DISPATCH

Transposition: A CRISPR way to get around

Tatiana Dimitriu[1], Ben Ashby[2], Edze R. Westra[1]

[1]Environment and Sustainability Institute, Biosciences, University of Exeter, Cornwall Campus, Penryn, Cornwall TR10 9FE, UK
[2]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, Somerset, BA2 7AY, UK

Email: E.R.Westra@exeter.ac.uk

Summary

CRISPR-Cas systems provide sequence-specific immunity against selfish genetic elements in prokaryotes. Now, two studies show that transposon-encoded variants can guide sequence-specific transposition. These findings have important practical implications but also raise questions of why and how this strategy would benefit transposons.

Although CRISPR-Cas systems are well known for their classical function as adaptive immune systems in bacteria and archaea [1], they can have other intriguing roles as well [2]. Microbiology textbooks will explain how CRISPR-Cas systems protect their host by inserting 'spacer' sequences from phages and plasmids into CRISPR arrays in order to defend the same lineage against re-infection, using Cas proteins loaded with processed CRISPR transcripts (crRNA) to detect and destroy the invading genome (Figure 1A). However, bioinformatic analyses have revealed the presence of CRISPR-Cas systems encoded on some transposons, potentially with a role unrelated to host defense [3,4]. These suspicions were confirmed by two recent studies that teased apart how these CRISPR-Cas variants guide sequence-specific transposition [5,6] (Figure 1B).

Transposons (Tn) are DNA sequences with the ability to excise and insert themselves into new locations within a host genome or into other DNA molecules that are present in the same cell. Some transposons insert themselves randomly across the genome (a property used for creating mutant libraries), whereas others insert into a specific site that is usually conserved across host species. Tn7 is a well-studied transposon best characterized in *Escherichia coli*. Tn7 encodes two transposition pathways: one depends on the site-specific DNA-binding protein TnsD, which targets the *attTn7* insertion site found in many bacteria; the other utilizes the protein TnsE and allows for more random transposition biased towards plasmids, as it recognizes lagging-strand DNA synthesis during replication [7]. Recently, some Tn7-like transposons were found to encode CRISPR-Cas systems [3,4]. Intriguingly, these Tn7 variants lack TnsE homologues, and their associated CRISPR-Cas systems lack both the Cas proteins needed to acquire novel spacers and the nucleolytic activity to cleave targets. However, these 'minimal' systems retain the target recognition genes, suggesting that they may have been repurposed for guiding transposition to targets that are defined by the spacers in the CRISPR arrays [3,4].

Two teams have now experimentally demonstrated exactly this, using independent model systems: Strecker *et al* [5] show that in cyanobacteria, a type V-K CRISPR effector complex guides the associated transposase towards its insertion site, and Klompe *et al* [6] show that a *Vibrio cholerae* transposon has co-opted a variant type I-F CRISPR system to guide transposition. In both cases, crRNAs trigger transposon

integration downstream from the target site, leaving the original target site intact (Figure 1B). Integration itself is mediated by the Tn7 machinery, yielding the characteristic 5-bp duplication 'scar' of Tn7 insertion. Both CRISPR-guided systems therefore require *tns* genes that encode the transposase, namely the *tnsB* gene that catalyses the DNA insertion, the *tnsC* ATPase that regulates target-site selection, and a *tnsD* homologue, known as *tniQ*. In the transposon-encoded type I-F variant, three Cas proteins and a crRNA form a so-called 'Cascade ribonucleoprotein complex' [8,9] that directly interacts with TniQ, hence guiding the transposon machinery to DNA sites that have a sequence complementary to that of the crRNA spacer [6]. For the V-K variant, a similar interaction is predicted to occur between the Cas12k protein loaded with a tracrRNA and a crRNA [5].

The practical implications of this discovery are important: control over transposition target sites by simply changing the crRNA sequence means that transposon integration can easily be programmed using synthetic spacers. With both systems, the authors show high frequency CRISPR-guided integration in *E. coli*, with no need for selection. CRISPR-guided transposition might thus form an extremely useful expansion of the CRISPR toolbox already at our disposal. A key advantage over many of the existing tools is that it does not rely on cleavage and repair of genomic DNA, which risks introducing undesired mutations, for example due to off-target cleavage events [10].

Apart from these applications, CRISPR-guided transposition is also fascinating from an evolutionary perspective. This strategy evolved independently at least three times in Tn7-like elements [3,4], suggesting that using CRISPR spacer information to select an integration site can be an adaptive strategy for transposons. The fact that this strategy is not ubiquitous, however, also suggests that it is costly under certain conditions. What would the costs and benefits of CRISPR-Cas guided transposition be?

One hypothesis is that these Tn7-encoded CRISPR-Cas systems could rapidly acquire spacers in response to infections by mobile genetic elements in order to enhance their horizontal spread. However, as mentioned above, these are minimal CRISPR-Cas systems that lack the ability to insert new spacers, and the existing CRISPR arrays have a limited spacer repertoire, with typically only two spacers per array for the Type I-F CRISPR-Cas variants [3]. Yet, some of these spacers target mobile genetic elements

such as plasmids and phages, consistent with the idea that they may guide transposition into these elements in order to facilitate their horizontal transfer [4]. For transposons, hitchhiking in this way may be essential for their long-term success, especially for parasitic elements that do not benefit the host cells [11]. Indeed, preference for transposition towards mobile genetic elements has been observed: for example, the TnsE-dependent non-specific pathway of Tn7 transposition is heavily biased towards incoming conjugating plasmids [12] and filamentous phages [13]. Given that CRISPR-guided Tn7 transposition is associated with the absence of TnsE homologues, it stands to reason that the ancestral non-specific transposition mechanism has been replaced by a CRISPR-Cas-dependent mechanism in which biased transposition into mobile genetic elements is based on the sequence specificity of the CRISPR memory [4].

However, given that these CRISPR-Cas systems lack the genes needed to capture new spacers, the question remains how they identify rapidly evolving mobile genetic elements. Perhaps these CRISPR spacers target highly conserved sequences in these elements and therefore do not need to be updated? This seems unlikely, since most spacers do not match any sequences present in databases, but this view may change as more sequences of phages and plasmids become available. Alternatively, these systems may take advantage of a host-encoded CRISPR-Cas machinery (Figure 2) in two possible ways. First, host CRISPR adaptation enzymes may capture spacers and insert them into Tn7-encoded CRISPR arrays, as suggested by Strecker *et al* [5]. Second, host spacers might be used by the transposition machinery in *trans*. Host-encoded crRNA are intrinsically enriched for mobile genetic element sequences [14], and may therefore provide valuable information that can be exploited by the transposon to enhance its horizontal spread.

However, relying on active host CRISPR systems for adaptation or transposition might be risky for Tn7-like elements, as there is the obvious danger that the mobile genetic element that has been selected for integration based on host-encoded spacers is at the same time being targeted for destruction by the host CRISPR system. This is particularly relevant given that the transposition event leaves the original target site intact (Figure 1B). The evolutionary success of CRISPR-associated transposition might in this scenario depend on imperfect targeting of mobile genetic elements by the host CRISPR system, for instance because of inhibition by anti-CRISPR proteins [15] or

because of mismatches between spacer and target sequences due to coevolution between mobile genetic elements and CRISPR-Cas immune systems. Mismatches in the spacer sequence reduce the efficiency of transposition significantly in the variant I-F system [6]; however even low rates of transposition may still provide sufficient advantage to transposons to make it worthwhile, and at a strongly reduced risk of cleavage by the host CRISPR-Cas immune system [16,17]. Furthermore, transposition may in this model be naturally biased towards DNA elements with imperfect matches, since perfectly matching targets have a limited lifespan inside a cell.

The ground breaking work of Strecker *et al* [5] and Klompe *et al* [6] provides key insights into the molecular biology and applications of CRISPR-guided transposons, however, future theoretical and experimental work will be critical to explain their existence in nature. For example, it will be fascinating to know if and when transposon-encoded CRISPR arrays are updated *in trans* by host enzymes. It will also be equally important to know whether transposon-encoded CRISPR-Cas systems can be guided by host-derived crRNAs, and if so, what level of sequence dissimilarity in the crRNA repeat sequences is tolerated and how this affects transposition efficiency. Another point to consider is the effect of imperfect targeting by CRISPR-Cas, due to anti-CRISPR mechanisms [18,19] or sequence mismatches [16,17]: do these mechanisms specifically inhibit CRISPR-mediated cleavage by host immune systems, or will they also interfere with CRISPR-guided transposition by transposon-encoded systems? These are important considerations as CRISPR-guided transposition might be particularly adaptive when the infected host cannot defend properly, analogous to a terminal investment strategy. Lastly, if CRISPR-guided transposition does indeed increase horizontal spread of the transposons via mobile genetic elements — and hence reduces dependence on vertical transmission in the host — does this select for more parasitic transposons? In general, horizontal transmission should select for greater parasitism when compared to vertical transmission [20], and if CRISPR-Cas enhances rates of horizontal transmission over that of transposons that rely solely on TnsE, they would be expected to drive evolution of more selfish transposons. Such experimental data will help to generate models that may shed light on the conditions where one would expect selection for and against CRISPR-guided transposition in hosts with or without CRISPR-Cas immune systems. These models will need to take into consideration that spacers encoded by the host immune system may on the one hand provide a reliable

signal for successful mobile genetic elements, but on the other hand also provide a signal for imminent cleavage of that element's genome and high levels of resistance in the wider population.

References

1. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.
2. Westra, E.R., Buckling, A., and Fineran, P.C. (2014). CRISPR-Cas systems: beyond adaptive immunity. Nat. Rev. Microbiol. *12*, 317–326.
3. Faure, G., Shmakov, S.A., Yan, W.X., Cheng, D.R., Scott, D.A., Peters, J.E., Makarova, K.S., and Koonin, E.V. (2019). CRISPR-Cas in mobile genetic elements: counter-defence and beyond. Nat. Rev. Microbiol. *17*, 513–525.
4. Peters, J.E., Makarova, K.S., Shmakov, S., and Koonin, E.V. (2017). Recruitment of CRISPR-Cas systems by Tn7-like transposons. Proc. Natl. Acad. Sci. USA *114*, E7358–E7366.
5. Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J.L., Makarova, K.S., Koonin, E.V., and Zhang, F. (2019). RNA-guided DNA insertion with CRISPR-associated transposases. Science *365*, 48–53.
6. Klompe, S.E., Vo, P.L.H., Halpin-Healy, T.S., and Sternberg, S.H. (2019). Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. Nature *571*, 219–225.
7. Parks, A.R., and Peters, J.E. (2009). Tn7 elements: engendering diversity from chromosomes to episomes. Plasmid *61*, 1–14.
8. Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J., *et al.* (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. USA *108*, 10092–10097.
9. Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964.

10. Pawelczak, K.S., Gavande, N.S., VanderVere-Carozza, P.S., and Turchi, J.J. (2018). Modulating DNA repair pathways to improve precision genome engineering. ACS Chem. Biol. *13*, 389–396.
11. Condit, R., Steward, F., and Levin, B.R. (1988). The population biology of bacterial transposons: A priori conditions for maintenance as parasitic DNA. Am. Nat. *132*, 129–147.
12. Wolkow, C.A., DeBoy, R.T., and Craig, N.L. (1996). Conjugating plasmids are preferred targets for Tn7. Genes Dev. *10*, 2145–2157.
13. Finn, J.A., Parks, A.R., and Peters, J.E. (2007). Transposon Tn7 directs transposition into the genome of filamentous bacteriophage M13 using the element-encoded TnsE protein. J. Bacteriol. *189*, 9122–9125.
14. Shmakov, S.A., Sitnik, V., Makarova, K.S., Wolf, Y.I., Severinov, K.V., and Koonin, E.V. (2017). The CRISPR spacer space is dominated by sequences from species-specific mobilomes. MBio *8*, e01397-17.
15. Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature *493*, 429–432.
16. Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J. Bacteriol. *190*, 1390–1400.
17. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. USA *108*, 10098–10103.
18. Landsberger, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, H., Buckling, A., Westra, E.R., and van Houte, S. (2018). Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. Cell *174*, 908–916.
19. Borges, A.L., Zhang, J.Y., Rollins, M.F., Osuna, B.A., Wiedenheft, B., and Bondy-Denomy, J. (2018). Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. Cell *174*, 917–925.

20. Lipsitch, M., Siller, S., and Nowak, M.A. (1996). The evolution of virulence in pathogens with vertical and horizontal transmission. Evolution *50*, 1729–1741.

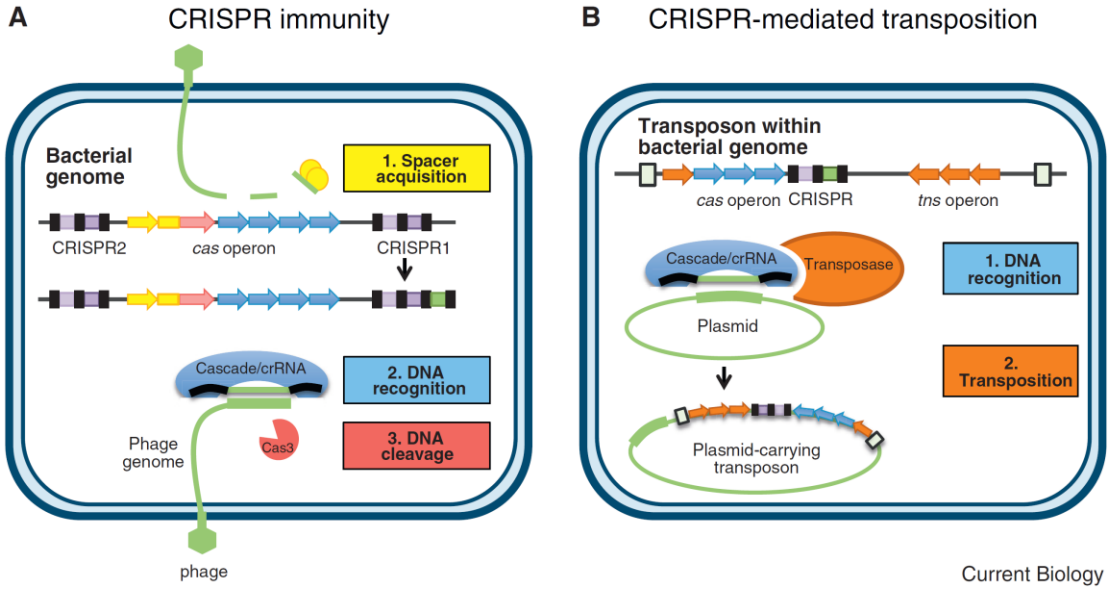

Figure 1. Model of CRISPR-mediated targeting for adaptive immunity and transposition.

A) Model of Type I-F CRISPR-mediated adaptive immunity. Cas proteins mediate spacer acquisition. Mobile genetic elements carrying the target sequence are recognized by the Cascade-crRNA complex. Subsequent recruitment of a Cas nuclease triggers DNA cleavage. B) Model of Type I-F CRISPR-guided transposition. The transposon-encoded Cascade-crRNA complex recruits the Tn7 transposase to a target site on a mobile genetic element that is defined by the crRNA sequence. This triggers transposon integration within the mobile element, close to the target site.

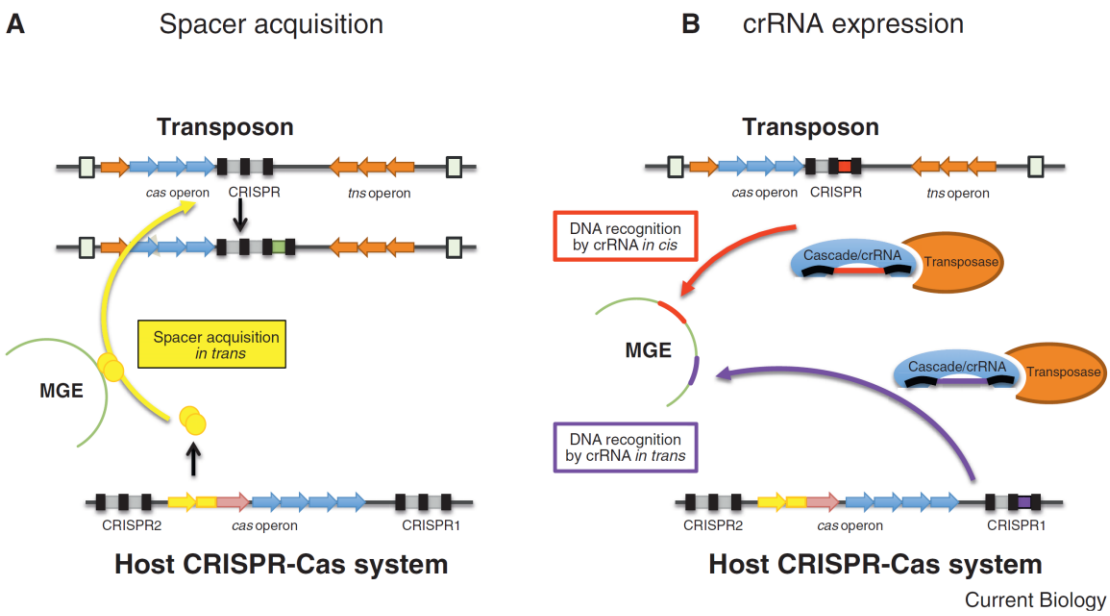

Figure 2. Minimal CRISPR-Cas systems encoded by transposons could make use of host-encoded Cas proteins or CRISPR memory sequences.

A) Host-encoded spacer acquisition enzymes could integrate new spacers into a CRISPR array that is located on the transposon — a functionality that is lacking from the minimal CRISPR-Cas systems of the transposon. B) CRISPR-guided transposition could use spacers from arrays that are part of the host CRISPR-Cas immune system as well as those encoded by the transposon CRISPR array. MGE, mobile genetic element.