

Transposable Element Loads in a Bacterial Symbiont of Weevils Are Extremely Variable^{∇†‡}

Kevin M. Dougherty and Gordon R. Plague*

Louis Calder Center—Biological Field Station, Department of Biological Sciences, Fordham University, P.O. Box 887, Armonk, New York 10504

Received 9 May 2008/Accepted 21 October 2008

Not only are transposable elements profuse in the bacterial endosymbiont of maize weevils, but we found that their quantities also vary ~10-fold among individual weevils. Because multicopy elements can facilitate homologous recombination, this insertion sequence (IS) load variability suggests that these essentially asexual bacteria may exhibit substantial intraspecific genomic variation.

Insertion sequences (ISs) are transposable genetic elements in bacterial genomes. ISs only code for enzymes responsible for their own mobility and are therefore generally considered genomic parasites (5, 23). Although ISs are common among bacteria, they are usually rare within the genomes of free-living species (36, 37), presumably because natural selection efficiently purges high-IS-load genotypes from their large populations (21). However, transposable elements often proliferate after bacteria transition from a free-living to an intracellular symbiotic lifestyle (2, 21). This is probably due to relaxed natural selection resulting from two unique aspects of intracellularity (21). First, the environment within eukaryotic cells is relatively stable and nutrient rich, so intracellular bacteria do not have to synthesize many of their own required metabolites (e.g., nucleotides and amino acids). Therefore, many biosynthetic genes may be superfluous and therefore selectively neutral territory for IS insertion (22). Second, intracellular bacteria generally experience a population bottleneck whenever they colonize a new host (7), so their effective population sizes are small relative to free-living bacteria (19). Because genetic drift is more pervasive in small than in large populations, natural selection is less effective at purging slightly deleterious genotypes from intracellular bacterial populations (20). Consequently, IS elements may be able to expand in small intracellular populations simply because natural selection cannot efficiently purge these genotypes.

IS elements represent an integral source of genomic instability in bacteria (28). Not only can they move within genomes, but multicopy elements can also be loci for homologous recombination (24, 33, 40). Consequently, genomic fluidity should positively correlate with IS element quantity within a genome (11). Such fluidity is generally detrimental, although transposable elements occasionally generate beneficial mutations (12, 32). Therefore, profuse ISs may be a vital source of

genetic variation for intracellular bacteria that rarely, if ever, recombine with other bacteria (11, 35, 39). Despite their importance, very little is known about the level of IS-mediated genomic variation within and among populations of intracellular bacteria.

The nutritive symbiont within *Sitophilus zeamais* weevils (called *S. zeamais* primary endosymbiont, or SZPE) exhibits the most extreme case of IS proliferation of any known bacterium (27 and see reference 6). Specifically, SZPE harbors about 10× more IS elements than any other known bacterium, with >5,000 IS256 and ~60 IS903 copies per chromosome (27). As a first step to understanding how this extreme IS proliferation may affect intraspecific genomic variability, we quantified the relative IS256 and IS903 loads within natural SZPE populations.

We collected *S. zeamais* larvae from infested corn from four locations across the United States in September and October 2006: LaPorte County, IN; Riley County, KS; Platte County, NE; and Lebanon County, PA. We isolated SZPE from weevil larvae because the bacteriomes are relatively easy to identify and dissect—more so than in adults—thus allowing us to acquire relatively pure SZPE DNA. However, no dichotomous key exists to identify *Sitophilus* larvae, and *Sitophilus oryzae* weevils (which harbor their own unique bacteriome-associated endosymbiont) can also infest stored corn (34). Therefore, we used two strategies to confirm the identity of our collected weevils. First, we identified a total of 50 emergent adults from the field-collected corn of each of the four populations (10). All of these were *S. zeamais*. Therefore, if any *S. oryzae* weevils were coinfecting our sampled corn, they were relatively rare. Second, we used diagnostic PCR primers to amplify a portion of the nuclear ribosomal DNA which discriminates between *S. oryzae* and *S. zeamais* in Taiwan (25). As expected, these primers produced positive PCR products in all of the bacteriome DNA templates used in this study, while the *S. oryzae* negative control template did not amplify. Therefore, we are confident that all DNA templates used in this study were from SZPE. Also, we used diagnostic PCRs to confirm that *Wolbachia* was not coinfecting the bacteriomes of any analyzed larvae (41). Further details are given in the supplemental material, and PCR primers are listed in Table S1 in the supplemental material.

* Corresponding author. Mailing address: Louis Calder Center, Fordham University, P.O. Box 887, Armonk, NY 10504. Phone: (914) 273-3078, ext. 20. Fax: (914) 273-2167. E-mail: plague@fordham.edu.

‡ Contribution no. 239 of the Louis Calder Center—Biological Field Station, Fordham University.

† Supplemental material for this article may be found at <http://aem.asm.org/>.

[∇] Published ahead of print on 24 October 2008.

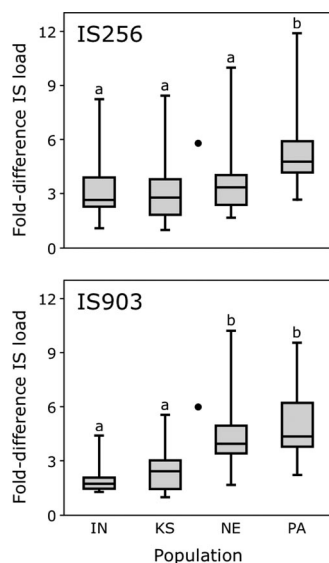


FIG. 1. Estimates of the relative IS256 and IS903 loads per genome for SZPE isolates across four populations. (All values are relative to the isolate with the lowest quantity of each IS.) The boxes represent the median and the 25th and 75th percentiles for each population, the bars represent the range, and the dots represent the relative estimates for the pooled SZPE template used by Plague et al. (27). Similar letters above the bars denote populations that are not significantly different ($P < 0.05$, Scheffe's test).

When transposing, IS256 forms an extrachromosomal circular molecule before inserting into a new location (18, 27, 29). However, we were interested in estimating strictly the chromosomal IS loads, so we extracted total genomic DNA from single bacteriomes using a modified cetyltrimethylammonium bromide method (38) that enriches for chromosomal DNA (30). We then estimated the relative abundances of IS256 and IS903 elements within these SZPE DNA isolates using real-time quantitative PCR on an Opticon (MJ Research, Waltham, MA). We used primers IS256F1 and IS256R1 (1,262-bp amplicon) and IS903F1 and IS903R1 (84 bp) to estimate the relative quantity of each IS element in comparison to the single-copy chromosomal gene *murA*, determined using primers *murAF4* and *murAR3* (131 bp). Further details are given in the supplemental material, and PCR primers are listed in Table S1 in the supplemental material.

We found that IS loads are extremely variable among individual weevils. Specifically, IS256 and IS903 both exhibit >10-fold copy number variation among field-collected weevils (Fig. 1). Because we used relative quantitative PCR (26) to estimate IS load variability, we do not know the absolute IS256 and IS903 quantities within each weevil. However, our analysis also included the pooled SZPE template from the *S. zeamais* laboratory culture used by Plague et al. (27), which allows us to roughly estimate absolute IS quantities based on this template's absolute estimates. The relative IS256 and IS903 estimates of this pooled template fall within the observed range for our field-collected templates (Fig. 1) and suggest an absolute range of ~1,000 to 10,000 IS256 copies and ~10 to 100 IS903 copies among the field-collected weevils.

This study provides a unique glimpse at IS load variability among natural bacterial isolates. Because few species have had

more than one isolate's genome completely sequenced (see, e.g., references 3, 4, and 17), comparative genomics currently offers limited insight into intraspecific variability. To our knowledge, the only other extensive analysis of intraspecific IS load variability involved DNA hybridizations with the 71 natural isolates in the *Escherichia coli* Reference Collection (ECOR) (8, 31). Like in SZPE, most IS elements exhibit ~10-fold variation among ECOR isolates (not including strains that contain zero copies of an element). However, unlike SZPE, most ECOR isolates harbor fewer than five copies of any particular element, so SZPE's observed ~10-fold intraspecific variation spans a much greater quantitative range. Because IS elements can greatly enhance genomic fluidity, particularly when they are abundant (24, 33, 40), this remarkable IS load variability among SZPE isolates may reflect substantial intraspecific variability in genome content and therefore metabolic capability of this nutritional symbiont. Furthermore, the extensive IS copy variation within weevil populations (Fig. 1) suggests that IS loads may diverge quickly among weevil lineages. Relatively rapid transposition and excision rates (36), coupled with serially occurring population bottlenecks each time SZPE cells are provisioned to weevil eggs (1, 9), could potentially lead to IS loads changing within a few weevil generations, and possibly even within the lifetime of a weevil, drifting higher in some lineages and lower in others.

We also found that IS loads differ significantly among weevil populations (Fig. 1). Specifically, the PA population harbors significantly more IS256 copies than the other populations ($P < 0.05$), and the NE and PA populations harbor significantly more IS903 copies than the IN and KS populations (Fig. 1). This may not be surprising given the potential lack of gene flow among geographically isolated weevil populations (e.g., 14, 15, 16) and thus the possibility for differential genetic drift. However, unique selective pressures among populations could also play a role. For example, weevils infesting grain with elevated nutrient concentrations (e.g., high-lysine corn [13]) may be less dependent on their endosymbionts for certain metabolites, thereby potentially increasing the number of selectively neutral SZPE genes available for IS insertion. Regardless of whether drift or selection is a stronger force shaping this IS load polymorphism, IS256 and IS903 are both probably important sources of novel genetic variation within and among essentially asexual SZPE populations.

We thank Paul Flinn, Robert Hooley III, Bill Klug, Linda Mason, Tom Phillips, and Vern Schafer for advice and assistance in weevil collection and identification and Jacob Russell for providing the *Wolbachia* positive control. We also thank Tom Daniels and Amy Tuininga for comments on a previous version of the manuscript.

This work was partially supported by a faculty research grant from Fordham University to G.R.P.

REFERENCES

1. Anselme, C., A. Vallier, S. Balmant, M.-O. Fauvarque, and A. Heddi. 2006. Host *PGRP* gene expression and bacterial release in endosymbiosis of the weevil *Sitophilus zeamais*. *Appl. Environ. Microbiol.* **72**:6766–6772.
2. Bordenstein, S. R., and W. S. Reznikoff. 2005. Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* **3**:688–699.
3. Chain, P. S. G., P. Hu, S. A. Malfatti, L. Radnedge, F. Larimer, L. M. Vergez, P. Worsham, M. C. Chu, and G. L. Andersen. 2006. Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: evidence of gene reduction in an emerging pathogen. *J. Bacteriol.* **188**:4453–4463.
4. Deng, W., S.-R. Liou, G. Plunkett III, G. F. Mayhew, D. J. Rose, V. Burland, V. Kodoyianni, D. C. Schwartz, and F. R. Blattner. 2003. Comparative

- genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. *J. Bacteriol.* **185**:2330–2337.
5. Doolittle, W. F., and C. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**:601–603.
 6. Gil, R., E. Belda, M. J. Gosalbes, L. Delaye, A. Vallier, C. Vincent-Monégat, A. Heddi, F. J. Silva, A. Moya, and A. Latorre. 2008. Massive presence of insertion sequences in the genome of SOPE, the primary endosymbiont of the rice weevil *Sitophilus oryzae*. *Int. Microbiol.* **11**:41–48.
 7. Hacker, J., U. Hentschel, and U. Dobrindt. 2003. Prokaryotic chromosomes and disease. *Science* **301**:790–793.
 8. Hall, B. G., L. L. Parker, P. W. Betts, R. F. DuBose, S. A. Sawyer, and D. L. Hartl. 1989. IS103, a new insertion element in *Escherichia coli*: characterization and distribution in natural populations. *Genetics* **121**:423–431.
 9. Heddi, A. 2003. Endosymbiosis in the weevil of the genus *Sitophilus*: genetic, physiological, and molecular interactions among associated genomes, p. 67–82. *In* K. Bourtzis and T. A. Miller (ed.), *Insect symbiosis*. CRC Press, Boca Raton, FL.
 10. Hidayat, P., T. W. Phillips, and R. H. French-Constant. 1996. Molecular and morphological characters discriminate *Sitophilus oryzae* and *S. zeamais* (Coleoptera: Curculionidae) and confirm reproductive isolation. *Ann. Entomol. Soc. Am.* **89**:645–652.
 11. Hughes, D. 2000. Evaluating genome dynamics: the constraints on rearrangements within bacterial genomes. *Genome Biol.* **1**:reviews0006.1–0006.8.
 12. Kidwell, M. G., and D. Lisch. 1997. Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA* **94**:7704–7711.
 13. Ladely, S. R., R. A. Stock, T. J. Klopfenstein, and M. H. Sindt. 1995. High-lysine corn as a source of protein and energy for finishing calves. *J. Anim. Sci.* **73**:228–235.
 14. Laffin, R. D., L. M. Dossdall, and F. A. H. Sperling. 2005. Population structure and phylogenetic relationships of *Ceutorhynchus neglectus* (Coleoptera: Curculionidae). *Can. Entomol.* **137**:672–684.
 15. Laffin, R. D., L. M. Dossdall, and F. A. H. Sperling. 2005. Population structure of the cabbage seedpod weevil, *Ceutorhynchus obstrictus* (Marsham) (Coleoptera: Curculionidae): origins of North American introductions. *Environ. Entomol.* **34**:504–510.
 16. Laffin, R. D., D. W. Langor, and F. A. H. Sperling. 2004. Population structure and gene flow in the white pine weevil, *Pissodes strobi* (Coleoptera: Curculionidae). *Ann. Entomol. Soc. Am.* **97**:949–956.
 17. Lanie, J. A., W.-L. Ng, K. M. Kazmierczak, T. M. Andrzejewski, T. M. Davidsen, K. J. Wayne, H. Tettelin, J. I. Glass, and M. E. Winkler. 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. *J. Bacteriol.* **189**:38–51.
 18. Loessner, I., K. Dietrich, D. Dittrich, J. Hacker, and W. Ziebuhr. 2002. Transposase-dependent formation of circular IS256 derivatives in *Staphylococcus epidermidis* and *Staphylococcus aureus*. *J. Bacteriol.* **184**:4709–4714.
 19. Mira, A., and N. A. Moran. 2002. Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. *Microb. Ecol.* **44**:137–143.
 20. Moran, N. 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA* **93**:2873–2878.
 21. Moran, N. A., and G. R. Plague. 2004. Genomic changes following host restriction in bacteria. *Curr. Opin. Genet. Dev.* **14**:627–633.
 22. Moran, N. A., and J. J. Wernegreen. 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol. Evol.* **15**:321–326.
 23. Orgel, L. E., and F. H. C. Crick. 1980. Selfish DNA: the ultimate parasite. *Nature* **284**:604–607.
 24. Parkhill, J., M. Sebaihia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, M. T. G. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M. Cerdeno-Tarraga, L. Temple, K. James, B. Harris, M. A. Quail, M. Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T. Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A. Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule, H. Norberczak, S. O'Neill, D. Ormond, C. Price, E. Rabinowitsch, S. Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G. Barrell, and D. J. Maskell. 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis*, and *Bordetella bronchiseptica*. *Nat. Genet.* **35**:32–40.
 25. Peng, W. K., H. C. Lin, C. N. Chen, and C. H. Wang. 2003. DNA identification of two laboratory colonies of the weevils, *Sitophilus oryzae* (L.) and *S. zeamais* Motschulsky (Coleoptera: Curculionidae) in Taiwan. *J. Stored Prod. Res.* **39**:225–235.
 26. Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**:2002–2007.
 27. Plague, G. R., H. E. Dunbar, P. L. Tran, and N. A. Moran. 2008. Extensive proliferation of transposable elements in heritable bacterial symbionts. *J. Bacteriol.* **190**:777–779.
 28. Posfai, G., G. Plunkett III, T. Feher, D. Frisch, G. M. Keil, K. Umenhoffer, V. Kolisnychenko, B. Stahl, S. S. Sharma, M. de Arruda, S. W. Harcum, and F. R. Blattner. 2006. Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**:1044–1046.
 29. Prudhomme, M., C. Turlan, J.-P. Claverys, and M. Chandler. 2002. Diversity of Tn4001 transposition products: the flanking IS256 elements can form tandem dimers and IS circles. *J. Bacteriol.* **184**:433–443.
 30. Sambrook, J., and D. W. Russell. 2001. *Molecular cloning: a laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
 31. Sawyer, S. A., D. E. Dykhuizen, R. F. DuBose, L. Green, T. Mutangadura-Mhlanga, D. F. Wolczyk, and D. L. Hartl. 1987. Distribution and abundance of insertion sequences among natural isolates of *Escherichia coli*. *Genetics* **115**:51–63.
 32. Schneider, D., and R. E. Lenski. 2004. Dynamics of insertion sequence elements during experimental evolution of bacteria. *Res. Microbiol.* **155**:319–327.
 33. Song, Y., Z. Tong, J. Wang, L. Wang, Z. Guo, Y. Han, J. Zhang, J. Pei, D. Zhou, H. Qin, X. Pang, Y. Han, J. Zhai, M. Li, B. Cui, Z. Qi, L. Jin, R. Dai, F. Chen, S. Li, C. Ye, Z. Du, W. Lin, J. Wang, J. Yu, H. Yang, J. Wang, P. Huang, and R. Yang. 2004. Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to humans. *DNA Res.* **11**:179–197.
 34. Throne, J. E., and L. D. Cline. 1991. Seasonal abundance of maize and rice weevils (Coleoptera: Curculionidae) in South Carolina. *J. Agric. Entomol.* **8**:93–100.
 35. Tillier, E. R. M., and R. A. Collins. 2000. Genome rearrangement by replication-directed translocation. *Nat. Genet.* **26**:195–197.
 36. Wagner, A. 2006. Periodic extinctions of transposable elements in bacterial lineages: evidence from intragenomic variation in multiple genomes. *Mol. Biol. Evol.* **23**:723–733.
 37. Wagner, A., C. Lewis, and M. Bischel. 2007. A survey of bacterial insertion sequences using IScan. *Nucleic Acids Res.* **35**:5284–5293.
 38. Wilson, K. 1997. Preparation of genomic DNA from bacteria, p. 2.4.1–2.4.5. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), *Current protocols in molecular biology*. John Wiley and Sons, New York, NY.
 39. Wren, B. W. 2003. The yersiniae—a model genus to study the rapid evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **1**:55–64.
 40. Yang, F., J. Yang, X. Zhang, L. Chen, Y. Jiang, Y. Yan, X. Tang, J. Wang, Z. Xiong, J. Dong, Y. Xue, Y. Zhu, X. Xu, L. Sun, S. Chen, H. Nie, J. Peng, J. Xu, Y. Wang, Z. Yuan, Y. Wen, Z. Yao, Y. Shen, B. Qiang, Y. Hou, J. Yu, and Q. Jin. 2005. Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. *Nucleic Acids Res.* **33**:6445–6458.
 41. Zhou, W., F. Rousset, and S. O'Neill. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. R. Soc. Lond. B Biol. Sci.* **265**:509–515.