University of South Bohemia in České Budějovice

Faculty of Science

Department of Parasitology

Bachelor Thesis

Optimization of CRISPR/Cas9 technique for the *Ixodes* ticks genome editing

Daniel Kitzberger

Supervisor: Jan Štěrba, RNDr., Ph.D.

Co-Supervisor: Kateryna Kotsarenko, Mgr., Ph.D.

Institute of Parasitology, Academy of Sciences of the Czech Republic

České Budějovice, 8.12.2020

Kitzberger, D., 2020: Optimization of CRISPR/Cas9 technique for the *Ixodes* ticks genome editing. Bc. Thesis, in English – 48p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

Annotation

CRISPR/Cas9 technique was used for genome editing of several arthropod species, but it has not yet been used on Ixodes spp. tick species. Therefore, in my work, I performed a bioinformatic analysis of the I. ricinus transcriptome and I. scapularis genome to reveal the sequence of the target gene for the CRISPR/Cas9-mediated knock-out. We prepared single guide RNA and mixed it with the Cas9 endonuclease to produce the Ribonucleoprotein complex and cleave the target sequence of the I. ricinus gene beta-galactoside alpha-2,6-sialyltransferase.

Declaration

I hereby declare that I have worked on the submitted bachelor thesis independently. All additional sources are listed in the bibliography section. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full form to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages.

Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

České budějovice,…… ……..

Daniel Kitzberger

Acknowledgement

I would like to thank Mgr. Kateryna Kotsarenko PhD. for patient guidance and supervision during the work in the laboratory and writing of the thesis. I would also like to thank RNDr. Jan Štěrba, PhD for allowing me to work on this topic and for professional supervision. I would also like to thank Pavlina Vechtova, PhD for the assistance and providing transcriptome data necessary for this thesis. Next, I would like to thank Prof. RNDr. Libor Grubhoffer CSc. for supporting this student program.

Lastly, I would like to thank my friends and family for their support.

Abstract

CRISPR/Cas9 is a brand-new genome editing tool, which was used for genome editing in many species. CRISPR/Cas9 is able to create heritable mutations in genes, thanks to its ability to create double strand breaks in DNA, which are then repaired with HDR or NHEJ for insertions or deletions, respectively. This technique was used for genome editing of several arthropod species, but it has not yet been used on *Ixodes* spp. tick species. Therefore, the aim of my work was to analyse the conditions for editing of *Ixodes ricinus* genome by using CRISPR/Cas9 editing technique. During my bachelor's work I mastered several new methods, including isolation and purification of DNA from ticks and cultured cell lines, transformation and cultivation of bacterial cells, plasmid isolation, and RNA synthesis. Also, I used several programs for the bioinformatic analysis of the nucleotide and protein sequences and the design of the plasmid constructs. I analysed some of *I. ricinus* and *I. scapularis* genes and chose one – beta-galactoside alpha-2,6-sialyltransferase, which can be used for the targeted editing by using the CRISPR/Cas9 system. I designed and produced the sequence of sgRNA, which targets the exon region of beta-galactoside alpha-2,6-sialyltransferase gene. As a result, ribonucleoprotein complex, consisting of Cas9 enzyme and sgRNA, was assembled and prepared for the cleavage of the target DNA sequence in *I. ricinus* genomic DNA.

Table of contents

1. Introduction

1.1. History of CRISPR/Cas9 technology

Clustered regularly interspaced short palindromic repeats (CRISPR) and their associated nuclease Cas9 (CRISPR-associated protein 9) is a precise and adaptive mechanism that enables bacterial and archaeal resistance against invading viruses, phages and plasmids (Barrangou, 2015). CRISPR/Cas9 technique is a new and currently the most promising gene editing tool available (Jinek et al., 2012). However, before CRISPR was discovered, other endonucleases and their complexes were used, mainly zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). Nevertheless, both methods have some limitations, their implementation is relatively expensive and there are possibilities for many off-target effects, leading to cellular toxicity (Gaj et al., 2013).

First CRISPRs were found over 30 years ago by Y. Ishino in *Escherichia coli* (Ishino et al., 1987). The function of these repeats was then not determined. Still, similar structures were found in archaea (1993), such as *Haloferax mediterranei* (F J M Mojica et al., 1993). In the early 2000s, a link between the structure of spacer regions and sequences of bacteriophages was found, showing a possible interpretation of CRISPR's function as a prokaryotic immune system (Francisco J M Mojica et al., 2005).

Research, which was explaining possible usage of CRISPR in genome editing, was published in 2012 (Jinek et al., 2012), and since then, CRISPR has been used in various vertebrate (Blitz et al., 2013; Hwang et al., 2013; Oishi et al., 2016; Rasys et al., 2019) as well as invertebrate species (Bassett & Liu, 2014; Y. Liu et al., 2014; Reegan et al., 2017). In 2020, developers of CRISPR/Cas9 genome editing tool, Emmanuelle Charpentier and Jennifer A. Doudna were awarded with the Nobel prize. The primary purpose of CRISPR/Cas9 in genome editing was a targeted knock-out or knock-in genes to determine their functions and influence on the given species (Joung et al., 2017; Salsman & Dellaire, 2017), but nowadays, this editing tool is used for many other applications, for example activation of gene expression, visualization of gene localization, epigenetic regulation of gene expression, etc.

1.2. Mechanism of CRISPR/Cas9 genome editing

Prokaryotes need a functional “immune system” to defend against viral inviders. However, such a complex function as immunity needs to be coded on comparatively short prokaryotic DNA within a single non-specialized cell. CRISPR/Cas9 system recognizes invading DNA and further neutralizes it. CRISPR is a locus of bacterial and archaeal DNA that consists of a set of spacers and repeats, and it works as a database of previously encountered viral or plasmidic infections. Every spacer is a target-specific DNA sequence, while every repeat is a comma between them (Barrangou, 2015). Spacer is a DNA sequence derived from DNA or RNA of a virus that has previously infected the cell. When the bacterial cell meets the infection for the first time, viral nucleic acids are cut up using unspecific Cas enzymes. The nucleic acid fragment is then stored in the CRISPR locus of DNA as a spacer. The size of the spacer usually ranges between 25 and 60 bp (Grissa et al., 2007).

Protospacer adjacent motif (PAM) sequences are the most notable elements in the target sequence (F. J.M. Mojica et al., 2009). There is the theory that PAMs are important not only for recognizing target sequences during viral interference, but also for spacer acquisition. For instance, spacers most often contain both PAM and protospacer sequences right after acquisition (Swarts et al., 2012). Different PAM sequences are recognized by different Cas proteins. For instance, NGG sequences are recognized by Cas9 enzyme in Type II CRISPR/Cas system, and CTT (Glediztsch et al, 2019) is recognized by Cas1 or Cas2 enzymes in Type I CRISPR/Cas system (Rath et al., 2015).

When the infection by a similar DNA or RNA fragment happens again, the stored DNA spacer is transcribed to crRNA (CRISPR RNA), which then reacts dependently on its complementary Cas enzyme (Rath et al., 2015). There are several Cas proteins, each of them forms a complex with the crRNA differently. There are 3 main types of the Ribonucleoprotein (RNP) complex formation between Cas enzyme and RNA, and multiple subtypes of each type (Fig. 1).

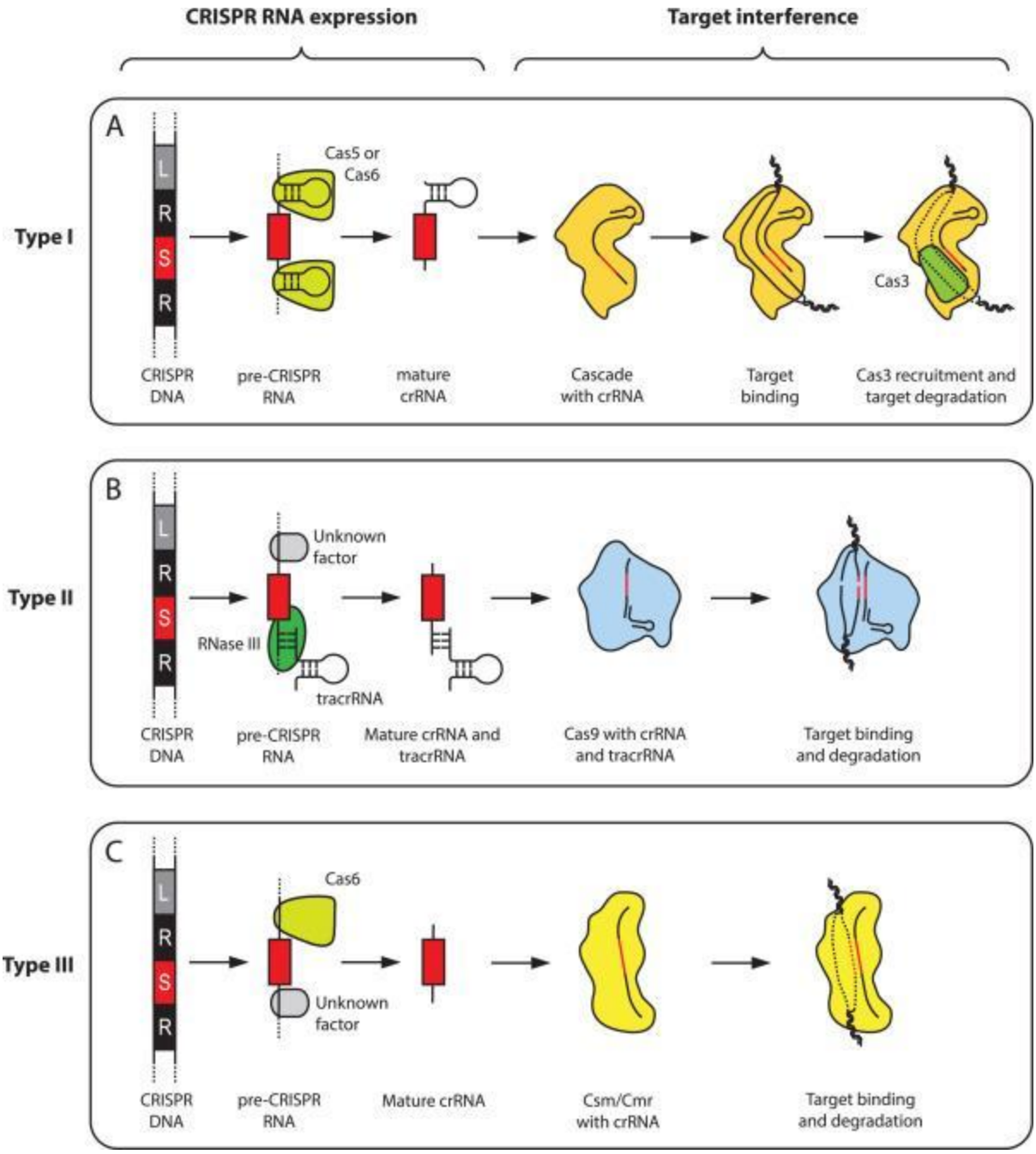


Figure 1. Types I, II and III of CRISPR/Cas systems (Rath et al., 2015).

Different subtypes vary from one another by the specific enzyme used for crRNA maturation or by spacer acquisition (Rath et al., 2015).

Type I includes a maturation of pre-crRNA using Cas5 or Cas6 protein (Figure 1A). The mature crRNA then binds to a Cascade (complex of multiple Cas proteins), which then binds to the Cas3 enzyme, which further performs the cut on the viral DNA (Bakr Shabbir et al., 2016).

Type II consists of binding the transcribed crRNA to a protein specific tracrRNA (Trans-activating CRISPR RNA), which then forms a gRNA (guide RNA) (Figure 1B). Cas9 enzyme binds the gRNA using the tracrRNA locus and then uses the crRNA to recognize invading nucleic acids and degrades them (Swarts et al., 2012).

Type III crRNA processing is performed by Cas6 enzyme and a so far unknown factor trimming the pre-crRNA. crRNA then binds to a Csm or Csr complex and targets invading DNA or RNA (Rath et al., 2015). Type III Cas proteins Cas10 are differentiated into Csm or Cmr complexes, which are similar to Cascade complexes (Rath et al., 2015). Csm-Cas10 subunit, which is subclassified as type IIIA; Cmr-Cas10 subunit, which is subclassified as type IIIB, depending on their specificity for DNA or RNA targets (Rouillon et al., 2013).

Recently a type IV system has been proposed, containing several Cascade genes, but with no Cas1 or Cas2 proteins, guided only by the protein-DNA interaction, without use of crRNA, working as a system with preset sequences to attack (Koonin & Krupovic, 2015).

1.3. CRISPR-associated proteins

In bacterial DNA, Cas enzymes are coded by Cas genes, which are placed downstream from the CRISPR repeat spacer sequences. As previously mentioned, there are several different enzymes with various functions for the bacterial immune system in different bacterial species (Jinek et al., 2012). Currently, only several Cas proteins are used for genome editing from all prokaryotic immune systems. Most notable of these are Cas9, Cas12a (formerly known as Cpf1), and Cas13 (formerly known as C2c2).

Cas9 is an RNA guided, site specific DNA binding protein, which can be programmed through use of sgRNA to specifically bind and cleave desired DNA sequence. It is found in several bacterial and archaeal species and SpCas9 (*Streptococcus pyogenes* Cas9) is currently the most widely used type (Francisco J.M. Mojica & Montoliu, 2016).

Cas9 protein consists of several domains, with each having a different function — the three main domains, which perform the tasks of locating and performing a double-strand break (DSB) in the target DNA site are HNH, RuvC, and PAM-interacting domains (Fig. 2).

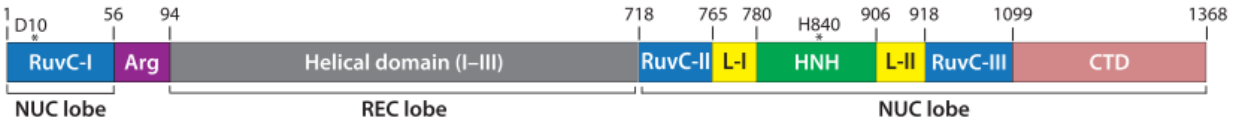


Figure 2. Domain structure of Cas9 protein structure, the PAM-interacting domain is localized near the C terminus of the NUC lobe (Jiang & Doudna, 2017).

The HNH domain cuts the strand of target DNA, which is complementary to the gRNA, while the RuvC domain performs a cut on the target DNA sequence, which is identical to the gRNA (Jiang & Doudna, 2017). The PAM-interacting domain recognizes the protospacer adjacent motifs on target sequence to which it stochastically binds until a match between the crRNA and the target sequence is found. In such a manner, the PAM-interacting domain significantly reduces the possible number of binding sites for the Cas9 protein. The PAM sequence for Cas9 is NGG, where N stands for any nucleotide (Sternberg et al., 2014).

Cas9 can be used to promote many targeted genome engineering applications. Wild type Cas9 has enabled targeted genome modification in a wide variety of species which were either difficult or impossible to affect through traditional means of genome editing, such as ZFNs and TALENs. The possibility to redirect Cas9 activity by simply changing the sgRNA allows for large scale experiments on gene functions and gene variants elucidation. Cas9 may further be modified into RNA guided homing devices (dCas9) by catalytic domain inactivation (P. Hsu, 2014).

Cas12 is a compact enzyme originally found in *Francisella novicida*, which creates staggered cuts in dsDNA (Yan et al. 2019). It relies on a thymine rich PAM sequence, which usually reads as TTTN or YTN, where Y means any pyrimidine base and N can mean any base. (F. Yan et al., 2019) The enzyme also shows higher precision during DNA cleavage than the popular Cas9 enzyme, as the Cas9 shows a strong bonding to DNA only after 8 base matches between crRNA and target. Cas12 on the other hand is strongly discriminatory against mismatches (Yan et al. 2019).

Cas13 is an RNA targeting Cas enzyme. After recognition of crRNA complementary ssRNA, the enzyme unleashes unspecific RNase activity, destroying all nearby RNA chains regardless of their complementarity to target sequence. It is considered for possible therapeutic use in gene expression regulation on RNA manipulation level (Yan et al. 2019). The enzyme is rather

precise in its ssRNA recognition as 1-2 mismatches on the spacer sequence were shown to significantly reduce RNA cleavage efficiency (Yan et al., 2019).

CRISPR/Cas9 system creates DSB in DNA, and then DNA has 2 possible repair mechanisms - homology directed repair (HDR) and non-homologous end joining (NHEJ). HDR requires a homologous sequence to guide the repair mechanism, while the NHEJ only uses small DNA sequences called microhomologies.

Homology directed repair is a naturally occurring DNA repair system, which can be used in many organisms to modify genomes. HDR is initiated by DSB presence (Liang et al. 1998), which can be created using CRISPR/Cas9 (Fig. 3).

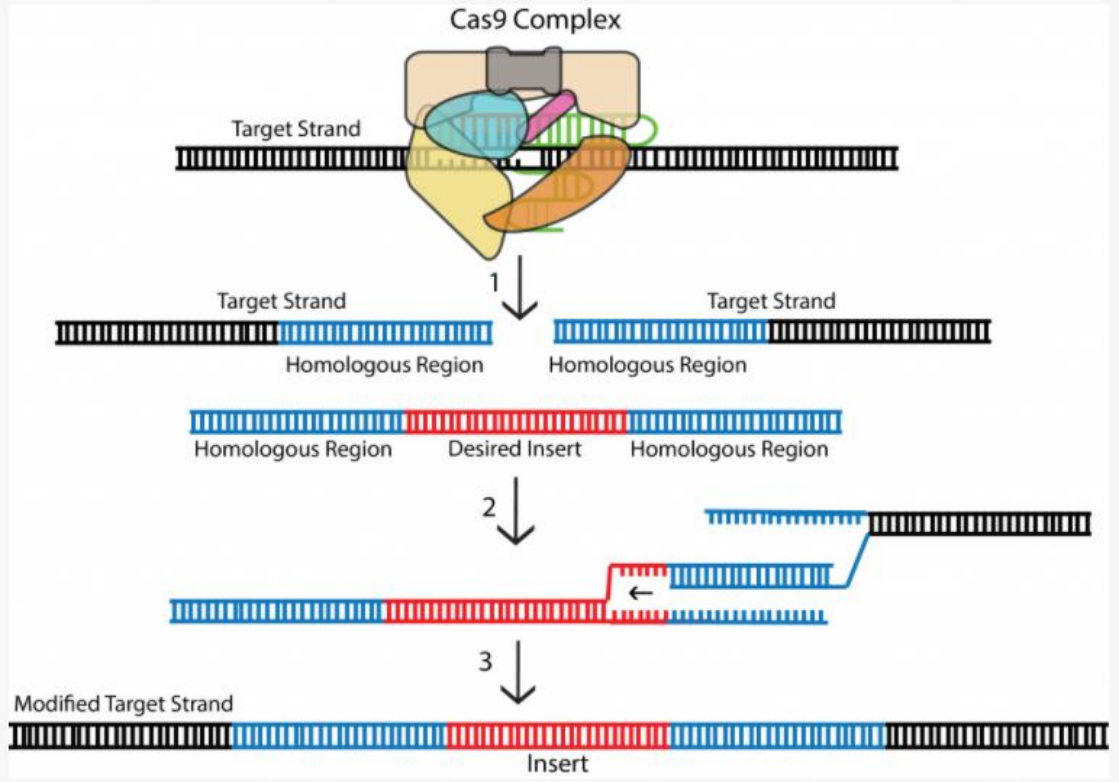


Figure 3. HDR mechanism. https://sites.tufts.edu/crispr/genome-editing/homology-directed-repair/)

Once the DNA is cleaved using a precise CRISPR/Cas9 RNA targeted system, the target cell can be given large quantities of the donor template, which contains desired insertion. The insertion must be flanked by DNA segments homologous to the cleavage site. The cellular DNA-

repair mechanism then inserts the desired sequence into the genome of the target with high precision (Mali et al. Feb 2013).

Non-homologous end joining is, when compared to HDR, prone to deletions and non-specific insertions leading to gene knock-outs (Fig. 4).

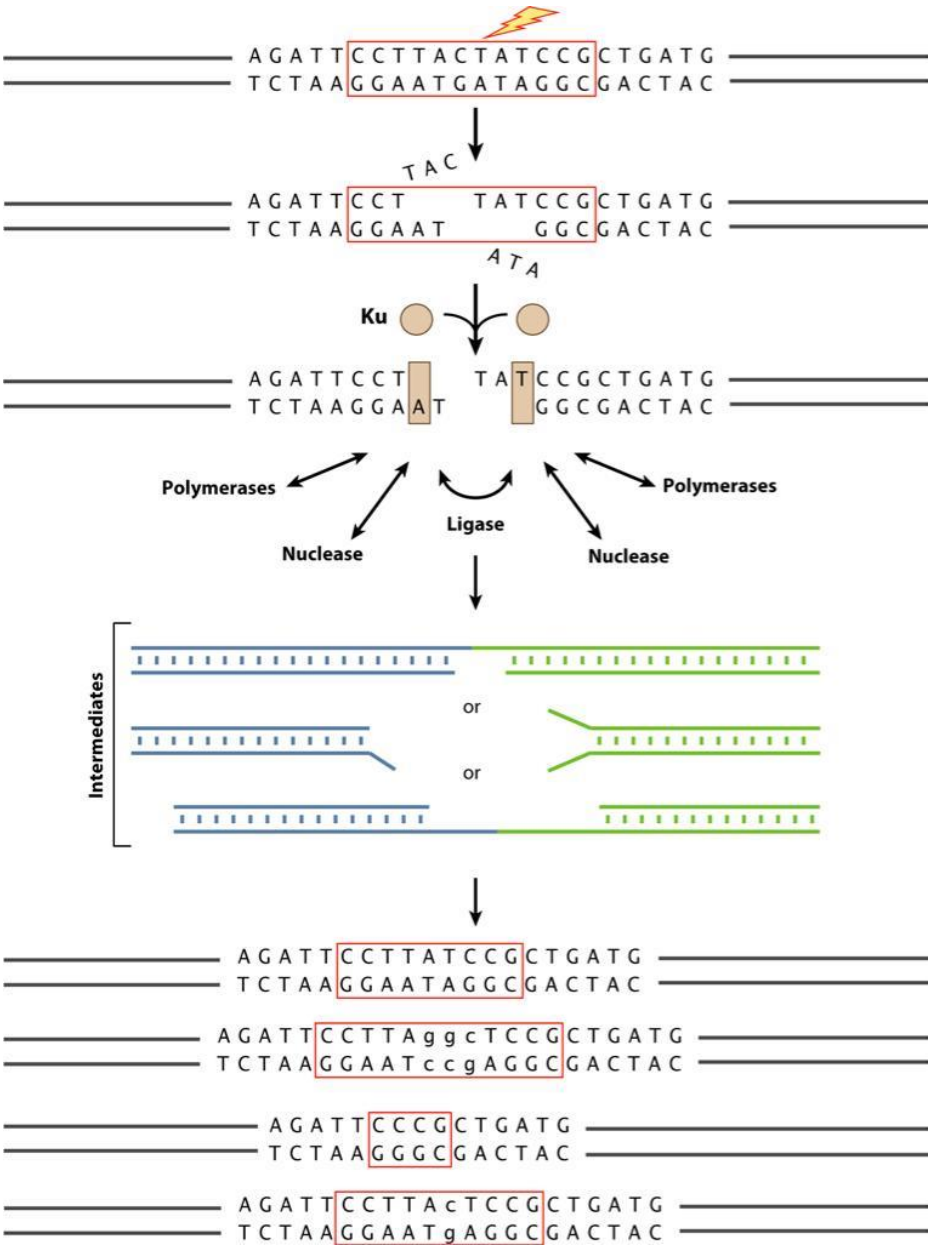


Figure 4. Mechanism of NHEJ repair of DNA

Compared to HDR, NHEJ does not require a homologous DNA template, which simplifies the gene editing process (Yan et al. 2020).

1. 4. Applications of CRISPR/Cas9 technique

As previously mentioned, the Cas proteins are coded downstream of the repeat spacer sequences on CRISPR locus. CRISPR/Cas systems I and III code multiple proteins in this area, which then assemble into complexes. Type I forms a Cascade complex, while the type III forms either Cmr or Csm complexes (Brouns et al., 2008). CRISPR/Cas type II on the other hand shows a significant reduction in the number of required proteins.

Cas9 itself comes from *Streptococcus pyogenes*, a pathogenic bacterial species from the Streptococci genus. In 2012 Jennifer Doudna and Emmanuelle Charpentier found a way to re-engineer a CRISPR/Cas9 system from a 4-component complex to a 2-component complex. They created a chimeric sgRNA from crRNA and tracrRNA, significantly reducing the difficulty of work with CRISPR/Cas9 complex (Jinek et al., 2013).

CRISPR/Cas9 is primarily used for knock-out or knock-in of a target gene. Knock-out is performed simply using Non-homologous end joining repair of double strand breaks, which leads to open reading frame disruption and thus disables the gene expression (Yan et al., 2020). Knock-in mutation is performed by supplying the cell with DNA repair template, containing the DSB containing sequence, as well as sequences upstream and downstream of the DSB, also known as right and left homology arms (Jacobi et al., 2017).

Homology directed repair has a rather low efficiency compared to the DSB formation, which leads to most of the DSBs to be repaired using NHEJ and thus deleting the target sites. There has recently been research studying ways to improve this efficiency and with it improving the knock-in success rates (M. Liu et al., 2019).

To overcome problems related to low efficiency of HDR, which is mediated by DSB after CRISPR/Cas9 editing, another method can be used. For instance, adenine base editors (ABEs) or cytosine base editors (CBEs) can be created by fusion of dCas9 (deactivated Cas9) with deaminases to perform the precise single-base mutations. This method doesn't show high efficiency now, but is currently being researched further, while searching for improvements of the method (Wang et al., 2020).

Due to dCas9 maintained ability to still bind to its sgRNA target sites, it may be used to repress transcription via blocking the initiation when bound to transcription start location. It has further been shown possible to fuse dCas9 with a repressor (CRISPR interference CRISPRi) or activator (CRISPRa) proteins (Qi et al., 2013). In a similar manner, CRISPR has been fused with epigenetic modifier to affect cytosine methylation in gene's promoter or by inducing histone acetylation or demethylation (Xie et al., 2018)

CRISPR/Cas has also been developed into a genome screening tool. Fusing dCas9 and GFP (Green fluorescent protein) or YFP (Yellow fluorescent protein) creates a customizable DNA labeler, which is compatible with fluorescence microscopy in living cells. This method was further upgraded using SunTag, which is a protein tagging system for signal amplification and mNeonGreen, a monomeric yellow-green fluorescent protein. The results proved superior than with common GFP or YFP (Ye et al., 2017).

Cas9 has also been re-engineered to create higher fidelity enzymes eSpCas9 and SpCas9-HF1 (Kleinstiver et al., 2016; Slaymaker et al., 2016). Two years later, HypaCas9 (Ikeda et al., 2019) enzyme has been created as well, with all 3 enzymes generating significantly smaller amounts of off-target mutations. Since then, other modified enzymes were engineered or found, such as evoCas9 (Casini et al., 2018).

To understand the role of DNA methylation in insects, dCas9 has been fused with ten-eleven translocation (TET) methylcytosine dioxygenase. This fusion protein can effectively demethylate DNA in *B. mori* species in a target specific manner (Liu et al. 2019).

1.5. Limitations of CRISPR/Cas9

Cas9 has been previously characterized as a mismatch-tolerating nuclease in a manner sensitive to the number, position and mismatch distribution. The findings indicate that while CRISPR shows high success rates in *in vitro* experiments, *in vivo* experiments show a high percentage of off-target cleavage and this would prove to be a problem in therapeutic application of CRISPR-Cas9.

To use RNA guided endonucleases (RGENs) in wide therapeutic and research applications, the understanding of off-target effects is essential (Fu et al., 2013; Hsu et al., 2013). The complications caused by off-target effects have already been subjected to extensive research in the past and several off-target cleavage affecting trends were shown.

Cas9 requires a high percentage of homology between crRNA and target sequence in order to perform the cleavage reaction. However, it can stay semi-transiently bound to a short sequence of complementary DNA. That would suggest that there is a large percentage of off-target binding sites, but a low percentage of off-target cleavage (Hsu, 2014; Wu et al., 2014).

Concentration of Cas9 enzyme is another important factor in off-target cleavage generation. Cas9 by itself is able to tolerate up to five mismatches inside the target site (Fu et al., 2013). The number of tolerated mismatches rises with concentration of Cas9 enzyme, leading to a higher off target activity (Hsu et al., 2013). Decreasing of Cas9 concentration then directly decreases the off-target mutations. The duration of Cas9 expression is another factor that seems to show changes in on to off-target ratio, but it is so far unclear how (Hsu, 2014).

Cas9 can be used to promote many targeted genome engineering applications. Wild type Cas9 has enabled targeted genome modification in a wide variety of species which were either difficult or impossible to affect through traditional means of genome editing, such as ZFNs and TALENs. The possibility to redirect Cas9 activity by simply changing the sgRNA allows for large scale experiments on gene functions and gene variants elucidation. Cas9 may further be modified into RNA guided homing devices (dCas9) by catalytic domain inactivation (P. Hsu, 2014).

1.6. CRISPR genome editing in arthropods

The first arthropod species, which was edited by the CRISPR-Cas9 system, was the insect *Drosophila melanogaster*. After this introduction to arthropods it has since been used in many other insect and non-insect arthropods, such as flies, mosquitos, moths, butterflies and others (Sun et al., 2017). *Drosophila melanogaster* has been used as a model insect for genome editing for decades now and success in the genome editing of *Drosophilae* has promoted use of CRISPR-Cas9 in other insects (Bassett & Liu, 2014).

There have been several approaches to induce mutagenesis in *Drosophila* species. First attempt consisted of coinjection of two plasmids into blastoderm embryo of *D. melanogaster* embryo. One plasmid expressing the sgRNA under Pol III promoter, the other plasmid expressing Cas9 gene under Hsp70 promoter. The experiment was overall successful, but has shown a low efficiency with only 5,9 % of injected flies producing a mutant offspring (Gratz et al., 2013).

Another approach was described as co-injection of mRNA, coding Cas9, and sgRNA into early stages of embryos (Bassett & Liu, 2014). High efficiency of this method was also confirmed by (Bassett et al., 2013; Yu et al., 2013).

The lower success rate of the method using plasmids is likely caused by expression levels of Cas9 and sgRNA or possibly by the timing of expression relative to specification of germ cells in the embryo (Bassett & Liu, 2014).

Recently, the possibility of regulation of gene expression by CRISPR/Cas9 editing was shown in *D. melanogaster* (Mendoza-García et al., 2017). Further, CRISPR/Cas9 mediated mutagenesis has been performed on insects *Aedes aegypti* (Taning et al., 2017), *Bombyx mori* (Xu et al., 2017; Zhang et al., 2015), *Spodoptera litura* (Bi et al., 2016; Zhu et al., 2016), *Plutella xylostella* (Douris et al., 2016; Huang et al., 2016), and other arthropod species from orders *Orthoptera, Coleoptera, Hymenoptera, Acarina, Decapoda*.

1.7. CRISPR genome editing in arachnids

When this thesis is being written, there has so far been only one article describing the use of CRISPR/Cas9 in arachnid species - in the plant-feeding spider mite *Tetranychus urticae* (Dermauw et al., 2020). Spider mites are herbivore species, able to feed on up to 1100 plant species, making them an important crop pest worldwide.

The research consisted of designing two sgRNAs based on the phytoene desaturase (tetur01g11270) from *T. urticae genome.* Cas9 was ordered as well as both sgRNA sequences. The CRISPR/Cas9 knock-out was first shown *in vitro* and further *in vivo*. The CRISPR/Cas9 was introduced into living *T. urticae* unfertilized females via microinjection of RNP complex.

The research has shown proof of concept, that CRISPR/Cas9 can be used in mite species to knock out genes *in vivo* by inducing two separate mutagenesis. This research shows a possible approach of using CRISPR/Cas9 in mites (Dermauw et al., 2020) and by extent possibly in all other chelicerates, including the tick species.

1.8. *Ixodes* spp. ticks

Ixodes spp. ticks are vectors for the various human pathogens, which cause such diseases as Lyme disease, tick-borne encephalitis, anaplasmosis, and others. *Ixodes* ticks have also shown to spread widely over the last few decades due to the urbanization of their natural habitats (Eisen, 2018; Rizzoli et.al.. 2014). Therefore, intensive study of the tick genome will allow scientists to control the physiological processes of the tick and, as a consequence, prevent the spread of ticks.

Genomes and transcriptomes of *Ixodes* spp. ticks are not fully sequenced, however, there are some publicly-available data in the GenBank database. The GenBank database contains the sequences of *I. scapularis* genome: accession ID: ABJB010000000 (Gulia-Nuss et al., 2016) and partial transcriptome: accession ID: GBBN01000001.1 (Egekwu et al., 2014), and *I. ricinus* transcriptomes: accession ID: GBIH00000000.1 (Kotsyfakis et al., 2015); accession ID: GCJO00000000.1 (Cramaro et al., 2015); accession ID: GANP00000000.1 (Schwarz et al., 2014), and newly published results of *I. ricinus* transcriptome, prepared by the group of Dr. Jan Sterba

(Vechtova et al., 2020) for different life stages of *I. ricinus*: eggs, larvae, partially fed nymphs, females, and partially fed females.

2. Aim

Aim of my bachelor work was to analyse the conditions for the editing of *Ixodes ricinus* genome by using CRISPR/Cas9 editing technique, determine the target gene, synthetize the site-specific sgRNA and prepare the Ribonucleoprotein complex for the *in vitro* digestion of the target gene.

3. Methods

3.1. Bioinformatic analysis

For the analysis of *I. ricinus* and *I. scapularis* DNA and RNA sequences we used programs BLASTn and BLASTp. CHOPCHOP program (https://chopchop.cbu.uib.no) was used for the design of sgRNA, primers and plasmid construction was performed using Geneious 10.0.5 software.

sgRNA sequence was designed by using of Geneious 10.0.5. Great attention was given to making sure that there is no other NGG sequence in the desired locus and there is low possibility of the off-target effects.

The synthesis of sgRNA was ordered from GenScript company. It was cloned it into pUC57 plasmid with SmaI restriction site downstream of the desired sequence. Restriction site allowed for later linearization of the plasmid and transcription of the target sequence.

3.2. Genomic DNA isolation from *Ixodes* cells

We used NucleoSpin Tissue DNA isolation kit (Macherey-Nagel, Duren, Germany) for genomic DNA isolation from the tick cell cultures and tick tissues.

Tick cell cultures were collected from the cultivation tubes, centrifuged and pellet was then lysed in a buffer with proteinase K, according to the NucleoSpin Tissue (Macherey-Nagel, Duren,

Germany) protocol. DNA was eluted by an elution buffer and concentration was measured on NanoPhotometer Pearl (Implen, Munchen, Germany).

3.3. PCR

We performed PCR for amplification of chosen loci of studied target genes: DAMT, DNMT1, DNMT3, Sia 2,6, and Sia 2,3 using the primers, listed in Table 1.

Table 1. PCR primers **A –** primers for preliminary PCR reactions;
B – PCR of sequences used for *in vitro* editing

A

Primer name	Sequence 5’ - 3’	T m, ^{o}C	Length of the product
DAMT_F	GGGTGGTTTGTATCTCATGAT	55	138
DAMT_R	CGCAGAATCCATTTGAAAAGG	56	
DNMT1_F	CGCAGCACAAGGTGACGC	60.5	140
DNMT1_R	GGCAGACGGGTCCTCC	59.5	
DNMT3_F	CACAAGCACCACAAGCACAAG G	62	222
DNMT3_R	GCTCGCCAGCAGGTCG	59.5	
Sia 2,6_F	GCACGAGGACCTGGGATGC	63	132
Sia 2,6_R	CCTCCGCCAAGAGGAAGACG	62.5	
Sia 2,3_F	TCGTCGGTTTCTCTGGACAC	60	144
Sia 2,3_R	CGTCGATCAGCCTGAAGGTT	60	

B

Primer name	Sequences 5’ – 3’	
SiaT_rt_F1	ACTTCCTCGACAGGGAGAACCC	1040
SiaT07911_I.ric_3R	CCTCCGCCAAGAGGAAGACG	
SiaT_rt_F2	GCACGAGGACCTGGGATGC	132
SiaT07911_I.ric_3R	CCTCCGCCAAGAGGAAGACG	

Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswitch, Massachusetts, USA) kit was used. The PCR mix was mixed on ice according to order and ratios, mentioned in Table 2.

Table 2. PCR component ratios

Component	Volume (µl)	Final Concentration
2mM dNTPs	2,5	200 µM
5X Q5 Reaction Buffer	5	1X
Nuclease-Free Water	13,75	
Q5 High-Fidelity DNA Polymerase	0,25	0,02 U/µl
10 µM Forward Primer	1,25	500nM
10 µM Reverse Primer	1,25	500nM
Template DNA	1	400ng

After thorough mixing, the prepared samples were inserted into the PCR machine, which was set for values, described in Table 3.

Table 3. PCR conditions

Step	Temp	Time
Initial Denaturation	98°C	30 s
Denaturation	98°C	
Annealing	Variable	Variable
Elongation	72°C	Variable
Final Elongation	72°C	2 minutes
Hold	4°C	Unlimited

3.4. Agarose gel electrophoresis

Agarose gels were prepared by dissolving agarose in a TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) in the microwave, cooled down to 50 °C and then cast into the specific tray. Well combs were then added to the tray and the gel is left to solidify in RT. After the gel polymerization the combs were taken out and the gel was put with the tray into the electrophoresis chamber. The chamber was filled with a TAE buffer and was ready to be loaded with samples and markers.

PCR products were mixed with 6x Loading dye (10 mM Tris-HCl,pH 7.6, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 60 % glycerol 60 mM EDTA) (ThermoFisher Scientific, Waltham, Massachusetts, USA), containing SYBR Green and loaded into the wells prepared in the gel. GeneRuler DNA Ladder (ThermoFisher Scientific, Waltham, Massachusetts, USA) with SYBR Green dye was loaded on the gels as well.

After the separation is finished, the gel is taken out of the chamber, placed under UV light and observed using the gel documentation system SynGene (Cambridge, United Kingdom).

3.5. Transformation of bacteria

Plasmid pUC57 with the target sequence insert was prepared by the GenScript company. Lyophilized plasmids then were diluted in sterile milliQ water to reach the concentration of 20 ng/μl. Diluted plasmid solution (2 μl) was added to NovaBlue Singles competent cells (Novagen, Merck, Darmstadt, Germany), which were then placed for 30 min on ice. Afterwards, the cells were heat shocked for 30s at 42 °C and incubated with 250 μl of SOC media at 37 °C for 1 hour. The cell suspensions (50 and 150 μl) were then spread on the LB agar with ampicillin. 24 hours after bacterial transformation, agar plates were examined, single colonies were extracted from the plates, placed into test tubes containing Luria-Bertani (LB) broth and let grow overnight in a 37 °C incubator at 200 rpm.

3.6. Plasmid isolation

Overnight cell suspension was centrifuged for 15 min at 4500 rpm, and pellet was used for plasmid isolation with the High Pure Plasmid Isolation Kit (Roche, Mannheim, Germany). The concentration and purity of extracted plasmids was measured using NanoPhotometer Pearl and samples were examined via agarose gel electrophoresis to ensure all extracted plasmids were of the correct size.

3.7. Plasmid linearization

Extracted plasmids were then digested with SmaI restriction enzyme according to the manufacturer's procedure, New England Biolabs (Ipswitch, Massachusetts, USA). Site of restriction 5'...CCC'GGG...3'/ 3'...GGG'CCC...5'. The digestion was performed for 1 hour at temperature 25 ^{o}C as follows.

Components were mixed in a total volume of 50µl. The digestion mix consisted of a 10x CutSmart buffer, isolated plasmid DNA, deionized DEPC-treated water and SmaI restriction enzyme. The reaction was conducted at 25 °C for 1 hour before enzyme deactivation at 65 °C.

Full plasmid digestion would then be checked via agarose gel electrophoresis, samples would then be excised from the gel and the DNA would be separated from the gel using NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Duren, Germany). The concentration of DNA was measured by using NanoPhotometer Pearl.

3.8. sgRNA synthesis

sgRNA was synthesized from the pUC57 plasmids using HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs, Ipswich, MA, USA).

The synthesis reaction mix contained linearized plasmids as template DNA, T7 polymerase and reaction buffer (Table 4).

Table 4. sgRNA synthesis reaction mixture

Nuclease free water	up to 20 µl
NTP r mix	8µl (10mM each NTP final)
Template DNA	(1 µg)
T7 RNA polymerase mix	2µl
Total reaction volume	20µl

The reaction mix was then thoroughly mixed using a microcentrifuge and then incubated at 37°C for 2 hours. Firefly luciferase gene was used as a DNA template for the positive control reaction. Synthesized RNA was then DNAse treated to remove DNA template.

The sgRNA was afterwards purified according to an RNA blue protocol from Top-bio (Vestec, Czech Republic) with some modifications. The sgRNA was first mixed with RNABlue (1ml) reagent and incubated at RT for 5 minutes and afterwards chloroform (0.2ml) was added. The test tube was shaken for 15s and incubated at RT for another 5 minutes. The sample was then centrifuged at 12000g for 10 minutes at 4°C. The aqueous phase was transferred into a new test tube, where the RNA was precipitated from the solution using isopropanol (0.5ml). The sample was then incubated at 4°C for 10 minutes and centrifuged at 12000g for 10 minutes at 4°C. The supernatant was removed and RNA was cleaned using 75% EtOH (1ml), shaken and centrifuged at 12000g for 5 minutes at 4°C. The sample was then air-dried for 7 minutes and RNA was then dissolved in DEPC-treated H_2O under increased temperature for better dissolution. The proper size of the product was then checked using polyacrylamide gel electrophoresis with urea, according to protocol from Denaturing urea polyacrylamide gel electrophoresis (Summer et. al. 2009). The RNA ladder used was RiboRuler High Range RNA Ladder from Thermo Fisher Scientific, Waltham, Massachusetts, USA

3.9. *In vitro* gDNA digestion by using RNP complex

The newly synthesized and purified sgRNA was mixed with the Cas9 protein (M0386) purchased from New England Biolabs (Ipswich, MA, USA) and other required components (Table 5) according to the provider's procedure with slight modifications.

Table 5. *In Vitro* DNA digestion with Cas9 RNP

Component	Volume
Water	33,3µl
NEBuffer 3.1 (10x)	5µl
RNA (300nM)	5µl
Cas9 (900nM)	1,7µl
Incubate at 25°C for 10 minutes	
DNA (30nM)	5 µl
Total reaction volume	50µl

The mixture was afterwards mixed and pulse-spun in a microcentrifuge and incubated at 37°C for 15 minutes. Afterwards 1µl of Proteinase K was added and the mixture was pulse-spun again. The mixture was then incubated at RT for 10 minutes and the DNA fragments were analysed using agarose gel electrophoresis.

4. Results

4.1. Presence of the target genes in the genomic DNA of *Ixodes* ticks

We analyzed several genes from *I. ricinus* transcriptome, provided by the group of Dr. Jan Sterba and compared their sequence with the publicly available data of *I. scapularis* genome. Thus, we decided to analyze the following genes: DAMT (MH926033), DNMT1 (MH795945), DNMT3 (MH926034), beta-galactoside alpha-2,6-sialyltransferase (HM480299), and beta-galactoside alpha-2,3-sialyltransferase (EEC07371.1). First, we isolated genomic DNA from *I. scapularis*

tissues and cell lines ISE6 and ISE18 and then we performed PCR to confirm the presence of the chosen genes, the expected sizes of products were around 130 bp. (Fig. 5).

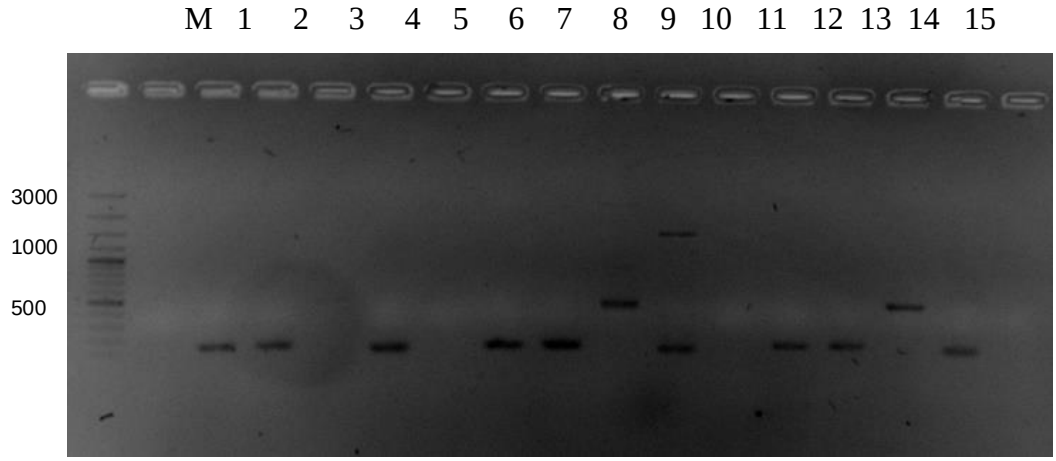


Figure 5. Results of PCR analysis of the presence of the target genes in the genomic DNA of *I. scapularis* ticks and their cell lines: M - 100 bp DNA ladder, *I. scapularis:* 1 - DAMT, 2 - DNMT1, 3 - DNMT3, 4 - Sia 2,3, 5 - Sia2,6; cell line ISE 6: 6 - DAMT, 7 - DNMT1, 8 - DNMT3, 9 - Sia 2,3, 10 - Sia 2,6; cell line ISE 18: 11 - DAMT, 12 - DNMT1, 13 - DNMT3, 14 - Sia 2,3, 15 - Sia 2,6. The picture is here presented as a negative for better contrast.

Results of PCR analysis revealed that gDNA of *I. scapularis* contains the sequence of genes DNMT1, DNMT3, and Sia 2,6. However, we did not identify the PCR products with the primers to DAMT.. Moreover, in the case of sample 10 (PCR product from ISE6 cell line gDNA with the primers to Sia 2,6 gene) we observed two bands on the agarose gel with the expected size and band with the size approx. 2000 bp. However, in the gDNA from *I. scapularis* and cell line ISE18, only PCR products with correct length (132 bp) were detected. The difference in the PCR products between cell lines ISE6 and ISE18 can be possibly explained by the karyotype differences between long-time cultivated tick cell lines, as it was described previously (Kotsarenko et al., 2020), or

unspecific primers binding. Nevertheless,e chose the gene Sia 2,6 for further study and analysed the presence of this gene in the gDNA of *I. ricinus* ticks. We used two pairs of primers for PCR (Table 1 B), which produce products with size 132 bp and 1040 bp long (Fig. 6).

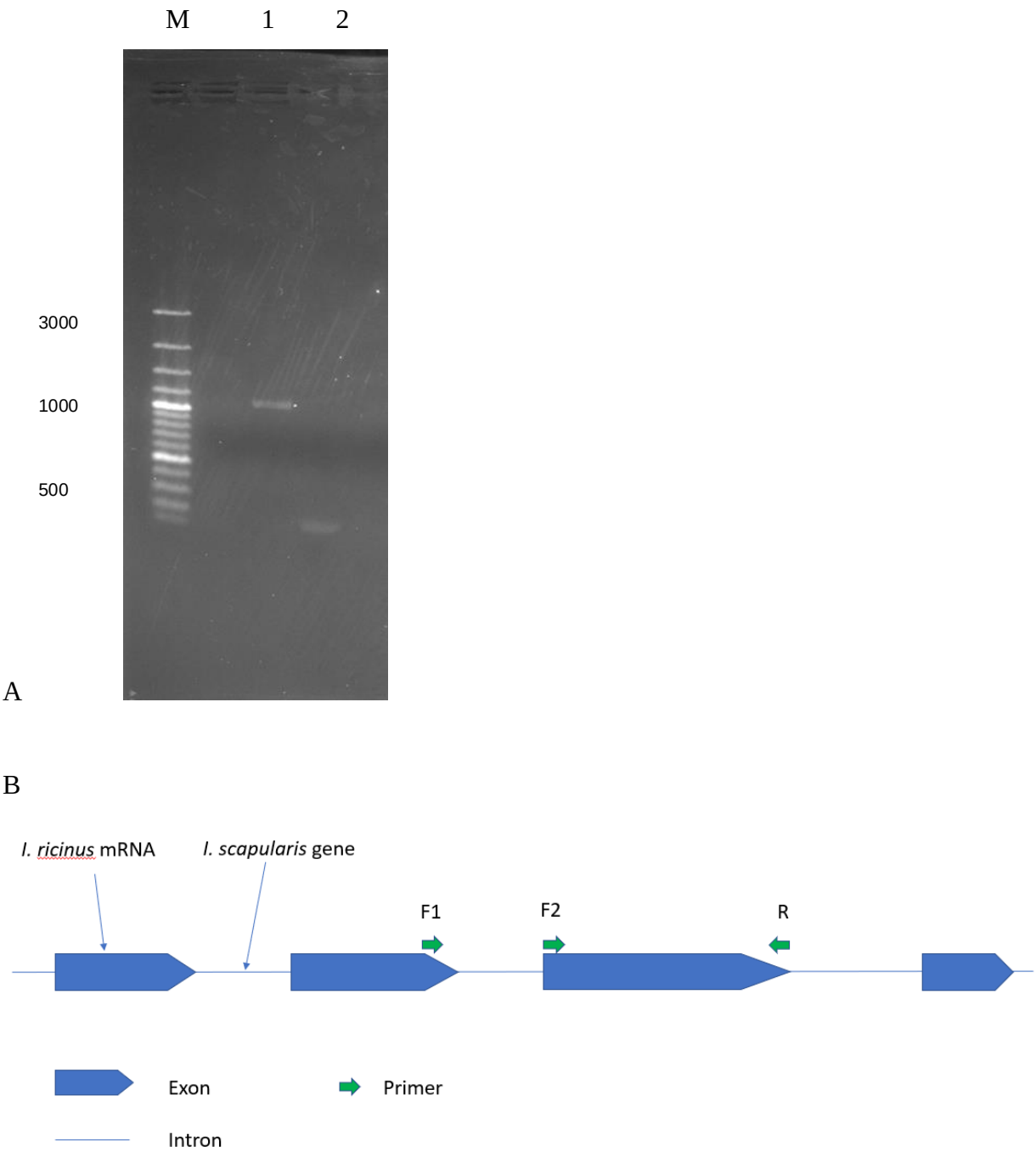


Figure 6. Presence of the Sia 2,6 gene in the gDNA of *I. ricinus* ticks. A. PCR amplification of **1** - long (1040 bp) and **2** - short (132 bp) fragments of Sia 2,6 in the gDNA of *I. ricinus*. B. The scheme of the exon-intron region of Sia 2,6 gene. The sequence of Sia 2,6 mRNA (*I. ricinus*) was

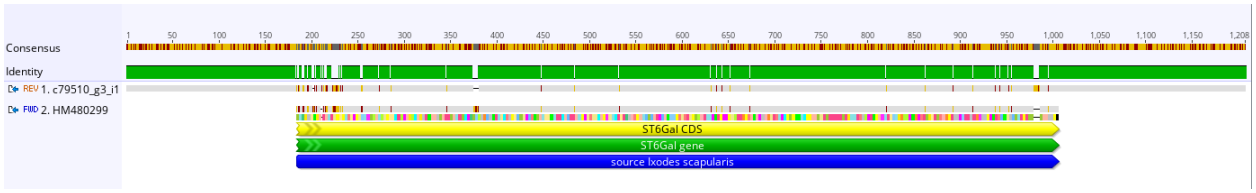
aligned with the sequence of Sia 2,6 gene (*I. scapularis)* to distinguish exon-intron boundaries. Primers are shown as green arrows.

We analysed the results of PCR on the 1 % agarose gel and found the PCR products with expected size. Therefore, we decided to use a sequence of the third exon of Sia 2,6 gene (*I. ricinus*) for the design of sgRNA.

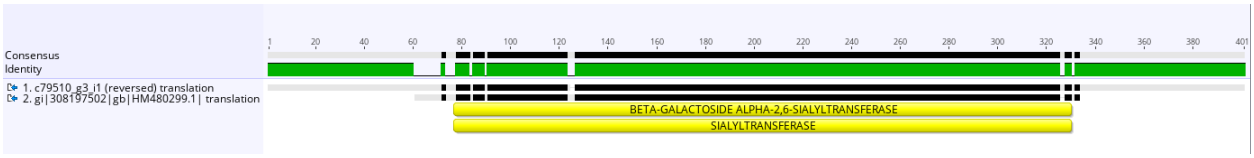
4.2. Design of sgRNA

We compared the sequence of the Sia 2,6 gene from the *I. scapularis* genome (HM480299) and *I. ricinus* transcriptome (contig 79510_g3_i1, provided by Dr. Pavlina Vechtova) (Fig. 7).

A



B



C

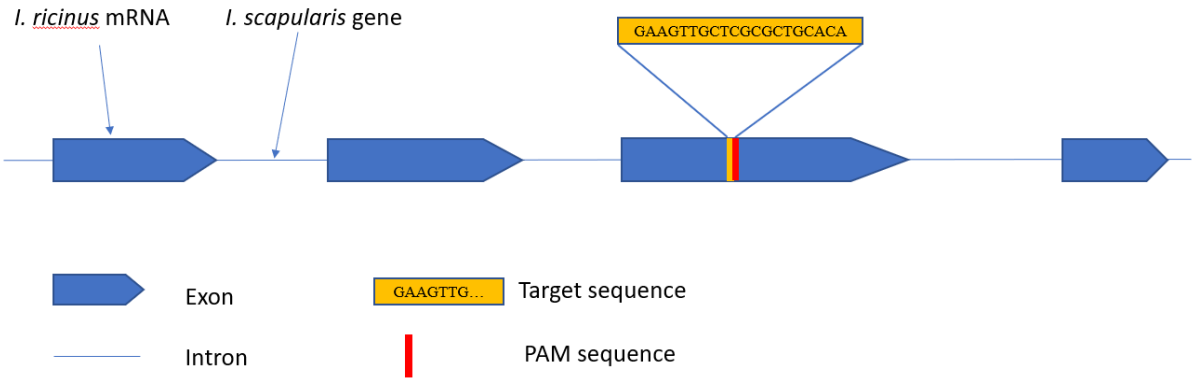


Figure 7. Sequence of Sia 2,6 gene. Alignment of coding sequence (CDS) (A) and translated sequence (B) of Sia 2,6 from *I. scapularis* (HM480299) and *I. ricinus* (contig 79510_g3_i1, provided by Dr. Pavlina Vechtova). The scheme of the exon-intron region of Sia 2,6 gene and the target sequence for sgRNA-Cas9 recognition. The sequence of Sia 2,6 mRNA (*I. ricinus*) was

aligned with the sequence of Sia 2,6 gene sequence of (*I. scapularis)* to distinguish exon-intron boundaries.

Amino acid sequences of both *I. ricinus* and *I. scapularis* Sia 2,6 genes showed high similarity (Fig. 7B), however, nucleotide sequences differed in several nucleotides. Therefore, we focused on the sequence of Sia 2,6 from *I. ricinus* tick and designed sgRNA which targeted the third exon (Fig. 7C).

Finally, we determined the sequences of sgRNA for beta-galactoside alpha-2,6-sialyltransferase gene, confirmed that this sequence has the lowest off-targets number using CHOPCHOP program and used it for further cloning into a pUC57 plasmid (Fig. 8), that was ordered from GenScript company.

A

TTCTAATACGACTCACTATAGGAAGTTGCTCGCGCTGCACAGTTTAGAGCTAGAAAT AGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTT

B

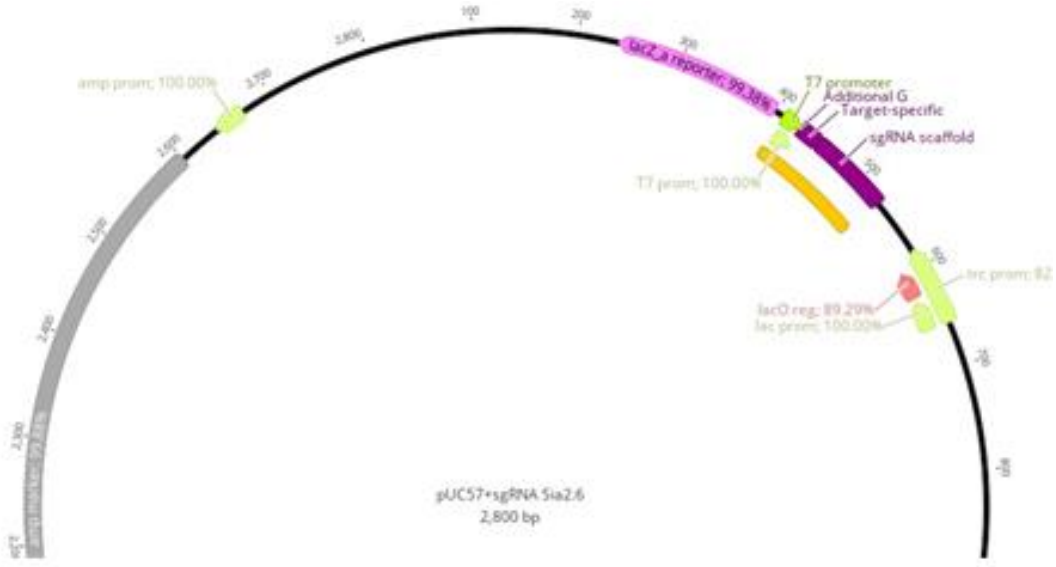


Figure. 8. sgRNA for the gene beta-galactoside alpha-2,6-sialyltransferase, *I.ricinus*. A. The sequence of sgRNA: orange colour for the T7-promoter sequence, blue colour for target specific part of sgRNA, and red colour for sgRNA scaffold. B. The scheme of pUC57 plasmid containing the T7 promoter, sgRNA sequence and sgRNA scaffold.

4.3. Plasmid isolation and linearization

NovaBlue single cells were transformed with the plasmid pUC57+ sgRNA to amplify the plasmid. Then we isolated the plasmid from the bacteria and linearized it using site specific enzyme. Fig. 8).

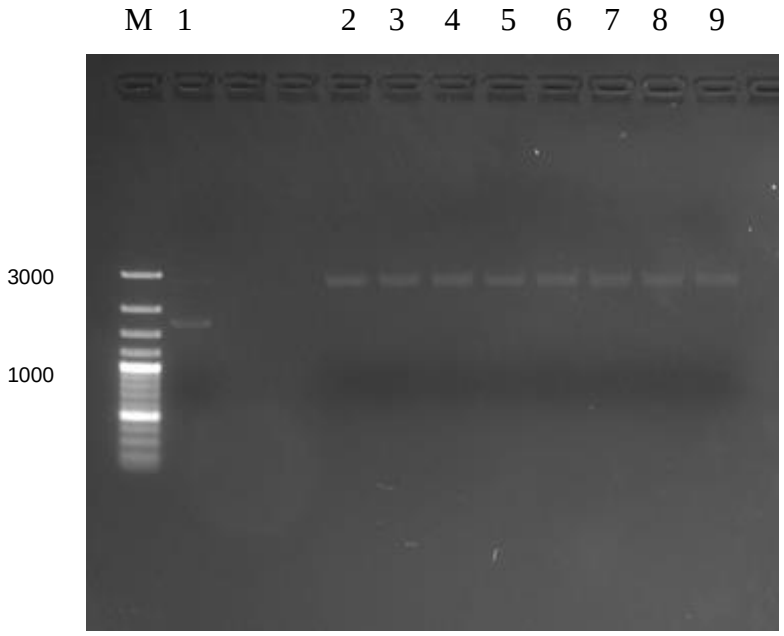


Figure. 9. Agarose gel electrophoresis of the SmaI digested plasmid pUC57+sgRNAn. 1 - undigested plasmid, 2-9 - linearized plasmid.

We linearized the plasmid to produce RNA transcript of a defined length. For this we used SmaI restriction enzyme, which recognizes 5'...CCC'GGG...3'/ 3'...GGG'CCC...5'and generates

blunt ends downstream of the insert to be transcribed. Circular plasmid templates will generate long heterogeneous RNA transcripts in higher quantities because of high processivity of T7 RNA polymerase, therefore, circular plasmid is not suitable template for the sgRNA synthesis.

4.4. Synthesis of gRNA and formation of RNP complex

After plasmid linearization, sgRNA was transcribed overnight using HiScribe T7 RNA Polymerase. Concentrations of the synthesizes RNA are presented in Table 6.

Table 6. Concentration of synthesized RNA

Sample	Concentration ng/µl	280/260	260/230
sgRNA	2834	2.057	2.508
Positive control - RNA of Firefly luciferase gene	1248	2.039	1.559

After synthesis of sgRNA we purified of RNA according to the protocol Purification of Synthesized RNA (E2040) from NEB (https://international.neb.com/protocols/0001/01/01/purification-of-synthesized-rna-e2040), in order to remove proteins, salts and most of the free nucleotides left from RNA synthesis. Then we measure concentration and purity of the RNA again (Tab. 7).

Table 7. Concentration of the synthesized RNA after purification

Sample	Concentration ng/µl	280/260	260/230
Positive control - RNA of Firefly luciferase gene	171	2.023	1.559
sgRNA for Sia 2,6 gene (sample 1)	98	2.000	2.174

sgRNA for Sia 2,6 gene (sample 2)	814	1.986	2.196
sgRNA for Sia 2,6 gene (sample 3)	1000	2.018	2.429

Then we performed the denaturing urea polyacrylamide gel electrophoresis of the synthetized and purified RNA to determine the size and purity of the samples (Fig. 10).

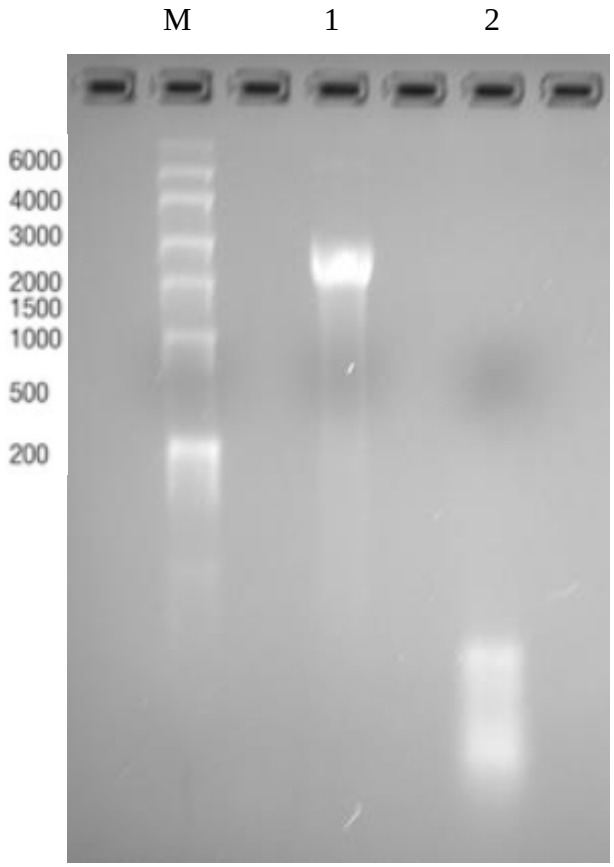


Figure 10. Denaturing urea polyacrylamide gel electrophoresis of the synthesized RNA of Firefly luciferase gene and sgRNA: 1 - RNA ladder, 2 - RNA of Firefly luciferase gene, 3 - sgRNA for Sialyltransferase gene (sample 2).

We observed the clear band of the control RNA, synthetized from Firefly luciferase gene, and several bands, which looked like smear on the gel, belonging to the synthetized sgRNA.

However, the expected length of the sgRNA was 100 bp. Therefore, we decided to compare these data with the measurement of sgRNA concentration and resolution on the gel using 2100 Bioanalyzer (Agilent, USA). The measurement on Bioanalyzer was performed by Dr. Pavlina Vechtova using the microfluidic chip specific for RNA (Fig. 11).

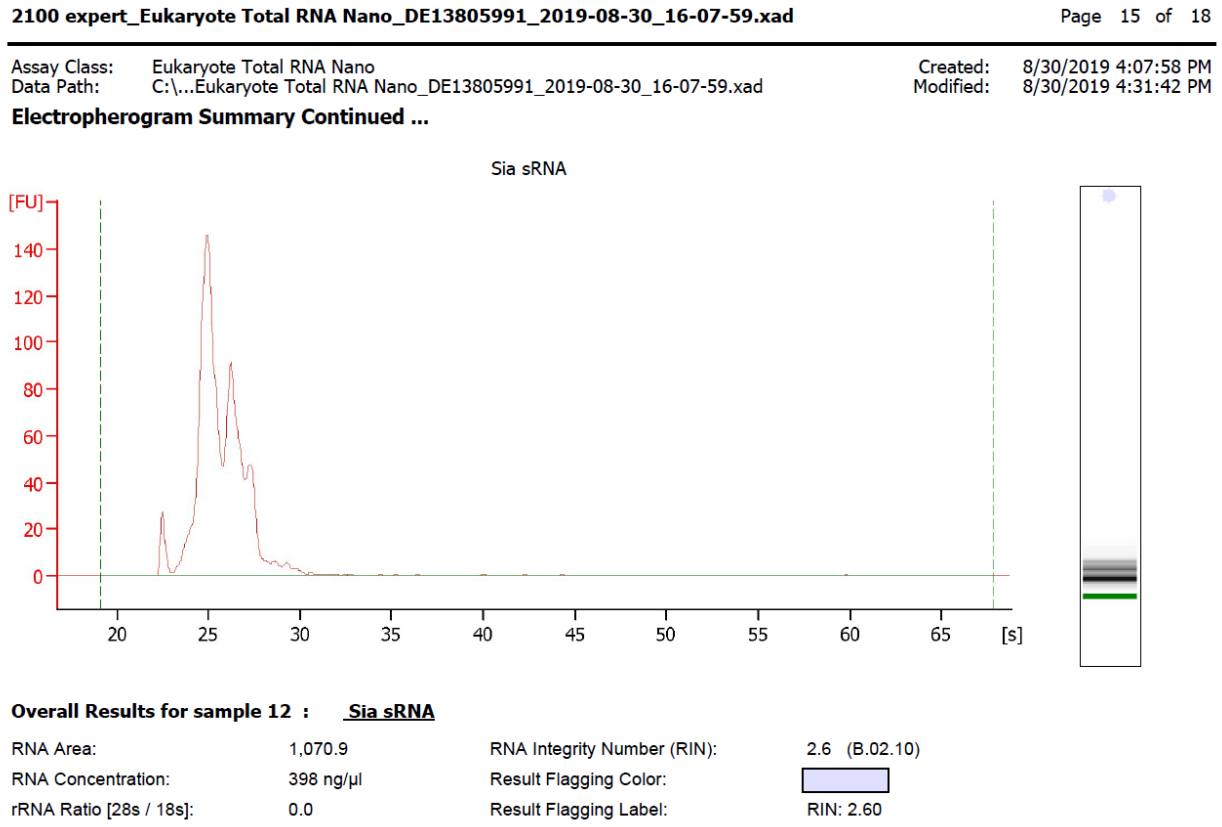


Figure 11. Measurement of the concentration of synthesized sgRNA for Sia 2,6 gene (sample 2).

A similar RNA electrophoregram was shown for both the denaturing urea polyacrylamide gel and the microfluidic chip. Thus, we found that synthesis of sgRNA using HiScribe T7 kit resulted in multiple RNA products with different length. Using Bioanalyzer, we also observed a lower RNA concentration, than it was measured by NanoPhotometer (Table 7). However, we expect that this will not have an influence on the efficiency of RNP complex formation.

4.5. Digestion of DNA *in vitro* using RNP complex

We prepared RNP complex by mixing of Cas9 enzyme with the purified sgRNA (sample 2) for the digestion of 132 bp DNA target sequence (Fig. 6). Due to technical circumstances (the Loading buffer with SYBR Green was probably contaminated and we did not observe a signal in either the DNA samples or the DNA ladder) we did not obtain the results for the DNA digest. Therefore, we prepared the schematic image of the expected results (Fig. 12).

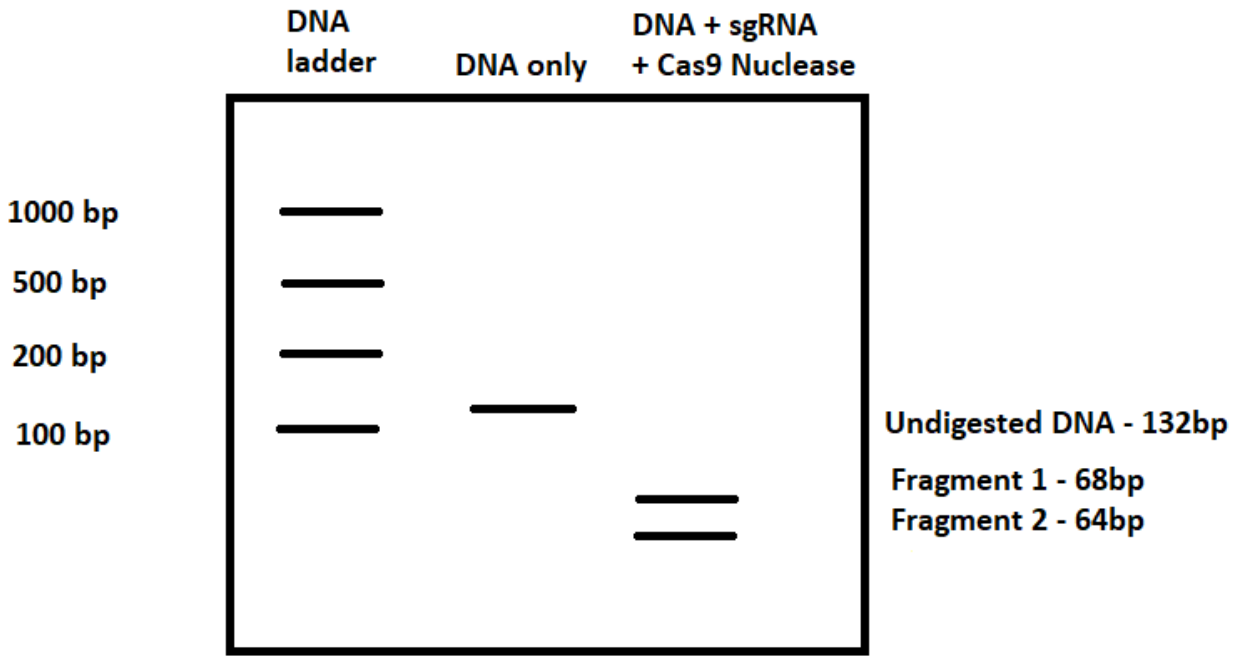


Figure 12. Expected results of the DNA *in vitro* digestion.

We expect to have digested DNA in the site 5’ NNNNGAAGTTGCTCGCGCTGC/ACA TGGNNN 3’. Due to the absence of a repair system in *in vitro* conditions, DNA will not be repaired and, therefore, we Cas9 should cut dsDNA into two pieces with the length 68 bp and 64 bp.

5. Discussion

The aim of my work was to determine and optimize the conditions for the use of the CRISPR/Cas9 system in *Ixodes* spp. ticks and attempt the first *in vitro* genome editing in the *I. ricinus* species. The absence of a fully sequenced genome for *I. ricinus* and transcriptome for *I. scapularis* made the choice of target gene for the CRISPR-Cas9 editing rather problematic, as this system is very precise, and its efficiency depends on the off-target effects. Moreover, the target cutting site should be located within the protein-coding region (exon), to have the effect on the protein synthesis due to the reading frame shift. Therefore, we compared *I. ricinus* transcriptome with *I. scapularis* genome to find the exon-intron boundaries. The target region also had to show a low off-target effect for sgRNA. Not all sgRNA prediction programs contain the data on *Ixodes* genomes, so they cannot be used for the prediction of the off-targets. Program CHOPCHOP contains the *I. scapularis* genome data and we used this program to predict the possible off-targets for both *I. scapularis* and *I. ricinus* genes. Several *I. scapularis* genes were aligned against *I. ricinus* transcriptome, until several genes of interest (DAMT, DNMT1, DNMT3, Sia 2,3 and Sia 2,6) were chosen, from which beta-galactoside alpha-2,6-sialyltransferase (primers for which were provided by Pavlina Vechtova PhD.) was chosen as the best match for our work, due to its presence in all ISE cell lines, *I. scapularis* and *I. ricinus* ticks.

In the conserved domain of the gene a 20nt sequence before PAM (NGG) was chosen and was ordered from NEB, cloned into a pUC57 plasmid with sgRNA scaffold. The RNA was transcribed from the plasmid and purified. We ordered a standard Cas9 protein from NEB to introduce DSBs into the gene of interest. In the case of *in vitro* digestion of DNA, we expected to see two bands on the gel, which correspond to two pieces of DNA after cutting before the PAM sequence. However, in the case of *in vivo* editing, the gene would be repaired via NHEJ or HDR repair mechanisms and cells will receive mutation at the place of double strand break. We prepared an RNP complex of Cas9 enzyme with sgRNA to cut the target DNA sequence *in vitro*, however, due to some technical circumstances we did not receive the results of these experiments. Nevertheless, our calculations and preparation of sgRNA for *I. ricinus* ticks will be useful for other researchers in the Laboratory of Applied Biochemistry, JCU, which works on the genome study of *Ixodes* spp. ticks.

6. Conclusions

Using the CRISPR/Cas9 genome editing technique for the *Ixodes ricinus* ticks is limited due to the absence of a fully-sequences genome for *I. ricinus.* Nevertheless, beta-galactoside alpha-2,6-sialyltransferase was chosen as the best candidate gene for the target digestion by Cas9 enzyme. Also, we designed the sgRNA sequence, determined the target site in the exon of the gene and calculated all conditions for the RNP complex formation and CRISPR/Cas9 editing of the *I. ricinus* gene.

7. List of references

Barrangou, R. (2015). The roles of CRISPR–Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology*, *32*, 36–41. https://doi.org/https://doi.org/10.1016/j.coi.2014.12.008

Bassett, A. R., & Liu, J.-L. (2014). CRISPR/Cas9 and genome editing in *Drosophila*. *Journal of Genetics and Genomics*, *41*(1), 7–19. https://doi.org/10.1016/j.jgg.2013.12.004

Bassett, A. R., Tibbit, C., Ponting, C. P., & Liu, J. L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Reports*, *4*(1), 220–228. https://doi.org/10.1016/j.celrep.2013.06.020

Bi, H. L., Xu, J., Tan, A. J., & Huang, Y. P. (2016). CRISPR/Cas9-mediated targeted gene mutagenesis in *Spodoptera litura*. *Insect Science*, *23*(3), 469–477. https://doi.org/10.1111/1744-7917.12341

Blitz, I. L., Biesinger, J., Xie, X., & Cho, K. W. Y. (2013). Biallelic genome modification in F(0) *Xenopus tropicalis* embryos using the CRISPR/Cas system. *Genesis (New York, N.Y. : 2000)*, *51*(12), 827–834. https://doi.org/10.1002/dvg.22719

Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science.* 2008 Aug 15;321(5891):960-4. doi: 10.1126/science.1159689. PMID: 18703739; PMCID: PMC5898235.

Casini A, Olivieri M, Petris G, Montagna C, Reginato G, Maule G, Lorenzin F, Prandi D, Romanel A, Demichelis F, Inga A, Cereseto A. A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat Biotechnol.* 2018 Mar;36(3):265-271. doi: 10.1038/nbt.4066. Epub 2018 Jan 29. PMID: 29431739; PMCID: PMC6066108.

Cramaro, W. J., Revets, D., Hunewald, O. E., Sinner, R., Reye, A. L., & Muller, C. P. (2015). Integration of *Ixodes ricinus* genome sequencing with transcriptome and proteome annotation of the naïve midgut. *BMC Genomics*, *16*(1), 1–15. https://doi.org/10.1186/s12864-015-1981-

Dermauw, W., Jonckheere, W., Riga, M., Livadaras, I., Vontas, J., & Van Leeuwen, T. (2020). Targeted mutagenesis using CRISPR-Cas9 in the chelicerate herbivore *Tetranychus urticae*. *Insect Biochemistry and Molecular Biology*, *120* , 103347. https://doi.org/10.1016/j.ibmb.2020.103347

Douris, V., Steinbach, D., Panteleri, R., Livadaras, I., Pickett, J. A., Van Leeuwen, T., Nauen, R., & Vontas, J. (2016). Resistance mutation conserved between insects and mites unravels the benzoylurea insecticide mode of action on chitin biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(51), 14692–14697. https://doi.org/10.1073/pnas.1618258113

Eisen RJ, Eisen L. The blacklegged tick, *Ixodes scapularis*: An increasing public health concern. *Trends Parasitol*. 2018;34(4):295-309. doi:10.1016/j.pt.2017.12.006

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *International Review on Modelling and Simulations*, *4*(2), 851–857. https://doi.org/10.1038/nbt.2623.High

Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, *31*(7), 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004

Gleditzsch D, Pausch P, Müller-Esparza H, Özcan A, Guo X, Bange G, Randau L. PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. RNA Biol. 2019 Apr;16(4):504-517. doi: 10.1080/15476286.2018.1504546. Epub 2018 Sep 18. PMID: 30109815; PMCID: PMC6546366.

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J., & O'connor-Giles, K. M. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, *194*(4), 1029–1035. https://doi.org/10.1534/genetics.113.152710

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M., & Oconnor-Giles, K. M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics*, *196*(4), 961–971. https://doi.org/10.1534/genetics.113.160713

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, *35*(SUPPL.2), 52–57. https://doi.org/10.1093/nar/gkm360

Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014 Jun 5;157(6):1262-1278. doi: 10.1016/j.cell.2014.05.010. PMID: 24906146; PMCID: PMC4343198.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G., & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*(9), 827–832. https://doi.org/10.1038/nbt.2647

Huang, Y., Chen, Y., Zeng, B., Wang, Y., James, A. A., Gurr, G. M., Yang, G., Lin, X., Huang, Y., & You, M. (2016). CRISPR/Cas9 mediated knockout of the abdominal-A homeotic gene in the global pest, diamondback moth (*Plutella xylostella*). *Insect Biochemistry and Molecular Biology*, *75*, 98–106. https://doi.org/10.1016/j.ibmb.2016.06.004

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J., & Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology*, *31*(3), 227–229. https://doi.org/10.1038/nbt.2501

Ikeda, A., Fujii, W., Sugiura, K., & Naito, K. (2019). High-fidelity endonuclease variant HypaCas9 facilitates accurate allele-specific gene modification in mouse zygotes. *Communications Biology*, *2*(1), 1–7. https://doi.org/10.1038/s42003-019-0627-8

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429 LP – 5433.

https://doi.org/10.1128/jb.169.12.5429-5433.1987

Jacobi, S. A., Jacobi, A. M., Collingwood, M. A., Rettig, G. R., Vakulskas, C. A., Bode, N. M., Dobosy, J. R., Cannieux, M., & Behlke, M. A. (2017). Efficient homology-directed repair using single-stranded DNA templates. http://sfvideo.blob.core.windows.net/sitefinity/docs/default-source/poster/efficient-homology-directed-repair-using-ssdna.pdf?sfvrsn=dbac3707_4

Jiang, F., & Doudna, J. A. (2017). CRISPR – Cas9 Structures and Mechanisms. *Annual Review of Biophysics.46*, 505–531.https://doi.org/10.1146/annurev-biophys-062215-010822.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., & Doudna, J. (2013). RNA-programmed genome editing in human cells. *ELife*, *2013*(2), 1–9. https://doi.org/10.7554/eLife.00471

Joung, J., Konermann, S., Gootenberg, J. S., Abudayyeh, O. O., Platt, R. J., Brigham, M. D., Sanjana, N. E., & Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols*, *12*(4), 828–863. https://doi.org/10.1038/nprot.2017.016

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Keith Joung, J. (2016). High-fidelity CRISPR-Cas9 variants with undetectable genome- wide off-targets. *Nature*, *528*(7587), 490–495. https://doi.org/10.1038/nature16526.High-fidelity

Koonin, E. V., & Krupovic, M. (2015). Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nature Reviews Genetics*, *16*(3), 184–192. https://doi.org/10.1038/nrg3859

Kotsarenko, K., Vechtova, P., Lieskovska, J., Füssy, Z., Cabral-de-Mello, D. C., Rego, R. O. M., Alberdi, P., Collins, M., Bell-Sakyi, L., Sterba, J., & Grubhoffer, L. (2020). Karyotype changes in long-term cultured tick cell lines. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-70330-5

Kotsyfakis, M., Schwarz, A., Erhart, J., & Ribeiro, J. M. C. (2015). Tissue- and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Scientific Reports*, *5*(1), 9103. https://doi.org/10.1038/srep09103

Liang F, Han M, Romanienko P.J., & Jasin M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proceedings of the National Academy of Sciences*, **95**(9), 5172-5177 (1998).

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., & Ma, W. (2019). Methodologies for improving HDR efficiency. *Frontiers in Genetics*, *10*(JAN), 1–9. https://doi.org/10.3389/fgene.2018.00691

Liu, Y., Ma, S., Chang, J., Zhang, T., Chen X., Liang Y., Xia Q., (2019). Programmable targeted epigenetic editing using CRISPR system in *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, *110*, 105–111. https://doi.org/10.1016/j.ibmb.2019.04.013

Mali P, Yang L, Esvelt K.M., Aach J, Guell M, DiCarlo J.E., Norville J.E., Church G.M. RNA-Guided Human Genome Engineering via Cas9. *Science***339**(6121), 823-826 (2013).

Mendoza-García, P., Hugosson, F., Fallah, M., Higgins, M. L., Iwasaki, Y., Pfeifer, K., Wolfstetter, G., Varshney, G., Popichenko, D., Gergen, J. P., Hens, K., Deplancke, B., & Palmer, R. H. (2017). The Zic family homologue Odd-paired regulates Alk expression in *Drosophila*. *PLoS Genetics*, *13*(4), 1–27. https://doi.org/10.1371/journal.pgen.1006617

Yan, M.-Y., Li, S.-S., Ding, X.-Y., Guo, X.-P., Jin, Q., & Sun, Y.-C. (2020). A CRISPR-Assisted Nonhomologous End-Joining Strategy for Efficient Genome Editing in *Mycobacterium tuberculosis*. *MBio*, *11*(1), e02364-19. https://doi.org/10.1128/mBio.02364-19

Mojica, F. J.M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), 733–740. https://doi.org/10.1099/mic.0.023960-0

Mojica, F J M, Juez, G., & Rodriguez-Valera, F. (1993). Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Molecular*

Microbiology, *9*(3), 613–621. https://doi.org/10.1111/j.1365-2958.1993.tb01721.x

Mojica, Francisco J.M., & Montoliu, L. (2016). On the origin of CRISPR-Cas technology: from prokaryotes to mammals. *Trends in Microbiology*, *24*(10), 811–820. https://doi.org/10.1016/j.tim.2016.06.005

Mojica, Francisco J M, Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60*(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Oishi, I., Yoshii, K., Miyahara, D., Kagami, H., & Tagami, T. (2016). Targeted mutagenesis in chicken using CRISPR/Cas9 system. *Scientific Reports*, *6*, 23980. https://doi.org/10.1038/srep23980

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Rasys, A. M., Park, S., Ball, R. E., Alcala, A. J., Lauderdale, J. D., & Menke, D. B. (2019). CRISPR-Cas9 gene editing in lizards through microinjection of unfertilized oocytes. *Cell Reports*, *28*(9), 2288-2292.e3. https://doi.org/10.1016/j.celrep.2019.07.089

Rath, D., Amlinger, L., Rath, A., & Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie*, *117*, 119–128. https://doi.org/10.1016/j.biochi.2015.03.025

Reegan, A. D., Ceasar, S. A., Paulraj, M. G., Ignacimuthu, S., & Al-Dhabi, N. A. (2017). Current status of genome editing in vector mosquitoes: A review. *Bioscience Trends*, *10*(6), 424–432. https://doi.org/10.5582/bst.2016.01180

Rizzoli A, Silaghi C, Obiegala A, et al. *Ixodes ricinus* and its transmitted pathogens in urban and peri-urban areas in Europe: new hazards and relevance for public health. *Front Public Health*. 2014;2:251. Published 2014 Dec 1. doi:10.3389/fpubh.2014.00251

Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., Graham, S., Robinson, C. V., Spagnolo, L., & White, M. F. (2013). Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Molecular Cell*, *52*(1), 124–134. https://doi.org/10.1016/j.molcel.2013.08.020

Salsman, J., & Dellaire, G. (2017). Precision genome editing in the CRISPR era. *Biochemistry and Cell Biology*, *95*(2), 187–201. https://doi.org/10.1139/bcb-2016-0137

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science*, *351*(6268), 84–88. https://doi.org/10.1126/science.aad5227

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014 Mar 6;507(7490):62-7. doi: 10.1038/nature13011. Epub 2014 Jan 29. PMID: 24476820; PMCID: PMC4106473.

Summer H, Grämer R, Dröge P. Denaturing urea polyacrylamide gel electrophoresis (Urea PAGE). *J Vis Exp*. 2009;(32):1485. Published 2009 Oct 29. doi:10.3791/1485

Sun, D., Guo, Z., Liu, Y., & Zhang, Y. (2017). Progress and prospects of CRISPR/Cas systems in insects and other arthropods. *Frontiers in Physiology*, *8*(SEP). https://doi.org/10.3389/fphys.2017.00608

Swarts, D. C., Mosterd, C., van Passel, M. W. J., & Brouns, S. J. J. (2012). CRISPR interference directs strand specific spacer acquisition. *PloS One*, *7*(4), e35888–e35888. https://doi.org/10.1371/journal.pone.0035888

Taning, C. N. T., Van Eynde, B., Yu, N., Ma, S., & Smagghe, G. (2017). CRISPR/Cas9 in insects: Applications, best practices and biosafety concerns. *Journal of Insect Physiology*, *98*, 245–257. https://doi.org/10.1016/j.jinsphys.2017.01.007

Wang, F., Zhang, C., Xu, W., Yuan, S., Song, J., Li, L., Zhao, J., & Yang, J. (2020). Developing high-efficiency base editors by combining optimized synergistic core components with new types of nuclear localization signal peptide. *Crop Journal*, *8*(3), 408–417. https://doi.org/10.1016/j.cj.2020.01.003

Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, Jaenisch R, Zhang F, Sharp PA. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol. 2014 Jul;32(7):670-6. doi: 10.1038/nbt.2889. Epub 2014 Apr 20. PMID: 24752079; PMCID: PMC4145672.

Xie, N., Zhou, Y., Sun, Q., & Tang, B. (2018). Novel epigenetic techniques provided by the CRISPR/Cas9 System. *Stem Cells International*, *2018*. https://doi.org/10.1155/2018/7834175

Xu, J., Chen, S., Zeng, B., James, A. A., Tan, A., & Huang, Y. (2017). *Bombyx mori* P-element somatic inhibitor (BmPSI) is a key auxiliary factor for silkworm male sex determination. *PLoS Genetics*, *13*(1), 1–17. https://doi.org/10.1371/journal.pgen.1006576

Yan, F., Wang, W., & Zhang, J. (2019). CRISPR-Cas12 and Cas13: the lesser known siblings of CRISPR-Cas9. *Cell Biology and Toxicology*, *35*(6), 489–492. https://doi.org/10.1007/s10565-019-09489-1

Yan, M., Li, S., Ding, X., Guo, X., Jin, Q., & Sun, Y. (2020). Molecular Biology and Physiology Efficient Genome Editing in Mycobacterium tuberculosis. *Molecular Biology and Physiology*. 1–14. DOI: 10.1128/mBio.02364-19.

Ye, H., Rong, Z., & Lin, Y. (2017). Live cell imaging of genomic loci using dCas9-SunTag system and a bright fluorescent protein. *Protein and Cell*, *8*(11), 853–855. https://doi.org/10.1007/s13238-017-0460-0

Yu, Z., Chen, H., Liu, J., Zhang, H., Yan, Y., Zhu, N., Guo, Y., Yang, B., Chang, Y., Dai, F., Liang, X., Chen, Y., Shen, Y., Deng, W. M., Chen, J., Zhang, B., Li, C., & Jiao, R. (2014). Various applications of TALEN- and CRISPR/Cas9-mediated homologous recombination to modify the *Drosophila* genome. *Biology Open*, *3*(4), 271–280. https://doi.org/10.1242/bio.20147682

Yu, Z., Ren, M., Wang, Z., Zhang, B., Rong, Y. S., Jiao, R., & Gao, G. (2013). Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics*, *195*(1), 289–291. https://doi.org/10.1534/genetics.113.153825

Zhang, Z., Aslam, A. F. M., Liu, X., Li, M., Huang, Y., & Tan, A. (2015). Functional analysis of *Bombyx* Wnt1 during embryogenesis using the CRISPR/Cas9 system. *Journal of Insect Physiology*, *79*(June), 73–79. https://doi.org/10.1016/j.jinsphys.2015.06.004

Zhu, G. H., Xu, J., Cui, Z., Dong, X. T., Ye, Z. F., Niu, D. J., Huang, Y. P., & Dong, S. L. (2016). Functional characterization of SlitPBP3 in *Spodoptera litura* by CRISPR/Cas9 mediated genome editing. *Insect Biochemistry and Molecular Biology*, *75*, 1–9. https://doi.org/10.1016/j.ibmb.2016.05.006

8. List of abbreviations

CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CAS	Crispr-Associated protein
TAE	Tris-acetate EDTA
PAM	Protospacer Adjacent Motif
dNTP	Deoxyribose Nucleoside TriPhospate
HNH	Histidine and asparagine rich Cas9 domain
RuvC	Endonuclease domain named for an *E. coli* protein involved in DNA repair
crRNA	CRISPR RNA
bp	Base pair
RNP	Ribonucleoprotein
NHEJ	Non homologous end joining
HDR	Homology directed repair

DSB	Double strand break
DAMT	DNA adenine methyltransferase
DNMT	DNA (cytosine-5) methyltransferase
Sia 2,6	Beta-galactoside alpha-2,6-sialyltransferase
Sia 2,3	Beta-galactoside alpha-2,3-sialyltransferase