Supplemental Table S2. Genes containing SNPs that correlate best (top 1%) with the Clade I/Clade II division

Gene[a]	**Variable positions**[b]	**Coding region (bases)**	**Top 1% SNPs**[c]	**SNPs causing amino acid change**	**Putative Gene Product**
DNO_0033	44	741	4	3	extracellular solute-binding family protein
DNO_0246*	76	2504	2	0	leucyl-tRNA synthetase
DNO_0249	21	777	1	1	hypothetical protein
DNO_0297	88	1755	51	34	hypothetical protein
DNO_0298	12	756	2	1	uroporphyrinogen-III synthase family protein HemD
DNO_0299	18	918	3	1	porphobilinogen deaminase
DNO_0307	12	621	6	0	ribonuclease T
DNO_0309*	25	1164	6	1	N-acetylglucosamine-6-phosphate deacetylase
DNO_0311	85	1455	10	5	N-acetylmuramoyl-L-alanine amidase
DNO_0330*	117	1842	5	0	competence family protein
DNO_0421	27	1344	1	0	acetyl-CoA carboxylase; biotin carboxylase
DNO_0423	32	957	1	0	dihydrouridine synthase
DNO_0426	21	1290	5	0	hypothetical protein
DNO_0530	157	1620	3	1	sulfatase
DNO_0534	7	456	2	1	OstA-like family protein
DNO_0538	21	696	1	0	ubiquinone biosynthesis O-methyltransferase
DNO_0605	23	1812	6	2	basic extracellular subtilisin-like protease BprV
DNO_0619	58	1371	9	2	L-serine dehydratase
DNO_0620*	15	192	4	1	ABC iron transporter permease
DNO_0622	90	3909	1	1	ATP-dependent helicase HrpA
DNO_0688	86	3528	1	0	DNA polymerase III subunit alpha
DNO_0798	17	1035	1	1	isopentenyl pyrophosphate isomerase
DNO_0817*	13	816	4	1	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase

DNO_0818*	17	474	3	1	GTP cyclohydrolase I
DNO_0819*	48	1236	7	3	two-component phosphate sensor PhoR
DNO_0820*	19	675	7	3	hypothetical membrane protein
DNO_0822	15	1206	2	0	response regulator receiver domain-containing protein
DNO_0823*	31	1197	6	0	phosphopantothenoylcysteine decarboxylase/phosphopantothenate-cysteine ligase
DNO_0825	9	564	1	0	co-chaperone GrpE
DNO_0869	57	2208	10	1	two-component sensor kinase
DNO_0870	51	732	3	0	glycosyl transferase family protein
DNO_0871	34	1185	1	1	macrolide-specific ABC-type efflux protein
DNO_0872	30	1935	4	1	macrolide-specific ABC-type efflux carrier protein
DNO_0874	23	1401	8	0	aspartate ammonia-lyase
DNO_0899	29	1128	1	0	D-alanyl-D-alanine carboxypeptidase
DNO_0917	22	777	2	0	exodeoxyribonuclease III
DNO_0929	13	834	1	1	pyridoxal kinase
DNO_0930	57	1779	6	4	hypothetical protein
DNO_0979	27	933	1	1	D-alanine-D-alanine ligase
DNO_1015	15	675	1	0	hypothetical protein
DNO_1031	37	1008	3	1	heptosyl transferase I
DNO_1034*	80	1998	37	10	carboxyl-terminal protease family protein
DNO_1047	49	2082	4	3	polynucleotide phosphorylase/polyadenylase
DNO_1070	6	369	1	0	glycine cleavage system protein H
DNO_1071	19	537	1	0	adenylate kinase
DNO_1135	23	819	1	0	formamidopyrimidine-DNA glycosylase
DNO_1165	93	3597	3	1	exonuclease V; beta subunit
DNO_1166	1	357	1	0	hypothetical protein
DNO_1167*	29	1806	11	5	acidic extracellular subtilisin-like protease AprV2
DNO_1191	75	3141	1	1	mutidrug efflux protein

DNO_1200	26	1455	1	1	pyruvate kinase
DNO_1202	25	585	12	3	N-acetyl-anhydromuranmyl-L-alanine amidase
DNO_1203*	18	867	10	2	hypothetical protein
DNO_1217	50	1539	1	1	glutamate--cysteine ligase
DNO_1340	4	288	1	0	co-chaperonin GroES
DNO_1343	31	750	12	9	extracellular solute-binding protein
DNO_1344	21	645	3	2	phosphoglycerate mutase family protein
DNO_1345	12	738	1	0	extracellular solute-binding family protein
Non-coding regions	1156		15	-	

[a] Locus tag from the VCS1703A genome annotation.

[b] Positions where at least one of the 101 isolates has a SNP

[c] SNPs were sorted using Plink in an order that ranged from exactly matching the Clade I/Clade II division through to most unrelated. The top 1% of SNPs that exactly matched and nearly matched were used for the top 1% SNP set

* indicates genes included in Table 2