The Dynamics of Genetic Interactions Between *Vibrio metoecus* and *Vibrio cholerae*, Two Close Relatives Co-Occurring in the Environment

SUPPLEMENTARY TABLES AND FIGURES

Fabini D. Orata[1], Paul C. Kirchberger[1], Raphaël Méheust[2],
E. Jed Barlow[3], Cheryl L. Tarr[4], Yan Boucher[1]

[1] Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada
[2] Unité Mixte de Recherche 7138, Evolution Paris-Seine, Institut de Biologie Paris-Seine, Université Pierre et Marie Curie, Paris, France
[3] Department of Computing Science, University of Alberta, Edmonton, AB, Canada
[4] National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA

Table S1

Source and year of isolation of the *V. metoecus* and *V. cholerae* strains used in this study. The genome sequences are available from the DDBJ, EMBL, and GenBank databases.

Species	Strain	Geographical Origin	Source of Isolation	Year of Isolation	Accession Number/s	Reference/s
Vibrio metoecus	06-2478	Mississippi, USA	Clinical (stool)	2006	LCUD00000000	This study
Vibrio metoecus	07-2435	North Carolina, USA	Clinical (leg wound)	2007	LCUE00000000	This study
Vibrio metoecus	08-2459	North Carolina, USA	Clinical (blood)	2008	LCUF00000000	This study
Vibrio metoecus	2010V-1005	USA	Clinical (ear)	2010	LCUG00000000	This study
Vibrio metoecus	YB4D01	Oyster Pond, Massachusetts, USA	Water	2009	LBGO00000000	This study
Vibrio metoecus	YB5B04	Oyster Pond, Massachusetts, USA	Water	2009	LBGP00000000	This study
Vibrio metoecus	YB5B06	Oyster Pond, Massachusetts, USA	Water	2009	LBGQ00000000	This study
Vibrio metoecus	YB9D03	Oyster Pond, Massachusetts, USA	Water	2009	LBGR00000000	This study
Vibrio metoecus	OP3H	Oyster Pond, Massachusetts, USA	Water	2006	JJMN00000000	Kirchberger et al. 2014
Vibrio metoecus	RC341	Chesapeake Bay, Maryland, USA	Water	1998	ACZT00000000	Haley et al. 2010a
Vibrio cholerae	YB1A01	Oyster Pond, Massachusetts, USA	Water	2009	LBCL00000000	This study
Vibrio cholerae	YB1G06	Oyster Pond, Massachusetts, USA	Water	2009	LBFV00000000	This study
Vibrio cholerae	YB2A05	Oyster Pond, Massachusetts, USA	Water	2009	LBFW00000000	This study
Vibrio cholerae	YB2A06	Oyster Pond, Massachusetts, USA	Water	2009	LBFX00000000	This study
Vibrio cholerae	YB2G01	Oyster Pond, Massachusetts, USA	Water	2009	LBFY00000000	This study
Vibrio cholerae	YB2G05	Oyster Pond, Massachusetts, USA	Water	2009	LBFZ00000000	This study
Vibrio cholerae	YB2G07	Oyster Pond, Massachusetts, USA	Water	2009	LBGA00000000	This study
Vibrio cholerae	YB3B05	Oyster Pond, Massachusetts, USA	Water	2009	LBGB00000000	This study
Vibrio cholerae	YB3G04	Oyster Pond, Massachusetts, USA	Water	2009	LBGC00000000	This study
Vibrio cholerae	YB4B03	Oyster Pond, Massachusetts, USA	Water	2009	LBGD00000000	This study
Vibrio cholerae	YB4C07	Oyster Pond, Massachusetts, USA	Water	2009	LBGE00000000	This study
Vibrio cholerae	YB4F05	Oyster Pond, Massachusetts, USA	Water	2009	LBGF00000000	This study
Vibrio cholerae	YB4G05	Oyster Pond, Massachusetts, USA	Water	2009	LBGG00000000	This study
Vibrio cholerae	YB4G06	Oyster Pond, Massachusetts, USA	Water	2009	LBGH00000000	This study
Vibrio cholerae	YB4H02	Oyster Pond, Massachusetts, USA	Water	2009	LBGI00000000	This study
Vibrio cholerae	YB5A06	Oyster Pond, Massachusetts, USA	Water	2009	LBGJ00000000	This study
Vibrio cholerae	YB6A06	Oyster Pond, Massachusetts, USA	Water	2009	LBGK00000000	This study
Vibrio cholerae	YB7A06	Oyster Pond, Massachusetts, USA	Water	2009	LBGL00000000	This study
Vibrio cholerae	YB7A09	Oyster Pond, Massachusetts, USA	Water	2009	LBGM00000000	This study
Vibrio cholerae	YB8E08	Oyster Pond, Massachusetts, USA	Water	2009	LBGN00000000	This study
Vibrio cholerae	12129(1)	Australia	Water	1985	ACFQ00000000	Chun et al. 2009
Vibrio cholerae	1587	Lima, Peru	Clinical	1994	AAUR00000000	Chun et al. 2009
Vibrio cholerae	2010EL-1786	Artibonite, Haiti	Clinical	2010	CP003069; CP003070	Reimer et al. 2011
Vibrio cholerae	2740-80	Gulf Coast, USA	Water	1980	AAUT00000000	Chun et al. 2009
Vibrio cholerae	623-39	Bangladesh	Water	2002	AAWG00000000	Chun et al. 2009
Vibrio cholerae	877-163	Bangladesh	Water	2002	LBNV00000000	This study
Vibrio cholerae	AM-19226	Bangladesh	Clinical	2001	AATY00000000	Chun et al. 2009
Vibrio cholerae	BX 330286	Australia	Water	1986	ACIA00000000	Chun et al. 2009; Haley et al. 2010b
Vibrio cholerae	CIRS101	Dhaka, Bangladesh	Clinical	2002	ACVW00000000	Chun et al. 2009; Grim CJ, et al. 2010
Vibrio cholerae	MAK757	Sulawesi, Indonesia	Clinical	1937	AAUS00000000	Chun et al. 2009
Vibrio cholerae	MJ-1236	Matlab, Bangladesh	Clinical	1994	CP001485; CP001486	Chun et al. 2009; Grim CJ, et al. 2010
Vibrio cholerae	MO10	Madras, India	Clinical	1992	AAKF00000000	Chun et al. 2009
Vibrio cholerae	MZO-2	Bangladesh	Clinical	2001	AAWF00000000	Chun et al. 2009
Vibrio cholerae	MZO-3	Bangladesh	Clinical	2001	AAUU00000000	Chun et al. 2009
Vibrio cholerae	N16961	Bangladesh	Clinical	1971	AE003852; AE003853	Heidelberg et al. 2000
Vibrio cholerae	O395	India	Clinical	1965	CP001235; CP001236	Feng et al. 2008
Vibrio cholerae	RC385	Chesapeake Bay, Maryland, USA	Plankton	1998	AAKH00000000	Chun et al. 2009
Vibrio cholerae	TM 11079-80	Brazil	Sewage	1980	ACHW00000000	Chun et al. 2009
Vibrio cholerae	TMA21	Brazil	Seawater	1982	ACHY00000000	Chun et al. 2009
Vibrio cholerae	V51	USA	Clinical	1987	AAKI00000000	Chun et al. 2009
Vibrio cholerae	V52	Sudan	Clinical	1968	AAKJ00000000	Chun et al. 2009
Vibrio cholerae	VL426	Maidstone, Kent, UK	Water	Unknown	ACHV00000000	Chun et al. 2009

Literature Cited

Chun J, et al. 2009. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc Natl Acad Sci USA. 106:15442-15447.

Feng L, et al. 2008. A recalibrated molecular clock and independent origins for the cholera pandemic clones. PLoS One. 3:e4053.

Grim CJ, et al. 2010. Genome sequence of hybrid *Vibrio cholerae* O1 MJ-1236, B-33, and CIRS101 and comparative genomics with *V. cholerae*. J Bacteriol. 192:3524-3533.

Haley BJ, et al. 2010a. Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to *V. cholerae*. BMC Microbiol. 10:154.

Haley BJ, et al. 2010b. The pre-seventh pandemic *Vibrio cholerae* BX 330286 El Tor genome: evidence for the environment as a genome reservoir. Environ Microbiol Rep. 2:208-216.

Heidelberg JF, et al. 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature. 406:477-483.

Kirchberger PC, et al. 2014. *Vibrio metoecus* sp. nov., a close relative of *Vibrio cholerae* isolated from coastal brackish ponds and clinical specimens. Int J Syst Evol Microbiol. 64:3208-3214.

Reimer AR, et al. 2011. Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. Emerg Infect Dis. 17:2113-2121.

Table S2

Percentage of genes shared and the ANI of the core genome (from 1,560 single-copy core genes, ≈ 1.42 mbp) and the chromosomal integron region from intra- and interspecies pairwise comparison of *V. cholerae* (Vc) and *V. metoecus* (Vm) strains.

Strain 1	Strain 2	Unique Genes 1[a]	Unique Genes 2[b]	Core Genes	% Shared 1[c]	% Shared 2[d]	ANI Core (%)	ANI Integron (%)
Vc MZO-3	Vc N16961	405	634	2,985	88.05	82.48	98.51	96.40
Vc MZO-3	Vc O395	359	582	3,031	89.41	83.89	98.50	94.83
Vc MZO-3	Vc RC385	413	588	2,977	87.82	83.51	98.00	94.19
Vc MZO-3	Vc TM 11079-80	436	532	2,954	87.14	84.74	98.27	94.17
Vc MZO-3	Vc VL426	439	412	2,951	87.05	87.75	98.02	93.33
Vc MZO-3	Vc YB1A01	457	492	2,933	86.52	85.64	98.12	93.68
Vc MZO-3	Vc YB1G06	401	486	2,989	88.17	86.01	98.17	93.53
Vc MZO-3	Vc YB2A06	361	612	3,029	89.35	83.19	98.28	94.05
Vc MZO-3	Vc YB4B03	333	455	3,057	90.18	87.04	98.51	93.04
Vc MZO-3	Vc YB4C07	245	455	3,145	92.77	87.36	98.66	95.36
Vc MZO-3	Vm 06-2478	543	734	2,847	83.98	79.50	86.55	93.78
Vc MZO-3	Vm 07-2435	600	557	2,833	82.52	83.57	87.09	93.46
Vc MZO-3	Vm 08-2459	852	560	2,830	76.86	83.48	86.34	94.21
Vc MZO-3	Vm 2010V-1005	512	782	2,878	84.90	78.63	86.49	93.69
Vc MZO-3	Vm OP3H	527	755	2,863	84.45	79.13	86.67	93.18
Vc MZO-3	Vm RC341	567	612	2,823	83.27	82.18	86.51	93.20
Vc MZO-3	Vm YB4D01	524	680	2,866	84.54	80.82	86.54	93.74
Vc MZO-3	Vm YB5B04	691	475	2,915	80.84	85.99	86.52	92.84
Vc MZO-3	Vm YB5B06	663	583	2,807	80.89	82.80	86.48	92.69
Vc MZO-3	Vm YB9D03	553	683	2,837	83.69	80.60	86.47	95.17
Vc N16961	Vc O395	300	294	3,319	91.71	91.86	99.36	99.19
Vc N16961	Vc RC385	497	551	3,068	86.06	84.77	98.02	95.32
Vc N16961	Vc TM 11079-80	379	512	3,107	89.13	85.85	98.26	94.79
Vc N16961	Vc VL426	327	583	3,036	90.28	83.89	98.09	92.90
Vc N16961	Vc YB1A01	538	344	3,081	85.13	89.96	98.22	95.23
Vc N16961	Vc YB1G06	385	529	3,090	88.92	85.38	98.23	94.92
Vc N16961	Vc YB2A06	483	461	3,158	86.73	87.26	98.33	94.90
Vc N16961	Vc YB4B03	373	480	3,139	89.38	86.74	98.49	94.86
Vc N16961	Vc YB4C07	419	438	3,181	88.36	87.90	98.57	96.24
Vc N16961	Vm 06-2478	587	625	2,994	83.61	82.73	86.49	92.70
Vc N16961	Vm 07-2435	458	644	2,975	86.66	82.21	87.04	94.69
Vc N16961	Vm 08-2459	698	635	2,984	81.04	82.45	86.31	94.44
Vc N16961	Vm 2010V-1005	656	615	3,004	82.08	83.01	86.45	94.83
Vc N16961	Vm OP3H	638	637	2,981	82.37	82.39	86.63	94.27
Vc N16961	Vm RC341	524	708	2,911	84.75	80.44	86.49	93.81
Vc N16961	Vm YB4D01	560	633	2,986	84.21	82.51	86.49	94.95
Vc N16961	Vm YB5B04	609	622	2,997	83.11	82.81	86.48	95.18
Vc N16961	Vm YB5B06	486	635	2,984	85.99	82.45	86.48	91.89
Vc N16961	Vm YB9D03	521	620	2,999	85.20	82.87	86.42	95.43
Vc O395	Vc RC385	448	496	3,117	87.43	86.27	98.01	93.90
Vc O395	Vc TM 11079-80	272	399	3,214	92.20	88.96	98.25	93.70
Vc O395	Vc VL426	274	524	3,089	91.85	85.50	98.07	92.74
Vc O395	Vc YB1A01	287	475	3,138	91.62	86.85	98.21	93.85
Vc O395	Vc YB1G06	328	466	3,147	90.56	87.10	98.23	93.51
Vc O395	Vc YB2A06	390	362	3,251	89.29	89.98	98.31	94.35
Vc O395	Vc YB4B03	321	422	3,191	90.86	88.32	98.48	94.29
Vc O395	Vc YB4C07	350	363	3,250	90.28	89.95	98.57	95.12
Vc O395	Vm 06-2478	549	581	3,032	84.67	83.92	86.47	92.80
Vc O395	Vm 07-2435	404	584	3,029	88.23	83.84	87.02	93.18
Vc O395	Vm 08-2459	657	588	3,025	82.16	83.73	86.28	93.26
Vc O395	Vm 2010V-1005	622	575	3,038	83.01	84.09	86.42	94.29
Vc O395	Vm OP3H	606	601	3,012	83.25	83.37	86.60	93.67
Vc O395	Vm RC341	491	669	2,944	85.71	81.48	86.47	92.89
Vc O395	Vm YB4D01	516	583	3,030	85.45	83.86	86.47	94.14
Vc O395	Vm YB5B04	565	572	3,041	84.33	84.17	86.45	94.35
Vc O395	Vm YB5B06	430	573	3,040	87.61	84.14	86.45	92.79
Vc O395	Vm YB9D03	485	578	3,035	86.22	84.00	86.39	94.56
Vc RC385	Vc TM 11079-80	492	413	3,073	86.20	88.15	97.89	93.49
Vc RC385	Vc VL426	236	438	3,127	92.98	87.71	97.95	93.38
Vc RC385	Vc YB1A01	399	259	3,166	88.81	92.44	98.29	94.29
Vc RC385	Vc YB1G06	365	275	3,200	89.76	92.09	98.25	93.33
Vc RC385	Vc YB2A06	439	515	3,126	87.69	85.86	98.04	93.40
Vc RC385	Vc YB4B03	375	428	3,137	89.32	87.99	98.08	92.10
Vc RC385	Vc YB4C07	487	452	3,113	86.47	87.32	98.00	93.36
Vc RC385	Vm 06-2478	583	599	2,982	83.65	83.27	86.52	93.83
Vc RC385	Vm 07-2435	504	636	2,929	85.32	82.16	86.90	92.51
Vc RC385	Vm 08-2459	745	628	2,937	79.77	82.38	86.27	93.74
Vc RC385	Vm 2010V-1005	556	651	3,009	84.40	82.21	86.41	93.10
Vc RC385	Vm OP3H	638	691	2,927	82.10	80.90	86.57	92.86

Vc RC385	Vm RC341	502	632	2,933	85.39	82.27	86.44	91.49
Vc RC385	Vm YB4D01	578	559	2,987	83.79	84.24	86.47	93.42
Vc RC385	Vm YB5B04	614	573	2,992	82.97	83.93	86.46	93.08
Vc RC385	Vm YB5B06	534	629	2,936	84.61	82.36	86.43	92.51
Vc RC385	Vm YB9D03	614	569	2,951	82.78	83.84	86.40	94.39
Vc TM 11079-80	Vc VL426	326	449	3,037	90.31	87.12	97.94	93.46
Vc TM 11079-80	Vc YB1A01	432	371	3,054	87.61	89.17	98.06	94.17
Vc TM 11079-80	Vc YB1G06	404	415	3,071	88.37	88.10	98.07	94.42
Vc TM 11079-80	Vc YB2A06	519	364	3,122	85.75	89.56	98.16	95.07
Vc TM 11079-80	Vc YB4B03	388	362	3,124	88.95	89.62	98.31	93.72
Vc TM 11079-80	Vc YB4C07	473	359	3,127	86.86	89.70	98.29	95.07
Vc TM 11079-80	Vm 06-2478	599	504	2,982	83.27	85.54	86.47	93.75
Vc TM 11079-80	Vm 07-2435	446	499	2,987	87.01	85.69	86.97	94.09
Vc TM 11079-80	Vm 08-2459	735	539	2,947	80.04	84.54	86.22	93.79
Vc TM 11079-80	Vm 2010V-1005	672	498	2,988	81.64	85.71	86.43	93.81
Vc TM 11079-80	Vm OP3H	517	649	2,969	85.17	82.06	86.57	93.20
Vc TM 11079-80	Vm RC341	563	614	2,872	83.61	82.39	86.39	90.77
Vc TM 11079-80	Vm YB4D01	554	494	2,992	84.38	85.83	86.44	91.32
Vc TM 11079-80	Vm YB5B04	612	492	2,994	83.03	85.89	86.42	91.93
Vc TM 11079-80	Vm YB5B06	477	493	2,993	86.25	85.86	86.39	93.77
Vc TM 11079-80	Vm YB9D03	539	505	2,981	84.69	85.51	86.39	95.18
Vc VL426	Vc YB1A01	236	298	3,127	92.98	91.30	98.20	93.67
Vc VL426	Vc YB1G06	203	315	3,160	93.96	90.94	98.15	95.22
Vc VL426	Vc YB2A06	262	540	3,101	92.21	85.17	97.99	94.88
Vc VL426	Vc YB4B03	397	248	3,115	88.70	92.63	98.04	94.15
Vc VL426	Vc YB4C07	486	249	3,114	86.50	92.60	98.02	92.52
Vc VL426	Vm 06-2478	432	650	2,931	87.15	81.85	86.42	95.30
Vc VL426	Vm 07-2435	525	455	2,908	84.71	86.47	86.92	94.35
Vc VL426	Vm 08-2459	787	468	2,895	78.63	86.08	86.21	94.31
Vc VL426	Vm 2010V-1005	415	712	2,948	87.66	80.55	86.36	94.92
Vc VL426	Vm OP3H	448	703	2,915	86.68	80.57	86.54	93.09
Vc VL426	Vm RC341	552	480	2,883	83.93	85.73	86.36	91.69
Vc VL426	Vm YB4D01	408	591	2,955	87.87	83.33	86.41	93.94
Vc VL426	Vm YB5B04	647	404	2,959	82.06	87.99	86.37	93.00
Vc VL426	Vm YB5B06	551	444	2,919	84.12	86.80	86.36	93.99
Vc VL426	Vm YB9D03	450	607	2,913	86.62	82.76	86.33	95.70
Vc YB1A01	Vc YB1G06	253	203	3,222	92.72	94.07	98.64	94.48
Vc YB1A01	Vc YB2A06	494	278	3,147	86.43	91.88	98.25	94.16
Vc YB1A01	Vc YB4B03	377	290	3,135	89.27	91.53	98.25	93.59
Vc YB1A01	Vc YB4C07	453	278	3,147	87.42	91.88	98.17	93.96
Vc YB1A01	Vm 06-2478	625	469	2,956	82.55	86.31	86.50	92.35
Vc YB1A01	Vm 07-2435	491	483	2,942	85.70	85.90	86.96	92.34
Vc YB1A01	Vm 08-2459	738	481	2,944	79.96	85.96	86.27	92.71
Vc YB1A01	Vm 2010V-1005	702	467	2,958	80.82	86.36	86.42	93.95
Vc YB1A01	Vm OP3H	480	673	2,945	85.99	81.40	86.60	93.17
Vc YB1A01	Vm RC341	543	533	2,892	84.19	84.44	86.43	92.68
Vc YB1A01	Vm YB4D01	575	454	2,971	83.78	86.74	86.48	93.42
Vc YB1A01	Vm YB5B04	634	453	2,972	82.42	86.77	86.44	93.27
Vc YB1A01	Vm YB5B06	515	470	2,955	85.16	86.28	86.43	90.84
Vc YB1A01	Vm YB9D03	562	467	2,958	84.03	86.36	86.41	94.30
Vc YB1G06	Vc YB2A06	472	306	3,169	87.04	91.19	98.22	93.79
Vc YB1G06	Vc YB4B03	333	296	3,179	90.52	91.48	98.26	94.02
Vc YB1G06	Vc YB4C07	437	312	3,163	87.86	91.02	98.21	94.58
Vc YB1G06	Vm 06-2478	482	588	2,993	86.13	83.58	86.49	93.55
Vc YB1G06	Vm 07-2435	487	529	2,946	85.81	84.78	86.99	93.17
Vc YB1G06	Vm 08-2459	728	521	2,954	80.23	85.01	86.29	95.07
Vc YB1G06	Vm 2010V-1005	462	647	3,013	86.71	82.32	86.47	94.13
Vc YB1G06	Vm OP3H	518	661	2,957	85.09	81.73	86.64	93.30
Vc YB1G06	Vm RC341	518	558	2,917	84.92	83.94	86.46	93.78
Vc YB1G06	Vm YB4D01	473	544	3,002	86.39	84.66	86.53	93.06
Vc YB1G06	Vm YB5B04	588	457	3,018	83.69	86.85	86.46	92.57
Vc YB1G06	Vm YB5B06	507	512	2,963	85.39	85.27	86.45	92.10
Vc YB1G06	Vm YB9D03	534	489	2,986	84.83	85.93	86.43	95.18
Vc YB2A06	Vc YB4B03	250	379	3,262	92.88	89.59	98.38	93.92
Vc YB2A06	Vc YB4C07	366	407	3,234	89.83	88.82	98.35	93.54
Vc YB2A06	Vm 06-2478	609	549	3,032	83.27	84.67	86.58	92.12
Vc YB2A06	Vm 07-2435	406	614	3,027	88.17	83.14	87.09	92.66
Vc YB2A06	Vm 08-2459	660	619	3,022	82.07	83.00	86.39	92.59
Vc YB2A06	Vm 2010V-1005	587	606	3,054	83.88	83.44	86.50	93.99
Vc YB2A06	Vm OP3H	607	584	3,034	83.33	83.86	86.71	93.63
Vc YB2A06	Vm RC341	477	683	2,958	86.11	81.24	86.55	92.21
Vc YB2A06	Vm YB4D01	605	510	3,036	83.38	85.62	86.57	93.42
Vc YB2A06	Vm YB5B04	563	598	3,043	84.39	83.58	86.50	92.91
Vc YB2A06	Vm YB5B06	439	610	3,031	87.35	83.25	86.54	93.46

Vc YB2A06	Vm YB9D03	457	578	3,063	87.02	84.13	86.48	94.26
Vc YB4B03	Vc YB4C07	375	287	3,225	89.58	91.83	98.51	94.72
Vc YB4B03	Vm 06-2478	454	523	3,058	87.07	85.40	86.52	92.53
Vc YB4B03	Vm 07-2435	415	494	3,018	87.91	85.93	87.02	92.76
Vc YB4B03	Vm 08-2459	680	510	3,002	81.53	85.48	86.31	93.65
Vc YB4B03	Vm 2010V-1005	468	616	3,044	86.67	83.17	86.46	93.33
Vc YB4B03	Vm OP3H	509	615	3,003	85.51	83.00	86.65	93.72
Vc YB4B03	Vm RC341	473	550	2,962	86.23	84.34	86.48	90.46
Vc YB4B03	Vm YB4D01	452	486	3,060	87.13	86.29	86.51	93.31
Vc YB4B03	Vm YB5B04	532	438	3,074	85.25	87.53	86.48	91.69
Vc YB4B03	Vm YB5B06	461	503	3,009	86.71	85.68	86.47	90.91
Vc YB4B03	Vm YB9D03	500	508	3,012	85.76	85.57	86.44	94.02
Vc YB4C07	Vm 06-2478	575	556	3,025	84.03	84.47	86.53	92.40
Vc YB4C07	Vm 07-2435	414	581	3,019	87.94	83.86	87.11	93.71
Vc YB4C07	Vm 08-2459	674	592	3,008	81.69	83.56	86.32	93.08
Vc YB4C07	Vm 2010V-1005	568	628	3,032	84.22	82.84	86.45	94.34
Vc YB4C07	Vm OP3H	544	562	3,056	84.89	84.47	86.64	94.52
Vc YB4C07	Vm RC341	484	649	2,951	85.91	81.97	86.50	92.73
Vc YB4C07	Vm YB4D01	562	508	3,038	84.39	85.67	86.53	95.14
Vc YB4C07	Vm YB5B04	528	522	3,078	85.36	85.50	86.51	94.35
Vc YB4C07	Vm YB5B06	457	587	3,013	86.83	83.69	86.50	91.59
Vc YB4C07	Vm YB9D03	570	490	3,030	84.17	86.08	86.43	93.75
Vm 06-2478	Vm 07-2435	266	414	3,167	92.25	88.44	93.61	93.51
Vm 06-2478	Vm 08-2459	439	338	3,243	88.08	90.56	97.59	93.04
Vm 06-2478	Vm 2010V-1005	366	287	3,294	90.00	91.99	97.89	94.62
Vm 06-2478	Vm OP3H	344	381	3,237	90.39	89.47	97.48	93.02
Vm 06-2478	Vm RC341	289	435	3,146	91.59	87.85	97.81	92.15
Vm 06-2478	Vm YB4D01	281	316	3,265	92.08	91.18	97.69	93.91
Vm 06-2478	Vm YB5B04	281	256	3,325	92.21	92.85	98.17	94.09
Vm 06-2478	Vm YB5B06	251	362	3,219	92.77	89.89	97.68	90.99
Vm 06-2478	Vm YB9D03	249	310	3,271	92.93	91.34	97.78	93.49
Vm 07-2435	Vm 08-2459	537	288	3,145	85.42	91.61	93.62	93.34
Vm 07-2435	Vm 2010V-1005	280	507	3,153	91.84	86.15	93.48	92.83
Vm 07-2435	Vm OP3H	290	475	3,143	91.55	86.87	93.71	93.33
Vm 07-2435	Vm RC341	404	406	3,029	88.23	88.18	93.54	91.71
Vm 07-2435	Vm YB4D01	274	387	3,159	92.02	89.09	93.79	93.73
Vm 07-2435	Vm YB5B04	434	261	3,172	87.96	92.40	93.77	93.56
Vm 07-2435	Vm YB5B06	309	272	3,161	91.10	92.08	93.79	90.56
Vm 07-2435	Vm YB9D03	255	342	3,178	92.57	90.28	93.70	94.12
Vm 08-2459	Vm 2010V-1005	404	382	3,278	89.03	89.56	97.70	93.36
Vm 08-2459	Vm OP3H	447	383	3,235	87.86	89.41	97.46	92.64
Vm 08-2459	Vm RC341	588	341	3,094	84.03	90.07	97.66	92.29
Vm 08-2459	Vm YB4D01	492	356	3,190	86.64	89.96	97.58	92.90
Vm 08-2459	Vm YB5B04	410	486	3,196	88.63	86.80	97.63	93.12
Vm 08-2459	Vm YB5B06	302	514	3,168	91.30	86.04	97.65	90.24
Vm 08-2459	Vm YB9D03	481	319	3,201	86.94	90.94	97.70	93.52
Vm 2010V-1005	Vm OP3H	422	380	3,238	88.47	89.50	97.57	93.00
Vm 2010V-1005	Vm RC341	293	518	3,142	91.47	85.85	97.93	92.66
Vm 2010V-1005	Vm YB4D01	268	382	3,278	92.44	89.56	97.78	93.26
Vm 2010V-1005	Vm YB5B04	348	402	3,258	90.35	89.02	97.86	93.71
Vm 2010V-1005	Vm YB5B06	256	446	3,214	92.62	87.81	97.78	93.10
Vm 2010V-1005	Vm YB9D03	279	419	3,241	92.07	88.55	97.89	94.61
Vm OP3H	Vm RC341	344	527	3,091	89.99	85.43	97.56	92.37
Vm OP3H	Vm YB4D01	250	322	3,296	92.95	91.10	97.97	93.76
Vm OP3H	Vm YB5B04	310	322	3,296	91.40	91.10	97.46	93.79
Vm OP3H	Vm YB5B06	203	351	3,267	94.15	90.30	97.84	92.75
Vm OP3H	Vm YB9D03	277	375	3,243	92.13	89.64	97.61	93.74
Vm RC341	Vm YB4D01	270	381	3,165	92.14	89.26	97.74	92.93
Vm RC341	Vm YB5B04	418	247	3,188	88.41	92.81	97.79	94.70
Vm RC341	Vm YB5B06	378	343	3,092	89.11	90.01	97.72	91.94
Vm RC341	Vm YB9D03	324	409	3,111	90.57	88.38	97.86	94.34
Vm YB4D01	Vm YB5B04	292	232	3,314	91.90	93.46	97.74	93.45
Vm YB4D01	Vm YB5B06	179	255	3,291	94.84	92.81	98.02	90.71
Vm YB4D01	Vm YB9D03	240	266	3,280	93.18	92.50	97.83	93.91
Vm YB5B04	Vm YB5B06	340	204	3,266	90.57	94.12	97.73	91.65
Vm YB5B04	Vm YB9D03	322	236	3,284	91.07	93.30	97.76	94.37
Vm YB5B06	Vm YB9D03	203	253	3,267	94.15	92.81	97.88	94.25

[a] Unique genes in strain 1
[b] Unique genes in strain 2
[c] Percentage of genes shared by strain 1
[d] Percentage of genes shared by strain 2

Table S3

Integron regions extracted from *V. metoecus* and *V. cholerae* genomes.

Species	Strain	Size (bp)	GC (%)	Total ORFs	Hypothetical ORFs (%)
V. metoecus	06-2478	91,284	41.1	115	52.2
V. metoecus	07-2435	101,971	41.9	147	57.8
V. metoecus	08-2459	140,984	41.3	208	65.9
V. metoecus	2010V-1005	121,408	41.7	156	62.2
V. metoecus	OP3H	117,537	41.4	156	64.1
V. metoecus	RC341	71,256	42.0	85	56.5
V. metoecus	YB4D01	128,849	40.7	179	69.3
V. metoecus	YB5B04	86,814	41.5	121	71.9
V. metoecus	YB5B06	111,519	41.0	147	72.1
V. metoecus	YB9D03	117,178	41.3	188	69.7
V. cholerae	MZO-3	132,885	42.5	214	48.1
V. cholerae	N16961	125,669	42.2	187	75.4
V. cholerae	O395	116,524	41.4	199	80.4
V. cholerae	RC385	125,710	42.0	219	42.9
V. cholerae	TM 11079-80	62,779	42.9	106	38.7
V. cholerae	VL426	56,384	42.0	57	47.4
V. cholerae	YB1A01	90,204	41.4	109	64.2
V. cholerae	YB1G06	71,516	42.0	90	50.0
V. cholerae	YB2A06	116,704	41.6	164	64.0
V. cholerae	YB4B03	95,111	41.7	119	56.3
V. cholerae	YB4C07	67,955	42.4	97	48.5

Table S4

Gene function prediction and HGT inference for the *V. metoecus*-*V. cholerae* 1,947 core genes. Locus tags are based from the *V. cholerae* N16961 genome annotation. Functional roles for genes were predicted based on the COG database. Some genes have multiple COG hits (indicated by * on locus tag). Duplication was called if at least one of the strains has more than one copy of a gene in an orthologous gene family. HGT was inferred if all members of *V. metoecus* and *V. cholerae* could not be split into distinct clades in a phylogenetic tree.

Family Name[a]	Locus Tag	Duplication?	HGT?	COG Number	Description	Class	Class Description
fam4_fq	VC_0004	No duplication	No	COG0706	Preprotein translocase subunit YidC	U	Intracellular trafficking, secretion, and vesicular transport
fam4_fp	VC_0013	No duplication	HGT	COG0592	DNA polymerase sliding clamp subunit (PCNA homolog)	L	Replication, recombination and repair
fam4_fo	VC_0014	No duplication	HGT	COG1195	Recombinational DNA repair ATPase (RecF pathway)	L	Replication, recombination and repair
fam4_fn	VC_0015	No duplication	No	COG0187	Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), B subunit	L	Replication, recombination and repair
fam4_fm	VC_0016	No duplication	No	No hit			
fam4_fl	VC_0018	No duplication	HGT	COG0071	Molecular chaperone (small heat shock protein)	O	Posttranslational modification, protein turnover, chaperones
fam1_jn	VC_0019	Duplication	HGT	COG3977	Alanine-alpha-ketoisovalerate (or valine-pyruvate) aminotransferase	E	Amino acid transport and metabolism
fam4_fk	VC_0020	No duplication	No	COG0751	Glycyl-tRNA synthetase, beta subunit	J	Translation, ribosomal structure and biogenesis
fam4_fj	VC_0021	No duplication	HGT	COG0752	Glycyl-tRNA synthetase, alpha subunit	J	Translation, ribosomal structure and biogenesis
fam4_fi	VC_0022	No duplication	HGT	COG2119	Predicted membrane protein	S	Function unknown
fam4_fh	VC_0023	No duplication	HGT	No hit			
fam4_ff	VC_0024	No duplication	No	COG0425	Predicted redox protein, regulator of disulfide bond formation	O	Posttranslational modification, protein turnover, chaperones
fam4_fe	VC_0025	No duplication	No	No hit			
fam4_fd	VC_0026	No duplication	No	COG0604	NADPH:quinone reductase and related Zn-dependent oxidoreductases	CR	Multiple classes
fam4_fc	VC_0027	No duplication	No	COG1171	Threonine dehydratase	E	Amino acid transport and metabolism
fam4_fb	VC_0028	No duplication	No	COG0129	Dihydroxyacid dehydratase/phosphogluconate dehydratase	EG	Multiple classes
fam4_fa	VC_0029	No duplication	HGT	COG0115	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase	EH	Multiple classes
fam4_ez	VC_0030	No duplication	HGT	COG3978	Acetolactate synthase (isozyme II), small (regulatory) subunit	S	Function unknown
fam4_ey	VC_0031	No duplication	HGT	COG0028	Thiamine pyrophosphate-requiring enzymes	EH	Multiple classes
fam4_ex	VC_0032	No duplication	No	COG0606	Predicted ATPase with chaperone activity	O	Posttranslational modification, protein turnover, chaperones
fam1_pi	VC_0034	Duplication	No	COG1651	Protein-disulfide isomerase	O	Posttranslational modification, protein turnover, chaperones
fam4_ew	VC_0035	No duplication	No	COG2334	Putative homoserine kinase type II (protein kinase fold)	R	General function prediction only
fam4_ev	VC_0037	No duplication	No	COG3084	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_eu	VC_0039	No duplication	No	COG4326	Sporulation control protein	R	General function prediction only
fam4_et	VC_0040	No duplication	No	COG1272	Predicted membrane protein, hemolysin III homolog	R	General function prediction only
fam4_es	VC_0041	No duplication	No	No hit			
fam1_ie	VC_0042	Duplication	No	COG0168	Trk-type K+ transport systems, membrane components	P	Inorganic ion transport and metabolism
fam4_er	VC_0043	No duplication	No	COG0569	K+ transport systems, NAD-binding component	P	Inorganic ion transport and metabolism
fam4_eq	VC_0044	No duplication	No	COG0144	tRNA and rRNA cytosine-C5-methylases	J	Translation, ribosomal structure and biogenesis
fam4_ep	VC_0045	No duplication	No	COG0223	Methionyl-tRNA formyltransferase	J	Translation, ribosomal structure and biogenesis
fam4_eo	VC_0046	No duplication	No	COG0242	N-formylmethionyl-tRNA deformylase	J	Translation, ribosomal structure and biogenesis
fam4_en	VC_0047	No duplication	No	COG1652	Uncharacterized protein containing LysM domain	S	Function unknown
fam4_em	VC_0048	No duplication	No	COG0758	Predicted Rossmann fold nucleotide-binding protein involved in DNA uptake	LU	Multiple classes
fam4_el	VC_0049	No duplication	No	COG2922	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_ek	VC_0054	No duplication	No	COG0009	Putative translation factor (SUA5)	J	Translation, ribosomal structure and biogenesis
fam4_ej	VC_0055	No duplication	HGT	COG0408	Coproporphyrinogen III oxidase	H	Coenzyme transport and metabolism
fam4_ei	VC_0056	No duplication	No	COG0169	Shikimate 5-dehydrogenase	E	Amino acid transport and metabolism
fam4_eh	VC_0057	No duplication	No	No hit			
fam4_eg	VC_0058	No duplication	No	COG0663	Carbonic anhydrases/acetyltransferases, isoleucine patch superfamily	R	General function prediction only
fam3_iq	VC_0061	No duplication	No	COG0422	Thiamine biosynthesis protein ThiC	H	Coenzyme transport and metabolism
fam3_fg	VC_0083	No duplication	HGT	COG2226	Methylase involved in ubiquinone/menaquinone biosynthesis	H	Coenzyme transport and metabolism
fam3_ex	VC_0087	No duplication	No	COG1826	Sec-independent protein secretion pathway components	U	Intracellular trafficking, secretion, and vesicular transport
fam1_gy	VC_0111	Duplication	No	COG0218	Predicted GTPase	R	General function prediction only
fam3_ce	VC_0113	No duplication	No	COG2227	2-polyprenyl-3-methyl-5-hydroxy-6-metoxy-1,4-benzoquinol methylase	H	Coenzyme transport and metabolism
fam3_ca	VC_0114	No duplication	No	COG3078	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_bs	VC_0115	No duplication	No	No hit			
fam3_bn	VC_0116	No duplication	No	COG0635	Coproporphyrinogen III oxidase and related Fe-S oxidoreductases	H	Coenzyme transport and metabolism
fam3_bl	VC_0117	No duplication	No	COG3071	Uncharacterized enzyme of heme biosynthesis	H	Coenzyme transport and metabolism
fam3_bj	VC_0118	No duplication	No	COG2959	Uncharacterized enzyme of heme biosynthesis	H	Coenzyme transport and metabolism
fam3_bc	VC_0119	No duplication	No	COG1587	Uroporphyrinogen-III synthase	H	Coenzyme transport and metabolism
fam3_at	VC_0120	No duplication	No	COG0181	Porphobilinogen deaminase	H	Coenzyme transport and metabolism
fam2_sy	VC_0125	No duplication	No	COG0019	Diaminopimelate decarboxylase	E	Amino acid transport and metabolism
fam2_so	VC_0126	No duplication	No	COG0253	Diaminopimelate epimerase	E	Amino acid transport and metabolism
fam2_sa	VC_0128	No duplication	No	COG4973	Site-specific recombinase XerC	L	Replication, recombination and repair
fam2_rs	VC_0129	No duplication	No	COG1011	Predicted hydrolase (HAD superfamily)	R	General function prediction only
fam1_mg	VC_0130*	Duplication	No	COG3287	Uncharacterized conserved protein	S	Function unknown
fam1_mg	VC_0130*	Duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam1_mg	VC_0130*	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_rb	VC_0132	No duplication	No	COG3650	Predicted membrane protein	S	Function unknown
fam2_qr	VC_0134	No duplication	No	COG0561	Predicted hydrolases of the HAD superfamily	R	General function prediction only
fam2_qk	VC_0135	No duplication	No	COG2267	Lysophospholipase	I	Lipid transport and metabolism
fam2_qd	VC_0136	No duplication	No	COG1280	Putative threonine efflux protein	E	Amino acid transport and metabolism

fam2_pt	VC_0137	No duplication	No	COG3434	Predicted signal transduction protein containing EAL and modified HD-GYP domains	T	Signal transduction mechanisms
fam2_pm	VC_0139	No duplication	No	COG0783	DNA-binding ferritin-like protein (oxidative damage protectant)	P	Inorganic ion transport and metabolism
fam2_pe	VC_0142	No duplication	No	No hit			
fam2_of	VC_0144	No duplication	No	COG3714	Predicted membrane protein	S	Function unknown
fam2_nh	VC_0147	No duplication	No	COG0552	Signal recognition particle GTPase	U	Intracellular trafficking, secretion, and vesicular transport
fam2_mv	VC_0148	No duplication	No	COG2884	Predicted ATPase involved in cell division	D	Cell cycle control, cell division, chromosome partitioning
fam4_np	VC_0161	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam4_no	VC_0162	No duplication	No	COG0059	Ketol-acid reductoisomerase	EH	Multiple classes
fam4_nn	VC_0163	No duplication	HGT	COG2960	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ai	VC_0164	Duplication	HGT	COG0841	Cation/multidrug efflux pump	V	Defense mechanisms
fam4_nm	VC_0165	No duplication	No	COG0845	Membrane-fusion protein	M	Cell wall/membrane/envelope biogenesis
fam4_nl	VC_0170	No duplication	No	COG1123	ATPase components of various ABC-type transport systems, contain duplicated ATPase	R	General function prediction only
fam1_qb	VC_0171	Duplication	No	COG0747	ABC-type dipeptide transport system, periplasmic component	E	Amino acid transport and metabolism
fam4_nk	VC_0172	No duplication	No	COG0601	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	EP	Multiple classes
fam4_nj	VC_0173	No duplication	No	COG1173	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	EP	Multiple classes
fam4_ni	VC_0174	No duplication	HGT	No hit			
fam4_nh	VC_0186	No duplication	No	COG1249	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	C	Energy production and conversion
fam4_ng	VC_0187	No duplication	No	COG2961	Protein involved in catabolism of external DNA	R	General function prediction only
fam4_nf	VC_0188	No duplication	No	COG0339	Zn-dependent oligopeptidases	E	Amino acid transport and metabolism
fam4_ne	VC_0190	No duplication	HGT	COG0210	Superfamily I DNA and RNA helicases	L	Replication, recombination and repair
fam4_nd	VC_0191	No duplication	HGT	COG1280	Putative threonine efflux protein	E	Amino acid transport and metabolism
fam4_nc	VC_0192	No duplication	HGT	COG4977	Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	K	Transcription
fam4_nb	VC_0195	No duplication	HGT	COG2962	Predicted permeases	R	General function prediction only
fam4_na	VC_0196	No duplication	HGT	COG0514	Superfamily II DNA helicase	L	Replication, recombination and repair
fam4_mz	VC_0197	No duplication	HGT	No hit			
fam1_ax	VC_0199	Duplication	HGT	COG1132	ABC-type multidrug transport system, ATPase and permease components	V	Defense mechanisms
fam4_mo	VC_0204	No duplication	HGT	COG1737	Transcriptional regulators	K	Transcription
fam4_mb	VC_0207*	No duplication	HGT	COG1263	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	G	Carbohydrate transport and metabolism
fam4_mb	VC_0207*	No duplication	HGT	COG1264	Phosphotransferase system IIB components	G	Carbohydrate transport and metabolism
fam4_lw	VC_0208	No duplication	No	COG2862	Predicted membrane protein	S	Function unknown
fam4_lg	VC_0209	No duplication	No	COG1561	Uncharacterized stress-induced protein	S	Function unknown
fam4_ky	VC_0211	No duplication	HGT	COG0461	Orotate phosphoribosyltransferase	F	Nucleotide transport and metabolism
fam4_kp	VC_0212	No duplication	HGT	COG1560	Lauroyl/myristoyl acyltransferase	M	Cell wall/membrane/envelope biogenesis
fam4_kh	VC_0213	No duplication	HGT	COG1560	Lauroyl/myristoyl acyltransferase	M	Cell wall/membrane/envelope biogenesis
fam1_dw	VC_0214	Duplication	HGT	COG1309	Transcriptional regulator	K	Transcription
fam4_jx	VC_0215	No duplication	HGT	COG0452	Phosphopantothenoylcysteine synthetase/decarboxylase	H	Coenzyme transport and metabolism
fam4_iv	VC_0219	No duplication	No	COG0267	Ribosomal protein L33	J	Translation, ribosomal structure and biogenesis
fam4_is	VC_0220	No duplication	No	No hit			
fam4_ik	VC_0221	No duplication	No	COG0266	Formamidopyrimidine-DNA glycosylase	L	Replication, recombination and repair
fam4_hs	VC_0224	No duplication	HGT	COG0463	Glycosyltransferases involved in cell wall biogenesis	M	Cell wall/membrane/envelope biogenesis
fam1_pp	VC_0225	Duplication	HGT	COG0859	ADP-heptose:LPS heptosyltransferase	M	Cell wall/membrane/envelope biogenesis
fam4_fr	VC_0233	No duplication	HGT	COG1519	3-deoxy-D-manno-octulosonic-acid transferase	M	Cell wall/membrane/envelope biogenesis
fam4_fg	VC_0236	No duplication	HGT	COG0859	ADP-heptose:LPS heptosyltransferase	M	Cell wall/membrane/envelope biogenesis
fam4_dt	VC_0240	No duplication	HGT	COG0451	Nucleoside-diphosphate-sugar epimerases	MG	Multiple classes
fam1_ia	VC_0264	Duplication	HGT	COG1236	Predicted exonuclease of the beta-lactamase fold involved in RNA processing	J	Translation, ribosomal structure and biogenesis
fam3_st	VC_0265	No duplication	HGT	COG0668	Small-conductance mechanosensitive channel	M	Cell wall/membrane/envelope biogenesis
fam3_sg	VC_0267	No duplication	HGT	COG0111	Phosphoglycerate dehydrogenase and related dehydrogenases	HE	Multiple classes
fam1_os	VC_0268	Duplication	HGT	COG3068	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ri	VC_0271	No duplication	No	COG1253	Hemolysins and related proteins containing CBS domains	R	General function prediction only
fam3_qy	VC_0273	No duplication	No	COG0776	Bacterial nucleoid DNA-binding protein	L	Replication, recombination and repair
fam3_qn	VC_0274	No duplication	No	No hit			
fam3_qh	VC_0275	No duplication	HGT	COG0151	Phosphoribosylamine-glycine ligase	F	Nucleotide transport and metabolism
fam3_pz	VC_0276	No duplication	HGT	COG0138	AICAR transformylase/IMP cyclohydrolase PurH (only IMP cyclohydrolase domain in Aful)	F	Nucleotide transport and metabolism
fam3_pj	VC_0277	No duplication	HGT	COG0789	Predicted transcriptional regulators	K	Transcription
fam3_pd	VC_0278	No duplication	HGT	COG3710	DNA-binding winged-HTH domains	K	Transcription
fam3_ox	VC_0280	No duplication	HGT	COG0531	Amino acid transporters	E	Amino acid transport and metabolism
fam3_os	VC_0281	No duplication	HGT	COG1982	Arginine/lysine/ornithine decarboxylases	E	Amino acid transport and metabolism
fam1_oe	VC_0282	Duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam3_mg	VC_0293	No duplication	HGT	COG2264	Ribosomal protein L11 methylase	J	Translation, ribosomal structure and biogenesis
fam1_nv	VC_0297	Duplication	No	COG0757	3-dehydroquinate dehydratase II	E	Amino acid transport and metabolism
fam1_nt	VC_0299	Duplication	No	COG0847	DNA polymerase III, epsilon subunit and related 3'-5' exonucleases	L	Replication, recombination and repair
fam3_ku	VC_0300	No duplication	No	COG2905	Predicted signal-transduction protein containing cAMP-binding and CBS domains	T	Signal transduction mechanisms
fam3_kj	VC_0302	No duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam1_hm	VC_0303*	Duplication	No	COG0591	Na+/proline symporter	ER	Multiple classes
fam1_hm	VC_0303*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_hm	VC_0303*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam3_js	VC_0304	No duplication	No	COG0248	Exopolyphosphatase	FP	Multiple classes
fam1_hl	VC_0305	Duplication	No	COG0513	Superfamily II DNA and RNA helicases	LKJ	Multiple classes

fam3_jh	VC_0306	No duplication	No	COG3118	Thioredoxin domain-containing protein	O	Posttranslational modification, protein turnover, chaperones
fam3_jg	VC_0307	No duplication	No	COG1158	Transcription termination factor	K	Transcription
fam3_jd	VC_0308	No duplication	No	No hit			
fam1_np	VC_0309	Duplication	No	COG0043	3-polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases	H	Coenzyme transport and metabolism
fam1_qz	VC_0318	Duplication	No	COG0812	UDP-N-acetylmuramate dehydrogenase	M	Cell wall/membrane/envelope biogenesis
fam1_qw	VC_0328	Duplication	No	COG0085	DNA-directed RNA polymerase, beta subunit/140 kD subunit	K	Transcription
fam1_qv	VC_0329	Duplication	No	COG0086	DNA-directed RNA polymerase, beta' subunit/160 kD subunit	K	Transcription
fam5_bj	VC_0330	No duplication	No	COG3160	Regulator of sigma D	K	Transcription
fam5_bi	VC_0331	No duplication	No	COG2816	NTP pyrophosphohydrolases containing a Zn-finger, probably nucleic-acid-binding	L	Replication, recombination and repair
fam1_jc	VC_0332	Duplication	No	COG0407	Uroporphyrinogen-III decarboxylase	H	Coenzyme transport and metabolism
fam5_bh	VC_0333	No duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam1_qu	VC_0334	Duplication	No	No hit			
fam5_bg	VC_0335	No duplication	No	COG4942	Membrane-bound metallopeptidase	D	Cell cycle control, cell division, chromosome partitioning
fam5_bf	VC_0336	No duplication	No	COG0696	Phosphoglyceromutase	G	Carbohydrate transport and metabolism
fam1_at	VC_0338	Duplication	HGT	COG0471	Di- and tricarboxylate transporters	P	Inorganic ion transport and metabolism
fam5_at	VC_0339	No duplication	HGT	COG0688	Phosphatidylserine decarboxylase	I	Lipid transport and metabolism
fam1_qn	VC_0344*	Duplication	No	COG0860	N-acetylmuramoyl-L-alanine amidase	M	Cell wall/membrane/envelope biogenesis
fam1_qn	VC_0344*	Duplication	No	COG1388	FOG: LysM repeat	M	Cell wall/membrane/envelope biogenesis
fam1_qk	VC_0348	Duplication	No	COG2262	GTPases	R	General function prediction only
fam4_rl	VC_0349	No duplication	No	COG0330	Membrane protease subunits, stomatin/prohibitin homologs	O	Posttranslational modification, protein turnover, chaperones
fam4_rj	VC_0351	No duplication	No	COG3242	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_rd	VC_0352	No duplication	No	COG2900	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_qu	VC_0353	No duplication	No	COG2319	FOG: WD40 repeat	R	General function prediction only
fam1_av	VC_0361	Duplication	HGT	COG0480	Translation elongation factors (GTPases)	J	Translation, ribosomal structure and biogenesis
fam2_ix	VC_0371	No duplication	No	COG0305	Replicative DNA helicase	L	Replication, recombination and repair
fam2_ku	VC_0372	No duplication	No	COG0787	Alanine racemase	M	Cell wall/membrane/envelope biogenesis
fam4_bw	VC_0374	No duplication	No	COG0166	Glucose-6-phosphate isomerase	G	Carbohydrate transport and metabolism
fam5_ql	VC_0376	No duplication	No	COG2220	Predicted Zn-dependent hydrolases of the beta-lactamase fold	R	General function prediction only
fam4_gi	VC_0377	No duplication	No	COG1406	Predicted inhibitor of MCP methylation, homolog of CheC	N	Cell motility
fam4_ko	VC_0379	No duplication	No	COG0042	tRNA-dihydrouridine synthase	J	Translation, ribosomal structure and biogenesis
fam1_ip	VC_0381	Duplication	No	No hit			
fam4_ns	VC_0383	No duplication	HGT	No hit			
fam1_iq	VC_0384	Duplication	No	COG0369	Sulfite reductase, alpha subunit (flavoprotein)	P	Inorganic ion transport and metabolism
fam4_oc	VC_0385	No duplication	HGT	COG0155	Sulfite reductase, beta subunit (hemoprotein)	P	Inorganic ion transport and metabolism
fam4_gh	VC_0389	No duplication	No	COG0025	NhaP-type Na+/H+ and K+/H+ antiporters	P	Inorganic ion transport and metabolism
fam1_pn	VC_0390*	Duplication	No	COG1410	Methionine synthase I, cobalamin-binding domain	E	Amino acid transport and metabolism
fam1_pn	VC_0390*	Duplication	No	COG0646	Methionine synthase I (cobalamin-dependent), methyltransferase domain	E	Amino acid transport and metabolism
fam4_gj	VC_0391	No duplication	HGT	COG0527	Aspartokinases	E	Amino acid transport and metabolism
fam1_if	VC_0392	Duplication	No	COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	E	Amino acid transport and metabolism
fam4_gk	VC_0393	No duplication	No	COG3307	Lipid A core - O-antigen ligase and related enzymes	M	Cell wall/membrane/envelope biogenesis
fam4_gl	VC_0394	No duplication	No	COG0178	Excinuclease ATPase subunit	L	Replication, recombination and repair
fam4_gm	VC_0395	No duplication	No	COG1210	UDP-glucose pyrophosphorylase	M	Cell wall/membrane/envelope biogenesis
fam4_gn	VC_0396	No duplication	No	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes
fam4_go	VC_0397	No duplication	No	COG0629	Single-stranded DNA-binding protein	L	Replication, recombination and repair
fam4_gp	VC_0398*	No duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam4_gp	VC_0398*	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam4_gq	VC_0399	No duplication	No	COG3166	Tfp pilus assembly protein PilN	NU	Multiple classes
fam4_gr	VC_0400	No duplication	No	COG3167	Tfp pilus assembly protein PilO	NU	Multiple classes
fam4_gs	VC_0401	No duplication	No	No hit			
fam4_gt	VC_0402	No duplication	No	COG1450	Type II secretory pathway, component PulD	NU	Multiple classes
fam4_gu	VC_0403	No duplication	No	COG3267	Type II secretory pathway, component ExeA (predicted ATPase)	U	Intracellular trafficking, secretion, and vesicular transport
fam4_gv	VC_0404	No duplication	No	COG5010	Flp pilus assembly protein TadD, contains TPR repeats	U	Intracellular trafficking, secretion, and vesicular transport
fam4_gw	VC_0405	No duplication	No	COG2804	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	NU	Multiple classes
fam1_cp	VC_0406	Duplication	No	COG1459	Type II secretory pathway, component PulF	NU	Multiple classes
fam4_gx	VC_0408	No duplication	No	COG2165	Type II secretory pathway, pseudopilin PulG	NU	Multiple classes
fam4_gy	VC_0410	No duplication	No	COG2165	Type II secretory pathway, pseudopilin PulG	NU	Multiple classes
fam4_gz	VC_0411	No duplication	No	COG2165	Type II secretory pathway, pseudopilin PulG	NU	Multiple classes
fam4_ha	VC_0412	No duplication	No	COG4966	Tfp pilus assembly protein PilW	NU	Multiple classes
fam1_cq	VC_0414	Duplication	HGT	No hit			
fam4_hb	VC_0415	No duplication	No	COG1077	Actin-like ATPase involved in cell morphogenesis	D	Cell cycle control, cell division, chromosome partitioning
fam4_hc	VC_0416	No duplication	No	COG1792	Cell shape-determining protein	M	Cell wall/membrane/envelope biogenesis
fam4_hd	VC_0417	No duplication	No	COG2891	Cell shape-determining protein	M	Cell wall/membrane/envelope biogenesis
fam4_he	VC_0419	No duplication	No	COG1530	Ribonucleases G and E	J	Translation, ribosomal structure and biogenesis
fam1_po	VC_0420	Duplication	No	COG3164	Predicted membrane protein	S	Function unknown
fam4_hf	VC_0421	No duplication	No	COG0388	Predicted amidohydrolase	R	General function prediction only
fam4_hg	VC_0422	No duplication	No	COG0312	Predicted Zn-dependent proteases and their inactivated homologs	R	General function prediction only
fam4_hh	VC_0423	No duplication	No	COG2235	Arginine deiminase	E	Amino acid transport and metabolism
fam4_hi	VC_0424	No duplication	No	COG3076	Uncharacterized protein conserved in bacteria	S	Function unknown

fam4_hj	VC_0425	No duplication	No	No hit			
fam4_hk	VC_0426	No duplication	No	COG3205	Predicted membrane protein	S	Function unknown
fam4_hl	VC_0428	No duplication	No	COG0589	Universal stress protein UspA and related nucleotide-binding proteins	T	Signal transduction mechanisms
fam1_ci	VC_0429	Duplication	No	COG4666	TRAP-type uncharacterized transport system, fused permease components	R	General function prediction only
fam4_hm	VC_0430	No duplication	No	COG2358	TRAP-type uncharacterized transport system, periplasmic component	R	General function prediction only
fam4_hn	VC_0431	No duplication	No	COG1438	Arginine repressor	K	Transcription
fam4_ho	VC_0432	No duplication	No	COG0039	Malate/lactate dehydrogenases	C	Energy production and conversion
fam4_hp	VC_0433	No duplication	No	COG0531	Amino acid transporters	E	Amino acid transport and metabolism
fam4_hq	VC_0434	No duplication	No	COG0142	Geranylgeranyl pyrophosphate synthase	H	Coenzyme transport and metabolism
fam4_hr	VC_0437	No duplication	No	COG0536	Predicted GTPase	R	General function prediction only
fam4_ht	VC_0451	No duplication	HGT	COG2924	Fe-S cluster protector protein	CO	Multiple classes
fam1_ig	VC_0452	Duplication	HGT	COG1194	A/G-specific DNA glycosylase	L	Replication, recombination and repair
fam4_hu	VC_0453	No duplication	HGT	COG0220	Predicted S-adenosylmethionine-dependent methyltransferase	R	General function prediction only
fam1_pq	VC_0454	Duplication	No	COG2066	Glutaminase	E	Amino acid transport and metabolism
fam4_hv	VC_0455	No duplication	HGT	COG0635	Coproporphyrinogen III oxidase and related Fe-S oxidoreductases	H	Coenzyme transport and metabolism
fam4_hw	VC_0456	No duplication	HGT	COG0127	Xanthosine triphosphate pyrophosphatase	F	Nucleotide transport and metabolism
fam4_hx	VC_0457	No duplication	No	No hit			
fam4_hy	VC_0458	No duplication	No	COG1872	Uncharacterized conserved protein	S	Function unknown
fam4_hz	VC_0459	No duplication	No	COG0762	Predicted integral membrane protein	S	Function unknown
fam4_ia	VC_0460	No duplication	HGT	COG0345	Pyrroline-5-carboxylate reductase	E	Amino acid transport and metabolism
fam4_ib	VC_0461	No duplication	No	COG0325	Predicted enzyme with a TIM-barrel fold	R	General function prediction only
fam4_ic	VC_0462	No duplication	No	COG2805	Tfp pilus assembly protein, pilus retraction ATPase PilT	NU	Multiple classes
fam4_id	VC_0463	No duplication	No	COG5008	Tfp pilus assembly protein, ATPase PilU	NU	Multiple classes
fam4_ie	VC_0464	No duplication	HGT	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes
fam1_du	VC_0465	Duplication	HGT	COG0162	Tyrosyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam4_if	VC_0466	No duplication	HGT	COG0816	Predicted endonuclease involved in recombination (possible Holliday junction resolvase in Mycoplasmas and B. subtilis)	L	Replication, recombination and repair
fam1_pr	VC_0467	Duplication	HGT	COG1678	Putative transcriptional regulator	K	Transcription
fam4_ig	VC_0468	No duplication	HGT	COG0189	Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase)	HJ	Multiple classes
fam4_ih	VC_0469	No duplication	HGT	COG1385	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_ii	VC_0470	No duplication	HGT	COG2356	Endonuclease I	L	Replication, recombination and repair
fam4_ij	VC_0471	No duplication	HGT	COG3091	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_il	VC_0476	No duplication	HGT	COG0057	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	G	Carbohydrate transport and metabolism
fam1_ps	VC_0477	Duplication	No	COG0126	3-phosphoglycerate kinase	G	Carbohydrate transport and metabolism
fam4_im	VC_0478	No duplication	No	COG0191	Fructose/tagatose bisphosphate aldolase	G	Carbohydrate transport and metabolism
fam4_in	VC_0480	No duplication	No	COG0668	Small-conductance mechanosensitive channel	M	Cell wall/membrane/envelope biogenesis
fam4_io	VC_0481	No duplication	No	COG1279	Lysine efflux permease	R	General function prediction only
fam4_ip	VC_0482	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam4_iq	VC_0483	No duplication	No	COG2968	Uncharacterized conserved protein	S	Function unknown
fam4_ir	VC_0484	No duplication	No	No hit			
fam4_it	VC_0485	No duplication	No	COG0469	Pyruvate kinase	G	Carbohydrate transport and metabolism
fam1_ih	VC_0486	Duplication	No	COG1349	Transcriptional regulators of sugar metabolism	KG	Multiple classes
fam4_iu	VC_0487	No duplication	No	COG0449	Glucosamine 6-phosphate synthetase, contains amidotransferase and phosphosugar isomerase domains	M	Cell wall/membrane/envelope biogenesis
fam1_as	VC_0488	Duplication	HGT	COG4663	TRAP-type mannitol/chloroaromatic compound transport system, periplasmic component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam1_an	VC_0489	Duplication	HGT	COG3176	Putative hemolysin	R	General function prediction only
fam1_ad	VC_0514	Duplication	HGT	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam4_iw	VC_0522	No duplication	No	COG0596	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	R	General function prediction only
fam4_ix	VC_0523	No duplication	No	COG0344	Predicted membrane protein	S	Function unknown
fam4_iy	VC_0524	Duplication	No	COG1539	Dihydroneopterin aldolase	H	Coenzyme transport and metabolism
fam4_iz	VC_0525	No duplication	No	COG0801	7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase	H	Coenzyme transport and metabolism
fam4_ja	VC_0526	No duplication	No	COG1968	Uncharacterized bacitracin resistance protein	V	Defense mechanisms
fam4_jb	VC_0528	No duplication	No	COG1211	4-diphosphocytidyl-2-methyl-D-erithritol synthase	I	Lipid transport and metabolism
fam4_jc	VC_0529	No duplication	No	COG0245	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	I	Lipid transport and metabolism
fam4_jd	VC_0530	No duplication	No	COG0585	Uncharacterized conserved protein	S	Function unknown
fam4_je	VC_0531	No duplication	No	COG0496	Predicted acid phosphatase	R	General function prediction only
fam4_jf	VC_0532	No duplication	No	COG2518	Protein-L-isoaspartate carboxylmethyltransferase	O	Posttranslational modification, protein turnover, chaperones
fam4_jg	VC_0533	No duplication	No	COG0739	Membrane proteins related to metalloendopeptidases	M	Cell wall/membrane/envelope biogenesis
fam4_jh	VC_0535	No duplication	No	COG0249	Mismatch repair ATPase (MutS family)	L	Replication, recombination and repair
fam4_ji	VC_0537	No duplication	HGT	COG0031	Cysteine synthase	E	Amino acid transport and metabolism
fam4_jj	VC_0538	No duplication	HGT	COG4150	ABC-type sulfate transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam4_jk	VC_0539	No duplication	HGT	COG0555	ABC-type sulfate transport system, permease component	O	Posttranslational modification, protein turnover, chaperones
fam4_jl	VC_0540	No duplication	HGT	COG4208	ABC-type sulfate transport system, permease component	P	Inorganic ion transport and metabolism
fam4_jm	VC_0541	No duplication	HGT	COG1118	ABC-type sulfate/molybdate transport systems, ATPase component	P	Inorganic ion transport and metabolism
fam4_jn	VC_0542	No duplication	HGT	COG1546	Uncharacterized protein (competence- and mitomycin-induced)	R	General function prediction only
fam4_jo	VC_0543	No duplication	HGT	COG0468	RecA/RadA recombinase	L	Replication, recombination and repair
fam4_jp	VC_0544	No duplication	HGT	COG2137	Uncharacterized protein conserved in bacteria	R	General function prediction only
fam1_ii	VC_0545	Duplication	HGT	COG0013	Alanyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam1_ao	VC_0550*	Duplication	HGT	COG5016	Pyruvate/oxaloacetate carboxyltransferase	C	Energy production and conversion
fam1_ao	VC_0550*	Duplication	HGT	COG0511	Biotin carboxyl carrier protein	I	Lipid transport and metabolism

fam4_jq	VC_0551	No duplication	No	COG1883	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, beta subunit	C	Energy production and conversion
fam4_jr	VC_0552	No duplication	No	COG0604	NADPH:quinone reductase and related Zn-dependent oxidoreductases	CR	Multiple classes
fam4_js	VC_0553	No duplication	HGT	COG1238	Predicted membrane protein	S	Function unknown
fam1_dv	VC_0554*	Duplication	No	COG0612	Predicted Zn-dependent peptidases	R	General function prediction only
fam1_dv	VC_0554*	Duplication	No	COG1025	Secreted/periplasmic Zn-dependent peptidases, insulinase-like	O	Posttranslational modification, protein turnover, chaperones
fam4_jt	VC_0556	No duplication	No	COG2918	Gamma-glutamylcysteine synthetase	H	Coenzyme transport and metabolism
fam4_ju	VC_0557	No duplication	No	COG1854	LuxS protein involved in autoinducer AI2 synthesis	T	Signal transduction mechanisms
fam4_jv	VC_0558	No duplication	No	COG4536	Putative Mg2+ and Co2+ transporter CorB	P	Inorganic ion transport and metabolism
fam4_jw	VC_0559	No duplication	No	COG4137	ABC-type uncharacterized transport system, permease component	R	General function prediction only
fam4_jy	VC_0585	No duplication	No	COG0634	Hypoxanthine-guanine phosphoribosyltransferase	F	Nucleotide transport and metabolism
fam4_jz	VC_0586	No duplication	No	COG0288	Carbonic anhydrase	P	Inorganic ion transport and metabolism
fam4_ka	VC_0587	No duplication	No	COG0659	Sulfate permease and related transporters (MFS superfamily)	P	Inorganic ion transport and metabolism
fam4_kb	VC_0589	No duplication	No	COG1131	ABC-type multidrug transport system, ATPase component	V	Defense mechanisms
fam4_kc	VC_0590	No duplication	No	COG1682	ABC-type polysaccharide/polyol phosphate export systems, permease component	GM	Multiple classes
fam4_kd	VC_0591	No duplication	HGT	COG0414	Panthothenate synthetase	H	Coenzyme transport and metabolism
fam4_ke	VC_0593	No duplication	No	COG0801	7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase	H	Coenzyme transport and metabolism
fam4_kf	VC_0597	No duplication	No	COG1489	DNA-binding protein, stimulates sugar fermentation	R	General function prediction only
fam5_sf	VC_0600	No duplication	No	No hit			
fam1_ik	VC_0601	Duplication	No	COG1643	HrpA-like helicases	L	Replication, recombination and repair
fam4_kg	VC_0602	No duplication	No	COG0744	Membrane carboxypeptidase (penicillin-binding protein)	M	Cell wall/membrane/envelope biogenesis
fam1_il	VC_0603*	Duplication	No	COG0729	Outer membrane protein	M	Cell wall/membrane/envelope biogenesis
fam1_il	VC_0603*	Duplication	No	COG4667	Predicted esterase of the alpha-beta hydrolase superfamily	R	General function prediction only
fam1_pv	VC_0604	Duplication	No	COG1049	Aconitase B	C	Energy production and conversion
fam4_kj	VC_0608	No duplication	No	COG1840	ABC-type Fe3+ transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam4_kl	VC_0611	No duplication	HGT	COG1109	Phosphomannomutase	G	Carbohydrate transport and metabolism
fam1_pw	VC_0612	Duplication	HGT	COG3459	Cellobiose phosphorylase	G	Carbohydrate transport and metabolism
fam4_km	VC_0613	No duplication	No	COG3525	N-acetyl-beta-hexosaminidase	G	Carbohydrate transport and metabolism
fam1_aq	VC_0614	Duplication	HGT	COG2971	Predicted N-acetylglucosamine kinase	G	Carbohydrate transport and metabolism
fam1_aq	VC_0615	Duplication	HGT	No hit			
fam4_kn	VC_0616	No duplication	No	COG0444	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	EP	Multiple classes
fam4_kq	VC_0617	No duplication	No	COG0444	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	EP	Multiple classes
fam4_kr	VC_0618	No duplication	No	COG1173	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	EP	Multiple classes
fam4_ks	VC_0619	No duplication	No	COG0601	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	EP	Multiple classes
fam4_kt	VC_0620	No duplication	No	COG0747	ABC-type dipeptide transport system, periplasmic component	E	Amino acid transport and metabolism
fam1_im	VC_0622*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_im	VC_0622*	Duplication	No	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam4_ku	VC_0623	No duplication	No	COG2813	16S RNA G1207 methylase RsmC	J	Translation, ribosomal structure and biogenesis
fam4_kv	VC_0624	No duplication	No	COG0628	Predicted permease	R	General function prediction only
fam4_kw	VC_0626	No duplication	No	COG0001	Glutamate-1-semialdehyde aminotransferase	H	Coenzyme transport and metabolism
fam4_kx	VC_0627	No duplication	No	COG0316	Uncharacterized conserved protein	S	Function unknown
fam1_in	VC_0628	Duplication	No	COG0845	Membrane-fusion protein	M	Cell wall/membrane/envelope biogenesis
fam4_la	VC_0641	No duplication	No	COG0779	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_lb	VC_0642	No duplication	No	COG0195	Transcription elongation factor	K	Transcription
fam4_lc	VC_0644	No duplication	No	COG0858	Ribosome-binding factor A	J	Translation, ribosomal structure and biogenesis
fam4_ld	VC_0645	No duplication	No	COG0130	Pseudouridine synthase	J	Translation, ribosomal structure and biogenesis
fam4_le	VC_0646	No duplication	No	COG0184	Ribosomal protein S15P/S13E	J	Translation, ribosomal structure and biogenesis
fam4_lf	VC_0647	No duplication	No	COG1185	Polyribonucleotide nucleotidyltransferase (polynucleotide phosphorylase)	J	Translation, ribosomal structure and biogenesis
fam4_lh	VC_0648	No duplication	No	COG4785	Lipoprotein NlpI, contains TPR repeats	R	General function prediction only
fam4_li	VC_0649	No duplication	No	COG1321	Mn-dependent transcriptional regulator	K	Transcription
fam1_px	VC_0650	Duplication	No	COG0534	Na+-driven multidrug efflux pump	V	Defense mechanisms
fam4_lj	VC_0651	No duplication	No	COG0826	Collagenase and related proteases	O	Posttranslational modification, protein turnover, chaperones
fam4_lk	VC_0652	No duplication	No	COG0826	Collagenase and related proteases	O	Posttranslational modification, protein turnover, chaperones
fam4_ll	VC_0653*	No duplication	No	COG5001	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	T	Signal transduction mechanisms
fam4_ll	VC_0653*	No duplication	No	COG2202	FOG: PAS/PAC domain	T	Signal transduction mechanisms
fam4_lm	VC_0655	No duplication	No	COG3153	Predicted acetyltransferase	R	General function prediction only
fam4_ln	VC_0656	No duplication	No	COG3050	DNA polymerase III, psi subunit	L	Replication, recombination and repair
fam4_lo	VC_0657	No duplication	No	COG0456	Acetyltransferases	R	General function prediction only
fam4_lp	VC_0658	No duplication	No	COG5001	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	T	Signal transduction mechanisms
fam4_lq	VC_0659	No duplication	No	COG4108	Peptide chain release factor RF-3	J	Translation, ribosomal structure and biogenesis
fam4_lr	VC_0660	No duplication	No	COG0513	Superfamily II DNA and RNA helicases	LKJ	Multiple classes
fam4_ls	VC_0661	No duplication	No	COG4123	Predicted O-methyltransferase	R	General function prediction only
fam4_lt	VC_0662	No duplication	No	COG1114	Branched-chain amino acid permeases	E	Amino acid transport and metabolism
fam4_lu	VC_0664	No duplication	No	COG1190	Lysyl-tRNA synthetase (class II)	J	Translation, ribosomal structure and biogenesis
fam4_lv	VC_0665	No duplication	No	COG1221	Transcriptional regulators containing an AAA-type ATPase domain and a DNA-binding domain	KT	Multiple classes
fam4_lx	VC_0667	No duplication	No	COG0667	Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	C	Energy production and conversion
fam4_ly	VC_0668	No duplication	No	COG3066	DNA mismatch repair protein	L	Replication, recombination and repair
fam4_lz	VC_0672	No duplication	No	COG3605	Signal transduction protein containing GAF and PtsI domains	T	Signal transduction mechanisms
fam4_ma	VC_0673	No duplication	No	COG0730	Predicted permeases	R	General function prediction only

fam1_py	VC_0674	Duplication	No	COG0682	Prolipoprotein diacylglyceryltransferase	M	Cell wall/membrane/envelope biogenesis
fam4_mc	VC_0675	No duplication	No	COG0207	Thymidylate synthase	F	Nucleotide transport and metabolism
fam4_md	VC_0676	No duplication	No	COG1283	Na+/phosphate symporter	P	Inorganic ion transport and metabolism
fam4_mf	VC_0681	No duplication	No	COG0196	FAD synthase	H	Coenzyme transport and metabolism
fam4_mg	VC_0682	No duplication	No	COG0060	Isoleucyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam4_mh	VC_0683	No duplication	HGT	COG0597	Lipoprotein signal peptidase	MU	Multiple classes
fam4_mi	VC_0687	No duplication	HGT	COG1966	Carbon starvation protein, predicted membrane protein	T	Signal transduction mechanisms
fam4_mj	VC_0688	No duplication	No	No hit			
fam4_mk	VC_0690	No duplication	HGT	COG2103	Predicted sugar phosphate isomerase	R	General function prediction only
fam4_ml	VC_0691	No duplication	HGT	COG2377	Predicted molecular chaperone distantly related to HSP70-fold metalloproteases	O	Posttranslational modification, protein turnover, chaperones
fam4_mm	VC_0692	No duplication	HGT	COG1472	Beta-glucosidase-related glycosidases	G	Carbohydrate transport and metabolism
fam4_mn	VC_0694	No duplication	HGT	COG3275	Putative regulator of cell autolysis	T	Signal transduction mechanisms
fam4_mp	VC_0695	No duplication	HGT	COG0722	3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase	E	Amino acid transport and metabolism
fam4_mq	VC_0696*	No duplication	HGT	COG0287	Prephenate dehydrogenase	E	Amino acid transport and metabolism
fam4_mq	VC_0696*	No duplication	HGT	COG1605	Chorismate mutase	E	Amino acid transport and metabolism
fam4_mr	VC_0697	No duplication	HGT	No hit			
fam1_io	VC_0698	Duplication	HGT	COG0488	ATPase components of ABC transporters with duplicated ATPase domains	R	General function prediction only
fam1_dx	VC_0700	Duplication	No	COG0741	Soluble lytic murein transglycosylase and related regulatory proteins (some contain LysM/invasin domains)	M	Cell wall/membrane/envelope biogenesis
fam4_ms	VC_0701	No duplication	No	COG2973	Trp operon repressor	K	Transcription
fam4_mt	VC_0703*	No duplication	No	COG5001	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	T	Signal transduction mechanisms
fam4_mt	VC_0703*	No duplication	No	COG3850	Signal transduction histidine kinase, nitrate/nitrite-specific	T	Signal transduction mechanisms
fam4_mu	VC_0704	No duplication	No	COG0687	Spermidine/putrescine-binding periplasmic protein	E	Amino acid transport and metabolism
fam4_mv	VC_0705*	No duplication	No	COG0077	Prephenate dehydratase	E	Amino acid transport and metabolism
fam4_mv	VC_0705*	No duplication	No	COG1605	Chorismate mutase	E	Amino acid transport and metabolism
fam4_mw	VC_0706	No duplication	No	COG1544	Ribosome-associated protein Y (PSrp-1)	J	Translation, ribosomal structure and biogenesis
fam4_mx	VC_0709	No duplication	No	COG0564	Pseudouridylate synthases, 23S RNA-specific	J	Translation, ribosomal structure and biogenesis
fam4_my	VC_0710	No duplication	HGT	COG1496	Uncharacterized conserved protein	S	Function unknown
fam4_ef	VC_0714	No duplication	HGT	No hit			
fam4_ee	VC_0715	No duplication	HGT	COG0778	Nitroreductase	C	Energy production and conversion
fam4_ed	VC_0716	No duplication	HGT	COG2768	Uncharacterized Fe-S center protein	R	General function prediction only
fam4_ec	VC_0717	No duplication	HGT	COG0826	Collagenase and related proteases	O	Posttranslational modification, protein turnover, chaperones
fam4_eb	VC_0718	No duplication	HGT	COG2974	DNA recombination-dependent growth factor C	L	Replication, recombination and repair
fam4_ea	VC_0719	No duplication	HGT	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam4_dz	VC_0720	No duplication	HGT	COG5002	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam4_dy	VC_0721	No duplication	No	COG0226	ABC-type phosphate transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam1_ph	VC_0723	Duplication	No	COG0855	Polyphosphate kinase	P	Inorganic ion transport and metabolism
fam4_dw	VC_0724	No duplication	No	COG4590	ABC-type uncharacterized transport system, permease component	R	General function prediction only
fam4_dv	VC_0725*	No duplication	No	COG4985	ABC-type phosphate transport system, auxiliary component	P	Inorganic ion transport and metabolism
fam4_dv	VC_0725*	No duplication	No	COG0581	ABC-type phosphate transport system, permease component	P	Inorganic ion transport and metabolism
fam4_du	VC_0726	No duplication	No	COG1117	ABC-type phosphate transport system, ATPase component	P	Inorganic ion transport and metabolism
fam4_ds	VC_0727	No duplication	No	COG0704	Phosphate uptake regulator	P	Inorganic ion transport and metabolism
fam4_dr	VC_0728	No duplication	No	COG2326	Uncharacterized conserved protein	S	Function unknown
fam4_dq	VC_0730	No duplication	No	COG3142	Uncharacterized protein involved in copper resistance	P	Inorganic ion transport and metabolism
fam4_dp	VC_0731	No duplication	No	COG0450	Peroxiredoxin	O	Posttranslational modification, protein turnover, chaperones
fam4_do	VC_0732	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam4_dn	VC_0734	No duplication	HGT	COG2225	Malate synthase	C	Energy production and conversion
fam4_dm	VC_0736	No duplication	No	COG2224	Isocitrate lyase	C	Energy production and conversion
fam4_dl	VC_0737	No duplication	No	COG0517	FOG: CBS domain	R	General function prediction only
fam4_dk	VC_0739	No duplication	HGT	COG0809	S-adenosylmethionine:tRNA-ribosyltransferase-isomerase (queuine synthetase)	J	Translation, ribosomal structure and biogenesis
fam4_dj	VC_0742	No duplication	HGT	COG1862	Preprotein translocase subunit YajC	U	Intracellular trafficking, secretion, and vesicular transport
fam4_di	VC_0745	No duplication	No	COG0483	Archaeal fructose-1,6-bisphosphatase and related enzymes of inositol monophosphatase family	G	Carbohydrate transport and metabolism
fam4_dh	VC_0746	No duplication	No	COG0565	rRNA methylase	J	Translation, ribosomal structure and biogenesis
fam4_dg	VC_0747	No duplication	No	COG1959	Predicted transcriptional regulator	K	Transcription
fam4_df	VC_0748	No duplication	No	COG1104	Cysteine sulfinate desulfinase/cysteine desulfurase and related enzymes	E	Amino acid transport and metabolism
fam4_de	VC_0750	No duplication	HGT	COG0316	Uncharacterized conserved protein	S	Function unknown
fam4_dd	VC_0751	No duplication	No	COG1076	DnaJ-domain-containing proteins 1	O	Posttranslational modification, protein turnover, chaperones
fam4_dc	VC_0752	No duplication	No	COG0443	Molecular chaperone	O	Posttranslational modification, protein turnover, chaperones
fam4_db	VC_0753	No duplication	HGT	COG0633	Ferredoxin	C	Energy production and conversion
fam4_da	VC_0755	No duplication	HGT	COG0260	Leucyl aminopeptidase	E	Amino acid transport and metabolism
fam4_cz	VC_0756	No duplication	No	COG0105	Nucleoside diphosphate kinase	F	Nucleotide transport and metabolism
fam4_cy	VC_0757	No duplication	No	COG0820	Predicted Fe-S-cluster redox enzyme	R	General function prediction only
fam4_cx	VC_0759	No duplication	No	COG0821	Enzyme involved in the deoxyxylulose pathway of isoprenoid biosynthesis	I	Lipid transport and metabolism
fam4_cw	VC_0760	No duplication	No	COG0124	Histidyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam4_cv	VC_0761	No duplication	No	COG2976	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_cu	VC_0762	No duplication	No	COG1520	FOG: WD40-like repeat	S	Function unknown
fam4_ct	VC_0763	No duplication	No	COG1160	Predicted GTPases	R	General function prediction only
fam4_cs	VC_0764	No duplication	No	COG2872	Predicted metal-dependent hydrolases related to alanyl-tRNA synthetase HxxxH domain	R	General function prediction only
fam1_pf	VC_0765	Duplication	No	No hit			

fam4_cr	VC_0767*	No duplication	No	COG0516	IMP dehydrogenase/GMP reductase	F	Nucleotide transport and metabolism
fam4_cr	VC_0767*	No duplication	No	COG0517	FOG: CBS domain	R	General function prediction only
fam4_cq	VC_0769*	No duplication	No	COG3979	Uncharacterized protein contain chitin-binding domain type 3	R	General function prediction only
fam4_cq	VC_0769*	No duplication	No	COG0810	Periplasmic protein TonB, links inner and outer membranes	M	Cell wall/membrane/envelope biogenesis
fam4_cp	VC_0770	No duplication	No	COG1584	Predicted membrane protein	S	Function unknown
fam4_co	VC_0784	No duplication	No	COG1115	Na+/alanine symporter	E	Amino acid transport and metabolism
fam4_cn	VC_0786	No duplication	HGT	COG0665	Glycine/D-amino acid oxidases (deaminating)	E	Amino acid transport and metabolism
fam4_cm	VC_0788	No duplication	No	COG3805	Aromatic ring-cleaving dioxygenase	Q	Secondary metabolites biosynthesis, transport and catabolism
fam4_cl	VC_0789	No duplication	No	No hit			
fam4_ck	VC_0791	No duplication	HGT	COG3290	Signal transduction histidine kinase regulating citrate/malate metabolism	T	Signal transduction mechanisms
fam4_cj	VC_0792	No duplication	HGT	COG1883	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, beta subunit	C	Energy production and conversion
fam4_ci	VC_0794	No duplication	HGT	COG3630	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, gamma subunit	C	Energy production and conversion
fam1_id	VC_0796	Duplication	HGT	COG3053	Citrate lyase synthetase	C	Energy production and conversion
fam4_ch	VC_0797	No duplication	HGT	COG3052	Citrate lyase, gamma subunit	C	Energy production and conversion
fam1_am	VC_0798	Duplication	HGT	COG2301	Citrate lyase beta subunit	G	Carbohydrate transport and metabolism
fam1_am	VC_0799	Duplication	HGT	COG3051	Citrate lyase, alpha subunit	C	Energy production and conversion
fam4_cg	VC_0800	No duplication	HGT	COG3697	Phosphoribosyl-dephospho-CoA transferase (holo-ACP synthetase)	HI	Multiple classes
fam4_cf	VC_0801	No duplication	HGT	COG1767	Triphosphoribosyl-dephospho-CoA synthetase	H	Coenzyme transport and metabolism
fam4_ce	VC_0803	No duplication	HGT	COG0566	rRNA methylases	J	Translation, ribosomal structure and biogenesis
fam4_cd	VC_0804	No duplication	HGT	COG1146	Ferredoxin	C	Energy production and conversion
fam1_pe	VC_0806	Duplication	HGT	COG2509	Uncharacterized FAD-dependent dehydrogenases	R	General function prediction only
fam4_cc	VC_0807	No duplication	HGT	No hit			
fam4_cb	VC_0848	No duplication	HGT	COG0691	tmRNA-binding protein	O	Posttranslational modification, protein turnover, chaperones
fam4_ca	VC_0849	No duplication	HGT	COG2867	Oligoketide cyclase/lipid transport protein	I	Lipid transport and metabolism
fam4_bz	VC_0850	No duplication	HGT	COG2914	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_by	VC_0851	No duplication	HGT	COG2913	Small protein A (tmRNA-binding)	J	Translation, ribosomal structure and biogenesis
fam1_ic	VC_0852	Duplication	No	COG0497	ATPase involved in DNA repair	L	Replication, recombination and repair
fam4_bx	VC_0853	No duplication	No	COG0061	Predicted sugar kinase	G	Carbohydrate transport and metabolism
fam4_bv	VC_0856	No duplication	No	COG0484	DnaJ-class molecular chaperone with C-terminal Zn finger domain	O	Posttranslational modification, protein turnover, chaperones
fam4_bu	VC_0857	No duplication	HGT	COG4968	Tfp pilus assembly protein PilE	NU	Multiple classes
fam4_bt	VC_0859	No duplication	HGT	COG4795	Type II secretory pathway, component PulJ	U	Intracellular trafficking, secretion, and vesicular transport
fam1_dt	VC_0860	Duplication	HGT	COG3156	Type II secretory pathway, component PulK	U	Intracellular trafficking, secretion, and vesicular transport
fam4_bs	VC_0861	No duplication	No	COG4967	Tfp pilus assembly protein PilV	NU	Multiple classes
fam4_br	VC_0863	No duplication	No	COG2978	Putative p-aminobenzoyl-glutamate transporter	H	Coenzyme transport and metabolism
fam4_bq	VC_0864	No duplication	No	COG0590	Cytosine/adenosine deaminases	FJ	Multiple classes
fam4_bp	VC_0865	No duplication	No	No hit			
fam1_pd	VC_0866	Duplication	No	COG4623	Predicted soluble lytic transglycosylase fused to an ABC-type amino acid-binding protein	M	Cell wall/membrane/envelope biogenesis
fam1_pc	VC_0869*	Duplication	No	COG0046	Phosphoribosylformylglycinamidine (FGAM) synthase, synthetase domain	F	Nucleotide transport and metabolism
fam1_pc	VC_0869*	Duplication	No	COG0047	Phosphoribosylformylglycinamidine (FGAM) synthase, glutamine amidotransferase domain	F	Nucleotide transport and metabolism
fam4_bo	VC_0871	No duplication	HGT	No hit			
fam4_bn	VC_0872	No duplication	HGT	No hit			
fam4_bm	VC_0875	No duplication	No	COG0442	Prolyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam4_bl	VC_0876	No duplication	No	COG1720	Uncharacterized conserved protein	S	Function unknown
fam4_bk	VC_0881	No duplication	No	COG1671	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_bj	VC_0882	No duplication	No	No hit			
fam4_bi	VC_0884	No duplication	No	COG1670	Acetyltransferases, including N-acetylases of ribosomal proteins	J	Translation, ribosomal structure and biogenesis
fam4_bh	VC_0886	No duplication	No	No hit			
fam4_bg	VC_0887	No duplication	No	COG3098	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_bf	VC_0888	No duplication	No	COG0564	Pseudouridylate synthases, 23S RNA-specific	J	Translation, ribosomal structure and biogenesis
fam1_pb	VC_0889	Duplication	No	COG1154	Deoxyxylulose-5-phosphate synthase	HI	Multiple classes
fam4_be	VC_0890	No duplication	No	COG0142	Geranylgeranyl pyrophosphate synthase	H	Coenzyme transport and metabolism
fam4_bd	VC_0891	No duplication	No	COG1722	Exonuclease VII small subunit	L	Replication, recombination and repair
fam4_bc	VC_0892	No duplication	No	COG1291	Flagellar motor component	N	Cell motility
fam1_ds	VC_0893	Duplication	No	COG1360	Flagellar motor protein	N	Cell motility
fam1_ib	VC_0894*	Duplication	No	COG0301	Thiamine biosynthesis ATP pyrophosphatase	H	Coenzyme transport and metabolism
fam1_ib	VC_0894*	Duplication	No	COG0607	Rhodanese-related sulfurtransferase	P	Inorganic ion transport and metabolism
fam4_bb	VC_0896	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam4_ba	VC_0897	No duplication	No	COG2363	Uncharacterized small membrane protein	S	Function unknown
fam4_az	VC_0898	No duplication	No	COG0258	5'-3' exonuclease (including N-terminal domain of PolI)	L	Replication, recombination and repair
fam4_ay	VC_0899	No duplication	No	COG1611	Predicted Rossmann fold nucleotide-binding protein	R	General function prediction only
fam4_ax	VC_0900	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_pa	VC_0901	Duplication	No	No hit			
fam4_aw	VC_0902*	No duplication	No	COG2904	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_aw	VC_0902*	No duplication	No	COG0780	Enzyme related to GTP cyclohydrolase I	R	General function prediction only
fam1_dr	VC_0903	Duplication	HGT	No hit			
fam4_av	VC_0904	No duplication	No	COG5595	Zn-ribbon-containing, possibly nucleic-acid-binding protein	R	General function prediction only
fam4_au	VC_0905	No duplication	HGT	COG1464	ABC-type metal ion transport system, periplasmic component/surface antigen	P	Inorganic ion transport and metabolism
fam4_at	VC_0906	No duplication	HGT	COG2011	ABC-type metal ion transport system, permease component	P	Inorganic ion transport and metabolism

fam4_as	VC_0907	No duplication	HGT	COG1135	ABC-type metal ion transport system, ATPase component	P	Inorganic ion transport and metabolism
fam4_ar	VC_0908	No duplication	No	COG0241	Histidinol phosphatase and related phosphatases	E	Amino acid transport and metabolism
fam1_bs	VC_0911	Duplication	HGT	COG0366	Glycosidases	G	Carbohydrate transport and metabolism
fam4_aq	VC_0913	No duplication	No	COG0845	Membrane-fusion protein	M	Cell wall/membrane/envelope biogenesis
fam1_ai	VC_0914	Duplication	HGT	COG0841	Cation/multidrug efflux pump	V	Defense mechanisms
fam4_ap	VC_0916	No duplication	No	COG0394	Protein-tyrosine-phosphatase	T	Signal transduction mechanisms
fam1_bp	VC_0917	Duplication	HGT	COG0381	UDP-N-acetylglucosamine 2-epimerase	M	Cell wall/membrane/envelope biogenesis
fam1_ct	VC_0918	Duplication	HGT	COG0677	UDP-N-acetyl-D-mannosaminuronate dehydrogenase	M	Cell wall/membrane/envelope biogenesis
fam1_co	VC_0925	Duplication	HGT	COG0438	Glycosyltransferase	M	Cell wall/membrane/envelope biogenesis
fam4_ao	VC_0927	No duplication	No	COG1922	Teichoic acid biosynthesis proteins	M	Cell wall/membrane/envelope biogenesis
fam1_ac	VC_0930	Duplication	HGT	No hit			
fam4_an	VC_0931	No duplication	No	COG3307	Lipid A core - O-antigen ligase and related enzymes	M	Cell wall/membrane/envelope biogenesis
fam1_oz	VC_0932	Duplication	No	No hit			
fam1_oz	VC_0933	Duplication	No	No hit			
fam1_kf	VC_0934	Duplication	No	COG2148	Sugar transferases involved in lipopolysaccharide synthesis	M	Cell wall/membrane/envelope biogenesis
fam4_am	VC_0935	No duplication	No	COG5338	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_al	VC_0936	No duplication	No	COG1596	Periplasmic protein involved in polysaccharide export	M	Cell wall/membrane/envelope biogenesis
fam1_oy	VC_0937*	Duplication	No	COG3206	Uncharacterized protein involved in exopolysaccharide biosynthesis	M	Cell wall/membrane/envelope biogenesis
fam1_oy	VC_0937*	Duplication	No	COG0489	ATPases involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam4_ak	VC_0938	No duplication	No	No hit			
fam4_aj	VC_0939	No duplication	No	No hit			
fam4_ai	VC_0941	No duplication	No	COG0112	Glycine/serine hydroxymethyltransferase	E	Amino acid transport and metabolism
fam4_ah	VC_0944	No duplication	No	COG0321	Lipoate-protein ligase B	H	Coenzyme transport and metabolism
fam4_ag	VC_0950	No duplication	No	COG0768	Cell division protein FtsI/penicillin-binding protein 2	M	Cell wall/membrane/envelope biogenesis
fam4_af	VC_0951	No duplication	No	COG1576	Uncharacterized conserved protein	S	Function unknown
fam4_ae	VC_0952	No duplication	No	COG0799	Uncharacterized homolog of plant Iojap protein	S	Function unknown
fam4_ac	VC_0954	No duplication	HGT	COG2980	Rare lipoprotein B	M	Cell wall/membrane/envelope biogenesis
fam1_ox	VC_0956	Duplication	HGT	COG0495	Leucyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam4_ab	VC_0957	No duplication	HGT	No hit			
fam4_aa	VC_0958	No duplication	HGT	COG0815	Apolipoprotein N-acyltransferase	M	Cell wall/membrane/envelope biogenesis
fam3_tf	VC_0959	No duplication	No	COG4535	Putative Mg2+ and Co2+ transporter CorC	P	Inorganic ion transport and metabolism
fam3_te	VC_0960	No duplication	HGT	COG0319	Predicted metal-dependent hydrolase	R	General function prediction only
fam3_td	VC_0961	No duplication	HGT	COG1702	Phosphate starvation-inducible protein PhoH, predicted ATPase	T	Signal transduction mechanisms
fam3_tc	VC_0970	No duplication	No	COG3115	Cell division protein	D	Cell cycle control, cell division, chromosome partitioning
fam1_ow	VC_0971	Duplication	No	COG0272	NAD-dependent DNA ligase (contains BRCT domain type II)	L	Replication, recombination and repair
fam3_ta	VC_0974	No duplication	No	COG0789	Predicted transcriptional regulators	K	Transcription
fam3_sz	VC_0976	No duplication	No	COG0330	Membrane protease subunits, stomatin/prohibitin homologs	O	Posttranslational modification, protein turnover, chaperones
fam3_sy	VC_0977	No duplication	No	COG3118	Thioredoxin domain-containing protein	O	Posttranslational modification, protein turnover, chaperones
fam3_sx	VC_0979	No duplication	No	COG0300	Short-chain dehydrogenases of various substrate specificities	R	General function prediction only
fam3_sw	VC_0981	No duplication	No	COG0622	Predicted phosphoesterase	R	General function prediction only
fam3_sv	VC_0982	No duplication	No	COG3526	Uncharacterized protein conserved in bacteria	O	Posttranslational modification, protein turnover, chaperones
fam3_su	VC_0983	No duplication	No	No hit			
fam3_ss	VC_0984	No duplication	No	COG3710	DNA-binding winged-HTH domains	K	Transcription
fam1_hz	VC_0985	Duplication	No	COG0326	Molecular chaperone, HSP90 family	O	Posttranslational modification, protein turnover, chaperones
fam3_sr	VC_0986	No duplication	No	COG0563	Adenylate kinase and related kinases	F	Nucleotide transport and metabolism
fam3_sq	VC_0987	No duplication	No	COG0276	Protoheme ferro-lyase (ferrochelatase)	H	Coenzyme transport and metabolism
fam3_sp	VC_0988	No duplication	No	COG3104	Dipeptide/tripeptide permease	E	Amino acid transport and metabolism
fam3_so	VC_0990	No duplication	No	COG0250	Transcription antiterminator	K	Transcription
fam1_ou	VC_0991	Duplication	No	COG0367	Asparagine synthase (glutamine-hydrolyzing)	E	Amino acid transport and metabolism
fam1_ot	VC_0992*	Duplication	No	COG4651	Kef-type K+ transport system, predicted NAD-binding component	P	Inorganic ion transport and metabolism
fam1_ot	VC_0992*	Duplication	No	COG1226	Kef-type K+ transport systems, predicted NAD-binding component	P	Inorganic ion transport and metabolism
fam3_sn	VC_0993	No duplication	No	COG1940	Transcriptional regulator/sugar kinase	KG	Multiple classes
fam3_sm	VC_0994	No duplication	No	COG1820	N-acetylglucosamine-6-phosphate deacetylase	G	Carbohydrate transport and metabolism
fam3_sl	VC_0995*	No duplication	No	COG1263	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	G	Carbohydrate transport and metabolism
fam3_sl	VC_0995*	No duplication	No	COG1264	Phosphotransferase system IIB components	G	Carbohydrate transport and metabolism
fam1_hy	VC_0997	Duplication	No	COG0008	Glutamyl- and glutaminyl-tRNA synthetases	J	Translation, ribosomal structure and biogenesis
fam1_ch	VC_0998*	Duplication	No	COG3170	Tfp pilus assembly protein FimV	NU	Multiple classes
fam1_ch	VC_0998*	Duplication	No	COG5271	AAA ATPase containing von Willebrand factor type A (vWA) domain	R	General function prediction only
fam3_sk	VC_0999	No duplication	No	COG0101	Pseudouridylate synthase	J	Translation, ribosomal structure and biogenesis
fam1_hx	VC_1000	Duplication	HGT	COG0777	Acetyl-CoA carboxylase beta subunit	I	Lipid transport and metabolism
fam3_sj	VC_1001	No duplication	No	COG0285	Folylpolyglutamate synthase	H	Coenzyme transport and metabolism
fam3_si	VC_1002	No duplication	No	COG3147	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_sh	VC_1003	No duplication	No	COG1286	Uncharacterized membrane protein, required for colicin V production	R	General function prediction only
fam3_sf	VC_1004	No duplication	No	COG0034	Glutamine phosphoribosylpyrophosphate amidotransferase	F	Nucleotide transport and metabolism
fam3_se	VC_1005	No duplication	No	COG2056	Predicted permease	R	General function prediction only
fam3_sd	VC_1006	No duplication	No	COG0847	DNA polymerase III, epsilon subunit and related 3'-5' exonucleases	L	Replication, recombination and repair
fam3_sc	VC_1016	No duplication	No	COG2878	Predicted NADH:ubiquinone oxidoreductase, subunit RnfB	C	Energy production and conversion
fam3_sb	VC_1017	No duplication	No	COG4657	Predicted NADH:ubiquinone oxidoreductase, subunit RnfA	C	Energy production and conversion

fam1_oq	VC_1021	Duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam3_sa	VC_1022	No duplication	HGT	No hit			
fam3_rz	VC_1023	No duplication	No	COG0391	Uncharacterized conserved protein	S	Function unknown
fam3_ry	VC_1024	No duplication	No	COG2896	Molybdenum cofactor biosynthesis enzyme	H	Coenzyme transport and metabolism
fam3_rx	VC_1025	No duplication	HGT	COG0521	Molybdopterin biosynthesis enzymes	H	Coenzyme transport and metabolism
fam3_rw	VC_1026	No duplication	No	COG0315	Molybdenum cofactor biosynthesis enzyme	H	Coenzyme transport and metabolism
fam3_rv	VC_1027	No duplication	No	COG1977	Molybdopterin converting factor, small subunit	H	Coenzyme transport and metabolism
fam3_ru	VC_1028	No duplication	No	COG0314	Molybdopterin converting factor, large subunit	H	Coenzyme transport and metabolism
fam3_rt	VC_1029	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_jr	VC_1031	Duplication	No	COG2524	Predicted transcriptional regulator, contains C-terminal CBS domains	K	Transcription
fam1_op	VC_1033	Duplication	No	COG2217	Cation transport ATPase	P	Inorganic ion transport and metabolism
fam3_rs	VC_1034	No duplication	No	COG2820	Uridine phosphorylase	F	Nucleotide transport and metabolism
fam3_rr	VC_1035	No duplication	No	No hit			
fam3_rq	VC_1036*	No duplication	No	COG0143	Methionyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam3_rq	VC_1036*	No duplication	No	COG0073	EMAP domain	R	General function prediction only
fam3_rp	VC_1037	No duplication	HGT	COG0489	ATPases involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam3_ro	VC_1038	No duplication	HGT	COG0572	Uridine kinase	F	Nucleotide transport and metabolism
fam1_hw	VC_1039	Duplication	HGT	COG2982	Uncharacterized protein involved in outer membrane biogenesis	M	Cell wall/membrane/envelope biogenesis
fam3_rn	VC_1040	No duplication	HGT	COG2109	ATP:corrinoid adenosyltransferase	H	Coenzyme transport and metabolism
fam3_rm	VC_1041	No duplication	HGT	COG0394	Protein-tyrosine-phosphatase	T	Signal transduction mechanisms
fam3_rl	VC_1042	No duplication	HGT	COG2067	Long-chain fatty acid transport protein	I	Lipid transport and metabolism
fam3_rk	VC_1043	No duplication	HGT	COG2067	Long-chain fatty acid transport protein	I	Lipid transport and metabolism
fam3_rj	VC_1044	No duplication	No	No hit			
fam3_rh	VC_1045	No duplication	No	COG1595	DNA-directed RNA polymerase specialized sigma subunit, sigma24 homolog	K	Transcription
fam3_rg	VC_1046	No duplication	No	COG0183	Acetyl-CoA acetyltransferase	I	Lipid transport and metabolism
fam3_rf	VC_1047*	No duplication	No	COG1250	3-hydroxyacyl-CoA dehydrogenase	I	Lipid transport and metabolism
fam3_rf	VC_1047*	No duplication	No	COG1024	Enoyl-CoA hydratase/carnithine racemase	I	Lipid transport and metabolism
fam3_re	VC_1048	No duplication	HGT	COG0778	Nitroreductase	C	Energy production and conversion
fam3_rd	VC_1049	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam3_rc	VC_1050*	No duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam3_rc	VC_1050*	No duplication	No	COG2208	Serine phosphatase RsbU, regulator of sigma subunit	TK	Multiple classes
fam3_rb	VC_1051	No duplication	No	No hit			
fam3_ra	VC_1052	No duplication	No	COG2832	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_qz	VC_1053	No duplication	No	COG0503	Adenine/guanine phosphoribosyltransferases and related PRPP-binding proteins	F	Nucleotide transport and metabolism
fam1_oo	VC_1062	Duplication	No	COG1522	Transcriptional regulators	K	Transcription
fam3_qw	VC_1063	No duplication	No	COG1946	Acyl-CoA thioesterase	I	Lipid transport and metabolism
fam3_qv	VC_1064	No duplication	No	COG3126	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_qu	VC_1065	No duplication	No	COG3695	Predicted methylated DNA-protein cysteine methyltransferase	L	Replication, recombination and repair
fam3_qt	VC_1066	No duplication	No	No hit			
fam3_qs	VC_1067*	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam3_qs	VC_1067*	No duplication	No	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam3_qr	VC_1068	No duplication	HGT	COG0640	Predicted transcriptional regulators	K	Transcription
fam3_qq	VC_1069	No duplication	HGT	COG0057	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	G	Carbohydrate transport and metabolism
fam3_qp	VC_1070	No duplication	HGT	COG2453	Predicted protein-tyrosine phosphatase	T	Signal transduction mechanisms
fam3_qo	VC_1071	No duplication	HGT	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam1_af	VC_1073	Duplication	HGT	COG3325	Chitinase	G	Carbohydrate transport and metabolism
fam1_hv	VC_1074	Duplication	No	No hit			
fam3_qm	VC_1075	No duplication	HGT	COG1881	Phospholipid-binding protein	R	General function prediction only
fam3_ql	VC_1076	No duplication	No	COG4977	Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	K	Transcription
fam3_qk	VC_1077	No duplication	No	No hit			
fam3_qj	VC_1079	No duplication	No	COG0009	Putative translation factor (SUA5)	J	Translation, ribosomal structure and biogenesis
fam1_kk	VC_1080	Duplication	No	COG2198	FOG: HPt domain	T	Signal transduction mechanisms
fam3_qi	VC_1081*	No duplication	No	COG1639	Predicted signal transduction protein	T	Signal transduction mechanisms
fam3_qi	VC_1081*	No duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam1_on	VC_1082	Duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam3_qg	VC_1083	No duplication	No	COG1776	Chemotaxis protein CheC, inhibitor of MCP methylation	NT	Multiple classes
fam3_qf	VC_1084	No duplication	No	COG4191	Signal transduction histidine kinase regulating C4-dicarboxylate transport system	T	Signal transduction mechanisms
fam3_qe	VC_1085	No duplication	No	COG4191	Signal transduction histidine kinase regulating C4-dicarboxylate transport system	T	Signal transduction mechanisms
fam3_qd	VC_1086*	No duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam3_qd	VC_1086*	No duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam3_qc	VC_1087	No duplication	No	COG3437	Response regulator containing a CheY-like receiver domain and an HD-GYP domain	KT	Multiple classes
fam3_qb	VC_1089	No duplication	No	COG3221	ABC-type phosphate/phosphonate transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam3_qa	VC_1091	No duplication	No	COG4166	ABC-type oligopeptide transport system, periplasmic component	E	Amino acid transport and metabolism
fam1_om	VC_1092	Duplication	No	COG0601	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	EP	Multiple classes
fam3_py	VC_1093	No duplication	No	COG1173	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components	EP	Multiple classes
fam3_px	VC_1094	No duplication	No	COG0444	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	EP	Multiple classes
fam3_pw	VC_1095	No duplication	No	COG4608	ABC-type oligopeptide transport system, ATPase component	E	Amino acid transport and metabolism
fam3_pv	VC_1096	No duplication	No	COG0435	Predicted glutathione S-transferase	O	Posttranslational modification, protein turnover, chaperones

fam3_pu	VC_1097*	No duplication	No	COG0280	Phosphotransacetylase	C	Energy production and conversion
fam3_pu	VC_1097*	No duplication	No	COG0857	BioD-like N-terminal domain of phosphotransacetylase	R	General function prediction only
fam3_pt	VC_1098	No duplication	No	COG0282	Acetate kinase	C	Energy production and conversion
fam3_ps	VC_1099	No duplication	No	COG3092	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ol	VC_1100	Duplication	HGT	COG4628	Uncharacterized conserved protein	S	Function unknown
fam3_pr	VC_1101	No duplication	No	COG2984	ABC-type uncharacterized transport system, periplasmic component	R	General function prediction only
fam3_pq	VC_1102	No duplication	No	COG4120	ABC-type uncharacterized transport system, permease component	R	General function prediction only
fam3_pp	VC_1103	No duplication	No	COG1101	ABC-type uncharacterized transport system, ATPase component	R	General function prediction only
fam3_po	VC_1104	No duplication	HGT	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam3_pn	VC_1105	No duplication	HGT	COG0042	tRNA-dihydrouridine synthase	J	Translation, ribosomal structure and biogenesis
fam3_pm	VC_1106	No duplication	HGT	COG3318	Predicted metal-binding protein related to the C-terminal domain of SecA	R	General function prediction only
fam3_pl	VC_1107	No duplication	HGT	COG2834	Outer membrane lipoprotein-sorting protein	M	Cell wall/membrane/envelope biogenesis
fam3_pk	VC_1108	No duplication	HGT	COG2256	ATPase related to the helicase subunit of the Holliday junction resolvase	L	Replication, recombination and repair
fam1_ok	VC_1110	Duplication	HGT	COG0172	Seryl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam1_hu	VC_1111	Duplication	HGT	COG0161	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	H	Coenzyme transport and metabolism
fam1_oj	VC_1112	Duplication	HGT	COG0502	Biotin synthase and related enzymes	H	Coenzyme transport and metabolism
fam1_oi	VC_1115	Duplication	No	COG0132	Dethiobiotin synthetase	H	Coenzyme transport and metabolism
fam3_ph	VC_1118	No duplication	No	COG3631	Ketosteroid isomerase-related protein	R	General function prediction only
fam3_pg	VC_1119	No duplication	HGT	COG0300	Short-chain dehydrogenases of various substrate specificities	R	General function prediction only
fam3_pf	VC_1120	No duplication	HGT	COG2907	Predicted NAD/FAD-binding protein	R	General function prediction only
fam1_oh	VC_1122	Duplication	No	COG2230	Cyclopropane fatty acid synthase and related methyltransferases	M	Cell wall/membrane/envelope biogenesis
fam3_pe	VC_1123	No duplication	No	No hit			
fam3_pc	VC_1124	No duplication	No	No hit			
fam5_pd	VC_1125	No duplication	No	No hit			
fam1_ht	VC_1126	Duplication	No	COG0015	Adenylosuccinate lyase	F	Nucleotide transport and metabolism
fam3_pb	VC_1127	No duplication	No	COG2915	Uncharacterized protein involved in purine metabolism	R	General function prediction only
fam3_pa	VC_1128	No duplication	No	COG0482	Predicted tRNA(5-methylaminomethyl-2-thiouridylate) methyltransferase, contains the PP-loop ATPase domain	J	Translation, ribosomal structure and biogenesis
fam1_ar	VC_1129	Duplication	HGT	COG0524	Sugar kinases, ribokinase family	G	Carbohydrate transport and metabolism
fam3_oz	VC_1130	No duplication	No	COG2916	DNA-binding protein H-NS	R	General function prediction only
fam1_aj	VC_1131	Duplication	HGT	COG1757	Na+/H+ antiporter	C	Energy production and conversion
fam3_oy	VC_1132	No duplication	HGT	COG0040	ATP phosphoribosyltransferase	E	Amino acid transport and metabolism
fam1_cm	VC_1133	Duplication	HGT	COG0141	Histidinol dehydrogenase	E	Amino acid transport and metabolism
fam1_fm	VC_1134	Duplication	HGT	COG0079	Histidinol-phosphate/aromatic aminotransferase and cobyric acid decarboxylase	E	Amino acid transport and metabolism
fam1_fl	VC_1135*	Duplication	HGT	COG0131	Imidazoleglycerol-phosphate dehydratase	E	Amino acid transport and metabolism
fam1_fl	VC_1135*	Duplication	HGT	COG0241	Histidinol phosphatase and related phosphatases	E	Amino acid transport and metabolism
fam1_fk	VC_1136	Duplication	HGT	COG0118	Glutamine amidotransferase	E	Amino acid transport and metabolism
fam1_cl	VC_1137	Duplication	HGT	COG0106	Phosphoribosylformimino-5-aminoimidazole carboxamide ribonucleotide (ProFAR) isomerase	E	Amino acid transport and metabolism
fam1_fj	VC_1138	Duplication	HGT	COG0107	Imidazoleglycerol-phosphate synthase	E	Amino acid transport and metabolism
fam3_ow	VC_1139*	No duplication	HGT	COG0139	Phosphoribosyl-AMP cyclohydrolase	E	Amino acid transport and metabolism
fam3_ow	VC_1139*	No duplication	HGT	COG0140	Phosphoribosyl-ATP pyrophosphohydrolase	E	Amino acid transport and metabolism
fam1_og	VC_1140	Duplication	HGT	COG1187	16S rRNA uridine-516 pseudouridylate synthase and related pseudouridylate synthases	J	Translation, ribosomal structure and biogenesis
fam3_ov	VC_1141	No duplication	No	COG2838	Monomeric isocitrate dehydrogenase	C	Energy production and conversion
fam3_ou	VC_1142	No duplication	No	COG1278	Cold shock proteins	K	Transcription
fam3_ot	VC_1143	No duplication	No	COG2127	Uncharacterized conserved protein	S	Function unknown
fam1_of	VC_1144	Duplication	No	COG0542	ATPases with chaperone activity, ATP-binding subunit	O	Posttranslational modification, protein turnover, chaperones
fam3_or	VC_1145	No duplication	No	COG2985	Predicted permease	R	General function prediction only
fam3_oq	VC_1146	No duplication	No	COG0695	Glutaredoxin and related proteins	O	Posttranslational modification, protein turnover, chaperones
fam3_op	VC_1147	No duplication	No	COG1454	Alcohol dehydrogenase, class IV	C	Energy production and conversion
fam3_oo	VC_1148	No duplication	No	COG1737	Transcriptional regulators	K	Transcription
fam3_on	VC_1149	No duplication	No	COG0076	Glutamate decarboxylase and related PLP-dependent proteins	E	Amino acid transport and metabolism
fam3_om	VC_1150	No duplication	No	No hit			
fam3_ol	VC_1151	No duplication	No	COG2431	Predicted membrane protein	S	Function unknown
fam3_ok	VC_1152	No duplication	No	COG1639	Predicted signal transduction protein	T	Signal transduction mechanisms
fam3_oj	VC_1153	No duplication	No	COG3070	Regulator of competence-specific genes	K	Transcription
fam1_hs	VC_1156	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_hs	VC_1157	Duplication	No	No hit			
fam3_oi	VC_1158	No duplication	No	COG5339	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_hr	VC_1159	Duplication	No	COG1932	Phosphoserine aminotransferase	HE	Multiple classes
fam3_oh	VC_1169	No duplication	No	COG0159	Tryptophan synthase alpha chain	E	Amino acid transport and metabolism
fam3_og	VC_1170	No duplication	No	COG0133	Tryptophan synthase beta chain	E	Amino acid transport and metabolism
fam1_oc	VC_1171*	Duplication	HGT	COG0134	Indole-3-glycerol phosphate synthase	E	Amino acid transport and metabolism
fam1_oc	VC_1171*	Duplication	HGT	COG0135	Phosphoribosylanthranilate isomerase	E	Amino acid transport and metabolism
fam3_of	VC_1172	No duplication	HGT	COG0547	Anthranilate phosphoribosyltransferase	E	Amino acid transport and metabolism
fam3_oe	VC_1173	No duplication	HGT	COG0512	Anthranilate/para-aminobenzoate synthases component II	EH	Multiple classes
fam3_od	VC_1174	No duplication	HGT	COG0147	Anthranilate/para-aminobenzoate synthases component I	EH	Multiple classes
fam3_oc	VC_1177	No duplication	No	COG0613	Predicted metal-dependent phosphoesterases (PHP family)	R	General function prediction only
fam3_ob	VC_1178	No duplication	No	COG0009	Putative translation factor (SUA5)	J	Translation, ribosomal structure and biogenesis
fam3_oa	VC_1179	No duplication	No	COG1187	16S rRNA uridine-516 pseudouridylate synthase and related pseudouridylate synthases	J	Translation, ribosomal structure and biogenesis

fam3_nz	VC_1180	No duplication	No	COG4987	ABC-type transport system involved in cytochrome bd biosynthesis, fused ATPase and permease components	CO	Multiple classes
fam1_ob	VC_1181	Duplication	No	COG4988	ABC-type transport system involved in cytochrome bd biosynthesis, ATPase and permease components	CO	Multiple classes
fam3_ny	VC_1182	No duplication	No	COG0492	Thioredoxin reductase	O	Posttranslational modification, protein turnover, chaperones
fam3_nx	VC_1183	No duplication	No	COG4340	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_nw	VC_1184	No duplication	No	COG0520	Selenocysteine lyase	E	Amino acid transport and metabolism
fam3_nv	VC_1185	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam3_nu	VC_1186	No duplication	No	COG2949	Uncharacterized membrane protein	S	Function unknown
fam3_nt	VC_1187	No duplication	No	No hit			
fam3_ns	VC_1188	No duplication	No	COG0281	Malic enzyme	C	Energy production and conversion
fam3_nr	VC_1189	No duplication	No	No hit			
fam3_nq	VC_1190	No duplication	No	COG0152	Phosphoribosylaminoimidazolesuccinocarboxamide (SAICAR) synthase	F	Nucleotide transport and metabolism
fam3_np	VC_1191	No duplication	No	No hit			
fam3_no	VC_1194*	No duplication	No	COG0790	FOG: TPR repeat, SEL1 subfamily	R	General function prediction only
fam3_no	VC_1194*	No duplication	No	COG0484	DnaJ-class molecular chaperone with C-terminal Zn finger domain	O	Posttranslational modification, protein turnover, chaperones
fam3_nn	VC_1195	No duplication	No	COG2885	Outer membrane protein and related peptidoglycan-associated (lipo)proteins	M	Cell wall/membrane/envelope biogenesis
fam3_nm	VC_1196	No duplication	No	No hit			
fam3_nl	VC_1197	No duplication	No	No hit			
fam3_nk	VC_1198*	No duplication	No	COG0277	FAD/FMN-containing dehydrogenases	C	Energy production and conversion
fam3_nk	VC_1198*	No duplication	No	COG0247	Fe-S oxidoreductase	C	Energy production and conversion
fam1_oa	VC_1199	Duplication	HGT	COG2020	Putative protein-S-isoprenylcysteine methyltransferase	O	Posttranslational modification, protein turnover, chaperones
fam1_hq	VC_1203	Duplication	HGT	COG2987	Urocanate hydratase	E	Amino acid transport and metabolism
fam3_nj	VC_1205	No duplication	HGT	COG1228	Imidazolonepropionase and related amidohydrolases	Q	Secondary metabolites biosynthesis, transport and catabolism
fam3_ni	VC_1206	No duplication	HGT	COG2188	Transcriptional regulators	K	Transcription
fam3_nh	VC_1208	No duplication	HGT	COG3099	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ng	VC_1209	No duplication	HGT	COG0231	Translation elongation factor P (EF-P)/translation initiation factor 5A (eIF-5A)	J	Translation, ribosomal structure and biogenesis
fam3_nf	VC_1213	No duplication	HGT	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes
fam1_nz	VC_1214	Duplication	No	COG0322	Nuclease subunit of the excinuclease complex	L	Replication, recombination and repair
fam3_ne	VC_1229	No duplication	No	COG0679	Predicted permeases	R	General function prediction only
fam3_nd	VC_1231	No duplication	No	COG0295	Cytidine deaminase	F	Nucleotide transport and metabolism
fam3_nc	VC_1232	No duplication	No	COG1346	Putative effector of murein hydrolase	M	Cell wall/membrane/envelope biogenesis
fam3_nb	VC_1233	No duplication	No	COG1380	Putative effector of murein hydrolase LrgA	R	General function prediction only
fam3_na	VC_1234	No duplication	No	COG2925	Exonuclease I	L	Replication, recombination and repair
fam3_mz	VC_1235	No duplication	No	COG1823	Predicted Na+/dicarboxylate symporter	R	General function prediction only
fam3_my	VC_1237	No duplication	No	COG2038	NaMN:DMB phosphoribosyltransferase	H	Coenzyme transport and metabolism
fam3_mx	VC_1238	No duplication	No	COG0368	Cobalamin-5-phosphate synthase	H	Coenzyme transport and metabolism
fam3_mw	VC_1239	No duplication	No	COG2087	Adenosyl cobinamide kinase/adenosyl cobinamide phosphate guanylyltransferase	H	Coenzyme transport and metabolism
fam3_mv	VC_1240	No duplication	No	COG0406	Fructose-2,6-bisphosphatase	G	Carbohydrate transport and metabolism
fam3_mu	VC_1242	No duplication	No	COG2988	Succinylglutamate desuccinylase	E	Amino acid transport and metabolism
fam3_mt	VC_1244	No duplication	No	COG4139	ABC-type cobalamin transport system, permease component	H	Coenzyme transport and metabolism
fam3_ms	VC_1246	No duplication	No	No hit			
fam3_mr	VC_1247	No duplication	No	No hit			
fam3_mq	VC_1248	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam3_mp	VC_1249	No duplication	No	COG2716	Glycine cleavage system regulatory protein	E	Amino acid transport and metabolism
fam3_mo	VC_1250	No duplication	No	No hit			
fam1_hp	VC_1256	Duplication	No	COG0209	Ribonucleotide reductase, alpha subunit	F	Nucleotide transport and metabolism
fam3_mn	VC_1257	No duplication	No	COG2227	2-polyprenyl-3-methyl-5-hydroxy-6-metoxy-1,4-benzoquinol methylase	H	Coenzyme transport and metabolism
fam3_mm	VC_1259	No duplication	No	COG1054	Predicted sulfurtransferase	R	General function prediction only
fam3_ml	VC_1260	No duplication	No	COG4977	Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	K	Transcription
fam1_ny	VC_1261	Duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam3_mk	VC_1262	No duplication	No	No hit			
fam3_mj	VC_1263	No duplication	No	COG0807	GTP cyclohydrolase II	H	Coenzyme transport and metabolism
fam3_mh	VC_1269	No duplication	No	COG3108	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_as	VC_1273	Duplication	HGT	COG4663	TRAP-type mannitol/chloroaromatic compound transport system, periplasmic component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam3_mf	VC_1274	No duplication	No	COG4665	TRAP-type mannitol/chloroaromatic compound transport system, small permease component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam3_me	VC_1275	No duplication	No	COG4664	TRAP-type mannitol/chloroaromatic compound transport system, large permease component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam3_md	VC_1276	No duplication	No	COG4564	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam3_mc	VC_1277	No duplication	No	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes
fam3_mb	VC_1278	No duplication	No	COG1846	Transcriptional regulators	K	Transcription
fam3_ma	VC_1279	No duplication	No	COG1292	Choline-glycine betaine transporter	M	Cell wall/membrane/envelope biogenesis
fam1_nw	VC_1287	Duplication	No	COG2943	Membrane glycosyltransferase	M	Cell wall/membrane/envelope biogenesis
fam1_ek	VC_1288	Duplication	No	COG3131	Periplasmic glucans biosynthesis protein	P	Inorganic ion transport and metabolism
fam3_ly	VC_1289	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam3_lx	VC_1290	No duplication	No	COG0847	DNA polymerase III, epsilon subunit and related 3'-5' exonucleases	L	Replication, recombination and repair
fam3_lw	VC_1291	No duplication	No	COG2905	Predicted signal-transduction protein containing cAMP-binding and CBS domains	T	Signal transduction mechanisms
fam3_lv	VC_1293	No duplication	No	COG1448	Aspartate/tyrosine/aromatic aminotransferase	E	Amino acid transport and metabolism
fam3_lu	VC_1296	No duplication	No	COG0351	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	H	Coenzyme transport and metabolism
fam3_lt	VC_1297	No duplication	No	COG0017	Aspartyl/asparaginyl-tRNA synthetases	J	Translation, ribosomal structure and biogenesis
fam3_ls	VC_1298	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes

fam3_lr	VC_1300	No duplication	No	COG1760	L-serine deaminase	E	Amino acid transport and metabolism
fam3_lq	VC_1302	No duplication	No	COG0494	NTP pyrophosphohydrolases including oxidative damage repair enzymes	LR	Multiple classes
fam3_lp	VC_1303	No duplication	No	COG0147	Anthranilate/para-aminobenzoate synthases component I	EH	Multiple classes
fam3_lo	VC_1304*	No duplication	No	COG1951	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	C	Energy production and conversion
fam3_lo	VC_1304*	No duplication	No	COG1838	Tartrate dehydratase beta subunit/Fumarate hydratase class I, C-terminal domain	C	Energy production and conversion
fam3_ln	VC_1305	No duplication	No	No hit			
fam3_lm	VC_1306	No duplication	No	COG3106	Predicted ATPase	R	General function prediction only
fam3_ll	VC_1308	No duplication	No	COG3283	Transcriptional regulator of aromatic amino acids metabolism	KE	Multiple classes
fam3_lk	VC_1309	No duplication	No	COG1670	Acetyltransferases, including N-acetylases of ribosomal proteins	J	Translation, ribosomal structure and biogenesis
fam3_lj	VC_1310	No duplication	No	No hit			
fam3_li	VC_1311	No duplication	No	COG5006	Predicted permease, DMT superfamily	R	General function prediction only
fam3_lh	VC_1312	No duplication	No	COG0787	Alanine racemase	M	Cell wall/membrane/envelope biogenesis
fam3_lg	VC_1313	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_at	VC_1314	Duplication	HGT	COG0471	Di- and tricarboxylate transporters	P	Inorganic ion transport and metabolism
fam3_lf	VC_1315*	No duplication	No	COG4191	Signal transduction histidine kinase regulating C4-dicarboxylate transport system	T	Signal transduction mechanisms
fam3_lf	VC_1315*	No duplication	No	COG0664	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	T	Signal transduction mechanisms
fam3_le	VC_1316	No duplication	No	COG3437	Response regulator containing a CheY-like receiver domain and an HD-GYP domain	KT	Multiple classes
fam3_ld	VC_1317	No duplication	No	COG2990	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_lc	VC_1318	No duplication	No	COG3713	Outer membrane protein V	M	Cell wall/membrane/envelope biogenesis
fam1_nu	VC_1319	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam3_lb	VC_1320	No duplication	No	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam3_la	VC_1321	No duplication	No	COG0744	Membrane carboxypeptidase (penicillin-binding protein)	M	Cell wall/membrane/envelope biogenesis
fam3_kz	VC_1323	No duplication	No	COG1433	Uncharacterized conserved protein	S	Function unknown
fam3_ky	VC_1325	No duplication	No	COG1879	ABC-type sugar transport system, periplasmic component	G	Carbohydrate transport and metabolism
fam1_ho	VC_1327	Duplication	No	COG1129	ABC-type sugar transport system, ATPase component	G	Carbohydrate transport and metabolism
fam3_kx	VC_1328	No duplication	No	COG4211	ABC-type glucose/galactose transport system, permease component	G	Carbohydrate transport and metabolism
fam3_kw	VC_1332	No duplication	No	COG3333	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_kv	VC_1333	No duplication	No	No hit			
fam3_kt	VC_1334	No duplication	No	COG3181	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ks	VC_1335	No duplication	HGT	COG1802	Transcriptional regulators	K	Transcription
fam3_kr	VC_1336	No duplication	HGT	COG2513	PEP phosphonomutase and related enzymes	G	Carbohydrate transport and metabolism
fam3_kq	VC_1337	No duplication	HGT	COG0372	Citrate synthase	C	Energy production and conversion
fam1_hn	VC_1338	Duplication	HGT	COG1048	Aconitase A	C	Energy production and conversion
fam3_kp	VC_1339	No duplication	HGT	COG2828	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ko	VC_1341	No duplication	No	COG0456	Acetyltransferases	R	General function prediction only
fam3_kn	VC_1342	No duplication	No	COG0494	NTP pyrophosphohydrolases including oxidative damage repair enzymes	LR	Multiple classes
fam3_km	VC_1343	No duplication	No	COG2195	Di- and tripeptidases	E	Amino acid transport and metabolism
fam1_fd	VC_1344	Duplication	No	COG3185	4-hydroxyphenylpyruvate dioxygenase and related hemolysins	ER	Multiple classes
fam1_fe	VC_1345	Duplication	HGT	COG3508	Homogentisate 1,2-dioxygenase	Q	Secondary metabolites biosynthesis, transport and catabolism
fam1_ff	VC_1346	Duplication	No	COG0179	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	Q	Secondary metabolites biosynthesis, transport and catabolism
fam1_fg	VC_1347	Duplication	No	COG0625	Glutathione S-transferase	O	Posttranslational modification, protein turnover, chaperones
fam1_fh	VC_1348	Duplication	No	COG3437	Response regulator containing a CheY-like receiver domain and an HD-GYP domain	KT	Multiple classes
fam1_ce	VC_1349*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_ce	VC_1349*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam1_ce	VC_1349*	Duplication	No	COG2198	FOG: HPt domain	T	Signal transduction mechanisms
fam1_ce	VC_1349*	Duplication	No	COG4753	Response regulator containing CheY-like receiver domain and AraC-type DNA-binding domain	T	Signal transduction mechanisms
fam3_kl	VC_1350	No duplication	No	COG0678	Peroxiredoxin	O	Posttranslational modification, protein turnover, chaperones
fam1_dq	VC_1353*	Duplication	No	COG3292	Predicted periplasmic ligand-binding sensor domain	T	Signal transduction mechanisms
fam1_dq	VC_1353*	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam3_kk	VC_1354	No duplication	No	COG1092	Predicted SAM-dependent methyltransferases	R	General function prediction only
fam3_ki	VC_1355	No duplication	No	COG1254	Acylphosphatases	C	Energy production and conversion
fam3_kh	VC_1356	No duplication	No	COG2920	Dissimilatory sulfite reductase (desulfoviridin), gamma subunit	P	Inorganic ion transport and metabolism
fam3_kg	VC_1358	No duplication	No	COG0670	Integral membrane protein, interacts with FtsH	R	General function prediction only
fam3_kf	VC_1360	No duplication	No	COG0765	ABC-type amino acid transport system, permease component	E	Amino acid transport and metabolism
fam3_ke	VC_1361	No duplication	No	COG4597	ABC-type amino acid transport system, permease component	E	Amino acid transport and metabolism
fam3_kd	VC_1362	No duplication	No	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam3_kc	VC_1363*	No duplication	No	COG1648	Siroheme synthase (precorrin-2 oxidase/ferrochelatase domain)	H	Coenzyme transport and metabolism
fam3_kc	VC_1363*	No duplication	No	COG0007	Uroporphyrinogen-III methylase	H	Coenzyme transport and metabolism
fam3_kb	VC_1364	No duplication	No	COG0561	Predicted hydrolases of the HAD superfamily	R	General function prediction only
fam3_ka	VC_1365	No duplication	No	COG0602	Organic radical activating enzymes	O	Posttranslational modification, protein turnover, chaperones
fam3_jz	VC_1366	No duplication	No	COG0603	Predicted PP-loop superfamily ATPase	R	General function prediction only
fam1_ns	VC_1367*	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_ns	VC_1367*	Duplication	No	COG2976	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_jy	VC_1368	No duplication	No	No hit			
fam3_jx	VC_1369	No duplication	No	COG0715	ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components	P	Inorganic ion transport and metabolism
fam3_jw	VC_1370	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam3_jv	VC_1371	No duplication	No	No hit			
fam3_ju	VC_1372	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms

fam3_jt	VC_1373	No duplication	No	COG0443	Molecular chaperone	O	Posttranslational modification, protein turnover, chaperones
fam1_nr	VC_1374	Duplication	No	COG0443	Molecular chaperone	O	Posttranslational modification, protein turnover, chaperones
fam3_jr	VC_1375	No duplication	No	No hit			
fam3_jq	VC_1376	No duplication	No	COG3614	Predicted periplasmic ligand-binding sensor domain	T	Signal transduction mechanisms
fam3_jp	VC_1377	No duplication	No	COG5018	Inhibitor of the KinA pathway to sporulation, predicted exonuclease	R	General function prediction only
fam3_jo	VC_1379	No duplication	HGT	COG4103	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_jn	VC_1380	No duplication	HGT	No hit			
fam1_cb	VC_1382	Duplication	HGT	COG1643	HrpA-like helicases	L	Replication, recombination and repair
fam3_jm	VC_1384	No duplication	No	No hit			
fam3_jl	VC_1386	No duplication	No	COG0443	Molecular chaperone	O	Posttranslational modification, protein turnover, chaperones
fam3_jk	VC_1388	Duplication	No	COG0095	Lipoate-protein ligase A	H	Coenzyme transport and metabolism
fam3_jj	VC_1390	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam1_nq	VC_1391	Duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam3_ji	VC_1393	No duplication	No	COG2076	Membrane transporters of cations and cationic drugs	P	Inorganic ion transport and metabolism
fam3_jf	VC_1414	No duplication	No	COG2317	Zn-dependent carboxypeptidase	E	Amino acid transport and metabolism
fam1_ab	VC_1416	Duplication	HGT	COG3501	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_je	VC_1417	No duplication	HGT	No hit			
fam1_bk	VC_1420	Duplication	HGT	No hit			
fam3_jc	VC_1421	No duplication	HGT	COG1280	Putative threonine efflux protein	E	Amino acid transport and metabolism
fam3_jb	VC_1422	No duplication	HGT	COG1115	Na+/alanine symporter	E	Amino acid transport and metabolism
fam3_ja	VC_1423	No duplication	No	COG3264	Small-conductance mechanosensitive channel	M	Cell wall/membrane/envelope biogenesis
fam3_iz	VC_1424	No duplication	No	COG0687	Spermidine/putrescine-binding periplasmic protein	E	Amino acid transport and metabolism
fam3_iy	VC_1429	No duplication	No	No hit			
fam3_ix	VC_1430	No duplication	No	COG2992	Uncharacterized FlgJ-related protein	R	General function prediction only
fam3_iw	VC_1433	No duplication	No	COG0589	Universal stress protein UspA and related nucleotide-binding proteins	T	Signal transduction mechanisms
fam3_iv	VC_1434	No duplication	No	COG0664	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	T	Signal transduction mechanisms
fam1_kc	VC_1436	Duplication	No	COG3197	Uncharacterized protein, possibly involved in nitrogen fixation	P	Inorganic ion transport and metabolism
fam1_cv	VC_1437	Duplication	No	COG2217	Cation transport ATPase	P	Inorganic ion transport and metabolism
fam1_kb	VC_1439	Duplication	No	COG2010	Cytochrome c, mono- and diheme variants	C	Energy production and conversion
fam1_ka	VC_1441	Duplication	HGT	COG2993	Cbb3-type cytochrome oxidase, cytochrome c subunit	C	Energy production and conversion
fam1_jz	VC_1442	Duplication	No	COG3278	Cbb3-type cytochrome oxidase, subunit 1	O	Posttranslational modification, protein turnover, chaperones
fam1_cu	VC_1445*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_cu	VC_1445*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam3_ir	VC_1508	No duplication	No	COG1666	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_hi	VC_1552	Duplication	HGT	COG3839	ABC-type sugar transport systems, ATPase components	G	Carbohydrate transport and metabolism
fam3_ik	VC_1553	No duplication	No	COG1349	Transcriptional regulators of sugar metabolism	KG	Multiple classes
fam3_ij	VC_1554	No duplication	No	COG0584	Glycerophosphoryl diester phosphodiesterase	C	Energy production and conversion
fam3_ii	VC_1555	No duplication	No	COG2731	Beta-galactosidase, beta subunit	G	Carbohydrate transport and metabolism
fam3_ih	VC_1556	No duplication	HGT	COG0251	Putative translation initiation inhibitor, yjgF family	J	Translation, ribosomal structure and biogenesis
fam1_nl	VC_1557	Duplication	HGT	COG1609	Transcriptional regulators	K	Transcription
fam3_ig	VC_1558	No duplication	HGT	COG2723	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	G	Carbohydrate transport and metabolism
fam3_if	VC_1560	No duplication	No	COG0376	Catalase (peroxidase I)	P	Inorganic ion transport and metabolism
fam3_ie	VC_1563	No duplication	No	COG0845	Membrane-fusion protein	M	Cell wall/membrane/envelope biogenesis
fam1_hh	VC_1565	Duplication	HGT	COG1538	Outer membrane protein	MU	Multiple classes
fam1_do	VC_1567	Duplication	HGT	COG0577	ABC-type antimicrobial peptide transport system, permease component	V	Defense mechanisms
fam3_id	VC_1568	No duplication	HGT	COG1136	ABC-type antimicrobial peptide transport system, ATPase component	V	Defense mechanisms
fam5_lv	VC_1570	No duplication	HGT	COG1294	Cytochrome bd-type quinol oxidase, subunit 2	C	Energy production and conversion
fam3_ic	VC_1571	Duplication	HGT	COG1271	Cytochrome bd-type quinol oxidase, subunit 1	C	Energy production and conversion
fam3_ib	VC_1574	No duplication	No	COG3169	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ia	VC_1575	No duplication	HGT	COG3305	Predicted membrane protein	S	Function unknown
fam3_hz	VC_1594	No duplication	No	COG2017	Galactose mutarotase and related enzymes	G	Carbohydrate transport and metabolism
fam3_hy	VC_1595	No duplication	No	COG0153	Galactokinase	G	Carbohydrate transport and metabolism
fam1_bw	VC_1596	Duplication	HGT	COG1085	Galactose-1-phosphate uridylyltransferase	C	Energy production and conversion
fam3_hx	VC_1597	No duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam3_hw	VC_1599	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam3_hv	VC_1600	No duplication	No	COG0551	Zn-finger domain associated with topoisomerase type I	L	Replication, recombination and repair
fam3_hu	VC_1601	No duplication	No	No hit			
fam3_ht	VC_1603	No duplication	No	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam3_hs	VC_1604	No duplication	No	COG4565	Response regulator of citrate/malate metabolism	KT	Multiple classes
fam1_ap	VC_1606	Duplication	HGT	COG1538	Outer membrane protein	MU	Multiple classes
fam1_ap	VC_1607	Duplication	HGT	COG1566	Multidrug resistance efflux pump	V	Defense mechanisms
fam3_hr	VC_1608	No duplication	No	COG0842	ABC-type multidrug transport system, permease component	V	Defense mechanisms
fam3_hq	VC_1614	No duplication	HGT	COG3129	Predicted SAM-dependent methyltransferase	R	General function prediction only
fam3_hp	VC_1615	No duplication	HGT	COG1451	Predicted metal-dependent hydrolase	R	General function prediction only
fam1_jk	VC_1616	Duplication	HGT	COG0695	Glutaredoxin and related proteins	O	Posttranslational modification, protein turnover, chaperones
fam3_ho	VC_1617	No duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam3_hn	VC_1618	No duplication	HGT	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam3_hm	VC_1619	No duplication	HGT	COG3310	Uncharacterized protein conserved in bacteria	S	Function unknown

fam3_hl	VC_1621	No duplication	No	COG1538	Outer membrane protein	MU	Multiple classes
fam3_hk	VC_1622	No duplication	No	COG2885	Outer membrane protein and related peptidoglycan-associated (lipo)proteins	M	Cell wall/membrane/envelope biogenesis
fam3_hj	VC_1624	No duplication	No	COG1748	Saccharopine dehydrogenase and related proteins	E	Amino acid transport and metabolism
fam1_nk	VC_1627	Duplication	No	COG3004	Na+/H+ antiporter	P	Inorganic ion transport and metabolism
fam1_nj	VC_1629	Duplication	No	COG4591	ABC-type transport system, involved in lipoprotein release, permease component	M	Cell wall/membrane/envelope biogenesis
fam3_hi	VC_1630	No duplication	No	COG1136	ABC-type antimicrobial peptide transport system, ATPase component	V	Defense mechanisms
fam3_hh	VC_1634	No duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam3_hg	VC_1635	No duplication	No	COG1187	16S rRNA uridine-516 pseudouridylate synthase and related pseudouridylate synthases	J	Translation, ribosomal structure and biogenesis
fam3_hf	VC_1636	No duplication	No	COG1061	DNA or RNA helicases of superfamily II	KL	Multiple classes
fam3_he	VC_1638	No duplication	No	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam3_hc	VC_1640	No duplication	No	COG1825	Ribosomal protein L25 (general stress protein Ctc)	J	Translation, ribosomal structure and biogenesis
fam3_hb	VC_1642	No duplication	No	No hit			
fam3_ha	VC_1643	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam5_kx	VC_1649	No duplication	HGT	COG5640	Secreted trypsin-like serine protease	O	Posttranslational modification, protein turnover, chaperones
fam1_jo	VC_1650	Duplication	No	No hit			
fam1_hg	VC_1653*	Duplication	HGT	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_hg	VC_1653*	Duplication	HGT	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam1_hg	VC_1653*	Duplication	HGT	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam1_hg	VC_1653*	Duplication	HGT	COG2198	FOG: HPt domain	T	Signal transduction mechanisms
fam3_gy	VC_1655	No duplication	HGT	COG2239	Mg/Co/Ni transporter MgtE (contains CBS domain)	P	Inorganic ion transport and metabolism
fam3_gx	VC_1656	No duplication	No	No hit			
fam3_gw	VC_1658	No duplication	No	COG0814	Amino acid permeases	E	Amino acid transport and metabolism
fam3_gu	VC_1664	No duplication	No	COG4134	ABC-type uncharacterized transport system, periplasmic component	R	General function prediction only
fam1_by	VC_1665	Duplication	No	COG4135	ABC-type uncharacterized transport system, permease component	R	General function prediction only
fam3_gt	VC_1666	No duplication	No	COG4136	ABC-type uncharacterized transport system, ATPase component	R	General function prediction only
fam3_gs	VC_1667	No duplication	No	COG0558	Phosphatidylglycerophosphate synthase	I	Lipid transport and metabolism
fam3_gr	VC_1668	No duplication	No	COG0564	Pseudouridylate synthases, 23S RNA-specific	J	Translation, ribosomal structure and biogenesis
fam3_gq	VC_1670	No duplication	No	COG1502	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthases and related enzymes	I	Lipid transport and metabolism
fam3_gp	VC_1671	No duplication	No	COG0626	Cystathionine beta-lyases/cystathionine gamma-synthases	E	Amino acid transport and metabolism
fam3_go	VC_1672	No duplication	No	COG2818	3-methyladenine DNA glycosylase	L	Replication, recombination and repair
fam1_aw	VC_1673	Duplication	HGT	COG0841	Cation/multidrug efflux pump	V	Defense mechanisms
fam3_gn	VC_1676	No duplication	No	COG1983	Putative stress-responsive transcriptional regulator	KT	Multiple classes
fam3_gm	VC_1677	No duplication	No	No hit			
fam3_gl	VC_1678	No duplication	No	COG1842	Phage shock protein A (IM30), suppresses sigma54-dependent transcription	KT	Multiple classes
fam3_gk	VC_1679	No duplication	No	COG1221	Transcriptional regulators containing an AAA-type ATPase domain and a DNA-binding domain	KT	Multiple classes
fam3_gj	VC_1680	No duplication	No	COG4166	ABC-type oligopeptide transport system, periplasmic component	E	Amino acid transport and metabolism
fam3_gi	VC_1681	No duplication	No	COG4168	ABC-type antimicrobial peptide transport system, permease component	V	Defense mechanisms
fam3_gh	VC_1682	No duplication	No	COG4171	ABC-type antimicrobial peptide transport system, permease component	V	Defense mechanisms
fam3_gg	VC_1683	No duplication	No	COG4170	ABC-type antimicrobial peptide transport system, ATPase component	V	Defense mechanisms
fam1_hf	VC_1684	Duplication	No	COG4167	ABC-type antimicrobial peptide transport system, ATPase component	V	Defense mechanisms
fam1_ni	VC_1685	Duplication	HGT	No hit			
fam3_gf	VC_1686	No duplication	HGT	No hit			
fam3_ge	VC_1687	No duplication	No	COG1227	Inorganic pyrophosphatase/exopolyphosphatase	C	Energy production and conversion
fam3_gd	VC_1690	No duplication	No	COG3345	Alpha-galactosidase	G	Carbohydrate transport and metabolism
fam1_al	VC_1692	Duplication	HGT	COG0243	Anaerobic dehydrogenases, typically selenocysteine-containing	C	Energy production and conversion
fam3_gc	VC_1695	No duplication	No	COG2116	Formate/nitrite family of transporters	P	Inorganic ion transport and metabolism
fam3_gb	VC_1696	No duplication	No	No hit			
fam3_ga	VC_1697	No duplication	No	COG0451	Nucleoside-diphosphate-sugar epimerases	MG	Multiple classes
fam3_fz	VC_1698	No duplication	No	COG1434	Uncharacterized conserved protein	S	Function unknown
fam3_fy	VC_1699	No duplication	No	No hit			
fam3_fx	VC_1700	No duplication	No	COG2917	Intracellular septation protein A	D	Cell cycle control, cell division, chromosome partitioning
fam3_fw	VC_1701	No duplication	No	COG1607	Acyl-CoA hydrolase	I	Lipid transport and metabolism
fam1_nh	VC_1704	Duplication	No	COG0620	Methionine synthase II (cobalamin-independent)	E	Amino acid transport and metabolism
fam3_fv	VC_1706	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam3_fu	VC_1707	No duplication	HGT	No hit			
fam3_ft	VC_1708	No duplication	No	No hit			
fam1_ng	VC_1709	Duplication	No	COG0612	Predicted Zn-dependent peptidases	R	General function prediction only
fam1_nf	VC_1710*	Duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam1_nf	VC_1710*	Duplication	No	COG2202	FOG: PAS/PAC domain	T	Signal transduction mechanisms
fam1_ne	VC_1711	Duplication	HGT	COG1032	Fe-S oxidoreductase	C	Energy production and conversion
fam3_fs	VC_1713	No duplication	HGT	COG1522	Transcriptional regulators	K	Transcription
fam1_he	VC_1714	Duplication	HGT	COG3096	Uncharacterized protein involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam3_fr	VC_1715	No duplication	HGT	COG3095	Uncharacterized protein involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam3_fq	VC_1716	No duplication	HGT	COG3006	Uncharacterized protein involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam3_fp	VC_1717	No duplication	HGT	COG2227	2-polyprenyl-3-methyl-5-hydroxy-6-metoxy-1,4-benzoquinol methylase	H	Coenzyme transport and metabolism
fam3_fo	VC_1718	No duplication	No	COG1434	Uncharacterized conserved protein	S	Function unknown
fam3_fn	VC_1721	No duplication	No	COG1609	Transcriptional regulators	K	Transcription
fam3_fm	VC_1722	No duplication	No	COG3070	Regulator of competence-specific genes	K	Transcription

fam1_hd	VC_1723	Duplication	No	COG0398	Uncharacterized conserved protein	S	Function unknown
fam3_fl	VC_1724	No duplication	No	No hit			
fam3_fk	VC_1725	No duplication	No	COG0596	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	R	General function prediction only
fam1_nd	VC_1726	Duplication	No	COG0297	Glycogen synthase	G	Carbohydrate transport and metabolism
fam1_nc	VC_1727	Duplication	No	COG0448	ADP-glucose pyrophosphorylase	G	Carbohydrate transport and metabolism
fam1_nb	VC_1730*	Duplication	No	COG0550	Topoisomerase IA	L	Replication, recombination and repair
fam1_nb	VC_1730*	Duplication	No	COG0551	Zn-finger domain associated with topoisomerase type I	L	Replication, recombination and repair
fam3_fj	VC_1731	No duplication	No	No hit			
fam3_fi	VC_1732	No duplication	No	COG0128	5-enolpyruvylshikimate-3-phosphate synthase	E	Amino acid transport and metabolism
fam3_fh	VC_1735	No duplication	No	COG2360	Leu/Phe-tRNA-protein transferase	O	Posttranslational modification, protein turnover, chaperones
fam3_ff	VC_1736	No duplication	No	COG2935	Putative arginyl-tRNA:protein arginylyltransferase	O	Posttranslational modification, protein turnover, chaperones
fam3_fe	VC_1737	No duplication	No	COG0361	Translation initiation factor 1 (IF-1)	J	Translation, ribosomal structure and biogenesis
fam3_fd	VC_1738	No duplication	No	COG3007	Uncharacterized paraquat-inducible protein B	S	Function unknown
fam1_ah	VC_1740	Duplication	HGT	COG1960	Acyl-CoA dehydrogenases	I	Lipid transport and metabolism
fam3_fc	VC_1741	No duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam3_fb	VC_1743	No duplication	No	COG2515	1-aminocyclopropane-1-carboxylate deaminase	E	Amino acid transport and metabolism
fam3_fa	VC_1744	No duplication	No	No hit			
fam1_df	VC_1745	Duplication	No	COG1012	NAD-dependent aldehyde dehydrogenases	C	Energy production and conversion
fam3_ez	VC_1746	No duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam5_kd	VC_1755	No duplication	No	COG3009	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ey	VC_1756	No duplication	No	COG0845	Membrane-fusion protein	M	Cell wall/membrane/envelope biogenesis
fam1_aw	VC_1757	Duplication	HGT	COG0841	Cation/multidrug efflux pump	V	Defense mechanisms
fam3_ew	VC_1828	No duplication	HGT	No hit			
fam1_na	VC_1831*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_na	VC_1831*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam1_mz	VC_1832	Duplication	No	COG3318	Predicted metal-binding protein related to the C-terminal domain of SecA	R	General function prediction only
fam3_ev	VC_1833	No duplication	No	COG0379	Quinolinate synthase	H	Coenzyme transport and metabolism
fam3_eu	VC_1834	No duplication	No	COG1729	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_et	VC_1835	No duplication	No	COG2885	Outer membrane protein and related peptidoglycan-associated (lipo)proteins	M	Cell wall/membrane/envelope biogenesis
fam3_ep	VC_1860	No duplication	No	COG0708	Exonuclease III	L	Replication, recombination and repair
fam3_eo	VC_1861	No duplication	No	COG4160	ABC-type arginine/histidine transport system, permease component	E	Amino acid transport and metabolism
fam3_en	VC_1862	No duplication	No	COG4215	ABC-type arginine transport system, permease component	E	Amino acid transport and metabolism
fam3_em	VC_1863	No duplication	No	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam1_my	VC_1864	Duplication	No	COG4598	ABC-type histidine transport system, ATPase component	E	Amino acid transport and metabolism
fam1_hc	VC_1865	Duplication	No	No hit			
fam3_el	VC_1866	No duplication	No	COG1882	Pyruvate-formate lyase	C	Energy production and conversion
fam1_mx	VC_1867	Duplication	No	COG3528	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ek	VC_1868	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_mw	VC_1869	Duplication	No	COG1180	Pyruvate-formate lyase-activating enzyme	O	Posttranslational modification, protein turnover, chaperones
fam3_ej	VC_1870	No duplication	No	COG3235	Predicted membrane protein	S	Function unknown
fam3_ei	VC_1871	No duplication	No	COG3013	Uncharacterized conserved protein	S	Function unknown
fam3_eh	VC_1874	No duplication	No	COG2719	Uncharacterized conserved protein	S	Function unknown
fam3_eg	VC_1875	No duplication	No	COG1212	CMP-2-keto-3-deoxyoctulosonic acid synthetase	M	Cell wall/membrane/envelope biogenesis
fam3_ef	VC_1876	No duplication	No	COG2835	Uncharacterized conserved protein	S	Function unknown
fam3_ee	VC_1877	No duplication	No	COG1663	Tetraacyldisaccharide-1-P 4'-kinase	M	Cell wall/membrane/envelope biogenesis
fam1_hb	VC_1878	Duplication	No	COG1132	ABC-type multidrug transport system, ATPase and permease components	V	Defense mechanisms
fam1_mv	VC_1879	Duplication	No	COG2333	Predicted hydrolase (metallo-beta-lactamase superfamily)	R	General function prediction only
fam1_ha	VC_1880	Duplication	No	COG3216	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ed	VC_1882	No duplication	No	COG4591	ABC-type transport system, involved in lipoprotein release, permease component	M	Cell wall/membrane/envelope biogenesis
fam3_ec	VC_1883	No duplication	No	COG1136	ABC-type antimicrobial peptide transport system, ATPase component	V	Defense mechanisms
fam3_eb	VC_1884	No duplication	No	COG4591	ABC-type transport system, involved in lipoprotein release, permease component	M	Cell wall/membrane/envelope biogenesis
fam3_ea	VC_1885	No duplication	No	No hit			
fam1_gz	VC_1886	Duplication	No	COG1197	Transcription-repair coupling factor (superfamily II helicase)	LK	Multiple classes
fam3_dz	VC_1887	No duplication	No	No hit			
fam1_ac	VC_1888	Duplication	HGT	No hit			
fam3_dx	VC_1889	No duplication	No	COG1670	Acetyltransferases, including N-acetylases of ribosomal proteins	J	Translation, ribosomal structure and biogenesis
fam3_dw	VC_1890	No duplication	No	COG1252	NADH dehydrogenase, FAD-containing subunit	C	Energy production and conversion
fam3_dv	VC_1892	No duplication	No	COG3150	Predicted esterase	R	General function prediction only
fam3_du	VC_1893	No duplication	No	COG0510	Predicted choline kinase involved in LPS biosynthesis	M	Cell wall/membrane/envelope biogenesis
fam1_jt	VC_1894	Duplication	No	COG3417	Collagen-binding surface adhesin SpaP (antigen I/II family)	R	General function prediction only
fam3_dt	VC_1895	No duplication	No	COG5633	Predicted periplasmic lipoprotein	R	General function prediction only
fam1_mu	VC_1898*	Duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_mu	VC_1898*	Duplication	No	COG5278	Predicted periplasmic ligand-binding sensor domain	T	Signal transduction mechanisms
fam3_ds	VC_1899	No duplication	No	No hit			
fam3_dr	VC_1900	No duplication	No	COG2186	Transcriptional regulators	K	Transcription
fam3_dq	VC_1902	No duplication	No	COG1495	Disulfide bond formation protein DsbB	O	Posttranslational modification, protein turnover, chaperones
fam3_do	VC_1912	No duplication	No	COG2956	Predicted N-acetylglucosaminyl transferase	G	Carbohydrate transport and metabolism
fam3_dn	VC_1925	No duplication	No	COG4191	Signal transduction histidine kinase regulating C4-dicarboxylate transport system	T	Signal transduction mechanisms

fam3_dm	VC_1926	No duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam3_dl	VC_1927	No duplication	No	COG1593	TRAP-type C4-dicarboxylate transport system, large permease component	G	Carbohydrate transport and metabolism
fam3_dk	VC_1928	No duplication	No	COG3090	TRAP-type C4-dicarboxylate transport system, small permease component	G	Carbohydrate transport and metabolism
fam3_dj	VC_1929	No duplication	No	COG1638	TRAP-type C4-dicarboxylate transport system, periplasmic component	G	Carbohydrate transport and metabolism
fam3_di	VC_1931	No duplication	No	COG0397	Uncharacterized conserved protein	S	Function unknown
fam3_dg	VC_1938	No duplication	No	COG0824	Predicted thioesterase	R	General function prediction only
fam3_df	VC_1939	No duplication	No	COG1280	Putative threonine efflux protein	E	Amino acid transport and metabolism
fam3_de	VC_1940	Duplication	No	COG0560	Phosphoserine phosphatase	E	Amino acid transport and metabolism
fam3_dd	VC_1941	No duplication	No	COG4645	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_dc	VC_1942	No duplication	No	COG0190	5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase	H	Coenzyme transport and metabolism
fam1_al	VC_1950	Duplication	HGT	COG0243	Anaerobic dehydrogenases, typically selenocysteine-containing	C	Energy production and conversion
fam5_iy	VC_1951	No duplication	No	COG3005	Nitrate/TMAO reductases, membrane-bound tetraheme cytochrome c subunit	C	Energy production and conversion
fam3_db	VC_1953	No duplication	HGT	COG1972	Nucleoside permease	F	Nucleotide transport and metabolism
fam3_da	VC_1955	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam3_cz	VC_1959	No duplication	No	COG0850	Septum formation inhibitor	D	Cell cycle control, cell division, chromosome partitioning
fam3_cy	VC_1960	No duplication	No	COG2894	Septum formation inhibitor-activating ATPase	D	Cell cycle control, cell division, chromosome partitioning
fam3_cx	VC_1963	No duplication	No	COG3492	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ms	VC_1964	Duplication	No	COG4395	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_cw	VC_1965	No duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam1_mr	VC_1966	Duplication	No	COG0204	1-acyl-sn-glycerol-3-phosphate acyltransferase	I	Lipid transport and metabolism
fam3_cu	VC_1971	No duplication	HGT	COG0318	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	IQ	Multiple classes
fam3_ct	VC_1972	No duplication	HGT	COG1441	O-succinylbenzoate synthase	H	Coenzyme transport and metabolism
fam3_cs	VC_1973	No duplication	HGT	COG0447	Dihydroxynaphthoic acid synthase	H	Coenzyme transport and metabolism
fam3_cr	VC_1974	No duplication	No	COG2267	Lysophospholipase	I	Lipid transport and metabolism
fam3_cq	VC_1975	No duplication	No	COG1165	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase	H	Coenzyme transport and metabolism
fam3_cp	VC_1976	No duplication	No	COG1169	Isochorismate synthase	HQ	Multiple classes
fam3_co	VC_1977	No duplication	No	COG0436	Aspartate/tyrosine/aromatic aminotransferase	E	Amino acid transport and metabolism
fam3_cn	VC_1978	No duplication	No	COG1896	Predicted hydrolases of HD superfamily	R	General function prediction only
fam3_cm	VC_1979	No duplication	No	COG0232	dGTP triphosphohydrolase	F	Nucleotide transport and metabolism
fam3_cl	VC_1980	No duplication	HGT	COG3148	Uncharacterized conserved protein	S	Function unknown
fam1_mq	VC_1981	Duplication	HGT	COG0705	Uncharacterized membrane protein (homolog of Drosophila rhomboid)	R	General function prediction only
fam3_ck	VC_1985	No duplication	HGT	COG0318	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	IQ	Multiple classes
fam3_cj	VC_1986	No duplication	HGT	COG0596	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	R	General function prediction only
fam3_ci	VC_1987	No duplication	HGT	COG3065	Starvation-inducible outer membrane lipoprotein	M	Cell wall/membrane/envelope biogenesis
fam3_ch	VC_1988	No duplication	No	No hit			
fam3_cg	VC_1990	No duplication	No	COG1199	Rad3-related DNA helicases	KL	Multiple classes
fam1_mp	VC_1992	Duplication	No	COG0788	Formyltetrahydrofolate hydrolase	F	Nucleotide transport and metabolism
fam3_cf	VC_1993*	No duplication	No	COG1902	NADH:flavin oxidoreductases, Old Yellow Enzyme family	C	Energy production and conversion
fam3_cf	VC_1993*	No duplication	No	COG0446	Uncharacterized NAD(FAD)-dependent dehydrogenases	R	General function prediction only
fam3_cd	VC_2001	No duplication	No	COG0676	Uncharacterized enzymes related to aldose 1-epimerase	G	Carbohydrate transport and metabolism
fam3_cc	VC_2004	No duplication	No	COG3016	Uncharacterized iron-regulated protein	S	Function unknown
fam1_mo	VC_2005	Duplication	No	No hit			
fam3_cb	VC_2006*	No duplication	No	COG0835	Chemotaxis signal transduction protein	NT	Multiple classes
fam3_cb	VC_2006*	No duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam3_bz	VC_2007	No duplication	No	COG1940	Transcriptional regulator/sugar kinase	KG	Multiple classes
fam3_by	VC_2008	No duplication	No	COG0469	Pyruvate kinase	G	Carbohydrate transport and metabolism
fam3_bx	VC_2009	No duplication	No	COG4667	Predicted esterase of the alpha-beta hydrolase superfamily	R	General function prediction only
fam3_bw	VC_2011	No duplication	No	No hit			
fam3_bv	VC_2012	No duplication	No	COG0733	Na+-dependent transporters of the SNF family	R	General function prediction only
fam1_gx	VC_2013*	Duplication	No	COG1263	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	G	Carbohydrate transport and metabolism
fam1_gx	VC_2013*	Duplication	No	COG1264	Phosphotransferase system IIB components	G	Carbohydrate transport and metabolism
fam3_bu	VC_2014	No duplication	No	COG0084	Mg-dependent DNase	L	Replication, recombination and repair
fam3_bt	VC_2015	No duplication	No	COG0470	ATPase involved in DNA replication	L	Replication, recombination and repair
fam3_br	VC_2016	No duplication	No	COG0125	Thymidylate kinase	F	Nucleotide transport and metabolism
fam3_bq	VC_2017	No duplication	No	COG1559	Predicted periplasmic solute-binding protein	R	General function prediction only
fam3_bp	VC_2018	No duplication	No	COG0115	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase	EH	Multiple classes
fam3_bo	VC_2019	No duplication	No	COG0304	3-oxoacyl-(acyl-carrier-protein) synthase	IQ	Multiple classes
fam3_bm	VC_2037	No duplication	No	COG1757	Na+/H+ antiporter	C	Energy production and conversion
fam1_mm	VC_2039	Duplication	No	COG3081	Nucleoid-associated protein	R	General function prediction only
fam3_bk	VC_2046	No duplication	No	No hit			
fam3_bi	VC_2047	No duplication	No	COG1028	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	IQR	Multiple classes
fam3_bh	VC_2048	No duplication	No	COG2853	Surface lipoprotein	M	Cell wall/membrane/envelope biogenesis
fam3_bg	VC_2051	No duplication	No	COG0526	Thiol-disulfide isomerase and thioredoxins	OC	Multiple classes
fam3_bf	VC_2052	No duplication	No	COG1138	Cytochrome c biogenesis factor	O	Posttranslational modification, protein turnover, chaperones
fam3_be	VC_2053	No duplication	No	COG2332	Cytochrome c-type biogenesis protein CcmE	O	Posttranslational modification, protein turnover, chaperones
fam3_bd	VC_2054	No duplication	No	COG3114	Heme exporter protein D	U	Intracellular trafficking, secretion, and vesicular transport
fam1_mk	VC_2055	Duplication	No	COG0755	ABC-type transport system involved in cytochrome c biogenesis, permease component	O	Posttranslational modification, protein turnover, chaperones
fam3_bb	VC_2056	No duplication	No	COG2386	ABC-type transport system involved in cytochrome c biogenesis, permease component	O	Posttranslational modification, protein turnover, chaperones

fam3_ba	VC_2057	No duplication	No	COG4133	ABC-type transport system involved in cytochrome c biogenesis, ATPase component	O	Posttranslational modification, protein turnover, chaperones
fam3_az	VC_2058	No duplication	No	No hit			
fam3_ay	VC_2059	No duplication	No	COG0835	Chemotaxis signal transduction protein	NT	Multiple classes
fam1_gw	VC_2060	Duplication	No	No hit			
fam3_ax	VC_2061	No duplication	No	COG1192	ATPases involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam3_aw	VC_2062	No duplication	No	COG2201	Chemotaxis response regulator containing a CheY-like receiver domain and a methylesterase domain	NT	Multiple classes
fam3_av	VC_2064	No duplication	No	COG3143	Chemotaxis protein	NT	Multiple classes
fam3_au	VC_2065	No duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam3_as	VC_2066	No duplication	No	COG1191	DNA-directed RNA polymerase specialized sigma subunit	K	Transcription
fam3_ar	VC_2067	No duplication	No	COG0455	ATPases involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning
fam3_aq	VC_2068	No duplication	No	COG1419	Flagellar GTP-binding protein	N	Cell motility
fam1_kl	VC_2069	Duplication	No	COG1298	Flagellar biosynthesis pathway, component FlhA	NU	Multiple classes
fam1_mj	VC_2070	Duplication	No	COG2062	Phosphohistidine phosphatase SixA	T	Signal transduction mechanisms
fam1_gv	VC_2072	Duplication	No	COG1025	Secreted/periplasmic Zn-dependent peptidases, insulinase-like	O	Posttranslational modification, protein turnover, chaperones
fam3_ap	VC_2073	No duplication	No	COG3102	Uncharacterized protein conserved in bacteria	S	Function unknown
fam3_ao	VC_2074	No duplication	No	COG0018	Arginyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam3_an	VC_2075	No duplication	No	COG0537	Diadenosine tetraphosphate (Ap4A) hydrolase and other HIT family hydrolases	FGR	Multiple classes
fam3_am	VC_2077	No duplication	No	COG0370	Fe2+ transport system protein B	P	Inorganic ion transport and metabolism
fam3_al	VC_2078	No duplication	No	COG1918	Fe2+ transport system protein A	P	Inorganic ion transport and metabolism
fam1_cg	VC_2080*	Duplication	No	COG3292	Predicted periplasmic ligand-binding sensor domain	T	Signal transduction mechanisms
fam1_cg	VC_2080*	Duplication	No	COG2207	AraC-type DNA-binding domain-containing proteins	K	Transcription
fam3_ak	VC_2081	No duplication	No	COG4531	ABC-type Zn2+ transport system, periplasmic component/surface adhesin	P	Inorganic ion transport and metabolism
fam3_aj	VC_2082	No duplication	No	COG1121	ABC-type Mn/Zn transport systems, ATPase component	P	Inorganic ion transport and metabolism
fam3_ai	VC_2083	No duplication	No	COG1108	ABC-type Mn2+/Zn2+ transport systems, permease components	P	Inorganic ion transport and metabolism
fam3_ah	VC_2084	No duplication	No	COG0074	Succinyl-CoA synthetase, alpha subunit	C	Energy production and conversion
fam3_ag	VC_2085	No duplication	No	COG0045	Succinyl-CoA synthetase, beta subunit	C	Energy production and conversion
fam3_af	VC_2086	No duplication	No	COG0508	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	C	Energy production and conversion
fam1_dm	VC_2087	Duplication	No	COG0567	2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, and related enzymes	C	Energy production and conversion
fam3_ad	VC_2088	No duplication	No	COG0479	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	C	Energy production and conversion
fam3_ac	VC_2091	No duplication	No	COG2009	Succinate dehydrogenase/fumarate reductase, cytochrome b subunit	C	Energy production and conversion
fam3_ab	VC_2092	No duplication	No	COG0372	Citrate synthase	C	Energy production and conversion
fam3_aa	VC_2094	No duplication	No	COG3782	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_mi	VC_2095	Duplication	No	COG0033	Phosphoglucomutase	G	Carbohydrate transport and metabolism
fam2_tf	VC_2096	No duplication	HGT	COG3057	Negative regulator of replication initiationR	L	Replication, recombination and repair
fam1_mh	VC_2110*	Duplication	No	COG4121	Uncharacterized conserved protein	S	Function unknown
fam1_mh	VC_2110*	Duplication	No	COG0665	Glycine/D-amino acid oxidases (deaminating)	E	Amino acid transport and metabolism
fam2_te	VC_2111	No duplication	HGT	COG0431	Predicted flavoprotein	R	General function prediction only
fam2_td	VC_2113	No duplication	HGT	COG3101	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_tc	VC_2115	No duplication	No	COG2860	Predicted membrane protein	S	Function unknown
fam2_tb	VC_2116	No duplication	No	COG0082	Chorismate synthase	E	Amino acid transport and metabolism
fam2_ta	VC_2118	No duplication	HGT	COG2890	Methylase of polypeptide chain release factors	J	Translation, ribosomal structure and biogenesis
fam2_sz	VC_2120	No duplication	No	COG1377	Flagellar biosynthesis pathway, component FlhB	NU	Multiple classes
fam2_sx	VC_2121	No duplication	No	COG1684	Flagellar biosynthesis pathway, component FliR	NU	Multiple classes
fam2_sw	VC_2122	No duplication	No	COG1987	Flagellar biosynthesis pathway, component FliQ	NU	Multiple classes
fam2_sv	VC_2123	No duplication	No	COG1338	Flagellar biosynthesis pathway, component FliP	NU	Multiple classes
fam2_su	VC_2124	No duplication	No	COG3190	Flagellar biogenesis protein	N	Cell motility
fam2_st	VC_2125	No duplication	No	COG1886	Flagellar motor switch/type III secretory pathway protein	NU	Multiple classes
fam2_ss	VC_2126	No duplication	No	COG1868	Flagellar motor switch protein	N	Cell motility
fam2_sr	VC_2127	No duplication	No	COG1580	Flagellar basal body-associated protein	N	Cell motility
fam1_gu	VC_2128	Duplication	No	COG3144	Flagellar hook-length control protein	N	Cell motility
fam2_sq	VC_2129	No duplication	No	COG2882	Flagellar biosynthesis chaperone	NUO	Multiple classes
fam2_sp	VC_2130	No duplication	No	COG1157	Flagellar biosynthesis/type III secretory pathway ATPase	NU	Multiple classes
fam2_sn	VC_2131	No duplication	No	COG1317	Flagellar biosynthesis/type III secretory pathway protein	NU	Multiple classes
fam2_sm	VC_2133	No duplication	No	COG1766	Flagellar biosynthesis/type III secretory pathway lipoprotein	NU	Multiple classes
fam2_sl	VC_2134	No duplication	No	COG1677	Flagellar hook-basal body protein	NU	Multiple classes
fam2_sk	VC_2136	No duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam2_sj	VC_2137	No duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam2_si	VC_2138	No duplication	No	COG1516	Flagellin-specific chaperone FliS	NUO	Multiple classes
fam2_sh	VC_2139	No duplication	No	No hit			
fam2_sg	VC_2140	No duplication	No	COG1345	Flagellar capping protein	N	Cell motility
fam2_sf	VC_2141	No duplication	No	COG1334	Uncharacterized flagellar protein FlaG	N	Cell motility
fam1_aa	VC_2142	Duplication	HGT	COG1344	Flagellin and related hook-associated proteins	N	Cell motility
fam1_aa	VC_2143	Duplication	HGT	COG1344	Flagellin and related hook-associated proteins	N	Cell motility
fam1_gt	VC_2145	Duplication	No	COG2837	Predicted iron-dependent peroxidase	P	Inorganic ion transport and metabolism
fam2_se	VC_2146	No duplication	No	No hit			
fam2_sd	VC_2147	No duplication	No	No hit			
fam2_sc	VC_2149	No duplication	No	COG4519	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_sb	VC_2150	No duplication	No	COG1393	Arsenate reductase and related proteins, glutaredoxin family	P	Inorganic ion transport and metabolism

fam2_rz	VC_2153	No duplication	No	COG1876	D-alanyl-D-alanine carboxypeptidase	M	Cell wall/membrane/envelope biogenesis
fam2_ry	VC_2154	No duplication	No	No hit			
fam2_rx	VC_2156	No duplication	No	COG3317	Uncharacterized lipoprotein	M	Cell wall/membrane/envelope biogenesis
fam2_rw	VC_2157	No duplication	No	COG0329	Dihydrodipicolinate synthase/N-acetylneuraminate lyase	EM	Multiple classes
fam2_rv	VC_2159	No duplication	No	COG2716	Glycine cleavage system regulatory protein	E	Amino acid transport and metabolism
fam2_ru	VC_2160	No duplication	No	COG1225	Peroxiredoxin	O	Posttranslational modification, protein turnover, chaperones
fam1_ad	VC_2161	Duplication	HGT	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_gs	VC_2162	Duplication	No	COG0628	Predicted permease	R	General function prediction only
fam2_rt	VC_2163	No duplication	No	COG0425	Predicted redox protein, regulator of disulfide bond formation	O	Posttranslational modification, protein turnover, chaperones
fam2_rr	VC_2164	No duplication	No	COG4783	Putative Zn-dependent protease, contains TPR repeats	R	General function prediction only
fam2_rq	VC_2165	No duplication	No	COG1393	Arsenate reductase and related proteins, glutaredoxin family	P	Inorganic ion transport and metabolism
fam2_rp	VC_2166	No duplication	No	COG0655	Multimeric flavodoxin WrbA	R	General function prediction only
fam2_ro	VC_2167	No duplication	No	COG3308	Predicted membrane protein	S	Function unknown
fam2_rn	VC_2168	No duplication	No	COG3249	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_rm	VC_2171	No duplication	No	COG2233	Xanthine/uracil permeases	F	Nucleotide transport and metabolism
fam2_rl	VC_2172	No duplication	No	COG2606	Uncharacterized conserved protein	S	Function unknown
fam2_rk	VC_2174	No duplication	No	COG0737	5'-nucleotidase/2',3'-cyclic phosphodiesterase and related esterases	F	Nucleotide transport and metabolism
fam1_mf	VC_2178	Duplication	No	COG2890	Methylase of polypeptide chain release factors	J	Translation, ribosomal structure and biogenesis
fam1_me	VC_2180	Duplication	No	COG0373	Glutamyl-tRNA reductase	H	Coenzyme transport and metabolism
fam1_aa	VC_2187	Duplication	HGT	COG1344	Flagellin and related hook-associated proteins	N	Cell motility
fam2_rj	VC_2190	No duplication	No	COG1344	Flagellin and related hook-associated proteins	N	Cell motility
fam2_ri	VC_2191	No duplication	No	COG1256	Flagellar hook-associated protein	N	Cell motility
fam2_rh	VC_2192	No duplication	No	COG1705	Muramidase (flagellum-specific)	NU	Multiple classes
fam2_rg	VC_2193	No duplication	No	COG1706	Flagellar basal-body P-ring protein	N	Cell motility
fam2_rf	VC_2194	No duplication	No	COG2063	Flagellar basal body L-ring protein	N	Cell motility
fam2_re	VC_2195	No duplication	No	COG4786	Flagellar basal body rod protein	N	Cell motility
fam2_rd	VC_2196	No duplication	No	COG4787	Flagellar basal body rod protein	N	Cell motility
fam2_rc	VC_2197	No duplication	No	COG1749	Flagellar hook protein FlgE	N	Cell motility
fam2_ra	VC_2198	No duplication	No	COG1843	Flagellar hook capping protein	N	Cell motility
fam1_js	VC_2199	Duplication	No	COG1558	Flagellar basal body rod protein	N	Cell motility
fam2_qz	VC_2200	No duplication	No	COG1815	Flagellar basal body protein	N	Cell motility
fam2_qy	VC_2201	No duplication	No	COG1352	Methylase of chemotaxis methyl-accepting proteins	NT	Multiple classes
fam2_qx	VC_2202*	No duplication	No	COG0835	Chemotaxis signal transduction protein	NT	Multiple classes
fam2_qx	VC_2202*	No duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam2_qw	VC_2203	No duplication	No	COG1261	Flagellar basal body P-ring biosynthesis protein	NO	Multiple classes
fam2_qv	VC_2204	No duplication	No	COG2747	Negative regulator of flagellin synthesis (anti-sigma28 factor)	KNU	Multiple classes
fam2_qu	VC_2205	No duplication	No	COG3418	Flagellar biosynthesis/type III secretory pathway chaperone	NUO	Multiple classes
fam2_qt	VC_2206	No duplication	No	COG3018	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_qs	VC_2207	No duplication	No	COG5616	Predicted integral membrane protein	S	Function unknown
fam2_qq	VC_2208	No duplication	No	No hit			
fam2_qp	VC_2212	No duplication	No	No hit			
fam2_qo	VC_2213	No duplication	No	COG2885	Outer membrane protein and related peptidoglycan-associated (lipo)proteins	M	Cell wall/membrane/envelope biogenesis
fam2_qm	VC_2215	No duplication	HGT	COG2217	Cation transport ATPase	P	Inorganic ion transport and metabolism
fam2_ql	VC_2216	No duplication	HGT	COG3019	Predicted metal-binding protein	R	General function prediction only
fam2_qj	VC_2221	No duplication	HGT	No hit			
fam2_qi	VC_2222	No duplication	HGT	COG2840	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_qh	VC_2223	No duplication	No	COG1187	16S rRNA uridine-516 pseudouridylate synthase and related pseudouridylate synthases	J	Translation, ribosomal structure and biogenesis
fam5_hk	VC_2224	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_qg	VC_2225	No duplication	No	COG0035	Uracil phosphoribosyltransferase	F	Nucleotide transport and metabolism
fam2_qf	VC_2226	No duplication	HGT	COG0150	Phosphoribosylaminoimidazole (AIR) synthetase	F	Nucleotide transport and metabolism
fam2_qe	VC_2227	No duplication	HGT	COG0299	Folate-dependent phosphoribosylglycinamide formyltransferase PurN	F	Nucleotide transport and metabolism
fam2_qc	VC_2229	No duplication	HGT	COG0121	Predicted glutamine amidotransferase	R	General function prediction only
fam2_qb	VC_2230	No duplication	HGT	COG0279	Phosphoheptose isomerase	G	Carbohydrate transport and metabolism
fam1_ah	VC_2231	Duplication	HGT	COG1960	Acyl-CoA dehydrogenases	I	Lipid transport and metabolism
fam2_qa	VC_2232	No duplication	No	No hit			
fam2_pz	VC_2233	No duplication	No	COG0847	DNA polymerase III, epsilon subunit and related 3'-5' exonucleases	L	Replication, recombination and repair
fam2_py	VC_2234	No duplication	No	COG0328	Ribonuclease HI	L	Replication, recombination and repair
fam2_px	VC_2235	No duplication	No	COG2226	Methylase involved in ubiquinone/menaquinone biosynthesis	H	Coenzyme transport and metabolism
fam2_pw	VC_2236	No duplication	No	COG0491	Zn-dependent hydrolases, including glyoxylases	R	General function prediction only
fam2_pv	VC_2237	No duplication	No	COG0741	Soluble lytic murein transglycosylase and related regulatory proteins (some contain LysM/invasin domains)	M	Cell wall/membrane/envelope biogenesis
fam2_pu	VC_2238	No duplication	No	COG3021	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_ps	VC_2239	No duplication	No	COG0347	Nitrogen regulatory protein PII	E	Amino acid transport and metabolism
fam2_pr	VC_2241	No duplication	No	COG2863	Cytochrome c553	C	Energy production and conversion
fam2_pq	VC_2242	No duplication	No	COG0037	Predicted ATPase of the PP-loop superfamily implicated in cell cycle control	D	Cell cycle control, cell division, chromosome partitioning
fam2_pp	VC_2244	No duplication	No	COG0825	Acetyl-CoA carboxylase alpha subunit	I	Lipid transport and metabolism
fam2_po	VC_2245	No duplication	No	COG0587	DNA polymerase III, alpha subunit	L	Replication, recombination and repair
fam2_pn	VC_2247	No duplication	No	COG0763	Lipid A disaccharide synthetase	M	Cell wall/membrane/envelope biogenesis
fam1_md	VC_2248	Duplication	No	COG1043	Acyl-[acyl carrier protein]--UDP-N-acetylglucosamine O-acyltransferase	M	Cell wall/membrane/envelope biogenesis

fam2_pl	VC_2249	No duplication	No	COG0764	3-hydroxymyristoyl/3-hydroxydecanoyl-(acyl carrier protein) dehydratases	I	Lipid transport and metabolism
fam2_pk	VC_2250	No duplication	No	COG1044	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	M	Cell wall/membrane/envelope biogenesis
fam2_pj	VC_2252	No duplication	No	COG4775	Outer membrane protein/protective antigen OMA87	M	Cell wall/membrane/envelope biogenesis
fam2_ph	VC_2261	No duplication	No	COG0024	Methionine aminopeptidase	J	Translation, ribosomal structure and biogenesis
fam2_pg	VC_2262	No duplication	No	COG2844	UTP:GlnB (protein PII) uridylyltransferase	O	Posttranslational modification, protein turnover, chaperones
fam2_pd	VC_2269*	No duplication	No	COG0108	3,4-dihydroxy-2-butanone 4-phosphate synthase	H	Coenzyme transport and metabolism
fam2_pd	VC_2269*	No duplication	No	COG0807	GTP cyclohydrolase II	H	Coenzyme transport and metabolism
fam2_pc	VC_2270	No duplication	No	COG0307	Riboflavin synthase alpha chain	H	Coenzyme transport and metabolism
fam2_pb	VC_2271*	No duplication	No	COG0117	Pyrimidine deaminase	H	Coenzyme transport and metabolism
fam2_pb	VC_2271*	No duplication	No	COG1985	Pyrimidine reductase, riboflavin biosynthesis	H	Coenzyme transport and metabolism
fam2_pa	VC_2272	No duplication	HGT	COG1327	Predicted transcriptional regulator, consists of a Zn-ribbon and ATP-cone domains	K	Transcription
fam2_oz	VC_2273	No duplication	No	COG0014	Gamma-glutamyl phosphate reductase	E	Amino acid transport and metabolism
fam1_mb	VC_2274	Duplication	No	COG0263	Glutamate 5-kinase	E	Amino acid transport and metabolism
fam2_oy	VC_2278	No duplication	No	COG2252	Permeases	R	General function prediction only
fam2_ox	VC_2279	No duplication	No	COG2195	Di- and tripeptidases	E	Amino acid transport and metabolism
fam2_ow	VC_2280	No duplication	No	COG4067	Uncharacterized protein conserved in archaea	O	Posttranslational modification, protein turnover, chaperones
fam2_ov	VC_2281	No duplication	No	COG0189	Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase)	HJ	Multiple classes
fam2_ou	VC_2282	No duplication	No	COG3608	Predicted deacylase	R	General function prediction only
fam2_ot	VC_2283	No duplication	No	COG0733	Na+-dependent transporters of the SNF family	R	General function prediction only
fam2_os	VC_2285*	No duplication	No	COG0733	Predicted periplasmic ligand-binding sensor domain	T	Signal transduction mechanisms
fam2_os	VC_2285*	No duplication	No	COG3322	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_or	VC_2286	No duplication	No	No hit			
fam2_oq	VC_2287	No duplication	No	COG0389	Nucleotidyltransferase/DNA polymerase involved in DNA repair	L	Replication, recombination and repair
fam2_op	VC_2288	No duplication	No	COG2991	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_oo	VC_2289	No duplication	No	COG1477	Membrane-associated lipoprotein involved in thiamine biosynthesis	H	Coenzyme transport and metabolism
fam2_on	VC_2290	No duplication	No	COG2871	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrF	C	Energy production and conversion
fam2_om	VC_2292	No duplication	No	COG1347	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrD	C	Energy production and conversion
fam2_ol	VC_2293	No duplication	No	COG2869	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrC	C	Energy production and conversion
fam2_ok	VC_2294	No duplication	No	COG1805	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrB	C	Energy production and conversion
fam2_oj	VC_2295	No duplication	No	COG1726	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrA	C	Energy production and conversion
fam2_oi	VC_2296	No duplication	No	COG0271	Stress-induced morphogen (activity unknown)	T	Signal transduction mechanisms
fam1_ma	VC_2297	Duplication	No	COG2813	16S RNA G1207 methylase RsmC	J	Translation, ribosomal structure and biogenesis
fam2_oh	VC_2298	No duplication	No	COG3056	Uncharacterized lipoprotein	M	Cell wall/membrane/envelope biogenesis
fam2_og	VC_2299	No duplication	No	COG0652	Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	O	Posttranslational modification, protein turnover, chaperones
fam2_oe	VC_2300	No duplication	No	No hit			
fam2_od	VC_2301	No duplication	No	COG3806	Anti-sigma factor	T	Signal transduction mechanisms
fam1_lz	VC_2302	Duplication	No	COG1595	DNA-directed RNA polymerase specialized sigma subunit, sigma24 homolog	K	Transcription
fam2_oc	VC_2305	No duplication	HGT	COG3248	Nucleoside-binding outer membrane protein	M	Cell wall/membrane/envelope biogenesis
fam2_ob	VC_2307	No duplication	HGT	COG1893	Ketopantoate reductase	H	Coenzyme transport and metabolism
fam2_oa	VC_2309	No duplication	HGT	COG0520	Selenocysteine lyase	E	Amino acid transport and metabolism
fam2_nz	VC_2310	No duplication	HGT	COG2166	SufE protein probably involved in Fe-S center assembly	R	General function prediction only
fam2_ny	VC_2311	No duplication	HGT	COG1179	Dinucleotide-utilizing enzymes involved in molybdopterin and thiamine biosynthesis family 1	H	Coenzyme transport and metabolism
fam2_nw	VC_2314	No duplication	No	No hit			
fam1_gr	VC_2316*	Duplication	HGT	COG0548	Acetylglutamate kinase	E	Amino acid transport and metabolism
fam1_gr	VC_2316*	Duplication	HGT	COG1246	N-acetylglutamate synthase and related acetyltransferases	E	Amino acid transport and metabolism
fam2_nv	VC_2319	No duplication	HGT	COG0507	ATP-dependent exoDNAse (exonuclease V), alpha subunit - helicase superfamily I member	L	Replication, recombination and repair
fam1_ly	VC_2320	Duplication	HGT	COG1074	ATP-dependent exoDNAse (exonuclease V) beta subunit (contains helicase and exonuclease domains)	L	Replication, recombination and repair
fam1_gq	VC_2322	Duplication	HGT	COG1330	Exonuclease V gamma subunit	L	Replication, recombination and repair
fam2_nu	VC_2323	No duplication	HGT	COG1275	Tellurite resistance protein and related permeases	P	Inorganic ion transport and metabolism
fam2_nt	VC_2324	No duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam2_ns	VC_2326	No duplication	HGT	COG3141	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_nq	VC_2329	No duplication	HGT	COG2171	Tetrahydrodipicolinate N-succinyltransferase	E	Amino acid transport and metabolism
fam1_gp	VC_2330	Duplication	HGT	COG3213	Uncharacterized protein involved in response to NO	P	Inorganic ion transport and metabolism
fam2_np	VC_2332	No duplication	HGT	COG0456	Acetyltransferases	R	General function prediction only
fam2_no	VC_2333	No duplication	HGT	COG0189	Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase)	HJ	Multiple classes
fam2_nn	VC_2334	No duplication	HGT	COG0598	Mg2+ and Co2+ transporters	P	Inorganic ion transport and metabolism
fam2_nm	VC_2335	No duplication	HGT	No hit			
fam2_nl	VC_2336	No duplication	HGT	COG0073	EMAP domain	R	General function prediction only
fam2_nk	VC_2337	No duplication	HGT	COG1609	Transcriptional regulators	K	Transcription
fam2_ni	VC_2339	No duplication	No	COG2059	Chromate transport protein ChrA	P	Inorganic ion transport and metabolism
fam2_ng	VC_2340	No duplication	No	COG2206	HD-GYP domain	T	Signal transduction mechanisms
fam2_nf	VC_2341	No duplication	No	COG1022	Long-chain acyl-CoA synthetases (AMP-forming)	I	Lipid transport and metabolism
fam1_av	VC_2342	Duplication	HGT	COG0480	Translation elongation factors (GTPases)	J	Translation, ribosomal structure and biogenesis
fam1_go	VC_2344	Duplication	No	No hit			
fam2_ne	VC_2345	No duplication	No	COG0560	Phosphoserine phosphatase	E	Amino acid transport and metabolism
fam2_nd	VC_2347	No duplication	HGT	COG0813	Purine-nucleoside phosphorylase	F	Nucleotide transport and metabolism
fam2_nc	VC_2348	No duplication	HGT	COG1015	Phosphopentomutase	G	Carbohydrate transport and metabolism
fam2_nb	VC_2349	No duplication	HGT	COG0213	Thymidine phosphorylase	F	Nucleotide transport and metabolism

fam2_na	VC_2350	No duplication	HGT	COG0274	Deoxyribose-phosphate aldolase	F	Nucleotide transport and metabolism
fam2_mz	VC_2352	No duplication	No	COG1972	Nucleoside permease	F	Nucleotide transport and metabolism
fam2_mx	VC_2357	No duplication	No	No hit			
fam2_mw	VC_2358	No duplication	No	COG3945	Uncharacterized conserved protein	S	Function unknown
fam2_mu	VC_2360	No duplication	HGT	COG0648	Endonuclease IV	L	Replication, recombination and repair
fam2_mt	VC_2363	No duplication	HGT	COG0083	Homoserine kinase	E	Amino acid transport and metabolism
fam1_gn	VC_2364*	Duplication	HGT	COG0527	Aspartokinases	E	Amino acid transport and metabolism
fam1_gn	VC_2364*	Duplication	HGT	COG0460	Homoserine dehydrogenase	E	Amino acid transport and metabolism
fam2_ms	VC_2368	No duplication	No	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam1_lx	VC_2369*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_lx	VC_2369*	Duplication	No	COG2198	FOG: HPt domain	T	Signal transduction mechanisms
fam1_lx	VC_2369*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam1_lx	VC_2369*	Duplication	No	COG2202	FOG: PAS/PAC domain	T	Signal transduction mechanisms
fam2_mr	VC_2370	No duplication	No	COG3706	Response regulator containing a CheY-like receiver domain and a GGDEF domain	T	Signal transduction mechanisms
fam1_gl	VC_2371	Duplication	No	COG1242	Predicted Fe-S oxidoreductase	R	General function prediction only
fam1_ag	VC_2373*	Duplication	HGT	COG0069	Glutamate synthase domain 2	E	Amino acid transport and metabolism
fam1_ag	VC_2373*	Duplication	HGT	COG0067	Glutamate synthase domain 1	E	Amino acid transport and metabolism
fam1_ag	VC_2373*	Duplication	HGT	COG0070	Glutamate synthase domain 3	E	Amino acid transport and metabolism
fam2_mq	VC_2374	No duplication	No	COG0493	NADPH-dependent glutamate synthase beta chain and related oxidoreductases	ER	Multiple classes
fam1_ag	VC_2376*	Duplication	HGT	COG0069	Glutamate synthase domain 2	E	Amino acid transport and metabolism
fam1_ag	VC_2376*	Duplication	HGT	COG0067	Glutamate synthase domain 1	E	Amino acid transport and metabolism
fam1_ag	VC_2376*	Duplication	HGT	COG0070	Glutamate synthase domain 3	E	Amino acid transport and metabolism
fam2_mp	VC_2377	No duplication	No	COG0493	NADPH-dependent glutamate synthase beta chain and related oxidoreductases	ER	Multiple classes
fam2_mo	VC_2378	No duplication	No	COG4446	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_mn	VC_2379	No duplication	No	COG0775	Nucleoside phosphorylase	F	Nucleotide transport and metabolism
fam2_mm	VC_2380	No duplication	No	COG1270	Cobalamin biosynthesis protein CobD/CbiB	H	Coenzyme transport and metabolism
fam2_ml	VC_2381	No duplication	No	COG0614	ABC-type Fe3+-hydroxamate transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam2_mk	VC_2382	No duplication	No	COG2860	Predicted membrane protein	S	Function unknown
fam2_mj	VC_2384	No duplication	No	COG0730	Predicted permeases	R	General function prediction only
fam2_mi	VC_2389*	No duplication	HGT	COG0458	Carbamoylphosphate synthase large subunit (split gene in MJ)	EF	Multiple classes
fam2_mi	VC_2389*	No duplication	HGT	COG0439	Biotin carboxylase	I	Lipid transport and metabolism
fam2_mi	VC_2389*	No duplication	HGT	COG0138	AICAR transformylase/IMP cyclohydrolase PurH (only IMP cyclohydrolase domain in Aful)	F	Nucleotide transport and metabolism
fam2_mh	VC_2391	No duplication	HGT	COG0289	Dihydrodipicolinate reductase	E	Amino acid transport and metabolism
fam2_mg	VC_2392	No duplication	No	COG1051	ADP-ribose pyrophosphatase	F	Nucleotide transport and metabolism
fam1_lw	VC_2394	Duplication	No	COG0653	Preprotein translocase subunit SecA (ATPase, RNA helicase)	U	Intracellular trafficking, secretion, and vesicular transport
fam2_mf	VC_2395	No duplication	HGT	COG4701	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_me	VC_2396	No duplication	HGT	COG0774	UDP-3-O-acyl-N-acetylglucosamine deacetylase	M	Cell wall/membrane/envelope biogenesis
fam2_md	VC_2398	No duplication	No	COG0849	Actin-like ATPase involved in cell division	D	Cell cycle control, cell division, chromosome partitioning
fam2_mc	VC_2399	No duplication	No	COG1589	Cell division septal protein	M	Cell wall/membrane/envelope biogenesis
fam1_lv	VC_2400	Duplication	No	COG0773	UDP-N-acetylmuramate-alanine ligase	M	Cell wall/membrane/envelope biogenesis
fam2_mb	VC_2401	No duplication	No	COG0707	UDP-N-acetylglucosamine:LPS N-acetylglucosamine transferase	M	Cell wall/membrane/envelope biogenesis
fam2_ma	VC_2403	No duplication	No	COG0771	UDP-N-acetylmuramoylalanine-D-glutamate ligase	M	Cell wall/membrane/envelope biogenesis
fam2_lz	VC_2404	No duplication	HGT	COG0472	UDP-N-acetylmuramyl pentapeptide phosphotransferase/UDP-N-acetylglucosamine-1-phosphate transferase	M	Cell wall/membrane/envelope biogenesis
fam2_ly	VC_2405	No duplication	HGT	COG0770	UDP-N-acetylmuramyl pentapeptide synthase	M	Cell wall/membrane/envelope biogenesis
fam2_lx	VC_2406	No duplication	HGT	COG0769	UDP-N-acetylmuramyl tripeptide synthase	M	Cell wall/membrane/envelope biogenesis
fam2_lw	VC_2407	No duplication	HGT	COG0768	Cell division protein FtsI/penicillin-binding protein 2	M	Cell wall/membrane/envelope biogenesis
fam2_lv	VC_2408	No duplication	HGT	COG3116	Cell division protein	D	Cell cycle control, cell division, chromosome partitioning
fam2_lu	VC_2409	No duplication	HGT	COG0275	Predicted S-adenosylmethionine-dependent methyltransferase involved in cell envelope biogenesis	M	Cell wall/membrane/envelope biogenesis
fam2_lt	VC_2412	No duplication	HGT	COG1249	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	C	Energy production and conversion
fam2_ls	VC_2413*	No duplication	HGT	COG0508	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	C	Energy production and conversion
fam2_ls	VC_2413*	No duplication	HGT	COG0511	Biotin carboxyl carrier protein	I	Lipid transport and metabolism
fam2_lr	VC_2414	No duplication	HGT	COG2609	Pyruvate dehydrogenase complex, dehydrogenase (E1) component	C	Energy production and conversion
fam2_lq	VC_2415	No duplication	No	COG2186	Transcriptional regulators	K	Transcription
fam2_lp	VC_2418	No duplication	No	COG1651	Protein-disulfide isomerase	O	Posttranslational modification, protein turnover, chaperones
fam2_lo	VC_2419	No duplication	No	COG4974	Site-specific recombinase XerD	L	Replication, recombination and repair
fam2_ln	VC_2420	No duplication	No	COG0716	Flavodoxins	C	Energy production and conversion
fam5_cs	VC_2422	No duplication	HGT	COG0157	Nicotinate-nucleotide pyrophosphorylase	H	Coenzyme transport and metabolism
fam5_cu	VC_2424	No duplication	No	COG2804	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	NU	Multiple classes
fam1_rs	VC_2425	No duplication	No	COG1459	Type II secretory pathway, component PulF	NU	Multiple classes
fam1_sc	VC_2426	No duplication	HGT	COG1989	Type II secretory pathway, prepilin signal peptidase PulO and related peptidases	NOU	Multiple classes
fam1_sj	VC_2427	No duplication	HGT	COG0237	Dephospho-CoA kinase	H	Coenzyme transport and metabolism
fam1_fx	VC_2438	Duplication	No	COG1391	Glutamine synthetase adenylyltransferase	OT	Multiple classes
fam2_bk	VC_2440	No duplication	No	COG3025	Uncharacterized conserved protein	S	Function unknown
fam2_br	VC_2441	No duplication	No	COG1392	Phosphate transport regulator (distant homolog of PhoU)	P	Inorganic ion transport and metabolism
fam2_ej	VC_2452	No duplication	No	COG2265	SAM-dependent methyltransferases related to tRNA (uracil-5-)-methyltransferase	J	Translation, ribosomal structure and biogenesis
fam2_eq	VC_2453*	No duplication	No	COG4999	Uncharacterized domain of BarA-like signal transduction histidine kinases	T	Signal transduction mechanisms
fam2_eq	VC_2453*	No duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam2_eq	VC_2453*	No duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms

fam2_eq	VC_2453*	No duplication	No	COG2198	FOG: HPt domain	T	Signal transduction mechanisms
fam1_ld	VC_2454	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_fe	VC_2455	No duplication	No	No hit			
fam1_di	VC_2456	Duplication	No	No hit			
fam2_fn	VC_2457	No duplication	No	COG0736	Phosphopantetheinyl transferase (holo-ACP synthase)	I	Lipid transport and metabolism
fam2_fp	VC_2458	No duplication	No	COG0854	Pyridoxal phosphate biosynthesis protein	H	Coenzyme transport and metabolism
fam2_fq	VC_2459	No duplication	HGT	COG1381	Recombinational DNA repair protein (RecF pathway)	L	Replication, recombination and repair
fam2_fr	VC_2460	No duplication	HGT	COG1159	GTPase	R	General function prediction only
fam2_fs	VC_2461	No duplication	HGT	COG0571	dsRNA-specific ribonuclease	K	Transcription
fam2_ft	VC_2462	No duplication	No	COG0681	Signal peptidase I	U	Intracellular trafficking, secretion, and vesicular transport
fam1_gc	VC_2463	Duplication	No	COG0481	Membrane GTPase LepA	M	Cell wall/membrane/envelope biogenesis
fam2_fu	VC_2464	No duplication	No	COG3086	Positive regulator of sigma E activity	T	Signal transduction mechanisms
fam2_ga	VC_2465	No duplication	No	COG3026	Negative regulator of sigma E activity	T	Signal transduction mechanisms
fam2_gc	VC_2466	No duplication	No	COG3073	Negative regulator of sigma E activity	T	Signal transduction mechanisms
fam2_hs	VC_2475	No duplication	No	COG0654	2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases	HC	Multiple classes
fam2_ih	VC_2479	No duplication	No	COG0212	5-formyltetrahydrofolate cyclo-ligase	H	Coenzyme transport and metabolism
fam2_ip	VC_2480	No duplication	No	COG0120	Ribose 5-phosphate isomerase	G	Carbohydrate transport and metabolism
fam2_is	VC_2481	No duplication	No	COG0111	Phosphoglycerate dehydrogenase and related dehydrogenases	HE	Multiple classes
fam2_it	VC_2482	No duplication	No	COG0440	Acetolactate synthase, small (regulatory) subunit	E	Amino acid transport and metabolism
fam2_iu	VC_2483	No duplication	No	COG0028	Thiamine pyrophosphate-requiring enzymes	EH	Multiple classes
fam2_iv	VC_2484	No duplication	No	COG1022	Long-chain acyl-CoA synthetases (AMP-forming)	I	Lipid transport and metabolism
fam2_iw	VC_2485	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam2_iy	VC_2487	No duplication	No	COG1819	Glycosyl transferases, related to UDP-glucuronosyltransferase	GC	Multiple classes
fam2_iz	VC_2488	No duplication	No	COG0671	Membrane-associated phospholipid phosphatase	I	Lipid transport and metabolism
fam2_jj	VC_2490	No duplication	No	COG0119	Isopropylmalate/homocitrate/citramalate synthases	E	Amino acid transport and metabolism
fam2_jp	VC_2491	No duplication	No	COG0473	Isocitrate/isopropylmalate dehydrogenase	CE	Multiple classes
fam2_ju	VC_2492	No duplication	HGT	COG0065	3-isopropylmalate dehydratase large subunit	E	Amino acid transport and metabolism
fam2_jy	VC_2493	No duplication	HGT	COG0066	3-isopropylmalate dehydratase small subunit	E	Amino acid transport and metabolism
fam2_kb	VC_2494	No duplication	HGT	No hit			
fam5_ac	VC_2501	No duplication	HGT	COG0260	Leucyl aminopeptidase	E	Amino acid transport and metabolism
fam1_qo	VC_2504	Duplication	HGT	COG1052	Lactate dehydrogenase and related dehydrogenases	CHR	Multiple classes
fam5_aa	VC_2505	No duplication	HGT	COG0564	Pseudouridylate synthases, 23S RNA-specific	J	Translation, ribosomal structure and biogenesis
fam1_ix	VC_2506	Duplication	HGT	COG0553	Superfamily II DNA/RNA helicases, SNF2 family	KL	Multiple classes
fam4_tf	VC_2507	No duplication	HGT	COG1875	Predicted ATPase related to phosphate starvation-inducible protein PhoH	T	Signal transduction mechanisms
fam4_te	VC_2508	No duplication	HGT	COG0078	Ornithine carbamoyltransferase	E	Amino acid transport and metabolism
fam4_td	VC_2510	No duplication	HGT	COG0540	Aspartate carbamoyltransferase, catalytic chain	F	Nucleotide transport and metabolism
fam4_tc	VC_2511	No duplication	HGT	COG1781	Aspartate carbamoyltransferase, regulatory subunit	F	Nucleotide transport and metabolism
fam4_tb	VC_2512	No duplication	HGT	COG0251	Putative translation initiation inhibitor, yjgF family	J	Translation, ribosomal structure and biogenesis
fam4_ta	VC_2513	No duplication	HGT	COG0204	1-acyl-sn-glycerol-3-phosphate acyltransferase	I	Lipid transport and metabolism
fam4_sz	VC_2514	No duplication	HGT	COG0766	UDP-N-acetylglucosamine enolpyruvyl transferase	M	Cell wall/membrane/envelope biogenesis
fam4_sy	VC_2515	No duplication	HGT	COG5007	Predicted transcriptional regulator, BolA superfamily	K	Transcription
fam4_sx	VC_2516	No duplication	HGT	COG3113	Predicted NTP binding protein (contains STAS domain)	R	General function prediction only
fam4_sw	VC_2517	No duplication	HGT	COG2854	ABC-type transport system involved in resistance to organic solvents, auxiliary component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam4_sv	VC_2518	No duplication	HGT	COG1463	ABC-type transport system involved in resistance to organic solvents, periplasmic component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam4_su	VC_2519	No duplication	HGT	COG0767	ABC-type transport system involved in resistance to organic solvents, permease component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam4_st	VC_2520	No duplication	HGT	COG1127	ABC-type transport system involved in resistance to organic solvents, ATPase component	Q	Secondary metabolites biosynthesis, transport and catabolism
fam4_ss	VC_2522	No duplication	No	COG0530	Ca2+/Na+ antiporter	P	Inorganic ion transport and metabolism
fam4_sr	VC_2523*	No duplication	No	COG0794	Predicted sugar phosphate isomerase involved in capsule formation	M	Cell wall/membrane/envelope biogenesis
fam4_sr	VC_2523*	No duplication	No	COG2905	Predicted signal-transduction protein containing cAMP-binding and CBS domains	T	Signal transduction mechanisms
fam4_sq	VC_2524	No duplication	No	COG1778	Low specificity phosphatase (HAD superfamily)	R	General function prediction only
fam4_sp	VC_2525	No duplication	No	COG3117	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_so	VC_2527	No duplication	No	COG1934	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_sn	VC_2528	No duplication	No	COG1137	ABC-type (unclassified) transport system, ATPase component	R	General function prediction only
fam1_dz	VC_2529	Duplication	No	COG1508	DNA-directed RNA polymerase specialized sigma subunit, sigma54 homolog	K	Transcription
fam1_jl	VC_2530	Duplication	No	COG1544	Ribosome-associated protein Y (PSrp-1)	J	Translation, ribosomal structure and biogenesis
fam4_sm	VC_2531	No duplication	No	COG1762	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)	GT	Multiple classes
fam1_iw	VC_2532	Duplication	No	COG1660	Predicted P-loop-containing kinase	R	General function prediction only
fam4_sl	VC_2534	No duplication	HGT	COG2239	Mg/Co/Ni transporter MgtE (contains CBS domain)	P	Inorganic ion transport and metabolism
fam4_sk	VC_2535	No duplication	HGT	COG0312	Predicted Zn-dependent proteases and their inactivated homologs	R	General function prediction only
fam4_sj	VC_2536	No duplication	HGT	COG3028	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_si	VC_2537	No duplication	HGT	COG3840	ABC-type thiamine transport system, ATPase component	H	Coenzyme transport and metabolism
fam4_sh	VC_2538	No duplication	HGT	COG1178	ABC-type Fe3+ transport system, permease component	P	Inorganic ion transport and metabolism
fam4_sg	VC_2539	No duplication	HGT	COG4143	ABC-type thiamine transport system, periplasmic component	H	Coenzyme transport and metabolism
fam4_sf	VC_2541	No duplication	HGT	COG0163	3-polyprenyl-4-hydroxybenzoate decarboxylase	H	Coenzyme transport and metabolism
fam4_se	VC_2542	No duplication	HGT	COG0773	UDP-N-acetylmuramate-alanine ligase	M	Cell wall/membrane/envelope biogenesis
fam4_sd	VC_2544	No duplication	No	COG0158	Fructose-1,6-bisphosphatase	G	Carbohydrate transport and metabolism
fam4_sc	VC_2545	No duplication	HGT	COG0221	Inorganic pyrophosphatase	C	Energy production and conversion
fam4_sb	VC_2546	No duplication	HGT	COG2105	Uncharacterized conserved protein	S	Function unknown

fam1_iv	VC_2547	Duplication	No	COG2911	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_qm	VC_2548	Duplication	No	COG0729	Outer membrane protein	M	Cell wall/membrane/envelope biogenesis
fam4_sa	VC_2549	No duplication	No	COG0225	Peptide methionine sulfoxide reductase	O	Posttranslational modification, protein turnover, chaperones
fam4_rz	VC_2550	No duplication	No	COG3054	Predicted transcriptional regulator	R	General function prediction only
fam4_ry	VC_2551	No duplication	No	No hit			
fam4_rx	VC_2552	No duplication	No	No hit			
fam4_rw	VC_2553	No duplication	No	COG1136	ABC-type antimicrobial peptide transport system, ATPase component	V	Defense mechanisms
fam4_rv	VC_2554	No duplication	No	COG4591	ABC-type transport system, involved in lipoprotein release, permease component	M	Cell wall/membrane/envelope biogenesis
fam4_ru	VC_2555	No duplication	No	COG3495	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ql	VC_2556	Duplication	No	No hit			
fam1_ql	VC_2557	Duplication	No	No hit			
fam4_rt	VC_2558	No duplication	No	COG0529	Adenylylsulfate kinase and related kinases	P	Inorganic ion transport and metabolism
fam4_rs	VC_2559	No duplication	No	COG2895	GTPases - Sulfate adenylate transferase subunit 1	P	Inorganic ion transport and metabolism
fam4_rr	VC_2560	No duplication	No	COG0175	3'-phosphoadenosine 5'-phosphosulfate sulfotransferase (PAPS reductase)/FAD synthetase and related enzymes	EH	Multiple classes
fam4_rq	VC_2561	No duplication	No	COG0007	Uroporphyrinogen-III methylase	H	Coenzyme transport and metabolism
fam4_rp	VC_2562	No duplication	No	COG0737	5'-nucleotidase/2',3'-cyclic phosphodiesterase and related esterases	F	Nucleotide transport and metabolism
fam4_ro	VC_2563	No duplication	No	COG0664	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	T	Signal transduction mechanisms
fam4_rn	VC_2564	No duplication	No	COG0513	Superfamily II DNA and RNA helicases	LKJ	Multiple classes
fam4_rm	VC_2565	No duplication	No	COG2153	Predicted acyltransferase	R	General function prediction only
fam1_bi	VC_2599	Duplication	HGT	COG0557	Exoribonuclease R	K	Transcription
fam4_ri	VC_2603	No duplication	No	COG1446	Asparaginase	E	Amino acid transport and metabolism
fam1_qj	VC_2604	Duplication	No	COG1047	FKBP-type peptidyl-prolyl cis-trans isomerases 2	O	Posttranslational modification, protein turnover, chaperones
fam1_qi	VC_2606*	Duplication	No	COG0475	Kef-type K+ transport systems, membrane components	P	Inorganic ion transport and metabolism
fam1_qi	VC_2606*	Duplication	No	COG0569	K+ transport systems, NAD-binding component	P	Inorganic ion transport and metabolism
fam4_rh	VC_2608	No duplication	HGT	COG0488	ATPase components of ABC transporters with duplicated ATPase domains	R	General function prediction only
fam4_rg	VC_2609	No duplication	No	COG5589	Uncharacterized conserved protein	S	Function unknown
fam4_rf	VC_2610	No duplication	HGT	COG0429	Predicted hydrolase of the alpha/beta-hydrolase fold	R	General function prediction only
fam4_re	VC_2612	No duplication	HGT	COG3089	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_rc	VC_2614	No duplication	HGT	COG0664	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	T	Signal transduction mechanisms
fam4_rb	VC_2615	No duplication	No	No hit			
fam4_ra	VC_2616	No duplication	HGT	COG1012	NAD-dependent aldehyde dehydrogenases	C	Energy production and conversion
fam4_qz	VC_2617	No duplication	No	COG3138	Arginine/ornithine N-succinyltransferase beta subunit	E	Amino acid transport and metabolism
fam4_qy	VC_2618	No duplication	No	COG4992	Ornithine/acetylornithine aminotransferase	E	Amino acid transport and metabolism
fam4_qx	VC_2619	No duplication	No	COG0512	Anthranilate/para-aminobenzoate synthases component II	EH	Multiple classes
fam4_qw	VC_2620	No duplication	No	COG0621	2-methylthioadenine synthetase	J	Translation, ribosomal structure and biogenesis
fam4_qv	VC_2621	No duplication	No	COG2374	Predicted extracellular nuclease	R	General function prediction only
fam1_el	VC_2622	Duplication	No	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam4_qt	VC_2623	No duplication	No	COG0180	Tryptophanyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam4_qs	VC_2624	No duplication	No	COG0546	Predicted phosphatases	R	General function prediction only
fam4_qr	VC_2625	No duplication	No	COG0036	Pentose-5-phosphate-3-epimerase	G	Carbohydrate transport and metabolism
fam1_iu	VC_2626	Duplication	No	COG0338	Site-specific DNA methylase	L	Replication, recombination and repair
fam1_cr	VC_2627*	Duplication	No	COG3267	Type II secretory pathway, component ExeA (predicted ATPase)	U	Intracellular trafficking, secretion, and vesicular transport
fam1_cr	VC_2627*	Duplication	No	COG3266	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_qq	VC_2630	No duplication	No	COG4796	Type II secretory pathway, component HofQ	U	Intracellular trafficking, secretion, and vesicular transport
fam4_qp	VC_2632	No duplication	No	COG3167	Tfp pilus assembly protein PilO	NU	Multiple classes
fam4_qo	VC_2633	No duplication	No	COG3166	Tfp pilus assembly protein PilN	NU	Multiple classes
fam1_cd	VC_2634	Duplication	No	COG4972	Tfp pilus assembly protein, ATPase PilM	NU	Multiple classes
fam4_qn	VC_2635	No duplication	No	COG5009	Membrane carboxypeptidase/penicillin-binding protein	M	Cell wall/membrane/envelope biogenesis
fam4_qm	VC_2636	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam4_ql	VC_2637*	No duplication	No	COG0678	Peroxiredoxin	O	Posttranslational modification, protein turnover, chaperones
fam4_ql	VC_2637*	No duplication	No	COG0695	Glutaredoxin and related proteins	O	Posttranslational modification, protein turnover, chaperones
fam4_qk	VC_2638	No duplication	HGT	COG1249	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	C	Energy production and conversion
fam4_qj	VC_2640	No duplication	HGT	No hit			
fam4_qi	VC_2641	No duplication	HGT	COG0165	Argininosuccinate lyase	E	Amino acid transport and metabolism
fam4_qh	VC_2642	No duplication	HGT	COG0137	Argininosuccinate synthase	E	Amino acid transport and metabolism
fam4_qg	VC_2644	No duplication	HGT	COG0002	Acetylglutamate semialdehyde dehydrogenase	E	Amino acid transport and metabolism
fam4_qf	VC_2645	No duplication	HGT	COG0624	Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases	E	Amino acid transport and metabolism
fam1_qh	VC_2646	Duplication	No	COG2352	Phosphoenolpyruvate carboxylase	C	Energy production and conversion
fam4_qe	VC_2647	No duplication	No	COG1695	Predicted transcriptional regulators	K	Transcription
fam4_qd	VC_2651	No duplication	No	COG0240	Glycerol-3-phosphate dehydrogenase	C	Energy production and conversion
fam4_qc	VC_2653	No duplication	No	COG1952	Preprotein translocase subunit SecB	U	Intracellular trafficking, secretion, and vesicular transport
fam4_qb	VC_2654	No duplication	No	COG0607	Rhodanese-related sulfurtransferase	P	Inorganic ion transport and metabolism
fam1_kj	VC_2655	Duplication	No	COG2269	Truncated, possibly inactive, lysyl-tRNA synthetase (class II)	J	Translation, ribosomal structure and biogenesis
fam4_qa	VC_2657	No duplication	No	COG0479	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	C	Energy production and conversion
fam4_pz	VC_2658	No duplication	No	COG3029	Fumarate reductase subunit C	C	Energy production and conversion
fam4_py	VC_2659	No duplication	No	COG3080	Fumarate reductase subunit D	C	Energy production and conversion
fam4_px	VC_2660	No duplication	No	COG0231	Translation elongation factor P (EF-P)/translation initiation factor 5A (eIF-5A)	J	Translation, ribosomal structure and biogenesis
fam4_pw	VC_2661	No duplication	No	COG1509	Lysine 2,3-aminomutase	E	Amino acid transport and metabolism

fam4_pv	VC_2662	No duplication	HGT	No hit			
fam1_bc	VC_2664	Duplication	HGT	COG0459	Chaperonin GroEL (HSP60 family)	O	Posttranslational modification, protein turnover, chaperones
fam4_pu	VC_2665	No duplication	No	COG0234	Co-chaperonin GroES (HSP10)	O	Posttranslational modification, protein turnover, chaperones
fam4_pt	VC_2666	No duplication	No	COG2345	Predicted transcriptional regulator	K	Transcription
fam4_ps	VC_2667	No duplication	No	No hit			
fam4_pr	VC_2668	No duplication	No	COG3152	Predicted membrane protein	S	Function unknown
fam4_pq	VC_2669	No duplication	No	COG3232	5-carboxymethyl-2-hydroxymuconate isomerase	E	Amino acid transport and metabolism
fam4_pp	VC_2670	No duplication	No	COG0149	Triosephosphate isomerase	G	Carbohydrate transport and metabolism
fam4_po	VC_2671	No duplication	No	COG0534	Na+-driven multidrug efflux pump	V	Defense mechanisms
fam4_pn	VC_2672	No duplication	No	COG0684	Demethylmenaquinone methyltransferase	H	Coenzyme transport and metabolism
fam1_it	VC_2673	Duplication	No	COG1575	1,4-dihydroxy-2-naphthoate octaprenyltransferase	H	Coenzyme transport and metabolism
fam4_pm	VC_2674	No duplication	No	COG1220	ATP-dependent protease HslVU (ClpYQ), ATPase subunit	O	Posttranslational modification, protein turnover, chaperones
fam4_pl	VC_2675	No duplication	No	COG5405	ATP-dependent protease HslVU (ClpYQ), peptidase subunit	O	Posttranslational modification, protein turnover, chaperones
fam4_pk	VC_2676	No duplication	No	COG3087	Cell division protein	D	Cell cycle control, cell division, chromosome partitioning
fam1_qg	VC_2678	Duplication	No	COG1198	Primosomal protein N' (replication factor Y) - superfamily II helicase	L	Replication, recombination and repair
fam4_pj	VC_2681	No duplication	No	COG0281	Malic enzyme	C	Energy production and conversion
fam4_pi	VC_2682	No duplication	No	COG3060	Transcriptional regulator of met regulon	KE	Multiple classes
fam4_ph	VC_2683	No duplication	HGT	COG0626	Cystathionine beta-lyases/cystathionine gamma-synthases	E	Amino acid transport and metabolism
fam4_pf	VC_2685	No duplication	No	COG0685	5,10-methylenetetrahydrofolate reductase	E	Amino acid transport and metabolism
fam4_pe	VC_2686	No duplication	HGT	COG3074	Uncharacterized protein conserved in bacteria	S	Function unknown
fam4_pd	VC_2688	No duplication	No	COG1494	Fructose-1,6-bisphosphatase/sedoheptulose 1,7-bisphosphatase and related proteins	G	Carbohydrate transport and metabolism
fam4_pc	VC_2689	No duplication	No	COG0205	6-phosphofructokinase	G	Carbohydrate transport and metabolism
fam4_pb	VC_2690	No duplication	No	COG0053	Predicted Co/Zn/Cd cation transporters	P	Inorganic ion transport and metabolism
fam4_pa	VC_2691	No duplication	No	COG3678	P pilus assembly/Cpx signaling pathway, periplasmic inhibitor/zinc-resistance associated protein	UNTP	Multiple classes
fam4_oz	VC_2693	No duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam4_oy	VC_2695	No duplication	No	COG0219	Predicted rRNA methylase (SpoU class)	J	Translation, ribosomal structure and biogenesis
fam4_ox	VC_2696	No duplication	No	COG3030	Protein affecting phage T7 exclusion by the F plasmid	R	General function prediction only
fam1_is	VC_2697	Duplication	HGT	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_ir	VC_2698	Duplication	No	COG1027	Aspartate ammonia-lyase	E	Amino acid transport and metabolism
fam4_ow	VC_2699	No duplication	No	COG2704	Anaerobic C4-dicarboxylate transporter	R	General function prediction only
fam4_ov	VC_2701	No duplication	No	COG4232	Thiol:disulfide interchange protein	OC	Multiple classes
fam4_ou	VC_2702	No duplication	HGT	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes
fam4_ot	VC_2704	No duplication	No	COG4327	Predicted membrane protein	S	Function unknown
fam4_os	VC_2705	No duplication	No	COG4147	Predicted symporter	R	General function prediction only
fam4_or	VC_2706	No duplication	No	COG1738	Uncharacterized conserved protein	S	Function unknown
fam4_oq	VC_2708	No duplication	No	COG0194	Guanylate kinase	F	Nucleotide transport and metabolism
fam4_op	VC_2709	No duplication	No	COG1758	DNA-directed RNA polymerase, subunit K/omega	K	Transcription
fam4_oo	VC_2710	No duplication	No	COG0317	Guanosine polyphosphate pyrophosphohydrolases/synthetases	TK	Multiple classes
fam4_on	VC_2711	No duplication	HGT	COG1200	RecG-like helicase	LK	Multiple classes
fam1_qf	VC_2712	Duplication	No	COG2233	Xanthine/uracil permeases	F	Nucleotide transport and metabolism
fam4_om	VC_2713	No duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam4_ol	VC_2714	No duplication	No	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam4_ok	VC_2715	No duplication	HGT	COG0782	Transcription elongation factor	K	Transcription
fam1_qe	VC_2716	Duplication	HGT	COG2183	Transcriptional accessory protein	K	Transcription
fam4_oj	VC_2717	No duplication	HGT	No hit			
fam4_oi	VC_2718	No duplication	HGT	COG0596	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	R	General function prediction only
fam4_oh	VC_2719	No duplication	HGT	COG1040	Predicted amidophosphoribosyltransferases	R	General function prediction only
fam4_og	VC_2720*	No duplication	HGT	COG0316	Uncharacterized conserved protein	S	Function unknown
fam4_og	VC_2720*	No duplication	HGT	COG0694	Thioredoxin-like proteins and domains	O	Posttranslational modification, protein turnover, chaperones
fam4_of	VC_2721	No duplication	HGT	COG0494	NTP pyrophosphohydrolases including oxidative damage repair enzymes	LR	Multiple classes
fam4_oe	VC_2722	No duplication	HGT	COG1218	3'-Phosphoadenosine 5'-phosphosulfate (PAPS) 3'-phosphatase	P	Inorganic ion transport and metabolism
fam4_od	VC_2723	No duplication	HGT	No hit			
fam4_ob	VC_2724	No duplication	No	COG3149	Type II secretory pathway, component PulM	U	Intracellular trafficking, secretion, and vesicular transport
fam1_ck	VC_2725	Duplication	HGT	COG3297	Type II secretory pathway, component PulL	U	Intracellular trafficking, secretion, and vesicular transport
fam1_ey	VC_2726	Duplication	No	COG3156	Type II secretory pathway, component PulK	U	Intracellular trafficking, secretion, and vesicular transport
fam1_cx	VC_2727	Duplication	No	COG4795	Type II secretory pathway, component PulJ	U	Intracellular trafficking, secretion, and vesicular transport
fam1_ez	VC_2728	Duplication	No	COG2165	Type II secretory pathway, pseudopilin PulG	NU	Multiple classes
fam1_fa	VC_2729	Duplication	No	COG2165	Type II secretory pathway, pseudopilin PulG	NU	Multiple classes
fam1_fb	VC_2730	Duplication	No	COG2165	Type II secretory pathway, pseudopilin PulG	NU	Multiple classes
fam1_fc	VC_2731	Duplication	No	COG1459	Type II secretory pathway, component PulF	NU	Multiple classes
fam1_cy	VC_2732	Duplication	No	COG2804	Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	NU	Multiple classes
fam4_oa	VC_2734	No duplication	No	COG3031	Type II secretory pathway, component PulC	U	Intracellular trafficking, secretion, and vesicular transport
fam4_nz	VC_2735	No duplication	No	COG1188	Ribosome-associated heat shock protein implicated in the recycling of the 50S subunit (S4 paralog)	J	Translation, ribosomal structure and biogenesis
fam4_ny	VC_2736	No duplication	No	COG1281	Disulfide bond chaperones of the HSP33 family	O	Posttranslational modification, protein turnover, chaperones
fam1_qd	VC_2738	Duplication	No	COG1866	Phosphoenolpyruvate carboxykinase (ATP)	C	Energy production and conversion
fam4_nx	VC_2739	No duplication	No	No hit			
fam4_nw	VC_2740	No duplication	No	COG2153	Predicted acyltransferase	R	General function prediction only
fam4_nv	VC_2742	No duplication	No	COG1295	Predicted membrane protein	S	Function unknown

fam4_nu	VC_2743	No duplication	No	COG3911	Predicted ATPase	R	General function prediction only
fam4_nt	VC_2744	No duplication	No	COG1217	Predicted membrane GTPase involved in stress response	T	Signal transduction mechanisms
fam4_nr	VC_2746	No duplication	No	COG0174	Glutamine synthetase	E	Amino acid transport and metabolism
fam1_qc	VC_2747	Duplication	No	No hit			
fam4_nq	VC_2748	No duplication	No	COG3852	Signal transduction histidine kinase, nitrogen specific	T	Signal transduction mechanisms
fam1_jq	VC_2749	Duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam1_dy	VC_2750*	Duplication	No	COG5001	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	T	Signal transduction mechanisms
fam1_dy	VC_2750*	Duplication	No	COG2770	FOG: HAMP domain	T	Signal transduction mechanisms
fam4_gg	VC_2757	No duplication	HGT	COG1739	Uncharacterized conserved protein	S	Function unknown
fam1_pm	VC_2758*	Duplication	HGT	COG1250	3-hydroxyacyl-CoA dehydrogenase	I	Lipid transport and metabolism
fam1_pm	VC_2758*	Duplication	HGT	COG1024	Enoyl-CoA hydratase/carnithine racemase	I	Lipid transport and metabolism
fam4_gf	VC_2759	No duplication	HGT	COG0183	Acetyl-CoA acetyltransferase	I	Lipid transport and metabolism
fam4_ge	VC_2760	No duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam4_gd	VC_2761	No duplication	HGT	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam4_gc	VC_2762	No duplication	No	COG1207	N-acetylglucosamine-1-phosphate uridyltransferase (contains nucleotidyltransferase and I-patch acetyltransferase domains)	M	Cell wall/membrane/envelope biogenesis
fam4_gb	VC_2763	No duplication	No	COG0355	F0F1-type ATP synthase, epsilon subunit (mitochondrial delta subunit)	C	Energy production and conversion
fam1_pl	VC_2764	Duplication	No	COG0055	F0F1-type ATP synthase, beta subunit	C	Energy production and conversion
fam4_ga	VC_2765	No duplication	No	COG0224	F0F1-type ATP synthase, gamma subunit	C	Energy production and conversion
fam4_fz	VC_2766	No duplication	No	COG0056	F0F1-type ATP synthase, alpha subunit	C	Energy production and conversion
fam4_fy	VC_2767	No duplication	No	COG0712	F0F1-type ATP synthase, delta subunit (mitochondrial oligomycin sensitivity protein)	C	Energy production and conversion
fam4_fx	VC_2768	No duplication	No	COG0711	F0F1-type ATP synthase, subunit b	C	Energy production and conversion
fam4_fw	VC_2769	No duplication	No	COG0636	F0F1-type ATP synthase, subunit c/Archaeal/vacuolar-type H+-ATPase, subunit K	C	Energy production and conversion
fam1_pk	VC_2770	Duplication	No	COG0356	F0F1-type ATP synthase, subunit a	C	Energy production and conversion
fam4_fv	VC_2771	No duplication	No	COG3312	F0F1-type ATP synthase, subunit I	C	Energy production and conversion
fam4_fs	VC_2774	No duplication	No	COG0357	Predicted S-adenosylmethionine-dependent methyltransferase involved in bacterial cell division	M	Cell wall/membrane/envelope biogenesis
fam2_an	VC_A0002	No duplication	No	No hit			
fam2_am	VC_A0003	No duplication	No	No hit			
fam2_ak	VC_A0004	No duplication	No	No hit			
fam2_aj	VC_A0005	No duplication	No	No hit			
fam2_ai	VC_A0006	No duplication	No	COG0217	Uncharacterized conserved protein	S	Function unknown
fam2_ah	VC_A0007	No duplication	No	COG2084	3-hydroxyisobutyrate dehydrogenase and related beta-hydroxyacid dehydrogenases	I	Lipid transport and metabolism
fam2_ag	VC_A0008	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_af	VC_A0010	No duplication	No	No hit			
fam1_kv	VC_A0011	Duplication	No	COG2909	ATP-dependent transcriptional regulator	K	Transcription
fam2_ae	VC_A0013	No duplication	No	COG0058	Glucan phosphorylase	G	Carbohydrate transport and metabolism
fam2_ad	VC_A0014	No duplication	No	COG1640	4-alpha-glucanotransferase	G	Carbohydrate transport and metabolism
fam2_ac	VC_A0016	No duplication	No	COG0296	1,4-alpha-glucan branching enzyme	G	Carbohydrate transport and metabolism
fam1_ab	VC_A0018	Duplication	HGT	COG3501	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_ab	VC_A0026	No duplication	No	COG3110	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ku	VC_A0027*	Duplication	No	COG3325	Chitinase	G	Carbohydrate transport and metabolism
fam1_ku	VC_A0027*	Duplication	No	COG3979	Uncharacterized protein contain chitin-binding domain type 3	R	General function prediction only
fam2_aa	VC_A0029	No duplication	HGT	COG1414	Transcriptional regulator	K	Transcription
fam1_tf	VC_A0030	No duplication	No	No hit			
fam1_fw	VC_A0033	Duplication	No	COG3211	Predicted phosphatase	R	General function prediction only
fam1_te	VC_A0035	No duplication	No	COG0671	Membrane-associated phospholipid phosphatase	I	Lipid transport and metabolism
fam1_fv	VC_A0036	Duplication	No	COG3633	Na+/serine symporter	E	Amino acid transport and metabolism
fam1_td	VC_A0040	No duplication	No	COG2035	Predicted membrane protein	S	Function unknown
fam1_tc	VC_A0041	No duplication	No	COG0786	Na+/glutamate symporter	E	Amino acid transport and metabolism
fam1_tb	VC_A0042	No duplication	No	No hit			
fam1_ta	VC_A0043	No duplication	No	COG2388	Predicted acetyltransferase	R	General function prediction only
fam1_cf	VC_A0044	Duplication	No	COG4946	Uncharacterized protein related to the periplasmic component of the Tol biopolymer transport system	S	Function unknown
fam1_cf	VC_A0045	Duplication	No	COG0793	Periplasmic protease	M	Cell wall/membrane/envelope biogenesis
fam1_fu	VC_A0049	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_fu	VC_A0050	Duplication	No	No hit			
fam1_sz	VC_A0051	No duplication	No	No hit			
fam1_sy	VC_A0056	No duplication	No	COG0789	Predicted transcriptional regulators	K	Transcription
fam1_sx	VC_A0057	No duplication	No	COG0415	Deoxyribodipyrimidine photolyase	L	Replication, recombination and repair
fam1_sw	VC_A0063	Duplication	No	COG1770	Protease II	E	Amino acid transport and metabolism
fam1_ft	VC_A0064	Duplication	No	COG1629	Outer membrane receptor proteins, mostly Fe transport	P	Inorganic ion transport and metabolism
fam5_dl	VC_A0065	No duplication	No	COG4935	Regulatory P domain of the subtilisin-like proprotein convertases and other proteases	O	Posttranslational modification, protein turnover, chaperones
fam5_dk	VC_A0066	No duplication	No	No hit			
fam5_dj	VC_A0067	No duplication	No	No hit			
fam1_su	VC_A0070	No duplication	No	COG0226	ABC-type phosphate transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam1_st	VC_A0071	No duplication	No	COG0573	ABC-type phosphate transport system, permease component	P	Inorganic ion transport and metabolism
fam1_ss	VC_A0072	No duplication	No	COG0581	ABC-type phosphate transport system, permease component	P	Inorganic ion transport and metabolism
fam1_sr	VC_A0073	No duplication	No	COG1117	ABC-type phosphate transport system, ATPase component	P	Inorganic ion transport and metabolism
fam1_kt	VC_A0074	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_sq	VC_A0075	No duplication	No	COG3915	Uncharacterized protein conserved in bacteria	S	Function unknown

fam1_sp	VC_A0077	No duplication	No	COG0659	Sulfate permease and related transporters (MFS superfamily)	P	Inorganic ion transport and metabolism
fam1_so	VC_A0078	No duplication	No	No hit			
fam1_sn	VC_A0079	No duplication	No	COG0739	Membrane proteins related to metalloendopeptidases	M	Cell wall/membrane/envelope biogenesis
fam1_kr	VC_A0080	Duplication	No	COG5001	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	T	Signal transduction mechanisms
fam1_sm	VC_A0082	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam1_sl	VC_A0083	No duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam1_sk	VC_A0087	No duplication	No	No hit			
fam1_kq	VC_A0098	Duplication	No	COG1488	Nicotinic acid phosphoribosyltransferase	H	Coenzyme transport and metabolism
fam1_kp	VC_A0099	Duplication	No	COG0673	Predicted dehydrogenases and related proteins	R	General function prediction only
fam1_si	VC_A0100	No duplication	No	COG0607	Rhodanese-related sulfurtransferase	P	Inorganic ion transport and metabolism
fam1_sh	VC_A0101	No duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam1_sg	VC_A0102	No duplication	No	COG0637	Predicted phosphatase/phosphohexomutase	R	General function prediction only
fam1_sf	VC_A0103	No duplication	No	COG0659	Sulfate permease and related transporters (MFS superfamily)	P	Inorganic ion transport and metabolism
fam1_ko	VC_A0104	Duplication	No	COG0564	Pseudouridylate synthases, 23S RNA-specific	J	Translation, ribosomal structure and biogenesis
fam1_kn	VC_A0105	Duplication	No	COG4104	Uncharacterized conserved protein	S	Function unknown
fam1_se	VC_A0106	No duplication	No	No hit			
fam1_sd	VC_A0107	No duplication	No	COG3516	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_fs	VC_A0108	Duplication	No	COG3517	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_sb	VC_A0109	No duplication	No	COG3518	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_sa	VC_A0110	No duplication	No	COG3519	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_rz	VC_A0111	No duplication	No	COG3520	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ry	VC_A0112	No duplication	No	COG3456	Uncharacterized conserved protein, contains FHA domain	T	Signal transduction mechanisms
fam1_rx	VC_A0113	No duplication	No	COG3521	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_rw	VC_A0114	No duplication	No	COG3522	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_rv	VC_A0115	No duplication	No	COG3455	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ru	VC_A0117	No duplication	No	COG3829	Transcriptional regulator containing PAS, AAA-type ATPase, and DNA-binding domains	KT	Multiple classes
fam1_rt	VC_A0118	No duplication	No	No hit			
fam5_cw	VC_A0119	No duplication	No	COG3515	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_rf	VC_A0120	Duplication	No	COG3523	Uncharacterized protein conserved in bacteria	S	Function unknown
fam5_cv	VC_A0121	No duplication	No	COG3515	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ab	VC_A0123*	Duplication	HGT	COG3501	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ab	VC_A0123*	Duplication	HGT	COG3409	Putative peptidoglycan-binding domain-containing protein	M	Cell wall/membrane/envelope biogenesis
fam1_ea	VC_A0161	Duplication	No	COG3033	Tryptophanase	E	Amino acid transport and metabolism
fam5_cr	VC_A0163	No duplication	No	COG3530	Uncharacterized protein conserved in bacteria	S	Function unknown
fam5_cq	VC_A0164	No duplication	No	COG3034	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_jh	VC_A0165	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam5_cp	VC_A0166	No duplication	No	COG1278	Cold shock proteins	K	Transcription
fam5_co	VC_A0167	No duplication	No	COG2859	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_jg	VC_A0168	Duplication	No	No hit			
fam1_jg	VC_A0169	Duplication	No	No hit			
fam1_jg	VC_A0170	Duplication	No	No hit			
fam1_ju	VC_A0171*	Duplication	No	COG2304	Uncharacterized protein containing a von Willebrand factor type A (vWA) domain	R	General function prediction only
fam1_ju	VC_A0171*	Duplication	No	COG5624	Transcription initiation factor TFIID, subunit TAF12 (also component of histone acetyltransferase SAGA)	K	Transcription
fam5_cn	VC_A0172	No duplication	No	No hit			
fam5_cm	VC_A0173	No duplication	No	No hit			
fam5_ck	VC_A0175	No duplication	No	COG0714	MoxR-like ATPases	R	General function prediction only
fam5_cj	VC_A0176	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam5_ci	VC_A0177	No duplication	No	No hit			
fam1_rd	VC_A0178	Duplication	No	COG2761	Predicted dithiol-disulfide isomerase involved in polyketide biosynthesis	Q	Secondary metabolites biosynthesis, transport and catabolism
fam5_ch	VC_A0179	No duplication	No	COG1972	Nucleoside permease	F	Nucleotide transport and metabolism
fam1_rc	VC_A0180	Duplication	No	COG2195	Di- and tripeptidases	E	Amino acid transport and metabolism
fam5_cg	VC_A0181	No duplication	No	COG0668	Small-conductance mechanosensitive channel	M	Cell wall/membrane/envelope biogenesis
fam1_rb	VC_A0182	Duplication	No	COG3604	Transcriptional regulator containing GAF, AAA-type ATPase, and DNA binding domains	KT	Multiple classes
fam5_cf	VC_A0183*	No duplication	No	COG1017	Hemoglobin-like flavoprotein	C	Energy production and conversion
fam5_cf	VC_A0183*	No duplication	No	COG1018	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1	C	Energy production and conversion
fam5_ce	VC_A0186	No duplication	No	No hit			
fam5_cd	VC_A0189*	No duplication	No	COG1776	Chemotaxis protein CheC, inhibitor of MCP methylation	NT	Multiple classes
fam5_cd	VC_A0189*	No duplication	No	COG3706	Response regulator containing a CheY-like receiver domain and a GGDEF domain	T	Signal transduction mechanisms
fam5_cc	VC_A0190	No duplication	No	No hit			
fam5_cb	VC_A0191*	No duplication	No	COG0063	Predicted sugar kinase	G	Carbohydrate transport and metabolism
fam5_cb	VC_A0191*	No duplication	No	COG0062	Uncharacterized conserved protein	S	Function unknown
fam5_ca	VC_A0192	No duplication	No	COG1052	Lactate dehydrogenase and related dehydrogenases	CHR	Multiple classes
fam5_bz	VC_A0193	No duplication	No	COG1757	Na+/H+ antiporter	C	Energy production and conversion
fam1_ra	VC_A0194	Duplication	No	COG1226	Kef-type K+ transport systems, predicted NAD-binding component	P	Inorganic ion transport and metabolism
fam5_by	VC_A0195	No duplication	No	No hit			
fam5_bx	VC_A0204	No duplication	HGT	COG0513	Superfamily II DNA and RNA helicases	LKJ	Multiple classes
fam5_bw	VC_A0205	No duplication	HGT	COG2704	Anaerobic C4-dicarboxylate transporter	R	General function prediction only
fam5_bv	VC_A0207	No duplication	HGT	COG0171	NAD synthase	H	Coenzyme transport and metabolism

fam1_qy	VC_A0208	Duplication	No	COG1057	Nicotinic acid mononucleotide adenylyltransferase	H	Coenzyme transport and metabolism
fam5_bu	VC_A0210	No duplication	No	COG3437	Response regulator containing a CheY-like receiver domain and an HD-GYP domain	KT	Multiple classes
fam5_bt	VC_A0211	No duplication	No	COG4191	Signal transduction histidine kinase regulating C4-dicarboxylate transport system	T	Signal transduction mechanisms
fam1_aj	VC_A0213	Duplication	HGT	COG1757	Na+/H+ antiporter	C	Energy production and conversion
fam5_br	VC_A0214	No duplication	No	COG2814	Arabinose efflux permease	G	Carbohydrate transport and metabolism
fam1_qx	VC_A0218	Duplication	No	COG3240	Phospholipase/lecithinase/hemolysin	IR	Multiple classes
fam1_em	VC_A0240	Duplication	HGT	No hit			
fam1_en	VC_A0241	Duplication	HGT	COG3623	Putative L-xylulose-5-phosphate 3-epimerase	G	Carbohydrate transport and metabolism
fam1_eo	VC_A0242	Duplication	HGT	COG0269	3-hexulose-6-phosphate synthase and related proteins	G	Carbohydrate transport and metabolism
fam1_ep	VC_A0243	Duplication	HGT	COG0561	Predicted hydrolases of the HAD superfamily	R	General function prediction only
fam1_eq	VC_A0244	Duplication	HGT	COG0235	Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	G	Carbohydrate transport and metabolism
fam1_er	VC_A0245	Duplication	HGT	COG1762	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)	GT	Multiple classes
fam1_es	VC_A0246*	Duplication	HGT	COG3037	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_es	VC_A0246*	Duplication	HGT	COG3414	Phosphotransferase system, galactitol-specific IIB component	G	Carbohydrate transport and metabolism
fam1_et	VC_A0247	Duplication	HGT	COG1349	Transcriptional regulators of sugar metabolism	KG	Multiple classes
fam1_eu	VC_A0248	Duplication	HGT	COG2220	Predicted Zn-dependent hydrolases of the beta-lactamase fold	R	General function prediction only
fam1_kd	VC_A0249	Duplication	HGT	COG3038	Cytochrome B561	C	Energy production and conversion
fam1_bv	VC_A0254	Duplication	HGT	No hit			
fam5_bp	VC_A0265	No duplication	HGT	No hit			
fam5_bo	VC_A0269	Duplication	No	COG0076	Glutamate decarboxylase and related PLP-dependent proteins	E	Amino acid transport and metabolism
fam5_bn	VC_A0271	No duplication	No	No hit			
fam1_jf	VC_A0274	Duplication	No	COG3338	Carbonic anhydrase	P	Inorganic ion transport and metabolism
fam5_bl	VC_A0277	No duplication	HGT	COG0509	Glycine cleavage system H protein (lipoate-binding)	E	Amino acid transport and metabolism
fam1_je	VC_A0279	Duplication	HGT	COG1396	Predicted transcriptional regulators	K	Transcription
fam1_kn	VC_A0284	Duplication	No	COG4104	Uncharacterized conserved protein	S	Function unknown
fam1_jd	VC_A0287	Duplication	HGT	COG0441	Threonyl-tRNA synthetase	J	Translation, ribosomal structure and biogenesis
fam1_jb	VC_A0510	Duplication	No	COG1524	Uncharacterized proteins of the AP superfamily	R	General function prediction only
fam1_ja	VC_A0511	Duplication	No	COG1328	Oxygen-sensitive ribonucleoside-triphosphate reductase	F	Nucleotide transport and metabolism
fam5_be	VC_A0512	No duplication	No	COG0602	Organic radical activating enzymes	O	Posttranslational modification, protein turnover, chaperones
fam5_bd	VC_A0514	No duplication	No	COG4067	Uncharacterized protein conserved in archaea	O	Posttranslational modification, protein turnover, chaperones
fam1_jp	VC_A0516*	Duplication	No	COG1299	Phosphotransferase system, fructose-specific IIC component	G	Carbohydrate transport and metabolism
fam1_jp	VC_A0516*	Duplication	No	COG1445	Phosphotransferase system fructose-specific component IIB	G	Carbohydrate transport and metabolism
fam5_bc	VC_A0517	No duplication	No	COG1105	Fructose-1-phosphate kinase and related fructose-6-phosphate kinase (PfkB)	G	Carbohydrate transport and metabolism
fam5_bb	VC_A0518*	No duplication	No	COG4668	Mannitol/fructose-specific phosphotransferase system, IIA domain	G	Carbohydrate transport and metabolism
fam5_bb	VC_A0518*	No duplication	No	COG1925	Phosphotransferase system, HPr-related proteins	G	Carbohydrate transport and metabolism
fam5_ba	VC_A0519	No duplication	No	COG1609	Transcriptional regulators	K	Transcription
fam5_ay	VC_A0520	No duplication	No	COG0420	DNA repair exonuclease	L	Replication, recombination and repair
fam1_qt	VC_A0521	Duplication	No	COG0419	ATPase involved in DNA repair	L	Replication, recombination and repair
fam1_iz	VC_A0522*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_iz	VC_A0522*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam5_ax	VC_A0523	No duplication	No	COG0156	7-keto-8-aminopelargonate synthetase and related enzymes	H	Coenzyme transport and metabolism
fam5_aw	VC_A0524	No duplication	No	COG1275	Tellurite resistance protein and related permeases	P	Inorganic ion transport and metabolism
fam5_av	VC_A0526	No duplication	No	COG0038	Chloride channel protein EriC	P	Inorganic ion transport and metabolism
fam5_au	VC_A0527	No duplication	HGT	COG0523	Putative GTPases (G3E family)	R	General function prediction only
fam1_qs	VC_A0528	Duplication	HGT	COG1161	Predicted GTPases	R	General function prediction only
fam5_ar	VC_A0546	No duplication	No	COG0861	Membrane protein TerC, possibly involved in tellurium resistance	P	Inorganic ion transport and metabolism
fam5_aq	VC_A0554	No duplication	No	COG2223	Nitrate/nitrite transporter	P	Inorganic ion transport and metabolism
fam5_ap	VC_A0556	No duplication	No	No hit			
fam5_ao	VC_A0558	No duplication	No	COG0405	Gamma-glutamyltransferase	E	Amino acid transport and metabolism
fam5_an	VC_A0559	No duplication	No	COG4206	Outer membrane cobalamin receptor protein	H	Coenzyme transport and metabolism
fam5_am	VC_A0560	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam5_al	VC_A0561	No duplication	No	No hit			
fam5_ak	VC_A0562	No duplication	No	COG5453	Uncharacterized conserved protein	S	Function unknown
fam5_aj	VC_A0564	No duplication	HGT	COG1282	NAD/NADP transhydrogenase beta subunit	C	Energy production and conversion
fam5_ai	VC_A0565	No duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam5_ah	VC_A0566	No duplication	No	COG0745	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	TK	Multiple classes
fam1_qp	VC_A0567	Duplication	No	No hit			
fam5_ag	VC_A0568	No duplication	No	No hit			
fam1_ki	VC_A0569	Duplication	No	COG3140	Uncharacterized protein conserved in bacteria	S	Function unknown
fam5_af	VC_A0570	No duplication	No	COG0023	Translation initiation factor 1 (eIF-1/SUI1) and related proteins	J	Translation, ribosomal structure and biogenesis
fam5_ae	VC_A0573	No duplication	No	COG3266	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_iy	VC_A0574*	Duplication	No	COG1042	Acyl-CoA synthetase (NDP forming)	C	Energy production and conversion
fam1_iy	VC_A0574*	Duplication	No	COG1247	Sortase and related acyltransferases	M	Cell wall/membrane/envelope biogenesis
fam5_ad	VC_A0575	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam1_gk	VC_A0576	Duplication	HGT	COG1629	Outer membrane receptor proteins, mostly Fe transport	P	Inorganic ion transport and metabolism
fam2_lm	VC_A0578	No duplication	No	COG4533	ABC-type uncharacterized transport system, periplasmic component	R	General function prediction only
fam2_ll	VC_A0580	No duplication	No	COG2050	Uncharacterized protein, possibly involved in aromatic compounds catabolism	Q	Secondary metabolites biosynthesis, transport and catabolism
fam2_lk	VC_A0581	No duplication	No	COG4783	Putative Zn-dependent protease, contains TPR repeats	R	General function prediction only

fam2_lj	VC_A0582	No duplication	No	COG2501	Uncharacterized conserved protein	S	Function unknown
fam1_dl	VC_A0583	Duplication	No	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam2_li	VC_A0584	No duplication	No	COG0625	Glutathione S-transferase	O	Posttranslational modification, protein turnover, chaperones
fam2_lh	VC_A0585	No duplication	No	COG0625	Glutathione S-transferase	O	Posttranslational modification, protein turnover, chaperones
fam2_lg	VC_A0586	No duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam2_lf	VC_A0587	No duplication	No	COG1661	Predicted DNA-binding protein with PD1-like DNA-binding motif	R	General function prediction only
fam1_ae	VC_A0588	Duplication	HGT	COG4172	ABC-type uncharacterized transport system, duplicated ATPase component	R	General function prediction only
fam2_le	VC_A0589	No duplication	No	COG4239	ABC-type uncharacterized transport system, permease component	R	General function prediction only
fam1_ak	VC_A0590	Duplication	HGT	COG4174	ABC-type uncharacterized transport system, permease component	R	General function prediction only
fam1_ak	VC_A0591	Duplication	HGT	COG4166	ABC-type oligopeptide transport system, periplasmic component	E	Amino acid transport and metabolism
fam2_ld	VC_A0592	No duplication	No	COG1051	ADP-ribose pyrophosphatase	F	Nucleotide transport and metabolism
fam1_lu	VC_A0593	Duplication	No	COG0618	Exopolyphosphatase-related proteins	R	General function prediction only
fam2_lc	VC_A0595	No duplication	HGT	No hit			
fam2_lb	VC_A0599	No duplication	HGT	No hit			
fam2_la	VC_A0600	No duplication	HGT	COG2188	Transcriptional regulators	K	Transcription
fam2_kw	VC_A0604	No duplication	HGT	COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	E	Amino acid transport and metabolism
fam2_kv	VC_A0609	No duplication	HGT	No hit			
fam1_lt	VC_A0610	Duplication	HGT	COG3155	Uncharacterized protein involved in an early stage of isoprenoid biosynthesis	Q	Secondary metabolites biosynthesis, transport and catabolism
fam2_kt	VC_A0612	No duplication	HGT	COG1970	Large-conductance mechanosensitive channel	M	Cell wall/membrane/envelope biogenesis
fam1_dk	VC_A0614	Duplication	HGT	COG2759	Formyltetrahydrofolate synthetase	F	Nucleotide transport and metabolism
fam2_ks	VC_A0615*	No duplication	HGT	COG0225	Peptide methionine sulfoxide reductase	O	Posttranslational modification, protein turnover, chaperones
fam2_ks	VC_A0615*	No duplication	HGT	COG0229	Conserved domain frequently associated with peptide methionine sulfoxide reductase	O	Posttranslational modification, protein turnover, chaperones
fam1_gj	VC_A0616	Duplication	HGT	COG0302	GTP cyclohydrolase I	H	Coenzyme transport and metabolism
fam1_gi	VC_A0617	Duplication	HGT	COG0303	Molybdopterin biosynthesis enzyme	H	Coenzyme transport and metabolism
fam2_kr	VC_A0618	No duplication	No	COG0476	Dinucleotide-utilizing enzymes involved in molybdopterin and thiamine biosynthesis family 2	H	Coenzyme transport and metabolism
fam2_kq	VC_A0619	No duplication	No	No hit			
fam2_kp	VC_A0620	No duplication	HGT	COG2897	Rhodanese-related sulfurtransferase	P	Inorganic ion transport and metabolism
fam2_ko	VC_A0621	No duplication	HGT	COG2390	Transcriptional regulator, contains sigma factor-related N-terminal domain	K	Transcription
fam2_km	VC_A0623	No duplication	HGT	COG0176	Transaldolase	G	Carbohydrate transport and metabolism
fam2_kl	VC_A0627	No duplication	HGT	COG0219	Predicted rRNA methylase (SpoU class)	J	Translation, ribosomal structure and biogenesis
fam2_kk	VC_A0629	No duplication	HGT	COG3180	Putative ammonia monooxygenase	R	General function prediction only
fam1_gh	VC_A0630	Duplication	HGT	COG0111	Phosphoglycerate dehydrogenase and related dehydrogenases	HE	Multiple classes
fam1_gg	VC_A0631	Duplication	HGT	COG3981	Predicted acetyltransferase	R	General function prediction only
fam2_kj	VC_A0632	No duplication	HGT	COG1051	ADP-ribose pyrophosphatase	F	Nucleotide transport and metabolism
fam2_ki	VC_A0633	No duplication	HGT	COG3136	Uncharacterized membrane protein required for alginate biosynthesis	R	General function prediction only
fam2_kh	VC_A0634	No duplication	HGT	COG2384	Predicted SAM-dependent methyltransferase	R	General function prediction only
fam2_kg	VC_A0636	No duplication	HGT	COG1359	Uncharacterized conserved protein	S	Function unknown
fam1_jv	VC_A0637	Duplication	HGT	COG0778	Nitroreductase	C	Energy production and conversion
fam2_kf	VC_A0638	No duplication	No	COG0841	Cation/multidrug efflux pump	V	Defense mechanisms
fam1_ls	VC_A0639	Duplication	No	COG0845	Membrane-fusion protein	M	Cell wall/membrane/envelope biogenesis
fam2_ke	VC_A0640	No duplication	No	COG2391	Predicted transporter component	R	General function prediction only
fam2_kd	VC_A0641	No duplication	No	COG2391	Predicted transporter component	R	General function prediction only
fam2_kc	VC_A0642	No duplication	No	COG0640	Predicted transcriptional regulators	K	Transcription
fam2_ka	VC_A0644*	No duplication	No	COG0446	Uncharacterized NAD(FAD)-dependent dehydrogenases	R	General function prediction only
fam2_ka	VC_A0644*	No duplication	No	COG0607	Rhodanese-related sulfurtransferase	P	Inorganic ion transport and metabolism
fam1_an	VC_A0646	Duplication	HGT	COG3176	Putative hemolysin	R	General function prediction only
fam2_jz	VC_A0652	No duplication	HGT	No hit			
fam5_gh	VC_A0653*	No duplication	HGT	COG1263	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	G	Carbohydrate transport and metabolism
fam5_gh	VC_A0653*	No duplication	HGT	COG1264	Phosphotransferase system IIB components	G	Carbohydrate transport and metabolism
fam5_gg	VC_A0654	No duplication	HGT	COG1609	Transcriptional regulators	K	Transcription
fam5_gf	VC_A0655	No duplication	HGT	COG1621	Beta-fructosidases (levanase/invertase)	G	Carbohydrate transport and metabolism
fam5_ge	VC_A0656	No duplication	HGT	COG0524	Sugar kinases, ribokinase family	G	Carbohydrate transport and metabolism
fam2_jx	VC_A0657	No duplication	HGT	COG0578	Glycerol-3-phosphate dehydrogenase	C	Energy production and conversion
fam1_dj	VC_A0658*	Duplication	HGT	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_dj	VC_A0658*	Duplication	HGT	COG2202	FOG: PAS/PAC domain	T	Signal transduction mechanisms
fam2_jw	VC_A0659	No duplication	No	COG2885	Outer membrane protein and related peptidoglycan-associated (lipo)proteins	M	Cell wall/membrane/envelope biogenesis
fam2_jv	VC_A0661	No duplication	No	COG0545	FKBP-type peptidyl-prolyl cis-trans isomerases 1	O	Posttranslational modification, protein turnover, chaperones
fam2_jt	VC_A0662	No duplication	No	COG0637	Predicted phosphatase/phosphohexomutase	R	General function prediction only
fam2_js	VC_A0663	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_jr	VC_A0665	No duplication	No	COG3069	C4-dicarboxylate transporter	C	Energy production and conversion
fam2_jq	VC_A0666	No duplication	No	COG1760	L-serine deaminase	E	Amino acid transport and metabolism
fam2_jo	VC_A0675	No duplication	No	COG3850	Signal transduction histidine kinase, nitrate/nitrite-specific	T	Signal transduction mechanisms
fam2_jn	VC_A0676	No duplication	No	COG1142	Fe-S-cluster-containing hydrogenase components 2	C	Energy production and conversion
fam2_jm	VC_A0677	No duplication	HGT	COG3062	Uncharacterized protein involved in formation of periplasmic nitrate reductase	P	Inorganic ion transport and metabolism
fam1_gf	VC_A0678	Duplication	No	COG0243	Anaerobic dehydrogenases, typically selenocysteine-containing	C	Energy production and conversion
fam2_jl	VC_A0679	No duplication	No	COG3043	Nitrate reductase cytochrome c-type subunit	C	Energy production and conversion
fam2_jk	VC_A0680	No duplication	No	COG3005	Nitrate/TMAO reductases, membrane-bound tetraheme cytochrome c subunit	C	Energy production and conversion
fam1_lr	VC_A0682	Duplication	HGT	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes

fam2_ji	VC_A0683	No duplication	HGT	COG3851	Signal transduction histidine kinase, glucose-6-phosphate specific	T	Signal transduction mechanisms
fam2_jh	VC_A0684	No duplication	HGT	COG2271	Sugar phosphate permease	G	Carbohydrate transport and metabolism
fam2_jg	VC_A0685	No duplication	No	COG1840	ABC-type Fe3+ transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam1_lq	VC_A0686	Duplication	HGT	COG1178	ABC-type Fe3+ transport system, permease component	P	Inorganic ion transport and metabolism
fam2_jf	VC_A0687	No duplication	HGT	COG3842	ABC-type spermidine/putrescine transport systems, ATPase components	E	Amino acid transport and metabolism
fam1_lp	VC_A0688	Duplication	HGT	COG3243	Poly(3-hydroxyalkanoate) synthetase	I	Lipid transport and metabolism
fam2_je	VC_A0689	No duplication	No	No hit			
fam2_jd	VC_A0690	No duplication	No	COG0183	Acetyl-CoA acetyltransferase	I	Lipid transport and metabolism
fam2_jc	VC_A0691	No duplication	No	COG1028	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	IQR	Multiple classes
fam1_lo	VC_A0693	Duplication	No	COG0342	Preprotein translocase subunit SecD	U	Intracellular trafficking, secretion, and vesicular transport
fam2_jb	VC_A0695	No duplication	No	No hit			
fam1_dd	VC_A0697	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_ja	VC_A0699	No duplication	No	COG0448	ADP-glucose pyrophosphorylase	G	Carbohydrate transport and metabolism
fam1_af	VC_A0700*	Duplication	HGT	COG3325	Chitinase	G	Carbohydrate transport and metabolism
fam1_af	VC_A0700*	Duplication	HGT	COG3979	Uncharacterized protein contain chitin-binding domain type 3	R	General function prediction only
fam1_fr	VC_A0702	Duplication	No	COG1979	Uncharacterized oxidoreductases, Fe-dependent alcohol dehydrogenase family	C	Energy production and conversion
fam1_fq	VC_A0703	Duplication	No	COG3714	Predicted membrane protein	S	Function unknown
fam1_cn	VC_A0704	Duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam1_da	VC_A0705	Duplication	No	COG4192	Signal transduction histidine kinase regulating phosphoglycerate transport system	T	Signal transduction mechanisms
fam1_fp	VC_A0706	Duplication	HGT	COG1840	ABC-type Fe3+ transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam1_fo	VC_A0707	Duplication	No	COG2271	Sugar phosphate permease	G	Carbohydrate transport and metabolism
fam1_fn	VC_A0708	Duplication	HGT	COG0469	Pyruvate kinase	G	Carbohydrate transport and metabolism
fam2_ir	VC_A0711	No duplication	No	COG1803	Methylglyoxal synthase	G	Carbohydrate transport and metabolism
fam1_lm	VC_A0712	Duplication	No	COG1335	Amidases related to nicotinamidase	Q	Secondary metabolites biosynthesis, transport and catabolism
fam2_iq	VC_A0714	No duplication	HGT	COG0818	Diacylglycerol kinase	M	Cell wall/membrane/envelope biogenesis
fam2_io	VC_A0717	No duplication	No	COG0210	Superfamily I DNA and RNA helicases	L	Replication, recombination and repair
fam2_in	VC_A0718	No duplication	No	COG1289	Predicted membrane protein	S	Function unknown
fam2_im	VC_A0721	No duplication	No	No hit			
fam2_il	VC_A0722	No duplication	No	No hit			
fam2_ik	VC_A0723	No duplication	No	COG1257	Hydroxymethylglutaryl-CoA reductase	I	Lipid transport and metabolism
fam2_ij	VC_A0724	No duplication	No	COG4148	ABC-type molybdate transport system, ATPase component	P	Inorganic ion transport and metabolism
fam2_ii	VC_A0726	No duplication	No	COG0725	ABC-type molybdate transport system, periplasmic component	P	Inorganic ion transport and metabolism
fam2_ig	VC_A0727	No duplication	No	COG1492	Cobyric acid synthase	H	Coenzyme transport and metabolism
fam2_if	VC_A0732	No duplication	No	COG3111	Uncharacterized conserved protein	S	Function unknown
fam2_ie	VC_A0734	No duplication	No	No hit			
fam2_id	VC_A0735	No duplication	No	No hit			
fam1_dc	VC_A0736*	Duplication	No	COG0642	Signal transduction histidine kinase	T	Signal transduction mechanisms
fam1_dc	VC_A0736*	Duplication	No	COG0784	FOG: CheY-like receiver	T	Signal transduction mechanisms
fam2_ic	VC_A0737	No duplication	No	COG1879	ABC-type sugar transport system, periplasmic component	G	Carbohydrate transport and metabolism
fam1_ll	VC_A0738	Duplication	No	No hit			
fam2_ib	VC_A0740	No duplication	No	COG3132	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_ia	VC_A0741	No duplication	No	COG2926	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_hz	VC_A0743	No duplication	No	No hit			
fam2_hy	VC_A0744	No duplication	No	COG0554	Glycerol kinase	C	Energy production and conversion
fam2_hx	VC_A0747	No duplication	HGT	COG0578	Glycerol-3-phosphate dehydrogenase	C	Energy production and conversion
fam2_hw	VC_A0748	No duplication	HGT	COG3075	Anaerobic glycerol-3-phosphate dehydrogenase	E	Amino acid transport and metabolism
fam2_hv	VC_A0749	No duplication	HGT	COG0247	Fe-S oxidoreductase	C	Energy production and conversion
fam2_hu	VC_A0751	No duplication	HGT	COG0332	3-oxoacyl-[acyl-carrier-protein] synthase III	I	Lipid transport and metabolism
fam2_ht	VC_A0752	No duplication	HGT	COG3118	Thioredoxin domain-containing protein	O	Posttranslational modification, protein turnover, chaperones
fam1_ge	VC_A0754	Duplication	HGT	COG3675	Predicted lipase	I	Lipid transport and metabolism
fam2_hr	VC_A0758	No duplication	HGT	COG4215	ABC-type arginine transport system, permease component	E	Amino acid transport and metabolism
fam2_hq	VC_A0759	No duplication	HGT	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam2_hp	VC_A0760	No duplication	No	COG4161	ABC-type arginine transport system, ATPase component	E	Amino acid transport and metabolism
fam1_db	VC_A0762	Duplication	No	COG2425	Uncharacterized protein containing a von Willebrand factor type A (vWA) domain	R	General function prediction only
fam1_lk	VC_A0763	Duplication	No	COG0714	MoxR-like ATPases	R	General function prediction only
fam2_ho	VC_A0764	No duplication	No	COG0494	NTP pyrophosphohydrolases including oxidative damage repair enzymes	LR	Multiple classes
fam2_hn	VC_A0765	No duplication	No	COG2008	Threonine aldolase	E	Amino acid transport and metabolism
fam2_hm	VC_A0766	No duplication	No	COG2863	Cytochrome c553	C	Energy production and conversion
fam1_lj	VC_A0767	Duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam2_hl	VC_A0768	No duplication	No	COG0513	Superfamily II DNA and RNA helicases	LKJ	Multiple classes
fam2_hi	VC_A0770	No duplication	No	No hit			
fam2_hh	VC_A0772	No duplication	No	COG0814	Amino acid permeases	E	Amino acid transport and metabolism
fam2_hg	VC_A0773	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_hf	VC_A0774	No duplication	No	COG1087	UDP-glucose 4-epimerase	M	Cell wall/membrane/envelope biogenesis
fam2_he	VC_A0776	No duplication	No	COG1540	Uncharacterized proteins, homologs of lactam utilization protein B	R	General function prediction only
fam2_hd	VC_A0777	No duplication	No	COG2049	Allophanate hydrolase subunit 1	E	Amino acid transport and metabolism
fam2_hc	VC_A0779	No duplication	No	COG0644	Dehydrogenases (flavoproteins)	C	Energy production and conversion
fam2_hb	VC_A0783	No duplication	No	COG2755	Lysophospholipase L1 and related esterases	E	Amino acid transport and metabolism

fam2_ha	VC_A0784	No duplication	No	COG3007	Uncharacterized paraquat-inducible protein B	S	Function unknown
fam2_gz	VC_A0785*	No duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam2_gz	VC_A0785*	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_gz	VC_A0785*	No duplication	No	COG2203	FOG: GAF domain	T	Signal transduction mechanisms
fam2_gy	VC_A0786	No duplication	No	COG3044	Predicted ATPase of the ABC class	R	General function prediction only
fam2_gx	VC_A0788	No duplication	No	COG0484	DnaJ-class molecular chaperone with C-terminal Zn finger domain	O	Posttranslational modification, protein turnover, chaperones
fam2_gw	VC_A0789	No duplication	No	COG3647	Predicted membrane protein	S	Function unknown
fam1_li	VC_A0798	Duplication	No	COG0637	Predicted phosphatase/phosphohexomutase	R	General function prediction only
fam2_gv	VC_A0799	No duplication	No	No hit			
fam2_gu	VC_A0800	No duplication	No	COG3045	Uncharacterized protein conserved in bacteria	S	Function unknown
fam1_ar	VC_A0801	Duplication	HGT	COG0524	Sugar kinases, ribokinase family	G	Carbohydrate transport and metabolism
fam2_gt	VC_A0803	No duplication	HGT	COG5640	Secreted trypsin-like serine protease	O	Posttranslational modification, protein turnover, chaperones
fam2_gq	VC_A0809	No duplication	No	COG3046	Uncharacterized protein related to deoxyribodipyrimidine photolyase	R	General function prediction only
fam2_gp	VC_A0811	No duplication	No	COG3397	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_go	VC_A0817	No duplication	No	COG0668	Small-conductance mechanosensitive channel	M	Cell wall/membrane/envelope biogenesis
fam1_bc	VC_A0820	Duplication	HGT	COG0459	Chaperonin GroEL (HSP60 family)	O	Posttranslational modification, protein turnover, chaperones
fam2_gn	VC_A0827	No duplication	No	COG2154	Pterin-4a-carbinolamine dehydratase	H	Coenzyme transport and metabolism
fam2_gm	VC_A0828	No duplication	No	COG3186	Phenylalanine-4-hydroxylase	E	Amino acid transport and metabolism
fam1_lg	VC_A0829	Duplication	No	COG0365	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	I	Lipid transport and metabolism
fam5_fe	VC_A0833	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam5_fd	VC_A0834	No duplication	No	COG3965	Predicted Co/Zn/Cd cation transporters	P	Inorganic ion transport and metabolism
fam2_gl	VC_A0835	No duplication	No	No hit			
fam2_gk	VC_A0836	No duplication	No	COG0110	Acetyltransferase (isoleucine patch superfamily)	R	General function prediction only
fam2_gj	VC_A0837	No duplication	No	COG1253	Hemolysins and related proteins containing CBS domains	R	General function prediction only
fam2_gi	VC_A0838	No duplication	No	COG1765	Predicted redox protein, regulator of disulfide bond formation	O	Posttranslational modification, protein turnover, chaperones
fam2_gh	VC_A0840	No duplication	No	COG2131	Deoxycytidylate deaminase	F	Nucleotide transport and metabolism
fam2_gg	VC_A0843	No duplication	No	COG0057	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	G	Carbohydrate transport and metabolism
fam2_gf	VC_A0846	No duplication	HGT	COG1280	Putative threonine efflux protein	E	Amino acid transport and metabolism
fam2_ge	VC_A0847	No duplication	No	COG0531	Amino acid transporters	E	Amino acid transport and metabolism
fam1_bq	VC_A0849	Duplication	No	No hit			
fam1_gd	VC_A0859	Duplication	No	COG4989	Predicted oxidoreductase	R	General function prediction only
fam2_gb	VC_A0860	No duplication	No	COG0366	Glycosidases	G	Carbohydrate transport and metabolism
fam1_lf	VC_A0862	Duplication	No	COG2067	Long-chain fatty acid transport protein	I	Lipid transport and metabolism
fam2_fz	VC_A0863	No duplication	No	No hit			
fam2_fy	VC_A0865	No duplication	No	COG3227	Zinc metalloprotease (elastase)	E	Amino acid transport and metabolism
fam2_fx	VC_A0867	No duplication	No	COG3047	Outer membrane protein W	M	Cell wall/membrane/envelope biogenesis
fam2_fw	VC_A0868	No duplication	No	No hit			
fam2_fv	VC_A0870	No duplication	No	COG1686	D-alanyl-D-alanine carboxypeptidase	M	Cell wall/membrane/envelope biogenesis
fam2_fm	VC_A0885	No duplication	HGT	COG1063	Threonine dehydrogenase and related Zn-dependent dehydrogenases	ER	Multiple classes
fam2_fl	VC_A0886	No duplication	HGT	COG0156	7-keto-8-aminopelargonate synthetase and related enzymes	H	Coenzyme transport and metabolism
fam2_fk	VC_A0889	No duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam2_fj	VC_A0890	No duplication	HGT	COG0346	Lactoylglutathione lyase and related lyases	E	Amino acid transport and metabolism
fam2_fi	VC_A0891	No duplication	HGT	COG3182	Uncharacterized iron-regulated membrane protein	S	Function unknown
fam2_fh	VC_A0892	No duplication	HGT	No hit			
fam2_fg	VC_A0894	No duplication	No	COG1387	Histidinol phosphatase and related hydrolases of the PHP family	ER	Multiple classes
fam1_le	VC_A0895	Duplication	HGT	COG2206	HD-GYP domain	T	Signal transduction mechanisms
fam1_kg	VC_A0896	Duplication	No	COG0364	Glucose-6-phosphate 1-dehydrogenase	G	Carbohydrate transport and metabolism
fam1_fi	VC_A0897	Duplication	HGT	COG0363	6-phosphogluconolactonase/Glucosamine-6-phosphate isomerase/deaminase	G	Carbohydrate transport and metabolism
fam1_cz	VC_A0898	Duplication	HGT	COG0362	6-phosphogluconate dehydrogenase	G	Carbohydrate transport and metabolism
fam5_ey	VC_A0901	No duplication	No	COG3531	Predicted protein-disulfide isomerase	O	Posttranslational modification, protein turnover, chaperones
fam2_ff	VC_A0902	No duplication	No	No hit			
fam2_fd	VC_A0903	No duplication	No	COG1929	Glycerate kinase	G	Carbohydrate transport and metabolism
fam2_fc	VC_A0906	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_fb	VC_A0908	No duplication	HGT	COG3721	Putative heme iron utilization protein	P	Inorganic ion transport and metabolism
fam2_fa	VC_A0909	No duplication	No	COG0635	Coproporphyrinogen III oxidase and related Fe-S oxidoreductases	H	Coenzyme transport and metabolism
fam2_ez	VC_A0910	No duplication	No	COG0810	Periplasmic protein TonB, links inner and outer membranes	M	Cell wall/membrane/envelope biogenesis
fam2_ey	VC_A0911	No duplication	No	COG0811	Biopolymer transport proteins	U	Intracellular trafficking, secretion, and vesicular transport
fam2_ex	VC_A0914	No duplication	No	COG0609	ABC-type Fe3+-siderophore transport system, permease component	P	Inorganic ion transport and metabolism
fam2_ew	VC_A0917	No duplication	No	COG1309	Transcriptional regulator	K	Transcription
fam2_ev	VC_A0919	No duplication	No	COG3691	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_eu	VC_A0920	No duplication	No	No hit			
fam2_et	VC_A0921	No duplication	No	COG0204	1-acyl-sn-glycerol-3-phosphate acyltransferase	I	Lipid transport and metabolism
fam2_es	VC_A0922	No duplication	No	No hit			
fam1_ad	VC_A0923	Duplication	HGT	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_er	VC_A0924*	No duplication	No	COG3217	Uncharacterized Fe-S protein	R	General function prediction only
fam2_er	VC_A0924*	No duplication	No	COG1018	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1	C	Energy production and conversion
fam2_ep	VC_A0925	No duplication	HGT	COG0418	Dihydroorotase	F	Nucleotide transport and metabolism
fam2_eo	VC_A0929	No duplication	HGT	COG2265	SAM-dependent methyltransferases related to tRNA (uracil-5-)-methyltransferase	J	Translation, ribosomal structure and biogenesis

fam2_en	VC_A0930	No duplication	No	COG2354	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_em	VC_A0931	No duplication	No	COG2206	HD-GYP domain	T	Signal transduction mechanisms
fam2_el	VC_A0933	No duplication	No	COG1278	Cold shock proteins	K	Transcription
fam2_ek	VC_A0936	No duplication	No	COG3608	Predicted deacylase	R	General function prediction only
fam2_ei	VC_A0937	No duplication	No	COG2207	AraC-type DNA-binding domain-containing proteins	K	Transcription
fam2_eh	VC_A0939	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_eg	VC_A0940	No duplication	No	COG1349	Transcriptional regulators of sugar metabolism	KG	Multiple classes
fam2_ef	VC_A0941	No duplication	No	COG1607	Acyl-CoA hydrolase	I	Lipid transport and metabolism
fam2_ee	VC_A0943	No duplication	No	COG3833	ABC-type maltose transport systems, permease component	G	Carbohydrate transport and metabolism
fam2_ed	VC_A0944	No duplication	No	COG1175	ABC-type sugar transport systems, permease components	G	Carbohydrate transport and metabolism
fam2_ec	VC_A0945	No duplication	No	COG2182	Maltose-binding periplasmic proteins/domains	G	Carbohydrate transport and metabolism
fam2_eb	VC_A0946	No duplication	No	COG3839	ABC-type sugar transport systems, ATPase components	G	Carbohydrate transport and metabolism
fam2_ea	VC_A0947	No duplication	No	COG1670	Acetyltransferases, including N-acetylases of ribosomal proteins	J	Translation, ribosomal structure and biogenesis
fam2_dz	VC_A0948	No duplication	No	COG4568	Transcriptional antiterminator	K	Transcription
fam2_dy	VC_A0949	No duplication	No	COG3751	Predicted proline hydroxylase	O	Posttranslational modification, protein turnover, chaperones
fam2_dx	VC_A0951	No duplication	No	COG0393	Uncharacterized conserved protein	S	Function unknown
fam2_dw	VC_A0952	No duplication	No	COG2197	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	TK	Multiple classes
fam2_dv	VC_A0953	No duplication	No	COG0760	Parvulin-like peptidyl-prolyl isomerase	O	Posttranslational modification, protein turnover, chaperones
fam2_du	VC_A0955	No duplication	No	COG1846	Transcriptional regulators	K	Transcription
fam2_dt	VC_A0956	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_ds	VC_A0957	No duplication	No	COG2225	Malate synthase	C	Energy production and conversion
fam2_dq	VC_A0958	No duplication	No	COG1396	Predicted transcriptional regulators	K	Transcription
fam1_ga	VC_A0961	Duplication	No	COG3145	Alkylated DNA repair protein	L	Replication, recombination and repair
fam2_dp	VC_A0962	No duplication	No	No hit			
fam2_do	VC_A0963	No duplication	No	No hit			
fam2_dn	VC_A0964	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam2_dm	VC_A0965	No duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam2_dl	VC_A0966	No duplication	No	No hit			
fam2_dk	VC_A0969	No duplication	No	COG1741	Pirin-related protein	R	General function prediction only
fam2_dj	VC_A0970	No duplication	No	COG3123	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_di	VC_A0971	No duplication	No	No hit			
fam1_lb	VC_A0972	Duplication	No	COG2271	Sugar phosphate permease	G	Carbohydrate transport and metabolism
fam1_eh	VC_A0975	Duplication	No	COG1067	Predicted ATP-dependent protease	O	Posttranslational modification, protein turnover, chaperones
fam1_ei	VC_A0976	Duplication	No	No hit			
fam1_ae	VC_A0977	Duplication	HGT	COG4172	ABC-type uncharacterized transport system, duplicated ATPase component	R	General function prediction only
fam1_cs	VC_A0982	Duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam1_ej	VC_A0983	Duplication	HGT	COG1620	L-lactate permease	C	Energy production and conversion
fam2_dh	VC_A0984	No duplication	HGT	COG1304	L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid dehydrogenases	C	Energy production and conversion
fam1_dh	VC_A0985*	Duplication	HGT	COG0277	FAD/FMN-containing dehydrogenases	C	Energy production and conversion
fam1_dh	VC_A0985*	Duplication	HGT	COG0247	Fe-S oxidoreductase	C	Energy production and conversion
fam1_bt	VC_A0988	Duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_ew	VC_A0989	Duplication	No	COG0534	Na+-driven multidrug efflux pump	V	Defense mechanisms
fam1_ex	VC_A0990	Duplication	HGT	COG0513	Superfamily II DNA and RNA helicases	LKJ	Multiple classes
fam2_dg	VC_A0994	No duplication	HGT	COG4874	Uncharacterized protein conserved in bacteria containing a pentein-type domain	S	Function unknown
fam1_ax	VC_A0996	Duplication	HGT	COG1132	ABC-type multidrug transport system, ATPase and permease components	V	Defense mechanisms
fam2_df	VC_A0997	No duplication	HGT	No hit			
fam1_fz	VC_A0998	Duplication	HGT	COG1902	NADH:flavin oxidoreductases, Old Yellow Enzyme family	C	Energy production and conversion
fam5_en	VC_A0999	No duplication	HGT	COG1309	Transcriptional regulator	K	Transcription
fam2_dd	VC_A1000	No duplication	HGT	COG1280	Putative threonine efflux protein	E	Amino acid transport and metabolism
fam2_dc	VC_A1001	No duplication	HGT	COG2207	AraC-type DNA-binding domain-containing proteins	K	Transcription
fam2_db	VC_A1002	No duplication	HGT	COG1296	Predicted branched-chain amino acid permease (azaleucine resistance)	E	Amino acid transport and metabolism
fam2_da	VC_A1003	No duplication	HGT	COG4392	Predicted membrane protein	S	Function unknown
fam2_cz	VC_A1004	No duplication	HGT	COG0179	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	Q	Secondary metabolites biosynthesis, transport and catabolism
fam2_cy	VC_A1005	No duplication	No	COG1846	Transcriptional regulators	K	Transcription
fam2_cx	VC_A1006	No duplication	HGT	COG1764	Predicted redox protein, regulator of disulfide bond formation	O	Posttranslational modification, protein turnover, chaperones
fam2_cw	VC_A1008	No duplication	No	COG3203	Outer membrane protein (porin)	M	Cell wall/membrane/envelope biogenesis
fam2_cv	VC_A1010	No duplication	No	COG0697	Permeases of the drug/metabolite transporter (DMT) superfamily	GER	Multiple classes
fam2_cu	VC_A1011	No duplication	No	COG0695	Glutaredoxin and related proteins	O	Posttranslational modification, protein turnover, chaperones
fam2_ct	VC_A1013	No duplication	No	COG3314	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_cs	VC_A1015	No duplication	No	COG1055	Na+/H+ antiporter NhaD and related arsenite permeases	P	Inorganic ion transport and metabolism
fam2_cr	VC_A1016	No duplication	No	No hit			
fam2_cq	VC_A1017	No duplication	No	COG0350	Methylated DNA-protein cysteine methyltransferase	L	Replication, recombination and repair
fam2_cp	VC_A1018*	No duplication	No	COG2169	Adenosine deaminase	F	Nucleotide transport and metabolism
fam2_cp	VC_A1018*	No duplication	No	COG0122	3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase	L	Replication, recombination and repair
fam2_co	VC_A1019	No duplication	No	COG2259	Predicted membrane protein	S	Function unknown
fam2_cn	VC_A1020	No duplication	No	COG0583	Transcriptional regulator	K	Transcription
fam2_cm	VC_A1021	No duplication	No	COG3085	Uncharacterized protein conserved in bacteria	S	Function unknown
fam2_cl	VC_A1023	No duplication	No	COG3930	Uncharacterized protein conserved in bacteria	S	Function unknown

fam2_ck	VC_A1024	No duplication	No	No hit			
fam2_cj	VC_A1025	No duplication	No	COG0363	6-phosphogluconolactonase/Glucosamine-6-phosphate isomerase/deaminase	G	Carbohydrate transport and metabolism
fam2_ci	VC_A1026	No duplication	No	COG0790	FOG: TPR repeat, SEL1 subfamily	R	General function prediction only
fam2_ch	VC_A1027	No duplication	HGT	No hit			
fam1_fy	VC_A1029	Duplication	No	COG1523	Type II secretory pathway, pullulanase PulA and related glycosidases	G	Carbohydrate transport and metabolism
fam2_cg	VC_A1030	No duplication	No	No hit			
fam2_cf	VC_A1033	No duplication	No	COG1653	ABC-type sugar transport system, periplasmic component	G	Carbohydrate transport and metabolism
fam2_ce	VC_A1034	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_cd	VC_A1035	No duplication	No	No hit			
fam2_cc	VC_A1037	No duplication	HGT	COG1126	ABC-type polar amino acid transport system, ATPase component	E	Amino acid transport and metabolism
fam2_cb	VC_A1038	No duplication	HGT	COG0765	ABC-type amino acid transport system, permease component	E	Amino acid transport and metabolism
fam1_la	VC_A1039	Duplication	HGT	COG0834	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	ET	Multiple classes
fam2_ca	VC_A1040	No duplication	HGT	COG0765	ABC-type amino acid transport system, permease component	E	Amino acid transport and metabolism
fam2_bz	VC_A1041	No duplication	HGT	COG1109	Phosphomannomutase	G	Carbohydrate transport and metabolism
fam2_by	VC_A1048	No duplication	HGT	COG0673	Predicted dehydrogenases and related proteins	R	General function prediction only
fam2_bx	VC_A1050	No duplication	HGT	No hit			
fam2_bw	VC_A1051	No duplication	HGT	COG3304	Predicted membrane protein	S	Function unknown
fam2_bv	VC_A1052	No duplication	HGT	COG0730	Predicted permeases	R	General function prediction only
fam2_bu	VC_A1053	No duplication	HGT	COG0607	Rhodanese-related sulfurtransferase	P	Inorganic ion transport and metabolism
fam2_bt	VC_A1054	No duplication	HGT	COG2910	Putative NADH-flavin reductase	R	General function prediction only
fam2_bs	VC_A1055	No duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam2_bp	VC_A1057	No duplication	HGT	COG4221	Short-chain alcohol dehydrogenase of unknown specificity	R	General function prediction only
fam1_kz	VC_A1058	Duplication	HGT	COG0583	Transcriptional regulator	K	Transcription
fam2_bo	VC_A1059	No duplication	HGT	COG1901	Uncharacterized conserved protein	S	Function unknown
fam2_bn	VC_A1060	No duplication	HGT	COG0108	3,4-dihydroxy-2-butanone 4-phosphate synthase	H	Coenzyme transport and metabolism
fam2_bm	VC_A1062	No duplication	HGT	COG0531	Amino acid transporters	E	Amino acid transport and metabolism
fam2_bl	VC_A1063	No duplication	HGT	COG1982	Arginine/lysine/ornithine decarboxylases	E	Amino acid transport and metabolism
fam2_bj	VC_A1068	No duplication	HGT	COG1522	Transcriptional regulators	K	Transcription
fam2_bi	VC_A1069	No duplication	HGT	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam1_jw	VC_A1070	Duplication	HGT	No hit			
fam2_bh	VC_A1071	No duplication	HGT	COG0591	Na+/proline symporter	ER	Multiple classes
fam2_bg	VC_A1072	No duplication	HGT	COG4230	Delta 1-pyrroline-5-carboxylate dehydrogenase	C	Energy production and conversion
fam2_bf	VC_A1074	No duplication	HGT	COG2207	AraC-type DNA-binding domain-containing proteins	K	Transcription
fam2_be	VC_A1075	No duplication	No	COG1196	Chromosome segregation ATPases	D	Cell cycle control, cell division, chromosome partitioning
fam2_bc	VC_A1078*	No duplication	No	COG2771	DNA-binding HTH domain-containing proteins	K	Transcription
fam2_bc	VC_A1078*	No duplication	No	COG2202	FOG: PAS/PAC domain	T	Signal transduction mechanisms
fam2_bb	VC_A1079	No duplication	No	COG0259	Pyridoxamine-phosphate oxidase	H	Coenzyme transport and metabolism
fam1_ky	VC_A1080	Duplication	No	COG1566	Multidrug resistance efflux pump	V	Defense mechanisms
fam2_ba	VC_A1081	No duplication	No	COG3672	Predicted periplasmic protein	S	Function unknown
fam1_dg	VC_A1082	Duplication	No	COG2199	FOG: GGDEF domain	T	Signal transduction mechanisms
fam1_dg	VC_A1083	Duplication	No	COG2200	FOG: EAL domain	T	Signal transduction mechanisms
fam2_az	VC_A1084	No duplication	No	COG2274	ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain	V	Defense mechanisms
fam1_kx	VC_A1085	Duplication	No	COG1315	Predicted polymerase, most proteins contain PALM domain, HD hydrolase domain and Zn-ribbon domain	L	Replication, recombination and repair
fam1_ee	VC_A1086*	Duplication	No	COG2208	Serine phosphatase RsbU, regulator of sigma subunit	TK	Multiple classes
fam1_ee	VC_A1086*	Duplication	No	COG3437	Response regulator containing a CheY-like receiver domain and an HD-GYP domain	KT	Multiple classes
fam1_ed	VC_A1087	Duplication	No	COG1366	Anti-anti-sigma regulatory factor (antagonist of anti-sigma factor)	T	Signal transduction mechanisms
fam1_ec	VC_A1088	Duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_ay	VC_A1090	No duplication	No	COG1871	Chemotaxis protein; stimulates methylation of MCP proteins	NT	Multiple classes
fam2_ax	VC_A1091	No duplication	No	COG1352	Methylase of chemotaxis methyl-accepting proteins	NT	Multiple classes
fam2_aw	VC_A1092	No duplication	No	COG0840	Methyl-accepting chemotaxis protein	NT	Multiple classes
fam2_av	VC_A1094	No duplication	No	COG0835	Chemotaxis signal transduction protein	NT	Multiple classes
fam1_kw	VC_A1095	Duplication	No	COG0643	Chemotaxis protein histidine kinase and related kinases	NT	Multiple classes
fam2_au	VC_A1096	No duplication	No	COG2204	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	T	Signal transduction mechanisms
fam2_at	VC_A1097	No duplication	No	COG1366	Anti-anti-sigma regulatory factor (antagonist of anti-sigma factor)	T	Signal transduction mechanisms
fam2_aq	VC_A1113	No duplication	No	COG0687	Spermidine/putrescine-binding periplasmic protein	E	Amino acid transport and metabolism
fam2_ap	VC_A1114	No duplication	No	COG1475	Predicted transcriptional regulators	K	Transcription
fam2_ao	VC_A1115	No duplication	No	COG1192	ATPases involved in chromosome partitioning	D	Cell cycle control, cell division, chromosome partitioning

[a] Arbitrary gene family unique identifier

Table S5

HGT count for *V. metoecus* and *V. cholerae* strains based on 376 single-copy core genes with inferred HGT. An HGT event was hypothesized when a strain clustered with members of the other species in a phylogenetic tree, with reliable bootstrap support (>70%). Unequal variance *t*-test, *p* = 0.0053. Multiple *V. cholerae* YB strains from Oyster Pond (Massachusetts, USA) belong to the same clonal complex (CC), and only one strain from each clonal complex was included.

Species	Strain	Classification	HGT Count	% of Total
V. metoecus	07-2435	Clinical	151	23.05
V. metoecus	OP3H	Environmental	55	8.40
V. metoecus	06-2478	Clinical	48	7.33
V. metoecus	YB4D01	Environmental	43	6.56
V. metoecus	08-2459	Clinical	38	5.80
V. metoecus	YB5B06	Environmental	37	5.65
V. metoecus	2010V-1005	Clinical	31	4.73
V. metoecus	YB5B04	Environmental	30	4.58
V. metoecus	RC341	Environmental	29	4.43
V. metoecus	YB9D03	Environmental	27	4.12
V. cholerae	V51	non-O1/O139 clinical	21	3.21
V. cholerae	RC385	Environmental	17	2.60
V. cholerae	TMA21	Environmental	16	2.44
V. cholerae	YB2G01 (CC 5)	Environmental	16	2.44
V. cholerae	MZO-2	non-O1/O139 clinical	10	1.53
V. cholerae	YB4F05 (CC 3)	Environmental	9	1.37
V. cholerae	1587	non-O1/O139 clinical	8	1.22
V. cholerae	877-163	Environmental	7	1.07
V. cholerae	AM-19226	non-O1/O139 clinical	6	0.92
V. cholerae	MAK757	O1/O139 clinical	6	0.92
V. cholerae	2740-80	Environmental	5	0.76
V. cholerae	BX 330286	Environmental	4	0.61
V. cholerae	O395	O1/O139 clinical	4	0.61
V. cholerae	TM 11079-80	Environmental	4	0.61
V. cholerae	YB4B03 (CC 2)	Environmental	4	0.61
V. cholerae	12129(1)	Environmental	3	0.46
V. cholerae	2010EL-1786	O1/O139 clinical	3	0.46
V. cholerae	CIRS101	O1/O139 clinical	3	0.46
V. cholerae	MJ-1236	O1/O139 clinical	3	0.46
V. cholerae	MO10	O1/O139 clinical	3	0.46
V. cholerae	N16961	O1/O139 clinical	3	0.46
V. cholerae	623-39	Environmental	2	0.31
V. cholerae	MZO-3	non-O1/O139 clinical	2	0.31
V. cholerae	V52	non-O1/O139 clinical	2	0.31
V. cholerae	VL426	Environmental	2	0.31
V. cholerae	YB7A06 (CC 4)	Environmental	2	0.31
V. cholerae	YB3B05 (CC 1)	Environmental	1	0.15

Classification	HGT Count	% of Total
V. metoecus clinical	268	40.92
V. metoecus environmental	221	33.74
	489	**74.66**
V. cholerae environmental	92	14.05
V. cholerae non-O1/O139 clinical	49	7.48
V. cholerae O1/O139 clinical	25	3.82
	166	**25.34**
Total Count	**655**	**100**

Table S6

ANI of the core genome of *V. cholerae* strains from Oyster Pond (Massachusetts, USA) based on the pairwise comparison of the 1,560 single-copy core genes (≈ 1.42 mbp). The different clonal complexes are indicated by colour.

		Clonal Complex 1			Clonal Complex 2			Clonal Complex 3						Clonal Complex 4			Clonal Complex 5				
		YB3B05	YB4C07	YB8E08	YB2G07	YB4B03	YB4G05	YB1G06	YB2A05	YB2G05	YB4F05	YB6A06	YB7A09	YB1A01	YB5A06	YB7A06	YB2A06	YB2G01	YB3G04	YB4G06	YB4H02
1	**YB3B05**	100	100	100	99	99	99	98	98	98	98	98	98	98	98	98	98	98	98	98	98
	YB4C07	100	100	100	99	99	99	98	98	98	98	98	98	98	98	98	98	98	98	98	98
	YB8E08	100	100	100	99	99	99	98	98	98	98	98	98	98	98	98	98	98	98	98	98
2	**YB2G07**	99	99	99	100	100	100	98	98	98	98	98	98	98	98	98	98	98	98	98	98
	YB4B03	99	99	99	100	100	100	98	98	98	98	98	98	98	98	98	98	98	98	98	98
	YB4G05	99	99	99	100	100	100	98	98	98	98	98	98	98	98	98	98	98	98	98	98
3	**YB1G06**	98	98	98	98	98	98	100	100	100	100	100	100	99	99	99	98	98	98	98	98
	YB2A05	98	98	98	98	98	98	100	100	100	100	100	100	99	99	99	98	98	98	98	98
	YB2G05	98	98	98	98	98	98	100	100	100	100	100	100	99	99	99	98	98	98	98	98
	YB4F05	98	98	98	98	98	98	100	100	100	100	100	100	99	99	99	98	98	98	98	98
	YB6A06	98	98	98	98	98	98	100	100	100	100	100	100	99	99	99	98	98	98	98	98
	YB7A09	98	98	98	98	98	98	100	100	100	100	100	100	99	99	99	98	98	98	98	98
4	**YB1A01**	98	98	98	98	98	98	99	99	99	99	99	99	100	100	100	98	98	98	98	98
	YB5A06	98	98	98	98	98	98	99	99	99	99	99	99	100	100	100	98	98	98	98	98
	YB7A06	98	98	98	98	98	98	99	99	99	99	99	99	100	100	100	98	98	98	98	98
5	**YB2A06**	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	100	100	100	100	100
	YB2G01	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	100	100	100	100	100
	YB3G04	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	100	100	100	100	100
	YB4G06	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	100	100	100	100	100
	YB4H02	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	100	100	100	100	100

Legend:

🟥	Clonal complex 1
🟩	Clonal complex 2
🟪	Clonal complex 3
🟦	Clonal complex 4
🟧	Clonal complex 5

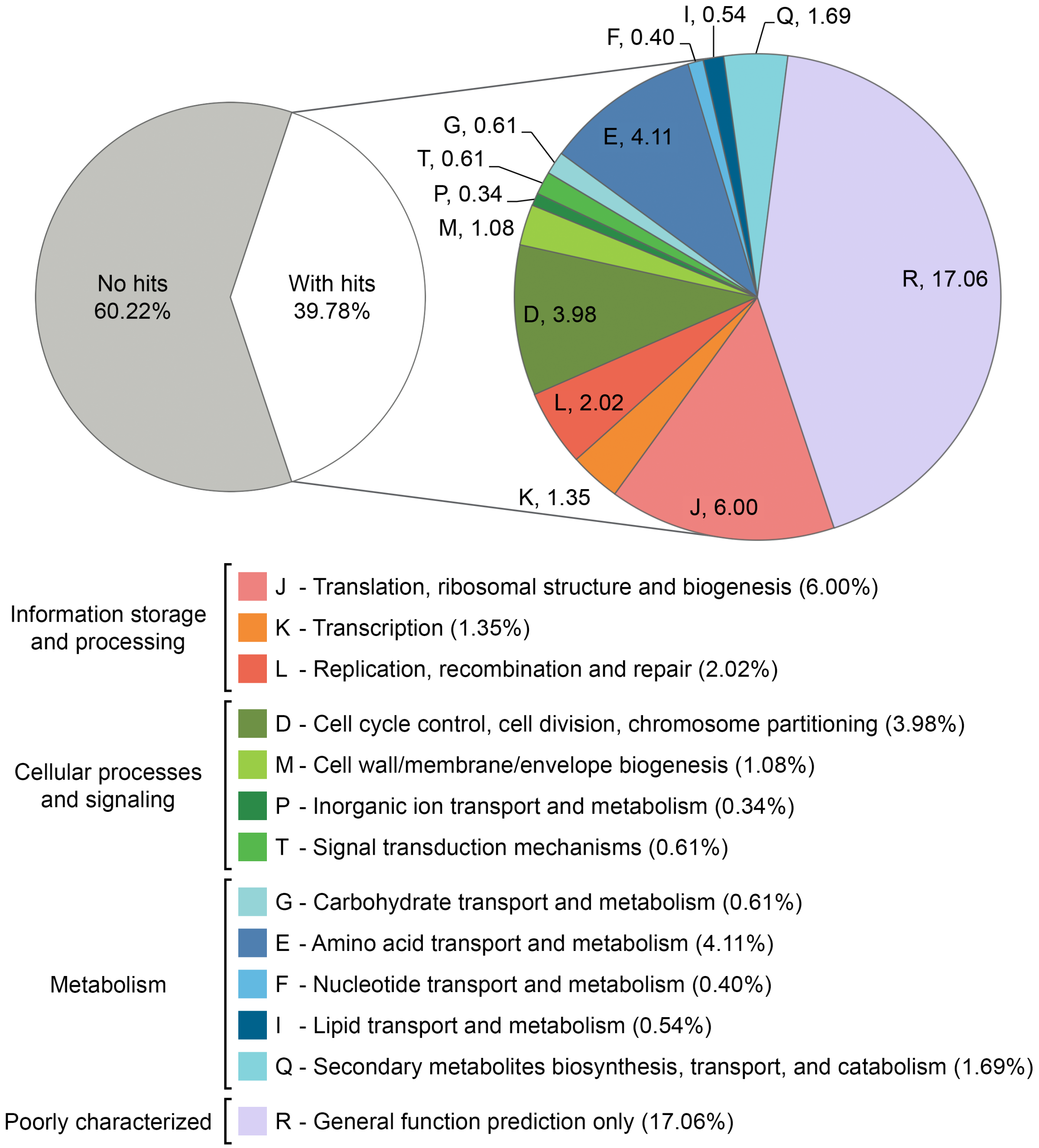


Fig. S1. – Predicted functional roles of the 1,452 *V. metoecus* and *V. cholerae* chromosomal integron gene cassettes from 116 gene families based on the COG database. The numbers on the pie chart represent percentage of genes in a category. Genes categorized as “function unknown” or not assigned into a COG category because there is no similarity to the database are combined (“no hits,” gray), whereas the genes with assigned COG categories (“with hits,” white) belong to four major groups: information storage and processing (reds), cellular processes and signaling (greens), metabolism (blues), and poorly characterized genes with general function prediction only (purple).

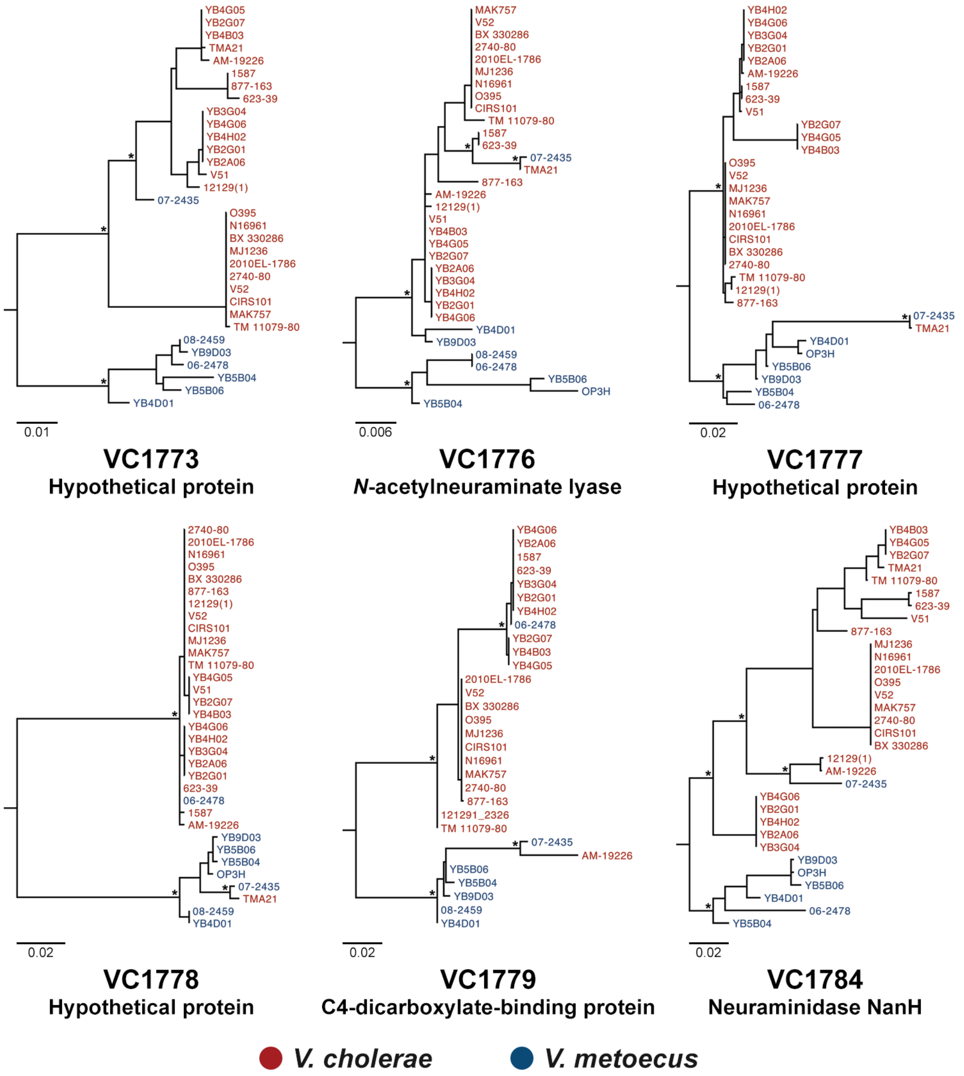


Fig. S2. – ML phylogenetic trees of orthologous gene families of islet B (*nan-nag* cluster, VC1773-VC1784) of the VPI-2 with reliable bootstrap support (6 out of 11 gene families). Relevant bootstrap support (>70%) is indicated with *. The scale bars represent nucleotide substitutions per site.

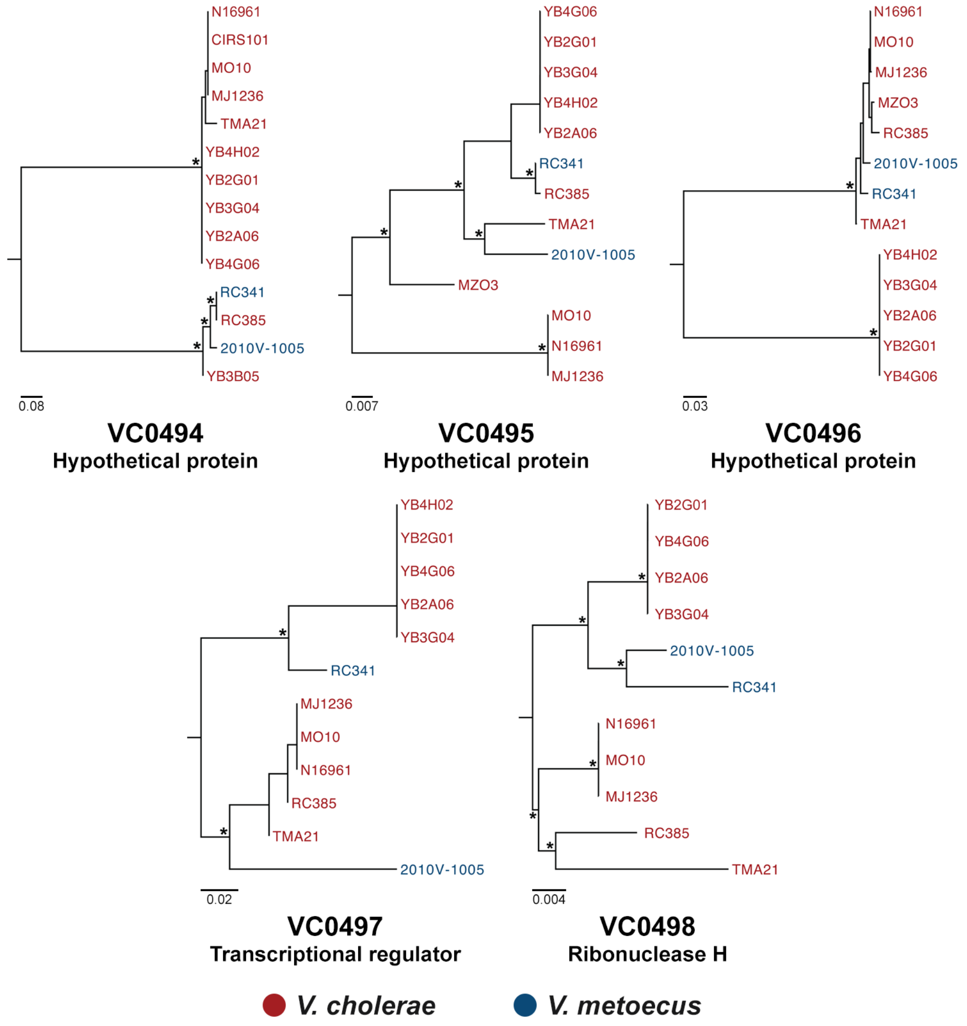


Fig. S3. – ML phylogenetic trees of orthologous gene families of the core VSP-2 (5' end, VC0494-VC0498) with reliable bootstrap support (5 out of 5 gene families). Relevant bootstrap support (>70%) is indicated with *. The scale bars represent nucleotide substitutions per site. The 3' end (VC0504-VC0510) did not produce trees with reliable bootstrap support.

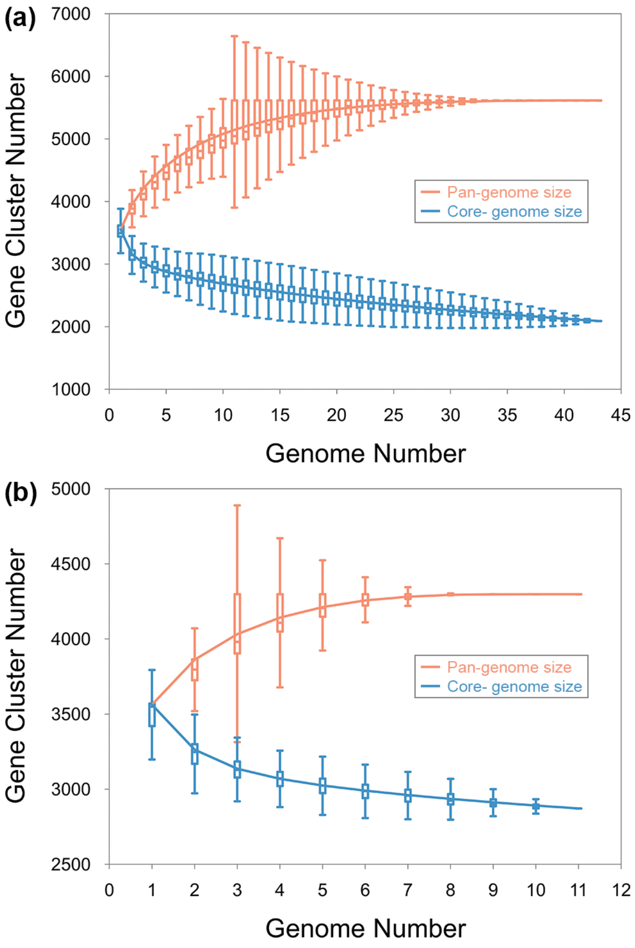


Fig. S4. – The pan-genome and core genome profiles for (*a*) *V. cholerae* based on 42 genomes and (*b*) *V. metoecus* based on 10 genomes. Graphs show the number of gene clusters (families) with every addition of a genome into the gene pool.

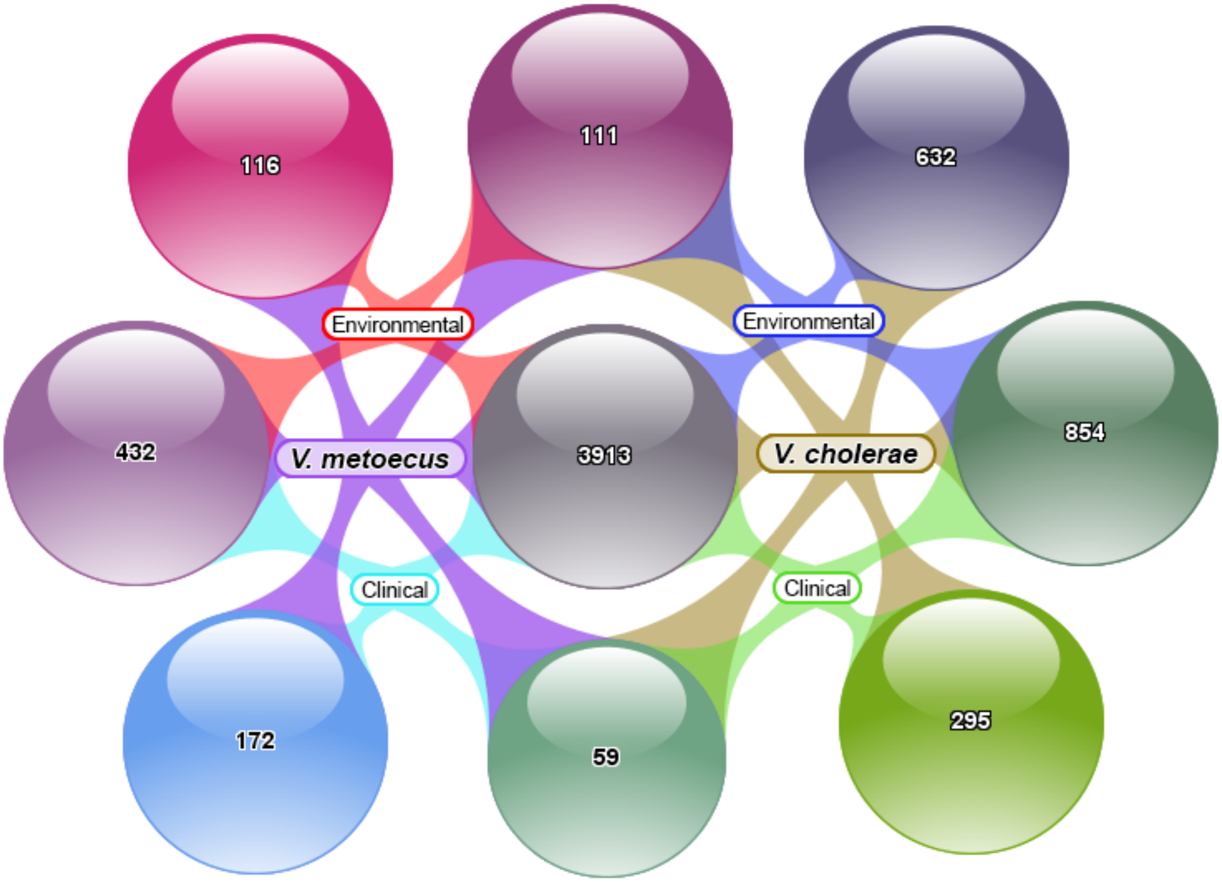
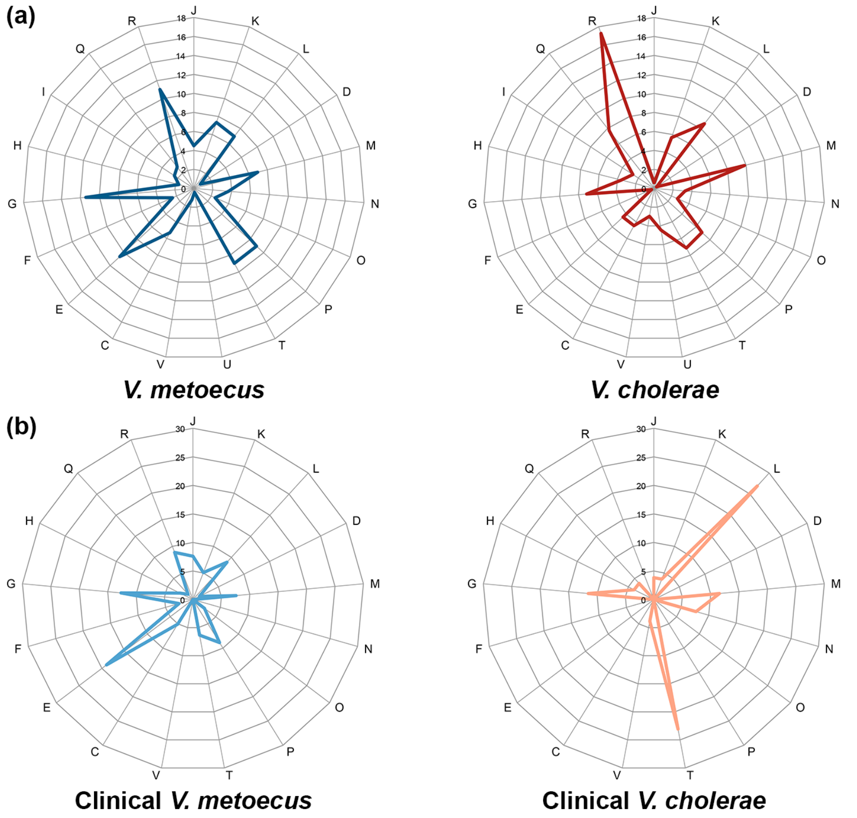


Fig. S5. – Cluster diagram of all orthologous gene families of the *V. metoecus*-*V. cholerae* pan-genome. Numbers indicate gene families containing members of various groups of strains. Clockwise from top left bubble: environmental *V. metoecus* only (116), environmental *V. metoecus* and *V. cholerae* (111), environmental *V. cholerae* only (632), environmental and clinical *V. cholerae* (854), clinical *V. cholerae* only (295), clinical *V. metoecus* and *V. cholerae* (59), clinical *V. metoecus* only (172), environmental and clinical *V. metoecus* (432). The gene families in the middle bubble contain environmental and clinical *V. metoecus* and *V. cholerae* members (3,913).



Information storage and processing
- J - Translation, ribosomal structure and biogenesis
- K - Transcription
- L - Replication, recombination and repair

Cellular processes and signaling
- D - Cell cycle control, cell division, chromosome partitioning
- M - Cell wall/membrane/envelope biogenesis
- N - Cell motility
- O - Post-translational modification, protein turnover, chaperones
- P - Inorganic ion transport and metabolism
- T - Signal transduction mechanisms
- U - Intracellular trafficking, secretion, and vesicular transport
- V - Defense mechanisms

Metabolism
- C - Energy production and conversion
- E - Amino acid transport and metabolism
- F - Nucleotide transport and metabolism
- G - Carbohydrate transport and metabolism
- H - Coenzyme transport and metabolism
- I - Lipid transport and metabolism
- Q - Secondary metabolites biosynthesis, transport, and catabolism

Poorly characterized
- R - General function prediction only

Fig. S6. – Comparison of orthologous gene families and their abundance, uniquely found in (*a*) *V. metoecus* or *V. cholerae* and (*b*) clinical *V. metoecus* or clinical *V. cholerae*. Functional roles for genes were predicted based on the COG database. The numbers on the charts represent percentage of genes in a COG category. Genes categorized as “function unknown” or not assigned into a COG category because there is no similarity to the database are not included.