View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Current Biology, Vol. 12, R276–R278, April 16, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0960-9822(02)00799-6

Selfish DNA: New Abode for Homing Endonucleases

Dispatch

David R. Edgell

A 30-year old conundrum concerning the genetics of T-even bacteriophages has at last been solved, the answer turning out to involve free-standing homologs of intron-encoded homing endonucleases.

The elucidation of the genetic code, the discovery of messenger RNA and the finding of bacterial introns are but a few seminal discoveries stemming from studies of *Escherichia coli* bacteriophage T4 [1]. Phage T4 and its relatives, the T-even phages, were among the first 'model' organisms, as modern-day molecular biologists might call them, and are still the system of choice for some experimental biologists. Recent work from the laboratory of David Shub [2] extends the long history of interesting observations on T-even phage biology, as it has revealed the molecular basis of a long-standing, but unexplained, phenomenon called marker exclusion [3,4]. In the process, the work uncovered an unexpected and fascinating link to self-splicing group I introns and the homing endonucleases encoded within them [5,6].

Marker exclusion was first identified in the 1970s, when the term partial exclusion was coined to describe the unequal inheritance of T2 genetic markers in the progeny resulting from co-infection with phage T4 (Figure 1A). Russell and Huskey [4] found that T2 markers were, at most, represented at a frequency of 30% in progeny, instead of the expected 50% (termed general or phage exclusion by Shub and colleagues [2]). More surprising was the finding that there were two regions of the T2 genome, centered on gene *32* and genes *39-49*, which are preferentially excluded and represented in less than 1% of progeny (localized marker exclusion [2]). Such strong exclusion of T2 markers was unexpected because T4 and T2 are very similar: mutations in genes of one phage can be complemented by genes of the other phage, the genetic maps are co-linear over much of the genome, and recombination is readily observed between T4 and T2 during co-infection.

Marker exclusion has also been described for other phages, notably SP01 and SP82 which infect *Bacillus subtilis* [7]. Previous work in the Shub lab [8–10] revealed that exclusion of SP01 markers from progeny phage of mixed infections with SP01 and SP82 is the result of an intron homing-like process initiated by an intron-encoded endonuclease. Inspired by the finding of endonuclease-mediated marker exclusion in gram-positive phages, Shub and colleagues considered the possibility that the localized marker exclusion

Molecular Genetics Program, Wadsworth Center, New York State Department of Health, Albany, NY 12201-2002, USA.
E-mail: edgell@wadsworth.org

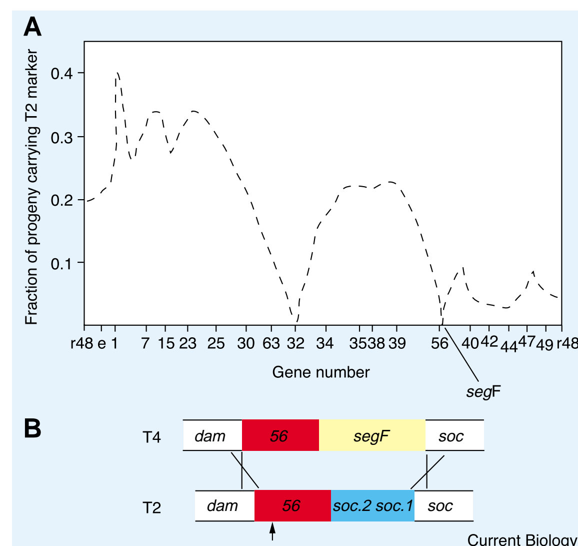

Figure 1. Unequal inheritance of T2 and T4 genetic markers in progeny phage of a mixed infection.

(A) This plot shows the frequency of T2 genetic markers in progeny phage resulting from mixed infection with T4. T4 gene *segF* is located in one of the strongest regions of exclusions, centered on gene *56*. (Adapted from [4].) (B) Schematic of the genomic region surrounding gene *56* of phages T2 and T4. T4 gene *segF* (yellow) cleaves T2 in gene *56* (red), as indicated by an arrow. The double-strand break is repaired using T4 DNA as a template, most likely initiated by strand invasion utilizing the conserved *dam* and *soc* genes.

phenomenon in T-even phage was mediated by a similar process. One of the strongest regions of exclusion mapped was to gene *56* of T2 [4]. Intriguingly, this region is polymorphic between the two phages; T4 has an adjacent gene *69* that is absent from T2, which instead has two non-homologous genes, *soc.2* and *soc.1* (Figure 1B). Database searches with gene *69* of T4 revealed weak similarity to the *seg* genes of phage T4, as well as intron-encoded endonucleases of the GIY-YIG family. Five *seg* — for 'similarity to endonucleases encoded by group I introns' — genes have been described in the T4 genome as free-standing open reading frames [11]. Two of the proteins encoded by these genes, SegA and SegE, have been shown to have double-strand DNA endonuclease activity, but given that they were not intron-encoded, the biological relevance of these endonucleases was unclear [12,13].

That T4 gene *69* — renamed *seg*F as the sixth member of the *seg* family — had similarity to known DNA endonucleases, and mapped to a position of strong exclusion, suggested that cleavage of T2 by SegF would initiate a localized gene conversion event similar to that of homing endonucleases of group I introns [5]. Repair of the double-strand break in the T2 genome using T4 DNA as a template would result in replacement of T2 markers by *seg*F and flanking T4

markers. In support of this hypothesis, Belle and co-workers were able to show that SegF is a double-strand DNA endonuclease that preferentially cleaves T2 gene *56* over T4 gene *56 in vitro* and *in vivo* [2]. Furthermore, mutation of *seg*F alleviated the exclusion of T2 markers as the frequency of T2 gene *56* rose in mixed infections with mutant T4 *seg*F phage. Taken together, these results implicate a free-standing homolog of an intron-encoded endonuclease as the active agent of marker exclusion.

The simple fact that *seg*F is not intron encoded but inserted within a polymorphic region has important implications for the recognition and cleavage sites of SegF. Intron-containing alleles are immune to cleavage by their own endonucleases because the endonuclease's recognition sequence is interrupted by the presence of the intron [14]. If SegF were to use a similar strategy as intron endonucleases, the SegF recognition sequence would include sequences up- and downstream of the *seg*F gene, analogous to exon sequences surrounding an intron. However, sequences downstream of the SegF insertion site — the optional *soc.1* and *soc.2* — are not ideal choices for a recognition sequence because they would be poorly conserved between T-even phages. Potential SegF recognition sites must instead lie within DNA sequences likely to be conserved between T-even phages, such as the coding regions of essential genes that flank the *seg*F insertion site. But by choosing a recognition site that is present in any T-even genome, SegF now faces the problem of cleaving its own T4 DNA. Although the SegF cleavage sites lie in a 31 base pair patch of homology between T2 and T4 gene *56*, SegF effectively discriminates between the two DNAs, possibly by recognizing nucleotide differences at either end of the conserved sequence block.

In addition to the six *seg* genes, T4 has an additional seven open reading frames encoding proteins that show sequence similarity to GIY-YIG or H-N-H family homing endonucleases [15]. Some of these open reading frames are intercistronic and positioned near regions of marker exclusion mapped by Russell and Huskey [4], with the implication that these endonucleases mediate a process similar to that catalysed by SegF, excluding T2 markers from the progeny of mixed infections. Belle *et al.* [2] suggest that cumulative affect of all thirteen endonucleases is responsible for the generalized exclusion of T2 markers by T4.

Surprisingly, many of the free-standing open reading frames which encode endonucleases in T4 are absent from T-even genomes, raising the interesting question of how and why T4 ended up with such a large collection of these genes. Regardless of the molecular pathways by which their coding sequences arrived in the T4 genome, the cumulative effect of thirteen endonucleases, each with the potential to cleave at multiple positions in the T4 genome, must have placed very strong selection on the endonucleases and T4 itself. Of course, all the endonucleases need not have arrived in the genome at once, and T4 may have bootstrapped itself to survive life with non-specific endonucleases by adapting to one or a few at a time.

This scenario is favored by Belle *et al.* [2], who suggest that T4 evolved an über-DNA repair system, efficiently repairing double-strand breaks in its genome in response to non-specific cleavages generated by these endonucleases [2]. Once in place, the efficient repair system would allow T4 to tolerate the presence of multiple intergenic endonucleases. Conversely, free-standing endonucleases might be tolerated because, as DNA replication in T4 occurs in a recombination-dependent context [16], non-specific double-strand breaks might actually be beneficial.

One cannot help but wonder if the presence of thirteen endonucleases is of some benefit to T4, as they (probably) all mediate a process similar to that catalysed by SegF, ensuring the spread of T4 alleles throughout the population. Questions such as this are problematic, because T-even phages appear to freely exchange large parts of their genomes — for instance, head and tail fibre genes [17,18] — suggesting that our current view of T2 or T4 is just a snapshot of a dynamically evolving genome. In this sense, T4 does not benefit from excluding T2 markers in progeny of mixed infections, because both phages are collections of genes that are shared between phages within a larger gene pool.

The fact that T4 does exist within a larger phage gene pool makes the current distribution of SegF and its relatives even more remarkable: why has SegF not been more successful in colonizing other T-even phage genomes, when it has evolved specific mechanisms to do so? Regardless of how one thinks of this problem, the beauty of using T4 as an experimental system is exactly that — the ease by which hypotheses can be addressed by rigorous experimentation. SegF and its free-standing relatives provide an excellent system for studying the diverse strategies that selfish genetic elements adopt to ensure their spread.

References

1. Karam, J.D. (1994). Foreword, In Molecular Biology of Bacteriophage T4, J.D. Karam, ed. (Washington, ASM Press), pp. xvii.
2. Belle, A., Landthaler, M. and Shub, D.A. (2002). Intronless homing: site-specific endonuclease SegF of bacteriophage T4 mediates localized marker exclusion analogous to homing endonucleases of group I introns. Genes Dev. *16*, 351–362.
3. Pees, E. and De Groot, B. (1970). Partial exclusion of genes of bacteriophage T2 with T4-glucosylated DNA in crosses with bacteriophage T4. Genetica *41*, 541–550.
4. Russell, R.L. and Huskey, R.J. (1974). Partial exclusion between T-even bacteriophages: an incipient genetic isolation mechanism. Genetics *78*, 989–1014.
5. Lambowitz, A.M. and Belfort, M. (1993). Introns as mobile genetic elements. Annu. Rev. Biochem. *62*, 587–622.
6. Shub, D.A., Coetzee, T., Hall, D.W. and Belfort M. (1994). The self-splicing introns of bacteriophage T4. In Molecular Biology of Bacteriophage T4, J.D. Karam, ed. (ASM Press, Washington), pp. 186–192.
7. Stewart, C.R. and Frank, M. (1981). Predominance of bacteriophage SP82 over bacteriophage SP01 in mixed infections of *Bacillus subtilis*. J. Virol. *38*, 1081–1083.
8. Goodrich-Blair, H. and Shub, D.A. (1994). The DNA polymerase genes of several HMU-bacteriophages have similar group I introns with highly divergent open reading frames. Nucleic Acids Res. *22*, 3715–3721.
9. Goodrich-Blair, H. and Shub, D.A. (1996). Beyond homing: competition between intron endonucleases confers a selective advantage on flanking genetic markers. Cell *84*, 211–221.
10. Edgell, D.R., Fast, N.M. and Doolittle, W.F. (1996). Selfish DNA: The best defense is a good offense. Curr. Biol. *6*, 385–388.

11. Sharma, M. and Hinton, D.M. (1992). Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of phage and fungi. Proc. Natl. Acad. Sci. U.S.A. *89*, 6658–6662.
12. Sharma, M. and Hinton, D.M. (1994). Purification and characterization of the SegA protein of bacteriophage T4, an endonuclease related to proteins encoded by group I introns. J. Bacteriol. *176*, 6439–6448.
13. Kadyrov, F.A., Shlyapnikov, M.G. and Kryukov, V.M. (1997). A phage T4 site-specific endonuclease, SegE, is responsible for a non-reciprocal genetic exchange between T-even-related phages. FEBS Lett. *415*, 75–80.
14. Jurica, M.S. and Stoddard, B.L. (1999). Homing endonucleases: structure, function, and evolution. Cell. Mol. Life Sci. *55*, 1304–1326.
15. Kutter, E., Gachechiladze, K., Pogalzov, A., Marusich, E., Sheider, M., Aronsson, P., Napuli, A., Porter, D. and Anzhinov, V.M. (1996). Evolution of T4 related phages. Virus Genes *11*, 285–297.
16. George, J.W., Stohr, B.A., Tomso, D.J. and Kreuzer, K.N. (2001). The tight linkage between DNA replication and double-strand break repair in bacteriophage T4. Proc. Natl. Acad. Sci. U.S.A. *98*, 8290–8297.
17. Tetart, F., Desplats, C., Kutateladze, M., Monod, C., Ackermann, H.W. and Krisch, H.M. (2001). Phylogeny of the major head and tail genes of the wide-ranging T4-type bacteriophages. J. Bacteriol. *183*, 358–366.
18. Monod, C., Repoila, F., Kutateladze, M., Tetart, F. and Krisch, H.M. (1997). The genome of the pseudo T-even bacteriophages, a diverse group that resembles T4. J. Mol. Biol. *267*, 237–249.