

SUPPLEMENTARY FIGURES AND TABLES

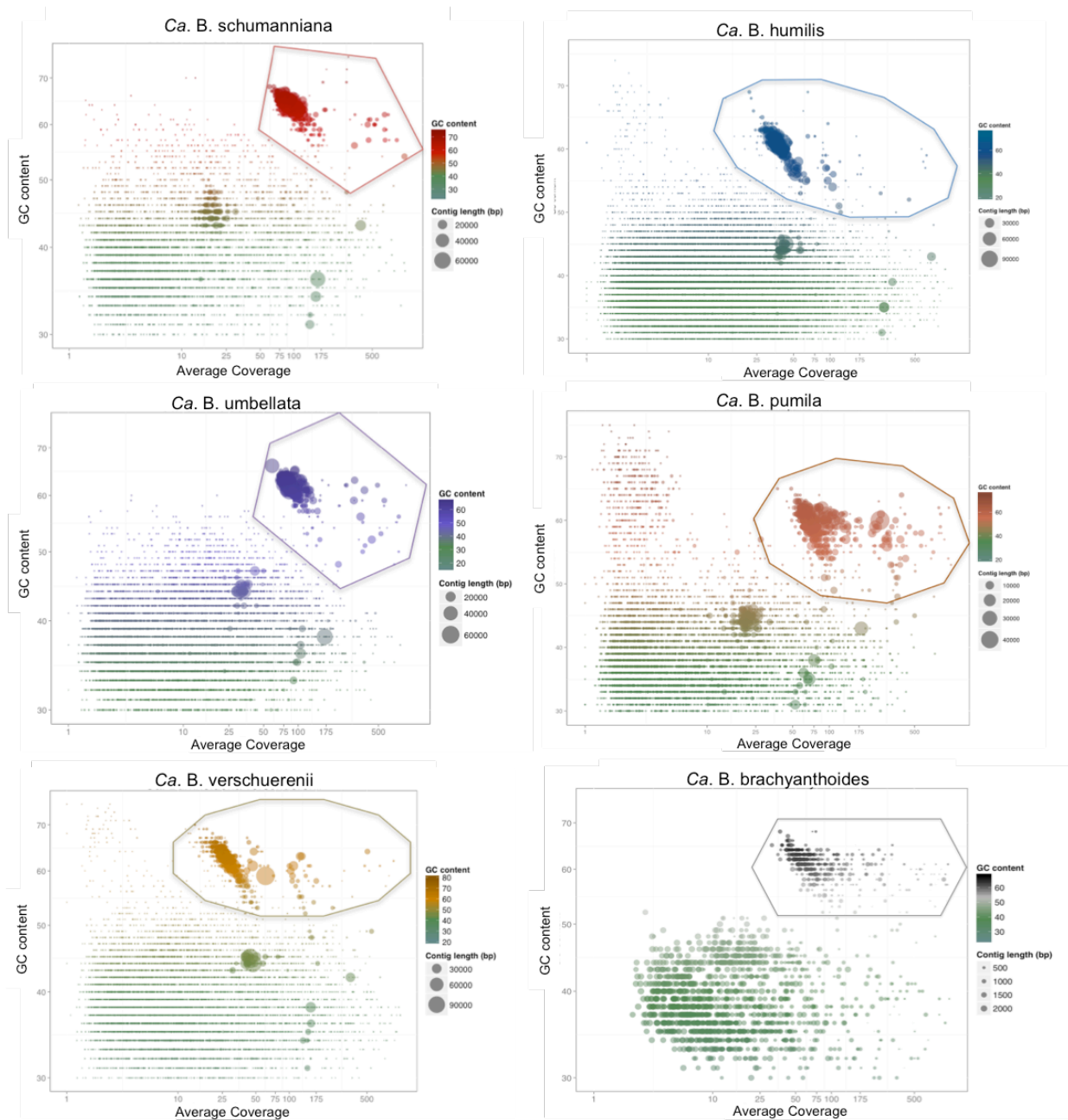


Fig. S1. Shotgun assembly of leaf nodule DNA. Contigs were plotted according to average coverage and GC content. Circles represent contigs, scaled by the square root of their length and colored by GC content. Contigs with high %GC and high coverage are bacterial in origin, while the rest comes from plant DNA. Contigs selected for bacteria genome assembly are circled. Large contigs with moderate coverage and low %GC comprise the chloroplast genome.

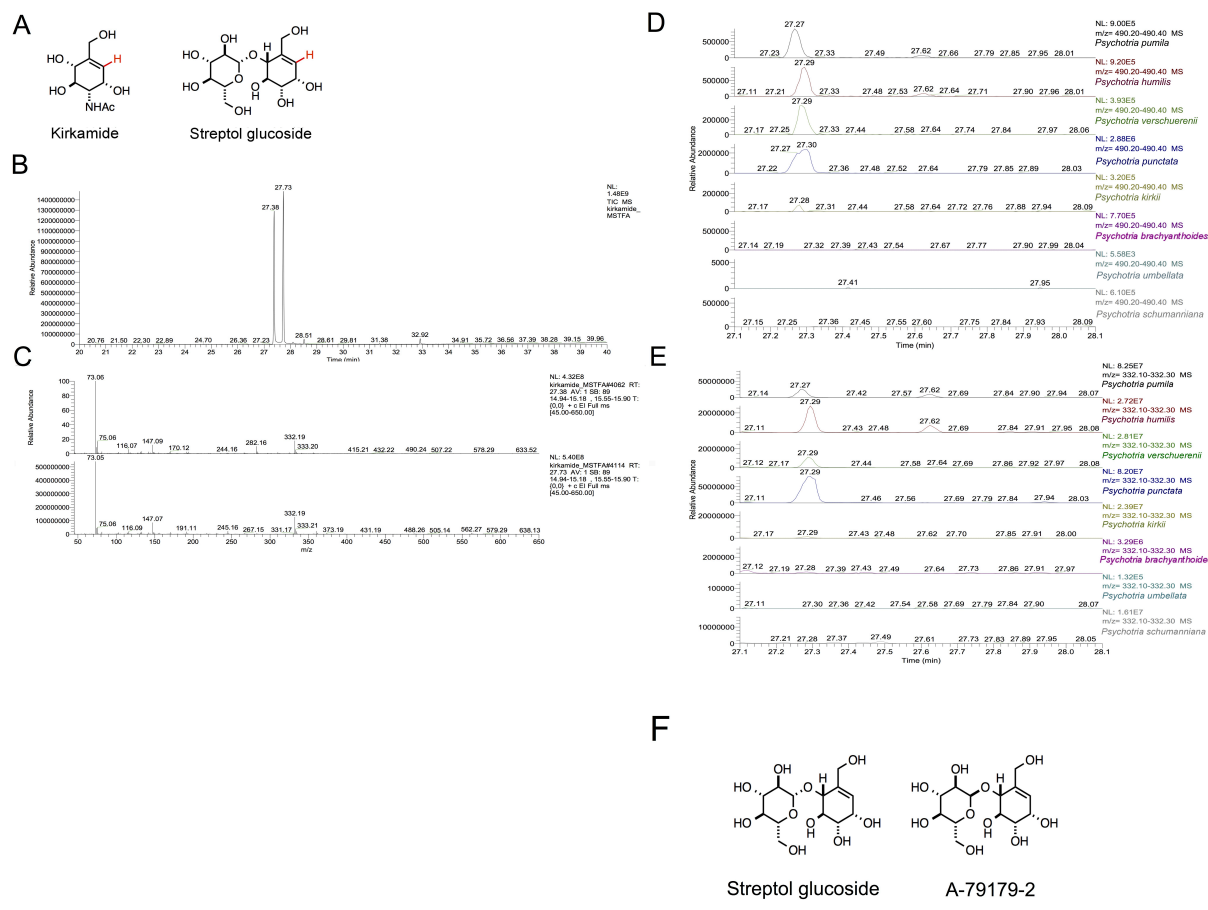
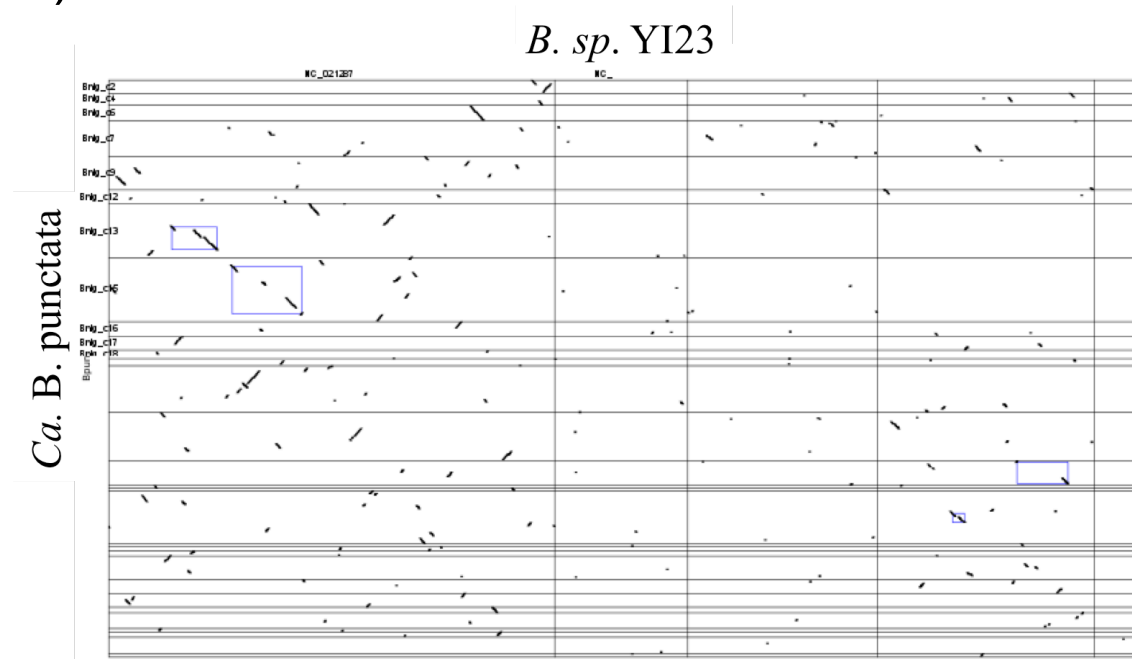


Fig. S2. Detection of kirkamide in extracts of *Psychotria* and *Pavetta* leaves by GC-MS. A. Natural products (kirkamide and streptol glucoside) quantified by NMR, the protons highlighted in red were used for the measurement. B. Total Ion Current (TIC) chromatogram of synthetic TMS-kirkamide. B. Mass spectra of synthetic TMS-kirkamide at the retention times of 27.38 min (kirkamide with 4 TMS) and 27.73 min (kirkamide with 5 TMS). C. GC-MS chromatograms of leaf extracts after MSTFA derivatization. Two diagnostic compounds with $m/z = 490$ Da and 577 Da were used to detect kirkamide in complex extracts. X-axis: Retention time of the analytes (min), y-axis= Relative abundance. F. Structural comparison between streptol glucoside and A-79179-2.

Table S1. Representation of the insertion sequence families diversity and predominance in the genomes of leaf nodule symbionts. The families highlighted in grey are represented by at least one full transposase copy. Pseudogenes are indicated by Ψ .

	Ca. B. kirkii	Ca. B. punctata	Ca. B. pumila	Ca. B. schumanniana	Ca. B. umbellata	Ca. B. brachyanthoides	Ca. B. humilis	Ca. B. verschuerenii
IS3	1/ 5 Ψ	36 Ψ		10 Ψ	1 Ψ	1/ 0 Ψ	1/ 1 Ψ	
IS4	4/ 18 Ψ	18 Ψ		1/ 9 Ψ			1 Ψ	1/ 1 Ψ
IS5	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 2 Ψ
IS200	1/ 4 Ψ	1/ 0 Ψ	1/ 0 Ψ	2 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ	1/ 0 Ψ
IS256					153 Ψ			2 Ψ
IS30			17 Ψ			1/ 0 Ψ		
IS66	9/ 38 Ψ	1/ 0 Ψ	9 Ψ	1/ 56 Ψ	1/ 2 Ψ	1/ 2 Ψ	1 Ψ	3 Ψ
IS110			4 Ψ	1/ 117 Ψ	1/ 167 Ψ		1/ 25 Ψ	
IS1182					88 Ψ			1 Ψ

A)



B)

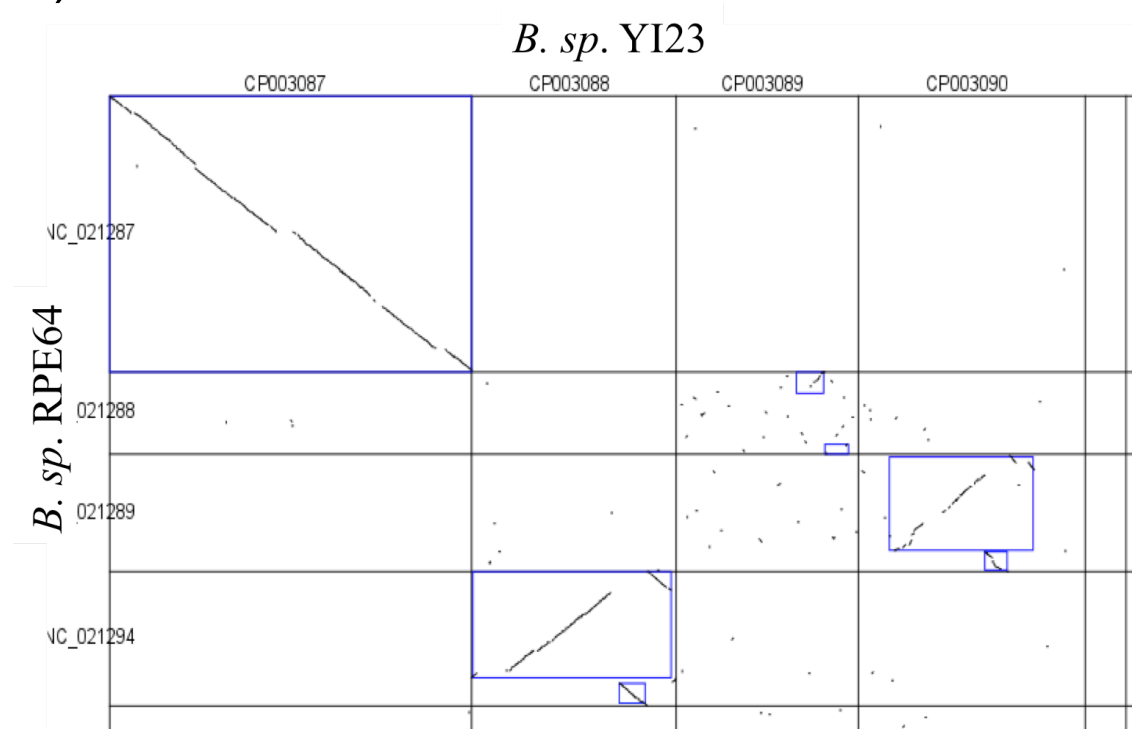


Fig. S3. Dot plot representation of Mummer alignment between A. *Ca. B. punctata* and *B. sp. YI23*. B. *B. sp. RPE64* and *B. sp. YI23* as a visual reference. The degrees of relatedness between *Ca. B. punctata* and *B. sp. YI23* and *B. sp. RPE64* and *B. sp. YI23* are similar (see text). Note the lack of large syntenic stretches between *Ca. B. punctata* and *B. sp. YI23*.

Table S2. Comparative GC-content between functional genes and pseudogenes in leaf nodule symbionts.

	Ca. B. kirkii	Ca. B. punctata	Ca. B. brachyanthoides	Ca. B. humilis	Ca. B. pumila	Ca. B. schumanniana	Ca. B. umbellata	Ca. B. verschuerenii
%GC functional	59.34	63.10	64.38	62.12	60.78	60.15	63.83	61.93
%GC pseudogenes	59.10	62.77	63.29	62.19	60.55	59.86	63.23	61.97
stdev functional	3.43	3.66	3.50	3.74	3.64	3.61	3.58	3.75
stdev pseudogenes	5.55	5.65	7.01	5.76	6.72	5.40	6.47	5.83

Table S3. Essential candidate genes in *B. phymatum* missing in leaf nodule symbionts.

Gene identifiers in <i>B. phymatum</i>	Predicted Product
gij186471077 ref YP_001862395.1	2-oxoglutarate dehydrogenase E1 component [Burkholderia phymatum STM815]
gij186471992 ref YP_001859334.1	phosphate ABC transporter substrate-binding protein [Burkholderia phymatum STM815]
gij186473727 ref YP_001861069.1	hexapaptide repeat-containing transferase [Burkholderia phymatum STM815]
gij186474513 ref YP_001863484.1	2-oxo-acid dehydrogenase E1 subunit, homodimeric type [Burkholderia phymatum STM815]
gij186475621 ref YP_001857091.1	hypothetical protein Bphy_0856 [Burkholderia phymatum STM815]
gij186476921 ref YP_001858391.1	hypothetical protein Bphy_2170 [Burkholderia phymatum STM815]
gij186477415 ref YP_001858885.1	redoxin domain-containing protein [Burkholderia phymatum STM815]

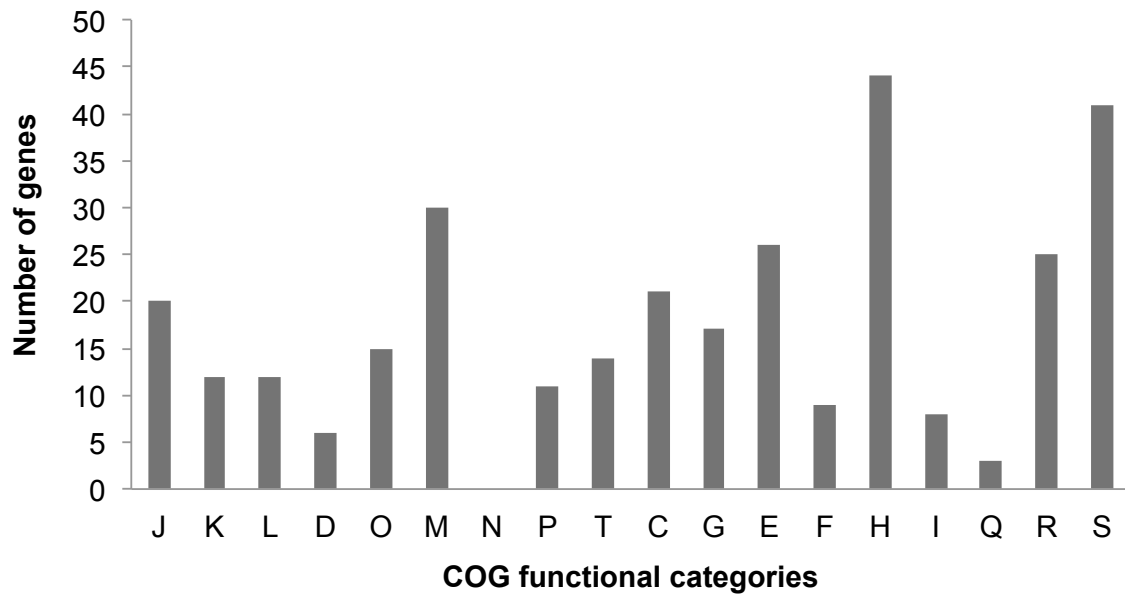


Fig. S4. Categories of non-essential genes conserved in leaf nodule symbionts. Cell envelope biogenesis and outer membrane (M) and coenzyme metabolism (H) are over-represented among the other gene categories.

Table S4. List of non-essential core genes conserved in leaf nodule symbionts. The genes involved in cell envelope biogenesis and outer membrane (M) and coenzyme metabolism (H) are shown in bold.

Gene locus tag	Product	Functional COG categories	Reactive Oxygen Species (ROS) detoxification	
Bkir_c10_6262	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha	Amino acid transport and metabolism	E	
Bkir_c12_2336	Diaminopimelate epimerase	Amino acid transport and metabolism	E	
Bkir_c14_3241	D-amino acid dehydrogenase small subunit	Amino acid transport and metabolism	E	
Bkir_c149_4879	3-dehydroquinate synthase	Amino acid transport and metabolism	E	
Bkir_c18_6035	Aspartate-semialdehyde dehydrogenase	Amino acid transport and metabolism	E	
Bkir_c18_6039	Phosphoribosylanthranilate isomerase	Amino acid transport and metabolism	E	
Bkir_c18_6040	Tryptophan synthase beta chain	Amino acid transport and metabolism	E	
Bkir_c18_6041	Tryptophan synthase alpha chain	Amino acid transport and metabolism	E	
Bkir_c2_4615	5-methyltetrahydrofolate--homocysteine methyltransferase	Amino acid transport and metabolism	E	
Bkir_c2_4616	5-methyltetrahydrofolate--homocysteine methyltransferase	Amino acid transport and metabolism	E	
Bkir_c31_1433	Glycine oxidase ThiO	Amino acid transport and metabolism	E	
Bkir_c31_1438	Glutamate synthase [NADPH] large chain	Amino acid transport and metabolism	E	
Bkir_c38_2423	putative cationic amino acid transporter	Amino acid transport and metabolism	E	
Bkir_c4_0303	2-isopropylmalate synthase	Amino acid transport and metabolism	E	
Bkir_c43_2113	Nitrogen regulatory protein P-II	Amino acid transport and metabolism	E	
Bkir_c51_2630	Shikimate kinase I	Amino acid transport and metabolism	E	
Bkir_c62_5030	L-threonine 3-O-phosphate decarboxylase	Amino acid transport and metabolism	E	
Bkir_c7_1468	N-acetyl-gamma-glutamyl-phosphate reductase	Amino acid transport and metabolism	E	
Bkir_c7_1477	Cysteine synthase B	Amino acid transport and metabolism	E	
Bkir_c7_1489	Cyclohexadienyl dehydrogenase	Amino acid transport and metabolism	E	
Bkir_c72_1660	Glycine cleavage system H protein	Amino acid transport and metabolism	E	
Bkir_c72_1662	Glycine dehydrogenase [decarboxylating](glycine cleavage system P protein)	Amino acid transport and metabolism	E	
Bkir_c77_5215	N-acetylglutamate synthase	Amino acid transport and metabolism	E	
Bkir_c132_1779	Phosphoadenylyl-sulfate reductase [thioredoxin] / Adenylyl-sulfate reductase [thioredoxin]	Amino acid transport and metabolism / Coenzyme metabolism	EH	Thioredoxin pathway
Bkir_c132_1780	Sulfate adenylyltransferase subunit 2	Amino acid transport and metabolism / Coenzyme metabolism	EH	
Bkir_c31_1437	Glutamate synthase [NADPH] small chain	Amino acid transport and metabolism / General function prediction only	ER	
Bkir_c1_2865	Alpha,alpha-trehalose-phosphate synthase[UDP-forming]	Carbohydrate transport and metabolism	G	
Bkir_c108_3931	Mannose-6-phosphate isomerase	Carbohydrate transport and metabolism	G	

Bkir_c128_1243	Fructose-bisphosphate aldolase class I	Carbohydrate transport and metabolism	G	
Bkir_c13_6194	Trehalose-6-phosphate phosphatase	Carbohydrate transport and metabolism	G	
Bkir_c15_5370	Phosphomannomutase /Phosphoglucomutase	Carbohydrate transport and metabolism	G	
Bkir_c20_3823	Phosphoenolpyruvate synthase	Carbohydrate transport and metabolism	G	
Bkir_c22_0076	Pyruvate kinase	Carbohydrate transport and metabolism	G	
Bkir_c44_1547	Polymyxin resistance protein PmrJ, predicted deacetylase	Carbohydrate transport and metabolism	G	
Bkir_c58_0780	Glucose-6-phosphate isomerase	Carbohydrate transport and metabolism	G	
Bkir_c81_1223	Glucose-6-phosphate 1-dehydrogenase	Carbohydrate transport and metabolism	G	
Bkir_c83_0844	Putative transport protein	Carbohydrate transport and metabolism	G	
Bkir_c79_3384	PTS system nitrogen-specific IIA component,PtsN	Carbohydrate transport and metabolism / Signal transduction mechanisms	GT	
Bkir_c11_4427	Chromosome partition protein smc	Cell division and chromosome partitioning	D	
Bkir_c2_4610	Cell division protein FtsN	Cell division and chromosome partitioning	D	
Bkir_c34_2546	Cell division protein FtsL	Cell division and chromosome partitioning	D	
Bkir_c47_5649	Cell division topological specificity factor MinE	Cell division and chromosome partitioning	D	
Bkir_c47_5650	Septum site-determining protein MinD	Cell division and chromosome partitioning	D	
Bkir_c47_5651	Septum site-determining protein MinC	Cell division and chromosome partitioning	D	
Bkir_c1_2894	Outer membrane protein (porin)	Cell envelope biogenesis, outer membrane	M	
Bkir_c106_6216	Membrane fusion protein of RND family multidrug efflux pump	Cell envelope biogenesis, outer membrane	M	
Bkir_c118_3289	Prolipoprotein diacylglycerol transferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c123_5759	Outer membrane lipoprotein LolB precursor	Cell envelope biogenesis, outer membrane	M	
Bkir_c15_5294	dTDP-4-dehydrorhamnose 3,5-epimerase	Cell envelope biogenesis, outer membrane	M	
Bkir_c15_5295	Glucose-1-phosphate thymidyltransferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c15_5368	Lipopolysaccharide heptosyltransferase I	Cell envelope biogenesis, outer membrane	M	
Bkir_c15_5374	Lipid carrier :UDP-N-acetylgalactosaminyltransferase /Alpha-1,3-N-acetylgalactosamine transferase PglA , Putative glycosyltransferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c15_5375	UDP-glucose 4-epimerase	Cell envelope biogenesis, outer membrane	M	
Bkir_c15_5376	Undecaprenyl-phosphate N-acetylglucosaminyl1-phosphate transferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c167_0275	Probable transmembrane protein	Cell envelope biogenesis, outer membrane	M	
Bkir_c17_4544	Rod shape-determining protein MreD	Cell envelope biogenesis, outer membrane	M	
Bkir_c2_4634	Cytidyltransferase-related domain protein	Cell envelope biogenesis, outer membrane	M	
Bkir_c22_0100	Lipid A core-O-antigen ligase and related enzymes	Cell envelope biogenesis, outer membrane	M	
Bkir_c22_0122	N-acetylmuramoyl-L-alanine amidase	Cell envelope biogenesis, outer membrane	M	
Bkir_c32_4006	Glycosyltransferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c32_4007	Lipopolysaccharide core biosynthesis glycosyltransferase, group 2 family protein	Cell envelope biogenesis, outer membrane	M	

Bkir_c34_2536	Cell division protein FtsQ	Cell envelope biogenesis, outer membrane	M	
Bkir_c34_2542	Phospho-N-acetylmuramoyl-pentapeptide- transferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c34_2543	UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diamino pimelate--D-alanyl-D-alanine ligase	Cell envelope biogenesis, outer membrane	M	
Bkir_c4_0300	Cell wall-associated hydrolases (invasion-associated proteins)	Cell envelope biogenesis, outer membrane	M	
Bkir_c44_1550	Polymyxin resistance protein ArnC, glycosyltransferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c44_1551	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarateaminotransferase	Cell envelope biogenesis, outer membrane	M	
Bkir_c44_1554	Polymyxin resistance protein ArnT, undecaprenylphosphate-alpha-L-Ara4N transferase, Melittin resistanceprotein PqaB	Cell envelope biogenesis, outer membrane	M	
Bkir_c7_1480	ADP-heptose synthase /D-glycero-beta-D-mannoheptose 7-phosphate kinase	Cell envelope biogenesis, outer membrane	M	
Bkir_c7_1481	UDP-glucose dehydrogenase	Cell envelope biogenesis, outer membrane	M	
Bkir_c27_1826	Nucleoside-diphosphate sugar epimerase/dehydratase	Cell envelope biogenesis, outer membrane / Carbohydrate transport and metabolism	MG	
Bkir_c27_1829	UDP-glucose 4-epimerase	Cell envelope biogenesis, outer membrane / Carbohydrate transport and metabolism	MG	
Bkir_c44_1548	Polymyxin resistance protein ArnA_DH,UDP-glucuronic acid decarboxylase	Cell envelope biogenesis, outer membrane / Carbohydrate transport and metabolism	MG	
Bkir_c7_1479	ADP-L-glycero-D-mannoheptose-6-epimerase	Cell envelope biogenesis, outer membrane / Carbohydrate transport and metabolism	MG	
Bkir_c10_6252	Ferrochelataze, protoheme ferro-lyase	Coenzyme metabolism	H	
Bkir_c105_1108	Chorismate--pyruvate lyase	Coenzyme metabolism	H	
Bkir_c132_1782	Uroporphyrinogen-III methyltransferase	Coenzyme metabolism	H	
Bkir_c168_2458	Thiamine-monophosphate kinase	Coenzyme metabolism	H	
Bkir_c17_4538	Queuosine biosynthesis QueD, PTPS-I	Coenzyme metabolism	H	
Bkir_c18_6043	Dihydrofolate synthase /Folypolyglutamate synthase	Coenzyme metabolism	H	
Bkir_c2_4689	Biotin synthesis protein bioC	Coenzyme metabolism	H	
Bkir_c206_5744	GTP cyclohydrolase II	Coenzyme metabolism	H	
Bkir_c25_4128	2-amino-4-hydroxy-6-hydroxymethylidihydropteridinepyrophosphokinase	Coenzyme metabolism	H	
Bkir_c25_4130	3-methyl-2-oxobutanoatehydroxymethyltransferase	Coenzyme metabolism	H	
Bkir_c26_5810	Riboflavin kinase / FMNadenylyltransferase	Coenzyme metabolism	H	
Bkir_c31_1430	Thiamin-phosphate pyrophosphorylase	Coenzyme metabolism	H	
Bkir_c31_1432	Sulfur carrier protein ThiS	Coenzyme metabolism	H	

Bkir_c32_4016	Nicotinate-nucleotide adenylyltransferase, bacterial NadD family	Coenzyme metabolism	H	
Bkir_c46_0476	Porphobilinogen deaminase	Coenzyme metabolism	H	
Bkir_c46_0478	Homolog of E. coli HemY protein	Coenzyme metabolism	H	
Bkir_c47_5610	Thiamin biosynthesis protein ThiC	Coenzyme metabolism	H	
Bkir_c49_6062	Pyridoxine 5'-phosphate synthase	Coenzyme metabolism	H	
Bkir_c52_4917	Cobyric acid A,C-diamide synthase	Coenzyme metabolism	H	
Bkir_c52_4918	Cob(I)alamin adenosyltransferase	Coenzyme metabolism	H	
Bkir_c52_4924	CobN component of cobalt chelatase involved inB12 biosynthesis	Coenzyme metabolism	H	
Bkir_c52_4925	ChII component of cobalt chelatase involved inB12 biosynthesis	Coenzyme metabolism	H	
Bkir_c52_4926	ChID component of cobalt chelatase involved inB12 biosynthesis	Coenzyme metabolism	H	
Bkir_c52_4927	Cobalamin biosynthesis protein CbiG /Cobalt-precorrin-3b C17-methyltransferase	Coenzyme metabolism	H	
Bkir_c52_4928	Cobalt-precorrin-2 C20-methyltransferase	Coenzyme metabolism	H	
Bkir_c52_4929	Cobalt-precorrin-8x methylmutase	Coenzyme metabolism	H	
Bkir_c52_4931	Cobalt-precorrin-6y C5-methyltransferase / Cobalt-precorrin-6y C15-methyltransferase[decarb oxylating]	Coenzyme metabolism	H	
Bkir_c52_4933	Cobalt-precorrin-6x reductase	Coenzyme metabolism	H	
Bkir_c52_4934	Cobalt-precorrin-4 C11-methyltransferase	Coenzyme metabolism	H	
Bkir_c62_5025	Nicotinate-nucleotide--dimethylbenzimidazolephosphoribosyltransferase	Coenzyme metabolism	H	
Bkir_c62_5026	Cobalamin synthase	Coenzyme metabolism	H	
Bkir_c62_5031	Adenosylcobinamide-phosphate synthase	Coenzyme metabolism	H	
Bkir_c62_5032	Adenosylcobinamide-phosphate guanylyltransferase	Coenzyme metabolism	H	
Bkir_c62_5033	Cobyric acid synthase	Coenzyme metabolism	H	
Bkir_c71_4327	Uroporphyrinogen III decarboxylase	Coenzyme metabolism	H	
Bkir_c82_5071	Phosphomethylpyrimidine kinase	Coenzyme metabolism	H	
Bkir_c91_2758	Octaprenyl-diphosphate synthase /Dimethylallyltransferase /Geranyltransferase / Geranylgeranyl pyrophosphate synthetase	Coenzyme metabolism	H	
Bkir_c95_1623	Glutamate--cysteine ligase	Coenzyme metabolism	H	Glutathione metabolism
Bkir_c96_2129	Nicotinate phosphoribosyltransferase	Coenzyme metabolism	H	
Bkir_c154_3763	2-octaprenyl-6-methoxyphenol hydroxylase	Coenzyme metabolism / Energy production and conversion	HC	
Bkir_c91_2759	1-deoxy-D-xylulose 5-phosphate synthase	Coenzyme metabolism / Lipid metabolism	HI	
Bkir_c2_4673	Glutathione synthetase	Coenzyme metabolism / Translation, ribosomal structure and biogenesis	HJ	Glutathione metabolism

Bkir_c180_0805	ABC-type multidrug transport system, ATPase component	Defense mechanisms		
Bkir_c180_0806	ABC-type multidrug transport system, permease component	Defense mechanisms		
Bkir_c6_3056	Lipoprotein releasing system ATP-binding protein LoID	Defense mechanisms		
Bkir_c13_6201	Excinuclease ABC subunit C	DNA replication, recombination, and repair	L	
Bkir_c170_1733	Excinuclease ABC subunit B	DNA replication, recombination, and repair	L	
Bkir_c178_0365	DNA-binding protein HU	DNA replication, recombination, and repair	L	
Bkir_c25_4120	DNA mismatch repair protein MutL	DNA replication, recombination, and repair	L	
Bkir_c28_2971	Uracil-DNA glycosylase, family 1	DNA replication, recombination, and repair	L	
Bkir_c49_6063	DNA recombination and repair protein RecO	DNA replication, recombination, and repair	L	
Bkir_c6_3053	Single-stranded-DNA-specific exonuclease RecJ	DNA replication, recombination, and repair	L	
Bkir_c6_3096	DNA mismatch repair protein MutS	DNA replication, recombination, and repair	L	
Bkir_c67_4304	DNA primase	DNA replication, recombination, and repair	L	
Bkir_c91_2750	DNA polymerase I	DNA replication, recombination, and repair	L	
Bkir_c119_5128	Putative periplasmic cytochrome oxidoreductase type-C signal peptide protein	Energy production and conversion	C	
Bkir_c119_5129	Putative periplasmic cytochrome oxidoreductase type-C signal peptide protein	Energy production and conversion	C	
Bkir_c12_2347	Ferredoxin--NADP(+) reductase	Energy production and conversion	C	
Bkir_c15_5310	Rubredoxin-NAD(+) reductase	Energy production and conversion	C	
Bkir_c170_1737	Ferredoxin	Energy production and conversion	C	
Bkir_c172_3903	Cobalamin biosynthesis protein BluB 5,6-dimethylbenzimidazole synthase, flavin destructase family	Energy production and conversion	C	
Bkir_c2_4633	FIG00453717: hypothetical protein	Energy production and conversion	C	
Bkir_c2_4692	Cytochrome c oxidase polypeptide I	Energy production and conversion	C	
Bkir_c2_4695	Cytochrome c oxidase polypeptide III	Energy production and conversion	C	
Bkir_c203_4881	Citrate synthase (si)	Energy production and conversion	C	
Bkir_c25_4135	Cytochrome c4	Energy production and conversion	C	
Bkir_c25_4136	Putative diheme cytochrome c-553	Energy production and conversion	C	
Bkir_c33_3893	ATP synthase B chain	Energy production and conversion	C	
Bkir_c60_2293	Succinate dehydrogenase hydrophobic membrane anchor protein	Energy production and conversion	C	
Bkir_c60_2294	Succinate dehydrogenase cytochrome b-556 subunit	Energy production and conversion	C	
Bkir_c92_0577	Malate dehydrogenase	Energy production and conversion	C	
Bkir_c92_0586	Aconitate hydratase @2-methylisocitrate dehydratase	Energy production and conversion	C	
Bkir_c10_6264	Cob(I)alamin adenosyltransferase PduO	Function unknown	S	
Bkir_c103_2561	FIG000557: hypothetical protein co-occurring with RecR	Function unknown	S	

Bkir_c103_2565	Inner membrane protein	Function unknown	S	
Bkir_c118_3279	Sirohydrochlorin cobaltochelataase	Function unknown	S	
Bkir_c133_4185	Predicted membrane protein	Function unknown	S	
Bkir_c15_5396	COG1565: Uncharacterized conserved protein	Function unknown	S	
Bkir_c171_0648	FIG003276: zinc-binding protein	Function unknown	S	
Bkir_c171_0649	FIG002842: hypothetical protein	Function unknown	S	
Bkir_c2_4669	FIG00453990: hypothetical protein	Function unknown	S	
Bkir_c20_3829	UPF0125 protein yjfF	Function unknown	S	
Bkir_c203_4882	YgfY COG2938	Function unknown	S	
Bkir_c211_0165	Probable transmembrane protein	Function unknown	S	
Bkir_c211_0166	Probable transmembrane protein	Function unknown	S	
Bkir_c24_4070	putative lipoprotein	Function unknown	S	
Bkir_c34_2548	Cell division protein MraZ	Function unknown	S	
Bkir_c44_1556	Membrane protein	Function unknown	S	
Bkir_c45_4856	ABC transporter ATP-binding protein	Function unknown	S	
Bkir_c53_5706	Transporter	Function unknown	S	
Bkir_c53_5710	Protein YigP (COG3165) clustered with ubiquinone biosynthetic genes	Function unknown	S	
Bkir_c53_5721	NADPH dependent preQ0 reductase	Function unknown	S	
Bkir_c59_5280	Copper metallochaperone, bacterial analog of Cox17 protein	Function unknown	S	
Bkir_c6_3054	Regulatory protein, RpfE type	Function unknown	S	
Bkir_c66_2507	Putative cytoplasmic protein	Function unknown	S	
Bkir_c68_5575	UPF0434 protein YcaR	Function unknown	S	
Bkir_c78_2064	Believed to be involved in assembly of Fe-S clusters	Function unknown	S	
Bkir_c79_3389	LptA, protein essential for LPS transport across the periplasm	Function unknown	S	
Bkir_c9_5446	Serine/threonine kinase	Function unknown	S	
Bkir_c90_6087	Uncharacterized protein conserved in bacteria	Function unknown	S	
Bkir_c163_1964	undefined product	Function unknown	S	
Bkir_c28_2972	undefined product	Function unknown	S	
Bkir_c10_6231	Zinc metalloprotease	General function prediction only	R	
Bkir_c10_6260	FIG003879: Predicted amidohydrolase	General function prediction only	R	
Bkir_c11_4436	Probable component of the lipoprotein assembly complex (forms a complex with YaeT, YfgL, and NlpB)	General function prediction only	R	
Bkir_c12_2381	DNA-binding protein H-NS	General function prediction only	R	
Bkir_c15_5372	Membrane protein involved in the export of O-antigen and teichoic acid	General function prediction only	R	
Bkir_c15_5373	Glycosyl transferase, family 2	General function prediction only	R	
Bkir_c19_4791	GTP-binding protein EngA	General function prediction only	R	
Bkir_c19_4792	RNA-binding protein Hfq	General function prediction only	R	
Bkir_c191_3654	putative, Molybdenum cofactor biosynthesis protein	General function prediction only	R	

	A			
Bkir_c2_4698	Probable transmembrane protein	General function prediction only	R	
Bkir_c22_0121	ATPase YjeE, predicted to have essential role in cell wall biosynthesis	General function prediction only	R	
Bkir_c24_4064	FOG: Ankyrin repeat	General function prediction only	R	
Bkir_c3_2200	hypothetical protein	General function prediction only	R	
Bkir_c31_1434	ABC transporter, fused ATPase and inner membrane subunits	General function prediction only	R	
Bkir_c36_3530	Anti-anti-sigma regulatory factor (antagonist of anti-sigma factor)	General function prediction only	R	
Bkir_c41_0736	membrane protein, putative	General function prediction only	R	
Bkir_c44_1560	ATPase components of ABC transporters with duplicated ATPase domains	General function prediction only	R	
Bkir_c51_2613	GTP-binding protein EngB	General function prediction only	R	
Bkir_c52_4923	CobW GTPase involved in cobalt insertion for B12 biosynthesis	General function prediction only	R	
Bkir_c57_5558	Inosine-5'-monophosphate dehydrogenase	General function prediction only	R	
Bkir_c82_5074	Ribosomal-protein-S18p-alanine acetyltransferase	General function prediction only	R	
Bkir_c91_2745	Homoserine kinase	General function prediction only	R	
Bkir_c1_2880	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	Inorganic ion transport and metabolism	P	
Bkir_c1_2881	ABC-type anion transport system, duplicated permease component	Inorganic ion transport and metabolism	P	
Bkir_c113_0534	Zinc ABC transporter, inner membrane permease protein ZnuB	Inorganic ion transport and metabolism	P	
Bkir_c113_0535	Zinc ABC transporter, ATP-binding protein ZnuC	Inorganic ion transport and metabolism	P	
Bkir_c113_0539	Zinc uptake regulation protein ZUR	Inorganic ion transport and metabolism	P	
Bkir_c170_1734	Periplasmic protein p19 involved in high-affinity Fe ²⁺ transport	Inorganic ion transport and metabolism	P	
Bkir_c170_1736	High-affinity iron permease	Inorganic ion transport and metabolism	P	
Bkir_c205_4849	Kup system potassium uptake protein	Inorganic ion transport and metabolism	P	
Bkir_c52_4930	Cobalamin biosynthesis protein CobG	Inorganic ion transport and metabolism	P	
Bkir_c66_2508	Transporter, monovalent cation:proton antiporter-2 (CPA2) family	Inorganic ion transport and metabolism	P	
Bkir_c66_2522	Superoxide dismutase [Cu-Zn]	Inorganic ion transport and metabolism	P	superoxide radicals degradation
Bkir_c4_0313	undefined product	Intracellular trafficking and secretion		
Bkir_c123_5758	4-diphosphocytidyl-2-C-methyl-D- erythritolkinase	Lipid metabolism	I	
Bkir_c126_1575	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	Lipid metabolism	I	
Bkir_c18_6042	Acetyl-coenzyme A carboxyl transferase betachain	Lipid metabolism	I	
Bkir_c20_3815	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase	Lipid metabolism	I	
Bkir_c49_6061	Holo-[acyl-carrier protein] synthase	Lipid metabolism	I	
Bkir_c6_3045	Phosphatidylserine decarboxylase	Lipid metabolism	I	

Bkir_c86_5986	Phytoene synthase	Lipid metabolism	I	
Bkir_c18_6046	Amidophosphoribosyltransferase	Nucleotide transport and metabolism	F	
Bkir_c25_4125	Phosphoribosylformylglycinamide cyclo-ligase	Nucleotide transport and metabolism	F	
Bkir_c26_5809	Phosphoribosylglycinamide formyltransferase	Nucleotide transport and metabolism	F	
Bkir_c5_4991	Nucleoside 5-triphosphatase RdgB (dHATP, dITP,XTP-specific)	Nucleotide transport and metabolism	F	
Bkir_c58_0784	Phosphoribosylformylglycinamide synthase,synthetase subunit	Nucleotide transport and metabolism	F	
Bkir_c7_1463	Orotate phosphoribosyltransferase	Nucleotide transport and metabolism	F	
Bkir_c7_1487	Cytidylate kinase	Nucleotide transport and metabolism	F	
Bkir_c36_3514	FIG146285: Diadenosine tetraphosphate (Ap4A) hydrolase and other HIT family hydrolases	Nucleotide transport and metabolism / Carbohydrate transport and metabolism / General function prediction only	FGR	
Bkir_c98_2690	tRNA-specific adenosine-34 deaminase	Nucleotide transport and metabolism / Translation, ribosomal structure and biogenesis	FJ	
Bkir_c17_4537	Queuosine Biosynthesis QueE Radical SAM	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c183_3540	Peptide methionine sulfoxide reductase MsrB	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c20_3801	[Protein-PII] uridylyltransferase	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c26_5819	ATP-dependent Clp protease ATP-binding subunit ClpA	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c36_3510	Outer membrane stress sensor protease DegQ,serine protease	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c38_2426	Alkyl hydroperoxide reductase subunit C-like protein	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c46_0483	Protein-L-isoaspartate O-methyltransferase	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c51_2586	Periplasmic thiol:disulfide oxidoreductase DsbB, required for DsbA reoxidation	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c6_3094	FKBP-type peptidyl-prolyl cis-trans isomerase SlyD	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c67_4307	YgjD/Kae1/Qri7 family, required for threonylcarbamoyladenosine formation in tRNA	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c82_5075	Inactive homolog of metal-dependent proteases,putative molecular chaperone	Posttranslational modification, protein turnover, chaperones	O	
Bkir_c90_6110	Glutaredoxin-related protein	Posttranslational modification, protein turnover, chaperones	O	glutaredoxin pathway
Bkir_c17_4512	Thiol-disulfide isomerase and thioredoxins	Posttranslational modification, protein turnover, chaperones / Energy production and conversion	OC	thioredoxin pathway
Bkir_c2_4664	Thiol:disulfide interchange protein tlpA	Posttranslational modification, protein turnover, chaperones / Energy production and conversion	OC	
Bkir_c40_1365	Cytochrome c-type biogenesis protein DsbD,protein-disulfide reductase	Posttranslational modification, protein turnover, chaperones / Energy production and conversion	OC	
Bkir_c14_3225	3'-to-5' oligoribonuclease (orn)	RNA processing and modification		
Bkir_c12_2342	Protein involved in biosynthesis of mitomycin antibiotics/polyketide fumonisin	Secondary metabolites biosynthesis, transport, and catabolism	Q	

Bkir_c36_3531	Uncharacterized ABC transporter, component YrbC auxiliary	Secondary metabolites biosynthesis, transport, and catabolism	Q	
Bkir_c15_5397	FoIM Alternative dihydrofolate reductase 1	Secondary metabolites biosynthesis, transport, and catabolism / General function prediction only	QR	
Bkir_c1_2868	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	Signal transduction mechanisms	T	
Bkir_c1_2875	Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	Signal transduction mechanisms	T	
Bkir_c102_5934	Signal transduction histidine kinase	Signal transduction mechanisms	T	
Bkir_c11_4471	FIG00974056: hypothetical protein	Signal transduction mechanisms	T	
Bkir_c127_0391	Signal transduction histidine kinase	Signal transduction mechanisms	T	
Bkir_c183_3543	Cell division protein BolA	Signal transduction mechanisms	T	
Bkir_c44_1545	Predicted ATPase related to phosphate starvation-inducible protein PhoH	Signal transduction mechanisms	T	
Bkir_c49_6070	Sigma factor RpoE negative regulatory protein RseB precursor	Signal transduction mechanisms	T	
Bkir_c64_0432	Nitrogen regulation protein NR(I)	Signal transduction mechanisms	T	
Bkir_c64_0433	FOG: PAS/PAC domain	Signal transduction mechanisms	T	
Bkir_c70_5965	GTP-binding protein TypA/BipA	Signal transduction mechanisms	T	
Bkir_c102_5933	Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	Signal transduction mechanisms / Transcription	TK	
Bkir_c126_1568	Two-component system response regulator OmpR	Signal transduction mechanisms / Transcription	TK	
Bkir_c1_2859	RNA polymerase sigma-54 factor RpoN	Transcription	K	
Bkir_c10_6280	Hydrogen peroxide-inducible genes activator	Transcription	K	
Bkir_c2_4636	Pantothenate kinase type III, CoaX-like	Transcription	K	
Bkir_c40_1409	DNA-directed RNA polymerase beta subunit	Transcription	K	
Bkir_c41_0725	Transcription termination protein NusB	Transcription	K	
Bkir_c49_6072	RNA polymerase sigma factor RpoE	Transcription	K	
Bkir_c6_3040	DNA-directed RNA polymerase specialized sigma subunit, sigma24-like	Transcription	K	
Bkir_c9_5407	3'-to-5' exoribonuclease RNase R	Transcription	K	
Bkir_c11_4435	DinG family ATP-dependent helicase YoaA	Transcription / DNA replication, recombination, and repair	KL	
Bkir_c37_3337	DNA-binding protein Fis	Transcription / DNA replication, recombination, and repair	KL	
Bkir_c10_6282	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	Translation, ribosomal structure and biogenesis	J	
Bkir_c105_1121	Putative translation initiation inhibitor, yjgF family	Translation, ribosomal structure and biogenesis	J	
Bkir_c14_3234	SSU ribosomal protein S16p	Translation, ribosomal structure and biogenesis	J	
Bkir_c160_5733	unnamed protein product	Translation, ribosomal structure and biogenesis	J	
Bkir_c18_6038	unnamed protein product	Translation, ribosomal structure and biogenesis	J	
Bkir_c22_0069	YrdC/Sua5 family protein, required for	Translation, ribosomal structure and biogenesis	J	

	threonylcarbamoyladenine (t(6)A) formation in tRNA			
Bkir_c3_2230	Translation initiation factor 1	Translation, ribosomal structure and biogenesis	J	
Bkir_c32_4012	Cytoplasmic axial filament protein CafA and Ribonuclease G	Translation, ribosomal structure and biogenesis	J	
Bkir_c40_1377	LSU ribosomal protein L36p	Translation, ribosomal structure and biogenesis	J	
Bkir_c40_1378	Translation initiation factor 1	Translation, ribosomal structure and biogenesis	J	
Bkir_c40_1388	LSU ribosomal protein L24p (L26e)	Translation, ribosomal structure and biogenesis	J	
Bkir_c57_5535	LSU ribosomal protein L35p	Translation, ribosomal structure and biogenesis	J	
Bkir_c57_5536	Translation initiation factor 3	Translation, ribosomal structure and biogenesis	J	
Bkir_c6_3085	Tryptophanyl-tRNA synthetase	Translation, ribosomal structure and biogenesis	J	
Bkir_c68_5583	SSU ribosomal protein S20p	Translation, ribosomal structure and biogenesis	J	
Bkir_c79_3385	Ribosome hibernation protein YhbH	Translation, ribosomal structure and biogenesis	J	
Bkir_c8_3439	LSU ribosomal protein L21p	Translation, ribosomal structure and biogenesis	J	
Bkir_c88_6120	Ribonuclease P protein component	Translation, ribosomal structure and biogenesis	J	
Bkir_c4_0290	Enoyl-[acyl-carrier-protein] reductase [NADH]	acyl-carrier-protein		
Bkir_c11_4433	Membrane protein involved in colicin uptake	no assignment		
Bkir_c118_3294	Chromosome segregation ATPases	no assignment		
Bkir_c121_2677	FIG01282171: hypothetical protein	no assignment		
Bkir_c13_6185	FIG00453615: hypothetical protein	no assignment		
Bkir_c14_3236	Probable transmembrane protein	no assignment		
Bkir_c170_1735	putative exported protein	no assignment		
Bkir_c181_1019	Nucleoside-diphosphate-sugar epimerases	no assignment		
Bkir_c2_4617	FIG00452673: hypothetical protein	no assignment		
Bkir_c2_4696	Probable transmembrane protein	no assignment		
Bkir_c211_0167	hypothetical protein	no assignment		
Bkir_c25_4143	FIG00453399: hypothetical protein	no assignment		
Bkir_c3_2202	COG0489: ATPases involved in chromosome partitioning	no assignment		
Bkir_c4_0328	Probable transmembrane protein	no assignment		
Bkir_c4_0355	FIG00453072: hypothetical protein	no assignment		
Bkir_c44_1552	Permease of the drug/metabolite transporter(DMT) superfamily	no assignment		
Bkir_c47_5652	Histone acetyltransferase HPA2 and related acetyltransferases	no assignment		
Bkir_c51_2595	Probable transmembrane protein	no assignment		
Bkir_c52_4915	hypothetical protein	no assignment		
Bkir_c52_4916	Predicted cobalt transporter CbtA	no assignment		
Bkir_c66_2513	FIG00454446: hypothetical protein	no assignment		
Bkir_c7_1493	Probable signal peptide protein	no assignment		

Bkir_c78_2075	hypothetical protein	no assignment		
Bkir_c82_5068	Probable transmembrane protein	no assignment		
Bkir_c9_5512	Permeases of the major facilitator superfamily	no assignment		
Bkir_c92_0584	Probable signal peptide protein	no assignment		
Bkir_c95_1644	General secretion pathway protein C	no assignment		
Bkir_c9_5513	undefined product	no assignment		

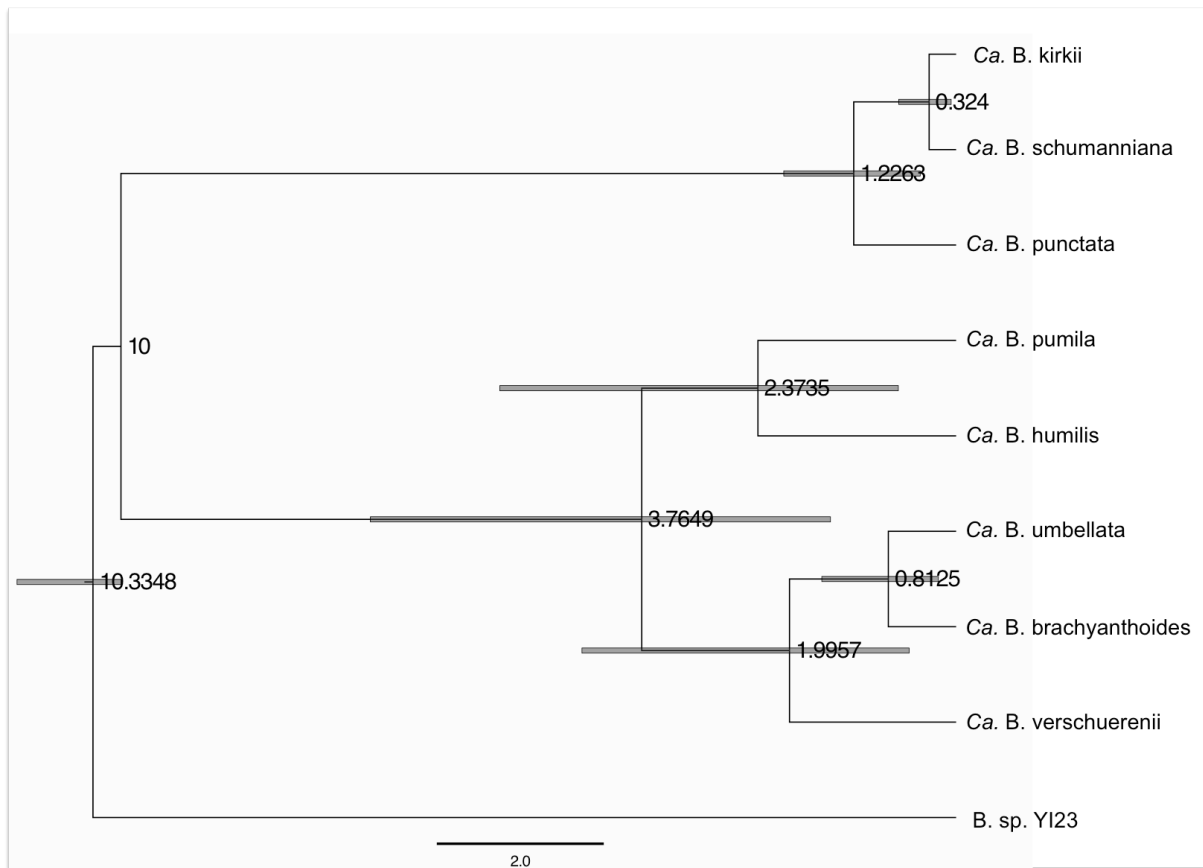


Fig. S5. Bayesian divergence time estimation of leaf nodule symbiosis in *Burkholderia* leaf nodule symbionts. Maximum likelihood phylogenetic reconstruction of the bacterial leaf nodule symbionts obtained with BEAST (2.2.0). The mean time ages estimated in Mya are specified in each node. Confidence intervals (95%) are highlighted in grey.

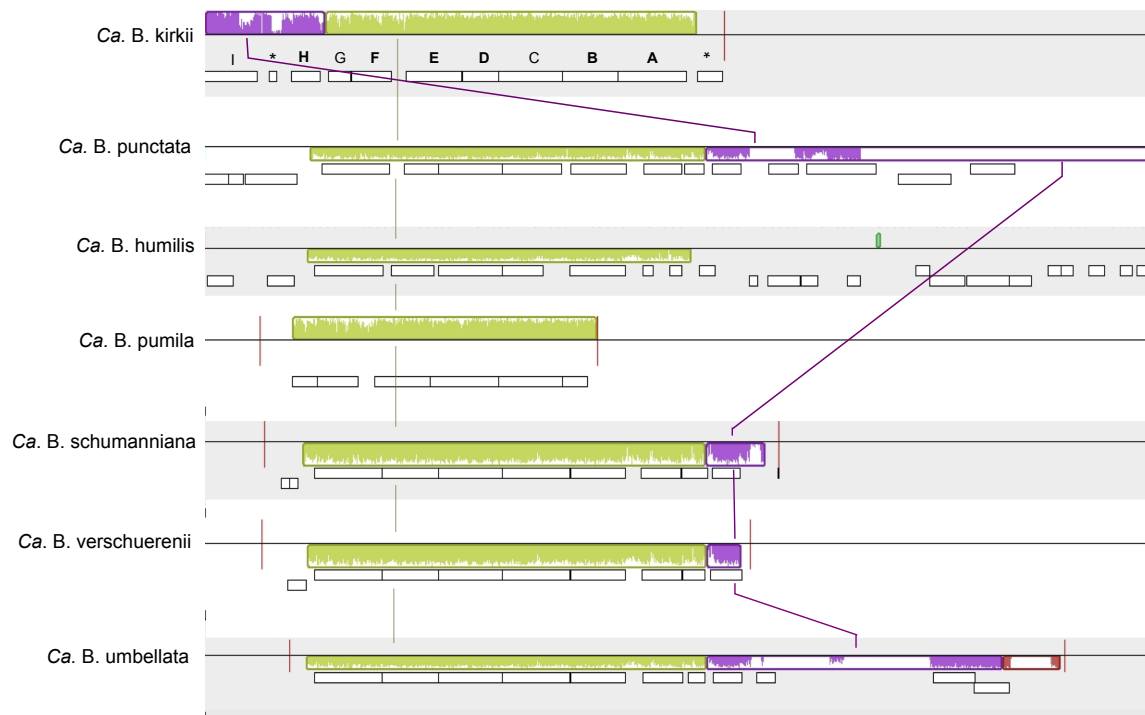


Fig. S6. Mauve alignment of gene clusters involved in secondary metabolite production in leaf nodule symbionts (*Ca. B. brachyanthoides* not shown). Horizontal lines represent the genomic region for each genome including the annotated coding regions (white boxes). *Ca. B. kirkii* is used as a reference. The annotation of the boxes corresponds to (A) 2-epi-5-epi-valiolone synthase Bkir_c149_4879, (B) N-acetylmannosamine kinase Bkir_c149_4878, (C) Aminosugar transaminase Bkir_c149_4877, (D) Glycoside hydrolase Bkir_c149_4876, (E) Sugar dehydrogenase Bkir_c149_4875, (F) HAD family hydrolase Bkir_c149_4874, (G) Sugar phosphate isomerase Bkir_c149_4873, (H) acetyltransferase, GNAT family Bkir_c149_4872. The genes conserved among all the leaf nodule symbionts are shown in bold. (*) represent hypothetical proteins. The colored boxes represent a similarity plot, the height of which is proportional to the level of sequence identity in that region.

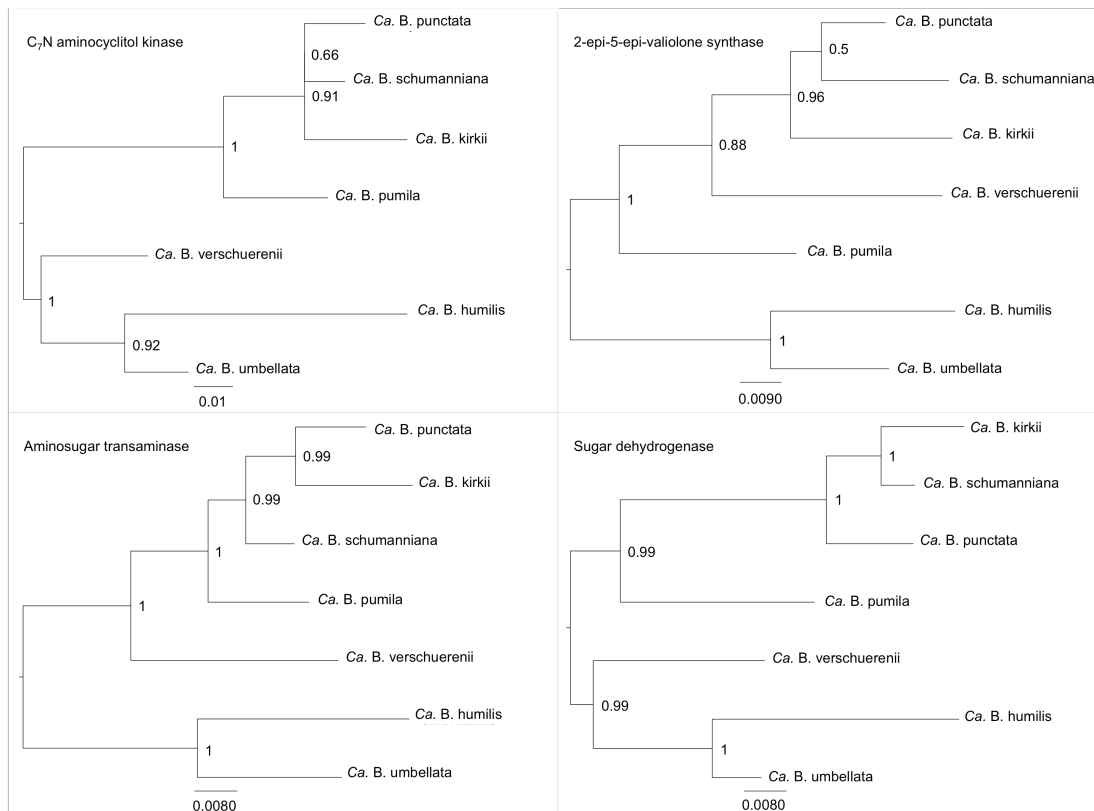


Fig. S7. Maximum parsimony phylogenetic reconstruction of the secondary metabolites genes conserved in all leaf nodule symbionts. Bootstrap values are shown as proportions (scale from 0 to 1). Homologous genes from *Pseudomonas* sp. CFT9 (2-epi-5-epi-valiolone synthase, c7-cyclitol-7-kinase) and *Pseudomonas fluorescens* A506 (putative dehydrogenase, aminotransferase) was used as outgroup (not shown).

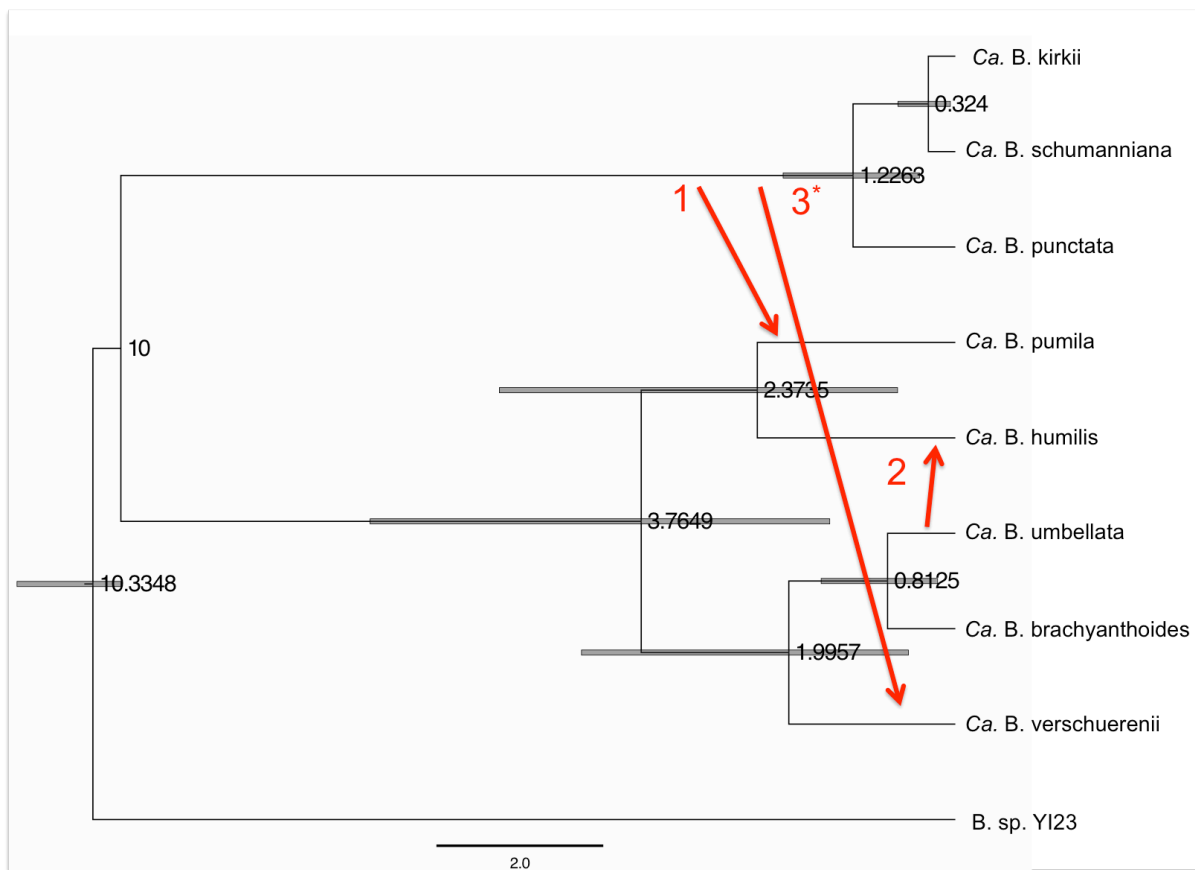


Fig. S8. Proposed horizontal gene transfer events affecting the genes involved in the biosynthesis of the secondary metabolites of the C7N aminocyclitol family. Arrows indicate the hypothetical direction of horizontal gene transfer events, inferred from the phylogenetic analysis presented on Fig. S7. Asterisk denotes a HGT event that affected only the 2-epi-5-epi-valiolone synthase gene.

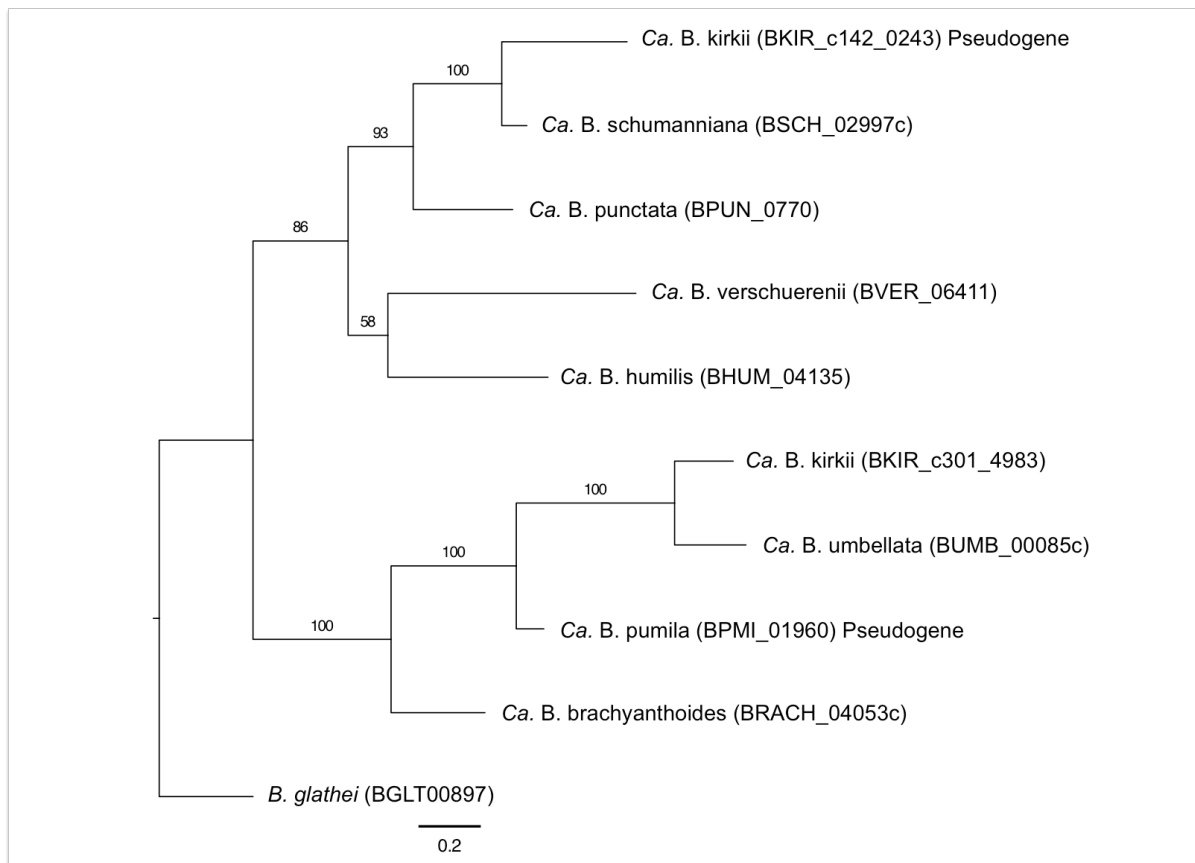


Fig. S9. Maximum likelihood phylogenetic reconstruction of the replication protein A conserved in most of the leaf nodule symbionts. The alignment also included pseudogene copies of the genes where indicated. Large gaps and poorly aligned regions were trimmed.