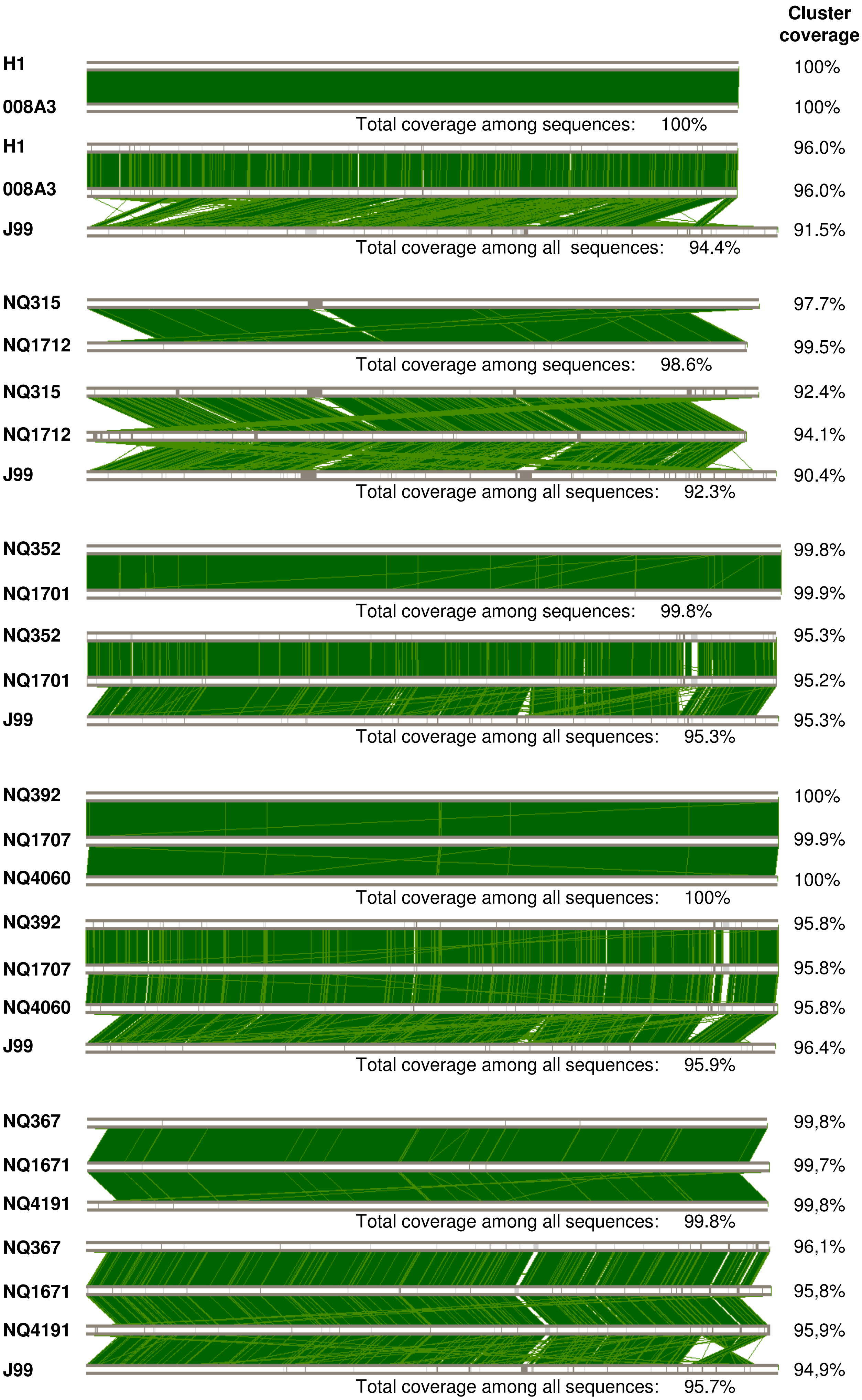
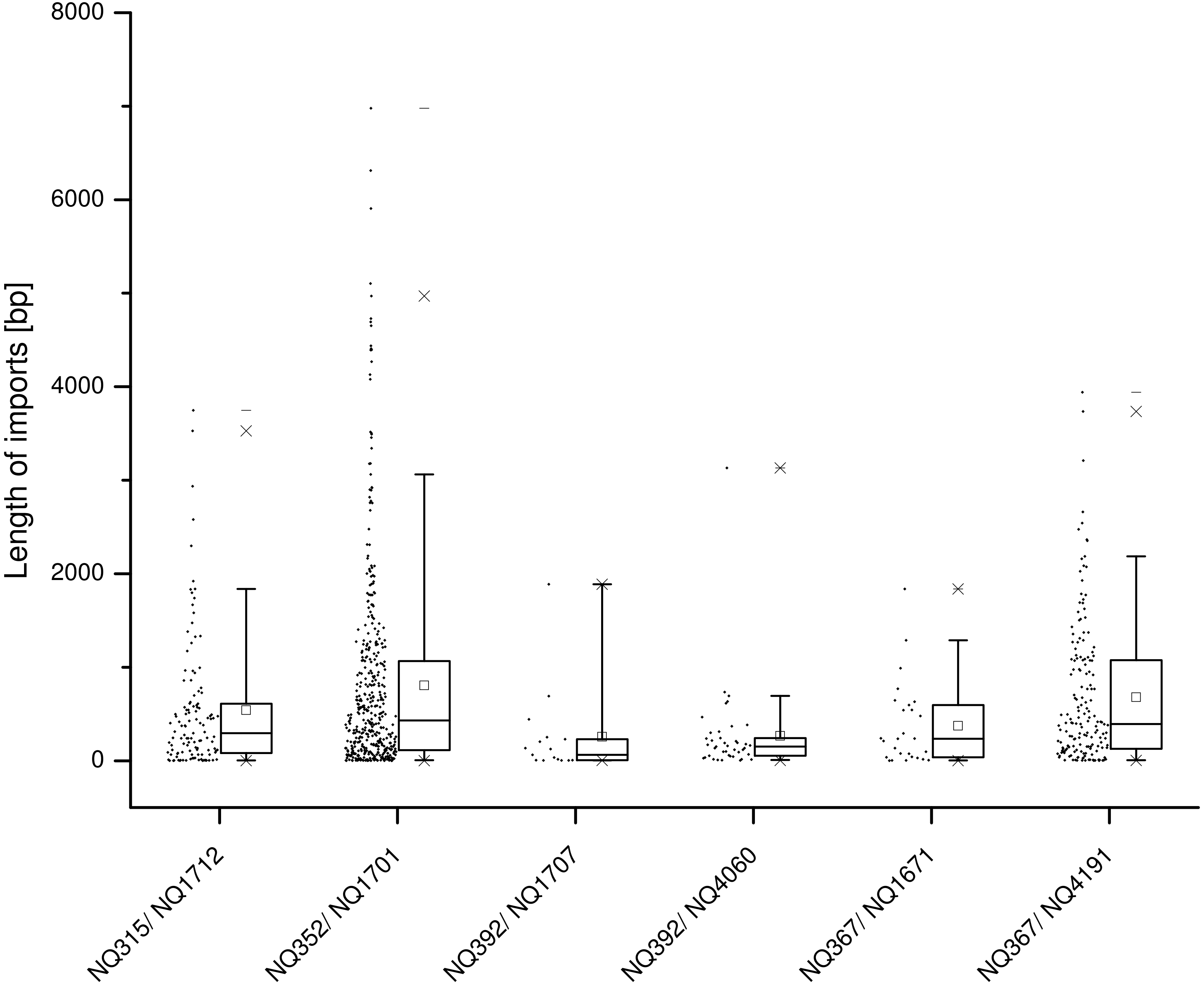
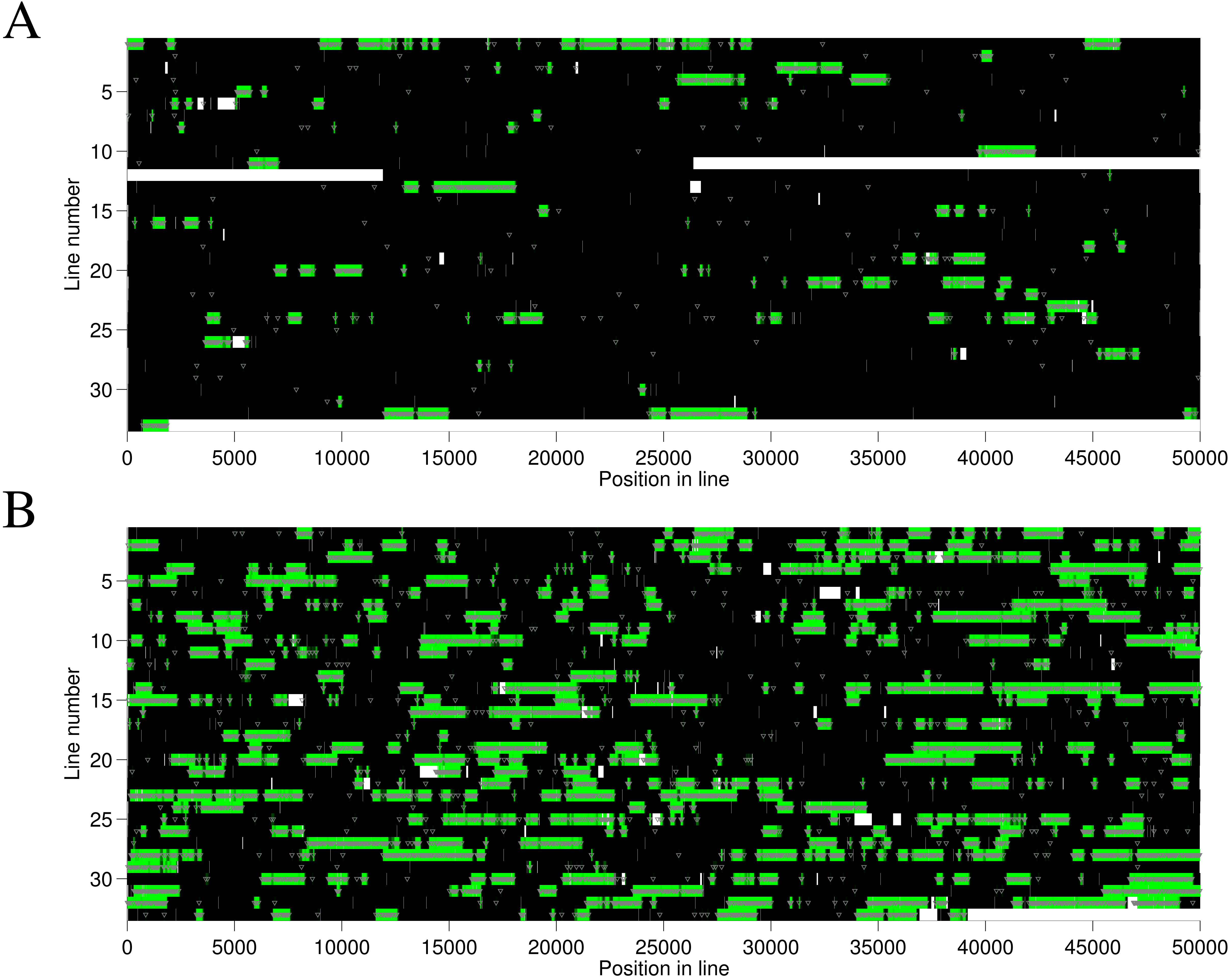
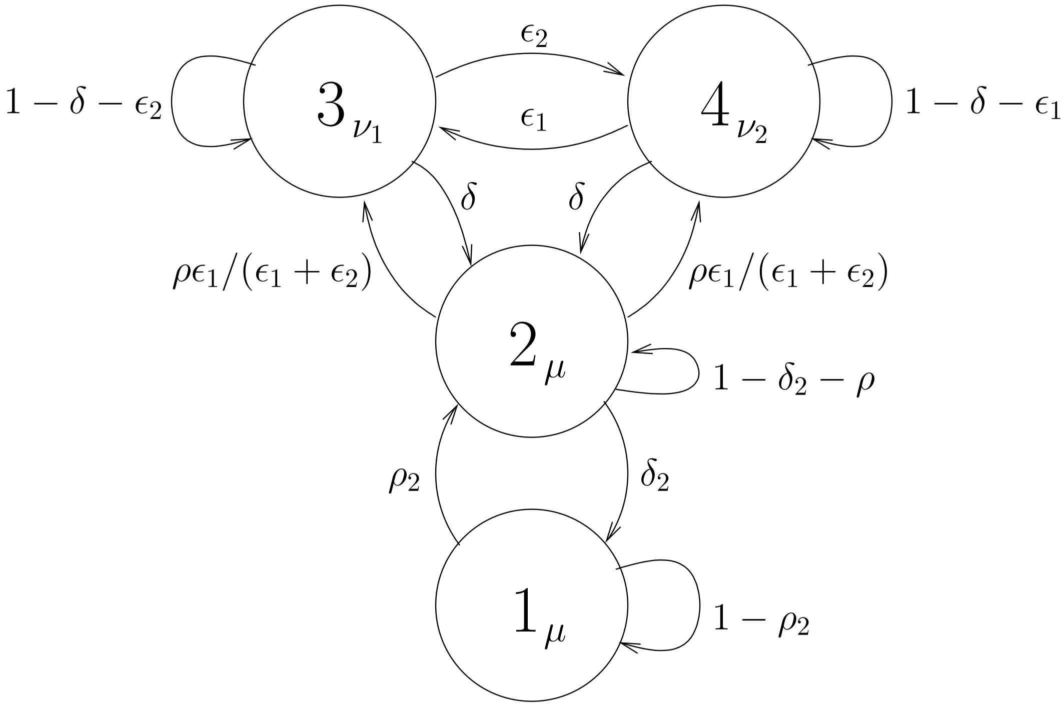
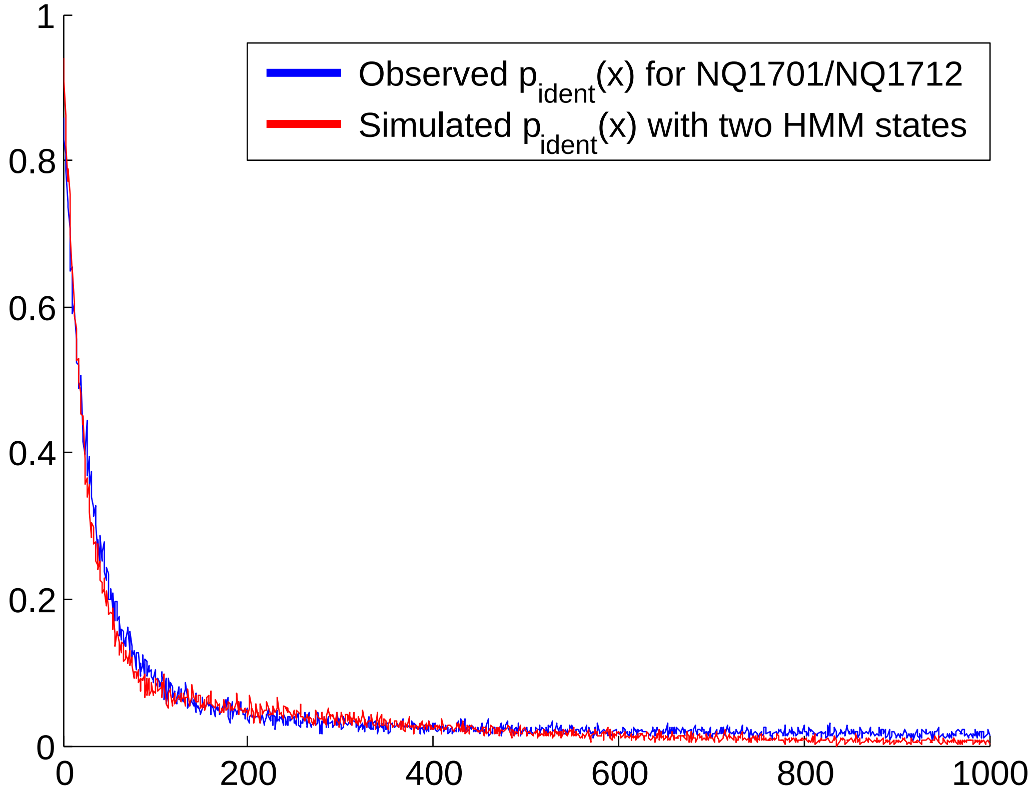

Cluster coverage
H1
008A3
100%
100%
Total coverage among sequences: 100%
H1
008A3
J99
96.0%
96.0%
91.5%
Total coverage among all sequences: 94.4%
NQ315
NQ1712
97.7%
99.5%
Total coverage among sequences: 98.6%
NQ315
NQ1712
J99
92.4%
94.1%
90.4%
Total coverage among all sequences: 92.3%
NQ352
NQ1701
99.8%
99.9%
Total coverage among sequences: 99.8%
NQ352
NQ1701
J99
95.3%
95.2%
95.3%
Total coverage among all sequences: 95.3%
NQ392
NQ1707
NQ4060
100%
99.9%
100%
Total coverage among all sequences: 100%
NQ392
NQ1707
NQ4060
J99
95.8%
95.8%
95.8%
96.4%
Total coverage among all sequences: 95.9%
NQ367
NQ1671
NQ4191
99,8%
99,7%
99,8%
Total coverage among all sequences: 99.8%
NQ367
NQ1671
NQ4191
J99
96,1%
95,8%
95,9%
94,9%
Total coverage among all sequences: 95.7%

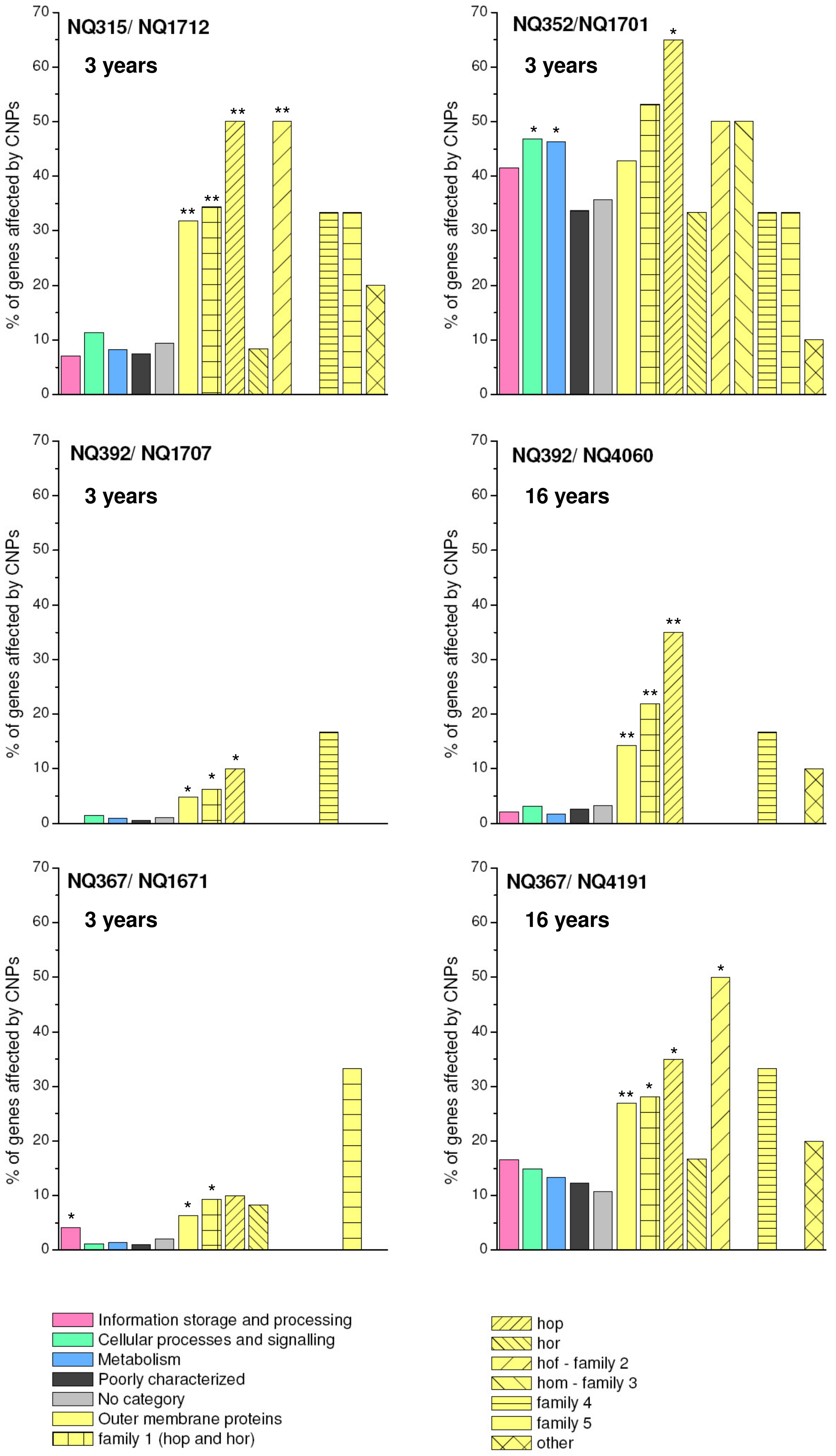


Length of imports [bp]
8000
6000
4000
2000
0
NQ315/ NQ1712
NQ352/ NQ1701
NQ392/ NQ1707
NQ392/ NQ4060
NQ367/ NQ1671
NQ367/ NQ4191


A
Line number
5
10
15
20
25
30
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
Position in line
B
Line number
5
10
15
20
25
30
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
Position in line


ϵ_2
$1-\delta-\epsilon_2$
3_{ν_1}
ϵ_1
4_{ν_2}
$1-\delta-\epsilon_1$
δ
δ
$\rho\epsilon_1/(\epsilon_1+\epsilon_2)$
$\rho\epsilon_1/(\epsilon_1+\epsilon_2)$
2_μ
$1-\delta_2-\rho$
ρ_2
δ_2
1_μ
$1-\rho_2$


1
0.8
0.6
0.4
0.2
0
0
200
400
600
800
1000
Observed $p_{ident}(x)$ for NQ1701/NQ1712
Simulated $p_{ident}(x)$ with two HMM states


NQ315/ NQ1712
3 years
NQ352/NQ1701
3 years
NQ392/ NQ1707
3 years
NQ392/ NQ4060
16 years
NQ367/ NQ1671
3 years
NQ367/ NQ4191
16 years
% of genes affected by CNPs
0
10
20
30
40
50
60
70
*
**
Information storage and processing
Cellular processes and signalling
Metabolism
Poorly characterized
No category
Outer membrane proteins
family 1 (hop and hor)
hop
hor
hof - family 2
hom - family 3
family 4
family 5
other

Supplementary Text: Supplementary Figure Legends

Fig. S1:

Alignments of paired isolates with each other and with the reference genome of *H. pylori* J99. Alignments were generated with M-GCAT. Grey bars represent the genomes, while green blocks indicate MUM-clusters (38).

Fig. S2:

Lengths of the clusters of nucleotide polymorphisms (CNPs) in isolate pairs. For better readability, two columns, one depicting the length of individual imports as dots, and one with a corresponding box and whiskers plot are drawn next to each other.

The arithmetic mean is indicated by a small square (□), the 99% percentile by a cross (X), and the maximum by a short horizontal line (-).

Fig S3:

Results of the ClonalFrame analysis of the aligned genomes of sequential isolate pairs NQ315/1712 (part A) and NQ352/1701 (part B). Each line represents 50 kb of sequence. Non-aligned regions are shown in white. Regions colored in black are clonal whereas regions in green have been imported. Polymorphism between the paired genomes is shown by grey triangles.

Fig S4:

Structure of the Hidden Markov Model (HMM) used in the analysis of the paired genomes. The HMM is made of 4 states represented by the circles. State 1 corresponds to clonal regions of the genomes, state 2 corresponds to clonal regions surrounding imports, and states 3 and 4 correspond to imported regions (which are well modelled using two HMM states, as shown in Fig. S5). The arrows represent the transitions from one state to another and the values next to the arrows represent the rate at which these transitions occur.

Fig S5:

The blue curve shows the observed $p_{ident}(x)$ function between the unrelated genomes NQ1701 and NQ1712. The red curve shows the best fit to the blue curve of a two-states Hidden Markov Model. Such a model has four parameters (the two rates of substitutions in each of the two states and the two rates of transitions from one state to another) and the values of these parameters were chosen to give the best fit of the simulated red curve to the observed blue curve.

Fig S6:

Proportion of genes affected by imports for different functional classes of encoded proteins. Import frequency for a category is calculated as the percentage of genes affected by imports. Color-coding is shown below. Asterisks indicate significantly increased frequency of imports (Fisher's exact test; * $p < 0.05$; ** $p < 0.01$). The composition of the group of outer membrane proteins (yellow bars) and the subfamilies (hatched yellow bars) is based on ref. (13).

SI Materials and Methods

DNA preparation. Whole genomic DNA for 454 sequencing was isolated using Qiagen Genomic-tip 100/G columns and the Genomic DNA Buffer Set (Qiagen). Small scale preparations of genomic DNA for sequencing of individual gene fragments were performed with the DNAeasy Blood & Tissue kit (Qiagen).

454 sequencing. Samples of genomic DNA were treated following the protocols provided by the manufacturer. The process is summarized briefly as follows: 3-5 µg of DNA were fragmented by nebulization. The DNA was then purified over MinElute columns following the PCR clean-up protocol from Qiagen. For FLX samples, an Ampure bead purification step was performed in order to remove small fragments and a quality assessment was performed with the Agilent Bioanalyser (DNA Chip7500). The purified, fragmented DNA was polished with T4 DNA Polymerase and T4 Polynucleotide Kinase (New England Biolabs, Beverly, MA/ Library Preparation Kit, Roche). After a half-hour reaction with kinase, the DNA fragments were purified again over MinElute columns, before blunt-end ligation (New England Biolabs Quick Ligation Kit/ Library Preparation Kit, Roche) to 454's proprietary double-stranded DNA adaptors A and B. The ligated DNA fragments were then bound to magnetic Streptavidin-coated beads (Dynal Biotech, Oslo, Norway/ Library Preparation Kit, Roche). Because of the non-phosphorylated 5'adaptor ends, the ligated fragments contain nicks, which were displaced with *Bst* polymerase (New England Biolabs/ Library Preparation Kit, Roche). Single-stranded DNA was melted away from the beads with 0.125 M NaOH and the single-stranded fragments were purified over Qiagen MinElute columns. The resulting purified DNA library was run on an Agilent 2100 BioAnalyzer using a RNA 6000 Pico LabChip® (Palo Alto, CA) to quantify the single stranded DNA concentration, ranging from 200-500 bases (GS20) or from 400-800 bases (FLX). The library was combined with a solution containing emulsion PCR reagents and DNA capture beads provided by Roche 454 (Branford, CT) and amplification was carried out as described elsewhere (9). After amplification, emulsion breaking, and enrichment of the library beads, the non-covalently bound strand was melted away, a sequencing primer was annealed to the covalently bound, amplified strands, and the primed DNA beads were sequenced on the Roche-454 Sequencing System. Raw data processing and assembly were performed using proprietary software provided by Roche/454 as listed in Table S1.

PCR and Sanger sequencing. For Sanger sequencing, gene fragments were obtained by PCR amplification as previously described (36) and PCR products were purified using the QIAquick PCR Purification kit (Qiagen). Double-stranded automated sequencing of PCR amplicons was performed with the Big Dye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) and a 3130*xl* capillary sequencer (Applied Biosystems).

Generation of virtual genomes. Virtual genomes were generated in order to obtain a dataset more convenient to handle than a large collection of contigs. For this purpose an assembly was performed with the Kodon software either using the contigs of one isolate or using all contigs of a pair of isolates. Additionally the published genome of *H. pylori* J99 (11) was used as a scaffold, conserving the integrity of the preassembled contigs and the contig order based on the pair wise assemblies created with the Kodon software. Colinearity of virtual genomes was checked with the programmes M-GCAT and Mauve (37, 38).

In the NQ392/1707/4060 dataset, some intra-isolate variability was detected, reflected by small contigs with variant sequences that were assembled to the same position in the virtual genome. In most cases, the contig with the higher similarity to the other isolates was included in the virtual genome. In the remaining cases, the contig with more reads was used for the virtual genome. Differences between the isolates reflecting this intra-isolate variability are summarized in Table S14.

Pair wise analysis of sequential isolates. The pair wise analysis of sequential isolates was performed using Kodon (Applied Maths) and Mauve (38). A list of polymorphisms exported from the software was used as a basis for the analysis, which was then curated manually. Indels in homopolymeric nucleotide tracts and 1 nucleotide indels were excluded from the analysis. Clusters of nucleotide polymorphisms (CNPs) were defined as groups of at least two polymorphisms with a distance of less than 200 bp between two consecutive polymorphic sites, flanked by identical sequences of at least 200 bp. Sequence differences not categorized as SNP or CNP are listed in Tables S15a-d.

The statistical model. We developed a Hidden Markov Model (HMM; ref. 39), the details of which are shown in Fig. S4. This model assumed that an alignment of two sequentially isolated genomes was made of parts that evolved clonally (states 1 and 2 in Fig. S4) and parts where recombination happened (states 3 and 4 in Fig. S4).

The differences observed between the two genomes in the clonal parts were caused by mutations and were assumed to be uniformly distributed. The substitutions observed in the recombined parts were introduced by recombination and their distribution was assumed to be the same as for two unrelated genomes, following ref. (4). This distribution can be characterized by the $p_{ident}(x)$ function, ie. the probability that an import of size x does not result in any substitution (4). The $p_{ident}(x)$ function was learnt by comparison of the two unrelated genomes NQ1701 and NQ1712, and was found (Fig. S5) to be well modelled using a mixture of two HMM states with different rates of polymorphism (labelled 3 and 4 in Fig. S4). Furthermore, our model allowed the recombined regions to be grouped through the use of an additional intermediate HMM state (labelled 2 in Fig. S4) occurring between clonal and recombined regions.

Bayesian inference. We performed Bayesian inference under the model described above by embedding the HMM into a Monte-Carlo Markov Chain (MCMC; (40, 41). Each iteration of the MCMC included a forward-backward move of the HMM (39) and a Gibbs update (42) for each of the parameters. The MCMC was run for 100,000 iterations. The first half was discarded as burn-in and the second half was sampled every 100 iterations to produce a posterior sample of size 500 for each parameter and hidden state. Convergence and mixing of the MCMC was verified by comparing runs with different starting conditions manually and using the method of Gelman and Rubin (43).

Assignment of imports to functional gene categories. The assignment of the location of imports to functional gene categories was mainly based on an automatic annotation of the virtual genomes generated with Kodon using the published genome annotation of *H. pylori* J99 as a reference. Ambiguities were resolved manually (e.g. when different annotation results for a CNP were obtained for the two isolates). These discrepancies could in most cases be attributed to disruptions of the ORF in one strain due to frameshifts within homopolymeric tracts. Predicted ORFs affected by CNP with no homologue in J99 were manually annotated by means of BLAST searches. COG classifications were used to assign genes affected by CNPs into functional groups. COG classifications were extracted from the *H. pylori* J99 GenBank file (as of May 2010), or, in rare cases, from the annotations of other published *H. pylori* genomes. If affected genes belonged to more than one functional group, they were counted once for each category.

Since frameshifts in homopolymeric repeats prevented a reliable automatic annotation of the virtual genomes, J99 was chosen as reference genome to calculate the total combined length and number of genes belonging to one functional category, which was used as the denominator in calculations of import frequencies. Genes assigned to more than one functional category were included in multiple categories. Since the outer membrane proteins of *H. pylori* have in most cases not been assigned to any functional category they have been grouped into an additional "functional group", based on a previous detailed analysis of all *H. pylori* outer membrane proteins (13).

References:

for References 1-36, see Main Article

37. Treangen TJ, Messeguer X (2006) M-GCAT: interactively and efficiently constructing large-scale multiple genome comparison frameworks in closely related species. *BMC Bioinformatics* 7:433.

38. Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394-1403.

39. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* 77:257-286.

40. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machines. *J Chem Phys* 21:1087-1091.

41. Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97-109.

42. Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans Pattern Analysis and Machine Intelligence* 6:721-741.

43. Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Stat Sci* 7:457-511.

Table S1: General features of the genomes analysed in this study compared with previously released finished *H. pylori* genomes

Isolate	Number of contigs in virtual genomes*	Size range of contigs (base pairs)	Size (base pairs)	Sequencing technology/ assembly software/ Reference
26695	-	-	1,667,867	(Tomb et al., 1997)
J99	-	-	1,643,831	(Alm et al., 1999)
HPAG1	-	-	1,596,366	(Oh et al., 2006)
Shi470	-	-	1,608,548	(Kersulyte et al., 2009)
G27	-	-	1,652,982	(Baltrus et al., 2009)
P12	-	-	1,673,813	(Fisher et al., 2010)
B38	-	-	1,576,758	(Thiberge et al., 2010)
H1	31	766 - 339,083	1,548,400	GS 20/ Newbler 1.0
8A3	44	631 - 339,064	1,547,179	GS 20/ Newbler 1.0
NQ315	57	610 - 142,862	1,601,938	GS 20/ Newbler 1.0
NQ1712[a]	62	636 - 288,173	1,572,582	GS 20/ Newbler 1.0
NQ352	61	705 - 167,051	1,639,151	GS 20/ Newbler 1.0
NQ1701[a]	78	715 - 106,414	1,639,623	GS 20/ Newbler 1.0
NQ392	70	102 - 206,825	1,650,953	FLX/ Newbler 2.0
NQ1707[a]	79	102 - 133,105	1,650,485	FLX/ Newbler 2.0
NQ4060[b]	59	104 - 215,738	1,649,736	FLX/ Newbler 2.0
NQ367	90	101 - 97,875	1,622,411	FLX/ Newbler 2.0
NQ1671[a]	38	105 - 424,440	1,626,950	FLX/ Newbler 2.0
NQ4191[b]	42 + 1**	100 - 192,652	1,624,247 + 4,818**	FLX/ Newbler 2.0

*	for GS 20 datasets only large contigs (>500bp) were used to generate virtual genomes
**	indicates a plasmid in NQ4191
a	isolated 3 years later than first isolate
b	isolated 16 years later than first isolate

Table S2: Isolated (non-clustered) single nucleotide polymorphisms (SNPs) in strain set NQ392/1707/4060. Positions refer to the position in the virtual genome of the first isolate NQ392.

SNPs			Annotation[a]						
No NQ1707	No NQ4060	Position in NQ392	Description	gene	HP-number	jhp-number	COG	OMP	type of mutation
	1	4989	type III rstriction enzyme	res_2	1521	1410	V		non-synonymous
	2	23927	type I restriction enzyme restriction subunit	hsdR_3	1402	1424	V		synonymous
	3	26130	hypothetical protein	-	1401	1425	R		non-synonymous
	4	39606	no annotation/ intergenic						intergenic
1	5	42126	hypothetical protein	-	1384	1441	-		synonymous
	6	49827	NA$^+$/H$^+$ antiporter 1	nhaA	1552	1447	P		synonymous
	7	50387	preprotein translocase subunit YajC	yajC	1551	1448	U		synonymous
	8	66523	no annotation/ intergenic						intergenic
	9	82435	putative outer membrane protein	-	1571	1479	M		non-synonymous
	10	99012	2-dehydro-3-deoxyphosphooctonate aldolase	kdsA	0003	0003	M		synonymous
	11	99437	carbonic anhydrase	icfA	0004	0004	P		non-synonymous
2		103155	putative outer membrane protein	-	0009	0007	-	hop	synonymous
3		103403	putative outer membrane protein	-	0009	0007	-	hop	non-synonymous
4	X	103801	putative outer membrane protein	-	0009	0007	-	hop	non-synonymous
5	X	104180	putative outer membrane protein	-	0009	0007	-	hop	non-synonymous
	12	114458	hypothetical protein	-	0018	0016	S		non-synonymous
	13	118477	lipid A phosphoethanolamine transferase	-	0022	0020	R		non-synonymous
	14	132112	ATP-dependent protease, ATP-binding subunit	clpA	0033	0029	O		non-synonymous
	15	142782	hydrogenase expression/ formation protein	hypE	0047	0040	O		synonymous
	16	166564	hypothetical protein	-	0066	0061	D		synonymous
	17	168282	urease accessory protein	ureG	0068	0063	OK		non-synonymous
	18	201686	hypothetical protein	-	0100	0092	S		synonymous
	19	205996	methyl-accepting chemotaxis protein (MCP)	-	0103	0095	NT		non-synonymous
	20	232258	hypothetical protein similar to HPG27_118				-		synonymous
6		242763	L-lactate permease	lldP_1	0140	0128	C		synonymous
	21	252500	hypothetical protein	-	0150	0138	-		non-synonymous
	22	252772	hypothetical protein	-	0150	0138	-		synonymous
	23	271343	molybdopterin biosynthesis protein	moeA	0172	0158	H		synonymous
	24	280539	hypothetical protein	-		0165	-		non-synonymous
	25	283061	apolipoprotein N-acyltransferase	lnt	0180	0168	M		synonymous
	26	338498	hypothetical protein	-	0233	0218	E		synonymous
7		352114	ATP-dependent RNA helicase DeaD	deaD	0247	0232	LKJ		non-synonymous
	27	395201	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		non-synonymous
	28	416372	peptide ABC transporter ATP-binding protein	dppF	0302	0287	R		synonymous
	29	431110	outer membrane protein-adhesin	babA	1243	0833	-	hop	synonymous
	30	437785	poly E-rich protein	-	0322	0305	-		synonymous
	31	442803	tetraacyldisaccharide 4'-kinase	lpxK	0328	0311	M		synonymous
	32	488241	copper-tranporting P-type ATPase	copA	1072	0353	P		non-synonymous
	33	492006	ribosomal protein L11 methyltransferase	prmA	1068	0357	J		non-synonymous
	34	506775	translation initiation factor IF-2	infB	1048	0377	J		non-synonymous
	35	546340	no annotation/ intergenic						intergenic
	36	550434	no annotation/ intergenic						intergenic
	37	578202	no annotation/ intergenic						intergenic
	38	582071	putative potassium channel protein	-	0490	0442	P		non-synonymous
	39	607495	ATP-dependent protease peptidase subunit	hslV	0515	0464	O		synonymous
	40	622821	cag island protein	orf13/14	0527	0476	U		non-synonymous
	41	627256	cag island protein	orf16	0529	0478	-		synonymous
	42	634027	cag island protein	cagM	0537	0485	-		non-synonymous
	43	639209	cag island protein	cagF	0543	0491	-		non-synonymous
	44	643103	cag island protein	cagC	0546	0494	-		synonymous

SNPs			Annotation[a]						
No NQ1707	No NQ4060	Position in NQ392	Description	gene	HP-number	jhp-number	COG	OMP	type of mutation
	45	670732	hypothetical protein	-	0575	0522	R		non-synonymous
	46	692326	penicillin-binding protein	-	0597	0544	M		non-synonymous
	47	696077	methyl-accepting chemotaxis protein (MCP)	-	0599	0546	NT		non-synonymous
	48	741680	hydrogenase, small subunit	hyaA	0631	0574	C		non-synonymous
	49	748043	putative outer membrane protein	-	0638	0581	-	hop	synonymous
	50	765780	protective surface antigen D15	-	0655	0600	M	other	synonymous
	51	770652	hypothetical protein	-	0659	0604	-		non-synonymous
	52	789244	putative, similar to jhp0914		-	-			synonymous
8		793478	putative lipopolysaccharide biosynthesis protein	-	0679	0620	R		non-synonymous
	53	794519	ribonucleoside-diphosphate reductase subnit alpha	nrdA	0680	0621	F		non-synonymous
	54	815312	3-oxoacid CoA-transferase, subunit B	scoB	0692	0636	I		non-synonymous
	55	826928	exinuclease ABC subunit A	uvrA	0705	0644	L		non-synonymous
	56	834649	hypothetical protein	-	0712/0713	0651	S		non-synonymous
	57	842109	putative outer membrane protein	-	0722	0659	-	hop	non-synonymous
	58	843257	anaerobic C4-dicarboxylate membrane transporter	dcuA	0724	0660	R		synonymous
	59	848057	putative outer membrane protein	-	0726	0663	-	other	synonymous
	60	858765	no annotation/ intergenic						intergenic
	61	864865	putative rod shape-determining protein	rodA_1	0743	0680	D		synonymous
	63	865362	putative rod shape-determining protein	rodA_1	0743	0680	D		non-synonymous
9		872082	hypothetical protein	-	0750	0687	D		non-synonymous
	63	877246	hypothetical protein	-	0757	0694	R		non-synonymous
	64	886675	hypothetical protein	-	0766	0704	-		non-synonymous
	65	887775	no annotation/ intergenic						intergenic
10		890307	flagellar biosynthesis protein FlhB	flhB	0770	0707	NU		synonymous
	66	894215	tyrosyl-tRNA synthetase	tyrS	0774	0711	J		non-synonymous
	67	904406	putative outer membrane protein	-	0782	0719	-	hof	non-synonymous
	68	905358	hypothetical protein	-	0783	0720	-		non-synonymous
	69	931368	flagellar basal body-associated protein FliL	fliL	0809	0745	N		non-synonymous
11		950800	dephospho-CoA kinase	coaE	0831	0770	H		non-synonymous
	70	972340	ABC transporter ATP-binding protein	-	0853	0789	R		non-synonymous
	71	973745	guanosine 5'-monophosphate oxidoreductase	guaC	0854	0790	F		non-synonymous
12		991398	putative iron-regulated outer membrane protein	frpB_1	0876	0810	P	family 4	non-synonymous
	72	991710	putative iron-regulated outer membrane protein	frpB_1	0876	0810	P	family 4	non-synonymous
	73	1006218	putative lipopolysaccharide biosynthesis protein	-		0820	M		non-synonymous
	74	1019447	phosphotransacetylase	pta	0904/0905	0841	C		synonymous
	75	1029645	outer membrane protein/ porin	hopC	0912	0848	-	hop	non-synonymous
	76	1064533	putative sodium/alanine symporter	-	0942	0877	E		synonymous
	77	1064761	putative sodium/alanine symporter	-	0942	0877	E		synonymous
	78	1065616	D-amino acid dehydrogenase	dadA	0943	0878	E		non-synonymous
	79	1074552	hypothetical protein	-	0953	0887	-		non-synonymous
13	**80**	1075188	putative aldehyde dehydrogenase	-	0954	0888	C		non-synonymous
	81	1081561	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase	gpsA	0961	0895	C		non-synonymous
	82	1089056	putative cation efflux system protein	czcA_1	0969	0903	P		synonymous
	83	1096768	hypothetical protein	-	0973	0907	-		non-synonymous
	84	1119086	hypothetical protein	-		0948	-		synonymous
	85	1138219	putative paralog of HpaA	-	0410	0971	-		non-synonymous
	86	1150734	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	ispH	0400	0981	IM		synonymous
	87	1158249	hypothetical protein	-	0394	0987	S		non-synonymous
	88	1168535	hypothetical protein	-	0383	0998	-		non-synonymous

SNPs			**Annotation**[a]						
No NQ1707	**No NQ4060**	**Position in NQ392**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**type of mutation**
14		1187083	spore coat polysaccharide biosynthesis protein C	-	0366	1015	M		non-synonymous
	89	1197693	phosphogluconate dehydratase	edd	1100	1026	EG		non-synonymous
	90	1200411	glucose-6-phosphate 1-dehydrogenase	zwf	1101	1027	G		non-synonymous
	91	1203651	putative lipopolysaccharide biosynthesis protein	-	1105	1031	M		non-synonymous
15	**92**	1207649	putative outer membrane protein	-	1107	1034	-	hor	synonymous
	93	1210318	pyruvate flavodoxin oxidoreductase subunit alpha	porA	1110	1037	C		synonymous
	94	1210945	ferrodoxin oxidoreductase beta subunit	porB	1111	1038	C		synonymous
16	**95**	1216199	hypothetical protein	-	1115	1042	-		non-synonymous
17		1219608	gamma-glutamyltranspeptidase	ggt	1118	1046	E		synonymous
	96	1243538	no annotation/ intergenic						intergenic
	97	1244795	no annotation/ intergenic						intergenic
18		1244954	no annotation/ intergenic						intergenic
	98	1245489	no annotation/ intergenic						intergenic
	99	1257928	putative outer membrane protein	-	1156	1083	-	hop	non-synonymous
	100	1259000	putative outer membrane protein	-	1157	1084	-	hop	non-synonymous
	101	1269777	no annotation/ intergenic						intergenic
	102	1278754	glucose/galactose transporter	gluP	1174	1101	G		synonymous
	103	1279384	hypothetical protein	-	1175	1102	R		non-synonymous
	104	1281948	putative outer membrane function	-	1177	1103	-	hop	non-synonymous
	105	1289074	no annotation/ intergenic						intergenic
19	**106**	1303176	elongation factor G	fusA	1195	1118	J		synonymous
20	**107**	1316754	transcription antitermination protein NusG	nusG	1203	1126	K		synonymous
	108	1322600	type II DNA modification enzyme (methyltransferase)	M.HpyI	1208	1131	L		synonymous
21	**109**	1324064	O-serine acetyltransferase	cysE	1210	1133	E		non-synonymous
	110	1342558	DNA polymerase III subunit delta'	holB	1231	1152	L		synonymous
22		1364004	putative peptide ABC transporter, ATP-binding protein	-	1251	1172	R		non-synonymous
	111	1390031	tryptophan synthase subunit alpha	trpA	1277	1198	E		synonymous
	112	1437123	putative outer membrane protein	-	1342	1261	-	hop	synonymous
	113	1447571	type II DNA modification enzyme (methyltransferase)	-	1352	1271	L		synonymous
	114	1451354	hypothetical protein	-	1354	1272	R		non-synonymous
	115	1469310	rod shape-determining protein MreB	mreB	1373	1287	D		non-synonymous
	116	1482965	putative type III restriction enzyme	res_1		1297	S		synonymous
	117	1485853	no annotation/ intergenic						intergenic
	118	1510541	hypothetical protein	-		0928	KL		non-synonymous
23	**119**	1535257	no annotation/ intergenic						intergenic
	120	1549909	no annotation/ intergenic						intergenic
	121	1559283	hypothetical protein	-	1430	1323	R		synonymous
24		1563588	no annotation/ intergenic						intergenic
	122	1569406	hypothetical protein	-	1437	1330	-		synonymous
	123	1577407	putative inner membrane protein translocase component YidC	-	1450	1343	U		non-synonymous
	124	1599375	DNA polymerase I	polA	1470	1363	L		synonymous
	125	1604226	putative type II modification enzyme (methyltransferase)	-	1472	1365	V		synonymous
25		1615479	hypothetical protein	-	1484	1377	S		synonymous
	126	1626025	UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase	murE	1494	1387	M		synonymous
26	**127**	1633569	putative component of cation transport for cbb3-type oxidase	fixI	1503	1396	P		non-synonymous
	128	1635696	hypothetical protein	-	1504	1397	R		non-synonymous
	129	1647093	transcription elongation factor NusA	nusA	1514	1407	K		synonymous
27		1647788	hypothetical protein	-		1408	-		synonymous

SNPs			Annotation[a]						
No NQ1707	No NQ4060	Position in NQ392	Description	gene	HP-number	jhp-number	COG	OMP	type of mutation

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S3: Isolated (non-clustered) single nucleotide polymorphisms (SNPs) in strain set NQ367/1671/4191. Positions refer to the position in the virtual genome of the first isolate NQ367.

SNPs			Annotation						
No NQ1671	No NQ4191	Position in NQ367	Description	gene	HP-number	jhp-number	COG	OMP	type of mutation
	1	11923	leucyl-tRNA synthetase	leuS	1547	1452	J		non-synonymous
	2	15829	hypothetical protein	-	1542	1457	M		non-synonymous
1		17286	transcription-repair coupling factor	mfd	1541	1458	LK		synonymous
	3	20440	ubiquinol cytochrome c oxidoreductase, cytochrome b subunit	petB	1539	1460	C		non-synonymous
	4	36296	hypothetical protein	-	1568	1476	S		synonymous
	5	37173	hypothetical protein	-	1569	1477	-		non-synonymous
	6	39285	regulatory protein DniR	dniR	1572	1480	M		non-synonymous
2	**7**	45059	putative undecaprenyl-phosphate-alpha-N-acetylglucosaminyltransferase	wecA	1581	1488	M		non-synonymous
	8	46436	pyridoxine 5'-phosphate synthase	pdxJ	1582	1489	H		synonymous
	9	49221	no annotation/ intergenic						intergenic
	10	55185	2-dehydro-3-deoxyphosphooctonate aldolase	kdsA	0003	0003	M		non-synonymous
	11	57528	pantoate-beta-alanine ligase	panC	0006	0006	H		non-synonymous
	12	59028	no annotation/ intergenic						intergenic
	13	60558	putative outer membrane protein	-	0009	0007	-	hop	non-synonymous
	14	111764	no annotation/ intergenic						intergenic
	15	115135	hypothetical protein	-	0060	0054	-		non-synonymous
	16	145641	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	-	0089	0082	F		synonymous
	17	154817	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		synonymous
3	**18**	159238	methyl-accepting chemotaxis protein (MCP)	-	0103	0095	NT		non-synonymous
	19	160033	methyl-accepting chemotaxis protein (MCP)	-	0103	0095	NT		non-synonymous
4	**20**	184348	threonyl-tRNA synthetase	thrS	0123	0113	J		non-synonymous
	21	188710	hypothetical protein	-	0130	0119	-		non-synonymous
5		189625	no annotation/ intergenic						intergenic
6	**Pseudo-cluster**	199784	no annotation/ intergenic						intergenic
7		206396	putative cytochrome oxidase (CBB3-type)	fixQ	0146	0134	O		synonymous
8		208156	hypothetical protein	-	0149	0137	-		non-synonymous
	22	212467	phosphopyruvate hydratase	eno	0154	0142	G		non-synonymous
9	**23**	289302	hypothetical protein	-	0228	0213	P		synonymous
	24	290330	outer membrane protein/ porin	hopA	0229	0214	-	hop	synonymous
	25	295089	hypothetical protein	-	0233	0218	E		synonymous
10	**26**	311374	ABC transporter, ATP-binding protein	-	0250	0235	R		non-synonymous
	27	318073	hypothetical protein	-	0257	0241	S		non-synonymous
	28	345592	hypothetical protein	-	0284	0269	M		synonymous
	29	347181	hypothetical protein	-	0285	0270	J		synonymous
11	***CNP 37***	356364	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		non-synonymous
	30	378619	hypothetical protein	-	0309	0294	R		synonymous
	31	385238	arginyl-tRNA synthetase	argS	0319	0302	J		non-synonymous
	32	388938	poly E-rich protein	-	0322	0305	-		non-synonymous
	33	392866	CMP-N-acetylneuraminic acid synthetase (neuA)	-	0326	0309	M		synonymous
	34	405014	no annotation/ intergenic						intergenic
	35	420138	hypothetical protein	-	1089	0336	L		non-synonymous
	36	420533	hypothetical protein	-	1089	0336	L		synonymous
	37	421372	hypothetical protein	-	1089	0336	L		non-synonymous
	38	427213	putative outer membrane protein	-	1083	0342	-	hof	non-synonymous

SNPs			**Annotation**						
No NQ1671	**No NQ4191**	**Position in NQ367**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**type of mutation**
	39	433191	hypothetical protein	-	1078	0347	-		non-synonymous
12	**40**	443110	putative ribosomal protein L11 methyltransferase	prmA	1068	0357	J		non-synonymous
	41	492297	7-alpha-hydroxysteroid dehydrogenase	-	1014	0409	IQR		synonymous
	42	500479	type I restriction enzyme modification subunit	hsdM_1	0463	0415	V		non-synonymous
	43	501865	fragment of type I restriction enzyme R protein (hsdR)	-	0464	0416	V		synonymous
	44	512195	glutathione-regulated potassium-efflux system protein	kefB	0471	0423	P		non-synonymous
	45	540177	putative transporter	-	0498	0450	R		synonymous
	46	542396	putative phospholipase A1	pldA	0499	0451	M		synonymous
	47	544434	DNA gyrase subunit B	gyrB	0501	0453	L		synonymous
	48	547876	no annotation/ intergenic						intergenic
	49	559924	ATP-dependent protease ATP-binding subunit HslU	hslU	0516	0465	O		synonymous
	50	571318	no annotation/ intergenic						intergenic
	51	594556	cag island protein	cagC	0546	0494	-		non-synonymous
13	**add. diff.**	595530	no annotation/ intergenic						intergenic
	52	595993	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		non-synonymous
	53	599076	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		non-synonymous
	54	612916	hypothetical protein	-	0565	0512	S		synonymous
	55	642423	hypothetical protein	-	0595	0542	O		non-synonymous
14	**56**	650958	flagellin A	flaA	0601	0548	N		non-synonymous
15	**57**	653984	hypothetical protein	-	0605	0552	MU	family 5	non-synonymous
	58	667355	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		non-synonymous
	59	669009	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		non-synonymous
	60	676811	putative lipopolysaccharide biosynthesis protein	-	0619	0563	M		non-synonymous
	61	687422	no annotation/ intergenic						intergenic
	62	717227	hypothetical protein	-	0656	0601	HR		synonymous
	63	725724	hypothetical protein	-	0664	0609	-		non-synonymous
	64	733564	no annotation/ intergenic						intergenic
	65	734893	aspartate aminotransferase	aspB	0672	0615	E		synonymous
	66	750460	hypothetical protein	-	0688	0628	L		non-synonymous
	67	760435	hypothetical protein	-	0693	0693	-		synonymous
	68	763421	acetyl-coA acetyltransferase	thl	0690	0638	I		synonymous
	69	766721	DNA gyrase subunit A	gyrA	0701	0641	L		non-synonymous
16	**CNP 66**	787687	putative outer membrane protein	-	0725	0662	-	hop	non-synonymous
17		808881	D-alanyl-D-alanine-adding enzyme	murF	0740	0677	M		non-synonymous
	70	819323	flagellar protein FlaG	-	0751	0688	N		non-synonymous
	71	821975	flagellar protein FliS	fliS	0753	0690	NUO		non-synonymous
18		827021	phosphodiesterase	-	0760	0697	R		synonymous
	72	829014	hypothetical protein	-	0762	0699	-		synonymous
	73	830553	hypothetical protein	-	0764	0701	R		non-synonymous
19		838927	N-acetylmuramoyl-L-alanine amidase	amiA	0772	0709	M		synonymous
	74	842877	guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase	spoT	0775	0712	TK		non-synonymous
	75	852693	outer-membrane lipoprotein carrier protein	lolA	0785	0722	M		non-synonymous
	76	865702	putative outer membrane protein	-	0796	0732	-	hor	non-synonymous
	77	866581	neuraminyllactose-binding hemagglutinin precursor	hpaA	0797	0733	-		synonymous
	78	868224	molybdenum cofactor biosynthesis protein	mogA	0799	0735	H		synonymous
	79	876036	4'-phosphopantetheinyl transferase	acpS	0808	0744	I		synonymous
20	**80**	878907	hypothetical protein	-	0812	0748	S		synonymous
	81	879397	no annotation/ intergenic						intergenic

SNPs			**Annotation**						
No NQ1671	**No NQ4191**	**Position in NQ367**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**type of mutation**
	82	908468	phosphomethylpyrimidine kinase	thiD	0844	0782	H		synonymous
	83	917662	putative ABC transporter, ATP-binding protein	-	0853	0789	R		non-synonymous
	84	922779	ADP-L-glycero-D-mannoheptose-6-epimerase	gmhD	0859	0793	MG		synonymous
21		933086	CDP-diacylglycerol pyrophosphatase	cdh	0871	0805	I		non-synonymous
	85	939808	holliday junction resolvase	ruvC	0877	0811	L		non-synonymous
22		940452	hypothetical protein	-	0878/ 0879	0812	S		synonymous
	86	949124	vacuolating cytotoxin	vacA	0887	0819	-		non-synonymous
	87	951127	vacuolating cytotoxin	vacA	0887	0819	-		synonymous
23	**Contig-border**	959262	outer membrane protein-adhesin	babA	1243	0833	-	hop	non-synonymous
	88	977307	outer membrane protein/porin	hopB	0913	0849	-	hop	non-synonymous
	89	979898	putative outer membrane protein	-	0914	0850	-	hof	non-synonymous
	90	983705	no annotation/ intergenic						intergenic
24		994244	putative vacuolating cytotoxin (VacA) paralog	-	0922	0856	N		synonymous
	91	999029	recombination protein RecR	recR	0925	0859	L		non-synonymous
	92	1002625	geranyltransferase	ispA	0929	0864	H		non-synonymous
	93	1004106	hypothetical protein	-	0931/ 0932	0866	-		synonymous
	94	1006555	no annotation/ intergenic						intergenic
	95	1023274	hypothetical protein	-	0953	0887	-		non-synonymous
25		1023617	no annotation/ intergenic						intergenic
	96	1023707	putative aldehyde dehydrogenase	-	0954	0888	C		non-synonymous
	97	1052173	no annotation/ intergenic						intergenic
	98	1054107	hypothetical protein	-		0914	-		synonymous
	99	1066885	hypothetical protein	HPP12_0455			-		non-synonymous
	100	1069023	no annotation/ intergenic						intergenic
	101	1070859	hypothetical protein HP1004	-	1004		U		synonymous
26		1077997	hypothetical protein	-		0928	KL		synonymous
	102	1081350	hypothetical protein	-		0928	KL		synonymous
	103	1099954	putative polysaccharide biosynthesis protein	-	0421	0963	M		non-synonymous
	104	1103854	methionyl-tRNA synthetase	metG	0417	0967	J		non-synonymous
	105	1122340	30S ribosomal protein S1	rpsA	0399	0982	J		non-synonymous
27		1125061	D-3-phosphoglycerate dehydrogenase	serA	0397	0984	HE		non-synonymous
28		1126527	hypothetical protein	-	0396	0985	H		synonymous
	106	1160510	protein-L-isoaspartate O-methyltransferase	pcm	0363	1017	O		non-synonymous
29	**107**	1163812	no annotation/ intergenic						intergenic
30		1177716	hypothetical protein	-	1106	1033	-		synonymous
	108	1203311	F0F1 ATP synthase subunit gamma	atpG	1133	1061	C		non-synonymous
31	**109**	1205901	F0F1 ATP synthase subunit B	atpF	1136	1064	C		synonymous
	110	1218988	no annotation/ intergenic						intergenic
32	**111**	1227398	putative outer membrane protein	-	1156	1083	-	hop	synonymous
	112	1230665	putative outer membrane protein	-	1157	1084	-	hop	synonymous
	113	1231365	putative outer membrane protein	-	1157	1084	-	hop	non-synonymous
	114	1237196	putative thioredoxin reductase	trxB_2	1164	1091	O		non-synonymous
	115	1241924	putative outer membrane protein	-	1167	1094	-	hof	non-synonymous
33	**116**	1246888	amino acid ABC transporter, ATP-binding protein	-	1171	1098	E		non-synonymous
	117	1257992	phosphopentomutase	deoB	1179	1105	G		non-synonymous
	118	1270337	ADP-heptose-LPS-heptossyltransferase II	waaF	1191	1116	M		synonymous
34	***CNP 126***	1278580	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		non-synonymous
	119	1283288	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		non-synonymous
	120	1284040	50S ribosomal protein L7/L12	rplL	1199	1122	J		synonymous

SNPs			Annotation						
No NQ1671	No NQ4191	Position in NQ367	Description	gene	HP-number	jhp-number	COG	OMP	type of mutation
	121	1286374	transcription antitermination protein NusG	nusG	1203	1126	K		synonymous
	122	1286689	preprotein translocase subunit SecE - HPAG1_1143	secE			U		non-synonymous
	123	1297830	hypothetical protein	-	1214	1137	R		non-synonymous
	124	1299778	putative role in outermembrane permeability	imp	1215/ 1216	1138	M		non-synonymous
	125	1317555	carbamoyl phosphate synthase small subunit	carA	1237	1158	EF		non-synonymous
	126	1335123	hypothetical protein	-	1252	1173	E		non-synonymous
	127	1348212	NADH dehydrogenase subunit H	nuoH	1267	1188	C		non-synonymous
	128	1354215	NADH dehydrogenase subunit N	nuoN	1273	1194	C		non-synonymous
	129	1371435	hypothetical protein	-	1288	1208	-		non-synonymous
	130	1374578	50S ribosomal protein L17	rplQ	1292	1212	J		non-synonymous
	131	1376505	30S ribosomal protein S11	rpsK	1295	1215	J		non-synonymous
35		1382435	50S ribosomal protein L5	rplE	1307	1227	J		non-synonymous
	132	1385746	50S ribosomal protein L2	rplB	1316	1236	J		synonymous
36		1389178	hypothetical protein	-	1321	1241	R		non-synonymous
	133	1393395	hypothetical protein	-	1327	1247	MU	family 5	non-synonymous
	134	1394004	hypothetical protein	-	1327	1247	MU	family 5	non-synonymous
	135	1405758	no annotation/ intergenic						intergenic
	136	1429530	DNA transfer protein	comEC	1361	1279	R		non-synonymous
	137	1455983	type I restriction enzyme (restriction subunit)	hsdR_3	1402	1424	V		non-synonymous
37	**138**	1461826	no annotation/ intergenic						intergenic
38	**139**	1465362	hypothetical protein	-	1530	1418	F		synonymous
	140	1468129	hypothetical protein	-	1527	1416	-		synonymous
	141	1473098	type III DNA modification enzyme (methyltransferse)	mod_2	1522	1411	L		synonymous
	142	1477319	type III restriction enzyme	res_2	1521	1410	V		synonymous
	143	1481089	putative type II DNA modification enzyme (methyltransferase)	-		1409	V		non-synonymous
	144	1484702	selenocysteine synthase	selA	1513	1406	E		non-synonymous
	145	1490103	ferrodoxin-like protein	-	1508	1401	C		non-synonymous
	146	1490533	ferrodoxin-like protein	-	1508	1401	C		non-synonymous
	147	1492251	sodium/glutamate symporter	gltS	1506	1399	E		non-synonymous
	148	1495962	putative component of cation transport for cbb3-type oxidase	fixI	1503	1396	P		non-synonymous
39		1497932	putative outer membrane protein	-	1501	1394	-	hor	synonymous
	149	1503338	transaldolase	tal	1495	1388	G		non-synonymous
	150	1512694	hypothetical protein	-	1486	1379	-		synonymous
	151	1513523	hypothetical protein	-	1486	1379	-		non-synonymous
	152	1527063	predicted ORF/ putative methylase				-		non-synonymous
40	**153**	1533559	branched-chain amino acid aminotransferase	ilvE	1468	1361	EH		synonymous
	154	1560192	hypothetical protein	-	1409	1301	S		synonymous
	155	1590101	hypothetical protein	-	1438	1331	-		synonymous
	156	1595407	no annotation/ intergenic						intergenic
	157	1597132	putative ribonuclease N	rbn	1407	1299	S		non-synonymous

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S4: Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ392/1707. Positions refer to the position in the virtual genome of the first isolate NQ392.

	Position in NQ392			**Annotation[a]**						
No	**start**	**end**	**Corresponding import in NQ4060**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
1	78038	78162		penicillin-binding protein	-	1565	1473	M		125
2	268448	268700		hypothetical protein	-	0170	0156	-		253
3	845988	846022		putative outer membrane protein	-	0725	0662	-	hop	35
4	846387	846389	CNP 22	putative outer membrane protein	-	0725	0662	-	hop	3
5	846996	847011	CNP 23	putative outer membrane protein	-	0725	0662	-	hop	16
6	913426	913629	*CNP 24 (partially)*	anti-codon nuclease masking agent	prrB	0790		V		204
7	1355279	1355281		outer membrane protein-adhesin	babB	0896	1164	-	hop	3
8	1616862	1616866		hypothetical protein	-	1486	1379	-		5
9	1619767	1621653		lipase-like protein	-	1489	1382	MU		1887
				hypothetical protein	-	1490	1383	R		
10	1621866	1622096		hypothetical protein	-	1490	1383	R		231
11	1622374	1622436		phosphate permease	-	1491	1384	P		63
12	1622646	1623336		phosphate permease	-	1491	1384	P		691
13	1639547	1639550		ferrodoxin-like protein	-	1508	1401	C		4
14	1644421	1644864		putative iron regulated outer membrane protein	frpB_3	1512	1405	P	family 4	444
15	1645672	1645677		selenocysteine synthase	selA	1513	1406	E		6
16	1650888	1651023		putative type II DNA modification enzyme (methyltransferase)	-		1409	V		136

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S5: **Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ367/1671. Positions refer to the position in the virtual genome of the first isolate NQ367.**

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ4191**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
1	184565	184568	CNP 15	translation initiation factor IF-3	infC	0124	0114	J		4
2	184985	185277	CNP 16	translation initiation factor IF-3	infC	0124	0114	J		293
				50S ribosomal protein L35	rpmI	0125	0115	J		
3	185526	186814	CNP 17	50S ribosomal protein L20	rplT	0126	0116	J		1289
				putative outer membrane protein	-	0127	0117	-	hor	
4	200401	200941	*CNP 19 (partially)*	L-lactate permease	lldp_2	0141	0129	C		541
5	201337	203173	*CNP 20 (partially)*	L-lactate permease	lldp_2	0141	0129	C		1837
				DNA glycosylase MutY	mutY	0142	0130	L		
				hypothetical protein	-	0143	0131	P		
6	203395	203470	*CNP 21 (partially)*	hypothetical protein	-	0143	0131	P		76
7	207219	207815	CNP22	cytochrome oxidase (CBB3-type)	fixP	0147	0135	C		597
				hypothetical protein	-	0148	0136	-		
				hypothetical protein	-	0149	0137	-		
8	216997	217075	*CNP23(partially)*	hypothetical protein	-	0160	0148	R		79
9	517734	517776	-	putative outer membrane protein	-	0923	0857	-	hop	43
10	597023	597257	*CNP 53 (partially)*	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		235
11	618437	619427	-	leucyl aminopeptidase	pepA	0570	0517	E		991
12	652785	653327	-	uroporphyrinogen decarboxylase	hemE	0604	0551	H		543
13	653538	653775	-	uroporphyrinogen decarboxylase	hemE	0604	0551	H		238
				hypothetical protein	-	0605	0552	MU	family 5	
14	654404	654537	-	hypothetical protein	-	0605	0552	MU	family 5	134
15	913449	913479	-	anti-codon nuclease masking agent (prrB)		0790		V		31
16	924377	924379	CNP 85	hypothetical protein	-	0861	0795	S		3
17	936705	937184	*CNP 87 (partially)*	catalase	katA	0875	0809	P		480
18	958445	958446	-	outer membrane protein-adhesin	babA	1243	0833	-	hop	2
19	1434040	1434055	CNP 157	putative transcriptional regulator	-	1365	1283	TK		16
20	1434393	1434430	CNP 158	no annotation/ intergenic						38
21	1466158	1466254	-	chromosomal replication initiator protein	dnaA	1529	1417	L		97
22	1468776	1468987	CNP 165	exodeoxyribonuclease III	exoA	1526	1415	L		212
23	1469762	1469768	-	hypothetical protein	-	1525	1414	-		7
24	1470071	1470311	-	hypothetical protein	-	1525	1414	-		241
				hypothetical protein	-	1524	1413	-		

No	Position in NQ367 start	end	Corresponding import in NQ4191	Annotation[a] Description	gene	HP-number	jhp-number	COG	OMP	Import length
25	1470536	1471306	-	ATP-dependent DNA helicase RecG	recG	1523	1412	LK		771
26	1471872	1472503	-	ATP-dependent DNA helicase RecG	recG	1523	1412	LK		632
				type III DNA modification enzyme (methyltransferase)	mod_2	1522	1411	L		
27	1473098	1473743	SNP 141	type III DNA modification enzyme (methyltransferase)	mod_2	1522	1411	L		646

a Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S6: Isolated (non-clustered) single nucleotide polymorphisms (SNPs) in strain pair NQ315/1712. Positions refer to the position in the virtual genome of the first isolate NQ315.

		Annotation^{a}						
No	**Position in NQ315**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP family**	**type of mutation**
1	17094	hypothetical protein	-	1587	1493	S		non-synonymous
2	19154	no annotation/ intergenic						intergenic
3	27657	no annotation/ intergenic						intergenic
4	32405	hypothetical protein	-	0013	0011	J		synonymous
5	52228	hypothetical protein	-	0031	0027	-		non-synonymous
6	62695	GDP-D-mannose dehydratase	gmd	0044	0038	M		non-synonymous
7	84409	hypothetical protein	-	0065	0060	-		non-synonymous
8	89517	urease accessory protein UreE	ureE	0070	0065	O		synonymous
9	93094	no annotation/ intergenic						intergenic
10	107925	RNA polymerase sigma factor RpoD	rpoD	0088	0081	K		synonymous
11	110372	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	-	0089	0082	F		synonymous
12	110659	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	-	0089	0082	F		non-synonymous
13	115841	2-hydroxyacid dehydrogenase	-	0096	0088	CHR		non-synonymous
14	120895	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		synonymous
15	124771	methyl-accepting chemotaxis protein (MCP)	-	0103	0095	NT		non-synonymous
16	127177	2',3'-cyclic-nucleotide 2'-phosphodiesterase	cpdB	0104	0096	F		non-synonymous
17	150410	putative outer membrane protein	-	0127	0117	-	hor	non-synonymous
18	152154	hypothetical protein	-	0130	0119	-		non-synonymous
19	165858	DNA glycosylase MutY	mutY	0142	0130	L		non-synonymous
20	193748	putative peptidyl-prolyl cis-trans isomerase	-	0175	0161	O		non-synonymous
21	202251	serine hydroxymethyltransferase	glyA	0183	0171	E		synonymous
22	213107	no annotation/ intergenic						intergenic
23	231502	putative lipopolysaccharide biosynthesis protein	-	0208	0194	M		synonymous
24	237256	succinyl-diaminopimelate desuccinylase	dapE	0212	0198	E		synonymous
25	241716	putative transporter	-	0214	0200	P		synonymous
26	253546	putative outer membrane protein	-	0227	0212	-	hop	synonymous
27	300061	hypothetical protein	-	0271	0256	-		non-synonymous
28	302203	hypothetical protein	-	0275	0260	G		synonymous
29	348307	hypothetical protein	-	0311	0296	-		non-synonymous
30	358143	flagellar basal body L-ring protein	flgH	0325	0308	N	other	synonymous
31	358434	CMP-N-acetylneuraminic acid synthetase (neuA)	-	0326	0309	M		non-synonymous
32	369763	putative single-stranded-DNA-specific exonuclease RecJ	recJ	0348	0322	L		synonymous
33	376843	1-deoxy-D-xylulose-5-phosphate synthase	dxs	0354	0328	HI		synonymous
34	378456	GTP-binding protein LepA	lepA	0355	0329	M		non-synonymous
35	384210	septum formation protein	ftsK	1090	0335	D		synonymous
36	429464	no annotation/ intergenic						intergenic
37	447944	hypothetical protein	-	1023	0401	-		non-synonymous
38	456049	phosphatidylglycerophosphate synthase	pgsA		0407	I		non-synonymous
38	456049	hypothetical protein	-	1015	0408	-		synonymous
39	473840	hypothetical protein	-	0469	0421	-		synonymous
40	499950	putative paralog of HpaA	-	0492	0444	-		synonymous
41	500580	phospho-N-acetylmuramoyl-pentapeptide-transferase	mraY	0493	0445	M		non-synonymous
42	594636	hypothetical protein	-	0586	0534	-		non-synonymous
43	609188	methyl-accepting chemotaxis protein (MCP)	-	0599	0546	NT		non-synonymous
44	611566	flagellin A	flaA	0601	0548	N		non-synonymous
45	630827	putative vacuolating cytotoxin (VacA) paralog	-	0609/0610	0556	N		non-synonymous
46	634570	NAD-dependent DNA ligase LigA	ligA	0615	0558	L		non-synonymous

		Annotation[a]						
No	**Position in NQ315**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP family**	**type of mutation**
47	654033	hydrogenase, small subunit	hyaA	0631	0574	C		synonymous
48	664574	NAD(P)H-flavin oxidoreductase	-	0642	0586	C		non-synonymous
49	676498	no annotation/ intergenic						intergenic
50	678159	protective surface antigen D15	-	0655	0600	M	other	non-synonymous
51	720162	no annotation/ intergenic						intergenic
52	740405	hypothetical protein	-	0711	0650	U		synonymous
53	745690	hypothetical protein	-	0718	0656	R		synonymous
54	761115	hypothetical protein	-	0734	0671	J		non-synonymous
55	777648	flagellar protein FlaG	-	0751	0688	N		non-synonymous
56	778951	flagellar capping protein	fliD	0752	0689	N		synonymous
57	817879	putative heavy-metal cation-transporting P-type ATPase	hmcT	0791	0727	P		synonymous
58	833655	iron (III) dicitrate transport protein	fecA_2	0807	0743	P	family 4	non-synonymous
59	836879	hypothetical protein	-	0812	0748	S		non-synonymous
60	853598	inositol-5-monophosphate dehydrogenase	guaB	0829	0768	F		non-synonymous
61	862180	putative outer membrane protein	-	0839	0777	I	other	non-synonymous
62	869033	type I restriction enzyme restriction subunit	hsdR_2	0846	0784	V		non-synonymous
63	885894	transcription elongation factor GreA	greA	0866	0800	K		non-synonymous
64	914113	hypothetical protein	-	0891	0824	I		non-synonymous
65	929855	hypothetical protein	-	0909	0845	-		synonymous
66	934844	outer membrane protein/porin	hopC	0912	0848	-	hop	non-synonymous
67	935824	outer membrane protein/porin	hopB	0913	0849	-	hop	synonymous
68	938193	putative outer membrane protein	-	0914	0850	-	hof	synonymous
69	938397	putative outer membrane protein	-	0914	0850	-	hof	synonymous
70	944766	carbamoyl phosphate synthase large subunit	carB	0919	0853	EF		synonymous
71	964086	hypothetical protein	-	0935	0869	R		non-synonymous
72	967048	proline/betaine transporter	proP	0936	0871	-		non-synonymous
73	984288	putative aldehyde dehydrogenase	-	0954	0888	C		non-synonymous
74	993255	hypothetical protein	-	0964 /0965	0898	-		synonymous
75	1034050	methionine-tRNA synthetase	metG	0417	0967	J		synonymous
76	1035874	cyclopocyclopropane fatty acid synthase	cfa	0416	0968	M		synonymous
77	1053154	30S ribosomal protein S1	rpsA	0399	0982	J		non-synonymous
78	1054048	30S ribosomal protein S1	rpsA	0399	0982	J		non-synonymous
79	1073021	glutamate dehydrogenase	gdhA	0380	1001	E		synonymous
80	1083586	biotin carboxylase	accC	0370	1011	I		synonymous
81	1092824	tRNA pseudouridylate synthase A	truA	0361	1019	J		non-synonymous
82	1114540	excinuclease ABC subunit B	uvrB	1114	1041	L		non-synonymous
83	1119124	hypothetical protein	-	1117	1045	R		non-synonymous
84	1123858	hypothetical protein	-	1120	1048	-		synonymous
85	1126685	hypothetical protein	-	1124	1053	S		non-synonymous
86	1148572	hypothetical protein	-	1150	1077	R		non-synonymous
87	1156618	putative outer membrane protein	-	1156	1083	-	hop	non-synonymous
88	1157140	putative outer membrane protein	-	1157	1084	-	hop	synonymous
89	1167418	glucose-6-phosphate isomerase	pgi	1166	1093	G		non-synonymous
90	1171386	hypothetical protein	-	1168	1095	T		synonymous
91	1172178	no annotation/ intergenic						intergenic
92	1177102	glucose/galactose transporter	gluP	1174	1101	G		synonymous
93	1177405	glucose/galactose transporter	gluP	1174	1101	G		synonymous
94	1193688	carbonic anhydrase	-	1186	1112	P		non-synonymous
95	1205094	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		non-synonymous
96	1207064	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		synonymous
97	1209897	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		non-synonymous
98	1242831	dihydropteroate synthase	folP	1232	1153	H		synonymous
99	1272916	putative NADH oxidoreductase I	nuoE	1264	1185	-		synonymous
100	1279465	NADH dehydrogenase subunit L	nuoL	1271	1192	CP		non-synonymous
101	1291301	no annotation/ intergenic						intergenic

No	Position in NQ315	Annotation[a] Description	gene	HP-number	jhp-number	COG	OMP family	type of mutation
102	1292464	hypothetical protein	-	1286	1206	S		synonymous
103	1295052	anthranilate synthase component I	trpE	1282	1203	EH		non-synonymous
104	1324652	hypothetical protein	-	1327	1247	MU	family 5	synonymous
105	1353392	nicotinate-nucleotide pyrophosphorylase	nadC	1355	1273	H		non-synonymous
106	1355563	phosphatidylserine decarboxylase	psd	1357	1275	I		synonymous
107	1392803	lipopolysaccharide biosynthesis protein	-	1416	1311	M		non-synonymous
108	1403349	no annotation/ intergenic						intergenic
109	1404484	hypothetical protein	-	1428	1325	R		synonymous
110	1450075	DNA-transfer protein	virB4_2		0917	U		synonymous
111	1458043	chromosomal replication initiation protein	dnaA	1529	1417	L		synonymous
112	1466320	type III DNA modification enzyme (methyltransferase)	mod_2	1522	1411	L		synonymous
113	1509477	hypothetical protein	-	1487	1380	S		synonymous
114	1566549	hypothetical protein	-	1391	1436	-		synonymous

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S7: Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ315/1712. Positions refer to the position in the virtual genome of the first isolate NQ315.

	Position in NQ315		**Annotation[a]**						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
1	1600893	705	putative outer membrane protein penicillin-binding protein	- -	1564 1565	1472 1473	P M		1921
2	1879	2173	hypothetical protein ribosome biogenesis GTP-binding protein	- engB	1566 1567	1474 1475	- R		295
3	8988	9949	ABC transporter ATP-binding protein ABC transporter, permease	metN -	1576 1577	1484 1485	P P		962
4	10787	12260	hypothetical protein putative undecaprenyl-phosphate-alpha-N-acetylglucosaminyltransferase	- wecA	1580 1581	1487 1488	I M		1474
5	12493	12519	pyridoxine 5'-phosphate synthase	pdxJ	1582	1489	H		27
6	12928	12955	pyridoxine 5'-phosphate synthase	pdxJ	1582	1489	H		28
7	13169	13234	4-hydroxythreonine-4-phosphate	pdxA	1583	1490	H		66
8	13760	13958	4-hydroxythreonine-4-phosphate	pdxA	1583	1490	H		199
9	14240	14480	putative DNA-binding/iron metalloprotein/AP endonuclease	ydiE	1584	1491	O		241
10	16796	16824	no annotation/ intergenic						29
11	18228	18244	no annotation/ intergenic						17
12	20255	20835	transcription antitermination protein NusB 6,7-dimethyl-8-ribityllumazine synthase	nusB ribH	0001 0002	0001 0002	K H		581
13	21101	22767	2-dehydro-3-deoxyphosphooctonate aldolase carbonic anhydrase orotidine 5'-phosphate decarboxylase	kdsA icfA pyrF	0003 0004 0005	0003 0004 0005	M P F		1667
14	23053	24311	orotidine 5'-phosphate decarboxylase pantoate-beta-alanine ligase	pyrF panC	0005 0006	0005 0006	F H		1259
15	24734	25432	putative outer membrane protein	-	0009	0007	-	hop	699
16	25867	26008	putative outer membrane protein	-	0009	0007	-	hop	142
17	26218	27078	putative outer membrane protein	-	0009	0007	-	hop	861
18	28112	28217	chaperonin GroEL	groEL	0010	0008	O		106
19	28598	29039	chaperonin GroEL	groEL	0010	0008	O		442
20	44640	46222	putative outer membrane protein type II citrate synthase	- gltA	0025 0026	0021 0022	- C	hop	1583
21	89824	90250	urea transporter	ureI	0071	0066	-		427
22	117199	117316	hypothetical protein	-	0097	0089	-		118
23	119120	119315	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		196
24	119598	119733	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		136
25	130290	132087	cysteine synthase hypothetical protein molecular chaperone DnaK	cysK - dnaK	0107 0108 0109	0099 0100 0101	E - O		1798
26	132333	133272	molecular chaperone DnaK 24kDa chaperone	dnaK grpE	0109 0110	0101 0102	O O		940
27	175649	178228	phosphopyruvate hydratase hypothetical protein hypothetical protein shikimate kinase hypothetical protein	eno - - aroK -	0154 0155 0156 0157 0158	0142 0143 0144 0145 0146	G - - E -		2580
28	178469	178707	hypothetical protein	-	0158	0146	-		239
29	180852	180917	hypothetical protein	-	0160	0148	R		66
30	183756	185494	putative histidine kinase sensor protein putative transkriptional regulator	- -	0164/ 0165 0166	0151 0152	T TK		1739
31	205109	205735	putative cardiolipin synthase	-	0190	0176	I		627
32	206291	206456	putative cardiolipin synthase	-	0190	0176	I		166
33	249233	249240	DNA repair protein RadA	radA	0223	0209	O		8
34	252094	252334	hypothetical protein	-	0226	0211	R		241
35	252758	252970	hypothetical protein	-	0226	0211	R		213

	Position in NQ315		Annotation[a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
36	255055	255057	putative outer membrane protein	-	0227	0212	-	hop	3
37	258701	259103	3-deoxy-manno-octulosonate	kdsB	0230	0215	M		403
38	274829	275197	ATP-dependent RNA helicase DeaD hypothetical protein	deaD -	0247 0248	0232 0233	LKJ O		369
39	278622	278838	ABC transporter, permease	-	0251	0236	R		217
40	279869	280244	putative outer membrane protein	-	0252	0237	-	hop	376
41	301155	301170	hypothetical protein	-	0274	0259	R		16
42	318949	319225	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		277
43	338860	338926	peptide ABC transporter ATP-binding protein	dppF	0302	0287	R		67
44	352433	352616	arginyl-tRNA synthetase	argS	0319	0302	J		184
45	359646	359647	CMP-N-acetylneuraminic acid synthetase	-	0326	0309	M		2
46	362522	362527	ketol-acid reductoisomerase	ilvC	0330	0313	EH		6
47	367780	368160	hypothetical protein	-	0347	0321	J		381
48	489698	489788	hypothetical protein	-	0485	0437	P		91
49	489993	492290	outer membrane protein putative outer membrane protein	- -	0486 0487	0438 0439	- -	hof hof	2298
50	505715	507040	putative transporter putative phospholipase A1	- pldA	0498 0499	0450 0451	R M		1326
51	595819	595825	hypothetical protein	-	0586	0534	-		7
52	612953	613565	flagellin A 3-methyladenine DNA glycosylase	flaA -	0601 0602	0548 0549	N L		613
53	614341	618088	hypothetical protein uroporphyrinogen decarboxylase hypothetical protein putative efflux transporter putative efflux transporter	- hemE - - -	0603 0604 0605 0606 0607	0550 0551 0552 0553 0554	- H MU M V	 family 5 	3748
54	719252	719624	putative outer membrane protein	-	0694	0634	S	other	373
55	737795	738322	hypothetical protein hypothetical protein	- -	0708 0709	0647 0648	- S		528
56	738675	738980	hypothetical protein	-	0709	0648	S		306
57	739784	740022	hypothetical protein	-	0711	0650	U		239
58	742057	742061	RNA polymerase factor sigma-54	rpoN	0714	0652	K		5
59	750402	750408	putative outer membrane protein	-	0725	0662	-	hop	7
60	751250	751764	putative outer membrane protein	-	0725	0662	-	hop	515
61	752744	753336	putative outer membrane protein putative transcriptional regulator	- -	0726 0727	0663 0664	- J	other 	593
62	753943	753949	putative transcriptional regulator	-	0727	0664	J		7
63	776178	776179	hypothetical protein	-	0749	0686	D		2
64	846122	846124	exinuclease ABC subunit C	urvC	0821	0760	L		3
65	894682	895098	catalase	katA	0875	0809	P		417
66	896275	896520	putative iron-regulated outer membrane	frpB_1	0876	0810	P	family 4	246
67	916524	916528	no annotation/ intergenic						5
68	936226	936769	outer membrane protein/porin	hopB	0913	0849	-	hop	544
69	937217	937787	no annotation/ intergenic						571
70	938613	939993	putative outer membrane protein putative iron-regulated outer membrane protein	- frpB_2	0914 0915/ 0916	0850 0851	- P	hof family 4	1381
71	957000	957456	recombination protein RecR tRNA pseudouridine synthase D	recR truD	0925 0926	0859 0860	L S		457
72	958172	958781	tRNA pseudouridine synthase D heat shock protein HtpX	truD htpX	0926 0927	0860 0861	S O		610
73	959808	960980	geranyltranstransferase stationary phase survival protein SurE	ispA surE	0929 0930	0864 0865	H R		1173
74	962939	963021	hypothetical protein	-	0934	0868	O		83
75	975992	976104	hypothetical protein	HPAG1_0928			-		113
76	976792	976867	hypothetical protein	-	0946	0880	C		76
77	977176	977179	hypothetical protein	-	0946	0880	C		4
78	1029285	1029325	putative polysaccharide biosynthesis protein	-	0421	0963	M		41
79	1030691	1030756	hypothetical protein	-	0420	0964	Q		66

	Position in NQ315		Annotation[a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
80	1031851	1032323	hypothetical protein	-	0418	0966	-		473
81	1032537	1033281	hypothetical protein	-	0418	0966	-		745
			methionyl-tRNA synthetase	metG	0417	0967	J		
82	1034403	1034881	methionyl-tRNA synthetase	metG	0417	0967	J		479
			cyclopocyclopropane fatty acid synthase	cfa	0416	0968	M		
83	1035088	1035586	cyclopocyclopropane fatty acid synthase	cfa	0416	0968	M		499
84	1038139	1039976	hypothetical protein	-	0412	0970	-		1838
			hypothetical protein	HPAG1 _0981			-		
			hypothetical protein		0411		-		
			putative paralog of HpaA	-	0410	0971	-		
			GMP synthase	guaA	0409	0972	F		
85	1040797	1041245	GMP synthase	guaA	0409	0972	F		449
86	1090623	1090934	protein-L-isoaspartate O-methyltransferase	pcm	0363	1017	O		312
87	1092005	1092498	hypothetical protein	-	0362	1018	R		494
			tRNA pseudouridine synthase A	truA	0361	1019	J		
88	1143006	1144837	hypothetical protein	-	1143	1071	S		1832
			hypothetical protein	-	1144	1072	-		
89	1153853	1154396	undecaprenyldiphospho-muramoylpentapeptide beta-N-	murG	1155	1082	M		544
90	1157630	1158191	putative outer membrane protein	-	1157	1084	-	hop	562
91	1159817	1159819	putative outer membrane protein	-	1157	1084	-	hop	3
92	1160618	1160822	pyrroline-5-carboxylate reductase	proC	1158	1085	E		205
93	1161504	1161505	putative cAMP-induced cell filamentation	fic	1159	1086	D		2
94	1165996	1166002	hypothetical protein	-	1165	1092	-		7
95	1167678	1168178	glucose-6-phosphate isomerase	pgi	1166	1093	G		501
96	1168438	1169432	putative outer membrane protein	-	1167	1094	-	hof	995
97	1179507	1179751	putative outer membrane function	-	1177	1103	-	hop	245
98	1180115	1180568	putative outer membrane function	-	1177	1103	-	hop	454
99	1187521	1188485	multidrug ABC transproter	-	1181	1107	G		965
			hypothetical protein	-	1182	1108	D		
100	1190201	1190315	putative NA$^+$/H$^+$ antiporter	-	1183	1109	P		115
101	1191019	1192351	hypothetical protein	-	1184	1110	V		1333
			sugar efflux transporter	-	1185	1111	G		
102	1193039	1193296	sugar efflux transporter	-	1185	1111	G		258
103	1194707	1195277	aspartate-semialdehyde dehydrogenase	asd	1189	1114	E		571
104	1253758	1254537	outer membrane protein - adhesin	babB	0896	1164	-	hop	780
105	1254749	1254913	outer membrane protein - adhesin	babB	0896	1164	-	hop	165
106	1255636	1255765	outer membrane protein -fragment	-	-	-	-		130
107	1338669	1338735	outer membrane protein	-	1342	1261	-	hop	67
108	1345403	1345492	hypothetical protein	-	1349	1268	-		90
109	1345742	1346600	hypothetical protein	-	1349	1268	-		859
			carboxyl-terminal protease	prc	1350	1269	M		
110	1346846	1347248	carboxyl-terminal protease	prc	1350	1269	M		403
111	1366508	1366614	type III restriction enzyme R protein	-	1371	1285	-		107
112	1366983	1366984	type III restriction enzyme R protein	-	1371	1285	-		2
113	1368025	1368037	type III restriction enzyme R protein	-	1371	1285	-		13
114	1474062	1474256	putative type II DNA modification enzyme (methyltransferase)	-		1409	V		195
115	1509998	1510114	hypothetical protein	-	1487	1380	S		117
116	1562153	1565089	hypothetical protein	-	1395	1432	-	hor	2937
			hypothetical protein	-	1394	1433	G		
			DNA repair protein (recombination protein N)	recN	1393	1434	L		
117	1574475	1575202	transcription-repair coupling factor	mfd	1541	1458	LK		728
118	1575490	1579017	transcription-repair coupling factor	mfd	1541	1458	LK		3528
			hypothetical protein	-	1542	1457	M		
			hypothetical protein	-	1543	1456	M		
			hypothetical protein	-	1544	1455	M		
119	1579404	1579413	hypothetical protein	-	1544	1455	M		10
120	1599197	1599672	hypothetical protein	-	1561	1469	P		476

	Position in NQ315		Annotation[a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
121	1599883	1599972	putative peroxidase	tsaA	1563	1471	O		90

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S8: Isolated (non-clustered) single nucleotide polymorphisms (SNPs) in strain pair NQ352/1701. Positions refer to the position in the virtual genome of the first isolate NQ352.

		Annotation[a]						
No	**Position in NQ352**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP family**	**type of mutation**
1	5016	putative recombination protein RecB	pcrA	1553	1446	L		non-synonymous
2	5352	putative recombination protein RecB	pcrA	1553	1446	L		non-synonymous
3	7065	Na^+/H^+ antiporter 1	nhaA	1552	1447	P		synonymous
4	9049	preprotein translocase subunit SecD	secD	1550	1449	U		non-synonymous
5	20696	ubiquinol cytochrome c oxidoreductase,	petB	1539	1460	C		non-synonymous
6	21909	ubiquinol cytochrome c oxidoreductase,	petC	1538	1461	C		non-synonymous
7	32469	putative outer membrane protein	-	1564	1472	P		non-synonymous
8	44465	hypothetical protein	-	1580	1487	I		non-synonymous
9	60836	putative outer membrane protein	-	0009	0007	-	hop	non-synonymous
10	65401	hypothetical protein	-	0013	0011	J		synonymous
11	72058	putative chemotaxis protein	cheV_1	0019	0017	NT		synonymous
12	80908	isocitrate dehydrogenase	icd	0027	0023	C		synonymous
13	89812	hypothetical protein	-	0037	0033	U		non-synonymous
14	121607	hypothetical protein	-	0063	0058	-		non-synonymous
15	135868	hypothetical protein	-	0080	0074	-		synonymous
16	173819	hypothetical protein	-	0114	0106	S		synonymous
17	178523	hypothetical protein	-	0117	0109	C		non-synonymous
18	204376	hypothetical protein	-	0143	0131	P		non-synonymous
19	216372	hypothetical protein	-	0158	0146	-		synonymous
20	218120	hypothetical protein	-	0160	0148	R		synonymous
21	218634	no annotation/ intergenic						intergenic
22	227062	peptide chain release factor 2	prfB	0171	0157	J		non-synonymous
23	230334	no annotation/ intergenic						intergenic
24	231356	putative peptidyl-prolyl cis-trans isomerase	-	0175	0161	O		synonymous
25	245164	fumarate reductase iron-sulfur subunit	sdhB	0191	0177	C		synonymous
26	249679	enoyl-(acyl carrier protein) reductase	fabI	0195	0181	I		synonymous
27	252179	S-adenosylmethionine synthetase	metK	0197	0183	H		non-synonymous
28	253168	50S ribosomal protein L32	rpmF	0200	0186	J		non-synonymous
29	255495	hypothetical protein	-	0203	0189	-		synonymous
30	262798	putative outer membrane protein	-	0209	0195	-	hof	non-synonymous
31	276412	cysteine desulfurase	-	0220	0206	E		synonymous
32	277860	hypothetical protein	-	0222	0208	-		non-synonymous
33	283578	putative outer membrane protein	-	0227	0212	-	hop	non-synonymous
34	286928	outer membrane protein/porin	hopA	0229	0214	-	hop	synonymous
35	288235	hypothetical protein	-	0231	0216	O		synonymous
36	288898	hypothetical protein	-	0231	0216	O		synonymous
37	292350	hypothetical protein	-	0235	0220	R		non-synonymous
38	292611	hypothetical protein	-	0235	0220	R		non-synonymous
39	297097	glutamyl-tRNA reductase	hemA	0239	0224	H		non-synonymous
40	297414	glutamyl-tRNA reductase	hemA	0239	0224	H		synonymous
41	298584	hypothetical protein	-	0241	0226	-		synonymous
42	309956	putative outer membrane protein	-	0252	0237	-	hop	synonymous
43	312639	adenylosuccinate synthetase	purA	0255	0239	F		synonymous
44	327806	hypothetical protein	-	0270	0255	S		non-synonymous
45	331622	hypothetical protein	-	0275	0260	G		synonymous
46	337445	queuine tRNA-ribosyltransferase	tgt	0281	0266	J		synonymous
47	337668	hypothetical protein	-	0282	0267	S		synonymous
48	340878	hypothetical protein	-	0284	0269	M		synonymous
49	345957	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		non-synonymous
50	346475	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		synonymous
51	356189	hypothetical protein	-	0292	0277	S		synonymous
52	364973	peptide ABC transporter permease	dppB	0299	0284	EP		non-synonymous
53	367619	peptide ABC transporter ATP-binding protein	dppF	0302	0287	R		synonymous
54	373079	hypothetical protein	-	0308	0293	-		synonymous
55	379229	no annotation/ intergenic						intergenic

		Annotation[a]						
No	**Position in NQ352**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP family**	**type of mutation**
56	383779	poly E-rich protein	-	0322	0305	-		synonymous
57	394021	holliday junction resolvase-like protein	-	0334	0317	L		non-synonymous
58	399232	CTP synthase	pyrG	0349	0323	F		synonymous
59	405518	1-deoxy-D-xylulose-5-phosphate synthase	dxs	0354	0328	HI		synonymous
60	420747	hypothetical protein	-	1085	0340	-		synonymous
61	436547	ATP-dependent zinc metallopeptidase	ftsH_2	1069	0356	O		synonymous
62	447694	hypothetical protein	-	1055	0370	-		non-synonymous
63	456536	hypothetical protein	-	1046	0379	S		non-synonymous
64	459585	hypothetical protein	-	1044	0380	R		non-synonymous
65	481279	response regulator	-	1021	0403	TK		non-synonymous
66	486886	phosphatidylglycerophosphate synthase	pgsA	1016	0407	I		synonymous
67	492261	polyphosphate kinase	ppk	1010	0413	P		synonymous
68	502244	hypothetical protein	-	0466	0418	-		synonymous
69	506783	no annotation/ intergenic						intergenic
70	515990	hypothetical protein	-	0480	0432	T		non-synonymous
71	523448	outer membrane protein	-	0486	0438	-	hof	synonymous
72	527295	hypothetical protein	-	0488	0440	-		non-synonymous
73	527714	hypothetical protein	-	0488	0440	-		synonymous
74	534579	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	murD	0494	0446	M		synonymous
75	536249	putative transporter	-	0497	0449	R		non-synonymous
76	545893	putative outer membrane protein	-	0506	0456	M	other	synonymous
77	551069	no annotation/ intergenic						intergenic
78	559384	hypothetical protein	-	0518	0467	S		non-synonymous
79	559784	hypothetical protein	-	0519	0468	R		synonymous
80	560022	hypothetical protein	-	0519	0468	R		non-synonymous
81	560304	hypothetical protein	-	0519	0468	R		non-synonymous
82	570696	cag island protein	orf13/14	0527	0476	U		synonymous
83	581265	cag island protein	cagP	0536	0484	-		non-synonymous
84	596066	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		non-synonymous
85	596285	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		non-synonymous
86	606389	3-oxoacyl-(acyl carrier protein) synthase II	fabB	0558	0505	IQ		synonymous
87	615170	GTP-dependent nucleic acid-binding protein EngD	-	0569	0516	J		non-synonymous
88	620073	signal peptidase I	lepB	0576	0523	U		non-synonymous
89	625859	dihydroorotase	pyrC_2	0581	0528	F		synonymous
90	629084	hypothetical protein	-	0585	0533	P		non-synonymous
91	632439	hypothetical protein	-	0587	0535	R		synonymous
92	634472	2-oxoglutarate acceptor oxidoreductase subunit OorA	oorA	0589	0537	C		synonymous
93	637862	hypothetical protein	-		0540	R		synonymous
94	651717	hypothetical protein	-	0605	0552	MU	family 5	non-synonymous
95	664920	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		non-synonymous
96	672162	aspartyl-tRNA synthetase	aspS	0617	0560	J		synonymous
97	680162	UDP-N-acetylmuramate-L-alanine ligase	murC	0623	0567	M		synonymous
98	704330	hypothetical protein	-	0647	0592	-		synonymous
99	708192	alpha(1,3)-fucosyltransferase	fucT	0651	0596	-		non-synonymous
100	711383	protective surface antigen D15	-	0655	0600	M	other	non-synonymous
101	712668	protective surface antigen D15	-	0655	0600	M	other	synonymous
102	715331	putative processing protease	-	0657	0602	R		non-synonymous
103	717953	hypothetical protein	-	0659	0604	-		synonymous
104	722529	hypothetical protein	-	0664	0609	-		synonymous
105	743850	iron (III) dicitrate transport protein	fecA_1	0686	0626	P	family 4	synonymous
106	747924	no annotation/ intergenic						intergenic

		Annotation[a]						
No	Position in NQ352	Description	gene	HP-number	jhp-number	COG	OMP family	type of mutation
107	748429	hypothetical protein	-	0688/ 0689	0628	L		non-synonymous
108	749207	acetyl-coA acetlytransferase	thl	0690	0638	I		synonymous
109	754350	putative hydrantoin utilization	-	0695	0633	EQ		non-synonymous
110	766304	excinuclease ABC subunit A	uvrA	0705	0644	L		synonymous
111	766599	excinuclease ABC subunit A	uvrA	0705	0644	L		synonymous
112	775987	RNA polymerase factor sigma-54	rpoN	0714	0652	K		synonymous
113	794425	hypothetical protein	-	0733	0670	R		non-synonymous
114	804387	hypothetical protein	-	0744	0681	R		non-synonymous
115	806140	hypothetical protein	-	0745	0682	J		non-synonymous
116	807444	hypothetical protein	-	0746	0683	R		synonymous
117	818871	hypothetical protein	-	0759	0696	V		non-synonymous
118	825968	molybdenum cofactor biosynthesis protein A	moaA	0768	0705	H		non-synonymous
119	826836	molybdenum cofactor biosynthesis protein A	moaA	0768	0705	H		non-synonymous
120	831271	hypothetical protein	-	0773	0710	R		synonymous
121	835342	guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase	spoT	0775	0712	TK		synonymous
122	837834	bifunctional aconitate hydratase 2/2-mehtylisocitrate dehydrogenase	acnB	0779	0716	C		synonymous
123	844793	outer-membrane lipoprotein carrier protein	lolA	0785	0722	M		synonymous
124	846604	preprotein translocase subunit SecA	secA	0786	0723	U		synonymous
125	862686	GTP cyclohydrolase II	ribA	0802	0738	H		synonymous
126	879707	excinuclease ABC subunit C	uvrC	0821	0760	L		synonymous
127	880049	excinuclease ABC subunit C	uvrC	0821	0760	L		non-synonymous
128	906675	putative type I restriction enzyme (specifity subunit)	hsdS_4		0726	V		synonymous
129	909006	no annotation/ intergenic						intergenic
130	915502	putative ADP-D-glycero-D-mannoheptose synthase	waaE	0858	0792	M		synonymous
131	922177	lipid-A-disaccharide synthase	lpxB	0867	0801	M		non-synonymous
132	926558	CDP-diacylglycerol pyrophosphatase	cdh	0871	0805	I		synonymous
133	951020	hypothetical protein	HPAG1_0319			-		synonymous
134	960481	phosphotransacetylase	pta	0904/ 0905	0841	C		synonymous
135	967638	putative ATP-dependent helicase	-	0911	0847	L		synonymous
136	991686	no annotation/ intergenic						intergenic
137	997492	hypothetical protein	-	0931/ 0932	0866	-		synonymous
138	998449	hypothetical protein	-	0934	0868	O		synonymous
139	1000301	putative outer membrane protein	-		0870	-	hom	synonymous
140	1012913	hypothetical protein	-	0946	0880	C		synonymous
141	1019310	putative aldehyde dehydrogenase	-	0954	0888	C		non-synonymous
142	1020026	prolipoprotein diacylglyceryl transferase	lgt	0955	0889	M		non-synonymous
143	1025647	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase	gpsA	0961	0895	C		synonymous
144	1027474	hypothetical protein	-	0964	0898	-		synonymous
145	1048258	no annotation/ intergenic						intergenic
146	1053300	DNA transfer protein	virB4_3		0918	U		synonymous
147	1061772	integrase/ recombinase (XERCD family)	-		0951	L		non-synonymous
148	1063997	hypothetical protein	-		0955	S		non-synonymous
149	1081534	putative S/N-oxide reductase	-	0407	0974	C		synonymous
150	1085282	nifS-like protein	-	0405	0976	E		synonymous

No	Position in NQ352	Annotation[a]						
		Description	gene	HP-number	jhp-number	COG	OMP family	type of mutation
151	1086325	phenylalanyl-tRNA synthetase subunit alpha	pheS	0403	0978	J		synonymous
152	1093112	30S ribosomal protein S1	rpsA	0399	0982	J		non-synonymous
153	1097541	hypothetical protein	-	0395	0986	R		synonymous
154	1097937	hypothetical protein	-	0395	0986	R		synonymous
155	1108640	hypothetical protein	-	0384	0997	-		synonymous
156	1110453	putative zinc-metallo protease	-	0382	0999	O		synonymous
157	1115513	putative cytochrome C-type biogenesis protein	-	0378	1003	O		non-synonymous
158	1141784	putative dehydrogenase	-	1102	1028	G		non-synonymous
159	1147375	putative outer membrane protein	-	1107	1034	-	hor	non-synonymous
160	1148308	pyruvate flavodoxin oxidoreductase subunit gamma	porG	1108	1035	C		non-synonymous
161	1151649	adenylosuccinate lyase	purB	1112	1039	F		synonymous
162	1163417	flagellar hook-associated protein FlgK	flgK	1119	1047	N		non-synonymous
163	1189103	hypothetical protein	-	1145/ 1146	1073	S		synonymous
164	1202342	putative outer membrane protein	-	1157	1084	-	hop	synonymous
165	1206053	flavodoxin FldA	fldA	1161	1088	C		synonymous
166	1210705	glucose-6-phosphate isomerase	pgi	1166	1093	G		synonymous
167	1212785	putative outer membrane protein	-	1167	1094	-	hof	non-synonymous
168	1218390	no annotation/ intergenic						intergenic
169	1223526	putative outer membrane function	-	1177	1103	-	hop	non-synonymous
170	1233696	sugar efflux transporter	-	1185	1111	G		synonymous
171	1245944	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		non-synonymous
172	1247758	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		synonymous
173	1250546	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		non-synonymous
174	1260347	ABC transporter ATP-binding protein	-	1206	1129	V		synonymous
175	1264038	ulcer-associated gene restriction endonuclease (iceA)	-	1209		-		non-synonymous
176	1264353	ulcer-associated gene restriction endonuclease (iceA)	-	1209		-		non-synonymous
177	1267662	polynucleotide phosphorylase/ polyadenylase	pnp	1213	1136	J		non-synonymous
178	1274727	D-lactate dehydrogenase	dld	1222	1143	C		non-synonymous
179	1282377	aspartate kinase	lysC	1229	1150	E		synonymous
180	1288843	carbamoyl phosphate synthase small subunit	carA	1237	1158	EF		non-synonymous
181	1292287	Maf-like protein	maf	1240	1161	D		synonymous
182	1292495	alanyl-tRNA synthetase	alaS	1241	1162	J		synonymous
183	1297979	30S ribosomal protein S18	rpsR	1244	1165	J		synonymous
184	1298560	single-strand DNA-binding protein	ssb	1245	1166	L		non-synonymous
185	1299550	DNA polymerase III subunit delta	-	1247	1168	L		synonymous
186	1337959	hypothetical protein	-	1285	1205	R		non-synonymous
187	1338296	hypothetical protein	-	1285	1205	R		non-synonymous
188	1343138	50S ribosomal protein L17	rplQ	1292	1212	J		synonymous
189	1347208	preprotein translocase subunit SecY	secY	1300	1220	U		non-synonymous
190	1356280	50S ribosomal protein L3	rplC	1319	1239	J		synonymous
191	1357816	hypothetical protein	-	1321	1241	R		non-synonymous
192	1360179	"putative" contains conserved domains sirD2				-		synonymous
193	1375835	no annotation/ intergenic						intergenic
194	1383853	hypothetical protein	-	1349	1268	-		synonymous
195	1389337	hypothetical protein	-	1354	1272	R		non-synonymous
196	1392502	nicotinate-nucleotide pyrophosphorylase	nadC	1355	1273	H		non-synonymous

		Annotation[a]						
No	**Position in NQ352**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP family**	**type of mutation**
197	1408261	hypothetical protein	-	0996	0942	U		non-synonymous
198	1409913	integrase/ recombinase (XERCD family)	-	0995	0941	L		synonymous
199	1421074	no annotation/ intergenic						intergenic
200	1421319	no annotation/ intergenic						intergenic
201	1427842	no annotation/ intergenic						intergenic
202	1434481	hypothetical protein	-		0928	KL		non-synonymous
203	1447227	VirB8 type IV secretion protein- HPP12-1328	virB8-2			U		non-synonymous
204	1458769	UDP-N-acetylglucosamine acyltransferase	lpxA	1375	1289	M		synonymous
205	1460833	hypothetical protein	-	1378	1292	R		synonymous
206	1463183	ATP-dependet protease LA	lon	1379	1293	O		non-synonymous
207	1472980	no annotation/ intergenic						intergenic
208	1474649	no annotation/ intergenic						intergenic
209	1476259	no annotation/ intergenic						intergenic
210	1482919	hypothetical protein	-	1414	1309	S		non-synonymous
211	1483500	tRNA delta(2)-isopentenylpyrophosphate transferase	miaA	1415	1310	J		non-synonymous
212	1492610	isoleucyl-tRNA synthetase	ileS	1422	1317	J		non-synonymous
213	1517551	putative outer membrane protein	-	1453	1346	-	hom	synonymous
214	1534066	ABC transporter ATP-binding protein	-	1465	1358	Q		synonymous
215	1540661	DNA polymerase I	polA	1470	1363	L		synonymous
216	1540909	DNA polymerase I	polA	1470	1363	L		non-synonymous
217	1553504	hypothetical protein	-	1481	1374	R		synonymous
218	1557027	hypothetical protein	-	1486	1379	-		synonymous
219	1567341	transaldolase	tal	1495	1388	G		non-synonymous
220	1581888	no annotation/ intergenic						intergenic
221	1583145	putative iron-regulated outer membrane protein	frpB_3	1512	1405	P	family 4	synonymous
222	1583643	putative iron regulated outer membrane protein	frpB_3	1512	1405	P	family 4	synonymous
223	1599827	hypothetical protein	-	1524	1413	-		synonymous
224	1601025	exodeoxyribonuclease III	exoA	1526	1415	L		non-synonymous
225	1601891	hypothetical protein	-	1527	1416	-		non-synonymous
226	1609045	putative type I restriction enzyme (specifity subunit)	hsdS_3a		1422	V		non-synonymous
227	1620421	amino acid permease - HPAG1-1471	HPAG1_1471			E		synonymous
228	1624985	hypothetical protein	-	1396	1431	S		non-synonymous
229	1626608	hypothetical protein	-	1394	1433	G		non-synonymous
230	1626948	hypothetical protein	-	1394	1433	G		synonymous
231	1627200	hypothetical protein	-	1394	1433	G		synonymous
232	1633337	DNA polymerase III subunit epsilon	-	1387	1438	L		non-synonymous

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S9: Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ352/1701. Positions refer to the position in the virtual genome of the first isolate NQ352.

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
1	7911	8580	preprotein translocase subunit SecD	secD	1550	1449	U		670
2	12791	12792	leucyl-tRNA synthetase	leuS	1547	1452	J		2
3	25061	25370	penicillin-binding protein	-	1556	1464	M		310
4	26435	28206	penicillin-binding protein	-	1556	1464	M		1772
			flagellar hook-basal body protein FliE	fliE	1557	1465	NU		
			flagellar basal body rod protein FlgC	flgC	1558	1466	N		
			flagellar basal body rod protein FlgB	flgB	1559	1467	N		
5	33208	33595	penicillin-binding protein	-	1565	1473	M		388
6	34667	34673	penicillin-binding protein	-	1565	1473	M		7
7	34889	35067	penicillin-binding protein	-	1565	1473	M		179
			hypothetical protein	-	1566	1474	-		
8	36416	37376	hypothetical protein	-	1568	1476	S		961
			hypothetical protein	-	1569	1477	-		
			hypothetical protein	-	1570	1478	R		
9	38431	38527	putative outer membrane protein	-	1571	1479	M		97
10	38993	39438	regulatory protein DniR	dniR	1572	1480	M		446
11	39993	40019	hypothetical protein	-	1573	1481	L		27
12	40559	40639	hypothetical protein	-	1573	1481	L		81
13	41785	43557	ABC transporter ATP-binding protein	metN	1576	1484	P		1773
			ABC transporter, permease	-	1577	1485	P		
			hypothetical protein	-	1579	1486	-		
14	47841	48258	putative DNA-binding/iron metalloprotein/AP endonuclease	ydiE	1584	1491	O		418
15	49422	51393	flagellar basal body rod protein FlgG	flgG	1585	1492	N		1972
			hypothetical protein	-	1587	1493	S		
16	60136	60460	putative outer membrane protein	-	0009	0007	-	hop	325
17	61895	62950	chaperonin GroEL	groEL	0010	0008	O		1056
			co-chaperonin GroES	groES	0011	0009	O		
18	64470	64874	DNA primase	dnaG	0012	0010	L		405
19	74583	74987	lipid A 1-phosphatase	-	0021	0019	I		405
			lipid A phosphoethanolamine transferase	-	0022	0020	R		
20	75482	76109	lipid A phosphoethanolamine transferase	-	0022	0020	R		628
21	76322	77912	putative outer membrane protein	-	0025	0021	-	hop	1591
22	78540	79048	putative outer membrane protein	-	0025	0021	-	hop	509
23	81187	82459	isocitrate dehydrogenase	icd	0027	0023	C		1273
			hypothetical protein	-	0028	0024	-		
24	83068	85153	dethiobiotin synthetase	bioD	0029	0025	H		2086
			hypothetical protein	-	0030	0026	-		
			hypothetical protein	-	0031	0027	-		
25	85536	86341	hypothetical protein	-	0032	0028	S		806
			ATP-dependent protease, ATP-binding subunit	clpA	0033	0029	O		
26	88247	89332	aspartate alpha-decarboxylase	panD	0034	0030	H		1086

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
			hypothetical protein	-	0035	0031	S		
			hypothetical protein	-	0036	0032	-		
27	95109	95692	phosphomannose isomerase/GDP-mannose pyrophosphorylase	manC	0043	0037	M		584
			GDP-D-mannose dehydratase	gmd	0044	0038	M		
28	95941	96265	GDP-D-mannose dehydratase	gmd	0044	0038	M		325
29	96725	97088	putative sugar nucleotide biosynthesis	-	0045	0039	MG		364
30	99088	99903	hydrogenase expression/ formation protein	hypE	0047	0040	O		816
			putative transcriptional regulator	-	0048	0041	O		
31	109356	111378	sodium/proline symporter	putP	0055	0047	ER		2023
			proline/pyrroline-5-carboxylate dehydrogenase	putA	0056	0048	C		
32	114441	114761	proline/pyrroline-5-carboxylate dehydrogenase	putA	0056	0048	C		321
33	114982	115248	hypothetical protein	-	0057	0049	-		267
34	121242	121334	hypothetical protein	-	0063	0058	-		93
35	122544	122608	hypothetical protein	-	0066	0061	D		65
36	125854	125923	urease accessory protein	ureG	0068	0063	OK		70
37	126202	127138	urease accessory protein	ureF	0069	0064	O		937
			urease accessory protein UreE	ureE	0070	0065	O		
38	127568	127589	urea transporter	ureI	0071	0066	-		22
39	132136	132139	phosphoglucosamine mutase	glmM	0075	0070	G		4
40	132469	132781	phosphoglucosamine mutase	glmM	0075	0070	G		313
41	133142	133399	30S ribosomal protein S20	rpsT	0076	0071	J		258
			peptide chain release factor 1	prfA	0077	0072	J		
42	133739	133907	peptide chain release factor 1	prfA	0077	0072	J		169
43	134120	134823	peptide chain release factor 1	prfA	0077	0072	J		704
			putative outer membrane protein	-	0078/ 0079	0073	-	hor	
44	135078	135507	putative outer membrane protein	-	0078/ 0079	0073	-	hor	430
45	136249	136978	hypothetical protein	-	0080	0074	-		730
46	137278	138488	hypothetical protein	-	0080	0074	-		1211
			methyl-accepting chemotaxis protein (MCP)	-	0082	0075	NT		
47	138749	140217	methyl-accepting chemotaxis protein (MCP)	-	0082	0075	NT		1469
48	140430	140608	30S ribosomal protein S9	rpsI	0083	0076	J		179
49	140821	142491	50S ribosomal protein L13	rplM	0084	0077	J		1671
			hypothetical protein	-	0085	0078	-		
			hypothetical protein	-	0086	0079	R		
50	143568	143850	hypothetical protein	-	0087	0080	M		283
51	151846	153050	hypothetical protein	-	0095	0087	-		1205
			2-hydroxyacid dehydrogenase	-	0096	0088	CHR		
52	155668	156054	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		387
53	156674	156684	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		11
54	157420	158271	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		852
			hypothetical protein	-	0100	0092	S		

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
55	167196	167224	hypothetical protein	-	0108	0100	-		29
56	170011	170012	24kDa chaperone	grpE	0110	0102	O		2
57	170389	170394	24kDa chaperone	grpE	0110	0102	O		6
58	170680	170692	heat-inducible transcription repressor	-	0111	0103	K		13
59	173294	173342	hypothetical protein	-	0114	0106	S		49
60	177148	178084	DNA topoisomerase I	topA_1	0116	0108	L		937
61	180467	183244	hypothetical protein	-	0118	0110	-		2778
			phosphoenolpyruvate synthase	ppsA	0121	0111	G		
62	183468	184108	hypothetical protein	-	0122	0112	-		641
			threonyl-tRNA synthetase	thrS	0123	0113	J		
63	185616	185835	translation initiation factor IF-3	infC	0124	0114	J		220
64	193037	197473	phospho-2-dehydro-3-deoxyheptonate aldolase	aroF	0134	0122	E		4437
			hypothetical protein	-	0135	0123	-		
			bactoferritin comigratory protein	bcp	0136	0124	O		
			hypothetical protein	-	0137	0125	S		
			putative iron-sulfur protein	-	0138	0126	C		
65	198794	200670	L-lactate permease	lldP_1	0140	0128	C		1877
			L-lactate permease	lldP_2	0141	0129	C		
66	201100	202285	L-lactate permease	lldP_2	0141	0129	C		1186
			DNA glycosylase MutY	mutY	0142	0130	L		
67	205554	208616	cytochrome oxidase (CBB3- type)	fixN	0144	0132	O		3063
			cytochrome oxidase (CBB3- type)	fixO	0145	0133	C		
			putative cytochrome oxidase (CBB3- type)	fixQ	0146	0134	O		
			cytochrome oxidase (CBB3-type)	fixP	0147	0135	C		
			hypothetical protein	-	0148	0136	-		
			hypothetical protein	-	0149	0137	-		
68	208833	209734	hypothetical protein	-	0150	0138	-		902
69	211562	212132	hypothetical protein	-	0152	0140	R		571
			recombinase A	recA	0153	0141	L		
70	213948	215838	phosphopyruvate hydratase	eno	0154	0142	G		1891
			hypothetical protein	-	0155	0143	-		
			hypothetical protein	-	0156	0144	-		
			shikimate kinase	aroK	0157	0145	E		
			hypothetical protein	-	0158	0146	-		
71	217026	217067	putative lipopolysaccharide biosynthesis protein	-	0159	0147	M		42
72	219566	219674	hypothetical protein	-	0162	0149	S		109
73	221645	222332	putative histidine kinase sensor protein	-	0164/ 0165	0151	T		688
74	243454	244816	putative cardiolipin synthase	-	0190	0176	I		1363
			fumarate reductase iron-sulfur subunit	frdB	0191	0177	C		
75	246701	247379	fumarate reductase flavoprotein subunit	frdA	0192	0178	C		679
76	256387	256672	hypothetical protein	-	0204	0190	-		286
77	257273	257559	hypothetical protein	-	0205	0191	E		287
78	263114	263601	putative outer membrane protein	-	0209	0195	-	hof	488
79	269183	269718	tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA	gidA	0213	0199	D		536

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
80	270802	270937	putative transporter	-	0214	0200	P		136
81	271573	272383	putative transporter	-	0214	0200	P		811
			CDP-diacylglycerol synthase	cdsA	0215	0201	I		
82	274110	274592	hypothetical protein	-	0218	0204	R		483
83	285396	286188	hypothetical protein	-	0228	0213	P		793
			outer membrane protein/porin	hopA	0229	0214	-	hop	
84	286464	286473	outer membrane protein/porin	hopA	0229	0214	-	hop	10
85	287741	287901	3-deoxy-manno-octulosonate cytidylyltransferase	kdsB	0230	0215	M		161
86	294608	295008	prolyl-tRNA synthetase	proS	0238	0223	J		401
87	295350	295365	prolyl-tRNA synthetase	proS	0238	0223	J		16
88	298003	298006	putative octaprenyl-diphosphate synthase	ispB	0240	0225	H		4
89	299272	299667	neutrophil-activating protein A	napA	0243	0228	P		396
90	300246	300599	putative histidine kinase sensor protein	-	0244	0229	T		354
91	306479	306749	ABC transporter, ATP-binding protein	-	0250	0235	R		271
92	308405	308498	no annotation / intergenic						94
93	308735	308911	putative outer membrane protein	-	0252	0237	-	hop	177
94	309253	309336	putative outer membrane protein	-	0252	0237	-	hop	84
95	309652	309718	putative outer membrane protein	-	0252	0237	-	hop	67
96	311044	311370	putative outer membrane protein	-	0253/ 0254	0238	-	hop	327
97	311581	311884	putative outer membrane protein	-	0253/ 0254	0238	-	hop	304
			adenylosuccinate synthetase	purA	0255	0239	F		
98	317478	317590	type II DNA modification enzyme (methyltransferase)	-	0260	0244	L		113
			hypothetical protein	-	0261	0245	-		
99	320312	321199	heat shock protein	clpB	0264	0249	O		888
100	324213	324869	dihydroorotase	pyrC_1	0266	0251	F		657
			chlorohydrolase	orf1	0267	0252	FR		
101	330540	330698	hypothetical protein	-	0275	0260	G		159
102	331972	332044	hypothetical protein	-	0276	0261	-		73
103	333504	335533	guanosine -5'-triphosphate,3'-diphosphate pyrophosphatase	gppA	0278	0263	FP		2030
			lipopolysaccharide heptosyltransferase-1	waaC	0279	0264	M		
			lipid A biosynthesis lauroyl acyltransferase	waaM	0280	0265	M		
104	341241	345633	hypothetical protein	-	0284	0269	M		4393
			hypothetical protein	-	0285	0270	J		
			ATP-dependent zinc metallopeptidase	ftsH_1	0286	0271	O		
			hypothetical protein	-	0287	0272	-		
			hypothetical protein	-	0288	0273	-		
			putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		
105	351116	351119	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		4
106	351455	351594	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		140
107	352305	353450	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		1146
108	354059	355309	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		1251

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
			diaminopimelate decarboxylase	lysA	0290	0275	E		
109	355560	355562	hypothetical protein	-	0291	0276	E		3
110	360771	360804	flagellar hook-associated protein FlgL	flgL	0295	0280	N		34
111	361251	362066	flagellar hook-associated protein FlgL	flgL	0295	0280	N		816
112	365805	367359	dipeptide ABC transporter permease	dppC	0300	0285	EP		1555
			peptide ABC transporter ATP-binding protein	dppD	0301	0286	EP		
113	370000	370532	hypothetical protein	-	0304	0289	-		533
114	375265	375472	hypothetical protein	-	0311	0296	-		208
115	379743	379890	hypothetical protein	-	0318	0301	P		148
116	380474	380565	hypothetical protein	-	0318	0301	P		92
			arginyl-tRNA synthetase	argS	0319	0302	J		
117	381519	382423	arginyl-tRNA synthetase	argS	0319	0302	J		905
			hypothetical protein	-	0320	0303	U		
118	384040	384475	poly E-rich protein	-	0322	0305	-		436
119	386824	386945	CMP-N-acetylneuraminic acid synthetase (neuA)	-	0326	0309	M		122
120	387588	392316	CMP-N-acetylneuraminic acid synthetase (neuA)	-	0326	0309	M		4729
			putative flagellar biosynthesis protein	flaG	0327	0310	J		
			tetraacyldisaccharide 4'-kinase	lpxK	0328	0311	M		
			NAD synthetase	nadE	0329	0312	H		
			ketol-acid reductoisomerase	ilvC	0330	0313	EH		
			cell division inhibitor	minD	0331	0314	D		
121	392874	392877	hypothetical protein	-	0333	0316	LU		4
122	394850	397160	hypothetical protein	-	0337	0319	-		2311
			hypothetical protein	-	0338	0320	-		
			hypothetical protein	-	0347	0321	J		
			putative single-stranded -DNA-specific exonuclease RecJ	recJ	0348	0322	L		
123	402842	403956	flagellar motor switch protein G	fliG	0352	0326	N		1115
			flagellar assembly protein H	fliH	0353	0327	NU		
124	404189	405272	1-deoxy-D-xylulose-5-phosphate synthase	dxs	0354	0328	HI		1084
125	416114	416197	hypothetical protein	-	1089	0336	L		84
126	416957	417258	hypothetical protein	-	1089	0336	L		302
			transketolase	tktA	1088	0337	G		
127	418878	419032	transketolase	tktA	1088	0337	G		155
128	421604	422834	aspartate carbamoyltransferase catalytic subunit	pyrB	1084	0341	F		1231
			putative outer membrane protein	-	1083	0342	-	hof	
129	423375	423380	putative outer membrane protein	-	1083	0342	-	hof	6
130	423898	424298	multi-drug resistance protein	msbA	1082	0343	V		401
131	431088	432499	hypothetical protein	-	1075	0350	R		1412
			hypothetical protein	-	1074	0351	-		
132	433742	434079	copper-transporting P-type ATPase	copA	1072	0353	P		338
133	434286	434832	copper-transporting P-type ATPase	copA	1072	0353	P		547
134	435608	435800	phosphatidylserine synthase	pssA	1071	0354	I		193
135	443583	443758	sec-independent translocase	-	1060	0365	U		176
136	448426	448720	hypothetical protein	-	1055	0370	-		295

	Position in NQ352		**Annotation [a]**						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
			hypothetical protein	-	1054	0371	M		
137	448947	449527	hypothetical protein	-	1054	0371	M		581
138	449729	449885	hypothetical protein	-	1054	0371	M		157
			septum formation inhibitor	minC	1053	0372	D		
139	450221	450680	septum formation inhibitor	minC	1053	0372	D		460
			UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase	lpxC	1052	0373	M		
140	451489	451737	hypothetical protein	-	1051	0374	O		249
141	454593	455784	translation initiation factor IF-2	infB	1048	0377	J		1192
142	457010	457227	acetyl-coA synthetase	-	1045		I		218
143	457704	457995	acetyl-coA synthetase	-	1045		I		292
144	459242	459299	hypothetical protein	-	1044	0380	R		58
145	460116	460701	putative transcriptional regulator	-	1043	0381	TK		586
146	463705	468397	flagellar biosynthesis protein FlhA	flhA	1041	0383	NU		4693
			30S ribosomal protein S15	rpsO	1040	0384	J		
			hypothetical protein	-	1039	0385	M		
			3-dehydroquinate dehydratase	aroD	1038	0386	E		
			putative proline peptidase	pepQ	1037	0387	E		
147	472642	472747	flagellar motor switch protein FliM	fliM	1031	0393	N		106
148	473083	474187	flagellar motor switch protein FliM	fliM	1031	0393	N		1105
			flagellar motor switch protein FliY	-	1030	0394	NT		
149	483537	483975	protease DO	htrA	1019	0405	O		439
150	485033	485896	amino acid permease	-	1017	0406	E		864
151	489111	492003	dihydrodipicolinate synthase	dapA	1013	0410	EM		2893
			putative zinc protease	-	1012	0411	R		
			dihydroorotate dehydrogenase 2	pyrD	1011	0412	F		
			polyphosphate kinase	ppk	1010	0413	P		
152	492803	493617	polyphosphate kinase	ppk	1010	0413	P		815
153	496592	499771	type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		3180
			type I restriction enzyme R protein (hsdR)	-	0464	0416	V		
154	500844	500857	hypothetical protein	-	0465	0417	S		14
155	502970	504228	hypothetical protein	-	0467	0419	-		1259
			hypothetical protein	-	0468	0420	C		
156	504629	504863	hypothetical protein	-	0469	0421	-		235
157	507323	507446	glutathione-regulated potassium-efflux system protein	kefB	0471	0423	P		124
158	507959	508835	glutathione-regulated potassium-efflux system protein	kefB	0471	0423	P		877
			putative outer membrane protein	-	0472	0424	-	hor	
159	511542	511746	glutamyl-tRNA synthetase	gltX	0476	0428	J		205
160	513657	514894	type II DNA modification enzyme (methyltransferase)	-	0478	0430	L		1238
			hypothetical protein	-	0479	0431	M		
161	521315	521323	hypothetical protein	-	0485	0437	P		9
162	548648	548654	putative glycolate oxidase	glcD	0509	0459	C		7
163	548873	550323	putative glycolate oxidase	glcD	0509	0459	C		1451
			dihydrodipicolinate reductase	dapB	0510	0460	E		
164	551883	551889	glutamine synthetase	glnA	0512	0461	E		7

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
165	555576	556850	50S ribosomal protein L9	rplI	0514	0463	J		1275
			ATP-dependent protease peptidase subunit	hslV	0515	0464	O		
			ATP-dependent protease ATP-binding subunit HslU	hslU	0516	0465	O		
166	567575	567930	cag island protein, DNA transfer protein	virB11_1	0525	0474	NU		356
			cag island protein	orf12	0526	0475	-		
167	590359	590383	DNA transfer protein	cagE	0544	0492	U		25
168	591958	592953	cag island protein	cagC	0546	0494	-		996
			cag pathogenicity island protein B	HPAG1_0523			-		
169	605363	605714	3-oxoacyl-(acyl carrier protein) synthase II	fabB	0558	0505	IQ		352
170	608934	610055	hypothetical protein	-	0563	0510	-		1122
			hypothetical protein	-	0564	0511	-		
			hypothetical protein	-	0565	0512	S		
171	620695	622759	methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	folD	0577	0524	H		2065
			hypothetical protein	-	0578	0525	M		
172	623267	623523	hypothetical protein	-	0578	0525	M		257
			hypothetical protein	-	0579	0526	-		
173	623766	623803	hypothetical protein	-	0579	0526	-		38
174	637187	637396	hypothetical protein	-		0540	R		210
175	650432	651140	uroporphyrinogen decarboxylase	hemE	0604	0551	H		709
176	659156	659186	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		31
177	660011	660034	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		24
178	662753	663776	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		1024
179	667090	667275	no annotation / intergenic						186
180	667623	671079	hypothetical protein	-	0614	0557	-		3457
			NAD-dependent DNA ligase LigA	ligA	0615	0558	L		
			putative chemotaxis protein	cheV_2	0616	0559	NT		
			aspartyl-tRNA synthetase	aspS	0617	0560	J		
181	675270	675471	putative lipopolysaccharide biosynthesis protein	-	0619	0563	M		202
182	679696	679863	UDP-N-acetylmuramate-L-alanine ligase	murC	0623	0567	M		168
183	681181	681488	hypothetical protein	-	0624	0568	E		308
184	683542	684608	putative 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase	dapD	0626	0570	E		1067
			hypothetical protein	-	0627/ 0628	0571	R		
185	685473	689879	hypothetical protein	-	0629	0572	S		4407
			hypothetical protein	-	0630	0573	R		
			hydrogenase, small subunit	hyaA	0631	0574	C		
			hydrogenase, large subunit	hyaB	0632	0575	C		
186	690347	696251	hydrogenase, large subunit	hyaB	0632	0575	C		5905
			hydrogenase, cytochrome subunit	hyaC	0633	0576	C		
			hydrogenase expression/formation protein	hyaD	0634	0577	C		
			hypothetical protein	-	0635	0578	-		
			hypothetical protein	-	0636	0579	-		

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
			hypothetical protein	-	0637	0580	-		
			putative outer membrane protein	-	0638	0581	-	hop	
			hypothetical protein	-	0639	0582	R		
187	697677	702328	polynucleotide adenylyltransferase	pcnB	0640	0583	J		4652
			hypothetical protein	-	0641	0584	-		
			putative 3-hydroxyacid dehydrogenase	-		0585	I		
			NAD(P)H-flavin oxidoreductase	-	0642	0586	C		
			glutamylglutaminyl-tRNA synthetase	gltX	0643	0588	J		
			hypothetical protein	-	0644	0589	-		
			lytic murein transglycosylase	-	0645	0590	M		
188	703021	703494	lytic murein transglycosylase	-	0645	0590	M		474
			UTP-glucose-1-phosphate uridylyltransferase	galU	0646	0591	M		
189	703710	703948	UTP-glucose-1-phosphate uridylyltransferase	galU	0646	0591	M		239
190	704791	704830	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	murA	0648	0593	M		40
191	705741	705751	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	murA	0648	0593	M		11
192	706861	707448	aspartate ammonia-lyase	aspA	0649	0594	E		588
			hypothetical protein	-	0650	0595	L		
193	709456	709522	nonheme iron-containing ferritin	pfr	0653	0598	P		67
194	709755	709808	no annotation / intergenic						54
195	713541	713565	protective surface antigen D15	-	0655	0600	M	other	25
196	713884	714598	hypothetical protein	-	0656	0601	HR		715
197	718155	718260	hypothetical protein	-	0659	0604	-		106
198	719991	721024	ribonuclease HI	rnhA	0661	0606	L		1034
			ribonuclease III	rnc	0662	0607	K		
199	721250	721269	chorismate synthase	aroC	0663	0608	E		20
200	722271	722292	hypothetical protein	-	0664	0609	-		22
201	723503	727019	coproporphyrinogen III oxidase	hemN_1	0665	0610	H		3517
			glycerol-3-phosphate dehydrogenase	glpC	0666	0611	C		
			hypothetical protein	-	0667/ 0668	0612	R		
			hypothetical protein	-	0669	0613	R		
202	727319	727672	hypothetical protein	-	0669	0613	R		354
203	733577	734125	putative, "similar to jhp0914"	-	-	-	-		549
204	740675	740735	ribonucleotide-diphosphate reductase subunit alpha	nrdA	0680	0621	F		61
205	741020	741248	UDP-N-acetylglucosamine pyrophosphorylase	glmU	0683	0624	M		229
206	741518	742870	UDP-N-acetylglucosamine pyrophosphorylase	glmU	0683	0624	M		1353
			flagellar biosynthesis protein FliP	fliP	0684/ 0685	0625	NU		
207	744804	745371	iron (III) dicitrate transport protein	fecA_1	0686	0626	P	family 4	568
208	746103	747348	ferrous iron transport protein B	feoB	0687	0627	P		1246
209	749718	749915	acetly-coA-acetyltransferase	thl	0690	0638	I		198
210	750741	750822	3-oxoacid coA-transferase subunit A	scoA	0691	0637	I		82
211	763252	765930	DNA gyrase subunit A	gyrA	0701	0641	L		2679
			hypothetical protein	-	0702	0642	-		

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
			putative transcriptional regulator	-	0703	0643	T		
			hypothetical protein	-	0704	-	-		
			excinuclease ABC subunit A	uvrA	0705	0644	L		
212	766915	772018	excinuclease ABC subunit A	uvrA	0705	0644	L		5104
			outer membrane protein/porin	hopE	0706	0645	-	hop	
			S-adenosyl-methyltransferase MraW	mraW	0707	0646	M		
			hypothetical protein	-	0708	0647	-		
			hypothetical protein	-	0709	0648	S		
213	776532	776586	ABC transporter ATP-binding protein	-	0715	0653	R		55
214	795129	795285	hypothetical protein	-	0734	0671	J		157
215	800155	800157	D-alanyl-D-alanine-adding enzyme	murF	0740	0677	M		3
216	800554	800555	D-alanyl-D-alanine-adding enzyme	murF	0740	0677	M		2
217	808062	808075	tRNA (guanine-N(7))-methyltransferase	trmB	0747	0684	R		14
218	813411	813413	flagellar capping protein	fliD	0752	0689	N		3
219	817999	818311	hypothetical protein	-	0759	0696	V		313
220	832244	832832	tyrosyl-tRNA synthetase	tyrS	0774	0711	J		589
221	835981	836094	uridylate kinase	pyrH	0777	0714	F		114
222	836821	837185	hypothetical protein	-	0778	0715	R		365
223	837429	837544	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	acnB	0779	0716	C		116
224	838209	839125	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	acnB	0779	0716	C		917
225	839890	841175	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	acnB	0779	0716	C		1286
			hypothetical protein	-	0780	0717	-		
			hypothetical protein	HPAG1_0766			-		
			hypothetical protein	-	0781	0718	L		
226	849156	849164	putative outer membrane protein	-	0788	0725	-	hof	9
227	854551	855196	putative heavy-metal cation-transporting P-type ATPase	hmcT	0791	0727	P		646
			hypothetical protein	-	0792	0728	O		
228	855553	857603	hypothetical protein	-	0792	0728	O		2051
			peptide deformylase	def	0793	0729	J		
			ATP-dependent Clp protease proteolytic subunit	clpP	0794	0730	OU		
229	863090	863109	GTP cyclohydrolase II	ribA	0802	0738	H		20
230	864621	864954	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein	ribBA	0804	0740	H		334
231	874313	874412	molybdopterin-synthase sulfurylase	moeB	0814	0750	H		100
			flagellar motor protein MotA	motA	0815	0751	N		
232	899033	899276	bifunctional phosphopantothenoylcysteine decarboxylase/ phosphopantothenate synthase	dfp	0841	0779	H		244
233	905767	906276	type I restriction enzyme (restriction subunit)	hsdR_2	0846	0784	V		510
			type I restrictionenzyme (specifity subunit)	hsdS_2	0848/0849	0785	V		
234	909604	910988	hypothetical protein	-	0851	0787	I		1385
			hypothetical protein	-	0852	0788	S		
			ABC transporter ATP-binding protein	-	0853	0789	R		

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
235	912230	912728	ABC transporter ATP-binding protein	-	0853	0789	R		499
236	914608	914611	putative ADP-D-glycero-D-mannoheptose synthase	waaE	0858	0792	M		4
237	916051	919553	ADP-L-glycero-D-mannoheptose-6-epimerase	gmhD	0859	0793	MG		3503
			hypothetical protein	-	0860	0794	E		
			hypothetical protein	-	0861	0795	S		
			pantothenate kinase	-	0862	0796	K		
			hypothetical protein	-	0863	0797	-		
238	922799	924044	lipid-A-disaccharide synthase	lpxB	0867	0801	M		1246
			hypothetical protein	-	0868	0802	-		
			hydrogenase nickel incorporation protein	hypA	0869	0803	R		
			flagellar hook protein FlgE	flgE	0870	0804	N		
239	924521	924647	flagellar hook protein FlgE	flgE	0870	0804	N		127
240	936697	941666	holliday junction DNA helicase RuvA	ruvA	0883	0815	L		4970
			hypothetical protein	-	0884	0816	-		
			hypothetical protein	-	0885	0817	R		
			cysteinyl-tRNA synthetase	cysS	0886	0818	J		
241	942627	942651	vacuolating cytotoxin	vacA	0887	0819	-		25
242	943900	945057	vacuolating cytotoxin	vacA	0887	0819	-		1158
243	946923	948696	putative lipopolysaccharide biosynthesis protein	-		0820	M		1774
			iron (III) dicitrate transport system ATP-binding protein	fecE	0888	0821	PH		
			iron (III) dicitrate transport system permease protein	fecD	0889	0822	P		
244	952101	953177	outer membrane protein - adhesin	babA	1243	0833	-	hop	1077
245	953394	953434	outer membrane protein - adhesin	babA	1243	0833	-	hop	41
246	953717	953792	outer membrane protein - adhesin	babA	1243	0833	-	hop	76
247	954466	954593	no annotation / intergenic						128
248	955257	957031	hypothetical protein	-		0834	O		1775
			hydrogenase expression/formation protein	hypD	0898	0835	O		
			hydrogenase expression/formation protein	hypC	0899	0836	O		
			hydrogenase expression/formation protein	hypB	0900	0837	OK		
249	958269	959802	acetate kinase	ackA	0903	0840	C		1534
			phosphotransacetylase	pta	0904/0905	0841	C		
250	963460	965937	flagellar basal bodyl rod modification protein	flgD	0907	0843	N		2478
			flagellar hook protein FlgE	flgE	0908	0844	N		
			hypothetical protein	-	0909	0845	-		
251	966424	967371	type II DNA modification enzyme (methyltransferase)	-	0910	0846	L		948
			putative ATP-dependent helicase	-	0911	0847	L		
252	967851	968433	putative ATP-dependent helicase	-	0911	0847	L		583
253	970250	970454	outer membrane protein/porin	hopC	0912	0848	-	hop	205
254	970884	970996	outer membrane protein/porin	hopC	0912	0848	-	hop	113
255	971499	971796	outer membrane protein/porin	hopC	0912	0848	-	hop	298
256	972555	972891	outer membrane protein/porin	hopB	0913	0849	-	hop	337
257	973113	974779	outer membrane protein/porin	hopB	0913	0849	-	hop	1667

	Position in NQ352		**Annotation [a]**						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
			putative outer membrane protein	-	0914	0850	-	hof	
258	977575	977726	putative iron-regulated outer membrane protein	frpB_2	0915/ 0916	0851	P	family 4	152
259	978893	979135	hypothetical protein	-	0918	0852	J		243
			carbamoyl-phosphate synthase large subunit	carB	0919	0853	EF		
260	985405	989484	putative vacuolating cytotoxin (VacA) paralog	-	0922	0856	N		4080
261	991367	991441	putative vacuolating cytotoxin (VacA) paralog	-	0922	0856	N		75
262	992299	992313	recombination protein RecR	recR	0925	0859	L		15
263	994617	994801	heat shock protein HtpX	htpX	0927	0861	O		185
			GTP cyclohydrolase I	folE	0928	0863	H		
264	995847	995985	geranyltranstransferase	ispA	0929	0864	H		139
265	998744	999571	hypothetical protein	-	0934	0868	O		828
			hypothetical protein	-	0935	0869	R		
266	1002962	1003540	proline/betaine transporter	proP	0936	0871	-		579
267	1003743	1004543	hypothetical protein	-	0937	0872	-		801
268	1010772	1010971	D-amino acid dehydrogenase	dadA	0943	0878	E		200
			hypothetical protein	-	0944	0879	J		
269	1011238	1011283	hypothetical protein	-	0944	0879	J		46
270	1014453	1015558	hypothetical protein	-	0948	0882	-		1106
271	1017194	1018660	putative recombination protein RecO	-	0951	0885	-		1467
			hypothetical protein	-	0952	0886	R		
			hypothetical protein	-	0953	0887	-		
272	1020422	1021594	prolipoprotein diacylglyceryl transferase	lgt	0955	0889	M		1173
			hypothetical protein	-	0956	0890	J		
			3-deoxy-D-manno-octulosonic-acid transferase	waaA	0957	0891	M		
273	1060686	1061010	hypothetical protein	-		0949	U		325
			hypothetical protein	-		0950	-		
274	1062754	1062833	integrase/recombinase (XERCD family)	-		0951	L		80
275	1066577	1067881	arginine decarboxylase	speA	0422	0962	E		1305
276	1070545	1070620	hypothetical protein	-	0419	0965	-		76
277	1070857	1071928	hypothetical protein	-	0419	0965	-		1072
			hypothetical protein	-	0418	0966	-		
			methionine-tRNA synthetase	metG	0417	0967	J		
278	1073408	1073487	methionine-tRNA synthetase	metG	0417	0967	J		80
279	1074410	1074910	cyclopocyclopropane fatty acid synthase	cfa	0416	0968	M		501
280	1075501	1076656	hypothetical protein	-	0415	0969	M		1156
281	1077408	1077702	hypothetical protein	-	0412	0970	-		295
282	1077963	1078006	no annotation / intergenic						44
283	1078273	1079007	hypothetical protein	-	0411		-		735
			putative paralogue of HpaA	-	0410	0971	-		
284	1079466	1080397	GMP synthase	guaA	0409	0972	F		932
285	1080926	1081014	hypothetical protein	-	0408	0973	-		89
286	1084351	1084366	hypothetical protein	-	0406	0975	-		16
287	1087588	1087623	phenylalanyl-tRNA synthetase subunit beta	pheT	0402	0979	J		36
288	1088184	1088211	phenylalanyl-tRNA synthetase subunit beta	pheT	0402	0979	J		28
289	1089468	1091103	phenylalanyl-tRNA synthetase subunit beta	pheT	0402	0979	J		1636

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
			3-phosphoshikimate 1-carboxyvinyltransferase	aroA	0401	0980	E		
			4-hydroxy-3-methylbut-2-enyl diphosphate reductase	ispH	0400	0981	IM		
290	1092076	1092184	30S ribosomal protein S1	rpsA	0399	0982	J		109
291	1092502	1092784	30S ribosomal protein S1	rpsA	0399	0982	J		283
292	1095098	1095233	D-3-phosphoglycerate dehydrogenase	serA	0397	0984	HE		136
293	1096652	1097225	hypothetical protein	-	0396	0985	H		574
294	1098838	1099479	hypothetical protein	-	0394	0987	S		642
			putative chemotaxis protein	cheV_3	0393	0988	NT		
295	1100234	1102217	histidine kinase	cheA	0392	0989	NT		1984
296	1102453	1103993	histidine kinase - MCP coupling protein	cheW	0391	0990	NT		1541
			thiol peroxidase	tpx	0390	0991	O		
			iron-dependent superoxide dismutase	sodF	0389	0992	P		
297	1104232	1106127	iron-dependent superoxide dismutase	sodF	0389	0992	P		1896
			hypothetical protein	-	0388	0993	H		
			primosome assembly protein PriA	priA	0387	0994	L		
298	1106507	1108214	primosome assembly protein PriA	priA	0387	0994	L		1708
			hypothetical protein	-	0386	0995	-		
			hypothetical protein	-	0385	0996	-		
			hypothetical protein	-	0384	0997	-		
299	1111534	1113187	protoporphyrinogen oxidase	hemG	0381	1000	J		1654
			glutamate dehydrogenase	gdhA	0380	1001	E		
			alpha-(1,3)-fucosyltransferase	fucU	0379	1002	-		
300	1113253	1113287	alpha-(1,3)-fucosyltransferase	fucU	0379	1002	-		35
301	1113523	1114226	putative cytochrome C-type biogenesis protein	-	0378	1003	O		704
302	1114714	1115283	putative cytochrome C-type biogenesis protein	-	0378	1003	O		570
303	1115847	1116351	putative cytochrome C-type biogenesis protein	-	0378	1003	O		505
			hypothetical protein	-	0377	1004	O		
304	1118119	1118120	hypothetical protein	-	0375	1006	-		2
305	1119408	1120243	putative outer membrane protein	-	0373	1008	-	hom	836
306	1120557	1122747	putative outer membrane protein	-	0373	1008	-	hom	2191
			deoxycytidine triphosphate deaminase	dcd	0372	1009	F		
			biotin carboxyl carrier protein	accB	0371	1010	I		
			biotin carboxylase	accC	0370	1011	I		
307	1124961	1128451	putative type II DNA modification enzyme (methyltransferase)	-	0369	1012	-		3491
			hypothetical protein	-	0368	1013	-		
			hypothetical protein	-	0367	1014	-		
			spore coat polysaccharide biosynthesis protein C	-	0366	1015	M		
			hypothetical protein	-	0365		-		
308	1129548	1130590	ribonucleotide-diphosphate reductase subunit beta	nrdF	0364	1016	F		1043
			protein-L-isoaspartate O-methyltransferase	pcm	0363	1017	O		
			hypothetical protein	-	0362	1018	R		
309	1152255	1152913	adenylosuccinate lyase	purB	1112	1039	F		659

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
			putative outer membrane protein	-	1113	1040	-	hor	
310	1153306	1153309	putative outer membrane protein	-	1113	1040	-	hor	4
311	1153542	1155458	excinuclease ABC subunit B	uvrB	1114	1041	L		1917
312	1165776	1165781	type II DNA modification enzyme (methyltransferase)	-	1121	1050	L		6
313	1173515	1174168	F0F1 ATP synthase subunit beta	atpD	1132	1060	C		654
			F0F1 ATP synthase subunit gamma	atpG	1133	1061	C		
314	1175397	1175961	F0F1 ATP synthase subunit alpha	atpA	1134	1062	C		565
315	1176400	1176530	F0F1 ATP synthase subunit delta	atpH	1135	1063	C		131
316	1180406	1181083	methionyl-tRNA formyltransferase	fmt	1141	1069	J		678
317	1181744	1184501	ATPase, HPP12_1108	-			PH		2758
			hypothetical protein	-	1143	1071	S		
318	1201754	1201882	putative outer membrane protein	-	1157	1084	-	hop	129
319	1202818	1202831	putative outer membrane protein	-	1157	1084	-	hop	14
320	1206831	1206920	hypothetical protein	-	1162	1089	S		90
321	1213235	1214027	hypothetical protein	-	1168	1095	T		793
322	1214836	1218177	hypothetical protein	-	1168	1095	T		3342
			amino acid ABC transporter permease	-	1169	1096	E		
			amino acid ABC transporter permease	-	1170	1097	E		
			amino acid ABC transporter ATP-binding protein	-	1171	1098	E		
			amino acid ABC transporter	-	1172	1099	ET		
323	1218694	1219204	hypothetical protein	-	1173	1100	-		511
324	1219841	1222739	glucose/galactose transporter	gluP	1174	1101	G		2899
			hypothetical protein	-	1175	1102	R		
			putative outer membrane function	-	1177	1103	-	hop	
325	1223060	1223250	putative outer membrane function	-	1177	1103	-	hop	191
326	1224452	1225024	hypothetical protein	HPAG1_1116			-		573
			purine nucleoside phosphorylase	deoD	1178	1104	F		
327	1225426	1225583	purine nucleoside phosphorylase	deoD	1178	1104	F		158
			phosphopentomutase	deoB	1179	1105	G		
328	1225908	1225965	phosphopentomutase	deoB	1179	1105	G		58
329	1226627	1226693	phosphopentomutase	deoB	1179	1105	G		67
330	1232892	1233172	hypothetical protein	-	1184	1110	V		281
			sugar efflux transporter	-	1185	1111	G		
331	1236944	1237605	aspartate-semialdehyde dehydrogenase	asd	1189	1114	E		662
332	1237875	1238518	aspartate-semialdehyde dehydrogenase	asd	1189	1114	E		644
			histidyl-tRNA synthetase	hisS	1190	1115	J		
333	1238773	1241697	histidyl-tRNA synthetase	hisS	1190	1115	J		2925
			ADP-heptose-LPS heptosyltransferase II	waaF	1191	1116	M		
			putative motility protein	-	1192	1117	-		
			aldo-keto reductase, putative	-	1193		C		
334	1241902	1241955	aldo-keto reductase, putative	-	1193		C		54
335	1242715	1243966	elongation factor G	fusA	1195	1118	J		1252
336	1246495	1247434	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		940
337	1250773	1250962	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		190

	Position in NQ352		**Annotation [a]**						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
338	1251682	1252936	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		1255
339	1256874	1257566	50S ribosomal protein L11	rplK	1202	1125	J		693
			transcription antitermination protein NusG	nusG	1203	1126	K		
			preprotein translocase subunit SecE	secE	1203	1126a	-		
340	1257809	1258301	50S ribosomal protein L33	rpmG	1204	1127	J		493
			elongation factor Tu	tuf	1205	1128	J		
341	1272475	1272732	glycinamide ribonucleotide synthetase	purD	1218	1140	F		258
342	1273152	1273353	hypothetical protein	-	1219		-		202
			putative ABC transporter, ATP-binding protein	-	1220	1141	V		
343	1279761	1280570	coproporphyrinogen III oxidase	hemN_2	1226	1147	H		810
			periplamic cytochrome C-553	-	1227	1148	C		
344	1281795	1281909	aspartate kinase	lysC	1229	1150	E		115
345	1284894	1285397	hypothetical protein	-	1233	1154	-		504
346	1287585	1288045	hypothetical protein	-	1235	1156	-		461
			hypothetical protein	-	1236	1157	S		
347	1290817	1291848	hypothetical protein	-	1239		-		1032
			hypothetical protein	-		1160	-		
			Maf-like protein	maf	1240	1161	D		
348	1293248	1293401	alanyl-tRNA synthetase	alaS	1241	1162	J		154
349	1294648	1295102	alanyl-tRNA synthetase	alaS	1241	1162	J		455
			hypothetical protein	-	1242	1163	S		
350	1295730	1297432	outer membrane protein - adhesin	babB	0896	1164	-	hop	1703
351	1300432	1300433	ribonuclease II family protein	vacB	1248	1169	K		2
352	1308505	1312772	preprotein translocase subunit SecG	secG	1255	1176	U		4268
			ribosome recycling factor	frr	1256	1177	J		
			orotate phosphoribosyltransferase	pyrE	1257	1178	F		
			hypothetical protein	-	1258	1179	-		
			hypothetical protein	-	1259	1180	K		
			NADH dehydrogenase subunit A	nuoA	1260	1181	C		
			NADH dehydrogenase subunit B	nuoB	1261	1182	C		
			NADH dehydrogenase subunit C	nuoC	1262	1183	C		
353	1313205	1313751	NADH dehydrogenase subunit D	nuoD	1263	1184	C		547
354	1314204	1315722	NADH dehydrogenase subunit D	nuoD	1263	1184	C		1519
			putative NADH oxidoreductase I	nuoE	1264	1185	-		
			putative NADH oxidoreductase I	nuoF	1265	1186	-		
			NADH dehydrogenase subunit G	nuoG	1266	1187	C		
355	1316818	1316868	NADH dehydrogenase subunit G	nuoG	1266	1187	C		51
356	1318466	1321283	NADH dehydrogenase subunit H	nuoH	1267	1188	C		2818
			NADH dehydrogenase subunit I	nuoI	1268	1189	C		
			NADH dehydrogenase subunit J	nuoJ	1269	1190	C		
			NADH dehydrogenase subunit K	nuoK	1270	1191	C		
			NADH dehydrogenase subunit L	nuoL	1271	1192	CP		
357	1331924	1333136	tryptophan synthase subunit beta	trpB	1278	1199	E		1213
			bifunctional indole-3-glycerol phosphate synthase/phosphoribosylanthranilate isomerase	trpC	1279	1200	E		

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
358	1333801	1333836	anthranilate phosphoribosyltranferase	trpD	1280	1201	E		36
359	1334352	1334953	anthranilate synthase component II	trpG	1281	1202	EH		602
			anthranilate synthase component I	trpE	1282	1203	EH		
360	1335272	1335668	anthranilate synthase component I	trpE	1282	1203	EH		397
361	1339139	1339788	hypothetical protein	-	1286	1206	S		650
			putative transcriptional regulator	-	1287	1207	K		
362	1340596	1341163	nicotinamide mononucleotide transporter	pnuC	1290	1210	H		568
363	1350302	1352465	30S ribosomal protein S14	rpsN	1306	1226	J		2164
			50S ribosomal protein L5	rplE	1307	1227	J		
			50S ribosomal protein L24	rplX	1308	1228	J		
			50S ribosomal protein L14	rplN	1309	1229	J		
			30S ribosomal protein S17	rpsQ	1310	1230	J		
			50S ribosomal protein L29	rpmC	1311	1231	-		
			50S ribosomal protein L16	rplP	1312	1232	J		
364	1352698	1353220	30S ribosomal protein S3	rpsC	1313	1233	J		523
365	1353433	1353554	50S ribosomal protein L22	rplV	1314	1234	J		122
366	1358149	1358578	hypothetical protein	-	1322	1242	-		430
367	1358803	1358804	ribonuclease HII	rnhB	1323	1243	L		2
368	1362051	1366180	fumarate hydratase	fumC	1325	1245	C		4130
			hypothetical protein	-	1326	1246	-		
			hypothetical protein	-	1327	1247	MU	family 5	
			cation efflux system protein	czcB_2	1328	1248	M		
369	1366428	1366752	cation efflux system protein	czcA_2	1329	1249	P		325
370	1371765	1371767	hypothetical protein	-	1333	1253	-		3
371	1377802	1377811	no annotation / intergenic						10
372	1378284	1378886	hypothetical protein	-	1343	1262	P		603
			magnesium/ cobalt transporter CorA	corA	1344	1263	P		
373	1379483	1381284	magnesium/ cobalt transporter CorA	corA	1344	1263	P		1802
			phosphoglycerate kinase	pgk	1345	1264	G		
			glyceraldehyde-3-phosphate dehydrogenase	gap_2	1346	1265	G		
374	1381700	1381996	glyceraldehyde-3-phosphate dehydrogenase	gap_2	1346	1265	G		297
			uracil-DNA glycosylase	ung	1347	1266	L		
375	1382211	1382629	uracil-DNA glycosylase	ung	1347	1266	L		419
376	1382889	1383499	1-acyl-SN-glycerol-3-phosphate acyltransferase	plsC	1348	1267	I		611
			hypothetical protein	-	1349	1268	-		
377	1384447	1384461	hypothetical protein	-	1349	1268	-		15
378	1385137	1385455	carboxyl-terminal protease	prc	1350	1269	M		319
379	1386062	1386447	hypothetical protein	-	1351	1270	-		386
380	1386697	1388100	hypothetical protein	-	1351	1270	-		1404
			type II DNA modification enzyme (methyltransferase)	-	1352	1271	L		
381	1388856	1388922	hypothetical protein	-	1353/ 1354	1272	R		67
382	1389555	1390977	hypothetical protein	-	1353/ 1354	1272	R		1423
383	1392902	1393053	quinolinate synthetase	nadA	1356	1274	H		152
384	1394193	1394678	phosphatidylserine decarboxylase	psd	1357	1275	I		486
			hypothetical protein	-	1358	1276	-		

	Position in NQ352		Annotation [a]						
No	**start**	**end**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
385	1395128	1396918	hypothetical protein	-	1359	1277	-		1791
			prenyltransferase	ubiA	1360	1278	H		
			DNA transfer protein	comEC	1361	1279	R		
386	1397203	1401597	DNA transfer protein	comEC	1361	1279	R		4395
			replicative DNA helicase	dnaB	1362	1280	L		
			hypothetical protein	-	1363	1281	S		
			putative histidine kinase sensor protein	-	1364	1282	T		
387	1401816	1402333	putative transcriptional regulator	-	1365	1283	TK		518
388	1402371	1402396	putative transcriptional regulator	-	1365	1283	TK		26
389	1418863	1418918	"putative" plasticity zone and typeIV secretion system tfs3 gene cluster	-			-		56
390	1420651	1420805	"putative" plasticity zone and typeIV secretion system tfs3 gene cluster	-			-		155
391	1421971	1422070	hypothetical protein	-		0935	D		100
392	1422336	1422413	hypothetical protein	-		0935	D		78
393	1427449	1427560	"putative"				-		112
394	1453798	1453888	type III restriciton enzyme R protein	-	1371	1285	-		91
395	1456398	1458370	rod shape-determining protein MreB	mreB	1373	1287	D		1973
			ATP-dependent protease ATP-binding subunit ClpX	clpX	1374	1288	O		
396	1459497	1460383	(3R)-hydroxymyristoyl-ACP dehydratase	fabZ	1376	1290	I		887
			flagellar assembly protein FliW	-	1377	1291	S		
397	1461076	1461286	hypothetical protein	-	1378	1292	R		211
398	1466174	1466782	putative type III DNA modification enzyme (methyltransferase)	mod_1		1296	L		609
399	1467123	1468136	putative type III DNA modification enzyme (methyltransferase)	mod_1		1296	L		1014
			putative type III restriction enzyme	res_1		1297	S		
400	1470438	1472749	putative type III restriction enzyme	res_1		1297	S		2312
			biotin synthase	bioB	1406	1298	H		
			putative ribonuclease N	rbn	1407	1299	S		
401	1472755	1472777	no annotation / intergenic						23
402	1473244	1473300	no annotation / intergenic						57
403	1477311	1477480	hypothetical protein	-	1409	1301	S		170
404	1478393	1479308	hypothetical protein	-	1409	1302	-		916
			hypothetical protein	-	1410	1303	L		
405	1479621	1480414	hypothetical protein	-	1410	1303	L		794
			hypothetical protein	-	1411	1304	-		
406	1481558	1481588	hypothetical protein	-	1411	1305	-		31
407	1487879	1488110	UDP-N-acetylenolpyruvoylglucosamine reductase	murB	1418	1313	M		232
			flagellar biosynthesis protein FliQ	fliQ	1419	1314	NU		
408	1488693	1489341	flagellum-specific ATP synthase	fliI	1420	1315	NU		649
409	1491132	1492189	isoleucyl-tRNA synthetase	ileS	1422	1317	J		1058
410	1494528	1494594	no annotation / intergenic						67
411	1496859	1499776	hypothetical protein	-	1430	1323	R		2918
			hypothetical protein	-	1429	1324	M		
			hypothetical protein	-	1428	1325	R		
412	1500464	1502546	hypothetical protein	-	1433	1326	-		2083

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
413	1509925	1510029	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	-	1443	1336	I		105
			SsrA-binding protein	smpB	1444	1337	O		
414	1515176	1515527	tRNA modification GTPase TrmE	trmE	1452	1345	R		352
415	1515777	1516593	tRNA modification GTPase TrmE	trmE	1452	1345	R		817
			putative outer membrane protein	-	1453	1346	-	hom	
416	1519381	1520100	hypothetical protein	-	1454	1347	-		720
			hypothetical protein	-	1455	1348	-		
417	1523726	1526902	DNA polymerase III subunit alpha	dnaE	1460	1353	L		3177
418	1545640	1551952	thymidylate kinase	tmk	1474	1367	F		6313
			phosphopantetheine adenylyltransferase	coaD	1475	1368	H		
			3-octaprenyl-4-hydroxybenzoate carboxy-lyase	ubiD	1476	1369	H		
			flagellar basal body P-ring biosynthesis protein FlgA	flgA	1477	1370	NO		
			putative ATP-dependent DNA helicase	rep	1478	1371	L		
			hypothetical protein	-	1479	1372	O		
419	1556074	1556274	hypothetical protein	-	1485	1378	S		201
420	1566606	1566607	transaldolase	tal	1495	1388	G		2
421	1571558	1572241	putative outer membrane protein	-	1501	1394	-	hor	684
422	1572450	1572576	hypothetical protein	-	1502	1395	S		127
423	1572798	1574085	hypothetical protein	-	1502	1395	S		1288
			putative component of cation transport for cbb3-type oxidase	fixI	1503	1396	P		
424	1574607	1574741	putative component of cation transport for cbb3-type oxidase	fixI	1503	1396	P		135
425	1578170	1580172	hypothetical protein	-	1507	1400	E		2003
			ferredoxin-like protein	-	1508	1401	C		
426	1581042	1581105	putative glycerol-3-phosphate acyltransferase PlsY	-	1509	1402	S		64
427	1581470	1581472	dihydroneopterin aldolase	folB	1510	1403	H		3
			hypothetical protein	-	1511	1404	-		
428	1584657	1587413	selenocysteine synthase	selA	1513	1406	E		2757
			transcription elongation factor NusA	nusA	1514	1407	K		
			hypothetical protein	-		1408	-		
429	1587834	1588309	hypothetical protein	-		1408	-		476
430	1591444	1591829	putative type II DNA modification enzyme (methyltransferaese)	-		1409	V		386
431	1592186	1599163	putative type II DNA modification enzyme (methyltransferaese)	-		1409	V		6978
			type III restriction enzyme	res_2	1521	1410	V		
			type III DNA modification enzyme (methyltransferase)	mod_2	1522	1411	L		
			ATP-dependent DNA helicase RecG	recG	1523	1412	LK		
432	1603386	1603643	chromosomal replication initiation protein	dnaA	1529	1417	L		258
433	1606960	1607708	glucosamine-fructose-6-phosphate aminotransferase	glmS	1532	1420	M		749
			FAD-dependent thymidylate synthase	thyX	1533	1421	F		
434	1611754	1612728	type I restriction enzyme modification subunit	hsdM_3	1403	1423	V		975
			type I restriciton enzyme restriction subunit	hsdR_3	1402	1424	V		
435	1618518	1618651	arginase	rocF	1399	1427	E		134

	Position in NQ352		Annotation [a]						
No	start	end	Description	gene	HP-number	jhp-number	COG	OMP	Import length
436	1619944	1620040	amino acid permease -HPAG1_1471	-		-	E		97
437	1621021	1621670	amino acid permease -HPAG1_1471	-		-	E		650
			L-alanine dehydrogenase	ald	1398	1428	E		
438	1627671	1629463	DNA repair protein (recombination protein N)	recN	1393	1434	L		1793
			fibronectin/ fibrinogen-binding protein	-	1392	1435	K		
439	1634151	1635370	ribulose-phosphate 3-epimerase	rpe	1386	1439	G		1220
			fructose-1,6-bisphosphatase	fbp	1385	1440	G		
			hypothetical protein	-	1384	1441	-		
440	1635627	1636901	type IIS restriction enzyme M2 protein (mod)	-	1368		L		1275
			type IIS restriction enzyme M1 protein (mod)	-	1367		L		
			type II restriction enzyme	-	1366	1442	-		
441	1638527	1639210	type II restriction enzyme	-	1366	1442	-		684

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S10: Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ392/4060. Positions refer to the position in the virtual genome of the first isolate NQ392.

	Position in NQ392			**Annotation**[a]						
No	**start**	**end**	**Corresponding import in NQ1707**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
1	181	340		putative type II DNA modification enzyme (methyltransferase)	-		1409	V		160
2	19264	19897		putative type I restriction enzyme specificity subunit	hsdS_3a		1422	V		634
				type I restriction enzyme modification subunit	hsdM_3	1403	1423	V		
3	21205	21303		type I restriction enzyme modification subunit	hsdM_3	1403	1423	V		99
4	103596	104211		putative outer membrane protein	-	0009	0007	-	hop	616
5	265304	265362		putative transcriptional regulator	-	0166	0152	TK		59
6	355503	355692		ABC transporter, ATP-binding protein	-	0250	0235	R		190
7	359083	359130		putative outer membrane protein	-	0253/ 0254	0238	-	hop	48
8	411601	411698		periplasmic dipeptide transport substrate-binding protein	dppA	0298	0283	E		98
9	411948	412642		periplasmic dipeptide transport substrate-binding protein	dppA	0298	0283	E		695
10	413618	413987		peptide ABC transporter permease	dppB	0299	0284	EP		370
11	430781	430785		no annotation/ intergenic						5
12	432110	432153		outer membrane protein-adhesin	babA	1243	0833	-	hop	44
13	432994	433236		hypothetical protein	-	0318	0301	P		243
14	433556	433674		hypothetical protein	-	0318	0301	P		119
15	704485	704795		putative efflux transporter	-	0606	0553	M		311
				putative efflux transporter	-	0607	0554	V		
16	705305	705512		putative efflux transporter	-	0607	0554	V		208
17	760307	760315		alpha (1,3)-fucosyltransferase	fucT	0651	0596	-		9
18	787411	787602		hypothetical protein	-		0616	-		192
19	789932	790082		"putative" - similar to jhp0914				-		151
20	799727	799819		iron(III) dicitrate transport protein	fecA_1	0686	0626	P	family 4	93
21	841413	841548		putative outer membrane protein	-	0722	0659	-	hop	136
22	846387	846389	CNP 4	putative outer membrane protein	-	0725	0662	-	hop	3
23	846996	847011	CNP 5	putative outer membrane protein	-	0725	0662	-	hop	16
24	913426	913441	*CNP 6 (partially)*	anti-codon nuclease masking agent	prrB	0790		V		16
25	936676	936891		flagellar motor protein MotB	motB	0816	0752	N		216
				hypothetical protein	-	0817	0753	-		
26	958047	958165		putative outer membrane protein	-	0839	0777	I	other	119
27	1076240	1079370		prolipoprotein diacylglyceryl transferase	lgt	0955	0889	M		3131
				hypothetical protein	-	0956	0890	J		
				3-deoxy-D-manno-octulosonic-acid transferase	waaA	0957	0891	M		

	Position in NQ392			**Annotation**[a]						
No	**start**	**end**	**Corresponding import in NQ1707**	**Description**	**gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
				hypothetical protein	-	0958	0892	R		
				hypothetical protein	-	0959	0893	S		
28	1079679	1079977		hypothetical protein	-	0959	0893	S		299
				glycyl-tRNA synthetase subunit alpha	glyQ	0960	0894	J		
29	1189463	1189597		protein-L-isoaspartate O-methyltransferase	pcm	0363	1017	O		135
30	1204822	1204874		no annotation/ intergenic						53
31	1244241	1244417		no annotation/ intergenic						177
32	1247046	1247215		hypothetical protein	-	1145 /1146	1073	S		170
33	1279671	1280404		hypothetical protein	-	1175	1102	R		734
34	1282180	1282562		putative outer membrane function	-	1177	1103	-	hop	383
35	1282853	1283089		putative outer membrane function	-	1177	1103	-	hop	237
36	1301534	1301602		putative motility protein	-	1192	1117	-		69
37	1413271	1413304		50S ribosomal protein L16	rplP	1312	1232	J		34
38	1419026	1419188		hypothetical protein	-	1322	1242	-		163
39	1437957	1437984		putative outer membrane protein	-	1342	1261	-	hop	28
40	1571118	1571131		no annotation/ intergenic						14
41	1646081	1646547		selenocysteine synthase	selA	1513	1406	E		467
				transcription elongation factor NusA	nusA	1514	1407	K		

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S11: **Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ367/4191. Positions refer to the position in the virtual genome of the first isolate NQ367.**

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
1	54489	54515	-	6,7-dimethyl-8-ribityllumazine synthase	ribH	0002	0002	H		27
2	88260	89334	-	hypothetical protein	-	0035	0031	S		1075
				hypothetical protein	-	0036	0032	-		
				hypothetical protein	-	0037	0033	U		
3	92434	93107	-	DNA transformation competancy	comB3	0041/ 0042	0036	U		674
4	110917	110919	-	proline/pyrroline-5-carboxylate dehydrogenase	putA	0056	0048	C		3
5	116116	116144	-	hypothetical protein	-	0060	0055	-		29
6	117842	118200	-	hypothetical protein	-	0063	0058	-		359
7	118486	118564	-	hypothetical protein	-	0063	0058	-		79
8	121924	122022	-	hypothetical protein	-	0066	0061	D		99
9	127317	129003	-	urease subunit beta	ureB	0072	0067	E		1687
				urease subunit alpha	ureA	0073	0068	E		
				lipoprotein signal peptidase	lspA	0074	0069	MU		
				phosphoglucosamine mutase	glmM	0075	0070	G		
10	136704	138632	-	methyl-accepting chemotaxis protein (MCP)	-	0082	0075	NT		1929
11	140648	140926	-	hypothetical protein	-	0086	0079	R		279
12	154047	154439	-	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		393
13	161882	161932	-	2',3'-cyclic-nucleotide 2'-phosphodiesterase	cpdB	0104	0096	F		51
14	178652	178824	-	hypothetical protein	-	0118	0110	-		173
15	184565	184568	CNP 1	translation initiation factor IF-3	infC	0124	0114	J		4
16	184985	185277	CNP 2	translation initiation factor IF-3	infC	0124	0114	J		293
				50S ribosomal protein L35	rpmI	0125	0115	J		
17	185526	186814	CNP 3	50S ribosomal protein L20	rplT	0126	0116	J		1289
				putative outer membrane protein	-	0127	0117	-	hor	
18	199704	199784	SNP	L-lactate permease	lldP_1	0140	0128	C		81
19	200401	201190	*CNP 4 (partially)*	L-lactate permease	lldp_2	0141	0129	C		790
20	201646	203016	*CNP 5 (partially)*	DNA glycosylase MutY	mutY	0142	0130	L		1371

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
				hypothetical protein	-	0143	0131	P		
21	203223	203470	*CNP 6 (partially)*	hypothetical protein	-	0143	0131	P		248
22	207219	207815	CNP 7	cytochrome oxidase (CBB3-TYPE)	fixP	0147	0135	C		597
				hypothetical protein	-	0148	0136	-		
				hypothetical protein	-	0149	0137	-		
23	216807	217075	*CNP 8 (partially)*	hypothetical protein	-	0160	0148	R		269
24	239750	240212	-	apolipoprotein N-acyltransferase	lnt	0180	0168	M		463
25	263546	263696	-	hypothetical protein	-	0205	0191	E		151
26	267351	267854	-	putative outer membrane protein	-	0209	0195	-	hof	504
27	279747	281472	-	hypothetical protein	-	0219	0205	-		1726
				cysteine desulfurase	-	0220	0206	E		
28	283552	284660	-	DNA repair protein RadA	radA	0223	0209	O		1109
				bifunctional methionine sulfoxide reductase A/B protein	-	0224	0210	O		
29	287257	287294	-	putative outer membrane protein	-	0227	0212	-	hop	38
30	287300	288386	-	putative outer membrane protein	-	0227	0212	-	hop	1087
31	291746	292044	-	3-deoxy-manno-oculosonate cytidylyltransferase	kdsB	0230	0215	M		299
32	293524	294489	-	hypothetical protein	-	0232	0217	-		966
				hypothetical protein	-	0233	0218	E		
33	296879	297905	-	hypothetical protein	-	0235	0220	R		1027
				hypothetical protein	-	0236	0221	-		
34	320772	322081	-	exodeoxyribonuclease VII large subunit	xseA	0259	0243	L		1310
				type II DNA modification enzyme (methyltransferase)	-	0260	0244	L		
				hypothetical protein	-	0261	0245	-		
35	323818	324442	-	type II DNA modification enzyme (methyltransferase)	-	0263	0248	L		625
				heat shock protein	clpB	0264	0249	O		
36	345358	345391	-	hypothetical protein	-	0284	0269	M		34
37	355391	357551	***SNP***	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		2161
38	368110	368533	-	periplasmic dipeptide transport substrate-binding protein	dppA	0298	0283	E		424
39	397362	397364	-	cell division inhibitor	minD	0331	0314	D		3
40	458034	460119	-	translation initiation factor IF-2	infB	1048	0377	J		2086
41	460375	460526	-	ribosome binding factor A	rbfA	1047	0378	J		152
42	496825	497803	-	polyphosphate kinase	ppk	1010	0413	P		979
43	498444	498560	-	no annotation/ intergenic						117

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
44	498927	499856	-	type I restriction enzyme S protein (hsdS)	-	0462		V		930
				type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		
45	500072	500082	-	type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		11
46	500829	501361	-	type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		533
				fragment of type I restriction enzyme R protein (hsdR)	-	0464	0416	V		
47	527296	527692	-	outer membrane protein	-	0486	0438	-	hof	397
48	541655	542116	-	putative phospholipase A1	pldA	0499	0451	M		462
49	542790	542797	-	DNA polymerase III subunit beta	dnaN	0500	0452	L		8
50	550896	551456	-	hypothetical protein	-	0508	0458	-		561
				putative glycolate oxidase	glcD	0509	0459	C		
51	576609	577181	-	cag island protein	orf15	0528	0477	U		573
52	577417	577428	-	cag island protein	orf16	0529	0478	-		12
53	597023	597257	*CNP 10 (partially)*	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		235
54	598485	598614	-	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		130
55	608002	608770	-	acetyl-coenzyme A carboxylase subunit alpha	accA	0557	0504	I		769
				3-oxoacyl-(acyl carrier protein) synthase II	fabB	0558	0505	IQ		
56	665272	665368	-	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		97
57	690907	692173	-	hydrogenase, small subunit	hyaA	0631	0574	C		1267
				hydrogenase, large subunit	hyaB	0632	0575	C		
58	696460	696612	-	hypothetical protein	-	0636	0579	-		153
59	736230	736459	-	integrase-recombinase protein (XERCD family)	-	0675	0617	L		230
60	736667	737772	-	integrase-recombinase protein (XERCD family)	-	0675	0617	L		1106
				methylated-DNA-protein-cysteine methyltransferase	ogt	0676	0618	L		
61	738035	739130	-	hypothetical protein	-	0677	0619	R		1096
				putative lipopolysaccharide biosynthesis protein	-	0679	0620	R		
62	739344	739594	-	putative lipopolysaccharide biosynthesis protein	-	0679	0620	R		251
63	739836	741206	-	ribonucleotide-diphosphate reductase subunit alpha	nrdA	0680	0621	F		1371
64	741821	741977	-	ribonucleotide-diphosphate reductase subunit alpha	nrdA	0680	0621	F		157
65	742558	742624	-	hypothetical protein	-	0681	0622	-		67
66	787373	788028	SNP	putative outer membrane protein	-	0725	0662	-	hop	656

	Position in NQ367			Annotation[a]						
No	start	end	Corresponding import in NQ1671	Description	Gene	HP-number	jhp-number	COG	OMP	Import length
67	788647	789174	-	putative outer membrane protein	-	0725	0662	-	hop	528
68	792285	794068	-	putative outer membrane protein	-	0722	0659	-	hop	1784
				putative outer membrane protein	-	0726	0663	-	other	
69	828238	828245	-	phosphodiesterase	-	0760	0697	R		8
				hypothetical protein	-	0761	0698	H		
70	849294	850422	-	hypothetical protein	-	0781	0718	L		1129
				putative outer membrane protein	-	0782	0719	-	hof	
71	870771	871722	-	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein	ribBA	0804	0740	H		952
72	872477	874100	-	putative lipopolysaccharide biosynthesis protein	-	0805	0741	M		1624
				hypothetical protein	-	0806	0742	R		
				iron(III) dicitrate transport protein	fecA_2	0807	0743	P	family 4	
73	880037	880209	-	molybdopterin-synthase sulfurylase	moeB	0814	0750	H		173
74	880733	881135	-	molybdopterin-synthase sulfurylase	moeB	0814	0750	H		403
				flagellar motor protein MotA	motA	0815	0751	N		
75	886549	888062	-	excinuclease ABC subunit C	uvrC	0821	0760	L		1514
				homoserine dehydrogenase	hom	0822	0761	E		
76	893161	894358	-	F0F1 ATP synthase subunit A	atpB	0828	0767	C		1198
				inosine-5'-monophosphate dehydrogenase	guaB	0829	0768	F		
77	895074	895082	-	aspartyl/glutamyl-tRNA amidotransferase subunit A	gatA	0830	0769	J		9
78	900054	902239	-	GTP-binding protein EngA	engA	0834	0773	R		2186
				DNA-binding protein Hu	-	0835	0774	L		
				hypothetical protein	-	0836/ 0837	0775	S		
				hypothetical protein	-	0838	0776	-		
79	910113	910373	-	type I restriction enzyme (restriction subunit)	hsdR_2	0846	0784	V		261
80	911288	913314	-	type I restriction enzyme (restriction subunit)	hsdR_2	0846	0784	V		2027
				anti-codon nuclease masking agent (prrB)	-	0790		V		
81	919474	921004	-	guanosine 5'-monophosphate oxidoreductase	guaC	0854	0790	F		1531
				phosphoheptose isomerase	gmhA	0857	0791	G		
				putative ADP-D-glycero-D-mannoheptose synthase"	waaE	0858	0792	M		
82	921214	922287	-	putative ADP-D-glycero-D-mannoheptose synthase	waaE	0858	0792	M		1074
				ADP-L-glycero-D-mannoheptose-6-epimerase	gmhD	0859	0793	MG		
83	923004	923007	-	ADP-L-glycero-D-mannoheptose-6-epimerase	gmhD	0859	0793	MG		4

	Position in NQ367			Annotation[a]						
No	start	end	Corresponding import in NQ1671	Description	Gene	HP-number	jhp-number	COG	OMP	Import length
84	923388	923794	-	hypothetical protein	-	0860	0794	E		407
				hypothetical protein	-	0861	0795	S		
85	924377	924379	CNP 16	hypothetical protein	-	0861	0795	S		3
86	929445	929501	-	hypothetical protein	-	0868	0802	-		57
87	936705	937184	*CNP 17 (partially)*	catalase	katA	0875	0809	P		480
88	946383	946827	-	hypothetical protein	-	0885	0817	R		445
89	947061	947064	-	cysteinyl-tRNA synthetase	cysS	0886	0818	J		4
90	959612	959874	-	outer membrane protein-adhesin	babA	1243	0833	-	hop	263
91	963928	964230	-	acetate kinase	ackA	0903	0840	C		303
92	1018022	1018691	-	hypothetical protein	-	0946	0880	C		670
				hypothetical protein	-	0947	0881	-		
93	1019663	1020471	-	hypothetical protein	-	0948	0882	-		809
				rRNA large subunit methyltransferase	-	0949	0883	S		
94	1020704	1020750	-	rRNA large subunit methyltransferase	-	0949	0883	S		47
95	1078887	1080577	-	hypothetical protein	-		0928	KL		1691
96	1097117	1097486	-	arginine decarboxylase	speA	0422	0962	E		370
97	1128133	1129204	-	hypothetical protein	-	0395	0986	R		1072
				hypothetical protein	-	0394	0987	S		
98	1131411	1133183	-	histidine kinase	cheA	0392	0989	NT		1773
				histidine kinase-MCP coupling protein	cheW	0391	0990	NT		
99	1143361	1143508	-	alpha-(1,3)-fucosyltransferase	fucU	0379	1002	-		148
				putative cytochrome C-type biogenesis protein	-	0378	1003	O		
100	1155688	1155792	-	putative type II DNA modification enzyme (methyltransferase)	-	0369	1012	-		105
101	1156094	1156244	-	putative type II DNA modification enzyme (methyltransferase)	-	0369	1012	-		151
				hypothetical protein	-	0368	1013	-		
102	1172779	1173423	-	glucokinase	glk	1103	1029	G		645
				zinc-dependent alcohol dehydrogenase	-	1104	1030	R		
103	1173626	1174328	-	zinc-dependent alcohol dehydrogenase	-	1104	1030	R		703
				putative lipopolysaccharide biosynthesis protein	-	1105	1031	M		
104	1175249	1175573	-	putative lipopolysaccharide biosynthesis protein	-		1032	M		325
105	1175829	1175989	-	putative lipopolysaccharide biosynthesis protein	-		1032	M		161
106	1184522	1185290	-	excinuclease ABC subunit B	uvrB	1114	1041	L		769

	Position in NQ367			Annotation[a]						
No	start	end	Corresponding import in NQ1671	Description	Gene	HP-number	jhp-number	COG	OMP	Import length
107	1185512	1186486	-	excinuclease ABC subunit B	uvrB	1114	1041	L		975
				hypothetical protein	-	0489	0441	-		
108	1187919	1189421	-	hypothetical protein	-	0488	0440	-		1503
				hypothetical protein	-	1117	1045	R		
109	1189928	1190997	-	gamma-glutamyltranspeptidase	ggt	1118	1046	E		1070
110	1197408	1197823	-	putative outer membrane protein	-	1125	1054	M	other	416
111	1198037	1198532	-	translocation protein TolB	tolB	1126	1055	U		496
112	1220571	1220726	-	hypothetical protein	rimM	1149	1076	J		156
113	1221616	1221743	-	putative signal recognition particle protein	ffh	1152	1079	U		128
114	1221980	1224054	-	putative signal recognition particle protein	ffh	1152	1079	U		2075
				valyl-tRNA synthetase	valS	1153	1080	J		
115	1224519	1224528	-	valyl-tRNA synthetase	valS	1153	1080	J		10
116	1225274	1225588	-	valyl-tRNA synthetase	valS	1153	1080	J		315
				flagellar assembly protein FliW	-	1154	1081	S		
117	1226381	1226382	-	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	murG	1155	1082	M		2
118	1229561	1229964	-	putative outer membrane protein	-	1157	1084	-	hop	404
119	1230879	1230887	-	putative outer membrane protein	-	1157	1084	-	hop	9
120	1231763	1234424	-	putative outer membrane protein	-	1157	1084	-	hop	2662
				pyrroline-5-carboxylate reductase	proC	1158	1085	E		
				putative cAMP-induced cell filamentation protein	fic	1159	1086	D		
121	1234716	1234826	-	putative metalloprotease	-	1160	1087	R		111
122	1240516	1240763	-	putative outer membrane protein	-	1167	1094	-	hof	248
123	1252393	1253748	-	putative outer membrane function	-	1177	1103	-	hop	1356
124	1263637	1263642	-	hypothetical protein	-	1184	1110	V		6
125	1275912	1276329	-	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		418
126	1277452	1278612	***SNP***	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		1161
127	1288692	1289961	-	ABC transporter ATP-binding protein	-	1206	1129	V		1270
128	1290230	1292705	-	ABC transporter ATP-binding protein	-	1206	1129	V		2476
				hypothetical protein	-	1207	1130	R		
				type II DNA modification enzyme (methyltransferase)	M.HpyI	1208	1131	L		
129	1293895	1293979	-	O-serine acetyltransferase	cysE	1210	1133	E		85
130	1294180	1296543	-	hypothetical protein	-	1211	1134	-		2364

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
				F0F1 ATP synthase subunit C	atpE	1212	1135	C		
				polynucleotide phosphorylase/polyadenylase	pnp	1213	1136	J		
131	1302752	1306486	-	putative ABC transporter, ATP-binding protein	-	1220	1141	V		3735
				hypothetical protein	-	1221	1142	I		
				D-lactate dehydrogenase	dld	1222	1143	C		
132	1307038	1307234	-	uroporphyrinogen-III synthase	hemD	1224	1145	H		197
133	1310170	1311087	-	dinucleoside polyphosphate hydrolase	mutT	1228	1149	LR		918
				aspartate kinase	lysC	1229	1150	E		
134	1324969	1325063	-	outer membrane protein-adhesin	babB	0896	1164	-	hop	95
135	1325416	1326846	-	outer membrane protein-adhesin	babB	0896	1164	-	hop	1431
136	1335980	1336391	-	hypothetical protein	-	1252	1173	E		412
				tryptophanyl-tRNA synthetase	trpS	1253	1174	J		
137	1360346	1360473	-	no annotation/ intergenic						128
138	1360679	1360818	-	tryptophan synthase subunit alpha	trpA	1277	1198	E		140
139	1361124	1361264	-	tryptophan synthase subunit alpha	trpA	1277	1198	E		141
140	1382301	1382304	-	50S ribosomal protein L5	rplE	1307	1227	J		4
141	1385966	1386073	-	50S ribosomal protein L2	rplB	1316	1236	J		108
142	1386322	1386610	-	50S ribosomal protein L23	rplW	1317	1237	J		289
				50S ribosomal protein L4	rplD	1318	1238	J		
143	1389012	1389099	-	hypothetical protein	-	1321	1241	R		88
144	1390389	1390394	-	ribonuclease HII	rnhB	1323	1243	L		6
145	1396455	1396945	-	cation efflux system protein	czcA_2	1329	1249	P		491
146	1404898	1405288	-	hypothetical protein	-	1337	1256	H		391
				nickel responsive regulator	-	1338	1257	K		
147	1406007	1406202	-	biopolymer transport protein	exbB_2	1339	1258	U		196
148	1406735	1407856	-	biopolymer transport EXBD protein	exbD_2	1340	1259	U		1122
				siderophore-mediated iron transport protein	tonB_2	1341	1260	M		
				putative outer membrane protein	-	0227	0212	-	hop	
149	1408174	1409258	-	putative outer membrane protein	-	0227	0212	-	hop	1085
150	1409776	1409809	-	putative outer membrane protein	-	0227	0212	-	hop	34
151	1415785	1416999	-	hypothetical protein	-	1349	1268	-		1215
				carboxyl-terminal protease	prc	1350	1269	M		
152	1417494	1417824	-	carboxyl-terminal protease	prc	1350	1269	M		331

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
153	1418038	1418277	-	no annotation/ intergenic						240
154	1418715	1420306	-	hypothetical protein	-	1351	1270	-		1592
				type II DNA modification enzyme (methyltransferase)	-	1352	1271	L		
				hypothetical protein	-	1354	1272	R		
155	1420959	1421010	-	hypothetical protein	-	1354	1272	R		52
156	1421293	1422216	-	hypothetical protein	-	1354	1272	R		924
157	1434040	1434055	CNP 19	putative transcriptional regulator	-	1365	1283	TK		16
158	1434393	1434430	CNP 20	no annotation/ intergenic						38
159	1442119	1444661	-	DNA repair protein (recombination protein N)	recN	1393	1434	L		2543
				hypothetical protein	-	1394	1433	G		
				hypothetical protein	-	1395	1432	-	hor	
160	1446949	1447717	-	L-alanine dehydrogenase	ald	1398	1428	E		769
161	1447967	1448266	-	amino acid permease	-	1017	0406	E		300
162	1448470	1448845	-	amino acid permease	-	1017	0406	E		376
163	1456339	1459549	-	type I restriction enzyme (restriction subunit)	hsdR_3	1402	1424	V		3211
				type I restriction enzyme (modification subunit)	hsdM_3	1403	1423	V		
164	1465978	1466118	-	chromosomal replication initiaton protein	dnaA	1529	1417	L		141
165	1468776	1468987	CNP 22	exodeoxyribonuclease III	exoA	1526	1415	L		212
166	1485283	1487636	-	putative iron-regulated outer membrane protein	frpB_3	1512	1405	P	family 4	2354
167	1536937	1537102	-	hypothetical protein	-	1464	1357	Q		166
168	1596264	1596339	-	no annotation/ intergenic						76
169	1601621	1602000	-	putative type III DNA modification enzyme (methyltransferase)	mod_1		1296	L		380
170	1615236	1619177	-	type III restriction enzyme R protein	-	1371	1285	-		3942
				type II DNA modification enzyme (methyltransferase)	-		1284	L		

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S11: **Clusters of nucleotide polymorphisms (CNPs)/ imports in strain pair NQ367/4191. Positions refer to the position in the virtual genome of the first isolate NQ367.**

	Position in NQ367			Annotation[a]						
No	start	end	Corresponding import in NQ1671	Description	Gene	HP-number	jhp-number	COG	OMP	Import length
1	54489	54515	-	6,7-dimethyl-8-ribityllumazine synthase	ribH	0002	0002	H		27
2	88260	89334	-	hypothetical protein	-	0035	0031	S		1075
				hypothetical protein	-	0036	0032	-		
				hypothetical protein	-	0037	0033	U		
3	92434	93107	-	DNA transformation competancy	comB3	0041/ 0042	0036	U		674
4	110917	110919	-	proline/pyrroline-5-carboxylate dehydrogenase	putA	0056	0048	C		3
5	116116	116144	-	hypothetical protein	-	0060	0055	-		29
6	117842	118200	-	hypothetical protein	-	0063	0058	-		359
7	118486	118564	-	hypothetical protein	-	0063	0058	-		79
8	121924	122022	-	hypothetical protein	-	0066	0061	D		99
9	127317	129003	-	urease subunit beta	ureB	0072	0067	E		1687
				urease subunit alpha	ureA	0073	0068	E		
				lipoprotein signal peptidase	lspA	0074	0069	MU		
				phosphoglucosamine mutase	glmM	0075	0070	G		
10	136704	138632	-	methyl-accepting chemotaxis protein (MCP)	-	0082	0075	NT		1929
11	140648	140926	-	hypothetical protein	-	0086	0079	R		279
12	154047	154439	-	methyl-accepting chemotaxis protein (MCP)	-	0099	0091	NT		393
13	161882	161932	-	2',3'-cyclic-nucleotide 2'-phosphodiesterase	cpdB	0104	0096	F		51
14	178652	178824	-	hypothetical protein	-	0118	0110	-		173
15	184565	184568	CNP 1	translation initiation factor IF-3	infC	0124	0114	J		4
16	184985	185277	CNP 2	translation initiation factor IF-3	infC	0124	0114	J		293
				50S ribosomal protein L35	rpmI	0125	0115	J		
17	185526	186814	CNP 3	50S ribosomal protein L20	rplT	0126	0116	J		1289
				putative outer membrane protein	-	0127	0117	-	hor	
18	199704	199784	SNP	L-lactate permease	lldP_1	0140	0128	C		81
19	200401	201190	*CNP 4 (partially)*	L-lactate permease	lldp_2	0141	0129	C		790
20	201646	203016	*CNP 5 (partially)*	DNA glycosylase MutY	mutY	0142	0130	L		1371

	Position in NQ367			Annotation[a]						
No	start	end	Corresponding import in NQ1671	Description	Gene	HP-number	jhp-number	COG	OMP	Import length
				hypothetical protein	-	0143	0131	P		
21	203223	203470	*CNP 6 (partially)*	hypothetical protein	-	0143	0131	P		248
22	207219	207815	CNP 7	cytochrome oxidase (CBB3-TYPE)	fixP	0147	0135	C		597
				hypothetical protein	-	0148	0136	-		
				hypothetical protein	-	0149	0137	-		
23	216807	217075	*CNP 8 (partially)*	hypothetical protein	-	0160	0148	R		269
24	239750	240212	-	apolipoprotein N-acyltransferase	lnt	0180	0168	M		463
25	263546	263696	-	hypothetical protein	-	0205	0191	E		151
26	267351	267854	-	putative outer membrane protein	-	0209	0195	-	hof	504
27	279747	281472	-	hypothetical protein	-	0219	0205	-		1726
				cysteine desulfurase	-	0220	0206	E		
28	283552	284660	-	DNA repair protein RadA	radA	0223	0209	O		1109
				bifunctional methionine sulfoxide reductase A/B protein	-	0224	0210	O		
29	287257	287294	-	putative outer membrane protein	-	0227	0212	-	hop	38
30	287300	288386	-	putative outer membrane protein	-	0227	0212	-	hop	1087
31	291746	292044	-	3-deoxy-manno-oculosonate cytidylyltransferase	kdsB	0230	0215	M		299
32	293524	294489	-	hypothetical protein	-	0232	0217	-		966
				hypothetical protein	-	0233	0218	E		
33	296879	297905	-	hypothetical protein	-	0235	0220	R		1027
				hypothetical protein	-	0236	0221	-		
34	320772	322081	-	exodeoxyribonuclease VII large subunit	xseA	0259	0243	L		1310
				type II DNA modification enzyme (methyltransferase)	-	0260	0244	L		
				hypothetical protein	-	0261	0245	-		
35	323818	324442	-	type II DNA modification enzyme (methyltransferase)	-	0263	0248	L		625
				heat shock protein	clpB	0264	0249	O		
36	345358	345391	-	hypothetical protein	-	0284	0269	M		34
37	355391	357551	***SNP***	putative vacuolating cytotoxin (VacA) paralog	-	0289	0274	N		2161
38	368110	368533	-	periplasmic dipeptide transport substrate-binding protein	dppA	0298	0283	E		424
39	397362	397364	-	cell division inhibitor	minD	0331	0314	D		3
40	458034	460119	-	translation initiation factor IF-2	infB	1048	0377	J		2086
41	460375	460526	-	ribosome binding factor A	rbfA	1047	0378	J		152
42	496825	497803	-	polyphosphate kinase	ppk	1010	0413	P		979
43	498444	498560	-	no annotation/ intergenic						117

No	Position in NQ367		Corresponding import in NQ1671	Annotation[a]				COG	OMP	Import length
	start	end		Description	Gene	HP-number	jhp-number			
44	498927	499856	-	type I restriction enzyme S protein (hsdS)	-	0462		V		930
				type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		
45	500072	500082	-	type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		11
46	500829	501361	-	type I restriction enzyme (modification subunit)	hsdM_1	0463	0415	V		533
				fragment of type I restriction enzyme R protein (hsdR)	-	0464	0416	V		
47	527296	527692	-	outer membrane protein	-	0486	0438	-	hof	397
48	541655	542116	-	putative phospholipase A1	pldA	0499	0451	M		462
49	542790	542797	-	DNA polymerase III subunit beta	dnaN	0500	0452	L		8
50	550896	551456	-	hypothetical protein	-	0508	0458	-		561
				putative glycolate oxidase	glcD	0509	0459	C		
51	576609	577181	-	cag island protein	orf15	0528	0477	U		573
52	577417	577428	-	cag island protein	orf16	0529	0478	-		12
53	597023	597257	*CNP 10 (partially)*	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		235
54	598485	598614	-	cag island protein, cytotoxicity associated immunodominant antigen	cagA	0547	0495	-		130
55	608002	608770	-	acetyl-coenzyme A carboxylase subunit alpha	accA	0557	0504	I		769
				3-oxoacyl-(acyl carrier protein) synthase II	fabB	0558	0505	IQ		
56	665272	665368	-	putative vacuolating cytotoxin (VacA) paralog	-	0609/ 0610	0556	N		97
57	690907	692173	-	hydrogenase, small subunit	hyaA	0631	0574	C		1267
				hydrogenase, large subunit	hyaB	0632	0575	C		
58	696460	696612	-	hypothetical protein	-	0636	0579	-		153
59	736230	736459	-	integrase-recombinase protein (XERCD family)	-	0675	0617	L		230
60	736667	737772	-	integrase-recombinase protein (XERCD family)	-	0675	0617	L		1106
				methylated-DNA-protein-cysteine methyltransferase	ogt	0676	0618	L		
61	738035	739130	-	hypothetical protein	-	0677	0619	R		1096
				putative lipopolysaccharide biosynthesis protein	-	0679	0620	R		
62	739344	739594	-	putative lipopolysaccharide biosynthesis protein	-	0679	0620	R		251
63	739836	741206	-	ribonucleotide-diphosphate reductase subunit alpha	nrdA	0680	0621	F		1371
64	741821	741977	-	ribonucleotide-diphosphate reductase subunit alpha	nrdA	0680	0621	F		157
65	742558	742624	-	hypothetical protein	-	0681	0622	-		67
66	787373	788028	SNP	putative outer membrane protein	-	0725	0662	-	hop	656

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
67	788647	789174	-	putative outer membrane protein	-	0725	0662	-	hop	528
68	792285	794068	-	putative outer membrane protein	-	0722	0659	-	hop	1784
				putative outer membrane protein	-	0726	0663	-	other	
69	828238	828245	-	phosphodiesterase	-	0760	0697	R		8
				hypothetical protein	-	0761	0698	H		
70	849294	850422	-	hypothetical protein	-	0781	0718	L		1129
				putative outer membrane protein	-	0782	0719	-	hof	
71	870771	871722	-	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein	ribBA	0804	0740	H		952
72	872477	874100	-	putative lipopolysaccharide biosynthesis protein	-	0805	0741	M		1624
				hypothetical protein	-	0806	0742	R		
				iron(III) dicitrate transport protein	fecA_2	0807	0743	P	family 4	
73	880037	880209	-	molybdopterin-synthase sulfurylase	moeB	0814	0750	H		173
74	880733	881135	-	molybdopterin-synthase sulfurylase	moeB	0814	0750	H		403
				flagellar motor protein MotA	motA	0815	0751	N		
75	886549	888062	-	excinuclease ABC subunit C	uvrC	0821	0760	L		1514
				homoserine dehydrogenase	hom	0822	0761	E		
76	893161	894358	-	F0F1 ATP synthase subunit A	atpB	0828	0767	C		1198
				inosine-5'-monophosphate dehydrogenase	guaB	0829	0768	F		
77	895074	895082	-	aspartyl/glutamyl-tRNA amidotransferase subunit A	gatA	0830	0769	J		9
78	900054	902239	-	GTP-binding protein EngA	engA	0834	0773	R		2186
				DNA-binding protein Hu	-	0835	0774	L		
				hypothetical protein	-	0836/ 0837	0775	S		
				hypothetical protein	-	0838	0776	-		
79	910113	910373	-	type I restriction enzyme (restriction subunit)	hsdR_2	0846	0784	V		261
80	911288	913314	-	type I restriction enzyme (restriction subunit)	hsdR_2	0846	0784	V		2027
				anti-codon nuclease masking agent (prrB)	-	0790		V		
81	919474	921004	-	guanosine 5'-monophosphate oxidoreductase	guaC	0854	0790	F		1531
				phosphoheptose isomerase	gmhA	0857	0791	G		
				putative ADP-D-glycero-D-mannoheptose synthase"	waaE	0858	0792	M		
82	921214	922287	-	putative ADP-D-glycero-D-mannoheptose synthase	waaE	0858	0792	M		1074
				ADP-L-glycero-D-mannoheptose-6-epimerase	gmhD	0859	0793	MG		
83	923004	923007	-	ADP-L-glycero-D-mannoheptose-6-epimerase	gmhD	0859	0793	MG		4

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
84	923388	923794	-	hypothetical protein	-	0860	0794	E		407
				hypothetical protein	-	0861	0795	S		
85	924377	924379	CNP 16	hypothetical protein	-	0861	0795	S		3
86	929445	929501	-	hypothetical protein	-	0868	0802	-		57
87	936705	937184	*CNP 17 (partially)*	catalase	katA	0875	0809	P		480
88	946383	946827	-	hypothetical protein	-	0885	0817	R		445
89	947061	947064	-	cysteinyl-tRNA synthetase	cysS	0886	0818	J		4
90	959612	959874	-	outer membrane protein-adhesin	babA	1243	0833	-	hop	263
91	963928	964230	-	acetate kinase	ackA	0903	0840	C		303
92	1018022	1018691	-	hypothetical protein	-	0946	0880	C		670
				hypothetical protein	-	0947	0881	-		
93	1019663	1020471	-	hypothetical protein	-	0948	0882	-		809
				rRNA large subunit methyltransferase	-	0949	0883	S		
94	1020704	1020750	-	rRNA large subunit methyltransferase	-	0949	0883	S		47
95	1078887	1080577	-	hypothetical protein	-		0928	KL		1691
96	1097117	1097486	-	arginine decarboxylase	speA	0422	0962	E		370
97	1128133	1129204	-	hypothetical protein	-	0395	0986	R		1072
				hypothetical protein	-	0394	0987	S		
98	1131411	1133183	-	histidine kinase	cheA	0392	0989	NT		1773
				histidine kinase-MCP coupling protein	cheW	0391	0990	NT		
99	1143361	1143508	-	alpha-(1,3)-fucosyltransferase	fucU	0379	1002	-		148
				putative cytochrome C-type biogenesis protein	-	0378	1003	O		
100	1155688	1155792	-	putative type II DNA modification enzyme (methyltransferase)	-	0369	1012	-		105
101	1156094	1156244	-	putative type II DNA modification enzyme (methyltransferase)	-	0369	1012	-		151
				hypothetical protein	-	0368	1013	-		
102	1172779	1173423	-	glucokinase	glk	1103	1029	G		645
				zinc-dependent alcohol dehydrogenase	-	1104	1030	R		
103	1173626	1174328	-	zinc-dependent alcohol dehydrogenase	-	1104	1030	R		703
				putative lipopolysaccharide biosynthesis protein	-	1105	1031	M		
104	1175249	1175573	-	putative lipopolysaccharide biosynthesis protein	-		1032	M		325
105	1175829	1175989	-	putative lipopolysaccharide biosynthesis protein	-		1032	M		161
106	1184522	1185290	-	excinuclease ABC subunit B	uvrB	1114	1041	L		769

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
107	1185512	1186486	-	excinuclease ABC subunit B	uvrB	1114	1041	L		975
				hypothetical protein	-	0489	0441	-		
108	1187919	1189421	-	hypothetical protein	-	0488	0440	-		1503
				hypothetical protein	-	1117	1045	R		
109	1189928	1190997	-	gamma-glutamyltranspeptidase	ggt	1118	1046	E		1070
110	1197408	1197823	-	putative outer membrane protein	-	1125	1054	M	other	416
111	1198037	1198532	-	translocation protein TolB	tolB	1126	1055	U		496
112	1220571	1220726	-	hypothetical protein	rimM	1149	1076	J		156
113	1221616	1221743	-	putative signal recognition particle protein	ffh	1152	1079	U		128
114	1221980	1224054	-	putative signal recognition particle protein	ffh	1152	1079	U		2075
				valyl-tRNA synthetase	valS	1153	1080	J		
115	1224519	1224528	-	valyl-tRNA synthetase	valS	1153	1080	J		10
116	1225274	1225588	-	valyl-tRNA synthetase	valS	1153	1080	J		315
				flagellar assembly protein FliW	-	1154	1081	S		
117	1226381	1226382	-	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	murG	1155	1082	M		2
118	1229561	1229964	-	putative outer membrane protein	-	1157	1084	-	hop	404
119	1230879	1230887	-	putative outer membrane protein	-	1157	1084	-	hop	9
120	1231763	1234424	-	putative outer membrane protein	-	1157	1084	-	hop	2662
				pyrroline-5-carboxylate reductase	proC	1158	1085	E		
				putative cAMP-induced cell filamentation protein	fic	1159	1086	D		
121	1234716	1234826	-	putative metalloprotease	-	1160	1087	R		111
122	1240516	1240763	-	putative outer membrane protein	-	1167	1094	-	hof	248
123	1252393	1253748	-	putative outer membrane function	-	1177	1103	-	hop	1356
124	1263637	1263642	-	hypothetical protein	-	1184	1110	V		6
125	1275912	1276329	-	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		418
126	1277452	1278612	***SNP***	DNA-directed RNA polymerase subunit beta/beta'	rpoB	1198	1121	K		1161
127	1288692	1289961	-	ABC transporter ATP-binding protein	-	1206	1129	V		1270
128	1290230	1292705	-	ABC transporter ATP-binding protein	-	1206	1129	V		2476
				hypothetical protein	-	1207	1130	R		
				type II DNA modification enzyme (methyltransferase)	M.HpyI	1208	1131	L		
129	1293895	1293979	-	O-serine acetyltransferase	cysE	1210	1133	E		85
130	1294180	1296543	-	hypothetical protein	-	1211	1134	-		2364

	Position in NQ367			Annotation[a]						
No	start	end	Corresponding import in NQ1671	Description	Gene	HP-number	jhp-number	COG	OMP	Import length
				F0F1 ATP synthase subunit C	atpE	1212	1135	C		
				polynucleotide phosphorylase/polyadenylase	pnp	1213	1136	J		
131	1302752	1306486	-	putative ABC transporter, ATP-binding protein	-	1220	1141	V		3735
				hypothetical protein	-	1221	1142	I		
				D-lactate dehydrogenase	dld	1222	1143	C		
132	1307038	1307234	-	uroporphyrinogen-III synthase	hemD	1224	1145	H		197
133	1310170	1311087	-	dinucleoside polyphosphate hydrolase	mutT	1228	1149	LR		918
				aspartate kinase	lysC	1229	1150	E		
134	1324969	1325063	-	outer membrane protein-adhesin	babB	0896	1164	-	hop	95
135	1325416	1326846	-	outer membrane protein-adhesin	babB	0896	1164	-	hop	1431
136	1335980	1336391	-	hypothetical protein	-	1252	1173	E		412
				tryptophanyl-tRNA synthetase	trpS	1253	1174	J		
137	1360346	1360473	-	no annotation/ intergenic						128
138	1360679	1360818	-	tryptophan synthase subunit alpha	trpA	1277	1198	E		140
139	1361124	1361264	-	tryptophan synthase subunit alpha	trpA	1277	1198	E		141
140	1382301	1382304	-	50S ribosomal protein L5	rplE	1307	1227	J		4
141	1385966	1386073	-	50S ribosomal protein L2	rplB	1316	1236	J		108
142	1386322	1386610	-	50S ribosomal protein L23	rplW	1317	1237	J		289
				50S ribosomal protein L4	rplD	1318	1238	J		
143	1389012	1389099	-	hypothetical protein	-	1321	1241	R		88
144	1390389	1390394	-	ribonuclease HII	rnhB	1323	1243	L		6
145	1396455	1396945	-	cation efflux system protein	czcA_2	1329	1249	P		491
146	1404898	1405288	-	hypothetical protein	-	1337	1256	H		391
				nickel responsive regulator	-	1338	1257	K		
147	1406007	1406202	-	biopolymer transport protein	exbB_2	1339	1258	U		196
148	1406735	1407856	-	biopolymer transport EXBD protein	exbD_2	1340	1259	U		1122
				siderophore-mediated iron transport protein	tonB_2	1341	1260	M		
				putative outer membrane protein	-	0227	0212	-	hop	
149	1408174	1409258	-	putative outer membrane protein	-	0227	0212	-	hop	1085
150	1409776	1409809	-	putative outer membrane protein	-	0227	0212	-	hop	34
151	1415785	1416999	-	hypothetical protein	-	1349	1268	-		1215
				carboxyl-terminal protease	prc	1350	1269	M		
152	1417494	1417824	-	carboxyl-terminal protease	prc	1350	1269	M		331

	Position in NQ367			Annotation[a]						
No	**start**	**end**	**Corresponding import in NQ1671**	**Description**	**Gene**	**HP-number**	**jhp-number**	**COG**	**OMP**	**Import length**
153	1418038	1418277	-	no annotation/ intergenic						240
154	1418715	1420306	-	hypothetical protein	-	1351	1270	-		1592
				type II DNA modification enzyme (methyltransferase)	-	1352	1271	L		
				hypothetical protein	-	1354	1272	R		
155	1420959	1421010	-	hypothetical protein	-	1354	1272	R		52
156	1421293	1422216	-	hypothetical protein	-	1354	1272	R		924
157	1434040	1434055	CNP 19	putative transcriptional regulator	-	1365	1283	TK		16
158	1434393	1434430	CNP 20	no annotation/ intergenic						38
159	1442119	1444661	-	DNA repair protein (recombination protein N)	recN	1393	1434	L		2543
				hypothetical protein	-	1394	1433	G		
				hypothetical protein	-	1395	1432	-	hor	
160	1446949	1447717	-	L-alanine dehydrogenase	ald	1398	1428	E		769
161	1447967	1448266	-	amino acid permease	-	1017	0406	E		300
162	1448470	1448845	-	amino acid permease	-	1017	0406	E		376
163	1456339	1459549	-	type I restriction enzyme (restriction subunit)	hsdR_3	1402	1424	V		3211
				type I restriction enzyme (modification subunit)	hsdM_3	1403	1423	V		
164	1465978	1466118	-	chromosomal replication initiaton protein	dnaA	1529	1417	L		141
165	1468776	1468987	CNP 22	exodeoxyribonuclease III	exoA	1526	1415	L		212
166	1485283	1487636	-	putative iron-regulated outer membrane protein	frpB_3	1512	1405	P	family 4	2354
167	1536937	1537102	-	hypothetical protein	-	1464	1357	Q		166
168	1596264	1596339	-	no annotation/ intergenic						76
169	1601621	1602000	-	putative type III DNA modification enzyme (methyltransferase)	mod_1		1296	L		380
170	1615236	1619177	-	type III restriction enzyme R protein	-	1371	1285	-		3942
				type II DNA modification enzyme (methyltransferase)	-		1284	L		

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches. COG classifications are primarily based on J99 as of May 2010.

Table S12: Mean estimates and 95% credibility intervals from the Bayesian analyses of the paired *H. pylori* isolates obtained sequentially during chronic infection.

	NQ315/ NQ1712	NQ352/ NQ1701	NQ392/ NQ1707	NQ392/ NQ4060	NQ367/ NQ1671	NQ367/ NQ4191
Length of imports (bp)	402 [327;496]	629 [571;699]	283 [150;468]	286 [207;398]	261 [176;375]	519 [440;610]
Length of groups of imports (bp)	3362 [2561;4291]	3686 [3179;4250]	7922 [1977;18176]	10088 [4772;20469]	2599 [1460;4663]	3359 [2556;4359]
Number of imports per group	2.41 [2.27;2.58]	2.45 [2.35;2.55]	2.6 [2.1;3.6]	3.1 [2.6;4.0]	2.54 [2.23;2.97]	2.69 [2.51;2.91]
Distance between two imports in a group (bp)	700 [504;916]	621 [504;770]	1933 [391;4289]	2207 [1045;4217]	542 [267;1020]	531 [367;761]
Distance between two groups of imports (bp)	19438 [14374;27048]	3157 [2598;3801]	134916 [47060;293286]	30973 [15633;49392]	72170 [40906;12480]	8469 [6449;10646]
Mutation rate [a]	2.4 [1.9;2.9]	6.5 [5.3;7.5]	0.5 [0.3;0.7]	0.5 [0.4;0.6]	0.7 [0.5;1.0]	0.7 [0.6;0.8]
Import rate [b]	4.0 [3.0;5.8]	16.5 [14.0;19.0]	0.5 [0.2;1.2]	0.2 [0.1;0.3]	1.0 [0.4;1.8]	1.0 [0.8;1.1]
Ratio of rates of imports and mutations	1.7 [1.1;2.6]	2.5 [2.0;3.2]	1.2 [0.4;3.0]	0.4 [0.2;0.6]	1.4 [0.5;2.7]	1.5 [1.1;2.0]
Ratio of effects of imports and mutations	26.7 [15.9;41.0]	62.0 [47.5;78.0]	12.4 [3.5;35.0]	4.3 [2.2;7.4]	13.7 [5.2;28.3]	29.5 [20.7;39.6]

[a] (per site, per year, x10^{-5})
[b] (per initiation site, per year, x10^{-5})

Table S13: Frequency of imports /CNPs within different functional gene categories

		No. of genes in J99	combined length of genes in category (kb)	No. of genes affected in NQ315/ NQ1712	% of genes affected by imports in NQ315/ NQ1712	Fisher's exact test P-value	No. of genes affected in NQ352/ NQ1701	% of genes affected by imports in NQ352/ NQ1701	Fisher's exact test P-value
Information storage and processing		241	262.864	17	7.05	0.958	100	41.49	0.544
Translation, ribosomal structure and biogenesis	J	121	102.705	6	4.96	0.985	47	38.84	0.770
Transcription	K	35	47.313	5	14.29	0.255	15	42.86	0.506
Replication, recombination and repair	L	85	112.846	6	7.06	0.858	38	44.71	0.314
Cellular processes and signalling		344	417.178	39	11.34	0.166	161	46.80	***0.017***
Cell cycle control, cell divison, cnromosome partitioning	D	23	24.042	4	17.39	0.183	7	30.43	0.907
Defense mechanisms	V	25	41.565	3	12.00	0.451	12	48.00	0.324
Signal transduction mechanisms	T	27	30.753	3	11.11	0.503	15	55.56	0.100
Cell wall/ membrane/ envelope biogenesis	M	97	111.006	15	15.46	0.046	47	48.45	0.097
Cell motility	N	49	76.56	3	6.12	0.875	29	59.18	**0.009**
Intracellular Traffiking, secretion, and vesicular transport	U	59	66.967	2	3.39	0.985	16	27.12	0.994
Posttranslational modification, protein turnover, chaperones	O	64	66.285	9	14.06	0.168	35	54.69	***0.022***
Metabolism		414	397.179	34	8.21	0.915	192	46.38	***0.014***
Energy production and conversion	C	69	74.643	2	2.90	0.994	39	56.52	**0.008**
Carbohydrate transport and metabolism	G	39	41.499	5	12.82	0.335	15	38.46	0.713
Amino acid transport and metabolism	E	89	102.735	5	5.62	0.949	42	47.19	0.162
Nucleotide transport and metabolism	F	38	38.820	2	5.26	0.901	18	47.37	0.286
Coenzyme transport and metabolism	H	75	65.046	7	9.33	0.616	32	42.67	0.470
Lipid transport and metabolism	I	40	34.686	2	5.00	0.917	18	45.00	0.388
Inorganic ion transport and metabolism	P	53	67.458	10	18.87	0.029	27	50.94	0.105
Secondary metabolites biosynthesis, transport and catabolism	Q	11	11.112	1	9.09	0.680	1	9.09	0.997
Poorly characterized		187	171.912	14	7.49	0.901	63	33.69	0.993
General function prediction only	R	117	112.485	9	7.69	0.831	35	29.91	0.998
Function unknown	S	70	59.427	5	7.14	0.835	28	40.00	0.655
No category	-	392	302.798	37	9.44	0.644	140	35.71	0.997
Outer membrane proteins		63	95.283	20	31.75	**0.000**	27	42.86	0.468
family 1 (hop and hor)		32	44.949	11	34.38	**0.000**	17	53.13	0.125
hop		20	35.628	10	50.00	**0.000**	13	65.00	***0.029***
hor		12	9.321	1	8.33	0.712	4	33.33	0.808
hof - family 2		8	11.571	4	50.00	**0.005**	4	50.00	0.444
hom - family 3		4	8.472	0	0.00	1.000	2	50.00	0.553
family 4		6	14.673	2	33.33	0.110	2	33.33	0.792
family 5		3	3.969	1	33.33	0.267	1	33.33	0.801
other		10	11.649	2	20.00	0.256	1	10.00	0.995

		No. of genes in J99	combined length of genes in category (bp)	No. of genes affected in NQ392/ NQ1707	% of genes affected by imports in NQ392/ NQ1707	Fisher's exact test P-value	No. of genes affected in NQ392/ NQ4060	% of genes affected by imports in NQ392/ NQ4060	Fisher's exact test P-value
Information storage and processing		241	262.864	0	0.00	1.000	5	2.07	0.880
Translation, ribosomal structure and biogenesis	J	121	102.705	0	0.00	1.000	3	2.48	0.728
Transcription	K	35	47.313	0	0.00	1.000	2	5.71	0.289
Replication, recombination and repair	L	85	112.846	0	0.00	1.000	0	0.00	1.000
Cellular processes and signalling		344	417.178	5	1.45	0.275	11	3.20	0.484
Cell cycle control, cell divison, cnromosome partitioning	D	23	24.042	0	0.00	1.000	0	0.00	1.000
Defense mechanisms	V	25	41.565	2	8.00	***0.026***	5	20.00	**0.001**
Signal transduction mechanisms	T	27	30.753	0	0.00	1.000	1	3.70	0.569
Cell wall/ membrane/ envelope biogenesis	M	97	111.006	2	2.06	0.266	3	3.09	0.577
Cell motility	N	49	76.560	0	0.00	1.000	1	2.04	0.785
Intracellular Traffiking, secretion, and vesicular transport	U	59	66.967	1	1.69	0.465	0	0.00	1.000
Posttranslational modification, protein turnover, chaperones	O	64	66.285	0	0.00	1.000	1	1.56	0.867
Metabolism		414	397.179	4	0.97	0.656	7	1.69	0.983
Energy production and conversion	C	69	74.643	1	1.45	0.520	0	0.00	1.000
Carbohydrate transport and metabolism	G	39	41.499	0	0.00	1.000	0	0.00	1.000
Amino acid transport and metabolism	E	89	102.735	1	1.12	0.614	3	3.37	0.517
Nucleotide transport and metabolism	F	38	38.820	0	0.00	1.000	0	0.00	1.000
Coenzyme transport and metabolism	H	75	65.046	0	0.00	1.000	0	0.00	1.000
Lipid transport and metabolism	I	40	34.686	0	0.00	1.000	1	2.50	0.714
Inorganic ion transport and metabolism	P	53	67.458	2	3.77	0.102	3	5.66	0.217
Secondary metabolites biosynthesis, transport and catabolism	Q	11	11.112	0	0.00	1.000	0	0.00	1.000
Poorly characterized		187	171.912	1	0.53	0.874	5	2.67	0.691
General function prediction only	R	117	112.485	1	0.85	0.717	3	2.56	0.706
Function unknown	S	70	59.427	0	0.00	1.000	2	2.86	0.639
No category	-	392	302.798	4	1.02	0.607	13	3.32	0.415
Outer membrane proteins		63	95.283	3	4.76	***0.025***	9	14.29	**0.000**
family 1 (hop and hor)		32	44.949	2	6.25	***0.042***	7	21.88	**0.000**
hop		20	35.628	2	10.00	***0.017***	7	35.00	**0.000**
hor		12	9.321	0	0.00	1.000	0	0.00	1.000
hof - family 2		8	11.571	0	0.00	1.000	0	0.00	1.000
hom - family 3		4	8.472	0	0.00	1.000	0	0.00	1.000
family 4		6	14.673	1	16.67	0.061	1	16.67	0.170
family 5		3	3.969	0	0.00	1.000	0	0.00	1.000
other		10	11.649	0	0.00	1.000	1	10.00	0.267

		No. of genes in J99	combined length of genes in category (bp)	No. of genes affected in NQ367/ NQ1671	% of genes affected by imports in NQ367/ NQ1671	Fisher's exact test P-value	No. of genes affected in NQ367/ NQ4191	% of genes affected by imports in NQ367/ NQ4191	Fisher's exact test P-value
Information storage and processing		241	262.864	10	4.15	***0.020***	40	16.60	0.114
Translation, ribosomal structure and biogenesis	J	121	102.705	3	2.48	0.464	15	12.40	0.730
Transcription	K	35	47.313	2	5.71	0.162	4	11.43	0.739
Replication, recombination and repair	L	85	112.846	5	5.88	***0.028***	21	24.71	***0.004***
Cellular processes and signalling		344	417.178	4	1.16	0.948	51	14.83	0.314
Cell cycle control, cell divison, cnromosome partitioning	D	23	24.042	0	0.00	1.000	3	13.04	0.640
Defense mechanisms	V	25	41.565	1	4.00	0.410	10	40.00	***0.001***
Signal transduction mechanisms	T	27	30.753	1	3.70	0.434	5	18.52	0.318
Cell wall/ membrane/ envelope biogenesis	M	97	111.006	1	1.03	0.877	14	14.43	0.484
Cell motility	N	49	76.560	0	0.00	1.000	7	14.29	0.532
Intracellular Traffiking, secretion, and vesicular transport	U	59	66.967	1	1.69	0.716	8	13.56	0.589
Posttranslational modification, protein turnover, chaperones	O	64	66.285	0	0.00	1.000	4	6.25	0.985
Metabolism		414	397.179	6	1.45	0.894	55	13.29	0.688
Energy production and conversion	C	69	74.643	2	2.90	0.423	12	17.39	0.242
Carbohydrate transport and metabolism	G	39	41.499	0	0.00	1.000	5	12.82	0.649
Amino acid transport and metabolism	E	89	102.735	1	1.12	0.853	17	19.10	0.100
Nucleotide transport and metabolism	F	38	38.820	0	0.00	1.000	4	10.53	0.796
Coenzyme transport and metabolism	H	75	65.046	1	1.33	0.800	6	8.00	0.962
Lipid transport and metabolism	I	40	34.686	0	0.00	1.000	3	7.50	0.932
Inorganic ion transport and metabolism	P	53	67.458	2	3.77	0.301	6	11.32	0.768
Secondary metabolites biosynthesis, transport and catabolism	Q	11	11.112	0	0.00	1.000	2	18.18	0.465
Poorly characterized		187	171.912	2	1.07	0.914	23	12.30	0.780
General function prediction only	R	117	112.485	1	0.85	0.921	17	14.53	0.462
Function unknown	S	70	59.427	1	1.43	0.776	6	8.57	0.940
No category	-	392	302.798	8	2.04	0.587	42	10.71	0.987
Outer membrane proteins		63	95.283	4	6.35	***0.039***	17	26.98	***0.004***
family 1 (hop and hor)		32	44.949	3	9.38	***0.027***	9	28.13	***0.025***
hop		20	35.628	2	10.00	0.063	7	35.00	***0.014***
hor		12	9.321	1	8.33	0.223	2	16.67	0.513
hof - family 2		8	11.571	0	0.00	1.000	4	50.00	***0.016***
hom - family 3		4	8.472	0	0.00	1.000	0	0.00	1.000
family 4		6	14.673	0	0.00	1.000	2	33.33	0.198
family 5		3	3.969	1	33.33	0.061	0	0.00	1.000
other		10	11.649	0	0.00	1.000	2	20.00	0.415

Table S14: Positions with intra-isolate sequence heterogeneity in NQ392 and NQ1707 (see Materials and Methods for details)

A Heterogeneity in NQ392

	Position in NQ392					Annotation[a]			
No	**start**	**end**	**Contigs[b]**	**Number of polymorphic sites**	**Variant present in other isolates[c]**	**Description**	**gene**	**HP-number**	**jhp-number**
1	9468	9640	*68* and 73	6	no	putative		1524	1413
2	431720	432681	*63* and 78	106	partially, SNP_X_1, X_2 and X_3 (3/106)	outer membrane protein-adhesin	babA	1243	0833
3	847766	847913	*32* and 22	9	CNP_X_1 (9/9)	putative outer membrane protein		0726	0663
4	913066	913150	*57* and 54	2	CNP_X_2 (2/2)	type I restriction enzyme (specificity subunit)	hsdS_4		0726
5	913183	913273	*57* and 54	5	corresponding variability in NQ1707	type I restriction enzyme (specificity subunit)	hsdS_4		0726
6	913281	913388	*47* and 77	14	partially, CNP_X_3 (10/14)	type I restriction enzyme (specificity subunit)	hsdS_4		0726
7	913781	913936	*4* and 5	8	CNP_X_4 (8/8)	anti-codon nuclease masking agent (prrB)		0790	
8	933182	933632	*75* and 70	36	no	putative		0812	0748
9	967799	968370	*74* and 60	46	no	type I restriction enzyme (modification subunit)	hsdM_2	0850	0786
10	1246422	1246696	*53* and 35	34	CNP_X_5 (34/34)	intergenic			
11	1650240	1651165	*49* and 14	71	partially, CNP_X_6 (6/71)	putative type II DNA modification enzyme (methyltransferase)			1409

B Heterogeneity in NQ1707

No	Position in NQ392		Contigs[b]	Number of polymorphic sites	Variant present in other isolates[c]	Annotation[a]			
	start	end				Description	gene	HP-number	jhp-number
1	99583	99956	58 and 57	23	no	carbonic anhydrase	icfA	0004	0004
2	266915	267129	*34* and 86	21	no	putative		0168	0154
3	335622	335904	*51* and 52	16	no	3-deoxy-manno-octulosonate cytidylyltransferase	kdsB	0230	0215
4	721820	722174	*32* and 69	11	no	DNA-ligase	lig	0615	0558
5	913183	913273	*95* and 66	6	corresponding variability in NQ392	typeI restriction enzyme (specificity subunit)	hsdS_4		0726
6	913450	913552	*19* and 66	5	partially, SNP_X_4 (1/5)	anti-codon nuclease masking agent (prrB)		0790	
7	974494	974653	*76* and 90	16	no	phosphoheptose isomerase	gmhA	0857	0791
8	1621733	1621835	*84* and 21	12	no	putative		1490	1383
9	102594	102984	*50* and 100	31	no	putative outer membrane protein		0009	0007
10	88600	89056	*102* and 65	21	no	putative undecaprenyl-phosphate-alpha-N-acetylglucosaminyltransferaseE	wecA	1581	1488
11	87542	88742	*102* and 94	72	no	putative		1579	1486
						putative		1580	1487
						putative undecaprenyl-phosphate-alpha-N-acetylglucosaminyltransferaseE	wecA	1581	1488
12	in NQ1707 622412	in NQ1707 622526	*33* and 74	5	no	cag island protein	orf13/14	0527	0476
13	in NQ1707 622139	in NQ1707 622250	*33* and 72	4	no	cag island protein	orf13/14	0527	0476

C Variability with variants present in other isolates

No	Position in NQ392 start	Position in NQ392 end	Variant present in isolate: NQ392	NQ1707	NQ4060	Annotation[a] Description	gene	HP-number	jhp-number
SNP_X_1	431720		A/B	A	B	outer membrane protein-adhesin	babA	1243	0833
SNP_X_2	432168		A/B	A	B	outer membrane protein-adhesin	babA	1243	0833
SNP_X_3	432460		A/B	A	B	outer membrane protein-adhesin	babA	1243	0833
CNP_X_1	847766	847913	A/B	B	A	putative outer membrane protein		0726	0663
CNP_X_2	913066	913150	A/B	B	A	type I restriction enzyme (specificity subunit)	hsdS_4		0726
CNP_X_3	913281	913305	A/B	A	B	type I restriction enzyme (specificity subunit)	hsdS_4		0726
SNP_X_4	913450		A	A/B	B	anti-codon nuclease masking agent (prrB)		0790	
CNP_X_4	913781	913936	A/B	B	A	anti-codon nuclease masking agent (prrB)		0790	
CNP_X_5	1246422	1246696	A/B	A	B	intergenic			
CNP_X_6	1650912	1651162	A/B	B	A	putative type II DNA modification enzyme (methyltransferase)			1409

[a] Annotation is based on an automatic annotation, generated with the Kodon software using the J99 annotation as a reference. The annotation of predicted ORFs with no homologue in J99 was done manually by means of BLAST searches.

[b] Italics indicate contigs used for the virtual genomes

[c] Correspondence of heterogeneity in other isolates is described in more detail in part C

Table S15 a: Differences in strain pair NQ315/1712 not categorized as SNPs or CNPs

No	Position in NQ315		type of difference	sequence	gene affected by difference					
	start	end					HP-number	jhp-number	COG	comment
1	232464	232479	difference in 2nt repeat 8 X in NQ315 9 X in NQ1712	`ct`	putative lipopolysaccharide biosynthesis protein	rfaJ	0208	0194		phase variation NQ315 ON NQ1712 OFF
2	456503	456619	117nt indel in NQ1712	`aaaatagccc cacacaaagg ggcgaaaaaa tagcccaaca agcatttttt aatgaattgc aactctttgt tttggtggat gtggataatg tttttcaaaa aaaatcttta tttaaaa`	putative		1015	0408	-	
					putative short chain dehydrogenase		1014	0409	IQR	
3	526346	561870	loss of cagPAI							
4	955124	955126	2 additional contigs in NQ1712		outer membrane protein - adhesin	babA	1243	0833	-	
5	1022600	1022656	difference in 19nt repeat 3 X in NQ315 4 X in NQ1712	`aaaatttaga gtttgaaag`	no annotation/ intergenic					
6	1317764	1317827	difference in 8nt repeat 8 X in NQ315 10 X in NQ1712	`ttatgtat`	no annotation/ intergenic					
7	1394038	1394057	difference in 2nt repeat 10 X in NQ315 9 X in NQ1712	`ga`	hypothetical protein		1417	1312		
8	1412552	1412563	difference in 6nt repeat 2 X in NQ315 4 X in NQ1712	`aagtct`	no annotation/ intergenic					
9	1414357	1414359	2 additional contigs in NQ1712	`position in NQ1712V: 1568115 – 1570317`	(BLAST: hypothetical protein HP9810_898g1[H. pylori 98-10])					
10	1460004	1460021	difference in 9nt repeat 2 X in NQ315 4 X in NQ1712	`aaaacaatg`	putative	-	1527	1416	-	
11	1528236	1528325	difference in 18nt repeat 5 X in NQ315 2 X in NQ1712	`cctatattag atagcgtg`	DNA polymerase I	polA	1470	1363	L	

Table S15 b: Differences in strain pair NQ352/1701 not categorized as SNPs or CNPs

No	Position in NQ352		type of difference	sequence	gene affected by difference					comment
	start	end					HP-number	jhp-number	COG	
1	124571	124648	difference in 12nt repeat 6 X in NQ352 4 X in NQ1701	gagattaaaa aa	no annotation/ intergenic					
2	223721	223768	difference in 12nt repeat 4 X in NQ352 3 X in NQ1701	ctgtattttc tt	putative		0167	0153	-	
3	413843	413876	difference in 10nt repeat 3 X in NQ352 2 X in NQ1701	gtgggttttc	septum formation protein	ftsK	1090	0335	D	
4	495338	495403	difference in 33nt repeat 2 X in NQ352 1 X in NQ1701	ttcatcaaaa gaaaccaaag cgcaaaaatc cgc	type I restriction enzyme modification subunit	hsdM_1	0463	0415	V	
5	572684	572737	difference in 18nt repeat 3 X in NQ352 2 X in NQ1701	ttcttgatac tcattatt	cag island protein	orf13/14 cagY	0527	0476	U	
6	733421	733429	difference in 9nt repeat 1X in NQ352 2X in NQ1701	gctttataa	no annotation / intergenic					
7	951712	951741	difference in 13nt repeat 2 X in NQ352 1X in NQ1701	aaagaaaagg gagtt	no annotation/ intergenic					
8	969921	969939	18nt indel	tagattttag actctagct	no annotation/ intergenic					
9	1056588	1056589	3nt indel	aaa	comB9 like competence protein	HPG27_966				
10	1093596	1093637	difference in 21nt repeat 2 X in NQ352 1X in NQ1701	tctcaaacca atgaaagccc c	putative	-	0398	0983	-	
11	1233888	1234680	additional conitg in NQ352		sugar efflux transporter carbonic anhydrase	- -	1185 1186	1111 1112	G P	

No	Position in NQ352		type of difference	sequence	gene affected by difference					comment
	start	end					HP-number	jhp-number	COG	
12	1236256	1236309	54nt indel repeat 1 X in NQ352 2 X inNQ1701	`tatattgctt gtcttttgtt tttcttgttc tacttttatt tgattgttag tctc`	putative	-	1187	1113	-	
13	1297760	1297849	difference in 10nt repeat 9 X in NQ352 8 X in NQ1701	`aagattaaac`	no annotation/ intergenic					
14	1305366	1305379	14nt indel	`cagggctttttctt`	hypothetical protein	-	1252	1173	E	
15	1357965	1358061	difference in 8nt repeat 12 X in NQ352 13X in NQ352	`tatgtatt`	no annotation/ intergenic					
16	1419103	1419126	difference in 12nt repeat 2 X in NQ352 1 X in NQ1701	`gctctttttctt`	probable protein encoding sequence (Kodon)	(BLAST: hypothetical protein HPP12_1343 [H. pylori P12])				
17	1476690	1476763	74nt indel	`aagataaaac catcattaaa aaaacatttc acatttcaca ttcataagat aaacatttca aagataaa`	no annotation/ intergenic					
18	1476863	1476888	difference in repeat	`caaccattca ccattcaatc ttca`	hypothetical protein		1408	1300	-	
19	1495286	1495297	difference in 3nt repeat 4 X in NQ352 3 X in NQ1701	`gtt`	putative histidine and glutamine-rich metal-binding protein	hpn	1432	1321	-	

Table S15 c: Differences in strain set NQ392/1707/4060 not categorized as SNPs or CNPs

No	Position in NQ392		type of differenc	sequence	gene affected by difference				COG	comment
	start	end					HP-number	jhp-number		
1	20160	20939	additional contig (C50) in NQ392		type I restriciton enzyme (modification subunit)	hsdM_3	1403	1423	V	missing in NQ4060 - not associated with Contigborder
2	103912	103916	difference in 3nt repeat 7 X in NQ392 6 X in NQ1707 CNP in NQ4060	`tag`	outer membrane protein		0009	0007	-	
3	104644	104645	2nt indel in NQ1707	`tt`	no annotation / intergenic					
4	309513	309514	difference in 2nt repeat 8 X in NQ392 9 X in NQ1707 9 X in NQ4060	`ct`	putative lipopolysaccharide biosynthesis protein		0208	0194	M	phase variable repeat
5	623709	623741	33nt indel missing in NQ4060	`tttctgtggg gttgtcaaga tgattgtctg agc`	cag island protein	orf13/14	0527	0476	U	
6	864243	864244	difference in 14nt repeat 4 X in NQ392 5 X in NQ1707 3 X in NQ4060	`agcagttaaa aagg`	no annotation / intergenic					
7	887675	887676	2nt indel missing in NQ1707	`cc`	no annotation / intergenic					
8	912602	912623	difference in 11nt repeat 3 X in NQ392 Contigborder in NQ1707 2X in NQ4060	`aaacaaccatt`	no annotation/ intergenic					
9	913274	913276	CNP between NQ1707 and NQ4060		putative type I restriction enzyme (specificity subunit)	hsdS_4		0726	V	

No	Position in NQ392		type of differenc	sequence	gene affected by difference					comment
	start	end					HP-number	jhp-number	COG	
10	1013253	1013266	differnce in 2nt repeat 7 X in NQ392 8 X in NQ1707 8 X in NQ4060	`ct`	outer membrane protein	babB	0896	1164	-	phase variable repeat
11	1029272	1029280	difference in 9nt repeat 3 X in NQ392 3 X in NQ1707 2 X in NQ4060	`aatggaact`	outer membrane protein/ porin	hopC	0912	0848	-	
12	1296875	1296876	difference in slightly degenerated 21nt repeat 5 X in NQ392 Contigborder in NQ1707 6 X in NQ4060	`ctgtcttttg tttttcttgt t`	hypothetical protein		1187	1113	-	
13	1501421	1501422	2nt indel in NQ4060	`ta`	no annotation / intergenic					
14	1563604	1563923	318bp deletion in NQ4060		no annotation / intergenic					
15	1614309	1614310	difference in 9nt repeat 2 X in NQ392 2 X in NQ1707 3 X in NQ4060	`gtttttagt`	hypothetical protein		1482	1375		

Table S15 d: Differences in strain set NQ367/1671/4191 not categorized as SNPs or CNPs

No	Position in NQ367		type of difference	sequence	gene affected by difference				COG	comment
	start	end					HP-number	jhp-number		
1	4477	4478	difference in 12 nt repeat 6 X in NQ367 7 X in NQ1671 8 X in NQ4191	aaaaaagaaa gc	putative recombination protein RecB (addA?)	pcrA	1553	1446	L	
2	53724	53725	difference in 7 nt repeat 7 X in NQ367 9 X in NQ1671 10 X in NQ4191	agtgatt	no annotation/ intergenic					
3	74334	74345	difference in 13 nt repeat 3 X in NQ367 3 X in NQ1671 2 X in NQ4191	attaaggttt ttg	no annotation/ intergenic					
4	91725	91726	difference in 9 nt repeat 8 X in NQ367 7 X in NQ1671 11 X in NQ4191	aaagaaaat	DNA transformation compentancy	comB2	0039/ 0040	0035	U	
5	199955	199956	2 nt indel missing in NQ1671	gt	L-lactate permease	lldP_2	0141	0129	C	
6	216595	216596	difference in 6 nt repeat 1 X in NQ367 1 X in NQ1671 2 X in NQ4191	aatagc	no annotation/ intergenic					
7	570023	570024	2 nt indel missing in NQ1671 and NQ4191	ac	cag island protein, DNA transfer protein	virB11_1	0525	0474	NU	
8	595528	595529	2nt indel , missing in NQ4060	ag	no annotation/ intergenic					
9	677463	677529	difference in 21 nt repeat 4 X in NQ367 4 X in NQ1671 2 X in NQ4191	aaaatacgat gatctcacag	putative lipopolysaccharide biosynthesis protein		0619	0563	M	

No	Position in NQ367		type of difference	sequence	gene affected by difference				COG	comment
	start	end					HP-number	jhp-number		
10	758109	758110	14 nt indel in NQ4191	`aatagctaaa ctct`	putative hydantoin utilization		0695	0633	EQ	
11	961006	961007	134 nt indel in NQ4191 Contigborder in NQ1671		no annotation/ intergenic					
12	978744	978748	difference in 9 nt repeat Contigborder in NQ367 12 X in NQ1707 8 X in NQ4191	`tttattatg`	no annotation/ intergenic					
13	1109538	1109541	difference in 4 nt repeat 3 X in NQ367 3 X in NQ1671 2 X in NQ4191	`cttt`	no annotation/ intergenic					
14	1142005	1142006	difference in 2 nt repeat 2 X in NQ367 3 X in NQ1671 3 X in NQ4191	`tg`	glutamate dehydrogenase	gdhA	0380	1001	E	
15	1327216	1327217	difference in 13 nt repeat 3 X in NQ367 3 X in NQ1671 4 X in NQ4191	`tttgttgttt caa`	no annotation/ intergenic					
16	1389417	1389418	difference in 8 nt repeat Contigborder in NQ367 12 X in NQ1671 7 X in NQ4191	`tttatgta`	no annotation/ intergenic					
17	1592570	1592618	difference in 8 nt repeat 7 X in NQ367 Contigborder in N1671 9 X in NQ4191	`atggttga`	hypothetical protein		1408	1300	-	