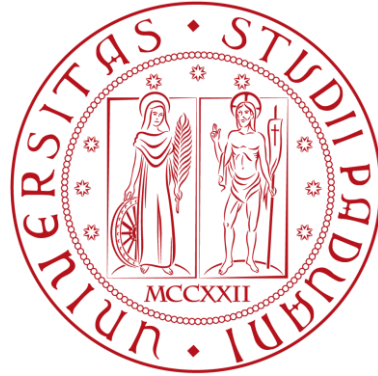


Università degli Studi di Padova
Dipartimento di Biologia
Corso di Laurea in Biologia



Elaborato di Laurea
**NUOVE TECNOLOGIE PER IL GENOME
EDITING: IL SISTEMA CRISPR/CAS9**

Tutor: Prof. Pietro Benedetti
Dipartimento di Biologia

Laureando: Riccardo Carangelo

Anno Accademico 2016/2017

Indice

Introduzione	1
Struttura ed elementi di base del sistema CRISPR/Cas	2
Funzionamento generale del sistema CRISPR/Cas	4
Adattamento	4
Biogenesi.....	4
Interferenza	5
Il sistema CRISPR/Cas9	6
Struttura.....	6
Funzionamento.....	7
CRISPR/Cas9 per il genome editing.....	9
Ingegnerizzazione dell'RNA guida.....	12
Ingegnerizzazione di Cas9	12
Applicazioni di dCas9 e spCas9.....	12
Gestione di dCas9 e spCas9	13
Metodologie di controllo chimico di Cas9.....	13
Metodologie di controllo optogenetico di Cas9	15
Applicazioni e studi futuri.....	17
Implicazioni etiche di CRISPR/Cas9	18
Bibliografia	21

Introduzione

Nell'insieme di scoperte scientifiche rivoluzionarie avvenute nel corso della storia spicca con distinzione il contributo degli studi relativi a CRISPR. Questo meccanismo biologico è stato protagonista negli ultimi anni di una rivoluzione caratterizzata da una rapidissima ascesa. Le sue potenzialità nell'ambito della ricerca medica e molecolare, così come gli importanti dibattiti bioetici sollevati, hanno reso CRISPR un caso mediatico rilevante.

La storia di CRISPR (acronimo di Clustered Regularly Interspaced Short Palindromic Repeats) ha avuto inizio nel 1993, quando vennero scoperte per caso delle strutture peculiari mai osservate fino a quel momento: copie multiple di una specifica sequenza separate da spaziatori. Solo a seguito di un lungo studio operato da più gruppi di ricerca, fu possibile identificare le misteriose strutture in molti organismi filogeneticamente distanti tra loro, fatto quest'ultimo che ne confermava l'importanza funzionale.

Nella ricerca di questa fantomatica importante funzione furono presi in analisi gli spaziatori, decifrandone la funzione e riuscendo a dimostrare che erano implicati in una sorta di ancestrale "meccanismo immunitario" nei procarioti. La portata della scoperta rese famoso questo peculiare meccanismo, al punto che un numero sempre maggiore di ricercatori iniziava ad approcciarsi al suo studio, mettendone sempre più in luce l'ampia gamma di situazioni alle quali esso è applicabile.

Dopo poco tempo dalla sua scoperta, la fama di CRISPR andò quindi incontro ad una rapida ascesa, dando il via ad un periodo di studio mirato ad una maggiore caratterizzazione del suo funzionamento e delle strutture che lo riguardano. Ciò che emerse da questi studi risultò essere un meccanismo intracellulare di precisione straordinaria, dotato di enormi potenzialità applicative in molti campi.

Un sempre più folto gruppo di ricercatori da ogni parte del mondo indirizza la propria ricerca nello studio di CRISPR. Quello che 20 anni prima era un oscuro meccanismo microbico scoperto in Spagna per puro caso, si è ormai tramutato in un punto di riferimento per le riviste scientifiche e divulgative, per i congressi e le conferenze, una vera e propria rivoluzione per la ricerca scientifica e per il progresso.

Struttura ed elementi di base del sistema CRISPR/Cas

Nel corso degli anni successivi alla scoperta di CRISPR, è stato possibile osservare un incremento sempre maggiore del numero di studi mirati a comprendere quali fossero gli elementi protagonisti del meccanismo associato alle sequenze CRISPR e come questi elementi fossero strutturati e tra loro organizzati (Lander, 2016). Lo studio effettuato su una serie di organismi batterici e di archei, ha condotto alla scoperta di varie tipologie di sistemi associati a CRISPR, che si differenziano tra loro, presentando peculiarità strutturali e funzionali più o meno marcate (Jiang & Doudna, 2015). Nonostante l'inaspettata e notevole diversità osservata, tutti i sistemi finora analizzati hanno rivelato la presenza di caratteristiche di base che risultano invarianti nei confronti del sistema a cui si fa riferimento. In questo capitolo viene descritta la struttura di base della sequenza CRISPR e vengono inoltre presentate le componenti basilari del sistema CRISPR/Cas, condivise da ogni variante finora scoperta (Hille & Charpentier, 2016).

In generale, il sistema CRISPR/Cas, come indica il nome stesso, consiste di un locus CRISPR e di una serie di proteine Cas (acronimo di CRISPR – associated system) (Jiang & Doudna, 2015).

Il locus CRISPR è composto da tre componenti principali (Jiang & Doudna, 2015) (*Figura 1*):

- Una sequenza CRISPR;
- Un operone cas;
- Una sequenza leader ricca in A/T.

La sequenza CRISPR produce RNA non codificanti (Lander, 2016) e al suo interno è possibile individuare due tipologie di elementi (Barrangou & Marraffini, 2015):

- Segmenti di DNA costituiti da brevi sequenze ripetute, lunghi solitamente da 28 a 37 paia di basi;
- Segmenti di DNA detti spaziatori, lunghi solitamente da 32 a 38 paia di basi.

Le sequenze ripetute e gli spaziatori si organizzano tra loro secondo un preciso schema di alternanza, pertanto tra due sequenze ripetute figura sempre uno spaziatore (Hille & Charpentier, 2016) (*Figura 1*).

Gli spaziatori rappresentano l'elemento chiave per comprendere la funzione del sistema CRISPR/Cas. Difatti ogni spaziatore è dotato di un corrispettivo all'interno di un preciso fago o di un preciso plasmide (Bolotin *et al.*, 2005; Mojica *et al.*, 2005). Tale corrispettivo possiede la medesima sequenza dello spaziatore dato e prende il nome di protospaziatore. Questa perfetta corrispondenza di sequenze non è ovviamente casuale: ogni spaziatore visibile nella sequenza CRISPR è stato infatti prodotto grazie all'incorporazione del corrispettivo protospaziatore nella fase di primo contatto con l'agente infettante (Jiang & Doudna, 2015). Gli spaziatori incorporati nel genoma sono in forma inattiva e divengono funzionali a seguito di un processo di trascrizione e maturazione, alla fine del quale viene prodotto per ogni spaziatore un preciso crRNA (CRISPR RNA) non codificante, costituito dalla sequenza completa di uno spaziatore fiancheggiata da frammenti provenienti dalle

sequenze ripetute (Jiang & Doudna, 2015). Il crRNA lega specifiche proteine Cas, attivandole e guidandole verso un preciso protospaziatore (Jiang & Doudna, 2015). L'organismo è poi in grado di attuare meccanismi di vario genere per distinguere spaziatori (self) e protospaziatori (non self) (Hille & Charpentier, 2016).

In prossimità della sequenza CRISPR è presente una serie di geni (*Figura 1*) che costituiscono l'operone cas (CRISPR – associated system) e codificano due tipologie di enzimi (Hille & Charpentier, 2016):

- Enzimi in grado di tagliare in maniera mirata e precisa DNA o RNA esogeni, grazie alla loro attività catalitica e al crRNA che funge da attivatore enzimatico e da guida nel localizzare la sequenza corretta all'interno del materiale genetico esogeno;
- Enzimi che, in occasione del primo contatto con un agente infettante, sono in grado di individuare la sequenza del protospaziatore, escidendolo ed incorporandolo permanentemente nel genoma procariotico, in modo da consentire il riconoscimento del medesimo agente nei successivi casi di infezione.

In funzione dell'organismo considerato, l'operone cas può presentare variazioni in merito al proprio set di geni e alla loro organizzazione. Esistono infatti ben 93 geni cas raggruppati in 35 famiglie diverse (Makaraova *et al.*, 2015). Tra due sistemi considerati i geni cas possono essere tutti completamente diversi, ma vi sono 10 geni (da cas1 a cas10) che compongono il nucleo cas. Ogni sistema CRISPR/Cas completo possiede almeno uno dei geni inclusi nel nucleo cas (Makaraova *et al.*, 2015). Come si vedrà nel prossimo capitolo, l'operone cas risulta essere il discriminante per classificare due qualsiasi sistemi CRISPR/Cas a confronto, poiché codifica varie tipologie di enzimi, influenzando drasticamente i meccanismi di funzionamento per ogni tipologia di sistema considerata.

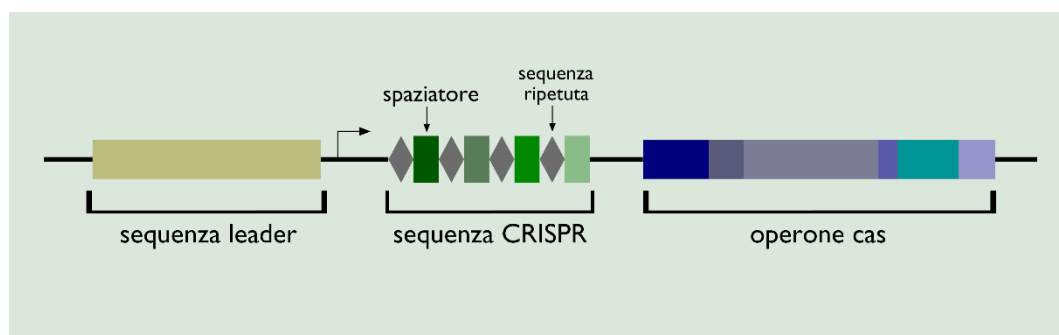


Figura 1 – organizzazione generale di un locus CRISPR. In varie tonalità di blu/viola i geni cas dell'operone; nella sequenza CRISPR si possono osservare gli spaziatori in varie tonalità di verde che ne indicano la diversità (ogni spaziatore fa riferimento ad un preciso agente infettante); dello stesso colore figurano invece le brevi sequenze ripetute che si alternano agli spaziatori; a monte si osserva, in beige, una sequenza leader ricca in AT. L'ordine delle tre strutture può variare di locus in locus.

Funzionamento generale del sistema CRISPR/Cas

Il sistema CRISPR/Cas presenta un funzionamento analogo a quello della RNAi negli eucarioti. Esso sfrutta una o più proteine in grado di individuare DNA e RNA esogeni ed effettuare su di essi precisi tagli endonucleolitici, agendo come enzimi di restrizione programmabili (Marraffini & Sontheimer, 2008) ed utilizzando specifici crRNA come guida per il riconoscimento di particolari sequenze (Brouns *et al.*, 2008). Il meccanismo operativo di CRISPR/Cas può essere riassunto in tre tappe distinte, presentate in ordine cronologico (Hille & Charpentier, 2016):

1. Adattamento, fase di acquisizione degli spaziatori;
2. Biogenesi, fase in cui avviene l'espressione della sequenza CRISPR;
3. Interferenza, fase in cui ha luogo l'effettiva interferenza con il materiale genetico esogeno.

Di seguito vengono presentate e caratterizzate le singole tappe (opportunamente riassunte nel diagramma in *Figura 2*).

Adattamento

In questa fase iniziale avvengono il primo contatto con uno specifico agente infettante e la conseguente acquisizione dei suoi protospaziatori, i quali vengono accorpati alla sequenza CRISPR, ampliandone il pool (Hille & Charpentier, 2016). Tra le varie tipologie di sistemi CRISPR/Cas figurano varie strategie tra loro differenti per consentire al procariote la distinzione delle sequenze appartenenti agli spaziatori (costituenti il self), dalle medesime sequenze appartenenti ai protospaziatori (costituenti il non-self) (Hille & Charpentier, 2016). Lo spaziatore in sé non può essere utilizzato per discriminare self e non-self, in quanto possiede una sequenza pienamente complementare al corrispondente protospaziatore, dal quale proviene. Il processo di adattamento consente pertanto all'organismo procariote di memorizzare informazioni sull'agente infettante, immagazzinandone parte del materiale genetico sotto forma di spaziatore, e mette in evidenza la natura adattativa dell'immunità mediata da CRISPR/Cas. (Barrangou *et al.*, 2017).

Ogni tipologia di sistema CRISPR/Cas possiede un suo specifico pool di geni che generano le proteine utili al processo integrativo dei protospaziatori. Tuttavia vi è una coppia di geni finora riscontrati quasi in ogni sistema preso in analisi: cas1 e cas2, responsabili rispettivamente della codifica delle proteine Cas1 e Cas2. Queste ultime svolgono un ruolo fondamentale nel processo di acquisizione degli spaziatori (Hille & Charpentier, 2016).

Biogenesi

In questa fase avviene la trascrizione della sequenza CRISPR contenente tutti gli spaziatori acquisiti durante le varie fasi di adattamento. La trascrizione produce un lungo filamento di pre-crRNA. Questo filamento, non ancora maturo e inattivo, necessita d'essere processato. Il processamento prevede una serie di tagli endonucleolitici atti a separare tra loro gli spaziatori (Haurwitz *et al.*, 2017). Il processo di separazione genera una serie di filamenti di crRNA maturo. Ogni filamento contiene un singolo spaziatore e diviene disponibile al legame con la proteina Cas specifica.

Interferenza

Questa ultima fase comprende più tappe. È presente anzitutto una fase di legame del crRNA maturo con una proteina Cas specifica. Questo assemblaggio porta alla formazione del complesso ribonucleoproteico crRNP (acronimo di CRISPR Ribo-Nucleoprotein Particle), il quale è in grado di intercettare il giusto sito di taglio sequenza-specifico utilizzando il crRNA come guida (van der Oost *et al.*, 2014). Risulta qui evidente l'importanza del crRNA al fine di attivare e guidare la relativa proteina Cas verso la sequenza di uno specifico agente infettante. Il complesso crRNP possiede ora la capacità di distinguere il non-self e, in particolare, di riconoscere le sequenze infettanti di interesse, alle quali il crRNA risulta complementare. Questo riconoscimento è seguito da una fase di ibridazione del protospaziatore con il corrispettivo crRNA del complesso ribonucleoproteico. L'ibridazione consente la formazione di un R-loop che induce un cambiamento conformazionale nella componente proteica del complesso crRNP. Questa transizione strutturale funge da input per il reclutamento di nucleasi specifiche (nel caso in cui il complesso crRNP non sia dotato di attività nucleasica) o per l'attivazione della componente nucleasica intrinseca (qualora essa sia già presente all'interno del complesso crRNP) (van der Oost *et al.*, 2014).

Per ogni tipologia di sistema CRISPR/Cas vi sono varie tipologie di proteine Cas che operano nella fase di interferenza. Possono presentarsi più proteine che lavorano sinergicamente oppure una sola grande proteina dotata di più caratteristiche funzionali (Jiang & Doudna, 2015).

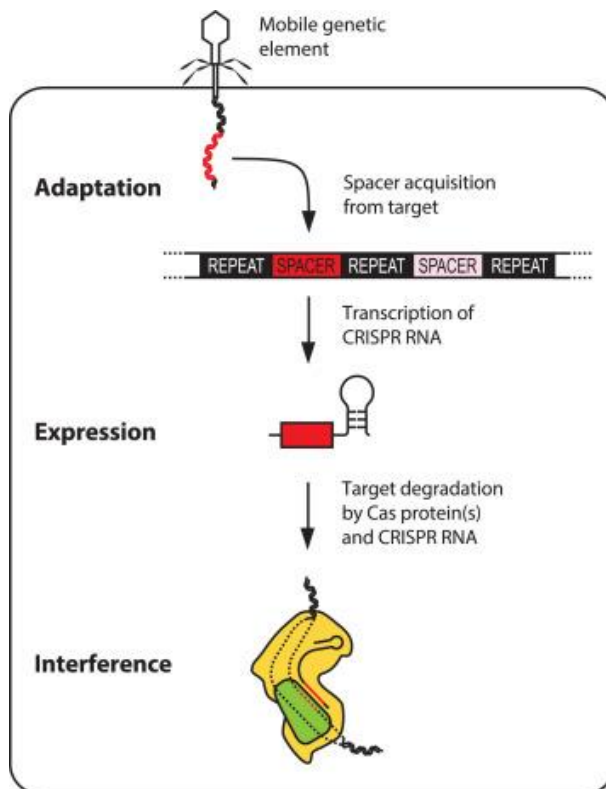


Figura 2 – schema riassuntivo del meccanismo di funzionamento di un sistema CRISPR/Cas. Si possono osservare la fase di adattamento (in alto), nella quale un protospaziatore (in rosso) viene incluso nella sequenza CRISPR, la fase di biogenesi (al centro), in cui si ha la trascrizione e la maturazione del crRNA, e la fase di interferenza, nella quale una specifica proteina Cas (in giallo), si serve di uno specifico filamento di crRNA per effettuare un taglio in uno specifico punto del DNA esogeno (Rath *et al.*, 2015).

Il sistema CRISPR/Cas9

In questo capitolo viene presentato uno specifico sistema CRISPR/Cas, il sistema CRISPR/Cas9, particolarmente importante per via dei promettenti risultati applicativi da esso forniti in merito al genome editing (Doudna & Charpentier, 2014). CRISPR/Cas9, pur possedendo le proprie peculiarità strutturali e funzionali, risponde a tutte le caratteristiche di ogni generico sistema CRISPR/Cas illustrate nei due capitoli precedenti. In questo capitolo si parlerà pertanto delle peculiarità strutturali e funzionali di questo meccanismo.

Struttura

A livello genomico, oltre alle canoniche strutture presenti in ogni sistema CRISPR/Cas (sequenza leader, operone cas e sequenza CRISPR), si osserva la presenza di una sequenza peculiare, responsabile della trascrizione del tracrRNA (trans-encoded small RNA), un RNA non codificante che può avere svariate lunghezze e che risulta essenziale nei processi di biogenesi e di interferenza (Deltcheva *et al.*, 2011).

L'operone cas è dotato di un preciso pool di geni che contraddistinguono CRISPR/Cas9 dagli altri sistemi. Tipicamente l'operone è piuttosto minimale e si distinguono solamente quattro geni: cas9, cas1, cas2 e csn2. Il primo è responsabile della codifica della proteina Cas9, protagonista nel processo di interferenza, gli altri sono invece coinvolti nel processo di acquisizione degli spaziatori durante la fase di adattamento (Jiang & Doudna, 2017).

I vari sistemi CRISPR/Cas possiedono, come già spiegato nel capitolo sul funzionamento generale, meccanismi di vario genere per discriminare le sequenze appartenenti ad uno spaziatore dalle sequenze appartenenti ad un protospaziatore. Nel caso di CRISPR/Cas9 entrano in gioco delle particolari sequenze dette PAM (acronimo di Protospacer Adjacent Motif). Una sequenza PAM è lunga dalle due alle cinque paia di basi ed è posizionata all'estremità 3' della sequenza del protospaziatore, ma è assente negli spaziatori. Pertanto è possibile reperire le sequenze PAM soltanto all'interno del genoma dell'agente infettante e non all'interno del genoma del procariote. Questa esclusiva presenza all'interno del solo genoma infettante consente al procariote l'essenziale discriminazione tra spaziatori e protospaziatori, utile sia durante il processo di adattamento di un protospaziatore, sia per attuare il processo di interferenza (Shah *et al.*, 2013) (Jiang & Doudna, 2017).

L'enzima Cas9 comprende tipicamente una serie di domini e svolge il ruolo di una nucleasi polifunzionale (Jiang & Doudna, 2017). Esso è in grado di tagliare il dsDNA esogeno esattamente 3 paia di basi a monte della sequenza PAM mediante l'utilizzo di due distinti domini nucleasici (Jiang & Doudna, 2017):

- Un dominio nucleasico di tipo HNH che taglia la porzione di DNA complementare alla sequenza del RNA guida;
- Un dominio nucleasico di tipo RuvC responsabile del taglio della porzione di DNA opposta a quella complementare.

L'apoenzima Cas9 presenta due strutture lobate distinte strettamente connesse da due segmenti di adesione:

- Il lobo REC (recognition lobe) (*Figura 4*), contenente l'alfa elica di riconoscimento e composto da tre domini Hel (Helical domain) ad alfa elica (Hel-I, Hel-II e Hel-III) (Jiang & Doudna, 2017);
- Il lobo NUC (nuclease lobe) (*Figura 4*), contenente la porzione responsabile dell'attività nucleasica di Cas9 e dove è possibile riscontrare i due domini HNH e RuvC (Jiang & Doudna, 2017) e la porzione responsabile dell'interazione con la sequenza PAM (Nishimasu *et al.*, 2014).

Funzionamento

Oltre a partecipare alla fase di interferenza, l'enzima Cas9 partecipa anche alla biogenesi (contribuendo alla maturazione del crRNA) e alla fase di adattamento (collaborando con le altre proteine di acquisizione) (Jiang & Doudna, 2017). Nonostante Cas9 ricopra più ruoli funzionali, il suo ruolo primario e più rilevante rimane quello dell'enzima di interferenza.

Durante i processi di biogenesi e interferenza il tracrRNA svolge ruoli fondamentali. Esso è infatti implicato nel processo di maturazione del crRNA (Jiang & Doudna, 2015) e risulta essenziale per l'interferenza mediata dalla proteina Cas9, in quanto consente a quest'ultima di legarsi con il crRNA per attivarsi. Al processo di maturazione contribuisce, mediante tagli endonucleolitici sul pre-crRNA, la RNasi III. Una volta maturo, il crRNA rimane legato al tracrRNA ed inizia la sua fase di assemblaggio con l'apoenzima Cas9 (Karvelis *et al.*, 2013). Durante questo assemblaggio, la proteina Cas9 va incontro a cambiamenti conformazionali al fine di giungere ad uno stato funzionale attivo per acquisire la possibilità di riconoscere il DNA esogeno (anche grazie alla specificità fornita dal crRNA). (Jiang & Doudna, 2017). Il legame con l'RNA guida corrisponde al cambiamento conformazionale più rilevante di Cas9 e segna il passaggio tra lo stato inattivo e lo stato attivo dell'enzima; l'entità di questo cambio conformazionale dimostra che l'RNA guida costituisce un regolatore chiave per la funzione di Cas9 (Jiang & Doudna, 2017). Il cambiamento conformazionale più rilevante avviene a livello del lobo REC, a seguito del legame con l'RNA guida. Si ha in particolare che Hel-III si muove di 65 Å verso il dominio HNH, rendendo Cas9 in grado di riconoscere la sequenza PAM (Jiang & Doudna, 2017). Questo segna la fine dell'assemblaggio di quello che prende il nome di complesso di sorveglianza, una struttura ternaria composta da Cas9, dal crRNA e dal tracrRNA. Questa struttura, finalmente attiva, è in grado di intercettare uno specifico protospaziatore per eseguire la sua funzione di interferenza (Jiang & Doudna, 2017).

Il tracrRNA possiede un ruolo anche nel processo di interferenza. In questa fase il lobo REC è essenziale per il legame con il DNA target e con l'RNA guida. Quando il complesso di sorveglianza entra in contatto con la sequenza PAM e con il protospaziatore si osservano dei cambiamenti conformazionali che sono meno rilevanti rispetto a quello che avviene a seguito del legame tra Cas9 e l'RNA guida (Jiang & Doudna, 2015; 2017) (Nishimasu *et al.*, 2014).

Una volta raggiunto il sito di interferenza, il complesso di sorveglianza interagisce in prima istanza con la sequenza PAM, creando un saldo legame (Jiang & Doudna, 2015). Dopo essersi legato ad una sequenza PAM, il complesso attraversa una fase di analisi della sequenza fiancheggiante per valutare la presenza di complementarità con l'RNA guida (Jiang & Doudna, 2015). In caso di complementarità, avverrà il legame del complesso di sorveglianza con il DNA del protospaziatore, con un ulteriore cambio di conformazione di Cas9 (Jiang & Doudna, 2015). Il riconoscimento del protospaziatore corretto induce la formazione di un R-loop (*Figura 4*) che comporta uno svolgimento del DNA esogeno, con conseguente svincolamento dei due filamenti: il filamento complementare all'RNA guida forma un ibrido RNA-DNA, lasciando l'altro filamento libero; ultimata la formazione di questa struttura, è possibile attuare la catalisi operata dal complesso di sorveglianza, con conseguente taglio endonucleolitico del DNA infettante (Jiang *et al.*, 2016). L'intero funzionamento del sistema CRISPR/Cas9 è opportunamente riassunto in *Figura 3*.

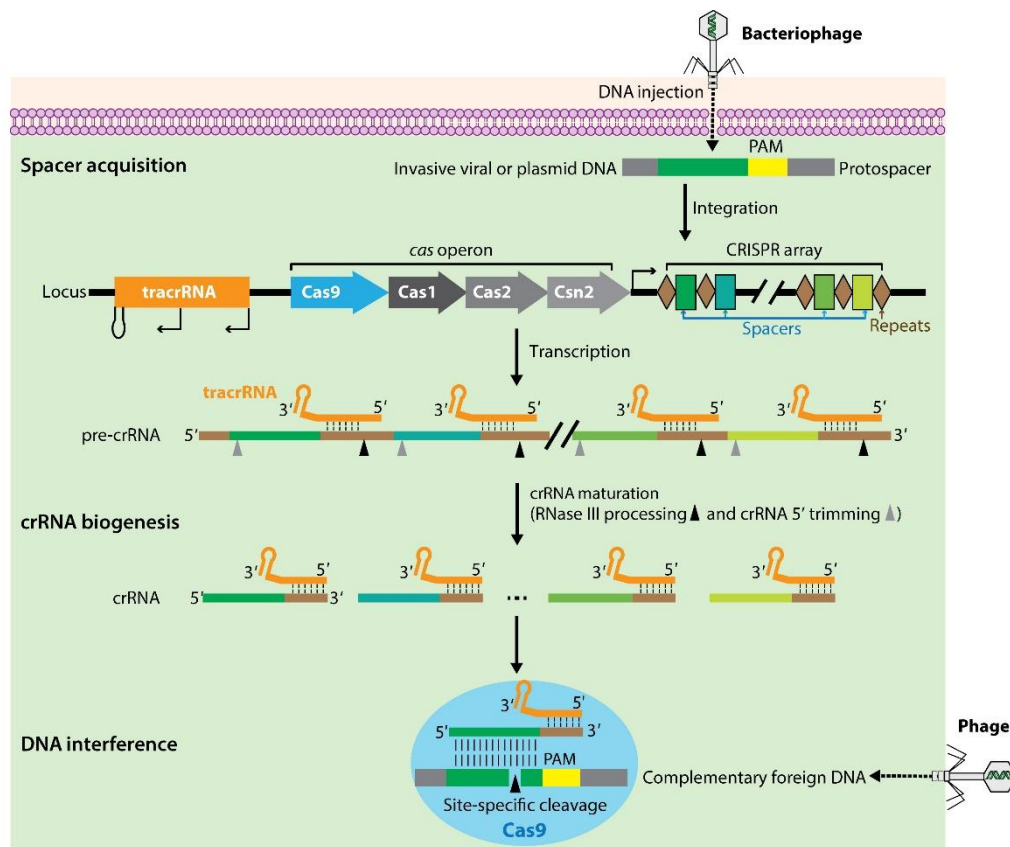


Figura 3 – schema riassuntivo del funzionamento di un sistema CRISPR/Cas9. Si osserva in alto la fase di adattamento, al centro la fase di biogenesi e in basso la fase di interferenza. In giallo è rappresentata la sequenza PAM, al cui fianco è presente, in verde, il protospaziatore d'interesse; in varie tonalità di verde e di forma rettangolare sono rappresentati gli spaziatori; si osservano l'operone *cas* affiancato dalla sequenza che codifica il *tracrRNA*. Nello schema è apprezzabile la struttura del pre-crRNA, il quale viene maturato, formando vari filamenti separati di crRNA. Infine, nella fase di interferenza, si osserva il taglio operato dal complesso di sorveglianza (Jiang & Doudna, 2015).

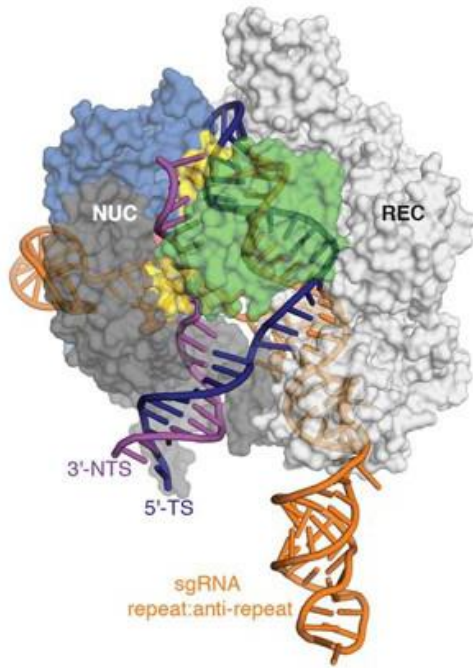


Figura 4 (a sinistra) – struttura cristallografica del complesso di sorveglianza legato al dsDNA esogeno. È possibile osservare il lobo NUC (a sinistra), il lobo REC (a destra) e il R-loop formato dal dsDNA esogeno e dal sgRNA con la sua architettura dettagliata; in particolare è possibile osservare l'RNA guida (in arancione) e i due filamenti del DNA esogeno svolti (uno blu, che si ibrida con l'RNA guida, ed uno viola, che rimane libero); di particolare interesse risulta l'osservazione della distorsione che il complesso di sorveglianza impone al dsDNA, distorcendone l'andamento con la formazione del R-loop (Jiang *et al.*, 2016).

CRISPR/Cas9 per il genome editing

La modifica sito-specifica del DNA è sempre risultata una questione complessa, a causa dell'intrinseca difficoltà nell'ottenere strumenti muniti dell'adeguata precisione nell'individuare il sito d'azione e nell'effettuare un taglio preciso. Con il procedere del progresso scientifico, l'editing a livello del genoma si è avvalso di un certo numero di tecniche col tempo sempre più efficienti. Piccole molecole appositamente scelte e introni dotati di self-splicing venivano utilizzati come strumenti precoci per introdurre modificazioni all'interno del genoma (Doudna & Charpentier, 2017). Solo recentemente è stata scoperta la maggiore efficacia degli enzimi per modificare il DNA. In particolare è stato di notevole impatto l'utilizzo di nucleasi in grado di produrre DSBs (acronimo di Double Strand Breaks) in punti specifici del genoma (Doudna & Charpentier, 2017). Prima dell'avvento di CRISPR/Cas9 vi erano tre principali famiglie di enzimi utilizzati per il genome editing:

- Meganucleasi, enzimi batterici in grado di identificare e tagliare in maniera molto specifica estese sequenze di DNA (da 14 a 40 paia di basi) (Stoddard, 2006);
- Zinc Finger Nucleases (ZFNs), enzimi di restrizione artificiali generati dalla fusione di domini nucleasici (in grado di tagliare il DNA) con domini muniti di motivi a dita di zinco (in grado di legare il DNA) (Carroll, 2011);
- Transcription Activator-Like Effector Nucleases (TALENs), enzimi batterici recentemente utilizzati per l'editing genomico e il controllo trascrizionale; si sono dimostrati particolarmente pratici da ingegnerizzare (Nemudryi *et al.*, 2014).

Le nucleasi sono in grado di individuare specifici siti all'interno del genoma ed effettuare tagli precisi in corrispondenza di tali siti. Grazie a questi tagli è possibile

aprire la strada a due meccanismi di riparazione, che producono mutazioni mirate (Nuñez *et al.*, 2016):

- Non-Homologous End Joining (NHEJ) (*Figura 5*), in cui il taglio viene riparato e le estremità vengono riunite senza che sia necessaria la presenza di un filamento di DNA donatore omologo; in questo caso è possibile che il meccanismo produca inserzioni o delezioni in corrispondenza del taglio riparato;
- Homology Directed Repair (HDR) (*Figura 5*), sfrutta il DSB generato dall'attività della nucleasi per inserire nuovo DNA all'interno del genoma, pertanto necessita di un filamento donatore omologo, appositamente scelto per fiancheggiare le estremità libere originatesi dal taglio.

A differenza delle proteine Cas, le quali si affidano al crRNA, le nucleasi appartenenti alle famiglie sopra descritte sono già in possesso di tutte le informazioni necessarie per la loro attività e non si affidano pertanto a molecole o stimoli esterni per attivarsi e compiere la loro funzione. Per questo motivo ogni modifica effettuata per ingegnerizzare questi enzimi deve interessare in maniera diretta la proteina e risulta pertanto limitante, essendo una struttura proteica complessa da gestire (Doudna & Charpentier, 2017).

A differenza degli enzimi appena presentati, una proteina Cas è in grado di tagliare in maniera molto specifica frammenti di DNA sotto le istruzioni di un frammento di RNA guida generato dall'unione di crRNA e tracrRNA. Ingegnerizzando il suo RNA è possibile quindi indirizzare la proteina Cas verso un qualsiasi sito di taglio, senza dover agire direttamente su quest'ultima. Essendo l'RNA uno strumento molto pratico da ingegnerizzare, il sistema CRISPR/Cas risulta molto più efficace per il genome editing di qualsiasi altro strumento mai utilizzato, fornendo all'ingegneria genetica una tecnica veloce, versatile e relativamente semplice da utilizzare per modificare il DNA in maniera sito-specifica (Le Cong *et al.*, 2013) (Jinek *et al.*, 2013) (Mali *et al.*, 2013). Tra gli enzimi Cas si è rivelato di particolare utilità l'enzima Cas9 estratto da *S. pyogenes*. Esso è infatti l'enzima Cas maggiormente utilizzato per le manipolazioni genetiche (Nuñez *et al.*, 2016).

Nei successivi due capitoli saranno trattate in maniera più specifica l'ingegnerizzazione dell'RNA guida di Cas9 e l'ingegnerizzazione di Cas9.

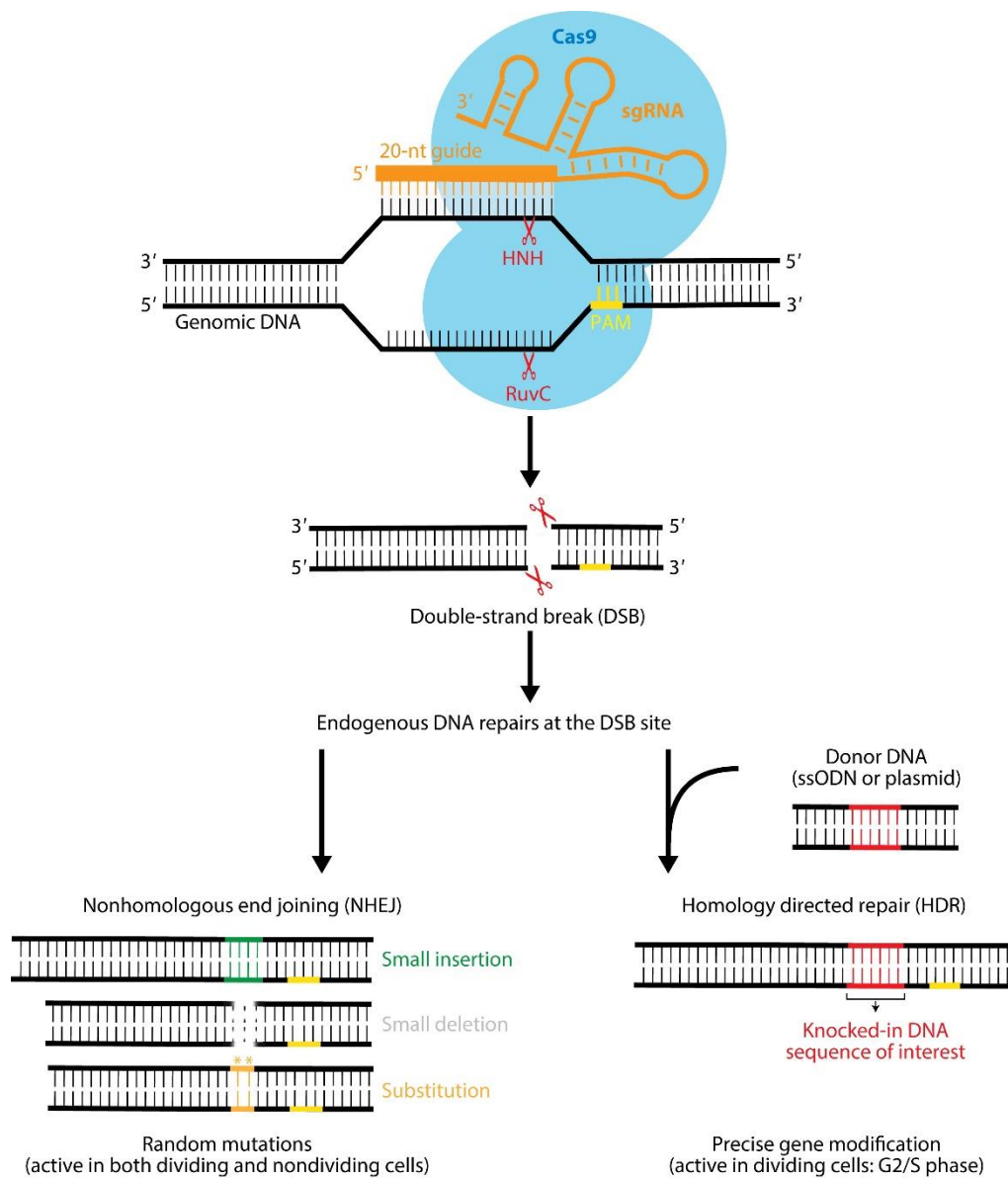


Figura 5 – meccanismo di taglio operato da Cas9 sotto la guida dell'sgRNA. Si osserva come questo evento porti alla possibilità di effettuare riparazioni di due tipologie differenti. La via NHEJ genera inserzioni, delezioni e sostituzioni di piccolo calibro, mentre la via HDR consente l'inserimento di sequenze di maggiori dimensioni, sotto la guida di uno specifico filamento donatore (Jiang & Doudna, 2015).

Ingegnerizzazione dell'RNA guida

L'RNA guida costituisce lo strumento base per attivare ed indirizzare Cas9 e rappresenta quindi l'oggetto di interesse primario nella fase di manipolazione del sistema CRISPR/Cas9. L'RNA guida è stato ottimizzato mediante processi di ingegnerizzazione, producendo un sgRNA (acronimo di single guide RNA). Esso corrisponde ad un singolo filamento di RNA formato da due strutture (Doudna & Charpentier, 2017):

- Una sequenza all'estremità 5' che individua e lega il sito di taglio sul DNA mediante un appaiamento di tipo Watson – Crick;
- Una sequenza all'estremità 3' che lega Cas9.

La produzione di sgRNA viene effettuata semplicemente legando l'estremità 3' del crRNA all'estremità 5' del tracrRNA, creando quindi una struttura molto simile all'ibrido naturale tracrRNA:crRNA. L'sgRNA, in termini di efficienza, consente di ottenere risultati analoghi a quelli della sua variante naturale, pertanto risulta potenzialmente in grado di guidare efficientemente Cas9 verso qualsiasi sito di taglio permettendo la creazione di un meccanismo efficace, versatile ed ampiamente programmabile (Jinek *et al.*, 2017).

Ingegnerizzazione di Cas9

Per ottemperare a specifici obiettivi di ricerca o al fine di ottenere una versione maggiormente efficace di Cas9, sono state apportate modificazioni specifiche alla struttura intrinseca di questo enzima. Ad oggi si possono osservare due principali versioni ingegnerizzate di Cas9:

- L'enzima dCas9 (deactivated Cas9), ovvero una variante ingegnerizzata di Cas9 che risulta cataliticamente inattiva da un punto di vista nucleasico, ma in grado comunque di individuare specifiche sequenze di DNA se dotata di un RNA guida. L'inattivazione è operata mediante mutazioni puntiformi che impediscono l'attività nucleasica tipica di Cas9 (Qi *et al.*, 2013);
- L'enzima spCas9 (split-Cas9), generato dalla separazione dei lobi REC e NUC di Cas9. Un vantaggio offerto dalla naturale struttura di Cas9 sta infatti nella separazione delle attività di riconoscimento e taglio sui due diversi lobi. Ciò consente, mediante separazione dei due lobi, di controllarne l'attività. A questo proposito è stata pertanto creata una versione a lobi separati di Cas9, detta spCas9, concepita allo scopo di indurre l'attività di Cas9 solo mediante unione controllata dei due lobi (Nuñez *et al.*, 2016).

Gli utilizzi e i metodi per la gestione delle due versioni ingegnerizzate di Cas9 saranno esposti nei successivi capitoli.

Applicazioni di dCas9 e spCas9

Come è stato appena visto, dCas9 è stato privato di una parte della sua funzionalità, pertanto non può essere utilizzato da solo per l'editing del genoma. L'utilizzo di dCas9 è alla base delle tecniche denominate CRISPRi (CRISPR interference) e CRISPRa (CRISPR activation), nelle quali l'enzima dCas9 viene utilizzato per la regolazione dell'espressione genica, agendo direttamente sulla trascrizione genica.

CRISPRi opera mediante inibizione reversibile della trascrizione. Tale inibizione viene attuata mediante la produzione di un blocco sterico che arresta la fase di allungamento trascrizionale operata dalla RNA polimerasi; ne risulta la repressione del gene di interesse. La tecnica CRISPRi viene attuata mediante la produzione di dCas9 chimeriche, ovvero fuse a domini regolatori, i quali fungono da repressori nei confronti dei siti individuati da dCas9. La tecnica è applicabile a livello di cellule di organismi eucarioti e procarioti (Qi *et al.*, 2013) (Larson *et al.*, 2013).

CRISPRa, al contrario, opera mediante induzione della trascrizione, grazie alla produzione di dCas9 chimeriche, cioè fuse a domini regolatori che fungono da attivatori trascrizionali nei confronti dei siti intercettati da dCas9 (Dominiguez *et al.*, 2016).

Le proprietà di intercettazione di dCas9 non sono applicate solamente alla CRISPRi, ma possono essere utilizzate per individuare specifici loci in cellule vive, mediante la tecnica di dynamic imaging applicata al genoma, al fine di condurre studi sul comportamento di specifici siti in determinati intervalli temporali (Chen *et al.*, 2013).

Mentre dCas9 è stato privato di parte della sua funzione, spCas9, una volta uniti i due lobi REC e NUC, è in grado di espletare a pieno la propria attività, pertanto gli utilizzi di spCas9 sono mirati al vero e proprio editing genomico.

Gestione di dCas9 e spCas9

Una delle principali questioni tecniche legate ad un utilizzo esteso di CRISPR/Cas9 all'interno di un organismo riguarda il problema della specificità di Cas9 nell'individuare il corretto sito di taglio. Le attività di taglio che avvengono fuori dal corretto sito d'azione possono risultare pericolose per il genoma e rendono Cas9 uno strumento ancora inadatto ad un efficiente editing del genoma (Nuñez *et al.*, 2016). Determinati studi hanno dimostrato che con l'innalzarsi dei livelli di espressione di Cas9 si ha anche un innalzamento della quantità di tagli fuori bersaglio (Hsu *et al.*, 2013). Al fine di gestire i livelli di espressione di Cas9 sono stati sviluppati dei sistemi per la modulazione dell'attività nelle cellule di mammifero, mediante utilizzo di metodi chimici ed ottici in grado di inibire o indurre Cas9 o le sue forme ingegnerizzate.

Metodologie di controllo chimico di Cas9

Induzione di spCas9 mediante inteine

In questa tecnica vengono sfruttate delle inteine, proteine naturali in grado di effettuare splicing, che vengono legate ai due lobi di spCas9. Le due inteine presenti sui due lobi di spCas9 risultano complementari e sono in grado di riconoscersi, unirsi tra loro e distaccarsi da spCas9, comportando la formazione di una Cas9 integra e funzionale. Questo distaccamento, così come l'unione dei due lobi di spCas9, è permesso grazie alle proprietà di splicing delle due inteine. I risultati ottenuti con questa tecnica hanno efficacia paragonabile a quella di una Cas9 non ingegnerizzata (Shah & Muir, 2014) (*Figura 6*).

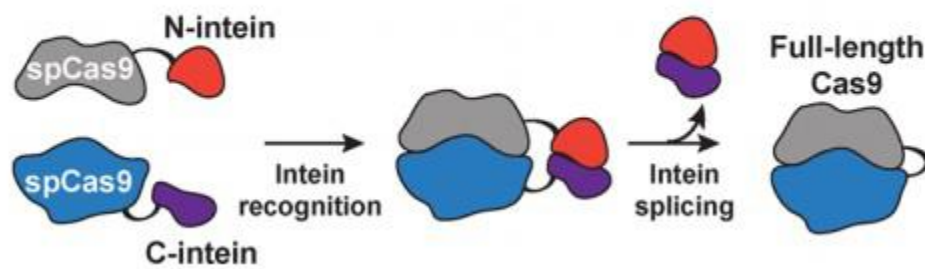


Figura 6 – percorso di attivazione di una SpCas9 mediante l'utilizzo di inteine, si possono osservare i due lobi (REC e NUC) di spCas9 separati, la fase di riconoscimento delle due inteine, la loro unione e l'allontanamento mediante splicing, che permette ai due lobi di condensarsi e formare così una Cas9 attiva (Nuñez *et al.*, 2016).

Induzione di spCas9 mediante rapamicina

In questa tecnica viene utilizzata la rapamicina, un macrolide. A livello strutturale si ha che:

- Uno dei due lobi viene legato ad una proteina FKBP (FK506 binding protein 12) e ad una sequenza NLS (Nuclear Localization Sequence); tutta la struttura viene mantenuta al di fuori del nucleo, a livello del citoplasma;
- L'altro lobo viene legato ad una proteina FRB (FKBP Rapamycin Binding protein) e ad una sequenza NES (Nuclear Export Sequence); tutta la struttura viene mantenuta all'interno del nucleo.

La rapamicina funge da attivatore e viene quindi utilizzata nel momento in cui è necessario indurre l'attività di Cas9. La rapamicina è in grado di legare la proteina FRB e di indurre la dimerizzazione di FKBP e FRB. Tramite questa dimerizzazione si ha l'avvicinamento dei due lobi di spCas9 e la conseguente attivazione di quest'ultima. Le due sequenze amminoacidiche NEL e NLS servono per mantenere separati i due lobi (NES mantiene un lobo all'esterno del nucleo, NLS mantiene l'altro lobo all'interno del nucleo) (Nuñez *et al.*, 2016) (Banaszynski *et al.*, 2004) (Figura 7).

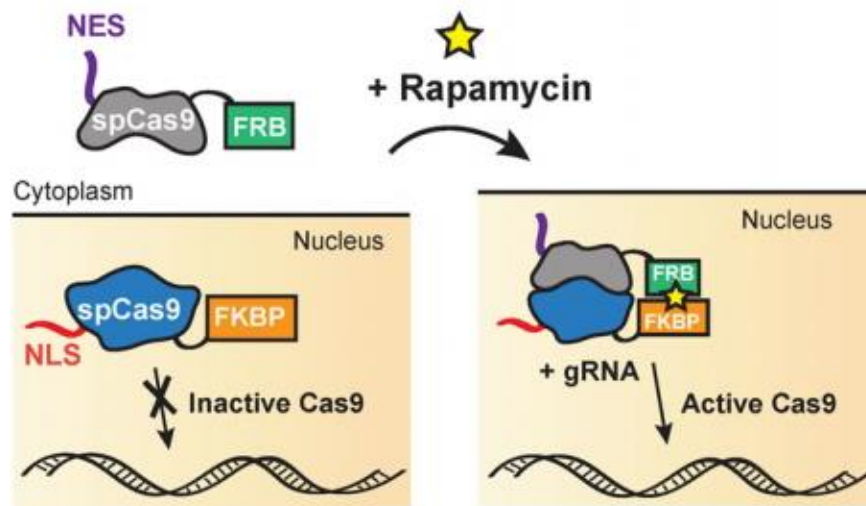


Figura 7 – percorso di attivazione di una SpCas9 mediante l'utilizzo di rapamicina e proteine DKBP e FRB, si osservano i due lobi (REC e NUC) di spCas9 separati e legati alle strutture supplementari; si osserva inoltre la fase di aggiunta della rapamicina, con conseguente importazione nel nucleo del lobo di SpCas9 contenente la struttura FRB-NES e la sua unione con il lobo di SpCas9 contenente la struttura FKBP-NLS; ne risulta una forma attiva di Cas9 contenente le due sequenze NES e NLS e le due proteine ormai dimerizzate FRB e FKBP (Nuñez *et al.*, 2016).

Metodologie di controllo optogenetico di Cas9

Le metodologie di controllo chimico di Cas9 presentano delle limitazioni, dovute all'irreversibilità dell'innescio e ad una limitazione intrinseca nella possibilità di gestire questa tipologia di tecnica. A tal proposito sono state create nuove metodologie basate sull'optogenetica. Le tecniche di induzione ottica che ne sono derivate risultano reversibili, più raffinate e maggiormente controllabili.

Induzione ottica di spCas9

Questa tecnica di induzione ottica sfrutta spCas9, mediante opportune procedure di ingegnerizzazione. Determinati studi sono stati in grado di sviluppare una paCas9 (photoactivable Cas9), ovvero una forma di spCas9 indotta dalla luce ed inibita dall'assenza di luce (entrambi i processi sono reversibili). Questa tecnica viene attuata mediante l'utilizzo di domini proteici fotoinducibili detti Mag (Magnets) in grado di formare dimeri se esposti ad una luce blu non nociva. Ognuno dei due lobi di spCas9 è legato ad uno di questi domini Mag (i due domini risultano complementari), i quali dimerizzano grazie ad interazioni elettrostatiche sotto l'effetto della luce, per poi distaccarsi in condizioni di carenza di luce. Il meccanismo di induzione ottica si è dimostrato efficace in una quantità significativa di casi (Figura 8) (Nihongaki *et al.*, 2015).

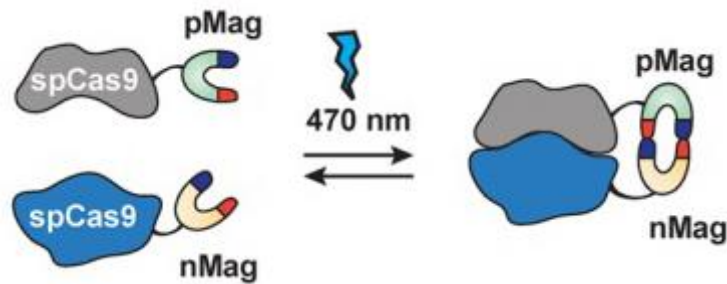


Figura 8 – percorso di attivazione di una SpCas9 mediante l'utilizzo di domini proteici Magnets, sono evidenziate la lunghezza d'onda della luce blu e la reversibilità della reazione (doppia freccia); nell'immagine è possibile osservare i due lobi di spCas9 recanti i domini Mag, sia separati (a sinistra) che in condizioni di unione (a destra), a seguito della dimerizzazione dei due domini Mag (Nuñez *et al.*, 2016).

Induzione ottica di dCas9

La tecnica di induzione ottica può essere applicata anche a dCas9 per indurre e controllare l'attivazione della trascrizione nell'ambito di CRISPRa. L'enzima dCas9 risulta legato ad una proteina CIB1 ed è posizionato in corrispondenza di uno specifico sito di taglio, in base all'istruzione dell'sgRNA. La struttura appena descritta non possiede l'attivatore trascrizionale. Esso viene infatti legato ad una proteina CRY2 libera. Quest'ultima è in grado di formare un dimero con CIB1 sotto induzione luminosa, portando l'attivatore a contatto con il sito a cui è legata dCas9 ed avviando così la trascrizione; l'intero processo risulta del tutto reversibile (Figura 9) (Nihongaki *et al.*, 2015).

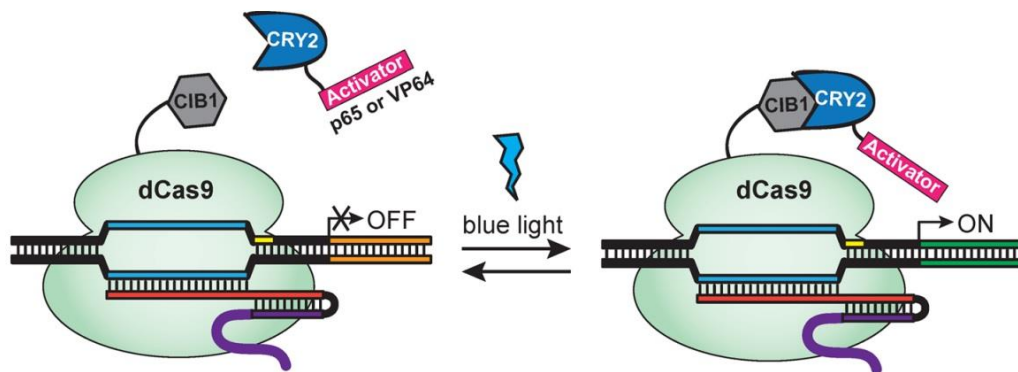


Figura 9 – percorso di attivazione di una dCas9 mediante l'utilizzo delle proteine dimerizzanti CIB1 e CRY2 (quest'ultima reca anche l'attivatore, costituito da un dominio p65 o da un dominio VP64); si osserva in basso l'sgRNA legato a dCas9 e il processo reversibile (doppia freccia) di attivazione operata da luce blu (Nuñez *et al.*, 2016).

Applicazioni e studi futuri

Le tecniche che la biologia molecolare ha sviluppato grazie al sistema CRISPR/Cas9 consentono una grande varietà di applicazioni funzionali in molti campi. Rilevante risulta la maggior parte dei risultati portati alla luce da studi effettuati mediante l'utilizzo di CRISPR/Cas9 su un certo numero di specie di eucarioti (esseri umani compresi); numerose risultano pertanto le prove sperimentali a supporto dell'efficacia di tale meccanismo (Hsu *et al.*, 2014). Le applicazioni in campo biologico del sistema vanno dalla ricerca di base fino alla ricerca biomedica (Doudna & Charpentier, 2017). I meccanismi del sistema CRISPR/Cas9 sono ancora da comprendere a pieno, tuttavia il suo potenziale e la sua efficienza porteranno senza dubbio ad una trasformazione della ricerca scientifica e industriale, così come porteranno probabilmente allo sviluppo di terapie molecolari efficaci per il trattamento delle malattie umane (Sander & Joung, 2014).

Uno dei campi in cui la tecnologia CRISPR/Cas9 risulta promettente è quello della terapia genica. Cas9 potrebbe essere utilizzata nel trattamento di malattie determinate da alleli recessivi o dominanti, al fine di correggere le mutazioni responsabili, così come potrebbe essere applicata per la correzione delle problematiche legate alla duplicazione genica, nelle quali Cas9 potrebbe essere sfruttata per la delezione di elementi duplicati (Hsu *et al.*, 2014).

Recenti studi hanno rivelato la possibilità di CRISPR/Cas9 di operare su un numero multiplo di geni, mediante inserimento, all'interno del plasmide da trasfettare, di più copie differenti di sgRNA. Questa caratteristica può risultare molto utile per una futura applicazione del meccanismo per lo studio e il trattamento di malattie poligeniche (ad esempio come il diabete, la schizofrenia e l'autismo) (Hsu *et al.*, 2014) (Le Cong *et al.*, 2013).

Per quanto riguarda le applicazioni terapeutiche, sussistono ancora delle problematiche relative al fenomeno dei tagli fuori bersaglio, i quali possono risultare in reazioni oncogene. Un'ulteriore complicazione risulta essere quella relativa alla risposta immunitaria (faccenda apparentemente risolvibile mediante l'utilizzo di immunosoppressori specifici da somministrare in corso di trattamento) (Mali *et al.*, 2013).

Un'interessante applicazione della tecnologia CRISPR/Cas9 sta nella generazione di modelli animali e cellulari. CRISPR/Cas9 ha permesso infatti di effettuare una efficiente modifica genomica in una grande varietà di tipologie cellulari rilevanti per la ricerca biomedica, così come ha consentito l'editing genomico all'interno di svariati organismi modello. La produzione di modelli cellulari può essere operata mediante introduzione nelle cellule, attraverso tecniche di trasfezione, di plasmidi contenenti Cas9 e sgRNA. In tal modo è possibile applicare tecniche di genome editing, così come è possibile applicare tecniche di interferenza (ad esempio CRISPRi e CRISPRa) e di imaging per loci specifici. Per quanto riguarda i modelli animali, Cas9 può essere direttamente iniettato all'interno degli zigoti, in modo da ottenere modifiche ereditabili del genoma in uno o più alleli. Un problema notevole riguardante l'iniezione di plasmidi di CRISPR/Cas9 all'interno di uno zigote sta nel

manifestarsi di fenomeni di mosaicismo genetico, i quali provocano la comparsa di un numero multiplo di linee genetiche diverse, espresse contemporaneamente all'interno dello stesso organismo, il tutto a causa di un'attività non omogenea di Cas9 all'interno delle cellule dell'organismo oggetto. Si tratta di un altro problema al quale la ricerca farà fronte con il perfezionamento del meccanismo (Hsu *et al.*, 2014) (Sander & Joung, 2014).

In particolar modo, CRISPRi e CRISPRa potranno essere applicate per la regolazione genica negli eucarioti e in particolare nei mammiferi. Tale meccanismo di regolazione risulterebbe notevolmente efficace, in quanto non presente negli eucarioti come meccanismo naturale di interferenza (a differenza della RNAi). Grazie a questa sua naturale assenza, l'interferenza mediante CRISPR è esente da eventuali fenomeni di competizione con i meccanismi endogeni di interferenza preesistenti all'interno degli organismi eucarioti (Doudna & Charpentier, 2017).

Per concludere, CRISPR/Cas9 si sta rivelando importante nella modifica del genoma di piante e funghi e potrebbe rivoluzionare il corso della ricerca agricola. I risultati finora emersi circa l'editing genomico mediante Cas9 su piante e funghi si stanno rivelando di particolare significatività. La ricerca è stata in grado di produrre modificazioni capaci di essere trasmesse alle generazioni successive, senza che vi siano livelli significativamente preoccupanti di tagli fuori bersaglio. Si avranno presto a portata di mano strumenti capaci di generare piante e funghi resistenti ai parassiti e alle malattie (Doudna & Charpentier, 2017) (Zhang *et al.*, 2014).

Implicazioni etiche di CRISPR/Cas9

La questione etica fondamentale che sta interessando la tecnologia CRISPR/Cas9 riguarda la possibilità di applicare quest'ultima alla linea germinale umana. In generale, la modifica della linea somatica umana viene eticamente accettata, qualora risulti chiaro in termini quantitativi il bilancio tra rischi e benefici e qualora si presenti consensualità da parte del soggetto interessato. Il discorso cambia nel momento in cui si parla di linea germinale.

Recenti preoccupazioni (da un punto di vista etico) sono sorte a seguito di lavori di ricerca sul genome editing che hanno interessato linee germinali umane; oltre all'evidente tema della controversia sulla consensualità, sorgono ulteriori problematiche legate al rischio di trasmissione di mutazioni impreviste alle generazioni future, essendo la tecnologia basata sul sistema CRISPR/Cas9 ancora in una fase di perfezionamento. Nonostante la maggior parte degli studi attuali sul genome editing a livello umano abbiano coinvolto cellule somatiche, un recente esperimento realizzato in Cina nel 2015 (Liang *et al.*, 2015) è stato attuato su zigoti umani. Tuttavia, i risultati ottenuti, si sono rivelati tutt'altro che di supporto per poter affermare che la tecnologia di CRISPR sia sufficientemente matura per essere utilizzata in sicurezza ed efficacemente su linee germinali umane (cosa quest'ultima che viene anche affermata nelle conclusioni dell'articolo di Liang e colleghi) (Baltimore *et al.*, 2015).

La scelta di studio del gruppo di ricerca cinese ha destato non poche preoccupazioni, fomentando un dibattito già esistente, il quale, nello stesso anno di pubblicazione dell'articolo di Liang e colleghi, ha condotto all'organizzazione di un incontro ufficiale, atto a discutere delle principali controversie etiche di uno strumento tanto potente quanto rischioso. In conclusione di tale incontro si è sancito come irresponsabile e non ammissibile qualsiasi intervento sulla linea germinale umana mediante l'utilizzo di CRISPR/Cas9, giudicando questa tecnologia ancora non sufficientemente perfezionata per essere utilizzata in maniera adeguata.

Il problema dell'inadeguatezza dello strumento discusso finora non è il solo problema esistente; Se un giorno la tecnologia CRISPR/Cas9 dovesse raggiungere un livello significativamente ammissibile di efficacia e sicurezza, tali da consentire un adeguato utilizzo sulla linea germinale umana, sorgeranno comunque nuove questioni di carattere etico e sociale. Saranno necessarie nuove discussioni riguardanti il dilemma dell'abuso della tecnologia, al fine di determinare la legittimità di eventuali utilizzi non terapeutici di CRISPR/Cas9, finalizzati esclusivamente al miglioramento della condizione fenotipica di un individuo. Sarà necessario chiarire in quali circostanze l'utilizzo di CRISPR/Cas9 su un essere umano rappresenti un uso legittimo dell'ingegneria genetica. Il dialogo coinvolge un numero costantemente in crescita di esperti di bioetica, esperti in legge, genetisti e ricercatori a livello internazionale (Baltimore *et al.*, 2015) (Rodriguez *et al.*, 2016).

Bibliografia

- Baltimore, David, Paul Berg, Michael Botchan, Dana Carroll, R. Alta Charo, George Church, Jacob E. Corn et al. "A prudent path forward for genomic engineering and germline gene modification." *Science* 348, no. 6230 (2015): 36-38.
- Banaszynski, Laura A., Corey W. Liu, and Thomas J. Wandless. "Characterization of the FKBP-Rapamycin-FRB Ternary Complex." *Journal of the American Chemical Society* 127, no. 13 (2005): 4715-4721.
- Barrangou, Rodolphe, and Luciano A. Marraffini. "CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity." *Molecular cell* 54, no. 2 (2014): 234-244.
- Barrangou, Rodolphe, Christophe Fremaux, H  l  ne Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A. Romero, and Philippe Horvath. "CRISPR provides acquired resistance against viruses in prokaryotes." *Science* 315, no. 5819 (2007): 1709-1712.
- Brouns, Stan JJ, Matthijs M. Jore, Magnus Lundgren, Edze R. Westra, Rik JH Slijkhuis, Ambrosius PL Snijders, Mark J. Dickman, Kira S. Makarova, Eugene V. Koonin, and John Van Der Oost. "Small CRISPR RNAs guide antiviral defense in prokaryotes." *Science* 321, no. 5891 (2008): 960-964.
- Carroll, Dana. "Genome engineering with zinc-finger nucleases." *Genetics* 188, no. 4 (2011): 773-782.
- Chen, Baohui, Luke A. Gilbert, Beth A. Cimini, Joerg Schnitzbauer, Wei Zhang, Gene-Wei Li, Jason Park *et al.* "Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system." *Cell* 155, no. 7 (2013): 1479-1491.
- Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu *et al.* "Multiplex genome engineering using CRISPR/Cas systems." *Science* 339, no. 6121 (2013): 819-823.
- Deltcheva, Elitza, Krzysztof Chylinski, Cynthia M. Sharma, Karine Gonzales, Yanjie Chao, Zaid A. Pirzada, Maria R. Eckert, J  rg Vogel, and Emmanuelle Charpentier. "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." *Nature* 471, no. 7340 (2011): 602-607.
- Dominguez, Antonia A., Wendell A. Lim, and Lei S. Qi. "Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation." *Nature Reviews Molecular Cell Biology* 17, no. 1 (2016): 5-15.
- Doudna, Jennifer A., and Emmanuelle Charpentier. "The new frontier of genome engineering with CRISPR-Cas9." *Science* 346, no. 6213 (2014).
- Haurwitz, Rachel E., Martin Jinek, Blake Wiedenheft, Kaihong Zhou, and Jennifer A. Doudna. "Sequence-and structure-specific RNA processing by a CRISPR endonuclease." *Science* 329, no. 5997 (2010): 1355-1358.
- Hille, Frank, and Emmanuelle Charpentier. "CRISPR-Cas: biology, mechanisms and relevance." *Phil. Trans. R. Soc. B* 371, no. 1707 (2016).

Hsu, Patrick D., David A. Scott, Joshua A. Weinstein, F. Ann Ran, Silvana Konermann, Vineeta Agarwala, Yinqing Li *et al.* "DNA targeting specificity of RNA-guided Cas9 nucleases." *Nature biotechnology* 31, no. 9 (2013): 827-832.

Hsu, Patrick D., Eric S. Lander, and Feng Zhang. "Development and applications of CRISPR-Cas9 for genome engineering." *Cell* 157, no. 6 (2014): 1262-1278.

Jiang, Fuguo, and Jennifer A. Doudna. "CRISPR-Cas9 Structures and Mechanisms." *Annual Review of Biophysics* 0 (2016).

Jiang, Fuguo, and Jennifer A. Doudna. "The structural biology of CRISPR-Cas systems." *Current opinion in structural biology* 30 (2015): 100-111.

Jiang, Fuguo, David W. Taylor, Janice S. Chen, Jack E. Kornfeld, Kaihong Zhou, Aubri J. Thompson, Eva Nogales, and Jennifer A. Doudna. "Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage." *Science* 351, no. 6275 (2016): 867-871.

Jinek, Martin, Alexandra East, Aaron Cheng, Steven Lin, Enbo Ma, and Jennifer Doudna. "RNA-programmed genome editing in human cells." *elife* 2 (2013)

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." *Science* 337, no. 6096 (2012): 816-821.

Karvelis, Tautvydas, Giedrius Gasiunas, Algirdas Miksys, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys. "crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*." *RNA biology* 10, no. 5 (2013): 841-851.

Lander, Eric S. "The heroes of CRISPR." *Cell* 164, no. 1 (2016): 18-28.

Larson, Matthew H., Luke A. Gilbert, Xiaowo Wang, Wendell A. Lim, Jonathan S. Weissman, and Lei S. Qi. "CRISPR interference (CRISPRi) for sequence-specific control of gene expression." *Nature protocols* 8, no. 11 (2013): 2180-2196.

Liang, Puping, Yanwen Xu, Xiya Zhang, Chenhui Ding, Rui Huang, Zhen Zhang, Jie Lv *et al.* "CRISPR/Cas9-mediated gene editing in human trippronuclear zygotes." *Protein & cell* 6, no. 5 (2015): 363-372.

Makarova, Kira S., Yuri I. Wolf, Omer S. Alkhnbashi, Fabrizio Costa, Shiraz A. Shah, Sita J. Saunders, Rodolphe Barrangou *et al.* "An updated evolutionary classification of CRISPR-Cas systems." *Nature Reviews Microbiology* 13, no. 11 (2015): 722-736.

Mali, P., Esvelt, K. M., and Church, G. M. "Cas9 as a versatile tool for engineering biology. *Nature methods*.", 10 no. 10 (2013): 957-963.

Mali, Prashant, John Aach, P. Benjamin Stranges, Kevin M. Esvelt, Mark Moosburner, Sriram Kosuri, Luhan Yang, and George M. Church. "CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering." *Nature biotechnology* 31, no. 9 (2013): 833-838.

Mali, Prashant, Luhan Yang, Kevin M. Esvelt, John Aach, Marc Guell, James E. DiCarlo, Julie E. Norville, and George M. Church. "RNA-guided human genome engineering via Cas9." *Science* 339, no. 6121 (2013): 823-826.

- Marraffini, Luciano A., and Erik J. Sontheimer. "CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA." *science* 322, no. 5909 (2008): 1843-1845.
- Nemudryi, A. A., K. R. Valetdinova, S. P. Medvedev, and S. M. Zakian. "TALEN and CRISPR/Cas genome editing systems: tools of discovery." *Acta Naturae (англоязычная версия)* 6, no. 3 (22) (2014).
- Nihongaki, Yuta, Fuun Kawano, Takahiro Nakajima, and Moritoshi Sato. "Photoactivatable CRISPR-Cas9 for optogenetic genome editing." *Nature biotechnology* 33, no. 7 (2015): 755-760.
- Nihongaki, Yuta, Shun Yamamoto, Fuun Kawano, Hideyuki Suzuki, and Moritoshi Sato. "CRISPR-Cas9-based photoactivatable transcription system." *Chemistry & biology* 22, no. 2 (2015): 169-174.
- Nishimasu, Hiroshi, F. Ann Ran, Patrick D. Hsu, Silvana Konermann, Soraya I. Shehata, Naoshi Dohmae, Ryuichiro Ishitani, Feng Zhang, and Osamu Nureki. "Crystal structure of Cas9 in complex with guide RNA and target DNA." *Cell* 156, no. 5 (2014): 935-949.
- Nuñez, James K., Lucas B. Harrington, and Jennifer A. Doudna. "Chemical and biophysical modulation of Cas9 for tunable genome engineering." *ACS chemical biology* 11, no. 3 (2016): 681-688.
- Rath, Devashish, Lina Amlinger, Archana Rath, and Magnus Lundgren. "The CRISPR-Cas immune system: biology, mechanisms and applications." *Biochimie* 117 (2015): 119-128.
- Rodriguez, E. "Ethical issues in genome editing using Crispr/Cas9 system." *J. Clin. Res. Bioeth* 7, no. 266 (2016): 10-4172.
- Sander, J. D., & Joung, J. K. "CRISPR-Cas systems for editing, regulating and targeting genomes." *Nature biotechnology*, 32 no. 4 (2014): 347-355.
- Shah, Neel H., and Tom W. Muir. "Inteins: nature's gift to protein chemists." *Chemical science* 5, no. 2 (2014): 446-461.
- Shah, Shiraz A., Susanne Erdmann, Francisco JM Mojica, and Roger A. Garrett. "Protospacer recognition motifs: mixed identities and functional diversity." *RNA biology* 10, no. 5 (2013): 891-899.
- Stoddard, Barry L. "Homing endonuclease structure and function." *Quarterly reviews of biophysics* 38, no. 1 (2005): 49-95.
- Van Der Oost, John, Edze R. Westra, Ryan N. Jackson, and Blake Wiedenheft. "Unravelling the structural and mechanistic basis of CRISPR-Cas systems." *Nature reviews Microbiology* 12, no. 7 (2014): 479-492.
- Zhang, Hui, Jinshan Zhang, Pengliang Wei, Botao Zhang, Feng Gou, Zhengyan Feng, Yanfei Mao *et al.* "The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation." *Plant biotechnology journal* 12, no. 6 (2014): 797-807.

