bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Title: Temperature-induced transcriptional responses of the deep-biosphere bacterium, *Kosmotoga olearia*, illuminate its adaptation to growth from 20ºC to 79ºC

Running title: Transcriptional response to temperature in *K. olearia*

Stephen M. J. Pollo[a], Abigail A. Adebusuyi[a*], Timothy J. Straub[b], Julia M. Foght[a], Olga Zhaxybayeva[b,c] and Camilla L. Nesbø[a,d#]

[a]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

[b]Department of Biological Sciences, Dartmouth College, Hanover, NH, USA

[c]Department of Computer Science, Dartmouth College, Hanover, NH, USA

[d]Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, Oslo, Norway

Address correspondence to nesbo@ualberta.ca

[*]Current address: Centre for Research & Innovation, Grande Prairie Regional College, Grande Prairie, AB

Word count for the abstract: 249 + 138 Word count for the text: 6056

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Temperature is an important parameter determining the distribution of organisms, and ambient temperature change is a physiological challenge commonly faced by living cells. Most organisms are adapted to growing within a temperature range that rarely exceeds ~ 30°C, but the anaerobic thermophilic bacterium *Kosmotoga olearia* is capable of growing over an extremely wide temperature range (20°C - 79°C). We used transcriptomic and comparative genomic analyses to elucidate the mechanisms enabling this extraordinary trait. When transcriptomes of *K. olearia* grown at 30°C, 40°C, and 77°C were compared to those from its optimal growth at 65°C, 573 of 2,224 genes (25%) were significantly differentially expressed. We find that *K. olearia* remodels its metabolism dramatically at different temperatures, with increased expression of genes involved in energy and carbohydrate metabolism at high temperatures and up-regulation of amino acid metabolism at lower temperatures. At sub-optimal temperatures, many transcriptional changes were similar to those observed in mesophilic bacteria at physiologically low temperatures, including up-regulation of genes encoding enzymes for fatty acid synthesis, typical cold stress genes, and ribosomal proteins. In comparison to other Thermotogae, *K. olearia* has multiple copies of some cold-associated genes, suggesting that an increase in gene copy number is a strategy for cold adaptation. Many of these cold response genes are predicted to be laterally acquired, highlighting the role of gene exchange in bacterial thermoadaptation. Notably, at 77°C one third of the up-regulated genes encode proteins with hypothetical functions, indicating that many features of adaptations to high temperature growth are still unknown.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Importance

The subsurface is arguably the largest habitat on Earth, and insights into the composition, adaptation, and evolution of its microbial communities is likely to advance our knowledge of the biosphere and of global element cycling. Some subsurface systems, such as oil reservoirs, are thought to be so stable that any changes in environmental conditions, including temperature, occur over thousands of years. Yet a member of the oil reservoir microbial communities, the bacterium *Kosmotoga olearia*, is capable of growing over a range of 59°C. Our finding of coordinated temperature-specific gene expression patterns, and by extension temperature specific metabolism, suggests that *Kosmotoga* populations encounter variable environments, probably through migration. This raises the question: are deep subsurface microbial communities more dynamic than currently perceived? The large number of identified temperature-responsive genes also indicates that temperature response is a complex polygenic trait.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Microorganisms are capable of growing over an impressive temperature range, at least from -15°C to 122°C (1, 2), and temperature is one of the most important physical factors determining their distribution, diversity, and abundance (3). However, individual microbial species grow only within a much narrower temperature interval. For example, *Escherichia coli* O157:H7 thrives in the laboratory between 19°C and 41°C (4), while *Geobacillus thermoleovorans* has a growth range of 37°C to 70°C (5). Microorganisms with temperature ranges >50°C are rare, and research into the few with ranges >40°C has focused on psychrophiles (e.g., (2)). *Kosmotoga olearia* TBF 19.5.1 (hereafter referred to as *K. olearia*) is an anaerobic thermophile from the bacterial phylum Thermotogae with a growth range that spans almost 60°C (6). How does this lineage achieve such physiological flexibility and what are the evolutionary advantages and implications of having this capability?

Fluctuations in temperature induce broad physiological changes in cells, including alterations to cell wall and membrane composition, translation, and energy metabolism (3, 7, 8). These physiological changes can be classified into two broad types of cellular response. Cold or heat *shock* designates the changes observed *immediately* after the shift of a culture to a lower or higher temperature, while *prolonged growth* at a specific lower or higher temperature elicits an *acclimated* low- or high-temperature response (7). Most studies on prokaryotes have focused on temperature shock responses. Among the Thermotogae, responses to both heat shock and prolonged growth at high temperatures have been studied in the hyperthermophile *Thermotoga maritima*, which can grow between 55°C and 90°C (9, 10). During prolonged high temperature growth *T. maritima* strongly up-regulates central carbohydrate metabolism genes and expresses a few typical heat shock protein genes (10). Little is known about how *T. maritima* responds to

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sub-optimal temperatures, although it encodes some genes implicated in cold shock response. For example, its family of cold shock proteins (Csp), which are nucleic acid chaperones known to be induced during cold shock and cold acclimation in mesophilic bacteria (7, 11), exhibits nucleic acid melting activity at physiologically low temperatures (12). Similarly, responses to cold shock in a few other thermophiles involve many of the genes implicated in mesophiles' cold shock response (e.g., (13, 14)). In this study we systematically assess bacterial physiological changes associated with response to prolonged growth at both high and low temperature using *K. olearia* as a model system.

The *K. olearia* genome (NC_012785) has 2,302,126 bp and is predicted to encode 2,224 genes (15). Within the Thermotogae, genome size, intergenic region size, and number of predicted coding regions correlate with the optimal growth temperature of an isolate (16), with hyperthermophilic Thermotogae genomes being the most compact. Phylogenetically, the Thermotogae order Kosmotogales comprises the genera *Kosmotoga* and *Mesotoga* spp., the latter being the only described mesophilic Thermotogae lineage (8). Assuming a hyperthermophilic last common ancestor of the Thermotogae (17), the Kosmotogales can be hypothesized to have acquired wide growth temperature tolerance secondarily by expanding its gene repertoire. Moreover, it is likely that the ability of the Kosmotogales common ancestor to grow at low temperatures made the evolution of mesophily in *Mesotoga* possible (8).

Such adaptations of lineages to new environments can be greatly facilitated by lateral gene transfer (LGT), since genes already "adapted" to the new conditions are readily available in the microbial communities of the new environment (18). For instance, LGT has been implicated in adaptation to high temperature growth in hyperthermophilic bacteria, including *Thermotoga* spp., and to low temperature growth in Archaea (8, 18, 19). Genome analysis of the mesophilic

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mesotoga prima revealed that it laterally acquired 32% of its genes after it diverged from other Thermotogae lineages (16). Many of the predicted gene donors are mesophiles, supporting the importance of lateral acquisition of genes already adapted to mesophilic conditions in the evolution of *Mesotoga*.

To further gain insights into mechanisms of bacterial temperature response we sequenced 17 transcriptomes from isothermal and temperature-shifted cultures of *K. olearia* and examined transcriptional differences at temperatures spanning its wide growth range. Additionally, through comparative genomic and phylogenetic analyses of identified temperature responsive genes and their homologs in two newly sequenced *Kosmotoga* isolates, as well as in genomes of other thermophilic and mesophilic Thermotogae, we investigated the importance of within-lineage evolution through LGT and gene duplication for adaptation of *K. olearia* to growth over a wide temperature range.

Results and Discussion

Temperature shifts and isothermic conditions elicit different growth patterns in *K. olearia*. Under laboratory conditions in liquid anaerobic medium we observed growth of *K. olearia* at temperatures as low as 25°C and as high as 79°C, with optimal growth at 65°C (Fig. 1 and Fig. S1 in the supplemental material). Using a non-linear regression model (20) we estimate a growth-permissive temperature range of 20.2 – 79.3°C, consistent with the previously reported wide growth range of this isolate (6). Interestingly, we were not able to cultivate *K. olearia* at temperatures near its range boundaries (30°C and 77°C) by direct transfer from 65°C cultures. Instead, the growth temperature had to be changed sequentially in ≤10°C increments. Particularly

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

at the extremes, even small temperature shifts caused both a longer lag phase and a slower growth rate compared to isothermal cultures (Fig. 1 and Fig. S1 in the supplemental material). This phenomenon has also been noted for mesophilic bacteria, especially for transitions from high to low temperature (21). Our observations suggest that cells shifted to a new temperature need to undergo large physiological changes that require time (i.e. an 'acclimation' period (7)) and that these physiological challenges are too great to overcome when temperature changes are large. To illuminate *K. olearia*'s transcriptional responses to changes in temperature we sequenced 17 transcriptomes of replicate mid- to late-log cultures grown isothermally at 30°C, 40°C, 65°C, and 77°C, and of two 30°C cultures shifted to 25°C and 4°C (see Table S1 in Dataset S1 and Text S1 in the supplemental material).

Architecture of the *K. olearia* transcriptome. Analysis of transcription start and stop sites predicted a minimum of 916 transcriptional units (TU) in *K. olearia* (Text S1 and Table S2 in Dataset S1 in the supplemental material), 52% of which consist of a single gene. This fraction of single-gene TUs lies between the 43% recorded for *T. maritima*, which has been shown to have a streamlined genome and a low-complexity transcriptome (22), and the 65% reported for *E. coli* (23). The average TU length of ~2.39 genes in *K. olearia* is less than the 3.3 genes per transcript of *T. maritima* (22) but closer to 2.2 genes per transcript in the mesophilic firmicute *Geobacter sulfurreducens* (24) and 1-2 genes per transcript in Bacteria in general (e.g. (23)). Given that the *K. olearia* genome has more intergenic DNA than *T. maritima*'s genome (the ratio of the nucleotides located in non-coding vs. coding regions is 0.13 in *K. olearia* and 0.06 in *T. maritima*), the shorter TU lengths in *K. olearia* may point to more flexible transcriptional regulation and may be linked to *Kosmotoga*'s ability to grow under more variable conditions.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Consistent energy generation across different temperature conditions. *K. olearia* produces ATP from pyruvate using a biochemically well-understood fermentation pathway that generates hydrogen, carbon dioxide and acetate ((6); Fig. 2 and data not shown). Given that pyruvate was the carbon and energy source in all experiments, we surveyed 51 genes predicted to be involved in pyruvate catabolism and identified 15 genes with consistently high expression in all temperature treatments (Fig. 2 Table S3 in Dataset S1 in the supplemental material). In addition to indirectly validating the previously known functional annotations of these genes, we furthermore propose that genes Kole_1509 – 1513 encode a pyruvate ABC transporter (Fig. 2). Their current annotation as a peptide ABC transporter may be erroneous since most of the peptide ABC transporters predicted in *T. maritima* using bioinformatics have been shown instead to bind and transport sugars (25). Our findings also indicate that the enzymes involved in the pyruvate fermentation pathway are versatile enzymes that are expressed and are capable of functioning across an extremely wide temperature range.

Identification of temperature-related transcriptional responses in *K. olearia.* Based on hierarchical clustering, transcriptome replicates at the same temperature group together (Fig. S2 and Text S1 in the supplemental material), suggesting that the observed changes in transcription are due to the culture growth temperature. Principal Component Analysis (PCA) clearly separated the transcriptomes into quadrants corresponding to optimal (65°C), intermediate (40°C), low (30°C, 25°C and 4°C) and high (77°C) growth temperatures (Fig. 3). Several genes with a high correlation between their expression level and a specific growth temperature (vectors in Fig. 3, Table S4 in Dataset S1 in the supplemental material) are known to be involved in temperature response. For example, expression of the protease Kole_1599 positively correlated with the 77°C transcriptomes, where high expression of proteases was

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expected based on their involvement in heat shock response in *T. maritima* (9). Similarly, expression of the cold shock protein genes Kole_0109 and Kole_2064 positively correlated with low temperature growth. Lastly, some observed changes presumably were due to the expected decreased metabolic activity of the culture at a non-optimal temperature. This can be exemplified by the high expression and strong correlation of the central carbon metabolism gene glyceraldehyde-3-phosphate dehydrogenase (Kole_2020) with the 65°C transcriptomes. Overall, the observed differential expression of known temperature-responsive genes implies that the remaining detected transcriptional changes also likely reflect temperature-related trends in gene expression.

Detailed analysis of changes in gene expression in response to prolonged growth at different temperatures. Putative temperature-responsive genes were identified by pairwise comparisons of each isothermic temperature treatment to the optimal growth at 65°C (i.e., 30°C vs 65°C, 40°C vs 65°C, and 77°C vs 65°C). Across all comparisons 573 genes fulfilled our criteria for temperature responsiveness (i.e., ≥ 2-fold difference in expression, > 20 reads per transcript, False Discovery Rate < 0.05) with 430, 115, and 169 genes detected in the 30°C vs 65°C, 40°C vs 65°C, and 77°C vs 65°C comparisons respectively (Table S5 in Dataset S1 in the supplemental material). Most of these genes were up-regulated (Fig. S3 in the supplemental material) with the exception of down-regulation of many genes involved in carbohydrate and energy metabolism at 30°C (Clusters of Orthologous Groups [COG] categories G and C). The latter probably reflects the very slow growth rate at this temperature. Curiously, despite the slower growth rate at 77°C, genes from COG C and G categories were over-represented and up-regulated at this temperature (Fig. 4 and Fig. S3 in the supplemental material; discussed in detail below).

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In all transcriptomes the list of putative temperature-responsive genes was depleted of genes involved in translation (COG category J) and nucleotide metabolism (COG category F) (Fig. 4 and Fig. S3 in the supplemental material) and enriched in genes involved in replication, recombination and repair (COG category L, particularly at 30°C), and signal transduction (COG category T). Most of the identified COG category L genes are either mobile elements or CRISPR-associated proteins. Movement of mobile genetic elements is a common feature of stress responses (26) and the up-regulation of CRISPR-associated genes could therefore be a response to the proliferation of these elements. Differential expression of the signal transduction genes suggests the importance of these systems for regulating cellular responses at all tested temperatures. Additionally, at both 30°C and 77°C many genes encoding transcription regulators (COG category K, transcription) are up-regulated, suggesting that prolonged growth at sub- and supra-optimal temperatures results in detectable changes in transcriptional gene regulation in *K. olearia*. Below we discuss the identified temperature-responsive gene expression patterns in more detail.

i) At 40°C there are pronounced differences in membrane fatty acid composition but no signs of cold stress. Although the growth rate of *K. olearia* at 40°C is only one-third of that at the optimum 65°C (Fig. 1 and Fig. S1 in the supplemental material), clustering analysis suggested that the 40°C transcriptome was most similar to that at 65°C (Fig. 3 and Fig. S2 in the supplemental material). The slower growth rate was reflected by the four most highly expressed temperature responsive genes at 40°C showing significantly lower expression than at 65°C, including growth-related genes like the toga protein Kole_1501 (Table S5 in Dataset S1 in the supplemental material). Yet, 94 of 115 putative temperature responsive genes were up-regulated

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Table S5 in Dataset S1 in the supplemental material), suggesting that slower metabolism is not the only explanation for the observed transcriptional response to growth at 40°C.

Lipid metabolism (COG category I) appears to be particularly important at 40°C. For instance, all of the predicted fatty acid synthesis genes showed the highest expression at 40°C (Table S5 in Dataset S1 and Fig. S4 in the supplemental material), with two genes involved in synthesis of unsaturated fatty acids (Kole_0968) and initiation of fatty acid synthesis (Kole_0969) having significantly higher expression. Biochemical analyses of total fatty acids at 40°C and 65°C showed a much greater diversity of fatty acids at 40°C (Table S6 in Dataset S1 in the supplemental material), which may explain the higher demand for these genes at lower temperatures. Interestingly, at 40°C in particular there was increased expression of a phosphate ABC transporter (Kole_0707 – Kole_0711, Table S5 in Dataset S1 in the supplemental material), which may be linked to the increased production of polar membrane lipids at moderately low temperatures. Maintenance of a functional cell membrane is crucial for survival, and bacteria respond to changes in temperature by altering the membrane's fatty acid composition (27). The observation that lipid metabolism genes were among the highly expressed genes at low temperature, despite the lower growth rate, suggests that changes to the cell membrane composition are one of the most important adaptations for survival of *K. olearia* at lower temperatures.

Proper protein folding at a lower temperature is another physiological challenge that may require enzymatic assistance. For example, proline isomerization happens spontaneously at high temperatures, but at lower temperatures (e.g., 37°C) the reaction needs a catalyzing enzyme - peptidylprolyl isomerase (PPIase) (28). Not surprisingly, *K. olearia* has three temperature-responsive PPIase genes: two PpiC-type genes (Kole_1682 and Kole_0383) that are both highly

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expressed at 40°C, and one FKBP-type gene (Kole_1745), which shows high expression at all temperatures except 77°C (Table S5 in Dataset S1 in the supplemental material). These expression patterns suggest PPIase is particularly important at moderately low temperatures where the cells are still relatively active. However, the enzymes known to assist protein folding in cellular stress responses, chaperones (e.g., GroEL and Hsp) and protease Do, were significantly down-regulated at 40°C (Table S5 in Dataset S1 in the supplemental material). Among other typical cold stress related proteins, only one of *K. olearia*'s three cold shock proteins (Kole_0109) showed significantly higher expression at 40°C and its up-regulation was merely moderate when compared to its expression levels at 30°C (Table S5 in Dataset S1 in the supplemental material). This overall lack of induction of typical stress-related genes, especially when compared to 30°C and 77°C (see below), suggests that 40°C is still within the "Goldilocks" temperature range for *K. olearia.*

ii) *K. olearia* is in cold stress at 30°C. Transcriptomes from 30°C, 25°C, and 4°C cultures were very similar to each other (Fig. 3 and Fig. S2 in the supplemental material). Overall, the gene expression differences observed at 30°C were even more pronounced at 25°C and 4°C (Table S5 in Dataset S1 in the supplemental material). However, due to adjustments in culture handling required to obtain enough biomass at lower temperatures (see Text S1 in the supplemental material), some gene expression patterns at 25°C and 4°C may be due to the cells either responding to fresh medium or displaying an immediate cold shock response. Therefore, we focused our further analyses on genes differentially expressed at 30°C, while the 25°C and 4°C transcriptomes were used to confirm the patterns observed at 30°C.

Two of three Csp-encoding genes in *K. olearia* (Kole_0109 and Kole_2064, Table S5 in Dataset S1 in the supplemental material) were among the three most highly expressed up-

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

regulated genes at low temperatures, suggesting that the cells were in a cold-stressed state during growth at ≤30°C. Further support for this hypothesis comes from significant up-regulation of genes linked to bacterial cold response (7): a DEAD/DEAH -box RNA helicase (Kole_0922), RbfA (Kole_2103), and NusA (Kole_1529). Hence, the thermophile *K. olearia* uses homologs of the cold response genes employed by mesophilic bacteria at physiologically low temperatures.

With decreasing temperature, we observed up-regulation of several ribosomal proteins (Fig. 3). Some (L10 (Kole_1840) and L7/L12 (Kole_1839)) have already been linked to both cold shock and prolonged low temperature growth responses in bacteria (e.g., (29)). The most dramatic differential expression, however, was observed for a ribosomal protein gene not yet connected to cold response (L34; Kole_0258). L34, a bacteria-specific ribosomal protein hypothesized to be a relatively recent addition to the evolving ribosome (30), is required for proper ribosome formation (31). A *Bacillus subtilis* mutant lacking the L34 gene showed slow growth at low temperature (32), suggesting a role for L34 in this condition. Many ribosomal proteins are recruited for extra-ribosomal functions (33), hence some of the up-regulated ribosomal proteins may have alternative roles in response to low temperature that are unrelated to the ribosome itself. However, genes encoding ribosomal RNA (rRNA) methyltransferases, rmlH (Kole_1718) and rmlL (Kole_0897), were also significantly up-regulated, and methylation of rRNAs has been associated with responses to environmental stress, including temperature (34). Combined with observations that ribosomes need to be fine-tuned to function properly at low temperature (7), we hypothesize that *K. olearia* modifies its ribosome by changing stoichiometry of its components and by methylating rRNA. Time required for such ribosomal adjustments could also explain the longer lag phase following temperature shifts (Fig. S1 in the supplemental material).

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To detect a decrease in environmental temperature and elicit an appropriate regulatory response, some bacteria have evolved two-component cold sensors (27). These signal transduction systems consist of a sensor, a membrane-integrated protein with a kinase domain that detects changes in the fluidity of the cell membrane, and the cytoplasmic response regulator, a protein that induces expression of cold-responsive genes. In *K. olearia*, a histidine kinase with two predicted transmembrane domains (Kole_1017) and two response regulators (Kole_1015 and Kole_1016) showed a steady increase in expression as temperatures decreased from 65°C, but no significant change in expression at 77°C (Table S5 in Dataset S1 in the supplemental material), leading us to hypothesize that these genes encode a cold-sensing two-component system.

iii) Increased amino acid metabolism at sub-optimal temperatures. At lower growth temperatures (and especially at and below 30°C) we observed an over-representation of genes involved in amino acid metabolism (COG category E). At 30°C, and to a lesser extent at 40°C, a peptide ABC transporter gene (Kole_2046 – Kole_2050) and several genes in the arginine (Kole_0092 – Kole_0097) and lysine (Kole_0104 – Kole_0107, 30°C only) biosynthesis pathways were up-regulated, suggesting the potential for accumulation of peptides and amino acids (or their intermediates) at lower temperatures. At 30°C there was also significant up-regulation of a citrate synthase gene (Kole_1230). Intriguingly, in *Staphylococcus aureus* citrate was shown to accumulate during prolonged cold stress (29), which could also be the case for *K. olearia*. Alternatively, citrate synthase, together with isocitrate dehydrogenase (Kole_1227), may be involved in converting pyruvate or acetyl-CoA to 2-oxoglutarate, a precursor for several amino acids including arginine. Accumulation of both arginine and lysine was observed during low temperature growth of *Clostridium botulinum*, where these amino acids were suggested to

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

act as compatible solutes (35). Interestingly, while the cells may accumulate peptides at 30°C, at 40°C there was increased expression of an oligo-peptidase (Kole_1190) and genes involved in lysine degradation (Kole_0958, Kole_0963 – Kole_0966). Such distinguishably different metabolic responses to moderately low (40°C) and low (≤30°C) temperatures suggest a fine-tuned temperature-dependent peptide turnover.

Two paralogs of ornithine carbamoyl-transferase (*ArgF*; Kole_1433 and Kole_2071) showed significantly lower expression at both 40°C and 30°C. The amino acid ornithine is an intermediate of arginine synthesis and lower expression of *ArgF* suggests that ornithine, rather than arginine, may accumulate at sub-optimal temperatures. However, ornithine has also been implicated in biofilm formation and species cross-talk (36), suggesting a possible alternative role of this amino acid in cellular responses to low temperature. Our unpublished observation of increased clumping of *K. olearia* cells grown at 30°C indirectly supports this hypothesis.

Re-modelling of amino acid metabolism at low temperatures has also been observed in other bacteria (e.g., (35, 37)). Interestingly, the genome of *M. prima* encodes more genes involved in amino acid metabolism than the genomes of *K. olearia* and other Thermotogae (16). Perhaps the mesophilic *Mesotoga* spp. have adapted to the increased need for peptides by expanding existing and acquiring new amino acid metabolism gene families. Amino acid metabolism genes are also among the most numerous bacterial genes laterally acquired by mesophilic Archaea, which was hypothesized to reflect their adaptation to low temperature growth (19).

iv) *K. olearia* is in heat stress at 77°C. Both the multivariate (Fig. 3) and clustering (Fig. S2 in the supplemental material) analyses showed that the 65°C and 77°C transcriptomes are distinct. Since 77°C is almost the upper limit for *K. olearia* growth, we hypothesize that the

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

observed differences in expression profiles at 77°C reflect a cell-wide heat stress response. Of the 169 differentially expressed genes, 119 showed increased expression at 77°C (Table S5 in Dataset S1 in the supplemental material). Hypothetical proteins made up a sizeable fraction (41 genes; 34%) of the 119 genes, indicating that adaptation to growth at sustained high temperature remains largely uncharacterized. Only one of the known heat shock response genes (9), the extreme heat stress sigma factor-24 (rpoE, Kole_2150), was up-regulated. Among the most highly expressed genes were the structural RNAs ffs (Kole_R0010), ssrA (Kole_R0006), and rnpB (Kole_R0049) (Fig. 3), suggesting an increased rate of RNA turnover at supra-optimal temperature. As mentioned earlier, carbohydrate and energy metabolism genes (COG category C and G) were also up-regulated. It is unclear, however, if the underlying cause is the increased turnover of enzymes at elevated temperatures, or a demand for more enzymes due to increased carbohydrate catabolism. Notably, two genes predicted to produce amino sugars for cell surface polysaccharides (Kole_1281 and Kole_1282) were also moderately up-regulated at 40°C, suggesting that cell surface polysaccharides may contribute to temperature adaptation at both sub- and supra-optimal temperatures. Increased carbohydrate metabolism in response to prolonged growth at supra-optimal temperature has been observed previously in *T. maritima* (10) and therefore may be a common adaptation to high temperature growth in the Thermotogae. The prolonged supra-optimal temperature growth of *T. maritima* also did not involve up-regulation of typical heat-shock response proteins (10). This highlights the difference between cellular response to an immediate heat-shock and to prolonged growth at supra-optimal temperature, and in general justifies classifying the cellular response to temperature into these two distinct categories.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

v) General stress response genes. Since we hypothesize that at 77°C and 30°C *K. olearia* cells are under stress, genes that are significantly up-regulated at both temperatures are candidates for a general temperature-stress response. There are 25 such genes, three of which were also significantly up-regulated at 40°C (Table S5 in Dataset S1 in the supplemental material). Among the most highly expressed of the 25 genes were Kole_2091, a gene with a distantly related homolog in only two other Thermotogae, and Kole_0418, a gene that within the Thermotogae has homologs only in *Mesotoga* spp. Both genes encode proteins of unknown function. Given such limited distribution within the Thermotogae, these genes may be involved in *Kosmotoga*-specific adaptation to a wide growth temperature range.

Conservation of *K. olearia*'s temperature-responsive genes across Kosmotogales. All genes that are required for adaptation and response of *K. olearia* to a wide range of growth temperatures are expected to be present in other *K. olearia* isolates, whereas some may be absent from *Kosmotoga* species having a narrower spectrum of growth temperature. Therefore, we compared the *K. olearia* genome to the genomes of *Kosmotoga* sp. DU53 and *Kosmotoga arenicorallina* (38). *Kosmotoga* sp. DU53 has a similar growth temperature range (observed range 25°C - 79°C, Table S7 in Dataset S1 in the supplemental material) and >99% average nucleotide identity (ANI) when compared to *K. olearia*, while *K. arenicorallina* exhibits a narrower growth temperature range (observed range 35°C - 70°C, Table S7 in Dataset S1 in the supplemental material) and has only 84% ANI when compared to *K. olearia.*

Indeed, the *Kosmotoga* sp. DU53 genome lacks only 10 of the 573 *K. olearia* putative temperature-responsive genes (BLASTP and TBLASTN searches, E-value < 10^{-3}, Table S5 in Dataset S1 in the supplemental material). All 10 genes were expressed in *K. olearia* at relatively low levels (the highest average expression value of 453 is for Kole_0200 at 77°C), suggesting

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

that they are unlikely to be essential for high or low temperature growth. On the other hand, the *K. arenicorallina* genome does not have detectable homologs of 103 of the 573 putative temperature-responsive genes in *K. olearia* (BLASTP and TBLASTN searches, E-value < 10^{-3}) (Table S5 in Dataset S1 in the supplemental material). The list of absent genes includes several of the arginine and lysine biosynthesis genes that are up-regulated in *K. olearia* during growth at ≤30°C, and seven of the hypothetical proteins up-regulated at 77°C. Therefore, we hypothesize that these 103 genes play a role in extending the growth range of *K. olearia* to ≤35°C and ≥70°C.

Role of lateral gene transfer in thermoadaptation of *K. olearia*. Obtaining "pre-adapted" genes from other genomes is one way prokaryotes adjust to new environmental conditions (18). Using HGTector (39) we predicted that 354 of *K. olearia*'s 2,118 protein coding genes have been acquired laterally by *K. olearia* or the Kosmotogales (i.e., *Kosmotoga* and *Mesotoga*), presumably representing LGT events occurring after the divergence of Kosmotogales from other Thermotogae (Table S8 in Dataset S1 in the supplemental material). Eighty-eight of the 354 genes were temperature responsive (Table S5 in Dataset S1, Fig. S5A and S5B in the supplemental material), including several already discussed highly expressed genes (Table 1 and Table S5 in Dataset S1 in the supplemental material). Notably, LGT appears to be especially important in *K. olearia*'s adaptation to the lower growth temperatures. Thirty-eight of the 88 temperature-responsive laterally acquired genes are shared with the strictly mesophilic *Mesotoga*, and most of them were highly expressed at lower temperatures, 30°C in particular (Fig. S5C in the supplemental material). Among these are the previously discussed rRNA methyltransferases (Kole_1718 and Kole_0897). The fatty acid synthesis genes (Kole_0969-Kole_0973) that are up-regulated at 40°C, as well as their Kosmotogales and *Mesotoga* homologs, form a distantly related sister clade to other Thermotogae lineages (Fig. S6A in the

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

supplemental material), suggesting that these genes may have been acquired from an un-sampled lineage. Similarly, the Csp-encoding gene highly expressed at 30°C (Kole_0109) is placed outside of the Thermotogae clade (Fig. S6B in the supplemental material). Predicted acquisition of the fatty acid synthesis and Csp genes by (now mesophilic) Archaea (19) additionally argues for the importance of these genes in adaptation to low temperature growth.

It is notable that some putative lateral acquisitions by *K. olearia* do not have homologs in *Mesotoga*. These include genes encoding the predicted cold temperature sensor (Kole_1015 – Kole_1017), one of the PPIase genes (Kole_1745), as well as the canonical cold response enzyme DEAD/DEAH box RNA helicase (Kole_0922). Lack of these genes in *Mesotoga* suggests their potential importance for *K. olearia*'s ability to grow over a wide temperature range.

Role of gene family expansion and lineage-specific gene evolution in thermoadaptation. Expansion of cold-responsive gene families may represent a common strategy for low temperature adaptation, as has been noted in many bacteria, especially in psychrophiles (e.g. (40)). *K. olearia* exhibits the same trend. For example, when compared to other Thermotogae, all three analyzed *Kosmotoga* genomes harboured more copies of Csp-encoding genes (Table S9 in Dataset S1 in the supplemental material). Additionally, *K. olearia* has extra homologs (Kole_0111 and Kole_0110) of the putative cold sensor system discussed above. The observed gene family expansions might be important not only for low temperature growth, but also for growth over a wide temperature interval. For example, *Mesotoga* functions with only a single Csp gene, demonstrating that having more copies of this gene is not required for low temperature growth. Having several copies of these genes, however, allow *K. olearia* the opportunity to regulate them differently at different temperatures. Similarly, the additional cold

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sensor homologs do not show co-ordinated temperature response: Kole_0110 is up-regulated at 40°C, while Kole_0111 is up-regulated at 77°C (Table 1). Therefore, these additional homologs may represent sensors tuned to different temperatures.

Gene family expansions can be achieved via within-lineage gene duplication or through LGT. A combination of these mechanisms appears to be at work in *K. olearia*, as demonstrated by the phylogenetic analyses of Csp genes (Fig. S6B in the supplemental material). Similarly, even though several Thermotogae genomes contain as many copies of PPIase genes as do *Kosmotoga* genomes (Table S9 in Dataset S1 in the supplemental material), phylogenetic analysis suggests that in the Kosmotogales this gene family has only recently been expanded by both LGT (the FKBP-type, Table 1) and duplication (the PpiC-type, Fig. S6C in the supplemental material).

However, the role of within-lineage evolution of specific genes in response to changing environmental conditions should not be neglected. For example, typical cold response genes RbfA (Kole_2103) and NusA (Kole_1529) were not laterally acquired, but nevertheless show high expression only at 30°C. Deciphering adaptive changes that occurred in such genes compared to thermophilic homologs may elucidate molecular mechanisms of low temperature adaptation.

Why maintain the capacity for growth over such a wide temperature range? Most bacteria are under selection to eradicate extraneous DNA (and genes) from their genomes (41), and among free-living bacteria Thermotogae in general have very compact genomes. Kosmotogales, however, have notably larger genomes than other thermophilic Thermotogae (8, 16), raising the possibility that expanded genomes are advantageous in *K. olearia*'s habitat. As discussed above, many of the genes in *K. olearia*, such as the cold-sensor system, were

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expressed only at specific sub- or supra-optimal temperatures, but do not seem to be important for growth at other temperatures (Table 1 and Table S5 in Dataset S1 in the supplemental material). The regulated response to low temperatures and the preservation of the laterally acquired genes specifically expressed at 40°C and 30°C suggest that *K. olearia* encounters environments with very different temperatures frequently enough to maintain these genes in its genome. Such environments may include oil reservoirs located at different depths, as well as marine sediments influenced by the mixing of cold deep sea water and hydrothermal fluids (42). As a result, this lineage was likely selected to become a temperature generalist. This conjecture is supported by the environmental conditions of the subsurface environments and marine hydrothermal vents from which *Kosmotoga* spp. have been isolated (6, 43, 44). *K. olearia* was isolated from a deep subsurface oil reservoir with *in situ* temperature of 68°C (6), but its 16S rRNA sequences also have been detected in many oil fields having *in situ* temperatures of 20°C–50°C (45). *Kosmotoga* sp. DU53, which is most similar to *K. olearia*, was isolated from an oil reservoir with an *in situ* temperature of ~50°C, while *K. arenicorallina* was isolated from hydrothermal sediments with a temperature of ~40°C (43). Notably, *K. olearia* was also identified as a major constituent in a metagenome from a deep subsurface oil reservoir with *in situ* temperature of 85°C and pressure of 25MPa (46). While the reservoir temperature is higher than the maximum *K. olearia* growth temperature reported here, elevated pressure could extend *K. olearia*'s temperature maximum, as has been demonstrated for some Archaea (e.g. (1)). Therefore, *K. olearia*'s growth temperature range under natural conditions may be even broader than 20-79°C.

Concluding Remarks. The present study demonstrates that a bacterium with a relatively small genome can use transcriptional changes to respond effectively to large changes in

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

temperature. We showed that *K. olearia*'s response to sustained exposure to a non-optimal temperature includes up-regulation of hundreds of genes. A substantial fraction of these genes have been acquired laterally, suggesting that LGT is an evolutionarily successful strategy for expansion of temperature tolerance; however, gene duplication and subsequent sub-functionalization of the paralogs also plays an important adaptive role.

The ability of *K. olearia* to inhabit both high and low temperature environments suggests that members of this lineage encounter environments with large temperature fluctuations and/or frequently migrate across ecological niches within the deep biosphere (e.g., between deep and shallow subsurface oil reservoirs). Therefore, the subsurface environments, as well as their microbial populations, might be viewed as a connected archipelago instead of isolated islands. As a corollary, we speculate that *K. olearia*-like ecological generalists could also facilitate LGT among seemingly isolated deep biosphere microbial communities adapted to a narrower ecological niche. For example, we have previously demonstrated high levels of gene flow among hyperthermophilic *Thermotoga* populations in subsurface oil reservoirs and marine hydrothermal vents (47), environments that are separated by non-thermophilic surroundings. The mechanism of such gene flow is not yet known, but *K. olearia*-like Thermotogae capable of growing both in subsurface oil reservoirs and adjacent marine sediments could serve as mediators of gene exchange.

Although some of the identified 573 temperature-responsive genes are already known to be expressed in Bacteria and Archaea grown at high or low temperatures, most of the up-regulated genes have not previously been implicated in temperature response and are in need of better functional and biochemical characterization. For example, the majority of the *K. olearia* genes responsive to elevated temperature encode proteins of unknown functions. Versatile

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

proteins that work across a broad range of temperatures also warrant further biochemical and evolutionary analyses, as understanding of their enzymatic flexibility can aid the design of commercially important thermostable proteins.

Materials and Methods

Bacterial culturing, and RNA and DNA isolation. *K. olearia* was grown at different temperatures (4°C, 25°C, 30°C, 40°C, 65°C, and 77°C), but otherwise optimal conditions, as described in (6) and Text S1 in the supplemental material.

For each temperature treatment, RNA was extracted in either mid-log phase or late-log phase, using the Zymo Research Fungal/Bacterial RNA MiniPrep Kit (Cedarlane Laboratories, Ltd.; Burlington, Ontario) and following the manufacturer's protocols (Table S1 in Dataset S1 in the supplemental material). The extracted RNA was sequenced on either an Ion Torrent PGM (RNA-Seq kit V2) or an Illumina MiSeq (TruSeq RNASeq v2 2x100 bp) from the libraries constructed following the manufacturer's instructions (Table S1 in Dataset S1 in the supplemental material). The transcriptomes are available in the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under the accession number SRP075860.

RNA-Seq analysis. For each transcriptome, sequenced reads were analyzed using the RNA-Seq module in CLC Genomics Workbench version 7.0.4 (http://www.clcbio.com/, CLC bio, Århus, Denmark), resulting in RPKM (Reads Per Kilobase of transcript per Million mapped reads) values for each gene, as described in *SI Materials and Methods*. RPKM values for all genes are listed in Table S4 in Dataset S1 in the supplemental material. Differentially expressed genes were identified by doing pairwise comparisons of the transcriptomes of the isothermically grown cultures at 30°C, 40°C, and 77°C to the cultures grown at the optimal temperature of

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

65°C. The analyses used the "Empirical Analysis of DGE" function, which employs the "Exact Test" for two-group comparisons (48). A gene was considered differentially expressed in a pairwise comparison if it had (*i*) > 20 reads in at least one of the two transcriptomes, (*ii*) a statistically significant difference in the RPKM values (corrected for multiple testing using False Discovery Rate [FDR] < 0.05), and (*iii*) a difference in RPKM values at least two-fold in magnitude. Principal Component Analysis (PCA) and biplot visualization were performed using R packages *ade4* and *bpca* respectively (49, 50). Transcription start and stop sites and number of transcripts were predicted using Rockhopper (51). For detailed descriptions see Text S1 in the supplemental material.

Comparative analyses of three *Kosmotoga* spp. genomes. The genome of *K. olearia* (accession number CP001634, (15)) was compared to genomes of *Kosmotoga* sp. DU53 (accession number JFHK00000000) and *K. arenicorallina* (accession number JGCK00000000) (38) using the IMG portal (52) and Geneious v.9. Pairwise Average Nucleotide Identity (ANI) (53) was calculated using the Enveomics Toolbox (54). Protein-coding genes in each genome were classified as putatively laterally transferred using a customized version of HGTector (39). For detailed descriptions see Text S1 in the supplemental material.

Data availability: The transcriptomics data is available in NCBI's Public Sequence Read Archive (SRA) database under the accession number SRP075860.

Funding information

This work was supported by a Research council of Norway award (project no. 180444/V40) to C.L.N.; a Simons Investigator award from the Simons Foundation, a Dartmouth College Walter

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and Constance Burke Research Initiation Award, and Dartmouth College start-up funds to O.Z.; and an NSERC Alexander Graham Bell Canada Graduate Scholarship CGS-M to S.M.J.P. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Acknowledgements

We thank Rhianna Charchuk for technical assistance, Dr. Amanda J. Lohan for advice and discussion of RNA-Seq methodology, Dr. Karen Budwill for head-space gas analyses, and Jennifer Franks for initial discussions of lateral gene transfer detection methods.

References

1. **Takai, K, Nakamura, K, Toki, T, Tsunogai, U, Miyazaki, M, Miyazaki, J, Hirayama, H, Nakagawa, S, Nunoura, T, Horikoshi, K.** 2008. Cell proliferation at 122°C and isotopically heavy CH_4 production by a hyperthermophilic methanogen under high-pressure cultivation. Proc. Natl. Acad. Sci. U. S. A. **105:**10949-10954. doi: 10.1073/pnas.0712334105.

2. **Mykytczuk, NCS, Foote, SJ, Omelon, CR, Southam, G, Greer, CW, Whyte, LG.** 2013. Bacterial growth at -15°C; molecular insights from the permafrost bacterium *Planococcus halocryophilus* Or1. The ISME Journal. **7:**1211-1226. doi: 10.1038/ismej.2013.8.

3. **Schumann, W.** 2009. Temperature sensors of eubacteria. Advances in Applied Microbiology. **67:**213-256. doi: 10.1016/S0065-2164(08)01007-1.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

4. **Raghubeer, EV, Matches, JR.** 1990. Temperature range for growth of *Escherichia coli* serotype 0157:H7 and selected coliforms in *E. coli* medium. J. Clin. Microbiol. **28:**803-805.

5. **Dinsdale, AE, Halket, G, Coorevits, A, van Landschoot, A, Busse, HJ, de Vos, P, Logan, NA.** 2011. Emended descriptions of *Geobacillus thermoleovorans* and *Geobacillus thermocatenulatus*. Int. J. Syst. Evol. Microbiol. **61:**1802-1810. doi: 10.1099/ijs.0.025445-0.

6. **DiPippo, JL, Nesbø, CL, Dahle, H, Doolittle, WF, Birkland, N, 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. doi: 10.1099/ijs.0.008045-0.

7. **Barria, C, Malecki, M, Arraiano, CM.** 2013. Bacterial adaptation to cold. Microbiology. **159:**2437-2443. doi: 10.1099/mic.0.052209-0.

8. **Pollo, SMJ, Zhaxybayeva, O, Nesbø, CL.** 2015. Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum *Thermotogae*. Can. J. Microbiol. **61:**655-670.

9. **Pysz, MA, Ward, DE, Shockley, KR, Montero, CI, Conners, SB, Johnson, MR, Kelly, RM.** 2004. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium *Thermotoga maritima*. Extremophiles. **8:**209-217. doi: 10.1007/s00792-004-0379-2.

10. **Wang, Z, Tong, W, Wang, Q, Bai, X, Chen, Z, Zhao, J, Xu, N, Liu, S.** 2012. The temperature dependent proteomic analysis of *Thermotoga maritima*. PloS One. **7:**e46463. doi: 10.1371/journal.pone.0046463.

11. **Phadtare, S.** 2004. Recent developments in bacterial cold-shock response. Curr. Issues Mol. Biol. **6:**125-136.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

12. **Phadtare, S, Hwang, J, Severinov, K, Inouye, M.** 2003. CspB and CspL, thermostable cold-shock proteins from *Thermotoga maritima*. Genes to Cells. **8:**801-810. doi: 10.1046/j.1365-2443.2003.00675.x.

13. **Boonyaratanakornkit, BB, Simpson, AJ, Whitehead, TA, Fraser, CM, El-Sayed, NMA, Clark, DS.** 2005. Transcriptional profiling of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* in response to lethal heat and non-lethal cold shock. Environ. Microbiol. **7:**789-797. doi: 10.1111/j.1462-2920.2005.00751.x.

14. **Mega, R, Manzoku, M, Shinkai, A, Nakagawa, N, Kuramitsu, S, Masui, R.** 2010. Very rapid induction of a cold shock protein by temperature downshift in *Thermus thermophilus.* Biochem. Biophys. Res. Commun. **399:**336-340. doi: 10.1016/j.bbrc.2010.07.065.

15. **Swithers, KS, DiPippo, JL, Bruce, DC, Detter, C, Tapia, R, Han, S, Goodwin, LA, Han, J, Woyke, T, Pitluck, S, Pennacchio, L, Nolan, M, Mikhailova, N, Land, ML, Nesbø, CL, Gogarten, JP, Noll, KM.** 2011. Genome sequence of *Kosmotoga olearia* strain TBF 19.5.1, a thermophilic bacterium with a wide growth temperature range, isolated from the Troll B oil platform in the North Sea. J. Bacteriol. **193:**5566-5567. doi: 10.1128/JB.05828-11.

16. **Zhaxybayeva, O, Swithers, KS, Foght, J, Green, AG, Bruce, D, Detter, C, Han, S, Teshima, H, Han, J, Woyke, T, Pitluck, S, Nolan, M, Ivanova, N, Pati, A, Land, ML, Dlutek, M, Doolittle, WF, Noll, KM, Nesbø, CL.** 2012. Genome sequence of the mesophilic Thermotogales bacterium *Mesotoga prima* MesG1.Ag.4.2 reveals the largest Thermotogales genome to date. Genome Biol. Evol. **4:**700-708. doi: 10.1093/gbe/evs059.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

17. **Zhaxybayeva, O, Swithers, KS, Lapierre, P, Fournier, GP, Bickhart, DM, DeBoy, RT, Nelson, KE, Nesbø, CL, Doolittle, WF, Gogarten, JP, Noll, KM.** 2009. On the chimeric nature, thermophilic origin, and phylogenetic placement of the *Thermotogales*. Proc. Natl. Acad. Sci. U. S. A. **106:**5865-5870. doi: 10.1073/pnas.0901260106.

18. **Boucher, Y, Douady, CJ, Papke, RT, Walsh, DA, Boudreau, MER, Nesbø, CL, Case, RJ, Doolittle, WF.** 2003. Lateral gene transfer and the origins of prokaryotic groups. Annu. Rev. Genet. **37:**283-328. doi: 10.1146/annurev.genet.37.050503.084247.

19. **López-García, P, Zivanovic, Y, Deschamps, P, Moreira, D.** 2015. Bacterial gene import and mesophilic adaptation in archaea. Nature Reviews.Microbiology. **13:**447-456. doi: 10.1038/nrmicro3485.

20. **Ratkowsky, DA, Lowry, RK, McMeekin, TA, Stokes, AN, Chandler, RE.** 1983. Model for bacterial culture growth rate throughout the entire biokinetic temperature range. J. Bacteriol. **154:**1222-1226.

21. **Swinnen, IAM, Bernaerts, K, Dens, EJJ, Geeraerd, AH, Van Impe J. F.** 2004. Predictive modelling of the microbial lag phase: A review. Int. J. Food Microbiol. **94:**137-159. doi: 10.1016/j.ijfoodmicro.2004.01.006.

22. **Latif, H, Lerman, JA, Portnoy, VA, Tarasova, Y, Nagarajan, H, Schrimpe-Rutledge, AC, Smith, RD, Adkins, JN, Lee, D, Qiu, Y, Zengler, K.** 2013. The genome organization of *Thermotoga maritima* reflects its lifestyle. PLoS Genetics. **9:**e1003485. doi: 10.1371/journal.pgen.1003485.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

23. **Cho, B, Zengler, K, Qiu, Y, Park, YS, Knight, EM, Barrett, CL, Gao, Y, Palsson, BØ.** 2009. The transcription unit architecture of the *Escherichia coli* genome. Nat. Biotechnol. **27:**1043-1049. doi: 10.1038/nbt.1582.

24. **Qiu, Y, Cho, BK, Park, YS, Lovley, D, Palsson, B, Zengler, K.** 2010. Structural and operational complexity of the *Geobacter sulfurreducens* genome. Genome Res. **20:**1304-1311. doi: 10.1101/gr.107540.110.

25. **Nanavati, DM, Thirangoon, K, Noll, KM.** 2006. Several archaeal homologs of putative oligopeptide-binding proteins encoded by *Thermotoga maritima* bind sugars. Appl. Environ. Microbiol. **72:**1336-1345. doi: 10.1128/AEM.72.2.1336-1345.2006.

26. **Foster, PL.** 2007. Stress-induced mutagenesis in bacteria. Crit Rev Biochem Mol Biol. **42:**373-397. doi: 10.1080/10409230701648494.Stress-Induced.

27. **de Mendoza, D.** 2014. Temperature sensing by membranes. Annu. Rev. Microbiol. **68:**101-116. doi: 10.1146/annurev-micro-091313-103612.

28. **Godin-Roulling, A, Schmidpeter, PAM, Schmid, FX, Feller, G.** 2015. Functional adaptations of the bacterial chaperone trigger factor to extreme environmental temperatures. Environ. Microbiol. **17:**2407-2420. doi: 10.1111/1462-2920.12707.

29. **Alreshidi, MM, Dunstan, RH, Macdonald, MM, Smith, ND, Gottfries, J, Roberts, TK.** 2015. Metabolomic and proteomic responses of *Staphylococcus aureus* to prolonged cold stress. Journal of Proteomics. **121:**44-55. doi: 10.1016/j.jprot.2015.03.010.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

30. **Fox, GE.** 2010. Origin and evolution of the ribosome, p. 2:a003483. *In* D. Deamer and J. W. Szostak (eds.), The Origins of Life. Cold Spring Harb Perspect Biol.

31. **Akanuma, G, Kobayashi, A, Suzuki, S, Kawamura, F, Shiwa, Y, Watanabe, S, Yoshikawa, H, Hanai, R, Ishizuka, M.** 2014. Defect in the formation of 70S ribosomes caused by lack of ribosomal protein L34 can be suppressed by magnesium. J. Bacteriol. **196:**3820-3830. doi: 10.1128/JB.01896-14.

32. **Akanuma, G, Nanamiya, H, Natori, Y, Yano, K, Suzuki, S, Omata, S, Ishizuka, M, Sekine, Y, Kawamura, F.** 2012. Inactivation of ribosomal protein genes in *Bacillus subtilis* reveals importance of each ribosomal protein for cell proliferation and cell differentiation. J. Bacteriol. **194:**6282-6291. doi: 10.1128/JB.01544-12.

33. **Bhavsar, RB, Makley, LN, Tsonis, PA.** 2010. The other lives of ribosomal proteins. Hum Genomics. **4:**327-344. doi: 10.1186/1479-7364-4-5-327.

34. **Baldridge, KC, Contreras, LM.** 2014. Functional implications of ribosomal RNA methylation in response to environmental stress. Crit. Rev. Biochem. Mol. Biol. **49:**69-89. doi: 10.3109/10409238.2013.859229.

35. **Dahlsten, E, Isokallio, M, Somervuo, P, Lindström, M, Korkeala, H.** 2014. Transcriptomic analysis of (group I) *Clostridium botulinum* ATCC 3502 cold shock response. PloS One. **9:**e89958. doi: 10.1371/journal.pone.0089958.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

36. **Sakanaka, A, Kuboniwa, M, Takeuchi, H, Hashino, E, Amano, A.** 2015. Arginine-ornithine antiporter ArcD controls arginine metabolism and interspecies biofilm development of *Streptococcus gordonii*. J. Biol. Chem. **290:**21185-21198. doi: 10.1074/jbc.M115.644401.

37. **Ghobakhlou, A, Johnston, A, Harris, L, Antoun, H, Laberge, S.** 2015. Microarray transcriptional profiling of Arctic *Mesorhizobium* strain N33 at low temperature provides insights into cold adaption strategies. BMC Genomics. **16:**383-396. doi: 10.1186/s12864-015-1611-4.

38. **Pollo, SMJ, Charchuk, R, Nesbø, CL.** 2016. Draft genome sequences of *Kosmotoga* sp. DU53 and *Kosmotoga arenicorallina* S304. GenomeA. in press.

39. **Zhu, Q, Kosoy, M, Dittmar, K.** 2014. HGTector: An automated method facilitating genome-wide discovery of putative horizontal gene transfers. BMC Genomics. **15:**717. doi: 10.1186/1471-2164-15-717.

40. **Piette, F, D'Amico, S, Struvay, C, Mazzucchelli, G, Renaut, J, Tutino, ML, Danchin, A, Leprince, P, Feller, G.** 2010. Proteomics of life at low temperatures: trigger factor is the primary chaperone in the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. Mol. Microbiol. **76:**120-132. doi: 10.1111/j.1365-2958.2010.07084.x.

41. **Graur, D.** 2016. Prokaryotic Genome Evolution, p. 451. *In* Anonymous Molecular and Genome Evolution. Sinauer Associates, Sunderland, MA.

42. **Sievert, S, Vetriani, C.** 2012. Chemoautotrophy at Deep-Sea Vents: Past, Present, and Future. Oceanography. **25:**218-233. doi: 10.5670/oceanog.2012.21.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

43. **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. doi: 10.1007/s00203-010-0611-7.

44. **L'Haridon, S, Jiang, L, Alain, K, Chalopin, M, Rouxel, O, Beauverger, M, Xu, H, Shao, Z, Jebbar, M.** 2014. *Kosmotoga pacifica* sp. nov., a thermophilic chemoorganoheterotrophic bacterium isolated from an East Pacific hydrothermal sediment. Extremophiles. **18:**81-88. doi: 10.1007/s00792-013-0596-7.

45. **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. doi: 10.1128/AEM.02846-09.

46. **Kotlar, HK, Lewin, A, Johansen, J, Throne-Holst, M, Haverkamp, T, Markussen, S, Winnberg, A, Ringrose, P, Aakvik, T, Ryeng, E, Jakobsen, K, Drabløs, F, Valla, S.** 2011. High coverage sequencing of DNA from microorganisms living in an oil reservoir 2.5 kilometres subsurface. Environmental Microbiology Reports. **3:**674-681. doi: 10.1111/j.1758-2229.2011.00279.x.

47. **Nesbø, CL, Swithers, KS, Dahle, H, Haverkamp, TH, Birkeland, N, Sokolova, T, Kublanov, I, Zhaxybayeva, O.** 2015. Evidence for extensive gene flow and *Thermotoga*

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

subpopulations in subsurface and marine environments. The ISME Journal. **9:**1532-1542. doi: 10.1038/ismej.2014.238.

48. **Robinson, MD, Smyth, GK.** 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics. **9:**321-332. doi: 10.1093/biostatistics/kxm030.

49. **Dray, S, Dufour, AB, Chessel, D.** 2007. The ade4 Package - II : Two-table and K-table methods. R News. **7:**47-52. doi: 10.1159/000323281.

50. **Faria, JC, Demétrio, CGB, Allaman, IB.** 2013. bpca: Biplot of multivariate data based on principal components analysis. UESC and ESALQ, Ilheus, Bahia, Brasil and Piracicaba, Sao Paulo, Brasil.

51. **McClure, R, Balasubramanian, D, Sun, Y, Bobrovskyy, M, Sumby, P, Genco, Ca, Vanderpool, CK, Tjaden, B.** 2013. Computational analysis of bacterial RNA-Seq data. Nucleic Acids Res. **41:**e140. doi: 10.1093/nar/gkt444.

52. **Markowitz, VM, Chen, IA, Palaniappan, K, Chu, K, Szeto, E, Pillay, M, Ratner, A, Huang, J, Woyke, T, Huntemann, M, Anderson, I, Billis, K, Varghese, N, Mavromatis, K, Pati, A, Ivanova, NN, Kyrpides, NC.** 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res. **42:**D560-7. doi: 10.1093/nar/gkt963.

53. **Goris, J, Konstantinidis, KT, Klappenbach, JA, Coenye, T, Vandamme, P, Tiedje, JM.** 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. **57:**81-91. doi: 10.1099/ijs.0.64483-0.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

54. **Rodriguez-R, LM, Konstantinidis, KT.** 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. Peer J Preprints. **4:**e19001.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Legends:

FIG 1. Growth rate of *K. olearia* as a function of temperature. Isothermic growth curves were generated at each temperature from an inoculum grown at that temperature for at least three transfers (except for 25°C and 80°C, for which an inoculum from the same temperature could not be generated; see Text S1 in the supplemental material.). Up-shifted and down-shifted growth curves were generated from an inoculum that was grown at lower and higher temperatures, respectively. Red squares, growth temperature up-shifted from 65°C to 77°C or from 40°C to 65°C; Blue circles, growth temperature down-shifted from 77°C to 65°C, 65°C to 40°C, or 40°C to 30°C. Data points represent the mean of replicate cultures (see Text S1 in the supplemental material); error bars represent standard error.

FIG 2. Model of energy generation pathway in *K. olearia* during growth on pyruvate. The model includes genes likely involved in conversion of pyruvate to acetate, CO_2, H_2, and ATP. The genes were selected from the list of genes highly expressed across all temperature conditions (Table S3 in Dataset S1 in the supplemental material). Acetate transport is not shown. The dashed box indicates hydrogenase activity. The two highly expressed hydrogenases are shown, but their potential interactions with each other or with the membrane are not known. Increased expression of citrate synthase at low temperature, which could redirect acetyl-CoA away from acetate production, is shown in grey. The model also explains the observed lower ratio of carbon dioxide to hydrogen produced by growth on maltose vs. pyruvate (not shown), as during growth on maltose reduced electron carriers would be generated from the conversions of maltose to pyruvate.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

FIG 3. Biplot of the principal component analysis of 12 transcriptomes. Each transcriptome is denoted by a point, while genes are represented by vectors. Genes that point into a specific "temperature quadrant" are up-regulated at the growth temperature(s) of that quadrant, and the five longest (i.e., most highly correlated) gene vectors pointing to each quadrant are shown. Co-ordinates and vector length for all genes can be found in Table S4 in Dataset S1 in the supplemental material. It should be noted that the *ffs* (Kole_R0010) transcript is only 115 nt, and may not have been fully represented in every transcriptome due to our isolation protocol which selects against small RNA (<200 nucleotides). Also, the high expression of the alcohol dehydrogenase (Kole_0742) is probably due to the RNA isolation method (see Text S1 in the supplemental material.).

FIG 4. Difference between observed and expected number of temperature responsive genes across functional categories. Functional categories were assigned using the Clusters of Orthologous Groups (COG) database as implemented in IMG (52) and are denoted by one-letter abbreviations along the X-axis (see Fig. S3 legend for notations). NC, for "no category", denotes genes not assigned to a functional category. For each temperature treatment (30°C, 40°C and 77°C) only the temperature-responsive fraction of the *K. olearia* genome was considered. If the temperature-responsive genes were randomly distributed across functional categories we would expect the same fraction of temperature-responsive genes in each COG category. The difference (in percent) between the observed and expected number of temperature responsive genes is plotted on the Y-axis with positive and negative values referring to over- and under-

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

representation of the temperature-responsive genes, respectively. For actual number of genes in each COG category see Fig. S3 in the supplemental material.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TABLE 1. Gene expression in selected laterally acquired temperature-responsive genes. At each temperature, the listed RPKM values represent the average expression levels across replicates. Values that are significantly different from 65°C are shown in bold font.

Locus Tag	Functional annotation	30°C[a]	40°C	65°C	77°C	Identified by
Kole_0109	Cold shock protein	**5602**	**892**	222	119	Phylogenetic analysis
Kole_0110	Histidine kinase	175	**333**	144	312	Phylogenetic analysis
Kole_0111	Response regulator	166	204	173	**446**	HGTector
Kole_0505	Glycerol dehydrogenase	721	**2668**	752	1242	HGTector
Kole_0506	Hypothetical protein	559	**2037**	461	783	Phylogenetic analysis
Kole_0507	Hypothetical protein	555	**2193**	521	809	HGTector
Kole_0508	Poly (3-hydroxybutyrate) depolymerase-like protein	212	423	200	314	HGTector
Kole_0897	Ribosomal RNA methyltransferase, rmlL	**503**	498	228	232	HGTector
Kole_0922	DEAD/DEAH box helicase	**755**	**288**	89	102	HGTector
Kole_0969	3-oxoacyl-ACP synthase III , FabH	**2386**	**3063**	939	1424	HGTector
Kole_0970	enoyl-ACP reductase II , fabK	2226	3243	1486	1641	HGTector

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Kole_0971	malonyl CoA-acyl carrier protein transacylase, fabD	2304	4211	2303	2647	HGTector
Kole_0972	acyl carrier protein	6531	12601	4850	4241	HGTector
Kole_0973	3-oxoacyl-ACP synthase II, fabF	4815	9257	4753	4498	HGTector
Kole_1015	Response regulator	**1289**	**515**	95	130	HGTector
Kole_1016	Response regulator	**783**	**280**	54	72	HGTector
Kole_1017	Histidine kinase	**697**	**275**	59	90	Phylogenetic analysis
Kole_1281	N-acylneuraminate-9-phosphate synthase	482	699	315	**840**	HGTector
Kole_1282	N-acylneuraminate cytidylyltransferase	244	**283**	128	**315**	Phylogenetic analysis
Kole_1718	Ribosomal RNA methyltransferase, rmlH	**531**	332	211	203	HGTector
Kole_1745	PPIase FKBP-type	2783	2382	1541	**430**	HGTector

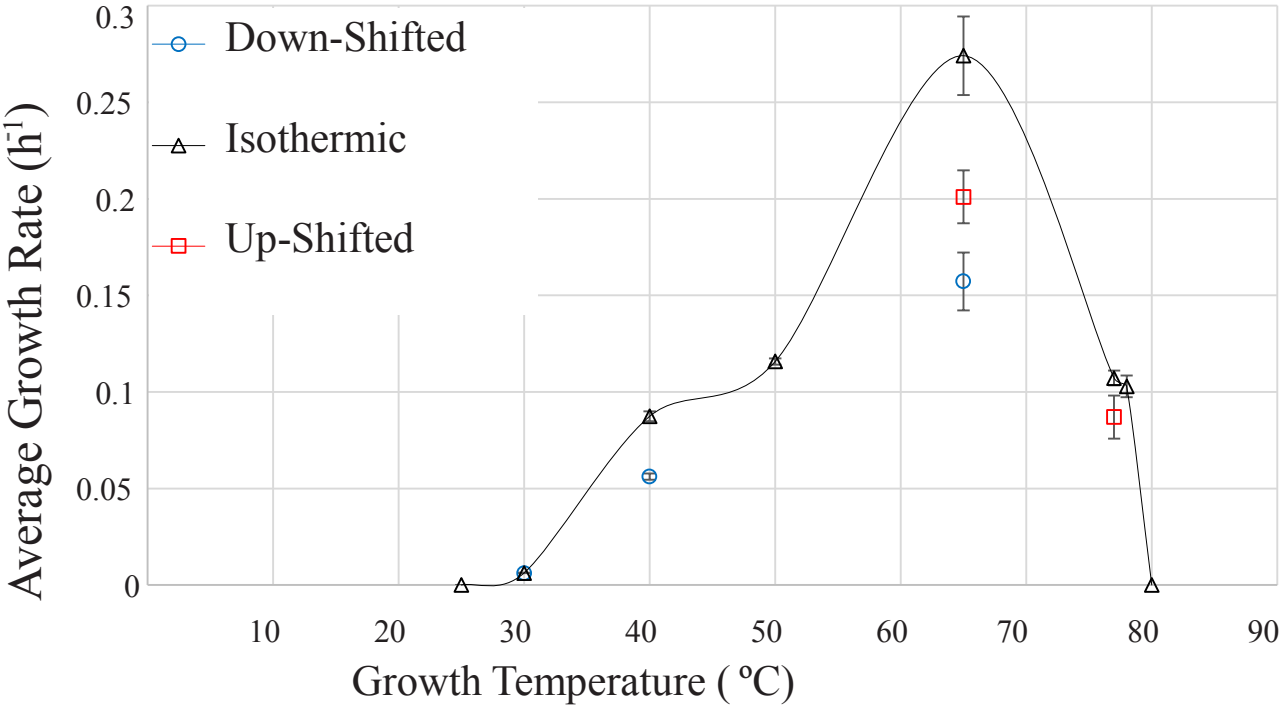
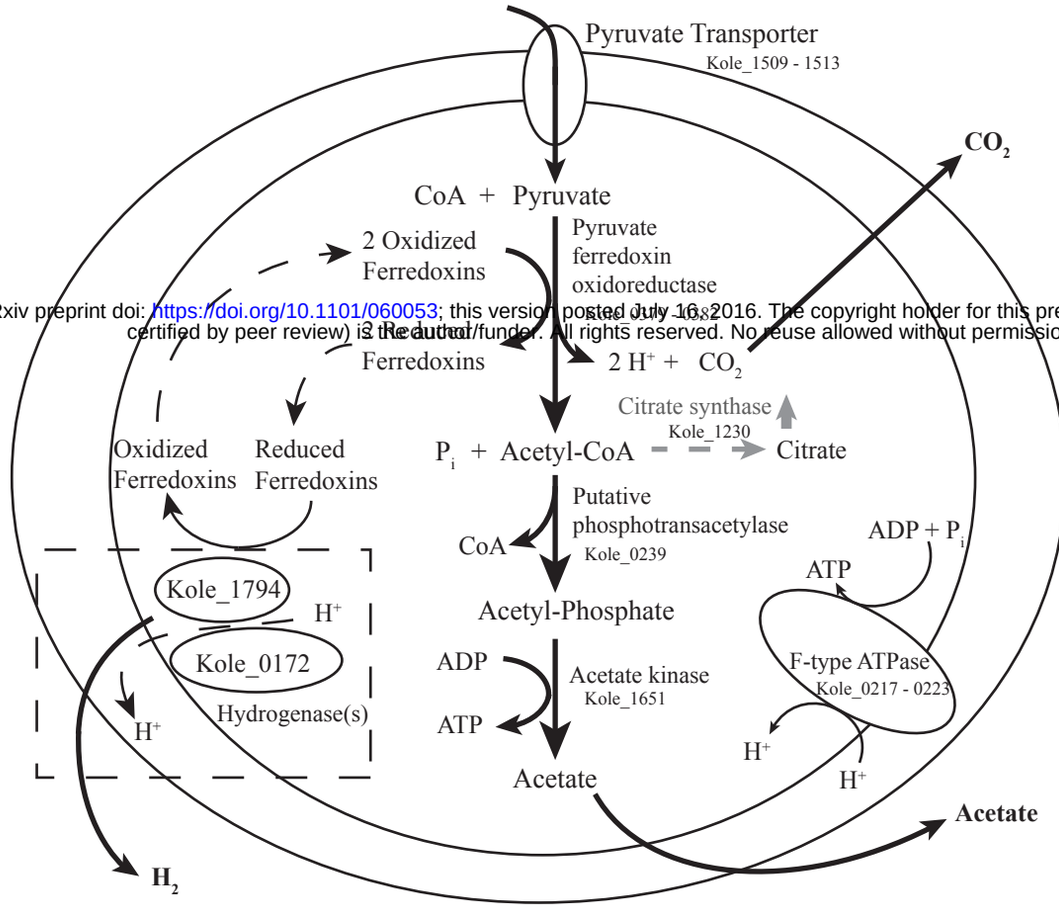


FIG 1. Growth rate of K. olearia as a function of temperature. Isothermic growth curves were generated at each temperature from an inoculum grown at that temperature for at least three transfers (except for 25°C and 80°C, for which an inoculum from the same temperature could not be generated; see Text S1 in the supplemental material.). Up-shifted and down-shifted growth curves were generated from an inoculum that was grown at lower and higher temperatures, respectively. Red squares, growth temperature up-shifted from 65°C to 77°C or from 40°C to 65°C; Blue circles, growth temperature down-shifted from 77°C to 65°C, 65°C to 40°C, or 40°C to 30°C. Data points represent the mean of replicate cultures (see Text S1 in the supplemental material); error bars represent standard error.



bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted July 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

FIG 2. Model of energy generation pathway in K. olearia during growth on pyruvate. The model includes genes likely involved in conversion of pyruvate to acetate, CO2, H2, and ATP. The genes were selected from the list of genes highly expressed across all temperature conditions (Table S3 in Dataset S1 in the supplemental material). Acetate transport is not shown. The dashed box indicates hydrogenase activity. The two highly expressed hydrogenases are shown, but their potential interactions with each other or with the membrane are not known. Increased expression of citrate synthase at low temperature, which could redirect acetyl-CoA away from acetate production, is shown in grey. The model also explains the observed lower ratio of carbon dioxide to hydrogen produced by growth on maltose vs. pyruvate (not shown), as during growth on maltose reduced electron carriers would be generated from the conversions of maltose to pyruvate.

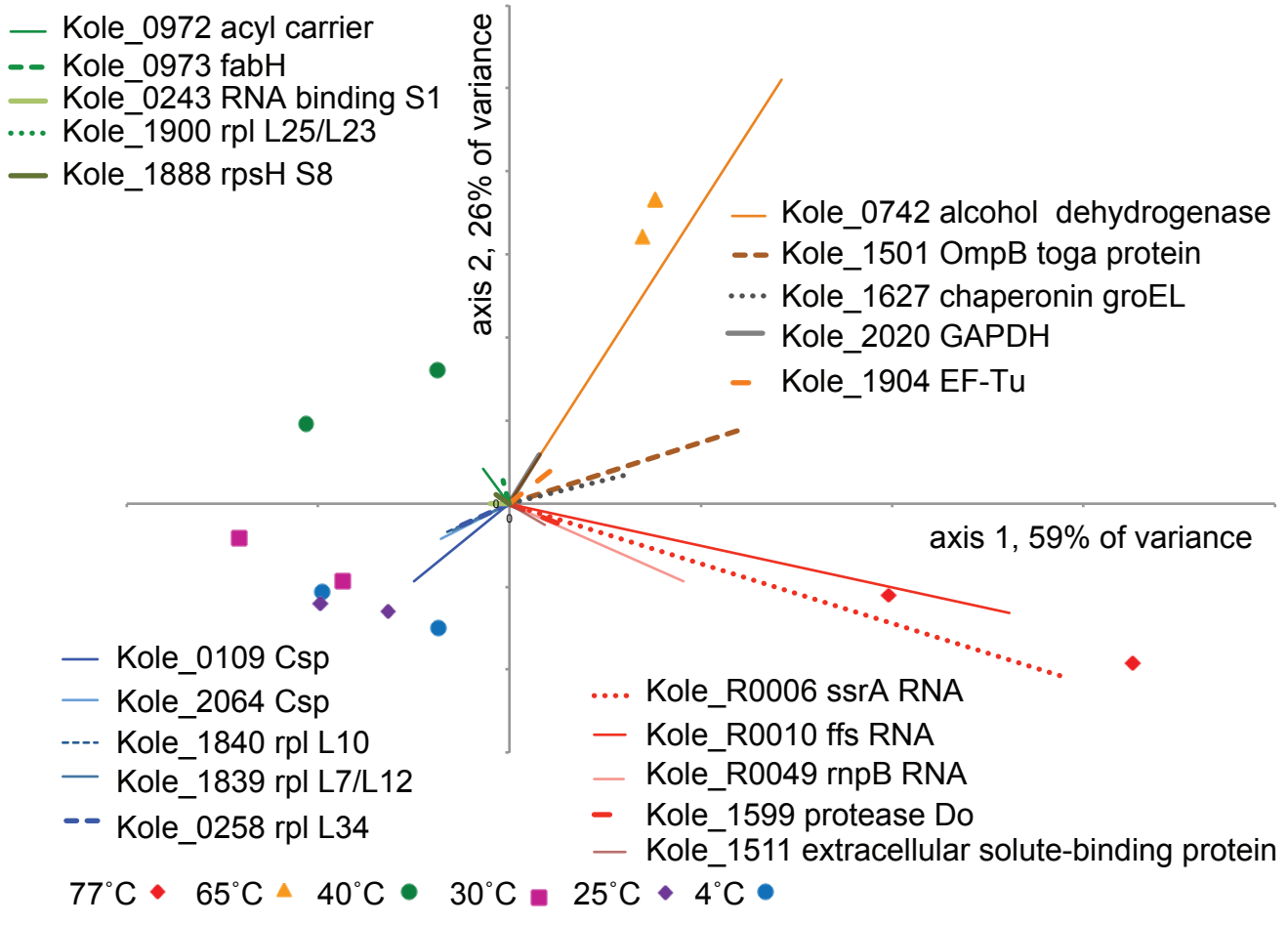


FIG 3. Biplot of the principal component analysis of 12 transcriptomes. Each transcriptome is denoted by a point, while genes are represented by vectors. Genes that point into a specific "temperature quadrant" are up-regulated at the growth temperature(s) of that quadrant, and the five longest (i.e., most highly correlated) gene vectors pointing to each quadrant are shown. Co-ordinates and vector length for all genes can be found in Table S4 in Dataset S1 in the supplemental material. It should be noted that the ffs (Kole_R0010) transcript is only 115 nt, and may not have been fully represented in every transcriptome due to our isolation protocol which selects against small RNA (<200 nucleotides). Also, the high expression of the alcohol dehydrogenase (Kole_0742) is probably due to the RNA isolation method (see Text S1 in the supplemental material.).

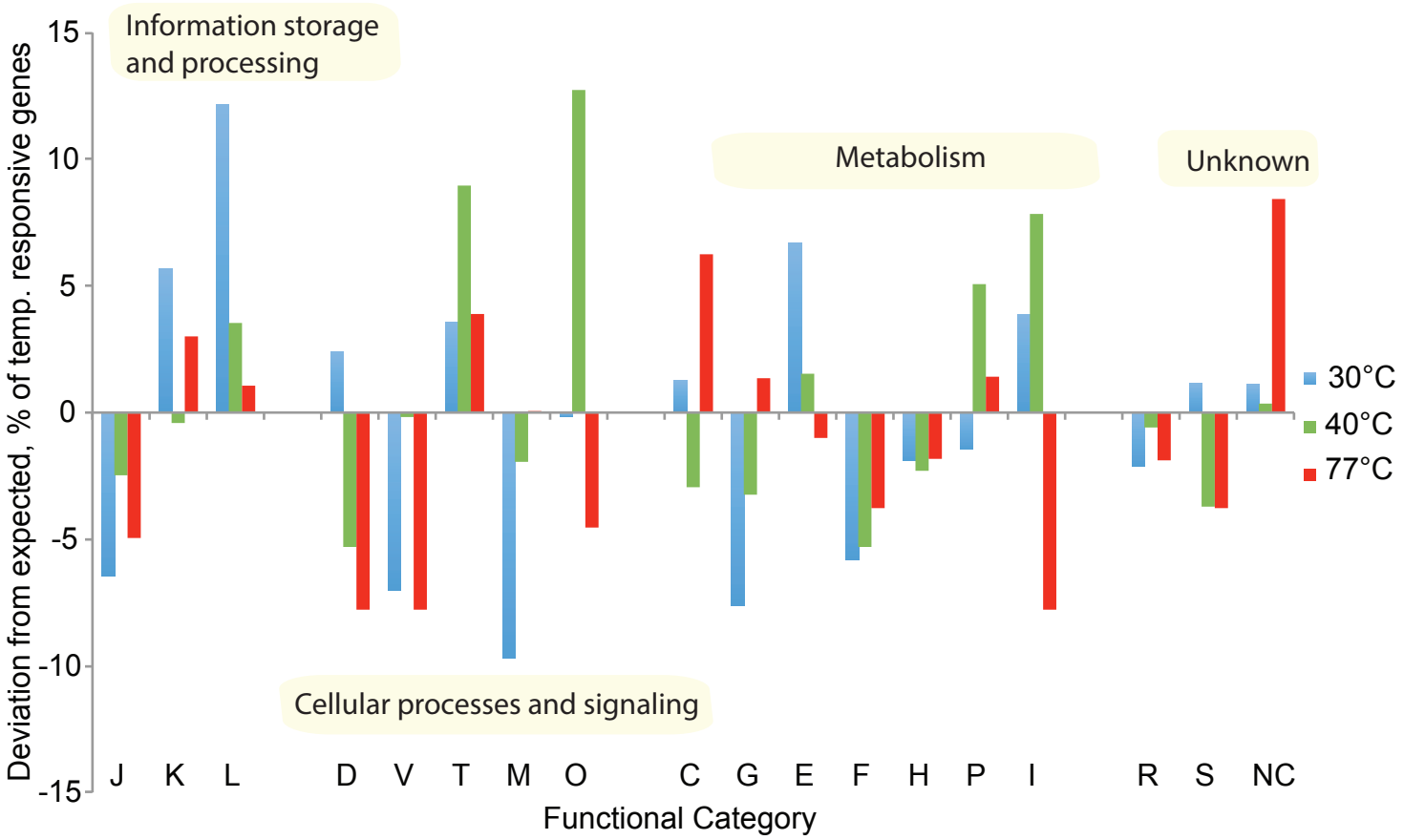


FIG 4. Difference between observed and expected number of temperature responsive genes across functional categories. Functional categories were assigned using the Clusters of Orthologous Groups (COG) database as implemented in IMG (52) and are denoted by one-letter abbreviations along the X-axis (see Fig. S3 legend for notations). NC, for "no category", denotes genes not assigned to a functional category. For each temperature treatment (30°C, 40°C and 77°C) only the temperature-responsive fraction of the K. olearia genome was considered. If the temperature-responsive genes were randomly distributed across functional categories we would expect the same fraction of temperature-responsive genes in each COG category. The difference (in percent) between the observed and expected number of temperature responsive genes is plotted on the Y-axis with positive and negative values referring to over- and under-representation of the temperature-responsive genes, respectively. For actual number of genes in each COG category see Fig. S3 in the supplemental material.