Metaproteomics Reveals Abundant Transposase Expression in Mutualistic Endosymbionts

Manuel Kleiner,[a] Jacque C. Young,[b,c] Manesh Shah,[b] Nathan C. VerBerkmoes,[b]* Nicole Dubilier[a]

Symbiosis Group, Max Planck Institute for Marine Microbiology, Bremen, Germany[a]; Oak Ridge National Laboratory, Chemical Sciences Division, Oak Ridge, Tennessee, USA[b]; Graduate School for Genome Science and Technology, University of Tennessee, Knoxville, Tennessee, USA[c]

* Present address: Nathan C. VerBerkmoes, New England Biolabs, Inc., Ipswich, Massachusetts, USA.
M.K. and J.C.Y. contributed equally to this work.

ABSTRACT **Transposases, enzymes that catalyze the movement of mobile genetic elements, are the most abundant genes in nature. While many bacteria encode an abundance of transposases in their genomes, the current paradigm is that the expression of transposase genes is tightly regulated and generally low due to its severe mutagenic effects. In the current study, we detected the highest number of transposase proteins ever reported in bacteria, in symbionts of the gutless marine worm *Olavius algarvensis* with metaproteomics. At least 26 different transposases from 12 different families were detected, and genomic and proteomic analyses suggest that many of these are active. This high expression of transposases indicates that the mechanisms for their tight regulation have been disabled or no longer exist.**

IMPORTANCE **The expansion of transposable elements (TE) within the genomes of host-restricted symbionts and pathogens plays an important role in their emergence and evolution and might be a key mechanism for adaptation to the host environment. However, little is known so far about the underlying causes and evolutionary mechanisms of this TE expansion. The current model of genome evolution in host-restricted bacteria explains TE expansion within the confines of the paradigm that transposase expression is always low. However, recent work failed to verify this model. Based on our data, we hypothesize that increased transposase expression, which has not previously been described, may play a role in TE expansion, and could be one explanation for the sometimes very rapid emergence and evolution of new obligate symbionts and pathogens from facultative ones.**

Received 28 March 2013 **Accepted** 22 May 2013 **Published** 18 June 2013

Citation Kleiner M, Young JC, Shah M, VerBerkmoes NC, Dubilier N. 2013. Metaproteomics reveals abundant transposase expression in mutualistic endosymbionts. mBio 4(3): e00223-13. doi:10.1128/mBio.00223-13.

Invited Editor Colleen Cavanaugh, Harvard University **Editor** Mary Ann Moran, University of Georgia

Copyright © 2013 Kleiner et al. This is an open-access article distributed under the terms of the Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license, which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to Manuel Kleiner, mkleiner@mpi-bremen.de.

Transposases are enzymes that catalyze the movement of mobile genetic elements in and between genomes and are the most abundant and ubiquitous genes in nature (1). Most often, transposases are part of transposable elements (TEs), which encode only the transposase gene and some short flanking sequences necessary for transposition. These basic TEs are called insertion sequence (IS) elements. Classically, TEs are considered to be selfish genetic elements or parasitic DNA with no purpose other than reproducing themselves (2–4). However, in more recent years, it has become clear that TEs are not always parasites, but can also have beneficial effects that increase the fitness of their bacterial host (for reviews of the ongoing debate, see references 3–5). TEs (especially IS elements) are involved in gene deletions, gene duplications, genome rearrangements, gene regulation, and horizontal gene transfer (for a review, see references 6 and 7), all of which can have beneficial effects on the host population by generating genomic diversity and thus enabling adaptation to environmental changes (8–10). However, TEs can also be detrimental if they disrupt important functional genes. Therefore, transposase expression and, thus, transpositional activity are usually very low (11), because the mutagenic effects of transposases would drive their hosts into extinction, thereby also eradicating their own existence (12). Accordingly, a large variety of mechanisms for the tight regulation of transposase expression exists at both the transcriptional and translational levels (11).

TEs and, thus, transposase genes are particularly enriched in the genomes of some mutualistic symbionts (here called "symbionts") and some pathogens that have recently transitioned or are transitioning to an obligate, host-associated lifestyle (5, 13). In contrast, TEs are absent in the reduced genomes of most obligate, host-restricted bacteria that have been associated with their host over long evolutionary time periods (TEs absent in 65% of the genomes from symbionts classified as obligately intracellular in Newton and Bordenstein [13]). There are, however, some ancient obligate intracellular bacteria, such as *Wolbachia pipientis w*Mel and "*Candidatus* Amoebophilus asiaticus," that have high numbers of TEs in their genomes (14, 15) (TEs present in 35% of the genomes from symbionts classified as obligately intracellular in Newton and Bordenstein [13]).

It has been shown that TE expansion in bacteria in transition to an obligate, host-associated lifestyle plays a crucial role in the emergence and early evolution of pathogens (8, 16–18) and sym-

bionts (19, 20). Currently, there is much uncertainty about the factors that lead to high TE loads in host-restricted bacteria, and several hypotheses have been put forth (reviewed in reference 3). The two main hypotheses, which represent partially opposing views, are (i) the selective advantage hypothesis and (ii) the relaxed natural selection hypothesis. In (i), TE expansion is selected for because it is beneficial for symbionts and pathogens transitioning to an obligate lifestyle, for example, by providing enhanced genomic plasticity for faster adaptation to the host environment (8, 9, 21). In (ii), temporary TE expansion in the genomes of host-restricted bacteria is due to a reduced effectiveness of natural selection against deleterious transpositions (5). The relaxed natural selection, according to this second hypothesis, is caused by genetic drift due to small population sizes during transmission of symbionts from one host generation to the next, and the fact that when symbionts reside within a host, many of their genes become superfluous and can thus act as neutral integration sites for TEs.

These two hypotheses can also explain the evolutionary processes involved in TE expansion in ancient obligate intracellular bacteria that have a reduced genome, yet still have high TE numbers in their genomes. For these ancient symbionts, the "intracellular arena hypothesis" explains how TEs first reenter their genomes: host switching events bring these symbionts into contact with other bacteria during coinfection within the same host and thus enable the uptake of foreign TEs (13, 14). These newly acquired TEs then multiply in the genomes of the ancient symbionts either (i) because they confer a selective advantage or (ii) due to reduced effectiveness of natural selection in the intracellular environment (5, 22).

A recent study that tested hypothesis (ii) by subjecting *Escherichia coli* for 4,000 generations to simulated conditions of relaxed natural selection raised doubts as to whether relaxed natural selection alone can account for TE expansion (21), because no TE expansion occurred under the tested conditions. Although Plague et al. (21) noted that "there are several possible reasons why [their] experiment may not have adequately tested the [relaxed natural selection] hypothesis," they hypothesized that other factors, including increased transposase activity, might have enabled the massive TE expansion observed in host-restricted bacteria. Common to both hypotheses (i) and (ii) is that they explain TE expansion in host-restricted bacteria within the confines of the paradigm that transposase expression and, thus, transpositional activity are always low.

High numbers of transposase genes were recently described in endosymbionts of the gutless marine oligochaete *Olavius algarvensis* (23). *O. algarvensis* inhabits shallow water sediments in the Mediterranean and lacks both a digestive and an excretory system, relying instead for nutrition and waste recycling on a symbiotic community of two gammaproteobacterial sulfur oxidizers (γ1 and γ3 symbionts), two deltaproteobacterial sulfate reducers (δ1 and δ4 symbionts), and a spirochete (23, 24). The symbiotic bacteria cooccur in an extracellular space just below the worm's cuticle and above the host's epidermal cells where they are in direct contact with each other and have access to solutes $<$70 kDa from the environment that can easily diffuse through the worm's cuticle (25). Metagenomic analyses of *O. algarvensis* showed that the γ1 symbiont has a remarkably high percentage of transposases in its genome, at nearly 21% of all genes, followed by the γ3 symbiont with 7.5% and the δ1 symbiont with 2.3% (23, 26). Nothing is currently known about the factors that have led to such high TE numbers in some of the *O. algarvensis* symbionts. We assume that the association with the γ1 symbiont is ancient because these symbionts occur in many gutless oligochaete species from around the world, indicating that the common ancestor of all gutless oligochaetes may have already had the γ1 symbiont (25). However, the genome of the γ1 symbiont shows no signs of reduction (23). The associations with the γ3 and deltaproteobacterial symbionts may be more recent, as these symbionts are found in some but not all gutless oligochaetes (27); the genomes of these symbionts are also not reduced (23). The transmission mode of symbionts in gutless oligochaetes has been examined only in two host species from Bermuda and appears to be vertical (28). However, occasional horizontal uptake of bacteria from the environment during egg deposition in the surrounding sediment cannot be excluded. Thus, the *O. algarvensis* symbionts could acquire new TEs from both environmental sources and cooccurring symbionts.

The aim of this study was to examine how the high percentage of TEs in the *O. algarvensis* symbiont genomes affects their expression. Interestingly, the relationship between high TE loads and transposase expression and activity has, to our knowledge, so far not been explored. Our metaproteomic analyses revealed that the γ1 and δ1 symbionts express a surprisingly high number of their transposase proteins, many of which are intact and possibly active, and we hypothesize that abundant transposase expression plays a key role in TE expansion.

Methods. Worms were collected as described previously (24) and either frozen immediately or symbionts were enriched via isopycnic centrifugation by layering the worm homogenate on top of a HistoDenz (Sigma) multistep density gradient. Symbionts were separated from each other and from host tissue by 1 h of centrifugation at 4°C as described by Kleiner et al. (24).

High proteome coverage was achieved, i.e., 2,265 symbiont proteins were detected, via automated 24-h two-dimensional liquid chromatography followed by tandem mass spectrometry (MS/MS) with a hybrid linear ion trap-Orbitrap (Thermo Fischer Scientific) as described in detail by Kleiner et al. (24). Protein databases, peptide and protein identifications, and all MS/MS spectra are available from http://compbio.ornl.gov/olavius_algarvensis_symbiont_metaproteome/, and detailed data on expressed transposases are available from http://compbio.ornl.gov/olavius_algarvensis_symbiont_transposases/.

Transposase numbers and abundance in *O. algarvensis* symbionts. We detected the highest number of transposase proteins ever reported in bacteria in the γ1 symbiont of *O. algarvensis* (22 proteins) (Table 1). Additionally, we detected two transposase proteins in the δ1 symbiont and two transposase proteins that could not be unambiguously assigned to a given symbiont (Table 1). We did not obtain good proteome coverage for the γ3 symbiont due to its low abundance in *O. algarvensis* and were therefore unable to determine if any of its numerous transposase genes were expressed. To compare our data to other studies, we searched for reports of transposase expression at the protein level and found that the highest numbers were found in free-living bacteria and not in host-associated bacteria. The highest numbers of expressed transposases were seven in a proteome of the cultured bacterium *Deinococcus radiodurans* (29) and eight in a metaproteome of the uncultured *Leptospirillum* group II bacterium from acid mine drainage (30). To our knowledge, transposase expression has so far not been examined at the protein level in symbionts or pathogens with high TE numbers in their genomes.

TABLE 1 Overview of all expressed transposases grouped according to shared peptide matches

Accession no.[a]	Symbiont	Transposase group[b]	No. of transposases in each group	No. of transposases in each group required to explain all peptide matches	Transposase classification[c]	Intact[d]	Premature translation termination possible
2004221906	γ1	1	38	4	IS*630*	No	No
2004222763	γ1	1			IS*630*	No	No
2004222856	γ1	1			IS*630*	Yes	No
2004222872	γ1	1			IS*630*	No	No
2004223397	γ1	2	19	3	IS*110*	Yes	No
2004222027	γ1	2			IS*110*	No	No
2004222937	γ1	2			IS*110*	No	No
Symbiont _37746	Unknown[e]	3	7	1	Unclassified	Unknown[f]	Unknown
2004223511	γ1	4	1	1	Unclassified	Yes	Unknown
Symbiont _28062	Unknown[e]	5	1	1	IS*1634*	Yes	No
2004221868	γ1	6	37	7	IS*481*	Yes	No
2004221830	γ1	6			IS*3*	No	Yes
2004222310	γ1	6			IS*481*	Yes	No
2004222544	γ1	6			IS*481*	No	No
2004222874	γ1	6			IS*481*	No	No
2004223428	γ1	6			IS*5*	No	No
2004223468	γ1	6			IS*481*	No	No
2004207437	δ1	7	2	1	IS*4*	Yes	No
2004223208	γ1	8	1	1	IS*1595*	Yes	No
2004212411	δ1	9	3	1	IS*21*	Yes	Yes
2004222138	γ1	10	16	2	IS*1*	Yes	Yes
2004222867	γ1	10			IS*1*	Yes	Yes
2004222325	γ1	11	7	3	IS*5*	Yes	No
2004223125	γ1	11			IS*630*	No	No
2004223697	γ1	11			IS*630*	No	No
2004221963	γ1	12	3	1	Unclassified	Yes	Unknown
Total			135[g]	26			

[a] Accession numbers refer to the JGI IMG/M database (http://img.jgi.doe.gov/cgi-bin/m/main.cgi). Sequences on unassigned metagenome fragments can be found at http://compbio.ornl.gov/olavius_algarvensis_symbiont_metaproteome.
[b] See Table S4 in the supplemental material for details on grouping.
[c] IS element family, according to IS Finder. Cutoff amino acid identity of >30%.
[d] Similar length and identical Pfam domain structures compared to homologous transposases. Some of the transposases that were identified as not intact may actually be intact and only appear to be fragmented due to the incomplete nature of the symbiont metagenome.
[e] Identified with an unassigned symbiont metagenome fragment.
[f] Only very short contigs available.
[g] This number is bigger than the actual number, 134, because one transposase fell into two groups.

Transposases comprised up to 1.95% of the total γ1 symbiont protein (see Table S1 in the supplemental material) and up to 0.084% of the total δ1 symbiont protein (Table S2). The abundance of transposases (as estimated from label-free proteomic quantitation) in the γ1 symbiont was greater than that of some of its most abundant housekeeping proteins, such as the ATPase B subunit (1.15%), malate dehydrogenase (0.38%), and 6-phosphofructokinase (0.35%) (24). Within the context of natural microbial communities, highly abundant transposase protein expression has, so far, only been reported from a microbial biofilm in acid mine drainage; however, relative abundances were estimated at less than half of the amounts observed here (30).

Since many transposase genes are present in multiple, nearly identical copies in the symbiont metagenomes (23, 26), the 26 groups of transposase genes (Table 1) could be encoded by as many as 134 transposase gene sequences in the metagenome (see Table S3 in the supplemental material). Given that proteins encoded by almost-identical sequences with identical tryptic peptides cannot be distinguished from each other by mass spectrometry-based proteomic analyses, we were not able to identify which of the 134 transposase genes were expressed and if multiple copies of identical genes were expressed. Therefore, we assigned the transposases that were identified with similar sets of peptides to 12 different groups (Table 1; see also Table S4 in the supplemental material) and used this grouping to identify the minimal set of transposases that must be expressed to explain all transposase-related peptides. This nonredundant set consisted of the above-mentioned 26 transposases.

We classified the 26 nonredundant transposases into IS element families using BLASTp against the curated IS Finder database (31) (http://www-is.biotoul.fr/) and found that they belong to at least 10 different IS element families in the γ1 symbiont and 2 families in the δ1 symbiont (Table 1). This clearly shows that the expressed transposase genes originated from multiple unrelated IS elements.

Could abundant transposase expression be caused by stress? Previous studies have reported increased transposase expression in response to stressful conditions (29, 32). However, relative abundances were much lower than those detected in our study. To exclude the possibility that transposase expression in the *O. algarvensis* symbionts was caused by stressful conditions during the 1 hour long symbiont enrichment procedure, we did a qualitative comparison of their proteomes with those of symbionts that were frozen in whole worms immediately following removal from the

sediment. We also measured high levels of transposase expression in these immediately frozen symbionts (see Tables S1 and S3 in the supplemental material). Thus, we conclude that the observed transposase expression is not due to stressful sampling and enrichment conditions but rather reflects expression under natural environmental conditions.

Are the expressed transposases active? We inferred that some of the expressed transposases are active by excluding the two main reasons for potential inactivity: (i) transposase genes could be in the process of gene degradation so that their expression would lead to incomplete and potentially inactive transposases and (ii) the expression of some transposase genes is regulated through programmed translational frameshifting, which would lead to the translation of a truncated, nonfunctional version of the transposase protein (6, 11, 33).

First, we checked for indications that the expressed transposase genes were intact rather than in a state of gene degradation. We compared the gene sizes and protein sequences to those of closely related transposases in the IS Finder database and compared their protein domain structures with the domain structure of similar transposases using the "domain organization" feature available on Pfam (http://pfam.sanger.ac.uk/search). We found that around half of the expressed transposase genes were intact, whereas for the other half, we were not able to exclude the possibility that they were in some stage of degeneration (Table 1). Second, we checked the literature and found that only the IS families of four out of the 26 detected transposases are known to be regulated by programmed translational frameshifting (11) (Table 1). Therefore, we assume that the majority of the expressed transposases are translated to full-length proteins.

Additional evidence for the expression of full-length transposase proteins comes from our proteomic data. For some of the identified transposases, we detected peptides not only in the beginning part of the protein but also in the middle or at the end of the protein, a finding which indicates that the protein was translated from beginning to end (http://compbio.ornl.gov/olavius_algarvensis_symbiont_transposases/).

Conclusion. Our results show that the paradigm that transposase expression in bacteria must be tightly regulated and generally low to prevent the host population from going extinct does not always hold true. We present evidence at the protein level that transposases are abundantly expressed in beneficial symbionts with high TE numbers in their genomes. This high expression of transposases indicates that the mechanisms typically found in bacteria for their tight regulation have been disabled or no longer exist, for example, possibly through mutations in proteins that are involved in transposase regulation (20, 34, 35). The fixation of such mutations may be enabled by the relaxed purifying selection suggested by Moran et al. for symbionts and pathogens that recently transitioned to an obligate, host-associated lifestyle (20).

Currently, it is not possible to determine if abundant transposase expression is present in other symbionts and pathogens with high TE numbers, because no comparable proteomic datasets exist for these bacteria. However, many recent studies have shown high transcription of transposase genes in symbionts and pathogens (15, 36–40), which may indicate that abundant transposase expression is common in these bacteria. However, the presence of these transcripts does not represent conclusive evidence for transposase expression because it is possible that they are not translated into proteins due to regulatory mechanisms (11, 15, 37).

Based on the high numbers of transposase proteins in the *O. algarvensis* γ1 symbiont and abundant transposase transcription in other symbionts and pathogens, we speculate that high transposase expression is one of the key factors in TE expansion in host-restricted bacteria. As discussed above, an experiment that simulated conditions of relaxed natural selection failed to cause TE expansion after 4,000 generations in *E. coli*. If, as we speculate, high transposase expression is the major catalyst for TE expansion, the mutations that lead to increased transposase expression might have simply not occurred yet in the Plague et al. (21) experiment. Additional studies that investigate transposase expression in pathogens and symbionts are needed because high transposase expression may be an important factor in the sometimes very rapid emergence and evolution of new obligate symbionts and pathogens from facultative ones.

The remarkably high expression of transposases in the *O. algarvensis* γ1 symbiont raises the question of how the symbiont can function over evolutionary time periods given the likelihood that high transposase expression leads to an increase in deleterious mutations. In other organisms that are confronted with frequent transpositions, genome rearrangements and disruptions, it has been suggested that polyploidy buffers against the detrimental effects of the factors that lead to these genome disruptions such as transposable elements, introns, heat and ionizing radiation (41–44). Polyploidy has recently been shown for several symbionts, including sulfur-oxidizing symbionts of clams (45), and it is possible that it also plays a role in the *O. algarvensis* γ1 symbiont.

To gain a better understanding of how the *O. algarvensis* γ1 symbiont deals with high transposase expression, it would be crucial to know specifically how high the transpositional activity of the abundant transposases is. Does it lead to transpositions in every second symbiont or in every millionth? Do the symbionts have multiple genome copies that buffer against the deleterious effects of transposase activity? To answer these questions, we are currently sequencing the genomes of single symbiont cells to examine how common transposition events are within the same individual host, within the host population, and between host populations.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00223-13/-/DCSupplemental.

Table S1, XLS file, 0.1 MB.
Table S2, XLS file, 0.1 MB.
Table S3, XLS file, 0.1 MB.
Table S4, XLS file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Stephan Schmitz-Esser for stimulating scientific discussions and helpful comments on the manuscript. We also thank the editor and two anonymous reviewers for detailed and thoughtful feedback that helped improve the manuscript significantly.

M. Kleiner was supported by a Studienstiftung des Deutschen Volkes scholarship. Funding for this study was provided by the Max Planck Society and by the Laboratory Directed Research and Development support at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy.

REFERENCES

1. **Aziz RK, Breitbart M, Edwards RA.** 2010. Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res. **38:** 4207–4217.
2. **Doolittle WF, Sapienza C.** 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature **284:**601–603.

3. **Touchon M, Rocha EP.** 2007. Causes of insertion sequences abundance in prokaryotic genomes. Mol. Biol. Evol. **24:**969–981.
4. **Werren JH.** 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc. Natl. Acad. Sci. U. S. A. **108**(Suppl 2): 10863–10870.
5. **Moran NA, Plague GR.** 2004. Genomic changes following host restriction in bacteria. Curr. Opin. Genet. Dev. **14:**627–633.
6. **Mahillon J, Chandler M.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725–774.
7. **Frost LS, Leplae R, Summers AO, Toussaint A.** 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. **3:**722–732.
8. **Leavis HL, Willems RJ, van Wamel WJ, Schuren FH, Caspers MP, Bonten MJ.** 2007. Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E.* faecium. PLoS Pathog. **3:**e7. http://dx.doi.org/10.1371/journal.ppat.0030007.
9. **Edwards RJ, Brookfield JF.** 2003. Transiently beneficial insertions could maintain mobile DNA sequences in variable environments. Mol. Biol. Evol. **20:**30–37.
10. **Chao L, McBroom SM.** 1985. Evolution of transposable elements: an IS10 insertion increases fitness in *Escherichia coli*. Mol. Biol. Evol. **2:**359–369.
11. **Nagy Z, Chandler M.** 2004. Regulation of transposition in bacteria. Res. Microbiol. **155:**387–398.
12. **Doolittle WF, Kirkwood TB, Dempster MA.** 1984. Selfish DNAs with self-restraint. Nature **307:**501–502.
13. **Newton IL, Bordenstein SR.** 2011. Correlations between bacterial ecology and mobile DNA. Curr. Microbiol. **62:**198–208.
14. **Bordenstein SR, Reznikoff WS.** 2005. Mobile DNA in obligate intracellular bacteria. Nat. Rev. Microbiol. **3:**688–699.
15. **Schmitz-Esser S, Penz T, Spang A, Horn M.** 2011. A bacterial genome in transition—an exceptional enrichment of IS elements but lack of evidence for recent transposition in the symbiont *Amoebophilus asiaticus*. BMC Evol. Biol. **11:**270.
16. **Song H, Hwang J, Yi H, Ulrich RL, Yu Y, Nierman WC, Kim HS.** 2010. The early stage of bacterial genome-reductive evolution in the host. PLoS Pathog. **6:**e1000922. http://dx.doi.org/10.1371/journal.ppat.1000922.
17. **Yang F, Yang J, Zhang X, Chen L, Jiang Y, Yan Y, Tang X, Wang J, Xiong Z, Dong J, Xue Y, Zhu Y, Xu X, Sun L, Chen S, Nie H, Peng J, Xu J, Wang Y, Yuan Z, Wen Y, Yao Z, Shen Y, Qiang B, Hou Y, Yu J, Jin Q.** 2005. Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. Nucleic Acids Res. **33:**6445–6458.
18. **Ochiai H, Inoue Y, Takeya M, Sasaki A, Kaku H.** 2005. Genome sequence of *Xanthomonas oryzae* pv. *oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. Jpn. Agric. Res. Q. **39:**275–287.
19. **Thomas T, Rusch D, DeMaere MZ, Yung PY, Lewis M, Halpern A, Heidelberg KB, Egan S, Steinberg PD, Kjelleberg S.** 2010. Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. ISME J. **4:**1557–1567.
20. **Plague GR, Dunbar HE, Tran PL, Moran NA.** 2008. Extensive proliferation of transposable elements in heritable bacterial symbionts. J. Bacteriol. **190:**777–779.
21. **Plague GR, Dougherty KM, Boodram KS, Boustani SE, Cao H, Manning SR, McNally CC.** 2011. Relaxed natural selection alone does not permit transposable element expansion within 4,000 generations in *Escherichia coli*. Genetica **139:**895–902.
22. **Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, Brownlie JC, McGraw EA, Martin W, Esser C, Ahmadinejad N, Wiegand C, Madupu R, Beanan MJ, Brinkac LM, Daugherty SC, Durkin AS, Kolonay JF, Nelson WC, Mohamoud Y, Lee P, Berry K, Young MB, Utterback T, Weidman J, Nierman WC, Paulsen IT, Nelson KE, Tettelin H, O'Neill SL, Eisen JA.** 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. PLoS Biol. **2:**e69. http://dx.doi.org/10.1371/journal.pbio.0020069.
23. **Woyke T, Teeling H, Ivanova NN, Huntemann M, Richter M, Gloeckner FO, Boffelli D, Anderson IJ, Barry KW, Shapiro HJ, Szeto E, Kyrpides NC, Mussmann M, Amann R, Bergin C, Ruehland C, Rubin EM, Dubilier N.** 2006. Symbiosis insights through metagenomic analysis of a microbial consortium. Nature **443:**950–955.
24. **Kleiner M, Wentrup C, Lott C, Teeling H, Wetzel S, Young J, Chang YJ, Shah M, Zarzicky J, Fuchs G, Market S, Hempal K, Hecker M, Dubilier N.** 2012. Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. Proc. Natl. Acad. Sci. U. S. A. **109:**E1173–E1182.
25. **Dubilier N, Blazejak A, Rühland C.** 2006. Symbiosis between bacteria and gutless marine oligochaetes, p 251–275. *In* Overmann J (ed), Molecular basis of symbiosis. Springer Verlag, Berlin, Germany.
26. **Kleiner M, Woyke T, Ruehland C, Dubilier N.** 2011. The *Olavius algarvensis* metagenome revisited: lessons learned from the analysis of the low diversity microbial consortium of a gutless marine worm, p 321–334. *In* Bruijn FJ (ed) Handbook of molecular microbial ecology II: metagenomics in different habitats. John Wiley & Sons, Hoboken, NJ.
27. **Ruehland C, Blazejak A, Lott C, Loy A, Erséus C, Dubilier N.** 2008. Multiple bacterial symbionts in two species of co-occurring gutless oligochaete worms from Mediterranean Sea grass sediments. Environ. Microbiol. **10:**3404–3416.
28. **Giere O, Langheld C.** 1987. Structural organisation, transfer and biological fate of endosymbiotic bacteria in gutless oligochaetes. Mar. Biol. **93:** 641–650.
29. **Schmid AK, Lipton MS, Mottaz H, Monroe ME, Smith RD, Lidstrom ME.** 2005. Global whole-cell FTICR mass spectrometric proteomics analysis of the heat shock response in the radioresistant bacterium *Deinococcus radiodurans*. J. Proteome Res. **4:**709–718.
30. **Ram RJ, VerBerkmoes NC, Thelen MP, Tyson GW, Baker BJ, Blake RC, Shah M, Hettich RL, Banfield JF.** 2005. Community proteomics of a natural microbial biofilm. Science **308:**1915–1920.
31. **Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M.** 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. **34:**D32–D36.
32. **Goodchild A, Raftery M, Saunders NF, Guilhaus M, Cavicchioli R.** 2004. Biology of the cold adapted archaeon, *Methanococcoides burtonii* determined by proteomics using liquid chromatography-tandem mass spectrometry. J. Proteome Res. **3:**1164–1176.
33. **Chandler M, Fayet O.** 1993. Translational frameshifting in the control of transposition in bacteria. Mol. Microbiol. **7:**497–503.
34. **Valle J, Vergara-Irigaray M, Merino N, Penadés JR, Lasa I.** 2007. σ^B regulates IS*256*-mediated *Staphylococcus aureus* biofilm phenotypic variation. J. Bacteriol. **189:**2886–2896..
35. **Roberts D, Hoopes BC, McClure WR, Kleckner N.** 1985. IS10 transposition is regulated by DNA adenine methylation. Cell **43:**117–130.
36. **Williams DL, Slayden RA, Amin A, Martinez AN, Pittman TL, Mira A, Mitra A, Nagaraja V, Morrison NE, Moraes M, Gillis TP.** 2009. Implications of high level pseudogene transcription in *Mycobacterium leprae*. BMC Genomics **10:**397.
37. **Bickhart DM, Benson DR.** 2011. Transcriptomes of *Frankia* sp. strain CcI3 in growth transitions. BMC Microbiol. **11:**192.
38. **Mitchell HL, Dashper SG, Catmull DV, Paolini RA, Cleal SM, Slakeski N, Tan KH, Reynolds EC.** 2010. Treponema denticola biofilm-induced expression of a bacteriophage, toxin-antitoxin systems and transposases. Microbiology **156:**774–788.
39. **Egas C, Pinheiro M, Gomes P, Barroso C, Bettencourt R.** 2012. The transcriptome of *Bathymodiolus azoricus* gill reveals expression of genes from endosymbionts and free-living deep-sea bacteria. Mar. Drugs **10:** 1765–1783.
40. **Radax R, Rattei T, Lanzen A, Bayer C, Rapp HT, Urich T, Schleper C.** 2012. Metatranscriptomics of the marine sponge *Geodia barretti*: tackling phylogeny and function of its microbial community. Environ. Microbiol. **14:**1308–1324.
41. **Soltis DE, Soltis PS.** 1999. Polyploidy: recurrent formation and genome evolution. Trends Ecol. Evol. **14:**348–352.
42. **Salman V, Amann R, Shub DA, Schulz-Vogt HN.** 2012. Multiple self-splicing introns in the 16S rRNA genes of giant sulfur bacteria. Proc. Natl. Acad. Sci. U. S. A. **109:**4203–4208.
43. **Slade D, Lindner AB, Paul G, Radman M.** 2009. Recombination and replication in DNA repair of heavily irradiated *Deinococcus radiodurans*. Cell **136:**1044–1055.
44. **Ohtani N, Tomita M, Itaya M.** 2010. An extreme thermophile, *Thermus thermophilus*, is a polyploid bacterium. J. Bacteriol. **192:**5499–5505.
45. **Caro A, Gros O, Got P, De Wit R, Troussellier M.** 2007. Characterization of the population of the sulfur-oxidizing symbiont of *Codakia orbicularis* (Bivalvia, Lucinidae) by single-cell analyses. Appl. Environ. Microbiol. **73:**2101–2109.