Extensive Genome Rearrangements and Multiple Horizontal Gene Transfers in a Population of *Pyrococcus* Isolates from Vulcano Island, Italy^{∇†}

James R. White,¹ Patricia Escobar-Paramo,¹ Emmanuel F. Mongodin,^{2‡}
Karen E. Nelson,² and Jocelyne DiRuggiero^{1*}

University of Maryland, College Park, Maryland,¹ and J. Craig Venter Institute, Rockville, Maryland²

Received 7 May 2008/Accepted 14 August 2008

The extent of chromosome rearrangements in *Pyrococcus* isolates from marine hydrothermal vents in Vulcano Island, Italy, was evaluated by high-throughput genomic methods. The results illustrate the dynamic nature of the genomes of the genus *Pyrococcus* and raise the possibility of a connection between rapidly changing environmental conditions and adaptive genomic properties.

Environmental genomic surveys have increasingly demonstrated the remarkable diversity of natural microbial communities and have had a significant impact on our understanding of microbial ecology and evolution (10, 34). For example, data sets of environmental microbial sequences and fine-scale population studies revealed the high level of genetic diversity between strains or ribotypes in natural populations, as seen at the levels of sequence divergence, gene content, and genome organization (1, 27, 31, 38). Genomic recombination and the presence of mobile genetic elements resulting in insertion, deletion, and rearrangements have been reported as important mechanisms for genotypic diversity in natural populations of both *Bacteria* and *Archaea* (2, 12, 37, 40). In a previous study, we showed that *Pyrococcus* sp. strains isolated from the same shallow hydrothermal vent system, with <1% divergence at the level of their 16S rRNA sequences, showed extensive gene rearrangements and insertions/deletions associated with transposition events (12). These results raise the question of the role of nucleotide substitutions and genome rearrangements as driving forces in generating genetic diversity and in altering physiological properties and, therefore, in microbial adaptation to available resources.

We used comparative genomic hybridization (CGH) with *Pyrococcus furiosus* and seven isolates from Vulcano Island (Table 1) (12) to determine the extent of chromosome rearrangements within the *Pyrococcus* genus. Genomic DNAs from *Pyrococcus* sp. environmental isolates were hybridized to a microarray consisting of 2,060 unique PCR products, representing the genome of *P. furiosus* DSM 3638 (GenBank accession no. AE009950) as described before (23). Statistical analysis of the data was performed using GACK analysis software (17a). Genomic DNA from all environmental *Pyrococcus* isolates hybridized against most of the open reading frames

(ORFs) present on the *P. furiosus* whole-genome array, reflecting a high level of sequence identity and genome conservation across these isolates. A significant finding from this study is the identification of six chromosomal regions, also called genomic islands (GIs), that were highly variable among the eight *Pyrococcus* genomes (Table 2). These regions contained between 5 and 126 ORFs, with 70% of the ORFs annotated as hypothetical or conserved hypothetical proteins in the genome of *P. furiosus*. No known essential or housekeeping genes were found among the variable ORFs, and the G+C content was not different than that of the rest of the genome of *P. furiosus*. Whereas CGH identifies ORFs from *P. furiosus* that are present or absent in the query isolates, it does not give information regarding genome location or gene order in the query isolates. This is particularly true for transposase sequences with nearly identical nucleotide sequences. However, an alignment of the GIs between *P. furiosus*, *Pyrococcus abyssi*, and *Pyrococcus horikoshii* substantiated the authenticity of those genomic regions (Fig. 1) (see Table S1 in the supplemental material). Patterns of presence and absence of groups of ORFs within these regions of synteny between the three *Pyrococcus* reference genomes, and the absence of those ORFs from the Vulcano isolates, characterized those regions as potential hot spots for chromosomal rearrangements. Two GIs, regions IV and VI, were associated with two to five insertion sequence (IS) elements of the IS6 family that is typically found in the *P. furiosus* genome (21), indicating possible mobile element-mediated horizontal gene transfer (HGT). This hypothesis was strengthened by the discovery of a 16-kb fragment in region V that is present only in *P. furiosus* and the closely related hyperthermophile *Thermococcus litoralis* but is absent from *P. abyssi*, *P. horikoshii*, and all the other Vulcano isolates (12). This genomic fragment was flanked by IS elements in *P. furiosus*.

These GIs have also been reported for *Prochlorococcus marinus*, a marine cyanobacterium, and *Haloquadratum walsbyi*, a halophilic archaeon (8, 9). In those organisms, GIs were associated with phage-like sequences, indicating possible phage-mediated HGTs. The gene pool associated with the *Pyrococcus* GIs is highly variable compared to those described in previous studies of GIs from environmental isolates or metagenomes (8, 9). One constant is the high number of hypothetical genes within those regions and the association with sequences from mobile genetic elements. In *Pyrococcus*, a small fraction of

* Corresponding author. Present address: Johns Hopkins University, Department of Biology, 3400 N. Charles Street, 127 Mudd Hall, Baltimore, MD 21218. Phone: (410) 516-8498. Fax: (410) 516-5213. E-mail: jdiruggiero@jhu.edu.

‡ Present address: University of Maryland School of Medicine, Institute for Genome Sciences, Baltimore, MD.

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

[∇] Published ahead of print on 22 August 2008.

TABLE 1. Strains used in this study

Strain(s)	Origin	Source or reference
<i>P. furiosus</i>	Vulcano Island ^c	13
VB81 ^a	Vulcano Island, site 20	12
VB82 ^{a,b}	Vulcano Island, site 20	12
VB85 ^a	Vulcano Island, site 20	12
VB86 ^a	Vulcano Island, site 20	12
VB91 ^a	Vulcano Island, site 20	12
VB93 ^a	Vulcano Island, site 20	12
VB112 ^a	Vulcano Island, site 2	12
12/1 ^b	East Pacific Ridge, Genesis	19
32/4 ^b	East Pacific Ridge, La Chainette	19
JT6 ^b	Juan de Fuca Ridge	Tuttle and DiRuggiero ^d

^a Strain used for CGH.^b Strain used for SSH.^c *P. furiosus* strain 3638 was obtained from the DSMZ culture collection.^d J. H. Tuttle and J. DiRuggiero, unpublished data.

ORFs were involved in transport systems for carbohydrates (ABC maltose transport system, region VI), metals or inorganic ions (sulfate transport, region VI), detoxification systems (peroxidases, region IV; daunorubicin membrane protein and ATPase, region I), peptide degradation (region IV), all functions potentially conferring an ecological advantage or a specific adaptation to a new source of nutrients. Similar findings were reported for *H. walsbyi* and *P. marinus* (8, 9), in particular at the level of transporters, supporting the idea that, in a dynamic natural environment, adaptive genomic properties such as chromosomal rearrangements are responsible for the maintenance of the microbial phenotypic diversity necessary to respond to rapidly changing environmental conditions. Many pathogenic bacteria have evolved to have phase variation, a spontaneous stochastic switching through genetic reorganization, to cope with unpredictably changing environments (18, 25). This strategy of phenotypic randomization is well known to ecologists as bet-hedging (33). Similarly, in a dynamic natural environment, the presence of different subsets of the microbial population that are well adapted to different types of environments will increase the adaptive success of the overall population.

The association of genomic “hot spots” for genome rearrangements with mobile genetic elements suggests a role for these elements in HGT. Evidence of HGT in the *Pyrococcus*

genus has been reported previously by our group and others, often in association with the presence of IS elements (11, 12, 17, 26). Differential patterns of distribution of IS elements among the Vulcano *Pyrococcus* isolates (12, 17) suggest that these IS elements are potentially mobile, as recently demonstrated for *Sulfolobus solfataricus* (22, 30). Viruses and plasmids have also been isolated from *Pyrococcus* species, suggesting a variety of processes for the transfer of genetic material (29, 39). In the present study, we used suppressive subtractive hybridization (SSH) (14, 28) to identify the complements of genes found in the genome of four *Pyrococcus* isolates but absent from the genome of *P. furiosus*, thus indicating that those genes might be the result of HGT events (Table 3). The genomic DNA from each *Pyrococcus* environmental isolate was hybridized against the *P. furiosus* genomic DNA by using a PCR-Select bacterial genome subtraction kit (Clontech, Mountain View, CA). High-throughput automated sequencing of clone inserts was performed at The Institute for Genomic Research, now the J. Craig Venter Institute, as previously described (23). Sequences have been made available in GenBank. We obtained 1,013 sequences read from four SSH libraries, totaling 564,139 bp of sequence (Table 3). BLASTX was used to align all reads to a local version of the NCBI nonredundant (nr) protein database for taxonomic and functional assignment (3). Hypothetical functional annotation also included a database of 4,873 clusters of orthologous groups (COGs). We observed a significant number of false positives (25%) among the SSH clones we sequenced, i.e., clone sequences with BLASTX hits to *Pyrococcus* sequences, requiring a large number of clones to be sequenced. Another considerable number of clones were similar to sequences from other *Pyrococcus* species (63%), indicating a significant sequence divergence of the isolates' nucleotide sequences from that of *P. furiosus*. Although the large number of SSH fragments matching *Pyrococcus* species sequences might be a limitation of the method, our analysis resulted in the characterization of 38 sequences unique to the *Pyrococcus* isolates, with no hit to any known *Pyrococcus* sequences currently available in public databases (Table 3) (see Table S2 in the supplemental material). We also found 80 sequences that had no known matches in existing databases, underlying the ability of the method to enrich for unique DNA fragments. A large number of unknown sequences was also reported in other studies, such as the *Sorcerer II* global ocean sampling expedition (31) and a study of fossil genes from

TABLE 2. ORF contents for highly variable chromosome regions I to VI

Chromosome region	Loci in <i>P. furiosus</i>	No. of transposons associated	No. of ORFs in corresponding regions ^a									
			Pfu	VB81	VB82	VB85	VB86	VB91	VB93	VB112	Pab	Pho
I	PF0025–PF0042	0	18	1	0	0	1	0	0	18	6	7
II	PF0148–PF0154	0	7	2	0	0	1	0	1	1	6 ^b	5 ^c
III	PF0289–PF0299	0	11	4	0	0	2	0	0	1	10 ^d	10 ^d
IV	PF0681–PF0807	5	126	62	18	18	64	50	50	126	ND	ND
V	PF1337–PF1341	0	5	0	0	0	0	0	1	5	5	5 ^e
VI	PF1737–PF1751	2	15	0	0	0	0	0	0	0	0	0

^a Pfu, *P. furiosus*; Pab, *P. abyssi*; Pho, *P. horikoshii*; ND, not determined.^b Five ORFs at another location.^c Four ORFs at another location.^d Eight ORFs at other locations.^e Three ORFs at another location.

FIG. 1. Highly variable chromosomal regions in *Pyrococcus* strains and environmental isolates. (A) Region I (PF0025 to PF0042; 18 ORFs). Colored arrows represent ORFs present in the three *Pyrococcus* genomes and VB112, gray arrows represent ORFs present only in *P. furiosus* and VB112, and striped arrows represent ORFs present in some environmental isolates. (B) Region V (PF1337 to PF1341; five ORFs) (boxed). Colored arrows represent ORFs present in the three *Pyrococcus* genomes and VB112, black arrows represent ORFs not in synteny in region V of *P. horikoshii*, and striped arrows represent ORFs present in some environmental isolates. (C) Region IV (PF0681 to PF0807, 126 ORFs). Colored bars indicate ORFs present (red bars) or absent (green bars); black bars indicate transposase genes that are present in *P. furiosus* and putatively present in the isolates.

ancient ice (5), reflecting the limited representation of environmental sequences in current databases.

SSH fragments from both *Archaea* (75%) and *Bacteria* (25%), but none from *Eukarya*, were identified from the four SSH libraries. Most hits to the *Archaea* were to members of the order *Thermococcales*, followed by methanogens, bacteria mostly from the firmicutes and gammaproteobacterium phyla, halophilic archaea, and various anaerobic/high-temperature archaea. ORF functions were again dominated by

hypothetical genes and by so-called operational genes, which are involved in cell growth and maintenance (see Table S2 in the supplemental material). Although recent work (36) showed that there is no metabolic category of genes resilient to transfer, the authors of the study also found that the frequency of informational gene transfer was much lower than that of operational gene transfer. This might explain the small number of those gene functions in our SSH fragments and those of others (14). Sequence reads

TABLE 3. BLASTX analysis of four SSH libraries

Libraries	Organism	No. of sequence reads	Total no. of bp	Avg bp length	No. of BLASTX hits in searches against nr database			Vector	No BLASTX hit ^e (%)
					Pfu (false positive) ^d (%)	<i>Pyrococcus</i> spp. other than Pfu ^e (%)	Other organisms ^f (%)		
SSH1	1/21 ^a	372	197,787	531	74 (19.9)	257 (69.1)	14 (3.8)	1	26 (7.0)
SSH2	JT6 ^b	275	157,930	574	58 (21.1)	179 (65.1)	11 (4.0)	1	26 (9.5)
SSH3	VB82 ^c	91	56,143	616	55 (60.4)	24 (26.4)	8 (8.8)	1	3 (3.3)
SSH7	32/4 ^a	275	152,279	553	66 (24.0)	179 (65.1)	5 (1.8)	0	25 (9.1)
Total		1,013	564,139	568.5	253 (25.0)	639 (63.1)	38 (3.8)	3	80 (7.9)

^a East Pacific Ridge.
^b Juan de Fuca Ridge.
^c Vulcano Island.
^d Top hits for *P. furiosus* (Pfu).
^e Any *Pyrococcus* sequence in 10 top hits (hits for *P. furiosus* [Pfu] were already removed); genomic DNA.
^f Significant hits for organisms other than *Pyrococcus*; *e* value > -5; manually curated.
^g BLASTN search against nr database for any *Pyrococcus* hit; possible intergenic regions.

were inspected for clustered regularly interspaced short palindromic repeat (CRISPR) elements using the program PilerCR (11a). CRISPRs have been found in 50% of the bacterial and most of the archaeal genomes sequenced so far and have been implicated in an RNA interference type of defense mechanism directly primarily against viruses (4, 20, 35). Several of the SSH clones we analyzed contained CRISPR sequence elements, suggesting a possible mode of propagation for those sequences between strains and across lineages (15). A CGH study that compared nine *Thermotoga* strains from different geographic locations to the sequenced *Thermotoga maritima* strain MSB8 also revealed significant differences across the strains that could be related to substrate utilization, and there were many examples of deletion/insertion events with complete cassettes of genes and gene rearrangements (24).

In the present study of the genotypic diversity in a population of *Pyrococcus* species, hyperthermophilic archaea isolated from hydrothermal vents in Vulcano Island, Italy, we showed extensive genome rearrangements, revealing a possible mechanism by which microbial phenotypic diversity is maintained in this ecosystem. The shallow vents of Vulcano Island are characterized by extensive variations in physicochemical conditions (6, 32), raising the question of whether adaptive genomic properties, such as genome rearrangements and transposition and lateral gene transfer events, are a necessary adaptation to rapidly changing environmental conditions.

Nucleotide sequence accession numbers. Sequences are available under GenBank accession numbers EU922594 to EU923604.

This research was funded under a grant from the U.S. Department of Energy (DE-FG02-01ER63133) to K.E.N. and J.D.

We thank Patrick Forterre and Evelyne Marguet for providing strains 12/1 and 32/4.

REFERENCES

- Acinas, S. G., V. Klepac-Ceraj, D. E. Hunt, C. Phario, I. Ceraj, D. L. Distel, and M. F. Polz. 2004. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**:551–554.
- Allen, E. E., and J. F. Banfield. 2005. Community genomics in microbial ecology and evolution. *Nat. Rev. Microbiol.* **3**:489–498.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**:1709–1712.
- Bidle, K. D., S. Lee, D. R. Marchant, and P. G. Falkowski. 2007. Fossil genes and microbes in the oldest ice on Earth. *Proc. Natl. Acad. Sci. USA* **104**:13455–13460.
- Capasso, G., R. S. Favara, S. Francoforte, and S. Inguaggiato. 1999. Chemical and isotopic variations in fumarolic discharge and thermal waters at Vulcano Island (Aeolian Islands, Italy). *J. Volcanol. Geothermal Res.* **88**:167–175.
- Reference deleted.
- Coleman, M. L., M. B. Sullivan, A. C. Martiny, C. Steglich, K. Barry, E. F. Delong, and S. W. Chisholm. 2006. Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* **311**:1768–1770.
- Cuadros-Orellana, S., A. B. Martin-Cuadrado, B. Legault, G. D'Auria, O. Zhaxybayeva, R. T. Papke, and F. Rodriguez-Valera. 2007. Genomic plasticity in prokaryotes: the case of the square haloarchaeon. *ISME J.* **1**:235–245.
- DeLong, E. F., C. M. Preston, T. Mincer, V. Rich, S. J. Hallam, N.-U. Frigaard, A. Martinez, M. B. Sullivan, R. Edwards, B. R. Brito, S. W. Chisholm, and D. M. Karl. 2006. Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**:496–503.
- DiRuggiero, J., D. Dunn, D. L. Maeder, R. Holley-Shanks, J. Chatard, R. Horlacher, F. T. Robb, W. Boos, and R. B. Weiss. 2000. Evidence of recent lateral gene transfer among hyperthermophilic archaea. *Mol. Microbiol.* **38**:684–693.
- Edgar, R. 2007. PILER-CR: fast and accurate identification of CRISPR repeats. *BMC Bioinform.* **8**:18.
- Escobar-Paramo, P., S. Ghosh, and J. DiRuggiero. 2005. Evidence for genetic drift in the diversification of a geographically isolated population of the hyperthermophilic archaeon *Pyrococcus*. *Mol. Biol. Evol.* **22**:2297–2303.
- Fiala, G., and K. O. Stetter. 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100°C. *Arch. Microbiol.* **145**:56–61.
- Galbraith, E. A., D. A. Antonopoulos, and B. A. White. 2004. Suppressive subtractive hybridization as a tool for identifying genetic diversity in an environmental metagenome: the rumen as a model. *Environ. Microbiol.* **6**:928–937.
- Grissa, I., G. Vergnaud, and C. Pourcel. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinform.* **8**:172.
- Reference deleted.
- Kanoksilapatham, W., J. M. González, D. L. Maeder, J. DiRuggiero, and F. T. Robb. 2004. A proposal to rename the hyperthermophile *Pyrococcus woesei* as *Pyrococcus furiosus* subsp. *woesei*. *Archaea* **1**:275–281.
- Kim, C. C., E. A. Joyce, K. Chan, and S. Falkow. 2002. Improved analytical methods for microarray-based genome-composition analysis. *Genome Biol.* **3**:research0065.1–0065.17.
- Kussell, E., and S. Leibler. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**:2075–2078.
- Lepage, E., E. Marguet, C. Geslin, O. Matte-Tailliez, W. Zillig, P. Forterre, and P. Tailliez. 2004. Molecular diversity of new *Thermococcales* isolates from a single area of hydrothermal deep-sea vents as revealed by randomly amplified polymorphic DNA fingerprinting and 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.* **70**:1277–1286.
- Lillestøl, R. K., P. Redder, R. Garrett, and K. Brugger. 2006. A putative viral defense mechanism in archaeal cells. *Archaea* **2**:59–72.
- Maeder, D. L., R. B. Weiss, D. M. Dunn, J. L. Cherry, J. M. Gonzalez, J. DiRuggiero, and F. T. Robb. 1999. Divergence of the hyperthermophilic archaea *Pyrococcus furiosus* and *P. horikoshii* inferred from complete genomic sequences. *Genetics* **152**:1299–1305.
- Martusewitsch, E., C. W. Sensen, and C. Schleper. 2000. High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements. *J. Bacteriol.* **182**:2574–2581.
- Mongodin, E., J. Finan, M. W. Climo, A. Rosato, S. Gill, and G. L. Archer. 2003. Microarray transcription analysis of clinical *Staphylococcus aureus* isolates resistant to vancomycin. *J. Bacteriol.* **185**:4638–4643.
- Mongodin, E. F., I. R. Hance, R. T. Deboy, S. R. Gill, S. Daugherty, R. Huber, C. M. Fraser, K. Stetter, and K. E. Nelson. 2005. Gene transfer and genome plasticity in *Thermotoga maritima*, a model hyperthermophilic species. *J. Bacteriol.* **187**:4935–4944.
- Moxon, E. R., R. E. Lenski, and P. B. Rainey. 1998. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Perspect. Biol. Med.* **42**:154–155.
- Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser. 1999. Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**:323–329.
- Nesbø, C. L., M. Dlutek, and W. F. Doolittle. 2006. Recombination in *Thermotoga*: implications for species concepts and biogeography. *Genetics* **172**:759–769.
- Nesbø, C. L., K. E. Nelson, and W. F. Doolittle. 2002. Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima*. *J. Bacteriol.* **184**:4475–4488.
- Prieur, D., G. Erauso, C. Geslin, S. Lucas, M. Gaillard, A. Bidault, A. Mattenet, K. Rouault, D. Flament, P. Forterre, and M. Le Romancer. 2004. Genetic elements of *Thermococcales*. *Biochem. Soc. Trans.* **32**:184–187.
- Redder, P., and R. A. Garrett. 2006. Mutations and rearrangements in the genome of *Sulfolobus solfataricus* P2. *J. Bacteriol.* **188**:4198–4206.
- Rusch, D. B., A. L. Halpern, G. Sutton, K. B. Heidelberg, S. Williamson, S. Yooseph, D. Wu, J. A. Eisen, J. M. Hoffman, K. Remington, K. Beeson, B. Tran, H. Smith, H. Baden-Tillson, C. Stewart, J. Thorpe, J. Freeman, C. Andrews-Pfannkoch, J. E. Venter, K. Li, S. Kravitz, J. F. Heidelberg, T. Utterback, Y.-H. Rogers, L. I. Falcón, V. Souza, G. Bonilla-Rosso, L. E. Eguarte, D. M. Karl, S. Sathyendranath, T. Platt, E. Bermingham, V. Gallardo, G. Tamayo-Castillo, M. R. Ferrari, R. L. Strausberg, K. Neelson, R. Friedman, M. Frazier, and J. C. Venter. 2007. The *Sorcerer II* global ocean sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* **5**:e77.
- Sedwick, P., and D. Stueben. 1996. Chemistry of shallow submarine warm springs in an arc-volcanic setting: Vulcano Island, Aeolian Archipelago, Italy. *Mar. Chem.* **53**:147–161.
- Seger, J., and H. J. Brockmann. 1987. What is bet-hedging?, p. 182–211. *In*

- P. H. Harvey and L. Partridge (ed.), Oxford surveys in evolutionary biology, vol. 4. Oxford University Press, Oxford, United Kingdom.
34. **Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl.** 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proc. Natl. Acad. Sci. USA* **103**:12115–12120.
35. **Sorek, R., V. Kunin, and P. Hugenholtz.** 2008. CRISPR: a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* **6**:181–186.
36. **Sorek, R., Y. Zhu, C. J. Creevey, M. P. Francino, P. Bork, and E. M. Rubin.** 2007. Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **318**:1449–1452.
37. **Thompson, J. R., S. Pacocha, C. Pharino, V. Klepac-Ceraj, D. E. Hunt, J. Benoit, R. Sarma-Rupavtarm, D. L. Distel, and M. F. Polz.** 2005. Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**:1311–1313.
38. **Tyson, G. W., J. Chapman, P. Hugenholtz, E. E. Allen, R. J. Ram, P. M. Richardson, V. V. Solovyev, E. M. Rubin, D. S. Rokhsar, and J. F. Banfield.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**:37–43.
39. **Ward, D. E., I. M. Revet, R. Nandakumar, J. H. Tuttle, W. M. de Vos, J. van der Oost, and J. DiRuggiero.** 2002. Characterization of plasmid pRT1 from *Pyrococcus* sp. strain JT1. *J. Bacteriol.* **184**:2561–2566.
40. **Whitaker, R. J., D. W. Grogan, and J. W. Taylor.** 2005. Recombination shapes the natural population structure of the hyperthermophilic archaeon *Sulfolobus islandicus*. *Mol. Biol. Evol.* **22**:2354–2361.