UNIVERSITY OF GRONINGEN

CRISPR, a prokaryotic adaptive defense mechanism against invading phages

Author: Tom van lente (1610155), Supervisor: Oscar Kuipers

5-7-2013

Content

Abstract

A recent discovered defense system protects prokaryotes from invading phages. This system makes use of RNA guided protein complexes to cleave invading DNA. This so called CRISPR has the ability to constantly adjust itself on a genomic level to obtain immunity for phages that were encountered earlier. The system is composed of repeated DNA sequences that are separated by variable spacer sequences obtained from phage DNA and a set of CRISPR associated (Cas) genes. CRISPR defense is obtained in three steps: (i) adaptation of spacers from phage DNA, (ii) expression and maturation of the CRISPR RNA (crRNA) and (iii) interference of invading phage DNA by crRNA and Cas proteins. Recognition of invading phages is done through a short sequence (PAM) that is located adjacent to the targeted spacer on phage DNA. There are other possible applications regarding the CRISPR interference mechanism. The Cas9 protein can be used in combination with designed crRNA to cleave a specific location of chromosomal DNA in a wide variety of organisms. This is a useful tool for genome editing and induction of mutations.

Introduction

Prokaryotic organisms have developed many strategies to defend their genetic heritage. Several of these strategies are responses to invading prokaryotic viruses (phages). These antiviral mechanisms include adsorption inhibition and the restriction modification system (RMS). Adsorption inhibition is a process where bacteria hide or modify their receptors in order to escape viral particles from binding onto them (Labrie *et al*., 2010). The other mechanism, restriction modification, uses a process in which the bacterial DNA is methylated. Foreign DNA which is not methylated is cleared by RMS restriction nuclease (Wilson, 1991). A new recently discovered defense system, CRISPR-Cas, uses RNA as a tool to recognize invading DNA. Evidence related to this system was seen 25 years ago for the first time (Ishino *et al*., 1987).

Discovery of the CRISPR-Cas system

In 1987, a short repetitive sequence was discovered downstream of the iap gene in *E. coli* that did not have an explicit function (Ishino *et al*., 1987). This sequence consists of 29 nucleotide repeated parts interspaced by 32 nucleotide spacer parts that seemed quite random. These repeated sequences were later found in other organisms with slightly different lengths (Groenen *et al*., 1993). These spacers were used for a new method termed 'spoligotyping', genotyping of spacers (Goyal *et al*., 1997). This was first used in hospitals where the variety of *M. tuberculosis* spacers were analyzed to find out where the roots were located. This can be quite helpful in determining the origin of a pathogen outbreak. Spacer oligonucleotide typing can be used for phylogenetic and evolutionary studies as well.

In 2002 a new name was suggested after several names were used for the same sequence. The term: clustered regularly interspaced short palindromic repeats (CRISPR) was introduced and is now the rightful name. The same year 4 genes were discovered and linked to the CRISPR sequence. These Cas genes were only seen in CRISPR positive organisms in the direct vicinity of CRISPR arrays (Jansen *et al*., 2002). In the upcoming years this number was expanded by the addition of 42 new Cas genes from several organisms (Haft *et al*., 2005). The same year a very important discovery was made. Several research groups found that the origin of the spacers were extra chromosomal. The spacer sequences were homologous with certain parts of phage and plasmid DNA (Bolotin *et al*., 2005, Mojica *et al*., 2005, Pourcel *et al*., 2005). This suggests that the spacers are obtained from incoming viruses.

After this discovery it did not take long before the CRISPR was linked to immunity. Bacteria isolated after phage invasion and bacteria where the spacer content was changed to match the specific phage, showed a decreased sensitivity towards infection. The CRISPR sequences were analyzed after infection and revealed that new spacers resembled a part of the phage DNA (Barrangou *et al*., 2007). More research on this topic confirmed that this is a prokaryotic antiviral immune system based on the implementation of viral DNA.

Research on CRISPRs now mainly focuses on the mechanism of acquiring and interfering the invasion of phage DNA. Not all the CRISPR-Cas systems work in a similar way and thus different functioning CRISPRs were classified. In this paper, a closer look on the composition and the mechanism of different CRISPR types will be explained. CRISPRs can also be used for a wide variety of applications. Some of these applications that were discovered during the last couple of years will be reviewed as well.

Composition of the CRISPR locus

The CRISPR locus consists of short repeated sequences that are separated with unique similar in length spacers (figure 1). These repeats are conserved except for small differences in CRISPR loci of other organisms. Some of the repeats have a palindromic nature allowing them to form secondary hairpin structures while others are unstructured (Kunin *et al*., 2007). A conserved sequence, GAAA(G/C), is present in most repeats at the 3' end for binding of proteins (Kunin *et al*., 2007, Godde *et al*., 2006, Mojica *et al*., 2000). The spacers, obtained from viral DNA, are inserted at one end of the locus between a repeated sequence and a leader sequence (Barrangou *et al*., 2007, Andersson *et al*., 2008, Horvath *et al*., 2010, Pourcel *et al*., 2005, Lillestøl *et al*., 2009). The leader sequence has promoter elements and regulatory binding sites (Pougach *et al*., 2010, Pul *et al*., 2010). Finally, the locus has a set of Cas genes that have a variety of functions needed during the process of viral immunity. The overall architecture of CRISPR loci in different organisms is the same. There is however a wide diversity in CRISPR length, repeat sequences and Cas genes (Kunin *et al*., 2007).

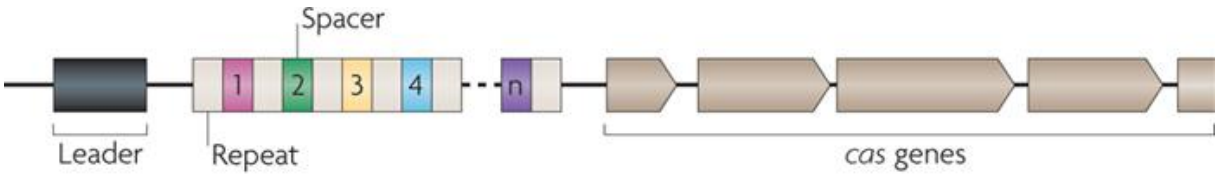

Figure 1) **Typical CRISPR locus.** *The construct has a leader sequence preceding the CRISPRs that has promoter and regulatory elements. The CRISPR part contains up to a few hundred different spacers that are separated by conserved repeat sequences. A set of Cas genes precedes or follows the repeats* (Marraffini *et al*., 2010 (2)).

Considering the differences in CRISPR loci, three major types can be assigned that follow a different mechanism. These types consist of several subtypes that have different sets of Cas genes within them resulting in a total of ten different CRISPR subtypes. The type 1 CRISPR-Cas system has six different subtypes, 1A to 1F. This type was found in both archaea and bacteria. Type 2 and 3 CRISPR-Cas systems only have two subtypes each, A and B. Type 2 is only found in bacteria and type 3 can only be found in archaea. Each type of CRISPR-cas system has a distinguished signature Cas gene. Type 1, 2 and 3 systems have the signature gene Cas3, Cas9 and Cas10 respectively (Makarova *et al*., 2011). These Cas genes are involved in the degradation of invading DNA. Cas3 is a helicase that promotes strand separation and cuts DNA during interference (Brouns *et al*., 2008). Cas9 forms a complex with the CRISPR RNA and acts as a nuclease (Jinek *et al*., 2012). Cas10 cuts single strands DNA and might play a role in strand separation (Cocozaki *et al*., 2012). From all the Cas genes, Cas1 and Cas2 are the only universally conserved proteins (Haft *et al*., 2005, Makarova *et al*., 2006). Cas1 is involved during the integration of new spacer DNA (Yosef *et al*., 2012), Cas2 facilitates spacer selection and might be involved in the processing of CRISPR RNA during the invasion of phage DNA (Brouns *et al*., 2008). The rest of the Cas proteins are present in specific subtypes or are specie-specific (Makarova *et al*., 2011). The function of the major Cas proteins is known (table 1) but the exact mechanism still remains unclear for some of them. Based on this knowledge some suggestions have been made on the mechanism of CRISPR defense.

Table 1) **Functionalities of most occurring Cas proteins (Cas1-10).** *Cas1 and 2 (blue) are universal proteins, Cas3, 9 and 10 (red) are signature proteins and Cas 4-8 are type specific proteins (*Table from Bhaya *et al.,* 2011*).*

Protein	Distribution	Process	Function
Cas1	Universal	Spacer acquisition	DNAse, not sequence specfic, can bind RNA; present in all Types; structure available for several Cas 1 proteins
Cas2	Universal	Spacer acquisition	Small RNAse specific to U-rich regions; present in all Types; structure available from *Thermus thermophilus* and *Sulfolobus solfataricus* and others
Cas3	Type I signature	Target interference	DNA helicase; most proteins have a fusion to HD nuclease
Cas4	Type I, II	Spacer acquisition	RecB-like nuclease with exonuclease activity homologous to RecB
Cas5	Type I	crRNA expression	RAMP protein, endoribonuclease involved in crRNA biogenesis; part of CASCADE
Cas6	Type I, III	crRNA expression	RAMP protein, endoribonuclease involved in crRNA biogenesis; part of CASCADE; structure available from *P. furiosus*
Cas7	Type I	crRNA expression	RAMP protein, endoribonuclease involved in crRNA biogenesis; part of CASCADE
Cas8	Type I	crRNA expression	Large protein with McrA/HNH-nuclease domain and RuvC-like nuclease; part of CASCADE
Cas9	Type II signature	Target interference	Large multidomain protein with McrA-HNH nuclease domain and RuvC-like nuclease domain; necessary for interference and target cleavage
Cas10	Type III signature	crRNA expression and interference	HD nuclease domain, palm domain, Zn ribbon; some homologies with CASCADE elements

The CRISRP mechanism

Step 1: Adaptation of CRISPRs

The mechanism of CRISPR interference is a very ingenious and somewhat complex one. This process can be divided into three stages (van der Oost *et al.*, 2009, Wiedenheft *et al.*, 2012). The first step, adaptation, is also the least understood compared to the other steps (figure 2). During this step, new spacers are obtained from invading DNA (protospacers) and placed in the CRISPR locus. There are indications which types of Cas proteins are involved during the acquisition of new spacers (Yosef *et al.*, 2012) and it is clear that new spacers are inserted at the leader end of the CRISPR (Barrangou *et* al., 2007, Horvath *et al.*, 2010, Pourcel *et al.*, 2005). How this step varies between subtypes in terms of mechanism is not clear.

The universal proteins Cas1 and Cas2 are both endonucleases (Wiedenheft *et al.*, 2009, Beloglazova *et al.*, 2008) that affect new spacer acquisition in the CRISPR locus (Makarova *et al.*, 2006, Brouns *et al.*, 2008). Experiments have shown that overexpression of both Cas1 and Cas2 genes result in more spacer acquisition (Yosef *et al.*, 2012), while no spacers are obtained when only one of these Cas genes is expressed (Datsenko *et al.*, 2012). In vitro experiments with Cas1 showed that this protein is able to cleave double stranded DNA (Wiedenheft *et al.*, 2009) to fragments of approximately 80 basepairs long (Garneau *et al.*, 2010). Since only a specific protospacer

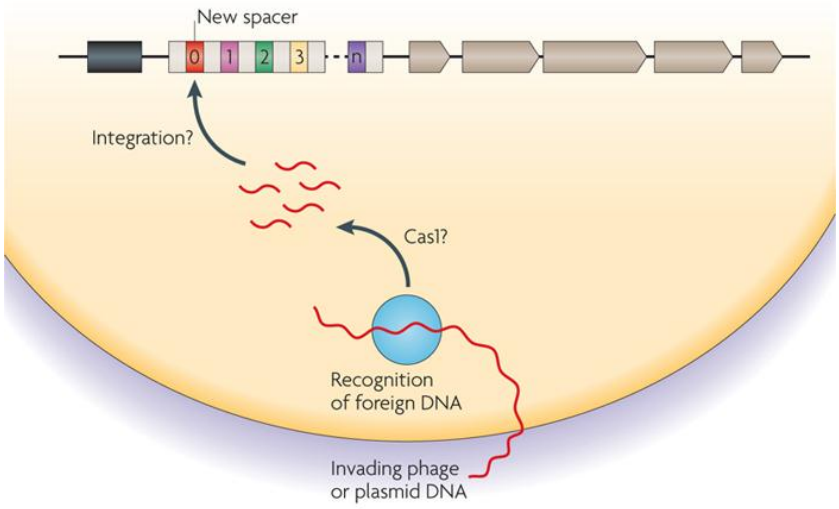

Figure 2) **Acquisition of new spacers in the CRISPR locus.** *Spacers are obtained from invading phage or plasmid DNA. The process of spacer acquisition is not known. It is thought that the nuclease abilities of Cas1 are used to implement new spacers* (Marraffini *et al.*, 2010 (2)).

is targeted during the adaptation stage, it was suggested that other subtype specific Cas proteins and Cas2 act with Cas1 to obtain new spacers selectively (Babu *et al*., 2011).

Selectivity of spacer acquisition

The CRISPR does not implement its own DNA in the locus, so what is driving this selective acquiring of spacer DNA? By screening the protospacer sequences from phage DNA, a short conserved part was seen adjacent to the protospacer (Bolotin *et al*., 2005). This sequence, also known as the protospacer-adjacent motif (PAM), was quickly suggested to be important for viral DNA detection and acquisition (Mojica *et al*., 2005). PAMs coincide with repeat types, thus different CRISPR classes recognize different PAMs (Kunin *et al*., 2007). This is not the case for type 3 CRISPRs that recognize protospacers without a PAM sequence (Marraffini *et al*., 2010 (1)). The PAM motif can be found at the 3' or the 5' end of the protospacer (van der Ploeg, 2009). This sequence is necessary for acquisition of new spacers as well as the silencing of invading DNA. Mutations within this small sequence in invading DNA result in avoidance of CRISPR interference (van der Oost *et al*., 2009).

Step 2: Expression of CRISPR RNA

The next stage is the expression of the CRISPR locus and maturation of CRISPR RNA (crRNA). The whole CRISPR locus is transcribed as a long piece of RNA (pre-crRNA) from a promoter sequence in the leader (figure 3). Pre-crRNA is processed further into smaller pieces of CRISPR RNA (crRNA) that are used for protection against invading DNA (Gesner *et al*., 2011, Sashital *et al*., 2011). The mechanism varies slightly depending on the CRISPR type. Between the subtypes the mechanism is similar, except that it is obtained with a different set of Cas proteins that may act in a slightly other way.

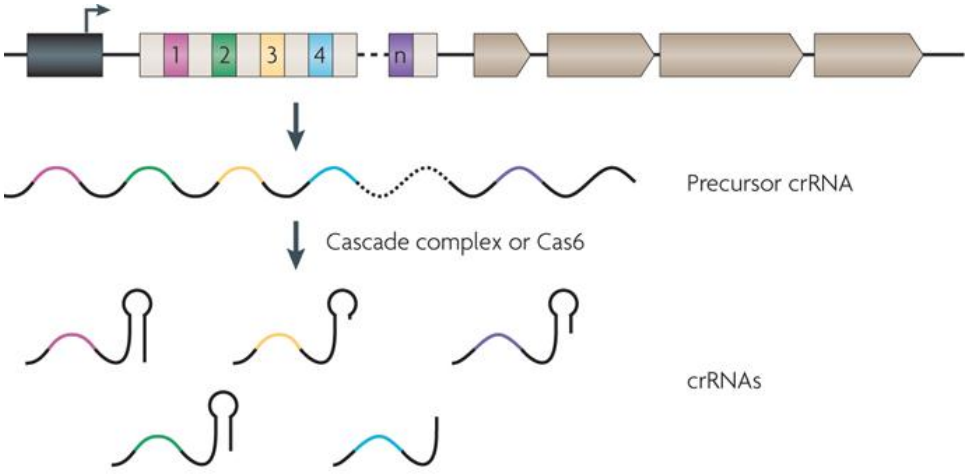

Figure 3) **Transcription and maturation of type 1 crRNA.** *The CRISPR is transcribed as a long piece of pre-crRNA. The pre-crRNA is then processed further by Cascade or specific Cas proteins to crRNA. The pre-crRNA is cleaved at the 3' end of the spacer resulting in crRNA that has part of the repeat sequence at both sides* (Marraffini *et al.*, 2010 (2)).

The subtype 1 specific repeats differ in sequence and bind different endoribonucleases to cleave the pre-crRNA into smaller single repeat-spacer units. These subtypes are conserved but may have a slightly different mechanism (Sorek *et al*., 2013). Some repeat sequences of type 1 CRISPRs are palindromic while others are unstructured. The pre-crRNA is cleaved in such a way that each CRISPR unit has a small part of the repeat sequence (8-10 nt) at the 5' end which is known as the 5' end handle. The repeat sequence with the hairpin loop or the unstructured RNA is located at the 3' end (3' end handle). These handles may serve as a conserved binding region for subunits of the CRISPR-associated complex for antiviral defense (Cascade) complex (Figure 4a). This protein complex is required for silencing viral DNA in the next stage (figure 4b) (Brouns *et al*.,2008). When the crRNA is obtained, Cascade binds onto it to protect the crRNA from degrading. The Cascade complex contains several sub-units at various locations on the crRNA. The sub-units on the spacer sequence (6x Cas7) are the same for all the type 1 CRISPRs. The sub-units located on the 5' and 3' handle differ in each sub-type (Sorek *et al*., 2013).

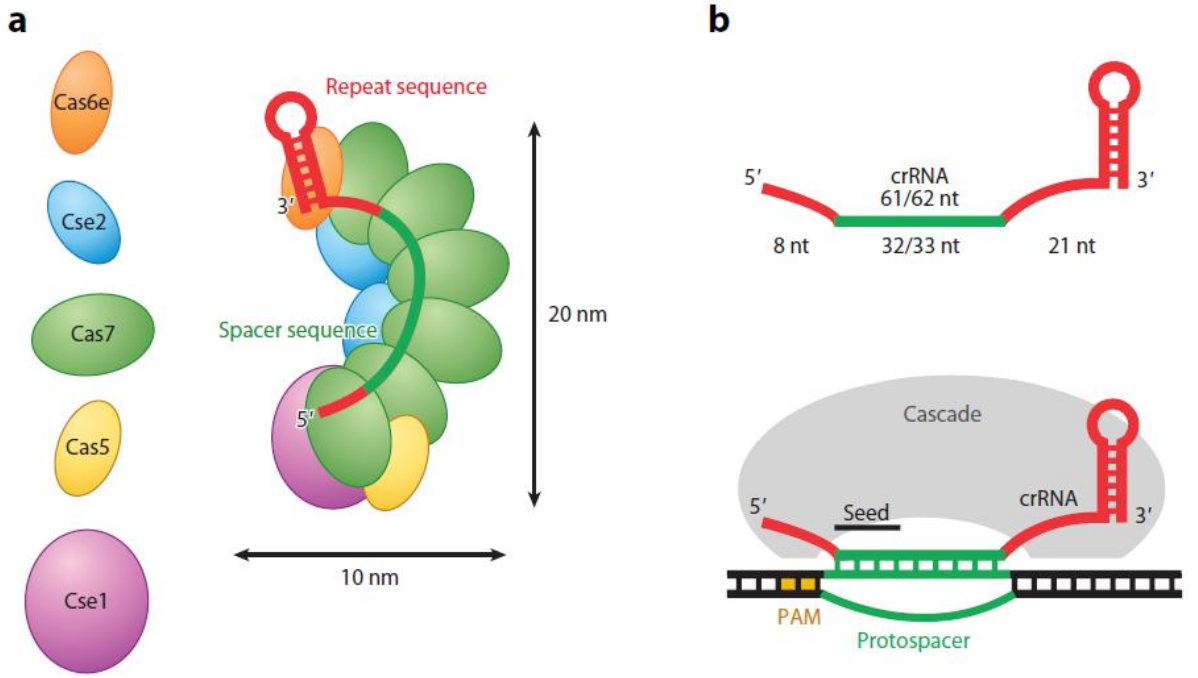

Figure 4) **Schematic representation of Cascade and Cascade-crRNA interaction with target DNA. (a)** *Cascade consists of 5 different subunits that form a structure around the crRNA. Cas7 is attached to the spacer and forms the backbone of the complex.* **(b)** *Cascade-crRNA bound to the targeted DNA. A double strand is formed between the crRNA and the targeted DNA resulting in a single stranded protospacer* (Westra *et al*., 2013 (2)).

Maturation of the type 2 CRISPRs is quite different from the other types (figure 5a). The process requires an extra strand of RNA that is transcribed from the CRISPR locus as well. This RNA sequence is a *trans*-acting CRISPR-associated RNA, also known as tracr-RNA, needed for the binding of RNase III, a cellular ribonuclease. Part of the tracr-RNA is complementary to the repeat RNA sequence and forms a short double stranded RNA (Deltcheva *et al*., 2011, Jinek *et al*., 2012). The 3' end is trimmed by RNase III resulting in a small piece of spacer RNA at the 5' end of the adjacent crRNA. This 5' end repeat sequence is removed by further trimming of the crRNA which requires Cas9. The final mature crRNA has the spacer at the 5' end and the repeat sequence with the tracrRNA at the 3' end. The complex with Cas9 is maintained to protect the crRNA from degradation (Jinek *et al*., 2012).

The maturation of the type 3 CRISPRs is obtained through cutting pre-crRNA with Cas6 (figure 5b) (Gesner *et al*., 2011, Sashital *et al*., 2011). The repeat sequences of these CRISPR types do not have a specific secondary structure (Wang *et al*., 2011). The 3' end is then trimmed further at different locations in the spacer sequence, resulting in different lengths of mature crRNA and the removal of the 3' end repeat sequence. The difference in length is constant for each organism (Hale *et al*., 2012, Carte *et al*., 2010). The trimmed crRNA is then imbedded in a protein complex also known as the Cas repeat-associated mysterious proteins (RAMP) module complex

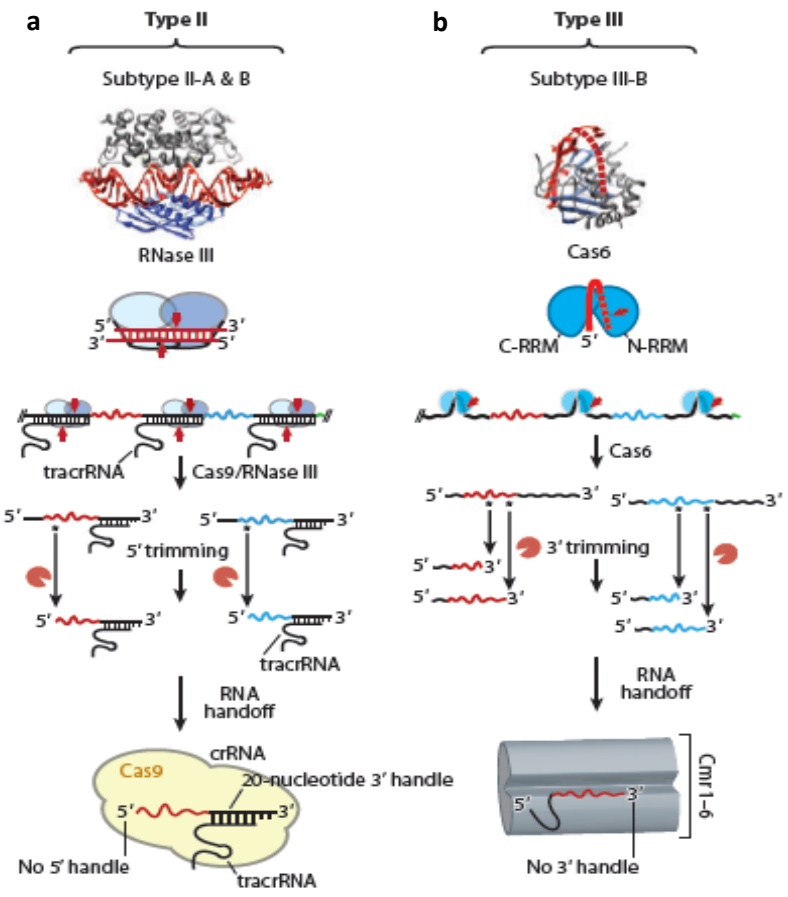

Figure 5) **Expression and maturation of type 2 and 3 CRISPRs. a)** *Type 2 pre-crRNA is combined with tracrRNA to form a partly double stranded RNA complex. Pre-crRNA is cleaved by Cas9 and RNase III to obtain mature crRNA. The crRNA is further trimmed by RNAse III. The crRNA combined with Cas9 is now ready to interfere invading DNA.* ***b)*** *Type 3 pre-crRNA is cleaved by Cas6 to obtain smaller pieces of RNA. The 3' end is trimmed further in the spacer sequence. The crRNA is combined with the RAMP module complex and is now able to interfere invading DNA* (Sore *et al.*, 2013).

(Hale *et al.*, 2009). It is still unknown why the difference in mature CRISPR length is always the same. Apparently there is a conserved sequence in the protospacer that is recognized and cleaved by Cas6.

Step 3: Interference of invading DNA

Invading DNA is silenced through interactions with crRNA combined with certain Cas proteins. The mechanism behind CRISPR interference for all the different subtypes is not yet understood. The Cascade complex in type 1 CRISPRs binds preferably to long strands of dsDNA, such as phage or plasmid DNA. Once the Cascade is bound, it will move along the strand until a specific subunit at the 5' end interacts with the PAM (figure 6). The double stranded DNA is altered by the Cascade complex resulting in a partly untwined strand also known as a g-loop. The spacer part of the crRNA then binds with the complementary protospacer (Westra *et al.*, 2012 (1)). Once the crRNA is bound to the target, the single strand is attacked by Cas3, a protein with nuclease and helicase abilities. The DNA is cleaved and the Cascade complex detaches. The still bound Cas3 nuclease will cleave the other strand (Makarova *et al.*, 2006).

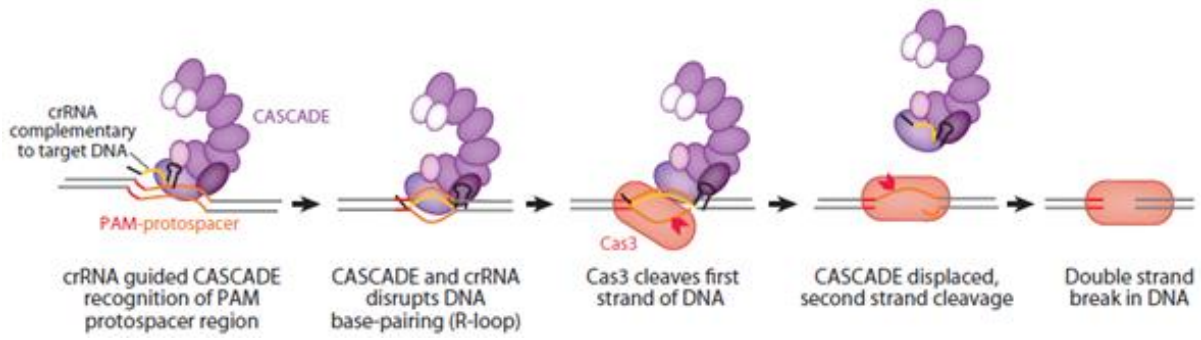

Figure 6) **Interference of the Cascade complex.** *Cascade follows the targeted DNA sequence until it encounters a PAM. This interactions promotes strand separation, allowing crRNA to bind with its complementary strand. This is followed by binding of Cas3 to the single target DNA strand. The strand is cleaved and Cascade detaches from the DNA. Then, the other strand is cleaved by Cas3, resulting in destruction of the invading DNA* (Bhaya *et al.*, 2011).

Interference of type 2 CRISPRs is obtained with a complex of Cas9, crRNA and tracrRNA. Cas9 is thought to be involved during the production of crRNA and the silencing of invading DNA (Jinek *et al*., 2012). It has different predicted nuclease domains, necessary for DNA cleavage (Makarova *et al*., 2011). One Cas9 domain cleaves the target strand while another domain can cleave the non-complementary strand (Jinek *et al*., 2012). Both strands are cleaved at the specific protospacer and result in a blunt-end double DNA strand (Garneau *et al*., 2010). The tracrRNA is enhancing the ability of Cas9 to cleave targeted DNA. This is likely caused by an increased binding of crRNA to the targeted DNA when tracrRNA binds crRNA. The part of tracrRNA that does not bind to the crRNA might have interactions with Cas9 or the targeted DNA. This may favor Cas9 binding as well (Jinek *et al*., 2012).

The Cas10 protein is thought to be involved during crRNA expression and interference in type 3 CRISPRs (Makarova *et al*., 2011). There are more type 3 subtype specific proteins that play a role during CRISPR interference. For type 3A CRISPRs these are Csr proteins and for type 3B CRISPRs these are the Cmr proteins (Wang *et al*., 2011). These proteins form complexes around the crRNA that resemble the structure of the Cascade complex in type 1 CRISPRs (Hale *et al*., 2009). Type 3A Csr proteins target DNA in a similar way as type 1 CRISPRs while type 3B Cmr proteins are thought to target RNA. RNA cleavage was observed *in vitro* but the mechanism is still unknown (Marraffini *et al.*, 2010 (2)). Recent results from Deng *et al*. suggest that some Cmr modules are capable of targeting DNA as well (Deng *et al*., 2013). As mentioned earlier, type 3 CRISPRs target invading DNA/RNA that do not have a PAM sequence. How DNA is targeted is not known. The pre-crRNA is trimmed at the spacer part and results in the same length for different spacers, suggesting a selection based on a sequence within the protospacer.

CRISPR design prevents autoimmunity

It is possible for an organism to target its own DNA when a part is implemented in the CRISPR loci (Stern *et al*., 2010). This is always the case for the spacer within the CRISPR loci. The crRNA will target the complementary spacer but does not cleave it for two reasons. The first prevention is that the CRISPR loci does not have PAMs. It was shown that the PAM is necessary for preventing self-targeting and degradation in type 1 and 2 CRISPRs (Jinek *et al*., 2012). The second prevention is due to the CRISPR design. If the crRNA binds with its own DNA on the CRISPR locus it overlaps perfectly with the spacer and the repeat part, while binding with phage DNA does not result in a complete overlap due to the 5' end repeat sequence (figure 7). This difference in binding is a way for the CRISPR mechanism to prevent self-targeting. The same mechanism prevents type 3A CRISPRs from autoimmunity (Marraffini *et al*., 2010 (1)). Since the other type 3 CRISPRs target RNA instead of DNA, no such mechanism is needed to prevent self targeting.

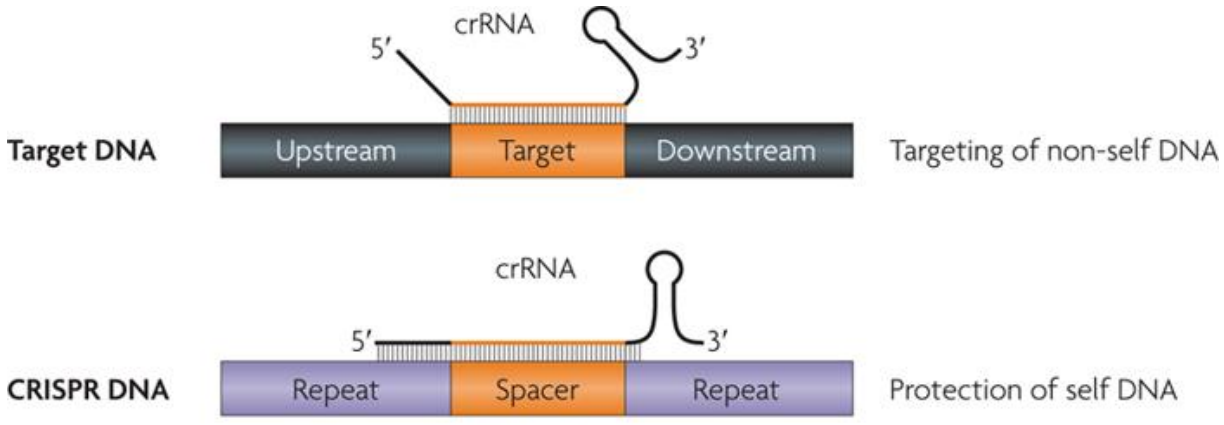

Figure 7) **crRNA targeting self and non-self DNA.** *CRISPRs are able to distinguish their own DNA from invading DNA. When it binds with invading DNA, only the spacer part overlaps. When it binds with its own CRISPR locus, it results in complete overlapping of the crRNA and the repeat-spacer. The latter prevents cleavage of the targeted DNA* (Marraffini *et al.*, 2010 (2)).

CRISPR applications

The understanding of the mechanism behind CRISPR-Cas interference is rapidly growing. Six years have passed since the link was made between CRISPRs and immunity (Barrangou *et al*., 2007). During this relatively small area of time, a wide array of other CRISPR applications were speculated and even tested that could improve or replace already existing techniques. A quite simple but effective application is the implementation of CRISPR-Cas loci in bacteria used in the dairy industry (Barrangou *et al.,* 2007). Phage infections can alter normal fermentation cycles, leading to the disruption of production and a decrease in the quality of the desired product (Mc Grath *et al*., 2007). Bacteria with new spacers can be isolated and used to replace the original strain. With the same properties and no genetic modifications, this is an easy way to prevent production losses due to phage infections (Marraffini *et al*., 2010 (2)).

Perhaps a more interesting feature, discovered recently, is the usage of the type 2 CRISPR interference mechanism for genome editing. The type 2 complex only uses Cas9 for cleaving the invading DNA (Jinek *et al*., 2012). This method can be applied to target specific DNA by modifying the crRNA to make it complementary. This new crRNA, also termed guide RNA (gRNA), can guide Cas9 to the desired location in the genome in order to cleave DNA and trigger deletions or insertions (Hwang *et al*., 2013). This technique was already applied in several organisms: *Streptococcus pneumoniae* and *Escherichia coli* (Jiang *et al*., 2013), *Sacharomyces cerevisiae* (Dicarlo *et al*., 2013), zebrafish embryos (Chang *et al*., 2013, Hwang *et al*., 2013) and even human cells (Cho *et* al., 2013, Mali *et al*., 2013).

One problem regarding this application is the need of a PAM adjacent to the targeted DNA, thus limiting the locations available. One way to solve this problem is by introducing other Cas9 genes that cleave DNA with different PAM sequences. This results in a wider range of specific locations. This might not be necessary because the targeting of PAM sequences with some Cas genes is not so strict. Hwang *et al*, performed divers mutations in the specific PAM sequence to determine if cleavage still takes place. With deep sequencing they found that only three base pairs are necessary for that specific Cas gene to work. Sequences that have the form 5′-GG-N18-NGG-3′ can be targeted. Even a shift of one base pair results in cleavage of the desired DNA. This sequence can be found once every 128 basepairs of a random DNA sequence. The GG part at the 5' end of the targeted sequence is needed because the T7 promoter requires this for gRNA transcription. The promoter site is dependent on the organisms' polymerase, so different sequences can be targeted when other polymerases are used. Finding this sequence in the genome is very high although it depends slightly on the G/C content. With this in mind, almost every location on the genome can be targeted with a designed gRNA (Hwang *et al*., 2013).

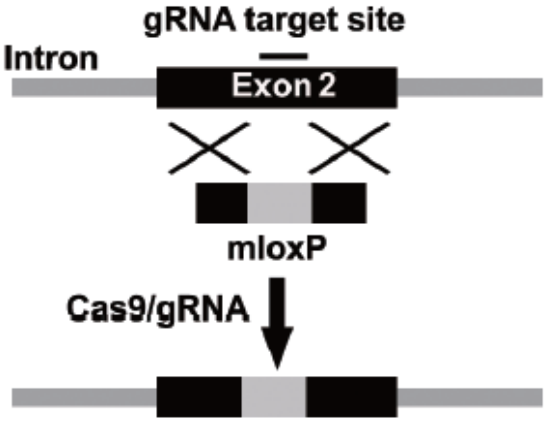

Figure 8) **Insertion of mloxP with Cas9 and gRNA.** *Cas9 is combined with gRNA to target Exon 2. Cas9 will then cleave Exon 2 which is repaired by mloxP donor DNA that has overlapping parts with Exon 2* (Chang *et al*., 2013).

Other than cleaving specific DNA, it was also accomplished to insert additional DNA sequences. Chang *et al*, successfully inserted a mutant *loxP* in zebrafish embryos, which in turn can either be used for knocking out genes or replace genes from another vector with the addition of Cre recombinase (figure 8). The insertion was obtained by cleaving the DNA, using Cas9 with gRNA, and repairing it with a donor strand containing loxP through homologous recombination. If de overlapping parts contain a sequence that is targeted by gRNA, they will be cleaved. This is something to be alert of when designing a gRNA (Chang *et al*., 2013).

There already exist different strategies to cleave double stranded DNA such as zinc finger nucleases (ZFNs) (Urnov *et al*., 2010) and transcription activator-like effector nucleases (TALENs) (Bogdanove *et al*., 2011). Both consist of DNA binding proteins that, combined with nucleases, cleave DNA at a desired location. In order to use these systems, some protein engineering is

required to allow binding at the specific DNA location. Compared to gRNA, where the specific sequence must be synthesized, protein engineering is more time consuming (Jiang *et al*., 2013).

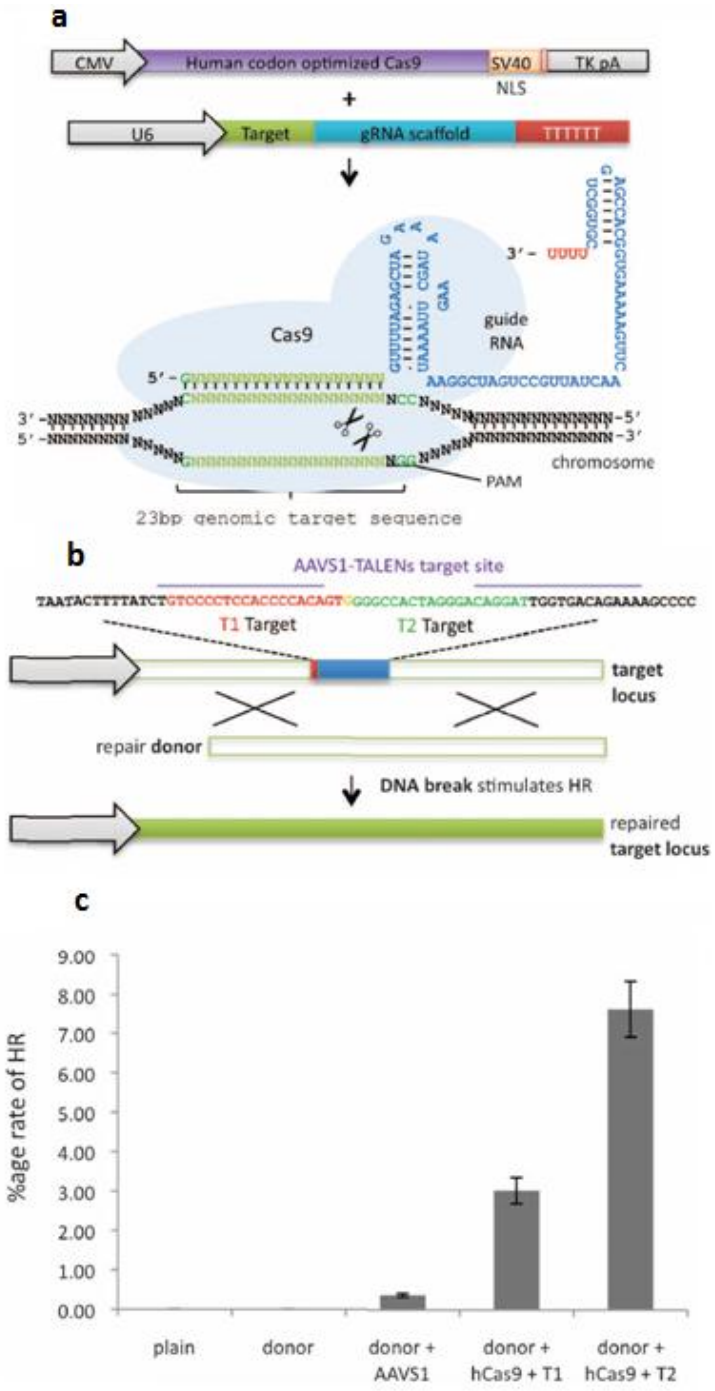

Figure 9) **Comparison of Cas9 and TALEN through repair of mutant GFP. a)** *Human codon optimized Cas9 with SV40 nuclear localization signal and designed gRNA are transcribed. Cas9 unwinds the target strand and cleaves both target strands when the gRNA recognizes the specific target.* **b)** *The targeted gene is transcribed with the gRNA. This gene transcribes a mutant GFP that has a stop codon in the middle. Once cleaved, the gene will be repaired with wt GFP. Two different gRNAs T1 and T2 were tested on this gene.* **c)** *The amount of fluorescence was measured for the each gRNA and TALEN (AAVS1). These results indicate that Cas9 with both gRNAs, T1 and T2, are more efficient than the TALEN* (Mali *et al., 2013*).

Cas9 with gRNA is easier to use but is it efficient enough to compete with the previous mentioned methods? A comparative result was performed by Mali *et al*, to see how the efficiency varies between two different gRNA, that target the same site but at a different location, and a TALEN (figure 9). When the targeted DNA is cleaved an additional repair donor overlaps the gap, altering the DNA to code GFP. These results show that both the gRNA-Cas9 complexes were more efficient in cleaving the target DNA compared to the TALEN (Mali *et al*., 2013).

gRNA-Cas9 complexes have proven to work in a wide variety of organisms capable of receiving plasmid DNA (Jiang *et al*., 2013). However, in most cases there was a certain percentage of cells that were lost due to toxic effects. CRISPR mediated cleavage of chromosomal DNA results in cell death in many bacteria and archaea (Marraffini *et al*., 2010 (1)). Adding a template to recombine with the targeted DNA results in the prevention of cell death through cleavage.

Conclusions

The understanding of CRISPR mechanisms is growing but there are still some gaps left. It is clear that CRISPR defense is obtained in three steps: adaptation, expression followed by maturation, and interference. Furthermore, invading DNA is only cleaved when a specific PAM is present, except for type 3 CRISPRs. The mechanism behind CRISPR adaptation and interference is partly understood. Several classes have evolved that follow a different mechanism. There are some proposed applications with the use of CRISPRs that were speculated or tested. Organisms used in industrial processes can be selected on their spacer content to prevent certain phage infections. Another application is the use of Cas9 for genome engineering. This was done successfully in several

organisms. With designed crRNAs it is possible to induce mutations and even insert new parts of DNA by adding a donor with overlapping regions. This application is more efficient and easier to use compared to other engineering tools such as TALENs and ZFs.

Future prospects

CRISPRs seem promising with many applications. The hard to understand mechanism of adaptation might give new insights in finding new ways of genome engineering. This can, for instance, result in an easier way to transfer DNA from a plasmid into a host. Another application that might provide new methods for biological engineering is PAM recognition. This small selective region is recognized by certain Cas proteins. Once known how this recognition works, it might be used in other applications to guide biological machineries to specific parts on the DNA.

There is some speculation on the usefulness of CRISPRs in reducing the amount of antibiotic resistant organisms. Plasmids that contain an antibiotic resistance gene can transfer from one organism to another, resulting in a growing population of resistant bacteria. If a spacer matches the plasmid, transfer will likely not happen. A recent study on antibiotic resistance with regard to CRISPRs showed that there is no correlation between the presence of plasmids, integrons, acquired antibiotic resistance and CRISPRs (Touchon *et al*., 2012). It is not clear how some plasmids can still invade organisms when a matching spacer is present. A better understanding of this process might lead to an improved method of CRISPR interference and perhaps a new way to slow down antibiotic resistant plasmids from spreading.

References

Andersson, A.F., Banfield, J.F. (2008). "Virus population dynamics and acquired virus resistance in natural microbial communities." Science **320**: 1047-1050.

Babu, M., Beloglazova, N., Flick, .R, Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A. et al. (2011). "A dual function of the CRISPR–Cas system in bacterial antivirus immunity and DNA repair." Mol Microbiol. **79**: 484-502.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). "CRISPR provides acquired resistance against viruses in prokaryotes." Science **315**: 1709-1712.

Beloglazova, N., Brown, G., Zimmerman, M.D., Proudfoot, M., Makarova, K.S., et al. (2008). "A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats." J. Biol. Chem. **283**: 20361-20371.

Bhaya, D., Davison, M., Barrangou, R. (2011). "CRISP-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Reguation." Annu. Rev. Genet. **45**: 273-297.

Bogdanove, A.J., Voytas, D.F. (2011). "TAL effectors: customizable proteins for DNA targeting." Science **30;333(6051)**: 1843-1846.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin." Microbiology **151:** 2551-2561.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., et al. (2008). "Small CRISPR RNAs guide antiviral defense in prokaryotes." Science. **321**: 960-964.

Carte, J., Pfister, N.T., Compton, M.M., Terns, R.M., Terns, M.P. (2010). "Binding and cleavage of CRISPR RNA by Cas6." RNA. **16**: 2181-2188.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.W., Xi, J.J. (2013). "Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos." Cell Res. **23**: 465-472.

Cho, S.W., Kim, S., Kim, J.M., Kim, J.S. (2013). "Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease." Nat. Biotechnol. **31**: 230-232.

Cocozaki, A.I., Ramia, N.F., Shao, Y., Hale, C.R., Terns, R.M., et al. (2012). "Structure of the Cmr2 subunit of the CRISPR-Cas RNA silencing complex." Structure **20**: 545-553.

Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., Semenova, E. (2012). "Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system." Nat. Commun. **3**: 945.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., et al. (2011). "CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III." Nature. **471**: 602-607.

Deng, L., Garrett, R.A., Shah, S.A., Peng, X., She, Q. (2013). "A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus". Mol. Microbiol. Mar; **87**: 1088-1099.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M. (2013). "Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems." Nucleic Acids Res. **41**: 4336-4343.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyava,l P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S. (2010). "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." Nature. **468**: 67-71.

Gesner, E.M., Schellenberg, M.J., Garside, E..L, George, M.M., MacMillan, A.M. (201). "Recognition and maturation of effector RNAs in a CRISPR interference pathway." Nat. Struct. Mol. Biol. **18**: 688-692.

Godde, J.S., Bickerton, A. (2006). "The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes." J. Mol. Evol. **62**: 718-729.

Goyal, M., Saunders, N.A., van Embden, J.D., Young, D.B., and Shaw, R.J. (1997). "Differentiation of *Mycobacterium tuberculosis* isolates by spoligotyping and IS6110 restriction fragment length polymorphism." J. Clin. Microbiol. **35**: 647-651.

Groenen, P.M., Bunschoten, A.E., van Soolingen, D., and van Embden, J.D. (1993). "Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method." Mol. Microbiol. **10**: 1057-1065.

Haft, D.H., Selengut, J., Mongodin, E.F., and Nelson, K.E. (2005). "A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes." PLoS Comput. Biol. **1**: 474-483.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., et al. (2009). "RNA-guided RNAcleavage by a CRISPR RNA-Cas protein complex." Cell. **139**: 945-956.

Hale, C.R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., et al. (2012). "Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs." Mol. Cell. **45**: 292-302.

Horvath, P., Barrangou, R. (2010). "CRISPR/Cas, the immune system of bacteria and archaea." Science **327**: 167-170.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., Joung, J.K. (2013). "Efficient genome editing in zebrafish using a CRISPR-Cas system." Nat. Biotechnol. **31**: 227-229.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. (1987) " Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product." J. Bacteriol. **169**: 5429-5433.

Jansen, R., Embden, J.D., Gaastra, W., and Schouls, L.M. (2002). "Identification of genes that are associated with DNA repeats in prokaryotes." Mol. Microbiol. **43**: 1565-1575.

Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, L.A. (2013). "RNA-guided editing of bacterial genomes using CRISPR-Cas systems." Nat. Biotechnol. **31**: 233-239.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. (2012). "A programmable dual- RNA-guided DNA endonuclease in adaptive bacterial immunity." Science. **337**: 816-821.

Kunin, V., Sorek, R., Hugenholtz, P. (2007). "Evolutionary conservation of sequence and secondary structures in CRISPR repeats." Genome Biol. **8**: R61.1-R61.7.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). "Bacteriophage resistance mechanisms" Nat. Rev. Microbiol. **8**: 317-327.

Lillestøl, R.K., Shah, S.A., Brügger, K., Redder, P., Phan, H., et al. (2009). "CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties." Mol. Microbiol. **72**: 259-272.

Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., Koonin, E.V. (2006). "A putative RNA-interference based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action." Biol. Direct **1**: 7.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., et al. (2011). "Evolution and classification of the CRISPR-Cas systems." Nat. Rev. Microbiol. **9**: 467-477.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M. (2013). "RNA-guided human genome engineering via Cas9." Science. **339**: 819-823.

Marraffini, L.A., Sontheimer, E.J. (2010) (1). "Self versus non-self discrimination during CRISPR RNA-directed immunity." Nature. **463**: 568-571.

Marraffini, L.A,, Sontheimer, E.J. (2010) (2). "CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea." Nat. Rev. Genet. **11**: 181-190.

Mc Grath, S., Fitzgerald, G.F., Sinderen van, D. (2007). “Bacteriophages in dairy products: pros and cons.” **2**: 450-455.

Mojica, F.J., D´ıez-Villase ˜ nor, C., Soria, E., Juez, G. (2000). “Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria.” Mol. Microbiol. **36**: 244-246.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). “Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements.” J. Mol. Evol. **60**: 174-182.

Oost van der, J., Jore, M.M., Westra, E.R., Lundgren, M., Brouns, S.J. (2009). “CRISPR-based adaptive and heritable immunity in prokaryotes.” Trends Biochem Sci. **34**: 401-407.

Ploeg van der, J.R. (2009). “Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages.” Microbiology. **155**: 1966-1976.

Pougach, K., Semenovam, E., Bogdanova, E., Datsenko, K.A., Djordjevic, M., et al. (2010). “Transcription, processing and function of CRISPR cassettes in *Escherichia coli*.” Mol. Microbiol. **77**: 1367-1379.

Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). “CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies.” Microbiology **151**: 653-663.

Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., Wagner, R. (2010). “Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS.” Mol. Microbiol. **75**: 1495-1512.

Sashita,l D.G., Jinek, M., Doudna, JA. (2011). “An RNA induced conformational change required for CRISPR RNA cleavage by the endonuclease Cse3.” Nat. Struct. Mol. Biol. **18**: 680-687.

Sorek, R., Lawrence, M., Wiedenheft, B. (2013). “CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea”. Annu. Rev. Biochem. **82**: 11.1-11.30.

Stern, A., Keren, L., Wurtzel, O., Amita,i G., Sorek, R. (2010) ” Self-targeting by CRISPR: gene regulation or autoimmunity?” Trends Genet **26**: 335-340.

Touchon, M. Charpentier, S., Pognard, D., Picard, B., Arlet, G., Rocha, E.P.C., Denamur, E., Branger, C. (2012). “Antibiotic resistance plasmids spread among natural isolates of Escherischia coli in spite of CRISPR elements.” Microbiology. **158**: 2997-3004.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., Gregory, P.D. (2010). “Genome editing with engineered zinc finger nucleases.” Nat Rev Genet. **11**: 636-646.

Wang, R., Preamplume, G., Terns, M.P., Terns, R.M., Li, H. (2011). “Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage.” Structure. **19**: 257-264.

Westra, E.R., van Erp, P.B.G., Künne, T., Won,g S.P., Staals, R.H.J., et al. (2012). “CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3.” Mol. Cell. **46**: 595-605 (1).

Westra, E.R., Swarts, D.C., Staals, R.H.J., Jore, M.M., Brouns, S.J.J., Oost van der, J. (2012).” The CRISPRs, They Are A-Changin’: How Prokaryotes Generate Adaptive Immunity.” Annu. Rev. Genet. **46**: 311-339 (2).

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S.M., Ma W, Doudna J.A. (2009). “Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated antiviral defense.” Structure **17**: 904-912.

Wiedenheft, B., Sternberg, S.H., Doudna, J.A. (2012). “RNA-guided genetic silencing systems in bacteria and archaea.” Nature **482**: 331-338.

Wilson, G.G. (1991). “Organization of restriction-modification systems.” Nucleic Acids Res. **19**: 2539-2566.

Yosef, I., Goren, M.G., Qimron, U. (2012). “Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*.” Nucleic Acids Res. **40**: 5569-5576.

.