DOI: 10.1534/genetics.105.049312

Recombination in Thermotoga: Implications for Species Concepts and Biogeography

Camilla L. Nesbø,[1] Marlena Dlutek and W. Ford Doolittle

Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 1X5, Canada and Genome Atlantic, Halifax, Nova Scotia B3J 1S5, Canada

Manuscript received August 8, 2005
Accepted for publication November 4, 2005

ABSTRACT

Here we characterize regions of the genomes of eight members of the hyperthermophilic genus Thermotoga. These bacteria differ from each other physiologically and by 3–20% in gene content and occupy physically distinct environments in widely disparate regions of the globe. Among the four different lineages (represented by nine different strains) that we compare, no two are closer than 96% in the average sequences of their genes. By most accepted recent definitions these are different "ecotypes" and different "species." And yet we find compelling evidence for recombination between them. We suggest that no single prokaryotic species concept can accommodate such uncoupling of ecotypic and genetic aspects of cohesion and diversity, and that without a single concept, the question of whether or not prokaryotic species might in general be cosmopolitan cannot be sensibly addressed. We can, however, recast biogeographical questions in terms of the distribution of genes and their alleles.

TWO of the most contentious questions in prokaryote biology are "What are prokaryotic species?" and "Are such species cosmopolitan in their distribution?" (Fenchel 2003; Finlay and Fenchel 2004). They are linked, and we address both here, in the context of a comparative genomics survey of Thermotoga isolates from around the world.

The first question, about the nature of prokaryotic species, has practical and theoretical aspects. (A recent review by Gevers *et al.* 2005 aptly summarizes problems and prospects.) There is a pressing practical need for more or less uniform criteria by which to decide when two isolates should be given the same specific name. Such criteria would constitute what is often called a *species definition*. DNA–DNA reassociation values of 70% or ribosomal RNA sequence similarity >97% currently are the most frequently used criteria (Rossello-Moran and Amann 2001). Konstantinidis and Tiedje (2005) very recently have proposed that an average nucleotide identity (ANI) value for protein-coding genes of >99% should be required—a highly stringent criterion motivated by an appreciation of the extent to which isolates very similar or even identical at the rRNA sequence level may differ in the content of their genes (Thompson *et al.* 2005). But however we choose to define species, if the groupings of organisms so designated are to be understood as "natural," we need something more. We need a *species concept*: a theory about biological processes affecting genetic cohesion within and divergence between species on the basis of which our definition can be rationalized.

The second contentious question, about prokaryotic distribution, is best encapsulated in the aphorism "everything is everywhere (but) the environment selects," generally attributed to L. G. M. Baas-Becking (Baas-Becking 1934) and made popular in particular by T. Fenchel and B. J. Finlay (Fenchel 2003; Finlay and Fenchel 2004). According to this way of thinking, microbial species—because of the tiny size of individuals and the enormous size of populations—are easily dispersed and difficult to eradicate. Favorable local conditions will of course be necessary for any microbe to become abundant, but there will always be a few nongrowing individuals within dispersal range to capitalize on such conditions, if they occur. Thus microbial species should, unlike those of bulkier and less numerous organisms such as mammals, be globally distributed (or cosmopolitan).

This "everything is everywhere" view might be supported by the finding of prokaryotes (and of protists) of very similar or identical appearance and physiology, wherever on earth suitable growth conditions occur. But it cannot really be tested as an hypothesis without some further agreement about just what are the "species" that are supposed to be everywhere. Biogeographers cannot avoid addressing this question, since they need both a species definition to devise experimental tests and a species concept to make biological sense of the results.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AJ810485, AJ811017–AJ811437, AJ872266–AJ872273, and AM072709–AM072729.

[1]*Corresponding author:* Department of Biochemistry and Molecular Biology, 5850 College St., Dalhousie University, Halifax B3H 1X5, Nova Scotia, Canada. E-mail: cnesbo@dal.ca

If neither can be achieved, then a different conceptual framework within which to address the interplay of divergence and dispersal in any geographical mapping of microbial diversity is called for.

In this article we present comparative genomics data obtained by suppressive subtractive hybridization and fosmid sequencing with thermophilic bacteria of the genus Thermotoga, including isolates from widely separated locations expected to exhibit "island biogeography"—if indeed any prokaryotes can be said to have island biogeography (Fenchel 2003). We show that both gene acquisition (lateral gene transfer) and recombination are important in the evolution of these bacteria and that recombination can occur between lineages that are sufficiently different phenotypically to be considered different species. We suggest that at least for these and likely for many other prokaryotes, no principled species definition or concept will be achievable. Biogeographical questions about distribution and dispersal will have to be addressed at the level of genes.

MATERIALS AND METHODS

Sources of DNA: DNA from *Thermotoga neapolitana* LA4, *T. neapolitana* LA10, *Thermotoga* sp. strain SG1, *Thermotoga* sp. strain KOL6, *Thermotoga* sp. strain RQ7, and *Thermotoga* sp. strain RQ2 were extracted from a cell mass donated by Karl O. Stetter using the protocol of Charbonnier and Forterre (1995). DNA from *T. petrophila* RKU1 and *T. naphthophila* RKU10 were gifts of Yoh Takahata.

Suppressive subtractive hybridization: Genomic subtraction of *T. petrophila* RKU-1 was carried out using the PCR-select bacterial genome subtraction kit (CLONTECH, Palo Alto, CA) as in Nesbø *et al.* (2002). Two different restriction enzymes, *Rsa*I and *Alu*I, were used, and two different hybridization temperatures were used for the *Rsa*I-restricted DNA (55° and 63°) while for the *Alu*I-restricted DNA 55° was used. PCR products obtained after suppressive subtractive hybridization (SSH) were cloned into TOPO 2.1 (Invitrogen, San Diego) and sequenced using ABI Big Dye-terminator chemistry on an ABI 377 instrument. The sequences were trimmed and analyzed using Sequencher 4.2.2. The sequences were used as probes to search the *T. maritima* MSB8 genome at The Institute for Genome Research (TIGR) Blast (http://tigrblast.tigr.org/cmr-blast/) and GenBank at NCBI–Blast (http://www.ncbi.nlm.nih.gov/BLAST/). The sequences were categorized and the library size was estimated as in Nesbø *et al.* (2002). For a more detalied description see supplemental material at http://www.genetics.org/supplemental/. The sequences have been submitted to the EMBL Nucleotide Sequence Database with accession nos. AJ810485 (plasmid pRKU1) and AJ811017–AJ811437.

PCR amplification of TM0938, TM0967, and TM1022: These genes were amplified using the following primers: TM0938.FOR (TCG CAC GAG TTT TYC AYG AR), TM0938.REV (CTT CCT CGC AGC RCA CCA NA), TM0967.for (CCA TCA CTC CAG ARA CNT GG), TM0967.rev (GTG GTG GAG ATR TTN GCR TG), TM1022.FOR (TTT CTT GAG GGA GGN AAY GA), and TM1022.REV (CAT TCC CCC TTCA RCC AYT C). In most cases PCR products were directly sequenced, but in a few cases the PCR product was cloned before sequencing. At least two independent PCR products were sequenced for each strain, except for *T. petrophila* RKU1 TM1022 where only one clone from a cloned PCR product gave the right sequence. The sequences have been submitted to the EMBL Nucleotide Sequence Database with accession nos. AM072709–AM072729.

Fosmid library construction and fosmid clone-end sequencing: Fosmid libraries were constructed using the CopyControl fosmid library production kit from Epicentre (Madison, WI) following the protocol of the manufacturer. Fosmid clones were miniprepped using either alkaline lysis prep on grown cultures with GeneMachine robotics (Genomic Solutions) or the R.E.A.L. Prep 96 plasmid kit (QIAGEN, Chatsworth, CA). End sequencing of miniprepped fosmid clones was done using the DYEnamic ET dye terminator kit (MegaBACE) and run on a MegaBACE 1000 (Amersham, Buckinghamshire, UK). Clones containing rRNA genes were identified by PCR screening the clones for the 23S rRNA intron (Nesbø and Doolittle 2003a) or by identifying the neighboring genes from end sequences. Subcloning of fosmids was done using the TOPO Shotgun subcloning kit (Invitrogen), and each fosmid was sequenced to >6× coverage (6.6–14.4× coverage). Low-quality regions and gaps were fixed by PCR. The fosmids were assembled using phredPhrap, ORFs were identified using Glimmer (Delcher *et al.* 1999), and tRNAs were identified using tRNAscan-SE (Lowe and Eddy 1997). For *Thermotoga* sp. strain RQ2 and *T. petrophila* RKU1, the contigs were assembled from two overlapping fosmid clones. The sequences have been submitted to the EMBL Nucleotide Sequence Database with accession nos. AJ872266–AJ872273.

Phylogenetic and recombination analyses: Sequences from *T. maritima* MSB8 and the unfinished *T. neapolitana* NS-E genome were obtained from the TIGR website (http://www.tigr.org/). Phylogenetic analyses of the ORFs were done using PAUP* 4.0b10 (Swofford 2001). The approximately unbiased (AU) tests were performed using CONSEL (Shimodaira and Hasegawa 2001; Shimodaira 2002). Recombination was detected using splits-decomposition analyses (Huson 1998; D. H. Huson and D. Bryant, unpublished results), the informative-sites tests (Worobey 2001), and several different analyses in the RDP version 2 β08 program package (Martin *et al.* 2005): GENECONV (Padidam *et al.* 1999), RDP, and MaxChi (Posada and Crandall 2001).

RESULTS

Thermotoga isolates differ substantially in phenotype and genotype: Members of the genus Thermotoga are obligately anaerobic heterotrophs, with optimal growth between 66° and 80° (Huber and Hannig 2005). Isolates employed in this study are listed in Table 1. They derive from geothermally heated sites around the globe and have been described as members of at least four species. *T. petrophila* RKU-1 and *T. naphthophila* RKU-10 were recovered from production fluids of the Kubiki Oil Reservoir, near the coast of the Sea of Japan (Takahata *et al.* 2001). Although showing ~99% rRNA sequence identity (for both 16S and 23S) with *T. maritima*, these isolates exhibit <30% DNA reassociation with an identified *T. maritima* strain, occupy different environments, grow at different minimal temperatures, and respond differently to several sugars, rifampicin, elemental sulfur, and thiosulfate. *T. neapolitana* has traditionally been differentiated from *T. maritima* by several phenotypic characteristics, including a lower G + C content (41% *vs.* 46%), and differ by at least 20% in ANI.

TABLE 1

Thermotoga strains used in this study

Strain	Habitat	Temperature[e]	Accession nos.[f]
T. maritima MSB8[a]	Geothermally heated seafloor, Vulcano, Italy	55°–90°(80)	
T. sp RQ2[a]	Geothermally heated seafloor, Ribeira Quente, the Azores	76°–82°	AJ872273, AM072713, AM072717, AM072728
T. naphthophila RKU-10[b]	Deep subterranean oil reservoir in Niigata, Japan	48°–86° (80)	AJ872268, AM072712, AM072724
T. petrophila RKU-1[b]	Deep subterranean oil reservoir in Niigata, Japan	47°–88° (80)	AJ810485, AJ811017–AJ811437, AJ872269, AM072716, AM072722, AM072723
T. neapolitana NS-E[c]	Shallow submarine hotspring, Naples, Italy	55°–90° (80)	
T. neapolitana LA4[d]	Shore of Lac Abbé, Djibouti, Africa	82°	AJ872266, AM072711, AM072719, AM072726
T. neapolitana LA 10[d]	Shore of Lac Abbé, Djibouti, Africa	87°	AJ872267, AM072710, AM072718, AM072727
T. sp RQ7[a]	Geothermally heated seafloor, Ribeira Quente, the Azores	76°–82°	AJ872270, AM072709, AM072720, AM072729
T. sp SG1[d]	Boiling beach during volcanic eruption, Sangeang Island, Indonesia	85°	AJ872271, AM072715, AM072721
T. sp KOL6[d]	Submarine hydrothermal system, Kolbeinsey ridge, north of Iceland	90°	AJ872272, AM072714, AM072725

[a] Described in Huber *et al.* (1986). Cell mass was a gift from Karl O. Stetter.
[b] Described in Takahata *et al.* (2001). DNA was a gift from Yoh Takahata.
[c] Described in Jannasch *et al.* (1988). Cell mass was a gift from Karl O. Stetter.
[d] Personal communication from Karl O. Stetter. For these strains, the temperature at the isolation site is given. Cell mass was a gift from Karl O. Stetter.
[e] Temperature refers to the growth temperature range for characterized strains and the numbers in parentheses refer to the optimal temperature. For strains in which only a single number is given, the number refers to the temperature at the sampling site (K. O. Stetter, personal communication).
[f] Accession numbers for sequences obtained in this study. Sequences from the *T. maritima* MSB8 genome and the *T. neapolitana* NS-E unfinished genome were obtained from the TIGR website (http://www.tigr.org/).

We prepared fosmid libraries for eight of the strains listed in Table 1 and used end sequencing to obtain a rough estimate of average nucleotide sequence similarity with the ninth, *T. maritima* MSB8 (see supplemental Table 1 at http://www.genetics.org/supplemental/). Similarity values estimated from 22 to 122 end sequences ranged from 75 to 95%. Thus none of the eight are likely to be conspecific with MSB8 by the criterion of Konstantinidis and Tiedje (2005), although *Thermotoga* sp. RQ2, *T. petrophila* RKU1, and *T. napthophila* RKU10 clearly cluster with it, to the exclusion of the designated *T. neapolitana* strains LA4 and LA10 and *Thermotoga* spp. SG1, RQ7, and KOL6. Among the 512 ORFs that were end sequenced, 40 had no match to MSB8 at the DNA level. These ORFs are listed in supplemental Table 2 at http://www.genetics.org/supplemental/: of the 21 with a functionally identified match in GenBank, six were sugar transporters and five might be predicted to be involved with polysaccharide degradation. Such a result is consistent with the differentiation of Thermotoga species by carbohydrate utilization patterns in Bergey's Manual (Garrity 2001); the fact that the *T. maritima* MSB8 genome boasts a record number of genes involved in sugar metabolism, which it regulates in complex ways (Nelson *et al.* 1999; Chhabra *et al.* 2003; Nguyen *et al.* 2004); and the rich substrate mix available in many known Thermotoga habitats, in particular, petroleum reservoirs.

Suppressive subtractive hybridization of *T. petrophila* RKU1 against *T. maritima* MSB8 produced a more comprehensive list of genes that differentiate these two organisms in particular. Of 593 clones obtained from *T. petrophila* RKU-1 by this method, 175 were derived from an 846-bp plasmid, which we were thus able to assemble (pRKU1, GenBank accession no. AJ10485) and found to be 99% identical in DNA sequence to pRQ7 from *Thermotoga* sp. strain RQ7 (Harriott *et al.* 1994; Yu and Noll 1997). Since the two strains are not closely related in the rRNA tree and show only 72% identity among shared protein-coding sequences (19,597 bp), we consider this to be evidence for a very recent transfer event. This plasmid has also been found in Thermotoga strains isolated from the Kuril Island region (Akimkina *et al.* 1999). An additional 131 nonplasmid clones had no DNA match in *T. maritima* MSB8, and of these, 96 did match a sequence in GenBank other than from *Thermotoga* sp. strain RQ2. Since several clones identified the same GenBank gene, the total number of *T. petrophila* RKU1-specific genes identified in this way was 57. From the frequency of independent recovery of

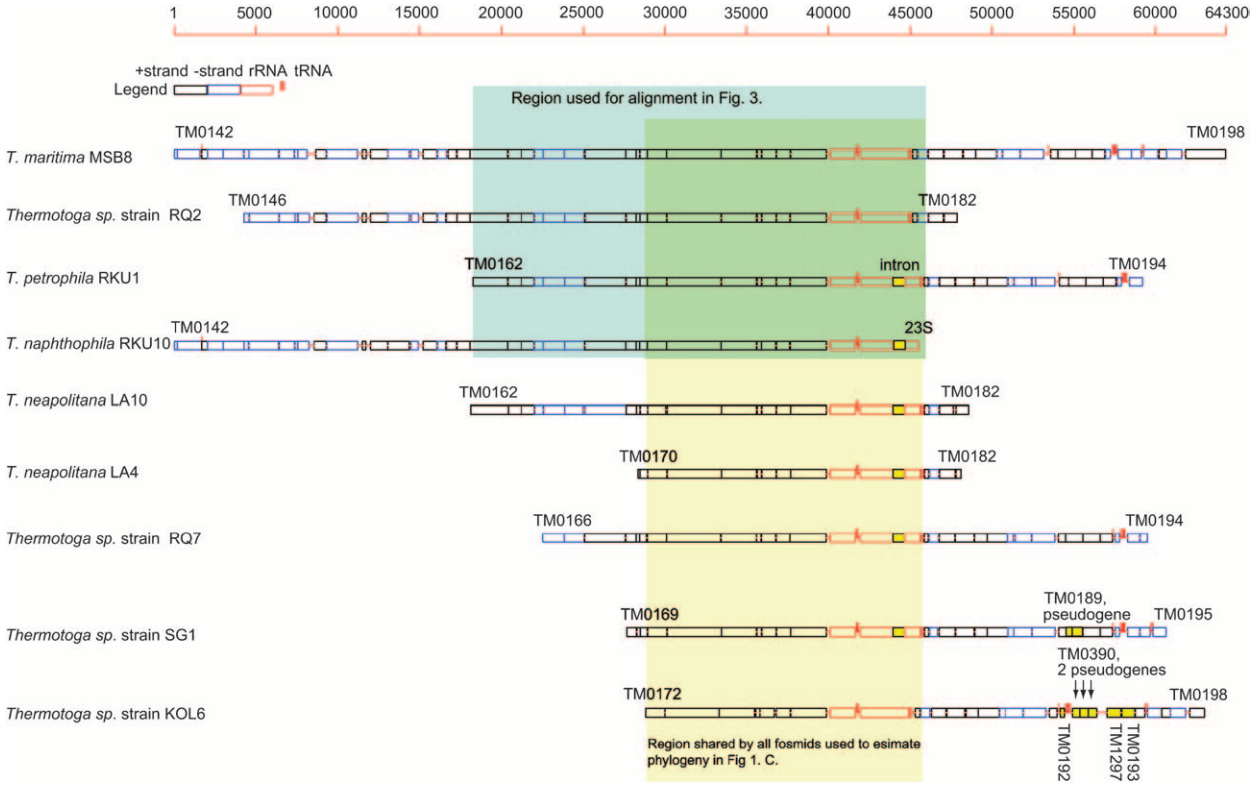

Figure 1.—Overview of the eight fosmids sequenced. The *T. maritima* MSB8 sequence is given as reference. ORFs that differ (in position or presence) from *T. maritima* MSB8 are given in yellow. We have indicated the sequence regions used in phylogenetic analysis in Figure 2B in yellow and for the alignment of variable sites in Figure 3 in blue. Genome organization in this chromosomal region showed only a few differences from the *T. maritima* MSB8 genome: (i) As observed previously (Nesbø and Doolittle 2003a), a group I intron is inserted into the 23S rRNA of some of the strains. (ii) In *Thermotoga* sp. strain RQ7 and *Thermotoga* sp. strain SG1, tRNA Ala is located between TM0191 and TM0192 instead of between TM0187 and TM0188 as in *T. maritima* MSB8. (iii) In *Thermotoga* sp. strain SG1 there is a stop codon in TM0189, turning this ORF into a pseudogene. (iv) In *T. petrophila* RKU1, TM0187 is shortened by 246 bp in the 3′-end due to a frameshift mutation (a deletion of 4 bp). (v) Three ORFs, TM0189–TM0191, are missing in *Thermotoga* sp. strain KOL6. These ORFs all encode components of an iron (III) ABC transporter. (vi) tRNA Ala is located between TM0188 and TM0192, not between TM0187 and TM0188, in *Thermotoga* sp. strain KOL6, and there are four extra ORFs positioned between TM0192 and TM0193, all of which have homologs in other positions in the *T. maritima* MSB8 genome. Three of these show similarity to TM0390 (a hypothetical protein), two of which are pseudogenes. The fourth ORF shows similarity to TM1297, a putative oxidoreductase. Phylogenetic analysis of TM0189–TM0191 showed that the Thermotoga genes cluster with high support as a sister group of *Pyrococcus* spp. (96–100 bootstrap support in maximum-likelihood protein distance trees), while TM0188 clusters as a sister to *Archaeblobus fuldigus* (100% bootstrap support). Only TM0192 has significant homologs in bacteria (Firmicutes and Thermus–Deinococcus). This suggests that TM0189–TM0191 were transferred into the Thermotoga lineage from Pyrococcus, and the most parsimonious scenario would be that this happened once after the separation of *Thermotoga* sp. strain KOL6 and the other Thermotoga strains.

clones from the same genes, we estimated that 7–9% of the genes in the genome of *T. petrophila* RKU1 should have no match in that of *T. maritima* MSB8. Supplemental Table 3 at http://www.genetics.org/supplemental/ lists those with functions that can be assigned by comparison to sequences in GenBank. Again, there is a clear preponderance of genes involved with the uptake and metabolism of carbohydrates.

Prokaryotic species definitions variously emphasize sequence similarity of individual marker genes (rRNA), co-occurrence and/or similarity of multiple genes (DNA–DNA reassociation), overall sequence similarity (ANI), or, more traditionally, gene-based phenotypic differences expected to underlie ecological ("ecotypic") differentiation (Rossello-Moran and Amann 2001). By all criteria but the first (rRNA sequence), only *Thermotoga* sp. RQ2 might be considered conspecific with *T. maritima* MSB8, and even in this case our previous suppressive subtractive hybridization studies with RQ2 (Nesbø *et al.* 2002) identified many genes that should at the least differentiate these as "ecotypes" (see below). Nevertheless, for convenience, and in agreement with the ribosomal RNA gene analyses, we refer to these two isolates as " the *maritima* strains" and to *T. petrophila* RKU1 and *T. naphthophila* RKU10 as "the Japanese strains."

Recombination can be detected in fosmids bearing the ribosomal RNA gene cluster: Fosmids containing rRNA genes were subcloned and fully sequenced for each of the eight libraries (Figure 1: gene designations TM0142–TM0198 are based on the *T. maritima* MSB8 genome). With only a few exceptions (see Figure 1 legend)

the genome organization in this chromosomal region was identical to that of *T. maritima* MSB8.

Fosmid sequence alignments were used to estimate the level of recombination between strains. The tree estimated from the 15,787 bp shared by all the fosmids spanning TM0172 to 23S rRNA (yellow rectangle in Figure 1) is shown in Figure 2B and differs from that found with the rRNA gene clusters (Figure 2A), since *T. petrophila* RKU1 clusters with *T. maritima* MSB8 and *Thermotoga* sp. strain RQ2, and not with *T. naphtophila* RKU10. Visual inspection of this alignment revealed at least three large segments (1640, 4119, and 3262 bp) where the similarity patterns among the *maritima* and Japanese strains differed from what would be expected from the rRNA tree. AU tests of the trees estimated from these regions separately showed that they were significantly different from the rRNA tree ($P = 0.02$, <0.001, and <0.001), while the phylogenetic tree excluding these fragments had the same topology as the rRNA tree, suggesting a conflicting signal due to recombination.

Recombination was further examined by phylogenetic analysis of single ORFs, splits-decomposition analyses, several recombination detection programs (see MATERIALS AND METHODS in supplemental material at http://www.genetics.org/supplemental/), calculation of the informative sites index (ISI), and visual inspection of the alignments of variable sites. The results are detailed in supplemental Tables 4 and 5 at http://www.genetics.org/supplemental/. When correcting for multiple tests (significance level at $P = 0.002$, 0.05/25 tests when including additional tests mentioned below), only one ORF (TM0178, $P = 0.002$) showed a significantly different topology from the rRNA genes. However, for five additional ORFs, both bootstrap analyses and AU tests supported an alternative phylogeny even though the AU test was not significant (supplemental Table 4 at http://www.genetics.org/supplemental/). Phylogenetic trees of TM0178 as well as a splits-decomposition graph of one gene are shown in Figure 2, C–D. The results from recombination detection programs in supplemental Table 5 at http://www.genetics.org/supplemental/ supported these findings and suggest a complex pattern with numerous recombination events detected in regions corresponding well with the phylogenetic analyses. These analyses also detected recombination events in regions where phylogenetic methods could not be used, for instance, where we have sequences from only three strains: TM0146–TM0162 (supplemental Table 5 at http://www.genetics.org/supplemental/).

Phylogenetic methods rely on altered topologies as indicators of recombination and detect only a subset of recombination events; in particular, they do not detect the presumably most frequent kind, involving closely related taxa for which we have no genomic information. All the likely recombination events detected using the phylogenetic methods involve the "*maritima* strains" (*T. maritima* MSB8 and *Thermotoga* sp. strain RQ2) and the "Japanese strains" (*T. petrophila* RKU1 and *T. naphthophila* RKU10). Thus we examined the extended region of overlap between these four fosmids in greater detail. The alignment of the 1695 variable sites extracted from the 26,517-bp alignment of these four strains covering TM0162–23S rRNA is shown in Figure 3. The ISI (Worobey 2001) calculated from the alignment covering TM0162 up to the 23S rRNA including only these fosmids was 0.66 ($P = 0.001$). In comparison, the ISI estimated from a *Helicobacter pylori* alignment was 0.85 and from a DEN-1 virus data set was 0.49: the level of recombination between these Thermotoga lineages was thus comparable to that in lineages considered frequently recombining.

Regions where four or more consecutive sites show a similarity pattern that differs in a consistent way from that expected from the rRNA phylogeny (for instance, where *Thermotoga* sp. strain RQ2 is more similar to *T. petrophila* RKU1 than to *T. maritima* MSB8) are boxed (and hereafter referred to as "potentially recombinant fragments"). Excluding these boxed regions in Figure 3 from the alignment resulted in a tree with the same topology as that we observed with rRNA. The uncorrected nucleotide sequence distance between the "*maritima* strains" (again excluding boxed regions) was 0.01826 and between the "Japanese strains" was 0.00796, while the between-group distances were ~0.044 (0.04432, 0.04510, 0.04385, 0.04432).

Thirty-four potentially recombinant fragments were identified, ranging in size from 10 to 4131 bp (average 488 bp, SD = 905). One large potential recombination event involves 3276 bp (spanning part of TM0177, including TM0178 and extending at least up to position 217 in the 16S rRNA gene). Over this region *Thermotoga* sp. strain RQ2 and *T. petrophila* RKU1 are identical except for two positions (in the 16S gene), while *T. maritima* MSB8 and *T. naphtophila* RKU10 differ from both by 53 and 236 positions, respectively—the ratios of nonsynonymous-to-synomymous site mutations in these regions (0.43 and 0.32) being consistent with purifying selection. Explaining this without invoking recombination requires that *Thermotoga* sp. strain RQ2 and *T. petrophila* RKU1 have independently experienced 50-fold reductions in the rate of mutation (2 *vs.* an expected 118 substitutions) in this region and, indeed, an absolute suppression of synonymous mutations in protein-coding sequences since diverging from the common ancestor of all four strains (supplemental Table 4 at http://www.genetics.org/supplemental/ and Figure 3). A similar argument can be made for an additional large fragment, not shown in Figure 3, which involves >5 kb of sequence covering TM0156–TM0161. In this region *T. naphthophila* RKU10 is almost identical (99% identity) in sequence to *T. maritima* MSB8 and *Thermotoga* sp. strain RQ2, showing only 33 and 11 differences between them, when ~200 would be expected. This region was also identified by the

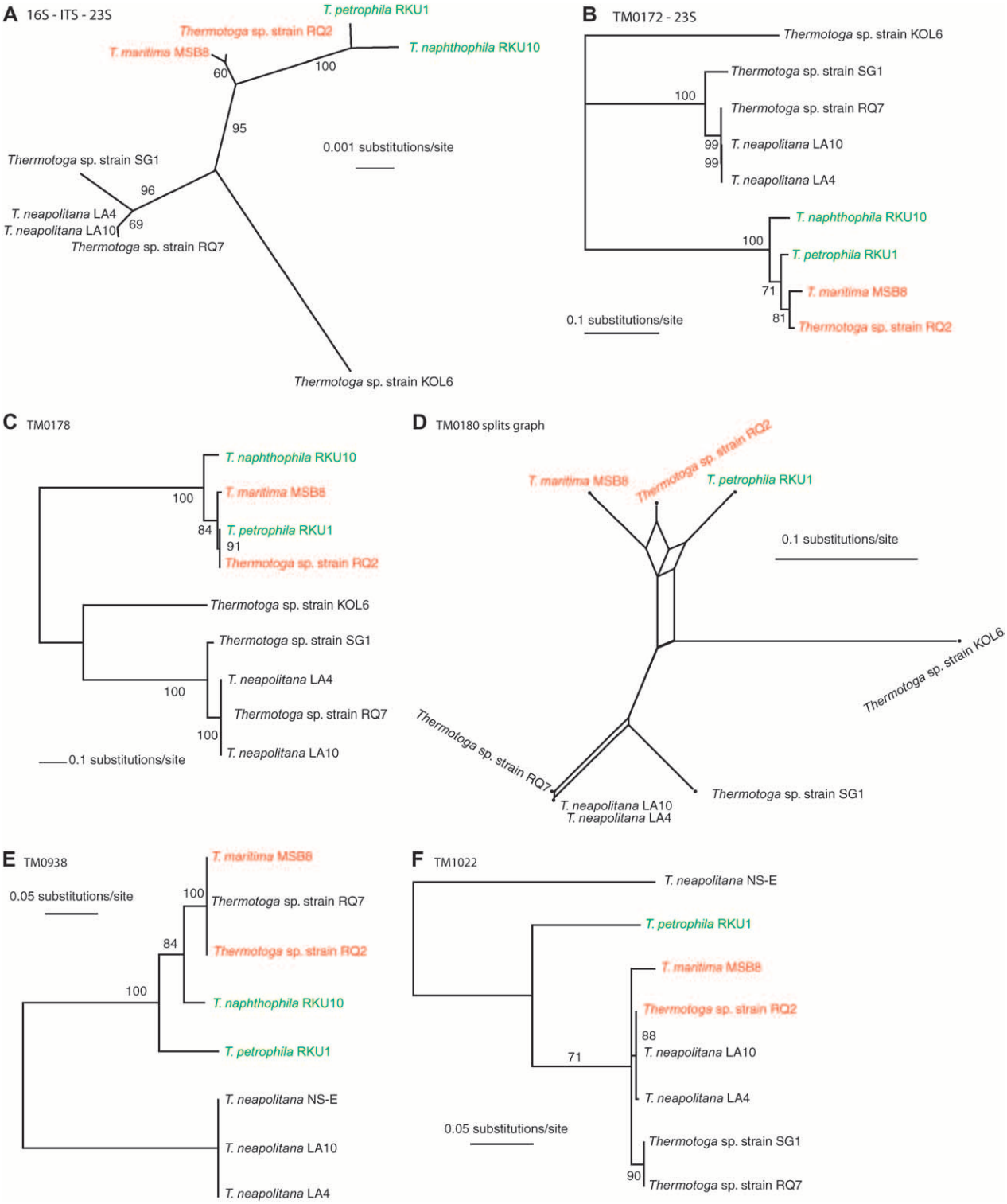

FIGURE 2.—(A) Maximum-likelihood tree of the 4645 bp corresponding to the 16S, ITS, and 23S from the strains from which we have fosmid sequences (B) Maximum-likelihood tree of the 15,787 bp shared by all the fosmid clones. (C) Maximum-likelihood tree of TM0178. (D) Splits-decomposition graph of TM0180. Hamming distances were used and the fit was 96.9 (D. H. HUSON and D. BRYANT, unpublished results). (E) Maximum-likelihood tree of TM0938. (F) Maximum-likelihood tree of TM1022. For the maximum-likelihood trees in A–C and E and F, a GTR + G + I model was used. Numbers on branches indicate percentage occurrence in 100 bootstrap replicates.

FIGURE 3.—The 1695 variable sites extracted from the alignment (26,517 bp) of the fosmid clone sequences from *Thermotoga* sp. strain RQ2, *T. petrophila* RKU1, and *T. naphthophila* RKU10 to the *T. maritima* MSB8 sequence. Regions where four or more consecutive bases show a similarity pattern that disagrees with the rRNA phylogeny in the same way are indicated by boxes. For potential recombinant regions consisting of >10 variable sites (which contained phylogenetically informative sites), we constructed four-taxon maximum-likelihood trees. Red boxes indicate fragments where the topology of the four-taxon tree differs from the rRNA phylogeny, while a blue box indicates regions where the fragment gives the same topology as the rRNA tree and we cannot distinguish between recombination among the lineages in the alignment and a recombination with an unknown Thermotoga strain. The significance of four-taxon trees that differed from the rRNA topology was tested using bootstrap analysis (1000 replicates with 10 random additions per replicate) and the AU test on the three possible topologies: rRNA [(TM, RQ2) (RKU1, RKU10)], Alt1 [(RKU1, RQ2) (TM, RKU10)], and Alt2 [(RQ2, RKU10) (TM, RQ2)]. The results from the AU test and the bootstrap analysis are shown on the side. We also included as red boxes shorter-than-10-variable-sites regions where the monophyly of the two 16S-rRNA clades was broken (box 1, box 4, box 10, box 12, and box 25). The distribution of the variable sites in the full 26,517-bp alignment is given in the insert at the bottom.

recombination detection software, as detailed in supplemental Table 5 at http://www.genetics.org/supplemental/. Both the analyses in supplemental Table 5 and the alignment in Figure 3 reveal that most of the likely recombination events do not correspond with boundaries of ORFs—nor would this be expected given the extensive synteny and conservative nature (mostly synonymous substitutions) of differences observed. Thus several competing phylogenetic signals may be present in each ORF, and it may be unsurprising that only one whole-ORF tree (that for TM0178) showed significant difference from the rRNA tree when multiple tests were considered.

To test if any of the recombination events that we detected might have occurred between the lineages that we have studied here (or their close relatives), rather than with other Thermotogales outside the clade represented by these eight fosmida, and for which we lack data, we made four-taxon trees for 16 potentially recombinant regions with >10 variable sites (which also

contain phylogenetically informative sites). If the four-taxon topology is different from the four-taxon phylogeny predicted by the 16S rRNA, then the most likely explanation is a recombination event involving two of the lineages represented here, or their close relatives. Of these 16 regions, 11 showed a topology different from the 16S rRNA phylogeny. These regions are in red boxes in Figure 3. In the red boxes we also included shorter-than-10-variable-sites regions where the monophyly of the two 16S-rRNA clades was broken (for instance, box 1 in Figure 3). Six of the 11 regions with alternative phylogenies gave trees that were supported by >90% bootstrap support (1000 bootstrap replicates) and three regions, including the 3276-bp stretch discussed above, showed significantly different topologies in AU tests when correcting for multiple tests (significance level at $P = 0.002$) (Figure 3). For many of the fragments where the AU test was not significant, additional recombination events involving these and/or other Thermotoga lineages may have occurred, obscuring the signal (see, for instance, the region covered by box 26 and box 29). It is notable that most of these fragments support the alternative phylogeny where Thermotoga strain RQ2 and *T. petrophila* RKU1 cluster together, and in many cases these strains have almost identical sequences, suggesting several recent recombination events between them.

With the fosmids shown in Figure 1, no recombination events were detected involving the four *T. neapolitana*-related strains (*T. neapolitana* LA4 and LA10 and *Thermotoga* sp. RQ7 and SG-1) by the phylogenetic method, and fewer were found by recombination detection analysis (supplemental Tables 4 and 5 at http://www.genetics.org/supplemental/). This may be due to the high similarity among three of the strains. For instance, only 53 mutations separate *Thermotoga* sp. strain RQ7 from *T. neapolitana* LA10 within the 26,001-bp shared region. A preliminary analysis of the unfinished genome of *T. neapolitana* NS-E (at http://tigrblast.tigr.org/ufmg/), a strain isolated from a submarine thermal vent near Naples, revealed similar high levels of sequence conservation (RQ7–NS-E, 98–99% identity; LA10–NS-E, 100% identity). However, such strong similarity makes it reasonable to assume that when clustered sequence differences are found among these strains, they have arisen through recombination, not point mutation. For instance, 5 of the above-mentioned 53 differences between *Thermotoga* sp. strain RQ7 and *T. neapolitana* LA10 occur within 16 bp in TM180, while 30 of the remaining 46 are found in TM0166, TM0167, and TM0168 (covering <20% of the alignment).

Recombination between *maritima* and *neapolitana* clusters: In the rRNA gene-containing fosmids (Figure 1) no recombination events were detected between the *T. neapolitana* or the KOL6 lineage and the *maritima* strains or the two Japanese strains. However, when we searched the unfinished *T. neapolitana* NS-E genome at the TIGR website for an integrase gene (in an independent search for integrons in different genomes), we found that it was identical in sequence to the *T. maritima* MSB8 copy. The flanking sequences around this gene revealed that this gene is part of an ~88-kb region where these two genomes have almost identical sequences (99% DNA identity spanning from TM0938 to TM1022, these terminal genes being chimeric). Analysis of a total of 84 genes sampled from three different comparably large (88 kb) regions randomly selected from the *T. maritima* MSB8 genome gave an average DNA identity of 77.8% to *T. neapolitana* NS-E genes in blastn searches against the unfinished genome. A *t*-test showed that these regions are significantly different from the integrase gene-bearing region ($P =$ 1.73e-62), suggesting that the high similarity in the TM0938–TM1022 region is due to a recent transfer or recombination event between these lineages. We therefore made primers for three genes in this region—TM0938, TM0967, and TM1022—to see if we would detect additional recombination events involving the *T. neapolitana* and the *T. maritima* lineage. TM0938 and TM1022 genes could not be amplified from all the strains that we tested (Figure 2, E and F), and the distribution pattern varied between the genes. The d_s/d_n values for the three genes all showed evidence of purifying selection (average d_s/d_n for TM0938, 11.98; TM0967, 12.92; and TM1022, 12.96). This region was also reported as one of the variable *Thermotoga* sp. strain RQ2 regions investigated in Mongodin *et al.* (2005), where, for instance, TM0969 in strain RQ2 is much larger than in *T. maritima* MSB8, with a unique N terminus. Moreover, in a previous study (Nesbø and Doolittle 2003b), we found that two archaeal-type Mut-S homologs have been inserted into the RQ2 genome between TM0962 and TM0963. Hence, this region likely represents a highly variable part of the Thermotoga genomes.

For TM0938, a conserved hypothetical protein, we obtained PCR products from all strains except *Thermotoga* sp. SG1 and *Thermotoga* sp. KOL6. The final 89 bp of the 567 amplified bases (extending into the 88-kb recombinant region) are identical between *T. maritima* MSB8 and *T. neapolitana* NS-E and represent the 5′-end point of the recombination. The phylogenetic tree revealed an additional recombination event involving *Thermotoga* sp. strain RQ7 [P (rRNA topology) < 0.001], which was identical in sequence to *T. maritima* MSB8 and *Thermotoga* sp. strain RQ2.

The integrase gene TM0967 was amplified from all strains. The phylogeny agrees with the rRNA phylogeny with the exception that *T. maritima* MSB8 clusters within the *T. neapolitana* clade [not shown; AU test: P (rRNA topology) < 0.001], which would suggest that the *neapolitana* lineage was the donor in this recombination event. However, the distance (~5%) between the *T. maritima* MSB8–*neapolitana* sequences and the *Thermotoga* sp.

RQ2 is more similar to what we would expect between the *T. maritima* lineage and the Japanese strains. This is probably not due to this gene being particularly highly conserved since *Thermotoga* sp. KOL6 shows expected levels of divergence (37% compared to 33–35% for the reverse gyrase gene TM0173). Hence, the *maritima* lineage likely donated this sequence to the *neapolitana* lineage, and there have been additional recombination events involving *Thermotoga* sp. RQ2 and the *T. petrophila* RKU1–*T. naphthophila* RKU10 lineage.

We could amplify TM1022 (an esterase) from all the strains except *Thermotoga* sp. KOL6 and *T. naphthophila* RKU10, although the sequence obtained from *T. petrophila* RKU1 was ~130 bp shorter in the 3′-end than the other sequences (and was therefore not included in the AU test). TM1022 contains the 3′-end point of the recombination, and *T. maritima* MSB8 and *T. neapolitana* NS-E are identical in sequence for the first 117 bp of the amplified fragment. The phylogenetic tree again suggests additional recombination events involving the *maritima* lineage and the *neapolitana* lineage, as all the *neapolitana* strains except for NS-E cluster with *T. maritima* MSB8 and *Thermotoga* sp. RQ2 [P (rRNA topology) < 0.001].

Taken together, the 88-kb fragment was probably donated to the *T. neapolitana* lineage from the *T. maritima* MSB8 lineage. The introduction of this large high-similarity region has thus created the opportunity for additional recombination events between the *maritima* strains and the *neapolitana* strains.

DISCUSSION

Two *species concept*s currently attract the attention of prokaryotic microbiologists. They focus on distinct processes that might maintain genotypic similarity within species, while promoting differentiation between species. The "ecotype concept," championed most enthusiastically by F. W. Cohan, while accepting a possible role for recombination, sees periodic selection as the force of cohesion (Cohan 2002). "An ecotype," Cohan writes (p. 467), "is a set of strains using the same or similar ecological resources, such that an adaptive mutant from within the ecotype out-competes to extinction all other strains of the same ecotype; an adaptive mutant does not, however, drive to extinction strains from other ecotypes." Thus different ecotypes, using different resources, will independently fix different fitter types, and inevitably diverge. This kind of ecologically constrained speciation is perhaps all that is available for clones of asexual organisms, which prokaryotes were long thought to be.

The biological species concept, on the other hand, attributes genotypic cohesion over the long term to homologous recombination within species, while genetic barriers to exchange between species participate in their divergence (Dykhuizen and Green 1991; Hanage *et al.* 2005). That prokaryotic species might thus be analogous to those of animals, for which this concept was initially developed, is a discovery of only the last 10 or 15 years. It stems from renewed attention to modeling the population genetics of prokaryotes and new techniques (in particular, multi-locus sequence typing) that faciiltate the detection and quantitation of recombination (Spratt *et al.* 2001; Papke *et al.* 2004; Vinuesa *et al.* 2005).

No realistic ecotype model would deny that between-lineage recombination can be the source of fitter-type mutations (Cohan 2002), and any realistic biological species model would accept that most species exhibit some degree of clonality. Nor would sensible versions of either model ignore the importance of useful genes transferred laterally from phylogenetically distant sources; however, these might then be fixed within populations. Thus the principal difference between models might be seen to lie in what they take as the barrier to exchange between ecotypes (or species)—ecological specialization in the ecotype model and limitations on recombination in the biological species model. These limitations might reflect specificities of agents of exchange such as phages, ecological/geographical isolation, or dramatic reduction in the frequency of homologous recombination due to sequence divergence.

Most work supporting recombination as a major force of cohesion has addressed human pathogens (Feil *et al.* 2000; Feil and Spratt 2001; Feil and Enright 2004). In several cases it is possible to infer that species defined by the ability to share genes through recombination are indeed biogeographically cosmopolitan. Their shared gene pools are global, and the mobility of their hosts helps to keep these pools stirred. But there may be no barriers to exchange that clearly demark species boundaries, while extensive ecological differentiation through acquisition of novel laterally transferred genes (producing novel "ecotypes") does occur between freely recombining lineages. There may be no reliable coupling of ecotype and genotype that would lend support to a meaningful species concept, and thus no possibility of a principled species definition. As Hanage *et al.* (2005, p. 6) observe for the genotypic clusters ("fuzzy species") of Neisseria, "the point at which such a group is described as a species is more a matter of human interest and attention than any intrinsic evolutionary process."

A similar expression of skepticism about "species" seems justified by the results described here. As nonpathogens presumably indifferent to the movements of human populations and restricted in their growth to hot anaerobic habitats (mostly >50°), Thermotoga lineages would be expected to exhibit "island biogeography." Local adaptations and genomic divergence through periodic selection should produce ecologically differentiated populations incapable of "driving to extinction" their cousins at distant sites—should they somehow be transported there—while accumulated sequence differences should reduce the frequency of recombination between such populations. No doubt

sufficiently prolonged separation could achieve this end. But our data show that strains with enough accumulated differences in phenotype, gene content, and gene sequences to be considered different ecotypes—and indeed different species by almost all criteria—still give evidence of genetic information exchange through recombination. We cannot claim to have shown, without exhaustive documentation of the diversity of Thermotoga lineages at each site, that such genetic exchanges have occurred across vast distances of inhospitably cold water from a petroleum deposit on the coast of the Sea of Japan to a beach on the Italian coast: it could be that all strains occur at all sites. But we may safely argue that ecotypic and biological species concepts will produce different understandings of the relationships between these organisms and that no single species definition applied to them could be taken as "natural."

Our data point out, furthermore, that species concepts based on the ability to share information through homologous recombination must take on board a still-greater dimension of complexity. We found little evidence for recombination between *T. neapolitana* or its relatives and the cluster of recombining *maritima* and Japanese strains in the region of their ribosomal RNA gene clusters. Quite possibly the decrease in DNA sequence similarity (93–95% to 75–80%) enormously reduces its frequency. However, an 88-kb stretch of the *T. neapolitana* genome, which was likely received as a lateral transfer from a relative of the *maritma* strains, can undergo further recombination with genomes of that lineage. As shown in Figure 4, different parts of a genome can belong to different biological species, if our species concept is based on the ability to share information through homologous recombination.

Ecological and genetic forces and processes that give rise to the diversity of microbial phenotypes and genotypes (the magnitude of which we have only just begun to appreciate) are many and complex, but they can in principle be understood at their own levels. Higher-order concepts like "species" may do no real work in helping us to understand these forces and processes, in spite of the fact that we need them in the practice of microbiology and may think we need them in the documentation of diversity. We suggest that the question of whether or not microbial species are cosmopolitan in their distribution cannot be answered in any general way because microbes (or least many prokaryotes) do not form natural clusters to which the term "species" can be universally and sensibly applied. We can, however, commit ourselves to a study of biogeography on the basis of the global distribution of genes and their alleles and their patterns of divergence and dispersal. This should be a central guiding principle for the new science of metagenomics.

We thank E. Bapteste for critical reading of the manuscript as well as J. Palmer and D. Posada for helpful comments on an earlier version of this manuscript. We also want to thank Yoh Takahata and Karl O. Stetter for donating Thermotoga DNA and cell masses. Preliminary sequence data from *T. neapolitana* NS-E were obtained from The Institute for Genomic Research website at http://www.tigr.org. Sequencing of *T. neapolitana* NS-E was accomplished with support from the Department of Energy. This work was supported by funds from the Canadian Institutes for Health Research (MOP 4467) and Genome Canada (Genome Atlantic).

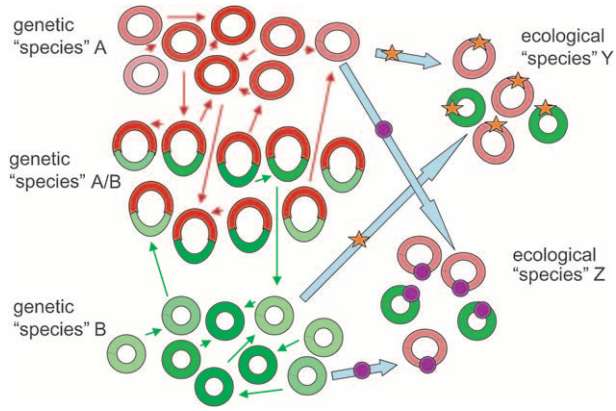

FIGURE 4.—A simplified model representing genetic and ecological "speciation" within a group of closely related prokaryotes (perhaps designated as species in a bacterial genus). "Genetic species A" comprises organisms whose genomes are similar enough in sequence that recombination (small arrows) can occur frequently for all genes, albeit at varying rates. "Genetic species B" is defined similarly, but B genomes are sufficiently different from A genomes in sequence that recombination between them is infrequent. Although infrequent, A-to-B recombination (or homology-independent gene transfer events) does occur, producing chimeric genomes whose different regions behave as if they are parts of different gene pools. Such chimerism might be the rule, not the exception, in many prokaryotic groups, such as Thermotogales, and precludes precise delimitation of genetic species. Lateral transfer of genes (or more likely suites of genes), represented by yellow stars and purple circles, allows some members of genetic species A and B to act as a single "ecotype" or ecological species Y and other transfers to create ecological species Z.

LITERATURE CITED

AKIMKINA, T., P. IVANOV, S. KOSTROV, T. SOKOLOVA, E. BONCH-OSMOLOVSKAYA *et al.*, 1999 A highly conserved plasmid from the extreme thermophile *Thermotoga maritima* MC24 is a member of a family of plasmids distributed worldwide. Plasmid **42:** 236–240.

BAAS-BECKING, L., 1934 *Geobiologie of Inleiding Tot de Milieukunde.* W. P. van Stockum & Zoon N. V., The Hague, The Netherlands.

CHARBONNIER, F., and P. FORTERRE, 1995 Protocol 12: purification of plasmids from thermophilic and hyperthermophilic Archaea, pp. 87–90 in *Archaea: A Laboratory Manual—Thermophiles,* edited by F. T. ROBB and A. R. PLACE. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

CHHABRA, S. R., K. R. SHOCKLEY, S. B. CONNERS, K. L. SCOTT, R. D. WOLFINGER *et al.*, 2003 Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima.* J. Biol. Chem. **278:** 7540–7552.

COHAN, F. M., 2002 What are bacterial species? Annu. Rev. Microbiol. **56:** 457–487.

Delcher, A. L., D. Harmon, S. Kasif, O. White and S. L. Salzberg, 1999 Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:** 4636–4641.

Dykhuizen, D. E., and L. Green, 1991 Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. **173:** 7257–7268.

Feil, E. J., and M. C. Enright, 2004 Analyses of clonality and the evolution of bacterial pathogens. Curr. Opin. Microbiol. **7:** 308–313.

Feil, E. J., and B. G. Spratt, 2001 Recombination and the population structures of bacterial pathogens. Annu. Rev. Microbiol. **55:** 561–590.

Feil, E. J., M. C. Enright and B. G. Spratt, 2000 Estimating the relative contributions of mutation and recombination to clonal diversification: a comparison between *Neisseria meningitidis* and *Streptococcus pneumoniae*. Res. Microbiol. **151:** 465–469.

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

Finlay, B. J., and T. Fenchel, 2004 Cosmopolitan metapopulations of free-living microbial eukaryotes. Protist **155:** 237–244.

Garrity, G., 2001 *Bergey's Manual of Systematic Bacteriology*, Ed. 2. Springer-Verlag, New York.

Gevers, D., F. M. Cohan, J. G. Lawrence, B. G. Spratt, T. Coenye *et al.*, 2005 Opinion: re-evaluating prokaryotic species. Nat. Rev. Microbiol. **3:** 733–739.

Hanage, W. P., C. Fraser and B. G. Spratt, 2005 Fuzzy species among recombinogenic bacteria. BMC Biol. **3:** 6.

Harriott, O. T., R. Huber, K. O. Stetter, P. W. Betts and K. M. Noll, 1994 A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. J. Bacteriol. **176:** 2759–2762.

Huber, R., and M. Hannig, 2005 Thermotogales, in *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, Ed. 3, Release 3.14, edited by M. Dworkin. Springer-Verlag, New York (http://link.springer-ny.com/link/service/books/10125/).

Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese *et al.*, 1986 *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol. **144:** 324–333.

Huson, D. H., 1998 SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics **14:** 68–73.

Jannasch, H., R. Huber, S. Belkin and K. Stetter, 1988 *Thermotoga neapolitana* sp. nov. of the extremely thermophilic, eubacterial genus Thermotoga. Arch. Microbiol. **150:** 103–104.

Konstantinidis, K. T., and J. M. Tiedje, 2005 Genomic insights that advance the species definition for prokaryotes. Proc. Natl. Acad. Sci. USA **102:** 2567–2572.

Lowe, T. M., and S. R. Eddy, 1997 tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:** 955–964.

Martin, D. P., C. Williamson and D. Posada, 2005 RDP2: recombination detection and analysis from sequence alignments. Bioinformatics **21:** 260–262.

Mongodin, E. F., I. R. Hance, R. T. Deboy, S. R. Gill, S. Daugherty *et al.*, 2005 Gene transfer and genome plasticity in *Thermotoga maritima*, a model hyperthermophilic species. J. Bacteriol. **187:** 4935–4944.

Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson *et al.*, 1999 Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature **399:** 323–329.

Nesbø, C. L., and W. F. Doolittle, 2003a Active self-splicing group I introns in the 23S rRNA genes of hyperthermophilic bacteria, derived from introns in eukaryotic organelles. Proc. Natl. Acad. Sci. USA **100:** 10806–10811.

Nesbø, C. L., and W. F. Doolittle, 2003b Targeting clusters of transferred genes in *Thermotoga maritima*. Environ. Microbiol. **5:** 1144–1154.

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.

Nguyen, T. N., A. D. Ejaz, M. A. Brancieri, A. M. Mikula, K. E. Nelson *et al.*, 2004 Whole-genome expression profiling of *Thermotoga maritima* in response to growth on sugars in a chemostat. J. Bacteriol. **186:** 4824–4828.

Padidam, M., S. Sawyer and C. M. Fauquet, 1999 Possible emergence of new geminiviruses by frequent recombination. Virology **265:** 218–225.

Papke, R. T., J. E. Koenig, F. Rodriguez-Valera and W. F. Doolittle, 2004 Frequent recombination in a saltern population of *Halorubrum*. Science **306:** 1928–1929.

Posada, D., and K. A. Crandall, 2001 Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. USA **98:** 13757–13762.

Rossello-Moran, R., and R. Amann, 2001 The species concept for prokaryotes. FEMS Microbiol. Rev. **25:** 39–67.

Shimodaira, H., 2002 An approximately unbiased test of phylogenetic tree selection. Syst. Biol. **51:** 492–508.

Shimodaira, H., and M. Hasegawa, 2001 CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17:** 1246–1247.

Spratt, B. G., W. P. Hanage and E. J. Feil, 2001 The relative contributions of recombination and point mutation to the diversification of bacterial clones. Curr. Opin. Microbiol. **4:** 602–606.

Swofford, D. L., 2001 *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, MA.

Takahata, Y., M. Nishijima, T. Hoaki and T. Maruyama, 2001 Thermotoga petrophila sp. nov. and *Thermotoga naphthophila* sp. nov., two hyperthermophilic bacteria from the Kubiki oil reservoir in Niigata, Japan. Int. J. Syst. Evol. Microbiol. **51:** 1901–1909.

Thompson, J. R., S. Pacocha, C. Pharino, V. Klepac-Ceraj, D. E. Hunt *et al.*, 2005 Genotypic diversity within a natural coastal bacterioplankton population. Science **307:** 1311–1313.

Vinuesa, P., C. Silva, D. Werner and E. Martinez-Romero, 2005 Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in Bradyrhizobium species cohesion and delineation. Mol. Phylogenet. Evol. **34:** 29–54.

Worobey, M., 2001 A novel approach to detecting and measuring recombination: new insights into evolution in viruses, bacteria, and mitochondria. Mol. Biol. Evol. **18:** 1425–1434.

Yu, J. S., and K. M. Noll, 1997 Plasmid pRQ7 from the hyperthermophilic bacterium *Thermotoga* species strain RQ7 replicates by the rolling-circle mechanism. J. Bacteriol. **179:** 7161–7164.

Communicating editor: M. W. Feldman