Evolutionary genome engineering using a restriction–modification system

Yoko Asakura[1,2,*], Hiroyuki Kojima[1] and Ichizo Kobayashi[2,3,4]

[1]Ajinomoto Co., Inc., 1-1 Suzuki-cho, Kawasaki-ku, Kawasaki-shi, Kanagawa, [2]Institute of Medical Science, [3]Department of Medical Genome Sciences, Graduate School of Frontier Science and [4]Graduate Program in Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, Tokyo, Japan

ABSTRACT

Modification of complex microbial cellular processes is often necessary to obtain organisms with particularly favorable characteristics, but such experiments can take many generations to achieve. In the present article, we accelerated the experimental evolution of *Escherichia coli* populations under selection for improved growth using one of the restriction–modification systems, which have shaped bacterial genomes. This resulted in faster evolutionary changes in both the genome and bacterial growth. Transcriptome/genome analysis at various stages enabled prompt identification of sequential genome rearrangements and dynamic gene-expression changes associated with growth improvement. The changes were related to cell-to-cell communication, the cell death program, as well as mass production and energy consumption. These observed changes imply that improvements in microorganism population growth can be achieved by inactivating the cellular mechanisms regulating fraction of active cells in a population. Some of the mutations were shown to have additive effects on growth. These results open the way for the application of evolutionary genome engineering to generate organisms with desirable properties.

INTRODUCTION

Organisms have evolved to adapt to their environments by changing their genomes and transcriptomes. This adaptation involves the re-coordination of complex intra- and inter-cellular processes, and elucidation of these processes is one of the goals of genetics. Such analyses and the generation of organisms with desired properties are mutually dependent, as clearly seen in the synthetic biology of microorganisms.

The design of optimal bacterial genomes with desirable properties has been attempted by bioinformatics-based modeling (1,2) and in the construction of minimal genomes (3–5), which are expected to serve as a basic genetic framework for the addition of genetic elements. These rational approaches are, however, limited by current knowledge. In contrast, evolutionary methods can be applied even before the genetic elements and their global interactions required for optimal performance by an organism are understood. Such evolutionary approaches could utilize multiple cycles of mutations, including genome rearrangements, and selection for adaptation to an environment, as in natural evolution (6–8).

Continuous cultivation of a clonal population in the absence of mutation induction over many generations has been used for bacterial experimental evolution (9–14). In this way, populations acquire increased fitness in a selective environment, and clones with desirable phenotypes can be isolated to study the genetic changes responsible (15). However, this approach is time-consuming, and there are many difficulties involved in linking genome changes to adaptive phenotypes.

Meanwhile, an unique role in genome evolution has been elucidated for restriction–modification (RM) systems (16–19). These are composed of a restriction enzyme and a modification enzyme that methylates restriction sites to prevent cleavage. Genome analysis has provided ample evidence that RM systems have shaped prokaryotic genomes (16), and their mode of action (see model in Supplementary Figure S1) is thought to involve an imbalance between the two enzyme activities that occurs when the persistence or expression of an RM system is somehow disturbed. This leads to chromosomal restriction breaks, which might eventually cause cell death (16–18,20,21). The actions of various proteins on the DNA breaks might generate a variety of rearranged genomes, in addition to the restored genome. If RM genes are properly expressed in one of these genomes, methylation might resume and the restriction attack will cease. An RM system can

therefore select for the persistence both of itself and of its favored host genome variants (21).

Here, we used an RM system *in vivo* to accelerate bacterial adaptive evolution in the population-cultivation procedure. We analyzed the mechanism of growth-phenotype improvement during adaptation, and were successful in rapidly identifying responsible genome changes.

MATERIALS AND METHODS

Strains

Bacterial strains and plasmids are listed in Supplementary Table S1. All strains are derivatives of *Escherichia coli* K-12 wild-type MG1655 (= CGSC#6300, F-, LAM-, *rph*-1) strain (22), a gift from Dr Biek (National Cancer Institute, Bethesda, MD, USA). The YA027 parental strain was constructed from MG1655 as described previously (21). YA027 had a 524-kb-long chromosomal partial duplicate between IS2F (*yi22_6*) and IS2A (*yi22_1*), which was verified by Southern hybridization and nucleotide-sequencing analysis, as described in Supplementary Methods. PaeR7I RM genes linked with a kanamycin resistance (Km^r) gene (21) had been inserted into one of the duplicated regions, while the other had a chloramphenicol resistance (Cm^r) gene insertion. The isogenic r$^-$m$^+$ (YA074) strain was constructed as described in Supplementary Methods. The identified mutations were re-introduced into YA027, MG1655 and related strains as described in Supplementary Methods.

Medium, culture conditions and growth analysis

Escherichia coli cells were cultured at 37°C in an L-shaped test tube containing 10 ml Davis minimum medium supplied with 20 amino acids each at 200 mg/l. They experienced constant shaking at 70 r.p.m. using a TN-2612 rocking incubator (Advantec, Dublin, CA, USA), which monitored growth automatically by measuring the OD$_{660}$ every 12 min. Km and Cm were included at concentrations of 25 and 12.5 µg/ml, respectively. In evolution experiments, cells were grown for 24 h as described above, then 100 µl aliquots were used to inoculate 10 ml of fresh medium for the next culture. For the passage 0 of the evolution experiment and growth analysis of evolved clones and reconstructed strains, cells were grown to an OD$_{660}$ value of 0.1–0.15, then used to inoculate a main culture at an OD$_{660}$ value of 0.01. For transcriptome analysis, cells were harvested when the OD$_{660}$ reached 0.1.

Test of RM phenotype

Plaque assays were performed using λ *cI71* as described previously (23).

Statistical confirmation of restriction-accelerated evolution in growth

Three independent experiments were performed, in each of which six populations of r$^+$m$^+$ and r$^-$m$^+$ cells were serially propagated until passage 5. We confirmed that the increase in initial growth of the r$^+$m$^+$ population was larger than that of the r$^-$m$^+$ population based on a one-sided Mann–Whitney *U*-test ($\alpha = 0.01$).

Southern hybridization analysis using insertion sequence element probes

Chromosomal DNA was digested with EcoRV and other restriction enzymes, and probed with IS*1*, IS*2*, IS*3* or IS*5* fragments, which were polymerase chain reaction (PCR)-amplified as described in Supplementary Methods.

PFGE

Genomic DNA was digested with NotI or BlnI using a CHEF Bacterial Plug kit (Bio-Rad, Hercules, CA, USA). PFGE was performed in a 0.6% SeaKem Gold (BioWhittaker, Walkersville, MD, USA) 1× Tris–acetate–EDTA (TAE buffer) agarose gel using the CHEF-DRIII PFGE system (Bio-Rad). The running conditions were as follows: switching range, 0.1–5.0 s; temperature, 14°C; running time, 10 h; voltage, 6 V/cm; angle, 120°.

Transcriptome and statistical analyses

Transcriptome analyses were performed at OD$_{660}$ 0.1 (early log phase) to minimize the influence of the growth phase. All gene-expression measurements were made in duplicate or triplicate with total RNA and an Affymetrix (Santa Clara, CA, USA) *E. coli* anti-sense genome array according to the manufacturer's instructions (http://www.affymetrix.com/index.affx). The statistical significance of differences in transcript abundance between strains was determined by an analysis of variance (ANOVA) at a significance level of $P < 0.05$. Genes that were reliably differentially expressed between strains were identified by the Student–Newman–Keuls *post-hoc* test. Multiple testing corrections using the Benjamini and Hochberg false discovery rate method were performed to analyze evolved clones (see Supplementary Materials). Gene-expression data and other relevant information were deposited in the Gene Expression Omnibus (GEO) repository under accession number GSE21869. The transcriptional changes of 12 selected genes were validated using quantitative real-time reverse-transcription (RT)–PCR (Supplementary Materials and Supplementary Table S2).

Identification of mutations conferring growth difference among passage 84 clones

Among the 268 genes with higher expression in clones 3-84-4 and 3-84-10 compared to clones 3-84-2 and 3-84-6 and the 261 genes with lower expression (Supplementary Table S3 and Supplementary Materials), genes with the eight highest and 18 lowest expression levels and the regulatory genes of those that showed expression changes (*rpoS*, *rpoB*, *purR*, *lrp* and *hns)* were sequenced, together with 500-bp fragments upstream and downstream of their open-reading frames.

Detection of identified mutations in evolving populations

Ten clones were isolated from each population at passages 11, 84 and 172. Chromosomal regions around *tnaLAB*, *flh-che*, *cydA*, *e14*, *rph-pyrE* and *lacY* were amplified by PCR using the corresponding sets of oligonucleotide primers, shown in Supplementary Table S4, and fragment lengths were analyzed by agarose gel electrophoresis. When the length of PCR fragments differed from those already identified in the representative clones, nucleotide sequences were determined. Variants in insertion sequence (IS) elements and insertion positions were scored as mutant forms.

RESULTS

Adaptive evolution experiment

The parental *E. coli* strain for experimental evolution unstably harbors RM genes on its chromosome, such that the host genome is continuously damaged by the restriction enzyme. This might allow multiple cycles of genome rearrangements, and selection of the cells that harbored the RM genes. The strain can be cultured for generations, during which time cells with a competitive growth advantage are selected.

The parental strain (YA027, Figure 1A), constructed from the *E. coli* wild-type MG1655 strain, carried *Pae*R7I RM genes linked with a Km^r gene (21) on a chromosomal partial duplicate, while the other duplicate had a Cm^r gene insertion (Figure 1A). As the r$^+$m$^+$ strain (YA027) showed no difference in the growth rate with or without $12.5\,\mu g/ml$ Cm plus $25\,\mu g/ml$ Km, and as the growth rate of the isogenic r$^-$m$^+$ strain (YA074) was the same with that of the r$^+$m$^+$ strain under the latter condition, we added $12.5\,\mu g/ml$ Cm and $25\,\mu g/ml$ Km in the experiments here. Transcriptome analysis of this strain suggested that the PaeR7I restriction endonuclease damaged the chromosome and induced the SOS response, because similar r$^+$-dependent transcription was observed to that previously seen in cells dying due to chromosomal breakage after the loss of plasmid *Pae*R7I RM genes (24), albeit to a lesser extent (detailed in Supplementary Table S5). The parental strain stably maintained the PaeR7I RM-Km^r allele, even in the absence of antibiotic selection, as expected (Figure 1B).

Four populations of the r$^+$m$^+$ parental strain (YA027) and two control populations of an isogenic r$^-$m$^+$ strain (YA074) were grown and serially propagated using 100-fold dilutions in an amino-acid-rich medium with Cm and Km (Figure 1C). One passage corresponded to six or seven generations in the beginning of the experiment, although this may have changed later. Growth was monitored daily (Figure 1D).

The growth curve shifted upward with passage number in every population (Figure 1D), representing improved growth in terms of both initial rate and saturation cell density. At passage 172 (the last passage), adaptive evolution had resulted in an approximate 7-fold increase in the initial hourly growth rate [indicated by increased optical density (OD_{660}) in the first 4 h; Figure 2A] and a 2-fold increase in the saturation cell density in all populations.

Restriction-accelerated evolutionary changes in bacterial growth and the genome

We examined whether the evolutionary change was accelerated in an r$^+$-gene-dependent manner for the adaptive phenotype and the genotype.

Initial hourly growth rates (initial growth rate) for each passage over the course of the evolutionary experiment are shown in Figure 2A. We could not see difference in the growth rates at passage 0 between the r$^+$ and the r$^-$ populations. For the first five passages, the increase in initial growth (slope of the graph) of the r$^+$m$^+$ population (0.0019 increased OD_{660} per hour per passage for an average of four populations) was much larger than that of the r$^-$m$^+$ population (average, 0.0002 per passage) (Figure 2A). This was statistically confirmed by three further independent tests (0.0031 ± 0.0005, 0.0022 ± 0.0006 and 0.0022 ± 0.0002 for the r$^+$m$^+$ population; 0.0010 ± 0.0003, 0.0008 ± 0.0005 and 0.0012 ± 0.0006 for the r$^-$m$^+$ population; average $\pm$ SD for each experiment. See 'Materials and Methods' section for detail.). From passage 6 to 172, the averages of the slopes calculated for each successive 10 passages were larger in the r$^+$m$^+$ population than in the r$^-$m$^+$ population, based on a one-sided χ^2-test ($P < 1.1 \times 10^{-191}$; Figure 2A).

Genome rearrangements, as measured by IS restriction fragment length polymorphisms (RFLPs), appeared more extensive in the r$^+$ strain than the r$^-$ strain over the first 11 passages (Figure 2B). Ten clones were isolated from each population at passage 11 and were examined by genomic Southern hybridization with an IS*1*, IS*2* or IS*3* probe. For each clone, the difference score was obtained as a sum of the number of newly appearing or disappearing bands compared with the parental strain (Supplementary Figure S2), the averages of which (per clone) for the r$^+$ and r$^-$ genotypes were 2.5 and 0.2, respectively (Figure 2B); these were significantly different according to the one-tailed Welch's *t*-test ($P = 0.00397$, $\alpha = 0.05$; Figure 2B).

Thus, the r$^+$ strains evolved faster than the r$^-$ strains with respect to the growth rate phenotype and the IS polymorphism genotype. Bacteriocidal antibiotics, such as kanamycin are known to ultimately lead to radical production and DNA damage (25), which could influence the evolution rate. In the present study, we compared the evolution rates between the r$^+$ and r$^-$ populations under the same condition regarding the concentration of the antibiotics. Our previous transcriptome study suggests reactive oxygen species production might be triggered by the DNA damage resulting from the loss of the *Pae*R7I RM genes (24). In the r$^+$ strain, the influence by kanamycin could be additive to that by the r$^+$ gene, or could lead to more complex cellular response interacting with the responses resulting from the r$^+$-dependent DNA damage, which might influence the evolution rate differently in the r$^+$ strain from in the r$^-$ strain.
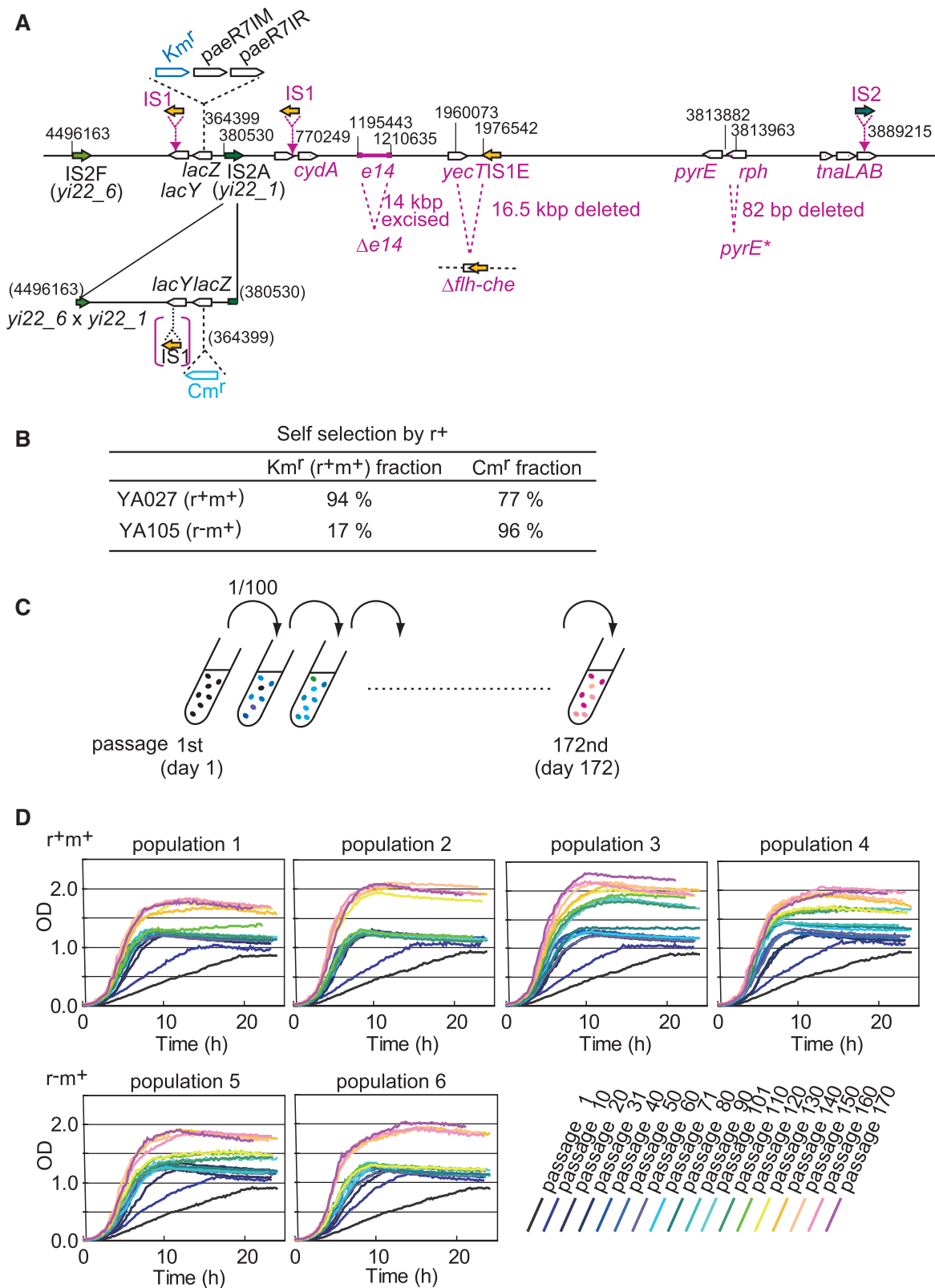
Figure 1. Experimental evolution for growth improvement. (**A**) Genome map of parental strain YA027 with changes (shown in magenta) identified after experimental evolution. (**B**) Self-selection by RM system. From a single YA027 colony, 10^{8-9} cells were suspended in LB liquid medium without antibiotics and grown overnight. The culture was diluted 1000-fold and allowed to grow overnight. This procedure was repeated twice more. Resulting clones were tested for antibiotic resistance. Data represent an average of two independent experiments each with 50 clones. (**C**) Experimental procedure. Four populations of r^+m^+ strain and two populations of an isogenic r^-m^+ strain were cultured with shaking in amino acid-rich medium with Cm and Km. They were serially propagated by 100-fold dilution daily until passage 172, and growth was monitored. (**D**) Growth curves after every 10 passages. r^+ phenotype, as measured by λ-assay (23), was maintained by 100% (30/30) of clones at passages 42, 100 and 172 in all r^+ populations.

Dynamic transcriptome changes during adaptive evolution

To characterize evolutionary changes associated with growth improvements, we focused on population 3 (Figure 1D). Genome changes should improve host growth in combination with the effects of changes inherited from earlier passages. We noticed two sharp rises in the growth curve: one around passage 10 and the other around passage 80 (Figure 1D), at which points a genome change(s) responsible for the growth improvement may have occurred. The population at this passage was expected to comprise a mixture of clones that had or had not yet experienced such a change. Indeed, the 10 clones isolated at passages 11 and 84 could be split

into two or more groups with respect to growth level (Figure 3).

It has been reported that genetic background of clones in an evolving population is diverse and that genetic diversity in a population fluctuates along with evolution (14,26). Phenotypic diversity might be basically dependent on the genotypic diversity, but might be smaller because synonymous mutations are possible for a phenotype. Extent of diversity we can see depends on the timing when we take the samples. We saw phenotypic diversity among the isolated clones also at passage 172.

We chose representative clones [labeled in Figure 3 as (population number) – (passage number) – (clone number)]
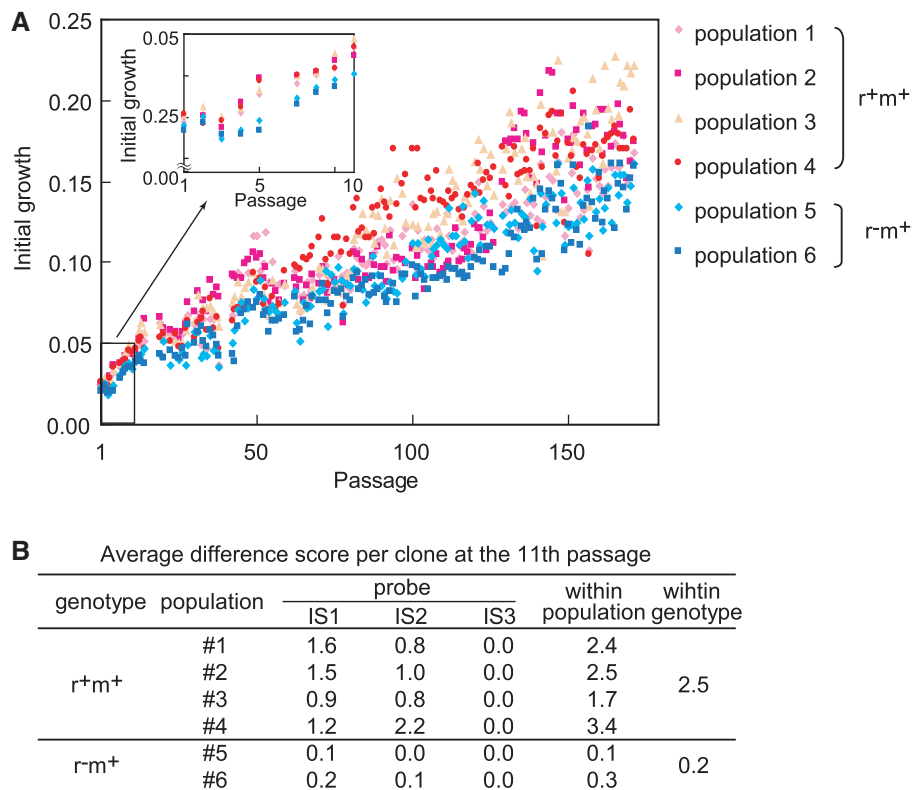
Figure 2. Acceleration of evolution by RM system. (**A**) Change in initial growth, as measured by OD increase per h in first 4 h, during experimental evolution. (**B**) Genome rearrangements at passage 11 as measured by IS element polymorphism. Ten clones from each of six populations examined. For each population, difference score obtained as sum of number of newly appearing or disappearing bands compared with parental strain. Averages of difference scores for r^+m^+ and r^-m^+ genotypes are shown.
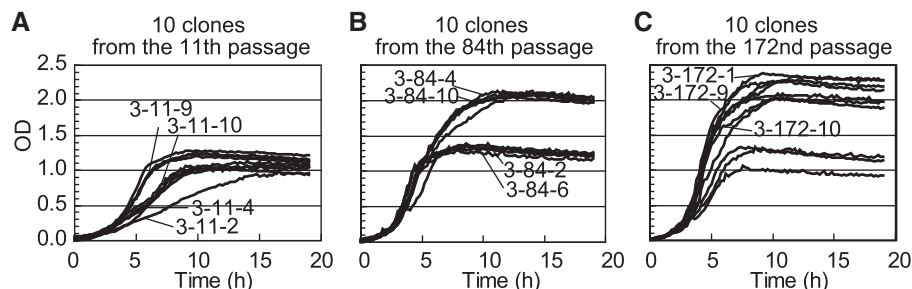
Figure 3. Properties of clones isolated from population at different evolutionary stages. Clones are labeled [population (culture) number] – [passage number] – [clone number].

for transcriptome analysis (Supplementary Figure S3). Genes with significantly different transcript abundance compared with the parent were identified in each of these evolved clones ($P < 0.05$, ANOVA, Supplementary Tables S6–S8). Their numbers, functional category distribution, and hierarchical analysis are shown in Figure 4 and Supplementary Table S9. The clones isolated in a particular generation (passage number) shared a common functional category distribution for the affected genes (Figure 4A), and this pattern appeared unchanged throughout the evolution experiment, especially until passage 84. This implied that expression changes that realized growth improvements were conserved with respect to cellular functions.

By passage 84, genes involved in translation, cell-wall and -membrane synthesis, and nucleotide and coenzyme metabolism were significantly activated (Figure 4A, Supplementary Table S9), which might contribute to larger cell mass production. Genes for sugar catabolism and amino-acid biosynthesis were down-regulated, which seemed reasonable as the medium did not contain sugars and was rich in amino acids. Genes involved in cell motility and signal transduction were inactivated (Figure 4A and B), which might be expected as there is no need for movement in a well-mixed culture, and is consistent with deletion of the flagella and chemotaxis genes explained below. At passage 172, genes involved in translation were no longer strongly enriched among the affected genes (Figure 4A and Supplementary Table S9).

Among the genes with transcript changes common to all three clones at passage 172 (Figure 4C), more than half had already shown changes in the same direction in all clones isolated from passages 11 and 84, suggesting that major changes had occurred by passage 11 that persisted through passage 172. The patterns of global gene expression appeared to diversify and converge repeatedly during the evolution experiment (Figure 4B).

Identification of relevant genome changes

We next tried to identify genome changes responsible for growth improvements. Rather than enumerating all mutations and testing each in turn, we sought clues in the above transcriptome changes accompanying the step-wise phenotype changes, using representative clones of passages 11 and 84 (Figure 3 and Supplementary Figure S3). RFLP, pulsed-field gel electrophoresis (PFGE) and sequence analysis (Supplementary Materials) were performed for confirmation. Six genome changes were identified, as described in Figures 1A and 5 and Supplementary Materials.

Comparing the transcriptome of four clones from passage 11 (detailed in Supplementary Table S10) revealed that two of them (3-11-9 and 3-11-10; Figure 3) had a deletion of a 16 454-bp region, on which genes related to flagella and chemotaxis are present (Figure 5B, Δ*flh-che*). The comparison also revealed an IS*2* insertion in the *tnaB* gene of two clones (3-11-9 and 3-11-4), which most likely affected the *tnaLAB* messenger RNA quantity (Figure 5A, *tnaAB*::IS). Among the 10 clones from passage 11 (Figure 3), the presence/absence of these two

genome changes was consistent with the three levels in the growth curve (Figures 3 and 5E). The *tnaAB*::IS and Δ*flh-che* mutations were both present in all 10 clones from passage 84 (Figure 5E). Two of these clones (3-84-4 and 3-84-10) had a growth advantage over two others (3-84-2 and 3-84-6; Figure 3). Comparing the transcriptome of the former and latter two clones (ANOVA, $P < 0.05$, see Materials and methods section, Supplementary Materials and Supplementary Table S3) revealed the presence of *pyrE** (Figure 5C and Supplementary Figure S4) and *lacY*::IS mutations (Figure 1A) in the former two clones (Figure 5E). Δ*e14* (27,28) and IS-*cydA* mutations (Figures 2A and 5) were identified through transcriptome analysis and IS*1* Southern hybridization of passage 84 clones.

No other large deletion (>10 bp) or genome rearrangement (>10 bp) than the Δ*flh-che* and Δ*e14* mutations was found in any of the 10 clones from passages 84 and 172 in PFGE.

Distribution of mutations and reconstruction

To further explore the question of mutual interaction between the mutations, we investigated how these changes occurred within and across populations during adaptation. Ten clones were isolated from each of the six populations at passages 11, 84 and 172, and we tested for the six identified genome changes by PCR (Figure 5F).

Looking at the four r$^+$ populations, the *tnaAB*::IS and IS-*cydA* mutations came to dominate a population once they had appeared. Δ*flh-che* showed a similar pattern, with the exception of population 4 where it was lost. *pyrE** increased in frequency in two out of four populations by passage 84, but declined thereafter; it might have been replaced by a lineage containing a mutation with a comparable effect. *lacY*::IS and Δ*e14* were variable among the populations, which was consistent with their lack of effect on growth. These results suggest that adaptive evolution was realized by the different populations through different pathways with respect to mutations, although the pathways shared common genetic changes.

Regarding the influence of the r$^+$ genotype, frequency of the Δ*flh-che* deletion appeared higher in the r$^+$ populations than in the r$^-$ populations at the early stage of the evolution. This suggested that some changes might occur more frequently in the r$^+$ background, though more precise measurements with larger and balanced numbers of cultures with r$^+$m$^+$, r$^-$m$^+$ and r$^-$m$^-$ backgrounds to confirm this point.

To examine adaptive effects of the six identified mutations, they were re-introduced into the parental strain either alone or in combination (Supplementary Materials). The *tnaAB*::IS, Δ*flh-che*, and *pyrE** mutations each improved growth, and their effects were additive (Figure 6). IS-*cydA* and *lacY*::IS had no detectable effect on growth, even when combined with any of the other identified mutations (Figure 6 and data not shown). On growth, Δ*e14* also had no effect. Similar results were obtained in MG1655 used for construction of the parental strain (Supplementary Figure S6). A combination

A

Functional category	up (evolved > parent)											down (evolved < parent)											
	3-11-				3-84-				3-172-			3-11-				3-84-				3-172-		
	2	4	10	9	2	6	4	10	10	1	9	2	4	10	9	2	6	4	10	10	1	9
J Translation																						
J'* tRNA and rRNA																						
K Transcription																						
L Replication, recombination and repair																						
D Cell cycle control, mitosis and meiosis																						
V Defense mechanisms																						
T Signal transduction mechanisms																						
M Cell wall/membrane biogenesis																						
N Cell motility																						
U Intracellular trafficking and secretion																						
O Posttranslational modification*																						
C Energy production and conversion																						
G Carbohydrate transport and metabolism																						
E Amino acid transport and metabolism																						
F Nucleotide transport and metabolism																						
H Coenzyme transport and metabolism																						
I Lipid transport and metabolism																						
P Inorganic ion transport and metabolism																						
Q Secondary metabolites**																						
R General function prediction only																						
S Function unknown																						
not assigned																						
number of the gene	654	697	701	694	616	654	815	776	369	370	328	909	944	1012	999	807	824	965	1016	288	304	328
	(601)				(450)				(286)			(825)				(646)				(253)		

*: and protein turnover, chaperons
**: biosynthesis, metabolism and catabolism

strongly enriched ($p < 0.01$)
significantly enriched ($0.01 < p < 0.05$)
enriched ($0.05 < p$)
not observed or not enriched

Figure 4. Transcriptome changes during adaptive evolution. (**A**) Functional category distribution. Genes differentially expressed compared with parental strain were classified according to function using Clusters of Orthologous Groups (COG) codes (http://www.ncbi.nlm.nih.gov/COG/new/) with an additional J' category. Colors show significance of category over-representation based on χ^2-test. Gene numbers shown at bottom of each column, and numbers of genes commonly up- or down-regulated in all three or four clones selected at each passage shown in parentheses. (**B**) Dynamics. (i) Genes with differential expression in either of selected clones compared with parental strain hierarchically clustered. Transcript level of each gene normalized to that of parental strain (yellow). (ii) Standard correlations of transcriptome measurements between each clone and parental strain. (**C**) Genes with significantly more (up-regulated) or fewer (down-regulated) transcripts shared by clones at passages 11, 84 and 172. (i) Venn diagram. Each circle shows number of shared genes with significantly more (up-regulated) or fewer (down-regulated) transcripts compared with parental strain for clones at passages 11, 84 and 172, respectively. Among genes with shared transcript changes in three clones at passage 172, transcription of more than half had already shown similar changes in all clones isolated from passage 11 (I). (ii) Examples of genes shown in Venn diagram (for details see Supplementary Table S8).

of *tnaAB*::IS, Δ*flh-che* and *pyrE** mutations further improved growth of the parental strain, but the level attained was not as high as that observed in the evolved clones (Figure 3), which suggests involvement of other, unidentified mutations in the adaptation. Some of these will be dealt with in the next section.

Roles of mutations

The *tnaAB* genes are involved in indole production from tryptophan. Thus, as expected, disruption of *tnaAB* in the present study resulted in loss of tryptophan consumption (Supplementary Figures S7 and S8). Indole works as a signal molecule in cell-to-cell communication (quorum sensing) and reduces both the growth rate during the exponential phase and the saturation cell density during the stationary phase of *E. coli* (29,30). This might explain the occurrence of the *tnaAB*::IS mutation. A selective advantage of the *tnaAB* mutation has also been implied by the disruption of *tnaAB* reported in a laboratory strain (31).

We observed a decrease in the transcript level compared with the parental for Lsr transporter genes involved in the intake of another quorum-sensing signal, AI-2, from passages 11 through 172 (Supplementary Table S8). Under stressful conditions, these and other signal molecules induce changes in gene expression that lead to central metabolism inactivation, cell-division reduction (29,32), biofilm formation (33) and subpopulation death (34) for population resistance (35). Most of these changes ultimately result in reduction of the population growth as a whole; therefore, freedom from the quorum-sensing control seems to be a credible strategy for improving population growth.

The *cydAB* genes encode cytochrome *bd* oxidase, which is involved in ATP generation at the end of the respiratory
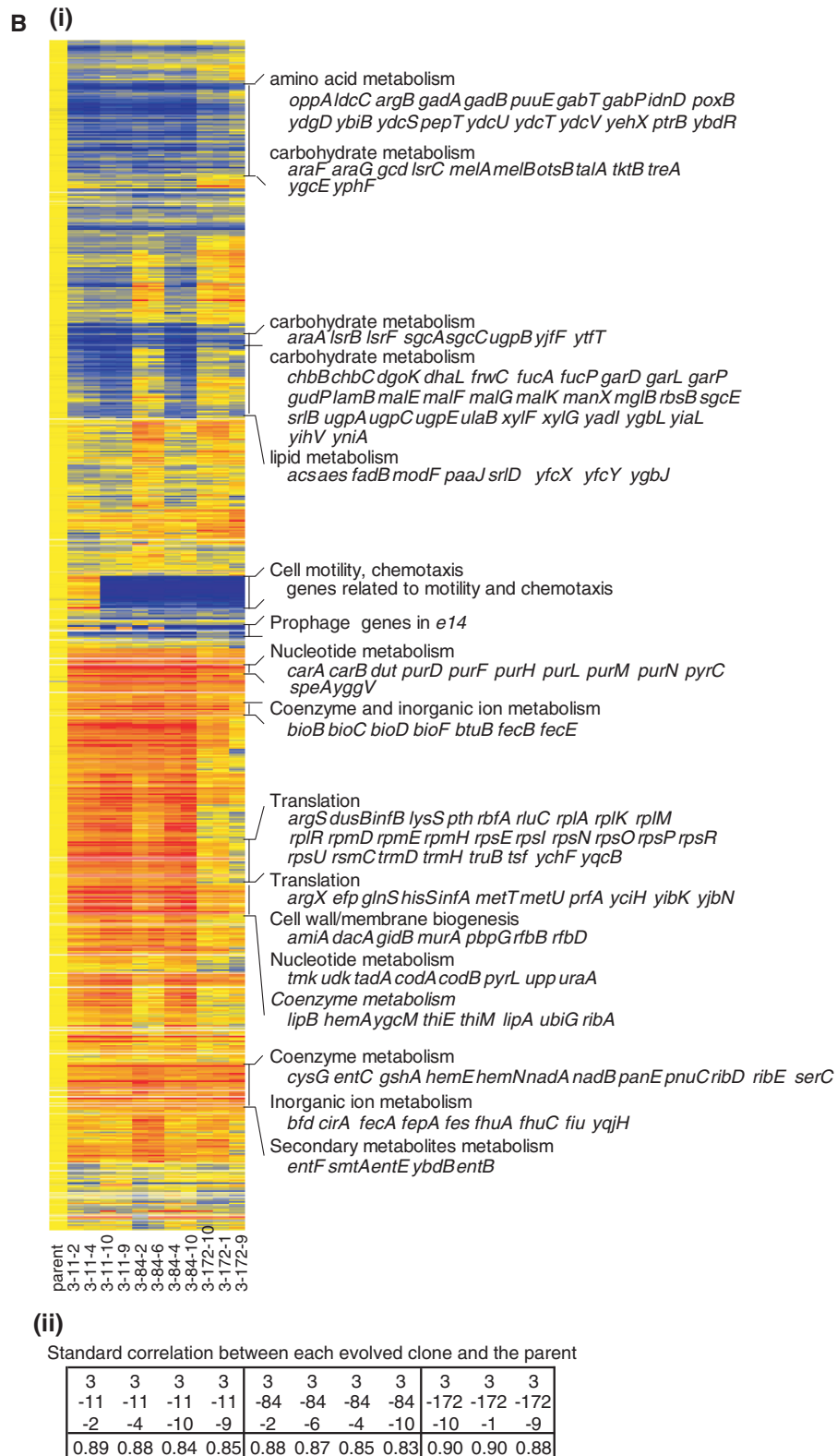
B **(i)**

amino acid metabolism
  *oppA ldcC argB gadA gadB puuE gabT gabP idnD poxB
  ydgD ybiB ydcS pepT ydcU ydcT ydcV yehX ptrB ybdR*
carbohydrate metabolism
  *araF araG gcd lsrC melA melB otsB talA tktB treA
  ygcE yphF*

carbohydrate metabolism
  araA lsrB lsrF sgcA sgcC ugpB yjfF ytfT
carbohydrate metabolism
  *chbB chbC dgoK dhaL frwC fucA fucP garD garL garP
  gudP lamB malE malF malG malK manX mglB rbsB sgcE
  srlB ugpA ugpC ugpE ulaB xylF xylG yadI ygbL yiaL
  yihV yniA*
lipid metabolism
  acs aes fadB modF paaJ srlD yfcX yfcY ygbJ

Cell motility, chemotaxis
  genes related to motility and chemotaxis

Prophage genes in *e14*

Nucleotide metabolism
  *carA carB dut purD purF purH purL purM purN pyrC
  speA yggV*
Coenzyme and inorganic ion metabolism
  bioB bioC bioD bioF btuB fecB fecE

Translation
  *argS dusB infB lysS pth rbfA rluC rplA rplK rplM
  rplR rpmD rpmE rpmH rpsE rpsI rpsN rpsO rpsP rpsR
  rpsU rsmC trmD trmH truB tsf ychF yqcB*
Translation
  argX efp glnS hisS infA metT metU prfA yciH yibK yjbN
Cell wall/membrane biogenesis
  amiA dacA gidB murA pbpG rfbB rfbD
Nucleotide metabolism
  tmk udk tadA codA codB pyrL upp uraA
Coenzyme metabolism
  lipB hemA ygcM thiE thiM lipA ubiG ribA

Coenzyme metabolism
  cysG entC gshA hemE hemN nadA nadB panE pnuC ribD ribE serC
Inorganic ion metabolism
  bfd cirA fecA fepA fes fhuA fhuC fiu yqjH
Secondary metabolites metabolism
  entF smtA entE ybdB entB

parent
3-11-2
3-11-4
3-11-10
3-11-9
3-84-2
3-84-6
3-84-4
3-84-10
3-172-10
3-172-1
3-172-9

(ii)

Standard correlation between each evolved clone and the parent

3 -11 -2	3 -11 -4	3 -11 -10	3 -11 -9	3 -84 -2	3 -84 -6	3 -84 -4	3 -84 -10	3 -172 -10	3 -172 -1	3 -172 -9
0.89	0.88	0.84	0.85	0.88	0.87	0.85	0.83	0.90	0.90	0.88

Figure 4. Continued.

chain for electron transport and has been suggested to play a role in programmed cell death by generating super-oxide (36–38). Their expression is regulated by ArcA and Fnr through upstream binding (39). In our present study, the evolved clones had IS insertions in the regions of ArcA and Fnr binding (Figure 5D), which likely affects the
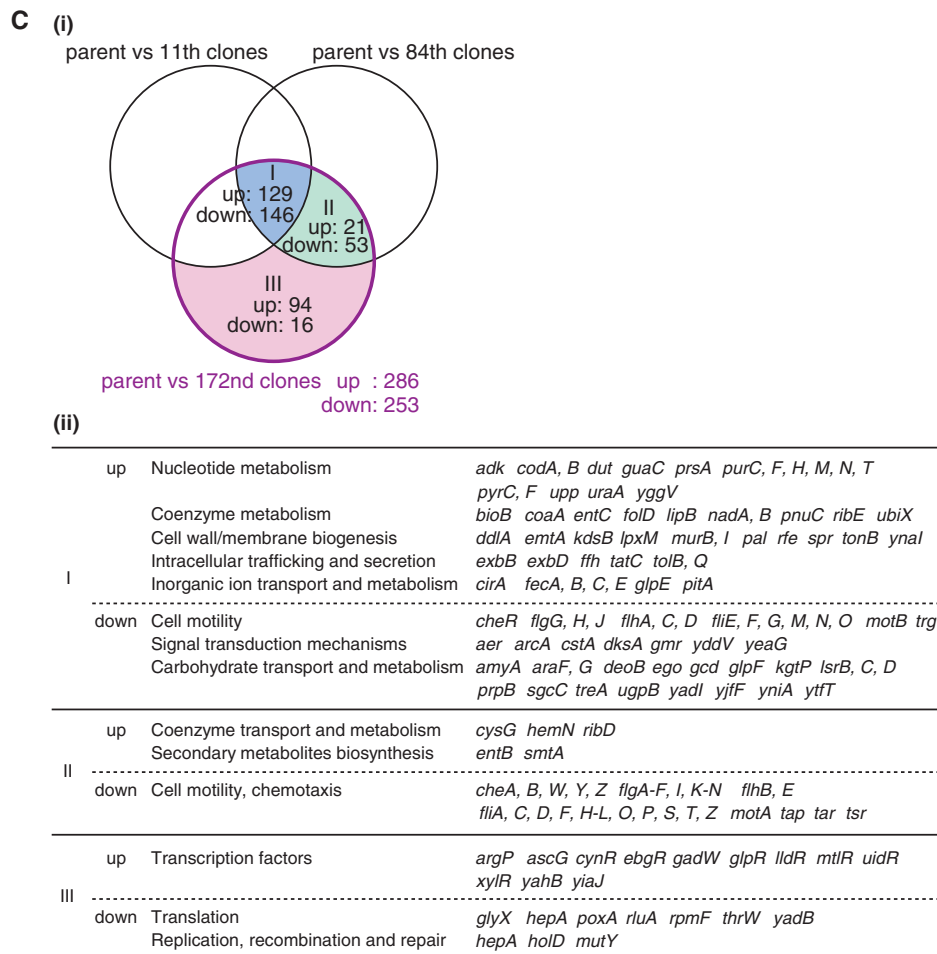
C (i)

(ii)

I	up	Nucleotide metabolism	*adk codA, B dut guaC prsA purC, F, H, M, N, T pyrC, F upp uraA yggV*
		Coenzyme metabolism	*bioB coaA entC folD lipB nadA, B pnuC ribE ubiX*
		Cell wall/membrane biogenesis	*ddlA emtA kdsB lpxM murB, I pal rfe spr tonB ynaI*
		Intracellular trafficking and secretion	*exbB exbD ffh tatC tolB, Q*
		Inorganic ion transport and metabolism	*cirA fecA, B, C, E glpE pitA*
	down	Cell motility	*cheR flgG, H, J flhA, C, D fliE, F, G, M, N, O motB trg*
		Signal transduction mechanisms	*aer arcA cstA dksA gmr yddV yeaG*
		Carbohydrate transport and metabolism	*amyA araF, G deoB ego gcd glpF kgtP lsrB, C, D prpB sgcC treA ugpB yadI yjfF yniA ytfT*
II	up	Coenzyme transport and metabolism	*cysG hemN ribD*
		Secondary metabolites biosynthesis	*entB smtA*
	down	Cell motility, chemotaxis	*cheA, B, W, Y, Z flgA-F, I, K-N flhB, E fliA, C, D, F, H-L, O, P, S, T, Z motA tap tar tsr*
III	up	Transcription factors	*argP ascG cynR ebgR gadW glpR lldR mtlR uidR xylR yahB yiaJ*
	down	Translation	*glyX hepA poxA rluA rpmF thrW yadB*
		Replication, recombination and repair	*hepA holD mutY*

Figure 4. Continued.

cydAB-Arc-Fnr regulatory system. The role of the mutation in growth improvement and in changes to transcript abundance is under study. It appears reasonable that we obtained a mutation in genes implicated in programmed cell death because selection for better growth would coincide with selection against death.

The Δ*flhE-che* mutation deletes cellular machinery involved in chemotaxis and motility. Loss of the flagella may well be advantageous in our well-mixed culture to save energy. Indeed, many *E. coli* laboratory strains are non-flagellate, and spontaneous non-flagellate mutants were reported to rapidly dominate a motile population in a stirred culture (40). In our present study, the MG1655 Δ*flhE-che* strain had a 9.3 ± 0.1% (average of three independent experiments) larger cell mass compared to the MG1655 strain when grown on glucose minimal medium. In addition, wild-type *E. coli* has a regulatory system for reducing cell division upon serine depletion, but mutants of the *flhD* gene, which is in the deleted region, are insensitive to this regulation and undergo five more cycles of division before entering the stationary phase (41). This might have contributed to the advantageous phenotype, as serine was one of the preferentially consumed amino acids and was depleted within 2 h in the parental strain culture (Supplementary Figure S8).

The *pyrE* gene encodes orotate phosphoribosyltransferase for pyrimidine nucleotide synthesis. This activity is reportedly limiting in *E. coli* MG1655, and its increased expression improves growth rate on glucose and glycerol (42). Amplification of the *rph-pyrE* operon or the *pyrE* gene alone by a plasmid in the MG1655 strain improved growth (Supplementary Figure S9). The *pyrE** mutation is likely to be responsible for the increase in the evolved strain *pyrE* transcript, as previously suggested (15,43).

DISCUSSION

The present study examined the adaptive experimental evolution of *E. coli* populations driven by an RM system. We identified sequential changes in both the genome and the transcriptome during this process, and partially reconstructed the adaptive phenotype by introducing mutations into the parental strain.

We showed that experimental evolution was accelerated by an RM system. Chemical mutagens and other mutagen such as UV irradiation are known to increase mutation rate. In breeding, they are used under conditions that induce mutations at a much higher frequency. However, they would also induce deleterious mutation that would slow down bacterial growth. These mutagens might also
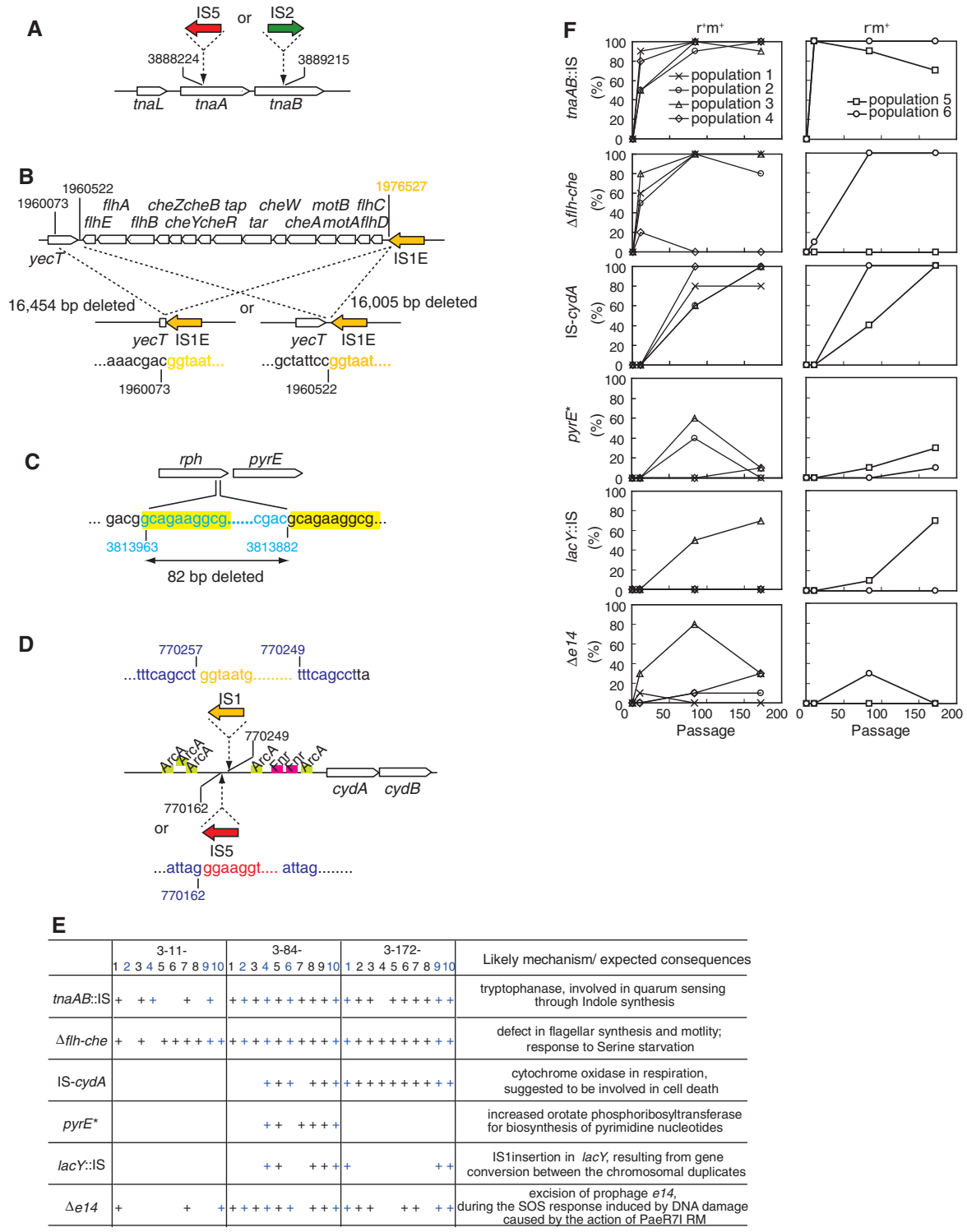
Figure 5. Genome changes. *tnaAB*::IS, Δ*flh-che* and IS-*cydA* showed variation in IS elements and positions of insertions as shown in (**A**, **B** and **D**). Variants scored as a mutant form in (**E** and **F**). (**A**) *tnaAB*::IS. Insertion of IS element into *tnaLAB* operon for synthesis of indole. Two variants were found: IS2 insertion into *tnaB* (clones 3-11-1, 3, 4, 7, 9 and 3-84-1∼5, 7∼10), and IS5 insertion into *tnaA* (3-84-6). (**B**) Δ*flh-che*. Deletion of region including genes involved in flagella synthesis and chemotaxis, resulting from IS1E-mediated adjacent deletion. From middle of *yecT* to IS1E, 16 454-bp deletion (clones 3-11-1,3, 5, 6, 7, 8, 9, 10 and 3-84-1–10), and 16 055-bp deletion from upstream of *flhE* to IS1E (clones 6-11-7 and 6-84-1, 2) were found. (**C**) *pyrE**. In *rph-pyrE* operon, 82-bp deletion (light blue) lies within *rph* gene and upstream of *pyrE* attenuator, and probably resulted from recombination involving two 10-bp regions of sequence identity (yellow box). (**D**) IS-*cydA*. Insertion of IS element upstream of *cydAB* operon. IS1 insertion (3-84-4, 5, 8, 9, 10), and IS5 insertion (3-84-6) were found. (**E**) Distribution of six mutations among clones of population 3. Representative clones used in transcriptome and further analysis shown in blue. (**F**) Temporal and inter-population distribution of six mutations. Ten clones analyzed for each population to obtain frequency.
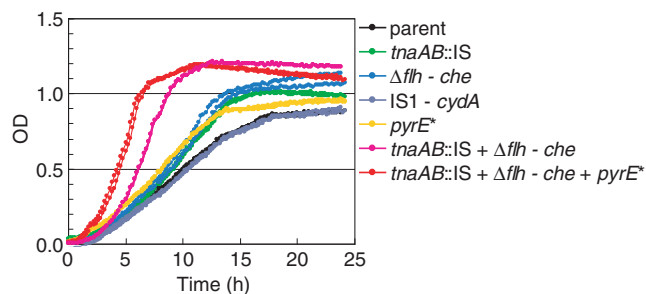
Figure 6. Growth of strains constructed by introduction of identified mutations into parental strain. Results of two independent experiments are shown.

be used to moderately accelerate experimental evolution in a similar procedure, but bacterial resistance to these mutagens is usually increased through the process of mutation and selection (44). Sensitivity of a bacterial population to UV irradiation has been reported to be decreased in 80 cycles of irradiation and growth (45). RM systems are expected to be stably maintained due to their selfish behavior (21). Indeed, the r^+ phenotype was stably maintained until the end of the present experiment (passage 172). Mutations that make the host genome less sensitive to the restriction attack might take place, which could limit the potential for further adaptive evolution.

Most of those mutagens are known to generate DNA damages and induce special DNA polymerases, which generate mutations at these damages in the act of translesion synthesis (46). The *paeR7I* RM system induces genes involved in DNA double-strand break repair and other interactions with DNA as well as genes involved in programmed death (24). These might generate and select a variety of genome rearrangements in addition to base substitutions, though r^+ dependence of the mutation spectrum is to be more precisely addressed.

The overall pattern of the evolutionary changes in the phenotype was similar to those reported earlier (26,47); the increase in the growth rate was not constant throughout the experimental period but rather stepwise (Figure 2). Extent of increase in the growth advantage, which is realized by a new mutation interacting with the inherited genetic background, should determine how the mutation is sustained in the population.

We analyzed the dynamic changes in transcriptome during the adaptive evolution by using the clones isolated at various stages of the evolution experiment and found that transcriptional change was uni-directional in terms of cellular function. There could be many possible combinations of multiple mutations to realize movement in one direction.

In the analysis of growth-phenotype improvements during adaptation, we found that they likely relied upon the coordinated action of multiple mutations and transcriptional changes related to cell–cell communication, programmed cell death, amino acid deprivation response, cell mass production and energy efficiency.

Modifications of genes regulating the ratio of active cells in a population, including the cell death program, were of particular interest, because this implied that such

mechanisms should be less activated for improved population growth during a relatively short period. However, they may be important for maintenance of a population over a longer period (48).

Microbes have been used in a wide range of fields, including food, medicine and the environment, to yield products such as amino acids, antibiotics and fuels. Their breeding has long been exploited and metabolic engineering has been successfully applied to optimize each step of the metabolic pathways related to products of interest (49–51). From the viewpoint of biotechnology, it is important to optimize the activity of an entire working population for certain periods of time under expected conditions, together with finely optimizing metabolisms; this requires genome modifications with additive/synergistic effects, including those that are silent by themselves. Here, we demonstrated that the sequential genome changes that occurred during adaptive evolution provide important information for re-engineering genomes.

Our approach could be used to rapidly rearrange and refine a genome to create a beneficial phenotype, and to understand the underlying genetic mechanisms for genome engineering applications. It could also be used for various microbes, cell lines and other organisms, particularly those with industrial applications. These may include organisms with synthetic (52) and chimeric genomes (53) carrying desired genetic elements, which have been created by recent advances in synthetic biology but might be maladapted to their industrial environments.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Tyo,K.E.J., Kocharin,K. and Nielsen,J. (2010) Toward design-based engineering of industrial microbes. *Curr. Opin. Microbiol.*, **13**, 255–262.
2. Ideker,T., Galitski,T. and Hood,L. (2001) A new approach to decoding life: systems biology. *Annu. Rev. Genomics. Hum. Genet.*, **2**, 343–372.
3. Glass,J.I., Assad-Garcia,N., Alperovich,N., Yooseph,S., Lewis,M.R., Maruf,M., Hutchison,C.A. III, Smith,H.O. and

Venter,J.C. (2006) Essential genes of a minimal bacterium. *Proc. Natl Acad. Sci. USA*, **103**, 425–430.

4. Posfai,G., Plunkett,G. III, Feher,T., Frisch,D., Keil,G.M., Umenhoffer,K., Kolisnychenko,V., Stahl,B., Sharma,S.S., de Arruda,M. *et al.* (2006) Emergent properties of reduced-genome *Escherichia coli. Science*, **312**, 1044–1046.

5. Hashimoto,M., Ichimura,T., Mizoguchi,H., Tanaka,K., Fujimitsu,K., Keyamura,K., Ote,T., Yamakawa,T., Yamazaki,Y., Mori,H. *et al.* (2005) Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. *Mol. Microbiol.*, **55**, 137–149.

6. Homma,K., Fukuchi,S., Nakamura,Y., Gojobori,T. and Nishikawa,K. (2007) Gene cluster analysis method identifies horizontally transferred genes with high reliability and indicates that they provide the main mechanism of operon gain in 8 species of gamma-Proteobacteria. *Mol. Biol. Evol.*, **24**, 805–813.

7. Naas,T., Blot,M., Fitch,W.M. and Arber,W. (1994) Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. *Genetics*, **136**, 721–730.

8. Fischer,G., James,S.A., Roberts,I.N., Oliver,S.G. and Louis,E.J. (2000) Chromosomal evolution in Saccharomyces. *Nature*, **405**, 451–454.

9. Elena,S.F. and Lenski,R.E. (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.*, **4**, 457–469.

10. Ibarra,R.U., Edwards,J.S. and Palsson,B.O. (2002) *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature*, **420**, 186–189.

11. Cooper,T.F., Rozen,D.E. and Lenski,R.E. (2003) Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli. Proc. Natl Acad. Sci. USA*, **100**, 1072–1077.

12. Fong,S.S., Joyce,A.R. and Palsson,B.O. (2005) Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res.*, **15**, 1365–1372.

13. Papadopoulos,D., Schneider,D., Meier-Eiss,J., Arber,W., Lenski,R.E. and Blot,M. (1999) Genomic evolution during a 10,000-generation experiment with bacteria. *Proc. Natl Acad. Sci. USA*, **96**, 3807–3812.

14. Barrick,J.E., Yu,D.S., Yoon,S.H., Jeong,H., Oh,T.K., Schneider,D., Lenski,R.E. and Kim,J.F. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli. Nature*, **461**, 1243–1247.

15. Herring,C.D., Raghunathan,A., Honisch,C., Patel,T., Applebee,M.K., Joyce,A.R., Albert,T.J., Blattner,F.R., van den Boom,D., Cantor,C.R. *et al.* (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat. Genet.*, **38**, 1406–1412.

16. Kobayashi,I. (2001) Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res.*, **29**, 3742–3756.

17. Ishikawa,K., Fukuda,E. and Kobayashi,I. (2010) Conflicts targeting epigenetic systems and their resolution by cell death: novel concepts for methyl-specific and other restriction systems. *DNA Res.*, **17**, 325–342.

18. Fukuda,E., Kaminska,K.H., Bujnicki,J.M. and Kobayashi,I. (2008) Cell death upon epigenetic genome methylation: a novel function of methyl-specific deoxyribonucleases. *Genome Biol.*, **9**, R163.

19. Kita,K., Kawakami,H. and Tanaka,H. (2003) Evidence for horizontal transfer of the EcoT38I restriction-modification gene to chromosomal DNA by the P2 phage and diversity of defective P2 prophages in *Escherichia coli* TH38 strains. *J. Bacteriol.*, **185**, 2296–2305.

20. Naito,T., Kusano,K. and Kobayashi,I. (1995) Selfish behavior of restriction-modification systems. *Science*, **267**, 897–899.

21. Handa,N., Nakayama,Y., Sadykov,M. and Kobayashi,I. (2001) Experimental genome evolution: large-scale genome rearrangements associated with resistance to replacement of a chromosomal restriction-modification gene complex. *Mol. Microbiol.*, **40**, 932–940.

22. Guyer,M., Reed,R.E., Steitz,T. and Low,K.B. (1981) Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spr. Harb. Symp. Quant. Biol.*, **45**, 135–140.

23. Korona,R. and Levin,B.R. (1993) Phage-mediated selection and the evolution and maintenance of restriction-modification. *Evolution*, **47**, 556–575.

24. Asakura,Y. and Kobayashi,I. (2009) From damaged genome to cell surface: transcriptome changes during bacterial cell death triggered by loss of a restriction-modification gene complex. *Nucleic Acids Res.*, **37**, 3021–3031.

25. Kohanski,M.A., Dwyer,D.J., Hayete,B., Lawrence,C.A. and Collins,J.J. (2007) A common mechanism of cellular death induced by bactericidal antibiotics. *Cell*, **130**, 797–810.

26. Woods,R.J., Barrick,J.E., Cooper,T.F., Shrestha,U., Kauth,M.R. and Lenski,R.E. (2011) Second-order selection for evolvability in a large *Escherichia coli* population. *Science*, **331**, 1433–1436.

27. Greener,A. and Hill,C.W. (1980) Identification of a novel genetic element in *Escherichia coli* K-12. *J. Bacteriol.*, **144**, 312–321.

28. Hill,C.W., Gray,J.A. and Brody,H. (1989) Use of the isocitrate dehydrogenase structural gene for attachment of *e14* in *Escherichia coli* K-12. *J. Bacteriol.*, **171**, 4083–4084.

29. Chattoraj,D.K. (2007) Tryptophanase in sRNA control of the *Escherichia coli* cell cycle. *Mol. Microbiol.*, **63**, 1–3.

30. Chant,E.L. and Summers,D.K. (2007) Indole signalling contributes to the stable maintenance of *Escherichia coli* multicopy plasmids. *Mol. Microbiol.*, **63**, 35–43.

31. Hayashi,K., Morooka,N., Yamamoto,Y., Fujita,K., Isono,K., Choi,S., Ohtsubo,E., Baba,T., Wanner,B.L., Mori,H. *et al.* (2006) Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol. Syst. Biol.*, **2**, 0007.

32. Withers,H.L. and Nordstrom,K. (1998) Quorum-sensing acts at initiation of chromosomal replication in *Escherichia coli. Proc. Natl Acad. Sci. USA*, **95**, 15694–15699.

33. Park,S., Wolanin,P.M., Yuzbashyan,E.A., Silberzan,P., Stock,J.B. and Austin,R.H. (2003) Motion to form a quorum. *Science*, **301**, 188.

34. Kolodkin-Gal,I., Hazan,R., Gaathon,A., Carmeli,S. and Engelberg-Kulka,H. (2007) A linear pentapeptide is a quorum-sensing factor required for *mazEF*-mediated cell death in *Escherichia coli. Science*, **318**, 652–655.

35. Lee,H.H., Molla,M.N., Cantor,C.R. and Collins,J.J. Bacterial charity work leads to population-wide resistance. *Nature*, **467**, 82–85.

36. Strauss,B., Kelly,K., Dincman,T., Ekiert,D., Biesieda,T. and Song,R. (2004) Cell death in *Escherichia coli dnaE*(Ts) mutants incubated at a nonpermissive temperature is prevented by mutation in the *cydA* gene. *J. Bacteriol.*, **186**, 2147–2155.

37. Davies,B.W., Kohanski,M.A., Simmons,L.A., Winkler,J.A., Collins,J.J. and Walker,G.C. (2009) Hydroxyurea induces hydroxyl radical-mediated cell death in *Escherichia coli. Mol. Cell*, **36**, 845–860.

38. Kohanski,M.A., Dwyer,D.J. and Collins,J.J. (2010) How antibiotics kill bacteria: from targets to networks. *Nat. Rev. Microbiol.*, **8**, 423–435.

39. Cotter,P.A., Melville,S.B., Albrecht,J.A. and Gunsalus,R.P. (1997) Aerobic regulation of cytochrome d oxidase (*cydAB*) operon expression in *Escherichia coli*: roles of Fnr and ArcA in repression and activation. *Mol. Microbiol.*, **25**, 605–615.

40. Macnab,R.M. (1996) Flagella and Motility. In Ingraham,J.L. and Neidhardt,F.C. (eds), Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*. American Society for Microbiology, Washington, DC, pp. 123–145.

41. Pruss,B. and Matsumura,P. (1996) A regulator of the flagellar regulon of *Escherichia coli*, *flhD*, also affects cell division. *J. Bacteriol.*, **178**, 668–674.

42. Jensen,K.F. (1993) The *Escherichia coli* K-12 'wild types' W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels. *J. Bacteriol.*, **175**, 3401–3407.

43. Conrad,T.M., Joyce,A.R., Applebee,M.K., Barrett,C.L., Xie,B., Gao,Y. and Palsson,B.O. (2009) Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol.*, **10**, R118.

44. Wright,S.J. and Hill,E.C. (1968) The development of radiation-resistant cultures of *Escherichia coli* I by a process of 'growth-irradiation cycles'. *J. Gen. Microbiol.*, **51**, 97–106.

45. Alcantara-Diaz,D., Brena-Valle,M. and Serment-Guerrero,J. (2004) Divergent adaptation of *Escherichia coli* to cyclic ultraviolet light exposures. *Mutagenesis*, **19**, 349–354.
46. Jarosz,D.F., Beuning,P.J., Cohen,S.E. and Walker,G.C. (2007) Y-family DNA polymerases in *Escherichia coli*. *Trends Microbiol*, **15**, 70–77.
47. Lenski,R.E. (1995) Evolution in experimental populations of bacteria. In Baumbergs,S. (ed.), *Proceedings of the Population Genetics of Bacteria: Fifty-second Sysmposium of the Society for General Microbiology*. Press Syndicate of the University of Cambridge, New York, NY, USA, pp. 195–218.
48. Rosenberg,S.M. (2009) Life, death, differentiation, and the multicellularity of bacteria. *PLoS genetics*, **5**, e1000418.
49. Kimura,E. (2003) Metabolic engineering of glutamate production. *Adv. Biochem. Eng. Biotechnol.*, **79**, 37–57.
50. Asakura,Y., Kimura,E., Usuda,Y., Kawahara,Y., Matsui,K., Osumi,T. and Nakamatsu,T. (2007) Altered metabolic flux due to deletion of *odhA* causes L-glutamate overproduction in *Corynebacterium glutamicum*. *Appl. Env. Microbiol.*, **73**, 1308–1319.
51. Santos,C.N. and Stephanopoulos,G. (2008) Combinatorial engineering of microbes for optimizing cellular phenotype. *Curr. Opin. Chem. Biol.*, **12**, 168–176.
52. Gibson,D.G., Glass,J.I., Lartigue,C., Noskov,V.N., Chuang,R.Y., Algire,M.A., Benders,G.A., Montague,M.G., Ma,L., Moodie,M.M. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**, 52–56.
53. Itaya,M., Fujita,K., Kuroki,A. and Tsuge,K. (2008) Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat. Methods.*, **5**, 41–43.