

## سیستم CRISPR از مواجهه با یک توالی تکراری مرموز برای تکنولوژی، تا ویرایش ژنوم

مجتبی سهرابی<sup>۱</sup>، حامد سلمانی<sup>۲</sup>، زهرا سادات منزوی<sup>۳</sup>، زهرا عبدالهی<sup>۴</sup>، سمیه دهقانی سانج<sup>۵</sup>، عباس مروتی<sup>۶\*</sup>

### چکیده

تناوب‌های کوتاه پالیندروم فاصله‌دار منظم خوشه‌ای ( *clustered regularly interspaced short palindromic repeats* ) و سیستم‌های CRISPR-Cas، سیستم‌های ایمنی به دست آمده شناخته شده‌ای است که در آرکیا و باکتری‌ها گسترده است. نوکلن‌های RNA راهنما از سیستم‌های CRISPR-Cas، در حال حاضر به عنوان ابزار قابل اعتماد برای ویرایش و مهندسی ژنوم در نظر گرفته شده است. نخستین اشاره به آن‌ها در سال ۱۹۸۷ بود؛ زمانی که در ژنوم اشرشیا کلای در طول تجزیه و تحلیل ژن‌های دخیل در متابولیسم فسفات یک توالی DNA ی تکراری غیر معمول و پس از آن به عنوان یک CRISPR تعریف شد. الگوهای توالی مشابه، پس از آن در طیف وسیعی از باکتری‌های دیگر و همچنین در آرکی باکتری‌های نمک دوست گزارش شده است، که نشان می‌دهد برای چنین خوشه‌های تکاملی حفظ شده از توالی‌های تکراری دارای نقش مهمی است. یک گام مهم به منظور مشخص کردن ویژگی سیستم CRISPR-Cas، شناسایی ارتباط بین CRISPR ها و پروتئین‌های مرتبط با Cas بود و در ابتدا فرض بر این بود که در تعمیرات DNA در آرکی‌های گرمادوست دخالت داشته باشد. زیست‌شناسی ساختاری و بیوشیمی پیشرفته می‌تواند نه تنها در ویرایش ژنوم بر پایه CRISPR-Cas<sup>۹</sup> و دیگر سیستم‌های CRISPR-Cas کلاس II دست به دست هم دهد؛ بلکه درک منشأ و تکامل این سیستم از عناصر متحرک ژنتیکی، نشان دهنده عناصر متحرک ژنتیکی<sup>۲</sup> است.

واژگام کلیدی: آرکیا، عناصر ژنتیکی متحرک، اصلاح ژنی، سکانس‌های تکراری.

۱. استادیار، ژنتیک، گروه میکروبیولوژی، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

۲. دانشجو، دانشجوی بیوتکنولوژی، گروه بیوتکنولوژی، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

۳. مربی، کارشناس ارشد، گروه میکروبیولوژی، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

۴. مربی، کارشناس ارشد، گروه میکروبیولوژی، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

۵. مربی، دانشجوی دکتری و ویروس‌شناسی، گروه میکروبیولوژی، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

۶. \*نویسنده مسئول: عباس مروتی، مربی، گروه میکروبیولوژی، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

موبایل: ۰۹۱۲۵۵۳۵۳۹۵ ایمیل: Abbasmorovvati@gmail.com

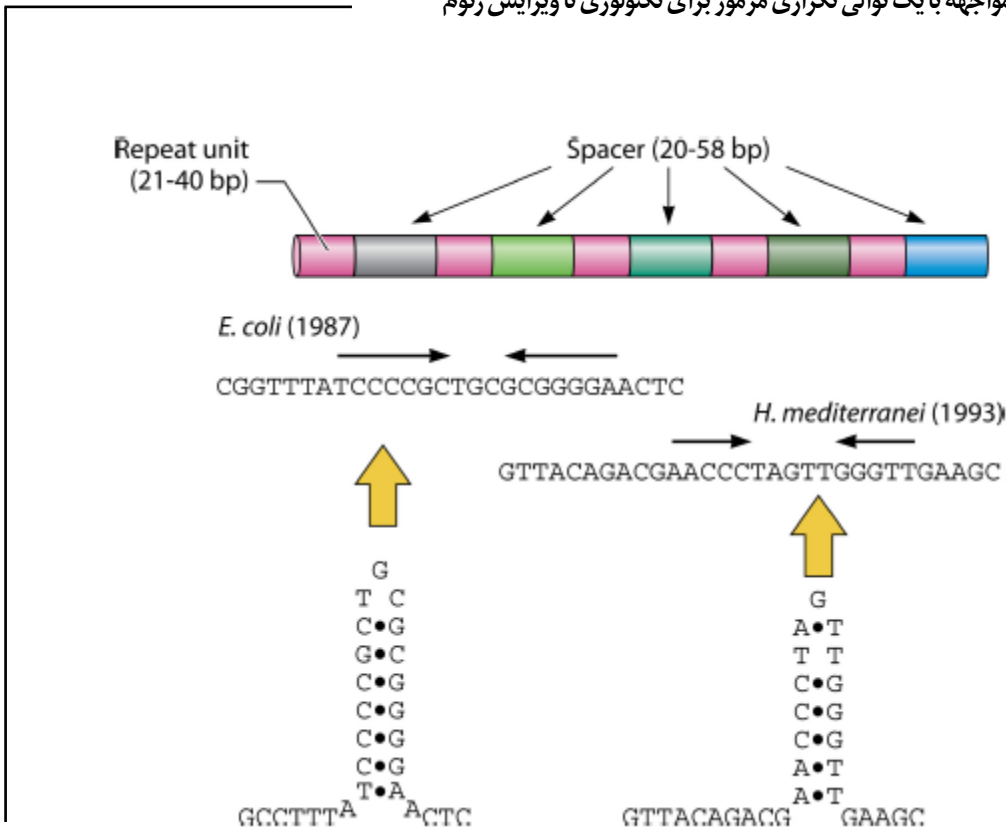
## مقدمه

تناوب‌های کوتاه پالیندروم فاصله‌دار منظم خوشه‌ای (Crispr)<sup>۱</sup> در حال حاضر، کانون توجه پژوهش‌های فعال در زیست‌شناسی هستند. اولین سیستم CRISPR سی سال پیش توسط دانشمندی به نام "ایشنو" در باکتری اشرشیا کلای، در جریان تجزیه و تحلیل ژن مسئول تبدیل ایزوآنزیم آلکالین فسفاتاز شناسایی شد (ویژگی‌های ساختاری CRISPR در تصویر شماره ۱ نشان داده شده است). در آن زمان، به سختی امکان داشت عملکرد بیولوژیکی این توالی‌های تکراری غیرمعمول را به دلیل عدم اطلاعات کافی توالی DNA، به ویژه برای عناصر متحرک ژنتیکی پیش‌بینی کرد. عملکرد واقعی این توالی منحصر به فرد، تا اواسط سال ۲۰۰۰ مبهم باقی ماند. در سال ۱۹۹۳، اولین بار سیستم CRISPR، در آرکی باکتری مشاهده (به خصوص در سویه هالوفراکس مدیترانه‌ای) و پس از آن، در تعداد بیش‌تری از ژنوم‌های باکتری و آرکی باکتری‌ها تشخیص داده شد. حفظ این توالی‌ها در دو مورد از سه دامنه زندگی، برای درک اهمیت آن‌ها بسیار مهم بود. در اوایل دهه ۲۰۰۰، کشف شباهت توالی بین مناطق جداکننده<sup>۲</sup> از CRISPR و توالی‌های باکتریوفاژ، ویروس‌ها، آرکی باکتری‌ها و پلاسمیدها در نهایت سبب شد تا عملکرد CRISPR به عنوان یک سیستم ایمنی روشن شود. این کشف حیرت‌انگیز توسط "موجیکا" و دیگران، در آن زمان به شدت ناشناخته بود (۱) و به طور مستقل، در سال ۲۰۰۵، توسط سه گروه تحقیقاتی منتشر شد (۲). به موازات، چندین ژن پیشنهاد شده برای رمزکردن پروتئین‌های ترمیم DNA، به ویژه برای آرکی‌های گرمادوست افراطی<sup>۳</sup> شدیداً به عنوان مرتبط با CRISPR شناخته شد و ژن‌های مربوط به Cas تعیین شده بودند (۳). بنابراین، تجزیه و تحلیل ژنوم مقایسه‌ای نشان داد که پروتئین‌های CRISPR و Cas (محصولات ژن Cas) در واقع با یکدیگر کار می‌کنند و یک سیستم ایمنی به دست آمده، مشابه سیستم RNAی مداخله‌گر (RNAi) یوکاریوتی را برای محافظت از سلول‌های پروکاریوتی در برابر ویروس‌های مهاجم و پلاسمیدها تشکیل می‌دهد (۴).

<sup>۱</sup> clustered regularly interspaced short palindromic repeats

<sup>۲</sup> spacer regions

<sup>۳</sup> Hyperthermophili archaea

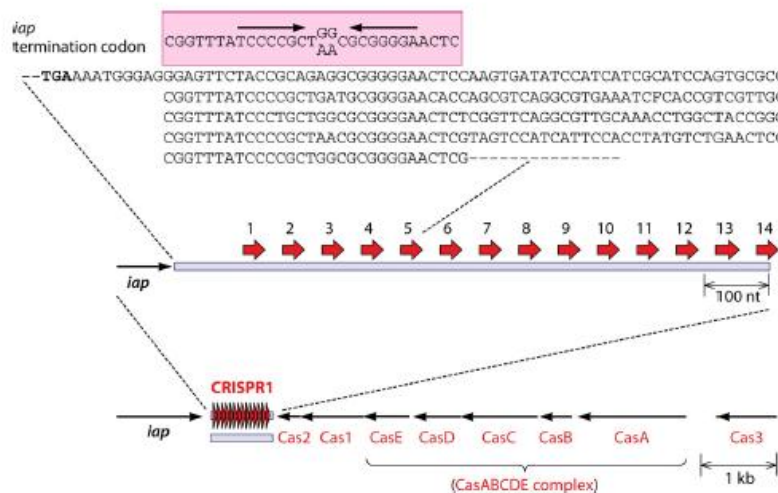


۳

تصویر شماره ۱. ویژگی‌های ساختاری CRISPR. توالی تکرار با طول ثابت، به طور کلی تقارن dyad را برای ساختار پالیندرومیک تشکیل می‌دهند (توسط فلش نشان داده شده است). دو نمونه از اولین CRISPR شناسایی شده از *E. coli* (باکتریایی) و *H. mediterranei* (آرکی باکتریایی) نشان داده شده است. مناطق فاصله‌انداز ثابت و بدون توالی همولوگ هستند.

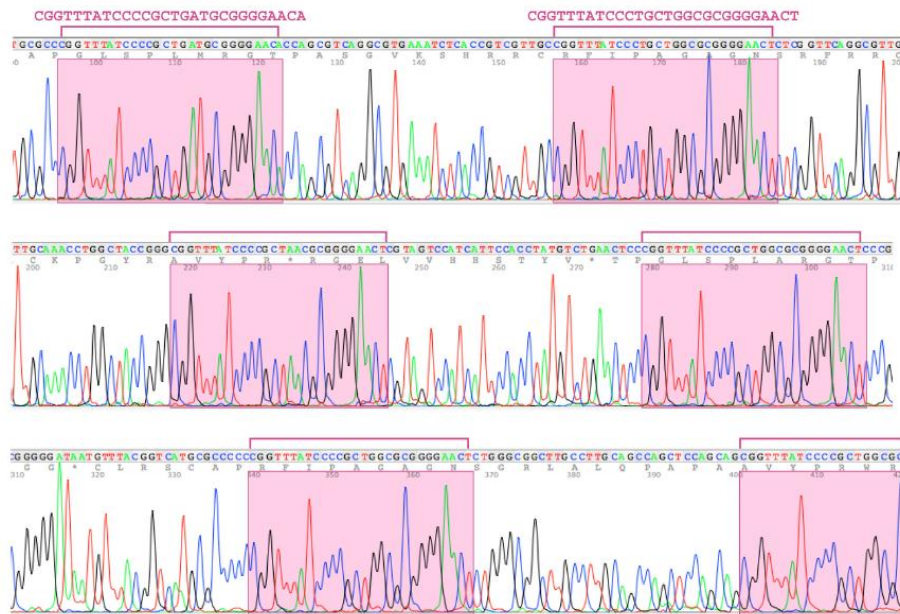
در اواسط دهه ۱۹۸۰، هنگام مطالعه تبدیل ایزوآنزیم آلکالین فسفاتاز (AP)، توسط "ایشنو"، ایزوآنزیم AP توسط آنالیزهای بیوشیمیایی و ژنتیکی تشخیص داده شد (۵-۶). در آن زمان، برای توالی متداول دی دنوکسی M۱۳، DNA الگوی تک رشته باید توسط کلونینگ DNA هدف به یک وکتور M۱۳ تولید شود؛ در حالی که واکنش انتهای زنجیره دی دنوکسی، توسط قطعه کلنو از پلی مراز I باکتری اشرشیاکلائی انجام شد. محصولات واکنش با الحاق  $dATP$  ( $\alpha$ - $^{32}P$ ) نشان‌دار شدند و تصاویر توالی سایز مارکر توسط اتورادیوگرافی به دست آمدند. برای توالی، قطعه DNA کلون شده باید پس از تجزیه شدن به قطعات کوتاه، به وکتور M۱۳mp۱۸ و M۱۳mp۱۹ کلون شود (برای رشته‌های coding و noncoding). در طی توالی قطعه DNA حاوی *iap*، توالی مشابه چندین بار در کلون‌های مختلف ظاهر شد. علاوه بر این، دشوار بود که توالی‌های تکراری دقیقاً با استفاده از قطعه کلنو در ۳۷ درجه سانتی‌گراد، به دلیل انتها غیراختصاصی واکنش‌های الحاق دی دنوکسی نوکلئوتید برای DNA الگو، به علت تشکیل ساختار ثانویه توسط توالی پالیندرومیک خوانده شود. به همین دلیل، چند ماه طول کشید تا توالی‌یی از منطقه CRISPR دقیقاً در سال ۱۹۸۷ مشخص شود. یک توالی تکراری خاص در پایین دست

کدون پایان ترجمه، برای ژن *iap* (شکل ۲) شناسایی شد. قابل توجه است که دقیقاً همان منطقه را می توان فقط یک روز با استفاده از تکنولوژی فعلی با تقویت منطقه هدف توسط PCR، به طور مستقیم از ژنوم، به دنبال نشاندارکردن فلورسنت و توالی یابی چرخه در دمای ۷۲ درجه سانتی گراد (شکل ۳) توالی قرار داد. ویژگی توالی تکراری بسیار مبهم و غیرمنتظره بود (که در بحث ذکر شده است)؛ حتی اگر عملکرد آن درک نشده بود (۵). به طور قابل توجهی، همان توالی دارای تقارن دیا<sup>۱</sup> ۱۴ جفت باز، پنج بار پراکنده بین تکرارها با یک توالی ۳۲ نوکلئوتیدی متغیر تکرار شد (شکل ۲). توالی های نوکلئوتیدی خوب حفظ شده (که شامل تقارن دیا<sup>۱</sup> هستند)، تحت عنوان "توالی های پالیندرومی خارج ژنی تکراری" (۷)؛ قبلاً در اشرشیاکلا<sup>۱</sup> و سالمونلا انتریکا سرووار تیفی موریوم یافت و برای تثبیت mRNA پیشنهاد شد. با این حال، بین توالی خارج ژنی تکراری و توالی های تکراری تشخیص داده شده در پایین دست ژن *iap* هیچ شباهتی یافت نشد. در واقع، این توالی در آن زمان، در پایگاه داده های توالی، منحصر به فرد بود. همان طور که بعداً معلوم شد، این اولین مواجهه با توالی CRISPR بود. بعد از آن، توالی های مشابه با آنالیز هیبریداسیون ساترن بلات در سایر گونه های اشرشیاکلا<sup>۱</sup> (C600 and Ymel) و در دو عضو دیگر انتروباکتریاسه، شیگلادیسانتتری و سالمونلا انتریکا سرووار تیفی موریوم شناسایی شدند (۸) و پس از آن، توالی های تکراری مشابه در اعضای شاخه های اکتینوباکتریاسه، مانند مایکوباکتریوم توبرکلوزیس یافت شدند (۹)؛ اما در سویه نزدیک به مایکوباکتریوم لیره استفاده از این توالی های تکراری متغیر را برای تعیین تیپ باعث نمی شود (۱۰).



شکل ۲ اولین سیستم CRISPR در *E. coli*

<sup>۱</sup> dyad symmetry



شکل ۳. اولین توالی CRISPR در *E. coli*.

## کشف CRISPR در آرکی باکتری

اصلاً پیشرفت عمده زمانی حاصل شده که توالی‌های تکراری مشابه در آرکی باکتر هالوفراکس مدیترانه‌ای توسط "موجیکا" و همکارانش در طول تحقیقات بر روی مکانیسم‌های تنظیمی آرکی باکتری بسیار نمک دوست، که برای انطباق با محیط‌های با نمک بالا صورت گرفت؛ شناسایی شد (۲). رونویسی مناطق ژنومیک حاوی توالی‌های تکراری توسط نورترن بلات، نشان داده شده است (۲)؛ اما برای پردازش رونوشت‌ها به چند محصول مختلف RNA، تنها اخیراً شواهد قانع کننده‌ای نشان داده شده است (۱۱). در این مطالعه، نویسندگان ابتدا عنوان کردند که این توالی‌های تکراری می‌تواند در تنظیم بیان ژن دخالت داشته باشد و احتمالاً تسهیل تبدیل DNA ی دورشته از فرم B به Z را برای اتصال خاص پروتئین تنظیم کننده موجب شود. در حقیقت، اغلب در آن زمان مطرح شد که مقدار بالای GC ژنوم‌های هالوفیلیک می‌تواند این انتقال B به Z را برای اهداف تنظیمی در غلظت بالای نمک داخل سلولی مشخصه هالو آرکیا تسهیل کند. با این حال، چنین توضیحی برای باکتری‌ها معتبر نیست. کمی بعد از آن، همان نویسندگان، یک توالی تکراری در هالوفراکس ولکانی یافتند و فرض کردند که این توالی‌های تکراری می‌توانند در دسته‌بندی رپلیکون دخالت داشته باشند (۱۲). در عین حال، اختراع دستگاه‌های تعیین توالی خودکار و توسعه روش‌های کارآمد برای تعیین توالی DNA، طی دهه ۹۰؛ دانشمندان را اولین بار به دسترسی به توالی ژنوم کامل رهنمون گردید. شروع با هموفیلوس آنفولانزا (۱۳) و سپس متانوکوکوس جاناشی (۱۴) و ساکارومایسس سروزیه، در واقع، تمام سه دامنه زندگی وارد دوره ژنومیک شدند. سپس توالی‌های تکراری غیر معمول همراه با توالی‌های

غیر محافظت شده، ابتدا در اشرشیا کلای و هالوفراکس مدیترانه‌ای و سپس در تعداد زیادی از ژنوم‌های باکتری و آرکی باکتری شناسایی و با استفاده از نام‌های مختلف توسط نویسندگان مختلف شرح داده شدند؛ مانند تکرارهای فاصله تنظیمی کوتاه (short regularly spaced repeats ( SRSRs )) (۲)؛ تکرارهای مستقیم پراکنده فاصله‌ها (spacers interspersed direct repeats (SPIDRs))؛ خوشه بزرگ از تکرارهای پشت سر هم (large cluster of tandem repeats (LCTRs)) (۲۵). در آرکیای گرمادوست پیروکوکوس آبیاسی و پیروکوکوس هوریکوشی، دو مجموعه توالی LCTR به صورت متقارن در هر طرف از مبدا همانندسازی قرار گرفتند؛ که این دوباره نقش احتمالی دسته‌بندی کروموزوم را نشان می‌دهد. با این حال، آن‌ها با تعدادی بیش‌تر و پیچیده‌تر در ژنوم پیروکوکوس فوریزوس قرار داشتند.

اصطلاح CRISPR، توسط "جانسن" و همکارانش در سال ۲۰۰۲ پیشنهاد شد (۷). مطالعات ژنتیکی مقایسه‌ای ویژگی‌های مشترک CRISPR را روشن کرد؛ یعنی ۱. آن‌ها در مناطق بین ژنی قرار دارند؛ ۲. آن‌ها حاوی چند تکرار مستقیم کوتاه با تغییرات توالی بسیار کمی هستند؛ ۳. تکرارها با توالی‌های غیرمحافظت شده پراکنده‌اند؛ ۴. یک توالی الگوی مشترک چند صد جفت باز در یک طرف خوشه تکرار قرار دارد.

این واقعیت که این توالی‌های مبهم در دو دامنه مختلف زندگی حفظ شدند، به نقش کلی‌تر این توالی‌ها اشاره دارد. توالی‌های CRISPR در تقریباً تمام ژنوم آرکی باکتری‌ها و حدود نیمی از ژنوم‌های باکتریایی یافت می‌شوند. ارائه آن‌ها به طور خانواده گسترده‌ای است از توالی‌های تکراری که در پروکاریوت‌ها توزیع می‌شود. تاکنون توالی‌های CRISPR در هیچ ژنوم یوکاریوتی یافت نشده است.

## شناسایی ژن‌های Cas

تجمع توالی ژنومی، در آغاز قرن حاضر، دانشمندان را قادر ساخت که زمینه ژنومیک مناطق CRISPR را در بسیاری از ارگانیزم‌ها مقایسه کنند، که این، به کشف چهار ژن محافظت شده که به طور منظم در مجاورت مناطق CRISPR حضور دارند، منجر شد. ژن‌ها از ژن‌های ۱ تا ۴ مرتبط با CRISPR (cas ۱ به cas ۴) تعیین شدند (۷). برای Cas ۱ و Cas ۲، به دامنه‌های عملکردی پروتئین شناخته شده هیچ شباهتی ملاحظه نمی‌شود. در مقابل، Cas ۳ حاوی هفت خصوصیات موتیف است که از هلیکازهای سوپر خانواده ۲ تشکیل شده است؛ در حالی که Cas ۴ که با اگزونوکلازهای RecB مرتبط بود و به عنوان بخشی از کمپلکس RecBCD برای برش انتهایی از شکست دو رشته برای شروع نوترکیبی همولوگ کار می‌کرد، یافت شد. بنابراین، پیش‌بینی شده بود که Cas ۳ و Cas ۴ در متابولیسم DNA، از جمله تعمیر و نوترکیبی DNA، تنظیم رونویسی و جداسازی کروموزوم دخالت داشته باشند و با توجه به ارتباط آن‌ها با CRISPR، عنوان شد که پروتئین‌های Cas در پیدایش جایگاه CRISPR درگیر هستند (۷).

تقریباً همزمان، "ماکارووا" و همکارانش، به طور مستقل و سیستماتیک، زمینه‌های حفظ شده ژن در تمام ژنوم پروکاریوتی قابل دسترس در آن زمان را تجزیه و تحلیل کردند و خوشه‌های متعددی از ژن‌های مربوط به ژن‌های cas در ژنوم آرکی باکترهای گرمادوست افراطی و در دو باکتری گرمادوست با توالی‌های ژنوم موجود، (آکوفیکس و ترموتوگا) را شناسایی کردند. این ژن‌های محافظت شده در آن زمان، در آرکی باکتری‌ها و باکتری‌های مزوفیلیک و ترموفیلیک متوسط، یافت نشد. بر اساس این مشاهدات، پیش بینی شده بود که این پروتئین‌ها می‌توانند بخشی از یک سیستم تعمیر DNA مشخص نشده و مبهم مخصوص ارگانسیم‌های ترموفیلیک می‌باشد (۸).

## کشف عملکرد CRISPR

در آغاز عصر ژنوم، بیش تر توالی ژنوم آرکی باکتریایی از آن دسته از ارگانسیم‌های ترموفیلیک و هیپرترموفیلیک بودند. همچنین، آرکی باکتری‌های ترموفیلیک، علاوه بر باکتری‌های هیپرترموفیلیک مانند آکوفیکس آنولیکوس<sup>۱</sup> و ترموتوگا ماریتیمی<sup>۲</sup> نسبت به ارگانسیم‌های مزوفیلیک CRISPR بیش تر و بزرگ تر دارند (۷). بر اساس این مشاهدات در اوایل مطرح شد که عملکرد یک CRISPR ممکن است به سازگاری ارگانسیم‌ها با درجه حرارت بالا مربوط باشد. با این حال، با در نظر گرفتن بیش تر توالی‌های در دسترس، معلوم شد که این همبستگی قوی نیست و بسیاری از ارگانسیم‌های مزوفیلیک نیز شامل توالی CRISPR بودند. زمان ابراز پیروزی از اکتشاف وقتی ایجاد شد که "فرانسیسکو موجیکا" در آلکانتا و "کریستین پورسل" در اورسای به طور مستقل متوجه شدند که مناطق فاصله‌انداز بین توالی‌های تکراری، با توالی‌های باکتریوفاژها، پروفاژها و پلاسمیدها همولوگ هستند (۱). بدیهی است براساس بررسی ادبیات، آن‌ها اشاره کردند که فاژها و پلاسمیدها، سویه‌های میزبان را که حاوی توالی‌های فاصله‌انداز همولوگ در CRISPR هستند؛ آلوده نمی‌کنند. از این مشاهدات، آن‌ها به طور مستقل مطرح کردند که توالی‌های CRISPR در چارچوب یک سیستم دفاعی زیستی، برای محافظت از سلول‌ها از ورود این عناصر متحرک ژنتیکی خارجی شبیه سیستم RNAi یوکاریوتی عمل می‌کنند. دو گروه همچنین پیشنهاد کردند که CRISPR می‌تواند به نحوی باعث جذب قطعات DNA مهاجم خارجی شود تا خاطره‌ای از حمله ژنتیکی گذشته ایجاد کند (۱). در سومین مقاله، "بولتین" و همکارانش این مشاهدات را تأیید کردند و بیش تر ارتباط بین تعداد فاصله‌اندازها از مبدأ فاژ و درجه مقاومت به عفونت فاژ را متوجه شدند. آن‌ها پیشنهاد کردند که از CRISPR می‌توان برای تولید RNA آنتی سنس استفاده کرد (۱۴). در سال ۲۰۰۵، نویسندگان، اکتشاف ژن‌های cas را تصدیق کردند، که نشان

<sup>۱</sup> Aquifex aeolicus

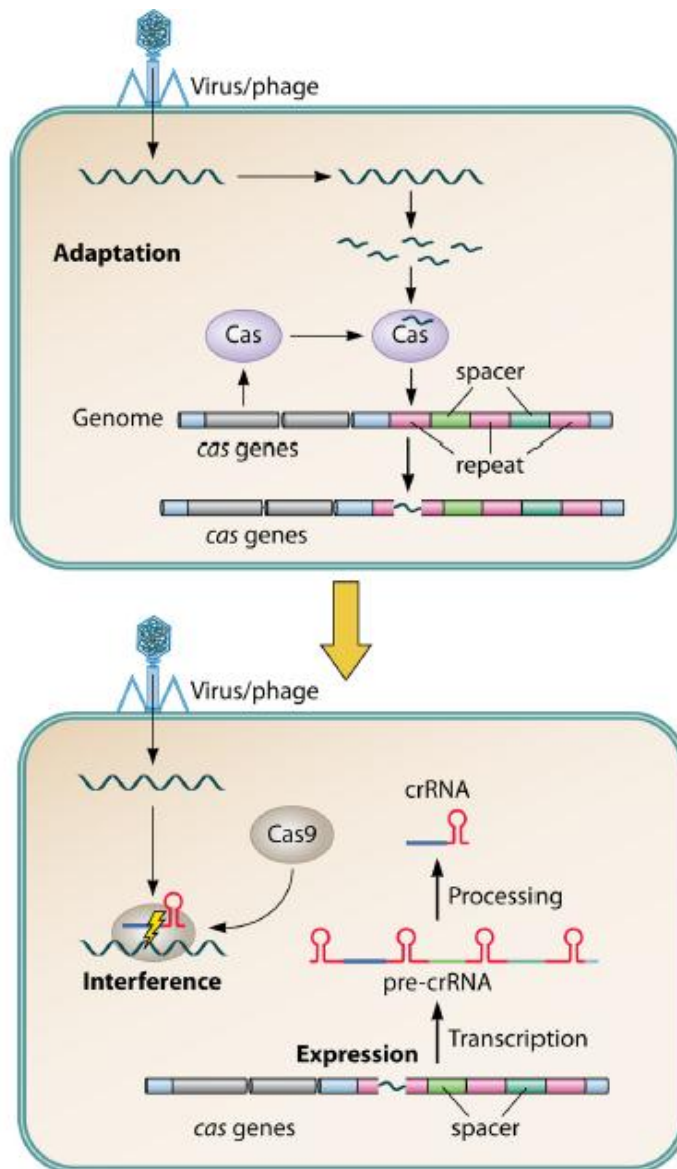
<sup>۲</sup> Thermotoga maritima

می دهد پروتئین هایی که توسط این ژن ها رمزگذاری شده اند، باید در عملکرد این سیستم ایمنی پیش بینی کننده پروکاریوتی قرار گیرند (۲). نقش پیش بینی کننده پروتئین های Cas بر اثر ایمنی پروکاریوتی، یک سال پس از آن در یک مقاله تحلیلی جامع منتشر شده توسط "ماکارووا" و همکارانش تأکید شد (۴). ماکارووا و همکارانش بر اساس کار قبلی خود، تجزیه و تحلیل دقیقی از توالی های پروتئین Cas انجام دادند و تلاش کردند عملکرد آن ها را در مکانیسم شبیه سیستم RNAi یوکاریوتی پیش بینی کنند (۴). قابل ذکر است که در بسیاری از موارد، این پیش بینی های عملکردی غیر معمول، همان طور که در مورد اینتگراز Cas1، به طور آزمایشگاهی چند سال بعد تأیید شد؛ همچنان به هدایت تحقیقات تجربی روی سیستم CRISPR-CAS کمک می کند. مهم این است که آن ها مشخص کرده اند که سیستم CRISPR-CAS با اجزای حافظه آن، به سیستم ایمنی سازگار با مهره داران شبیه است؛ زیرا تفاوت اساسی این است که سیستم ایمنی حیوانات ارثی نیست. با توجه به تنوع سیستم CRISPR-CAS، توزیع نامنظم آن ها با توجه به تحرک زیاد و در همه جای آرکی باکتری ها، ماکارووا و همکارانش عنوان کردند که سیستم CRISPR-CAS در آرکی باکتری ها ظاهر شده و به صورت افقی به باکتری ها گسترش یافته است. همچنین پیشنهاد کردند که سیستم های CRISPR-CAS بتوانند برای خاموش کردن ژن های موجود در موجودات کد کننده پروتئین های Cas، مورد استفاده قرار گیرند (۴).

عملکرد سیستم CRISPR-CAS، به عنوان سیستم ایمنی به دست آمده از پروکاریوت به طور آزمایشگاهی در سال ۲۰۰۷، با استفاده از باکتری استرپتوکوکوس ترموفیلوس در همان سال ثابت شد (۱۸). قرارداد توالی فاژ در ناحیه فاصله انداز CRISPR باکتری استرپتوکوکوس ترموفیلوس (که این سویه به فاژ مربوط مقاوم می باشد)، ساخته شده است. از سوی دیگر، این مقاومت باکتریایی به عفونت فاژ زمانی ناپدید شد که توالی DNA خارجی مربوط از ژنوم فاژ حذف شد. علاوه بر این، به طور تجربی نشان داده شده است که CRISPR-Cas ترانسفورماسیون پلاسمیدهای حامل توالی های مطابق فاصله اندازهای CRISPR را در برمی گیرد (۱۹). سپس گروه Van der Oost سیستم ایمنی را با استفاده از CRISPR باکتری اشرشیاکلائی (که در ابتدا در سال ۱۹۸۷ کشف شد)، بازسازی کردند. آن ها نشان دادند که مولکول های RNA ی پردازش شده از رونویسی منطقه CRISPR با همکاری پروتئین های Cas تولید شده از ژن های واقع در کنار CRISPR عمل می کنند (۲۰). تقریباً در همان زمان، آنالیز متاژنومیکس آرکی باکتری ها توسط Andersson و Banfield، تغییرات پویا توالی های در جایگاه CRISPR در یک دوره زمانبندی از ماه و توالی های جدید فاصله انداز مربوط به فاژهای در جوامع مشابه ظاهر شده را نشان دادند (۲۱). پس از آن، سیستم CRISPR-Cas باکتری استرپتوکوکوس ترموفیلوس بیان شده در اشرشیاکلائی، محافظت هترولوگوسی در برابر ترانسفورماسیون پلاسمید و عفونت فاژ توسط سیستم CRISPR-Cas9 بازسازی شده از



باکتری استرپتوکوکوس ترموفیلوس را نشان می‌دهد (۲۲). این کار همچنین نشان داد که Cas9 در این مورد، تنها ژن لازم برای CRISPR رمزگذاری شده می‌باشد. کمی بعد از آن، ثابت شد کمپلکس خالص-Cas9 (CRISPR RNA (crRNA)، به برش DNA هدف در *in vitro* قادر می‌باشد (۲۳). سپس سیستم CRISPR-Cas باکتری استرپتوکوکوس پیورنز برای انجام ویرایش ژنوم در سلول‌های عصبی انسان و کلیه موش، اعمال شد (۲۴). بنابراین، CRISPR-Cas به طور گسترده به عنوان سیستم ایمنی به دست آمده از پروکاریوت شناخته شد (۲۵). (مراحل مختلف اساسی عملکرد این سیستم در شکل ۴ نشان داده شده است). پروتئین‌های گوناگون و بسیار متنوعی در مراحل مختلف ایمنی CRISPR دخیل هستند. آن‌ها انواع فعالیت‌های دست کاری اسید نوکلئیک پیش‌بینی شده، مانند نوکلئازها، هلیکازها و پلی‌مرازها را نمایش می‌دهند (۲۶). به طور خلاصه، Cas1 و Cas2 در سراسر انواع شناخته شده‌ترین انواع سیستم CRISPR-Cas محافظت می‌شوند و کمپلکسی را تشکیل می‌دهند که این، نشان‌دهنده مدل آدپتاسیون مورد نیاز برای الحاق فاصله‌اندازهای جدید در آرایه‌های CRISPR است. در مرحله بیان، جایگاه CRISPR رونویسی و رونوشت pre-crRNA توسط اندونوکلازهای تیپ‌های خاص Cas به شکل crRNA بالغ پردازش می‌شود. در مرحله تداخل، crRNA ها توسط اندونوکلازهای مؤثر Cas محدود می‌شوند و کمپلکس‌های مربوط به منظور برش DNA یا RNAی هدف به روش وابسته به توالی به کار گرفته می‌شوند (شکل ۴). به طور مشخص، بر خلاف مدل آدپتاسیون، آنزیم‌های Cas که در مراحل بیان و تداخل دخیل‌اند، از تیپ CRISPR-Cas با یکدیگر متفاوت هستند و آنزیم‌های مشابه می‌توانند در هر دو مرحله از ایمنی شرکت کنند.

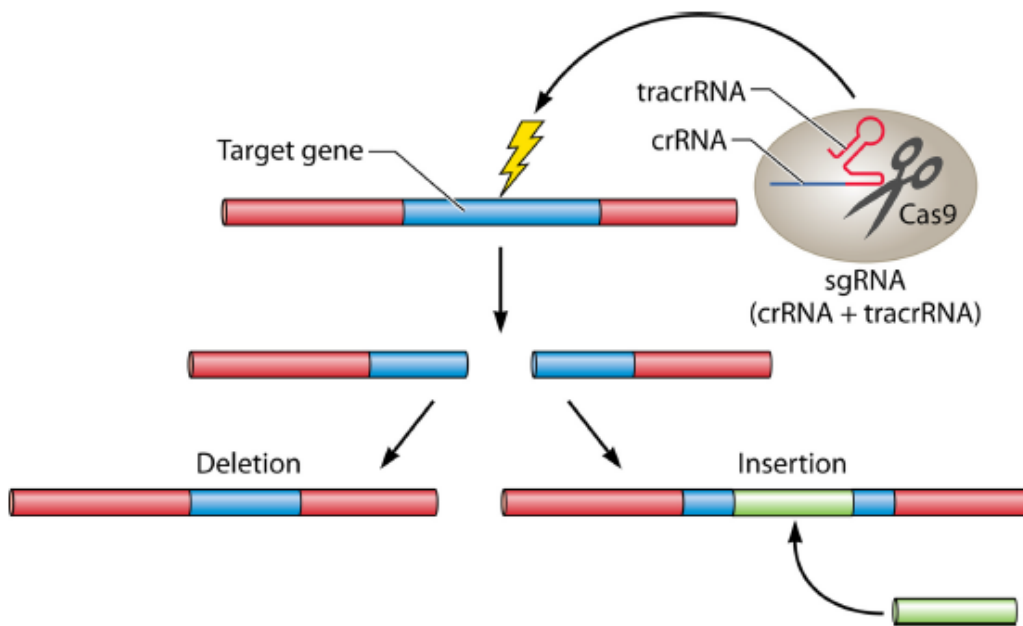


شکل ۴. فرایند CRISPR-Cas سیستم ایمنی را به دست آورد. تصویر بالا (آداپتاسیون) DNA مهاجم توسط پروتئین‌های CAS شناسایی شده، قطعه قطعه شده و به درون مناطق فاصله انداز CRISPR گنجاده شده، و در ژنوم ذخیره می‌شود. تصویر پایین (بیان). Pre-crRNA توسط رونویسی منطقه CRISPR تولید می‌شود و به واحدهای کوچک تر RNA به نام crRNA تبدیل می‌شود. (پایین) تداخل. با بهره‌گیری از homology توالی spacer موجود در crRNA، DNA خارجی گرفته شده است، و یک کمپلکس با پروتئین Cas دارای فعالیت نوکلئازی DNA را برش می‌دهد.

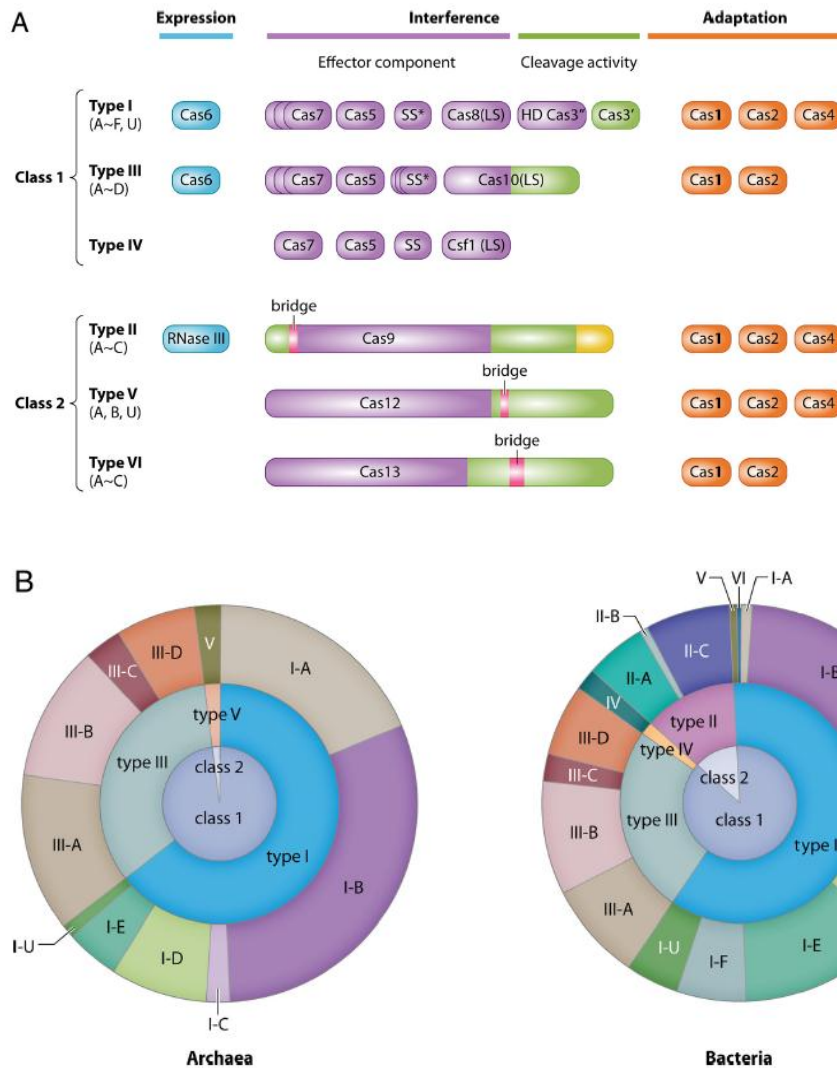
## تنوع و طبقه‌بندی CRISPR-Cas

قابل توجه است که سویه‌های نزدیک به هم، در محتوا و توزیع CRISPR به طور قابل توجهی متفاوت هستند. به عنوان مثال، در جنس مایکوباکتریوم، در گونه توبرکلوزیس CRISPR وجود دارد؛ اما در لپره وجود ندارد. از سوی دیگر، اشرشیاکلاهی و مایکوباکتریوم آویوم و همچنین متانوتر موباکترترموتروفیکوس و آرکیاگلوباس فولجیدوس به طور فیلوژنتیکی، دارای توالی‌های تقریباً یکسان CRISPR هستند (۷). تعداد آرایه‌های CRISPR در یک ژنوم از ۱ تا ۱۸ متغیر و تعداد واحدهای تکرار در یک آرایه CRISPR از ۲ تا ۳۷۴ متفاوت است (۲۷). بر اساس پایگاه داده CRISPR (<http://crispr.i2bc.paris-saclay.fr>) تا ماه May ۲۰۱۷، CRISPR در توالی کامل ژنوم ۲۰۲، از ۲۳۲ گونه‌های آرکی باکتری‌های آنالیز شده (۸۷٪) و ۳۰۵۹ از ۶۷۸۲ گونه‌های باکتریایی (۴۵٪) شناسایی شدند. جالب توجه است بر اساس بررسی از ۱۷۲۴ طرح اولیه ژنوم، سیستم CRISPR-Cas در جوامع میکروبی محیطی (۴/۱۰٪ در باکتری‌ها و ۱/۱۰٪ در آرکی باکتری‌ها) بسیار کم‌تر شایع هستند. به نظر می‌رسد این تفاوت بزرگ بین شیوع تخمین زده شده از ژنوم کامل میکروب‌های کشت داده شده به نسبت آن‌هایی که کشت داده نشده‌اند؛ به دلیل عدم سیستم CRISPR-Cas در میان دودمان‌های باکتری‌های اصلی است که دارای نمایندگان کشت نشده بودند (۲۸). همان‌طور که در شکل ۵ نشان داده شده است، آخرین طبقه‌بندی سیستم CRISPR-Cas، شامل دو کلاس ۱ و ۲ بر اساس پروتئین‌های مؤثر (effector) کد شده است (۲۵). سیستم‌های کلاس ۱- CRISPR-Cas با کمپلکس‌های مؤثر چند زیر واحدی (که شامل ۴ تا ۷ پروتئین Cas موجود در استوکیومتری غیر یکنواخت است)، کار می‌کنند. این سیستم در باکتری‌ها و آرکی باکتری‌ها، از جمله در تمام هایپرترموفیل‌ها، شامل تقریباً ۹۰٪ از همه جایگاه‌های مشخص شده CRISPR-cas گسترده است. باقی مانده (حدود ۱۰٪ به کلاس ۲ متعلق است که از یک پروتئین مؤثر چند دمینی استفاده می‌کنند و تقریباً در باکتری‌ها یافت می‌شوند (۲۶). هر کلاس، در حال حاضر، شامل سه تیپ، یعنی تیپ‌های I، III و IV در کلاس ۱ و تیپ‌های II، V، و VI در کلاس ۲ می‌باشد. تیپ‌های I، II و III، به واسطه حضور پروتئین‌هایی با اثر منحصر به فرد، به راحتی قابل تشخیص هستند: Cas3 برای نوع I، Cas9 برای نوع II و Cas10 برای نوع III. کمپلکس‌های مؤثر مولتی متریک سیستم‌های تیپ I و III که به ترتیب به عنوان کمپلکس CRISPR-associated برای دفاع ضد ویروسی و کمپلکس‌های Csm/Cmr شناخته می‌شوند؛ از لحاظ معماری مشابه و تکاملی مرتبط هستند (۲۷). بر خلاف دیگر سیستم‌های CRISPR-Cas شناخته شده، سیستم‌های تیپ IV با عملکرد نامشخص، شامل مدل آدپتاسیون دارای نوکلئازهای Cas1 و Cas2 نیست (۲۷). بدیهی است مدل‌های اثرگذار سیستم‌های زیرگروه III-B، برای استفاده از فاصله اندازهای تولید شده توسط سیستم‌های تیپ I شناخته شده‌اند که این، مقایسه‌ای بودن سیستم‌های CRISPR-Cas را تصدیق می‌کند (۲۸)؛

اگرچه بسیاری از ژنوم‌های کدگذاری شده در سیستم‌های نوع IV جایگاه قابل شناسایی CRISPR را نشان نمی‌دهند. از این رو، سیستم‌های نوع IV مشابه سیستم‌های زیرگروه III-B، در استفاده کردن از crRNA هایی که از آرایه‌های مختلف CRISPR زمانی که این‌ها در دسترس است؛ مستثنا نمی‌باشد (۲۳). در نهایت، هر تیپ به چندین زیر گروه طبقه‌بندی شده است ( IA به F و U و III-A به D در کلاس ۱؛ و II-A به C ، V-A به E و U و VI-A به C در کلاس ۲ ) ( ۲۹ ). شکل ۶ B ، توزیع سیستم‌های CRISPR-Cas را در آرکی باکتری‌ها و باکتری‌ها نشان می‌دهد.



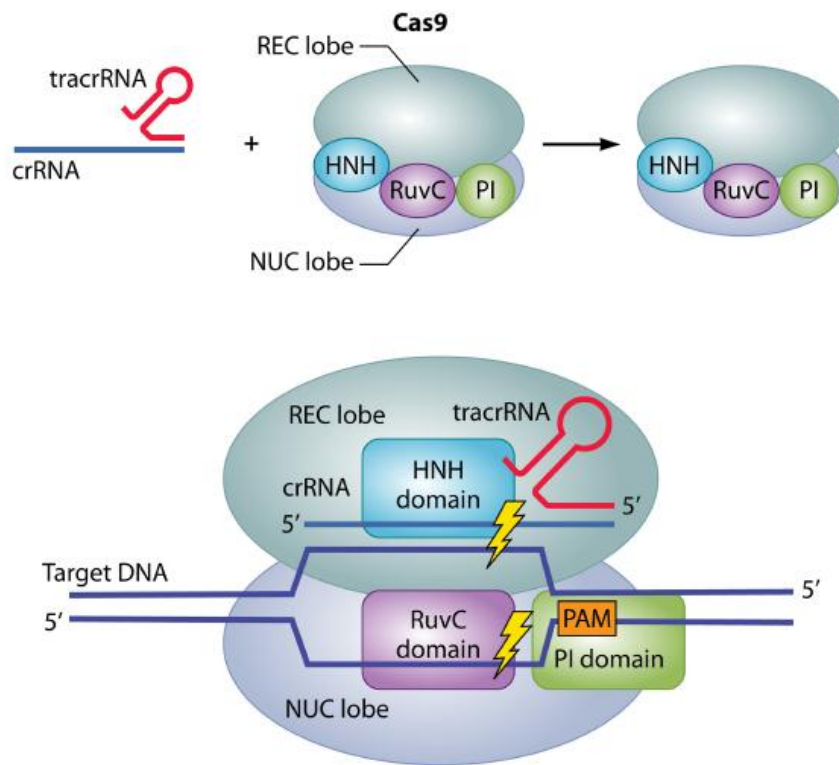
شکل ۵. ویرایش ژن توسط CRISPR-Cas9. اصل ویرایش ژنوم، برش DNA ی دو رشته در موقعیت هدف بر روی ژنوم است. تیپ II ، به عنوان ساده ترین نوکلئاز هدفمند در میان سیستم‌های CRISPR-Cas است. CRISPR RNA (crRNA) ، دارای توالی همولوگ با سایت هدف ، و trans-activating CRISPR RNA (tracrRNA) برای به دست آوردن نوکلئاز Cas9 به سایت هدف ، به اندازه کافی هستند . اتصال مصنوعی crRNA و tracrRNA به یک زنجیره RNA ( single-guide RNA [sgRNA] ) بر عملکرد اثر ندارد. هنگامی که کمپلکس Cas9-sgRNA ژن هدف را برش می‌دهد ، اختلال عملکرد ژن با جهش deletion یا insertion آسان است .



شکل ۶. جدیدترین طبقه‌بندی سیستم‌های ایمنی CRISPR-CAS. (A) بر اساس آنالیز دقیق توالی و سازمان ژن پروتئین Cas، CRISPR-Cas، بسته به این که یک افکتور کمپلکس مرکب از پروتئین‌های متعدد Cas یا یک افکتور واحد باشد؛ به دو کلاس عمده طبقه‌بندی شد. به علاوه تیپ‌های مرسوم I، II و III، تیپ‌های IV و V به ترتیب به کلاس‌های ۱ و ۲ اضافه شدند. تیپ‌های IV و V آن‌هایی‌اند که Cas ۱ و Cas ۲ برای فرایند آدپتاسیون در همان جایگاه CRISPR لازم‌اند. تیپ VI اخیراً به کلاس ۲ اضافه شده است. (B) نمودار نشان‌دهنده نسبت‌های مشخص شده جایگاه CRISPR-cas، در کل ژنوم باکتریایی و آرکی باکتریایی از نوشته‌های منتشر شده می‌باشد. سهم جایگاه‌هایی را که سیستم‌های ناقص را رمزگذاری می‌کنند و یا نمی‌توانند به طور یکنواخت طبقه‌بندی شوند، شامل نمی‌شوند.

## سیستم‌های کلاس ۲، مناسب برای تکنولوژی ویرایش ژن

معماری ساده از کمپلکس‌های اثرگذار، سیستم‌های کلاس ۲، CRISPR-Cas را برای توسعه یک نسل جدید از تکنولوژی‌های ویرایش ژن انتخاب مناسبی ساخته است (شکل ۷). چندین اثرگذاری‌های متمایز کلاس ۲، از جمله Cas۹ در تیپ II، Cas۱۲a (قبلاً Cpf۱)، Cas۱۲b (C۲c۱) در تیپ V و Cas۱۲a (C۲c۲) و Cas۱۳b و Cas۱۳a (C۲c۳) در تیپ VI گزارش شده است (۲۹). شایع‌ترین و بهترین پروتئین اثرگذار چند دمینی مورد مطالعه Cas۹ است؛ یک اندونوکلیاز وابسته به crRNA، متشکل از دو دمین نوکلئاز غیر مرتبط، (HNH و RuvC)، که به ترتیب مسئول برش رشته‌های جابه‌جا شده (غیر هدف) و DNAی هدف در کمپلکس crRNA-target DNA می‌باشد. جایگاه تیپ CRISPR-cas II همچنین یک trans-activating crRNA (tracrRNA) را که ممکن است از CRISPR مربوط تکامل یافته باشد، کدگذاری کند. همچنین مولکول tracrRNA برای پردازش pre-crRNA و شناسایی هدف در سیستم‌های تیپ II ضروری است. مکانیسم مولکولی برش DNAی هدف، توسط کمپلکس Cas۹-crRNA، که در شکل ۷ نشان داده شده است، در سطح اتمی به وسیله آنالیز ساختار کریستال از کمپلکس DNA-Cas۹-crRNA کشف شده است (۳۰). یک ژن در ابتدا به نام *cpfA* در چندین ژنوم باکتری‌ها و آرکی باکتری‌ها شناخته شده است؛ جایی که در مجاورت *cas۱*، *cas۲* و یک آرایه CRISPR وجود دارد (۲۵). Cas۱۲a (Cpf۱)، نمونه اولیه افکتورهای تیپ V، حاوی دو دمین RuvC-like و فاقد دمین HNH می‌باشد. با این حال، آنالیز ساختاری اخیر از کمپلکس Cas۱۲a-crRNA-target DNA دمین نوکلئاز ثانویه، با یک تاخوردگی منحصر به فرد نشان می‌دهد که به طور عملکردی با دمین HNH Cas۹ آنالوگ است (۳۰). Cas۱۲a یک نوکلئاز RNA راهنما است که به tracrRNA که برای فعالیت Cas۹ ضروری است؛ نیاز ندارد (۳۱). همچنین پروتئین با Cas۹ در الگوی برش و تشخیص موتیف مجاور توالی‌های خارجی وارد شده آن، (که رشته‌های هدف را تعیین می‌کند)، متفاوت است.



شکل ۷. مکانیسم برش DNA هدف، به وسیله Cas9-tracrRNA-crRNA کمپلکس Cas9-crRNA، tracrRNA به DNA خارجی حاوی PAM متصل می‌شود؛ در حالی که باندها و مبدهای Cas9، DNA خارجی دو رشته‌ای را برای تحریک تشکیل دوتایی crRNA و DNA خارجی باز می‌کند. Cas9 شامل دو ناحیه به نام‌های لوب REC (recognition) و لوب NUC (nuclease) می‌باشد. لوب REC مسئول تشخیص اسید نوکلئیک است. لوب NUC حاوی دامین‌های HNH و نوکلئازی RuvC و یک منطقه C-terminal حاوی یک دامین PAM-interacting (PI) است. دامین HNH و دامین RuvC، به ترتیب رشته DNA را برش می‌دهند؛ تشکیل یک دوبلکس crRNA با دیگر رشته DNA، به طوری که شکست دو رشته در DNA هدف اتفاق می‌افتد.

کشف دو پروتئین افکتور کلاس ۲، Cas9 و Cas12a که با هم ارتباط دوری دارند، مطرح می‌کند که ممکن است انواع متمایز دیگری از چنین سیستم‌هایی وجود داشته باشند. در واقع اخیراً؛ Cas12b (تیپ V)، Cas13a و Cas13b (نوع VI)، که از Cas9 و Cas12a متمایز هستند؛ از طریق جست‌وجوی بیوانفورماتیک هدایت شده برای افکتورهای کلاس II کشف شده‌اند و فعالیت آن‌ها تأیید شده است (۳۲). افکتورهای تیپ V، همانند Cas9، به یک tracrRNA برای فعالیت هدفمند نیاز دارند. بسیاری از سیستم‌های CRISPR-Cas، به طور مشخص شده‌ای تاکنون DNA را هدف قرار داده‌اند و تنها سیستم‌های چند ترکیبی تیپ III-A و III-B علاوه بر این، به RNA هدف اختصاص داده شده‌اند. در مقابل، افکتورهای تیپ VI Cas13a و Cas13b به طور خاص RNA را هدف قرار

می دهند و به این ترتیب، RNAی مداخله گر، میانجی گری می کند. برخلاف افکتور های تیپ II و V، Cas13a و Cas13b فاقد مشخصه های دمین های نوکلئاز RuvC-like هستند و به جای آن حاوی یک جفت دمین HEPN<sup>1</sup> می باشند. کشف افکتورهای جدید کلاس ۲، به احتمال زیاد برای استفاده از سیستم های CRISPR به تکنولوژی مهندسی ژنوم فرصت های جدیدی را ارائه خواهد کرد (۳۳).

## کاربرد سیستم CRISPR/Cas9 در صنعت کشاورزی

دامداران قبلاً مارکرهای کروموزومی مرتبط با صفات شناخته شده را به عنوان صفات کمی شناخته اند و از مارکرها استفاده می کنند تا ویژگی های ارزشمندی را به طور انتخابی پیشگویی کنند. این روند با استفاده از فناوری های ویرایش ژن، به تازگی در گاوهای لبنی، به ترتیب به منظور محافظت در برابر ویروس ها و حذف شاخ ها، نشان داده شده است. کاربرد دیگر این سیستم در حیوانات، مهندسی تولید محصولات پزشکی یا تولید بافت است. مهندسی فعال کریسپر در محصولات تجاری و مدل های آزمایشگاهی برای افزایش عملکرد ایده آل، افزایش تحمل نسبت به خشکی و افزایش رشد در شرایط محدود مواد غذایی و تولید محصولات با خواص تغذیه ای بهبود یافته، مورد استفاده قرار می گیرد. استفاده از تکنولوژی کریسپر در ذرت و سویا نشان دهنده سرعت پذیرش تکنولوژی کریسپر خارج از آزمایشگاه است. هدف قرار دادن ژن مبتنی بر کریسپر نیز می تواند برای مبارزه با بیماری های گیاهی مورد استفاده قرار گیرد؛ همان طور که این نکته برای ویروس Curl leaf زرد گوجه فرنگی در *Nicotiana benthamian* نشان داده شده است (۳۴).

## کاربرد سیستم کریسپر در مواد غذایی و بیوتکنولوژی صنعتی

تاکنون کاربرد سیستم های کریسپر در باکتری ها، شامل ژنوتیپ، کشت محصولات صنعتی علیه ویروس ها، کنترل جذب و انتشار ژن های مقاوم به آنتی بیوتیک توسط باکتری ها و کشت های پروبیوتیک مهندسی را شامل می شود. موفقیت تجاری سیستم ایمنی طبیعی کریسپر برای واکسیناسیون استرپتوکوکوس ترموفیلوس (که از آن در فرآورده های لبنی مانند ماست و پنیر) استفاده می شود) راه را برای این سیستم در تغذیه هموار کرده است. کار اخیر نیز ثابت کرده که مفهوم تولید باکتری مفید این است که در برابر جذب و انتشار ژن هایی که مقاومت آنتی بیوتیکی را رمزگذاری می کنند؛ ایمن می شود. فناوری کریسپر بر تمامی صنایع مربوط به باکتری ها، قارچ ها و مخمرها تأثیر گسترده ای خواهند داشت؛ چرا که ما در صدد استفاده گسترده از این ویرایش کننده ژنوم در این ارگان ها هستیم.

<sup>1</sup> Higher eukaryote and prokaryote nucleotide binding



احتمالاً CRISPR/Cas9 برای مهندسی باکتری‌ها، مخمرها و قارچ‌های صنعتی برای تولید مواد شیمیایی سبز، از جمله سوخت‌های زیستی و مواد بیولوژیکی استفاده می‌شود (۳۵).

## نتیجه‌گیری

انقلاب ژنتیکی در اواسط دهه ۱۹۹۰، همراه با توسعه وسیع ابزارهای بیوانفورماتیک قدرتمند، در نهایت روشن ساختن عملکردهای CRISPR را فراهم کرد. آرایه‌های CRISPR و پروتئین‌های Cas به طور گسترده‌ای در ژنوم پروکاریوت‌ها، به خصوص در آرکی باکتری‌ها، توزیع شده و اکنون به عنوان سیستم ایمنی به دست آمده بسیار کارآمد، تشکیل شده‌اند. اگر چه کشف CRISPR-CAS به خودی خود نوعی ویژگی عالی از زیست‌شناسی بنیادی بود؛ به توسعه ابزار نسل بعدی برای مهندسی ژنتیک منجر شد. توسعه تکنولوژی ویرایش ژنوم توسط CRISPR Cas9، یادآور زمانی است که PCR متولد شد. هنگامی که در شرایط آزمایشگاهی، تکنیک‌های مهندسی ژنتیک، با استفاده از آنزیم‌های محدود کننده و آنزیم‌های اصلاح کننده نوکلئیک اسید ایجاد شد؛ باز هم اغلب یک کار پیچیده بود تا یک ژن واحد را کلون کنند. با این حال، این مشکل توسط اختراع PCR با استفاده از DNA پلی‌مراز مقاوم به حرارت کاهش یافت که عمیقاً باعث افزایش کاربرد تکنیک‌های مهندسی ژنتیک در تمام آزمایشگاه‌های زیستی در سراسر جهان شد. کشف DNA پلی‌مراز مقاوم به حرارت برای "انقلاب PCR" حیاتی بود؛ زیرا طراحی یک دستگاه PCR را برای استفاده عملی امکان‌پذیر ساخت. به طور مشابه، در مورد ویرایش ژنوم، انقلاب CRISPR با شناسایی سیستم آنزیمی مناسب (Cas9) امکان‌پذیر شد، که این، می‌تواند متدولوژی را برای بهره‌برداری از پتانسیل سیستم CRISPR-CAS آسان کند. کنجکاوی یک توالی تکراری مرموز و یک پرسش مستمر برای توضیح عملکرد آن، کشفیات بزرگ را به وجود آورد.

## رفرانس

۱. Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* ۲۰۰۵, ۶۰: ۱۷۴-۱۸۲. <https://doi.org/10.1007/s00239-004-0046-3>.
۲. Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV, A. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res* ۲۰۰۲. ۳۰: ۴۸۲-۴۹۶. <https://doi.org/10.1093/nar/30.2.482>.
۳. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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* ۲۰۰۶, ۱: ۷. <https://doi.org/10.1186/1745-6150-1-7>.
۴. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* ۱۹۸۷. ۱۶۹: ۵۴۲۹-۵۴۳۳. <https://doi.org/10.1128/jb.169.12.5429-5433.1987>.
۵. Nakata A, Amemura M, Makino K. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-۱۲ chromosome. *J Bacteriol* ۱۹۸۹, ۱۷۱: ۳۵۵۳-۳۵۵۶. <https://doi.org/10.1128/jb.171.6.3553-3556.1989>.
۶. Hermans PW, van Soolingen D, Bik EM, de Haas PE, Dale JW, van Embden JD. Insertion element IS<sup>987</sup> from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. *Infect Immun* ۱۹۹۱, ۵۹: ۲۶۹۵-۲۷۰۵.
۷. Groenen PM, Bunschoten AE, van Soolingen D, van Embden JD. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol Microbiol* ۱۹۹۳, ۱۰: ۱۰۵۷-

۱۰۶۵. <https://doi.org/10.1111/j.1365-2958.1993.tb00976.x>.
۸. Mojica FJ, Rodriguez-Valera F. The discovery of CRISPR in archaea and bacteria. FEBS J ۲۰۱۶, ۲۸۳: ۳۱۶۲-۳۱۶۹. <https://doi.org/10.1111/febs.13766>.
۹. Mojica FJ, Ferrer C, Juez G, Rodriguez-Valera F. Long stretches of short tandem repeats are present in the largest replicons of the archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. Mol Microbiol ۱۹۹۵, ۱۷: ۸۵-۹۳. [https://doi.org/10.1111/j.1365-2958.1995.mmi\\_17.1.0085.x](https://doi.org/10.1111/j.1365-2958.1995.mmi_17.1.0085.x).
۱۰. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM. Whole-genome randomsequencing and assembly of *Haemophilus influenzae* Rd. Science ۱۹۹۵, ۲۶۹: ۴۹۶-۵۱۲. <https://doi.org/10.1126/science.7۵۴۲۸۰۰>.
۱۱. Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science ۱۹۹۶. ۲۷۳: ۱۰۵۸-۱۰۷۳. <https://doi.org/10.1126/science.2۷۳.۵۲۷۸.1058>.
۱۲. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P۲. Proc Natl Acad Sci U S A ۲۰۰۱, ۹۸: ۷۸۳۵-۷۸۴۰. <https://doi.org/10.1073/pnas.141222098>.
۱۳. Morange M. What history tells us XXXVII. CRISPR-Cas: the discovery of an immune system in prokaryotes. J Biosci ۲۰۱۵, ۴۰: ۲۲۱-۲۲۳.
۱۴. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P. CRISPR provides acquired resistance against viruses in prokaryotes. Science ۲۰۰۷. ۳۱۵: ۱۷۰۹-۱۷۱۲. <https://doi.org/10.1126/science.1138140>.
۱۵. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science ۲۰۰۸. ۳۲۲: ۱۸۴۳-۱۸۴۵. <https://doi.org/10.1126/science.1160771>.

۱۶. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*. ۲۰۰۸;۳۲۱:۹۶۰ – ۹۶۴. <https://doi.org/10.1126/science.1159689>.
۱۷. Andersson AF, Banfield JF. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* ۲۰۰۸. ۳۲۰:۱۰۴۷–۱۰۵۰. <https://doi.org/10.1126/science.1157358>.
۱۸. Sapranaukas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* ۲۰۱۱, ۳۹:۹۲۷۵–۹۲۸۲. <https://doi.org/10.1093/nar/gkr606>.
۱۹. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas<sup>9</sup>-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* ۲۰۱۲. ۱۰۹:E۲۵۷۹ –E۲۵۸۶. <https://doi.org/10.1073/pnas.1208507109>.
۲۰. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas<sup>9</sup>. *Science* ۲۰۱۳. ۳۳۹:۸۲۳–۸۲۶. <https://doi.org/10.1126/science.1232033>.
۲۱. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* ۲۰۱۰. ۳۲۷:۱۶۷–۱۷۰. <https://doi.org/10.1126/science.1179555>.
۲۲. Jackson SA, McKenzie RE, Fagerlund RD, Kieper SN, Fineran PC, Brouns SJ. CRISPR-Cas: adapting to change. *Science* ۳۵۶ eaal۵۰۵۶. <https://doi.org/10.1126/science.aal۵۰۵۶>.
۲۳. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* ۲۰۱۰. ۱۱:۱۸۱–۱۹۰. <https://doi.org/10.1038/nrg۲۷۴۹>.
۲۴. Burstein D, Sun CL, Brown CT, Sharon I, Anantharaman K, Probst AJ. Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nat Commun* ۲۰۱۶, ۷:۱۰۶۱۳. <https://doi.org/10.1038/ncomms10613>.

۲۵. Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, Abudayyeh OO, Diversity and evolution of class ۲ CRISPR-Cas systems. *Nat Rev Microbiol* . ۲۰۱۷. ۱۵:۱۶۹-۱۸۲. <https://doi.org/10.1038/nrmicro.2016.184>.
۲۶. Burstein D, Harrington LB, Strutt SC, Probst AJ, Anantharaman K, Thomas BC, Doudna JA, Banfield JF. New CRISPR-Cas systems from uncultivated microbes. *Nature* ۲۰۱۷, ۵۴۲:۲۳۷-۲۴۱. <https://doi.org/10.1038/nature21059>.
۲۷. Rouillon C, Zhou M, Zhang J, Politis A, Beilsten-Edmands V, Cannone G, Graham S, Robinson CV, Spagnolo L, White MF. Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Mol Cell* ۲۰۱۳, ۵۲:۱۲۴-۱۳۴. <https://doi.org/10.1016/j.molcel.2013.08.020>.
۲۸. Garrett RA, Vestergaard G, Shah SA. Archaeal CRISPR-based immune systems: exchangeable functional modules. *Trends Microbiol* ۲۰۱۱, ۱۹: ۵۴۹-۵۵۶. <https://doi.org/10.1016/j.tim.2011.08.002>.
۲۹. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. Crystal structure of Cas<sup>9</sup> in complex with guide RNA and target DNA. *Cell* ۲۰۱۴, ۱۵۶:۹۳۵-۹۴۹. <https://doi.org/10.1016/j.cell.2014.02.001>.
۳۰. Yamano T, Nishimasu H, Zetsche B, Hirano H, Slaymaker IM, Li Y, Fedorova I. Crystal structure of Cpf<sup>1</sup> in complex with guide RNA and target DNA. *Cell* ۲۰۱۶. ۱۶۵:۹۴۹-۹۶۲. <https://doi.org/10.1016/j.cell.2016.04.003>.
۳۱. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Cpf<sup>1</sup> is a single RNA-guided endonuclease of a class ۲ CRISPR-Cas system. *Cell* ۲۰۱۵. ۱۶۳:۷۵۹-۷۷۱. <https://doi.org/10.1016/j.cell.2015.09.038>.
۳۲. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Discovery and functional characterization of diverse class ۲ CRISPR-Cas systems. *Mol Cell* ۲۰۱۵, ۶۰:۳۸۵-۳۹۷. <https://doi.org/10.1016/j.molcel.2015.10.008>.
۳۳. Murugan K, Babu K, Sundaresan R, Rajan R, Sashital DG. The revolution continues: newly discovered systems expand the CRISPR-Cas toolkit. *Mol Cell*

بررسی ارتباط بین پلی مورفیسم در ژن فیبریلین نوع ۳ در بیماران مبتلا به سندرم تخمدان پلی کیستیک

۲۰۱۷, ۶۸: ۱۵-۲۵. <https://doi.org/10.116/j.molcel.2017.09.007>.

۳۴. Ali, Z. et al. Efficient virus-mediated genome editing in plants using the CRISPR/Cas<sup>9</sup> system. *Mol. Plant* ۸, ۱۲۸۸-۱۲۹۱ (۲۰۱۵).

۳۵. Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, Bhattacharyya S, Shelton JM, Bassel-Duby R, Olson EN. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy

## CRISPR System from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology

Sohrabi M<sup>۱</sup>, Salmani H<sup>۲</sup>, Monzavi ZS<sup>۳</sup>, Abdolahi Z<sup>۴</sup>, Dehghani Sanij S<sup>۳</sup>, Morovvati A<sup>۴\*</sup>

۱. Department of microbiology, Qom Branch, Islamic Azad University, Qom, Iran

۲. student, MS in Biotechnology, Department of Biotechnology, Qom Branch, Islamic Azad University, Qom, Iran

۳. Lecture, MsC in Microbiology, Department of microbiology, Qom Branch, Islamic Azad University, Qom, Iran.

۴. Lecture, Phd students in Virology, Department of microbiology, Qom Branch, Islamic Azad University, Qom, Iran.

۲۳

**ABSTRACT** Clustered regularly interspaced short palindromic repeat (CRISPR)-Cas systems are well-known acquired immunity systems that are widespread in archaea and bacteria. The RNA-guided nucleases from CRISPR-Cas systems are currently regarded as the most reliable tools for genome editing and engineering. The first hint of their existence came in ۱۹۸۷, when an unusual repetitive DNA sequence, which subsequently was defined as a CRISPR, was discovered in the *Escherichia coli* genome during an analysis of genes involved in phosphate metabolism. Similar sequence patterns were then reported in a range of other bacteria as well as in halophilic archaea, suggesting an important role for such evolutionarily conserved clusters of repeated sequences. A critical step toward functional characterization of the CRISPR-Cas systems was the recognition of a link between CRISPRs and the associated Cas proteins, which were initially hypothesized to be involved in DNA repair

in hyperthermophilic archaea. Comparative genomics, structural biology, and advanced biochemistry could then work hand in hand, not only culminating in the explosion of genome editing tools based on CRISPR-Cas<sup>9</sup> and other class II CRISPR-Cas systems but also providing insights into the origin and evolution of this system from mobile genetic elements denoted casposons. To celebrate the ۳۰<sup>th</sup> anniversary of the discovery of CRISPR, this minireview briefly discusses the fascinating history of CRISPR-Cas systems, from the original observation of an enigmatic sequence in *E. coli* to genome editing in humans.

**KEYWORDS** archaea, casposon, genome editing, repeated sequence