

<b>Locus Tag</b>	<b>Identity<sup>c</sup></b>	<b>Closest relationship<sup>a</sup></b>	<b>Identity<sup>c</sup></b>	<b>Identity with BH-01-0142 (HS3/13/50)</b>
HS3.01	MeOPN biosynthesis	Cj1415c (cysC)	168/170 (98%)	170/170 (100%)
HS3.02	MeOPN biosynthesis	CJJ81176_1414	253/253 (100%)	253/253 (100%)
HS3.03	MeOPN biosynthesis	CJJ81176_1416	199/200 (99%)	200/200 (100%)
HS3.04	MeOPN biosynthesis	CJJ81176_1417	776/779 (99%)	766/766 (100%)
HS3.05	Unknown	CJJ81176_1418	258/259 (99%)	262/262 (100%)
HS3.06*	Unknown	CJJ81176_1419	254/257 (98%)	267/267 (100%)
HS3.07*	putative MeOPN transferase	Cj8421_1470	323/632 (51%)	574/577 (99%)
HS3.08	putative sugar-phosphate nucleotidyltransferase	hddC	212/224 (94%)	224/224 (100%)
HS3.09	sedoheptulose 7-phosphate isomerase	gmhA2	162/166 (97%)	194/194 (100%)
HS3.10*	putative sugar transferase	HS41.09	182/358 (50%)	335/335 (100%)
HS3.11	D-glycero-Dmanno-heptose 1-phosphate kinase	hddA	339/339 (100%)	339/339 (100%)
HS3.12	putative sugar-nucleotide epimerase/dehydratase	Cj8421_1474	311/312 (99%)	311/312 (99%)
HS3.13*	GDP-mannose 4,6-dehydratase	dmhA	357/358 (99%)	358/358 (100%)
HS3.14	GDP-fucose synthetase	fcl	338/346 (97%)	346/346 (100%)
HS3.15	putative nucleotide-sugar epimerase/dehydratase	Cj8486_1471c	176/181 (97%)	181/181 (100%)
HS3.16	putative sugar transferase	hddD	253/593 (42%)	598/598 (100%)
HS3.17*	putative sugar transferase	HS41.06	205/597 (35%)	517/517(100%)
HS3.18	putative sugar transferase	JJD26997_1424	94/96 (97%)	97/97 (100%)
HS3.19	NAD dependent epimerase/dehydratase family	JJD26997_1425	194/199 (97%)	207/207 (100%)
HS3.20	putative sugar transferase	JJD26997_1426	431/449 (95%)	450/450 (100%)
HS3.21	putative sugar transferase	CJJ26094_1492	168/471 (35%)	453/453 (100%)
HS3.22	putative sugar transferase	ggaB	728/796 (91%)	804/804 (100%)
HS3.23	putative glycosyl transferase	Cj8486_1479c	512/542 (94%)	542/542 (100%)

<b>Locus Tag</b>	<b>Putative function<sup>a</sup></b>	<b>Closest relationship<sup>a</sup></b>	<b>Identity<sup>c</sup></b>
HS4.01	MeOPN biosynthesis	Cj8486_1458c	168/170 (98%)
HS4.02	MeOPN biosynthesis	Cj8486_1459c	253/253 (100%)
HS4.03	MeOPN biosynthesis	Cj8486_1460c	197/200 (98%)
HS4.04	MeOPN biosynthesis	Cj8486_1461c	776/779 (99%)
HS4.05	Unknown	Cj8486_1462c	249/253 (98%)
HS4.06*	Unknown	Cj8486_1463c	256/257 (99%)
HS4.07*	putative MeOPN transferase	Cj8486_1464c	91/359 (25%)
HS4.08	putative sugar-phosphate nucleotidyltransferase	hddC	217/224 (97%)
HS4.09	sedoheptulose 7-phosphate isomerase	gmhA2	193/201 (96%)
HS4.10	D-glycero-Dmanno-heptose 1-phosphate kinase	hddA	333/339 (98%)
HS4.11	putative sugar-nucleotide epimerase/dehydratease	Cj8486_1468c	312/312 (100%)
HS4.12*	GDP-mannose 4,6-dehydratase	dmhA	343/344 (99%)
HS4.13	GDP-fucose synthetase	fcl	337/347 (98%)

HS4.14	putative nucleotide-sugar epimerase/dehydratase	Cj8486_1471c	176/181 (97%)
HS4.15	putative sugar transferase	Cj8486_1472c <sup>d</sup>	1227/1241(99%)
HS4.16*	putative MeOPN transferase	Cj8486_1475	332/628 (53%)
HS4.17	putative sugar transferase	ggaB	773/838 (92%)
HS4.18	putative glycosyltransferase	Cj8486_1479c	514/542 (95%)

<b>Locus Tag</b>	<b>Putative function<sup>a</sup></b>	<b>Closest relationship<sup>b</sup></b>	<b>Identity<sup>c</sup></b>
HS10.01	MeOPN biosynthesis	Cj8486_1458c	168/170 (98%)
HS10.02	MeOPN biosynthesis	Cj8486_1459c	253/253 (100%)
HS10.03	MeOPN biosynthesis	Cj8486_1460c	199/200 (99%)
HS10.04	MeOPN biosynthesis	Cj8486_1461c	748/752 (99%)
HS10.05	Unknown	Cj8486_1462c	252/253 (99%)
HS10.06*	Unknown	Cj8486_1463c	254/257 (98%)
HS10.07*	MeOPN transferase	HS41.05	232/633 (37%)
HS10.08	heptosyltransferase	Cj1431c (hddD)	180/589 (30%)
HS10.09	putative sugar-phosphate nucleotidyltransferase	hddC	181/198 (91%)
HS10.10	sedoheptulose 7-phosphate isomerase	gmhA2	193/201 (96%)
HS10.11	D-glycero-D-manno-heptose 7-phosphate kinase	hddA	330/339 (97%)
HS10.12	GDP-mannose 4,6-dehydratase	dmhA	327/343 (95%)
HS10.13	GDP-fucose synthetase	fcl	342/353 (96%)
HS10.14*	GDP-fucose synthetase	CJJ26094_1481 (fcl2)	352/357 (98%)
HS10.15	putative nucleotide-sugar epimerase-dehydratase	CJJ26094_1482	175/181 (96%)
HS10.16	putative nucleotidyl-sugar pyranose mutase	CJJ26094_1483 (glF)	354/406 (88%)
HS10.17	putative sugar transferase	CJJ81176_1433	171/487 (36%)
HS10.18	putative sugar transferase	CJJ26094_1490	33/102 (32%)
HS10.19	putative sugar transferase	CJJ26094_1492	93/287 (32%)
HS10.20	UDP-galactopyranose mutase	CJJ26094_1485 (glF3)	200/375 (53%)
HS10.21	putative molybdenum cofactor biosynthesis protein	CJJHB9313_1449	205/341 (60%)
HS10.22	glucose-1-phosphate cytidyltransferase	CJJ26094_1489	167/239 (69%)
HS10.23	putative sugar transferase	ggaB	719/813 (89%)
HS10.24	putative sugar transferase	Cj8486_1479c	117/121 (96%)
HS10.25	putative sugar transferase	Cj8486_1479c	404/419 (96%)

<b>Locus Tag</b>	<b>Putative function<sup>a</sup></b>	<b>Closest relationship<sup>b</sup></b>	<b>Identity<sup>c</sup></b>
HS15.01	MeOPN biosynthesis	Cj8486_1458c	168/170 (98%)
HS15.02	MeOPN biosynthesis	Cj8486_1459c	253/253 (100%)
HS15.03	MeOPN biosynthesis	Cj8486_1460c	199/200 (99%)
HS15.04	MeOPN biosynthesis	Cj8486_1461c (ppsA)	768/779 (98%)

HS15.05	Unknown	Cj8486_1462c	251/253 (99%)
HS15.06*	Unknown	Cj8486_1463c	254/257 (99%)
HS15.07*	MeOPN transferase	Cj8486_1464c	496/523 (94%)
HS15.08	putative sugar-phosphate nucleotidyltransferase	hddC	209/224 (93%)
HS15.09	sedoheptulose 7-phosphate isomerase	gmhA2	193/193 (100%)
HS15.10	D-glycero-Dmanno-heptose 1-phosphate kinase	hddA	330/339 (98%)
HS15.11	GDP-mannose 4,6-dehydratase	dmhA	328/343 (95%)
HS15.12	GDP-fucose synthetase	fcl	198/353 (56%)
HS15.13	putative nucleotide-sugar epimerase/dehydratase	Cj8486_1471c	80/99 (81%)
HS15.14	putative sugar transferase	CJJ26094_1484	168/549 (30%)
HS15.15	putative glycosyl transferase	CJM1_1382	132/282 (47%)
HS15.16	UDP-galactopyranose mutase	CJJ26094_1493 (glF3)	356/370 (97%)
HS15.17	UDP-glucose 4-epimerase	CJJ26094_1494 (galE1)	319/324 (98%)
HS15.18	UDP-glucose 6-dehydrogenase	CJJ26094_1495 (udg)	408/432 (94%)
HS15.19	NAD dependent epimerase/dehydratase family	CJJ26094_1496 (galE2)	307/318 (96%)
HS15.20	glycosyl transferase	CJJ26094_1497	444/460 (96%)
HS15.21	putative glycosyltransferase	CJJ26094_1498	327/347 (94%)
HS15.22	putative glycosyltransferase	CJJ26094_1499	513/543 (94%)

Locus Tag	Putative function <sup>a</sup>	Closest relationship <sup>b</sup>	Identity <sup>c</sup>	Identity with ATCC 43436 (HS8)
HS17.01	putative Na <sup>+</sup> /H <sup>+</sup> antiporter	HS41.01	410/417 (99%)	417/417 (100%)
HS17.02	Pyruvate kinase	HS41.02	317/319 (99%)	319/319 (100%)
HS17.03	putative sulfate adenylyltransferase	HS41.03	343/348 (99%)	348/348 (100%)
HS17.04	putative adenylylsulfate kinase	HS41.04	173/177 (98%)	177/177 (100%)
HS17.05*	putative sugar transferase	HS41.05	324/605 (54%)	603/604 (99%)
HS17.06	putative sugar-phosphate nucleotidyltransferase	hddC	224/224 (100%)	186/200 (93%)
HS17.07	putative phosphoheptose isomerase	gmhA2	171/178 (97%)	169/177 (95%)
HS17.08	putative sugar kinase	hddA	335/339 (99%)	335/339 (99%)
HS17.09	GDP-mannose 4,6-dehydratase	dmhA	331/343 (97%)	351/357 (98%)
HS17.10	putative fucose synthetase	fcl	337/346 (98%)	357/357 (100%)
HS17.11*	putative fucose synthetase	fcl2	343/348 (99%)	395/396 (99%)
HS17.12*	hypothetical protein	Cj1429c	304/308 (99%)	309/310 (99%)
HS17.13	putative dTDP-4-dehydrorhamnose 3,5-epimerase	Cj1430c (rbfC)	173/176 (99%)	181/181 (100%)
HS17.14	hypothetical protein	CJE1603	209/671 (32%)	620/622 (99%)
HS17.15	putative sugar transferase	CJJ26094_1474	131/417 (32%)	422/424 (99%)
HS17.16	putative CDP glycerol glycerophosphotransferase	HS1.09 (tagF)	464/1096 (43%)	1570/1577 (99%)
HS17.17	hypothetical protein	HS1.10	366/381 (97%)	381/385 (99%)
HS17.18	putative Glycerol-3-phosphate	HS1.11 (tagD)	127/129 (99%)	128/129 (99%)

Locus Tag	Putative function <sup>a</sup>	Closest relationship <sup>b</sup>	Identity <sup>c</sup>
HS42.01	putative Na <sup>+</sup> /H <sup>+</sup> antiporter	HS41.01	410/417 (98%)
HS42.02	Pyruvate kinase	HS41.02	316/319 (99%)
HS42.03	putative sulfate adenylyltransferase	HS41.03	343/348 (98%)
HS42.04	putative adenylylsulfate kinase	HS41.04	173/177 (97%)
HS42.05*	putative sugar transferase	HS41.05	323/605 (54%)
HS42.06	putative sugar-phosphate nucleotidyltransferase	hddC	203/224 (90%)
HS42.07	sedoheptulose 7-phosphate isomerase	gmhA2	197/201 (98%)
HS42.08	D-glycero-Dmanno-heptose 1-phosphate kinase	hddA	333/339 (98%)
HS42.09	GDP-mannose 4,6-dehydratase	dmhA	332/343 (96%)
HS42.10	GDP-fucose synthetase	fcl	343/353 (97%)
HS42.11	putative nucleotide-sugar epimerase/dehydratase	Cj1430c	160/176 (90%)
HS42.12	putative sugar transferase	HS41.16	170/525 (32%)
HS42.13	putative sugar transferase	HS41.22	126/345 (36%)
HS42.14	putative sugar transferase	C8J_1342	71/248 (28%)
HS42.15	nucleotidyl-sugar pyranose mutase	HS41.24 (glf3)	355/364 (97%)
HS42.16	UDP-glucose 4-epimerase	HS41.25 (galE1)	312/324 (96%)
HS42.17	putative UDP-glucose 6-dehydrogenase	HS41.26 (udg)	411/432 (95%)
HS42.18	nucleotidyl-sugar pyranose mutase	HS41.27 (galE2)	308/318 (96%)
HS42.19	putative sugar transferase	HS41.28	432/460 (93%)
HS42.20	putative sugar transferase	HS41.29	325/349 (93%)
HS42.21	putative sugar transferase	HS41.30	506/543 (93%)

Sup. Table 1: Capsule biosynthesis genes of HS3, HS4, HS10, HS15, HS17 and HS42. (a) Function attribution of genes and (b) closest relationship are the best BLASTP match on the non-redundant protein sequences database.

(c) In parenthesis is the percentage of identity between the protein and its best match in the protein database.

(d) Compared to the published sequence of HS4/13/64, the HS4 type strain appeared to have a fusion of the homolog of Cj8486\_1472c and Cj8486\_1474c. Sequence analysis confirmed the fusion of those genes in CG8486 as well. The sequence has been updated in the NCBI database as Cj8486\_1472c.

(\*) Homopolymeric GC tract.