Cargo capacity of phages and plasmids and other factors influencing horizontal transfers of prokaryote transposable elements

Sébastien Leclercq,[1,2] Clément Gilbert[1] and Richard Cordaux[1,*]

[1]Université de Poitiers; UMR CNRS 7267 Ecologie et Biologie des Interactions; Equipe Ecologie Evolution Symbiose; Poitiers, France;
[2]Max Planck Institute for Evolutionary Anthropology; Leipzig, Germany

Horizontal transfer of transposable elements (TEs) plays a key role in prokaryote genome evolution. Most TEs do not encode the enzymatic machinery allowing them to transfer between host cells and it is widely assumed in the literature that horizontal transfer of prokaryote TEs is mediated by other mobile genetic elements such as phages and plasmids. In a recent study, we have shown that phages are less tolerant to insertion sequences (IS, the most frequent class of prokaryote TEs) and therefore have a lower cargo capacity than plasmids. Consequently, while our analysis confirmed the crucial role of plasmids as efficient vehicles of IS horizontal transfer, we concluded that phages are unlikely to efficiently shuttle IS elements between prokaryotes. Here, we discuss whether or not the distribution pattern observed for IS elements in phages and plasmids also holds for other TEs, such as transposons and mobile introns. We also further explore various factors that may impact the relative capacity of phages and plasmids to mediate TE horizontal transfer among prokaryotes.

Introduction

Insertion sequences (IS) are short prokaryote transposable elements (TEs) characterized by a simple structure, often consisting of a single transposase-encoding open reading frame flanked by terminal inverted repeats.[1] Because their proliferation can induce variation in genome size, activate and inactivate gene expression and trigger various types of genomic rearrangements, IS have a profound influence on prokaryote genome evolution.[2,3] The general lack of congruence between IS and host phylogenies, together with their overall patchy distribution among hosts, suggest that IS distribution among prokaryotes is strongly shaped by horizontal transfer.[4,5] Unlike IS, bacteriophages and plasmids encode the enzymatic machinery allowing them to move between cells. These two types of mobile genetic elements (MGEs) have traditionally been considered as the main vectors of IS horizontal transfer.[3] However the relative contribution of phages and plasmids to IS horizontal transfer has never been thoroughly evaluated.

We recently reported a comprehensive survey of the abundance and distribution of IS elements integrated in 900 phage and plasmid annotated genomes available in the ACLAME database.[6] This study revealed that while over half of plasmids contain at least one IS (with a density of one copy every 19 kb on average), the vast majority (92%) of phages are completely devoid of IS insertion (one copy every 346 kb on average). We further showed that the density of deleterious insertion sites is significantly higher in phages than in plasmids. Consequently, acquisition of IS elements is more strongly counter-selected in phages than in plasmids. Based on these results, we concluded that phages are less tolerant to IS insertions than plasmids, i.e., phages have a lower cargo capacity than plasmids. Therefore, while our analysis confirmed the classically held view that plasmids are efficient

vectors of IS horizontal transfer, we concluded that phages may be rather poor vectors of IS horizontal transfer among prokaryotes. If phages indeed do possess a lower cargo capacity than plasmids do, then we expect the distribution pattern observed for IS elements to also hold for other TEs. Does the available evidence support this prediction?

Do Phages and Plasmids Shuttle TEs Other than IS Among Prokaryotes?

Transposons are certainly the most thoroughly studied bacterial TEs. They comprise elements with various modes of mobility: (1) composite (formerly class I) transposons are genetic sequences flanked by IS elements which use the IS transposase for their mobility, (2) unit transposons (formerly type II/Tn3 family) which use specific transposase/resolvase genes for their mobility, and (3) conjugative transposons/ICEs which can transfer from a bacterial cell to another through conjugation.[7] Transposons are often found in plasmids.[8] On the contrary, to our knowledge, there has been no report of a transposon being carried by an active bacteriophage. The distribution of transposons in plasmids and phages therefore seems fully consistent with higher selective constrains on phages compared with plasmids.

Mobile introns are other bacterial TEs which are able to self-splice from the transcript they are inserted in. They include two unrelated types of elements: group I introns (which move through a homologous recombination process termed homing) and group II introns (which move through reverse-transcription and integration of an RNA intermediate in a process termed retrohoming).[9] Group II introns are widespread in the bacterial world[10,11] and frequently observed within plasmid sequences.[12,13] To our knowledge, no group II intron has yet been reported in an active bacteriophage sequence. The distribution of group II introns in plasmids and phages thus apparently reflects that of IS elements.[6] Recent studies suggest that group II intron insertions may be deleterious for hosts despite their splicing ability.[3] This is in agreement with their apparent absence in phages. While the greater tolerance of plasmids to group II introns is likely due to lower selective constraints acting on their genomes compared with phages genomes, the presence of some group II introns in plasmids may also be explained by their insertional preferences. For example, the well-studied group II intron Ll.ltrB performs specific retrohoming into relaxase genes, which are parts of the plasmid mobility machinery.[14] Analogously, members of the GIIC-attC intron class target integrons themselves often integrated within plasmids.[15] Finally, the targeting of IS elements by various group II introns[16,17] may also account for group II intron preferential presence in plasmids compared with phages.

In contrast with group II introns, group I introns are preferentially found in phages rather than plasmids. Many group I introns have indeed been detected in phage sequences[18-20] but to our knowledge, only two copies have been reported in a plasmid.[13] Furthermore, these two plasmid-borne group I introns are inserted within a prophage which is itself inserted into the plasmid genome. Therefore, there is as yet no evidence that a group I intron has jumped into a plasmid genome. The pattern of group I intron distribution in MGEs is puzzling as it is contradictory to that observed for other TEs, but it may be explained by the specific mobility mechanism used by group I introns. Indeed, group I introns move from an intron-containing locus to a highly similar intron-free locus (generally another allele of the same locus) through a recombination-mediated process. They produce a homing endonuclease which specifically cleaves the target allele. The break is then repaired by the host homologous recombination machinery, which uses the intron-carrying allele as a template.[21] As homologous recombination proceeds only on highly similar regions, the spread of group I introns relies on their ability to invade conserved regions. Accordingly, chromosomal group I introns are usually found in tRNA and rRNA genes or other essential genes, while most phage-borne group I introns are found in genes important for phage biology.[13,21] In this context, the absence of group I intron in plasmids might simply reflect the fact that these site-specific elements never evolved to target plasmid genes.

In sum, this brief overview of the distribution of the main bacterial TEs in plasmids and phages globally supports our previous observations on IS elements.[6] It is fully consistent with higher selective constraints applying to phage genomes compared with plasmid genomes.

Factors Influencing Phage and Plasmid Ability to Mediate Horizontal Transfers of Transposable Elements

In addition to the density of deleterious insertion sites, other factors may also contribute to the basis for strong negative selection against TE insertions in phages and plasmids, and thus their cargo capacity. For example, growth rates of plasmid-carrying bacteria are negatively correlated with plasmid genome size.[22] Accumulation of TEs should thus be detrimental to plasmids. However, the decrease in growth rate induced by plasmid genome size increase is negligible compared with that induced by highly expressed genes.[22] Thus, a small plasmid with a single highly expressed gene may be much more counter-selected than a large plasmid with many lowly expressed TEs. Thus, constraints on genome size may not be too severe for plasmids. By contrast, genome size is a much more salient constraint in phages because capsids (i.e., phage container particles) cannot package more than 105–106% of the wild type genome size.[23] With a median size of 42 kb for phage genomes,[6] this means that half of phage genomes cannot package more than 2.5 kb of additional sequence, which translates into a couple of IS elements or one group II intron, at most. Thus, constraints on genome size also argue for a larger cargo capacity for plasmids relative to phages.

Limitation on cargo capacity provides a basis for strong purifying selection of TE insertions. However, whatever its strength, purifying selection has to be efficient in order to effectively purge MGE genomes from deleterious TE insertions. Population genetics theory indicates that selection efficiency depends on effective population size: the higher the effective population

size, the more efficient the selection. Phages are the most numerous organisms on Earth and their effective population size is presumably enormous.[24] Thus, in addition to the fact that their genome contains a higher proportion of deleterious sites, the efficiency of purifying selection in phages is likely to be much higher than in plasmids: for an equally deleterious insertion site, purifying selection is expected to remove a TE quicker from a phage population than from a plasmid population. On the other hand, many more low-frequency neutral TE insertions are expected in (large) phage populations than in (smaller) plasmid populations. Even though the lifespan of individual insertions is very short, they may collectively represent a dynamic reservoir of TEs that might be transferable between bacterial cells. The magnitude of such a phenomenon is currently difficult to estimate, but the availability of an increasing amount of metagenomic data is likely to provide new insights into this question in the near future. If this hypothesis was verified, it would argue for a potentially more important role of phages in the horizontal transfer of TEs than the current evidence suggests. Nevertheless, another population genetics parameter to keep in mind for a horizontal transfer to be successful is not only vector fitness, but also bacterial host fitness. Indeed, a large number of potential transfers initiated by phages simply cannot lead to successful TE horizontal transfers because they result in the death of the bacterial hosts. On the contrary, plasmid transfers can bring new genes without jeopardizing bacterial host survival.

Another factor that might affect the efficiency with which phages and plasmids can shuttle TEs between bacteria is the breadth of MGE host range. While a substantial number of broad host-range plasmids and phages have been characterized,[25-30] the majority of plasmids and phages are believed to have a relatively narrow host range and to be capable of transferring genetic material only between closely related species of the same genus.[31,32] This trend is indeed consistent with the observation that horizontal transfer of both IS and non-IS DNA is more frequent between closely related bacteria than between distantly related bacterial lineages.[4,33] Overall, the available data on MGE host range do not seem to indicate any dramatic difference between typical phage and plasmid host ranges. However, studies of MGE host ranges remain relatively scarce and are often plagued with various pitfalls and limitations.[32] A method was recently developed to predict plasmid host range based on genomic signatures expected to be shared (or not) between plasmids and bacterial chromosomes, depending on plasmid host range breadth.[34] This method, which exclusively relies on sequence data, was first validated on various plasmids belonging to six well-studied incompatibility groups (Inc.F, Inc.H, Inc.I, Inc.N, Inc.P and Inc.W) and was then applied to other less studied plasmids (Inc.Q, Inc.U, PromA, Inc.A/C and Inc.P-9) for which host range was mostly inferred to be wide. Given the rate at which new MGEs and bacterial genomes are currently added to public databases, it will certainly help to better delineate the host range of numerous MGEs. It is however important to keep in mind that simple entry of a TE-carrying MGE into a new bacterial cell is potentially sufficient for TE horizontal transfer to happen. In other words, MGEs can facilitate TE horizontal transfer even in the absence of successful phage replication or plasmid maintenance. Therefore, the host range relevant to a particular MGE (i.e., the number of bacterial species in which a phage can successfully replicate or in which a plasmid can be maintained) might be narrower than the host range truly relevant to TE horizontal transfer.[29]

Conclusion

Our comparison of IS abundance in phages and plasmids has provided the first formal test of the relative contributions of phages and plasmids to IS horizontal transfer between bacterial cells, suggesting that plasmids are better vectors of IS horizontal transfer than phages.[6] This result seems to also hold for other TEs, such as transposons and group II introns. In fact, the better capacity of plasmids to shuttle TEs compared with phages is probably not due to an intrinsic property of TEs, as a recent study showed that this trend seems to apply to all types of DNA sequences.[35] We have summarized a number of potential underlying causes for the differential implication of phages and plasmids in the horizontal transfer of TEs. While strength and efficiency of selection are clearly major players in this game, a number of other parameters (e.g., low-frequency insertions in phages and MGE host range) may have a non-trivial impact as well. However, the actual magnitude of the different factors currently remains difficult to evaluate with accuracy. There is no doubt that the ever-increasing amount of genomic data being made available will soon allow for refined estimates of these parameters.

References

1. Siguier P, Filée J, Chandler M. Insertion sequences in prokaryotic genomes. Curr Opin Microbiol 2006; 9:526-31; PMID:16935554; http://dx.doi.org/10.1016/j.mib.2006.08.005

2. Treangen TJ, Abraham AL, Touchon M, Rocha EP. Genesis, effects and fates of repeats in prokaryotic genomes. FEMS Microbiol Rev 2009; 33:539-71; PMID:19396957; http://dx.doi.org/10.1111/j.1574-6976.2009.00169.x

3. Cerveau N, Leclercq S, Bouchon D, Cordaux R. Evolutionary dynamics and genomic impact of prokaryote transposable elements. In: Pontarotti P, ed. Evolutionary Biology: Concepts, Biodiversity, Macroevolution and Genome Evolution. Berlin: Springer, 2011:291-312.

4. Wagner A, de la Chaux N. Distant horizontal gene transfer is rare for multiple families of prokaryotic insertion sequences. Mol Genet Genomics 2008; 280:397-408; PMID:18751731; http://dx.doi.org/10.1007/s00438-008-0373-y

5. Cerveau N, Leclercq S, Leroy E, Bouchon D, Cordaux R. Short- and long-term evolutionary dynamics of bacterial insertion sequences: insights from Wolbachia endosymbionts. Genome Biol Evol 2011; 3:1175-86; PMID:21940637; http://dx.doi.org/10.1093/gbe/evr096

6. Leclercq S, Cordaux R. Do phages efficiently shuttle transposable elements among prokaryotes? Evolution 2011; 65:3327-31; PMID:22023596; http://dx.doi.org/10.1111/j.1558-5646.2011.01395.x

7. Roberts AP, Chandler M, Courvalin P, Guédon G, Mullany P, Pembroke T, et al. Revised nomenclature for transposable genetic elements. Plasmid 2008; 60:167-73; PMID:18778731; http://dx.doi.org/10.1016/j.plasmid.2008.08.001

8. Tan HM. Bacterial catabolic transposons. Appl Microbiol Biotechnol 1999; 51:1-12; PMID:10077818; http://dx.doi.org/10.1007/s002530051356

9. Edgell DR, Chalamcharla VR, Belfort M. Learning to live together: mutualism between self-splicing introns and their hosts. BMC Biol 2011; 9:22; PMID:21481283; http://dx.doi.org/10.1186/1741-7007-9-22

10. Candales MA, Duong A, Hood KS, Li T, Neufeld RA, Sun R, et al. Database for bacterial group II introns. Nucleic Acids Res 2012; 40(Database issue):D187-90; PMID:22080509; http://dx.doi.org/10.1093/nar/gkr1043

11. Leclercq S, Giraud I, Cordaux R. Remarkable abundance and evolution of mobile group II introns in Wolbachia bacterial endosymbionts. Mol Biol Evol 2011; 28:685-97; PMID:20819906; http://dx.doi.org/10.1093/molbev/msq238

12. Klein JR, Dunny GM. Bacterial group II introns and their association with mobile genetic elements. Front Biosci 2002; 7:d1843-56.

13. Tourasse NJ, Kolstø AB. Survey of group I and group II introns in 29 sequenced genomes of the Bacillus cereus group: insights into their spread and evolution. Nucleic Acids Res 2008; 36:4529-48; PMID:18587153; http://dx.doi.org/10.1093/nar/gkn372

14. Staddon JH, Bryan EM, Manias DA, Dunny GM. Conserved target for group II intron insertion in relaxase genes of conjugative elements of gram-positive bacteria. J Bacteriol 2004; 186:2393-401; PMID:15060042; http://dx.doi.org/10.1128/JB.186.8.2393-2401.2004

15. Léon G, Roy PH. Potential role of group IIC-attC introns in integron cassette formation. J Bacteriol 2009; 191:6040-51; PMID:19633079; http://dx.doi.org/10.1128/JB.00674-09

16. Dai L, Zimmerly S. The dispersal of five group II introns among natural populations of Escherichia coli. RNA 2002; 8:1294-307; PMID:12403467; http://dx.doi.org/10.1017/S1355838202023014

17. Biondi EG, Toro N, Bazzicalupo M, Martínez-Abarca F. Spread of the group II intron RmInt1 and its insertion sequence target sites in the plant endosymbiont Sinorhizobium meliloti. Mob Genet Elements 2011; 1:2-7; PMID:22016840; http://dx.doi.org/10.4161/mge.1.1.15316

18. Kwan T, Liu J, DuBow M, Gros P, Pelletier J. The complete genomes and proteomes of 27 Staphylococcus aureus bacteriophages. Proc Natl Acad Sci U S A 2005; 102:5174-9; PMID:15788529; http://dx.doi.org/10.1073/pnas.0501140102

19. Lazarevic V, Soldo B, Düsterhöft A, Hilbert H, Mauël C, Karamata D. Introns and intein coding sequence in the ribonucleotide reductase genes of Bacillus subtilis temperate bacteriophage SPbeta. Proc Natl Acad Sci U S A 1998; 95:1692-7; PMID:9465078; http://dx.doi.org/10.1073/pnas.95.4.1692

20. Edgell DR, Gibb EA, Belfort M. Mobile DNA elements in T4 and related phages. Virol J 2010; 7:290; PMID:21029434; http://dx.doi.org/10.1186/1743-422X-7-290

21. Raghavan R, Minnick MF. Group I introns and inteins: disparate origins but convergent parasitic strategies. J Bacteriol 2009; 191:6193-202; PMID:19666710; http://dx.doi.org/10.1128/JB.00675-09

22. Diaz Ricci JC, Hernández ME. Plasmid effects on Escherichia coli metabolism. Crit Rev Biotechnol 2000; 20:79-108; PMID:10890453; http://dx.doi.org/10.1080/07388550008984167

23. Nurmemmedov E, Castelnovo M, Catalano CE, Evilevitch A. Biophysics of viral infectivity: matching genome length with capsid size. Q Rev Biophys 2007; 40:327-56; PMID:18423102; http://dx.doi.org/10.1017/S0033583508004666

24. Breitbart M, Rohwer F. Here a virus, there a virus, everywhere the same virus? Trends Microbiol 2005; 13:278-84; PMID:15936660; http://dx.doi.org/10.1016/j.tim.2005.04.003

25. Götz A, Pukall R, Smit E, Tietze E, Prager R, Tschäpe H, et al. Detection and characterization of broad-host-range plasmids in environmental bacteria by PCR. Appl Environ Microbiol 1996; 62:2621-8; PMID:8779598

26. Hoffmann A, Thimm T, Dröge M, Moore ER, Munch JC, Tebbe CC. Intergeneric transfer of conjugative and mobilizable plasmids harbored by Escherichia coli in the gut of the soil microarthropod Folsomia candida (Collembola). Appl Environ Microbiol 1998; 64:2652-9; PMID:9647844

27. Jensen EC, Schrader HS, Rieland B, Thompson TL, Lee KW, Nickerson KW, et al. Prevalence of broad-host-range lytic bacteriophages of Sphaerotilus natans, Escherichia coli, and Pseudomonas aeruginosa. Appl Environ Microbiol 1998; 64:575-80; PMID:9464396

28. Beumer A, Robinson JB. A broad-host-range, generalized transducing phage (SN-T) acquires 16S rRNA genes from different genera of bacteria. Appl Environ Microbiol 2005; 71:8301-4; PMID:16332816; http://dx.doi.org/10.1128/AEM.71.12.8301-8304.2005

29. Smorawinska M, Szuplewska M, Zaleski P, Wawrzyniak P, Maj A, Plucienniczak A, et al. Mobilizable narrow host range plasmids as natural suicide vectors enabling horizontal gene transfer among distantly related bacterial species. FEMS Microbiol Lett 2012; 326:76-82; PMID:22092700; http://dx.doi.org/10.1111/j.1574-6968.2011.02432.x

30. Kim JH, Son JS, Choi YJ, Choresca CH, Jr., Shin SP, Han JE, et al. Complete genome sequence and characterization of a broad-host range T4-like bacteriophage phiAS5 infecting Aeromonas salmonicida subsp. salmonicida. Vet Microbiol 2012; 157:164-71; PMID:22226819; http://dx.doi.org/10.1016/j.vetmic.2011.12.016

31. Sørensen SJ, Bailey M, Hansen LH, Kroer N, Wuertz S. Studying plasmid horizontal transfer in situ: a critical review. Nat Rev Microbiol 2005; 3:700-10; PMID:16138098; http://dx.doi.org/10.1038/nrmicro1232

32. Hyman P, Abedon ST. Bacteriophage host range and bacterial resistance. Adv Appl Microbiol 2010; 70:217-48; PMID:20359459; http://dx.doi.org/10.1016/S0065-2164(10)70007-1

33. Andam CP, Gogarten JP. Biased gene transfer in microbial evolution. Nat Rev Microbiol 2011; 9:543-55; PMID:21666709; http://dx.doi.org/10.1038/nrmicro2593

34. Suzuki H, Yano H, Brown CJ, Top EM. Predicting plasmid promiscuity based on genomic signature. J Bacteriol 2010; 192:6045-55; PMID:20851899; http://dx.doi.org/10.1128/JB.00277-10

35. Halary S, Leigh JW, Cheaib B, Lopez P, Bapteste E. Network analyses structure genetic diversity in independent genetic worlds. Proc Natl Acad Sci U S A 2010; 107:127-32; PMID:20007769; http://dx.doi.org/10.1073/pnas.0908978107