The type III-B CRISPR-Cas System Affects Energy Metabolism and Adaptation in the Archaeon *Saccharolobus solfataricus*

Authors:

Erika Wimmer[1*], Logan H. Hodgskiss[1*], Melina Kerou[1], Isabelle A. Zink[1,2#], Christa Schleper[1]

Affiliations:

[1] Department of Functional and Evolutionary Ecology, Faculty of Life Sciences, University of Vienna

[2] Laboratory of Microbiology, Wageningen University and Research, Stippeneng 4, 6708 WE Wageningen, The Netherlands.

[*]these authors contributed equally to this work

[#]corresponding author

ABSTRACT

Type III CRISPR-Cas immune systems that recognize and cleave extrachromosomal RNA when active, are particularly widespread in archaea. Mechanistically, these systems have the potential to regulate gene expression of host genes on a post-transcriptional level, but very little is known about any potential accessory roles of type III-B systems beyond immunity. We have created knockout mutants of a type III-B CRISPR-Cas complex in the thermoacidophilic archaeon *Saccharolobus solfataricus* to investigate potential secondary functions of the type III-B system. Deletion mutants exhibited an accelerate growth but were less quickly adaptable to changes in carbon sources in their growth media. In line with this phenotype, upregulated genes were significantly enriched in functional categories of energy production and conversion, as well as with carbohydrate or amino acid transport and metabolism in RNAseq studies. Generally, a significant accumulation of genes encoding transmembrane proteins in the upregulated proportion of the transcriptome suggests interconnections between the type III-B CRISPR-Cas system and various membrane-associated processes. Notably, the deletion mutants did not lose their general virus- or plasmid defense activities indicating that this particular system might have been partially adopted for cellular regulatory roles.

INTRODUCTION

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR associated proteins) immune systems are present in roughly a third of fully sequenced bacterial and the majority of archaeal genomes[1]. Despite the impressive diversity of known CRISPR-Cas systems, the principle first steps of the immune reaction remain: A small crRNA (CRISPR RNA) is transcribed from a so-called spacer in a CRISPR array in the host genome. Spacers are DNA snippets of previously encountered viruses and plasmids that have been stored in the host genome during the process of CRISPR adaptation[2,3]. A CRISPR-Cas protein effector forms a complex with one small crRNA to identify complementary nucleic acid targets (so-called protospacers) on invading genetic elements, such as viruses or plasmids[4]. Dependent on the CRISPR-Cas system, the bound nucleic acid target is then either directly cleaved by the Cas enzyme or abortive infection pathways are activated to clear the invader from the population[5].

While most CRISPR-Cas effectors target double stranded DNA, type III and type VI CRISPR-Cas systems exhibit dedicated RNA-targeting mechanisms[6]. Type III CRISPR-Cas systems are frequently found in various archaeal phyla and are particularly abundant among Thermoproteota (formerly Crenarchaeota)[5]. These systems are especially intriguing because they are multi-subunit complexes that execute their immune reaction in three different ways: Upon recognition and binding of a specific RNA target (i.e. nascent mRNA transcribed from a plasmid), the Cas7 backbone of the type III complex cleaves the RNA protospacer. Furthermore, binding to the protospacer activates the Cas10 subunit of the complex which cleaves single stranded DNA in transcription bubbles promiscuously via its HD domain[7–10]. Concomitantly, the cyclase (palm) domain of Cas10 synthesizes cyclic molecules from ATP (cOAs)[11–14] which in turn activate ancillary CARF-domain containing proteins which can be either RNases, DNases or proteases that indiscriminately degrade their substrates in the cell[15,16]. This leads to cell death or cell dormancy which curtails the spread of the invader in the population[17–23].

Interestingly, the nucleolytic and abortive infection pathways of the type III CIRSPR-Cas complex can be inactivated, thereby only allowing specific cleavage of the RNA target. Complementarity of a sequence stretch at the 3’ end of the protospacer (the so called antitag) to the 5’ part of the crRNA or mismatched bases in certain parts of the crRNA:target duplex

allosterically inhibit the catalytic activities of Cas10[8,9,11–14,24–26]. Therefore, type III CRISPR-Cas effector complexes can specifically cleave mRNA targets, which raises the question as to whether they also contribute to the regulation of host gene expression. Indeed, there are some indications for additional roles of type III CRISPR-Cas complexes apart from immunity. For instance, type III loci often contain high numbers of accessory genes potentially constituting a functional connection to cellular processes other than defense[27,28]. Also, in many cases type III CRISPR-*cas* operons are not genomically linked to CRISPR arrays or modules involved in the acquisition of new spacers, thus they are reliant on co-occurring CRISPR-Cas systems for a canonical immune response[5,29]. Furthermore, we find type III CRISPR-Cas systems to be among the most highly expressed genes in *Saccharolobus solfataricus*, also when no plasmid or virus invader is present.

Sa. solfataricus P1 belongs to the archaeal phylum Thermoproteota and contains two type III CRISPR-Cas complexes, a type III-D (Csm) and a type III-B (Cmr-ß), as well as three DNA-targeting type I-A complexes[5,30]. In vitro studies on the Cmr-ß complex from a close relative, *Sa. islandicus* (SisCmr-ß), revealed that catalytic activities of the complex are modulated by the additional subunit Cmr7[31,32] as its presence coincided with starkly reduced cOA production by the Cas10 subunit[33]. In this study, the endogenous type I CRISPR-Cas system was harnessed to generate a *cmr-ß* knockout strain of *Sa. solfataricus* in order to discern its putative links to other processes and alternative roles besides immunity. A wide range of effects of the genomic deletion on energy metabolism and transport mechanisms indicate a general impact on the host's physiology as revealed by RNAseq. In addition, knockout strains were more sensitive to switches from complex into single sugars media, indicating a role of the Cmr-ß complex in metabolic adaptability. Furthermore, we did not observe any effect on plasmid immunity upon type III CRISPR-Cas deletion, indicating that the Cmr-ß might not be a main contributor to type III CRISPR-Cas immune defense in *Sa. solfataricus*.

RESULTS

Type IA-CRISPR-Cas-mediated knockout of the type III-B CRISPR-*cas* gene module in *Sa. solfataricus* P1

The endogenous DNA-targeting type I-A CRISPR-Cas system was harnessed for deletion of the genes encoding the type III-B CRISPR-Cas complex Cmr-ß (SSOP1_RS10085 - SSOP1_RS10115) (Fig. 1A). Two different 37-bp protospacers with 3' protospacer adjacent

105 motif (PAM) 5'-NGG[34] were selected within each of the three targeted genes (*cmr7, cmr6,*
106 *cmr3/cas5*) on the *Sa. solfataricus* chromosome (Fig. 1A). In total, six different mini CRISPR
107 arrays (miniCR) were assembled as described previously[35] containing one spacer or a
108 combination of two to three spacers targeting the respective genes on either coding or
109 template strand (cs or ts, respectively). The miniCRs and two approximately 800-bp donor
110 sequences for DNA repair corresponding to upstream and downstream flanking regions of
111 the *cmr-ß* locus were cloned into the plasmid-based vector pIZ-GW[36] (Fig. 1B). *Sa.*
112 *solfataricus* M18 cells transformed with the genome-editing plasmids were enriched under
113 selection in liquid medium. Although all constructs yielded the intended deletion as detected
114 by culture PCR, one knockout culture was chosen for plating to isolate single colonies (Suppl.
115 Fig. S1). The genomes of two purified knockout mutant cultures (KO1 and KO2) grown from
116 two single colonies were sequenced showing that in comparison to the deposited reference
117 (NZ_LT549890.1) the *cmr-ß* genes and their upstream basal promoter elements had been
118 scarlessly excised from the genome in both mutants (Fig. 2). Interestingly, KO2 had
119 additionally acquired a larger deletion (over 37 kb) between the genes SSOP1_RS14925 and
120 SSOP1_RS15080, both of which are encoding IS5 family transposases (Suppl. tab. 1, Fig. 3).
121
122 **No difference in plasmid immunity between *cmr-ß* knockout mutants and wildtype strains**
123 To investigate the contribution of Cmr-ß to the immune defense against foreign genetic
124 elements in *Sa. solfataricus* P1, both wildtype and *cmr-ß* knockout mutant were challenged
125 with two plasmids generating a transcript of a protospacer (i.e. target) complementary to a
126 crRNA encoded in CRISPR array D (Fig. 4A and B). We have previously demonstrated that
127 cells transfected with a viral vector carrying the same protospacer flanked by a PAM could
128 not survive, even under selective conditions[24], indicating an active type I CRISPR-Cas
129 immune response. To solely study the immune response of Cmr-ß and/or Csm by
130 circumventing type I CRISPR-mediated interference, the PAM of construct D2.109 was now
131 disrupted by nucleotide insertion at the 3' end of the protospacer. Upon recognition of this
132 protospacer transcript, the available type III effector(s) should elicit a multi-layered immune
133 response causing detrimental effects to the cells due to abortive infection. As mentioned
134 above, abortive infection can be abolished by the presence of a sequence stretch at the 3'
135 end of the protospacer (antitag) complementary to the 5' tag of the crRNA while the bound
136 target is still degraded[7–9,11–14,37,38]. Therefore, we equipped the same protospacer with a 3'
137 flanking antitag in construct D2.109-a (Fig. 4A), which in our previous study yielded a type III-
138 CRISPR-Cas-mediated reduction in protospacer transcript levels while DNA interference was
139 not observed[24]. In addition, two different controls were used, namely cells transformed with

140 plasmid A53 lacking a matching protospacer[24,39] and untransformed cells not harboring any
141 plasmid (negative control, NC) (Fig. 4A, 4B). After introduction of the different plasmids into
142 both wildtype and deletion mutant, it became evident that growth of cells harboring D2.109
143 was negatively affected to a similar extent in both wildtype and *cmr-ß* knockout mutants (Fig.
144 4D, Suppl. Fig. S2 and S3). In both strains, cells transformed with A53 and D2.109-a were
145 still growing in a 10fold higher dilution on plates compared to cells carrying plasmid D2.109.
146 The same overall trend could also be observed when following their growth in liquid culture
147 (Fig. 4C). Consequently, the absence of the Cmr-ß complex had no obvious effect on the
148 extent of the observed type III CRISPR-Cas-mediated immune response in our experimental
149 setting.
150
151 **Knockout mutants exhibit faster growth and significant changes in their transcriptional**
152 **landscape compared to the wildtype**
153 To test whether growth of *cmr-ß* knockout mutants was affected compared to the wildtype
154 in complex medium, four replicate cultures for each strain were inoculated to the same initial
155 OD_{600} (0.027 ± 5.88E-04). In comparison to the wildtype, similarly accelerated growth was
156 observed for both knockout strains KO1 and KO2, the latter being the strain that attained the
157 additional genomic deletion (Fig. 5A). Samples were taken at the start of exponential growth
158 at an OD_{600} = 0.2 and RNA was extracted to perform RNA-seq. Principal component analysis
159 (PCA) revealed a distinct shift in the transcriptomic landscape upon deletion of Cmr-ß (Fig
160 5B). All four replicates of each strain grouped together, thereby indicating high reproducibility.
161 Additionally, the two deletion strains were overlapping whereas the wildtype was clearly
162 clustering separately. One principal component (PC1) could explain ~39% of the variance
163 between knockout strains and wildtype.
164
165 Changes in gene expression with high statistical significance (cutoff for adjusted *P* value
166 (padj) < 0.001) were taken into consideration in the downstream analysis even in case of
167 smaller log2 fold changes (FC) between deletion mutant and control. This approach is in
168 accordance with previous transcriptomic studies in other archaea that have reported lower
169 log2 fold changes of statistical significance for genes of physiological relevance[40–42]. Of the
170 3002 coding genes in *Sa. solfataricus* P1, roughly 30 % were differentially expressed in both
171 KO1 and KO2 when compared to the wildtype (Fig. 5E and F, respectively). The majority of
172 significantly up- and downregulated genes was shared between the two knockout mutants
173 KO1 and KO2 (Fig. 5C and D, respectively). The top 25 upregulated genes for KO1 and KO2
174 are listed in Supplementary Tables S2 and S3, respectively. Observed differences were

175 seemingly not caused by a polar effect on the up- and downstream genes surrounding the
176 site where the targeted deletion had occurred, as significant changes in transcriptional levels
177 of genes were dispersed throughout the genome (Fig. 5G and H). A functional enrichment
178 analysis based on archaeal clusters of orthologous genes (arCOGs) was performed to
179 determine which general biological functions were affected by the deletion (Fig. 6A and B).
180 Genes involved in energy production and conversion (category C), amino acid transport and
181 metabolism (category E), and nucleotide transport and metabolism (category F) were
182 significantly enriched among the significantly upregulated genes in both KO1 and KO2. In
183 KO1, genes associated with carbohydrate transport and metabolism were additionally found
184 to be overrepresented among the significantly upregulated group of genes (Fig. 5A). In
185 contrast, there were no arCOG categories specifically enriched among the downregulated
186 group of genes.
187
188 **Absence of Cmr-ß results in significant changes in core metabolic modules**
189 In accordance with the functional enrichment analysis and the accelerated growth observed
190 in *cmr-ß* knockout strains, several key components of the energy metabolism were
191 upregulated in both knockout strains (all pathways discussed in this section are summarized
192 in Fig. 7, Supplementary Notes). Regarding the electron transport chain (ETC), most of the
193 genes encoding subunits of complex I (NADH dehydrogenase) were upregulated apart from
194 NADH dehydrogenase subunit B (NuoB), which is encoded elsewhere in the genome. In
195 Sulfolobales, c-type cytochromes are absent and the functions of complex III and IV of the
196 ETC are combined within quinone oxidoreductase supercomplexes[43]. These unusual terminal
197 oxidase complexes accept electrons directly from quinones, then in turn reduce molecular
198 oxygen and pump protons to generate a proton motive force. Transcription levels of two out
199 of three terminal oxidases in *Sa. solfataricus* were increased in *cmr-ß* knockout strains. While
200 all genes encoding for terminal oxidase SoxEFGHIM were upregulated, only half of the genes
201 encoding for terminal oxidase SoxABCDL showed significantly higher expression. The
202 hydrogenase-related subunits of a putative formate hydrogenlyase complex (FHL) present in
203 the genome were significantly upregulated in knockout strains, while the expression levels of
204 the putative formate dehydrogenase subunit (FdhA, a molybdenum-dependent enzyme) were
205 unchanged, or very slightly downregulated. In accordance with that, the periplasmic
206 component of the ABC molybdate transport system located downstream of the FdhA was
207 significantly downregulated.
208

209 Another process that was seemingly impacted on a transcriptional level by the *cmr-ß*
210 knockout was the transport of nutrients, with several different types of transporters being
211 significantly upregulated. Regarding organic carbon compounds, the arabinose ABC
212 transporter (3.A.1.1.14-AraS) and several putative peptide ABC transporters (3.A.1.5-PepT)
213 were clearly upregulated in KO1 and KO2, while the glucose ABC transporter (3.A.1.1.13-
214 GlcS) exhibited mildly increased transcript levels[44,45]. The gene encoding transcriptional
215 regulator XylR (SSOP1_RS15310), which in *S. acidocaldarius* was shown to positively
216 regulate the expression of pentose transporters as well as metabolizing genes[46], was
217 significantly upregulated. Accordingly, all but one gene with the corresponding predicted
218 binding site for XylR[47] were upregulated, among them *araS* and *glcS* encoding the substrate-
219 binding proteins of the respective ABC-transporters (Fig. 7). Interestingly, some genes of the
220 XylR regulon were recently reported to be part of an independently modulated set of genes
221 that also included the *soxEFGHIM* gene cluster, indicating that their expression might be
222 governed by a shared unknown regulator[48]. Transporters putatively involved in the uptake of
223 amino acids (2.A.1.15-The Aromatic Acid:H+ Symporter Family, 2.A.3.11-The
224 Aspartate/Glutamate Transporter Family) were also highly upregulated in both mutants.
225
226 A predicted thiamine transporter[49] (ThiT) was highly upregulated in *cmr-ß* knockout strains,
227 as well as at least one homolog of ThiC (HMP phosphate (HMP-P) synthase) catalyzing the
228 first step in the synthesis of the pyrimidine moiety of thiamine (HMP-PP), and Thi4 (thiamine
229 thiazole synthase) catalyzing the formation of the thiazole ring (HET-P)[50,51]. Additionally,
230 transcript levels of at least one out of two TenA homologs (aminopyrimidine aminohydrolase)
231 implicated in the salvage of thiamine precursors was slightly higher in both mutant strains
232 compared to wildtype. These genes are part of a proposed thiamine regulon in *Sa.*
233 *solfataricus*, which includes genes involved in the biosynthesis pathway and transport of
234 thiamine and/or precursors[49].
235
236 *Sa. solfataricus* employs a reversed monophosphate pathway for the synthesis of ribose-5-
237 phosphate, which serves as pentose moiety in nucleotide biosynthesis[52]. The genes encoding
238 the enzymes partaking in purine de novo biosynthesis until the formation of the intermediate
239 5′-phosphoribosyl-5-aminoimidazole (aminoimidazole ribotide, AIR) were upregulated in KO1
240 and KO2. Interestingly, this intermediate compound is used as a precursor during the
241 generation of HMP-PP, establishing a link to the observations regarding thiamine
242 biosynthesis[49]. In addition, slightly increased expression levels could also be observed for
243 enzymes participating in purine salvage. In contrast, the genes encoding enzymes that

244 catalyze the reaction from cytosine to uracil in the process of pyrimidine salvage were
245 significantly downregulated, likely due to the supplementation of uracil in the medium.
246
247 Genes encoding TatC (SSOP1_RS15345) and SecY (SSOP1_RS03615), components of two
248 different protein translocation pathways[53–56], were also significantly upregulated in *cmr-ß*
249 knockout strains. Interestingly, insertion of TatC itself into the membrane was shown to
250 require the Sec pathway together with translocase YidC in E. coli[57], which we also found
251 significantly upregulated in our dataset (SSOP1_RS17510). Overall, a significant enrichment
252 of genes encoding transmembrane domain proteins could be observed among the
253 upregulated fraction of the transcriptomes of both KO1 and KO2 (Fig. 8A and B, respectively).
254
255 A number of genes involved in the isoprenoid lipid biosynthesis were significantly upregulated
256 in the *cmr-ß* knockout strains. Specifically, the genes responsible for the synthesis of
257 intermediates mevalonate (thiolase - *aact*, hydroxymethylglutaryl-CoA synthase - *hmgs*,
258 hydroxymethylglutaryl-CoA reductase - *hmgr*), glycerol-1-phosphate (glycerol-1-phosphate
259 dehydrogenase) and the ether bond formation step to geranylgeranylglyceryl phosphate
260 (GGGP) synthase), as well as geranylgeranyl reductase (*ggr*) and tetraether synthase (*tes*),
261 leading to the formation of the membrane-spanning glycerol dibiphytanyl glycerol tetraethers
262 (GDGTs). Interestingly, the modifying enzyme calditol synthase (*cds*) responsible for the
263 synthesis of the cyclopentenyl head group calditol[58] was significantly upregulated, while
264 GDGT ring synthases *grsA*, *grsB1* and *grsB2*[59] were downregulated, the latter two
265 significantly.
266
267 Lastly, 28-32 putative transcriptional regulators (arCOG functional category K), characterized
268 and uncharacterized, were significantly differentially expressed in the knockout strains,
269 indicating a strong involvement of the Cmr-ß complex to various regulatory circuits. However,
270 the CRISPR-associated Csa3 transcription factors involved in regulation of various CRISPR-
271 Cas system components and coordinating the immune response[60–64] were not differentially
272 expressed according to our cutoffs.
273
274 **Response of CRISPR-Cas defense systems and transposases to the deletion of Cmr-ß**
275 Next, we wanted to investigate whether the absence of Cmr-ß had influenced the expression
276 of other genes implicated in antiviral defense. Looking at the expression levels of CRISPR-
277 *cas* modules in our wildtype samples, it became evident that the first out of three type I-A
278 *cas* modules (SSOP1_RS07365-SSOP1_RS07390) as well as the *cmr-ß* module were most

279 highly expressed. The second and third type I-A module were considerably lower and
280 basically not expressed, respectively. In comparison, the expression level of the *csm* module
281 was more than ten times lower than the one of the *cmr-ß* module under standard growth
282 conditions. Upon *cmr-ß* knockout, we could detect an upregulation of the majority of genes
283 of the first type I-A *cas* module as well as increased expression levels for the *csm* module
284 (Fig. 7).
285
286 Several transposase genes were significantly downregulated in absence of Cmr-ß. The genus
287 *Saccharolobus* generally carries a remarkable number of insertion sequence (IS) elements[65–
288 68]. For *Sa. solfataricus* P2 specifically it is estimated that IS elements make up approximately
289 10% of its total genome sequence (in bp)[69]. In the P1 reference genome, almost 170
290 functional transposases are currently annotated (based on eggnog annotations of
291 NZ_LT549890.1). Several transposases of the most prominent IS4 (ISC1225, ISC1359,
292 ISC1439) and IS5 (ISC1058, ISC1212, ISC1234) families were found to be downregulated in
293 both deletion mutants. Transposase genes belonging to the less abundant IS630 (ISC1048,
294 ISC1078), ISC1217 and IS110 (ISC1190, ISC1229) elements were also exhibiting lower
295 transcript levels in absence of the Cmr-ß complex. In contrast, although some transposases
296 of the IS200/IS605 and IS607 families were found to be downregulated, certain
297 representatives were upregulated (especially in KO2).
298
299 **Adaptation to minimal media is affected in *cmr-ß* knockout mutant**
300 Since we observed several genes related to transport being differentially expressed in
301 complex (sucrose-tryptone) media, we were interested in examining to which extent this
302 phenotype would still be observable in defined media. Wildtype and KO2 strain were adapted
303 to grow in media containing either glucose or sucrose as sole carbon source. The switch from
304 complex to defined media caused a lag phase in both strains, however it was more
305 pronounced in the *cmr-ß* knockout strain (Fig. 9A). The switch to sucrose resulted in a lag
306 phase of about 48 hours (hr) for the wildtype strain, after which it assumed a growth rate
307 similar to the one exhibited under complex media and reached stationary after 100 hr, while
308 the knockout endured both a prolonged lag phase of about 72-96 hr as well as an overall
309 reduced growth rate with linear growth for the next 96 hr, reaching stationary phase after 192
310 hr. The switch to glucose media proved more cumbersome for both strains and resulted a
311 slightly different profile. The wildtype endured a lag phase of about 72 hours, after which it
312 followed linear growth for the next 138 hours, while the knockout exhibited a very long lag
313 phase of 264 hr, after which it showed exponential growth with a growth rate only slightly

314 lower than in complex media. After adaptation, the growth curves of wildtype and KO2
315 ultimately converged in glucose medium, but not in sucrose, where the wildtype grew slower
316 than the knockout (Fig. 9B and C, respectively). Of the differentially expressed genes in
317 glucose medium, 234 (~8 % of protein-coding genes) were upregulated and 260 (~9 % of
318 protein-coding genes) were downregulated in KO2 (not shown). In sucrose medium, more
319 than twice as many genes were found to be significantly differentially expressed in KO2,
320 among those 621 (~20 %) were upregulated and 560 (~19 %) were downregulated (not
321 shown). The complex media condition served as an internal control in this second
322 experiment, and we found 287 (~10 %) and 254 (~9 %) genes to be up- and downregulated
323 in KO2, respectively (not shown). Here, the vast majority of upregulated genes was consistent
324 with our previous dataset (~73 %) (Fig 10A and Supplementary Information). Interestingly, 20
325 genes were similarly upregulated in all three conditions according to DESeq analysis (Suppl.
326 Tab. S4). Together with genes exhibiting a strong correlation with all *cmr-ß* knockout samples
327 in combined PCA analyses (Fig. 10B, Suppl. Tab. S5), the overarching theme again points
328 towards perturbations in energy metabolism, transport and lipid biosynthesis following the
329 *cmr-ß* knockout. The higher expression of these genes, regardless of medium composition,
330 strongly implicate a role of the Cmr-ß in the regulation of metabolism either directly or
331 indirectly.
332
333
334 **DISCUSSION**
335
336 **Cmr-ß might not be a major contributor to CRISPR-Cas-mediated plasmid defense in *Sa.***
337 ***solfataricus***
338 The co-existence of two different type III CRISPR-Cas complexes in *Sa. solfataricus* and other
339 Sulfolobales[70] might seem redundant at first, however is most likely indicative of different
340 mechanistic specialization which enables complementation in response to varying load and
341 types of MGEs[71,72]. In this study, we set out to investigate the specific roles of Cmr-ß in the
342 immune response of *Sa. solfataricus* P1. Interestingly, while *cmr-ß* knockout mutants were
343 readily obtained, similar attempts at deleting the *csm* locus caused large genomic deletions
344 (Wimmer unpublished) and isolated mutants could therefore not be tested. Our plasmid
345 challenge assay indicated that the transcription-dependent targeting effect of a protospacer-
346 carrying plasmid in *cmr-ß* knockout mutant was comparable to the wildtype. Hence, *Sa.*
347 *solfataricus* P1 Csm-mediated immunity leads to a similar outcome irrespective of the
348 presence of the Cmr-ß system (see Fig. 4). In analogy to SisCmr-ß[33], one could assume that

349 Cmr-ß of *Sa. solfataricus* P1 does not synthesize significant amounts of cOA molecules and
350 is possibly not a major contributor to the canonical multi-faceted type III CRISPR - mediated
351 defense.
352
353
354 **Transcriptional response to the *cmr-ß* knockout supports accelerated growth phenotype**
355 **and interaction with regulatory networks**
356 The findings of our transcriptomic analysis show that the deletion of Cmr-ß had significant
357 effects on the transcriptional landscape and growth behavior. In accordance with accelerated
358 growth following the deletion, we could observe upregulation of genes involved in distinct
359 functional categories of energy production and conversion, as well as metabolism and
360 transport of amino acids, nucleotides, and carbohydrates. The upregulation of complex I
361 (NADH dehydrogenase) and two out of three complex IVs is interesting regarding their
362 importance in maintaining pH homeostasis and generating a proton motif force in
363 thermoacidophilic organisms[43]. This observation is even more compelling considering that
364 complex I is capable of pumping protons as opposed to complex II (succinate
365 dehydrogenase)[73] and the bb3-type terminal oxidase has a higher electron to proton pumping
366 ratio than the other terminal oxidases (aa3-type SoxABCDL and DoxBCE) due to the
367 presence of two proton pumping sites[74,75]. Moreover, its subunit SoxE was suggested to act
368 as a pH sensor and regulate electron flow[74,76]. The activity of characterized FHL complexes
369 has also been connected to counteracting cytoplasm acidification through removal of
370 formate and H2 evolution[77,78]. Moreover, being membrane-bound complexes with similarities
371 to complex I, it has long been hypothesized that they might be capable of proton pumping
372 and therefore contribute to proton motive force generation, a capacity strongly supported by
373 the formate-driven growth in *Thermococcus onnurineus*[79,80]. Both these activities might
374 explain the upregulation profile of the hydrogenase component of the FHL in the knockout
375 strains.
376
377 Upregulation of ABC-type and other transporters for the uptake of substrates such as sugars,
378 peptides/amino acids and phosphate further supports the accelerated growth phenotype of
379 the knockout strains. The additional upregulation of thiamine transport and biosynthesis as
380 well as nucleotide biosynthesis seemingly underlines the need for higher ATP synthesis and
381 energy generation. The increased incorporation of membrane-anchored proteins could be
382 supported by the upregulation of pivotal components of protein translocation pathways,
383 namely *tatC, secY* and *yidC*. As central component of the Tat pathway, TatC captures folded

384 proteins destined for translocation such as subunits of terminal oxidases that assemble their
385 cofactors prior to translocation in the cytoplasm[81,82]. In contrast, ABC transporters of CUT
386 class of *Sa. solfataricus* harbor type IV pilin like signal peptides and are presumed to be
387 secreted via the Sec pathway[44,45,83].
388
389 As summarized above, the upregulation of the ETC complexes and the FHL, as well as the
390 increased organic carbon availability implied by the upregulation of the respective
391 transporters supports the increase in growth rate observed in the deletion mutants under
392 complex media. The observed upregulation in main lipid metabolism genes as well as
393 downregulation of the ring synthases would also agree with this phenotype, as it would be
394 the expected profile under logarithmic growth without energy limitations[84–87]. The observed
395 profile in the *cmr-ß* knockout could also indicate that *Sa. solfataricus* is experiencing (or
396 displaying the transcriptional profile of) a mild case of acid stress, as suggested by the
397 upregulation of calditol synthase[58]. The general transcriptional profile however is not
398 reminiscent of a classic case of acid stress, as documented in *Sa. islandicus*, which displayed
399 decreased energy production, upregulation of the proton-pumping ATPase, upregulation of
400 *grsB*, upregulation of repair and turnover pathways and increased anabolism[88]. As always
401 though, transcriptomic profiles need to be interpreted with caution, as it has been shown that
402 the observed profiles do not always translate into matching proteomic or lipidomic profiles[88].
403 An unclear profile in our case is nevertheless not surprising, as the actual imposed condition
404 (the absence of the *cmr-ß* module) is neither an environmental stressor nor a manipulation of
405 nutrient levels, and therefore the observed phenotypic changes cannot be interpreted in a
406 strict framework. Rather, we are getting a glimpse of the regulatory networks and processes
407 that the Cmr-ß is involved in in the absence of external stimulants, thereby sometimes leading
408 to contradictory profiles.
409
410 **Functional link between CRISPR systems and cell envelope processes**
411 As outlined above, our transcriptomic analysis of type III-B CRISPR-Cas deletion mutants
412 repeatedly pointed towards an involvement with membrane-associated processes.
413 Accordingly, genes encoding for transmembrane proteins were significantly enriched among
414 the upregulated fraction in type III-B deletion mutants under all conditions examined (Fig. 8).
415 In general, a connection between CRISPR-Cas systems and membrane-associated
416 processes seems intuitive as it would allow for quick responses upon entry of infecting
417 viruses.
418

419 First hints at the functional link between the cell envelope and CRISPR-Cas were already
420 gathered relatively early on, when it was shown that the incorporation of misfolded proteins
421 translocated by the Tat system likely caused proton leakage at the membrane prompting the
422 activation of the type I CRISPR-Cas system in *E.coli*[89]. In the human pathogen *Francisella*
423 *novicida*, Cas9 complexed with a small CRISPR-associated RNA (scaRNA) was found to
424 transcriptionally repress Blp, a membrane-exposed lipoprotein usually recognized by the
425 hosts immune cells[90,91]. In the gram-negative bacterium *Serratia*, the Rcs signaling pathway
426 involved in outer membrane stress response[92] downregulates type I and type III CRISPR-Cas
427 defense while simultaneously enhancing innate immunity, e.g through repression of viral cell-
428 surface receptors[93]. Genome analyses have also pointed out the enriched presence of various
429 predicted membrane proteins in the genomic neighborhood of type III CRISPR-Cas systems,
430 indicating a close association with the cell envelope[27,28], although the exact nature of this
431 proposed functional connections largely remains to be investigated *in vivo*.
432
433 **The metabolic burden of maintaining CRISPR-Cas systems**
434 Some of the observed changes in the transcriptional landscape upon deletion of the *cmr-ß*
435 gene cluster could also be explained by the reduced metabolic burden and cost associated
436 with the maintenance of defense systems. In *Streptococcus thermophilus* DGCC7710
437 harboring multiple CRISPR-Cas systems (2 x type II, type I and type III), the disruption of one
438 of the *cas9* genes gives mutants a competitive advantage in competition with wildtype cells,
439 possibly due to a reduced lag phase[94]. In contrast, *Pseudomonas aeruginosa* PA14
440 possesses only one type I-F CRISPR-Cas system, and its deletion caused reduced
441 comparative fitness in competition assays[95]. In general, the studies in *P. aeruginosa* suggest
442 that constitutive fitness costs associated with the presence of functional CRISPR-Cas
443 systems are low, however conditional costs arise due to the elicitation of the immune
444 response upon MGE exposure[95–97]. Similar to what has been observed for *S. thermophilus*[94],
445 the deletion of one out of several co-existing *cas* gene modules also caused an accelerated
446 growth phenotype in *Sa. solfataricus* P1. However, the significant congruence between
447 transcriptomes from two different deletion mutant strains despite the presence of a second
448 large genomic deletion in only one of them suggests that we are, at least in part, witnessing
449 valid functional connections with certain degree of specificity.
450
451 **Are the observed changes the result of direct regulation by Cmr-ß or secondary effects?**
452 The observed profiles of anabolic pathways and membrane composition could be interpreted
453 either as direct regulatory targets or indirectly affected due to the increase in available energy

454 and resulting increased growth rate. The amount of (both characterized and uncharacterized)
455 differentially regulated transcription factors in our dataset also indicates the entanglement of
456 CRISPR-Cas systems with global regulatory networks. On this note, a LysR-type
457 transcriptional regulator involved in control of secondary metabolites and motility was shown
458 to repress type III-A immunity in the gram-negative bacterium *Serratia*[98]. Moreover, the
459 observed differential expression of transposase genes in response to the Cmr-ß deletion
460 could also influence host gene expression and are a known source of various ncRNAs in *Sa.*
461 *solfataricus*[99–101] and other archaea[102–106]. Former studies analyzing the content of isolated
462 Cmr-ß and Csm protein complexes have revealed that both complexes are sequestering a
463 proportion non-canonical guides in addition to canonical crRNAs *in vivo*[32,107], some of which
464 could match potential cellular targets, providing the possibility for direct CRISPR-mediated
465 regulation. However, we also expect that many effects are downstream events of primary
466 direct key regulation targets such as transcription factors, or reaction to the physiological
467 changes invoked by presumed direct effects (e.g. direct upregulation of ETC transcripts
468 would result in a change in the energy budget of the cell). Our current experimental design
469 does not allow for an investigation into the molecular mechanisms by which the Cmr-ß could
470 convey the observed regulatory effects, but we are working on elucidating the direct and
471 indirect effects in the future.
472
473 **CONCLUSION**
474
475 The deletion of the type III system consistently leads to an upregulation of genes encoding
476 for membrane proteins and membrane-associated processes, the most important of which
477 concern energy metabolism, the establishment of a proton motive force and essential nutrient
478 acquisition, accompanied by an accelerated growth phenotype. As these results are in the
479 absence of MGEs, we conclude that the type III system is exerting an influence on the greater
480 metabolic functions, and therefore fitness, of the organism irrespective of its role in cellular
481 defense.
482
483
484 **ACKNOWLEDGEMENTS**
485 This research was funded in part by the Austrian Science Fund (FWF) [J 4611-B]. We thank
486 John van der Oost, Raymond Staals, Sonja V. Albers and Xu Peng for scientific discussions.
487
488

489 **DATA AVAILABILITY**

490 All datasets and supplementary tables will be made available upon publication of the
491 manuscript.

492

493

494

495

496 **MATERIALS & METHODS**

497

498 **Design of genetic constructs and cloning.** The synthetic mini CRISPR arrays (miniCRs)
499 containing one to three targeting spacers were assembled and cloned into pENTRY vectors
500 as described previously[35]. Two 800 to 900 bp sequences homologous to the upstream and
501 downstream regions of the *cmr-ß* locus were introduced into each vector to direct DNA repair
502 mechanisms. Flank 1 corresponding to the upstream region was amplified using primers
503 flank1_fw_SalI and flank1_rv, flank 2 corresponding to the downstream region was amplified
504 using primers flank2_fw and flank2_rv_EagI. The sequences of all primers used for
505 construction of genome editing constructs can be found in Suppl. Tab. S6. The resulting PCR
506 products were checked by gel-electrophoresis and purified (Monarch PCR&DNA Cleanup
507 Kit, NEB) according to manufacturer's instructions. Flanks were phosphorylated using the T4
508 Polynucleotide Kinase (NEB) and cleaved with corresponding restriction enzyme SalI-HF or
509 EagI-HF (NEB). pENTRY-miniCR vectors were double digested with SalI-HF and EagI-HF
510 (NEB). Linearized pENTRY-miniCR vectors and both flanks were fused using the Quick
511 Ligation™ Kit (NEB). Correct sequences of miniCRs and flanks were verified by Sanger
512 sequencing and consequently transferred from pENTRY vectors into destination vector pIZ-
513 GW[36,108] using the Gateway™ LR Clonase™ II Enzyme mix (Invitrogen™).

514

515 **Culturing, transformation and isolation of knockout mutants.** Uracil-auxotrophic derivate
516 M18 of *Sa. solfataricus* P1 (DSM 1616, ATCC 35091) was cultivated in Brock salts medium
517 supplemented with D(+)-Sucrose (0.2 %, w/v) (Serva), Tryptone (0.1%, w/v) (Roth) and Uracil
518 (0.0125 mg/mL) (Sigma Aldrich), in short Brock medium S/T/U. Media pH was initially
519 adjusted to 3.5 with H_2SO_4 and cultures were incubated under shaking conditions at 77 °C
520 (Zebec 2016). Growth was assessed by measuring optical density (OD) at a wavelength of
521 600nm in a spectrophotometer. Competent cells for electroporation were prepared as
522 described previously[109]. Per construct, 50 µl of M18 competent cells were mixed with 150 ng
523 of plasmid DNA and electroporated using 1-mm cuvettes in a Gene Pulser Xcell (Biorad) with

524 the following settings: 1240 V, 1000 Ω, 25 µF. Electroporated cells were transferred to equal
525 volumes of recovery solution[110] and incubated for 1 hour at 77 °C. After recovery,
526 electroporated cells were inoculated directly into 50mL selective medium containing D(+)-
527 Sucrose (0.2%, w/v) and NZ-Amine (0.1%, w/v) (Fluka Analytics). For plasmid
528 counterselection, cells harboring pIZ-cmr73ts were transferred to medium containing
529 Sucrose, Tryptone, Uracil and 5-Fluoroorotic acid (0.05 mg/mL) (5-FOA, Sigma Aldrich).
530 Consequently, dilution streaks on gellan gum plates (0.64% w/v) (Gelrite, Roth) with the same
531 media composition were performed. Deletion mutant colonies finally underwent three
532 consecutive rounds of plating (without 5-FOA) to ensure the presence of a single genotype.
533 As a control, a purified isolate of untransformed *Sa. solfataricus* M18 (wildtype) was
534 generated using the same approach. Plates were incubated under high humidity at 77°C.
535
536 **Culture PCR.** The presence of the *cmr-ß* genes was checked by adding 2 µl aliquots from
537 OD samples or colony suspensions as template in PCR reaction with primer pairs cmr-KO-
538 check_fw/rv or cmr-middle_fw/cmr-KO-check_rv. Successful removal of plasmid and
539 persistence of IS-element conferring uracil auxotrophy were verified using primers
540 ORF904_fw/rv and pyrBE_fw/rv, respectively. All primers used for culture PCR are listed in
541 Suppl. Tab S7. For all culture PCRs, full reaction volumes were loaded on 1% agarose gels
542 together with 1 kb Plus DNA ladder (Thermo Scientific). Knockout of *cmr-ß* indicated by
543 smaller bands in gel-electrophoresis of PCR products was verified by sequencing of DNA
544 from gel-purified bands (Monarch DNA extraction Kit, NEB).
545
546 **Plasmid challenge assay.** ORF406 of conjugative plasmid pNOB8 was amplified using
547 primers ORF406_fw/rv. Matching overhangs were introduced into pENTRY vectors
548 downstream of a constitutive promoter (i.e truncated version of leader sequence of CRISPR
549 array D in *Sa. solfataricus* P2) with primers pENTRY-ORF406_fw/rv. The resulting product
550 was used as template in a touchdown PCR reaction (57°C, TΔ-0,2) for fusion with the ORF406
551 amplicon. ORF406 harbors a protospacer with a 5'NGA PAM exhibiting partial
552 complementarity to spacer A53 of *Sa. solfataricus* P2[39] which is absent from P1. Using
553 primers OE-D2.109_fw/rv, this protospacer region was exchanged to match spacer 109 of
554 CRISPR array D2 (D2.109) in P1 as described previously[24]. The insertion of an additional
555 nucleotide directly downstream of the protospacer rendered the original PAM sequence
556 dysfunctional impeding type I-A CRISPR-mediated interference. Additionally, we constructed
557 a protospacer D2.109-a with an antitag (5' CTTTCAAT) matching to the 5' tag of the D2.109
558 crRNA using primer OE-D2.109-a_fw. The sequences of all primers used for construction of

559 ORF406 target constructs are listed in Suppl. Tab. S6. Transformation of wildtype and
560 deletions strain was carried out as described above, with the following alterations: prior to
561 electroporation competent cells from both strains were adjusted to the same optical densities
562 with 20 mM Sucrose, then three replicates of 50 μL cells mixed with 350 ng of plasmid DNA
563 were set up for each construct. Cells without the addition of plasmid DNA were used as
564 control. Following 1 hour recovery, each replicate was inoculated in 50 mL selective medium.
565 After 41 hours (i.e. before cultures had started to grow), 45 μl of serial dilutions from 10^1 to
566 10^{-5} from all cultures were spotted on selective gellan gum plates containing D(+)-Sucrose
567 (0.2%, w/v) and NZ-Amine (0.1%, w/v).
568
569 **Adaptation experiment.** M18 wildtype and KO2 were inoculated from glycerol stocks into
570 prewarmed Brock medium S/T/U, 50 mL, pH 3.5, and incubated under shaking conditions at
571 77°C. As a preadaptation step, aliquots of both strains were then transferred to Brock
572 medium G/T/U supplemented with D(+)-Glucose (Sigma Aldrich) instead of Sucrose at the
573 same final concentration. A second aliquot was again transferred to Brock S/T/U. For the
574 adaptation growth curve, 4 replicate cultures were set up for the following media
575 compositions: Brock medium Sucrose (0.3%, w/v) with Uracil, Brock medium Glucose (0.3%,
576 w/v) with Uracil, and standard complex Brock medium S/T/U. When cultures had reached
577 late exponential growth, one replicate was transferred to 4 replicate flasks with fresh medium
578 of the same composition until growth had stabilized for two consecutive transfers. At that
579 point, 5 mL samples were taken at OD_{600}=0.2 for RNA preparation. Throughout the
580 experiment, cultures were inoculated to the same initial OD_{600}.
581
582 **DNA preparation and genome mutation analysis.** DNA was extracted as described in ref[24].
583 The sequencing was performed by the Next Generation Sequencing Facility at at Vienna
584 BioCenter Core Facilities (VBCF), member of the Vienna BioCenter (VBC), Austria. Samples
585 were sequenced (150 PE) sequencing on an Illumina MiSeq platform with a Westburg DNA
586 preparation kit resulting in 300,000-500,000 reads per sample. Raw reads were downloaded
587 from the VBCF website and checked using md5sum. FastQC[111] version 0.11.8 and MultiQC[112]
588 version 1.8 were used to initially analyze sequences and the samtools package[113] version
589 1.10 (using htslib 1.10[114]) along with bedtools [115] version 2.29.0 was used to convert file types.
590 Trimmomatic[116] version 0.36 trimmed adapter sequences and cropped reads using the
591 options: SLIDINGWINDOW:5:20 LEADING:5 TRAILING:5 MINLEN:50 HEADCROP:12. The
592 stand-alone program PrinSeq Lite[117] version 0.20.4 was used to remove sequences < 38 bp
593 and/or with quality scores < 30. Next, Bowtie2[118] version 2.3.5.1 was used to align the

594 remaining RNA sequences to the *Sa. solfataricus* strain P1 reference genome
595 (NZ_LT549890.1) downloaded from the National Center for Biotechnology Information
596 (NCBI). Mutations, insertions, and deletions were then found using breseq[119,120] version
597 0.32.1 with default settings using the NCBI RefSeq annotation from 02/19/2020.
598
599 **RNA preparation and sequencing analysis.** RNA was extracted from 5 mL of culture
600 sampled at OD_{600}=0.2 as described in ref[24]. Sample quality control, library preparation (150
601 PE) from rRNA-depleted samples (to approx. 50%) and sequencing on an Illumina NovaSeq
602 platform was performed by Novogene (UK) Co., Ltd. Raw reads were downloaded from the
603 sequencing facilities' website and checked using md5sum. Adapter sequences were trimmed
604 using Trimmomatic[116] version 0.39 using the following settings: HEADCROP:12
605 SLIDINGWINDOW:4:15 LEADING:3 TRAILING:3 MINLEN:30. Reads were filtered for a
606 minimum mean quality score of 30 using PrinSeq Lite[117] version 0.20.4. Reads mapping to
607 ribosomal RNA were removed from total reads with the sortmeRNA tool[121] version 4.3.6 by
608 supplying a curated reference file containing *Sa. solfataricus* P1 rRNA sequences. The
609 remaining reads were aligned to the *Sa. solfataricus* P1 reference genome (NCBI:
610 NZ_LT549890.1) using Hisat2[122] version 2.2.1 with parameter --rna-strandness RF. The
611 resulting sequence alignment files were converted to binary format using the samtools
612 package[113] version 1.20. featureCounts from the Subread package [123] version 2.0.6 assigned
613 mapped reads to genomic features (feature and attribute type specified as "gene" and "ID",
614 respectively) using a General Feature Format (GFF) annotation file with RefSeq annotations
615 downloaded from the NCBI website on as a reference. To assess expression of transposase
616 genes, a separate run with featureCounts was performed to inlcude reads mapping at
617 multiple positions within the genome ('-M' option for multimapping reads). The output count
618 matrix was used to perform differential expression analysis with the DESseq2 package[124]
619 version 1.42.1 in R[125] version 4.3.3 with Rstudio[126] version 2023.12.1+402. Default p-value
620 adjustment in DESeq2 (Benjamini-Hochberg method) was used to determine relevant p-
621 values (referred to as p-adj in result tables). Principal component analysis (PCA) was
622 performed as described previoulsy[40].
623
624 **Assignment of transporter families, arCOGs and transmembrane domains, and functional**
625 **enrichment analysis.** The BLAST tool integrated into the curated Transporter Classification
626 Database (TCDB)[127] was used to classify transporters. Archaeal Clusters of Orthologous
627 Groups (arCOGs)[128] were assigned to genes according to database version 2018. Genes
628 without an arCOG assignment were annotated using eggNOG-mapper[129] version 2.1.12.

Genes encooding proteins with predicted transmembrane domains were identified using Phobius[130]. In both cases, a modified statistical hypergeometric test to determine functional categories significantly enriched for *Sa. solfataricus* (*P*-value < 0.05) among the up- or downregulated genes was conducted using the phyper function in R[40].

References

1. Pourcel, C. *et al.* CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res* (2019) doi:10.1093/nar/gkz915.
2. Jackson, S. A. *et al.* CRISPR-Cas: Adapting to change. *Science (1979)* **356**, (2017).
3. McGinn, J. & Marraffini, L. A. Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nat Rev Microbiol* **17**, 7–12 (2019).
4. Shmakov, Sergey; Sitnik, Vassilii; Makarova, Kira; Wolf, Yuri; Severinov, Konstantin; Koonin, E. The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. *mBio* **8**, 1–18 (2017).
5. Makarova, K. S. *et al.* Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* **18**, (2020).
6. van Beljouw, S. P. B., Sanders, J., Rodríguez-Molina, A. & Brouns, S. J. J. RNA-targeting CRISPR–Cas systems. *Nat Rev Microbiol* **21**, 21–34 (2023).
7. Elmore, J. R. *et al.* Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR–Cas system. *Genes Dev* **30**, 447–459 (2016).
8. Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č. & Siksnys, V. Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. *Mol Cell* **62**, 295–306 (2016).
9. Estrella, M. A., Kuo, F.-T. & Bailey, S. RNA-activated DNA cleavage by the Type III-B CRISPR–Cas effector complex. *Genes Dev* **30**, 460–470 (2016).
10. Liu, T. Y., Liu, J. J., Aditham, A. J., Nogales, E. & Doudna, J. A. Target preference of Type III-A CRISPR-Cas complexes at the transcription bubble. *Nat Commun* **10**, (2019).
11. Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G. & Siksnys, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science (1979)* **357**, 605–609 (2017).
12. Niewoehner, O. *et al.* Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543 (2017).
13. Han, W. *et al.* A Type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding. *Nucleic Acids Res* **46**, 10319–10330 (2018).
14. Rouillon, C., Athukoralage, J. S., Graham, S., Grüschow, S. & White, M. F. Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* **7**, 1–25 (2018).
15. Steens, J. A., Salazar, C. R. P. & Staals, R. H. J. The diverse arsenal of type III CRISPR–Casassociated CARF and SAVED effectors. *Biochemical Society Transactions* vol. 50 1353–1364 Preprint at https://doi.org/10.1042/BST20220289 (2022).
16. Makarova, K. S. *et al.* Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antivirus defense. *Nucleic Acids Res* **48**, 31 (2020).

677 17. Steens, J. A. *et al.* Type III-B CRISPR-Cas cascade of proteolytic cleavages. *Science*
678 *(1979)* **383**, 512–519 (2024).
679 18. Rostøl, J. T. & Marraffini, L. A. Non-specific degradation of transcripts promotes
680 plasmid clearance during type III-A CRISPR–Cas immunity. *Nat Microbiol* **4**, 656–662
681 (2019).
682 19. Rostøl, J. T. *et al.* The Card1 nuclease provides defence during type III CRISPR
683 immunity. *Nature* **590**, 624–629 (2021).
684 20. McMahon, S. A. *et al.* Structure and mechanism of a Type III CRISPR defence DNA
685 nuclease activated by cyclic oligoadenylate. *Nat Commun* **11**, 1–11 (2020).
686 21. Mayo-Muñoz, D. *et al.* Type III CRISPR-Cas provides resistance against nucleus-
687 forming jumbo phages via abortive infection. *Mol Cell* **82**, 4471-4486.e9 (2022).
688 22. Hogrel, G. *et al.* Cyclic nucleotide-induced helical structure activates a TIR immune
689 effector. *Nature* 1–5 (2022) doi:10.1038/s41586-022-05070-9.
690 23. Rouillon, C. *et al.* Antiviral signalling by a cyclic nucleotide activated CRISPR protease.
691 *Nature* **614**, 168–174 (2023).
692 24. Zebec, Z., Manica, A., Zhang, J., White, M. F. & Schleper, C. CRISPR-mediated
693 targeted mRNA degradation in the archaeon Sulfolobus solfataricus. *Nucleic Acids*
694 *Res* **42**, 5280–5288 (2014).
695 25. Pan, S. *et al.* A seed motif for target RNA capture enables efficient immune defence by
696 a type III-B CRISPR-Cas system. *RNA Biol* **16**, 1166–1178 (2019).
697 26. Steens, J. A. *et al.* SCOPE enables type III CRISPR-Cas diagnostics using flexible
698 targeting and stringent CARF ribonuclease activation. *Nat Commun* **12**, 1–12 (2021).
699 27. Shmakov, S. A., Makarova, K. S., Wolf, Y. I., Severinov, K. V. & Koonin, E. V.
700 Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene
701 neighborhood analysis. *Proceedings of the National Academy of Sciences* **115**,
702 E5307–E5316 (2018).
703 28. Shah, S. A. *et al.* Comprehensive search for accessory proteins encoded with archaeal
704 and bacterial type III CRISPR-cas gene cassettes reveals 39 new cas gene families.
705 *RNA Biol* **00**, 1–13 (2018).
706 29. Silas, S. *et al.* Type III CRISPR-Cas systems can provide redundancy to counteract
707 viral escape from type I systems. *Elife* **6**, (2017).
708 30. Garrett, R. *et al.* CRISPR-Cas Adaptive Immune Systems of the Sulfolobales:
709 Unravelling Their Complexity and Diversity. *Life* **5**, 783–817 (2015).
710 31. Deng, L., Garrett, R. A., Shah, S. A., Peng, X. & She, Q. A novel interference
711 mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol Microbiol* **87**, 1088–
712 1099 (2013).
713 32. Zhang, J. *et al.* Structure and Mechanism of the CMR Complex for CRISPR-Mediated
714 Antiviral Immunity. *Mol Cell* **45**, 303–313 (2012).
715 33. Sofos, N. *et al.* Structures of the Cmr-β Complex Reveal the Regulation of the
716 Immunity Mechanism of Type III-B CRISPR-Cas. *Mol Cell* **79**, 741-757.e7 (2020).
717 34. Gudbergsdottir, S. *et al.* Dynamic properties of the Sulfolobus CRISPR/Cas and
718 CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and
719 protospacers. *Mol Microbiol* **79**, 35–49 (2011).
720 35. Wimmer, E., Zink, I. A. & Schleper, C. Reprogramming CRISPR-Mediated RNA
721 Interference for Silencing of Essential Genes in Sulfolobales. in *Methods in Molecular*
722 *Biology* vol. 2522 177–201 (Humana Press Inc., 2022).
723 36. Zink, I. A. *et al.* CRISPR-mediated gene silencing reveals involvement of the archaeal
724 S-layer in cell division and virus infection. *Nat Commun* **10**, 1–14 (2019).
725 37. Samai, P. *et al.* Co-transcriptional DNA and RNA cleavage during type III CRISPR-cas
726 immunity. *Cell* **161**, 1164–1174 (2015).
727 38. Han, W. *et al.* A type III-B CRISPR-Cas effector complex mediating massive target
728 DNA destruction. *Nucleic Acids Res* **45**, 1983–1993 (2017).

39. Manica, A., Zebec, Z., Teichmann, D. & Schleper, C. In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol Microbiol* **80**, 481–491 (2011).
40. Reyes, C. *et al.* Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon Nitrososphaera viennensis upon exposure to copper limitation. *ISME Journal* **14**, 2659–2674 (2020).
41. Grünberger, F. *et al.* Uncovering the temporal dynamics and regulatory networks of thermal stress response in a hyperthermophile using transcriptomics and proteomics. *mBio* **14**, (2023).
42. Esteves, A. M. *et al.* Combined transcriptomics–metabolomics profiling of the heat shock response in the hyperthermophilic archaeon Pyrococcus furiosus. *Extremophiles* **23**, 101–118 (2019).
43. Auernik, K. S. & Kelly, R. M. Identification of components of electron transport chains in the extremely thermoacidophilic crenarchaeon Metallosphaera sedula through iron and sulfur compound oxidation transcriptomes. *Appl Environ Microbiol* **74**, 7723–7732 (2008).
44. Albers, S.-V. *et al. Glucose Transport in the Extremely Thermoacidophilic Sulfolobus Solfataricus Involves a High-Affinity Membrane-Integrated Binding Protein. JOURNAL OF BACTERIOLOGY* vol. 181 https://journals.asm.org/journal/jb (1999).
45. Elferink, M. G. L., Albers, S. V., Konings, W. N. & Driessen, A. J. M. Sugar transport in Sulfolobus solfataricus is mediated by two families of binding protein-dependent ABC transporters. *Mol Microbiol* **39**, 1494–1503 (2001).
46. van der Kolk, N. *et al.* Identification of XylR, the Activator of Arabinose/Xylose Inducible Regulon in Sulfolobus acidocaldarius and Its Application for Homologous Protein Expression. *Front Microbiol* **11**, 1066 (2020).
47. Brouns, S. J. J. *et al.* Identification of the Missing Links in Prokaryotic Pentose Oxidation Pathways. *Journal of Biological Chemistry* **281**, 27378–27388 (2006).
48. Chauhan, S. M. *et al.* Machine Learning Uncovers a Data-Driven Transcriptional Regulatory Network for the Crenarchaeal Thermoacidophile Sulfolobus acidocaldarius. *Front Microbiol* **12**, (2021).
49. Rodionov, D. A., Leyn, S. A., Li, X. & Rodionova, I. A. A novel transcriptional regulator related to thiamine phosphate synthase controls thiamine metabolism genes in Archaea. *J Bacteriol* **199**, 743–759 (2017).
50. Rodionov, D. A., Vitreschak, A. G., Mironov, A. A. & Gelfand, M. S. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *Journal of Biological Chemistry* **277**, 48949–48959 (2002).
51. Palmer, L. D. & Downs, D. M. The thiamine biosynthetic enzyme thic catalyzes multiple turnovers and is inhibited by S-Adenosylmethionine (AdoMet) metabolites. *Journal of Biological Chemistry* **288**, 30693–30699 (2013).
52. Soderberg, T. Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. *Archaea* **1**, 347–352 (2005).
53. Albers, S. V., Szabó, Z. & Driessen, A. J. M. Protein secretion in the Archaea: Multiple paths towards a unique cell surface. *Nat Rev Microbiol* **4**, 537–547 (2006).
54. Van Den Berg, B. *et al. X-Ray Structure of a Protein-Conducting Channel.* www.nature.com/nature (2003).
55. Hamsanathan, S., Anthonymuthu, T. S., Bageshwar, U. K. & Musser, S. M. A Hinged Signal Peptide Hairpin Enables Tat-Dependent Protein Translocation. *Biophys J* **113**, 2650–2668 (2017).
56. Szabo, Z. & Pohlschroder, M. Diversity and subcellular distribution of archaeal secreted proteins. *Front Microbiol* **3**, (2012).

57. Zhu, L., Klenner, C., Kuhn, A. & Dalbey, R. E. Both yidc and SecYEG are required for translocation of the periplasmic loops 1 and 2 of the multispanning membrane protein tatC. *J Mol Biol* **424**, 354–367 (2012).
58. Zeng, Z., Liu, X. L., Wei, J. H., Summons, R. E. & Welander, P. V. Calditol-linked membrane lipids are required for acid tolerance in Sulfolobus acidocaldarius. *Proc Natl Acad Sci U S A* **115**, 12932–12937 (2018).
59. Zeng, Z. *et al.* GDGT cyclization proteins identify the dominant archaeal sources of tetraether lipids in the ocean. *Proc Natl Acad Sci U S A* **116**, 22505–22511 (2019).
60. Liu, Z. *et al.* A CRISPR-associated factor Csa3a regulates DNA damage repair in Crenarchaeon Sulfolobus islandicus. *Nucleic Acids Res* **48**, 9681–9693 (2020).
61. Charbonneau, A. A., Eckert, D. M., Gauvin, C. C., Lintner, N. G. & Lawrence, C. M. Cyclic tetra-adenylate (Ca4) recognition by csa3; implications for an integrated class 1 crispr-cas immune response in saccharolobus solfataricus. *Biomolecules* **11**, (2021).
62. Ye, Q. *et al.* CRISPR-Associated Factor Csa3b Regulates CRISPR Adaptation and Cmr-Mediated RNA Interference in Sulfolobus islandicus. *Front Microbiol* **11**, (2020).
63. He, F., Vestergaard, G., Peng, W., She, Q. & Peng, X. CRISPR-Cas type I-A Cascade complex couples viral infection surveillance to host transcriptional regulation in the dependence of Csa3b. *Nucleic Acids Res* **45**, 1902–1913 (2017).
64. Xia, P., Dutta, A., Gupta, K., Batish, M. & Parashar, V. Structural basis of cyclic oligoadenylate binding to the transcription factor Csa3 outlines cross talk between type III and type I CRISPR systems. *Journal of Biological Chemistry* **298**, 101591 (2022).
65. Guo, L. *et al.* Genome analyses of icelandic strains of Sulfolobus islandicus, model organisms for genetic and virus-host interaction studies. *J Bacteriol* **193**, 1672–1680 (2011).
66. Brügger, K. *et al.* Mobile elements in archaeal genomes. *FEMS Microbiol Lett* **206**, 131–141 (2002).
67. Filée, J., Siguier, P. & Chandler, M. Insertion Sequence Diversity in Archaea. *Microbiology and Molecular Biology Reviews* **71**, 121–157 (2007).
68. Medvedeva, S. *et al.* New insights into the diversity and evolution of the archaeal mobilome from three complete genomes of Saccharolobus shibatae. *Environ Microbiol* **23**, 4612–4630 (2021).
69. She, Q. *et al.* The complete genome of the crenarchaeon Sulfolobus solfataricus P2. *Proc Natl Acad Sci U S A* **98**, 7835–7840 (2001).
70. Zink, I. A., Wimmer, E. & Schleper, C. Heavily Armed Ancestors: CRISPR Immunity and Applications in Archaea with a Comparative Analysis of CRISPR Types in Sulfolobales. *Biomolecules* **10**, 1–41 (2020).
71. Peng, W., Feng, M., Feng, X., Liang, Y. X. & She, Q. An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res* **43**, 406–417 (2015).
72. Zhang, J., Graham, S., Tello, A., Liu, H. & White, M. F. Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. *Nucleic Acids Res* **44**, 1789–1799 (2016).
73. Schäfer, G., Engelhard, M. & Müller, V. Bioenergetics of the Archaea. *Microbiology and Molecular Biology Reviews* **63**, 570–620 (1999).
74. Komorowski, L., Verheyen, W. & Schäfer, G. The Archaeal Respiratory Supercomplex SoxM from S. acidocaldarius Combines Features of Quinole and Cytochrome c Oxidases. *Biol. Chem.* **383**, 1791–1799 (2002).
75. Gleißner, M., Kaiser, U., Antonopoulos, E. & Schäfer, G. The Archaeal SoxABCD Complex Is a Proton Pump in Sulfolobus acidocaldarius. *Journal of Biological Chemistry* **272**, 8417–8426 (1997).

76. Komorowski, L., Anemüller, S. & Schäfer, G. First expression and characterization of a recombinant CuA-containing subunit II from an archaeal terminal oxidase complex. *J Bioenerg Biomembr* **33**, 27–34 (2001).
77. Böck, A. & Sawers, G. Fermentation. in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (ed. Neidhardt, F. C.) 262–282 (American Society for Microbiology Press, Washington, DC, 1996).
78. Takács, M. *et al.* Formate hydrogenlyase in the hyperthermophilic archaeon, Thermococcus litoralis. *BMC Microbiol* **8**, 88 (2008).
79. Kim, Y. J. *et al.* Formate-driven growth coupled with H2 production. *Nature* **467**, 352–355 (2010).
80. Peters, K. & Sargent, F. Formate hydrogenlyase, formic acid translocation and hydrogen production: dynamic membrane biology during fermentation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1864**, 148919 (2023).
81. Berks, B. C., Palmer, T. & Sargent, F. Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Current Opinion in Microbiology* vol. 8 174–181 Preprint at https://doi.org/10.1016/j.mib.2005.02.010 (2005).
82. Berks, B. C. A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**, 393–404 (1996).
83. Albers, S.-V., Konings, W. N. & Driessen, A. J. M. A unique short signal sequence in membrane-anchored proteins of Archaea. *Mol Microbiol* **31**, 1595–1596 (1999).
84. Zhou, A. *et al.* Energy flux controls tetraether lipid cyclization in Sulfolobus acidocaldarius. *Environ Microbiol* **22**, 343–353 (2020).
85. Quehenberger, J., Pittenauer, E., Allmaier, G. & Spadiut, O. The influence of the specific growth rate on the lipid composition of Sulfolobus acidocaldarius. *Extremophiles* **24**, 413–420 (2020).
86. Cobban, A. *et al.* Multiple environmental parameters impact lipid cyclization in Sulfolobus acidocaldarius. *Environ Microbiol* **22**, 4046–4056 (2020).
87. Jensen, S. *et al.* The Effects of Temperature and Growth Phase on the Lipidomes of Sulfolobus islandicus and Sulfolobus tokodaii. *Life* **5**, 1539–1566 (2015).
88. Chiu, B. K. *et al.* Membrane lipid and expression responses of Saccharolobus islandicus REY15A to acid and cold stress. *Front Microbiol* **14**, (2023).
89. Perez-Rodriguez, R. *et al.* Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in Escherichia coli. *Mol Microbiol* **79**, 584–599 (2011).
90. Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y.-L. & Weiss, D. S. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* **497**, 254–257 (2013).
91. Ratner, H. K. *et al.* Catalytically Active Cas9 Mediates Transcriptional Interference to Facilitate Bacterial Virulence. *Mol Cell* **75**, 498-510.e5 (2019).
92. Meng, J., Young, G. & Chen, J. The Rcs System in Enterobacteriaceae: Envelope Stress Responses and Virulence Regulation. *Frontiers in Microbiology* vol. 12 Preprint at https://doi.org/10.3389/fmicb.2021.627104 (2021).
93. Smith, L. M. *et al.* The Rcs stress response inversely controls surface and CRISPR–Cas adaptive immunity to discriminate plasmids and phages. *Nat Microbiol* **6**, 162–172 (2021).
94. Vale, P. F. *et al.* Costs of CRISPR-Cas-mediated resistance in Streptococcus thermophilus. *Proceedings of the Royal Society B: Biological Sciences* **282**, (2015).
95. Westra, E. R. *et al.* Parasite exposure drives selective evolution of constitutive versus inducible defense. *Current Biology* **25**, 1043–1049 (2015).
96. Meaden, S. *et al.* Phage gene expression and host responses lead to infection-dependent costs of CRISPR immunity. *ISME Journal* **15**, 534–544 (2021).
97. Alseth, E. O. *et al.* Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature* **574**, 549–552 (2019).

881 98. Smith, L. M. *et al.* CRISPR-Cas immunity is repressed by the LysR-type transcriptional
882 regulator PigU. *Nucleic Acids Res* **52**, 755–768 (2024).
883 99. Zago, M. A., Dennis, P. P. & Omer, A. D. The expanding world of small RNAs in the
884 hyperthermophilic archaeon Sulfolobus solfataricus. *Mol Microbiol* **55**, 1812–1828
885 (2005).
886 100. Tang, T. H. *et al.* Identification of novel non-coding RNAs as potential antisense
887 regulators in the archaeon Sulfolobus solfataricus. *Mol Microbiol* **55**, 469–481 (2005).
888 101. Wurtzel, O. *et al.* A single-base resolution map of an archaeal transcriptome. *Genome*
889 *Res* **20**, 133–141 (2010).
890 102. Straub, J. *et al.* Small RNAs in haloarchaea: Identification, differential expression and
891 biological function. *RNA Biol* **6**, 281–292 (2009).
892 103. Jäger, D. *et al.* Deep sequencing analysis of the Methanosarcina mazei Gö 1
893 transcriptome in response to nitrogen availability. *Proc Natl Acad Sci U S A* **106**,
894 21878–21882 (2009).
895 104. Jäger, D., Förstner, K. U., Sharma, C. M., Santangelo, T. J. & Reeve, J. N. Primary
896 transcriptome map of the hyperthermophilic archaeon Thermococcus kodakarensis.
897 *BMC Genomics* **15**, 684 (2014).
898 105. Gelsinger, D. R. & DiRuggiero, J. Transcriptional landscape and regulatory roles of
899 small noncoding RNAs in the oxidative stress response of the haloarchaeon Haloferax
900 volcanii. *J Bacteriol* **200**, 1–17 (2018).
901 106. Babski, J. *et al.* Genome-wide identification of transcriptional start sites in the
902 haloarchaeon Haloferax volcanii based on differential RNA-Seq (dRNA-Seq). *BMC*
903 *Genomics* **17**, 629 (2016).
904 107. Rouillon, C. *et al.* Structure of the CRISPR interference complex CSM reveals key
905 similarities with cascade. *Mol Cell* **52**, 124–134 (2013).
906 108. Zink, I. A. Development of a plasmid shuttle vector system for CRISPR studies in the
907 archaeon Sulfolobus solfataricus. (University of Vienna, Vienna, Austria, 2013).
908 109. Manica, A., Zebec, Z., Steinkellner, J. & Schleper, C. Unexpectedly broad target
909 recognition of the CRISPR-mediated virus defence system in the archaeon Sulfolobus
910 solfataricus. *Nucleic Acids Res* **41**, 10509–10517 (2013).
911 110. Wagner, M. *et al.* Versatile genetic tool box for the crenarchaeote Sulfolobus
912 acidocaldarius. *Front Microbiol* **3**, 1–12 (2012).
913 111. Andrews, S. FastQC: a quality control tool for high throughput sequence data. Preprint
914 at (2010).
915 112. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis
916 results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048
917 (2016).
918 113. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
919 2078–2079 (2009).
920 114. Bonfield, J. K. *et al.* HTSlib: C library for reading/writing high-throughput sequencing
921 data. *Gigascience* **10**, (2021).
922 115. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic
923 features. *Bioinformatics* **26**, 841–842 (2010).
924 116. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
925 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
926 117. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic
927 datasets. *Bioinformatics* **27**, 863–864 (2011).
928 118. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*
929 *Methods* **9**, 357–359 (2012).
930 119. Deatherage, D. E. & Barrick, J. E. Identification of Mutations in Laboratory-Evolved
931 Microbes from Next-Generation Sequencing Data Using breseq. in *Engineering and*

Analyzing Multicellular Systems. Methods in Molecular Biology (eds. Sun, L. & Shou, W.) vol. 1151 165–188 (Humana Pres, New York, NY, 2014).

120. Barrick, J. E. *et al.* Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. *BMC Genomics* **15**, 1039 (2014).
121. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
122. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907–915 (2019).
123. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
124. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
125. R Foundation for Statistical Computing. A language and environment for statistical computing. Preprint at (2016).
126. Team Rs. RStudio: Integrated Development for R. Preprint at (2016).
127. Saier, M. H. *et al.* The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res* **49**, D461–D467 (2021).
128. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Archaeal clusters of orthologous genes (arCOGs): An update and application for analysis of shared features between thermococcales, methanococcales, and methanobacteriales. *Life* **5**, 818–840 (2015).
129. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol Biol Evol* **38**, 5825–5829 (2021).
130. Käll, L., Krogh, A. & Sonnhammer, E. L. L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *J Mol Biol* **338**, 1027–1036 (2004).

FIGURES

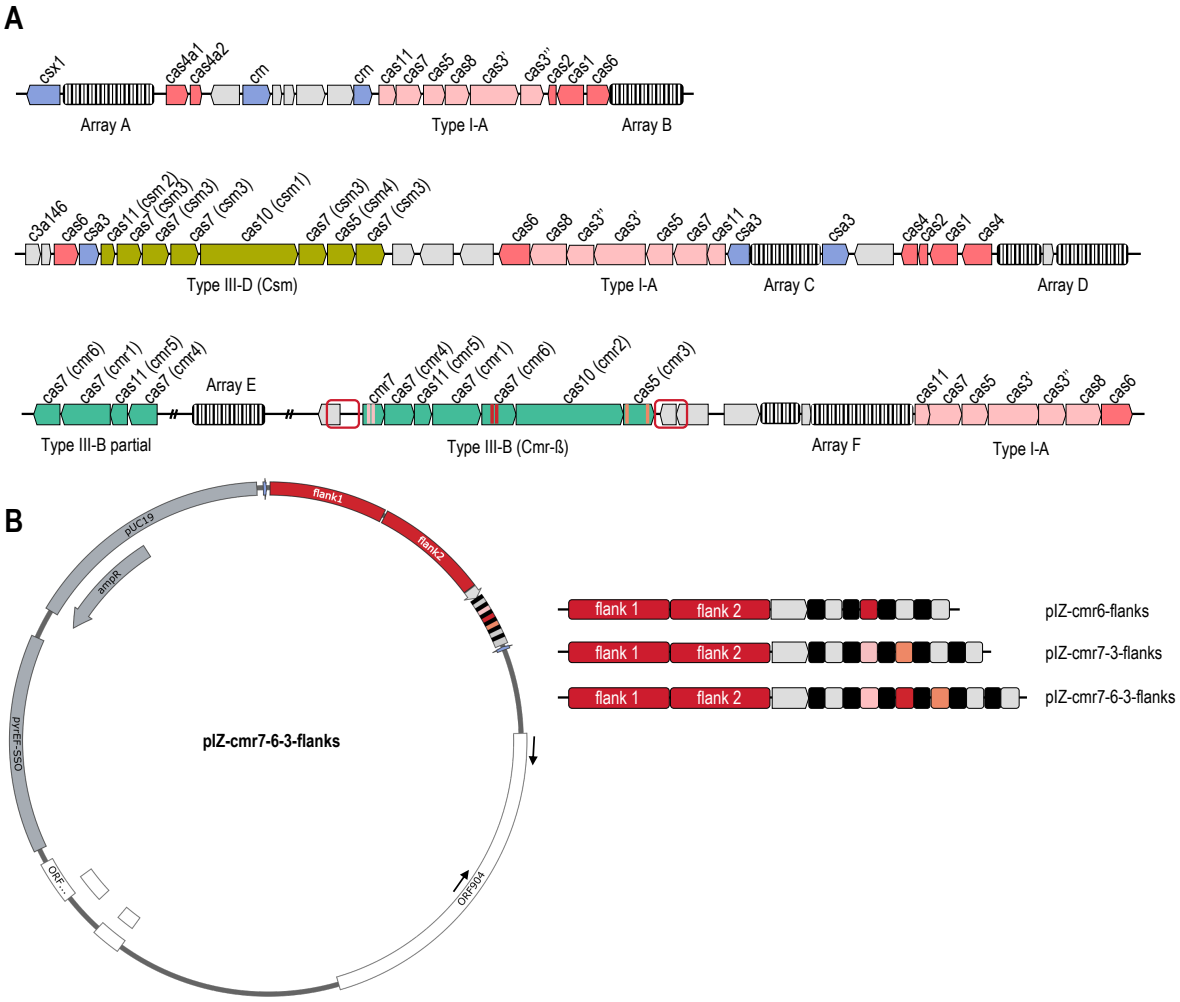


Figure 1. (A) Overview of the CRISPR-Cas systems in *Sa. solfataricus* P1. Three type I-A (pink), one type III-D (lime green), and one type III-B (green) system are present. In addition, a partial type III-B *cas* gene cluster can be detected. The strain harbors six CRISPR arrays (striped boxes), two are split through the insertion of fragments, in the case of array F stemming from a pNOB8-like conjugative plasmid. Six different protospacers that were chosen for the type I-A CRISPR-mediated deletion of the *cmr-ß* gene module are indicated. Two protospacers each are located in genes *cmr7* (pink), *cmr6* (red) and *cmr3* (salmon). Flanking sequences of the *cmr-ß* genes (red boxes) are provided as donor for the DNA repair process (see B). **(B)** Different combinations of artificial spacers (single to triple) all targeting either template or coding strand of the respective genes were cloned into a miniCR array on the plasmid-based vector pIZ, preceded by donor DNA sequences (red). Color code for miniCR: self-targeting spacers – equivalent to (A); non-targeting spacers (used for cloning procedure)– grey; repeats – black.

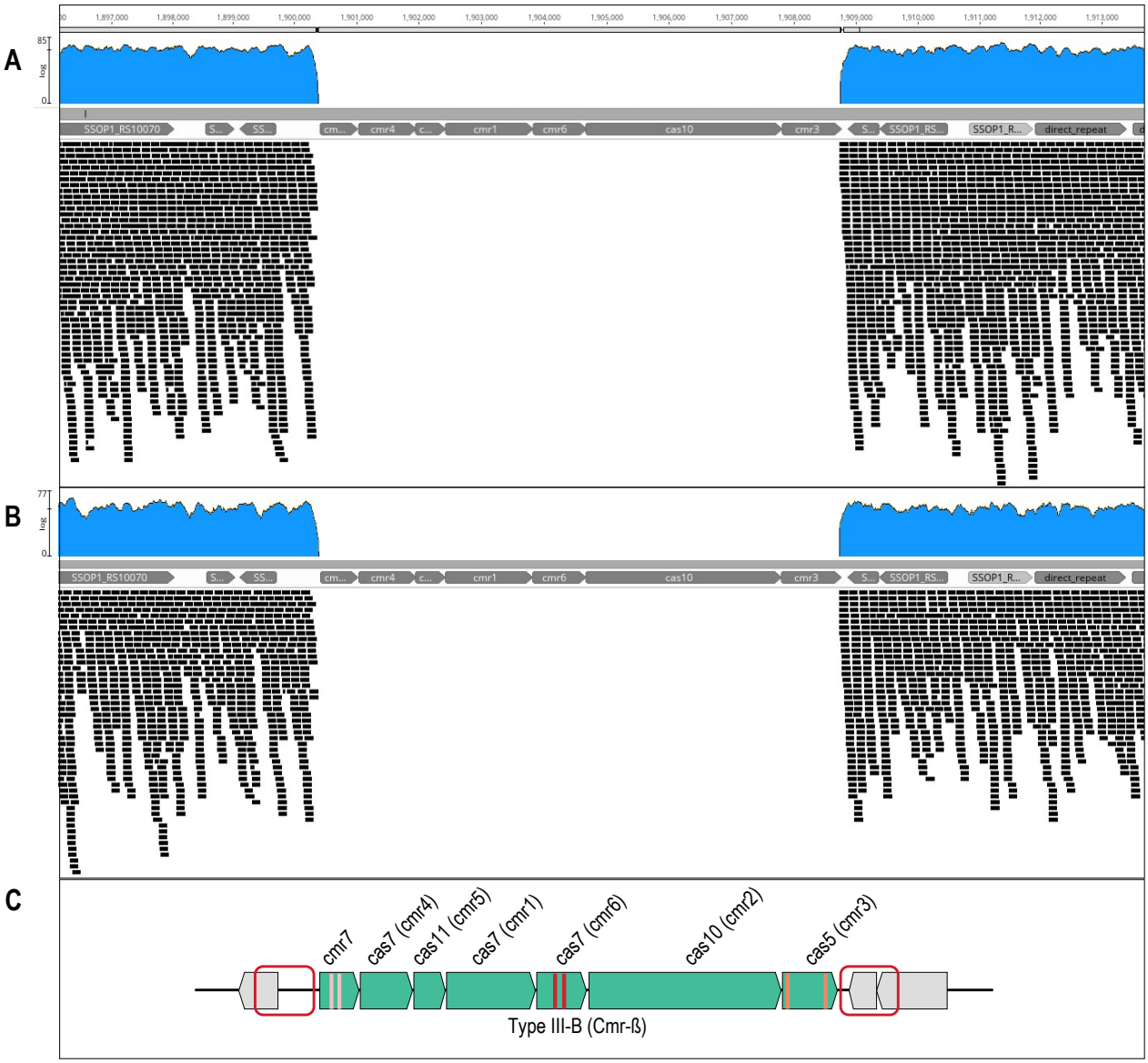


Figure 2. Genome viewer snapshot of *cmr-ß* knockout mutants. Paired-end reads (150 bp) generated from **(A)** KO1 and **(B)** KO2 were mapped to the *Sa. solfataricus* P1 NCBI reference genome (NZ_LT549890.1). Reference genes are displayed as grey arrows. The upper panel shows read coverage (blue) and the lower panel shows aligned reads (black). The excision of the *cmr-ß* gene cluster and its promoter as depicted in **(C)** was confirmed to have happened as intended for both isolated mutant strains.

Figure 3. Genome viewer snapshot of second deletion that occurred spontaneously in *cmr-ß* mutant KO2. Paired-end reads (150 bp) were mapped to the *Sa. solfataricus* P1 NCBI reference genome (NZ_LT549890.1). Reference genes are displayed as grey arrows. The upper panel shows read coverage (blue) and the lower panel shows aligned reads (black). The deleted region indicated by a red arrow spans over 37 kb and is borderd by two IS5 family transposase (pseudo-) genes. Read alignments within the indicated deleted region correspond to copies of different IS4 family transposase genes.

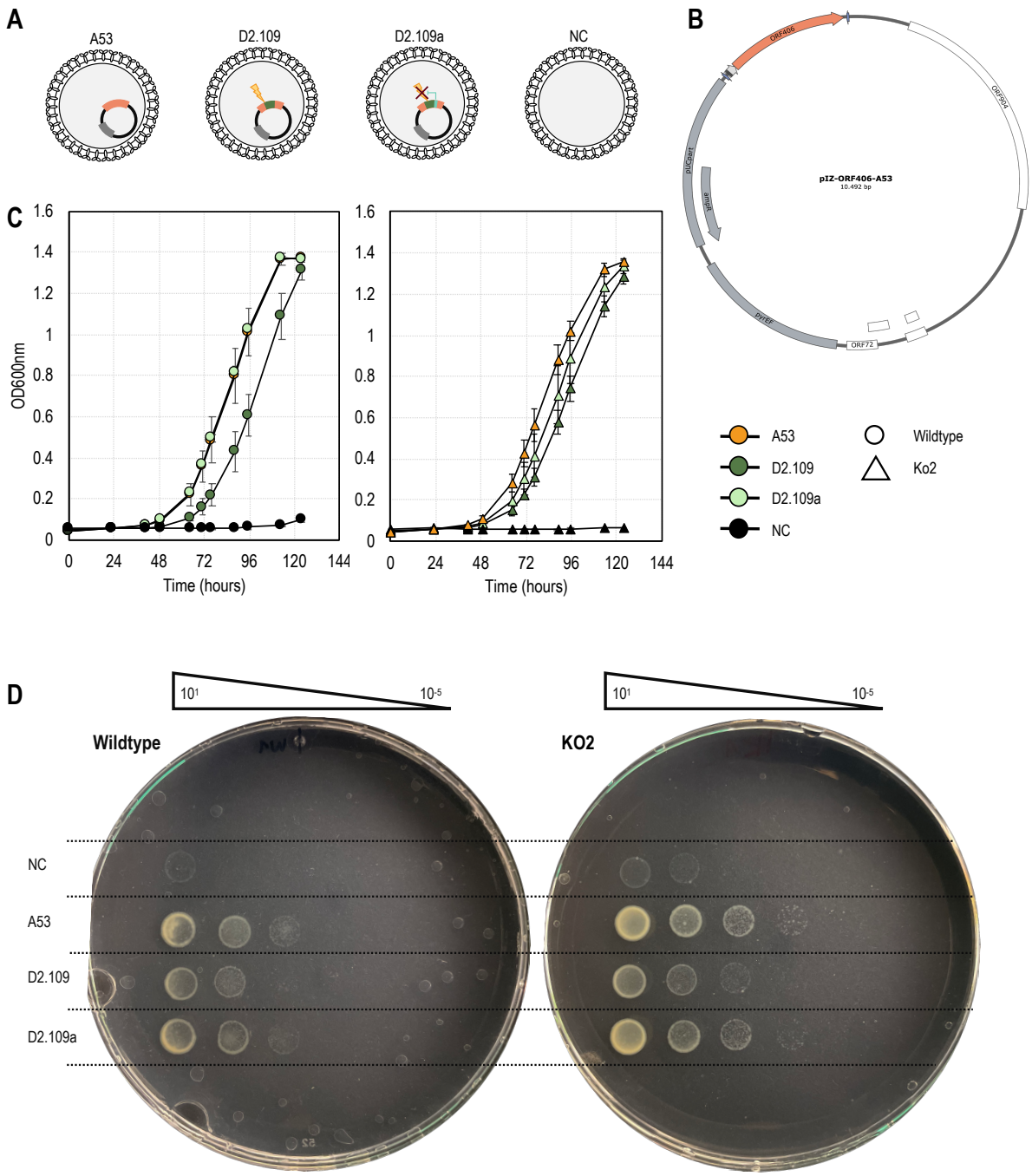
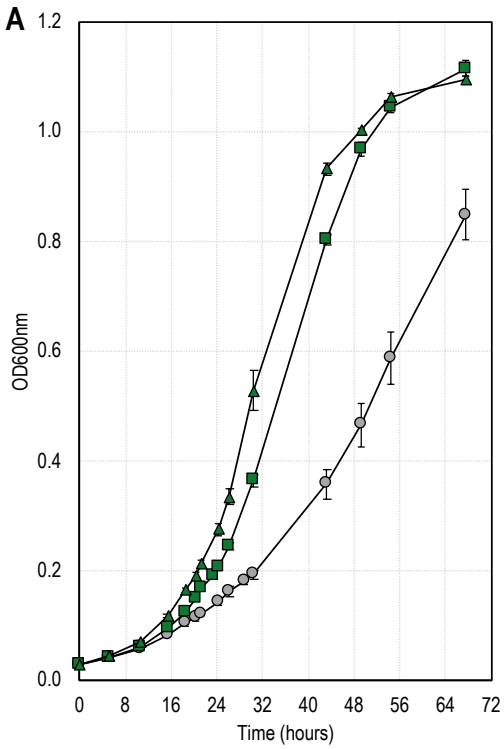
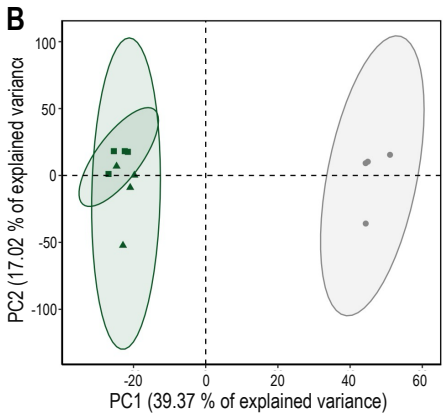
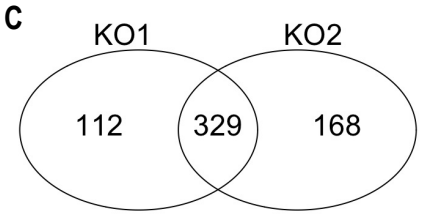
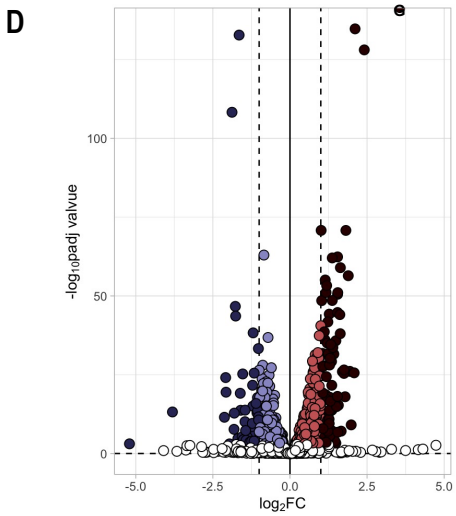


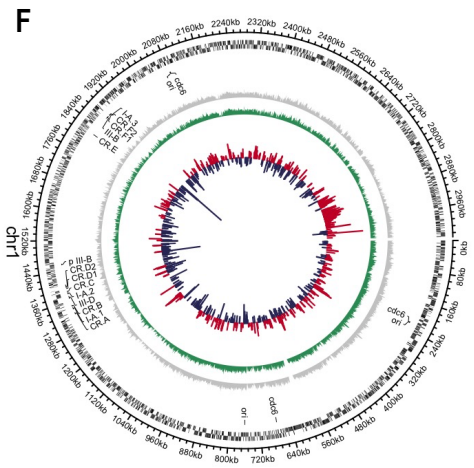
Figure 4. Plasmid challenge experiment. (A) Wildtype and deletion mutant KO2 were transformed with plasmids harboring different protospacers: A53 is not recognized by the CRISPR immune system due the lack of a matching spacer; D2.109 matches an endogenous spacer in CRISPR locus D eliciting type III CRISPR-mediated immunity; D2.109a harbors the same protospacer as in D2.109 and a matching 3' antitag to inhibit abortive infection. Negative control (NC) cells did not receive a plasmid and therefore lack an intact *pyrEF* locus. **(B)** Plasmid map of pIZ-ORF406-A53. **(C)** Growth curves for control (left) and KO2 (right) were generated by measuring optical density at 600nm, samples are color coded as follows: NC - black; A53 - orange; D2.109 - dark green; D2.109a - light green. Error bars: mean ± SD (n = 3, i.e. three separate inoculations after electroporation). **(D)** After transformation, cells were recovered in selective media before performing 10-fold serial dilutions (10^1 to 10^{-5}) and spotted on plates. One representative plate per strain is shown, the remaining plates are provided in the supplementary material (Suppl. Fig. S2 and S3).

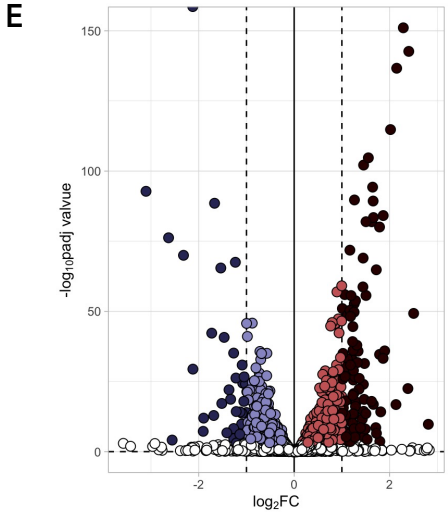

A
OD600nm
Time (hours)


B
PC2 (17.02 % of explained variance)
PC1 (39.37 % of explained variance)


C
KO1
KO2
112
329
168


D
-log10padj valvue
log2FC


F
chr1


E
-log10padj valvue
log2FC

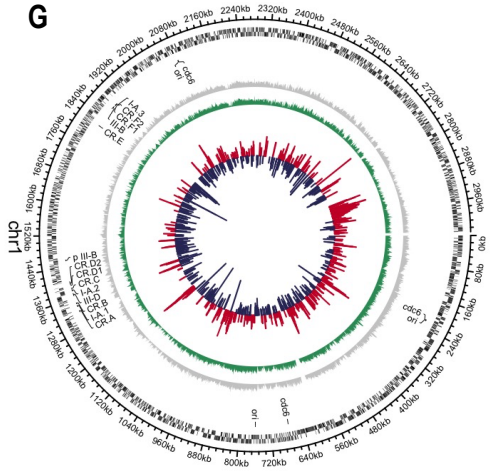

G
chr1

Figure 5. Knockout of the *cmr-ß* genes leads to growth accelaration and significant changes on the transcriptomic level. **(A)** Initial growth curve following isolation of two separate *cmr-ß* knockout mutants in comparison to wildtype undergoing the same procedure for the isolation of a single colony. Growth was followed by measuring optical density at 600nm. Color code: wildtype control – grey; deletion mutant 1 (KO1) – dark green; deletion mutant 2 (KO2) – light green. Error bars, mean ± SD (n = 4, i.e. four separate inoculations). **(B)** Principal component analysis of log transformed transcriptomic data from early exponential growth phase. Ellipses indicate 95% confidence interval. Same color code as in growth curve is used. **(C)** Venn diagram showing overlap between KO1 and KO2 regarding significantly upregulated genes in comparison to wildtype. **(D)** and **(E)** Volcano plots displaying log2-fold changes (x-axis) and associated significance (-log10 adjusted p value, y-axis), comparing KO1 and KO2, respectively. Color code indicates genes that are strongly upregulated (log2FC > 1, padj ≤ 0.001, dark red), upregulated (log2FC < 1, padj ≤ 0.001, red), strongly downregulated (log2FC < -1, padj ≤ 0.001, dark blue), downregulated (log2FC > -1, padj ≤ 0.001, blue). **(F)** and **(G)** Circular genome-wide plot for KO1 and KO2, respectively, showing *Sa. solfataricus* genome with position 0 to the right. Position of origins of replication (ori) and corresponding *cdc6* genes, as well as CRISPR-Cas systems are indicated. From outside inwards, rings indicate annotated coding sequences (black), log2-transformed TPM values for control (grey) and knockout mutant (green). Log2-fold changes in expression levels are displayed for significantly upregulated (padj ≤ 0.001, red) and downregulated (padj ≤ 0.001, blue) genes.

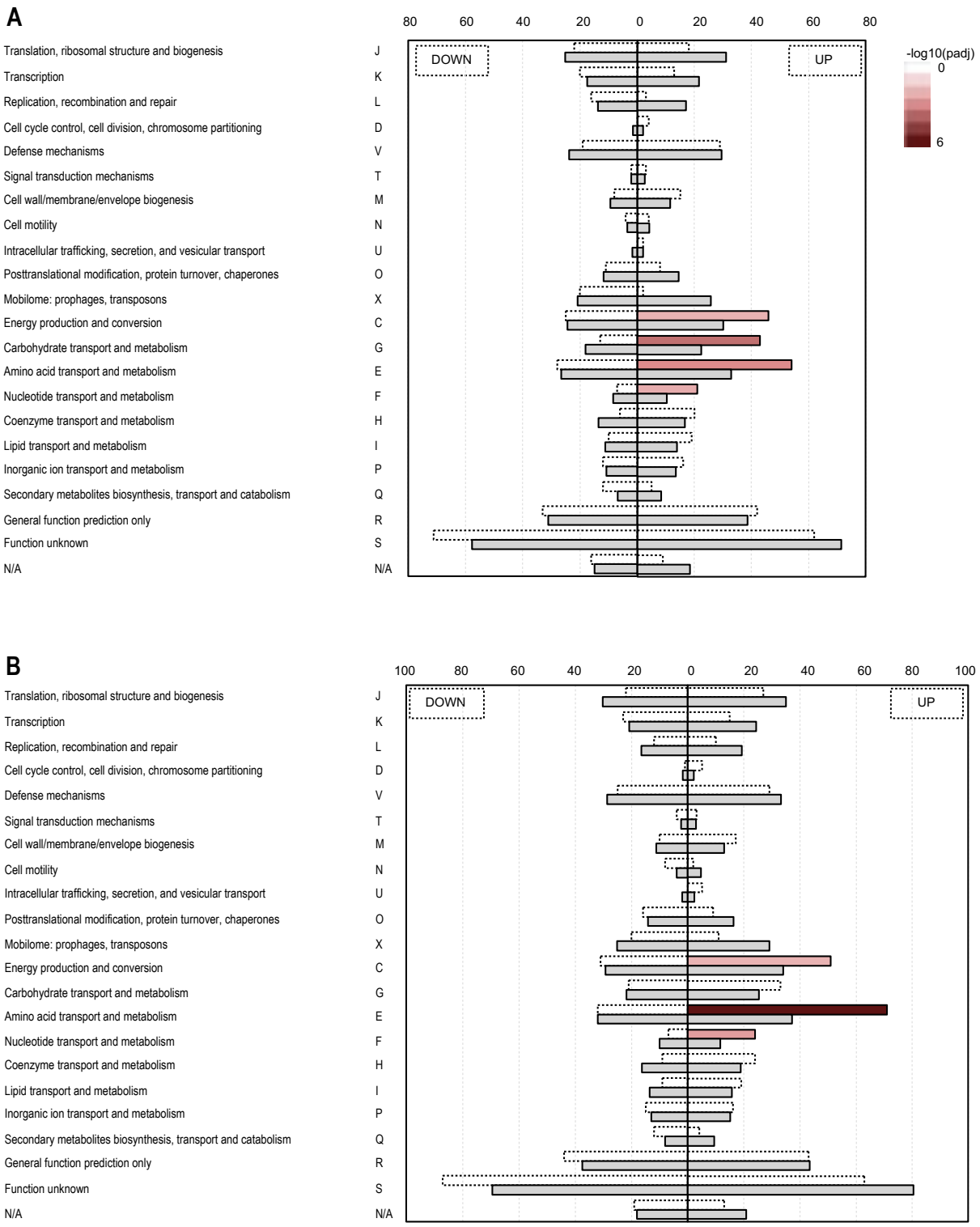
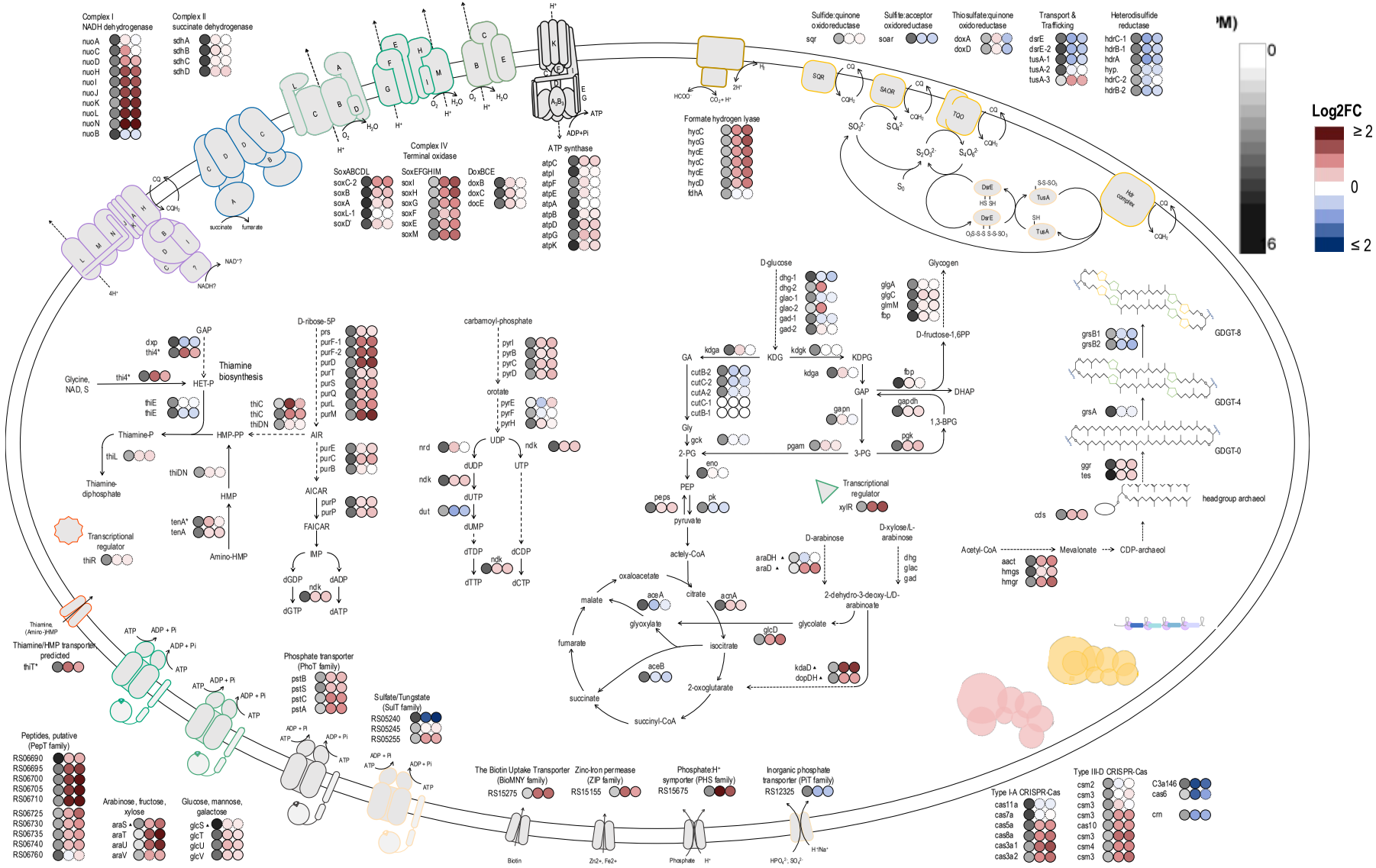


Figure 6. Functional enrichment analysis based archaeal clusters of orthologous genes (arCOGs) for **(A)** KO1 and **(B)** KO2. Left and right side of the plot show results for significantly downregulated and upregulated genes, respectively. Grey bar size represents the number of genes expected for each category based on a hypergeometric test, whereas neighboring bars represent the found number of genes for each category. Categories with padj-value < 0.05 are considered significantly overrepresented and highlighted by solid lines around the bar. Significance levels (-log10 transformed adjusted p-value) are indicated by color scale from white to dark red.



Complex I
NADH dehydrogenase
nuoA
nuoC
nuoD
nuoH
nuoI
nuoJ
nuoK
nuoL
nuoN
nuoB
Complex II
succinate dehydrogenase
sdhA
sdhB
sdhC
sdhD
Complex IV
Terminal oxidase
SoxABCDL
soxC-2
soxB
soxA
soxL-1
soxD'
SoxEFGHIM
soxI
soxH
soxG
soxF
soxE
soxM
DoxBCE
doxB
doxC
docE
ATP synthase
atpC
atpI
atpF
atpE
atpA
atpB
atpD
atpG
atpK
Formate hydrogen lyase
hycC
hycG
hycE
hycC
hycE
hycD
fdhA
Sulfide:quinone oxidoreductase
sqr
Sulfite:acceptor oxidoreductase
soar
Thiosulfate:quinone oxidoreductase
doxA
doxD
Transport & Trafficking
dsrE
dsrE-2
tusA-1
tusA-2
tusA-3
Heterodisulfide reductase
hdrC-1
hdrB-1
hdrA
hyp.
hdrC-2
hdrB-2
Log2FC
≥ 2
0
≤ 2
D-glucose
dhg-1
dhg-2
glac-1
glac-2
gad-1
gad-2
Glycogen
glgA
glgC
glmM
fbp
D-fructose-1,6PP
GAP
dxp
thi4*
Thiamine biosynthesis
Glycine, NAD, S
HET-P
thiE
thiE
Thiamine-P
HMP-PP
thiL
Thiamine-diphosphate
thiDN
HMP
tenA*
tenA
Amino-HMP
Transcriptional regulator
thiR
D-ribose-5P
prs
purF-1
purF-2
purD
purT
purS
purQ
purL
purM
thiC
thiC
thiDN
AIR
purE
purC
purB
AICAR
purP
purP
FAICAR
IMP
dGDP
dADP
dGTP
dATP
carbamoyl-phosphate
pyrI
pyrB
pyrC
pyrD
orotate
pyrE
pyrF
pyrH
UDP
nrd
ndk
dUDP
UTP
ndk
dUTP
dut
dUMP
dTDP
dCDP
dTTP
dCTP
GA
kdga
KDG
kdgk
KDPG
cutB-2
cutC-2
cutA-2
cutC-1
cutB-1
Gly
gck
2-PG
eno
PEP
peps
pk
pyruvate
actely-CoA
oxaloacetate
citrate
acnA
aceA
malate
glyoxylate
isocitrate
glcD
fumarate
aceB
2-oxoglutarate
succinate
succinyl-CoA
D-arabinose
araDH
araD
D-xylose/L-arabinose
dhg
glac
gad
2-dehydro-3-deoxy-L/D-arabinoate
glycolate
kdaD
dopDH
Transcriptional regulator
xylR
GAP
gapn
pgam
3-PG
pgk
1,3-BPG
gapdh
DHAP
fbp
GDGT-8
grsB1
grsB2
GDGT-4
grsA
GDGT-0
ggr
tes
cds
headgroup archaeol
Acetyl-CoA
Mevalonate
CDP-archaeol
aact
hmgs
hmgr
Thiamine (Amino-)HMP
Thiamine/HMP transporter predicted
thiT*
Phosphate transporter (PhoT family)
pstB
pstS
pstC
pstA
Sulfate/Tungstate (SulT family)
RS05240
RS05245
RS05255
Peptides, putative (PepT family)
RS06690
RS06695
RS06700
RS06705
RS06710
RS06725
RS06730
RS06735
RS06740
RS06760
Arabinose, fructose, xylose
araS
araT
araU
araV
Glucose, mannose, galactose
glcS
glcT
glcU
glcV
The Biotin Uptake Transporter (BioMNY family)
RS15275
Biotin
Zinc-Iron permease (ZIP family)
RS15155
Zn2+, Fe2+
Phosphate:H+ symporter (PHS family)
RS15675
Phosphate
H+
Inorganic phosphate transporter (PiT family)
RS12325
Type I-A CRISPR-Cas
cas11a
cas7a
cas5a
cas8a
cas3a1
cas3a2
Type III-D CRISPR-Cas
csm2
csm3
csm3
csm3
cas10
csm3
csm4
csm3
C3a146
cas6
crn

Figure 7. Response in metabolism to deletion of type III-B CRISPR-Cas system. First circle from left to right represents log2-transformed TPM values as an approximation for expression levels in wildtype control, indicated by color bar from white (not expressed) to black (highly expressed). Log2FC in transcript levels found in KO1 and KO2 are represented in that order by the two following circles, indicated by color scale from dark blue (strongly downregulated) to dark red (strongly upregulated). Significant change (padj ≤ 0.001) is indicated by solid outlines around circles. Genes with predicted binding sites for transcription factors ThiR or XylR are indicated by star or triangle symbols, respectively.

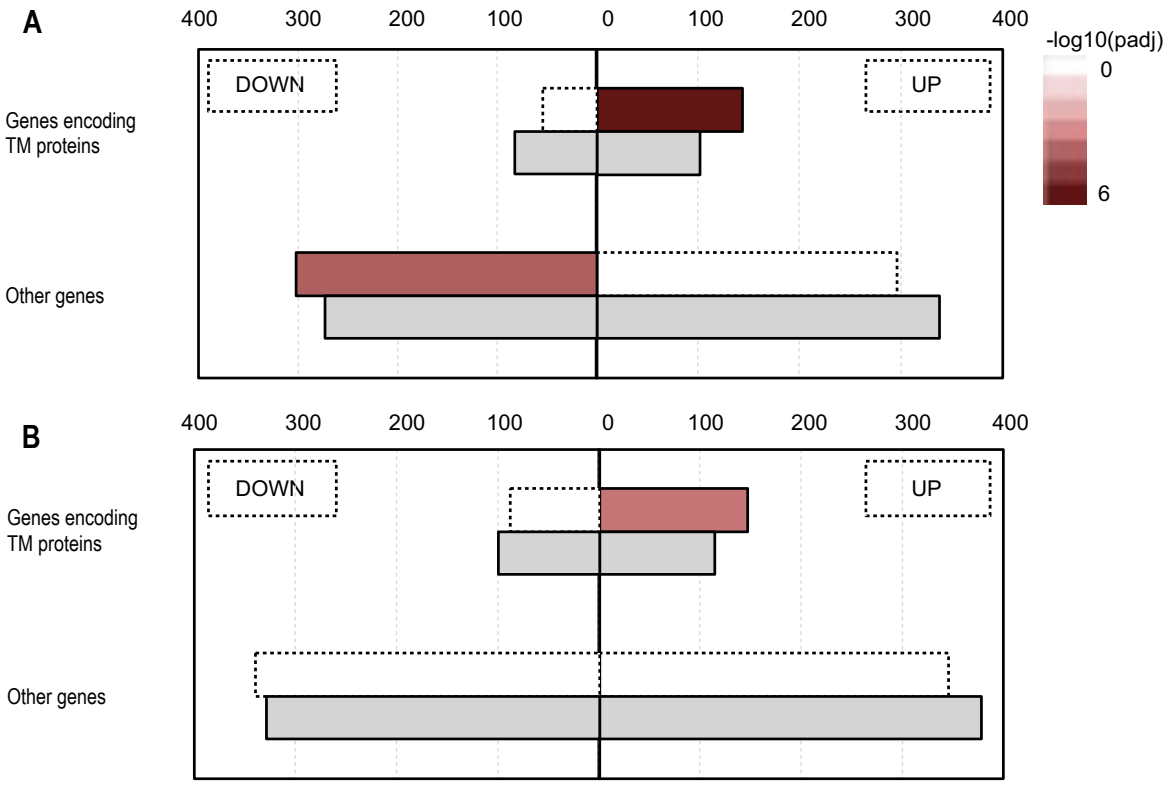


Figure 8. Enrichment analysis of genes encoding transmembrane-domain proteins as predicted for protein products by phobius. for **(A)** KO1 and **(B)** KO2. Left and right side of the plot show results for significantly downregulated and upregulated genes, respectively. Grey bar size represents the number of genes expected for each category based on a hypergeometric test, whereas neighboring bars represent the found number of genes. Categories with padj-value < 0.05 are considered significantly overrepresented and highlighted by solid lines around the bar. Significance levels (-log10 transformed adjusted p-value) are indicated by color bar from white to dark red.

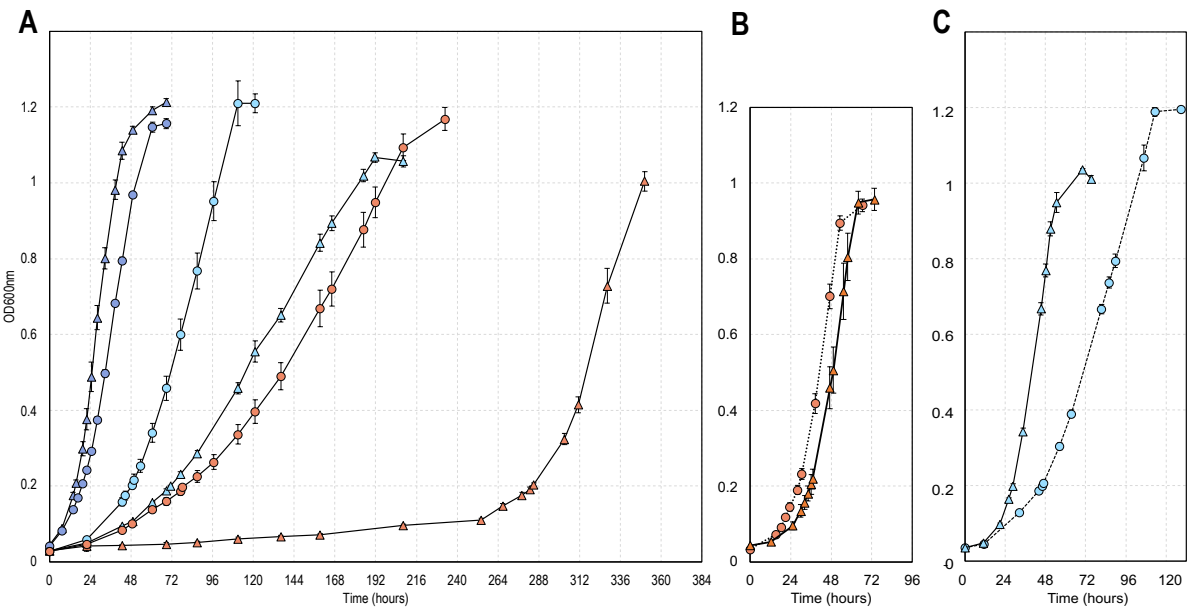


Figure 9. Adaptation Experiment Growth Curves. (A) Switching wildtype control and deletion mutant from complex to defined carbon sources sucrose and glucose causes lag phase during adaptation. Growth was followed by measuring optical density at 600nm. Wildtype control and Ko2 samples are represented by circles and triangles, respectively. Samples are color coded according to media composition as follows: Sucrose-tryptone – blue; sucrose – light blue; glucose – orange. Error bars, mean ± SD (n = 4, i.e. four separate inoculations). **(B)** and **(C)** show growth of KO2 and wildtype in glucose and sucrose media, respectively, after adaptation.

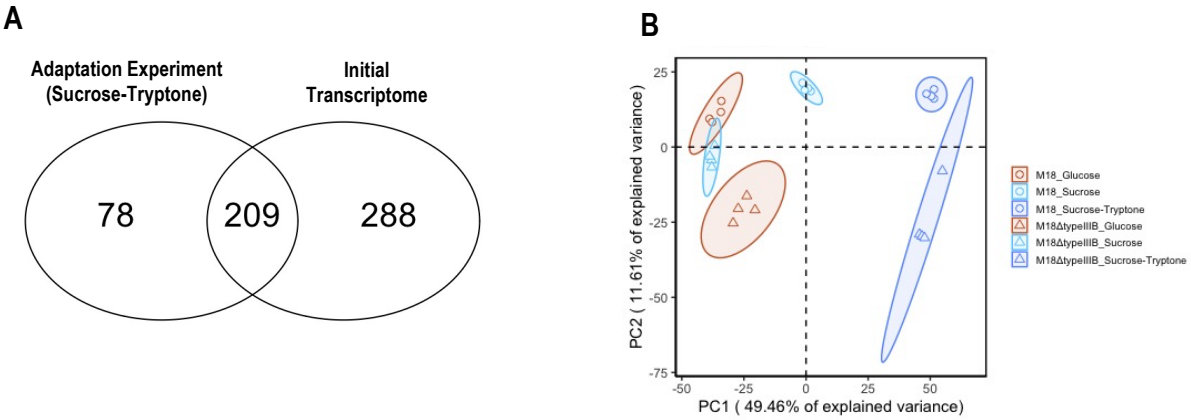


Figure 10. Transcriptomic Analysis of Cultures Adapted to Glucose and Sucrose Medium. (A) Venn diagram showing overlap in significantly upregulated genes between initial transcriptomic dataset and sucrose-tryptone condition of adaptation experiment. **(B)** Principal component analysis of log transformed transcriptomic data from early exponential growth phase. Ellipses indicate 95% confidence interval. Wildtype control and KO2 samples are represented by circles and triangles, respectively. Samples are color coded according to media composition as follows: Sucrose-tryptone – blue; sucrose – light blue; glucose – orange.