ORIGINAL PAPER

Evolution of temperature optimum in *Thermotogaceae* and the prediction of trait values of uncultured organisms

Håkon Dahle · Bjarte Hannisdal · Bjørn Olav Steinsbu ·
Hege Ommedal · Jørn Einen · Sigmund Jensen ·
Øyvind Larsen · Lise Øvreås · Svein Norland

Abstract Quantitative characterization of the mode and rate of phenotypic evolution is rarely applied to prokaryotes. Here, we present an analysis of temperature optimum (T_{opt}) evolution in the thermophilic family *Thermotogaceae*, which has a large number of cultured representatives. We use log-rate-interval analysis to show that T_{opt} evolution in *Thermotogaceae* is consistent with a Brownian motion (BM) evolutionary model. The properties of the BM model are used to a establish confidence intervals on the unknown phenotypic trait value of an uncultured organism, given its distance to a close relative with known trait value. Cross-validation by bootstrapping indicates that the predictions are robust.

Keywords *Thermotogaceae* · Phenotypic evolution · Temperature optimum · Microbial ecology · Trait prediction · Thermophiles

Communicated by F. Robb.

H. Dahle · B. Hannisdal · B. O. Steinsbu · L. Øvreås
Centre for Geobiology, University of Bergen,
Allegaten 41, 5007 Bergen, Norway

H. Dahle (✉) · S. Jensen · L. Øvreås · S. Norland
Department of Biology, University of Bergen,
5007 Bergen, Norway
e-mail: Hakon.Dahle@bio.uib.no

B. Hannisdal · B. O. Steinsbu
Department of Earth Science, University of Bergen,
Allegaten 41, 5007 Bergen, Norway

H. Ommedal
Department of Chemistry, University of Bergen,
Allegaten 41, 5007 Bergen, Norway

J. Einen
Weyland AS, Espehaugen 25, 5258 Blomsterdalen,
Norway

Ø. Larsen
Department of Molecular Biology, University of Bergen,
Thormøhlensgate 55, 5008 Bergen, Norway

Introduction

Temperature optimum (T_{opt}) is arguably one of the most conserved phenotypic traits. For a cell to achieve a specific T_{opt}, most or all of its genes need to be functional at this temperature, e.g., all the essential genes of a thermophilic organism must encode thermostable proteins. Furthermore, although genomes may evolve quickly through for example frequent events of horizontal gene transfer, this does not mean that phenotypes in general, and particularly not T_{opt}, change rapidly through successive generations. On the contrary, it seems reasonable to assume that T_{opt}, or more specifically protein thermostability and cell-membrane thermostability, represents a barrier to fixation of horizontally acquired genes. Although frequent horizontal gene transfer may result in genomes being a mosaic of genes from various lineages, the changes in T_{opt} over the same time-frame may be very small.

How fast is the evolution of T_{opt}? Addressing this question requires a measure of how T_{opt} changes as a function of evolutionary distance. The 16S rRNA gene has for decades been a useful marker for measuring evolutionary relatedness among prokaryotes (Woese and Fox 1977; Woese 1987) and is a cornerstone of present day prokaryotic taxonomy (Coenye et al. 2005; Stackebrandt et al. 2002). Phylogenetic clusters derived from 16S rRNA gene

sequences commonly contain organisms with similar T_{opt} values, suggesting that although the 16S rRNA gene evolves extremely slowly, its resolution as a phylogenetic marker is sufficient to make inferences about T_{opt} evolution.

Phenotypic evolution can be divided into two components: the rate of evolution (tempo) and the mechanisms driving the rates (mode) (Kinnison and Hendry 2001; Simpson 1944). Quantitative characterization of these components is not only important for understanding the evolution, but may also have practical implications for microbial ecology. For example, knowing the mode of T_{opt} evolution in a microbial lineage may help constrain the expected T_{opt} values of uncultured organisms. This would be particularly useful for the organisms within thermophilic lineages as these organisms are typically detected in hydrothermal environments characterized by steep temperature gradients. For example, a few grams of a sample from the wall of a marine hydrothermal chimney may comprise micro niches with temperatures ranging from 0 to more than 300°C (Pagé et al. 2008). Thus, predicting T_{opt} of detected organisms from a temperature measurement of a bulk sample, if available at all, would be uninformative.

Here, we analyze the mode of T_{opt} evolution within the *Thermotogaceae* family, which has a large number of cultured type strains distributed in 7 described genera. Members of *Thermotogacea* are moderately thermophilic to hyperthermophilic, fermentative bacteria, typically isolated from hydrothermal systems. Our findings are consistent with the hypothesis that T_{opt} evolves according to a Brownian motion evolutionary model. We show how phenotypes of uncultured members of *Thermotogaceae* can be predicted based on known trait values of cultured relatives.

Data and methods

Database construction

We developed the software SPOT (sequence and phenotype organizing tool) to construct a database of 16S rRNA gene sequence information and T_{opt} information. SPOT, along with a users's manual and the dataset analyzed here, can be downloaded from http://webber.uib.no/geobio/spot. Phenotypic data were compiled from the literature as follows: *Fervidobacterium pennavorans* (Friedrich and Antranikian 1996), *Fervidobacterium changbaicum* (Cai et al. 2007), *Fervidobacterium gondwanense* (Andrews and Patel 1996), *Fervidobacterium islandicum* (Huber et al. 1990), *Fervidobacterium nodosum* (Patel et al. 1985), *Geotoga petraea* (Davey et al. 1993), *Geotoga subterranea* (Davey et al. 1993), *Kosmotoga olearia* (DiPippo et al. 2009), *Marinitoga camini* (Wery et al. 2001), *Marinitoga*

hydrogenitolerans (Postec et al. 2005), *Marinitoga piezophila* (Alain et al. 2002), *Marinitoga okinawensis* (Nunoura et al. 2007), *Petrotoga mexicana* (Miranda-Tello et al. 2004), *Petrotoga miotherma* (Davey et al. 1993), *Petrotoga mobilis* (Lien et al. 1998), *Petrotoga olearia* (L'Haridon et al. 2002), *Petrotoga sibirica* (L'Haridon et al. 2002), *Petrotoga halophila* (Miranda-Tello et al. 2007), *Thermosipho africanus* (Huber et al. 1989), *Thermosipho atlanticus* (Urios et al. 2004), *Thermosipho geolei* (L'Haridon et al. 2001), *Thermosipho japonicus* (Takai and Horikoshi 2000), *Thermosipho melanesiensis* (Antoine et al. 1997), *Thermotoga elfii* (Ravot et al. 1995), *Thermotoga hypogea* (Fardeau et al. 1997), *Thermotoga lettingae* (Balk et al. 2002), *Thermotoga maritima* (Huber et al. 1986), *Thermotoga naphthophila* (Takahata et al. 2001), *Thermotoga neapolitana* (Jannasch et al. 1988), *Thermotoga petrophila* (Takahata et al. 2001), *Thermotoga subterranea* (Jeanthon et al. 1995), *Thermotoga thermarum* (Windberger et al. 1989), and *Kosmotoga shengliensis* (Feng et al. 2010; Nunoura et al. 2010). Accession numbers of 16S rRNA gene sequences together with corresponding organism names and T_{opt} values are given in Supplementary Table 1.

Brownian motion simulation

Temporal evolution of a single trait can be modeled as a Gaussian random walk, which assumes that the long-term dynamics of an evolving population is governed by the mean and variance of the distribution of evolutionary steps. In the limit of very small steps, the random walk approaches a one-dimensional Wiener process, and the latter is commonly referred to as Brownian motion model in the evolutionary biology literature (Felsenstein 1985). We simulated one-dimensional Brownian motions (BM) by calculating the cumulative sum of random deviates drawn from a normal distribution, whose mean and standard deviation are the governing parameters, describing the rate (standard deviation) and directionality (mean) of evolution in a phenotypic trait. We define a BM model as equivalent to an unbiased random walk with a mean of zero. Modeling evolutionary changes as random steps is motivated by our uncertainty regarding the role of different microevolutionary processes and the need to minimize mechanistic assumptions, but does not imply that evolution is random with respect to underlying causal factors.

Linking T_{opt} variability with evolutionary distance

SPOT links an evolutionary distance matrix with phenotypic variability and produces a tab delimited output file including organism names, distances, and differences in T_{opt} (ΔT_{opt}) for all unique pairs of organisms. Whenever

T_{opt} values were reported as a range, differences in T_{opt} values between two strains were based on midpoint values. 16S rRNA distance matrices were constructed using the following approach: 16S rRNA genes sequences from all strains were aligned using the SINA webaligner (Pruesse et al. 2007) (http://www.arb-silva.de/aligner/). A distance matrix was generated from the alignment (*E. coli* positions 43–1371) using DNADIST with Jukes–Cantor correction and otherwise default settings as implemented in ARB (Ludwig et al. 2004). Comparisons involving branch lengths obtained from phylogenetic trees were done using in-house perl scripts.

Construction of phylogenetic trees and calculation of independent contrasts

The SINA-generated 16S rRNA gene sequence alignment (see above) was exported to ARB (version 5.0) (Ludwig et al. 2004), where trees were produced using three different algorithms: (1) maximum likelihood by PhyML (Guindon and Gascuel 2003) applying the Hasegawa–Kishono–Yano nucleotide substitution model and where a discrete-gamma model (Yang et al. 1994) is implemented to accommodate rate variation among sites (four substitution rate categories were used and the gamma distribution parameter was estimated by maximizing the likelihood of the phylogeny), (2) maximum parsimony by Phylip DNAPARS (Felsenstein 1989), and (3) neighbor joining (Saitou and Nei 1987) with Jukes–Cantor (Jukes and Cantor 1969) correction. All trees were constructed from the same alignment (*E. coli* position 43–1371) and by applying a bacterial positional variability filter (pos_var_Bacteria_102), leaving 1357 valid columns. For other parameters in the tree constructions, default settings were used.

Trees were exported from ARB in newick format and converted to nexus trees using in-house perl scripts. Nexus trees were imported by Mesquite (version 2.72) (http://mesquiteproject.org/mesquite/mesquite.html), where phylogenetically independent contrasts (IC) (Felsenstein 1985) were calculated using the PDAP (Garland et al. 1999) module (version 1.14). In PDAP, phenotypic data from tip organisms in a phylogenetic tree are used to estimate the phenotypes of internal nodes. IC are calculated as $(T_{opt(node1)} - T_{opt(node2)})([L_{branch}]^{-0.5})$, where node1 and node2 are two sister nodes in the tree and L_{branch} is a corrected branch length between the sister nodes (see Garland et al. (2005) for a worked example).

A possible method for prediction of trait values of uncultured organisms

Generally, if we let X_i be the value of a continuous phenotypic trait X in strain i, then a BM evolutionary model predicts that at a certain evolutionary distance (D), independent observations of ($X_i - X_j$) are normally distributed with expectation zero and variance proportional to D (Felsenstein 1985). Thus, observations of ($X_i - X_j$) at different distances can be standardized by dividing by the square root of the distance, obtaining values of $S = (X_i - X_j)$ $(D^{-0.5})$, which are also normally distributed with expectation 0 and variance σ^2. $|S|$ will have a half-normal distribution with a standard deviation, s, which can be estimated from a sample of independent pairs of organisms with known trait values and known distances. A given value of s can further be used to estimate σ from the relationship $s^2 = \sigma^2(1 - 2/\pi)$. Given a 95% confidence interval $(-1.96\sigma, 1.96\sigma)$ of S, the confidence interval of $(X_i - X_j)$ is $[-1.96\sigma(D^{0.5}), 1.96\sigma(D^{0.5})]$. Therefore, if we have a pair of organisms (i, j) with known evolutionary distance D, but where only one of the trait values (X_j) is known, the 95% confidence interval for the unknown trait value (X_i) can be expressed as

$$X_i = X_j \pm 1.96\sigma(D^{0.5}) \tag{1}$$

Cross-validation

Cross-validation of each predictive equation was performed by iterative re-sampling (5,000 bootstrap replicates) of the original dataset, randomly assigning species into a training set (17 species) and a validation set (16 species). In each iteration, σ values were estimated from the random training set and used in Eq. 1. Next, each organism in the validation set was paired with a random organism in the training set and their observed T_{opt} difference (ΔT_{opt}) was compared with the predicted value. A prediction was considered to be erroneous when $1.96\sigma(D^{0.5}) < \Delta T_{opt}$. Error rates were defined as the total proportion of erroneous predictions.

Results and discussion

The mode of T_{opt} evolution in *Thermotogaceae*

Following Gingerich (1993, 2009), we use the distribution of log rate against log distance to assess the mode of evolution (e.g., random, directional, or static). In the case of BM simulations (Fig. 1a–c), the slope (Fig. 1c) in the LRI plot is close to the theoretical expectation of −0.5, and the intercept value is close to the true rate used to generate the time series. On a normal scale (Fig. 1b), rates of change over time appear low and almost constant across a wide range of distances, but with a sharp increase when the time of divergence is very low (i.e., close relatives). This pattern is expected under a BM model, because short-time intervals can show rapid directional change, whereas long time
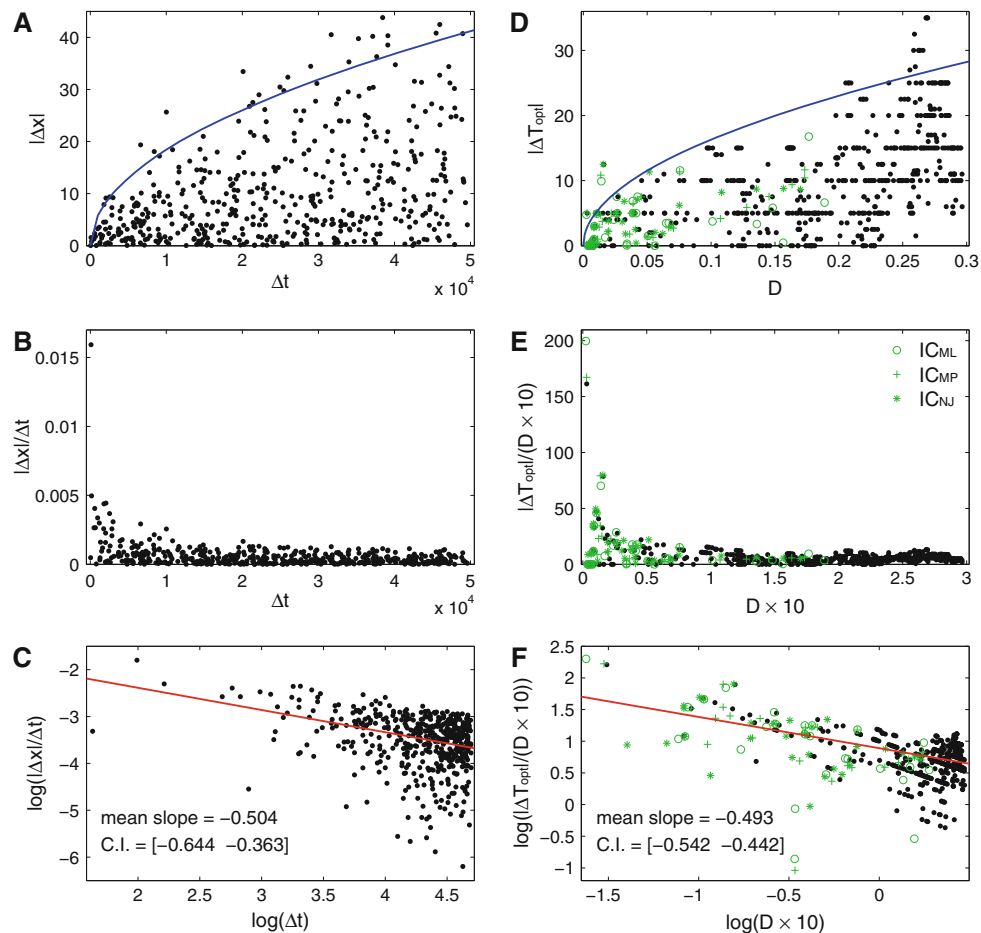
Fig. 1 Log-rate-interval analysis of evolutionary mode. **a** Distribution of absolute differences ($|\Delta x|$) between simulated trait values as a function of time interval (Δt) over which the difference is measured, from a set of 100 simulated evolutionary trajectories. Each trajectory is modeled as a one-dimensional Brownian motion over 50,000 time steps (generations) by calculating the cumulative sum of random deviates drawn from a Gaussian distribution with a mean of zero and standard deviation of 0.1. Pairwise differences are sampled randomly to match the number of pairs in the *Thermotogaceae* dataset. *Blue line* represents the 95% confidence interval as derived from Eq. 1. **b** The same data as in **a**, but plotted as rates (change per time interval) against time interval, showing the characteristic abrupt increase in rates at very short-time intervals. **c** On a log–log scale, the rates tend to follow a straight line. If we estimate the slope of that line repeatedly under iterated random sampling (5,000 replicates), we obtain a mean slope (−0.5; *red*) close to the theoretical expectation for a Brownian motion (−0.5), with a 95% confidence interval (CI) of [−0.644 −0.363]. **d** Absolute pairwise differences in optimal growth temperature (ΔT_{opt}) in the *Thermotogaceae*, as a function of

evolutionary distance ($D = 16S_{dist}$) between pairs (*black*). Independent contrasts (*green*) are plotted as a function of distance (D = branch length) based on three different algorithms: maximum likelihood (IC_{ML}), maximum parsimony (IC_{MP}), and neighbor joining (IC_{NJ}). *Blue line* represents 95% confidence interval as derived from Eq. 1. **e** Same data as in *panel D*, but plotted as rates against evolutionary distance. As in the simulated data (**b**), rates increase abruptly at short-time intervals. **f** LRI plot of the data in *panel E*, omitting pairs for which the ΔT_{opt} value is zero. To account for inaccuracy and binning of the reported T_{opt} values, we added random noise on the interval [−2.5 2.5] to T_{opt} and estimated the mean (−0.493; *red*) and 95% confidence interval ($CI = [−0.542 −0.442]$) of the resulting distribution of LRI slopes (5,000 replicates). The range of estimated slopes is consistent with a Brownian motion process (**c**). The three different independent contrasts reconstructions give different slopes ($IC_{ML} = −0.808$, $IC_{MP} = −0.671$, $IC_{NJ} = −0.397$; not shown), all of which fall within the 95% confidence interval of a Brownian motion [−0.948 −0.067], for the same number of contrasts ($N = 32$)

intervals incorporate more fluctuations and reversals, which tend to reduce the net change, and thus the apparent rate (Gingerich 1993; Sadler 1981). The precision of the slope estimate is a function of the number of observations, and to establish a significance criterion we calculated the mean and 95% confidence interval of the BM slope by repeated random walk simulations (5,000 bootstrap

replicates), given the same number of observations as the *Thermotogaceae* dataset. The resulting confidence interval on the BM slope [−0.644 −0.363] can be used to test the null hypothesis that the observed T_{opt} evolution does not significantly deviate from that expected under a BM model.

Our dataset contains 33 members of *Thermotogaceae*, yielding 528 unique pairs of organisms, with corresponding

DNADIST-derived evolutionary distances ($16S_{dist}$) and differences in T_{opt} (ΔT_{opt}) (Fig. 1d). For each pair, we calculated an average rate of T_{opt} evolution defined as $\Delta T_{opt}/16S_{dist}$, which shows the characteristic abrupt increase in rate at small distances (Fig. 1e). Inaccuracies in the reported T_{opt} values, many of which are binned into 5 degree intervals, would pose a serious problem only if that inaccuracy varies in a systematic manner on the phylogeny, e.g., if T_{opt} is measured differently within one subgroup of *Thermotogaceae* than in other subgroups. We do not have any reason to suspect that this is the case. Instead, we assume that over- and underestimation of T_{opt} is randomly distributed over the phylogeny. However, binning may have an effect on the LRI analysis by underestimating T_{opt} variance, potentially biasing the slope estimates toward stasis. In our view, true T_{opt} values are likely to fall within a 5-degree interval around the reported value. To account for this uncertainty, we added noise to the T_{opt} values in the form of uniformly random deviates drawn from the interval [−2.5 2.5]. This allows us to calculate a mean LRI slope (−0.493) and 95% confidence interval [−0.542 −0.442] by bootstrapping (5,000 replicates), which more realistically accounts for the inaccuracy of the T_{opt} data (Fig. 1f). The range of estimated LRI slopes for *Thermotogaceae* falls well within the 95% confidence interval for the BM model, consistent with the hypothesis that T_{opt} evolution occurs as BM.

A problem with applying LRI to all possible pairs is that different pairs of organisms are not independent because related organisms share parts of their evolutionary history (phylogenetic autocorrelation). In order to evaluate whether the LRI results are biased by such autocorrelation, we performed the LRI analysis on independent contrasts (IC) generated by three different algorithms (Fig. 1d–f). Since the number of IC is much smaller than the DNADIST-based dataset, the BM 95% confidence interval is correspondingly much wider [−0.948 −0.067]. The LRI slopes for all three IC reconstructions fall within that interval. Moreover, analyses in PDAP (Garland et al. 1999) showed no significant correlation between IC and the square root of branch length, further supporting our conclusion.

A framework for predicting T_{opt} in *Thermotogaceae*

The distribution in Fig. 1d can be used to quantitatively constrain T_{opt} predictions. For example, all 74 pairs observed to have $16S_{dist} < 0.10$ show $\Delta T_{opt} < 10°C$. Intuitively, one might infer that the T_{opt} of an uncultured organism is likely to differ by $< 10°C$ from the T_{opt} of a cultured close relative given that $16S_{dist} < 0.10$. The goal of this study is to assign quantitative measures of confidence on such predictions. LRI analysis shows that the evolution of T_{opt} in *Thermotogaceae* is consistent with a BM model. We can therefore take advantage of the properties of the BM model to obtain a general prediction of T_{opt} according to Eq. 1 by letting X represent T_{opt} values. As explained earlier, σ can be estimated in two ways: (1) a non-phylogenetic approach based on the values of $(X_k - X_l)(D^{-0.5})$, where X_k and X_l represent known T_{opt} values for two strains k and l, respectively, and D is $16S_{dist}$. Based on all 528 unique pairs, we found an estimated value of $\sigma = 26.3$. (2) A phylogenetic approach, whereby σ can be estimated from the square root of the variance of IC values. Depending on the algorithm used to obtain the evolutionary distances, this approach yields σ estimates in the range 29.8–33.3 (Table 1). The consistency between the different σ estimates is encouraging given the different assumptions and underlying distance measures.

Table 1 Overview of T_{opt} prediction equations generated from four different approaches to measuring evolutionary distance

	Methods requiring phylogenetic reconstruction			Non-phylogenetic method
Algorithm used to obtain evolutionary distances (D)	Maximum likelihood	Maximum parsimony	Neighbor-joining	DNADIST
D defined as	Branch length	Branch length	Branch length	$16S_{dist}$
Variance of $[\Delta T_{opt}(D^{-0.5})]$				251.1
σ	29.8	32.1	33.3	26.3
Prediction	$X_i = X_j \pm 58.4(D^{0.5})$	$X_i = X_j \pm 62.9(D^{0.5})$	$X_i = X_j \pm 65.3(D^{0.5})$	$X_i = X_j \pm 51.5(D^{0.5})$
Cross-validation				
Mean of σ	25.3	27.9	27.5	26.3
Standard deviation of σ	6.1	6.7	5.6	0.6
Median error rate	0	0	0.06	0
Mean error rate	0.06	0.04	0.07	0.04
95th percentile error rate	0.25	0.25	0.25	0.13

X_i unknown T_{opt} of a given organism i, X_j known T_{opt} of a given organism j, D evolutionary distance between i and j

Table 2 Predicted confidence intervals of T_{opt} from selected members of 'mesotogas'

Representative organism (accession number)	Lineage	Predicted T_{opt} (°C)			
		Maximum likelihood	Maximum parsimony	Neigbor-joining	DNADIST
HM003102	M1	44–86	42–88	43–87	46–84
HM003104	M2	43–87	43–87	43–87	46–84
EU735623	M4	33–96	28–97	32–88	36–84
CU921025	M4	34–95	31–98	38–92	41–89
HM003101	M5	21–79	22–78	23–76	27–73

Representatives from each mesotoga lineage were taken from Nesbø et al. (2010). See Table 1 for prediction equations

Cross-validation

Cross-validation by bootstrapping shows that the sigma values are robust to subsampling and stable regardless of the method used (Table 1). Furthermore, median T_{opt} prediction error rates are generally lower than the theoretical expectation of 0.05 (Table 1). Note that the error rate distributions are highly skewed (hence we report median values) with rare instances of large error rates (>0.25).

Constraining the T_{opt} in 'mesotogas'

Currently, all cultivated isolates of the *Thermotogaceae* family are either moderate thermophiles, thermophiles, or hyperthermophiles. However, 16S rRNA gene sequences from members of this family have been detected in mesothermic environments or enrichment cultures. The corresponding lineages have informally been designated as 'mesotogas' (Nesbø et al. 2006, 2010). In the absence of a mesophilic isolate, their existence remains unproven. Nevertheless, we used the equations in Table 1 to address the following questions: are uncultured mesotogas unlikely to be mesophiles given their relatedness to thermophilic organisms? Are all *Thermotogaceae* detected in mesothermic environments likely to be mesophiles? Here, we analyzed a dataset recently published by Nesbø et al. (2010) describing five putative lineages of mesotogas. Four of these (M1, M2, M4, and M5) are represented by at least one near full-length 16S rRNA gene sequence. Using the equations in Table 1, we predicted confidence intervals of T_{opt} values of mesotoga members given their evolutionary distance to the closest cultured relative (Table 2). The predictions are of limited precision because the mesotogas are only distantly related to the described species (DNADIST and branch length values were in the range of 0.11–0.26). Yet, our results indicate that it is reasonable to assume that mesotoga lineages M4 and M5 comprise mesophilic organisms, whereas lineages M1 and M2 seem to be at least moderately thermophilic. It is still possible that lineages M1 and M2 are mesophilic and have

temperature optima lower than predicted (Table 2), but if they are, the rate of T_{opt} evolution in the branches leading to these lineages are unexpectedly high under a BM model.

Conclusion

The evolution of T_{opt} within *Thermotogaceae* is consistent with a Brownian motion model of phenotypic evolution. Based on this model, we developed a general method for T_{opt} prediction of uncultured members of this family. Cross-validation shows that the predictions are accurate (low error rates), stable under different phylogenetic reconstructions, and robust to taxonomic sampling. Similar analyses can be performed on any continuous characteristic (e.g., regarding pressure, pH, salinity, growth rate, genome size, and % GC content) and within any microbial group, which would greatly enhance the value of functional studies and provide valuable insight into evolutionary dynamics.

References

Alain K, Marteinsson VT, Miroshnichenko ML, Bonch-Osmolovskaya EA, Prieur D, Birrien J-L (2002) *Marinitoga piezophila* sp. nov., a rod-shaped, thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. Int J Syst Evol Microbiol 52:1331–1339

Andrews KT, Patel BKC (1996) *Fervidobacterium gondwanense* sp. nov., a new thermophilic anaerobic bacterium isolated from nonvolcanically heated geothermal waters of the Great Artesian Basin of Australia. Int J Syst Bacteriol 46:265–269

Antoine E, Cilia V, Meunier JR, Guezennec J, Lesongeur F, Barbier G (1997) *Thermosipho melanesiensis* sp. nov., a new thermophilic anaerobic bacterium belonging to the order Thermotogales, isolated from deep-sea hydrothermal vents in the Southwestern Pacific Ocean. Int J Syst Bacteriol 47:1118–1123

Balk M, Weijma J, Stams AJM (2002) *Thermotoga lettingae* sp. nov., a novel thermophilic, methanol-degrading bacterium isolated from a thermophilic anaerobic reactor. Int J Syst Evol Microbiol 52:1361–1368

Cai J, Wang Y, Liu D, Zeng Y, Xue Y, Ma Y, Feng Y (2007) *Fervidobacterium changbaicum* sp. nov., a novel thermophilic anaerobic bacterium isolated from a hot spring of the Changbai Mountains, China. Int J Syst Evol Microbiol 57:2333–2336

Coenye T, Gevers D, de Peer YV, Vandamme P, Swings J (2005) Towards a prokaryotic genomic taxonomy. FEMS Microbiol Rev 29:147–167

Davey ME, Wood WA, Key R, Nakamura K, Stahl DA (1993) Isolation of three species of *Geotoga* and *Petrotoga*: two new genera, representing a new lineage in the bacterial line of descent distantly related to the "Thermotogales". Syst Appl Microbiol 16:191–200

DiPippo JL, Nesbø CL, Dahle H, Doolittle WF, Birkland N-K, Noll KM (2009) *Kosmotoga olearia* gen. nov., sp. nov., a thermophilic, anaerobic heterotroph isolated from an oil production fluid. Int J Syst Evol Microbiol 59:2991–3000

Fardeau ML, Ollivier B, Patel BKC, Magot M, Thomas P, Rimbault A, Rocchiccioli F, Garcia JL (1997) *Thermotoga hypogea* sp. nov., xylanolytic, thermophilic bacterium from an oil-producing well. Int J Syst Bacteriol 47:1013–1019

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Felsenstein J (1989) PHYLIP—phylogeny inference package (version 3.2). Cladistics 5:925–926

Feng Y, Cheng L, Zhang X, Li X, Deng Y, Zhang H (2010) Thermococcoides shengliensis gen. nov., sp. nov., a new member of the order Thermotogales isolated from oil-production fluid. Int J Syst Evol Microbiol 60:932–937

Friedrich AB, Antranikian G (1996) Keratin degradation by *Fervidobacterium pennavorans*, a novel thermophilic anaerobic species of the order Thermotogales. Appl Environ Microbiol 62:2875–2882

Garland T Jr, Midford PE, Ives AR (1999) An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. Am Zool 39:374–388

Garland T, Bennett AF, Rezende EL (2005) Phylogenetic approaches in comparative physiology. J Exp Biol 208:3015–3035

Gingerich PD (1993) Quantification and comparison of evolutionary rates. Am J Sci 293A:453–478

Gingerich PD (2009) Rates of evolution. Annu Rev Ecol Evol Syst 40:657–675

Guindon Sp, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Huber R, Langworthy TA, Konig H, Thomm M, Woese CR, Sleytr UB, Stetter KO (1986) *Thermotoga maritima*, sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch Microbiol 144:324–333

Huber R, Woese CR, Langworthy TA, Fricke H, Stetter KO (1989) *Thermosipho africanus*, gen. nov., represents a new genus of thermophilic eubacteria within the Thermotogales. Syst Appl Microbiol 12:32–37

Huber R, Woese CR, Langworthy TA, Kristjansson JK, Stetter KO (1990) *Fervidobacterium islandicum* sp. nov., a new extremely thermophilic eubacterium belonging to the "Thermotogales". Arch Microbiol 154:105–111

Jannasch HW, Huber R, Belkin S, Stetter KO (1988) *Thermotoga neapolitania* sp. nov. of the extremely thermophilic eubacterial genus Thermotoga. Arch Microbiol 150:103–104

Jeanthon C, Reysenbach A-L, L'Haridon S, Gambacorta A, Pace NR, Glenat P, Prieur D (1995) *Thermotoga subterranea* sp. nov., a new thermophilic bacterium isolated from a continental oil reservoir. Arch Microbiol 164:91–97

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 211–232

Kinnison MT, Hendry AP (2001) The pace of modern life II: from rates of contemporary microevolution to pattern and process. Genetica 112–113:145–164

L'Haridon S, Miroshnichenko ML, Hippe H, Fardeau ML, Bonch-Osmolovskaya E, Stackebrandt E, Jeanthon C (2001) *Thermosipho geolei* sp. nov., a thermophilic bacterium isolated from a continental petroleum reservoir in Western Siberia. Int J Syst Evol Microbiol 51:1327–1334

L'Haridon S, Miroshnichenko ML, Hippe H, Fardeau ML, Bonch-Osmolovskaya EA, Stackebrandt E, Jeanthon C (2002) *Petrotoga olearia* sp. nov. and *Petrotoga sibirica* sp. nov., two thermophilic bacteria isolated from a continental petroleum reservoir in Western Siberia. Int J Syst Evol Microbiol 52:1715–1722

Lien T, Madsen M, Rainey FA, Birkeland NK (1998) *Petrotoga mobilis* sp. nov., from a North Sea oil-production well. Int J Syst Bacteriol 48:1007–1013

Ludwig W, Strunk O, Westram R et al (2004) ARB: a software environment for sequence data. Nucl Acids Res 32:1363–1371

Miranda-Tello E, Fardeau ML, Thomas P, Ramirez F, Casalot L, Cayol JL, Garcia JL, Ollivier B (2004) *Petrotoga mexicana* sp. nov., a novel thermophilic, anaerobic and xylanolytic bacterium isolated from an oil-producing well in the Gulf of Mexico. Int J Syst Evol Microbiol 54:169–174

Miranda-Tello E, Fardeai M-L, Joullan C, Magot M, Thomas P, Tholozan J-L, Ollivier B (2007) *Petrotoga halophila* sp. nov., a thermophilic, moderately halophilic, fermentative bacterium isolated from an offshore oil well in Congo. Int J Syst Evol Microbiol 57:40–44

Nesbø CL, Dlutek M, Zhaxybayeva O, Doolittle WF (2006) Evidence for existence of "mesotogas" members of the order Thermotogales adapted to low-temperature environments. Appl Environ Microbiol 72:5061–5068

Nesbø CL, Kumaraswamy R, Dlutek M, Doolittle WF, Foght J (2010) Searching for mesophilic Thermotogales bacteria: 'Mesotogas' in the wild. Appl Environ Microbiol 76:4896–4900

Nunoura T, Oida H, Miyazaki M, Suzuki Y, Takai K, Horikoshi K (2007) *Marinitoga okinawensis* sp. nov., a novel thermophilic and anaerobic heterotroph isolated from a deep-sea hydrothermal field, Southern Okinawa Trough. Int J Syst Evol Microbiol 57:467–471

Nunoura T, Hirai M, Imachi H, Miyazaki M, Makita H, Hirayama H, Furushima Y, Yamamoto H, Takai K (2010) *Kosmotoga arenicorallina* sp. nov a thermophilic and obligately anaerobic heterotroph isolated from a shallow hydrothermal system occurring within a coral reef, southern part of the Yaeyama Archipelago, Japan, reclassification of *Thermococcoides shengliensis* as *Kosmotoga shengliensis* comb. nov., and emended description of the genus Kosmotoga. Arch Microbiol 192:811–819

Pagé A, Tivey MK, Stakes DS, Reysenbach AL (2008) Temporal and spatial archaeal colonization of hydrothermal vent deposits. Environ Microbiol 10:874–884

Patel BKC, Morgan HW, Daniel RM (1985) *Fervidobacterium nodosum* gen. nov and spec. nov, a new chemoorganotrophic, caldoactive, anaerobic bacterium. Arch Microbiol 141:63–69

Postec A, Le Breton C, Fardeau M-L, Lesongeur F, Pignet P, Querellou J, Ollivier B, Godfroy A (2005) *Marinitoga hydrogenitolerans* sp. nov., a novel member of the order Thermotogales isolated from a black smoker chimney on the Mid-Atlantic Ridge. Int J Syst Evol Microbiol 55:1217–1221

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucl Acids Res 35:7188–7196

Ravot G, Magot M, Fardeau ML, Patel BKC, Prensier G, Egan A, Garcia JL, Ollivier B (1995) *Thermotoga elfii* sp. nov., a novel thermophilic bacterium from an African oil-producing well. Int J Syst Bacteriol 45:308–314

Sadler PM (1981) Sediment accumulation rates and the completeness of stratigraphic sections. J Geol 89:569–584

Saitou N, Nei M (1987) The neighbor-joining method—a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Simpson GC (1944) Tempo and mode in evolution. Columbia University Press, New York

Stackebrandt E, Frederiksen W, Garrity GM et al (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 52:1043–1047

Takahata Y, Nishijima M, Hoaki T, Maruyama T (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

Takai K, Horikoshi K (2000) *Thermosipho japonicus* sp. nov., an extremely thermophilic bacterium isolated from a deep-sea hydrothermal vent in Japan. Extremophiles 4:9–17

Urios L, Cueff-Gauchard V, Pignet P, Postec A, Fardeau ML, Ollivier B, Barbier G (2004) *Thermosipho atlanticus* sp. nov., a novel member of the Thermotogales isolated from a Mid-Atlantic Ridge hydrothermal vent. Int J Syst Evol Microbiol 54:1953–1957

Wery N, Lesongeur F, Pignet P, Derennes V, Cambon-Bonavita MA, Godfroy A, Barbier G (2001) *Marinitoga camini* gen. nov., sp. nov., a rod-shaped bacterium belonging to the order Thermotogales, isolated from a deep-sea hydrothermal vent. Int J Syst Evol Microbiol 51:495–504

Windberger E, Huber R, Trincone A, Fricke H, Stetter KO (1989) *Thermotoga thermarum* sp. nov. and *Thermotoga neapolitana* occuring in African continental solfataric springs. Arch Microbiol 151:506–512

Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221–271

Woese CR, Fox GE (1977) Phylogenetic structure of prokaryotic domain—primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090

Yang ZH, Goldman N, Friday A (1994) Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. Mol Biol Evol 11:316–324