

# RNAi & CRISPR/Cas

## Einführung

Mit RNAi und CRISPR/Cas9 ist die genetische Forschung in den letzten Jahren um zwei ganz wichtige Werkzeuge bereichert worden. Diese beiden RNA-basierten Methoden erlauben es, die genetische Information eines Organismus mit bisher unbekannter Effizienz und Spezifität zu manipulieren. In RNAi-Experimenten wird dabei die Expression von Genen beeinflusst, wohingegen in CRISPR/Cas9-Experimenten die DNA-Sequenz selber verändert wird.

Die Spezifität dieser Werkzeuge wird in beiden Fällen durch RNA-Moleküle gesteuert und die Zielsequenz ist gemäss Watson-Crick-Basenpaarungsregeln bestimmt. Dadurch ist es - zumindest theoretisch - möglich, jeden beliebigen Locus im Genom direkt anzusteuern.

Beiden Methoden ist auch gemein, dass sie auf natürlichen, zellulären Verteidigungsmechanismen gegen fremde Nucleinsäuren basieren.

In dieser Lerneinheit werden Sie die ursprünglichen biologischen Funktionen von RNAi und CRISPR/Cas9 kennenlernen, sehen wie diese Systeme für ihre Funktion als experimentelle Werkzeuge adaptiert wurden und welche genetische Experimente sie nun ermöglichen.

## RNA-Interferenz (RNAi)

### RNAi: ein weitverzweigter zellulärer Mechanismus zur post-translationalen Kontrolle von Genaktivität

RNA-Interferenz (kurz RNAi) ist ein zellulärer Prozess in dem kurze regulatorische RNA-Moleküle mit der Translation von mRNAs interferieren und so die Expression von Genen beeinflussen, meist unterdrücken.

Die ersten Beispiele für RNAi wurden erst in den 1990er Jahren beobachtet. Seitdem hat sich die Forschung zur RNAi aber explosionsartig entwickelt. Inzwischen ist klar, dass RNAi ein ganz zentraler regulatorischer Mechanismus ist, der praktisch alle biologischen Prozesse beeinflusst.

Durch diese Forschung ist auch klargeworden, dass RNA-Interferenz ein sehr komplexer und mit anderen regulatorischen Pathways eng verknüpfter Prozess ist. So beruht RNA-Interferenz nicht auf einem einzelnen, linearen Pathway, sondern auf einer Mehrzahl teils unabhängiger, teils überlappender Pathways, die jeweils wieder mit andern regulatorischen Pathways verknüpft sind.

Grundsätzlich kann man aber drei RNAi-Pathways unterscheiden, die jeweils auf einer anderen Art regulatorischer RNA beruhen: miRNA (*micro*), siRNA (*small inhibitory*) und piRNA (*piwi-interacting*). Von diesen drei versteht man den piRNA-basierten Pathway noch am wenigsten, dieser hebt sich zudem auch mechanistisch deutlich von den beiden anderen ab und wird hier deshalb nicht weiter besprochen.

Die beiden anderen siRNA- und miRNA-basierten Pathways teilen viele mechanistische Aspekte, unterscheiden sich aber voneinander in drei grundlegenden Punkten:

- *Ursprung der regulatorischen RNA Sequenzen:*  
miRNAs basieren auf endogenen (also aus den Zellen selber stammenden) nicht protein-kodierenden Sequenzen, die von der Zelle in speziellen Genen kodiert sind. siRNAs sind hingegen exogenen Ursprungs, stammen also z.B. aus viralen Genomen oder aus Transposons.
- *Komplementarität zwischen regulatorischer RNA und target RNA:*  
siRNAs haben in der Regel eine perfekte Komplementarität zu den von ihnen getargeteten RNA-Sequenzen, wohingegen miRNAs oft nur eine recht begrenzte Komplementarität zu ihren Zielsequenzen besitzen.
- *Zelluläre Funktion:*  
siRNAs sind Teil eines Verteidigungsmechanismus gegen fremde RNA und dafür verantwortlich, dass eine ganz bestimmte Zielsequenz ausgeschaltet wird. miRNAs sind hingegen an vielen unterschiedlichen biologischen Prozessen beteiligt. Eine miRNA moduliert dabei oft die Aktivität einer grossen Anzahl von Genen.

### miRNA-basierte RNA-Interferenz: eine zusätzliche Kontrollebene der eukaryotischen Genexpression

Neben den klassischen Transkriptionsregulatoren wurde die miRNA-basierte RNA-Interferenz (RNAi) inzwischen als ein ganz zentraler Mechanismus zur Regulation der eukaryotischen Genexpression anerkannt. Diese miRNAs entfalten ihre Wirkung vorwiegend auf post-transkriptioneller Ebene. miRNAs sind kleine (ca. 22 Nukleotide lange), einsträngige, nicht-kodierende RNAs, die zuerst im Fadenwurm *Caenorhabditis elegans* entdeckt wurden. Sie werden in den Genomen von fast allen Eukaryonten und einigen Viren kodiert. Diese miRNAs kontrollieren die

Expression von Zielgenen, indem sie die entsprechende mRNA spezifisch inhibieren. Diese hemmende Wirkung kann entweder durch Transkript-Destabilisierung, translationale Hemmung oder beides erfolgen. Bioinformatische Vorhersagen und experimentelle Ansätze zeigen, dass eine miRNA mehr als hundert verschiedene mRNAs inhibieren kann. Tatsächlich lassen Studien vermuten, dass die Aktivität von mehr als 60% aller protein-kodierender Gene im Menschen durch miRNAs gesteuert wird. Mittlerweile konnte gezeigt werden, dass miRNAs an der Steuerung einer Vielzahl von biologischen Funktionen beteiligt sind, wie z.B. Zellwachstum, Proliferation, Entwicklung, Differenzierung, Metabolismus und Apoptose.

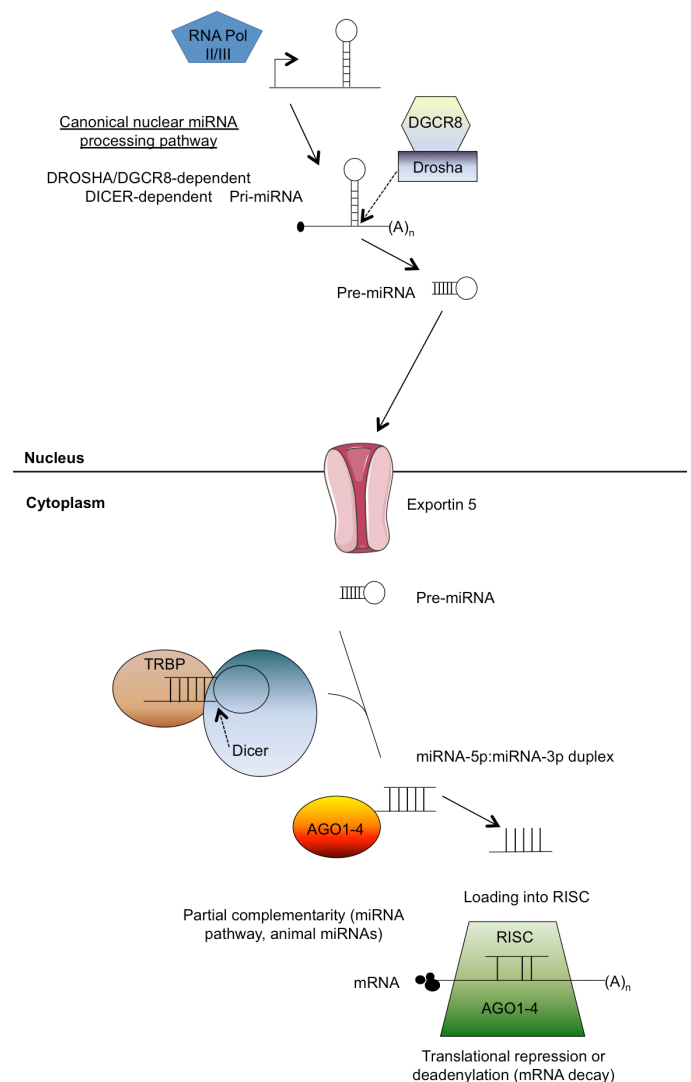
## Die molekularen Mechanismen der miRNA-basierten RNA-Interferenz

Für miRNA kodierenden Gene finden sich häufig, jedoch nicht ausschliesslich, in Intronsequenzen und zu meist in nächster Nachbarschaft zu den Genen die sie regulieren. Die DNA-Sequenzen für miRNAs erstrecken sich teilweise über mehrere Kilobasen. Diese Sequenzen werden zunächst in ein grosses Primärtranskript, die sogenannte Pri-miRNA, transkribiert. Ähnlich wie mRNAs werden diese Pri-miRNA prozessiert: Es wird eine Polyadenylinkette angehängt, das 5'-Cap wird angeknüpft und Introns werden durch Splicing entfernt.

Die eigentlichen miRNA-Sequenzen enthalten *inverted repeats*, die miteinander hybridisieren und so Haarnadelstrukturen formen (Abbildung 1). Diese Haarnadelstrukturen werden von einem grossen Proteinkomplex, dem Microprocessor, erkannt und gebunden. Dieser Microprocessor setzt sich hauptsächlich aus zwei Proteinen zusammen: Drosha und DGCR8. Das Protein Drosha schneidet nun den miRNA-Stem-Loop aus der Pri-miRNA heraus, sodass die ca. 60-80 Nukleotide lange Sequenz der pre-miRNA isoliert wird.

Die pre-miRNA wird anschliessend über Exportin 5 ins Cytoplasma exportiert. Das Protein Dicer, welches doppelsträngige RNA-Moleküle erkennt, wird nun aktiv. Es bindet die pre-miRNA und schneidet sie so, dass die doppelsträngige miRNA-Sequenz (miRNA:miRNA\*-Duplex) entsteht. Von dem miRNA-Duplex wird nun die nicht benötigte Sequenz abgetrennt, sodass die einzelsträngige miRNA (reife miRNA) in der Zelle vorliegt und an den RISC-Komplex (RNA-induced silencing complex) binden kann.

Der entstandene RISC-miRNA-Komplex reguliert nun die Expression der mRNA über zwei verschiedene Mechanismen: Ist die miRNA exakt komplementär zur mRNA, so bindet der RISC-Komplex die mRNA und schneidet sie in zwei Stücke. Die zertrennte mRNA wird dann von der Zelle als defekt erkannt und abgebaut.



**Abbildung 1** Kanonischer Pathway der miRNA-basierten RNA-Interferenz.

Sind die miRNA und die mRNA nicht exakt komplementär, so bindet der RISC-Komplex. Dieser zertrennt die mRNA aber nicht, sondern bleibt an der mRNA gebunden und reduziert dadurch die Effizienz mit der diese mRNA translatiert wird (*translational block*). Insbesondere in Säugetierzellen wird diese zweite Art der Regulierung weitaus häufiger beobachtet als die Zertrennung und der Abbau der getargeteten mRNA. Dadurch, dass miRNAs nicht exakt binden müssen um wirken zu können, wird auch deutlich, wie viele verschiedene Gene durch eine einzige miRNA-Sequenz reguliert werden können: Man geht derzeit davon aus, dass eine einzelne miRNA-Sequenz die Translation von bis zu mehreren Hundert mRNAs beeinflussen kann.

## siRNA-basierte RNA-Interferenz: Vom Verteidigungsmechanismus gegen fremde RNA zum molekularen Werkzeug der Genetik

Wie oben beschrieben fungiert miRNA-basierte RNA-Interferenz als ganz allgemeiner Mechanismus für die Regulation der Genexpression und die Zelle produziert dabei die regulatorischen RNAs selbst.

Seinen molekularen Ursprung hat der RNA-Interferenz-Mechanismus aber anscheinend in einem zellulären Verteidigungsmechanismus gegen fremde RNA-Sequenzen. Dieser Mechanismus ist vor allem in Pflanzen, die ja kein klassisches Immunsystem besitzen, weiterhin ein besonders wichtiger Verteidigungsmechanismus. Die entsprechenden molekularen Pathways sind aber auch in vielen höheren Eukaryoten (z.B. dem Menschen) weiterhin aktiv.

Wie am Anfang erwähnt, wird dieser Verteidigungsmechanismus, der nicht auf die Modulation, sondern auf die komplette Zerstörung einer RNA angelegt ist, durch siRNAs reguliert. Anders als die miRNAs sind die siRNAs stets perfekt komplementär zu der getargeteten mRNA.

### Gen-Knockdown-Experimente

Nach der Entdeckung des siRNA-Mechanismus wurde dessen Potential für zellbiologische Experimente schnell erkannt. Durch die Herstellung entsprechender siRNA-Sequenzen und deren Einschleusung in Zellen sollte man den Abbau einer jeden beliebige mRNA auslösen, und somit die Aktivität eines jeden beliebigen Gens ausschalten können.

Man bezeichnet solche Experimente in denen die von einem Gen transkribierte RNA inaktiviert wird gemeinhin als Gen-Knockdown-Experiment und unterscheidet sie damit von Gen-Knockout-Experimenten, in denen ein Gen dadurch ausgeschaltet wird, dass die DNA-Sequenz des Gens selbst inaktiviert wird.

Für Gen-Knockdown-Experimente existieren im Prinzip zwei Möglichkeiten (Abbildung 2). Entweder kann man die fertige doppelsträngige siRNA chemisch herstellen und direkt in die Zelle einschleusen oder man benutzt einen biologischen Vektor (z.B. einen Plasmid) mit dem man die siRNA-Sequenz durch entsprechende molekularbiologische Methoden auf DNA-Ebene einführt und dann die Transkriptionsmaschinerie der Zelle nutzt um diese DNA in RNA zu übersetzen.

Im letzteren Fall verwendet man dabei eine DNA-Sequenz (Abbildung 3), welche die siRNA-Sequenz als *inverted repeats* beinhaltet, die durch einen kurzen Linker verknüpft sind. Die so transkribierte RNA formt dann eine shRNA (für *short hairpin*), die durch das Dicer-Enzym in eine doppelsträngige siRNA umgewandelt wird.

Während die auf chemisch synthetisierter RNA beruhende Methode immer nur einen vorübergehenden Knockdown verursacht (die siRNA wird mit der Zeit abgebaut), kann ein mit einem Vektor eingeführtes shRNA-Gen permanent in einer Zelle etabliert werden, wodurch kontinuierlich neue siRNA produziert wird und das Gen so permanent ausgeschaltet wird.

### Design einer optimalen siRNA

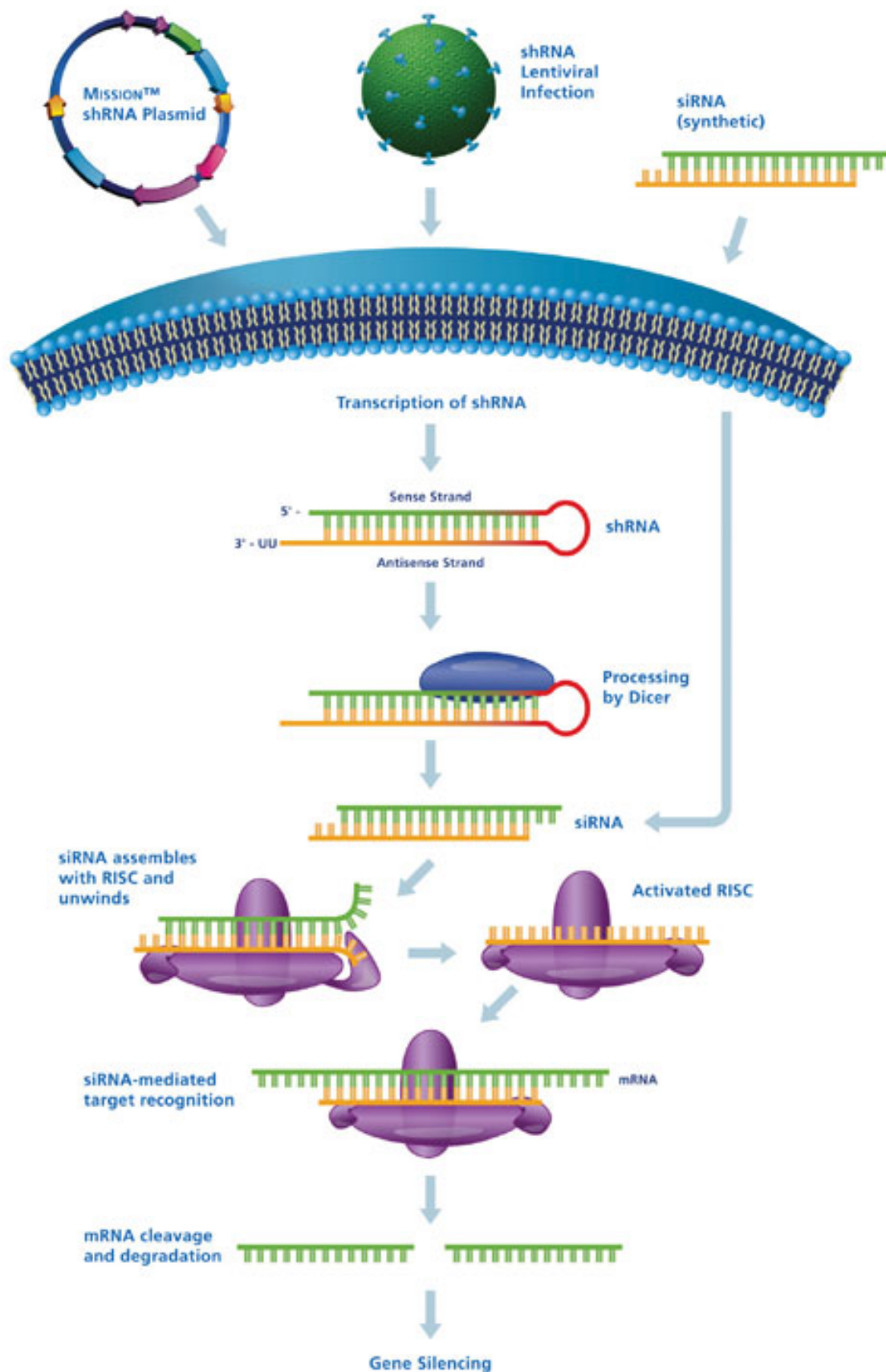
Die getargeteten mRNAs sind wesentlich länger als die siRNAs. Es stellt sich also die Frage welche Sequenz innerhalb dieser mRNA-Sequenz für die siRNA verwendet werden sollte. Durch die Erfahrungen aus vielen tausenden siRNA-Experimenten haben sich einige Faustregeln für das Design von siRNA-Sequenzen ergeben. Die 21 bp langen Sequenzen sollten z.B. in den ersten 50-100 Basen der mRNA liegen, mit zwei A's beginnen, einen GC-Gehalt von ca. 50% haben und keine Möglichkeit für interne Hybridisierung besitzen. Letzteres würde mit der Ausbildung der doppelsträngigen Struktur der siRNA interferieren.

Die Suche nach siRNA-Sequenzen, die diese Kriterien erfüllen, ist inzwischen in entsprechenden Bioinformatik-Tools automatisiert worden.

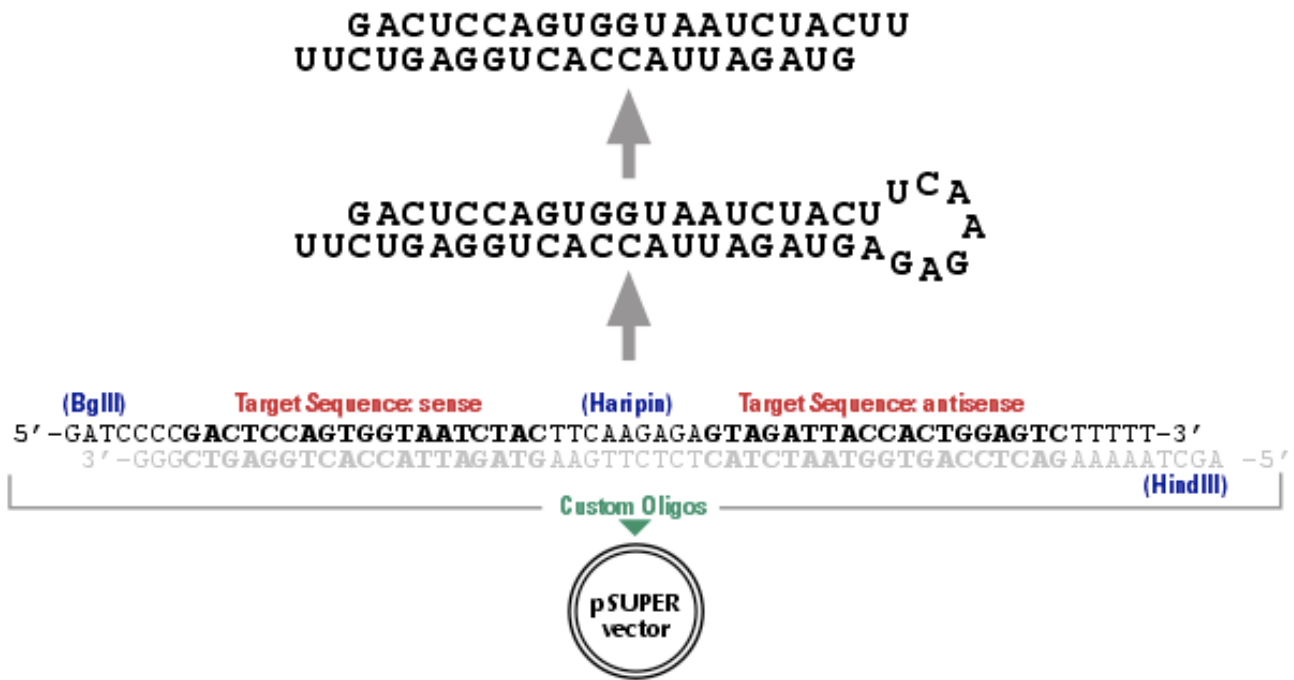
Für biomedizinisch wichtige Organismen (z.B. Mensch und Maus) werden selbst diese Softwaretools nicht mehr benötigt, denn dort sind inzwischen kommerzielle siRNA-Libraries (sowohl in RNA- als auch in Vektorform) erhältlich mit denen die Expression eines jeden beliebigen Gens unterdrückt werden kann.

### Risiko für off-target Effekte & Kontrollexperimente

Wie wir im Abschnitt über die miRNAs gesehen haben, können kurze doppelsträngige RNAs auch dann einen Effekt auf die Translation eines Gens haben, wenn nur ein partieller Match zwischen der regulatorischen RNA und der regulierten mRNA besteht. Wenn wir also eine siRNA verwenden, um eine bestimmte mRNA auszuschalten, besteht immer das Risiko, dass diese siRNA-Sequenz einen partiellen Match mit einer anderen mRNA ergibt und so deren Translation über einen miRNA-ähnlichen Mechanismus beeinflusst. Ein phänotypischer Effekt, der in einem siRNA-Experiment beobachtet wird, könnte also auch durch miRNA-artige Effekte dieser RNAs verursacht werden welche die Expression ganz anderer mRNAs beeinflusst.



**Abbildung 2 Methoden zur Einschleusung von siRNA-Sequenzen für Gen-Knockdown-Experimente.** Die siRNA-Sequenzen können entweder direkt als synthetisch hergestellte siRNAs in die Zelle eingeführt werden. Alternativ kann die, das siRNA Vorgängermolekül (shRNA) enthaltende RNA durch einen Lentivirus injiziert werden. Ausserdem besteht die Möglichkeit die siRNA-Sequenzen als DNA-Plasmid in die Zelle einzuschleusen. Dort wird diese Sequenz von der Zelle zunächst als shRNA transkribiert und dann zur siRNA prozessiert. In allen drei Fällen ist das Endresultat der durch die siRNA gesteuerte Abbau der getargeteten mRNA.



**Abbildung 3** Doppelsträngige siRNAs können in der Zelle in situ erzeugt werden. Dazu wird eine der siRNA entsprechende DNA-Sequenz, bestehend aus der Zielsequenz, einem Hairpin-Linker und der „antisense target“-Sequenz, in einem Vektor hinter einen Promoter inseriert. In der Zelle wird diese Sequenz als shRNA exprimiert. Nach Abspaltung des hairpin loops durch den Dicer-Komplex entsteht die aktive doppelsträngige siRNA. (Abbildungsquelle Oligo Engine)

Um solche *off-target* Effekte zu vermeiden bzw. von den wirklichen siRNA-basierten Effekten zu unterscheiden, verwendet man hauptsächlich zwei Strategien. Zum einen benutzt man bioinformatische Methoden (siehe oben) um siRNA-Sequenzen so auszuwählen, dass die erwarteten *off-target* Effekte minimal sind.

Zum anderen ist man dazu übergegangen, für Knockdown-Experiment nicht nur eine, sondern mindestens 3 siRNAs zu verwenden, die zwar von der Sequenz her unabhängig, aber auf dieselbe mRNA gerichtet sind. Zeigen die Experimente mit allen drei siRNAs denselben Phänotyp, so ist es unwahrscheinlich, dass dieser Phänotyp durch *off-target* Effekte zu erklären ist.

In einem weiteren Kontrollexperiment verwendet man eine siRNA für den Knockdown eines anderen Gens. Dieses Experiment stellt sicher, dass der beobachtete Phänotyp nicht durch die Injektion einer beliebigen siRNA ausgelöst wird (z.B. durch eine nichtspezifische Immunreaktion).

Darüberhinaus ist es üblich, die Konzentration der getargeteten mRNA zu messen und so den Erfolg des Knockdowns zu bestätigen.

## Das CRISPR/Cas9-System

### Biologische Ursprünge des CRISPR/Cas9-Systems

Bakterielle Zellen besitzen Mechanismen, die es ihnen erlauben fremde DNA in der Form von Plasmiden aufzunehmen und zu nutzen. Dies kann für das Bakterium einen grossen evolutionären Vorteil bedeuten. Z.B. können Bakterien durch die Aufnahme von Plasmiden schnell Antibiotikaresistenzen erwerben.

Andererseits kann die Aufnahme von Plasmiden auch gefährlich sein. So können z.B. aufgenommene Plasmide besonders aggressive Transposonelemente enthalten, die dann vielleicht das bakterielle Genom zerstören.

Zudem sind Bakterien Phagen ausgesetzt, die ihre DNA in die Bakterien injizieren. Diese Phagen-DNA nutzt dann die bakterielle DNA- und Proteinsynthesefunktionen um so viele Kopien des Phagen zu produzieren, dass die bakterielle Zelle am Ende platzt.

Die Fähigkeit sich, ähnlich dem menschlichen Immunsystem, an vorausgegangene „Infektionen“ zu „erinnern“ und sich gegen solche „aggressiven“ DNA-Sequenzen zu verteidigen, sollte also für Bakterien einen entscheidenden evolutionären Vorteil bedeuten.



Das CRISPR/Cas-System (Abbildung 4) erfüllt genau diese Funktion eines molekularen Gedächtnisses mit dem sich bakterielle Zellen an gefährliche DNA-Sequenzen erinnern und diese zerstören können. Wie aber funktioniert dieses bakterielle „Immunsystem“ für schädliche DNA-Sequenzen auf der molekularen Ebene?

## Der molekulare Mechanismus des CRISPR/Cas-Systems

CRISPR steht für *Clustered Regularly Interspaced Short Palindromic Repeats* und beschreibt Bereiche im bakteriellen Genom, in denen in regelmässigen Abständen eine bestimmte (oft palindromische) Repeat-Sequenz (von engl. *repeat* = wiederholen) wiederholt wird. Zwischen diesen Repeat-Sequenzen liegen kurze Spacer-Sequenzen (von engl. *spacer* = Abstandhalter). Diese Spacer-Sequenzen stammen offenbar von invasiven Phagen- bzw. Plasmidsequenzen ab und es sind genau diese Sequenzen gegen die das CRISPR/Cas-System die bakterielle Zelle verteidigt.

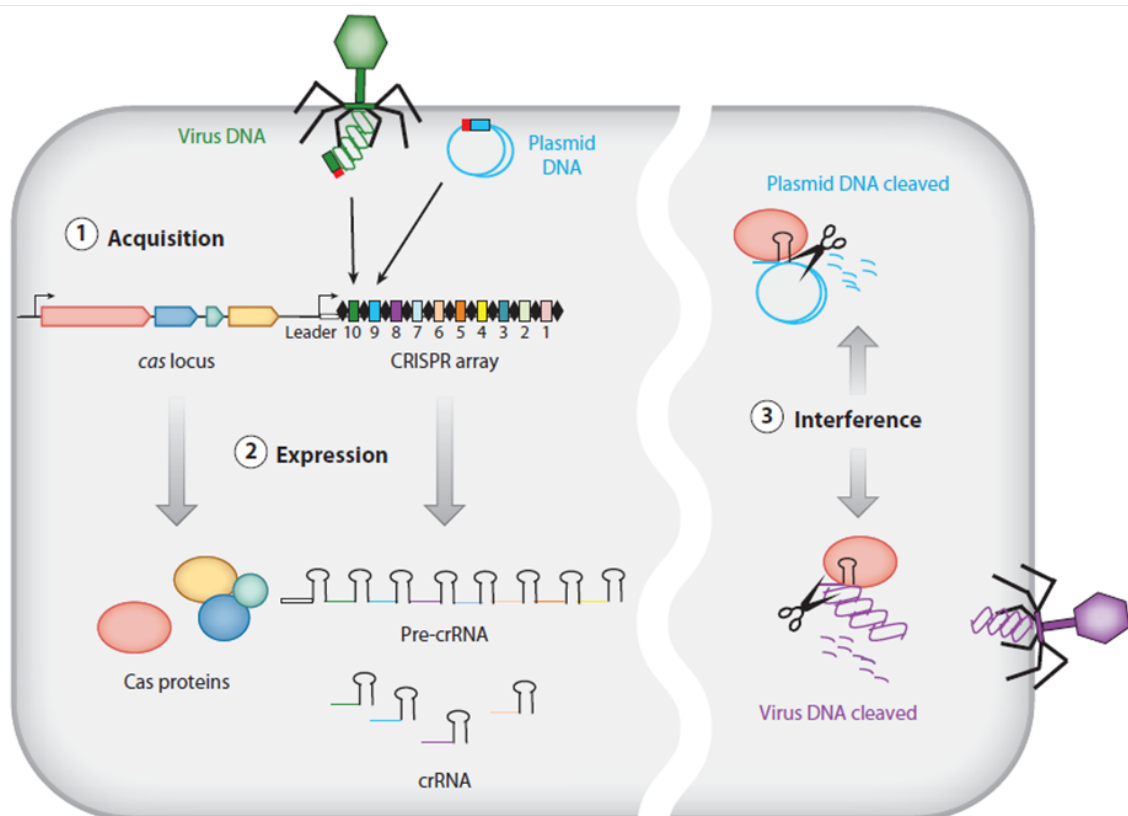
Man geht davon aus, dass diese Spacer-Sequenzen während einer vorhergehenden Infektion aus invasiven DNAs herausgeschnitten und in die CRISPR-Sequenz aufge-

nommen wurden. Diese Sichtweise wird durch Experimente unterstützt in denen kurze DNA-Sequenzen aus einem Phagengenom als künstliche Spacer-Sequenzen in eine bakterielle CRISPR-Sequenz integriert wurden. Die so modifizierten Bakterien entwickelten dann eine „Immunität“ gegen die entsprechenden Phagen.

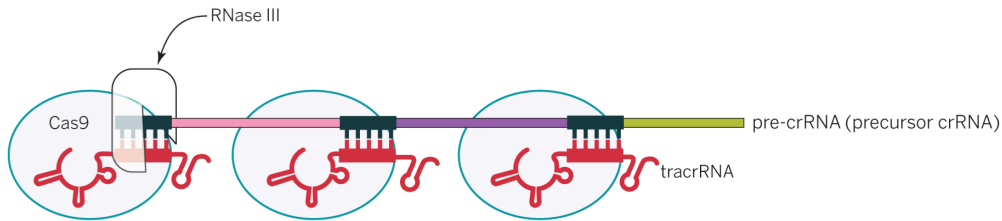
Bisher sind drei unterschiedliche Klassen von CRISPR-Systemen bekannt und die mechanistischen Details dieser drei Klassen unterscheiden sich recht deutlich voneinander. Exemplarisch ist hier der Mechanismus des CRISPR/Cas-Systems des Bakteriums *Streptococcus pyogenes* im Detail beschrieben. Dieses CRISPR/Cas-System ist von besonderer Bedeutung, weil die meisten biotechnologischen Anwendungen von CRISPR/Cas-Systemen auf ihm basieren (siehe weiter unten).

Der CRISPR/Cas-Verteidigungsmechanismus beginnt damit, dass der gesamte CRISPR-DNA-Abschnitt mit seinen alternierenden Spacer- und Repeat-Sequenzen zu einer langen pre-crRNA transkribiert wird (Abbildung 5).

Diese lange RNA wird dann später in die einzelnen crRNA-Abschnitte zerteilt, die jeweils aus einem Repeat- und einem Spacer-Abschnitt bestehen. Die Repeat-Sequenzen der pre-crRNA hybridisieren jeweils mit einer weiteren RNA, der sogenannten tracrRNA (für trans-



**Abbildung 4** Das CRISPR/Cas-System ermöglicht bakteriellen Zellen die Erkennung und Zerstörung schädlicher, fremder DNA aus Phagen und Plasmiden. Dazu werden Sequenzabschnitte von besonders invasiven DNA-Sequenzen in einem speziellen Abschnitt des bakteriellen Genoms gespeichert und dann als RNA exprimiert. Diese RNA-Sequenzen steuern dann die Spezifität von RNA-Protein-Komplexen welche die invasiven DNA-Sequenzen erkennen und zerschneiden.



**Abbildung 5 Aufbau des aktiven CRISPR/Cas-Protein-RNA-Komplexes.** Die gesamte CRISPR-Sequenz wird als eine lange RNA (pre-crRNA) transkribiert. In dieser pre-crRNA wechseln sich Repeat-Sequenzen (schwarz) mit Spacer-Sequenzen (rosa, violett und grün) ab. Diese Spacer-Sequenzen stammen von DNA-Sequenzen aus Phagen und invasiven Plasmiden ab. Die Repeat-Sequenzen besitzen hingegen eine für die jeweilige Bakterienart charakteristische Basenfolge. Diese Repeat-Sequenzen paaren sich mit kurzen von der Bakterienzelle produzierten tracrRNAs und werden von einem Proteinkomplex gebunden. Die so entstandenen ternären RNA-Protein-Komplexe werden durch Durchtrennung der pre-crRNA voneinander getrennt.

crRNA). Die so entstandene hetero-dimerische RNA wird dann von einem grossen multi-domänen Protein, dem Cas9-Protein (Cas steht für CRISPR-associated) gebunden. Dieses Cas-Protein enthält neben den RNA-Bindungsfunktionen auch zwei Proteindomänen mit Nukleasefunktion.

Der so entstandene CAS-9-Komplex vereint in sich also:

- Die Spacer-RNA-Sequenz, die beliebig veränderbar ist und die durch ihre Sequenz die anvisierte fremde DNA (target DNA) bestimmt.
- Eine RNA-Komponente (bestehend aus den tracr- und Repeat-Sequenzen), welche die Bindung der Spacer-RNA-Sequenz an die Proteinkomponente gewährleistet.
- Eine Proteinkomponente mit zwei Nukleasedomänen, welche die von dem Komplex erkannte DNA-Sequenz durchtrennen.

Dieser Komplex ist nun in der Lage spezifisch solche doppelsträngige DNA-Moleküle zu binden, die in ihrer Basenabfolge der Spacer-Sequenz entsprechen. Der Erkennungsmechanismus beruht dabei auf der Aufschmelzung des target DNA-Doppelstrangs und der Ausbildung von Watson-Crick-Basenpaaren zwischen der Spacer-RNA

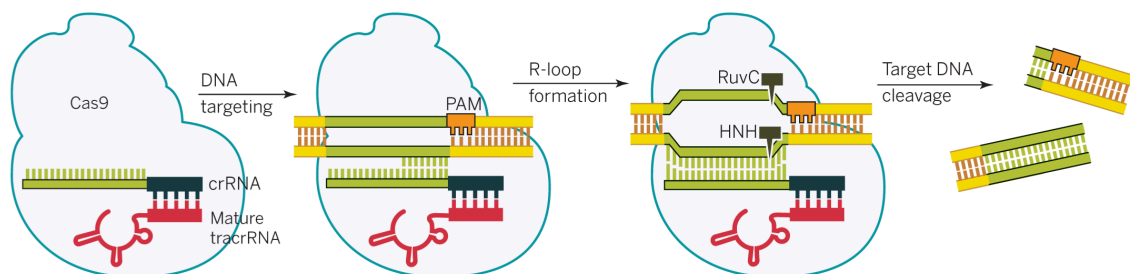
und dem komplementären Strang der target DNA (Abbildung 6).

Befindet sich direkt neben der so erkannten target DNA zusätzlich noch eine bestimmte kurze, sogenannte PAM-Sequenz (im Fall des CAS-9-Komplexes von *S. pyogenes* ist diese Sequenz NGG), dann durchtrennen die Nukleasedomänen des CAS-9-Proteins die target DNA und erzeugen so einen Doppelstrangbruch.

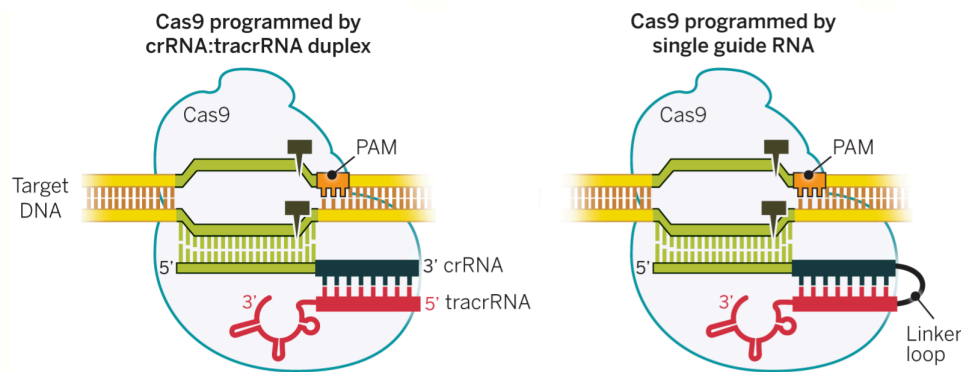
Taucht also in einer bakteriellen Zelle eine Virus- oder Plasmid-DNA auf, die im CRISPR-System durch eine entsprechende Spacer-Sequenz als schädlich „registriert“ ist, so wird diese DNA von dem entsprechenden CAS-Komplex erkannt und zerstört bevor sie Schaden anrichten kann.

## Vom bakteriellen Selbstverteidigungsmechanismus zum Werkzeug des genome engineering

Wie bereits erwähnt sind inzwischen mehrere unterschiedliche CRISPR-Systeme bekannt. Obwohl die grundlegende Funktionsweise dieser CRISPR-Systeme sehr ähnlich ist, unterscheiden sich ihre molekularen Me-



**Abbildung 6 Erkennung und Zertrennung einer target DNA durch den CAS-Komplex.** Der CAS-Komplex erkennt die von ihm „anvisierte“ target DNA durch klassische Watson-Crick-Basenpaarung. Dabei wird die DNA-Doppelhelix der target DNA vom CAS-Komplex aufgeschmolzen und es bildet sich zwischen der Spacer-Sequenz der crRNA und der target DNA ein RNA-DNA-Doppelstrang. Befindet sich neben der so erkannten DNA-Sequenz zudem noch eine sehr kurze sogenannte PAM-Sequenz (von engl. *protospacer adjacent motif*) so werden die beiden Stränge der target DNA zertrennt. Diese PAM-Sequenz ist trotz ihrer Länge von nur 3 Nucleotiden sowohl für die erste Anbindung des Cas9-Komplexes an die DNA als auch für die Zertrennung der DNA absolut essentiell.



**Abbildung 7** Die zwei RNA-Moleküle (crRNA und tracrRNA) Cas9-Komplexes von *Streptococcus pyogenes* können zu einer sgRNA (für engl. single guide RNA) fusioniert werden. Für einen funktionsfähigen Cas9-Komplex ist nun nur noch die Expression von zwei Molekülen notwendig.

chanismen zum Teil deutlich. Vor allem variiert die Komplexität der makromolekularen Komplexe, welche die fremde DNA-erkennen und zerschneiden.

Das CRISPR/Cas-System des Bakteriums *Streptococcus pyogenes* sticht dabei durch seine relative Einfachheit hervor. Der makromolekulare Komplex der in *Streptococcus pyogenes* die fremde DNA erkennt und zerschneidet besteht aus nur einem einzigen Protein (dem Cas9-Protein) sowie aus zwei RNA-Molekülen (crRNA und tracrRNA).

Diese Einfachheit erlaubte es aus dem natürlichen Cas9-System von *Streptococcus pyogenes* in relativ kurzer Zeit eine hocheffektive, genetisch encodierbare und sequenz-spezifische Nuklease zu entwickeln.

Einer der ersten Durchbrüche in diese Richtung war die erfolgreiche Verknüpfung der crRNA und tracrRNA in eine kontinuierliche sgRNA (für engl. single guide RNA) (Abbildung 7).

Es hat sich schnell gezeigt, dass sich der aus dem Cas9-Protein und der sgRNA bestehende Komplex relativ gut in einer Vielzahl unterschiedlicher Zelllinien und Organismen exprimieren lässt und in diesen Zellen auch aktiv ist. Durch kontinuierliche Optimierung der Cas9- und sgRNA-Sequenzen wurde diese Effizienz noch weiter erhöht.

Inzwischen ist das Cas9-System so effizient, dass die Co-Transformation einer Zelllinie mit einem Cas9- und einem sgRNA-Expressionsvektor in >80% der transformierten Zellen Doppelstrangbrüche verursacht. Diese Brüche entstehen jeweils genau an der Stelle, die durch die Spacer-Sequenz der sgRNA spezifiziert wurde.

Glücklicherweise besitzen eukaryotische Zellen aber sehr effiziente Mechanismen für die Reparatur von Doppelstrangbrüchen (Abbildung 8). Ansonsten wären solche Doppelstrangbrüche letal. Diese Reparaturmechanismen stellen in den meisten Fällen aber nicht wieder exakt die ursprüngliche Sequenz her, sondern resultieren in der

Einfügung von mehr oder weniger langen Sequenzabschnitten.

## Praktische Durchführung eines CRISPR/Cas9-Geninaktivierungsexperiments

Da das *targeting* der anvisierten DNA-Sequenz auf direkter Watson-Crick-Basenpaarung mit der Spacer-Sequenz der sgRNA basiert, ist es prinzipiell denkbar einfach das CRISPR/Cas9-System auf eine bestimmte Stelle im Genom zu programmieren. Man sucht dazu im protein-codierenden Bereich des auszuschaltenden Gens nach einer PAM-Sequenz (5'-NGG-3') von *S. pyogenes*. Aufgrund ihrer Kürze ist diese Sequenz recht oft zu finden. Die 20 vorhergehenden (d.h. 5'-gelegenen) Nucleotide der Gensequenz werden dann als Spacer-Sequenz in das Gen einer sgRNA kloniert, das sich bereits zusammen mit einem entsprechenden Promoter in einem Expressionsvektor befindet. Dieser sgRNA-Expressionsvektor wird nun zusammen mit einem zweiten, für die Expression des Cas9-Proteins verantwortlichen Vektor in die gewünschte Zelllinie eingeschleust.

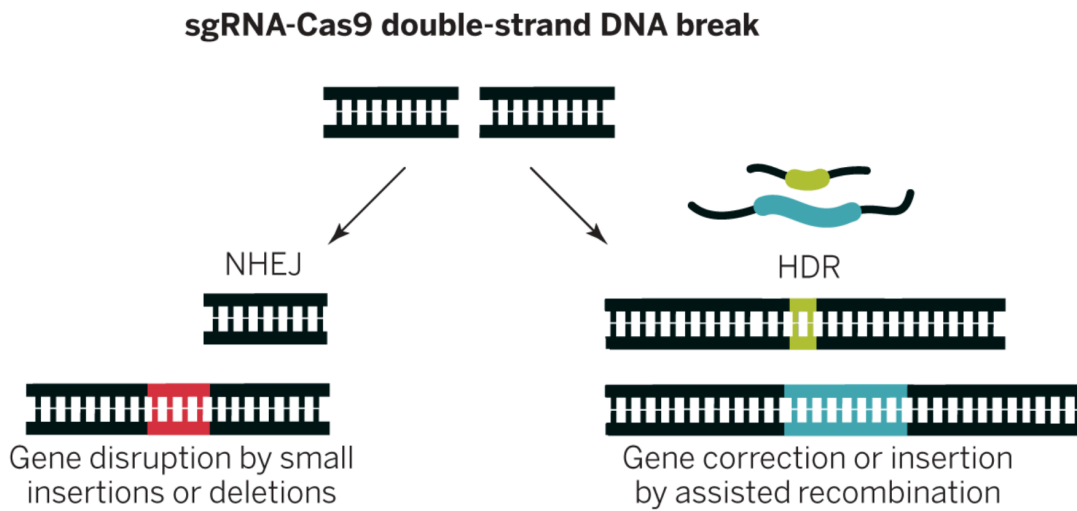
Durch die hohe Effizienz des CRISPR/Cas9-Systems werden in den so transfizierten Zellen zumeist beide Kopien des anvisierten Gens modifiziert. Es entstehen also in der Regel homozygote Knockouts.

Der Erfolg der Genmodifikation kann dann durch Sequenzieren des anvisierten Gens verifiziert werden.

Inzwischen wurden viele tausende CRISPR/Cas9-Geninaktivierungsexperimente durchgeführt. Aus diesen Erfahrungen hat man gelernt, dass nicht alle Spacer-Sequenzen gleich gut funktionieren oder gleich selektiv sind. Nicht-spezifische Doppelstrangbrüche stellten z.B. anfangs ein relativ schwerwiegendes Problem dar.

Ausserdem wurde auch entdeckt, dass wenn die PAM-Sequenz durch bestimmte Sequenzen flankiert wird, die Effizienz der Nukleaseaktivität weiter gesteigert werden kann.





**Abbildung 8 Reparatur von Doppelstrangbrüchen.** DNA-Doppelstrangbrüche werden von der Zelle entweder durch non-homologous end joining (NHEJ) (links) oder durch homology directed repair (HDR) (rechts) repariert. In beiden Fällen wird dabei die ursprüngliche DNA-Sequenz verändert. Verursacht man also mit einem CRISPR/Cas9-System einen Doppelstrangbruch in der protein-kodierenden Sequenz eines eukaryotischen Gens, so ist das Endresultat in der Regel eine Insertions-Mutation welche sehr oft die Funktion des codierten Proteins zerstört.

Alle diese Informationen fließen zunehmend in die Auswahl der optimalen *target* DNA-Sequenzen ein, wodurch die Effizienz und Spezifität des CRISPR/Cas9-Systems kontinuierlich verbessert wird.