

**Univerzita Palackého v Olomouci**

**Bakalářská práce**

**Olomouc 2022**

**Johana Richterová**

**Univerzita Palackého v Olomouci**  
**Přírodovědecká fakulta**  
**Katedra buněčné biologie a genetiky**



**Oxford Nanopore sekvenování a jeho  
využití pro analýzu struktury rostlinných  
genomů**

**Bakalářská práce**

**Johana Richterová**

Studijní program: Biologie

Studijní obor: Molekulární a buněčná biologie

Forma studia: Prezenční

**Olomouc 2022**

**Vedoucí práce: Mgr. Eva Hřibová, Ph.D.**

# UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2020/2021

## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: Johana RICHTEROVÁ  
Osobní číslo: R19853  
Studijní program: B1501 Biologie  
Studijní obor: Molekulární a buněčná biologie  
Téma práce: Oxford Nanopore sekvenování a jeho využití pro analýzu struktury rostlinných genomů  
Zadávající katedra: Katedra buněčné biologie a genetiky

### Zásady pro vypracování

8#x2022; Vypracování literární rešerše na zadané téma.  
8#x2022; Izolace vysokomolekulární DNA u vybraných rostlinných druhů.  
8#x2022; Příprava sekvenačních knihoven dlouhých fragmentů a Oxford Nanopore sekvenování.  
8#x2022; Základní analýza sekvenačních dat dostupnými programy (posouzení kvality a kvantity dat).

Rozsah pracovní zprávy:  
Rozsah grafických prací:  
Forma zpracování bakalářské práce: tištěná

### Seznam doporučené literatury:

8#x2022; Amarashinghe et al. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 21:30, 2020.  
8#x2022; Bleidorn C. Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity*, 2015. DOI: 10.1080/14772000.2015.1099575  
8#x2022; Dumshott et al. Oxford Nanopore sequencing: new opportunities for plant genomics? *J Exp Bot.* 71: 5313-5322, 2020.  
8#x2022; Oxford Nanopore Technologies: <https://nanoporetech.com/>  
8#x2022; Saxena et al. Structural variation in plant genomes. *Briefings in Functional Genomics* 138#x201d:296-307, 2014.  
8#x2022; Schiessl et al. The role of genomic structural variation in the genetic improvement of polyploid crops. *Crop J.* 7:127-140, 2019.

Vedoucí bakalářské práce: Mgr. Eva Hřibová, Ph.D.  
ÚEB AV ČR, Sokolovská 6, 772 00 Olomouc

Datum zadání bakalářské práce: 10. února 2021  
Termín odevzdání bakalářské práce: 31. července 2022

LS.

10-02-2021

UNIVERZITA PALACKÉHO V OLOMOUCI  
PŘÍRODOVĚDECKÁ FAKULTA  
KATEDRA DĚDEČNÉ BILOGIE A GENETIKY  
Šlechtitelů 21, 785 02 Olomouc, e-mail: [genetika@upol.cz](mailto:genetika@upol.cz)  
tel.: 585 431 634 901

---

doc. RNDr. Martin Kubala, Ph.D.  
děkan

---

prof. RNDr. Zdeněk Dvořák, DrSc.  
vedoucí katedry

## **Bibliografické údaje**

**Jméno a příjmení autora:** Johana Richterová

**Název práce:** Oxford Nanopore sekvenování a jeho využití pro analýzu struktury rostlinných genomů

**Typ práce:** bakalářská

**Pracoviště:** Katedra buněčné biologie a genetiky, PřF UP v Olomouci

**Vedoucí práce:** Mgr. Eva Hřibová, Ph.D.

**Rok obhajoby práce:** 2022

**Klíčová slova:** sekvenování, vysokomolekulární DNA, izolace DNA

**Počet stran:** 75

**Počet příloh:** 0

**Jazyk:** český

## ABSTRAKT

Rozvoj v oblasti sekvenačních technologií umožnil studovat organizaci a strukturu genomů, sestavovat celogenomové sekvence organismů charakteristických velkým obsahem DNA a identifikovat jednonukleotidové mutace. V dnešní době můžeme sekvenační technologie rozdělit do tří generací dle jejich povahy a výstupu, který poskytují, přičemž tu nejnovější třetí generaci představují metody, které umožňují sekvenování dlouhých molekul. Hlavní předností sekvenačních technologií třetí generace je skutečnost, že poskytují čtení dlouhá až několik desítek kilobází. Mezi metody třetí generace patří sekvenační metoda od společnosti PacBio – „Sekvenování jedné molekuly v reálném čase“ (SMRT) a Oxford Nanopore sekvenování od společnosti Oxford Nanopore Technologies (ONT).

Metoda ONT je založena na technologii nanopórů a obecně lze konstatovat, že produkuje sekvenační čtení, která jsou v průměru dlouhá 25–50 kb, s frakcí čtení dlouhých až několik stovek bází. ONT sekvenační data mají široké uplatnění, ale nejrozšířenější je jejich využití pro *de novo* sestavování genomů a pro determinaci dlouhých repetitivních oblastí. Důležitým krokem vedoucím k získání dlouhých čtení je izolace kvalitní vysokomolekulární DNA. Některé rostlinné druhy obsahují vysoké množství polysacharidů nebo rozličných sekundárních metabolitů a izolace vysokomolekulární DNA je tak obtížná.

Cílem této bakalářské práce byla izolace vysokomolekulární DNA z cizrny beraní (*Cicer arietinum*), která se vyznačuje vyšší přítomností polysacharidů. Izolace byla provedena pomocí dvou přístupů, za použití komerčně dostupného kitu a klasickou metodou založenou na CTAB nebo MATAB izolaci. Izolace buněčných jader cizrny následovaná izolací vysokomolekulární DNA pomocí komerčně dostupného kitu (Nanobind Plant Nuclei Big DNA Kit) se ukázala jako neúspěšná, kdy po deproteinaci jader nedošlo k navázání DNA na magnetické disky. Úspěšná byla izolace DNA pomocí klasického přístupu založeného na izolaci DNA pomocí CTAB metody. Metoda byla prováděna za použití lyofilizovaného/čerstvého rostlinného materiálu a byly testovány různé druhy pufřů. DNA izolovaná z čerstvého vstupního materiálu byla kvalitnější a všechny pufry vedly k získání DNA o téměř stejné velikosti. Získané fragmenty měly délku ~ 48–194 kb. K sekvenování byly vybrány vzorky izolované ze dvou genotypů cizrny, které byly izolovány z čerstvého vstupního materiálu za použití pufřů MATAB. Ačkoliv sekvenování probíhalo na průtokových komůrkách, které na začátku obsahovaly pouze cca 800 aktivních nanopórů, podařilo se ověřit, že DNA měla dostatečnou kvalitu a při sekvenování nedocházelo k rychlé degradaci/ucpávání nanopórů. Také se podařilo získat sekvence s délkou čtení N50 ~ 14–17 kb, kdy nejdelší přečtené sekvence měly délku ~ 112–559 kb.

## **Bibliographical identification**

**Author's first name and surname:** Johana Richterová

**Title:** Application of Oxford Nanopore sequencing for genome structure analysis in plants

**Type of thesis:** bachelor

**Department:** Department of Cell Biology and Genetics, Faculty of Science Palacký  
University Olomouc

**Supervisor:** Mgr. Eva Hřibová, Ph.D.

**The year of presentation:** 2022

**Keywords:** sequencing, high-molecular weight DNA, DNA isolation

**Number of pages:** 75

**Number of appendices:** 0

**Language:** Czech

## SUMMARY

The development in the field of sequencing technologies has made it possible to study the organization and structure of genomes, assemble whole-genome sequences of organisms characteristic for large content of DNA, and identify single-nucleotide mutations. Nowadays, sequencing technologies can be divided into three generations according to their nature and the output they provide, with the latest third generation featuring methods that allow sequencing of long molecules. The main advantage of third-generation sequencing technologies is that they provide reading lengths up to several tens of kilobases long. Third-generation methods include the PacBio sequencing method – " Single molecule real time sequencing " (SMRT) and Oxford Nanopore sequencing from Oxford Nanopore Technologies (ONT).

The ONT method is based on the nanopore technology and generally produces sequencing read lengths that are on average 25–50 kb long, with read lengths fractions up to several hundred bases long. ONT sequencing data are widely used, but most widely used for *de novo* genome assemblies and for determining long repetitive regions. An important step in obtaining long reads is the isolation of high-quality, high-molecular weight DNA. Some plant species contain high levels of polysaccharides or diverse secondary metabolites, and isolating high-molecular weight DNA is then difficult.

The aim of this bachelor thesis was to isolate high-molecular weight DNA from chickpea (*Cicer arietinum*), which is characterised by a higher presence of polysaccharides. The isolation was done using two approaches, using a commercially available kit and a classic method based on CTAB or MATAB isolation. The isolation of chickpea cell nuclei followed by the isolation of high-molecular weight DNA using a commercially available kit (Nanobind Plant Nuclei Big DNA Kit) proved to be unsuccessful, with no DNA binding to magnetic disks after deproteination of the nuclei. DNA isolation was successful applying a classical approach based on DNA isolation using the CTAB method. The method was performed using lyophilised/fresh plant material and various types of buffers were tested. The DNA isolated from the fresh input material was of better quality, and all the buffers resulted in DNA of nearly the same size. The obtained fragments were ~ 48–194 kb in length. For sequencing, samples isolated from two chickpea genotypes were selected and they were isolated from fresh input material using the MATAB buffer. Although sequencing took place on flow cells that initially contained only about 800 active nanopores, it was possible to verify that the DNA was of sufficient quality and the sequencing did not result in rapid degradation/clogging of the nanopores. It was able to obtain sequences with a read length N50 ~ 14–17 kb, when the longest read sequences were ~ 112–559 kb.



Prohlašuji, že jsem bakalářskou práci vypracovala samostatně pod vedením Mgr. Evy Hřibové,  
Ph.D. s použitím uvedených literárních zdrojů.

V Olomouci dne .....

.....

Johana Richterová

### **Poděkování:**

Ráda bych poděkovala vedoucí mé práce Mgr. Evě Hříbové, Ph.D. za odborné vedení, za cenné podnětné rady a za trpělivost a ochotu, kterou mi v průběhu zpracování bakalářské práce věnovala.

## Obsah

1	Úvod.....	1
2	Cíle práce.....	2
3	Literární přehled.....	3
3.1	Struktura jaderných genomů rostlin .....	3
3.1.1	Metody studia genomů .....	4
3.2	Počátky sekvenování .....	5
3.2.1	Maxam-Gilbertova metoda .....	6
3.2.2	Sangerovo sekvenování.....	7
3.3	Metody sekvenování druhé generace .....	9
3.3.1	454 pyrosekvenování.....	11
3.3.2	Ion Torrent.....	13
3.3.3	Illumina .....	15
3.4	Sekvenování dlouhých molekul.....	17
3.4.1	PacBio sekvenování .....	17
3.4.2	Oxford Nanopore sekvenování.....	19
3.5	Využití Oxford Nanopore sekvenování pro analýzu rostlinných genomů .....	22
3.6	Příprava vysokomolekulární DNA .....	23
4	Materiál a metody.....	26
4.1	Biologický materiál .....	26
4.2	Použité chemikálie, soupravy a roztoky .....	26
4.3	Seznam použitých přístrojů a zařízení.....	28
4.4	Použité experimentální a vyhodnocovací postupy .....	29
4.4.1	Izolace vysokomolekulární (HMW) DNA cizrný beraní.....	29
4.4.2	Izolace vysokomolekulární DNA pomocí Nanobind Plant Nuclei Big DNA Kitu .....	30
4.4.3	Izolace vysokomolekulární DNA pomocí CTAB a MATAB.....	31
4.4.4	Pulzní gelová elektroforéza.....	32
4.4.5	Příprava HMW DNA pro přípravu sekvenačních knihoven na Oxford Nanopore sekvenování.....	33
4.4.6	Příprava sekvenační knihovny .....	33
5	Výsledky.....	38
5.1	Izolace HMW DNA pomocí Nanobind Plant Nuclei Big DNA Kitu.....	38
5.2	Izolace HMW DNA pomocí klasické CTAB metody .....	38
5.3	Oxford Nanopore sekvenování .....	42
6	Diskuse.....	45
7	Závěr.....	48
8	Literatura .....	49

## Seznam symbolů a zkratek

<b>ASIC</b>	‘application specific integrated circuit sensors’
<b>BAC</b>	‘bacterial artificial chromosome’
<b>CCD kamera</b>	‘charge-coupled device’
<b>CCS</b>	‘circular consensus sequencing’
<b>CMOS</b>	doplňkový polovodič na bázi kovu a oxidu
<b>ddNTP</b>	dideoxynukleotidy
<b>dNTP</b>	deoxynukleotidy
<b>DSB</b>	dvouvláknové zlomy DNA
<b>emPCR</b>	emulzní polymerázová řetězová reakce
<b>FISH</b>	fluorescenční <i>in situ</i> hybridizace
<b>FRET</b>	fluorescenční (Försterův) rezonanční přenos energie
<b>G-banding</b>	Giemsa banding
<b>HGP</b>	‘Human Genome Project’
<b>HMW DNA</b>	vysokomolekulární DNA
<b>CHEF</b>	„upnuté“ homogenní elektrické pole
<b>ICRISAT</b>	‘International Crops Research Institute for the Semi-Arid Tropics’
<b>ISFET</b>	iontově selektivní tranzistor s efektem pole
<b>ITS1, ITS2</b>	mezerníky, ‘internal transcribe spacer’
<b>LTR</b>	‘long terminal repeat retrotransozons’
<b>ONT</b>	‘Oxford Nanopore Technologies’
<b>PacBio</b>	‘Pacific Biosciences’
<b>PCR</b>	polymerázová řetězová reakce
<b>PFGE</b>	pulzní gelová elektroforéza
<b>PTP</b>	pikotitrační deska
<b>R-banding</b>	Reverzní banding
<b>rDNA</b>	ribozomální DNA
<b>rRNA</b>	ribozomální RNA
<b>SBH</b>	sekvenování hybridizací
<b>SBS</b>	sekvenování syntézou
<b>SGS</b>	sekvenování druhé generace
<b>SMRT</b>	„Sekvenování jedné molekuly v reálném čase“
<b>SNP</b>	‘single nucleotide polymorphism’
<b>SNV</b>	‘single nucleotide variants’
<b>SOLiD</b>	‘Sequencing by Oligonucleotide Ligation and Detection’
<b>SV</b>	strukturní varianty
<b>tRNA</b>	transferová RNA
<b>WGD</b>	duplikace celého genomu
<b>ZMW</b>	‘zero-mode waveguides’

## Seznam obrázků

Obrázek 1: Schéma Maxam-Gilbertova sekvenování.....	7
Obrázek 2: Průběh Sangerova sekvenování.....	9
Obrázek 3: Schéma sekvenování hybridizací.....	10
Obrázek 4: Schéma 454 pyrosekvenování .....	12
Obrázek 5: Schématické znázornění jedné jamky Ion Torrent sekvenačního čipu.....	14
Obrázek 6: Schéma Illumina sekvenování. ....	15
Obrázek 7: Schéma SMRT/PacBio sekvenování .....	18
Obrázek 8: Schéma ONT .....	21
Obrázek 9: Vložení průtokové komůrky do sekvenátoru MinION.....	36
Obrázek 10: Části průtokové komůrky k sekvenátoru MinION .....	37
Obrázek 11: Pulzní gelová elektroforéza vyizolované DNA cizrný beraní .....	39
Obrázek 12: Pulzní gelová elektroforéza DNA cizrný beraní po fragmentaci a přečištění .....	42
Obrázek 13: Histogramy znázorňující rozložení ONT čtení.....	44

## Seznam tabulek

Tabulka 1: Typy sekvenátorů společnosti Illumina .....	16
Tabulka 2: Porovnání sekvenátorů od společnosti ONT.....	20
Tabulka 3: Použité genotypy cizrný beraní.....	26
Tabulka 4: Příprava reakčního mixu k opravě DNA a přípravě konců fragmentů DNA.....	34
Tabulka 5: Příprava reakčního mixu k navázání adaptérů na připravené konce.....	35
Tabulka 6: Příprava sekvenační knihovny .....	36
Tabulka 7: Naměřené hodnoty vzorků DNA pomocí spektrofotometru NanoDrop a fluorometru Qubit.....	40
Tabulka 8: Statistika sekvenování genotypu ICC 1882 .....	43
Tabulka 9: Statistika sekvenování genotypu ICC 8261 .....	43

# 1 Úvod

Rozvoj v genomice nastal díky sekvenačním technologiím, které se v poslední době extrémně rozvinuly. Jedná se o metody, které umožňují určit přesné pořadí nukleotidů v molekule nukleové kyseliny. Sekvenační technologie se dělí do tří generací dle jejich povahy a výstupu, který poskytují, přičemž tu nejnovější třetí generaci představují metody, které umožňují sekvenování dlouhých molekul bez jejich předešlé amplifikace a jejich hlavní předností je, že poskytují dlouhá čtení. Mezi metody třetí generace patří sekvenační metoda od společnosti PacBio – „Sekvenování jedné molekuly v reálném čase“ (SMRT) a Oxford Nanopore sekvenování od společnosti Oxford Nanopore Technologies (ONT).

Oxford Nanopore sekvenování využívá k determinaci nukleotidů v řetězci DNA technologii nanopórů a nevyužívá procesu syntézy DNA. Oxford Nanopore sekvenování se díky svým vlastnostem velice často využívá ke zkoumání velkých sekvenačních změn jako jsou inserce, delece, duplikace, inverze a translokace. Také se díky svým dlouhým čtením, která překlenou celé zkoumané oblasti, využívá k identifikaci struktury a velikosti mnohých repetitivních segmentů. Dalším využitím je *de novo* sestavování prokaryotických a eukaryotických genomů.

K dosažení kvalitních výsledků pomocí technologií sekvenování dlouhých molekul a pro získání dlouhých čtení je důležitým krokem izolace kvalitní vysokomolekulární DNA, která je u rostlin obtížně izolovatelná z důvodu vysokého obsahu sekundárních metabolitů rostlinných buněk a také kvůli přítomnosti pevných buněčných stěn.

## **2 Cíle práce**

Vypracování literární rešerše.

Izolace vysokomolekulární DNA u vybraných rostlinných druhů.

Příprava sekvenačních knihoven dlouhých fragmentů a Oxford Nanopore sekvenování.

Základní analýza sekvenačních dat dostupnými programy (posouzení kvality a kvantity dat).



### 3 Literární přehled

#### 3.1 Struktura jaderných genomů rostlin

Rostlinný genom obsahuje veškerý genetický materiál organismu, který je uložen v jádře a mimojaderných organelách, v případě rostlin v mitochondriích a chloroplastech. Rostlinné genomy jsou charakteristické vysokou variabilitou z hlediska struktury, komplexnosti, heterozygotnosti a velikosti genomu. V mnoha případech velikost genomu, tedy celkový počet párů bází DNA přítomný v jedné kopii haploidního genomu (1C hodnota), neodráží komplexnost rostlin, a to zejména vyšších rostlin. Tento fenomén je známý jako paradox C-hodnoty, nebo dříve také jako záhada hodnoty C (Thomas, 1971; Gregory, 2001). Velikosti jaderných genomů vyšších rostlin se liší více než dva tisíckrát. Největší i nejmenší jaderné genomy byly identifikovány u krytosemenných rostlin. Dosud nejmenší objevený genom má rostlina *Genlisea tuberosa* (61 Mb) a naopak extrémně velký genom byl objeven u rostliny *Paris japonica* (148,852 Mb) (Pellicer *et al.*, 2010; Fleischmann *et al.*, 2014).

Jaderný genom obsahuje kódující DNA, která je reprezentována geny a nekódující DNA, která je z velké části složena z repetitivních sekvencí DNA. Počet kódujících genů je relativně podobný u všech eukaryot, i když se do určité míry může lišit a to především v souvislosti s polyploidií. Polyploidie (duplikace celého genomu – WGD) je označení pro přítomnost několika identických (autopolyploidie) nebo rozdílných (allopolyploidie) kopií chromozomových sad v jednom složeném genomu (Devos *et al.*, 2002). Polyploidie je běžná především u kvetoucích rostlin a kaprad'orostů, ale jen vzácně se vyskytuje u nahosemenných rostlin a mechorostů (Van Straalen *et al.*, 2011). Při nepřítomnosti polyploidie je repetitivní DNA bezpochyby hlavní příčinou zodpovědnou za paradox C-hodnoty.

Repetitivní DNA reprezentují sekvence, které se v genomu opakují stokrát až tisíckrát a tvoří velkou část všech genomů (Devos *et al.*, 2002). Z hlediska uspořádání repetitivních jednotek se repetitivní DNA dělí na dvě kategorie: tandemové repetice a rozptýlené mobilní elementy.

Mobilní elementy zahrnují především různé typy DNA transpozonů a retrotranspozonů. LTR ('long terminal repeat') retrotranspozony jsou nejčastěji se vyskytující elementy ve většině rostlinných genomů a jsou distribuovány po celých chromozomech. Transponovatelné elementy mohou indukovat přestavby genomu a to pomocí transpozice, inserce, delece, chromozomálních zlomů a ektopické rekombinace (Bennetzen, 2002). DNA transpozony mohou sloužit také jako templáty při reparaci dvouvláknových zlomů (DSB) DNA (Vu *et al.*, 2015; Schubert a Vu, 2016). Proto mají rostlinné genomy vyvinutý obranný mechanismus

sloužící k epigenetickému umlčení aktivity transponovatelných elementů pomocí RNA interference nebo pomocí methylace DNA (Hollister a Gaut, 2009; Roessler *et al.*, 2018).

Tandemové repetice se vyznačují upořádáním repetitivních jednotek těsně za sebou (v tandemu) a mezi jejich typické zástupce patří tzv. satelitní DNA, ale také např. mikrosatelity. Satelitní DNA vytváří dlouhé oblasti opakujících se monomerních sekvencí, které se nachází především v centromerických a telomerických oblastech, ale můžeme je najít také v intersticiálních oblastech chromozomů. Tyto repetice mohou tvořit podstatnou část rostlinného genomu (u některých druhů rostlin více než 30 %; *Olea europaea*, *Frittilaria*) a často jsou také druhově specifické (Melters *et al.*, 2013; Barghini *et al.*, 2014; Garrido-Ramos, 2015; Kelly *et al.*, 2015). Povahu tandemové repetice mají také geny kódující ribozomální RNA (5S a 45S rRNA geny). Oproti transponovatelným elementům se tandemové repetice nacházejí na specifických místech v genomu, na diskretních lokusech chromozomů. Oba lokusy rDNA slouží jako užitečné fylogenetické a cytogenetické markery a zejména 35S rDNA se běžně používá jako model pro zkoumání procesů homogenizace a epigenetické genové regulace (Kovařík *et al.*, 2008; Zozomová-Lihová *et al.*, 2014, Navrátilová *et al.* 2022).

### 3.1.1 Metody studia genomů

Varianty na úrovni DNA (v kódujících a nekódujících oblastech) sekvencí se pohybují v rozmezí změn jednoho nukleotidu (tzv. SNP, 'single nucleotide polymorphism') až po velké, několik megabázi dlouhé, mikroskopicky viditelné chromozomální přestavby. K jejich detekci lze využít různých genetických metod a nástrojů v závislosti na typu genomových změn (Gasperskaja a Kučinskas, 2017).

K identifikaci jednonukleotidových změn se využívají výhradně metody založené na sekvenování, v dnešní době především na nových sekvenačních technologiích (např. Illumina sekvenování), zatímco k identifikaci dlouhých chromozomálních přestaveb se kromě sekvenačních technologií dají využít také metody cytogenetické. Tyto jsou založeny na vizualizaci specifických úseků na chromozomech, či v jádrech, a jejich analýze světelnou, respektive fluorescenční či konfokální mikroskopií.

Mezi cytogenetické metody patří tzv. pruhovací techniky, jako např. G-banding a R-banding. Jsou to způsoby, které umožňují identifikaci delece, duplikace nebo přestavby genomického materiálu na základě jedné buňky pomocí specifického proužkování chromozomů (Comings, 1975, 1978; Latt, 1977). G- i R-banding preferenčně barví euchromatin, který obsahuje velké množství genů. Při G-bandingu neboli Giemsa bandingu dochází nejprve k enzymatickému předpůsobení trypsinem. Jedná se o proteolytický enzym,

který natráví chromozomové proteiny. Následně dochází k barvení Giemsou. Po obarvení lze pozorovat vznik pruhů, kdy tmavé pruhy jsou bohaté na adenin a thymin a světlé pruhy na cytosin a guanin (Dutrillaux *et al.*, 1972). R-banding neboli reverzní banding je inverzní vůči G-bandingu. Chromozomy jsou vystaveny specifickým vlivům (inkubaci v horkém SSC pufru), aby došlo k denaturaci DNA (především oblastí bohatých na AT-páry) a následně dochází k barvení Giemsou. Dojde ke vzniku pruhů, které jsou opačné k pruhům G-bandingu, jelikož zdenaturované oblasti se barví méně. Kde jsou tedy G-pruhy světlé, jsou R-pruhy tmavé a naopak (Verma a Lubs, 1975). Tyto techniky se často využívaly hlavně při identifikaci chromozomálních změn v karyotypech člověka a jiných příbuzných druhů, u rostlin však mají nízké uplatnění. Obecně lze říci, že pruhovací techniky se provádí na metafázních chromozomech a mají spíše nízké rozlišení, kdy není možné spolehlivě potvrdit submikroskopické chromozomové aberace.

Pokrok v cytogenetické analýze chromozomálních změn přinesla aplikace tzv. molekulární cytogenetiky, která je založena na přípravě značených úseků genomu a jejich následné hybridizaci na chromozomy či intaktní jádra. Technika využívá fluorescenční *in situ* hybridizaci (FISH) se sondami jednokopiovými nebo repetitivními a umožňuje mnohem přesnější strukturní analýzu jednotlivých lokusů nejen na chromozomech, ale i v interfázních jádrech (Liehr *et al.*, 2009; Gelmann *et al.*, 2013). I přes nedávný značný rozvoj tzv. chromozomálního malování ('chromosome painting'; Jiang, 2019) u rostlinných druhů – tato technika neumožňuje identifikaci přesných míst translokačních událostí, nebo identifikaci jiných typů chromozomálních přestaveb, např. inverzí (Bishop, 2010; Strachan *et al.*, 2011; Šimoníková *et al.*, 2020; Zhao *et al.*, 2021).

Oproti metodám molekulární cytogenetiky, které se využívají pro identifikaci delších strukturních změn v genomu, celogenomové (re)sekvenování se využívá pro identifikaci všech typů variant v genomové DNA (jednonukleotidové změny, ale také delší inserce a delece, respektive transolace; Gasperskaja a Kučinskas, 2017). Pro identifikaci dlouhých strukturních změn v genomech je tak možné využít sekvenování dlouhých molekul (PacBio/SMRT sekvenování nebo Oxford Nanopore sekvenování), které v poslední době prošly rychlým vývojem (Jung *et al.*, 2019).

### **3.2 Počátky sekvenování**

V roce 1953 Watson a Crick publikovali klíčový článek odhalující strukturu dvojité šroubovice DNA (Watson a Crick, 1953). Ve svojí práci zdůraznili její roli při uchování genetické informace a také důležitost determinace nukleotidových bází v řetězci DNA. Ve

stejném roce byly také sekvenovány první biologické molekuly. Pomocí metody separační chromatografie byly sekvenovány dva řetězce inzulinového proteinu (Sanger a Thompson, 1953a; 1953b). Proteiny byly sekvenovány dříve než nukleové kyseliny, ale mnoho principů zůstalo stejných, což připravilo cestu k modernímu sekvenování DNA. Sekvenování proteinů slouží k určení pořadí nukleových kyselin v polynukleotidovém řetězci a dochází k získání sekvence DNA. Díky sekvencím je možné určit, jaký typ genetické informace nese daný segment DNA (Heather a Chain, 2016).

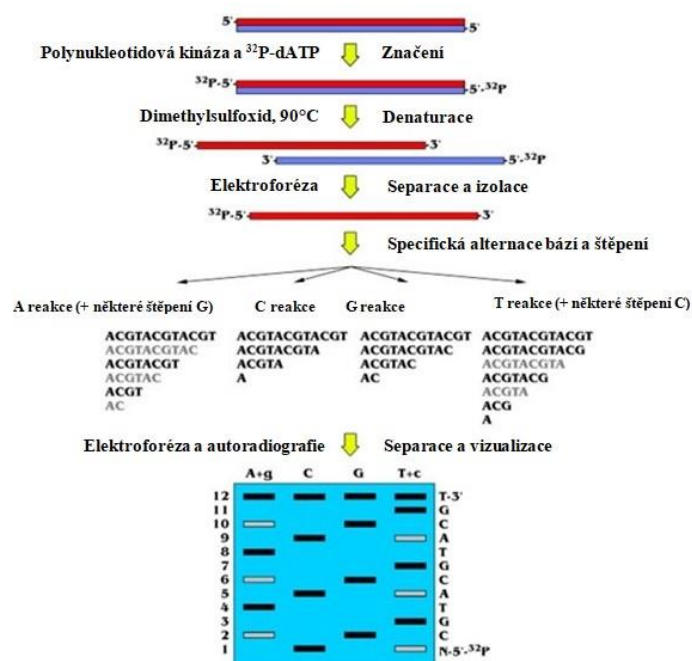
Ribonukleová kyselina (RNA) byla poprvé sekvenována v roce 1965 a jednalo se o 76 bází dlouhou alaninovou tRNA z kvasinky *Saccharomyces cerevisiae*. Časový rozestup mezi sekvenováním RNA a DNA (deoxynukleová kyselina) byl dán dostupností metod štěpení fragmentů RNA, které byly vypracovány ve 40. letech 20. století (Holley *et al.*, 1965). Mezi objevem dvoušroubovice DNA (1953) a prvním sekvenováním DNA (1968) uplynulo patnáct let (Hutchinson, 2007). Prvními sekvenovanými molekulami DNA byly 12 bází dlouhé komplementární konce fága lambda (Wu a Kaiser, 1968). V roce 1975 Sanger a Coulson vynalezli „plus a minus“ metodu k sekvenování DNA a použili ji k determinaci dvou krátkých oblastí jednovláknové DNA fága  $\phi$ X174. Za použití této metody Sanger sestavil 5368 bází velký genom fága  $\phi$ X174, což byl také první sekvenovaný genom (Sanger a Coulson, 1975).

Moderní sekvenování DNA začalo v roce 1977 s rozvojem chemické metody Maxama a Gilberta a dideoxy metody Sangera, Nicklena a Coulsona a také s první kompletní sekvencí DNA (fág  $\phi$ X174), která prokázala, že sekvence DNA poskytuje informace o genetické organizaci. Maxam-Gilbertova metoda byla první technika, která byla široce využívána. Nicméně hlavním průlomem, který změnil pokrok v sekvenování DNA, byla tzv. dideoxy terminační reakce, často nazývána také jako Sangerova metoda (Sanger a Nicklen, 1977). Sangerovo sekvenování se stalo díky své přesnosti, robustnosti a snadnému použití nejběžněji používanou technologií sekvenování DNA, a v následujících letech docházelo k jeho vylepšování (Heather, 2016). Metody sekvenování první generace produkovaly čtení o něco kratší, než je jedna kilobáze (Staden, 1979).

### **3.2.1 Maxam-Gilbertova metoda**

Maxam a Gilbert vyvinuli novou techniku pro sekvenování DNA. Proces určuje nukleotidovou sekvenci pomocí radioaktivně značené molekuly DNA a to tak, že ji specificky štěpí v místech výskytu adeninu, guaninu, cytosinu nebo thyminu, pomocí chemických činidel. DNA molekuly mohou být jednovláknové i dvouvláknové a musí být značené buď na 3' nebo na 5' konci pomocí radioaktivního fosforu ( $^{32}\text{P}$ ). Specifické štěpení v každé bázi vede

k produkci sady radioaktivně značených fragmentů, prodlužujících se od značeného konce vždy o jeden nukleotid (Maxam a Gilbert, 1977). Provedením polyakrylamidové elektroforézy dochází k determinaci délek vzniklých fragmentů. Autoradiograf gelu odhalí vzorec pruhů, ze kterého může být přímo determinována sekvence (Obrázek 1) (Heather a Chain, 2016).



Obrázek 1: Schéma Maxam-Gilbertova sekvenování. Převzato a upraveno z: Dorádo *et al.*, 2019.

Tato metoda umožňovala sekvenování asi 100 bází z konců jakéhokoliv radioaktivně (terminálně) značeného fragmentu DNA. I když byla tato metoda mnoho let populární, zlepšení Sangerovy metody terminace řetězce vedlo k jejímu upadnutí do pozadí a to hlavně kvůli složitosti a používání nebezpečných chemikálií (Heather a Chain, 2016).

### 3.2.2 Sangerovo sekvenování

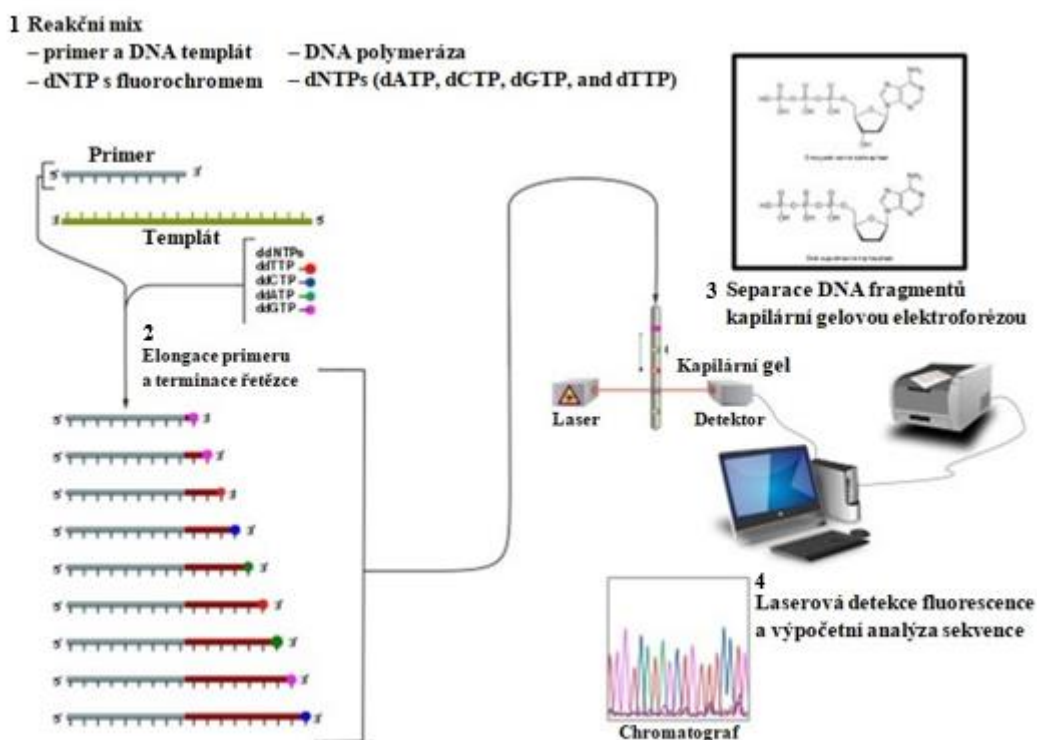
Na počátku 80. let 20. století bylo sekvenování prováděno pomocí původní Sangerovy metody, která zahrnovala čtyři sekvenační reakce. Reakce probíhaly ve čtyřech oddělených zkumavkách za použití radioaktivně značených primerů a produkty byly následně naneseny na polyakrylamidový gel ve čtyřech oddělených řadách. V roce 1985 Smith a Hood syntetizovali fluorescenční DNA primery a použili je k částečné automatizaci Sangerova sekvenování (Smith *et al.*, 1985; Smith *et al.*, 1986). V této metodě se používaly čtyři barviva, rozlišitelná dle jejich emisního spektra fluorescence. Tato barviva byla kovalentně navázána k oligonukleotidovým primerům a zapojená do rozdílných reakcí dideoxy extenzí. Produkty byly pak společně naneseny na polyakrylamidový gel a ke čtení sekvence se využíval fluorescenční (Försterův)

rezonanční přenos energie (FRET = 'Förster resonance energy transfer'). V roce 1987 Prober popsal nový přístup založený na čtyřech ukončujících dideoxynukleotidech, kdy každý nesl fluorescenční barvivo s jiným emisním spektrem (Prober *et al.*, 1987). Různé fluorescenční signály umožnily, aby reakce mohla probíhat v jedné zkumavce. V roce 1986 se společnost Applied Biosystems (založená v roce 1981) začala soustředit na nukleové kyseliny a představila první komerčně dostupný automatizovaný DNA sekvenátor (ABI 370A), který využíval dideoxy terminační reakce a následné analýzy pomocí kapilární elektroforézy. Skenovací systém umožnil sekvenování několika vzorků najednou pomocí automatické identifikace bází (tzv. 'base-calling') na počítači (Tabor a Richardson, 1987).

Pokroky Sangerovy metody byly také urychleny díky celogenomovému sekvenačnímu projektu člověka – Human Genome Project (HGP), jehož cílem byla produkce genetických map, fyzických map a především sestavení kompletní sekvence lidských chromozomů (Giani *et al.*, 2019). Kompletní sekvence byla publikována po 13 letech v roce 2003. Samotný projekt využíval novější dideoxy sekvenátory, které umožňovaly simultánní sekvenování stovek vzorků (Obrázek 2) (Smith, 1986; Ansorge, 2009).

Obrovský pokrok sekvenačních technologií vedl také k prudkému rozvoji analýz organizace a struktury rostlinných genomů. První sekvenování a sestavování celogenomových sekvencí využívalo původní metodu Sangerova sekvenování, která byla časově i pracovně náročná. Nukleotidová sekvence prvního kompletního rostlinného genomu byla publikována v roce 2000 a jednalo se o genom *Arabidopsis thaliana*. *Arabidopsis thaliana* byla prvním rostlinným druhem určeným k *de novo* sekvenování genomu, jelikož byla vhodným modelovým organismem k výzkumu rostlin a měla malý genom (125 megabází) (Arabidopsis Genom Initiative, 2000). Do konce roku 2020 bylo sekvenováno a publikováno 1031 genomů od 788 různých druhů rostlin (včetně poddruhů) a toto číslo neustále stoupá. Kromě genomu *Arabidopsis thaliana* bylo Sangerovo sekvenování využito pro sestavení celogenomových sekvencí genomů rýže (*Oryza sativa*), vinné révy (*Vitis vinifera*), kukuřice (*Zea mays*) či banánovníku (*Musa spp.*) (Goff *et al.*, 2002; Yu *et al.*, 2002; Jaillon *et al.*, 2007; Schnable *et al.*, 2009; D'hont *et al.*, 2012).





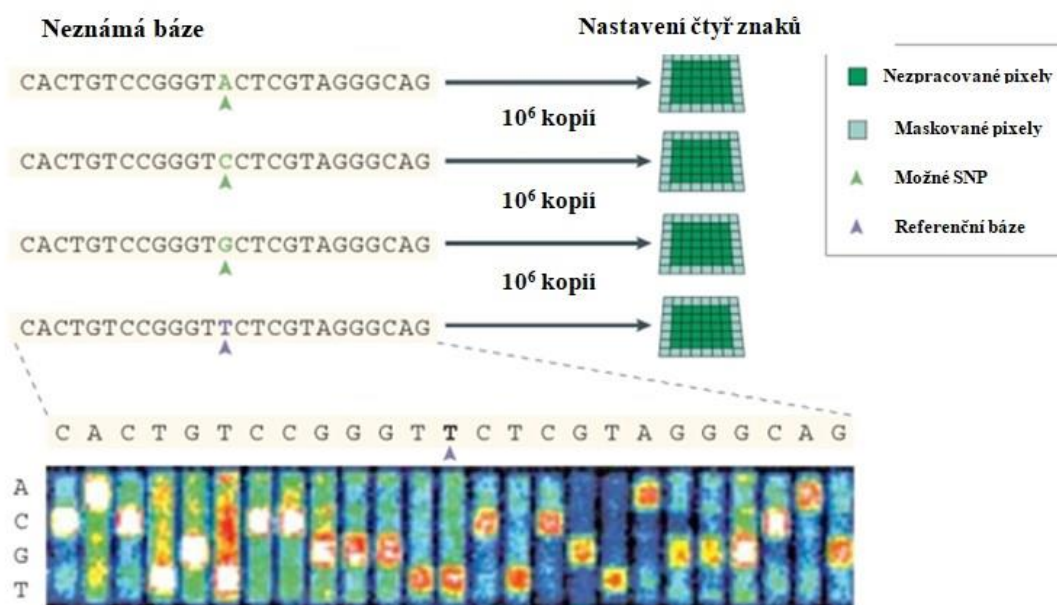
Obrázek 2: Průběh Sangerova sekvenování. 1. Reakce vyžaduje templát DNA, primery, DNA polymerázu, deoxynukleotidy (dNTP-dATP, dCTP, dGTP a dTTP) a různě fluorescenčně značené dideoxynukleotidy (ddNTP-ddATP, ddCTP, ddGTP a ddTTP), které jsou přítomné jen v nízkých koncentracích. 2. Polymeráza katalyzuje elongaci řetězce DNA, pomocí náhodné inkorporace přítomných nukleotidů. Navázáním fluorescenčně značeného ddNTP, dojde k terminaci řetězce (dideoxynukleotidy neobsahují 3'hydroxylovou skupinu, potřebnou k vytvoření fosfodiesterové vazby s dalším nukleotidem). Dochází ke generování fragmentů DNA různých délek, které jsou značené různými fluorofory. 3. Značené DNA fragmenty jsou separovány na základě molekulové hmotnosti pomocí kapilární gelové elektroforézy a jejich fluorescence je detekována pomocí fluorescenčního detektoru a laseru. 4. Vzniklý chromatogram je tvořen na základě detekovaných píků fluorescence, které odpovídají fluorescenčně značeným ddNTP inkorporovaným v jednotlivých fázích do řetězce DNA. Převzato a upraveno z: Let's Talk Science, 2022.

### 3.3 Metody sekvenování druhé generace

Termín „sekvenování druhé generace“ (SGS) označuje sekvenační technologie, které jsou druhou generací po původních Sangerových metodách. Vývoj mnohých technologií druhé generace spustila potřeba vyšší výkonnosti k sekvenování velkých genomů a požadavek na nižší náklady (Slatko *et al.*, 2018).

Většina přístupů SGS se opírá o přípravu knihovny s použitím nativní nebo amplifikované DNA. V klasickém protokolu jsou po fragmentaci DNA a selekci velikosti fragmentů navázány adaptéry na konce každého fragmentu. Poté zpravidla následuje krok amplifikace DNA. Výsledná knihovna je nanesena na průtokovou komůrku a sekvenována v reakci masivně paralelního sekvenování (Head *et al.*, 2014). Sekvenační metody druhé generace mohou být rozděleny do dvou hlavních kategorií, a to na sekvenování hybridizací (SBH) a sekvenování syntézou (SBS).

Sekvenování hybridizací je nadějná alternativa klasických přístupů sekvenování DNA (Hannenhalli *et al.*, 1996). SBH je založeno na použití DNA čipu (nebo DNA array), který obsahuje pevný počet sond. Jednovláknový fragment cílové DNA hybridizuje se sondou (krátký  $k$ -nukleotidový fragment nebo  $k$ -tice) pokud je řetězec cílové DNA k sondě komplementární. V klasickém schématu SBH experimentu čip umístí všechny  $4^k$  sondy o délce  $k$ . Hybridizací DNA fragmentu s tímto čipem může být neznámá cílová DNA testovaná a je určeno sestavení všech jejích  $k$ -tic, tedy její spektrum (Obrázek 3).



Obrázek 3: Schéma sekvenování hybridizací. Pro sekvenování dané báze jsou na DNA čipu přítomny čtyři oligonukleotidy. Oligonukleotidy se liší pouze v centrální pozici, kde se nachází všechny čtyři možné nukleotidy (A, C, G a T). Genotypové údaje každé báze jsou získány diferenciací hybridizací genomové DNA na každou ze čtyř sond. Převzato a upraveno z: Cutler *et al.*, 2001.

SBH poskytuje informace o  $k$ -ticích přítomných v cílové DNA, ale neposkytuje informace o pozicích těchto  $k$ -tic. Což vytváří problém při rekonstrukci cílové DNA z těchto dat (Zhang *et al.*, 2002). Hlavní problém nastává, když několik sekvencí má stejné spektrum. Potom neexistuje způsob, jak determinovat skutečnou sekvenci (Pe'er *et al.*, 2002). Nově vyvinuté metody SBH používají ligaci DNA (Drmanac *et al.*, 2002). SOLiD sekvenování využívá sekvenování ligací, ve kterém dochází ke čtení sekvence templátu DNA komplementárním vázáním sond na vznikající řetězec.

Metody SBS jsou dalším vývojem Sangerova sekvenování, ale bez využití dideoxy terminátorů. Metody se od přístupu Sangerova sekvenování liší hlavně v tom, že využívají mnohem kratší čtení. SBS metody využívají masivně paralelní sekvenování, kdy dochází k sekvenaci miliónů až miliard krátkých sekvencí DNA (50–300 nukleotidů) a následně se získává cílová sekvence na základě identifikace konsenzuální sekvence. Na první pohled se tyto

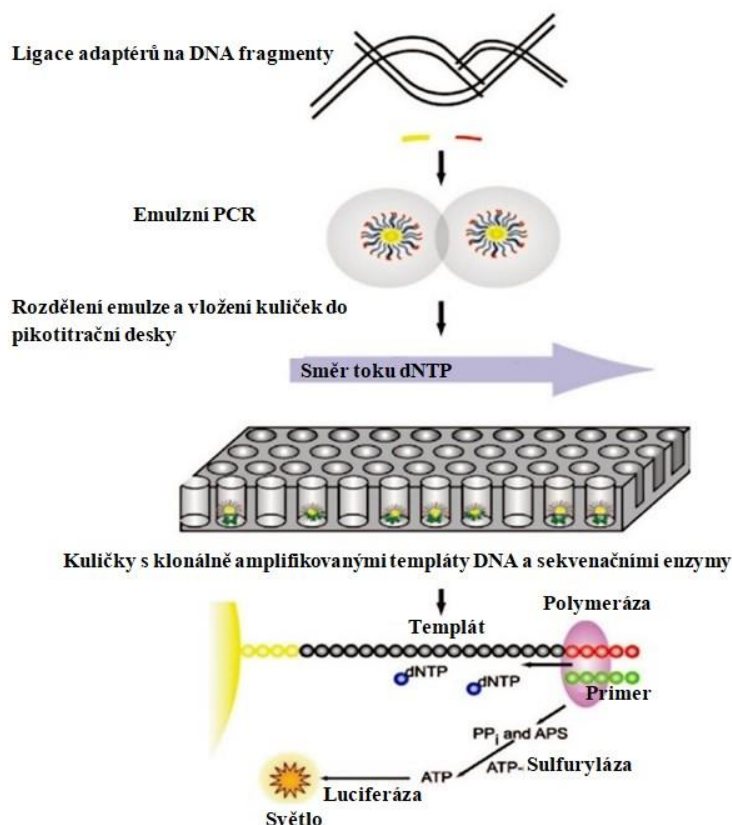


metody mohou zdát drahé, ale reakce probíhají paralelně v objemech nanolitrů, pikolitrů nebo zeptolitrů, a proto cena za osekvenovanou bázi je adekvátně nižší (Slatko *et al.*, 2018). Běžnou strategií SBS metod je použití enzymů, např. DNA polymerázy nebo ligázy k paralelnímu prodlužování mnoha řetězců DNA. Nukleotidy nebo krátké oligonukleotidy jsou poskytovány buď jednou za čas, nebo jsou modifikovány pomocí identifikačních značek, aby mohlo docházet k determinaci typu báze inkorporovaného nukleotidu nebo oligonukleotidu během extenze (Fuller *et al.*, 2009). Mezi SBS metody patří 454 pyrosekvenování, Ion Torrent a Illumina sekvenování (Ambardar *et al.*, 2016).

### **3.3.1 454 pyrosekvenování**

První z masivně paralelních metod, která se stala komerčně dostupnou byla vyvinuta společností 454 Life Sciences a používá radikálně jiný přístup, než je terminace řetězce (Margulies *et al.*, 2005; Slatko *et al.*, 2019).

Je založena na technice pyrosekvenování (Nyrén *et al.*, 1993; Ronaghi *et al.*, 1996), kdy pyrosekvenační technologie 454 spočívá v cyklickém proudění nukleotidových činidel (opakovaně proudící T, A, C nebo G) přes pikotitrační destičku ('picotiter plate' – PTP). Destička se skládá z přibližně jednoho milionu jamek a každá jamka obsahuje nanejvýš jednu kuličku nesoucí kopii jedinečného jednovláknového fragmentu DNA, který má být sekvenován (Slatko *et al.*, 2019). Pokud je přítomný nukleotid komplementární k templátovému vlákně v jamce, je přítomné vlákno DNA v jamce rozšířeno o další nukleotid pomocí polymerázy. Tento krok má za následek reakci, která generuje pozorovatelný světelný signál, který je zaznamenán kamerou (Obrázek 4) (Margulies *et al.*, 2005).



Obrázek 4: Schéma 454 pyrosekvenování. DNA templát je fragmentován, dojde k opravě vzniklých konců a k ligaci adaptérů. Následuje klonální amplifikace emulzní PCR. Po amplifikaci jsou kuličky vloženy do jamek pikotitrační destičky společně se sekvenačními enzymy. Pikotitrační destička funguje jako průtoková komůrka, kde probíhá opakované pyrosekvenování. Inkorporace nukleotidů vede k uvolnění pyrofosfátu (PP<sub>i</sub>) a k luminiscenci lokalizované v jamce. Převzato a upraveno z: Voelkerding *et al.*, 2009.

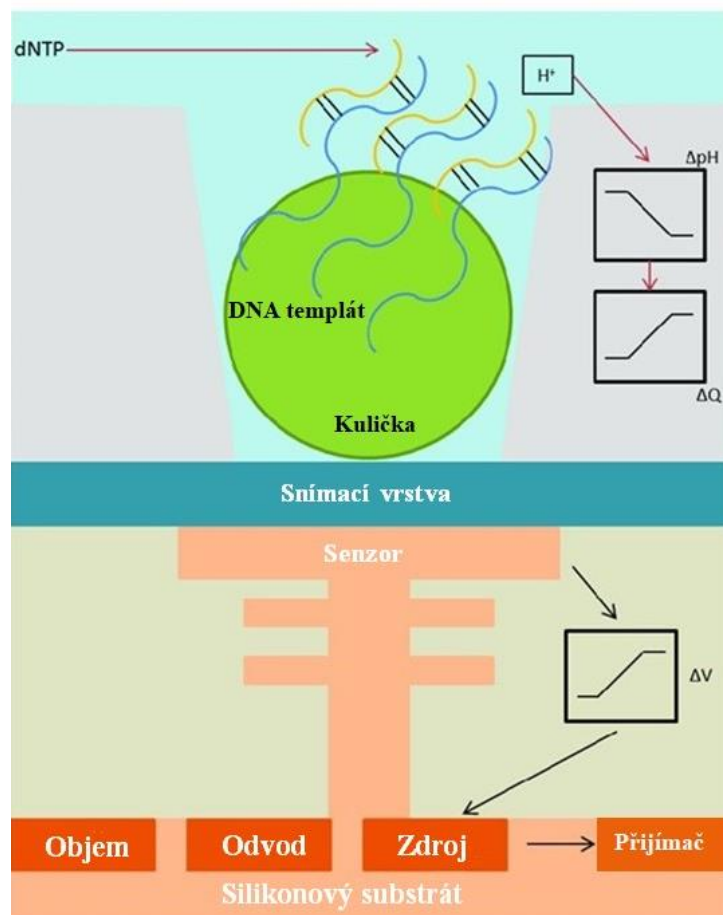
Jednotlivé nukleotidy nejsou značené, místo toho dochází ke čtení sekvence pomocí detekce záblesků chemiluminiscence. Chemiluminiscenci generuje enzym sulfuryláza, který je z molekuly pyrofosfátu uvolněn pokaždé, kdy DNA polymeráza přidá deoxynukleotid na 3' konec rostoucího řetězce. Záblesk chemiluminiscence tedy signalizuje úspěšné zkopírování jedné pozice v molekule templátu. Pokud by byly přidány všechny čtyři deoxynukleotidy najednou, záblesky světla by byly viditelné neustále a nedošlo by k získání užitečné sekvenační informace. Z tohoto důvodu je každý deoxynukleotid přidáván samostatně, jeden po druhém. Do reakční směsi se dále přidává enzym nukleotidáza. Pokud deoxynukleotid není inkorporován do polynukleotidu, dojde pomocí tohoto enzymu k jeho rychlé degradaci předtím, než je přidán další. Tento proces umožňuje sledování pořadí, v jakém jsou deoxynukleotidy inkorporovány do rostoucího řetězce.

454 pyrosekvenování poskytuje průměrnou délku čtení 700 bp s přesností 99,997 % (Liu *et al.*, 2012). 454 sekvenátory se přestaly používat v roce 2013, i když potřebné sloučeniny jsou stále dostupné u několika dodavatelů (Slatko *et al.*, 2019).

### 3.3.2 Ion Torrent

Ion Torrent technologie přímo převádí nukleotidovou sekvenci na digitální informace pomocí polovodičového čipu. U jiných metod sekvenování se užívá detekce nepřímá pomocí laserový skenerů, CCD kamer nebo zdroje světla (Rothberg *et al.*, 2011).

Sekvenční reakce Ion Torrent je prováděna ve velkém množství jamek, které pokrývají polovodičový čip. Polovodičový čip obsahuje miliony pixelů, které přeměňují chemické informace na sekvenační informace. Pro zahájení procesu je DNA fragmentována na fragmenty, které jsou vázány k adaptérům. Fragmenty DNA jsou připojeny ke kuličkám díky komplementárním sekvencím na kuličkách a adaptérech. Fragmenty jsou poté na kuličkách amplifikovány pomocí emulzní PCR (emPCR). Tento proces umožňuje, aby každá z miliónů kuliček měla více kopií jedné sekvence DNA. Kuličky se pak přelévají přes čip obsahující jamky tak, že pouze jedna kulička může vstoupit do každé jamky. Když poté sekvenační činidla, která obsahují vždy pouze jeden ze 4 nukleotidů, protékají přes jamky a dojde k inkorporaci příslušného nukleotidu, uvolní se vodíkový iont (Slatko *et al.*, 2019). Tím se změní pH roztoku, které může být zaznamenáno doplňkovým (komplementárním) polovodičem na bázi kovu a oxidu (CMOS – 'complementary metal-oxide semiconductor') a iontově selektivním tranzistorem s efektem pole (ISFET), který funguje podobně jako pH metr (Obrázek 5).



Obrázek 5: Schématické znázornění jedné jamky Ion Torrent sekvenačního čipu. V jamce je uložena kulička, která obsahuje templát DNA. Když dojde k inkorporaci nukleotidu dojde k uvolnění protonu ( $H^+$ ) a změní se pH jamky ( $\Delta pH$ ). Snímací vrstva detekuje tuto změnu nábojů ( $\Delta Q$ ) a převádí chemický signál na digitální signál ( $\Delta V$ ). Převzato a upraveno z: Juarez P., 2017.

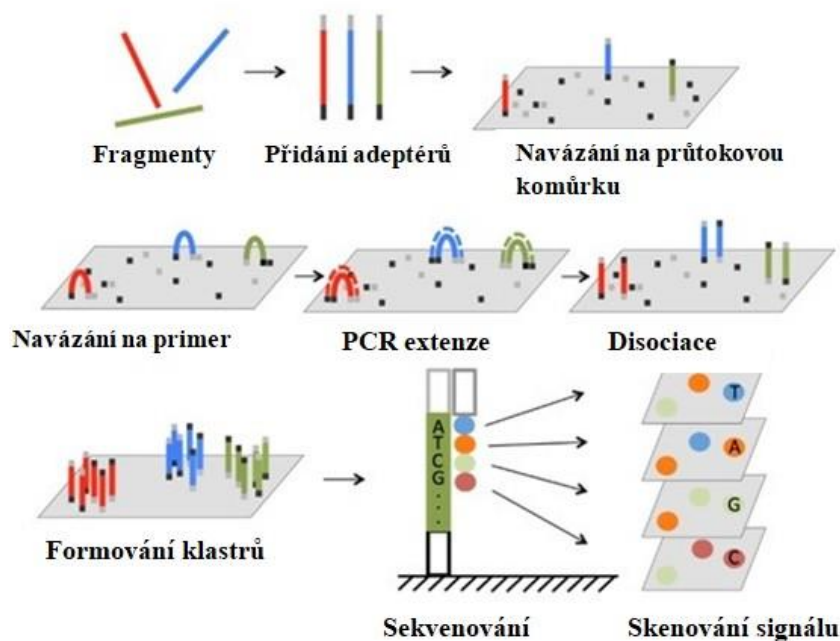
Změna pH detekována senzorem je nedokonale úměrná počtu detekovaných nukleotidů, což způsobuje omezenou přesnost při měření homopolymerů (Goodwin *et al.*, 2016). Dlouhé homopolymerní řetězce stejného nukleotidu jsou tedy obtížně rozeznatelné a sekvenování je zde náchylné k chybám, jelikož repeticity budou inkorporovány v jednom kole, a to povede k vzniku vyššího elektronického signálu kvůli uvolnění velkého počtu vodíkových iontů. Pokud nedojde k inkorporaci žádného nukleotidu nedochází ke změně pH (Slatko *et al.*, 2019).

Ion Torrent platforma nabízí několik druhů čipů, kdy výkonnost těchto čipů se pohybuje od ~ 50 Mb do 15 Gb a kdy doba jednoho běhu se liší od 2 až po 7 hodin, takže je rychlejší než většina ostatních současných platforem (Goodwin *et al.*, 2016). Sekvenování se provádí v reálném čase s čtecí délkou 200 nukleotidů. Technika dále pracuje za nízkých nákladů, jelikož nevyžaduje modifikované báze a využívá jediný polymerázový enzym k elongaci řetězce (Meldrum *et al.*, 2011).

### 3.3.3 Illumina

Technologie Illumina je hlavním zástupcem sekvenování druhé generace (Slatko *et al.*, 2019). Illumina dominuje v odvětví sekvenování s krátkým čtením díky své nízké chybovosti (přesnost je > 99,5 %), kompatibilitě napříč platformami a široké škále platform (Bentley *et al.*, 2008). Sada dostupných přístrojů sahá od MiniSeq s nízkou výkonností až po NovaSeq s ultra vysokou výkonností.

Pouze platformy Illumina umožňují sekvenování párových konců, a to díky tzv. můstkové amplifikaci DNA fragmentů (Obrázek 6). Sekvenování párových konců umožňuje sekvenování fragmentů z obou konců, což vede k získání vyššího počtu čtení (vyššímu pokrytí na genom, tzv. 'coverage') a dat oproti platformám, u kterých dochází k sekvenování pouze z jednoho konce.



Obrázek 6: Schéma Illumina sekvenování. 1. krokem je fragmentace DNA a většinou následná selekce fragmentů vhodných délek. 2. krokem je ligace adaptérů k oběma koncům fragmentu DNA. 3. krokem je hybridizace DNA fragmentů s komplementárním oligonukleotidem, který je kovalentně navázaný na povrch průtokové destičky. Druhý konec každé jednovláknové DNA hybridizuje se sousedícím komplementárním oligonukleotidem. Následně dochází k amplifikaci fragmentů můstkovou PCR. V dalším kole PCR je templátový řetězec a nově nasynthetizovaný komplementární řetězec denaturován, aby mohla začít další amplifikace. Po několika kolech amplifikace dojde ke generování miliónů klastrů v každém pruhu průtokové destičky. Následně je destička připravena k sekvenování syntézou pomocí reverzibilních terminátorů, které jsou značené. Převzato a upraveno z: Biomarker Technologies, 2022.

Proces Illumina sekvenování se skládá ze čtyř kroků, kterým předchází extrakce DNA (Goodwin *et al.*, 2016). Prvním krokem je příprava DNA knihovny. DNA knihovna je zpracovaný materiál vzorku, který slouží jako vstupní materiál pro sekvenační metody. Příprava

knihovny zahrnuje náhodnou fragmentaci genomové DNA, buď mechanicky – pomocí sonikace nebo nebulizace, nebo s využitím náhodných enzymatických metod k získání fragmentů požadované délky. Následně dochází k navázání oligonukleotidových sekvenčních adaptérů k oběma koncům DNA fragmentů.

Druhým krokem je amplifikace „shluků“ (klastřů). Připravená knihovna DNA je nanesena na amplifikační destičku a DNA fragmenty se vážou na primery na jejím povrchu. Neznačené nukleotidové báze a DNA polymeráza jsou přidány k zahájení můstkové amplifikace za použití primerů komplementárních k sekvencím adaptérů.

Třetím krokem je sekvenování syntézou a analýza obrazu. Primery, DNA polymeráza a fluorescenčně značené nukleotidy jsou přidány do amplifikační destičky. Každý nukleotid obsahuje reverzibilní terminátor, který zastavuje syntézu DNA, aby zabránil několikanásobné inkorporaci v jednom kole syntézy (Illumina, 2022). Terminátorové molekuly blokuji ribózovou 3'-OH skupinu nukleotidu a tím zabraňují dalšímu prodlužování řetězce (Brown, 2016). Pouze jedna báze je začleněna v každém kole. Každý ze čtyř nukleotidů má kromě terminátorové molekuly i specifickou značku, která je excitována pomocí laserů, aby došlo k emisi světla charakteristické vlnové délky. Po každém kole je fluorescence detekována pomocí kamery a počítač zaznamenává, která báze byla inkorporována (Illumina, 2022). Jakmile dojde k identifikaci inkorporovaného nukleotidu, dochází k odpojení terminátorové molekuly a fluoroforu, takže zůstane standardní nukleotid. Následně může proběhnout další kolo syntézy (Brown, 2016). Posledním krokem je postsekvenční zpracování a analýza dat (Illumina, 2022).

V sekvenování párových konců jsou generovány dvě krátká čtení z opačných konců, obvykle ~ 500 bp dlouhého fragmentu. Z každého fragmentu DNA lze tedy sekvenovat celkem 50–300 bází dlouhý produkt, v závislosti na zvoleném typu sekvenčního kitu a použitém přístroji (Tabulka 1) (Erguner *et al.*, 2015).

Tabulka 1: Typy sekvenátorů společnosti Illumina. Převzato a upraveno z: Illumina, 2022.

Typ sekvenátoru	iSeq	MiniSeq	MiSeq	NextSeq 550 Series	NextSeq 1000 a 2000	NovaSeq 6000
<b>Maximální výtěžek [Gb]</b>	1,2	7,5	15	120	360	6000
<b>Maximální délka čtení [bp]</b>	2 × 150	2 × 150	2 × 300	2 × 150	2 × 150	2 × 250

### 3.4 Sekvenování dlouhých molekul

Metody třetí generace se v současné době aktivně vyvíjejí. Nukleotidové sekvence jsou sekvenovány na úrovni jedné molekuly, na rozdíl od stávajících metod, které vyžadují rozbití dlouhých vláken DNA na malé fragmenty. Místo sekvenování amplifikovaného templátu, dochází k sekvenování jediného templátu DNA, což vede k minimálnímu použití biochemických látek a k miniaturizaci celého procesu do nanoměřítko a k eliminaci chyb vlivem PCR (Dewey *et al.*, 2012; Heather a Chain, 2016).

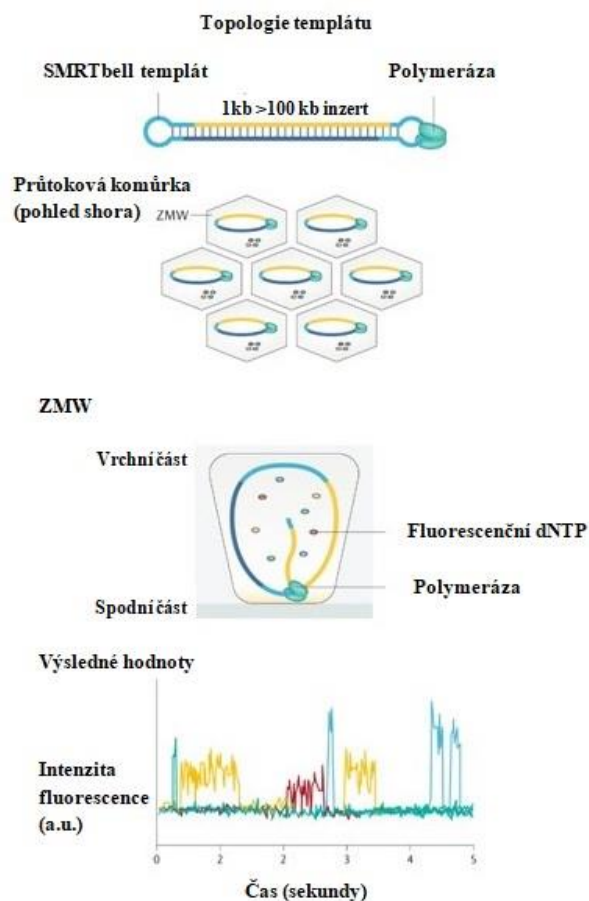
První technologie třetí generace byla komercializována společností Helicos BioSciences a pracovala podobným způsobem jako Illumina, ale bez jakékoliv můstkové amplifikace. Přestože se jednalo o relativně nákladnou a pomalou technologii (produkující relativně krátká čtení) byla to první technologie, která umožnila sekvenování DNA bez předchozí amplifikace a tím omezila vznik souvisejících chyb a odchylek (Schadt *et al.*, 2010; Pareek *et al.*, 2011). Proces byl založen na sekvenování jedné molekuly pomocí cyklické inkorporace terminátorových nukleotidů (Pushkarev *et al.*, 2009). Společnost Helicos vyhlásila bankrot počátkem roku 2012 a vedení v sekvenování třetí generace převzaly jiné společnosti (GenomeWeb, 2012).

Dostupné sekvenční platformy třetí generace jsou „Sekvenování jedné molekuly v reálném čase“ (‘Single Molecule Real-Time’, SMRT) od společnosti Pacific Biosciences (PacBio) a sekvenační platforma Oxford Nanopore Technologies (ONT), které generují čtení obvykle dlouhá několik desítek kilobází (maximálně 150 kb pro PacBio, a 2,3 Mb pro ONT) (Dewey *et al.*, 2012; Payne *et al.*, 2019).

#### 3.4.1 PacBio sekvenování

Oproti sekvenování druhé generace, sekvenování PacBio/SMRT překonává především délku čtení, a to bez nutnosti PCR amplifikace. Templátem je dvouvláknová molekula DNA s nalogovanými vlásenkovými adaptéry, které vytvářejí uzavřený kruh, což umožňuje opakované sekvenování uměle vytvořené kruhové DNA (Travers *et al.*, 2010). Základem PacBio/SMRT sekvenování je nepřerušovaná templátem řízená syntéza prováděná DNA polymerázou za použití čtyř různě fluorescenčně značených nukleotidů (Obrázek 7) (Eid *et al.*, 2009).





Obrázek 7: Schéma SMRT/PacBio sekvenování. DNA je ligována s adaptéry za vzniku kruhové molekuly. Templát a polymeráza jsou imobilizovány na dně komůrky a jsou přidány fluorescenčně značené dNTP. DNA polymeráza syntetizuje nový řetězec a inkorporované fluorescenčně značené dNTP jsou ozářeny světelným pulzem a dojde k excitaci fluoroforu. Emitované světlo z fluoroforu je detekováno kamerou. Sekvence DNA je determinována emisí fluorescence, která se mění a je zaznamenávána v každé ZMW, kdy každá barva odpovídá určité bázi DNA. Převzato a upraveno z: Logsdon *et al.*, 2020.

Fluorescenčně značené nukleotidy mají u SMRT sekvenování molekulu barviva vázanou na fosfátový řetězec nukleotidu, a ne k bázi, jak je tomu u ostatních metod využívajících fluorescenční značení bází. Při inkorporaci nukleotidu pak dochází k přirozenému štěpení fosfátového řetězce DNA polymerázou a uvolnění molekuly barviva z nukleotidu. Po uvolnění se molekula barviva rychle rozptýlí a zanechá zcela přirozený řetězec DNA (Schadt *et al.*, 2010).

Sekvenování probíhá v průtokové komůrce se 150 000 pikolitrovými jamkami. V každé této jamce (označované také jako 'zero-mode waveguides', ZMW) je jedna molekula DNA polymerázy ukotvena ke spodnímu skleněnému povrchu díky interakci biotin-streptavidin, v celkovém objemu ~ 20 zeptolitřů, kdy lze snadno detekovat aktivitu DNA polymerázy při



inkorporaci jednotlivých nukleotidů (Schadt *et al.*, 2010; Mardis, 2013). Nukleotidy (každý typ je označený jiným fluoroforem) jsou na průtokovou komůrku nanoseny v požadované koncentraci, kdy při replikačním procesu dochází při začlenění fluorescenčně značeného nukleotidu k uvolnění fluorescenčního signálu a vyzáření barevného světla. Sekvenační přístroj detekuje tyto záblesky, jejichž barvy odpovídají identitě bází, v reálném čase. Po inkorporaci se signál okamžitě vrátí na výchozí hodnotu a proces se opakuje, přičemž DNA polymeráza nadále inkorporuje několik bází za sekundu (Schadt *et al.*, 2010).

Sekvenování SMRT produkuje v průměru obvykle ~ 10 kb dlouhé čtení, s tím, že délky některých čtení se mohou blížit až k 100 kbp (Chin *et al.*, 2016). Relativně vysoká míra chybovosti původního PacBio/SMRT sekvenování se pohybuje kolem 15 % a je způsobena tím, že v aktivním místě enzymu zůstávají některé nukleotidy dostatečně dlouho, aby došlo k jejich detekci, ale nakonec nedochází k jejich inkorporaci do vlákna DNA (Mardis, 2013).

V nedávné době přišlo PacBio s dalším vylepšením technologie SMRT, tzv. HiFi sekvenováním neboli CCS ('circular consensus sequencing'). Tato metoda generuje vysoce přesné sekvenační data a umožňuje dlouhá čtení s průměrnými délkami 10–25 kb s přesností větší než 99,5 % (Hon *et al.*, 2020). Templátem je kruhová molekula, na kterou nasedá polymeráza, která opakovaně čte fragmenty DNA. Získaná data jsou pak porovnána a je identifikována finální sekvence DNA. Dochází tak k získání vyšší přesnosti, která je srovnatelná s přesností Illumina sekvenování (PacBio, 2022).

### **3.4.2 Oxford Nanopore sekvenování**

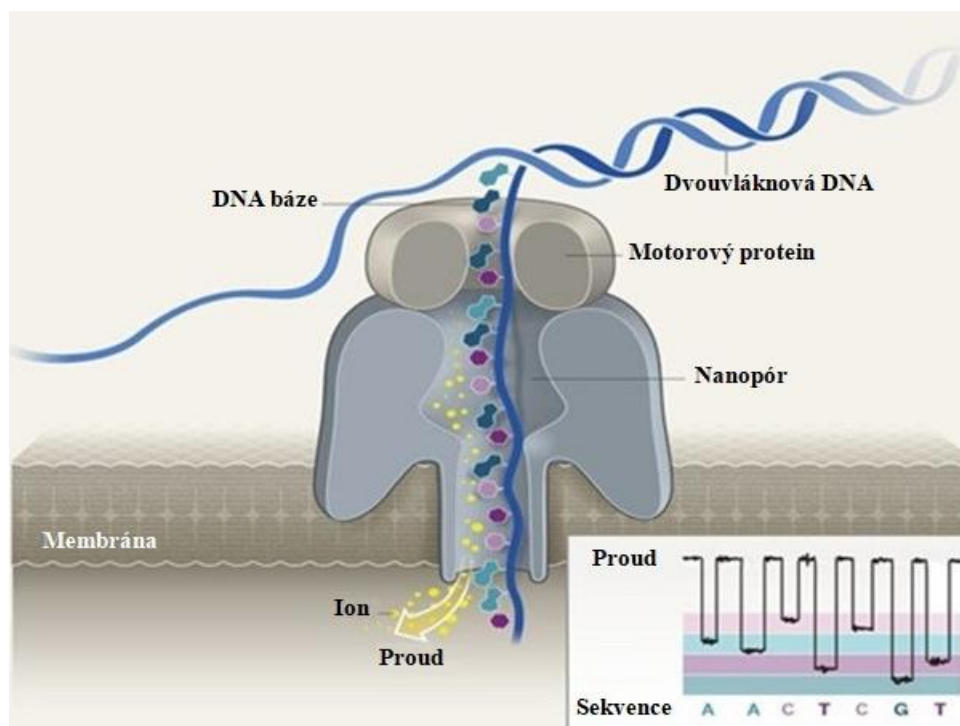
Další technologií sekvenování dlouhých čtení je Oxford Nanopore sekvenování, které bylo vyvinuto společností Oxford Nanopore Technologies (ONT; <https://nanoporetech.com/>). První komerčně dostupný sekvenátor MinION využíval k samotnému sekvenování technologii nanopórů. Později představené sekvenátory společnosti ONT jsou GridION a PromethION, které umožňují paralelní sekvenování až 48 průtokových cel, a také Flongle sekvenátor, což je menší adaptér (velikosti flash klíčenky) umožňující provádění menších experimentů (Tabulka 2).

Tabulka 2: Porovnání sekvenátorů od společnosti ONT. Převzato a upraveno z: Oxford Nanopore Technologies, 2022.

	<b>MinION Mk1B/1C</b>	<b>GridION Mk1</b>	<b>PromethION 2/2 Solo</b>	<b>PromethION 24</b>	<b>PromethION 48</b>
<b>Délka čtení</b>	Závislá na kvalitě vstupní DNA. Obecně, průměrná délka čtení ~ 35–50 kb.				
<b>Počet cel na zařízení</b>	1	5	2	24	48
<b>Počet kanálu na celu</b>	512	512	2675	2675	2675
<b>Možné výstupy [Gb]</b>	10-30-50	10-30-50	100-200-300	100-200-300	100-200-300

Nanopore sekvenování umožňuje uživatelům sekvenování ultra dlouhých molekul DNA nebo RNA v reálném čase. Oxford Nanopore sekvenování na rozdíl od jiných metod nevyužívá zobrazovací zařízení k detekci nukleotidů, což umožnilo zmenšení celého sekvenančního systému. Cena zařízení je také o dost nižší v porovnání s ostatními masivně paralelními sekvenátory. Výzvou v čtecích délkách zde není samotná sekvenační technologie, ale především potřeba extrémně kvalitní vysokomolekulární (HMW) DNA k přípravě knihovny (Jain *et al.*, 2018).

Při ONT sekvenování dochází k aplikaci proudu na nanopór pro řízení iontového toku. Každá molekula vstupující do póru narušuje tok iontů, a proto indukuje charakteristickou a měřitelnou změnu proudu. Před sekvenováním dochází k přípravě sekvenační knihovny. Při její přípravě je na konec řetězce DNA ligován vedoucí adaptér, na kterém se nachází motorový protein. Adaptér přináší DNA k póru a následně se zasune na začátek vnitřní části nanopóru. Jeho funkcí je udržení bází na místě tak dlouho, aby mohlo dojít k jejich elektronické identifikaci. Funkcí motorového proteinu je rozplétání dvouvláknové molekuly tak, aby nanopórem procházel jeden neporušený řetězec DNA (Obrázek 8).



Obrázek 8: Schéma ONT. Nanopóry jsou zakotveny v syntetické membráně uvnitř průtokové cely. Motorový protein rozplétá dvouvláknovou DNA a řídí její translokaci přes nanopór. Na membránu je aplikováno napětí, což pohání negativně nabitou vláknou DNA přes nanopór. Průchod jednotlivých bází přes nanopór způsobuje změnu proudu, která je měřena pomocí senzoru, který je integrován uvnitř průtokové cely. Převzato a upraveno z: MedGadget, 2022.

Motorový protein dále společně s elektrickým proudem pohání negativně nabitou jednovláknovou molekulu DNA přes pór kontrolovanou rychlostí. U DNA se rychlost pohybuje kolem 450 bp za sekundu a u RNA kolem 70 bp za sekundu. Součástí systému je nanopórový senzor, který byl navržen tak, aby se jeho zúžená část nacházela uvnitř nanopóru. Při průchodu DNA nanopórem senzor čte kombinace nukleotidů, které vytvářejí charakteristické změny proudu iontů – změnu napětí, jelikož jednotlivé báze blokují průtok iontů specifickým způsobem. Jakmile řetězec kompletně projde nanopórem, tak se pór stává dostupným pro analýzu dalšího řetězce. Oxford Nanopore sekvenování také umožňuje sekvenování obou vláken dvouvláknové molekuly DNA pomocí vlásenkového adaptéru. Vlášenkový adaptér kovalentně připojuje jedno vlákno k druhému, takže nejdříve dojde k průchodu jednoho vlákna nanopórem a proces se pak opakuje pro komplementární vlákno. Toto umožňuje zpřesnění analýzy (Jain *et al.*, 2016). Jediná průtoková cela obsahuje až 2 048 nanopórů, které jsou kontrolovány ve skupinách po 512 nanopórech pomocí specifických senzorů ASIC ('Application specific integrated circuit sensors').

Hlavní nevýhodou nanopórového sekvenování, stejně jako u většiny metod s čtením dlouhých molekul, je relativně vysoký podíl chyb oproti sekvenování s krátkými čtecími délkami. Současná chybovost se pohybuje od 5 % do 20 % na základě typu analyzovaných

molekul a metod přípravy knihovny. Chyby zahrnují jak inserce, tak delece (Rang *et al.*, 2018). Na rozdíl od SMRT sekvenování (PacBio) to vypadá, že u Nanopore sekvenování se nacházejí systematické chyby (abnormální kolísání proudu nebo nesprávné translace získaných elektrických signálů do sekvence DNA), takže oprava chyb většinou vyžaduje přídatné sekvenování s krátkými čtecími délkami. Aktivní vývoj v této oblasti však vede ke zlepšení těchto problémů (Jain *et al.*, 2018).

Novinkou ONT, která má být dostupná od poloviny května tohoto roku (2022), je kombinace nového typu průtokové komůrky (např. MinION R10.4.1) a nového typu ligačního kitu (Kit 14). Díky této novince bude i ONT technologie poskytovat dlouhá sekvenační čtení s kvalitou více než 99 %, a bude tak činit alternativu pro výše zmíněné HiFi čtení technologie PacBio (Oxford Nanopore Technologies, 2022).

### **3.5 Využití Oxford Nanopore sekvenování pro analýzu rostlinných genomů**

Sekvenování, která generují krátké čtení, jsou vhodná k identifikaci jednonukleotidových variant (SNV, 'single nucleotide variants') a malých inzercí a delecí (tzv. indelů), ale nejsou vhodná k rychlé detekci dlouhých genomových přestaveb (Mahmoud *et al.*, 2019). Mezi delší strukturní varianty (SV) patří inserce, delece, duplikace, inverze nebo translokace, které postihují  $\geq 50$  bp (Mills *et al.*, 2011). Sekvenační technologie dlouhých čtení tak představují velmi užitečný nástroj k identifikaci takto dlouhých strukturních změn (Mitsubishi a Matsumoto, 2020; Sakamoto *et al.*, 2020).

ONT sekvenování dále umožňuje identifikaci struktury a velikosti mnohých repetitivních oblastí v genomech, jelikož dlouhá čtení snáze překlenou segmenty repetitivní DNA a je tak možná přesná determinace délky, složení a počtu repetic. Především je tak možné získat detailnější informace o organizaci a struktuře dlouhých oblastí tandemově uspořádaných repetic, jako jsou např. geny kódující rDNA, především 45S rRNA geny, jejichž jedna jednotka je dlouhá několik kilobází a je složena z dlouhého mezerníku, 18S, 5,8S a 26S rRNA genů, které jsou vzájemně odděleny krátkými transkribovanými mezerníky (ITS1 a ITS2), nebo dalších typů tzv. satelitní DNA. Nedávné studie využívající dlouhých sekvenačních čtení, respektive kombinaci dlouhých sekvenačních čtení s informací z Bionano optické mapy, tak poskytly první ucelený obraz toho, jak jsou jednotlivé opakující se jednotky organizovány v rámci dlouhých tandemově uspořádaných oblastí na chromozomech (Kapustová *et al.*, 2019; Belser *et al.*, 2021; Vondrak *et al.*, 2021; Navrátilová *et al.*, 2022).

Kromě určení velikosti a uspořádání repetitivních jednotek se sekvenování pomocí nanopórů také využívá k detekci epigenetických modifikací nativní DNA. Detekce modifikací,

jako je metylace DNA, je umožněna schopností platformy sekvenovat nativní nukleové kyseliny bez nutnosti použití speciálních chemikálií, které mohou poškozovat DNA (Flusberg *et al.*, 2010). Sekvenování umožňuje rozlišení mezi 5-methylcytosinem a 5-hydroxymethylcytosinem nebo detekci N<sup>6</sup>-methyladenosinu (Rand *et al.*, 2017; Simpson *et al.*, 2017; Sun *et al.*, 2019). Podobně mohou být detekovány různé modifikace bází přítomné v nativních RNA molekulách, jelikož modifikované DNA a RNA báze vedou k charakteristickým změnám v profilech proudu při průchodu nanopórem a z těchto změn můžou být následně modifikace detekovány (Garade *et al.*, 2018).

Dalším důležitým využitím Oxford Nanopore sekvenování je *de novo* sestavování prokaryotických a eukaryotických genomů (Dijk *et al.*, 2018). Zejména u polyploidních organismů s komplexními genomy, jako je např. pšenice (allohexapolyloidní druh), a u regionů s nízkou komplexitou umožňují dlouhá čtení správné sestavování sekvenačních čtení do dlouhých kontinuálních contigů (Schmidt *et al.*, 2017; Schmid *et al.*, 2018; Genova *et al.*, 2019; Kapustová *et al.*, 2019; Shin *et al.*, 2019; Wang *et al.*, 2019; Belser *et al.*, 2021; Mascher *et al.*, 2021; Bredeson *et al.*, 2022). ONT bylo využito k sekvenování malých prokaryotních genomů, např. genomu bakterie *Escherichia coli*, stejně jako k sekvenování velkých a repetitivních rostlinných a živočišných genomů (Loman *et al.*, 2015). Příklady zahrnují lidský genom a rostlinné genomy v rozsahu od malého ~ 119.5 Mbp genomu *Arabidopsis thaliana*, středních genomů *Musa spp.* (523 Mbp) a *Solanum pennellii* (1–1,2 Gb) až po velký 2,53 Gbp genom od *Chrysanthemum nankingense* (Schmidt *et al.*, 2017; Jain *et al.*, 2018; Michael *et al.*, 2018; Song *et al.*, 2018; Belser *et al.*, 2021). Dlouhá čtení generovaná pomocí ONT průtokových komůrek jsou užitečná při ověřování sestavených heterozygotních genomů. Wang *et al.* (2020) sekvenovali a sestavili vysoce heterozygotní genom blahovičnicku (*Eucalyptus*), za použití kombinace dlouhých čtení generovaných z ONT a z dat krátkých čtení od Illuminy.

### 3.6 Příprava vysokomolekulární DNA

Pro získání dlouhých sekvenačních čtení pomocí Oxford Nanopore sekvenování (stejně jako pro PacBio/SMRT technologii) je potřeba izolace neporušené a velmi čisté vysokomolekulární DNA (HMW DNA, 'high-molecular weight DNA'). Obecně je izolace HMW DNA z živočišných tkání relativně jednoduchá. Naproti tomu izolace HMW DNA z rostlinných materiálů je považována za obtížnou, částečně kvůli pevným buněčným stěnám a také díky vysokému obsahu metabolitů rostlinných buněk. Existují různé metody izolace DNA, které jsou založené na různých přístupech. Obecně by se mělo při izolaci HMW DNA vyhnout metodám extrakce DNA za použití centrifugačních kolonek, které jsou hojně

využívány komerčními kity pro izolaci genomové DNA, jelikož tyto metody mají tendenci lámat DNA na fragmenty o velikosti  $< 8$  kb. Vyšší úspěšnost byla získána pomocí metod založených na purifikaci pomocí magnetických kuliček, avšak i při této metodě může dojít k získání DNA s průměrnou délkou  $< 30$  kb. Obecně nejvíc úspěšné metody bývají ty, které jsou založené na extrakci pomocí CTAB (cetyltrimethylammonium bromid) pufru v kombinaci s navíjením DNA na skleněný háček. Tyto přístupy poskytují excelentní kvalitu DNA o vysoké molekulární hmotnosti, ale často vyžadují větší množství rostlinného materiálu, než kity založené na magnetických kuličkách (Jung *et al.*, 2019). Další metoda, kterou je možné využít pro izolaci vysokomolekulární DNA, je založena na izolaci jader a jejich dalším přečištění. Tato metoda se původně používala např. pro izolaci HMW DNA vhodné pro klonování dlouhých inzertů (u rostlin především do vektoru BAC – ‘bacterial artificial chromosome’) a vytváření genomových knihoven.

Kvalita HMW DNA – určení celistvosti/neporušenosti izolovaného vzorku DNA se provádí pomocí separace na pulzní gelové elektroforéze. Standardní agarózová gelová elektroforéza se obecně nedoporučuje, jelikož nedochází k dostatečnému rozdělení molekul DNA větších než 10 kb. Běžně se využívá pulzní elektroforéza nebo CHEF (‘contour-clamped homogeneous electric field’ = „upnuté“ homogenní elektrické pole) a doporučuje se také TapeStation nebo fragmentační analyzátoři ve spojení s kity analyzujícími HMW DNA (Jung *et al.*, 2019).

Dalšími velmi důležitými charakteristikami HMW DNA je její čistota – především přítomnost proteinů, která je nežádoucí. Kvalita se stanovuje velmi často pomocí NanoDrop<sup>TM</sup> spektrofotometru (Thermo Fisher Scientific) a pro přípravu knihoven pro sekvenování dlouhých čtení se doporučuje použít pouze takovou DNA, která má při absorpenci 260 nm / 280 nm (čistotu DNA) hodnotu  $\sim 1,8$ . Hodnotu absorbance může ovlivňovat pH vzorku, přítomnost proteinů, fenolu nebo polyfenolických metabolitů nebo dalších nečistot, které absorbují světlo v blízkosti 280 nm. Absorbance při 260 nm / 230 nm se využívá jako sekundární ukazatel čistoty DNA a měla by být vyšší než hodnota pro A260/280, tedy většinou by se u čisté DNA měla pohybovat v rozmezí hodnot 2,0–2,2. Pokud je poměr menší, je způsoben přítomností látek, které absorbují světlo při 230 nm – 270 nm, např. EDTA, fenol či TRIzol.

Vysoce kvalitní HMW DNA pak může být použita pro přípravu sekvenačních knihoven a následně pro samotné sekvenování pomocí technologií dlouhých čtení – ONT i PacBio. ONT poskytuje různé kity a protokoly k přípravě sekvenační knihovny a jsou optimalizovány pro

velký rozsah přístupů. ONT nabízí tři možnosti přípravy sekvenační knihovny: Rapid Sequencing kit, Ligation Sequencing kit a VolTRAX. U všech tří přístupů dochází k upravení konců DNA fragmentů a k navázání specifických adaptérů na tyto konce, které zajistí samotný průběh sekvenování – navedení a průchod skrze nanopór (Oxford Nanopore Technologies, 2022).



## 4 Materiál a metody

### 4.1 Biologický materiál

V praktické části bakalářské práce byly použity 4 genotypy cizrny beraní (*Cicer arietinum*), které byly vypěstovány ze semen a uchovány ve skleníku. Semena všech genotypů byla získána z Mezinárodního výzkumného institutu plodin semiaridních tropů (International Crops Research Institute for the Semi-Arid Tropics, ICRISAT) sídlícím v Patancheru, Indie.

Tabulka 3: Použité genotypy cizrny beraní.

Rod	Druh	Typ	Kód genotypu	Počet chromozomů
<i>Cicer</i>	<i>arietinum</i>	Kabuli	---	$2n = 2x = 16$
<i>Cicer</i>	<i>arietinum</i>	Kabuli	ICC 8261	$2n = 2x = 16$
<i>Cicer</i>	<i>arietinum</i>	Desi	ICC 1882	$2n = 2x = 16$
<i>Cicer</i>	<i>arietinum</i>	Desi	ICC 283	$2n = 2x = 16$

### 4.2 Použité chemikálie, soupravy a roztoky

#### Použité chemikálie

- β-merkptoethanol (Sigma-Aldrich, kat. č. M3148)
- Bromfenolová modř (Sigma-Aldrich, kat. č. 114405)
- CTAB (Sigma-Aldrich, kat.č. 8.14119)
- Destilovaná voda
- EDTA (Sigma-Aldrich, kat. č. E5134)
- Ethanol (Lach-ner, kat. č. 20025)
- Ethidium bromid (Sigma-Aldrich, kat. č. E8751)
- Glycerol (Sigma-Aldrich, kat. č. G5516)
- Gold Agarose (SeaKem, kat. č. 50150)
- H<sub>3</sub>BO<sub>3</sub> (Sigma-Aldrich, kat. č. B7660)
- HCl (Lach-ner, kat. č. 10033)
- Chloroform (PENTA, kat. č. 17110)
- Isopropanol (Carl Roth, kat. č. 6752.3)
- Izoamylalkohol (Lach-ner, kat. č. 20035)
- Lambda PFG Ladder (BioLabs, kat. č. N0341S)
- MATAB (Sigma-Aldrich, kat.č. M7635)
- Midrange PFG Marker (BioLabs, kat. č. N0342S)
- NaCl (Lach-ner, kat. č. 30093)
- PMSF (Sigma-Aldrich, kat. č. 78830)
- PVP (Sigma-Aldrich, kat. č. P0930)
- Sacharóza (Lach-ner, kat. č. 40135)
- Tekutý dusík
- Tris-base (Sigma-Aldrich, kat. č. T8524)
- Triton X (Sigma-Aldrich, kat. č. X100)
- Xylene cyanol (Sigma-Aldrich, kat. č. 335940)



## **Použité soupravy**

- AMPure XP bead (Beckman Coulter, A63881)
- Ligation Sequencing Kit (Oxford Nanopore technologies, SQK-LSK110)
- Nanobind Plant Nuclei Big DNA Kit (Circulomics, NB-900-801-01)
- NEBNext Companion Module for Oxford Nanopore Technologies (BioLabs, E7180S)
- Qubit<sup>TM</sup> dsDNA BR Assay Kit (Thermo Fisher Scientific, Q32853)

## **Použité roztoky a jejich příprava**

### Tris-HCl:

Rozpustit 12,1 g 1M Tris-base v malém množství destilované vody. Upravit pH přidáním požadovaného množství 15% HCl a doplnit destilovanou vodou na celkový objem 100 ml.

### 0,5M EDTA:

Rozpustit 18,61 g EDTA v 80 ml destilované vody a během rozpouštění upravovat pH pomocí NaOH. Doplnit destilovanou vodou na celkový objem 100 ml.

### 5M NaCl:

Rozpustit 29,22 g NaCl ve 100 ml destilované vody.

### 2x CTAB:

Rozpustit 2 g CTABu v destilované vodě (v 50 ml) za stálého míchání a zahřívání roztoku. Jakmile je CTAB dobře rozpuštěn a zchlazen na pokojovou teplotu, přidat 10 ml 1M roztoku Tris-HCl, 4 ml 0,5M roztoku EDTA a 28 ml 4M roztoku NaCl. Doplnit destilovanou vodou na celkový objem 100 ml.

### 2x CTAB-PVP:

Rozpustit 2 g CTABu v destilované vodě (v 20 ml) za stálého míchání a zahřívání roztoku. Rozpustit 2 g PVP v destilované vodě (v 20 ml) za stálého míchání a zahřívání roztoku. Jakmile je CTAB a PVP dobře rozpuštěno a zchlazeno na pokojovou teplotu, slít dohromady a přidat 10 ml 1M roztoku Tris-HCl, 4 ml 0,5M roztoku EDTA a 28 ml 5M roztoku NaCl. Doplnit destilovanou vodou na celkový objem 100 ml.

### 2x MATAB:

Rozpustit 2 g MATABu v destilované vodě (v 50 ml) za stálého míchání a zahřívání roztoku. Jakmile je MATAB dobře rozpuštěn a zchlazen na pokojovou teplotu, přidat 10 ml 1M roztoku Tris-HCl, 4 ml 0,5M roztoku EDTA a 28 ml 4M roztoku NaCl. Doplnit destilovanou vodou na celkový objem 100 ml.

### 2x MATAB-PVP:

Rozpustit 2 g MATABu v destilované vodě (v 20 ml) za stálého míchání a zahřívání roztoku. Rozpustit 2 g PVP v destilované vodě (v 20 ml) za stálého míchání a zahřívání roztoku. Jakmile je MATAB a PVP dobře rozpuštěno a zchlazeno na pokojovou teplotu, slít dohromady a přidat 10 ml 1M roztoku Tris-HCl, 4 ml 0,5M roztoku EDTA a 28 ml 5M roztoku NaCl. Doplnit destilovanou vodou na celkový objem 100 ml.

#### TE pufr:

Smíchat 1 ml 1M roztoku Tris-HCl s 200  $\mu$ l 0,5M roztoku EDTA a doplnit destilovanou vodou na celkový objem 100 ml.

#### Chloroform:izoamylalkohol (24:1):

Smíchat 96 ml chloroformu s 4 ml izoamylalkoholu.

#### 70% Ethanol:

Smíchat 36,46 ml 96% ethanolu s 13,54 ml destilované vody.

#### 1M sacharóza:

Rozpustit 68,48 g sacharózy ve 200 ml destilované vody.

#### 1x H pufr:

Smíchat 30 ml 10x HB, 150 ml 1M sacharózy, 1,5 ml Tritonu a 5,22 ml PMSF. Doplnit destilovanou vodou na celkový objem 300 ml.

#### 5x TBE pufr:

54 g Tris-base a 27,5 g kyseliny borité a 20 ml 0,5M EDTA rozpustit v destilované vodě a doplnit destilovanou vodou na celkový objem 1 l.

#### 0,05x TBE pufr:

Smíchat 10 ml 5x TBE pufru s 990 ml destilované vody.

#### 6x STOP C:

Smíchat 2 ml 0,5M roztoku EDTA, 1 ml 10% SDS, 4,3 ml 99,9% glycerolu s 5 mg bromfenolové modři a 5 mg xylene cyanolu. Rozpustit a doplnit destilovanou vodou na celkový objem 10 ml.

#### TC pufr:

Do 90 ml destilované vody přidat 5 ml 1M Tris-HCl (pH 7,5), 3,75 ml 2M NaCl, 600  $\mu$ l 1M roztoku  $MgCl_2$  a 100  $\mu$ l 0,1M roztoku  $CaCl_2$ . Doplnit destilovanou vodou do celkového objemu 100 ml.

### **4.3 Seznam použitých přístrojů a zařízení**

- Analytické váhy (Sartorius)
- Centrifuga CR3i (Thermo Electron Corporation)
- Centrifuga MINI (Labnet)
- Centrifuga STAR 60 (VWR)
- Digestoř (Merci)
- DynaMag<sup>TM</sup>-2 (Thermo Fisher Scientific)
- Elektroforetická komůrka k pulzní elektroforéze (Bio-Rad)
- Fluorometr Qubit 3 (Thermo Fisher Scientific)

- Hlubokomrazící mrazák (Thermo Scientific)
- HulaMixér™ Sample Mixer (Thermo Fisher Scientific)
- Chladící zařízení k pulzní elektroforéze (Bio-Rad)
- Injekční stříkačky Micro-Fine Plus 0,5 ml (Becton Dickinson)
- Inkubační třepačka C24 (New Brunswick Scientific)
- Inkubátor Incucell V55 (BMT Medical Technology s.r.o.)
- Lyofilizátor CoolSafe (LaboGene)
- Magnetická míchačka (IKA)
- Mikrovlnná trouba EMN (Zanussi 2015)
- Orbitální třepačka DuoMax 1030 (Heidolph)
- Oscilační mlýn MM 301 (Retsch)
- Průtoková cela (Oxford Nanopore Technologies)
- Pulzní gelová elektroforéza s pumpou (Bio-Rad)
- Sekvenátor MinION Mk1B (Oxford Nanopore Technologies)
- Spektrofotometr NanoDrop ND – 1000 (Thermo Fisher Scientific)
- Termocycler C1000 Touch (Bio-Rad)
- Thermomixer C (Eppendorf)
- UV transiluminátor InGenius (Syngene)
- Vodní lázeň SM 30 (Grant)
- Vortex (Heidolph)
- Výrobník ledu MF 26 (Scotsman)

#### **4.4 Použité experimentální a vyhodnocovací postupy**

##### **4.4.1 Izolace vysokomolekulární (HMW) DNA cizrný beraní**

Pro získání vysoce kvalitní vysokomolekulární DNA cizrný bylo využito dvou přístupů: 1) izolace jader následovaná izolací HMW DNA pomocí komerčně dostupného kitu Nanobind Plant Nuclei Big DNA Kit (Circulomics), využívajícím magnetické disky; 2) klasický přístup založený na izolaci DNA pomocí CTAB metody.

Cílem praktické části bylo zjistit, zda oba protokoly poskytnou srovnatelné výsledky z hlediska kvality (délka fragmentů a čistota HMW DNA) a kvantitivy vyzolované HMW DNA.

Pro první pokusy izolace HMW DNA byly využity nově lyofilizované listy cizrný a v případě, že použitý protokol poskytl pozitivní výsledky, byla metoda použita na čerstvý

rostlinný materiál homogenizovaný v tekutém dusíku.

#### **4.4.2 Izolace vysokomolekulární DNA pomocí Nanobind Plant Nuclei Big DNA Kitu**

Před samotným použitím Nanobind Plant Nuclei Big DNA Kitu je nutné vyizolovat buněčná jádra, která jsou vstupním materiálem pro izolaci vysokomolekulární DNA pomocí Nanobind terčků.

##### 1. Izolace buněčných jader cizrny

0,5 g lyofilizovaných listů cizrny bylo zhomogenizováno pomocí oscilačního mlýnu (7 min, 27 Hz). Homogenizovaný rostlinný materiál byl přenesen do kádinky s 35 ml pufru H, která byla umístěna na ledu. Směs byla opatrně, ale důsledně promíchána a následně inkubována 5 minut při pokojové teplotě. Po inkubaci byly ze směsi odstraněny hrubé nečistoty (větší kusy rostlinných pletiv) přefiltrováním přes nylonové sítko (Miracloth, s velikostí pórů 48  $\mu\text{m}$ ) do nových sterilních kyvet. Po vyvážení kyvetek (přidáním ledového pufru H) byla směs centrifugována 15 minut při 200 g, 4 °C, za použití pomalého brzdění (level 2) tak, aby byly ze vzorku eliminovány intaktní buňky a buněčné zbytky. Supernatant byl následně opatrně vylit (nebo odpipetován), aby nedošlo k utrnutí peletu. K peletu bylo opatrně napipetováno 35 ml vychlazeného pufru H, a směs byla zhomogenizována – opatrným promícháním, pomalým otáčením kyvety. Dle potřeby byl předešlý krok opakován (přibližně 4x), s cílem získat pelet světlé barvy, což indikuje odstranění převážné části plastidové DNA. Následně byl pelet resuspendován v 20 ml ledového TC pufru. Roztok byl centrifugován 5 minut při 400 g, 4 °C (pomalé brzdění). Supernatant byl opatrně odstraněn a peletka buněčných jader byla využita pro izolaci HWM DNA s využitím Nanobind terčků (Nanobind Plant Nuclei Big DNA Kit od firmy Circulomics; viz níže).

##### 2. Izolace HMW DNA z vyizolovaných jader cizrny kitem Nanobind Plant Nuclei Big DNA Kit (Circulomics)

Ke vzorku izolovaných jader cizrny bylo přidáno 80  $\mu\text{l}$  pufru PL1 a 30  $\mu\text{l}$  pufru CS. Směs byla opatrně promíchána pipetou se zastřiženou špičkou a inkubována v termomixéru při 55 °C, po dobu 15 minut, za stálého míchání (900 rpm). Pro odstranění RNA, bylo ke vzorku přidáno 20  $\mu\text{l}$  RNásy A a vzorek byl inkubován 3 minuty při pokojové teplotě. Po inkubaci s RNásou byl vzorek centrifugován 5 minut při 16 000 g za pokojové teploty. Supernatant byl přepipetován do nové sterilní mikrozkuhavky (typ Eppendorf), kam byly přidány Nanobind disky a 120  $\mu\text{l}$  isopropanolu. Vzorek byl promíchán opatrným převrácením mikrozkuhavky a umístěn na HulaMixér na 20 minut k inkubaci při pokojové teplotě. Mikrozkuhavka se vzorkem byla následně umístěna do magnetického stojánku a byl odstraněn supernatant.

Ke vzorku bylo přidáno 500 µl pufru PW1 a vzorek byl promíchán obrácením mikrozkušavky, umístěn na magnetický stojánek a supernatant byl opět odpipetován. Tento krok byl opakován a zkušavka byla krátce centrifugována k odstranění zbytkového supernatantu. Ke vzorku bylo následně přidáno 150 µl pufru EB+ a následovala inkubace při pokojové teplotě po dobu 10 minut. Takto získaná DNA byla přepipetována do nové mikrozkušavky. Pro kontrolu vyizolované DNA byla provedena pulzní elektroforéza.

#### **4.4.3 Izolace vysokomolekulární DNA pomocí CTAB a MATAB**

Vysokomolekulární DNA byla izolována z listů cizrný beraní (*Cicer arietinum*), které byly lyofilizované nebo čerstvé, a izolace byla prováděna za použití čtyř různých pufrů (MATAB, MATAB-PVP, CTAB a CTAB-PVP) ke zjištění, jaký postup vede k získání nejkvalitnější vysokomolekulární DNA.

Izolace byla provedena dle následujícího postupu:

Bylo naváženo 0,5 g lyofilizovaného rostlinného materiálu nebo 6 g čerstvého (zamraženého) rostlinného materiálu. Lyofilizovaný rostlinný materiál byl homogenizován pomocí oscilačního mlýnu (7 minut, 27 Hz) a čerstvý rostlinný materiál byl homogenizován pomocí tekutého dusíku a třecí misky s tloučkem.

Podrcený rostlinný materiál byl následně přidán do kádinky s 15 ml 2x MATAB/2x MATAB-PVP (nebo 2x CTAB/2x CTAB-PVP) pufru, 300 µl β-merkptoethanolu a opatrně homogenizován mícháním na ledu. Homogenizovaný roztok byl přelit do falkony a inkubován ve vodní lázni 30 minut při 65 °C, za občasného promíchání obrácením falkony.

Po inkubaci byl přidán stejný objem (15 ml) směsi chloroform:izoamylalkohol (24:1), roztok byl opatrně promíchán obrácením falkony a centrifugován 10 minut při 4 000 g, 4 °C. Po centrifugaci byla vodní fáze (horní) odebrána do nové falkony a byl přidán stejný objem směsi chloroform:izoamylalkoholu (24:1). Roztok byl opatrně promíchán obrácením falkony a opět centrifugován 10 minut při 4 000 g, 4 °C. Tento krok byl opakován, dokud nebylo rozhraní čisté.

Po dosažení požadované čistoty byla odebrána vodní fáze po 1 ml do nových 2ml mikrozkušavek a k vodní fázi byl přidán 1 ml směsi chloroform:izoamylalkoholu (24:1). Roztok byl opět centrifugován 10 minut při 4 000 g, 4 °C. Tento krok byl opět opakován, dokud nebylo rozhraní čisté.

Po dosažení požadované čistoty rozhraní dvou fází, byla vodní fáze ze všech mikrozkušavek přepipetována do nové 12ml zkušavky a po stěně bylo přidáno

0,6 objemu isopropanolu. Po promíchání roztoku překlápěním zkumavky byla pozorována srážející se DNA, která byla buďto namotána na skleněnou kličku, anebo centrifugována 3 minuty při 1500 g, 4 °C, aby došlo k usazení vysrážené DNA na dně zkumavky. Supernatant byl po centrifugaci vylit a pelet byl opláchnut v 70% ledovém ethanolu (2 x 5 minut, při pokojové teplotě).

Po oschnutí byl pelet přes noc rozpouštěn v TE pufru (pH 8) na třepačce při pokojové teplotě. Objem pufru byl zvolen dle velikosti peletu (obvykle 200–500 µl).

Kvalita a kvantita vyizolované DNA byla zkontrolována pomocí spektrofotometru NanoDrop, fluorometru Qubit a pulzní gelové elektroforézy.

#### **4.4.4 Pulzní gelová elektroforéza**

Kvalita a kvantita vyizolované HMW DNA pomocí dvou různých přístupů byla zkontrolována pomocí pulzní gelové elektroforézy (PFGE), jelikož umožňuje separaci molekul DNA větších než 15 kb.

##### Příprava 1,2% agarózového gelu v 0,05x TBE pufru:

Do Erlenmeyerovy baňky bylo naváženo 3,6 g agarózy s nízkým bodem tání (Gold Agarose). Ve skleněném válci bylo odměřeno 300 ml 0,05x TBE pufru a pufr byl následně přilít do baňky k agaróze. Roztok byl krouživými pohyby promíchán a vložen do mikrovlnné trouby, kde byl ohříván, dokud nedošlo k varu a k úplnému rozpuštění agarózy v TBE pufru. Baňka byla opatrně vyjmuta z mikrovlnné trouby a chlazena pod tekoucí vodou, dokud roztok neměl požadovanou teplotu. Do připravené nalévací vany s hřebínkem byl následně přelit rozehřátý agar tak, aby došlo ke vzniku co nejmenšího množství bublinek. Případné vzniklé bublinky byly odstraněny špičkou od pipety. Po 30 minutách gel zatuhnul a byl vložen do elektroforetické vaničky naplněné 0,05x TBE pufrem. 5 µl vzorků DNA bylo smícháno s 5 µl STOP-C. Do první a poslední jamky byl vložen marker molekulové hmotnosti (Midrange PFG Marker, Lambda PFG Ladder) a do ostatních jamek bylo nanášeno 10 µl směsi vzorků s 5 µl STOP-C. Program PFGE byl následující: proud 6 V; doba přepnutí (tzv. 'switch time') 1–50 s; úhel 120°; a celková doba běhu separace 14 hodin při 14 °C. Po ukončení elektroforetické separace byl gel vložen do ethidium bromidu na 20 minut, poté opláchnut v destilované vodě a vizualizován pomocí transiluminátoru.

#### **4.4.5 Příprava HMW DNA pro přípravu sekvenačních knihoven na Oxford Nanopore sekvenování**

Na základě výsledků z PFGE bylo zjištěno, že vyizolovaná vysokomolekulární DNA byla příliš dlouhá, a tak byla provedena fragmentace pomalým přetahováním přes injekční stříkačku Micro-fine plus (Becton Dickinson). Podle PFGE byla DNA cizrný přes stříkačku přetažena 15–20krát. Poté byla přečištěna pomocí AMPure XP kuliček (Beckman Coulter) a opět analyzována pomocí PFGE.

##### Přečištění HMW DNA pomocí AMPure XP kuliček:

Ke 160 µl vzorku vyizolované HMW DNA cizrný bylo přidáno ustríženou špičkou 160 µl magnetických kuliček AMPure XP (Beckman Coulter). Mikrozkušavky se vzorky byly následně vloženy do HulaMixéru a inkubovány 10 minut při pokojové teplotě, za stálého míchání. Po inkubaci byly mikrozkušavky umístěny na magnetický stojánek a třikrát promyty 1 ml 70% ledového ethanolu. Po odpipetování ethanolu bylo ke kuličkám s navázanou DNA cizrný přidáno 200 µl TE pufru, roztok byl jemně promíchán poklepáním na zkušavky, nebo pomalým pipetováním ustríženou špičkou. Následně byly mikrozkušavky inkubovány 30 minut při 37 °C, kdy došlo k separaci DNA z kuliček. Po inkubaci byly mikrozkušavky opět vloženy na magnetický stojánek a vzorek (supernatant s přečištěnou DNA) zbavený kuliček byl napipetován do nových mikrozkušavek. Ke kontrole velikosti fragmentů byla následně provedena pulzní gelová elektroforéza.

#### **4.4.6 Příprava sekvenační knihovny**

Sekvenační knihovna byla připravena ze tří vzorků vyizolované vysokomolekulární DNA a to z genotypu ICC 8261 (fragmentace 15x), ICC 8261 (fragmentace 20x) a z genotypu ICC 1882 (fragmentace 15x). Příprava sekvenační knihovny byla prováděna za použití ligačního sekvenačního kitu (Ligation Sequencing Kit, Oxford Nanopore Technologies).

##### Příprava vzorků HMW DNA k přípravě sekvenační knihovny:

Koncentrace vzorků byla nejprve stanovena za použití Qubit<sup>TM</sup> dsDNA BR Assay kitu a vyhodnocena na fluorometru Qubit. Dle změřené koncentrace DNA byl vypočítán objem DNA potřebný pro reakci (2500 ng DNA/reakce). Objem DNA byl následně doplněn destilovanou vodou na požadovaný objem 47 µl. Všechny potřebné komponenty z kitu byly připraveny dle instrukcí výrobce.

DNA CS (DCS) byl rozmrazen při pokojové teplotě, centrifugován, promíchán pipetováním a umístěn na led. Ostatní činidla byla také umístěna na led. Reakční mix byl



připraven na ledu do tenkostěnné PCR zkumavky dle Tabulky 4.

Tabulka 4: Příprava reakčního mixu k opravě DNA a přípravě konců fragmentů DNA.

Činidlo	Objem [μl]
DNA CS	1
DNA + H <sub>2</sub> O	47
NEBNext FFPE DNA Repair Buffer	3,5
NEBNext FFPE DNA Repair Mix	2
Ultra II End-prep reaction buffer	3,5
Ultra II End-prep enzyme mix	3
Celkem	60

Připravený reakční mix byl jemně promíchán a krátce centrifugován. Vzorek byl vložen do termocykleru a inkubován 5 minut při 20 °C a následně 5 minut při 65 °C, aby došlo k opravě DNA a k přípravě konců fragmentů DNA, aby na ně v dalším kroku mohly být navázány adaptéry.

Adapter Mix F (AMX-F) a Quick T4 Ligase byly krátce centrifugovány a umístěny na led. Ligation Buffer (LNB) byl rozmrazen při pokojové teplotě, krátce centrifugován a promíchán pipetováním. Po rozmrazení a promíchání byl okamžitě umístěn na led. Elution Buffer (EB) byl rozmrazen při pokojové teplotě, vortexován, krátce centrifugován a následně umístěn na led. K zachování fragmentů DNA o velikosti 3 kb nebo větší byl použit a rozmrazen za pokojové teploty Long Fragment Buffer (LFB). Následně byl vortexován, krátce centrifugován a umístěn na led. V nové mikrozkuhavce byl z připravených činidel připraven reakční mix 2, aby mohlo dojít k navázání adaptérů k připraveným koncům fragmentů DNA (Tabulka 5).



Tabulka 5: Příprava reakčního mixu k navázání adaptérů na připravené konce.

Činidlo	Objem [μl]
Vzorek DNA z předešlého kroku	60
Ligation Buffer (LNB)	25
NEBNext Quick T4 DNA ligase	10
Adapter Mix F (AMX-F)	5
Celkem	100

Připravený mix byl promíchán, krátce centrifugován a následně inkubován přibližně 8 minut při pokojové teplotě.

#### Přečištění vzorku pomocí AMPure XP kuliček:

Po inkubaci bylo k směsi přidáno 150 μl resuspendovaných AMPure XP kuliček k odstranění fragmentů DNA nevhodných délek. Směs s kuličkami byla inkubována na HulaMixéru 5 minut při pokojové teplotě, aby došlo k navázání DNA. Po inkubaci a centrifugaci byl vzorek umístěn na magnetický stojánek a byl odpipetován supernatant. Následně byly kuličky promyty přidáním 250 μl Long Fragment Buffer (LFB), který slouží k zachování fragmentů DNA o velikosti 3 kb nebo větší, a resuspendovány. Po stočení byla mikrozkušavka vrácena na magnetický stojánek, aby došlo k vzniku peletu a mohl být odpipetován supernatant, ve kterém zůstaly fragmenty DNA nevhodných délek. Celý tento krok byl opakován ještě jednou. Po odpipetování supernatantu byl vzorek stočen a opět umístěn na magnetický stojánek, aby mohl být odpipetován zbytkový supernatant. Pelet byl sušen 30 sekund a následně odebrán z magnetického stojánku a resuspendován v 15 μl Elution Buffer (EB). Vzorek byl krátce centrifugován a následně inkubován 10 minut při 37 °C, aby došlo k uvolnění fragmentů DNA z kuliček do EB pufru. Následně byl vzorek opět vložen do magnetického stojánku a 15 μl eluátu s DNA bylo odebráno a napipetováno do nové mikrozkušavky. Kuličky byly vyhozeny. Ke kvantifikaci vzorku byl použit 1 μl vzorku a množství DNA bylo stanoveno pomocí fluorometru Qubit.

#### Příprava sekvenční knihovny a její nanesení na průtokovou komůrku:

Sequencing Buffer II (SBII), Loading Beads II (LBII), Flush Tether (FLT) byly rozpuštěny za pokojové teploty, vortexovány a centrifugovány. K přípravě tzv. priming mixu pro průtokovou komůrku bylo přidáno 30 μl FLT pufru přímo do zkumavky s FB pufrům.

Sekvenátor MinION byl otevřen a byla do něj vložena průtoková komůrka (viz Obrázek 9). Byl otevřen kryt Priming portu a k odstranění případných bublin bylo z Priming portu odsáto pár mikrolitrů roztoku nastavením automatické pipety (pro objem 1000  $\mu\text{l}$ ) na 200  $\mu\text{l}$ , vložení špičky do Priming portu a upravováním jejího objemu. Objem byl upravován, dokud se ve špičce neobjevilo malé množství kapaliny.



Obrázek 9: Vložení průtokové komůrky do sekvenátoru MinION. Převzato z: Oxford Nanopore Technologies, 2022.

Dle stanovené koncentrace vzorku DNA pomocí fluorometru Qubit byl vypočítán potřebný objem DNA pro samotné sekvenování – tak aby se ve vzorku nacházelo 650 ng DNA v celkovém objemu 12  $\mu\text{l}$ . Přes Priming port bylo do průtokové komůrky naneseno 800  $\mu\text{l}$  předem připraveného priming mixu a během následujících pěti minut byla připravena sekvenační knihovna dle Tabulky 6.

Tabulka 6: Příprava sekvenační knihovny.

Činidlo	Objem [ $\mu\text{l}$ ]
Sequencing Buffer II (SBII)	37,5
Loading Beads II (LBII)	25,5
DNA knihovna	12
Celkem	75

Po přípravě sekvenační knihovny byl otevřen SpotON sample port cover a do Priming portu bylo naneseno zbylých 200  $\mu\text{l}$  priming mixu. Jednotlivé části průtokové komůrky jsou ukázány na Obrázku 10. Následně bylo po kapkách naneseno 75  $\mu\text{l}$  sekvenační knihovny do SpotON sample portu. SpotON sample cover a Priming port byly pak uzavřeny a bylo uzavřeno také víko sekvenátoru MinION, který byl následně připraven k sekvenování.



Obrázek 10: Části průtokové komůrky k sekvenátoru MinION. Převzato a upraveno z: Oxford Nanopore Technologies, 2022.

## 5 Výsledky

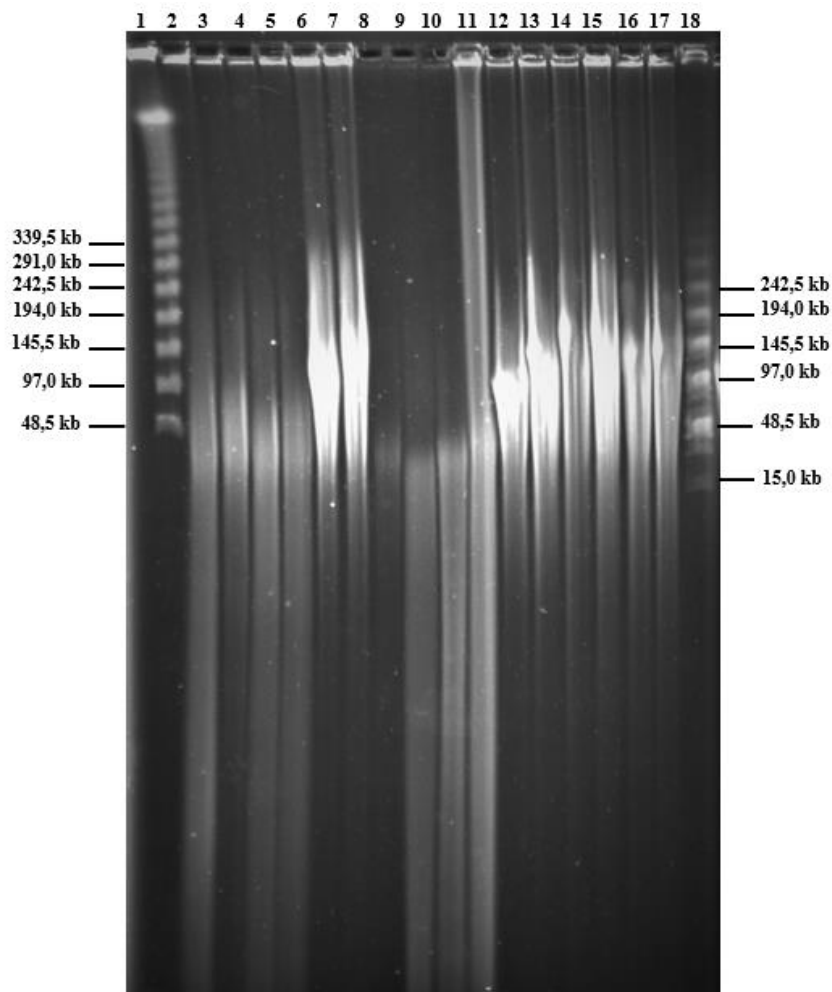
Cílem experimentální části bakalářské práce byla optimalizace protokolu sloužícího k izolaci vysokomolekulární DNA z rostlin, která je potřebná pro Oxford Nanopore sekvenování. Vysokomolekulární DNA byla izolována z cizrny beraní (*Cicer Arietinum*) za použití dvou různých přístupů. Prvním přístupem byla izolace jader a následná izolace vysokomolekulární DNA za použití komerčně dostupného kitu Nanobind Plant Nuclei Big DNA Kit (Circulomics) využívajícím magnetické disky. A druhým přístupem byl klasický přístup za použití CTAB (MATAB) metody. Výsledky izolace byly kontrolovány pomocí pulzní elektroforézy a získaná HMW DNA byla následně použita k přípravě sekvenční knihovny pro Oxford Nanopore sekvenování. Následně bylo provedeno samotné sekvenování jako důkaz, že daný postup funguje.

### 5.1 Izolace HMW DNA pomocí Nanobind Plant Nuclei Big DNA Kitu

Vyizolovaná DNA pomocí komerčně dostupného kitu byla podrobena pulzní elektroforéze. Elektroforetogram gelu ukázal, že přístup za využití Nanobind Plant Nuclei Big DNA Kitu byl neúspěšný – nedošlo ke vzniku produktu. Izolace jader, která předcházela izolaci HMW DNA pomocí kitu byla úspěšná, došlo ke vzniku peletky. Nedošlo tedy k navázání DNA na disky. I přesto, že pokus s čerstvě lyofilizovaným materiálem byl neúspěšný, proces byl opakován i s čerstvým rostlinným materiálem homogenizovaným za použití tekutého dusíku. Bohužel i tento pokus nebyl úspěšný, nedošlo tedy k navázání DNA na magnetické disky.

### 5.2 Izolace HMW DNA pomocí klasické CTAB metody

Izolace HMW DNA za použití klasické metody CTAB byla nejprve optimalizována na lyofilizovaném vstupním materiálu, který byl homogenizován pomocí oscilačního mlýnu. Cílem bylo porovnat, který pufr poskytne lepší výtěžnost a kvalitu vysokomolekulární DNA, aby následně tento pufr mohl být využit k izolaci za použití čerstvého vstupního materiálu. Úspěšnost izolace HMW DNA z lyofilizovaného rostlinného materiálu byla zkontrolována pomocí PFGE. Výsledek pulzní elektroforézy ukázal, že izolace DNA byla v tomto případě úspěšná a bylo zjištěno, že všechny testované pufrы (MATAB, MATAB-PVP, CTAB, CTAB-PVP) poskytly DNA o stejné (velmi podobné) velikosti (viz Obrázek 11). Jelikož se ukázalo, že použitý postup funguje, tak byl proces opakován pro čerstvý rostlinný materiál, který byl homogenizován pomocí tekutého dusíku. Výsledky pulzní elektroforézy porovnávající izolace HMW DNA za použití lyofilizovaného nebo čerstvého rostlinného materiálu a za použití různých pufrů můžeme vidět na Obrázku 11.



Obrázek 11: Pulzní gelová elektroforéza vyizolované DNA cizrný beraní.

Legenda:

1) Lambda PFG Ladder; 2) DNA izolovaná pomocí MATAB pufru z lyofilizovaných listů cizrný beraní; 3) DNA izolovaná pomocí CTAB pufru z lyofilizovaných listů cizrný beraní; 4) DNA izolovaná pomocí CTAB-PVP pufru z lyofilizovaných listů cizrný beraní; 5) DNA izolovaná pomocí MATAB-PVP pufru z lyofilizovaných listů cizrný beraní; 6) a 7) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní; 8) DNA izolovaná pomocí CTAB pufru z lyofilizovaných listů cizrný beraní; 9) DNA izolovaná pomocí CTAB-PVP pufru z lyofilizovaných listů cizrný beraní, genotypu ICC 8261; 10) DNA izolovaná pomocí CTAB pufru z lyofilizovaných listů cizrný beraní, genotypu ICC 8261; 11) DNA izolovaná pomocí MATAB-PVP pufru z lyofilizovaných listů cizrný beraní, genotypu ICC 8261; 12) a 13) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní, genotypu ICC 8261; 14) a 15) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní, genotypu ICC 1882; 16) a 17) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní, genotypu ICC 283; 18) Midrange PFG Marker.

Elektroforetogram pulzní elektroforézy ukázal, že izolace HMW DNA proběhla v pořádku u obou postupů, ale k získání viditelně kvalitnější vysokomolekulární DNA vede izolace z čerstvého vstupního materiálu. Kvalita vyizolované DNA byla také zkontrolována pomocí spektrofotometru NanoDrop a fluorometru Qubit (Tabulka 7). Všechny vzorky DNA měly poměr A260/A280 vyšší než 1,8 (takže nebyly kontaminované proteiny). Vzorky izolované z čerstvého materiálu měly vyšší koncentraci DNA než vzorky z lyofilizovaného materiálu. U všech vzorků byl poměr A260/230 vyšší než 1,8.

Tabulka 7: Naměřené hodnoty vzorků DNA pomocí spektrofotometru NanoDrop a fluorometru Qubit.

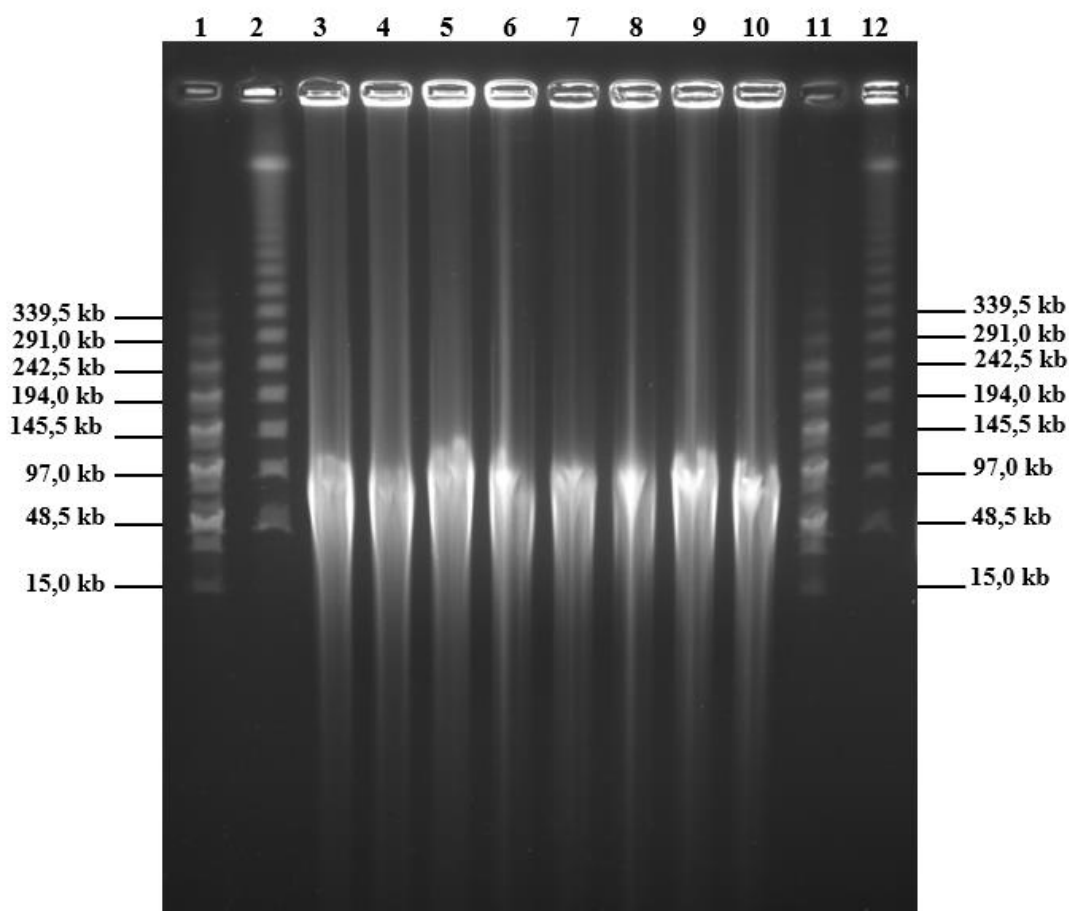
Genotyp	Vstupní materiál	Pufr	Koncentrace [ng/μl]	NanoDrop		Qubit
				A260/A280	A260/A230	Koncentrace [ng/μl]
-	Lyofilizovaný	MATAB	1 864,3	2,18	2,37	61,8
-	Lyofilizovaný	CTAB	1 730,6	2,16	2,35	138
-	Lyofilizovaný	CTAB-PVP	2 273,1	2,21	2,22	177
-	Lyofilizovaný	MATAB-PVP	1 875,8	2,15	2,13	80,4
-	Čerstvý	MATAB	4 967,5	2,15	2,09	208
-	Čerstvý	MATAB	3 800, 2	2,14	2,19	112

Tabulka 7 (pokračování): Naměřené hodnoty vzorků DNA pomocí spektrofotometru NanoDrop a fluorometru Qubit.

Genotyp	Vstupní materiál	Pufr	Konzentrace [ng/μl]	NanoDrop		Qubit
				A260/A280	A260/A230	Konzentrace [ng/μl]
ICC 8261	Lyofilizovaný	CTAB-PVP	2 600,0	2,06	1,99	238
ICC 8261	Lyofilizovaný	CTAB	5 743,6	2,19	2,35	195
ICC 8261	Lyofilizovaný	MATAB-PVP	6 275,4	1,90	2,00	240
ICC 8261	Čerstvý	MATAB	3 585,3	2,12	2,18	140
ICC 8261	Čerstvý	MATAB	3 391,3	2,12	2,19	244
ICC 1882	Čerstvý	MATAB	4 230,0	2,11	2,14	122
ICC 1882	Čerstvý	MATAB	4 312,1	2,12	2,13	141
ICC 283	Čerstvý	MATAB	5 633,5	2,19	2,17	204
ICC 283	Čerstvý	MATAB	6 061,3	2,16	2,18	150

Výsledky analýzy velikosti HMW DNA cizrný beraní ukázaly, že získaná DNA je příliš dlouhá pro ONT sekvenování a mohla by tak způsobit velmi rychlé ucpávání nanopórů a negativně tak ovlivnit výtěžnost sekvenační reakce. V předkládané bakalářské práci nebylo využito speciálního kitu pro přípravu extrémně dlouhých čtení, a proto byla provedena fragmentace HMW DNA a její následné přečištění. Ke kontrole, zda fragmentace vedla k získání fragmentů DNA optimálních délek pro přípravu ONT sekvenační knihovny, byla opět provedena pulzní elektroforéza (Obrázek 12), která ukázala, že „jemná“ fragmentace DNA vedla k získání HMW DNA vhodné pro ONT sekvenování, kdy převážná většina fragmentů

měla délku 48–97 kb, jak je patrné z Obrázku 12.



Obrázek 12: Pulzní gelová elektroforéza DNA cizrný beraní po fragmentaci a přečištění.

Legenda: 1) Midrange PFG Marker; 2) Lambda PFG Ladder; 3) a 4) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní; 5) a 6) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní, genotypu ICC 8261; 7) a 8) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní, genotypu ICC 1882; 9) a 10) DNA izolovaná pomocí MATAB pufru z čerstvých listů cizrný beraní, genotypu ICC 283; 11) Midrange PFG Marker; 12) Lambda PFG Ladder.

### 5.3 Oxford Nanopore sekvenování

HMW DNA byla následně použita pro přípravu sekvenačních knihoven a samotné ONT sekvenování. Cílem sekvenace bylo zjistit, zda je získaná DNA (čistota a délka fragmentů DNA) vhodná pro Oxford Nanopore sekvenování a zda DNA neobsahuje některé jiné sekundární metabolity, které by mohly negativně ovlivnit průběh sekvenování, např. vysokou rychlostí ucpávání pórů.

Na základě výsledků analýzy kvality a velikosti vyizolované DNA cizrný beraní byly pro přípravu sekvenační knihovny a následné Oxford Nanopore sekvenování vybrány tři vzorky



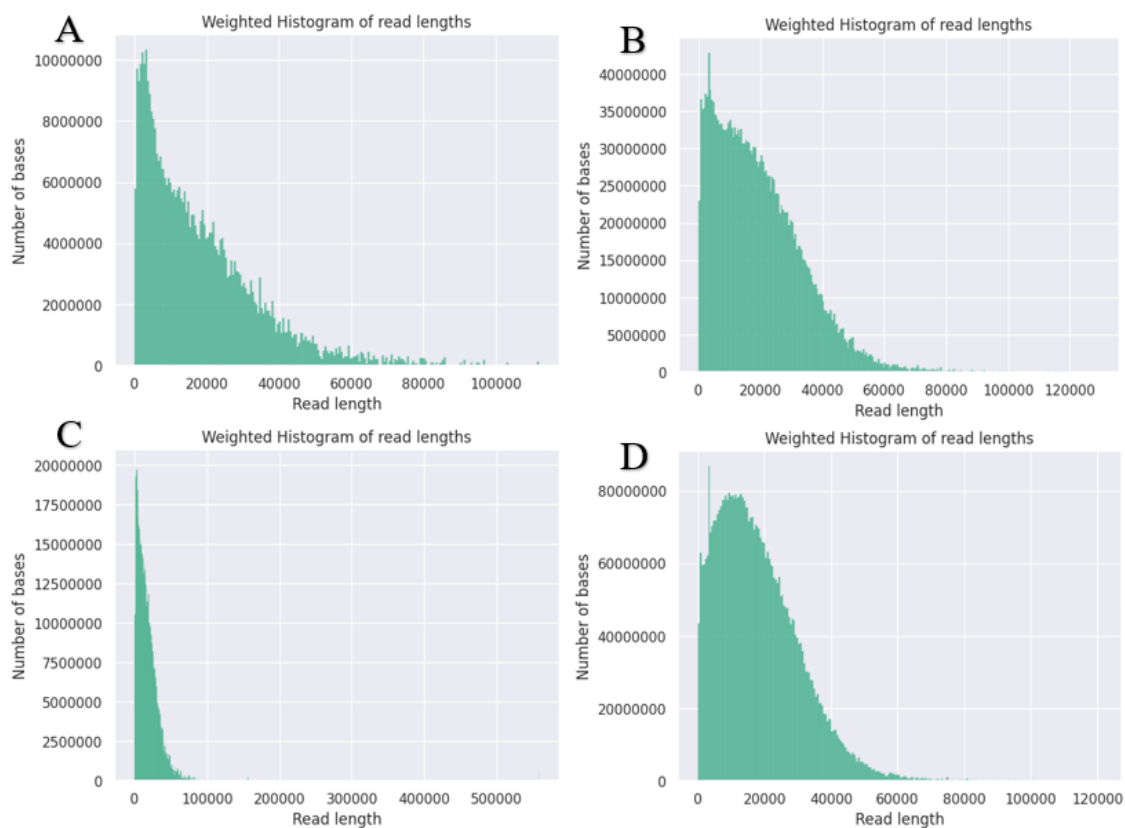
DNA, izolované z genotypů ICC 8261 (2x) a ICC 1882. Jednalo se o vzorky izolované z čerstvého vstupního materiálu za použití pufru MATAB. K přípravě sekvenační knihovny byl využit ligační kit (Ligation Sequencing Kit, Oxford Nanopore Technologies), který je optimalizován pro dlouhé čtení. K sekvenování byly využity sekvenační komůrky, které obsahovaly ~ 400–800 aktivních nanopórů. Sekvenování probíhalo 48 hodin pro vzorek s genotypem ICC 1882 a 1 den 22 hodin pro vzorky s genotypem ICC 8261. Analýza získaných čtení byla provedena pomocí programu Nanoplot. Shrnutí dat získaných ze sekvenování vzorku cizrný genotypu ICC 1882 můžeme vidět v Tabulce 8 a shrnutí dat pro vzorky cizrný genotypu ICC 8261 v Tabulce 9. Histogramy znázorňující rozložení ONT čtení můžeme vidět na Obrázku 13.

Tabulka 8: Statistika sekvenování genotypu ICC 1882.

ICC 1882	Fail	Pass
Průměrná délka čtení	4 958,9 báží	6 492,0 báží
Průměrná kvalita čtení	6,8	11,1
Počet čtení	82 911	349 538
N50 délky čtení	14 232 báží	17 047 báží
Nejdelší čtení	112 004 báží	129 431 báží
Celkem báží	411 149 928	2 269 184 919

Tabulka 9: Statistika sekvenování genotypu ICC 8261.

ICC 8261	Fail	Pass
Průměrná délka čtení	5 242,3 báží	7 197,0 báží
Průměrná kvalita čtení	6,7	11,0
Počet čtení	166 700	650 783
N50 délky čtení	13 881 báží	16 191 báží
Nejdelší čtení	559 039 báží	121 183 báží
Celkem báží	873 899 453	4 683 654 001



Obrázek 13: Histogramy znázorňující rozložení ONT čtení.

Legenda: A. Histogram pro vzorek DNA cizrný beraní genotypu ICC 1882 (fail); B. Histogram pro vzorek DNA cizrný beraní genotypu ICC 1882 (pass); C. Histogram pro vzorek DNA cizrný beraní genotypu ICC 8261 (fail); D. Histogram pro vzorek DNA cizrný beraní genotypu ICC 8261 (pass).

## 6 Diskuse

Oxford Nanopore sekvenování je technika sekvenování dlouhých molekul, která může produkovat ultra-dlouhá čtení, která překonávají problémy spojené s použitím krátkých čtení, jako je charakterizace dlouhých strukturních variant a vysoce repetitivních oblastí. ONT identifikuje báze a modifikace nativní DNA a RNA pomocí detekce změn signálů iontového proudu, které charakteristicky odpovídají každému nukleotidu procházejícímu přes nanopór.

Hlavní oblastí, ve které se ONT sekvenování uplatňuje je *de novo* sestavování genomů, kde dlouhá čtení usnadňují generování dlouhých, souvislých contigů. Teoreticky ONT sekvenování může produkovat neomezeně dlouhé sekvenační čtení, prakticky je však délka čtení omezena délkou vyizolovaných DNA fragmentů. Izolace HMW DNA je relativně rutinní pro bakterie a savčí buněčné kultury, nicméně u rostlin přítomnost buněčných stěn, které se skládají z polysacharidů a ligninu a široký rozptyl specializovaných metabolitů, zahrnujících fenoly, terpeny, alkaloidy a flavonoidy, může stěžovat a bránit izolaci vysoce-kvalitní nebo/a vysokomolekulární DNA. Přítomnost těchto látek v rostlinných pletivech, které se používají pro izolaci může negativně ovlivnit nejen délku vyizolovaných fragmentů, ale především kvalitu DNA, která je nutná pro samotné sekvenování pomocí nanopórů. Pokud jsou tyto látky přítomny během přípravy sekvenační knihovny nebo sekvenování, tyto nečistoty často vedou k rychlé degradaci nanopórů a mohou negativně ovlivnit výnos sekvenační reakce (Healey *et al.*, 2014).

U rostlin se pro izolaci vysokomolekulární DNA nejčastěji využívá metod založených na izolaci jader a jejich následném uchování v agarózových bločcích, kde je jaderná DNA deproteinizována s cílem získat HMW DNA. Tento přístup se dříve používal např. pro izolaci DNA pro přípravu knihoven dlouhých inzertů (např. Peterson *et al.*, 2000; Ming *et al.*, 2001; Lin *et al.*, 2011; Figueira *et al.*, 2012; Zhang *et al.*, 2016). Velmi často využívanou metodou izolace HMW DNA pro sekvenační technologie třetí generace je modifikovaná CTAB extrakce následovaná dalším přečištěním přes kolonu nebo pomocí magnetických kuliček (Belser *et al.*, 2018; Deschamps *et al.*, 2018; Michael *et al.*, 2018). Výhodou tohoto přístupu izolace vysokomolekulární DNA z rostlin je použití extrakčního pufu CTAB (nebo jeho alternativy MATAB), který usnadňuje separaci polysacharidů z DNA (Saghai-Marooof *et al.*, 1984). Další výhodou je získání většího množství DNA z relativně malého vstupního materiálu v porovnání s přístupem založeném na izolaci jader.

V praktické části bakalářské práce byla vysokomolekulární DNA vhodná pro Oxford Nanopore sekvenování (a SMRT/PacBio sekvenování) izolována pomocí jednoduchého

protokolu založeného na CTAB metodě. Většina extrakčních metod DNA z listů rostlin je prováděna původní CTAB metodou (Doyle a Doyle, 1973). Od zavedení CTAB metody bylo publikováno mnoho různých modifikací této metody. Protokol použitý v bakalářské práci byl původně vytvořen pro izolaci HMW DNA z rostlinných druhů, které obsahují vysoký podíl především polyfenolických sekundárních metabolitů (banánovník a yam), u kterých bylo velmi omezené počáteční množství DNA (Hřibová, nepublikováno). Úspěšnost této jednoduché izolace vysoce kvalitní HMW DNA pro ONT sekvenování byla prokázána i v případě cizrný beraní, rostlinného druhu, který obsahuje vysoký podíl polysacharidů. Nejlepší výsledky z hlediska délky vyizolovaných fragmentů byly získány nejšetnějším protokolem, tedy homogenizací čerstvého rostlinného materiálu v tekutém dusíku a následném navinutí DNA na skleněný háček. V některých případech, např. při malém vstupním množství materiálu nebo u rostlinných druhů, které se vyznačují malými velikostmi genomů, se nemusí vždy podařit namotat vysráženou HMW DNA na skleněný háček. V takovém případě je tedy možné vysráženou DNA centrifugovat při nízkých otáčkách. Tímto postupem je sice získaná HMW DNA obecně kratší než v případě namotávání na skleněný háček, ale stále z hlediska délek fragmentů i kvality vhodná pro sekvenování třetí generace.

Dalším problematickým krokem, který je spojen s přítomností různých sekundárních metabolitů je využití látek, které zajistí jejich inhibici – např.  $\beta$ -merkapt ethanol, PVP nebo Triton. K odstranění (zmírnění) problémů spojených s fenoly a polysacharidy byl protokol modifikován za použití PVP a vysokých koncentrací solí k izolaci genomové DNA. Bylo zjištěno, že přídavek PVP k zavedené metodě ne vždy ovlivňuje efektivitu protokolu. PVP tedy není vyžadováno k zabránění fenolické oxidace a to pravděpodobně díky přítomnosti  $\beta$ -merkapt ethanolu a centrifugaci po inkubaci při 65 °C (Healey *et al.*, 2014).

Naopak izolace HMW DNA z cizrný beraní pomocí izolace jader a následné izolace HMW DNA pomocí komerčně dostupného kitu Nanobind Plant Nuclei Big DNA Kit (Circulomics), využívajícím magnetické disky, byla neúspěšná. Ačkoliv se v prvním kroku podařilo vyizolovat buněčná jádra, ve druhém kroku – izolaci HMW DNA pomocí Nanobind kitu se však po deproteinaci jader nepodařilo navázat DNA na magnetické disky, což bylo s největší pravděpodobností způsobeno přítomností sekundárních metabolitů nebo polysacharidů, které znemožnily vytvořit optimální prostředí pro navázání DNA na magnetické disky. Podobně, Vaillancourt a Buell (2019) testovali různé extrakční metody pro izolaci HMW DNA a jedna z nich byla izolace za použití Nanobind kitu. U rostliny *S. tuberosum* provedli úspěšnou extrakci buněčných jader dle protokolu Workmana *et al.* (2018). Jádra byla následně využita jako vstupní materiál pro Nanobind kit a následně k Oxford Nanopore sekvenování.

Proces byl úspěšný a došlo k získání průměrné hodnoty N50 42,5 kb. Stejný postup byl následně použit k izolaci DNA z *L. angustifolia*, což je druh známý pro produkci sekundárních metabolitů, ale izolace byla neúspěšná. Pravděpodobně tedy použití Nanobind kitu je vhodné u rostlin s nižším obsahem sekundárních metabolitů, ale je obtížné jeho využití k izolaci HMW DNA z rostlin, které obsahují velké množství sekundárních metabolitů a polysacharidů. Tuto skutečnost podporují nejen další negativní pokusy vyizolovat HMW DNA z jiných rostlinných druhů, zahrnujících např. různé genotypy banánovníku, yam či čirok, které probíhaly v laboratoři školícího pracoviště, ale především fakt, že v publikovaných pracích, ve kterých byla izolována rostlinná DNA pro ONT sekvenování, nebylo využito Nanobind kitu, ale tradičních přístupů (např. Belser *et al.*, 2021; Bibi *et al.*, 2021; Mascher *et al.*, 2021; Vondrak *et al.*, 2021; Bredeson *et al.*, 2022; Kawash *et al.*, 2022; Miao *et al.*, 2021).

Vzhledem k tomu, že pro ONT sekvenování bylo použito standardních kitů pro přípravu sekvenačních knihoven, byla vyizolovaná DNA z čerstvého rostlinného materiálu na základě PFGE vyhodnocena jako příliš dlouhá, což by mohlo vést k rychlému ucpávání nanopórů. Před samotnou přípravou ONT knihoven byla HMW DNA tedy jemně fragmentována. Získané fragmenty měly velikost kolem 48–97 kb, kdy optimální velikost fragmentů HMW DNA je > 50 kb (Mayjonade *et al.*, 2016). K přípravě sekvenační knihovny a k následnému sekvenování byly vybrány dva genotypy a to ICC 1882 a ICC 8261, které byly izolovány z čerstvého rostlinného materiálu (homogenizace tekutým dusíkem) za použití MATAB pufru. K sekvenování byly vybrány sekvenační komůrky, které obsahovaly menší počet aktivních pórů (~ 400–800), čemuž odpovídal i výnos sekvenace. Při samotné sekvenaci nedocházelo k rychlé degradaci nanopórů a pro oba genotypy byly získány sekvence s délkou čtení N50 ~ 14–17 kb, kdy nejdelší přečtené sekvence měly délku ~ 112–559 kb.

Sekvenační data potvrdila, že rychlý protokol využívající CTAB/MATAB izolace DNA je vhodný pro získání vysokomolekulární DNA vhodné pro Oxford Nanopore sekvenování u rostlinných druhů s vysokým obsahem polysacharidů. Navíc velikost vyizolovaných fragmentů naznačuje využití protokolu pro sekvenování extrémně dlouhých sekvencí pomocí Oxford Nanopore technologie.

## 7 Závěr

V rámci bakalářské práce se podařilo optimalizovat protokol k izolaci vysokomolekulární DNA z cizrny beraní (*Cicer arietinum*) pro Oxford Nanopore sekvenování. HMW DNA je potřebná pro získání dlouhých čtení pomocí sekvenačních technologií třetí generace, ale její izolace z rostlin bývá často obtížná.

Izolace HMW DNA byla prováděna dvěma postupy, kdy pouze jeden byl úspěšný. Izolace buněčných jader a následná izolace vysokomolekulární DNA pomocí komerčně dostupného kitu (Nanobind Plant Nuclei Big DNA Kit, Circulomics) nebyla úspěšná – nedošlo k navázání DNA na disky. Použití rychlé CTAB metody, která byla za účelem získání HMW DNA pro ONT technologie z rostlin s vysokým poměrem polyfenolických sekundárních metabolitů využívána v laboratoři školitele, bylo úspěšné i pro cizrnu beraní a vedlo k získání kvalitní vysokomolekulární DNA.

Izolace byla prováděna jak z lyofilizovaného, tak z čerstvého rostlinného materiálu za použití různých pufků (MATAB, MATAB-PVP, CTAB, CTAB-PVP). Bylo zjištěno, že k získání kvalitnější HMW DNA vede izolace z čerstvého vstupního materiálu, homogenizovaného tekutým dusíkem. Volba pufku měla minimální vliv na kvalitu získané DNA – všechny pufky vedly k získání téměř stejně dlouhých molekul DNA (~ 48–194 kb). Po jemné fragmentaci a přečištění došlo k získání fragmentů o velikosti ~ 48–97 kb, které byly využity pro přípravu sekvenačních knihoven a následné Oxford Nanopore sekvenování.

K přípravě sekvenační knihovny a k Oxford Nanopore sekvenování byly vybrány vzorky izolované ze dvou genotypů cizrny beraní, které byly izolovány z čerstvého rostlinného materiálu za použití MATAB pufku. Při sekvenování nedocházelo k rychlému ucpávání/degradaci pórů a sekvenační data potvrdila, že izolací došlo k získání vysokomolekulární DNA, která měla dostatečnou kvalitu. Podařilo se získat sekvence s délkou čtení N50 ~ 14–17 kb, kdy nejdelší přečtené sekvence měly délku ~ 112–559 kb.

## 8 Literatura

- Ambardar S., Gupta R., Trakroo D., Lal R., Vakhlu J. (2016): High Throughput Sequencing: An Overview of Sequencing Chemistry. *Indian journal of microbiology*, 56(4): 394–404.
- Ansorge W. J. (2009): Next-generation DNA sequencing techniques. *New biotechnology*, 25(4): 195–203.
- Arabidopsis Genome Initiative (2000): Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408(6814): 796–815.
- Barghini E., Natali L., Cossu R. M., Giordani T., Pindo M., Cattonaro F., Scalabrin S., Velasco R., Morgante M., Cavallini A. (2014): The peculiar landscape of repetitive sequences in the olive (*Olea europaea* L.) genome. *Genome biology and evolution*, 6(4): 776–791.
- Belser C., Baurens F. C., Noel B., Martin G., Cruaud C., Istace B., Yahiaoui N., Labadie K., Hřibová E., Doležel J., Lemainque A., Wincker P., D'Hont A., Aury J. M. (2021): Telomere-to-telomere gapless chromosomes of banana using nanopore sequencing. *Communications biology*, 4(1): 1047.
- Belser C., Istace B., Denis E., Dubarry M., Baurens F. C., Falentin C., Genete M., Berrabah W., Chèvre A. M., Delourme R., Deniot G., Denoeud F., Duffé P., Engelen S., Lemainque A., Manzanares-Dauleux M., Martin G., Morice J., Noel B., Vekemans X., D'Hont A., Rousseau-Gueutin M., Barbe V., Cruaud C., Wincker P., Aury J. M. (2018): Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nature plants*, 4(11): 879–887.
- Bennetzen J. L. (2000): Transposable element contributions to plant gene and genome evolution. *Plant molecular biology*, 42(1): 251–269.
- Bentley D. R., Balasubramanian S., Swerdlow H. P., Smith G. P., Milton J., Brown C. G., Hall K. P., Evers D. J., Barnes C. L., Bignell H. R., Boutell J. M., Bryant J., Carter R. J., Keira Cheetham R., Cox A. J., Ellis D. J., Flatbush M. R., Gormley N. A., Humphray S. J., Irving L. J., Karbelashvili M. S., Kirk S. M., Li H., Liu X., Maisinger K. S., Murray L. J., Obradovic B., Ost T., Parkinson M. L., Pratt M. R., Rasolonjatovo I. M., Reed M. T., Rigatti R., Rodighiero C., Ross M. T., Sabot A., Sankar S. V., Scally A., Schroth G. P., Smith M. E., Smith V. P., Spiridou A., Torrance P. E., Tzonev S. S., Vermaas E. H., Walter K., Wu X., Zhang L., Alam M. D., Anastasi C., Aniebo I. C., Bailey D. M.,



Bancarz I. R., Banerjee S., Barbour S. G., Baybayan P. A., Benoit V. A., Benson K. F., Bevis C., Black P. J., Boodhun A., Brennan J. S., Bridgham J. A., Brown R. C., Brown A. A., Buermann D. H., Bundu A. A., Burrows J. C., Carter N. P., Castillo N., Chiara E., Catenazzi M., Chang S., Neil Cooley R., Crake N. R., Dada O. O., Diakoumakos K. D., Dominguez-Fernandez B., Earnshaw D. J., Egbujor U. C., Elmore D. W., Etchin S. S., Ewan M. R., Fedurco M., Fraser L. J., Fuentes Fajardo K. V., Scott Furey W., George D., Gietzen K. J., Goddard C. P., Golda G. S., Granieri P. A., Green D. E., Gustafson D. L., Hansen N. F., Harnish K., Haudenschild C. D., Heyer N. I., Hims M. M., Ho J. T., Horgan A. M., Hoschler K., Hurwitz S., Ivanov D. V., Johnson M. Q., James T., Huw Jones T. A., Kang G. D., Kerelska T. H., Kersey A. D., Khrebtukova I., Kindwall A. P., Kingsbury Z., Kokko-Gonzales P. I., Kumar A., Laurent M. A., Lawley C. T., Lee S. E., Lee X., Liao A. K., Loch J. A., Lok M., Luo S., Mammen R. M., Martin J. W., McCauley P. G., McNitt P., Mehta P., Moon K. W., Mullens J. W., Newington T., Ning Z., Ling Ng B., Novo S. M., O'Neill M. J., Osborne M. A., Osnowski A., Ostadan O., Paraschos L. L., Pickering L., Pike A. C., Pike A. C., Chris Pinkard D., Pliskin D. P., Podhasky J., Quijano V. J., Raczy C., Rae V. H., Rawlings S. R., Chiva Rodriguez A., Roe P. M., Rogers J., Rogert Bacigalupo M. C., Romanov N., Romieu A., Roth R. K., Rourke N. J., Ruediger S. T., Rusman E., Sanches-Kuiper R. M., Schenker M. R., Seoane J. M., Shaw R. J., Shiver M. K., Short S. W., Sizto N. L., Sluis J. P., Smith M. A., Ernest Sohna Sohna J., Spence E. J., Stevens K., Sutton N., Szajkowski L., Tregidgo C. L., Turcatti G., Vandevondele S., Verhovsky Y., Virk S. M., Wakelin S., Walcott G. C., Wang J., Worsley G. J., Yan J., Yau L., Zuerlein M., Rogers J., Mullikin J. C., Hurles M. E., McCooke N. J., West J. S., Oaks F. L., Lundberg P. L., Klenerman D., Durbin R., Smith A. J. (2008): Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218): 53–59.

Bibi A. C., Kollias A., Astrinaki M., Vassou D., Kafetzopoulos D., Kalantidis K., Moschou P. N. (2021): DNA Preparation in *Vitis Vinifera* L. For Third Generation Sequencing. *ResearchSquare*.

Bishop R. (2010): Applications of fluorescence in situ hybridization (FISH) in detecting genetic aberrations of medical significance. *Bioscience Horizons*, 3(1): 85-95.

Bredeson J. V., Lyons J. B., Oniyinde I. O., Okereke N. R., Kolade O., Nnabue I., Nwadili C. O., Hřibová E., Parker M., Nwogha J., Shu S., Carlson J., Kariba R., Muthemba S., Knop K., Barton G. J., Sherwood A. V., Lopez-Montes A., Asiedu R., Jamnadass,



- Muchugi A., Goodstein D., Egesi C. N., Featherston J., Asfaw A., Simpson G. G., Doležel J., Hendre P. S., Van Deynze A., Kumar P. L., Obidiegwu J. E., Bhattacharjee R., Rokhsar D. S. (2022): Chromosome evolution and the genetic basis of agronomically important traits in greater yam. *Nature communications*, 13(1): 2001.
- Comings D. E. (1975): Mechanisms of chromosome banding. VIII. Hoechst 33258-DNA interaction. *Chromosoma*, 52(3): 229–243.
- Comings D. E. (1978): Mechanisms of chromosome banding and implications for chromosome structure. *Annual review of genetics*, 12: 25–46.
- Cutler D. J., Zwick M. E., Carrasquillo M. M., Yohn C. T., Tobin K. P., Kashuk C., Mathews D. J., Shah N. A., Eichler E. E., Warrington J. A., Chakravarti A. (2001): High-throughput variation detection and genotyping using microarrays. *Genome research*, 11(11): 1913–1925.
- Deschamps S., Zhang Y., Llaca V., Ye L., Sanyal A., King M., May G., Lin H. (2018): A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping. *Nature communications*, 9(1): 4844.
- Devos K. M., Brown J. K., Bennetzen J. L. (2002): Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. *Genome research*, 12(7): 1075–1079.
- Dewey C. M., Cenik B., Sephton C. F., Johnson B. A., Herz J., Yu G. (2012): TDP-43 aggregation in neurodegeneration: are stress granules the key? *Brain research*, 1462: 16–25.
- D'Hont A., Denoeud F., Aury J. M., Baurens F. C., Carreel F., Garsmeur O., Noel B., Bocs S., Droc G., Rouard M., Da Silva C., Jabbari K., Cardi C., Poulain J., Souquet M., Labadie K., Jourda C., Lenggellé J., Rodier-Goud M., Alberti A., Bernard M., Correa M., Ayyampalayam S., Mckain M. R., Leebens-Mack J., Burgess D., Freeling M., Mbéguié-A-Mbéguié D., Chabannes M., Wicker T., Panaud O., Barbosa J., Hřibová E., Heslop-Harrison P., Habas R., Rivallan R., Francois P., Poirion C., Kilian A., Burthia D., Jenny C., Bakry F., Brown S., Guignon V., Kema G., Dita M., Waalwijk C., Joseph S., Dievert A., Jaillon O., Leclercq J., Argout X., Lyons E., Almeida A., Jeridi M., Dolezel J., Roux N., Risterucci A. M., Weissenbach J., Ruiz M., Glaszmann J. C., Quétier F., Yahiaoui N., Wincker P. (2012): The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, 488(7410): 213–217.

- Dijk E. L., Jaszczyszyn Y., Naquin D., Thermes C. (2018): The Third Revolution in Sequencing Technology. *Trends in genetics: TIG*, 34(9): 666–681.
- Doyle J. J., Doyle J. L. (1990): Isolation of plant DNA from fresh tissue. *Focus*, 12: 13–15.
- Drmanac R., Drmanac S., Chui G., Diaz R., Hou A., Jin H., Jin P., Kwon S., Lacy S., Moeur B., Shafto J., Swanson D., Ukrainczyk T., Xu C., Little D. (2002): Sequencing by hybridization (SBH): advantages, achievements, and opportunities. *Advances in biochemical engineering/biotechnology*, 77: 75–101.
- Dutrillaux B., Couturier J., Rotman J., Salat J., Lejeune J. (1972): Stérilité et translocation familiale t (1q-;Xq+ [Sterility and familial t (1q-;Xq+) translocation]). *Comptes rendus hebdomadaires des seances de l'Academie des sciences. Serie D: Sciences naturelles*, 274(24): 3324–3327.
- Eid J., Fehr A., Gray J., Luong K., Lyle J., Otto G., Peluso P., Rank D., Baybayan P., Bettman B., Bibillo A., Bjornson K., Chaudhuri B., Christians F., Cicero R., Clark S., Dalal R., Dewinter A., Dixon J., Foquet M., Gaertner A., Hardenbol P., Heiner C., Hester K., Holden D., Kearns G., Kong X., Kuse R., Lacroix Y., Lin S., Lundquist P., Ma C., Marks P., Maxham M., Murphy D., Park I., Pham T., Phillips M., Roy J., Sebra R., Shen G., Sorenson J., Tomaney A., Travers K., Trulson M., Vieceli J., Wegener J., Wu D., Yang A., Zaccarin D., Zhao P., Zhong F., Korlach J., Turner S. (2009): Real-time DNA sequencing from single polymerase molecules. *Science (New York, N.Y.)*, 323(5910): 133–138.
- Ergüner B., Ustek D., Sagiroglu M. S. (2015): Performance comparison of Next Generation sequencing platforms. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2015*: 6453–6456.
- Figueira T. R., Okura V., Rodrigues da Silva F., Jose da Silva M., Kudrna D., Ammiraju J. S., Talag J., Wing R., Arruda P. (2012): A BAC library of the SP80-3280 sugarcane variety (*saccharum* sp.) and its inferred microsynteny with the sorghum genome. *BMC Research Notes*, 5(1), 185.
- Fleischmann A., Michael T. P., Rivadavia F., Sousa A., Wang W., Temsch E. M., Greilhuber J., Müller K. F., Heubl G. (2014): Evolution of genome size and chromosome number in the carnivorous plant genus *Genlisea* (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. *Annals of botany*, 114(8): 1651–1663.

- Flusberg B. A., Webster D. R., Lee J. H., Travers K. J., Olivares E. C., Clark T. A., Korlach J., Turner, S. W. (2010): Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature methods*, 7(6): 461–465.
- Fuller C. W., Middendorf L. R., Benner S. A., Church G. M., Harris T., Huang X., Jovanovich S. B., Nelson J. R., Schloss J. A., Schwartz D. C., Vezenov D. V. (2009): The challenges of sequencing by synthesis. *Nature biotechnology*, 27(11): 1013–1023.
- Garalde D. R., Snell E. A., Jachimowicz D., Sipos B., Lloyd J. H., Bruce M., Pantic N., Admassu T., James P., Warland A., Jordan M., Ciccone J., Serra S., Keenan J., Martin S., McNeill L., Wallace E. J., Jayasinghe L., Wright C., Blasco J., Young S., Brocklebank D., Juul S., Clarke J., Heron A. J., Turner D. J. (2018): Highly parallel direct RNA sequencing on an array of nanopores. *Nature methods*, 15(3): 201–206.
- Garrido-Ramos M. A. (2015): Satellite DNA in Plants: More than Just Rubbish. *Cytogenetic and genome research*, 146(2): 153–170.
- Gasperskaja E., Kučinskas V. (2017): The most common technologies and tools for functional genome analysis. *Acta medica Lituanica*, 24(1): 1–11.
- Gelman J. S., Sironi J., Berezniuk I., Dasgupta S., Castro L. M., Gozzo F. C., Ferro E. S., Fricker L. D. (2013): Alterations of the intracellular peptidome in response to the proteasome inhibitor bortezomib. *PLoS one*, 8(1): e53263.
- Giani A. M., Gallo G. R., Gianfranceschi L., Formenti G. (2019): Long walk to genomics: History and current approaches to genome sequencing and assembly. *Computational and structural biotechnology journal*, 18: 9–19.
- Goff S. A., Ricke D., Lan T. H., Presting G., Wang R., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H., Hadley D., Hutchison D., Martin C., Katagiri F., Lange B. M., Moughamer T., Xia Y., Budworth P., Zhong J., Miguel T., Paszkowski U., Zhang S., Colbert M., Sun W. L., Chen L., Cooper B., Park S., Wood T. C., Mao L., Quail P., Wing R., Dean R., Yu Y., Zharkikh A., Shen R., Sahasrabudhe S., Thomas A., Cannings R., Gutin A., Pruss D., Reid J., Tavtigian S., Mitchell J., Eldredge G., Scholl T., Miller R. M., Bhatnagar S., Adey N., Rubano T., Tusneem N., Robinson R., Feldhaus J., Macalma T., Oliphant A., Briggs S. (2002): A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science (New York, N.Y.)*, 296(5565): 92–100.
- Goodwin S., McPherson J. D., McCombie W. R. (2016): Coming of age: ten years of next-generation sequencing technologies. *Nature reviews. Genetics*, 17(6): 333–351.

- Gregory T. R. (2001): Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological reviews of the Cambridge Philosophical Society*, 76(1): 65–101.
- Hannenhalli S. (1996): Polynomial algorithm for computing translocation distance between genomes. *Discrete Applied Mathematics*, 71: 137–151.
- Head S. R., Komori H. K., LaMere S. A., Whisenant T., Van Nieuwerburgh F., Salomon D. R., Ordoukhanian P. (2014): Library construction for next-generation sequencing: overviews and challenges. *BioTechniques*, 56(2): 61–passim.
- Healey A., Furtado A., Cooper T., Henry R. J. (2014): Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant methods*, 10, 21.
- Heather J. M., Chain B. (2016): The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1): 1–8.
- Holley R. W., Apgar J., Everett G. A., Madison J. T., Marquisee M., Merrill S. H., Penswick J. R., Zamir A. (1965): STRUCTURE OF A RIBONUCLEIC ACID. *Science (New York, N.Y.)*, 147(3664): 1462–1465.
- Hollister J. D., Gaut B. S. (2009): Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome research*, 19(8): 1419–1428.
- Hutchison C. A., 3rd (2007): DNA sequencing: bench to bedside and beyond. *Nucleic acids research*, 35(18): 6227–6237.
- Chin C. S., Peluso P., Sedlazeck F. J., Nattestad M., Concepcion G. T., Clum A., Dunn C., O'Malley R., Figueroa-Balderas R., Morales-Cruz A., Cramer G. R., Delledonne M., Luo C., Ecker J. R., Cantu D., Rank D. R., Schatz M. C. (2016): Phased diploid genome assembly with single-molecule real-time sequencing. *Nature methods*, 13(12): 1050–1054.
- Jaillon O., Aury J. M., Noel B., Policriti A., Clepet C., Casagrande A., Choisne N., Aubourg S., Vitulo N., Jubin C., Vezzi A., Legeai F., Huguency P., Dasilva C., Horner D., Mica E., Jublot D., Poulain J., Bruyère C., Billault A., Segurens B., Gouyvenoux M., Ugarte E., Cattonaro F., Anthouard V., Vico V., Del Fabbro C., Alaux M., Di Gaspero G., Dumas V., Felice N., Paillard S., Juman I., Moroldo M., Scalabrin S., Canaguier A., Le Clainche I., Malacrida G., Durand E., Pesole G., Laucou V., Chatelet P., Merdinoglu

- D., Delledonne M., Pezzotti M., Lecharny A., Scarpelli C., Artiguenave F., Pè M. E., Valle G., Morgante M., Caboche M., Adam-Blondon A. F., Weissenbach J., Quétier F., Wincker P., French-Italian Public Consortium for Grapevine Genome Characterization (2007): The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449(7161): 463–467.
- Jain M., Koren S., Miga K. H., Quick J., Rand A. C., Sasani T. A., Tyson J. R., Beggs A. D., Dilthey A. T., Fiddes I. T., Malla S., Marriott H., Nieto T., O'Grady J., Olsen H. E., Pedersen B. S., Rhie A., Richardson H., Quinlan A. R., Snutch T. P., Tee L., Paten B., Phillippy A. M., Simpson J. T., Loman N. J., Loose M. (2018): Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature biotechnology*, 36(4): 338–345.
- Jiang J. (2019): Fluorescence in situ hybridization in plants: recent developments and future applications. *Chromosome research: an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 27(3): 153–165.
- Jung I., Schmitt A., Diao Y., Lee A. J., Liu T., Yang D., Tan C., Eom J., Chan M., Chee S., Chiang Z., Kim C., Masliah E., Barr C. L., Li B., Kuan S., Kim D., Ren B. (2019): A compendium of promoter-centered long-range chromatin interactions in the human genome. *Nature genetics*, 51(10): 1442–1449.
- Kapustová V., Tulpová Z., Toegelová H., Novák P., Macas J., Karafiátová M., Hřibová E., Doležel J., Šimková H. (2019): The Dark Matter of Large Cereal Genomes: Long Tandem Repeats. *International journal of molecular sciences*, 20(10): 2483.
- Kawash J., Colt K., Hartwick N. T., Abramson B. W., Vorsa N., Polashock J. J., Michael T. P. (2022): Contrasting a reference cranberry genome to a crop wild relative provides insights into adaptation, domestication, and breeding. *PloS one*, 17(3): e0264966.
- Kelly L. J., Renny-Byfield S., Pellicer J., Macas J., Novák P., Neumann P., Lysak M. A., Day P. D., Berger M., Fay M. F., Nichols R. A., Leitch A. R., Leitch I. J. (2015): Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. *The New phytologist*, 208(2): 596–607.
- Kovarik A., Dadejova M., Lim Y. K., Chase M. W., Clarkson J. J., Knapp S., Leitch A. R. (2008): Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. *Annals of botany*, 101(6): 815–823.

- Latt S. A. (1977): Fluorometric detection of deoxyribonucleic acid synthesis; possibilities for interfacing bromodeoxyuridine dye techniques with flow fluorometry. *The journal of histochemistry and cytochemistry: official journal of the Histochemistry Society*, 25(7): 913–926.
- Liehr T., Ewers E., Kosyakova N., Klaschka V., Rietz F., Wagner R., Weise A. (2009): Handling small supernumerary marker chromosomes in prenatal diagnostics. *Expert review of molecular diagnostics*, 9(4): 317–324.
- Liu L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L., Law M. (2012): Comparison of next-generation sequencing systems. *Journal of biomedicine & biotechnology*, 251364.
- Logsdon G. A., Vollger M. R., Eichler E. E. (2020): Long-read human genome sequencing and its applications. *Nature reviews. Genetics*, 21(10): 597–614.
- Loman N. J., Quick J., Simpson J. T. (2015): A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nature methods*, 12(8): 733–735.
- Mahmoud M., Gobet N., Cruz-Dávalos D. I., Mounier N., Dessimoz C., Sedlazeck F. J. (2019): Structural variant calling: the long and the short of it. *Genome biology*, 20(1): 246.
- Manen J. F., Sinitsyna O., Aeschbach L., Markov A. V., Sinitsyn A. (2005): A fully automatable enzymatic method for DNA extraction from plant tissues. *BMC plant biology*, 5: 23.
- Mardis E. R. (2013): Next-generation sequencing platforms. *Annual review of analytical chemistry (Palo Alto, Calif.)*, 6: 287–303.
- Margulies M., Egholm M., Altman W. E., Attiya S., Bader J. S., Bemben L. A., Berka J., Braverman M. S., Chen Y. J., Chen Z., Dewell S. B., Du L., Fierro J. M., Gomes X. V., Godwin B. C., He W., Helgesen S., Ho C. H., Irzyk G. P., Jando S. C., Alenquer M. L., Jarvie T. P., Jirage K. B., Kim J. B., Knight J. R., Lanza J. R., Leamon J. H., Lefkowitz S. M., Lei M., Li J., Lohman K. L., Lu H., Makhijani V. B., McDade K. E., McKenna M. P., Myers E. W., Nickerson E., Nobile J. R., Plant R., Puc B. P., Ronan M. T., Roth G. T., Sarkis G. J., Simons J. F., Simpson J. W., Srinivasan M., Tartaro K. R., Tomasz A., Vogt K. A., Volkmer G. A., Wang S. H., Wang Y., Weiner M. P., Yu P., Begley R. F., Rothberg J. M. (2005): Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057): 376–380.
- Martorelli Di Genova B., Wilson S. K., Dubey J. P., Knoll L. J. (2019): Intestinal delta-6-

- desaturase activity determines host range for *Toxoplasma* sexual reproduction. *PLoS biology*, 17(8): e3000364.
- Mascher M., Wicker T., Jenkins J., Plott C., Lux T., Koh C. S., Ens J., Gundlach H., Boston L. B., Tulpová Z., Holden S., Hernández-Pinzón I., Scholz U., Mayer K., Spannagl M., Pozniak C. J., Sharpe A. G., Šimková H., Moscou M. J., Grimwood J., Schmutz J., Stein N. (2021): Long-read sequence assembly: a technical evaluation in barley. *The Plant cell*, 33(6): 1888–1906.
- Maurya S., Singh U. P., Singh D. P., Singh K. P., Srivastava J. S. (2005): Secondary metabolites of chickpea (*Cicer arietinum*) and their role in pathogenesis after infection by *Sclerotium rolfsii*/Sekundärmetaboliten der Kichererbse (*Cicer arietinum*) und ihre Rolle während der Pathogenese nach einem Befall mit *Sclerotium rolfsii*. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz/Journal of Plant Diseases and Protection*, 118-123.
- Maxam A. M., Gilbert W. (1977): A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74(2): 560–564.
- Mayjonade B., Gouzy J., Donnadiou C., Pouilly N., Marande W., Callot C., Langlade N., Muñoz S. (2016): Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *BioTechniques*, 61(4): 203–205.
- Melters D. P., Bradnam K. R., Young H. A., Telis N., May M. R., Ruby J. G., Sebra R., Peluso P., Eid J., Rank D., Garcia J. F., DeRisi J. L., Smith T., Tobias C., Ross-Ibarra J., Korff I., Chan S. W. (2013): Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome biology*, 14(1): R10.
- Miao J., Feng Q., Li Y., Zhao Q., Zhou C., Lu H., Fan D., Yan J., Lu Y., Tian Q., Li W., Weng Q., Zhang L., Zhao Y., Huang T., Li L., Huang X., Sang T., Han B. (2021): Chromosome-scale assembly and analysis of biomass crop *Miscanthus lutarioriparius* genome. *Nature communications*, 12(1): 1-13.
- Michael T. P., Jupe F., Bemm F., Motley S. T., Sandoval J. P., Lanz C., Loudet O., Weigel D., Ecker J. R. (2018): High contiguity *Arabidopsis thaliana* genome assembly with a single nanopore flow cell. *Nature communications*, 9(1): 541.
- Mills R. E., Walter K., Stewart C., Handsaker R. E., Chen K., Alkan C., Abyzov A., Yoon S. C., Ye K., Cheetham R. K., Chinwalla A., Conrad D. F., Fu Y., Grubert F., Hajirasouliha I., Hormozdiari F., Iakoucheva L. M., Iqbal Z., Kang S., Kidd J. M., Konkel M. K.,



- Korn J., Khurana E., Kural D., Lam H. Y., Leng J., Li R., Li Y., Lin C. Y., Luo R., Mu X. J., Nemesh J., Peckham H. E., Rausch T., Scally A., Shi X., Stromberg M. P., Stütz A. M., Urban A. E., Walker J. A., Wu J., Zhang Y., Zhang Z. D., Batzer M. A., Ding L., Marth G. T., McVean G., Sebat J., Snyder M., Wang J., Ye K., Eichler E. E., Gerstein M. B., Hurles M. E., Lee C., McCarroll S. A., Korbel J. O., 1000 Genomes Project (2011): Mapping copy number variation by population-scale genome sequencing. *Nature*, 470(7332): 59–65.
- Ming R., Moore P. H., Zee F., Abbey C. A., Ma H., Paterson A. H. (2001): Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. *TAG Theoretical and Applied Genetics*, 102(6-7): 892–899.
- Mitsuhashi S., Matsumoto N. (2020): Long-read sequencing for rare human genetic diseases. *Journal of human genetics*, 65(1): 11–19.
- Navrátilová P., Toegelová H., Tulpová Z., Kuo Y. T., Stein N., Doležel J., Houben A., Šimková H., Mascher M. (2022): Prospects of telomere-to-telomere assembly in barley: Analysis of sequence gaps in the MorexV3 reference genome. *Plant biotechnology journal*, 10.1111/pbi.13816. Advance online publication.
- Nyrén P., Pettersson B., Uhlén M. (1993): Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Analytical biochemistry*, 208(1): 171–175.
- Pareek C. S., Smoczynski R., Tretyn A. (2011): Sequencing technologies and genome sequencing. *Journal of applied genetics*, 52(4): 413–435.
- Payne A., Holmes N., Rakyan V., Loose M. (2019): BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files. *Bioinformatics (Oxford, England)*, 35(13): 2193–2198.
- Pe'er I., Arbili N., Liu Y., Enck C., Gelfand C. A., Shamir R. (2003): Advanced computational techniques for re-sequencing DNA with polymerase signaling assay arrays. *Nucleic acids research*, 31(19): 5667–5675.
- Pellicer J., Fay M. F., Leitch I. J. (2010): The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society*, 164(1): 10-15.
- Peterson D. G., Tomkins J. P., Frisch D. A., Wing R. A., Paterson A. H. (2000): Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. *Journal of Agricultural genomics*, 5: 1-100.



- Prober J. M., Trainor G. L., Dam R. J., Hobbs F. W., Robertson C. W., Zagursky R. J., Cocuzza A. J., Jensen M. A., Baumeister K. (1987): A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science (New York, N.Y.)*, 238(4825): 336–341.
- Pushkarev D., Neff N. F., Quake S. R. (2009): Single-molecule sequencing of an individual human genome. *Nature biotechnology*, 27(9): 847–850.
- Rand A. C., Jain M., Eizenga J. M., Musselman-Brown A., Olsen H. E., Akeson M., Paten B. (2017): Mapping DNA methylation with high-throughput nanopore sequencing. *Nature methods*, 14(4): 411–413.
- Rang F. J., Kloosterman W. P., de Ridder J. (2018): From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome biology*, 19(1): 90.
- Roessler K., Bousios A., Meca E., Gaut B. S. (2018): Modeling Interactions between Transposable Elements and the Plant Epigenetic Response: A Surprising Reliance on Element Retention. *Genome biology and evolution*, 10(3): 803–815.
- Ronaghi M., Karamohamed S., Pettersson B., Uhlén M., Nyrén P. (1996): Real-time DNA sequencing using detection of pyrophosphate release. *Analytical biochemistry*, 242(1): 84–89.
- Rothberg J. M., Hinz W., Rearick T. M., Schultz J., Mileski W., Davey M., Leamon J. H., Johnson K., Milgrew M. J., Edwards M., Hoon J., Simons J. F., Marran D., Myers J. W., Davidson J. F., Branting A., Nobile J. R., Puc B. P., Light D., Clark T. A., Huber M., Branciforte J. T., Stoner I. B., Cawley S. E., Lyons M., Fu Y., Homer N., Sedova M., Miao X., Reed B., Sabina J., Feierstein E., Schorn M., Alanjary M., Dimalanta E., Dressman D., Kasinskas R., Sokolsky T., Fidanza J. A., Namsaraev E., McKernan K. J., Williams A., Roth G. T., Bustillo J. (2011): An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, 475(7356): 348–352.
- Saghai-Marouf M. A., Soliman K. M., Jorgensen R. A., Allard R. W. (1984): Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 81(24): 8014–8018.
- Sakamoto Y., Xu L., Seki M., Yokoyama T. T., Kasahara M., Kashima Y., Ohashi A., Shimada Y., Motoi N., Tsuchihara K., Kobayashi S. S., Kohno T., Shiraishi Y., Suzuki

- A., Suzuki Y. (2020): Long-read sequencing for non-small-cell lung cancer genomes. *Genome research*, 30(9): 1243–1257.
- Sanger F., Coulson A. R. (1975): A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology*, 94(3): 441–448.
- Sanger F., Nicklen S., Coulson A. R. (1977): DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12): 5463–5467.
- Sanger F., Thompson E. O. (1953a): The amino-acid sequence in the glyceryl chain of insulin. I. The identification of lower peptides from partial hydrolysates. *The Biochemical journal*, 53(3): 353–366.
- Sanger F., Thompson E. O. (1953b): The amino-acid sequence in the glyceryl chain of insulin. II. The investigation of peptides from enzymic hydrolysates. *The Biochemical journal*, 53(3): 366–374.
- Sasani T. A., Tyson J. R., Beggs A. D., Dilthey A. T., Fiddes I. T., Malla S., Marriott H., Nieto T., O'Grady J., Olsen H. E., Pedersen B. S., Rhie A., Richardson H., Quinlan A. R., Snutch T. P., Tee L., Paten B., Phillippy A. M., Simpson J. T., Loman N. J., Loose M. (2018): Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature biotechnology*, 36(4): 338–345.
- Shin S. J., You S. C., Park Y. R., Roh J., Kim J. H., Haam S., Reich C. G., Blacketer C., Son D. S., Oh S., Park R. W. (2019): Genomic Common Data Model for Seamless Interoperation of Biomedical Data in Clinical Practice: Retrospective Study. *Journal of medical Internet research*, 21(3): e13249.
- Schadt E. E., Turner S., Kasarskis A. (2010): A window into third-generation sequencing. *Human molecular genetics*, 19(R2): R227–R240.
- Schmid M., Frei D., Patrignani A., Schlapbach R., Frey J. E., Remus-Emsermann M., Ahrens C. H. (2018): Pushing the limits of de novo genome assembly for complex prokaryotic genomes harboring very long, near identical repeats. *Nucleic acids research*, 46(17): 8953–8965.
- Schmidt K., Mwaigwisya S., Crossman L. C., Doumith M., Munroe D., Pires C., Khan A. M., Woodford N., Saunders N. J., Wain J., O'Grady J., Livermore D. M. (2017): Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *The Journal of antimicrobial*

*chemotherapy*, 72(1): 104–114.

Schnable P. S., Ware D., Fulton R. S., Stein J. C., Wei F., Pasternak S., Liang C., Zhang J., Fulton L., Graves T. A., Minx P., Reily A. D., Courtney L., Kruchowski S. S., Tomlinson C., Strong C., Delehaunty K., Fronick C., Courtney B., Rock S. M., Belter E., Du F., Kim K., Abbott R. M., Cotton M., Levy A., Marchetto P., Ochoa K., Jackson S. M., Gillam B., Chen W., Yan L., Higginbotham J., Cardenas M., Waligorski J., Applebaum E., Phelps L., Falcone J., Kanchi K., Thane T., Scimone A., Thane N., Henke J., Wang T., Ruppert J., Shah N., Rotter K., Hodges J., Ingenthron E., Cordes M., Kohlberg S., Sgro J., Delgado B., Mead K., Chinwalla A., Leonard S., Crouse K., Collura K., Kudrna D., Currie J., He R., Angelova A., Rajasekar S., Mueller T., Lomeli R., Scara G., Ko A., Delaney K., Wissotski M., Lopez G., Campos D., Braidotti M., Ashley E., Golser W., Kim H., Lee S., Lin J., Dujmic Z., Kim W., Talag J., Zuccolo A., Fan C., Sebastian A., Kramer M., Spiegel L., Nascimento L., Zutavern T., Miller B., Ambroise C., Muller S., Spooner W., Narechania A., Ren L., Wei S., Kumari S., Faga B., Levy M. J., McMahan L., Van Buren P., Vaughn M. W., Ying K., Yeh C. T., Emrich S. J., Jia Y., Kalyanaraman A., Hsia A. P., Barbazuk W. B., Baucom R. S., Brutnell T. P., Carpita N. C., Chaparro C., Chia J. M., Deragon J. M., Estill J. C., Fu Y., Jeddelloh J. A., Han Y., Lee H., Li P., Lisch D. R., Liu S., Liu Z., Nagel D. H., McCann M. C., SanMiguel P., Myers A. M., Nettleton D., Nguyen J., Penning B. W., Ponnala L., Schneider K. L., Schwartz D. C., Sharma A., Soderlund C., Springer N. M., Sun Q., Wang H., Waterman M., Westerman R., Wolfgruber T. K., Yang L., Yu Y., Zhang L., Zhou S., Zhu Q., Bennetzen J. L., Dawe R. K., Jiang J., Jiang N., Presting G. G., Wessler S. R., Aluru S., Martienssen R. A., Clifton S. W., McCombie W. R., Wing R. A., Wilson R. K. (2009): The B73 maize genome: complexity, diversity, and dynamics. *Science (New York, N.Y.)*, 326(5956): 1112–1115.

Schubert I., Vu G. (2016): Genome Stability and Evolution: Attempting a Holistic View. *Trends in plant science*, 21(9): 749–757.

Simpson J. T., Workman R. E., Zuzarte P. C., David M., Dursi L. J., Timp W. (2017): Detecting DNA cytosine methylation using nanopore sequencing. *Nature methods*, 14(4): 407–410.

Slatko B. E., Gardner A. F., Ausubel F. M. (2018): Overview of Next-Generation Sequencing Technologies. *Current protocols in molecular biology*, 122(1): e59.

- Smith L. M., Fung S., Hunkapiller M. W., Hunkapiller T. J., Hood L. E. (1985): The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic acids research*, 13(7): 2399–2412.
- Smith L. M., Sanders J. Z., Kaiser R. J., Hughes P., Dodd C., Connell C. R., Heiner C., Kent S. B., Hood L. E. (1986): Fluorescence detection in automated DNA sequence analysis. *Nature*, 321(6071): 674–679.
- Song X., Beck C. R., Du R., Campbell I. M., Coban-Akdemir Z., Gu S., Breman A. M., Stankiewicz P., Ira G., Shaw C. A., Lupski J. R. (2018): Predicting human genes susceptible to genomic instability associated with *Alu/Alu*-mediated rearrangements. *Genome research*, 28(8): 1228–1242.
- Staden R. (1979): A strategy of DNA sequencing employing computer programs. *Nucleic acids research*, 6(7): 2601–2610.
- Stark J. R., Cardon Z. G., Peredo E. L. (2020): Extraction of high-quality, high-molecular-weight DNA depends heavily on cell homogenization methods in green microalgae. *Applications in plant sciences*, 8(3): e11333.
- Strachan B. C., Sloane H. S., Lee J. C., Leslie D. C., Landers J. P. (2015): Investigation of the DNA target design parameters for effective hybridization-induced aggregation of particles for the sequence-specific detection of DNA. *The Analyst*, 140(6): 2008–2015
- Sun K., Jiang P., Cheng S. H., Cheng T., Wong J., Wong V., Ng S., Ma B., Leung T. Y., Chan S. L., Mok T., Lai P., Chan H., Sun H., Chan K., Chiu R., Lo Y. (2019): Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. *Genome research*, 29(3): 418–427.
- Šimoníková D., Němečková A., Čížková J., Brown A., Swennen R., Doležel J., Hříbová E. (2020): Chromosome Painting in Cultivated Bananas and Their Wild Relatives (*Musa* spp.) Reveals Differences in Chromosome Structure. *International journal of molecular sciences*, 21(21): 7915.
- Tabor S., Richardson C. C. (1987): DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 84(14): 4767–4771.
- Thomas C. A., Jr (1971): The genetic organization of chromosomes. *Annual review of genetics*, 5: 237–256.

- Travers K. J., Chin C. S., Rank D. R., Eid J. S., Turner S. W. (2010): A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic acids research*, 38(15): e159.
- Vaillancourt B., Buell C. R. (2019): High molecular weight DNA isolation method from diverse plant species for use with Oxford Nanopore sequencing. *BioRxiv*, 783159.
- Verma R. S., Lubs H. A. (1975): A simple R banding technic. *American journal of human genetics*, 27(1): 110–117.
- Voelkerding K. V., Dames S. A., Durtschi J. D. (2009): Next-generation sequencing: from basic research to diagnostics. *Clinical chemistry*, 55(4): 641–658.
- Vondrak T., Oliveira L., Novák P., Koblížková A., Neumann P., Macas J. (2021): Complex sequence organization of heterochromatin in the holocentric plant *Cuscuta europaea* elucidated by the computational analysis of nanopore reads. *Computational and structural biotechnology journal*, 19: 2179–2189.
- Vu V., Verster A. J., Schertzberg M., Chuluunbaatar T., Spensley M., Pajkic D., Hart G. T., Moffat J., Fraser A. G. (2015): Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. *Cell*, 162(2): 391–402.
- Wang F., Huang S., Gao R., Zhou Y., Lai C., Li Z., Xian W., Qian X., Li Z., Huang Y., Tang Q., Liu P., Chen R., Liu R., Li X., Tong X., Zhou X., Bai Y., Duan G., Zhang T., Xu X., Wang J., Yang H., Liu S., He Q., Jin X., Liu L. (2020): Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *Cell discovery*, 6(1): 83.
- Wang O., Chin R., Cheng X., Wu M., Mao Q., Tang J., Sun Y., Anderson E., Lam H. K., Chen D., Zhou Y., Wang L., Fan F., Zou Y., Xie Y., Zhang R. Y., Drmanac S., Nguyen D., Xu C., Villarosa C., Gablenz S., Barua N., Nguyen S., Tian W., Liu J. S., Wang J., Liu X., Qi X., Chen A., Wang H., Dong Y., Zhang W., Alexeev A., Yang H., Wang J., Kristiansen K., Xu X., Drmanac R., Peters B. A. (2019): Efficient and unique cobarcodeing of second-generation sequencing reads from long DNA molecules enabling cost-effective and accurate sequencing, haplotyping, and de novo assembly. *Genome research*, 29(5): 798–808.
- Watson J. D., Crick F. H. (1953): Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356): 737–738.
- Workman R., Timp W., Fedak R., Kilburn D., Hao S., Liu K. (2018): High molecular

weight DNA extraction from recalcitrant plant species for third generation sequencing. *ResearchSquare*.

- Wu R., Kaiser A. D. (1968): Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. *Journal of molecular biology*, 35(3): 523–537.
- Yu J., Hu S., Wang J., Wong G. K., Li S., Liu B., Deng Y., Dai L., Zhou Y., Zhang X., Cao M., Liu J., Sun J., Tang J., Chen Y., Huang X., Lin W., Ye C., Tong W., Cong L., Geng J., Han Y., Li L., Li W., Hu G., Huang X., Li W., Li J., Liu Z., Li L., Liu J., Qi Q., Liu J., Li L., Li T., Wang X., Lu H., Wu T., Zhu M., Ni P., Han H., Dong W., Ren X., Feng X., Cui P., Li X., Wang H., Xu X., Zhai W., Xu Z., Zhang J., He S., Zhang J., Xu J., Zhang K., Zheng X., Dong J., Zeng W., Tao L., Tan J., Ren X., Chen X., He J., Liu D., Tian W., Tian C., Xia H., Bao Q., Li G., Gao H., Cao T., Wang J., Zhao W., Li P., Chen W., WANG X., Zhang Y., Hu J., Wang J., Liu S., Yang J., Zhang G., Xiong Y., Li Z., Mao L., Zhou C., Zhu Z., Chen R., Hao B., Zheng W., Chen S., Guo W., Li G., Liu S., Tao M., Wang J., Zhu L., Yuan L., Yang, H. (2002): A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science (New York, N.Y.)*, 296(5565): 79–92.
- Zhang Q., Hu W., Zhu F., Wang L., Yu Q., Ming R., Zhang J. (2016): Structure, phylogeny, allelic haplotypes and expression of sucrose transporter gene families in *Saccharum*. *BMC Genomics*, 17(1).
- Zhang Z., Qiu W., Quinn S. J., Conigrave A. D., Brown E. M., Bai M. (2002): Three adjacent serines in the extracellular domains of the CaR are required for L-amino acid-mediated potentiation of receptor function. *The Journal of biological chemistry*, 277(37): 33727–33735.
- Zhao Q., Meng Y., Wang P., Qin X., Cheng C., Zhou J., Yu X., Li J., Lou Q., Jahn M., Chen J. (2021): Reconstruction of ancestral karyotype illuminates chromosome evolution in the genus *Cucumis*. *The Plant journal: for cell and molecular biology*, 107(4): 1243–1259.
- Zozomová-Lihová J., Mandáková T., Kovaříková A., Mühlhausen A., Mummenhoff K., Lysak M. A., Kovařík A. (2014): When fathers are instant losers: homogenization of rDNA loci in recently formed *Cardamine × schulzii* trigenomic allopolyploid. *The New phytologist*, 203(4): 1096–1108.

**Knihy:**

Brown T. A. (2016): Gene cloning & DNA analysis: an introduction. Wiley-Blackwell, University of Manchester, Manchester.

Van Straalen N. M., Roelofs D. (2011): Introduction to Ecological Genomics. Oxford University Press, Oxford, UK.

**Závěrečné práce:**

Juarez P. (2017): Regulatory mechanisms of mexEF-oprN efflux operon in *Pseudomonas aeruginosa*: from mutations in clinical isolates to its induction as response to electrophilic stress. Bacteriology, Université Bourgogne Franche-Comté.

**Internetové zdroje:**

Biomarker Technologies, [online] [navštíveno 10.3.2022]. Dostupné z:

<http://en.biomarker.com.cn/platforms/illumina>.

GenomeWeb, [online] [navštíveno 7.3.2022]. Dostupné z:

<https://www.genomeweb.com/sequencing/helicos-biosciences-files-chapter-11-bankruptcy-protection#.Ymbsl9pBzIU>

Illumina, [online] [navštíveno 10.3.2022]. Dostupné z:

<http://en.biomarker.com.cn/platforms/illumina>

Let's Talk Science, [online] [navštíveno 20.2.2022]. Dostupné z:

<https://letstalkscience.ca/educational-resources/backgrounders/sanger-sequencing>

MedGadget, [online] [navštíveno 10.3.2022]. Dostupné z:

<https://www.medgadget.com/2020/04/nanopore-sequencing-market-revolutionary-trends-2020-2026-by-industry-statistics.html>

Oxford Nanopore Technologies, [online] [navštíveno: 29.3.2022]. Dostupné z:

<https://nanoporetech.com/q20plus-chemistry>

PacBio, [online] [navštíveno: 28.3. 2022]. Dostupné z:

<https://www.pacb.com/technology/hifi-sequencing/>