



UPPSALA  
UNIVERSITET

# CRISPR/cas-systemet - ett vaccinationskort för prokaryoter?

Lars Nordling

---

Independent Project in Biology  
Självständigt arbete i biologi, 15 hp, höstterminen 2012  
Institutionen för biologisk grundutbildning, Uppsala universitet

# CRISPR/cas-systemet – ett vaccinationskort för prokaryoter?

Lars Nordling

Självständigt arbete i biologi 2012

## Sammandrag

CRISPR (eng. Clustered Regularly Interspaced Short Palindromic Repeats) är ett relativt nyupptäckt adaptivt immunsystem hos prokaryoter som delvis riktar sig mot faginfektioner, men också mot plasmider överförda via horisontell genöverföring. CRISPR finns hos cirka 40% av alla sekvenserade bakterier, och så gott som alla sekvenserade arkéer.

CRISPR är ett RNA-guidat system som integrerar delar av främmande nukleinsyror i ett specifikt CRISPR-lokus. Lokuset är uppdelat i ett antal funktionella operon som är essentiella för CRISPRs funktion, däribland cas-generna (eng. CRISPR associated genes). När lokuset transkriberas bildas korta RNA-strängar kallade crRNA (CRISPR RNA) och cas-proteiner. Tillsammans bildar de proteinkomplexet Cascade (eng. CRISPR-associated complex for antiviral defence), som basparar crRNA till angripande nukleinsyror och oskadliggör dem.

CRISPR har stor potential inom ett flertal användningsområden och används idag bland annat vid mejeriproduktion och artbestämning av prokaryoter. I framtiden hoppas man att CRISPR dessutom ska kunna användas som för att minska spridningen av multiresistens hos bakterier inom sjukhusmiljöer, men än så länge är det mycket som kvarstår att förklaras innan CRISPRs mekanismer och funktioner är helt förstådda.

## Inledning

Prokaryoter är bland de mest anpassliga och utspridda organismerna på vår planet och återfinns i alla tänkbara miljöer: från vulkaner och djupa hav till insidan av oss människor. I antal räknat är de endast överskuggade av virus (Beritbard & Rohwer 2005). Detta enorma antal virus ger såklart upphov till en otrolig diversitet, som genom sitt parasitiska levnadssätt i sin tur skapar en enorm diversitet hos sina värddjur. Ett virus kan infektera prokaryota celler genom att injicera delar av sitt egna DNA i sin värd, som därefter integreras och gör att cellen blir en veritabel virusfabrik vilket slutligen leder till att cellen lyseras. Man tror att så mycket som upp till 50% av alla prokaryoter dör som resultat av faginfektion (Beritbard & Rohwer 2005).

Både prokaryoter och eukaryoter har genom selektion och evolution utvecklat avancerade immunförsvar mot infektioner. Att eukaryoter har ett komplext och välutvecklat

immunförsvar som dessutom har kapacitet för anpassning har varit känt länge. För prokaryoter har man länge lagt fokus väldokumenterade restriktionssystem där speciella metyltransferaser metylerar bakteriens eget DNA vid specifika platser. Detta gör att det egna genomet skyddas mot cellegna restriktionsenzymer, medan främmande DNA från virus/plasmider saknar rätt metylering, och bryts därmed ned (Labrie & Samson 2010). En annan vanlig försvarsmekanism är abortiv infektion: genom att stoppa sin egen translation kan en infekterad cell förhindra vidare produktion och spridning av virus (Labrie & Samson 2010). Nyare studier har dock mer och mer riktat fokus mot ett anpassningsbart nedärvt system, CRISPR (eng. Clustered Regularly Interspaced Short Palindromic Repeats), som med hjälp av RNA söker upp och eliminerar främmande gener (Makarova *et al.* 2011). CRISPR är slående likt ett eukaryot försvarssystem, RNA interferens, men i takt med att forskningen gjort framsteg har också stora skillnader trätt fram (Makarova *et al.* 2006).

CRISPR är ett lokus med korta upprepade sekvenser som kan spara främmande DNA inför framtida attacker, och fungerar på så sätt som en sorts vaccination. CRISPR lokuset kan därefter transkriberas till ett stort antal så kallad CRISPR RNA (crRNA) som i princip agerar antikroppar (Maraffinini & Sontheimer 2010). Varje crRNA överensstämmer med sekvenser från en tidigare påträffad angripande nukleinsyra, antingen från fager eller från en plasmid. Tillsammans kan crRNA bilda ett enormt komplex som patrullerar cellens inre, i jakt på inkräktare (Brouns *et al.* 2011).

CRISPR observerades för första gången hos *Escherichia coli* 1987, men man hade då ingen aning om vad man egentligen hittat, utan såg bara ett mönster av regelbundna upprepningar med oregelbundna sekvenser mellan (Ishino *et al.* 1987). Samma mönster uppmärksammades hos andra prokaryoter och fick namnet CRISPR år 2002 (Jansen 2002). Det var dock först när man insåg att de oregelbundna sekvenserna överensstämde med sekvenser funna hos virus och plasmider som man kom med hypotesen att CRISPR fungerade som ett adaptivt immunsystem hos prokaryoterna (Bolotin 2005).

Som resultat fullkomligt exploderade forskarvärlden med forskning kring CRISPR och fler och fler fick liknande resultat. Så småningom började funktionen och omkringliggande mekanismer träda fram. *cas*-generna (eng. CRISPR associated genes) – en grupp med hjälpproteinerna som är grundläggande för CRISPRs alla funktioner och mekanismer – definierades 1999 (Overbeek 1999), och i takt med ökad förståelse och framförallt sekvensering börjar bilden klarna. Man räknar med att drygt 40% av de sekvenserade bakterierna och så gott som alla arkéer har CRISPR. Dock måste det tilläggas att denna siffra inte nödvändigtvis är representativ, då sekvensering främst görs på den minimala del bakterier som kan odlas i labb (Kunin *et al.* 2007).

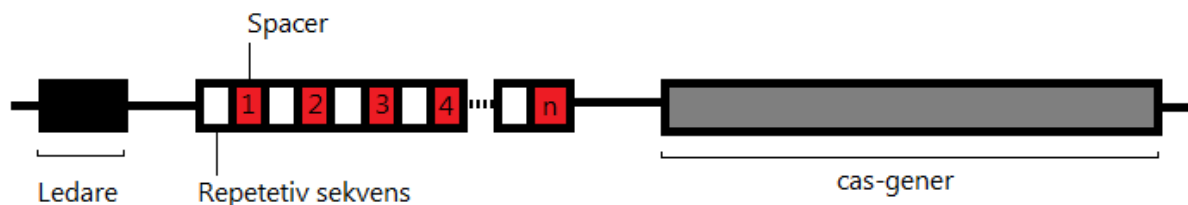
CRISPR är ett ämne som fortfarande är mycket aktuellt, då stora delar av systemet och dess funktioner fortfarande är ett mysterium. Syftet med följande sammanfattande artikel är att sprida grundläggande och överblickande kunskap om CRISPR/cas-systemets uppbyggnad, funktion och praktiska användningsområden.

## CRISPR lokus

CRISPR består av en serie upprepade och till viss del palindromiska repetitiva sekvenser, skilda av främmande sekvenser från virus, oftast kallade spacers (Garneau *et al.* 2010). Det repetitiva segmentet är oftast mellan 23 och 47 baspar långt, medan en spacer är 21 till 72 bp (Jansen & Embden 2002). Uppbyggnaden och storleken på det repetitiva sekvenserna och CRISPR-lokuset skiljer sig ofta från organism till organism, men principen är alltid densamma (figur 1).

Besläktade organismer har ofta CRISPR-lokus av liknande storlek och uppbyggnad, dock har man funnit en relativt stor variation generellt sett (Kunin *et al.* 2007). Vanligtvis hittar man ett tjugotal och spacersegment i varje lokus. En viss palindromitet hos de repetitiva sekvenserna ger upphov till stabila sekundärstrukturer i form av en hårnåls-/öglestruktur (Kunin & Sorek 2007). Terminalt på 3' har man dessutom sett tecken på en vanlig återkommande sekvens (GAA), som eventuellt kan fungera som bindningsställe för de cas-proteiner som driver CRISPR (Grissa 2007).

En sekvens med promotorlik funktion ligger nära intill spacer-segmentet och kallas för ledaränden (Bult *et al.* 1996). Den ligger alltid i änden på lokuset och det är här som immunförsvaret så småningom byggs på med nya virus-/plasmidsekvenser. Ledaränden består av en okodande AT rik sekvens som kan vara hundratals baspar lång. Dess funktion gör den till en grundläggande och viktig del av CRISPR systemet, då CRISPR-lokus utan ledaroperonet saknar förmågan att lägga till nya spacers (Lillestol *et al.* 2006).



**Figur 1.** Schematisk figur över den generella uppbyggnaden och principen för ett CRISPR-lokus. Ledarsekvensen har en promotorlik funktion, det är i denna ände som CRISPR-lokuset växer i takt med att nya repetitiva och spacersegment läggs till. Spacersekvensen ligger mellan två identiska repetitiva sekvenser och innehåller genetiskt material överensstämmande med en invaderande nukleinsyra. *cas*-operonet kodar för så kallade cas-protein, vilka är essentiella för systemets funktion (Garneau *et al.* 2010).

Generellt hittar man CRISPR-lokuset på kromosomen, men har också observerats på plasmider, vilket alltså skulle indikera på en spridning mellan individer likväl som arter (Kunin & Sorek 2007). Ytterligare har man också kunnat påvisa hur det finns potential till att ha ett stort antal skilda CRISPR-lokus. Till exempel har *Methanocaldococcus jannaschii* hela 18 olika CRISPR lokus som tillsammans står för 1% av dess totala genom (Lillestol *et al.* 2006).

## ***cas*-generna**

När man först stötte på den täta klunga av gener vi idag kallar *cas*-gener fann man dem proximalt till CRISPR-lokuset (Makarova *et al.* 2002). Initialt trodde man att deras funktion var bunden till DNA-reparation (Makarova *et al.* 2002), men när man så småningom insåg att *cas*-generna endast påträffades i samband med CRISPR lokus (Jansen *et al.* 2002) tog det inte lång tid innan de kopplades samman. Till en början kunde man endast identifiera fyra *cas*-gener: *cas* 1-4 (Jansen *et al.* 2002). Vidare forskning av flera forskargrupper kunde snart därefter avslöja ytterligare 25 (Haft *et al.* 2005, Makarova *et al.* 2006).

Även om deras funktion fortfarande inte är helt fastställd, är det klart att de är essentiella för systemets funktion (tabell 1). Aktuella studier av CRISPR-komplexet visar att *cas*-proteiner är starkt kopplade till igenkännandet av komplementära sekvenser för crRNA (Wiedenheft 2011) och 2005 identifierades sekvenser hos *cas*-generna som oftast associeras med nukleaser, helikaser, polymeraser och polynukleotid-bindningsproteiner (Haft 2005).

*cas*-1 till och med *cas*-6 utgör kärnan bakom *cas*-proteinerna. *cas*-1 och *cas*-2 återfinns i så gott som samtliga CRISPR-system, och används i stor utsträckning som markörer för CRISPR i ett genom. Vidare delar man upp proteinerna beroende delvis på vilken organism de tillhör, men till viss del också vilken funktion de har (Makarova *et al.* 2006).

**Tabell 1.** Detaljerad beskrivning av de enskilda *cas*-proteinerna, samt deras uppskattade funktion. Sammanfattad från en mer detaljerad och utförlig tabell ur Matthijs *et al.* 2011.

|                          | <b>Namn</b>   | <b>Aktivitet</b>   |
|--------------------------|---|--|
| <b>Spacerintegrering</b> | <i>cas</i> 1  | DNA endonukleas <sup>1</sup> , RNA och DNA bindning <sup>2</sup> |
|                          | <i>cas</i> 2  | Ribonukleas <sup>3</sup>   |
|                          | <i>cas</i> 4  |  |
| <b>crRNA-syntes</b>      | <i>cas</i> E (Cascade), <i>cas</i> 6<br>cse-komplex (Cascade) | pre-crRNA klyvning <sup>4</sup>                                  |
|                          | <i>cas</i> 3  | Nukleas <sup>5</sup>   |
| <b>Interferens</b>       | cmr-komplex   | Klyver RNA komplementärt till crRNA <sup>6</sup>                 |

<sup>1</sup> Wiedenheft *et al.* 2009, <sup>2</sup> Han *et al.* 2009, <sup>3</sup> Beloglazova *et al.* 2008, <sup>4</sup> Brouns *et al.* 2008, Carte *et al.* 2008, <sup>5</sup> Han & Krauss 2009, <sup>6</sup> Hale *et al.* 2008.

*cas*-1 tros vara ett nukleas/integras med förmågan att klyva DNA som dessutom har en roll i integreringen av nya spacers i CRISPR-lokuset. *cas*-2 har visat sig vara ett ribonukleas som klyver enkelsträngat RNA vid U-rika sekvenser (Makarova *et al.* 2006). Observationer av *cas*-1 och *cas*-2 har visat på hur de inte är närvarande vid försvarsstadiet av CRISPR (Hale *et al.* 2009), utan tycks snarare vara endonukleaser inblandade i bildningen av crRNA (Carte *et al.* 2008)

cas-5 och cas-6 är så kallade RAMP-proteiner (eng. Repeat-associated mysterious proteins). RAMP är en relativt splittrad grupp. De består av proteiner som inte nödvändigtvis sitter tätt bundna vid CRISPR lokuset, men endast påträffats i genom som har ett CRISPR-lokus (Makarova *et al.* 2006).

Utöver de mer generella *cas*-generna har åtta underkategorier definierats (Csa, Csd, Cse, Csh, Csm, Csn, Cst samt Csy), beroende på kompositionen av operonet och vilken organism de påträffats i – Cse står exempelvis för *cas*-gener från *E. coli* (Haft *et al.* 2005). Man har också sett tecken på att vissa organismer kan ha flera sorters CRISPR, vilket skulle antyda att dessa kan samarbeta och jobba unisont (Makarova *et al.* 2011).

## CRISPR - Mekanismer och funktion

CRISPR-systemets funktion blir oftast indelad i tre separata faser: till att börja med byggs och ansamlas spacers för att etablera resistens. Därefter transkriberas CRISPR-lokusset innehållande resistensgenerna och klyvs till mindre RNA-segment. RNA-segmenten utnyttjas i sista steget, då stora CRISPR-komplex använder dem för att baspara mot främmande DNA (Barrangou *et al.* 2008).

### ”Vaccinering”

2008 kunde Barrangou *et al.* presentera de första bevisen på att CRISPR är en del av det prokaryota immunförsvaret, och dessutom att det är adaptivt. Genom att exponera *Streptococcus thermophilus* för fager såg man hur den större delen av kulturen dog. Den överlevande delen som återstod hade mycket riktigt tagit upp nytt fag-DNA. Samma studie visar också att resistensen är beroende av spacers. Genom splicing kunde man ta bort samma spacer från de resistenta *S. thermophilus* och observera att den nyfunna resistensen också försvann (Barrangou *et al.* 2008).

Främmande DNA tillförs till prokaryoterna genom transduktion från en invaderande fag. Den egentliga mekanismen är fortfarande okänd, men en begränsad del kallad proto-spacer av den främmande DNA-strängen inkorporeras i CRISPR lokusets ledarände. När spacersekvensen lagts till byggs en ny repetitiv segment i förberedelse för nästa spacer-segment (Barrangou *et al.* 2007).

Mekanismen för uppbyggnaden av den nya repetitiva sekvensen är fortfarande okänd, men studier gjorda på *S. thermophilus* indikerar en stark koppling med *cas*-generna (Deveau *et al.* 2008). Flera studier har dock sett tecken på att cas-1 och cas-2 är inblandade i spacertillägget (Hale 2006, Babu 2011).

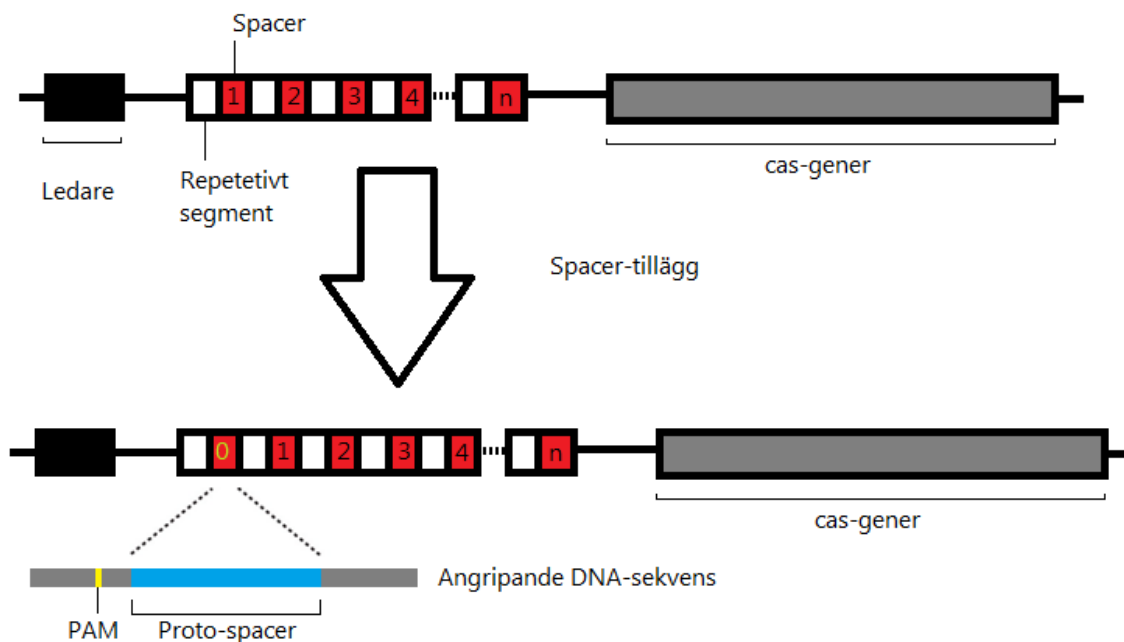
### PAM

Vilken del av den främmande sekvensen som väljs ut till att bli en protospacer bestäms av en grupp sekvenser kallade PAM: (eng. Protospacer adjacent motif). Genom studier av *Streptococcus thermophilus* kunde man se ett samband mellan de protospacers som valdes och närliggande sekvenser (Deveau *et al.* 2008). Snarlika resultat dök upp i liknande undersökningar (Boltin *et al.* 2005), och senare studier av arkéefamiljen *Sulfolobales*

(Lillestol *et al.* 2009) visar också på tydliga preferenser för valet av protospacers, som sammanfaller med närvaron av specifika PAM-sekvenser.

Varje PAM tycks vara en kort sekvens på endast ett par baspar, och man tror att PAM-sekvensen är specifik för varje unikt ledaroperon. Denna hypotes har uppstått ur studier av *E. coli* (Mojica *et al.* 2009) där man sett tecken på att liknande ledarsekvenser i stor grad tycks föredra samma PAM. Samtidigt kunde man också observera att det saknades en preferens för kodande eller icke-kodande RNA-strängar, och endast närvaron av en överensstämmande PAM var avgörande (figur 2) (Mojica *et al.* 2009).

Det förefaller alltså som att varje CRISPR-system har en mycket specifik utgångspunkt vid valet av protospacers, vilket i fallet med flera CRISPR-typer i en individ skulle ge ett relativt brett försvarsspektrum.



**Figur 2.** En schematisk bild över processen av en spacertillägning. En ny spacer adderas på ledarsidan av CRISPR lokuset, och består utav en proto-spacer. Varje protospacer känns igen av CRISPR systemet genom att binda till en ledar specifik PAM-sekvens, som ligger tätt i anslutning till varje proto-spacer (Mojica *et al.* 2009).

### crRNA syntes

Nästa steg i försvarssystemet är att CRISPR-lokuset transkriberas till så kallade CRISPR-RNA (crRNA). Studier gjorda på *Archaeoglobus fulgidus* (Tang *et al.* 2002) och *Sulfolobus solfataricus* (Tang *et al.* 2005) har genom analys av de repetitiva segmenten visat tecken på att CRISPR först transkriberas till en lång förupplaga, som sedan klyvs mitt i upprepningen, för att på så sätt bilda individuella crRNA innehållande unika spacersekvenser från olika plasmider eller virus.

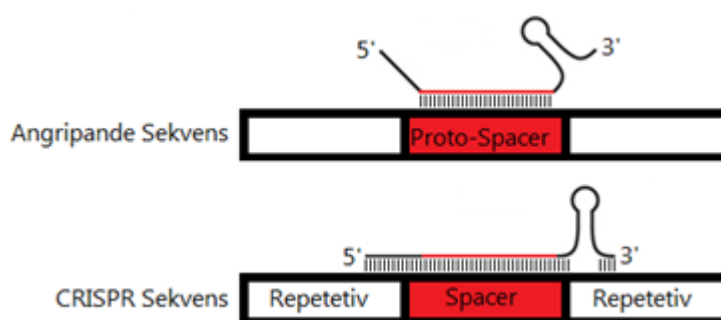
Ännu kan man inte säkert säga hur syntesen går till, men en och en börjar mekanismerna träda fram. Initialt binder cas-6e till de repetitiva sekvenserna och klyver dem till mindre segment. De repetitiva segmenten är delvis palindromiska och bildar positivt laddade hårnålsstrukturer (Sashital 2011). cas-proteinet känner igen hårnålsstrukturen och kan genom att ändra konformationen på de längst understa basparet, klyva isär de repetitiva segmentet och bilda crRNA. Samtidigt fäster sig proteinet vid 3' änden av öglan som bildats, och man tror att proteinet då får möjligheten att binda till andra crRNA vid bildningen av Cascade-komplex (eng. CRISPR-associated complex for antiviral defence) (Sashital 2011).

Stora delar av mekanismerna bakom crRNA är fortfarande oklara, mycket eftersom man än inte helt lyckats utröna funktionen hos cas-generna.

### Cascade-komplex

Det verkliga försvaret ser vi när crRNA går samman för att bilda stora CRISPR ribonukleoproteinkomplex: Cascade-komplex. Det är i detta stadiet som spacersekvenserna utnyttjas för att känna igen invaderande DNA. Genom basparning kan spacersekvensen användas som en mall för att känna igen främmande nukleinsyror (Garneau *et al.* 2010). På så sätt fungerar CRISPR lokuset som ett slags vaccinationskort, innehållande spacers från alla fager den infekterats av.

Bindandet till det komplementära DNA-strängen är ett ömtåligt system, och en enda punktmutation har visat sig kapabel att helt ha slå ut CRISPRs funktion (Semenova *et al.* 2011). Effektiviteten hos komplexet är dessutom beroende av antalet spacers som svarar för varje angripare: fler spacers som kodar för gener hos ett specifikt virus ökar chansen att den främmande gensekvensen kan kännas igen (Barrangou *et al.* 2007).



**Figur 3.** Schematisk bild som föreställer den mekanism som Cascade utnyttjar för att skilja celleget DNA från främmande. Den angripande DNA-sekvensen saknar de repetitiva segment som omringar en spacer hos det cellegeta genom. Då crRNA till viss del består av rester från de repetitiva sekvenserna kan hela crRNA binda till den cellegeta sekvensen, men inte den angripande, som endast har ett proto-spacer segment (Marraffini & Sontheimer 2010).

För att CRISPR ska kunna fungera utan att angripa sitt eget genom måste systemet kunna särskilja de spacersekvenser som finns sparade i CRISPR-lokuset ifrån främmande DNA. Cascade-komplexet känner igen det cellegeta genom genom att hela crRNA-sekvensen kan binda, inkluderat resterna från det repetitiva segment som fortfarande återstår. De främmande DNA-sekvenserna

innehåller endast en proto-spacer och saknar de omkringliggande repetitiva segmenten, därmed kan endast spacer-segmentet av crRNA binda (figur 3) (Marraffini & Sontheimer 2010).



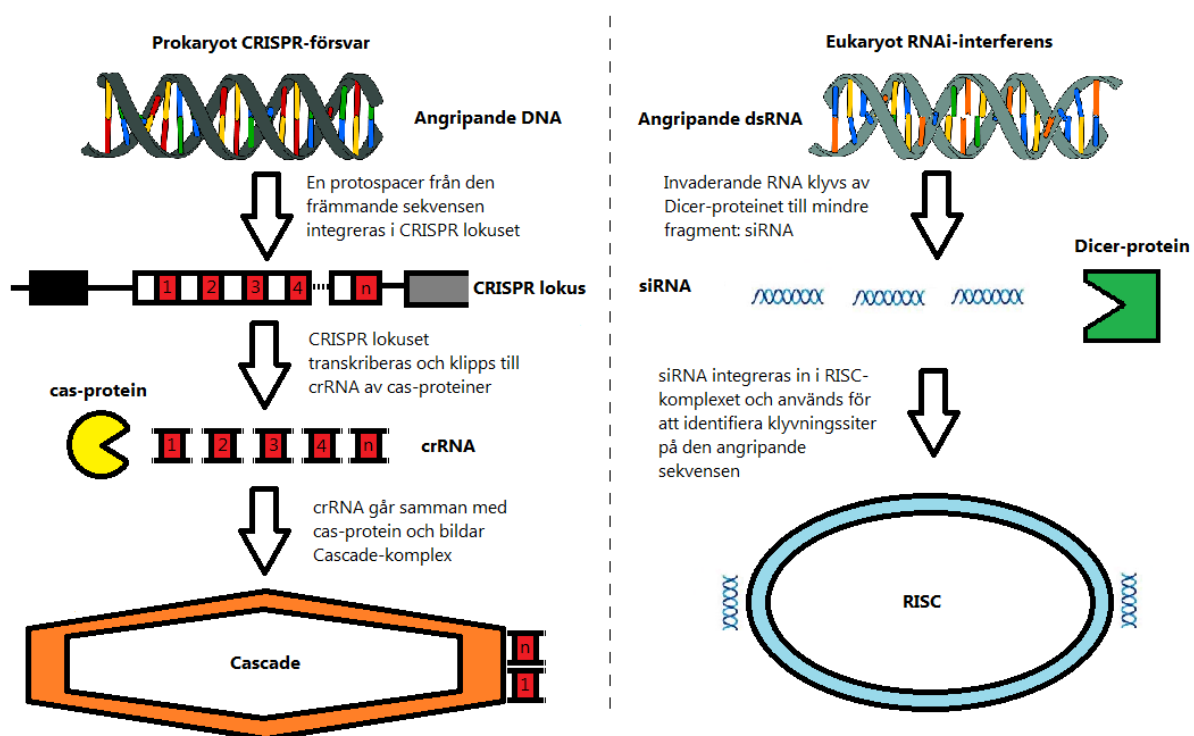
Studier av *Pyrococcus furiosus* har visat på hur vissa varianter av CRISPR har valt att fokusera på RNA istället för att angripa DNA. Detta medför att systemet för att skilja eget från främmande DNA är överflödigt då RNA transkriberas ensträngt, och alltså inte producerar en komplementär sekvens att angripa. Detta skulle då innebära att denna typ av CRISPR kan vara fördelaktig i försvar mot RNA-baserade virus (Hale *et al.* 2009).

## CRISPR är ett system under konstant utveckling

CRISPR är ofta relativt olik hos närbesläktade organismer. Spridningen av spacers kan vara enorm, även inom en enda biofilm, samtidigt som plasticiteten är stor, då de äldsta ofta tas bort för att göra plats för nya (Pourcel 2005). I samma takt som CRISPR evolverar och lägger till nya spacersekvenser utvecklas samtidigt fagera. En evig kapprustning pågår, där endast en part kan stå som segrare. Eftersom en fag som upptäcks av CRISPR oskadliggörs är det selektiva trycket på virusen högt. Enskilda punktmutationer i en protospacer eller PAM leder ofta till att CRISPR måste börja om på nytt (Devau *et al.* 2008).

## Hur skiljer CRISPR sig från eukaryot RNA interferens?

Hos eukaryota organismer finns ett system som till synes är slående likt prokaryoternas CRISPR. RNA interferens bygger på samma princip som CRISPR. Ett protein kallat Dicer klyver dubbelsträngat RNA som injicerats i den eukaryota cellen. Resultatet blir ett stort antal små dubbelsträngiga RNA-segment. Dessa används av ett stort proteinkomplex, RISC (eng. RNA induced silencing complex). RISC utnyttjar den kodande delen av RNA-strängen för att baspara med liknande sekvenser och klyva dem (Rand *et al.* 2005)



**Figur 4.** En jämförande schematisk bild av det prokaryota CRISPR systemet ställt mot det eukaryota RNAi-systemet. Likheterna är tillsynes många. Båda bildar stora patrullerande komplex som använder små RNA fragment med sekvenser ursprungligen från angriparen för att genom basparning binda och känna igen en angripande nukleinsyra (Barrangou *et al.* 2007, Rand *et al.* 2005)

Trots att likheterna med CRISPR (figur 4) är stora kan man inte hitta något genetiskt släktskap mellan de två systemen (Makarova *et al.* 2006). Och allt eftersom mera forskning som görs på området, desto fler skillnader träder dessutom fram: CRISPR är ett adaptivt nedärvt försvar, där crRNA transkriberas direkt från kromosomen. RNAi måste vid varje tillfälle klyva nya RNA utifrån den invaderande strängen. Dessutom har RNAi valt att lägga sitt fokus på att söka upp och bryta ned RNA. CRISPR siktar generellt på att bryta ned DNA (Makarova *et al.* 2006).

### **CRISPR:s praktiska tillämpningar**

CRISPR är och kan vara ett kraftfullt verktyg inom både medicin, mikrobiell forskning och till och med mejeriindustrin. Än så länge är visionerna och potentialerna stora, i väntan på att vidare forskning ska öppna upp för framtiden.

#### *Horisontell genöverföring*

Stor del av den variation vi ser hos prokaryoter har sin grund det fenomen som kallas horisontell genöverföring (HGT). Genom transformation, transduktion och konjugation sprider och delar bakterier och arkéer konstant genetiskt material sinsemellan varandra (Nakamura 2004). Detta är dock både på gott och ont, då endast en bråkdel av det överförda materialet är fördelaktigt för organismen (Thomas 2005).

HGT är en av de viktigaste faktorerna vad det gäller spridningen av virulensfaktorer så som antibiotikaresistens inom sjukhusmiljöer, exempelvis hos resistenta *Staphylococcus aureus* där man visat på hur resistensen är direkt kopplad till plasmidöverföring via konjugation (Thomas 2003). Man har med hjälp av CRISPR kunnat minimera överföringen av specifika gener via HGT hos *S. aureus*, och hoppas nu att framtiden kommer medföra kliniska metoder som kan hindra spridningen av multiresistens (Maraffini & Sontheimer 2008).

Studier gjorda på *Streptococcus pyrogenes* visar på hur många av de gener som kodar för virulensfaktorer har sitt ursprung hos virus, men har genom HGT spridits vidare till bakterier och på så sätt gjort dem mera patogena (Banks *et al.* 2002). De bakteriegenom som har CRISPR innehåller oftast inte några virulensgener från virus, medan man samtidigt sett tecken på att de bakterier som saknar CRISPR i större utsträckning har virala virulensgener. Detta visar på hur CRISPR ytterligare kan fungera som ett försvar mot HGT, inte bara genom att förhindra faglysering, utan dessutom genom att förhindra att virala virulensgener integreras i genomet (Grissa 2007).

#### *Genreglering*

Teoretiskt sett skulle CRISPR kunna användas som verktyg inom mikrobiell forskning. Genom att programmera CRISPR skulle man kunna slå ut en eller flera specifika gener med hög noggrannhet, utan att manipulera det ursprungliga genomet. Genom att föra in den utvalda genen eller generna som en spacer i ett CRISPR system, och därefter överföra samma CRISPR lokus till den valda organismen via transformation, skulle man kunna hitta ett mycket specifikt och användbart verktyg för framtida forskning (Sorek *et al.* 2008).

## *Spoligotyping*

Eftersom CRISPR-lokuset är nedärvt kommer det i stort sett vara mycket liket inom en och samma bakteriestam. Denna egenskap gör att CRISPR-lokuset kan utnyttjas för att artbestämma bakterier och till och med skilja stammar och populationer från varandra. Metoden kallas för Spoligotyping (eng. Spacer oligonucleotide typing), och har visat sig vara mycket specifikare än många andra bestämningsmetoder (Kamerbeek *et al.* 1997).

## *Mejeriförurning*

En av de stora aktörerna vad det gäller forskning inom fagresistens hos prokaryoter är mejeriindustrin (Sturino 2006). *S. thermophilus* är en av de mest essentiella arkeéstammarna inom mejeriproduktion, där den är står för förurningen av mjölkprodukter. En faginfektion kan i denna bransch innebära stora ekonomiska förluster genom att störa fermentationsprocessen, försämra slutprodukten och förhindra produktionskedjan. Därför är just mejeribranschen, och däribland företaget DANISCO, en av de mest engagerade och aktiva vad det gäller just forskning angående immunförsvar hos prokaryoter. Att kunna välja och vraka bland de prokaryoter som ger högst resistens mot de vanligaste fagera eller själva konstruera immuna stammar är ett drömscenario (Mc Grath 2007).

## **Diskussion**

CRISPR är ett komplext system som återfinns hos en stor del av alla prokaryota organismer (Beritbard & Rohwer 2005). Dess funktion tycks vara att identifiera och oskadliggöra främmande angripande nukleinsyror, både i form av plasmider och fager (Makarova *et al.* 2011). Mekanismerna bakom CRISPR är till stor del fortfarande okända, men forskningen går snabbt framåt.

CRISPR fungerar genom att integrera delar av den angripande DNA-sekvensen i sitt eget genom, i form av så kallade spacers (Garneau *et al.* 2010). CRISPR lokuset transkriberas och klyvs till små crRNA-segment som utnyttjas av ett stort Cascade-komplex för att binda till motsvarande sekvenser hos det angripande DNA:t (Tang *et al.* 2005). Essentiella för processen är de så kallade cas-proteinerna, en grupp av diverse sorters proteiner med olika funktion i systemet (Haft *et al.* 2005, Makarova *et al.* 2006).

Som tidigare nämnts i artikeln muteras och utvecklas virus i takt med att deras genom integreras i CRISPR. En ständig dragkamp tycks pågå, där det selektiva trycket hela tiden gör att båda parter är under konstant utveckling. Vad händer då CRISPR integrerar en gen som är livsnödvändig för virusets överlevnad till en spacer?

En annan stor fråga är varför det tycks finnas en så stor skillnad i förekomsten av CRISPR mellan prokaryoterna själva? Eftersom systemet har en skyddande funktion förefaller det självklart att det också borde gynnas av selektion, vilket också återspeglas i det faktum att så gott som alla arkéer som sekvenserats har CRISPR. Hur kommer det sig då att endast drygt hälften av de sekvenserade bakterierna har CRISPR (Beritbard & Rohwer 2005)?

CRISPR är ett intressant ämne med många outforskade frågetecken. Mycket måste fortfarande fastställas innan vi kan få en tydlig och heltäckande bild över hur de olika mekanismerna fungerar och samspelar.

## Tack

Jag vill tacka mina medstudenter Linnea Blomberg och Bedreddin Acar samt min handledare Jan Andersson för återkoppling och förslag under arbetets gång.

## Referenser

- Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, Gagarinova A, Pogoutse O, Brown G, Binkowski A, Phanse S, Joachimiak A, Koonin E, Savchenko A, Emili A, Greenblatt J, Edwards AM, Yakunin AF. 2011. A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Molecular Microbiology* **79**: 484–502.
- Banks D J , Beres SB, Musser JM. 2002. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. *Trends Microbiology* **10**: 515–521.
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–1712.
- Beloglazova N, Brown G, Zimmerman MD, Proudfoot M, Makarova KS, Kudritska M, Kochinyan S, Wang S, Chruszcz M, Minor W, Koonin EV, Edwards AM, Savchenko A, Yakunin AF. 2008. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *Journal of Biological Chemistry* **283**: 20361–20371.
- Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**: 2551–2561.
- Breitbart M, Rohwer F. 2005. Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology* **6**: 278-284.
- Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**: 960–964.
- Bult, CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, , Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL,

- Geoghagen NS, Venter JC. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**: 1058–1073.
- Carte J, Wang R, Li H, Terns RM, Terns MP. 2008. Cas6 is an endoribonuclease that generates guideRNAs for invader defense in prokaryotes. *Genes Dev* **22**: 3489–3496.
- Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* **190**: 1390–1400.
- Garneau J E, Dupuis MÈ, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**: 67–71.
- Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8**: 172.
- Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology*, doi:10.1371/journal.pcbi.0010060.
- Hale C, Kleppe K, Terns R M, Terns MP. 2008. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **14**, 2572–2579.
- Hale C, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP. 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**: 945–956.
- Han D, Lehmann K, Krauss G. 2009. SSO1450--a CAS1 protein from *Sulfolobus solfataricus* P2 with high affinity for RNA and DNA. *FEBS Letters* **583**: 1928–1932.
- Hu X, Van der Auwera G, Timmery S, Zhu L, Mahillon J. 2009. Distribution, diversity, and potential mobility of extrachromosomal elements related to the *Bacillus anthracis* pXO1 and pXO2 virulence plasmids. *Applied and Environmental Microbiology* **75**: 3016–3028.
- 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**: 5429–5433.
- Jansen R, Embden JD, Gastra W, Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* **43**: 1565–1575.
- Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J. 1997. Simultaneous Detection and Strain Differentiation of *Myobacterium tuberculosis* for Diagnosis and Epidemiology. *Journal of Clinical Microbiology* **35**: 907-914.
- Kunin V, Sorek R, Hugenholtz P. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biology* **8**: doi:10.1186/gb-2007-8-4-r61.
- Labrie SJ, Samson JE, Moineau S. 2010. Bacteriophage resistance mechanisms. *Nature Reviews Microbiology* **8**: 317–327.
- Lillestol RK, Redder P, Garrett RA, Brugger K. 2006. A putative viral defence mechanism in archaeal cells. *Archaea* **2**: 59–72.

- Lillestol RK, Shah SA, Brugger K, Redder P, Phan H, Christiansen J, Garrett RA. 2009. CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Molecular Microbiology* **72**: 259–272.
- Marraffini LA, Sontheimer EJ. 2008. CRISPR interference limits horizontal gene transfer in *staphylococci* by targeting DNA. *Science* **322**: 1843–1845.
- Marraffini LA, Sontheimer EJ. 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**: 568–57.
- Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. 2002. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research* **30**: 482–496.
- Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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. *Biology Direct* **1**: 7.
- Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, van der Oost J, Koonin EV. 2011. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology* **9**, 467–477.
- Matthijs M, Jore S, Brouns JJ, van der Oost J. 2012. RNA in Defense: CRISPRs Protect Prokaryotes against Mobile Genetic Elements. *Cold Spring Harb Perspectives in Biology*, doi: 10.1101/cshperspect.a003657.
- Mc Grath S, Fitzgerald GF, van Sinderen D. 2007. Bacteriophages in dairy products: pros and cons. *Biotechnology Journal* **2**: 450–455.
- Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**: 733–740.
- Nakamura Y, Itoh T, Matsuda H, Gojobori T. 2004. Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nature Genetics* **36**: 760–766.
- Overbeek R, Fonstein M, D'Souza M, Pusch GD, Maltsev N. 1999. The use of gene clusters to infer functional coupling. *Proceedings of the National Academy of Sciences* **96**: 2896–2901.
- Pougach K, Semekova E, Boqdanova E, Datsenko KA, Djordjevic M, Wanner, BL, Severinov K. 2010. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Molecular Microbiology*. **77**: 1367–1379.
- Pourcel C, Salvignol G, 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.
- Rand TA, Petersen S, Du F, Wang X. 2005. Argonaute2 cleaves the antitarget strand of siRNA during RISC activation. *Cell* **123**: 621–629.
- Sashital DG, Jinek M, Doudna JA. 2011. An RNA induced conformational change required for CRISPR RNA cleavage by the endonuclease Cse3. *Nature Structural & Molecular Biology* **18**: 680–687.
- Semenova E, Jore MM, Datsenko KA, Semenkova A, Westra ER, Wanner B, van der Oost J, Brouns SJJ, Severinov K. 2011. Interference by clustered regularly interspaced short

- palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences* **108**: 10098–10103.
- Sorek R, Kunin V, Hugenholz P. 2008. CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology* **6**: 181–186.
- Sturino JM, Klaenhammer TR. 2006. Engineered bacteriophage-defence systems in bioprocessing. *Nature Reviews Microbiology* **4**: 395–404.
- Tang T H, Bachelier JP, Rozhdestvensky T, Bortolin ML, Huber H, Drungowski M, Elge T, Brosius J, Hüttenhofer A. 2002. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proceedings of the National Academy of Sciences* **99**: 7536–7541.
- Tang TH, Polacek N, Zywicki M, Huber H, Brugger K, Garrett R, Bachelier JP, Hüttenhofer A. 2005. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. *Molecular Microbiology* **55**: 469–481.
- Thomas CM, Nielsen KM. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology* **3**: 711–721.
- Weigel LM, Clewell DB, Gill SR, Clark NC, McDougal LK, Flannagan SE, Kolonay JF, Shetty J, Killgore GE, Tenover FC. 2003. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science* **302**: 1569–1571.
- Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA. 2009. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated antiviral defense. *Structure* **17**: 904–912.
- Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, Westphal W, Heck AJR, Boekema EJ, Dickman MJ, Doudna JA. 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proceedings of the National Academy of Sciences* **108**: 10092–10097.