

# CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA

Luciano A. Marraffini and Erik J. Sontheimer\*

Horizontal gene transfer (HGT) in bacteria and archaea occurs through phage transduction, transformation, or conjugation, and the latter is particularly important for the spread of antibiotic resistance. Clustered, regularly interspaced, short palindromic repeat (CRISPR) loci confer sequence-directed immunity against phages. A clinical isolate of *Staphylococcus epidermidis* harbors a CRISPR spacer that matches the *nickase* gene present in nearly all staphylococcal conjugative plasmids. Here we show that CRISPR interference prevents conjugation and plasmid transformation in *S. epidermidis*. Insertion of a self-splicing intron into *nickase* blocks interference despite the reconstitution of the target sequence in the spliced mRNA, which indicates that the interference machinery targets DNA directly. We conclude that CRISPR loci counteract multiple routes of HGT and can limit the spread of antibiotic resistance in pathogenic bacteria.

Clustered, regularly interspaced, short palindromic repeat (CRISPR) loci are present in ~40% of eubacterial genomes and nearly all archaeal genomes sequenced to date and consist of short (~24 to 48 nucleotides) repeats separated by similarly sized unique spacers (1, 2). They are generally flanked by a set of CRISPR-associated (*cas*) protein-coding genes (3–5). The CRISPR spacers and repeats are transcribed and processed into small CRISPR RNAs (crRNAs) (4, 6–8) that specify acquired immunity against bacteriophage infection by a mechanism that relies on the strict identity between CRISPR spacers and phage targets (3, 4).

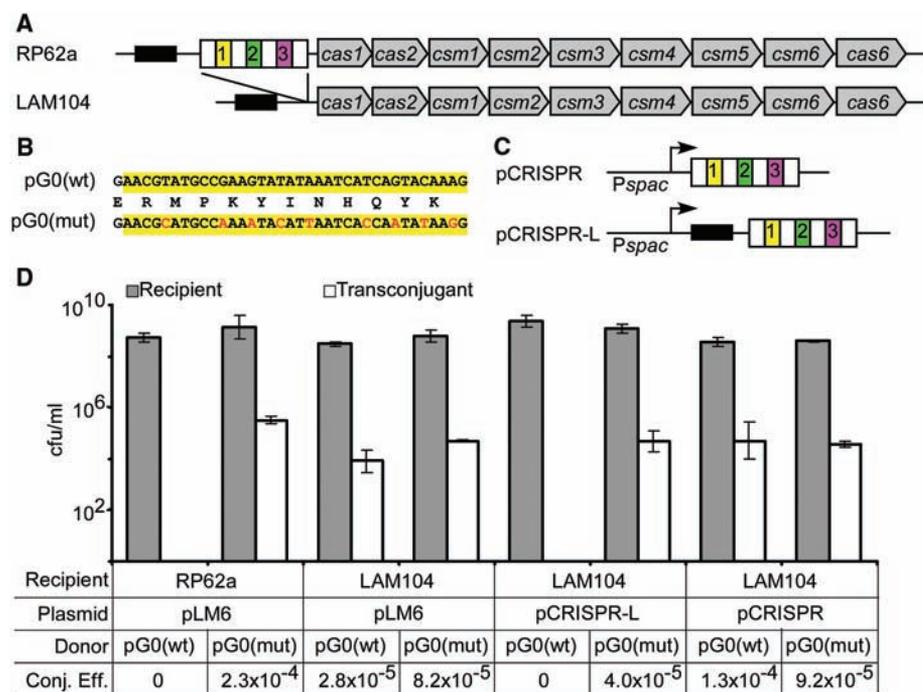
The rise of hospital- and community-acquired methicillin- and vancomycin-resistant *Staphylococcus aureus* (MRSA and VRSA, respectively) is directly linked to the horizontal transfer of antibiotic resistance genes by plasmid conjugation (9, 10). *S. aureus* and *S. epidermidis* strains are the most common causes of nosocomial infections (11–13), and conjugative plasmids can spread from one species to the other. Although the *S. epidermidis* strain American Type Culture Collection (ATCC) 12228 (14) lacks CRISPR sequences, the clinically isolated strain RP62a (15) contains a CRISPR locus (Fig. 1A and fig. S1A) that includes a spacer (*spc1*) that is homologous to a region of the *nickase* (*nes*) gene found in all sequenced staphylococcal conjugative plasmids (fig. S1B), including those from MRSA and VRSA strains (9, 16, 17).

To test whether *spc1* prevents plasmid conjugation into *S. epidermidis* RP62a, we disrupted the sequence match by introducing nine silent mutations into the *nes* target in the conjugative plasmid pG0400 (18), generating pG0(mut) (Fig.

1B) (19). We tested whether both wild-type and mutant pG0400 transferred from *S. aureus* strain RN4220 (20) into either of the two *S. epidermidis* strains (Fig. 1D and fig. S1C). Although the conjugation frequency of both plasmids was similar for the CRISPR-negative ATCC 12228

strain, only pG0(mut) was transferred into the CRISPR-positive RP62a strain and with a frequency similar to that of wild-type pG0400 in the control ATCC 12228 strain. These results indicate that CRISPR interference can prevent plasmid conjugation in a manner that is specified by sequence identity between a spacer and a plasmid target sequence.

To test this conclusion more rigorously, and to determine whether the CRISPR sequences themselves are responsible for the observed interference, we deleted the four repeats and three spacers present in the RP62a locus in order to generate the isogenic  $\Delta$ *crispr* strain LAM104 (Fig. 1A) and tested its ability to act as a recipient for the conjugative transfer of pG0400. Again, wild-type RP62a was refractory to pG0400 transfer, whereas the conjugation efficiency for the LAM104 strain was similar to that obtained for *S. epidermidis* ATCC 12228 (Fig. 1D and fig. S1C). pG0(mut) transfer was similar in both strains. To restore interference in the  $\Delta$ *crispr* mutant, we transformed LAM104 with plasmid pCRISPR or pCRISPR-L (Fig. 1C). Both of these plasmids contain the RP62a CRISPR repeats and spacers downstream of an isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG)-inducible pro-



**Fig. 1.** A CRISPR locus provides immunity against plasmid conjugation in *S. epidermidis*. (A) Organization of the RP62a CRISPR locus. Repeats and spacers (colored boxes) are followed by CRISPR-associated genes (*cas1*, *cas2*, and *cas6*) and *cas* subtype *Mycobacterium tuberculosis* genes (*csml* to *csm6*) (5). An AT-rich “leader” sequence precedes the repeat-spacer region (black box). LAM104 is an isogenic  $\Delta$ *crispr* strain lacking only the repeat and spacer sequences. (B) The staphylococcal conjugative plasmid pG0400 *spc1* target sequence [pG0(wt), highlighted in yellow] is shown on the top. This sequence was altered to introduce synonymous mutations to generate pG0(mut), with changes shown in red. (C) To restore interference in strain LAM104, two plasmids were introduced: pCRISPR and pCRISPR-L. (D) Conjugation was carried out by filter-mating in triplicate; the colony-forming units (CFU) per milliliter values (mean  $\pm$  SD) obtained for recipients and transconjugants are shown. Recipient strains, complementing plasmids, and donor conjugative plasmids are indicated. Conjugation efficiency (Conj. Eff.) was calculated as the recipients/transconjugants ratio.

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, 2205 Tech Drive, Evanston, IL 60208, USA.

\*To whom correspondence should be addressed. E-mail: erik@northwestern.edu

moter, but they differ in the amount of upstream flanking sequence included. The appearance of the *spc1* crRNA in pCRISPR-L-containing LAM104 cells (fig. S2) was IPTG-independent, which indicates that the insert includes a natural CRISPR promoter. pCRISPR-L restored interference in the  $\Delta$ *crispr* strain, but pCRISPR

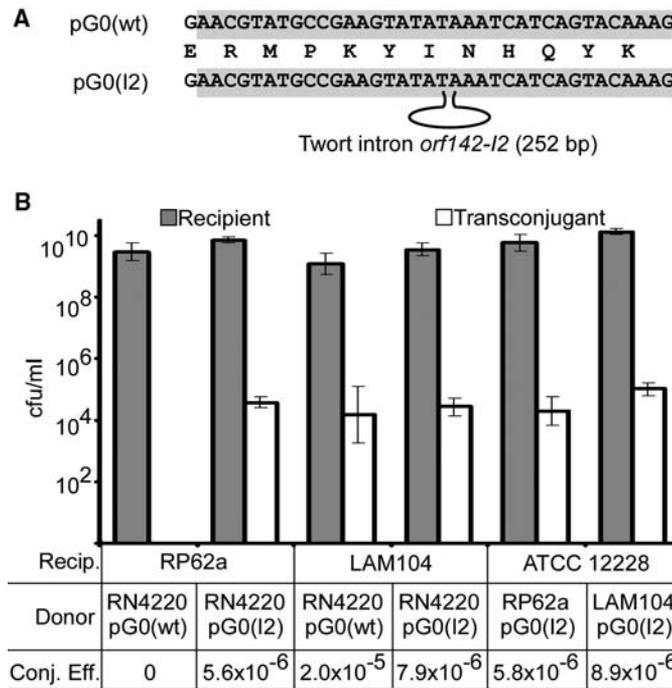
did not, even in the presence of IPTG (fig. S1D). This suggests a role for the leader sequence (Fig. 1A) or other upstream sequences in cis during CRISPR interference, perhaps for processing of the crRNA precursor by Cascade proteins (4). Introduction of pCRISPR-L into strain ATCC 12228 did not alter the conjugation ef-

iciency of pG0400 (fig. S1D) as expected because of the lack of *cas* genes in this strain.

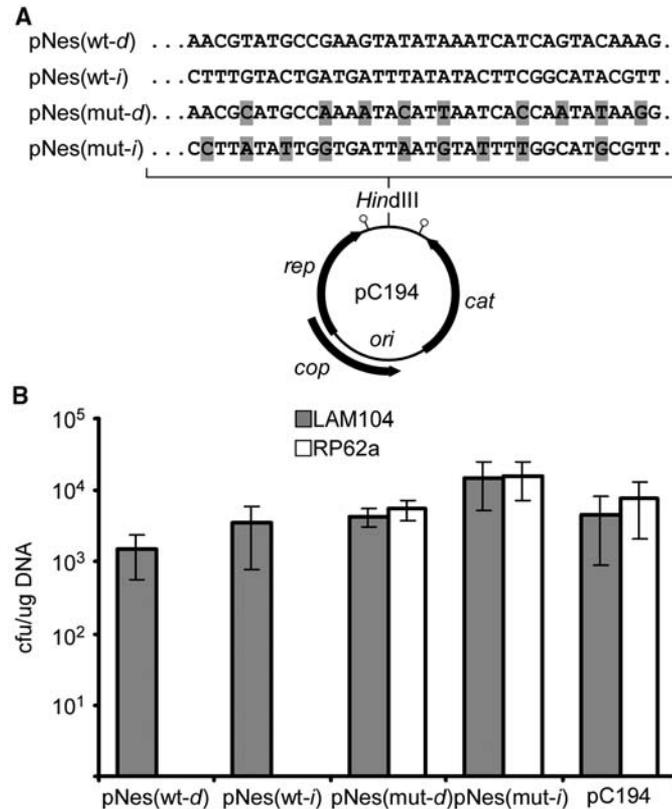
The nature of the crRNA targeting event (RNA-RNA or RNA-DNA) is not known. The requirement for nickase activity only in the donor cell (16) implies that interference with *nes* mRNA and protein expression should block the ability of *S. epidermidis* RP62a to function as a donor. Instead, our data (Fig. 1 and fig. S1) indicate that the CRISPR locus limits the ability of *S. epidermidis* RP62a to act as a plasmid recipient and therefore suggest that *spc1*-directed interference does not target the *nes* transcript. Consistent with this, the orientation of *spc1* leads to the expression of a crRNA that is identical with, rather than complementary to, the *nes* mRNA target sequence, and we have found no evidence for the expression of RNA from the opposite strand (fig. S2) (4). Alternatively, the *spc1* crRNA may target the incoming DNA. To test this, we interrupted the target sequence of the pG0400 *nes* gene with the *orf142-I2* self-splicing group I intron from the staphylococcal Twort phage (21, 22). The mutant conjugative plasmid pG0(I2) lacks an intact *spc1* target DNA sequence (Fig. 2A), but the *spc1* target sequence is regenerated in the *nes* mRNA after transcription and rapid splicing. When tested in conjugation assays, pG0(I2) was transferred to wild-type and  $\Delta$ *crispr* strains with equal efficiencies (Fig. 2B). This observation reflects an evasion of CRISPR activity when the intron is present in the plasmid and therefore indicates that an intact target site is required in the *nes* DNA, but not mRNA, for interference to occur. To confirm that splicing is required for *nes* function, we tested the conjugation of pG0(dI2), a derivative of pG0(I2) containing a three-nucleotide deletion within the intron that inactivates self-splicing (21) (fig. S3A). The pG0(dI2) plasmid was unable to transfer to the RP62a strain (fig. S3B), which indicates a lack of nickase activity in the presence of the unspliced intron.

The requirement for *nes* transcription, splicing, and translation in the donor cell during conjugation (16), and our ability to obtain RP62a transconjugants with the intron-containing plasmid, allowed us to test the capacity of the CRISPR system to target intact spliced *nes* mRNA by using RP62a as a pG0(I2) donor. pG0(I2) conjugative transfer was just as efficient from RP62a as from the isogenic  $\Delta$ *crispr* strain LAM104 (Fig. 2), which indicates that spliced functional *nes* mRNA (which must be present for conjugation to occur) is not targeted during CRISPR interference. Reverse transcription polymerase chain reaction assays confirmed the splicing of the *nes* pre-mRNA in RP62a cells carrying pG0(I2) (fig. S3C). Although our observations do not formally exclude an RNA targeting event that is somehow restricted to nascent, transient, unspliced transcripts, they provide strong evidence that DNA rather than mRNA is the likely crRNA target during CRISPR interference.

**Fig. 2.** CRISPR interference requires an intact target sequence in plasmid DNA but not mRNA. (A) Disruption of the pG0400 *nes* target sequence with the *orf142-I2* self-splicing intron, generating the conjugative plasmid pG0(I2). (B) Conjugation efficiency was measured as in Fig. 1D with RP62a and the  $\Delta$ *crispr* mutant LAM104 as recipients for pG0(wt) and pG0(I2). To test for interference with spliced *nes* mRNA, RP62a and LAM104 transconjugants were also used as donors of pG0(I2) to *S. epidermidis* ATCC 12228.



**Fig. 3.** Plasmid transformation is subject to CRISPR interference. (A) Introduction of the wild-type and mutant *nes* target sequences (mutations highlighted in gray) into the plasmid pC194 (*d*, direct insertion; *i*, inverted insertion). The origin of replication (*ori*) and protein-coding genes are indicated. Stem loops denote the *rep* and *cat* transcriptional terminators. (B) RP62a and the  $\Delta$ *crispr* mutant LAM104 were transformed in triplicate, with the plasmids described in (A). Transformation efficiency was calculated as CFU/ $\mu$ g of DNA (mean  $\pm$  SD).



CRISPR activity against phage and conjugative plasmid DNA molecules suggests that CRISPR systems may also prevent plasmid DNA transformation. We therefore introduced pG0(wt) and pG0(mut) *nes*-target and -flanking sequences (200 base pairs) in either orientation into the staphylococcal plasmid pC194 (23), generating pNes(wt) and pNes(mut), respectively (Fig. 3A). Flanking DNA was included in the inserts to ensure the presence of any sequences outside of the target that may contribute to CRISPR interference (24). Plasmids were transformed by electroporation into wild-type RP62a and isogenic  $\Delta$ crispr LAM104 strains. pC194 and both pNes(mut) plasmids were transformed into both strains, whereas the pNes(wt) plasmids were transformed only into the  $\Delta$ crispr mutant (Fig. 3B). We also performed pNes(wt)/pNes(mut) mixed transformations of RP62a or LAM104 strains to test interference in an internally controlled fashion. Again, only pNes(mut) plasmids were recovered from RP62a transformants, whereas pNes(wt) and pNes(mut) plasmids were found in LAM104 transformant colonies (fig. S4). It remains to be established whether natural transformation, which involves the uptake of a single DNA strand (25), is subject to CRISPR interference. Nonetheless, our experiments suggest that CRISPR systems can counteract multiple routes of plasmid transfer.

These transformation data provide additional evidence that crRNAs target DNA molecules. First, interference occurred regardless of the insert orientation in pNes(wt); this, combined with the lack of compelling evidence for CRISPR-derived double-stranded RNA (fig. S2) (4, 6, 7), is consistent with *spc1* targeting either DNA strand rather than a unidirectional transcript. Second, the target sites in the pNes(wt) and pNes(mut) plasmids are located between the transcriptional terminators of the *rep* and *cat* genes (Fig. 3A) (23, 26, 27). This minimizes the likelihood that this region of the plasmid is even transcribed, which is consistent with its dispensability for plasmid maintenance (23, 28).

Altogether, these data provide strong functional evidence that CRISPR interference acts at the DNA level and therefore differs fundamentally from the RNA interference (RNAi) phenomenon observed in eukaryotes and with which CRISPR activity was originally compared (29). A DNA targeting mechanism for CRISPR interference implies a means to prevent its action at the encoding CRISPR locus itself, as well as other potential chromosomal loci, such as prophage sequences. Little information exists to suggest how crRNAs would avoid targeting “self” DNA, although the role of flanking sequences during CRISPR interference (24) could contribute to target specificity. From a practical standpoint, the ability to direct the specific addressable destruction of DNA that contains any given 24- to 48-nucleotide target sequence could have considerable functional utility, especially if the system can function outside of its native bacterial or archaeal context. Furthermore, our results demon-

strate that CRISPR function is not limited to phage defense, but instead encompasses a more general role in the prevention of HGT and the maintenance of genetic identity, as with restriction-modification systems. A primary difference between restriction-modification and CRISPR interference is that the latter can be programmed by a suitable effector crRNA. If CRISPR interference could be manipulated in a clinical setting, it would provide a means to impede the ever-worsening spread of antibiotic resistance genes and virulence factors in staphylococci and other bacterial pathogens.

#### References and Notes

1. I. Grissa, G. Vergnaud, C. Pourcel, *BMC Bioinformatics* **8**, 172 (2007).
2. R. Sorek, V. Kunin, P. Hugenholtz, *Nat. Rev. Microbiol.* **6**, 181 (2008).
3. R. Barrangou *et al.*, *Science* **315**, 1709 (2007).
4. S. J. Brouns *et al.*, *Science* **321**, 960 (2008).
5. D. H. Haft, J. Selengut, E. F. Mongodin, K. E. Nelson, *PLoS Comput. Biol.* **1**, e60 (2005).
6. T. H. Tang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7536 (2002).
7. T. H. Tang *et al.*, *Mol. Microbiol.* **55**, 469 (2005).
8. C. Hale, K. Kleppe, R. M. Terns, M. P. Terns, *RNA*, **10**, 1261/rna.1246808 (2008).
9. L. M. Weigel *et al.*, *Science* **302**, 1569 (2003).
10. E. Y. Furuya, F. D. Lowy, *Nat. Rev. Microbiol.* **4**, 36 (2006).
11. S. M. Lim, S. A. Webb, *Anaesthesia* **60**, 887 (2005).
12. F. D. Lowy, *N. Engl. J. Med.* **339**, 520 (1998).
13. C. von Eiff, G. Peters, C. Heilmann, *Lancet Infect. Dis.* **2**, 677 (2002).
14. Y. Q. Zhang *et al.*, *Mol. Microbiol.* **49**, 1577 (2003).
15. S. R. Gill *et al.*, *J. Bacteriol.* **187**, 2426 (2005).
16. M. W. Climo, V. K. Sharma, G. L. Archer, *J. Bacteriol.* **178**, 4975 (1996).

17. B. A. Diep *et al.*, *Lancet* **367**, 731 (2006).
18. T. M. Morton, J. L. Johnston, J. Patterson, G. L. Archer, *Antimicrob. Agents Chemother.* **39**, 1272 (1995).
19. Materials and methods are available as supporting material on Science Online.
20. B. N. Kreiswirth *et al.*, *Nature* **305**, 709 (1983).
21. B. L. Golden, H. Kim, E. Chase, *Nat. Struct. Mol. Biol.* **12**, 82 (2005).
22. M. Landthaler, D. A. Shub, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7005 (1999).
23. S. Horinouchi, B. Weisblum, *J. Bacteriol.* **150**, 815 (1982).
24. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
25. I. Chen, D. Dubnau, *Nat. Rev. Microbiol.* **2**, 241 (2004).
26. W. H. Byeon, B. Weisblum, *J. Bacteriol.* **158**, 543 (1984).
27. M. F. Gros, H. te Riele, S. D. Ehrlich, *EMBO J.* **8**, 2711 (1989).
28. M. F. Gros, H. te Riele, S. D. Ehrlich, *EMBO J.* **6**, 3863 (1987).
29. K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, E. V. Koonin, *Biol. Direct* **1**, 7 (2006).
30. We are indebted to O. Schneewind (University of Chicago) for reagents and experimental assistance, B. Golden (Purdue University) for *orf142-2* intron constructs and advice, and members of our laboratory for discussions and comments on the manuscript. Results described here are being used in support of a patent filing by Northwestern University. L.A.M. is a fellow of The Jane Coffin Childs Memorial Fund for Medical Research. This work was supported by a grant (GM072830) from NIH to E.J.S.

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/322/5909/1843/DC1  
Materials and Methods  
Figs. S1 to S4  
Table S1  
References

10 September 2008; accepted 14 November 2008  
10.1126/science.1165771

## Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters

Leighton J. Core,\* Joshua J. Waterfall,\* John T. Lis†

RNA polymerases are highly regulated molecular machines. We present a method (global run-on sequencing, GRO-seq) that maps the position, amount, and orientation of transcriptionally engaged RNA polymerases genome-wide. In this method, nuclear run-on RNA molecules are subjected to large-scale parallel sequencing and mapped to the genome. We show that peaks of promoter-proximal polymerase reside on ~30% of human genes, transcription extends beyond pre-messenger RNA 3' cleavage, and antisense transcription is prevalent. Additionally, most promoters have an engaged polymerase upstream and in an orientation opposite to the annotated gene. This divergent polymerase is associated with active genes but does not elongate effectively beyond the promoter. These results imply that the interplay between polymerases and regulators over broad promoter regions dictates the orientation and efficiency of productive transcription.

Transcription of coding and noncoding RNA molecules by eukaryotic RNA polymerases requires their collaboration with hundreds of transcription factors to direct and control polymerase recruitment, initiation, elongation, and termination. Whole-genome microarrays and ultra-high-throughput sequencing technologies enable efficient mapping of the distribution of transcription factors, nucleosomes, and their modi-

fications, as well as accumulated RNA transcripts throughout genomes (1, 2), thereby providing a global correlation of factors and transcription states. Studies using the chromatin immunoprecipitation assay coupled to genomic DNA microarrays (ChIP-chip) or to high-throughput sequencing (ChIP-seq) indicate that RNA polymerase II (Pol II) is present at disproportionately higher amounts near the 5' end of many eukaryotic