bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder.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.

1 # Genomic insights into temperature-dependent transcriptional
2 responses of *Kosmotoga olearia*, a deep-biosphere bacterium that can
3 grow from 20°C to 79°C

4 **Stephen M. J. Pollo[1], Abigail A. Adebusuyi[1#], Timothy J. Straub[2], Julia M. Foght[1], Olga**
5 **Zhaxybayeva[2,3] and Camilla L. Nesbø[1,4*]**

6 [1]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

7 [2]Department of Biological Sciences, Dartmouth College, Hanover, NH, USA

8 [3]Department of Computer Science, Dartmouth College, Hanover, NH, USA

9 [4]Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, Oslo,
10 Norway

11 *** Correspondence:**
12 Corresponding Author
13 nesbo@ualberta.ca

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

16

17 **Keywords: Thermotogae, RNA-Seq, lateral gene transfer, temperature response, cold response,**
18 **thermal adaptation**

19

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder.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.

20 **Abstract**

21 Temperature is one of the defining parameters of an ecological niche. Most organisms thrive within a
22 temperature range that rarely exceeds ~ 30°C, but the deep subsurface bacterium *Kosmotoga olearia*
23 can grow over a temperature range of ~60°C (20°C - 79°C). To identify genes correlated with this
24 flexible phenotype, we compared transcriptomes of *K. olearia* cultures grown at its optimal 65°C to
25 those at 30°C, 40°C, and 77°C. The temperature treatments affected expression of 573 of 2,224 *K.*
26 *olearia* genes. Notably, this transcriptional response elicits re-modeling of the cellular membrane and
27 changes in metabolism, with increased expression of genes involved in energy and carbohydrate
28 metabolism at high temperatures and up-regulation of amino acid metabolism at lower temperatures.
29 At sub-optimal temperatures, many transcriptional changes were similar to those observed in
30 mesophilic bacteria at physiologically low temperatures, including up-regulation of typical cold
31 stress genes and ribosomal proteins. At 77°C one third of the up-regulated genes are of hypothetical
32 function, indicating that many features of high temperature growth are unknown. Comparative
33 genomic analysis of additional Thermotogae genomes, indicate that one of *K. olearia*'s strategies for
34 low temperature growth is increased copy number of some typical cold response genes through
35 duplication and/or lateral acquisition.

36

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

37 **Introduction**

38 Microorganisms are capable of growing over an impressive temperature range, at least from -15°C to
39 122°C (Takai et al. 2008;Mykytczuk et al. 2013), and temperature is one of the most important
40 physical factors determining their distribution, diversity, and abundance (Schumann 2009). However,
41 individual microbial species grow only within a much narrower temperature interval. For example,
42 *Escherichia coli* O157:H7 thrives in the laboratory between 19°C and 41°C (Raghubeer and Matches
43 1990), while *Geobacillus thermoleovorans* has a growth range of 37°C to 70°C (Dinsdale et al.
44 2011). Microorganisms with temperature ranges >50°C are rare and, to date, research into the few
45 having ranges >40°C has focused on psychrophiles (e.g. (Mykytczuk et al. 2013)). *Kosmotoga*
46 *olearia* TBF 19.5.1 (hereafter referred to as *K. olearia*) is an anaerobic thermophile from the bacterial
47 phylum Thermotogae with a growth range that spans almost 60°C (DiPippo et al. 2009). How does
48 this organism achieve such physiological flexibility, and what are the evolutionary advantages and
49 implications of this capability?

50 Fluctuations in temperature induce broad physiological changes in cells, including growth
51 rate, alterations to cell wall and membrane composition, translation, and energy metabolism (Barria
52 et al. 2013;Pollo et al. 2015;Schumann 2009). These physiological changes can be classified into two
53 broad types of cellular response. Cold or heat *shock* designates the changes observed *immediately*
54 after the shift of a culture to a lower or higher temperature, while *prolonged growth* at a specific
55 lower or higher temperature elicits an *acclimated* low- or high-temperature response (Barria et al.
56 2013;Schumann 2009). Most studies of prokaryotes have focused on temperature shock responses
57 rather than acclimated growth. Among the Thermotogae, responses to both heat shock and prolonged
58 growth at high temperatures have been studied in the hyperthermophile *Thermotoga maritima*, which
59 can grow between 55°C and 90°C (Pysz et al. 2004;Wang et al. 2012). During prolonged high-
60 temperature growth *T. maritima* strongly up-regulates central carbohydrate metabolism genes and

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

61 expresses a few typical heat shock protein genes (Wang et al. 2012). Little is known about how *T.*
62 *maritima* responds to sub-optimal temperatures, although it encodes some genes implicated in cold
63 shock response. For example, its family of cold shock proteins (Csp), which are nucleic acid
64 chaperones known to be induced during cold shock and cold acclimation in mesophilic bacteria
65 (Barria et al. 2013;Phadtare 2004), exhibits nucleic acid melting activity at physiologically low
66 temperatures (Phadtare et al. 2003). Similarly, responses to cold shock in a few other thermophiles
67 involve many of the genes implicated in mesophilic cold shock response (e.g. (Boonyaratanakornkit
68 et al. 2005;Mega et al. 2010)). In this study we systematically assessed bacterial physiological
69 changes associated with response to prolonged growth at both high and low temperature using *K.*
70 *olearia* as a model system. Such changes can reflect not only response to temperature itself, but also
71 growth rate effects and general responses to stress. By conflating these factors, our study examined
72 overall changes in *K. olearia*'s gene expression in the environments defined by a specific
73 temperature.

74 The *K. olearia* genome (NC_012785) has 2,302,126 bp and is predicted to encode 2,224
75 genes (Swithers et al. 2011). Within the Thermotogae, genome size, intergenic region size, and
76 number of predicted coding regions correlate with the optimal growth temperature of an isolate
77 (Zhaxybayeva et al. 2012), with hyperthermophilic Thermotogae genomes being the most compact.
78 Phylogenetically, the Thermotogae order Kosmotogales comprises the genera *Kosmotoga* and
79 *Mesotoga* spp., the latter being the only described mesophilic Thermotogae lineage (Pollo et al.
80 2015). Assuming a hyperthermophilic last common ancestor of the Thermotogae (Zhaxybayeva et al.
81 2009), the Kosmotogales can be hypothesized to have acquired wide growth temperature tolerance
82 secondarily by expanding its gene repertoire. Moreover, it is likely that the ability of the
83 Kosmotogales common ancestor to grow at low temperatures enabled the evolution of mesophily in
84 *Mesotoga* (Pollo et al. 2015).

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

85 Such adaptations to novel environments can be greatly facilitated by lateral gene transfer
86 (LGT), since genes already "adapted" to the new conditions are readily available in the microbial
87 communities of the new environment (Boucher et al. 2003). For instance, LGT has been implicated
88 in adaptation to high temperature growth in hyperthermophilic bacteria, including *Thermotoga* spp.,
89 and to low temperature growth in archaea (López-García et al. 2015;Pollo et al. 2015;Boucher et al.
90 2003). Genome analysis of the mesophilic *Mesotoga prima* revealed that it laterally acquired 32% of
91 its genes after it diverged from other Thermotogae (Zhaxybayeva et al. 2012). Many of the predicted
92 gene donors are mesophiles, supporting the importance of lateral acquisition of genes already adapted
93 to mesophilic conditions in the evolution of *Mesotoga*.
94 To further gain insights into mechanisms of bacterial temperature response we sequenced 13
95 transcriptomes from isothermal and temperature-shifted cultures of *K. olearia* and examined
96 transcriptional differences at temperatures spanning its wide growth range. We also investigated the
97 importance of gene family expansion for adaptation of *K. olearia* to growth over a wide temperature
98 range via comparative genomic and phylogenetic analyses of identified temperature responsive genes
99 and their homologs in two newly sequenced *Kosmotoga* isolates, as well as in genomes of other
100 thermophilic and mesophilic Thermotogae.

101

102 Results

103 Temperature shifts and isothermic conditions elicit different growth patterns in *K. olearia*.

104 Under laboratory conditions in liquid anaerobic medium we observed growth of *K. olearia* at
105 temperatures as low as 25°C and as high as 79°C, with optimal growth at 65°C, defined as the
106 temperature affording the fastest growth rate (Fig. 1 and Fig. S1). Using a non-linear regression
107 model (Ratkowsky et al. 1983) we estimate a growth-permissive temperature range of 20.2 – 79.3°C,

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

108 consistent with the previously reported wide growth range of this isolate (DiPippo et al. 2009).
109 Interestingly, we were not able to cultivate *K. olearia* at temperatures near its range boundaries (30°C
110 and 77°C) by direct transfer from 65°C cultures. Instead, the growth temperature had to be changed
111 sequentially in ≤10°C increments. Particularly at the extremes, even small temperature shifts caused
112 both a longer lag phase and a slower growth rate compared to isothermal cultures (Fig. 1 and Fig.
113 S1). This phenomenon has been previously noted for mesophilic bacteria, especially for transitions
114 from high to low temperature (Swinnen et al. 2004). Our observations suggest that cells shifted to a
115 new temperature need to undergo large physiological changes that require time (i.e. an 'acclimation'
116 period (Barria et al. 2013)), and that these physiological challenges are too great to overcome when
117 temperature changes are large. To illuminate *K. olearia*'s transcriptional responses across its
118 temperature range we sequenced RNA from 13 cultures grown isothermally at 30°C, 40°C, 65°C, and
119 77°C (see Table S1 and Supplemental material).

120

121 **Architecture of the *K. olearia* transcriptome**

122 Analysis of transcription start and stop sites predicted a minimum of 916 transcriptional units (TU) in
123 *K. olearia* (Supplemental material and Table S2), 52% of which consist of a single gene. This
124 fraction of single-gene TUs lies between the 65% reported for *E. coli* (Cho et al. 2009) and the 43%
125 recorded for *T. maritima*, which has also been shown to have a streamlined genome and a low-
126 complexity transcriptome (i.e. few sub-operonic transcripts and few genes with multiple start sites)
127 (Latif et al. 2013). The average TU length of ~2.39 genes in *K. olearia* is less than the 3.3 genes per
128 transcript of *T. maritima* (Latif et al. 2013) but closer to 2.2 genes per transcript in the mesophilic
129 firmicute *Geobacter sulfurreducens* (Qiu et al. 2010) and 1-2 genes per transcript in bacteria in
130 general (e.g. (Cho et al. 2009)). Given that the *K. olearia* genome has more intergenic DNA than *T.*

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

131 *maritima*'s genome (the ratio of the nucleotides located in non-coding vs. coding regions is 0.13 in *K.*
132 *olearia* and 0.06 in *T. maritima*), the shorter TU lengths in *K. olearia* may point to more flexible
133 transcriptional regulation.

134

135 **Consistent energy generation across different temperature conditions**

136 *K. olearia* produces ATP from pyruvate using a biochemically well-understood fermentation
137 pathway that generates hydrogen, carbon dioxide and acetate ((DiPippo et al. 2009); Fig. 2 and data
138 not shown). Since pyruvate was the carbon and energy source provided in all experiments, we
139 surveyed 51 genes predicted to be involved in core energy metabolism during growth on pyruvate.
140 Based on expression levels, we constructed a model that accounts for all major end products during
141 growth at 65°C (Fig. 2) and contains 15 genes with consistently high expression across all
142 temperature treatments (Fig. S2a and Table S3). In addition to indirectly validating the previously
143 known functional annotations of these genes, we furthermore propose that the most highly expressed
144 ABC-transporter gene cluster, Kole_1509 – 1513, encodes a pyruvate transporter (Fig. 2). Its current
145 annotation as a peptide ABC transporter may be erroneous, since most of the peptide ABC
146 transporters predicted in *T. maritima* using bioinformatics have been shown instead to bind and
147 transport sugars (Nanavati et al. 2006). Further functional studies of the transporter (e.g. binding
148 assays, expression with different substrates) are needed to confirm this hypothesis.

149

150 **Identification of temperature-related transcriptional responses in *K. olearia***

151 Hierarchical clustering and Principal Component Analysis (PCA) separated the transcriptomes
152 according to temperature treatment (Fig. 3, Fig. S3 and Supplemental Results and Discussion),

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

153 suggesting that the observed changes in transcription are due to the culture growth temperature.
154 Several genes with a high correlation between their expression level and a specific growth
155 temperature (Table S4 and vectors in Fig. 3) are known to be involved in temperature response (Pollo
156 et al. 2015). For example, expression of the heat shock serine protease Kole_1599 positively
157 correlated with the 77°C transcriptomes (Fig. 3), where high expression of this protease was expected
158 based on its involvement in heat shock response in *T. maritima* (Pysz et al. 2004). Similarly,
159 expression of the cold shock protein genes Kole_0109 and Kole_2064 positively correlated with low
160 temperature growth (Fig. 3). Lastly, some observed changes presumably were due to the expected
161 decreased metabolic activity of the culture at sub- and supra-optimal temperatures. This can be
162 exemplified by the high expression and strong correlation of the alcohol dehydrogenase Kole_0742
163 (see Supplemental material for further discussion of the expression pattern of this gene) and the
164 central carbon metabolism gene glyceraldehyde-3-phosphate dehydrogenase (Kole_2020) with the
165 65°C transcriptomes.

166 Putative temperature-responsive genes were identified by pairwise comparisons of Illumina
167 transcriptomes from each temperature treatment to the optimal condition of 65°C (i.e. 30°C vs 65°C,
168 40°C vs 65°C, and 77°C vs 65°C). Across all comparisons 573 genes fulfilled our criteria for
169 temperature responsiveness (≥ 2-fold difference in expression, > 20 reads per transcript, False
170 Discovery Rate < 0.05) with 430, 115, and 169 genes detected in the 30°C vs 65°C, 40°C vs 65°C,
171 and 77°C vs 65°C comparisons, respectively (Table S5). Expression of 306 of the 573 putative
172 temperature-responsive genes correlated with growth rate ($r > |0.7|$; Table S5). However, growth rate
173 was a significant factor for only 60 of the 306 genes (ANCOVA, $p < 0.05$; Table S5). Moreover, for
174 26 of these 60 genes temperature was also a significant factor, and the simpler growth-rate-only
175 model was rejected for all but 13 of the 60 genes (Likelihood Ratio Test, FDR-corrected p-value <
176 0.05; Table S5). In an alternative comparison based on relative effect size of temperature vs. growth

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

177 rate, growth rate had larger influence on expression of 51 of the 306 genes (Table S5). Among these
178 genes were the glyceraldehyde-3-phosphate dehydrogenase (Kole_2020) and the iron-containing
179 alcohol dehydrogenase (Kole_0742) mentioned above. Therefore, although we retained the
180 designation of "putatively temperature-responsive" for all 573 genes, below we focus on genes not
181 affected by growth rate.

182

183 **Distribution of temperature responsive genes across functional categories suggests a regulated**
184 **response to temperature**

185 Most of the temperature-responsive genes were up-regulated when compared to the expression at
186 65°C (155 genes are down-regulated and 559 up-regulated; Table S5). One notable exception to this
187 overall trend were genes involved in carbohydrate and energy metabolism (Clusters of Orthologous
188 Groups [COG] categories C and G) where 32 genes were down-regulated at 30°C compared to 15
189 genes up-regulated. In transcriptomes at all non-optimal temperatures the list of putative temperature-
190 responsive genes was depleted in genes involved in translation (COG category J) and nucleotide
191 metabolism (COG category F) (Fig. 4) and conversely enriched in genes involved in signal
192 transduction (COG category T), and replication, recombination and repair (COG category L,
193 particularly at 30°C). Differential expression of the signal transduction genes (COG category T)
194 suggests the importance of these systems for regulating cellular responses at all tested temperatures.
195 Most of the identified COG category L genes are either mobile elements or CRISPR-associated
196 proteins, hinting at an increased activity of selfish genetic elements – a known common feature of
197 stress responses (Foster 2007). Additionally, at both 30°C and 77°C many genes encoding
198 transcription regulators (COG category K) are up-regulated, implying that prolonged growth at sub-

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

199 and supra-optimal temperatures results in detectable changes in transcriptional gene regulation in *K.*
200 *olearia*.
201

202 **At 40°C there are pronounced differences in membrane fatty acid composition but no signs of**
203 **cold stress**

204 Although the growth rate of *K. olearia* at 40°C is only one-third of that at the optimum 65°C (Fig. 1
205 and Fig. S1), clustering analysis suggested that the 40°C transcriptome was most similar to that at
206 65°C (Fig. 3 and Fig. S3). However, at 40°C even the four most highly expressed temperature-
207 responsive genes, including the growth-rate dependent alcohol dehydrogenase (Kole_0742), had
208 significantly lower expression levels (Table S5), reflecting the slower metabolic rate at lower
209 temperature growth. Yet, 94 of 115 putative temperature responsive genes were up-regulated (Table
210 S5, Fig. S2b), suggesting that slower metabolism is not the only explanation for the observed
211 transcriptional response to growth at 40°C.

212 Lipid metabolism (COG category I) appears to be particularly important at 40°C. For
213 instance, all the predicted fatty acid synthesis genes showed the highest expression at 40°C (Fig. S4),
214 with significantly higher expression of two genes involved in synthesis of unsaturated fatty acids
215 (Kole_0968) and initiation of fatty acid synthesis (Kole_0969). Biochemical analyses of total fatty
216 acids at 40°C and 65°C showed a much greater diversity of fatty acids at 40°C (Table S6), which
217 may explain the higher demand for the products of these genes at lower temperatures. Interestingly,
218 there was increased expression of a phosphate ABC transporter (Kole_0707 – Kole_0711, Table S5,
219 Fig. S2b), which may be linked to increased production of polar membrane lipids at moderately low
220 temperatures.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

221 An enrichment of differentially expressed genes in "post-translational modification, protein
222 turnover, and chaperone function" (COG category O; the category that harbors genes related to
223 cellular stress) was due to both up- and down-regulation of the involved genes (Fig. 4). For instance,
224 *K. olearia* has three temperature-responsive peptidylprolyl isomerase (PPIase) genes: two PpiC-type
225 genes (Kole_1682 and Kole_0383) that are both highly expressed at 40°C, and one FKBP-type gene
226 (Kole_1745) that shows high expression at all temperatures except 77°C (Table S5). At lower
227 temperatures (e.g. 37°C), these enzymes catalyze proline isomerization, which happens
228 spontaneously at higher temperatures (Godin-Roulling et al. 2015). However, the enzymes known to
229 assist protein folding under cellular stress, chaperones (GroEL [Kole_1627] and Hsp70 [Kole_0886])
230 and protease Do (Kole_1599), were significantly down-regulated at 40°C (Table S5 and Fig. S2b).
231 Among other typical cold stress-related proteins, only one of *K. olearia*'s three cold shock proteins
232 (Kole_0109) showed significantly higher expression at 40°C, but its up-regulation was merely
233 moderate when compared to its expression levels at 30°C (Table S5 and Fig. S2b). Taken together,
234 the observed expression patterns of known cold stress-related genes are consistent with the cells
235 being a non-stressed state at 40°C.
236

237 ***K. olearia* is in cold stress at 30°C**

238 Among the three most highly expressed up-regulated genes at 30°C are two Csp-encoding genes
239 (Kole_0109 and Kole_2064, Table S5, Fig. S2c), suggesting that the cells were in a cold-stressed
240 state during growth at 30°C. In support of this there was also significant up-regulation of other genes
241 previously linked to bacterial cold response (e.g (Barria et al. 2013)(Alreshidi et al. 2015)), including
242 a DEAD/DEAH-box RNA helicase (Kole_0922), *rbfA* (Kole_2103), *nusA* (Kole_1529) and
243 ribosomal proteins L10 (Kole_1840) and L7/L12 (Kole_1839). Genes encoding several additional
244 ribosomal proteins and ribosomal RNA (rRNA) methyltransferases, *rmlH* (Kole_1718) and *rmlL*
245 (Kole_0897), were also up-regulated (Fig. 3 and Table S5).

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

246 To detect a decrease in environmental temperature and elicit an appropriate regulatory
247 response, some bacteria have evolved two-component cold sensors (de Mendoza 2014). These signal
248 transduction systems consist of a sensor, a membrane-integrated protein with a kinase domain that
249 detects changes in the fluidity of the cell membrane, and the cytoplasmic response regulator, a
250 protein that induces expression of cold-responsive genes. In *K. olearia*, a histidine kinase with two
251 predicted transmembrane domains (Kole_1017) and two response regulators (Kole_1015 and
252 Kole_1016) showed a steady increase in expression as temperatures decreased from 65°C, but no
253 significant change in expression at 77°C (Table S5), leading us to hypothesize that these genes
254 encode a cold-sensing two-component system.

255 At 30°C, and to a lesser extent at 40°C, we also observed an over-representation of highly
256 expressed genes involved in amino acid metabolism (COG category E). Specifically, several genes in
257 the arginine (Kole_0092 – Kole_0097) and lysine (Kole_0104 – Kole_0107, 30°C only) biosynthesis
258 pathways were up-regulated, suggesting the potential for accumulation of peptides and amino acids
259 (or their intermediates) at lower temperatures. Interestingly, while the cells may accumulate peptides
260 at 30°C, at 40°C there was increased expression of an oligo-peptidase (Kole_1190) and genes
261 involved in lysine degradation (Kole_0958, Kole_0963 – Kole_0966). Such distinguishably different
262 metabolic responses to moderately low (40°C) and low (30°C) temperatures suggest a fine-tuned
263 temperature-dependent peptide turnover.

264

265 ***K. olearia* is in heat stress at 77°C**

266 Since 77°C is near the upper limit for *K. olearia* growth under our laboratory conditions, we
267 hypothesize that the observed differences in expression profiles at 65 and 77°C would reflect a cell-
268 wide heat stress response. Of the 169 significantly differentially expressed genes, 119 showed

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

269 increased expression at 77°C (Table S5 and Fig. S2d). Among the most highly expressed genes were
270 those encoding the structural RNAs *ffs* (Kole_R0010), *ssrA* (Kole_R0006), and *rnpB* (Kole_R0049)
271 (Fig. 3), suggesting an increased rate of RNA turnover at supra-optimal temperature. Moreover,
272 genes involved in carbohydrate and energy metabolism (COG C and G categories) were over-
273 represented and up-regulated at 77°C (Fig. 4, Table S5). However, only two of the known heat *shock*
274 response genes (Pysz et al. 2004), the extreme heat stress sigma factor-24 (*rpoE*, Kole_2150) and the
275 heat shock protease (Kole_1599), were up-regulated, and 41 of 119 genes (34%) are annotated as
276 "hypothetical proteins", indicating that adaptation to growth at sustained high temperature remains
277 largely uncharacterized. Putative functions could be inferred for some of the encoded proteins of
278 these 41 genes of unknown function. Twenty-three of the 41 proteins have one or more predicted
279 trans-membrane regions, suggesting they are either membrane-associated or secreted. Kole_0652
280 carries a PrcB-domain, which interacts with and stabilizes PrtP protease (Godovikova et al. 2010).
281 Kole_1314 (and its paralog Kole_1297) contains an AbiEii-toxin domain, and may be part of a toxin-
282 antitoxin system. Furthermore, the 41 genes could be grouped into 34 transcriptional units, each
283 containing at most three of these genes (Table S2) and scattered across the *K. olearia* genome. Two
284 of the genes, Kole_1430 and Kole_1431, are co-transcribed with genes from a two-component
285 system (Kole_1428 and Kole_1429), suggesting they may be involved in environmental sensing or
286 signaling. Three other co-transcribed genes (Kole_1266, Kole_1267, Kole_1270) are found in a
287 cluster containing CRISPR-genes, and two of them (Kole_1266 and Kole_1270) contain RAMP-
288 domains, suggesting CRISPR-related function (Makarova et al. 2011). The majority of the 41 genes
289 of unknown function have homologs in genomes of other *Kosmotoga* spp. (N=38), *Mesotoga* spp.
290 (N=23), and other Thermotogae (N=26).

291

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

292 **Global regulators of temperature response**

293 The transcriptional changes seen at the sub- and supra-optimal temperatures are likely to be
294 controlled by one or a few global regulators (Balleza et al. 2009), such as some temperature sensors
295 (de Mendoza 2014) and sigma factors needed for transcription initiation (Buck et al. 2000). The two-
296 component cold sensor up-regulated at low temperatures (Kole_1015 - Kole_1017) may represent
297 one such global regulator. Two different sigma factors (Kole_2150 and (Kole_1408) were
298 significantly up-regulated at 77°C and at 30°C and 40°C, respectively (Table S5), hinting at existence
299 of temperature-specific sigma factors. Kole_2150 belongs to the sigma-24 ECF subfamily, which is
300 activated in response to environmental stress (Balleza et al. 2009). Kole_1408 belongs to the sigma-
301 54 family, which is involved in enhancer-dependent transcription (Buck et al. 2000), introducing the
302 possibility that this sigma factor may be a target of the two-component cold sensor. In general, at
303 both 30°C and 77°C the differentially expressed genes were enriched in genes involved in
304 transcriptional regulation (COG category K) (Fig. 4), leaving a possibility of additional global
305 regulators.

306

307 **Conservation of *K. olearia*'s temperature-responsive genes within *Kosmotoga* genus**

308 All genes that are required for adaptation and response of *K. olearia* strain TBF 19.5.1 to a wide
309 range of growth temperatures are expected to be present in other *K. olearia* isolates, whereas some
310 may be absent from *Kosmotoga* species that have a narrower spectrum of growth temperature.
311 Therefore, we contrasted the *K. olearia* genome to the genomes of *Kosmotoga* sp. DU53 and
312 *Kosmotoga arenicorallina* (Pollo et al. 2016). When compared to *K. olearia*, *Kosmotoga* sp. DU53
313 has a similar growth temperature range (25°C - 79°C, Table S7) and >99% average nucleotide

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

314 identity (ANI), while *K. arenicorallina* exhibits a narrower growth temperature range (35°C - 70°C,
315 Table S7) and has only 84% ANI.

316 Indeed, the *Kosmotoga* sp. DU53 genome lacks only 10 of the 573 *K. olearia* putative
317 temperature-responsive genes (BLASTP and TBLASTN searches, E-value <10^{-3}, Table S5). All 10
318 genes were expressed in *K. olearia* at relatively low levels, suggesting that they are unlikely to be
319 essential for high or low temperature growth. On the other hand, the *K. arenicorallina* genome does
320 not have detectable homologs of 103 of the 573 *K. olearia*'s putative temperature-responsive genes
321 (BLASTP and TBLASTN searches, E-value <10^{-3}; Table S5). The list of absent genes includes
322 several of the arginine and lysine biosynthesis genes that are up-regulated in *K. olearia* during
323 growth at 30°C, and seven of the genes of unknown function up-regulated at 77°C. Therefore, we
324 hypothesize that a subset of these 103 genes may play a role in extending the growth range of *K.*
325 *olearia* to ≤35°C and ≥70°C.

326

327 **Many *key* temperature-responsive genes are laterally acquired**

328 Obtaining "pre-adapted" genes from other genomes is one way prokaryotes adjust to new
329 environmental conditions (Boucher et al. 2003). Using HGTector (Zhu et al. 2014) we predicted that
330 354 of *K. olearia*'s 2,118 protein coding genes have been acquired laterally by *K. olearia* or the
331 Kosmotogales (i.e. *Kosmotoga* and *Mesotoga*), presumably representing LGT events occurring after
332 the divergence of Kosmotogales from other Thermotogae (Table S8). Eighty-eight of the 354 genes
333 were temperature responsive (Table S5, Fig. S5A and S5B), including several above-discussed
334 highly expressed genes (Table 1), and 37 of the 88 genes are shared with *Mesotoga* (Table S5).
335 Notably, 76% of theses latter 37 genes are upregulated at 30°C (Fig. S5C), suggesting that their
336 acquisition may have been important in adaptation to low temperature growth. Among these are the
337 earlier mentioned rRNA methyltransferase genes (Kole_1718 and Kole_0897). The fatty acid

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

338 synthesis genes (Kole_0969 - Kole_0973) that are up-regulated at 40°C, as well as their
339 Kosmotogales homologs, form a distantly related sister clade to other Thermotogae lineages (Fig.
340 S6A), implying that these genes may have been acquired from an un-sampled lineage. Similarly, the
341 Csp-encoding gene highly expressed at 30°C (Kole_0109) is placed outside of the Thermotogae
342 clade (Fig. S6B).

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

348

349 **Discussion**

350 **How can *K. olearia* grow at such a wide range of temperature?**

351 Examination of *K. olearia*'s transcriptional response to sustained exposure to a non-optimal
352 temperature revealed both high expression of core metabolic genes and differential expression of
353 hundreds of other genes, with selected features highlighted in Fig. 5. At each tested temperature, core
354 metabolism genes for pyruvate utilization show high relative expression, which strongly suggests that
355 *K. olearia* uses the same proteins for its core energy metabolism and that these proteins can function
356 across its wide growth temperature range. In contrast, genes involved in regulatory functions showed
357 significant changes in expression at all temperatures, particularly at 30°C and 77°C (COG category K
358 and T in Fig. 4), suggesting that regulating gene expression is important in response to a temperature
359 shift. Among these genes we identified putative global temperature regulators: the two-component
360 cold sensor and temperature-specific sigma factors.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

361 Close to *K. olearia*'s growth temperature maximum of 77°C, carbohydrate and energy
362 metabolism genes (COG categories C and G) were up-regulated (Fig. 4). It is unclear, however, if the
363 underlying cause is the increased turnover of enzymes at elevated temperatures, or a demand for
364 more enzymes due to increased carbohydrate catabolism. Increased carbohydrate metabolism in
365 response to prolonged growth at supra-optimal temperature has been observed previously in *T.*
366 *maritima* (Wang et al. 2012), and therefore may be a common adaptation to high temperature growth
367 in the Thermotogae. As observed for *K. olearia,* the prolonged supra-optimal temperature growth of
368 *T. maritima* also did not involve up-regulation of typical heat-shock response proteins (Wang et al.
369 2012). This highlights the difference between cellular response to an immediate heat-shock and to
370 prolonged growth at supra-optimal temperature (Balleza et al. 2009), and in general justifies
371 classifying the cellular response to temperature into these two distinct categories.
372 At the moderately sub-optimal growth temperature of 40°C, *K. olearia* cells face
373 physiological challenges of proper protein folding (Godin-Roulling et al. 2015), and of maintenance
374 of a functional cell membrane (de Mendoza 2014). Our observation that at 40°C, despite the lower
375 growth rate, lipid metabolism genes were among the most highly expressed genes suggests that
376 changes to the cell membrane composition are one of the most important adaptations for survival of
377 *K. olearia* at lower temperatures. Proper protein folding may require enzymatic assistance (Godin-
378 Roulling et al. 2015), which may be achieved via *K. olearia*'s three PPIases. The significant up-
379 regulation of the PPIase genes at 40°C suggest that they are particularly important at moderately low
380 temperatures where the cells are still relatively active. However, the overall lack of induction of
381 typical stress-related genes at 40°C, especially when compared to 30°C and 77°C, suggests that 40°C
382 is within the "Goldilocks" temperature range for *K. olearia.*
383 At 30°C *K. olearia* is clearly under cold stress (Fig. 5), as evidenced by expression of genes
384 known to be implicated in cold response. One of the strategies for maintenance of proper cellular
385 function at non-optimal temperatures (Pollo et al. 2015) is accumulation of compatible solutes.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

386 Specifically, re-modelling of amino acid metabolism and possible accumulation of amino acids as
387 compatible solutes at low temperatures has been observed in bacteria (e.g. (Dahlsten et al.
388 2014;Ghobakhlou et al. 2015)). The up-regulation of many genes involved in amino acid metabolism
389 suggests that *K. olearia* may also accumulate amino acids or their intermediates for this purpose,
390 especially at 30°C. At 30°C there was also significant up-regulation of a citrate synthase gene
391 (Kole_1230), suggesting that *K. olearia* cells may accumulate citrate, as was observed in
392 *Staphylococcus aureus* during prolonged cold stress (Alreshidi et al. 2015). Alternatively, citrate
393 synthase, together with isocitrate dehydrogenase (Kole_1227), may be involved in converting
394 pyruvate or acetyl-CoA to 2-oxoglutarate, a precursor for several amino acids including arginine,
395 which has been suggested to accumulate in for instance *Clostridium* (Dahlsten et al. 2014).
396 Interestingly, the genome of strictly mesophilic *M. prima* encodes more genes involved in amino acid
397 metabolism than the genomes of *K. olearia* and other thermophilic Thermotogae (Zhaxybayeva et al.
398 2012). This supports the notion that the shift towards more amino acid metabolism in *K. olearia* may
399 be an adaptation to low temperature growth and suggests that the mesophilic *Mesotoga* have taken
400 this a step further by also *acquiring* more amino acid metabolism genes. Moreover, amino acid
401 metabolism genes are among the most numerous bacterial genes laterally acquired by mesophilic
402 archaea, which is hypothesized to reflect archaeal adaptation to low temperature growth (López-
403 García et al. 2015).

404 Ribosomes also need to be fine-tuned to function properly at low temperature (Barria et al.
405 2013). Consistently, we observed a change in expression of several genes encoding ribosomal
406 proteins. The most dramatic differential expression, however, was observed for a ribosomal protein
407 gene not yet connected to cold response (L34; Kole_0258, Fig. 3, Table S5). L34, a bacteria-specific
408 ribosomal protein hypothesized to be a relatively recent addition to the evolving ribosome (Fox
409 2010), is required for proper ribosome formation (Akanuma et al. 2014). A *Bacillus subtilis* mutant
410 lacking the L34 gene showed particularly slow growth at low temperature (Akanuma et al. 2012),

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

411 suggesting a role for L34 in this condition. We also observed significant increased expression of
412 rRNA methyltransferases, and methylation of rRNAs has been associated with responses to
413 environmental stress, including temperature (Baldridge and Contreras 2014). Therefore, we
414 hypothesize that *K. olearia* modifies its ribosome by changing stoichiometry of the ribosome
415 components and by methylating rRNA. Time required for such ribosomal adjustments could also
416 explain the longer lag phase following temperature shifts (Fig. S1).

417 Expansion of cold-responsive gene families may represent a common strategy for low
418 temperature adaptation, as has been noted in many bacteria, especially in psychrophiles (e.g. (Piette
419 et al. 2010). *K. olearia* exhibits the same trend. For instance, when compared to other Thermotogae,
420 all three analyzed *Kosmotoga* genomes harboured more copies of Csp-encoding genes (Table S9).
421 The observed gene family expansions, however, might be important not necessarily for low
422 temperature growth alone, but instead for growth over a wide temperature interval. *Mesotoga*
423 functions with only a single *csp* gene, demonstrating that having more copies of this gene is not
424 required for low temperature growth. Having several versions of a gene could make differential
425 regulation under different growth temperatures easier. For example, extra homologs (Kole_0111 and
426 Kole_0110) of the earlier-discussed putative cold sensor system do not show coordinated temperature
427 responses: Kole_0110 is up-regulated at 40°C, while Kole_0111 is up-regulated at 77°C (Table 1).
428 Therefore, these additional homologs may encode sensors tuned to different temperatures.
429

430 **Evolutionary mechanisms that drive adaptive changes in Kosmotogales genomes**

431 Gene family expansions can be achieved via within-lineage gene duplication or through LGT, and a
432 combination of these mechanisms appears to be at work in *K. olearia*. For example, even though
433 several Thermotogae genomes contain as many copies of PPIase genes as do *Kosmotoga* genomes
434 (Table S9), phylogenetic analysis suggests that in the Kosmotogales this gene family has only

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

435 recently been expanded by both LGT (the FKBP-type, Table 1) and duplication (the PpiC-type, Fig.
436 S6C). Similar conclusions can be made from the phylogenetic analyses of *csp* genes (Fig. S6B).
437 LGT appears to be a significant factor in evolution of low temperature growth in
438 Kosmotogales. In *K. olearia* although the proportion of transferred genes among the temperature
439 responsive genes is similar to what is observed in the genome (Fig. S5), many of the genes identified
440 here as key for low temperature growth have been acquired by LGT. For instance, the fatty acid
441 synthesis genes with high expression at 40°C were acquired by LGT into the Kosmotogales and are
442 likely important for low temperature growth in both *Mesotoga* and *Kosmotoga*. Other key LGT genes
443 including the DEAD/DEAH box RNA helicase, the temperature sensor, and the low temperature
444 induced *csp* genes, are only found in *Kosmotoga* spp. and may therefore be more important for its
445 wide temperature growth (Table 1). The predicted acquisition of fatty acid synthesis and *csp* genes by
446 (now mesophilic) archaea (López-García et al. 2015) additionally argues for the importance of these
447 genes in adaptation to low temperature growth.
448 Despite the importance of expanded gene families in low temperature adaptation, the role of
449 mutation and consequent natural selection on specific genes in response to changing environmental
450 conditions should not be neglected. For example, typical cold response genes *rbfA* (Kole_2103) and
451 *nusA* (Kole_1529) were not laterally acquired, but nevertheless show high expression only at 30°C.
452 Deciphering adaptive changes that occurred in such genes compared to thermophilic homologs may
453 elucidate molecular mechanisms of low temperature adaptation.
454

455 **Why maintain the capacity for growth over such a wide temperature range?**

456 Most bacteria are under selection to eradicate extraneous DNA (and genes) from their genomes
457 (Graur 2016), and among free-living bacteria Thermotogae in general have very compact genomes.
458 Kosmotogales, however, have notably larger genomes than other thermophilic Thermotogae (Pollo et

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

459 al. 2015;Zhaxybayeva et al. 2012), raising the possibility that expanded genomes are advantageous in
460 *K. olearia*'s habitat. As discussed above, many of the genes in *K. olearia*, such as the cold-sensor
461 system, were expressed only at specific sub- or supra-optimal temperatures, and do not seem to be
462 important for growth at other temperatures (Table 1 and Table S5). The regulated response to low
463 temperatures and the preservation of the laterally acquired genes specifically expressed at 40°C and
464 30°C suggest that *K. olearia* encounters environments with very different temperatures frequently
465 enough to maintain these genes in its genome. Such environments may include oil reservoirs located
466 at different depths, as well as marine sediments influenced by the mixing of cold deep-sea water and
467 hydrothermal fluids (Sievert and Vetriani 2012). Perhaps, *K. olearia* migrates between such locations
468 via subsurface fluids and, as a result, may have been selected to become a temperature generalist.
469 Indeed, the environmental conditions of the subsurface environments and marine hydrothermal vents
470 from which *Kosmotoga* spp. have been isolated vary substantially (DiPippo et al. 2009;Nunoura et al.
471 2010;L'Haridon et al. 2014). For example, *K. olearia* originated from a deep subsurface oil reservoir
472 with *in situ* temperature of 68°C (DiPippo et al. 2009), but its 16S rRNA sequences have also been
473 detected in many oil fields having *in situ* temperatures of 20°C–50°C (Nesbø et al. 2010). *Kosmotoga*
474 sp. DU53, which is most similar to *K. olearia*, was isolated from an oil reservoir with an *in situ*
475 temperature of ~50°C, while *K. arenicorallina* was obtained from hydrothermal sediments with a
476 temperature of ~40°C (Nunoura et al. 2010). Notably, *K. olearia* was also identified as a major
477 constituent of a metagenome from a deep subsurface oil reservoir with *in situ* temperature of 85°C
478 and pressure of 25MPa (Kotlar et al. 2011). While the reservoir temperature is higher than the
479 maximum *K. olearia* growth temperature reported here, elevated pressure could extend *K. olearia*'s
480 temperature maximum, as has been demonstrated for some archaea (e.g. (Takai et al. 2008)).
481 Therefore, *K. olearia*’s growth temperature range under natural conditions may be even broader than
482 20-79°C.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

483 **Conclusions**

484 The present study demonstrates that even bacteria with relatively small genomes can use
485 transcriptional changes to respond effectively to large changes in temperature (Fig. 5). We showed
486 that *K. olearia*'s response to sustained exposure to a non-optimal temperature includes up-regulation
487 of hundreds of genes. Several key genes with known temperature-related functions apparently have
488 been acquired laterally, suggesting that LGT is an evolutionarily successful strategy for expansion of
489 temperature tolerance. However, gene duplication and subsequent sub-functionalization of the
490 paralogs likely also play an important adaptive role.

491 The ability of *K. olearia* to inhabit both high and low temperature environments suggests that
492 members of this lineage encounter environments with large temperature fluctuations and/or migrate
493 across ecological niches within the deep biosphere (e.g. between deep and shallow subsurface oil
494 reservoirs). Therefore, the subsurface environments, as well as their microbial populations, might be
495 viewed as a connected archipelago instead of isolated islands. As a corollary, we speculate that *K.*
496 *olearia*-like ecological generalists could also facilitate LGT among seemingly isolated deep
497 biosphere microbial communities adapted to a narrower ecological niche. For example, we have
498 previously demonstrated high levels of gene flow among hyperthermophilic *Thermotoga* populations
499 in subsurface oil reservoirs and marine hydrothermal vents (Nesbø et al. 2015), environments that are
500 separated by non-thermophilic surroundings but are hydrologically linked. The mechanism of such
501 gene flow is not yet known, but *K. olearia*-like Thermotogae capable of growing both in subsurface
502 oil reservoirs and adjacent marine sediments could serve as mediators of gene exchange.

503 Although some of the identified 573 temperature-responsive genes are already known to be
504 expressed in bacteria and archaea grown at high or low temperatures, most of the up-regulated genes
505 have not previously been implicated in temperature response and are in need of better functional and
506 biochemical characterization. Moreover, the majority of the *K. olearia* genes responsive to elevated
507 temperature encode proteins of unknown functions. Versatile proteins that work across a broad range

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

508 of temperatures also warrant further biochemical and evolutionary analyses, as understanding of their
509 enzymatic flexibility can aid the design of commercially important thermostable proteins.

510 Finally, other regulatory mechanisms (e.g. DNA methylation, post-transcriptional
511 modifications) may reveal additional proteins important in *K. olearia*'s temperature responses, and
512 should be targeted in future studies.

513

514 Material and Methods

515 Bacterial culturing

516 *K. olearia* TBF 19.5.1 (DSM 21960(T), ATCC BAA-1733(T), Genbank accession number
517 NC_012785) was grown at different temperatures (30°C, 40°C, 65°C, and 77°C), but otherwise
518 optimal conditions, in *Kosmotoga* medium (KTM) using pyruvate as growth substrate as described in
519 (DiPippo et al. 2009). Cultures used as inocula were stored at 4°C, except those used for experiments
520 at ≥77°C. Actively growing cultures at temperatures ≥77°C had to be used directly as inoculum
521 because the cultures would not grow from inocula stored at either 4°C or room temperature (~22°C).
522 Replicate cultures received the same volume of inoculum; however, variable inoculum volume was
523 used at different temperatures (Table S1), as larger inoculum volumes were required to achieve
524 growth at the non-optimal temperature treatments.

525 Measurement of *K. olearia* growth at different temperatures

526 Growth curves were constructed from optical density measurements at 600 nm (OD_{600}) using an
527 Ultrospec 3100 pro. For cultures grown at 40°C, 65°C, 77°C, and 78°C two sets of triplicate bottles,
528 inoculated from the same inoculum 12 h apart, were monitored for a 12 h period per day to generate
529 the growth curves. The cultures for isothermic growth at 40°C, 65°C, 77°C, and 78°C were

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

530 monitored hourly, while the cultures for shifted growth at these temperatures were monitored every
531 1-5 hours. At 30°C one set of triplicate bottles was monitored once daily. Isothermic growth curves
532 were calculated from six replicates, except for the curves for 30°C and 77°C that had three and 12
533 replicates, respectively. All shifted growth curves consisted of six replicate cultures, except for 40°C
534 and 30°C which had four and three replicates, respectively. To determine growth rates (Fig. 1), for
535 each culture the $\ln OD_{600}$ was plotted against growth time and the curve was fitted with a linear trend
536 line. The growth rate was defined as the slope at the log phase. To determine the time span of each
537 growth phase (Fig. S1), full composite growth curves were constructed using pooled replicate data.
538 For each curve, OD_{600} for all replicates was plotted against growth time and a polynomial regression
539 trend line was fitted.

540

541 **Cultivation of *Kosmotoga* sp. DU53 and *K. arenicorallina,* and confirmation of their growth**
542 **temperature ranges**

543 *Kosmotoga* sp. DU53 was grown in KTM as described above for *K. olearia*. *K. arenicorallina* was
544 also grown in KTM; however, maltose was used as substrate (2.5 mL and 0.5 mL 10% maltose was
545 added to serum bottles and Hungate tubes, respectively). One mL of culture was used as inoculum for
546 all cultures (bottles and tubes). The temperature range of each strain was assessed by examination of
547 culture turbidity as a proxy for growth. Starting from cultures grown at optimal growth temperature
548 (~ 65°C for *Kosmotoga* sp. DU53 and 60°C for *K. arenicorallina*), new cultures were shifted in
549 ≤10°C increments. If growth was observed after a shift, then that culture was used to initiate a new
550 culture. The shifting procedure was terminated when growth was no longer evident at a given
551 temperature.

552

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

553 **RNA isolation and processing**

554 Cultures used for RNA extraction were inoculated from cultures that had been grown under the same
555 temperature conditions for at least three transfers. The time at which a culture was expected to be in a
556 desired growth phase was determined from the composite growth curves (Fig. S1), and was used as a
557 cell harvesting time (listed in Table S1). This procedure avoided exposure of the cultures to the lower
558 ambient temperatures in the laboratory during subsampling. For the 30°C cultures, OD_{600} was
559 additionally measured 24 h before harvesting to ensure the culture was in mid log phase. In order to
560 stabilize the transcripts and to avoid any transcriptional response to the change in temperature, an
561 equal volume of "stop solution" (10% phenol in ethanol) was added to the sealed cultures via syringe
562 immediately upon removal from the incubator. For each temperature treatment, RNA was extracted
563 in mid-log to late-log phase, using the Zymo Research Fungal/Bacterial RNA MiniPrep Kit
564 (Cedarlane Laboratories, Ltd.; Burlington, Ontario) and following the manufacturer's protocols
565 (Table S1).

566 Following recommendations in (Haas et al. 2012), we aimed to sequence ~3 million non-
567 ribosomal-RNA reads per sample. Ribosomal RNA (rRNA) depletion was performed on all samples
568 using the Ribo-Zero rRNA Removal Kit (Gram-Positive Bacteria) Magnetic Kit (MRZGP126,
569 Epicentre). On average, two rRNA depletions were needed to generate sufficient input RNA (200-
570 500 ng for Ion Torrent PGM and 10-400 ng for Illumina MiSeq), although some samples required as
571 many as five rRNA depletions. Quality and quantity of the total RNA, as well as efficiency of rRNA
572 depletion, were assessed on an Agilent 2100 Bioanalyzer RNA Nano chip or RNA Pico chip
573 following the manufacturer's instructions for "Prokaryote Total RNA". RNA successfully depleted of
574 rRNA was used to construct RNA-Seq libraries following the manufacturer's instructions, and
575 sequenced on either an Ion Torrent PGM (RNA-Seq kit V2; transcriptomes are labeled with an "IT"
576 suffix) or an Illumina MiSeq (TruSeq RNASeq v2 2x100 bp). The platform and RNA extraction

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

577 technique used for each transcriptome are summarized in Table S1. The transcriptomes are available
578 in the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under the accession number
579 SRP075860.

580

581 **RNA-Seq analysis**

582 For each transcriptome**,** sequenced reads were analyzed using the "RNA-Seq" function in CLC
583 Genomics Workbench version 7.0.4. Briefly, reads were first trimmed using an ambiguous limit of 2
584 and a quality limit of 0.05. To remove sequences that matched remaining rRNA transcripts, the
585 trimmed reads were subjected to a relaxed mapping protocol: only rRNA genes were used as a
586 “reference genome”, reads were mapped only to the reference sequences using similarity fraction of
587 0.5; length fraction was set to 0.5; all other parameters were set to default (maximum number of hits
588 for a read = 10, map to both strands, mismatch cost = 2, insertion cost = 3, deletion cost = 3, auto-
589 detect paired distances, color space alignment, color error cost = 3). The mapped reads were
590 designated as rRNA and were removed from further analysis.

591 The remaining reads were subjected to an RNA-Seq protocol with strict mapping parameters
592 (allowing mapping to intergenic regions, similarity fraction = 0.95; length fraction = 0.95; all other
593 default settings as described above) using the *K. olearia* annotated genome as a reference. Unmapped
594 reads were discarded. Expression levels for every gene were estimated using "Reads Per Kilobase of
595 transcript per Million mapped reads" (RPKM) values.

596 RPKM values for all genes are listed in Table S4. Differentially expressed genes were
597 identified by doing pairwise comparisons of Illumina transcriptomes of the isothermically grown
598 cultures at 30°C, 40°C, and 77°C to the cultures grown at the optimal temperature of 65°C. The
599 analyses used the “Empirical Analysis of DGE” function, which employs the “Exact Test” for two-

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

600 group comparisons (Robinson and Smyth 2008). A gene was considered differentially expressed in a
601 pairwise comparison if it had (*i*) > 20 reads in at least one of the two transcriptomes, (*ii*) a statistically
602 significant difference in the RPKM values (corrected for multiple testing using False Discovery Rate
603 [FDR] < 0.05), and (*iii*) a difference in RPKM values at least two-fold in magnitude. Principal
604 Component Analysis (PCA) and biplot visualization were performed using R packages *ade4* and
605 *bpca* respectively (Dray et al. 2007;Faria et al. 2013). Each gene was assigned to a Clusters of
606 Orthologous Groups (COG; (Galperin et al. 2015)) functional category using the Integrated
607 Microbial Genomes (IMG) portal (Markowitz et al. 2014). Genes assigned to more than one COG
608 category were counted in all assigned categories.

609 Batch culture cannot selectively discern gene expression that is exclusively influenced by
610 temperature from expression that is solely growth rate-dependent. In theory, continuous culture
611 conducted at a single growth rate could afford such discrimination. However, given the extremely
612 slow growth of *K. olearia* near its temperature maxima and minima (Fig. 1), it was not feasible to use
613 anaerobic bioreactors to cultivate cells at this marginal growth rate across the temperature range.
614 Hence, in order to assess how differences in growth rates influence the expression patterns, we
615 examined correlations of the expression pattern of the putative temperature responsive genes with
616 growth rates calculated at each temperature. Pearson correlation of expression values and growth
617 rates were calculated in Microsoft Excel. ANCOVA, linear regression, and likelihood ratio tests were
618 carried out in R. The "growth rate only" model was the linear regression (expression = growth rate).
619 In the ANCOVA model growth rates was set as the quantitative variable and temperature as
620 qualitative variable (expression = growth rate + temperature). The growth rates used were 0.006 for
621 30°C, 0.087 for 40°C, 0.274 for 65°C and 0.107 for 77°C (see Fig. 1). When comparing the most
622 significant temperature coefficient to the growth rate coefficient, the latter was scaled by the average

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

623 growth rate. The growth rate effect was defined as being greater than the temperature rate effect if
624 |temperature/(growth rate * 0.115)| < 1.

625

626 **Identification of genes involved in growth on pyruvate**

627 *K. olearia* genes predicted to be involved in pathways for pyruvate conversion to acetate, CO_2, H_2
628 and ATP were retrieved from the KEGG (Kanehisa et al. 2014) and BioCyc (Caspi et al. 2014)
629 databases. Genes encoding hydrogenases were taken from (Schut et al. 2012), and genes encoding the
630 F-type ATPase subunits were identified using IMG (Markowitz et al. 2014).

631

632 **Fatty acids analysis**

633 Total lipids were extracted from *K. olearia* grown at 40°C to early stationary phase and 65°C to mid-
634 log, early stationary, late stationary and death phase by using methanol-chloroform (1:1 v/v). Fatty
635 acid methyl esters (FAME) were prepared from total lipids extracts using mild alkaline methanolysis
636 (Guckert et al. 1985). Dried FAME samples were re-dissolved in 300 µl chloroform (HPLC grade,
637 Fisher Scientific) and analyzed by gas chromatography with mass spectrometry (GC-MS) on an
638 Agilent 6890N gas chromatograph with a model 5973 inert mass selective detector (Agilent) fitted
639 with an Agilent HP-5MS capillary column (30 m × 0.25 mm ID, 0.25 µm film thickness; J + W
640 Scientific). Helium was used as the carrier gas with a temperature program of 130°C increasing at
641 3°C min^{-1} to 290°C and held for 2 min. Sample peaks were identified by comparison to Bacterial
642 Acid Methyl Ester Mix standards (Supelco, Sigma Aldrich) or on the basis of mass spectra and
643 expressed as a percentage of the total FAME quantified in each sample.

644

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

645 **Comparative analyses of three *Kosmotoga* spp. genomes**

646 The genome of *K. olearia* was compared to genomes of *Kosmotoga* sp. DU53 (accession number
647 JFHK00000000) and *K. arenicorallina* (accession number JGCK00000000) (Pollo et al. 2016) using
648 the IMG portal (Markowitz et al. 2014) and Geneious v.9. Specifically, genes were declared
649 homologous if they were significantly similar in BLASTP and TBLASTN (Altschul et al. 1997)
650 searches (E-value < 10^{-3}). For phylogenetic analyses, additional homologs of *K. olearia* genes were
651 retrieved from the NCBI non-redundant (*nr*) protein database and the IMG databases via BLASTP
652 searches, retaining the 100 top-scoring matches with E-value < 10^{-3}. Sequences were aligned using
653 MAFFT (Katoh et al. 2002), and phylogenetic trees were reconstructed using RAxML (Stamatakis
654 2006), as implemented in Geneious v. 9.1.3. Homologs from the recently released genome of
655 *Kosmotoga pacifica* (NZ_CP011232) (L'Haridon et al. 2014) were included in gene-specific
656 phylogenetic analyses. Pairwise Average Nucleotide Identity (ANI) (Goris et al. 2007) was
657 calculated using the Enveomics Toolbox (Rodriguez-R and Konstantinidis 2016).

658

659 **Detection of laterally transferred genes**

660 A customized version of HGTector (Zhu et al. 2014) (available through https://github.com/ecg-
661 lab/hgtector) was used to identify putatively transferred protein-coding genes in the *K. olearia*
662 genome. Homologs of each annotated protein-coding open reading frame (ORF) in the NC_012785
663 GenBank record were retrieved from a local copy of the NCBI *nr* database (downloaded November
664 21, 2014) using the BLASTP program from BLAST 2.2.28+ (Altschul et al. 1997). Sequences were
665 first filtered for low complexity regions using the *seg* algorithm. Then, only matches with E-value
666 $<10^{-5}$ and sequence coverage $\geq 70\%$ were retained. Database matches were expanded according to
667 the *MultispeciesAutonomousProtein2taxname* file from RefSeq release 68. This was necessary as

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

668 some genes across various taxonomic ranks were combined into a single entry in RefSeq, which
669 artificially decreased the representation of these genes in Close and Distal groups (see below), and
670 would confound downstream analysis. Taxonomic affiliation of each match was assigned using the
671 NCBI Taxonomy database (downloaded on November 21, 2014). Only 500 top-scoring matches
672 (after filtering for sequence coverage) were used as input for HGTector. The "Self" group was
673 defined as TaxID 651456 (genus *Kosmotoga*), and the "Close" group was defined as either TaxID
674 1184396 (genus *Mesotoga*, a sister group) or TaxID 2419 (order Thermotogales, comprising
675 *Thermotoga*, *Mesotoga*, and *Kosmotoga*). In either case, the “Distal” group comprised the remaining
676 taxonomic groups. The conservative cutoff (the median between the zero peak and the first local
677 minimum) was used for both the “Close” and “Distal” groups. A gene was designated as putatively
678 transferred if its "Close" score was below the cutoff and its "Distal" score was above the cutoff.
679 Putatively transferred genes with no top-scoring match in Thermotogae were designated as recent
680 transfer events into *K. olearia* (labelled "K" in Table S8). Putatively transferred genes for which the
681 difference between Close(Thermotoga) and Close(Mesotoga) scores was <1 were designated as gene
682 transfer events into Kosmotogales (i.e. *Kosmotoga* and *Mesotoga*; labelled "K+M" in Table S8).

683

684 **Acknowledgments**

685 This work was supported by a Research council of Norway award (project no. 180444/V40) to
686 C.L.N.; a Simons Investigator award from the Simons Foundation, a Dartmouth College Walter and
687 Constance Burke Research Initiation Award, and Dartmouth College start-up funds to O.Z.; and an
688 NSERC Alexander Graham Bell Canada Graduate Scholarship CGS-M to S.M.J.P. We thank
689 Rhianna Charchuk for technical assistance, Dr. Amanda J. Lohan for advice and discussion of RNA-

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

690 Seq methodology, Dr. Karen Budwill for head-space gas analyses, and Jennifer Franks for initial

691 discussions of lateral gene transfer detection methods.

692

693 **References**

694 Akanuma G, Nanamiya H, Natori Y, Yano K, Suzuki S, Omata S, Ishizuka M, Sekine Y, Kawamura

695 F (2012) Inactivation of ribosomal protein genes in *Bacillus subtilis* reveals importance of each

696 ribosomal protein for cell proliferation and cell differentiation. J Bacteriol 194: 6282-6291

697 Akanuma G, Kobayashi A, Suzuki S, Kawamura F, Shiwa Y, Watanabe S, Yoshikawa H, Hanai R,

698 Ishizuka M (2014) Defect in the formation of 70S ribosomes caused by lack of ribosomal protein L34

699 can be suppressed by magnesium. J Bacteriol 196: 3820-3830

700 Alreshidi MM, Dunstan RH, Macdonald MM, Smith ND, Gottfries J, Roberts TK (2015)

701 Metabolomic and proteomic responses of *Staphylococcus aureus* to prolonged cold stress. Journal of

702 proteomics 121: 44-55

703 Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped

704 BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res

705 25: 3389-3402

706 Baldridge KC, Contreras LM (2014) Functional implications of ribosomal RNA methylation in

707 response to environmental stress. Crit Rev Biochem Mol Biol 49: 69-89

708 Balleza E, López-Bojorquez LN, Martínez-Antonio A, Resendis-Antonio O, Lozada-Chávez I,

709 Balderas-Martínez YI, Encarnación S, Collado-Vides J (2009) Regulation by transcription factors in

710 bacteria: beyond description. FEMS Microbiol Rev 33: 133-151

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

711 Barria C, Malecki M, Arraiano CM (2013) Bacterial adaptation to cold. Microbiology 159: 2437-
712 2443

713 Boonyaratanakornkit BB, Simpson AJ, Whitehead TA, Fraser CM, El-Sayed NMA, Clark DS (2005)
714 Transcriptional profiling of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* in
715 response to lethal heat and non-lethal cold shock. Environ Microbiol 7: 789-797

716 Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau MER, Nesbø CL, Case RJ, Doolittle WF
717 (2003) Lateral gene transfer and the origins of prokaryotic groups. Annu Rev Genet 37: 283-328

718 Buck M, Gallegos M, Studholme DJ, Guo Y, Gralla JD (2000) The bacterial enhancer-dependent
719 σ^{54} (σ^N) transcription factor. J Bacteriol 182: 4129-4136

720 Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM,
721 Kothari A, Kubo A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Subhraveti P,
722 Weaver DS, Weerasinghe D, Zhang P, Karp PD (2014) The MetaCyc database of metabolic
723 pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res
724 42: D459-D471

725 Cho B, Zengler K, Qiu Y, Park YS, Knight EM, Barrett CL, Gao Y, Palsson BØ (2009) The
726 transcription unit architecture of the *Escherichia coli* genome. Nat Biotechnol 27: 1043-1049

727 Dahlsten E, Isokallio M, Somervuo P, Lindström M, Korkeala H (2014) Transcriptomic analysis of
728 (group I) *Clostridium botulinum* ATCC 3502 cold shock response. PloS one 9: e89958

729 de Mendoza D (2014) Temperature sensing by membranes. Annu Rev Microbiol 68: 101-116

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

730 Dinsdale AE, Halket G, Coorevits A, van Landschoot A, Busse HJ, de Vos P, Logan NA (2011)
731 Emended descriptions of *Geobacillus thermoleovorans* and *Geobacillus thermocatenulatus*. Int J
732 Syst Evol Microbiol 61: 1802-1810

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

736 Dray S, Dufour AB, Chessel D (2007) The ade4 Package - II : Two-table and K-table methods. R
737 News 7: 47-52

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

740 Foster PL (2007) Stress-induced mutagenesis in bacteria. Crit Rev Biochem Mol Biol 42: 373-397

741 Fox GE (2010) Origin and evolution of the ribosome. In: Deamer D, Szostak JW (eds) The Origins
742 of Life. Cold Spring Harb Perspect Biol, pp 2:a003483

743 Galperin MY, Makarova KS, Wolf YI, Koonin EV (2015) Expanded microbial genome coverage and
744 improved protein family annotation in the COG database. Nucleic Acids Res 43: D261-D269

745 Ghobakhlou A, Johnston A, Harris L, Antoun H, Laberge S (2015) Microarray transcriptional
746 profiling of Arctic *Mesorhizobium* strain N33 at low temperature provides insights into cold adaption
747 strategies. BMC Genomics 16: 383-396

748 Godin-Roulling A, Schmidpeter PAM, Schmid FX, Feller G (2015) Functional adaptations of the
749 bacterial chaperone trigger factor to extreme environmental temperatures. Environ Microbiol 17:
750 2407-2420

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

751 Godovikova V, Wang HT, Goetting-Minesky MP, Ning Y, Capone RF, Slater CK, Fenno JC (2010)
752 *Treponema denticola* PrcB is required for expression and activity of the PrcA-PrtP (dentilisin)
753 complex. J Bacteriol 192: 3337-3344

754 Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007) DNA-DNA
755 hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol
756 Microbiol 57: 81-91

757 Graur D (2016) Prokaryotic Genome Evolution. In: Anonymous Molecular and Genome Evolution.
758 Sinauer Associates, Sunderland, MA, pp 451

759 Guckert JB, Antworth CP, Nichols PD, White DC (1985) Phospholipid, ester-linked fatty acid
760 profiles as reproducible assays for changes in prokaryotic community structure of estuarine
761 sediments. FEMS Microbiol Lett 31: 147-158

762 Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J (2012) How deep is deep enough for RNA-Seq
763 profiling of bacterial transcriptomes? BMC Genomics 13: 734-745

764 Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information,
765 knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 42: D199-D205

766 Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence
767 alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059-3066

768 Kotlar HK, Lewin A, Johansen J, Throne-Holst M, Haverkamp T, Markussen S, Winnberg A,
769 Ringrose P, Aakvik T, Ryeng E, Jakobsen K, Drabløs F, Valla S (2011) High coverage sequencing of
770 DNA from microorganisms living in an oil reservoir 2.5 kilometres subsurface. Environmental
771 Microbiology Reports 3: 674-681

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

772 Latif H, Lerman JA, Portnoy VA, Tarasova Y, Nagarajan H, Schrimpe-Rutledge AC, Smith RD,
773 Adkins JN, Lee D, Qiu Y, Zengler K (2013) The genome organization of *Thermotoga*
774 *maritima* reflects its lifestyle. PLoS genetics 9: e1003485

775 L'Haridon S, Jiang L, Alain K, Chalopin M, Rouxel O, Beauverger M, Xu H, Shao Z, Jebbar M
776 (2014) *Kosmotoga pacifica* sp. nov., a thermophilic chemoorganoheterotrophic bacterium isolated
777 from an East Pacific hydrothermal sediment. Extremophiles 18: 81-88

778 López-García P, Zivanovic Y, Deschamps P, Moreira D (2015) Bacterial gene import and mesophilic
779 adaptation in archaea. Nature reviews.Microbiology 13: 447-456

780 Makarova KS, Aravind L, Wolf YI, Koonin EV (2011) Unification of Cas protein families and a
781 simple scenario for the origin and evolution of CRISPR-Cas systems. Biology direct 6: 38-65

782 Markowitz VM, Chen IA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T,
783 Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC
784 (2014) IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic
785 Acids Res 42: D560-7

786 Mega R, Manzoku M, Shinkai A, Nakagawa N, Kuramitsu S, Masui R (2010) Very rapid induction
787 of a cold shock protein by temperature downshift in *Thermus thermophilus*. Biochem Biophys Res
788 Commun 399: 336-340

789 Mykytczuk NCS, Foote SJ, Omelon CR, Southam G, Greer CW, Whyte LG (2013) Bacterial growth
790 at -15°C; molecular insights from the permafrost bacterium *Planococcus halocryophilus* Or1. The
791 ISME journal 7: 1211-1226

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

792 Nanavati DM, Thirangoon K, Noll KM (2006) Several archaeal homologs of putative oligopeptide-
793 binding proteins encoded by *Thermotoga maritima* bind sugars. Appl Environ Microbiol 72: 1336-
794 1345

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

797 Nesbø CL, Swithers KS, Dahle H, Haverkamp TH, Birkeland N, Sokolova T, Kublanov I,
798 Zhaxybayeva O (2015) Evidence for extensive gene flow and *Thermotoga* subpopulations in
799 subsurface and marine environments. The ISME Journal 9: 1532-1542

800 Nunoura T, Hirai M, Imachi H, Miyazaki M, Makita H, Hirayama H, Furushima Y, Yamamoto H,
801 Takai K (2010) *Kosmotoga arenicorallina* sp. nov. a thermophilic and obligately anaerobic
802 heterotroph isolated from a shallow hydrothermal system occurring within a coral reef, southern part
803 of the Yaeyama Archipelago, Japan, reclassification of *Thermococcoides shengliensis* as *Kosmotoga*
804 *shengliensis* comb. nov., and emended description of the genus *Kosmotoga*. Arch Microbiol 192:
805 811-819

806 Phadtare S (2004) Recent developments in bacterial cold-shock response. Curr Issues Mol Biol 6:
807 125-136

808 Phadtare S, Hwang J, Severinov K, Inouye M (2003) CspB and CspL, thermostable cold-shock
809 proteins from *Thermotoga maritima*. Genes to Cells 8: 801-810

810 Piette F, D'Amico S, Struvay C, Mazzucchelli G, Renaut J, Tutino ML, Danchin A, Leprince P,
811 Feller G (2010) Proteomics of life at low temperatures: trigger factor is the primary chaperone in the
812 Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. Mol Microbiol 76: 120-132

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

813 Pollo SMJ, Charchuk R, Nesbø CL (2016) Draft genome sequences of *Kosmotoga* sp. DU53 and
814 *Kosmotoga arenicorallina* S304. GenomeA : in press

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

817 Pysz MA, Ward DE, Shockley KR, Montero CI, Conners SB, Johnson MR, Kelly RM (2004)
818 Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium
819 *Thermotoga maritima*. Extremophiles 8: 209-217

820 Qiu Y, Cho BK, Park YS, Lovley D, Palsson B, Zengler K (2010) Structural and operational
821 complexity of the *Geobacter sulfurreducens* genome. Genome Res 20: 1304-1311

822 Raghubeer EV, Matches JR (1990) Temperature range for growth of *Escherichia coli* serotype
823 0157:H7 and selected coliforms in *E. coli* medium. J Clin Microbiol 28: 803-805

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

826 Robinson MD, Smyth GK (2008) Small-sample estimation of negative binomial dispersion, with
827 applications to SAGE data. Biostatistics 9: 321-332

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

830 Schumann W (2009) Temperature sensors of eubacteria. Advances in applied microbiology 67: 213-
831 256

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

832 Schut GJ, Boyd ES, Peters JW, Adams MWW (2012) The modular respiratory complexes involved
833 in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their
834 evolutionary implications. FEMS Microbiol Rev 37: 182-203

835 Sievert S, Vetriani C (2012) Chemoautotrophy at Deep-Sea Vents: Past, Present, and Future.
836 Oceanography 25: 218-233

837 Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with
838 thousands of taxa and mixed models. Bioinformatics 22: 2688-2690

839 Swinnen IAM, Bernaerts K, Dens EJJ, Geeraerd AH, Van Impe J. F. (2004) Predictive modelling of
840 the microbial lag phase: A review. Int J Food Microbiol 94: 137-159

841 Swithers KS, DiPippo JL, Bruce DC, Detter C, Tapia R, Han S, Goodwin LA, Han J, Woyke T,
842 Pitluck S, Pennacchio L, Nolan M, Mikhailova N, Land ML, Nesbø CL, Gogarten JP, Noll KM
843 (2011) Genome sequence of *Kosmotoga olearia* strain TBF 19.5.1, a thermophilic bacterium with a
844 wide growth temperature range, isolated from the Troll B oil platform in the North Sea. J Bacteriol
845 193: 5566-5567

846 Takai K, Nakamura K, Toki T, Tsunogai U, Miyazaki M, Miyazaki J, Hirayama H, Nakagawa S,
847 Nunoura T, Horikoshi K (2008) Cell proliferation at 122°C and isotopically heavy CH_4 production
848 by a hyperthermophilic methanogen under high-pressure cultivation. Proc Natl Acad Sci U S A 105:
849 10949-10954

850 Wang Z, Tong W, Wang Q, Bai X, Chen Z, Zhao J, Xu N, Liu S (2012) The temperature dependent
851 proteomic analysis of *Thermotoga maritima*. PloS one 7: e46463

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

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

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

859 Zhu Q, Kosoy M, Dittmar K (2014) HGTector: An automated method facilitating genome-wide
860 discovery of putative horizontal gene transfers. BMC Genomics 15: 717

861

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

862 **Table 1. Gene expression in selected laterally acquired temperature-responsive genes**. At each
863 temperature, the listed RPKM values represent the average expression levels across replicates.
864 Values that are significantly different from 65°C are shown in bold font. None of the temperature
865 responsive expression patterns show significant effect of growth rate and, where tested (i.e. |R| > 0.7),
866 the growth-rate-only model was rejected.

		RPKM values				
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[a]
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	**2386**	**3063**	939	1424	HGTector

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

	III , FabH					
Kole_0970	enoyl-ACP reductase II , fabK	2226	3243	1486	1641	HGTector[a]
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_1718	Ribosomal RNA methyltransferase, rmlH	**531**	332	211	203	HGTector
Kole_1745	PPIase FKBP-type	2783	2382	1541	**430**	HGTector

867 a) Phylogenetic tree is shown in Supplemental Fig. 6.

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

868 **Fig. Legends:**

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

877

878 **Fig. 2. Model of energy generation pathway in *K. olearia* during growth on pyruvate.** The model
879 includes genes likely involved in conversion of pyruvate to acetate, CO_2, H_2, and ATP. The genes
880 were selected from the list of genes highly expressed across all temperature conditions (Fig. S2a,
881 Table S3). Acetate transport is not shown. The dashed box indicates hydrogenase activity. The two
882 highly expressed hydrogenases are shown, but their potential interactions with each other or with the
883 membrane are not known. Increased expression of citrate synthase at low temperature, which could
884 redirect acetyl-CoA away from acetate production, is shown in grey. The assignment of the ABC
885 transporter as pyruvate transporter (Kole_1509-1513) is based on its high expression level, but
886 experiments are needed to confirm its involvement in pyruvate transport. The model also explains the
887 observed lower ratio of carbon dioxide to hydrogen produced by growth on maltose vs. pyruvate (not
888 shown), because during growth on maltose reduced electron carriers would be generated from the
889 conversions of maltose to pyruvate.

890

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

891 **Fig. 3. Biplot of the principal component analysis of 8 transcriptomes.** The plot is based on
892 expression values from all genes. Each transcriptome is denoted by a point and genes are represented
893 by vectors. The 20 longest (i.e. most highly correlated) gene vectors are shown. Coordinates and
894 vector length for all genes can be found in Table S4. It should be noted that the *ffs* (Kole_R0010)
895 transcript is only 115 nt, and may not have been fully represented in every transcriptome due to our
896 isolation protocol which selects against small RNA (<200 nucleotides). Also, the high expression of
897 the alcohol dehydrogenase (Kole_0742) is probably due to the addition of stop solution before RNA
898 isolation (see Supplemental Results and Discussion).

899

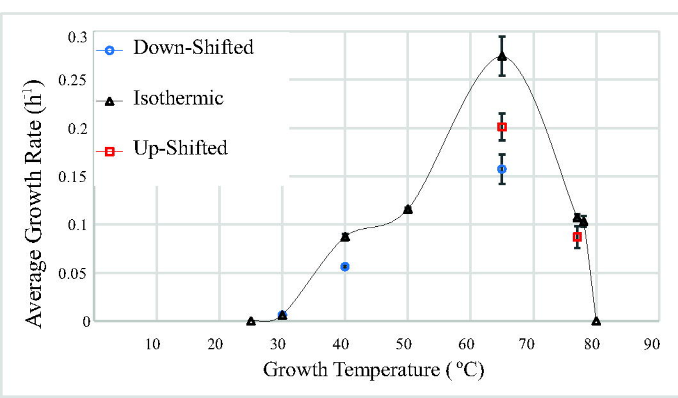
900 **Fig. 4. Difference between observed and expected number of temperature responsive genes**
901 **across functional categories.** Functional categories were assigned using the Clusters of Orthologous
902 Groups (COG) database as implemented in IMG (Markowitz et al. 2014) and are denoted by one-
903 letter abbreviations along the X-axis. For each temperature treatment (30°C, 40°C and 77°C) only the
904 temperature-responsive fraction of the *K. olearia* genome was considered. If the temperature-
905 responsive genes were randomly distributed across functional categories we would expect the same
906 fraction of temperature-responsive genes in each COG category. The difference (in percent) between
907 the observed and expected number of temperature responsive genes is plotted on the Y-axis with
908 positive and negative values referring to over- and under-representation of the temperature-
909 responsive genes, respectively. COG category abbreviations: J: Translation, ribosomal structure and
910 biogenesis, K: Transcription, L: Replication, recombination and repair, D: Cell cycle control, cell
911 division, chromosome partitioning, V: Defense mechanisms, T: Signal transduction mechanisms, M:
912 Cell wall/membrane/envelope biogenesis, U: Intracellular trafficking, secretion, and vesicular
913 transport, O: Post-translational modification, protein turnover, and chaperones, C: Energy production
914 and conversion, G: Carbohydrate transport and metabolism, E: Amino acid transport and metabolism,

bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.

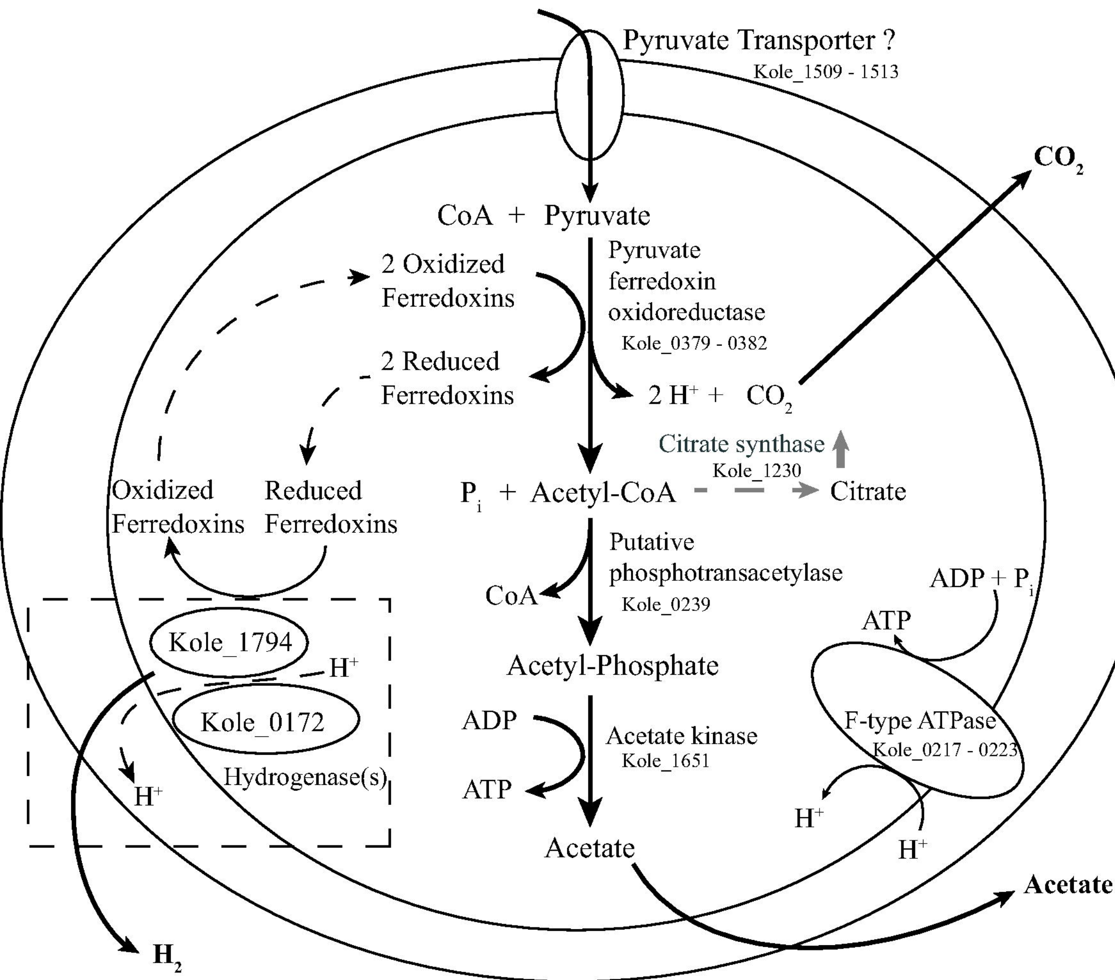
915 F: Nucleotide transport and metabolism, H: Coenzyme transport and metabolism, I: Lipid transport
916 and metabolism, P: Inorganic ion transport and metabolism, Q: Secondary metabolites biosynthesis,
917 transport, and catabolism, R: General function prediction only, S: Function unknown, NC: Not in
918 COG database.

919 **Fig. 5. Schematic of *K. olearia*'s major temperature-induced transcriptional responses.** Major
920 responses outlined in the text occurring in the three states observed are shown. The number of
921 chaperone, protease, and cold-shock symbols reflects their relative expression at each temperature.
922 Chaperones include groEL (Kole_1627), groES (Kole_1626), and dnaK (Kole_0886). Proteases refer
923 to protease Do (Kole_1599) and protease La (Kole_0536). Cold-shock genes consist of Kole_2064
924 (dark green squares) and Kole_0109 (green trapezoids). The different coloured ribosomes represent
925 changes in ribosomal protein composition at sub-optimal temperatures (see text for discussion). The
926 putative two-component regulatory system (Kole_1015 – 1017) that had a dramatic increase in
927 expression under cold stress is shown. Prominent changes in functional categories (COGs) for each
928 temperature condition are also shown. Arrows indicate relative gene expression when compared to
929 growth at optimal temperature. Core functions occurring at all temperatures are shown in the white
930 box. Basic membrane lipoprotein and the major toga anchor protein refer to Kole_1554 and
931 Kole_1500, respectively. See Fig. 2 for genes involved in pyruvate metabolism.

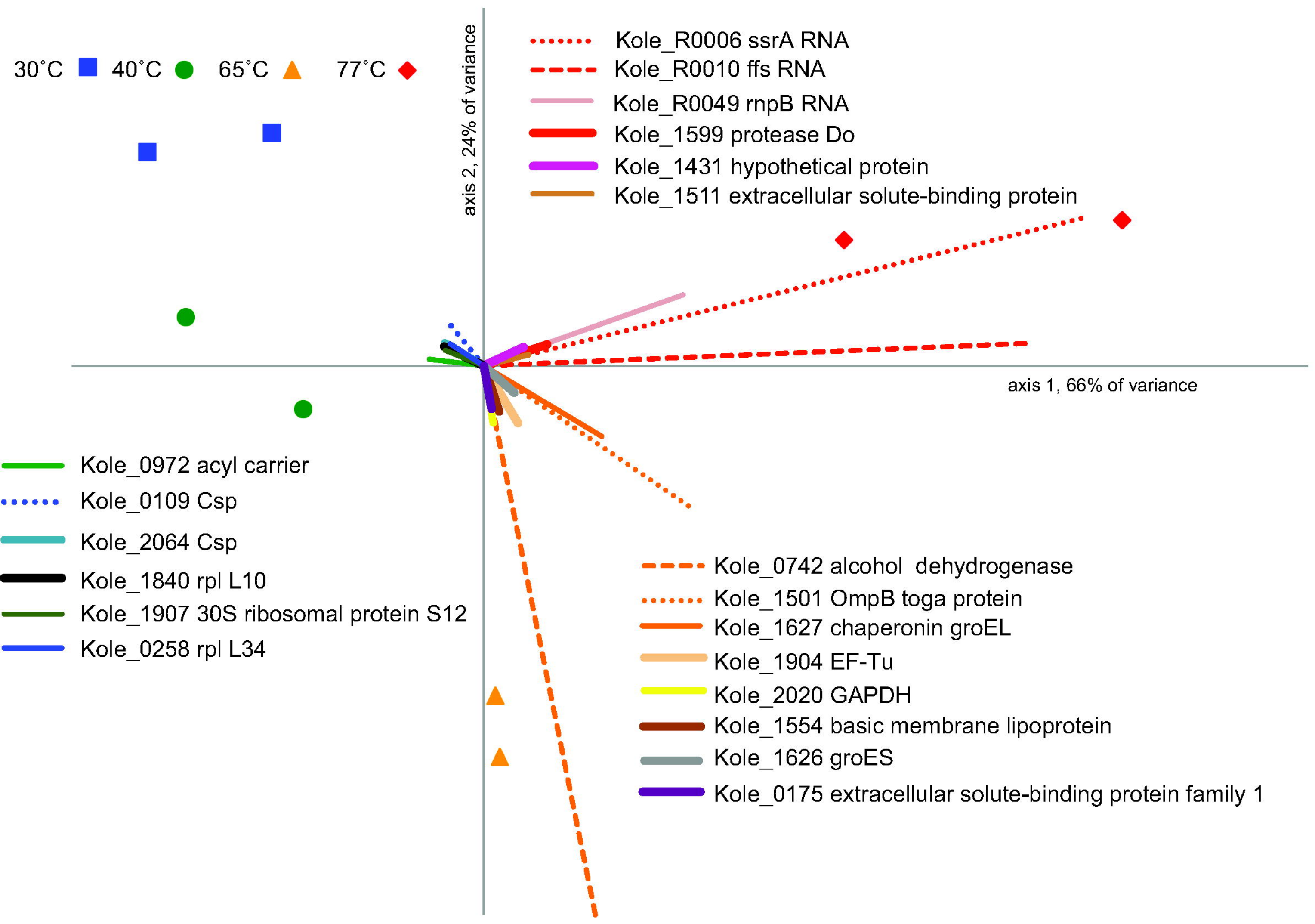
bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



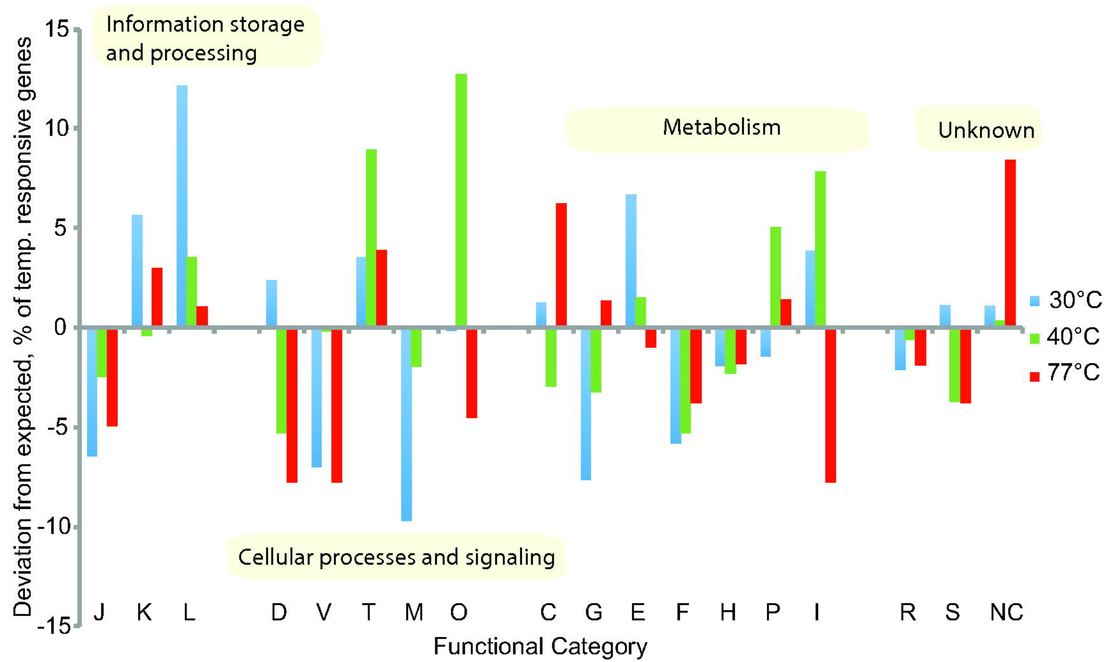
bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.



bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.



bioRxiv preprint doi: https://doi.org/10.1101/060053; this version posted June 7, 2017. 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.