Functional Genomics of the Rapidly Replicating Bacterium Vibrio Natriegens by CRISPRi

Share Your Story

Functional genomics of the rapidly replicating bacterium *Vibrio natriegens* by CRISPRi

Henry H. Lee[1*,†], Nili Ostrov[1*], Brandon G. Wong[2], Michaela A. Gold[1], Ahmad S. Khalil[2,3], George M. Church[1,3,†]

Author affiliations

[1] Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

[2] Department of Biomedical Engineering and Biological Design Center, Boston University, Boston, Massachusetts 02215, USA.

[3] Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02115, USA.

[*] Equal contribution

[†] Correspondence: gchurch@genetics.med.harvard.edu, hhlee@genetics.med.harvard.edu

Main text

The fast growing Gram-negative bacterium *Vibrio natriegens* is an attractive microbial system for molecular biology and biotechnology due to its remarkable short generation time[1,2] and metabolic prowess[3,4]. However, efforts to uncover and utilize the mechanisms underlying its rapid growth are hampered by the scarcity of functional genomic data. Here, we develop a pooled genome-wide CRISPR interference (CRISPRi) screen to identify a minimal set of genes required for rapid wild-type growth. Targeting 4,565 (99.7%) of predicted protein-coding genes, our screen uncovered core genes composed of putative essentials and growth-supporting genes which are enriched for respiratory pathways. We found 96% of core genes to be located on the larger chromosome 1, with growth-neutral duplicates of core gene located primarily on chromosome 2. Our screen also refines metabolic pathway annotations by distinguishing functional biosynthetic enzymes from those predicted based on comparative genomics. Taken together, this work provides a broadly applicable platform for high-throughput functional genomics to accelerate biological studies and engineering of *V. natriegens*.

We began by exploring laboratory culturing conditions which support rapid and consistent *V. natriegens* growth. We opted for Lysogeny Broth supplemented with 3% (w/v) final sodium chloride (LB3), which approximates ocean salinity and provides consistent growth (**Supplementary Figure 1, Supplementary Table 1**). At 37°C, we measured *V. natriegens*' generation time to be 15.2 minutes in bulk culture and 14.8 minutes using single-cell, time-lapse microscopy (**Figure 1a-b, Supplementary Movie 1, Supplementary Figure 1, Supplementary Figure 2**). Longer generation time (30.8 minutes) was observed in minimal glucose media supplemented with 2% (w/v) sodium chloride (**Supplementary Figure 1**). These generation times are comparable with other studies[3–5].

41

42 We then produced two closed circular chromosomes of 3.24 Mb (chromosome 1; chr1) and 1.92

43 Mb (chromosome 2; chr2) by *de novo* sequence assembly, followed by annotation using RAST[7]

44 (**RefSeq NZ_CP009977-8, Figure 1c, Supplementary Table 2, Supplementary Table 3,**

45 **Methods**)[6]. We identified 4,578 open reading frames (ORFs), of which ~63% reside on chr1 and

46 ~37% on chr2. Consistent with the broad metabolic capacity described for *Vibrio* spp.[8], nearly

47 half of all annotated ORFs are involved in carbohydrates, RNA and protein metabolism (**Figure**

48 **1d**). *V. natriegens* genome carries 11 rRNA operons and 129 tRNA genes, more than *V. cholerae*

49 N16961 (8 and 98, respectively)[9] or *E. coli* MG1655 (7 and 87, respectively)[10]. We further

50 identified 36,599 putative methylated adenine residues at GATC motifs. Finally, we mapped the

51 origin and terminus of each chromosome by quantifying sequencing coverage of actively

52 dividing cells[11], and found a single bidirectional origin on each chromosome (**Figure 1e**) which

53 displayed high sequence similarity to other *Vibrio* spp. (**Supplementary Figure 3**)[9].

54

55 We next sought to develop genetic tools to enable high-throughput assessment of gene

56 functionality. While transposon mutagenesis is commonly used to generate single gene mutant

57 libraries[12] , low insertion efficiency prohibited scalable saturation mutagenesis in *V. natriegens*

58 (**Supplementary Table 8, Supplementary Discussion**). Instead, we turned to CRISPR/Cas9

59 which has been demonstrated in well-studied model organisms[13–16] but has not been scaled for

60 functional genomics in emerging microbial systems.

61

62 To test CRISPR/Cas9 functionality, we created a reporter *V. natriegens* strain expressing GFP,

63 and tested the *Streptococcus pyogenes* Cas9 nuclease with or without GFP-targeting guide RNA

64 (gRNA) (plasmid pRSF-paraB-spCas9-gRNA, **Supplementary Figures 5-7, Supplementary**

65 **Table 4**). No transformants were detected following electroporation of Cas9 nuclease with a

66 GFP-targeting gRNA, even without induction of Cas9 expression (**Supplementary Figure 9**).

67 This result indicates significant toxicity due to double-stranded breaks and undetectable levels of

68 non-homologous end joining (NHEJ). We thus concluded that the CRISPR/Cas9 system alone

69 cannot be used for efficient generation of mutants by NHEJ.

70

71 Transcriptional inhibition of gene expression by CRISPR inhibition (CRISPRi) relies on the

72 nuclease-deficient variant dCas9 to block RNA polymerase activity[15]. Based on design rules

73 elucidated in other model bacteria[14,15], we designed guides targeting the template and non-

74 template strand proximal to GFP transcriptional start site and measured fluorescence in the

75 presence of dCas9 (plasmid pRSF-paraB-dCas9-gRNA, **Supplementary Table 4**). Consistent

76 with previous studies, we observed greater inhibition when targeting the non-template strand

77 (>13-fold, **Figure 2a**). As with Cas9 nuclease, induction of dCas9 was not required for activity

78 and led to longer lag phase indicative of marginal toxicity (**Supplementary Figure 9**). We thus

79 opted to use the presence and absence of the dCas9 plasmid in lieu of induction for subsequent

80 experiments.

81

82   For high-throughput screening of single-gene inhibition by CRISPRi, we transformed a pool of
83   gRNA plasmids such that each cell carries dCas9 and a single gRNA targeting inhibition of one
84   protein coding gene. The library was grown as a competitive culture and the abundance of each
85   gRNA was determined by sequencing of gRNA cassettes. Since the abundance of gRNA reflects
86   the fitness cost of reduced gene expression, we expect rapid depletion of gRNAs which target
87   genes that, when inhibited, impair cellular growth.

88

89   We initially used a small library of five gRNAs to test the feasibility of a pooled CRISPRi
90   approach. gRNAs were selected for inhibition of one putative essential gene ($lptF_{Vn}$, a homolog
91   of an essential *E. coli* gene), one putative non-essential *V. natriegens* gene ($flgC_{Vn}$), and two
92   control genes absent from the genome (GFP and *E.coli* gene *lptF*). All gRNAs plasmids (vector
93   pCTX-R6K-gRNA, **Supplementary Figure 8**) were co-transformed into *V. natriegens* bearing
94   dCas9 (plasmid pdCas9-bacteria[15]). Plasmids were extracted and sequenced to determine the
95   abundance of each gRNA over three hours of competitive growth. We observed a significant
96   reduction in abundance of the gRNA targeting the putative essential $lptF_{Vn}$ gene, while all other
97   guides remained relatively unchanged (**Figure 2b, Supplementary Figure 10**). These results
98   demonstrate that gRNA abundance can be used as proxy for a gene's fitness in pooled
99   competitive growth.

100

101   Next, we scaled up our gRNA library to cover all predicted ORFs in the *V. natriegens* genome.
102   Transcriptional knockdown of a gene by CRISPRi can be achieved by targeting either its
103   transcriptional start site, which blocks initiation, or its coding region, which blocks elongation[15].
104   As transcriptional start sites have not yet been mapped in this organism, we designed gRNAs
105   targeting the non-template strand proximal to each gene's coding region. We created a library of
106   13,567 unique gRNAs to target 4,565 (99.7%) RAST-predicted protein-coding genes. Our
107   library includes up to three unique gRNAs per gene, with 88% of all guides within the first 20%
108   of the coding sequence (**Figure 2c, Supplementary Table 5**). Library screens were performed in
109   LB3, M9-Glucose, or M9-Sucrose for 8 hours (**Figure 2d**, **Supplementary Figure 11**). To
110   estimate fitness for each gene from its targeting gRNAs, we generated a beta score and statistical
111   significance using MAGeCK's maximum likelihood estimation (MLE)[17] (**Supplementary**
112   **Figure 5**). A negative beta score indicates depletion of a gRNA under the tested condition.

113

114   Using sucrose as the sole carbon source, we identified fourteen unique genes with a significant
115   negative beta scores (FDR≤0.05) (**Figure 3a, Supplementary Table 5**). These include
116   phosphoglucomutase, several membrane transport proteins, and sucrose-6-phosphate hydrolase,
117   a key component of intracellular sucrose catabolism. As only one of two annotated sucrose-6-
118   phosphate hydrolase genes was identified in our screen, we generated in-frame deletion of both
119   genes and measured each under non-competitive growth conditions. Consistent with CRISPRi
120   results, we found the sucrose-6-phosphate hydrolase located on chr1 (PEG.1381, β_{suc} = -1.6907),

121　but not the one located on chr2 (PEG.3071, β_{suc} = 0.12), was required for growth in M9-sucrose
122　media (**Figure 3b**). This demonstrates the utility of our approach for identifying functionally
123　relevant genes.
124
125　Overall, our glucose and sucrose minimal media screen yielded 143 unique genes (FDR≤0.05)
126　enriched in RAST subcategories for amino acid and vitamin biosynthesis (**Figure 3c,**
127　**Supplementary Figure 12a**). When mapped onto the respective pathways, this set can be used
128　to distinguish functional genes from a larger set of computationally assigned biosynthetic
129　enzymes[7,18]. For example, of all predicted L-leucine biosynthetic genes, only one of two
130　acetolactate synthase complexes (PEGs 2236, 2237 and 2615, 2616) and one of two
131　isopropylmalate synthases (PEGs 2243 and 3818) was identified as functional in our screen
132　(**Figure 3d**). Similar results were obtained for tryptophan and isoleucine biosynthetic pathways.
133　Moreover, we show functional enzymes are not always the most differentially expressed (**Figure**
134　**3d**, **Supplementary Figure 12b,c**). Thus, while transcriptional enrichment corroborates the
135　importance of amino acid biosynthesis pathways in minimal media (**Supplementary Table 7**),
136　expression level is not sufficient to infer gene functionality.
137
138　We next sought to identify a minimal set of genes necessary for rapid wild type growth in rich
139　media (LB3). In our first experiment, we found 646 genes with significant negative beta score
140　(**Figure 3a**). When repeated, the same screen resulted in a slightly larger number of significant
141　genes (735), though further passaging of the culture yielded only 648 significant negative scoring
142　genes. Of these, 587 (92.4%) overlap between the two experiment (**Figure 4a, Supplementary**
143　**Figure 13**). Given the high overlap, these data indicate that passaging reduces variability and
144　provides more stringent results. We thus define 587 genes identified in both experiments as 'core
145　genes' which are required for rapid growth in rich media (**Supplementary Table 6**).
146
147　We further analyzed the passaging experiments for evidence of CRISPRi escape by examining
148　100 genes which were identified as significant in the first, but not the last passage
149　(**Supplementary Figure 13e, Supplementary Table 5**). We expect CRISPRi escape to result in
150　increased beta score over passages. However, the majority of these genes (73%) displayed a
151　decrease in beta score, suggesting that rather than escaping dCas9 inhibition, they likely lost
152　significance due to increase in the total number of low scoring genes (**Supplementary Figure**
153　**13f**). Of the remaining 27 genes which increase in beta score over passaging, only seven were
154　altered by more than 20% (**Supplementary Figure 14a**). For these, we found one of three gRNA
155　was repeatedly elevated in the third passage (**Supplementary Figure 14b**), indicating that
156　increase in beta score is due to an inefficient or promiscuous gRNA sequence rather than escape.
157
158　Core genes comprise 12.8% of 4,565 predicted coding genes in *V. natriegens*, and the vast
159　majority are located on chr1 (96.0%, 564 genes) (**Supplementary Table 6**). Only 16 core genes
160　(2.7%) were annotated as hypothetical proteins. We found the median expression for core genes

161   was 5.6-fold higher than non-core genes (160.71 and 28.56 median transcripts per million -
162   TPM, respectively). To assess essentiality, we compared the core set to genes previously
163   identified by high-density transposon mutagenesis in the well-studied *Vibrio cholerae* genome[19].
164   We found 278 homologous genes, corresponding to 84.4% of the 329 mappable essential genes
165   in *V. cholerae*, which we designate putative essential genes in *V. natriegens* (**Figure 4b**,
166   **Supplementary Table 6**).
167
168   Similar to other *Vibrio* spp.[20], the vast majority of putative essentials are located on chr1 (272 of
169   278 genes). Putative essentials were statistically enriched for RAST categories including 'Protein
170   metabolism' and 'Cell division and cell cycle', which include ribosomal proteins, tRNA
171   synthetases, and DNA polymerase (**Figure 4c**, **Supplementary Figure 15**). For example, 49 of
172   55 total ribosomal proteins are included in the core set, of which 44 (89.8%) are in agreement
173   with essentiality in *V. cholerae* and 37 (75.5%) are in agreement with essentiality in *E .coli*, as
174   determined by transposon mutagenesis[19] and gene deletions[21], respectively (**Figure 4d**,
175   **Supplementary Table 6**). Other notable essentials include DNA adenine methylase *dam*[22] and
176   *nadE*, a NAD+ synthase gene implicated in the conversion of chr2 from a megaplasmid into a
177   critical component of the *Vibrio* genome[23]. Putative essential genes, which were all actively
178   transcribed, showed lower beta scores and higher transcription levels than other core genes
179   (**Figure 4e-f**, **Supplementary Figure 15a**), in line with reports of essential gene expression in
180   other organisms[24].
181
182   Non-essential core genes (309 of 587 genes) were uniquely enriched for the RAST category
183   'Respiration' (**Figure 4c**, **Supplementary Figure 15b**). This set of 'growth-supporting' genes
184   include the Rnf electron transport complex, cytochrome C oxidase biogenesis enzymes, and the
185   Na+-NQR respiration complex implicated in aerobic respiration in halophilic bacteria[25]
186   (**Supplementary Table 6**). These pathways are likely responsible for the high oxygen uptake
187   and highly active electron transport chain activity[3] required to maintain rapid cell division.
188   Others, such as quorum sensing regulator *hfq*[26], may determine the rapid growth rate observed in
189   highly dilute cultures[4]. Further investigation of growth-supporting genes will refine our
190   understanding of the mechanisms underlying rapid *V. natriegens* growth.
191
192   Interestingly, we found that 14.3% of core genes (84 of 587 genes) - including Tyrosyl-tRNA
193   synthetase, heat shock proteins GroES/EL, DNA polymerase I and chaperone DnaK - share
194   identical annotation with one or more genes. Duplicates are distributed similarly between
195   essential and growth-supporting core genes (39 and 45, respectively) and most have non-
196   significant negative beta scores, suggesting they are unlikely to be functionally redundant
197   (**Supplementary Figure 16a**). Only five genes were found where both duplicates are
198   independently identified in our screen, including EF-Tu, Geranylgeranyl pyrophosphate
199   synthetase, MotA/TolQ/ExbB proton channel family protein, Ferric iron ABC transporter ATP-
200   binding protein, and a benzoquinol hydroxylase (**Supplementary Table 6**). We found 60.3% of

201 core genes with duplicate (53 of 84 genes) are located on chr1 with the duplicate located on chr2
202 (**Figure 4g**), contributing to the relatively large size of chr2 in *V. natriegens* compared to other
203 *Vibrio spp.*[20]. Duplicated genes, previously reported in *V. cholerae*[19], may represent
204 heterofunctional homologs or pseudogenes. Transcriptional profiling shows the median TPM for
205 core genes is ~17-fold higher than that of duplicates (125.6 and 7.4, respectively), suggesting
206 that expression of duplicates may be silenced under tested conditions (**Supplementary Figure**
207 **16b**). Further insight into the genesis and role of these gene duplications may benefit from
208 studies of interchromosomal interactions[27] in *Vibrio* genomes.
209
210 Interpretation of pooled CRISPRi results should consider potential caveats. First, metabolite
211 cross feeding, such as amino acid exchange, may obscure the effect of inhibition for critical
212 pathway enzymes[28]. Second, as CRISPRi relies on reducing RNA polymerase elongation,
213 inhibition of upstream genes in an operon may result in erroneous determination of essentiality
214 for downstream genes[14]. Identification of transcriptional start sites and operon structures,
215 currently unavailable for *V. natriegens*, will help to better address polar effects. Third, changes to
216 the statistical significance threshold may alter the list of functionally relevant genes. Lastly,
217 screen results may be affected by dCas9-specific toxicity, as observed for *recQ* helicase
218 (**Supplementary Discussion, Supplementary Figure 17**). Additional studies using in-frame
219 deletions are warranted to stringently evaluate the phenotype of single gene knockout.
220
221 Nevertheless, screening sensitivity and specificity may be further increased by improvements to
222 the gRNA library and dCas9 protein. While three gRNAs were used to mitigate inefficient gene
223 knockdown, we cannot exclude the possibility that inhibition levels vary between genes since the
224 knockdown efficiency for each gRNA was not measured. Future studies may utilize individually
225 optimized gRNAs to ensure equal inhibition between genes or increase the number of gRNAs
226 per gene for robust statistical inference. The library can also be expanded to target non-coding
227 elements, such as RNA or intergenic regions, to reveal functional genomic features not captured
228 by computational annotation. Additional design principles for CRISPRi libraries[16,29] may be used
229 to improve gRNA binding specificity or reduce potential sequence-specific toxicity[29].
230 Furthermore, gene dosage effects can be explored by tuning transcriptional knockdown via
231 engineering of *V. natriegens* promoters or synthetic dCas9 with reduced toxicity[30]. Overall, this
232 pooled approach can be utilized to screen for genotypes underlying any growth-linked phenotype
233 to accelerate our understanding of this organism. More broadly, we envision the use of pooled
234 genome-wide CRISPRi screens will enable rapid and high-throughput functional annotation of
235 diverse bacteria.
236
237 Taken together, we provide initial genome-wide experimental annotations for *V. natriegens*
238 which can inform targeted studies of its fast growth and metabolic capacity. These findings may
239 also provide a starting point for large-scale genome engineering. For example, core genes can be
240 used as a basis for probing the limits of codon reassignment. Furthermore, beta scores can

241 provide valuable experimental guidance for construction of a fast growing minimal genome by
242 bottom up construction or large-scale deletion[2]. With only 2.1% of putative essential genes
243 located on chr2 (**Figure 4e**, **Supplementary Table 6**), spatial distribution of core genes presents
244 intriguing opportunities for rational genome design, such as consolidating of functional genes to
245 chromosome 1 and re-purposing of chromosome 2 as an artificial chromosome. Further
246 investigation of this unique organism will facilitate its advancement as a versatile bacterial
247 system for research and biotechnology.
248

253

254 Correspondence to gchurch@genetics.med.harvard.edu, hhlee@genetics.med.harvard.edu

255

266

267 **Author contributions**
268 H.H.L. and N.O. designed, performed experiments, analyzed data, and wrote the paper. B.W. and
269 A.S.K. designed and performed single-cell microfluidics experiments, and provided input on the
270 paper. M.A.G. contributed to the electroporation experiments and formulated recovery media.
271 G.M.C. supervised the project.

272

273 **Competing financial interests**
274 H.H.L., N.O., and G.M.C. have filed patents related to this work.

Materials and Methods

<u>Strains used in this study</u>
Wild type *V. natriegens* (ATCC 14048) was used in this study. GFP reporter strain was constructed by genomic integration of a pLtetO-gfp-kanamycin cassette using the mariner transposon system described under methods.

<u>Growth Media</u>
Unless denoted, LB3, Lysogeny Broth with 3% (w/v) final NaCl, is used as standard rich media. We prepare this media by adding 20 grams of NaCl to 25 grams of LB Broth - Miller (Fisher BP9723-500). Media are formulated according to manufacturer instructions and supplemented with 1.5% (w/v) final Ocean Salts (Aquarium System, Inc.) to make high salt versions of Brain Heart Infusion (BHIO), Nutrient Broth (NBO), and Lysogeny Broth (LBO). No additional salts were added to Marine Broth (MB). Minimal M9 media was prepared according to manufacturer instruction. For culturing *V. natriegens*, 2% (w/v) NaCl was added to M9 (in addition to the pre-included 0.5 g/L NaCl) based on the screen of NaCl concentrations in rich media. Carbon sources were added as indicated to 0.4% (v/v) final. SOC3 media is composed of 5 grams of yeast extract, 20 grams tryptone, 30 grams sodium chloride, 2.4 grams magnesium sulfate, and 0.4% (w/v) final glucose. Antibiotic concentrations used for plasmid selection in *V. natriegens*: ampicillin/carbenicillin 100 µg/ml, kanamycin 75 µg/ml, chloramphenicol 5µg/ml, spectinomycin 100 µg/ml. *E. coli* experiments were performed in standard LB media and M9.

<u>Culturing and glycerol stock conditions</u>
An inoculation of -80°C frozen stock of *V. natriegens* can reach stationary phase after 5 hours when incubated at 37°C. Prolonged overnight culturing (>15 hours) at 37°C may lead to an extended lag phase upon subculturing. Routine overnight culturing of *V. natriegens* was performed for 8-15 hours at 37°C or 12-24 hours at room temperature. Unless otherwise indicated, *E. coli* cells used in this study were K-12 subtype MG1655 cultured overnight (>10 hours) at 37°C. *V. cholerae* O395 was cultured overnight (>10 hours) in LB at 30°C or 37°C in a rotator drum at 150rpm. To prepare *V. natriegens* cells for -80°C storage, an overnight culture of *V. natriegens* is washed in fresh media before storing in glycerol. Cultures were centrifuged for 1 minute at 20,000 rcf and the supernatant was removed. The cell pellet was resuspended in fresh LB3 media and glycerol was added to 20% final concentration. The stock is quickly vortexed and stored at -80°C. Bacterial glycerol stocks stored in this manner are viable for at least 5 years.

<u>Bulk measurements of generation time</u>
Growth was measured by kinetic growth monitoring (Biotek H1, H4, or Eon plate reader) in 96-well plates with continuous orbital shaking and optical density (OD) measurement at 600nm taken every 2 minutes. Overnight cells were washed once in fresh growth media, then subcultured by at least 1:100 dilution. To assay *V. natriegens* growth in different rich media,

315 cells were cultured overnight from frozen stock into the respective media. To assay *V. natriegens*
316 and *E. coli* growth in minimal media, cells were cultured overnight in LB3 and LB respectively,
317 and subcultured in the appropriate test media. Generation times were calculated by linear
318 regression of the log-transformed OD across at least 3 data points when growth was in
319 exponential phase. To avoid specious determination of growth rates due to measurement noise,
320 the minimal OD considered for analysis was maximized and the ODs were smoothed with a
321 moving average window of 3 data points for conditions that were challenging for growth.
322 Maximal growth rates were computed from generation times. Apparent lag times were estimated
323 with a fitted model-free smoothed spline using "grofit" R package.

324

325 Construction of microfluidics device or single cell imaging

326 Microfluidic devices were used as tools to measure and compare growth rates of *E. coli* and *V.*
327 *natriegens*. In these devices, cells are grown in monolayer and segmented/tracked in high
328 temporal resolution using time-lapse microscopy. The cells are constricted for imaging using
329 previously described Tesla microchemostat device designs, in which cell traps have heights that
330 match the diameters of the cells, minimizing movement and restricting growth in a monolayer [31].
331 Different trapping heights of 0.8 μm and 1.1 μm were used for *E. coli* and *V. natriegens,*
332 respectively. Microfluidic devices were fabricated with polydimethylsiloxane (PDMS/Sylgard
333 184, Dow Corning) using standard soft lithographic methods. Briefly, microfluidic devices were
334 fabricated by reverse molding from a silicon wafer patterned with two layers of photoresist (one
335 for the cell trap, another for flow channels). First, the cell trap layer was fabricated by spin
336 coating SU-8 2 (MicroChem Corp.) negative resist at 7000 RPM and 6800 RPM for *E. coli* and
337 *V. natriegens,* respectively, and patterned using a high resolution photomask (CAD/Art Services,
338 Inc.). Next, AZ4620 positive photoresist (Capitol Scientific, Inc.) was spun onto the silicon
339 wafer and aligned with another photomask for fabrication of ~8 μm tall flow channels (same for
340 both organisms). Reverse-molded PDMS devices were punched and bonded to No. 1.5 glass
341 coverslips (Fisher Scientific).

342

343 Time-lapse microscopy and image analysis

344 Cells were diluted down to 0.1 OD$_{600}$ from an overnight culture at optimal growth conditions and
345 allowed to grow for an hour in the corresponding media conditions (e.g. temperature, salt
346 concentration) before loading onto the device. Next, cells were loaded and grown on the device
347 in the corresponding environmental conditions until the cell trap chambers filled. Temperature
348 was maintained with a Controlled Environment Microscope Incubator (Nikon Instruments, Inc.).
349 Media flow on device was maintained by a constant pressure of 5 psi over the course of the
350 experiment after cell loading. During the experiment, phase contrast images were acquired every
351 minute with a 100x objective (Plan Apo Lambda 100X, NA 1.45) using an Eclipse Ti-E inverted
352 microscope (Nikon Instruments, Inc.), equipped with the "Perfect Focus" system, a motorized
353 stage, and a Clara-E charge-coupled device (CCD) camera (Andor Technology). After the
354 experiment, individual cells were segmented from the image time course using custom

355 MATLAB (Mathworks, Natick, MA) software. Doubling time of cells was scored well before
356 the density of the chamber impacted tracking and growth of cells. Results from repeat
357 experiments on different days and devices were consistent (Data not shown).
358
359 <u>Genome sequencing by Pacific Bioscience sequencing, *de novo* assembly, and annotation</u>
360 *V. natriegens* (ATCC 14048) was cultured for 24 hours at 30°C in Nutrient Broth with 1.5%
361 NaCl according to ATCC instructions. Genomic DNA was purified (Qiagen Puregene
362 Yeast/Bact. Kit B) and sequenced on a Single Molecule Real Time (SMRT) Pacific Biosciences
363 RS II system (University of Massachusetts Medical School Deep Sequencing Core) using 120
364 minute movies on 3 SMRTCells. SMRTanalysis v2.1 on Amazon Web Services was used to
365 process and assemble the sequencing data. The mean read length, after default quality filtering,
366 was 4,407bp. HGAP3 with default parameters was used to assemble the reads which yielded 2
367 contigs. The contigs were visualized with Gepard and manually closed. The two closed
368 chromosomes annotated using RAST under ID 691.12. The annotated genome is deposited in
369 NCBI under Biosample SAMN03178087, GenBank CP009977-8, RefSeq NZ_CP009977-8.
370 Base modification detection was performed on SMRTanalysis v2.1 with default setting and the
371 closed genome as reference. Codon usage was calculated using EMBOSS cusp.
372
373 <u>Mapping chromosomal origins and termini by Oxford Nanopore sequencing</u>
374 *V. natriegens* was cultured in LB3 and *E. coli* was cultured in LB. Both cultures were grown
375 overnight at 37°C. For stationary phase samples, 1mL of each culture was collected for genomic
376 DNA extraction. For exponential phase samples, each culture was subcultured and grown to
377 OD_{600} ~0.4 and 10mL of each was collected for genomic DNA extraction. Genomic DNA was
378 purified (Qiagen Puregene Yeast/Bact. Kit B). To maximize read length, ~1µg of genomic DNA
379 for each sample was used as input. 1D sequencing libraries were prepared, barcoded (SQK-
380 RAD002 and SQK-RBK001), and sequenced on the MinION with SpotON R9.4 flow cells for
381 48 hours. Cloud base-calling and sample demultiplexing was performed on Metrichor 1.4.5 and
382 FASTQ files prepared from FAST5 HDF files with a custom python script. Sequences were
383 aligned to the reference genome using GraphMap 0.5.1. Coverage was computed with bedtools
384 2.26.0 and GC-skew computed using the iRep package.
385
386 <u>Transcriptome profiling</u>
387 Triplicate *V. natriegens* cultures were grown overnight from -80°C stocks for each condition to
388 be assayed: 30°C in LB3, 37°C in LB3 and 37°C in M9 high-salt media supplemented with 2%
389 (w/v) final sodium chloride and 0.4% (w/v) glucose. Each culture was subcultured in the desired
390 conditions and grown to exponential phase (OD_{600} 0.3-0.6). To collect RNA, 10mL of each
391 culture was stabilized with Qiagen RNAprotect Bacteria Reagent and frozen at -80°C. RNA
392 extraction was performed with Qiagen RNeasy Mini Kit and rRNA depleted with Illumina Ribo-
393 Zero rRNA Removal Kit (Bacteria). Samples were spot-checked for RNA sample quality on an
394 Agilent 2100 RNA 6000 Nano Kit to ensure that the RNA Integrity Number (RIN) was above 9.

395   Sequencing libraries were prepared with the NEXTflex Rapid Directional qRNA-Seq Kit. Each
396   sample was barcoded and amplified with cycle-limited real-time PCR with KAPA SYBR FAST.
397   Resulting libraries were sequenced with MiSeq v3 150 to obtain paired end reads.
398   Sequences were trimmed with cutadapt. Transcripts were quantified with Salmon 0.8.1 and
399   counts were summarized with tximport for differential expression analysis with DESeq2.

400

401   <u>Construction and analysis of transposon mutant libraries</u>
402   To facilitate transposon mutagenesis, we engineered a suicide mariner-based transposon vector
403   modified for insertion mapping by high-throughput sequencing [12,32]. Our conjugative suicide
404   mariner transposon plasmid was propagated in BW29427, an *E. coli* with diaminopimelic acid
405   (DAP) auxotrophy. BW29427 growth requires 300μM of DAP even when cultured in LB.
406   Importantly, BW29427 does not grow in the absence of DAP, which simplifies counterselection
407   of this host strain following biparental mating with *V. natriegens*. For conjugation from *E. coli* to
408   *V. natriegens*, 24mL of each strain was grown to OD 0.4, spun down, resuspended and plated on
409   LB2 plates (Lysogeny Broth with 2% (w/v) final of sodium chloride) and incubated at 37°C for
410   60 minutes (see **Supplementary Figure 4** for optimization of conjugation conditions). This
411   conjugation time was chosen to minimize clonal amplification, based on optimization
412   experiments using 100uL of each strain. The cells are recovered from the plate in 1mL of LB3
413   media. The resulting cell resuspension is washed once in fresh LB3, resuspended to a final
414   volume of 1mL, and plated on 245mm x 245mm kanamycin selective plates (Corning). Plates
415   were incubated at 30°C for 12 hours to allow the formation of *V. natriegens* colonies. Colonies
416   were scraped from each plate with 3mL of LB3, gently vortexed, and stored as glycerol stock as
417   previously described. No colonies were detected in control experiments with only BW29427
418   donor cells. A similar protocol was used to generate an *E. coli* transposon mutant library, except
419   LB was used as the media at all steps.

420

421   For analysis of transposon mutant library, genomic DNA was extracted (Qiagen DNeasy Blood
422   & Tissue Kit), and digested with MmeI. To enrich for the fragment corresponding to the
423   kanamycin transposon fragment, the digested genomic DNA was electrophoresed on a 1% TAE
424   gel and an area of the gel corresponding to approximately 1.2kb was extracted. The resulting
425   DNA fragment was sticky-end ligated to an adapter. PCR was used to selectively amplify the
426   region around the transposon mosaic end and to add the required Illumina adapters. These
427   amplicons were sequenced 1x50bp on a MiSeq. Since properly prepared amplicons contain 16 or
428   17bp of genomic DNA and 32 or 33bp of the ligated adapter, only those sequencing reads with
429   the presence of the adapter were further analyzed. All adapters were trimming and the resulting
430   genomic DNA sequences were aligned to the reference genome with Bowtie. Statistical
431   enrichment of RAST categories were computed with the hypergeometric test and resulting p-
432   values were adjusted with Benjamini-Hochberg correction. Transposon mutagenesis data is
433   provided in **Supplementary Table 8**.

434

435 For the *E. coli* Himar1 mutant library, we isolated 1.1×10^6 transconjugants, prepared Tn-Seq
436 fragments as previously described, and analyzed by MiSeq [12]. We obtained 6.9×10^6 total reads,
437 of which 1.6% mapped to the transposon plasmid; 98.3% of filtered reads were mapped to the
438 genome. These insertions represent 107,723 unique positions, where >10 unique insertions were
439 present in 3,169 out of 4,917 features. For the *V. natriegens* Himar1 mutant library we isolated
440 8.6×10^5 mutants. We obtained 6.1×10^6 reads, of which 36.4% mapped to the transposon plasmid;
441 97.2% of filtered reads were mapped to the genome. These insertions represent 4,530 unique
442 positions, proportionally distributed between the two chromosomes where >1 unique insertions
443 were found in 2,357 out of 4,940 features.

444

445 Electroporation protocol for DNA transformation of *V. natriegens*
446 An overnight *V. natriegens* culture was pelleted, washed once in fresh media, and diluted 1:100
447 into growth media. Cells were harvested at $OD_{600}\sim0.4$ (1 hour growth when incubated at 37°C at
448 225rpm) and pelleted by centrifugation at 3500rpm for 5min at 4°C. The pellet is washed three
449 times using 1ml of cold 1M sorbitol and centrifuged at 20,000rcf for 1 minute at 4°C. The final
450 cell pellet was resuspended in 1M sorbitol at a 200-fold concentrate of the initial culture. For
451 long term storage, the concentrated competent cells were aliquoted in 50μL shots in pre-chilled
452 tubes, snap frozen in dry ice and ethanol, and stored in -80°C for future use. To transform, 50ng
453 of plasmid DNA was added to 50μL of concentrated cells in 0.1mm cuvettes and electroporated
454 using Bio-Rad Gene Pulser electroporator at 0.4kV, 1kΩ, 25μF and recovered in 1mL LB3 or
455 SOC3 media for 45 minutes at 37°C at 225rpm, and plated on selective media. Plates were
456 incubated at least 6 hours at 37°C or at least 12 hours at room temperature.

457

458 Plasmid construction
459 Routine cloning was performed by PCR of desired DNA fragments, assembly with NEB Gibson
460 Assembly or NEBuilder HiFi DNA Assembly, and propagation in *E. coli* unless otherwise
461 indicated. See **Supplementary Table 4** for plasmids used in this work. Plasmid maps are
462 provided in **Supplementary Figure 5**. We used pRSF for the majority of our work since it
463 carries all of its own replication machinery and should be minimally dependent on host factors.
464 For transformation optimizations, we constructed pRSF-pLtetO-gfp which constitutively
465 expresses GFP due to the absence of the tetR repressor in both *E. coli* and *V. natriegens*. We
466 engineered pCTX-R6K shuttle plasmid by fusing the pCTX-Km replicon, comprised of genes
467 RstR and RstA, with the pir-dependent conditional replicon, R6k. After electroporation of this
468 shuttle plasmid into *V. natriegens*, we were able to extract plasmid from a standard miniprep,
469 demonstrating plasmid replication. To construct the conjugative suicide mariner transposon, we
470 replaced the Tn5 transposase and Tn5 mosaic ends in pBAM1 with the mariner C9 transposase
471 and the mariner mosaic ends from pTnFGL3[33]. Our payload, the transposon DNA, consisted
472 solely of the minimal kanamycin resistance gene required for transconjugant selection. We next
473 performed site-directed mutagenesis on both transposon mosaic ends to introduce an MmeI cut-
474 site, producing the plasmid pMarC9 which is also based on the pir-dependent conditional

475 replicon, R6k. We also constructed a transposon plasmid capable of integrating a constitutively
476 expressing GFP cassette in the genome by inserting pLtetO-GFP with either kanamycin or
477 spectinomycin in the transposon DNA. All plasmids carrying the R6k origin was found to
478 replicate only in either BW29427 or EC100D pir$^+$/pir-116 *E. coli* cells. Induction systems were
479 cloned onto the pRSF backbone. For CRISPR/Cas9 experiments, a single RSF1010 plasmid
480 carried both *Streptococcus pyogenes* Cas9 and guide RNA. dCas9 was cloned under the control
481 of *E. coli* arabinose induction genes and the guide RNA under control of the constitutive J23100
482 promoter.
483
484 <u>DNA yield</u>
485 pRSF-pLtetO-gfp was transformed via electroporation into *E. coli* MG1655 and *V. natriegens*. *E.*
486 *coli* plates were incubated at 37°C and *V. natriegens* were incubated at room temperature for an
487 equivalent time to yield approximately similar colony sizes. Three colonies from each plate was
488 picked and grown for 5 hours at 37°C in 3mL of selective liquid culture (LB for *E. coli* and LB3
489 for *V. natriegens*) at 225rpm. Plasmid DNA was extracted from 3mL of culture (Qiagen Plasmid
490 Miniprep Kit).
491
492 <u>CTX vibriophage infection</u>
493 *V. cholerae* O395 carrying the replicative form of CTX, pCTX-Km , was cultured overnight in
494 LB without selection in a rotator drum at 150rpm at 30°C. Virions were purified from cell-free
495 supernatant (0.22μm filtered) of overnight cultures. Replicative forms were extracted from the
496 cells by standard miniprep (Qiagen). To test infectivity of the virions, naive *V. cholerae* O395
497 and *V. natriegens* were subcultured 1:1000 in LB and LB3 respectively and mixed gently with
498 ~10^6 virions. After static incubation for 30 minutes at 30°C, the mixture was plated on selective
499 media and incubated overnight for colony formation. Replicative forms were electroporated into
500 host strains using described protocols.
501
502 <u>Targeted gene perturbation by Cas9</u>
503 All Cas9 experiments were performed using a single pRSF plasmid (pRSF-paraB-Cas9-gRNA)
504 carrying Cas9 gene under the control of arabinose promoter, with or without GFP-targeting guide
505 RNA. All plasmids carry carbenicillin selective marker. Wild-type *V. natriegens* or strain
506 carrying genomically integrated GFP construct were grown at 37°C overnight (LB3 or
507 LB3+100μg/ml kanamycin, respectively) and transformed with 50ng of plasmid DNA using the
508 optimized transformation protocol described above. Following 1-hour recovery in LB3 at 37°C,
509 cells were plated on LB3+100μg/mL carbenicillin plates and incubated overnight at 37°C. No
510 arabinose induction was used for Cas9 experiments, as we observed low level of baseline
511 expression using arabinose promoter.
512
513 <u>Repression of chromosomally-encoded GFP with dCas9</u>

514 We transformed this engineered *V. natriegens* strain with a CRISPRi plasmid (pRSF-paraB-
515 dCas9-gRNA) carrying dCas9 under arabinose promoter and gRNA targeting GFP. To test the
516 repression of the chromosomally-encoded GFP with CRISPRi, we subcultured an overnight
517 cultures 1:1000 in fresh media supplemented with or without 1mM arabinose. We kinetically
518 measured OD_{600} and fluorescence of each culture over 12 hours in a microplate with orbital
519 shaking at 37°C (BioTek H1 or H4). Under these conditions, all cultures grew equivalently by
520 OD_{600}.
521
522 Pooled CRISPRi screen - five-member gRNA library
523 We used dCas9 (plasmid pdCas9-bacteria was a gift from Stanley Qi; Addgene plasmid # 44249)
524 under the control of tetracycline promoter. Guide RNA was expressed under constitutive J23100
525 promoter (plasmid pCTX-R6K-gRNA). Five pCTX-R6K plasmids (spectinomycin selective
526 marker) each carrying a gRNA were used for targeted inhibition of the following genes: *V.
527 natriegens* targeting genes $lptF_{Vn}$ and $flgC_{Vn}$; three control targets that do not exist in the host:
528 *E.coli* gene $lptF_{Ec}$ and two GFP guides. All guides were designed to target the non-template
529 strand. An equal mix of all five plasmids, each 20ng, was co-transformed into a dCas9
530 expressing *V. natriegens* strain. The transformation was recovered in 1mL SOC3 media for 45
531 minutes at 37°C at 225rpm and plated on 245mm x 245mm plates (Corning) with appropriate
532 antibiotics. After 13 hours at 37°C, colonies were scraped in LB3. Growth competition was
533 performed by subculturing this library 1:1000 in LB3 at 37°C for 3 hours in baffled 250mL
534 flasks (Corning). At each time point, gRNA plasmid was extracted from 3mL of culture (Qiagen
535 Plasmid Miniprep Kit). Barcoded Illumina sequencing libraries were prepared by cycle-limited
536 PCR with real-time PCR and sequenced with MiSeq v3 150. Resulting sequences were trimmed
537 for the promoter and gRNA scaffold and the count of each guide sequence was first normalized
538 by the number of sequences per time point, then expressed as a fraction of the sequence before
539 growth competition.
540
541 Construction, testing, and analysis of genome-wide gRNA library
542 A custom python script was used to select gRNA sequences targeting the non-template strand of
543 each RAST predicted protein-coding gene. Starting at the 5' end of the gene, 20bp sequences
544 with a terminal Cas9 NGG motif on the reverse complement strand were selected. Up to 3 targets
545 were selected for each RAST predicted gene features; each guide sequence was prefixed with a
546 promoter and suffixed with part of the gRNA scaffold. This sequence was synthesized by the
547 OLS process (Agilent Technologies) as an oligo library. The OLS pool was amplified by cycle-
548 limited real-time PCR, and assembled into the pCTX-R6K-gRNA backbone (NEBuilder HiFi) at
549 5-fold molar excess with 18bp overlap arms. 6μL of the assembled product was mixed with
550 300μL TransforMax EC100D pir+ *E. coli* (Epicentre) and 51μL aliquots of this mix was
551 electroporated in 0.1mm cuvettes with a Bio-Rad Gene Pulser electroporator at 1.8kV, 200Ω,
552 25μF. These *E. coli* transformants were recovered in 6x 1mL SOC media for 60 minutes at 37°C
553 at 225rpm, and plated on 245mm x 245mm spectinomycin selective plates (Corning). After 13

554 hours at 37°C, ~1.4x10^6 colonies were scraped and plasmid DNA extracted (Qiagen HiSpeed
555 Plasmid Maxi).
556
557 The plasmid gRNA library was transformed into wild type *V. natriegens,* or a wild type strain
558 expressing dCas9 (plasmid pdCas9-bacteria). No inducer was added for dCas9 expression, as
559 previous results indicate leaky expression of dCas9 is sufficient for inhibition (**Supplementary**
560 **Figure 9**) Briefly, ~600ng of the plasmid library was mixed with 300µL of electrocompetent
561 cells and 53.5µL of this mix was electroporated in 0.1mm cuvettes with a Bio-Rad Gene Pulser
562 electroporator at 0.4kV, 1kΩ, 25µF. Each transformation was recovered in 1mL SOC3 media for
563 45 minutes at 37°C at 225rpm and plated on 245mm x 245mm plates (Corning) with appropriate
564 antibiotics. After 13 hours at 37°C, colonies were scraped in LB3 and stored at -80°C as library
565 master stocks. Pooled screens for control and test libraries (with or without the dCas9 plasmid)
566 were performed in duplicate, starting with an initial population of <10^7 cells in 25mL of media in
567 baffled 250mL flasks (Corning). For media experiments, plasmids were extracted after 8 hours
568 of growth at 37°C in LB3, M9+0.4% (w/v) Glucose, or M9+0.4% (w/v) Sucrose. For serial
569 passaging experiments, plasmids were extracted after 8 hours of growth at 37°C in LB3 (Passage
570 1) and the culture is then diluted 1:1000 in 25mL of fresh LB3, yielding ~10^6 cells, and grown at
571 37°C for 4 hours (Passage 2). Subsequent passaging were performed in 4 hour increments.
572 Plasmids were extracted from >3mL of culture (Qiagen Plasmid Miniprep Kit). Barcoded
573 Illumina sequencing libraries were prepared by cycle-limited real-time PCR and sequenced with
574 NextSeq v2 High Output 500/550. Resulting sequences were trimmed for the promoter and 5'-
575 end of the gRNA scaffold. Sequencing was used to verify coverage of the gRNA library.
576 Sequencing was used to verify coverage of the gRNA library. For media experiments, the initial
577 no dCas9 and dCas9 library contains 15,539 (99.65%) and 13,579 (99.9%) of 13,587 total
578 guides, respectively. For serial passaging experiments, the initial no dCas9 and dCas9 library
579 contains 13,567 (99.94%) and 13,584 (99.98%), 13,587 total guides, respectively. After
580 sequencing all samples, the count of each guide sequence was normalized by the number of reads
581 per sequencing library to yield reads per million (RPM) per guide. The median gRNA copy
582 number for the initial no dCas9 and dCas9 media library is 70.3 and 62.3 reads per million,
583 respectively. The median gRNA copy number for the initial no dCas9 and dCas9 passaging
584 library is 70.1 and 62.1 reads per million, respectively. These counts were input into MAGeCK
585 and analyzed using the 'mle' module with default settings[17] to obtain a fitness value (Beta score
586 and FDR-adjusted p-value) for each gene. Essential genes from *E. coli* and *V. cholerae* were
587 mapped to *V. natriegens* via bactNOG or COG using eggnog-mapper based on eggNOG 4.5
588 orthology data[19,21,34]. Enrichments are calculated as a one-sided hypergeometric test assuming
589 independence between functional groups with multiple-hypothesis correction with FDR
590 (Benjamini-Hochberg).
591
592 <u>Construction of in-frame genomic knockouts by natural competency</u>

593 The endogenous tfoX gene (PEG.1425) was subcloned on to the RSF1010 backbone plasmid
594 with an *E. coli* IPTG-inducible promoter. For each gene target, a deletion cassette, consisting of
595 an antibiotic resistance marker with 500bp homology arms upstream and downstream, was
596 created by overlap extension PCR. Generation of deletion mutants are performed as previously
597 described [35]. Briefly, the strain carrying tfoX is grown overnight in 1mM final IPTG in a rotating
598 drum at 30°C, then subcultured 1:100 in 350μL of 2XIO (28g/L Ocean Salts) with 1mM IPTG
599 with >100ng of the deletion cassette. The culture is incubated statically at 30°C for 5 hours, then
600 recovered with 1mL of LBv2[2] at 30°C for 2 hours with shaking (225-250rpm). Following
601 recovery, cultures are plated on antibiotic-selective plates. Resulting transformants are screened
602 by PCR to confirm gene deletion.

603

604

605 **Data Availability**
606 Genome sequences are available in NCBI (GenBank: CP009977, CP009978; RefSeq:
607 NZ_CP009977, NZ_CP009978). Sequencing data for gRNA counts are available in NCBI SRA
608 under BioProject PRJNA511728 (SRR8369136, SRR8369137, SRR8369138, SRR8369139) and
609 transcriptome data is available via GSE126544 (GSM3603279, GSM3603280, GSM3603281,
610 GSM3603282, GSM3603283, GSM3603284). All other data are available in the Supplementary
611 Information, or by request.

612

613 **Code Availability**
614 Custom code is available at https://github.com/citizenlee/vnat_glib or will be made available
615 upon request.

616

617

Figure legends

620   **Figure 1.** Profiling of *V. natriegens* growth and genome. (a) Visualization of *V. natriegens* (Vn)
621   and *E. coli* (Ec) in a microfluidic chemostat (Cultures grown at 37°C in LB3 and LB,
622   respectively). Designated cell trapping heights were used to keep each species growing in a
623   monolayer. Scale bar 10µm (N=3 independent experiments). (b) Single-cell growth measurement
624   at 37°C (LB3 and LB, respectively). Data shown are mean±SD (N=18 *E. coli* cells and N=21 *V.
625   natriegens* cells). (c) *V. natriegens* chromosomes. From outside inward: two outer circles
626   represent protein-coding genes on the plus and minus strand, respectively. Color coding by
627   RAST annotation, as in panel d. The third circle represents G+C content relative to mean G+C
628   content of the respective chromosome, using a sliding window of 3,000 bp. tRNA and rRNA
629   genes are shown in the fourth and fifth circles, respectively. (d) Fraction of genes in each RAST
630   category on chr1 (dark colors) and chr2 (light yellow). (e) Filtered sequence coverage (black)
631   and GC-skew (green) for each chromosome, as measured for exponentially growing *V.
632   natriegens* in LB3 at 37°C. Origin (red) and terminus (blue) are denoted.

633

634   **Figure 2.** CRISPRi screen in rich and minimal media. (a) dCas9 inhibition of chromosomally-
635   integrated GFP, using gRNA targeting the template (T) or non-template (NT) strand. Inhibition
636   was observed with (black) or without induction (gray) of dCas9. Data are shown as mean±SD
637   (N=4 biological replicates). (b) Small-scale pooled CRISPRi screen. gRNAs include putative
638   essential gene $lptF_{Vn}$ (solid black line), putative neutral genes $flgC_{Vn}$ (dashed black line), GFP
639   (dashed gray line), and *E. coli* gene $lptF_{Ec}$ (solid gray line, target gene absent from genome).
640   Library was grown in LB3 at 37°C in the presence of dCas9. Fold change in gRNA abundance is
641   shown. (c) genome-wide gRNA library. Number (gray bars) and percentage (dotted line) of
642   gRNAs across open reading frames are shown. (d) Schematic overview of pooled genome-wide
643   CRISPRi screen.

644

645   **Figure 3**. CRISPRi inhibition in rich and minimal media. (a) Distribution of gene scores in rich
646   and minimal media screen are shown: blue dots represent genes with significantly negative score
647   in both rich and minimal media; red dots represent genes with significantly low score only in
648   minimal media; gray dots represent all other genes in the population. Histogram (gray) shows
649   distribution of all targeted genes. Inset: The number of genes with significantly negative core in
650   each media category. (b) Single gene knock-out of sucrose-6-phosphate hydrolase located on
651   chr1 (ΔPEG.1381) or its paralog located on chr2 (ΔPEG.3017) in LB3 (gray), M9-glucose
652   (black) or M9-sucrose (red). Data are shown as mean±SD (N=4 biological replicates) (c) Top ten
653   RAST categories enriched in genes that are uniquely depleted in minimal media (N=143 genes).
654   Br: Branched-Chain Amino Acid Biosynthesis, His: Histidine Biosynthesis, Bio: Biotin
655   biosynthesis, Trp: Tryptophan synthesis, Ch: Chorismate, Au: Auxin biosynthesis, Leu: Leucine
656   Biosynthesis, Pe: Peptide ABC transport system, Lys: Lysine Biosynthesis DAP Pathway, CoA:
657   Co-A Biosynthesis. Statistics are derived from one-tailed hypergeometric test for enrichment

assuming independence between subsystems. False discovery rate (FDR) controlled with Benjamini-Hochberg. (d) Leucine (L-leu) biosynthetic pathway starting from 2-pyruvate. Beta score (top) and change in expression between LB3 and M9 (fold change, bottom) are shown for each protein encoding genes (PEG number); functional genes identified by CRISPRi in minimal media are shown in red, other computationally assigned genes are shown in grey. Leucine biosynthetic genes numbered as follow: 1: acetoacetate synthase; 2: ketol-acid reductoisomerase; 3: dihydroxy-acid dehydratase; 4: isopropylmalate synthase; 5: 3-isopropylmalate dehydratase; 6: 3-isopropylmalate dehydrogenase; 7: aminotransferase.

Figure 4. Analysis of *V. natriegens* core genes. (a) Distribution of beta scores for all significantly depleted genes in rich media experiments (FDR≤0.05). Core genes, shown in red, display consistently negative beta score in both experiments. Negative scoring genes which were not included in core set are shown in blue (media experiment) and green (passage experiment). (b) Core genes were divided into putative essentials (yellow) and growth-supporting genes (blue). All other genes shown in gray. (c) Distribution of all targeted genes across RAST categories. Single genes, sorted by beta score, are displayed as gray lines in the respective category. Core genes shown in red. Histogram displays over-represented (filled bars) or under-represented categories (empty bars) within all core gene set (red), putative essential genes (yellow), and growth-supporting genes (blue). Statistics are derived from one-tailed hypergeometric test for enrichment and depletion assuming independence between categories. False discovery rate (FDR) controlled with Benjamini-Hochberg. N=587 core genes. Rast category names as shown in Figure 1b. (d) Ribosomal proteins essentiality as identified in *V. natriegens* (Vn) screen. Black square denotes essential gene. Empty square denone non-essential gene. Essentiality shown for *E.coli* (Ec)[21] and *V. cholerae* (Vc)[19]. LSU: Large Subunit; SSU: Small Subunit. (e) Beta score as calculated from CRISPRi screen for all gene (grey), putative essential genes (PE) and growth supporting genes (GS) are shown as a letter-value plot with median. (f) RNA-seq transcript counts as TPM (transcripts per million) for all genes (grey) and for the set of core genes, as measured at 37°C in LB3 medium. Putative essentials (PE) shown in yellow, growth-supporting genes (GS) shown in blue. Representative counts for a single biological replicate (out of three independent biological replicates) are shown as a letter-value plot with median. (g) Duplicated annotations of core genes. Core genes are shown as hash marks on chr1 (blue) and chr2 (green). Lines connect core genes with the respective duplicate annotation gene. Blue lines indicate core is on chr1 and duplicate on chr2. Green lines indicate core gene on chr2 and duplicate on chr1.
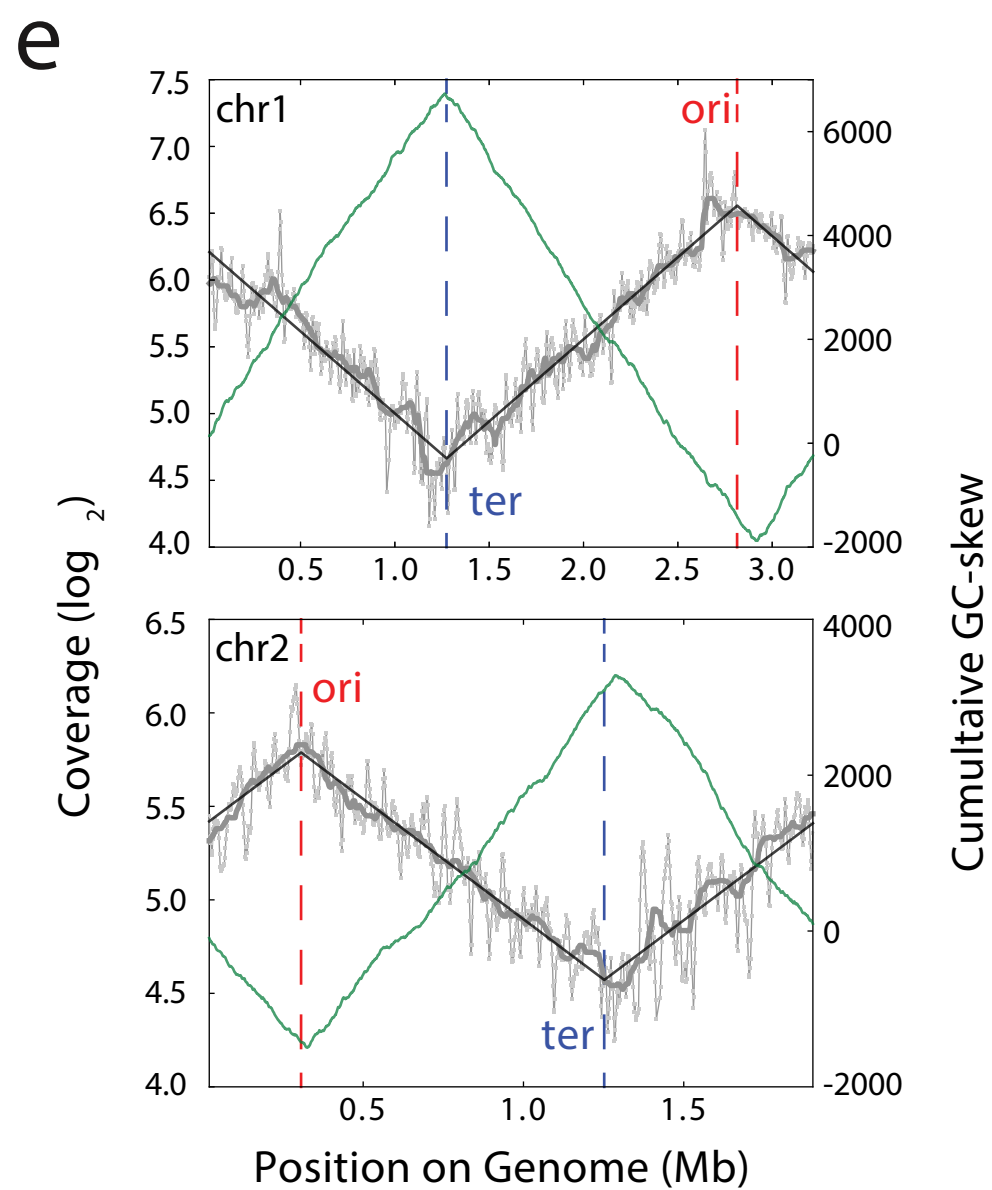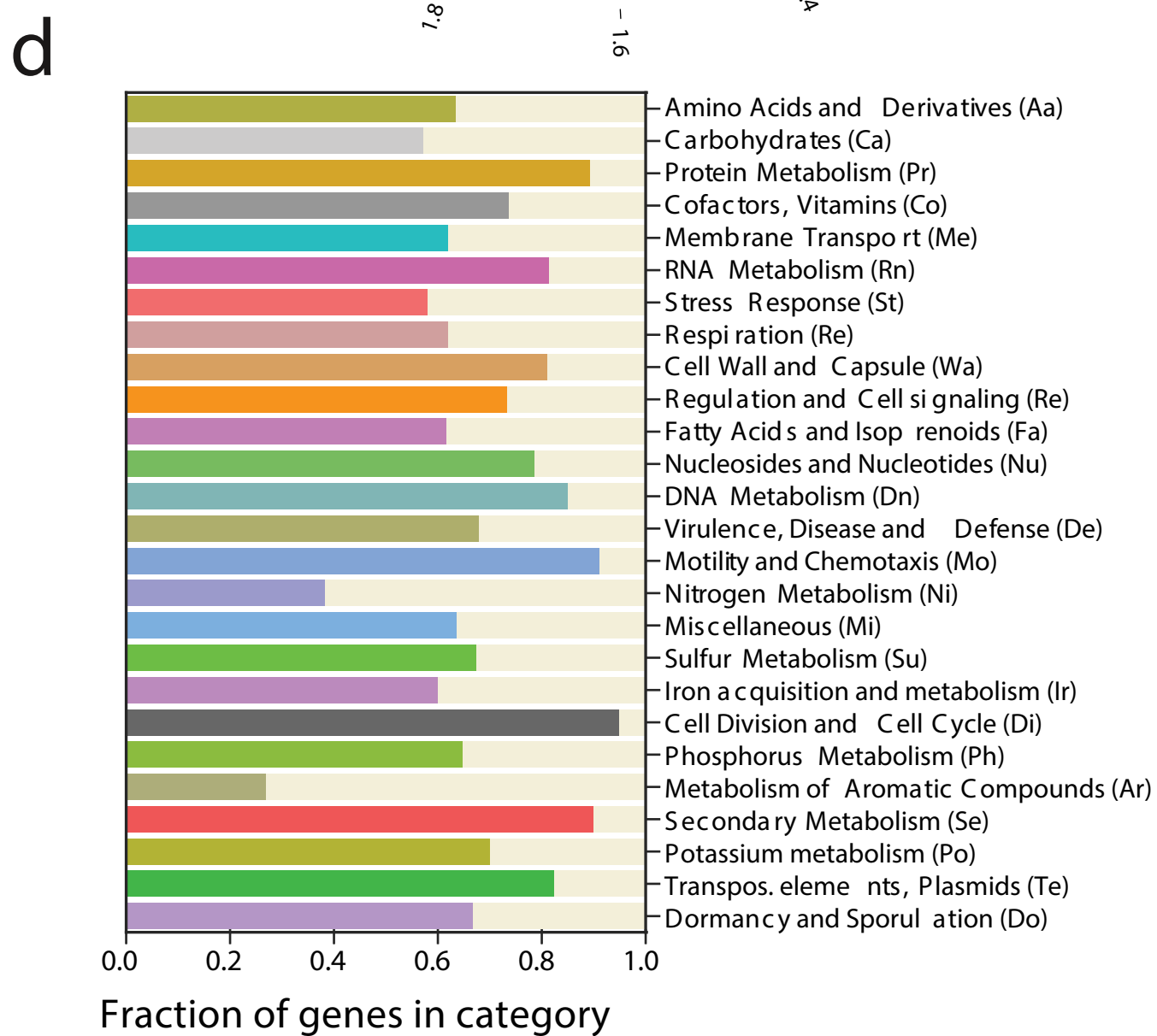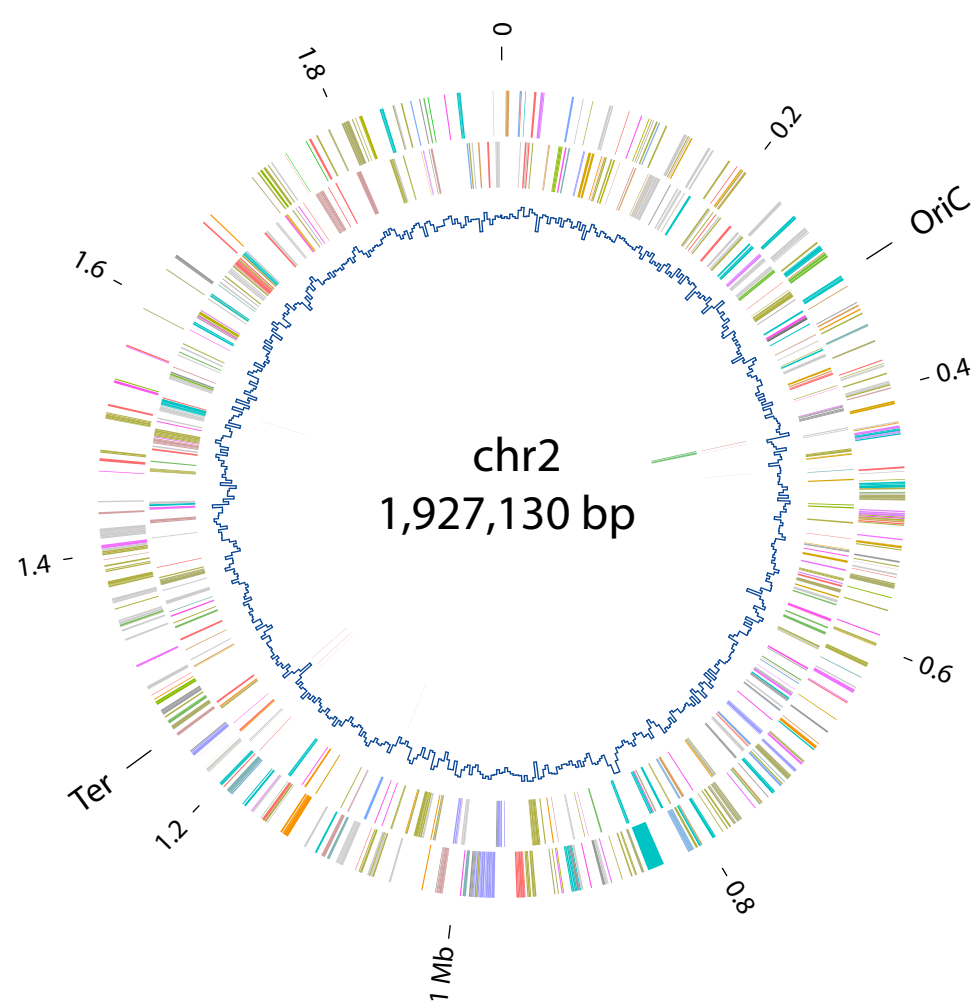
694 **References**
695

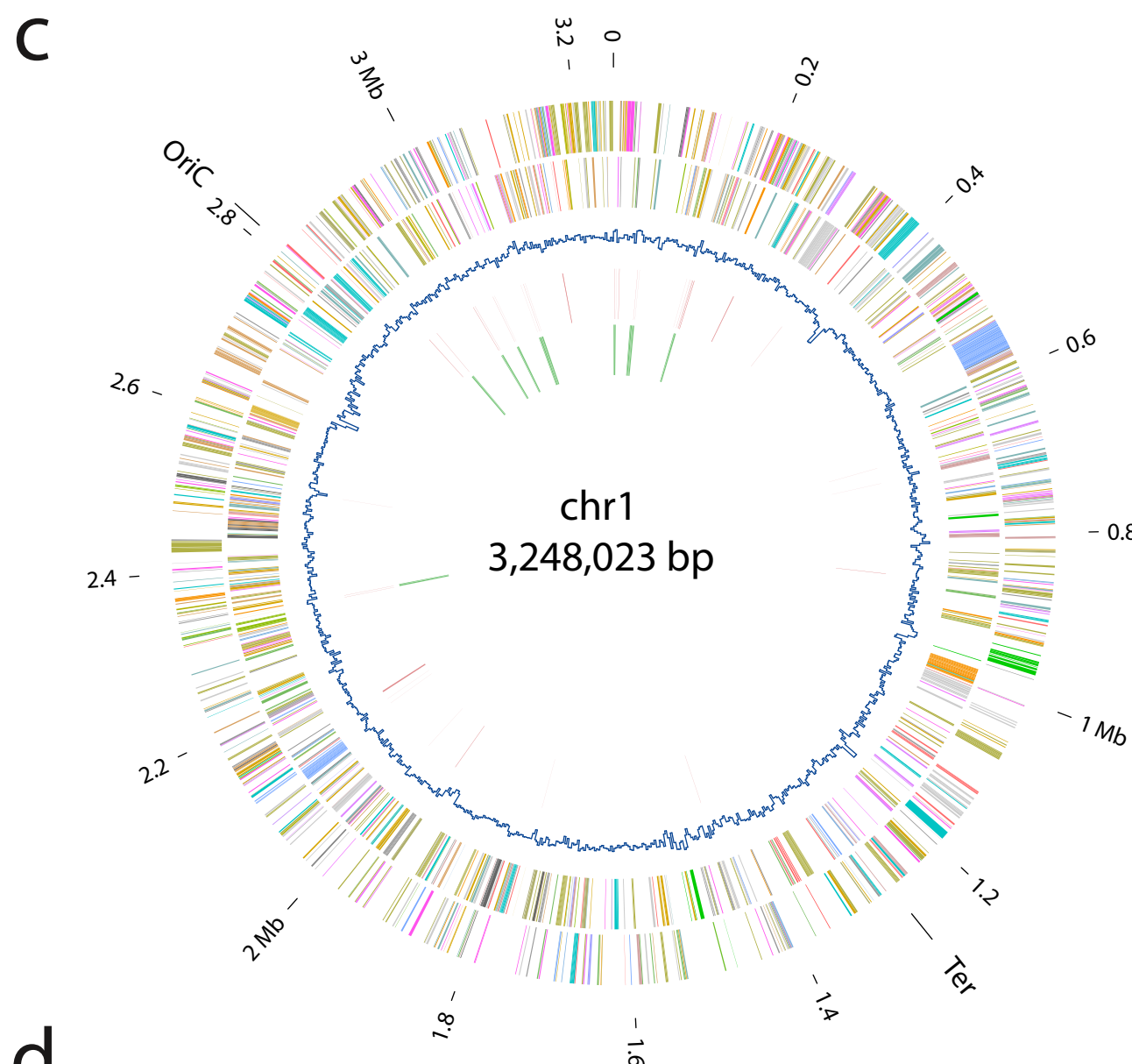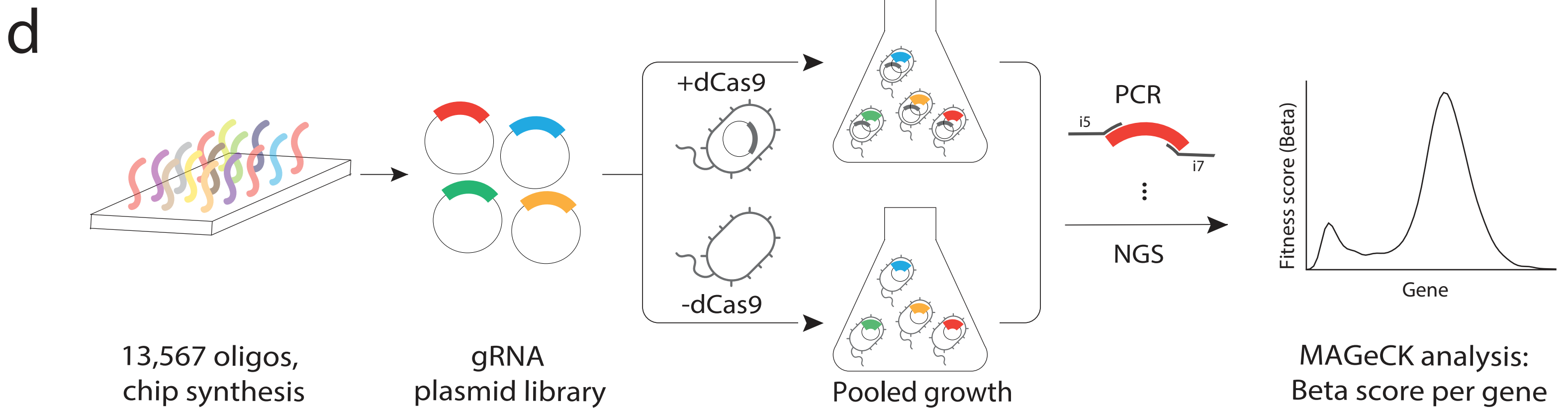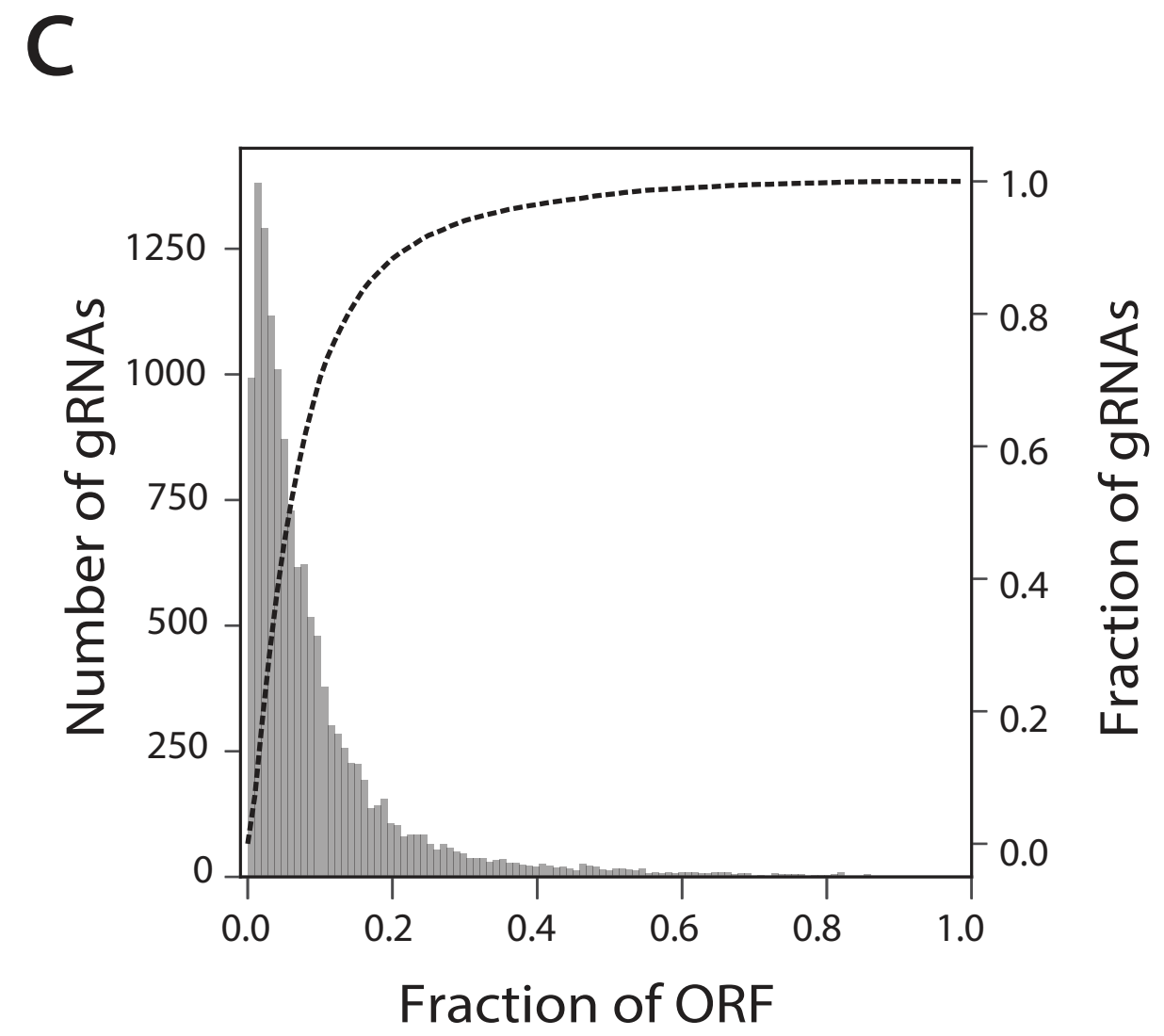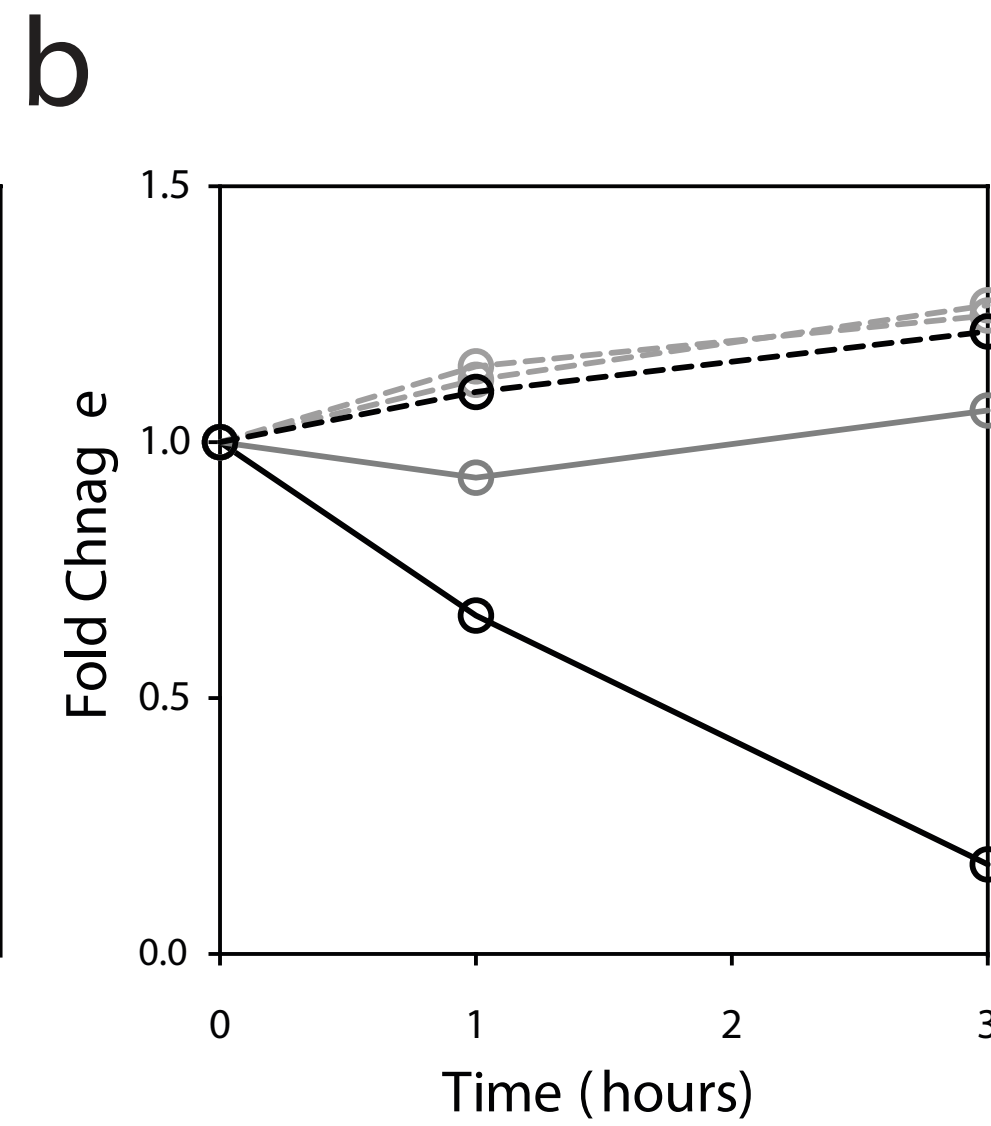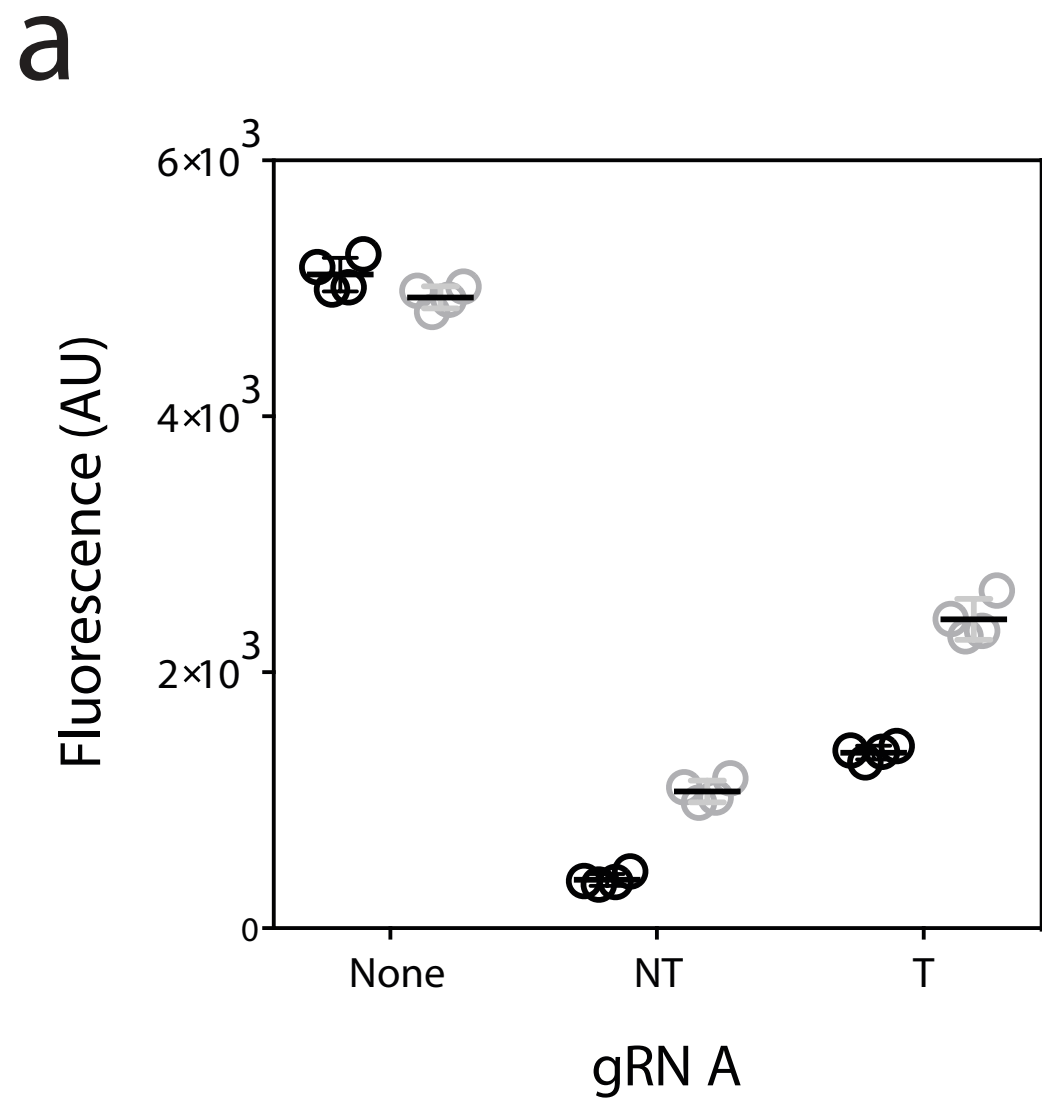696 1. Eagon, R. G. Pseudomonas natriegens, a marine bacterium with a generation time of less than 10 minutes. *J.*
697 *Bacteriol.* **83**, 736–737 (1962).

698 2. Weinstock, M. T., Hesek, E. D., Wilson, C. M. & Gibson, D. G. Vibrio natriegens as a fast-growing host for
699 molecular biology. *Nat. Methods* **13**, 849–851 (2016).

700 3. Long, C. P., Gonzalez, J. E., Cipolla, R. M. & Antoniewicz, M. R. Metabolism of the fast-growing bacterium
701 Vibrio natriegens elucidated by (13)C metabolic flux analysis. *Metab. Eng.* **44**, 191–197 (2017).

702 4. Hoffart, E. *et al.* High substrate uptake rates empower Vibrio natriegens as production host for industrial
703 biotechnology. *Appl. Environ. Microbiol.* (2017). doi:10.1128/AEM.01614-17

704 5. Aiyar, S. E., Gaal, T. & Gourse, R. L. rRNA promoter activity in the fast-growing bacterium Vibrio natriegens.
705 *J. Bacteriol.* **184**, 1349–1358 (2002).

706 6. Maida, I. *et al.* Draft Genome Sequence of the Fast-Growing Bacterium Vibrio natriegens Strain DSMZ 759.
707 *Genome Announc.* **1**, e00648–13–e00648–13 (2013).

708 7. Overbeek, R. *et al.* The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology
709 (RAST). *Nucleic Acids Res.* **42**, D206–D214 (2013).

710 8. Thompson, J. R. & Polz, M. F. Dynamics of Vibrio populations and their role in environmental nutrient
711 cycling. in *The biology of vibrios* 190–203 (American Society of Microbiology, 2006).

712 9. Heidelberg, J. F. *et al.* DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. *Nature*
713 **406**, 477–483 (2000).

714 10. Chan, P. P. & Lowe, T. M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence.
715 *Nucleic Acids Res.* **37**, D93–7 (2009).

716 11. Brown, C. T., Olm, M. R., Thomas, B. C. & Banfield, J. F. Measurement of bacterial replication rates in
717 microbial communities. *Nat. Biotechnol.* **34**, 1256–1263 (2016).

718 12. van Opijnen, T. & Camilli, A. Genome-wide fitness and genetic interactions determined by Tn-seq, a high-
719 throughput massively parallel sequencing method for microorganisms. *Curr. Protoc. Microbiol.* **Chapter 1**,
720 Unit1E.3 (2010).

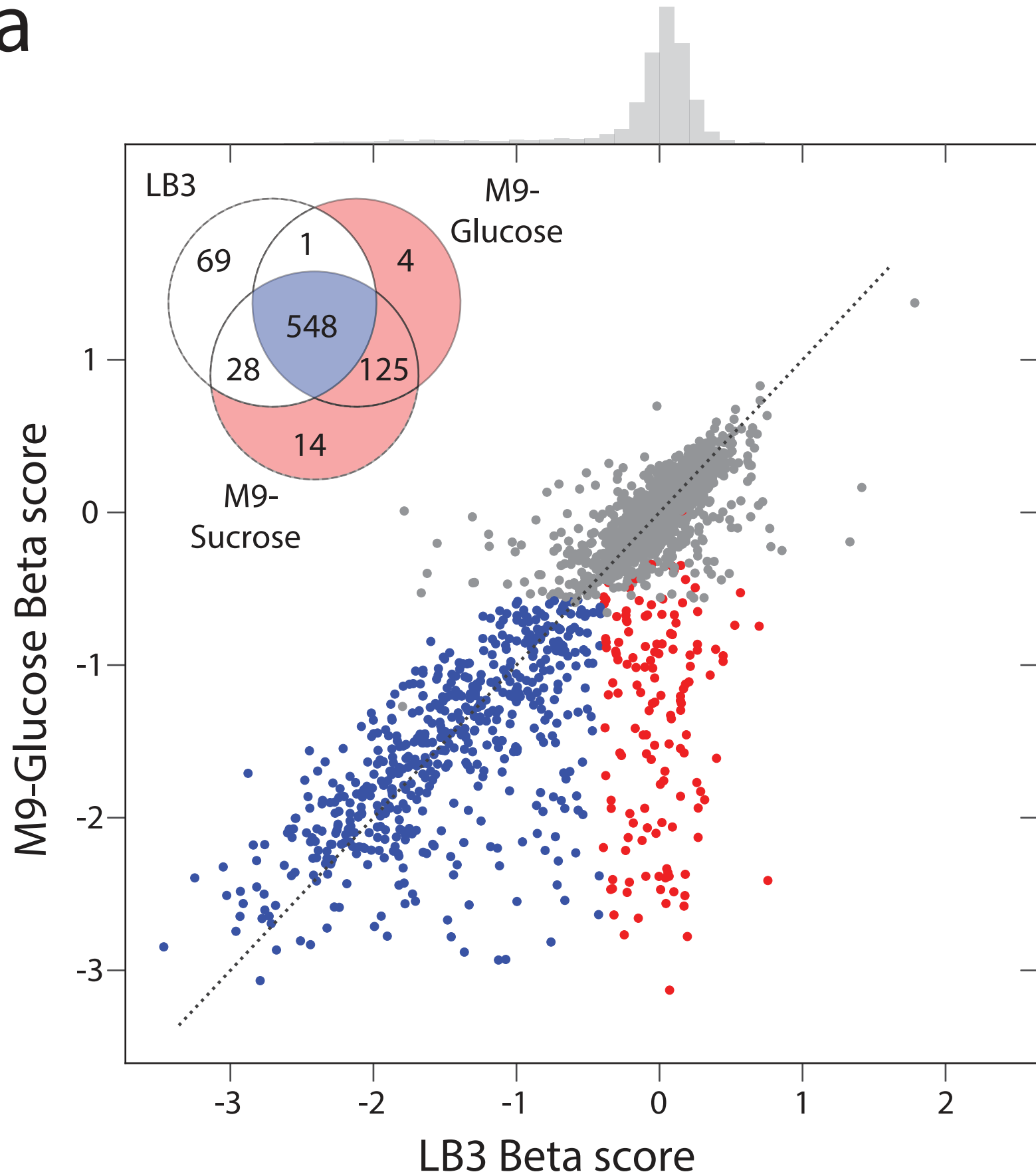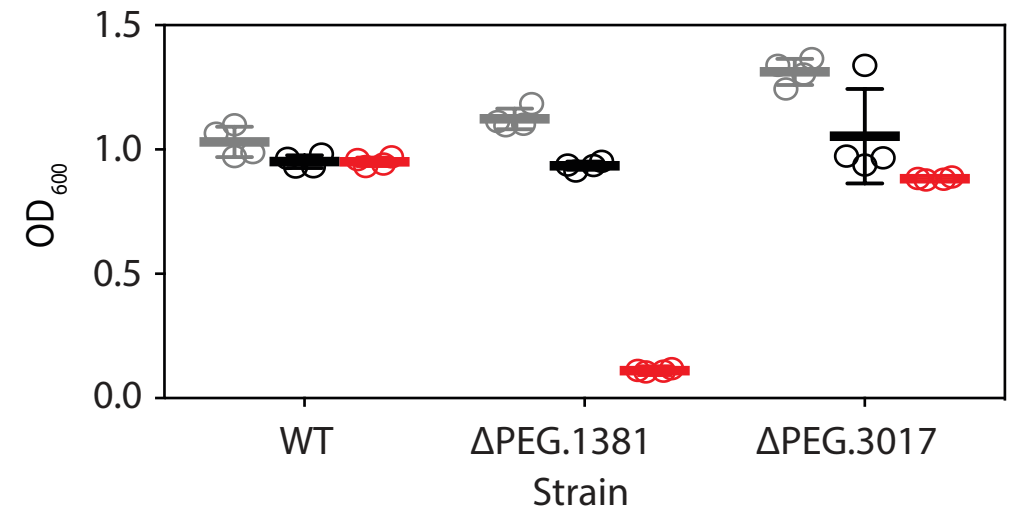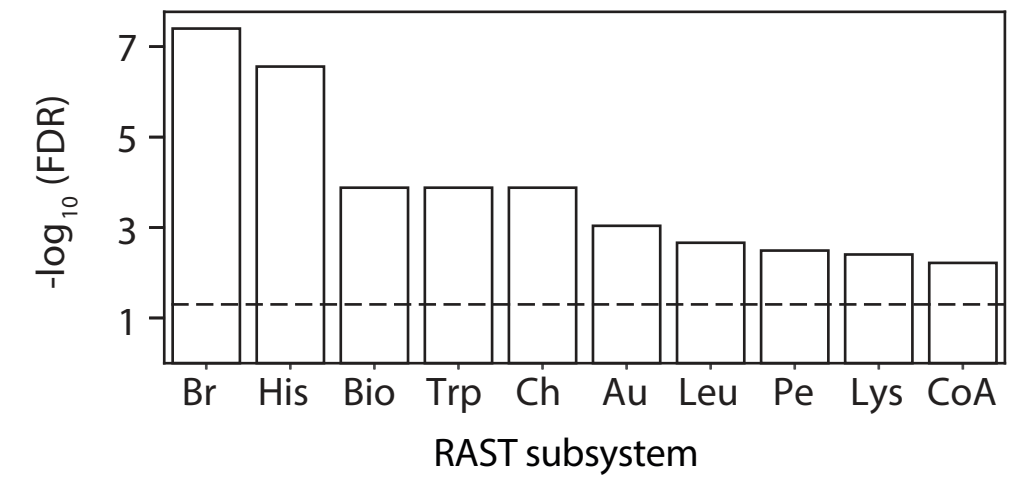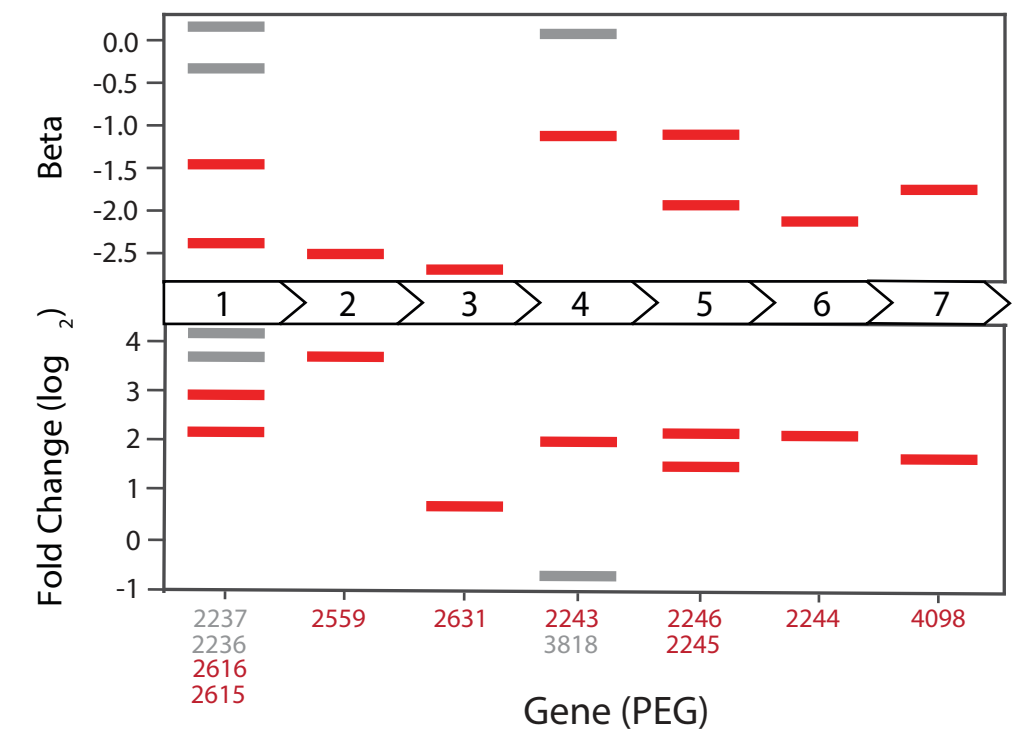721 13. Rock, J. M. *et al.* Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR

722   interference platform. *Nat Microbiol* **2**, 16274 (2017).

723   14. Peters, J. M. *et al.* A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell*

724   **165**, 1493–1506 (2016).

725   15. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene

726   expression. *Cell* **152**, 1173–1183 (2013).

727   16. Liu, X. *et al.* High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus

728   pneumoniae. *Mol. Syst. Biol.* **13**, 931 (2017).

729   17. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9

730   knockout screens. *Genome Biol.* **15**, 554 (2014).

731   18. Tatusova, T. *et al.* NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614–6624 (2016).

732   19. Chao, M. C. *et al.* High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov

733   model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res.* **41**, 9033–9048 (2013).

734   20. Okada, K., Iida, T., Kita-Tsukamoto, K. & Honda, T. Vibrios commonly possess two chromosomes. *J.*

735   *Bacteriol.* **187**, 752–757 (2005).

736   21. Yamazaki, Y., Niki, H. & Kato, J.-I. Profiling of Escherichia coli Chromosome database. *Methods Mol. Biol.*

737   **416**, 385–389 (2008).

738   22. Julio, S. M. *et al.* DNA Adenine Methylase Is Essential for Viability and Plays a Role in the Pathogenesis of

739   Yersinia pseudotuberculosis andVibrio cholerae. *Infect. Immun.* **69**, 7610–7615 (2001).

740   23. Egan, E. S., Fogel, M. A. & Waldor, M. K. Divided genomes: negotiating the cell cycle in prokaryotes with

741   multiple chromosomes. *Mol. Microbiol.* **56**, 1129–1138 (2005).

742   24. Wang, T. *et al.* Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–

743   1101 (2015).

744   25. Schleicher, L. *et al.* Vibrio natriegens as Host for Expression of Multisubunit Membrane Protein Complexes.

745   *Front. Microbiol.* **9**, 2537 (2018).

746   26. Lenz, D. H. *et al.* The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio

747   harveyi and Vibrio cholerae. *Cell* **118**, 69–82 (2004).

748   27. Val, M.-E. *et al.* A checkpoint control orchestrates the replication of the two chromosomes of Vibrio cholerae.
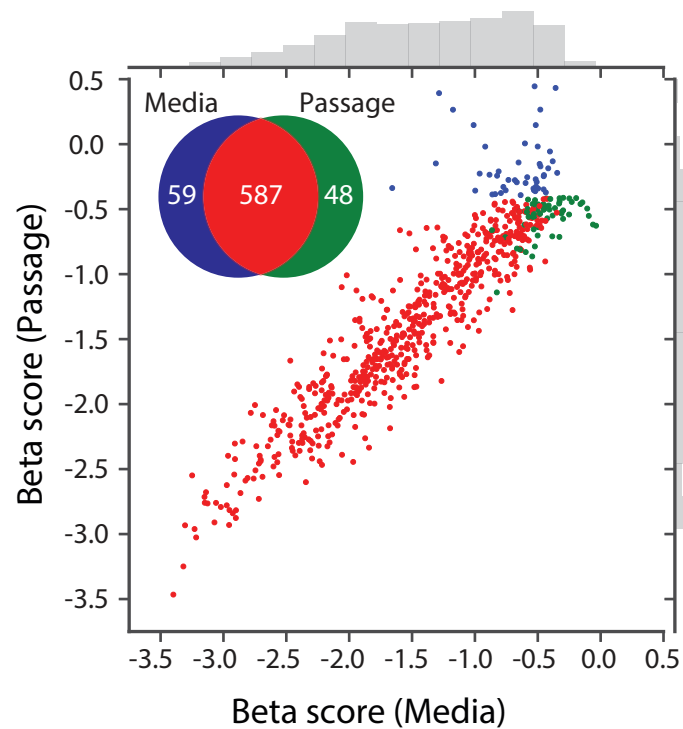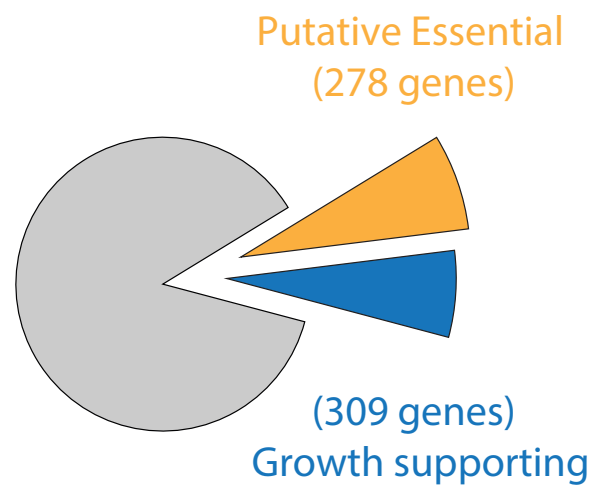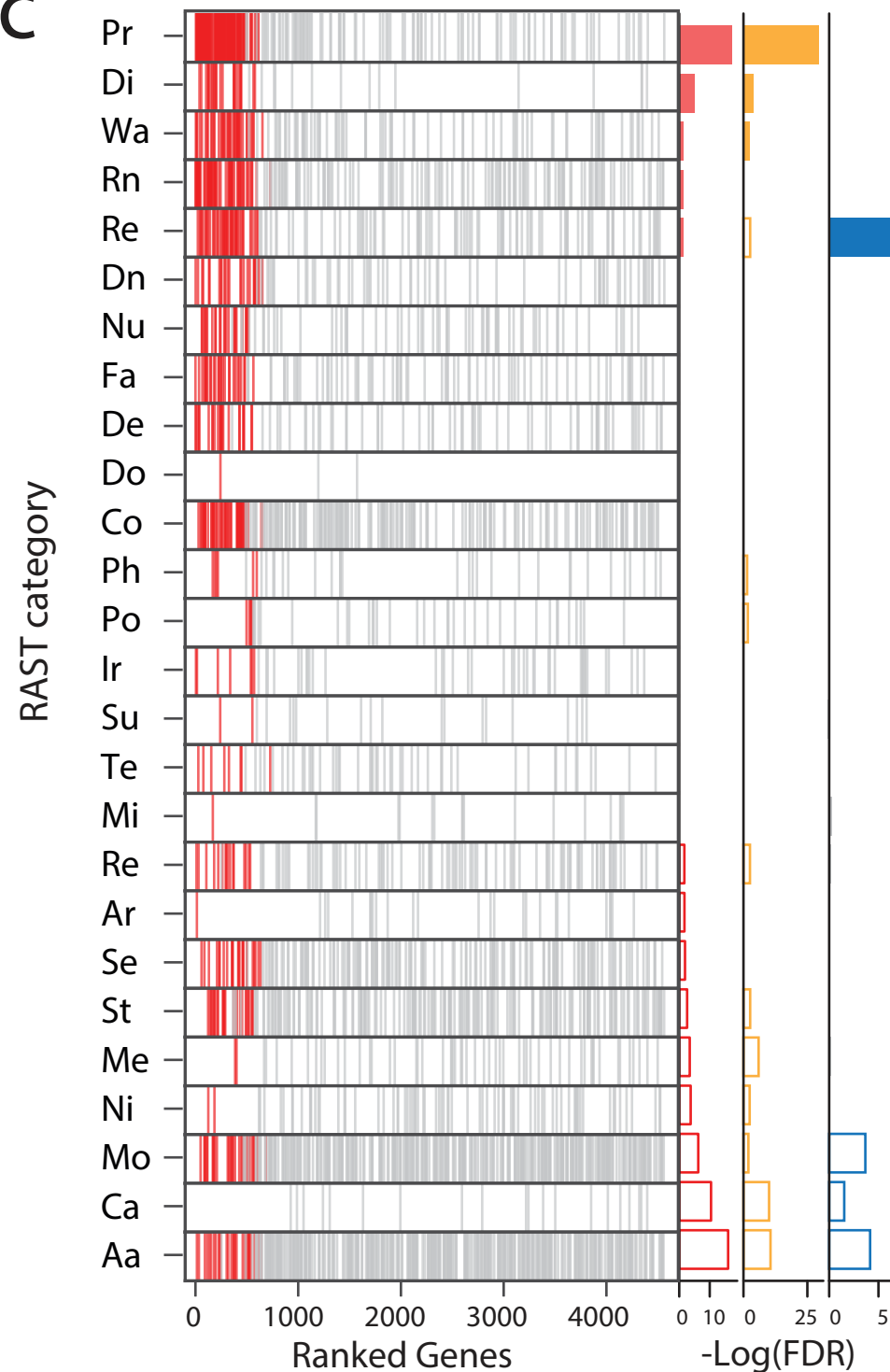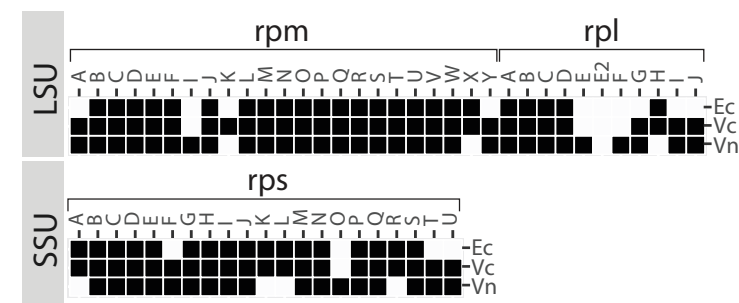
749   *Sci Adv* **2**, e1501914 (2016).

750   28. Mee, M. T., Collins, J. J., Church, G. M. & Wang, H. H. Syntrophic exchange in synthetic microbial

751   communities. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E2149–56 (2014).

752   29. Cui, L. *et al.* A CRISPRi screen in E. coli reveals sequence-specific toxicity of dCas9. *Nat. Commun.* **9**, 1912

753   (2018).

754   30. Zhang, S. & Voigt, C. A. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit

755   design. *Nucleic Acids Res.* **46**, 11115–11125 (2018).

756   31. Vega, N. M., Allison, K. R., Khalil, A. S. & Collins, J. J. Signaling-mediated bacterial persister formation. *Nat.*

757   *Chem. Biol.* **8**, 431–433 (2012).

758   32. Martínez-García, E., Calles, B., Arévalo-Rodríguez, M. & de Lorenzo, V. pBAM1: an all-synthetic genetic tool

759   for analysis and construction of complex bacterial phenotypes. *BMC Microbiol.* **11**, 38 (2011).

760   33. Cameron, D. E., Urbach, J. M. & Mekalanos, J. J. A defined transposon mutant library and its use in identifying

761   motility genes in Vibrio cholerae. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8736–8741 (2008).

762   34. Huerta-Cepas, J. *et al.* Fast genome-wide functional annotation through orthology assignment by eggNOG-

763   mapper. *Mol. Biol. Evol.* (2017). doi:10.1093/molbev/msx148

764   35. Dalia, T. N. *et al.* Multiplex Genome Editing by Natural Transformation (MuGENT) for Synthetic Biology in

765   Vibrio natriegens. *ACS Synth. Biol.* (2017). doi:10.1021/acssynbio.7b00116

766

a — Time-lapse phase contrast microscopy of *V. natriegens* (Vn) and *E. coli* (Ec) at 0 min, 30 min, and 60 min.

b — Generation Time (min) for V. nat and E. coli.

c — Circular genome maps of chr1 (3,248,023 bp) and chr2 (1,927,130 bp), showing OriC and Ter positions.

d — Fraction of genes in category:
- Amino Acids and Derivatives (Aa)
- Carbohydrates (Ca)
- Protein Metabolism (Pr)
- Cofactors, Vitamins (Co)
- Membrane Transport (Me)
- RNA Metabolism (Rn)
- Stress Response (St)
- Respiration (Re)
- Cell Wall and Capsule (Wa)
- Regulation and Cell signaling (Re)
- Fatty Acids and Isoprenoids (Fa)
- Nucleosides and Nucleotides (Nu)
- DNA Metabolism (Dn)
- Virulence, Disease and Defense (De)
- Motility and Chemotaxis (Mo)
- Nitrogen Metabolism (Ni)
- Miscellaneous (Mi)
- Sulfur Metabolism (Su)
- Iron acquisition and metabolism (Ir)
- Cell Division and Cell Cycle (Di)
- Phosphorus Metabolism (Ph)
- Metabolism of Aromatic Compounds (Ar)
- Secondary Metabolism (Se)
- Potassium metabolism (Po)
- Transpos. elements, Plasmids (Te)
- Dormancy and Sporulation (Do)

e — Coverage ($\log_2$) and Cumulative GC-skew vs Position on Genome (Mb) for chr1 and chr2, showing ori and ter positions.

a

Fluorescence (AU) vs gRN A (None, NT, T)

b

Fold Chnag e vs Time (hours)

c

Number of gRNAs / Fraction of gRNAs vs Fraction of ORF

d

13,567 oligos, chip synthesis → gRNA plasmid library → +dCas9 / −dCas9 → Pooled growth → PCR → NGS → MAGeCK analysis: Beta score per gene

a

b

c

d

a

b

Putative Essential
(278 genes)

(309 genes)
Growth supporting

c

RAST category: Pr, Di, Wa, Rn, Re, Dn, Nu, Fa, De, Do, Co, Ph, Po, Ir, Su, Te, Mi, Re, Ar, Se, St, Me, Ni, Mo, Ca, Aa

Ranked Genes

-Log(FDR)

d

LSU — rpm, rpl

SSU — rps

-Ec
-Vc
-Vn

e

Beta score — PE, GS, All

Gene group

f

TPM (mRNA) — PE, GS, All

Gene group

g

chr2

chr1