SHORT COMMUNICATION

Genomic evidence of rapid, global-scale gene flow in a *Sulfolobus* species

Dominic Mao and Dennis Grogan

Department of Biological Sciences, University of Cincinnati, Cincinnati, OH, USA

Local populations of *Sulfolobus islandicus* diverge genetically with geographical separation, and this has been attributed to restricted transfer of propagules imposed by the unfavorable spatial distribution of acidic geothermal habitat. We tested the generality of genetic divergence with distance in *Sulfolobus* species by analyzing genomes of *Sulfolobus acidocaldarius* drawn from three populations separated by more than 8000 km. In sharp contrast to *S. islandicus*, the geographically diverse *S. acidocaldarius* genomes proved to be nearly identical. We could not link the difference in genome conservation between the two species to a corresponding difference in genome stability or ecological factors affecting propagule dispersal. The results provide the first evidence that genetic isolation of local populations does not result primarily from properties intrinsic to *Sulfolobus* and the severe discontinuity of its geothermal habitat, but varies with species, and thus may reflect biotic interactions that act after propagule dispersal.

Genetic divergence with distance in *Sulfolobus* species

The extremely thermoacidophilic archaeon *Sulfolobus islandicus* colonizes acidic hot springs throughout the Northern Hemisphere and has diversified extensively, both within local populations and as a function of the distance between them (Whitaker *et al.*, 2003). The observed divergence with distance implies that immigration leading to gene flow between local populations is severely restricted. This contrasts sharply with the efficient passive dispersal of most micro-organisms (Finlay 2002), and has made *S. islandicus* an influential model of prokaryotic allopatric diversification (Martiny *et al.*, 2006).

However, the pattern of genetic divergence does not, in itself, define the basis of gene-flow restriction. The *S. islandicus* barrier is commonly attributed (sometimes implicitly) to the fact that the environmental conditions required by *Sulfolobus* species (Huber and Prangishvili, 2006) occur in geothermal sites that are extremely limited in area and separated by extremely large distances (Fenchel, 2003; Hahn and Pöckl, 2005; Knittel *et al.*, 2005; Cohan, 2006; Ramette and Tiedje, 2007; Pagaling *et al.*, 2009).

Successful immigration thus requires *Sulfolobus* cells to (i) be shed from relatively small populations, (ii) retain viability during prolonged transport and (iii) enter small islands of geothermal habitat. Although the hypothesis thus seems plausible that genetic isolation of *S. islandicus* populations results from excessive dilution and death of propagules (reflecting the interaction of organismal physiology and physical geography), this has not been tested. Importantly, it predicts that all *Sulfolobus* species will diverge similarly with distance, as the growth requirements that define the genus also define similar geothermal habitats having a similar spatial distribution for all its species.

Three geographically distinct strains of *S. acidocaldarius* have been reported to date: (i) the type strain 98-3 isolated from the Norris Geyser Basin of Yellowstone National Park (Brock *et al.*, 1972), (ii) strain N8 isolated from the Jigokudani thermal field near Norboribetsu, on the Japanese island of Hokkaido (Kurosawa *et al.*, 1995) and (iii) strain Ron 12/I cultured from self-heating mining waste near Ronneburg in eastern Germany (Fuchs *et al.*, 1995). The average distance between these sites (that is, 8209 km) exceeds the largest distances represented by sequenced *S. islandicus* genomes by about 2000 km (Reno *et al.*, 2009), and thus should measure primarily the impact of geographical separation, with only a minor contribution from local *Sulfolobus* diversity (Whitaker *et al.*, 2003; Reno *et al.*, 2009). We therefore investigated divergence by distance in *S. acidocaldarius* by

determining the complete genome sequences of *Sulfolobus* strains N8 (NBRC 15159) and Ron 12/I. We re-purified the strains on plates, extracted genomic DNAs, and submitted the DNAs for 76 extension cycles on an Illumina (San Diego, CA, USA) GAII instrument (DNA Analysis Facility, Iowa State University). About six million reads per genome were assembled (220-fold average coverage) and the resulting alignments were analyzed, using Genomics Workbench (CLC bio, Aarhus, Denmark). We filled small assembly gaps at repeats and low-coverage sites by Sanger sequencing of PCR products, to yield closed circular genome sequences.

Conservation of the *S. acidocaldarius* genome

Despite the global-scale distances represented, the genomes of strains N8, Ron12/I and the type strain 98-3 proved to be nearly identical, averaging fewer than 26 differences (polymorphisms) per pair (Table 1). Using *S. islandicus* as a reference, the average pairwise nucleotide divergence for shared genes of strains from Kamchatka vs those from North America is 1.11×10^{-2}, whereas the average for strains from within each of the two regions is 2.7×10^{-3} (Reno *et al.*, 2009). This latter divergence (24% of the total) thus estimates local diversity, and subtracting it from the total divergence yields the divergence of *S. islandicus* that can be attributed specifically to large geographical separation, that is, 8.4×10^{-3}.

The average nucleotide divergence of the *S. acidocaldarius* genomes is 1.12×10^{-5}. Assuming the same relative contribution of local diversity as in *S. islandicus* (24%) yields 8.5×10^{-6} as the divergence of *S. acidocaldarius* genomes that can be attributed to long-range geographical separation. When this value is normalized with respect to separation, the rate of *S. acidocaldarius* divergence with distance is 1300 times lower than that in *S. islandicus* (1.03×10^{-6} vs 1.37×10^{-3} per 1000 km).

We could not attribute the much stricter conservation of *S. acidocaldarius* genomes to a correspondingly higher replication accuracy or stronger purifying selection relative to *S. islandicus*.

In *S. acidocaldarius*, replication errors occur at the *pyrE* and *pyrF* genes at the rate of about 3×10^{-7} events per cell division (Grogan *et al.*, 2001), as compared with $0.4–16 \times 10^{-7}$ (median $= 4.9 \times 10^{-7}$) in strains of *S. islandicus* (Blount and Grogan, 2005). Furthermore, nearly all of the polymorphisms seen in *S. acidocaldarius* affect the presence or structure of protein, and thus should not be selectively neutral (about half are insertion/deletion events (Table 1), whereas most of the substitution mutations are nonsynonymous ($dS/dN = 3/18 = 0.1667$)). These data imply that *S. acidocaldarius* genomes are free to accumulate a variety of mutations, and, accordingly, that their conservation world-wide requires relatively rapid gene flow across large distances. We also considered the question of whether the three strains could be highly similar because they all diverged from a common laboratory culture. The available evidence argues against this possibility, but extensive efforts of multiple researchers may be required to define the pattern of natural divergence more rigorously (see Supplementary Information).

Mechanisms potentially producing an inter-specific difference in gene flow

The rapid gene flow in *S. acidocaldarius* indicated by our data contrasts sharply with the genetic isolation documented repeatedly for *S. islandicus* (Whitaker *et al.*, 2003; Grogan *et al.*, 2008; Reno *et al.*, 2009) and raises questions regarding what enforces this barrier in the latter species. Such a large difference between physiologically similar species suggests that biotic factors, rather than environmental ones, have critical roles, and the limited data regarding ecological parameters that we could obtain so far remain consistent with this. Properties that should facilitate the transfer of propagules among populations of a species include (i) large local populations, which would generate high fluxes of propagules from simple concentration (mass-action) effects, (ii) a large number of local populations, creating 'stepping-stone' networks with short distances between populations or (iii) large areas of habitat, which would capture

Table 1 Measures of *S. acidocaldarius* genome divergence

Strain pair	Separation (km)	Polymorphisms (total number)	Insertion/deletions (% of polymorphisms)	Nucleotide divergence
98-3 vs N8	8006	9	44.4	3.68×10^{-6}
N8 vs Ron12/I	8546	30	46.7	1.30×10^{-5}
Ron12/I vs 98-3	8075	38	44.7	1.66×10^{-5}
Average	8209	25.7	45.3	1.11×10^{-5}

Nucleotide divergences of *S. acidocaldarius* were calculated from the number of relatively small polymorphisms consistent with replication errors, and normalized by the number of shared nucleotides. Divergences of *S. islandicus* genomes were calculated as $1 - (bANI)$ (binary average nucleotide identity), using the bANI values reported by Reno *et al.* (2009). For comparison, *S. islandicus* strains drawn from Kamchatka vs the western United States represent an average separation of 6136 km and an average nucleotide divergence of 1.11×10^{-2} (Reno *et al.*, 2009).

dispersing propagules efficiently. We know of few field measurements that help define these parameters for different *Sulfolobus* species, but the available data generally predict that these parameters would favor propagule dispersal of *S. islandicus* (Supplementary Information).

One parameter that we could measure experimentally relates to the survival during dispersal. Because temperatures in the environmental matrix between geothermal islands preclude metabolism, growth and reproduction of *Sulfolobus*, survival between local populations would seem critical for gene flow. We found empirically that cell suspensions that were held in acidic solutions at 4 °C lost viability over a period of days to weeks, thus providing a basis for measuring relative durability. Under these conditions, *S. acidocaldarius* cells died markedly faster than cells of three geographically matched *S. islandicus* strains (Table 2). This demonstrates that the intrinsic durability of propagules does differ among *Sulfolobus* species, but, like the other parameters we considered, the difference observed would not explain the observed difference between species with respect to gene flow.

We are aware of one factor that would seem to explain the observed difference in gene flow between *S. islandicus* and *S. acidocaldarius* in the light of available data, namely, antagonism mediated by viruses and other mobile genetic elements; however, such antagonism would act only after a propagule is deposited into a new local population, and would not seem to depend directly on spatial distribution of habitat. Genetic elements capable of mediating severe and specific antagonisms can become established in microbial communities and create strong reciprocal selections that promote rapid coevolution of the element and host (Pal *et al.*, 2007; Paterson *et al.*, 2010). *S. islandicus* seems fully vulnerable to this, in terms of both its intrinsic ability to harbor diverse elements (Prangishvili *et al.*, 2001) and the relatively high density of its natural populations, which favors their proliferation; *S. acidocaldarius*, in contrast, fails both criteria. (For further discussion of these and other mechanisms, see Supplementary Information).

Conclusion

In the time since local populations of *S. islandicus* were first shown to diverge with distance (Whitaker *et al.*, 2003), several free-living mesophilic micro-organisms have been found to exhibit similar divergence, demonstrating that restricted gene flow does not require severely discontinuous habitat (Bass *et al.*, 2007; Kuehne *et al.*, 2007; Vos and Velicer, 2008). The evidence of global conservation of *S. acidocaldarius* genomes now argues, conversely, that severely discontinuous habitat is not sufficient *per se* to restrict gene flow of unicellular micro-organisms. Therefore, although the processes that restrict gene flow among local *S. islandicus* populations remain to be identified, they seem likely to have counterparts in other micro-organisms that exhibit allopatric diversification.

Table 2 Long-term survival of *Sulfolobus* cells at 4 °C

| Species[a] | Observed half-life of cell viability (days+s.d.)[b] | | | |
| | Growing cells[c] | | Resting cells[c] | |
	Buffer	HSF	Buffer	HSF
S. acidocaldarius	0.85 ± 0.04	0.18 ± 0.01	1.04 ± 0.13	0.23 ± 0.02
S. islandicus	5.17 ± 1.04	0.24 ± 0.11	4.66 ± 1.71	0.40 ± 0.11

Abbreviation: HSF, hot-spring fluid.
[a]One *S. islandicus* isolate was drawn arbitrarily from each of the three sites (western Italy, Kamchatka and YNP) close to those of the *S. acidocaldarius* sites (for underlying data, see Supplementary Information).
[b]Viability of cells suspended in sterile buffer or HSF was monitored by plating appropriate dilutions on dextrin-tryptone medium and counting the resulting colonies. Data for each species were averaged over the three strains, and the half-life of viability ($t_{\frac{1}{2}}$) was calculated from the slope of linear regression of log (fraction surviving) vs days for each strain in three independent trials (see Supplementary Information, Supplementary Figure S1); the resulting values were averaged for each species.
[c]Cultures of the six strains were grown in parallel in liquid dextrin-tryptone medium (Blount and Grogan, 2005); growing cells were withdrawn at a density of about 1.5×10^8 cells ml^{-1}, and resting cells after growth had ceased (1×10^9 cells ml^{-1}). Cells were harvested by centrifugation and suspended in either of two media: sterile dilution buffer 'Sdil', pH 3.5 (Grogan *et al.*, 2001), or acidic hot spring fluid (pH 2.7–3.3) drawn from acidic hot springs in YNP and clarified by centrifugation.

References

Bass D, Richards TA, Matthai L, Marsh V, Cavalier-Smith T. (2007). DNA evidence for global dispersal and probable endemicity of protozoa. *BMC Evol Biol* **7**: 162.

Blount ZD, Grogan DW. (2005). New insertion sequences of *Sulfolobus*: functional properties and implications for genome evolution in hyperthermophilic archaea. *Mol Microbiol* **55**: 312–325.

Brock TD, Brock KM, Belly RT, Weiss RL. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Microbiol* **84**: 54–68.

1616

Cohan FM. (2006). Towards a conceptual and operational union of bacterial systematics, ecology, and evolution. *Philos Trans R Soc Lond B Biol Sci* **361**: 1985–1996.

Fenchel T. (2003). Microbiology. Biogeography for bacteria. *Science* **301**: 925–926.

Finlay BJ. (2002). Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061–1063.

Fuchs T, Huber H, Teiner K, Burggraf S, Stetter KO. (1995). *Metallosphaera prunae*, sp. nov., a novel metal-mobilizing, thermoacidophilic archaeum, isolated from a uranium mine in Germany. *Syst Appl Microbiol* **18**: 560–566.

Grogan DW, Carver GT, Drake JW. (2001). Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. *Proc Natl Acad Sci USA* **98**: 7928–7933.

Grogan DW, Ozarzak MA, Bernander R. (2008). Variation in gene content among geographically diverse *Sulfolobus* isolates. *Environ Microbiol* **10**: 137–146.

Hahn MW, Pöckl M. (2005). Ecotypes of planktonic actinobacteria with identical 16S rRNA genes adapted to thermal niches in temperate, subtropical, and tropical freshwater habitats. *Appl Environ Microbiol* **71**: 766–773.

Huber H, Prangishvili D. (2006). Chapter 3. Sulfolobales. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K, Stackebrandt E (eds). *The Prokaryotes: A Handbook on the Biology of Bacteria*, 3rd edn. Springer Science+Business Media, LLC: New York, NY, pp 28–51.

Knittel K, Losekann T, Boetius A, Kort R, Amann R. (2005). Diversity and distribution of methanotrophic archaea at cold seeps. *Appl Environ Microbiol* **71**: 467–479.

Kuehne HA, Murphy HA, Francis CA, Sniegowski PD. (2007). Allopatric divergence, secondary contact, and genetic isolation in wild yeast populations. *Curr Biol* **17**: 407–411.

Kurosawa N, Sugai A, Fukuda I, Itoh T, Horiuchi T, Itoh YH. (1995). Characterization and identification of thermoacidophilic archaebacteria isolated in Japan. *J Appl Microbiol* **41**: 43–52.

Martiny JB, Bohannan BJ, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al.* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.

Pagaling E, Wang H, Venables M, Wallace A, Grant WD, Cowan DA *et al.* (2009). Microbial biogeography of six salt lakes in Inner Mongolia, China, and a salt lake in Argentina. *Appl Environ Microbiol* **75**: 5750–5760.

Pal C, Macia MD, Oliver A, Schachar I, Buckling A. (2007). Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450**: 1079–1081.

Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR *et al.* (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* **464**: 275–278.

Prangishvili D, Stedman K, Zillig W. (2001). Viruses of the extremely thermophilic archaeon *Sulfolobus*. *Trends Microbiol* **9**: 39–43.

Ramette A, Tiedje JM. (2007). Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microb Ecol* **53**: 197–207.

Reno ML, Held NL, Fields CJ, Burke PV, Whitaker RJ. (2009). Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc Natl Acad Sci USA* **106**: 8605–8610.

Vos M, Velicer GJ. (2008). Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. *Curr Biol* **18**: 386–391.

Whitaker RJ, Grogan DW, Taylor JW. (2003). Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**: 976–978.