Genomics Data 7 (2016) 1–3

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

The *Escherichia coli* transcriptome linked to growth fitness

Bei-Wen Ying [a,*], Kazuma Yama [b], Kazuki Kitahara [b], Tetsuya Yomo [b]

[a] University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan
[b] Osaka University, Suita, Osaka 565-0871, Japan

ARTICLE INFO

Article history:
Received 4 November 2015
Accepted 7 November 2015
Available online 10 November 2015

Keywords:
Transcriptome
Growth fitness
Escherichia coli
Microarray

ABSTRACT

A series of *Escherichia coli* strains with varied genomic sequences were subjected to high-density microarray analyses to elucidate the fitness-correlated transcriptomes. Fitness, which is commonly evaluated by the growth rate during the exponential phase, is not only determined by the genome but is also linked to growth conditions, *e.g.*, temperature. We previously reported genetic and environmental contributions to *E. coli* transcriptomes and evolutionary transcriptome changes in thermal adaptation. Here, we describe experimental details on how to prepare microarray samples that truly represent the growth fitness of the *E. coli* cells. A step-by-step record of sample preparation procedures that correspond to growing cells and transcriptome data sets that are deposited at the GEO database (GSE33212, GSE52770, GSE61739) are also provided for reference.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications	
Organism/cell line/tissue	*Escherichia coli* K12 strains MG1655, MDS42, DH1ΔleuB::gfp-Km^r
Sex	N/A
Sequencer or array type	High-density DNA tilling array (EcFs)
Data format	Raw (CEL file)
Experimental factors	Genome: wild type (MG1655) vs reduced (MDS42) Temperature: regular (37 °C) vs high (41–45 °C) Mutation: rpoH (Type II) vs groESLp (Type III) Evolution: thermal adaptive (41B, 43B, 45 L) vs ancestor (Anc)
Experimental features	Microarray gene expression profiling of a series of thermal adaptive *E. coli* cells growing under the regular and high temperatures
Consent	N/A
Sample source location	Tsukuba and Osaka, Japan

1. Direct link to deposited data

The data set is available at the GEO database under:
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33212.
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE52770.
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61749.

2. Experimental design, materials and methods

The key purpose of bacterial transcriptome analyses is to strictly correlate a gene expression pattern, *i.e.*, microarray data, and cellular physiology, which is generally represented by the growth fitness [1–3]. In particular, evolution studies focus on evolutionary changes in gene expression, which must be precisely linked to increased fitness. Our studies critically consider the correlation between cell growth rate and the transcriptome. Thus, the following experimental procedures, which were excluded from the Materials and Methods of the original papers [4–6], were performed. The details described here can be used to reproduce our deposited transcriptome data sets and the corresponding analytical results.

2.1. Strains and cell cultures

The genetic backgrounds of the *Escherichia coli* (*E. coli*) strains used for the transcriptome analyses are in previous reports [4–6]. An overview of these strains is shown in the *Specifications* table in the *Data in Brief*. All cell cultures were carried out in 5 mL in glass culture tubes (φ16 mm × 180 mm) with a silicon cap. *E. coli* cells were commonly grown in the minimal medium M63 (62 mM K_2HPO_4, 39 mM KH_2PO_4, 15 mM $(NH_4)_2SO_4$, 2 μM $FeSO_4 \cdot 7H_2O$, 15 μM thiamine hydrochloride, 203 μM MgSO4·7H2O and 22 mM glucose) for the precise control of the components for cell growth. The culture media (M63 or in the supplemental chemicals if required) and temperatures as well as the incubators and rotational (shaking) speeds are described in the original papers [4–6] and in the data summaries deposited at the GEO database.

* Corresponding author.

http://dx.doi.org/10.1016/j.gdata.2015.11.011
2213-5960/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

2.2. Serial transfer

To determine the constant growth dynamics, the *E. coli* cells were serially transferred for a few days until they reached a relatively constant growth rate. Cell stocks stored with 10–15% glycerol at −80 °C were used to inoculate the corresponding media. An overnight culture of the *E. coli* cells typically reached saturated growth. This overnight culture mixture was subsequently transferred to fresh media through a series of dilutions, *i.e.*, 10^3-, 10^4- and 10^5-fold, and cultured overnight again. The culture with a cell concentration (or OD_{600} value) in the range of 10^7–10^8 cells/mL (roughly equivalent to $OD_{600} = 0.01$–0.2) among the diluted cultures was used for the following serial transfer. The dilution rates were changed accordingly; for instance, if the 10^4 dilution culture was selected, then the following serial transfer featured the dilution factors 10^4, 10^5 and 10^6. The growth rate (μ) was calculated using the following equation (Eq. (1)).

$$\mu = \mathrm{LN}\left(\frac{C_j}{C_i}\right)/(t_j - t_i) \quad (1)$$

Here, C_i and C_j represent the initial and final cell concentrations (OD_{600} values), and t_i and t_j indicate the initial and final cell culture time points, respectively. A serial transfer was performed repeatedly until equivalent growth rates were detected in two continuous days (transfers), which often required three to five days. The final cell culture in the early or middle exponential growth phase was stored in 10–20 tubes at a small dose (50–200 μL) per tube for future analyses. The newly prepared glycerol stocks (single tubes) were used once, and the remainder was discarded.

2.3. Precise measurement of the growth curve

If a cell population was supposed to be heterogeneous, this property would disappear upon repeated serial transfers. Thus, we carefully measured growth curves of the cell stocks, to determine the timing for the exponential growth phase of the cell populations. After inoculating the fresh media with the glycerol stock, the initial cell concentration was measured using a flow cytometer (FACSCanto™ II, Becton, Dickinson and Company) or a cell counter (Multisizer™ 3 coulter counter, Beckman). Temporal changes in the cell concentration was monitored every 2–3 h until the cell culture approached a high density (approximately 5×10^8 cells/mL) close to the saturated concentration (approximately 2×10^9 cells/mL). Using the growth curve, we determined the exponential growth phase, and, consequently, we estimated the timing of cell collection for the microarray analyses. Based on the estimated culture time, the cells were cultured until the cell density reached the expected magnitude, and the cells were collected for transcriptome analyses. The initial and final cell concentrations as well as the growth rates that correspond to the deposited GEO data sets are summarized in Table 1.

Table 1
Details on the cell samples used for the transcriptome analyses. The samples are indicated by the names of the *E. coli* strains used in the transcriptome studies and the number of replicas. The Temp., Initial conc., Final conc., and μ represent the temperatures for cell growth, the initial and final cell concentrations of the cultures, and the growth rate calculated using the initial and final cell concentrations as well as culture times, respectively. The Total cells indicate the number of cells used for the total RNA isolation, and the RNAs represent the quantity of total RNA purified from the cells. The cDNAs indicate the quantity of cDNAs that were reverse transcribed from every 10 μg of the total RNA.

Samples	Temp.	Initial conc.	Final conc.	μ	Total cells	RNAs	cDNAs
(Replica no.)	(°C)	(Cells/mL)	(Cells/mL)	(h^{-1})	(cells)	(μg)	(μg)
MDS42 (1)	37	1.0e4	4.8e8	0.76	8.0e8	28.9	5.5
MDS42 (2)	37	1.0e4	2.1e8	0.71	1.1e9	32.8	4.7
MDS42 (3)	37	1.0e4	4.3e8	0.74	7.2e8	25.8	5.0
MDS42 (4)	37	1.0e4	1.3e8	0.73	6.5e8	22.2	4.4
MDS42 (5)	37	1.0e4	4.3e8	0.73	8.5e8	39.3	4.9
MDS42 (6)	37	1.0e4	2.7e8	0.76	1.4e9	35.3	4.2
MDS42 (7)	37	1.0e4	1.4e8	0.74	7.0e8	21.8	5.8
MG1655 (1)	37	1.0e4	9.8e7	0.64	4.9e8	20.4	5.8
MG1655 (2)	37	1.0e4	1.3e8	0.65	1.3e9	51.3	5.8
MG1655 (3)	37	1.0e4	1.2e8	0.70	5.9e8	22.8	5.1
MG1655 (4)	37	1.0e4	1.2e8	0.70	5.9e8	24.4	4.3
MG1655 (5)	37	1.0e4	1.3e8	0.70	6.3e8	25.7	5.1
MG1655 (6)	37	1.0e4	1.4e8	0.71	6.9e8	27.6	4.9
MG1655 (7)	37	1.0e4	1.2e8	0.70	6.1e8	24.1	5.0
Type II (1)	36.9	1.0e4	1.4e8	0.46	6.9e8	26.3	5.3
Type II (2)	36.9	1.0e4	1.0e8	0.51	5.1e8	21.5	5.2
Type II (3)	36.9	1.0e4	9.2e7	0.51	4.6e8	18.2	4.8
Type II (1)	44.7	1.0e7	1.6e8	0.39	7.9e8	22.6	6.2
Type II (2)	44.7	9.0e6	1.1e8	0.23	5.4e8	22.1	4.9
Type II (3)	44.7	9.0e6	1.4e8	0.25	6.9e8	20.1	4.6
Type III (1)	36.9	1.0e4	1.0e8	0.52	5.0e8	12.7	4.1
Type III (2)	36.9	1.0e4	1.0e8	0.57	5.2e8	17.4	3.7
Type III (3)	36.9	1.0e4	1.0e8	0.57	5.2e8	16.6	3.9
Type III (1)	44.7	9.0e6	9.4e7	0.29	4.7e8	18.5	5.0
Type III (2)	44.7	9.0e6	1.1e8	0.31	5.6e8	16.8	4.8
Type III (3)	44.7	9.0e6	1.0e8	0.30	5.2e8	13.5	4.4
Anc (1)	36.9	1.0e4	5.8e8	0.48	2.9e9	164.5	5.9
Anc (2)	36.9	1.0e4	3.0e8	0.45	1.5e9	102.8	6.6
Anc (3)	36.9	1.0e4	1.1e8	0.41	6.0e8	26.1	6.4
41B (1)	36.9	1.0e4	1.1e8	0.66	5.3e8	18.1	4.5
41B (2)	36.9	1.0e4	1.5e8	0.68	7.3e8	16.2	4.5
41B (3)	36.9	1.0e4	1.9e8	0.70	9.3e8	17.4	4.6
41B (1)	41.2	1.0e4	1.5e8	0.64	7.7e8	54.8	6.4
41B (2)	41.2	1.0e4	2.1e8	0.66	1.0e9	40.9	6.7
41B (3)	41.2	1.0e4	2.0e8	0.66	9.8e8	36.9	6.3
43B (1)	36.9	1.0e4	6.7e8	0.69	1.5e9	81.1	5.6
43B (2)	36.9	1.0e4	1.4e8	0.59	1.5e9	82.1	5.3
43B (3)	36.9	1.0e4	1.3e8	0.59	1.4e9	74.3	5.2
43B (1)	43.2	1.0e4	1.2e8	0.49	5.8e9	33.7	6.1
43B (2)	43.2	1.0e4	1.6e8	0.51	8.2e8	46.2	6.0
43B (3)	43.2	1.0e4	2.9e8	0.54	1.4e9	35.0	6.1
45 L (1)	36.9	1.0e4	2.3e8	0.63	1.2e9	132.8	5.2
45 L (2)	36.9	1.0e4	1.3e8	0.59	6.3e8	42.6	6.6
45 L (3)	36.9	1.0e4	1.4e8	0.60	6.9e8	37.6	6.2
45 L (1)	44.8	1.0e4	3.5e8	0.48	1.8e9	101.8	6.1
45 L (2)	44.8	1.0e4	1.6e8	0.44	8.0e8	27.8	5.2
45 L (3)	44.8	1.0e4	1.6e9	0.44	7.9e8	29.9	4.5

2.4. Cell collection and sample preparation

Exponentially growing cells were immediately collected for RNA isolation. A time-consuming cell collection procedure will disturb the RNA abundance because changes in mRNA often occur over a few minutes. The cell cultures in the exponential growth phase were interrupted by placing them directly into a cold phenol–ethanol solution (1 g of phenol in 10 mL of ethanol) that was prepared in advance. Immediately, the cells were collected using centrifugation at 16,000 ×*g* for 5 min at 4 °C, and the pelleted cells were stored at −80 °C prior to use. The total RNA was extracted using an RNeasy kit (Qiagen) in accordance with the manufacturer's instructions. The volume of the final elution buffer was 30 μL, which was used to generate a relatively high concentration of the total RNA (approx. 500–1000 μg/μL). The quantity of the purified total RNA was first determined based on the absorbance at 260 nm using a NanoDrop ND-1000 (Thermo Fisher Scientific Inc.). Second, a precise evaluation was performed using an Agilent 2100 Bioanalyzer with an RNA 6000 Nano kit (Agilent Technologies) and electrophoresis with an agarose gel under reducing conditions. Every 10 μg of the purified total RNAs that passed quality control were used as templates for reverse transcription to generate the cDNA. The synthesized cDNA was purified using a MinElute kit (Qiagen) with a final elution of 10 μL. The final productivities of the total RNA purification and cDNA preparation are summarized in Table 1. Every 3 μg of cDNA was used for the following microarray gene expression analysis, which was performed using an Affymetrix GeneChip system. Fragmentation, labelling and hybridization of cDNA were carried out in accordance with the Affymetrix GeneChip Expression Analysis Technical Manual. We slightly modified this method by increasing the incubation time to 2 h to

improve the labelling efficiency. Three to seven independent experiments (from cell culture to sample preparation) were performed for each condition.

References

[1] E. Soupene, W.C. van Heeswijk, J. Plumbridge, V. Stewart, D. Bertenthal, H. Lee, G. Prasad, O. Paliy, P. Charernnoppakul, S. Kustu, Physiological studies of *Escherichia coli* strain MG1655: growth defects and apparent cross-regulation of gene expression. J. Bacteriol. 185 (18) (2003) 5611–5626.

[2] R. Nahku, K. Valgepea, P.J. Lahtvee, S. Erm, K. Abner, K. Adamberg, R. Vilu, Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures. J. Biotechnol. 145 (1) (2010) 60–65.

[3] Y. Matsumoto, Y. Murakami, S. Tsuru, B.W. Ying, T. Yomo, Growth rate-coordinated transcriptome reorganization in bacteria. BMC Genomics 14 (2013) 808.

[4] B.W. Ying, Y. Matsumoto, K. Kitahara, S. Suzuki, N. Ono, C. Furusawa, T. Kishimoto, T. Yomo, Bacterial transcriptome reorganization in thermal adaptive evolution. BMC Genomics 16 (1) (2015) 802.

[5] K. Yama, Y. Matsumoto, Y. Murakami, S. Seno, H. Matsuda, K. Gotoh, D. Motooka, S. Nakamura, B.W. Ying, T. Yomo, Functional specialization in regulation and quality control in thermal adaptive evolution. Genes Cells 20 (11) (2015) 943–955.

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