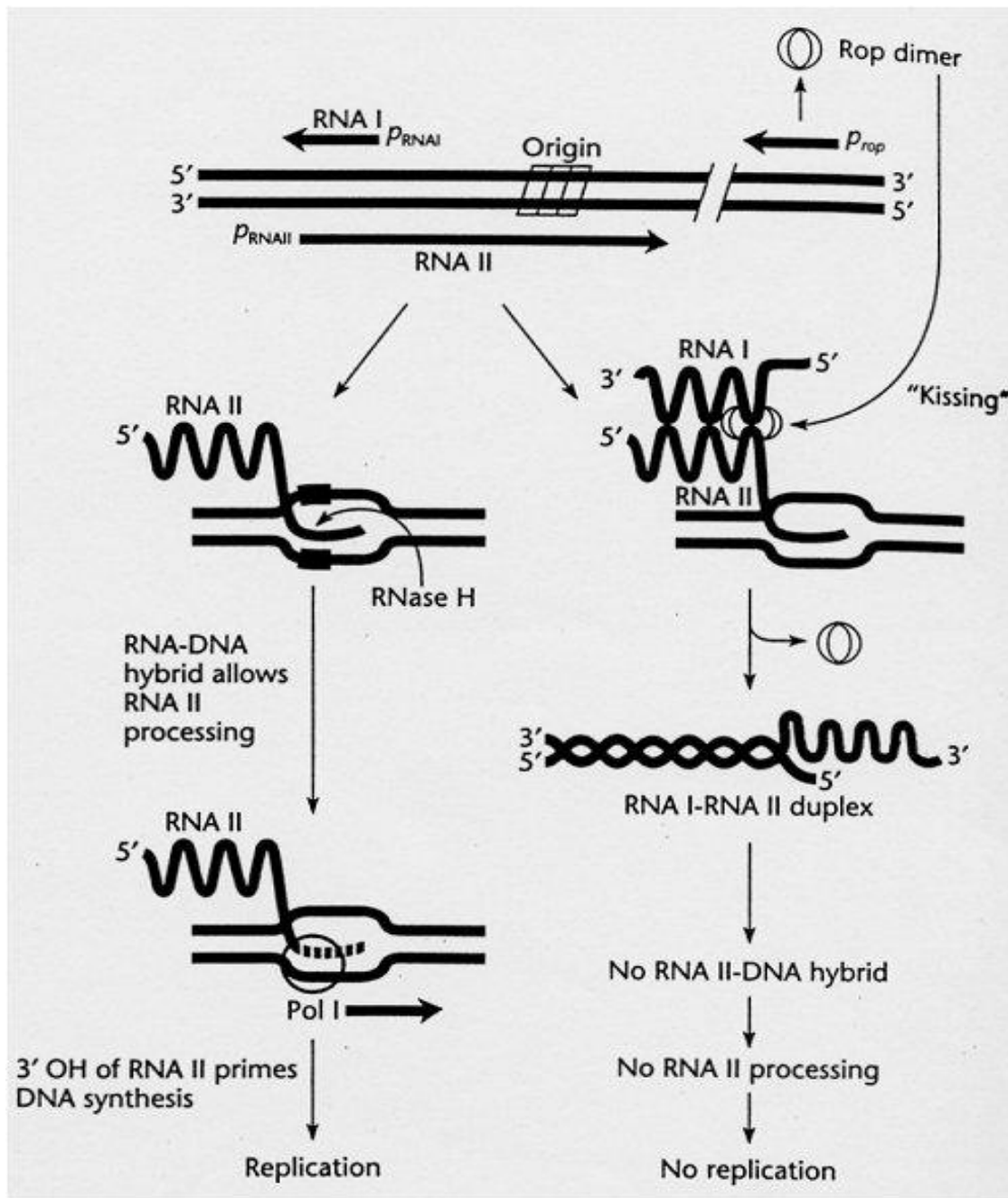


Small RNAs in bacteria

Discovered in 1981

Synthesized as discrete transcripts
with dedicated promoters and terminators

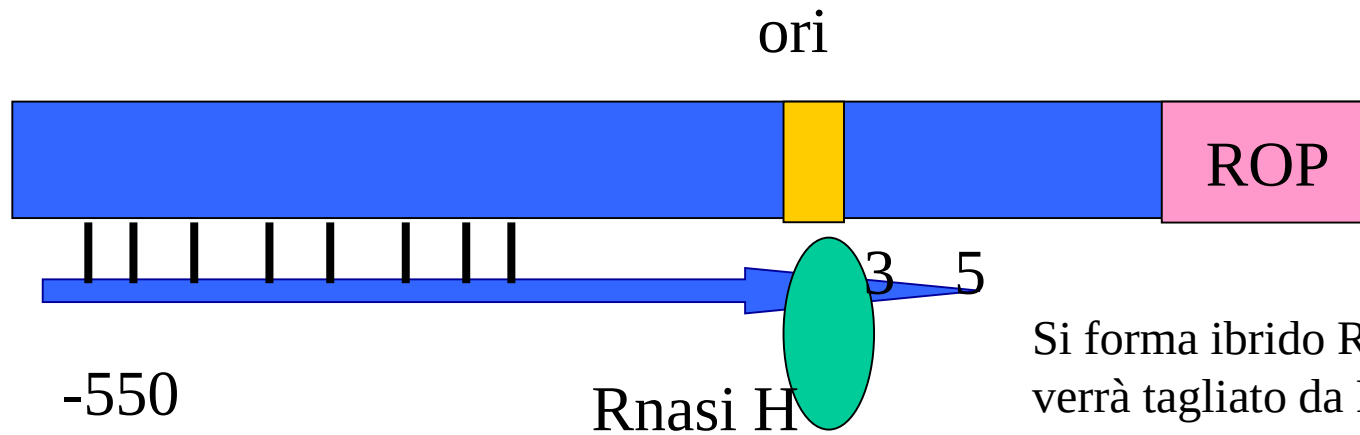
Act as regulators through base pairing with mRNAs



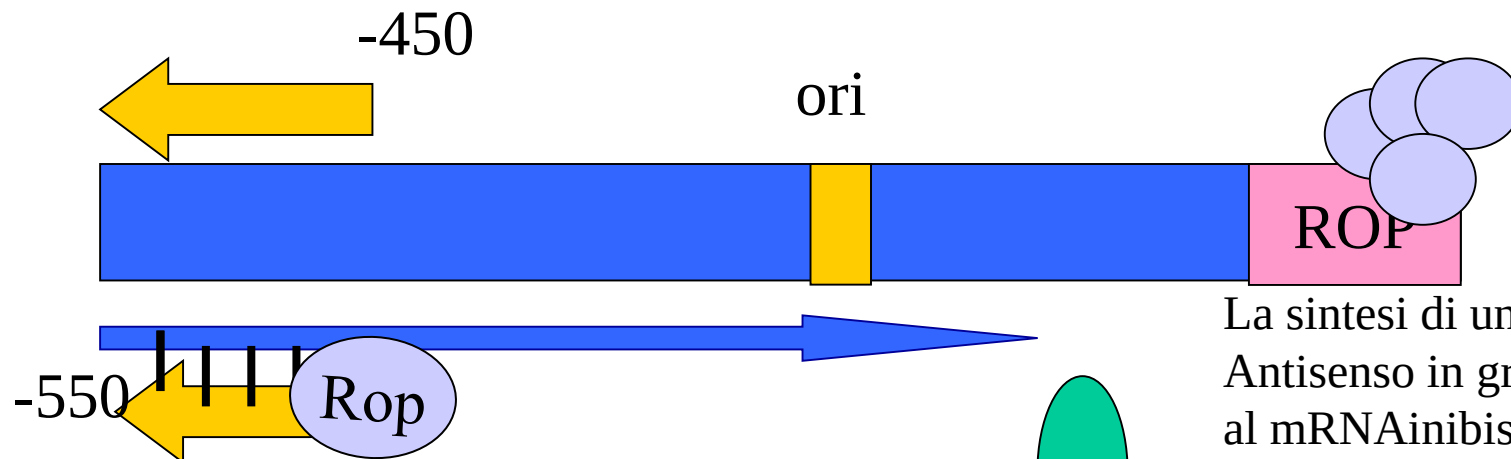
Controllo della replicazione nei plasmidi di tipo ColE1.

L'ibrido RNAII-DNA viene riconosciuto da RNase H che taglia creando l'innesco per la replicazione da parte di PolI

Se viene sintetizzato mRNAI si crea un ibrido RNAI -RNAII che impedisce il riconoscimento da parte di RNAII del filamento di DNA



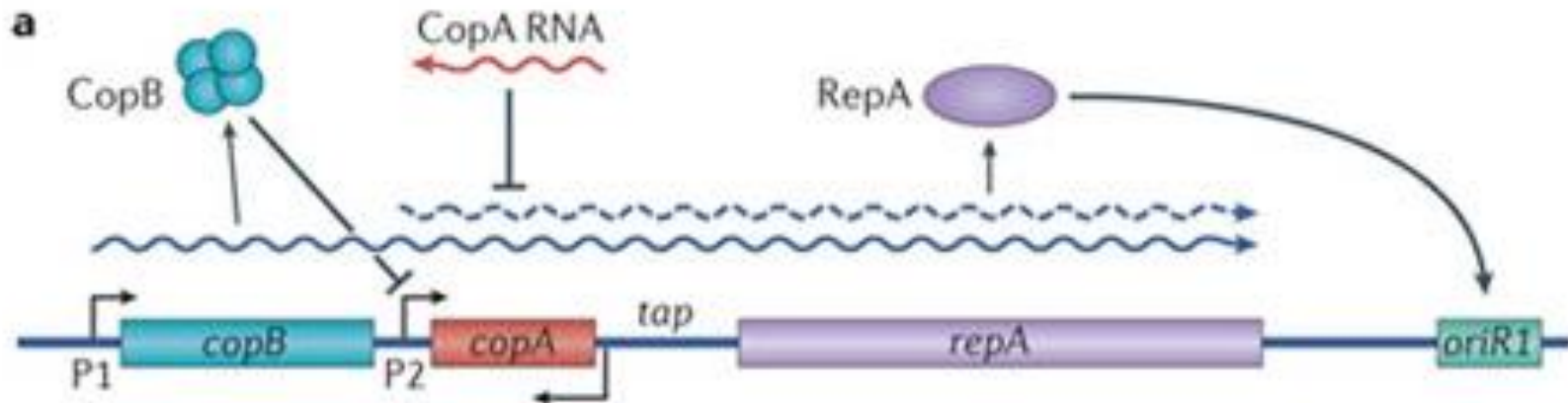
Si forma ibrido RNA-DNA che verrà tagliato da RNAsiH a livello di ori generando un primer per la replicazione



La sintesi di un RNA Antisense in grado di legarsi al mRNA inibisce la formazione dell'ibrido RNA-DNA

Rop stabilizza i legami RNA-RNA

La replicazione di R1



Il plasmide R1 (R100) costituisce un modello ben studiato di replicazione plasmidica .

In questo plasmide la proteina necessaria per la replicazione RepA si lega all'origine (*oriR1*) localizzata a valle del gene.

Il gene *repA* è espresso a partire da 2 promotori (P1 e P2) ed è sottoposto ad una duplice forma di controllo negativo.

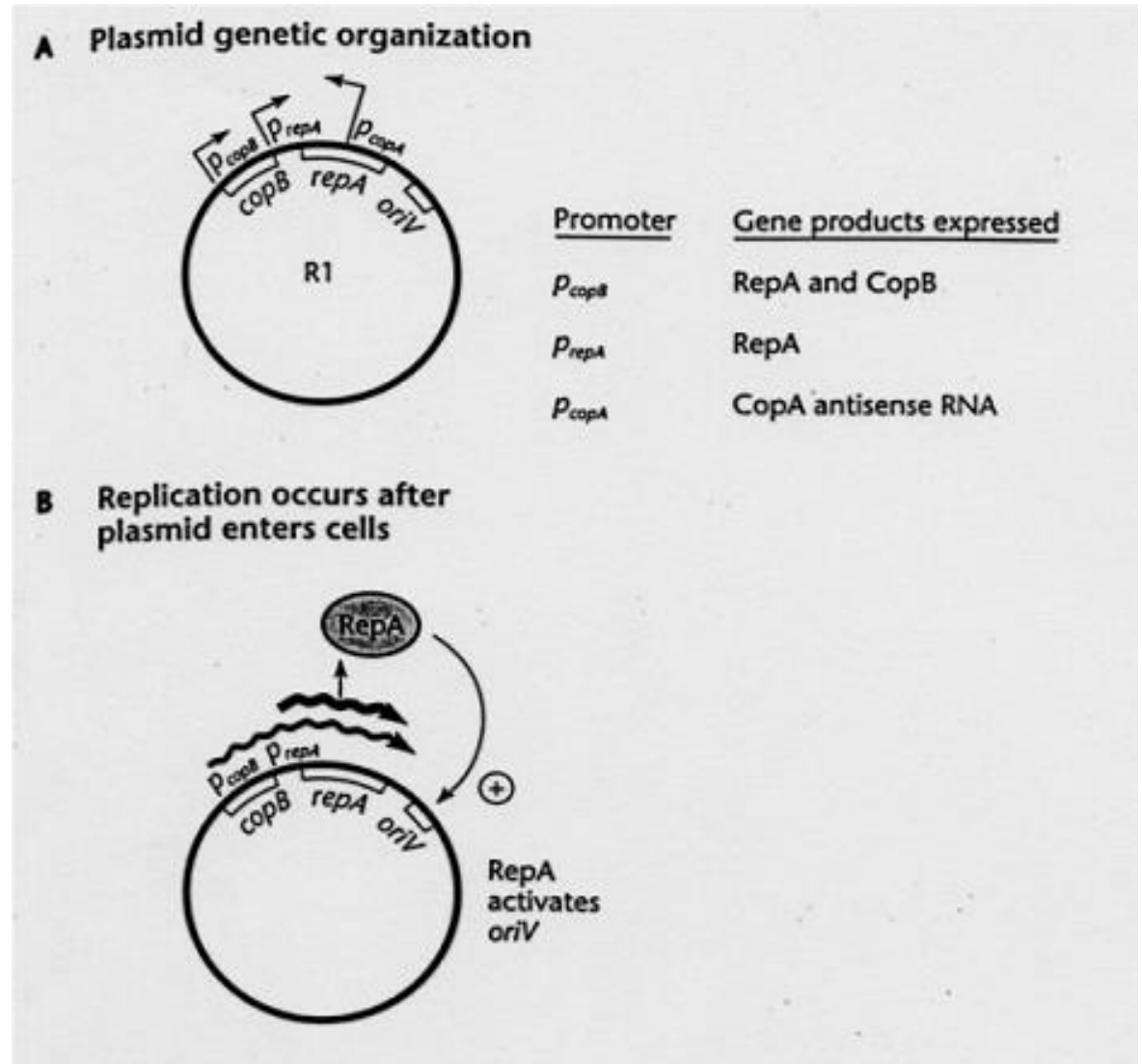
Il primo repressore è CopB che viene trascritto assieme al gene *repA* quando la trascrizione parte dal promotore P1. CopB reprime la trascrizione di *repA* a partire dal promotore P2 che quindi rimane silente in condizioni normali. Oltre a CopB, il livello di *repA* è controllato anche da un piccolo RNA chiamato CopA che agisce a livello post trascrizionale legandosi mRNA di *repA*, alterandone la struttura in modo prevenirne la traduzione

Replicazione di R1(o R100)

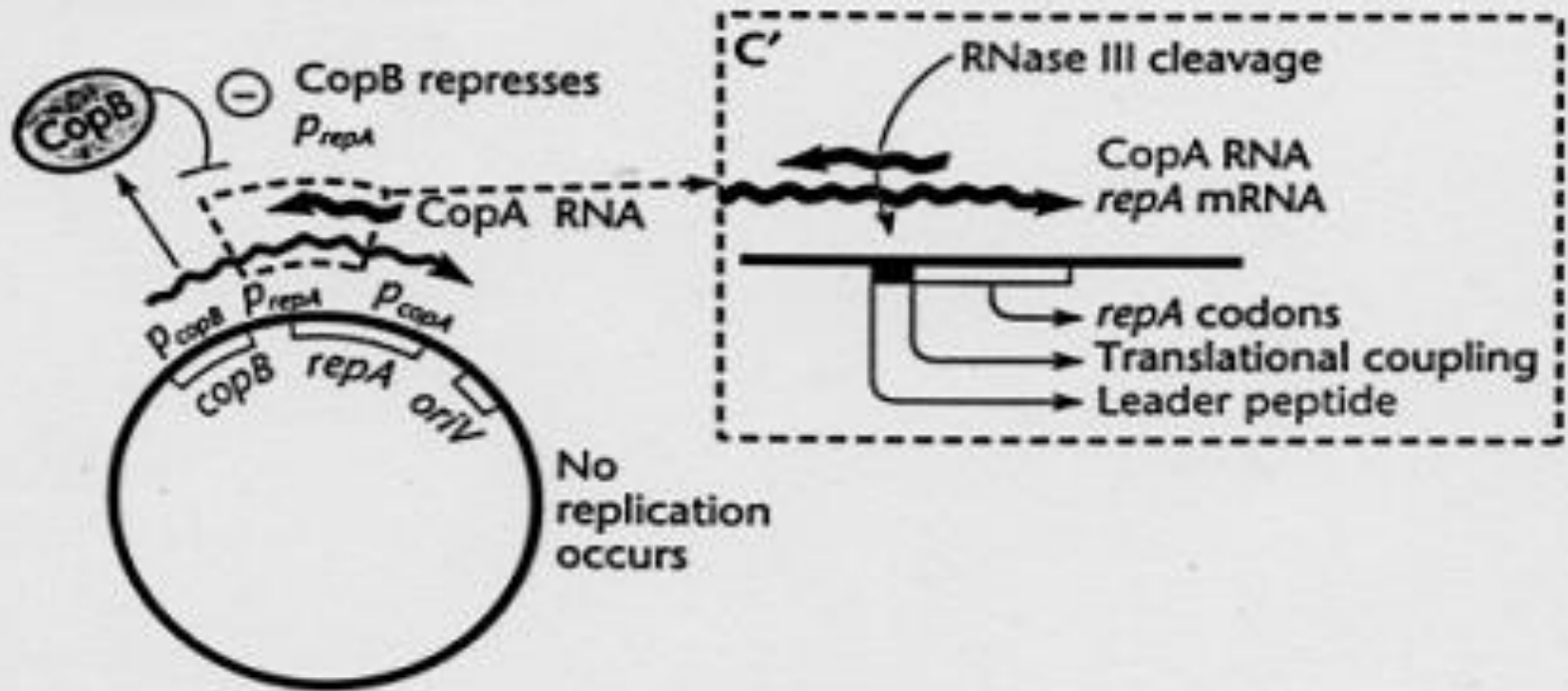
La proteina RepA necessaria per attivare la replicazione può essere trascritta a partire dal promotore

- *P_{copB}*
- *P_{repA}*

In assenza di CopB la proteina viene trascritta da entrambi i promotori:
Quando succede?
Quando il plasmide entra in una cellula o dopo la divisione cellulare per diluizione di CopB



C Replication shutdown



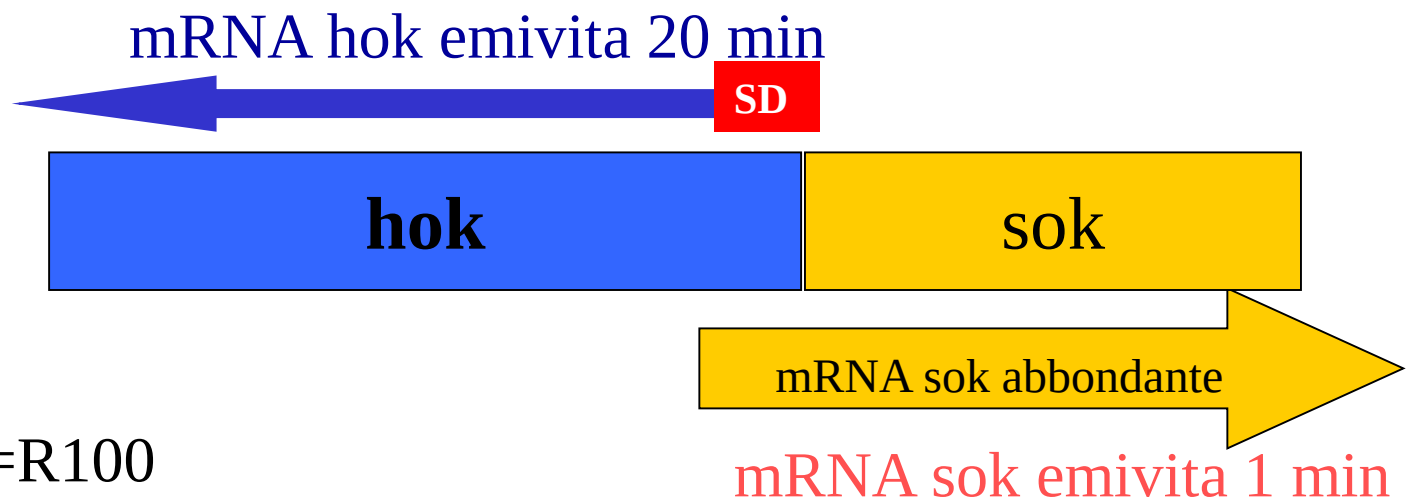
In presenza di CopB

- CopB reprime il *PrepA*.
- *repA* viene trascritto solo dal *P_{copB}*
- viene trascritto sull'elica complementare il mRNA *copA* che agisce da antisenso sul mRNA *copB-repA*.
- L'ibrido RNA-RNA viene digerito da RNase III e non si ha traduzione di RepA

Sistema hok -sok

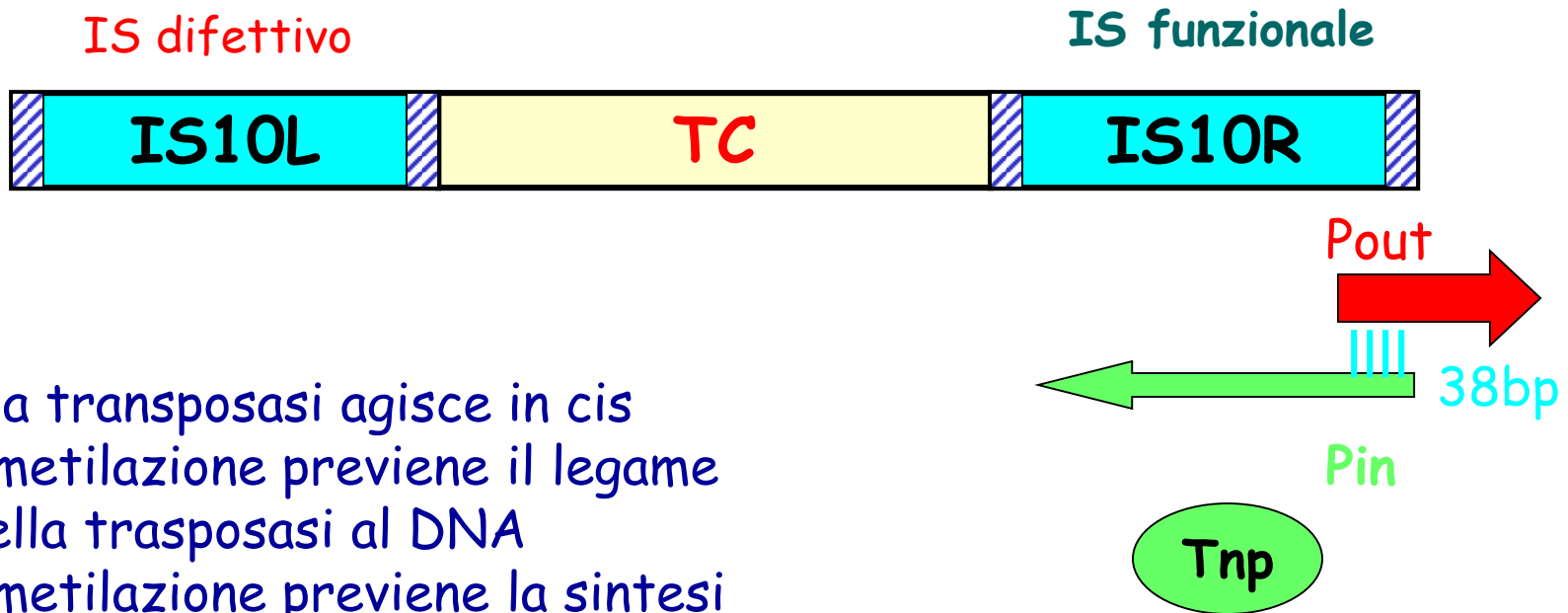
Il plasmide R1(o R100) porta un gene letale *hok* (host cell killing) che codifica per una tossina in grado di provocare depolimerizzazione delle membrana.

Sull'elica complementare del DNA di *hok* viene trascritta il mRNA del gene *sok* che ha una una regione di 128 nt complementare con la regione SD di *hok*. I 2 RNA hanno diversa emivita 20 min e 1 min. Hok non viene mai tradotto per azione del mRNA di sok e la cellula con R1 rimane pertanto vitale. Se una cellula non eredita R1 in seguito a divisione allora mRNAsok che ha una lunga emivita verrà tradotto perchè mRNA sok avendo un emivita più breve non sarà più presente.

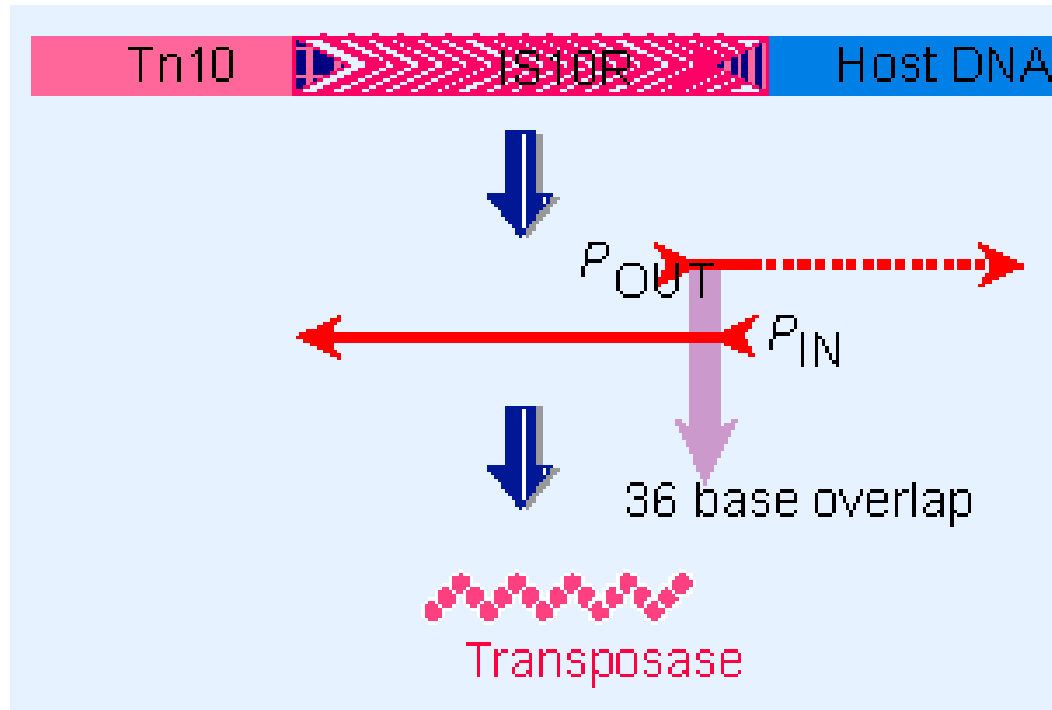


Plasmide R1=R100

Struttura di Tn10



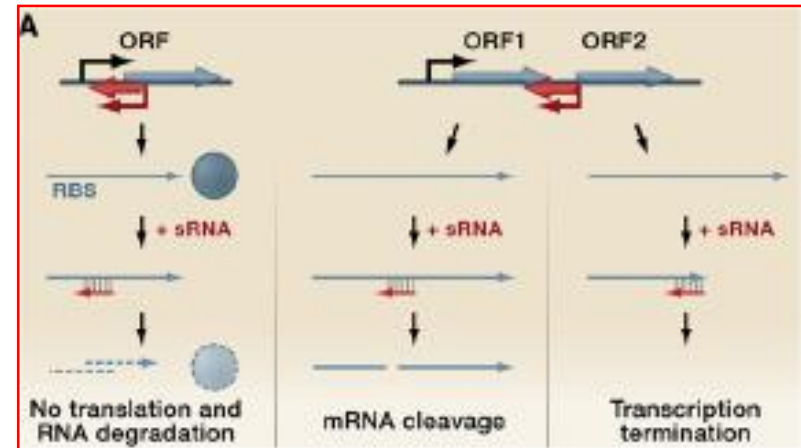
- la trasposasi agisce in cis
- metilazione previene il legame della trasposasi al DNA
- metilazione previene la sintesi della trasposasi



Tn10 è caratterizzato dalla presenza di 2 promotori localizzati all'estremità dell'IS10R. Il promotore forte P_{OUT} promuove la trascrizione verso l'esterno dell'elemento ovvero verso sequenze dell'ospite. Il promotore più debole P_{IN} promuove la trascrizione di un mRNA interno all' IS10R che sarà tradotto nella trasposasi. Il trascritto OUT è un mRNA di 70 nt espresso ad un livello 100 volte superiore rispetto mRNA IN ed è anche molto più stabile di mRNA IN.

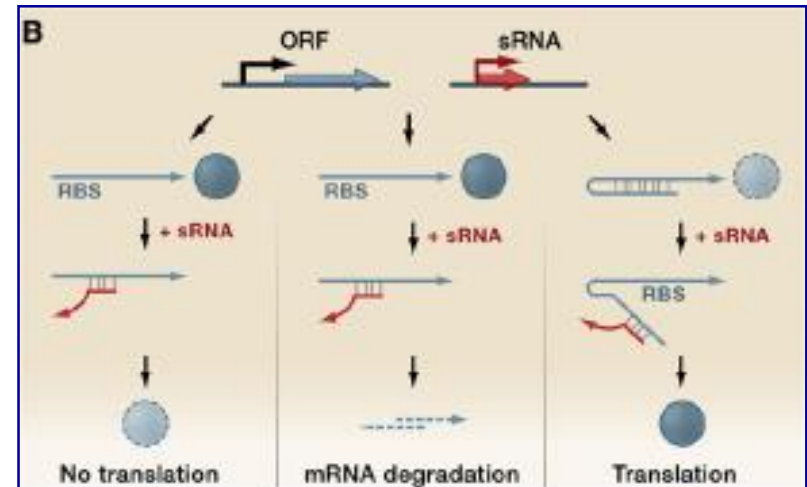
Cis encoded sRNAs

- are encoded on the DNA strand opposite the target RNA
- share extended regions of complete complementarity



Trans encoded sRNAs

- are encoded far from their targets
- share only limited complementarity with their target mRNAs
- require RNA chaperones to facilitate the RNA-RNA interactions.



sRNA cis encoded o asRNA

Numerosi sRNA sono codificati sull'elica opposta rispetto alla sequenza che codifica mRNA bersaglio e sono definiti RNA antisenso (asRNA). Questi asRNA si ritrovano implicati nella traduzione, e/o nella stabilità dei messaggeri del gene "senso" codificati dall'elica complementare.

In alcuni casi la regolazione sul mRNA target può essere effettuata dalla trascrizione stessa dell'asRNA piuttosto che dal asRNA

sRNA trans encoded

Gli sRNA studiati in maggior dettaglio sono definiti sRNA trans encoded e sono molecole che regolano mRNA tramite appaiamenti di basi brevi ed imperfetti.

Molti di questi trans RNAs si legano nel "ribosome binding site" (RBS o Sequenza Shine-Delgarno) o in sequenze adiacenti al RBS dei loro geni bersaglio e bloccano la traduzione del mRNA impedendo l'accesso ai ribosomi

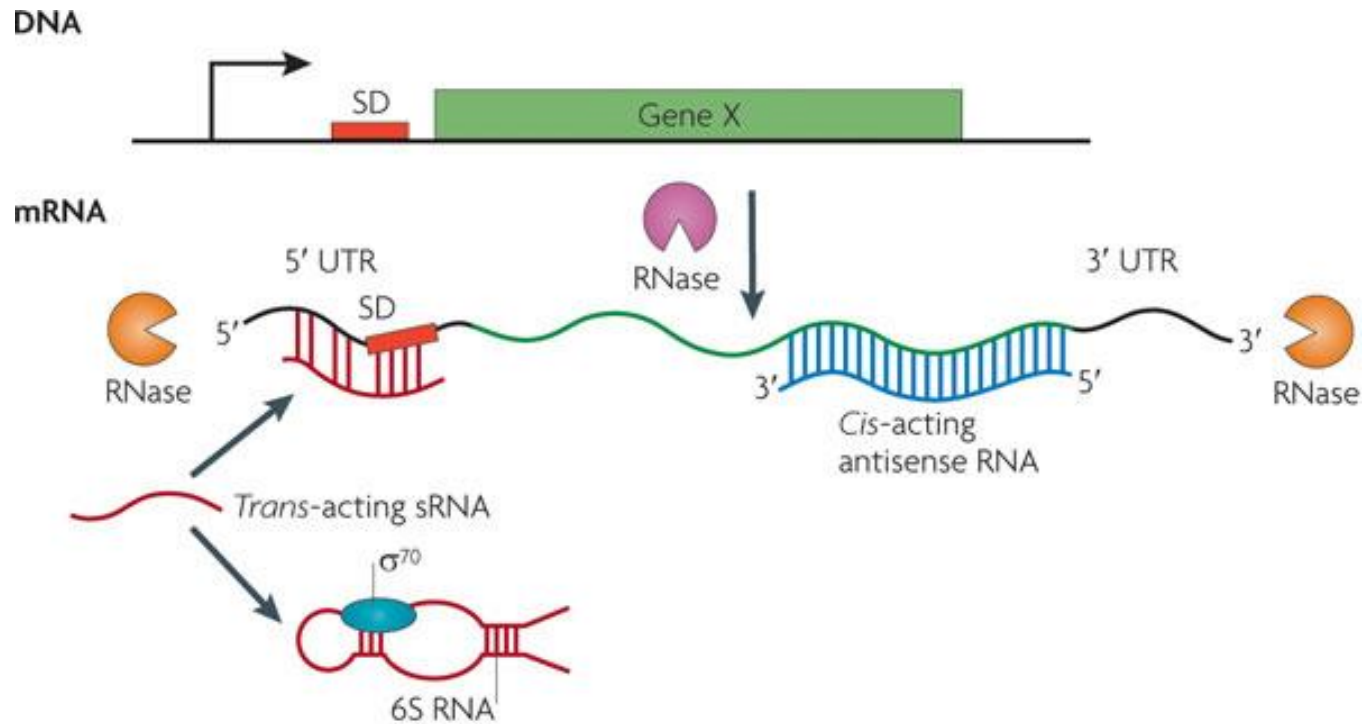
All'interno di questa famiglia si trovano sRNA che pur legandosi in regioni diverse (generalmente più lontane dal RBS) bloccano sempre il legame dei ribosomi o in qualche caso aumentano il legame dei ribosomi prevenendone così la formazione di strutture secondarie o aumentano/ diminuiscono la stabilità degli mRNA.

Nei Batteri Gram- la proteina Hfq è richiesta per la funzione e stabilità di questa famiglia di sRNA.

Piccoli RNA come molecole regolatrici nei procarioti

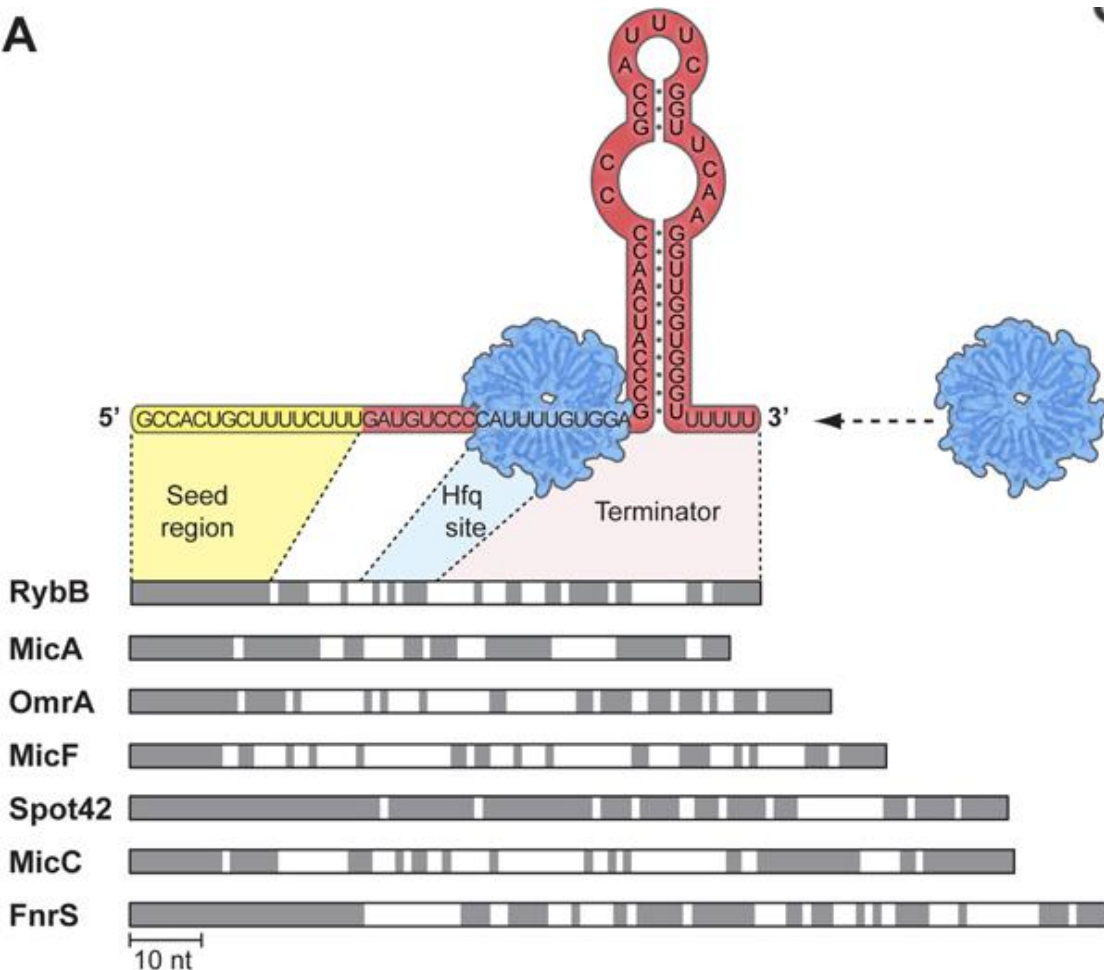
- Identificati in numerosi batteri molecole di **RNA non codificanti** di piccole dimensioni 40-500 nucleotidi
- La maggior parte implicati nel controllo del processo di traduzione mediante appaiamento con la regione leader del mRNA del gene bersaglio
- Sono complementari a mRNA del gene bersaglio ma trascritti sull'altra elica, funzionano da **RNA antisenso**
- **Le interazioni RNA-RNA** favorite da una proteina molto abbondante chaperon degli RNA definita **Hfq**

Controllo di un mRNA nei batteri



Struttura modulare dei sRNA trans encoded

A



Una sequenza all'estremità 3' seguita da una sequenza di poliT promuove la terminazione Rho-indipendente del piccolo RNA e lo protegge dalle esonucleasi 3' (Terminator)

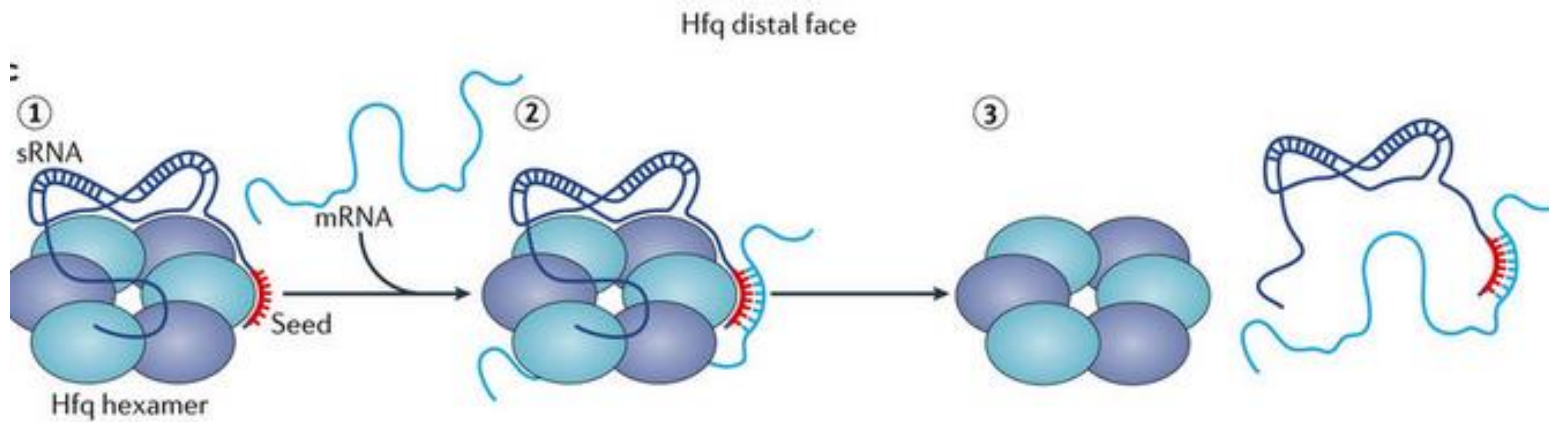
Un secondo dominio è quello di legame ad Hfq che è richiesta per la funzione e stabilizzazione di molti RNA (Hfq site) .

Una terza regione è utilizzata per l'appiamento con RNA bersaglio (seed region) .

La funzione degli sRNA trans nell'appaiamento al mRNA bersaglio dipende generalmente dalla proteina Hfq considerata un chaperone per gli RNA.

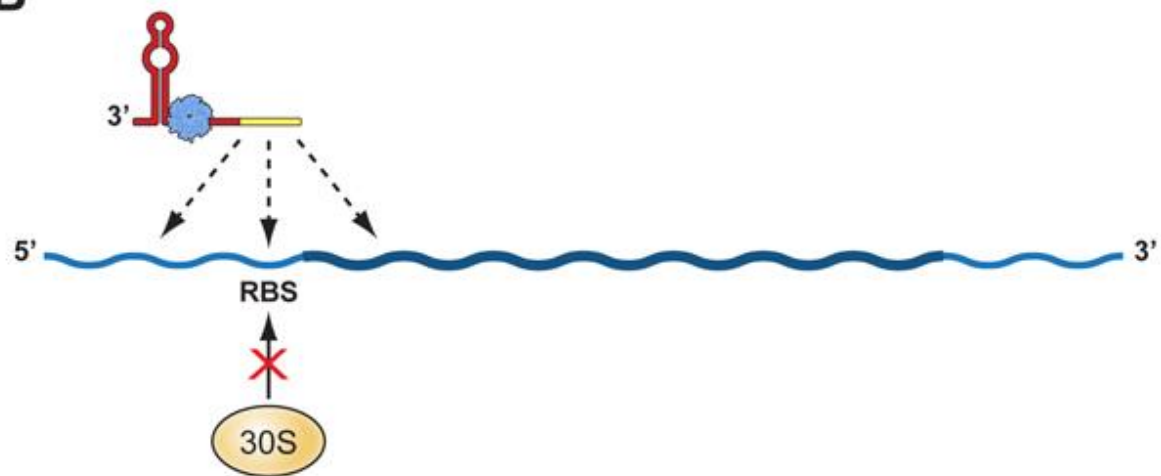
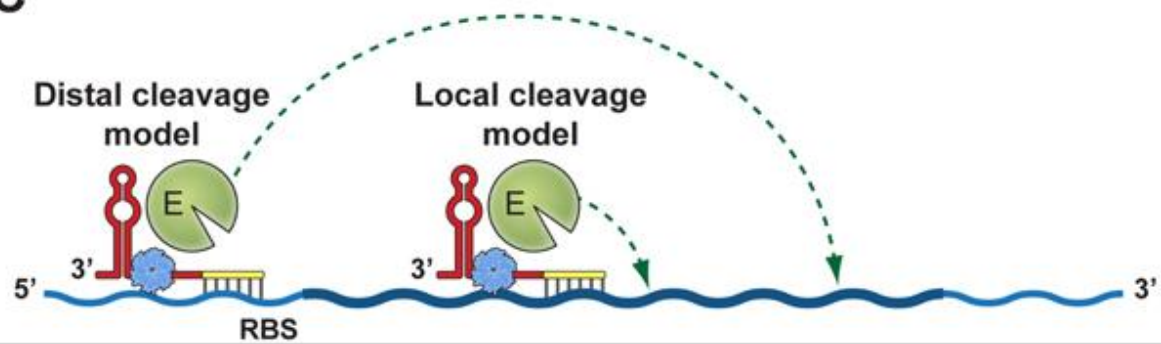
Studi in vitro suggeriscono che Hfq si leghi ad una regione a singola elica ricca in A/U localizzata nelle adiacenze di una struttura a stem-loop .

Le predizioni sono in ogni caso non molto attendibili dato il basso livello di conservazione nella sequenza primaria dei sRNA e la presenza di siti multipli ricchi in A/U. Inoltre dati recenti indicano che le sequenze di poliU all'estremità 3' degli sRNAs potrebbero essere riconosciute da Hfq come sito di legame.

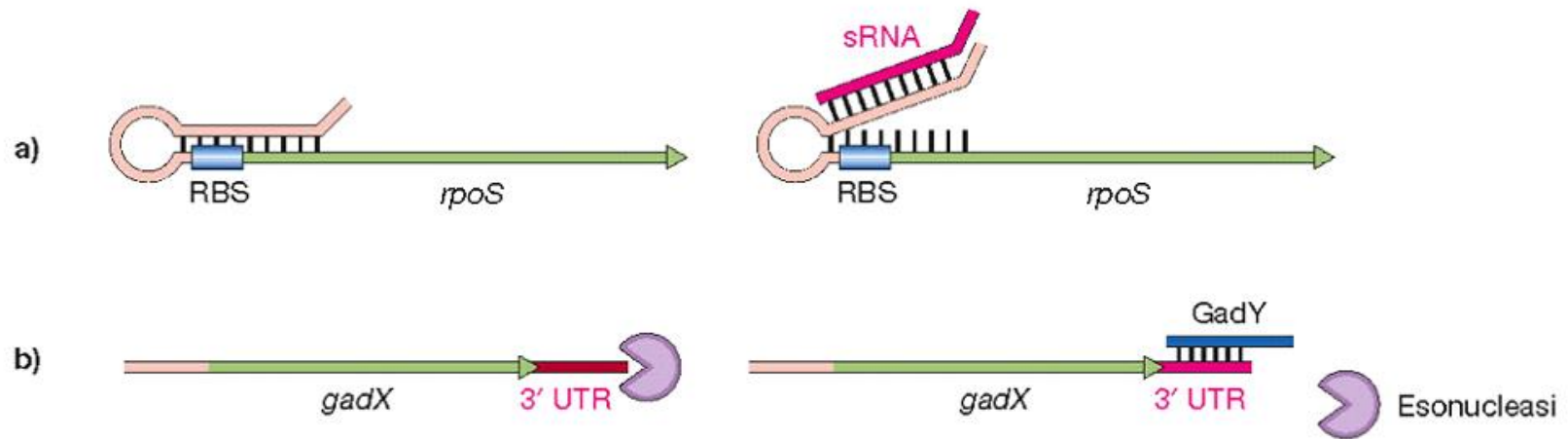


Nonostante il ruolo di Hfq come RNA chaperon sia ampiamente accettato i dettagli dei meccanismi tramite i quali promuove degli incontri "produttivi" tra le coppie di RNA rimane non definito.

La struttura ad alta di risoluzione di Hfq insieme a saggi di legame di Hfq e di suoi mutanti hanno permesso di evidenziare che vi sono vari siti di legame all'RNA sull'anello esamerico di Hfq, confermando il fatto che Hfq sia in grado di legare simultaneamente più RNA.

B**C**

Regolazione positiva mediata dai piccoli RNA



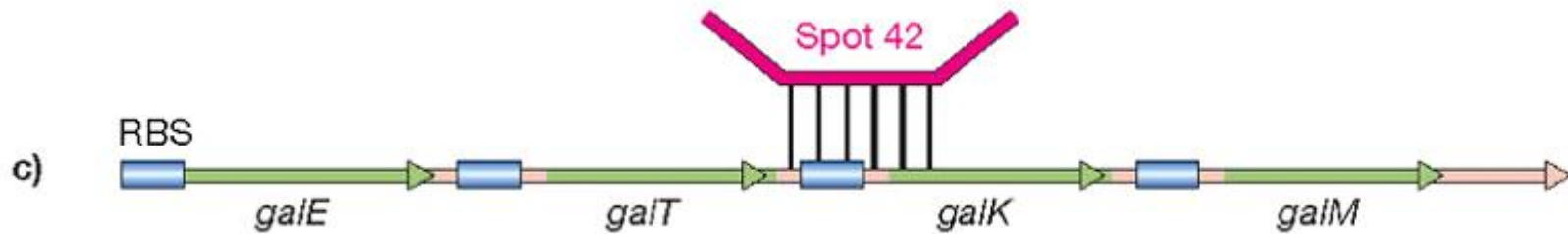
a) Nel caso del trascritto per gene per σ^s il sito di legame dei ribosomi (RBS) forma una struttura secondaria con una regione di mRNA a monte impedendo così l'accesso ai ribosomi.

Il legame di sRNA alla regione di appaiamento con la sequenza RBS apre la struttura permettendo la traduzione del mRNA σ^s .

b) GadY un piccolo RNA non codificante impedisce la degradazione del trascritto per *gadX* legandosi nella regione 3'UTR. Il sistema gad è importante per la sopravvivenza in stress acido.

Regolazione negativa mediata dai piccoli RNA

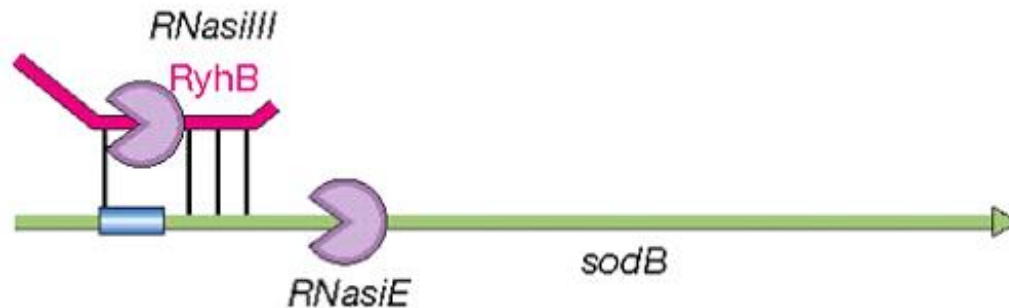
L'operone galETK



In presenza di glucosio un piccolo RNA Spot42 è espresso ad alto livello e si lega al sito d'inizio della traduzione del gene *galK* inibendone la traduzione.

In questo modo si destabilizza mRNA dell'intero operone trascritto dal Promotore P2. Si avrà una ridotta traduzione di *galT* mentre l'effetto su *galE* sarà minore

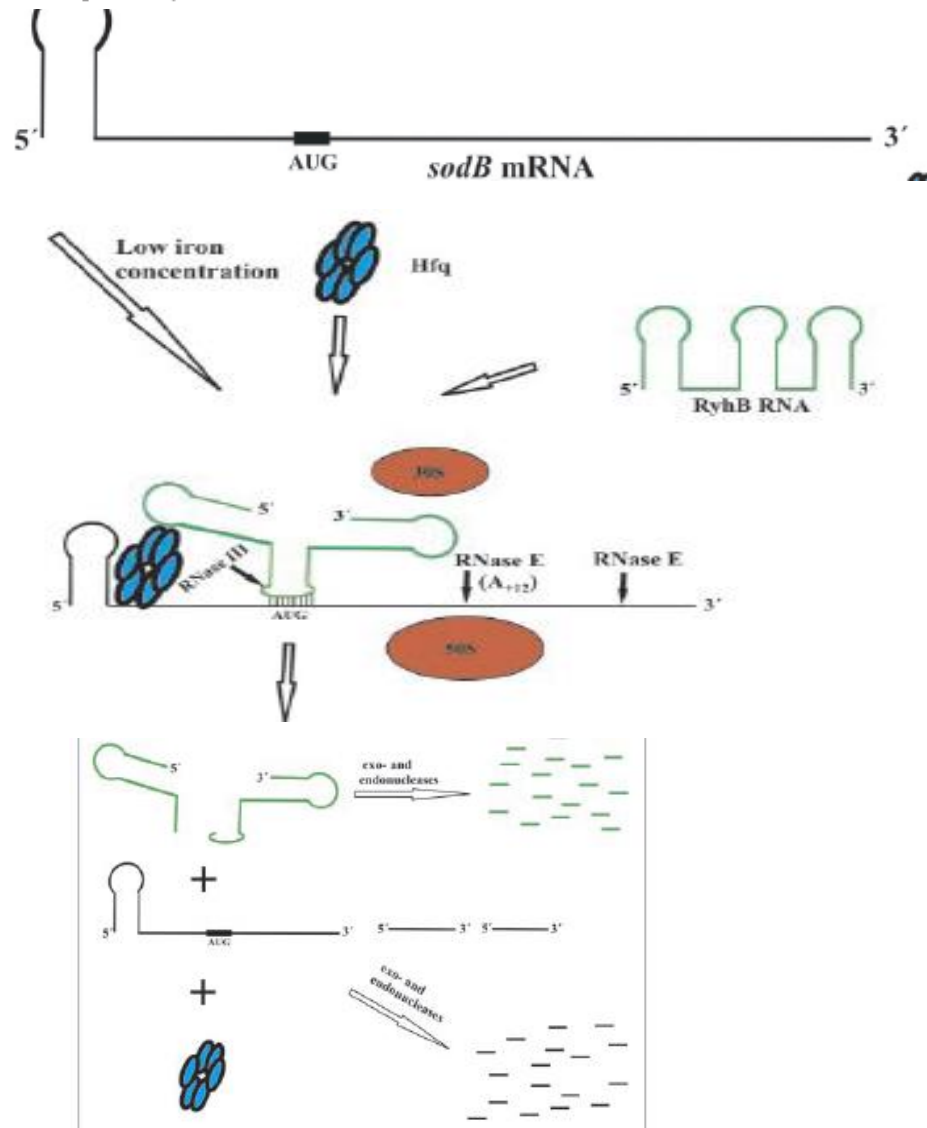
Controllo negativo mediato da RyhB



Il piccolo RNA RyhB

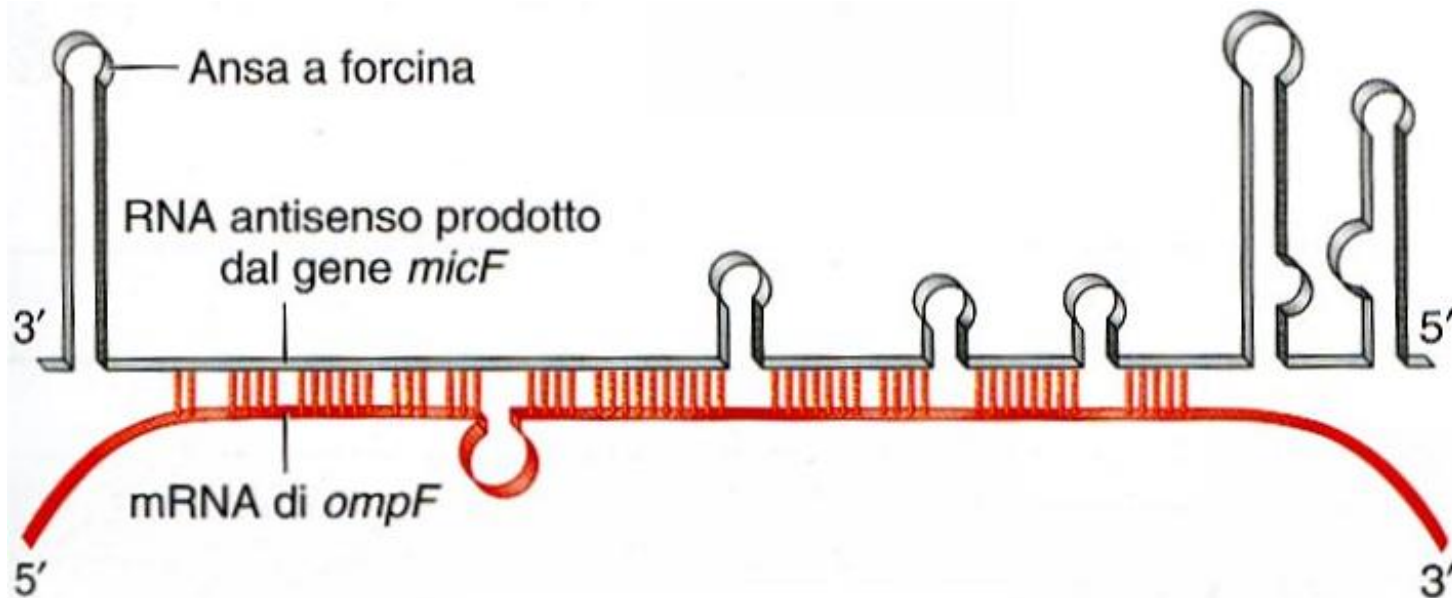
- è trascritto ad alta efficienza in in carenza di ferro
- si lega al mRNA del gene *sodB* favorendone la degradazione da parte della RNasi III
- viene a sua volta degradato da RNasiIII

lower affinity for 30S ribosomal subunits, translation of the truncated *sodB* mRNA is less efficient, thereby allowing RNase E cleavage at downstream site(s). The latter results in the subsequent loss of ribosomal subunits and degradation of the intermediate ribosome-free RNA fragments by endo- and exonucleases. The iron-dependent inactivation of the *sodB* transcript, which is initiated by the small regulatory RNA RyhB and Hfq, is shown on the right. The base pairing with RyhB, which is known to cause structural rearrangements within the *sodB* 5'-UTR (28), inhibits translation and induces RNase E cleavage at the downstream site A₊₁₂, whereas the coordinated decay of RyhB is initiated by RNase III cleavage at U₄₆ (for details, see Figure 7). Similar to the general pathway, the degradation of the intermediate products is accomplished by exo- and endoribonucleases.



sRNA e regolazione della porina OmpF

Oltre alla proteina OmpR l'espressione del gene *ompF* è regolata da un RNA antisenso chiamato MicF (mRNA interfering complementary RNA).



RNA MicF è complementare alla sequenza d'inizio della traduzione del mRNA di *ompF*. Quando RNA MicF si appaia all'm RNA di *ompF* ne impedisce la traduzione

sRNA che agiscono in trans :RNAIII

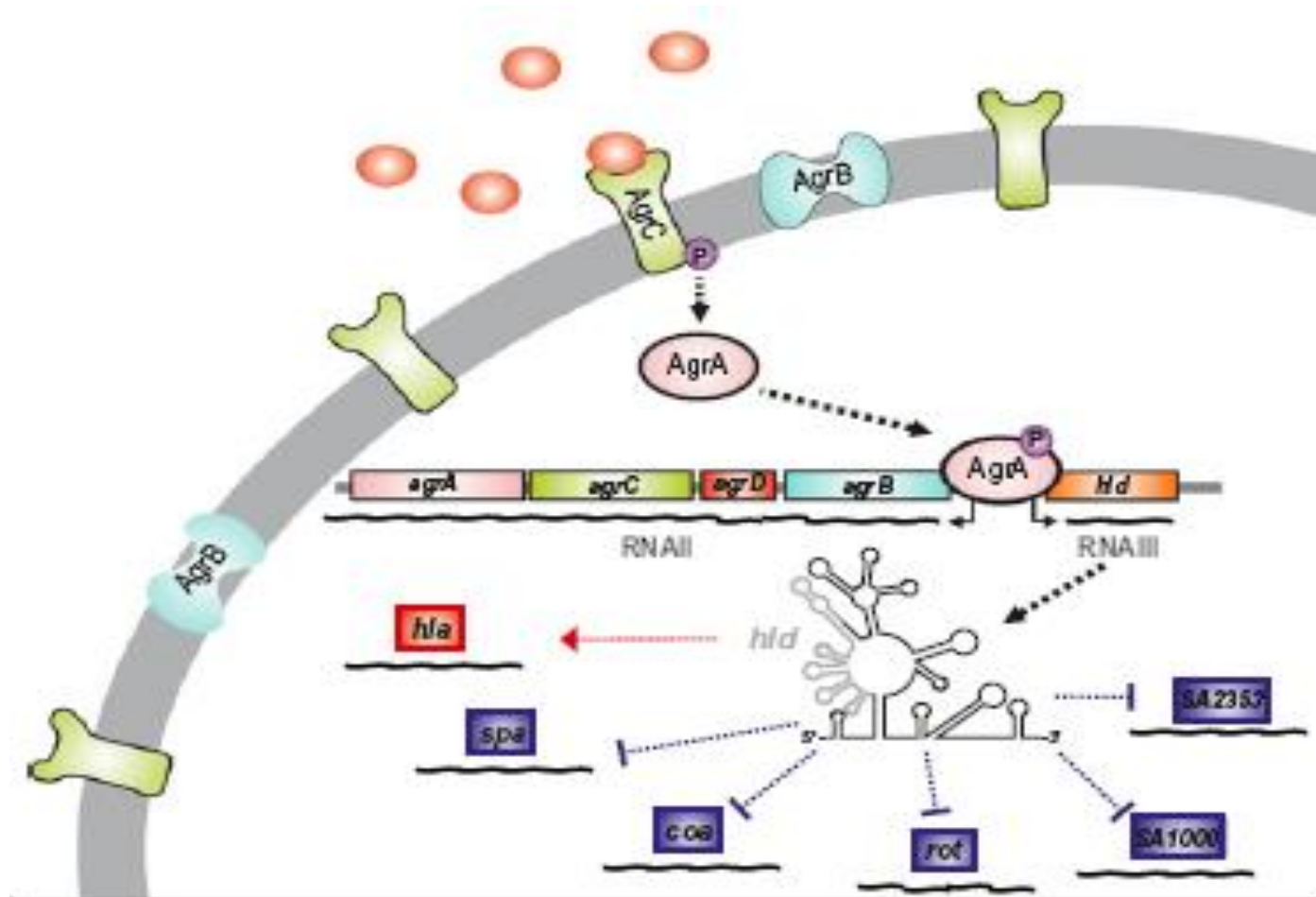
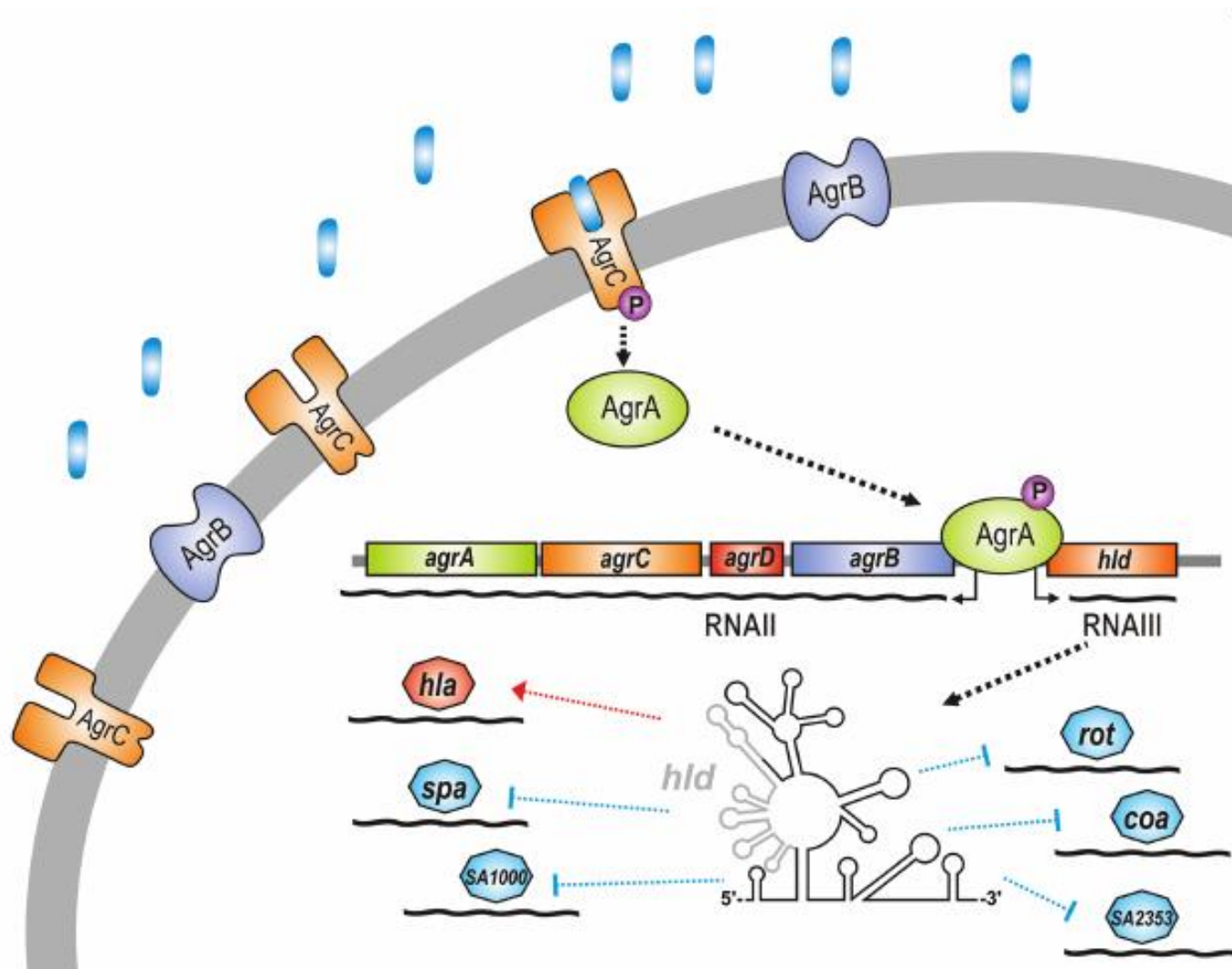


Figure 2. Quorum Sensing and RNAIII-Controlled Gene Expression

S. aureus produce an autoinducing peptide that accumulates in the medium and is sensed by a histidine kinase (AgrC). Sensing of the autoinducing peptide by AgrC leads to phosphorylation of the response regulator AgrA, which in turn is a transcriptional activator of the bifunctional RNAIII. RNAIII harbors the *hld* gene (coding for δ -hemolysin) but also acts as a posttranscriptional regulator of several target mRNAs, most of which with profound impact on virulence. While *spa*, *coa*, *rot*, SA1000, and SA2353 mRNAs are repressed, the *hla* mRNA is activated by RNAIII.

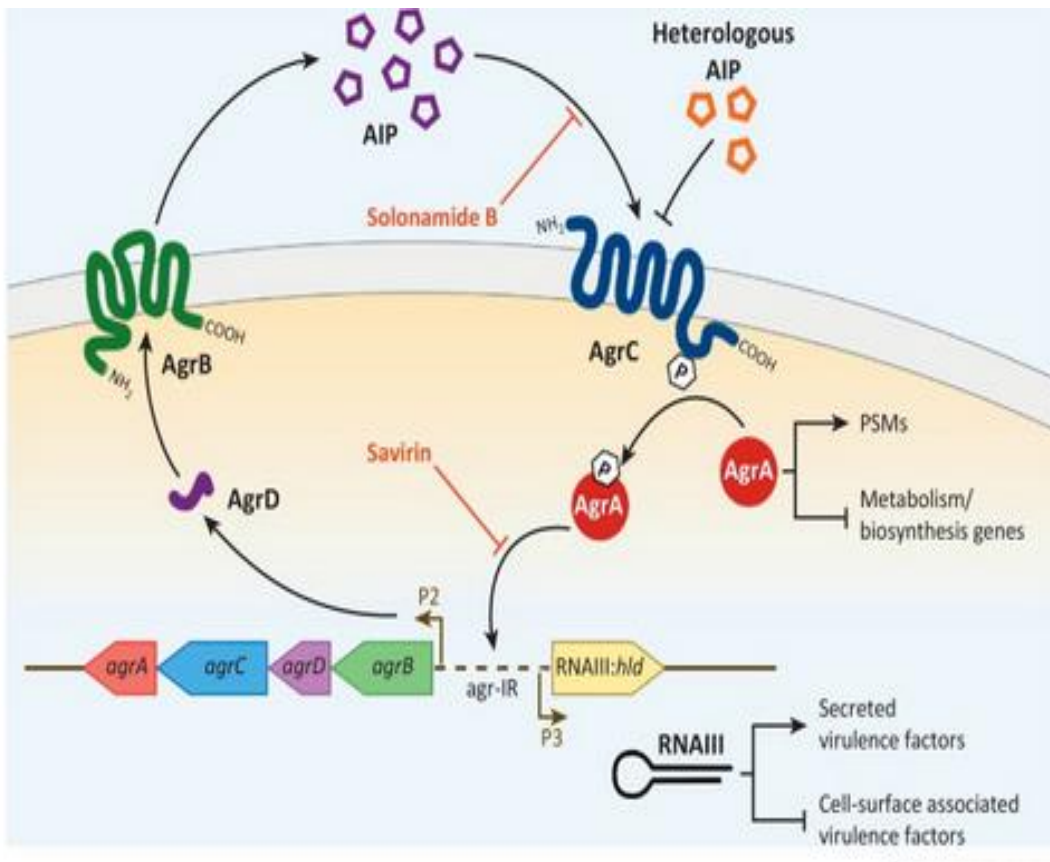


RnaIII una molecola tante funzioni

Staphylococcus aureus contiene il locus *agr* che codifica per un sistema di quorum sensing composto da due unità di trascrizione che sono trascritte da due operoni divergenti P2 e P3. P2 controlla l'espressione dei geni coinvolti nel QS (*agrBDCA*)

Da P3 nasce un trascritto bifunzionale: è un sRNA trans di 514 RNAIII e un RNA codificante per la δ emolisina.

L'espressione di RNAIII è regolata da AgrA che è il regolatore della risposta attivato in seguito a fosforilazione mediata dal sensore AgrC

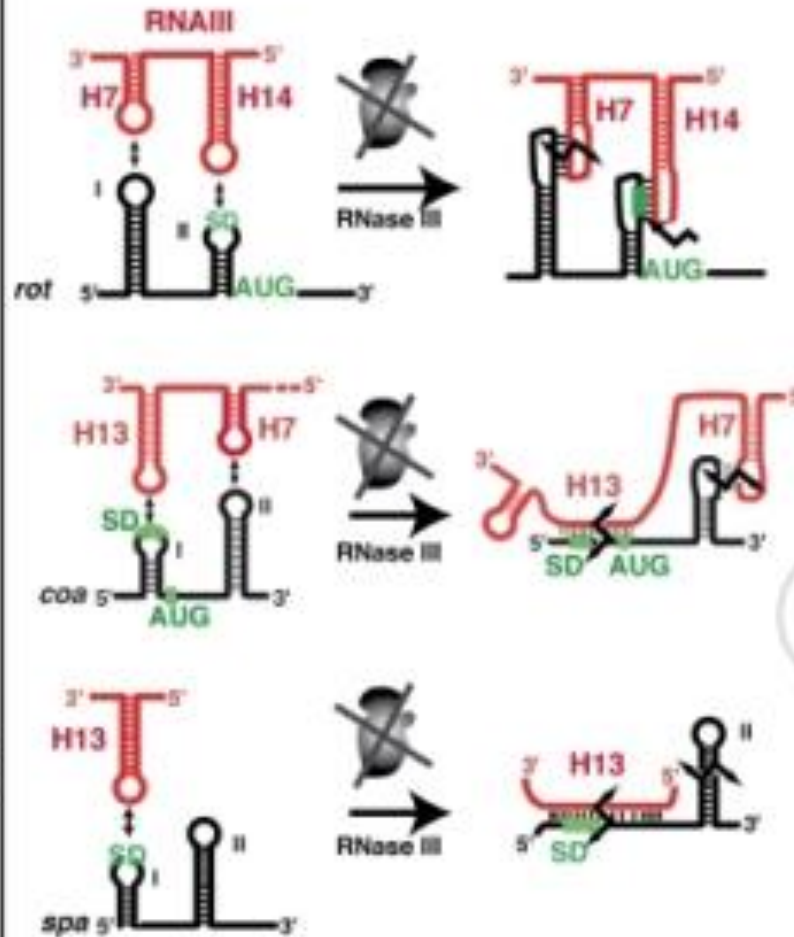
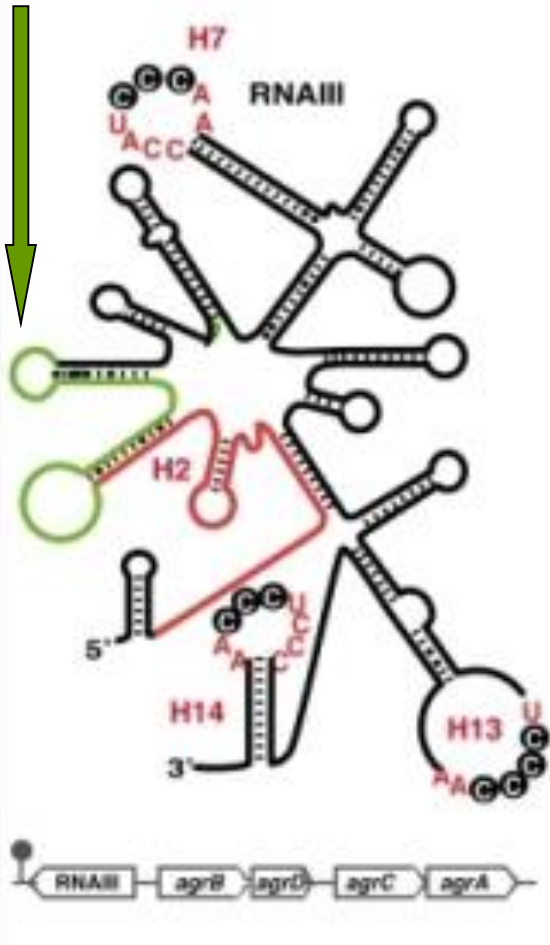


RNAIII può funzionare sia come attivatore che come repressore:
Regola l'espressione del trascritto per l'emolisina (hla mRNA)
legandosi ad una regione a monte che in genere sequestra la regione
SD inibendo la traduzione,
Il legame di RNAIII rilascia la sequenza SD di hla
e la traduzione può aver luogo.

La struttura di RNAIII è piuttosto complessa:
La regione centrale è codificante,
tramite la regione al 5' regola *hla*
tramite la regione 3' regola negativamente il
regolatore *rot*, la proteina che lega il
fibrinogeno, e altri fattori di virulenza.



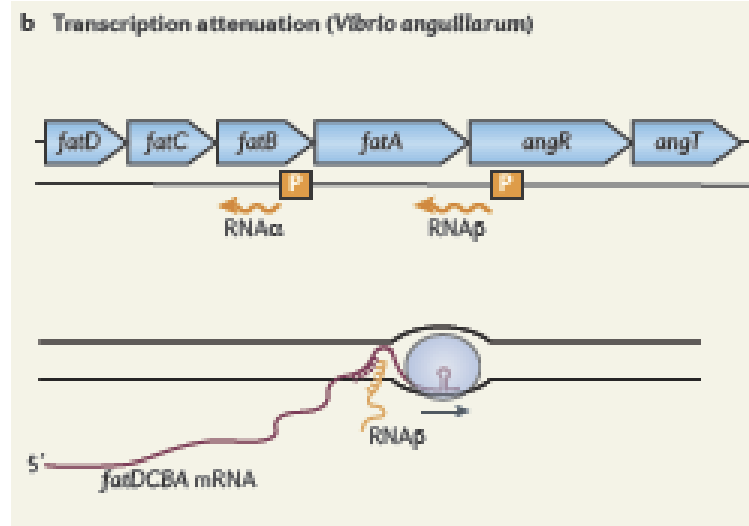
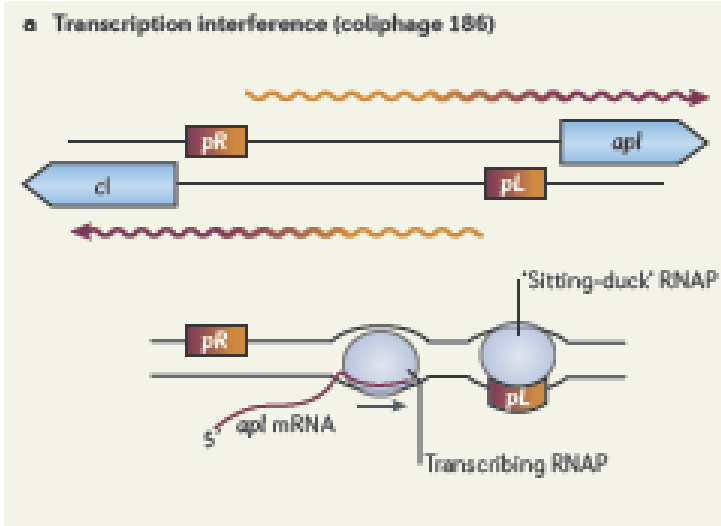
Sequenza codificante



Nel caso di *rot* RNAIII si lega al trascritto tramite 3'UTR (H14) e un dominio centrale (H7), inibendo la trascrizione e promuovendo la degradazione RNasi mediata

Nel caso di *coa* l'interazione di RnaIII avviene con SD e con parte della codante e promuove degradazione del trascritto e di RNAIII stesso

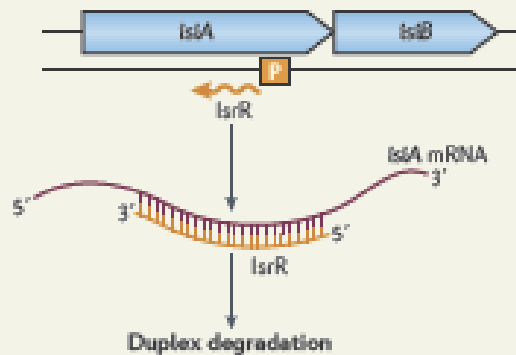
Gli sRNA antisense : diverse strategie d'azione



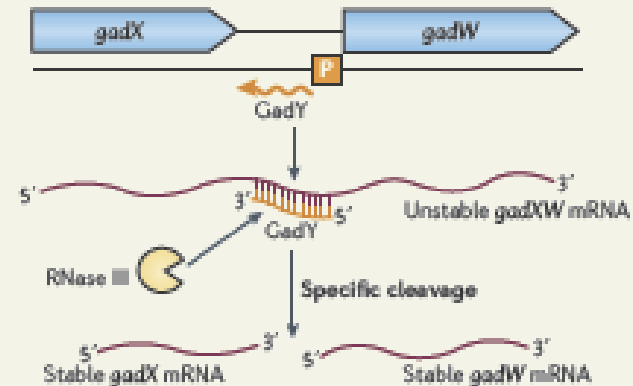
Regolazione della trascrizione

- a) Interferenza trascrizionale
b) Attenuazione della trascrizione

c) Degradation of sense transcript (*Synechocystis* sp. PCC 6803)



d) Stabilization of sense transcript (*Escherichia coli*)



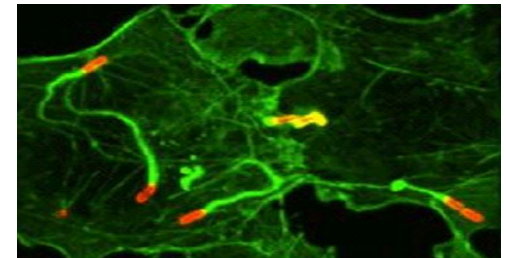
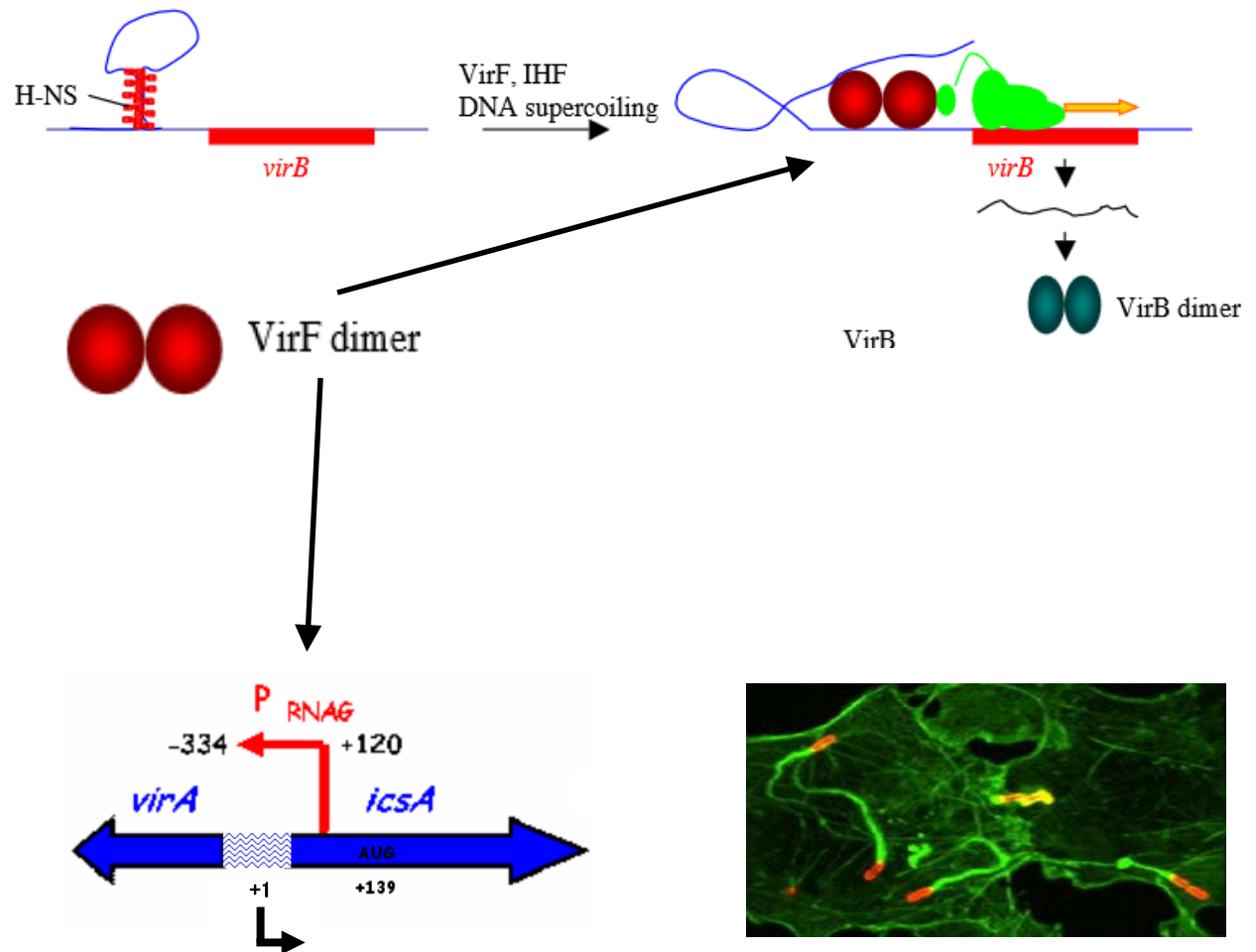
c) Degradazione del trascritto senso

d) Taglio e stabilizzazione del trascritto policistronico

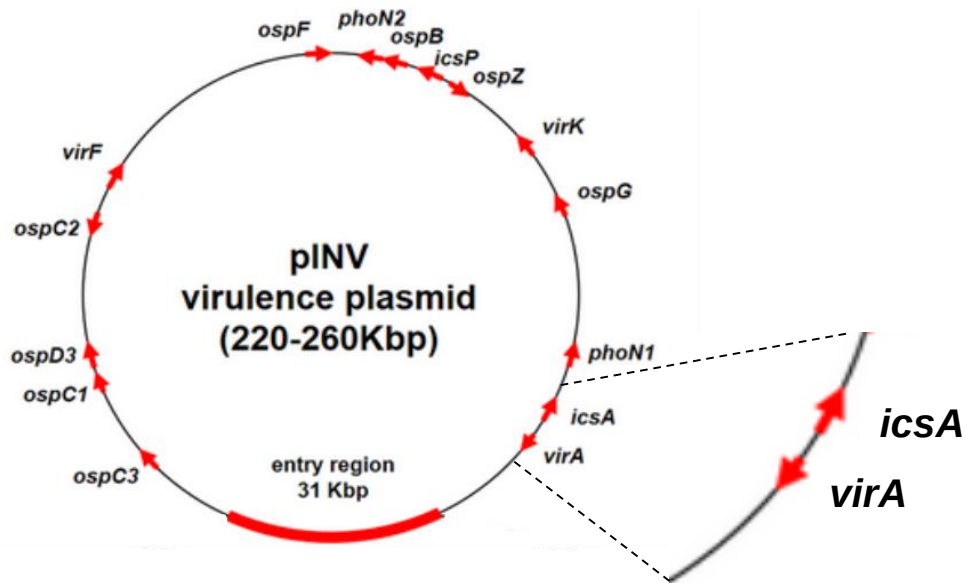
VirF as regulator of virulence genes

VirF ...

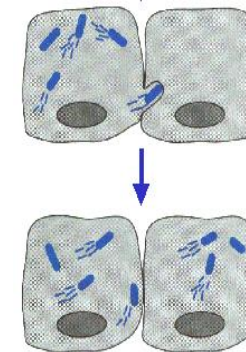
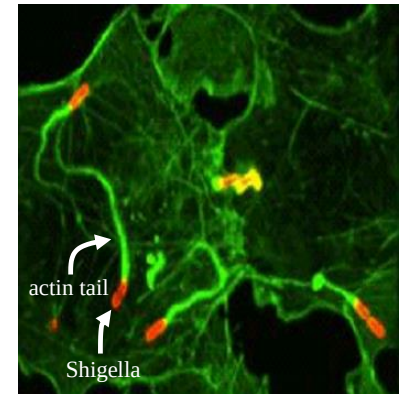
- acts as anti H-NS factor on the P_{virB} promoters
- is required for the *virB* expression
- controls the expression of *icsA* (encoding the protein involved in *Shigella* movement)



Small RNAs are known key regulators of virulence gene expression in bacteria.
Does this apply also to *Shigella*?



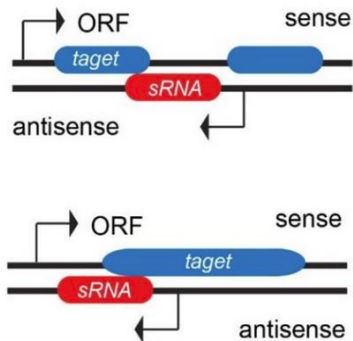
icsA encodes a large protein (110KDa) which induces polymerization of F-actin at bacterial cell poles



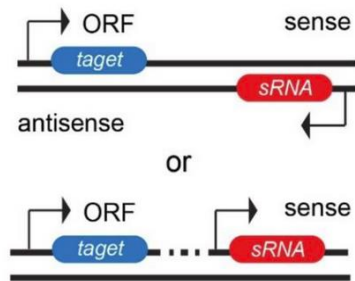
IcsA is responsible for the movement of the bacterium through the host cytoplasm and into adjacent cells.

RnaG: a small RNA transcribed from the complementary strand of the *icsA* mRNA

Cis-encoded small RNAs

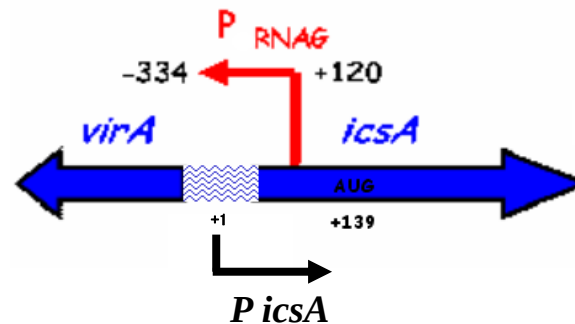


Trans-encoded small RNAs

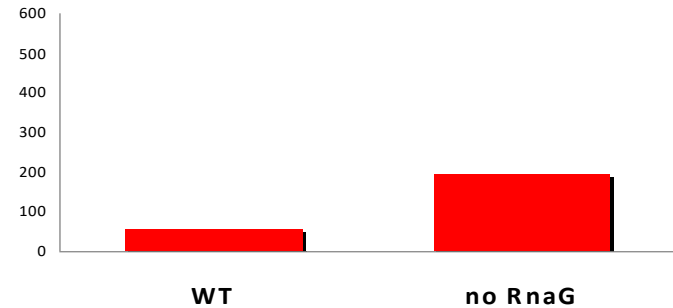
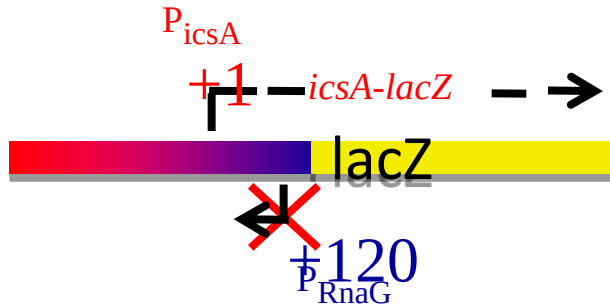


RnaG ...

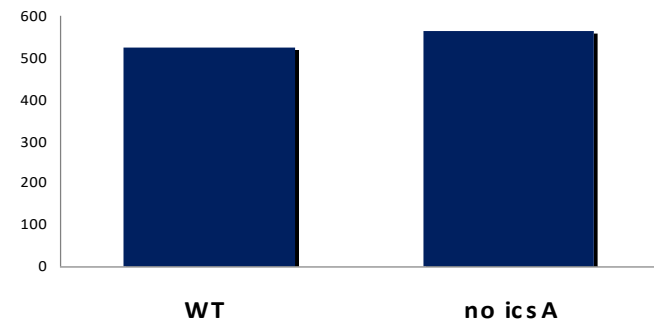
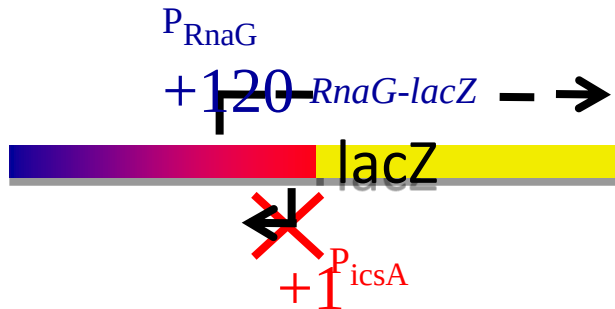
- ... is a non coding sRNA;
- ... is actually not that small: 450 nucleotides (+ 120 to -334);
- ... is the first pINV-encoded sRNA found in *Shigella*.



RnaG downregulates *icsA* transcription



Silencing of the P_{RnaG} promoter induces a 3.5-fold βgal increase in *icsA-lacZ* fusions

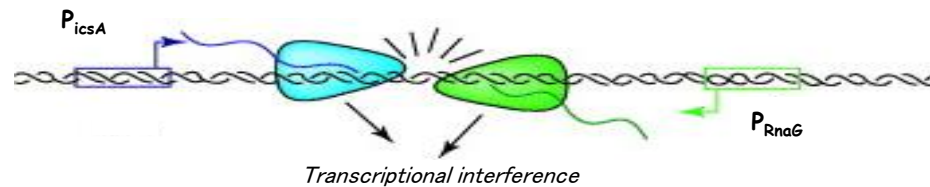
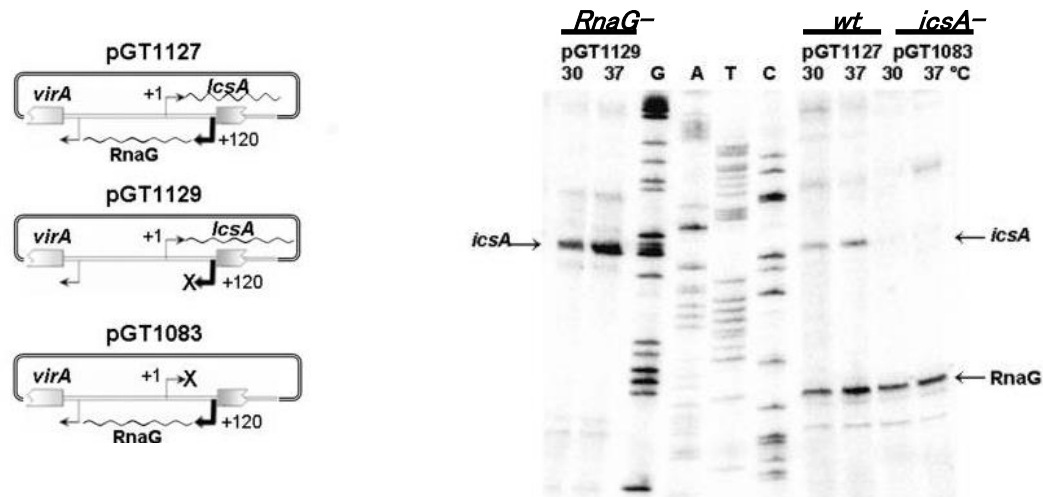


Silencing of the P_{icsA} promoter induces only a slight βgal increase in *RnaG-lacZ* fusions

P_{RnaG} is stronger (3-fold) as compared to P_{icsA}

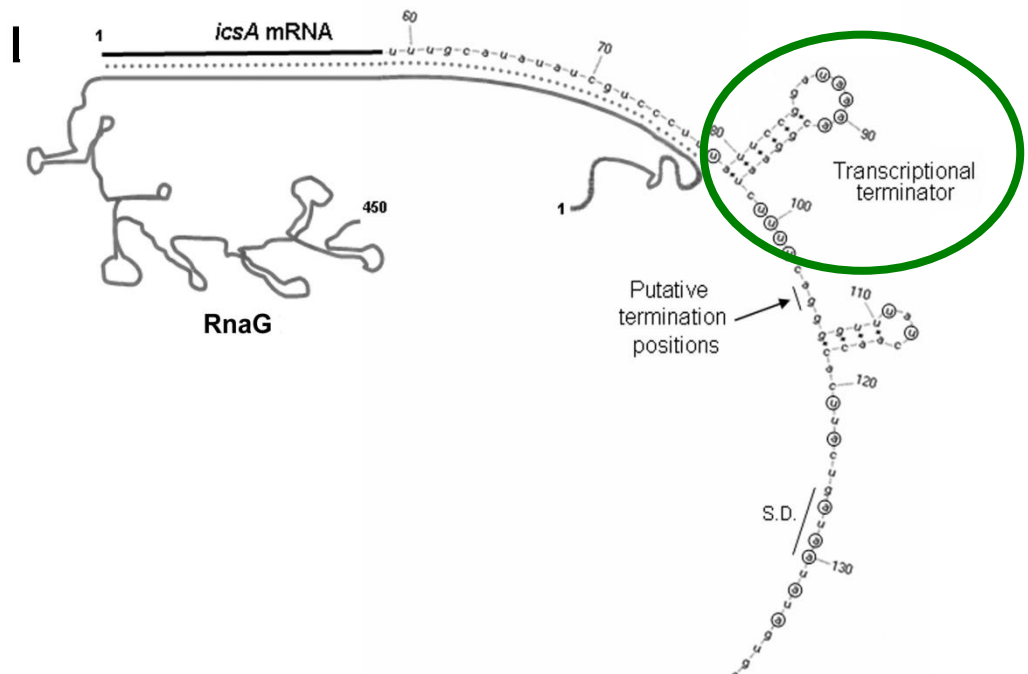
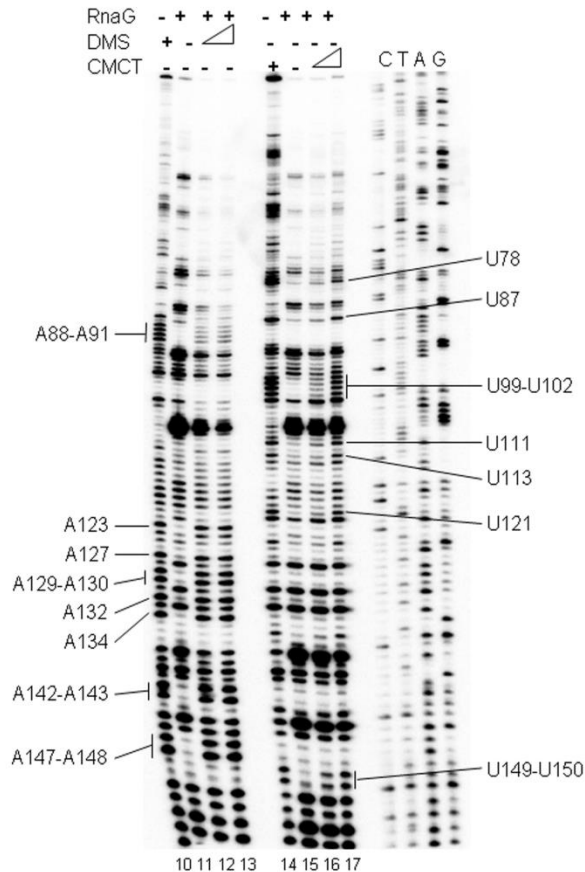
RnaG downregulates *icsA* transcription by two independent mechanisms:

1. Transcriptional interference from P_{RnaG} on P_{icsA}



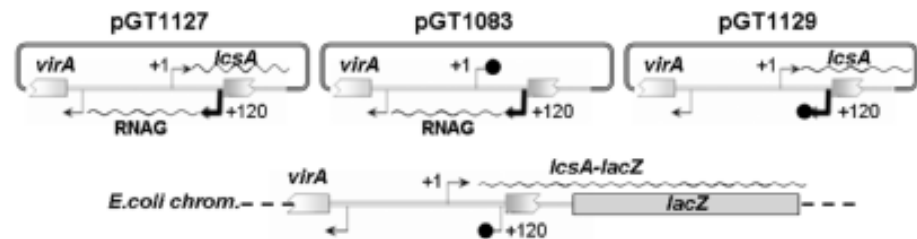
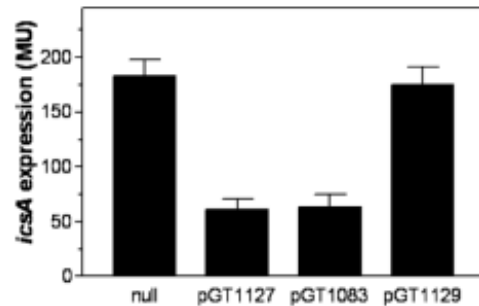
Upon addition of RnaG ...

- AH1 and AH2 hairpins are unfolded
- *icsA* mRNA adopts an alternate structure which leads to premature termination

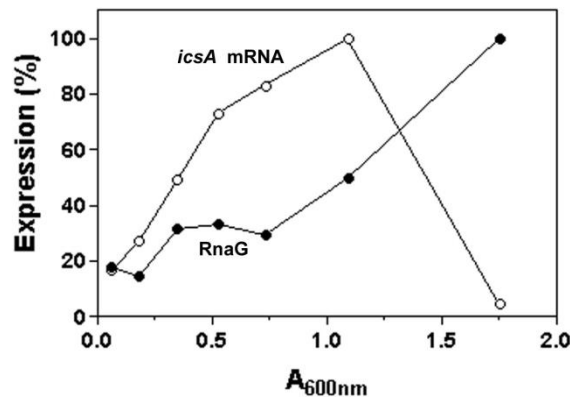


RnaG hampers *icsA* expression

In vivo



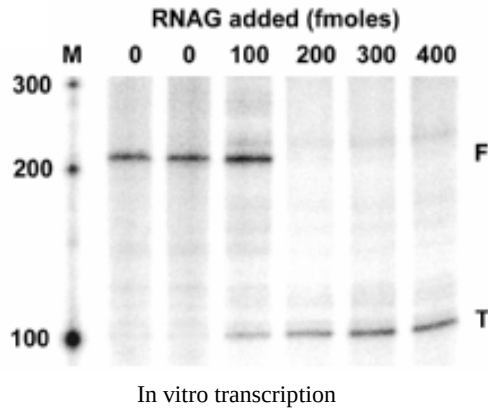
The introduction of plasmids encoding RnaG decreases the expression of a P_{icsA} -*lacZ* fusion.



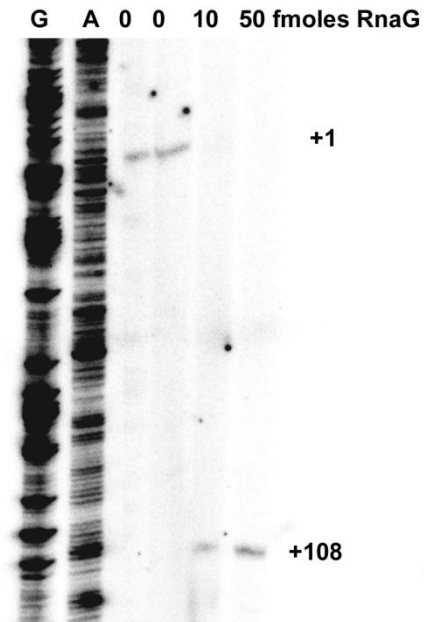
Approaching the stationary phase the increase of RnaG expression is immediately followed by an abrupt decrease of *icsA* expression

In vitro

The addition of RnaG molecules induces:



1. the disappearance of the *icsA* transcript
2. the appearance of an abortive short transcript, suggesting the presence of a potential terminator

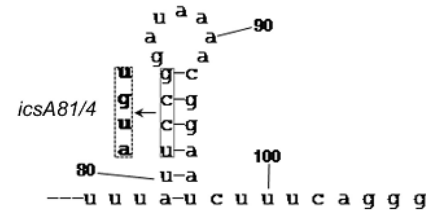


Primer extension

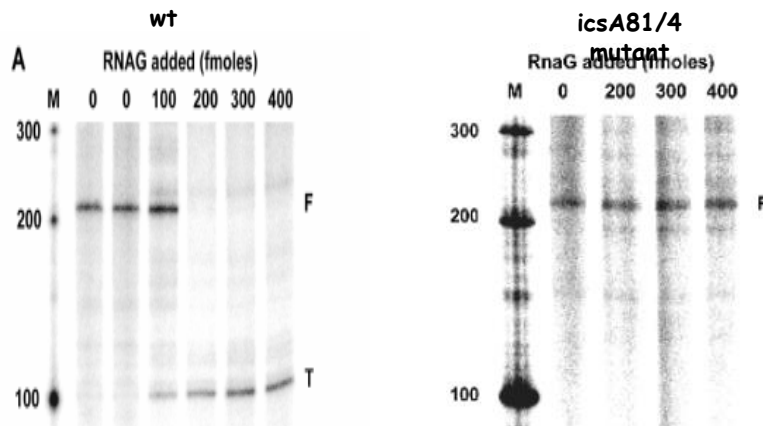
3. the displacement of the start point signal of *icsA* from +1 to +108, suggesting the formation of a RnaG-mRNA duplex

Towards a model for transcriptional attenuation

If the structure of the transcriptional terminator is altered ...

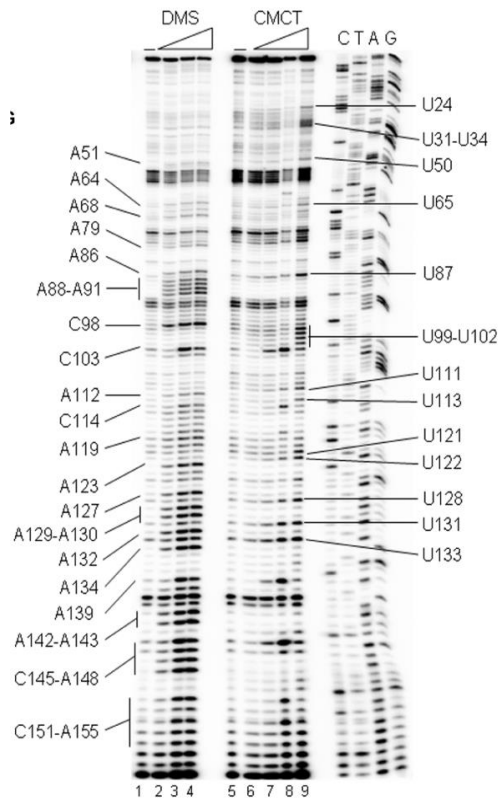


icsAm RNA in the presence of RnaG

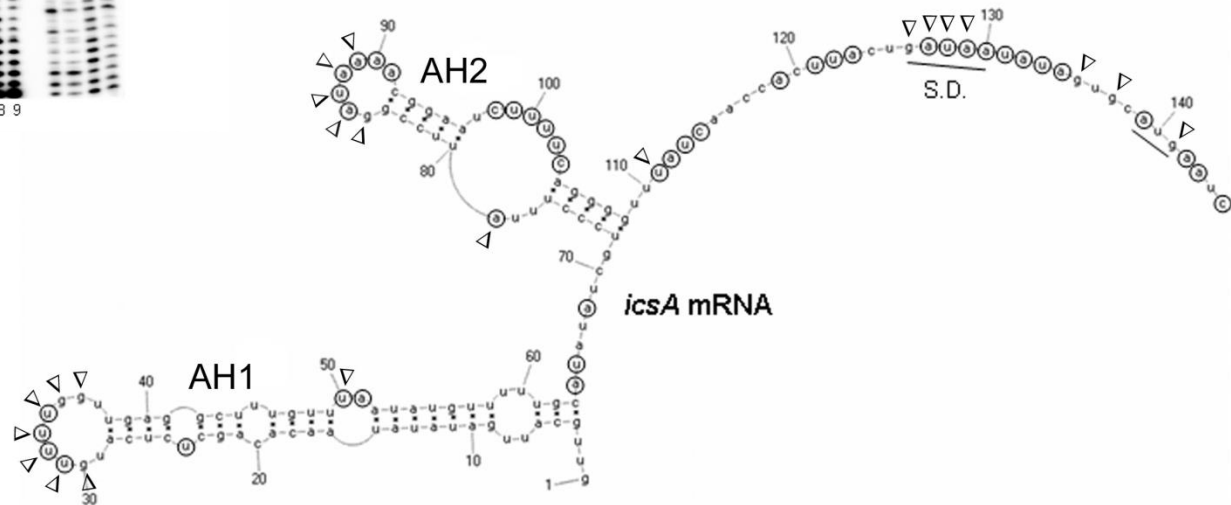


In vitro transcription in presence or absence of Rna G molecules

... RnaG becomes unable to repress *icsA* transcription in mutants harboring a defective transcription terminator

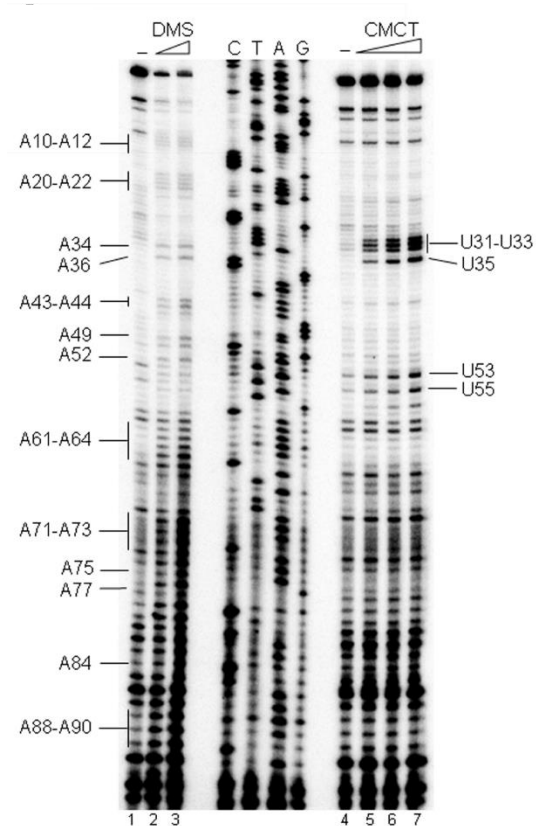
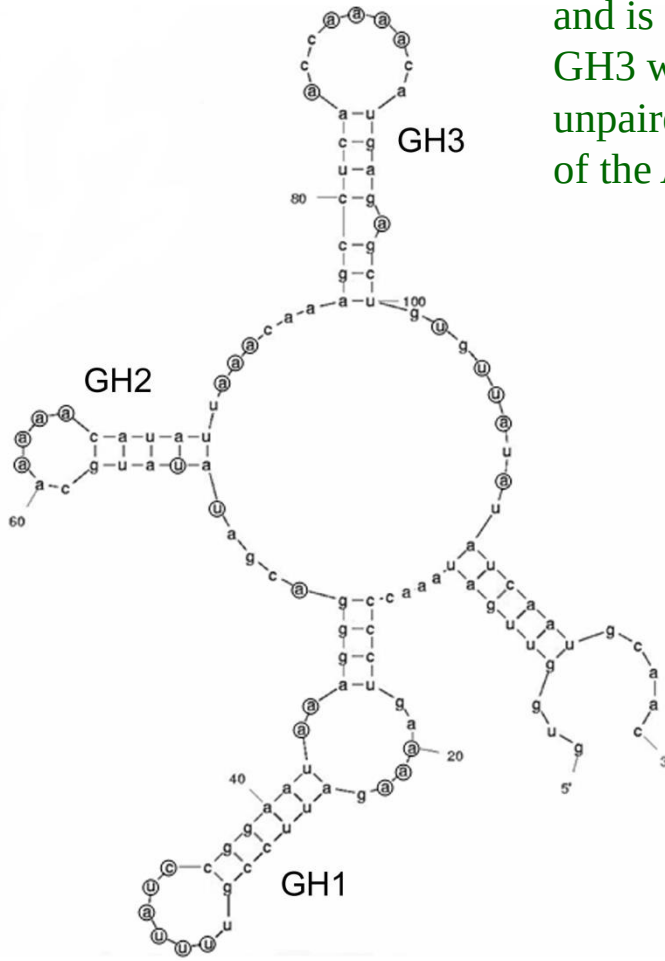


RNA probing reveals the presence of two hairpin-loop structures (AH1 and AH2) in the 5' UTR *icsA* transcript



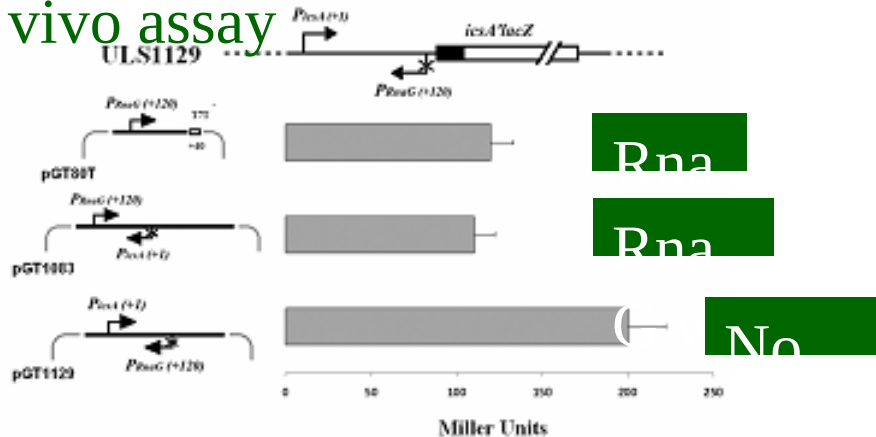
Secondary structure of the antisense region of RnaG

The antisense region (120nt at the 5' end) is highly structured and is characterized by three stem-loop motifs GH1, GH2 and GH3 which might associate with unpaired bases in the apical loop and in the basal bulge of the AH1-AH2 structures of the *icsA* transcript



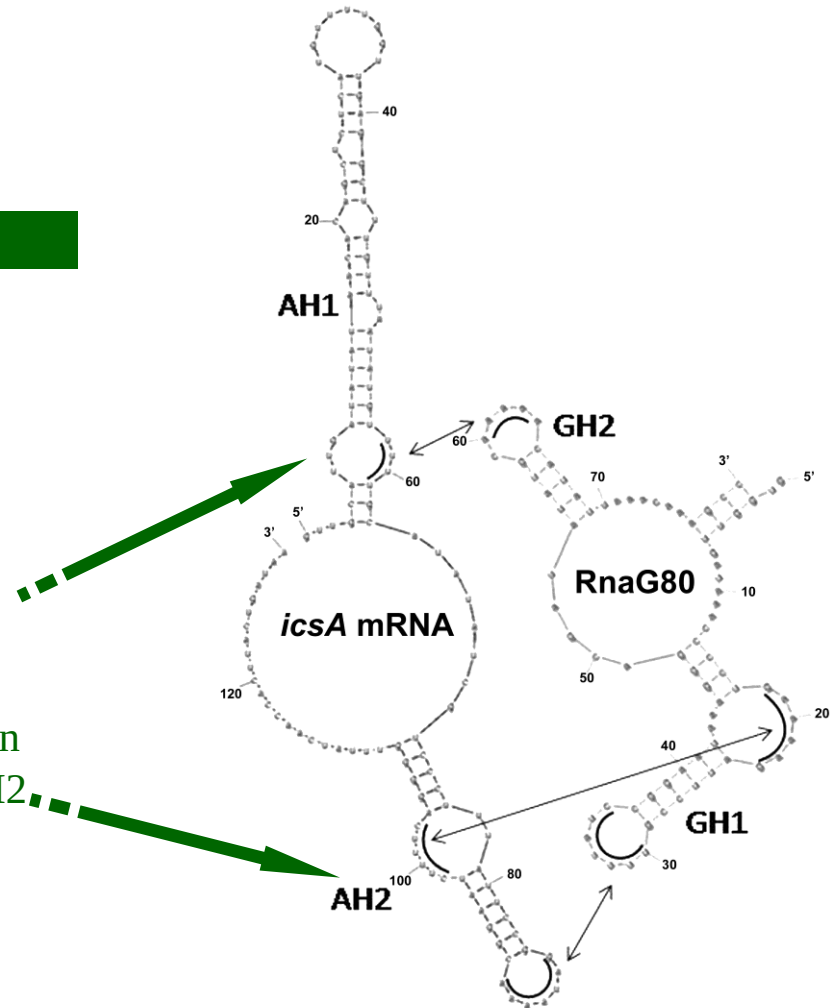
Out of the 120 nt at the 5' end, the initial 80 represent the minimum stretch able to repress *icsA* transcription

In vivo assay

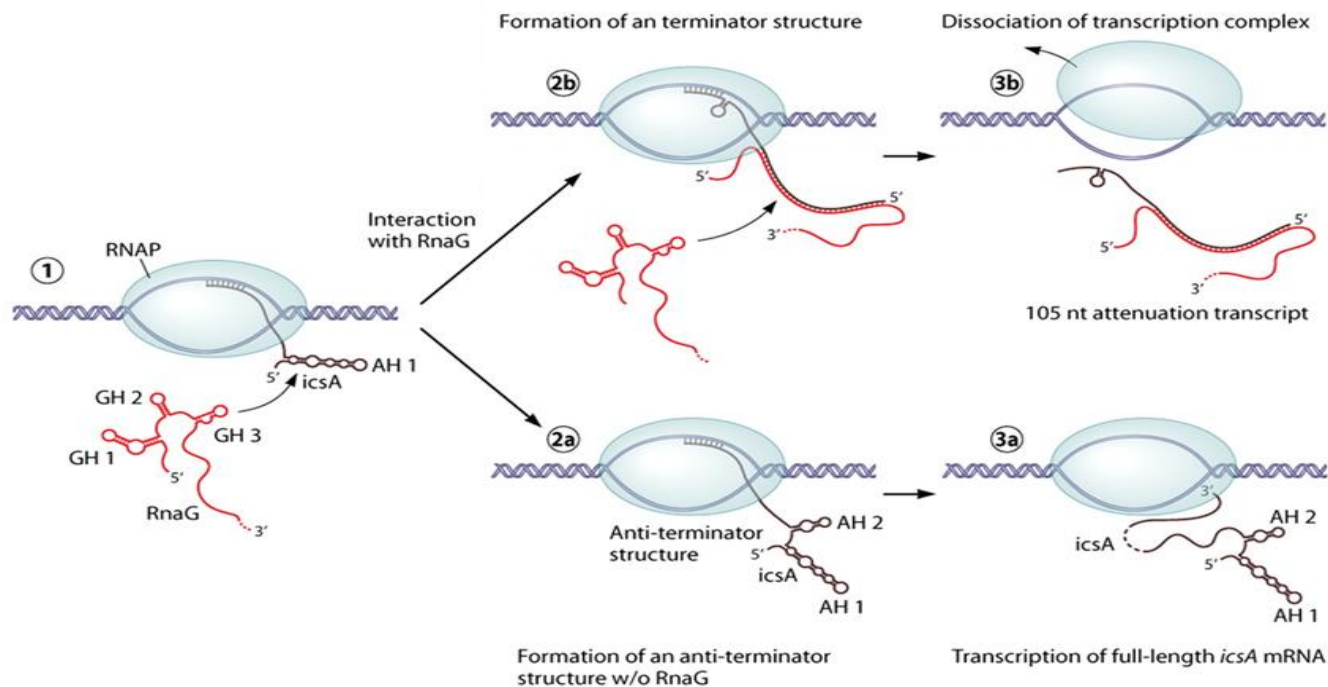


Using mutated RnaG molecules we observed :

1. The apical loop of GH2 can pair with the basal buldge of AH1
2. The apical loop and the internal bulge of GH1 can anneal with the structural similar motifs of AH2



Model for the involvement of RNAG in transcriptional attenuation



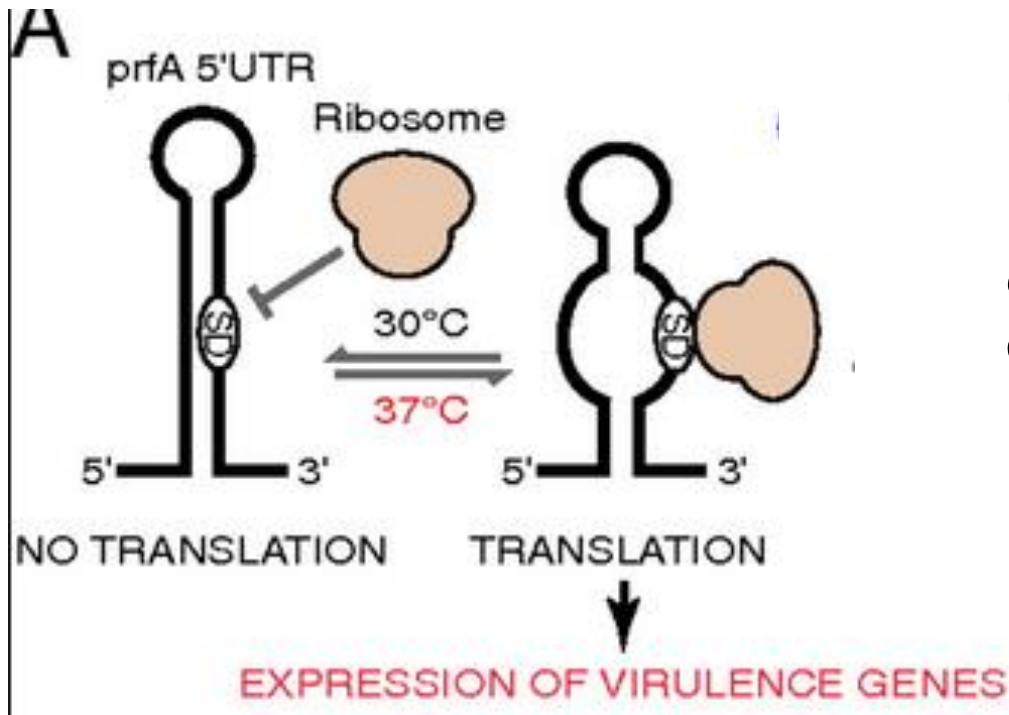
La regione 5'UTR

Regione molto importante per la regolazione nei mRNA è la regione 5'UTR

Regione variabile da poche basi a centinaia di basi

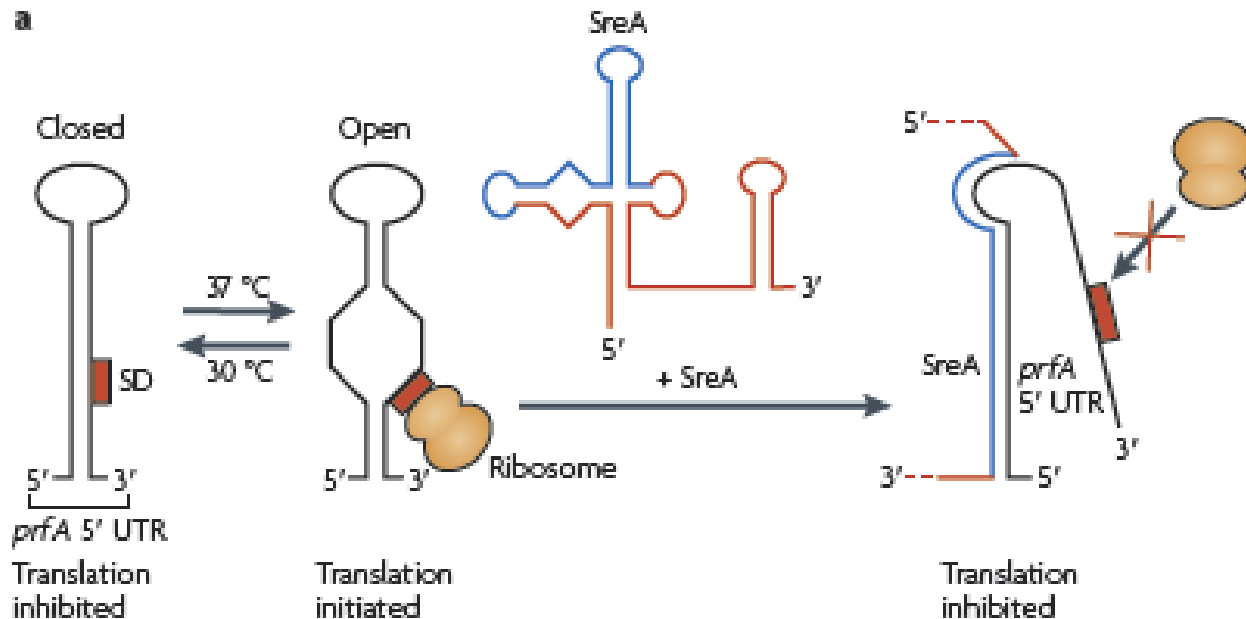
La regione 5'UTR può essere utilizzata per modificare l'espressione genica in base a cambi di T, pH o presenza di metaboliti.

RNA Termosensori : il caso di *prfA*



PrfA è un gene importante per la patogenicità di *Listeria monocytogenes* in quanto codifica per un attivatore trascrizionale di molti geni di virulenza

La regione 5'UTR di 116 nucleotidi forma una struttura secondaria a bassa T°C mascherando la sequenza SD ed inibendo la traduzione. A 37°C si forma una struttura alternativa che espone la SD e permette la traduzione di *prfA* che codifica per attivatore trascrizionale dei geni di virulenza.

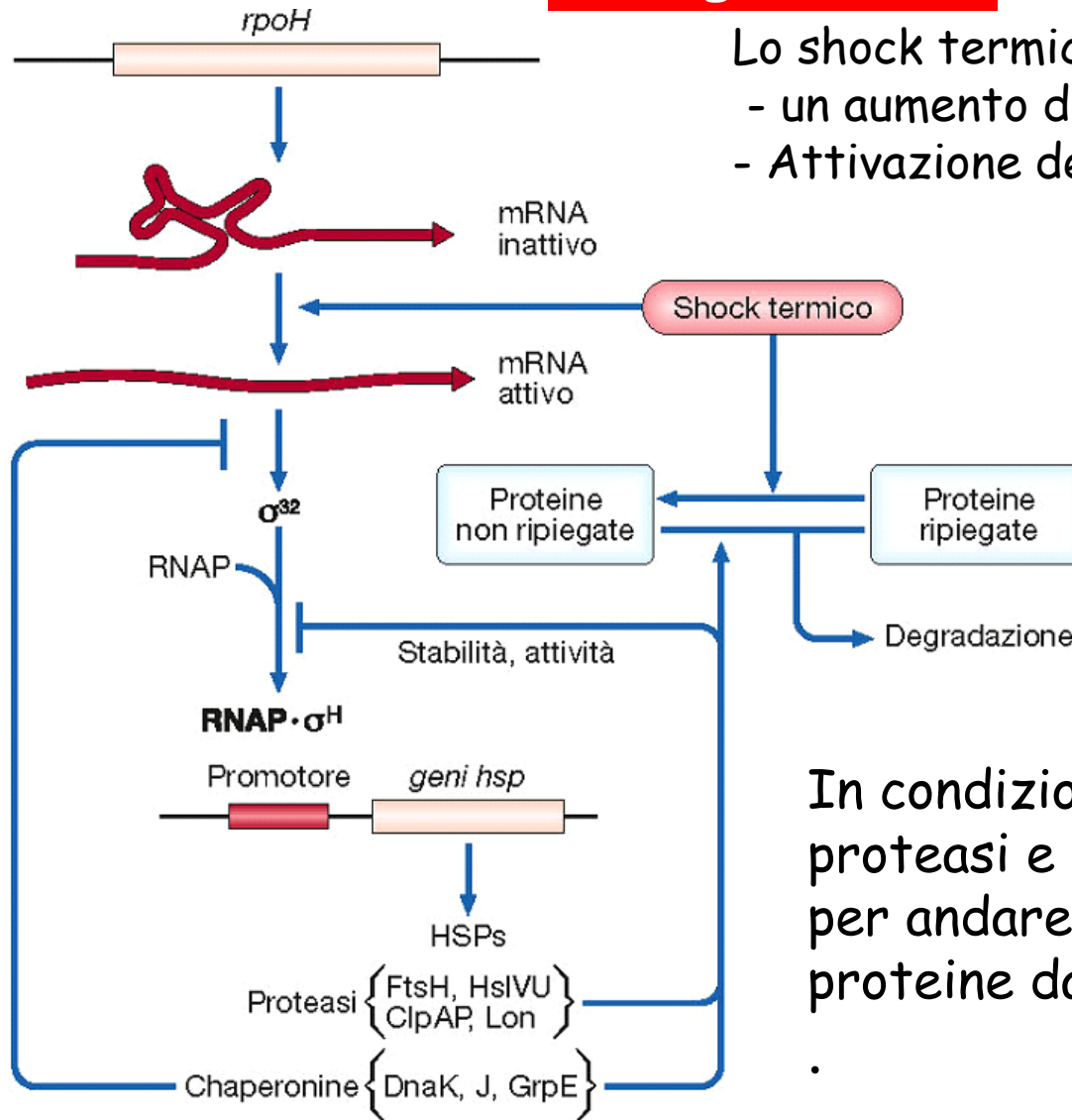


La regolazione di *prfA* è controllata ulteriormente da un piccolo RNA trans che legandosi alla regione 5'UTR ne modifica la struttura provocando instabilità dei ribosomi alla SD e diminuzione dei livelli dell'espressione di *prfA*

Il regulone σ^H

Lo shock termico induce sia

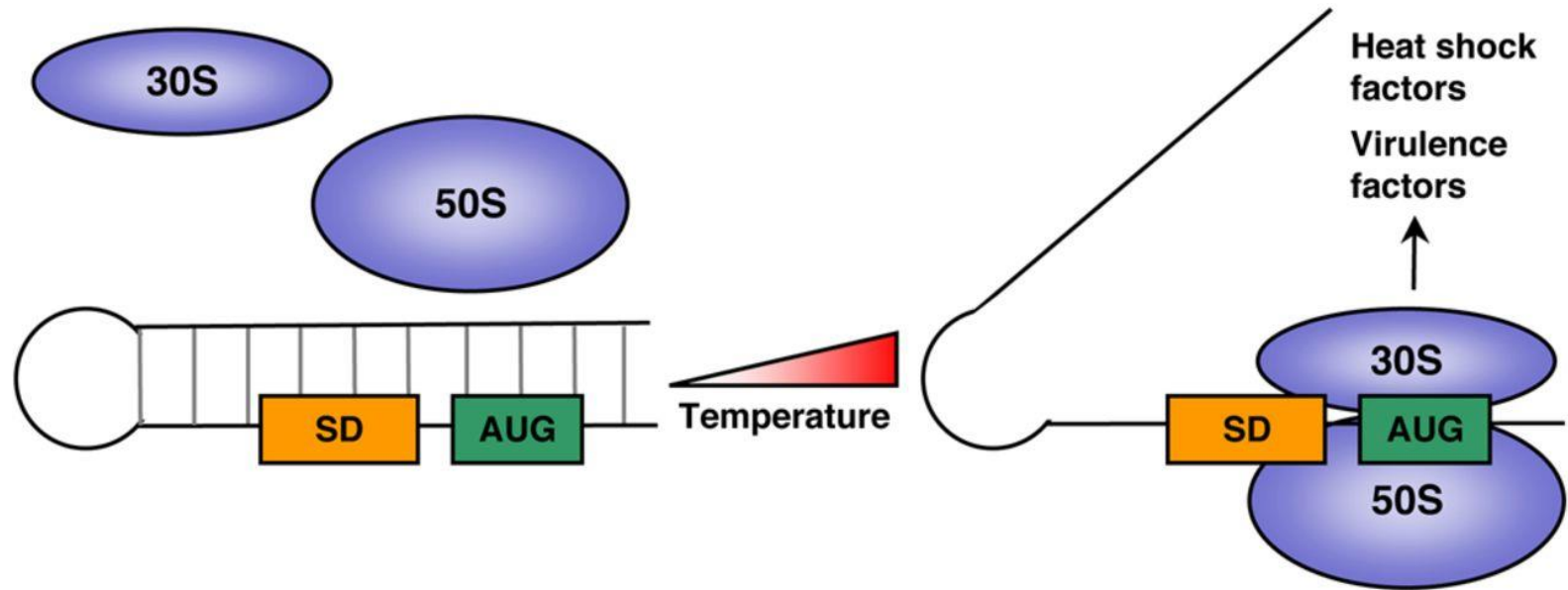
- un aumento della stabilità di σ^H
- Attivazione della traduzione del mRNA di σ^H



A bassa temperatura il livello di σ^H è mantenuto basso ad opera di proteasi che lo degradano e di chaperonine (DnaK-DnaJ-GrpE) che interagendo con σ^H gli impediscono di legarsi alla RNA polimerasi.

In condizione di shock termico le proteasi e ciaperonine liberano σ^H per andare a degradare o rifoldare le proteine danneggiate.

.

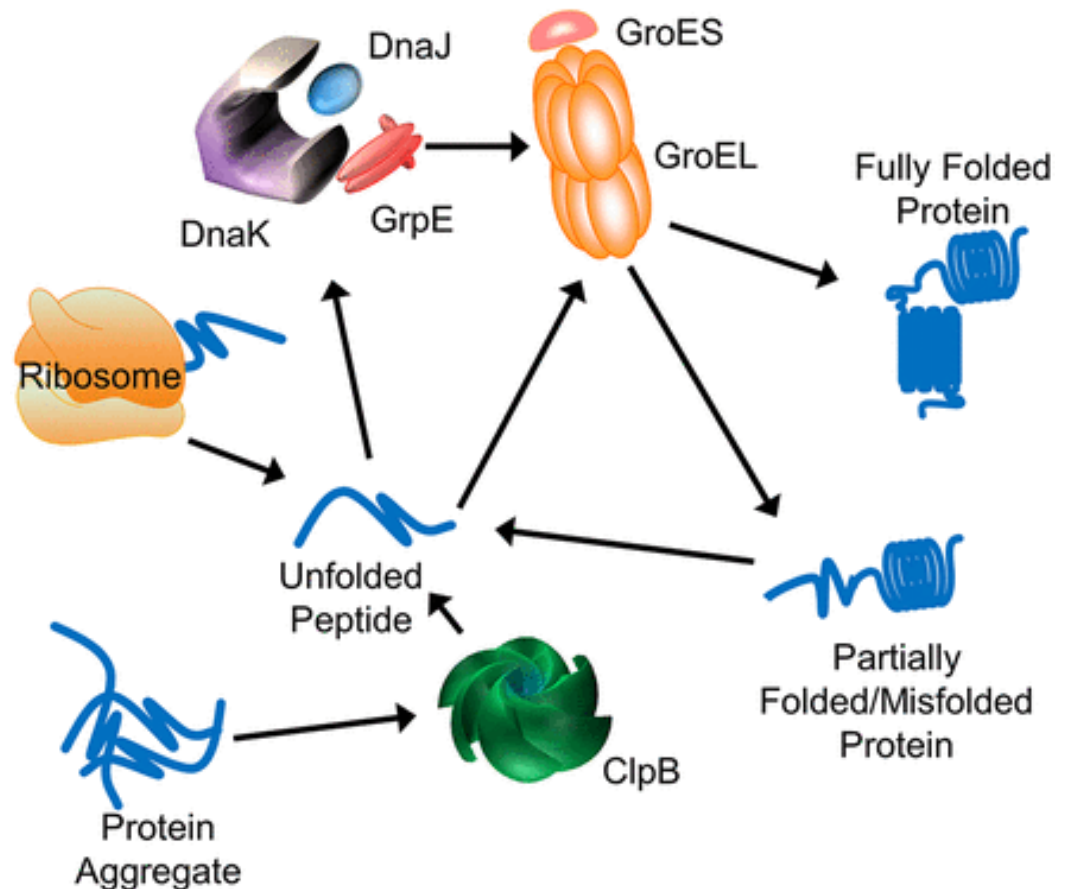


La traduzione del mRNA di σ^H sarebbe facilitata dalla perdita di una struttura secondaria del mRNA (5'UTR) che funge da termosensore

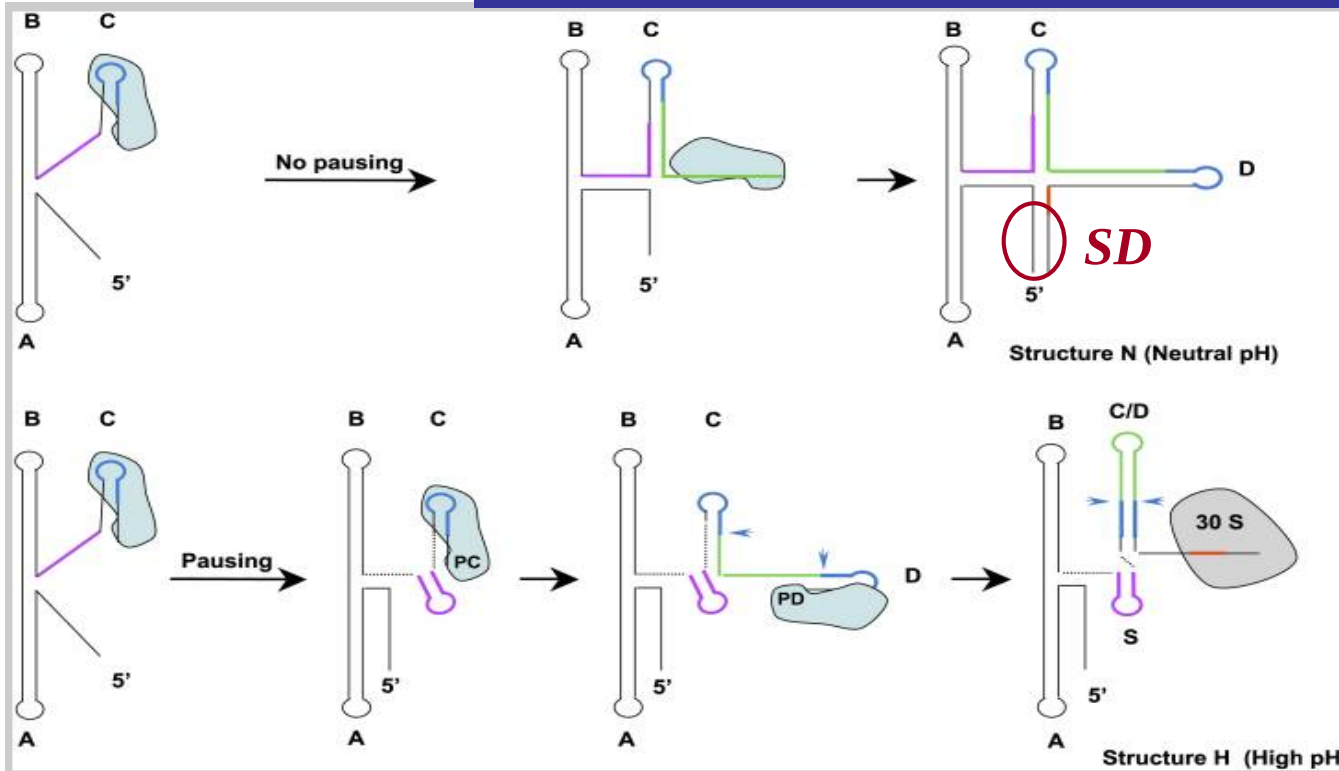
In assenza di shock termico i geni per le proteine heat shock sono espressi a basso livello

Rapido aumento della sintesi dopo stress

Spegnimento della espressione circa 10 minuti dopo cessazione stimolo.



Regolazione in risposta al pH



pH neutro OFF

pH alcalino ON

Un trasportatore implicato nella resistenza alla tellurite viene espresso solo in condizioni alcaline (gene *alx*) in *E. coli*.

A pH normale la trascrizione della regione 5'UTR procede ininterrotta formando una struttura inattiva per la traduzione.

In condizione alcaline si viene formare una struttura alternativa dovuta al rallentamento della RNA pol. in due diversi siti del 5' UTR.

A model for PRE-*alx* regulation by alkaline conditions. Under conditions of normal growth, PRE-*alx* forms a translationally inactive structure (structure N), whereas under extreme alkaline conditions when folding is kinetically controlled by RNA polymerase pausing, the newly synthesized RNA adopts an active structure (structure H) in which the ribosome-binding site (in red) is exposed for 30S ribosomal subunit binding. Formation of the active structure involves pausing of RNA polymerase at two sites located at hairpin C (PC) and hairpin D (PD). Upon pausing of RNA polymerase (shaded in light blue) at PC, the formation of hairpin C is prevented, and instead formation of hairpin S (indicated by a purple line) is promoted. Pausing at PD inhibits the formation of hairpin D, allowing base-pairing between the complementary sequences in hairpins C and D (in blue) and formation of the apical loop of hairpin C/D (in green).

Open issues

How does RnaG-mediated regulation integrate into other virulence gene systems?

What is the role of the 3' region of RnaG?

Are there other sRNAs-encoding genes on the pINV plasmid of *Shigella*?

CRISPR

Clustered Regularly Interspaced Short Palindromic Repeats



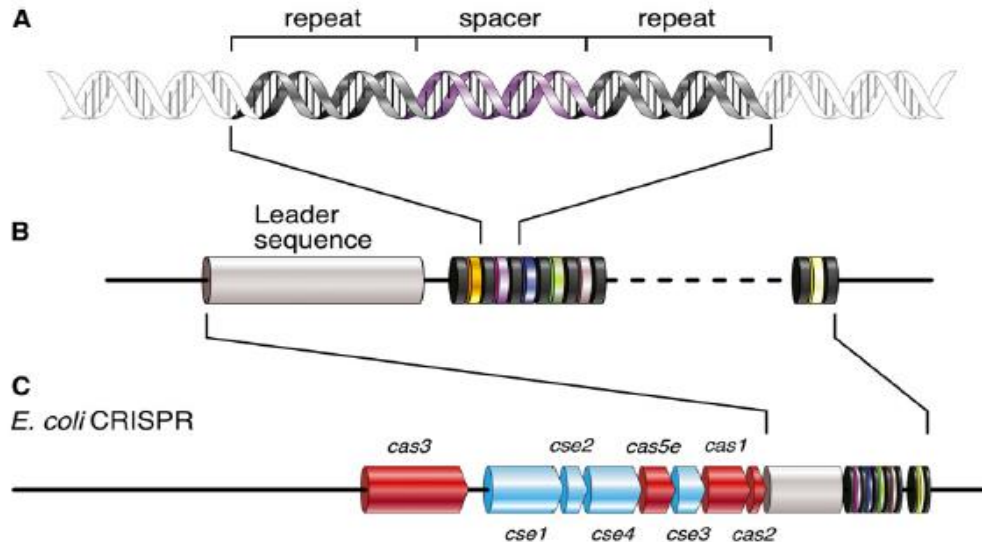
Le regioni CRISPR sono essenzialmente delle banche di memoria di sequenze fagiche ostili.

UNA CRISPR è costituita dalla ripetizione di diversi frammenti fagici (SPACERS) alternate con sequenze ripetute identiche (REPEATS) .

Il sistema CRISPR conferisce resistenza ai fagi che contengono nel proprio genoma una sequenza identica o strettamente correlata a quella contenuta nei CRISPR.

Sono quindi regioni presenti nel genoma della gran parte dei batteri (70% nei Batteri, 90% degli Archea)

Elementi caratteristici di un locus CRISPR



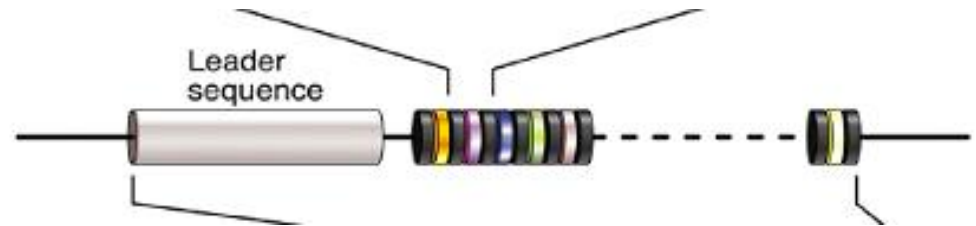
In genere si ritrova un locus CRISPR per genoma anche se non è raro trovare genomi con diversi CRISPR loci.

Un locus CRISPR è caratterizzato dall'allineamento di sequenze ripetute **REPEATS** praticamente identiche all'interno di un locus separate da sequenze **SPACERS** di lunghezza identica ma di contenuto in basi molto diverso.

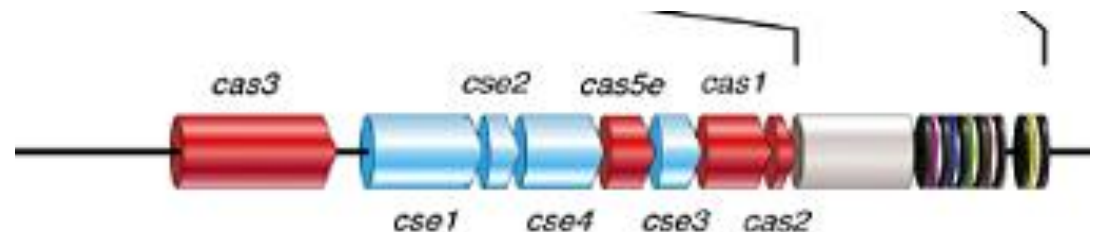
I **REPEATS** variano in lunghezza da specie a specie dalle 21 alle 47 basi (con una media di 32 basi) mentre la lunghezza degli **SPACERS** varia tra i 20 e 72 bp. I **REPEATS** sono caratterizzati dalla capacità di formare un'ansa (sono sequenze parzialmente palindromiche). Il numero di unità ripetute di **SPACERS/REPEATS** per ogni locus è di circa 25 (con oscillazioni da 1 a 300!)

Oltre ai REPEATS and SPACERS il locus CRISPR è caratterizzato:

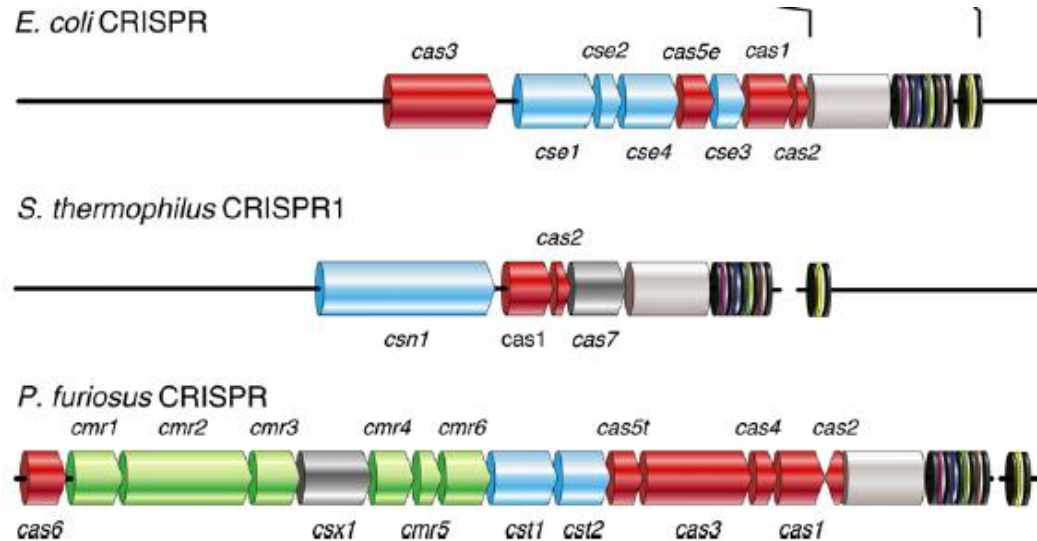
- dalla presenza di una regione definita LEADER lunga diverse centinaia di basi e non codificante posizionata sempre e solo ad una estremità del locus.



- Dalle proteine associate ai CRISPR (CAS) che sono codificate da geni localizzati a monte delle sequenze CRISPR



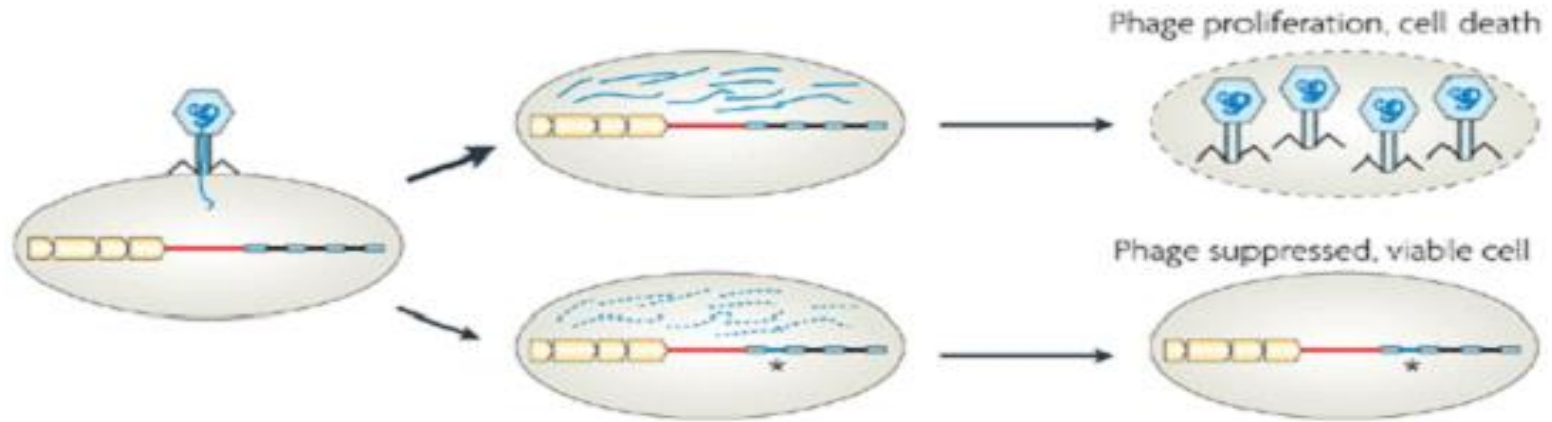
Le proteine CAS (CRISPR Associated Proteins)



I CRISPR sono caratterizzati dalla presenza di una serie di proteine CAS che variano in numero e disposizione da locus a locus.

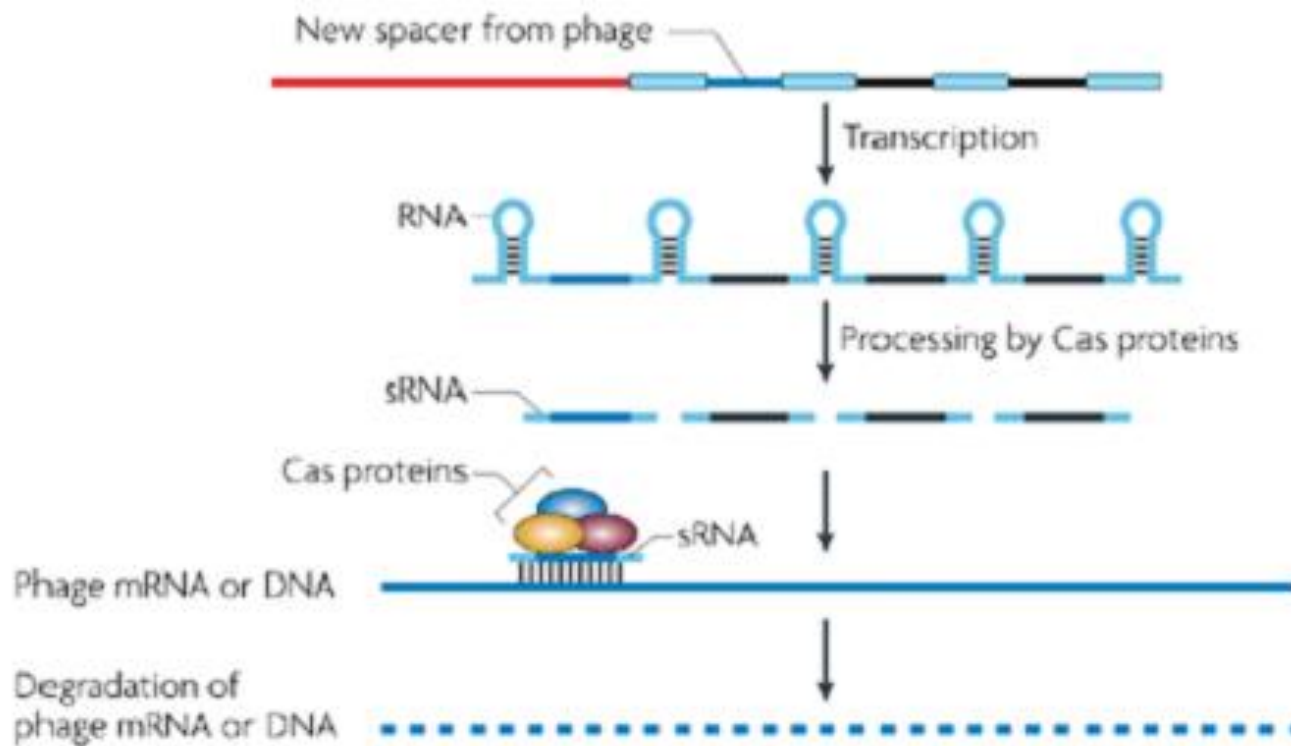
L'analisi comparativa di diversi loci ha permesso di identificare 4 tipi di proteine CAS (da 1 a 4) che costituiscono il core del sistema anche se spesso non sono tutte presenti.

Come fanno le regioni CRISPR ad acquisire le regioni di DNA (SPACERS) che si intervallano tra le sequenze ripetute (REPEATS)?



In seguito all'attacco da parte di un fago, gli acidi nucleici del fago si moltiplicano nella cellula e vengono prodotte nuove particelle fagiche secondo un classico ciclo litico portando a morte le cellule sensibili.

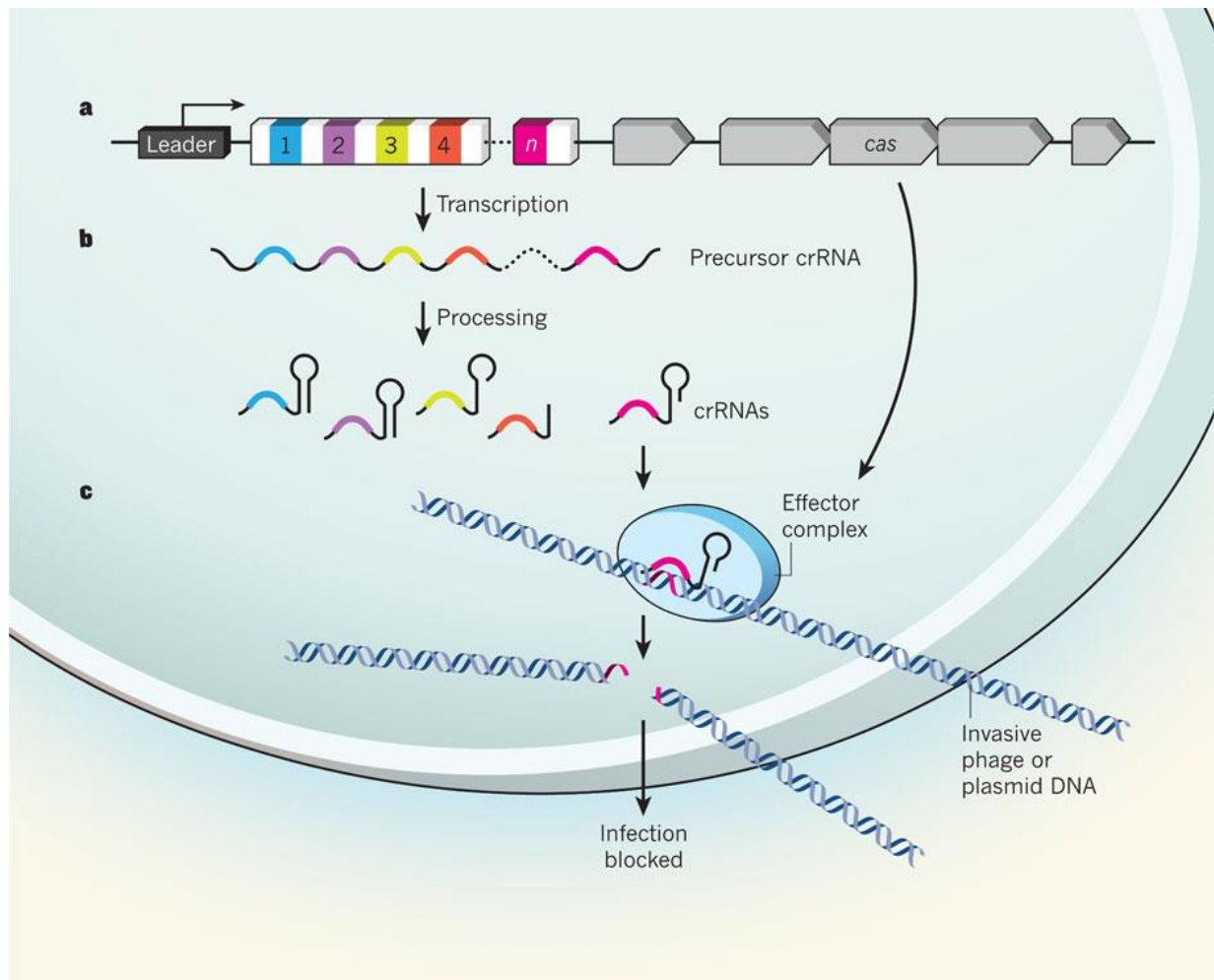
Un piccolo numero di batteri riesce ad acquisire gli "spaziatori fagici" (SPACERS) che permetteranno la sopravvivenza della cellula batterica tramite una degradazione CRISPR-mediata del DNA o mRNA fagico



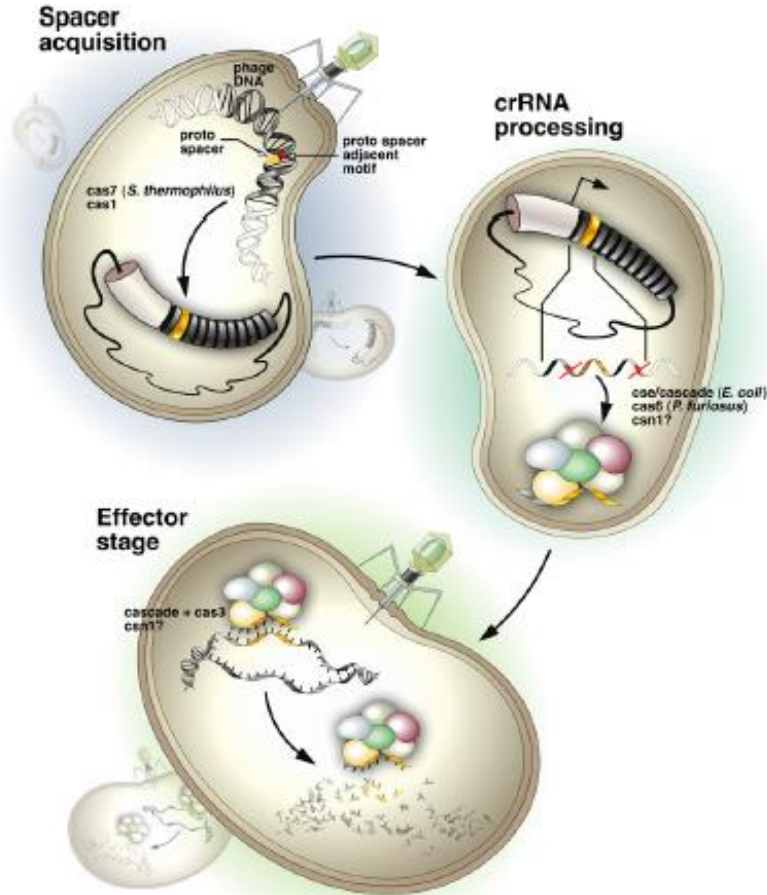
Modello semplificato : la serie di SPACERS-REPEATS è trascritta in un lungo RNA ed i REPEATS assumono una struttura secondaria.

Le proteine Cas riconoscono la struttura secondaria e tagliano l'RNA in modo da produrre dei sRNAs ognuno dei quali contiene una sequenza SPACERS e 2 mezzi REPEATS. Gli sRNA complessandosi alle proteine CAS si appaiano al DNA fagico provocandone la degradazione.

Questo processo è mediato da 1 o più proteine CAS



Le 3 fasi del processo di acquisizione dell'immunità

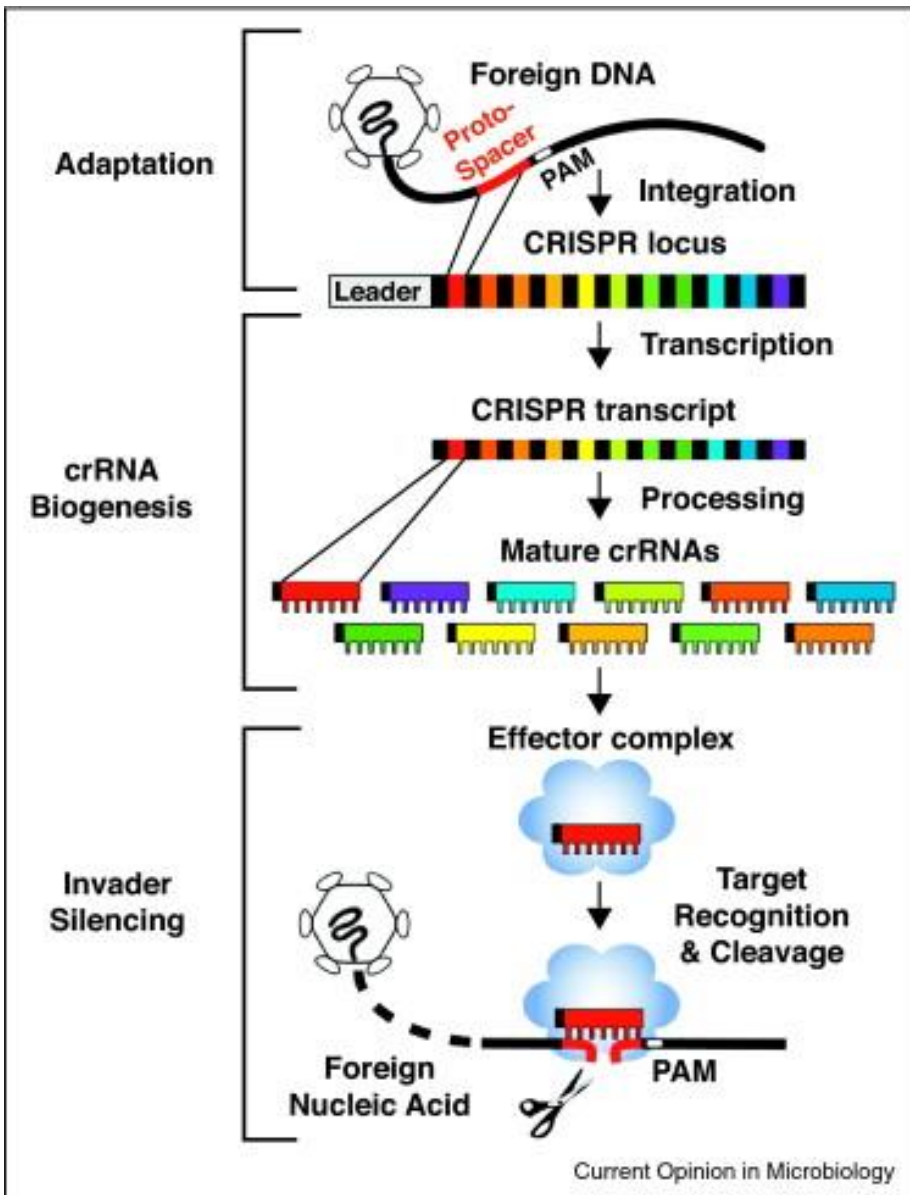


Adattamento: Acquisizione degli SPACERS: durante l'infezione fagica segmenti di acidi nucleici dell'elemento invasore vengono incorporati a valle della sequenza leader del CRISPR.

Espressione: Nella fase di processamento il locus è trascritto e processato in crRNA maturo contenente una sequenza REPEAT con un segnale di 8 nt ed un singola sequenza spacer.

Interferenza: Durante la fase effettrice i crRNA complessati alle proteine CAS portano alla degradazione il DNA complementare

Il sistema CRISPR-Cas per la difesa da DNA invasore



Nella fase di adattamento un piccolo frammento di DNA estraneo definito **protospacer** viene acquisito dall'elemento invasore all'interno del locus CRISPR in posizione adiacente al leader.

In seguito a questo processo di acquisizione di sequenze SPACERS dal DNA di elementi invasori, il locus CRISPR risulta costituito da brevi sequenze di sequenze ripetute separate da sequenze diverse (in colore).

Successivamente i trascritti del locus CRISPR sono processati per rilasciare dei piccoli crRNA ognuno in grado di riconoscere una sequenza specifica. I crRNA maturi contengono tipicamente alcuni nucleotidi delle sequenze ripetute che sono il segnale di riconoscimento dei crRNA

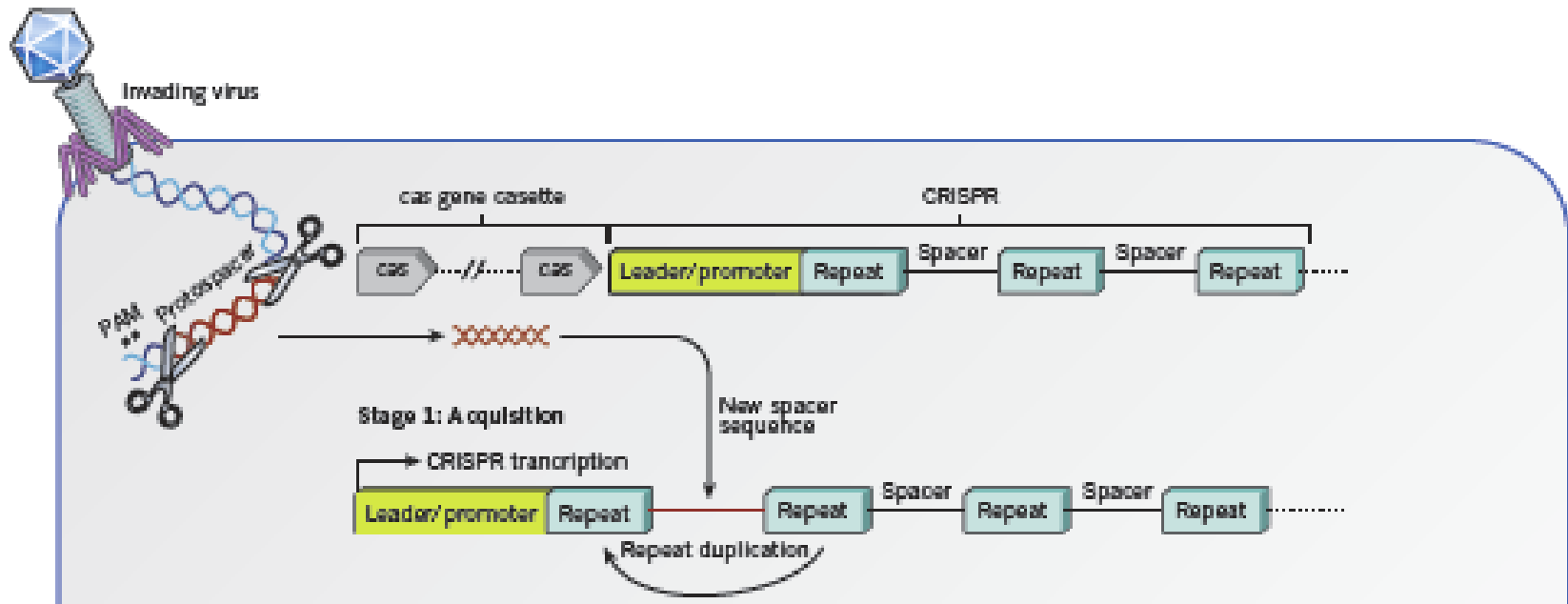
Nella fase di silenziamento il complesso crRNA-CAS riconosce il DNA (o RNA) estraneo tramite appaiamento con i piccoli crRNA.

I sistemi Cmr e Csn colpiscono il taglio dell'RNA target o DNA.

I **PAM Protospacers Adjacent Motif** sono localizzati vicino alle sequenze del DNA "invasore" selezionato per l'integrazione nei CRISPR forniscono ulteriori segnali per il riconoscimento delle sequenze "invasori"

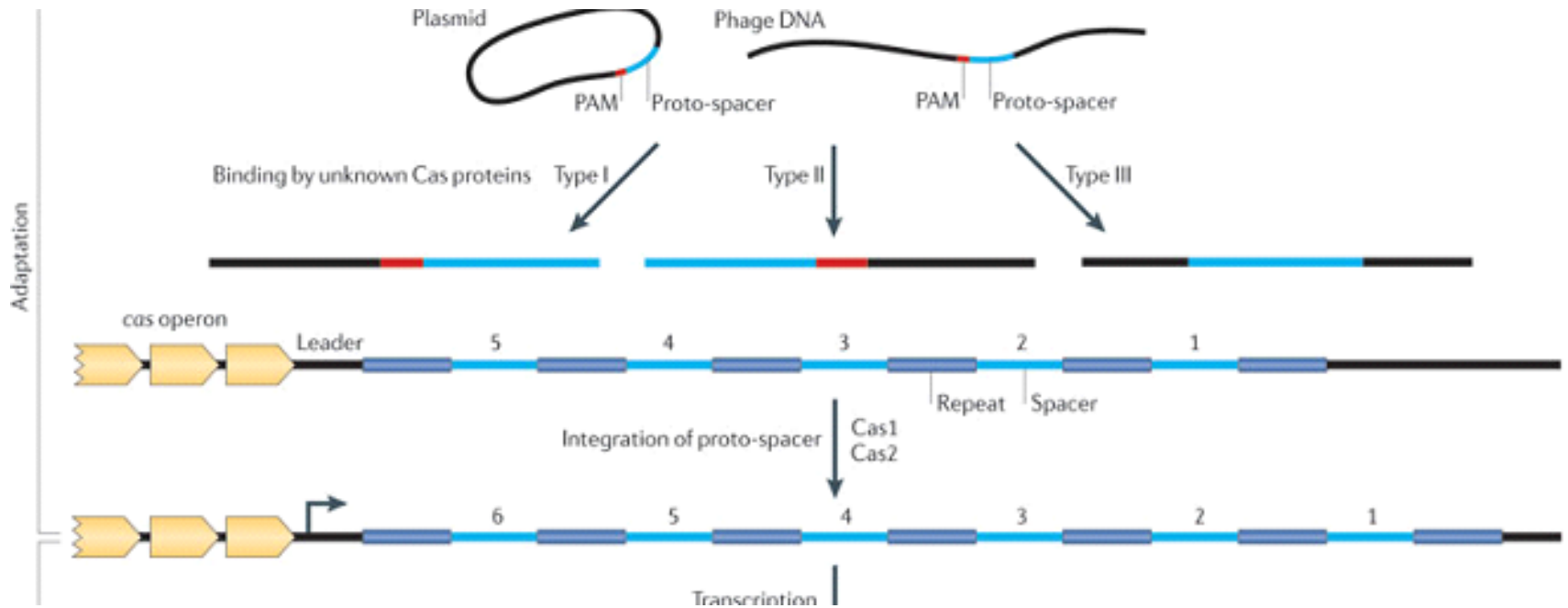
Diversità nei meccanismi CRISPR.

1. I processi iniziali sono in comune



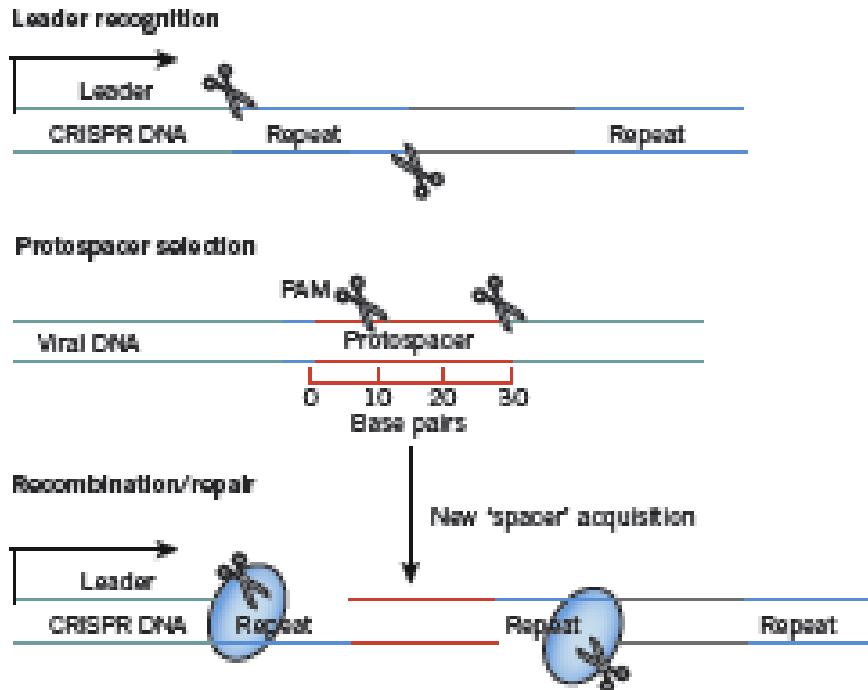
La leader sequence potrebbe contenere un elemento di riconoscimento in grado di reclutare l'apparato di integrazione. Si può anche ipotizzare che l'integrazione avvenga tramite una sola elica di DNA divenuta disponibile in seguito al processo di trascrizione. Il meccanismo di trascrizione/integrazione potrebbe associare la fase di integrazione a quella di trascrizione e garantire che gli spacers dei fagi più recenti siano quelli trascritti per primi

Fase 1: Adattamento



Nei CRISPR di tipo I e di tipo II il riconoscimento dei PROTOPACERS negli acidi nucleici "invasori" dipende proprio dai motivi PAM ma come avvenga questo tipo di riconoscimento non è ancora stato chiarito.

Dopo il riconoscimento iniziale il protospacers viene incorporato grazie alle proteine Cas1 e Cas2 per formare gli spacers.



La fase di adattamento avviene tramite integrazione di frammenti di acido nucleico estraneo vicino alla sequenza leader. I Protospacers non sono scelti a caso ma vengono selezionati tra le regioni che fiancheggiano le sequenze PAM (Protospacers Adjacent Motifs).

In seguito ad un taglio coordinato del DNA esogeno avviene l'inserimento della sequenza di PROTOPACER nella locus CRISPR attraverso un meccanismo che duplica la REPEAT SEQUENCE mantenendo l'architettura del locus (Repeat-Spacer-Repeat).

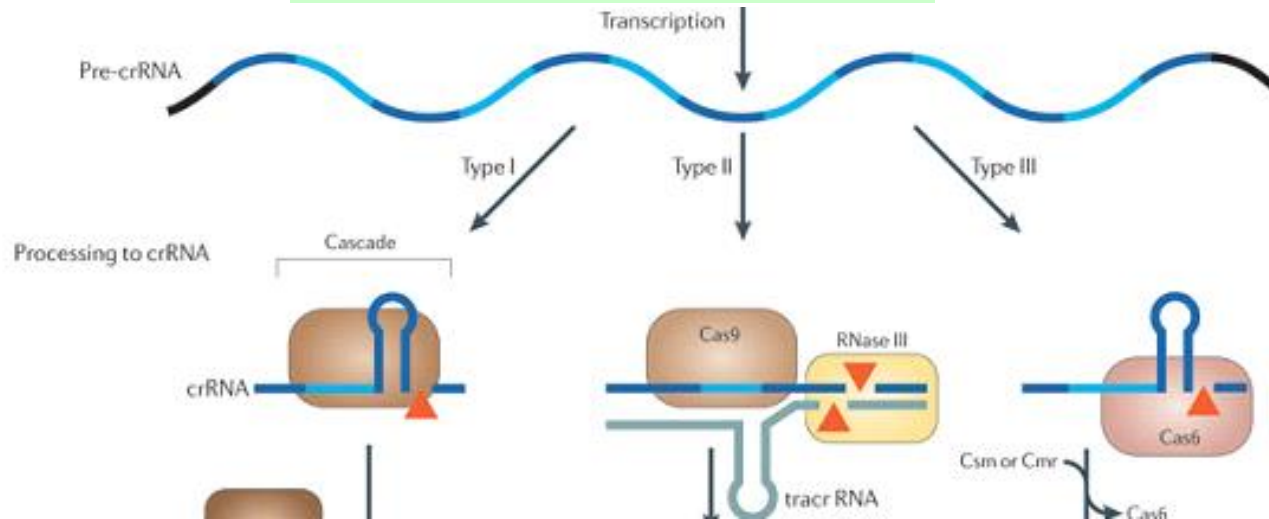
Il meccanismo non è stato ancora bene definito ma coinvolge la proteina CAS1 e altri fattori implicati nei processi di riparo e di ricombinazione generale

Le regioni CRISPR evolvono rapidamente e sono spesso ipervariabili anche tra ceppi correlati.

I nuovi spacers sono inseriti sempre all'estremità 5' vicini alla sequenza leader e costituiscono regioni più variabili rispetto ai "vecchi" spacer localizzati al 3'. Si osservano frequentemente delezioni di unità spacer-repeat necessarie per prevenire un ingrandimento eccessivo del locus : non è chiaro se avvengano per ricombinazione omologa o per un processo attivo.

Generalmente i CRISPR loci sono diversi tra le diverse specie batteriche anche se si possono trovare dei repeats simili tra specie diverse ad indicare eventi possibili di HGT mediato da megaplasmidi che possono contenere loci CRISPR

Fase 2: Espressione



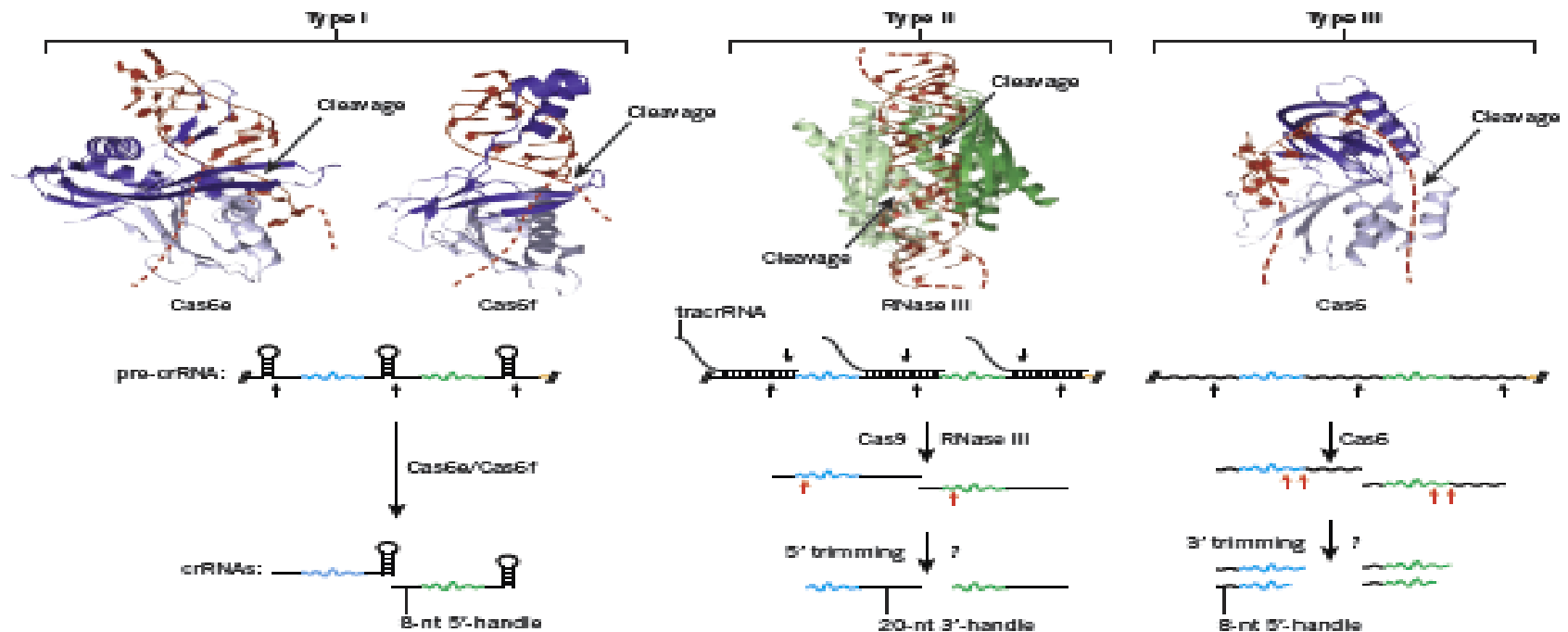
Durante la fase di espressione il locus CRISPR contenente gli spacers viene espresso dando luogo ad un lungo trascritto definito pre-crRNA.

Nel tipo I il complesso di proteine CRISPR associate (CASCADE) si lega al pre- crRNA che viene poi tagliato dalle subunità Cas6e e Cas6f dando luogo a piccoli crRNA costituiti da una tipica sequenza di 8 nucleotidi all'estremità 5' e dalla restante porzione del repeat che forma poi una struttura ad ansa al 3'

Nel tipo II un *trans encoded RNA* si lega al segmento repeat del pre crRNA e questo appaiamento RNA-RNA viene riconosciuto dalla RNase III dell'ospite in presenza delle proteine Cas.

Nel tipo III viene maturato dal Cas6 e poi trasferito a altri complessi proteici Csm e Cmr

Il processamento del crRNA.

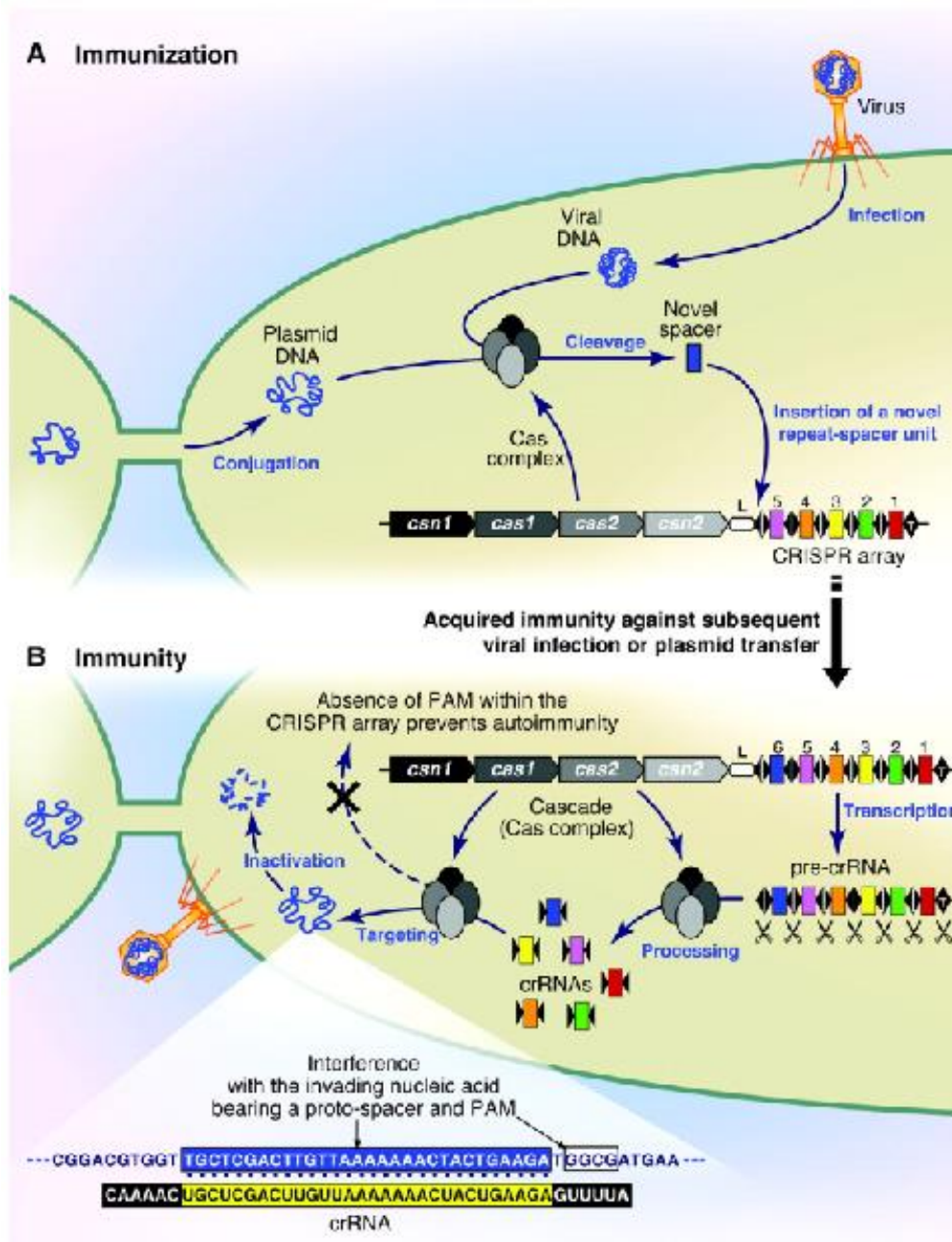


Tipo I Cas6e e C6f riconoscono l'ansa maggiore della struttura stem-loop del crRNA tramite interazioni elettrostatiche .Il taglio avviene nella giunzione tra RNA DS -SS, lasciando 8 basi come maniglia al 5'

TipoII. Interviene un tracrRNA (trans acting antisense RNA) codificato sull'altra elica con una complementarità di sequenza con le sequenze REPEAT del CRISPR RNA per formare un duplex che è riconosciuto e tagliato da RNaseIII in associazione a Cas9. Questo intermedio è successivamente processato al 5' dando luogo a cr RNA di circa 40 nucleotidi con una maniglia di circa 20nt al 3'

TipoIIIb .Troviamo che Cas6 è in grado di tagliare precrRNA 8 nucleotidi a monte della giunzione spacer/repeat. Un ulteriore maturazione avviene con un processamento al 3'.

Ruolo delle CAS protein



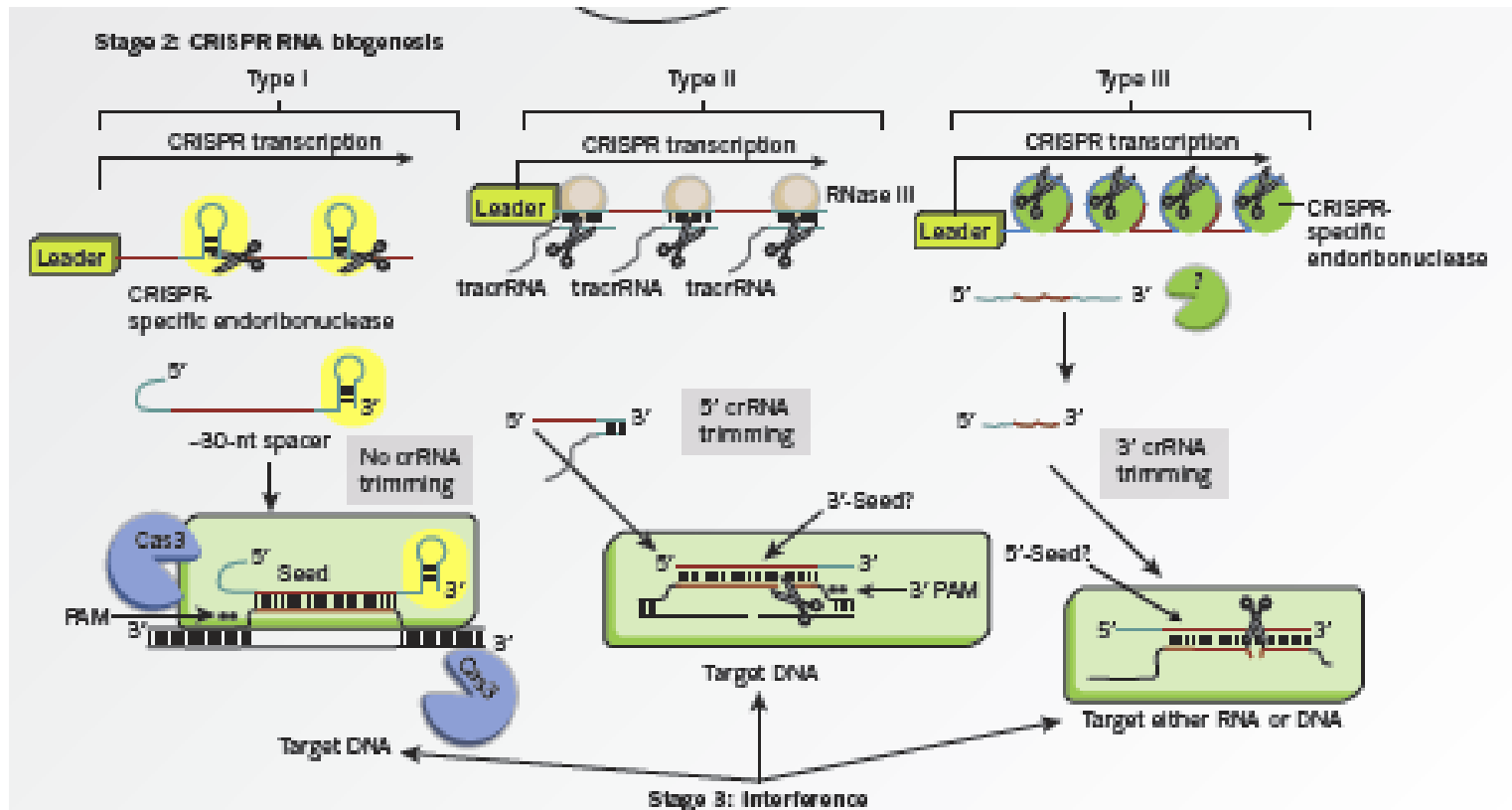
Le protein eCAS sono coinvolte nell'acquisizione di nuovi spacers ma contribuiscono a conferire resistenza al l'elemento genetico invasivo che può essere un fago o un plasmide

Nella fase di immunità ritroviamo le proteine CAS che sono essenziali per il processamento del trascritto pre crRNA.

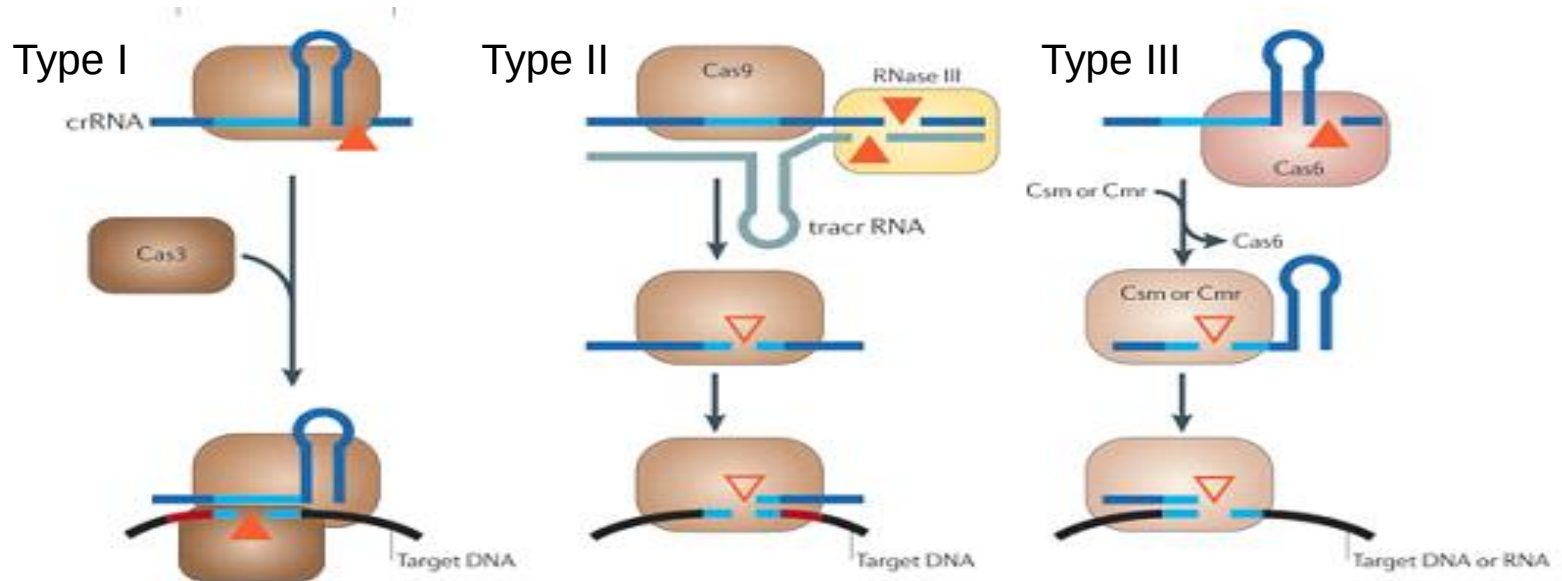
Infine nella fase di interferenza il Complesso CAS può indirizzare i crRNA verso il target.

Nell'appaiamento svolge un ruolo importante la sequenza PAM

2. La diversità nella fase di adattamento :3 strategie per la generazione dei crRNA maturi



Fase 3: Interferenza

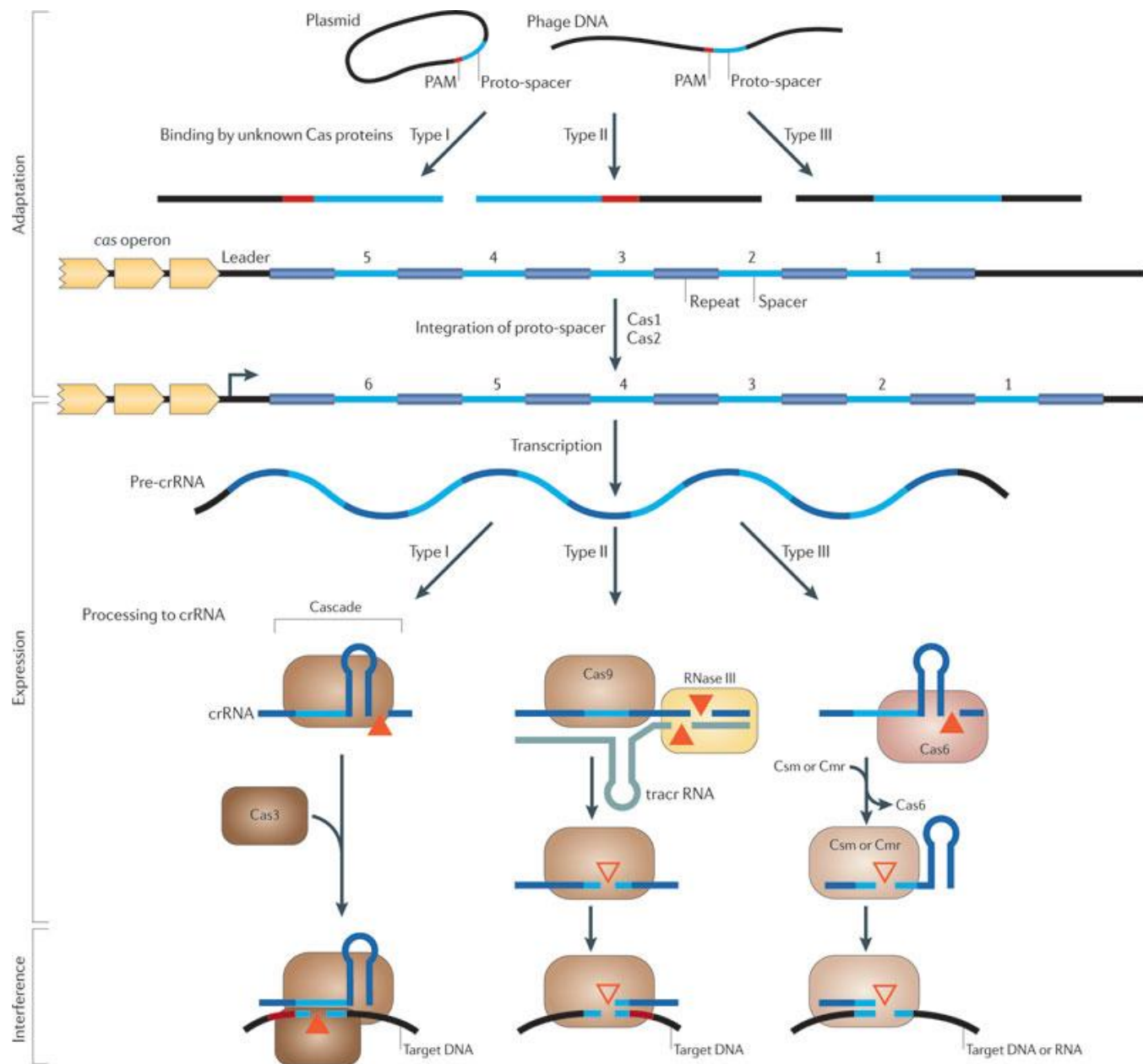


Durante la fase di interferenza l'acido nucleico "invasore" viene tagliato ma con strategie diverse.

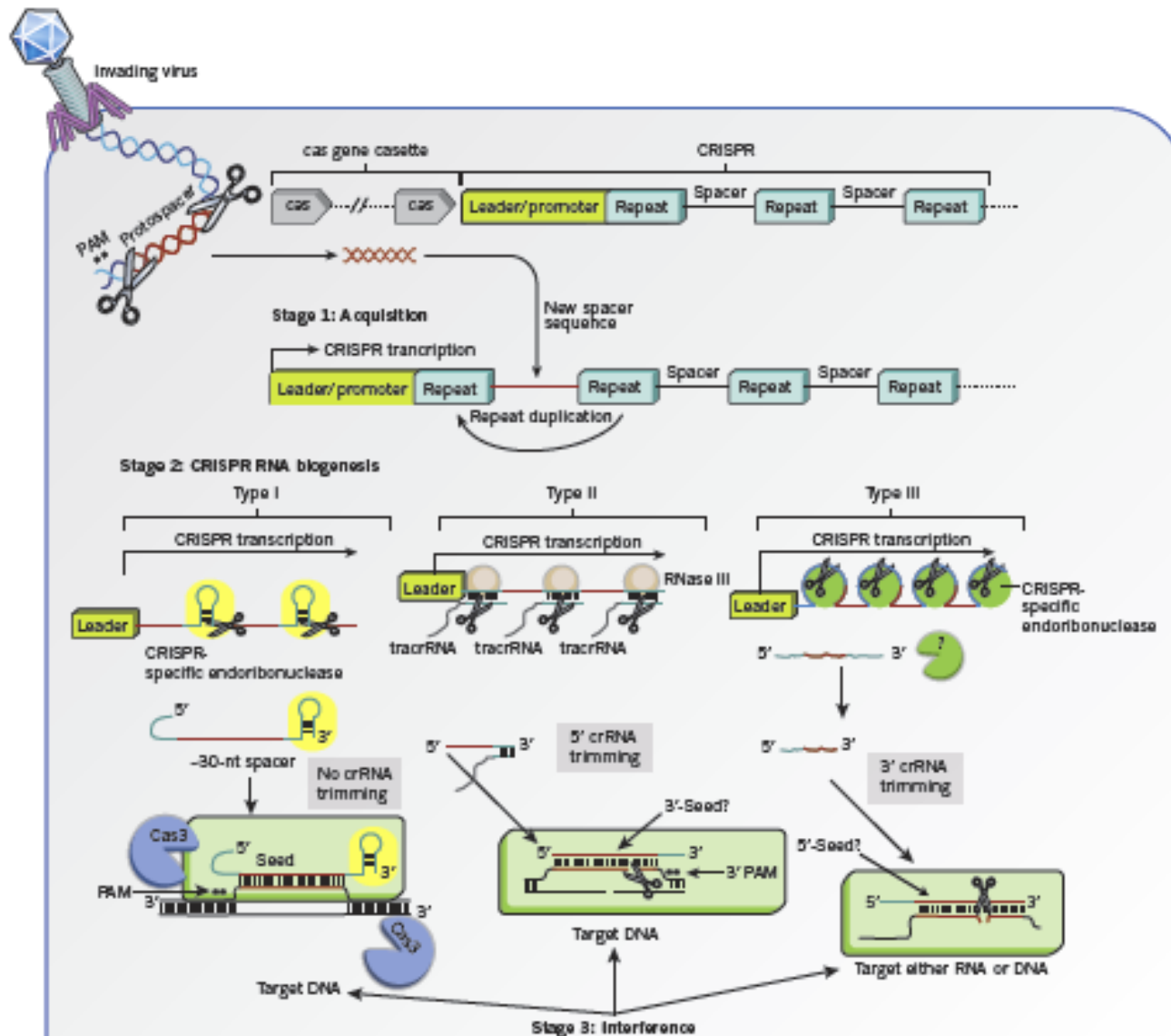
Nel Tipo I il crRNA guida il complesso Cascade verso le regioni bersaglio che contengono il DNA complementare che verrà poi tagliato probabilmente dal Cas3. Un ruolo importante per il riconoscimento viene svolto dalle PAM.

Nel Tipo II Cas9 contenente crRNA punta direttamente il DNA bersaglio con l'aiuto delle PAM.

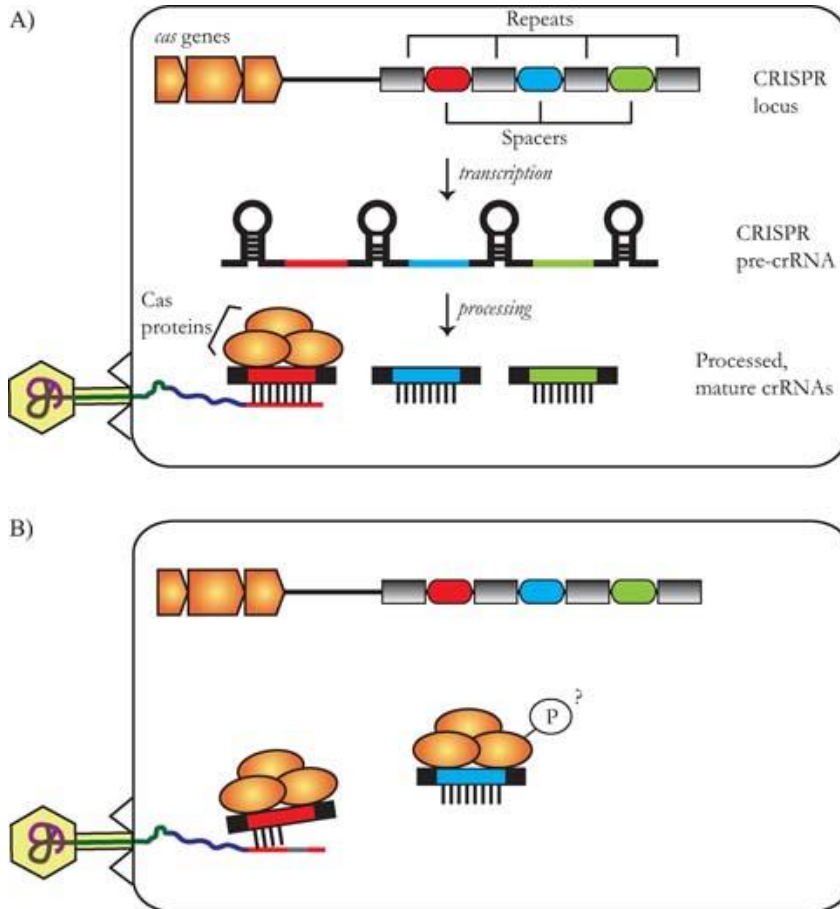
Nel Tipo III si può avere il riconoscimento sia del DNA che dell'RNA da parte di 2 diversi complessi (Csn o Cmr) ed il successivo taglio da parte di enzimi non ancora identificati



Visione d'insieme del processo di immunità CRISPR mediato.



The phage-host arms race: Shaping the evolution of microbes



Mechanism of action: transcription from the repeat-spacer CRISPR locus generates a long non-coding RNA, with repeats that may sometimes assume a secondary structure. Cleavage of the repeat sequences by the Cas proteins generates crRNAs that target the phage DNA or RNA, and interfere with phage infection.

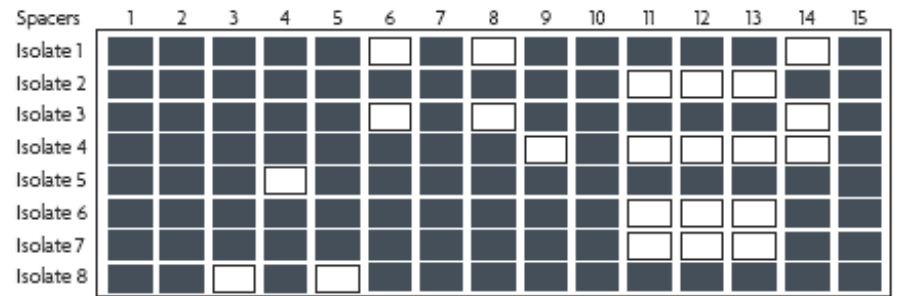
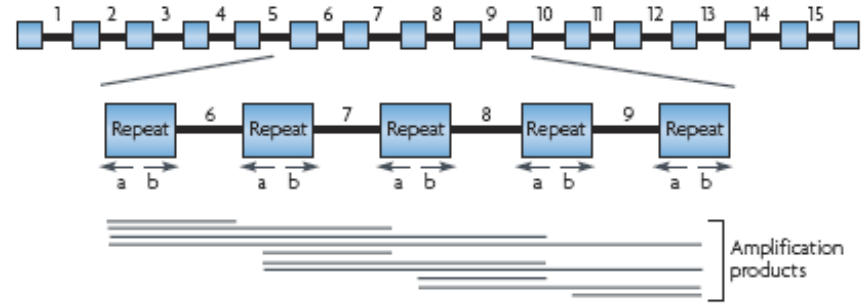
I fagi possono evadere il interferenza CRISPR con mutazioni o ricombinazioni nella sequenza protospacer. Oppure potrebbero fosforilare le CAS

Applicazioni dei CRISPR

1. SPOLIGOTYPING.

Prima ancora di capire quale fosse il ruolo funzionale dei CRISPR ,questi loci vennero utilizzati per la classificazione dei ceppi di *M.tuberculosis* .

Questa tecnica si basa sull'utilizzazione di oligo disegnati sulla base dei repeat in modo da amplificare le regioni CRISPR. I probes che riconoscono spacer noti vengono poi fissati su una membrana ed ibridati con i diversi prodotti di amplificazione. La presenza / assenza di ibridazione permette di raggruppare i diversi ceppi. Utilizzata per tipizzare *Corynebacterium diphtheriae* e recentemente *Lactobacillus*

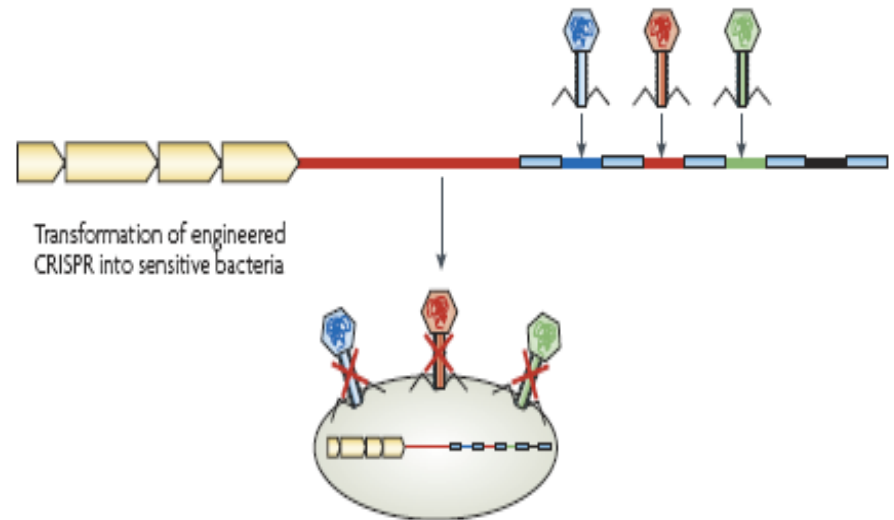


2. Costruzione di ceppi fago-resistenti

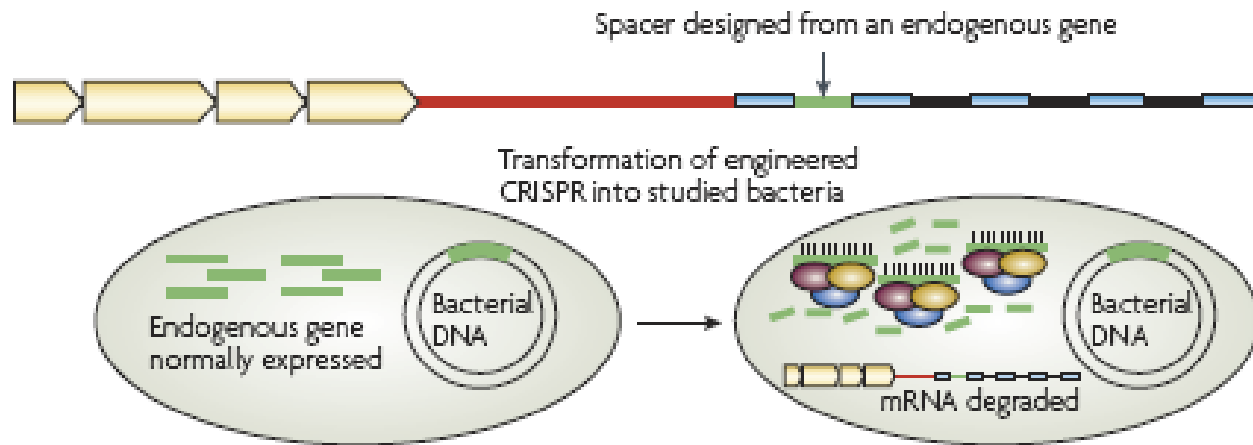
Per molte industrie che utilizzano batteri quali l'industria alimentare e del vino i fagi costituiscono un serio problema che provoca danni ingenti.

Ingegnerizzando i loci CRISPR tramite l'aggiunta in vitro di spacers con sequenze fagiche e poi trasformando questi loci si può indurre resistenza a diversi fagi in batteri sensibili.

Con questa tecnica si può indurre l'immunità nei batteri da utilizzare come starter

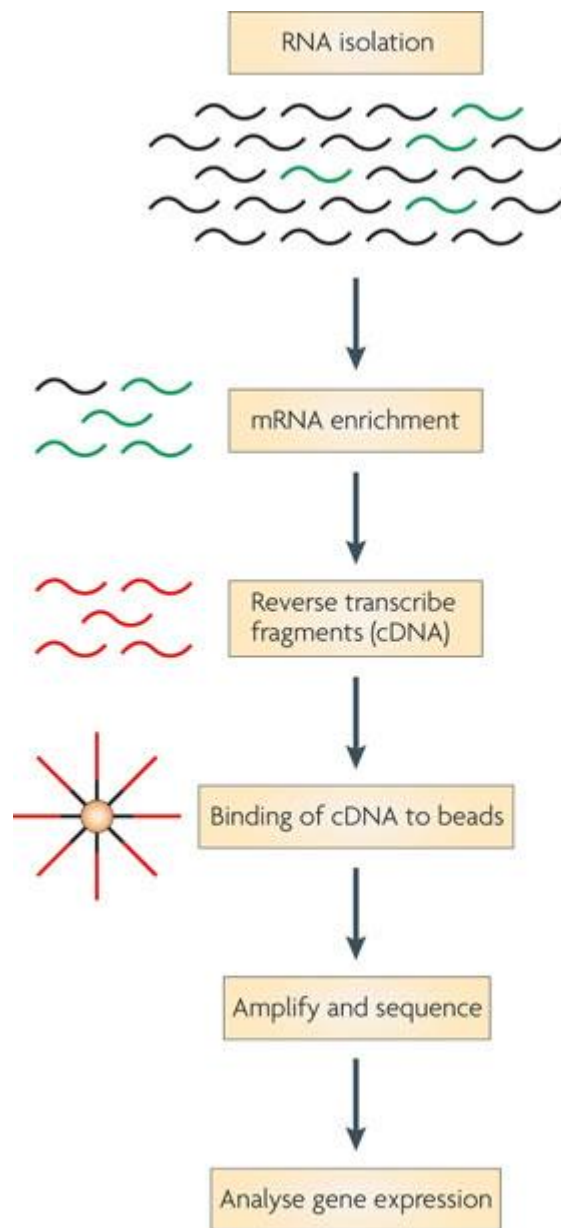


3. Silenziamento selettivo di geni



Tenendo conto delle analogie tra il sistema CRISPR ed il sistema di RNA interference degli eucarioti si può pensare di introdurre una determinata sequenza all'interno di un locus CRISPR e poi di silenziare il relativo mRNA tramite il sistema CRISPR/CAS.

Si potrebbero costruire dei plasmidi ricombinanti contenenti vari loci CRISPR con spacer idonei per il silenziamento di più geni diversi



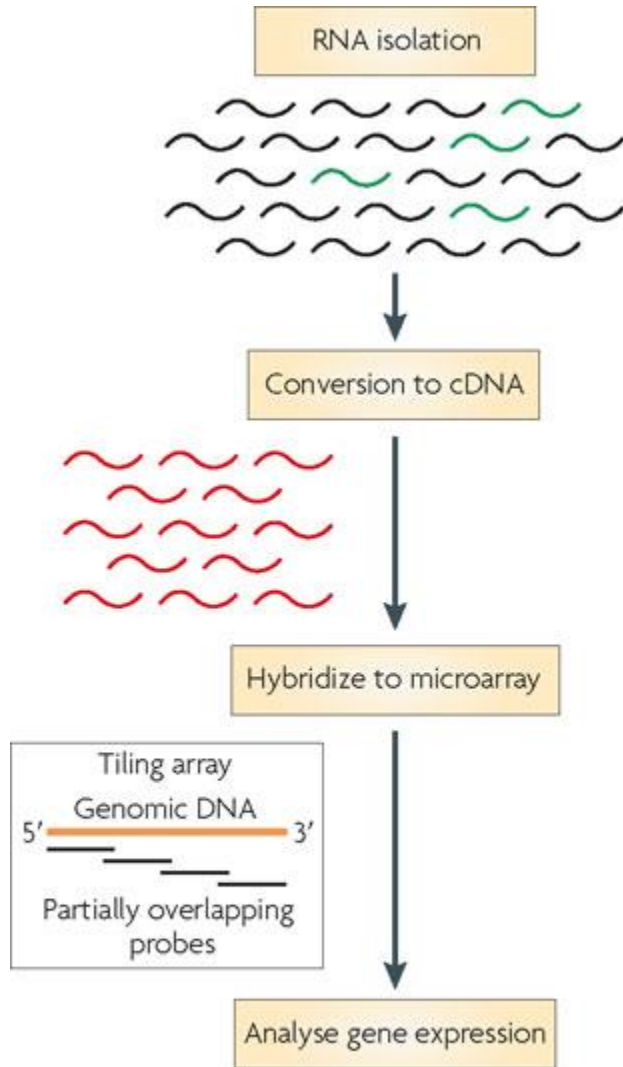
RNA isolato viene arricchito per mRNA tramite la rimozione degli rRNA ribosomiali e dei tRNA mediante varie tecniche. La cattura degli rRNA coinvolge l'uso di probe che legano RNA 16 e 23 S. RNA processato viene degradato da esonucleasi 5-3' che riconoscono le estremità 5' monofosforilate in questo modo vengono degradati preferenzialmente questi RNA mentre i mRNA portano estremità 5' trifosforilate.

L'arricchimento può avvenire anche tramite l'aggiunta di code di poliA artificiali seguite da un ripescaggio tramite oligo T.

Gli mRNA sono poi convertiti in cDNA tramite la reverse trascrittasi. Dal momento che non ci sono code di poli A vi sono metodi alternativi per il priming quali l'aggiunta di code di poli A seguita dal priming con oligo ricchi in T,

Il cDNA viene poi amplificato generalmente legato a piccole biglie e sequenziato tramite 454, Illumina e Solid .T

CRISPR–Cas (clustered regularly interspaced short palindromic repeats–CRISPR-associated proteins) systems act in three stages: adaptation, expression and interference. In type I and type II CRISPR–Cas systems, but not in type III systems, the selection of proto-spacers in invading nucleic acid probably depends on a proto-spacer-adjacent motif (PAM)^{22, 30, 31}, but how the PAM or the nucleic acid is recognized is still unclear. After the initial recognition step, Cas1 and Cas2 most probably incorporate the proto-spacers into the CRISPR locus to form spacers. During the expression stage, the CRISPR locus containing the spacers is expressed, producing a long primary CRISPR transcript (the pre-crRNA). The CRISPR-associated complex for antiviral defence (Cascade) complex binds the pre-crRNA, which is then cleaved by the Cas6e or Cas6f subunits (in subtype I-E or I-F, respectively), resulting in crRNAs with a typical 8-nucleotide repeat fragment on the 5' end and the remainder of the repeat fragment, which generally forms a hairpin structure, on the 3' flank. Type II systems use a trans-encoded small RNA (tracrRNA) that pairs with the repeat fragment of the pre-crRNA, followed by cleavage within the repeats by the housekeeping RNase III in the presence of Cas9 (formerly known as Csn1 or Csx12). Subsequent maturation might occur by cleavage at a fixed distance within the spacers²⁵, probably catalysed by Cas9. In type III systems, Cas6 is responsible for the processing step, but the crRNAs seem to be transferred to a distinct Cas complex (called Csm in subtype III-A systems and Cmr in subtype III-B systems). In subtype III-B systems, the 3' end of the crRNA is trimmed further²⁸. During the interference step, the invading nucleic acid is cleaved. In type I systems, the crRNA guides the Cascade complex to targets that contain the complementary DNA, and the Cas3 subunit is probably responsible for cleaving the invading DNA²¹. The PAM probably also plays an important part in target recognition in type I systems. In type II and type III systems, no Cas3 orthologue is involved ([Table 2](#)). In type II systems, Cas9 loaded with crRNA probably directly targets invading DNA, in a process that requires the PAM²⁶. The two subtypes of CRISPR–Cas type III systems target either DNA (subtype III-A systems³¹) or RNA (subtype III-B systems²⁸). In type III systems, a chromosomal CRISPR locus and an invading DNA fragment are distinguished by either base pairing to the 5' repeat fragment of the mature crRNA (resulting in no interference) or no base pairing (resulting in interference)³⁰. Filled triangles represent experimentally characterized nucleases, and unfilled triangles represent nucleases that have not yet been identified.



Tiling array (allineamento a pavimento di mattonelle)
Ogni array contiene circa 105 oligonucleotidi RNA totale viene isolato e convertito in cDNA marcato prima dell'ibridazione