



www.sciencemag.org/cgi/content/full/311/5768/1730/DC1

Supporting Online Material for

Howard Ochman* and Liliana M. Davalos

*To whom correspondence should be addressed. E-mail: hochman@email.arizona.edu

Published 24 March, *Science* **311**, 1730 (2006)

DOI: 10.1126/science.1119966

This PDF file includes:

Tables S1 to S5

References

Table S1. List of predicted disrupted pseudogenes in the 2002 annotation of the *E. coli* MG1655 genome (*I*).

Gene	Beginning	End	Strand	Description	Essential*	Function†	Original Blattner ID	New Blattner ID‡
<i>afuB</i>	278038	278400	-1	putative transport system permease protein	N	?	b0263	
<i>eaeH</i>	313581	314468	1	attaching and effacing protein pathogenesis factor	N	N	b0297	
<i>ykgA</i>	315674	316393	-1	putative ARAC-type regulatory protein	X	N	b0300	
<i>yaiT</i>	389475	390935	1	orf hypothetical protein	N	N	b0371	
<i>yaiU</i>	392239	393642	1	putative flagellin structural protein	N	?	b0374	
<i>b0499</i>	527173	527883	1	orf hypothetical protein	N	?	b0499	
<i>ybcY</i>	581375	581959	-1	orf hypothetical protein	N	N	b0562	
<i>ybfD</i>	737315	738076	1	putative DNA ligase	N	?	b0706	
<i>ycdN</i>	1080570	1080689	1	orf hypothetical protein	N	N	b1016	b4490
<i>ycdN</i>	1080570	1080689	1	putative cytochrome	N	N	b1017	b4490
<i>ydaW</i>	1420725	1421336	1	orf hypothetical protein	N	?	b1361	
<i>lomR</i>	1425482	1425637	1	orf hypothetical protein	X	N	b1369	
<i>gapC_2</i>	1487737	1487988	-1	glyceraldehyde-3-phosphate dehydrogenase (second fragment)	N	N	b1417	b4493
<i>gapC_1</i>	1487985	1488389	-1	glyceraldehyde 3-phosphate dehydrogenase C interrupted	?	N	b1416	b4493
<i>yncl</i>	1528610	1529356	1	orf hypothetical protein	N	?	b1458	
<i>yncl</i>	1529400	1529600	1	orf hypothetical protein	?	?	b1459	
<i>yddL</i>	1543762	1544052	-1	putative outer membrane porin protein	E	?	b1472	
<i>ydfX</i>	1645370	1645660	-1	orf hypothetical protein	N	?	b1568	
<i>arpB</i>	1801118	1803017	1	orf hypothetical protein	N	?	b1720	b4494
<i>arpB</i>	1801118	1803017	1	orf hypothetical protein	N	?	b1721	b4494
<i>yedS</i>	2032045	2032560	1	putative outer membrane protein	N	N	b1964	b4496
<i>yedS</i>	2032045	2032560	1	orf hypothetical protein	?	N	b1965	b4496
<i>yedS</i>	2032045	2032560	1	putative outer membrane protein	N	N	b1966	b4496
<i>yeeL</i>	2050300	2051352	-1	orf hypothetical protein	?	N	b1979	b4497
<i>yeeL</i>	2050300	2051352	-1	orf hypothetical protein	X	N	b1980	b4497
<i>yoeA</i>	2066632	2067051	1	orf hypothetical protein	X	?	b1995	
b1998	2068268	2068528	1	orf hypothetical protein	X	?	b1998	
<i>yeeP</i>	2068525	2069235	1	putative histone	N	?	b1999	
<i>molR</i>	2194496	2198291	1	molybdate metabolism regulator first fragment	N	N	b2115	b4499
<i>molR</i>	2194496	2198291	1	molybdate metabolism regulator second fragment	N	N	b2116	b4499
<i>molR</i>	2194496	2198291	1	molybdate metabolism regulator third fragment	N	N	b2117	b4499
<i>yehQ</i>	2207098	2208966	1	orf hypothetical protein	N	?	b2122	
<i>yfaR</i>	2327820	2332334	-1	orf hypothetical protein	N	N	b2227	b4500
<i>yfaS</i>	2327820	2332334	-1	putative membrane protein	X	N	b2228	b4500
<i>yfaH</i>	2347709	2347915	1	orf hypothetical protein	N	?	b2238	
<i>yfcT</i>	2450378	2451274	-1	putative outer membrane protein	N	?	b2337	

F: experimentally confirmed function

<i>tfaS</i>	2468783	2469127	1	orf hypothetical protein	N	?	b2353	
<i>ypjB</i>	2781660	2782451	-1	orf hypothetical protein	X	?	b2649	
<i>ypjC</i>	2782551	2783033	-1	orf hypothetical protein	E	?	b2650	
<i>ygaQ</i>	2784419	2784751	1	orf hypothetical protein	X	?	b2654	
<i>ygaR</i>	2784770	2785456	1	orf hypothetical protein	X	?	b2655	b4462
<i>ygaR</i>	2784770	2785456	1	orf hypothetical protein	X	?	b2656	b4462
<i>yqaC</i>	2785628	2786260	1	putative enzyme	X	?	b2657	
<i>yqaD</i>	2786399	2786671	1	orf hypothetical protein	N	?	b2658	
<i>ygaY</i>	2806338	2806604	1	orf hypothetical protein	X	?	b2680	
<i>b268I</i>	2806598	2807515	1	putative transport protein	N	?	b2681	
<i>ygeP</i>	2995711	2996010	-1	orf hypothetical protein	X	?	b2862	
<i>ygeQ</i>	2996056	2996892	-1	orf hypothetical protein	X	?	b2863	
<i>ygfQ</i>	3025143	3026510	1	orf hypothetical protein	N	N	b2884	b4464
<i>ygfR</i>	3025143	3026510	1	putative oxidoreductase	N	N	b2885	b4464
<i>yghX</i>	3144878	3145288	-1	orf hypothetical protein	N	?	b2999	
<i>yghX</i>	3145288	3145713	-1	orf hypothetical protein	X	?	b3000	
<i>ygiQ</i>	3156949	3159168	-1	orf hypothetical protein	N	?	b3015	b4469
<i>ygiR</i>	3156949	3159168	-1	orf hypothetical protein	N	?	b3016	b4469
<i>yhaM</i>	3253363	3254673	-1	orf hypothetical protein	N	?	b3108	b4470
<i>yhaN</i>	3253363	3254673	-1	orf hypothetical protein	N	?	b3109	b4470
<i>agaW</i>	3278723	3279124	1	PTS system N-acetylgalactosamine-specific IIC component 2	E	?	b3134	
<i>agaA</i>	3279144	3279647	1	putative N-acetylgalactosamine-6-phosphate deacetylase	N	N	b3135	
<i>yhdP</i>	3390480	3394280	-1	orf hypothetical protein	N	?	b3245	b4472
<i>yhdR</i>	3390480	3394280	-1	orf hypothetical protein	N	?	b3246	b4472
<i>frlC</i>	3500362	3501192	1	orf hypothetical protein	N	?	b3372	b4474
<i>frlC</i>	3500362	3501192	1	orf hypothetical protein	N	?	b3373	b4474
<i>gntU</i>	3573744	3575084	-1	low-affinity gluconate transport permease protein interrupted	N	F	b3435	b4476
<i>gntU</i>	3573744	3575084	-1	low-affinity gluconate transport permease protein fragment 2	N	F	b3436	b4476
<i>yhiK</i>	3630875	3631246	-1	orf hypothetical protein	N	?	b3489	
<i>yhiL</i>	3631243	3632481	-1	orf hypothetical protein	N	?	b3490	
<i>yhiS</i>	3649314	3650096	1	orf hypothetical protein	N	N	b3504	
<i>yibJ</i>	3765244	3765945	1	orf hypothetical protein	N	?	b3595	
<i>glvG</i>	3859372	3860010	-1	probable 6-phospho-beta-glucosidase	X	?	b3681	
<i>glvB</i>	3860010	3860495	-1	PTS system arbutin-like IIB component	X	?	b3682	
<i>glvC</i>	3860520	3861887	-1	PTS system arbutin-like IIC component	N	?	b3683	
<i>yidX</i>	3873461	3874117	1	putative replicase EC 2.7.-	N	?	b3696	
<i>trkD</i>	3929339	3931207	1	low affinity potassium transport system	N	F	b3747§	
<i>ilvG</i>	3948583	3950227	1	acetolactate synthase II large subunit cryptic interrupted	N	N	b4488	
b3776	3958035	3958310	-1	orf hypothetical protein	X	?	b3776	
<i>yifN</i>	3958265	3958525	-1	orf hypothetical protein	N	?	b3777	
<i>rffT</i>	3975548	3976627	1	orf conceptual translation in SwissProt is fused with b4405	N	F	b4404	b4481
<i>rffT</i>	3975548	3976627	1	orf conceptual translation in	N	F	b4405	b4481

SwissProt is fused with b4404								
<i>yigE</i>	3998315	3999079	-1	orf hypothetical protein	N	?	b3814	b4482
<i>yigE</i>	3998315	3999079	-1	orf hypothetical protein	N	?	b3815	b4482
<i>yigL</i>	4008223	4009023	1	orf hypothetical protein	N	?	b3826	
<i>tatD</i>	4021577	4022359	1	orf hypothetical protein	N	F	b3841	b4483
<i>tatD</i>	4021577	4022359	1	orf hypothetical protein	N	F	b3842	b4483
<i>trkH</i>	4031168	4032619	1	potassium uptake requires TrkE	N	F	b3849	
<i>cpxP</i>	4103843	4104343	1	orf hypothetical protein	N	F	b3913	b4484
<i>cpxP</i>	4103843	4104343	1	orf hypothetical protein	N	F	b3914	b4484
<i>yjbI</i>	4248978	4250306	1	orf hypothetical protein	?	?	b4038	
<i>yjbL</i>	4258344	4258598	1	orf hypothetical protein	E	?	b4047	
<i>yjbM</i>	4258622	4259329	1	orf hypothetical protein	E	?	b4048	
<i>b4103</i>	4320466	4320687	-1	orf hypothetical protein	X	?	b4103	
<i>phnE</i>	4320684	4321304	-1	membrane channel protein component of Pn transporter	N	F	b4104	
<i>ytfI</i>	4435730	4436668	1	orf hypothetical protein	E	?	b4215§	
<i>yjiP</i>	4567021	4567332	1	orf hypothetical protein	N	?	b4338	
<i>yjiQ</i>	4567381	4567941	1	orf hypothetical protein	N	?	b4339	

* Essential genes are those required for aerobic growth in nutrient-rich conditions, as assayed in (29).

Notations are as follows: E, essential; N, not essential, X, not evaluated; ?, ambiguous results.

† Gene function, as annotated in Ecocyc (2) and Echobase (3).

‡ New ID assigned to correct mis-sequencing or misannotation in *E. coli* MG1655 genome; original ID has been discontinued.

§ Originally missequenced or misannotated in *E. coli* MG1655 genome, and corrected without changing Blattner ID.

Table S2. List of predicted disrupted pseudogenes in the 2004 annotation of the *E. coli* MG1655 genome (4,5). Functional assignments follow the notation used in Table S1.

Gene	Beginning	End	Strand	Description	Essential	Function	Original Blattner ID†	New Blattner ID
<i>yabP</i>	58474	59124	1	Unknown function	N	?	b0056	
<i>yabQ</i>	59121	59279	1	Unknown function	N	?	b0057	
<i>ykfC</i>	272086	273216	1	hypothetical protein	N	?	b0258	
<i>yagP</i>	296605	297015	-1	putative LysR-like transcriptional regulator	N	?	b0282	
<i>yaiS</i>	383283	383693	-1	Putative conserved protein	E	?	b0364‡	
<i>ylbE</i>	547838	548839	1	hypothetical protein	N	?	b0519	
<i>emrE</i>	567538	567870	1	EmrE protein	N	F	b0543§	
<i>ylcE</i>	582098	582283	1	hypothetical protein	?	?	b0563	
<i>ybeM</i>	657478	658041	1	putative amidase	?	?	b0626	
<i>ybfO</i>	733443	734876	1	Unknown function	N	?	b0703	
<i>flgD</i>	1131077	1131772	1	flagellar biosynthesis, hook protein	N	?	b1075	
<i>flgE</i>	1131797	1133005	1	flagellar biosynthesis, hook protein	N	?	b1076	
<i>acpP</i>	1150838	1151074	1	acyl carrier protein	?	F	b1094‡	
<i>ymfH</i>	1200292	1200603	-1	hypothetical protein	N	?	b1142	
<i>ycgI</i>	1222918	1223130	1	Putative conserved protein	X	?	b1173	
b1228	1285932	1286207	1	hypothetical protein	E	?	b1228§	
b1354	1417192	1417368	1	hypothetical protein	N	?	b1354	
<i>yddJ</i>	1542408	1542743	-1	Unknown function	N	?	b1470	
<i>yddK</i>	1542782	1543738	-1	putative glycoprotein	E	?	b1471	
<i>ydeT</i>	1586877	1588025	-1	putative fimbrial usher protein	N	?	b1505	
<i>yneF</i>	1608931	1609878	-1	hypothetical protein	N	?	b1522	
<i>gnsB</i>	1635633	1635809	-1	GnsB protein	?	F	b1550	
<i>ydfA</i>	1646532	1646687	1	hypothetical protein	?	?	b1571	
<i>celC</i>	1817479	1817829	-1	PEP-dependent phosphotransferase enzyme III for cellobiose, arbutin, and salicin	E	F	b1736	
<i>araH</i>	1980578	1981567	-1	ABC-type arabinose transport system, permease component	N	F	b1898	b4460
<i>araH</i>	1980578	1981567	-1	ABC-type arabinose transport system, permease component	N	F	b1899	b4460
<i>yeeR</i>	2072797	2074335	1	hypothetical protein	N	?	b2001	
<i>yeeS</i>	2074332	2074778	1	hypothetical protein	N	?	b2002	
<i>yeeU</i>	2075136	2075504	1	hypothetical protein	N	?	b2004	
<i>gatR</i>	2167717	2169757	-1	Galactitol utilization operon repressor	N	N	b2087	b4498

<i>gatR</i>	2167717	2169757	-1	Galactitol utilization operon repressor	N	N	b2090	b4498
<i>yohG</i>	2225345	2226541	-1	Putative channel/filament proteins	N	F	b2138§	
<i>yfdP</i>	2471542	2471988	1	hypothetical protein	N	?	b2359	
<i>yffD</i>	2746820	2748082	1	hypothetical protein	N	?	b2612	b4461
<i>yffD</i>	2746820	2748082	1	hypothetical protein	N	?	b2613	b4461
<i>yffU</i>	2769862	2770176	-1	arsenate reductase	N	?	b2638	
b2639	2770189	2770707	-1	putative pump protein	N	?	b2639	
<i>ygcU</i>	2895986	2897440	-1	Putative conserved protein	N	?	b2772	b4463
<i>ygcU</i>	2895986	2897440	-1	Putative conserved protein	N	?	b2773	b4463
<i>mltA</i>	2944103	2945200	-1	Membrane-bound lytic murein transglycosylase A precursor	N	F	b2813§	
<i>yggP</i>	3074201	3075478	-1	hypothetical protein	N	?	b2931	b4465
<i>yggP</i>	3074201	3075478	-1	hypothetical protein	N	?	b2932	b4465
<i>yghQ</i>	3129363	3130430	-1	hypothetical protein	N	?	b2983	
<i>tdcG</i>	3256307	3257671	-1	L-serine dehydratase	N	F	b3111	b4471
<i>tdcG</i>	3256307	3257671	-1	L-serine dehydratase	N	F	b3112	b4471
<i>smf</i>	3430458	3431582	-1	hypothetical protein	N	?	b3285	b4473
<i>smf</i>	3430458	3431582	-1	hypothetical protein	N	?	b3286	b4473
<i>rtcA</i>	3553855	3554871	-1	Chain B, Phosphate Cyclase	N	F	b3419	b4475
<i>rtcA</i>	3553855	3554871	-1	Chain B, Phosphate Cyclase	N	F	b3420	b4475
<i>arsR</i>	3646551	3646904	1	transcriptional repressor for arsenical resistance (ArsR family)	?	?	b3501‡	
<i>dgoR</i>	3872494	3873183	-1	galactonate operon transcriptional repressor	N	?	b3694	b4479
<i>dgoR</i>	3872494	3873183	-1	galactonate operon transcriptional repressor	N	?	b3695	b4479
<i>hdfR</i>	3945151	3945990	-1	Transcription regulator	N	F	b3762	b4480
<i>hdfR</i>	3945151	3945990	-1	Transcription regulator	N	F	b3763	b4480
<i>rhtC</i>	4005780	4006400	1	threonine efflux protein	N	F	b3823	
<i>rhtB</i>	4006462	4007082	-1	homoserine-homoserine lactone efflux protein	N	F	b3824	
<i>yiiE</i>	4077287	4077532	1	hypothetical protein	N	?	b3889	
<i>ytfA</i>	4425792	4426118	1	hypothetical protein	X	?	b4205	
<i>ytfR</i>	4449081	4450583	1	putative ATP-binding component of ABC transporter	E	?	b4228	b4485
<i>ytfR</i>	4449081	4450583	1	putative ATP-binding component of ABC transporter	E	?	b4229	b4485
<i>cybC</i>	4457576	4457878	1	Soluble cytochrome	?	F	b4236§	

b562 precursor							
<i>yjgM</i>	4477057	4477560	-1	hypothetical protein	X	?	b4256¶
<i>yjgW</i>	4497622	4497957	1	hypothetical protein	N	?	b4274
<i>yjgX</i>	4498066	4498512	-1	hypothetical protein	X	?	b4275
<i>yjgY</i>	4498455	4498904	-1	hypothetical protein	X	?	b4276
<i>lplA</i>	4621124	4622140	-1	lipoate-protein ligase A	E	F	b4386§
<i>ytjB</i>	4622168	4622812	-1	hypothetical protein	N	?	b4387
Z0020	19661	19795	-1	Unknown function	X	?	
Z0039	34049	34265	-1	Unknown function	X	?	
c0184	167332	167511	-1	hypothetical protein	X	?	
Z0362	310963	311189	-1	Unknown function	X	?	
Z0377	316663	316788	1	putative dehydrogenase	X	?	
c0500	407792	407944	-1	hypothetical protein	X	?	
c1472	581200	581343	-1	hypothetical protein	X	?	
<i>kdpE*</i>	715947	716093	-1	regulator of kdp operon (transcriptional effector)	X	-	
c0824	779463	779756	-1	hypothetical protein	X	?	
c1294	1096978	1097129	-1	hypothetical protein	X	?	
c1315	1112638	1112860	1	hypothetical protein	X	?	
<i>icdA*</i>	1210633	1210797	1	isocitrate dehydrogenase	X	-	
SF1157	1221294	1221468	1	hypothetical protein	X	-	
c1620	1223176	1223357	1	hypothetical protein	X	?	
c1675	1268929	1269200	-1	hypothetical protein	X	?	
c1702	1292240	1292365	-1	hypothetical protein	X	?	
SF1333	1388749	1388886	-1	hypothetical protein	X	-	
SF1821	1438985	1439345	1	hypothetical protein	X	-	
c1403	1647497	1648039	-1	hypothetical protein	X	?	
Z3023	2009242	2009379	-1	putative secreted protein	X	?	
SF2066	2076467	2076892	1	hypothetical protein	X	-	
c2541	2085269	2085403	-1	hypothetical protein	X	?	
c2624	2176643	2176786	1	hypothetical protein	X	?	
c2993	2580697	2580954	1	hypothetical protein	X	?	
c3016	2618854	2619195	1	hypothetical protein	X	?	
c3045	2651343	2651495	1	hypothetical protein	X	?	
c3369	2931900	2932034	1	hypothetical protein	X	?	
c3421	2962188	2962387	1	hypothetical protein	X	?	
c3433	2974326	2974538	1	hypothetical protein	X	?	
<i>yibA*</i>	3621518	3621868	1	putative lyase	X	-	
c4352	3697527	3697661	1	hypothetical protein	X	?	
c4359	3703382	3703751	1	hypothetical protein	X	?	
c4363	3705689	3706094	-1	hypothetical protein	X	?	
Z5002	3748548	3748736	1	hypothetical protein	X	?	
Z5148	3837773	3838132	1	Unknown function	X	?	
SF3908	4013719	4013822	-1	putative enzyme	?	?	¶ b3830
Z5428	4076339	4076494	1	Unknown function	X	?	
c4868	4106353	4106487	1	hypothetical protein	X	?	
<i>yijI</i>	4139880	4140113	1	hypothetical protein	X	-	¶

c4992	4233385	4233447	-1	hypothetical protein	X	?
SF4150	4266227	4266630	-1	hypothetical protein	X	-
c5055	4275920	4276154	1	hypothetical protein	X	?
c5065	4285224	4285419	1	hypothetical protein	X	?
c5129	4346390	4346545	1	hypothetical protein	X	?
c5270	4409863	4410009	-1	hypothetical protein	X	?
c5271	4409877	4410026	-1	hypothetical protein	X	?
c5334	4456927	4457308	-1	hypothetical protein	X	?
Z5892	4533843	4534081	-1	Unknown function	X	?

*Functional paralog is also present in the *E. coli* MG1655 genome.

†Pseudogenes without an original Blattner ID were matches to intergenic regions in (4,5).

‡Query genome has two paralogous copies, and the longer one was used to identify the *E. coli* MG1655 disrupted pseudogene.

§One query genome has a copy of equal length to the *E. coli* MG1655 gene, while another has a longer copy used to identify the putative pseudogene.

||Missequenced in original query genome.

¶Originally missequenced or misannotated in *E. coli* MG1655 genome, and corrected without changing Blattner ID.

Table S3. List of predicted K_a/K_s pseudogenes in *E. coli* MG1655 genome. Functional

assignments follow the notation used in Table S1.

Evidence at protein
level.

Gene	Beginning	End	Strand	Description	Essential	Function	Blattner ID
<i>yobF</i>	1905472	1905615	-1	unknown CDS	?	?	b1824
<i>yqgB</i>	3083942	3084088	-1	orf hypothetical protein	?	?	b2939
<i>yqgC</i>	3084209	3084424	1	orf hypothetical protein	?	?	b2940
<i>ygiA</i>	3177618	3177878	1	orf hypothetical protein	N	?	b3036

Table S4. List of predicted SIFT pseudogenes in *E. coli* MG1655 genome. Functional assignments follow the notation used in Table S1.

Gene	Beginning	End	Strand	Description	Essential	Function	Blattner ID
<i>thiE</i>	4191592	4192227	-1	thiamin-phosphate synthase (thiamine-phosphate pyrophosphorylase)	N	F	b3993
<i>tam</i>	1605370	1606128	1	trans-aconitate 2-methyltransferase	N	F	b1519
<i>yfdH</i>	2466236	2467156	1	bactoprenol glucosyl transferase homolog from prophage CPS-53	X	?	b2351
<i>ybiX</i>	837753	838466	-1	putative enzyme	N	?	b0804
<i>ybgA</i>	738224	738733	1	orf, hypothetical protein	X	?	b0707

Table S5. List of *E. coli* MG1655 genes without detectable transcripts in (6-8). Functional assignments follow the notation used in Table S1.

Gene	Beginning	End	Strand	Description	Essential	Function	Blattner ID
<i>fliA</i>	248358	250097	-1	flagellar biosynthesis	N	?	b0229
<i>ykfA</i>	267321	268187	-1	putative GTP-binding protein	N	?	b0253
<i>yagG</i>	284619	286001	1	putative permease	E	?	b0270
<i>yagH</i>	286013	287623	1	putative beta-xylosidase (EC 3.2.1.37)	E	?	b0271
<i>yagR</i>	297960	300158	-1	orf hypothetical protein	N	?	b0284
<i>yagS</i>	300155	301111	-1	orf hypothetical protein	E	?	b0285
<i>yahG</i>	337549	338967	1	orf hypothetical protein	N	?	b0321
<i>yahJ</i>	340349	341731	1	putative deaminase	N	?	b0324
<i>ylbE</i>	547838	548839	1	orf hypothetical protein	N	?	b0519
<i>nfrA</i>	587205	590177	-1	bacteriophage N4 receptor outer membrane protein	N	F	b0568
<i>citG</i>	645854	646732	-1	orf hypothetical protein	N	?	b0613
<i>rhcC</i>	728806	732999	1	rhcC protein in rhs element	N	N	b0700
<i>ssuD</i>	994066	995211	-1	orf hypothetical protein	N	F	b0935
<i>hyaE</i>	1035577	1035975	1	processing of HyaA and HyaB proteins	N	F	b0976
<i>etp</i>	1043453	1043911	-1	putative phosphatase	N	F	b0982
<i>ycdI</i>	1069588	1070178	-1	putative enzyme	N	?	b1008
<i>ycdJ*</i>	1070188	1070988	-1	putative acetyltransferase	N	?	b1009
<i>ycdL</i>	1071394	1072128	-1	putative synthetase	E	?	b1011
<i>flgE</i>	1131797	1133005	1	flagellar biosynthesis hook protein	N	?	b1076
<i>flgI</i>	1135497	1136594	1	homolog of Salmonella P-ring of flagella basal body	N	?	b1080
<i>ymfR</i>	1204772	1204954	1	orf hypothetical protein	?	?	b1150
<i>ycgV</i>	1252308	1255175	-1	putative adhesion and penetration protein	N	?	b1202
<i>aldH</i>	1360767	1362254	1	aldehyde dehydrogenase prefers NADP over NAD	N	F	b1300
<i>ycjS</i>	1374856	1375911	1	putative dehydrogenase	N	F	b1315
<i>abgB</i>	1399834	1401279	-1	aminobenzoyl-glutamate utilization protein	N	F	b1337
<i>ydaL</i>	1404003	1404566	1	orf hypothetical protein	E	?	b1340
<i>paaA</i>	1451951	1452880	1	phenylacetic acid degradation protein	N	F	b1388
<i>paaF</i>	1455521	1456288	1	probable enoyl-CoA hydratase	N	F	b1393
<i>paal</i>	1458495	1458917	1	orf hypothetical protein	N	?	b1396
<i>narZ</i>	1536874	1540614	-1	cryptic nitrate reductase 2 alpha subunit	N	F	b1468
<i>uidC</i>	1689610	1690863	-1	membrane-associated protein	N	?	b1615
<i>ydiS</i>	1779419	1780708	1	flavoprotein; probably electron transport	N	?	b1699
<i>tap</i>	1967407	1969008	-1	methyl-accepting chemotaxis protein IV peptide sensor receptor	N	F	b1885
<i>fliI</i>	2014578	2015951	1	flagellum-specific ATP synthase	N	F	b1941
<i>yedQ</i>	2024347	2026056	-1	orf hypothetical protein	N	?	b1956

<i>yeeT</i>	2074841	2075062	1	orf hypothetical protein	N	?	b2003
<i>cpsB</i>	2121108	2122544	-1	mannose-1-phosphate guanylttransferase	N	F	b2049
<i>wza</i>	2134128	2135267	-1	putative polysaccharide export protein	N	?	b2062
<i>mdtB</i>	2153287	2156409	1	orf hypothetical protein	N	F	b2075
<i>rihB</i>	2252267	2253208	-1	pyrimidine specific nucleoside hydrolase	N	F	b2162
<i>yfaS</i>	2327820	2332334	-1	putative membrane protein	N	N	b2228
<i>ypeC</i>	2509023	2509349	1	orf hypothetical protein	?	?	b2390
<i>yffO</i>	2560133	2560549	1	orf hypothetical protein	E	?	b2446
<i>eutD</i>	2570511	2571527	-1	ethanolamine utilization; homolog of Salmonella acetyl/butyryl P transferase	N	?	b2458
<i>eutP</i>	2573000	2573479	-1	orf hypothetical protein	N	?	b2461
<i>hyfB</i>	2599840	2601858	1	hydrogenase 4 membrane subunit	N	F	b2482
<i>hyfD</i>	2602833	2604272	1	hydrogenase 4 membrane subunit	N	F	b2484
b2641	2770998	2771114	-1	orf hypothetical protein	X	?	b2641
<i>gutQ</i>	2827874	2828800	1	orf hypothetical protein	N	?	b2708
<i>norW</i>	2831934	2833067	1	putative oxidoreductase	N	F	b2711
<i>hycG</i>	2841465	2842232	-1	hydrogenase activity	N	F	b2719
<i>ygfF</i>	3043180	3043923	-1	putative oxidoreductase	X	?	b2902
<i>fadH</i>	3229687	3231705	1	2,4-dieonyl-CoA reductase, FMN- linked	N	F	b3081
<i>uxaA</i>	3239849	3241336	-1	altronate hydrolase	N	F	b3091
<i>yhbV</i>	3300493	3301389	1	orf hypothetical protein	N	?	b3159
<i>nanK</i>	3367497	3368405	-1	ManNAc kinase	N	F	b3222
<i>hofQ</i>	3517487	3518725	-1	putative transport portein	N	F	b3391
<i>uhpB</i>	3846657	3848162	-1	sensor histidine protein kinase phosphorylates UhpA	N	F	b3668
<i>phnM</i>	4314105	4315241	-1	phosphonate metabolism	N	F	b4095
<i>phnK</i>	4316029	4316787	-1	ATP-binding component of phosphonate transport	X	F	b4097
<i>fumB</i>	4343703	4345349	-1	fumarase B= fumarate hydratase Class I; anaerobic isozyme	N	F	b4122
<i>deoC</i>	4615346	4616125	1	2-deoxyribose-5-phosphate aldolase	N	F	b4381

* Single gene without a detectable transcript in (9).

References to Supplemental Tables

1. F. R. Blattner, G. Plunkett 3rd, C. A. Bloch, N. T. Perna, V. Burland, *et al.*, *Science* **277**, 1453 (1997).
2. I. M. Keseler, J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I. T. Paulsen, M. Peralta-Gil, P. D. Karp. *Nucl. Acids Res.* **33**, D334 (2005).
3. G. T. Thomas. *Bioinformatics* **15**, 860 (1999)
4. K. Homma, S. Fukuchi, T. Kawabata, M. Ota, K. Nishikawa, *Gene* **294**, 25 (2002).
5. E. Lerat, H. Ochman, *Genome Res.* **14**, 2273 (2004).
6. A. Lobner-Olesen, M. G. Marinus, F. G. Hansen, *Proc. Nat. Acad. Sci. USA* **100**, 4672 (2003).
7. R. W. Corbin, O. Paliy, F. Yang, J. Shabanowitz, M. Platt, *et al.*, *Proc. Nat. Acad. Sci. USA* **100**, 9232 (2003).
8. S. J. Arends, D. S. Weiss, *J. Bacteriol.* **186**, 880 (2004).
9. D. W. Selinger, K. J. Cheung, R. Mei, E. M. Johansson, C. S. Richmond, *et al.*, *Nat. Biotechnol.* **18**, 1262 (2000).