Sulfolobus islandicus: a model system for evolutionary genomics

Changyi Zhang, David J. Krause and Rachel J. Whitaker[1]
Department of Microbiology, University of Illinois at Urbana-Champaign, 601 S. Goodwin Ave, Urbana, IL 61801, U.S.A.

Abstract

Sulfolobus islandicus **has been developed as a model system for combining approaches of evolutionary and molecular biology in Archaea. We describe how the application of this interdisciplinary approach can lead to novel hypotheses derived from patterns of natural variation that can be tested in the laboratory when combined with a diversity of natural variants and versatile genetic markers. We review how this approach has highlighted the importance of recombination as an evolutionary parameter and provided insight into a molecular mechanism of recombination that may be unique in the archaeal domain. We review the development and improvement of the model system *S. islandicus* that will enable us to study the mechanism and genomic architecture of recombination guided by evolutionary genomic analysis of Nature's ongoing experiments in wild populations.**

Background

Although historically studied as unique and independent disciplines, the fusion of evolutionary and molecular biology has led to great advances in our understanding of the interactions among different scales of life from cells to populations of organisms (see [1] for a review of two examples). Research at the interface of these two fields is able to utilize molecular approaches to validate evolutionary hypotheses obtained from population genomics studies, while also lending important evolutionary perspectives to findings at a molecular level [2]. The synthesis of evolutionary and molecular biology has currently found favour in model eukaryotes such as *Drosophila* [3] and *Saccharomyces* [4], yet any organism with clearly defined patterns of natural variation among closely related individuals and a tractable genetic system could be a model system for research at this junction. Although tractable genetic systems are commonplace in the Bacteria and Archaea, complementary studies of natural variation in members of these domains are scarce [5–8]. Rarely have micro-organisms been exposed to this powerful synthesis. In the present review, we highlight the development of *Sulfolobus islandicus* as a model system for research at the molecular–evolutionary junction. We focus on a demonstration of the power of these tools in combination to study genetic exchange in Archaea.

Hypotheses developed from patterns of natural variation

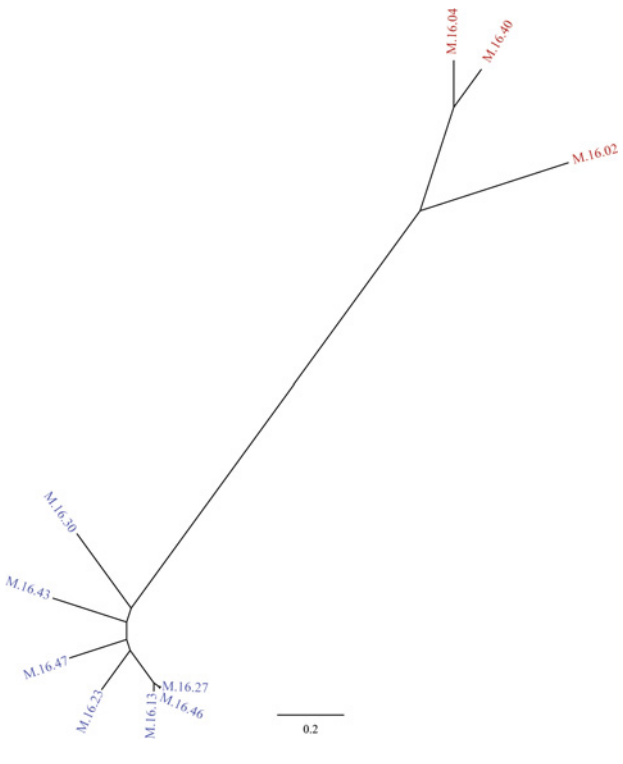
Traditional molecular microbiology seeks to eliminate natural variation, restricting variables to only those introduced in the laboratory. Although this approach was successful in the 20th Century, our understanding of the evolutionary process is limited to hypotheses generated *a priori.* Population genomics, made possible by high-throughput genome sequencing and environmental genomics, has now revealed both novel mechanisms of evolution [9] and new molecular systems [10] that could not have been imagined *a priori* by examining a few individual captive strains in the laboratory.

To gain new insights into evolutionary processes, patterns of natural variation must be examined at a population scale, focusing on closely related strains and therefore on recent evolutionary events. This level of evolutionary resolution eases the attribution of evolutionary and phenotypic consequences to specific genomic differences and greatly simplifies linking genotype and phenotype in the laboratory by decreasing the number of variables changed between strains. In addition, this evolutionary scale allows inference about the interaction between fundamental evolutionary parameters, such as mutation, selection, gene flow and neutral genetic drift. A precaution for these types of studies, however, is that careful consideration of sampling strategies is essential to answer specific questions [5].

As an example of using population genomics to examine and identify essential evolutionary processes, we review the development of *S. islandicus* as a model system to study recombination in Archaea. The Euryarchaea have received some attention for the effects of recombination in natural populations both in *Halorubrum* [11] and *Ferroplasma* [12], as well as empirical evidence for recombination among species of halophilic archaea [13]. In the Crenarchaea, recombination has been explored extensively at the molecular level using *Sulfolobus acidocaldarius* as a model system [14–18]. We were interested in recombination in *S. islandicus*, because it is a fundamental parameter that defines the process of evolution [19], and because we and others hypothesize that the study of this process in Crenarchaea may help to identify molecular

Key words: chromosomal marker exchange, evolution, recombination, *Sulfolobus islandicus*.
Abbreviations used: 5-FOA, 5-fluoro-orotic acid; X-gal, 5-bromo-4-chloroindol-3-yl β-D-galactopyranoside.
[1]To whom correspondence should be addressed (email rwhitaker@life.illinois.edu).

Figure 1 | **Phylogeny of the Red and Blue species of *S. islandicus***
The Red and Blue species are two sympatric divergent lineages defined by low rates of interspecies genetic exchange as described further in [5]. Tree built from an alignment of whole genome core nucleotide positions using ClonalFrame. Scale bar, units of substitutions per 1000 sites.



systems present in the common ancestor of Archaea and Eukaryotes [20–23].

Evolutionary analysis of a single isolated population

Population studies showed that a geographically isolated *S. islandicus* population from the Mutnovsky volcano in Kamchatka, Russia [24], was more diverse than North American *S. islandicus* populations [25]. In addition, multi-locus sequence analysis demonstrated evidence of recombination in this population [26]. Evolutionary hypotheses for this pattern included that the Mutnovsky population was older, more stable or more highly recombining than the North American populations. To better test these hypotheses, we conducted an additional population genomic analysis of diversity within a single hot spring from the Mutnovsky region. This study revealed a fourth, unexpected, pattern in which diversity within the Mutnovsky population was maintained by the coexistence of two sympatric species of *S. islandicus* [5] (Figure 1). Species are defined because genome analysis shows higher rates of recombination within than between them [5,27,28]. While solving one mystery about diversity in the Mutnovsky population, this study opened up new questions and hypotheses about mechanisms of speciation that demand genetic tools to investigate in the laboratory.

Two mechanisms maintain coexisting species: intrinsic barriers to gene transfer and ecological barriers. *Sulfolobus* cells have been shown to aggregate in response to UV irradiation through UV-inducible type IV pili, and this response mediates the exchange of DNA between cells in the divergent species *S. acidocaldarius* [16,29]. A barrier to recombination could be an inability of cells to successfully generate mating pairs and exchange DNA, an inability to recombine exchanged DNA or synthetic lethality from the combination of new alleles. Ecological speciation would occur if two species are "haunting different stations" [30], or utilizing different resources, within a single hot spring. Owing to the difficulty in recognizing biologically relevant environmental heterogeneity at the microbial scales of time and space, distinguishing between these hypotheses must be resolved in the laboratory using a genetic system that works for a diversity of strains.

Building versatile tools for wild *S. islandicus* strains

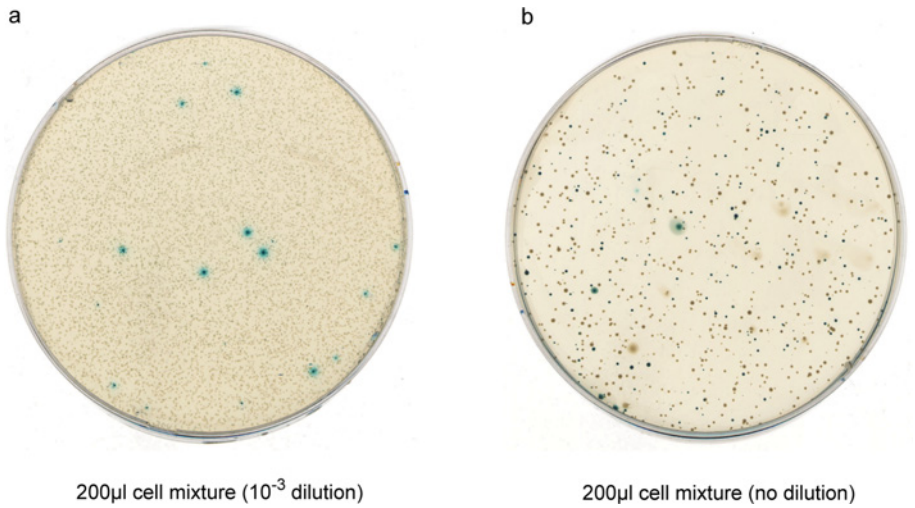
Great strides have been made recently in establishing *S. islandicus* as a model system for studying the biology of the Archaea from both biochemical and genetic standpoints. There now exist multiple selectable markers, efficient electroporation-based transformation techniques, homologous-recombination-based gene-knockout systems, and shuttle-vector-based gene reporter and overexpression systems [31–35]. We have been exploiting versatile genetic markers, especially efficient markers for positive selection, as well as developing new experimental tools to investigate the processes of marker exchange and recombination in *S. islandicus*, thereby testing hypotheses developed from observations of natural populations.

Our aim is to empirically test hypotheses regarding marker exchange in *S. islandicus* in order to elucidate the role of genetic exchange in microbial speciation. Specifically, if intrinsic barriers exist, strains of two different sympatric species will cross at a lower frequency than strains of the same species in the laboratory. Initial efforts focused on isolating various genotypes of uracil auxotrophs in strains of *S. islandicus* that fit into the two distinct species, which is readily achieved by spreading wild-type cells on to nutritive plates containing 5-FOA (5-fluoro-orotic acid) and uracil. Attempts to conduct a selection for uracil prototroph recombinants similar to those carried out in *S. acidocaldarius* failed [18], owing to the fact that even uracil auxotrophs can grow on plates lacking uracil. This was not surprising, as this 'leaky' phenomenon of uracil-based selection has been widely reported in genetic manipulations of several hyperthermophiles. It is still unclear why uracil selection is exhibited stringently in *S. acidocaldarius* and *Pyrococcus furiosus*, yet has no utility in *S. islandicus*, *S. solfataricus* or *Thermococcus kodakarensis* [34,36–39].

Because selection for uracil auxotrophs failed in *S. islandicus*, an additional genetic marker was needed. The

Figure 2 | Mating assay for studying marker exchange in *S. islandicus*

(**a**) Tester strain *pyrEF*$^-$ *lacS*$^-$ was crossed with the M.16.04 wild-type strain. (**b**) Tester strain *pyrEF*$^-$ *lacS*$^-$ *argDC*$^-$ was crossed with the M.16.27 wild-type strain. Presence of *lacS* activity causes hydrolysis of X-gal, resulting in blue colonies. White colonies are *lacS*$^-$.



assay was modified by introducing a mutation of the reporter gene *lacS*, which encodes β-galactosidase. A genetically stable uracil auxotrophic strain with a precise deletion of *pyrEF* (encoding orotidine-5′-monophosphate pyrophosphorylase and orotidine-5′ monophosphate decarboxylase respectively) was constructed in the M.16.04 wild-type *S. islandicus* strain using a gene-knockout system based on simvastatin selection, in which an overexpression cassette of the HMG-CoA (3-hydroxy-3-methylglutaryl-CoA) reductase gene (*hmgR*) was used as a selectable marker [34]. This marker-free *pyrEF* deletion mutant was used further as a background strain to delete the *lacS* gene using the same strategy, generating *pyrEF* and *lacS* double-deletion mutants (C. Zhang and R.J. Whitaker, unpublished work). Using the double mutant as a tester strain, a mating system was set up in which wild-type M.16.04 was mixed with the tester strain with or without UV irradiation and then spread on to 5′-FOA-containing plates. Two apparent genotypes arise from this process: the tester strain *pyrEF*$^-$ *lacS*$^-$, or recombinant *pyrEF*$^-$ *lacS*$^+$ (Figure 2a). We estimated a rate of recombination using this system of approximately 10^{-3} and 10^{-4} with and without UV irradiation respectively, once corrected for spontaneous mutation in the *pyrEF* locus of the wild-type strain.

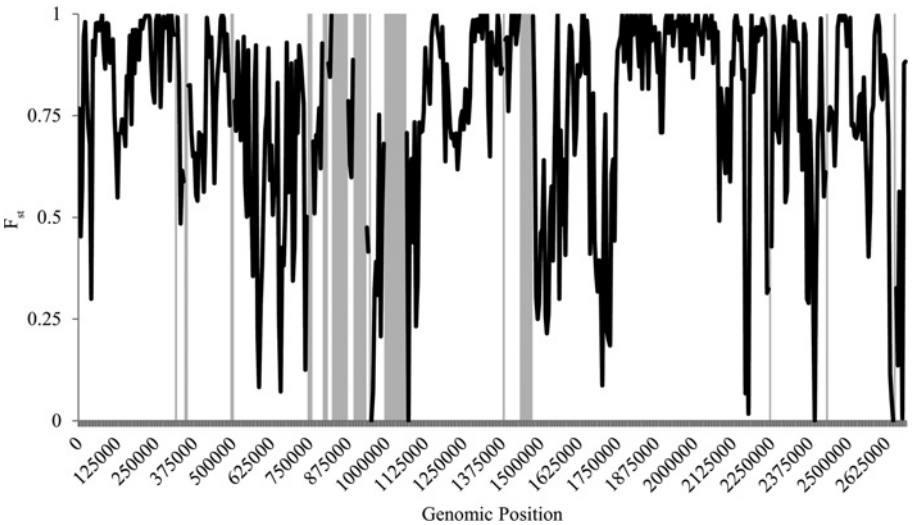
The necessity for blue/white screening limits the detection power for finding recombinants on plates when recombination rates are very low ($<10^{-5}$). Also, it was difficult to isolate and purify blue colonies from the tens of thousands of surrounding white colonies. Therefore a more sophisticated exchange system relying on multiple efficient selection markers was required. The *pyrEF*$^-$ *lacS*$^-$ double mutant was used in conjunction with a hybrid *pyrEF*/*lacS*-selectable marker to delete the arginine decarboxylase-encoding gene, *argDC*, via marker-less gene deletion, generating a triple mutant. Growth analysis showed that the triple mutant was unable to grow in liquid medium or on plates without agmatine; however, normal growth was restored upon supplementation with agmatine. Furthermore, genetic complementation of the triple mutant with a functional copy of *argDC in trans* can convert mutants from agmatine auxotrophs into prototrophs (C. Zhang and R.J. Whitaker, unpublished work). This clear phenotype made it possible to apply *argDC* as a reliable marker. Unlike the previous system with dominant growth of the tester strain, this modified system only allows the growth of recombinants and spontaneous mutants (Figure 2b). Similarly, there are also two genotypes on these plates, *pyrEF*$^-$ *lacS*$^-$ *argDC*$^+$ and *pyrEF*$^-$ *lacS*$^+$ *argDC*$^+$; however, low frequencies can be estimated with greater confidence. The colonies can be analysed by X-gal (5-bromo-4-chloroindol-3-yl β-D-galactopyranoside) staining and PCR diagnosis. Genetic tools now provide a means to test the hypothesis and identify the genetic basis for intrinsic barriers to recombination among closely related strains, should they exist.

Additional questions generated by population genomic studies

Additional information about how the mechanism and genome structure contribute to patterns of recombination can be derived from population genomic data. Figure 3 shows population differentiation between the two species measured throughout the genome using a sliding-window approach to see how different regions varied in their signal of fixation between species. This analysis showed that there is a mosaic of speciation in which a few large regions of the genome are highly differentiated between species and do not share

Figure 3 | Differentiation between species occurs in large genomic continents

F_{st} (fixation index) values from comparisons of the Blue and Red species are plotted for 10 kb windows with 5 kb steps. The line represents the 10 kb average of F_{st}. Grey bars represent windows with less than 5 kb of core nucleotide positions and are not analysed. High F_{st} values indicate increased differentiation, and low F_{st} values indicate reduced differentiation.



variation, whereas other regions of the chromosome were less differentiated, implying that variation is shared [40,41]. This mosaic of speciation could be driven by variation in selection, mutation or recombination across the chromosome. Previous studies have suggested that variation in genetic diversity similarly exists in more divergent isolated strains also driven by one of these three mechanisms [42,43]. Population genomic analyses of a single population allow us to test how genomic variation in these three processes contributes to these patterns. These results can be confirmed by investigating patterns of recombination in laboratory crosses using natural SNPs (single nucleotide polymorphisms) as genetic markers and high-throughput genomic sequencing. Remaining questions as to the mechanisms that cause variation in recombination, mutation or selection around the chromosome and whether the mechanism of genetic exchange contributes to the pattern that we observe can now be tested in the laboratory. Resolving the answers to these questions requires a molecular approach that explores a diversity of strains as described above.

Concluding remarks

S. islandicus is poised to become a model system in which one can combine evolutionary inference with molecular tools. We have demonstrated how a synthetic approach can provide insight into the molecular mechanisms of recombination within archaeal populations. Future studies hold great promise to define these mechanisms and how recombination is distributed around the genome. Application of a similar synthetic approach has been used to identify and experimentally test essential rapidly evolving genomic loci and the selection pressures that are imposed on them, evolutionary trajectories defined by geographic isolation, genetic drift and speciation, and the basis of virus–host co-evolutionary interactions. In addition, a synthetic approach can reveal a great deal about molecular mechanisms such as mutation, transposition and horizontal gene transfer. Applying this sort of synthesis to organisms from the third domain is sure to uncover novel insights into evolutionary and molecular biology that could not have been imagined without looking outside the laboratory.

Acknowledgements

We thank Christa Schleper for her suggestion of using *pyrEF*$^-$/*lacS*$^-$ mutants to assess recombination frequencies; Sonja Albers for helpful suggestions on our crossing protocol; Qunxin She and Yulong Shen for assistance with developing the simvastatin-based gene deletion system; Dennis Grogan and Alfa Herrera for assistance with strain isolation; Hinsby Cadillo-Quiroz for genome assembly and Alvaro Hernandez at the UIUC Keck Center for help with genome sequencing.

Funding

This work was supported by the National Aeronautics and Space Administration (NASA) [grant number NNX09AM92G].

References

1 Barrett, R.D. and Hoekstra, H.E. (2011) Molecular spandrels: tests of adaptation at the genetic level. Nat. Rev. Genet. **12**, 767–780

2 Dean, A.M. and Thornton, J.W. (2007) Mechanistic approaches to the study of evolution: the functional synthesis. Nat. Rev. Genet. **8**, 675–688
3 Presgraves, D.C., Balagopalan, L., Abmayr, S.M. and Orr, H.A. (2003) Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature **423**, 715–719
4 Delneri, D., Colson, I., Grammenoudi, S., Roberts, I.N., Louis, E.J. and Oliver, S.G. (2003) Engineering evolution to study speciation in yeasts. Nature **422**, 68–72
5 Cadillo-Quiroz, H., Didelot, X., Held, N.L., Herrera, A., Darling, A., Reno, M.L., Krause, D.J. and Whitaker, R.J. (2012) Patterns of gene flow define species of thermophilic Archaea. PLoS Biol. **10**, e1001265
6 Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O. et al. (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. **5**, e1000344
7 Haven, J., Vargas, L.C., Mongodin, E.F., Xue, V., Hernandez, Y., Pagan, P., Fraser-Liggett, C.M., Schutzer, S.E., Luft, B.J., Casjens, S.R. and Qiu, W.G. (2011) Pervasive recombination and sympatric genome diversification driven by frequency-dependent selection in *Borrelia burgdorferi*, the Lyme disease bacterium. Genetics **189**, 951–966
8 Shapiro, B.J., Friedman, J., Cordero, O.X., Preheim, S.P., Timberlake, S.C., Szabo, G., Polz, M.F. and Alm, E.J. (2012) Population genomics of early events in the ecological differentiation of bacteria. Science **336**, 48–51
9 Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S. et al. (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U.S.A. **102**, 13950–13955
10 Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I. and Koonin, E.V. (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct **1**, 7
11 Papke, R.T., Koenig, J.E., Rodriguez-Valera, F. and Doolittle, W.F. (2004) Frequent recombination in a saltern population of *Halorubrum*. Science **306**, 1928–1929
12 Eppley, J.M., Tyson, G.W., Getz, W.M. and Banfield, J.F. (2007) Genetic exchange across a species boundary in the archaeal genus *Ferroplasma*. Genetics **177**, 407–416, (2007)
13 Naor, A., Lapierre, P., Mevarech, M., Papke, R.T. and Gophna, U. (2012) Low species barriers in halophilic archaea and the formation of recombinant hybrids. Curr. Biol. **22**, 1444–1448
14 Grogan, D.W. and Rockwood, J. (2010) Discontinuity and limited linkage in the homologous recombination system of a hyperthermophilic archaeon. J. Bacteriol. **192**, 4660–4668
15 Mao, D. and Grogan, D.W. (2012) Heteroduplex formation, mismatch resolution, and genetic sectoring during homologous recombination in the hyperthermophilic archaeon *Sulfolobus acidocaldarius*. Front. Microbiol. **3**, 192
16 Ajon, M., Frols, S., van Wolferen, M., Stoecker, K., Teichmann, D., Driessen, A.J., Grogan, D.W., Albers, S.-V. and Schleper, C. (2011) UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. Mol. Microbiol. **82**, 807–817
17 White, M.F. (2011) Homologous recombination in the archaea: the means justify the ends. Biochem. Soc. Trans. **39**, 15–19
18 Grogan, D.W. (1996) Exchange of genetic markers at extremely high temperatures in the archaeon *Sulfolobus acidocaldarius*. J. Bacteriol. **178**, 3207–3211
19 Smith, J.M. (1993) The role of sex in bacterial evolution. J. Hered. **84**, 326–327
20 Yutin, N., Makarova, K.S., Mekhedov, S.L., Wolf, Y.I. and Koonin, E.V. (2008) The deep archaeal roots of eukaryotes. Mol. Biol. Evol. **25**, 1619–1630
21 Barry, E.R. and Bell, S.D. (2006) DNA replication in the archaea. Microbiol. Mol. Biol. Rev. **70**, 876–887
22 Cann, I.K. and Ishino, Y. (1999) Archaeal DNA replication: identifying the pieces to solve a puzzle. Genetics **152**, 1249–1267
23 Samson, R.Y. and Bell, S.D. (2011) Cell cycles and cell division in the archaea. Curr. Opin. Microbiol. **14**, 350–356
24 Whitaker, R.J., Grogan, D.W. and Taylor, J.W. (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science **301**, 976–978
25 Reno, M.L., Held, N.L., Fields, C.J., Burke, P.V. and Whitaker, R.J. (2009) Biogeography of the *Sulfolobus islandicus* pan-genome. Proc. Natl. Acad. Sci. U.S.A. **106**, 8605–8610
26 Whitaker, R.J., Grogan, D.W. and Taylor, J.W. (2005) Recombination shapes the natural population structure of the hyperthermophilic archaeon *Sulfolobus islandicus*. Mol. Biol. Evol. **22**, 2354–2361
27 Dykhuizen, D.E. and Green, L. (1991) Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. **173**, 7257–7268
28 Mayr, E. (1942) Systematics and the Origin of Species from the Viewpoint of a Zoologist, Columbia University Press, New York
29 Frols, S., Ajon, M., Wagner, M., Teichmann, D., Zolghadr, B., Folea, M., Boekema, E.J., Driessen, A.J., Schleper, C. and Albers, S.-V. (2008) UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation. Mol. Microbiol. **70**, 938–952
30 Darwin, C. (1859) On the Origin of Species by Means of Natural Selection, J. Murray, London
31 Deng, L., Zhu, H., Chen, Z., Liang, Y.X. and She, Q. (2009) Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. Extremophiles **13**, 735–746
32 Zhang, C., Guo, L., Deng, L., Wu, Y., Liang, Y., Huang, L. and She, Q. (2010) Revealing the essentiality of multiple archaeal *pcna* genes using a mutant propagation assay based on an improved knockout method. Microbiology **156**, 3386–3397
33 Zheng, T., Huang, Q., Zhang, C., Ni, J., She, Q. and Shen, Y. (2012) Development of a simvastatin selection marker for a hyperthermophilic acidophile, *Sulfolobus islandicus*. Appl. Environ. Microbiol. **78**, 568–574
34 Zhang, C. and Whitaker, R.J. (2012) A broadly applicable gene knockout system for the thermoacidophilic archaeon *Sulfolobus islandicus* based on simvastatin selection. Microbiology **158**, 1513–1522
35 Peng, N., Deng, L., Mei, Y., Jiang, D., Hu, Y., Awayez, M., Liang, Y. and She, Q. (2012) A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon *Sulfolobus islandicus*. Appl. Environ. Microbiol. **78**, 5630–5637
36 Sato, T., Fukui, T., Atomi, H. and Imanaka, T. (2003) Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J. Bacteriol. **185**, 210–220
37 Wagner, M., Berkner, S., Ajon, M., Driessen, A.J., Lipps, G. and Albers, S.-V. (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. Biochem. Soc. Trans. **37**, 97–101
38 Wagner, M., van Wolferen, M., Wagner, A., Lassak, K., Meyer, B.H., Reimann, J. and Albers, S.-V. (2012) Versatile genetic tool box for the crenarchaeote *Sulfolobus acidocaldarius*. Front. Microbiol. **3**, 214
39 Lipscomb, G.L., Stirrett, K., Schut, G.J., Yang, F., Jenney, Jr, F.E., Scott, R.A., Adams, M.W. and Westpheling, J. (2011) Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. Appl. Environ. Microbiol. **77**, 2232–2238
40 Via, S. (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow, Philos. Trans. R. Soc. London Ser. B **367**, 451–460
41 Nosil, P. and Feder, J.L. (2012) Genomic divergence during speciation: causes and consequences. Philos. Trans. R. Soc. London Ser. B **367**, 332–342
42 Flynn, K.M., Vohr, S.H., Hatcher, P.J. and Cooper, V.S. (2010) Evolutionary rates and gene dispensability associate with replication timing in the archaeon *Sulfolobus islandicus*. Genome Biol. Evol. **2**, 859–869
43 Andersson, A.F., Pelve, E.A., Lindeberg, S., Lundgren, M., Nilsson, P. and Bernander, R. (2010) Replication-biased genome organisation in the crenarchaeon *Sulfolobus*. BMC Genomics **11**, 454

Received 15 November 2012
doi:10.1042/BST20120338