ,  $p = 0.002$ ), CCL2, and CCL11 ( $r \geq 0.261$ ,  $p \leq 0.02$ ) correlated positively with disease activity. In line with other reports, a higher cumulative dosage of glucocorticoids was registered in the patients with  $\text{SDI} \geq 1$  (mean of 30.6 g, min–max 2.6–79.2 g) compared with those without damage (12.8, 0–54.0). Regarding association of disease duration and serum levels of studied proteins, we observed only mild association for CCL11 ( $r = 0.230$ ,  $p = 0.047$ ). The disease duration in SLE patients correlated with SDI ( $r = 0.298$ ,  $p = 0.009$ ).

### **Protein pattern of active lupus nephritis and other clinical subsets**

To investigate the serum patterns associated with active LN, we compared subgroups of SLE patients with/without biopsy–proven LN and subgroups of patients with LN classified by the renal SLEDAI as active (renal SLEDAI  $\geq 4$ ) or inactive renal disease at the day of sampling. Moreover, we assessed serum patterns associated with other clinical subsets of SLE as neurological, haematological, cardiovascular, skin and musculoskeletal involvements, antiphospholipid syndrome, and

renal disorder.

The analysis in biopsy-proven LN patients with active renal disease revealed elevated protein levels of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF compared to those without LN (Appendix F, Table 2, Figure 5). Although the serum levels of GDNF did not differ between the control group and SLE as a whole, its level was enhanced in the patients with LN in comparison to those without LN and the control group. The serum protein pattern associated with active LN and the changes in protein levels between the SLE patients without LN and active LN are shown in Figure 12.

When LN patients with active renal disease was compared to inactive LN subgroup, elevation of sIL15RA, CSF1, bNGF, sIL18R1, sCD40, sCX3CL1, and caspase 8 ( $p < 0.05$ , Appendix F, Table 2, Figure 5), but not GDNF, in active LN patients was observed.

In the other studied clinical subsets no differences in the serum pattern were detected. The subanalysis confirmed that no candidate biomarker for SLE, organ damage and/or LN are influenced by the gender ( $p_{corr} > 0.05$ ).

To investigate the utility of the serum levels of phenotype-associated proteins for the identification of patients with a high probability of severe phenotypes, we constructed probability plots for phenotype-associated proteins based on a Bayesian statistical approach. Additionally, we constructed ROC (receiver operating characteristic curve) curves for the proteins associated with organ damage and active LN.

In organ damage, the best predictive model was observed for the serum levels of CCL11 and MMP10, followed by CCL2, whereas IL6 and IL8 were not informative (Figure 13). Higher serum levels of CCL11 and MMP10 correspond to a higher probability of organ damage. For the analytes associated with organ damage, the ROC curve analysis showed that the area under the curve (AUC) of IL8, CCL2, IL6, CCL11, FGF21, MMP10, IL18, CCL3, FGF5, and FGF23 was 0.784, 0.738, 0.731, 0.727, 0.723, 0.706, 0.697, 0.691, 0.689, and 0.676, respectively.

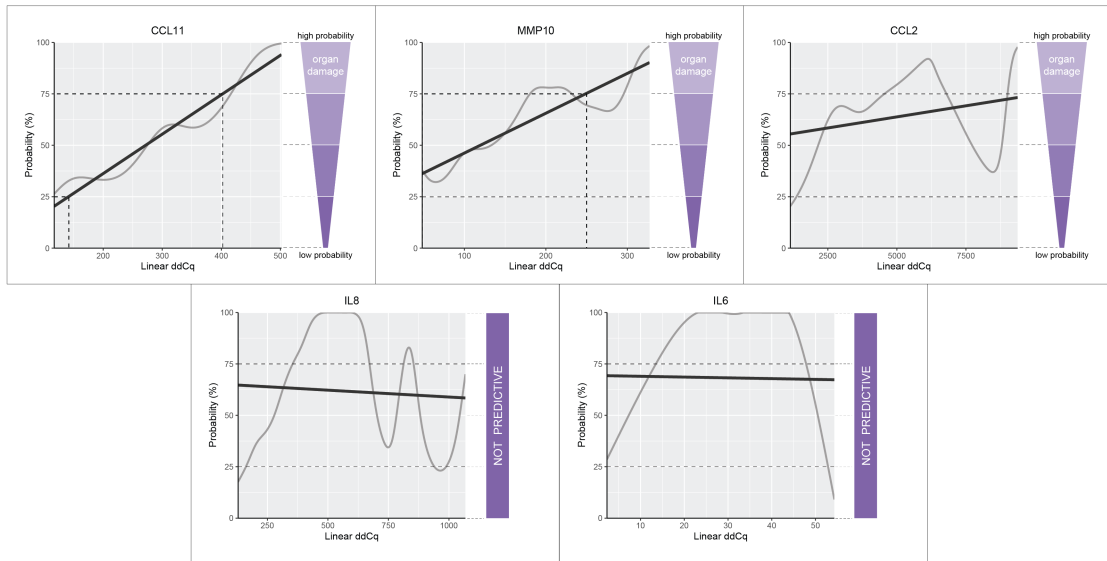


Figure 13: Probability plots of serum analytes associated with organ damage in SLE patients.

The grey curve represents a simulated model based on the individual patient serum levels and the black line represents overall trend calculated by the Bayesian statistical approach. The increasing overall trend the higher probability of organ damage. Higher serum levels of CCL11 and MMP10 correspond to higher probability of organ damage, lower serum levels of these analytes to lower probability of organ damage. IL8 and IL6 serum levels were not informative for organ damage prediction.

In active LN, the best predictive value was observed for CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, and sIL18R1 (Figure 14). Higher serum levels of all these analytes correspond to a higher probability of the presence of active LN. For the analytes associated with active LN, the ROC curve analysis showed that the AUC of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF were 0.873, 0.857, 0.854, 0.832, 0.798, 0.783, 0.780, and 0.778, respectively. Moreover, we observed great sensitivity and specificity for proteins sIL15RA (AUC: 0.879, sensitivity: 100%, specificity: 64.3%), CSF1 (0.813, 84.6, 78.6), sIL18R1 (0.810, 84.6, 78.6), and bNGF (0.805, 69.2, 100) showing good discrimination between active and inactive renal disease in LN patient subgroup. Inactive LN patients do not differ from patients without LN, except for GDNF (Appendix F, Figure 5), suggesting that serum GDNF level remains elevated even when LN is inactive.

All nominated biomarkers associated with organ damage and active LN showed better discrimination ability in our cohort than the classical markers. The only

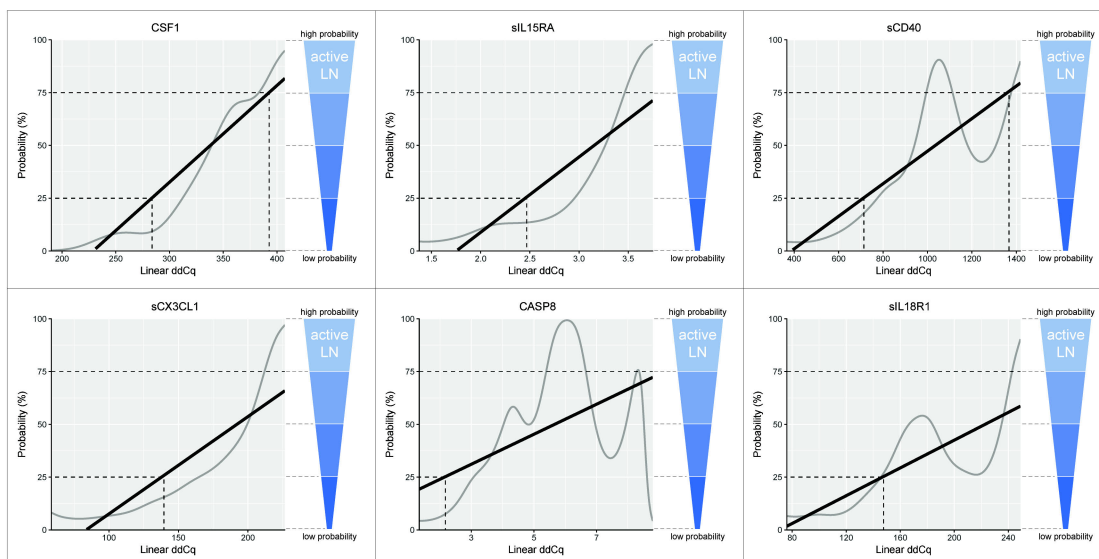


Figure 14: Probability plots of serum analytes associated with active lupus nephritis (LN) in SLE patients.

The grey curve represents a simulated model based on the individual patient serum levels and the black line represents overall trend calculated by the Bayesian statistical approach. The increasing overall trend the higher probability of active LN. Higher serum levels correspond to higher probability of active LN, lower serum levels of these analytes to lower probability of active LN. The best predictive value was observed for CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8 (CASP8), and sIL18R1.

exception was proteinuria (AUC 0.869), one of the criteria for renal SLEDAI classification.

### 4.3 Gene expression biomarkers in autoimmune diseases

In order to characterize innate immune signature in studied diseases, the expression profiles of selected innate immune genes between patients and healthy controls in all diseases were compared.

RA differed from controls by the upregulated expression of *TLR2*, *TLR3*, *TLR5*, *TLR8*, *IL1B*, *IL18*, *IL18R1*, *IL1RN*, *IL1RAP*, and *SIGIRR/IL1R8* ( $p_{corr} \leq 0.05$ ). In patients treated with anti-TNF- $\alpha$  therapy, a trend to lower *TLR5* levels in our RA patients was observed ( $p = 0.07$ ). In SLE, downregulation of *TLR10* was observed when compared to healthy controls ( $p = 0.02$ ); however, it did not reach significance after the correction for multiple comparisons. SSc differed from controls by the upregulated expression of *IL1RN*, *IL18*, and *CXCL8* and downregulated expression

of *IL1RAP* and *IL18R1* ( $p_{corr} \leq 0.05$ ).

To investigate the disease-specific innate immune gene expression pattern, we compared RA, SLE, and SSc patients to each other. RA differed from SLE and SSc by the upregulated expression of *TLR5* and *SIGIRR* ( $p_{corr} < 0.02$ ). RA further differed from SLE by the upregulated expression of *TLR2* ( $p_{corr} = 0.02$ ) and from SSc by the upregulation of *TLR3*, *IL1RAP*, and *IL18R1* genes ( $p_{corr} < 0.007$ ). In SSc, the upregulated expression of *IL1R1* ( $p_{corr} = 0.005$ ) was observed when compared to SLE (Figure 15, Appendix E, Table 2).

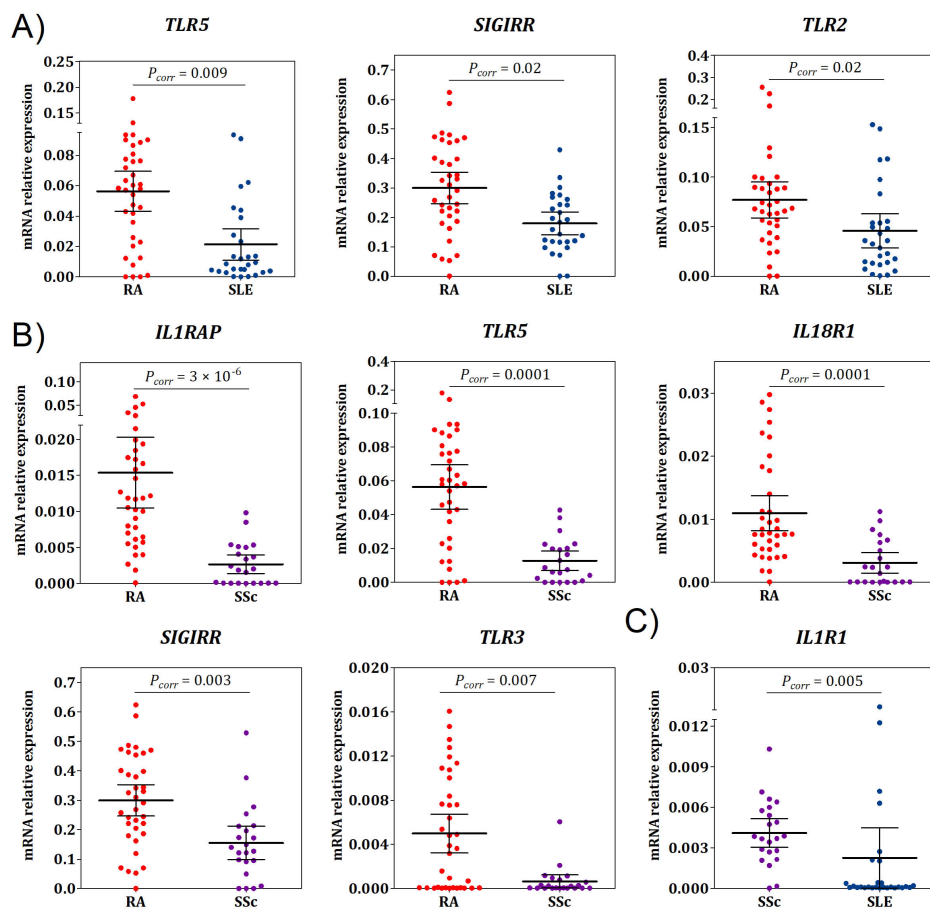


Figure 15: Relative mRNA expression levels of genes differentially expressed in A) RA vs. SLE, B) RA vs. SSc, and C) SSc vs. SLE. Group means are indicated by horizontal bars; error bars indicate 95% CI.

To investigate the disease-associated gene expression pattern, Andrews curves were used to visualize the differences between particular diseases using a set of significantly deregulated genes and the whole set of studied genes. First, we assessed

the differences in the innate expression pattern of genes revealed by classical statistics. Although a good separation of Andrews curves on the basis of significant genes was observed, better separation of the studied diseases was obtained when a whole set of studied genes was used (Figure 16).

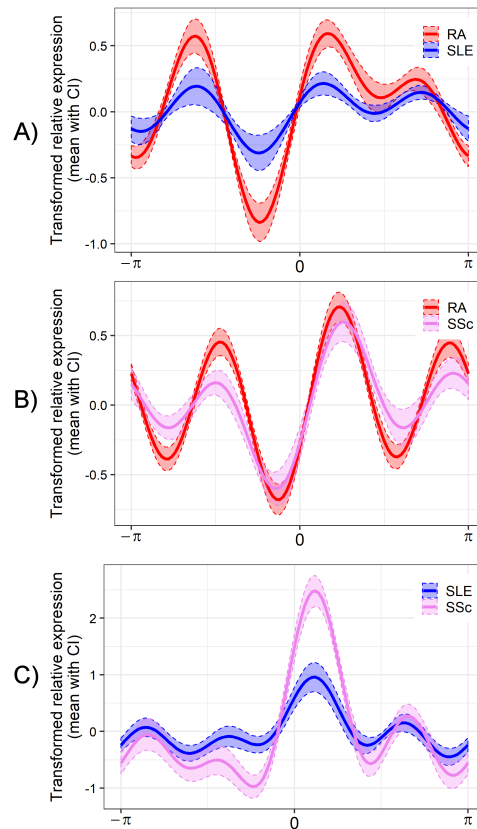


Figure 16: Differential innate gene expression analysis by Andrews curves between A) RA vs. SLE, B) RA vs. SSc, and C) SLE vs. SSc—representative examples. The Andrews curves were calculated for various combinations of gene expression values from the whole set of studied genes. Examples show the results of the Andrews curve analysis for the combination of (a) *TLR3*, *TLR7*, *TLR8*, *IL1R1*, *IL1RN*, and *IL18R1*; (b) *TLR3*, *TLR4*, *TLR6*, *TLR10*, *IL1B*, *IL1R1*, and *SIGIRR*; and (c) *TLR4*, *TLR6*, *TLR7*, *TLR8*, *IL1R1*, *IL1RN*, and *IL18*. For those sets of genes, a good separation of diseases was observed as visualized by separation of the curve’s amplitudes and phase shift. Full lines represent the mean values, the dashed lines 95% confidence intervals.

Next, we applied the association rule analysis to identify rules (set of genes including their expression levels) describing a certain disease within the three studied diseases. Based on the results from the Andrews curves, association rule analysis was performed using the whole gene set. For RA, six rules were identified, thus



showing high heterogeneity within this group of patients when compared to SLE and SSc (Figure 17), where for each of them, three rules were identified. In RA, a high level of *TLR3* and *IL1RAP* mRNA was identified in three and two rules, respectively. In SLE, low expression levels of *IL1RN* and *IL18R1* appeared in two rules, and in SSc, a low level of *TLR5* and *IL18R1* mRNA occurred in three and two rules, respectively. The obtained association rules and their support and confidence values deciphered for RA, SLE, and SSc patients are listed in Appendix E, Table 3. The accuracy of classification by using these rules for RA, SLE, and SSc was 83%, 78%, and 77%, respectively. Comparison of rules for each disease revealed that *TLR3*, *TLR5*, *IL18*, *IL18R1*, and *IL1R1* genes occurred in rules for all studied diseases, showing good discriminant power among studied autoimmune diseases.

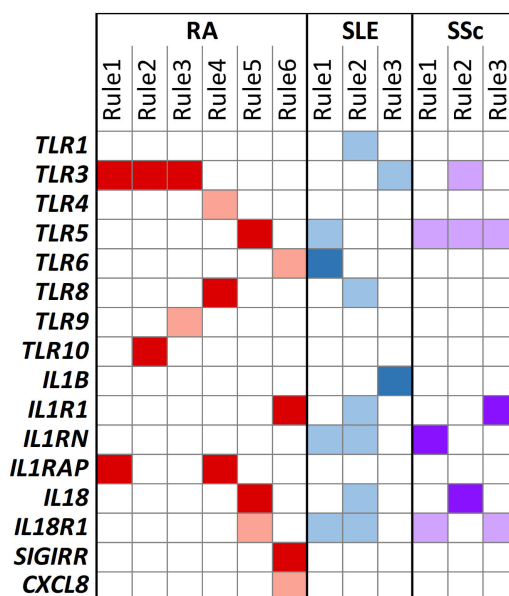


Figure 17: Association rules describing RA, SLE, and SSc.

Association rule analysis revealed a minimum of six rules for RA, three rules for SLE, and three rules for SSc, able to discriminate among all studied diseases with the accuracy above 77%. Columns represent individual rules (combinations of genes and its expression levels characterizing the particular disease). Dark/light color means high/low gene expression levels (cut-off: mean gene expression of the whole data set).

To further explore the heterogeneity in innate signature in RA, we performed multivariate analysis based on utilising patient similarity networks in RA patients only, also including the inactive patients in this analysis. When comparing active

and inactive RA, upregulated *TLR2*, *TLR4*, *TLR6*, and *TLR8* and downregulated *TLR10* expression was associated with the disease activity ( $p < 0.04$ , Appendix C, Figure 2). The multivariate analysis revealed the existence of four patient's subsets (clusters) based on different *TLR8* and *IL1RN* expression profiles, two in active and two in inactive RA (Appendix C, Figure 5). Moreover, neural network analysis identified two main gene sets describing active RA within an activity-related innate signature (*TLR1*, *TLR2*, *TLR3*, *TLR7*, *TLR8*, *CXCL8/IL8*, *IL1RN*, *IL18R1*).

## 5 DISCUSSION

### 5.1 Genetic biomarkers in haemato-oncology diseases

#### 5.1.1 Rare germline *ATM* variants in CLL

Our study revealed the underestimation of rare pathogenic germline *ATM* variants in multiple cancers. We showed an example of CLL that half of the rare *ATM* germline variants, currently classified as VUS, are pathogenic. Moreover, they predispose to cancer and del(11q) acquirement, leading to biallelic *ATM* inactivation (Figure 18). Our data highlight the need for the implementation of rare germline *ATM* variants into diagnostics and clinical decision-making.

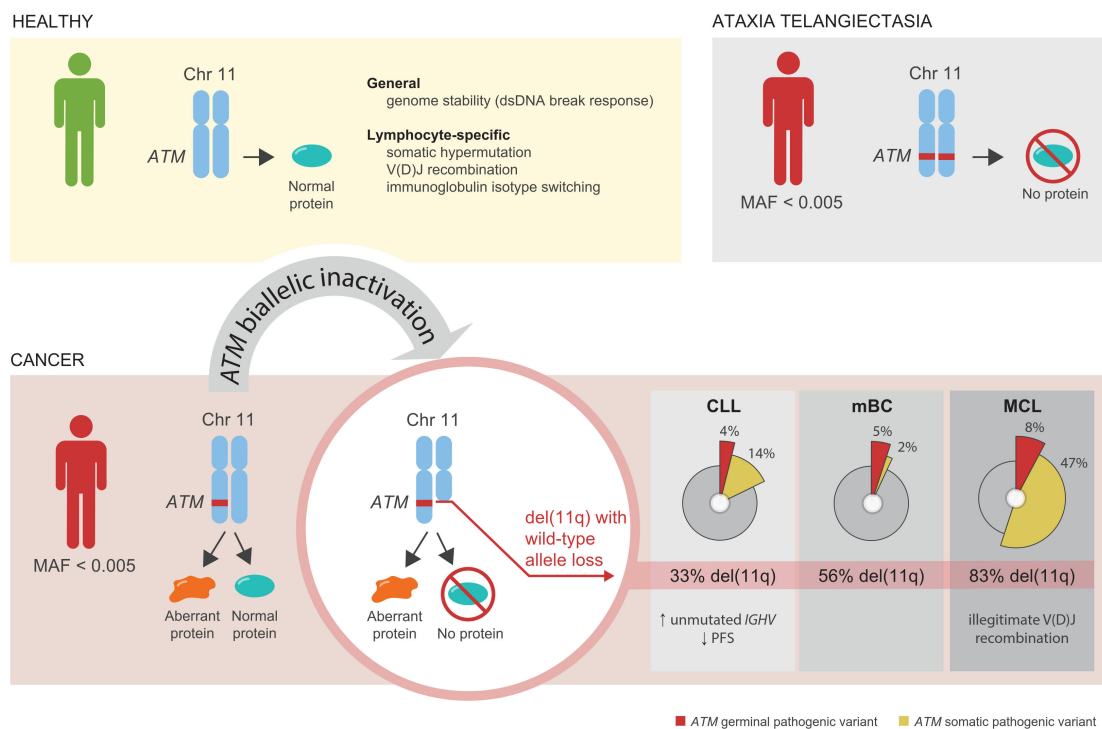


Figure 18: Involvement of *ATM* rare germline pathogenic/predicted pathogenic variants in cancer.

By analysing a panel of genes associated with CLL in our diagnostic NGS testing, we detected a number of variants in the *ATM* gene, classified as VUS by clinical databases in patients with CLL. Interestingly, the occurrence of VUS in *ATM* was significantly higher than in other investigated genes, where VUS were detected very

rarely (<1%). To gain deeper insight into these variants, we analysed matched germline samples and revealed that the majority of VUS are rare germline variants occurring in less than 0.5% in any population, based on databases. These variants occurred in  $\sim 7\%$  of patients. Of them, two-thirds were annotated as pathogenic (1.5%) or VUS with a pathogenic prediction (3.0%) and one-third with a benign prediction (2.4%). Together, germline *ATM* pathogenic/predicted pathogenic variants were detected in 4.5% of CLL patients, all as heterozygous in the germline.

Compared with previous studies on CLL, the inconsistency in the reported prevalence of rare germline variants in the *ATM* gene points to the challenges in the interpretation of these variants. One study reported only pathogenic variants known to cause ataxia telangiectasia occurring in 2.2% of CLL patients [62], while another study included all protein-altering variants with both pathogenic and benign predictions found in 26.3% of CLL patients, but also occurring in 16.6% of healthy controls [51]. To confirm our findings, we reanalysed the CLL public dataset on 445 paired tumour-normal samples investigated previously for recurrent somatic mutations [166]. We detected the same prevalence (3.4%) of rare cases of germline *ATM* pathogenic/predicted pathogenic variants as in our cohort. Our data confirmed that rare germline *ATM* pathogenic/predicted pathogenic variants predispose to CLL (odds ratio 9.1), as others have already shown [51,62]. Moreover, these *ATM* variants were associated with a family history of cancer (88 vs 43%) but not with age at diagnosis.

Generally, VUS are difficult to translate into the clinical meaning. Looking at the *ATM* gene,  $\sim 60\%$  of submitted variants in genetic databases are classified as VUS, of them,  $\sim 80\%$  are missense. A similar proportion of VUS was also found in other cancer-related genes (*TP53*, *BRCA1*, *BRCA2* and *CHEK2*) in databases (total classified variants compiled from UniProt, ClinVar, VarSome and PubMed databases [167]). The high proportion of VUS in *ATM* and other genes may be explained by the rarity of the individual variants, lack of causality with particular cancer types, challenging interpretations and primarily a lack of functional data.

To prove the pathogenicity of VUS-predicted pathogenic variants detected in our CLL cohort, we performed a function study in ex vivo primary cells. Indeed, *ATM* pathogenic and VUS-predicted pathogenic variants decreased the ATM kinase activity slightly ( $\sim 10\text{--}20\%$ , depending on the variant type, however, additional hit by del(11q) led to biallelic inactivation with complete loss of ATM activity in all CLL samples. Our data points to the importance of functionally characterizing the VUS-predicted pathogenic variants, and in diagnostics, to refer to them at least as VUS-predicted pathogenic variants until their function is further clarified. Our data revealed that 33% of CLL patients with rare pathogenic/predicted pathogenic *ATM* variants acquire del(11q), as shown in both investigated CLL cohorts. In these patients, del(11q) is three times more likely to occur than in those with wild type *ATM*. Similarly, 83% of patients with MCL and 56% with metastatic breast cancer with rare germline *ATM* pathogenic/predicted pathogenic variants had concurrent del(11q). Importantly, patients with *ATM* pathogenic/predicted pathogenic allele and del(11q) always lost the wild type allele. From this perspective, we suggest that rare germline *ATM* pathogenic/predicted pathogenic variants behave like somatic *ATM* variants, which are strongly associated with del(11q), thus accelerating the leukemia progression [168].

Our data further highlight the need for the implementation of rare germline *ATM* variants into clinical decision-making, not only in CLL. Nevertheless, germline variants with concurrent del(11q) may look like somatic mutations due to a shift in VAF from the expected 40–60% for a heterozygous allele to 60–100%, or it might be overseen if tumour-normal filtering is used in diagnostics [57]. Moreover, our study demonstrated that rare germline *ATM* pathogenic/predicted pathogenic variants have the same clinical impact as somatic *ATM* and/or *TP53* disruption, resulting in reduced PFS in treated patients, even on novel agents.

Furthermore, all CLL patients from both investigated cohorts carrying truncating or missense variants in the ATM kinase domain developed CLL with unmutated *IGHV* status, a strongly unfavourable factor in CLL [169]. Despite the small number

of patients, the association of these variants with unmutated *IGHV* status is supported by a recent study in mice, where the loss of *ATM* leads to a decreased rate of somatic hypermutation [170]. This study also demonstrated that *ATM* influences germinal center integrity in secondary lymphoid organs, where somatic hypermutation occurs in developing lymphocytes. As unmutated *IGHV* CLL has a more aggressive course associated with shorter survival [169], therefore the observed association of *ATM* pathogenic/predicted variants with unmutated *IGHV* deserves further investigation.

Besides CLL, there is increasing evidence that rare heterozygous germline *ATM* pathogenic variants increase the risk of other cancers [53, 59–61], while rare homozygous germline *ATM* pathogenic variants cause autosomal recessive disorder ataxia telangiectasia. This is consistent with observations in mouse models, where one mutated *ATM* allele had a heightened susceptibility to cancer [171] and caused more genomic instability than the complete loss of the *ATM*, leading to an ataxia-like phenotype [171, 172].

Among all cancers, the highest prevalence of germline *ATM* pathogenic variants was reported in pancreatic cancer (3%), but this study did not include *ATM* variants annotated as VUS [61]. Given the inconsistency of interpretations of rare germline *ATM* variants in CLL, we reanalysed the public dataset of i) metastatic breast cancer and ii) MCL, where somatic *ATM* mutations are most frequent among cancers [173], and germline variants have not been studied yet. In both datasets, we detected rare germline *ATM* pathogenic variants (breast cancer/MCL: 2%/4%) and the same proportion (2%/4%) of germline variants annotated as VUS but were predicted as pathogenic variants. Overall, the germline pathogenic/predicted pathogenic *ATM* variant in CLL, MCL and breast cancer reach such prevalence as *BRCA1* and *BRCA2* germline mutations in breast cancer ( $\sim 4\%$  each gene) [174, 175]. In contrast to the causal association of *BRCA1* and *BRCA2* with breast cancer, *ATM* variants are not associated with a specific type of cancer but occur in multiple malignancies. Our systemic review demonstrates that the rare *ATM* pathogenic/predicted

pathogenic variants are more common in cancer but often missed in diagnostics and highlights the role of the interplay of germline and somatic variation in cancer pathogenesis.

Haematological B-cell malignancies showed the highest prevalence of pathogenic germline *ATM* variants in our study. In this context, it is interesting to note that the initial event in MCL, translocation t(11;14), is a result of illegitimate V(D)J recombination [176] and in CLL, decreased rate of somatic hypermutation determines the CLL type with dismal clinical outcome [169]. As *ATM* has already been demonstrated to be involved in V(D)J recombination and somatic hypermutation processes in lymphocytes [170, 177, 178], one may suggest that pathogenic germline *ATM* variants may contribute to the onset and adverse phenotype of these diseases. How pathogenic germline *ATM* variants interfere with processes ongoing in lymph nodes deserves further investigation.

In conclusion, this study reveals that half of the rare germline *ATM* variants classified as VUS are pathogenic and behave in the same manner as *ATM* somatic mutations, at least in CLL. As shown by our own and public datasets, the prevalence of rare pathogenic/predicted pathogenic *ATM* variants is underestimated across cancers. This study highlights the importance of implementing rare *ATM* germline variants with pathogenic prediction in clinical diagnostics and decision making, not only in CLL.

### 5.1.2 Applicability of the whole bone marrow for analysis of *TP53* mutations in MM

Our study demonstrated that only three-quarters of diagnostic BM samples was possible to enrich due to low recovery and/or sample amount. The enrichment of PCs increased the mutation VAF on average 23%, which was not proportional to the increase of the tumour PCs fraction in the sample achieved by PC enrichment (on average 71%). This phenomenon may be explained by: i) high clonal heterogeneity of malignant myeloma PCs with variable immunophenotypes, and ii) presence of

non-malignant PCs or various B cell precursors in the enriched sample that are also positive for Syndecan-1 (CD138), although this marker might be of a lower surface density [179], which was excluded in our enriched samples (<1% as assessed by immunophenotyping).

In our present cohort, 28% of patients had *TP53* disruption (6% at diagnosis). In newly diagnosed MM, *TP53* mutations are rare and are associated with more aggressive disease and treatment resistance [69, 180]. During the disease course, the *TP53* mutations contribute to the disease progression and the biallelic inactivation of *TP53* has been reported in 21–26% patients at relapse [181]. Regarding other genes, *KRAS* mutations were detected in 19% patients, *NRAS* mutations in 13%, and *BRAF* mutations in 9%. Mutations in *KRAS* and *NRAS* were mutually exclusive; this phenomenon was already reported in myeloma cell lines [182].

The minimum percentage of cancer clonal fraction harbouring a del(17p), as well as *TP53* mutations indicative of poor prognosis in MM, is still under investigation. The European Myeloma Network recommends a 20% positive cut-off level for numerical abnormalities [183], another study suggests a 55% threshold for prognostic evaluation of del(17p) in newly diagnosed MM patients [184]. However, others demonstrated an independent association between subclonal *TP53* deletions and MM outcome [185].

In conclusion, our present analysis in a real-world diagnostic cohort demonstrates the utility of the whole BM for *TP53* mutation analysis by deep targeted NGS in MM when enriched PCs are not available, while obtaining diagnostic information comparable to enriched samples. Furthermore, employing novel, highly sensitive sequencing techniques will help to ensure the required sensitivity for mutation detection from the whole BM [100, 101]. Moreover, our present data highlight the importance of the assessment of *TP53* mutations in patients with MM, as they may occur regardless of del(17p), as well as the need for standardisation of PC enrichment in diagnostics. In the current era of precision medicine, routine screening for *TP53* mutations in MM can enhance patient risk stratification.



### 5.1.3 Standardization of sequencing coverage depth in diagnostic NGS

Although diagnostic NGS has gained prominence in clinical settings for the assessment of somatic mutations in cancer, insufficient standardization of sequencing parameters still limits its implementation in clinical practice [90], mainly for variants present at low allele frequencies [93]. We, therefore, addressed the technical question of correctly determining the sequencing depth in diagnostic NGS in order to obtain confident and reproducible detections of low VAF variants. In particular, we performed theoretical calculations to determine the optimum depth of coverage for the desired probability of detection of variants at low allele frequencies, taking into account the sequencing error rate. Moreover, we confirmed these theoretical calculations by conducting dilution experiments. Based on these observations, we recommend a depth of coverage of 1,650 or higher (together with the respective threshold of at least 30 mutated reads) to call  $\leq 3\%$  variants to achieve a 99.9% probability of variant detection, using the conventional NGS sequencing error only. Variants in the 1–3% VAF range can only be called if the obtained sequence data is of high quality (average Q30 > 90%) and/or when the variants are confirmed by replication or the orthogonal method [90, 100, 186]. We are also providing a simple, user-friendly theoretical calculator (software) to assist laboratories with resolving the correct sequencing depth and the corresponding minimum number of variant reads while taking into account the sequencing error rate. Our simple calculator may help to minimize the false positive and false negative results in diagnostic NGS.

Nevertheless, correct sequencing depth is also influenced by assay-specific factors [90]. Errors can occur at many stages during DNA processing and library preparation. The most common are amplification errors introduced during NGS library preparation [90, 162, 187]. Other common sources of errors have to do with library complexity (the number of independent DNA molecules analyzed), DNA quality, and target region complexity etc. All potential assay-specific errors should be addressed through test design, method validation, and quality control.

Currently, emerging error correction strategies, both computational and experimental, are being developed in order to mitigate the high error rates in diagnostic NGS [100]. So far, among the most promising error correction methods are UMI (unique molecular identifiers), which correct for PCR errors [188], and signal-to-noise correction approaches [100]. These advances attempt to reduce the LOD, thereby increasing sequencing accuracy needed for future opportunities in NGS diagnosis.

There is currently no consensus on the minimum required coverage in a clinical setting using deep targeted resequencing by NGS, and so each laboratory has to set its own parameters in order to meet sufficient quality [90,94]. To date, only a few studies have recommended the minimum coverage criteria for deep targeted NGS in clinical oncology: 500 depth of coverage and a LOD of 5% [91], 300–500 depth of coverage without defying the LOD [92], 250 depth and a LOD of 5% with threshold adjustment to 1,000 depth of coverage is required in cases of heterogeneous variants in low tumour cellularity samples [90], and 100 depth with at least 10 variant reads and a LOD of 10% [99]. According to the binominal data distribution, a coverage depth of 250 should indeed be sufficient to detect 5% VAF with a threshold of variant supporting reads  $\geq 5$ . On the other hand, NGS analysis with a coverage depth of 100 along with a requirement of at least 10 variant supporting reads as recommended by the ERIC consortium [99] would result in a false negative of 45% for samples with a LOD of 10%. The false negative rate is often underestimated in targeted NGS. A recent study investigating inter-laboratory results of somatic variant detection with VAFs between 15 and 50% in 111 laboratories with reported LODs of 5–15% shows that major errors in diagnostic NGS may arise from false negative results, even in samples with high mutation loads [95]. Of three concurrent false positive results, all variants were correctly detected but mischaracterised. Since laboratories have not been asked to report coverage depth for other regions than the identified variants [95], we may only assume that low coverage or high variant calling thresholds contributed to the false negative results. These results further highlight

the need for standardized coverage depth parameters in diagnostic NGS, taking into account sequencing errors as well as assay-specific errors.

In order to evaluate the occurrence of low VAF in real-world settings, we observed 52.6% frequency of *TP53* mutations with VAF at 10% or lower in our diagnostic cohort of CLL patients positive for *TP53* mutations. In line with our observations, a recent study [97] reported the presence of 63 and 84% low burden (Sanger negative) *TP53* mutations in CLL patients at the time of diagnosis and at the time of treatment, respectively, and confirmed the negative impact on the overall survival of *TP53* mutations above 1% VAF at the time of treatment.

In order to improve the standardization in diagnostic NGS, the estimation of correct coverage depth is a recommended starting point when assessing thresholds surrounding a particular NGS assay. Nevertheless, there is still lack of published guidance regarding the minimum technical requirements and its reporting in NGS, particularly important in detection of clonal and subclonal mutations in cancer diagnostics. This is mainly due to the broad range of library preparation approaches, and numerous variables playing a role in each specific NGS assay, that are difficult to standardize, together with inter-laboratory variability. Therefore, the definition of minimum technical requirements and its reporting in NGS is highly desirable. Based on our experience in diagnostic NGS in haemato-oncology, we suggest to report at least following technical parameters: LOD, overall error of NGS assay (or at least sequencing error rate), the amount of DNA input, source, and quality of DNA, minimum coverage depth and the percentage of targeted bases sequenced at this minimum depth, total number of target reads covering variant region and number of reads supporting the variant. Special emphasis should be given to NGS standardization of the formalin-fixed paraffin-embedded (FFPE) samples [189,190].

Taken together, our study highlights the importance of correct sequencing depth and the minimum number of reads required for reliable and reproducible detection of variants with low VAF in diagnostic NGS. The calculation of correct sequencing depth for a given error rate using our user-friendly theoretical calculator (software)

may help to minimize the false positive and false negative results in diagnostic NGS, in situations related to subclonal mutations among others. The rigorous testing and standardized minimum requirements for diagnostic NGS is particularly desirable to ensure correct results in clinical settings.

## 5.2 Serum biomarkers in SLE

Using innovative highly sensitive multiplex PEA analysis on 92 inflammation-related proteins, we identified the serum protein pattern associated with SLE, with many proteins not yet reported in this disease. Moreover, we identified the serum patterns associated with irreversible organ damage and active LN and identified proteins showing utility for the identification of patients at risk of these severe disease manifestations.

This serum protein study in SLE patients revealed the deregulation of 30 proteins in SLE. The majority of the upregulated proteins were known inflammatory mediators: IL6, IL10 [191], IL18 [192], CX3CL1 [193], CCL2 [194], CCL3, CCL7, CCL19 [195], and FGF23 [196] already reported in SLE previously. Interestingly, the most upregulated proteins—sirtuin 2 and caspase 8—were not associated with SLE or even with any autoimmune disease. However, recent reports in animal models and cell lines support their involvement in inflammation and autoimmunity. Regarding sirtuin 2, macrophages expressing this protein produced more inducible nitric oxide synthase/nitric oxide upon lipopolysaccharide (LPS) stimulation than those with depleted sirtuin 2 [197]. This result was also confirmed *in vivo*, where wild type mice responded to LPS by increased nitric oxide levels and a higher amount of M1-macrophages compared to sirtuin 2 knockout mice [197]. Elevated sirtuin 2 also contributed to prolonged hypoinflammation in a septic murine model [198]. Regarding caspase 8, a protein widely recognized for its role in apoptosis, recent reports identify this enzyme as a crucial regulator of inflammation through NF $\kappa$ B activation and cleavage of pro-IL1 $\beta$  and/or pro-IL18, similarly to caspase 1 [199, 200]. These observations lead us to suggest that caspase 8 may also promote autoimmunity by

stimulating IL17 production by T cells, as shown for caspase 1 [201]. Moreover, the therapeutic potential of caspase 8 is supported by the observation of attenuated retinal ischemic damage resulting from the inhibition of caspase 8, resulting in the blockade of IL1 $\beta$  production [202]. However, there is evidence about the pleiotropic effects of sirtuin 2 and caspase 8, and thus future studies on their role in SLE are needed.

Further highly upregulated proteins, IL18 and sulfotransferase 1A1, were already reported in autoimmunity. An elevated IL18 serum level was reported in SLE [203], especially in LN patients [204]. Regarding sulfotransferase 1A1, higher activity was found in autoimmune thyroid disease glands compared to normal thyroids [205], but no information yet exists in SLE. Interestingly, we did not detect any elevation of the serum level of the previously reported SLE-associated factor TWEAK and IFN $\gamma$  [206, 207]. Despite the reported association of the IFN gene expression signature with disease activity in SLE [164, 165], we did not confirm either elevated levels of the IFN-regulated chemokines CCL8, CXCL9, CXCL10 or strong correlation of the IFN protein signature with disease activity at the protein level in the sera of our patients. Our observation is in line with others [164], thus supporting the opinion that cytokine levels in serum are a less sensitive readout for activation of the IFN pathway than the gene expression signature.

Despite tremendous efforts, the greatest challenges still remain in the management of SLE patients with severe organ damage and active LN. Thus, there is a need to identify novel biomarkers that will better facilitate the assessment of organ involvement and disease activity. In our study, SLE patients with organ damage had elevated serum levels of IL8, CCL2, IL6, CCL11, FGF21, MMP10, IL18, CCL3, FGF5, and FGF23 compared to those without organ damage. Of these, enhanced levels of CCL11, MMP10, and CCL2 were informative for the identification of patients with organ damage. Importantly, CCL11, MMP10, and CCL2 also correlated with disease activity. The elevation of the chemokine CCL11 was already associated with damage to various organs, as shown in idiopathic retroperitoneal fibrosis [208]

and liver cirrhosis patients [209]. Moreover, in murine models of lung fibrosis [210], as well as of eosinophilic myocarditis [211], the blockade of the CCL11-CCR3 pathway prevented organ damage. Similarly, MMP10 was linked to renal damage [212] and tissue destruction in arthritis [213]. Elevation of MMP10 was already reported in SLE patients [214] and in a murine LN model with glomerulonephritis [215]. Another protein associated with organ damage, CCL2, was already reported in kidney damage in lupus murine models [216] and in SLE patients with irreversible renal damage [217]. Although IL6 and IL8, cytokines involved in the pathogenesis of SLE, were also enhanced in our patients with organ damage, our analysis did not support their predictive value for this severe phenotype. The usefulness of CCL11, MMP10, and CCL2 as biomarkers or possible treatment targets needs to be elucidated in future studies.

Lupus nephritis is considered another challenging SLE phenotype from the point of view of its prediction and preemptive diagnostics. Renal biopsy is still the gold standard to assess the renal involvement of SLE and its severity and pathological category [115]. The search for non-invasive biomarkers in serum and urine reflecting the renal disease activity is therefore a major focus of interest. Our serum protein analysis in LN patients with active renal disease revealed upregulated levels of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF compared to those without LN. All these markers showed excellent discrimination for active LN, significantly better than the classical markers as shown by us and others [116, 117]. Moreover, we observed good discrimination between active and inactive renal disease in LN patient subgroup for all markers, except for GDNF. Apart from caspase 8 and sIL15RA, emerging evidence of the active involvement of these proteins in LN already exists. Regarding CSF1, elevated serum levels in patients with SLE were shown to reflect kidney histopathology and to predict renal disease activity [218]. Moreover, CSF1 deficiency protected against LN in murine models [219]. Enhanced protein and gene expression of IL15RA was detected in leucocytes from SLE patients [220, 221], probably as a results of hydroxymethyla-

tion in promoter region of this gene in SLE [221]. There is also evidence about the crucial role of the CD40-CD40L system in the development, progression and outcome of SLE [222]. Enhanced CD40L protein level was detected in sera from SLE patients [222] as well as class III and IV LN and other inflammatory renal diseases [223]. Moreover, CD40 gene silencing reduced the progression of experimental LN [224]. Regarding sCX3CL1, elevated expression was reported in proliferative LN [225] and the administration of a CX3CL1 antagonist to mice delayed the initiation and ameliorated the progression of LN [226]. Also enhanced expression of IL18R1 has already been reported in SLE patients [227] as well as in peripheral plasmacytoid dendritic cells in active LN patients [228]. Similarly, increased levels of NGF, a complex of 3 subunits—*a*NGF, *b*NGF, and *g*NGF, has been reported in the sera of SLE patients [229] and various renal disorders [230]. Regarding GDNF, a high expression of this protein was detected in renal biopsies from patients with proteinuric nephropathy [231] and increased plasma levels of GDNF were reported in patients with chronic renal diseases [232]. This mesangial autocrine growth factor was shown to play a pivotal role in mesangial cell proliferation, which is essential for the progression of various glomerular diseases [233]. Our study did not confirm IL18 as a useful biomarker to assess the activity of renal disease, as reported by others [203]. On the other hand, our results nominated spectrum of novel biomarkers of renal involvement for further confirmation studies.

Although relatively high sensitivity and specificity was obtained for each individual marker in our LN and organ damage subgroups, we believe that using rather a panel of multiple biomarkers and/or combination with other clinical and laboratory parameters would be an appropriate approach in the identification of patients with these severe manifestations.

This exploratory study revealed many novel proteins associated with SLE for future immunopathogenesis studies, as well as nominating candidate biomarkers for irreversible organ damage and active lupus nephritis. Future studies on larger cohorts with well-defined phenotypes as well as the longitudinal follow-up during

disease development are needed to prove the suitability of these proteins or their combinations as biomarkers for organ damage and lupus nephritis, with special emphasis on disease activity.

### 5.3 Gene expression biomarkers in autoimmune diseases

This study focused on the innate immunity gene signature among major autoimmune diseases: RA, SLE, and SSc, showing heterogeneity in the innate signature among and within these diseases. This first cross-disease study showed the highest diversity and abundance in the innate signature in RA when compared to SLE and SSc.

Innate immunity plays a key role in the pathogenesis of autoimmune rheumatic diseases as evidenced from numerous studies on individual members of innate immunity pathways [136, 137]. However, little is known about the similarities and differences in the innate signature at the molecular level between and within these diseases. Therefore, we investigated the differential expression of key innate genes in RA, SLE, and SSc. Importantly, our study was restricted only to the cases with active disease in order to exclude heterogeneity due to the activity and inactivity of the diseases. To obtain a more complex picture, the multivariate analysis was applied to assess the complexity of the differential innate signature having an advantage over classical statistical approaches due to taking into account the intrinsic characteristics of gene expression data and assessing the relationships between studied molecules.

Firstly, we applied Andrews curve analysis for assessment of differences and similarities in the gene innate signature between studied diseases, an approach particularly useful for visualization of the structure in multidimensional data [158]. When using combination of genes reaching statistical significance as well as using the whole gene set, we confirmed the diversity among innate profiles in RA, SLE, and SSc by Andrews curve analysis. Upregulated expression of *TLR3*, *TLR5*, and *SIGIRR* was characteristic for RA when compared to both SLE and SSc. An intracellular receptor TLR3 recognizing dsRNA has been shown to be involved in the RA pathogenesis:



necrotic synovial fluid cells release RNA that can activate TLR3 in RA synovial fibroblasts [234]. TLR5, a surface receptor highly upregulated in our RA patients, recognizes bacterial flagellin. However, their endogenous ligand(s) in synovial fluid able to activate TLR5 in RA is(are) still unknown [235,236]. In line with our results, increased TLR5 in peripheral blood myeloid cells correlated with RA disease activity and TNF- $\alpha$  levels [237]. There is also evidence that anti-TNF- $\alpha$  therapy markedly suppress TLR5 expression in RA monocytes [238], a trend which was also observed in our study. Also, the next highly upregulated SIGIRR (IL1R8/TIR8), an orphan receptor required for the anti-inflammatory effects of IL37, has been reported in RA synovial tissue previously [239].

Also, other genes such as *TLR2*, *IL1RAP*, and *IL18R1* from the differential innate signature associated with RA revealed by our analysis were reported in autoimmune conditions previously. In line with our results, abundant TLR2 on monocyte subsets in active RA produced a spectrum of proinflammatory cytokines after stimulation [240]. TLR2 recognizes a wide range of conserved microbial products, probably due to its cooperation with TLR1 or TLR6, as well as its hypothetic ligand HMGB1 released from dying and activated cells [241]. Regarding *IL1RAP* and *IL18R1*, their upregulated expression in RA was reported recently [145] and their downregulation in SSc we report here for the first time. Finally, SSc was characterized by an increase in *IL1R1* in comparison to SLE. The first evidence about critical involvement of IL1R1, an essential mediator for proinflammatory IL1 signaling [242], in fibrotic processes has been already reported in a murine lung injury model [243]. Importantly, data from our cross-disease analysis are in line with previous studies on individual innate members and basic statistical analysis and further highlight the activation of innate immunity in RA when compared to SLE and SSc. The infectious agents and endogenous ligands activating innate receptors leading to a self-sustaining inflammatory loop responsible for chronic and destructive progression in RA need to be further elucidated.

Next, we investigated the differential innate signature among and within the

studied diseases by association rule analysis, a method commonly used to uncover the most frequently purchased combinations of items in a market basket analysis. It has been shown that this analysis is highly convenient for gene expression datasets and gives additional information due to preservation of the causality between the gene expression level and phenotype [244]. For RA, six rules were identified, thus showing high heterogeneity within this group of patients when compared to SLE and SSc, where three rules were identified for each of them. In RA, the association rules most frequently included high expression of *TLR3* and/or *IL1RAP/IL1R3*, thus again highlighting activation of the innate system in active RA. In SLE, a low expression of *IL1RN* and *IL18R1* and in SSc, a low level of *TLR5* and *IL18R1* occurred often in the rules. Applying association rules (combinations of genes describing a particular disease), excellent confidence and accuracy above 77% was achieved for all investigated diseases.

Interestingly, about half of the patients in each disease were characterized by multiple rules, while others were typical by only one gene expression pattern rule. The existence of several innate profile subgroups within RA patients lets us suggest that the heterogeneity in the innate pattern in RA may contribute to various clinical disease manifestations [135, 145], thus deserving future investigation. We also hypothesize that observed heterogeneity in the innate signature may contribute to the heterogeneity in the IFN signature recently reported in RA [135]. Our data further highlighted the application of advanced multivariate data analysis especially for diseases such as SLE, where many clinical phenotypes exist. This may be reflected in the high variability in the expression pattern which might be underestimated by univariate statistics, especially in the case of low abundant genes. Finally, our data points out the involvement of various key innate molecules as well as the different interplay between individual innate receptors in the studied diseases.

To gain a more complete picture of the innate signature in autoimmune diseases, we report also the differential profile of the innate signature in studied diseases compared to healthy controls. This comparison revealed the upregulation of four mem-

bers of TLR (*TLR2*, *TLR3*, *TLR5*, and *TLR8*) and six members of the *IL1/IL1R* family (*IL1B*, *IL1RN*, *IL1RAP*, *IL18R1*, *IL18*, and *SIGIRR*) in RA when compared to healthy controls. In line with our results, deregulation of these genes or their protein products was already registered in RA [145, 237, 239, 245–250]. In SLE, this study showed for the first time downregulation of *TLR10*, a broad negative regulator of TLR signaling [251, 252]. The first evidence about the possible involvement of TLR10 in autoimmunity has been already observed: downregulated *TLR10* expression was reported in PBMC of patients with microscopic polyangiitis [253]. In contrast to the murine models of SLE [254], we did not observe increased *TLR7* and *TLR9* expression in our SLE patients. In SSc, our study revealed upregulation of *IL1RN*, *IL18*, and *CXCL8* and downregulation of *IL1RAP* and *IL18R1*. In line with our results, upregulated *IL1RN* mRNA [255], increased *IL18* expression in skin biopsies [256], and elevated serum IL8 in patients with scleroderma [257] were reported. Here, we report for the first time downregulation of *IL1RAP* and *IL18R1* in SSc. IL1RAP (IL1R3) is a coreceptor of IL1R1 and is indispensable for the transmission of IL1 signaling [242]. Regarding *IL18R1* gene, it encodes the  $\alpha$  subunit of the IL18 receptor responsible for IL18 binding. The activated receptor then initiates the same signaling pathway as IL1 to activate NF- $\kappa$ B [258]. How these proteins contribute to the SSc pathogenesis deserves future investigations.

We further investigated the heterogeneity of the innate signature in RA patients with a particular focus on active and inactive RA patients. When active and inactive RA were compared, the upregulation of *TLR2*, *TLR4*, *TLR6*, and *TLR8* and downregulation of *TLR10* were revealed in active RA, nominating them as predictive biomarkers of disease activity. Regarding *TLR10*, an association of the I473T allelic variant (rs11466657) with disease severity and a low response to infliximab has been reported [259]. Functional studies have shown that the *TLR10* I473T variant lacks inhibitory activity on the NF- $\kappa$ B inflammatory pathway in comparison to the wild type allele [259]. Similarly, downregulation of *TLR10* was also observed in our patients with active disease. These observations further nominate TLR10 as a

candidate target molecule able to attenuate the inflammation in active RA.

## 6 CONCLUSION

This thesis summarizes the results of research on new biomarkers and refinement of the analysis of existing biomarkers using novel molecular techniques in haemato-oncological and autoimmune diseases, which were published in nine original articles.

In particular, our studies revealed nomination of i) pathogenic germline variants in *ATM* gene as biomarkers of cancer susceptibility and predictive biomarkers of adverse outcome in chronic lymphocytic leukemia, same as somatic *ATM* mutations, ii) somatic mutations in *TP53* gene from whole bone marrow as predictive biomarker in multiple myeloma and MCL, iii) refinement of next-generation analysis of somatic predictive and prognostic biomarkers in haemato-oncology by standardisation of sequencing coverage, iv) candidate serum biomarkers for irreversible organ damage and active lupus nephritis in systemic lupus erythematosus, v) innate immune gene expression biomarkers in major autoimmune diseases, vi) chromosome 1 abnormalities as candidate biomarker of extramedullary disease in multiple myeloma.

In conclusion, our studies have uncovered many candidate biomarkers for the diseases studied, demonstrating their potential to tailor medical decisions and treatments as required by the era of precision medicine.

## 7 Abbreviations

**AUC** area under the curve

**BCR** B–cell receptor

**BM** bone marrow

**CLL** chronic lymphocytic leukemia

**CNA** copy number aberration

**DAS28** disease activity score in 28 joints

**DLBCL** diffuse large B–cell lymphoma

**dsDNA** double–stranded deoxyribonucleic acid

**ELISA** enzyme–linked immuno sorbent assay

**EMM** extramedullary multiple myeloma

**ERIC** European research initiative on chronic lymphocytic leukemia

**EUSTAR** European scleroderma trials and research group

**FFPE** formalin–fixed paraffin–embedded

**FISH** fluorescence in situ hybridisation

**HL** Hodgkin lymphoma

**iPCR** immuno polymerase chain reaction

**LN** lupus nephritis

**LOD** limit of detection

**LPS** lipopolysaccharide

**MAF** minor allele frequency

**MCL** mantle cell lymphoma

**MM** multiple myeloma

**mRNA** messenger ribonucleic acid

**NGS** next–generation sequencing

**OS** overall survival

**PBMC** peripheral blood mononuclear cells

**PBS** phosphate–buffered saline

**PC** plasma cell

<b>PCR</b>	polymerase chain reaction
<b>PEA</b>	proximity extension immunoassay
<b>PET/CT</b>	positron emission tomography/computed tomography
<b>PFS</b>	progression-free survival
<b>RA</b>	rheumatoid arthritis
<b>ROC</b>	receiver operating characteristic
<b>RT</b>	Richter transformation
<b>RT-qPCR</b>	quantitative reverse transcription polymerase chain reaction
<b>SDI</b>	systemic lupus international collaborating clinics/American college of rheumatology damage index
<b>SLE</b>	systemic lupus erythematosus
<b>SLEDAI</b>	SLE disease activity index
<b>SSc</b>	systemic sclerosis
<b>SV</b>	structural variation
<b>TLR</b>	Toll-like receptor
<b>UMI</b>	unique molecular identifiers
<b>UTR</b>	untranslated region
<b>VAF</b>	variant allele frequency
<b>VUS</b>	variant of uncertain significance
<b>WES</b>	whole exome sequencing
<b>WGS</b>	whole genome sequencing

## 8 References

- [1] The Precision Medicine Initiative [(accessed on 8 October 2021)]; Available online: <https://www.nih.gov/sites/default/files/research-training/initiatives/pmi/pmi-infographic.pdf/>
- [2] Abrahams E. Right drug-right patient-right time: personalized medicine coalition. *Clin Transl Sci.* 2008;1(1):11-12.
- [3] Ghasemi M et al. Precision medicine and molecular imaging: new targeted approaches toward cancer therapeutic and diagnosis. *Am J Nucl Med Mol Imaging.* 2016;6(6):310-327.
- [4] Pritzker K. Biomarker imprecision in precision medicine. *Expert Rev Mol Diagn.* 2018;18(8):685-687.
- [5] Dalton WS et al. Cancer biomarkers—an invitation to the table. *Science.* 2006;312(5777):1165-1168.
- [6] Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001;69(3):89-95.
- [7] Puntmann VO. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. *Postgrad Med J.* 2009;85(1008):538-545.
- [8] Nalejska E et al. Prognostic and predictive biomarkers: tools in personalized oncology. *Mol Diagn Ther.* 2014;18(3):273-284.
- [9] Fey MF et al. Molecular diagnosis of haematological neoplasms. *Blood Rev.* 1988;2(2):78-87.
- [10] Sokolenko AP et al. Molecular Diagnostics in Clinical Oncology. *Front Mol Biosci.* 2018;5:76.
- [11] Ciardiello F et al. Delivering precision medicine in oncology today and in future—the promise and challenges of personalised cancer medicine: a position paper by the European Society for Medical Oncology (ESMO). *Ann Oncol.* 2014;25(9):1673-1678.
- [12] Braggio E et al. Lessons from next-generation sequencing analysis in hematological malignancies. *Blood Cancer J.* 2013;3(7):e127.
- [13] Kohlmann A et al. Next-generation sequencing - feasibility and practicality in haematology. *Br J Haematol.* 2013;160(6):736-753.
- [14] Black JS et al. The impact of next generation sequencing technologies on haematological research – A review. *Pathogenesis.* 2015;2(1-2):9-16.
- [15] Bench AJ et al. Molecular genetic analysis of haematological malignancies: I. Acute leukaemias and myeloproliferative disorders. *Clin Lab Haematol.* 2005;27(3):148-171.



- 
- [16] Eichhorst B et al. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2021;32(1):23-33.
- [17] Döhner H et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med.* 2000;343(26):1910-1916.
- [18] Zenz T et al. TP53 mutation and survival in chronic lymphocytic leukemia. *J Clin Oncol.* 2010;28(29):4473-4479.
- [19] Landau DA et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell.* 2013;152(4):714-726.
- [20] Wang L et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med.* 2011;365(26):2497-2506.
- [21] Bretones G et al. Altered patterns of global protein synthesis and translational fidelity in RPS15-mutated chronic lymphocytic leukemia. *Blood.* 2018;132(22):2375-2388.
- [22] Villamor N et al. NOTCH1 mutations identify a genetic subgroup of chronic lymphocytic leukemia patients with high risk of transformation and poor outcome. *Leukemia.* 2013;27(5):1100-1106.
- [23] Stilgenbauer S et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood.* 2014;123(21):3247-3254.
- [24] Herling CD et al. Complex karyotypes and KRAS and POT1 mutations impact outcome in CLL after chlorambucil-based chemotherapy or chemoimmunotherapy. *Blood.* 2016;128(3):395-404.
- [25] Baliakas P et al. Cytogenetic complexity in chronic lymphocytic leukemia: definitions, associations, and clinical impact. *Blood.* 2019;133(11):1205-1216.
- [26] Rose-Zerilli MJ et al. ATM mutation rather than BIRC3 deletion and/or mutation predicts reduced survival in 11q-deleted chronic lymphocytic leukemia: data from the UK LRF CLL4 trial. *Haematologica.* 2014;99(4):736-742.
- [27] Petrackova A et al. Revisiting Richter transformation in the era of novel CLL agents. *Blood Rev.* 2021;49:100824.
- [28] Sedlarikova L et al. *Front Oncol.* 2020;10:894.
- [29] Byrd JC et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med.* 2013;369(1):32-42.
- [30] Barr PM et al. Sustained efficacy and detailed clinical follow-up of first-line ibrutinib treatment in older patients with chronic lymphocytic leukemia: extended phase 3 results from RESONATE-2. *Haematologica.* 2018;103(9):1502-1510.
- [31] Burger JA et al. Ibrutinib as Initial Therapy for Patients with Chronic Lymphocytic Leukemia. *N Engl J Med.* 2015;373(25):2425-2437.

- 
- [32] Byrd JC et al. Long-term follow-up of the RESONATE phase 3 trial of ibrutinib vs ofatumumab. *Blood*. 2019;133(19):2031-2042.
- [33] Kaur V et al. Ibrutinib in CLL: a focus on adverse events, resistance, and novel approaches beyond ibrutinib. *Ann Hematol*. 2017;96(7):1175-1184.
- [34] Puła B et al. Overcoming Ibrutinib Resistance in Chronic Lymphocytic Leukemia. *Cancers (Basel)*. 2019;11(12):1834.
- [35] Woyach JA et al. Targeted therapies in CLL: mechanisms of resistance and strategies for management. *Blood*. 2015;126(4):471-477.
- [36] Woyach JA et al. Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. *N Engl J Med*. 2014;370(24):2286-2294.
- [37] Ahn IE et al. Clonal evolution leading to ibrutinib resistance in chronic lymphocytic leukemia. *Blood*. 2017;129(11):1469-1479.
- [38] Woyach JA et al. BTKC481S-Mediated Resistance to Ibrutinib in Chronic Lymphocytic Leukemia. *J Clin Oncol*. 2017;35(13):1437-1443.
- [39] Kadri S et al. Clonal evolution underlying leukemia progression and Richter transformation in patients with ibrutinib-relapsed CLL. *Blood Adv*. 2017;1(12):715-727.
- [40] Cheng S et al. Functional characterization of BTK(C481S) mutation that confers ibrutinib resistance: exploration of alternative kinase inhibitors. *Leukemia*. 2015;29(4):895-900.
- [41] Hamasy A et al. Substitution scanning identifies a novel, catalytically active ibrutinib-resistant BTK cysteine 481 to threonine (C481T) variant. *Leukemia*. 2017;31(1):177-185.
- [42] Liu TM et al. Hypermorphic mutation of phospholipase C,  $\gamma 2$  acquired in ibrutinib-resistant CLL confers BTK independency upon B-cell receptor activation. *Blood*. 2015;126(1):61-68.
- [43] Burger JA et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun*. 2016;7:11589.
- [44] Maddocks KJ et al. Etiology of Ibrutinib Therapy Discontinuation and Outcomes in Patients With Chronic Lymphocytic Leukemia. *JAMA Oncol*. 2015;1(1):80-87.
- [45] Maffei R et al. Targeting neoplastic B cells and harnessing microenvironment: the "double face" of ibrutinib and idelalisib. *J Hematol Oncol*. 2015;8:60.
- [46] Stilgenbauer S et al. Venetoclax in relapsed or refractory chronic lymphocytic leukaemia with 17p deletion: a multicentre, open-label, phase 2 study. *Lancet Oncol*. 2016;17(6):768-778.

- 
- [47] Jones JA et al. Venetoclax for chronic lymphocytic leukaemia progressing after ibrutinib: an interim analysis of a multicentre, open-label, phase 2 trial. *Lancet Oncol.* 2018;19(1):65-75.
- [48] Roberts AW et al. Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. *N Engl J Med.* 2016;374(4):311-322.
- [49] Tausch E et al. Venetoclax resistance and acquired BCL2 mutations in chronic lymphocytic leukemia. *Haematologica.* 2019;104(9):e434-e437.
- [50] Blombery P et al. Acquisition of the Recurrent Gly101Val Mutation in BCL2 Confers Resistance to Venetoclax in Patients with Progressive Chronic Lymphocytic Leukemia. *Cancer Discov.* 2019;9(3):342-353.
- [51] Tiao G et al. Rare germline variants in ATM are associated with chronic lymphocytic leukemia. *Leukemia.* 2017;31(10):2244-2247.
- [52] Ji X et al. Protein-altering germline mutations implicate novel genes related to lung cancer development. *Nat Commun.* 2020;11:2220.
- [53] Esai Selvan et al. Inherited Rare, Deleterious Variants in ATM Increase Lung Adenocarcinoma Risk. *J Thorac Oncol.* 2020;15(12):1871-1879.
- [54] Pomerantz MM et al. The association between germline BRCA2 variants and sensitivity to platinum-based chemotherapy among men with metastatic prostate cancer. *Cancer.* 2017;123(18):3532-3539.
- [55] Singh M et al. Emerging role of NUDT15 polymorphisms in 6-mercaptopurine metabolism and dose related toxicity in acute lymphoblastic leukaemia. *Leuk Res.* 2017;62:17-22.
- [56] Crona DJ et al. Genetic variants of VEGFA and FLT4 are determinants of survival in renal cell carcinoma patients treated with sorafenib. *Cancer Res.* 2019;79(1):231-241.
- [57] Koboldt DC. Best practices for variant calling in clinical sequencing. *Genome Med.* 2020;12(1):91.
- [58] Laake K et al. Characterization of ATM mutations in 41 Nordic families with ataxia telangiectasia. *Hum Mutat.* 2000;16(3):232-246.
- [59] Paulo P et al. Targeted next generation sequencing identifies functionally deleterious germline mutations in novel genes in early-onset/familial prostate cancer. *PLoS Genet.* 2018;14(4):e1007355.
- [60] Zhao ZL et al. ATM rs189037 (G<A) polymorphism increased the risk of cancer: an updated meta-analysis. *BMC Med Genet.* 2019;20(1):28.
- [61] Hall MJ et al. Germline Pathogenic Variants in the Ataxia Telangiectasia Mutated (ATM) Gene are Associated with High and Moderate Risks for Multiple Cancers. *Cancer Prev Res (Phila).* 2021;14(4):433-440.

- 
- [62] Skowronska A et al. ATM germline heterozygosity does not play a role in chronic lymphocytic leukemia initiation but influences rapid disease progression through loss of the remaining ATM allele. *Haematologica*. 2012;97(1):142-146.
- [63] Zingone A et al. Pathogenesis of monoclonal gammopathy of undetermined significance and progression to multiple myeloma. *Semin Hematol*. 2011;48(1):4-12.
- [64] Palumbo A et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. *J Clin Oncol*. 2015;33(26):2863-2869.
- [65] Avet-Loiseau H et al. Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). *J Clin Oncol*. 2010;28(30):4630-4634.
- [66] Bolli N et al. Analysis of the genomic landscape of multiple myeloma highlights novel prognostic markers and disease subgroups. *Leukemia*. 2018;32(12):2604-2616.
- [67] Morgan GJ et al. The genetic architecture of multiple myeloma. *Nat Rev Cancer*. 2012;12(5):335-348.
- [68] Walker BA et al. Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. *Leukemia*. 2014;28(2):384-390.
- [69] Manier S et al. Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol*. 2017;14(2):100-113.
- [70] Kumar SK et al. Multiple Myeloma, Version 3.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2017;15(2):230-269.
- [71] Mansilla C et al. Combined Selection System to Lower the Cutoff for Plasma Cell Enrichment Applied to iFISH Analysis in Multiple Myeloma. *Transl Oncol*. 2018;11(3):647-652.
- [72] Romano A et al. Minimal Residual Disease Assessment Within the Bone Marrow of Multiple Myeloma: A Review of Caveats, Clinical Significance and Future Perspectives. *Front Oncol*. 2019;9:699.
- [73] Petrackova A et al. Diagnostic deep-targeted next-generation sequencing assessment of TP53 gene mutations in multiple myeloma from the whole bone marrow. *Br J Haematol*. 2020;189(4):e122-e125.
- [74] Klener P. Advances in Molecular Biology and Targeted Therapy of Mantle Cell Lymphoma. *Int J Mol Sci*. 2019;20(18):4417.
- [75] Malarikova D et al. Concurrent TP53 and CDKN2A Gene Aberrations in Newly Diagnosed Mantle Cell Lymphoma Correlate with Chemoresistance and Call for Innovative Upfront Therapy. *Cancers (Basel)*. 2020;12(8):2120.

- 
- [76] Obr A et al. A high TP53 mutation burden is a strong predictor of primary refractory mantle cell lymphoma. *Br J Haematol.* 2020;191(5):e103-e106.
- [77] Sanger F et al. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol.* 1975;94(3):441-448.
- [78] Sanger F et al. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74(12):5463-5467.
- [79] Maxam AM et al. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 1980;65(1):499-560.
- [80] Slatko BE et al. Overview of Next-Generation Sequencing Technologies. *Curr Protoc Mol Biol.* 2018;122(1):e59.
- [81] Hood L et al. The Human Genome Project: big science transforms biology and medicine. *Genome Med.* 2013;5(9):79.
- [82] van Dijk EL et al. The Third Revolution in Sequencing Technology. *Trends Genet.* 2018;34(9):666-681.
- [83] Schloss JA. How to get genomes at one ten-thousandth the cost. *Nat Biotechnol.* 2008;26(10):1113-1115.
- [84] Liu L et al. Comparison of next-generation sequencing systems. *J Biomed Biotechnol.* 2012;2012:251364.
- [85] Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet.* 2008;24(3):133-141.
- [86] Kumar KR et al. Next-Generation Sequencing and Emerging Technologies. *Semin Thromb Hemost.* 2019;45(7):661-673.
- [87] Bamshad MJ et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet.* 2011;12(11):745-755.
- [88] Lupski JR et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med.* 2010;362(13):1181-1191.
- [89] Nigro V et al. Next-generation sequencing approaches for the diagnosis of skeletal muscle disorders. *Curr Opin Neurol.* 2016;29(5):621-627.
- [90] Jennings LJ et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn.* 2017;19(3):341-365.
- [91] D'Haene N et al. Clinical Validation of Targeted Next Generation Sequencing for Colon and Lung Cancers. *PLoS One.* 2015;10(9):e0138245.
- [92] Deans ZC et al. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN Path ASBL. *Virchows Arch.* 2017;470(1):5-20.

- 
- [93] Ivanov M et al. Towards standardization of next-generation sequencing of FFPE samples for clinical oncology: intrinsic obstacles and possible solutions. *J Transl Med.* 2017;15(1):22.
- [94] Bacher U et al. Challenges in the introduction of next-generation sequencing (NGS) for diagnostics of myeloid malignancies into clinical routine use. *Blood Cancer J.* 2018;8(11):113.
- [95] Merker JD et al. Proficiency Testing of Standardized Samples Shows Very High Interlaboratory Agreement for Clinical Next-Generation Sequencing-Based Oncology Assays. *Arch Pathol Lab Med.* 2019;143(4):463-471.
- [96] Rossi D et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood.* 2014;123(14):2139-2147.
- [97] Brieghel C et al. Deep targeted sequencing of TP53 in chronic lymphocytic leukemia: clinical impact at diagnosis and at time of treatment. *Haematologica.* 2019;104(4):789-796.
- [98] Campo E et al. TP53 aberrations in chronic lymphocytic leukemia: an overview of the clinical implications of improved diagnostics. *Haematologica.* 2018;103(12):1956-1968.
- [99] Malcikova J et al. ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia-update on methodological approaches and results interpretation. *Leukemia.* 2018;32(5):1070-1080.
- [100] Salk JJ et al. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet.* 2018;19(5):269-285.
- [101] Petrackova A et al. Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. *Front Oncol.* 2019;9:851.
- [102] Chan EKF et al. Optical mapping reveals a higher level of genomic architecture of chained fusions in cancer. *Genome Res.* 2018;28(5):726-738.
- [103] Neveling K et al. Next generation cytogenetics: comprehensive assessment of 48 leukemia genomes by genome imaging. *bioRxiv.* 2020.02.06.935742.
- [104] Kriegova E et al. Whole-genome optical mapping of bone-marrow myeloma cells reveals association of extramedullary multiple myeloma with chromosome 1 abnormalities. *Sci Rep.* 2021;11(1):14671.
- [105] Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med.* 2002;347(12):911-920.
- [106] Chatenoud L. Precision medicine for autoimmune disease. *Nat Biotechnol.* 2016;34:930-932.
- [107] Ceccarelli F et al. Genetic Factors of Autoimmune Diseases 2017. *J Immunol Res.* 2017;2017:2789242.

- [108] Gladman DD et al. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index for Systemic Lupus Erythematosus International Comparison. *J Rheumatol.* 2000;27(2):373-376.
- [109] Nossent J et al. Current causes of death in systemic lupus erythematosus in Europe, 2000–2004: relation to disease activity and damage accrual. *Lupus.* 2007;16(5):309-317.
- [110] Chambers SA et al. Damage and mortality in a group of British patients with systemic lupus erythematosus followed up for over 10 years. *Rheumatology (Oxford).* 2009;48(6):673-675.
- [111] Mok CC et al. Damage accrual in southern Chinese patients with systemic lupus erythematosus. *J Rheumatol.* 2003;30(7):1513-1519.
- [112] Danila MI et al. Renal damage is the most important predictor of mortality within the damage index: data from LUMINA LXIV, a multiethnic US cohort. *Rheumatology (Oxford).* 2009;48(5):542-545.
- [113] Doria A et al. Optimizing outcome in SLE: treating-to-target and definition of treatment goals. *Autoimmun Rev.* 2014;13(7):770-777.
- [114] Cameron JS. Lupus nephritis. *J Am Soc Nephrol.* 1999;10(2):413-424.
- [115] Misra R et al. Biomarkers in lupus nephritis. *Int J Rheum Dis.* 2015;18(2):219-232.
- [116] Balow JE. Clinical presentation and monitoring of lupus nephritis. *Lupus.* 2005;14(1):25-30.
- [117] Enghard P et al. Immunology and the diagnosis of lupus nephritis. *Lupus.* 2009;18(4):287-290.
- [118] Susianti H et al. Analysis of urinary TGF- $\beta$ 1, MCP-1, NGAL, and IL-17 as biomarkers for lupus nephritis. *Pathophysiology.* 2015;22(1):65-71.
- [119] Kulkarni O et al. Chemokines in lupus nephritis. *Front Biosci.* 2008;13:3312-3320.
- [120] López P et al. A pathogenic IFN $\alpha$ , BLyS and IL-17 axis in Systemic Lupus Erythematosus patients. *Sci Rep.* 2016;6:20651.
- [121] Schwartz N et al. Urinary TWEAK as a biomarker of lupus nephritis: a multicenter cohort study. *Arthritis Res Ther.* 2009;11(5):R143.
- [122] Ding H et al. Insulin-like growth factor binding protein-2 as a novel biomarker for disease activity and renal pathology changes in lupus nephritis. *Clin Exp Immunol.* 2016;184(1):11-18.
- [123] Abdallah E et al. Diagnostic performance of urinary osteoprotegerin as a novel biomarker for early detection of lupus nephritis activity. *Life Sci J.* 2015;12:75-81.

- 
- [124] Li Y et al. Biomarker profiling for lupus nephritis. *Genomics Proteomics Bioinformatics*. 2013;11(3):158-165.
- [125] Benjachat T et al. Biomarkers for Refractory Lupus Nephritis: A Microarray Study of Kidney Tissue. *Int J Mol Sci*. 2015;16(6):14276-14290.
- [126] Petrackova A et al. Serum protein pattern associated with organ damage and lupus nephritis in systemic lupus erythematosus revealed by PEA immunoassay. *Clin Proteomics*. 2017;14:32.
- [127] Assarsson E et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One*. 2014;9(4):e95192.
- [128] Sano T et al. Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science*. 1992;258(5079):120-122.
- [129] Greenwood C et al. Proximity assays for sensitive quantification of proteins. *Biomol Detect Quantif*. 2015;4:10-16.
- [130] Niemeyer CM et al. Fluorometric polymerase chain reaction (PCR) enzyme-linked immunosorbent assay for quantification of immuno-PCR products in microplates. *Anal Biochem*. 1997;246(1):140-145.
- [131] Solier C et al. Antibody-based proteomics and biomarker research - current status and limitations. *Proteomics*. 2014;14(6):774-783.
- [132] Cooles FAH et al. The interferon gene signature is increased in patients with early treatment-naïve rheumatoid arthritis and predicts a poorer response to initial therapy. *J Allergy Clin Immunol*. 2018;141(1):445-448.e4.
- [133] Laurent P et al. Innate Immunity in Systemic Sclerosis Fibrosis: Recent Advances. *Front Immunol*. 2018;9:1702.
- [134] Pollard KM et al. Induction of Systemic Autoimmunity by a Xenobiotic Requires Endosomal TLR Trafficking and Signaling from the Late Endosome and Endolysosome but Not Type I IFN. *J Immunol*. 2017;199(11):3739-3747.
- [135] Rodríguez-Carrio J et al. Heterogeneity of the Type I Interferon Signature in Rheumatoid Arthritis: A Potential Limitation for Its Use As a Clinical Biomarker. *Front Immunol*. 2018;8:2007.
- [136] Joosten LA et al. Toll-like receptors and chronic inflammation in rheumatic diseases: new developments. *Nat Rev Rheumatol*. 2016;12(6):344-357.
- [137] Dinarello CA. Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol Rev*. 2018;281(1):8-27.
- [138] Santegoets KC et al. Toll-like receptors in rheumatic diseases: are we paying a high price for our defense against bugs? *FEBS Lett*. 2011;585(23):3660-3666.
- [139] Celhar T et al. Toll-like receptors in systemic lupus erythematosus: potential for personalized treatment. *Front Pharmacol*. 2014;5:265.



- 
- [140] Wu YW et al. Toll-like receptors: potential targets for lupus treatment. *Acta Pharmacol Sin.* 2015;36(12):1395-1407.
- [141] Clancy RM et al. Endosomal Toll-like receptors in clinically overt and silent autoimmunity. *Immunol Rev.* 2016;269(1):76-84.
- [142] Pretorius E et al. Major involvement of bacterial components in rheumatoid arthritis and its accompanying oxidative stress, systemic inflammation and hypercoagulability. *Exp Biol Med (Maywood).* 2017;242(4):355-373.
- [143] Bhattacharyya S et al. Emerging roles of innate immune signaling and toll-like receptors in fibrosis and systemic sclerosis. *Curr Rheumatol Rep.* 2015;17(1):474.
- [144] Petrackova A et al. Cross-Disease Innate Gene Signature: Emerging Diversity and Abundance in RA Comparing to SLE and SSc. *J Immunol Res.* 2019;2019:3575803.
- [145] Petrackova A et al. Revealed heterogeneity in rheumatoid arthritis based on multivariate innate signature analysis. *Clin Exp Rheumatol.* 2020;38(2):289-298.
- [146] Hallek M et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood.* 2008;111(12):5446-5456.
- [147] Genome Aggregation Database (gnomAD) [(accessed on 20 August 2021)]; Available online: <http://gnomad.broadinstitute.org/>
- [148] Rajkumar SV et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol.* 2014;15(12):e538-e548.
- [149] Mlynarcikova M et al. Molecular Cytogenetic Analysis of Chromosome 8 Aberrations in Patients With Multiple Myeloma Examined in 2 Different Stages, at Diagnosis and at Progression/Relapse. *Clin Lymphoma Myeloma Leuk.* 2016;16(6):358-365.
- [150] Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;40(9):1725.
- [151] Bombardier C et al. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. *Arthritis Rheum.* 1992;35(6):630-640.
- [152] Schneiderova P et al. Serum protein fingerprinting by PEA immunoassay coupled with a pattern-recognition algorithms distinguishes MGUS and multiple myeloma. *Oncotarget.* 2016;8(41):69408-69421.
- [153] Aletaha D et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis.* 2010;69(9):1580-1588.

- 
- [154] van den Hoogen F et al. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis.* 2013;72(11):1747-1755.
- [155] Tomankova T et al. Comparison of periprosthetic tissues in knee and hip joints: differential expression of CCL3 and DC-STAMP in total knee and hip arthroplasty and similar cytokine profiles in primary knee and hip osteoarthritis. *Osteoarthritis Cartilage.* 2014;22(11):1851-1860.
- [156] Falkenberg VR et al. Identification of Phosphoglycerate Kinase 1 (PGK1) as a reference gene for quantitative gene expression measurements in human blood RNA. *BMC Res Notes.* 2011;4:324.
- [157] Ochodkova E et al. Graph construction based on local representativeness. In: Cao Y., Chen J., editors. *Computing and Combinatorics. COCOON 2017. Lecture Notes in Computer Science.* 2017;10392:654-665.
- [158] Moustafa RE. Andrews curves. *WIREs Computational Statistics.* 2011;3(4):373-382.
- [159] Andrews DF. Plots of high-dimensional data. *Biometrics.* 1972;28(1):125-136.
- [160] García-Osorio C et al. Visualization of high-dimensional data via orthogonal curves. *J Univers. Comput Sci.* 2005;11:1806-1819.
- [161] Agrawal R et al. Mining association rules between sets of items in large databases. *SIGMOD '93 Proceedings of the 1993 ACM SIGMOD international conference on Management of data; May 1993; Washington, DC, USA; 207-216.*
- [162] Quail MA et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics.* 2012;13:341.
- [163] Tan G et al. Long fragments achieve lower base quality in Illumina paired-end sequencing. *Sci Rep.* 2019;9(1):2856.
- [164] Baechler EC et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U S A.* 2003;100(5):2610-2615.
- [165] Feng X et al. Identification of interferon-inducible genes as diagnostic biomarker for systemic lupus erythematosus. *Clin Rheumatol.* 2015;34(1):71-79.
- [166] Puente XS et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature.* 2015;526(7574):519-524.
- [167] VarSome Database [(accessed on 14 July 2021)]; Available online: <https://varsome.com>
- [168] Skowronska A et al. Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom Leukemia Research Fund Chronic Lymphocytic Leukemia 4 trial. *J Clin Oncol.* 2012;30(36):4524-4532.

- 
- [169] Yun X et al. Recent progress of prognostic biomarkers and risk scoring systems in chronic lymphocytic leukemia. *Biomark Res.* 2020;8:40.
- [170] Nicolas L et al. Cutting Edge: ATM Influences Germinal Center Integrity. *J Immunol.* 2019;202(11):3137-3142.
- [171] Spring K et al. Mice heterozygous for mutation in *Atm*, the gene involved in ataxia-telangiectasia, have heightened susceptibility to cancer. *Nat Genet.* 2002;32(1):185-190.
- [172] Yamamoto K et al. Kinase-dead ATM protein is highly oncogenic and can be preferentially targeted by Topo-isomerase I inhibitors. *Elife.* 2016;5:e14709.
- [173] Choi M et al. ATM Mutations in Cancer: Therapeutic Implications. *Mol Cancer Ther.* 2016;15(8):1781-1791.
- [174] Antoniou A et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet.* 2003;72(5):1117-1130.
- [175] Foulkes WD. Inherited susceptibility to common cancers. *N Engl J Med.* 2008;359(20):2143-2153.
- [176] Jares P et al. Molecular pathogenesis of mantle cell lymphoma. *J Clin Invest.* 2012;122(10):3416-3423.
- [177] Pan-Hammarström Q et al. Disparate roles of ATR and ATM in immunoglobulin class switch recombination and somatic hypermutation. *J Exp Med.* 2006;203(1):99-110.
- [178] Yin B et al. ATM prevents unattended DNA double strand breaks on site and in generations to come. *Cancer Biol Ther.* 2007;6(12):1837-1839.
- [179] Coffey DG et al. Ultradeep, targeted sequencing reveals distinct mutations in blood compared to matched bone marrow among patients with multiple myeloma. *Blood Cancer J.* 2019;9(10):77.
- [180] Walker BA et al. A high-risk, Double-Hit, group of newly diagnosed myeloma identified by genomic analysis. *Leukemia.* 2019;33(1):159-170.
- [181] Ryland GL et al. Novel genomic findings in multiple myeloma identified through routine diagnostic sequencing. *J Clin Pathol.* 2018;71(10):895-899.
- [182] Vikova V et al. Comprehensive characterization of the mutational landscape in multiple myeloma cell lines reveals potential drivers and pathways associated with tumor progression and drug resistance. *Theranostics.* 2019;9(2):540-553.
- [183] Ross FM et al. Report from the European Myeloma Network on interphase FISH in multiple myeloma and related disorders. *Haematologica.* 2012;97(8):1272-1277.
- [184] Thakurta A et al. High subclonal fraction of 17p deletion is associated with poor prognosis in multiple myeloma. *Blood.* 2019;133(11):1217-1221.

- 
- [185] Lakshman A et al. Natural history of multiple myeloma with de novo del(17p). *Blood Cancer J.* 2019;9(3):32.
- [186] Shin HT et al. Prevalence and detection of low-allele-fraction variants in clinical cancer samples. *Nat Commun.* 2017;8(1):1377.
- [187] Ma X et al. Analysis of error profiles in deep next-generation sequencing data. *Genome Biol.* 2019;20(1):50.
- [188] Smith T et al. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* 2017;27(3):491-499.
- [189] McDonough SJ et al. Use of FFPE-derived DNA in next generation sequencing: DNA extraction methods. *PLoS One.* 2019;14(4):e0211400.
- [190] Ascierio PA et al. Preanalytic Variables and Tissue Stewardship for Reliable Next-Generation Sequencing (NGS) Clinical Analysis. *J Mol Diagn.* 2019;21(5):756-767.
- [191] Gröndal G et al. Cytokine production, serum levels and disease activity in systemic lupus erythematosus. *Clin Exp Rheumatol.* 2000;18(5):565-570.
- [192] Favilli F et al. IL-18 activity in systemic lupus erythematosus. *Ann N Y Acad Sci.* 2009;1173:301-309.
- [193] Yajima N et al. Elevated levels of soluble fractalkine in active systemic lupus erythematosus: potential involvement in neuropsychiatric manifestations. *Arthritis Rheum.* 2005;52(6):1670-1675.
- [194] Hrycek E et al. Serum levels of selected chemokines in systemic lupus erythematosus patients. *Rheumatol Int.* 2013;33(9):2423-2427.
- [195] Bauer JW et al. Elevated serum levels of interferon-regulated chemokines are biomarkers for active human systemic lupus erythematosus. *PLoS Med.* 2006;3(12):e491.
- [196] Masi L et al. P8 - Measurement of Fibroblast Growth Factor-23 (FGF23) in the Serum of Patients Affected by Juvenile Systemic Lupus Erythematosus: A Possible Marker of Kidney Damage. *Clin Cases Miner Bone Metab.* 2010;7(3):214.
- [197] Lee AS et al. SIRT2 ameliorates lipopolysaccharide-induced inflammation in macrophages. *Biochem Biophys Res Commun.* 2014;450(4):1363-1369.
- [198] Wang X et al. Sirtuin-2 Regulates Sepsis Inflammation in ob/ob Mice. *PLoS One.* 2016;11(8):e0160431.
- [199] Lemmers B et al. Essential role for caspase-8 in Toll-like receptors and NF $\kappa$ B signaling. *J Biol Chem.* 2007;282(10):7416-7423.
- [200] Gurung P et al. FADD and caspase-8 mediate priming and activation of the canonical and noncanonical Nlrp3 inflammasomes. *J Immunol.* 2014;192(4):1835-1846.

- 
- [201] Lalor SJ et al. Caspase-1-processed cytokines IL-1 $\beta$  and IL-18 promote IL-17 production by  $\gamma\delta$  and CD4 T cells that mediate autoimmunity. *J Immunol.* 2011;186(10):5738-5748.
- [202] Chi W et al. Caspase-8 promotes NLRP1/NLRP3 inflammasome activation and IL-1 $\beta$  production in acute glaucoma. *Proc Natl Acad Sci U S A.* 2014;111(30):11181-11186.
- [203] Sahebari M et al. Correlation between serum concentrations of soluble Fas (CD95/Apo-1) and IL-18 in patients with systemic lupus erythematosus. *Rheumatol Int.* 2012;32(3):601-606.
- [204] Calvani N et al. Up-regulation of IL-18 and predominance of a Th1 immune response is a hallmark of lupus nephritis. *Clin Exp Immunol.* 2004;138(1):171-178.
- [205] Ebmeier CC et al. Human thyroid phenol sulfotransferase enzymes 1A1 and 1A3: activities in normal and diseased thyroid glands, and inhibition by thyroid hormones and phytoestrogens. *J Clin Endocrinol Metab.* 2003;89(11):5597-5605.
- [206] Choe JY et al. Serum TWEAK as a biomarker for disease activity of systemic lupus erythematosus. *Inflamm Res.* 2016;65(6):479-488.
- [207] al-Janadi M et al. Cytokine profile in systemic lupus erythematosus, rheumatoid arthritis, and other rheumatic diseases. *J Clin Immunol.* 1993;13(1):58-67.
- [208] Mangieri D et al. Eotaxin/CCL11 in idiopathic retroperitoneal fibrosis. *Nephrol Dial Transplant.* 2012;27(10):3875-3884.
- [209] Tacke F et al. Up-regulated eotaxin plasma levels in chronic liver disease patients indicate hepatic inflammation, advanced fibrosis and adverse clinical course. *J Gastroenterol Hepatol.* 2007;22(8):1256-1264.
- [210] Huaux F et al. Role of Eotaxin-1 (CCL11) and CC chemokine receptor 3 (CCR3) in bleomycin-induced lung injury and fibrosis. *Am J Pathol.* 2005;167(6):1485-1496.
- [211] Diny NL et al. Macrophages and cardiac fibroblasts are the main producers of eotaxins and regulate eosinophil trafficking to the heart. *Eur J Immunol.* 2016;46(12):2749-2760.
- [212] Friese RS et al. Matrix metalloproteinases: discrete elevations in essential hypertension and hypertensive end-stage renal disease. *Clin Exp Hypertens.* 2009;31(7):521-533.
- [213] Barksby HE et al. Matrix metalloproteinase 10 promotion of collagenolysis via procollagenase activation: implications for cartilage degradation in arthritis. *Arthritis Rheum.* 2006;54(10):3244-3253.
- [214] Wu T et al. Antibody-Array-Based Proteomic Screening of Serum Markers in Systemic Lupus Erythematosus: A Discovery Study. *J Proteome Res.* 2016;15(7):2102-2114.

- [215] Wang H et al. Glomerular transcriptional profiles reveal the candidate biomarkers diagnostic for the progression of lupus nephritis from acute to chronic stages. *Arthritis Rheum.* 2008;58:S317.
- [216] Tesch GH et al. Monocyte chemoattractant protein 1-dependent leukocytic infiltrates are responsible for autoimmune disease in MRL-Fas(lpr) mice. *J Exp Med.* 1999;190(12):1813-1824.
- [217] Chan RW et al. Intrarenal cytokine gene expression in lupus nephritis. *Ann Rheum Dis.* 2007;66(7):886-892.
- [218] Menke J et al. Colony-stimulating factor-1: a potential biomarker for lupus nephritis. *J Am Soc Nephrol.* 2015;26(2):379-389.
- [219] Menke J et al. Circulating CSF-1 promotes monocyte and macrophage phenotypes that enhance lupus nephritis. *J Am Soc Nephrol.* 2009;20(12):2581-2592.
- [220] Baranda L et al. IL-15 and IL-15R in leucocytes from patients with systemic lupus erythematosus. *Rheumatology (Oxford).* 2005;44(12):1507-1513.
- [221] Zhao M et al. Increased 5-hydroxymethylcytosine in CD4(+) T cells in systemic lupus erythematosus. *J Autoimmun.* 2016;69:64-73.
- [222] Goules A et al. Elevated levels of soluble CD40 ligand (sCD40L) in serum of patients with systemic autoimmune diseases. *J Autoimmun.* 2006;26(3):165-171.
- [223] Yellin MJ et al. Immunohistologic analysis of renal CD40 and CD40L expression in lupus nephritis and other glomerulonephritides. *Arthritis Rheum.* 1997;40(1):124-134.
- [224] Ripoll E et al. CD40 gene silencing reduces the progression of experimental lupus nephritis modulating local milieu and systemic mechanisms. *PLoS One.* 2013;8(6):e65068.
- [225] Nakatani K et al. Fractalkine expression and CD16+ monocyte accumulation in glomerular lesions: association with their severity and diversity in lupus models. *Am J Physiol Renal Physiol.* 2010;299(1):F207-F216.
- [226] Inoue A et al. Antagonist of fractalkine (CX3CL1) delays the initiation and ameliorates the progression of lupus nephritis in MRL/lpr mice. *Arthritis Rheum.* 2005;52(5):1522-1533.
- [227] Han GM et al. Analysis of gene expression profiles in human systemic lupus erythematosus using oligonucleotide microarray. *Genes Immun.* 2003;4(3):177-186.
- [228] Tucci M et al. Glomerular accumulation of plasmacytoid dendritic cells in active lupus nephritis: role of interleukin-18. *Arthritis Rheum.* 2008;58(1):251-262.
- [229] Bracci-Laudiero L et al. Increased levels of NGF in sera of systemic lupus erythematosus patients. *Neuroreport.* 1993;4(5):563-565.

- [230] Antonucci MT et al. Nerve growth factor and its monocyte receptors are affected in kidney disease. *Nephron Clin Pract.* 2009;111(1):c21-c28.
- [231] Morigi M et al. A previously unrecognized role of C3a in proteinuric progressive nephropathy. *Sci Rep.* 2016;6:28445.
- [232] Onodera H et al. Increased plasma GDNF levels in patients with chronic renal diseases. *Nephrol Dial Transplant.* 1999;14(6):1604-1605.
- [233] Kalechman Y et al. Production of the novel mesangial autocrine growth factors GDNF and IL-10 is regulated by the immunomodulator AS101. *J Am Soc Nephrol.* 2003;14(3):620-630.
- [234] Brentano F et al. RNA released from necrotic synovial fluid cells activates rheumatoid arthritis synovial fibroblasts via Toll-like receptor 3. *Arthritis Rheum.* 2005;52(9):2656-2665.
- [235] Kim SJ et al. Angiogenesis in rheumatoid arthritis is fostered directly by toll-like receptor 5 ligation and indirectly through interleukin-17 induction. *Arthritis Rheum.* 2013;65(8):2024-2036.
- [236] Elshabrawy HA et al. TLRs, future potential therapeutic targets for RA. *Autoimmun Rev.* 2017;16(2):103-113.
- [237] Chamberlain ND et al. TLR5, a novel and unidentified inflammatory mediator in rheumatoid arthritis that correlates with disease activity score and joint TNF- $\alpha$  levels. *J Immunol.* 2012;189(1):475-483.
- [238] Kim SJ et al. Ligation of TLR5 promotes myeloid cell infiltration and differentiation into mature osteoclasts in rheumatoid arthritis and experimental arthritis. *J Immunol.* 2014;193(8):3902-3913.
- [239] Cavalli G et al. Treating experimental arthritis with the innate immune inhibitor interleukin-37 reduces joint and systemic inflammation. *Rheumatology (Oxford).* 2016;55(12):2220-2229.
- [240] Lacerte P et al. Overexpression of TLR2 and TLR9 on monocyte subsets of active rheumatoid arthritis patients contributes to enhance responsiveness to TLR agonists. *Arthritis Res Ther.* 2016;18:10.
- [241] Aucott H et al. Ligation of free HMGB1 to TLR2 in the absence of ligand is negatively regulated by the C-terminal tail domain. *Mol Med.* 2018;24(1):19.
- [242] Dinarello CA. Immunological and inflammatory functions of the interleukin-1 family. *Annu Rev Immunol.* 2009;27:519-550.
- [243] Gasse P et al. IL-1R1/MyD88 signaling and the inflammasome are essential in pulmonary inflammation and fibrosis in mice. *J Clin Invest.* 2007;117(12):3776-3799.
- [244] Chen SC et al. Dynamic association rules for gene expression data analysis. *BMC Genomics.* 2015;16:786.

- [245] Iwahashi M et al. Expression of Toll-like receptor 2 on CD16+ blood monocytes and synovial tissue macrophages in rheumatoid arthritis. *Arthritis Rheum.* 2004;50(5):1457-1467.
- [246] Chamberlain ND et al. Ligation of TLR7 by rheumatoid arthritis synovial fluid single strand RNA induces transcription of TNF $\alpha$  in monocytes. *Ann Rheum Dis.* 2013;72(3):418-426.
- [247] Edwards CK et al. Combined anti-tumor necrosis factor- $\alpha$  therapy and DMARD therapy in rheumatoid arthritis patients reduces inflammatory gene expression in whole blood compared to DMARD therapy alone. *Front Immunol.* 2012;3:366.
- [248] Ramírez-Pérez S et al. High expression of interleukine-1 receptor antagonist in rheumatoid arthritis: Association with IL1RN\*2/2 genotype. *Autoimmunity.* 2017;50(8):468-475.
- [249] Shao XT et al. Expression of interleukin-18, IL-18BP, and IL-18R in serum, synovial fluid, and synovial tissue in patients with rheumatoid arthritis. *Clin Exp Med.* 2009;9(3):215-221.
- [250] Huang Q et al. Increased macrophage activation mediated through toll-like receptors in rheumatoid arthritis. *Arthritis Rheum.* 2007;56(7):2192-2201.
- [251] Oosting M et al. Human TLR10 is an anti-inflammatory pattern-recognition receptor. *Proc Natl Acad Sci U S A.* 2014;111(42):E4478-E4484.
- [252] Jiang S et al. TLR10 Is a Negative Regulator of Both MyD88-Dependent and -Independent TLR Signaling. *J Immunol.* 2016;196(9):3834-3841.
- [253] Lai Y et al. Expression profiles of toll-like receptor signaling pathway related genes in microscopic polyangiitis in Chinese people. *Int J Clin Exp Pathol.* 2016;9:5515-5524.
- [254] Celhar T et al. Toll-Like Receptor 9 Deficiency Breaks Tolerance to RNA-Associated Antigens and Up-Regulates Toll-Like Receptor 7 Protein in Sle1 Mice. *Arthritis Rheumatol.* 2018;70(10):1597-1609.
- [255] Tan FK et al. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. *Rheumatology (Oxford).* 2006;45(6):694-702.
- [256] Martínez-Godínez MA et al. Expression of NLRP3 inflammasome, cytokines and vascular mediators in the skin of systemic sclerosis patients. *Isr Med Assoc J.* 2015;17(1):5-10.
- [257] Ihn H et al. Demonstration of interleukin 8 in serum samples of patients with localized scleroderma. *Arch Dermatol.* 1994;130(10):1327-1328.
- [258] Kaplanski G. Interleukin-18: Biological properties and role in disease pathogenesis. *Immunol Rev.* 2018;281(1):138-153.



- [259] Torices S et al. A functional variant of TLR10 modifies the activity of NFkB and may help predict a worse prognosis in patients with rheumatoid arthritis. *Arthritis Res Ther.* 2016;18(1):221.
- [260] Liu Y et al. Rare deleterious germline variants and risk of lung cancer. *NPJ Precis Oncol.* 2021;5(1):12.
- [261] Huang KL et al. Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell.* 2018;173(2):355-370.e14.
- [262] Mosquera Orgueira A et al. Detection of Rare Germline Variants in the Genomes of Patients with B-Cell Neoplasms. *Cancers (Basel).* 2021;13(6):1340.

## 9 Appendix - full text publications related to the thesis

### APPENDIX A

**Petrackova A**, Savara J, Turcsanyi P, Gajdos P, Papajik T, Kriegova E. Rare germline pathogenic *ATM* variants are underestimated in cancer. – manuscript under review

**Rare germline pathogenic *ATM* variants are underestimated in cancer**

Anna Petrackova<sup>1</sup>, Jakub Savara<sup>2</sup>, Peter Turcsanyi<sup>3</sup>, Petr Gajdos<sup>2</sup>, Tomas Papajik<sup>3</sup>, Eva Kriegova<sup>1\*</sup>

<sup>1</sup>Department of Immunology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czech Republic

<sup>2</sup>Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava, Ostrava, Czech Republic

<sup>3</sup>Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czech Republic

\*Corresponding author: Eva Kriegova, Department of Immunology, Palacky University and University Hospital, Olomouc, Hnevotinska 3, 775 15 Olomouc, Czech Republic; e-mail: eva.kriegova@email.cz, phone: +420 585 632 280.

**Keywords:** rare germline variants; 11q deletion; ataxia-telangiectasia mutated gene; diagnostic next-generation sequencing; variants of uncertain significance

**Running head:** *ATM* germline variants in cancer

## **Abstract**

**Background:** Despite the emerging role of rare pathogenic germline variations in tumour suppressor genes in cancer pathogenesis, they are still neglected in diagnostics as they are often referred to as variants of uncertain significance (VUS) or overlooked based on tumour-normal filtering. Herein, we focus on the *ATM* (ataxia-telangiectasia mutated) gene, whose somatic disruption is common in solid and haemato-oncological cancers.

**Patients and methods:** A complete *ATM* coding sequence was analysed in own real-world cohort of 336 patients with chronic lymphocytic leukaemia (CLL) and public tumour-normal datasets (445 CLL, 75 mantle cell lymphoma, 216 metastatic breast cancer, 104 lung cancer patients). All next-generation sequencing datasets of tumour-normal samples were evaluated for rare germline and somatic variants, as well as copy number aberration (CNA).

**Results:** Analysis of own and public datasets together with a systemic review of published data revealed that the prevalence of rare pathogenic *ATM* variants is underestimated in cancer, as about two-thirds of these variants are missed due to variant interpretation issues, increasing the prevalence up to 8%. These variants predispose to the acquiring of an 11q deletion (del(11q)), leading to biallelic *ATM* inactivation, which is present in 40–80% of patients with *ATM* pathogenic/VUS-predicted pathogenic variants in evaluated datasets, always with a loss of the wild type allele. Functional study in *ex vivo* primary CLL cells showed that VUS-predicted pathogenic *ATM* variants slightly decrease the ATM kinase activity, however, concurrent del(11q) resulted in a complete loss of ATM activity in tumour cells in the same manner as somatic *ATM* variants. Similar to somatic *ATM* and/or *TP53* disruptions, CLL patients with germline *ATM* variants had reduced progression-free survival, even on novel agents.

**Conclusion:** Our data highlights the need to implement rare germline pathogenic *ATM* variants, including VUS-predicted pathogenic, into diagnostics and clinical decision-making.

## Introduction

Rare germline variations, occurring in the population with less than 0.5% frequency, have been recently revealed to have a crucial role in cancer aetiology, especially when they occur in tumour suppressor genes<sup>1,2,3</sup>. Multiple lines of evidence showed that rare germline variants affect drug sensitivity, enhance the likelihood of additional somatic aberrations and accelerate cancer progression<sup>4,5,6</sup>. However, in cancer diagnostics, they are often i) referred to as variants of uncertain significance (VUS), as there is limited knowledge of their clinical and functional impact, or ii) overseen if paired tumour-normal variant analysis is used to filter tumour-only variants<sup>7</sup>. One of the genes, where rare, protein-altering germline variants occur, is *ATM* (ataxia-telangiectasia mutated), a tumour suppressor essential for genome stability by regulating the DNA double-strand break response. Despite the known association of inherited rare variations in the *ATM* gene with the autosomal recessive disorder ataxia telangiectasia<sup>8</sup>, there is growing evidence of their role in cancer pathogenesis<sup>3,9,10,11</sup>, including chronic lymphocytic leukaemia (CLL)<sup>1,12</sup>. With the exception of truncating alleles causing ataxia telangiectasia<sup>8</sup>, the clinical significance of most rare *ATM* germline variants is not fully elucidated as the majority of data is available only for *ATM* somatic disruption in cancer. In CLL, somatic *ATM* disruption (mutations/deletions) is recognised as a negative prognostic event, comparable to *TP53* abnormalities<sup>13</sup>.

Our motivation for investigating rare germline, protein-altering variants in the *ATM* gene in CLL was the number of VUS variants in *ATM* that we have reported in our diagnostic tests of a panel of CLL-associated genes (*TP53*, *ATM*, *BIRC3*, *NOTCH1*, *SF3B1*, *POT1*, *MYD88*, *FBXW7*, *XPO1*, *EGR2*, *NFKBIE*, *RPS15*). We aimed here to elucidate the clinical and functional impact of rare germline *ATM* variants in CLL patients, particularly those on novel agents, and compared them with adverse somatic aberrations in CLL (del(11q), *ATM* mutation, del(17p), *TP53* mutation, *IGHV* status). Additionally, we analysed the annotation of rare

germline *ATM* pathogenic/predicted pathogenic variants in diagnostic databases (ClinVar, VarSome) and their frequency in public whole exome sequencing (WES) tumour-normal datasets of CLL, mantle cell lymphoma (MCL), metastatic breast and lung cancer patients. A comparison of published data on multiple cancers showed that the prevalence of rare *ATM* pathogenic/predicted pathogenic variants is underestimated across all cancers.

## **Patients and Methods**

A real-world cohort of 336 patients with CLL was recruited between 2016–2020 in a single tertiary haematological centre at University Hospital Olomouc. All enrolled patients were diagnosed in accordance with the international criteria<sup>14</sup> for clinical and genetic characteristics see Table 1. The median follow-up was 25 months (min–max 1–193 months). Additionally, 198 healthy controls (113/83 female/male, median age 78) with no history of cancer or autoimmune diseases were recruited based on the patients' records.

A complete *ATM* coding sequence was analysed by deep next-generation sequencing (NGS) of whole blood DNA as reported previously<sup>15,16,17</sup>. A germline/somatic origin of detected *ATM* variants was confirmed by analysis of patients' germline saliva DNA (available for 93% of patients). Rare variants were defined as those having a minor allele frequency (MAF) <0.5%, according to the gnomAD database in any population<sup>18</sup>. Only protein-altering variants (missense, nonsense, frameshift and splice region) were investigated and annotated by VarSome and ClinVar databases; for missense variants, Sift and PolyPhen prediction tools were used. *ATM* activity was assessed in CLL and T cells from cryopreserved peripheral blood mononuclear cells based on the phosphorylation of *ATM*-specific substrate KAP1 and *ATM* autophosphorylation of Ser-1981, after exposure to etoposide (Data Supplement).

All patients and controls provided written informed consent about the usage of biological materials for the purpose of this study, which was performed in accordance with the

Helsinki Declaration and approved by the Ethics Committee of the University Hospital and Palacký University Olomouc.

Moreover, following public whole exome/genome sequencing, datasets on tumour-normal samples in CLL (public datasets EGAD00001001464, EGAD00001001466), MCL (EGAD00001006159), metastatic breast (EGAD00001002747) and lung (EGAD00001004027, EGAD00001003960) cancers were evaluated for rare germline and somatic variants, as well as copy number aberration (CNA) in the 11q region using the VarScan (v.2.4.4) tool. Statistical analyses were performed using R software.

## **Results**

### ***Prevalence and annotation of rare germline pathogenic/predicted pathogenic ATM variants in CLL***

Four per cent of CLL patients carried rare, protein-altering germline *ATM* pathogenic/predicted pathogenic variants, as demonstrated in our patients (4.5%, 15/336, cohort A) and published datasets (3.4%, 15/445, cohort B, public datasets EGAD00001001464, EGAD00001001466<sup>19</sup>). Of the detected rare *ATM* germline pathogenic/predicted pathogenic variants in CLL cohorts, 59% and 62% were classified as a VUS according to ClinVar and VarSome databases, respectively (Figure 1). In the ClinVar and VarSome databases, respectively, 14% and 10% of detected variants did not have rs numbers or were not listed. Besides *ATM* pathogenic/predicted pathogenic variants, rare variants with benign predictions were also detected in ~3% of CLL patients (A: 2.4%, 8/336; B: 3.6%, 16/445) and 1.5% of healthy controls (3/196).

*ATM* pathogenic/predicted pathogenic variants were associated with an increased risk of CLL (odds ratio 9.1, 95% CI: 1.2–69.5) when compared with healthy controls (1/196, 0.5%,  $p = 0.03$ ), but not with age at diagnosis and time to the first treatment when compared with CLL patients without these variants ( $p > 0.05$ ). All patients carried a single variant across the whole

*ATM* gene with no apparent hotspot, except for one case with two variants. Variants were predominantly the missense type in both CLL cohorts (A, B) (69.0%, 20/29, Figure 1, Table S1). Truncating *ATM* alleles (nonsense, frameshift) known to cause ataxia telangiectasia were less frequent (17.2%, 5/29). The minor allele frequency (MAF) of *ATM* rare pathogenic/predicted pathogenic variants detected in CLL cohorts varied from extremely rare variants (singletons that were not found in population databases) to the highest MAF 0.005 observed in any population according to gnomAD. The majority (86%) of variants had MAF  $\leq 0.001$ , of which 10 were singletons. A family history of cancer was higher in patients with rare *ATM* pathogenic/predicted pathogenic variants (87.5%) when compared with patients with wild type *ATM* (43.1%,  $p = 0.02$ ), based on the available self-reports from CLL cohort A.

#### ***Rare germline pathogenic/predicted pathogenic ATM variants in multiple cancers***

In other analysed cancers, rare *ATM* pathogenic/predicted pathogenic variants occurred in 8% of MCL (6/75, public dataset EGAD00001006159<sup>20</sup>), 5% of metastatic breast cancer (11/216, EGAD00001002747<sup>21</sup>) and 6% of lung cancer (2/36, EGAD00001004027, cancer subtype undefined) patients (Figure 1). In patients with squamous cell lung cancer, rare *ATM* pathogenic/predicted pathogenic variants were not detected (0/104, EGAD00001003960). All rare germline *ATM* pathogenic/predicted pathogenic variants in all cohorts evaluated were detected as heterozygous in the paired germline sample, and each individual had a different variant.

Additionally, we systematically reviewed the prevalence of rare germline *ATM* pathogenic/predicted pathogenic variants in multiple cancers and compared it with our data, as shown in Figure 1.



***Association of rare germline pathogenic/predicted pathogenic ATM variants with adverse somatic aberrations in CLL***

Patients with rare germline *ATM* pathogenic/predicted pathogenic variants also had del(11q), 40% (6/15) in our CLL cohort A, and 27% (4/15) in cohort B. In both CLL cohorts, patients with rare germline *ATM* pathogenic/predicted pathogenic variants were ~3 times more likely to have del(11q), always with a loss of the wild type allele, when compared with patients without these variants and *ATM* somatic mutations (A: relative risk (RR): 2.46, 95% CI: 1.25–4.84,  $p = 0.009$ ; B: 3.87, 1.55–9.66,  $p = 0.004$ ). In patients with del(11q), the median variant allele frequency (VAF) of rare germline *ATM* variant in tumor samples reached 84% (A, min-max 50–93%) and 71% (B, 52–100%), respectively. Rare *ATM* variants with a benign prediction did not associate with the del(11q) ( $p > 0.05$ ). Similar to CLL, 83% (5/6) of MCL and 56% (6/11) of metastatic breast cancer patients with rare germline *ATM* pathogenic/predicted pathogenic variants had concurrent del(11q), as detected by CNA analysis in WES datasets.

Regarding somatic *ATM* variants, the vast majority of patients with rare *ATM* pathogenic/predicted pathogenic variants did not have additional somatic *ATM* variants in both CLL cohorts (A: 93.3%, 14/15; B: 93.3%, 14/15) (Figure 2). Rare *ATM* pathogenic/predicted pathogenic variants were different from somatic *ATM* variants, except for one variant type (p.C2488Y), as shown in Figure 3 (Table S2).

Regarding the association with *IGHV* status, rare germline *ATM* pathogenic/predicted pathogenic variants associated with unmutated *IGHV* that was detected in 93.3% of patients (14/15) in cohort A (A: RR: 1.50, 95% CI: 1.28–1.77,  $p < 0.0001$ , Figure 2). However, this association was not observed in cohort B when all variants, irrespective of position and variant type, were evaluated ( $p > 0.05$ ); 40% (6/15) of patients carrying rare germline *ATM* pathogenic/predicted pathogenic variants had unmutated *IGHV* status. Subanalysis in patients carrying only truncating or missense variants in the *ATM* kinase domain (aa2712-2962) (A: n

= 5, B: n = 4) revealed that all these patients developed CLL with unmutated *IGHV* (A: RR: 1.62, 95% CI: 1.49–1.77,  $p < 0.0001$ ; B: RR: 2.87, 95% CI: 2.52–3.27,  $p < 0.0001$ ).

Regarding *TP53* disruption (del(17p) and/or *TP53* mutation), the majority of patients carrying rare germline *ATM* pathogenic/predicted pathogenic variants did not acquire a *TP53* disruption (A: 80.0%, 12/15; B: 86.7%, 13/15; Figure 2). Of note, two patients harbouring rare germline *ATM* pathogenic/predicted pathogenic variants who, during the disease course, also acquired del(11q), del(17p) and *TP53* mutations were refractory to idelalisib-rituximab, ibrutinib and venetoclax, and one of those patients had a history of breast cancer and developed a Richter transformation eight months later.

*ATM* rare pathogenic/predicted pathogenic variants were not associated with other adverse somatic aberrations (trisomy 12, complex karyotype, mutated *SF3B1*, *BIRC3* and *NOTCH1*;  $p > 0.05$ ) in our CLL patients.

#### ***Functional analysis of rare germline ATM pathogenic/predicted pathogenic variants in CLL***

To evaluate the functional impact of rare germline *ATM* pathogenic/predicted pathogenic variants, we analysed a decrease in ATM activity in CLL and T cells obtained from 25 CLL patients, as assessed by KAP1 phosphorylation and ATM autophosphorylation of Ser-1981 (Figure 4, Table S3). All analysed rare germline *ATM* pathogenic/predicted pathogenic variants, except for one variant (R2443\*), were annotated by ClinVar and VarSome as VUS (missense variants: K224E, L480F, R717P, Y1442H, R2032K, Y2755S; splice region variant: c.4236+4A>C). For comparison, the following somatic *ATM* variants (VAF min–max 8–94%) were analysed: L439H, L1283fs, C1674W, L2698F, Q2714del and T2773I. In samples with del(11q), this aberration was present in more than 50% of CLL cells as assessed by FISH analysis<sup>22</sup>.

CLL cells from patients carrying rare *ATM* germline pathogenic/predicted pathogenic variants, together with del(11q), completely lost ATM activity, same as in patients with concurrent somatic *ATM* pathogenic variant and del(11q). Heterozygous rare germinal *ATM* variants pathogenic/predicted pathogenic alone resulted in only a partial decrease in ATM activity (reduction of 10–20%) in both CLL and T cells, and the activity varied depending on the individual variant type (Figure 4, Table S3).

***Progression-free survival of treated CLL patients carrying rare germline ATM pathogenic/predicted pathogenic variants***

We evaluated the impact of rare germline *ATM* pathogenic/predicted pathogenic variants on progression-free survival (PFS) of patients treated with novel agents and chemoimmunotherapy. Patients on novel agents carrying rare germline *ATM* pathogenic/predicted pathogenic variants had shorter PFS (median 24 months) than patients with wild type *ATM* and/or somatic *ATM* and/or *TP53* disruption (49/40 months,  $p < 0.05$ ) and did not differ from patients with *ATM* and/or *TP53* somatic disruption (40 months,  $p > 0.05$ ) (Figure 2). Similar results were obtained for patients on chemoimmunotherapy (Figure 2). Subanalysis in only patients with unmutated *IGHV* revealed the same negative association of rare germline *ATM* pathogenic/predicted pathogenic variants on PFS in patients treated with novel agents (Figure S1).

**Discussion**

This study revealed the underestimation of rare pathogenic germline *ATM* variants in multiple cancers. We showed an example of CLL that half of the rare *ATM* germline variants, currently classified as VUS, are pathogenic. Moreover, they predispose to cancer and del(11q) acquirement, leading to biallelic *ATM* inactivation (Figure 5). Our data highlight the need for

the implementation of rare germline *ATM* variants into diagnostics and clinical decision-making.

By analysing a panel of genes associated with CLL in our diagnostic NGS testing, we detected a number of variants in the *ATM* gene, classified as VUS by clinical databases in patients with CLL. Interestingly, the occurrence of VUS in *ATM* was significantly higher than in other investigated genes, where VUS were detected very rarely (<1%). To gain deeper insight into these variants, we analysed matched germline samples and revealed that the majority of VUS are rare germline variants occurring in less than 0.5% in any population, based on databases. These variants occurred in ~7% of patients. Of them, two-thirds were annotated as pathogenic (1.5%) or VUS with a pathogenic prediction (3.0%) and one-third with a benign prediction (2.4%). Together, germline *ATM* pathogenic/predicted pathogenic variants were detected in 4.5% of CLL patients, all as heterozygous in the germline.

Compared with previous studies on CLL, the inconsistency in the reported prevalence of rare germline variants in the *ATM* gene points to the challenges in the interpretation of these variants. One study reported only pathogenic variants known to cause ataxia telangiectasia occurring in 2.2% of CLL patients<sup>12</sup>, while another study included all protein-altering variants with both pathogenic and benign predictions found in 26.3% of CLL patients, but also occurring in 16.6% of healthy controls<sup>1</sup>. To confirm our findings, we reanalysed the CLL public dataset on 445 paired tumour-normal samples investigated previously for recurrent somatic mutations<sup>19</sup>. We detected the same prevalence (3.4%) of rare cases of germline *ATM* pathogenic/predicted pathogenic variants as in our cohort. Our data confirmed that rare germline *ATM* pathogenic/predicted pathogenic variants predispose to CLL (odds ratio 9.1), as others have already shown<sup>1,12</sup>. Moreover, these *ATM* variants were associated with a family history of cancer (88 vs 43%) but not with age at diagnosis.

Generally, VUS are difficult to translate into the clinical meaning. Looking at the *ATM* gene, ~60% of submitted variants in genetic databases are classified as VUS (Table S4), of them, ~80% are missense. A similar proportion of VUS was also found in other cancer-related genes (*TP53*, *BRCA1*, *BRCA2* and *CHEK2*) in databases (Table S4). The high proportion of VUS in *ATM* and other genes may be explained by the rarity of the individual variants, lack of causality with particular cancer types, challenging interpretations and primarily a lack of functional data. To prove the pathogenicity of VUS-predicted pathogenic variants detected in our CLL cohort, we performed a function study in *ex vivo* primary cells. Indeed, *ATM* pathogenic and VUS-predicted pathogenic variants decreased the ATM kinase activity slightly (~10–20%, depending on the variant type, however, additional hit by del(11q) led to biallelic inactivation with complete loss of ATM activity in all CLL samples. Our data points to the importance of functionally characterizing the VUS-predicted pathogenic variants, and in diagnostics, to refer to them at least as VUS-predicted pathogenic variants until their function is further clarified.

Our data revealed that 33% of CLL patients with rare pathogenic/predicted pathogenic *ATM* variants acquire del(11q), as shown in both investigated CLL cohorts. In these patients, del(11q) is three times more likely to occur than in those with wild type *ATM*. Similarly, 83% of patients with MCL and 56% with metastatic breast cancer with rare germline *ATM* pathogenic/predicted pathogenic variants had concurrent del(11q). Importantly, patients with *ATM* pathogenic/predicted pathogenic allele and del(11q) always lost the wild type allele. From this perspective, we suggest that rare germline *ATM* pathogenic/predicted pathogenic variants behave like somatic *ATM* variants, which are strongly associated with del(11q), thus accelerating the leukaemia progression<sup>23</sup>.

Our data further highlight the need for the implementation of rare germline *ATM* variants into clinical decision-making, not only in CLL. Nevertheless, germline variants with

concurrent del(11q) may look like somatic mutations due to a shift in VAF from the expected 40–60% for a heterozygous allele to 60–100%, or it might be overseen if tumour-normal filtering is used in diagnostics<sup>7</sup>. Moreover, our study demonstrated that rare germline *ATM* pathogenic/predicted pathogenic variants have the same clinical impact as somatic *ATM* and/or *TP53* disruption, resulting in reduced PFS in treated patients, even on novel agents.

Moreover, all CLL patients from both investigated cohorts carrying truncating or missense variants in the *ATM* kinase domain developed CLL with unmutated *IGHV* status, a strongly unfavourable factor in CLL<sup>24</sup>. Despite the small number of patients, the association of these variants with unmutated *IGHV* status is supported by a recent study in mice, where the loss of *ATM* leads to a decreased rate of somatic hypermutation<sup>25</sup>. This study also demonstrated that *ATM* influences germinal center integrity in secondary lymphoid organs, where somatic hypermutation occurs in developing lymphocytes. As unmutated *IGHV* CLL has a more aggressive course associated with shorter survival<sup>24</sup>, therefore the observed association of *ATM* pathogenic/predicted variants with unmutated *IGHV* deserves further investigation.

Besides CLL, there is increasing evidence that rare heterozygous germline *ATM* pathogenic variants increase the risk of other cancers<sup>3,9,10,11</sup>, while rare homozygous germline *ATM* pathogenic variants cause autosomal recessive disorder ataxia telangiectasia. This is consistent with observations in mouse models, where one mutated *ATM* allele had a heightened susceptibility to cancer<sup>26</sup> and caused more genomic instability than the complete loss of the *ATM*, leading to an ataxia-like phenotype<sup>26,27</sup>. Among all cancers, the highest prevalence of germline *ATM* pathogenic variants was reported in pancreatic cancer (3%), but this study did not include *ATM* variants annotated as VUS<sup>11</sup>. Given the inconsistency of interpretations of rare germline *ATM* variants in CLL, we reanalysed the public dataset of i) metastatic breast cancer and ii) MCL, where somatic *ATM* mutations are most frequent among cancers<sup>28</sup>, and germline variants have not been studied yet. In both datasets, we detected rare germline *ATM* pathogenic

variants (breast cancer/MCL: 2%/4%) and the same proportion (2%/4%) of germline variants annotated as VUS but were predicted as pathogenic variants. Overall, the germline pathogenic/predicted pathogenic *ATM* variant in CLL, MCL and breast cancer reach such prevalence as *BRCA1* and *BRCA2* germline mutations in breast cancer (~4% each gene)<sup>29,30</sup>. In contrast to the causal association of *BRCA1* and *BRCA2* with breast cancer, *ATM* variants are not associated with a specific type of cancer but occur in multiple malignancies. Our systemic review demonstrates that the rare *ATM* pathogenic/predicted pathogenic variants are more common in cancer but often missed in diagnostics and highlights the role of the interplay of germline and somatic variation in cancer pathogenesis.

Haematological B cell malignancies showed the highest prevalence of pathogenic germline *ATM* variants in our study. In this context, it is interesting to note that the initial event in MCL, translocation t(11;14), is a result of illegitimate V(D)J recombination<sup>31</sup> and in CLL, decreased rate of somatic hypermutation determines the CLL type with dismal clinical outcome<sup>24</sup>. As *ATM* has already been demonstrated to be involved in V(D)J recombination and somatic hypermutation processes in lymphocytes<sup>25,32,33</sup>, one may suggest that pathogenic germline *ATM* variants may contribute to the onset and adverse phenotype of these diseases. How pathogenic germline *ATM* variants interfere with processes ongoing in lymph nodes deserves further investigation.

In conclusion, this study reveals that half of the rare germline *ATM* variants classified as VUS are pathogenic and behave in the same manner as *ATM* somatic mutations, at least in CLL. As shown by our own and public datasets, the prevalence of rare pathogenic/predicted pathogenic *ATM* variants is underestimated across cancers. This study highlights the importance of implementing rare *ATM* germline variants with pathogenic prediction in clinical diagnostics and decision making, not only in CLL.

**Acknowledgements:** The authors thank Dr Helena Urbánková from the Olomouc University Hospital for cytogenetic analysis and Dr Zuzana Kubová from the Olomouc University Hospital for help with clinical data collection. This study was supported by Internal grant agency of Palacky University (IGA\_LF\_2021\_015, IGA\_LF\_2021\_001), and in part by Ministry of Health of Czech Republic (MH CZ – DRO FNOL, 00098892).

This manuscript has an **Online-only Supplementary Material**.

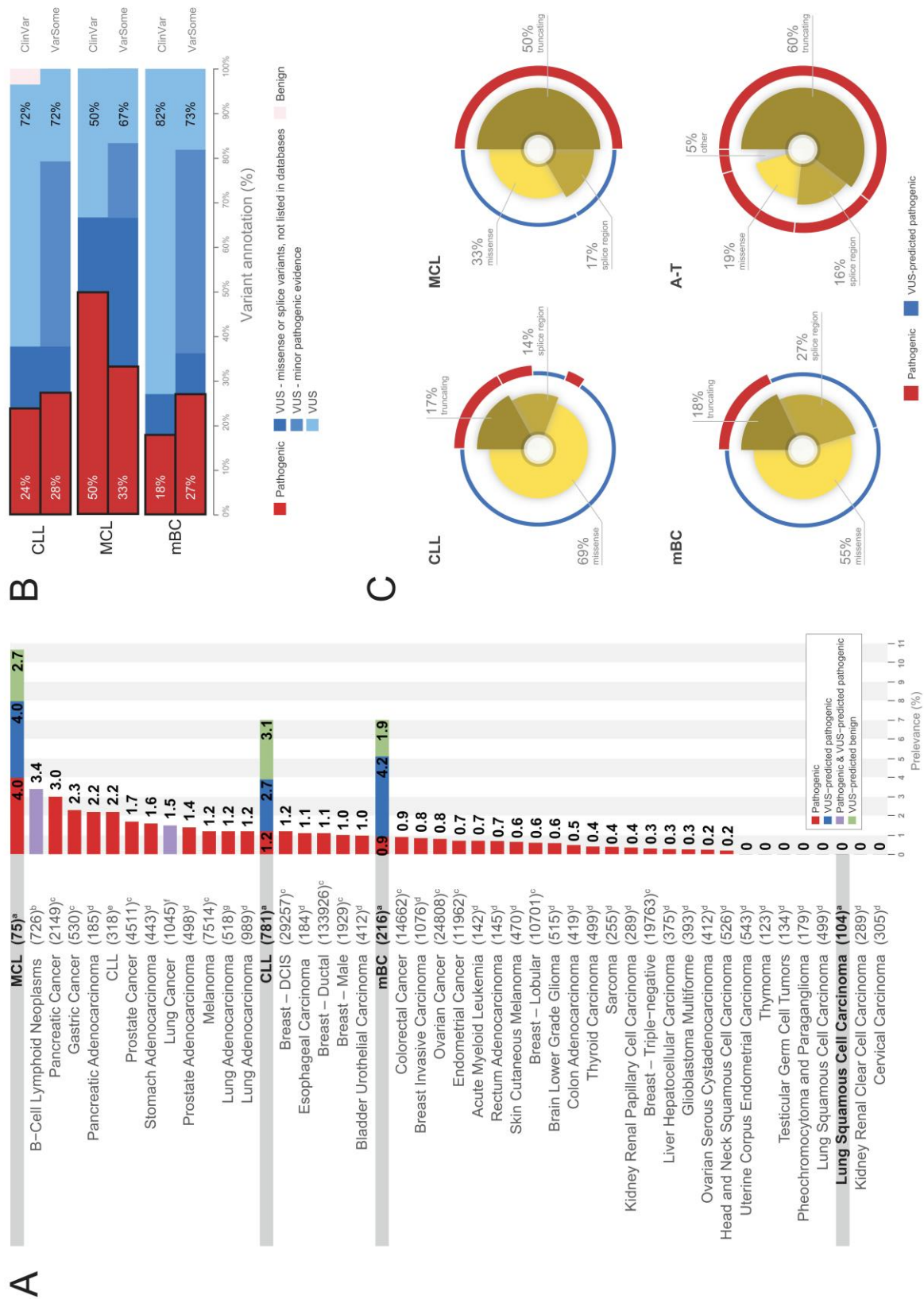
**Conflict of interest:** The authors declare no competing financial interests.



**Table 1** Patient's characteristics of enrolled 336 chronic lymphocytic leukaemia (CLL) cohort A patients.

Characteristics	Patients with rare germinal <i>ATM</i> pathogenic/predicted pathogenic variants (n=15)	Patients with wild type <i>ATM</i> (n=321)	p
Male/Female	8/7	180/141	0.819
Median age, years (range)	68 (59–78)	68 (27–86)	0.803
Median age at diagnosis, years (range)	61 (56–78)	64 (27–86)	0.814
Binet stage: A/B/C	9/4/2	189/89/43	>0.620
Bulky lymphadenopathy $\geq 5$ cm: % (yes/no/NA)	46.2% (6/7/2)	28.8% (77/222/22)	0.112
Splenomegaly: % (yes/no/NA)	61.5% (8/5/2)	32.8% (98/201/22)	<b>0.031</b>
Treatment: yes/no	10/5	192/129	0.588
Novel agents	6	110	0.852
ibrutinib, zanubrutinib, acalabrutinib	5	78	0.527
idelalisib+rituximab	0	22	0.226
venetoclax, venetoclax+rituximab	1	10	0.515
treatment-naïve/Relapsed or refractory	0/6	9/101	0.473
Chemoimmunotherapy	4	82	0.852
Genetics			
unmutated <i>IGHV</i> : % (yes/no/NA)	93.3% (14/1/0)	62.1% (195/119/7)	<b>0.014</b>
del(11q): % (yes/no/NA)	40.0% (6/9/0)	23.1% (74/246/1)	0.134
del(17p): % (yes/no)	13.3% (2/13)	14.3% (46/275)	0.914
del(13q): % (yes/no)	60.0% (9/6)	55.8% (179/142)	0.749
Trisomy 12: % (yes/no)	6.7% (1/14)	9.4% (30/291)	0.791
Complex karyotype: % (yes/no/NA)	20.0% (3/12/0)	11.2% (36/282/3)	0.285
<i>TP53</i> mutated: % (yes/no)	20.0% (3/12)	21.8% (70/251)	0.855
<i>ATM</i> mutated: % (yes/no)	6.7% (1/14)	13.4% (43/278)	0.453
<i>NOTCH1</i> mutated: % (yes/no)	13.3% (2/13)	18.1% (58/263)	0.636
<i>SF3B1</i> mutated: % (yes/no)	26.7% (4/11)	17.1% (55/266)	0.340
<i>BIRC3</i> mutated: % (yes/no)	0% (0/15)	5.0% (16/305)	0.376

**Figures:**

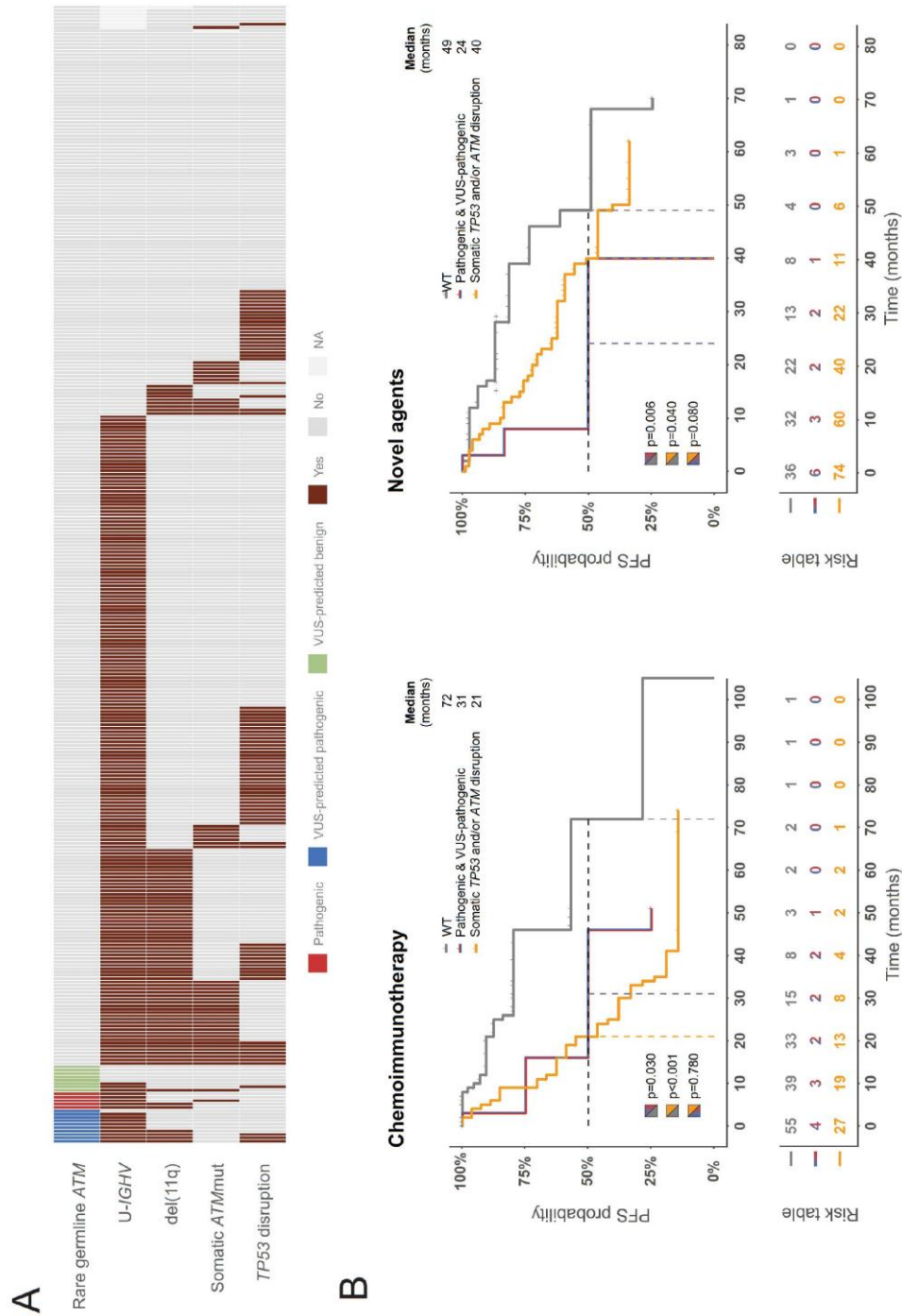


**Figure 1** The prevalence and types of *ATM* rare germline variants in multiple cancers based on our data and published studies.

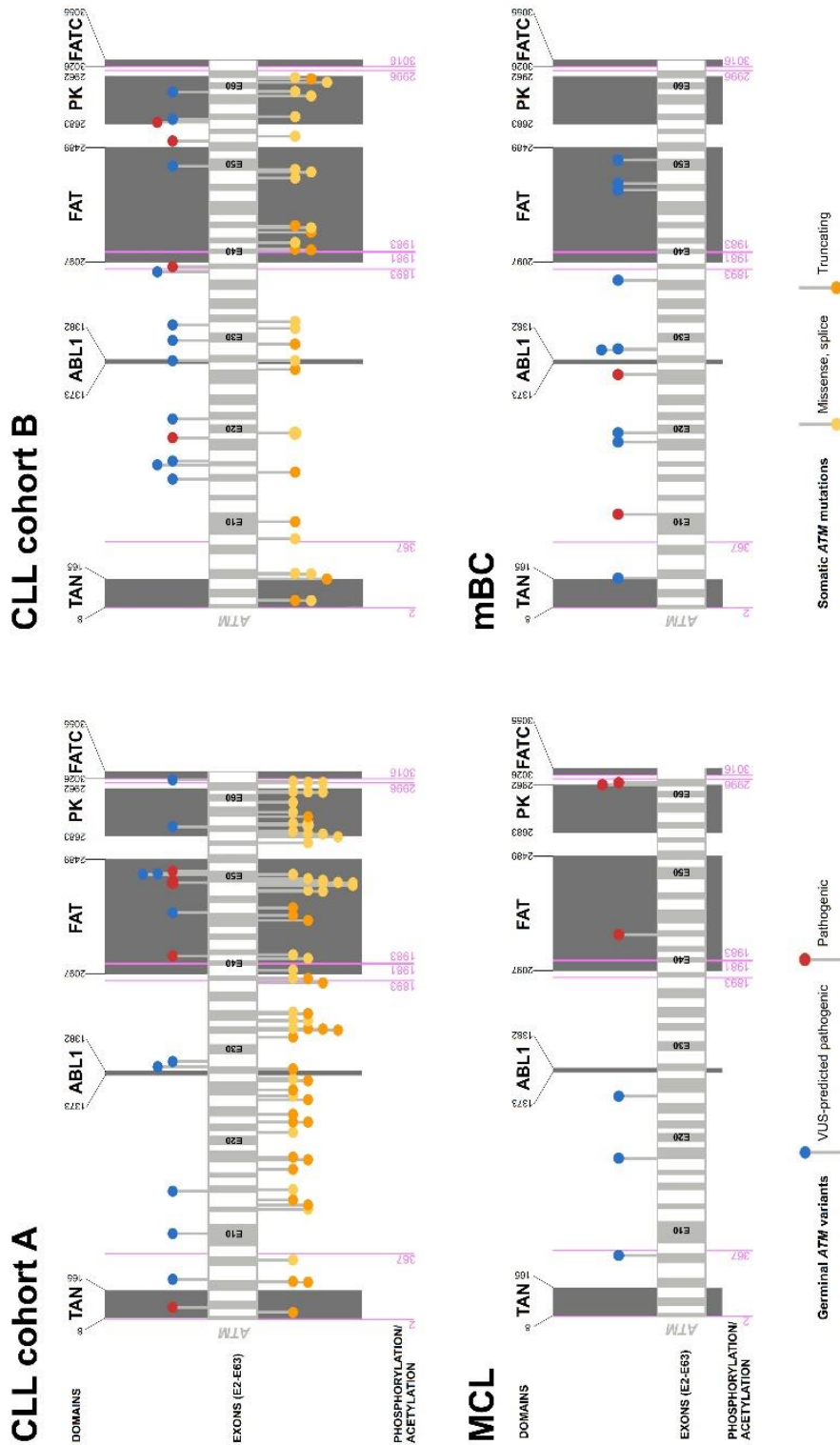
A) The prevalence of *ATM* rare germline variants across cancers. Patients with rare *ATM* germline pathogenic variants are red<sup>#</sup>. Patients with *ATM* germline variants of uncertain significance (VUS)-predicted pathogenic variants are blue. Patients with either pathogenic or VUS-predicted pathogenic *ATM* rare variants are violet, as some studies combine these variants. Patients with *ATM* germline VUS-predicted benign variants are green. Reviewed cohorts and studies: (a) reanalysed public datasets, chronic lymphocytic leukaemia (CLL) cohort A and cohort B are merged, (b)<sup>34</sup>, (c)<sup>11</sup>, (d)<sup>35</sup>, (e)<sup>12</sup>, (f)<sup>36</sup>, (g)<sup>3</sup>. <sup>#</sup>Four studies (c, d, e, g) excluded patients with VUS from analysis.

B) Annotation of rare germline *ATM* pathogenic/predicted pathogenic variants detected in CLL, metastatic breast cancer (mBC) and mantle cell lymphoma (MCL) by ClinVar and VarSome databases.

C) Spectrum of *ATM* rare germline pathogenic/predicted pathogenic variant types detected in CLL, MCL, mBC and ataxia telangiectasia (A-T). In CLL, cohort A and cohort B are merged.



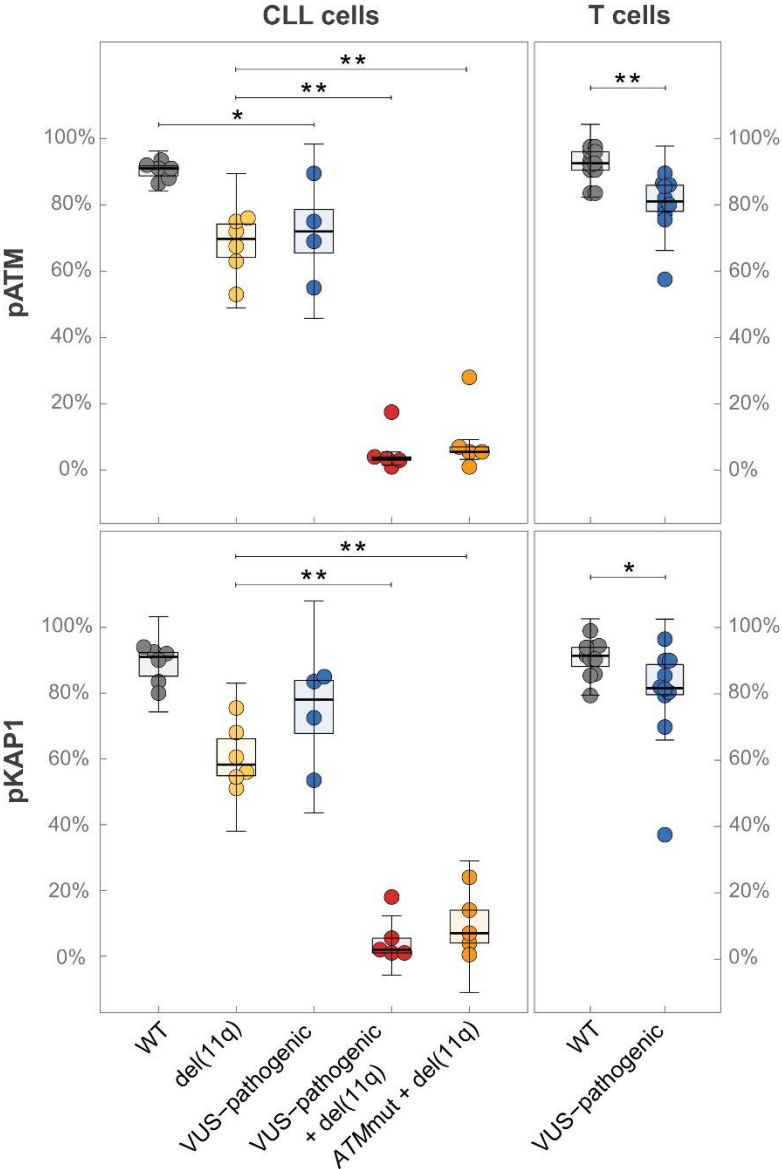
**Figure 2** A) Frequency and distribution of rare germline *ATM* pathogenic/predicted pathogenic variants with adverse somatic aberrations in our chronic lymphocytic leukaemia (CLL) patient cohort A. Each column represents one patient. Rare germline *ATM* variants are coloured according to their annotation in the databases: pathogenic (red), variants of uncertain significance (VUS)-predicted pathogenic (blue) and VUS-predicted benign (green). B) Progression-free survival of patients with chronic lymphocytic leukaemia (CLL) treated with novel agents and chemoimmunotherapy stratified by the presence of rare germline *ATM* pathogenic/predicted pathogenic variants and somatic *ATM* and/or *TP53* disruption.



**Figure 3** Distribution of rare germline *ATM* pathogenic/predicted pathogenic and somatic variants in our chronic lymphocytic leukaemia (CLL) patient cohort A, cohort B, metastatic breast cancer (mBC) and mantle cell lymphoma (MCL). In mBC and MCL, only rare germline *ATM* variants are shown.

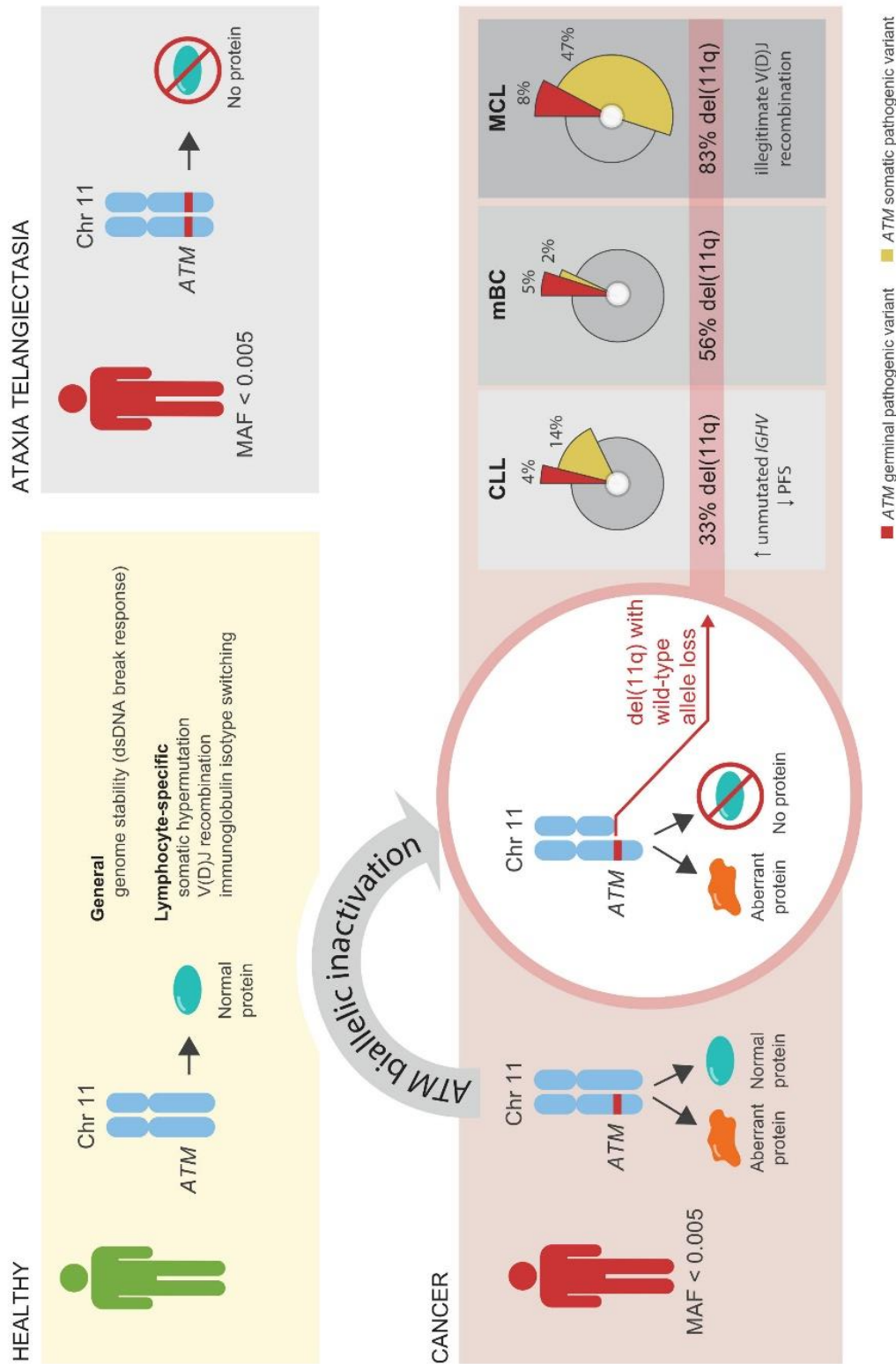
*ATM* variants are coloured as follows: rare germline pathogenic (red), rare germline VUS-predicted pathogenic (blue), somatic truncating (orange) and somatic missense or splice

variants (yellow). The amino acid positions of domains, phosphorylation and acetylation sites in *ATM* are according to the NCBI Protein and UniProt databases. Lollipop size is exponentially proportional to the number of times the variant has been observed in individual patients. Figure was created using the Lollipops (v.1.5.2) tool.



**Figure 4** ATM activity in CLL and T cells obtained from chronic lymphocytic leukaemia (CLL) patients with rare germline *ATM* pathogenic/predicted pathogenic variants, *ATM* wild type (WT) and somatic *ATM* disruption.

In a few patients with a high percentage of CLL cells in the peripheral blood, T cell analysis was not performed due to insufficient cell count.



**Figure 5** Involvement of *ATM* rare germline pathogenic/predicted pathogenic variants in cancer.

## References:

- [1] Tiao G, Improgo MR, Kasar S, et al. Rare germline variants in ATM are associated with chronic lymphocytic leukemia. *Leukemia* 2017; 31(10): 2244-2247.
- [2] Ji X, Mukherjee S, Landi MT, et al. Protein-altering germline mutations implicate novel genes related to lung cancer development. *Nat Commun* 2020; 11: 2220(2020).
- [3] Esai Selvan M, Zauderer MG, Rudin CM, et al. Inherited Rare, Deleterious Variants in ATM Increase Lung Adenocarcinoma Risk. *J Thorac Oncol* 2020; 15(12): 1871-1879.
- [4] Pomerantz MM, Spisák S, Jia L, et al. The association between germline BRCA2 variants and sensitivity to platinum-based chemotherapy among men with metastatic prostate cancer. *Cancer* 2017; 123(18): 3532-3539.
- [5] Singh M, Bhatia P, Khera S, Trehan A. Emerging role of NUDT15 polymorphisms in 6-mercaptopurine metabolism and dose related toxicity in acute lymphoblastic leukaemia. *Leuk Res* 2017; 62: 17-22.
- [6] Crona DJ, Skol AD, Leppanen VM, et al. Genetic variants of VEGFA and FLT4 are determinants of survival in renal cell carcinoma patients treated with sorafenib. *Cancer Res* 2019; 79(1): 231-241.
- [7] Koboldt DC. Best practices for variant calling in clinical sequencing. *Genome Med* 2020; 12(1): 91.
- [8] Laake K, Jansen L, Hahnemann JM, et al. Characterization of ATM mutations in 41 Nordic families with ataxia telangiectasia. *Hum Mutat* 2000; 16(3): 232-246.
- [9] Paulo P, Maia S, Pinto C, et al. Targeted next generation sequencing identifies functionally deleterious germline mutations in novel genes in early-onset/familial prostate cancer. *PLoS Genet* 2018; 14(4): e1007355.
- [10] Zhao ZL, Xia L, Zhao C, Yao J. ATM rs189037 (G > A) polymorphism increased the risk of cancer: an updated meta-analysis. *BMC Med Genet* 2019; 20(1): 28.
- [11] Hall MJ, Bernhisel R, Hughes E, et al. Germline Pathogenic Variants in the Ataxia Telangiectasia Mutated (ATM) Gene are Associated with High and Moderate Risks for Multiple Cancers. *Cancer Prev Res (Phila)* 2021; 14(4): 433-440.
- [12] Skowronska A, Austen B, Powell JE, et al. ATM germline heterozygosity does not play a role in chronic lymphocytic leukemia initiation but influences rapid disease progression through loss of the remaining ATM allele. *Haematologica* 2012; 97(1): 142-6.
- [13] Rose-Zerilli MJ, Forster J, Parker H, et al. ATM mutation rather than BIRC3 deletion and/or mutation predicts reduced survival in 11q-deleted chronic lymphocytic leukemia: data from the UK LRF CLL4 trial. *Haematologica* 2014; 99(4): 736-742.
- [14] Hallek M, Cheson BD, Catovsky D, et al. International Workshop on Chronic Lymphocytic Leukemia. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* 2008; 111(12): 5446-5456.
- [15] Petrackova A, Vasinek M, Sedlarikova L, et al. Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. *Front Oncol* 2019; 9: 851.



- [16] Petrackova A, Minarik J, Sedlarikova L, et al. Diagnostic deep-targeted next-generation sequencing assessment of TP53 gene mutations in multiple myeloma from the whole bone marrow. *Br J Haematol* 2020; 189(4): e122-e125.
- [17] Malarikova D, Berkova A, Obr A, et al. Concurrent TP53 and CDKN2A Gene Aberrations in Newly Diagnosed Mantle Cell Lymphoma Correlate with Chemoresistance and Call for Innovative Upfront Therapy. *Cancers (Basel)* 2020; 12(8): 2120.
- [18] Genome Aggregation Database (gnomAD) [(accessed on 20 August 2021)]; Available online: <http://gnomad.broadinstitute.org/>
- [19] Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* 2015; 526(7574): 519-524.
- [20] Pararajalingam P, Coyle KM, Arthur SE, et al. Coding and noncoding drivers of mantle cell lymphoma identified through exome and genome sequencing. *Blood* 2020; 136(5): 572-584.
- [21] Lefebvre C, Bachelot T, Filleron T, et al. Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med* 2016; 13(12): e1002201.
- [22] Kruzova L, Schneiderova P, Holzerova M, et al. Complex karyotype as a predictor of high-risk chronic lymphocytic leukemia: A single center experience over 12 years. *Leuk Res* 2019; 85: 106218.
- [23] Skowronska A, Parker A, Ahmed G, et al. Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom Leukemia Research Fund Chronic Lymphocytic Leukemia 4 trial. *J Clin Oncol* 2012; 30(36): 4524-4532.
- [24] Yun X, Zhang Y, Wang, X. Recent progress of prognostic biomarkers and risk scoring systems in chronic lymphocytic leukemia. *Biomark Res* 2020; 8: 40.
- [25] Nicolas L, Cols M, Smolkin R, et al. Cutting Edge: ATM Influences Germinal Center Integrity. *J Immunol* 2019; 202(11): 3137-3142.
- [26] Spring K, Ahangari F, Scott SP, et al. Mice heterozygous for mutation in *Atm*, the gene involved in ataxia-telangiectasia, have heightened susceptibility to cancer. *Nat Genet* 2002; 32(1): 185-190.
- [27] Yamamoto K, Wang J, Sprinzen L, et al. Kinase-dead ATM protein is highly oncogenic and can be preferentially targeted by Topo-isomerase I inhibitors. *Elife* 2016; 5: e14709.
- [28] Choi M, Kipps T, Kurzrock R. ATM Mutations in Cancer: Therapeutic Implications. *Mol Cancer Ther* 2016; 15(8): 1781-1791.
- [29] Antoniou A, Pharoah PD, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 2003; 72(5): 1117-1130.
- [30] Foulkes WD. Inherited susceptibility to common cancers. *N Engl J Med* 2008; 359(20): 2143-2153.
- [31] Jares P, Colomer D, Campo E. Molecular pathogenesis of mantle cell lymphoma. *J Clin Invest* 2012; 122(10): 3416-23.
- [32] Pan-Hammarström Q, Lähdesmäki A, Zhao Y, et al. Disparate roles of ATR and ATM in immunoglobulin class switch recombination and somatic hypermutation. *J Exp Med* 2006; 203(1): 99-110.
- [33] Yin B, Savic V, Bassing CH. ATM prevents unattended DNA double strand breaks on site and in generations to come. *Cancer Biol Ther* 2007; 6(12): 1837-1839.

- [34] Mosquera Orgueira A, Cid López M, Peleteiro Raíndo A, et al. Detection of Rare Germline Variants in the Genomes of Patients with B-Cell Neoplasms. *Cancers (Basel)* 2021; 13(6): 1340.
- [35] Huang KL, Mashl RJ, Wu Y, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell* 2018; 173(2): 355-370.e14.
- [36] Liu Y, Xia J, McKay J, et al. Rare deleterious germline variants and risk of lung cancer. *NPJ Precis Oncol* 2021; 5(1): 12.

## APPENDIX B

**Petrackova A**, Minarik J, Sedlarikova L, Libigerova T, Hamplova A, Krhovska P, Balcarkova J, Pika T, Papajik T, Kriegova E. Diagnostic deep-targeted next-generation sequencing assessment of *TP53* gene mutations in multiple myeloma from the whole bone marrow. *Br J Haematol.* 2020;189(4):e122-e125. (IF 2020: 6.998; Q1)

## References

- Antony-Debré, I., Bluteau, D., Itzykson, R., Baccini, V., Renneville, A., Boehlen, F., Morabito, M., Droin, N., Deswarte, C., Chang, Y., Leverger, G., Solary, E., Vainchenker, W., Favier, R. & Raslova, H. (2012) MYH10 protein expression in platelets as a biomarker of RUNX1 and FLI1 alterations. *Blood*, **120**, 2719–2722.
- Antony-Debré, I., Duployez, N., Bucci, M., Geffroy, S., Micol, J.-B., Renneville, A., Boissel, N., Dhédin, N., Réa, D., Nelken, B., Berthon, C., Leblanc, T., Mozziconacci, M.-J., Favier, R., Heller, P.G., Abdel-Wahab, O., Raslova, H., Latger-Cannard, V. & Preudhomme, C. (2016) Somatic mutations associated with leukemic progression of familial platelet disorder with predisposition to acute myeloid leukemia. *Leukemia*, **30**, 999–1002.
- Bluteau, D., Glembofsky, A.C., Raimbault, A., Balayn, N., Gilles, L., Rameau, P., Nurden, P., Alessi, M.C., Debili, N., Vainchenker, W., Heller, P.G., Favier, R. & Raslova, H. (2012) Dysmegakaryopoiesis of FPD/AML pedigrees with constitutional RUNX1 mutations is linked to myosin II deregulated expression. *Blood*, **120**, 2708–2718.
- Bluteau, D., Balduini, A., Balayn, N., Currao, M., Nurden, P., Deswarte, C., Leverger, G., Noris, P., Perrotta, S., Solary, E., Vainchenker, W., Debili, N., Favier, R. & Raslova, H. (2014) Thrombocytopenia-associated mutations in the ANKRD26 regulatory region induce MAPK hyperactivation. *The Journal of Clinical Investigation*, **124**, 580–591.
- Chisholm, K.M., Denton, C., Keel, S., Geddis, A.E., Xu, M., Appel, B.E., Cantor, A.B., Fleming, M.D. & Shimamura, A. (2019) Bone marrow morphology associated with germline RUNX1 mutations in patients with familial platelet disorder with associated myeloid malignancy. *Pediatric and Developmental Pathology*, **22**, 315–328.
- Duployez, N., Lejeune, S., Renneville, A. & Preudhomme, C. (2016) Myelodysplastic syndromes and acute leukemia with genetic predispositions: a new challenge for hematologists. *Expert Review of Hematology*, **9**, 1189–1202.
- Duployez, N., Abou Chahla, W., Lejeune, S., Marceau-Renaut, A., Letizia, G., Boyer, T., Geffroy, S., Peyrouze, P., Gardel, N., Nelken, B., Michel, G., Bertrand, Y. & Preudhomme, C. (2018) Detection of a new heterozygous germline ETV6 mutation in a case with hyperdiploid acute lymphoblastic leukemia. *European Journal of Haematology*, **100**, 104–107.
- Hock, H., Meade, E., Medeiros, S., Schindler, J.W., Valk, P.J.M., Fujiwara, Y. & Orkin, S.H. (2004) Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *Genes & Development*, **18**, 2336–2341.
- Kanagal-Shamanna, R., Loghavi, S., DiNardo, C.D., Medeiros, L.J., Garcia-Manero, G., Jabbour, E., Routbort, M.J., Luthra, R., Bueso-Ramos, C.E. & Khoury, J.D. (2017) Bone marrow pathologic abnormalities in familial platelet disorder with propensity for myeloid malignancy and germline RUNX1 mutation. *Haematologica*, **102**, 1661–1670.
- Latger-Cannard, V., Philippe, C., Jonveaux, P., Lecompte, T. & Favier, R. (2011) Dysmegakaryopoiesis, a clue for an early diagnosis of familial platelet disorder with propensity to acute myeloid leukemia in case of unexplained inherited thrombocytopenia associated with normal-sized platelets. *Journal of Pediatric Hematology/Oncology*, **33**, e264–e266.
- Noris, P., Perrotta, S., Seri, M., Pecci, A., Gnan, C., Loffredo, G., Pujol-Moix, N., Zecca, M., Scognamiglio, F., Rocco, D.D., Punzo, F., Melazzini, F., Scianguetta, S., Casale, M., Marconi, C., Pippucci, T., Amendola, G., Notarangelo, L.D., Klersy, C., Civaschi, E., Balduini, C.L. & Savoia, A. (2011) Mutations in ANKRD26 are responsible for a frequent form of inherited thrombocytopenia: analysis of 78 patients from 21 families. *Blood*, **117**, 6673–6680.
- Perez Botero, J., Oliveira, J.L., Chen, D., Reichard, K.K., Viswanatha, D.S., Nguyen, P.L., Pruthi, R.K., Majerus, J., Gada, P., Gangat, N., Tefferi, A. & Patnaik, M.M. (2015) ASXL1 mutated chronic myelomonocytic leukemia in a patient with familial thrombocytopenia secondary to germline mutation in ANKRD26. *Blood Cancer Journal*, **5**, e315.
- Poggi, M., Canault, M., Favier, M., Turro, E., Saultier, P., Ghalloussi, D., Baccini, V., Vidal, L., Mezzapesa, A., Chelghoum, N., Mohand-Oumoussa, B., Falaise, C., Favier, R., Ouwehand, W.H., Fiore, M., Peiretti, F., Morange, P.E., Saut, N., Bernot, D., Greinacher, A., BioResource, N., Nurden, A.T., Nurden, P., Freson, K., Trégouët, D.-A., Raslova, H. & Alessi, M.C. (2017) Germline variants in ETV6 underlie reduced platelet formation, platelet dysfunction and increased levels of circulating CD34<sup>+</sup> progenitors. *Haematologica*, **102**, 282–294.
- Steenma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P. & Ebert, B.L. (2015) Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*, **126**, 9–16.
- Zhang, M.Y., Churpek, J.E., Keel, S.B., Walsh, T., Lee, M.K., Loeb, K.R., Gulsuner, S., Pritchard, C.C., Sanchez-Bonilla, M., Delrow, J.J., Basom, R.S., Forouhar, M., Gyurkocza, B., Schwartz, B.S., Neistadt, B., Marquez, R., Mariani, C.J., Coats, S.A., Hofmann, I., Lindsley, R.C., Williams, D.A., Abkowitz, J.L., Horwitz, M.S., King, M.C., Godley, L.A. & Shimamura, A. (2015) Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nature Genetics*, **47**, 180–185.

## Diagnostic deep-targeted next-generation sequencing assessment of *TP53* gene mutations in multiple myeloma from the whole bone marrow

Among the important markers of poor prognosis in multiple myeloma (MM) are aberrations in *TP53* gene (Manier *et al.*, 2017), caused by the deletion and/or mutation in *TP53* gene (*TP53mut*). The detection of del(17p) is a part of the recommended risk assessment in newly diagnosed MM (Kumar *et al.*, 2017). However, *TP53* mutation analysis in MM is not

widely performed in diagnostics, particularly due to technical limitations regarding sample collection and plasma cell (PC) enrichment. The pitfalls are mainly connected with inter-individual variability in the sample amount, PC infiltration in bone marrow (BM), PC immunophenotypes and time-dependent losses of surface markers as well as haemodilution,

patchy or site varied PC distribution, aspirate pull order and aggregation of PC in aspirated BM (Mansilla *et al.*, 2018; Romano *et al.*, 2019), all together resulting in low PC recovery in some patients.

Therefore, there is a need to standardise the pre-processing of BM in clinical settings, as well as to have an alternative, when enriched tumour PCs are not available. In the present study, we explored the applicability of the whole BM for analysis of *TP53* mutations, together with *NRAS*, *KRAS* and *BRAF* hotspots, by deep-targeted next-generation sequencing (NGS) in diagnostics.

## Patients and Methods

Diagnostic BM samples were obtained from 54 patients with MM (Table I), diagnosed according to the International Myeloma Working Group criteria (2014). All patients provided written informed consent about the usage of BM for the purpose of this study, which was performed in accordance with the Helsinki Declaration and approved by the Ethics Committee of the University Hospital and Palacký University Olomouc.

The sample collection and PC enrichment are described in Data S1. NGS mutation assessment of the *TP53* full coding sequence and *NRAS*, *KRAS*, and *BRAF* hotspots was performed with a minimum target read depth of 5000×, as reported previously (Obr *et al.*, 2018; Petrackova *et al.*, 2019). The detection limit was 1%, the variants within the range 1–3% were confirmed by replication. Cytogenetic abnormalities were assessed by fluorescence *in situ* hybridisation (FISH) with immunophenotyping, as reported previously (Mlynarcikova *et al.*, 2016). Details are described in Data S1.

## Results and discussion

In our present cohort, only in three-quarters of diagnostic BM samples was it possible to enrich with at least 100 000 cells needed to perform deep-targeted NGS. Of the 54 BM samples [median (range) 4.5 (1.0–12.0) ml] with PC infiltration (10.5%, range 2.0–67.0%), 72.9% samples were successfully enriched (PC purity: 90.9%, range 39.0–99.1%; PC absolute counts:  $1.0 \times 10^6$ , range  $0.1\text{--}35.30 \times 10^6$ ; yield of PC: 32.2%, range 7.0–100%). Lower PC recovery was observed using Ficoll gradient separation (24.1%, range 7.3–46.5%) compared to red-blood-cell lysis (43.0%, range 6.7–95.2%). Enrichment in 27.1% samples failed due to low recovery and/or sample amount.

To evaluate the utility of whole BM for mutation NGS analysis, we investigated the whole BM and matched enriched PCs ( $n = 27$ ). There was concordance between *TP53*, *NRAS*, *KRAS*, and *BRAF* mutations in 85.2% patients (Fig 1A). Four samples, in which mutations were found only in enriched samples, had PC infiltration in BM of 4.0%, 7.0%, 20.8%, and 26.5%. The variant allele frequency (VAF) of detected mutations from the whole BM corresponded or

were lower than the PC infiltration in BM (Fig 1B). Overall, the enrichment of PCs increased the mutation VAF on average 23.3% (range 3.0–53.0%), which is not proportional to the increase of the tumour PCs fraction in the sample achieved by PC enrichment (on average 70.7%) (Fig 1C). This phenomenon may be explained by: i) high clonal heterogeneity of malignant myeloma PCs with variable immunophenotypes, and ii) presence of non-malignant PCs or various B cell precursors in the enriched sample that are also positive for Syndecan-1 (CD138), although this marker might be of a lower surface density (Coffey *et al.*, 2019), which was excluded in our enriched samples (<1% as assessed by immunophenotyping using CD19/CD38/CD45/CD56/CD138).

In our present cohort, 28% (15/54) of patients had *TP53* disruption and of them three patients at diagnosis [one only *TP53*mut, two patients del(17p)]. Only *TP53*mut were detected in 47% (7/15), *TP53*mut together with del(17p) in 33% (5/15) and only del(17p) in 20% (3/15) of patients with *TP53* disruption. Of 22% (12/54) of patients carrying *TP53*mut, five had  $\geq 2$  (maximum 7) *TP53*mut. Additionally, three patients at diagnosis had chr17 trisomy with FISH confirmation of three *TP53* gene copies, of them one patient also carried *TP53*mut. In newly diagnosed MM, *TP53*mut are rare and are associated with more aggressive disease and

Table I. Characteristics of MM patient cohort.

Demographic and clinical features	MM (n=54)
Male/female, <i>n</i>	22/32
Age, median (range), years	70 (41–89)
PC infiltration in BM determined by flow cytometry, median (range), %	10.5 (2.0–67.0)
Sampling at, <i>n</i> (%):	
Diagnosis	30 (55.6)
Relapse or disease progression	22 (40.7)
MRD or during the disease course without treatment	2 (3.7)
ISS staging, <i>n</i> (%)	
ISS I	19 (35.2)
ISS II	16 (29.6)
ISS III	19 (35.2)
Chromosomal aberrations*, <i>n</i> (%)	
del(17p13)	8 (14.8)
trisomy of chromosome 17 <sup>†</sup>	3 (5.6)
del(13q14)	32 (59.3)
t(4;14)	9 (16.7)
t(11;14)	5 (9.3)
1q21 gain	20 (37.0)
Hyperploidy <sup>‡</sup>	35 (64.8)

MRD, minimal residual disease; ISS, International Staging System.

\*The cut-off level of 10% was used for all regions; at least 100 PCs were counted for each probe.

<sup>†</sup>Three *TP53* copies confirmed by FISH.

<sup>‡</sup>Hyperdiploidy defined as presence  $\geq 1$  trisomy in  $\geq 60\%$  of PCs.

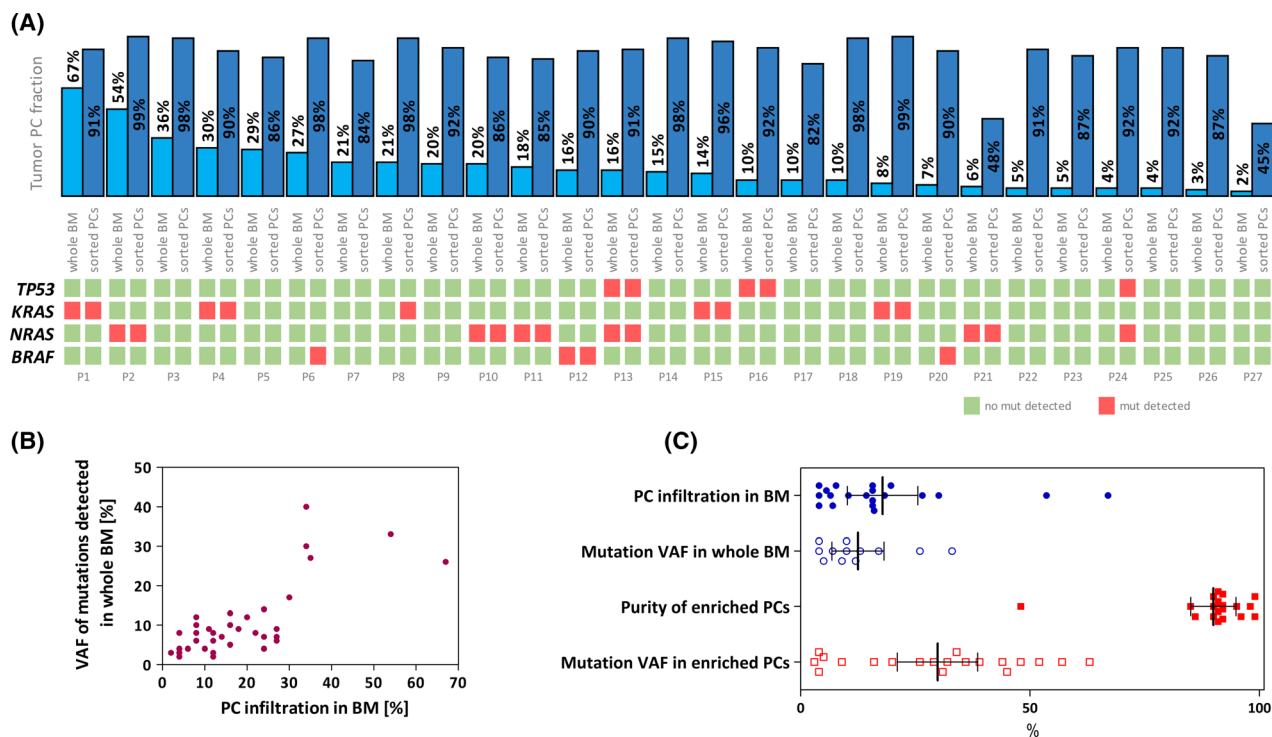


Fig 1. A) Concordance of mutations found in matched whole bone marrow (BM) and paired enriched tumour plasma cell (PC) samples. The heatmap represents detection of individual mutations in a series of paired whole BM and enriched tumour PC samples. Columns represent infiltration of tumour PCs in whole BM and purity of enriched PCs. B) Relationship between variant allele frequency (VAF) of detected mutations from the whole BM and the PC infiltration in the BM. C) Comparison of VAF of mutations found in matched whole BM and enriched tumour PC samples and their relationship to the PC fraction in the samples.

treatment resistance (Manier *et al.*, 2017; Walker *et al.*, 2019). During the disease course, the *TP53*mut contributes to the disease progression and the biallelic inactivation of *TP53* has been reported in 21–26% patients at relapse (Ryland *et al.*, 2018). Regarding other genes, *KRAS*mut were detected in 18.5% (10/54) patients, *NRAS*mut in 13.0% (7/54), and *BRAF*mut in 9.3% (5/54). Mutations in *KRAS* and *NRAS* were mutually exclusive; this phenomenon was already reported in myeloma cell lines (Vikova *et al.*, 2019).

The minimum percentage of cancer clonal fraction harbouring a del(17p), as well as *TP53*mut indicative of poor prognosis in MM, is still under investigation. The European Myeloma Network recommends a 20% positive cut-off level for numerical abnormalities (Ross *et al.*, 2012), another study suggests a 55% threshold for prognostic evaluation of del(17p) in newly diagnosed MM patients (Thakurta *et al.*, 2019). However, others demonstrated an independent association between subclonal *TP53* deletions and MM outcome (Lakshman *et al.*, 2019).

In conclusion, our present analysis in a real-world diagnostic cohort demonstrates the utility of the whole BM for *TP53*mut analysis by deep-targeted NGS in MM when enriched PCs are not available, while obtaining diagnostic information comparable to enriched samples. Furthermore, employing novel, highly sensitive sequencing techniques will

help to ensure the required sensitivity for mutation detection from the whole BM (Salk *et al.*, 2018; Petrackova *et al.*, 2019). Moreover, our present data highlight the importance of the assessment of *TP53*mut in patients with MM, as they may occur regardless of del(17p), as well as the need for standardisation of PC enrichment in diagnostics. In the current era of precision medicine, routine screening for *TP53*mut in MM can enhance patient risk stratification.

### Acknowledgements

Grant support: MZ CR VES16-32339A, in part by the MH CZ – DRO (FNOL, 00098892) and IGA LF UP\_2020\_16.

### Conflict of interest

The authors declare no conflict of interest.

### Authorship contributions

Anna Petrackova conceived the study, analysed the data and wrote the manuscript. Petra Krhovska, Tomas Pika and Tomas Papajik collected patients' samples. Jana Balcarkova performed cytogenetic analysis. Lenka Sedlarikova, Tereza Libigerova and Alzbeta Hamplova performed sequencing and

analysed data. Jiri Minarik collected patients' samples, provided clinical data and critically revised the manuscript. Eva Kriegova supervised the study and critically revised the manuscript. All authors had final approval of the submitted and published versions.

Anna Petrackova<sup>1</sup>  
 Jiri Minarik<sup>2</sup>   
 Lenka Sedlarikova<sup>1</sup>  
 Tereza Libigerova<sup>1</sup>  
 Alzbeta Hamplova<sup>1</sup>  
 Petra Krhovska<sup>2</sup>  
 Jana Balcarkova<sup>2</sup>  
 Tomas Pika<sup>2</sup>  
 Tomas Papajik<sup>2</sup>  
 Eva Kriegova<sup>1</sup> 

<sup>1</sup>Department of Immunology, Faculty of Medicine and Dentistry, Palacky University and University Hospital and <sup>2</sup>Department of

Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czech Republic.

E-mail: eva.kriegova@email.cz

**Keywords:** disruption of tumour protein p53 (*TP53*) gene, multiple myeloma, mutation, next-generation sequencing, plasma cell enrichment

First published online 4 March 2020

doi: 10.1111/bjh.16547

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Data S1.** Supporting Information

## References

- Coffey, D.G., Wu, Q.V., Towler, A.M.H., Ornelas, S., Morales, A.J., Xu, Y., Green, D.J. & Warren, E.H. (2019) Ultradeep, targeted sequencing reveals distinct mutations in blood compared to matched bone marrow among patients with multiple myeloma. *Blood Cancer Journal*, **9**, 77.
- Kumar, S.K., Callander, N.S., Alsina, M., Atanackovic, D., Biermann, J., Chandler, J.C., Costello, C., Faiman, M., Fung, H.C., Gasparetto, C., Godby, K., Hofmeister, C., Holmberg, L., Holstein, S., Huff, C., Kassim, A., Liedtke, M., Martin, T., Omel, J., Rajé, N., Reu, F.J., Singhal, S., Somlo, G., Stockerl-Goldstein, K., Treon, S.P., Weber, D., Yahalom, J., Shad, D.A. & Kumar, R. (2017) Multiple myeloma, version 3.2017, NCCN clinical practice guidelines in oncology. *Journal of the National Comprehensive Cancer Network*, **15**, 230–269.
- Lakshman, A., Painuly, U., Rajkumar, S.V., Ketterling, R.P., Kapoor, P., Greipp, P.T., Gertz, M.A., Buadi, F.K., Lacy, M.Q., Dingli, D., Dispenzieri, A., Fonder, A.L., Hayman, S.R., Hobbs, M.A., Gonsalves, W.I., Hwa, Y.L., Leung, N., Go, R.S., Lin, Y., Kourelis, T.V., Warsame, R., Lust, J.A., Russell, S.J., Zeldenrust, S.R., Kyle, R.A. & Kumar, S.K. (2019) Natural history of multiple myeloma with de novo del(17p). *Blood Cancer Journal*, **9**, 32.
- Manier, S., Salem, K.Z., Park, J., Landau, D.A., Getz, G. & Ghobrial, I.M. (2017) Genomic complexity of multiple myeloma and its clinical implications. *Nature Reviews Clinical Oncology*, **14**, 100–113.
- Mansilla, C., Soria, E., Vallejo, M., Valiente, A., Perez-Juana, A., Zabalza, A., Hurtado, G., Sala, F. & Ramirez, N. (2018) Combined selection system to lower the cutoff for plasma cell enrichment applied to iFISH analysis in multiple myeloma. *Translational oncology*, **11**, 647–652.
- Mlynarcikova, M., Balcarkova, J., Mickova, P., Scudla, V., Pika, T., Bacovsky, J., Minarik, J., Janousova, E. & Jarosova, M. (2016) Molecular cytogenetic analysis of chromosome 8 aberrations in patients with multiple myeloma examined in 2 different stages, at diagnosis and at progression/relapse. *Clinical Lymphoma Myeloma and Leukemia*, **16**, 358–365.
- Obr, A., Procházka, V., Jirkuvová, A., Urbánková, H., Kriegova, E., Schneiderová, P., Vatošková, M. & Papajik, T. (2018) TP53 mutation and complex karyotype portends a dismal prognosis in patients with mantle cell lymphoma. *Clinical Lymphoma Myeloma and Leukemia*, **18**, 762–768.
- Petrackova, A., Vasinek, M., Sedlarikova, L., Dyskova, T., Schneiderova, P., Novosad, T., Papajik, T. & Kriegova, E. (2019) Standardization of sequencing coverage depth in NGS: recommendation for detection of clonal and subclonal mutations in cancer diagnostics. *Frontiers in Oncology*, **9**, 851.
- Romano, A., Palumbo, G.A., Parrinello, N.L., Conticello, C., Martello, M. & Terragna, C. (2019) Minimal residual disease assessment within the bone marrow of multiple myeloma: a review of caveats, clinical significance and future perspectives. *Frontiers in oncology*, **9**, 699.
- Ross, F.M., Avet-Loiseau, H., Ameye, G., Gutiérrez, N.C., Liebisch, P., O'Connor, S., Dalva, K., Fabris, S., Testi, A.M., Jarosova, M., Hodgkinson, C., Collin, A., Kerndrup, G., Kuglik, P., Ladon, D., Bernasconi, P., Maes, B., Zemanova, Z., Michalova, K., Michau, L., Neben, K., Hermansen, N.E., Rack, K., Rocci, A., Protheroe, R., Chiecchio, L., Poirel, H.A., Sonneveld, P., Nyegaard, M. & Johnsen, H.E. (2012) Report from the European myeloma network on interphase FISH in multiple myeloma and related disorders. *Haematologica*, **97**, 1272–1277.
- Ryland, G.L., Jones, K., Chin, M., Markham, J., Aydoğan, E., Kankanige, Y., Caruso, M., Guinto, J., Dickinson, M., Prince, H.M., Yong, K. & Blombery, P. (2018) Novel genomic findings in multiple myeloma identified through routine diagnostic sequencing. *Journal of Clinical Pathology*, **71**, 895–899.
- Salk, J.J., Schmitt, M.W. & Loeb, L.A. (2018) Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nature Reviews Genetics*, **19**, 269–285.
- Thakurta, A., Ortiz, M., Bleuca, P., Towfic, F., Corre, J., Serbina, N.V., Flynt, E., Yu, Z., Yang, Z., Palumbo, A., Dimopoulos, M.A., Gutierrez, N.C., Goldschmidt, H., Sonneveld, P. & Avet-Loiseau, H. (2019) High subclonal fraction of 17p deletion is associated with poor prognosis in multiple myeloma. *Blood*, **133**, 1217–1221.
- Vikova, V., Jourdan, M., Robert, N., Requirand, G., Boireau, S., Bruyer, A., Vincent, L., Cartron, G., Klein, B., Elemento, O., Kassambara, A. & Moreaux, J. (2019) Comprehensive characterization of the mutational landscape in multiple myeloma cell lines reveals potential drivers and pathways associated with tumor progression and drug resistance. *Theranostics*, **9**, 540–553.
- Walker, B.A., Mavrommatis, K., Wardell, C.P., Ashby, T.C., Bauer, M., Davies, F.E., Rosenthal, A., Wang, H., Qu, P., Hoering, A., Samur, M., Towfic, F., Ortiz, M., Flynt, E., Yu, Z., Yang, Z., Rozelle, D., Obenaus, J., Trotter, M., Auclair, D., Keats, J., Bolli, N., Fulciniti, M., Szalat, R., Moreau, P., Durie, B., Stewart, A.K., Goldschmidt, H., Raab, M.S., Einsele, H., Sonneveld, P., San Miguel, J., Lonial, S., Jackson, G.H., Anderson, K.C., Avet-Loiseau, H., Munshi, N., Thakurta, A. & Morgan, G.J. (2019) A high-risk, double-hit, group of newly diagnosed myeloma identified by genomic analysis. *Leukemia*, **33**, 159–170.

APPENDIX C

**Petrackova A**, Horak P, Radvansky M, Fillerova R, Smotkova Kraiczova V, Kudelka M, Mrazek F, Skacelova M, Smrzova A, Kriegova E. Revealed heterogeneity in rheumatoid arthritis based on multivariate innate signature analysis. Clin Exp Rheumatol. 2020;38:289-298. (IF 2020: 4.473)

*Dean's award for the best student scientific publications in 2020*



# Revealed heterogeneity in rheumatoid arthritis based on multivariate innate signature analysis

A. Petrackova<sup>1</sup>, P. Horak<sup>2</sup>, M. Radvansky<sup>3</sup>, R. Fillerova<sup>1</sup>, V. Smotkova Kraiczova<sup>1</sup>, M. Kudelka<sup>3</sup>, F. Mrazek<sup>1</sup>, M. Skacelova<sup>2</sup>, A. Smrzova<sup>2</sup>, E. Kriegova<sup>1</sup>

<sup>1</sup>Department of Immunology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Czech Republic; <sup>2</sup>Department of Internal Medicine III - Nephrology, Rheumatology and Endocrinology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Czech Republic; <sup>3</sup>Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava, Czech Republic.

---

## Abstract

### Objective

A growing body of evidence highlights the persistent activation of the innate immune system and type I interferon (IFN) signature in the pathogenesis of rheumatoid arthritis (RA) and its association with disease activity. Since the recent study revealed heterogeneity in the IFN signature in RA, we investigated for the first time the heterogeneity in innate signature in RA.

---

### Methods

The innate gene expression signature (10 TLRs, 7 IL1/IL1R family members, and CXCL8/IL8) was assessed in peripheral blood mononuclear cells from RA patients (n=67), both with active (DAS28 $\geq$ 3.2, n=32) and inactive disease (DAS28<3.2, n=35), and in healthy control subjects (n=55).

---

### Results

Of the 13 deregulated innate genes (TLR2, TLR3, TLR4, TLR5, TLR8, TLR10, IL1B, IL1RN, IL18, IL18R1, IL1RAP, and SIGIRR/IL1R8) associated with RA, TLR10 and IL1RAP are being reported for the first time. Multivariate analysis based on utilising patient similarity networks revealed the existence of four patient's subsets (clusters) based on different TLR8 and IL1RN expression profiles, two in active and two in inactive RA. Moreover, neural network analysis identified two main gene sets describing active RA within an activity-related innate signature (TLR1, TLR2, TLR3, TLR7, TLR8, CXCL8/IL8, IL1RN, IL18R1). When comparing active and inactive RA, upregulated TLR2, TLR4, TLR6, and TLR8 and downregulated TLR10 (P<0.04) expression was associated with the disease activity.

---

### Conclusion

Our study on the comprehensive innate gene profiling together with multivariate analysis revealed a certain heterogeneity in innate signature within RA patients. Whether the heterogeneity of RA elucidated from diversity in innate signatures may impact the disease course and treatment response deserves future investigations.

---

### Key words

rheumatoid arthritis, heterogeneity, IL1 family, Toll-like receptors, disease activity

Anna Petrackova  
 Pavel Horak  
 Martin Radvansky  
 Regina Fillerova,  
 Veronika Smotkova Kraiczova  
 Milos Kudelka  
 Frantisek Mrazek  
 Martina Skacelova  
 Andrea Smrzova  
 Eva Kriegova

Please address correspondence to:

Eva Kriegova,  
 Department of Immunology,  
 Faculty of Medicine and Dentistry,  
 Palacky University Olomouc  
 and University Hospital Olomouc,  
 Hnevotinska 3,  
 775 15 Olomouc, Czech Republic.  
 E-mail: eva.kriegova@email.cz

Received on April 4, 2019; accepted  
 in revised form on May 27, 2019.

© Copyright CLINICAL AND  
 EXPERIMENTAL RHEUMATOLOGY 2020.

## Introduction

Rheumatoid arthritis (RA) is a chronic systemic inflammatory disease characterised by synovial inflammation and the progressive destruction of joint cartilage and bones (1). The pathogenesis of RA is complex with a growing body of evidence of a major impact of innate immunity and type I interferons (IFNs), respectively (2-4).

The major players in innate immunity are Toll-like receptors (TLRs) and members of the interleukin (IL)-1/IL-1R family, both of which share the same intracellular signalling Toll-IL-1-receptor (TIR) homology domain. Thus, a strong pro-inflammatory signal leading to NF- $\kappa$ B activation is indistinguishable in both the TLR and IL-1 ligands (5). The TLRs may be activated by i) *Proteus* infection of the urinary tract and oral and gut dysbiosis, ii) Epstein-Barr virus and parvovirus B19, and iii) endogenous TLR ligands such as the heat shock protein gp96 and tenascin in RA (reviewed in (6, 7)). Similarly, several members of the IL-1 family were found to be over-expressed in the synovial membrane in RA, making a substantial contribution to the alteration of cartilage and bone homeostasis (8). Importantly, RA was the first disease in which IL-1 inhibition was successfully applied, leading to reduced inflammation and articular damage (9).

Recently, a heterogeneity within genes regulated by IFN type I (IFN signature) has been reported in RA (10). Although there is a lack of knowledge of the exact mechanisms leading to aberrant IFN activation in autoimmunity, the activation of the IFN signature has been linked to TLRs and other innate genes (11, 12). The IFN signature is believed to prompt the tolerance breakdown and the subsequent autoimmune perpetuation (13). From the clinical point of view, an enhanced IFN signature has been associated with clinical outcome, treatment response and disease activity in RA (10, 14-16), although some controversy still exists concerning its clinical relevance (10).

Based on the existing linkage between IFN and innate signatures, we were wondering whether innate signature

shows the heterogeneity in RA. We, therefore, analysed the complex expression pattern of innate genes including *TLR1-10*, seven members of the *IL1/IL1R* family and interleukin 8 (*CXCL8/IL8*) in peripheral blood mononuclear cells (PBMCs) of patients with RA. Using the multivariate data mining analysis, we evaluated the diversity of the innate signatures in RA and its relationship to the disease activity.

## Materials and methods

### Study subjects

The study cohort consisted of 67 Czech patients who met the 2010 ACR/EULAR classification criteria for RA (17) and were recruited at a single tertiary rheumatology centre. All the patients were treated according to the national Czech guidelines and standard protocols (18); for the medication used, duration of the disease, and the demographic and clinical features see Table I. Subgroups were formed on the basis of the disease activity as assessed by means of the Disease Activity Score in 28 joints (DAS28), with a DAS28 of  $\geq 3.2$  being taken as active RA (inactive RA, n=35; active RA, n=32). The baseline demographic and clinical data, as well as a type of medications, its duration and cumulative steroid dosage, did not differ between subgroups of active and inactive patients ( $p > 0.05$ ). The age- and gender-matched healthy control subjects comprised 55 medical staff members or their relatives (mean age 54 yrs, range 41–90 yrs, female/male 45/10) in whom autoimmune and inflammatory diseases, recent vaccination, infection, and usage of immunosuppressive drugs were excluded by means of questionnaires.

The patients and control subjects provided written informed consent in accordance with the Helsinki Declaration about the use of peripheral blood for the purpose of this study, which was approved by the ethics committee of the University Hospital and Palacký University Olomouc.

### Real-time reverse transcription-polymerase chain reaction (qRT-PCR)

The PBMC were isolated from blood collected in EDTA tubes by Ficoll den-

*Funding: this work was supported by the Grant Agency of Ministry of Health of the Czech Republic (MZ CR VES15-28659A), in part MH CZ-DRO (FNOL, 00098892) and IGA UP\_2019\_014.*

*Competing interests: none declared.*

sity gradient centrifugation (Sigma-Aldrich, Germany), then lysed in Tri reagent (Sigma-Aldrich, Germany) and frozen at -80°C. Total RNA was extracted using a Direct-zol RNA kit (Zymo Research, USA) according to the manufacturer's recommendations. Reverse transcription was performed using a Transcriptor First Strand cDNA Synthesis Kit (Roche, Switzerland) as reported previously (19).

qPCR was performed using a high-throughput SmartChip Real-Time-qPCR system (WaferGen, USA) allowing 5,184 reactions per chip. The reactions were carried out in 100 nl reaction volume containing LightCycler 480 SYBR Green I Master mix (Roche, Switzerland) with 1.6 µM (each) of gene-specific exon-spanning primers and 0.27 ng of cDNA in quadruplicates. The primer sequences are listed in Table S1 (Integrated DNA Technologies, USA). Each run included a no-template control, in which RNA was replaced by water, and human universal reference RNA (Stratagene, USA) which was used in quadruplicates as a calibrator at a 0.27 ng/reaction mix. The thermal cycler parameters were as follows: one cycle of 95°C for 5 min followed by 40 cycles of 34 s at 95°C and 1 min 4 s at 60°C. Melting curve analysis was performed from 97°C to 60°C (0.4°C/step) immediately after amplification. The relative mRNA expression was calculated using Phosphoglycerate kinase 1 as the reference gene (20).

In order to assess the innate immunity gene expression signature in RA, we investigated the expression of *TLR* (*TLR1-10*), the *IL1/IL1R* family (21 members), and *CXCL8/IL8* in PBMC. On the basis of the pilot evaluation of qPCR assays on a cohort of 20 RA patients, 14 assays of *IL1/IL1R* family members (*IL1A*, *IL36RN*, *IL36A*, *IL36B*, *IL36G*, *IL37*, *IL38*, *IL33*, *IL1R2*, *IL18RAP*, *IL1RL1*, *IL1RL2*, *IL1RAPL1*, *IL1RAPL2*) were below the detection limit of the system and were, therefore, excluded from further analysis. The study continued with gene expression profiling of 18 innate immunity genes: *TLR1-10* and seven members of the *IL1/IL1R* family, together with *CXCL8/IL8*.

**Table I.** Demographic and clinical characteristics of RA patients.

Demographic and clinical features	RA (n=67)	Inactive RA (n=35)	Active RA (n=32)
Female/Male	56/11	31/4	25/7
Age (years) mean (min-max)	55 (27-80)	52 (27-73)	57 (44-61)
Age at the onset of the disease (years) mean (min-max)	39 (5-65)	40 (5-65)	38 (15-57)
Duration of the disease (years) mean (min-max)	16 (1-58)	13 (1-33)	19 (1-58)
Disease activity:			
DAS28 mean (min-max)	3.44 (0.60-6.70)	2.41 (0.60-3.14)	4.56 (3.20-6.70)
ESR (mm/hr) mean (min-max)	19 (2-116)	12 (2-40)	27 (3-116)
CRP (mg/l) mean (min-max)	8.0 (0.6-65.0)	3.0 (0.6-9.2)	13.4 (0.6-65.0)
ACPA positive, % (n)	76 (51)	77 (27)	75 (24)
RF positive, % (n)	66 (44)	69 (24)	63 (20)
Medications, % (n)			
Steroids	70 (47)	51 (18)	91 (29)
NSAIDs	64 (43)	51 (18)	78 (25)
Methotrexate	85 (57)	86 (30)	84 (27)
Other DMARDs*	19 (13)	8 (2)	34 (11)
Biologics <sup>#</sup>	46 (31)	49 (17)	44 (14)

ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; ACPA: anticitrullinated protein antibodies; RF: rheumatoid factor; NSAIDs: non-steroidal anti-inflammatory drugs; DMARDs: disease-modifying anti-rheumatic drugs.

\*Other DMARDs taken were hydroxychloroquine (n=3), leflunomide (n=7), sulfasalazine (n=2), and combination of leflunomide and sulfasalazine (n=1).

<sup>#</sup>Biologics taken were TNF-α inhibitors (n=18), tocilizumab (n=6), abatacept (n=4) and rituximab (n=3).

*Statistical analysis and data mining methods*

Statistical analysis (Mann-Whitney U-test, Kruskal-Wallis test, Benjamini-Hochberg correction, Shapiro-Wilk test) of relative gene expression values were calculated using GraphPad Prism 5.01 (GraphPad Software, USA) and the R statistical software package, a free software environment for statistical computing and graphics (<http://www.r-project.org/>). Spearman correlation between gene expression and continuous DAS28 values were performed using Genex (MultiD Analyses AB, Sweden). A *p*-value <0.05 was considered significant. Firstly, the LRNet algorithm (21) was used to construct a patient similarity network (PSN) to show the similarities of the gene expression profiles among individual patients. The nearest neighbours within the network have the highest similarity in terms of gene expression levels and colours distinguish the particular subgroups of patients with similar profiles. To obtain a set of the most characteristic genes, we constructed these networks based on different combinations of a small number of genes. The selection of the best gene

combination for active and inactive RA patients was evaluated by measured values of weighted modularity (the network partitioning ability) and silhouette (evaluation of the internal quality of clusters) (22, 23). For more details, see On-line supplementary file.

Secondly, a neural network-based algorithm (ANN), together with 10-fold cross-validation (Neuralnet package ([https://cran.r-project.org/package\\_neuralnet](https://cran.r-project.org/package_neuralnet)), from the R software) was applied to a learning set of 57 RA patients with known disease activity status. For the pre-selection of the most informative genes for ANN, the Random Forest machine learning classifier was applied. The selection of the best combination of ANN markers and ANN structure was performed on the basis of the root mean square error (RMSE) and classification error. The classification error for the top marker sets and final ANN was calculated on a validation cohort of 10 patients in whom the activity status was hidden from the bioinformaticians (MR, MK). A flowchart of the process is documented in Supplementary Fig. S1. For more details, see On-line supplementary file. Next,

**Table II.** Relative mRNA expression levels of genes differentially expressed between A) RA vs. healthy controls, B) active vs. inactive RA.

**A: RA vs. healthy controls**

Gene	Mean (95 % CI)		FC	p	P <sub>corr</sub>
	Healthy controls	RA			
SIGIRR	0.196 (0.167-0.225)	0.367 (0.329-0.405)	1.87	3.9 × 10 <sup>-10</sup>	0.1 × 10 <sup>-9</sup>
IL18	0.036 (0.031-0.042)	0.060 (0.054-0.067)	1.56	4.1 × 10 <sup>-8</sup>	3.7 × 10 <sup>-7</sup>
IL1RN	0.018 (0.013-0.024)	0.039 (0.034-0.044)	2.75	1.4 × 10 <sup>-7</sup>	8.6 × 10 <sup>-7</sup>
TLR5	0.029 (0.020-0.037)	0.060 (0.052-0.067)	3.20	4.4 × 10 <sup>-7</sup>	2.0 × 10 <sup>-6</sup>
IL18R1	0.006 (0.004-0.007)	0.011 (0.009-0.012)	1.99	3.4 × 10 <sup>-6</sup>	1.2 × 10 <sup>-5</sup>
TLR3	0.003 (0.002-0.004)	0.006 (0.005-0.007)	6.59	1.8 × 10 <sup>-5</sup>	5.4 × 10 <sup>-5</sup>
IL1RAP	0.008 (0.006-0.010)	0.014 (0.012-0.017)	2.08	4.2 × 10 <sup>-5</sup>	1.1 × 10 <sup>-4</sup>
TLR8	0.040 (0.032-0.049)	0.062 (0.053-0.071)	1.59	4.2 × 10 <sup>-4</sup>	9.5 × 10 <sup>-4</sup>
IL1B	0.035 (0.002-0.067)	0.062 (0.033-0.091)	1.79	8.2 × 10 <sup>-4</sup>	1.6 × 10 <sup>-3</sup>
TLR2	0.049 (0.035-0.062)	0.067 (0.057-0.077)	1.91	1.3 × 10 <sup>-3</sup>	2.3 × 10 <sup>-3</sup>
CXCL8/IL8	0.108 (0.025-0.191)	0.145 (0.096-0.195)	2.48	2.2 × 10 <sup>-3</sup>	3.7 × 10 <sup>-3</sup>
TLR10	0.007 (0.006-0.008)	0.010 (0.008-0.011)	1.41	2.1 × 10 <sup>-2</sup>	3.2 × 10 <sup>-2</sup>
TLR4	0.050 (0.043-0.057)	0.041 (0.036-0.046)	0.86	3.2 × 10 <sup>-2</sup>	4.5 × 10 <sup>-2</sup>

**B: Active vs. inactive RA**

Gene	Mean (95 % CI)		FC	p	P <sub>corr</sub>
	Inactive RA	Active RA			
TLR10	0.011 (0.009-0.013)	0.008 (0.005-0.011)	0.49	6.5 × 10 <sup>-3</sup>	1.2 × 10 <sup>-1</sup>
TLR8	0.057 (0.042-0.072)	0.067 (0.056-0.077)	1.37	1.4 × 10 <sup>-2</sup>	1.2 × 10 <sup>-1</sup>
TLR6	0.023 (0.017-0.028)	0.030 (0.024-0.036)	1.57	2.1 × 10 <sup>-2</sup>	1.3 × 10 <sup>-1</sup>
TLR2	0.057 (0.046-0.068)	0.078 (0.061-0.095)	1.40	3.3 × 10 <sup>-2</sup>	1.5 × 10 <sup>-1</sup>
TLR4	0.039 (0.031-0.048)	0.043 (0.037-0.049)	1.34	4.1 × 10 <sup>-2</sup>	1.5 × 10 <sup>-1</sup>

P<sub>corr</sub> value corrected for multiple comparisons (Benjamini-Hochberg correction)  
 FC (Fold change) between group medians of relative mRNA expression levels.

a gene expression similarity network was constructed by means of the LR-Net algorithm using nearest neighbour and representativeness analysis (21) in subgroups of active and inactive RA patients. The network vertices represent the individual genes and the size of each vertex corresponds to the local importance of the expression of a particular gene on the basis of the number of its nearest neighbours (= other genes). The links (edges) between vertices and their strength represent the similarities between pairs of vertices. For more details, see On-line supplementary file.

**Results**

*Innate immune gene expression signature of RA*

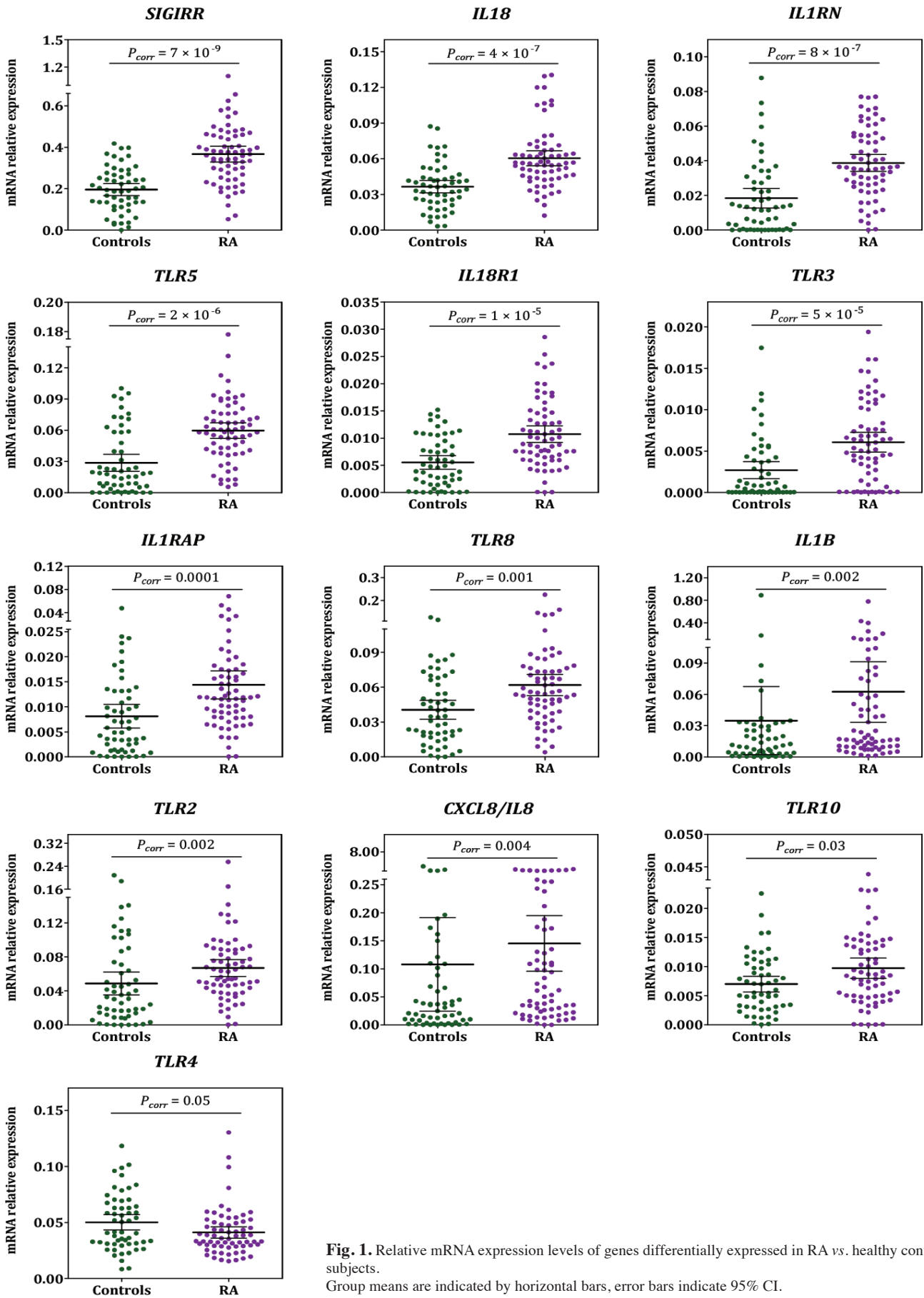
In order to gain a deeper insight into the innate immune system associated with RA, we investigated the innate immunity expression signature in the RA patients and healthy controls. Since our data did not meet the assumption of nor-

malty as assessed by the Shapiro-Wilk test, the non-parametric Mann-Whitney U-test was used for the comparison of data distribution between two groups. Of the thirteen deregulated genes in RA were six TLRs: upregulated *TLR2*, *TLR3*, *TLR5*, *TLR8*, and *TLR10*, and downregulated *TLR4* comparing to controls ( $p_{corr} < 0.05$ ; Table IIA; Fig. 1). Of IL-1/IL-1R family, six members were upregulated *IL1B*, *IL1RN*, *IL18*, *IL18R1*, *IL1RAP*, and *SIGIRR/IL1R8*, as well as upregulated chemokine *CXCL8/IL8* ( $P_{corr} < 0.05$ ; Table IIA; Fig. 1) in RA. The expression of *IL1R1*, *TLR1*, *TLR6*, *TLR7*, and *TLR9* were not different between RA and controls ( $p_{corr} > 0.05$ ; Suppl. Table S2A). To exclude the differences in gene profiles between patients treated with different drugs, we compared subgroups formed on the basis of the medications used. No difference was observed between the profiles in the subgroups of patients based on various medications ( $p > 0.05$ ).

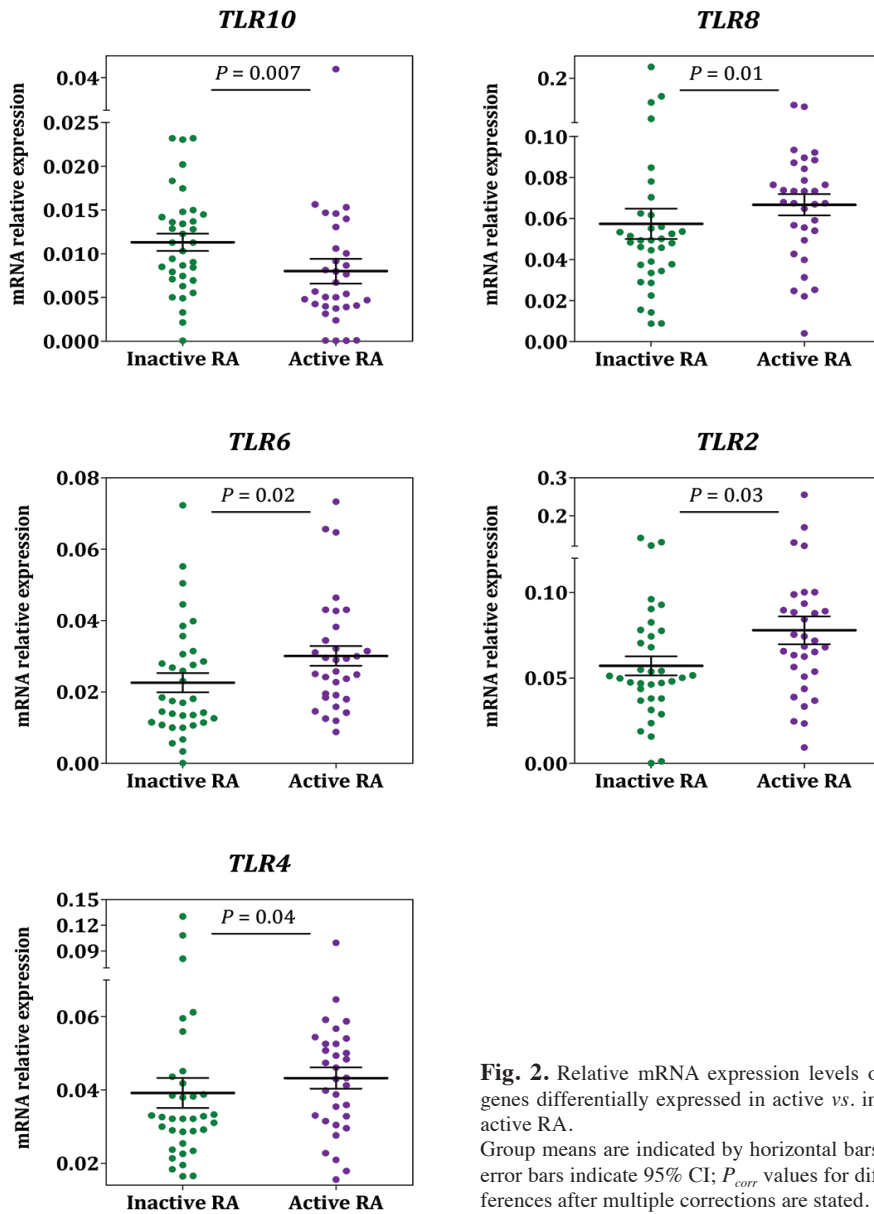
*Innate immune gene expression signature associated with active and inactive RA*

Next, we wondered which genes or their combinations characterise patients with active and inactive RA. When active and inactive RA were compared, upregulated expression of *TLR2*, *TLR4*, *TLR6*, and *TLR8* and downregulation of *TLR10* was observed in patients with active disease ( $p < 0.05$ ; Table IIB, Suppl. Table S2B, Fig. 2). Moreover, the DAS28 score correlated negatively with the expression of *TLR10* ( $r = -0.367$ ,  $p = 0.002$ , Fig. S2A) and positively with the expression of *TLR8* ( $r = 0.236$ ,  $p = 0.05$ , Fig. S2B). Although the mRNA expression of *TLR4* was downregulated in the RA patients as a whole when compared to the healthy controls, subanalysis in subgroups according to the disease activity revealed upregulation of *TLR4* in active RA (Fig. 3). Regarding *TLR10*, the mRNA expression of *TLR10* was upregulated in the RA patients as a whole when compared to the healthy controls, while subanalysis revealed *TLR10* mRNA downregulation in those patients with active RA (Fig. 3). Concerning the IL-1/IL-1R family, no difference was observed in the expression of the genes that were studied between the patients with active and inactive disease when basic statistics were applied (Suppl. Table S2B). Then, combinations of multiple genes for the discrimination of active and inactive RA were investigated by multivariate data analysis. The gene expression similarity network for active RA was characterised by the expression of *TLR2*, *TLR3*, *TLR8*, and *IL18R1*, and inactive RA was associated with the *TLR2*, *TLR5*, *TLR7*, *IL18R1*, and *IL1RAP* genes (Fig. 4). The selected genes had the highest representativeness in the individual networks. To investigate and visually assess the complex expression innate signatures in our patients, we performed analysis by utilising the abovementioned patient similarity networks. This analysis revealed the existence of four patient's subsets (clusters) based on different *TLR8* and *IL1RN* expression profiles, two in active and two in inactive RA. The high modularity and the good per-





**Fig. 1.** Relative mRNA expression levels of genes differentially expressed in RA vs. healthy control subjects. Group means are indicated by horizontal bars, error bars indicate 95% CI.



**Fig. 2.** Relative mRNA expression levels of genes differentially expressed in active vs. inactive RA. Group means are indicated by horizontal bars; error bars indicate 95% CI;  $P_{corr}$  values for differences after multiple corrections are stated.

formance of silhouette analysis were observed across the combinations that were tested (Fig. 5). To exclude the differences in gene profiles due to the different treatment regimen, the distributions of used drugs in particular subgroups were compared. As shown in Figure S3, the proportion of patients treated with a particular drug did not differ among revealed subsets (clusters). Further, using ANN we identified two combinations of genes: *TLR1*, *TLR2*, *TLR7*, *TLR8*, *IL1RN*, *IL18R1*, and *CXCL8/IL8*, and *TLR1*, *TLR2*, *TLR3*, *IL1RN*, and *IL18R1*, whose co-expression discriminates between patients with active and inactive RA. With these combinations used as an input to a clas-

sifier containing ten neural networks, 80% overall agreement was achieved for blinded patient data on the basis of five-fold cross-validation. Furthermore, two combinations of genes were needed for successful characterisation of the subgroups of patients, showing that within the active RA subgroup, there are at least two different gene expression signatures. When ANN was constructed for the combination of only TLR genes associated with disease activity on the basis of classical statistics (*TLR2*, *TLR4*, *TLR6*, *TLR8*, and *TLR10*), this combination reached only 40% overall agreement for blinded patient data based on five-fold cross-validation. Moreover, the observed exist-

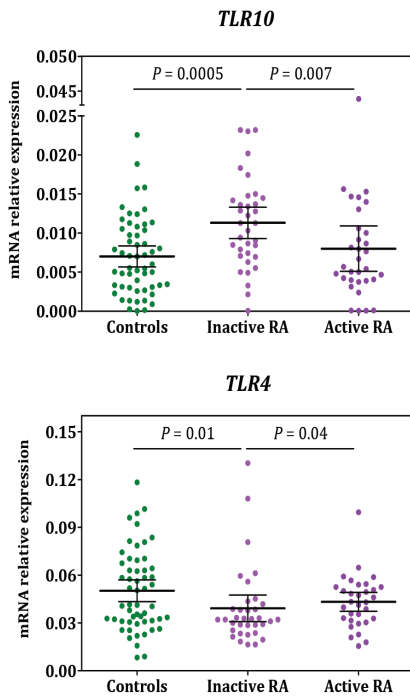
ence of two main subsets with different expression signatures within the active RA patients confirmed the result from patient similarity network analysis.

**Discussion**

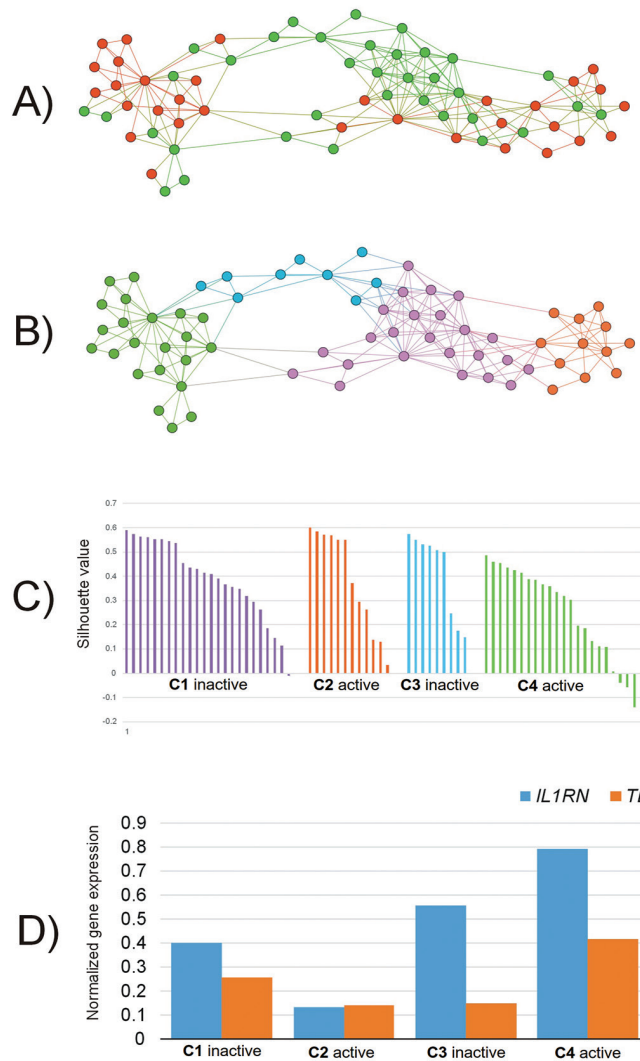
A growing body of evidence highlights a persistent activation of the innate immune system and IFN signature in the pathogenesis of RA as well as its relationship with the disease activity. In addition to recent studies that revealed heterogeneity in the IFN signature in RA (10, 14), our study for the first time also highlighted the heterogeneity in the innate signature within RA patients.

To analyse the innate signature in our patients, we used two multivariate data mining approaches that have excellent properties for analysing gene expression patterns. Firstly, gene expression and patient similarity networks exploration enables visual assessment of the most informative markers within a sample set and shows the relationship between patients with similar gene profiles (24-26). Also applied neural network approach takes into account the intrinsic characteristics of gene expression data (27, 28), confirms the most informative gene subsets, and improves classification accuracy with best parameters based on datasets (14, 24, 29, 30). Using patient similarity network analysis, four patient's subsets based on the innate signature were detected, two in active and two in inactive RA. The applied network exploration identified expression of *TLR8* and *IL1RN* as the most discriminant among detected subgroups. Importantly, the heterogeneity in RA patients was further supported by the neural network analysis which identified two main gene sets describing active RA within an innate signature (*TLR1*, *TLR2*, *TLR3*, *TLR7*, *TLR8*, *CXCL8/IL8*, *IL1RN*, *IL18R1*). Our data for the first time identified certain heterogeneity in innate signature in RA, which may have a significant impact on the disease course and treatment response, thus deserving future investigations.

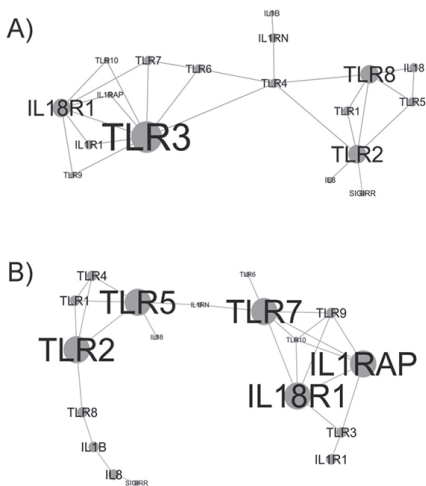
From our results, *TLR8* and *IL1RN* appear to be key genes whose expressions characterise diversity in RA and active and inactive RA subgroups. *TLR8* is



**Fig. 3.** Relative mRNA expression levels of *TLR4* and *TLR10* in active RA, inactive RA, and healthy control subjects. Group means are indicated by horizontal bars, error bars indicate 95 % CI. Kruskal-Wallis test revealed differences among all tested subgroups (*TLR4*:  $p=0.02$ ; *TLR10*:  $p=0.001$ ). The horizontal connecting lines show significant differences between two particular subgroups (controls vs. inactive RA, inactive RA vs. active RA, respectively); comparison between controls and active RA did not reach significance for both *TLR4* and *TLR10*.



**Fig. 5.** Patient similarity network analysis based on *TLR8* and *IL1RN* expression in RA patients **A)** with active (red) and inactive disease (green). **B)** Of four well-separated clusters, two clusters included predominantly inactive RA patients (C1 - violet, C3 - blue) and two predominantly active RA (C2 - orange, C4 - green). **C)** The silhouette analysis of detected clusters. The bars represent individual patients, and high values for them indicate that the patient is well matched to their own cluster and poorly matched to neighbouring clusters. **D)** Characteristics of observed clusters showing the normalised gene expression values of *IL1RN* and *TLR8*.



**Fig. 4.** Gene expression similarity network in **A)** active and **B)** inactive RA. Vertices represent the individual genes, and the size of each vertex corresponds to the local importance (representativeness) of the expression of a particular gene on the basis of the number of its nearest neighbours. Links (edges) between vertices and their strength represent similarities between pairs of vertices.

able to recognise viral single-stranded and bacterial RNAs and induce both NF- $\kappa$ B-dependent cytokines and type I IFNs (31). Studies on the role of the *TLR8* in arthritic inflammation showed that its increased expression correlates with the elevation of IL-1 $\beta$  levels and disease status (32). Another study proved that the activation of the *TLR8* signalling pathway in human blood results in a predominant pro-inflammatory gene signature (33). Importantly, recent studies demonstrated that the activation profile of individual TLRs may be influenced by the complex TLR-TLR interactions (34, 35). Regarding *TLR8*,

a distinct immune activation profile was observed by co-signalling of *TLR8* together with *TLR2* when compared to *TLR8* alone (36). Moreover, *TLR2* was shown to suppress IFN $\beta$  production induced by *TLR8* activation (37). Additionally, a cross-talk of *TLR8* with other endosomal TLRs has been identified crucially involved in the generation of autoimmunity (31). Next, key gene within innate signature was *IL1RN* which codifies IL-1 receptor antagonist (IL-1Ra) that blocks IL-1 signalling. The importance of *IL1RN* in the RA pathogenesis has been demonstrated by *Il1rn*(-/-) mice, which spontaneously

develop autoimmune arthritis that is dependent on TLR activation (38). In human, associations of sequence variants *IL1RN* VNTR (rs2234663) and +2018 SNP with RA disease activity was reported (39, 40). Other VNTR variant (allele *IL1RN\*2*) was found to influence not only the plasma levels of IL-1Ra, but also the response to infliximab therapy (41). However, contrary to the animal studies (38), we and others observed elevated expression of *IL1RN* in RA patients compared to healthy subjects (39). Our multivariate analysis revealed very high variability in *IL1RN* expression among our patient subgroups, irrespective of the disease activity. This led us to suggest that *IL1RN* may act in a dose-dependent manner and most likely in interaction with complex TLR interactions, thus deserving future investigations. Our data further supports the crucial role of IL-1Ra together with TLR8 in the RA pathogenesis, thus nominating them as candidates for future studies.

Also, other genes or their protein products of an innate signature identified from neural network analysis have been associated with RA disease activity in previous studies, such as TLR3 (42), IL18R1 (43), TLR7 (44), and CXCL8/IL8 (45). Although individual innate genes have been already linked to RA, a more complex picture may be observed when using multivariate analysis. Overall, our findings revealed heterogeneity in RA in innate signature including *TLR* and *IL1/IL1R* genes. Whether the heterogeneity of innate signature contributes to the reported variability in IFN signature in RA (11) deserves future investigations.

In our study, we also comprehensively explored the RA-associated signature when compared to healthy controls. Of 13 innate deregulated genes (*TLR2*, *TLR3*, *TLR4*, *TLR5*, *TLR8*, *TLR10*, *IL1B*, *IL1RN*, *IL18*, *IL18R1*, *IL1RAP*, and *SIGIRR/IL1R8*) associated with RA, *TLR10* and *IL1RAP* are being reported for the first time. In line with our results, the upregulation of *TLR2*, *TLR5*, and *TLR8* in peripheral blood monocytes (46-48) and *TLR3* in synovial tissue (49) was reported. Regarding *TLR4*, which is highly expressed in the synovium (50), we and others (51) ob-

served the downregulation of *TLR4* expression in RA PBMC. We also detected the upregulated expression of *TLR10* in the PBMC of RA patients as a whole for the first time. The first evidence about the possible involvement of TLR10 in RA already exists; it is based on the association of a *TLR10* I473T allelic variant with RA (52). However, there are controversies regarding its function. Some studies have demonstrated that TLR10 is a pro-inflammatory receptor activating NF- $\kappa$ B signalling (53, 54), while others have reported NF- $\kappa$ B inhibitory activity (52, 55) and inflammation suppression (56, 57). Further studies are needed to clarify its function, ligands, and the influence of the genetic background in RA on its regulation. Our study also confirmed the results from analyses of individual members of the IL-1 family: increased expression of the pro-inflammatory members *IL1B*, *IL1RN*, and *IL18* was demonstrated in the peripheral blood cells (39, 58) and of *IL18R1* and *SIGIRR* in the synovial tissue (59, 60) of RA patients. Here, we report for the first time upregulated gene expression of *IL1RAP* in RA. IL-1RAP is a co-receptor involved in several signalling pathways, including IL-1, IL-33, IL-36G, and SCF (61, 62), and a lack of IL-1RAP was shown to abrogate the cellular response to IL-1 (63). The contribution of this co-receptor to the RA pathogenesis deserves future investigation. Additionally, elevation in the expression of the chemokine *CXCL8/IL8* was observed in our study, which is in line with the reported elevation of IL-8 in synovial fluids and serum in RA patients (64). When active and inactive disease were compared, the upregulation of *TLR2*, *TLR4*, *TLR6*, and *TLR8* and downregulation of *TLR10* were revealed in active RA. Concerning TLR10, an association of the I473T allelic variant (rs11466657) with disease severity and a low response to infliximab has been reported (52). Functional studies have shown that the *TLR10* I473T variant lacks inhibitory activity on the NF- $\kappa$ B inflammatory pathway in comparison to the wild-type allele (52). Similarly, downregulation of *TLR10* was also observed in our patients with active disease. These observations

nominate TLR10 as a candidate target molecule able to attenuate the inflammation in active RA.

The authors are aware of some limitations. First, the study was performed in a real-world cohort of patients treated with different medications, however, the distribution of various medications, its duration and dosage did not differ between compared subgroups of active and inactive patients. Second, the innate gene signature should be completed on a protein level of functionally active cytokines in future studies. However, we believe that this multivariate approach highlighted for the first time the heterogeneity of innate molecules in RA and nominated combinations of key innate molecules for further functional studies. To conclude, our study on comprehensive innate gene profiling together with multivariate data mining analysis revealed a certain heterogeneity in innate signature within RA patients. Moreover, *TLR8* and *IL1RN* were identified as the key genes whose expressions contribute to the heterogeneity of innate signature in RA. The clinical consequences of the observed heterogeneity of innate signature in RA should be addressed in future studies. We believe that this integrated approach is likely to generate insights into the heterogeneity of innate signature in RA.

### Acknowledgments

We would like to thank to Radmila Bartova for her help with sample processing.

### References

1. MCINNES IB, SCHEFF G: The pathogenesis of rheumatoid arthritis. *N Engl J Med* 2011; 365: 2205-19.
2. RODRÍGUEZ-CARRIO J, LÓPEZ P, SUÁREZ A: Type I IFNs as biomarkers in rheumatoid arthritis: towards disease profiling and personalized medicine. *Clin Sci* 2015; 128: 449-64.
3. CALABRESI E, PETRELLI F, BONIFACIO AF, PUXEDDU I, ALUNNO A: One year in review 2018: pathogenesis of rheumatoid arthritis. *Clin Exp Rheumatol* 2018; 36: 175-84.
4. BARRERA P, FAURÉ S, PRUD'HOMME JF et al.: European genetic study on rheumatoid arthritis: is there a linkage of the interleukin-1 (IL-1), IL-10 or IL-4 genes to RA? *Clin Exp Rheumatol* 2001; 19: 709-14.
5. DINARELLO CA: Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol Rev* 2018; 281: 8-27.
6. JOOSTEN LA, ABDOLLAHI-ROODSAZ S, DIN-



- ARELLO CA, O'NEILL L, NETEA MG: Toll-like receptors and chronic inflammation in rheumatic diseases: new developments. *Nat Rev Rheumatol* 2016; 12: 344-57.
7. PRETORIUS E, AKEREDOLU OO, SOMA P, KELL DB: Major involvement of bacterial components in rheumatoid arthritis and its accompanying oxidative stress, systemic inflammation and hypercoagulability. *Exp Biol Med* 2017; 242: 355-73.
  8. BRENNAN FM, MCINNES IB: Evidence that cytokines play a role in rheumatoid arthritis. *J Clin Invest* 2008; 118: 3537-45.
  9. SCHEIT G, DAYER JM, MANGER B: Interleukin-1 function and role in rheumatic disease. *Nat Rev Rheumatol* 2016; 12: 14-24.
  10. RODRÍGUEZ-CARRIO J, ALPERI-LÓPEZ M, LÓPEZ P, BALLINA-GARCÍA FJ, SUÁREZ A: Heterogeneity of the type I Interferon signature in rheumatoid arthritis: a potential limitation for its use as a clinical biomarker. *Front Immunol* 2018; 8: 2007.
  11. LÓPEZ DE PADILLA CM, NIEWOLD TB: The type I interferons: basic concepts and clinical relevance in immune-mediated inflammatory diseases. *Gene* 2015; 576: 14-21.
  12. OON S, WILSON NJ, WICKS I: Targeted therapeutics in SLE: emerging strategies to modulate the interferon pathway. *Clin Transl Immunology* 2016; 5: e79.
  13. RÖNNBLUM L, ELORANTA ML: The interferon signature in autoimmune diseases. *Curr Opin Rheumatol* 2013; 25: 248-53.
  14. EL-SHERBINY YM, PSARRAS A, MD YUSOF MY *et al.*: A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. *Sci Rep* 2018; 8: 5793.
  15. COOLES FAH, ANDERSON AE, LENDREM DW *et al.*: The interferon gene signature is increased in patients with early treatment-naive rheumatoid arthritis and predicts a poorer response to initial therapy. *J Allergy Clin Immunol* 2017; 141: 445-8.
  16. DE JONG TD, VOSSLAMBER S, BLITS M *et al.*: Effect of prednisone on type I interferon signature in rheumatoid arthritis: consequences for response prediction to rituximab. *Arthritis Res Ther* 2015; 17: 78.
  17. ALETAHA D, NEOGI T, SILMAN AJ *et al.*: 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010; 69: 1580-8.
  18. ŠENOLT L, MANN H, ZÁVADA J, PAVELKA K, VENCOSKÝ J: 2017 recommendations of the Czech Society for Rheumatology for the pharmacological treatment of rheumatoid arthritis [in Czech]. *Česká Revmatologie* 2017; 25: 8-24.
  19. TOMANKOVA T, KRIEKOVA E, FILLEROVA R, LUZNA P, EHRMANN J, GALLO J: Comparison of periprosthetic tissues in knee and hip joints: differential expression of CCL3 and DC-STAMP in total knee and hip arthroplasty and similar cytokine profiles in primary knee and hip osteoarthritis. *Osteoarthritis Cartilage* 2014; 22: 1851-60.
  20. FALKENBERG VR, WHISTLER T, MURRAY JR, UNGER ER, RAJEEVAN MS: Identification of Phosphoglycerate Kinase 1 (PGK1) as a reference gene for quantitative gene expression measurements in human blood RNA. *BMC Res Notes* 2011; 4: 324.
  21. OCHODKOVA E, ZEHNALOVA S, KUDELKA M: Graph construction based on local representativeness. In: *Computing and Combinatorics*. Cham: Springer International Publishing 2017: 654-65. (Lecture Notes In Computer Science)
  22. BLONDEL VD, GUILLAUME JL, LAMBIOTTE R, LEFEBVRE E: Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* 2008; 10: P10008.
  23. ROUSSEEUW PJ: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987; 20: 53-65.
  24. FILLEROVA R, GALLO J, RADVANSKY M, KRAICZOVA V, KUDELKA M, KRIEKOVA E: Excellent diagnostic characteristics for ultrafast gene profiling of DEFA1-IL1B-LTF in detection of prosthetic joint infections. *J Clin Microbiol* 2017; 55: 2686-97.
  25. STAROSTKA D, KRIEKOVA E, KUDELKA M *et al.*: Quantitative assessment of informative immunophenotypic markers increases the diagnostic value of immunophenotyping in mature CD5-positive B-cell neoplasms. *Cytometry B Clin Cytom* 2018; 94: 576-87.
  26. TURCSANYI P, KRIEKOVA E, KUDELKA M *et al.*: Improving risk-stratification of patients with chronic lymphocytic leukemia using multivariate patient similarity networks. *Leuk Res* 2019; 79: 60-8.
  27. LIU B, CUI Q, JIANG T, MA S: A combinational feature selection and ensemble neural network method for classification of gene expression data. *BMC Bioinformatics* 2004; 5: 136.
  28. CHEN YC, CHANG YC, KE WC, CHIU HW: Cancer adjuvant chemotherapy strategic classification by artificial neural network with gene expression data: An example for non-small cell lung cancer. *J Biomed Inform* 2015; 56: 1-7.
  29. MOTEGHAED, NY, MAGHOOLI K, PIRHADI S, GARSHASBI M: Biomarker discovery based on hybrid optimization algorithm and artificial neural networks on microarray data for cancer classification. *J Med Signals Sens* 2015; 5: 88-96.
  30. MEHRIDEHNAVI A, ZIAEI L: Minimal gene selection for classification and diagnosis prediction based on gene expression profile. *Adv Biomed Res* 2013; 2: 26.
  31. CERVANTES JL, WEINERMAN B, BASOLE C, SALAZAR JC: TLR8: the forgotten relative revindicated. *Cell Mol Immunol* 2012; 9: 434-8.
  32. SACRE SM, LO A, GREGORY B *et al.*: Inhibitors of TLR8 reduce TNF production from human rheumatoid synovial membrane cultures. *J Immunol* 2008; 181: 8002-9.
  33. GUIDUCCI C, GONG M, CEPIKA AM *et al.*: RNA recognition by human TLR8 can lead to autoimmune inflammation. *J Exp Med* 2013; 210: 2903-19.
  34. UNDERHILL DM: Collaboration between the innate immune receptors dectin-1, TLRs, and Nods. *Immunol Rev* 2007; 219: 75-87.
  35. NEGISHI H, YANAI H, NAKAJIMA A *et al.*: Cross-interference of RLR and TLR signaling pathways modulates antibacterial T cell responses. *Nat Immunol* 2012; 13: 659-66.
  36. BÖSL K, GIAMBELLUCA M, HAUG M *et al.*: Coactivation of TLR2 and TLR8 in primary human monocytes triggers a distinct inflammatory signaling response. *Front Physiol* 2018; 9: 618.
  37. BERGSTRØM B, AUNE MH, AWUH JA *et al.*: TLR8 senses *Staphylococcus aureus* RNA in human primary monocytes and macrophages and induces IFN-beta production via a TAK1-IKKbeta-IRF5 signaling pathway. *J Immunol* 2015; 195: 1100-11.
  38. ABDOLLAHI-ROODSAZ S, JOOSTEN LA, KOENDERS MI *et al.*: Stimulation of TLR2 and TLR4 differentially skews the balance of T cells in a mouse model of arthritis. *J Clin Invest* 2007; 118: 205-16.
  39. AMÍREZ-PÉREZ S, DE LA CRUZ-MOSSO U, HERNÁNDEZ-BELO J *et al.*: High expression of interleukine-1 receptor antagonist in rheumatoid arthritis: Association with IL1RN\*2/2 genotype. *Autoimmunity* 2017; 50: 468-75.
  40. ISMAIL E, NOFAL OKJ, SAKTHISWARY R, SHAHARIR SS, SRIDHARAN R, JIANG Y: The clinical significance of interleukin-1 receptor antagonist +2018 polymorphism in rheumatoid arthritis. *PLoS One* 2016; 11: e0153752.
  41. TOLUSSO B, PIETRAPERTOSA D, MORELLI A *et al.*: IL-1B and IL-1RN gene polymorphisms in rheumatoid arthritis: relationship with protein plasma levels and response to therapy. *Pharmacogenomics* 2006; 7: 683-95.
  42. LASKA MJ, HANSEN B, TROLDORGB A *et al.*: A non-synonymous single-nucleotide polymorphism in the gene encoding Toll-like Receptor 3 (TLR3) is associated with seronegative Rheumatoid Arthritis (RA) in a Danish population. *BMC Res Notes* 2014; 7: 716.
  43. MATSUI K, TSUTSUI H, NAKANISHI K: Pathophysiological roles for IL-18 in inflammatory arthritis. *Expert Opin Ther Targets* 2005; 7: 701-24.
  44. KIM SJ, CHEN Z, ESSANI AB *et al.*: Identification of a novel TLR7 endogenous ligand in RA synovial fluid that can provoke arthritic joint inflammation. *Arthritis Rheumatol* 2016; 68: 1099-110.
  45. KLIMIUK PA, SIERAKOWSKI S, DOMYSŁAWSKA I, CHWIECKO J: Serum chemokines in patients with rheumatoid arthritis treated with etanercept. *Rheumatol Int* 2011; 31: 457-61.
  46. IWAHASHI M, YAMAMURA M, AITA T *et al.*: Expression of toll-like receptor 2 on CD16+ blood monocytes and synovial tissue macrophages in rheumatoid arthritis. *Arthritis Rheum* 2004; 50: 1457-67.
  47. CHAMBERLAIN ND, VILA OM, VOLIN MV *et al.*: TLR5, a novel and unidentified inflammatory mediator in rheumatoid arthritis that correlates with disease activity score and joint TNF- $\alpha$  levels. *J Immunol* 2012; 189: 475-83.
  48. CHAMBERLAIN ND, KIM SJ, VILA OM *et al.*: Ligation of TLR7 by rheumatoid arthritis synovial fluid single strand RNA induces transcription of TNF- $\alpha$  in monocytes. *Ann Rheum Dis* 2013; 72: 418-26.
  49. HUANG QQ, MA Y, ADEBAYO A, POPE RM: Increased macrophage activation mediated through toll-like receptors in rheumatoid arthritis. *Arthritis Rheum* 2007; 56: 2192-201.

50. OSPELT C, BRENTANO F, RENGEL Y *et al.*: Overexpression of toll-like receptors 3 and 4 in synovial tissue from patients with early rheumatoid arthritis: Toll-like receptor expression in early and longstanding arthritis. *Arthritis Rheum* 2008; 58: 3684-92.
51. KOWALSKI ML, WOLSKA A, GRZEGORCZYK J *et al.*: Increased responsiveness to Toll-like receptor 4 stimulation in peripheral blood mononuclear cells from patients with recent onset rheumatoid arthritis. *Mediators Inflamm* 2008; 2008: 132732.
52. TORICES S, JULIA A, MUÑOZ P *et al.*: A functional variant of TLR10 modifies the activity of NFκB and may help predict a worse prognosis in patients with rheumatoid arthritis. *Arthritis Res Ther* 2016; 18: 221.
53. HASAN U, CHAFFOIS C, GAILLARD C *et al.*: Human TLR10 is a functional receptor, expressed by B cells and plasmacytoid dendritic cells, which activates gene transcription through MyD88. *J Immunol* 2005; 174: 2942-50.
54. REGAN T, NALLY K, CARMODY R *et al.*: Identification of TLR10 as a key mediator of the inflammatory response to *Listeria monocytogenes* in intestinal epithelial cells and macrophages. *J Immunol* 2013; 191: 6084-92.
55. GUAN Y, RANOA DR, JIANG S *et al.*: Human TLRs 10 and 1 share common mechanisms of innate immune sensing but not signaling. *J Immunol* 2010; 184 :5094-103.
56. STAPPERS MH, OOSTING M, IOANA M *et al.*: Genetic variation in TLR10, an inhibitory Toll-like receptor, influences susceptibility to complicated skin and skin structure infections. *J Infect Dis* 2015; 212: 1491-9.
57. OOSTING M, CHENG SC, BOLSCHER JM *et al.*: Human TLR10 is an anti-inflammatory pattern-recognition receptor. *Proc Natl Acad Sci USA* 2014; 111: E4478-84.
58. EDWARDS CK, GREEN JS, VOLK HD *et al.*: Combined anti-tumor necrosis factor-α therapy and DMARD therapy in rheumatoid arthritis patients reduces inflammatory gene expression in whole blood compared to DMARD therapy alone. *Front Immunol* 2012; 3: 366.
59. SHAO XT, FENG L, GU LJ *et al.*: Expression of interleukin-18, IL-18BP, and IL-18R in serum, synovial fluid, and synovial tissue in patients with rheumatoid arthritis. *Clin Exp Med* 2009; 9: 215-21.
60. CAVALLI G, KOENDERS M, KALABOKIS V *et al.*: Treating experimental arthritis with the innate immune inhibitor interleukin-37 reduces joint and systemic inflammation. *Rheumatology* 2016; 55: 2220-9.
61. DINARELLO CA: Immunological and inflammatory functions of the interleukin-1 family. *Annu Rev Immunol* 2009; 27: 519-50.
62. DRUBE S, HEINK S, WALTER S *et al.*: The receptor tyrosine kinase c-Kit controls IL-33 receptor signaling in mast cells. *Blood* 2010; 115: 3899-906.
63. CULLINAN EB, KWEE L, NUNES P *et al.*: IL-1 receptor accessory protein is an essential component of the IL-1 receptor. *J Immunol* 1998; 161: 5614-20.
64. KOCH AE, KUNKEL SL, BURROWS JC *et al.*: Synovial tissue macrophage as a source of the chemotactic cytokine IL-8. *J Immunol* 1991; 147: 2187-95.

## APPENDIX D

**Petrackova A**, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, Papajik T, Kriegova E. Standardization of sequencing coverage depth in NGS: recommendation for detection of clonal and subclonal mutations in cancer diagnostics. *Front Oncol.* 2019;9:851. (IF 2019: 4.848)

*Dean's award for the best student scientific publications in 2019*



# Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics

Anna Petrackova<sup>1</sup>, Michal Vasinek<sup>2</sup>, Lenka Sedlarikova<sup>1</sup>, Tereza Dyskova<sup>1</sup>, Petra Schneiderova<sup>1</sup>, Tomas Novosad<sup>2</sup>, Tomas Papajik<sup>3</sup> and Eva Kriegova<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Ye Wang,  
Qingdao University Medical  
College, China

### Reviewed by:

Robert Charles Day,  
University of Otago, New Zealand  
Ole Winther,  
Technical University of  
Denmark, Denmark

### \*Correspondence:

Eva Kriegova  
eva.kriegova@email.cz

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Oncology

**Received:** 07 June 2019

**Accepted:** 19 August 2019

**Published:** 04 September 2019

### Citation:

Petrackova A, Vasinek M,  
Sedlarikova L, Dyskova T,  
Schneiderova P, Novosad T, Papajik T  
and Kriegova E (2019) Standardization  
of Sequencing Coverage Depth in  
NGS: Recommendation for Detection  
of Clonal and Subclonal Mutations in  
Cancer Diagnostics.  
*Front. Oncol.* 9:851.  
doi: 10.3389/fonc.2019.00851

<sup>1</sup> Department of Immunology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czechia, <sup>2</sup> Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava, Ostrava, Czechia, <sup>3</sup> Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czechia

The insufficient standardization of diagnostic next-generation sequencing (NGS) still limits its implementation in clinical practice, with the correct detection of mutations at low variant allele frequencies (VAF) facing particular challenges. We address here the standardization of sequencing coverage depth in order to minimize the probability of false positive and false negative results, the latter being underestimated in clinical NGS. There is currently no consensus on the minimum coverage depth, and so each laboratory has to set its own parameters. To assist laboratories with the determination of the minimum coverage parameters, we provide here a user-friendly coverage calculator. Using the sequencing error only, we recommend a minimum depth of coverage of 1,650 together with a threshold of at least 30 mutated reads for a targeted NGS mutation analysis of  $\geq 3\%$  VAF, based on the binomial probability distribution. Moreover, our calculator also allows adding assay-specific errors occurring during DNA processing and library preparation, thus calculating with an overall error of a specific NGS assay. The estimation of correct coverage depth is recommended as a starting point when assessing thresholds of NGS assay. Our study also points to the need for guidance regarding the minimum technical requirements, which based on our experience should include the limit of detection (LOD), overall NGS assay error, input, source and quality of DNA, coverage depth, number of variant supporting reads, and total number of target reads covering variant region. Further studies are needed to define the minimum technical requirements and its reporting in diagnostic NGS.

**Keywords:** next-generation sequencing, variant allele frequency (VAF), coverage depth calculator, sequencing error, small subclones, TP53 gene

## INTRODUCTION

Next-generation sequencing (NGS) has rapidly expanded into the clinical setting in haemato-oncology and oncology, as it may bring great benefits for diagnosis, selection of treatment, and/or prognostication for many patients (1). Recently, several articles about the validation of deep targeted NGS in clinical oncology were published (2, 3), including a comprehensive recommendation by the Association for Molecular Pathology and the College of American Pathologists (1). However, the lack of standardization of targeted NGS methods still limits their implementation in clinical practice (4).

One challenge in particular is the correct detection of mutations present at low variant allele frequencies (VAF) and standardization of sequencing coverage depth (1, 5, 6). This is especially important for mutations that have clinical impacts at subclonal frequencies (1) such as the case of *TP53* gene mutations (*TP53mut*) in chronic lymphocytic leukemia (CLL) (7, 8). *TP53* aberrations (*TP53mut* and/or chromosome 17p deletion) are among the strongest prognostic and predictive markers guiding treatment decisions in CLL (9). Nowadays, the European Research Initiative on Chronic Lymphocytic Leukemia (ERIC) recommends detecting *TP53mut* with a limit of detection (LOD) of at least 10% VAF (10), and a growing body of evidence exists dedicated to the clinical impact of small *TP53* mutated subclones in CLL (7, 8).

Sanger sequencing and deep targeted NGS are currently the techniques most used for *TP53mut* analysis (10) as well as for analysis of other genes with clinical impacts at low allele frequencies. Although Sanger sequencing provides a relatively accessible sequencing approach, it lacks the sensitivity needed to detect subclones due to its detection limit of 10–20% of mutated alleles (10). NGS-based analysis has thus gained prominence in diagnostic laboratories for the detection of somatic variants and various technical developments of error correction strategies, both computational and experimental, are being developed for the accurate identification of low-level genetic variations (11). We therefore address the importance of the correct determination of sequencing depth in diagnostic NGS in order to obtain a confident and reproducible detection, not only of low VAF variants. Finally, we performed a dilution experiment to confirm our theoretical calculations, and we close by discussing our experience with diagnostic detection of *TP53mut* in CLL patients and further perspectives about NGS standardization in cancer diagnostics.

## NGS SEQUENCING DEPTH AND ERROR RATE

NGS sequencing depth directly affects the reproducibility of variant detection: the higher the number of aligned sequence reads, the higher the confidence to the base call at a particular position, regardless of whether the base call is the same as the reference base or is mutated (1). In other words, individual sequencing error reads are statistically irrelevant when they are outnumbered by correct reads. Thus, the desired coverage depth

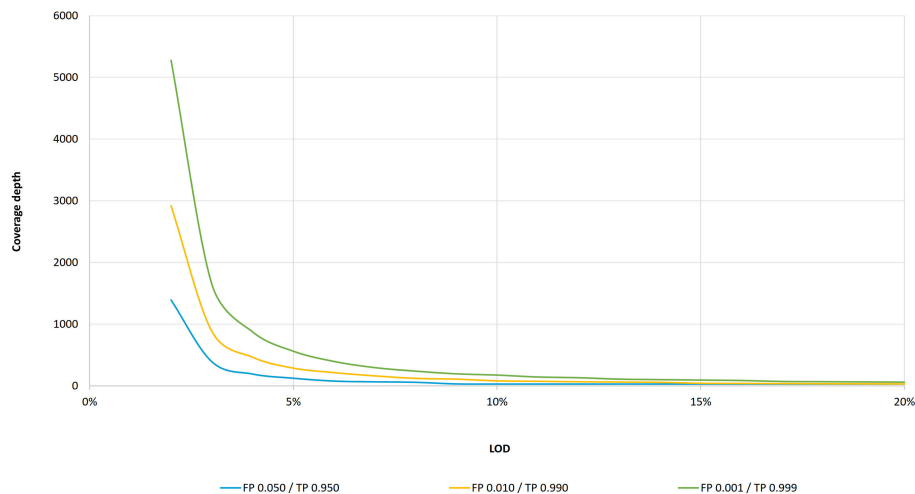
should be determined based on the intended LOD, the tolerance for false positive or false negative results, and the error rate of sequencing (1, 11).

Using a binomial distribution, the probability of false positive and false negative results for a given error rate as well as the intended LOD can be calculated, and the threshold for a variant calling for a given depth can be estimated (1). For example, given a sequencing error rate of 1%, a mutant allele burden of 10%, and a depth of coverage 250 reads, the probability of detecting 9 or fewer mutated reads is, according to the binomial distribution, 0.01%. Hence, the probability of detecting 10 or more mutated reads is 99.99% (100–0.01%), and the threshold for a variant calling can be defined. In other words, a coverage depth of 250 with a threshold of at least 10 mutated reads will have a 99.99% probability that 10% of the mutant allele load will not be missed by the variant calling (although it can be detected in a different proportion). In this way, the risk of a false negative result is greatly minimized. On the other hand, the probability of false positives heavily depends on the sequencing error rate (as the accuracy of all analytical measurements depends on the signal-to-noise ratio) (1, 11). In our example, the probability of a false positive result is 0.025%; however, the rate of false positives is not negligible when decreasing the LOD to the value close to the error rate. Conventional intrinsic NGS error rates range between 0.1 and 1% (Phred quality score of 20–30) (1, 11) depending on the sequencing platform, the GC content of the target regions (12), and the fragment length, as shown in Illumina paired-end sequencing (13). Therefore, the detection of variants at VAFs <2% is affected by a high risk of a false positive result, regardless of the coverage depth. It is also important to mention that the sequencing error rate applies only for errors produced by sequencing itself and does not include other errors introduced during DNA processing and library preparation, particularly during amplification steps, which further increase error rates (1, 11).

## MINIMUM SEQUENCING COVERAGE IN CLINICAL SETTINGS

There is currently no consensus on the minimum required coverage in a clinical setting using deep targeted resequencing by NGS, and so each laboratory has to set its own parameters in order to meet sufficient quality (1, 5). To date, only a few studies have recommended the minimum coverage criteria for deep targeted NGS in clinical oncology: 500 depth of coverage and a LOD of 5% (2), 300–500 depth of coverage without defying the LOD (3), 250 depth and a LOD of 5% with threshold adjustment to 1,000 depth of coverage is required in cases of heterogeneous variants in low tumor cellularity samples (1), and 100 depth with at least 10 variant reads and a LOD of 10% (10). According to the binomial data distribution, a coverage depth of 250 should indeed be sufficient to detect 5% VAF with a threshold of variant supporting reads  $\geq 5$  (Figure 1). On the other hand, NGS analysis with a coverage depth of 100 along with a requirement of at least 10 variant supporting reads as recommended by the ERIC consortium (10) would





**FIGURE 1 |** LOD as a function of coverage depth according to the binomial distribution. Coverage depth needed to maintain an intended LOD (within 3–20% VAF range) for three cumulative probability settings: for false positive probability of 0.001 and true positive of 0.999, a LOD of 20% is achieved at 61 coverage depth, a LOD of 10% at 175, a LOD of 5% at 562, and a LOD of 3% at 1,650. For the false positive probability of 0.010 and true positive of 0.990, a LOD of 20% is achieved at 31, a LOD of 10% at 81, a LOD of 5% at 288, and a LOD of 3% at 886 coverage depth, respectively. For the false positive probability of 0.050 and true positive of 0.950, a LOD of 20% is achieved at 30, a LOD of 10% at 30, a LOD of 5% at 124, and a LOD of 3% at 392 coverage depth, respectively.

result in a false negative of 45% for samples with a LOD of 10%. To confirm these theoretic calculations, we performed two independent dilution experiments to estimate the performance of *TP53* NGS analysis to detect 10% VAF at a depth of coverage of 100 reads. Indeed, we detected 30% of false negatives (5 positive samples of 7 true-positive samples and 9 positive samples of 13 true-positive samples) in two independent sequencing runs. Unfortunately, the false negative rate is often underestimated in targeted resequencing. Also, a recent study investigating inter-laboratory results of somatic variant detection with VAFs between 15 and 50% in 111 laboratories with reported LODs of 5–15% (6) shows that major errors in diagnostic NGS may arise from false negative results, even in samples with high mutation loads (6). Of three concurrent false positive results, all variants were correctly detected but mischaracterised (6). Since laboratories have not been asked to report coverage depth for other regions than the identified variants (6), we may only assume that low coverage or high variant calling thresholds contributed to the false negative results. These results further highlight the need for standardized coverage depth parameters in diagnostic NGS, taking into account sequencing errors as well as assay-specific errors.

## FREQUENCY OF *TP53* SUBCLONAL MUTATIONS IN CLL DETECTED THROUGH DIAGNOSTIC NGS

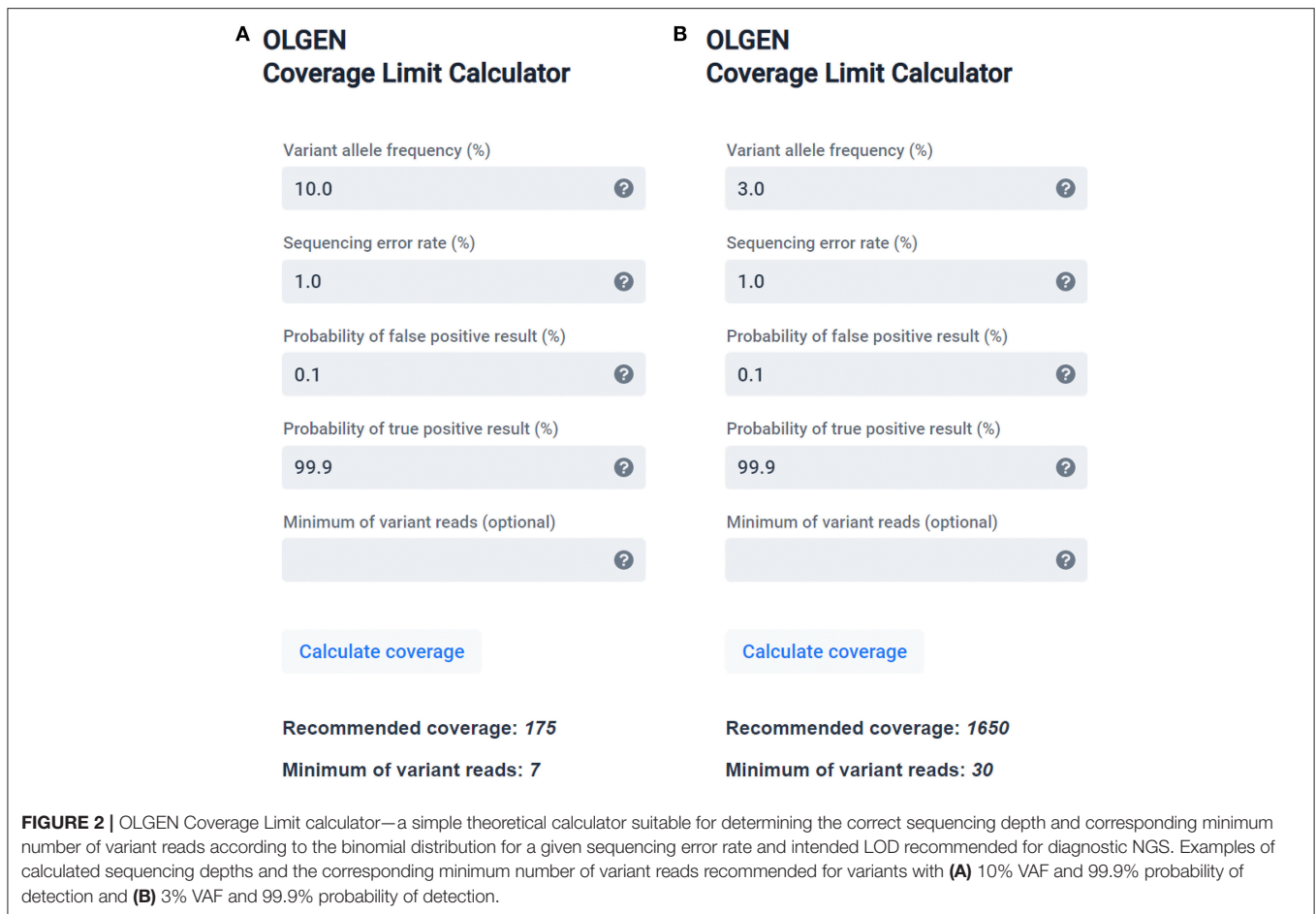
In order to evaluate the occurrence of low VAF in real-world settings, we reviewed our cohort of CLL patients examined for *TP53*mut in our diagnostic laboratory. The *TP53*mut were assessed as reported previously (14, 15). Briefly, *TP53* (exons 2–10 including 2 bp intronic overlap, 5' and 3'UTR) was analyzed using 100 ng gDNA per reaction. Amplicon-based libraries

were sequenced as paired-end on MiSeq (2x151, Illumina) with minimum target read depths of 5,000x. The LOD of *TP53*mut was set up to 1%, and the variants in the range 1–3% were confirmed by replication. Written informed consent was obtained from all the patients who were enrolled in accordance with the Helsinki Declaration, and the study was approved by the local ethical committee.

Of the diagnostic cohort of 859 CLL patients (April 2016–April 2019), 25% (215/859) were positive for *TP53*mut, and of those, 52.6% (113/215) carried variants with VAF at 10% or lower. In line with our observations, a recent study (8) reported the presence of 63 and 84% low burden (Sanger negative) *TP53*mut in CLL patients at the time of diagnosis and at the time of treatment, respectively, and confirmed the negative impact on the overall survival of *TP53*mut above 1% VAF at the time of treatment (8).

## CALCULATOR FOR DIAGNOSTIC NGS SETTINGS FOR DETECTION OF SUBCLONAL MUTATIONS

To assist laboratories with the determination of the minimum proper coverage parameters, we are providing a simple, user-friendly theoretical calculator (software) based on the binomial distribution (Figure 2), described in the Supplementary File. A web (or desktop) application and stand-alone source codes in R are accessible on Github: <https://github.com/mvasinek/olgen-coverage-limit>. Using this calculator, the correct parameters of sequencing depth and the corresponding minimum number of variant reads for a given sequencing error rate and intended LOD can easily be determined. Moreover, users can also take into account other errors by simply adding assay-specific errors to the sequencing error rate and using this overall error as an input



to the calculator. For example, in our case of *TP53* mutational analysis we calculated with the overall error of ~1.16%, thus we set up our minimum coverage depth requirements to 2,000 with threshold of minimum 40 reads for 3% VAF.

## DISCUSSION

Although diagnostic NGS has gained prominence in clinical settings for the assessment of somatic mutations in cancer, insufficient standardization of sequencing parameters still limits its implementation in clinical practice (1), mainly for variants present at low allele frequencies (4). We, therefore, addressed the technical question of correctly determining the sequencing depth in diagnostic NGS in order to obtain confident and reproducible detections of low VAF variants. In particular, we performed theoretical calculations to determine the optimum depth of coverage for the desired probability of detection of variants at low allele frequencies, taking into account the sequencing error rate. Moreover, we confirmed these theoretical calculations by conducting dilution experiments. Based on these observations, we recommend a depth of coverage of 1,650 or higher (together with the respective threshold of at least 30 mutated reads) to call  $\geq 3\%$  variants to achieve a 99.9% probability of variant detection, using the conventional NGS sequencing error only.

Variants in the 1–3% VAF range can only be called if the obtained sequence data is of high quality (average Q30 > 90%) and/or when the variants are confirmed by replication or the orthogonal method (1, 11, 16). We are also providing a simple, user-friendly theoretical calculator (software) to assist laboratories with resolving the correct sequencing depth and the corresponding minimum number of variant reads while taking into account the sequencing error rate. Our simple calculator may help to minimize the false positive and false negative results in diagnostic NGS.

Nevertheless, correct sequencing depth is also influenced by assay-specific factors (1). Errors can occur at many stages during DNA processing and library preparation. The most common are amplification errors introduced during NGS library preparation (1, 12, 17). Other common sources of errors have to do with library complexity (the number of independent DNA molecules analyzed), DNA quality, and target region complexity etc. All potential assay-specific errors should be addressed through test design, method validation, and quality control.

Currently, emerging error correction strategies, both computational and experimental, are being developed in order to mitigate the high error rates in diagnostic NGS (11). So far, among the most promising error correction methods are UMI (Unique Molecular Identifiers), which correct for

PCR errors (18), and signal-to-noise correction approaches (11). These advances attempt to reduce the LOD, thereby increasing sequencing accuracy needed for future opportunities in NGS diagnosis.

In order to improve the standardization in diagnostic NGS, the estimation of correct coverage depth is a recommended starting point when assessing thresholds surrounding a particular NGS assay. Nevertheless, there is still lack of published guidance regarding the minimum technical requirements and its reporting in NGS, particularly important in detection of clonal and subclonal mutations in cancer diagnostics. This is mainly due to the broad range of library preparation approaches, and numerous variables playing a role in each specific NGS assay, that are difficult to standardize, together with inter-laboratory variability. Therefore, the definition of minimum technical requirements and its reporting in NGS is highly desirable. Based on our experience in diagnostic NGS in haemato-oncology, we suggest to report at least following technical parameters: LOD, overall error of NGS assay (or at least sequencing error rate), the amount of DNA input, source, and quality of DNA, minimum coverage depth and the percentage of targeted bases sequenced at this minimum depth, total number of target reads covering variant region and number of reads supporting the variant. Special emphasis should be given to NGS standardization of the formalin-fixed paraffin-embedded (FFPE) samples (19, 20).

Taken together, our study highlights the importance of correct sequencing depth and the minimum number of reads required for reliable and reproducible detection of variants with low VAF in diagnostic NGS. The calculation of correct sequencing depth for a given error rate using our user-friendly theoretical calculator (software) may help to minimize the false positive and false negative results in diagnostic NGS, in situations related

to subclonal mutations among others. The rigorous testing and standardized minimum requirements for diagnostic NGS is particularly desirable to ensure correct results in clinical settings.

## DATA AVAILABILITY

The datasets generated for this study are available on reasonable request to the corresponding author.

## AUTHOR CONTRIBUTIONS

AP and EK designed the study, interpreted the results, and wrote the manuscript. AP, LS, TD, and PS performed NGS analysis. TP collected the patient samples and clinical data. MV performed bioinformatics analysis and wrote the calculator code. TN prepared web application. All authors read and approved the final version of manuscript.

## FUNDING

Grant support: MZ CR VES16-32339A, in part by the MH CZ—DRO (FNOI, 00098892).

## ACKNOWLEDGMENTS

We apologize to the many authors whose articles could not be cited because of reference limits.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.00851/full#supplementary-material>

## REFERENCES

- Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn.* (2017) 19:341–65. doi: 10.1016/j.jmoldx.2017.01.011
- D'Haene N, Le Mercier M, De Neve N, Blanchard O, Delaunoy M, El Housni H, et al. Clinical validation of targeted next generation sequencing for Colon and Lung cancers. *PLoS ONE.* (2015) 10:e0138245. doi: 10.1371/journal.pone.0138245
- Deans ZC, Costa JL, Cree I, Dequeker E, Edsjo A, Henderson S, et al. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN path ASBL. *Virchows Arch.* (2017) 470:5–20. doi: 10.1007/s00428-016-2025-7
- Ivanov M, Laktionov K, Breder V, Chernenko P, Novikova E, Telysheva E, et al. Towards standardization of next-generation sequencing of FFPE samples for clinical oncology: intrinsic obstacles and possible solutions. *J Transl Med.* (2017) 15:22. doi: 10.1186/s12967-017-1125-8
- Bacher U, Shumilov E, Flach J, Porret N, Joncourt R, Wiedemann G, et al. Challenges in the introduction of next-generation sequencing (NGS) for diagnostics of myeloid malignancies into clinical routine use. *Blood Cancer J.* (2018) 8:113. doi: 10.1038/s41408-018-0148-6
- Merker JD, Devereaux K, Iafrate AJ, Kamel-Reid S, Kim AS, Moncur JT, et al. Proficiency testing of standardized samples shows very high interlaboratory agreement for clinical next-generation sequencing-based oncology assays. *Arch Pathol Lab Med.* (2019) 143:463–71. doi: 10.5858/arpa.2018-0336-CP
- Rossi D, Khiabani H, Spina V, Ciardullo C, Brusca A, Fama R, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood.* (2014) 123:2139–47. doi: 10.1182/blood-2013-11-539726
- Brieghel C, Kinalis S, Yde CW, Schmidt AY, Jonson L, Andersen MA, et al. Deep targeted sequencing of TP53 in chronic lymphocytic leukemia: clinical impact at diagnosis and at time of treatment. *Haematologica.* (2019) 104:789–96. doi: 10.3324/haematol.2018.195818
- Campo E, Cymbalista F, Ghia P, Jager U, Pospisilova S, Rosenquist R, et al. TP53 aberrations in chronic lymphocytic leukemia: an overview of the clinical implications of improved diagnostics. *Haematologica.* (2018) 103:1956–68. doi: 10.3324/haematol.2018.187583
- Malcikova J, Tausch E, Rossi D, Sutton LA, Soussi T, Zenz T, et al. ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia-update on methodological approaches and results interpretation. *Leukemia.* (2018) 32:1070–80. doi: 10.1038/s41375-017-0007-7
- Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet.* (2018) 19:269–85. doi: 10.1038/nrg.2017.117
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent,



- Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*. (2012) 13:341. doi: 10.1186/1471-2164-13-341
13. Tan G, Opitz L, Schlapbach R, Rehrauer H. Long fragments achieve lower base quality in Illumina paired-end sequencing. *Sci Rep*. (2019) 9:2856. doi: 10.1038/s41598-019-39076-7
  14. Obr A, Prochazka V, Jirkuvova A, Urbankova H, Kriegova E, Schneiderova P, et al. TP53 mutation and complex karyotype portends a dismal prognosis in patients with mantle cell lymphoma. *Clin Lymphoma Myeloma Leuk*. (2018) 18:762–8. doi: 10.1016/j.clml.2018.07.282
  15. Turcsanyi P, Kriegova E, Kudelka M, Radvansky M, Kruzova L, Urbanova R, et al. Improving risk-stratification of patients with chronic lymphocytic leukemia using multivariate patient similarity networks. *Leuk Res*. (2019) 79:60–8. doi: 10.1016/j.leukres.2019.02.005
  16. Shin HT, Choi YL, Yun JW, Kim N, Kim SY, Jeon HJ, et al. Prevalence and detection of low-allele-fraction variants in clinical cancer samples. *Nat Commun*. (2017) 8:1377. doi: 10.1038/s41467-017-01470-y
  17. Ma X, Shao Y, Tian L, Flasch DA, Mulder HL, Edmonson MN, et al. Analysis of error profiles in deep next-generation sequencing data. *Genome Biol*. (2019) 20:50. doi: 10.1186/s13059-019-1659-6
  18. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res*. (2017) 27:491–9. doi: 10.1101/gr.209601.116
  19. McDonough SJ, Bhagwate A, Sun Z, Wang C, Zschunke M, Gorman JA, et al. Use of FFPE-derived DNA in next generation sequencing: DNA extraction methods. *PLoS ONE*. (2019) 14:e0211400. doi: 10.1371/journal.pone.0211400
  20. Ascierto PA, Bifulco C, Palmieri G, Peters S, Sidiropoulos N. Pre-analytic variables and tissue stewardship for reliable next-generation sequencing (NGS) clinical analysis. *J Mol Diagn*. (2019) 21:756–67. doi: 10.1016/j.jmoldx.2019.05.004

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Petrackova, Vasinek, Sedlarikova, Dyskova, Schneiderova, Novosad, Papajik and Kriegova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## APPENDIX E

**Petrackova A**, Horak P, Radvansky M, Skacelova M, Fillerova R, Kudelka M, Smrzova A, Mrazek F, Kriegova E. Cross-disease innate gene signature: emerging diversity and abundance in RA comparing to SLE and SSc. *J Immunol Res.* 2019;3575803. (IF 2019: 3.327)

## Research Article

# Cross-Disease Innate Gene Signature: Emerging Diversity and Abundance in RA Comparing to SLE and SSc

Anna Petrackova,<sup>1</sup> Pavel Horak,<sup>2</sup> Martin Radvansky,<sup>3</sup> Martina Skacelova,<sup>2</sup> Regina Fillerova,<sup>1</sup> Milos Kudelka,<sup>3</sup> Andrea Smrzova,<sup>2</sup> Frantisek Mrazek,<sup>1</sup> and Eva Kriegova <sup>1</sup>

<sup>1</sup>Department of Immunology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital, Olomouc, Czech Republic

<sup>2</sup>Department of Internal Medicine III-Nephrology, Rheumatology and Endocrinology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital, Olomouc, Czech Republic

<sup>3</sup>Faculty of Electrical Engineering and Computer Science, Department of Computer Science, VSB-Technical University of Ostrava, Ostrava, Czech Republic

Correspondence should be addressed to Eva Kriegova; [eva.kriegova@email.cz](mailto:eva.kriegova@email.cz)

Received 3 April 2019; Accepted 12 June 2019; Published 16 July 2019

Academic Editor: Dariusz Widera

Copyright © 2019 Anna Petrackova et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Overactivation of the innate immune system together with the impaired downstream pathway of type I interferon-responding genes is a hallmark of rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and systemic sclerosis (SSc). To date, limited data on the cross-disease innate gene signature exists among those diseases. We compared therefore an innate gene signature of Toll-like receptors (TLRs), seven key members of the interleukin (IL)1/IL1R family, and CXCL8/IL8 in peripheral blood mononuclear cells from well-defined patients with active stages of RA ( $n = 36$ , DAS28  $\geq 3.2$ ), SLE ( $n = 28$ , SLEDAI  $> 6$ ), and SSc ( $n = 22$ , revised EUSTAR index  $> 2.25$ ). Emerging diversity and abundance of the innate signature in RA patients were detected: RA was characterized by the upregulation of *TLR3*, *TLR5*, *IL1RAP/IL1R3*, *IL18R1*, and *SIGIRR/IL1R8* when compared to SSc ( $P_{\text{corr}} < 0.02$ ) and of *TLR2*, *TLR5*, and *SIGIRR/IL1R8* when compared to SLE ( $P_{\text{corr}} < 0.02$ ). Applying the association rule analysis, six rules (combinations and expression of genes describing disease) were identified for RA (most frequently included high *TLR3* and/or *IL1RAP/IL1R3*) and three rules for SLE (low *IL1RN* and *IL18R1*) and SSc (low *TLR5* and *IL18R1*). This first cross-disease study identified emerging heterogeneity in the innate signature of RA patients with many upregulated innate genes compared to that of SLE and SSc.

## 1. Introduction

Rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and systemic sclerosis (SSc) are systemic autoimmune diseases characterized by overactivation of the innate immune system together with impaired downstream pathway of type I interferon- (IFN-) responding genes (IFN signature). Nevertheless, a certain heterogeneity in the IFN signature among those diseases has been recognized, and some patients even lack its presence [1–4].

Although the emerging role of the innate immunity in the pathogenesis of RA, SLE, and SSc has been demonstrated, there is no data on the cross-disease innate gene signature as well as its heterogeneity among those diseases yet. Numerous studies on individual innate immunity members in RA, SLE, and SSc showed the crucial role of Toll-like receptors (TLRs) and IL1 family [5, 6]. Notable examples of common innate pathways are (i) the involvement of the adapter protein MyD88 which is required for signal transduction by TLRs and receptors of the IL1 family, (ii) the activation of

TABLE 1: Demographic and clinical characteristics of enrolled patients.

	RA ( <i>n</i> = 36)	SLE ( <i>n</i> = 28)	SSc ( <i>n</i> = 22)
Female/male	26/10	24/4	15/7
Age (years) mean (min-max)	57.5 (39-80)	40.1 (19-67)	58.0 (38-77)
Duration of the disease (years) mean (min-max)	18.1 (9-50)	10.0 (1-20)	5.4 (0-21)
Medications (% ( <i>n</i> ))			
Steroids	89 (32)	82 (23)	96 (21)
NSAIDs	78 (28)	14 (4)	0 (0)
Methotrexate	83 (30)	14 (4)	9 (2)
Other DMARDs*	36 (13)	100 (28)	73 (16)
Biologics	39 (14)	0 (0)	0 (0)
Relative white blood count (%)			
Lymphocytes (mean (95% CI))	24.9 (20.5-29.3)	22.9 (18.5-27.3)	21.4 (17.5-25.4)
Neutrophils (mean (95% CI))	62.9 (57.9-67.9)	67.1 (61.6-72.6)	67.3 (62.5-72.2)
Monocytes (mean (95% CI))	8.9 (7.9-9.9)	8.5 (7.1-9.9)	9.2 (7.9-10.4)

NSAIDs: nonsteroidal anti-inflammatory drugs; DMARDs: disease-modifying antirheumatic drugs; CI: confidence interval. \*Other DMARDs taken were hydroxychloroquine (RA/SLE/SSc; *n* = 3/26/0), leflunomide (8/0/0), sulfasalazine (2/0/0), azathioprine (0/8/12), mycophenolate mofetil (0/6/0), cyclophosphamide (0/3/3), and cyclosporine (0/1/1).

the type I IFN, and (iii) the presence of endogenous TLR ligands [7]. Besides shared innate pathways, disease-specific molecular and cellular mechanisms exist. In SLE, recent evidence has suggested a close relationship between the endosomal TLR activation and the disease onset [8, 9] with an essential role of endosomal TLRs in the generation of anti-nuclear antibodies and type I IFNs [10]. In RA, abundant activation of individual members of TLR and IL1 families was already evidenced with a proposed role for exogenous TLR ligands in the disease onset (i.e., *Proteus* infection of urinary tract, Epstein-Barr virus, and parvovirus B19) and for endogenous ligands in self-sustaining of the inflammatory loop [5, 11]. In SSc, signaling via TLR is increasingly recognized as a key player driving the persistent fibrotic response and is linked to the activity of TGF- $\beta$ ; however, the pathological role of TLRs and their ligands in SSc still remains unclear [12].

We undertook this study to elucidate the underlying differences in the innate immunity signature across three major autoimmune disorders using multivariate analysis. This first cross-disease analysis of the innate gene expression signature of 10 TLRs, 7 key members of the *IL1/IL1R* family, and interleukin 8 (*CXCL8*) in peripheral blood mononuclear cells (PBMC) from patients with active SLE, RA, and SSc revealed emerging diversity and abundance in RA compared to SLE and SSc. Our study contributes to further understanding of the innate signature underlying the immunopathology of major autoimmune diseases, with special emphases to discriminate shared and disease-specific expression patterns.

## 2. Materials and Methods

**2.1. Study Subjects.** The study cohort consisted of 86 Caucasian patients with autoimmune diseases from a single rheumatology center in Olomouc, Moravia region of Czech Republic. All enrolled RA/SLE/SSc patients met the 2010 ACR/EULAR classification criteria for RA [13], the ACR classification criteria for SLE [14], and the 2013 ACR/EULAR classification

criteria for SSc, respectively [15]. To exclude heterogeneity due to the activity and inactivity of the diseases, only cases with active phenotypes of the disease classified according to common activity scores (Disease Activity Score in 28 joints (DAS28), SLE Disease Activity Index (SLEDAI), and revised European Scleroderma Trials and Research group (EUSTAR) index) were included: RA (*n* = 36, DAS28  $\geq$  3.2), SLE (*n* = 28, SLEDAI > 6), and SSc (*n* = 22, revised EUSTAR index > 2.25).

The demographic and clinical features, used medication, duration of disease, and relative white blood count are described in Table 1. Distribution of lymphocyte, neutrophil, and monocyte counts did not differ between studied patient's groups ( $P > 0.05$ ). The healthy control cohort consisted of 77 subjects (mean age 51 yrs, min-max 24-90 yrs, female/male 58/19) out of which were formed three age-/gender-matched groups for each disease: 63 controls for RA (mean age 56 yrs, min-max 41-90 yrs, female/male 45/18), 33 controls for SLE (40, 24-50, 27/6, respectively), and 48 controls for SSc (58, 48-90, 34/14, respectively). In all healthy subjects, presence of inflammatory autoimmune diseases in first or second degree relatives, recent vaccination, infection, and usage of immunosuppressive drugs were excluded by questionnaire.

The patients and control subjects provided written informed consent about the usage of peripheral blood for the purpose of this study, which was approved by the ethics committee of the University Hospital and Palacký University Olomouc.

**2.2. Sample Processing and Real-Time Reverse Transcription-Polymerase Chain Reaction (qRT-PCR).** The PBMC were isolated from the peripheral blood collected in K<sub>3</sub>EDTA tubes by Ficoll density gradient centrifugation (Sigma-Aldrich, Germany) and stored in TRI Reagent (Sigma-Aldrich, Germany) at  $-80^{\circ}\text{C}$  until analysis. Total RNA was extracted using a Direct-zol RNA kit (Zymo Research, USA) according to the manufacturer's recommendations. After reverse transcription with a Transcriptor First Strand cDNA Synthesis

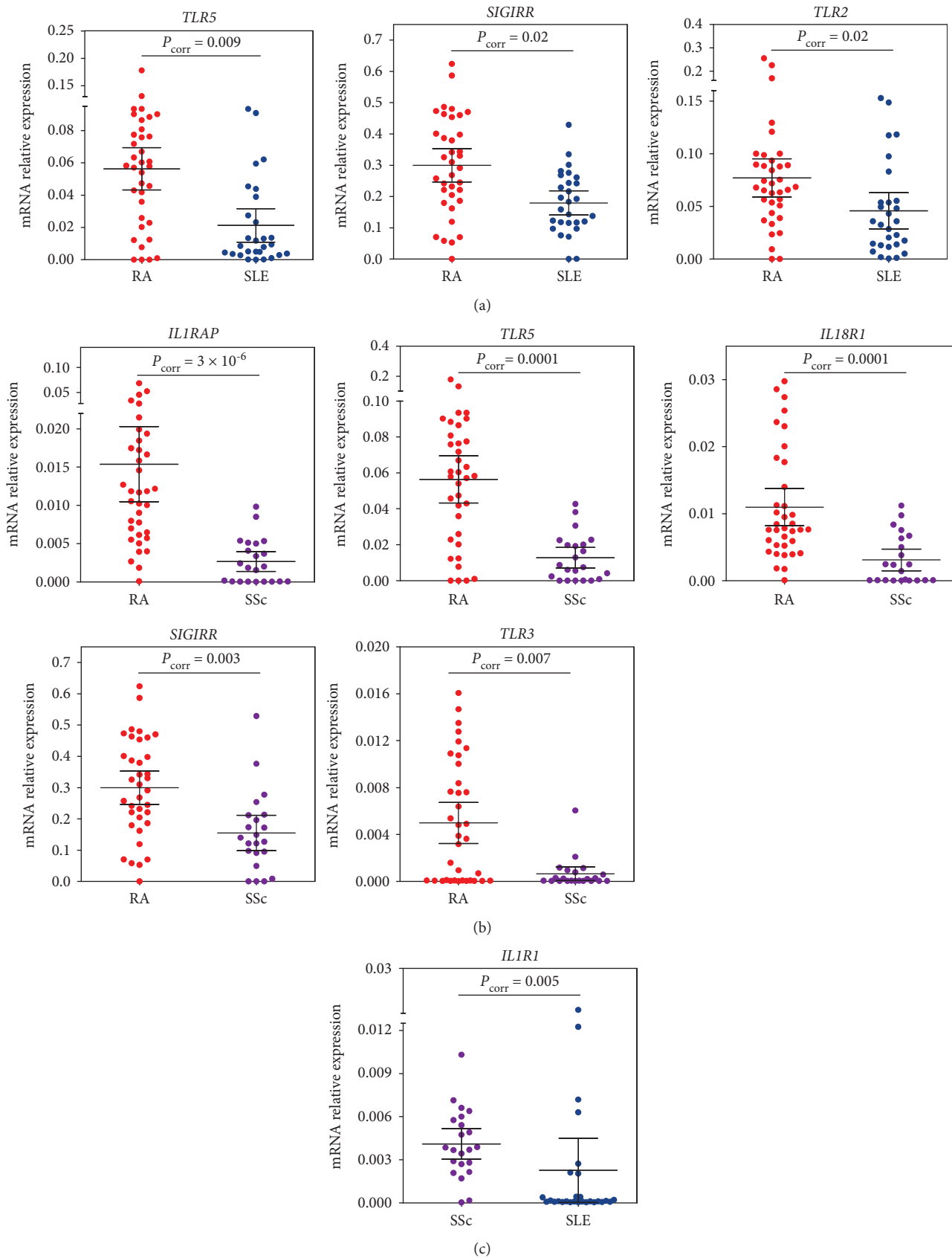


FIGURE 1: Relative mRNA expression levels of genes differentially expressed in (a) RA vs. SLE, (b) RA vs. SSc, and (c) SSc vs. SLE. Group means are indicated by horizontal bars; error bars indicate 95% CI.

TABLE 2: Relative mRNA expression levels of genes differentially expressed between (a) RA vs. SLE, (b) RA vs. SSc, (c) SSc vs. SLE.

(a) RA vs. SLE						
Gene	Mean (95% CI)		FC	P value	$P_{\text{corr}}$	
	RA	SLE				
<i>TLR5</i>	0.056 (0.043-0.070)	0.021 (0.011-0.032)	6.49	$5.2 \times 10^{-4}$	$9.3 \times 10^{-3}$	
<i>SIGIRR</i>	0.300 (0.247-0.353)	0.179 (0.141-0.218)	1.76	$2.0 \times 10^{-3}$	$2.0 \times 10^{-2}$	
<i>TLR2</i>	0.077 (0.059-0.095)	0.046 (0.029-0.063)	2.00	$3.7 \times 10^{-3}$	$2.2 \times 10^{-2}$	

(b) RA vs. SSc						
Gene	Mean (95% CI)		FC	P value	$P_{\text{corr}}$	
	RA	SSc				
<i>IL1RAP</i>	0.015 (0.011-0.020)	0.003 (0.001-0.004)	6.08	$1.7 \times 10^{-7}$	$3.0 \times 10^{-6}$	
<i>TLR5</i>	0.056 (0.043-0.070)	0.013 (0.007-0.019)	7.16	$1.1 \times 10^{-5}$	$9.8 \times 10^{-5}$	
<i>IL18R1</i>	0.011 (0.008-0.014)	0.003 (0.002-0.005)	4.08	$2.0 \times 10^{-5}$	$1.2 \times 10^{-4}$	
<i>SIGIRR</i>	0.300 (0.247-0.353)	0.155 (0.098-0.211)	2.26	$5.9 \times 10^{-4}$	$2.6 \times 10^{-3}$	
<i>TLR3</i>	0.005 (0.003-0.007)	0.001 ( $6.1 \times 10^{-5}$ -0.001)	28.5	$1.8 \times 10^{-3}$	$6.6 \times 10^{-3}$	

(c) SSc vs. SLE						
Gene	Mean (95% CI)		FC	P value	$P_{\text{corr}}$	
	SSc	SLE				
<i>IL1R1</i>	0.004 (0.003-0.005)	0.002 ( $3.1 \times 10^{-5}$ -0.004)	34.8	$2.7 \times 10^{-4}$	$4.8 \times 10^{-3}$	

$P_{\text{corr}}$  value corrected for multiple comparisons (Benjamini-Hochberg correction). FC (fold change) between group medians of relative mRNA expression levels.

Kit (Roche, Switzerland), qPCR was performed in a 100 nl reaction volume containing a LightCycler 480 SYBR Green I Master mix (Roche, Switzerland) using a high-throughput SmartChip Real-Time-qPCR System (WaferGen, USA) as reported previously [16, 17]. Primer sequences are listed in Table S1 (Integrated DNA Technologies, USA). The relative mRNA expression was calculated using phosphoglycerate kinase 1 as a reference gene [18].

In order to assess the innate immunity gene expression pattern, the expression of *TLR* (*TLR1-10*), *IL1/IL1R* family (21 members), and *CXCL8* was investigated in PBMC. Based on pilot evaluation of qPCR assays on a cohort of 20 RA, 20 SLE, and 20 SSc patients, 14 assays of *IL1/IL1R* family members (*IL1A*, *IL36RN*, *IL36A*, *IL36B*, *IL36G*, *IL37*, *IL38*, *IL33*, *IL1R2*, *IL18RAP*, *IL1RL1*, *IL1RL2*, *IL1RAPL1*, and *IL1RAPL2*) were below the limit of detection of the system and thus excluded from further analysis. The study continued therefore by expression profiling of 18 innate immunity genes: *TLR1-10*, 7 members of the *IL1/IL1R* family together with *CXCL8*.

**2.3. Statistical Analysis and Data-Mining Methods.** Statistical analysis (Mann-Whitney *U* test, Benjamini-Hochberg correction) of relative gene expression values was performed using Genex (MultiD Analyses AB, Sweden) and GraphPad Prism 5.01 (GraphPad Software, USA). *P* value < 0.05 was considered as significant.

In this study, a set of multivariate data-mining analyses to visualize and characterize the gene expression heterogeneity

between and within the diseases was applied. For a flowchart of the analysis process used, see Figure S1.

First, correlation networks using the LRNet algorithm [19] and Spearman's rank correlation coefficient were constructed and visualized to investigate the relationships between expressions of individual studied genes within the innate gene signature and to nominate the most representative molecules for the particular disease.

Second, Andrews curve analysis was applied for visualization of the structure in multidimensional expression data [20–23]. The relative gene expression values of individual patients were transformed using Andrews' formula (Equation S1); all calculations were performed by package Andrews from the R library [24]. The Andrews curves were plotted to visualize the differences between particular diseases using a set of significantly deregulated genes and the whole set of studied genes. The difference is demonstrated by separation of the Andrews curve's amplitudes and phase shift [20, 22, 23]. The curves of similar relative gene expression overlap between studied groups (Figure S2), while separation of curves demonstrates the differences in expression profiles (Figure S3) [20, 22, 23]. More detailed description of the Andrews curve analysis is stated in Supplementary File.

Third, we applied association rule mining, a technique for finding frequent patterns, correlations, or associations among the given data set [25] to investigate the heterogeneity within the diseases themselves. Firstly, each gene data set was divided into low/high expression groups by arithmetic means of relative gene expressions within the whole data

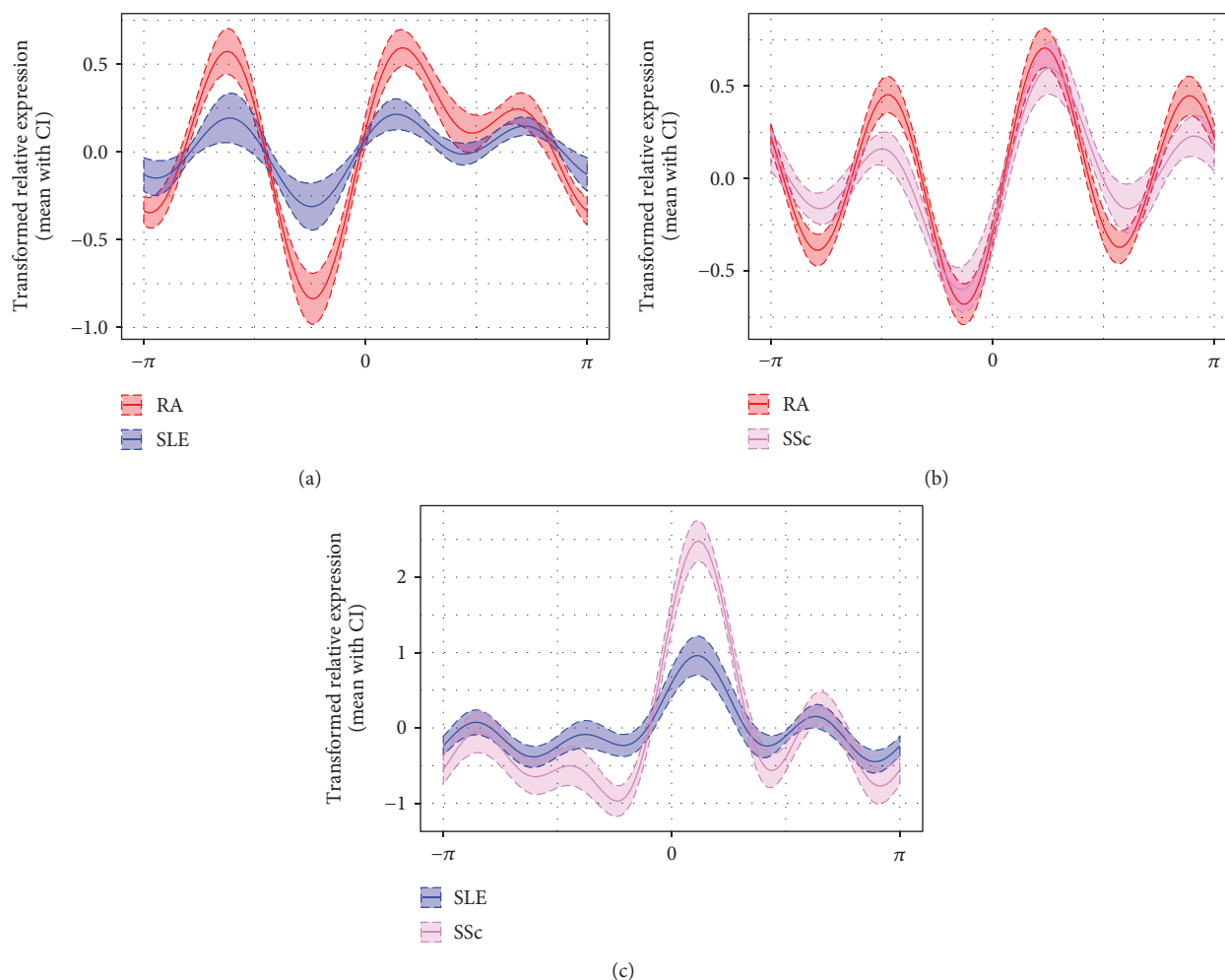


FIGURE 2: Differential innate gene expression analysis by Andrews curves between (a) RA vs. SLE, (b) RA vs. SSc, and (c) SLE vs. SSc—representative examples. The Andrews curves were calculated for various combinations of gene expression values from the whole set of studied genes. Examples show the results of the Andrews curve analysis for the combination of (a) *TLR3*, *TLR7*, *TLR8*, *IL1R1*, *IL1RN*, and *IL18R1*; (b) *TLR3*, *TLR4*, *TLR6*, *TLR10*, *IL1B*, *IL1R1*, and *SIGIRR*; and (c) *TLR4*, *TLR6*, *TLR7*, *TLR8*, *IL1R1*, *IL1RN*, and *IL18*. For those sets of genes, a good separation of diseases was observed as visualized by separation of the curve's amplitudes and phase shift. An example of combination of genes which does not discriminate between disease groups is shown in Figure S2. Full lines represent the mean values, the dashed lines 95% confidence intervals.

set. The applied package “arules” in the R system [26] was used to extract rules (combinations of genes and its expression levels associated with the particular disease). Only a minimum number of top ranked rules describing the particular disease with a good confidence (threshold 0.75) and support were used.

### 3. Results

**3.1. Innate Immune Gene Expression Pattern of RA, SLE, and SSc.** In order to characterize innate immune signature in studied diseases, the expression profiles of selected innate immune genes between patients and healthy controls in all diseases were compared.

To exclude the influence of age on the gene expression, the healthy controls were subdivided into age-matched subgroups despite no differences being observed in the expression profile of all investigated genes in the formed

subgroups ( $P_{\text{corr}} > 0.05$ ). RA differed from controls by the upregulated expression of *TLR2*, *TLR3*, *TLR5*, *TLR8*, *IL1B*, *IL18*, *IL18R1*, *IL1RN*, *IL1RAP*, and *SIGIRR/IL1R8* ( $P_{\text{corr}} \leq 0.05$ ; Figure S4A, Table S2A). In patients treated with anti-TNF- $\alpha$  therapy, a trend to lower *TLR5* levels in our RA patients was observed ( $P = 0.07$ ). In SLE, downregulation of *TLR10* was observed when compared to healthy controls ( $P = 0.02$ ); however, it did not reach significance after the correction for multiple comparisons (Figure S4B, Table S2B). SSc differed from controls by the upregulated expression of *IL1RN*, *IL18*, and *CXCL8* and downregulated expression of *IL1RAP* and *IL18R1* ( $P_{\text{corr}} \leq 0.05$ ; Figure S4C, Table S2C).

**3.2. Cross-Disease Analysis of Innate Pattern in RA, SLE, and SSc.** To investigate the disease-specific innate immune gene expression pattern, we compared RA, SLE, and SSc patients to each other. RA differed from SLE and SSc by the



	RA						SLE			SSc		
	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 1	Rule 2	Rule 3	Rule 1	Rule 2	Rule 3
TLR1												
TLR3	Dark	Dark	Dark					Light			Light	
TLR4				Light								
TLR5				Dark			Light			Light	Light	Light
TLR6						Light	Dark					
TLR8				Dark				Light				
TLR9			Light									
TLR10	Dark											
IL1B								Dark				
IL1R1					Dark		Light					Dark
IL1RN							Light	Light		Dark		
IL1RAP	Dark			Dark								
IL18				Dark			Light			Dark		
IL18R1				Light			Light	Light		Light	Dark	Light
SIGIRR						Dark						
CXCL8						Light						

FIGURE 3: Association rules describing RA, SLE, and SSc. Association rule analysis revealed a minimum of six rules for RA, three rules for SLE, and three rules for SSc, able to discriminate among all studied diseases with the accuracy above 77%. Columns represent individual rules (combinations of genes and its expression levels characterizing the particular disease). Dark/light color means high/low gene expression levels (cut-off: mean gene expression of the whole data set).

upregulated expression of *TLR5* and *SIGIRR* ( $P_{\text{corr}} < 0.02$ ; Figures 1(a) and 1(b), Tables 2(a) and 2(b), and Tables S3A and S3B). RA further differed from SLE by the upregulated expression of *TLR2* ( $P_{\text{corr}} = 0.02$ ; Figure 1(a), Tables 2(a) and S3A) and from SSc by the upregulation of *TLR3*, *IL1RAP*, and *IL18R1* genes ( $P_{\text{corr}} < 0.007$ ; Figure 1(b), Tables 2(b) and S3B). In SSc, the upregulated expression of *IL1R1* ( $P_{\text{corr}} = 0.005$ ; Figure 1(c), Tables 2(c) and S3C) was observed when compared to SLE.

**3.3. Visualization of Disease-Associated Gene Expression Pattern by Andrews Curves.** To investigate the disease-associated gene expression pattern, Andrews curves were used to visualize the differences between particular diseases using a set of significantly deregulated genes and the whole set of studied genes. First, we assessed the differences in the innate expression pattern of genes revealed by classical statistics. Although a good separation of Andrews curves on the basis of significant genes was observed (Figure S3), better separation of the studied diseases was obtained when a whole set of studied genes was used (Figure 2).

**3.4. Innate Pattern Characteristics of RA, SLE, and SSc.** Next, we applied the association rule analysis to identify rules (set of genes including their expression levels) describing a certain disease within the three studied diseases. Based on the

results from the Andrews curves, association rule analysis was performed using the whole gene set.

For RA, six rules were identified, thus showing high heterogeneity within this group of patients when compared to SLE and SSc (Figure 3), where for each of them, three rules were identified. In RA, a high level of *TLR3* and *IL1RAP* mRNA was identified in three and two rules, respectively. In SLE, low expression levels of *IL1RN* and *IL18R1* appeared in two rules, and in SSc, a low level of *TLR5* and *IL18R1* mRNA occurred in three and two rules, respectively. The obtained association rules and their support and confidence values deciphered for RA, SLE, and SSc patients are listed in Table 3. The accuracy of classification by using these rules for RA, SLE, and SSc was 83%, 78%, and 77%, respectively. Comparison of rules for each disease revealed that *TLR3*, *TLR5*, *IL18*, *IL18R1*, and *IL1R1* genes occurred in rules for all studied diseases, showing good discriminant power among studied autoimmune diseases as visualized by the Andrews curves (Figure S5).

## 4. Discussion

This study focused on the innate immunity gene signature among major autoimmune diseases: RA, SLE, and SSc, showing heterogeneity in the innate signature among and within these diseases. This first cross-disease study showed the highest diversity and abundance in the innate signature in RA when compared to SLE and SSc.

Innate immunity plays a key role in the pathogenesis of autoimmune rheumatic diseases as evidenced from numerous studies on individual members of innate immunity pathways [5, 6]. However, little is known about the similarities and differences in the innate signature at the molecular level between and within these diseases. Therefore, we investigated the differential expression of key innate genes in RA, SLE, and SSc. Importantly, our study was restricted only to the cases with active disease in order to exclude heterogeneity due to the activity and inactivity of the diseases. To obtain a more complex picture, the multivariate analysis was applied to assess the complexity of the differential innate signature having an advantage over classical statistical approaches due to taking into account the intrinsic characteristics of gene expression data and assessing the relationships between studied molecules.

Firstly, we applied Andrews curve analysis for assessment of differences and similarities in the gene innate signature between studied diseases, an approach particularly useful for visualization of the structure in multidimensional data [20, 21]. When using combination of genes reaching statistical significance as well as using the whole gene set, we confirmed the diversity among innate profiles in RA, SLE, and SSc by Andrews curve analysis. Upregulated expression of *TLR3*, *TLR5*, and *SIGIRR* was characteristic for RA when compared to both SLE and SSc. An intracellular receptor *TLR3* recognizing dsRNA has been shown to be involved in the RA pathogenesis: necrotic synovial fluid cells release RNA that can activate *TLR3* in RA synovial fibroblasts [27]. *TLR5*, a surface receptor highly upregulated in our RA patients, recognizes bacterial



TABLE 3: Association rules identified for (a) RA, (b) SLE, and (c) SSc.

No.	Rule	Support	Confidence	Number of patients identified
(a) RA				
1	<i>TLR3</i> high & <i>IL1RAP</i> high	0.13	1.00	11
2	<i>TLR3</i> high & <i>TLR10</i> high	0.12	1.00	10
3	<i>TLR3</i> high & <i>TLR9</i> low	0.12	1.00	10
4	<i>TLR4</i> low & <i>TLR8</i> high & <i>IL1RAP</i> high	0.14	1.00	12
5	<i>TLR5</i> high & <i>IL18</i> high & <i>IL18R1</i> low	0.14	1.00	12
6	<i>TLR6</i> low & <i>IL1R1</i> high & <i>SIGIRR</i> high & <i>CXCL8</i> low	0.12	0.91	10
(b) SLE				
1	<i>TLR5</i> low & <i>TLR6</i> high & <i>IL1RN</i> low & <i>IL18R1</i> low	0.10	0.90	9
2	<i>TLR1</i> low & <i>TLR8</i> low & <i>IL1R1</i> low & <i>IL1RN</i> low & <i>IL18</i> low & <i>IL18R1</i> low	0.13	0.85	11
3	<i>TLR3</i> low & <i>IL1B</i> high	0.10	0.82	9
(c) SSc				
1	<i>TLR5</i> low & <i>IL1RN</i> high & <i>IL18R1</i> low	0.10	1.00	9
2	<i>TLR5</i> low & <i>TLR3</i> low & <i>IL18</i> high	0.10	0.82	9
3	<i>TLR5</i> low & <i>IL1R1</i> high & <i>IL18R1</i> low	0.10	0.75	9

The data set for each gene was divided into low/high expression by means of a particular gene expression of the whole data set.

flagellin. However, their endogenous ligand(s) in synovial fluid able to activate TLR5 in RA is(are) still unknown [28, 29]. In line with our results, increased TLR5 in peripheral blood myeloid cells correlated with RA disease activity and TNF- $\alpha$  levels [30]. There is also evidence that anti-TNF- $\alpha$  therapy markedly suppress TLR5 expression in RA monocytes [31], a trend which was also observed in our study. Also, the next highly upregulated SIGIRR (*IL1R8/TIR8*), an orphan receptor required for the anti-inflammatory effects of IL37, has been reported in RA synovial tissue previously [32].

Also, other genes such as *TLR2*, *IL1RAP*, and *IL18R1* from the differential innate signature associated with RA revealed by our analysis were reported in autoimmune conditions previously. In line with our results, abundant TLR2 on monocyte subsets in active RA produced a spectrum of pro-inflammatory cytokines after stimulation [33]. TLR2 recognizes a wide range of conserved microbial products, probably due to its cooperation with TLR1 or TLR6, as well as its hypothetic ligand HMGB1 released from dying and activated cells [34]. Regarding *IL1RAP* and *IL18R1*, their upregulated expression in RA was reported recently [16] and their downregulation in SSc we report here for the first time. Finally, SSc was characterized by an increase in *IL1R1* in comparison to SLE. The first evidence about critical involvement of *IL1R1*, an essential mediator for proinflammatory IL1 signaling [35], in fibrotic processes has been already reported in a murine lung injury model [36]. Importantly, data from our cross-disease analysis are in line with previous studies on individual innate members and basic statistical analysis and further highlight the activation of innate immunity in RA when compared to SLE and SSc. The infectious agents and endogenous ligands activating innate receptors leading to a self-sustaining inflammatory loop responsible for chronic and destructive progression in RA need to be further elucidated.

Next, we investigated the differential innate signature among and within the studied diseases by association rule analysis, a method commonly used to uncover the most frequently purchased combinations of items in a market basket analysis. It has been shown that this analysis is highly convenient for gene expression datasets [37, 38] and gives additional information due to preservation of the causality between the gene expression level and phenotype [37]. For RA, six rules were identified, thus showing high heterogeneity within this group of patients when compared to SLE and SSc, where three rules were identified for each of them. In RA, the association rules most frequently included high expression of *TLR3* and/or *IL1RAP/IL1R3*, thus again highlighting activation of the innate system in active RA. In SLE, a low expression of *IL1RN* and *IL18R1* and in SSc, a low level of *TLR5* and *IL18R1* occurred often in the rules. Applying association rules (combinations of genes describing a particular disease), excellent confidence and accuracy above 77% was achieved for all investigated diseases.

Interestingly, about half of the patients in each disease were characterized by multiple rules, while others were typical by only one gene expression pattern rule. The existence of several innate profile subgroups within RA patients lets us suggest that the heterogeneity in the innate pattern in RA may contribute to various clinical disease manifestations [4, 16], thus deserving future investigation. We also hypothesize that observed heterogeneity in the innate signature may contribute to the heterogeneity in the IFN signature recently reported in RA [4]. Our data further highlighted the application of advanced multivariate data analysis especially for diseases such as SLE, where many clinical phenotypes exist. This may be reflected in the high variability in the expression pattern which might be underestimated by univariate statistics, especially in the case of low abundant genes. Finally, our data points out the involvement of various key innate molecules as

well as the different interplay between individual innate receptors in the studied diseases.

To gain a more complete picture of the innate signature in autoimmune diseases, we report also the differential profile of the innate signature in studied diseases compared to healthy controls. This comparison revealed the upregulation of four members of TLR (*TLR2*, *TLR3*, *TLR5*, and *TLR8*) and six members of the IL1/IL1R family (*IL1B*, *IL1RN*, *IL1RAP*, *IL18R1*, *IL18*, and *SIGIRR*) in RA when compared to healthy controls. In line with our results, deregulation of these genes or their protein products was already registered in RA [16, 30, 32, 39–44]. In SLE, this study showed for the first time downregulation of *TLR10*, a broad negative regulator of TLR signaling [45, 46]. The first evidence about the possible involvement of *TLR10* in autoimmunity has been already observed: downregulated *TLR10* expression was reported in PBMC of patients with microscopic polyangiitis [47] as well as RA patients with active disease [16]. In contrast to the murine models of SLE [48], we did not observe increased *TLR7* and *TLR9* expression in our SLE patients. In SSc, our study revealed upregulation of *IL1RN*, *IL18*, and *CXCL8* and downregulation of *IL1RAP* and *IL18R1*. In line with our results, upregulated *IL1RN* mRNA [49], increased IL18 expression in skin biopsies [50], and elevated serum IL8 in patients with scleroderma [51] were reported. Here, we report for the first time downregulation of *IL1RAP* and *IL18R1* in SSc. *IL1RAP* (*IL1R3*) is a coreceptor of *IL1R1* and is indispensable for the transmission of *IL1* signaling [35]. Regarding *IL18R1*, it encodes the  $\alpha$  subunit of the *IL18* receptor responsible for *IL18* binding. The activated receptor then initiates the same signaling pathway as *IL1* to activate *NF- $\kappa$ B* [52]. How these proteins contribute to the SSc pathogenesis deserves future investigations.

The authors are aware of some limitations. The study was performed as a cross-sectional analysis in a real-world setting of patients in different stages of the disease; however, the authors restricted analysis only to patients in the active disease stage in order to obtain a more homogenous cohort. Due to the small number of patients in the subgroups with particular gene patterns revealed by association analysis, the subanalysis of their association with clinical parameters was not performed. Future studies on larger cohorts with well-defined patients would be advisable to further confirm our results.

## 5. Conclusions

To conclude, this first cross-disease study highlighted the heterogeneous nature among and within RA, SLE, and SSc, with the identification of RA having the highest diversity and abundance in the innate signature when compared to SLE and SSc. Moreover, the results from applied data mining approaches show the importance of a multiple multivariate analysis for better understanding of relationships between individual molecules, especially in highly heterogeneous diseases.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Consent

All subjects have given their written informed consent.

## Conflicts of Interest

The authors have no conflicts of interest to declare.

## Authors' Contributions

A.P. and R.F. performed the measurements. M.Sk., A.S., and P.H. provided the samples and clinical data. A.P., M.R., and M.K. analyzed the data. A.P. and E.K. designed the study and wrote the paper. P.H. and F.M. helped with the discussion of the data and revising the paper. E.K. supervised the study. All authors were involved in reviewing the paper and had final approval of the submitted and published versions.

## Acknowledgments

This work was supported by the Grant Agency of the Czech Republic (Ministry of Health) (MZ CR VES15-28659A) and in part by MH CZ–DRO (FNOL, 00098892), IGA\_LF\_2019\_006, and IGA\_LF\_2019\_014.

## Supplementary Materials

Methods: description of Andrews curves analysis. Correlation analysis of gene patterns in RA, SLE, and SSc. Table S1: investigated genes and primers used for qRT-PCR. Table S2: relative mRNA expression levels of genes differentially expressed between (A) RA vs. healthy controls, (B) SLE vs. healthy controls, and (C) SSc vs. healthy controls. Table S3: relative mRNA expression levels of genes differentially expressed between (A) RA vs. SLE, (B) RA vs. SSc, and (C) SSc vs. SLE. Figure S1: algorithm flow chart of statistics and advanced data-mining methods used in this study. Figure S2: Andrews curve analysis using a gene expression data which does not discriminate between diseases—representative example. Figure S3: Andrews curves using a set of significantly deregulated genes between (A) RA vs. SLE, (B) RA vs. SSc, and (C) SLE vs. SSc. Figure S4: relative mRNA expression levels of genes differentially expressed in (A) RA vs. healthy controls, (B) SLE vs. healthy controls, and (C) SSc vs. healthy controls. Figure S5: Andrews curves using a set of genes revealed by association rules for discrimination of RA, SLE, and SSc. Figure S6: correlation network for studied innate genes in (A) RA, (B) SLE, and (C) SSc. Equation S1: formula used for computing and plotting of Andrews curves. (*Supplementary Materials*)

## References

- [1] F. A. H. Cooles, A. E. Anderson, D. W. Lendrem et al., “The interferon gene signature is increased in patients with early treatment-naive rheumatoid arthritis and predicts a poorer response to initial therapy,” *The Journal of Allergy and Clinical Immunology*, vol. 141, no. 1, pp. 445–448.e4, 2018.

- [2] P. Laurent, V. Sisirak, E. Lazaro et al., "Innate immunity in systemic sclerosis fibrosis: recent advances," *Frontiers in Immunology*, vol. 9, p. 1702, 2018.
- [3] K. M. Pollard, G. M. Escalante, H. Huang et al., "Induction of systemic autoimmunity by a xenobiotic requires endosomal TLR trafficking and signaling from the late endosome and endolysosome but not type I IFN," *Journal of Immunology*, vol. 199, no. 11, pp. 3739–3747, 2017.
- [4] J. Rodríguez-Carrio, M. Alperi-López, P. López, F. J. Ballina-García, and A. Suárez, "Heterogeneity of the type I interferon signature in rheumatoid arthritis: a potential limitation for its use as a clinical biomarker," *Frontiers in Immunology*, vol. 8, p. 2007, 2018.
- [5] L. A. B. Joosten, S. Abdollahi-Roodsaz, C. A. Dinarello, L. O'Neill, and M. G. Netea, "Toll-like receptors and chronic inflammation in rheumatic diseases: new developments," *Nature Reviews Rheumatology*, vol. 12, no. 6, pp. 344–357, 2016.
- [6] C. A. Dinarello, "Overview of the IL-1 family in innate inflammation and acquired immunity," *Immunological Reviews*, vol. 281, no. 1, pp. 8–27, 2018.
- [7] K. C. M. Santegoets, L. van Bon, W. B. van den Berg, M. H. Wenink, and T. R. D. J. Radstake, "Toll-like receptors in rheumatic diseases: are we paying a high price for our defense against bugs?," *FEBS Letters*, vol. 585, no. 23, pp. 3660–3666, 2011.
- [8] T. Celhar and A. M. Fairhurst, "Toll-like receptors in systemic lupus erythematosus: potential for personalized treatment," *Frontiers in Pharmacology*, vol. 5, p. 265, 2014.
- [9] Y. W. Wu, W. Tang, and J. P. Zuo, "Toll-like receptors: potential targets for lupus treatment," *Acta Pharmacologica Sinica*, vol. 36, no. 12, pp. 1395–1407, 2015.
- [10] R. M. Clancy, A. J. Markham, and J. P. Buyon, "Endosomal Toll-like receptors in clinically overt and silent autoimmunity," *Immunological Reviews*, vol. 269, no. 1, pp. 76–84, 2016.
- [11] E. Pretorius, O. O. Akeredolu, P. Soma, and D. B. Kell, "Major involvement of bacterial components in rheumatoid arthritis and its accompanying oxidative stress, systemic inflammation and hypercoagulability," *Experimental Biology and Medicine*, vol. 242, no. 4, pp. 355–373, 2017.
- [12] S. Bhattacharyya and J. Varga, "Emerging roles of innate immune signaling and toll-like receptors in fibrosis and systemic sclerosis," *Current Rheumatology Reports*, vol. 17, no. 1, p. 474, 2015.
- [13] D. Aletaha, T. Neogi, A. J. Silman et al., "2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative," *Annals of the Rheumatic Diseases*, vol. 69, no. 9, pp. 1580–1588, 2010.
- [14] M. C. Hochberg, "Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus," *Arthritis and Rheumatism*, vol. 40, no. 9, p. 1725, 1997.
- [15] F. van den Hoogen, D. Khanna, J. Franssen et al., "2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative," *Annals of the Rheumatic Diseases*, vol. 72, no. 11, pp. 1747–1755, 2013.
- [16] A. Petrackova, P. Horak, M. Radvansky et al., "Revealed heterogeneity in rheumatoid arthritis based on multivariate innate signature analysis," *Clinical and Experimental Rheumatology*, vol. 37, 2019(in press).
- [17] T. Tomankova, E. Kriegová, R. Fillerová, P. Luzná, J. Ehrmann, and J. Gallo, "Comparison of periprosthetic tissues in knee and hip joints: differential expression of CCL3 and DC-STAMP in total knee and hip arthroplasty and similar cytokine profiles in primary knee and hip osteoarthritis," *Osteoarthritis and Cartilage*, vol. 22, no. 11, pp. 1851–1860, 2014.
- [18] V. R. Falkenberg, T. Whistler, J. R. Murray, E. R. Unger, and M. S. Rajeevan, "Identification of phosphoglycerate kinase 1 (PGK1) as a reference gene for quantitative gene expression measurements in human blood RNA," *BMC Research Notes*, vol. 4, no. 1, p. 324, 2011.
- [19] E. Ochodkova, S. Zehnalova, and M. Kudelka, "Graph construction based on local representativeness," in *Computing and Combinatorics. COCOON 2017. Lecture Notes in Computer Science*, vol. 10392, Y. Cao and J. Chen, Eds., pp. 654–665, Springer, Cham, 2017.
- [20] R. E. Moustafa, "Andrews curves," *WIREs Computational Statistics*, vol. 3, no. 4, pp. 373–382, 2011.
- [21] P. Niedzielski, M. Mleczek, A. Budka et al., "A screening study of elemental composition in 12 marketable mushroom species accessible in Poland," *European Food Research and Technology*, vol. 243, no. 10, pp. 1759–1771, 2017.
- [22] D. F. Andrews, "Plots of high-dimensional data," *Biometrics*, vol. 28, no. 1, pp. 125–136, 1972.
- [23] C. García-Osorio and C. Fyfe, "Visualization of high-dimensional data via orthogonal curves," *Journal of Universal Computer Science*, vol. 11, pp. 1806–1819, 2005.
- [24] J. Myslivec, "andrews: Andrews curves. R package version 1.0," 2012, <http://CRAN.R-project.org/package=andrews>.
- [25] R. Agrawal, T. Imielinski, and A. N. Swami, "Mining association rules between sets of items in large databases," in *SIGMOD '93 Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, pp. 207–216, Washington, DC, USA, May 1993.
- [26] M. Hahsler, C. Buchta, B. Gruen, and K. Hornik, "arules: Mining Association Rules and Frequent Itemsets. R package version 1.6-1," 2018, <https://CRAN.R-project.org/package=arules>.
- [27] F. Brentano, O. Schorr, R. E. Gay, S. Gay, and D. Kyburz, "RNA released from necrotic synovial fluid cells activates rheumatoid arthritis synovial fibroblasts via Toll-like receptor 3," *Arthritis and Rheumatism*, vol. 52, no. 9, pp. 2656–2665, 2005.
- [28] S. J. Kim, Z. Chen, N. D. Chamberlain et al., "Angiogenesis in rheumatoid arthritis is fostered directly by Toll-like receptor 5 ligation and indirectly through interleukin-17 induction," *Arthritis and Rheumatism*, vol. 65, no. 8, pp. 2024–2036, 2013.
- [29] H. A. Elshabrawy, A. E. Essani, Z. Szekanecz, D. A. Fox, and S. Shahrara, "TLRs, future potential therapeutic targets for RA," *Autoimmunity Reviews*, vol. 16, no. 2, pp. 103–113, 2017.
- [30] N. D. Chamberlain, O. M. Vila, M. V. Volin et al., "TLR5, a novel and unidentified inflammatory mediator in rheumatoid arthritis that correlates with disease activity score and joint TNF- $\alpha$  levels," *Journal of Immunology*, vol. 189, no. 1, pp. 475–483, 2012.
- [31] S. Kim, Z. Chen, N. D. Chamberlain et al., "Ligation of TLR5 promotes myeloid cell infiltration and differentiation into mature osteoclasts in rheumatoid arthritis and experimental

- arthritis," *Journal of Immunology*, vol. 193, no. 8, pp. 3902–3913, 2014.
- [32] G. Cavalli, M. Koenders, V. Kalabokis et al., "Treating experimental arthritis with the innate immune inhibitor interleukin-37 reduces joint and systemic inflammation," *Rheumatology*, vol. 55, no. 12, pp. 2220–2229, 2016.
- [33] P. Lacerte, A. Brunet, B. Egarnes, B. Duchêne, J. P. Brown, and J. Gosselin, "Overexpression of TLR2 and TLR9 on monocyte subsets of active rheumatoid arthritis patients contributes to enhance responsiveness to TLR agonists," *Arthritis Research & Therapy*, vol. 18, no. 1, p. 10, 2016.
- [34] H. Aucott, A. Sowinska, H. E. Harris, and P. Lundback, "Ligation of free HMGB1 to TLR2 in the absence of ligand is negatively regulated by the C-terminal tail domain," *Molecular Medicine*, vol. 24, no. 1, p. 19, 2018.
- [35] C. A. Dinarello, "Immunological and inflammatory functions of the interleukin-1 family," *Annual Review of Immunology*, vol. 27, no. 1, pp. 519–550, 2009.
- [36] P. Gasse, C. Mary, I. Guenon et al., "IL-1R1/MyD88 signaling and the inflammasome are essential in pulmonary inflammation and fibrosis in mice," *The Journal of Clinical Investigation*, vol. 117, no. 12, pp. 3786–3799, 2007.
- [37] S. C. Chen, T. H. Tsai, C. H. Chung, and W. H. Li, "Dynamic association rules for gene expression data analysis," *BMC Genomics*, vol. 16, no. 1, p. 786, 2015.
- [38] S. Alagukumar and R. Lawrance, "A selective analysis of microarray data using association rule mining," *Procedia Computer Science*, vol. 47, pp. 3–12, 2015.
- [39] M. Iwahashi, M. Yamamura, T. Aita et al., "Expression of Toll-like receptor 2 on CD16+ blood monocytes and synovial tissue macrophages in rheumatoid arthritis," *Arthritis and Rheumatism*, vol. 50, no. 5, pp. 1457–1467, 2004.
- [40] N. D. Chamberlain, S. J. Kim, O. M. Vila et al., "Ligation of TLR7 by rheumatoid arthritis synovial fluid single strand RNA induces transcription of TNF $\alpha$  in monocytes," *Annals of the Rheumatic Diseases*, vol. 72, no. 3, pp. 418–426, 2013.
- [41] C. K. Edwards, J. S. Green, H. D. Volk et al., "Combined anti-tumor necrosis factor- $\alpha$  therapy and DMARD therapy in rheumatoid arthritis patients reduces inflammatory gene expression in whole blood compared to DMARD therapy alone," *Frontiers in Immunology*, vol. 3, p. 366, 2012.
- [42] S. Ramírez-Pérez, U. de la Cruz-Mosso, J. Hernández-Bello et al., "High expression of interleukine-1 receptor antagonist in rheumatoid arthritis: association with *IL1RN*\*2/2 genotype," *Autoimmunity*, vol. 50, no. 8, pp. 468–475, 2017.
- [43] X. T. Shao, L. Feng, L. J. Gu et al., "Expression of interleukin-18, IL-18BP, and IL-18R in serum, synovial fluid, and synovial tissue in patients with rheumatoid arthritis," *Clinical and Experimental Medicine*, vol. 9, no. 3, pp. 215–221, 2009.
- [44] Q. Q. Huang, Y. Ma, A. Adebayo, and R. M. Pope, "Increased macrophage activation mediated through Toll-like receptors in rheumatoid arthritis," *Arthritis and Rheumatism*, vol. 56, no. 7, pp. 2192–2201, 2007.
- [45] M. Oosting, S. C. Cheng, J. M. Bolscher et al., "Human TLR10 is an anti-inflammatory pattern-recognition receptor," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 42, pp. E4478–E4484, 2014.
- [46] S. Jiang, X. Li, N. J. Hess, Y. Guan, and R. I. Tapping, "TLR10 is a negative regulator of both MyD88-dependent and -independent TLR signaling," *Journal of Immunology*, vol. 196, no. 9, pp. 3834–3841, 2016.
- [47] Y. Lai, C. Xue, Y. Liao et al., "Expression profiles of toll-like receptor signaling pathway related genes in microscopic polyangiitis in Chinese people," *International Journal of Clinical and Experimental Pathology*, vol. 9, pp. 5515–5524, 2016.
- [48] T. Celhar, H. Yasuga, H. Y. Lee et al., "Toll-like receptor 9 deficiency breaks tolerance to RNA-associated antigens and up-regulates Toll-like receptor 7 protein in Sle1 mice," *Arthritis & Rheumatology*, vol. 70, no. 10, pp. 1597–1609, 2018.
- [49] F. K. Tan, X. Zhou, M. D. Mayes et al., "Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients," *Rheumatology*, vol. 45, no. 6, pp. 694–702, 2006.
- [50] M. A. Martínez-Godínez, M. D. Cruz-Domínguez, L. J. Jara et al., "Expression of NLRP3 inflammasome, cytokines and vascular mediators in the skin of systemic sclerosis patients," *The Israel Medical Association Journal*, vol. 17, pp. 5–10, 2015.
- [51] H. Ihn, S. Sato, M. Fujimoto, K. Kikuchi, and K. Takehara, "Demonstration of interleukin 8 in serum samples of patients with localized scleroderma," *Archives of Dermatology*, vol. 130, no. 10, pp. 1327–1328, 1994.
- [52] G. Kaplanski, "Interleukin-18: biological properties and role in disease pathogenesis," *Immunological Reviews*, vol. 281, no. 1, pp. 138–153, 2018.

## APPENDIX F

**Petrackova A**, Smrzova A, Gajdos P, Schubertova M, Schneiderova P, Kromer P, Snasel V, Skacelova M, Mrazek F, Zadrazil J, Horak P, Kriegova E. Serum protein pattern associated with organ damage and lupus nephritis in systemic lupus erythematosus revealed by PEA immunoassay. *Clinical Proteomics*. 2017;14:32. (IF 2017 3.516)

*Dean's award for the best student scientific publications in 2017*



RESEARCH

Open Access



# Serum protein pattern associated with organ damage and lupus nephritis in systemic lupus erythematosus revealed by PEA immunoassay

Anna Petrackova<sup>1</sup>, Andrea Smrzova<sup>2</sup>, Petr Gajdos<sup>3</sup>, Marketa Schubertova<sup>2</sup>, Petra Schneiderova<sup>1</sup>, Pavel Kromer<sup>3</sup>, Vaclav Snasel<sup>3</sup>, Martina Skacelova<sup>2</sup>, Frantisek Mrazek<sup>1</sup>, Josef Zadrazil<sup>2</sup>, Pavel Horak<sup>2</sup> and Eva Kriegova<sup>1\*</sup>

## Abstract

**Background:** Systemic lupus erythematosus (SLE) is a remarkably heterogeneous autoimmune disease. Despite tremendous efforts, our knowledge of serum protein patterns in severe SLE phenotypes is still limited. We investigated the serum protein pattern of SLE, with special emphasis on irreversible organ damage and active lupus nephritis (LN) as assessed by renal Systemic Lupus Erythematosus Disease Activity Index.

**Methods:** We used proximity extension immunoassay (PEA, Proseek Multiplex, Olink) to assess the serum levels of ninety-two inflammation-related proteins in Czech patients with SLE ( $n = 75$ ) and age-matched healthy control subjects ( $n = 23$ ). Subgroup analysis was carried out on the basis of organ damage (with/without, 42/33) and biopsy-proven LN (with/without, 27/48; active LN,  $n = 13$ ; inactive LN,  $n = 14$ ).

**Results:** Of thirty deregulated proteins between SLE and the healthy controls ( $P_{corr} < 0.05$ ), the top upregulated proteins in SLE were sirtuin 2, interleukin 18 (IL18), and caspase 8 ( $P_{corr} < 0.0006$ ). Of these, sirtuin 2 and caspase 8 had not yet been reported with SLE. Elevated levels of IL8, CCL2/MCP1, CCL11, and MMP10 ( $P_{corr} < 0.05$ ) were detected in patients with organ damage for which the serum levels of CCL11 and MMP10 were particularly informative in organ damage prediction. Comparing patients based on LN, elevated levels of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF ( $P_{corr} < 0.05$ ) were detected in active LN. Except GDNF, all LN-associated markers showed usefulness in prediction of active renal disease.

**Conclusions:** This highly sensitive PEA analysis identified the serum pattern of SLE, organ damage, and active LN, with many novel candidate proteins detected. Their exact role and suitability as biomarkers in SLE deserve further investigation.

**Keywords:** Serum pattern, Systemic lupus erythematosus, Proximity extension immunoassay, Organ damage, Lupus nephritis

## Background

Systemic lupus erythematosus (SLE) is a serious, complex, multi-system autoimmune rheumatic disease with significant variability in the phenotypes and severity of

the disease. The greatest challenges continue to be the prevention and management of irreversible organ damage and active lupus nephritis (LN), one of the most feared phenotypes in SLE.

Organ damage is a primary outcome in SLE, which is accrued not only during the disease course, but also by therapy itself [1]. Early damage is more likely to be linked to active inflammation, while late irreversible damage is often attributable to the side effects of drugs and

\*Correspondence: [eva.kriegova@email.cz](mailto:eva.kriegova@email.cz)

<sup>1</sup> Department of Immunology, Faculty of Medicine and Dentistry, Palacky University, Hnevotinska 3, 775 15 Olomouc, Czech Republic  
Full list of author information is available at the end of the article

especially to chronic and cumulative corticosteroid exposure [2]. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology SLICC/ACR Damage Index (SDI), divided into 38 items grouped in 12 organ systems, is a valid measure of irreversible organ damage in SLE [1]. Despite improvement in the survival of SLE patients in recent decades, significantly higher morbidity and mortality are reported in patients developing irreversible organ damage [1]. The patterns of organ damage vary among populations [3–5], but the musculoskeletal, cardiovascular, and renal systems are those most frequently affected [6]. Nowadays, prevention of irreversible damage is a major goal in the management of SLE patients and identification of the key molecules involved in the pathogenesis of organ damage is needed.

Lupus nephritis is a major manifestation associated with higher morbidity and mortality of SLE patients [7]. It has a considerable influence on treatment decisions, as well as long-term outcomes. The effective treatment of LN requires a correct diagnosis, timely intervention, and early treatment of any disease relapse. Renal biopsy is still the gold standard for diagnosis and deciding on therapy in LN but its invasive nature prevents it from being used repetitively in many cases [8]. Traditional clinical parameters such as proteinuria, glomerular filtration rate, urine sediments, anti-dsDNA antibodies, and complement levels are not sensitive or specific enough to detect activity and early relapse of LN [9, 10]. Novel serum and urinary biomarkers such as cytokines and chemokines CCL2 [11], CCL3, CCL5 [12], IL17 [11], BlyS, APRIL [13], growth factor TGF $\beta$  [11] and others (TWEAK [14], IGFBP2 [15], OPG [16]) have recently been nominated for diagnosis and monitoring of LN. Although intensively investigated [17, 18], only a few biomarkers have been assessed for prediction of renal activity or prognosis. Identification of novel and reliable biomarkers or their combinations for LN reflecting also disease activity is, therefore, highly desirable.

In this study we aimed to assess the serum protein pattern of SLE using a highly sensitive multiplex proximity extension immunoassay (PEA) on 92 inflammation-related proteins. Special emphasis was given to serum patterns associated with irreversible organ damage and LN reflecting the renal disease activity and their usefulness in the prediction of these severe phenotypes.

## Methods

### Study population and materials

Serum samples were obtained from 75 Czech SLE patients; all enrolled patients fulfilled the ACR classification criteria [19]. The samples were aliquoted and stored at  $-80^{\circ}\text{C}$  until further use. Organ damage was assessed by means of the SDI damage index (Systemic Lupus

International Collaborating Clinics/American College of Rheumatology Damage Index) [1] and disease activity was evaluated by means of SLEDAI (Systemic Lupus Erythematosus Disease Activity Index) [20]. Subgroups were formed on the basis of (1) the SDI (SDI = 0,  $n = 33$ ; SDI = 1,  $n = 17$ ; SDI  $\geq 2$ ,  $n = 25$ ), (2) the biopsy-proven presence of LN (no LN,  $n = 48$ ; LN,  $n = 27$ ), and (3) the renal SLEDAI within LN subgroup, where renal SLEDAI score of  $\geq 4$  was taken as an indicator of active LN (inactive LN,  $n = 14$ ; active LN,  $n = 13$ ). The renal SLEDAI consists of the four renal parameters: hematuria, pyuria, proteinuria, and urinary casts [20]. The mean of LN duration in active LN patients was 7 years (range 0–19 years) and in inactive LN patients 8 years (range 1–18 years). The demographic and clinical features are described in Table 1. The age-matched control group of healthy subjects comprised 23 medical staff members (mean age 40, range 26–73, female/male 15/8), who gave statements about their health status and excluded any medication used for SLE treatment (corticosteroids, antimalarials, immunosuppressant drugs). The patients and control subjects provided written informed consent about the usage of peripheral blood for the purpose of this study, which was approved by the ethics committee of the University Hospital and Palacky University Olomouc.

### Proximity extension immunoassay (PEA)

The serum levels of 92 inflammation-related proteins were simultaneously measured by a PEA using the Proseek Multiplex Inflammation kit I (Olink Bioscience, Sweden) according to the manufacturer's recommendation. Briefly, each analyte is recognized by a pair of oligonucleotide-labelled antibodies and when binding to their correct targets, they give rise to reporter amplicons which are amplified and quantified by microfluidic-based real-time PCR (BioMark<sup>TM</sup> HD System, Fluidigm Corporation). The data obtained is normalized and used for the relative quantification of the concentration of each analyte [26, 27]. The PEA kits offer the same level of performance as ELISA and comparable sensitivity to standard ELISA kits with much less sample and a higher dynamic range. For a panel description see Additional file 1: Table S1; for the sensitivity and specificity parameters of the PEA analysis see [26, 27].

### Statistics

All statistical analyses were performed on linearized data (linear ddCq) for each analyte. Statistical tests (Mann–Whitney–Wilcoxon test, Benjamini–Hochberg correction, Spearman correlations, Receiver Operating Characteristic (ROC) curve analysis, and Bayesian probability model) were performed using the R statistical software with the Caret package (<http://www.r-project.org/>;

**Table 1 Demographic and clinical characteristics of enrolled SLE patients and subgroups based on the presence of organ damage and LN**

Demographic and clinical features	SLE (n = 75)	Organ damage (n = 42)	No organ damage (n = 33)	LN (n = 27)	No LN (n = 48)
Female/Male	66/9	34/8	32/1	22/5	44/4
Age (years) mean (min–max)	40 (19–74)	44 (20–67)	35 (19–74)	35 (19–57)	46 (25–64)
Age at the onset of the disease (years) mean (min–max)	27 (11–58)	31 (11–58)	26 (12–56)	24 (12–55)	33 (11–58)
Duration of the disease (years) mean (min–max)	11 (1–38)	13 (1–38)	10 (1–31)	11 (1–20)	13 (1–38)
Organ damage (SDI $\geq$ 2/SDI = 1/SDI = 0)*	25/17/33	25/17/0	0/0/33	10/5/12	15/12/21
Organ damage: SDI mean (min–max)	1.2 (0–8)	2.2 (1–8)	0 (0–0)	1.1 (0–5)	1.3 (0–8)
Lupus nephritis, biopsy proven (Y/N)	27/48	15/27	12/21	27/0	0/48
Neurological involvement (Y/N) <sup>Ⓔ</sup>	22/53	15/27	7/26	9/18	13/35
Hematological involvement (Y/N) <sup>Ⓕ</sup>	19/56	15/27	4/29	5/22	14/34
Cardiovascular involvement (Y/N) <sup>Ⓖ</sup>	12/63	11/31	1/32	4/23	8/40
Skin and musculoskeletal involvement (Y/N) <sup>Ⓗ</sup>	56/19	28/14	28/5	6/21	35/13
Antiphospholipid syndrome (Y/N) <sup>Ⓙ</sup>	23/52	15/27	8/25	6/21	17/31
Renal disorder (Y/N) <sup>Ⓚ</sup>	35/40	19/23	12/21	27/0	9/39
Disease activity: SLEDAI mean (min–max)	7 (0–43)	8.8 (0–43)	4.7 (0–26)	10.3 (0–43)	5.2 (0–20)
Active/inactive renal disease <sup>Ⓛ</sup>	17/58	12/30	5/28	13/14	4/44
Mean of cumulative dose of glucocorticoids (g) (min–max)	22.8 (0–79.2)	30.6 (2.6–79.2)	12.8 (0–54.0)	27.2 (2.4–68.4)	20.3 (0–79.2)

\*SLICC/ACR Damage Index (SDI) was used as a measure of irreversible organ damage in SLE [1]

<sup>Ⓔ</sup> Defined by the ACR nomenclature [21]

<sup>Ⓕ</sup> Defined by the ACR classification criteria [19]

<sup>Ⓖ</sup> Defined as documented pericarditis or myocarditis with compromised left ventricular function or valvular disease

<sup>Ⓗ</sup> Skin involvement defined by Gillian’s criteria [22] and arthritis by ACR definition [19]

<sup>Ⓙ</sup> Defined by preliminary classification criteria for antiphospholipid syndrome [23]

<sup>Ⓚ</sup> Renal SLEDAI score of  $\geq$  4 was taken as an indicator of active LN [24, 25]

<http://topepo.github.io/caret/index.html>). The *P* value for each protein was adjusted for multiple comparisons using the False Discovery Rate by the Benjamini–Hochberg procedure. *P<sub>corr</sub>* value < 0.05 was considered significant.

## Results

### Protein pattern of SLE

In order to assess the serum protein fingerprint associated with SLE, we compared the serum protein levels obtained by PEA immunoassay in the SLE patients and healthy controls. Of 92 biomarkers that were analyzed, the levels of 14 analytes (IL1A, IL2, sIL2RB, IL4, IL5, IL13, IL20, sIL20RA, IL33, TSLP, ARTN, TNE, LIF, NRTN) were below the limit of detection in our sample set and therefore they were excluded from further analysis. Comparing SLE and the controls, 29 proteins were upregulated and sDNER downregulated in SLE (*P<sub>corr</sub>* < 0.05; Table 2a, Additional file 1: Table S2). The distribution of the serum levels of top-upregulated proteins (sirtuin 2, IL18, caspase 8, sCD40/sTNFRSF5, sSLAMF1, sTNFRSF9, axin 1, sulfotransferase 1A1, STAMBP, CCL19/MIP-3 $\beta$ , IL10, and CCL4/MIP-1 $\beta$ ; *P<sub>corr</sub>* < 0.003) is

shown in Fig. 1. For the serum protein pattern associated with SLE and the changes in protein levels between SLE and the controls for top-deregulated analytes see Figs. 2a and 3a.

Because of the suggested central role of the IFN pathway in SLE pathogenesis by promoting feedback loops progressively disrupting peripheral immune tolerance and driving disease activity [28, 29], we investigated the IFN protein “signature” of nine IFN-regulated cytokines. Because of the reported association of an increased IFN gene expression “signature” with disease activity [28, 29], we performed correlation analysis among the protein levels of IFN-regulated chemokines and disease activity as assessed by SLEDAI. The analysis revealed elevation of six IFN-regulated cytokines (IL6, CCL2/MCP1, CCL3/MIP-1 $\alpha$ , sCD40, CXCL11, and CCL19; *P<sub>corr</sub>*  $\leq$  0.01) in SLE and three (CCL8/MCP2, CXCL9, and CXCL10) did not reach significance (*P<sub>corr</sub>* > 0.05). Interestingly, only a mild positive correlation (*r* = 0.25, *P* = 0.03; Additional file 1: Table S3) was observed between the levels of IFN-regulated chemokines and disease activity as assessed by SLEDAI. Disease activity assessed by SLEDAI correlated



**Table 2 Serum levels of proteins differentiating between a healthy controls vs SLE, b SLE patients with organ damage (SDI ≥ 1) vs those without organ damage (SDI = 0), c patients with biopsy-proven active lupus nephritis (active LN) vs patients without lupus nephritis (no LN), d patients with biopsy-proven active lupus nephritis (active LN) vs patients with inactive biopsy-proven lupus nephritis (inactive LN)**

**a Healthy controls vs SLE**

Analyte	Mean linear ddCq (95% CI)		FC	P	P <sub>corr</sub>
	Healthy controls	SLE			
SIRT2	8.31 (6.49–10.1)	19.8 (15.5–24.0)	2.33	6.5 × 10 <sup>-6</sup>	5.1 × 10 <sup>-4</sup>
IL18	183 (155–212)	287 (257–316)	1.67	1.6 × 10 <sup>-5</sup>	6.2 × 10 <sup>-4</sup>
CASP8	2.04 (1.88–2.20)	2.99 (2.68–3.30)	1.37	2.5 × 10 <sup>-5</sup>	6.3 × 10 <sup>-4</sup>
sCD40	527 (466–588)	735 (639–831)	1.29	3.2 × 10 <sup>-5</sup>	6.3 × 10 <sup>-4</sup>
sSLAMF1	5.10 (4.0–6.19)	6.52 (6.0–7.05)	1.39	9.0 × 10 <sup>-5</sup>	1.1 × 10 <sup>-3</sup>
sTNFRSF9	87.8 (75.2–100)	141 (123–159)	1.54	1.1 × 10 <sup>-4</sup>	1.1 × 10 <sup>-3</sup>
ST1A1	3.36 (2.04–4.69)	8.04 (6.66–9.43)	2.41	1.3 × 10 <sup>-4</sup>	1.1 × 10 <sup>-3</sup>
STAMBP	12.1 (10.1–14.1)	18.9 (16.1–21.8)	1.42	1.5 × 10 <sup>-4</sup>	1.1 × 10 <sup>-3</sup>
CCL19	804 (394–1215)	1646 (1326–1966)	2.04	1.5 × 10 <sup>-4</sup>	1.1 × 10 <sup>-3</sup>
IL10	8.33 (6.94–9.72)	17.8 (10.4–25.1)	1.38	3.7 × 10 <sup>-4</sup>	2.6 × 10 <sup>-3</sup>
CCL4	77.7 (63.8–91.6)	123 (109–138)	1.46	4.2 × 10 <sup>-4</sup>	2.7 × 10 <sup>-3</sup>
IL12B	17.3 (13.3–21.4)	29.3 (25.1–33.5)	1.96	5.7 × 10 <sup>-4</sup>	3.4 × 10 <sup>-3</sup>
IL6	4.18 (3.28–5.09)	26.9 (–0.62–54.4)	1.67	7.0 × 10 <sup>-4</sup>	3.9 × 10 <sup>-3</sup>
CCL3	9.58 (4.47–14.7)	35.8 (–11.2–82.9)	1.51	7.6 × 10 <sup>-4</sup>	4.0 × 10 <sup>-3</sup>
CXCL11	187 (142–232)	343 (279–408)	1.58	1.1 × 10 <sup>-3</sup>	5.4 × 10 <sup>-3</sup>
sPDL1	3.43 (3.19–3.67)	4.24 (3.91–4.57)	1.28	1.2 × 10 <sup>-3</sup>	5.4 × 10 <sup>-3</sup>
sIL18R1	106 (90.1–122)	139 (126–152)	1.28	1.4 × 10 <sup>-3</sup>	6.0 × 10 <sup>-3</sup>
sCX3CL1	86.9 (74.8–99.0)	131 (108–153)	1.39	2.2 × 10 <sup>-3</sup>	9.0 × 10 <sup>-3</sup>
sDNER	170 (161–179)	150 (143–157)	0.91	2.6 × 10 <sup>-3</sup>	1.0 × 10 <sup>-2</sup>
sIL15RA	1.90 (1.68–2.13)	2.43 (2.19–2.66)	1.23	2.9 × 10 <sup>-3</sup>	1.1 × 10 <sup>-2</sup>
CSF1	241 (223–259)	281 (270–293)	1.11	3.3 × 10 <sup>-3</sup>	1.2 × 10 <sup>-2</sup>
sLIFR	8.35 (7.84–8.85)	10.3 (8.94–11.6)	1.04	3.4 × 10 <sup>-3</sup>	1.2 × 10 <sup>-2</sup>
IL8	236 (199–273)	383 (281–485)	1.27	3.9 × 10 <sup>-3</sup>	1.2 × 10 <sup>-2</sup>
CCL2	2133 (1801–2465)	3054 (2660–3449)	1.26	3.9 × 10 <sup>-3</sup>	1.2 × 10 <sup>-2</sup>
FGF23	2.77 (2.63–2.91)	4.32 (2.95–5.70)	1.01	5.8 × 10 <sup>-3</sup>	1.8 × 10 <sup>-2</sup>
LAP.TGFB1	131 (117–146)	158 (148–167)	1.28	6.6 × 10 <sup>-3</sup>	1.9 × 10 <sup>-2</sup>
sTRAIL	517 (470–565)	602 (568–636)	1.20	1.1 × 10 <sup>-2</sup>	3.1 × 10 <sup>-2</sup>
MMP10	98.5 (81.5–116)	156 (128–184)	1.34	1.4 × 10 <sup>-2</sup>	3.7 × 10 <sup>-2</sup>
CCL7	5.70 (4.57–6.83)	12.5 (7.31–17.7)	1.38	1.6 × 10 <sup>-2</sup>	4.1 × 10 <sup>-2</sup>

**b SDI = 0 vs SDI ≥ 1**

Analyte	Mean linear ddCq (95% CI)		FC	P	P <sub>corr</sub>
	SDI = 0	SDI ≥ 1			
IL8	286 (215–358)	459 (286–632)	1.32	2.7 × 10 <sup>-5</sup>	2.1 × 10 <sup>-3</sup>
CCL2	2485 (2011–2960)	3502 (2924–4080)	1.50	4.3 × 10 <sup>-4</sup>	1.6 × 10 <sup>-2</sup>
IL6	8.28 (3.95–12.6)	41.5 (–7.95–91.0)	1.98	6.5 × 10 <sup>-4</sup>	1.6 × 10 <sup>-2</sup>
CCL11	273 (247–299)	344 (315–373)	1.30	8.0 × 10 <sup>-4</sup>	1.6 × 10 <sup>-2</sup>
FGF21	56.1 (28.8–83.4)	257 (57.1–456)	2.43	1.0 × 10 <sup>-3</sup>	1.6 × 10 <sup>-2</sup>
MMP10	112 (93.4–131)	190 (144–237)	1.24	2.4 × 10 <sup>-3</sup>	3.1 × 10 <sup>-2</sup>
IL18	255 (208–302)	311 (274–349)	1.19	3.7 × 10 <sup>-3</sup>	4.1 × 10 <sup>-2</sup>
CCL3	9.70 (8.24–11.2)	56.4 (–28.7–142)	1.32	4.8 × 10 <sup>-3</sup>	4.4 × 10 <sup>-2</sup>
FGF5	2.25 (2.13–2.37)	2.72 (2.31–3.12)	1.08	5.1 × 10 <sup>-3</sup>	4.4 × 10 <sup>-2</sup>
FGF23	3.75 (2.51–5.00)	4.77 (2.47–7.06)	1.15	5.7 × 10 <sup>-3</sup>	4.4 × 10 <sup>-2</sup>

**Table 2 continued**

**c No LN vs active LN**

Analyte	Mean linear ddCq (95% CI)		FC	P	P <sub>corr</sub>
	No LN	Active LN			
CSF1	266 (254–278)	340 (310–370)	1.27	4.0 × 10 <sup>-5</sup>	2.7 × 10 <sup>-3</sup>
sIL15RA	2.17 (2.03–2.32)	3.65 (2.57–4.73)	1.43	9.0 × 10 <sup>-5</sup>	2.7 × 10 <sup>-3</sup>
sCD40	645 (595–695)	1116 (587–1645)	1.48	1.0 × 10 <sup>-4</sup>	2.7 × 10 <sup>-3</sup>
sCX3CL1	103 (93.2–112)	247 (134–359)	1.62	2.8 × 10 <sup>-4</sup>	5.4 × 10 <sup>-3</sup>
CASP8	2.80 (2.39–3.21)	3.84 (3.06–4.73)	1.49	1.1 × 10 <sup>-3</sup>	1.7 × 10 <sup>-2</sup>
sIL18R1	129 (119–139)	186 (121–251)	1.26	1.9 × 10 <sup>-3</sup>	2.3 × 10 <sup>-2</sup>
bNGF	2.66 (2.47–2.84)	3.66 (2.96–4.36)	1.41	2.2 × 10 <sup>-3</sup>	2.3 × 10 <sup>-2</sup>
GDNF	4.66 (4.37–4.96)	6.16 (4.98–7.33)	1.32	2.3 × 10 <sup>-3</sup>	2.3 × 10 <sup>-2</sup>

**d Inactive LN vs active LN**

Analyte	Mean linear ddCq (95% CI)		FC	P	P <sub>corr</sub>
	Inactive LN	Active LN			
sIL15RA	2.15 (1.82–2.48)	3.65 (2.57–4.73)	1.56	8.9 × 10 <sup>-4</sup>	6.9 × 10 <sup>-2</sup>
CSF1	280 (253–306)	340 (310–370)	1.28	4.7 × 10 <sup>-3</sup>	0.15
bNGF	2.64 (2.43–2.85)	3.66 (2.96–4.36)	1.29	4.7 × 10 <sup>-3</sup>	0.15
sIL18R1	73.4 (66.2–80.5)	186 (121–251)	1.31	6.6 × 10 <sup>-3</sup>	0.15
sCD40	688 (636–740)	1116 (587–1645)	1.32	1.2 × 10 <sup>-2</sup>	0.18
sCX3CL1	118 (93.6–143)	247 (134–359)	1.42	2.3 × 10 <sup>-2</sup>	0.25
CASP8	2.84 (2.42–3.26)	3.84 (3.06–4.73)	1.33	2.3 × 10 <sup>-2</sup>	0.25

P<sub>corr</sub> value corrected for multiple comparisons (Benjamini–Hochberg correction)

FC (fold-change) between group medians of linear ddCq

better with the following analytes: IL8, GDNE, CX3CL1/fractalkine ( $r \geq 0.403$ ,  $P \leq 0.0003$ ), and CCL7/MCP3, IL15RA, VEGFA, and MMP10 ( $r \geq 0.355$ ,  $P \leq 0.002$ ; Additional file 1: Table S3).

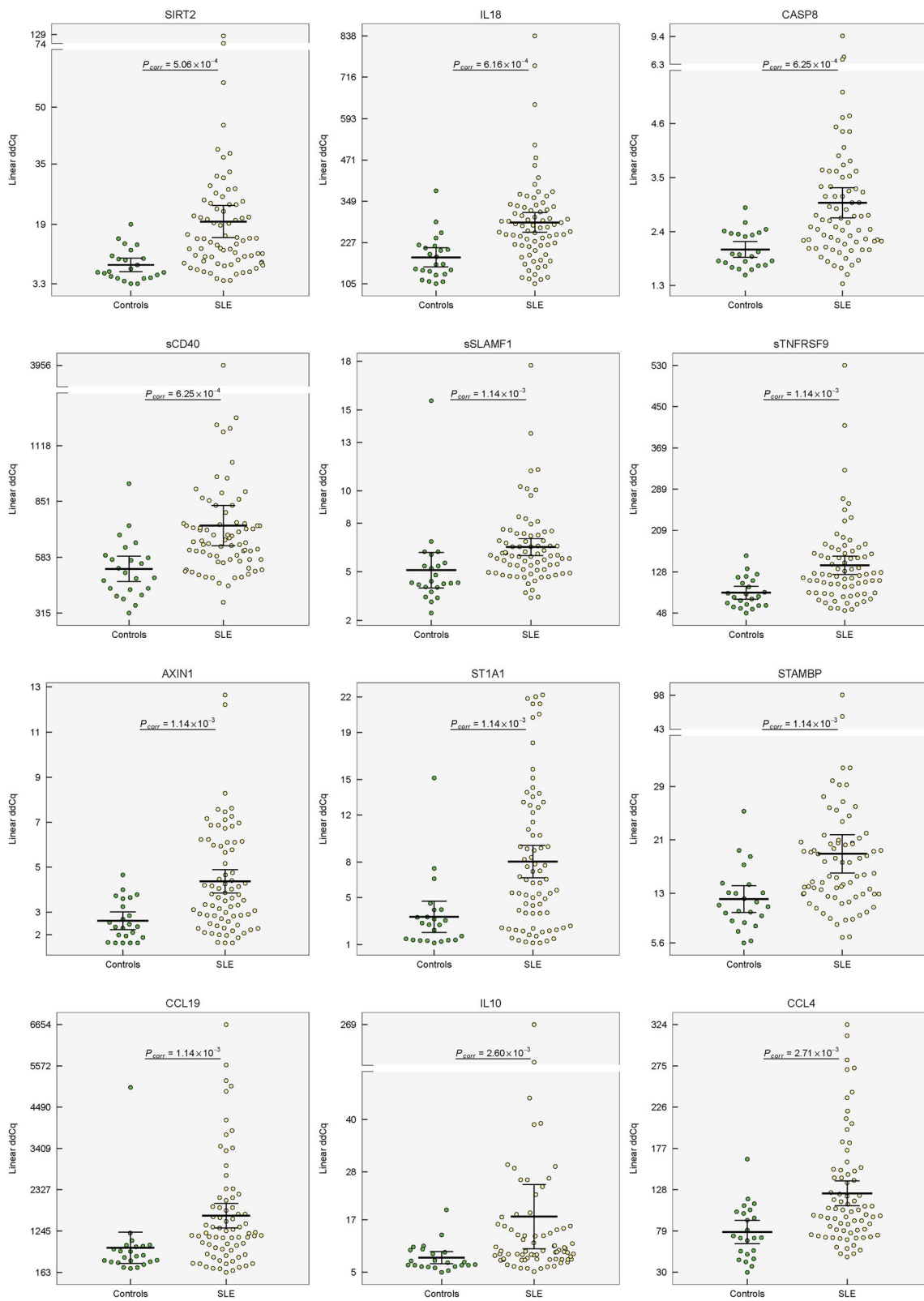
**Protein pattern of organ damage**

To obtain the protein pattern associated with organ damage, we compared the serum patterns from SLE patients with/without organ damage and subgroups according to the SDI ( $SDI \geq 2/SDI = 1/SDI = 0$ ).

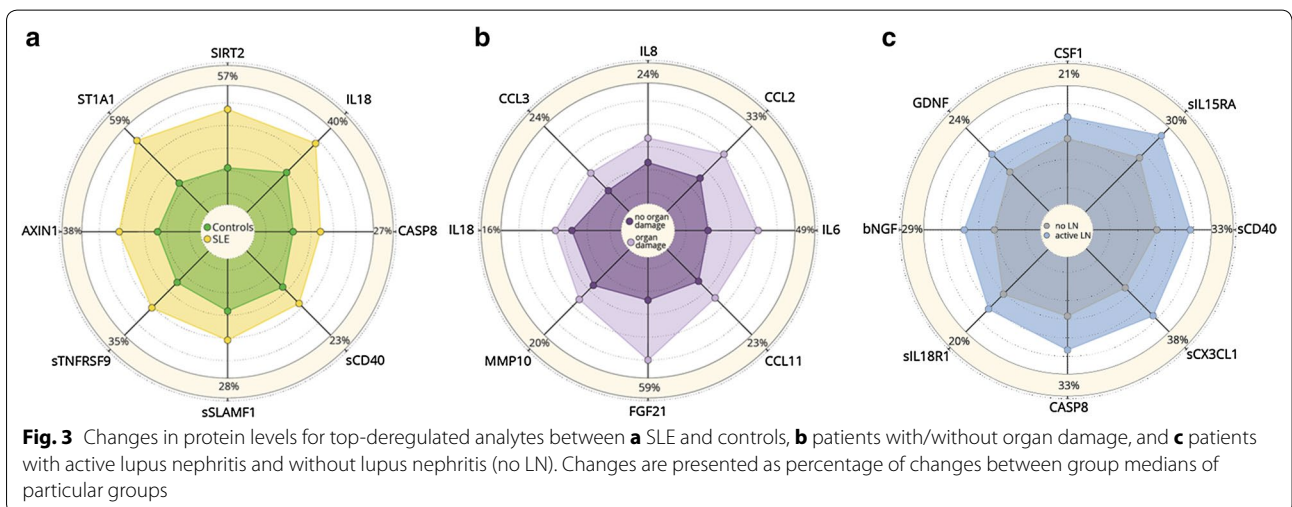
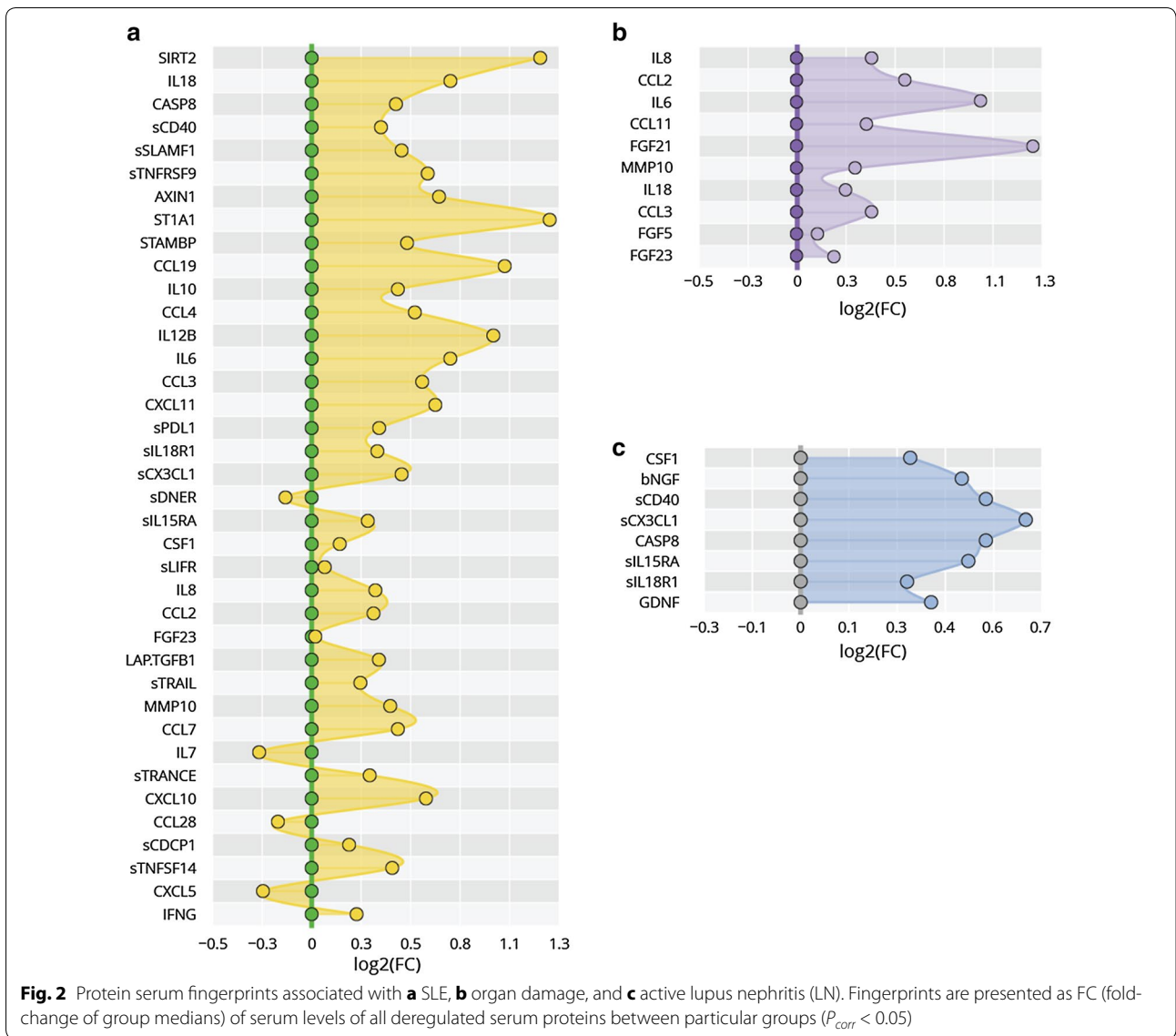
The distribution of damaged organs in our patient group and reported cohorts is shown in Additional file 1: Figure S1, Table S4. In the patients with organ damage ( $SDI \geq 1$ ), elevated serum levels of IL8, CCL2, IL6, CCL11/eotaxin, FGF21, MMP10, IL18, CCL3, FGF5, and FGF23 ( $P_{corr} < 0.05$ ) were detected (Table 2b, Fig. 4). The serum protein pattern associated with organ damage and the changes in protein levels between SLE patients with/without organ damage are shown in Figs. 2b and 3b. Although the serum level of CCL11 did not differ between the controls and SLE patients as a whole, the patients with organ damage had higher levels of CCL11 in comparison to those with no organ damage, as well as to the control

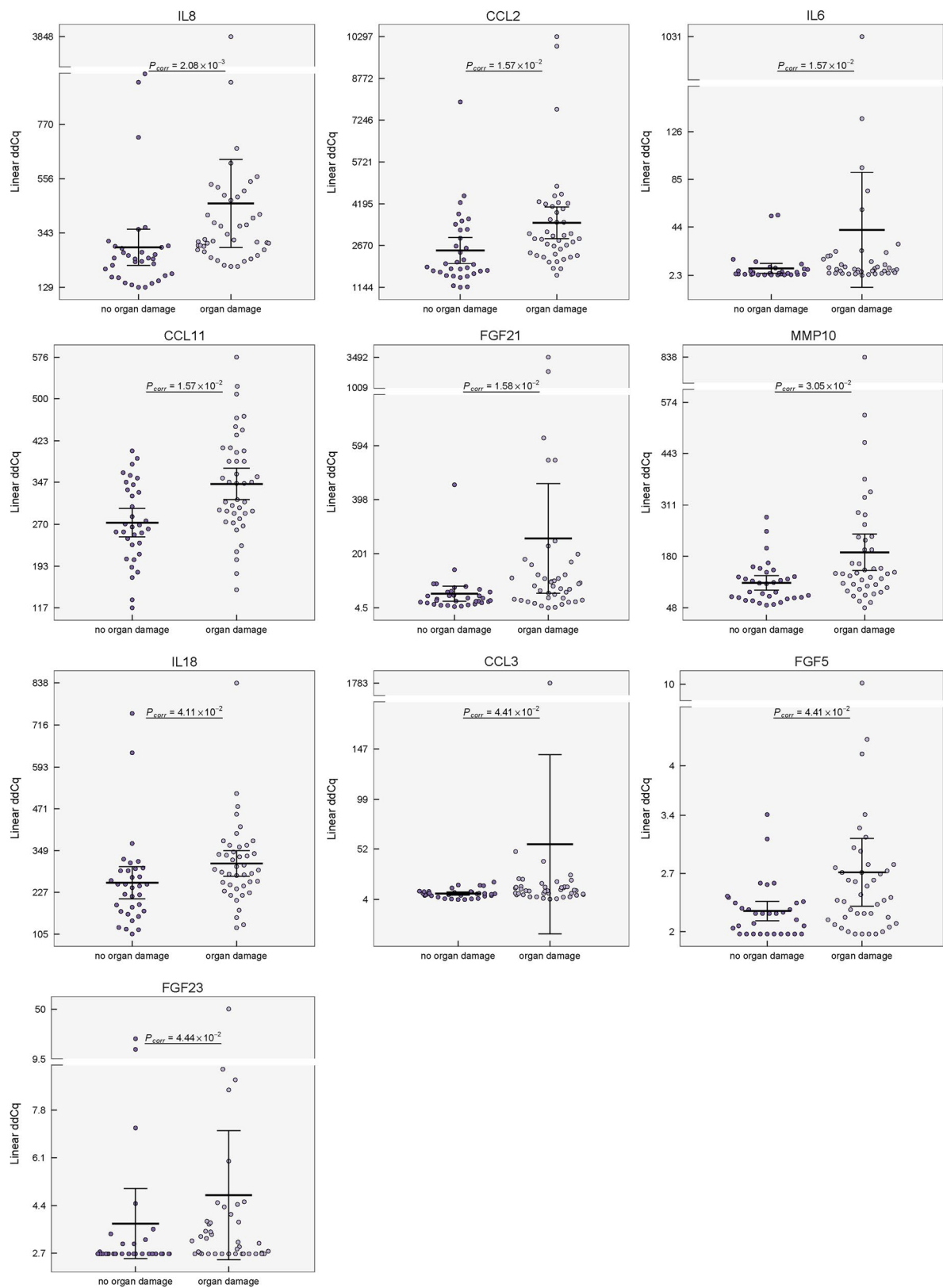
group (Additional file 1: Figure S2a). We did not observe differences in serum protein pattern between patients with  $SDI = 1$  and  $SDI \geq 2$  (data not shown).

Among organ damage associated analytes, the cumulative dose of glucocorticoids correlated positively with levels of IL8, CCL11 ( $r \geq 0.326$ ,  $P \leq 0.004$ ), CCL2 and MMP10 ( $r \geq 0.249$ ,  $P < 0.05$ ; Additional file 1: Table S3). Additionally, cumulative dose of glucocorticoids correlated with BDNE, CCL25, CXCL1, GDNF, IL17C, sADA, sCDCP1, sIL18R1, sSCE, and sTGFA ( $P < 0.05$ ; Additional file 1: Table S3). Moreover, IL8 ( $r = 0.416$ ,  $P = 0.0002$ ), MMP10 ( $r = 0.355$ ,  $P = 0.002$ ), CCL2, and CCL11 ( $r \geq 0.261$ ,  $P \leq 0.02$ ; Additional file 1: Table S3) correlated positively with disease activity. In line with other reports, a higher cumulative dosage of glucocorticoids was registered in the patients with  $SDI \geq 1$  (mean of 30.6 g, min–max 2.6–79.2 g) compared with those without damage (12.8, 0–54.0). Regarding association of disease duration and serum levels of studied proteins, we observed only mild association for CCL11 ( $r = 0.230$ ,  $P = 0.047$ ). The disease duration in SLE patients correlated with SDI ( $r = 0.298$ ,  $P = 0.009$ ).



**Fig. 1** Distribution of serum levels for top-deregulated proteins between healthy controls and SLE. Group means are indicated by horizontal bars, error bars indicate 95% CI;  $P_{corr}$  values after multiple corrections are stated





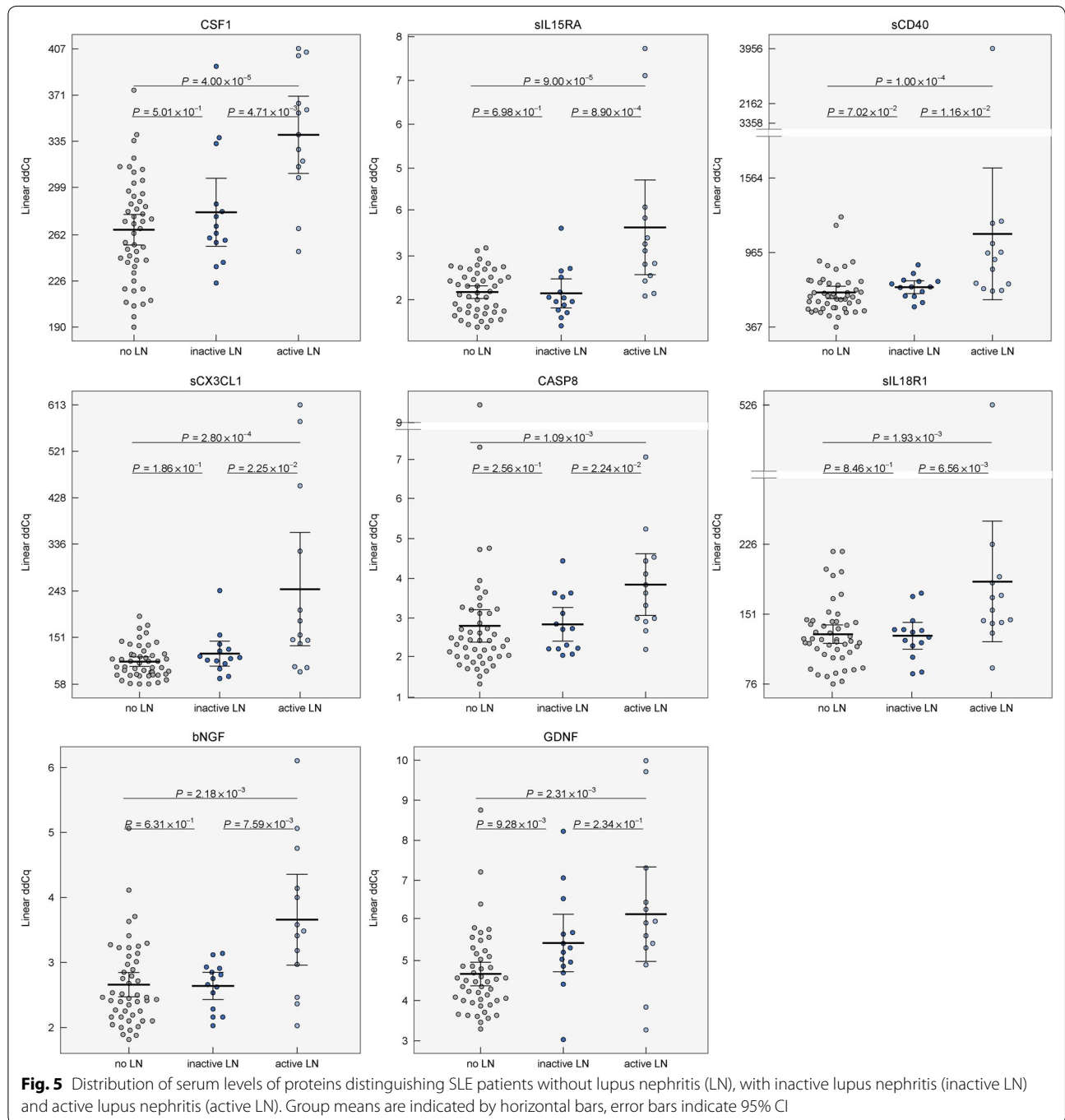
**Fig. 4** Distribution of serum levels of proteins distinguishing SLE patients with/without organ damage. Group means are indicated by horizontal bars, error bars indicate 95% CI;  $P_{corr}$  values for differences after multiple corrections are stated

**Protein pattern of active lupus nephritis and other clinical subsets of SLE**

To investigate the serum patterns associated with active LN, we compared subgroups of SLE patients with/without biopsy-proven LN and subgroups of patients with LN classified by the renal SLEDAI as active (renal SLEDAI  $\geq 4$ ) or inactive renal disease at the day of sampling. Moreover, we assessed serum patterns associated with

other clinical subsets of SLE as neurological, hematological, cardiovascular, skin and musculoskeletal involvements, antiphospholipid syndrome, and renal disorder.

The analysis in biopsy-proven LN patients with active renal disease revealed elevated protein levels of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF compared to those without LN (Table 2c, Fig. 5). Although the serum levels of GDNF did not differ





between the control group and SLE as a whole, its level was enhanced in the patients with LN in comparison to those without LN and the control group (Additional file 1: Figure S2b). The serum protein pattern associated with active LN and the changes in protein levels between the SLE patients without LN and active LN are shown in Figs. 2c and 3c.

When LN patients with active renal disease was compared to inactive LN subgroup, elevation of sIL15RA, CSF1, bNGF, sIL18R1, sCD40, sCX3CL1, and caspase 8 ( $P < 0.05$ , Table 2d, Fig. 5), but not GDNF, in active LN patients was observed.

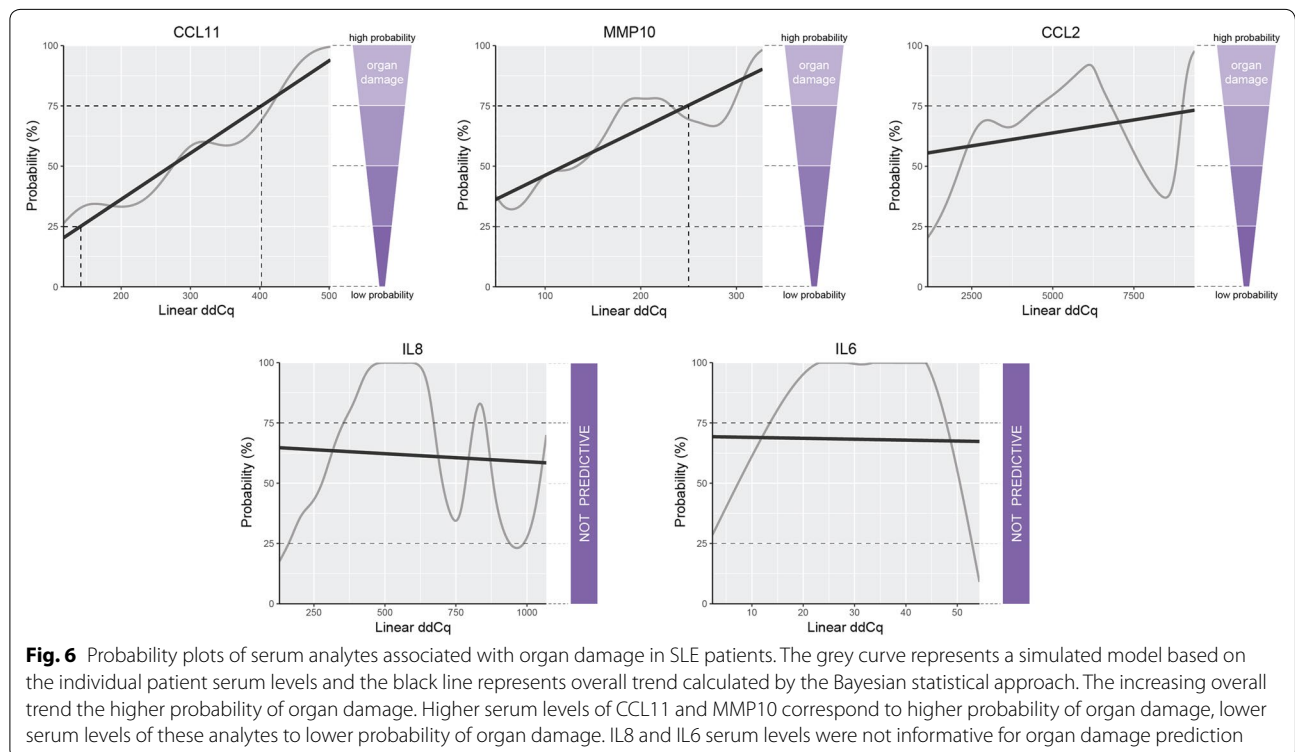
In the other studied clinical subsets no differences in the serum pattern were detected. The subanalysis confirmed that no candidate biomarker for SLE, organ damage and/or LN are influenced by the gender (data not shown).

**Identification of patients with a high probability of organ damage and active lupus nephritis**

To investigate the utility of the serum levels of phenotype-associated proteins for the identification of patients with a high probability of severe phenotypes, we constructed probability plots for phenotype-associated proteins based on a Bayesian statistical approach. Additionally, we constructed ROC curves for the proteins associated with organ damage and active LN.

In organ damage, the best predictive model was observed for the serum levels of CCL11 and MMP10, followed by CCL2, whereas IL6 and IL8 were not informative (Fig. 6). Higher serum levels of CCL11 and MMP10 correspond to a higher probability of organ damage. For the analytes associated with organ damage, the ROC curve analysis showed that the area under the curve (AUC) of IL8, CCL2, IL6, CCL11, FGF21, MMP10, IL18, CCL3, FGF5, and FGF23 was 0.784, 0.738, 0.731, 0.727, 0.723, 0.706, 0.697, 0.691, 0.689, and 0.676, respectively (Additional file 1: Figure S3a; for sensitivity, specificity, and other parameters see Additional file 1: Table S5a).

In active LN, the best predictive value was observed for CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, and sIL18R1 (Fig. 7). Higher serum levels of all these analytes correspond to a higher probability of the presence of active LN. For the analytes associated with active LN, the ROC curve analysis showed that the AUC of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF were 0.873, 0.857, 0.854, 0.832, 0.798, 0.783, 0.780, and 0.778, respectively (Additional file 1: Figure S3b, Table S5b). Moreover, we observed great sensitivity and specificity for proteins sIL15RA (AUC: 0.879, sensitivity: 100%, specificity: 64.3%), CSF1 (0.813, 84.6, 78.6), sIL18R1 (0.810, 84.6, 78.6), and bNGF (0.805, 69.2, 100) showing good discrimination between active and inactive renal disease in LN patient subgroup (Additional file 1:



**Fig. 6** Probability plots of serum analytes associated with organ damage in SLE patients. The grey curve represents a simulated model based on the individual patient serum levels and the black line represents overall trend calculated by the Bayesian statistical approach. The increasing overall trend the higher probability of organ damage. Higher serum levels of CCL11 and MMP10 correspond to higher probability of organ damage, lower serum levels of these analytes to lower probability of organ damage. IL8 and IL6 serum levels were not informative for organ damage prediction

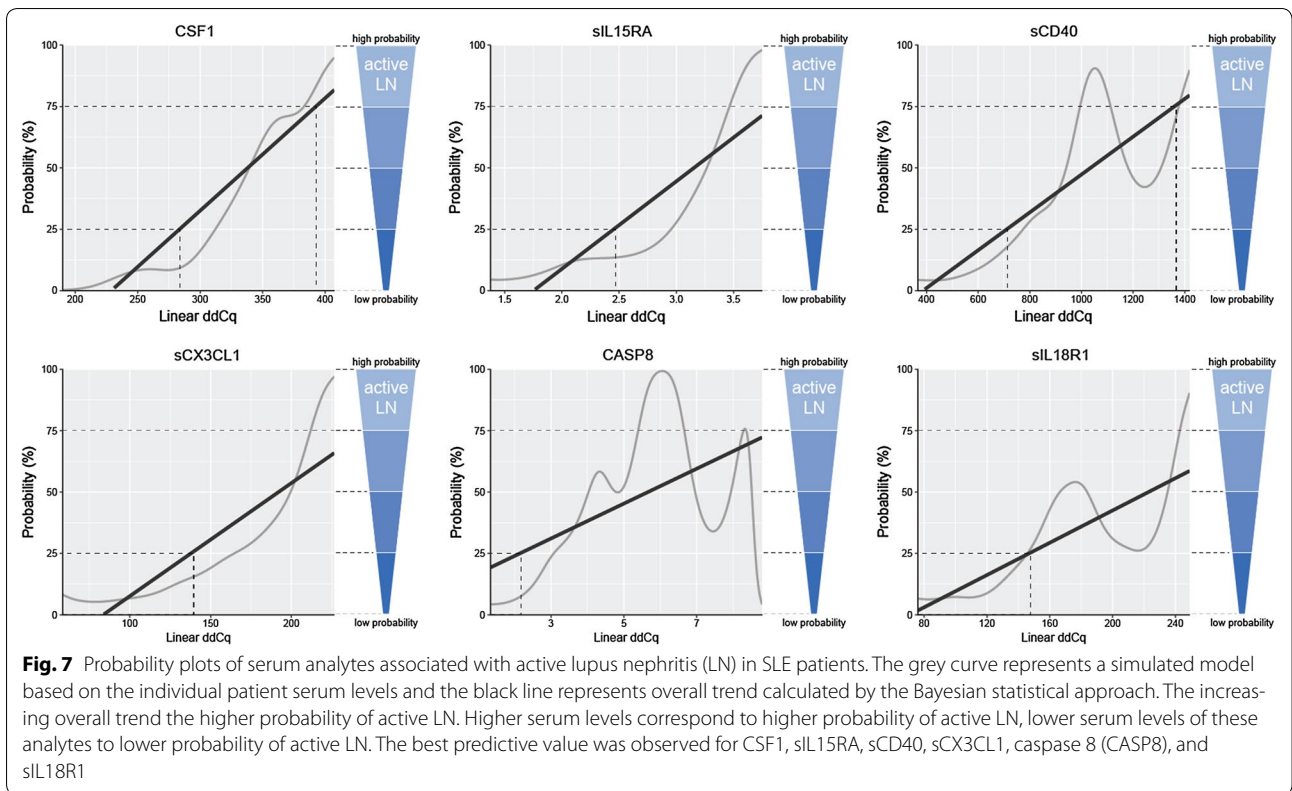


Figure S3c, Table S5c). Inactive LN patients do not differ from patients without LN, except for GDNF (Fig. 5), suggesting that serum GDNF level remains elevated even when LN is inactive.

All nominated biomarkers associated with organ damage and active LN showed better discrimination ability in our cohort than the classical markers (Additional file 1: Table S6). The only exception was proteinuria (AUC 0.869), one of the criteria for renal SLEDAI classification.

### Discussion

Using innovative highly sensitive multiplex PEA analysis on 92 inflammation-related proteins, we identified the serum protein pattern associated with SLE, with many proteins not yet reported in this disease. Moreover, we identified the serum patterns associated with irreversible organ damage and active LN and identified proteins showing utility for the identification of patients at risk of these severe disease manifestations.

This serum protein study in SLE patients revealed the deregulation of 30 proteins in SLE. The majority of the upregulated proteins were known inflammatory mediators: IL6, IL10 [30], IL18 [31], CX3CL1 [32], CCL2 [33], CCL3, CCL7, CCL19 [34], and FGF23 [35] already reported in SLE previously. Interestingly, the

most upregulated proteins—sirtuin 2 and caspase 8—were not associated with SLE or even with any autoimmune disease. However, recent reports in animal models and cell lines support their involvement in inflammation and autoimmunity. Regarding sirtuin 2, macrophages expressing this protein produced more iNOS/NO upon LPS stimulation than those with depleted sirtuin 2 [36]. This result was also confirmed in vivo, where WT mice responded to LPS by increased NO levels and a higher amount of M1-macrophages compared to sirtuin 2 KO mice [36]. Elevated sirtuin 2 also contributed to prolonged hypoinflammation in a septic murine model [37]. Regarding caspase 8, a protein widely recognized for its role in apoptosis, recent reports identify this enzyme as a crucial regulator of inflammation through NFκB activation and cleavage of pro-IL1β and/or pro-IL18, similarly to caspase 1 [38, 39]. These observations lead us to suggest that caspase 8 may also promote autoimmunity by stimulating IL17 production by T cells, as shown for caspase 1 [40]. Moreover, the therapeutic potential of caspase 8 is supported by the observation of attenuated retinal ischemic damage resulting from the inhibition of caspase 8, resulting in the blockade of IL1β production [41]. However, there is evidence about the pleiotropic effects of sirtuin 2 and caspase 8, and thus future studies on their role in SLE are needed.



Further highly upregulated proteins, IL18 and sulfotransferase 1A1, were already reported in autoimmunity. An elevated IL18 serum level was reported in SLE [42], especially in LN patients [43, 44]. Regarding sulfotransferase 1A1, higher activity was found in autoimmune thyroid disease glands compared to normal thyroids [45], but no information yet exists in SLE. Interestingly, we did not detect any elevation of the serum level of the previously reported SLE-associated factor TWEAK and IFN $\gamma$  [46, 47]. Despite the reported association of the IFN gene expression “signature” with disease activity in SLE [28, 29], we did not confirm either elevated levels of the IFN-regulated chemokines CCL8, CXCL9, CXCL10 or strong correlation of the IFN protein “signature” with disease activity at the protein level in the sera of our patients. Our observation is in line with others [28], thus supporting the opinion that cytokine levels in serum are a less sensitive readout for activation of the IFN pathway than the gene expression “signature”.

Despite tremendous efforts, the greatest challenges still remain in the management of SLE patients with severe organ damage and active LN. Thus, there is a need to identify novel biomarkers that will better facilitate the assessment of organ involvement and disease activity. In our study, SLE patients with organ damage had elevated serum levels of IL8, CCL2, IL6, CCL11, FGF21, MMP10, IL18, CCL3, FGF5, and FGF23 compared to those without organ damage. Of these, enhanced levels of CCL11, MMP10, and CCL2 were informative for the identification of patients with organ damage. Importantly, CCL11, MMP10, and CCL2 also correlated with disease activity. The elevation of the chemokine CCL11 was already associated with damage to various organs, as shown in idiopathic retroperitoneal fibrosis [48] and liver cirrhosis patients [49]. Moreover, in murine models of lung fibrosis [50], as well as of eosinophilic myocarditis [51], the blockade of the CCL11-CCR3 pathway prevented organ damage. Similarly, MMP10 was linked to renal damage [52] and tissue destruction in arthritis [53]. Elevation of MMP10 was already reported in SLE patients [54] and in a murine LN model with glomerulonephritis [55]. Another protein associated with organ damage, CCL2, was already reported in kidney damage in lupus murine models [56] and in SLE patients with irreversible renal damage [57]. Although IL6 and IL8, cytokines involved in the pathogenesis of SLE, were also enhanced in our patients with organ damage, our analysis did not support their predictive value for this severe phenotype. The usefulness of CCL11, MMP10, and CCL2 as biomarkers or possible treatment targets needs to be elucidated in future studies.

Lupus nephritis is considered another challenging SLE phenotype from the point of view of its prediction and

preemptive diagnostics. Renal biopsy is still the gold standard to assess the renal involvement of SLE and its severity and pathological category [8]. The search for non-invasive biomarkers in serum and urine reflecting the renal disease activity is therefore a major focus of interest. Our serum protein analysis in LN patients with active renal disease revealed upregulated levels of CSF1, sIL15RA, sCD40, sCX3CL1, caspase 8, sIL18R1, bNGF, and GDNF compared to those without LN. All these markers showed excellent discrimination for active LN, significantly better than the classical markers as shown by us and others [9, 10]. Moreover, we observed good discrimination between active and inactive renal disease in LN patient subgroup for all markers, except for GDNF. Apart from caspase 8 and sIL15RA, emerging evidence of the active involvement of these proteins in LN already exists. Regarding CSF1, elevated serum levels in patients with SLE were shown to reflect kidney histopathology and to predict renal disease activity [58]. Moreover, CSF1 deficiency protected against LN in murine models [59]. Enhanced protein and gene expression of IL15RA was detected in leucocytes from SLE patients [60, 61], probably as a result of hydroxymethylation in promoter region of this gene in SLE [61]. There is also evidence about the crucial role of the CD40-CD40L system in the development, progression and outcome of SLE [62]. Enhanced CD40L protein level was detected in sera from SLE patients [62, 63] as well as class III and IV LN and other inflammatory renal diseases [64]. Moreover, CD40 gene silencing reduced the progression of experimental LN [65]. Regarding sCX3CL1, elevated expression was reported in proliferative LN [66] and the administration of a CX3CL1 antagonist to mice delayed the initiation and ameliorated the progression of LN [67]. Also enhanced expression of IL18R1 has already been reported in SLE patients [68] as well as in peripheral plasmacytoid DCs in active LN patients [69]. Similarly, increased levels of NGF, a complex of 3 subunits—aNGF, bNGF, and gNGF, has been reported in the sera of SLE patients [70] and various renal disorders [71]. Regarding GDNF, a high expression of this protein was detected in renal biopsies from patients with proteinuric nephropathy [72] and increased plasma levels of GDNF were reported in patients with chronic renal diseases [73]. This mesangial autocrine growth factor was shown to play a pivotal role in mesangial cell proliferation, which is essential for the progression of various glomerular diseases [74]. Our study did not confirm IL18 as a useful biomarker to assess the activity of renal disease, as reported by others [42, 43]. On the other hand, our results nominated spectrum of novel biomarkers of renal involvement for further confirmation studies.

Although relatively high sensitivity and specificity was obtained for each individual marker in our LN and organ damage subgroups, we believe that using rather a panel of multiple biomarkers and/or combination with other clinical and laboratory parameters would be an appropriate approach in the identification of patients with these severe manifestations.

## Conclusions

This exploratory study revealed many novel proteins associated with SLE for future immunopathogenesis studies, as well as nominating candidate biomarkers for irreversible organ damage and active lupus nephritis. Future studies on larger cohorts with well-defined phenotypes as well as the longitudinal follow-up during disease development are needed to prove the suitability of these proteins or their combinations as biomarkers for organ damage and lupus nephritis, with special emphasis on disease activity.

## Additional file

**Additional file 1.** SLE serum pattern: Supplemental figures and tables.

## Abbreviations

4E-BP1: eukaryotic translation initiation factor 4E-binding protein 1; aNGF: alpha-nerve growth factor; APRIL: a proliferation-inducing ligand; ARTN: artemin; AXIN1: axin 1; BlyS: B lymphocyte stimulator; BDNF: brain-derived neurotrophic factor; bNGF: beta-nerve growth factor; CASP8: caspase 8; CCL2: C-C motif chemokine ligand 2, monocyte chemoattractant protein 1; CCL3: C-C motif chemokine ligand 3, macrophage inflammatory protein 1-alpha; CCL4: C-C motif chemokine ligand 4, macrophage inflammatory protein 1-beta; CCL7: C-C motif chemokine ligand 7, monocyte chemoattractant protein 3; CCL8: C-C motif chemokine ligand 8, monocyte chemoattractant protein 2; CCL11: C-C motif chemokine ligand 11, eotaxin-1; CCL13: C-C motif chemokine ligand 13, monocyte chemoattractant protein 4; CCL19: C-C motif chemokine ligand 19, macrophage inflammatory protein 3-beta; CCL20: C-C motif chemokine ligand 20, macrophage inflammatory protein 3-alpha; CCL23: C-C motif chemokine ligand 23, macrophage inflammatory protein 3; CCL25: C-C motif chemokine ligand 25; CCL28: C-C motif chemokine ligand 28; CD40L: cluster of differentiation 40 ligand; CCR3: C-C chemokine receptor type 3; CSF1: macrophage colony-stimulating factor 1; CST5: cystatin D; CXCL1: C-X-C motif chemokine ligand 1; CXCL5: C-X-C motif chemokine ligand 5; CXCL6: C-X-C motif chemokine ligand 6; CXCL9: C-X-C motif chemokine ligand 9; CXCL10: C-X-C motif chemokine ligand 10; CXCL11: C-X-C motif chemokine ligand 11; ENRAGE: extracellular newly identified receptor for advanced glycation end-products binding protein, protein S100-A12; FGF5: fibroblast growth factor 5; FGF19: fibroblast growth factor 19; FGF21: fibroblast growth factor 21; FGF23: fibroblast growth factor 23; gNGF: gamma-nerve growth factor; GDNF: glial cell line-derived neurotrophic factor; IFN: interferon; IFNG: interferon gamma; IGFBP2: insulin like growth factor binding protein 2; IL1A: interleukin-1 alpha; IL1B: interleukin-1 beta; IL2: interleukin-2; IL4: interleukin-4; IL5: interleukin-5; IL6: interleukin-6; IL7: interleukin-7; IL8: interleukin-8; IL10: interleukin-10; IL12B: interleukin-12 beta; IL13: interleukin-13; IL17: interleukin-17; IL17A: interleukin-17A; IL17C: interleukin-17C; IL18: interleukin-18; IL20: interleukin-20; IL24: interleukin-24; IL33: interleukin-33; iNOS/NO: inducible nitric oxide synthase/nitrogen oxide; KO: knock out; LAP.TGFB1: latency-associated peptide transforming growth factor beta-1; LIF: leukemia inhibitory factor; LN: lupus nephritis; LPS: lipopolysaccharide; MMP1: matrix metalloproteinase-1; MMP10: matrix metalloproteinase-10; NFkB: nuclear factor kappa B; NGF: nerve growth factor; NRTN: neurturin; NT3: neurotrophin-3; OSM: oncostatin-M; PCR: polymerase chain reaction; PEA: proximity extension immunoassay;

ROC: receiver operating characteristic; sADA: adenosine deaminase, soluble; sCD244: natural killer cell receptor 2B4, soluble; sCD40: cluster of differentiation 40, tumor necrosis factor receptor superfamily member 5, soluble; sCD5: cluster of differentiation 5, soluble; sCD6: cluster of differentiation 6, soluble; sCDPC1: CUB domain-containing protein 1, soluble; sCX3CL1: C-X3-C motif chemokine ligand 1, fractalkine, soluble; SDI: Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; sDNER: delta and notch-like epidermal growth factor-related receptor, soluble; sFlt3L: Fms-related tyrosine kinase 3 ligand, soluble; sHGF: hepatocyte growth factor; sIL10RA: interleukin-10 receptor subunit alpha, soluble; sIL10RB: interleukin-10 receptor subunit beta, soluble; sIL15RA: interleukin-15 receptor subunit alpha, soluble; sIL18R1: interleukin-18 receptor 1, soluble; sIL20RA: interleukin-20 receptor subunit alpha, soluble; sIL22RA1: interleukin-22 receptor subunit alpha-1, soluble; sIL2RB: interleukin-2 receptor subunit beta, soluble; SIRT2: sirtuin 2; SLE: systemic lupus erythematosus; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; sLIFR: leukemia inhibitory factor receptor, soluble; sOPG: osteoprotegerin, soluble; sPDL1: programmed cell death 1 ligand 1, soluble; sSCF: stem cell factor, soluble; sSLAMF1: signaling lymphocytic activation molecule 1, soluble; ST1A1: sulfotransferase 1A1; STAMPB: signal transducing adaptor molecule-binding protein; sTGFA: transforming growth factor alpha, soluble; sTNFB: tumor necrosis factor-beta, soluble; sTNFRSF9: tumor necrosis factor receptor superfamily member 9, soluble; sTNFSF14: tumor necrosis factor ligand superfamily member 14, soluble; sTRAIL: tumor necrosis factor-related apoptosis-inducing ligand, soluble; sTRANCE: tumor necrosis factor-related activation-induced cytokine, soluble; sTWEAK: tumor necrosis factor ligand superfamily member 12, soluble; TGFb: transforming growth factor beta; TNF: tumor necrosis factor; TSLP: thymic stromal lymphopoietin; uPA: urokinase-type plasminogen activator; VEGFA: vascular endothelial growth factor A; WT: wild type.

## Authors' contributions

All authors contributed substantially to the conception and design, analysis and interpretation of data, drafting the article or revising it critically for important intellectual content. Final approval of the version to be published was given by all the authors. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> Department of Immunology, Faculty of Medicine and Dentistry, Palacky University, Hnevotinska 3, 775 15 Olomouc, Czech Republic. <sup>2</sup> Department of Internal Medicine III - Nephrology, Rheumatology and Endocrinology, Faculty of Medicine and Dentistry, University Hospital, Palacky University, Olomouc, Czech Republic. <sup>3</sup> Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava, Ostrava, Czech Republic.

## Acknowledgements

We apologize to many authors whose important works could not be cited due to space limitations. We thank Martin Radvansky and Sarka Zehnalova for their kind help with preparing figures for Additional file.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

All data generated or analysed during this study are included in this published article and its additional file.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Patients and control subjects provided written informed consent about the usage of peripheral blood for the purpose of this study which was approved by the ethics committee of University Hospital and Palacky University Olomouc.

## Funding

Funding was obtained from the Ministry of Health of Czech Republic (MZ CR VES15-28659A).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 23 November 2016 Accepted: 18 September 2017

Published online: 03 October 2017

## References

- Gladman DD, Goldsmith CH, Urowitz MB, Bacon P, Fortin P, Ginzler E, et al. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) damage index for systemic lupus erythematosus international comparison. *J Rheumatol*. 2000;27:373–6.
- Nossent J, Cikes N, Kiss E, Marchesoni A, Nasonova V, Mosca M, et al. Current causes of death in systemic lupus erythematosus in Europe, 2000–2004: relation to disease activity and damage accrual. *Lupus*. 2007;16:309–17.
- Chambers SA, Allen E, Rahman A, Isenberg D. Damage and mortality in a group of British patients with systemic lupus erythematosus followed up for over 10 years. *Rheumatology (Oxford)*. 2009;48:673–5.
- Mok CC, Ho CT, Wong RW, Lau CS. Damage accrual in southern Chinese patients with systemic lupus erythematosus. *J Rheumatol*. 2003;30:1513–9.
- Danila MI, Pons-Estel GJ, Zhang J, Vilá LM, Reveille JD, Alarcón GS. Renal damage is the most important predictor of mortality within the damage index: data from LUMINA LXIV, a multiethnic US cohort. *Rheumatology (Oxford)*. 2009;48:542–5.
- Doria A, Gatto M, Zen M, Iaccarino L, Punzi L. Optimizing outcome in SLE: treating-to-target and definition of treatment goals. *Autoimmun Rev*. 2014;13:770–7.
- Cameron JS. Lupus nephritis. *J Am Soc Nephrol*. 1999;10:413–24.
- Misra R, Gupta R. Biomarkers in lupus nephritis. *Int J Rheum Dis*. 2015;18:219–32.
- Balow JE. Clinical presentation and monitoring of lupus nephritis. *Lupus*. 2005;14:25–30.
- Enghard P, Riemekasten G. Immunology and the diagnosis of lupus nephritis. *Lupus*. 2009;18:287–90.
- Susianti H, Iriane VM, Dharmasana S, Handono K, Widijanti A, Gunawan A, et al. Analysis of urinary TGF- $\beta$ 1, MCP-1, NGAL, and IL-17 as biomarkers for lupus nephritis. *Pathophysiology*. 2015;22:65–71.
- Kulkarni O, Anders HJ. Chemokines in lupus nephritis. *Front Biosci*. 2008;13:3312–20.
- López P, Rodríguez-Carrio J, Caminal-Montero L, Mozo L, Suárez A. A pathogenic IFN $\alpha$ , BlyS and IL-17 axis in Systemic Lupus Erythematosus patients. *Sci Rep*. 2016;6:20651.
- Schwartz N, Rubinstein T, Burkly LC, Collins CE, Blanco I, Su L, et al. Urinary TWEAK as a biomarker of lupus nephritis: a multicenter cohort study. *Arthritis Res Ther*. 2009;11:R143.
- Ding H, Kharboutli M, Saxena R, Wu T. Insulin-like growth factor binding protein-2 as a novel biomarker for disease activity and renal pathology changes in lupus nephritis. *Clin Exp Immunol*. 2016;184:11–8.
- Abdallah E, El-Shishtawy S, Sherif N, Abdelwahab MA. Diagnostic performance of urinary osteoprotegerin as a novel biomarker for early detection of lupus nephritis activity. *Life Sci J*. 2015;12:75–81.
- Li Y, Fang X, Li QZ. Biomarker profiling for lupus nephritis. *Genomics Proteomics Bioinform*. 2013;11:158–65.
- Benjachat T, Tongyoo P, Tantivayakul P, Somparn P, Hirankarn N, Prom-On S, et al. Biomarkers for refractory Lupus nephritis: a microarray study of kidney tissue. *Int J Mol Sci*. 2015;16:14276–90.
- Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*. 1997;40:1725.
- Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. *Arthritis Rheum*. 1992;35:630–40.
- The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. *Arthritis Rheum*. 1999;42:599–608.
- Gilliam JN, Sontheimer RD. Skin manifestations of SLE. *Clin Rheum Dis*. 1981;8:207–18.
- Miakis S, Lockshin MD, Atsumi T, Branch DW, Brey RL, Cervera R, et al. International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS). *J Thromb Haemost*. 2006;4:295–306.
- Singh S, Wu T, Xie C, Vanarsa K, Han J, Mahajan T, et al. Urine VCAM-1 as a marker of renal pathology activity index in lupus nephritis. *Arthritis Res Ther*. 2012;14:R164.
- Xuejing Z, Jiazhen T, Jun L, Xiangqing X, Shuguang Y, Fuyou L. Urinary TWEAK level as a marker of lupus nephritis activity in 46 cases. *J Biomed Biotechnol*. 2012;2012:359647.
- Assarsson E, Lundberg M, Holmquist G, Björkstén J, Thorsen SB, Ekman D, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE*. 2014;9:e95192.
- Schneiderova P, Pika T, Gajdos P, Fillerova R, Kromer P, Kudelka M, et al. Serum protein fingerprinting by PEA immunoassay coupled with a pattern-recognition algorithms distinguishes MGUS and multiple myeloma. *Oncotarget*. 2016 doi:10.18632/oncotarget.11242. [Epub ahead of print].
- Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci USA*. 2003;100:2610–5.
- Feng X, Huang J, Liu Y, Xiao L, Wang D, Hua B, et al. Identification of interferon-inducible genes as diagnostic biomarker for systemic lupus erythematosus. *Clin Rheumatol*. 2015;34:71–9.
- Gröndal G, Gunnarsson I, Rönnelid J, Rogberg S, Klareskog L, Lundberg I. Cytokine production, serum levels and disease activity in systemic lupus erythematosus. *Clin Exp Rheumatol*. 2000;18:565–70.
- Favilli F, Anzilotti C, Martinelli L, Quattroni P, De Martino S, Pratesi F, et al. IL-18 activity in systemic lupus erythematosus. *Ann NY Acad Sci*. 2009;1173:301–9.
- Yajima N, Kasama T, Isozaki T, Odai T, Matsunawa M, Negishi M, et al. Elevated levels of soluble fractalkine in active systemic lupus erythematosus: potential involvement in neuropsychiatric manifestations. *Arthritis Rheum*. 2005;52:1670–5.
- Hrycek E, Franek A, Błaszczak E, Dworak J, Hrycek A. Serum levels of selected chemokines in systemic lupus erythematosus patients. *Rheumatol Int*. 2013;33:2423–7.
- Bauer JW, Baechler EC, Petri M, Batliwalla FM, Crawford D, Ortmann WA, et al. Elevated serum levels of interferon-regulated chemokines are biomarkers for active human systemic lupus erythematosus. *PLoS Med*. 2006;3:e491.
- Masi L, Cavalli L, Falcini F, Franceschelli F, Leoncini G, Fossi C, et al. P8-measurement of fibroblast growth factor-23 (FGF23) in the serum of patients affected by juvenile systemic lupus erythematosus: a possible marker of kidney damage. *Clin Cases Miner Bone Metab*. 2010;7:214.
- Lee AS, Jung YJ, Kim D, Nguyen-Thanh T, Kang KP, Lee S, et al. SIRT2 ameliorates lipopolysaccharide-induced inflammation in macrophages. *Biochem Biophys Res Commun*. 2014;450:1363–9.
- Wang X, Buechler NL, Martin A, Wells J, Yoza B, McCall CE, et al. Sirtuin-2 regulates sepsis inflammation in ob/ob mice. *PLoS ONE*. 2016;11:e0160431.
- Lemmers B, Salmena L, Bidère N, Su H, Matsiyak-Zablocki E, Murakami K, et al. Essential role for caspase-8 in Toll-like receptors and NF- $\kappa$ B signaling. *J Biol Chem*. 2007;282:7416–23.
- Gurung P, Anand PK, Malireddi RK, Vande Walle L, Van Opdenbosch N, Dillon CP, et al. FADD and caspase-8 mediate priming and activation of the canonical and noncanonical Nlrp3 inflammasomes. *J Immunol*. 2014;192:1835–46.
- Lalor SJ, Dungan LS, Sutton CE, Basdeo SA, Fletcher JM, Mills KH. Caspase-1-processed cytokines IL-1 $\beta$  and IL-18 promote IL-17 production by gamma delta and CD4 T cells that mediate autoimmunity. *J Immunol*. 2011;186:5738–48.
- Chi W, Li F, Chen H, Wang Y, Zhu Y, Yang X, et al. Caspase-8 promotes NLRP1/NLRP3 inflammasome activation and IL-1 $\beta$  production in acute glaucoma. *Proc Natl Acad Sci USA*. 2014;111:11181–6.
- Sahebari M, Rezaieyazdi Z, Nakhjavani MJ, Hatem M, Mahmoudi M, Akhlaghi S. Correlation between serum concentrations of soluble Fas (CD95/Apo-1) and IL-18 in patients with systemic lupus erythematosus. *Rheumatol Int*. 2012;32:601–6.

43. Mohsen MA, Abdel Karim SA, Abbas TM, Amin M. Serum interleukin-18 levels in patients with systemic lupus erythematosus: relation with disease activity and lupus nephritis. *Egypt Rheumatol*. 2013;35:45–51.
44. Calvani N, Richards HB, Tucci M, Pannarale G, Silvestris F. Up-regulation of IL-18 and predominance of a Th1 immune response is a hallmark of lupus nephritis. *Clin Exp Immunol*. 2004;138:171–8.
45. Ebmeier CC, Anderson RJ. Human thyroid phenol sulfotransferase enzymes 1A1 and 1A3: activities in normal and diseased thyroid glands, and inhibition by thyroid hormones and phytoestrogens. *J Clin Endocrinol Metab*. 2004;89:5597–605.
46. Choe JY, Kim SK. Serum TWEAK as a biomarker for disease activity of systemic lupus erythematosus. *Inflamm Res*. 2016;65:479–88.
47. Al-Janadi M, Al-Balla S, Al-Dalaan A, Raziuddin S. Cytokine profile in systemic lupus erythematosus, rheumatoid arthritis and other rheumatic diseases. *J Clin Immunol*. 1993;13:58–67.
48. Mangieri D, Corradi D, Martorana D, Malerba G, Palmisano A, Libri I, et al. Eotaxin/CCL11 in idiopathic retroperitoneal fibrosis. *Nephrol Dial Transpl*. 2012;27:3875–84.
49. Tacke F, Trautwein C, Yagmur E, Hellerbrand C, Wiest R, Brenner DA, et al. Up-regulated eotaxin plasma levels in chronic liver disease patients indicate hepatic inflammation, advanced fibrosis and adverse clinical course. *J Gastroenterol Hepatol*. 2007;22:1256–64.
50. Huaux F, Gharaee-Kermani M, Liu T, Morel V, McGarry B, Ullenbruch M, et al. Role of Eotaxin-1 (CCL11) and CC chemokine receptor 3 (CCR3) in bleomycin-induced lung injury and fibrosis. *Am J Pathol*. 2005;167:1485–96.
51. Diny NL, Hou X, Barin JG, Chen G, Talor MV, Schaub J, et al. Macrophages and cardiac fibroblasts are the main producers of eotaxins and regulate eosinophil trafficking to the heart. *Eur J Immunol*. 2016; doi: [10.1002/eji.201646557](https://doi.org/10.1002/eji.201646557). [Epub ahead of print].
52. Friese RS, Rao F, Khandrika S, Thomas B, Ziegler MG, Schmid-Schönbein GW, et al. Matrix metalloproteinases: discrete elevations in essential hypertension and hypertensive end-stage renal disease. *Clin Exp Hypertens*. 2009;31:521–33.
53. Barksby HE, Milner JM, Patterson AM, Peake NJ, Hui W, Robson T, et al. Matrix metalloproteinase 10 promotion of collagenolysis via procollagenase activation: implications for cartilage degradation in arthritis. *Arthritis Rheum*. 2006;54:3244–53.
54. Wu T, Ding H, Han J, Arriens C, Wei C, Han W, et al. Antibody-array-based proteomic screening of serum markers in systemic lupus erythematosus: a discovery study. *J Proteome Res*. 2016;15:2102–14.
55. Wang H, Bagavant H, Deshmukh U. Glomerular transcriptional profiles reveal the candidate biomarkers diagnostic for the progression of lupus nephritis from acute to chronic stages. *Arthritis Rheum*. 2008;58:S317.
56. Tesch GH, Maifert S, Schwarting A, Rollins BJ, Kelley VR. Monocyte chemoattractant protein 1-dependent leukocytic infiltrates are responsible for autoimmune disease in MRL-Fas(lpr) mice. *J Exp Med*. 1999;190:1813–24.
57. Chan RW, Lai FM, Li EK, Tam L, Chow K, Lai K, et al. Intrarenal cytokine gene expression in lupus nephritis. *Ann Rheum Dis*. 2007;66:886–92.
58. Menke J, Amann K, Cavagna L, Blettner M, Weinmann A, Schwarting A, et al. Colony-stimulating factor-1: a potential biomarker for lupus nephritis. *J Am Soc Nephrol*. 2015;26:379–89.
59. Menke J, Rabacal WA, Byrne KT, Iwata Y, Schwartz MM, Stanley ER, et al. Circulating CSF-1 promotes monocyte and macrophage phenotypes that enhance lupus nephritis. *J Am Soc Nephrol*. 2009;20:2581–92.
60. Baranda L, de la Fuente H, Layseca-Espinosa E, Portales-Perez D, Nino-Moreno P, Valencia-Pacheco G, et al. IL-15 and IL-15R in leucocytes from patients with systemic lupus erythematosus. *Rheumatology*. 2005;44:1507–13.
61. Zhao M, Wang J, Liao W, Li D, Li M, Wu H, et al. Increased 5-hydroxymethylcytosine in CD4(+) T cells in systemic lupus erythematosus. *J Autoimmun*. 2016;69:64–73.
62. Goules A, Tzioufas AG, Manousakis MN, Kirou KA, Crow MK, Routsias JG. Elevated levels of soluble CD40 ligand (sCD40L) in serum of patients with systemic autoimmune diseases. *J Autoimmun*. 2006;26:165–71.
63. Ciferska H, Horak P, Hermanova Z, Ordeltova M, Zadrazil J, Tichy T, et al. The levels of sCD30 and of sCD40L in a group of patients with systemic lupus erythematosus and their diagnostic value. *Clin Rheumatol*. 2007;26:723–8.
64. Yellin MJ, D'Agati V, Parkinson G, Han AS, Szema A, Baum D, et al. Immunohistologic analysis of renal CD40 and CD40L expression in lupus nephritis and other glomerulonephritides. *Arthritis Rheum*. 1997;40:124–34.
65. Ripoll E, Merino A, Herrero-Fresneda I, Aran JM, Goma M, Bolanos N, et al. CD40 gene silencing reduces the progression of experimental lupus nephritis modulating local milieu and systemic mechanisms. *PLoS ONE*. 2013;8:e65068.
66. Nakatani K, Yoshimoto S, Iwano M, Asai O, Samejima K, Sakan H, et al. Fractalkine expression and CD16 + monocyte accumulation in glomerular lesions: association with their severity and diversity in lupus models. *Am J Physiol Renal Physiol*. 2010;299:F207–16.
67. Inoue A, Hasegawa H, Kohno M, Ito MR, Terada M, Imai T, et al. Antagonist of fractalkine (CX3CL1) delays the initiation and ameliorates the progression of lupus nephritis in MRL/lpr mice. *Arthritis Rheum*. 2005;52:1522–33.
68. Han GM, Chen SL, Shen N, Ye S, Bao CD, Gu YY. Analysis of gene expression profiles in human systemic lupus erythematosus using oligonucleotide microarray. *Genes Immun*. 2003;4:177–86.
69. Tucci M, Quattraro C, Lombardi L, Pellegrino C, Dammacco F, Silvestris F. Glomerular accumulation of plasmacytoid dendritic cells in active lupus nephritis: role of interleukin-18. *Arthritis Rheum*. 2008;58:251–62.
70. Bracci-Laudiero L, Aloe L, Levi-Montalcini R, Galeazzi M, Schilter D, Scully JL, et al. Increased levels of NGF in sera of systemic lupus erythematosus patients. *Neuroreport*. 1993;4:563–5.
71. Antonucci MT, Bonofiglio R, Papalia T, Caruso F, Caroleo MC, Mancuso D, et al. Nerve growth factor and its monocyte receptors are affected in kidney disease. *Nephron Clin Pract*. 2009;111:c21–8.
72. Morigi M, Locatelli M, Rota C, Buelli S, Corna D, Rizzo P, et al. A previously unrecognized role of C3a in proteinuric progressive nephropathy. *Sci Rep*. 2016;6:28445.
73. Onodera H, Nagata T, Kanazawa M, Taguma Y, Itoyama Y. Increased plasma GDNF levels in patients with chronic renal diseases. *Nephrol Dial Transplant*. 1999;14:1604–5.
74. Kalechman Y, Sredni B, Weinstein T, Freidkin I, Tobar A, Albeck M, et al. Production of the Novel Mesangial Autocrine Growth Factors GDNF and IL-10 Is Regulated by the Immunomodulator AS101. *J Am Soc Nephrol*. 2003;14:620–30.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



## APPENDIX G

Kriegova E, Fillerova R, Minarik J, Savara J, Manakova J, **Petrackova A**, Dihel M, Balcarkova, J, Krhovska P, Pika T, Gajdos P, Behalek M, Vasinek M, Papajik T. Whole-genome optical mapping of bone-marrow myeloma cells reveals association of extramedullary multiple myeloma with chromosome 1 abnormalities. *Sci Rep.* 2021;11:14671. (IF 2020: 4.379; Q1)



OPEN

## Whole-genome optical mapping of bone-marrow myeloma cells reveals association of extramedullary multiple myeloma with chromosome 1 abnormalities

Eva Kriegova<sup>1✉</sup>, Regina Fillerova<sup>1</sup>, Jiri Minarik<sup>2</sup>, Jakub Savara<sup>1,3</sup>, Jirina Manakova<sup>1</sup>, Anna Petrackova<sup>1</sup>, Martin Dihel<sup>1</sup>, Jana Balcarkova<sup>2</sup>, Petra Krhovska<sup>2</sup>, Tomas Pika<sup>2</sup>, Petr Gajdos<sup>3</sup>, Marek Behalek<sup>3</sup>, Michal Vasinek<sup>3</sup> & Tomas Papajik<sup>2</sup>

Extramedullary disease (EMM) represents a rare, aggressive and mostly resistant phenotype of multiple myeloma (MM). EMM is frequently associated with high-risk cytogenetics, but their complex genomic architecture is largely unexplored. We used whole-genome optical mapping (Saphyr, Bionano Genomics) to analyse the genomic architecture of CD138+ cells isolated from bone-marrow aspirates from an unselected cohort of newly diagnosed patients with EMM (n = 4) and intramedullary MM (n = 7). Large intrachromosomal rearrangements (> 5 Mbp) within chromosome 1 were detected in all EMM samples. These rearrangements, predominantly deletions with/without inversions, encompassed hundreds of genes and led to changes in the gene copy number on large regions of chromosome 1. Compared with intramedullary MM, EMM was characterised by more deletions (size range of 500 bp–50 kbp) and fewer interchromosomal translocations, and two EMM samples had copy number loss in the 17p13 region. Widespread genomic heterogeneity and novel aberrations in the high-risk *IGH/IGK/IGL*, 8q24 and 13q14 regions were detected in individual patients but were not specific to EMM/MM. Our pilot study revealed an association of chromosome 1 abnormalities in bone marrow myeloma cells with extramedullary progression. Optical mapping showed the potential for refining the complex genomic architecture in MM and its phenotypes.

### Abbreviations

EMM	Extramedullary multiple myeloma
MM	Multiple myeloma
BM	Bone marrow
HMW DNA	High molecular weight DNA
FISH	Fluorescence in situ hybridization
NGS	Next-generation sequencing
BMMC	Bone marrow mononuclear cell
SV	Structural variant
VAF	Variant allele frequency
CNV	Copy number variation

<sup>1</sup>Department of Immunology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Hnevotinska 3, 779 00 Olomouc, Czech Republic. <sup>2</sup>Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Olomouc, Czech Republic. <sup>3</sup>Department of Computer Science, Faculty of Electrical Engineering and Computer Science, VŠB-Technical University of Ostrava, Ostrava, Czech Republic. ✉email: eva.kriegova@email.cz



FICTION	Fluorescence immunophenotyping and interphase cytogenetics as a tool for investigation of neoplasms
CT	Chromosome territory

Multiple myeloma (MM) is a clonal plasma cell proliferative disorder usually limited to a bone marrow (BM) microenvironment. Rarely, patients present with extramedullary disease (EMM), in which myeloma cells spread to other organ systems<sup>1–3</sup>. This aggressive and mostly treatment-resistant sub-entity of MM can either accompany a newly diagnosed disease, occurring at a frequency of 3–18%<sup>4,5</sup>, or develop with disease progression or relapse, with a frequency of 6–20%<sup>4,6</sup>. Currently, little is known about the mechanisms leading to the development of EMM, stroma-independent growth and the survival of myeloma cells at extramedullary sites or the reasons for poor treatment responses. There is growing evidence that genetic factors may contribute to EMM pathogenesis and evolution<sup>1,4,5</sup>.

Genetic studies have shown that high-risk abnormalities, such as 1q21 gain and del(1p32) (detected in > 55% of EMM patients), t(4;14) (~ 52%), *MYC* overexpression (~ 38%), del(17p13) (~ 35%) and del(13q14) (~ 31%), are commonly associated with EMM<sup>1,4,5</sup>. The disruption of the *TP53* gene by del(17p) and/or mutations seems to be a crucial driver of EMM (EMM vs MM: 34.5% vs 11.9%)<sup>7,8</sup>. Mutations in the *RAS*<sup>9</sup>, *KRAS*, *PIK3CA*, *ATM* and *NFKB2*<sup>1</sup> genes have also been associated with the presence of EMM, including *CRBN* mutations leading to treatment resistance<sup>10</sup>. Other important aberrations in EMM include the activating mutations in the NF- $\kappa$ B pathway genes and the homozygous deletion of the genes encoding inhibitors of this pathway<sup>11</sup>. The resulting constitutive activation of NF- $\kappa$ B enhances the expression of adhesion molecules, such as integrin VLA-4, CD-44, P-selectin and numerous chemokines/receptors<sup>6,12</sup>, leading to the migration and stroma-independent growth of myeloma cells<sup>11</sup>. Additional genetic aberrations may occur in patients with extramedullary mass due to clonal evolution<sup>7,13</sup>. However, the complex genetic architecture in MM and EMM is still poorly understood, likely due to its complexity and heterogeneity.

Therefore, we applied novel whole-genome optical mapping to investigate the complex genomic architecture of BM myeloma cells in newly diagnosed MM and EMM patients. This method has an advantage in detecting small and large structural rearrangements as well as complex rearrangements across the whole genome that are undetectable by traditional methods, such as sequencing and cytogenetics<sup>14</sup>. The characterisation of genetic architecture in EMM could significantly contribute to the understanding of EMM pathogenesis with the potential to discover new prognostic and diagnostic biomarkers and improve the outcome of this MM entity. Moreover, a comparison of MM and EMM may help to elucidate genetic events, allowing the dissemination of myeloma cells from BM to blood and distant tissues.

## Materials and methods

**Subject enrolment.** BM aspirates were obtained from an unselected cohort of 11 newly diagnosed MM patients with EMM presentation (n = 4; median age: 77 years, min–max: 51–79; M/F: 3/1) and without EMM (MM, n = 7; 75 years, 62–82; 5/2). Patients were diagnosed according to the International Myeloma Working Group criteria<sup>15</sup>. The only criteria for patient enrolment were sampling at diagnosis and a sufficient number of sorted cells to perform all genetic analyses ( $\geq 2$  million myeloma cells). In our patients, all EMM sites were bone related, with two in the thoracic spine and two in the pelvis (one in the iliac bone and one in the acetabulum). Patient's clinical and demographic data are summarised in Table 1 and Table S1. For all patients, karyotype, FISH (fluorescence in situ hybridization, Table S2), arrayCGH (Table S3) and next-generation sequencing (NGS) for mutations in the *TP53*, *KRAS*, *NRAS* and *BRAF* genes (Table S2) were available.

All patients provided written informed consent about the usage of BM for this study, which was performed in accordance with the Helsinki Declaration and approved by the ethics committee of the University Hospital and Palacký University Olomouc.

**Collection of BM aspirates.** BM aspirates (2.5–10 ml) were collected in a 5 ml RPMI-1640 medium (Sigma-Aldrich, MO, USA) containing 5000 IU/ml heparin (Zentiva, Prague, Czech Republic). BM mononuclear cells (BMMCs) were collected after red blood cell lysis (155 mM NH<sub>4</sub>Cl, 10 mM KHCO<sub>3</sub>, 0.1 mM Na<sub>2</sub>EDTA, pH 7.3) by centrifugation (1000g, 5 min). After washing with phosphate-buffered saline containing 0.5 M EDTA (Sigma-Aldrich) and 2% FBS (Thermo Fisher Scientific, MA, USA), the total count of BMMCs and the infiltration of CD138+ cells were determined by BD FACSCanto II (BD Biosciences, CA, USA). CD138+ plasma cells were enriched using an EasySep Human CD138 positive Selection Kit II (STEMCELL Technologies, Vancouver, Canada), according to the manufacturer's instructions. The enriched myeloma cells were quantified by BD FACSCanto II (BD Biosciences, CA, USA) using a combination of CD19/CD38/CD45/CD56/CD138 antibodies (BioLegend, CA, USA). After centrifugation (2000g, 2 min), dry pellets of 0.6–2.5 million myeloma cells were stored at –80 °C for further analysis.

**Isolation of high molecular weight DNA, labelling and analysis.** Frozen myeloma cell pellets were processed following the Bionano Prep SP Frozen Cell Pellet DNA Isolation Protocol<sup>16</sup>. High molecular weight (HMW) genomic DNA was isolated using the SP Blood and Cell Culture DNA Isolation Kit (Bionano Genomics, CA, USA, #80030), according to the manufacturer's recommendations. DNA quantification was performed using the Qubit dsDNA BR assay kit (Thermo Fisher Scientific) with a Qubit 2.0 Fluorometer (Thermo Fisher Scientific).

A total of 750–1000 ng of HMW DNA was then labelled using the Bionano Prep Direct Label and Stain DLS DNA Kit (Bionano Genomics, #80005), according to the manufacturer's protocol<sup>17</sup>. The HMW-labelled DNA (within the recommended range of 8–25 labels/100 kbp) was loaded into the Saphyr Chip (Bionano Genomics,

Clinical features	All patients (n = 11)	EMM (n = 4)	MM (n = 7)
Male/female	8/3	3/1	5/2
Age (years), median (min–max)	77 (51–82)	77 (51–79)	75 (62–82)
<b>ISS staging, n (%)</b>			
ISS I	5 (45.5)	2 (50.0)	3 (42.9)
ISS II	1 (9.1)	0 (0.0)	1 (14.3)
ISS III	5 (45.5)	2 (50.0)	3 (42.9)
<b>Durie-Salmon stage, n (%)</b>			
IA	1 (9.1)	0 (0.0)	1 (14.3)
IIA	4 (36.4)	1 (25.0)	3 (42.9)
IIIA	4 (36.4)	2 (50.0)	2 (28.6)
IB	1 (9.1)	0 (0.0)	1 (14.3)
IIIB	1 (9.1)	1 (25.0)	0 (0.0)
<b>LC + FLC, n (%)</b>			
IgG kappa	5 (45.5)	3 (75.0)	2 (28.6)
IgA kappa	4 (36.4)	0 (0.0)	4 (57.1)
IgA lambda	2 (18.2)	1 (25.0)	1 (14.3)
<b>Cytogenetic analysis<sup>a</sup>, n (%)</b>			
t(4;14)	1 (9.1)	0 (0.0)	1 (14.3)
t(11;14)	1 (9.1)	0 (0.0)	1 (14.3)
Gain (1q21)	6 (54.5)	2 (50.0)	4 (57.1)
del(13q14)	4 (36.4)	1 (25.0)	3 (42.8)
del(1p32)	1 (9.1)	1 (25.0)	0 (0.0)
del(17p)	0 (0.0)	0 (0.0)	0 (0.0)
Monosomy	5 (45.5)	2 (50.0)	3 (42.8)
Trisomy	9 (81.8)	4 (100.0)	5 (71.4)
Tetrasomy	5 (45.5)	3 (75.0)	2 (28.6)
<b>NGS analysis<sup>b</sup>, n (%)</b>			
<i>TP53</i>	0 (0.0)	0 (0.0)	0 (0.0)
<i>KRAS</i>	2 (18.2)	1 (25.0)	1 (14.3)
<i>NRAS</i>	0 (0.0)	0 (0.0)	0 (0.0)
<i>BRAF</i>	3 (27.3)	1 (25.0)	2 (28.6)

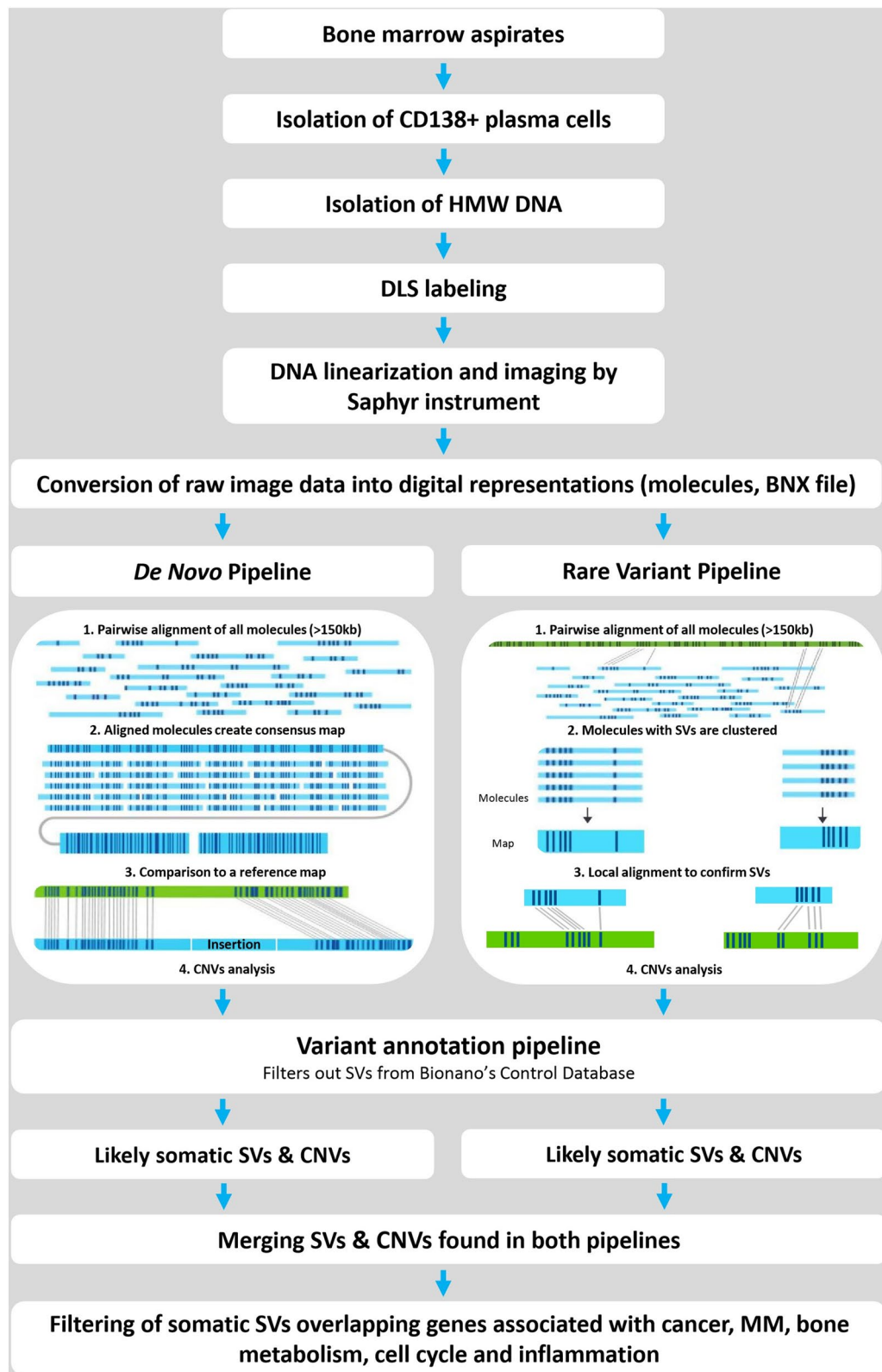
**Table 1.** Basic demographic and clinical characteristics of enrolled MM and EMM patients. ISS International Staging System, LC monoclonal protein's light chain, FLC free light chain. <sup>a</sup>10% positive cut-off level used. <sup>b</sup>The full coding sequence of the *TP53* gene (exons 2–11, plus 5' and 3'UTR; NM\_000546) and the hotspot regions in *NRAS* (exons 2–4; NM\_002524), *KRAS* (exons 2–4; NM\_004985) and *BRAF* (exons 11 and 15; NM\_004333) were sequenced.

#20319) flow cell at a concentration of 4–12 ng/μl and analysed using a Bionano Saphyr instrument, according to the manufacturer's instructions<sup>18</sup>, targeting 100–300× human genome coverage by collecting 500–1300 GB of data per sample.

**Data assembly, structural variant calling and the identification of breakpoint regions.** All data were analysed using Bionano Access software (v1.5) containing the Bionano Solve tool (v3.5) and featuring both de novo and rare variant bioinformatics pipelines (Fig. 1), according to the manufacturer's recommendations<sup>1–21</sup>. Only DNA molecules with a minimum length of 150 kbp were used for bioinformatics analysis along with a minimum of nine labels per molecule.

Briefly, the de novo pipeline's first assembly of all single molecules was based on the distinct distribution of sequence labels by pairwise alignment. The aligned molecules created consensus maps (contigs) in de novo genome maps, which were compared with the in silico DLE1 labelled human hg38 reference map. This pipeline revealed structural variants (SVs) from 500 bp to tens of Mbp long. In the rare-variant pipeline, all single molecules were pairwise aligned against the hg38 reference assembly; molecules with SVs were clustered, and the obtained maps were locally aligned to the hg38 reference sequence. This pipeline was sensitive enough to detect SVs from 5 kbp to tens of Mbp long at a variant allele frequency (VAF) as low as 5%. SVs were considered subclonal (i.e. low-allele frequency) when VAF was ≤ 25% and clonal (i.e. high-allele frequency) when VAF was > 25%, based on a cut-off value for neutral evolution in MM<sup>11</sup>. Additionally, both pipelines included copy number variation (CNV) analysis to detect the fractional copy number changes and chromosomal aneuploidy events. Specific hg38 masks concealing common structural variation in a human genome, N-base reference gaps and problematic sub-centromeric and sub-telomeric regions were used in both pipelines. To annotate the SV calls





**Figure 1.** Workflow of optical mapping and bioinformatics pipelines used. HMW DNA is isolated from CD138+ plasma cells of BM aspirates and labelled by DLS chemistry in specific sequences across entire genomes. Labelled DNA is loaded on the chip and linearised and visualised in a Saphyr instrument. Images are converted to BNX molecules. The architecture of the bioinformatics pipeline includes two pipelines (de novo and rare variant), constructing optical genome maps and comparing them with a human reference map (hg38), filtering detected variants for somatic SVs and merging data from both pipelines. The last step enables a comparison of the data with the gene panels created from NCBI gene datasets.

that were likely somatic variants, a variant annotation pipeline was applied to filter SVs out of the database of ethnically diverse, mapped control human genomes with no reported disease phenotypes.

In the next step, annotated SVs and CNVs from both pipelines were merged (Fig. 1), including aberrations sized 500 bp–5 Mbp (deletions, insertions, duplications and inversions) as well as inter- and intrachromosomal aberrations larger than 5 Mbp. The intrachromosomal rearrangements with breakpoints at least 5 Mbp apart, e.g. large deletions (supported by copy number loss), insertions (copy number gains) or inversions (no change in CNVs) were called intrachromosomal translocations by the Bionano software (Fig. S1). Only SVs with VAF > 5% and a minimum of ten self-molecules were further analysed in this study. Identified candidate SVs were confirmed by arrayCGH, FISH, breakpoint-specific PCR amplification and/or long-read whole-genome sequencing (TELL-Seq, Universal Sequencing Technology, CA, USA). For a comparison of optical mapping and long-read sequencing data, we developed our own tool, which is available at <http://olgen.cz/en/resources><sup>22</sup>.

Finally, the sample-specific SVs were compared with BED masks generated from the NCBI gene database (<https://www.ncbi.nlm.nih.gov/gene>) for gene panels associated with cancer (created using the keywords cancer, tumour suppressor and oncogene; panel of 10,812 genes), MM (696 genes), bone metabolism (osteolysis, cellular calcium signalling, bone metabolism; 1810 genes), cell cycle (cell signalling, cell division, apoptosis, cell cycle, DNA repair; 9750 genes) and inflammation (inflammation, cell migration, adhesion molecules, cytokine/receptor, chemokine/receptor; 4741 genes).

**NGS mutation assessment.** The full coding sequence of the *TP53* gene (exons 2–11, plus 5' and 3'UTR; NM\_000546) and the hotspot regions in *NRAS* (exons 2–4; NM\_002524), *KRAS* (exons 2–4; NM\_004985) and *BRAF* (exons 11 and 15; NM\_004333) were analysed by targeted, ultra-deep NGS, as reported previously<sup>23,24</sup>. Amplicon-based libraries were sequenced as paired ends on MiSeq (2 × 151 bp, Illumina, CA, USA), with a minimum target read depth of 5000×. The detection limit was set up to 1%, and the variants within 1–3% were confirmed by replication.

**Cytogenetic and molecular cytogenetic analysis.** After culturing the heparinised BM aspirates in the BM medium (Biological Industries, CN, USA) overnight with colcemid (Gibco, Thermo Fisher Scientific), the samples were processed as reported previously<sup>25</sup>, and at least ten metaphases were karyotyped. A combination of FISH with immunophenotyping, called fluorescence-immunophenotyping and interphase cytogenetics as a tool for investigation of neoplasms (FICTION), was used to assess the cytogenetic abnormalities using the following probes: LSI RB1 (Abbott Molecular, IL, USA), SPEC IGH, SPEC CKS1B/CDKN2C, TP53/c17, CCND1/IGH, FGFR3/IGH (Zytovision, Bremerhaven, Germany), XL MAF/IGH, CCND3/IGH, MAFB/IGH (MetaSystems, Altlußheim, Germany) and centromeric probes for chromosomes 7, 9, 11 and 15 (Cytocell, Cambridge, United Kingdom), as reported previously<sup>25</sup>. ArrayCGH was performed using SurePrint G3 CGH/CGH + SNP 4 × 180 K microarray (Agilent Technologies, CA, USA)<sup>26</sup>.

**Ethics declarations.** All patients provided written informed consent about the usage of bone marrow samples for this study, which was performed in accordance with the Helsinki Declaration and approved by the ethics committee of the University Hospital Olomouc and Palacký University Olomouc.

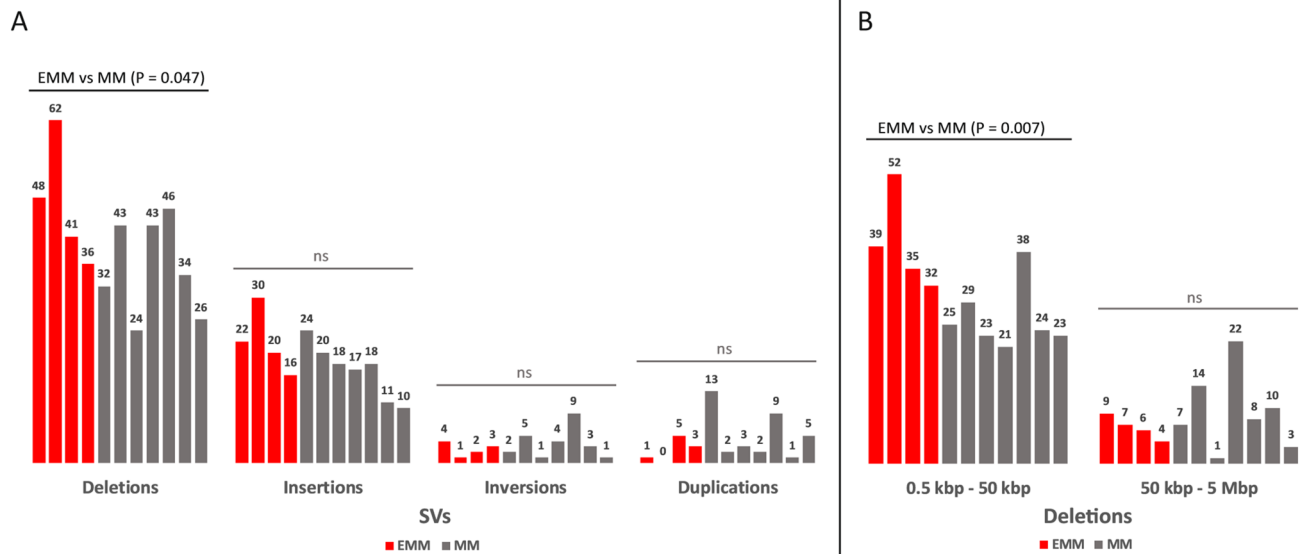
**Consent for publication.** This manuscript has been viewed and approved by all authors for publication.

## Results

**Sample analysis by optical mapping.** The infiltration of myeloma cells in BM aspirates based on immunophenotyping was highly variable in enrolled patients (3–36%); more than 10% infiltration of plasma cells was found in the BM smears of all enrolled patients. The inter-individual variability in the myeloma cell infiltration may be linked to patchy or site-varied myeloma cell distribution, haemodilution, aspirate pull order, the aggregation of myeloma cells in aspirated BM, myeloma cell immunophenotypes and time-dependent losses of surface markers<sup>23</sup>, as well as disease heterogeneity itself<sup>27</sup>. The infiltration of myeloma cells in all samples after enrichment was > 80% (81–96%). Optical mapping was performed in all enriched samples with the following run parameters: average effective coverage, 154× (min–max: 78–324×); collected data per sample, 699 GB (427–1710 GB); DNA molecule size (N50), 316 kbp (219–446 kbp); label density 17.3 labels per 100 kbp (14.1–22.6); and map rate, 74.4% (41.5–93.3%). The quality control parameters for each sample are summarised in Table S4.

**Detection of SVs and CNVs in myeloma samples.** The median number of SVs per patient was as follows: deletions, 1700 (min–max: 1583–1755); insertions, 4433 (4268–4550); inversions, 62 (44–75); duplications, 54 (48–79); chromosome translocations, 2 (0–8); and intrachromosomal rearrangements, 6 (0–24) (Table S5). After filtering only for likely somatic variants, the number of deletions per patient (41, 24–62) dominated over insertions (18, 10–30), inversions (3, 1–9) and duplications (3, 0–13) (Table S5, Fig. 2A), reaching high inter-individual variability. All detected chromosome translocations and intrachromosomal rearrangements were identified as somatic-like in all samples.

The EMM genome contained more deletions than the MM (median number of 45 vs 34,  $P = 0.05$ ), particularly small deletions of 500 bp–50 kbp (37 vs 24,  $P = 0.01$ ) (Fig. 2). The number of inversions and duplications did not differ between EMM and MM ( $P > 0.05$ ). The spectrum of SVs and affected genes and chromosomes displayed high inter-individual variability. In addition to the deletion of the *CCSER1* gene on chromosome 4 found in ~45% of our patients, the SVs in two patients covered *NKAIN2*, and two others covered the *EYS* gene, both within a commonly affected region, 6q.



**Figure 2.** (A) Distribution of SVs (deletions, insertions, inversions and duplications) and (B) deletions subdivided according to their size in EMM (red columns) and MM (grey columns) patients. Each column represents an individual patient and the column height the number of SVs detected.

Regarding CNVs, losses in copy numbers (CN = 1) (median per patient 13, min–max 5–38), as well as gains (CN = 3–25) (37, 4–56), were common in all patients. Except for two MM patients, the majority of patients had a mean of five regions of CN > 3 (range 1–16 per patient) in their genomes. The distribution of CNVs across the genome was highly variable in enrolled EMM and MM patients.

Optical mapping confirmed 98% of SV and CNV changes detected by diagnostic cytogenetic and arrayCGH assessments (Tables S2, Tables S3) and revealed numerous novel rearrangements in all enrolled patients.

**Interchromosomal translocations in MM and EMM.** In three MM patients, optical mapping detected translocations within *IGH/IGK/IGL* immunoglobulin loci, t(4;14) and t(11;14) (confirmed by diagnostic FISH), and one t(8;22)(q24;q11) translocation that was detected by mapping only (this region is not routinely assessed by FISH). In EMM patients, no translocations within *IGH/IGK/IGL* immunoglobulin loci were detected.

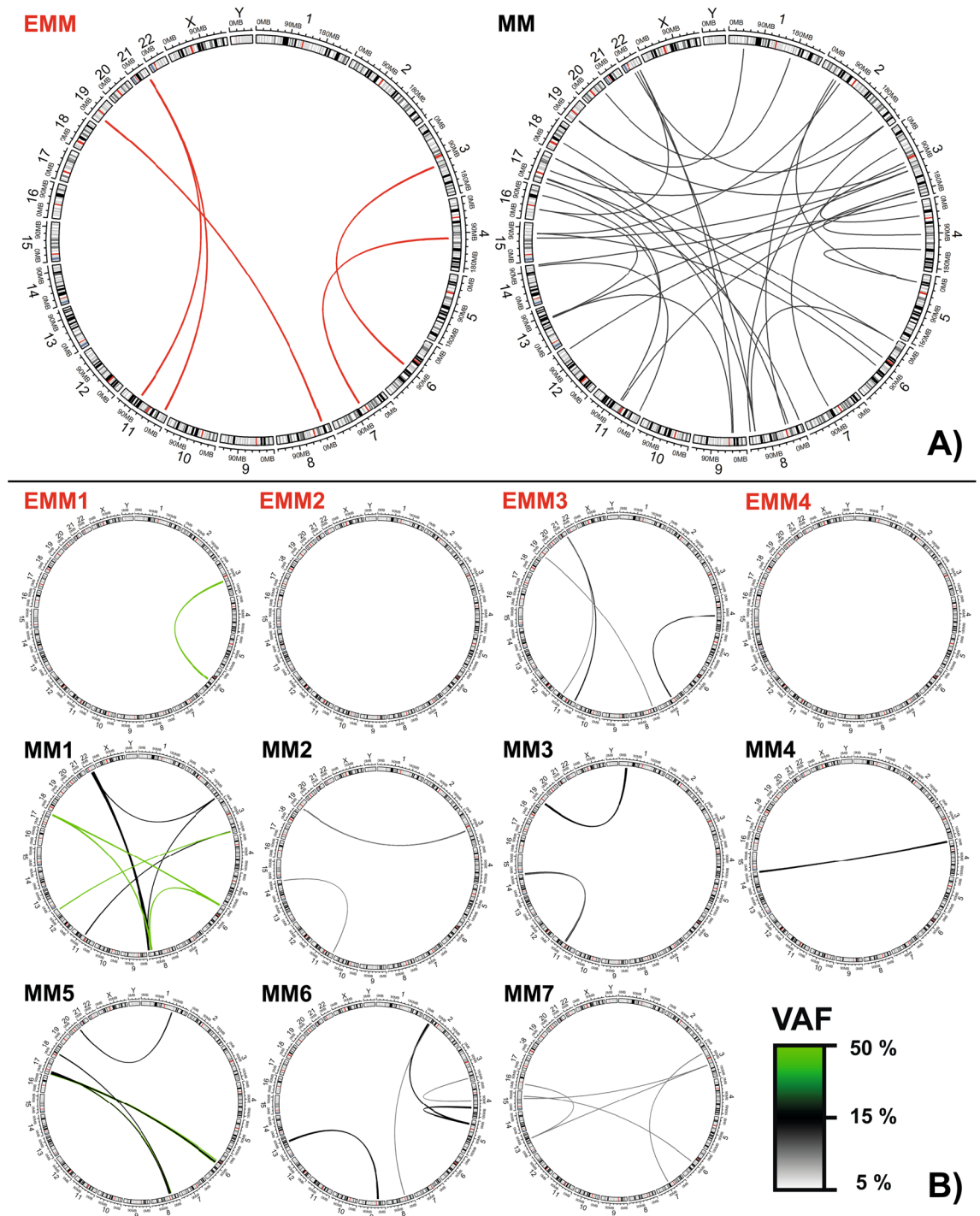
Additionally, numerous other translocations were detected across all MM patients, frequently affecting chromosomes 2, 3, 6 and 8 (Table S6). All MM patients carried at least two translocations, except for one MM patient with only t(4;14) (Table S6, Fig. 3). Complex chromosomal rearrangements involving three chromosomes were detected in four (57%) MM patients but not in any EMM patients (Table S6, Fig. 3). The translocations were present at clonal and subclonal levels (VAF 5–43%). The affected genes and putative fusion genes are shown in Table S6.

EMM genomes were associated with fewer translocations than MM; two EMM patients had no translocations, one EMM patient had one translocation and the only EMM patient that reached complete response after first-line therapy had four translocations. The translocations were present at clonal and subclonal levels (VAF 5–49%) (Fig. 3).

**Intrachromosomal rearrangements in MM and EMM.** Large chromosomal rearrangements encompassing regions longer than 5 Mbp on chromosome 1 were detected in all EMM genomes but not in any MM genomes (Fig. 4, Table 2). The large rearrangements, together with the small SVs (predominantly deletions), affected various regions across chromosome 1, often involving deletions and inversions accompanying the CNV changes. EMM1 had one large intrachromosomal rearrangement of 14.5 Mbp, encompassing 230 genes in the 1p36 region, and five deletions; EMM2 had three large intrachromosomal rearrangements of 47.5 Mbp, 57.9 Mbp and 21.5 Mbp, encompassing 1093 genes in the 1p35-p31, 1p32-p12 and 1p22-p13 regions, and an additional six deletions and one insertion. EMM3 had four rearrangements on chromosome 1 of 7.6 Mbp, 7.5 Mbp, 12.6 Mbp and 12.8 Mbp, encompassing 794 genes in the 1p35-p34, 1p22-p21 and 1p21-p13 regions, and two deletions. EMM4 had two large rearrangements of 36.1 Mbp and 12.0 Mbp, encompassing 564 genes in the 1p34-p31 and 1p34-1q23 regions, three deletions and five insertions (Fig. S1). The majority of the affected genes by intrachromosomal rearrangement across chromosome 1 in EMM were associated with cancer (~35%), cell cycle (~30%) and inflammation (~10%); very few affected genes were associated with MM (~10%) (Table S7).

In contrast, no intrachromosomal rearrangements, fewer deletions (2, 0–4) and more insertions and duplications (4, 0–6) on chromosome 1 were detected in MM compared with EMM. The number of affected genes was also low (2, 0–40).

Additionally, intrachromosomal rearrangements were distributed across other chromosomes in both MM and EMM (Table S8, Fig. S2). The typical patterns of intrachromosomal translocation were large deletions with partial inversion, accompanied by copy number loss. Multiple rearrangements within the same chromosome

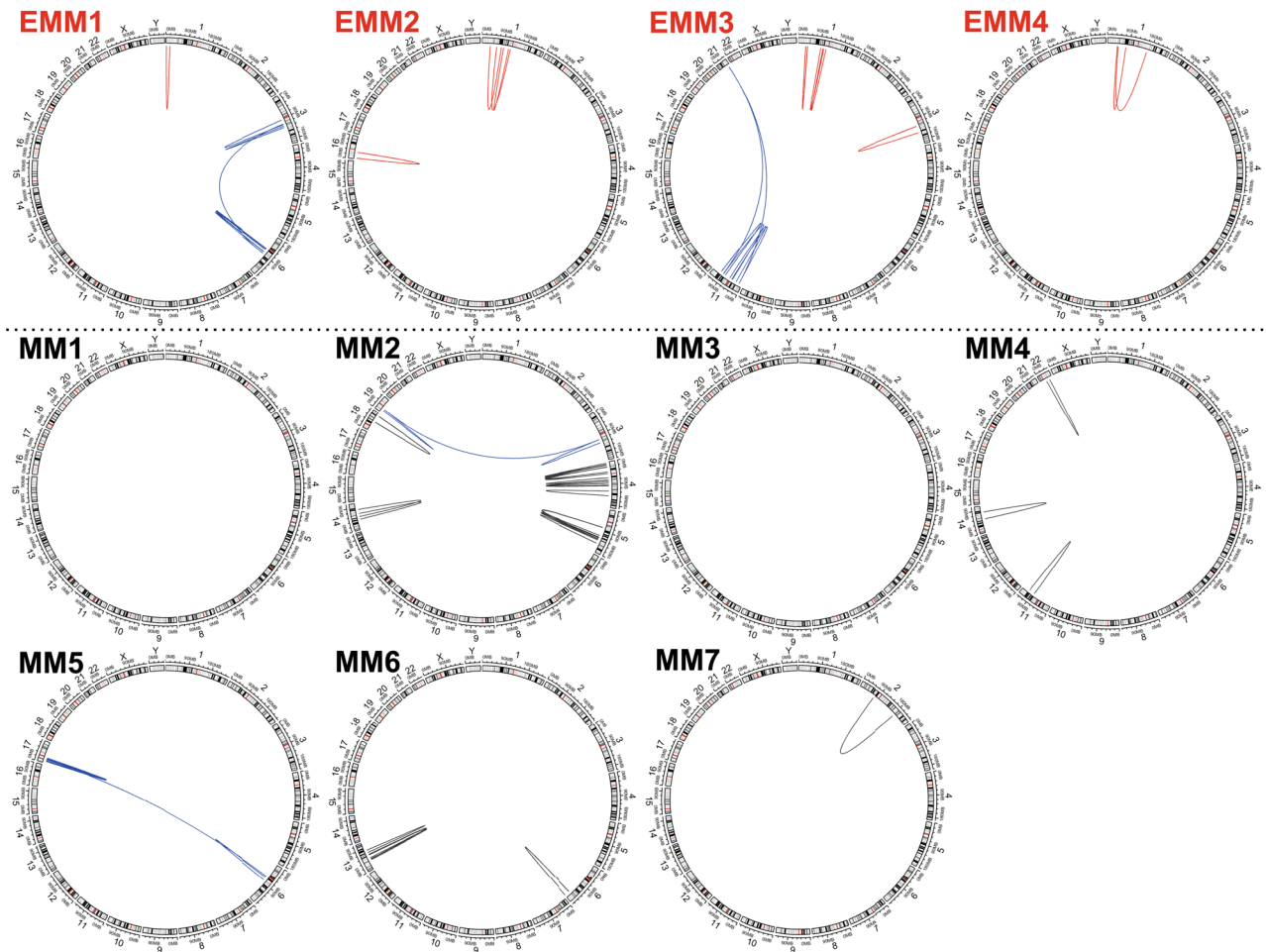


**Figure 3.** Distribution of chromosome translocations in EMM (red lines) and MM (black lines) patients. Large circles plots (A) show the sum of translocation in EMM and MM groups; (B) small circles show detected translocations in a particular patient. The VAF of each translocation is denoted by the thickness and colour of the line (key bottom right). SVs were visualised using circos plots<sup>28</sup>.

often occurred in some patients. In four patients, these rearrangements were part of the interchromosomal translocations (highlighted in blue in Fig. 4).

**SVs and CNVs in high-risk loci associated with MM/EMM.** In addition, we focused on SVs in high-risk regions such as *IGH/IGK/IGL* immunoglobulin loci, del(17p13), del(13q14), the 8q24 region, 1q21 gain and del(1p32).





**Figure 4.** Intrachromosomal rearrangements identified in EMM (red lines) and MM (black lines) patients. Blue lines represent complex rearrangements including translocations.

Regarding the *IGH* locus, optical mapping revealed  $t(4;14)$  and  $t(11;14)$  in three MM, which were confirmed by cytogenetics. In the majority (6/7) of MM samples, but not in any EMM sample, translocations involving immunoglobulin-associated chromosomes 2, 14 and 22 were detected. Additionally, a 0.4 Mbp inversion was detected in one EMM patient, and 1.2 Mbp and 0.8 Mbp duplications on chromosome 14 in two MM patients (Table S9). Also, somatic-like SVs within the *IGK* and *IGL* loci were detected: deletions in five patients (three EMM and two MM), insertions in two (two EMM) and duplication in one (MM) were identified (Table S9, Fig. S3).

Regarding *TP53* disruption, diagnostic analysis by FISH and NGS did not detect any abnormalities in enrolled patients. Nevertheless, optical mapping revealed copy number loss (CN=1) in the region overlapping the *TP53* gene in two EMM patients (Fig. 5).

Optical mapping confirmed  $del(13q14)$  identified by FISH in one EMM and three MM patients. Additionally, optical mapping detected a 1.1 Mbp deletion affecting the *RBI* gene, supported by copy number loss in the 13q14.2 region in one MM patient, which was not detected by FISH (Fig. 5).

Regarding the 8q24 locus, one EMM patient carried a deletion and one MM a duplication, detected by both mapping and cytogenetics. Optical mapping revealed additional changes within this locus associated with *MYC* gene amplification in three patients: one MM patient had a 0.6 Mbp insertion and three translocations,  $t(6;8)$ ,  $t(8;17)$  and  $t(8;22)$ ; one MM had an inversion; and one EMM patient had a novel 0.2 Mbp insertion (Table S10).

Regarding high-risk regions on chromosome 1 commonly affected in MM, we confirmed 1q21 gain in six patients (two EMM and four MM) and  $del(1p32)$  in one EMM patient (Table S11). On chromosome 1, 1.4 times more SVs within/outside the high-risk 1q21/1p32 regions were found in EMM than in MM. In EMM, deletions (50%) and intrachromosomal rearrangements (31%) were the most frequent, not duplications and translocations.

## Discussion

This study characterised genomes of BM myeloma cells in newly diagnosed EMM and MM patients using next-generation optical mapping. When comparing the EMM and MM genomes, EMM was associated with large intrachromosomal rearrangements across chromosome 1, fewer interchromosomal translocations and more deletions across the entire genome compared with MM. For high-risk loci, optical mapping revealed copy number

ID	SV type	SV size (kbp)	Cytobands	Chromosome A		Chromosome B		VAF (%)	Number of affected genes
				Nr	RefStart	Nr	RefEnd		
EMM1	Intra-chrom	14,502	1p36.33-p36.13	1	1451742	1	16010418	8	230
	Deletion	12.3	1p32.3	1	54437326	1	54468426	13	0
	Deletion	0.7	1p31.1	1	73986848	1	73989906	49	0
	Deletion	0.7	1q21.2	1	149255041	1	149297922	12	0
	Deletion	83.9	1q31.1	1	188862967	1	188948998	13	0
	Deletion	0.9	1q41	1	219227885	1	219230864	69	0
EMM2	Intra-chrom	46,876	1p35.1-p31.1	1	33649122	1	80528558	13	475
	Deletion	1,611	1p32.3	1	52146402	1	53764220	6	22
	Deletion	20.0	1p32.2	1	56859220	1	56897293	5	1
	Deletion	311	1p32.2	1	57042204	1	57356850	6	1
	Insertion	0.6	1p32.2	1	57981404	1	57987615	27	1
	Intra-chrom	57,947	1p32.1-p12	1	60494258	1	118466419	9	443
	Deletion	14.5	1p31.1	1	70404062	1	70438713	20	1
	Deletion	5.5	1p31.1	1	70766057	1	70784216	25	0
	Intra-chrom	21,501	1p22.2-p13.3	1	89446367	1	110953619	6	175
	Deletion	7.3	1q25.3	1	182529994	1	182549190	61	1
EMM3	Intra-chrom	7,575	1p35.2-p34.3	1	30221701	1	37801246	15	120
	Intra-chrom	7,465	1p35.2-p34.3	1	30860870	1	38353947	8	126
	Intra-chrom	12,639	1p22.1-p21.1	1	93051297	1	105699124	9	81
	Deletion	1.3	1p22.1	1	93329493	1	93335818	10	1
	Intra-chrom	12,839	1p21.1-p13.2	1	102272272	1	115119846	21	140
	Deletion	5.8	1q42.3	1	235342739	1	235355292	18	1
EMM4	Deletion	53.1	1p36.33	1	1679533	1	1743791	7	3
	Insertion	2.5	1p36.12	1	21983384	1	22006562	13	2
	Intra-chrom	36,076	1p34.2-p31.1	1	39891936	1	75973164	7	363
	Deletion	3,687	1p34.2-p34.1	1	42633541	1	46350766	8	102
	Intra-chrom	119,844	1p34.1-1q23.3	1	44054030	1	163946030	8	1175
	Insertion	2.5	1p13.2	1	111794239	1	111807945	11	1
	Insertion	6.1	1p13.1	1	115530277	1	115543191	14	0
	Deletion	0.8	1q31.3	1	195460336	1	195473816	16	0
	Insertion	18.6	1q32.2	1	207515921	1	207534396	12	2
	Insertion	6.0	1q32.3	1	213171761	1	213205644	12	1
MM1	Insertion	5.9	1p36.33	1	1590522	1	1654114	17	7
	Insertion	59.9	1q21.2	1	149365317	1	149390055	51	0
	Insertion	10.9	1q21.3	1	152289954	1	152296885	13	0
	Deletion	231.9	1q25.1	1	175947709	1	176185749	28	2
	Deletion	3176	1q32.3-q41	1	214386812	1	217572960	21	8
	Insertion	19.9	1q42.12	1	226337005	1	226338164	32	0
	Deletion	315.8	1q43	1	238186340	1	238513554	23	1
	Deletion	31.4	1q43	1	239651938	1	239694630	23	1
MM2	Insertion	5.4	1p36.12	1	20372589	1	20396493	12	1
	Deletion	14.6	1p31.1	1	83171695	1	83186312	19	0
	Duplication	95.5	1q21.2	1	148669395	1	148764931	46	1
	Insertion	2.4	1q23.2	1	161184787	1	161193876	13	1
	Deletion	4.4	1q25.2	1	179360227	1	179368480	28	1
MM3	t(1;19)		1p34.3-19p13.11	1	38291095	19	16838518	15	
	t(1;19)		1p34.3-19p13.12	1	39900616	19	15299192	18	
	Deletion	0.6	1q32.1	1	200212204	1	200225458	24	0
MM4	Insertion	14.6	1p36.31	1	5999446	1	6006996	20	0
	Insertion	1.8	1p12	1	119153160	1	119157652	27	0
	Duplication	95.5	1q21.2	1	148669395	1	148764931	49	1
	Insertion	2.4	1q23.3	1	161184787	1	161193876	28	1

Continued

ID	SV type	SV size (kbp)	Cytobands	Chromosome A		Chromosome B		VAF (%)	Number of affected genes
				Nr	RefStart	Nr	RefEnd		
MM5	Insertion	23.0	1p36.13	1	16040685	1	16054506	9	0
	Deletion	1.8	1p34.2	1	39074813	1	39085202	18	1
	Deletion	5.4	1p31.3	1	62279900	1	62311886	20	1
	Deletion	0.6	1p31.1	1	72996895	1	73015474	24	0
	Duplication	79.6	1q24.2	1	168300624	1	168380191	60	2
	Insertion	199.5	1q24.2	1	168300624	1	168380191	34	2
	t(1;20)		1q24.3-20q13.2	1	170346977	20	53385582	12	
Duplication	57.9	1q42.13	1	227147621	1	227205509	19	1	
MM6	Insertion	3.2	1p34.3	1	37899628	1	37903954	15	1
	Deletion	14.6	1p31.1	1	83171695	1	83190595	17	0
MM7	Insertion	5.7	1p36.33	1	1590522	1	1654114	5	0
	Insertion	3.8	1p36.32	1	4070359	1	4096038	5	0
	Deletion	23.2	1p36.12	1	21983384	1	22006562	17	2
	Insertion	3.4	1p34.1	1	44413885	1	44419138	7	1
	Duplication	47.2	1p21.3	1	99166139	1	99213371	7	0
	Insertion	55.4	1p21.3	1	99180335	1	99199495	6	0
	Deletion	721.6	1q21.1	1	143310164	1	144170341	25	9
	Duplication	633.7	1q22-q23.1	1	156251622	1	156885360	5	28
	Deletion	3.0	1q24.1	1	166768561	1	166777221	16	0
	Deletion	0.7	1q32.1	1	200212204	1	200225458	15	0

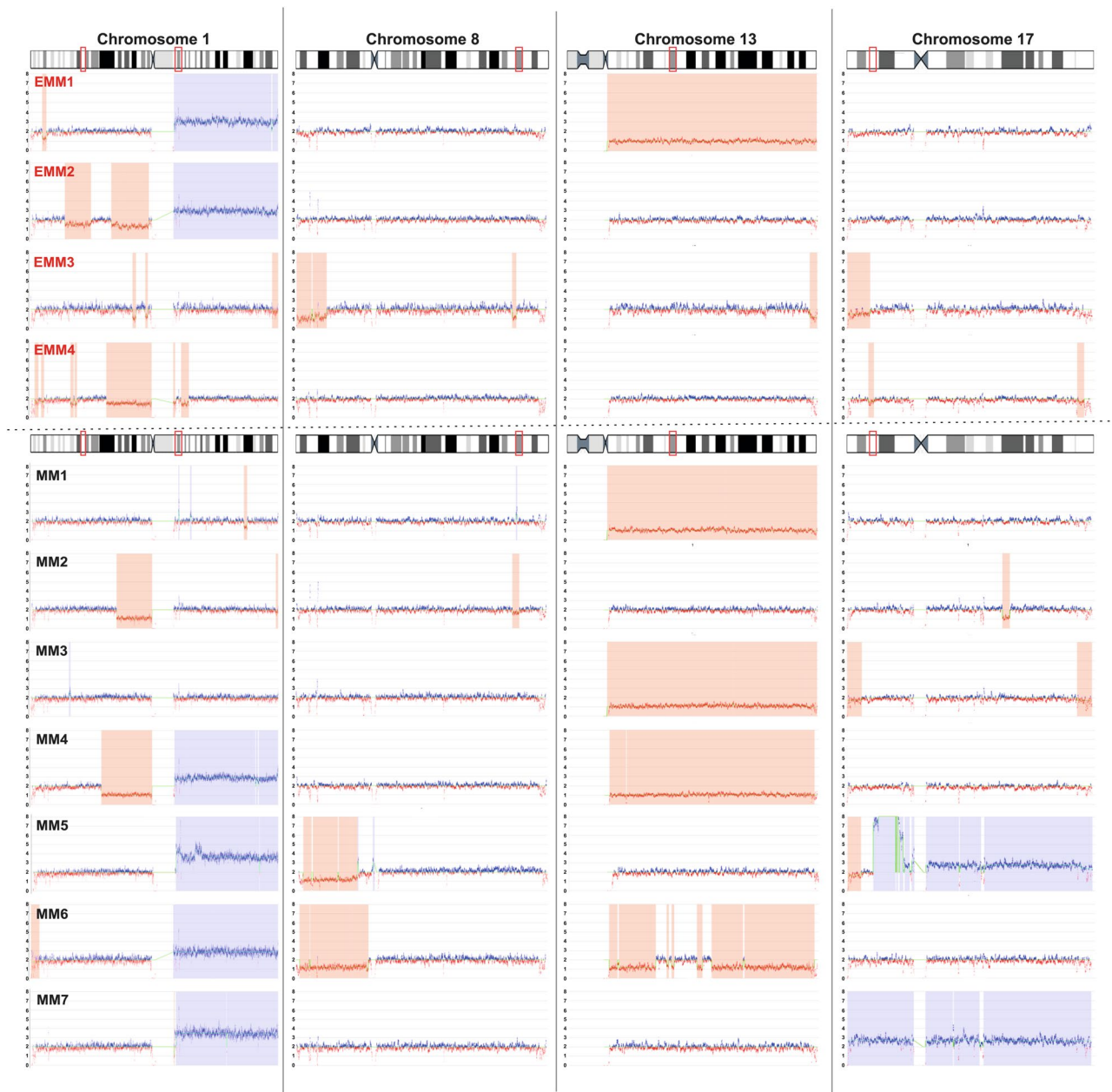
**Table 2.** SVs on chromosome 1 in enrolled EMM and MM patients. *intra-chrom* intrachromosomal rearrangements, *VAF* variant allele frequency, *SVs* structural variants.

loss in the 17p13 region in two EMMs, numerous SVs and CNVs in other high-risk 8q24 and 13q14 regions and *IGH/IGK/IGL* immunoglobulin loci that were not detected by diagnostic cytogenetic evaluation.

To date, the complex genomic architecture in MM and EMM has been poorly characterised, probably due to its complexity, heterogeneity and multiple levels of somatic mosaicism<sup>29,30</sup>. Therefore, we analysed EMM and MM genomes using innovative optical mapping that can detect small SVs and CNVs as well as complex large genomic rearrangements or chained fusions<sup>14,31–33</sup>, which are not recognisable by NGS and/or cytogenetics. The utility of this approach has been recently shown in leukaemia samples, where optical mapping confirmed the results of whole-genome sequencing and/or cytogenetic analysis and additionally revealed a large number of SVs not previously recognisable in analysed samples<sup>14,33</sup>. In this study, we used optical mapping for the first time to study the genome architecture of isolated myeloma cells from BM from newly diagnosed EMM and MM patients.

In line with the high degree of somatic genomic mosaicism and multiple levels of genetic variation in MM<sup>29,30</sup>, long-fragment mapping revealed simple and complex genomic rearrangements and CNVs in all samples. More interchromosomal translocations were detected in MM patients than EMM. Except for one patient with a high-risk 14q32 translocation, a common primary event in MM<sup>34</sup>, all MM patients had at least two other translocations. These often involved chromosomes 2, 3, 6 and 8, and many of them led to gene disruptions or the creation of putative gene fusions with at least one partner associated with cancer. Moreover, interchromosomal translocations in MM were often accompanied by intrachromosomal rearrangements located in the same chromosomal loci. On the contrary, three EMM patients had one or zero translocations, and one EMM patient with good treatment response had three translocations; the translocations occurred in our patients at subclonal and clonal levels (5–49%). Although the impact of clonal status on the prognostic value of SVs is unclear for most cancers, recent NGS studies in MM have shown that the clonality status of mutations does not influence survival but does impact the disease phenotype<sup>35</sup>. Experimental evidence also suggests that MM progression, both spontaneous in asymptomatic stages and at relapse after treatment, is linked to its heterogeneous subclonal composition<sup>36</sup>; thus, the direct measures of the clone size and its intrinsic biological features deserve future investigation. Optical mapping also revealed numerous complex translocations, involving three chromosomes in about half of the MM patients but not in the EMM patients. There are already reports about large chromosomal rearrangements, called chromothripsis, in MM<sup>29,30</sup>. Such complex structural changes, often accompanied by loss of heterozygosity<sup>37</sup>, are difficult to identify by other techniques and may escape attention. The presence of unusual rearrangements of numerous chromosomes in MM, but not EMM, deserves future investigation.

In addition to interchromosomal translocations, we detected numerous intrachromosomal rearrangements, which are rearrangements that involve loci located on the same chromosome. To date, few cancer types harbour both interchromosomal and intrachromosomal rearrangements; one of them is MM<sup>38</sup>. It has been suggested that the occurrence of intra- or interchromosomal recombinations depends on the spatial proximity between recombinogenic partners within the chromosome territories (CTs), a non-randomly formed, distinct space where each chromosome decondenses<sup>39,40</sup>. When loci are situated near the surface of their CTs, interchromosomal translocations occur, and when they are located deep in the CTs, intrachromosomal rearrangements occur<sup>41</sup>. There is already evidence that chromosomes involved in commonly occurring translocations – t(4;14), t(14;16)



**Figure 5.** The genome CNVs on chromosomes 1, 8, 13 and 17 in EMM (upper part) and MM (lower part) patients. Blue indicates gains and red indicates losses in gene copy numbers. The vertical bars represent detected copy number aberrations. The red boxes on the ideogram highlight the high-risk regions.

and t(11;14) – in MM are located within overlapping CTs<sup>42,43</sup>; however, the mechanisms of intrachromosomal rearrangements have not been investigated in MM.

Importantly, we detected EMM-specific intrachromosomal rearrangements encompassing several Mbp-long regions within chromosome 1, commonly including combinations of deletions and inversions and affecting hundreds of genes. These rearrangements were located across the whole of chromosome 1 and led to changes in the copy number of genes on large regions of this chromosome. The intrachromosomal rearrangements on chromosome 1 have already been reported in progressive, multi-drug refractory EMM<sup>10</sup> and EMM with soft tissue involvement at the time of MM diagnosis<sup>44</sup>. Interestingly, 80% (8/10) of patients with soft tissue EMM had chromosome 1 abnormalities, and an association between chromosome 1 abnormalities and soft tissue EMM was suggested<sup>44</sup>. Furthermore, 1p deletion and/or 1q gain were associated with the extramedullary plasmablastic transformation of MM in both BM and matched extramedullary tissue<sup>45</sup>. Other studies reported an association of chromosome 1 abnormalities in MM with the relapsed disease<sup>46</sup>. The affected patients have an exceedingly poor prognosis, short progression-free survival and overall survival, even in the era of novel therapies<sup>44,47,48</sup>. A recent study showed that the adverse impact of chromosome 1 abnormalities on survival is of similar magnitude to other high-risk chromosomal abnormalities<sup>47</sup>. The crucial role of chromosome 1 in MM pathogenesis is also



supported by the significant overrepresentation of genes derived from chromosome 1 in the high-risk signature in MM<sup>48</sup>. The occurrence of fewer interchromosomal translocations and more intrachromosomal rearrangements in EMM, particularly on chromosome 1, suggests that recombinations within loci deep in CTs may play a crucial role in MM pathogenesis, particularly influencing the phenotype of the disease. Furthermore, the observed chromosome 1 abnormalities may play a role on the required events that allow the dissemination of myeloma cells from BM to blood and distant tissues; this also deserves future investigation.

In addition to translocations and intrachromosomal rearrangements, we also detected tens of SVs in every EMM and MM genome. The most common were deletions distributed across all chromosomes. In particular, deletions ranging in size from 500 bp to 50 kbp occurred more frequently in EMM than in MM. An increased number of deletions in MM has already been associated with MM progression, as shown by comparing MM genomes at diagnosis and relapse<sup>49</sup>. Future studies should investigate the relationship of a higher deletion load in EMM compared with MM as well as prognosis. We also detected numerous novel SVs and CNVs within high-risk loci associated with MM not previously detected by sequencing and cytogenetics.

The most critical genetic factors that portend a poor prognosis for MM are translocations within the *IGH/IGK/IGL* loci<sup>50</sup>. Our study confirmed the common translocations t(4;14) and t(11;14) in three MM patients and revealed additional interchromosomal translocations involving chromosomes 2, 14 and 22, where immunoglobulin genes are located, in a majority (6/7) of MM patients. The functional consequences of the translocations outside the *IGH/IGK/IGL* loci needs to be clarified, as they may influence antibody expression and function and the mediation of disease phenotypes. Interestingly, we did not detect any translocation on the previously mentioned chromosomes in EMM, where deletions and inversions were predominantly found. Differences between MM and EMM in genetic rearrangements on immunoglobulin-associated chromosomes should be further investigated.

Next, we were interested in the disruption of the 17p13 locus overlapping the *TP53* gene, a driver aberration associated with EMM<sup>51,52</sup>, poor prognosis and low treatment response rates in MM patients<sup>53</sup>. The loss of *TP53* and other genetic aberrations may additionally occur in the extramedullary mass due to regional clonal evolution, as shown by comparing extramedullary tumours with their BM myeloma cells<sup>7,13</sup>. Although no *TP53* disruption was detected in myeloma cells from BM aspirates of enrolled patients by diagnostic cytogenetic and mutational analyses, optical mapping revealed copy number loss in the 17p13 region in two EMM patients. Our data further support the key role of *TP53* in EMM and emphasise the need to routinely incorporate SVs and CNVs, the major forms of genetic alterations in cancer, at many length scales to understand the MM genome more comprehensively.

Optical mapping also confirmed rearrangements at the *MYC/8q24* locus, a late tumour progression event associated with an increased expression of *MYC* and poor prognosis<sup>54</sup>, in about a third of patients. One MM patient had three translocations within this region, and two others had SVs within the *MYC/8q24* locus. Whether the changes at the *MYC/8q24* locus were EMM specific, as reported by others<sup>55</sup>, needs further investigation using larger cohorts.

This study has several limitations. First, we did not investigate extramedullary tumour mass because invasive biopsy was not feasible in enrolled patients. Second, due to the moderate number of patients included in this exploratory study, a sub-analysis based on clinical and laboratory parameters was not performed. Third, the proportion of the IgA subtype was higher in the MM cohort. However, there is growing evidence that adverse prognosis in patients with IgA MM versus non-IgA MM subtypes is more likely to be caused by the misclassification of disease response or the delayed detection of disease due to an underestimation of tumour burden<sup>56</sup> than changes in expression profile or cytogenetics<sup>57–59</sup>. Future studies on larger patient cohorts enabling a subanalysis of patients with particular clinical characteristics and stages of disease and the investigation of extramedullary tissue sites are warranted.

There is a growing body of evidence on the utility of optical mapping for comprehensive SV detection in haematology and solid tumours<sup>14,31–33</sup>. The aberrations detected by mapping have been confirmed by cytogenetics<sup>14</sup> or NGS<sup>33,60</sup>, particularly by long-read sequencing<sup>61,62</sup>, as also shown in our study. Optical mapping thus provides an ideal complement to sequencing for resolving complex genomic architecture in cancers<sup>31</sup>.

## Conclusion

Our pilot study using next-generation optical mapping revealed that in addition to known high-risk cytogenetic factors, chromosome 1 abnormalities in BM myeloma cells are associated with extramedullary progression. The detection of numerous novel, distinct genetic aberrations associated with EMM and MM shows the potential of optical mapping for the refinement of complex genomic architecture in MM and its phenotypes. The methodology and results described here represent a significant advance that may accelerate the introduction of genomics at long-length scales into clinical decisions for MM.

## Data availability

The data of this study are available from the corresponding author on reasonable request.

Received: 1 February 2021; Accepted: 24 June 2021

Published online: 19 July 2021

## References

1. Bhutani, M., Foureau, D. M., Atrash, S., Voorhees, P. M. & Usmani, S. Z. Extramedullary multiple myeloma. *Leukemia* **34**, 1–20 (2020).
2. Paquin, A. R. *et al.* Overall survival of transplant eligible patients with newly diagnosed multiple myeloma: Comparative effectiveness analysis of modern induction regimens on outcome. *Blood Cancer J.* **8**, 125 (2018).
3. Usmani, S. Z. *et al.* Extramedullary disease portends poor prognosis in multiple myeloma and is over-represented in high-risk disease even in the era of novel agents. *Haematologica* **97**, 1761–1767 (2012).

4. Qu, X. *et al.* Extramedullary manifestation in multiple myeloma bears high incidence of poor cytogenetic aberration and novel agents resistance. *Biomed. Res. Int.* **2015**, 787809 (2015).
5. Jagosky, M. H. & Usmani, S. Z. Extramedullary disease in multiple myeloma. *Curr. Hematol. Malig. Rep.* **15**, 62–71 (2020).
6. Blade, J. *et al.* Soft-tissue plasmacytomas in multiple myeloma: Incidence, mechanisms of extramedullary spread, and treatment approach. *J. Clin. Oncol.* **29**, 3805–3812 (2011).
7. Billecke, L. *et al.* Cytogenetics of extramedullary manifestations in multiple myeloma. *Br. J. Haematol.* **161**, 87–94 (2013).
8. Varga, C. *et al.* Development of extramedullary myeloma in the era of novel agents: No evidence of increased risk with lenalidomide-bortezomib combinations. *Br. J. Haematol.* **169**, 843–850 (2015).
9. de Haart, S. J. *et al.* Comparison of intramedullary myeloma and corresponding extramedullary soft tissue plasmacytomas using genetic mutational panel analyses. *Blood Cancer J.* **6**, e426 (2016).
10. Egan, J. B. *et al.* Extramedullary myeloma whole genome sequencing reveals novel mutations in Cereblon, proteasome subunit G2 and the glucocorticoid receptor in multi drug resistant disease. *Br. J. Haematol.* **161**, 748–751 (2013).
11. Furukawa, Y. & Kikuchi, J. Molecular basis of clonal evolution in multiple myeloma. *Int. J. Hematol.* **111**, 496–511 (2020).
12. Dahl, I. M. S., Rasmussen, T., Kauric, G. & Husebekk, A. Differential expression of CD56 and CD44 in the evolution of extramedullary myeloma. *Br. J. Haematol.* **116**, 273–277 (2002).
13. Rasche, L. *et al.* Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. *Nat. Commun.* **8**, 268 (2017).
14. Neveling, K. *et al.* Next generation cytogenetics: comprehensive assessment of 48 leukemia genomes by genome imaging. *bioRxiv*. <https://doi.org/10.1101/2020.02.06.935742> (2020).
15. Rajkumar, S. V. *et al.* International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol.* **15**, e538–e548 (2014).
16. Bionano Genomics. Bionano prep SP fresh cells dna isolation protocol (revision D). Document Number: 30257. <https://bionanogenomics.com/wp-content/uploads/2019/04/30257-Bionano-Prep-SP-Fresh-Cells-DNA-Isolation-Protocol.pdf> (2020).
17. Bionano Genomics. Bionano prep Direct Label and Stain (DLS) Protocol (revision F). Document Number: 30206. <https://bionanogenomics.com/wp-content/uploads/2018/04/30206-Bionano-Prep-Direct-Label-and-Stain-DLS-Protocol.pdf> (2019).
18. Bionano Genomics. Saphyr system user guide (revision C). Document Number: 30143. <https://bionanogenomics.com/wp-content/uploads/2017/10/30143-Saphyr-System-User-Guide.pdf> (2018).
19. Bionano Genomics. Introduction to copy number analysis (revision D). Document Number: 30210. <https://bionanogenomics.com/wp-content/uploads/2018/04/30210-Introduction-to-Copy-Number-Analysis.pdf> (2019).
20. Bionano Genomics. Bionano solve theory of operation: structural variant calling (revision J). Document Number: 30110. <https://bionanogenomics.com/wp-content/uploads/2018/04/30110-Bionano-Solve-Theory-of-Operation-Structural-Variant-Calling.pdf> (2020).
21. Bionano Genomics. Bionano solve theory of operation: variant annotation pipeline (revision H). Document Number: 30190. <https://bionanogenomics.com/wp-content/uploads/2018/04/30190-Bionano-Solve-Theory-of-Operation-Variant-Annotation-Pipeline.pdf> (2020).
22. Savara, J., Novosád, T., Gajdoš, P. & Kriegova, E. Comparison of structural variants detected by optical mapping with long-read next-generation sequencing. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btab359> (2021).
23. Petrackova, A. *et al.* Diagnostic deep-targeted next-generation sequencing assessment of TP53 gene mutations in multiple myeloma from the whole bone marrow. *Br. J. Haematol.* **189**, e122–e125 (2020).
24. Obr, A. *et al.* TP53 mutation and complex karyotype portends a dismal prognosis in patients with mantle cell lymphoma. *Clin. Lymphoma Myeloma Leuk.* **18**, 762–768 (2018).
25. Mlynarcikova, M. *et al.* Molecular cytogenetic analysis of chromosome 8 aberrations in patients with multiple myeloma examined in 2 different stages, at diagnosis and at progression/relapse. *Clin. Lymphoma Myeloma Leuk.* **16**, 358–365 (2016).
26. Kruzova, L. *et al.* Complex karyotype as a predictor of high-risk chronic lymphocytic leukemia: A single center experience over 12 years. *Leuk. Res.* **85**, 106218 (2019).
27. Lee, N. *et al.* Discrepancies between the percentage of plasma cells in bone marrow aspiration and BM biopsy: Impact on the revised IMWG diagnostic criteria of multiple myeloma. *Blood Cancer J.* **7**, e530 (2017).
28. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
29. Ye, C. J., Chen, J., Liu, G. & Heng, H. H. Somatic genomic mosaicism in multiple myeloma. *Front. Genet.* **11**, 388 (2020).
30. Lohr, J. G. *et al.* Widespread genetic heterogeneity in multiple myeloma: Implications for targeted therapy. *Cancer Cell* **25**, 91–101 (2014).
31. Chan, E. K. F. *et al.* Optical mapping reveals a higher level of genomic architecture of chained fusions in cancer. *Genome Res.* **28**, 726–738 (2018).
32. Mantere, T. *et al.* Next generation cytogenetics: Genome-imaging enables comprehensive structural variant detection for 100 constitutional chromosomal aberrations in 85 samples. *bioRxiv*. <https://doi.org/10.1101/2020.07.15.205245> (2020).
33. Xu, J. *et al.* An integrated framework for genome analysis reveals numerous previously unrecognizable structural variants in leukemia patients' samples. *bioRxiv*. <https://doi.org/10.1101/563270> (2019).
34. Walker, B. A. *et al.* Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. *Blood* **121**, 3413–3419 (2013).
35. Bolli, N. *et al.* Analysis of the genomic landscape of multiple myeloma highlights novel prognostic markers and disease subgroups. *Leukemia* **32**, 2604–2616 (2018).
36. Bolli, N. *et al.* Next-generation sequencing for clinical management of multiple myeloma: Ready for prime time?. *Front. Oncol.* **10**, 189 (2020).
37. Berry, N. K., Bain, N. L., Enjeti, A. K. & Rowlings, P. Genomic profiling of plasma cell disorders in a clinical setting: Integration of microarray and FISH, after CD138 selection of bone marrow. *J. Clin. Pathol.* **67**, 66–69 (2014).
38. Walker, B. A. Whole exome sequencing in multiple myeloma to identify somatic single nucleotide variants and key translocations involving immunoglobulin loci and MYC. *Methods Mol. Biol.* **1792**, 71–95 (2018).
39. Schardin, M., Cremer, T., Hager, H. D. & Lang, M. Specific staining of human chromosomes in Chinese hamster × man hybrid cell lines demonstrates interphase chromosome territories. *Hum. Genet.* **71**, 281–287 (1985).
40. Parada, L. & Misteli, T. Chromosome positioning in the interphase nucleus. *Trends Cell Biol.* **12**, 425–432 (2002).
41. Gandhi, M. S., Stringer, J. R., Nikiforova, M. N., Medvedovic, M. & Nikiforov, Y. E. Gene position within chromosome territories correlates with their involvement in distinct rearrangement types in thyroid cancer cells. *Genes Chromosom. Cancer.* **48**, 222–228 (2009).
42. Sathitruangsak, C. *et al.* Distinct and shared three-dimensional chromosome organization patterns in lymphocytes, monoclonal gammopathy of undetermined significance and multiple myeloma. *Int. J. Cancer.* **140**, 400–410 (2017).
43. Martin, L. D., Harizanova, J., Mai, S., Belch, A. R. & Pilarski, L. M. FGFR3 preferentially colocalizes with IGH in the interphase nucleus of multiple myeloma patient B-cells when FGFR3 is located outside of CT4. *Genes Chromosom. Cancer.* **55**, 962–974 (2016).
44. Neparidze, N. & Brown, J. E. Clinical outcomes of extramedullary multiple myeloma in the era of novel agents. *Blood* **130**, 5438 (2017).

45. Liu, Y. *et al.* Genetic basis of extramedullary plasmablastic transformation of multiple myeloma. *Am. J. Surg. Pathol.* **44**, 838–848 (2020).
46. Marzin, Y. *et al.* Chromosome 1 abnormalities in multiple myeloma. *Anticancer Res.* **26**, 953–959 (2006).
47. Giri, S. *et al.* Chromosome 1 abnormalities and survival of patients with multiple myeloma in the era of novel agents. *Blood Adv.* **4**, 2245–2253 (2020).
48. Shaughnessy, J. D. Jr. *et al.* A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* **109**, 2276–2284 (2007).
49. Gupta, A. *et al.* Single-molecule analysis reveals widespread structural variation in multiple myeloma. *Proc. Natl. Acad. Sci. USA* **112**, 7689–7694 (2015).
50. Barwick, B. G. *et al.* Multiple myeloma immunoglobulin lambda translocations portend poor prognosis. *Nat. Commun.* **10**, 1911 (2019).
51. Deng, S. *et al.* Features of extramedullary disease of multiple myeloma: high frequency of p53 deletion and poor survival: A retrospective single-center study of 834 cases. *Clin. Lymphoma Myeloma Leuk.* **15**, 286–291 (2015).
52. Katodritou, E. *et al.* Extramedullary (EMP) relapse in unusual locations in multiple myeloma: Is there an association with precedent thalidomide administration and a correlation of special biological features with treatment and outcome?. *Leuk. Res.* **33**, 1137–1140 (2009).
53. Dimopoulos, M. A. *et al.* Treatment of patients with relapsed/refractory multiple myeloma with lenalidomide and dexamethasone with or without bortezomib: Prospective evaluation of the impact of cytogenetic abnormalities and of previous therapies. *Leukemia* **24**, 1769–1778 (2010).
54. Misund, K. *et al.* MYC dysregulation in the progression of multiple myeloma. *Leukemia* **34**, 322–326 (2020).
55. Szabo, A. G. *et al.* Overexpression of c-myc is associated with adverse clinical features and worse overall survival in multiple myeloma. *Leuk. Lymphoma.* **57**, 2526–2534 (2016).
56. Visram, A. *et al.* Disease monitoring with quantitative serum IgA levels provides a more reliable response assessment in multiple myeloma patients. *Leukemia* **35**, 1428–1437 (2021).
57. Fonseca, R. *et al.* Clinical and biologic implications of recurrent genomic aberrations in myeloma. *Blood* **101**, 4569–4575 (2003).
58. Muddasani, R. *et al.* Association between immunoglobulin isotypes and cytogenetic risk groups in multiple myeloma. *Blood* **132**, 5585 (2018).
59. Nair, B. *et al.* Immunoglobulin isotypes in multiple myeloma: Laboratory correlates and prognostic implications in total therapy protocols. *Br. J. Haematol.* **145**, 134–137 (2009).
60. Jaratlersiri, W. *et al.* Next generation mapping reveals novel large genomic rearrangements in prostate cancer. *Oncotarget* **8**, 23588–23602 (2017).
61. Deschamps, S. *et al.* A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping. *Nat. Commun.* **9**, 4844 (2018).
62. Weissensteiner, M. H. *et al.* Combination of short-read, long-read, and optical mapping assemblies reveals large-scale tandem repeat arrays with population genetic implications. *Genome Res.* **27**, 697–708 (2017).

## Acknowledgements

We would like to thank Petra Sindelarova for administrative assistance.

## Author contributions

E.K., T.Pa. conceived the study, E.K., R.F. planned the experiments, interpreted the data, wrote the manuscript. R.F., J.Ma., M.D., A.P. performed the laboratory analysis. J.B. performed the cytogenetic analysis, A.P. performed NGS analysis, J.Mi., P.K., T.Pi., T.Pa. collected the patient samples and clinical data, J.S., P.G., M.B., M.V. performed the bioinformatics analysis. J.Mi., T.Pa., A.P. revised the manuscript. All authors read the final version of manuscript.

## Funding

The study was supported by research Grant Celgene (CZE\_102), Internal Grant Agency of Palacky University (IGA\_LF\_2021\_015), Ministry of Health of Czech Republic—RVO (FNOL, 00098892), and in part by NV18-03-00500.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-93835-z>.

**Correspondence** and requests for materials should be addressed to E.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

## APPENDIX H

Malarikova D, Berkova A, Obr A, Blahovcova P, Svaton M, Forsterova K, Kriegova E, Prihodova E, Pavlistova L, **Petrackova A**, Zemanova Z, Trneny M, Klener P. Concurrent TP53 and CDKN2A gene aberrations in newly diagnosed mantle cell lymphoma correlate with chemoresistance and call for innovative upfront therapy. *Cancers (Basel)*. 2020;12(8):2120. (IF 2020: 6.639; Q1)

Article

# Concurrent *TP53* and *CDKN2A* Gene Aberrations in Newly Diagnosed Mantle Cell Lymphoma Correlate with Chemoresistance and Call for Innovative Upfront Therapy

Diana Malarikova <sup>1,2</sup>, Adela Berkova <sup>1,3</sup>, Ales Obr <sup>4</sup> , Petra Blahovcova <sup>1</sup>, Michael Svaton <sup>5</sup> , Kristina Forsterova <sup>1</sup>, Eva Kriegova <sup>4</sup> , Eva Prihodova <sup>3</sup>, Lenka Pavlistova <sup>3</sup>, Anna Petrackova <sup>4</sup>, Zuzana Zemanova <sup>3</sup>, Marek Trneny <sup>1</sup> and Pavel Klener <sup>1,2,\*</sup> 

<sup>1</sup> First Department of Internal Medicine-Hematology, General University Hospital and First Faculty of Medicine, Charles University in Prague, 12808 Prague, Czech Republic; diana.tuskova@vfn.cz (D.M.); adela.berkova@vfn.cz (A.B.); blahovcova@lymphoma.cz (P.B.); kristina.forsterova@lf1.cuni.cz (K.F.); trneny@cesnet.cz (M.T.)

<sup>2</sup> Institute of Pathological Physiology, First Faculty of Medicine, Charles University in Prague, 12853 Prague, Czech Republic

<sup>3</sup> Center of Oncocytogenomics, Institute of Clinical Biochemistry and Laboratory Diagnostics, General University Hospital and First Faculty of Medicine, Charles University in Prague, 12853 Prague, Czech Republic; eva.prihodova@vfn.cz (E.P.); Lenka.Pavlistova@vfn.cz (L.P.); zuzana.zemanova@vfn.cz (Z.Z.)

<sup>4</sup> Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University and University Hospital Olomouc, 77515 Olomouc, Czech Republic; Ales.Obr@fnol.cz (A.O.); Eva.Kriegova@fnol.cz (E.K.); anna.petrackova@gmail.com (A.P.)

<sup>5</sup> CLIP, Department of Pediatric Hematology/Oncology, Second Faculty of Medicine and University Hospital in Motol, 15006 Prague, Czech Republic; michael.svaton@lfmotol.cuni.cz

\* Correspondence: pavel.klener2@lf1.cuni.cz

Received: 24 June 2020; Accepted: 27 July 2020; Published: 31 July 2020



**Abstract:** Mantle cell lymphoma (MCL) is a subtype of B-cell lymphoma with a large number of recurrent cytogenetic/molecular aberrations. Approximately 5–10% of patients do not respond to frontline immunochemotherapy. Despite many useful prognostic indexes, a reliable marker of chemoresistance is not available. We evaluated the prognostic impact of seven recurrent gene aberrations including tumor suppressor protein P53 (*TP53*) and cyclin dependent kinase inhibitor 2A (*CDKN2A*) in the cohort of 126 newly diagnosed consecutive MCL patients with bone marrow involvement  $\geq 5\%$  using fluorescent in-situ hybridization (FISH) and next-generation sequencing (NGS). In contrast to *TP53*, no pathologic mutations of *CDKN2A* were detected by NGS. *CDKN2A* deletions were found exclusively in the context of other gene aberrations suggesting it represents a later event (after translocation t(11;14) and aberrations of *TP53*, or ataxia telangiectasia mutated (*ATM*)). Concurrent deletion of *CDKN2A* and aberration of *TP53* (deletion and/or mutation) represented the most significant predictor of short EFS (median 3 months) and OS (median 10 months). Concurrent aberration of *TP53* and *CDKN2A* is a new, simple, and relevant index of chemoresistance in MCL. Patients with concurrent aberration of *TP53* and *CDKN2A* should be offered innovative anti-lymphoma therapy and upfront consolidation with allogeneic stem cell transplantation.

**Keywords:** mantle cell lymphoma; *TP53*; *CDKN2A*; prognostic markers; chemoresistance



## 1. Introduction

Mantle cell lymphoma (MCL) represents a subtype of non-Hodgkin lymphoma with poor prognosis, especially for patients who are resistant to front-line treatment [1–5]. Besides the translocation t(11;14), several recurrent cytogenetic aberrations have been reported [6–13]. Approximately 5–10% of patients with newly diagnosed MCL fail to achieve response to upfront immunochemotherapy regimen. Despite many prognostic markers including MCL international prognostic index (MIPI), proliferation marker Ki-67, deletion or mutation of *TP53*, deletion of cyclin dependent kinase 2A (*CDKN2A*), or blastoid morphology, a reliable marker of chemoresistance is not available at diagnosis [14–18].

Based on the so far published data we selected seven candidate genes (*TP53*, *CDKN2A*, ataxia teleangiectasia mutated (*ATM*), B-cell lymphoma 2 (*BCL2*), oncogene *MYC*, retinoblastoma protein 1 (*RBI*) and cyclin-dependent kinase 4 (*CDK4*)), and analyzed their aberrations using fluorescent in-situ hybridization FISH in 126 consecutive patients with newly diagnosed MCL with bone marrow (BM) involvement  $\geq 5\%$ . In 113 patients with available diagnostic DNA samples, *TP53* and *CDKN2A* mutation status was analyzed by next-generation sequencing (NGS) approach. The analyzed cytogenetic aberrations were correlated with overall response rate and survival (event-free survival—EFS, overall survival—OS).

## 2. Patients and Methods

### 2.1. Patients' Characteristics

An unselected cohort of 223 patients with newly diagnosed MCL was subject to analysis. Diagnostic samples obtained from 126 newly diagnosed MCL patients with BM infiltration  $\geq 5\%$  were analyzed by fluorescence in situ hybridization (FISH) between 1 January 2009 and 30 June 2018 at the First Department of Internal Medicine-Hematology, General University Hospital and First Faculty of Medicine, Charles University, Prague. The study was approved by University General Hospital Ethics Committee 63/16 from 22 June 2016. From these 126 patients, 113 were subject to mutational analysis of *TP53* and *CDKN2A* genes by next generation sequencing (NGS). In addition, we analyzed baseline characteristics and survival of other 97 patients with newly diagnosed MCL (from the same time period), who were not subject to FISH or NGS analysis due to low BM infiltration ( $< 5\%$ ) or lack of material.

### 2.2. Fluorescence in Situ Hybridization (FISH)

Cytogenetic and FISH analyses were implemented in the Center of Oncocytogenomics, General University Hospital, Charles University in Prague, Czech Republic, accredited according to ISO 15189. Detailed protocols and a list of FISH probes used are available in the Supplementary Materials.

### 2.3. *TP53* and *CDKN2A* Mutation Assessment by Next-Generation Sequencing (NGS)

NGS was implemented at the University Hospital Olomouc with the regular control of laboratory performance using the internal standards of known mutation load (2% and 5% variant allele frequency-VAF, respectively) in each run and periodic inspection by an external agency. The full coding sequence of the *TP53* (exons 2–11 including 2 bp intronic overlap, plus 5' and 3'UTR; NM\_000546) and *CDKN2A* (exons 1–3 including 2 bp intronic overlap, plus 5' and 3'UTR; NM\_000077) were analyzed by targeted ultra-deep NGS as reported previously [19,20]. Amplicon-based libraries were sequenced as a paired-end on MiSeq (2 × 151 bp, Illumina, San Diego, CA, USA) with minimum target read depth of 5000x. The detection limit was set up to 1%, and mutations within the range 1–3% were confirmed by replication [20]. All detected variants were manually inspected using the Integrative Genomics Viewer (IGV) and annotated using variant population databases (1000 Genomes, gnomAD, ExAC), clinical mutation databases (ClinVar, COSMIC, IARC *TP53* Database) and/or functional prediction tools (SIFT, PolyPhen-2). Only pathogenic or likely pathogenic variants were reported. Polymorphisms

were filtered out using variant population databases (1000 Genomes, gnomAD, ExAC). The data can be downloaded as Supplementary Materials from the journal site.

#### 2.4. Statistical Analysis

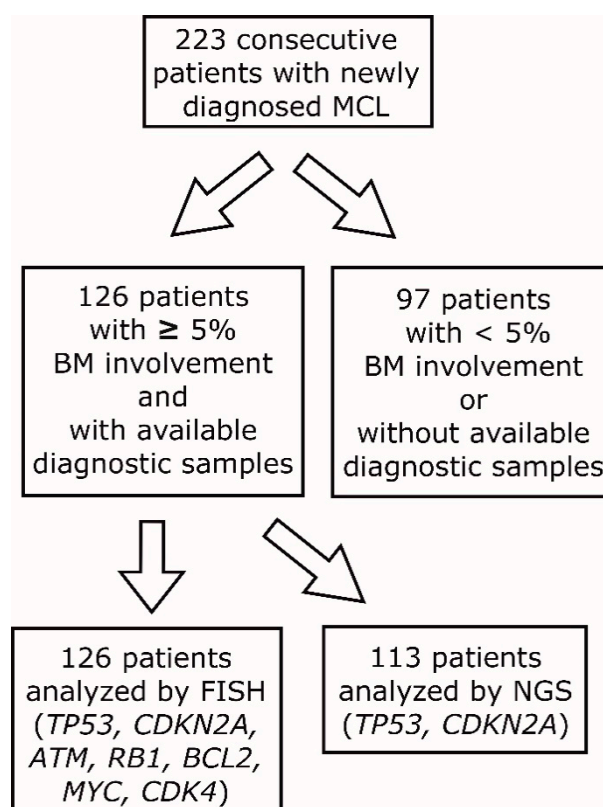
Categorical data were compared by chi-square tests, numerical data with Mann–Whitney U tests. Kaplan–Meier survival analysis was used to estimate EFS and OS, and the statistical significance between survival curves was assessed by a log-rank test. Cox regression multivariate analysis was used to calculate the effect of the variables upon EFS and OS. Data were analyzed using GraphPad Prism version 7.00 for Windows, GraphPad Software, La Jolla, CA, USA, and open-source RStudio, Boston, MA, USA.

We used the Random Forests for Survival, Regression, and Classification R-package to perform all random forests analysis. Variable importance (VIMP) was estimated based on the effect of random permutations on the prediction error. To eliminate the variance of VIMP, all calculations were repeated 100 times and mean values for each VIMP obtained [21].

### 3. Results

#### 3.1. Baseline Characteristics of the Analyzed Patients

We analyzed prognostic significance of molecular-cytogenetic aberrations of seven genes (*TP53*, *CDKN2A*, *ATM*, *BCL2*, *MYC*, *RB1*, and *CDK4*) using FISH in 126 consecutive patients with newly diagnosed MCL with BM involvement  $\geq 5\%$  between 1 January 2009 and 30 June 2018 at the Center of Oncocytogenomics, General University Hospital in Prague (Figure 1).



**Figure 1.** Flow chart of patient analysis. BM = bone marrow, FISH = fluorescent in-situ hybridization, NGS = next-generation sequencing.

In addition to FISH, mutational analysis of *TP53* and *CDKN2A* genes by NGS was implemented in 113 (87%) patients with available DNA. Baseline characteristics of the analyzed patients are displayed in Table 1.

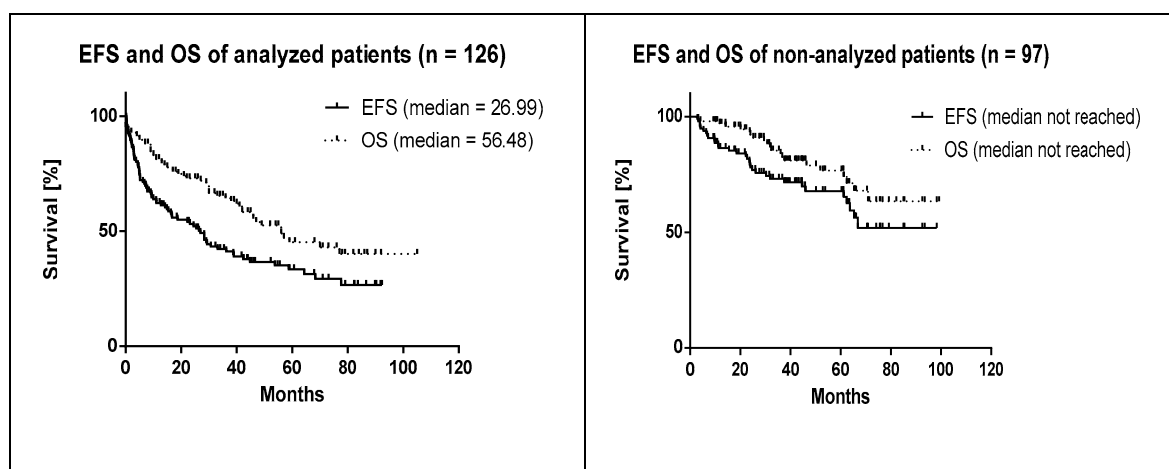
**Table 1.** Baseline characteristics and response to therapy of the analyzed and non-analyzed patients.

Cohort	126 Patients with Bone Marrow Involvement $\geq 5\%$		97 Patients with No Available Diagnostic Samples	
	N	%	N	%
Numbers (N) or percentages (%)				
All patients	126	57	97	43
M	88	70	73	75
F	38	30	24	25
Age (median; years)	68		66	
Age (range; years)	29–82		40–87	
<65 years	47	37	44	45
$\geq 65$ years	79	63	53	55
Stage I–II	0	0	7	7
Stage III	0	0	13	13
Stage IV	126	100	77	79
Ki-67 $\geq 30\%$ *	36	47	38	48
MIPI 1	19	15	25	26
MIPI 2	29	23	34	35
MIPI 3	78	62	38	39
B-symptoms	52	41	32	33
BM infiltration	126	100	73	75
BM infiltration $\geq 5\%$	126	100	25	26
Nodal involvement	108	86	88	91
Splenomegaly	89	71	46	47
Extra-hematological involvement	50	40	47	48
Bulky disease ( $\geq 5$ cm)	45	36	27	28
CNS involvement **	17	13	7	7
Intensified therapy	37	29	50	52
R-CHOP-like therapy	71	56	38	39
Palliative therapy	8	6	7	7
Watch and wait	7	6	1	1
Died before initiation of therapy	3	2	0	0
Died during induction	9	7	0	0
ORR (CR/PR)	90	71	92	95
CR	61	48	73	75
PR	29	23	19	20
SD	4	3	3	3
PD	15	12	0	0
Event	78	62	32	33
Relapse	48	38	18	19
Death **	55	44	21	22

M = male; F = female; MIPI = MCL international prognostic index; BM = bone marrow; CNS = central nervous system; ORR = overall response rate; CR = complete remission; PR = partial remission; SD = stable disease; PD = progressive disease; R-CHOP = R(ituximab + C(yclophosphamide) + H(ydroxydaunomycin) + O(ncovin) + P(rednisonone); response was assessed by international workshop criteria published by Cheson et al. in 1999 [22]. \* of the analyzed samples, \*\* anytime from diagnosis until database lock; differences  $>20\%$  between cohorts are highlighted in gray.



Because FISH was analyzed only on BM (with infiltration  $\geq 5\%$ ) or peripheral blood, but not on formalin-fixed paraffin-embedded tissue sections, the analyzed (FISH) patients ( $n = 126$ ) represented approx. 57% of all MCL patients ( $n = 223$ ) diagnosed at the First Dept. of Internal Medicine-Hematology, General University Hospital and First Faculty of Medicine, Charles University, Prague (Figure 1). From the 97 patients (43%) with no available FISH data, 72 patients (74%) had undetectable or low ( $<5\%$ ) infiltration of the BM. Besides that, FISH data were unavailable for 25 patients (26%) (no diagnostic samples available) (Table 1). The applied methodology thus inevitably led to over-representation of high-risk patients according to MIPI in the cohort analyzed by FISH and NGS because patients with  $<5\%$  infiltration of the BM were not analyzed. Indeed, while 5-year EFS and OS in the analyzed (FISH, NGS) cohort was 33.4 and 44.8%, respectively, 5-year EFS and OS of the non-analyzed cohort was 67.8% and 76.7%. (Figure 2). Median follow-up of the living patients reached 42 and 44 months in the analyzed and non-analyzed cohort, respectively.



**Figure 2.** Survival in patients with bone marrow involvement  $\geq 5\%$  compared to patients without available diagnostic samples. EFS = event-free survival, OS = overall survival.

### 3.2. Correlation of Baseline Clinical and Histopathological Parameters on Survival in the Cohort of 126 Patients with $\geq 5\%$ BM Involvement

From the analyzed clinical and histopathological factors, the following parameters correlated with EFS and OS in the cohort of 126 patients with  $\geq 5\%$  BM involvement: MIPI (high risk vs. intermediate risk vs. low risk  $p < 0.0001$ ), Ki-67 ( $\geq 30\%$  vs.  $<30\%$ ), B symptoms (present vs. absent), complex karyotype (yes vs. no), type of therapy (intensified vs. R-CHOP-based vs. palliative—for EFS  $p = 0.0007$  and for OS  $p = 0.0003$ ). Splenomegaly and bulky disease ( $\geq 5$  cm) correlated with EFS, but not OS (Table S1). Despite the fact that all patients had infiltrated BM, the extent of BM infiltration positively correlated with shorter survival (for EFS, HR = 1.009 for each 1% of increase of BM infiltration, 95% CI = 1.001–1.02,  $p = 0.0312$ ; for OS, HR = 1.016 for each 1% of increase of BM infiltration, 95% CI = 1.006–1.027,  $p = 0.002$ ).

Deletion of CDKN2A can be detected almost exclusively in the context of other recurrently found cytogenetic aberrations

Distribution of the analyzed gene aberrations is displayed in Figure 3 and Table S2.



(Table S1). As few as 25 patients (19.8%) had no detectable gene aberration (apart from the translocation t(11;14)). A single (isolated) gene aberration was observed in 39 patients (31%), and  $\geq 2$  aberrations were detected in 62 patients (49.2%). Two, three, four, five, and six aberrations were detected in 22 (17.5%), 16 (12.7%), 16 (12.7%), 5 (4%), and 3 (2.4%) patients, respectively. From 39 patients with an isolated gene aberration, 15 patients had *TP53* gene aberration (11.9%), 9 patients had *ATM* deletion (7.1%), and 5 patients had *RB1* deletion (4%). Isolated gene aberrations of *BCL2* ( $n = 3$ , 2.4%), *MYC* ( $n = 3$ , 2.4%), *CDK4* ( $n = 3$ , 2.4%), and *CDKN2A* ( $n = 1$ , 0.8%) were rare. In addition, the only patient with isolated *CDKN2A* deletion had chromosome 9 monosomy (Figure 3, Table S2). A Pearson chi-square test of seven analyzed aberrations revealed that *CDKN2A* gene deletion correlated with aberrations of *TP53*, *BCL2*, *RB1*, and *CDK4*, while aberration of *BCL2* correlated with *CDK4*. No other correlations were found among the seven analyzed genes (Table 2).

**Table 2.** Correlation between the analyzed gene aberrations.

	<i>CDK4</i>	<i>RB1</i>	<i>BCL2</i>	<i>ATM</i>	<i>TP53</i>	<i>CDKN2A</i>	<i>MYC</i>
<i>CDK4</i>	1	0.138	<b>&lt;0.001</b>	0.260	0.651	<b>0.016</b>	0.055
<i>RB1</i>		1	0.384	0.847	0.191	<b>&lt;0.001</b>	0.074
<i>BCL2</i>			1	0.091	0.965	<b>0.012</b>	0.055
<i>ATM</i>				1	0.164	0.05	0.706
<i>TP53</i>					1	<b>0.006</b>	0.071
<i>CDKN2A</i>						1	0.065
<i>MYC</i>							1

The table shows *p*-values of Pearson's chi-squared test. Statistically significant results are underlined and printed in bold fonts.

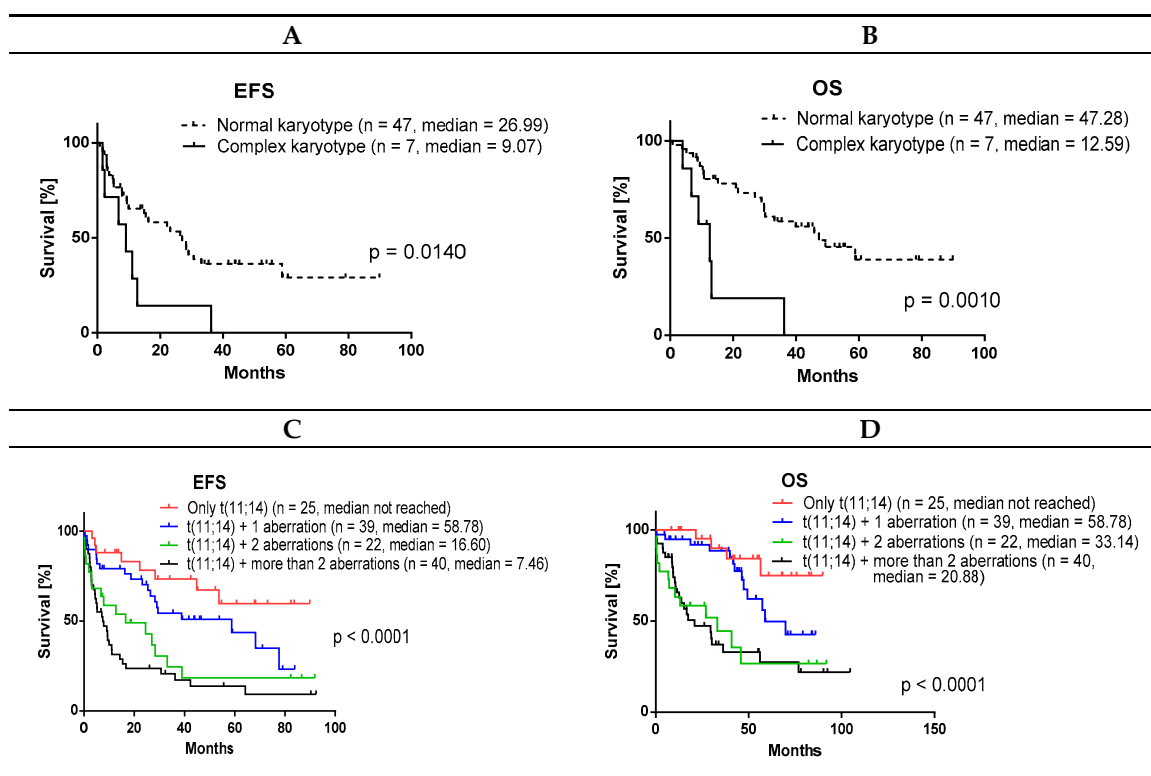
Distribution of the analyzed aberrations of 101 patients with at least one detected aberration including *TP53* mutation (except for the translocation t(11;14)). Each row represents one patient, gray squares represent aberrated genes, numbers represent type of aberration: 2 = monoallelic deletion (in case of *TP53* and/or mutation), 3 = biallelic deletion, 4 = monosomy, 5 = nullisomy, 6 = amplification, 7 = gain, 8 = trisomy, 9 = tetrasomy, 10 = *MYC* rearrangement; more numbers represent different subclones (e.g., 2,3 represent patients, in which both monoallelic and biallelic deletions of *CDKN2A* were detected).

*TP53* mutation and *TP53* deletion are both associated with adverse prognosis in MCL, while pathogenic *CDKN2A* mutations were not detected.

*TP53* gene aberration, either mutation, or deletion, was detected in 52 out of 126 patients (41.3%). Thirty-one of the analyzed patients (59.6%) with *TP53* aberration had both deletion and mutation of the *TP53* gene, while mutation without deletion, and deletion without mutation was detected in 12 (23.1%) and 5 (9.6%) patients. Another four patients (7.7%) had *TP53* deletion, but, due to lack of material, mutational analysis was not performed. Because *TP53* deletion and mutation significantly correlated with each other with respect to survival parameters (EFS and OS), we used *TP53* aberration (defined as *TP53* mutation and/or deletion) for all analyses (Tables S3 and S4, Figure S1). In addition to *TP53*, mutational analysis of *CDKN2A* gene was implemented by NGS, but no pathological mutations were identified.

A total number of gene aberrations is a strong predictor of outcome.

We confirmed that a complex karyotype significantly correlated with shorter survival (Figure 4A,B). Likewise, a total number of FISH aberrations (including *TP53* mutation) also correlated with EFS and OS. Interestingly, the biggest difference was observed between any two of the seven gene aberrations compared to any single (isolated) aberration, while three or more aberrations were not significantly worse predictors than two aberrations (Figure 4C,D). The total number of aberrations positively correlated with the male sex (Mann–Whitney U test  $p = 0.045$ ).



**Figure 4.** Complex karyotype and total number of gene aberrations correlates with shorter survival. (A,C) EFS; (B,D) OS. EFS = event-free survival; OS = overall survival.

Concurrent *TP53* aberration and *CDKN2A* deletion predicts chemoresistance.

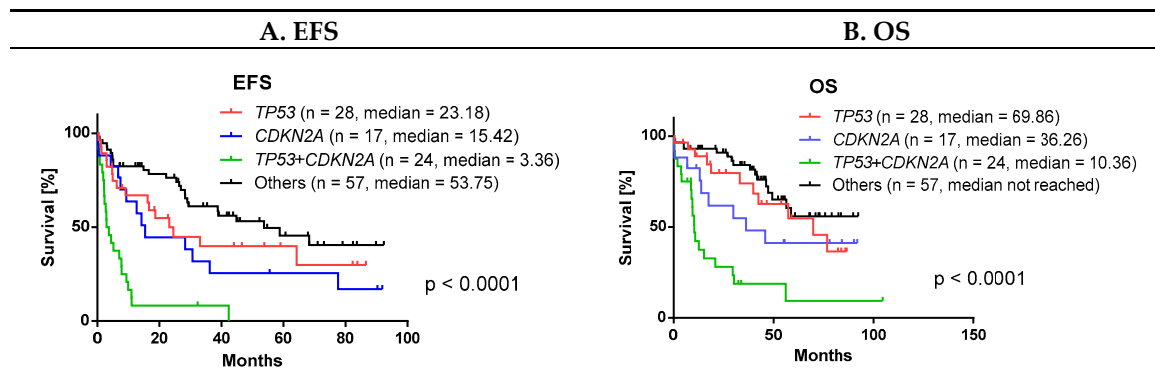
We asked which of the two analyzed gene aberrations belonged to the most relevant pretreatment markers associated with the worst clinical outcome. First, Cox regression revealed that both *TP53* and *CDKN2A* aberrations independently correlated with shorter EFS and OS, while *BCL2* aberration correlated with shorter OS (Table 3). Second, a random forest analysis of the seven gene aberrations analyzed revealed that *CDKN2A* deletion is the most important predictor of short EFS and OS (Figure S2A,B). Third, random forest analysis of all 21 possible aberration pairs unveiled that concurrent *TP53* gene aberration ( $TP53_{mut/del}$ ) and *CDKN2A* deletion ( $CDKN2A_{del}$ ) was the strongest predictor of short EFS, and together with concurrent *CDKN2A* and *BCL2* aberrations also predictor of short OS. (Figure S2C,D).

**Table 3.** Effect of the analyzed gene aberrations with survival parameters: multivariate analysis. (A) Event-Free Survival; (B) Overall Survival.

A. Event-Free Survival				B. Overall Survival			
Gene	HR	95% CI	<i>p</i>	Gene	HR	95% CI	<i>p</i>
<i>CDK4</i>	1.6	0.8–3.1	0.218	<i>CDK4</i>	1.7	0.8–3.7	0.205
<i>RB1</i>	0.9	0.5–1.6	0.803	<i>RB1</i>	1.2	0.6–2.2	0.645
<i>BCL2</i>	1.5	0.8–2.5	0.287	<i>BCL2</i>	2.6	1.4–4.8	0.004
<i>ATM</i>	1.1	0.7–1.9	0.667	<i>ATM</i>	1.0	0.6–2.0	0.921
<i>TP53</i>	2.3	1.4–3.6	0.001	<i>TP53</i>	2.2	1.2–3.8	0.008
<i>CDKN2A</i>	2.6	1.5–4.7	0.001	<i>CDKN2A</i>	2.5	1.2–4.9	0.011
<i>MYC</i>	1.6	1.0–2.6	0.06	<i>MYC</i>	1.2	0.7–2.2	0.507

Tables show Cox's proportional hazard model; HR = hazard ratio, CI = confidence interval  $p = p$ -value; statistically significant results are highlighted in gray.

Survival parameters of patients with concurrent  $TP53_{mut/del}$  and  $CDKN2A_{del}$  are shown in Figure 5.



**Figure 5.** Patients with concurrent *TP53* and *CDKN2A* aberrations have significantly shorter survival than isolated aberrations. Patients in groups *TP53* and *CDKN2A* are patients with *TP53* aberrations or *CDKN2A* deletions that are not contained in the *TP53+CDKN2A* subcohort. (A) EFS, (B) OS. EFS = event-free survival, OS = overall survival.

Subanalysis of patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> is given in Table S5. The biggest differences between the patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> and the remaining analyzed patients included frequency of CNS involvement (33% vs. 9%), B-symptoms (71% vs. 34%), MIPI (MIPI 3 in 79% vs. 58%) and Ki-67  $\geq 30\%$  (85% vs. 40% in remaining patients). Only 38% and 17% of patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> achieved response and complete response, respectively (compared to 79% and 56% of the patients without concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub>). At the time of database lock, 96% and 21% of patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> had an event and were alive, respectively (compared to 54% and 65% in the remaining analyzed patients).

#### 4. Discussion

As few as 25 patients (19.8%) had no detectable gene aberration (apart from the translocation t(11;14)). This is almost identical finding to that of Dalfau-Larue et al., who reported in their landmark analysis of 135 younger MCL patients from the European MCL Younger trial that only 24 patients (21%) displayed no copy number alteration (CNA) in any analyzed loci [18]. Dalfau-Larue et al. analyzed CNA of similar genes to our own selection including *TP53*, *CDKN2A*, *ATM*, *RB1*, *CDK4*, and *MYC* (but not *BCL2*). Remarkably, the distribution of CNAs of the analyzed genes in the report of Dalfau-Larue was similar to our own data including *RB1* deletions (26% compared to 33%), *ATM* deletions (25% compared to 28%), *CDKN2A* deletions (25% compared to 33%), *TP53* deletions (22% compared to 32% of *TP53* gene deletions detected by FISH), *MYC* aberrations (18% compared to 26%), and *CDK4* gains (8% compared to 12%) (Figure 3, Table S2). The overall lower incidence of all analyzed CNAs in the study of Dalfau-Larue et al. compared to our own data can be explained by differences in the analyzed cohorts of patients. The significantly higher number of high-risk MIPI patients in our cohort (68%) compared to that of Dalfau-Larue (25%) can be explained not only by different median age, but also by the fact that clinical trials usually do not enroll non-fit patients (slow go, no go) with poor performance status or serious comorbidities. In a recently published study by Wang et al., *MYC* rearrangements, but not *MYC* CNAs correlated with shorter OS independent of MIPI and Ki-67 [11]. In our study, 4 and 28 patients had *MYC* rearrangements and CNAs, respectively (one patient harbored both types of aberrations). While *MYC* aberrations correlated with shorter EFS in univariate analysis, its prognostic significance was lost in multivariate analysis (Table 3). In addition, *MYC* rearrangements were found exclusively in the context of other analyzed aberrations, suggesting it represents a later event similar to *CDKN2A* loss.

In the manuscript by Clot et al., 39 nodal MCL patients (62% belonging to high-risk MIPI group—exactly the same number as in our own analyzed cohort) were analyzed for *TP53* aberration and

*CDKN2A* deletion (besides other analyzed genes) [17]. Curiously, 36% and 33% of the analyzed patients had *TP53* aberration and *CDKN2A* deletion, respectively (compared to 41% and 33% in our study). In addition, 13% patients had both *TP53* aberration and *CDKN2A* deletion (compared to 19% in our own study). Thus far, published results thus clearly confirm that both *TP53* aberrations and *CDKN2A* deletions are frequent in MCL, and that both are associated with adverse outcome.

#### 4.1. *TP53* Deletions and Mutations

Eskelund et al. recently reported that *TP53* mutations (detected in 20 patients, 11%) had significantly worse outcomes compared to *TP53* deletions (detected in 29 patients, 16%) [16]. In his report, only nine patients (47% patients from the *TP53* mutated cases, and 31% patients from the *TP53* deleted cases) had both *TP53* aberrations. In the study by Obr et al., 50% of patients with *TP53* deletions (with available sample for NGS) also had *TP53* mutations [19]. In our study, from 126 patients analyzed by FISH (for *TP53* deletion), 113 patients were also analyzed by NGS for *TP53* mutation. *TP53* aberration was observed in 52 patients (41.3% from 126), 40 patients had *TP53* deletion (31.7% from 126) and 43 patients had *TP53* mutation (38.1% from 113 analyzed by NGS). Thirty-one patients had both *TP53* aberrations (77.5% and 72.1% of the patients analyzed by FISH and NGS, respectively). Mutation without deletion of *TP53* was detected in 12 out of 43 patients (27.9%) analyzed by NGS. Deletion without mutation of *TP53* was detected in 9 out of 40 patients (22.5%) analyzed by FISH, but only in 5 out of 43 (11.6%) patients analyzed also by NGS (Tables S2 and S3). Interestingly, similar results have been reported for chronic lymphocytic leukemia (CLL) patients, where >70% of patients with *TP53* deletion also carried a *TP53* mutation [23]. Explanation for the observed differences between our data and the study of Eskelund et al. [16] might include different study cohorts (transplant-eligible patients included in clinical trials versus unselected, predominantly elderly patients with BM infiltration), and different methods for detection of *TP53* deletions (droplet digital polymerase chain reaction versus conventional FISH). Notably, in the study of Eskelund et al., patients with *TP53* mutation without deletion had a higher incidence of concurrent *CDKN2A* deletion (58%) compared to patients with *TP53* deletion without mutation (41%), which might at least partially explain the different prognostic impact of *TP53* mutation compared to deletion reported by the Nordic group.

#### 4.2. Survival Analysis

Besides *TP53*, *CDKN2A* belongs to established prognostic markers in MCL [18]. *CDKN2A* deletion was observed in 41 out of 126 patients (32.5%) and represented the second most frequent aberration in the analyzed cohort. The data suggest that *CDKN2A* deletions represent late events in MCL because virtually all *CDKN2A* gene deletions were detected in the context of other genetic aberrations (Figure 3). Pearson's analysis confirmed that *CDKN2A* correlated with incidence of all analyzed gene aberrations except for *ATM* deletion. In addition, *CDKN2A* deletions also correlated with male sex, MIPI, proliferation index by Ki-67, B-symptoms, CNS disease, and complex karyotype. The higher incidence of analyzed aberrations (especially *CDKN2A* and *RB1* deletions) observed in men might at least partially explain their worse outcome compared to women.

#### 4.3. Concurrent Aberration of *TP53* and Deletion of *CDKN2A* Is Associated with Chemoresistance

Concurrent *CDKN2A* deletion and *TP53* aberration were associated with chemoresistance to currently used upfront immunochemotherapy. We thus confirmed the findings of Delfau-Larue et al. on a real-life cohort of predominantly elderly MCL patients (median age 68 compared to 56 years) with significantly higher proportion of high-risk MIPI (62% compared to 25%). Subanalysis of these patients is given in Table S5. Recently, Streich et al. reported that MCL with blastoid and pleomorphic morphology frequently harbor both *TP53* and *CDKN2A/B* aberrations, and that these cases are characterized by frequent chromothripsis [24]. Curiously, in sharp contrast to MCL, aberrations of *TP53* and *CDKN2A* were mutually exclusive in Burkitt lymphoma and were very rarely observed in diffuse large B cell lymphoma [25–27]. In indolent lymphoproliferative malignancies, namely CLL and



follicular lymphoma, deletions of *CDKN2A* are rare, and *CDKN2A* inactivation frequently correlates with the transformation to an aggressive lymphoma with adverse prognosis [28–31]. From this perspective, MCL patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> might be regarded as patients with “transformed” MCL and consequently with similarly adverse prognosis. Only three younger patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> were successfully treated with salvage therapy and allogeneic stem cell transplantation (allo-SCT).

## 5. Conclusions

The molecular/cytogenetic index *TP53*<sub>del/mut</sub> and *CDKN2A*<sub>del</sub> represents a novel, simple, and reliable pretreatment prognostic factor that identifies patients who do not profit from currently used therapies based on immunochemotherapy. Importantly, the combination of both aberrations represents a significantly more relevant prognostic marker of poor outcome compared to the isolated aberration of either gene. Patients with concurrent *TP53*<sub>mut/del</sub> and *CDKN2A*<sub>del</sub> might profit from innovative treatments, e.g., Bruton’s tyrosine kinase inhibitors, ideally in combination with other anti-lymphoma drugs (e.g., BH3-mimetics), and from upfront consolidation with allogeneic stem cell transplantation.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6694/12/8/2120/s1>, Supplementary Methods: Fluorescence in situ hybridization (FISH); Table S1: Univariate analysis: correlation of analyzed gene aberrations and selected clinical and laboratory parameters with survival; Table S2: Distribution of the analyzed gene aberrations; Table S3: TP53 mutation types and positions; Table S4: Univariate analysis of TP53 mutation and TP53 deletion; Table S5: Baseline characteristics and response to therapy of the patients with concurrent aberration of TP53 and CDKN2A (compared to remaining patients); Figure S1: Survival parameters in the TP53 mutation and TP53 deletion cohorts; Figure S2: Random Forest analysis of analyzed aberrations.

**Author Contributions:** Conceptualization, P.K., M.T., Z.Z., E.K.; methodology, M.S., P.B.; software, M.S., P.B.; validation, P.K., M.T.; formal analysis, D.M., P.B., M.S.; investigation, D.M., A.B., E.P., K.F., L.P., A.P.; resources, P.K., E.K., M.T., Z.Z.; data curation, D.M., A.B., E.K., Z.Z.; writing—original draft preparation, P.K., D.M.; writing—review and editing, P.K., M.T., A.O., E.K., Z.Z.; visualization, D.M.; supervision, P.K.; project administration, P.K.; funding acquisition, P.K., M.T., A.O., E.K., Z.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** P.K. was supported by Ministry of Health of the Czech Republic grant AZV 17-28980A, all rights reserved, Grant Agency of the Czech Republic GA20-25308S, Charles University Center of Excellence UNCE/MED/016 and Ministry of Education, Youth and Sports grants PROGRES Q26/LF1 and PROGRES Q28/LF1; Z.Z., E.P. and L.P. was supported by RVO-VFN64165; A.O., E.K. and A.P. were supported by MH CZ—DRO (FNOL, 00098892) and IGA\_LF\_2020\_002, IGA\_LF\_2020\_016.

**Conflicts of Interest:** The authors declare no competing financial interest.

## References

1. Klener, P. Advances in Molecular Biology and Targeted Therapy of Mantle Cell Lymphoma. *Int. J. Mol. Sci.* **2019**, *20*, 4417. [[CrossRef](#)] [[PubMed](#)]
2. Cheah, C.Y.; Seymour, J.F.; Wang, M.L. Mantle Cell Lymphoma. *J. Clin. Oncol.* **2016**, *34*, 1256–1269. [[CrossRef](#)] [[PubMed](#)]
3. Jain, P.; Wang, M. Mantle cell lymphoma: 2019 update on the diagnosis, pathogenesis, prognostication, and management. *Am. J. Hematol.* **2019**, *94*, 710–725. [[CrossRef](#)] [[PubMed](#)]
4. Maddocks, K. Update on mantle cell lymphoma. *Blood* **2018**, *132*, 1647–1656. [[CrossRef](#)] [[PubMed](#)]
5. Rule, S. The modern approach to mantle cell lymphoma. *Hematol. Oncol.* **2019**, *37*, 66–69. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, J.; Jima, D.; Moffitt, A.B.; Liu, Q.; Czader, M.; Hsi, E.D.; Fedoriw, Y.; Dunphy, C.H.; Richards, K.L.; Gill, J.I.; et al. The genomic landscape of mantle cell lymphoma is related to the epigenetically determined chromatin state of normal B cells. *Blood J. Am. Soc. Hematol.* **2014**, *123*, 2988–2996. [[CrossRef](#)]
7. Bea, S.; Valdes-Mas, R.; Navarro, A.; Salaverria, I.; Martín-García, D.; Jares, P.; Giné, E.; Pinyol, M.; Royo, C.; Nadeu, F.; et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18250–18255. [[CrossRef](#)]

8. Yang, P.; Zhang, W.; Wang, J.; Liu, Y.; An, R.; Jing, H. Genomic landscape and prognostic analysis of mantle cell lymphoma. *Cancer Gene Ther.* **2018**, *25*, 129–140. [[CrossRef](#)]
9. Ahmed, M.; Zhang, L.; Nomie, K.; Lam, L.; Wang, M. Gene mutations and actionable genetic lesions in mantle cell lymphoma. *Oncotarget* **2016**, *7*, 58638–58648. [[CrossRef](#)]
10. Kridel, R.; Meissner, B.; Rogic, S.; Boyle, M.; Telenius, A.; Woolcock, B.; Gunawardana, J.; Jenkins, C.; Cochrane, C.; Ben-Neriah, S.; et al. Whole transcriptome sequencing reveals recurrent NOTCH1 mutations in mantle cell lymphoma. *Blood* **2012**, *119*, 1963–1971. [[CrossRef](#)]
11. Wang, L.; Tang, G.; Medeiros, L.J.; Xu, J.; Huang, W.; Yin, C.C.; Wang, M.; Jain, P.; Lin, P.; Li, S. MYC rearrangement but not extra MYC copies is an independent prognostic factor in patients with mantle cell lymphoma. *Haematologica* **2020**, *105*. [[CrossRef](#)] [[PubMed](#)]
12. Ferrero, S.; Rossi, D.; Rinaldi, A.; Brusca, G.; Spina, V.; Eskelund, C.W.; Evangelista, A.; Moia, R.; Kwee, I.; Dahl, C. KMT2D mutations and TP53 disruptions are poor prognostic biomarkers in mantle cell lymphoma receiving high-dose therapy: A FIL study. *Haematologica* **2020**, *105*, 1604–1612. [[CrossRef](#)] [[PubMed](#)]
13. Hill, H.A.; Qi, X.; Jain, P.; Nomie, K.; Wang, Y.; Zhou, S.; Wang, M.L. Genetic mutations and features of mantle cell lymphoma: A systematic review and meta-analysis. *Blood Adv.* **2020**, *4*, 2927–2938. [[CrossRef](#)] [[PubMed](#)]
14. Hoster, E.; Dreyling, M.; Klapper, W.; Gisselbrecht, C.; Van Hoof, A.; Kluin-Nelemans, H.C.; Pfreundschuh, M.; Reiser, M.; Metzner, B.; Einsele, H.; et al. A new prognostic index (MIPI) for patients with advanced-stage mantle cell lymphoma. *Blood* **2008**, *111*, 558–565. [[CrossRef](#)]
15. Hoster, E.; Rosenwald, A.; Berger, F.; Bernd, H.W.; Hartmann, S.; Lodenkemper, C.; Barth, T.F.; Brousse, N.; Pileri, S.; Rymkiewicz, G.; et al. Prognostic Value of Ki-67 Index, Cytology, and Growth Pattern in Mantle-Cell Lymphoma: Results from Randomized Trials of the European Mantle Cell Lymphoma Network. *J. Clin. Oncol.* **2016**, *34*, 1386–1394. [[CrossRef](#)]
16. Eskelund, C.W.; Dahl, C.; Hansen, J.W.; Westman, M.; Kolstad, A.; Pedersen, L.B.; Montano-Almendras, C.P.; Husby, S.; Freiburghaus, C.; Ek, S.; et al. TP53 mutations identify younger mantle cell lymphoma patients who do not benefit from intensive chemoimmunotherapy. *Blood J. Am. Soc. Hematol.* **2017**, *130*, 1903–1910. [[CrossRef](#)]
17. Clot, G.; Jares, P.; Gine, E.; Navarro, A.; Royo, C.; Pinyol, M.; Martín-García, D.; Demajo, S.; Espinet, B.; Salar, A.; et al. A gene signature that distinguishes conventional and leukemic nonnodal mantle cell lymphoma helps predict outcome. *Blood* **2018**, *132*, 413–422. [[CrossRef](#)]
18. Delfau-Larue, M.H.; Klapper, W.; Berger, F.; Jardin, F.; Briere, J.; Salles, G.; Casasnovas, O.; Feugier, P.; Haioun, C.; Ribrag, V.; et al. High-Dose cytarabine does not overcome the adverse prognostic value of CDKN2A and TP53 deletions in mantle cell lymphoma. *Blood J. Am. Soc. Hematol.* **2015**, *126*, 604–611. [[CrossRef](#)]
19. Obr, A.; Prochazka, V.; Jirkuvova, A.; Urbánková, H.; Kriegova, E.; Schneiderová, P.; Vatošíková, M.; Papajík, T. TP53 Mutation and Complex Karyotype Portends a Dismal Prognosis in Patients with Mantle Cell Lymphoma. *Clin. Lymphoma Myeloma Leuk.* **2018**, *18*, 762–768. [[CrossRef](#)]
20. Petrackova, A.; Vasinek, M.; Sedlarikova, L.; Dyskova, T.; Schneiderova, P.; Novosad, T.; Papajik, T.; Kriegova, E. Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. *Front. Oncol.* **2019**, *9*, 851. [[CrossRef](#)]
21. Ishwaran, H.; Lu, M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Stat. Med.* **2019**, *38*, 558–582. [[CrossRef](#)] [[PubMed](#)]
22. Cheson, B.D.; Horning, S.J.; Coiffier, B.; Shipp, M.A.; Fisher, R.I.; Connors, J.M.; Lister, T.A.; Vose, J.; Grillo-López, A.; Hagenbeek, A.; et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. *J. Clin. Oncol.* **1999**, *17*, 1244. [[CrossRef](#)] [[PubMed](#)]
23. Zenz, T.; Krober, A.; Scherer, K.; Häbe, S.; Bühler, A.; Benner, A.; Denzel, T.; Winkler, D.; Edelmann, J.; Schwänen, C.; et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: Results from a detailed genetic characterization with long-term follow-up. *Blood J. Am. Soc. Hematol.* **2008**, *112*, 3322–3329. [[CrossRef](#)] [[PubMed](#)]
24. Streich, L.; Sukhanova, M.; Lu, X.; Chen, Y.H.; Venkataraman, G.; Mathews, S.; Zhang, S.; Kelemen, K.; Segal, J.; Gao, J.; et al. Aggressive morphologic variants of mantle cell lymphoma characterized with high genomic instability showing frequent chromothripsis, CDKN2A/B loss, and TP53 mutations: A multi-institutional study. *Genes Chromosomes Cancer* **2020**, *59*, 484–494. [[CrossRef](#)] [[PubMed](#)]



25. Sanchez-Beato, M.; Saez, A.I.; Navas, I.C.; Algara, P.; Mateo, M.S.; Villuendas, R.; Camacho, F.; Sánchez-Aguilera, A.; Sánchez, E.; Piris, M.A. Overall survival in aggressive B-Cell lymphomas is dependent on the accumulation of alterations in p53, p16, and p27. *Am. J. Pathol.* **2001**, *159*, 205–213. [[CrossRef](#)]
26. Lopez, C.; Kleinheinz, K.; Aukema, S.M.; Rohde, M.; Bernhart, S.H.; Hübschmann, D.; Wagener, R.; Toprak, U.H.; Raimondi, F.; Kreuz, M.; et al. Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. *Nat. Commun.* **2019**, *10*, 1459. [[CrossRef](#)]
27. Bolen, C.R.; Klanova, M.; Trneny, M.; Sehn, L.H.; He, J.; Tong, J.; Paulson, J.N.; Kim, E.; Vitolo, U.; Di Rocco, A.; et al. Prognostic impact of somatic mutations in diffuse large B-Cell lymphoma and relationship to cell-of-origin: Data from the phase III GOYA study. *Haematologica* **2019**, *105*, 227892. [[CrossRef](#)]
28. Rossi, D.; Spina, V.; Gaidano, G. Biology and treatment of Richter syndrome. *Blood* **2018**, *131*, 2761–2772. [[CrossRef](#)]
29. Chigrinova, E.; Rinaldi, A.; Kwee, I.; Rossi, D.; Rancoita, P.M.; Strefford, J.C.; Oscier, D.; Stamatopoulos, K.; Papadaki, T.; Berger, F.; et al. Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome. *Blood J. Am. Soc. Hematol.* **2013**, *122*, 2673–2682. [[CrossRef](#)]
30. Kwiecinska, A.; Ichimura, K.; Berglund, M.; Dinets, A.; Sulaiman, L.; Collins, V.P.; Larsson, C.; Porwit, A.; Lagercrantz, S.B. Amplification of 2p as a genomic marker for transformation in lymphoma. *Genes Chromosomes Cancer* **2014**, *53*, 750–768. [[CrossRef](#)]
31. Pasqualucci, L.; Khiabani, H.; Fangazio, M.; Vasishtha, M.; Messina, M.; Holmes, A.B.; Ouillette, P.; Trifonov, V.; Rossi, D.; Tabbò, F.; et al. Genetics of follicular lymphoma transformation. *Cell Rep.* **2014**, *6*, 130–140. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## APPENDIX I

Obr A, Klener P, Furst T, Kriegova E, Zemanova Z, Urbankova H, Jirkuvova A, **Petrackova A**, Malarikova D, Forsterova K, Cudova B, Sedlarikova L, Berkova A, Kasalova N, Papajik T, Trneny M. A high TP53 mutation burden is a strong predictor of primary refractory mantle cell lymphoma. *Br J Haematol.* 2020;191(5):e103-e106. (IF 2020: 6.998; Q1)

## References

- García-Porrúa C, González-Gay MA. Cutaneous vasculitis as a paraneoplastic syndrome in adults. *Arthritis Rheum.* 1998;41:1133–5.
- Oka S, Ono K, Nohgawa M. Multiple myeloma presenting as cutaneous leukocytoclastic vasculitis and eosinophilia disclosing a T helper type 1/T helper type 2 imbalance: a case report. *J Med Case Rep.* 2018;12:320.
- Sánchez NB, Canedo IF, García-Patos PE, Unamuno Pérez P, Benito AV, Pascual AM. Paraneoplastic vasculitis associated with multiple myeloma. *J Eur Acad Dermatol Venereol.* 2004;18:731–5.
- Richardson PG, Xie W, Mitsiades CS, Chanan-Khan AA, Lonial S, Hassoun H, et al. Single-agent bortezomib in previously untreated multiple myeloma: efficacy, characterization of peripheral neuropathy, and molecular correlations with response and neuropathy. *J Clin Oncol.* 2009;27:3518–25.
- Dalmau J, Rosenfeld MR. Paraneoplastic syndromes of the CNS. *Lancet Neurol.* 2008;7:327–40.
- Pittock SJ, Kryzer TJ, Lennon VA. Paraneoplastic antibodies coexist and predict cancer, not neurological syndrome. *Ann Neurol.* 2004;56:715–9.
- Iorio R, Lennon VA. Neural antigen-specific autoimmune disorders. *Immunol Rev.* 2012;248:104–21.
- Wingerchuk DM, Lennon VA, Lucchinetti CF, Pittock SJ, Weinshenker BG. The spectrum of neuromyelitis optica. *Lancet Neurol.* 2007;6:805–15.
- Pittock SJ, Lennon VA. Aquaporin-4 autoantibodies in a paraneoplastic context. *Arch Neurol.* 2008;65:629–32.
- Kater AP, Tonino SH, Egle A, Ramsay AG. How does lenalidomide target the chronic lymphocytic leukemia microenvironment? *Blood.* 2014;124:2184–9.

## A high *TP53* mutation burden is a strong predictor of primary refractory mantle cell lymphoma

The mantle cell lymphoma (MCL) International Prognostic Index is widely used as a strong prognostic stratifier, but it has rarely been used for the selection of therapeutic approaches.<sup>1</sup> New molecular prognostic predictors, which would reflect critical aspects of MCL biology and that would help in therapy decisions, have been under investigation.<sup>2</sup>

Recently, several key studies provided evidence of the poor prognostic impact of tumour protein p53 (*TP53*) gene aberrations or high p53 protein expression in patients with MCL.<sup>3–6</sup> There is a growing body of evidence that *TP53* aberrations could contribute to chemoresistance.<sup>7</sup>

In the present study, we analysed the prognostic impact of the extent of a *TP53* gene aberration (*TP53* mutation burden and deletion of 17p frequency) in a real-world cohort of patients with MCL.

This was a retrospective study of 114 consecutive unselected adult patients with newly diagnosed MCL between April 2006 and October 2016. Tumour tissues (bone marrow, peripheral blood, and lymph nodes) from all the patients who were analysed were obtained before the initiation of treatment. The samples were examined by fluorescence *in situ* hybridisation (FISH) and next-generation sequencing (NGS). The deletion frequency was defined as the percentage of nuclei carrying a deletion of 17p (del17) out of the total number of nuclei carrying the translocation t(11;14). The *TP53*mut burden was defined as the variant allele frequency (VAF) of the mutations that were detected or the highest VAF in the case of multiple mutations present in one patient. Patients with stable or progressive disease during induction or with relapse or progression within 6 months after the completion of induction were considered as primary refractory (PrR).

In all, 27 (23.7%) of the 114 patients who were analysed exhibited PrR. More detailed characteristics are summarised in Table I.

A total of 43 patients (37.7%) had some type of *TP53* disruption. The overall (OS) and progression-free survival (PFS) curves were not statistically different between the *TP53*mut, del17, and *TP53*mut + del17 cohorts (*TP53*mut/del17/*TP53*mut + del17) ( $P = 0.95$  for OS;  $P = 0.86$  for PFS) (Figure SA).

A total of 24 patients had both *TP53*mut and del17, 13 patients had isolated *TP53*mut, and five had del17. The association between *TP53*mut and del17 was significant, at  $P < 0.001$ . Moreover, in the PrR cohort, there was almost unit correlation ( $c = 0.90$ ,  $P < 0.001$ ) between the *TP53*mut burden and del17 frequency (Figure SB).

The *TP53*mut burden (mean 33%, range 3–85%) and del17 frequency (mean 47%, range 6–92%) varied significantly among the patients. As expected, a *TP53* aberration was a strong predictor of PrR, regardless of age and therapeutic approach. Interestingly, PrR patients without a *TP53* aberration tended to be older (Figure SC).

This led us to use age, *TP53*mut burden, and del17 frequency to predict the probability of PrR. The tight correlation between the del17 frequency and *TP53*mut burden meant that only *TP53*mut burden remained as a significant variable. Thus, the final model reads

$$\text{prob(PrR)} = \frac{1}{1 + \exp(-z)}, \quad (1)$$

with

$$z = -10.43 + 0.1245 * \text{age} + 0.0558 * \text{TP53mut.burden} \quad (2)$$

Both the coefficients are significant, at  $P < 0.001$ . The value  $z$  in equation (1) is a linear combination of age and the mutation burden. Such a classifier produces a receiver operating characteristic curve (Figure SD). Thus, for any cut-off value of

**Table I.** The descriptive statistics of the whole MCL cohort and the primary refractory (PrR) and non-PrR subgroups. The *P* value column concerns the difference in the respective characteristics between the PrR and the non-PrR group. A chi-square test or Fisher's exact (factorial) test was used. If the characteristics had more than two categories, they were grouped for the test. The grouping is explained in the brackets after the *P* value. Clinical stage categorised according to Ann Arbor scale. ECOG PS, Eastern Cooperative Oncology Group Performance Status. Morphology categorised as classical, blastic, or plasmablastic. MIPI, Mantle Cell Lymphoma International Prognostic Index. Induction regimen classified as CHOP (doxorubicin, cyclophosphamide, vincristine, and prednisone) or CHOP-like, intensive (high-dose AraC containing), or non-anthracyclin. Autologous stem cell transplant denoted as ASCT. Response evaluated after the first-line therapy as complete remission (CR), partial remission (PR), stable disease (SD), or disease progression (PD). Values of Kaplan–Meier estimate of the survival function at particular time-points quoted, *P* value of the corresponding log-rank test.

Variable	All patients, <i>n</i> = 114	Non-PrR patients, <i>n</i> = 87	PrR patients, <i>n</i> = 27	<i>P</i>
Male, <i>n</i> (%)	85 (74.6)	64 (73.6)	21 (77.8)	0.6
Age, years, median, (range)	66 (40–87)	66 (40–80)	70 (56–87)	<b>0.003</b>
Clinical stage (I/II/III/IV), <i>n</i> (%)	1 (0.9)/4 (3.5)/4 (3.5)/105 (92.1)	1 (1.1)/4 (4.6)/3 (3.4)/79 (90.8)	0/0/1 (3.7)/26 (96.3)	0.45 (IV vs other)
ECOG PS (0/1/2/3), <i>n</i> (%)	47 (41.2)/50 (43.9)/14 (12.3)/3 (2.6)	40 (46.0)/38 (43.7)/9 (10.3)/0	7 (25.9)/12 (44.4)/5 (18.5)/3 (11.1)	<b>0.03</b> (0–1 vs 2–3)
B-symptoms, <i>n</i> (%)	48 (42.1)	28 (32.2)	20 (74.1)	<b>&lt;0.001</b>
Bulky disease (≥10 cm), <i>n</i> (%)	19 (16.7)	13 (15.1)	6 (21.4)	0.36
CNS involvement, <i>n</i> (%)	8 (7.0)	1 (1.1)	7 (25.9)	<b>0.003</b>
Splenomegaly, <i>n</i> (%)	78 (68.4)	56 (64.4)	22 (81.5)	0.09
Ki-67 (≥30%/<30%), <i>n</i> (%)	37 (52.9)/33 (47.1)	23 (43.4)/30 (56.6)	14 (82.3)/3 (17.6)	<b>0.01</b>
Morphology (class./blast./pleom.), <i>n</i> (%)	39 (72.2)/6 (11.1)/9 (16.7)	33 (86.8)/1 (2.6)/4 (10.5)	6 (37.5)/5 (31.3)/5 (31.3)	<b>&lt;0.001</b> (class. vs other)
β <sub>2</sub> -microglobulin, <i>n</i> (%)	84 (80)	60 (75)	24 (96)	<b>0.02</b>
MIPI (low/inter./high), <i>n</i> (%)	21 (18.4)/32 (28.1)/61 (53.5)	19 (21.8)/25 (28.7)/43 (49.4)	2 (7.4)/7 (25.9)/18 (66.7)	0.17 (high vs other)
TP53 mutation, <i>n</i> (%)	37 (32.5)	18 (20.7)	19 (70.4)	<b>&lt;0.001</b>
Deletion 17p, <i>n</i> (%)	29 (25.4)	15 (17.2)	14 (51.9)	<b>&lt;0.001</b>
Induction (CHOP/HDAC/Non-A), <i>n</i> (%)	55 (48.2)/44 (38.6)/10 (8.8)	37 (42.5)/40 (46.0)/5 (5.7)	18 (66.7)/4 (14.8)/5 (18.5)	0.08 (CHOP vs other)
ASCT, <i>n</i> (%)	44 (38.6)	40 (46.0)	4 (14.8)	<b>0.007</b>
Maintenance rituximab, <i>n</i> (%)	59 (56.7)	55 (68.8)	4 (16.7)	<b>&lt;0.001</b>
Response (CR/PR/SD + PD), <i>n</i> (%)	60 (52.6)/31 (27.2)/13 (11.4)	54 (62.1)/21 (24.1)/6 (6.9)	6 (22.2)/10 (37.0)/7 (25.9)	<b>&lt;0.001</b> (CR vs. other)
1 year OS/1 year PFS, %	82/76	99/98	23/0	<b>&lt;0.001/&lt;0.001</b>
5 year OS/5 year PFS, %	54/39	65/51	10/0	<b>&lt;0.001/&lt;0.001</b>

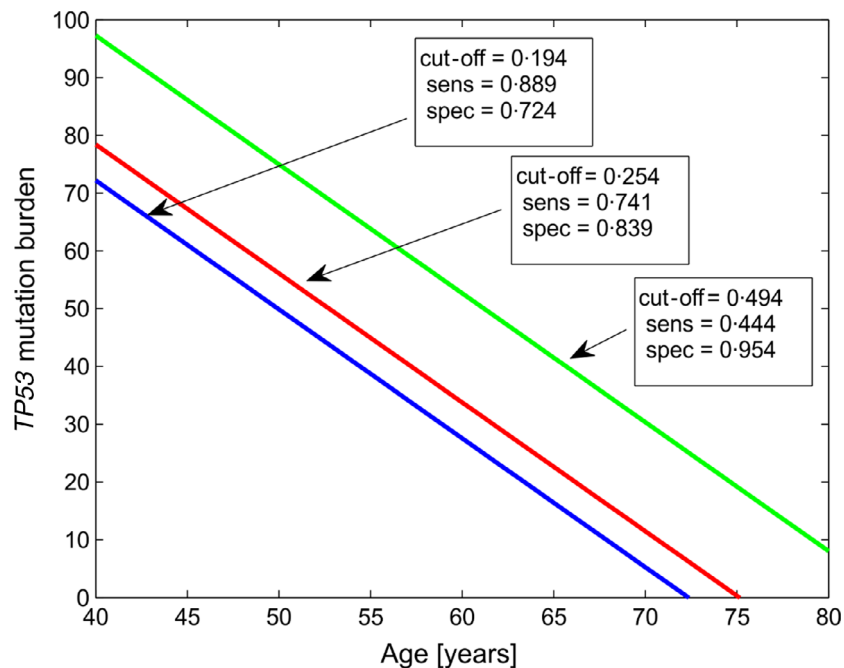


Fig 1. Visualisation of the decision boundary in the age-*TP53* mutation burden plane for three different values of sensitivity. See the text for an explanation. For example, for a sensitivity of 0.95, the prediction takes the following simple form: If  $2.23 \times \text{age} + TP53\text{mut\_load} > 146$ , the patient is predicted as PrR. Age is measured in years and the mutation burden in percentage. If a different sensitivity is required, the number 146 changes to a different threshold; otherwise the inequality remains the same.

the probability, the decision boundary has the form of a line in the age-burden plane (Fig 1). The patients above the decision line are predicted as PrR and the patients below the line as non-PrR. In our cohort, the patients with MCL with an aberration of the *TP53* gene had a significantly worse prognosis compared to those with the *TP53*-wild type. There was no difference in survival outcomes between those with *TP53*mut and del17. Our present findings thus further complement the results from Eskelund *et al.*<sup>3</sup>, who found lower OS and PFS in younger patients with MCL with *TP53* mutations. In contrast to that study, our calculations were performed on a whole real-life spectrum of younger and older patients with MCL.

The correlation between the *TP53*mut burden and del17 frequency in the PrR patients was almost perfect (Figure SB). Surprisingly, for the non-PrR patients, the del17 frequency tended to be higher than the *TP53*mut burden. We can hypothesise that an increasing *TP53*mut burden plays a more prominent role in the process of chemoresistance acquisition than del17. This is also supported by the observation that all the deletions were monoallelic. Partial preservation of the *TP53* gene function in patients with deleted *TP53* is therefore presumable, as previously discussed in another study.<sup>5</sup>

We found that the presence of a *TP53* aberration was a strong predictor of PrR regardless of age. Moreover, the PrR patients without any *TP53* aberration tended to be older (Figure SC). We noticed a strong correlation between older age and PrR.

The main goal of our present study was to identify PrR MCL patients. The usual cut-off age for transplant eligibility is 65 years. In our present cohort, six of the seven PrR patients aged <65 years had a *TP53*mut burden from 26% to 80%. Only seven of the 42 younger non-PrR patients had *TP53*mut, and the burden only exceeded 10% in three cases. Thus, we propose that all younger patients with a *TP53*mut burden >10% should be offered a novel treatment strategy.

In the elderly patients (aged >65 years), the association of *TP53*mut and PrR was not so close. With increasing age, other mutations plausibly complement *TP53* in mediating chemoresistance. In our present cohort, 12 of 15 elderly PrR patients had a *TP53*mut burden from 3% to 86%, but it was <50% in most cases. In all, 12 of 45 non-PrR patients had *TP53*mut, and in all but one the burden was <50%. In this group, it was not possible to set a burden threshold to discriminate PrR patients with sufficient sensitivity. Thus, we suggest all elderly patients with a *TP53*mut be indicated for an innovative up-front therapy.

To simplify the use of the predictive model (1)–(2), we introduce a nomogram in the form of a probability table (Figure SE).

Robust data suggesting appropriate therapy for patients with MCL with *TP53* alterations are still missing. Recently, Lin *et al.*<sup>8</sup> published promising results regarding younger patients treated with allogeneic stem cell transplants. The

authors found no significant differences in overall survival and the cumulative incidence of relapse among patients with and without *TP53* alterations. Another effective treatment option, also for elderly patients with *TP53* aberrations, might be the use of B-cell receptor (BCR) inhibitors with an acceptable toxicity profile.<sup>9</sup>

To conclude, we show that a high *TP53* mutation burden predicts chemoresistance in younger patients with newly diagnosed MCL regardless of the routinely used treatment strategy. We also demonstrate that age correlates positively with chemoresistance, irrespective of the type and frequency of *TP53* aberration. Although the results need to be validated in prospective clinical trials, we strongly support the implementation of a *TP53* mutation as a therapy classifier in all patients with newly diagnosed MCL.

## Acknowledgement



We would like to thank all the patients and their families for their goodwill and patience while participating in our study. Our research was supported by Ministry of Health and Ministry of Education, Youth and Sports of the Czech Republic and by the Charles University grants AZV 17-28980A, IGA\_LF\_2020\_002, IGA\_LF\_2020\_016, MH CZ – DRO (FNOL, 00098892), AZV16-32339A, AZV16-31092A, RVO-VFN64165, PROGRES Q26/LF1, and PROGRES Q28/LF1 [Correction was added on 14 September 2020, after first online publication: The Acknowledgement section has been updated with funding information in this current version.]

## Conflict of interest

The authors declare no conflict of interests.

## Data Availability Statement

The results of the FISH and NGS tests are available from the corresponding author upon reasonable request.

Ales Obr<sup>1</sup>   
 Pavel Klener<sup>2,3</sup>   
 Tomas Furst<sup>4</sup>  
 Eva Kriegova<sup>5</sup>  
 Zuzana Zemanova<sup>6</sup>  
 Helena Urbankova<sup>1</sup>  
 Andrea Jirkuvova<sup>1</sup>  
 Anna Petrackova<sup>5</sup>  
 Diana Malarikova<sup>2,3</sup>  
 Kristina Forsterova<sup>2</sup>  
 Barbora Cudova<sup>1</sup>  
 Lenka Sedlarikova<sup>5</sup>  
 Adela Berkova<sup>6</sup>  
 Nela Kasalova<sup>1</sup>  
 Tomas Papajik<sup>1</sup>  
 Marek Trneny<sup>2</sup>

<sup>1</sup>Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, <sup>2</sup>First Department of Medicine – Hematology, General University Hospital in Prague and First Faculty of Medicine, Charles University, Prague, <sup>3</sup>Institute of Pathological Physiology, First Faculty of Medicine, Charles University, Prague, <sup>4</sup>Department of Mathematical Analysis and Application of Mathematics, Faculty of Science, Palacky University, Olomouc, <sup>5</sup>Department of Immunology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc and <sup>6</sup>Center of Oncocytogenomics, Institute of Medical Biochemistry and Laboratory Diagnostics, General University Hospital and 1st Faculty of Medicine, Prague, Czech Republic.

E-mail: ales.obr@fnol.cz

The data were published in part at the 61st American Society of Hematology Annual Meeting, abstract no 3995.

**Keywords:** mantle cell lymphoma, chemoresistance, *TP53* aberration, mutation burden

First published online 30 August 2020

doi: 10.1111/bjh.17063

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Data S1.** Supplementary information

## References

1. Hoster E, Klapper W, Hermine O, Kluijn-Nelemans HC, Walewski J, van Hoof A, et al. Confirmation of the mantle-cell lymphoma international prognostic index in randomized trials of the European mantle-cell lymphoma network. *J Clin Oncol*. 2014;**32**:1338–46.
2. Klener P. Advances in molecular biology and targeted therapy of mantle cell lymphoma. *Int J Mol Sci*. 2019;**20**:4417.
3. Eskelund CW, Dahl C, Hansen JW, Westman M, Kolstad A, Pedersen LB, et al. *TP53* mutations identify younger mantle cell lymphoma patients who do not benefit from intensive chemoimmunotherapy. *Blood*. 2017;**130**:1903–10.
4. Delfau-Larue M-H, Klapper W, Berger F, Jardin F, Briere J, Salles G, et al. High-dose cytarabine does not overcome the adverse prognostic value of *CDKN2A* and *TP53* deletions in mantle cell lymphoma. *Blood*. 2015;**126**:604–11.
5. Obr A, Procházka V, Jirkuvová A, Urbánková H, Kriegova E, Schneiderová P, et al. *TP53* mutation and complex karyotype portends a dismal prognosis in patients with mantle cell lymphoma. *Clin Lymphoma Myeloma Leuk*. 2018;**18**:762–8.
6. Aukema SM, Hoster E, Rosenwald A, Canoni D, Delfau-Larue MH, Rymkiewicz G, et al. Expression of *TP53* is associated with the outcome of MCL independent of *MIPI* and *Ki-67* in trials of the European MCL Network. *Blood*. 2018;**131**:417–20.
7. Di Agostino S, Fontemaggi G, Strano S, Blandino G, D'Orazi G. Targeting mutant p53 in cancer: the latest insights. *J Exp Clin Cancer Res*. 2019;**38**:290.
8. Lin RJ, Ho C, Hilden PD, Barker JN, Giralt SA, Hamlin PA, et al. Allogeneic haematopoietic cell transplantation impacts on outcomes of mantle cell lymphoma with *TP 53* alterations. *Br J Haematol*. 2019;**184**:1006–10.
9. Hanna KS, Campbell M, Husak A, Sturm S. The role of Bruton's tyrosine kinase inhibitors in the management of mantle cell lymphoma. *J Oncol Pharm Pract*. 2020;**26**:1190–9.

APPENDIX J

**Petrackova A**, Turcsanyi P, Papajik T, Kriegova E. Revisiting Richter transformation in the era of novel CLL agents. *Blood Reviews*. *Blood Rev.* 2021;49:100824. (IF 2020: 8.250; Q1)



## Review

## Revisiting Richter transformation in the era of novel CLL agents

Anna Petrackova<sup>a</sup>, Peter Turcsanyi<sup>b</sup>, Tomas Papajik<sup>b</sup>, Eva Kriegova<sup>a,\*</sup><sup>a</sup> Department of Immunology, Faculty of Medicine and Dentistry, Palacký University and University Hospital Olomouc, Olomouc, Czech Republic<sup>b</sup> Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacký University and University Hospital Olomouc, Olomouc, Czech Republic

## ARTICLE INFO

## Keywords:

Chronic lymphocytic leukaemia  
Richter's transformation  
Ibrutinib  
Venetoclax  
TP53 disruption  
Genetics

## ABSTRACT

Richter transformation (RT) is the development of aggressive lymphoma – most frequently diffuse large B-cell lymphoma (DLBCL) and rarely Hodgkin lymphoma (HL) – arising on the background of chronic lymphocytic leukaemia (CLL). Despite recent advances in CLL treatment, RT also develops in patients on novel agents, usually occurring as an early event. RT incidence is lower in CLL patients treated with novel agents in the front line compared to relapsed/refractory cases, with a higher incidence in patients with *TP53* disruption. The genetic heterogeneity and complexity are higher in RT-DLBCL than CLL; the genetics of RT-HL are largely unknown. In addition to *TP53*, aberrations in *CDKN2A*, *MYC*, and *NOTCH1* are common in RT-DLBCL; however, no distinct RT-specific genetic aberration is recognised yet. RT-DLBCL on ibrutinib is frequently associated with *BTK* and *PLCG2* mutations. Here, we update on genetic analysis, diagnostics and treatment options in RT in the era of novel agents.

## 1. Introduction

Richter transformation (Richter syndrome, or RT) is defined by the World Health Organization (WHO) Classification of Tumours of Hematopoietic and Lymphoid Tissues as the development of an aggressive lymphoma arising on the background of chronic lymphocytic leukaemia (CLL) [1]. The first case was published in 1928 by Dr Maurice Richter, who reported that a patient with ‘reticular cell sarcoma of lymph nodes’ arising in ‘lymphatic leukaemia’ presented with rapidly fatal generalised lymphadenopathy and hepatosplenomegaly [2].

This life-threatening complication occurs in approximately 2–10% of CLL patients, more often during the disease course than at diagnosis [3,4]. Despite great advances in the treatment of CLL in recent years, RT also develops in patients treated with novel agents, as summarised in our review.

## 1.1. Forms of RT and clonality

The most common form of RT (~90%) is a diffuse large B-cell lymphoma (RT-DLBCL); Hodgkin lymphoma (RT-HL) is less frequent (~10%) [5]. Rarely, RT manifests as histiocytic/dendritic cell sarcoma, lymphoblastic lymphoma or other lymphomas (<1%) (Fig. 1) [6–9].

RT-DLBCL is further divided into two forms: clonally derived from CLL (~70%) and unrelated to the original CLL clone (~30%) [10–12].

With regard to RT-HL, the RT clonally derived from CLL occurs less frequently (~40%) than RT unrelated to the original CLL clone (~60%) [13] (Fig. 1).

RT clonally derived from CLL and RT unrelated to the original CLL represent distinct pathological and biological entities, with the first being more aggressive, resistant to chemotherapy, and having an extremely poor prognosis with an average survival time of six months, even in the era of novel agents [14]. RT in the form of clonally unrelated DLBCL responds to combination therapy used for *de novo* DLBCL, although it has a worse prognosis [15]. RT-DLBCL clonally derived from CLL compared to clonally unrelated RT-DLBCL also differ in the prevalence of genetic aberrations [14,16], as shown for *TP53* aberrations (60% vs 10–20%) [12,15].

Interestingly, clonally unrelated RT-HL usually develops on the background of *IGHV* unmutated CLL [13]. Clinical significance of the clonality in RT-HL is not known; however, RT-HL has an inferior prognosis when compared to *de novo* HL, but a better prognosis than RT-DLBCL [17].

The determination of the clonality of the RT to the underlying CLL is critical for proper management of RT, as patients with clonal RT may profit from novel treatment strategies and subsequent stem cell transplantation (SCT). The clonal relationship may be proven by molecular analysis of *IGHV* sequencing or analysis of clonal immunoglobulin rearrangement pattern; the restriction of the analysis to kappa or lambda

\* Corresponding author at: Department of Immunology, Palacký University and University Hospital, Hnevotinska 3, 775 15, Olomouc, Czech Republic.

E-mail addresses: [eva.kriegova@email.cz](mailto:eva.kriegova@email.cz), [eva.kriegova@fnol.cz](mailto:eva.kriegova@fnol.cz) (E. Kriegova).



light chains immunoglobulins is not sufficient [18]. Unfortunately, the clonal relationship between RT and the underlying CLL is currently not routinely assessed at most centres [19]. Therefore, the identification of novel markers easier to use in comparison with molecular techniques would be beneficial for clonality assessment. Recently, expression of programmed death 1 (PD-1) was nominated as a candidate clonality marker, as a correlation between high PD-1 expression on neoplastic large B-cells in RT tissue and clonal-relatedness to the underlying CLL was observed in RT-DLBCL, but not *de novo* DLBCL [20,21]. In addition to high PD-1 expression, clonally related RT-DLBCL showed at least partial expression of CD5 or CD23 which was not observed in clonally unrelated cases [21]. Further studies on a larger patient cohort are needed to confirm the utility of PD-1, CD5 and CD23 as potential markers for clonality assessment.

## 2. Morphological and clinical presentations of RT

### 2.1. Clinical characteristics that give rise to RT suspicion

Clinical symptoms associated with RT are variable and nonspecific [16,22]. Biopsy and histopathological evaluation are always necessary for a definitive RT diagnosis [22–26].

Suspicion of RT occurs in CLL patients who have persistent lymphadenomegaly, rapid clinical deterioration, and who develop ‘B symptoms’ on ongoing CLL treatment, or early after the end of treatment (Fig. 2). Common ‘B symptoms’ are fever, night sweats, weakness, weight loss, and fast progressive bulky lymphadenomegaly (>5 cm). Suspected RT is further supported by high serum lactate dehydrogenase levels and monoclonal gammopathy [27–29], and by less specific markers of paraneoplastic activity, such as hypercalcaemia and increased serum C-reactive protein in the absence of infection [30,31]. In about 50% of patients, decreased haemoglobin levels (<11 g/dL) and platelet count (<100,000/microL) are seen [29]. Patients developing RT on ibrutinib or venetoclax typically present with bulky nodal or

extranodal disease, similar to patients with highly aggressive B-cell lymphoma [3]. The current diagnostic algorithm for CLL patients with clinical RT suspicion is shown in Fig. 2, demonstrating the difficulties and complexity of RT diagnostics.

### 2.2. PET/CT of lesion and image-guided biopsy

In case of clinical RT suspicion, the fluorodeoxyglucose (<sup>18</sup>FDG) positron emission tomography/computed tomography (PET/CT) characterisation of lesions should be performed, and maximum standardised uptake value (SUVmax) evaluated (Fig. 2) [32]. The recommended cut-off SUVmax supporting RT diagnosis differs among studies, with the majority of them recommending an SUVmax≥5 threshold [22–24] and one study SUVmax≥10 [25]. The cut-off value of SUVmax≥5 reaches a very high negative predictive value (NPV) of 92–97% for RT [22–24], meaning that in the case of a negative <sup>18</sup>FDG PET/CT, the probability of RT is only 3–8%. However, cut-off SUVmax≥5 has a low specificity (47–80%) for RT detection [22–24], because PET-avid lesions may be also lymph nodes with expanded proliferation centres in cases of accelerated/aggressive CLL, other lymphoproliferative tumours, solid tumours metastases or lymph node infections [18].

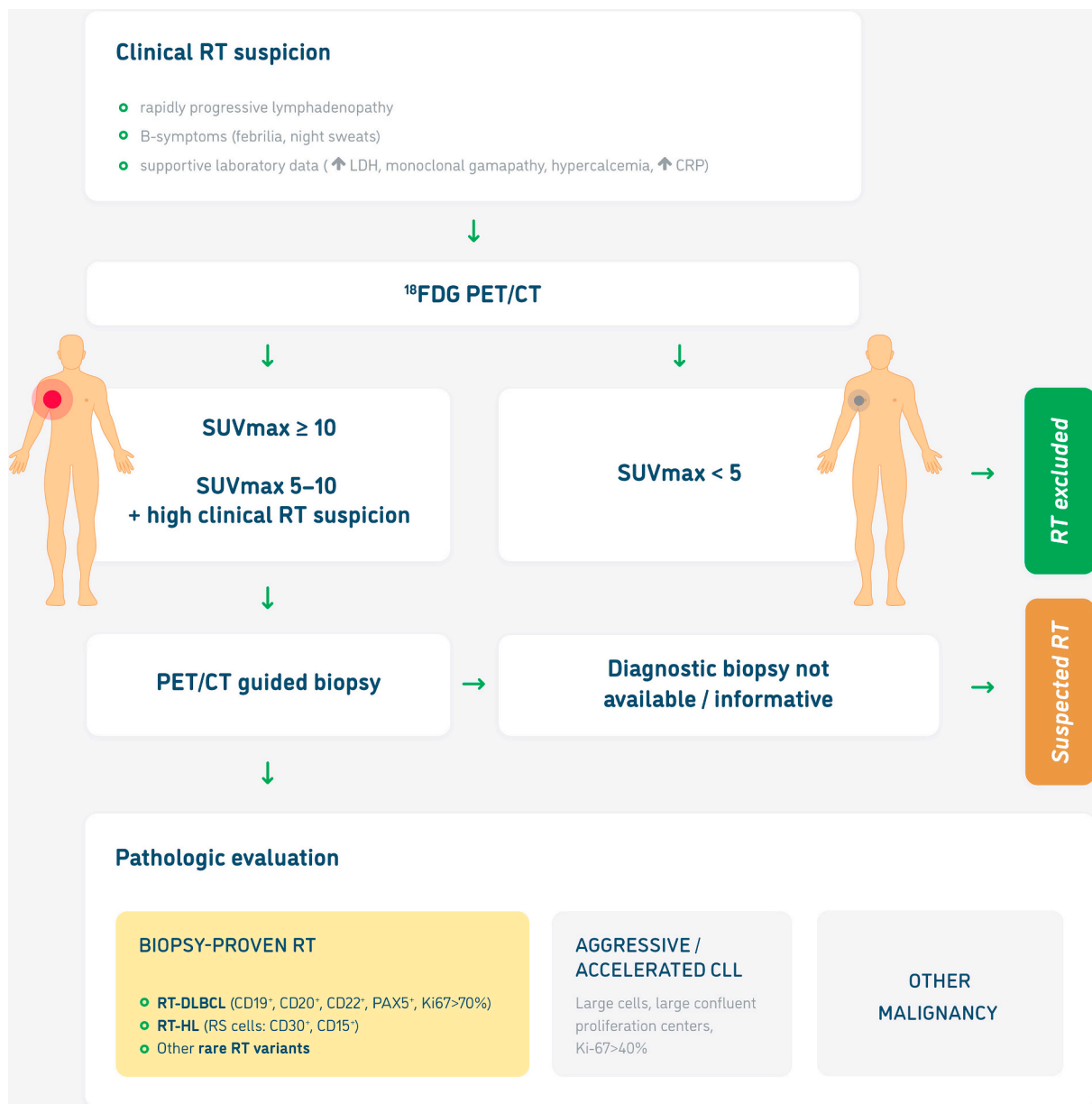
Regarding RT developed on novel agents, two studies investigated the utility of cut-off SUVmax≥10 in identifying RT among patients progressing on B-cell receptor (BCR) inhibitors [26,33]. Both studies emphasised the need to perform a tissue biopsy in patients with an SUVmax ≥5 and high clinical suspicion, as an SUVmax≥5 alone lacked both sensitivity and specificity to identify RT [26,33]. In 17 heavily pre-treated relapse/refractory (R/R) patients from three clinical trials [34–36] who developed RT on venetoclax, the reported median SUVmax was 15.5 (range 9.2–44) [37]. Examples of <sup>18</sup>FDG PET/CT scans of RT-DLBCL that developed at the time of CLL diagnosis and during ibrutinib treatment are shown in Fig. 3.

Collectively, current data support the opinion that the SUVmax≥10 threshold may identify RT with a higher specificity compared to the

# RICHTER TRANSFORMATION



Fig. 1. Forms of RT arising on the background of CLL. CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; HL: Hodgkin lymphoma; RT: Richter transformation.



**Fig. 2.** The diagnostic algorithm for CLL patients with clinical RT suspicion.

CLL: chronic lymphocytic leukaemia; CRP: C-reactive protein; DLBCL: diffuse large B-cell lymphoma; HL: Hodgkin lymphoma; LDH: lactate dehydrogenase; RS: Reed-Sternberg cells; RT: Richter transformation; SUVmax: maximum standardised uptake value; <sup>18</sup>F-FDG PET/CT: fluorodeoxyglucose positron emission tomography/computed tomography.

SUVmax $\geq$ 5 cut-off. However, some RT cases may occur with SUVmax values between 5–10, and therefore – if these patients have clinical features suggestive of RT – the performance of a diagnostic biopsy is also recommended (Fig. 2).

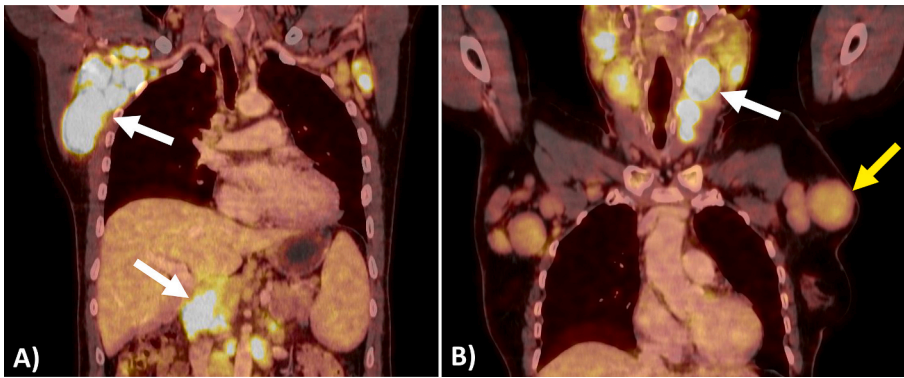
Importantly, in patients with suspected RT, the lesions that display the most avid <sup>18</sup>F-FDG uptake with the highest SUVmax should be selected for biopsy sampling, using core needle or lymph node excision, but not by fine-needle biopsy or aspiration [38]. However, biopsy is not feasible in some cases, or not informative: in these patients, the final diagnosis of RT by histological examination is not confirmed, and the patient is diagnosed with suspected RT (Fig. 2). Additionally, in patients with suspected RT that have bone lesions and severe cytopenia, the bone marrow biopsy is performed for complete patient staging [18]. Besides bone marrow involvement occasionally seen in RT, atypical neoplastic cells rarely occur also in the peripheral blood of RT patients [39,40]. Fig. 4 shows transformed cells in bone marrow aspirate together with an

occurrence of atypical neoplastic cells in peripheral blood from a single RT-DLBCL patient.

Recently, a non-invasive method of CT texture analysis for differentiation of RT-DLBCL and aggressive/accelerated CLL was published, demonstrating that lymph node architecture and vascularisation in RT-DLBCL differ significantly compared to aggressive/accelerated CLL [41]. Further efforts are needed to develop non-invasive techniques with sufficient specificity and sensitivity to identify RT patients, thus significantly improving RT diagnostics.

### 2.3. Morphology of RT

Histological confirmation is considered the gold standard for RT diagnosis [22–26,42]. It requires an experienced pathologist, as RT may be mimicked by many conditions, most frequently aggressive/accelerated CLL. A recent study demonstrated diagnostic misclassification in



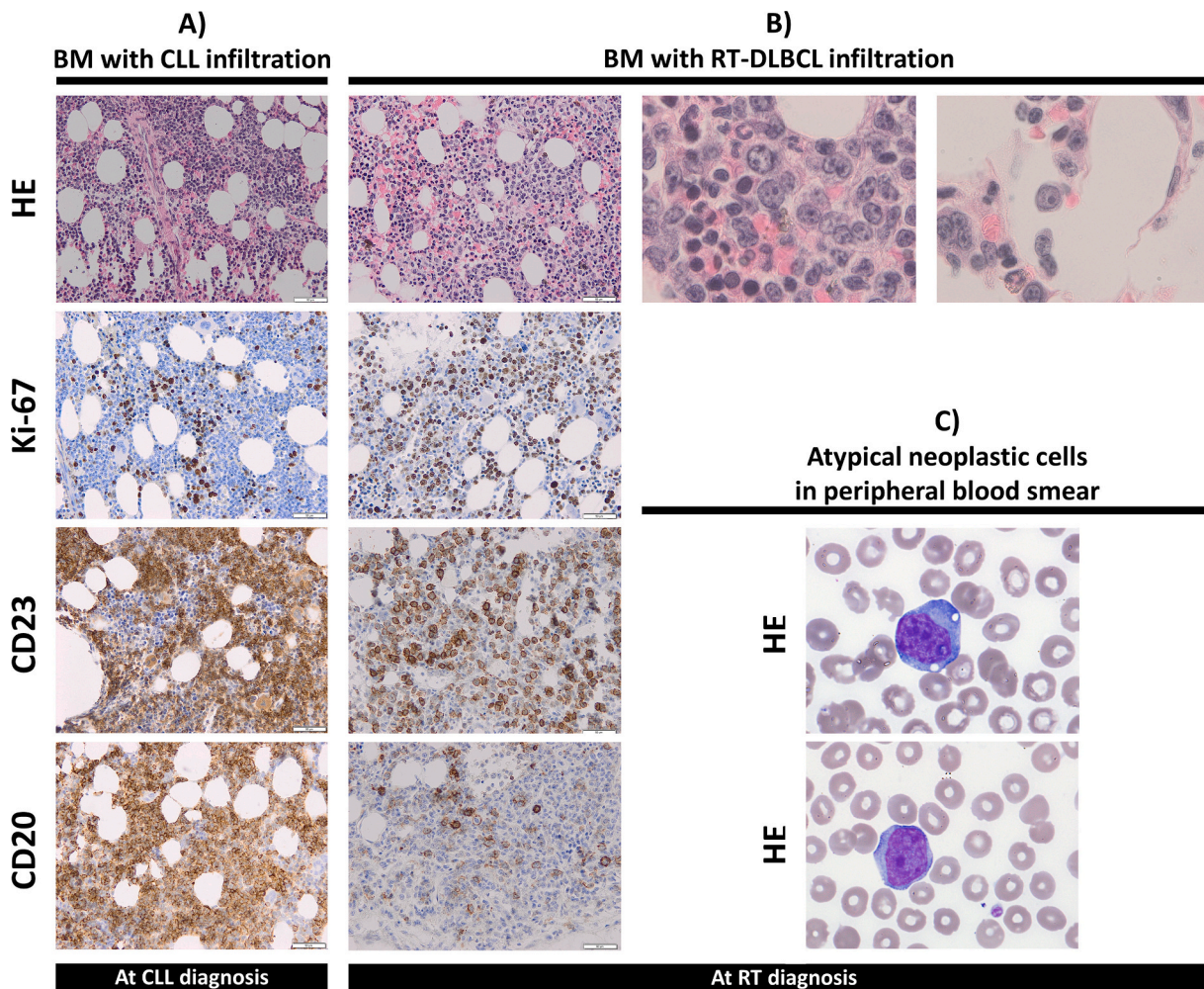
**Fig. 3.** PET/CT (<sup>18</sup>F) scans of patients with clonally related RT-DLBCL developed A) at the time of the CLL diagnosis and B) on ibrutinib treatment - representative examples.

A) In a patient who developed RT-DLBCL at the time of the CLL diagnosis, the CT scan shows dominant right lymphadenomegaly with maximal diameter 57x48 mm and SUVmax of 15.9, and abdominal lymphadenomegaly with SUVmax 16.7.

B) In a patient who developed RT-DLBCL on ibrutinib treatment, the CT scan shows numerous bilateral enlarged lymph nodes at the neck, with the largest one of 52x33 mm maximal diameter and SUVmax 24.5 (white arrow, for histology see Fig. 4B). Bilateral axillar lymphadenomegaly with 52x35 mm maximal diameter shows only mild <sup>18</sup>F uptake with SUVmax 3.4 (yellow arrow, for histology see Fig. 4A).

CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma.

large B-cell lymphoma; RT: Richter transformation; SUVmax: maximum standardized uptake value; <sup>18</sup>F PET/CT: fluorodeoxyglucose positron emission tomography/computed tomography.



**Fig. 4.** Histopathological assessment of bone marrow and peripheral blood smear in a CLL patient who developed RT-DLBCL on ibrutinib treatment. A) Bone marrow with CLL infiltration at the time of CLL diagnosis, B) bone marrow with RT-DLBCL infiltration from re-biopsy one year later at the time of RT diagnosis and C) atypical neoplastic cells in peripheral blood smear at the time of RT diagnosis; atypical cells have great irregular cores and rich, abundant basophilic cytoplasm. HE sections for bone marrow and peripheral blood together with immunohistochemistry analysis (Ki-67, CD23 and CD20) in bone marrow are shown. Original magnification for panels A)  $\times 200$ , B)  $\times 200$ ,  $\times 1000$ , C)  $\times 1000$ , respectively.

BM: bone marrow; CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; HE: hematoxylin and eosin; PB: peripheral blood; RT: Richter transformation.



18% of RT cases, therefore revision by a second expert pathologist is always recommended [43].

The morphology of RT-DLBCL is defined by WHO histological criteria [43,44] as follows: (i) large B-lymphoid cells with a nuclear size equal to or exceeding that of normal macrophage nuclei, or more than twice the size of a normal lymphocyte, and (ii) these cells must show a diffuse growth pattern, and not only present as small foci throughout the neoplasm. In most cases of RT-DLBCL, the diffuse effacement of sheets of large cells with centroblastic morphology are seen; those with immunoblastic features are rare. Common histology findings include mitotic figures, apoptotic bodies, as well as starry-sky pattern and tumour necrosis [45]. The majority of RT-DLBCL patients (80%) transform to an activated B-cell type of DLBCL (ABC; also known as non-germinal centre B-cell-like), whereas about 20% have a germinal centre B-cell-like (GCB) immunophenotype [12,46,47]. The representative examples of histopathological assessment of RT-DLBCL in lymph node biopsy (Fig. 5) and bone marrow aspirate (Fig. 4) are shown.

The most common condition that may mimic the RT is aggressive/accelerated CLL, which is characterised by expanded proliferation centres in lymph nodes that are broader than 20x field or becoming confluent [1,43]. Compared to RT, aggressive/accelerated CLL usually shows a monotonous proliferation of small lymphocytes with minimal mitotic activity and a low proliferation index, with varying sized proliferation centres and numbers of paraimmunoblasts [44]. Although data is limited, cases may also belong in this category when the Ki-67 proliferation index is >40%, or >2.4 mitoses in the proliferation centres [48,49]. These cases are reported to have an outcome intermediate between those of typical CLL and RT-DLBCL [24,48,49].

Regarding morphology of the RT-HL in the setting of CLL, two forms exist: i) type I, with isolated Reed–Sternberg (RS) cells in a rich background of CLL cells (Hodgkin-like lesion), and ii) type II, with RS cells dispersed in an inflammatory background typically seen in classical HL [50], thereby making it indistinguishable from *de novo* HL [46]. Type I may further transit to type II [13]. The clinical significance and biological features of these two morphological forms is unclear.

#### 2.4. Phenotype of RT

The phenotype of RT-DLBCL is characterised by the presence of large DLBCL cells, which carry B-cell markers CD19, CD20, CD22, PAX5, and monotypic surface immunoglobulin light chains. Positive expression of CD38, ZAP70, and CD49d is often seen, whereas CD5 and CD23 expression fluctuates [51]. The typical immunophenotype of ABC DLBCL is negative for CD10 and positive for MUM1/IRF4, whereas GCB immunophenotype typically shows positivity for CD10 and/or BCL6, with negativity for MUM1/IRF4 [12,46,47]. A high Ki-67 proliferation index (>70%) is common [52,53].

In RT-HL, RS cells are typical by CD30/CD15 phenotype, with variable CD20 expression, and occur on an appropriate polymorphous background of small T cells, epithelioid histiocytic cells, eosinophils, and plasma cells [13]. RT-HL cannot be diagnosed in the case of RS-like cells atypically expressing both CD30 and CD20 but lacking CD15, on the background of CLL [13,16]. The majority of RT-HL cases are EBV positive (~70%) [13].

### 3. Risk factors for RT-DLBCL

Currently, numerous genetic factors, and clinical and laboratory parameters have been nominated as candidate risk factors for RT [54], which may differ from risk factors for CLL progression [55].

Among key genetic risk factors associated with RT on chemoimmunotherapy are mutations and deletions of *TP53* and *CDKN2A* genes [3,16,55,56], as well as *NOTCH1* mutations [14,57]. Also, numerous other genetic abnormalities were associated with a high risk of RT: complex karyotype, 11q deletion, chromosome 12 trisomy, absence of deletion 13q [55], unmutated *IGHV* status, stereotyped BCR subset 8,

deregulated microRNA expression (miR-125a, miR-34a, miR-21, miR-146b, miR-181b, and miR-150) [58,59], and short telomere length of less than 5000 bp [3,10,12,60–64].

CLL patients with unmutated *IGHV* are at ~4-fold risk of RT-DLBCL relative to these with mutated *IGHV* [51,56]. Moreover, patients with unmutated stereotyped BCR subset 8 have a 70% probability of developing RT in the five-year horizon [64], and patients with mutations in *NOTCH1* have a 45% probability in the 15-year horizon [14,57]. High miR-125a and low miR-34a expression predicted RT development in ~50% of RT patients [58]. RT predisposition has also been described in patients with single nucleotide polymorphisms in *CD38*, *LRP4* and *BCL2* genes [4,65]. Among other parameters predisposing to RT are clinical factors (Binet stage B/C, lymphadenopathy) [55,56], biochemical factors (lactate dehydrogenase elevation) [55], high ZAP70, and CD38 expression [55]. In addition, a number of previous CLL treatment lines were associated with higher RT risk [29]. A summary of risk factors contributing to the RT-DLBCL is shown in Fig. 6. For RT-HL, the risk factors are largely unknown; however, prior fludarabine treatment has been associated with RT-HL risk [66].

### 4. Genetic pathogenesis of RT

Most information about genetic pathogenesis is known for the RT-DLBCL; no data exists regarding RT-HL. RT-DLBCL is characterised by a higher molecular heterogeneity and complexity than CLL, with no distinct unifying RT-specific genetic lesion [3,16]. Genetic aberrations in RT-DLBCL commonly involve *TP53* disruption (del(17p) and/or mutation), mutations in *NOTCH1* gene, loss of *CDKN2A* gene, and activation of *MYC* gene [18] (Fig. 7). At least one of these abnormalities/mutations is present in 90% of patients with RT [60]. In general, three molecular profiles based on these recurrent changes can be recognised: i) patients with *TP53* and *CDKN2A/B* aberration (~50% of cases), ii) patients with *NOTCH1* gene mutations (~30% of cases), and iii) a heterogeneous group of patients with further genetic aberrations (~20% of cases) [18]. Often, aberrant *MYC* gene activation correlates with the presence of *TP53* and *CDKN2A* aberrations [15,60].

#### 4.1. *TP53*, *MYC* and *CDKN2A* aberrations

Disruption of *TP53* (deletion and/or mutation) as well as *CDKN2A*, and eventually also *CDKN2B*, are among the most common genetic aberrations that occur in more than half of patients with RT clonally derived from CLL [18]. The *TP53* gene encodes the p53 protein and is one of the most prominent tumour suppressor genes. p53 has a central role in protecting the genomic integrity of a cell: it is activated by DNA damage and cellular stress (e.g. hypoxia, oncogene overexpression) and, as a transcription factor, it triggers the expression of many genes that further direct cell fate, either to the cell cycle arrest, DNA repair and/or apoptosis [67]. Impairment of *TP53* gene function (by deletion and/or mutation) is already well described as a significant negative prognostic factor in CLL [68,69]. More than 50% of patients with RT have a *TP53* aberration in the CLL clone before transformation [10,62]. A recent large study reported that *TP53* aberration is not only an independent risk factor for RT development, but can also be acquired during transformation [56].

The *TP53* disruption is often accompanied by *CDKN2A* aberrations that are frequently acquired at the time of transformation [12]. *CDKN2A* encodes the p16<sup>INK4A</sup> protein, which inhibits the activity of Cdk4 and 6 kinases, and thus negatively regulates cell cycle progression from G1 to the S phase [70]. It also encodes p14<sup>ARF</sup>, which is an MDM2 inhibitor [12]. The loss of *CDKN2A* is usually due to 9p21 deletion [15]. The *CDKN2B* gene encodes another negative cell cycle regulator, p15<sup>INK4B</sup>, and its loss is also relatively common in RT [15]. Interestingly, inactivating somatic mutations in *CDKN2A* are not described in RT in the literature, and the loss of function of these genes appears to be caused only by the deletion of the chromosomal part.



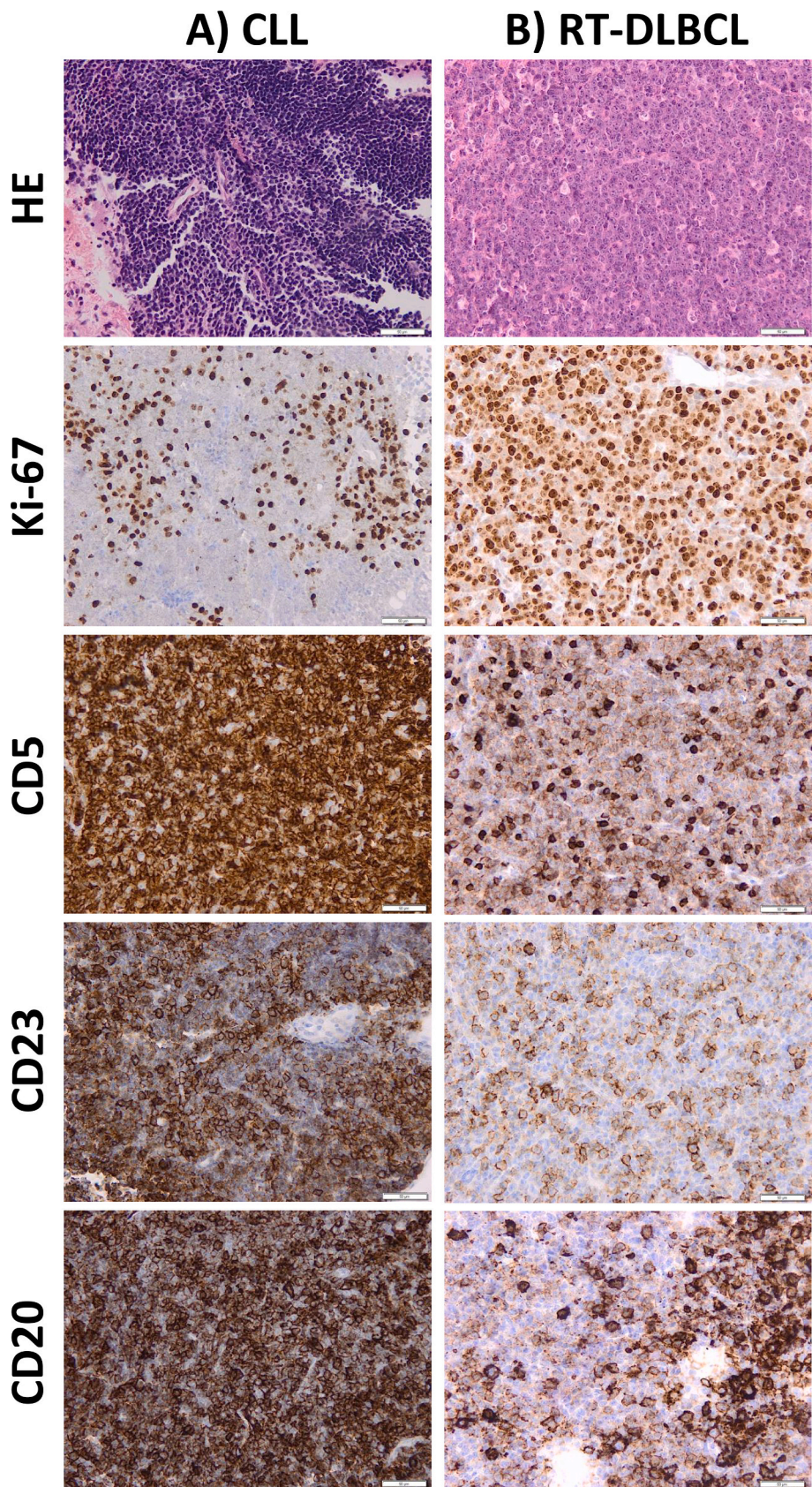
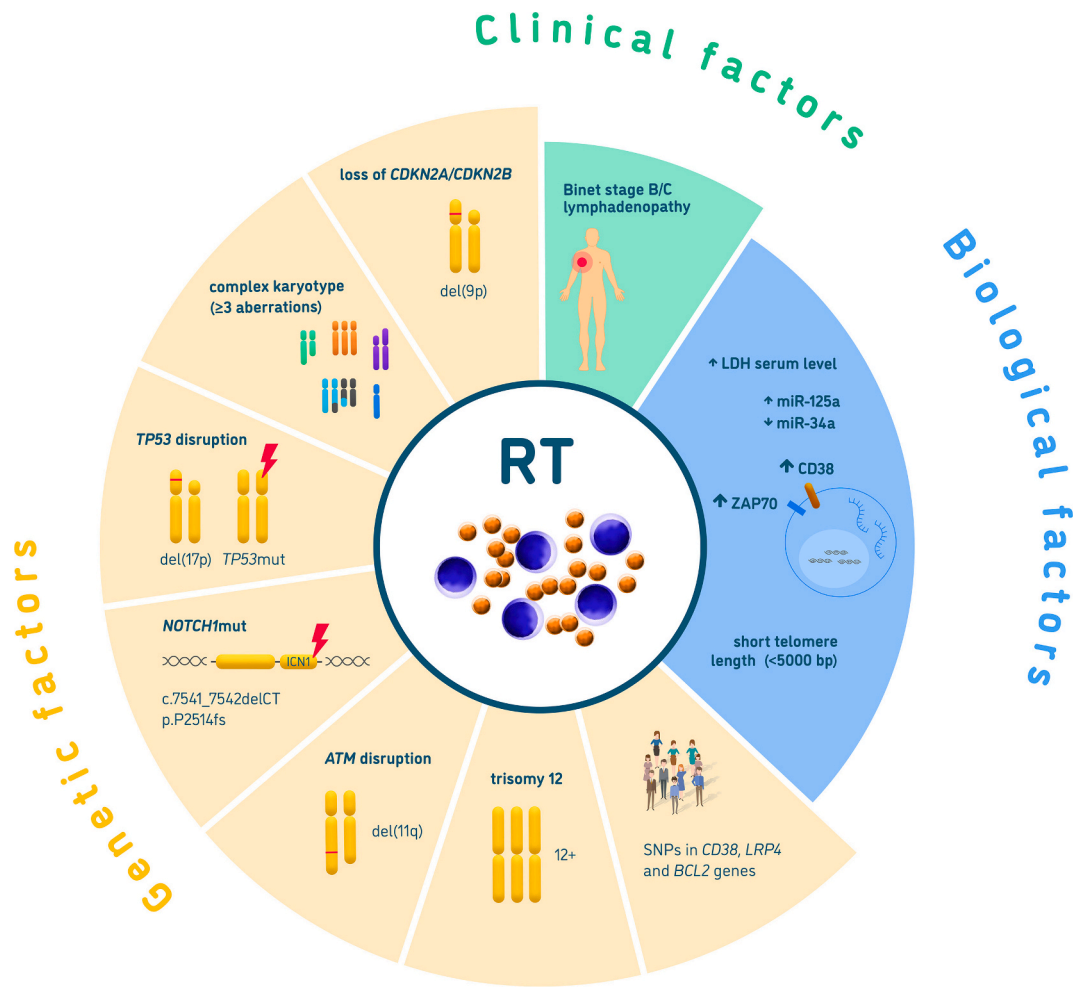


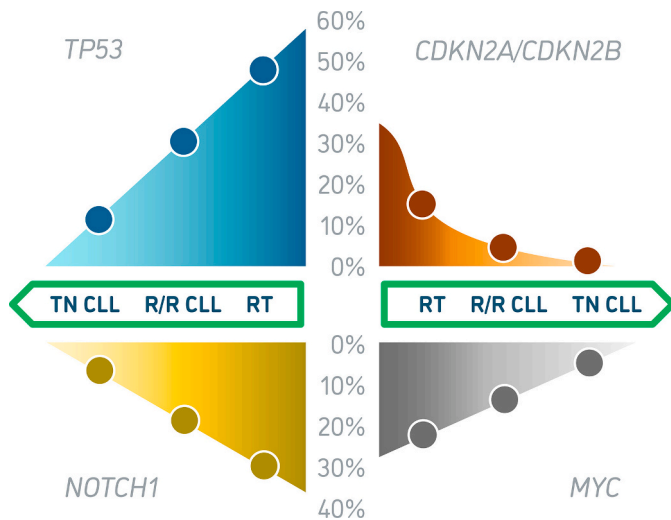
Fig. 5. Histopathological assessment of the tissue obtained from a patient with clonally related RT-DLBCL developed on idelalisib treatment - representative examples.

A) Lymph node tissue with CLL cell infiltrate and B) tumour tissue infiltrate of DLBCL cells. HE sections and immunohistochemistry analysis (Ki-67, CD5, CD23 and CD20) are shown. Original magnification  $\times 200$  for all panels.

CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; HE: hematoxylin and eosin; RT: Richter transformation.



**Fig. 6.** Risk factors associated with RT-DLBCL development in CLL. CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; LDH: lactate dehydrogenase; miR: microRNA; RT: Richter transformation; SNP: single nucleotide polymorphisms.



**Fig. 7.** Frequency of genetic abnormalities in *TP53*, *NOTCH1*, *CDKN2A* and *MYC* genes in RT-DLBCL compared to TN and R/R CLL. CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; R/R: relapsed/refractory; RT: Richter transformation; TN: treatment naïve.

In addition, aberrant *MYC* gene activation correlates with the presence of *TP53* and *CDKN2A* aberrations [15,60]. *MYC* acts as a transcription factor for an enormous number of genes (10–15% of all genes) [71], thereby significantly affecting cell growth and proliferation, regulating metabolism, adhesion, and mitochondrial function [72]. Aberrant activation of *MYC* is usually due to structural changes, e.g. translocation, where *MYC* comes under an active promoter (such as t(8;14)), or an amplification (8q24 amplification) [10,12,15]. Furthermore, it can be indirectly activated through deletions or mutations in the *MGA* gene that encodes a *MYC* antagonist [73,74].

**4.2. NOTCH1 mutations**

Activating mutations in the *NOTCH1* gene are found in approximately 30% of patients with RT clonally derived from CLL, and these patients do not simultaneously carry *TP53* and *CDKN2A* aberrations [18]. Mutations in *NOTCH1* are typically present in a CLL clone prior to RT, and these patients have a 45% chance of developing RT in the 15-year horizon [14,57,75]. Activating *NOTCH1* mutations in RT patients often occur concomitantly with chromosome 12 trisomy, unmutated *IGHV*, *ZAP-70* positivity and sometimes also with stereotyped BCR subset 8 [12,64,76,77].

*NOTCH1* encodes a transmembrane receptor that, upon ligand binding, undergoes conformational changes and proteolytic cleavage, resulting in the translocation of the cleaved short active fragment of ICN1 into the nucleus; ICN1 acts as a transcription factor for a number of



genes that promote cell proliferation (e.g. *CCND1*, *MYC*) and inhibit apoptosis (e.g. *BCL2*) [78–80]. The most common *NOTCH1* mutation, found in 80% of CLL patients, is the deletion of two bases (c.7541\_7542delCT) in the C-terminal PEST domain. This deletion causes a reading frameshift, leading to a premature stop codon [81]. Such a truncated protein lacks regulatory domains for its proteasomal degradation and hence its ‘lifetime’ is significantly extended over the normal state [81].

4.3. Genetic pathogenesis of RT-DLBCL developed on novel agents

First reports on genetic aberrations associated with RT-DLBCL that developed on novel agents include the same abnormalities in *TP53*, *CDKN2A*, *MYC*, and *NOTCH1* genes as RT that developed on chemotherapy (Fig. 8). In addition, the occurrence of *BTK* and *PLCG2* mutations was reported in patients developing RT on ibrutinib [60,82]. The largest body of information on the genetic nature of RT in the era of novel agents is available on patients treated with ibrutinib. Studies have shown that >70% of patients developing RT on ibrutinib had *TP53* abnormalities [82–87]. Among other detected abnormalities were: *ATM* disruption (by deletions and/or mutations); mutations in *SF3B1*, *NOTCH1* and *BIRC3* genes; loss of *CDKN2A*; trisomy 12; and *MYC* activation [82,88,89] (Fig. 9). In RT developed on venetoclax, *TP53* disruption was also present in >70% of patients [37,90].

Moreover, there is already early evidence about the cooperation between *TP53* and *CDKN2A/B* disruptions and BCR signalling, in an animal model of RT [91]. The authors demonstrated that genetic lesions associated with RT cooperate with BCR signals by downregulating the cell cycle negative regulators. Thus, this led us to suggest that the early manifestation of RT on ibrutinib may be associated with the cooperation of pre-existing genetic events predisposing to transformation, antigen-stimulated BCR and blocking of BCR signalling by the inhibitor. Further studies should clarify the influence of BCR signalling on the risk of driving towards RT in patients with *TP53* disruption, the most common genetic aberration associated with RT.

A very interesting study was conducted by Kadri et al. [82], who compared tumour RT tissue with a CLL clone in peripheral blood in six RT patients on ibrutinib. This study demonstrated frequent abnormalities of *MYC*, *CDKN2A*, *TP53*, and *NOTCH1* genes in RT, and most of these changes (60–95%) were present in both the CLL clone and RT tissue [82]. Mutations associated with the ibrutinib resistance in the *BTK* gene were detected in the CLL clone in four patients: in two patients, the same mutations were also confirmed in RT tumour tissue; in the third patient, one more *BTK* mutation than that found in the CLL clone was found in RT tissue; and, in the fourth patient, a resistance-associated mutation was found only in the CLL clone [82]. *TP53* abnormalities were detected in all patients with *BTK* mutations [82]. Based on the current studies that have analysed RT on ibrutinib [82,83,88,92,93], the

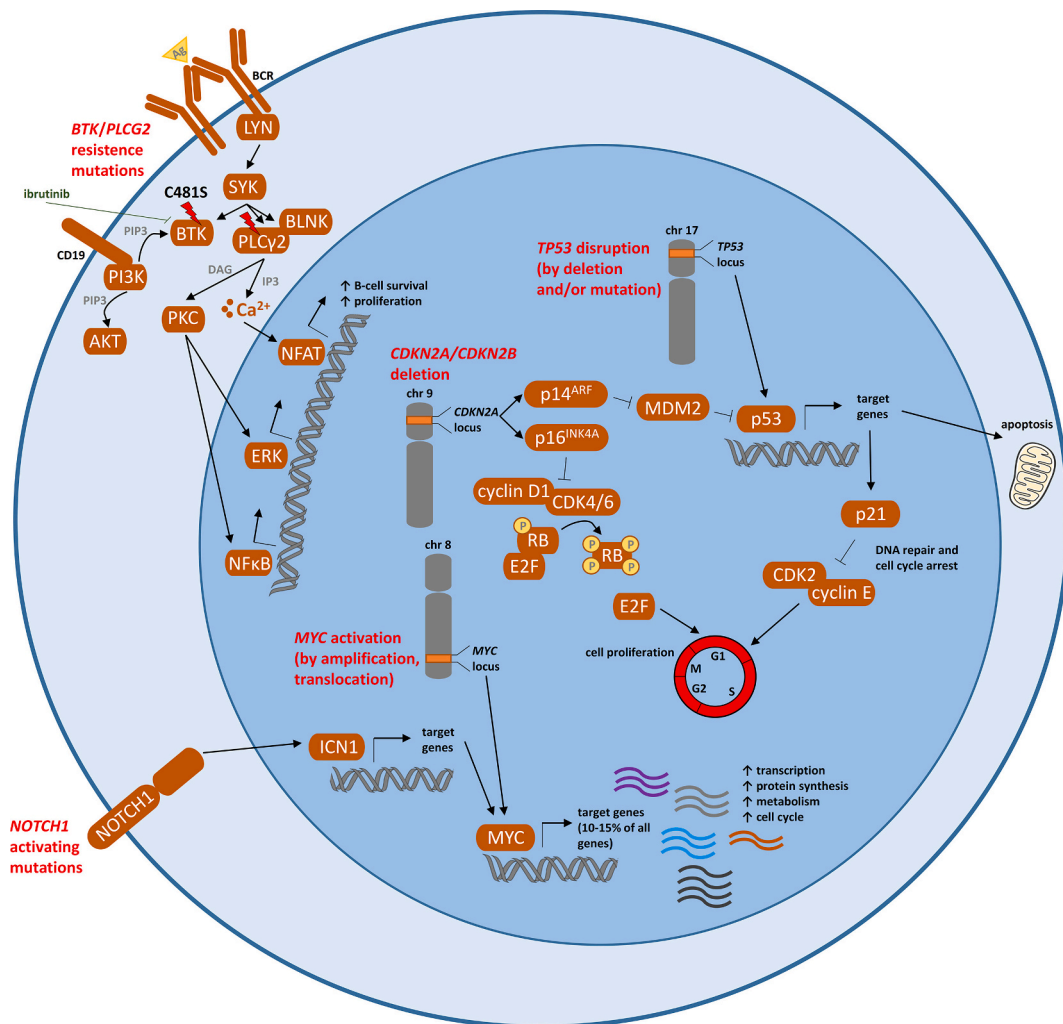
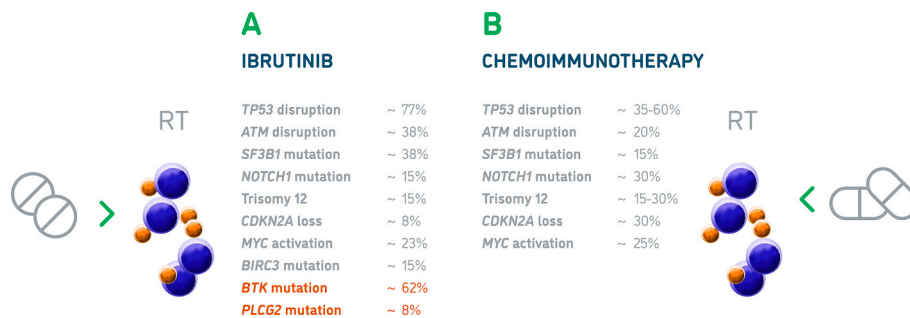
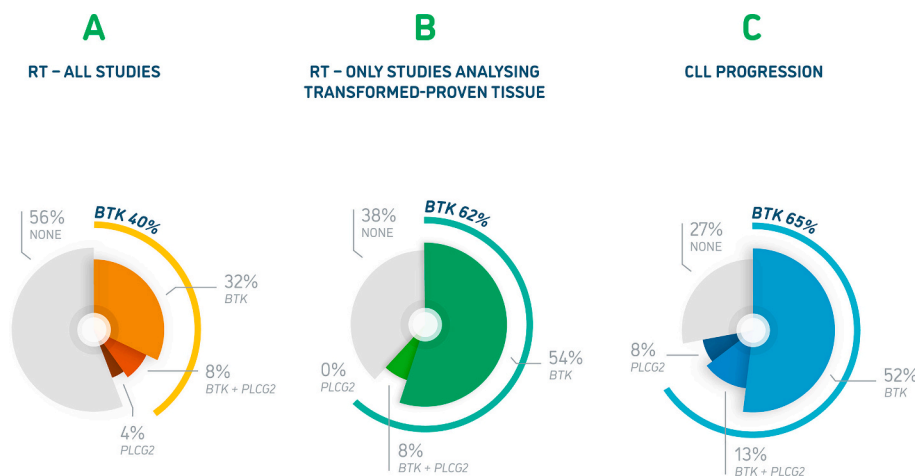


Fig. 8. Schematic representation of key genetic abnormalities with corresponding pathways associated with RT-DLBCL developed on novel agents. Key genetic abnormalities associated with RT-DLBCL are marked in red. DLBCL: diffuse large B-cell lymphoma; RT: Richter transformation.



**Fig. 9.** Overview of genetic aberrations associated with RT developed on A) ibrutinib and B) chemoimmunotherapy. Following studies were summarised: A) [82,88,89], B) [12,15,16,62,75,82,88,98,171]. RT: Richter transformation.



**Fig. 10.** Frequency of resistance-associated mutations in *BTK* and *PLCG2* genes in A) RT patients on ibrutinib from all studies reporting about resistance-associated mutations, B) RT patients on ibrutinib, when only studies analysing transformed tissues in all patients were included, and C) CLL patients experiencing progression on ibrutinib without transformation. Following studies were summarised: A) [82,83,88,89,92,93], B) [82,88,89], C) [82,83,88,92,93]. CLL: chronic lymphocytic leukaemia; RT: Richter transformation.

occurrence of *BTK* mutations has been demonstrated in 40% of patients, of whom 8% carried additional *PLCG2* mutation(s); *PLCG2* mutations alone were detected in only 4% of RT patients (Fig. 10). After the occurrence of *BTK* (C481) mutation, ibrutinib does not bind to BTK, causing the loss of its therapeutic effect and leading to BCR and NF- $\kappa$ B pathway activation; activating mutations in *PLCG2* gene may result in continuous BCR signalling, independently of BTK activation. Whether activated BCR signalling (as a result of resistance-associated mutations) together with *TP53* and *CDKN2A/B* disruptions are what drive the development of RT that has been reported in the animal RT model [91], deserves future investigations.

The occurrence of resistance-associated mutations in *BTK* and *PLCG2* genes on ibrutinib is reported as being lower in RT than in CLL progression, where these mutations have been detected in up to 73% of cases [82,83,92–95] (Fig. 10). The possible explanations for the lower incidence of *BTK/PLCG2* mutations in RT may lie in i) the different underlying biology of RT compared to CLL progression, and/or ii) difficulties in collection of transformed material for genetic analysis. However, when only the studies that analysed the RT-transformed tissue in all patients were evaluated [82,88,89], the incidence of *BTK/PLCG2* mutations in RT was similar to CLL progression (Fig. 10).

In RT, other resistance mechanisms that do not involve mutations in *BTK* or *PLCG2* may also exist, as shown in studies on *de novo* DLBCL without transformation. These studies reported that concomitant genetic lesions modified the effect of ibrutinib in ABC type of DLBCL [95–97]. *MYD88* mutations, when they present alone, were observed to confer primary resistance to ibrutinib, whereas when *CD79A* or *CD79B* mutations were present in addition to *MYD88* mutations, tumours generally exhibited responses to ibrutinib in ABC DLBCLs [95–97]. It has been suggested that *MYD88* mutations may mitigate the effect of

ibrutinib by providing an alternate means, independent of BTK, by which to activate NF- $\kappa$ B [96]. However, it is unclear why *MYD88* mutations in conjunction with *CD79A/B* mutations appear to render sensitivity to ibrutinib in ABC DLBCL [96].

## 5. Pitfalls of genetic analysis in RT

Regarding difficulties in the collection of RT-transformed material, only three studies used transformed biopsy-proven tissues for the genetic analysis of all patients who developed RT on ibrutinib [82,88,89]. This is particularly important as genetic aberrations that are present in the transformed tumour tissue can be absent in the circulating CLL clone [82,89]. Thus, peripheral blood sampling may provide inadequate information about the genetic architecture within the RT tissue [95]. Indeed, when only studies analysing transformed tissue in patients that developed RT on ibrutinib were taken into account [82,88,89], *BTK* mutations were detected in 62% of cases, 8% of which were combined with *PLCG2* mutations; no case with only *PLCG2* mutations was detected (Fig. 10). The frequency of *BTK* and *PLCG2* mutations was lower in RT on ibrutinib when also studies analysing no-RT-proven tissues were included (Fig. 10). Therefore, future studies on RT-proven tissues should confirm the association of *BTK* mutations and other genetic aberrations with RT on ibrutinib.

The next important consideration in the detection of resistance-conferring mutations is the required sensitivity of the method used for genetic testing, as mutations can be present at low variant allelic frequencies (VAFs) [82,88,92–94]. Usually, the mutant clone causing progression achieves dominance at the time of relapse, as shown in chronic myeloid leukaemia [95]. However, VAFs of *BTK* and *PLCG2* mutations at the time of transformation in the case of ibrutinib-treated



CLL patients vary substantially among studies, with the medians of detected VAFs reported as 98.5% [82], 63.2% [83], and 2.3% [88].

## 6. Detection of genetic aberrations associated with RT

To identify the patients at risk of developing RT, next-generation sequencing (NGS) and molecular cytogenetic analysis should be performed on appropriate biological samples, as different genetic architecture may be detected in RT tissue, peripheral blood, and bone marrow [82]. The genetic analysis should focus on aberrations in *TP53*, *NOTCH1*, *MYC*, and *CDKN2A* genes before initiating CLL treatment, as well as during progression [3,16]; in patients treated with ibrutinib/acalabrutinib, *BTK* and *PLGC2* genes should be analysed as well. Importantly, mutations in *TP53* and *NOTCH1*, as well as resistance-associated mutations, can also be detected at the subclonal level and might be later selected under CLL treatment pressure [98,99]. Using conventional targeted NGS, the sensitivity of up to 1–3% VAF can be reached when minimum technical requirements are applied in order to minimise the probability of false-positive and false-negative results [100]. Future studies should also elucidate how minor mutant subclones and their growth kinetics account for RT.

As an alternative, a non-invasive option for the monitoring of transformed clones in RT patients is the analysis of circulating tumour DNA (ctDNA) [89,101]. It has already been demonstrated that ctDNA analysis is able to identify RT-specific genetic aberrations in CLL [101], irrespective of the diseased compartments. Evidence on the usefulness of ctDNA analysis also arises from DLBCL studies, where it was used for the classification of transcriptionally defined tumour subtypes [102] and the detection of clonally represented somatic mutations [103]. However, ctDNA analysis is limited by the need for a sufficient amount of ctDNA in the sample, as the genetic changes belonging to ctDNA must be differentiated from the background of other cell-free DNA that highly correlate to changes found in white blood cells [104,105].

## 7. Update on RT incidence in the era of new agents

The RT incidence in CLL patients treated by chemoimmunotherapy ranges between 2–10% [3,4,11]. However, with the advent of novel agents used for CLL therapy, it is important to determine the occurrence of RT on these agents. With a growing number of studies, it is evident that the RT incidence on such agents is similar to chemoimmunotherapy. Here, we summarised nearly 40 studies reporting the incidence of RT in CLL/SLL patients treated with novel agents (Table 1). The median RT incidence in R/R CLL patients, as well as in treatment-naïve CLL patients with *TP53* disruption, was 6% on ibrutinib/acalabrutinib, as well as on venetoclax (Table 1, Fig. 11). The RT incidence is lower in patients who received novel agents in the front-line setting and were unselected for risk genetic factors, reaching the median incidence of 1% when treated with ibrutinib/acalabrutinib and 3% for venetoclax, respectively (Table 1, Fig. 11). When considering only heavily pre-treated R/R patients that experienced progressive disease on venetoclax, from three clinical trials [30–36], 21% developed RT-DLBCL and 4% RT-HL [37]. Regarding idelalisib/duvelisib, only a few studies were published, reporting a low incidence of RT of 1% (Table 1, Fig. 11).

RT in patients treated with novel agents usually occurs as an early event during the first 18 months of treatment, with a median overall survival (OS) of approximately six months [60,106]. Based on the published data, the median treatment duration on novel agents until RT development is 10 months (Table 1, Fig. 12). However, it is difficult to compare the time to RT development on novel agents to that on chemoimmunotherapy, as studies on RT on chemoimmunotherapy report only on the time from CLL diagnosis to RT. This has led to the common misperception that RT in CLL patients on chemoimmunotherapy is a late event [4]: when taking into account the published studies on RT on chemoimmunotherapy [19,51,56,64,107,108], the median time from CLL diagnosis to RT-DLBCL is ~2 years (Fig. 12). This indicates that RT

on chemoimmunotherapy is a relatively early complication of CLL, which may occur also in previously untreated patients [11]. Conversely, the RT-HL variant occurs later with a median of ~6 years from CLL diagnosis and mainly among previously treated patients [17,51].

Regarding CLL progression without transformation on novel agents, it occurs as a late event when compared to transformation, usually between the second and fourth year of treatment [92,94,106].

## 8. Treatment strategies for RT patients

Despite great advances in the treatment of CLL, the prognosis for patients with RT is extremely poor [3,16]. In addition, RT patients who received prior therapies have a worse prognosis and shorter survival when compared to previously untreated patients [19,56,109]. Patients developing RT on novel agents have particularly poor outcomes, with a highly aggressive disease course [16,18,83,110]. In RT on ibrutinib, median survival of ~4 months was reported [83,86,110]. In RT on venetoclax without prior ibrutinib treatment, median survival of ~12 months was reported in patients that were treated by ibrutinib for salvage therapy [37].

Treatment is always indicated for RT, and possible treatment options are considered in the context of the patient's performance status. However, treatment options for these patients are currently unsatisfactory [16,18].

### 8.1. Chemotherapy-based regimens with or without immunotherapy

In clonally unrelated RT-DLBCL, it is generally recommended to use the same treatment as in *de novo* DLBCL [16,60]. Patients with clonally related RT-DLBCL who are treated by conventional DLBCL chemoimmunotherapy only rarely reach complete remission (CR) and usually have short progression-free survival (PFS) and OS [109,111–117]. The chemoimmunotherapy treatment alone is palliative, and long-term survival can currently be achieved only by allogeneic or autologous SCT. Thus, it is strongly recommended that these patients should always be referred for clinical trials when available [3,16,18], followed by SCT if feasible. For cases with unknown clonality, clinical trials are also preferred. The treatment algorithm for RT is summarised in Fig. 13.

Among the chemoimmunotherapy options available for RT-DLBCL belongs the R-CHOP regimen of rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; this has an overall response rate (ORR) of 67%, a PFS and OS of 10 and 21 months, respectively, with low myelotoxicity, and a treatment related mortality of 3% [117]. Adding ofatumumab, instead of rituximab, did not improve the outcomes of patients with RT [109]. Other immunochemotherapy combinations – such as OFAR (oxaliplatin, fludarabine, ara-C, and rituximab), R-EPOCH (rituximab, etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin), hyper-CVAD (fractionated cyclophosphamide, vincristine, doxorubicin, and dexamethasone), DHAP (dexamethasone, cytarabine, and cisplatin), ESHAP (etoposide, methylprednisolone, cytarabine, and cisplatin), combination R+hyper-CVAD+GM-CSF/R+HDM-ara-C+GM-CSF (rituximab, fractionated cyclophosphamide, vincristine, liposomal daunorubicin, dexamethasone, and GM-CSF alternating with rituximab, methotrexate, ara-C and GM-CSF) – have shown more CRs in RT, but the toxicity was high and finally the OS was shorter (Table 2) [111–116].

For the RT-HL, it is recommended to use the same combination regimens as used in patients with advanced stage of HL, such as ABVD (doxorubicin, bleomycin, vinblastine, and dacarbazine). Patients who achieve a CR should be observed until progression; for those who do not achieve a CR, regimens for refractory HL should be used (Fig. 13). If RT-HL is treated as *de novo* HL, outcomes are not as good as in *de novo* HL patients [118–120]. The CR rates vary among retrospective studies from 17 to 37%, with an OS from 10 to 39.5 months (Table 2) [5,118,121].

**Table 1**  
Overview of studies reporting the incidence of RT in CLL/SLL patients treated with novel agents.

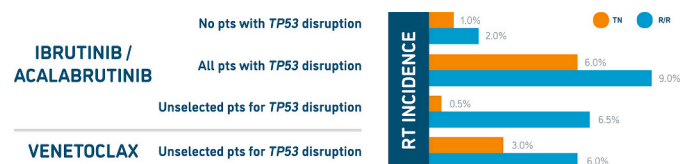
Author	Year published	Patient cohort	Number of subjects	Treatment	Del(17p) and/or TP53mut	Median follow-up (months)	Median duration of treatment with novel agents until RT (months, range)	Incidence of confirmed RT	RT variant
Ibrutinib/acalabrutinib									
O'Brien [140]	2014	TN	31	Ibrutinib	7%	22.1	9.6	1/31 (3%)	NA
Farooqui [141]	2015	TN	35	Ibrutinib	100%	24.0	8.1 (0.4-15.7)	2/35 (6%)	NA
Woyach [142]	2018	TN	182	Ibrutinib	del(17p) 5% TP53mut 9%	38.0	NA	0/181 (0%)	NA
Woyach [142]	2018	TN	182	Ibrutinib + rituximab	del(17p) 6% TP53mut 12%	38.0	NA	2/180 (1%)	NA
Mato [143]	2018	TN	80	Ibrutinib	del(17p) 37% TP53mut 12%	17.0	NA	1/80 (1%)	DLBCL: 1
O'Brien [144]	2019	TN	136	Ibrutinib	0%	36.0	NA	1/136 (1%)	NA
Dimou [145]	2019	TN	11	Ibrutinib	27%	24.0	NA	0/11 (0%)	NA
Moreno [146]	2019	TN	113	Ibrutinib + obinutuzumab	16%	31.3	NA	0/113 (0%)	NA
Burger [147]	2019	R/R, TN	208	Ibrutinib, Ibrutinib + rituximab	88%	36.0	NA	5/208 (2%)	DLBCL: 4; plasmablastic lymphoma: 1
Byrd [148]	2013	R/R	85	Ibrutinib	33%	20.9	NA	7/85 (8%)	NA
Farooqui [141]	2015	R/R	16	Ibrutinib	100%	24.0	7.2	1/16 (6%)	NA
UK CLL Forum [149]	2016	R/R	315	Ibrutinib	34%	16.0	13 pts within 12 months	18/315 (6%)	NA
O'Brien [150]	2016	R/R	144	Ibrutinib	100%	27.6	11 pts within 6 months and 6 pts within 25 months	17/144 (12%)	NA
Byrd [151]	2017	R/R	134	Acalabrutinib	23%	19.8	16.0 (2.0-16.0)	3/134 (2%)	NA
Mato [143]	2018	R/R	536	Ibrutinib	del(17p) 26% TP53mut 13%	17.0	NA	11/536 (2%)	DLBCL: 10; HL: 1
Huang [152]	2018	R/R	106	Ibrutinib	22%	17.8	NA	1/106 (1%)	DLBCL: 1
Nuttall [153]	2019	R/R	38	Ibrutinib	31%	23.0	NA	4/38 (11%)	DLBCL: 2; HL: 2
Byrd [154]	2019	R/R	195	Ibrutinib	del(17p) 32% TP53mut 51%	44.0	13.6 (1.7-27.8)	14/195 (7%)	DLBCL: 9; HL: 3; prolymphocytic lymphoma: 2
Winqvist [155]	2019	R/R	95	Ibrutinib	63%	30.0	14.0 (4.0-36.0)	12/95 (13%)	NA
Awan [156]	2019	R/R (all patients ibrutinib-intolerant)	33	Acalabrutinib	del(17p) 38% TP53mut 30%	19.0	NA	1/33 (3%)	NA
O'Brien [144]	2019	R/R	135	Ibrutinib	0%	44.0	NA	6/135 (4%)	NA
Fraser [157]	2019	R/R	289	Ibrutinib + bendamustine + rituximab	0%	34.8	NA	0/289 (0%)	NA
Dimou [145]	2019	R/R	47	Ibrutinib	22%	24.0	10.6 (1.0-35.9)	6/47 (13%)	NA
Idelalisib/duvelisib									
O'Brien [158]	2015	TN	64	Idelalisib + rituximab	14%	22.4	NA	0/64 (0%)	NA
Sharman [159]	2019	R/R	110	Idelalisib + rituximab	42%	18.0	NA	1/110 (1%)	NA
Zelenetz [160]	2017	R/R	207	Idelalisib + bendamustine + rituximab	33%	14.0	NA	4/207 (2%)	NA
Flinn [161]	2018	R/R	160	Duvelisib	19%	22.4	NA	0/160 (0%)	NA
Venetoclax									
Cramer [162]	2018	TN	34	Bendamustine + obinutuzumab + venetoclax	del(17p) 9% TP53mut 17%	16.0	NA	0/34 (0%)	NA
Flinn [163]	2019	TN	32		17%	26.7	NA	2/32 (6%)	DLBCL: 1; HL: 1

(continued on next page)

**Table 1** (continued)

Author	Year published	Patient cohort	Number of subjects	Treatment	Del(17p) and/or TP53mut	Median follow-up (months)	Median duration of treatment with novel agents until RT (months, range)	Incidence of confirmed RT	RT variant
Stilgenbauer [36]	2018	R/R, TN	158	Venetoclax + obinutuzumab		26.6	7.7 (0.4-27.8)	21/158 (13%)	NA
Roberts [164]	2016	R/R	116	Venetoclax		17.0	11 pts within 12 months	18/116 (16%)	NA
Seymour [35]	2017	R/R	49	Venetoclax + rituximab	del(17p) 19% TP53mut 31%	28.0	all pts within 9 months	5/49 (10%)	NA
Seymour [165]	2018	R/R	194	Venetoclax + rituximab	del(17p) 27% TP53mut 25%	23.8	NA	6/194 (3%)	NA
Jones [166]	2018	R/R (all patients after ibrutinib treatment)	91	Venetoclax	del(17p) 47% TP53mut 33%	14.0	17.0 (4.8-21.8)	5/91 (6%)	NA
Coutre [167]	2018	R/R (all patients after idelalisib treatment)	36	Venetoclax	31%	14.0	3.3 (2.5-4.0)	2/36 (6%)	NA
Cramer [162]	2018	R/R	29	Bendamustine + obinutuzumab + venetoclax	del(17p) 28% TP53mut 40%	16.0	NA	3/29 (10%)	NA
Rogers [168]	2018	R/R	12	Obinutuzumab + ibrutinib + venetoclax	13%	24.4	NA	0/12 (0%)	NA
Eyre [169]	2019	R/R (all patients after BTKi and PI3Ki treatment)	105	Venetoclax	48%	15.6	NA	9/98 (9%)	NA
Flinn [163]	2019	R/R	43	Venetoclax + obinutuzumab	55%	29.3	NA	1/43 (2%)	DLBCL: 1
Roeker [170]	2019	R/R	297	Venetoclax	45%	11.0	NA	13/297 (4%)	NA

BTKi: BTK inhibitors; CLL: Chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; HL: Hodgkin lymphoma; NA: not applicable; PI3Ki: PI3K inhibitors; R/R: relapsed/refractory, SLL: small lymphocytic lymphoma; TN: treatment-naïve; TP53mut: mutated TP53



**Fig. 11.** The incidence of RT in CLL patients treated with ibrutinib and venetoclax in TN and R/R CLL patients subdivided according to the TP53 disruption.

CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; R/R: relapsed/refractory; RT: Richter transformation; TN: treatment naïve.

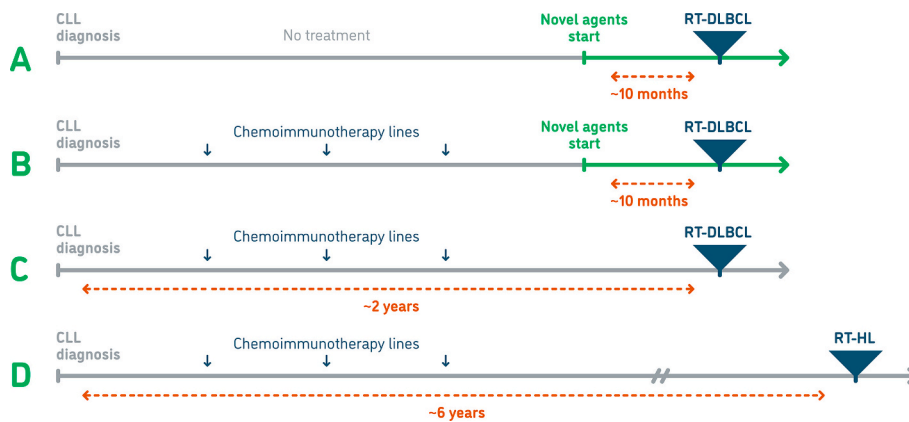
**8.2. Novel treatment options**

Several emerging treatment options have been suggested for RT patients, many of them in clinical trials (Fig. 13). So far, the most promising in the context of historical results, is combination therapy of venetoclax with dose-adjusted R-EPOCH [122]. The study on 27 RT patients enrolled showed an effectivity with an ORR of 59% (and 48% CRs). Of note, this study included six previously untreated RT patients, who generally respond better to R-CHOP and R-EPOCH type of therapy [19,56,62]. Only one patient with CR has progressed and eight underwent allogeneic SCT, with reported CR up to 2.5 years post SCT. The study had a median follow-up of 9.3 months (range 0.6-30) and median PFS and OS were both 16.3 months. According to these results,

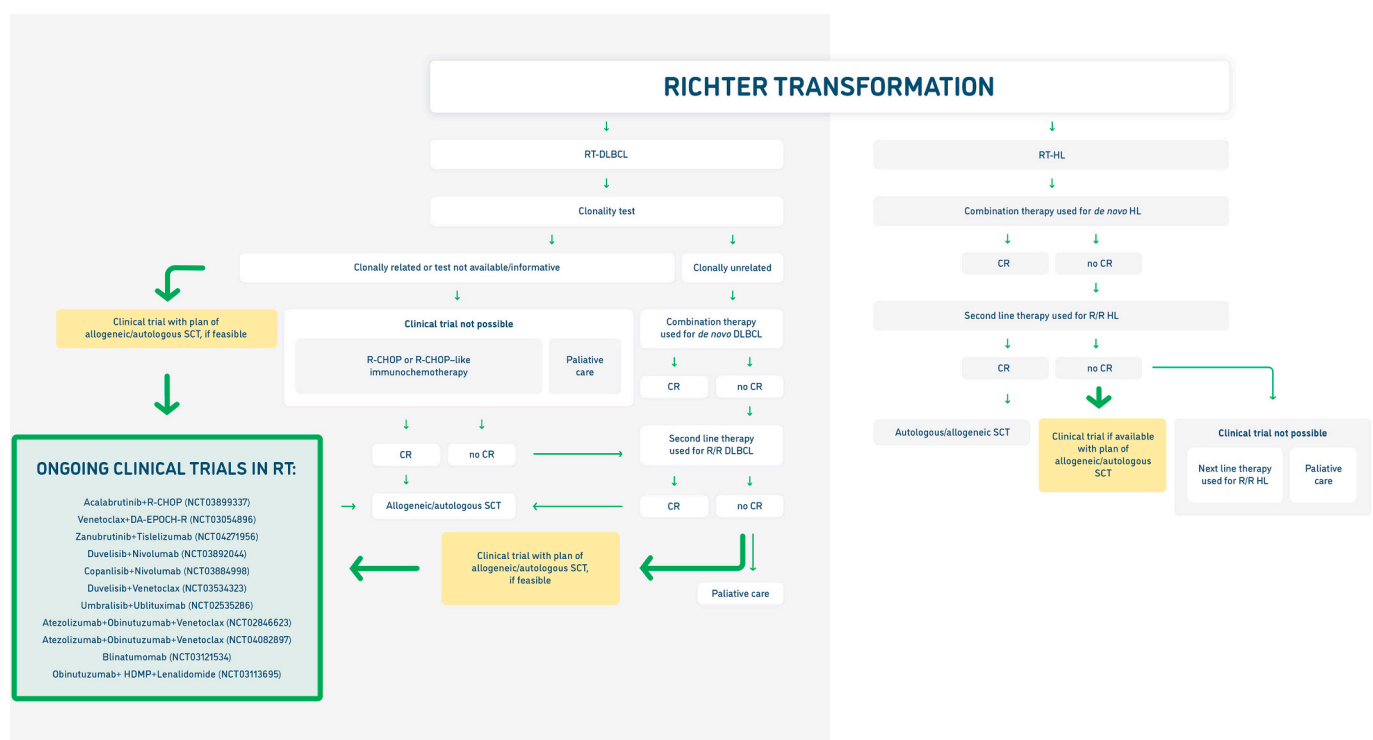
venetoclax with dose-adjusted R-EPOCH showed potential either as a bridge therapy to SCT or venetoclax maintenance.

Next, ibrutinib in monotherapy showed clinical activity in a small RT patient cohort with prior chemoimmunotherapy: of the four RT patients, one achieved CR, two achieved a partial response, and one showed clinical benefit [123]. Idelalisib demonstrated similar activity in RT as ibrutinib: ORR was reached in three, and CR in one out of four patients [124]. The activity of acalabrutinib – a second-generation BTK inhibitor – as a single agent, was studied in 29 heavily pre-treated patients with RT: the ORR achieved was 38%, with 14% CRs [125].

Another treatment approach suggested for RT, particularly for RT-HL, is blockade of PD-1. The expression of the PD-1 is increased in RT compared to *de novo* DLBCL [20,21]; moreover, there is first evidence that all high-risk RT patients with TP53 deletion exhibit high PD-1 expression [20]. A study in 9 RT patients of PD-1 inhibitor monotherapy, pembrolizumab, has demonstrated clinical efficacy, with an ORR of 44% (CR 11%), a PFS of 5.4 months, and an OS of 10.7 months [87]. To validate these findings, the activity of pembrolizumab was evaluated in 23 R/R RT patients (21 RT-DLBCL, 2 RT-HL) following one or more previous treatments [126]. The ORR was 13% with one CR (in a case of RT-HL) and two partial responses (one case of RT-HL, one RT-DLBCL). The RT-DLBCL patient who achieved a partial response received ibrutinib concomitantly with pembrolizumab. One responder subsequently underwent allogeneic SCT and remained in remission at data cut-off, whereas two others experienced disease progression after 2.7 and 6.2 months, respectively. This study demonstrated



**Fig. 12.** The timeline of RT development on novel agents and chemoimmunotherapy. CLL: chronic lymphocytic leukaemia; DLBCL: diffuse large B-cell lymphoma; HL: Hodgkin lymphoma; RT: Richter transformation.



**Fig. 13.** The treatment algorithm for RT patients and ongoing clinical trials. CLL: chronic lymphocytic leukaemia; CR: complete remission; DLBCL: diffuse large B-cell lymphoma; HL: Hodgkin lymphoma; R/R: relapsed/refractory; R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone; RT: Richter transformation; SCT: stem cell transplantation.

pembrolizumab activity in RT-HL, but not in RT-DLBCL patients, most of whom experienced disease progression early or died from adverse events associated with underlying malignancy. Another two patients subsequently underwent allogeneic SCT after additional therapy (venetoclax for one patient, and venetoclax and ibrutinib for the other patient): of them, one died following disease progression and one was alive at data cut-off [126]. Another group reported two RT patients who had high PD-1 expression and both showed objective response to the anti-PD-1 therapy [20]. The authors proposed that patients with high PD-1 expression would benefit most from therapy targeting PD-1/PD-L1.

It is possible that the combination of PD-1 inhibitor with other novel agents could improve outcomes of RT-DLBCL patients, a strategy that is currently being tested in several trials. The first promising results were shown for the triple combination of pembrolizumab with umbralisib, a PI3Kδ inhibitor, and the anti-CD20 monoclonal antibody ublituximab

[127]. The study on five RT patients, all ibrutinib refractory, reached an ORR of 40%; two patients reached durable CR (20+ and 12 months, respectively). Next, combination of PD-1 inhibitor nivolumab with ibrutinib, in a cohort of 20 patients, showed an ORR of 65%, with only 10% CR [128]. Most responses were transient, with a PFS of 5 months and an OS of 10.3 months [128]. Another trial of the nivolumab and ibrutinib combination, in a cohort of 23 RT patients (11 of whom had had previous exposure to ibrutinib), demonstrated an ORR of 43% with 35% CR, and the median duration of response was 9.3 months [129]. Also, atezolizumab – a humanised antibody targeting PD-L1 – when used alone or in combination with the CD20 monoclonal antibody obinutuzumab and venetoclax demonstrated safety and efficacy in RT-DLBCL [130].

Among other novel agents lies selinexor, a CRM1/XPO1 inhibitor that blocks nucleo-cytoplasmic shuttling of the tumour suppressor

**Table 2**  
Overview of published regimens for RT. Updated table published by Rossi et al (2018) [16].

Author	RT variant	Number of subjects	Regimen	ORR, %	CR, %	PFS/FFS, months	OS, months	Grade 3-4			TRM, %
								Neutropenia, %	Thrombocytopenia, %	Infection, %	
Chemoimmunotherapy regimens in RT-DLBCL											
Dabaja [111]	DLBCL	29 (3 pts were not RT)	Hyper-CVAD	41	38	NA	10	100	79	50	14
Durot [112]	DLBCL	28	DHAP, ESHAP	43	25	1	8	83	82	43	18
Tsimberidou [113]	DLBCL	30	R+hyper-CVAD+GM-CSF/R+HDM-ara-C+GM-CSF	43	38	NA	8	100	40	39	22
Tsimberidou [114]	DLBCL	35	OFAR1	50	20	3	8	85	95	8	3
Tsimberidou [115]	DLBCL	31	OFAR2	38	6	3	6	89	77	17	8
Langerbeins [117]	DLBCL	60	R-CHOP	67	7	10	21	55	65	28	3
Eyre [109]	DLBCL	37	O-CHOP	46	27	6	11	33	25	51	0
Rogers [116]	DLBCL	46	R-EPOCH	37	20	3.5	5.9	NA	NA	NA	NA
Novel agent therapies in RT-DLBCL											
Kuruvilla [131]	DLBCL	8	Selinexor	40	0	4	NA	NA	NA	NA	NA
Hillmen [125]	DLBCL	29	Acalabrutinib	38	14	3	NA	10	NA	NA	NA
Tsang [123]	DLBCL	4	ibrutinib	75	25	NA	NA	NA	NA	NA	NA
Visentin [124]	DLBCL	4	Idelalisib	75	25	NA	NA	NA	NA	NA	NA
Davids [110]	DLBCL	7	Venetoclax	43	0	NA	NA	NA	NA	NA	NA
Jain [129]	DLBCL	23	Nivolumab+ibrutinib	43	35	9.3	10.3	NA	NA	NA	NA
Younes [128]	DLBCL	20	Nivolumab+ibrutinib	65	10	5	10	NA	NA	NA	NA
Ding [87]	DLBCL	9	Pembrolizumab	44	11	5.4	10.7	NA	NA	NA	NA
Armand [126]	DLBCL	21	Pembrolizumab	4.8	0	NA	NA	NA	NA	NA	NA
Davids [122]	DLBCL	27	Venetoclax+R-EPOCH	59	48	16.3	16.3	58	50	52	11
Mato [127]	NA (all ibrutinib refractory)	5	Umbralisib + ublituximab + pembrolizumab	40%	40%	16.0	NA	NA	NA	NA	NA
Chemoimmunotherapy regimens in RT-HL											
Tsimberidou [29]	HL	18	ABVD (28%), CVPP (17%), CVPP/ABVD (5.5%), CHOP (5.5%), R-CHOP (5.5%), other (28%), no therapy (11%)	44	17	5	10	NA	NA	6	NA
Bockorny [5]	HL	86	ABVD (31%), MOPP (16%), CHOP (13%), other (40%)	NA	27	NA	19	NA	NA	NA	NA
Tadmor [121]	HL	16	ABVD (37.5%), MOPP+ABV (19%), BEACOPP (6%), Escalated BEACOPP (12%), ESHAP (6%), DVIP (12%), autoSCT (6%)	NA	37	NA	39.5	NA	NA	NA	NA
Novel agent therapies in RT-HL											
Armand [126]	HL	2	Pembrolizumab	100	50	NA	NA	NA	NA	NA	NA

autoSCT: autologous stem cell transplantation; ABVD: adriamycin, bleomycin, vinblastine, and dacarbazine; BEACOPP: bleomycin, etoposide, adriamycin, cyclophosphamide, oncovin, procarbazine, prednisone; CHOP: cyclophosphamide, doxorubicin, vincristine and prednisone; O-CHOP: CHOP plus ofatumumab; CR: complete response rate; CVPP: cyclophosphamide, vinblastine, procarbazine, and prednisone; DLBCL: diffuse large B-cell lymphoma; DHAP: dexamethasone, cytarabine, and cisplatin; DVIP: dexamethasone, etoposide, ifosfamide, cisplatin; ESHAP: etoposide, methylprednisolone, cytarabine, and cisplatin; FFS: failure-free survival; GM-CSF: granulocyte macrophage-colony-stimulating factor; HL: Hodgkin lymphoma; hyper-CVAD: fractionated cyclophosphamide, vincristine, liposomal daunorubicin, and dexamethasone; MOPP: mechlorethamine, vincristine, procarbazine, and prednisone; NA: not available; OFAR: oxaliplatin, fluradabine, ara-C, and rituximab; ORR: overall response rate; OS: overall survival; PFS: progression-free survival; R+hyper-CVAD+GM-CSF/R+HDM-ara-C+GM-CSF: rituximab, fractionated cyclophosphamide, vincristine, liposomal daunorubicin, dexamethasone and GM-CSF alternating with rituximab, methotrexate, ara-C and GM-CSF; R-CHOP: rituximab plus CHOP; R-EPOCH: rituximab, etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin; TRM: treatment-related mortality.

proteins and growth regulatory factors. In a study with eight enrolled RT patients, the ORR was 40%, with no CRs and a PFS of four months [131] (Table 2).

Moreover, successful treatment with bispecific antibodies has been

also reported in RT. Blinatumomab (anti-CD19/anti-CD3) resulted in CR in one R/R RT patient, allowing subsequent allogeneic SCT [132]. Also, treatment with another bispecific antibody, XmAb13676 (anti-CD20/anti-CD3), resulted in CR in one R/R RT patient [133].



Regarding a chimeric antigen receptor-modified T (CAR-T) cell therapy in RT patients, several studies demonstrated rather partial and transient responses [134–138]. Further studies on larger patient cohorts are needed to prove its usefulness in patients ineligible for SCT or as a complementary treatment, in addition to SCT [138].

### 8.3. Stem cell transplantation

RT patients with good clinical performance status can benefit from autologous or allogeneic SCT [3,132]. One study of allogeneic SCT in 17 RT patients demonstrated three years' OS in 75% of patients [29]. Another study of 25 RT patients undergoing allogeneic SCT (the majority with reduced-intensity conditioning), resulted in three years' OS in 36% of patients; better outcomes were seen in patients younger than 60 years and in those with reduced-intensity conditioning [139]. In addition, the crucial impact on survival was an achievement of CRs before SCT in both studies [29,139].

Autologous SCT was studied retrospectively in 34 patients: 3 years' OS was achieved in 56% of patients, of whom 82% were therapy-sensitive at the time of SCT [139].

Currently, long-term survival in RT-DLBCL can be achieved only by allogeneic or autologous SCT. However, most RT-DLBCL patients are ineligible for SCT either due to age/performance status or inadequate response to induction therapy for SCT [16,18]. Therefore, there is an urgent need to investigate novel treatment options in RT.

## 9. Conclusions and future directions

Despite significant improvements in CLL therapy with the introduction of novel agents, CLL patients continue to transform to RT with a similar incidence as previously. Moreover, patients who develop RT on novel agents have particularly poor outcomes, with an aggressive disease course and the transformation occurring early, within the first year of treatment. Currently, there are ongoing clinical trials testing combinations with venetoclax, BCR inhibitors, PD-1 inhibitors and CAR-T cells therapy. All of these modalities have already shown some clinical activity, either in case reports or in small patient cohorts. Although achieved remissions are not likely to be durable, these regimens may be offered as bridge therapies for SCT, which is currently the only treatment option that can achieve long-term survival.

### Practice points

- RT on novel agents used for CLL therapy develops as an early event usually occurring during the first year of treatment with novel drugs.
- The RT incidence on novel agents is similar to chemoimmunotherapy and is lower in patients treated with novel agents in the front line and without *TP53* disruption compared to relapsed/refractory cases.
- The diagnosis of RT is challenging when strictly based on the histologic confirmation of RT in a biopsy sample, as up to one-fifth of diagnostic samples may be misclassified; therefore, revisions by a second expert pathologist are always recommended to avoid misclassifications with conditions that mimic RT, mainly aggressive/accelerated CLL.
- The PET/CT-guided biopsy should always be directed at the most avid lesion with the highest SUVmax.
- The clonality of RT vs underlying CLL should always be determined in order to refer patients with clonally related RT for clinical trials when available as they have extremely poor outcomes with the standard combination chemotherapy/immunochemotherapy used for DLBCL.
- For the genetic analysis of RT, transformed-proven tissue must be used, as genetic aberrations present in the transformed tissue may be absent in the circulating tumour cells; hence, peripheral blood sampling may provide incorrect information.

- Studies and clinical trials reporting on RT should always refer to the RT variant (RT-DLBCL/RT-HL) and to the clonality because the outcome varies for these patients.
- Clinical trials should always report on any toxicity and adverse effects to RT patients.

### Research agenda

- Determining the markers/risk factors that would identify patients who will develop RT in multi-centre, collaborative studies.
- Understanding of the pathogenesis of RT development.
- Developing novel treatment approaches for RT, including either targeted therapy alone or combining targeted agents with chemoimmunotherapy.
- Developing novel non-invasive techniques with sufficient specificity and sensitivity to identify RT patients.
- Searching for novel markers for easy determination of clonality in RT tissues.
- Determining the clinical significance of the clonality in RT-HL and the genetic aberrations associated with this form.
- Employing novel, sensitive genetic methods for more comprehensive monitoring of RT.
- Elucidating the driving mechanism to RT in the context of novel agents, mainly how the selective pressure of treatment by novel agents shapes the genetics of RT.

### Conflict of interest statement

TP: Consultancy for Gilead, Janssen, Abbvie, travel support from Gilead, Janssen, Abbvie, PT: Travel support from Gilead, Janssen, Abbvie.

### Funding

This work was supported by the Ministry of Health, Czech Republic – conceptual development of research organization (FNOL, 0098892), and the Internal Grant Agency of Palacký University Olomouc [IGA\_LF\_2021\_001, IGA\_LF\_2021\_015].

### Acknowledgments

The authors thank Dr Lenka Henzlová from the Olomouc University Hospital for help in preparing PET/CT images, Dr Jana Janková from the Olomouc University Hospital for histopathological examination and help with image preparation, Dr Zuzana Prouzová from the 2nd Faculty of Medicine at Charles University for her advice in histopathology and Anna Lapčíková from the Olomouc University Hospital for providing images of blood smear and bone marrow.

### References

- [1] Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th ed. Lyon: IARC; 2017.
- [2] Richter MN. Generalized reticular cell sarcoma of lymph nodes associated with lymphatic leukemia. *Am J Pathol* 1928;4(4):285–92.
- [3] Ding W. Richter transformation in the era of novel agents. *Hematology Am Soc Hematol Educ Program* 2018;2018:256–63. <https://doi.org/10.1182/asheducation-2018.1.256>.
- [4] Parikh SA, Shanafelt TD. Risk factors for Richter syndrome in chronic lymphocytic leukemia. *Curr Hematol Malig Rep* 2014;9(3):294–9. <https://doi.org/10.1007/s11899-014-0223-4>.
- [5] Bockorny B, Codreanu I, Dasanu CA. Hodgkin lymphoma as Richter transformation in chronic lymphocytic leukaemia: a retrospective analysis of world literature. *Br J Haematol* 2012;156(1):50–66. <https://doi.org/10.1111/j.1365-2141.2011.08907.x>.
- [6] Trump DL, Mann RB, Phelps R, Roberts H, Conley CL. Richter's syndrome: diffuse histiocytic lymphoma in patients with chronic lymphocytic leukemia. A report of five cases and review of the literature. *Am J Med* 1980;68(4):539–48. [https://doi.org/10.1016/0002-9343\(80\)90300-9](https://doi.org/10.1016/0002-9343(80)90300-9).

- [7] Fraser CR, Wang W, Gomez M, Zhang T, Mathew S, Furman RR, et al. Transformation of chronic lymphocytic leukemia/small lymphocytic lymphoma to interdigitating dendritic cell sarcoma: evidence for transdifferentiation of the lymphoma clone. *Am J Clin Pathol* 2009;132(6):928–39. <https://doi.org/10.1309/AJCPWQ010DGBXBMHO>.
- [8] Shao H, Xi L, Raffeld M, Feldman AL, Ketterling RP, Knudson R, et al. Clonally related histiocytic/dendritic cell sarcoma and chronic lymphocytic leukemia/small lymphocytic lymphoma: a study of seven cases. *Mod Pathol* 2011;24(11):1421–32. <https://doi.org/10.1038/modpathol.2011.102>.
- [9] Chakhachiro Z, Yin CC, Abruzzo LV, Aladily TN, Barron LL, Banks HE, et al. B-lymphoblastic leukemia in patients with chronic lymphocytic leukemia: a report of four cases. *Am J Clin Pathol* 2015;144(2):333–40. <https://doi.org/10.1309/AJCPXESVMONMVLZ0>.
- [10] Rossi D, Spina V, Deambrogi C, Rasi S, Laurenti L, Stamatopoulos K, et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood* 2011;117(12):3391–401. <https://doi.org/10.1182/blood-2010-09-302174>.
- [11] Rossi D, Gaidano G. Richter syndrome: pathogenesis and management. *Semin Oncol* 2016;43(2):311–9. <https://doi.org/10.1053/j.seminoncol.2016.02.012>.
- [12] Chigrinova E, Rinaldi A, Kwee I, Rossi D, Rancoita PMV, Strefford JC, et al. Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome. *Blood* 2013;122(15):2673–82. <https://doi.org/10.1182/blood-2013-03-489518>.
- [13] Xiao W, Chen WW, Sorbara L, Davies-Hill T, Pittaluga S, Raffeld M, et al. Hodgkin lymphoma variant of Richter transformation: morphology, Epstein-Barr virus status, clonality, and survival analysis—with comparison to Hodgkin-like lesion. *Hum Pathol* 2016;55:108–16. <https://doi.org/10.1016/j.humpath.2016.04.019>.
- [14] Eyre TA, Schuh A. An update for Richter syndrome - new directions and developments. *Br J Haematol* 2017;178(4):508–20. <https://doi.org/10.1111/bjh.14700>.
- [15] Fabbri G, Khabanian H, Holmes AB, Wang J, Messina M, Mullighan CG, et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J Exp Med* 2013;210(11):2273–88. <https://doi.org/10.1084/jem.20131448>.
- [16] Rossi D, Spina V, Gaidano G. Biology and treatment of Richter syndrome. *Blood* 2018;131(25):2761–72. <https://doi.org/10.1182/blood-2018-01-791376>.
- [17] Parikh SA, Habermann TM, Chaffee KG, Call TG, Ding W, Leis JF, et al. Hodgkin transformation of chronic lymphocytic leukemia: incidence, outcomes, and comparison to de novo Hodgkin lymphoma. *Am J Hematol* 2015;90(4):334–8. <https://doi.org/10.1002/ajh.23939>.
- [18] Khan M, Siddiqi R, Thompson PA. Approach to Richter transformation of chronic lymphocytic leukemia in the era of novel therapies. *Ann Hematol* 2018;97(1):1–15. <https://doi.org/10.1007/s00277-017-3149-9>.
- [19] Wang Y, Ding W. Richter transformation of chronic lymphocytic leukemia in the era of novel agents. *Clin Adv Hematol Oncol* 2020;18(6):348–57.
- [20] Behdad A, Griffin B, Chen YH, Ma S, Kelemen K, Lu X, et al. PD-1 is highly expressed by neoplastic B-cells in Richter transformation. *Br J Haematol* 2019;185(2):370–3. <https://doi.org/10.1111/bjh.15514>.
- [21] He R, Ding W, Viswanatha DS, Chen D, Shi M, Van Dyke D, et al. PD-1 expression in Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma (CLL/SL) and large B-cell Richter Transformation (DLBCL-RT): A characteristic feature of DLBCL-RT and potential surrogate marker for clonal relatedness. *Am J Surg Pathol* 2018;42(7):843–54. <https://doi.org/10.1097/PAS.0000000000001077>.
- [22] Bruzzi JF, Macapinlac H, Tsimberidou AM, Truong MT, Keating MJ, Marom EM, et al. Detection of Richter's transformation of chronic lymphocytic leukemia by PET/CT. *J Nucl Med* 2006;47(8):1267–73.
- [23] Mauro FR, Chauvie S, Paoloni F, Biggi A, Cimino G, Rago A, et al. Diagnostic and prognostic role of PET/CT in patients with chronic lymphocytic leukemia and progressive disease. *Leukemia* 2015;29(6):1360–5. <https://doi.org/10.1038/leu.2015.21>.
- [24] Falchi L, Keating MJ, Marom EM, Truong MT, Schlette EJ, Sargent RL, et al. Correlation between FDG/PET, histology, characteristics, and survival in 332 patients with chronic lymphoid leukemia. *Blood* 2014;123(18):2783–90. <https://doi.org/10.1182/blood-2013-11-536169>.
- [25] Michallet AS, Sesques P, Rabe KG, Itti E, Tordot J, Tychyj-Pinel C, et al. An 18F-FDG-PET maximum standardized uptake value > 10 represents a novel valid marker for discerning Richter's Syndrome. *Leuk Lymphoma* 2016;57(6):1474–7. <https://doi.org/10.3109/10428194.2015.1099643>.
- [26] Mato AR, Wierda WG, Davids MS, Cheson BD, Coutre S, Choi M, et al. Analysis of PET-CT to identify Richter's transformation in 167 patients with disease progression following kinase inhibitor therapy. *Blood* 2017;130(S1):834. <https://doi.org/10.1182/blood.V130.Suppl.1.834.834>.
- [27] Robertson LE, Pugh W, O'Brien S, Kantarjian H, Hirsch-Ginsberg C, Cork A, et al. Richter's syndrome: a report on 39 patients. *J Clin Oncol* 1993;11(10):1985–9. <https://doi.org/10.1200/JCO.1993.11.10.1985>.
- [28] Harousseau JL, Flandrin G, Tricot G, Brouet JC, Seligmann M, Bernard J. Malignant lymphoma supervening in chronic lymphocytic leukemia and related disorders. Richter's syndrome: A study of 25 cases. *Cancer* 1981;48(6):1302–8. [https://doi.org/10.1002/1097-0142\(19810915\)48:6<1302::aid-cncr2820480609>3.0.co;2-q](https://doi.org/10.1002/1097-0142(19810915)48:6<1302::aid-cncr2820480609>3.0.co;2-q).
- [29] Tsimberidou AM, O'Brien S, Khouri I, Giles FJ, Kantarjian HM, Champlin R, et al. Clinical outcomes and prognostic factors in patients with Richter's syndrome treated with chemotherapy or chemoimmunotherapy with or without stem-cell transplantation. *J Clin Oncol* 2006;24(15):2343–51. <https://doi.org/10.1200/JCO.2005.05.0187>.
- [30] Beaudreuil J, Lortholary O, Martin A, Feuillard J, Guillemin L, Lortholary P, et al. Hypercalcemia may indicate Richter's syndrome: report of four cases and review. *Cancer* 1997;79(6):1211–5. [https://doi.org/10.1002/\(sici\)1097-0142\(19970315\)79:6<1211::aid-cncr21>3.0.co;2-1](https://doi.org/10.1002/(sici)1097-0142(19970315)79:6<1211::aid-cncr21>3.0.co;2-1).
- [31] Watanabe N, Yasuda H, Morishita S, Aota Y, Tomomatsu J, Tanaka M, et al. Richter's syndrome with hypercalcemia induced by tumor-associated production of parathyroid hormone-related peptide. *Case Rep Oncol* 2017;10(1):123–6. <https://doi.org/10.1159/000455913>.
- [32] Papajik T, Mysliveček M, Urbanová R, Buriánková E, Kapitáňová, Procházka V, et al. 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography examination in patients with chronic lymphocytic leukemia may reveal Richter transformation. *Leuk Lymphoma* 2014;55(2):314–9. <https://doi.org/10.3109/10428194.2013.802313>.
- [33] Wang Y, Rabe KG, Bold MS, Shi M, Hanson CA, Schwager SM, et al. The role of 18F-FDG-PET in detecting Richter's transformation of chronic lymphocytic leukemia in patients receiving therapy with a B-cell receptor inhibitor. *Haematologica* 2020. <https://doi.org/10.3324/haematol.2019.240564>. <https://doi.org/10.3324/haematol.2019.240564>. Online ahead of print.
- [34] Roberts AW, Ma S, Kipps TJ, Coutre SE, Davids MS, Eichhorst B, et al. Efficacy of venetoclax in relapsed chronic lymphocytic leukemia is influenced by disease and response variables. *Blood* 2019;134(2):111–22. <https://doi.org/10.1182/blood-2018882555>.
- [35] Seymour JF, Ma S, Brander DM, Choi MY, Barrientos J, Davids MS, et al. Venetoclax plus rituximab in relapsed or refractory chronic lymphocytic leukaemia: a phase 1b study. *Lancet Oncol* 2017;18(2):230–40. [https://doi.org/10.1016/S1470-2045\(17\)30012-8](https://doi.org/10.1016/S1470-2045(17)30012-8).
- [36] Stilgenbauer S, Eichhorst B, Schetelig J, Hillmen P, Seymour JF, Coutre S, et al. Venetoclax for patients with chronic lymphocytic leukaemia with 17p deletion: results from the full population of a Phase II pivotal trial. *J Clin Oncol* 2018;36(19):1973–80. <https://doi.org/10.1200/JCO.2017.76.6840>.
- [37] Anderson MA, Tam C, Lew TE, Juneja S, Juneja M, Westerman D, et al. Clinicopathological features and outcomes of progression of CLL on the BCL2 inhibitor venetoclax. *Blood* 2017;129(25):3362–70. <https://doi.org/10.1182/blood-2017-01-763003>.
- [38] Gascoyne RD. XIV. The pathology of transformation of indolent B cell lymphomas: transformation low-grade lymphoma. *Hematol Oncol* 2015;33(S1):75–9. <https://doi.org/10.1002/hon.2222>.
- [39] Rana C, Sharma S, Agarwal M. Prolymphocytic and Richter's transformation in peripheral blood: A case report and review of literature. *J Hematol* 2014;3(3):86–8. <https://doi.org/10.14740/jh121w>.
- [40] Hamilton RM, Leach M. Richter transformation in peripheral blood. *Blood* 2010;115(22):4330. <https://doi.org/10.1182/blood-2009-06-192153>.
- [41] Reinert CP, Federmann B, Hofmann J, Bösmüller H, Wirths S, Fritz J, et al. Computed tomography textural analysis for the differentiation of chronic lymphocytic leukemia and diffuse large B cell lymphoma of richter syndrome. *Eur Radiol* 2019;29(12):6911–21. <https://doi.org/10.1007/s00330-019-06291-9>.
- [42] Federmann B, Mueller MR, Steinhilber J, Horger MS, Fend F. Diagnosis of Richter transformation in chronic lymphocytic leukemia: histology tips the scales. *Ann Hematol* 2018;97(10):1859–68. <https://doi.org/10.1007/s00277-018-3390-x>.
- [43] Soilleux EJ, Wotherspoon A, Eyre TA, Clifford R, Cabes M, Schuh AH. Diagnostic dilemmas of high-grade transformation (Richter's syndrome) of chronic lymphocytic leukaemia: results of the phase II National Cancer Research Institute CHOP-OR clinical trial specialist haematology-pathology central review. *Histopathology* 2016;69(6):1066–76. <https://doi.org/10.1111/his.13024>.
- [44] Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. World Health Organization classification of tumours of haematopoietic and lymphoid tissues. *Lyon: IARC Press*; 2008.
- [45] Medeiros LJ, Miranda RN, Wang SA, BuesoRamos C. Diagnostic pathology: lymph nodes and spleen with extranodal lymphomas. *Altona, MB: Amirsys Publishing Inc*; 2011.
- [46] Mao Z, Quintanilla-Martinez L, Raffeld M, Richter M, Krugmann J, Burek C, et al. IgVH mutational status and clonality analysis of Richter's transformation: diffuse large B-cell lymphoma and hodgkin lymphoma in association with B-cell Chronic Lymphocytic Leukemia (B-CLL) represent 2 different pathways of disease evolution. *Am J Surg Pathol* 2007;31(10):1605–14. <https://doi.org/10.1097/PAS.0b013e31804bdaf8>.
- [47] Agbay RLMC, Jain N, Loghavi S, Medeiros LJ, Khoury JD. Histologic transformation of chronic lymphocytic leukemia/small lymphocytic lymphoma. *Am J Hematol* 2016;91(10):1036–43. <https://doi.org/10.1002/ajh.24473>.
- [48] Ciccone M, Agostinelli C, Rigolin GM, Piccaluga PP, Cavazzini F, Righi S, et al. Proliferation centers in chronic lymphocytic leukemia: correlation with cytogenetic and clinicobiological features in consecutive patients analyzed on tissue microarrays. *Leukemia* 2012;26(3):499–508. <https://doi.org/10.1038/leu.2011.247>.
- [49] Quintanilla-Martinez L. The 2016 updated WHO classification of lymphoid neoplasias. *Hematol Oncol* 2017;35(S1):37–45. <https://doi.org/10.1002/hon.2399>.
- [50] de Leval L, Vivario M, De Prieck B, Zhou Y, Boniver J, Harris NL, et al. Distinct clonal origin in two cases of Hodgkin's lymphoma variant of Richter's syndrome associated with EBV infection. *Am J Surg Pathol* 2004;28(5):679–86. <https://doi.org/10.1097/00000478-2004050000018>.
- [51] Parikh SA, Rabe KG, Call TG, Zent CS, Habermann TM, Ding W, et al. Diffuse large B-cell lymphoma (Richter syndrome) in patients with chronic lymphocytic leukaemia (CLL): a cohort study of newly diagnosed patients. *Br J Haematol* 2013;162(6):774–82. <https://doi.org/10.1111/bjh.12458>.

- [52] Giné E, Martínez A, Villamor N, López-Guillermo A, Camos M, Martínez D, et al. Expanded and highly active proliferation centers identify a histological subtype of chronic lymphocytic leukemia ("accelerated" chronic lymphocytic leukemia) with aggressive clinical behavior. *Haematologica* 2010;95(9):1526–33. <https://doi.org/10.3324/haematol.2010.022277>.
- [53] Cordone I, Matutes E, Catovsky D. Monoclonal antibody Ki-67 identifies B and T cells in cycle in chronic lymphocytic leukemia: correlation with disease activity. *Leukemia* 1992;6(9):902–6.
- [54] Hleuhel MH, Ben-Dali Y, Da Cunha-Bang C, Brieghel C, Clasen-Linde E, Niemann CU, et al. Risk factors associated with Richter's transformation in patients with chronic lymphocytic leukaemia: protocol for a retrospective population-based cohort study. *BMJ Open* 2019;9(3):e023566.
- [55] Rossi D, Cerri M, Capello D, Deambrogi C, Rossi FM, Zucchetto A, et al. Biological and clinical risk factors of chronic lymphocytic leukaemia transformation to Richter syndrome. *Br J Haematol* 2008;142(2):202–15. <https://doi.org/10.1111/j.1365-2141.2008.07166.x>.
- [56] Ben-Dali Y, Hleuhel MH, da Cunha-Bang C, Brieghel C, Poulsen CB, Clasen-Linde E, et al. Richter's transformation in patients with chronic lymphocytic leukaemia: a nationwide epidemiological study. *Leuk Lymphoma* 2020;61(6):1435–44. <https://doi.org/10.1080/10428194.2020.1719092>.
- [57] Rossi D, Rasi S, Spina V, Fangazio M, Monti S, Greco M, et al. Different impact of NOTCH1 and SF3B1 mutations on the risk of chronic lymphocytic leukemia transformation to Richter syndrome. *Br J Haematol* 2012;158:426–9. <https://doi.org/10.1111/j.1365-2141.2012.09155.x>.
- [58] Balatti V, Tomasello L, Rassenti LZ, Veneziano D, Nigita G, Wang HY, et al. MiR-125a and MiR-34a expression predicts Richter syndrome in chronic lymphocytic leukemia patients. *Blood* 2018;132:2179–82. <https://doi.org/10.1182/blood-2018-04-845115>.
- [59] Van Roosbroeck K, Bayraktar R, Calin S, Bloehdorn J, Dragomir MP, Okubo K, et al. The involvement of microRNA in the pathogenesis of Richter syndrome. *Haematologica* 2019;104:1004–15. <https://doi.org/10.3324/haematol.2018.203828>.
- [60] Allan JN, Furman RR. Current trends in the management of Richter's syndrome. *Int J Hematol Oncol* 2019;7:LJH09. <https://doi.org/10.2217/ijh-2018-0010>.
- [61] Rossi D, Lobetti Bodoni C, Genuardi E, Montillo L, Drandi D, Cerri M, et al. Telomere length is an independent predictor of survival, treatment requirement and Richter's syndrome transformation in chronic lymphocytic leukemia. *Leukemia* 2009;23:1062–72. <https://doi.org/10.1038/leu.2008.399>.
- [62] Wang Y, Tschautscher MA, Rabe KG, Call TG, Leis JF, Kenderian SS, et al. Clinical characteristics and outcomes of Richter transformation: experience of 204 patients from a single center. *Haematologica* 2019;224121. <https://doi.org/10.3324/haematol.2019.224121>.
- [63] Strati P, Abruzzo LV, Wierda WG, O'Brien S, Ferrajoli A, Keating MJ. Second cancers and Richter transformation are the leading causes of death in patients with trisomy 12 chronic lymphocytic leukemia. *Clin Lymphoma Myeloma Leuk* 2015;15:420–7. <https://doi.org/10.1016/j.clml.2015.02.001>.
- [64] Rossi D, Spina V, Cerri M, Rasi S, Deambrogi C, De Paoli L, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to richter syndrome. *Clin Cancer Res* 2009;15:4415–22. <https://doi.org/10.1158/1078-0432.CCR-08-3266>.
- [65] Rasi S, Spina V, Bruscazzin A, Vaisitti T, Tripodo C, Forconi F, et al. A variant of the LRP4 gene affects the risk of chronic lymphocytic leukaemia transformation to Richter syndrome. *Br J Haematol* 2011;152:284–94. <https://doi.org/10.1111/j.1365-2141.2010.08482.x>.
- [66] Fong D, Kaiser A, Spizzo G, Gastl G, Tzankov A. Hodgkin's disease variant of Richter's syndrome in chronic lymphocytic leukaemia patients previously treated with fludarabine. *Br J Haematol* 2005;129(2):199–205. <https://doi.org/10.1111/j.1365-2141.2005.05426.x>.
- [67] Catherwood MA, Gonzalez D, Donaldson D, Clifford R, Mills K, Thornton P. Relevance of TP53 for CLL diagnostics. *J Clin Pathol* 2019;72:343–6. <https://doi.org/10.1136/jclinpath-2018-205622>.
- [68] Zenz T, Eichhorst B, Busch R, Denzel T, Häbe S, Winkler D, et al. TP53 mutation and survival in chronic lymphocytic leukemia. *J Clin Oncol* 2010;28(29):4473–9. <https://doi.org/10.1200/JCO.2009.27.8762>.
- [69] Gonzalez D, Martinez P, Wade R, Hockley S, Oscier D, Matutes E, et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. *J Clin Oncol* 2011;29:2223–9. <https://doi.org/10.1200/JCO.2010.32.0838>.
- [70] Serrano M, Hannon GJ, Beach D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* 1993;366(6456):704–7. <https://doi.org/10.1038/366704a0>.
- [71] Knoepfler PS. Myc goes global: new tricks for an old oncogene. *Cancer Res* 2007;67(11):5061–3. <https://doi.org/10.1158/0008-5472.CAN-07-0426>.
- [72] Dang CV, O'Donnell KA, Zeller KI, Nguyen T, Osthus RC, Li F. The c-Myc target gene network. *Semin Cancer Biol* 2006;16:253–64. <https://doi.org/10.1016/j.semcancer.2006.07.014>.
- [73] Edelmann J, Holzmann K, Miller F, Winkler D, Bühler A, Zenz T, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. *Blood* 2012;120(24):4783–94. <https://doi.org/10.1182/blood-2012-04-423517>.
- [74] De Paoli L, Cerri M, Monti S, Rasi S, Spina V, Bruscazzin A, et al. MGA, a suppressor of MYC, is recurrently inactivated in high risk chronic lymphocytic leukemia. *Leuk Lymphoma* 2013;54(5):1087–90. <https://doi.org/10.3109/10428194.2012.723706>.
- [75] Fabbri G, Dalla-Favera R. The molecular pathogenesis of chronic lymphocytic leukaemia. *Nat Rev Cancer* 2016;16(3):145–62. <https://doi.org/10.1038/nrc.2016.8>.
- [76] Balatti V, Bottoni A, Palamarchuk A, Alder H, Rassenti LZ, Kipps TJ, et al. NOTCH1 mutations in CLL associated with trisomy 12. *Blood* 2012;119(2):329–31. <https://doi.org/10.1182/blood-2011-10-386144>.
- [77] Del Giudice I, Rossi D, Chiaretti S, Marinelli M, Tavolaro S, Gabrielli S, et al. NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. *Haematologica* 2012;97(3):437–41. <https://doi.org/10.3324/haematol.2011.060129>.
- [78] Kopan R, Ilagan MX. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* 2009;137:216–33. <https://doi.org/10.1016/j.cell.2009.03.045>.
- [79] Guruharsha KG, Kankel MW, Artavanis-Tsakonas S. The Notch signalling system: recent insights into the complexity of a conserved pathway. *Nat Rev Genet* 2012;13(9):654–66. <https://doi.org/10.1038/nrg3272>.
- [80] Fabbri G, Holmes AB, Viganotti M, Scuoppo C, Belver L, Herranz D, et al. Common nonmutational NOTCH1 activation in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 2017;114(14):E2911–9. <https://doi.org/10.1073/pnas.1702564114>.
- [81] Di Ianni M, Baldoni S, Rosati E, Ciunnelli R, Cavalli L, Martelli MF, et al. A new genetic lesion in B-CLL: a NOTCH1 PEST domain mutation. *Br J Haematol* 2009;146:689–91. <https://doi.org/10.1111/j.1365-2141.2009.07816.x>.
- [82] Kadri S, Lee J, Fitzpatrick C, Galanina N, Sukhanova M, Venkataraman G, et al. Clonal evolution underlying leukemia progression and Richter transformation in patients with ibrutinib-relapsed CLL. *Blood Adv* 2017;1:715–27. <https://doi.org/10.1182/bloodadvances.2016003632>.
- [83] Maddocks KJ, Ruppert AS, Lozanski G, Heerema NA, Zhao W, Abruzzo L, et al. Etiology of ibrutinib therapy discontinuation and outcomes in patients with chronic lymphocytic leukemia. *JAMA Oncol* 2015;1(1):80–7. <https://doi.org/10.1001/jamaoncol.2014.218>.
- [84] Ahn IE, Faraoui MZH, Tian X, Valdez J, Sun C, Soto S, et al. Depth and durability of response to ibrutinib in CLL: 5-year follow-up of a phase 2 study. *Blood* 2018;131:2357–66. <https://doi.org/10.1182/blood-2017-12-820910>.
- [85] Jain P, Keating M, Wierda W, Estrov Z, Ferrajoli A, Jain N, et al. Outcomes of patients with chronic lymphocytic leukemia after discontinuing ibrutinib. *Blood* 2015;125(13):2062–7. <https://doi.org/10.1182/blood-2014-09-603670>.
- [86] Jain P, Thompson PA, Keating M, Estrov Z, Ferrajoli A, Jain N, et al. Long-term outcomes for patients with chronic lymphocytic leukemia who discontinue ibrutinib. *Cancer* 2017;123:2268–73. <https://doi.org/10.1002/ncr.30596>.
- [87] Ding W, LaPlant BR, Call TG, Parikh SA, Leis JF, He R, et al. Pembrolizumab in patients with CLL and Richter transformation or with relapsed CLL. *Blood* 2017;129(26):3419–27. <https://doi.org/10.1182/blood-2017-02-765685>.
- [88] Kanagal-Shamanna R, Jain P, Patel KP, Roubort M, Bueso-Ramos C, Alhalouli T, et al. Targeted multigene deep sequencing of Bruton tyrosine kinase inhibitor-resistant chronic lymphocytic leukemia with disease progression and Richter transformation. *Cancer* 2019;125(4):559–74. <https://doi.org/10.1002/ncr.31831>.
- [89] Kiss R, Alpár D, Gángó A, Nagy N, Eyupoglu E, Aczél D, et al. Spatial clonal evolution leading to ibrutinib resistance and disease progression in chronic lymphocytic leukemia. *Haematologica* 2019;104(1):e38–41. <https://doi.org/10.3324/haematol.2018.202085>.
- [90] Herling CD, Abedpour N, Weiss J, Schmitt A, Jachimowicz RD, Merkel O, et al. Clonal dynamics towards the development of venetoclax resistance in chronic lymphocytic leukemia. *Nat Commun* 2018;9(1):727. <https://doi.org/10.1038/s41467-018-03170-7>.
- [91] Chakraborty S, Fortunati I, Martinez C, Yadav B, Dimishkovska M, Innocenti I, et al. A CRISPR/Cas9-generated murine model reveals cooperation between BCR signaling and CDKN2A/2B and TP53 disruption in Richter Syndrome. *Blood* 2019;134(S1):4278. <https://doi.org/10.1182/blood-2019-128425>.
- [92] Ahn IE, Underbayev C, Albitar A, Herman SEM, Tian X, Maric I, et al. Clonal evolution leading to ibrutinib resistance in chronic lymphocytic leukemia. *Blood* 2017;129(11):1469–79. <https://doi.org/10.1182/blood-2016-06-719294>.
- [93] Burger JA, Landau DA, Taylor-Weiner A, Bozic I, Zhang H, Sarosiek K, et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun* 2016;7:11589. <https://doi.org/10.1038/ncomms11589>.
- [94] Sedlarikova L, Petrackova A, Papajik T, Turcsanyi P, Kriegova E. Resistance-associated mutations in chronic lymphocytic leukaemia patients treated with novel agents. *Front Oncol* 2020. <https://doi.org/10.3389/fonc.2020.00894>. In press.
- [95] Lampson BL, Brown JR. Are BTK and PLCG2 mutations necessary and sufficient for ibrutinib resistance in chronic lymphocytic leukemia? *Expert Rev Hematol* 2018;11(3):185–94. <https://doi.org/10.1080/17474086.2018.1435268>.
- [96] Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med* 2015;21(8):922–6. <https://doi.org/10.1038/nm.3884>.
- [97] Wang YL. MYD88 mutations and sensitivity to ibrutinib therapy. *J Mol Diagn* 2018;20(2):264–6. <https://doi.org/10.1016/j.jmoldx.2017.11.006>.
- [98] Rosati E, Baldoni S, De Falco F, Del Papa B, Dorillo E, Rompietti C, et al. NOTCH1 aberrations in chronic lymphocytic leukemia. *Front Oncol* 2018;8:229. <https://doi.org/10.3389/fonc.2018.00229>.
- [99] Rossi D, Khiabanian H, Spina V, Ciardullo C, Bruscazzin A, Famà R, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood* 2014;123(14):2139–47. <https://doi.org/10.1182/blood-2013-11-539726>.



- [100] Petrackova A, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, et al. Standardization of sequencing coverage depth in NGS: recommendation for detection of clonal and subclonal mutations in cancer diagnostics. *Front Oncol* 2019;9:851. <https://doi.org/10.3389/fonc.2019.00851>.
- [101] Yeh P, Hunter T, Sinha D, Ftouni S, Wallach E, Jiang D, et al. Circulating tumour DNA reflects treatment response and clonal evolution in chronic lymphocytic leukaemia. *Nat Commun* 2017;8:14756. <https://doi.org/10.1038/ncomms14756>.
- [102] Scherer F, Kurtz DM, Diehn M, Alizadeh AA. High-throughput sequencing for noninvasive disease detection in hematologic malignancies. *Blood* 2017;130(4):440–52. <https://doi.org/10.1182/blood-2017-03-735639>.
- [103] Rossi D, Diop F, Spaccarotella E, Monti S, Zanni M, Rasi S, et al. Diffuse large B-cell lymphoma genotyping on the liquid biopsy. *Blood* 2017;129(14):1947–57. <https://doi.org/10.1182/blood-2016-05-719641>.
- [104] Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. *Genetics* 2000;156(1):297–304.
- [105] de Melo Gagliato D, Fontes Jardim DL. Noninvasive cancer biomarkers in solid malignancies: circulating tumor DNA-clinical utility, current limitations and future perspectives. *Ann Transl Med* 2018;6(11):233. <https://doi.org/10.21037/atm.2018.05.22>.
- [106] Woyach JA, Ruppert AS, Guinn D, Lehman A, Blachly JS, Lozanski A, et al. BTK (C481S)-mediated resistance to ibrutinib in chronic lymphocytic leukemia. *J Clin Oncol* 2017;35(13):1437–43. <https://doi.org/10.1200/JCO.2016.70.2282>.
- [107] Fan L, Wang L, Zhang R, Fang C, Zhu DX, Wang YH, et al. Richter transformation in 16 of 149 Chinese patients with chronic lymphocytic leukemia. *Leuk Lymphoma* 2012;53(9):1749–56. <https://doi.org/10.3109/10428194.2012.664845>.
- [108] Catovsky D, Richards S, Matutes E, Oscier D, Mjts Dyer, Bezares RF, et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. *Lancet* 2007;370(9583):230–9. [https://doi.org/10.1016/S0140-6736\(07\)61125-8](https://doi.org/10.1016/S0140-6736(07)61125-8).
- [109] Eyre TA, Clifford R, Bloor A, Boyle L, Roberts C, Cabes M, et al. NCI phase II study of CHOP in combination with ofatumumab in induction and maintenance in newly diagnosed Richter syndrome. *Br J Haematol* 2016;175(1):43–54. <https://doi.org/10.1111/bjh.14177>.
- [110] Davids MS, Huang Y, Rogers KA, Stern R, Brown JR, Thompson PA, et al. Richter's syndrome (RS) in patients with chronic lymphocytic leukemia (CLL) on novel agent therapy. *J Clin Oncol* 2017;35(S15):7505. [https://doi.org/10.1200/JCO.2017.35.15\\_suppl.7505](https://doi.org/10.1200/JCO.2017.35.15_suppl.7505).
- [111] Dabaja BS, O'Brien SM, Kantarjian HM, Cortes JE, Thomas DA, Albitar M, et al. Fractionated cyclophosphamide, vincristine, liposomal daunorubicin (Daunoxome), and dexamethasone (HyperCVXD) regimen in Richter's syndrome. *Leuk Lymphoma* 2001;42(3):329–37. <https://doi.org/10.3109/10428190109064589>.
- [112] Durot E, Michallet AS, Leprêtre S, Le QH, Leblond V, Delmer A. Platinum and high-dose cytarabine-based regimens are efficient in ultra high/high-risk chronic lymphocytic leukemia and Richter's syndrome: results of a French retrospective multicenter study. *Eur J Haematol* 2015;95(2):160–7. <https://doi.org/10.1111/ejh.12474>.
- [113] Tsimberidou AM, Kantarjian HM, Cortes J, Thomas DA, Faderl S, Garcia-Manero G, et al. Fractionated cyclophosphamide, vincristine, liposomal daunorubicin, and dexamethasone plus rituximab and granulocyte-macrophage-colony stimulating factor (GM-CSF) alternating with methotrexate and cytarabine plus rituximab and GM-CSF in patients with Richter syndrome or fludarabine-refractory chronic lymphocytic leukemia. *Cancer* 2003;97(7):1711–20. <https://doi.org/10.1002/cncr.11238>.
- [114] Tsimberidou AM, Wierda WG, Plunkett W, Kurzrock R, O'Brien S, Wen S, et al. Phase I-II study of oxaliplatin, fludarabine, cytarabine, and rituximab combination therapy in patients with Richter's Syndrome or Fludarabine-Refractory chronic lymphocytic leukemia. *J Clin Oncol* 2008;26(2):196–203. <https://doi.org/10.1200/JCO.2007.11.8513>.
- [115] Tsimberidou AM, Wierda WG, Wen S, Plunkett W, O'Brien S, Kipps TJ, et al. Phase I-II clinical trial of oxaliplatin, fludarabine, cytarabine, and rituximab therapy in aggressive relapsed/refractory chronic lymphocytic leukemia or Richter syndrome. *Clin Lymphoma Myeloma Leuk* 2013;13(5):568–74. <https://doi.org/10.1016/j.clml.2013.03.012>.
- [116] Rogers KA, Huang Y, Ruppert AS, Salem G, Stephens DM, Heerema NA, et al. A single-institution retrospective cohort study of first-line R-EPOCH chemioimmunotherapy for Richter syndrome demonstrating complex chronic lymphocytic leukaemia karyotype as an adverse prognostic factor. *Br J Haematol* 2018;180(2):259–66. <https://doi.org/10.1111/bjh.15035>.
- [117] Langerbeins P, Busch R, Anheier N, Dürig J, Bergmann M, Goebeler ME, et al. Poor efficacy and tolerability of R-CHOP in relapsed/refractory chronic lymphocytic leukemia and Richter transformation. *Am J Hematol* 2014;89(12):E239–43. <https://doi.org/10.1002/ajh.23841>.
- [118] Tsimberidou AM, O'Brien S, Kantarjian HM, Koller C, Hagemester FB, Fayad L, et al. Hodgkin transformation of chronic lymphocytic leukemia: the M. D. Anderson Cancer Center experience. *Cancer* 2006;107(6):1294–302. <https://doi.org/10.1002/cncr.22121>.
- [119] Al-Sawaf O, Robrecht S, Bahlo J, Fink AM, Cramer P, Tresckow JV, et al. Richter transformation in chronic lymphocytic leukemia (CLL) - a pooled analysis of German CLL Study Group (GCLLSG) front line treatment trials. *Leukemia* 2020; Mar 17. <https://doi.org/10.1038/s41375-020-0797-x>. Online ahead of print.
- [120] Abrisqueta P, Delgado J, Alcoceba M, Oliveira AC, Loscertales J, Hernández-Rivas JA, et al. Clinical outcome and prognostic factors of patients with Richter syndrome: real-world study of the Spanish Chronic Lymphocytic Leukemia Study Group (GELLC). *Br J Haematol* 2020;Jun 9. <https://doi.org/10.1111/bjh.16748>. Online ahead of print.
- [121] Tadmor T, Shvidel L, Goldschmidt N, Ruchlemer R, Fineman R, Bairey O, et al. Hodgkin's variant of Richter transformation in chronic lymphocytic leukemia: a retrospective study from the Israeli CLL study group. *Anticancer Res* 2014;34(2):785–90.
- [122] Davids MS, Rogers KA, Tyekuceva S, Pazienza S, Renner SK, Montegaard, et al. A multicenter phase II study of venetoclax plus dose-adjusted R-EPOCH (VR-EPOCH) for Richter's syndrome. *J Clin Oncol* 2020;38(15S):8004. [https://doi.org/10.1200/JCO.2020.38.15\\_suppl.8004](https://doi.org/10.1200/JCO.2020.38.15_suppl.8004).
- [123] Tsang M, Shanafelt TD, Call TG, Ding W, Chanan-Khan A, Leis JF, et al. The efficacy of ibrutinib in the treatment of Richter syndrome. *Blood* 2015;125(10):1676–8. <https://doi.org/10.1182/blood-2014-12-610782>.
- [124] Visentin A, Imbergamo S, Scomazzon E, Pravato S, Frezzato F, Bonaldi L, et al. BCR kinase inhibitors, idelalisib and ibrutinib, are active and effective in Richter syndrome. *Br J Haematol* 2019;185(1):193–7. <https://doi.org/10.1111/bjh.15440>.
- [125] Hillmen P, Schuh A, Eyre TA, Pagel JM, Brown JR, Ghia P, et al. Acalabrutinib Monotherapy in Patients with Richter Transformation from the Phase 1/2 ACE-CL-001 Clinical Study. *Blood* 2016;128(22):60. <https://doi.org/10.1182/blood.V128.22.60.60>.
- [126] Armand P, Murawski N, Molin D, Zain J, Eichhorst B, Gulbas Z, et al. Pembrolizumab in relapsed or refractory Richter syndrome. *Br J Haematol* 2020;190(2):e117–20. <https://doi.org/10.1111/bjh.16762>.
- [127] Mato AR, Svoboda J, Luning Prak ET, Schuster SJ, Tsao PY, Dorsey C, et al. Phase I/II study of umbralisib (TGR-1202) in combination with ublituximab (TG-1101) and pembrolizumab in patients with REL/REF CLL and Richter's transformation. *Hematol Oncol* 2019;37(S2):119–20.
- [128] Younes A, Brody J, Carpio C, Lopez-Guillermo A, Ben-Yehuda D, Ferhanoglu B, et al. Safety and activity of ibrutinib in combination with nivolumab in patients with relapsed non-Hodgkin lymphoma or chronic lymphocytic leukaemia: a phase 1/2a study. *Lancet Haematol* 2019;6(2):e67–78. [https://doi.org/10.1016/S2352-3026\(18\)30217-5](https://doi.org/10.1016/S2352-3026(18)30217-5).
- [129] Jain N, Basu S, Thompson PA, Ohanian M, Ferrajoli A, Pemmaraju N, et al. Nivolumab combined with ibrutinib for CLL and Richter transformation: a phase II trial. *Blood* 2016;128(22):59. <https://doi.org/10.1182/blood.V128.22.59.59>.
- [130] Montillo M, Rossi D, Zucca E, Frustaci AM, Pileri S, Cavalli F, et al. A multicenter, open label, uncontrolled, phase II trial evaluating safety and efficacy of venetoclax, atezolizumab and obinutuzumab in Richter transformation from CLL. *Hematol Oncol* 2019;37:557–8. <https://doi.org/10.1002/hon.42632>.
- [131] Kuruvilla J, Savona M, Baz R, Mau-Sorensen PM, Gabrail N, Garzon R, et al. Selective inhibition of nuclear export with selinexor in patients with non-Hodgkin lymphoma. *Blood* 2017;129(24):3175–83. <https://doi.org/10.1182/blood-2016-11-750174>.
- [132] Alderuccio JP, Mackrides N, Chapman JR, Vega F, Lossos IS. Rapid complete response to blinatumomab as a successful bridge to allogeneic stem cell transplantation in a case of refractory Richter syndrome. *Leuk Lymphoma* 2019;60(1):230–3. <https://doi.org/10.1080/10428194.2018.1461862>.
- [133] Patel K, Michot JM, Chanan-Khan AA, Salles GA, Cartron G, Peyrade F, et al. Preliminary safety and anti-tumor activity of XmAb13676, an anti-CD20 x anti-CD3 bispecific antibody, in patients with relapsed/refractory non-Hodgkin's lymphoma and chronic lymphocytic leukemia. *Blood* 2019;134(S1):4079. <https://doi.org/10.1182/blood-2019-128564>.
- [134] Cruz CRY, Micklethwaite KP, Savoldo B, Ramos CA, Lam S, Ku S, et al. Infusion of donor-derived CD19-redireted virus-specific T cells for B-cell malignancies relapsed after allogeneic stem cell transplant: a phase 1 study. *Blood* 2013;122(17):2965–73. <https://doi.org/10.1182/blood-2013-06-506741>.
- [135] Kochenderfer JN, Dudley ME, Kassim SH, Somerville RPT, Carpenter RO, Stetler-Stevenson M, et al. Chemotherapy-refractory diffuse large B-cell lymphoma and indolent B-cell malignancies can be effectively treated with autologous T cells expressing an Anti-CD19 chimeric antigen receptor. *J Clin Oncol* 2015;33:540–9. <https://doi.org/10.1200/JCO.2014.56.2025>.
- [136] Turtle CJ, Hay KA, Hanafi LA, Li D, Cheria S, Chen X, et al. Durable molecular remissions in chronic lymphocytic leukemia treated with CD19-specific chimeric antigen receptor-modified T cells after failure of ibrutinib. *J Clin Oncol* 2017;35(26):3010–20. <https://doi.org/10.1200/JCO.2017.72.8519>.
- [137] Gauthier J, Hirayama AV, Hay KA, Li D, Lypm J, Sheih A, et al. Comparison of efficacy and toxicity of CD19-specific chimeric antigen receptor T-Cells alone or in combination with ibrutinib for relapsed and/or refractory CLL. *Blood* 2018;132:299. <https://doi.org/10.1182/blood-2018-99-111061>.
- [138] Lemal R, Tournilhac O. State-of-the-art for CAR T-cell therapy for chronic lymphocytic leukemia in 2019. *J Immunother Cancer* 2019;7(1):202. <https://doi.org/10.1186/s40425-019-0686-x>.
- [139] Cwynarski K, van Biezen A, de Wreede L, Stilgenbauer S, Bunjes D, Metzner B, et al. Autologous and allogeneic stem-cell transplantation for transformed chronic lymphocytic leukemia (Richter's Syndrome): A retrospective analysis from the chronic lymphocytic leukemia subcommittee of the chronic leukemia working party and lymphoma working party of the European group for blood and marrow transplantation. *J Clin Oncol* 2012;30(18):2211–7. <https://doi.org/10.1200/JCO.2011.37.4108>.
- [140] O'Brien S, Furman RR, Coutre SE, Sharman JP, Burger JA, Blum KA, et al. Ibrutinib as initial therapy for elderly patients with chronic lymphocytic leukaemia or small lymphocytic lymphoma: an open-label, multicentre, phase 1b/2 trial. *Lancet Oncol* 2014;15(1):48–58. [https://doi.org/10.1016/S1470-2045\(13\)70513-8](https://doi.org/10.1016/S1470-2045(13)70513-8).

- [141] Farooqui MZ, Valdez J, Martyr S, Aue G, Saba N, Niemann CU, et al. Ibrutinib for previously untreated and relapsed or refractory chronic lymphocytic leukaemia with TP53 aberrations: a phase 2, single-arm trial. *Lancet Oncol* 2015;16(2):169–76. [https://doi.org/10.1016/S1470-2045\(14\)71182-9](https://doi.org/10.1016/S1470-2045(14)71182-9).
- [142] Woyach JA, Ruppert AS, Heerema NA, Zhao W, Booth AM, Ding W, et al. Ibrutinib regimens versus chemoimmunotherapy in older patients with untreated CLL. *N Engl J Med* 2018;379(26):2517–28. <https://doi.org/10.1056/NEJMoa1812836>.
- [143] Mato AR, Nabhan C, Thompson MC, Lamanna N, Brander DM, Hill B, et al. Toxicities and outcomes of 616 ibrutinib-treated patients in the United States: a real-world analysis. *Haematologica* 2018;103:874–9. <https://doi.org/10.3324/haematol.2017.182907>.
- [144] O'Brien SM, Byrd JC, Hillmen P, Coutre S, Brown JR, Barr PM, et al. Outcomes with ibrutinib by line of therapy and post-ibrutinib discontinuation in patients with chronic lymphocytic leukemia: Phase 3 analysis. *Am J Hematol* 2019;94(5):554–62. <https://doi.org/10.1002/ajh.25436>.
- [145] Dimou M, Iliakis T, Pardalis V, Bitsani C, Vassilakopoulos TP, Angelopoulou M, et al. Safety and efficacy analysis of long-term follow up real-world data with ibrutinib monotherapy in 58 patients with CLL treated in a single-center in Greece. *Leuk Lymphoma* 2019;60(12):2939–45. <https://doi.org/10.1080/10428194.2019.1620944>.
- [146] Moreno C, Greil R, Demirkan F, Tedeschi A, Anz B, Larratt L, et al. Ibrutinib plus obinutuzumab versus chlorambucil plus obinutuzumab in first-line treatment of chronic lymphocytic leukaemia (ILLUMINATE): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol* 2019;20(1):43–56. [https://doi.org/10.1016/S1470-2045\(18\)30788-5](https://doi.org/10.1016/S1470-2045(18)30788-5).
- [147] Burger JA, Sivina M, Jain N, Kim E, Kadia T, Estrov Z, et al. Randomized trial of ibrutinib vs ibrutinib plus rituximab in patients with chronic lymphocytic leukemia. *Blood* 2019;133(10):1011–9. <https://doi.org/10.1182/blood-2018-10-879429>.
- [148] Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, Blum KA, et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med* 2013;369(1):32–42. <https://doi.org/10.1056/NEJMoa1215637>.
- [149] UK CLL Forum. Ibrutinib for relapsed/refractory chronic lymphocytic leukemia: a UK and Ireland analysis of outcomes in 315 patients. *Haematologica* 2016;101(12):1563–72. <https://doi.org/10.3324/haematol.2016.147900>.
- [150] O'Brien S, Jones JA, Coutre SE, Mato AR, Hillmen P, Tam C, et al. Ibrutinib for patients with relapsed or refractory chronic lymphocytic leukaemia with 17p deletion (RESONATE-17): a phase 2, open-label, multicentre study. *Lancet Oncol* 2016;17(10):1409–18. [https://doi.org/10.1016/S1470-2045\(16\)30212-1](https://doi.org/10.1016/S1470-2045(16)30212-1).
- [151] Byrd JC, Wierda WG, Schuh A, Devereux S, Chaves JM, Brown JR, et al. Acalabrutinib monotherapy in patients with relapsed/refractory chronic lymphocytic leukemia: updated results from the phase 1/2 ACE-CL-001 study. *Blood*. 2017;130(S1):498. [https://doi.org/10.1182/blood.V130.Suppl\\_1.498.498](https://doi.org/10.1182/blood.V130.Suppl_1.498.498).
- [152] Huang X, Qiu L, Jin J, Zhou D, Chen X, Hou M, et al. Ibrutinib versus rituximab in relapsed or refractory chronic lymphocytic leukemia or small lymphocytic lymphoma: a randomized, open-label phase 3 study. *Cancer Med* 2018;7(4):1043–55. <https://doi.org/10.1002/cam4.1337>.
- [153] Nuttall E, Tung J, Trounce E, Johnston R, Chevassat T. Real-world experience of ibrutinib therapy in relapsed chronic lymphocytic leukemia: results of a single-center retrospective analysis. *J Blood Med* 2019;10:199–208. <https://doi.org/10.2147/JBM.S202286>.
- [154] Byrd JC, Hillmen P, O'Brien S, Barrientos JC, Reddy NM, Coutre S, et al. Long-term follow-up of the RESONATE phase 3 trial of ibrutinib vs ofatumumab. *Blood* 2019;133(19):2031–42. <https://doi.org/10.1182/blood-2018-08-870238>.
- [155] Wingqvist M, Andersson PO, Askild A, Karlsson K, Karlsson C, Lauri B, et al. Long-term real-world results of ibrutinib therapy in patients with relapsed or refractory chronic lymphocytic leukemia: 30-month follow up of the Swedish compassionate use cohort. *Haematologica* 2019;104(5):e208–10. <https://doi.org/10.3324/haematol.2018.198820>.
- [156] Awan FT, Schuh A, Brown JR, Furman RR, Pagel JM, Hillmen P, et al. Acalabrutinib monotherapy in patients with chronic lymphocytic leukemia who are intolerant to ibrutinib. *Blood Adv* 2019;3(9):1553–62. <https://doi.org/10.1182/bloodadvances.2018030007>.
- [157] Fraser G, Cramer P, Demirkan F, Silva RS, Grosicki S, Pristupa A, et al. Updated results from the phase 3 HELIOS study of ibrutinib, bendamustine, and rituximab in relapsed chronic lymphocytic leukemia/small lymphocytic lymphoma. *Leukemia* 2019;33:969–80. <https://doi.org/10.1038/s41375-018-0276-9>.
- [158] O'Brien SM, Lamanna N, Kipps TJ, Flinn I, Zelenetz AD, Burger JA, et al. A phase 2 study of idelalisib plus rituximab in treatment-naïve older patients with chronic lymphocytic leukemia. *Blood* 2015;126(25):2686–94. <https://doi.org/10.1182/blood-2015-03-630947>.
- [159] Sharman JP, Coutre SE, Furman RR, Cheson BD, Pagel JM, Hillmen P, et al. Final results of a randomized, phase III study of rituximab with or without idelalisib followed by open-label idelalisib in patients with relapsed chronic lymphocytic leukemia. *J Clin Oncol* 2019;37(16):1391–402. <https://doi.org/10.1200/JCO.18.01460>.
- [160] Zelenetz AD, Barrientos JC, Brown JR, Coiffier B, Delgado J, Egyed M, et al. Idelalisib or placebo in combination with bendamustine and rituximab in patients with relapsed or refractory chronic lymphocytic leukaemia: interim results from a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet Oncol* 2017;18(3):297–311. [https://doi.org/10.1016/S1470-2045\(16\)30671-4](https://doi.org/10.1016/S1470-2045(16)30671-4).
- [161] Flinn IW, Hillmen P, Montillo M, Nagy Z, Illés Á, Etienne G, et al. The phase 3 DUO trial: duvelisib vs ofatumumab in relapsed and refractory CLL/SLL. *Blood* 2018;132(23):2446–55. <https://doi.org/10.1182/blood-2018-05-850461>.
- [162] Cramer P, von Tresckow J, Bahlo J, Robrecht S, Langerbeins P, Al-Sawaf O, et al. Bendamustine followed by obinutuzumab and venetoclax in chronic lymphocytic leukaemia (CLL2-BAG): primary endpoint analysis of a multicentre, open-label, phase 2 trial. *Lancet Oncol* 2018;19(9):1215–28. [https://doi.org/10.1016/S1470-2045\(18\)30414-5](https://doi.org/10.1016/S1470-2045(18)30414-5).
- [163] Flinn IW, Gribben JG, Dyer MJS, Wierda W, Maris MB, Furman RR, et al. Phase 1b study of venetoclax-obinutuzumab in previously untreated and relapsed/refractory chronic lymphocytic leukemia. *Blood* 2019;133(26):2765–75. <https://doi.org/10.1182/blood-2019-01-896290>.
- [164] Roberts AW, Davids MS, Pagel JM, Kahl BS, Puvvada SD, Gerecitano JF, et al. Targeting BCL2 with venetoclax in relapsed chronic lymphocytic leukemia. *N Engl J Med* 2016;374(4):311–22. <https://doi.org/10.1056/NEJMoa1513257>.
- [165] Seymour JF, Kipps TJ, Eichhorst B, Hillmen P, D'Rozario J, Assouline S, et al. Venetoclax-rituximab in relapsed or refractory chronic lymphocytic leukemia. *N Engl J Med* 2018;378(12):1107–20. <https://doi.org/10.1056/NEJMoa1713976>.
- [166] Jones JA, Mato AR, Wierda WG, Davids MS, Choi M, Cheson BD, et al. Venetoclax for chronic lymphocytic leukaemia progressing after ibrutinib: an interim analysis of a multicentre, open-label, phase 2 trial. *Lancet Oncol* 2018;19(1):65–75. [https://doi.org/10.1016/S1470-2045\(17\)30909-9](https://doi.org/10.1016/S1470-2045(17)30909-9).
- [167] Coutre S, Choi M, Furman RR, Eradat H, Heffner L, Jones JA, et al. Venetoclax for patients with chronic lymphocytic leukemia who progressed during or after idelalisib therapy. *Blood* 2018;131(15):1704–11. <https://doi.org/10.1182/blood-2017-06-788133>.
- [168] Rogers KA, Huang Y, Ruppert AS, Awan FT, Heerema NA, Hoffman C, et al. Phase 1b study of obinutuzumab, ibrutinib, and venetoclax in relapsed and refractory chronic lymphocytic leukemia. *Blood* 2018;132(15):1568–72. <https://doi.org/10.1182/blood-2018-05-853564>.
- [169] Eyre TA, Kirkwood AA, Gohill S, Follows G, Walewska R, Walter H, et al. Efficacy of venetoclax monotherapy in patients with relapsed chronic lymphocytic leukaemia in the post-BCR inhibitor setting: a UK wide analysis. *Br J Haematol* 2019;185(4):656–69. <https://doi.org/10.1111/bjh.15802>.
- [170] Roeker LE, Fox CP, Eyre TA, Brander DM, Allan JN, Schuster SJ, et al. Tumor lysis, adverse events, and dose adjustments in 297 venetoclax-treated CLL patients in routine clinical practice. *Clin Cancer Res* 2019;25(14):4264–70. <https://doi.org/10.1158/1078-0432.CCR-19-0361>.
- [171] Gaidano G, Foà R, Dalla-Favera R. Molecular pathogenesis of chronic lymphocytic leukemia. *J Clin Invest* 2012;122(10):3432–8. <https://doi.org/10.1172/JCI64101>.

## APPENDIX K

Sedlarikova L\*, **Petrackova A\***, Papajik T, Turcsanyi P, Kriegova E. Resistance-associated mutations in chronic lymphocytic leukaemia patients treated with novel agents. *Front Oncol.* 2020;10:894. (IF 2020: 6.244) \*Contributed equally



# Resistance-Associated Mutations in Chronic Lymphocytic Leukemia Patients Treated With Novel Agents

Lenka Sedlarikova<sup>1†</sup>, Anna Petrackova<sup>1†</sup>, Tomas Papajik<sup>2</sup>, Peter Turcsanyi<sup>2</sup> and Eva Kriegova<sup>1\*</sup>

<sup>1</sup> Department of Immunology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czechia, <sup>2</sup> Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University and University Hospital, Olomouc, Czechia

## OPEN ACCESS

### Edited by:

Dimitar G. Efremov,  
International Centre for Genetic  
Engineering and Biotechnology, Italy

### Reviewed by:

Lydia Scarfò,  
San Raffaele Hospital, Italy  
Bartosz Michał Pula,  
Institute of Hematology and  
Transfusiology, Poland

### \*Correspondence:

Eva Kriegova  
eva.kriegova@email.cz

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Hematologic Malignancies,  
a section of the journal  
Frontiers in Oncology

Received: 07 March 2020

Accepted: 06 May 2020

Published: 25 June 2020

### Citation:

Sedlarikova L, Petrackova A,  
Papajik T, Turcsanyi P and Kriegova E  
(2020) Resistance-Associated  
Mutations in Chronic Lymphocytic  
Leukemia Patients Treated With Novel  
Agents. *Front. Oncol.* 10:894.  
doi: 10.3389/fonc.2020.00894

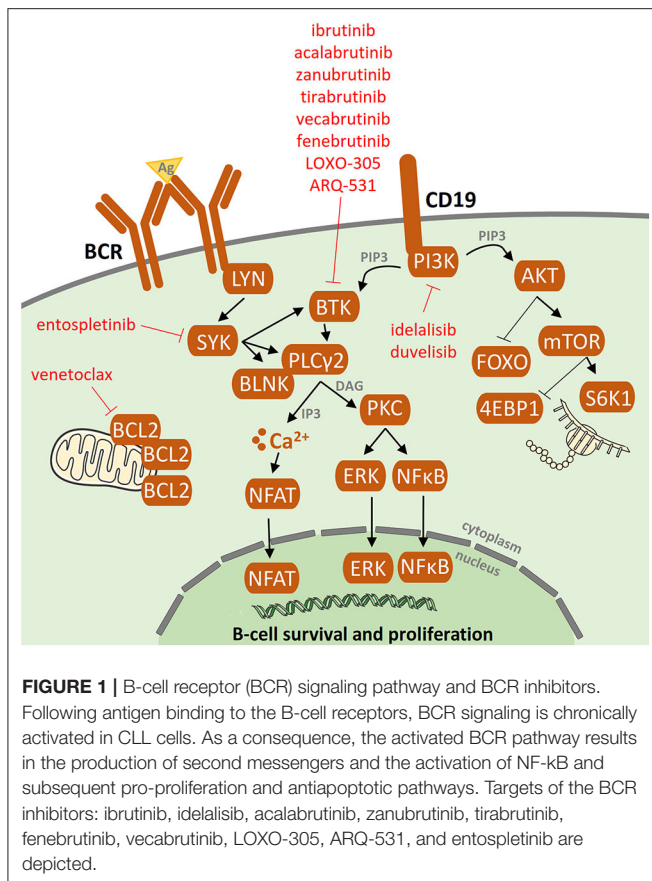
Inhibitors of B-cell receptor signaling, ibrutinib and idelalisib, and BCL-2 antagonist, venetoclax, have become the mainstay of treatment for chronic lymphocytic leukemia (CLL). Despite significant efficacy in most CLL patients, some patients develop resistance to these agents and progress on these drugs. We provide a state-of-the-art overview of the acquired resistance to novel agents. In 80% of patients with ibrutinib failure, acquired mutations in *BTK* and *PLCG2* genes were detected. No distinct unifying resistance-associated mutations or deregulated signaling pathways have been reported in idelalisib failure. Acquired mutations in the *BCL2* gene were detected in patients who had failed on venetoclax. In most cases, patients who have progressed on ibrutinib and venetoclax experience resistance-associated mutations, often present at low allelic frequencies. Resistance-associated mutations tend to occur between the second and fourth years of treatment and may already be detected several months before clinical relapse. We also discuss the development of next-generation agents for CLL patients who have acquired resistant mutations to current inhibitors.

**Keywords:** CLL, targeted therapy, resistance-associated mutations, treatment resistance, genetic aberrations

## INTRODUCTION

In the last decade, the great shift in the therapeutic management of chronic lymphocytic leukemia (CLL) is attributed to the approval of two B-cell receptor (BCR) signaling pathway inhibitors (ibrutinib and idelalisib) and the BCL-2 antagonist venetoclax. These novel agents have shown significant clinical efficacy in high-risk patients with relapsed/refractory (R/R) disease with 17p deletion and/or *TP53* mutation and complex karyotype as well as in previously untreated patients with/without poor-risk features (1–4). Despite the induction of long-term remission in most CLL patients, in some patients, the treatment fails. The number of patients who progress or develop clinical resistance is expected to increase with the growing number of patients indicated for this treatment and due to the long-term administration of these agents. Therefore, understanding the mechanisms driving resistance and identification of involved driver mutations and signaling pathways is a current need. These findings will help to design new targeted therapies to overcome resistance and to find drug combinations that will prevent the development of resistance or will help to avoid relapse. This manuscript provides an up-to-date overview of the acquired resistance-associated mutations in CLL patients treated with ibrutinib, idelalisib, and venetoclax. These mutations emerge during the treatment course





and predispose to the loss of function of the drug and disease progression. Moreover, we summarize recent findings in ways on how to manage the patients who have acquired resistant mutations to current inhibitors.

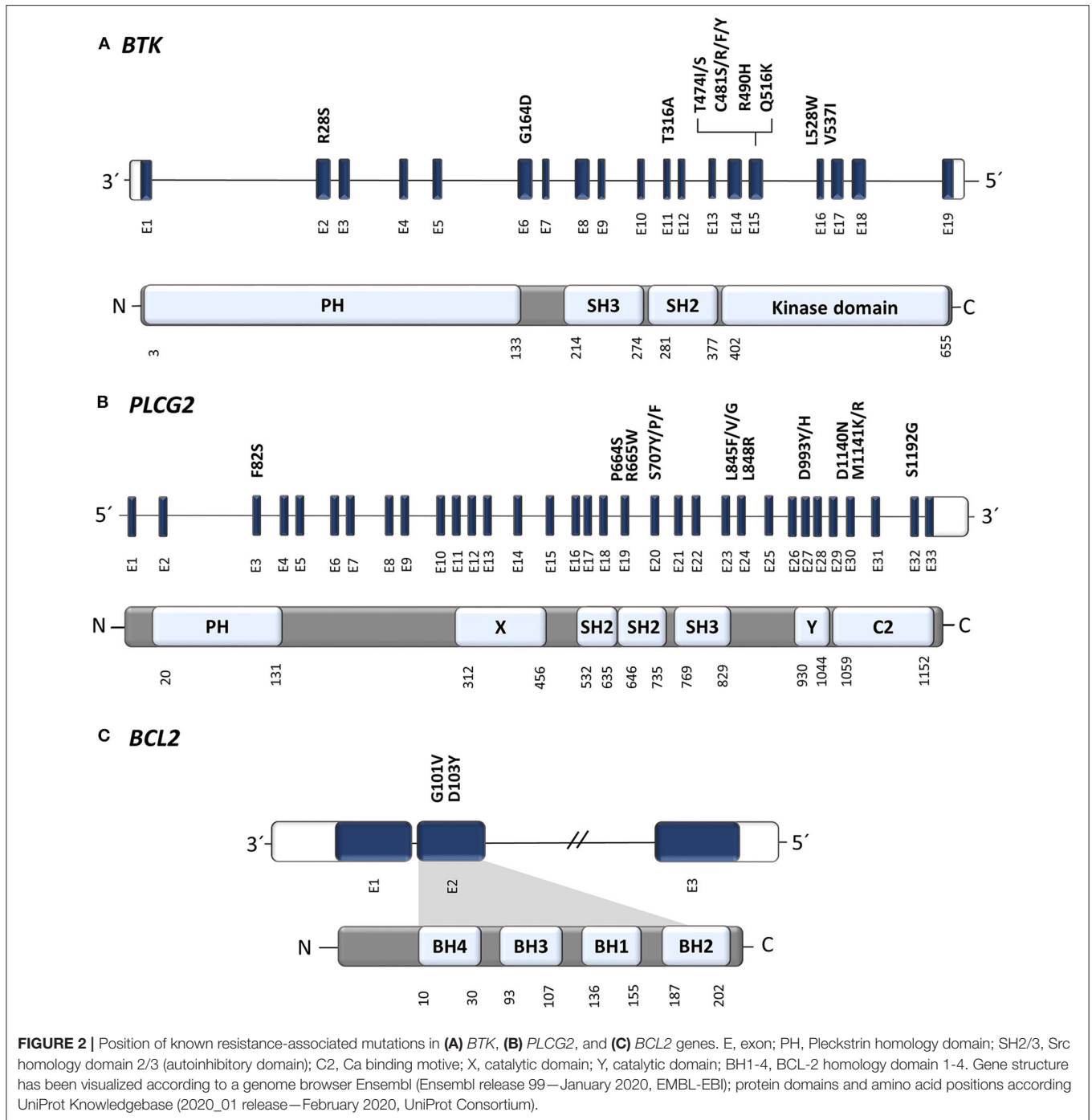
## RESISTANCE-ASSOCIATED MUTATIONS IN IBRUTINIB TREATMENT

So far, most information about resistance-associated mutations has been described for patients treated with ibrutinib, an oral agent inactivating Bruton's tyrosine kinase (BTK). This antiproliferative and proapoptotic agent, acting via formation of an irreversible covalent bond at the C481 position of BTK, inhibits both B-cell receptor (BCR) and NF-κB pathways [Figure 1; (5, 6)]. Ibrutinib has been shown to be highly effective not only in R/R CLL patients with del(17p) and/or TP53 mutation or complex karyotype (1, 2) but also in previously untreated patients with/without poor-risk features (3, 4). Primary resistance with no initial response to ibrutinib has been observed rarely and its mechanism is not yet understood (7–9). Acquired secondary resistance to ibrutinib occurs in 8–13% of CLL cases who responded well to the treatment initiation (7, 8) and are most commonly caused by the occurrence of resistant-associated mutations, as reported below (Figure 2).

In 2014, a study using whole-exome sequencing discovered acquired mutations within the *BTK* gene in 5/6 high-risk CLL patients relapsing on ibrutinib (10). A recent study on 30 CLL patients with residual lymphocytosis treated with ibrutinib for 3 years confirmed the presence of *BTK* mutations in 57% of CLL patients, and the presence of *BTK* mutations was associated with subsequent relapse (11). A number of studies have confirmed the presence of this mutation in CLL patients relapsing on ibrutinib and have shown that those mutations were not present prior to drug administration (12–14). The most common mutation (C481S) was found at the position of the binding site for ibrutinib thus reducing ibrutinib affinity for BTK (15, 16). More variants in the *BTK* gene, such as C481R, C481F, and C481Y, as well as less frequent variants at other gene positions (R28S, G164D, T316A, T474I/S, R490H, Q516K, L528W, and V537I), were revealed in later studies [Figure 2, Table 1; (12, 17–21)]. However, mutations outside of the kinase domain are rare (19). Functional characterization of these mutations has shown that the increase in ibrutinib dose is not sufficient to overcome the effect of the C481S/R/F/Y mutations (16). The *BTK* mutations usually develop between the second and fourth year of ibrutinib treatment (median 34.3 months, range 14–76.8 months) [Figure 3; (17)].

*PLCG2* gain-of-function mutations are the second most frequent mutations found in CLL patients who failed on the ibrutinib treatment (Figure 2). *PLCG2* is the gene encoding phospholipase C $\gamma$ 2, the protein immediately downstream of BTK. These mutations mostly have an activating effect resulting in continuous BCR signaling independently on BTK activation (10, 22). *PLCG2* mutation hotspots are located in several domains of the gene and often co-occur with hotspot *BTK* mutations (11–13, 21, 23). Moreover, different *PLCG2* mutations are usually found in multiple subclones with low allelic burden (12, 13, 17). A recent study confirmed *PLCG2* mutations in 13% of ibrutinib treated patients with residual lymphocytosis (11). However, the exact contribution of the *PLCG2* to the clinical resistance of CLL patients remains not fully understood (22). Similarly, as *BTK* mutations, also *PLCG2* mutations usually occur between the second and fourth year on ibrutinib (median 35.1 months, range 17.4–64.6 months) [Figure 3; (17)].

Although mutations in *BTK* and *PLCG2* genes are detected in ~80% of CLL patients who failed on ibrutinib (10, 12, 13), for 20% of patients, ibrutinib resistance-associated mutations remain unknown (13, 14, 23). These data further support the presence of alternative mechanisms of drug resistance other than *BTK/PLCG2* mutations in a subset of patients who are still under investigation. To date, several candidate loci/mutations that may contribute to resistance have been described. These include del(8p) and *SF3B1*, *PCLO*, *EP300*, *MLL2*, and *EIF2A* mutations (10, 14, 23). More candidate genetic factors associated with resistance for ibrutinib-treated patients involve *BCL6* rearrangements, *MYC* gene abnormalities, del(17p), del(18p), 2p gain, *XPO1* overexpression, complex karyotype, epigenetic changes, and changes in the cell microenvironment (1, 14, 18, 24). However, it remains unclear whether these aberrations contribute causally to clinical resistance.



## BTK AND/OR PLCG2 MUTATIONS CLONALITY

Studies of follow-up samples revealed that mutations within the *BTK* and *PLCG2* genes are often present several months before clinical relapse is observed [Figure 3; (11–13)]. Those resistance-associated mutations were detected as early as 9.3 months prior to clinical progression (13). Moreover, Ahn et al. revealed that

the resistance-associated mutations may be detected using highly sensitive approaches 15 months prior to disease progression (12).

The same *BTK* mutations have also been reported for CLL patients who experienced Richter transformation on ibrutinib treatment. In this group of patients, the resistance-associated mutations appeared in a smaller portion of patients (~40%) earlier, mostly within 15 months after ibrutinib treatment initiation [Figure 3; (13, 14)].

**TABLE 1** | Overview of selected studies describing the incidence of resistance-associated mutations in CLL patients treated with ibrutinib including selected studies describing resistance-associated mutations in patients with Richter transformation<sup>#</sup>.

References	Number of patients	Genetic characteristics of patients	Number of previous treatment lines, median (min-max)	Frequency of patients with resistance-associated mutations	BTK mutation		PLCG2 mutation	
					Variant(s)	VAF	Variant(s)	VAF
Woyach et al. (10)	6	3/6 del(17p) + CK 1/6 del(17p) + trisomy 12 1/6 CK 1/6 del(11q)	4 (2–9)	100% (6/6)	C481S	17–60%	R665W S707Y L845F	8–38%
Burger et al. (23)	5	4/5 del(17p) + del(13q) 1/5 del(11q)	4 (1–6)	40% (2/5)	C481S	NA	S707F D993H M1141K/R	12–35%
Sharma et al. (19) <sup>#</sup>	1*	1/1 del(17p)	1	100% (1/1)	T316A	75%	0	0
Ahn et al. (12) <sup>#</sup>	10	10/10 del(17p)/mutation TP53	NA	80% (8/10)	C481S/R	2–78%	P664S R665W S707Y L845F 6 nt del	0.1–18%
Woyach et al. (13)	54	40% del(17p)** 58% CK	3 (0–16)	89% (48/54)	C481S/R/F	0.2–100%	R665W S707Y/P/F L845F D993Y L845-846del	4–44%
Kadri et al. (14) <sup>#</sup>	9***	8/9 del(17p)	2 (1–4)	56% (5/9)	C481S/R T316A	3–90%	0	0
Quinquenel et al. (11)	30	15/30 mutation TP53 4/30 mutated IGHV	2 (NA-NA)	57% (17/30)	C481S/Y/R/G	0.2–73%	R665W L845G C849R D993H	1–11%
Gángó et al. (21) <sup>#</sup>	20****	7/20 del(17p) 10/20 del(13q) 6/20 trisomy 12 3/20 del(11q)	3 (1–5)	40% (8/20)	R28S G164D R490H C481S/Y Q516K	2.7–27.3%	F82S R694H D993H S1192G	2.6–4.9%

NA, not available; CK, complex karyotype; VAF, variant allele frequency.

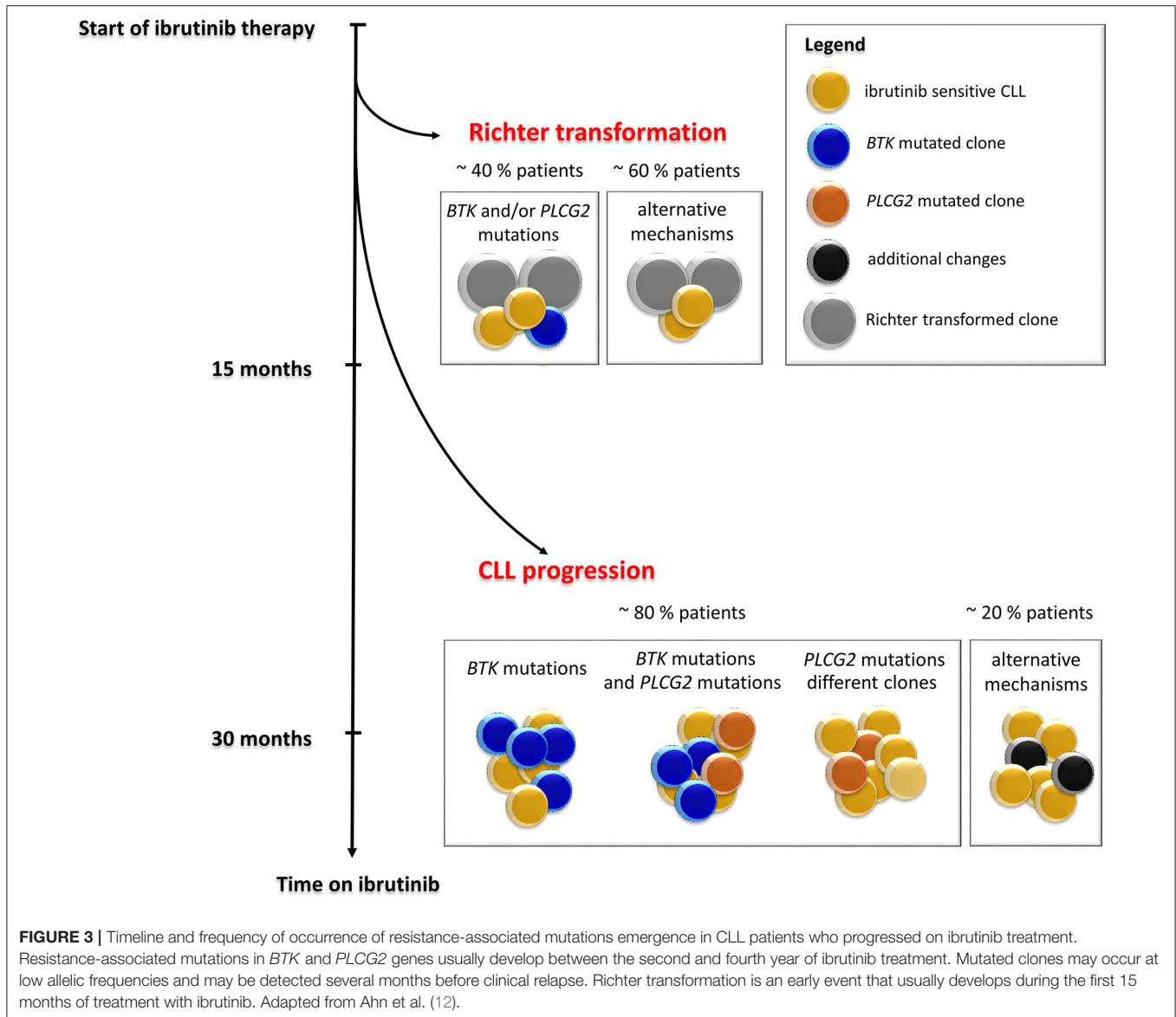
\*Mutation status of a patient with Richter transformation was described; resistance-associated mutation was found in peripheral blood, not in lymph nodes (19). \*\*Patient characteristics for a complete set of 308 patients (13). \*\*\*6/9 patients with Richter transformation were included (14). \*\*\*\*3/20 patients with Richter transformation were included (21).

Nearly half of the patients who progress on ibrutinib carry *BTK* mutations in minor subclones with low variant allele frequencies (VAF) below 10% (12, 13). Nevertheless, cases with variants >80% VAF were also no exception (25). Clonal development and subclonal heterogeneity of resistance-associated mutations, i.e., the detection of multiple independent subclones carrying different variants of resistance-associated mutations with distinct growth rates, are also frequently observed in ibrutinib-treated patients (12, 23). In particular, it applies to *PLCG2* gene variants with lower VAF. However, the precise mechanism by which *BTK* and/or *PLCG2* mutations drive clinical resistance when present at such low allelic frequencies has not been elucidated yet.

In the recent study of Gángó et al., an association between reduction or even elimination of *TP53* mutated clones and the presence of *BTK* mutations has been reported (21). Vice versa, in patients with persisting *TP53* mutated subclones, no *BTK* mutations were detected in this study. The authors speculate that

a longer duration of ibrutinib treatment could create conditions for the survival of subclones with *BTK* mutations by eliminating subclones with *TP53* mutations or that the elimination of *TP53* subclones enables the expansion of subclones harboring *BTK* mutations (21). The loss of preexisting mutations in *TP53* and *BIRC3* genes in patients who gained *BTK* mutations was also observed in another recent study (20). The relationship between *TP53* and *BTK* mutations should be confirmed on a larger patient cohort.

Another study showed that resistant mutated subclones and disease progression may occur only in specific compartments (13, 14). For some patients, Woyach et al. reported the presence of resistance-associated mutations only in lymph nodes without a presence in a corresponding peripheral blood sample (13). For these patients, the disease progression was also observed only in the lymph nodes. Therefore, it is necessary to choose the right collection material for the analysis of resistance-associated mutations, and even a negative result from



peripheral blood may not mean that the patient does not have a subclone with a resistance-associated mutation in another location (25).

## RESISTANCE-ASSOCIATED MUTATIONS IN IDELALISIB TREATMENT

Idelalisib is a phosphatidylinositol 3-kinase  $\delta$  isoform (PI3K $\delta$ ) inhibitor, a member of the BCR inhibitor family, which not only inhibits PI3K signaling but also selectively induces apoptosis in CLL cells [Figure 1; (26)]. Idelalisib was approved for the treatment of patients with R/R CLL and high-risk patients with *TP53* disruption (27, 28). Regardless of its clinical efficacy, disease progression during idelalisib treatment is observed in some CLL patients (29–31).

Despite intensive research, the biological mechanism of the disease progression in idelalisib-treated patients remains unknown and no resistance-associated mutations in specific gene(s) or signaling pathway alterations have been found so far (29–31). A whole-exome sequencing study in a small cohort of 13 CLL patients who progressed on idelalisib treatment revealed that no mutations occurred in the PI3K signaling pathway or in any related signaling pathway (31).

These results indicate that there is probably no single pathway or specific mutation associated with idelalisib resistance thus deserving future investigations. A recent study investigating *in vivo* mouse models of resistance to PI3K inhibitors identified upregulation of genes from the integrin receptor complex as a possible mechanism of resistance (30). This suggests that resistance to idelalisib may be more likely mediated by dysregulation of survival signaling rather than by recurrent mutations. Nevertheless, it remains to be determined to



what extent this mechanism of resistance plays a role in human CLL.

## RESISTANCE-ASSOCIATED MUTATIONS IN VENETOCLAX TREATMENT

Venetoclax is an oral BH3 mimetic and highly selective inhibitor of the BCL-2 antiapoptotic protein, capable of restoring apoptosis tumor cells with high overall response rates as well as in heavily pretreated, high-risk CLL patients. Venetoclax is primarily available for CLL patients with *TP53* disruption, for patients who failed on ibrutinib or were not suitable for the treatment with BCR signaling inhibitors, as well as for patients refractory to chemoimmunotherapy (32–34).

There is already growing evidence about the role of acquired mutations leading to the progression and failure of venetoclax. A recent study reported G101V mutation in the *BCL2* gene in 7 of 15 (47%) CLL patients progressing on venetoclax (35). Functional analysis demonstrated that G101V mutation disrupts the bond of venetoclax to BCL-2, therefore preventing the drug from competing with proapoptotic molecules to bind with BCL-2. The G101V mutation was absent at baseline, first detected 19–42 months after the initiation of venetoclax treatment and 25 months prior to clinical relapse, and persisted in five of seven patients for more than 6 months after the discontinuation of venetoclax therapy (35). The G101V mutation showed a high variability in VAF ranging from 1.4 to 70% (35).

Another recent study confirmed G101V in three of four CLL patients treated with venetoclax and found a second *BCL2* variant, D103Y, also within the BH3-binding site (36). The D103Y was acquired 39 months after the initiation of venetoclax therapy, at first with a low VAF of 7%, increasing to 18% over the next 5 months of venetoclax treatment in the reported patient. At a later date, this patient became positive also for G101V mutation with 14% VAF. NGS analysis showed that these two mutations exist as independent subclones with different growth dynamics, and they were not detected in the control group of 546 CLL patients not treated with venetoclax (36).

A whole-exome sequencing study in a small cohort of eight patients with del(17p) progressing on venetoclax identified a number of candidate resistance-associated aberrations, such as homozygous deletions of *CDKN2A/B* resulting in the loss of cell cycle control in three patients and mutations in the antiproliferative *BTG1* gene in two patients (37). Analysis of pretreatment samples revealed that these aberrations developed after the treatment initiation. In this study, the spectrum of other mutations in CLL-associated genes has been identified, such as *BRAF*, *NOTCH1*, *RBI*, *SF3B1*, and *TP53* mutations. Nevertheless, their causal relationship to resistance has not yet been established.

## CURRENT DEVELOPMENTS TO CIRCUMVENT FAILURE ON NOVEL AGENTS

Current studies have provided evidence that the early and sensitive detection of *BTK*, *PLCG2*, and *BCL2* mutations may be predictive of an impending relapse, as they occur several

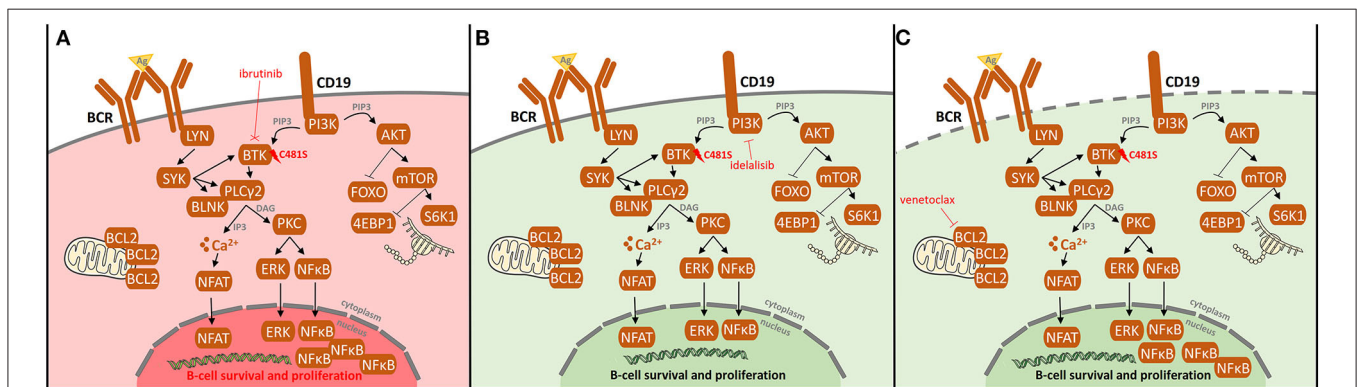
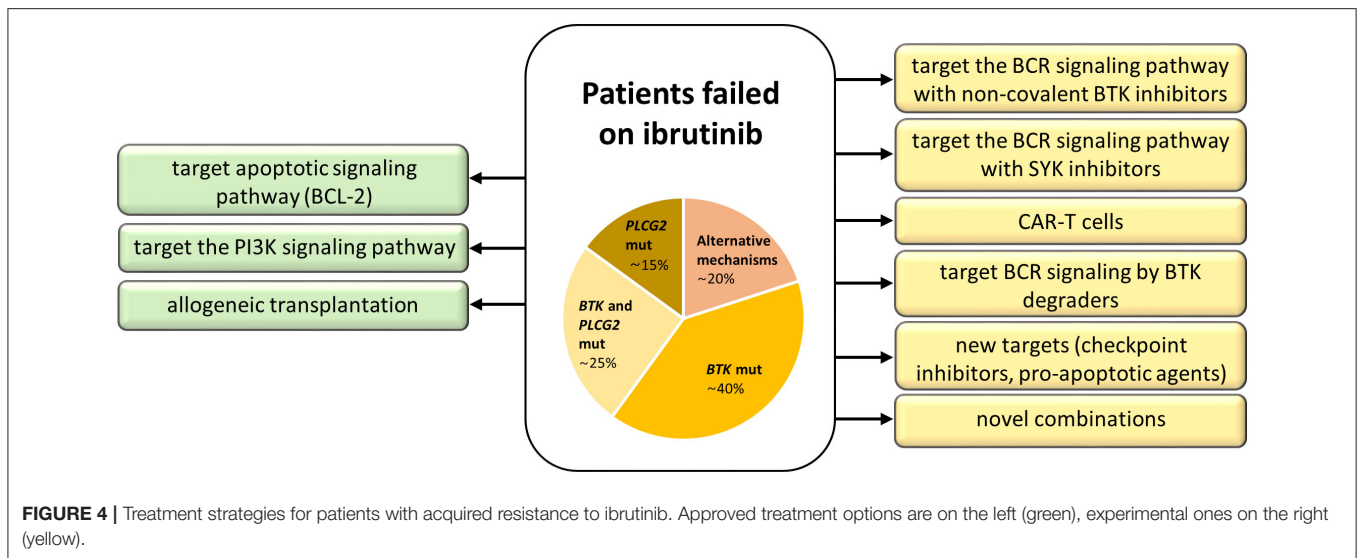
months before clinical relapse (35, 38). Therefore, the emergence of the resistance-associated mutations in patients receiving long-term treatment with BCR and BCL-2 inhibitors should be tested at regular intervals using highly sensitive ultradeep NGS approaches. As the resistance-associated mutations occur often at low VAF, it is necessary to standardize sequencing parameters in order to minimize the probability of false-positive and false-negative results (39–42).

In the case of a positive result for *BTK*, *PLCG2*, and *BCL2* mutations, the treatment should not be discontinued but rather different appropriate inhibitors or alternative combination therapies should be considered (43, 44). Therefore, understanding the nature of disease progression and emergence of clinical resistance has important implications, especially in the development of new treatment strategies that would help overcome resistance before it develops or as it emerges (43).

There are currently several options, clinically proven as well as in experimental settings, how to circumvent failure on ibrutinib and other novel agents (Figure 4). The most promising is to target alternative molecules or signaling pathways involved in the survival of malignant cells (Figure 5). Another therapeutic intervention could be the use of different combination therapies to reverse relapse (43). Very promising is the combination of ibrutinib with anti-CD20 antibodies (28). In addition, the combination of ibrutinib with venetoclax represents a great potential in the treatment of CLL, as there is increasing evidence that CLL cells previously treated with BCR inhibitors show an increased dependence on BCL-2 expression (45, 46). This is further supported by the observation of the changes in BCL-2 family protein levels in CLL cells when treated with ibrutinib (45) and acalabrutinib (47). Additionally, venetoclax complements ibrutinib and acalabrutinib-mediated apoptosis in CLL cells (45, 47). The synergic effect of both drugs acting by different mechanisms has led to a deep therapeutic effect in CLL (45, 46, 48); however, the exact mechanism of their interaction in CLL should be further elucidated.

## NEXT-GENERATION INHIBITORS AND EXPERIMENTAL TREATMENT OPTIONS

New covalent BTK inhibitors with high selectivity to BTK, such as acalabrutinib (ACP-196), zanubrutinib (BGB-3111), and tirabrutinib (ONO/GS-4059), have high therapeutic potential [Figure 1; (49–51)]. However, like ibrutinib, they covalently bind to BTK and are, therefore, not suitable for the treatment of patients with resistance-associated *BTK* hotspot mutation (C481S/R/F/Y). Compared to ibrutinib, the recently approved acalabrutinib (52) shows a more acceptable safety profile and modulation of BCL-2 family proteins contributing to cell death induced by venetoclax after acalabrutinib treatment in CLL (47). Similarly to ibrutinib, acquired resistance to acalabrutinib was mainly associated with *BTK* mutations (53). Zanubrutinib is the next potent and highly selective inhibitor of BTK, currently approved for mantle cell lymphoma treatment and tested in clinical trials for CLL (52). In four CLL patients progressing on zanubrutinib treatment, Handunnetti et al. identified *BTK* C481



**FIGURE 5 |** Resistance-associated *BTK* (C481) mutation and possibilities of its overcoming in ibrutinib-treated CLL patients. **(A)** After the occurrence of *BTK* (C481) mutation, ibrutinib does not bind to BTK causing the loss of its therapeutic effect leading to BCR and NF-κB pathway activation, both signaling pathways pivotal for maintenance and proliferation of CLL cells. Additionally, activating mutations in *PLCG2* gene may result in continuous BCR signaling independently on BTK activation. To overcome the resistance by mutations in *BTK/PLCG2* genes, other pathways may be targeted. **(B)** Idelalisib targets phosphatidylinositol 3-kinase (PI3K), resulting in the downregulation of PI3K/mTOR signaling pathway, irrespectively on the *BTK* mutation status. **(C)** Venetoclax targets BCL-2, thus leading to the apoptosis of CLL cells. As the previous treatment with BCR inhibitors results in an increased dependence on BCL-2 expression together with the fact that venetoclax complements ibrutinib-mediated apoptosis, a deep therapeutic effect on venetoclax is achieved.

and L528W mutations, both of them absent prior to zanubrutinib treatment (54).

Importantly, a new generation of non-covalent BTK inhibitors is currently being tested such as fenebrutinib (GDC-0853), ARQ-531, LOXO-305, and vecabrutinib (SNS-062) (51, 52, 55). The first clinical studies with fenebrutinib have shown that non-covalent selective inhibition of BTK may be effective in CLL patients with acquired resistance to ibrutinib therapy (56). But *in vitro* mutagenesis of the *BTK* gene has shown that mutations in the kinase domain (L512M, E513G, F517L, and L547P) reduce the effect of these new non-covalent BTK inhibitors (57). Also, ARQ-531, an ATP-competitive non-covalent reversible inhibitor, is able to overcome *BTK* (C481S) and *PLCG2* (R665W, S707P, S707F, R742P, and L845fs) mutations as shown in animal models (58). Similarly, the reversible BTK inhibitor LOXO-305 shows great potential in overcoming acquired resistance to irreversible

BTK inhibitors in preclinical CLL models; phase 1 clinical trial of LOXO-305 is currently ongoing (59).

Among experimental treatment options in patients who failed on ibrutinib belong SYK (Spleen tyrosine kinase) inhibitors, chimeric antigen receptor T (CAR-T) cell therapy, anti-PD-1 treatment, and BTK degraders (Figure 4). SYK is a kinase upstream of both BTK and PI3Kδ in the BCR signaling pathway. The SYK inhibitor entospletinib shows a clinical activity for R/R CLL patients who have relapsed on BTK or PI3Kδ inhibitors, even in the presence of *BTK* and *PLCG2* mutations (44). CAR-T cell therapy (60) represents another treatment opportunity for CLL patients failing BCR or BCL-2 inhibitor therapies, currently available only in clinical trials (52). Further, response to PD-1 blockade with pembrolizumab has been observed in patients with Richter transformation who had progression after prior therapy with ibrutinib, but not in R/R CLL patients (61). Another

emerging approach evidenced to have an effect on both wild-type and mutated BTK (C481S) in preclinical studies is the novel agent BTK degrader, MT-802, causing ubiquitination of BTK, and subsequent degradation through proteasome (62). A recent study on animal models has shown the possibility to overcome resistance to ibrutinib by preventing FOXO3a nuclear export and PI3K/AKT activation (63). These promising data and the development of new approaches to overcome acquired resistance to current BCR and BCL-2 inhibitors show great hope not only for CLL patients.

## CONCLUSION

With increasing experience in the treatment with novel agents, it has become extremely important to (i) detect genetic aberrations associated with resistance and progression and (ii) understand the mechanisms of resistance development and disease progression in patients treated with these agents. Resistance-associated mutations in *BTK*, *PLCG2*, and *BCL2* have the potential to be used as a biomarker for future relapse or disease progression and hence their detection could facilitate early therapeutic intervention therapy to prevent the relapse. In addition, understanding the mechanism of resistance may also help find a way on how to prevent resistance before it develops or to overcome as it emerges.

In the near future, we can expect an increase in the number of patients indicated for inhibitor therapy due to the excellent

performance of these agents, not only in CLL. Since patients are treated for longer with these drugs, it will certainly be reflected in the increased incidence of disease progression and failure on those agents; hence, the challenge of managing resistance and identifying which patients are at risk for relapse is of the highest importance. There is a long way to learn on how to manage a resistant disease that we will encounter more and more often in the era of treatment with novel targeted agents.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article.

## AUTHOR CONTRIBUTIONS

LS and EK wrote the manuscript. LS and AP prepared the figures and revisions. TP, PT, and AP revised the manuscript critically. EK contributed to conceptualization, study design, writing, and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Ministry of Health of the Czech Republic (MH CZ VES16-32339A), in part by FNOL 00098892, and the Internal Grant Agency of Palacky University (IGA UP\_2020\_016).

## REFERENCES

- Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med.* (2013) 369:32–42. doi: 10.1056/NEJMoa1215637
- Byrd JC, Hillmen P, O'Brien S, Barrientos JC, Reddy NM, Coutre S, et al. Long-term follow-up of the RESONATE phase 3 trial of ibrutinib vs ofatumumab. *Blood.* (2019) 133:2031–42. doi: 10.1182/blood-2018-08-870238
- Burger JA, Tedeschi A, Barr PM, Robak T, Owen C, Ghia P, et al. Ibrutinib as initial therapy for patients with chronic lymphocytic leukemia. *N Engl J Med.* (2015) 373:2425–37. doi: 10.1056/NEJMoa1509388
- Barr PM, Robak T, Owen C, Tedeschi A, Bairey O, Bartlett NL, et al. Sustained efficacy and detailed clinical follow-up of first-line ibrutinib treatment in older patients with chronic lymphocytic leukemia: extended phase 3 results from RESONATE-2. *Haematologica.* (2018) 103:1502–10. doi: 10.3324/haematol.2018.192328
- Herman SE, Mustafa RZ, Gyamfi JA, Pittaluga S, Chang S, Chang B, et al. Ibrutinib inhibits BCR and NF- $\kappa$ B signaling and reduces tumor proliferation in tissue-resident cells of patients with CLL. *Blood.* (2014) 123:3286–95. doi: 10.1182/blood-2014-02-548610
- Yosifov DY, Wolf C, Stiglbauer S, Mertens D. From biology to therapy: the CLL success story. *Hemasphere.* (2019) 3:e175. doi: 10.1097/HS9.000000000000175
- Kaur V, Swami A. Ibrutinib in CLL: a focus on adverse events, resistance, and novel approaches beyond ibrutinib. *Ann Hematol.* (2017) 96:1175–84. doi: 10.1007/s00277-017-2973-2
- Pula B, Gołos A, Górniak P, Jamrozik K. Overcoming ibrutinib resistance in chronic lymphocytic leukemia. *Cancers.* (2019) 11:E1834. doi: 10.3390/cancers11121834
- Woyach JA, Johnson AJ. Targeted therapies in CLL: mechanisms of resistance and strategies for management. *Blood.* (2015) 126:471–7. doi: 10.1182/blood-2015-03-585075
- Woyach JA, Furman RR, Liu TM, Ozer HG, Zapotka M, Ruppert AS, et al. Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. *N Engl J Med.* (2014) 370:2286–94. doi: 10.1056/NEJMoa1400029
- Quinquenel A, Fornecker LM, Letestu R, Ysebaert L, Fleury C, Lazarian G, et al. Prevalence of BTK and PLCG2 mutations in a real-life CLL cohort still on ibrutinib after 3 years: a FILO group study. *Blood.* (2019) 134:641–4. doi: 10.1182/blood.2019000854
- Ahn IE, Underbayev C, Albitar A, Herman SE, Tian X, Maric I, et al. Clonal evolution leading to ibrutinib resistance in chronic lymphocytic leukemia. *Blood.* (2017) 129:1469–79. doi: 10.1182/blood-2016-06-719294
- Woyach JA, Ruppert AS, Guinn D, Lehman A, Blachly JS, Lozanski A, et al. BTKC481S-mediated resistance to ibrutinib in chronic lymphocytic leukemia. *J Clin Oncol.* (2017) 35:1437–43. doi: 10.1200/JCO.2016.70.2282
- Kadri S, Lee J, Fitzpatrick C, Galanina N, Sukhanova M, Venkataraman G, et al. Clonal evolution underlying leukemia progression and Richter transformation in patients with ibrutinib-relapsed CLL. *Blood Adv.* (2017) 1:715–27. doi: 10.1182/bloodadvances.2016003632
- Cheng S, Guo A, Lu P, Ma J, Coleman M, Wang YL. Functional characterization of BTK(C481S) mutation that confers ibrutinib resistance: exploration of alternative kinase inhibitors. *Leukemia.* (2015) 29:895–900. doi: 10.1038/leu.2014.263
- Hamasy A, Wang Q, Blomberg KE, Mohammad DK, Yu L, Vihinen M, et al. Substitution scanning identifies a novel, catalytically active ibrutinib-resistant BTK cysteine 481 to threonine (C481T) variant. *Leukemia.* (2017) 31:177–85. doi: 10.1038/leu.2016.153
- Jones D, Woyach JA, Zhao W, Caruthers S, Tu H, Coleman J, et al. PLCG2 C2 domain mutations co-occur with BTK and PLCG2 resistance mutations in chronic lymphocytic leukemia undergoing ibrutinib treatment. *Leukemia.* (2017) 31:1645–7. doi: 10.1038/leu.2017.110
- Maddocks KJ, Ruppert AS, Lozanski G, Heerema NA, Zhao W, Abruzzo L, et al. Etiology of ibrutinib therapy discontinuation and outcomes in patients with chronic lymphocytic leukemia. *JAMA Oncol.* (2015) 1:80–7. doi: 10.1001/jamaoncol.2014.218



19. Sharma S, Galanina N, Guo A, Lee J, Kadri S, Van Slambrouck C, et al. Identification of a structurally novel BTK mutation that drives ibrutinib resistance in CLL. *Oncotarget*. (2016) 7:68833–41. doi: 10.18632/oncotarget.11932
20. Kanagal-Shamanna R, Jain P, Patel KP, Routbort M, Bueso-Ramos C, Alhalouli T, et al. Targeted multigene deep sequencing of Bruton tyrosine kinase inhibitor-resistant chronic lymphocytic leukemia with disease progression and Richter transformation. *Cancer*. (2019) 125:559–74. doi: 10.1002/cncr.31831
21. Gángó A, Alpár D, Galik B, Marosvári D, Kiss R, Fésüs V, et al. Dissection of subclonal evolution by temporal mutation profiling in chronic lymphocytic leukemia patients treated with ibrutinib. *Int J Cancer*. (2020) 146:85–93. doi: 10.1002/ijc.32502
22. Liu TM, Woyach JA, Zhong Y, Lozanski A, Lozanski G, Dong S, et al. Hypermorphic mutation of phospholipase C,  $\gamma 2$  acquired in ibrutinib-resistant CLL confers BTK independency upon B-cell receptor activation. *Blood*. (2015) 126:61–8. doi: 10.1182/blood-2015-02-626846
23. Burger JA, Landau DA, Taylor-Weiner A, Bozic I, Zhang H, Sarosiek K, et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun*. (2016) 7:11589. doi: 10.1038/ncomms11589
24. Maffei R, Fiorcari S, Martinelli S, Potenza L, Luppi M, Marasca R. Targeting neoplastic B cells and harnessing microenvironment: the “double face” of ibrutinib and idelalisib. *J Hematol Oncol*. (2015) 8:60. doi: 10.1186/s13045-015-0157-x
25. Lampson BL, Brown JR. Are BTK and PLCG2 mutations necessary and sufficient for ibrutinib resistance in chronic lymphocytic leukemia? *Expert Rev Hematol*. (2018) 11:185–94. doi: 10.1080/17474086.2018.1435268
26. Hoellenriegel J, Meadows SA, Sivina M, Wierda WG, Kantarjian H, Keating MJ, et al. The phosphoinositide 3'-kinase delta inhibitor, CAL-101, inhibits B-cell receptor signaling and chemokine networks in chronic lymphocytic leukemia. *Blood*. (2011) 118:3603–12. doi: 10.1182/blood-2011-05-352492
27. Brown JR, Byrd JC, Coutre SE, Benson DM, Flinn IW, Wagner-Johnston ND, et al. Idelalisib, an inhibitor of phosphatidylinositol 3-kinase p110 $\delta$ , for relapsed/refractory chronic lymphocytic leukemia. *Blood*. (2014) 123:3390–7. doi: 10.1182/blood-2013-11-535047
28. Burger JA, Wiestner A. Targeting B cell receptor signalling in cancer: preclinical and clinical advances. *Nat Rev Cancer*. (2018) 18:148–67. doi: 10.1038/nrc.2017.121
29. Arnason JE, Brown JR. Targeting B cell signaling in chronic lymphocytic leukemia. *Curr Oncol Rep*. (2017) 19:61. doi: 10.1007/s11912-017-0620-7
30. Scheffold A, Jebaraj BMC, Tausch E, Bloehdorn J, Ghia P, Yahiaoui A, et al. IGF1R as druggable target mediating PI3K- $\delta$  inhibitor resistance in a murine model of chronic lymphocytic leukemia. *Blood*. (2019) 134:534–47. doi: 10.1182/blood.2018881029
31. Ghia P, Ljungström V, Tausch E, Agathangelidis A, Scheffold A, Scarfo L, et al. Whole-exome sequencing revealed no recurrent mutations within the PI3K pathway in relapsed chronic lymphocytic leukemia patients progressing under idelalisib treatment. *Blood*. (2016) 128:2770. doi: 10.1182/blood.V128.22.2770.2770
32. Stilgenbauer S, Eichhorst B, Schetelig J, Coutre S, Seymour JF, Munir T, et al. Venetoclax in relapsed or refractory chronic lymphocytic leukaemia with 17p deletion: a multicentre, open-label, phase 2 study. *Lancet Oncol*. (2016) 17:768–78. doi: 10.1016/S1470-2045(16)30019-5
33. Jones JA, Mato AR, Wierda WG, Davids MS, Choi M, Cheso, et al. Venetoclax for chronic lymphocytic leukaemia progressing after ibrutinib: an interim analysis of a multicentre, open-label, phase 2 trial. *Lancet Oncol*. (2018) 19:65–75. doi: 10.1016/S1470-2045(17)30909-9
34. Roberts AW, Davids MS, Pagel JM, Kahl BS, Puvvada SD, Gerecitano JF, et al. Targeting BCL2 with venetoclax in relapsed chronic lymphocytic leukemia. *N Engl J Med*. (2016) 374:311–22. doi: 10.1056/NEJMoa1513257
35. Blombery P, Anderson MA, Gong JN, Thijssen R, Birkinshaw RW, Thompson ER, et al. Acquisition of the recurrent Gly101Val mutation in BCL2 confers resistance to venetoclax in patients with progressive chronic lymphocytic leukemia. *Cancer Discov*. (2019) 9:342–53. doi: 10.1182/blood-2018-120761
36. Tausch E, Close W, Dolnik A, Bloehdorn J, Chyla B, Bullinger L, et al. Venetoclax resistance and acquired BCL2 mutations in chronic lymphocytic leukemia. *Haematologica*. (2019) 104:e434–7. doi: 10.3324/haematol.2019.222588
37. Herling CD, Abedpour N, Weiss J, Schmitt A, Jachimowicz RD, Merkel O, et al. Clonal dynamics towards the development of venetoclax resistance in chronic lymphocytic leukemia. *Nat Commun*. (2018) 9:727. doi: 10.1038/s41467-018-03170-7
38. Sutton LA. Mechanisms of resistance to targeted therapies in chronic lymphocytic leukemia. *HemaSphere*. (2019) 3:40–3. doi: 10.1097/HS9.0000000000000240
39. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the association for molecular pathology and College of American Pathologists. *J Mol Diagn*. (2017) 19:341–65. doi: 10.1016/j.jmoldx.2017.01.011
40. Bacher U, Shumilov E, Flach J, Porret N, Joncourt R, Wiedemann G, et al. Challenges in the introduction of next-generation sequencing (NGS) for diagnostics of myeloid malignancies into clinical routine use. *Blood Cancer J*. (2018) 8:113. doi: 10.1038/s41408-018-0148-6
41. Merker JD, Devereaux K, Iafrate AJ, Kamel-Reid S, Kim AS, Moncur JT, et al. Proficiency testing of standardized samples shows very high interlaboratory agreement for clinical next-generation sequencing-based oncology assays. *Arch Pathol Lab Med*. (2019) 143:463–71. doi: 10.5858/arpa.2018-0336-CP
42. Petrackova A, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, et al. Standardization of sequencing coverage depth in NGS: recommendation for detection of clonal and subclonal mutations in cancer diagnostics. *Front Oncol*. (2019) 9:851. doi: 10.3389/fonc.2019.00851
43. Fürstenau M, Hallek M, Eichhorst B. Sequential and combination treatments with novel agents in chronic lymphocytic leukemia. *Haematologica*. (2019) 104:2144–54. doi: 10.3324/haematol.2018.208603
44. Awan FT, Thirman MJ, Patel-Donnelly D, Assouline S, Rao AV, Ye W, et al. Entospletinib monotherapy in patients with relapsed or refractory chronic lymphocytic leukemia previously treated with B-cell receptor inhibitors: results of a phase 2 study. *Leuk Lymphoma*. (2019) 60:1972–7. doi: 10.1080/10428194.2018.1562180
45. Deng J, Isik E, Fernandes SM, Brown JR, Letai A, Davids MS. Bruton's tyrosine kinase inhibition increases BCL-2 dependence and enhances sensitivity to venetoclax in chronic lymphocytic leukemia. *Leukemia*. (2017) 31:2075–84. doi: 10.1038/leu.2017.32
46. Bose P, Gandhi V. Recent therapeutic advances in chronic lymphocytic leukemia. *F1000Res*. (2017) 6:1924. doi: 10.12688/f1000research.11618.1
47. Patel VK, Lamothe B, Ayres ML, Gay J, Cheung JP, Balakrishnan K, et al. Pharmacodynamics and proteomic analysis of acalabrutinib therapy: similarity of on-target effects to ibrutinib and rationale for combination therapy. *Leukemia*. (2018) 32:920–30. doi: 10.1038/leu.2017.321
48. Tam CS, Siddiqi T, Allan JN, Kippes TJ, Flinn IW, Kuss BJ, et al. Ibrutinib (Ibr) plus Venetoclax (Ven) for first-line treatment of Chronic Lymphocytic Leukemia (CLL)/Small Lymphocytic Lymphoma (SLL): results from the MRD Cohort of the Phase 2 CAPTIVATE Study. *Blood*. (2019) 134:35. doi: 10.1182/blood-2019-121424
49. Byrd JC, Harrington B, O'Brien S, Jones JA, Schuh A, Devereux S, et al. Acalabrutinib (ACP-196) in relapsed chronic lymphocytic leukemia. *N Engl J Med*. (2016) 374:323–32. doi: 10.1056/NEJMoa1509981
50. Walter HS, Rule SA, Dyer MJ, Karlin L, Jones C, Cazin B, et al. A phase 1 clinical trial of the selective BTK inhibitor ONO/GS-4059 in relapsed and refractory mature B-cell malignancies. *Blood*. (2016) 127:411–9. doi: 10.1182/blood-2015-08-664086
51. Thompson PA, Burger JA. Bruton's tyrosine kinase inhibitors: first and second generation agents for patients with Chronic Lymphocytic Leukemia (CLL). *Expert Opin Investig Drugs*. (2018) 27:31–42. doi: 10.1080/13543784.2018.1404027
52. Iovino L, Shadman M. Novel therapies in chronic lymphocytic leukemia: a rapidly changing landscape. *Curr Treat Options Oncol*. (2020) 21:24. doi: 10.1007/s11864-020-0715-5
53. Woyach J, Huang Y, Rogers K, Bhat SA, Grever MR, Lozanski A, et al. Resistance to acalabrutinib in CLL is mediated primarily by BTK mutations. *Blood*. (2019) 134 (Suppl. 1):504. doi: 10.1182/blood-2019-127674

54. Handunnetti SM, Pek Sang Tang C, Nguyen T, Zhou X, Thompson E, Sun H, et al. BTK Leu528Trp - a potential secondary resistance mechanism specific for patients with chronic lymphocytic leukemia treated with the next generation BTK inhibitor. *Blood*. (2019) 134 (Suppl. 1):170. doi: 10.1182/blood-2019-125488
55. Crawford JJ, Johnson AR, Misner DL, Belmont LD, Castaneda G, Choy R, et al. Discovery of GDC-0853: a potent, selective, and noncovalent Bruton's tyrosine kinase inhibitor in early clinical development. *J Med Chem*. (2018) 61:2227–45. doi: 10.1021/acs.jmedchem.7b01712
56. Reiff SD, Muhowski EM, Guinn D, Lehman A, Fabian CA, Cheney C, et al. Noncovalent inhibition of C481S Bruton tyrosine kinase by GDC-0853: a new treatment strategy for ibrutinib-resistant CLL. *Blood*. (2018) 132:1039–49. doi: 10.1182/blood-2017-10-809020
57. Wang S, Mondal S, Zhao C, Berishaj M, Ghanakota P, Batlevi CL, et al. Noncovalent inhibitors reveal BTK gatekeeper and auto-inhibitory residues that control its transforming activity. *JCI Insight*. (2019) 4:e127566. doi: 10.1172/jci.insight.127566
58. Reiff SD, Mantel R, Smith LL, Greene JT, Muhowski EM, Fabian CA, et al. The BTK inhibitor ARQ 531 targets ibrutinib-resistant CLL and Richter transformation. *Cancer Discov*. (2018) 8:1300–15. doi: 10.1158/2159-8290.CD-17-1409
59. Gomez EB, Lippincott I, Rosendahl MS, Rothenberg SM, Andrews SW, Brandhuber BJ. Loxo-305, a highly selective and non-covalent next generation BTK inhibitor, inhibits diverse BTK C481 substitution mutations. *Blood*. (2019) 134(Suppl. 1):4644. doi: 10.1182/blood-2019-126114
60. Lemal R, Tournilhac O. State-of-the-art for CAR T-cell therapy for chronic lymphocytic leukemia in 2019. *J Immunother Cancer*. (2019) 7:202. doi: 10.1186/s40425-019-0686-x
61. Ding W, LaPlant BR, Call TG, Parikh SA, Leis JF, He R, et al. Pembrolizumab in patients with CLL and Richter transformation or with relapsed CLL. *Blood*. (2017) 129:3419–27. doi: 10.1182/blood-2017-02-765685
62. Buhimschi AD, Armstrong HA, Toure M, Jaime-Figueroa S, Chen TL, Lehman A, et al. Targeting the C481S ibrutinib-resistance mutation in Bruton's tyrosine kinase using PROTAC-mediated degradation. *Biochemistry*. (2018) 57:3564–75. doi: 10.1021/acs.biochem.8b00391
63. Kapoor I, Li Y, Sharma A, Zhu H, Bodo J, Xu W, et al. Resistance to BTK inhibition by ibrutinib can be overcome by preventing FOXO3a nuclear export and PI3K/AKT activation in B-cell lymphoid malignancies. *Cell Death Dis*. (2019) 10:924. doi: 10.1038/s41419-019-2158-0

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sedlarikova, Petrackova, Papajik, Turcsanyi and Kriegova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.