CRISPR in bacteria: a brief summary

50% of bacteria and 90% of archea have a CRISPR "adaptive" defence system. What happens is:
All the cell's DNA is scanned by the CRISPR system. When DNA is recognised as foreign (see below), it is cut to pieces by various enzymes (these are not restriction enzymes). These include cellular nucleases that are involved in homologous recombination. This generates fragments that are processed by the CAS1/2 complex and integrated into the CRISPR array. Each integrated sequence is separated from previous ones by a short spacer sequence. All the spacers are identical.
The CRISPR array, which is part of the genome, is transcribed into an RNA, which is processed by other CAS enzymes to make short RNA fragments that consist of a spacer and one of the unique sequences. These are assembled into a protein complex (one of the members of this is CAS9), which uses this "guide RNA" to scan all the DNA in the cell. If a perfect match is found, then the DNA is cut by CAS9.
The search for the target is extremely rapid (<1s).

Not all bacteria have this system, and even among those that have it, it is not activated all the time in some of them – why is unclear.

How do you avoid cutting the host DNA? This involves the presence of host sequences that are rare in viruses, and the manner in which the initial fragments are generated. The answer is complex, still incomplete, and certainly beyond the scope of this course!

"autoimmunity" – that is putting host DNA into the CRISPR array is, as you would expect, lethal.

CRISPR and restriction/modification systems work in parallel.

In the lab you supply the guide RNA and express CAS9 in the cells you wish to modify. This bypasses the creation of the CRISPR array, and the processing of the CRISPR array transcript.

50 % des bactéries et 90 % des archées ont un système de défense "adaptatif" CRISPR. Voici ce qui se passe :

Tout l'ADN de la cellule est analysé par le système CRISPR. Lorsque l'ADN est reconnu comme étranger (voir ci-dessous), il est coupé en morceaux par diverses enzymes (il ne s'agit pas d'enzymes de restriction). Il s'agit notamment de nucléases cellulaires impliquées dans la recombinaison homologue. Cela génère des fragments qui sont traités par le complexe CAS1/2 et intégrés dans le réseau CRISPR. Chaque séquence intégrée est séparée des précédentes par une courte séquence d'espacement. Tous les espaceurs sont identiques.

Le réseau CRISPR, qui fait partie du génome, est transcrit en ARN, qui est traité par d'autres enzymes CAS pour produire de courts fragments d'ARN constitués d'un espaceur et d'une des séquences uniques. Ces fragments sont assemblés en un complexe protéique (dont l'un des membres est CAS9), qui utilise cet "ARN guide" pour scanner tout l'ADN de la cellule. Si une correspondance parfaite est trouvée, l'ADN est coupé par CAS9.

La recherche pour la séquence cible est extrêmement rapide (<1s).

Toutes les bactéries ne possèdent pas ce système, et même parmi celles qui le possèdent, il n'est pas activé en permanence chez certaines d'entre elles - la raison n'en est pas claire.

Comment éviter de couper l'ADN de l'hôte ? Cela implique la présence de séquences de l'hôte qui sont rares dans les virus, et la manière dont les fragments initiaux sont générés. La réponse est complexe, encore incomplète, et dépasse certainement le cadre de ce cours !

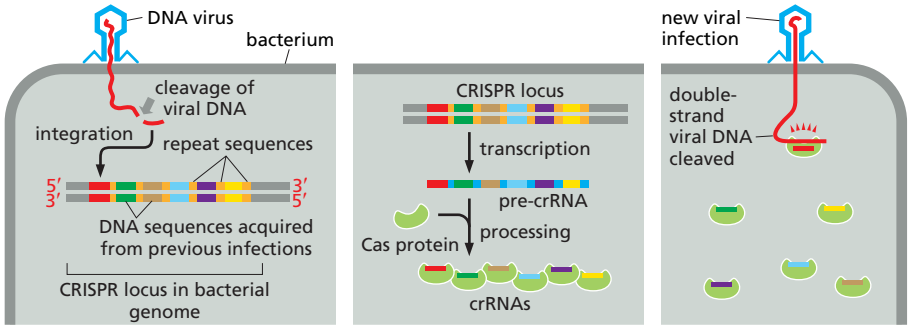
"L'auto-immunité", c'est-à-dire l'introduction de l'ADN de l'hôte dans le réseau CRISPR, est, comme on peut s'y attendre, mortelle.

Les systèmes CRISPR et de restriction/modification fonctionnent en parallèle.

En laboratoire, vous fournissez l'ARN guide et vous exprimez CAS9 dans les cellules que vous souhaitez modifier. Cela permet d'éviter la création du réseau CRISPR et le traitement de l'ARN produit à partir du réseau CRISPR.

Bacteria Use Small Noncoding RNAs to Protect Themselves from Viruses

Bacteria make up the vast majority of the Earth's biomass and, not surprisingly, viruses that infect bacteria greatly outnumber plant and animal viruses. These viruses generally have DNA genomes. A recent discovery revealed that many species of bacteria (and almost all species of archaebacteria) use a repository of



STEP 1: short viral DNA sequence is integrated into CRISPR locus

STEP 2: RNA is transcribed from CRISPR locus, processed, and bound to Cas protein

STEP 3: small crRNA in complex with Cas seeks out and destroys viral sequences

Figure 7–78 CRISPR-mediated immunity in bacteria and archaebacteria. After infection by a virus *(left panel)*, a small bit of DNA from the viral genome is inserted into the CRISPR locus. For this to happen, a small fraction of infected cells must survive the initial viral infection. The surviving cells, or more generally their descendants, transcribe the CRISPR locus and process the transcript into crRNAs *(middle panel)*. Upon reinfection with a virus that the population has already been "vaccinated" against, the incoming viral DNA is destroyed by a complementary crRNA *(right panel)*. For a CRISPR system to be effective, the crRNAs must not destroy the CRISPR locus itself, even though the crRNAs are complementary in sequence to it. In many species, in order for crRNAs to attack an invading DNA molecule, there must be additional short nucleotide sequences that are carried by the target molecule. Because these sequences, known as PAMs (protospacer adjacent motifs), lie outside the crRNA sequences, the host CRISPR locus is spared (see Figure 8–55).

small noncoding RNA molecules to seek out and destroy the DNA of the invading viruses. Many features of this defense mechanism, known as the **CRISPR** system, resemble those we saw above for miRNAs and siRNAs, but there are two important differences. First, when bacteria and archaea are first infected by a virus, they have a mechanism that causes short fragments of that viral DNA to become integrated into their genomes. These serve as "vaccinations," in the sense that they become the templates for producing small noncoding RNAs known as **crRNAs** (CRISPR RNAs) that will thereafter destroy the virus should it reinfect the descendants of the original cell. This aspect of the CRISPR system is similar in principle to adaptive immunity in mammals, in that the cell carries a record of past exposures that is used to protect against future exposures. The second distinguishing feature of the CRISPR system is that these crRNAs then become associated with special proteins that allow them to seek out and destroy double-stranded DNA molecules, rather than single-stranded RNA molecules.

Although many details of CRISPR-mediated immunity remain to be discovered, we can outline the general process in three steps (**Figure 7–78**). In the first, viral DNA sequences are integrated into special regions of the bacterial genome known as CRISPR (clustered regularly interspersed short palindromic repeat) loci, named for the peculiar structure that first drew the attention of scientists. In its simplest form, a CRISPR locus consists of several hundred repeats of a host DNA sequence interspersed with a large collection of sequences (typically 25–70 nucleotide pairs each) that has been derived from prior exposures to viruses and other foreign DNA. The newest viral sequence is always integrated at the 5′ end of the CRISPR locus, the end that is transcribed first. Each locus, therefore, carries a temporal record of prior infections. Many bacterial and archaeal species carry several large CRISPR loci in their genomes and are thus immune to a wide range of viruses.

In the second step, the CRISPR locus is transcribed to produce a long RNA molecule, which is then processed into the much shorter (approximately 30 nucleotides) crRNAs. In the third step, crRNAs complexed with *Cas* (*CRISPR-associated*) proteins seek out complementary viral DNA sequences and direct their destruction by nucleases. Although structurally dissimilar, Cas proteins are analogous to the Argonaute and Piwi proteins discussed above: they hold small single-stranded RNAs in an extended configuration that is optimized, in this case, for seeking and forming complementary base pairs with DNA.

We still have much to learn about CRISPR-based immunity in bacteria and archaebacteria. The mechanism through which viral sequences are first identified and integrated into the host genome is poorly understood, as is the way that the crRNAs find their complementary sequences in double-stranded DNA. Moreover, in different species of bacteria and archaebacteria, crRNAs are processed in different ways, and in some cases, the crRNAs can attack viral RNAs as well as DNAs.

We shall see in the following chapter that bacterial CRISPR systems have already been artificially "moved" into plants and animals, where they have become very powerful experimental tools for manipulating genomes.