1 **TITLE**

2 **Changes in the genetic requirements for microbial interactions with**

3 **increasing community complexity**

4 **Manon Morin[123], Emily C. Pierce[1], Rachel Dutton[1]**

5 **[1]** University of California San Diego, 95000 Gilman Dr, La Jolla, 92093, USA

6 **[2]** LISBP, Université de Toulouse, CNRS, INRA, INSA, Toulouse, France

7 **[3]** Inria, Univ. Grenoble-Alpes, Grenoble, France

8

9 **ABSTRACT**

10 Microbial community structure and function rely on complex interactions whose underlying

11 molecular mechanisms are poorly understood. To investigate these interactions in a simple

12 microbiome, we introduced *E. coli* into an experimental community based on a cheese rind and

13 identified the differences in *E. coli's* genetic requirements for growth in interactive and non-

14 interactive contexts using Random Barcode Transposon Sequencing (RB-TnSeq) and RNASeq.

15 *E. coli's* genetic requirements varied among pairwise growth conditions and between pairwise

16 and community conditions. Our analysis points to mechanisms by which growth conditions

17 change as a result of increasing community complexity and suggests that growth within a

18 community relies on a combination of pairwise and higher order interactions. Our work provides

19 a framework for using the model organism *E. coli* as a probe to investigate microbial interactions

20 regardless of the genetic tractability of members of the studied ecosystem.

21

22 **INTRODUCTION**

23 Microorganisms rarely grow as single isolated species but rather as part of diverse

24 microbial communities. In these communities, bacteria, archaea, protists, viruses and fungi can

25 coexist and perform complex functions impacting biogeochemical cycles and human health

26 (Falkowski, Fenchel, and Delong 2008; Flint et al. 2012). Deciphering microbial growth

27 principles within a community is challenging due to the intricate interactions between

28 microorganisms, and between microorganisms and their environment. While interest in

29 microbial communities has dramatically increased, our understanding of microbial interactions

30 within communities is lagging significantly behind our ability to describe the composition of a

31 given community. Further, the way in which these interactions are organized within a

32 community, such as whether they consist of predominantly pairwise or higher-order interactions,

33 is even less clear. A more precise understanding of microbial interactions, their underlying

1

34 mechanisms, and how interactions are structured within a community, are all necessary to
35 elucidate the principles by which a community is shaped. In this study, we combine genome-
36 scale genetic and transcriptomic approaches within an experimentally tractable model microbial
37 community to begin to address these questions.

38 Genome-scale approaches, such as transposon-mutagenesis coupled to next-
39 generation sequencing (TnSeq approaches) have been successfully used to quantify the
40 contribution and thus the importance of individual genes to a given phenotype (van Opijnen and
41 Camilli 2013). Recently, generation and introduction of unique random barcodes into transposon
42 mutant libraries made this approach more high-throughput and less laborious, enabling screens
43 of important genes within hundreds of conditions and for numerous genetically tractable
44 microorganisms (Wetmore et al. 2015; Price et al. 2016). In order to investigate the genetic
45 bases of microbial interactions, we have adapted this approach to allow us to identify and
46 compare genetic requirements in non-interactive and interactive conditions. To do so, we
47 introduced a large and diverse transposon library generated previously in the genetically-
48 tractable model bacterium *E. coli* to characterize the genetic requirements of interactions within
49 a model community based on the rind of cheese (such as Camembert) (Wetmore et al. 2015).
50 We (i) identify the set of important genes for *E. coli* when growing alone in the cheese
51 environment (ii) identify the set of important genes for *E. coli* growth in pairwise conditions with
52 each individual community member and (iii) identify the set of important genes for *E. coli* growth
53 with the complete community. Characterization of the functions or pathways associated with
54 growth in interactive versus non-interactive conditions can then be used to infer the biological
55 process involved in interactions. Additionally, we compared the set of important genes in
56 pairwise conditions with the ones important for growth in a community to investigate how
57 microbial interactions change depending on the complexity of the interactive context.

58 We further performed two complementary approaches to defining interactions within this
59 system. First, we measured changes in the transcriptional profile of *E. coli* during growth alone,
60 growth in pairwise conditions, and within the community, using RNAseq. This analysis revealed
61 a deep reorganization of gene expression whenever *E. coli* is in the presence of other species.
62 Next, as *E. coli* is a non-endogenous species in our model ecosystem, we generated a
63 transposon library in the endogenous species *Pseudomonas psychrophila* JB418, and
64 performed similar RB-TnSeq analysis during non-interactive and interactive conditions.

65 This work revealed numerous interactions between species, such as metabolic
66 competition for iron and nitrogen, as well as cross-feeding from fungal partners for certain amino
67 acids. It also highlighted the need for resistance to toxic compounds and osmotic stress as a

68 requirement for growth. Our analysis showed that most of the metabolic interactions
69 (competition and cross-feeding) observed in pairwise conditions are maintained and amplified
70 by addition of partners and similar pairwise interactions. However, around half of the genetic
71 requirements observed in pairwise conditions were no longer apparent in the community,
72 suggesting that higher order interactions emerge in a community.

73
74 **RESULTS**

75

76 **Identification of the basic genetic requirements for growth of the *E. coli* sensor species**
77 **in isolation**

78 We used the *E. coli* Keio_ML9 RB-TnSeq library from Wetmore et al., 2015, containing
79 152,018 different insertion mutants (covering 3728 of 4146 protein-coding genes), each
80 associated with a unique 20 nucleotide barcode. This library was originally generated in and
81 maintained on lysogeny broth medium (LB), and was used previously to identify genes required
82 for growth across a variety of conditions (Wetmore et al. 2015; Price et al. 2016). To determine
83 genes important for growth on our cheese-based medium, we grew the library by itself on sterile
84 cheese curd agar plates (CCA: 10% freeze-dried fresh cheese, 3% NaCl, 0.5% xanthan gum,
85 1.7% agar) (Figure 1A), the same medium used in all further experiments, and used previously
86 to demonstrate that cheese communities could be successfully reconstructed *in vitro* (Wolfe et
87 al. 2014). As genetic requirements are likely to change over the course of growth, we grew the
88 library on CCA for 1, 2 and 3 days. For each time point, we harvested the library from the
89 surface of the cheese plate, extracted genomic DNA, used PCR to amplify the barcoded regions
90 of the transposons, and then sequenced these products to measure the abundance (i.e the
91 number of sequencing reads associated with each barcode) of each transposon mutant over
92 time (see Materials and Methods). For each gene, there are on average 15 individual insertion
93 mutants present in the library. The fitness of each insertion mutant was calculated as the log2
94 ratio of its abundance at a given timepoint compared to T0. Then, we calculated the weighted
95 average of the fitness of all insertions in a given gene to determine the overall fitness associated
96 with each gene and calculate a corresponding t-score that accounts for the fitness consistency
97 within individual insertions of that gene and thus for the confidence we can attribute to that gene
98 fitness. All gene fitness values were normalized so that a gene with no effect on the growth
99 phenotype has a fitness value of 0 (see (Wetmore et al. 2015) for details).

100 At each timepoint, the fitness was calculated for 3298 protein-coding genes. First, we
101 removed genes in which the fitness value did not pass our confidence threshold (absolute t-

3

102   score >= 3). We then removed genes with positive fitness (day 1 n= 10, day 2 n=14 and day 3
103   n=16). Thus, we only retained the genes whose deletion leads to a consistent growth defect for
104   *E. coli* on CCA (hereafter referred as to essential genes). This filtering process left 160 genes
105   that were important for *E. coli* growth alone on CCA with a false discovery rate of 0.2% (n=86
106   for day 1, n=87 for day 2, n= 91 for day 3) (Figure 1B and Supplementary file 1).

107   To identify the functions that are associated with these 160 genes, we mapped them to
108   the KEGG Brite database (Figure 1C). 84 genes were assigned to KEGG modules and 64 of
109   them were associated with *E. coli* metabolism. Within these metabolic genes, we found 28
110   genes associated with amino acid metabolism, specifically the biosynthesis of all amino acids
111   except for proline, lysine and histidine. Quantification of free amino acids in our medium
112   highlighted very low concentrations of all amino acids (supplementary figure 1) suggesting that
113   a limited supply of free amino acids leads to a genetic requirement for amino acid biosynthesis.
114   This is supported by the observation that both *spoT* and *relA*, regulators of the stringent
115   response which can be triggered by amino acid starvation (Cashel and Rudd 1996), are also
116   essential. 19 essential genes were associated with energy metabolism and more specifically
117   with sulfur assimilation (n=7 genes) and respiration (n=8 genes). Here, we deduce that
118   essentiality of sulfur assimilation is directly correlated with the lack of the amino acids cysteine
119   and methionine which are the major pools of sulfur-containing compounds in the cell. As a non-
120   endogenous species, *E. coli* might not possess the adequate peptidases or proteases to
121   degrade and use the highly available protein casein. Identification of two of the three genes of
122   the Leloir pathway (*galE* and *galT*), involved in the uptake and conversion of galactose into
123   glucose, suggests that galactose might be a crucial nutrient for *E. coli* growth on CCA. Finally, 8
124   genes mapped to membrane transport and were associated with two specific pathways: ferric-
125   enterobactin transport and glycine-betaine transport. Ferric-enterobactin transport allows the
126   cells to scavenge iron in a low-iron environment (Raymond, Dertz, and Kim 2003; Hider and
127   Kong 2010). Iron is an essential micronutrient and cheese is known to be iron-limited (Albar et al.
128   2014). Glycine betaine is used by the cells as an osmoprotectant against high osmolarity
129   environments. During cheese curd processing, high concentrations of NaCl are added (Guinee
130   2004), and our CCA medium contains 3% NaCl to mimic these conditions. The importance for *E.*
131   *coli* to maintain its cell osmolarity is also suggested by the essentiality of genes responsible for
132   the transport of the ions sodium, potassium or zinc.

133   In summary, conditionally essential functions for *E. coli* to grow alone in our
134   experimental environment involved (i) response to low iron availability, (ii) response to osmotic
135   stress and (iii) response to low available nutrients (specifically free amino acids). These required

4

136 functions are consistent with recently published results about the requirements of the mammary

137 pathogenic *E. coli* (MPEC) during growth in milk (Olson et al. 2017) except for resistance to

138 osmotic stress which does not occur in milk.
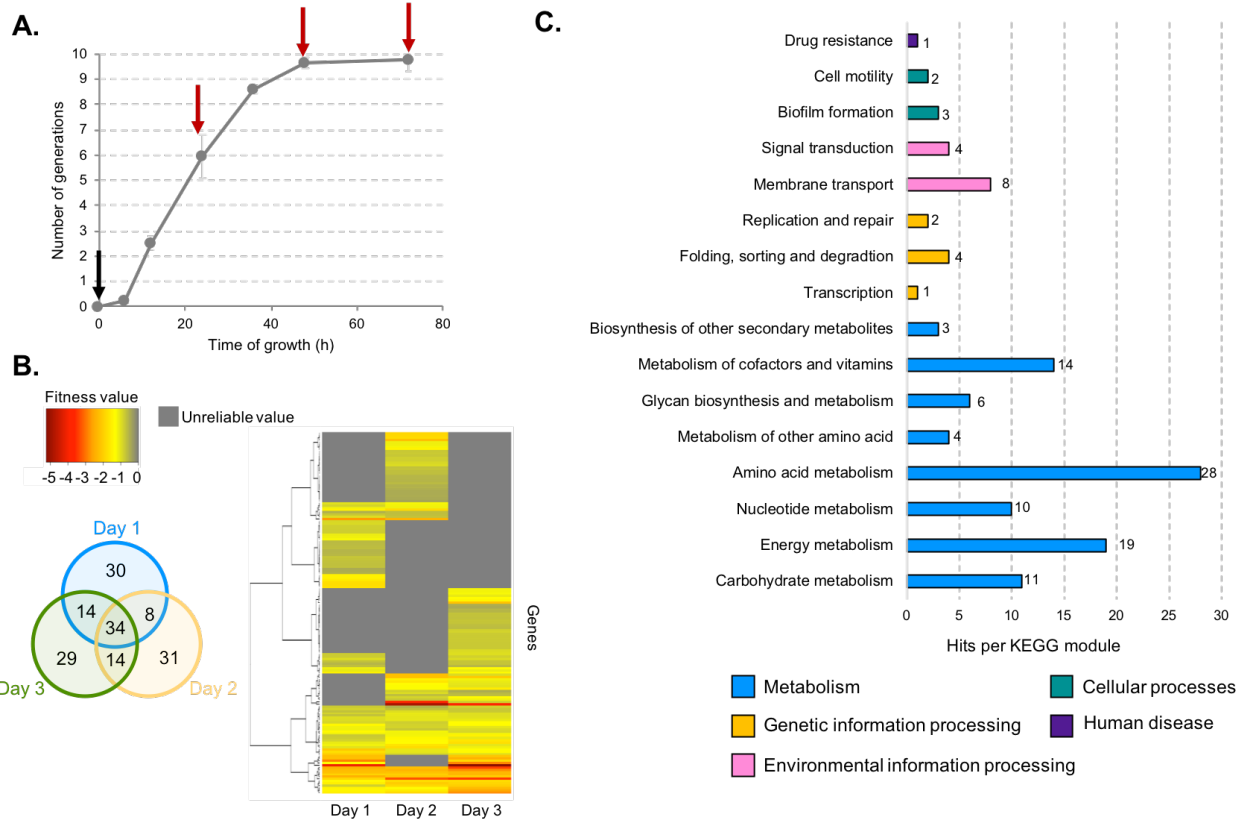
139

140

141

142 **Figure 1: Identification of genes essential for *E. coli* growth alone on cheese curd agar.** The E. coli RB-TnSeq
143 library Keio_ML9 (Wetmore et al., 2015) was grown for 3 days on cheese curd agar (CCA). Samples were harvested
144 at T=0h (black arrow), 24h, 48h and 72h (red arrows) for gene fitness determination **(A)**. Genes with a significant
145 negative fitness effect ("essential genes", abs(t-score) >= 3 and fitness < 0) were identified at day 1 (24h), day 2 (48h)
146 and day 3 (72h) **(B)**. Pooled together, they represented 160 essential genes for E. coli's growth on CCA. 84 of the
147 160 genes had hits when mapped to the KEGG BRITE database **(C)**.

148 To validate the results obtained with the transposon library, we measured the fitness of

149 individual knockout mutants from the *E. coli* Keio collection (Baba et al. 2006). We tested 25

150 knockout mutants corresponding to genes with a strong growth defect observed after one day of

151 growth. We carried out competitive assays between each knock-out mutant and the wild-type

152 strain on CCA. We calculated each knock-out mutant fitness as the log fold change of its

153 abundance after one day of growth. A z-score was also calculated to assess the confidence of

154 that fitness. 21 of 25 knock-out mutants displayed a fitness value lower than 0 with at least 95%

155 confidence (supplementary figure 2). The remaining 6 mutant strains (*brnQ, cysK, cysQ, serA,*

5

156 *trxA* and *waaP*) were associated with high fitness value variability across replicate experiments

157 and thus they were associated with a lower z-score. Altogether, this supports the reliability and

158 validity of RB-TnSeq results.

159

160 **Identification of genes essential for *E. coli* growth in pairwise conditions**

161 The growth of our *E. coli* library alone allowed us to determine the baseline set of genes

162 required for optimal growth in the model cheese environment. We next wanted to examine the

163 genes required for growth when another species is present. First, we analyzed the growth of *E.*

164 *coli* and the partner species. We grew *E. coli* for 3 days on CCA in the presence of either *H.*

165 *alvei*, *G. candidum* or *P. camemberti*. In addition to belonging to distinct domains or phyla, these

166 3 partners are the typical members of a bloomy rind cheese community (such as Brie or

167 Camembert). The presence of *E. coli* did not influence the growth of any species

168 (supplementary figure 3). However, *E. coli's* growth was reduced in the presence of each

169 partner (Figure 2A).

170 We then determined essential genes for *E. coli* growth in each pairwise condition using

171 RB-TnSeq (i.e any genes whose fitness value is negative and associated with an absolute t-

172 score greater than 3 in the pairwise condition). As performed above, barcode frequencies were

173 compared between T0 and after growth with each partner (at day 1,2 and 3). For each pairwise

174 condition, we pooled genes with a consistent fitness for at least one timepoint as a single set of

175 essential genes. We identified 145 conditionally essential genes for *E. coli* growth with *H. alvei*,

176 131 conditionally essential genes for its growth with *G. candidum* and 142 conditionally

177 essential genes for its growth with *P. camemberti* (Figure 2B and supplementary file 2). We saw

178 significant overlap between these gene sets, which altogether constitute a set of 153 genes that

179 were conditionally essential for growth in at least one pairwise culture. These genes are further

180 referred to as genes essential for growth in pairwise conditions.

181 Comparison of the set of genes identified when *E. coli* is grown alone with the genes

182 identified when *E. coli* is grown in pairwise conditions is expected to highlight differences

183 brought about by the presence of another species. Consistent presence of multiple genes of the

184 same pathway within a set of essential genes is likely to point out a pathway specifically

185 essential in one condition. Thus, we can infer possible interactions based on the different

186 essential pathways between interactive and non-interactive growth conditions.

187 We compared the 153 genes essential in pairwise conditions to the 160 essential genes

188 for *E. coli* growth alone (Figure 2C). Three groups of genes arose from that comparison: (i) the

189 core requirements: genes essential for both growth alone and in pairwise conditions (n=78), (ii)

190   the alone-specific requirements: any gene essential for *E. coli* growth alone that was not
191   identified as essential in the presence of at least one of the partner (n=82), and (iii) pairwise-
192   specific requirements: any gene essential in the presence of at least one of the partners but not
193   essential for growth alone (n=75). We further focused on the alone-specific and pairwise-
194   specific requirements as these groups contain genes potentially related to interactions.

195   Alone-specific requirements can highlight genes that are essential for growth alone but
196   no longer required due to the presence of a partner, thus suggesting interactions between *E.*
197   *coli* and the partner. In this specific case, these genes can be described as requirements which
198   are relieved by the presence of a partner. We mapped the 82 alone-specific genes (Figure 2D
199   and supplementary file 3) to the KEGG BRITE database to identify functions and pathways that
200   are no longer essential in the presence of a partner. 16 genes were associated with unknown or
201   predicted proteins and did not map to any field of the database. Of the remaining genes, 45
202   mapped to modules of the KEGG orthology hierarchy. 7 genes were associated with genetic
203   information processing, environmental information processing, cellular processes and human
204   disease modules. However, most of the alone-specific genes were associated with the KEGG
205   module metabolism, and are thus part of metabolic pathways. It is especially evident that
206   pairwise growth leads to major changes in the need for amino acid biosynthesis. For example, 6
207   out of the 8 genes of valine and isoleucine biosynthetic pathways are no longer required during
208   pairwise growth (Figure 3D and supplementary file 3). In addition, 2 genes of arginine
209   biosynthesis, 2 genes of methionine biosynthesis as well as final steps of homoserine, aspartate
210   and glutamate biosynthesis are no longer required. Moreover, *ilvY*, the transcriptional activator
211   of valine and isoleucine biosynthesis was also among the genes no longer required for pairwise
212   growth. Here, the dominant presence of amino acid biosynthesis genes in the alone-specific
213   requirements suggests cross-feeding of the pathway end-product or key intermediaries which
214   are either provided directly by the partner species or made more available in the environment as
215   a consequence of the partner's metabolic activity. Thus, our data suggest that pairwise growth
216   may allow cross feeding of the amino acids valine, isoleucine, arginine, methionine, homoserine,
217   aspartate and glutamate. Isoleucine or methionine are also intermediaries of cofactor
218   biosynthesis, and genes associated with their biosynthesis were also mapped to metabolism of
219   cofactors and vitamins.

220   To understand if the alone-specific requirements were related to a specific partner, we
221   investigated how each partner contributed to this gene set (supplementary file 3). Of the 82 total
222   genes, 36 were identified as non-essential in all pairwise cultures. They included genes
223   associated with amino acid metabolism specific to homoserine and methionine biosynthesis. Of

7

224 the remaining genes, 8 were specifically not required in the presence of *H. alvei*, 9 were

225 specifically not required in the presence of *G. candidum* and 9 were specifically not required in

226 the presence of *P. camemberti*. Alleviation of leucine and valine biosynthesis was observed with

227 all fungal partners, while biosynthesis of arginine appeared to be no longer required specifically

228 in the presence of *G. candidum*. Fungal species are known to secrete proteases that digest

229 small peptides and proteins (Kastman et al. 2016; Boutrou, Aziza, and Amrane 2006; Boutrou,

230 Kerriou, and Gassi 2006) and may lead to increased availability of amino acids in the

231 environment.

232 We then analyzed the 74 pairwise-specific genes (Figure 2D and supplementary file 3) in

233 order to identify functions or pathways that are newly essential due to the presence of a partner

234 compared. 33 genes mapped to KEGG Orthology terms. Among this gene set are pathways

235 associated with signal transduction, biofilm formation, and drug resistance. They were related to

236 3 major responses: metabolic switch (*creB*: carbon source responsive response regulator),

237 response to stress and toxic compounds (*cpxA*: sensory histidine kinase, *oxyR*: oxidative stress

238 regulator, *acrAB*: multidrug efflux) and biofilm formation (*rcsC* and *rcsB*: regulator of capsular

239 synthesis, *pgaC*: poly-*N*-acetyl-D-glucosamine synthase subunit). Biofilms are microbial

240 structures known to provide resistance to different stresses, including resistance to antibiotics,

241 and biofilm-inducing genes can be activated in presence of stress events (Landini 2009). The

242 transcriptional regulator *oxyR* and the transduction system *CpxA* and *CpxB* are known

243 coordinators of stress response and biofilm formation (Gambino and Cappitelli 2016; Dorel,

244 Lejeune, and Rodrigue 2006). While these genes represent only a small subset of the all the

245 pairwise specific gene set, they could suggest that partner species are producing toxic

246 compounds or oxidative stress-inducing compounds.

247 We again investigated if these responses were partner-specific (supplementary file 3).

248 Of the 74 pairwise-specific genes, 11 were found to be essential in the presence of all partners,

249 13 were specific to the presence of *H. alvei*, 24 were specific to the presence of *G. candidum*

250 and 11 were specific to the presence of *P. camemberti*. Despite involving different genes,

251 necessity of biofilm formation and response to toxic stress were associated with the presence of
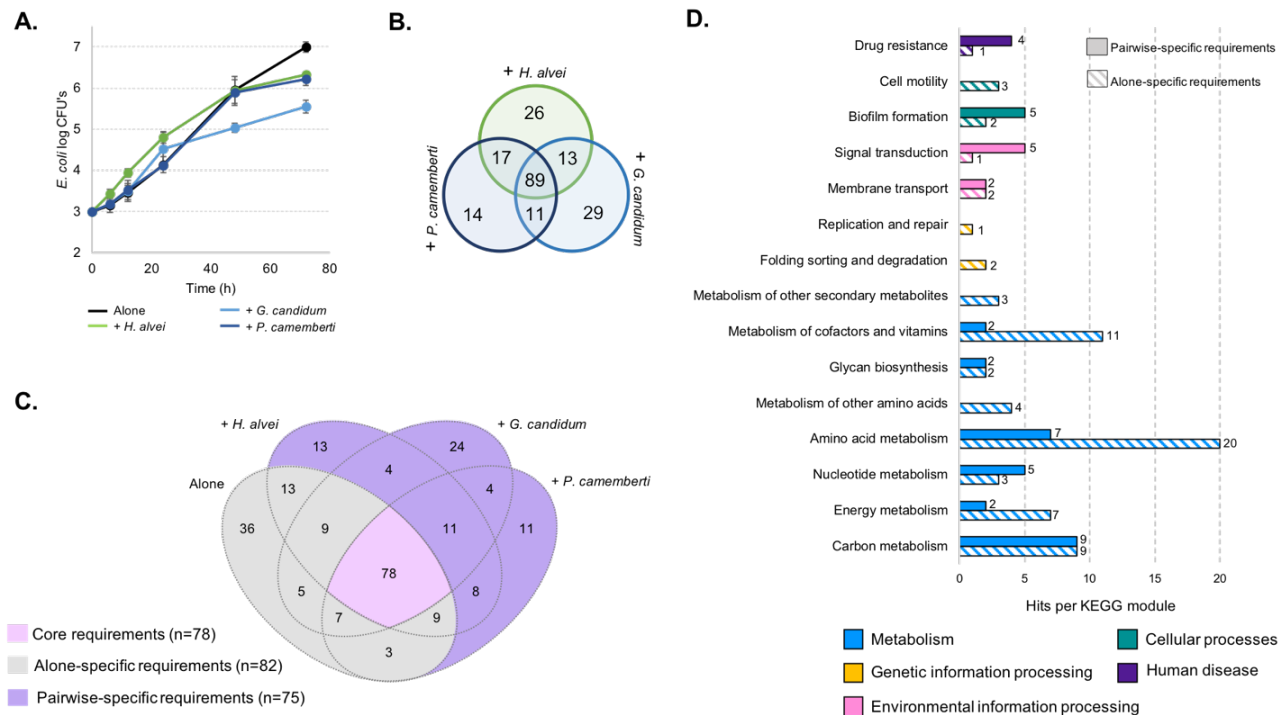
252 all partners.

253

Figure 2: Analysis of *E. coli* essential genes during growth in pairwise conditions. We grew *E. coli* in pairwise conditions with either with *H. alvei*, *G. candidum* or *P. camemberti* **(A)**. Using the *E. coli* RB-TnSeq library we identified the essential genes for growth on CCA in each pairwise condition. Requirements for growth with the different partners overlapped between the conditions. Altogether, it constitutes 153 essential genes **(B)**. Comparing these genes to the genes essential for *E. coli* growth alone, we identified 82 genetic requirements that were no longer required for growth in pairwise conditions as well as 75 genes that were added by pairwise growth **(C)**. We mapped the pairwise-specific and alone-specific genes to the KEGG BRITE database. 45 out of 82 genes and 33 out of 77 had hits **(D).** Most of the alone-specific requirements were mapped to metabolic pathways and especially amino-acid metabolism (n=20 genes) while pairwise-specific requirements mapped to metabolic functions but also to functions associated with response to stress (biofilm formation n=5, signal transduction n=5 and drug resistance n=4).

Identification of *E. coli* genes essential for growth within the community and comparison with genes essential for growth in pairwise conditions.

We next aimed to investigate the differences between genes essential for growth in a community (complex interactive condition) and genes essential for growth in associated pairwise conditions (simple interactive condition). We grew the *E. coli* library with the complete community composed of *H. alvei*, *G. candidum* and *P. camemberti* and identified 126 essential genes (supplementary file 2). *E. coli* final biomass was reduced by the presence of the community even more so than by a single partner. However, the growth of each community member was unaffected (Supplementary figure 3).

We first compared the genes essential for growth in the community with the genes essential for growth alone. We identified 89 genes that were essential for both conditions, 37 genes only essential with the community and 71 genes that were no longer essential to grow

278   with the community (essential only for growth alone) (Figure 3A). As for the presence of a single

279   partner, the presence of a complex community potentially introduces new essential genes for

280   growth while relieving some requirements.

281   By comparing the genes required for growth in interactive and non-interactive conditions,

282   we have identified the set of genes which are newly required during growth in a community

283   versus growth alone and the set of genes newly required for growth in pairwise conditions

284   versus growth alone. Comparing these two sets of additional genes can reveal if and how

285   community complexity modifies the genes that are essential in different interactives contexts

286   compared to growth alone. We identified 29 essential genes that were potentially newly

287   essential compared to growth alone in both pairwise conditions and community growth (Figure

288   3B and supplementary file 4). These new requirements are likely to be associated with pairwise

289   interactions which are maintained in a community context. These include genes associated with

290   oxidative stress and biofilm formation.

291   Meanwhile, we identified 46 genes that were essential only for growth in pairwise

292   conditions, but not with the community. These genes could be related to interactions that are

293   either alleviated or counteracted in a community, either by the presence of a specific species, or

294   the community as a whole. For example, some of the identified genes were associated with

295   antimicrobial resistance, and in a diverse community, other species could degrade the putative

296   antimicrobial molecule or prevent the producing species from secreting it. Consequently *E. coli*

297   would be exposed to a lesser level of antimicrobial, suppressing the necessity of a resistance

298   gene.

299   Finally, 8 essential genes appeared to be specifically required in the presence of the

300   community. These genes may be associated with specific interactions emerging from the

301   community context. 3 genes encode for uncharacterized proteins, and the others (*purK*, *purE*,

302   *damX*, *ftsX*, *secB*) are not associated with a single function or pathway. We therefore cannot

303   conclude whether specific interactions emerge from the community context or if these genes

304   appeared as specific because of the filtering process criteria.
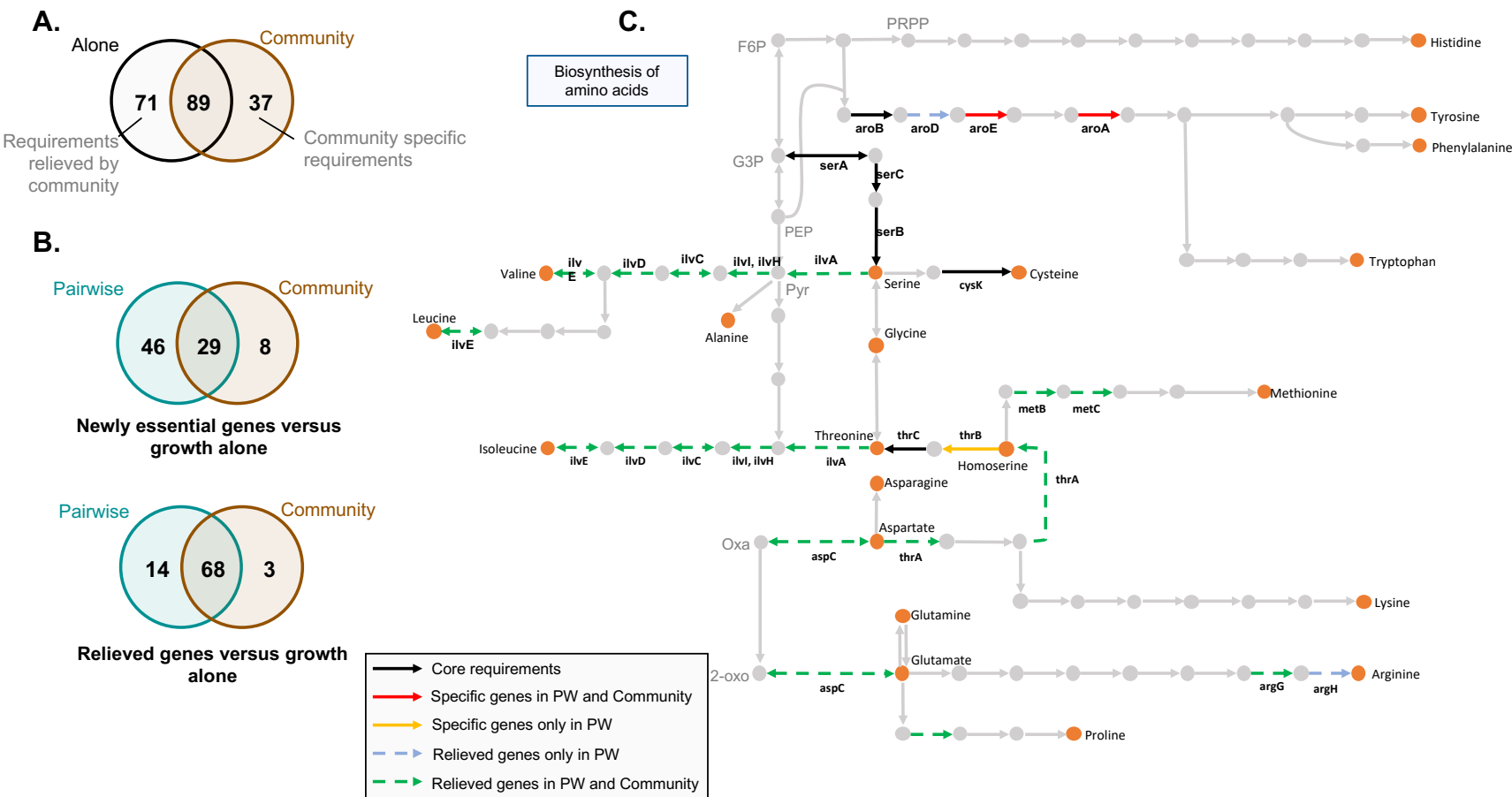
305   We next investigated if the interactions related to relieved requirements in community

306   growth versus alone and relieved requirements in pairwise conditions versus growth alone were

307   similar. Thus, we compared the alone specific requirements versus growth with the community

308   and the alone specific requirements versus growth in pairwise conditions (Figure 3B and

309   supplementary file 4). 68 genes were no longer required for both growth in pairwise conditions

310   and with the community compared to growth alone. These genes can also represent pairwise

311   interactions maintained in the community context. Amino acid biosynthesis was highly

10

312   represented within these genes and more specifically biosynthesis of valine, isoleucine,
313   methionine, homoserine, aspartate and glutamate (Figure 3C). This suggests that, despite the
314   presence of more species, these amino acids are still cross-fed.

315   We also identified 14 genes that were no longer required in pairwise conditions
316   compared to growth alone yet remained essential for growth with the community. Finally, 3
317   genes appeared to be specifically no longer required in the presence of the community. In both
318   cases, too few genes are involved for us to infer any hypothesis on the existence of specific
319   pairwise interactions not conserved in the community or the emergence of specific interactions
320   in the community.

321   To summarize, the genes that would be newly essential in the community compared to
322   growth alone were mostly maintained from newly essential genes identified in pairwise
323   conditions compared to growth alone. Similarly, the genes that were no longer essential for the
324   growth in the community were highly similar to the genes that were no longer essential in
325   pairwise condition. However, many newly essential genes for growth in pairwise condition
326   compared to growth alone were not found as essential for growth within the community.
327   Altogether, this highlights that part of the pairwise interactions are conserved within a
328   community while underlining the presence of higher order interactions related to the higher
329   complexity of the growth condition.

330

A.

Alone | Community

71 | 89 | 37

Requirements relieved by community

Community specific requirements

B.

Pairwise | Community

46 | 29 | 8

Newly essential genes versus growth alone

Pairwise | Community

14 | 68 | 3

Relieved genes versus growth alone

C. Biosynthesis of amino acids

F6P, PRPP → Histidine
aroB, aroD, aroE, aroA → Tyrosine, Phenylalanine
G3P, serA, serC, serB
PEP, Pyr
Valine, ilvE, ilvD, ilvC, ilvI, ilvH, ilvA, Serine, cysK → Cysteine
Tryptophan
Leucine, ilvE
Alanine, Glycine
metB, metC → Methionine
Threonine, thrC, thrB
Isoleucine, ilvE, ilvD, ilvC, ilvI, ilvH, ilvA, Homoserine, thrA
Asparagine
Aspartate, aspC, thrA
Oxa
Lysine
Glutamine
Glutamate, argG, argH → Arginine
2-oxo, aspC
Proline

Legend:
— Core requirements
— Specific genes in PW and Community
— Specific genes only in PW
-- Relieved genes only in PW
-- Relieved genes in PW and Community

331

Figure 3: Comparison of genes essential for *E. coli* growth within the community and in pairwise conditions. Using the *E. coli* RB-TnSeq library, we identified genes required to grow with the community (*H. alvei* + *G. candidum* + *P. camemberti*). Then, we compared the essential genes for growth alone and with the community **(A)** to identify the newly essential and relieved genes during growth with the community. Next, we compared the newly essential genes in community growth versus growth alone with the newly essential genes in pairwise growth (PW) versus growth alone. We did the same for relieved requirements **(B)**. Within the requirements that were relieved both in pairwise and community growth, genes associated with numerous amino acid biosynthetic pathways were identified **(C)**.

F6P: fructose-6-phosphate, PRPP: 5-phospho-ribose-1-di-phosphate, G3P: Glyceraldehyde-3-phosphate, PEP: phosphoenol-pyruvate, Pyr: Pyruvate, Oxa: Oxaloacetate, 2-oxo: 2-oxoglutarate

340 **E. coli *differential expression analysis in pairwise conditions versus growth alone***

341   So far, we used a genome-scale genetic approach to investigate potential microbial
342 interactions. As a complementary strategy, we generated transcriptomic data for *E. coli* during
343 growth in each previously discussed condition. Changes in transcriptional profiles can be a
344 powerful indicator of an organism's response to an environment and were previously used to
345 identify *E. coli* pathways involved in interactions (Croucher and Thomson 2010; McAdam,
346 Richardson, and Fitzgerald 2014; Galia et al. 2017).

347   To measure *E. coli* gene expression, we extracted and sequenced RNA from each
348 timepoint and condition of the samples used for RB-TNseq above (after 1, 2 and 3 days of
349 growth when grown alone, in pairwise conditions or with the community). Comparison of
350 transcriptional profiles suggests a strong reorganization of *E. coli* gene expression over time
351 and in response to the presence of a partner. Principal component analysis on the normalized
352 gene expression values (rlog transformation(Love et al. 2015)) revealed that samples clustered
353 based on the growth timepoint (separation on the horizontal axis) and whether *E. coli* is growing
354 alone or not (separation on the vertical axis) (Figure 4A). The different pairwise conditions and
355 the community conditions cluster together, underlining that the potential differences between
356 these conditions are less important than the differences between an alone and co-cultured
357 condition.

358   We next focused on the genes differentially expressed between pairwise and alone
359 growth and by calculating the fold change of gene expression between pairwise growth and
360 growth alone. We identified differentially expressed genes by screening for adjusted p-values
361 lower than 1% and absolute log of fold change (logFC) greater than 1. To remain consistent with
362 the analysis performed for the genetic requirements, we pooled the data at all time points after
363 identifying all of the up-regulated or down-regulated genes for each timepoint. We found a total
364 of 966 upregulated and 977 downregulated across all partners (482 upregulated genes and 478
365 downregulated genes in presence of *H. alvei*, 633 upregulated genes and 719 downregulated
366 genes in presence of *G. candidum*, 626 upregulated genes and 694 downregulated genes in
367 presence of *P. camemberti*, Figure 4B and supplementary file 5). 271 genes were up-regulated
368 in all pairwise conditions while 340 genes were always down-regulated. Meanwhile, a number of
369 genes were differentially expressed depending on the partner: 66 genes were specifically up-
370 regulated and 60 genes down-regulated with *H. alvei*, 213 up-regulated and 182 down-regulated
371 with *G. candidum*, 183 up-regulated and 161 down-regulated with *P. camemberti*.

372

373

374   Due to the larger gene set compared to RB-TNseq, we performed KEGG Pathway
375   enrichment analyses on the differentially expressed genes in pairwise conditions to determine
376   up-regulated functions and pathways depending on the partner species (Figure 4C). First,
377   certain functions are up-regulated regardless of the partner. This includes up-regulation of
378   almost all the aminoacyl-tRNA-synthetases and up-regulation of energy production. Interestingly
379   up-regulation of energy production through aerobic respiration and the TCA cycle happened
380   after 3 days of growth. Oxygen availability (Gunsalus 1992) and growth phase (Wackwitz et al.
381   1999) are the two known regulators of aerobic respiration. At day 3, *E. coli* was observed to be
382   in log phase when alone, whereas in the presence of a partner and especially with *P.*
383   *camemberti*, *E.coli* was observed to enter the stationary phase between day 2 and day 3
384   (Supplementary figure 5). Therefore, up-regulation of aerobic respiration is most likely
385   associated with the growth stage difference between *E. coli* alone and with a partner.

386   More genes were up-regulated in the presence of *G. candidum* than the other partners.
387   Several pathways associated with nucleotide biosynthesis (one-carbon pool by folate, purine
388   metabolism and pyrimidine metabolism) were specifically upregulated. This suggests that either
389   *E. coli* and *G. candidum* compete for nucleotide compounds from the environment or that
390   presence of *G. candidum* leads to an increase demand of nucleotide compounds for *E. coli's*
391   metabolism and growth.

392   We performed a similar KEGG pathway enrichment analysis on the down-regulated
393   genes in pairwise conditions. No pathway was enriched in the presence of *H. alvei.* Pathways
394   involved in the biosynthesis of amino acids, specifically tyrosine, phenylalanine, tryptophan,
395   methionine, lysine, arginine homoserine, leucine, glutamate, threonine and glycine, appeared to
396   be the principal down-regulated function in the presence of a fungal partner. Lysine biosynthesis
397   was specifically down-regulated with *G. candium* while arginine biosynthesis was specifically
398   downregulated with *P. camemberti*. Interestingly, some amino acid biosynthetic pathways were
399   up-regulated later in the growth (phenylalanine, tyrosine and leucine). Down-regulation of amino
400   acid biosynthesis suggests that the partner species generates amino acid cross feeding. Our
401   previous RB-TnSeq results previously suggested this particular interaction in the presence of
402   fungi. The observation of this interaction in the transcriptome data is consistent with our
403   interpretation of RB-TnSeq and reinforces the likelihood of such an interaction. However, late
404   up-regulation of some amino acid biosynthesis suggests that as the partner grows along with *E.*
405   *coli* they eventually end up competing for amino acids leading to biosynthesis up-regulation.
406

407   To summarize, presence of a single partner triggers a deep and dynamic reorganization

408   of *E. coli* gene expression. These modifications mostly rely on restructuring *E. coli* metabolic

409   activity whether it is to down-regulate specific pathways and activate new ones as a response to

410   potential metabolic competition and establishment of an appropriate metabolic strategy or to

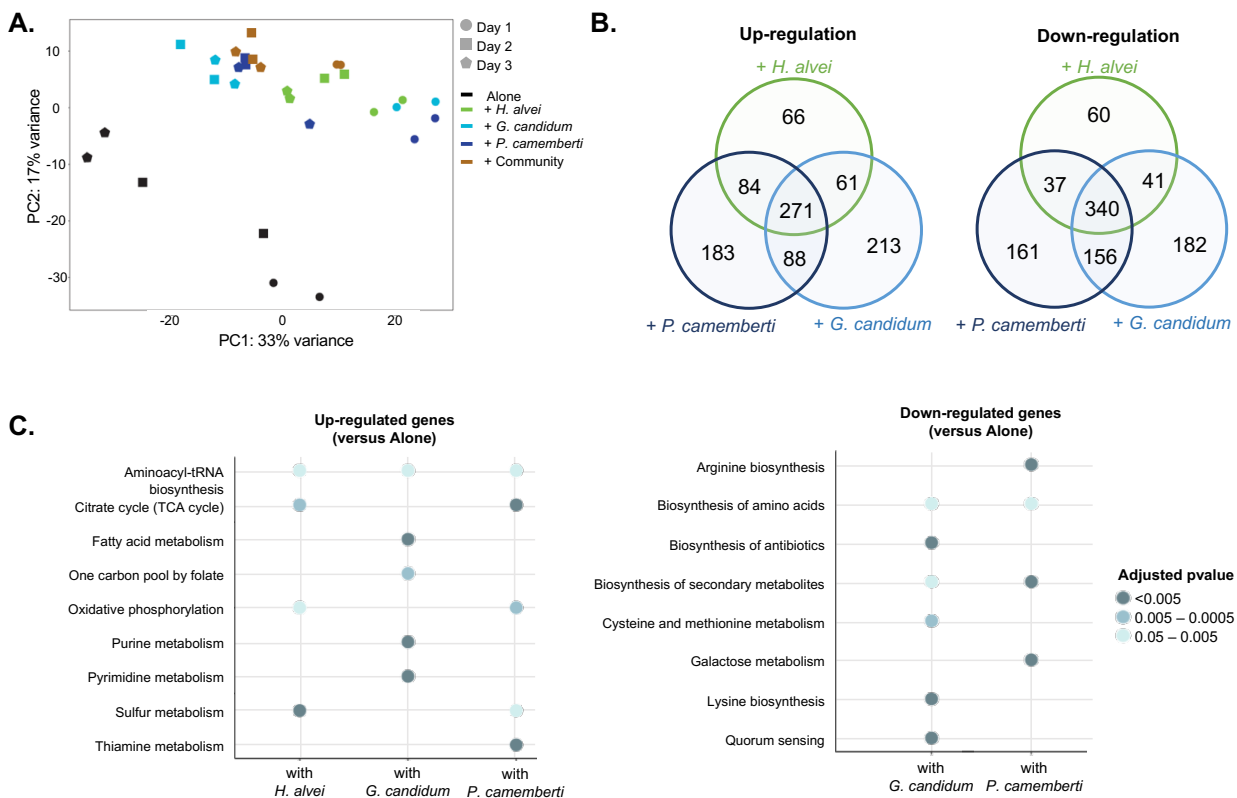411   down-regulate certain pathways and benefit from cross-feeding and common goods.

412

413

414   **Figure 4:** ***E. coli* gene expression analysis during growth on CCA alone, in pairwise conditions and with the community.** We
415   used RNASeq to investigate *E. coli* gene expression at three timepoints (1, 2 and 3 days) during growth on CCA
416   alone, in pairwise conditions (with *H. alvei*, *G. candidum* or *P. camemberti*) and with the community. We carried out a
417   principal component analysis on the rlog transformed expression data (Love et al., 2015) **(A)**. Using DESeq2 (Love et
418   al., 2015), we identified up and down-regulated genes in each pairwise condition at each timepoint compared to
419   growth alone. We kept only genes associated with a p-value lower than 1% and an absolute logFC higher than 1. For
420   each pairwise condition, we pooled together up-regulated genes at any timepoint and did the same for down-
421   regulated genes. Comparing the up- or down-regulated genes in the different pairwise conditions, we highlighted
422   overlapping response as well as specific responses depending on the partner **(B)**. Then, we carried out functional
423   enrichment analysis on the KEGG pathways for up- or down-regulated genes in each pairwise condition. To do so,
424   we used the R package clusterProfiler (Yu et al., 2012) as well as Benjamini-Hochberg multiple comparison
425   correction. Only the Kegg pathways enriched with a p-value lower than 5% were considered **(C)**.

426

427

428

429 **Comparison of differentially expressed genes in pairwise condition versus alone and**
430 **community growth versus alone.**

431   To determine whether *E. coli* gene expression significantly changes when grown with the
432 full community as compared to growth in pairwise conditions, we first calculated *E. coli* gene
433 logFC at each timepoint between growth with the community and growth alone. We further
434 analyzed genes with adjusted p-values lower than 1% and absolute logFC greater than 1. After
435 pooling across timepoints, we identified 465 up-regulated and 476 down-regulated genes in the
436 presence of the community versus growth alone (Supplementary file 5). We then compared
437 these genes to the 966 up-regulated genes and 977 down-regulated genes in pairwise
438 conditions versus growth alone (Figure 5A).

439   First, we identified 549 genes that were specifically up-regulated in pairwise conditions
440 versus growth alone and not up-regulated in community versus growth alone. KEGG pathway
441 enrichment analysis highlighted that these genes were mostly associated with quorum sensing,
442 fatty acid metabolism and oxidative phosphorylation (Figure 5B). This observation suggests that
443 addition of other species in the community counteracts or prevents pairwise specific interactions
444 from that partner. 416 genes were found to be up-regulated both in pairwise and community
445 growth versus growth alone. Enrichment analysis highlighted functions that were previously
446 described as up-regulated in most of the pairwise conditions: aminoacyl-tRNA-synthetase and
447 energy metabolism (Figure 5B). This suggests that certain interactions that *E. coli* experienced
448 in pairwise conditions are conserved in the community context. To investigate if the addition of
449 potentially similar interactions from different partners leads to an amplified response, we
450 explored if the magnitude of expression changes in these pathways is higher in the community.
451 We performed differential expression analysis on the genes comparably regulated in pairwise
452 conditions and with the community (Supplementary file 6). 50 of the 416 up-regulated were
453 significantly more up-regulated in community growth compared to pairwise growth. Among them,
454 sulfate assimilation genes were over-represented. This suggests that similar pairwise
455 interactions may be additive in the community leading to a stronger transcriptional response.

456   Finally, 49 genes were specifically up-regulated with community growth versus growth
457 alone. Despite a low number of genes, KEGG pathways were significantly enriched (Figure 5B).
458 These enriched groups include the biosynthesis of valine, leucine and isoleucine, pyrimidine
459 metabolism as well as arginine and proline metabolism. Up-regulation of certain amino acid
460 biosynthesis pathways suggests that despite potential cross feeding from individual partners,
461 addition of many partners eventually leads to competition. Up-regulation of pyrimidine, arginine
462 and proline metabolism however is part of a larger response; the response to nitrogen starvation

16

463   (Figure 5C). This response facilitates cell survival under nitrogen-limited conditions. Specifically,

464   up-regulated genes included all the genes involved in the regulatory loop of the transcriptional

465   regulator NtrC (*glnL*) and nitrogen utilization as well as NtrC transcriptional targets: the

466   transcriptional regulator Nac (*nac*), the operon *rutABCDEFG* involved in ammonium production

467   by uracil catabolism, the *astABCDE* operon constituting the arginine degradation pathway (AST

468   pathway) and the two regulators of the stringent response, *relA* and *spoT*. Thus, addition of

469   species in the community specifically triggers the activation of the response to nitrogen

470   starvation, which suggests a potential higher competition for nitrogen in the community context.

471   We performed a similar analysis on down-regulated genes in pairwise conditions and

472   with the community versus growth alone to investigate if transcriptional down-regulation in

473   pairwise and community conditions are similar (Figure 5A). 527 genes appeared specifically

474   down-regulated in pairwise conditions versus growth alone. Despite the large number of genes,

475   no KEGG pathway was significantly enriched among these genes. However, the *rutABCDEFG*

476   and *astABCDE* operons associated with the response to nitrogen starvation were down-

477   regulated in each pairwise conditions (Figure 5C). This observation strongly suggests that

478   presence of the community triggers a highly specific response that would be down-regulated in

479   the presence of only one species.

480   We identified 448 genes that were down-regulated during both pairwise and community

481   growth conditions versus growth alone. Enrichment analysis pointed to the down-regulation of

482   amino acid biosynthesis, lysine biosynthesis as well as cysteine and methionine metabolism.

483   Therefore, consistent with our RB-TnSeq data, this suggests that cross-feeding from a single

484   partner is maintained in a more complex context. Finally, 28 genes were specifically down-

485   regulated when *E. coli* was grown with the community. Individual analysis of these genes

486   highlighted 9 genes that code for uncharacterized proteins but did not highlight over-

487   represented functions.

488   To conclude, most of the changes in *E. coli* gene expression during growth with the

489   community were similar to those observed in pairwise conditions. Moreover, some of these

490   changes were amplified in the community compared to pairwise. This suggests that while a

491   large part of transcriptional regulation in community results from pairwise interactions, similar

492   interactions from different partners may be additive in the community and exert a stronger

493   impact on transcription. Also, the observed changes in nitrogen availability-related transcription

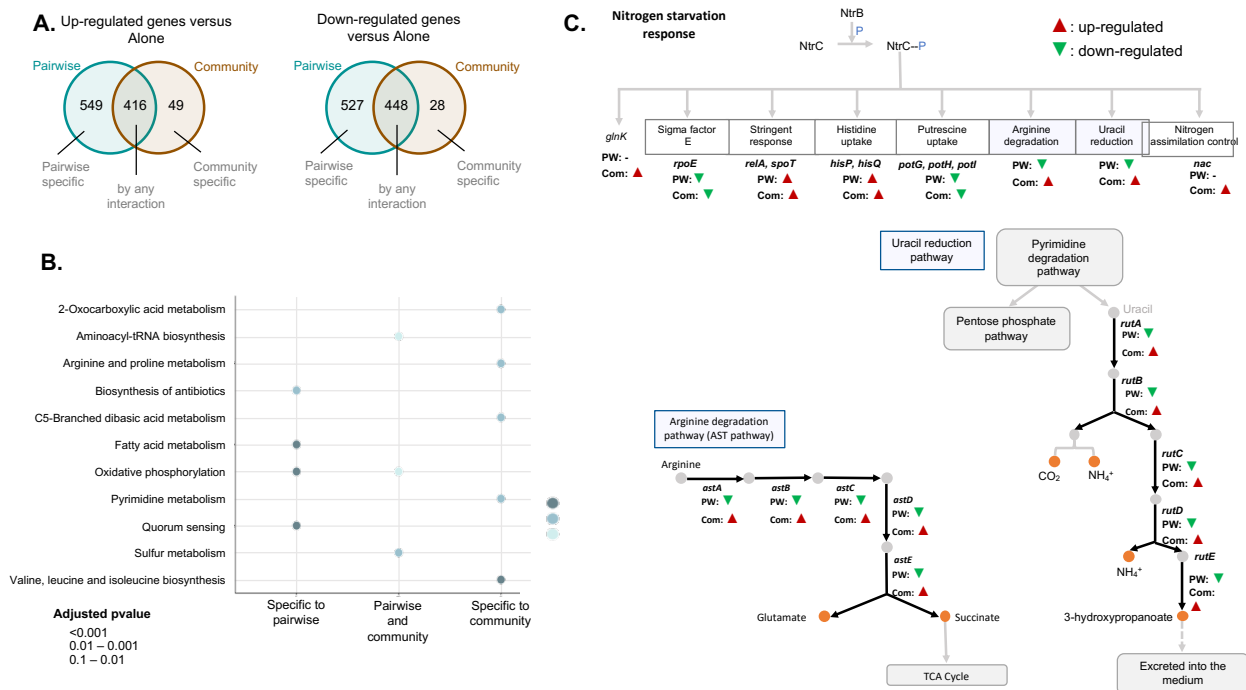494   suggest that community growth may induce new metabolic limitations.

495

496

497 **Figure 5: Comparison of differentially expressed genes in pairwise conditions versus alone and community**
498 **growth versus alone**. Using DESeq2 (Love et al., 2015), we identified up and down-regulated genes during growth
499 with the community compared to growth alone at each timepoint. We kept only genes associated with a p-value lower
500 than 1% and an absolute logFC higher than 1. We pooled together up-regulated genes at any timepoint and did the
501 same for down-regulated genes. Then, we compared up-regulated genes in pairwise conditions with genes up-
502 regulated during growth with the community and down-regulated genes in pairwise conditions with down-regulated
503 genes during growth with the community **(A)**. We further performed functional enrichment analysis on KEGG
504 pathways using the R package clusterProfiler (Yu et al., 2012) on the genes differentially expressed in different
505 situations. We used Benjamini-Hochberg multiple comparison correction and only the Kegg pathways enriched with a
506 p-value lower than 5% were considered **(B)**. Within the genes specifically up-regulated during growth with the
507 community, we observed the up-regulation of multiple genes associated with the nitrogen starvation response **(C)**.

508

509 **RB-TnSeq identification of genes essential for *Pseudomonas psychrophila* JB418 growth**

510 **with the bloomy rind cheese community**

511 Because *E. coli* is a well-characterized model organism, it provides a good starting point

512 for genetic analysis of interactions in this system. However, because *E. coli* is not typically found

513 in cheese, we wanted to characterize interactions using an endogenous species. We selected

514 *Pseudomonas psychrophila* JB418, which was previously isolated from a bloomy rind cheese

515 sample (Wolfe et al. 2014), and is genetically tractable,

516 We first sequenced, de novo assembled, and annotated the genome of this species (see

517 Materials and Methods). We used KEGG KOALA Blast (Kanehisa, Sato, and Morishima 2016)

518 to obtain KEGG identifiers for coding sequences. 2983 coding sequences (51.5%) were

519 assigned a KEGG identifier. We then used the same barcoded, transposon delivery library to

520 construct a RB-TnSeq library in this strain (see materials and methods for library construction

521 and characterization). We used TnSeq analysis to map barcode insertions to the annotated

522 genome. 272,329 insertions were mapped and 143,491 were centrally inserted (transposons

523 mapping to the first and last 10% of a gene were discarded) at 77,766 different locations. With

524 4811 genes with at least one central insertion, the library covers 83% of *P. psychrophila*

525 genome (Figure 6A). The genes with no central insertion are likely to be essential genes for

526 growth on LB.

527 We then grew the *P. psychrophila* library for 3 days on CCA either alone (supplementary

528 figure 4), in pairwise conditions (with *H. alvei*, *G. candidum* or *P. camemberti*) and with the full

529 community. Fitness values were successfully calculated for 3060 genes and the same statistical

530 screening (absolute t-value >=3) and fitness screening (fitness value < 0) were used to identify

531 essential genes. We first analyzed the 179 genes essential for *P. psychrophila* growth alone

532 (Figure 6B and supplementary file 7). Most of the essential genes for *P. psychrophila* growth

533 alone on CCA appeared to be associated with metabolism. More specifically, they are predicted

534 to be involved in amino acid biosynthesis (all amino acids except for lysine), biosynthesis of

535 cofactors and vitamins as well as energy metabolism (sulfate assimilation). Also, pathways

536 associated with membrane transport were related to iron-siderophore uptake and glycine

537 betaine transport. Essential functions for *P. psychrophila* growth alone were consistent with the

538 previous functions identified for *E. coli* growth alone.

539 Next, we identified 152 genes essential for *P. psychrophila* JB418 growth with *H. alvei*,

540 176 genes essential for growth with *G. candidum*, and 164 genes essential for growth with *P.*

541 *camemberti* (Supplementary file 7). Altogether, they represent 205 genes essential for *P.*

542 *psychrophila* growth in pairwise conditions. Compared to growth alone, 71 genes are no longer

543 required for *P. psychrophila* during pairwise growth (Figure 6Ci). We mapped these relieved

544 genes to the KEGG BRITE database to identify potentially relieved functions. Of the 71 genes,

545 only 27 were assigned a KEGG annotation. Here, no pathway appeared to be clearly relieved

546 (Supplementary figure 5).

547 We identified 74 newly essential genes for *P. psychrophila* growth in pairwise conditions

548 compared to growth alone (Figure 6Ci) and 34 of them were mapped to the KEGG Brite

549 DataBase (Supplementary figure 5). These genes included 11 associated with aromatic amino

550 acid biosynthesis. Since amino acid biosynthesis is also required for growth alone, these data

551 suggest that introduction of a partner leads to an amplified depletion of free amino acids, likely

552 due to competition for these nutrients. Additionally, 4 genes associated with DNA repair were

553 identified as essential in the presence of another species. This could suggest presence of a

554 toxic stress produced from the partner.

19

555   Finally, we identified 156 genes essential for *P. psychrophila* when grown with the
556   community (Supplementary file 7). Compared to growth alone, we identified 69 genes that were
557   no longer required (Figure 6Cii). Then, comparing these 69 relieved genes in community versus
558   growth alone with the relieved requirements in pairwise conditions versus growth alone (Figure
559   6Ciii), we concluded that only 8 genes were no longer required uniquely in the community. 4 of
560   these 8 genes were associated with valine and isoleucine biosynthesis, suggesting that addition
561   of more species might lead to cross-feeding of these amino acids similar to what was observed
562   with *E. coli*.

563   We identified 46 genes that were newly essential for *P. psychrophila* growth with the
564   community versus growth alone (Figure6Cii). Comparing these genes to the newly essential
565   genes identified in pairwise growth versus growth alone, we identified 8 genes uniquely
566   essential during community growth and not during any pairwise growth condition (Figure 6Civ).
567   Of these genes, 4 were annotated as either transporters or putative transporters. Here, as for *E.*
568   *coli*, we observed that newly essential and relieved requirements in the community context were
569   mostly maintained from pairwise conditions, while not all pairwise newly essential or relieved
570   requirements were maintained in the community.

571   To summarize, essential functions for *P. psychrophila* growth alone on CCA were
572   consistent with the ones identified for *E. coli*. While demonstrating important environmental
573   parameters to be considered for growth on CCA such as low available iron, low available free
574   amino acid and high osmolarity, it also underlines the robustness of the RB-TnSeq approach in
575   accurately identifying essential genes. Analysis of the functions of newly essential or relieved
576   genes in interactive contexts allowed us to examine potential *P. psychrophila* interactions with
577   the members of our model community. Interestingly, the possible interactions probed with *P.*
578   *psychrophila*, whether they are associated with relieved essential genes or newly essential
579   genes, appeared to differ from the ones identified with *E. coli*.
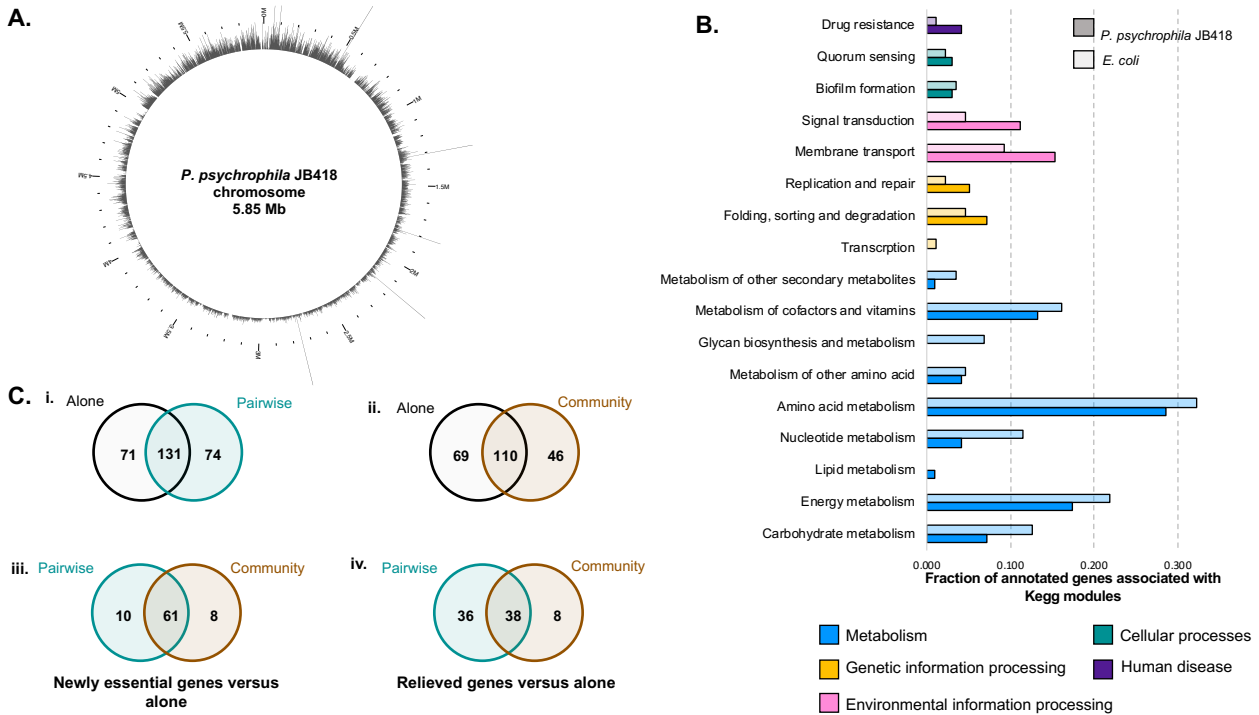
580

Figure 6: Identification of genes essential for *P. psychrophila* JB418 growth on CCA and with the model community. We built a barcoded-transposon library in the cheese isolate *P. psychrophila* JB418. 272,329 insertions were mapped to the genome. The library covers 83% of *P. psychrophila* JB418 genome. On the chromosome, each bar represents the number of insertions per kB **(A)**. We grew the library alone on CCA and identified 179 genes essential for *P. psychrofila* JB418. We annotated *P. psychrophila* JB418 using KEGG KOALA Blast (Kanehisa, Sato, and Morishima 2016). 98 of the 179 genes were attributed a KEGG annotation. To compare the relative importance of essential functions for growth alone observed for *P. psychrophila* JB418 and for *E. coli* we represented, per KEGG module, the ratio between the number of hits for the module and the number of genes with KEGG annotation **(B)**. We further identified *P. psychrophila JB418* essential genes for growth in the different pairwise conditions and with the community. Then, we determined which genes were newly essential or relieved in pairwise conditions compared to growth alone as well as newly essential or relieved genes with the community compared to growth alone. Finally, we compared newly essential genes in pairwise conditions versus alone with newly essential genes in the community versus alone as well as relieved genes in pairwise condition versus alone with relieved genes in community versus alone **(C)**.

DISCUSSION

In this work, we combined high-throughput, genome scale approaches to begin to define microbial interactions within a model community. Our integrative approach relies on the use of the model species *E. coli* as a genetic probe grown either alone, in pairwise culture with each community member or with the complete community. We first used RB-TnSeq to determine *E. coli's* genetic requirements in these different contexts associated with increasing levels of complexity. Our analysis showed that 3.86% of the genes in the *E. coli* genome were important to grow in our experimental environment. 53.1% of these genes appeared to be no longer required when grown with a partner but a distinct set of genes (comprising 1.98% of the *E. coli*

607 genome) appeared to be specifically required during interactive conditions. As identification of
608 differences for genes essential across conditions is based on a stringent filtering cutoff for
609 consistent fitness values, some differences might be attributed to false negative results. For
610 example, during growth alone, 11 genes that were found to be essential only at day 1 become
611 essential at all timepoints if we relax our t-score cutoff to absolute t-score >= 2 and increase our
612 false discovery rate to 2%. However, essentiality differences between timepoints can also point
613 out that genetic requirements do change over time. A method for quantitative comparison of
614 fitness values across conditions or timepoints will need to be developed to overcome that issue.
615 However, since we expect the false negative results to be randomly distributed across *E. coli*'s
616 genes, we consider that if genes belonging to the same pathway or function are specially
617 observed in a condition and not another, they are very likely to be involved in a microbial
618 interaction. To complement the RB-TnSeq approach, we used RNASeq and differential gene
619 expression analysis to investigate *E. coli* transcriptional response during growth in pairwise
620 culture and with the community compared to a growth alone. Here, we highlighted a deep
621 reorganization of gene expression whenever *E. coli* is in the presence of other species.

622 Our work illustrates for the first time the use and complementarity of RB-TnSeq and
623 RNASeq for microbial interaction analysis. These techniques lead to consistent interpretations
624 and each method compensated for the limitations of the other. RNASeq provides a
625 comprehensive view of interactions by looking at the transcriptional response of *E. coli* to an
626 interactive context. However, it fails to provide information on the actual requirement of a given
627 gene or pathway for an interactions. RB-TnSeq identifies the interactions that have a strong
628 impact on the *E. coli* growth phenotype including genes for which expression is constitutive in all
629 conditions and won't be highlighted with RNASeq analysis. On the other hand, gene
630 redundancy and the potential presence of "common good" molecules in the library are expected
631 to prevent identification of certain interactions by RB-TnSeq but not by RNASeq. This was
632 illustrated by the absence of essentiality of the biosynthesis and secretion of enterobactin but
633 the essentiality of enterobactin uptake system. It is very likely that the majority of the library
634 synthesizes and secretes enterobactin and thus the enterobactin synthesis mutants can uptake
635 this "common good" and survive. Interestingly, RNASeq analysis demonstrated upregulation of
636 the enterobactin biosynthesis pathway, suggesting enterobactin synthesis is necessary in an
637 interactive context. Similarly, differential expression analysis of the genes highlighted by RB-
638 TnSeq as essential in any condition allowed us to explore the presence of interactions
639 associated with these core functions. For example, it highlighted that iron-uptake was amplified
640 during interactions suggesting competition for that already limiting nutrient.

641   Together, interpretation of the functional analysis of RB-TnSeq and RNASeq results
642 pointed to possible microbial interaction mechanisms in the different growth contexts. Addition
643 of genetic requirements associated with metabolism and/or up-regulation of similar functions
644 can be interpreted as presence of metabolic competition toward associated compounds.
645 Conversely, alleviation of biosynthetic genes and/or downregulation of biosynthetic functions
646 can point out cross-feeding of associated compounds. Finally, production of a specific stress by
647 the partner can be deduced from addition of requirements associated with stress response
648 and/or up-regulation of similar functions.

649   Combining RB-TnSeq and RNASeq analysis, we identified potential interactions in
650 pairwise and community conditions. We have observed competition for iron in most interactive
651 conditions and in pairwise conditions, the partner appeared to produce possible toxic
652 compounds. Analysis of available genomes of *H. alvei* (Tan, Yin, and Chan 2014), *G. candidum*
653 (Polev et al. 2014) and *P. camemberti* (Cheeseman et al. 2014) in AntiSmash (Weber et al.
654 2015) highlighted the ability of each species to produce extracellular metabolites with potential
655 toxic properties including bacteriocins for *H. alvei* and diverse polyketides and terpenes for the
656 fungi. Analysis of extracellular metabolites will be necessary to characterize the molecular
657 signature of these stresses. Finally, we consistently observed possible amino acid cross feeding
658 from fungal partners, either in pairwise or community growth. This is very likely due to the fungal
659 capacities of secreting proteases and digesting the small peptides and proteins present in the
660 environment (Kastman et al. 2016; R. Boutrou, Aziza, and Amrane 2006; Rachel Boutrou,
661 Kerriou, and Gassi 2006). The observation of cross-feeding strengthens the statement that
662 cross-feeding events are widespread within communities (Goldford et al. 2017).

663   In our bottom-up approach, RNASeq and RB-TnSeq consistently highlighted that only a
664 fraction of pairwise interactions were conserved in the community. This highlights the existence
665 of higher order interactions counteracting simple pairwise interactions (Billick and Case 1994).
666 Higher order interactions could be compensatory pairwise interactions or upstream effects
667 ending in pairwise interaction cancellation. Although community structure is argued to be
668 predictable from pairwise interactions in specific cases, higher order interactions are believed to
669 be responsible for the general lack of predictability (Billick and Case 1994; Friedman, Higgins,
670 and Gore 2017; Momeni, Xie, and Shou 2016). Here, our work demonstrated the existence of
671 these higher order interactions even within a very simple community.

672   We generated data using both *E. coli*, a non-endogenous species, and *P. psychrophila*
673 JB418, a cheese endogenous species. We identified similar functions in both species as
674 important requirements to grow in the cheese environment. However, growth in interactive

contexts highlighted different interactions. We hypothesize that either *E. coli* and *P. psychrophila* do not similarly cope with the presence of a given partner or the partner species response depends on the probe-species as well as a combination of both scenarios. Altogether, this suggests important species-specific interaction patterns. However, our analysis was limited, especially for *P. psychrophila*, by genes which were poorly annotated or not characterized. Further investigation and characterization of these genes would improve interaction characterization and may highlight new microbial interactions pathways.

This study represents a novel approach to investigating microbial interactions in communities, and revealed the intricacy, redundancy and specificity of the many interactions governing a simple microbial community. The ability of *E. coli* to act as a probe for molecular interactions, the robustness of RB-TnSeq, and its complementarity with RNASeq open new paths for investigating molecular interactions in more complex communities, independently of their genetic tractability, and can thus contribute to a better understanding of the complexity and diversity of interactions within microbiomes.

MATERIAL AND METHODS

Strains and media

The following strains have been used to reconstruct the bloomy rind cheese community: *H. alvei* JB232 isolated previously from cheese (Wolfe et al. 2014) and two industrial cheese strains: *G. candidum* (*Geotrichum candidum* GEO13 LYO 2D, Danisco – CHOOZIT™, Copenhagen, Denmark) and *P. camemberti* (PC SAM 3 LYO 10D, Danisco - CHOOZIT™). The strain *P. psychrophila* JB418 was isolated from a sample of Robiola due latti (Italy) previously (Wolfe et al) and used for all the experiments involving *Pseudomonas*. All the *E. coli* strains used in this study shared the same genetic background of the initial strain *E. coli* K12 BW25113. The use of the different strains is described in Table 1.

Experiment	*E. coli* strain(s)	Ref
RB-TnSeq	*E. coli* Keio_ML9 library	(Wetmore et al. 2015)
Growth assays	*E. coli* "JW0024" strain (undisrupted mutant)	(Baba et al. 2006)
Competition assays	WT: Keio ME9062 Mutants: (supplementary figure 2)	(Baba et al. 2006)

707

708 **Table 1: *E. coli* strains used during the study**

709

710 **Growths assays on 10% cheese curd agar, pH7**

711 All growth assays have been carried out in at least triplicates on 10% cheese curd agar, pH7

712 (CCA) (10% freeze-dried Bayley Hazen Blue cheese curd (Jasper Hill Farm, VT), 3% NaCl,

713 0.5% xanthan gum and 1.7% agar). The pH of the CCA was buffered from 5.5 to 7 using 10M

714 NaOH.

715

716 *E. coli* and *P. psychrophila* JB418 growth on CCA:

717 Growth assays have been carried out for the *E. coli* "JW0024" strain (Baba et al., 2006) and *P.*

718 *psychrophila* JB418. For each species, $7*10^7$ cells were inoculated on the surface of a Petri dish

719 containing 20mL of CCA ($12.7*10^4$cells/cm^2) after an overnight pre-culture in liquid LB-

720 kanamycin (50μg/ml) at 37°C for *E. coli* or at room temperature (RT) for *P. psychrophila* JB418.

721 Growth on CCA was then carried out for 3 days at RT. Plugs of 0.44cm^2 were removed from the

722 Petri dish at T= 0h, 6h, 12h, 24h, 36h, 48h, 72h, 120h and 240h for *E. coli* and T=0h, 12h, 24h,

723 48h, 72h and 240h for *P. psychrophila* JB418. They were homogenized in 1 mL of PBS1X-

724 Tween0.05% and 3 dilutions were plated on LB-kanamycin (50μg/mL) for *E. coli* or LB for *P.*

725 *psychrophila*. Plates were incubated for 24h at 37°C for *E. coli* and at RT for *P. psychrophila*

726 JB418. After incubation, colony forming units (CFU) counting was used to estimate the number

727 of bacterial cells on the cheese curd agar plates.

728

729 *E. coli* growth in pairwise cultures:

730 Pairwise cultures were carried out on 96 well plates containing 200μL of CCA per well. *E. coli*

731 was co-inoculated with either *H. alvei*, *G. candidum* or *P. camemberti* at a ratio of 1:1 cells

732 (1000 cells of *E. coli* and 1000 cells of the partner). Single growth of *H. alvei*, *G. candidum* and

733 *P. camemberti* were also measured by inoculating 2000 cells of each species on individual wells.
734 At T=0h, 6h, 12h, 24h, 36h, 48h and 72h, a pellet of each condition was harvested,
735 homogenized in 1 mL of PBS1X-Tween0.05% and 3 dilutions were plated on LB, LB-kanamycin
736 (50µg/mL), LB-chloramphenicol (50µg/mL) and/or LB-kanamycin:cyclohexamide (50µg/mL and
737 10µg/mL) to follow the growth of each species (see Table 2). Plates were then either incubated
738 overnight at 37°C (LB, LB-kanamycin and LB-kanamycin:cyclohexamide plates) or for 2 days at
739 RT (LB-chloramphenicol plates) before determination of CFUs.
740
741 *E. coli* growth in the bloomy rind cheese community:
742 Cultures were grown on 96 well plates containing 10% cheese curd agar at pH7. The four
743 species (*E. coli*, *H. alvei* JB232, *G. candidum* and *P. camemberti*) have been co-inoculated at a
744 ratio of 10:10:10:1. At T=0h, 6h, 12h, 24h, 36h, 48h and 72h, a pellet was harvested,
745 homogenized in 1 mL of PBS1X-Tween0.05%, serially diluted and plated on LB-cyclohexamide
746 (10µg/mL), LB-chloramphenicol (50µg/mL) and LB-kanamycin:cyclohexamide (50µg/mL and
747 10µg/mL) (see Table 2) to follow the growth of each species. Plates were then either incubated
748 overnight at 37°C (LB-cyclohexamide and LB-kanamycin:cyclohexamide plates) or for 2 days at
749 RT (LB-chloramphenicol plates) before CFU counting. *P. camemberti* and *G. candidum* colonies
750 are morphologically distinct and CFUs of each species could be counted from the same plate.
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766

Growth assay	Plates used
E. coli + *H. alvei* JB232	LB (*E. coli* + *H. alvei* JB232 CFU's) LB-kanamycin (*E. coli* CFU's)
E. coli + *G. candidum*	LB-kanamycin:cyclohexamide (*E. coli* CFU's) LB-chloramphenicol (*G. candidum* CFU's)
E. coli + *P. camemberti*	LB-kanamycin:cyclohexamide (*E. coli* CFU's) LB-chloramphenicol (*P. camemberti* CFU's)
H. alvei	LB
G. candidum	LB-chloramphenicol
P. camemberti	LB-chloramphenocol
E. coli + Community	LB-cyclohexamide (*E. coli* and *H. alvei* JB232 CFU's), LB-kanamycin:cyclohexamide (*E. coli* CFU's), LB-chloramphenicol (*G. candidum* and *P. camemberti* CFU's)

Table 2: Organization of CFU's quantification for growth assays

P. psychrophila JB418 genome sequencing, assembly and annotation

P. psychrophila JB418 gDNA was sequenced using Pacific Biosciences (PacBio), Oxford Nanopore Minion (Oxford Nanopore, Oxford, UK), and Illumina sequencing. PacBio library preparation and sequencing were performed by the IGM Genomics Center at the University of California, San Diego. Nanopore library preparation and sequencing were done at the University of California, Santa Barbara as part of the KITP Quantitative Biology summer course. Illumina library preparation and sequencing were done at the Harvard University Center for Systems Biology. Canu was used to assemble the PacBio and nanopore reads (Koren et al. 2017). Pilon was then used to polish the assembly using Illumina genomic sequencing data (Walker et al. 2014). The assembled genome was annotated using the Integrated Microbial Genomes & Microbiomes (IMG/M) system (Markowitz et al. 2012). The *P. psychrophila* JB418 genome is 6,072,477 nucleotides long. It contains a single circular chromosome of 5.85 Mb and 4 plasmids

784 of 172.2 Kb, 37.7Kb, 5.8Kb and 2.4Kb. 6060 genes including 5788 open reading frames were
785 identified.

786 **Transposon mutant library construction in *P. psychrophila* JB418**

787 *P. psychrophila* JB418 was mutagenized by conjugation with *E. coli* strain APA766 (donor
788 WM3064 which carries the pKMW7 Tn5 vector library) (Wetmore et al. 2015). This donor strain
789 is auxotrophic for diaminopimelic acid (DAP). The full collection of the APA766 donor strain
790 (1mL) was grown up at 37°C overnight at 200 rpm. Four 25mL cultures (each started with
791 250µL of APA766 stock) were grown in LB-kanamycin:DAP (50µg/mL kanamycin and 60µg/mL
792 DAP). A 20mL culture was started from an individual *P. psychrophila* JB418 colony in LB broth
793 and grown at RT overnight at 200 rpm. *E. coli* donor cells were washed twice with LB and
794 resuspended in 25mL LB. Donor and recipient cells were then mixed at a 1:1 cell ratio based on
795 OD600 measurements, pelleted, and resuspended in 100µL. This was done separately for each
796 of the four *E. coli* cultures. 40µL were plated on nitrocellulose filters on LB plates with 60µg/mL
797 DAP. Two filters were used for each of the four conjugation mixtures (8 total conjugations). The
798 conjugations took place for 6 hours at RT. After 6 hours, the filters were each resuspended in
799 2mL of LB broth and then plated on LB:kanamycin (50µg/mL) for selection of transconjugants.
800 20 plates were plated of a 1:2 dilution for each conjugation (160 plates total). Transconjugants
801 were pooled and harvested after three days of growth on selection plates. The pooled mixture
802 was diluted back to 0.25 in 100mL of LB:kanamycin (50µg/mL). The culture was then grown at
803 RT to an OD600 of 1.3 before glycerol was added to 10% final volume and 1mL aliquots were
804 made and stored at -80°C for future use.

805

806 **TnSeq sequencing library preparation for *P. psychrophila* JB418 and TnSeq data analysis**
807 Library preparation was performed as in Wetmore, et al. 2015 with slight modifications
808 (Wetmore et al. 2015).

809

810 DNA extraction: DNA was extracted from the *P. psychrophila* JB418 RB-TnSeq library by
811 phenol:chloroform extraction. Briefly, the cell pellet was vortexed at maximum speed for 3
812 minutes in the presence of 500µL buffer B (200mM NaCl,20mM EDTA sterilized by filtration),
813 210µL of 20% SDS, a 1:1 mixture of 425-600µM and 150-212µm acid-washed beads, and
814 500µL of phenol:chloroform, pH 8. The sample was then centrifuged for 3 minutes at 4°C at

815   8000 rpm prior to removing the aqueous phase to a new tube. 1/10 of sample aqueous phase

816   volume of 3M sodium acetate was then added along with 1 aqueous phase volume of ice cold

817   isopropanol. The sample was then placed for ten minutes at -80°C before centrifugation for five

818   minutes at 4°C at 13000 rpm. The supernatant was discarded and 750µL of ice cold 70%

819   ethanol was added before another centrifugation for five minutes at 4°C at 13000 rpm. The

820   supernatant was discarded and the DNA pellet was allowed to air dry before resuspension in

821   50µL of nuclease-free water. DNA was quantified with Qubit double-stranded DNA high-

822   sensitivity assay kit (Invitrogen, Carlsbad, CA).

823

824   <u>DNA fragmentation and size selection</u>: 2µg of DNA was sheared with a Covaris E220 focused-

825   ultrasonicator with the following settings: 10% duty cycle, intensity 5, 200 cycles per burst, 150

826   seconds. DNA was split into 2 aliquots (1µg each) and samples were size-selected for

827   fragments of 300 bp using 0.85X Agencourt AMPure XP beads (Invitrogen) with a 1.4x ratio

828   following the manufacturer's instructions.

829

830   <u>Library preparation</u>: The entire 20µL volume of these two size-selected samples were then each

831   used as input into the NEBNext End Prep step 1.1 of the NEBNext Ultra DNA Library Prep Kit

832   for Illumina (New England Biolabs, Ipswich, MA) protocol. The remainder of the manufacturer's

833   protocol was then followed with the exception that for adapter ligation, we used 0.8µL of 15µM

834   double-stranded Y adapters. Adapters were prepared by first combining 5µL of 100µM

835   Mod2_TS_Univ and 5µL of 100µM Mod2_TruSeq. This mixture was then incubated in a

836   thermocycler for 30 min at 37°C, followed by ramping at 0.5°C per second to 97.5°C before a

837   hold at 97.5°C for 155 seconds. The temperature was then decreased by 0.1°C per five

838   seconds for 775 cycles, followed by a hold at 4°C. Annealed adapters were diluted to 15µM in

839   TE and stored at -80°C before use. AMPure XP ratios for a 200bp insert size were used as

840   recommended in Table 1.1 of the NEBNext Ultra DNA Library Prep Kit for Illumina manual.

841   To enrich for transposon-insertion sites, PCR amplification was done on the adapter-ligated

842   DNA with NEBNext Q5 Hot Start HiFi Master Mix and Nspacer_barseq_pHIMAR and

843   P7_MOD_TS_index3 primers (Wetmore et al. 2015) with the following program: 98°C 30 sec,

844   98°C 10 sec, 65°C 75 sec, repeat steps 2-3 24X, 65°C 5 min, and then maintained at

845   4°C. Following PCR and clean-up of step 1.5 of the NEBNext Ultra DNA Library Prep Kit for

846   Illumina manual, the two preps were pooled and the concentration was quantified with Qubit

847   double-stranded DNA high-sensitivity assay kit (Invitrogen). A second size selection clean-up

848 was performed by repeating step 1.5 of the NEBNext Ultra DNA Library Prep Kit for Illumina
849 manual.

850

851 <u>Library sequencing</u>: The sample was analyzed on an Agilent TapeStation and the average size
852 was 380bp and the concentration was 57pg/µL. This sample was then submitted for sequencing
853 on a HiSeq 2500 Rapid Run (150bp fragments, paired-end) at the UCSD IGM Genomics
854 Center:

855

856 <u>Library characterization</u>: TnSeq reads were analyzed with the perl script MapTnSeq.pl from
857 (Wetmore et al. 2015). That script maps each read to the *P. psychrophila* genome. The script
858 DesignRandomPool.pl (Wetmore et al. 2015) was used to generate the file containing the list of
859 barcodes that consistently map to a unique location as well as their location.

860

861 **RB-TnSeq experiments for *E. coli* and *P. psychrophila* JB418**
862 The *E. coli* barcoded transposon library Keio_ML9 and the *P. psychrophila* strain JB418 library
863 were used for fitness assays on CCA during growth alone, growth in pairwise condition with
864 each bloomy rind cheese community member and during growth with the full community.

865

866 <u>Library pre-culture</u>: Each library has to be initially amplified before use. One aliquot of each
867 library was thawed and inoculated into 25mL of liquid LB-kanamycin (50µg/mL). Once the
868 culture reached mid-log phase (OD=0.6-0.8), 5mL of that pre-culture was pelleted and stored as
869 the T0 reference for the fitness analysis. The remaining cells were used to inoculate the
870 different fitness assay conditions.

871

872 <u>Inoculations</u>: For each fitness assay, $7*10^6$ cells of the library pre-culture were inoculated on
873 10% cheese curd agar, pH 7 after having been washed in PBS1x-Tween0.05%. For each
874 pairwise assay, $7*10^6$ cells of the partner were co-inoculated with the library. For the community
875 assay, $7*10^6$ cells of H. lavei JB232 and *G. candidum* as well as $7*10^5$ cells of *P. camemberti*
876 were co-inoculated with the library. For each condition, assays were performed in triplicates.

877

878 <u>Harvest for gene fitness calculation</u>: Harvests were performed at T= 24h, 48h and 72h.
879 Sampling was done by flooding a plate with 1.5mL of PBS1X-Tween0.05% and gently scraping
880 the cells off. The liquid was then transferred into a 1.5mL microfuge tube and cells were pelleted

881 by centrifugation. After removing the supernatant, the cells were washed in 1 mL of RNA-protect
882 solution (Qiagen, Hilden, Germany), pelleted and stored at -80°C before further experiments.
883
884 gDNA and mRNA extraction: gDNA and mRNA were simultaneously extracted by a phenol-
885 chloroform extraction (pH 8) from samples of the competitive assays. For each extraction:
886 125µL of 425-600µm acid-washed beads and 125µL of 150-212µm acid-washed beads were
887 poured in a screw-caped 2mL tube. 500µL of 2X buffer B (200mM NaCl, 20mM EDTA) and
888 210µL of SDS 20% were added to the tube as well as the pellet and 500µL of
889 Phenol:Chloroform (pH 8). Cells were lysed by vortexing the tubes for 2 minutes at maximum
890 speed. Aqueous and organic phases were separated by centrifugation at 4°C, 8KRPM for 3
891 minutes and 450µL of the aqueous phase (upper phase) was recovered in a 1.5mL eppendorf
892 tube. 45µL of Sodium Acetate 3M and 450µL of ice cold isopropanol were added before
893 incubating the tubes at -80°C for 10 minutes. The tubes were then centrifuged for 5 minutes at
894 4°C at 13KRPM. The pellet was then washed in 750µL of 70% ice cold ethanol and re-
895 suspended in 50µL of DNAse/RNAse free water. Each sample was split into 2 times 25µL and
896 stored at -80°C until further analysis.
897
898 Library preparation and sequencing: After gDNA extraction, the 98°C BarSeq PCR as described
899 and suggested in Wetmore et al., 2015 was used to amplify only the barcoded region of the
900 transposons. Briefly, PCR was performed in a final volume of 50µL: 25µL of Q5 polymerase
901 master mix (New England Biolab), 10µL of GC enhancer buffer (New England Biolab), 2.5µL of
902 the common reverse primer (BarSeq_P1 – Wetmore et al., 2015) at 10µM, 2.5µL of a forward
903 primer from the 96 forwards primers (BarSeq_P2_ITXXX) at 10µM and 50ng to 2µg of gDNA.
904 For each triplicate, the PCR was performed with the same forward primer so all replicates of a
905 condition could be pooled and have the same sequencing multiplexing index. So, for *E. coli*
906 analysis, we performed 46 PCRs (1 T0 samples and 45 harvest samples) involving 16 different
907 multiplexing indexes. For *P. psychrophila* JB418 analysis, we performed 46 PCR (1 T0 and 45
908 harvest samples) involving 16 other multiplexing indexes. We used the following PCR program:
909 (i) 98°C - 4 minutes, (ii) 30 cycles of: 98°C – 30 seconds; 55°C – 30 seconds; 72°C – 30
910 seconds, (iii) 72°C – 5 minutes. After the PCR, 10µL of each of the 92 PCR products were
911 pooled together to create the BarSeq library(920µL) and 200µL of the pooled library were
912 purified using the MinElute purification kit (Qiagen). The final elution of the BarSeq library was
913 performed in 30µL in DNAse and RNAse free water.

914 The BarSeq library was then quantified using Qubit® dsDNA HS assay kit (Invitrogen) and
915 sequenced on HiSeq4000 (50bp, single-end reads), by the IGM Genomics Center at the
916 University of California, San Diego.

917

918 Data processing and fitness analysis: BarSeq data processing and gene fitness calculation was
919 performed separately for the *E. coli* and the *P. psychrophila* JB418 experiments. For each
920 library, BarSeq reads were processed using the custom perl script BarSeqTest.pl (Wetmore et
921 al. 2015). This script combines two perl scripts essential for the BarSeq data processing. After
922 the raw reads have been de-multiplexed into each conditions, the computational pipeline: (i)
923 identifies each barcode and the associated number of reads, (ii) calculates the strain fitness for
924 each insertion mutant as the log2 ratio of counts in the condition and in the T0 sample and (iii)
925 calculates each "gene fitness" as the average of strain fitness of all the insertion mutants
926 associated with that gene as well as calculation of a fitness t-score. The following rules were
927 applied during the fitness calculations: (i) only insertion mutants with barcode insertion within
928 the 10%-90% fraction of the genes were considered, (ii) barcodes with less than 3 reads in the
929 T0 were ignored and (iii) genes with less than 30 counts in T0 were ignored. For each library,
930 the pipeline uses a table where each barcode is mapped to a location in the genome. The Arkin
931 lab (Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley,
932 California, USA) kindly provided the TNSeq table for the *E. coli* library and we generated a
933 TNSeq table for *P. psychrophila* strain JB418.

934 We calculated *E. coli* and *P. psychrophila* JB418 genes fitnesses at T=24h, 48h and
935 72h in the following conditions: growth alone, growth with *H. alvei*, growth with *G. candidum*,
936 growth with *P. camemberti* and growth with the community.

937

Keio collection mutant competitive assays
939 We used mutants from the Keio collection to validate the genes identified by RB-TNseq as
940 having a significant fitness in *E. coli* (see list in supplementary data S3). Each mutant was
941 grown in a competition assay with the wild type (Keio ME9062 –(Baba et al. 2006)). 1000 cells
942 of a specific mutant were inoculated with 1000 cells of the wild type on the surface of the same
943 cheese plug in a 96 well plate containing 10% cheese curd agar, pH7. The number of the
944 mutant cells and the WT cells were calculated at T0 and day 1 after harvesting, homogenizing
945 the cheese plug, plating serial dilutions and counting CFUs. Experimental fitness of each mutant
946 was calculated as the log2 of the mutant abundance (mutant CFUs divided by total CFUs (WT +
947 mutant)) after 24 hours and its abundance at T0.

948

RNASeq and differential expression analysis

950 The RNASeq analysis was performed for the *E. coli* experiments only.

951

952 RNASeq libraries preparations: Libraries were prepared for duplicates of the following
953 conditions: *E. coli* growth alone, with *H. alvei*, with *G. candidum*, with *P. camemberti* and with
954 the community for T= 24h, 48h and 72h. RNA samples from the *E. coli* BarSeq experiment were
955 used to produce the RNASeq library.

956 Each library was prepared as follows. First, RNA samples were treated with DNAse using the
957 "Rigorous DNAse treatment" for the Turbi DNA-free kit (AMBIO – Life Technologies, USA) and
958 RNA concentration was measured by nucleic acid quantification in Epoch Microplate
959 Spectrophotometer (BioTek, Winooski, VT). Then, transfer RNAs and 5S RNA were removed
960 using the MEGAclear Kit Purification for Large Scale Transcription Reactions (AMBION, Life
961 Technologies, Waltham, MA) following manufacturer instructions. Absence of tRNA and 5S
962 RNA was verified by running 100ng of RNA on a 1.5% agarose gel and RNA concentration was
963 quantified by nucleic acid quantification in Epoch Microplate Spectrophotometer. Also, presence
964 of trace amounts of genomic DNA was assessed by PCR using universal bacterial 16S PCR
965 primers (Forward primer: AGAGTTTGATCCTGGCTCAG, Reverse Primer:
966 GGTTACCTTGTTACGACTT). Final volume of the PCR was 20µL: 10µL of Q5 polymerase
967 master mix (New England Biolab), 0.5µL of forward primer 10uM, 0.5µL of reverse primer 10uM
968 and 5µL of non-diluted RNA. PCR products were run on a 1.7% agarose gel and when genomic
969 DNA was amplified, another DNAse treatment was performed as well as a new verification of
970 absence of genomic DNA. Ribomosomal RNA depletion was performed using the Ribo-Zero
971 rRNA removal kit by Illumina (Illumina, San Diego, CA). Ribosomoal RNA depletion was
972 performed according to manufacturer instructions, we used 1µL of RiboGuard RNAse Inhibitor
973 in each sample as suggested, followed instructions for 1-2.5ug of RNA input and we used a 2:1
974 mix of bacterial Ribo-Zero Removal solution and yeast Ribo-Zero Removal solution. rRNA
975 depleted samples were recovered in 10µL after ethanol precipitation. Concentrations after
976 ribodepletion were measured using Qubit® RNA HS Assay Kits (Invitrogen). The RNASeq
977 library was produced using the NEBNext®UltraTM RNA Library Prep Kit for Illumina® for
978 purified mRNA or ribosome depleted RNA. We prepared a library with a fragment size of 300
979 nucleotides and used the 10uM NEBNext Multiplex Oligos for Illumina (Set 1, NEB #E7335) lot
980 0091412 and the NEBNext multiplex Oligos for Illumina (Set 2, NEB #E7500) lot 0071412. We
981 did the PCR product purification with 0.8X Agencourt AMPure XP Beads instead of 0.9X. Library

982   samples were quantified with Qubit® DNA HS Assay Kits before the quality and fragment size
983   were validated by TapeStation (HiSensD1000 ScreenTape). Library samples were pooled at a
984   concentration of 15nM each and were sequenced on HISeq4000 (50bp, single-end).

985

986   Differential expression analysis: RNASeq reads were mapped to the concatenated genome of
987   *Escherichia coli* K12 BW25113 (Grenier et al. 2014) and *H. alvei* using Geneious version R9
988   9.1.3 (http://www.geneious.com, (Kearse et al. 2012))). Only the reads that uniquely mapped to
989   a single location on the *E. coli* genome section were conserved. *E. coli* and *H. alvei* genome are
990   divergent enough so 50 nucleotides reads potentially originating from *H. alvei* mRNA would not
991   map to *E. coli* genome and few reads from *E. coli* would map on *H. alvei* genome.
992   *E. coli* expression analysis was performed using the R packages: Rsamtool (R package version
993   1.30.0), GenomeInfoDb (R package version 1.14.0.), GenomicFeatures (Lawrence et al. 2013),
994   GenomicAlignments , GenomicRanges (Lawrence et al. 2013) and DESeq2 (Love et al. 2015).
995   We followed the workflow described by Love et al. and performed the differential expression
996   analysis using the package DESeq2. Differentially expressed genes between two conditions
997   were selected with an adjusted p-value lower than 1% and an absolute log of fold change equal
998   or greater than 1.

999

1000   **KEGG pathway enrichment analysis**
1001   Functional enrichment analysis was performed using the R package clusterProfiler (Yu et al.
1002   2012). We used the latest version of the package org.EcK12.eg.db for *E. coli* annotations (R
1003   package version 3.5.0.). We used Benjamini-Hochberg for multiple comparison correction and
1004   only the KEGG pathways enriched with an adjusted p-value lower than 5% were considered.
1005   The *E. coli* database was used with the simplify function to avoid redundant GO Term and only
1006   display the GO Term with the highest significance within a cluster.

1007

1014

1015

1016 **COMPETING INTEREST:**
1017 The authors declare that no competing interests exist.
1018

1019

1020 **REFERENCES**

1021 Albar, Abdulgader H., Hussein A. Almehdar, Vladimir N. Uversky, and Elrashdy M. Redwan.
1022 2014. "Structural Heterogeneity and Multifunctionality of Lactoferrin." *Current Protein & Peptide*
1023 *Science* 15 (8): 778–97.
1024

1025 Baba, Tomoya, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A.
1026 Datsenko, Masaru Tomita, Barry L. Wanner, and Hirotada Mori. 2006. "Construction of
1027 Escherichia Coli K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection."
1028 *Molecular Systems Biology* 2 (February): 2006.0008.
1029

1030 Billick, Ian, and Ted J. Case. 1994. "Higher Order Interactions in Ecological Communities: What
1031 Are They and How Can They Be Detected?" *Ecology* 75 (6). Ecological Society of America:
1032 1529–43.
1033

1034 Boutrou, Rachel, Liliane Kerriou, and Jean-Yves Gassi. 2006. "Contribution of Geotrichum
1035 Candidum to the Proteolysis of Soft Cheese." *International Dairy Journal / Published in*
1036 *Association with the International Dairy Federation* 16 (7): 775–83.
1037

1038 Boutrou, R., M. Aziza, and A. Amrane. 2006. "Enhanced Proteolytic Activities of Geotrichum
1039 Candidum and Penicillium Camembertii in Mixed Culture." *Enzyme and Microbial Technology*
1040 39 (2): 325–31.
1041

1042 Cashel, Michael, and Kenneth E. Rudd. 1996. "The Stringent Response." In *Esherichia Coli and*
1043 *Salmonella; Cellular and Molecular Biology 2*, edited by Frederick C. Neidhart, John L.
1044 Ingraham, K. Brooks Low, Boris Magasanik, Moselio Schaechter, and H. Edwin Umbarger,
1045 1410–38. American Society for Microbiology.
1046

1047 Cheeseman, Kevin, Jeanne Ropars, Pierre Renault, Joëlle Dupont, Jérôme Gouzy, Antoine
1048 Branca, Anne-Laure Abraham, et al. 2014. "Multiple Recent Horizontal Transfers of a Large
1049 Genomic Region in Cheese Making Fungi." *Nature Communications* 5: 2876.
1050

1051 Croucher, Nicholas J., and Nicholas R. Thomson. 2010. "Studying Bacterial Transcriptomes
1052 Using RNA-Seq." *Current Opinion in Microbiology* 13 (5): 619–24.
1053

1054 Dorel, Corinne, Philippe Lejeune, and Agnès Rodrigue. 2006. "The Cpx System of Escherichia
1055 Coli, a Strategic Signaling Pathway for Confronting Adverse Conditions and for Settling Biofilm
1056 Communities?" *Research in Microbiology* 157 (4): 306–14.
1057

1058 Falkowski, Paul G., Tom Fenchel, and Edward F. Delong. 2008. "The Microbial Engines That
1059 Drive Earth's Biogeochemical Cycles." *Science* 320 (5879): 1034–39.
1060

1061 Flint, Harry J., Karen P. Scott, Petra Louis, and Sylvia H. Duncan. 2012. "The Role of the Gut
1062 Microbiota in Nutrition and Health." *Nature Reviews. Gastroenterology & Hepatology* 9 (10):
1063 577–89.

Friedman, Jonathan, Logan M. Higgins, and Jeff Gore. 2017. "Community Structure Follows Simple Assembly Rules in Microbial Microcosms." *Nature Ecology & Evolution* 1 (5): 109.

Galia, Wessam, Francoise Leriche, Stéphane Cruveiller, Cindy Garnier, Vincent Navratil, Audrey Dubost, Stéphanie Blanquet-Diot, and Delphine Thevenot-Sergentet. 2017. "Strand-Specific Transcriptomes of Enterohemorrhagic Escherichia Coli in Response to Interactions with Ground Beef Microbiota: Interactions between Microorganisms in Raw Meat." *BMC Genomics* 18 (1): 574.

Gambino, Michela, and Francesca Cappitelli. 2016. "Mini-Review: Biofilm Responses to Oxidative Stress." *Biofouling* 32 (2): 167–78.

Goldford, Joshua E., Nanxi Lu, Djordje Bajic, Sylvie Estrela, Mikhail Tikhonov, Alicia Sanchez-Gorostiaga, Daniel Segre, Pankaj Mehta, and Alvaro Sanchez. 2017. "Emergent Simplicity in Microbial Community Assembly." *bioRxiv*. https://doi.org/10.1101/205831.

Grenier, Frédéric, Dominick Matteau, Vincent Baby, and Sébastien Rodrigue. 2014. "Complete Genome Sequence of Escherichia Coli BW25113." *Genome Announcements* 2 (5). https://doi.org/10.1128/genomeA.01038-14.

Guinee, T. P. 2004. "Salting and the Role of Salt in Cheese." *International Journal of Dairy Technology* 57 (2-3). Blackwell Science Ltd: 99–109.

Gunsalus, R. P. 1992. "Control of Electron Flow in Escherichia Coli: Coordinated Transcription of Respiratory Pathway Genes." *Journal of Bacteriology* 174 (22): 7069–74.

Hider, Robert C., and Xiaole Kong. 2010. "Chemistry and Biology of Siderophores." *Natural Product Reports* 27 (5): 637–57.

Kanehisa, Minoru, Yoko Sato, and Kanae Morishima. 2016. "BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences." *Journal of Molecular Biology* 428 (4): 726–31.

Kastman, Erik K., Noelani Kamelamela, Josh W. Norville, Casey M. Cosetta, Rachel J. Dutton, and Benjamin E. Wolfe. 2016. "Biotic Interactions Shape the Ecological Distributions of Staphylococcus Species." *mBio* 7 (5). https://doi.org/10.1128/mBio.01157-16.

Kearse, Matthew, Richard Moir, Amy Wilson, Steven Stones-Havas, Matthew Cheung, Shane Sturrock, Simon Buxton, Alex Cooper, Sidney Markowitz, Chris Duran, Tobias Thierer, Bruce Ashton, Peter Meintjes and Alexei Drummond. 2012. "Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data." *Bioinformatics* 28 (12): 1647–49.

Koren, Sergey, Brian P. Walenz, Konstantin Berlin, Jason R. Miller, Nichols H. Bergman, and Adam M. Phillippy. 2017. "Canu: Scalable and Accurate Long-Read Assembly via Adaptive K-Mer Weighting and Repeat Separation." *Genome Research* 27 (5): 722–36.

Landini, Paolo. 2009. "Cross-Talk Mechanisms in Biofilm Formation and Responses to Environmental and Physiological Stress in Escherichia Coli." *Research in Microbiology* 160 (4): 259–66.

1114
1115 Lawrence, Michael, Wolfgang Huber, Hervé Pagès, Patrick Aboyoun, Marc Carlson, Robert
1116 Gentleman, Martin T. Morgan, and Vincent J. Carey. 2013. "Software for Computing and
1117 Annotating Genomic Ranges." *PLoS Computational Biology* 9 (8): e1003118.
1118
1119 Love, Michael I., Simon Anders, Vladislav Kim, and Wolfgang Huber. 2015. "RNA-Seq
1120 Workflow: Gene-Level Exploratory Analysis and Differential Expression." *F1000Research* 4
1121 (October): 1070.
1122
1123 Markowitz, Victor M., I-Min A. Chen, Krishna Palaniappan, Ken Chu, Ernest Szeto, Yuri
1124 Grechkin, Anna Ratner, et al. 2012. "IMG: The Integrated Microbial Genomes Database and
1125 Comparative Analysis System." *Nucleic Acids Research* 40 (Database issue): D115–22.
1126
1127 McAdam, Paul R., Emily J. Richardson, and J. Ross Fitzgerald. 2014. "High-Throughput
1128 Sequencing for the Study of Bacterial Pathogen Biology." *Current Opinion in Microbiology* 19
1129 (June): 106–13.
1130
1131 Momeni, Babak, Li Xie, and Wenying Shou. 2016. "The Validity of Pairwise Models in Predicting
1132 Community Dynamics." *bioRxiv*. https://doi.org/10.1101/060988.
1133
1134 Olson, Michael A., Timothy W. Siebach, Joel S. Griffitts, Eric Wilson, and David L. Erickson.
1135 2017. "Genome-Wide Identification of Fitness Factors in Mastitis-Associated Escherichia Coli."
1136 *Applied and Environmental Microbiology*, November. https://doi.org/10.1128/AEM.02190-17.
1137
1138 Opijnen, Tim van, and Andrew Camilli. 2013. "Transposon Insertion Sequencing: A New Tool for
1139 Systems-Level Analysis of Microorganisms." *Nature Reviews. Microbiology* 11 (7): 435–42.
1140
1141 Polev, Dmitrii E., Kirill S. Bobrov, Elena V. Eneyskaya, and Anna A. Kulminskaya. 2014. "Draft
1142 Genome Sequence of Geotrichum Candidum Strain 3C." *Genome Announcements* 2 (5).
1143 https://doi.org/10.1128/genomeA.00956-14.
1144
1145 Price, Morgan N., Kelly M. Wetmore, Robert Jordan Waters, Mark Callaghan, Jayashree Ray,
1146 Jennifer V. Kuehl, Ryan A. Melnyk, et al. 2016. "Deep Annotation of Protein Function across
1147 Diverse Bacteria from Mutant Phenotypes." *bioRxiv*. https://doi.org/10.1101/072470.
1148 Raymond, Kenneth N., Emily A. Dertz, and Sanggoo S. Kim. 2003. "Enterobactin: An Archetype
1149 for Microbial Iron Transport." *Proceedings of the National Academy of Sciences of the United*
1150 *States of America* 100 (7): 3584–88.
1151
1152 Tan, Jia-Yi, Wai-Fong Yin, and Kok-Gan Chan. 2014. "Gene Clusters of Hafnia Alvei Strain FB1
1153 Important in Survival and Pathogenesis: A Draft Genome Perspective." *Gut Pathogens* 6 (July):
1154 29.
1155
1156 Wackwitz, B., J. Bongaerts, S. D. Goodman, and G. Unden. 1999. "Growth Phase-Dependent
1157 Regulation of nuoA-N Expression in Escherichia Coli K-12 by the Fis Protein: Upstream Binding
1158 Sites and Bioenergetic Significance." *Molecular & General Genetics: MGG* 262 (4-5): 876–83.
1159
1160 Walker, Bruce J., Thomas Abeel, Terrance Shea, Margaret Priest, Amr Abouelliel, Sharadha
1161 Sakthikumar, Christina A. Cuomo, et al. 2014. "Pilon: An Integrated Tool for Comprehensive
1162 Microbial Variant Detection and Genome Assembly Improvement." *PloS One* 9 (11): e112963.
1163

1164 Weber, Tilmann, Kai Blin, Srikanth Duddela, Daniel Krug, Hyun Uk Kim, Robert Bruccoleri,
1165 Sang Yup Lee, et al. 2015. "antiSMASH 3.0-a Comprehensive Resource for the Genome Mining
1166 of Biosynthetic Gene Clusters." *Nucleic Acids Research* 43 (W1): W237–43.
1167
1168 Wetmore, Kelly M., Morgan N. Price, Robert J. Waters, Jacob S. Lamson, Jennifer He, Cindi A.
1169 Hoover, Matthew J. Blow, et al. 2015. "Rapid Quantification of Mutant Fitness in Diverse
1170 Bacteria by Sequencing Randomly Bar-Coded Transposons." *mBio* 6 (3): e00306–15.
1171
1172 Wolfe, Benjamin E., Julie E. Button, Marcela Santarelli, and Rachel J. Dutton. 2014. "Cheese
1173 Rind Communities Provide Tractable Systems for in Situ and in Vitro Studies of Microbial
1174 Diversity." *Cell* 158 (2): 422–33.
1175
1176 Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. "clusterProfiler: An R
1177 Package for Comparing Biological Themes among Gene Clusters." *Omics: A Journal of
1178 Integrative Biology* 16 (5): 284–87.
1179
1180

1181
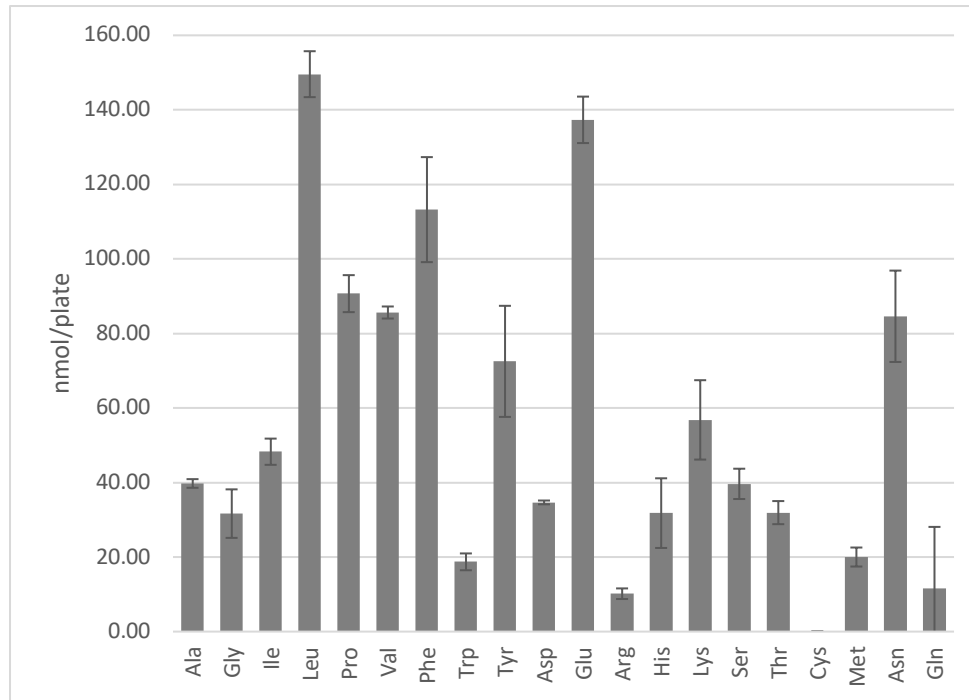
Supplementary figure 1: Quantification of free amino-acids in CCA. Free amino acid characterization and quantification have been carried out by the Proteomics & Mass Spectrometry Facility of the Donald Danforth Plant Science Center and each analysis has been performed in triplicate. Samples were prepared as according to the Proteomics & Mass Spectrometry Facility of the Donald Danforth Plant Science Center instructions. For free amino acid analysis 150 mg CCA were frozen in liquid nitrogen and grinded in the presence of 600 µL of water:chloroform:methanol (3:5:12 v/v). Tubes were then centrifugated at full speed for two minutes and supernatant was recovered in a new 2mL eppendorf tube. A second extraction with 600uL of water:chloroform:methanol was performed followed by a two minute centrifugation at full speed. Supernatant was then recovered and combined with the previous supernatant in a 2mL eppendorf tube. Then 300uL of chloroform and 450uL of water were added before centrifugation at full speed for two minutes. The upper phase was recovered and transferred to a new tube. Samples were dried in a speedvac overnight and stored at -20°C. The total concentration of free amino acids in CCA is 75.3 nmol/mL. Analysis of total amino acids was also performed by the Proteomics & Mass Spectrometry Facility of the Donald Danforth Plant Science Center. It highlights that free aminos are a very small fraction of total amino acid whose concentration is 16.5 ±2,97 µmol/mL, six times less than the LB concentration measured by Sezonov et al., 2012.

Ala: alanine, Gly: glycine, Ile: isoleucine, Leu: leucine, Pro: proline, Val: valine, Phe: phenylalanine, Trp: tryptophame, Tyr: tyrosine, Asp: aspartate, Glu: glutamate, Arg: arginine, His: histidine, Lys: lysine, Ser: serine, Thr: threonine, Cys: cysteine, Met: methionine, Asn: asparagine, Gln: glutamine.

Keio strain	Gene name	RB-TNSeq fitness	Experimental Fitness value	SEM of Experimental Fitness value	z-score
JW3932	argH	Negative	-0.61	0.02	24.60
JW0391	brnQ	Negative	-0.45	0.37	1.21
JW3789	corA	Negative	-4.37	0.30	14.58
JW2407	cysK	Negative	-1.72	0.91	1.90
JW2974	exbB	Negative	-1.48	0.05	27.25
JW2973	exbD	Negative	-3.25	0.08	43.16
JW5086	fepA	Negative	-1.97	0.46	4.28
JW0584	fepB	Negative	-1.60	0.13	12.42
JW0580	fepC	Negative	-3.11	0.74	4.21
JW0582	fepD	Negative	-3.13	0.64	4.85
JW0581	fepG	Negative	-2.07	1.02	2.03
JW0576	fes	Negative	-1.98	0.08	24.75
JW0742	galE	Negative	-3.69	0.41	8.89
JW2464	gcvR	Negative	-3.56	0.42	8.54
JW3377	nfuA	Negative	-4.09	0.29	14.07
JW0018	nhaA	Negative	-1.82	0.40	4.56
JW3627	recG	Negative	-0.63	0.28	2.23
JW2755	relA	Negative	-2.49	0.19	13.39
JW3624	rpoZ	Negative	-1.17	0.26	4.43
JW2880	serA	Negative	-0.24	0.30	0.79
JW4351	serB	Negative	-1.04	0.18	5.81
JW0890	serC	Negative	-1.01	0.15	6.85
JW5576	trkH	Negative	-0.57	0.10	5.49
JW5856	trxA	Negative	-0.75	0.54	1.39
JW3434	zntA	Negative	-0.64	0.19	3.37

1202
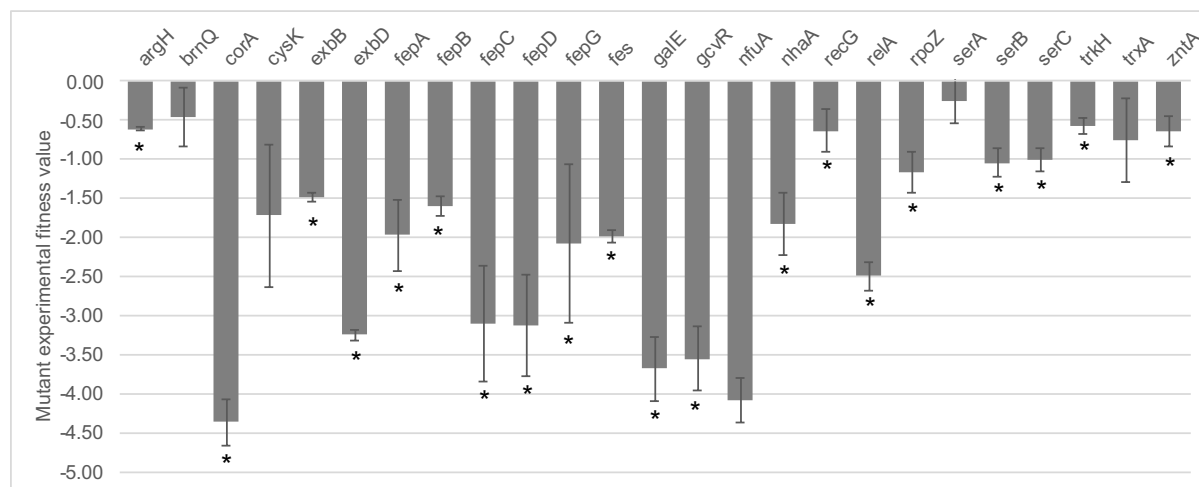
1203
1204
1205
1206 **Supplementary figure 2: Competitive assays of 35 mutants of the Keio collection (Baba et al., 2006).**
1207 Competition assays between single knock outs and the wild-type strain have been carried out for 35 strains
1208 associated with genes identified as essential for *E. coli* growth using RB-TnSeq (Significant fitness lower than -1 after
1209 1 day of growth). * highlights fitness values different from 0 with a confidence higher than 95%
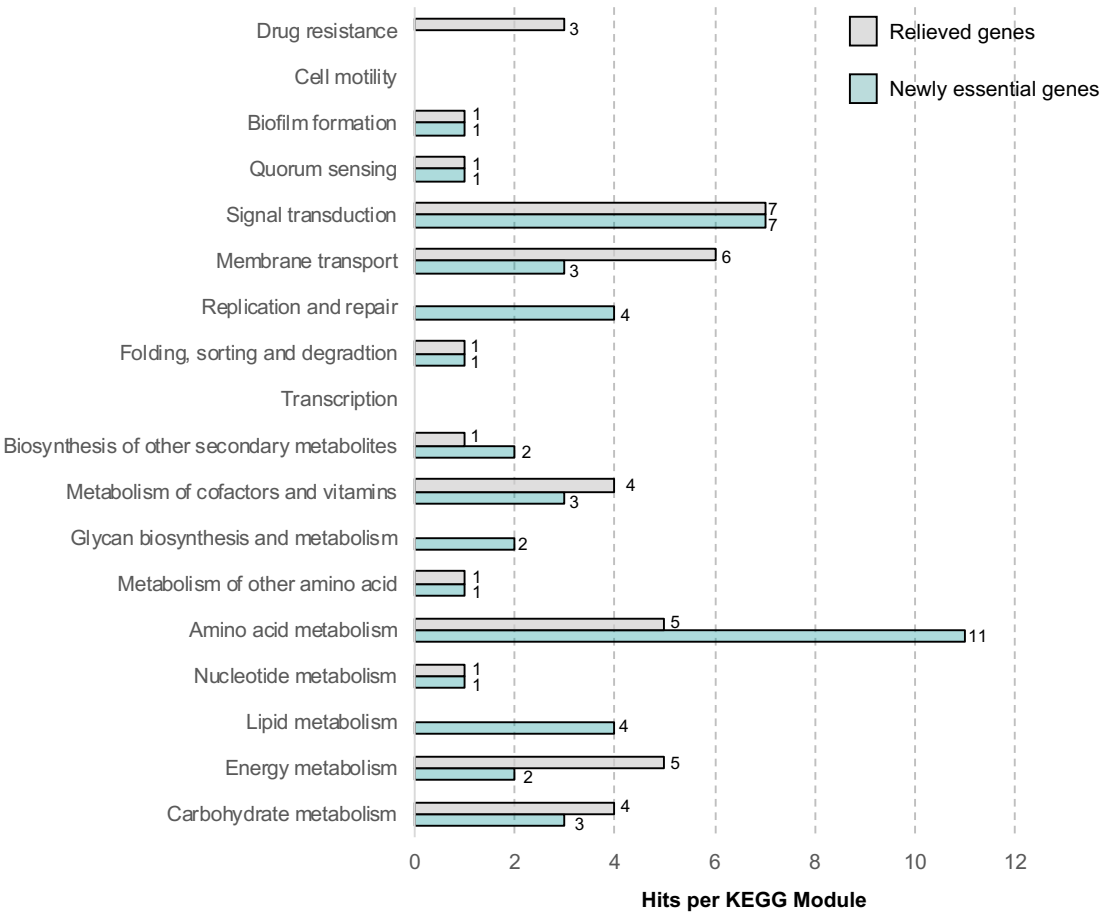
1210

1211

1212
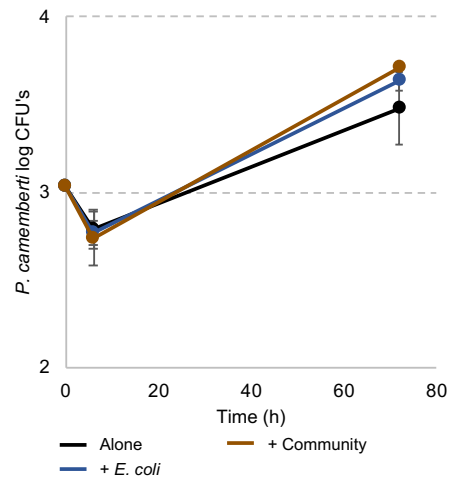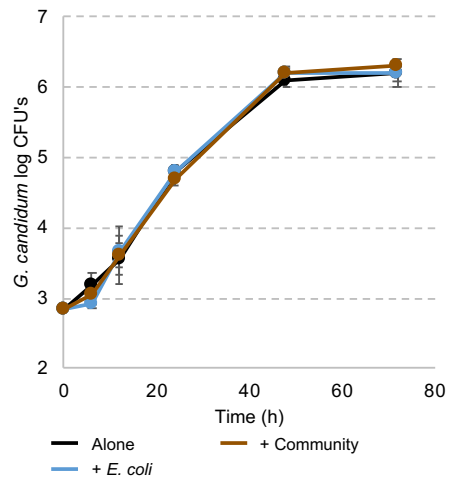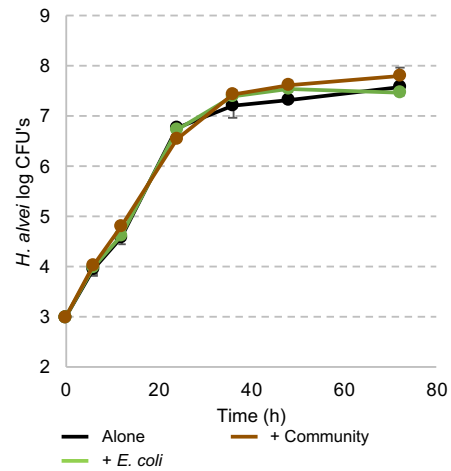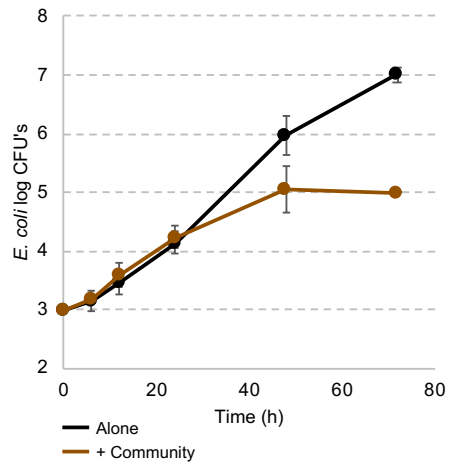
Supplementary figure 3: *E. coli* and community member growth curves alone, in pairwise conditions or during community growth. Each graph represents the growth over time of *E. coli*, *H. alvei*, *G. candidum* or *P. camameberti* alone, in pairwise growth or with the community.

1216

1217
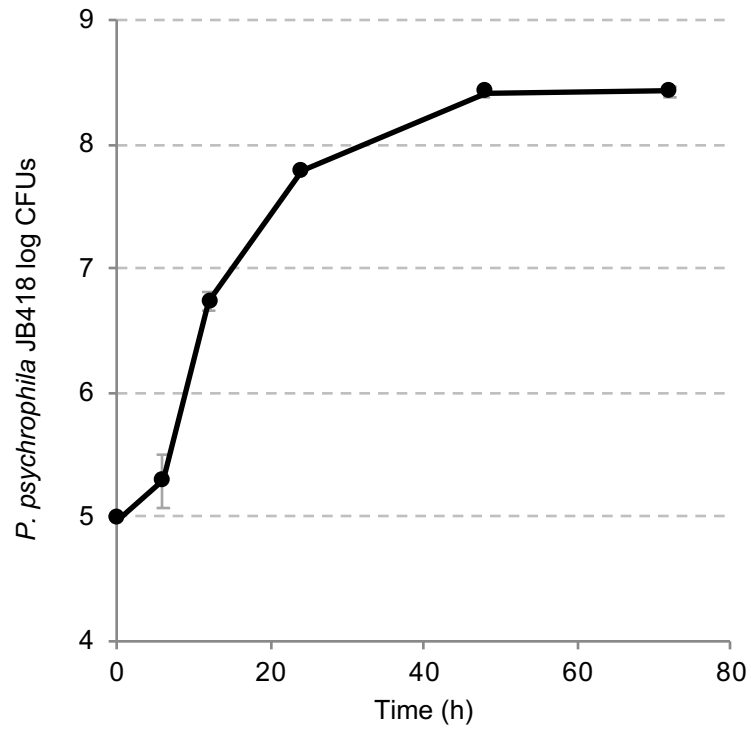
Supplementary figure 4: _P. psychrophila_ JB418 growth curve on CCA

1218
1219
1220

Supplementary figure 5: Functional analysis of newly essential genes and relieved genes for *P. psychrophila* JB418 growth in pairwise condition versus alone.