

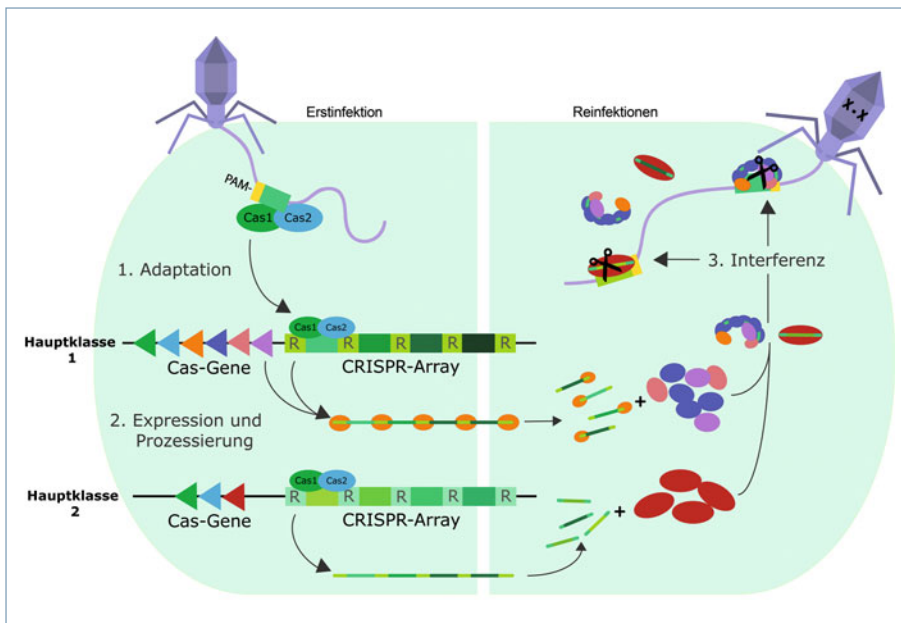
CRISPR-RNA-Biologie

Vielfältige Genscheren: natürliche Aktivitäten von CRISPR-Cas-Systemen

HANNA MÜLLER-ESPARZA, DANIEL GLEDITZSCH, LENNART RANDAU
MAX-PLANCK-INSTITUT FÜR TERRESTRICHE MIKROBIOLOGIE, MARBURG

Prokaryotic genomes often contain CRISPR-Cas systems that enable RNA-guided defense against viral attacks. Viral genomes were found to encode anti-CRISPR proteins that inactivate specific Cas protein functions. In the last years, a large variety of CRISPR-Cas system components and mechanisms has been uncovered. This diversity highlights the fact that CRISPR-Cas systems co-evolve with anti-CRISPR measures, but also indicates that they can adopt roles in the cell that extend beyond viral defense.

DOI: 10.1007/s12268-018-0978-6
© Die Autoren



▲ **Abb. 1:** Schematische Darstellung der Aktivitäten von CRISPR-Cas-Systemen. Die Aktivität von CRISPR-Cas-Systemen ist in drei Schritte unterteilt: (1) Adaptation: DNA-Fragmente von viralen Genomen (Protospacer) werden durch die Cas-Proteine Cas1 (grün) und Cas2 (blau) in CRISPR-Arrays der prokaryotischen Wirtszelle eingebaut. Dabei werden kurze Sequenzmotive in der Nähe des Protospacers (PAM-Sequenzen) erkannt, und der CRISPR-Array wächst um eine neue Einheit aus repetitiver DNA-Sequenz (*repeat*, R) und Protospacersequenz (Spacer). (2) Expression und Prozessierung: CRISPR-assoziierte (Cas-)Proteine werden produziert, und der gesamte CRISPR-Array wird transkribiert. Dieses Transkript wird durch Cas6-Proteine (orange) oder andere Wirtsfaktoren (z. B. RNase III) in einzelne CRISPR-RNAs (crRNAs) geschnitten. Jede crRNA enthält einen Spacer und terminale Fragmente der *repeat*-Sequenz. (3) Interferenz: Die crRNAs bilden Interferenzkomplexe mit einzelnen Cas-Proteinen (Klasse 2, z. B. Cas9 [rot]) oder mehreren Cas-Proteinen (Klasse 1, z. B. Cascade mit Cas6 [orange], Cas7 [blau], Cas5 [rosa] und großer Untereinheit [lila]). Die Spacersequenz der crRNA erkennt komplementäre Nukleinsäureabschnitte und bewirkt deren nukleolytischen Verdau durch Cas-Proteine.

■ Die CRISPR-Cas-Technologie erlaubt, genetisches Material einfach zu verändern. Sie hat als „Genschere“ mit bemerkenswerter Geschwindigkeit Einzug in den Laboralltag gehalten. Wird das Akronym CRISPR in jüngster Zeit auch als Verb benutzt („Ich CRISPR eine Maus“), ist damit meist die Mutation eines Gens nach gezielter DNA-Schneideaktivität durch ein Cas9-Protein aus *Streptococcus*-Stämmen gemeint. Die Erfolgsgeschichte von Cas9 ist ein deutliches Beispiel für die Bedeutung mikrobieller Grundlagenforschung. Die hohe Anzahl an CRISPR-Cas-Systemen mit vielen unterschiedlichen Cas-Proteinen in bakteriellen und archaeellen Wirten bietet ein großes Potenzial für zukünftige Entdeckungen neuer Mechanismen und für die Entwicklung von genetischen Werkzeugen.

Komponenten der CRISPR-Cas-Systeme

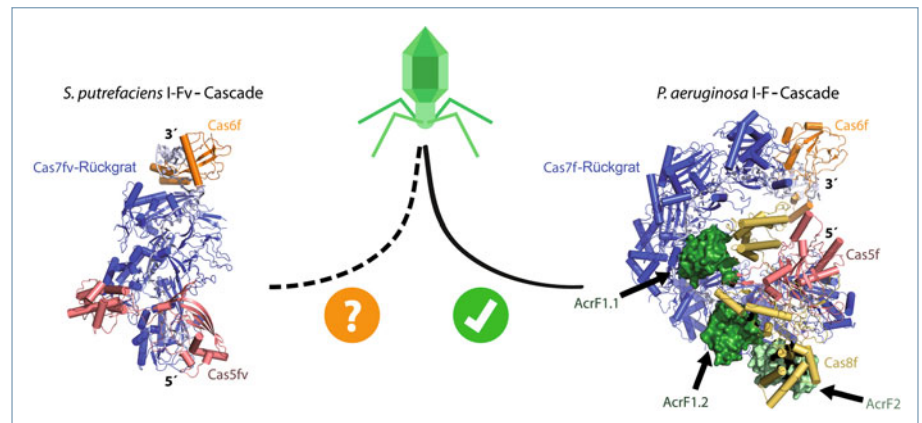
Die zwei Hauptkomponenten aller CRISPR-Cas-Systeme sind schon im Namen enthalten: CRISPR bezeichnet dabei Bereiche in prokaryotischen Genomen, die aus sich wiederholenden DNA-Abschnitten (den *repeats*) und zwischengeschalteten DNA-Bereichen mit variablen Sequenzen (den Spacern) bestehen. Diese Organisation beschreibt die englische Bezeichnung *clustered regularly interspaced short palindromic repeats* (CRISPR). Die Cas-Komponente fasst CRISPR-assoziierte Sequenzen (Cas) zusammen, die als Cas-Proteine die Aktivitäten des Immunsystems ausführen. Die CRISPR-Bereiche werden von RNA-Polymerasen abgeschrieben. Die entstehenden Transkripte prozessieren Cas-Proteine oder Wirtsenzyme zu CRISPR-RNA (crRNA)-Molekülen [1].

Jede crRNA besitzt einen einzelnen Spacer, der von *repeat*-Fragmenten flankiert wird. Die wichtige Information der CRISPR-Cas-Systeme liegt abgespeichert in den Sequenzen der Spacer. Im Jahr 2005 wurde erkannt, dass dieses „Speichersystem“ Fragmente viraler Genome und mobiler genetischer Elemente erfasst [2]. Einzelne crRNAs können den Befehl von Bak-

terien durch Viren, die deren Spacersequenz im Genom enthalten, verhindern. In den letzten Jahren wurde der Schritt des Speicherns von Spacersequenzen durch die universellen Cas-Proteine Cas1 und Cas2 aufgeklärt (**Abb. 1**, [3]). Die Ursprungsfragmente werden als Protospacer bezeichnet. Kurze DNA-Abschnitte, PAM (*protospacer adjacent motif*) genannt, dienen als Orientierungshilfen. Des Weiteren sorgen PAM-Sequenzen für die notwendige Unterscheidung von eigener und fremder DNA [4]. Das häufig vorkommende Cas4-Protein ist für die Anreicherung von Protospacern mit funktionellen PAM-Sequenzen zuständig [5]. Einzelne crRNAs bilden Komplexe mit Cas-Proteinen, um Nukleinsäuren nach Zielsequenzen absuchen zu können. Dabei erkennen diese Ribonukleoprotein-Komplexe die Zielsequenz über Basenkomplementarität zum crRNA-Spacer und fehlende Komplementarität zwischen deren *repeat*-Fragmenten und PAM-Sequenzen [6].

Zweiklassengesellschaft von CRISPR-Cas-Systemen

Die Erkennung von Ziel-DNA durch eine crRNA in einen Ribonukleoprotein-Komplex wird als Interferenz bezeichnet. In diesem Schritt unterscheiden sich die verschiedenen CRISPR-Cas-Systeme am deutlichsten. Mehrere überarbeitete Klassifikationen wurden benötigt, um gemeinsame Merkmale der Interferenzaktivitäten durch unterschiedliche Cas-Proteine beschreiben zu können (**Abb. 1**). Es werden sechs CRISPR-Cas-Typen charakterisiert, die wiederum in zwei Klassen eingeteilt sind [7]. Die Klasse 2 beinhaltet die populären Cas9-Proteine und andere Systeme, bei denen ein einzelnes großes Cas-Protein sowohl für die Erkennung der PAM-Sequenz als auch für die Entwindung und Spaltung der Ziel-DNA zuständig ist. Des Weiteren können manche dieser Cas-Proteine auch eine Rolle bei der Reifung der crRNA spielen. Die Synthese von künstlichen Spacersequenzen ermöglicht eine Programmierung von Cas-Proteinen als zielgerichtete DNA-Scheren, die in Geneditierungsapplikationen in Eukaryoten und Prokaryoten eingesetzt werden. Für diese Prozesse sind Cas-Proteine der Klasse 2 aufgrund ihrer geringen Anzahl an benötigten Komponenten die populärsten Werkzeuge. Cas9 wird dabei von Cas12 ergänzt, das andere DNA-Schnitte entstehen lässt und ohne eine zweite RNA-Komponente auskommt. Cas13 unterscheidet sich von diesen Proteinen, da es die spezifische Spaltung von RNA katalysiert und Ansätze



▲ **Abb. 2:** Struktur von Cascade-Komplexen und Bindung von Anti-CRISPR-Proteinen. Die Strukturen von Klasse-1-Interferenzkomplexen aus *Shewanella putrefaciens* (PDB-ID: 5O7H) und *Pseudomonas aeruginosa* (PDB-ID: 5ZU9) zeigen die Bindung der crRNA durch Cas-Proteine. Anti-I-F-CRISPR-Proteine (AcrF, grün) interagieren mit der großen Cascade-Untereinheit Cas8f (gelb), die bei der I-Fv-Cascade-Variante fehlt. Die Immunantwort durch I-F-Cascade kann so vom Phagen abgewehrt werden. Anti-CRISPR-Proteine gegen I-Fv-Cascade sind noch nicht bekannt.

zur Manipulation und sensitiven Detektion von RNA-Molekülen ermöglicht.

Die hohe Anwendbarkeit dieser Klasse-2-Systeme steht im Gegensatz zu ihrem relativ seltenen Vorkommen in prokaryotischen Wirtszellen. Diese setzen wesentlich öfter CRISPR-Cas-Systeme der Klasse 1 ein, die durch eine Zusammenarbeit von mehreren Cas-Proteinen mit klar definierten Rollen bei der Interferenz gekennzeichnet sind. In dieser Klasse fungieren Proteine der Cas6-Familie meist als RNA-Endonukleasen bei der crRNA-Reifung und bleiben oft am 3'-Ende der crRNA gebunden. Das 5'-Ende der crRNA wird durch Proteine der Cas5-Familie erkannt. Mehrere Kopien von Cas7-Proteinen binden am crRNA-Spacer und bilden so das helikale Rückgrat der Interferenzkomplexe (**Abb. 2**).

Ähnlichkeiten in der Struktur der RNA-Bindemotive weisen auf einen gemeinsamen Ursprung dieser drei Cas-Proteinfamilien hin. Die Organisation der Interferenzkomplexe wurde gut am Beispiel des Typ-I-E-Cascade (*CRISPR-associated complex for antiviral defense*)-Komplexes aus *Escherichia coli* untersucht, und Strukturen in unterschiedlichen Stadien der DNA-Erkennung wurden beschrieben [8]. Cascade enthält zwei zusätzliche Proteine, die als große Untereinheit bei der PAM-Bindung und als kleine Untereinheiten bei der Orientierung der DNA-Stränge beteiligt sind. In anderen Cascade-Varianten in Typ-I-C-Systemen übernimmt Cas5 die Rolle der RNA-Prozessierung; Cas6 wird nicht benötigt [9].

Das Genom von *Shewanella putrefaciens* und anderen Proteobakterien enthält eine

Variante des Typ-I-F-Systems ohne große Untereinheiten. In diesem Komplex übernimmt ein Cas5-Protein die Erkennung der PAM-Sequenz. Cas7-Proteine interagieren zusätzlich mit dem ungepaarten DNA-Strang der Ziel-DNA [10]. Die Struktur dieses minimalen Cascade-Komplexes zeigt, wie drei Cas-Proteine ausreichen, um eine crRNA-vermittelte Erkennung viraler DNA zu ermöglichen (**Abb. 2**). Dieser Prozess führt in Typ-I-E-Cascade zu einer Änderung der globalen Struktur des Komplexes, was die Rekrutierung der DNA-Schere Cas3 und den Abbau des viralen Genoms auslöst.

Andere Komplexe aus crRNA und Cas-Proteinen ähneln dem Aufbau von Cascade und werden als Typ-III-Komplexe zusammengefasst. Es war lange Zeit unklar, ob diese Komplexe RNA- oder DNA-Ziele erkennen und schneiden. Beide Aktivitäten sind miteinander vereinbar, wie kürzlich mehrere Forschungsgruppen zeigten, da sowohl DNA im Prozess der Transkription als auch die entstehenden Transkripte geschnitten werden. Dieser Mechanismus ermöglicht eine Interferenz gegen sich replizierende Viren, wohingegen lysogene Viren nicht sofort abgewehrt werden. Typ-III-Komplexe können zusätzlich unspezifische RNasen aktivieren und nutzen dazu zyklische Oligoadenylate als Signalmoleküle [11]. Die Verbreitung der Rezeptoren für diese Signalmoleküle und die Regulation dieses Signals werden zurzeit untersucht. In den letzten Jahren wurde deutlich, dass Bakterien und Archaeen CRISPR-Cas-Systeme für die RNA-gesteuerte Regulation anderer Prozesse „zweckentfremden“ können. Beispiele

sind crRNAs mit Spacern gegen das Wirtsgenom zur Genregulierung oder gegen verwandte Organismen zur Beeinflussung von horizontalem Gentransfer. Andere CRISPR-Cas-Systeme kommen ganz ohne erkennbare DNA-Scheren aus. Hierzu gehören alle Typ-IV-Systeme, deren funktionelle Rolle in den Wirtszellen noch nicht bekannt ist [12].

Virale Anti-CRISPR-Proteine und die Evolution von Anti-Anti-CRISPR-Maßnahmen

Die Koevolution von CRISPR-Cas-Immunsystemen und ihren viralen Zielscheiben trägt zur Diversifizierung der Cas-Proteine und deren Aktivitäten bei. Ähnlich lässt sich auch die große Anzahl an unterschiedlichen Restriktionsenzymen und Erkennungssequenzen erklären. Viren können Zielsequenzen und PAM-Motive mutieren oder modifizieren, sehen sich dabei in der Umwelt aber mit einer Vielzahl von heterogenen Spacern in den einzelnen Wirtszellen konfrontiert. Diese Heterogenität der viralen Zielsequenzen ermöglicht eine effiziente Immunabwehr der Bakterienpopulation. Da sich Viren so also nicht genügend durch Mutation der Protospacersequenzen wehren können, benötigen sie Mechanismen, um CRISPR-Interferenzkomplexe zu inaktivieren.

Viren produzieren dazu eine große Vielzahl an Anti-CRISPR-Proteinen, die die Aktivitäten von Cas-Proteinen modulieren können. Diese sofort nach der Infektion hergestellten Proteine imitieren z. B. PAM-Sequenzen in Größe und Oberflächenladung und binden an Interferenzkomplexe, um deren DNA-Erkennung zu unterbinden [13]. Wirtszellen benötigen also Gegenmaßnahmen gegen diese Anti-CRISPR-Proteine. Die große Cascade-Untereinheit ist eines der häufigeren Ziele von Anti-CRISPR-Proteinen. Die entdeckten Cascade-Varianten ohne große Untereinheiten bieten diesen Anti-CRISPR-Proteinen keine Bindemöglichkeiten und stellen einen selektiven Vorteil im Kampf gegen Anti-CRISPR-Proteine dar (**Abb. 2**). In diesem Wettrennen sorgen Viren so für die stete Evolution neuer CRISPR-Cas-Mechanismen, die uns ein enormes Potenzial an möglichen Entdeckungen und Anwendungen bieten.

Danksagung

Wir danken unserer Arbeitsgruppe und Kooperationspartnern. Die Arbeiten wurden gefördert von der Max-Planck-Gesellschaft und der Deutschen Forschungsgemeinschaft (DFG FOR 1680 und SPP 2141).

Literatur

- [1] Brouns SJ, Jore MM, Lundgren M et al. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–964
- [2] 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
- [3] Nunez JK, Kranzusch PJ, Noeske J et al. (2014) Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat Struct Mol Biol* 21:528–534
- [4] Westra ER, Semenova E, Datsenko KA et al. (2013) Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet* 9:e1003742
- [5] Kieper SN, Almendros C, Behler J et al. (2018) Cas4 facilitates PAM-compatible spacer selection during CRISPR adaptation. *Cell Rep* 22:3377–3384
- [6] Semenova E, Jore MM, Datsenko KA et al. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci USA* 108:10098–10103
- [7] Koonin EV, Makarova KS, Zhang F (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* 37:67–78
- [8] Xiao Y, Luo M, Hayes RP et al. (2017) Structure basis for directional R-loop formation and substrate handover mechanisms in type I CRISPR-Cas system. *Cell* 170:48–60

- [9] Nam KH, Haitjema C, Liu X et al. (2012) Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure* 20:1574–1584
- [10] Pausch P, Muller-Esparza H, Gleditzsch D et al. (2017) Structural variation of type I-F CRISPR RNA guided DNA surveillance. *Mol Cell* 67:622–632
- [11] Niewoehner O, Garcia-Doval C, Rostol JT et al. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* 548:543–548
- [12] Özcan A, Pausch P, Linden A et al. (2018) Type IV CRISPR RNA processing and effector complex formation in *Aromatoleum aromaticum*. *Nat Microbiol* (im Druck)
- [13] Pawluk A, Davidson AR, Maxwell KL (2018) Anti-CRISPR: discovery, mechanism and function. *Nat Rev Microbiol* 16:12–17

Open Access:

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Open access funding provided by Max Planck Society.

Korrespondenzadresse:

Dr. Lennart Randau
Prokaryotic Small RNA Biology
Max-Planck-Institut für terrestrische Mikrobiologie
Karl-von-Frisch-Straße 10
D-35043 Marburg
Tel.: 06421-178600
Fax: 06421-178599
lennart.randau@mpi-marburg.mpg.de

AUTOREN



Hanna Müller-Esparza

Jahrgang 1991. 2009–2015 Biochemiestudium an der Päpstlichen Katholischen Universität von Chile. Seit 2015 Promotion am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg.



Daniel Gleditzsch

Jahrgang 1990. 2010–2015 Biologiestudium an der TU Darmstadt. Seit 2016 Promotion am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg.



Lennart Randau

Jahrgang 1978. 2002–2006 Promotion am Institut für Mikrobiologie, TU Braunschweig, bei Prof. Dr. D. Jahn. 2006–2010 Postdoc an der Yale University, USA, Molecular Biophysics and Biochemistry, bei Prof. Dr. D. Söll. Seit 2010 Leiter einer unabhängigen Max-Planck-Forschungsgruppe am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg.