Supplementary Material S2

Genomic analysis of *Kiritimatiella glycovorans* L21-Fru-AB^T

1. Genome structure and evolution

The coding density of the genome is around 90% and of the 2,478 genes predicted, 2,418 were identified protein-coding genes and 60 RNAs including one complete ribosomal RNA operon (16S-23S-5S), 44 tRNAs and 13 RNAs of partly unknown functions. According to the annotation provided by the IMG/ER server around 80% of the protein-coding genes were assigned a putative function while the remaining ones were annotated as hypothetical proteins. More than 60% of protein-coding genes identified by the IMG/ER system could be assigned to COG functional categories (Figure S1a).

The number of 75 putative transposase genes detected in the L21-Fru-AB^T genome (according to the IMG/ER server) corresponds to 25.42 elements per Mb of the genome, which is relatively high compared to most other genomes of free-living species of *Verrucomicrobia*, which on an average have only 8.91 mobile elements/Mb (19 genomes analysed).

This could indicate that mobile elements represent a major driving force in the evolution of this strain. The observed high transposable element load could be the result of a population bottleneck associated with the colonization of the anoxic zone of the Kiritimati hypersaline microbial mat by an ancestor of strain L21-Fru-AB^T, because this niche represents a selective, relatively stable and isolated ecosystem. Hence, adaptation and specialization to this habitat may lead to similar genomic changes as observed in bacteria following host restriction (Moran and Plague, 2004). In addition, genome evolution by the rearrangement of mobile DNA elements or integration of foreign DNA is indicated by the detection of twelve tentative genomic islands as illustrated in Fig. S1b. With only one exception the identified genomic regions contained or were flanked by at least one transposase gene.

Stability and integrity of the L21-Fru-AB^T genome are protected by two sequence-directed defence mechanisms: host-controlled modification-restriction systems and arrays of clustered, regularly interspaced short palindromic repeats (CRISPR). Genes encoding subunits M (L21SP4_00661, L21SP4_01535), S (L21SP4_00662, L21SP4_01535) and R (L21SP4_00664, L21SP4_01534) of two distinct type I restriction complexes, a type III

restriction modification complex (L21SP4_00684, L21SP4_00685) and a putative type IV modification-dependent restriction system (L21SP4_00679 and L21SP4_00680) were detected. Three different recognition sequence motifs for methylation were identified by PacBio SMRT sequencing: CT**A**NNNNNNNTGC/GC**A**NNNNNNNTAG, **A**GCNNNNNGCT, and CAAGA**A**. In each case the modified base was identified as 6-methyladenosine (underlined).

The detection of two CRISPR loci within the genome sequence of L21-Fru-AB^T could indicate a frequent contact with phages, because it is thought that these genomic regions encode functions for the prevention of infection with foreign DNA (Marraffini and Sontheimer, 2010). However, no large prophage regions were found in a PHAST analysis (Zhou *et al.*, 2011). The identified CRISPR loci encode either an associated endonuclease of subtype I-E (L21SP4_02074) or subtype II (L21SP4_02005). In addition, several associated Cas proteins flanking the characteristic sequence repeat array were present. Both CRISPR regions are located in putative genomic islands adjacent to one or two transposase genes, which may indicate that they are mobile elements or were acquired by horizontal gene transfer.

Several phenotypic or molecular traits of strain L21-Fru-AB^T were deduced based on manually annotated genes listed in Supplementary Table S2. The significance of the annotated genes is explained below.

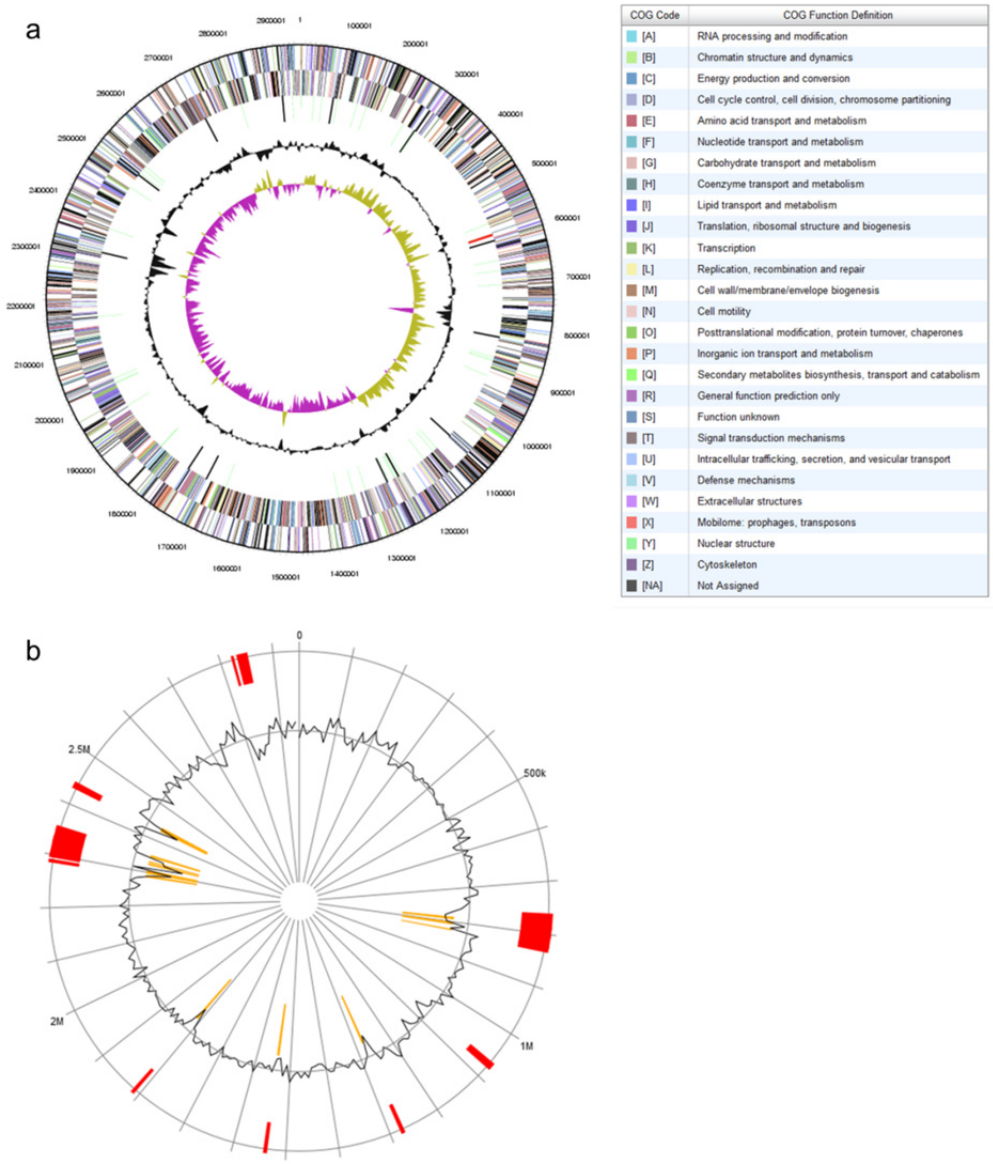


Figure S1. Circular maps of the chromosome of *K. glycovorans* L21-Fru-AB^{T}.

(a) Summary of genome features provided by the IMG/ER server. It is shown from outside to the centre: genes on forward strand (colour by COG categories), genes on reverse strand (colour by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, and GC skew. The colour code of function category for the top COG hit is shown on the right panel. **(b)** Predicted genomic islands mapped with IslandViewer 3. The orange centripetal lines indicate sites of putative horizontal origin predicted by the SIGI-HMM method, whereas the red blocks include also regions identified with the IslandPath-DIMOB algorithm. The black circular line graph symbolizes local deviations from the average GC value.

2. Cell biology

A large number of genes involved in pathways for the synthesis of lipopolysaccharides and peptidoglycan were identified in the L21-Fru-AB^T genome by the subsystems-based annotation tool provided by the RAST server (http://rast.nmpdr.org/). Examples of encoded enzymes involved in lipopolysaccharide synthesis are lipid-A-disaccharide synthase (L21SP4_02172), lipopolysaccharide core heptosyltransferase (L21SP4_01206) and lipid A biosynthesis lauroyl acyltransferase (L21SP4_01194). In addition, a cluster of genes located at L21SP4_00626 – 00637 seems to be involved in the production and export of lipopolysaccharides. Genes of key enzymes of peptidoglycan synthesis include D-alanine--D-alanine ligase (L21SP4_00463), UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase (L21SP4_01761), and peptidoglycan lipid II flippase (L21SP4_01757). The assumption of a typical Gram-negative cell plan is also supported by the presence of several predicted outer membrane and periplasmic proteins including TolB (L21SP4_01870) and Pal (L21SP4_01871), which anchor the outer membrane to the peptidoglycan layer (Jacquier *et al.*, 2015).

In strain L21-Fru-AB^T septum formation and binary fission seem to rely on the tubulin related protein FtsZ encoded at L21SP4_00466, which is a central component of the cell division complex in most bacteria. This represents a difference to chlamydia and planctomycetes, which use an alternative protein complex for cell division that lacks FtsZ (Pilhofer *et al.*, 2013; Jeske *et al.*, 2015). In addition, so far no *ftsZ* gene has been annotated in members of the *Lentisphaerae* phylum. However, the genome of *Lentisphaera araneosa* HTCC2155^T seems to encode two predicted proteins with a tubulin/FtsZ domain (LNTAR_12841 and LNTAR_23419), but it is unclear if they have the same function as FtsZ.

Several genes could be detected that are involved in the accumulation of the reserve polymers polyphosphate and glycogen. A gene encoding a polyphosphate kinase (L21SP4_01813) is probably involved in the intracellular accumulation of polyphosphates, whereas an exopolyphosphatase (L21SP4_01814) is used for the mobilization of phosphate and the generation of energy under starvation conditions. All necessary genes involved in the metabolism of glycogen as reserve material were detected in the genome of L21-Fru-AB^T. The accumulation of glycogen depends on glucose-1-phosphate adenylyltransferase (L21SP4_01507), glycogen synthase (L21SP4_00493 and L21SP4_01037) and a glycogen

branching enzyme (L21SP4_00494), whereas genes encoding a glycogen phosphorylase (L21SP4_00492 and L21SP4_01270) and a glycogen debranching enzyme (L21SP4_00491) are probably involved in the release of glucose-1-phosphate monomers from stored glycogen. Glycogen as reserve material was also detected in the verrucomicrobial methanotroph „*Candidatus* Methylacidiphilum fumariolicum“. In this bacterium accumulation of glycogen was induced by the depletion of a nitrogen source causing cessation of growth. During the subsequent starvation phase the stored glycogen was consumed thereby preventing loss of viability (Khadem *et al.*, 2012).

3. Nutrient uptake

Ammonium is probably transported into the cell by the AmtB channel proteins, which exist in two different isoforms in L21-Fru-AB^T encoded at L21SP4_01391 and L21SP4_01503. The primary intracellular acceptor for ammonium can be either 2-ketoglutarate which is converted to glutamate by glutamate dehydrogenase (L21SP4_00568) or glutamate which is converted to glutamine by the enzyme glutamine synthetase (L21SP4_01522). In hypersaline microbial mats both sulfate and reduced sources of sulfur are present in large amounts. Although, in this strain uptake and activation of sulfate is possible by a sulfate permease (L21SP4_02317) and a bifunctional sulfate adenylyltransferase/adenylylsulfate kinase (L21SP4_02316) further genes encoding enzymes of an assimilatory sulfate reduction pathway were missing, so that probably mainly reduced sources of sulfur, *e.g.* cysteine, are assimilated. In contrast to sulfate, phosphate concentrations are in the growth limiting range in most marine environments. Strain L21-Fru-AB^T seems to use at least two types of transporters for this essential nutrient: Either an inducible ATP-binding cassette (ABC)-type phosphate uptake system that is driven by the hydrolysis of ATP and encoded in a tentative operon (L21SP4_00290 - 00294) or a Na^+/phosphate symporter encoded at L21SP4_01162 that could be used for passive transport at higher phosphate concentrations. For the uptake of iron, which represents a major limiting factor for growth in many saline environments, strain L21-Fru-AB^T can utilize a range of TonB dependent transporters that selectively bind chelated ferric iron. All components of a ferric siderophore uptake system including an outer membrane receptor, an outer membrane transport energization complex, the periplasmic binding protein and the ABC-transporter are for example encoded in a putative operon at L21SP4_01108 - 01115. Alternatively, ATP- or GTP-dependent permeases for the translocation of unbound ferrous iron across the cytoplasmic membrane (*e.g.*, L21SP4_01608 - 01610) may be used. Furthermore, genes

encoding active uptake systems for the essential trace metals zinc (L21SP4_00297 – 00299), cobalt (L21SP4_01892, L21SP4_01897/01898) and copper (L21SP4_00567) were detected.

For the translocation of sugars into the cytoplasm probably mainly transporters are used that belong to the major facilitator superfamily including sugar porters with unknown specificity (*e.g.*, L21SP4_00133), several putative Na^+/melibiose symporter (*e.g.*, L21SP4_00407), and a fucose permease (L21SP4_00607). Furthermore, a predicted L-rhamnose-proton symporter (L21SP4_00931) belonging to the sugar-transport superfamily was detected. With the exception of one putative ABC-type transporter probably specific for N-acetylglucosamine (L21SP4_00327/00328; L21SP4_01748), no other genes encoding ATP-binding transporters for the uptake of monosaccharides form the environment were identified, which may indicate that substrate concentrations in the mat are not in the limiting range so that there is no need to use ATP for the highly specific active uptake of sugars.

Precursor metabolites may be imported into the cell from the environment using a peptide-specific ABC-type transporter (L21SP4_00561 – 00563 and L21SP4_02164/02165), two putative amino acid permeases (L21SP4_00756 and L21SP4_01661) and a tripartite ATP-independent periplasmic (TRAP)-type C4-dicarboxylate transporter (L21SP4_02424 - 02426), which offers an explanation for the stimulation of growth by supplementing cultivation media with yeast extract or peptone.

4. Secretion and resistance

Unfolded proteins may be exported by the Sec preprotein-translocation complex, which is encoded by several genes dispersed on the genome (see Supplementary Table S2). For the cotranslational Sec pathway the RNA containing signal recognition particle (SRP, L21SP4_00102) binds the signal peptide of the nascent protein chain and guides the ribosome to the SRP receptor FtsY (L21SP4_01780) of the Sec transmembrane complex. In the SecA (L21SP4_00353) dependent pathway the preprotein is completely synthesized in the cytoplasm before translocation across the cytoplasmic membrane. However, no gene encoding the SecB chaperone that is usually involved in this pathway and guides the preprotein to SecA could be identified. For the translocation of folded proteins across the cytoplasmic membrane the twin-arginine translocation (TAT) pathway is used, which is encoded by the genes *tatBCA* located in a tentative operon at L21SP4_01749 – 01751. A type II secretion system represents probably the main route for proteins transported across the

outer membrane. Most of the genes encoding the transmembrane secretion complex (*gpsFGHIJKLMD*) are located in a putative operon at L21SP4_00726 – 00734. In addition, genes encoding a cytoplasmic secretion ATPase (GpsE) were detected at several sites on the genome (*e.g.*, L21SP4_01644), but a gene for the inner membrane platform protein GspC that interacts with the outer membrane secretin GspD was lacking. It is worth mentioning that the *gspC* gene was also not present in other genomes of members of the PVC superphylum, so that in these bacteria the composition of the secretion complex may be different to the typical structure found in gammaproteobacteria (Korotkov *et al.*, 2012). Some proteins may be alternatively transported by the type V or autotransporter secretion system, which also depends on the Sec preprotein-translocation pathway. At least four predicted proteins contain an autotransporter-associated beta strand repeat domain forming a secretion pore in the outer membrane. In addition, the genes *bamA* (L21SP4_01180) and *bamD* (L21SP4_00777) representing two essential components of the Bam complex that assists in the insertion of autotransporter proteins in the outer membrane were identified.

Exopolysaccharides are produced and secreted in strain L21-Fru-AB^T likely via the Wzx/Wzy-dependent pathway (Schmid *et al.*, 2015). This biosynthesis pathway starts with the transfer of the first sugar unit to a phosphorylated polyprenyl lipid anchor at the inner membrane by a priming glycosyl transferase (*e.g.*, L21SP4_02279). Genes encoding a range of predicted glycosyltransferases (36), epimerases (14), and sulfotransferases (12) are present in the genome sequence according to the IMG/ER server and probably involved in the subsequent synthesis and decoration of the repeat units at the cytoplasmic face of the inner membrane before their transfer to the periplasmic space via a flippase and subsequent polymerization and export. Most of the genes involved in the polymerization and export of the produced exopolysaccharides seem to be organized in a putative operon located at L21SP4_00190 – 00202.

Resistance against antibiotics is probably based on several TolC dependent multidrug-resistance efflux pumps encoded for example at L21SP4_00769 – 00771 and at least one putative beta-lactamase (L21SP4_00553). In addition, resistance to toxic concentrations of arsenic is conferred by an arsenate reductase and efflux permease encoded in a tentative operon at L21SP4_00571 – 00574.

5. Response to environmental stress

According to the genome sequence the mechanisms for the detoxification of oxygen in strain L21-Fru-AB^T should be quite similar to the defence system proposed to operate in the aerotolerant spirochaete *Salinispira pacifica* L21-RPul-D2^T (Ben Hania *et al.*, 2015), which was isolated from the same habitat. In both species, oxygen is probably reduced by a cascade of reactions involving an unknown NADH oxidase, superoxide dismutase (L21SP4_02072), rubredoxin (L21SP4_00226), and rubrerythrin (L21SP4_00882). Interestingly, genes of a membrane-bound electron-transport complex were detected and annotated as Na^+-translocating NADH-quinone reductase (L21SP4_00578 - 00583), an enzyme which is usually only found in respiratory bacteria containing lipoquinones as electron acceptor for NADH oxidation. The function of this complex in an obligately fermentative bacterium is unknown, but it may be speculated that it is involved in an electron transport chain leading to the reduction of oxygen (Fig. 6). Small amounts of H_2O_2 may be removed by the enzyme alkylhydroperoxid reductase (L21SP4_01429) or by reaction with coenzyme A, which is assumed to represent the main thiol-redox buffer in this strain and could be reduced by a coenzyme A reductase (L21SP4_00052 and L21SP4_02384). Oxidative damage of lipids or proteins is probably repaired by the predicted enzymes peroxiredoxin (L21SP4_01626), alkylhydroperoxidase (L21SP4_00050), and methionine sulfoxide reductase A (L21SP4_01628).

Adaptation to various salt concentrations is likely mediated by the osmoprotectants glycine betaine and sucrose. Glycine betaine can be either synthesized from glycine by subsequent action of two methyltransferases encoded adjacently on the genome (L21SP4_02069 and L21SP4_02070) or is directly taken up from the environment by a Na^+/glycine betaine-symporter (L21SP4_00520) and an ABC-type transporter (L21SP4_02155 – 02157). In addition, genes of the key enzymes sucrose-phosphate synthase (L21SP4_02013) and sucrose-6F-phosphate phosphohydrolase (L21SP4_00973) were detected in the genome of strain L21-Fru-AB^T. The hydrolytic enzymes sucrose synthase (L21SP4_02012) and sucrose phosphorylase (L21SP4_02131) may function primarily in the control of the intracellular sucrose concentration in response to osmotic stress, because sucrose was not utilized as substrate in this strain.

In addition to osmolytes monovalent cation:proton antiporters have an important function in the regulation of intracellular pH, sodium homeostasis and cell volume at changing salinities and pH values. Genes encoding putative Na^+/H^+ exchangers were encoded at L21SP4_00167,

L21SP4_00232, L21SP4_00348, and L21SP4_02045. All of these genes were located adjacently to genes encoding a PTS IIA domain, which may indicate that phosphorylation plays a role in the regulation of the activity of these proteins (see below).

6. Central metabolic pathways

Several bioactive glycopolymers produced in microbial biofilms contain rhamnose as essential component. Strain L21-Fru-AB^T encodes a complete pathway for the degradation of rhamnose including the enzymes L-rhamnose mutarotase (L21SP4_02170), L-rhamnose isomerase (L21SP4_00462), rhamnulokinase (L21SP4_01498), and rhamnose utilisation protein (L21SP4_01948). These genes are not organized in an operon, but show a scattered distribution in the genome, which could indicate that they are constitutively expressed and essential for the metabolism of this strain. However, growth of strain L21-Fru-AB^T with L-rhamnose as sole carbon and energy source was not stable in laboratory experiments.

In contrast to rhamnose, glucose was a reliable substrate for growth under laboratory conditions. After the uptake of glucose into the cytoplasm it has to be phosphorylated to be further metabolized. No gene for a canonical hexokinase could be identified, but a putative glucokinase of the ROK family (L21SP4_02182) may activate glucose for further degradation. The following enzymes of the Embden-Meyerhof-Parnas pathway were encoded in the genome: glucose-6-phosphate isomerase (L21SP4_00854), 6-phosphofructokinase (L21SP4_01271), pyrophosphate--fructose 6-phosphate 1-phosphotransferase (L21SP4_01747), fructose-bisphosphate aldolase (L21SP4_00013), triosephosphate isomerase (L21SP4_00391), glyceraldehyde-3-phosphate dehydrogenase (L21SP4_00485), phosphoglycerate kinase (L21SP4_00558), phosphoglycerate mutase (L21SP4_01985), enolase (L21SP4_02068), and pyruvate phosphate dikinase (L21SP4_01945). In addition, an adenylate kinase (L21SP4_01695) could play a role in the regeneration of AMP required for the production of ATP and pyruvate from phosphoenolpyruvate catalysed by pyruvate phosphate dikinase (Fig. S2).

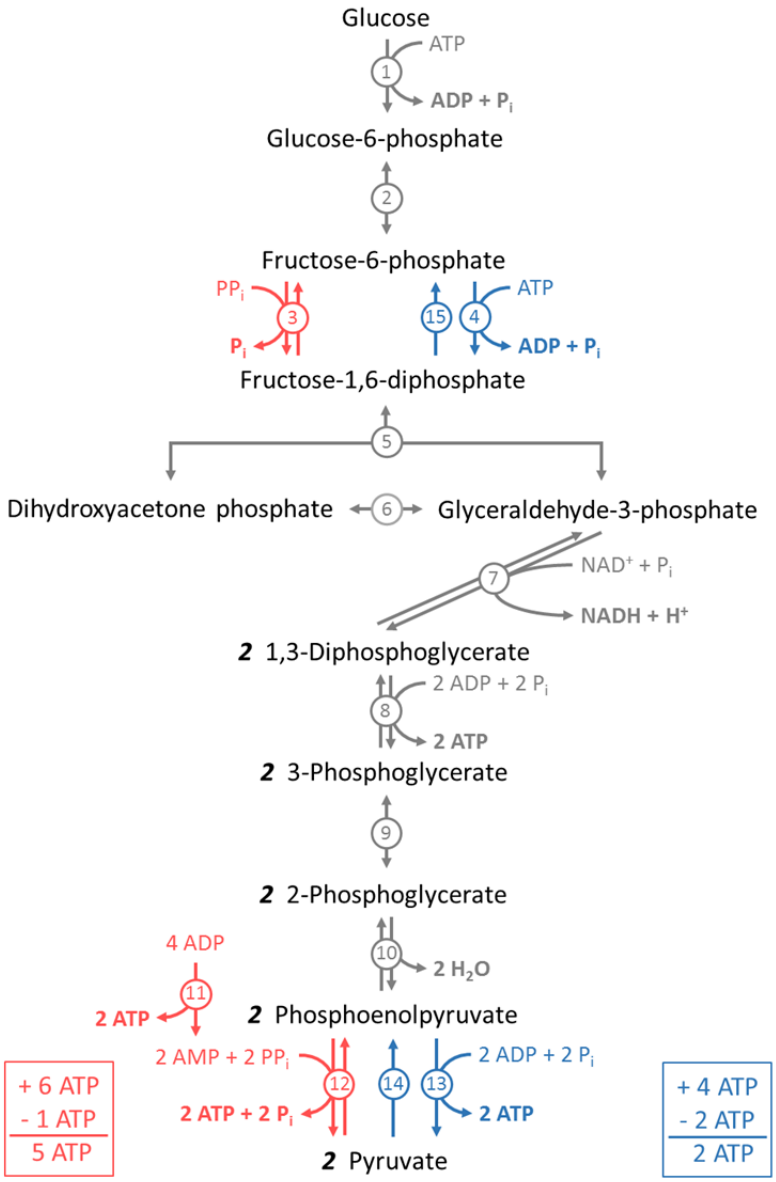


Figure S2. Proposed sequence of the Embden-Meyerhof-Parnas pathway in *K. glycovorans* L21-Fru-AB^T. Enzymes specific for the proposed pathway in L21-Fru-AB^T are labelled in red, while alternative enzymes of the canonical pathway, *e.g.* used in *Escherichia coli*, are in blue. Numbers denote the following enzymes: 1, glucokinase or hexokinase; 2, glucose-6-phosphate isomerase; 3, pyrophosphate--fructose 6-phosphate 1-phosphotransferase; 4, 6-phosphofructokinase; 5, fructose-bisphosphate aldolase; 6, triosephosphate isomerase; 7, glyceraldehyde-3-phosphate dehydrogenase; 8, phosphoglycerate kinase; 9, 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase; 10, enolase; 11, adenylate kinase; 12, pyruvate phosphate dikinase; 13, pyruvate kinase; 14, phosphoenolpyruvate synthase; 15, fructose-1,6-bisphosphatase.

In strain L21-Fru-AB^T the decarboxylation of pyruvate to acetyl-CoA is either catalysed by an oxygen tolerant pyruvate dehydrogenase complex (L21SP4_01431 - 01434) or the oxygen sensitive enzymes pyruvate:ferredoxin oxidoreductase (L21SP4_00338) and pyruvate:formate lyase (L21SP4_00415), which is located adjacent to its activating enzyme (L21SP4_00414). During fermentation formate was apparently not produced in significant amounts (below the detection limit of the used chemical assay) and no gene encoding a formate dehydrogenase was identified, so that formate produced by the enzyme pyruvate:formate lyase may be mainly funnelled into biosynthetic pathways by a formate--tetrahydrofolate ligase (L21SP4_02432).

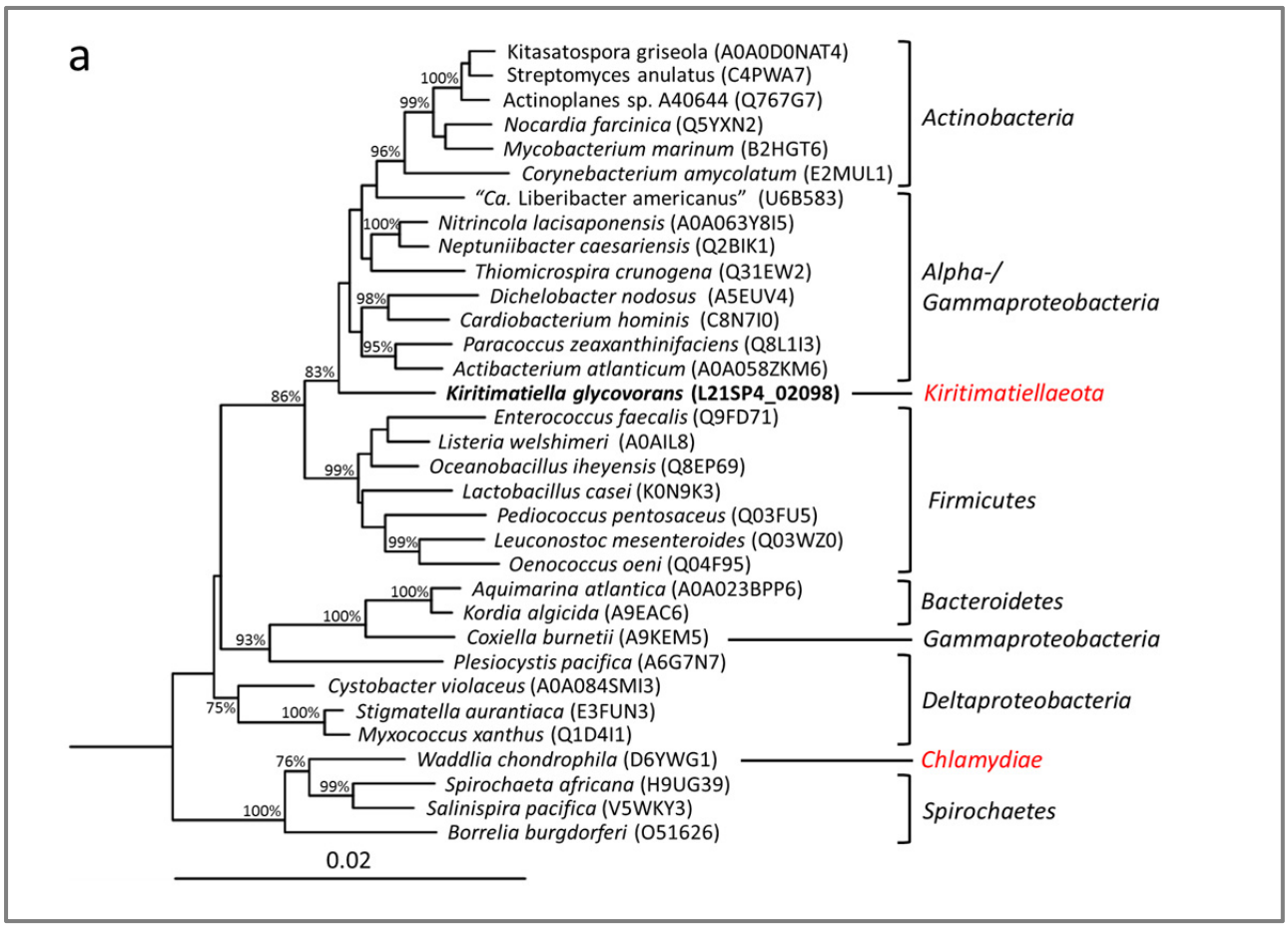
All genes encoding enzymes of known fermentation pathways leading to the products ethanol, acetate and lactate could be detected, including lactate dehydrogenase (L21SP4_02150), phosphate acetyltransferase (L21SP4_01446), acetate kinase (L21SP4_01445), and a combined aldehyde-alcohol dehydrogenase (L21SP4_02481). Hydrogen is probably produced during fermentation from ferredoxin and NADH by two distinct bifurcating cytoplasmic iron-only hydrogenases encoded at adjacent sites of the genome (L21SP4_01302 - 01308). The active hydrogenase complexes are probably trimeric and couple the exergonic release of hydrogen from reduced ferredoxin with the energetically unfavourable formation of hydrogen from NADH as described previously for the iron-only hydrogenase of *Thermotoga maritima* (Schut and Adams, 2009).

Besides substrate-level phosphorylation several membrane-bound enzyme complexes could be involved in the generation of energy using a chemiosmotic potential. A RNF electron transport complex encoded at L21SP4_01165 - 01170 is probably involved in generating a chemiosmotic gradient by using reduced ferredoxin as electron donor for NAD^+ reduction with concomitant translocation of protons or sodium ions across the cytoplasmic membrane as reported previously for the RNF complex of *Acetobacterium woodii* (Biegel *et al.*, 2011). The membrane potential can be used for the synthesis of ATP by a V-type ATP-synthase encoded in a tentative operon at L21SP4_01276 – 01282.

The required carbon skeletons and NADPH for most biosynthetic pathways can be supplied by glycolysis and the pentose phosphate pathway. Key enzymes of the pentose phosphate pathway encoded in the L21-Fru-AB^T genome were glucose-6-phosphate 1-dehydrogenase

(L21SP4_00987), 6-phosphogluconate dehydrogenase, decarboxylating (L21SP4_00648), transketolase (L21SP4_01787) and transaldolase (L21SP4_00829). In addition, oxaloacetate for the synthesis of the aspartate group of amino acids can be produced from pyruvate by the enzyme pyruvate carboxylase, which is split in two subunits that are encoded by genes located at different sites of the genome (L21SP4_00816 and L21SP4_00826). The large biotin carboxylase subunit (L21SP4_00816) can probably combine also with a biotin carboxyl carrier protein encoded adjacently at L21SP4_00815 and two subunits of carboxyltransferase (L21SP4_01026, L21SP4_01127) to form the acetyl-CoA carboxylase complex required for initiating fatty acid synthesis. The C5 precursor 2-ketoglutarate is required for the biosynthesis of the glutamate family of amino acids and produced from oxaloacetate by enzymes of the oxidative branch of the tricarboxylic acid cycle including citrate synthase (L21SP4_00845), aconitate hydratase (L21SP4_00761), and isocitrate dehydrogenase (L21SP4_00989). However, no genes encoding a 2-ketoglutarate dehydrogenase complex or enzymes of the reductive branch of the tricarboxylic acid cycle were identified, which is in line with the obligately fermentative metabolism of this strain.

Isopentenyl pyrophosphate is required for isoprenoid synthesis and synthesized in strain L21-Fru-AB^T via the mevalonate pathway. Several key enzymes involved in the mevalonate pathway including hydroxymethylglutaryl-CoA synthase, mevalonate kinase and diphosphomevalonate decarboxylase are encoded in a tentative operon located at L21SP4_02095 – L21SP4_02099. The gene of the enzyme mevalonate kinase appears to be duplicated with one paralog containing the N-terminal domain of the GHMP kinases superfamily (pfam00288) and the other the C-terminal domain of this superfamily (pfam08544), while usually both domains are present in genes encoding canonical mevalonate kinase. A similar gene arrangement can be found in the genomes of certain gammaproteobacteria, for example *Thiomicrospira crunogena* XCL-2. Most likely, in these organisms one of both genes encodes a phosphomevalonate kinase, which plays an essential role in the mevalonate pathway of bacteria. Based on the phylogenetic analyses presented in Fig. S3 it can be inferred that in L21-Fru-AB^T the genes at L21SP4_02096 and L21SP4_02097 encode for phosphomevalonate kinase and mevalonate kinase, respectively.


a
Kitasatospora griseola (A0A0D0NAT4)
Streptomyces anulatus (C4PWA7)
Actinoplanes sp. A40644 (Q767G7)
Nocardia farcinica (Q5YXN2)
Mycobacterium marinum (B2HGT6)
Corynebacterium amycolatum (E2MUL1)
Actinobacteria
"Ca. Liberibacter americanus" (U6B583)
Nitrincola lacisaponensis (A0A063Y8I5)
Neptuniibacter caesariensis (Q2BIK1)
Thiomicrospira crunogena (Q31EW2)
Dichelobacter nodosus (A5EUV4)
Cardiobacterium hominis (C8N7I0)
Paracoccus zeaxanthinifaciens (Q8L1I3)
Actibacterium atlanticum (A0A058ZKM6)
Alpha-/ Gammaproteobacteria
Kiritimatiella glycovorans (L21SP4_02098)
Kiritimatiellaeota
Enterococcus faecalis (Q9FD71)
Listeria welshimeri (A0AIL8)
Oceanobacillus iheyensis (Q8EP69)
Lactobacillus casei (K0N9K3)
Pediococcus pentosaceus (Q03FU5)
Leuconostoc mesenteroides (Q03WZ0)
Oenococcus oeni (Q04F95)
Firmicutes
Aquimarina atlantica (A0A023BPP6)
Kordia algicida (A9EAC6)
Bacteroidetes
Coxiella burnetii (A9KEM5)
Gammaproteobacteria
Plesiocystis pacifica (A6G7N7)
Cystobacter violaceus (A0A084SMI3)
Stigmatella aurantiaca (E3FUN3)
Myxococcus xanthus (Q1D4I1)
Deltaproteobacteria
Waddlia chondrophila (D6YWG1)
Chlamydiae
Spirochaeta africana (H9UG39)
Salinispira pacifica (V5WKY3)
Borrelia burgdorferi (O51626)
Spirochaetes
100%
99%
96%
100%
98%
95%
83%
86%
99%
99%
100%
100%
93%
75%
100%
76%
99%
100%
0.02

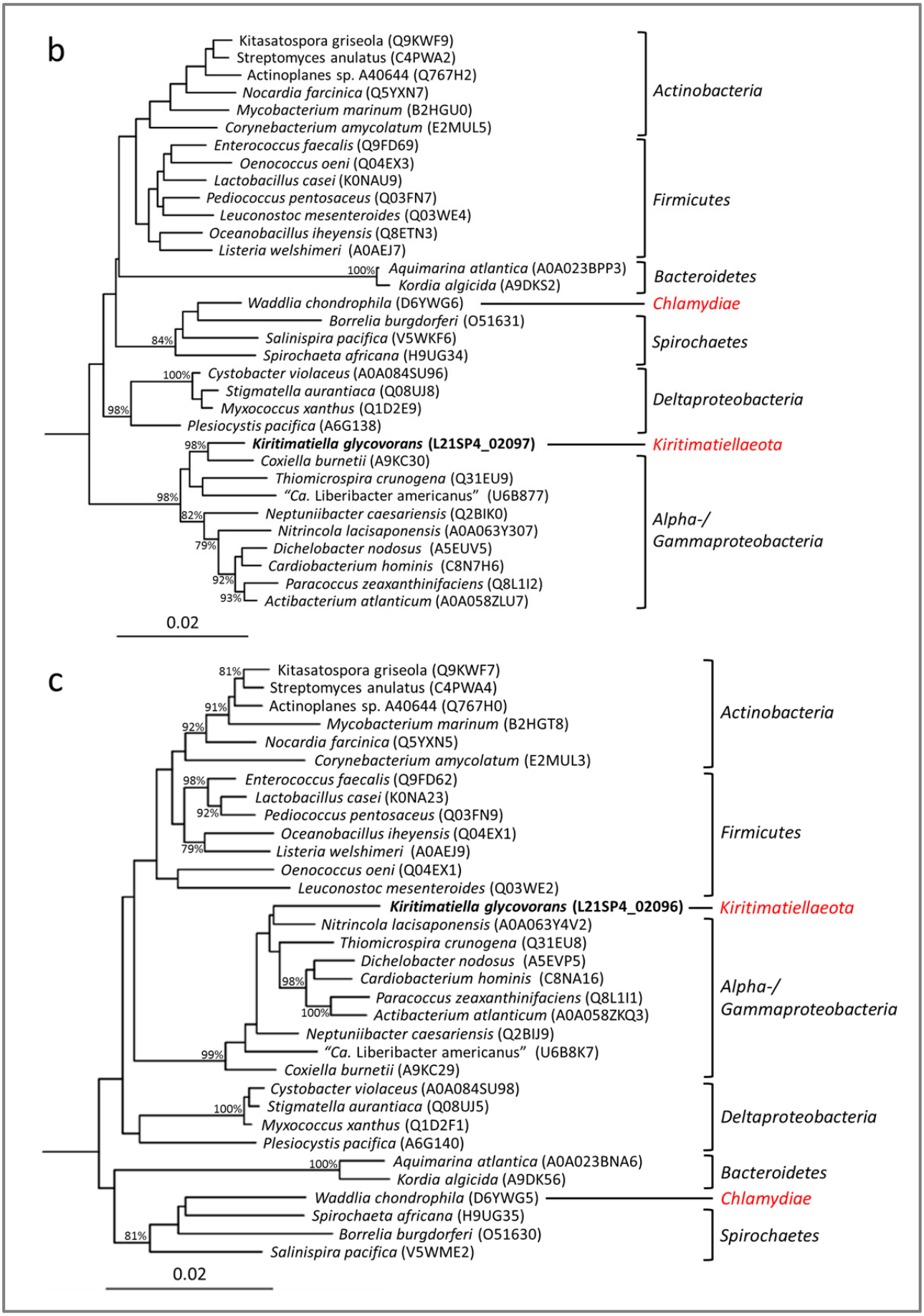

b
Kitasatospora griseola (Q9KWF9)
Streptomyces anulatus (C4PWA2)
Actinoplanes sp. A40644 (Q767H2)
Nocardia farcinica (Q5YXN7)
Mycobacterium marinum (B2HGU0)
Corynebacterium amycolatum (E2MUL5)
Enterococcus faecalis (Q9FD69)
Oenococcus oeni (Q04EX3)
Lactobacillus casei (K0NAU9)
Pediococcus pentosaceus (Q03FN7)
Leuconostoc mesenteroides (Q03WE4)
Oceanobacillus iheyensis (Q8ETN3)
Listeria welshimeri (A0AEJ7)
Aquimarina atlantica (A0A023BPP3)
Kordia algicida (A9DKS2)
Waddlia chondrophila (D6YWG6)
Borrelia burgdorferi (O51631)
Salinispira pacifica (V5WKF6)
Spirochaeta africana (H9UG34)
Cystobacter violaceus (A0A084SU96)
Stigmatella aurantiaca (Q08UJ8)
Myxococcus xanthus (Q1D2E9)
Plesiocystis pacifica (A6G138)
Kiritimatiella glycovorans (L21SP4_02097)
Coxiella burnetii (A9KC30)
Thiomicrospira crunogena (Q31EU9)
"Ca. Liberibacter americanus" (U6B877)
Neptuniibacter caesariensis (Q2BIK0)
Nitrincola lacisaponensis (A0A063Y307)
Dichelobacter nodosus (A5EUV5)
Cardiobacterium hominis (C8N7H6)
Paracoccus zeaxanthinifaciens (Q8L1I2)
Actibacterium atlanticum (A0A058ZLU7)
100%
98%
84%
100%
98%
98%
98%
82%
79%
92%
93%
Actinobacteria
Firmicutes
Bacteroidetes
Chlamydiae
Spirochaetes
Deltaproteobacteria
Kiritimatiellaeota
Alpha-/
Gammaproteobacteria
0.02
c
Kitasatospora griseola (Q9KWF7)
Streptomyces anulatus (C4PWA4)
Actinoplanes sp. A40644 (Q767H0)
Mycobacterium marinum (B2HGT8)
Nocardia farcinica (Q5YXN5)
Corynebacterium amycolatum (E2MUL3)
Enterococcus faecalis (Q9FD62)
Lactobacillus casei (K0NA23)
Pediococcus pentosaceus (Q03FN9)
Oceanobacillus iheyensis (Q04EX1)
Listeria welshimeri (A0AEJ9)
Oenococcus oeni (Q04EX1)
Leuconostoc mesenteroides (Q03WE2)
Kiritimatiella glycovorans (L21SP4_02096)
Nitrincola lacisaponensis (A0A063Y4V2)
Thiomicrospira crunogena (Q31EU8)
Dichelobacter nodosus (A5EVP5)
Cardiobacterium hominis (C8NA16)
Paracoccus zeaxanthinifaciens (Q8L1I1)
Actibacterium atlanticum (A0A058ZKQ3)
Neptuniibacter caesariensis (Q2BIJ9)
"Ca. Liberibacter americanus" (U6B8K7)
Coxiella burnetii (A9KC29)
Cystobacter violaceus (A0A084SU98)
Stigmatella aurantiaca (Q08UJ5)
Myxococcus xanthus (Q1D2F1)
Plesiocystis pacifica (A6G140)
Aquimarina atlantica (A0A023BNA6)
Kordia algicida (A9DK56)
Waddlia chondrophila (D6YWG5)
Spirochaeta africana (H9UG35)
Borrelia burgdorferi (O51630)
Salinispira pacifica (V5WME2)
81%
91%
92%
98%
92%
79%
98%
100%
99%
100%
100%
81%
Actinobacteria
Firmicutes
Kiritimatiellaeota
Alpha-/
Gammaproteobacteria
Deltaproteobacteria
Bacteroidetes
Chlamydiae
Spirochaetes
0.02

Figure S3. Phylogenetic trees of enzymes of the mevalonate pathway showing the position of *K. glycovorans* L21-Fru-AB^T among genome-sequenced bacteria having this pathway. Tree topologies were reconstructed under the maximum-likelihood criterion and rooted using respective sequences of the archaeon *Sulfolobus tokodaii* (not shown). Only maximum likelihood bootstrap values equal or above 75% are shown. Scale bar, 0.02 changes per nucleotide position. Phylogenetic trees of hydroxymethylglutaryl-CoA synthase **(a)**, mevalonate kinase **(b)** and phosphomevalonate kinase **(c)**. Note, that the gene encoding the phosphomevalonate kinase of *Plesiocystis pacifica* encodes also hydroxymethylglutaryl-CoA reductase, but for the phylogenetic tree reconstruction only the part of the protein was used that encodes phosphomevalonate kinase.

7. Regulation of central metabolic pathways

The genome of strain L21-Fru-AB^T harbours genes of a rudimental PEP:carbohydrate phosphotransferase system (PTS) used in some bacteria for the group translocation of sugars. Proteins involved in the phosphorylation cascade and encoded in the genome include phosphoenolpyruvate-protein phosphotransferase (L21SP4_00794 and L21SP4_01117), phosphocarrier histidine protein Hpr (L21SP4_00014 and L21SP4_00793), HPr kinase/phosphatase (L21SP4_00792) and six proteins containing putative Enzyme IIA domains (L21SP4_00166, L21SP4_00231, L21SP4_00347, L21SP4_01384, L21SP4_01661, and L21SP4_02044). However, genes of the specific PTS components II_B and II_C involved in the uptake and phosphorylation of sugars were not found, so that the phosphotransferase system in strain L21-Fru-AB^T is non-functional for the import of sugars.

Four of the six genes encoding a putative Enzyme II_A domain that is usually responsible for transferring of the phosphoryl group to a sugar phosphorylase (Enzyme II_B) were located adjacent to genes for monovalent cation:proton antiporters (Supplementary Table S2), which may indicate that these membrane-bound transporters represent terminal targets for phosphorylation and thus regulation at the protein level. The regulation of Na^+/H^+ exchangers for retaining cellular homeostasis of pH and sodium could be especially important during fermentative growth with Na^+ as coupling ion and therefore has to be tightly linked with the carbon flow in this strain. In addition, a HPr kinase/phosphatase that regulates the activity of

the histidine protein (HPr) transferring phosphoryl groups to Enzyme II_A could be involved in sensing the cellular energy status and controlling of the phosphorylation cascade leading to the activation of unknown transcription factors (Boël *et al.*, 2003).

Besides the PTS IIA domain most of the putative Enzyme II_A genes contain additional domains. Genes located at L21SP4_01384 and L21SP4_02044 contain also domains, which are suggested to bind nucleic acids, while the gene located at L21SP4_01661 encodes also an amino acid permease. Furthermore, the phosphoenolpyruvate-protein phosphotransferase encoded at L21SP4_01117 contains a regulatory GAF domain that binds cyclic nucleotides thereby allowing an interaction with the activity of a membrane-bound adenylate cyclase (L21SP4_01616) containing a periplasmic sensor domain.

References

Biegel E, Schmidt S, González JM, Müller V. (2011). Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell Mol Life Sci* **68**:613–634.

Boël G, Mijakovic I, Mazé A, Poncet S, Taha MK, Larribe M, *et al.* (2003). Transcription regulators potentially controlled by HPr kinase/phosphorylase in gram-negative bacteria. *J Mol Microbiol Biotechnol* **5**:206–215.

Ben Hania W, Joseph M, Schumann P, Bunk B, Fiebig A, Spröer C, *et al.* (2015). Complete genome sequence and description of *Salinispira pacifica* gen. nov., sp. nov., a novel spirochaete isolated form a hypersaline microbial mat. *Stand Genomic Sci* **10**:7.

Jacquier N, Frandi A, Viollier PH, Greub G. (2015). Disassembly of a medial transenvelope structure by antibiotics during intracellular division. *Chem Biol* **22**:1217–1227.

Jeske O, Schüler M, Schumann P, Schneider A, Boedeker C, Jogler M, *et al.* (2015). *Planctomycetes* do possess a peptidoglycan cell wall. *Nat Commun* **6**:7116.

Khadem AF, van Teeseling MCF, van Niftrik L, Jetten MSM, Op den Camp HJM, Pol A. (2012). Genomic and physiological analysis of carbon storage in the verrucomicrobial methanotroph '*Ca.* Methylacidiphilum fumariolicum' SolV. *Front Microbiol* **3**:345.

Korotkov KV., Sandkvist M, Hol WGJ. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* **10**:336–351.

Marraffini LA, Sontheimer EJ. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**:181–190.

Moran NA, Plague GR. (2004). Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev* **14**:627–633.

Pilhofer M, Aistleitner K, Biboy J, Gray J, Kuru E, Hall E, *et al.* (2013). Discovery of

chlamydial peptidoglycan reveals bacteria with murein sacculi but without FtsZ. *Nat Commun* **4**:2856.

Schmid J, Sieber V, Rehm B. (2015). Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front Microbiol* **6**:496.

Schut GJ, Adams MWW. (2009). The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: A new perspective on anaerobic hydrogen production. *J Bacteriol* **191**:4451–4457.

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res* **39**:W347–352.