Distinctive survival strategies, specialized in regulation and in quality control, were observed in thermal adaptive evolution with a laboratory *Escherichia coli* strain. The two specialists carried a single mutation either within rpoH or upstream of groESL, which led to the activated global regulation by sigma factor 32 or an increased amount of GroEL/ES chaperonins, respectively. Although both specialists succeeded in thermal adaptation, the common winner of the evolution was the specialist in quality control, that is, the strategy of chaperonin-mediated protein folding. To understand this evolutionary consequence, multilevel analyses of cellular status, for example, transcriptome, protein and growth fitness, were carried out. The specialist in quality control showed less change in transcriptional reorganization responding to temperature increase, which was consistent with the finding of that the two specialists showed the biased expression of molecular chaperones. Such repressed changes in gene expression seemed to be advantageous for long-term sustainability because a specific increase in chaperonins not only facilitated the folding of essential gene products but also saved cost in gene expression compared with the overall transcriptional increase induced by rpoH regulation. Functional specialization offered two strategies for successful thermal adaptation, whereas the evolutionary advantageous was more at the points of cost-saving in gene expression and the essentiality in protein folding.

**Introduction**

Investigating evolutionary principles of cellular adaptation is critical for understanding how the molecular mechanisms re-establish cellular homeostasis (Kultz 2005), which is maintained by upstream level control as gene expression regulation and downstream control as protein quality control. Both stress response to short-term perturbations and evolutionary adaptation to long-term environmental changes occur at the levels of these two controls. Although fitness recovery responding to changing environments is a common feature in adaptation, few studies have reported how molecular mechanisms functionally differentiated in evolution and contribute to improved fitness. As a representative case, living cells, in response to thermal stress, are mediated by two molecular mechanisms: regulation at the transcription level and quality control at the protein level (Mogk et al. 2011). The collaboration between the two mechanisms directs the cells to achieve a balanced physiological state and helps them survive environmental stress. A well-known example is the feedback mechanism attributed to sigma factor 32 (the gene product of rpoH), working as a transcriptional regulator (Yura et al. 1993), and the chaperonins (the gene products of groESL),
assisting in protein quality control (Hartl & Hayer-Hartl 2002). A negative feedback mechanism involved with the sigma factor 32 (Guisbert et al. 2008) and the chaperonins plays a crucial role in maintaining protein homeostasis in living cells. Whether evolutionary adaptation influences the two closely related mechanisms involved in the stress response is unknown.

Functional specialization could be happened in response to external disturbance. It is intriguing whether the evolutionary specialization would happen to the central mechanisms and whether there is any bias between them. Thus far, few studies have reported how the mechanisms of regulation and quality control contribute to fitness changes in an evolutionary view, whereas numerous studies have investigated the details of the individual mechanisms (e.g., ref. in (Hartl et al. 2011; Mogk et al. 2011)). Because an increased amount of either σ32 or chaperonins could perturb the feedback mechanism (Guisbert et al. 2008), the survival strategy initiated from either regulation or quality control may guide the cells for better growth. It is unknown whether there is any bias between the mechanisms when the cells face thermal stress. Experimental investigations are required to link the molecular mechanisms to fitness improvement in cellular life.

We analyzed thermally adapted *Escherichia coli* cells generated by thermal adaptive evolution carried out in the laboratory (Kishimoto et al. 2010). Transcriptome analysis of the endpoint strains found that the thermal adaptive evolution did not disturb the responsivity to heat shock but alter the gene expression at the steady states under both regular and high temperatures (Ying et al. submitted). This finding further triggered us to investigate the evolutionary dynamics of the cell populations. Fruitful stocks of the *E. coli* cells showing varied growth fitness and carrying different gene mutations were acquired. Among them, two genotypes with single mutations, either within *rpoH* or upstream of *groESL*, were found. An evolutionary competition between the two genotypes was investigated to examine both regulation and quality control. The two distinguished survival strategies caused only a slight difference in both cellular conditions (transcriptional reorganization and protein activity) and growth fitness but resulted in a significant difference in fate decision. The final choice preferred by the *E. coli* cells somehow indicated more economical benefits from the quality control at the protein level than the regulation for global reorganization in gene expression.

**Results and discussion**

**Genetic differentiation in regulation or quality control**

The single-cell isolation of the previously evolved *E. coli* cell populations in the rapid growth recovery phase (details in Experimental procedures) showed that the functional specialization in regulation and quality control was happened in the thermal adaptive evolution (Fig. S1 in Supporting Information). Genome resequencing analysis of four different single clones (Table S1 in Supporting Information) identified two genotypes in the cell populations in the early period of this rapid growth recovery phase (Figs 1A and S1 in Supporting Information). The two genotypes carried only a single different mutations: one was a nonsynonymous single nucleotide substitution in *rpoH* (type II, Fig. S2A in Supporting Information), and the other was a 5-bp deletion at the 5′-UTR of the *groESL* operon (type III, Fig. S2B in Supporting Information). As a result, the activity or abundance of σ32 (the gene product of *rpoH*) and the expression of chaperonins (GroEL and GroES) were supposed to be altered (Morigita et al. 1999a; Nonaka et al. 2006). Note that extensive single-cell isolation failed to acquire a genotype that contained both mutations.

Population dynamics based on the ratios of the nucleotide substitution determined by the Sanger method directly toward the stocked cell populations showed that a competition between them had occurred during the targeted evolutionary period (Fig. 1A, bottom). Growth tests showed that both type II and type III cells could grow independently at both regular and high temperatures (Fig. 1B C). The commonly increased growth fitness verified that both genotypes did adapt to the high temperature and indicated that both mutations were beneficial. Thus, the functional specialization of the *E. coli* cells in evolution shared the common purpose of fitness improvement relied on the changes in the molecular mechanisms of transcriptional regulation by *rpoH* and protein folding.

Nevertheless, type III specialized in quality control but not type II specialized in regulation won the evolutionary competition, although type II had transiently occupied the cell population more than type III (Fig. 1A, bottom). We noticed that type III cells
showed slightly improved fitness at the high temperature, that is, slightly shorter lag time (Fig. 1B) and slightly faster exponential growth at 45 °C, despite their equivalent growth rates at 37 and 43 °C (Fig. 1C). It implied that the thermal adaptive evolution preferred the specialist in quality control, although the functional specialization successfully born two specialists both adaptive to high temperatures. To understand this evolutionary consequence, the cellular status was investigated.

Differentially expressed genes in response to temperature increase

As the growth fitness largely correlated to the gene expression (Lopez-Maury et al. 2008; Matsumoto et al. 2013), transcriptome analyses of exponentially growing cells were carried out. Microarray analyses of cells growing at regular and high temperatures were carried out (Fig. S3 in Supporting Information) as previously described (Matsumoto et al. 2013; Ying...
et al. 2013). The two genotypes had the highly similar gene expression patterns at the identical growth temperatures and commonly differentiated in response to temperature increase (Fig. 2A). Rank product analysis (Breitling et al. 2004) showed that a total of 158 and 128 genes presented significant transcriptional changes (differentially expressed genes, DEGs, FDR < 0.05) in response to the raised growth temperature in the type II and III cells, respectively (Fig. 2B, details in Table S2 in Supporting Information). In addition to 88 overlaps, there were more DEGs in type II (70 genes) cells than in type III (40 genes) cells, indicating somehow more significant changes in gene expression in type II.

Enriched gene categories (Riley et al. 2006) exhibited the distinguished functions in the DEGs (Fig. 2C). The type II cells reorganized the genes functioning in sensing the environmental changes, such as the membrane and transport, but the type III cells reorganized genes functioning in RNA (most were tRNAs, particularly met-tRNAs and leu-tRNAs), which were responsible for translation.

Supplemental enrichment analysis showed that both the class III substrates of GroEL (Kerner et al. 2005; Fujiwara et al. 2010) and the gene products of intermediate solubility (40–50%) (Niwa et al. 2009) significantly (P < 0.05) presented in the overlaps (Fig. S4 in Supporting Information). Changes in the gene expression of the aggregate prones at the high temperature seemed to be common in both genotypes. These results implied that type II reorganized the expression of the genes sensitive to the environment but type III to the intracellular reaction, although both genotypes presented significant changes in gene expression when growing at the high temperature.

**Similarity and differentiation in the transcriptional changes of regulatory networks**

Besides the individual genes, the genes controlled by a certain transcriptional factor (TF) that presented significant reorganization were further determined using gene set enrichment analysis (GSEA) (Subramanian

---

**Figure 2** Differentially expressed genes (DEGs) in the two genotypes. (A) Dot plots of the gene expression. Gene expression of the two genotypes exponentially grown at the regular and high temperatures is plotted. II_37, II_45, III_37 and III_45 represent type II growing at 37 and 45 °C, type III growing at 37 and 45 °C, respectively. Gene expression is shown in the log-scale of mRNA concentration. The plots holding the red and blue circles are in correspondence to the Venn diagram in B. Correlation coefficients are indicated. (B) Venn diagram of DEGs commonly and specifically determined in the two genotypes. The numbers of DEGs of up- or down-regulation are indicated. Red and blue represented type III and type II, respectively. The details are described in the main text. (C) Differentiated gene categories. The significance of the DEGs enriched in the gene categories is shown. II, III and com represent the DEGs specifically in type II and type III, and common in both genotypes, respectively. The log-scale P-values are displayed as a heat map with gradations from dark blue to vivid yellow.
et al. 2005) (Table S3 in Supporting Information). Both genotypes showed roughly similar patterns of global reorganization when growth temperature increased, that is, the direction of transcriptional changes was comparable (Fig. 3). For example, the genes regulated by nac and purR were commonly repressed at the high temperature in both genotypes. However, the significantly induced and repressed TFs were different between them. Type III showed no significant up-regulated transcriptional networks but more repressed ones, such as those controlled by gadX and gadE. In contrast, type II cells showed more gene groups of significantly induced expression, such as those regulated by narP and fhlA, which were reported to be activated in the transient and short-term response to temperature increase (Matsumoto et al. 2013; Ying et al. 2013). The highly significant induction of these TFs in type II indicated that the cellular status was closer to the heat-shock state.

As type II harbored a mutation in rpoH (Fig. S2A in Supporting Information), how this mutated rpoH influenced the gene expression was further investigated. We found that the mutation in rpoH increased the protein abundance of itself (Fig. 4A) without changes in transcriptional level (Fig. S5 in Supporting Information). It indicated that the single nucleotide substitution in rpoH led to the changes in the binding affinity to DnaK/J (Gamer et al. 1992, 1996; Rodriguez et al. 2008). In a global view, an equal change in gene expression would lead to a slope of 1.0 between the two genotypes (Fig. 4B, broken line). As the slope of the linear regression became slightly smaller than 1.0 (Fig. 4B, pale pink), a relatively higher magnitude (~14%) of transcriptional fluctuation might occur in type II cells growing at the high temperature. As this increased changes in either the rpoH-regulated or the total genes (Fig. 4B) were slight, a rough estimation of the transcriptional changes in the total 295 genes regulated by rpoH was additionally carried out. It showed that larger changes were happened in type II than in type III (Fig. 4C). Taken together, the type II cells presented more significant changes than type III cells did, consistent with the DEG analysis (Fig. 2B), although both types showed somehow similarity in the directional changes when the growth temperature increased.

### Biased abundance of molecular chaperones for differentiated evolutionary directions

Such differentiated transcriptional reorganization was more evident in the major chaperones. Changes in the gene expression of the major chaperones clearly illustrated a compensative effect of the chaperonins (GroESL). The transcription of groL and groS was highly activated in type III cells, and other chaperones were activated in type II cells at both regular and high temperatures (Fig. 5A), although the mechanisms of heat-induced gene expression maintained regular in both genotypes (Fig. S5 in Supporting Information). The replacement of chaperones by chaperonins was additionally verified at the protein level. Western blot analysis clearly showed that the increase in GroEL (Fig. 5C) was accompanied by a decrease in DnaK (Fig. 5B) at the high temperature; nevertheless, their amounts were equivalent at the

---

**Figure 3** Fluctuation in regulatory networks. The regulators (transcriptional factors, TFs), whose downstream genes show significant fluctuation in response to raised temperature, are represented in dark purple (down-regulated) and dark green (up-regulated). The evaluation was based on GSEA. Highly significant TFs (P < 0.001) are indicated as white letters in colored circles.
The results not only confirmed that the 5-bp deletion upstream of \textit{groESL} (Fig. S2B in Supporting Information) caused the highly constitutive expression of chaperonins but also suggested that the increased amount of chaperonins compensated for other chaperones for quality control.

As the gene expression pattern of type II growing exponentially at the high temperature seemed to be close to the well-known heat-shock state, the expression levels of the representative chaperonins in the ancestor at the regular and the heat-shock conditions were evaluated and the transcriptional levels of the chaperonins under different states were compared (Fig. 5D). The expression levels of \textit{groL} and \textit{dnaK} were positively correlated to that of \textit{groS} and \textit{dnaJ}, respectively, and maintained the equivalent ratios independent of the genotypes or the growth conditions (Fig. 5D, middle and bottom). It indicated that the related regulatory mechanisms remained regular during the thermal adaptation. However, the ratio between \textit{dnaK} and \textit{groL} turned differentiated among the genotypes and growth conditions, that is, type II kept the ratio of \textit{groL}/\textit{dnaK} expression along the

Improved protein activity in the evolutionary direction of quality control

Whether such biased abundance of molecular chaperones caused the diversity in the cellular protein activity was further investigated. Although no significant difference was observed in the solubility of total protein between the two genotypes (Fig. 6A), better protein folding in type III cells was detected (Fig. 6B, C). Because the \textit{E. coli} cells carried a chromosomally incorporated reporter gene, \textit{gfp}, as a genetic marker, the folding efficiency of GFP could be evaluated by the green fluorescence using flow cytometry. The mean FITC/FS, used as a reference of cellular protein activity, was higher in type III than in type II cells \((P < 0.001)\) (Fig. 6C), regardless
of the comparable GFP abundance in both genotypes (Fig. 6B). The improved protein activity in type III was commonly detected in the cells growing at both the regular and high temperatures (Fig. 6C).

As GFP maturation relied on the chaperonins (Makino et al. 1997), this result triggered us to consider the chaperone dependency of protein folding. Type III cells relied more on essential genes because the GroEL substrates \( N = 250 \) (Kerner et al. 2005) presented significant enrichment in essential genes \( N = 42, P < 0.0001 \). Abundant chaperonins in type III specifically saved the essential proteins and may be beneficial in the competition against type II. In particular, the considerably high expression levels of the essential genes under all conditions (Fig. S6 in Supporting Information), consistent with previous reports (Ishihama et al. 2008), emphasized the essentiality of proper folding of these proteins. Thus, the highly induced chaperones in type II may prevent all misfolded proteins from aggregating, whereas type III specialized in those essential proteins.

**Figure 5** Biased abundance in chaperones. (A) Transcriptional changes in chaperones. Changes in gene expression of the major chaperones between the two genotypes growing at 37 °C (open) and 45 °C (filled) are shown. (B, C) Gene products of the representative chaperones. Repeated Western blot analyses were carried out to estimate the protein concentrations of DnaK and GroEL. Representative images of Western blot are shown (upper panels). The numbers of cells, varied from \( 5 \times 10^5 \) to \( 1 \times 10^7 \) cells, loaded on each lane are indicated. II_37, II_45, III_37 and III_45 represent type II growing at 37 and 45 °C, type III growing at 37 and 45 °C, respectively. The purified chaperones were used as the positive controls (the left lanes). The mean protein concentrations and standard errors \( N = 5 \) are shown (bottom panel). The color variation is as indicated. (D) Alternative strategies for more groL or more dnaK. Expression levels of the chaperones in the same genotypes are plotted on the y-axis and x-axis. Color variation represents the genotypes, as indicated. Open and closed circles indicate the expression levels of the cells exponentially growing at the regular and high temperatures, respectively, except the black circle, representing the heat-shock response. *P < 0.01.
As the analyses at the transcriptome and protein levels suggested the slight but significant priority of cellular status in type III, whether this evolutionary consequence, that is, type III as the final winner, was general, was confirmed by growth competition between the two genotypes. The competition tests showed that type III cells were the common winner, regardless of either the initial cell concentrations or the initial ratios between the two genotypes (Fig. 7). The results confirmed that type III cells were more viable at the high temperature.

Based on this common consequence, we proposed a cost-saving strategy for thermal adaptation based on the analytical results acquired in the present study (Fig. 8). Thermal adaptive evolution for better protein folding occurred in two directions for more GroEL or more DnaK (equivalent to rpoH). The direction of more GroEL triggered to a cost-saving path of gene expression for the cells growing at high temperatures.
temperature. Solely increasing chaperonins led to less change in global transcriptional reorganization than regulation. Type II cells fixed the transient induction of \( rpoH \), which largely influenced the expression of a large number of genes and enhanced the feed forward regulation. The type III genotype led to downstream changes within a limited number of genes, \( groL \) and \( groS \) only. This change improved both the feed forward and the feedback regulations, leading to a balanced change in cellular reactions. As the control of growth and stress response was balanced at gene expression cost in living cells (Lopez-Maury et al. 2008; Matsumoto et al. 2013), the cost-saving strategy in gene expression along with the essential folding may contribute to the slightly improved growth fitness and promise to win the evolutionary competition. The results at molecular levels agreed well with our previous findings of that the transcriptome evolution for thermal adaptation did not disturb the responsivity to temperature increase but fine-tune the steady states at both regular and high temperatures (submitted). The thermal adaptation preferring for quality control strategy maintained the highly functional responsive machinery and altered the downstream flux at protein level.

Two survival strategies specializing in regulation and quality control were used by the \( E. coli \) cells during thermal adaptive evolution, indicating evolutionary specialization with a functional division (Rueffler et al. 2012). A higher priority was observed on downstream quality control than upstream transcriptional regulation for long-term sustainability, likely due to the economics of maintaining cellular homeostasis. A global change in the transcription network (\( rpoH \) activated), comprising a number of genes, must be more energy-consuming than a localized change solely in chaperonin because gene expression (producing a new gene product) costs many more ATPs than does protein folding (Bogumil & Dagan 2012). Additionally, the special care of the essential proteins that occurred in type III may be a more efficient method for cells to overcome thermal stress. These results indicated that the survival strategy mediated by protein quality control facilitated cellular economy (Szekely et al. 2013).

**Experimental procedures**

**Strains, cell culture and growth rate**

The genetically engineered \( E. coli \) strain DH1DeluB::gfpuv5-kmr was used in the evolution experiments as previously described (Kishimoto et al. 2010). The two genotypes (II and III) were isolated during the 43B to 45A period (Fig. S1 in Supporting Information) and carried the mutations, as described in the Genome resequencing section. Cells were cultured in 5 mL of minimal medium M63 supplemented with 2 mM leucine (Wako) and 25 \( \mu \)g/mL of kanamycin sulfate (Wako). Cells were cultured using a water bath (Personal-11, Taitec) at 37, 43 and 45 °C and monitored using a platinum resistance thermometer 5615 (Fluke), which showed true temperatures of 36.9, 43.2 and 44.7 °C, respectively. Serial transfer was always carried out in the early exponential growth phase. The cell concentration was measured by flow cytometry as described in the FCM analysis section. The growth rate was calculated according to the following formula: \( \ln(C_t/C_0)/t \), where \( C_t \), \( C_0 \) and \( t \) represent the final and initial cell concentrations (cells/mL) within the exponential phase and the culture time (h), respectively.

**Genome resequencing and mutation analysis**

Four clones acquired from single-cell isolation were chosen for genome resequencing using a next-generation sequencer (Junior, Roche). Genomic DNA was purified and fragmented using the Covaris system as previously described (Kishimoto et al. 2010; Suzuki et al. 2011). Whole-genome shotgun sequencing by the 454 GS Junior platform (Roche) was applied according to the manufacturers’ instructions. The
sequence reads were assembled using Newbler 2.7 and aligned using the GS Reference Mapper software (ver. 2.6; Roche). Approximately 98% of the reads in each dataset were uniquely mapped to the DH1 genome (CP001637), and the statistics are summarized in Table S1 (Supporting Information). Mutation analysis was carried out using the GS Reference Mapper software. Candidates of mutations were detected as ‘HC (High Confidence) Differences’, ‘HC Structural Rearrangements’ and ‘HC Structural Variants’ by the software with the recommended parameter settings (system default). Mutation analysis confirmed that both genotypes carried the mutations reported in 43B and the three other mutations (oxyR, rnr, fre) reported in 45A. The analysis found that new mutations in lon occurred in both genotypes and that rpoH occurred solely in the type II background. In addition, the mutation upstream of groESL, which was known to be fixed in 45A, was detected in type III only. All mutations were verified by the Sanger method as previously reported (Kishimoto et al. 2010).

**Single-cell isolation**

Single-cell isolation of the thermal adaptive *E. coli* cells stocked during the evolution experiments was carried out to purify the genotypes among the populations. Cell concentrations of the stocked cell populations were measured by flow cytometry (as described in FCM analysis). The cell populations were diluted with the fresh media and inoculated to the 96-well plates (Corning) at concentrations of 0.1 or 0.05 cells/well and with 200 μL per well. The plates were incubated at 37 °C for 2 days. According to the Poisson distribution, the wells that became turbid were a result of single-cell-derived growth. Three different *E. coli* stocks were used for the single-cell isolation. A total of 2496 wells were inoculated, which led to a theoretical 192 wells with growing cells; 74 wells showed actual cell growth. Four of the wells were used for the genome resequencing analysis, and all others were examined by the Sanger method. We confirmed that the same genotypes

---

**Figure 8** Regulation and quality control in thermal adaptation. The hypothesis of the economy strategies of regulation and quality control is schematically drawn. In general, negative feedback mechanisms are involved with the regulator, sigma factor 32, and the major molecular chaperones, such as HSP70 (DnaK, DnaJ and GrpE) and HSP60 (GroEL and GroES) systems. The potential contributions of the mutations that occurred in type II and type III are highlighted in blue and red, respectively. The functions of activation and repression are illustrated as arrows and T-marks, respectively. Regulation initiated by rpoH causes transcriptional changes related to a cluster of genes (Nonaka et al. 2006). rpoH is a RNA thermometer (Morita et al. 1999b), and high temperatures initiate its translation, which in turn activates the transcription of downstream genes, such as heat-shock genes (Yura & Nakahigashi 1999). Regulation by rpoH affects the early steps in gene expression (protein biosynthesis) and contributes to a global gene network. Quality control by the chaperonins influences the final steps in protein biosynthesis (Hartl & Hayer-Hartl 2002). High temperatures (thermal stress) accelerate the binding of the chaperonins to the newly synthesized proteins that are aberrantly folded or aggregated (Guisbert et al. 2008). As a result, the chaperonins facilitate successful protein folding and trigger the cells into an adaptive state. The downstream optimization at the protein level facilitates the final growth fitness with economical cost in reorganization of gene expression.
showed identical growth fitness, despite being isolated from diverse populations.

**Population dynamics**

The mutations in *rpoH* and the *groESL* operon were detected by genome resequencing analysis using a next-generation sequencer (Junior, Roche) and further verified by the Sanger method (Applied Biosystems). Further details of the genome resequencing are supplied in Table S1 (Supporting Information). Single-cell isolation based on a Poisson distribution was carried out to isolate the two genotypes from the *E. coli* population that was stocked during the evolution experiment. The daily changes in population dynamics were determined according to the ratio of the peak values representing the wild-type and the substituted nucleotides in *rpoH*, as well as that of the peak values of G/C derived from the 5-bp deletion in the *groESL* operon, by the Sanger method.

**Growth competition**

The conditions of cell culture, measurement of cell concentrations and calculation of growth rates were carried out as previously described (Kishimoto et al. 2010; Ying et al. 2013). Exponentially growing cells were transferred to the prewarmed fresh media with ratios of 1:1 and 1:100 between the two genotypes. Cocultures of the mixed cell populations at 45 °C were carried out and serially transferred every 24 h for 5–7 passages. Every two lineages of each coculture (ratio) were carried out, with initial concentrations of both 10⁶ and 10⁷ cells/mL.

**FCM analysis**

Cell concentrations and cellular GFP folding efficiency were evaluated by a flow cytometer (CantoII; Becton, Dickinson and Company) equipped with a 488-nm argon laser and a 515–545-nm emission filter (FITC). The following PMT voltage settings were applied: forward scatter (FSC), 400; side scatter (SSC), 400; and FITC, 600. The flow rate for the sample measurements was set to ‘low.’ Cell concentrations were calculated according to the ratio of gated particles representing the number of *E. coli* cells carrying the reporter gene *gfp* and beads of known concentrations as previously described (Ying et al. 2014). The FITC intensity was normalized with the rainbow calibration particles (eight peaks, 3.0–3.4 μm) to avoid the noise or changes in the sensitivity of the laser in daily measurements. The mean value of the logarithmic ratio between FITC and FSC was used as the folding efficiency of GFP in exponentially growing cells (approximately 9000 cells per measurement). Repeated measurements of independent cultures on different days were taken, and the averaged mean values are presented in the main text.

**Western blot analysis**

The purified GFP was kindly provided by Dr. Kazuda (ERATO, JST), and the other purified proteins and antibodies were commercially available. The purified GroEL and DnaK proteins and the antibodies to GroEL, DnaK and σ32 were all purchased from Abcam, and the antibody to GFP was purchased from Funakoshi. The anti-mouse and anti-rabbit IgG-HRP conjugates were purchased from R&D Systems. The proteins were detected with the primary antibodies (0.1–0.5 μg/mL) derived from mice or rabbits, followed by the relevant IgG-HRP conjugates (1000– to 5000-fold dilution). After development with a chemiluminescent substrate (Advance Western Blotting Detection Kit, GE Healthcare), the proteins were detected using a luminescent imaging analyzer (Image Quant-t350, GE Healthcare). The SDS-PAGE was carried out with a precast gel (12% Mini-PROTEAN TGX Precast Gel, Bio-Rad), and the blotting was carried out with a semidy system (Trans-Blot Turbo Transfer System and mini PVDF Transfer Pack, Bio-Rad). The standard curves for quantitatively evaluation of cellular GFP, GroEL and DnaK ranged from 1 to 50, 5 to 75 and 25 to 200 ng per lane. The number of cells loaded per lane was varied from 1 × 10⁵ to 1 × 10⁷ cells according to the preliminary estimated sensitivity in detection.

**Protein solubility and activity**

Exponentially growing cells were collected by centrifugation at 10 000 g for 1 min. The final cell concentrations were all controlled at approximately 1 × 10⁸ cells/mL. The cell pellets were suspended with PBS buffer and sonicated. Soluble and insoluble proteins were separated by centrifugation at 16 000 g for 10–20 min at 4 °C. SDS-PAGE of the total (T), soluble (S) and insoluble (P) proteins was carried out. Repeated measurements of independent cultures were taken to calculate the total protein solubility, represented by the ratio of band intensities between S and S+P and evaluated using Image J. The protein activity was evaluated according to the FITC intensity of the *E. coli* cells and measured by flow cytometry (FCM). Three to five replicates were carried out to calculate the protein activity, and six to eight replicates were carried out to calculate the protein amounts.

**Microarray**

Exponentially growing cells (~10⁸ cells/mL) were collected for the analyses. The sample preparation of total RNA, microarray analysis with the Affymetrix GeneChip system and data extraction were carried out as previously described (Matsumoto et al. 2013; Ying et al. 2013). The results of the 12 arrays for the two strains under the two conditions were used (three replicates of exponential growth at both regular and high temperatures). The transcription levels were determined as the log-scale mRNA concentrations (pM). A total of 3140 genes were used for the analyses after removing those of low expression under the reliable detection level (P < 0.5), as previously described (Murakami et al. 2015). The expression data were deposited in the NCBI Gene Expression Omnibus database under accession number of GSE52770 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE52770).
Computational analyses

Gene set enrichment analysis (GSEA), which is an analytical approach to determine whether a group of genes has significant changes compared with other gene groups, was carried out according to the original report (Subramanian et al. 2005) using the available online tools (http://www.broadinstitute.org/gsea/index.jsp), as previously described (Matsumoto et al. 2013; Ying et al. 2013). The transcription and sigma factor gene regulation datasets were obtained from RegulonDB 8.0 (Salgado et al. 2012) (http://regulondb.ccg.unam.mx/index.jsp). The input gene lists for GSEA comprised genes filtered by the absolute value of their expression difference between type II and III at the respective temperatures. Forty-three transcriptional networks (comprising more than 15 regulated genes controlled by a regulator) were used in the analysis. The Bioconductor software package RankProd (Hong et al. 2006), which is based on the rank product method (Breitling et al. 2004), was used to identify the differential expression gene caused by high temperature. RankProd analysis was carried out using R software (Ihaka & Gentleman 1996) (http://www.r-project.org). The classification of gene categories was carried out in accordance with the original paper (Riley et al. 2006), and 23 gene categories (comprising more than 10 genes within each category) were used in the analysis. The list of 302 essential genes (Hashimoto et al. 2005) was downloaded from the Profiling of E. coli Chromosome website (http://www.shigen.nig.ac.jp/ecoli/pec/). A total of 300 of the 302 essential genes were successfully annotated in our datasets and used for the analysis. Classification of the GroEL substrates (Kerner et al. 2005; Furusawa et al. 2009) and protein solubility (Niwa et al. 2009) were carried out in accordance with the original papers.

Acknowledgements

This work was supported by Grants-in-Aid for Scientific Research (A) 23241061 (to TY) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and in part by the research grant of the Institute for Fermentation, Osaka, Japan (to BWY).

References

Bogunil, D. & Dagan, T. (2012) Cumulative impact of chaperone-mediated folding on genome evolution. Biochemistry 51, 9941–9953.
Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett. 573, 83–92.
Fujiwara, K., Ishihama, Y., Nakahigashi, K., Soga, T. & Taguchi, H. (2010) A systematic survey of in vivo obligate chaperonin-dependent substrates. EMBO J. 29, 1552–1564.
Furusawa, C., Ono, N., Suzuki, S., Agata, T., Shimizu, H. & Yomo, T. (2009) Model-based analysis of non-specific binding for background correction of high-density oligonucleotide microarrays. Bioinformatics 25, 36–41.
Gamer, J., Bujard, H. & Bukau, B. (1992) Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma 32. Cell 69, 833–842.
Gamer, J., Multhaup, G., Tomoyasu, T., McCarty, J.S., Rudiger, S., Schonfeld, H.J., Schirra, C., Bujard, H. & Bukau, B. (1996) A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the Escherichia coli heat shock transcription factor sigma32. EMBO J. 15, 607–617.
Guisbert, E., Yura, T., Rhodius, V.A. & Gross, C.A. (2008) Convergence of molecular, modeling, and systems approaches for an understanding of the Escherichia coli heat shock response. Microbiol. Mol. Biol. Rev. 72, 545–554.
Hartl, F.U., Bracher, A. & Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. Nature 475, 324–332.
Hartl, F.U. & Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295, 1852–1858.
Hashimoto, M., Ichimura, T., Mizoguchi, H., Tanaka, K., Fujimitsu, K., Keyamura, K., Ote, T., Yamakawa, T., Yamazaki, Y., Mori, H., Katayama, T. & Kato, J. (2005) Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome. Mol. Microbiol. 55, 137–149.
Hong, F., Breitling, R., McEntee, C.W., Wittner, B.S., Nemhauser, J.L. & Chory, J. (2006) RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. Bioinformatics 22, 2825–2827.
Ihaka, R. & Gentleman, R. (1996) R: a language for data analysis and graphics. J. Comput. Graph. Stat. 5, 299–314.
Ishihama, Y., Schmidt, T., Rappslber, J., Mann, M., Hartl, F.U., Kerner, M.J. & Frishman, D. (2008) Protein abundance profiling of the Escherichia coli cytosol. BMC Genom. 9, 102.
Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M. & Hartl, F.U. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell 122, 209–220.
Kishimoto, T., Iijima, L., Tatsumi, M., Ono, N., Oyake, A., Hashimoto, T., Matsuo, M., Okubo, M., Suzuki, S., Mori, K., Kashiwagi, A., Furusawa, C., Ying, B.W. & Yomo, T. (2010) Transition from positive to neutral in mutation fixation along with continuing rising fitness in thermal adaptive evolution. PLoS Genet. 6, e1001164.
Kultz, D. (2005) Molecular and evolutionary basis of the cellular stress response. Annu. Rev. Physiol. 67, 225–257.
Lopez-Maury, L., Marguerat, S. & Bahler, J. (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat. Rev. Genet. 9, 583–593.
Makino, Y., Amada, K., Taguchi, H. & Yoshida, M. (1997) Chaperonin-mediated folding of green fluorescent protein. J. Biol. Chem. 272, 12468–12474.
Matsumoto, Y., Murakami, Y., Tsuru, S., Ying, B.W. & Yomo, T. (2013) Growth rate-coordinated transcriptome reorganization in bacteria. BMC Genom. 14, 808.

Mogk, A., Huber, D. & Bukau, B. (2011) Integrating protein homeostasis strategies in prokaryotes. Cold Spring Harb. Perspect. Biol. 3, a004366.

Morita, M., Kanemori, M., Yanagi, H. & Yura, T. (1999a) Heat-induced synthesis of sigma32 in Escherichia coli: structural and functional dissection of rpoH mRNA secondary structure. J. Bacteriol. 181, 401–410.

Morita, M.T., Tanaka, Y., Kodama, T.S., Kyogoku, Y., Yanagi, H. & Yura, T. (1999b) Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. Genes Dev. 13, 655–665.

Murakami, Y., Matsumoto, Y., Tsuru, S., Ying, B.W. & Yomo, T. (2015) Global coordination in adaptation to gene rewiring. Nucleic Acids Res. 43, 1304–1316.

Niwa, T., Ying, B.W., Saito, K., Jin, W., Takada, S., Ueda, T. & Taguchi, H. (2009) Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of Escherichia coli proteins. Proc. Natl Acad. Sci. USA 106, 4201–4206.

Nonaka, G., Blankschiern, M., Herman, C., Gross, C.A. & Rhodius, V.A. (2006) Regulon and promoter analysis of the E. coli heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. Genes Dev. 20, 1776–1789.

Riley, M., Abe, T., Arnaud, M.B. et al. (2006) Escherichia coli K-12: a cooperatively developed annotation snapshot–2005. Nucleic Acids Res. 34, 1–9.

Rodriguez, F., Arsene-Ploetze, F., Rist, W., Rudiger, S., Schneider-Mergener, J., Mayer, M.P. & Bukau, B. (2008) Molecular basis for regulation of the heat shock transcription factor sigma32 by the DnaK and DnaJ chaperones. Mol. Cell 32, 347–358.

Rueffler, C., Hermisson, J. & Wagner, G.P. (2012) Evolution of functional specialization and division of labor. Proc. Natl Acad. Sci. USA 109, E326–E335.

Salgado, H., Peralta-Gil, M., Gama-Castro, S. et al. (2012) RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. Nucleic Acids Res. 41, D203–D213.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. & Mesirov, J.P. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl Acad. Sci. USA 102, 15545–15550.

Suzuki, S., Ono, N., Furusawa, C., Ying, B.W. & Yomo, T. (2013) Comparison of sequence reads obtained from three next-generation sequencing platforms. PLoS ONE 6, e19534.

Szekely, P., Sheftel, H., Mayo, A. & Alon, U. (2013) Evolutionary tradeoffs between economy and effectiveness in biological homeostasis systems. PLoS Comput. Biol. 9, e1003163.

Ying, B.W., Seno, S., Kaneko, F., Matsuda, H. & Yomo, T. (2013) Multilevel comparative analysis of the contributions of genome reduction and heat shock to the Escherichia coli transcriptome. BMC Genom. 14, 25.

Ying, B.W., Tsuru, S., Seno, S., Matsuda, H. & Yomo, T. (2014) Gene expression scaled by distance to the genome replication site. Mol. BioSyst. 10, 375–379.

Yura, T., Nagai, H. & Mori, H. (1993) Regulation of the heat-shock response in bacteria. Annu. Rev. Microbiol. 47, 321–350.

Yura, T. & Nakahigashi, K. (1999) Regulation of the heat-shock response. Curr. Opin. Microbiol. 2, 153–158.

Received: 21 June 2015
Accepted: 4 August 2015

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s web site:

Figure S1 Evolution experiments.
Figure S2 Details of the two mutations.
Figure S3 Box plots of gene expression.
Figure S4 Similarity in DEGs between type II and III cells.
Figure S5 Expression of major molecular chaperones.
Figure S6 Expression of the essential and total genes.
Table S1 Statistics of genome resequencing
Table S2 Differentially expressed genes (DEGs)
Table S3 Detailed results of the gene set enrichment analysis