

Statistical Comparison of Expression Profiles of Genes

Hijiri Maeno¹

hi-maeno@bs.aist-nara.ac.jp

Shin-ichi Matsusaki¹

s-matsu@bs.aist-nara.ac.jp

Yasushi Masuda²

yasuda@gtc.aist-nara.ac.jp

Taku Oshima²

taku@bs.aist-nara.ac.jp

Hiroko Asahi²

asahi@gtc.aist-nara.ac.jp

Shigehiko Kanaya²

skanaya@gtc.aist-nara.ac.jp

Hirotsada Mori²

hmori@gtc.aist-nara.ac.jp

¹ Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan

² CREST, JST (Japan Science and Technology Corporation), Honnmachi, Kawaguchi, Saitama 332-0012, Japan

Keywords: microarray, *Escherichia coli* K-12 (W3110), transcription profile, Genobase (ver 3.0)

1 Introduction

The recent advances of genome-scale sequencing and array technologies have made it possible to monitor the expression pattern of thousands of genes. In the present study, we have developed a novel method for determining transcription units based on comparing expression profiles between adjacent genes using microarray data for various experimental conditions. We basically examined expression profiles of 4390 genes in *Escherichia coli* K-12 (W3110).

2 Methods

We have calculated the correlation r_{ij} of transcription profiles between individual pairs of genes by the following equation. First, $X_i = (x_{i1}, x_{i2}, \dots, x_{ik}, \dots, x_{iM})$ represents expression profile for i th gene. Here, transcription intensity for k th ($k = 1, 2, \dots, M$) experiment for the i th gene was denoted by x_{ik} . Secondly, correlation coefficient r_{ij} is expressed by the following equation;

$$r_{ij} = \frac{\sum_{k=1}^M (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j)}{\sqrt{\sum_{k=1}^M (X_{ik} - \bar{X}_i)^2 \sum_{k=1}^M (X_{jk} - \bar{X}_j)^2}}$$

In general, some relative intensities denoted by x_{ik} cannot be obtained because reference intensities are very small. So we calculated r_{ij} for both i th and j th genes which could be obtained larger than $M/2$ -identical experiments, (which corresponds to subscript k).

3 Results and Discussion

We examined 51 experiments (time series data for 11 experiments, 40 disruptants for nucleoid genes and two-component systems). Correlation coefficients larger than 0.9 are obtained for 265 pairs of different genes. Of them, 117 pairs are involved in translation (Table 1 upper). In Table 1, “JW” represents the ID denoted by Japanese Genome Project (see Genobase 3.0 [1, 2]). This indicates that expression profiles for genes involved in translation are very correlated. We obtained the other 107 pairs of genes whose functions are known (Table 1 lower). Some of them are located in proximate regions of genome denoted by “cluster”, which are useful for estimating transcription units. We also obtained 41 pairs of function known genes and function-unknown genes such as y-genes. These may be useful for estimating molecular functions in view of expression profiles. Detail data will be discussed in GIW2001. In conclusion, this may be useful for determining transcription units. Furthermore, we are planning to examine relation between gene organization in genome and expression profiles of genes.

References

- [1] Mori, H., Isono, K., Horiuchi, T., and Miki, T., Functional genomics of *Escherichia coli* in Japan, *Res. Microbiol.*, 151(2):121–128, 2000.
- [2] <http://ecoli.aist-nara.ac.jp/>

Table 1: Profile correlations (0.90) between genes with translation.

JW	gene	JW	gene	Corr.	JW	gene	JW	gene	Corr.	JW	gene	JW	gene	Corr.
JW3268	rpsH	JW3270	rpIE	0.963	JW3263	rpIO	JW3267	rpIF	0.928	JW3303	rpsG	JW3950	rpoB	0.913
JW3275	rpIP	JW3280	rpIW	0.957	JW0164	rpsB	JW0165	tsf	0.927	JW3259	rpsK	JW3283	rpsJ	0.912
JW3267	rpIF	JW3269	rpsN	0.954	JW3268	rpsH	JW3270	rpIE	0.926	JW3268	rpsH	JW3303	rpsG	0.912
JW3259	rpsK	JW3260	rpsM	0.953	JW3258	rpsD	JW3272	rpIN	0.926	JW3268	rpsH	JW3283	rpsJ	0.912
JW3267	rpIF	JW3270	rpIE	0.952	JW3260	rpsM	JW3280	rpIW	0.925	JW3272	rpIN	JW3283	rpsJ	0.912
JW3270	rpIE	JW3272	rpIN	0.951	JW3267	rpIF	JW3302	fusA	0.924	JW3280	rpIW	JW3303	rpsG	0.911
JW0164	rpsB	JW0165	tsf	0.950	JW3946	rpIK	JW3947	rpIA	0.923	JW3278	rpsS	JW3280	rpIW	0.911
JW3268	rpsH	JW3950	rpoB	0.950	JW3276	rpsC	JW3281	rpID	0.923	JW3276	rpsC	JW3281	rpID	0.911
JW3283	rpsJ	JW3951	rpoC	0.949	JW3269	rpsN	JW3281	rpID	0.922	JW3258	rpsD	JW3260	rpsM	0.909
JW3270	rpIE	JW3303	rpsG	0.949	JW3256	rpIQ	JW3280	rpIW	0.922	JW3267	rpIF	JW3280	rpIW	0.909
JW3265	rpsE	JW3269	rpsN	0.946	JW3269	rpsN	JW3276	rpsC	0.922	JW3267	rpIF	JW3277	rpIV	0.908
JW3259	rpsK	JW3280	rpIW	0.945	JW3277	rpIV	JW3280	rpIW	0.922	JW3259	rpsK	JW3278	rpsS	0.907
JW3279	rpIB	JW3281	rpID	0.944	JW3271	rpIX	JW3303	rpsG	0.921	JW3268	rpsH	JW3272	rpIN	0.907
JW3269	rpsN	JW3302	fusA	0.943	JW3275	rpIP	JW3278	rpsS	0.921	JW3270	rpIE	JW3282	rpIC	0.907
JW3276	rpsC	JW3277	rpIV	0.942	JW3268	rpsH	JW3280	rpIW	0.921	JW3276	rpsC	JW3281	rpID	0.907
JW3256	rpIQ	JW3260	rpsM	0.941	JW3263	rpIO	JW3280	rpIW	0.920	JW3270	rpIE	JW3272	rpIN	0.906
JW3270	rpIE	JW3303	rpsG	0.940	JW3277	rpIV	JW3280	rpIW	0.920	JW3283	rpsJ	JW3950	rpoB	0.906
JW3262	secY	JW3277	rpIV	0.939	JW3260	rpsM	JW3263	rpIO	0.919	JW3256	rpIQ	JW3259	rpsK	0.905
JW3270	rpIE	JW3280	rpIW	0.939	JW3268	rpsH	JW3269	rpsN	0.918	JW3270	rpIE	JW3272	rpIN	0.905
JW3269	rpsN	JW3270	rpIE	0.939	JW3258	rpsD	JW3272	rpIN	0.918	JW3262	secY	JW3280	rpIW	0.905
JW3270	rpIE	JW3950	rpoB	0.937	JW3267	rpIF	JW3281	rpID	0.918	JW3266	rpIR	JW3302	fusA	0.905
JW3267	rpIF	JW3268	rpsH	0.937	JW3262	secY	JW3277	rpIV	0.918	JW3272	rpIN	JW3303	rpsG	0.905
JW3258	rpsD	JW3270	rpIE	0.935	JW3270	rpIE	JW3278	rpsS	0.918	JW3258	rpsD	JW3943	tufB	0.904
JW3263	rpIO	JW3268	rpsH	0.935	JW3269	rpsN	JW3303	rpsG	0.917	JW3258	rpsD	JW3270	rpIE	0.904
JW3265	rpsE	JW3302	fusA	0.934	JW3276	rpsC	JW3277	rpIV	0.916	JW3265	rpsE	JW3304	rpsL	0.903
JW3264	rpmD	JW3304	rpsL	0.934	JW3278	rpsS	JW3280	rpIW	0.916	JW3262	secY	JW3273	rpsQ	0.903
JW3266	rpIR	JW3281	rpID	0.934	JW3263	rpIO	JW3270	rpIE	0.916	JW3270	rpIE	JW3277	rpIV	0.903
JW3256	rpIQ	JW3260	rpsM	0.933	JW3270	rpIE	JW3303	rpsG	0.916	JW3946	rpIK	JW3947	rpIA	0.902
JW3946	rpIK	JW3947	rpIA	0.933	JW3272	rpIN	JW3280	rpIW	0.915	JW3268	rpsH	JW3278	rpsS	0.902
JW3262	secY/pr	JW3280	rpIW	0.931	JW3276	rpsC	JW3280	rpIW	0.915	JW3270	rpIE	JW3272	rpIN	0.902
JW3266	rpIR	JW3268	rpsH	0.930	JW3272	rpIN	JW3278	rpsS	0.915	JW3269	rpsN	JW3301	tufA	0.901
JW3271	rpIX	JW3277	rpIV	0.930	JW3278	rpsS	JW3951	rpoC	0.914	JW3265	rpsE	JW3269	rpsN	0.901
JW3270	rpIE	JW3280	rpIW	0.930	JW3268	rpsH	JW3282	rpIC	0.914	JW3263	rpIO	JW3283	rpsJ	0.901
JW3268	rpsH	JW3270	rpIE	0.929	JW3152	rpmA	JW3199	rpsI	0.914	JW3264	rpmD	JW3269	rpsN	0.901
JW3275	rpIP	JW3283	rpsJ	0.929	JW3263	rpIO	JW3269	rpsN	0.913	JW3301	tufA	JW3303	rpsG	0.901
JW3280	rpIW	JW3283	rpsJ	0.929	JW3301	tufA	JW3303	rpsG	0.913	JW3281	rpID	JW3302	fusA	0.901
JW3265	rpsE	JW3267	rpIF	0.928	JW3258	rpsD	JW3280	rpIW	0.913	JW3270	rpIE	JW3283	rpsJ	0.901
JW3265	rpsE	JW3281	rpID	0.928	JW3256	rpIQ	JW3258	rpsD	0.913	JW3256	rpIQ	JW3280	rpIW	0.900

JW	gene	JW	gene	Corr.	cluster	JW	gene	JW	gene	Corr.	cluster	JW	gene	JW	gene	Corr.	cluster
JW3994	malE	JW3996	lamB	0.978		JW0712	sdhD	JW0714	sdhB	0.931	sdhABCD	JW0013	dnaK	JW4102	mopB	0.912	
JW2273	nuoL	JW2275	nuoJ	0.972	nuo operon	JW0628	rlpA	JW2781	459#3	0.930		JW0114	acc2	JW0714	sdhB	0.912	
JW2074	ptkB	JW2078	(gatZ)	0.970		JW2273	nuoL	JW2280	nuoE	0.930	nuo operon	JW1068	figJ	JW1934	flrR	0.911	
JW2276	nuoI	JW2279	nuoF	0.970	nuo operon	JW2201	ada	JW2203	ompC	0.930		JW3023	508#7	JW3626	spoU	0.911	
JW2275	nuoJ	JW2277	nuoH	0.968	nuo operon	JW2278	nuoG	JW2280	nuoE	0.929	nuo operon	JW2498	hisS	JW2549	recO	0.910	
JW0973	226#7	JW1717	325#4	0.968		JW2940	477#7	JW2941	478#1	0.928		JW1064	figF	JW1068	figJ	0.910	fig operon
JW2276	nuoI	JW2281	nuoD	0.965	nuo operon	JW1061	figC	JW1063	figE	0.927	fig operon	JW1806	manX	JW1807	manY	0.909	
JW3864	fdoH	JW3865	fdoG	0.964	fdolHG	JW1426	tehB	JW1846	znuA	0.927		JW2279	nuoF	JW2280	nuoE	0.909	nuo operon
JW2276	nuoI	JW2279	nuoF	0.964	nuo operon	JW3709	uncC	JW3713	atpH	0.927		JW0714	sdhB	JW0716	sucB	0.908	
JW2277	nuoH	JW2280	nuoE	0.963	nuo operon	JW2276	nuoI	JW2278	nuoG	0.925	nuo operon	JW0231	phoE	JW0761	bioD	0.908	
JW3864	fdoH	JW3865	fdoG	0.962	fdolHG	JW3994	malE	JW3996	lamB	0.924		JW1236	oppB	JW1238	oppD	0.908	oppABCD
JW2273	nuoL	JW2275	nuoJ	0.961	nuo operon	JW2275	nuoJ	JW2282	nuoB	0.923	nuo operon	JW0714	sdhB	JW0718	sucD	0.908	
JW0738	gpmA	JW1846	znuA	0.958		JW1704	pheS	JW2276	nuoI	0.921		JW3994	malE	JW3997	malM	0.906	
JW2275	nuoJ	JW2280	nuoE	0.956	nuo operon	JW2276	nuoI	JW2278	nuoG	0.920	nuo operon	JW1064	figF	JW1923	flgG	0.906	
JW2276	nuoI	JW2281	nuoD	0.954	nuo operon	JW2203	ompC	JW2217	atoE	0.920		JW2271	nuoN	JW2273	nuoL	0.905	nuo operon
JW2273	nuoL	JW2280	nuoE	0.953	nuo operon	JW0711	sdhC	JW0713	sdhA	0.920	sdhABCD	JW0713	sdhA	JW0714	sdhB	0.905	sdhABCD
JW3995	malK	JW3997	malM	0.952		JW0711	sdhC	JW2571	kgtP	0.920		JW0024	ileS	JW1410	hrpA	0.905	
JW0013	dnaK	JW4103	cutA2	0.952		JW2273	nuoL	JW2276	nuoI	0.920	nuo operon	JW1602	tus	JW1867	flhE	0.905	
JW3993	malF	JW3997	malM	0.950	malEFG	JW1062	figD	JW1923	flgG	0.920		JW2275	nuoJ	JW2276	nuoI	0.904	nuo operon
JW2272	nuoM	JW2276	nuoI	0.947	nuo operon	JW3477	hdeB	JW3478	hdeA	0.919	hdeAB	JW3994	malE	JW3995	malK	0.904	
JW0717	sucC	JW0718	sucD	0.947	sucABCD	JW3995	malK	JW3997	malM	0.919		JW1071	rne	JW1076	plsX	0.904	
JW3993	malF	JW3995	malK	0.944		JW0717	sucC	JW0718	sucD	0.919	sucABCD	JW3995	malK	JW3997	malM	0.903	
JW2280	nuoE	JW2282	nuoB	0.944	nuo operon	JW1061	figC	JW1063	figE	0.918	fig operon	JW3996	lamB	JW3997	malM	0.903	
JW1854	ntpA	JW4349	lplA	0.942		JW0716	sucB	JW0718	sucD	0.918	sucABCD	JW2234	glpT	JW2236	glpB	0.903	
JW3477	hdeB	JW3478	hdeA	0.941	hdeAB	JW2859	prfB	JW3951	rpoC	0.917		JW3477	hdeB	JW3478	hdeA	0.903	hdeAB
JW1236	oppB	JW1238	oppD	0.941	oppABCD	JW0420	cyoC	JW0421	cyoB	0.917	cyoABCDE	JW2020	rfc	JW2022	rfsX	0.903	
JW3993	malF	JW3997	malM	0.938		JW1062	figD	JW1929	flmI	0.916		JW0712	sdhD	JW0714	sdhB	0.902	sdhABCD
JW0738	gpmA	JW1426	tehB	0.938		JW0714	sdhB	JW0718	sucD	0.916		JW0999	putA	JW3578	lldP	0.902	
JW2275	nuoJ	JW2280	nuoE	0.936	nuo operon	JW2082	gatZ	JW2083	gatY	0.915	gatYZ	JW0013	dnaK	JW4103	cutA2	0.902	
JW2280	nuoE	JW2282	nuoB	0.935	nuo operon	JW0716	sucB	JW0718	sucD	0.914	sucABCD	JW2074	ptkB	JW2078	(gatZ)	0.902	
JW1641	317#8	JW2769	exo	0.935		JW3995	malK	JW3996	lamB	0.914	malK-lamB	JW2280	nuoE	JW2281	nuoD	0.902	nuo operon
JW2203	ompC	JW2217	atoE	0.933		JW0714	sdhB	JW0716	sucB	0.914		JW1808	manZ	JW2076	ptkC	0.902	
JW1237	oppC	JW1238	oppD	0.933	oppABCD	JW1237	oppC	JW1238	oppD	0.913	oppABCD	JW2272	nuoM	JW2276	nuoI	0.901	nuo operon
JW3994	malE	JW3996	lamB	0.932		JW2275	nuoJ	JW2277	nuoH	0.913	nuo operon	JW1214	narK	JW1215	narG	0.900	
JW1237	oppC	JW1238	oppD	0.932	oppABCD	JW0420	cyoC	JW0422	cyoA	0.913	cyoABCDE	JW0179	dnaE	JW2548	pdxJ	0.900	
JW3994	malE	JW3996	lamB	0.931		JW2276	nuoI	JW2278	nuoG	0.912	nuo operon						