Table S3. Sequencing errors in the CB15 genome and the corrected versions represented in the NA1000 genome.

Error number[1]	CB15 position[2]	NA1000 position[3]	Gene annotation[4]	CB15 gene number[2]	NA1000 gene number[3]	Erroneous allele[5]	Corrected allele[5]	Type[6]
21	48812.5	48813	hypothetical cytosolic protein	CC_0046	CCNA_00044	–	T	I
22	56861	56861.5	tRNA 2-methylthioadenosine synthase	CC_0056	CCNA_00054	C	–	I
23	76045	76045	HemY domain membrane protein	CC_0075	CCNA_00073	T	G	S
24	162144.5	162145	heat-inducible transcription repressor, *hrcA*	CC_0153	CCNA_00152	–	G	I
25	162157	162157.5	heat-inducible transcription repressor, *hrcA*	CC_0153	CCNA_00152	C	–	I
26	361895	361895	Intergenic	CC_0342-0343	CCNA_00346-00347	C	T	I/NC
27	712701	738731	Intergenic	CC_0645-0646	CCNA_00682-00683	T	G	I/NC
28	723511	749541	PAS-family GGDEF/EAL protein	CC_0655	CCNA_00692	C	G	NS
29	824099.5	850130	hypothetical protein	CC_0749	CCNA_00786	–	T	I
30	1056600-1046653	1082630.5	hypothetical protein/IS element	CC_0949-CC_0950	CCNA_01000/CCNA_IS023	54 bp	–	I/NC
31	1268666	1294642.5	Intergenic	CC_1117-1118	CCNA_01174-01175	G	–	I/NC
32	1656828-1656839	1682979.5	Intergenic	CC_1497-1498	CCNA_01565-CCNA_01566	AAGAGAGAGGGA	–	I/NC
33	1958549	1984525	major facilitator superfamily transporter	CC_1774	CCNA_01852	C	G	S
34	2119218	2145194	Intergenic	CC_1921-1922	CCNA_01998-01999	T	G	NC
35	2119221	2145197	Intergenic	CC_1921-1922	CCNA_01998-01999	T	C	NC
36	2119229	2145205	Intergenic	CC_1921-1922	CCNA_01998-01999	A	G	NC
37	2136567	2162543	NADH-quinone oxidoreductase chain N	CC_1937	CCNA_02015	A	G	S
38	2403008	2428985-2428988	Intergenic	CC_2194-2195	CCNA_02277-02278	–	CGCG	I/NC
39	2411028	2437008	tRNA synthetase class I, catalytic domain protein	CC_2205	CCNA_02288	C	T	NS
40	2455419	2481400	Intergenic	CC_2258-2259	CCNA_02341	–	C	I/NC
41	2704184	2730166	cytochrome P450	CC_2494	CCNA_02579	C	G	NS
42	2761571	2787556.5	Intergenic	CC_2552-2553	CCNA_02635-02636	C	–	I/NC
43	2766758	2792741.5	phospho-N-acetylmuramoyl-pentapeptide transferase	CC_2557	CCNA_02640	C	–	I/NC
44	2815721	2841701	Intergenic	CC_2604-2605	CCNA_02687-02688	G	C	I/NC
45	2986972	3012952	chaperone protein, *dnaJ*	CC_2772	CCNA_02860	A	G	NS
46	3004624	3030603	Intergenic	CC_2791-2792	CCNA_02881-02882	G	C	NC
47	3004630	3030608.5	Intergenic	CC_2791-2792	CCNA_02881-02882	T	–	I/NC
48	3124497	3150478	Intergenic	CC_2900-2901	CCNA_02994-02995	–	G	I/NC
49	3158214	3184195	3-deoxy-manno-octulosonate cytidylyltransferase	CC_2934	CCNA_03029	A	G	S
50	3163968	3189949	tyrosine phenol-lyase	CC_2940	CCNA_03035	T	G	S
51	3163974	3189955	tyrosine phenol-lyase	CC_2940	CCNA_03035	C	G	S
52	3195323	3221304-3221407	Intergenic	CC_2975-2976	CCNA_03070-3071	–	CGCG	I/NC
53	3240112	3266097	hypothetical protein/IS element	CC_3021-3022	CCNA_03116	A	G	NC
54	3283569	3309555	hypothetical protein	CC_3057	CCNA_03153	–	C	I/NC
55	3351760	3377746	hypothetical protein	CC_3117	CCNA_03216	G	C	NS
56	3392424	3418413-3418415	Intergenic	CC_3146-3147	CCNA_03248-03249	–	GGC	I/NC
57	3570312	3596301	aldolase 1-epimerase	CC_3323	CCNA_03432	A	G	NS
58	3638746	3664738	Intergenic	CC_3395-3396	CCNA_03506-03507	–	G	I/NC
59	3687714	3713704	hypothetical protein	CC_3444	CCNA_03557	–	G	I
60	3730223	3756228.5	patatin family phospholipase domain protein	CC_3488	CCNA_03602	ATGTGCTGG	–	I
61	3762779	3788761	hypothetical protein	CC_3519	CCNA_03634	T	C	NS
62	3785933	3811915	ATP-dependent nuclease A subunit	CC_3538	CCNA_03652	G	T	NS
63	3796031	3822013	hypothetical protein	CC_3548	CCNA_03662	T	C	S
64	3960349	3986331	cell division protein, *ftsK*	CC_3704	CCNA_03819	A	C	NS
65	4003781	4029763	DNA polymerase III, delta subunit	CC_3751	CCNA_03867	G	C	NS

[1] Errors numbers are continued from Table S2.

[2] Postitions and gene numbers for CB15 are consistent with the published genome sequence, GenBank Accession AE005673.

[3] Postitions and gene numbers for NA1000 are consistent with the newly assembled and annotated genome, GenBank Accession CP001340.

[4] Gene annotations were determined using the Integrated Genomics gene annotation algorithm and comparison with protein sequences in GenBank.

[5] Erroneous and corrected alleles are represented in the CB15 (AE005673) and NA1000 (CP001340) genome sequences, respectively

[6] Type of mutation: I = insertion/deletion, NC = non-coding, S = synonymous, NS = non-synonymous