environmental microbiology

Environmental Microbiology (2019) **21**(11), 4151–4165 doi:10.1111/1462-2920.14768

Acetic acid bacteria encode two levansucrase types of different ecological relationship

Frank Jakob,[1*] **Yamira Quintero**,[2] **Alexis Musacchio**,[3] **Paulina Estrada-de los Santos**,[4] **Lázaro Hernández**[2*] **and Rudi F. Vogel**[1]

[1]*Lehrstuhl für Technische Mikrobiologie, Technische Universität München, Gregor-Mendel-Straße 4, 85354 Freising, Germany.*

[2]*Grupo Tecnología de Enzimas, Centro de Ingeniería Genética y Biotecnología (CIGB), Ave 31 entre 158 y 190, Apartado Postal 6162, Habana 10600, Cuba.*

[3]*Departamento de Biología de Sistemas, Centro de Ingeniería Genética y Biotecnología (CIGB), Ave 31 entre 158 y 190, Apartado Postal 6162, Habana 10600, Cuba.*

[4]*Instituto Politécnico Nacional, Escuela Nacional de Ciencias Biológicas, Prol. de Carpio y Plan de Ayala s/n, Col. Santo Tomás C.P., 11340 Cd. de México, Mexico.*

Summary

Acetic acid bacteria (AAB) are associated with plants and insects. Determinants for the targeting and occupation of these widely different environments are unknown. However, most of these natural habitats share plant-derived sucrose, which can be metabolized by some AAB via polyfructose building levansucrases (LS) known to be involved in biofilm formation. Here, we propose two LS types (T) encoded by AAB as determinants for habitat selection, which emerged from vertical (T1) and horizontal (T2) lines of evolution and differ in their genetic organization, structural features and secretion mechanism, as well as their occurrence in proteobacteria. T1-LS are secreted by plant-pathogenic α- and γ-proteobacteria, while *T2-LS* genes are common in diazotrophic, plant-growth-promoting α-, β- and γ-proteobacteria. This knowledge may be exploited for a better understanding of microbial ecology, plant health and biofilm formation by sucrase-secreting proteobacteria in eukaryotic hosts.

Received 16 April, 2019; revised 24 July, 2019; accepted 31 July, 2019. *For correspondence. E-mail frank.jakob@wzw.tum.de; Tel.: +49 8161 71 5273; fax: +49 8161 71 3327; lazaro.hernandez@cigb.edu.cu; Tel.: +53 72504361; fax: +53-72731779

Introduction

Acetic acid bacteria (AAB) are Gram-negative, aerobic rods, which belong to the class of α-proteobacteria. They are currently subdivided into 17 genera, which constitute the acetous group of *Acetobacteraceae* (Yamada, 2016). The genera with the highest numbers of described species are *Acetobacter*, *Komagataeibacter*, *Gluconobacter*, *Gluconacetobacter* and *Asaia*. The main characteristic of AAB is their oxidative metabolism enabling them to oxidize diverse alcohols and sugars to the corresponding acids via membrane-bound dehydrogenases, which are part of the respiratory chain and whose active centres are oriented into the periplasm (Matsushita *et al*., 1994; Deppenmeier and Ehrenreich, 2009). For most oxidation reactions, substrates are taken up into the periplasm, oxidized and subsequently released into the environment resulting in the acidification of the extracellular space. Most AAB can, in this way, cope with high alcohol or sugar concentrations, as energy can be generated without uptake of the respective osmolytes into the cytoplasm (Deppenmeier *et al*., 2002). Because of their acid and alcohol tolerance (Mullins *et al*., 2008) some AAB are commonly found coexisting with lactic acid bacteria and yeasts, for example, in traditionally fermented foods like kefir (Gulitz *et al*., 2011; Laureys and De Vuyst, 2014), kombucha, cocoa beans or coffee (De Roos and De Vuyst, 2018). *Acetobacter* and *Komagataeibacter* spp. are specialized on ethanol conversion to acetic acid via two successive oxidative steps and are thus used for vinegar manufacture (Gullo and Giudici, 2008; Raspor and Goranovič, 2008; Yakushi and Matsushita, 2010). By contrast, *Gluconobacter* spp. preferably oxidize glucose to gluconic acid(s) and usually occur in sugary environments (Prust *et al*., 2005). Most species of *Gluconacetobacter* are N_2-fixing, endophytic symbionts of plants like sugar cane and coffee (Cavalcante and Dobereiner, 1988; Pedraza, 2008), while *Asaia* spp. are beverage spoilers (Moore *et al*., 2002; Kregiel *et al*., 2012) and symbionts of malaria-transmitting mosquitoes (Favia *et al*., 2007; Damiani *et al*., 2010; Chouaia *et al*., 2012). The remaining AAB genera mostly comprise one single species being distinctly less frequently isolated from natural sources than strains of the five main genera. These more rarely

occurring species have been mostly isolated from sugary plants (Urakami *et al*., 1989; Lisdiyanti *et al*., 2002; Jojima *et al*., 2004; Loganathan and Nair, 2004; Yukphan *et al*., 2005; Yukphan *et al*., 2008; Yukphan *et al*., 2009; Yukphan *et al*., 2011; Ramírez-Bahena *et al*., 2013; Vu *et al*., 2013), at which the recently described genera *Bombella* and *Parasaccharibacter* appear to be symbiotically associated with honey bees (Corby-Harris *et al*., 2014; Li *et al*., 2015; Corby-Harris *et al*., 2016; Yun *et al*., 2017; Corby-Harris and Anderson, 2018). AAB are hence predominantly found on or in plants and their temporary visiting, sugar-feeding insects, which can be considered as conveyers of AAB (Crotti *et al*., 2010).

While the conversion of monosaccharides by AAB is well understood (Deppenmeier and Ehrenreich, 2009), the consumption of sucrose being the most abundant carbohydrate in photosynthetic plants (Avigad, 1982) has not been systematically described in AAB. In general, sucrose can be utilized by bacteria either in phosphorylated form by sucrose-6-phosphate hydrolases or non-phosphorylated form by β-fructosidases (invertases) or exo-fructanases (Reid and Abratt, 2005; Prechtl *et al*., 2018), all of them belonging to the glycoside hydrolase (GH) 32 family defined in the CAZy database (http://www.cazy.org/). Sucrose is also the natural substrate of secreted fructansucrases (GH68) and glucansucrases (GH70), responsible for the synthesis of fructan- and glucan-type homopolymers respectively (van Hijum *et al*., 2006). While the β-(2 → 6) linked polyfructan levan and the α-(1 → 6) linked polyglucan dextran are components of biofilms (Dogsa *et al*., 2013; Fels *et al*., 2018), the released monosaccharide (glucose by GH68 and fructose by GH70) is directly used for metabolic purposes. Besides their respective transfructosylation or transglucosylation, GH68 and GH70 enzymes also hydrolyze sucrose, as water often acts as the acceptor molecule (van Hijum *et al*., 2006). While dextransucrase (EC 2.4.1.5) is exclusively expressed by lactic acid bacteria (*Lactobacillales*), GH68 encoding genes are present among Gram-positive/-negative bacteria, archaea, fungi and plants (Öner *et al*., 2016). Diverse previous works revealed AAB from the genera *Gluconobacter*, *Gluconacetobacter*, *Komagataeibacter*, *Asaia*, *Neoasaia* and *Kozakia* as producers of active levansucrases (LS; EC 2.4.1.10) (Tajima *et al*., 1997; Arrieta *et al*., 2004; Kato *et al*., 2007; Jakob, 2014; Semjonovs *et al*., 2016).

Here, we provide a systematic overview about the diversity, genetic background, evolutionary connections, distinctive distribution, and modular properties of LSs from AAB based on own previous works and comparative genomics. We further summarize the presence of putative sucrose/fructan-hydrolyzing GH32 enzymes among AAB. In this way, we highlight the ecological and evolutionary relationship of AAB to other saccharolytic proteobacteria of diverse habitats.

Results and discussion

Two types of levansucrases are distinguishable in AAB and other proteobacteria

To get an overview about LSs from AAB, public available LS protein sequences were collected and used for calculation of a phylogenetic tree (Fig. 1A). Our database search revealed that LSs are especially abundant in the genera *Gluconobacter*, *Kozakia*, *Neoasaia*, *Asaia*, *Tanticharoenia* and *Gluconacetobacter* and that with few exceptions solely one levansucrase gene is encoded in these AAB. On the contrary, LSs are sporadically or not found in *Komagataeibacter*, *Acetobacter* and other AAB (Fig. 1B).

LSs encoded by a chromosomal gene from *Acetobacter aceti*, *Kozakia baliensis* and several species of *Gluconobacter* are phylogenetically next related to those of *Zymomonas mobilis* (α-proteobacteria, Sphingomonadaceae) and certain γ-proteobacteria, *inter alia* to the proven levan-producing species *Pseudomonas* (*P*.) *syringae* and *Erwinia* (*E*.) *amylovora*. By contrast, LSs from *Gluconacetobacter* and *Asaia* are next related to those of *Beijerinckia indica* (α-proteobacteria, Beijerinckiaceae), *Burkholderia*/*Paraburkholderia*/*Caballeronia*/*Cupriavidus* spp. (β-proteobacteria, Burkholderiaceae) and *Azotobacter* (γ-proteobacteria, Pseudomonadaceae). Considering that LSs from AAB are positioned in two separate phylogenetic clades (see also Supplementary File S1), they were termed as 'Type 1 (T1)' and 'Type 2 (T2)' LSs.

For further differentiation of T1- and T2-LS from AAB, we used the existing literature about intrinsic LS properties and levan production in/by AAB. While there are no reports about the activity of isolated T1-LS from AAB, strains carrying *T1-LS* genes produce considerable amounts of levan exhibiting very high molecular weight ($>10^9$ g mol^{-1}) (Jakob *et al*., 2013; Ua-Arak *et al*., 2017). On the contrary, T2-LS from the sugar cane endophyte *Gluconacetobacter* (*Ga*.) *diazotrophicus* yield high amounts of short-chain fructooligosaccharides and a lower level of polymerized levan (Trujillo *et al*., 2001). Furthermore, T2-LSs are most commonly synthesized as precursor proteins with a cleavable N-terminal sequence (Fig. 2, Supplementary File S2). The predicted 30-aa signal peptide of T2-LS from *Ga*. *diazotrophicus* was demonstrated to be cleaved off during the protein transport to the periplasmic space, where the N-terminal Gln is converted into cyclic pyroglutamate (pGlu) conferring protection against proteolytic degradation (Hernández *et al*., 1999). In a second step, the folded mature enzyme is transported across the outer membrane by the type-II secretory machinery (Arrieta *et al*., 2004).

Contrary, all T1-LSs are devoid of a predictable signal peptide (Fig. 2, Supplementary File S2), suggesting that their release to the external medium does not involve proteolytic cleavage. For instance, secretion of the entire

Genus	T1-LS	T2-LS	GH 32
Komagataeibacter	5/48	7/48	19/48
Kozakia	3/3	1/3	3/3
Neoasaia	1/1	1/1	1/1
Neokomagataea	0/1	0/1	1/1
Nguyenibacter	0	0	0
Saccharibacter	0/4	0/4	4/4
Swaminathania	0	0	0
Tanticharoenia	1/1	1/1	1/1
Acetobacter	3/106	0/106	15/106
Acidomonas	0/1	0/1	0/1
Ameyamaea	0	0	0
Asaia	0/9	5/9	9/9
Endobacter	0	0	0
Gluconacetobacter	0/5	4/5	4/5
Gluconobacter	41/41	1/41	41/41
Granulibacter	0/1	0/1	0/1
Bombella/Parasaccharibacter	0/6	0/6	6/6

Fig. 1. A. Phylogenetic tree calculated on the basis of representative LS from AAB and proteobacteria expressing similar LS using the neighbour-joining method. The scale-bar in A indicates number of changes per amino acid. Numbers following species names represent the accession numbers of the corresponding genes as deposited in GenBank. SacB from *Bacillus subtilis* (Firmicutes) was used as an outgroup. B. Abundance of T1-LS, T2-LS and GH 32 (Data analysis and processing) encoding sequences in genomes of the 17 currently described AAB genera. Depicted data per genus refer to the number of deposited genomes containing at least one of the respective genes/totally deposited genomes.

A

```
T1-LSs
Psyr  01 ------------------------------MSNSSSAVIQHKNSPLVGNIKYAPTVWSRADALKVNENDP
Eamy  01 ------------------------------------------------MSDYNYKPTLWTRADALKVHEDDP
Zmob  01 ----------------------------------------------MLNKAGIAEPSLWTRADAMKVHTDDP
Galb  01 ------------------------------------------------MRQTSRWTIADALKVHADDP
Kbal  01 ----------------------------------------MNLGISAKNQDVRTPSHWTIADAMKVHADDP

T2-LSs
Achr  28 ------------------------------ASLEPGPEPTVHTQEAFAPEGNFTAKWTRADARQIKRMSD
Gdia  31 QGNFSRQEAARMAHRPGVMPRGGPLFPGRSLAGVPGFPLPSIHTQQAYDPQSDFTARWTRADALQIKAHSD
Bind  29 ---------------------------------QSGYPIPTPHSGQAYDPFADFTAKWTRANARQIKAQSH
Bvie  32 ---------------------------------QTAGAPAPTPHTQQAHDPESNFTMRWTRADIRQIVAQSH
Pgra  36 ------------------------------QSSAAGFPAPTPHTQQAYDPDSSFTMRWTRADVRQIKAQSH
                                                                  *: *:  ::   .
```

B

```
T1-LSs
Psyr  248 VGSVELGPVPPG-------YEDVGGARFQVGCIGLAVAKD-------LTGEEWEILPPLVTAVGVNDQTERPH
Eamy  232 ITQAEMGNVPPG-------YEDVGGAKYQAGCVGLAVAKD-------LSGSEWQILPPLITAVGVNDQTERPH
Zmob  223 VGEEEIGPVPPK-------TETPDGARYCAAAIGIAQALN-------EARTEWKLLPPLVTAFGVNDQTERPH
Galb  228 IGANETGPIPPS-------YTVDPGAGYGAASIGLAILDDGAYEAGDFSRARWTQLGPIVSALGVNDQTERPH
Kbal  233 LHDTEIGAVPPG-------YTPAAGAQYGAAAIGIARLTDGAYEKGDFSRHNWTLLPALVTALGVNDQTERPH

T2-LSs
Achr  302 CTAEDLGYREGDPYAETVEQVNASGATYQIGNIGLAKAKN-------ADLTEWEFLPPILSANCVTDQTERPQ
Gdia  339 CTEADLGFRPNDPNAETLQEVLDSGAYYQKANIGLAIATD-------STLSKWKFLSPLISANCVNDQTERPQ
Bind  315 CTEADLGYSPNDPNKEDLNAVMDSGAIYQMANVGLAVATN-------DELTQWKFLPPILSGNCVNDQTERPQ
Bvie  306 CTEADLGYAPNDPYREDLNAVMNAGAVYQKANVGLAVATN-------PQLTEWKFLPPILSANCVDDQTERPQ
Pgra  313 CTEADLGYASNDPYKEDLNAVMNSGATYQKANVGLAVATN-------KQLTEWKFLPPLLSANCVNDQTERPQ
                 : *                ** :  . :*:*   :           .*  *  :::. * ******:
                                    S–S
```

Fig. 2. Partial sequence alignment of representative T1- and T2-levansucrases (LSs). A. N-terminal region of predicted mature proteins. B. Position of Cys residues forming the fold-stabilizing disulfide bond in T2-LSs. Numbers correspond to the precursor proteins. Sequences starting at Met 01 lack signal peptide. The non-catalytic signature motif W-T/S-R/I-A-D/N-A is boxed. Enzyme source and Protein IDs correspond to: Psyr, *Pseudomonas syringae* (WP_046266212); Eamy, *Erwinia amylovora* (WP_004161222); Zmob, *Zymomonas mobilis* (WP_011240294); Galb, *Gluconobacter albidus* (WP_082780180); Kbal, *Kozakia baliensis* (WP_083301879); Achr, *Azotobacter chroococcum* (WP_052264013); Gdia, *Gluconacetobacter diazotrophicus* (WP_012222901); Bind, *Beijerinckia indica* (WP_012384999); Bvie, *Burkholderia vietnamiensis* (WP_011882121); Pgra, *Paraburkholderia graminis* (WP_006050632). Symbols represent (*) identity, (:.) conservative changes.

coding region of T1-LS from *Gluconobacter* (*G.*) *albidus* TMW 2.1191 was demonstrated by Jakob (2014) via peptide sequencing of the extracellularly recovered enzyme. The signature motif W-T-R/I-A-D/N-A, which is part of the mature region of cleaved T2-LSs, is placed only few amino acids downstream of the initiating Met in T1-LSs (Fig. 2A), supporting the idea of a general lack of proteolytic processing in T1-LSs. The signal-peptide independent secretion mechanism of T1-LSs from AAB or the functionally characterized T1-LSs from other proteobacteria as *P. syringae* or *E. amylovora* has not been identified yet.

Further differences between T1- and T2-LS are derived from their intrinsic properties. T1-LSs from AAB have a lower molecular mass range (46–49 kDa) than proteolytically processed T2-LSs (56–62 kDa), due to the shorter length of some of the loops connecting the β-strands. The calculated isoelectric point ranges 4.7–5.3 for T1-LSs and 5.2–6.9 for mature T2-LSs (Supplementary File S2).

The most distinctive structural feature differentiating the two LS types in AAB is the formation of a unique intramolecular disulfide bridge in T2-LSs, which is absent in T1-LSs (Fig. 2B, Supplementary File S2). The Cys339—Cys395 bond in the crystal structure of T2-LS from *Ga. diazotrophicus* (PDB:1W18) stabilizes the protein fold by connecting the extended loop between β-strands IIIB and IIIC with the insertion located between blades III and IV (Fig. 3A) (Betancourt *et al*., 1999; Martínez-Fleites *et al*., 2005). An equivalent pair of Cys residues is strictly conserved in all T2-LSs from different origins (Fig. 2B, Supplementary File S2). The intramolecular disulfide bond in

T2-LS from *Ga. diazotrophicus*, which is preserved in β-fructofuranosidase (GH68) from *Microbacterium saccharophilum* K-1 (PDB:3VSR), is structurally complementary to the Ca^{2+}-cofactor binding site of the Cys-lacking LS from *Bacillus subtilis* (PDB:1OYG) (Meng and Fütterer, 2003) (Fig. 3B) and inulosucrase from *Lactobacillus johnsonii* (PDB:2YFR) (Pijning *et al*., 2011). Hence, it can be assumed that the disulphide bond formation in T2-LS from Proteobacteria and Actinobacteria and the cofactor binding in GH68 enzymes from Firmicutes have evolved at similar topological regions for a common function of fold stabilization.

A different scenario is observed in T1-type LSs that often contain several Cys residues but not at conserved positions. None of the four cysteines (positions 64, 103, 150, 256) in the crystal structure of T1-LS from *E. amylovora* (PDB:4D47) (Wuerges *et al*., 2015) are implicated in a disulfide bond (Fig. 3C). Similarly, the three cysteines (positions 142, 201, 213) of T1-LS from *G. albidus* (Seq Id. WP_082780180) reside at non-interacting distances in the constructed 3D structural model (Fig. 3D). LSs are released to the external medium as folded active enzymes. Differences in the fold stabilization mechanism between T1- and T2-LSs may thus have evolved during adaptation to the different natural environments of the host bacteria.

Analyses of gene sequence similarity and genomic organization suggest vertical T1-LS and horizontal T2-LS evolution in AAB

In Fig. 4 phylogenetic trees are depicted, which were calculated from multiple alignments based on *T1-LS* (Fig. 4A) and 16S rRNA gene sequences (Fig. 4B). Each available genome of *Gluconobacter*, *Kozakia* and *Zymomonas* harbours one *T1-LS* locus at a conserved region (Fig. 5A). Accordingly, PCR-screenings targeting on the *T1-LS* gene of *Gluconobacter* isolates from different habitats such as kefir or spoiled beer generally revealed positive amplicons (Fig. 4A). The synthesis of functional T1-LS proteins is supported by findings of Jakob *et al*. (2012a), in which each of the 12 tested *Gluconobacter* strains originating from diverse habitats produced levan from sucrose. In the study, a different extent of levan production was observed indicating a strain-specific regulation of levan biosynthesis.

Independently of the gene used for the tree construction (T1-LS vs. 16S rRNA), *Gluconobacter* strains can be subdivided into the known taxonomic *Gluconobacter oxydans* and *Gluconobacter cerinus* groups (Yamada and Yukphan, 2008), while *Gluconobacter morbifer* G707 retains its outstanding position (Fig. 4A and B). Moreover,

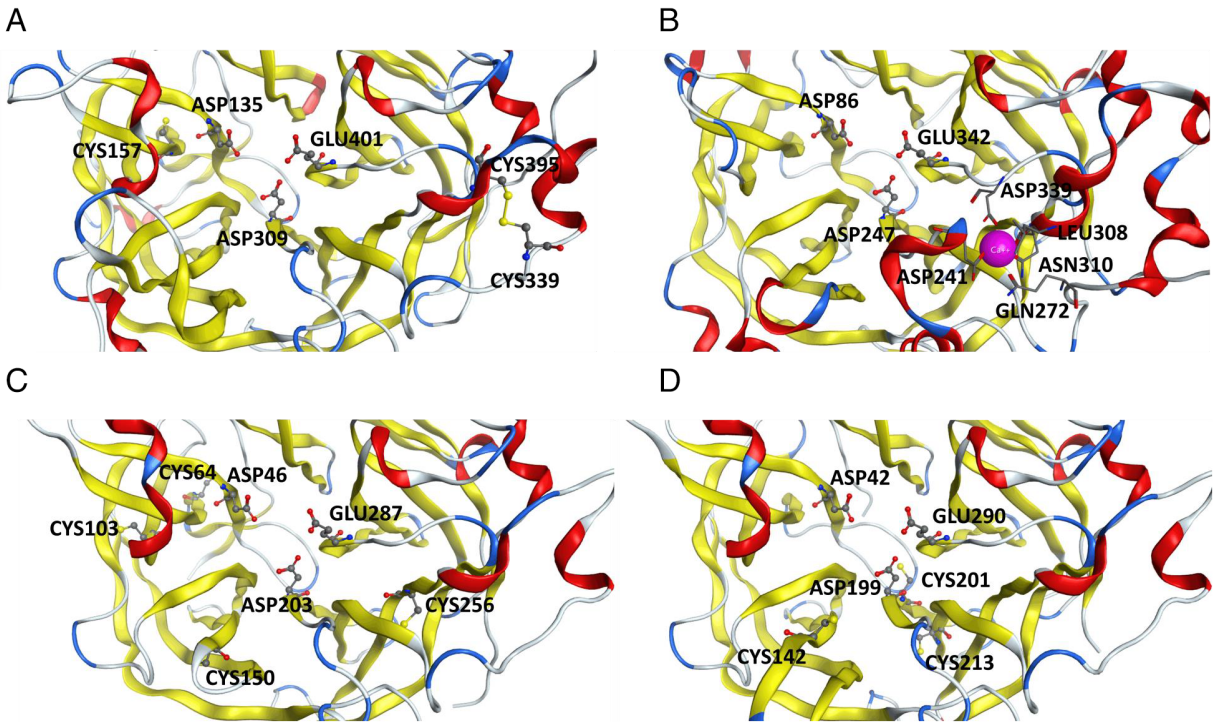

Fig. 3. Mapping of the fold-stabilizing disulfide bond (Cys339-Cys395) (A) T2-LS from *Gluconacetobacter diazotrophicus* (PDB:1W18) and the functionally/topologically complementary calcium binding site in (B) LS from *Bacillus subtilis* (PDB:1OYG). The Cys residues in the T1-LSs from (C) *Erwinia amylovora* (PDB:4D47), and (D) *Gluconobacter albidus* (WP_082780180) are not interacting with each other. The catalytic triad (Asp/Asp/Glu) at the bottom of the active site pocket is shown for each protein. The ribbon diagram of the determined or modelled 3D structures was prepared with PyMOL (http://www.pymol.org). [Color figure can be viewed at wileyonlinelibrary.com]

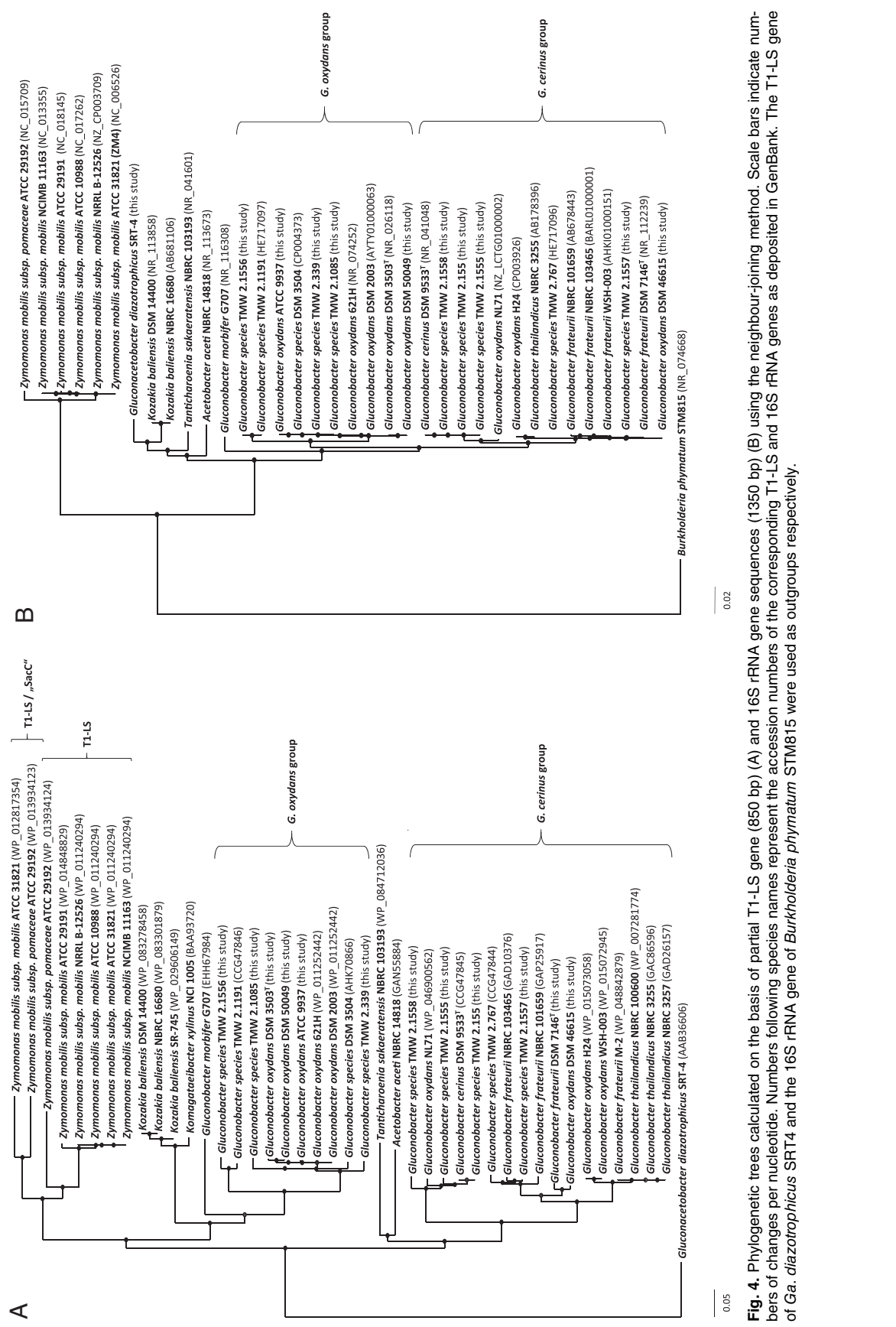

Fig. 4. Phylogenetic trees calculated on the basis of partial T1-LS gene (850 bp) (A) and 16S rRNA gene sequences (1350 bp) (B) using the neighbour-joining method. Scale bars indicate numbers of changes per nucleotide. Numbers following species names represent the accession numbers of the corresponding T1-LS and 16S rRNA genes as deposited in GenBank. The T1-LS gene of *Ga. diazotrophicus* SRT4 and the 16S rRNA gene of *Burkholderia phymatum* STM815 were used as outgroups respectively.

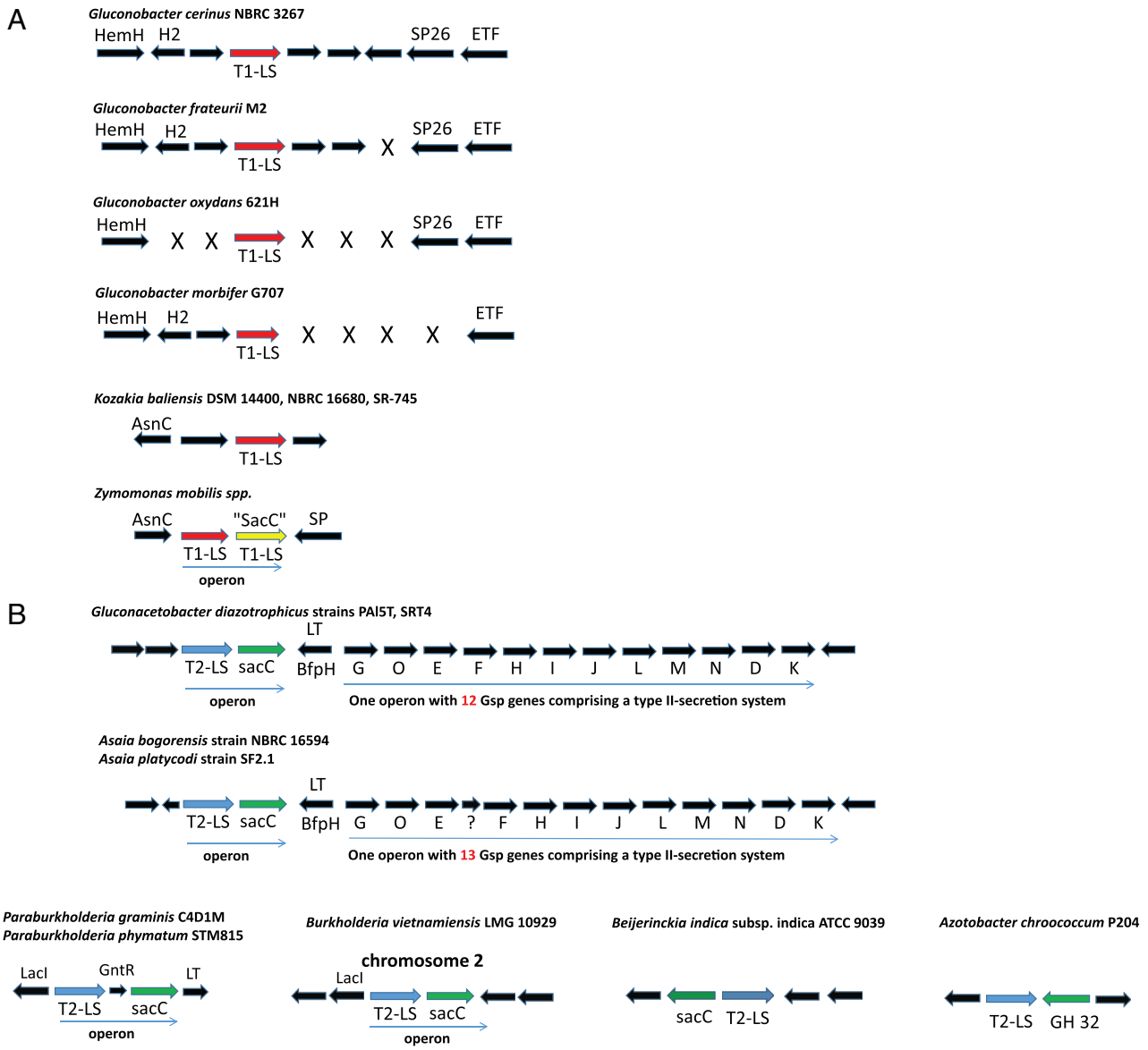

Fig. 5. Schematic overview of the genetic organization of T1- (A) and T2-LS (B) in AAB and related proteobacteria. 'X' indicate non-present, homologous genes. hemH, ferrochelatase; h, hypothetical protein; SP26, signal peptidase 26; ETF, electron transfer flavoprotein; AsnC, transcriptional regulator AsnC family; SP; signal peptidase; LT/BfpH, lytic transglycosylase; GntR, transcriptional regulator GntR family; LacI, transcriptional regulator LacI family. [Color figure can be viewed at wileyonlinelibrary.com]

T1-LS genes of *Gluconobacter* are located in a conserved genomic region, which slightly differs among the main groups *G. cerinus*, *G. frateurii*, *G. oxydans* and *G. morbifer* (Fig. 5A). Comparison of partial *T1-LS* genes (850 bp) allows a better *Gluconobacter* strain differentiation than obtained from partial 16S rRNA genes (1,350 bp) (see e.g., the subclusters comprising *G. species* TMW 2.155/2.1555/2.1558, *G. cerinus* DSM 9533T and *G. oxydans* NL71) (Fig. 4). *T1-LS* genes have thus evolved vertically within this genus and could be used as additional phylogenetic markers for differentiation of *Gluconobacter* strains.

Contrasting to the wide representation of T1-LS in the genus *Gluconobacter*, *LS* genes are scarcely found in species of *Acetobacter* and *Komagataeibacter* (Fig. 1B). The varied distribution of T1-LS among these three genera comprising the highest number of described AAB species clearly implies a diverse evolutionary adaptation in regard to the ability and way of metabolizing exogenous sucrose. *Gluconobacter* is in contrast to the preferably alcohol oxidizing genera *Acetobacter* and *Komagataeibacter* usually abundant in sugar-rich environments and generates most ATP via oxidation of glucose to (keto)-gluconate(s) by membrane-bound glucose/gluconate dehydrogenases

(Deppenmeier and Ehrenreich, 2009; Peters *et al*., 2013). The ability of *Gluconobacter* to synthesize levan while releasing glucose from sucrose by secreted T1-LS could thus be interpreted as an essential metabolic adaptation for the occupation of sucrose-rich niches, for instance plant and fruit surfaces from where *Gluconobacter* can usually be isolated (Barata *et al*., 2012; Komagata *et al*., 2014). The ecological abundance of *K. baliensis* is more unclear, because only few strains have been isolated from natural sources so far. The *T1-LS* locus is found at a conserved position in *K. baliensis*, which differs from that one of *Gluconobacter* (Fig. 5A). Moreover, some previous studies revealed that *K. baliensis* strains DSM 14400 and NBRC 16680 produce acetan-like heteropolysaccharides, whose structural features and biosynthesis are related to that of xanthan and amylovoran due to their biosynthesis via gum-like proteins (Brandt *et al*., 2016, 2017, 2018). These heteropolysaccharides are considered as virulence factors during colonization of plants in their producing species *Xanthomonas campestris* (Katzen *et al*., 1998) and *E. amylovora* (Koczan *et al*., 2009) respectively. Levan production via T1-LS as well as heteropolysaccharide production via gum-like proteins by AAB (Meneses *et al*., 2011) is thus related to that of plant-interactive species phylogenetically positioned in the γ-proteobacteria. The evolutionary relatedness of AAB (α-proteobacteria) to γ-proteobacteria is generally reflected by *Frateuria aurantia*, which was initially termed '*Acetobacter aurantius'* and is still considered as bispecific AAB due to its AAB-typical oxidative metabolism despite its phylogenetic position within the family Xanthomonadaceae (γ-proteobacteria) (Swings *et al*., 1980).

While the biological interactions between *T1-LS* expressing AAB and plants are still not understood, most of the *T2-LS* expressing *Gluconacetobacter* species are known N_2-fixing endophytes in host plants such as sugar cane, sugar beet, coffee, pineapple, maize, rice, sorghum and wheat, among others (Fuentes-Ramírez *et al*., 2001; Saravanan *et al*., 2008). Coincidentally, all eight PCR-screened isolates of *Ga. diazotrophicus* (strains CFN-Cf52 and CFN-Cf53), *Ga. azotocaptans* (strains CFN-Ca54^T, UAP-Ca97 and UAP-Ca99) and *Ga. johannae* (strains CFN-Cf55^T, CFN-Cf75 and UAP-Cf76) recovered from the rhizosphere or root tissue of coffee plants (Jimenez-Salgado *et al*., 1997; Fuentes-Ramírez *et al*., 2001) were found to carry the typical two-gene cluster encoding a T2-LS and an exolevanase (SacC; EC 3.2.1.65) (Fig. 5B, Supplementary File S3 and File S4). This operon organization (*T2LS-SacC*) was first identified in the sugarcane isolate *Ga. diazotrophicus* SRT4 (Arrieta *et al*., 1996) and then proven to be responsible for sucrose utilization and fructan metabolism in 14 *Ga. diazotrophicus* strains recovered from different host plants in diverse geographical regions (Hernández *et al*., 2000; Menéndez *et al*., 2002). In *Ga. diazotrophicus* the *T2LS-SacC* operon is functionally associated to the immediate downstream 12-gene cluster (G,O,E,F,H,I,J,L,M,N,D,K) encoding the components of the type-II general secretory pathway (Gsp) responsible for the transfer of both levansucrase and exolevanase across the outer membrane (Fig. 5B) (Arrieta *et al*., 2004). The two neighbour operons (*T2LS-SacC* and *Gsp*) are conserved in the chromosome of *Ga. liquefaciens* and some few strains of the genus *Asaia* (*A. bogorensis*, *A. platycodi*). Sequenced strains of at least four species of the genus *Komagataeibacter* (*K. oboediens*, *K. cocois*, *K. saccharivorans*, and *K. xylinus*) contain a chromosomal *T2LS-SacC* operon, but the Gsp components are encoded by two gene-clusters in opposite transcriptional directions placed upstream to *T2LS-SacC* (Supplementary File S4). A partial *Gsp* operon is found downstream of the *T2-LS* gene in sequenced plasmids of *K. baliensis* DSM 14400 and *G. frateurii* NBRC 103465 (Supplementary File S4). In both cases, the *SacC* gene is missing. In *K. baliensis* DSM 14400, T2-LS is not active due to a transposon insertion in the plasmid-encoded gene (Jakob, 2014; Brandt *et al*., 2016). The partial *T2-LS*/*SacC* and *Gsp* sequences on plasmids provide evidence of the horizontal acquisition and rearrangement of these genetic elements in some AAB. On the contrary, T1-LS encoding sequences have so far not been detected on plasmids in AAB, which again supports their vertical and ancient evolution within *Gluconobacter* and *Kozakia* as described above. To date, *T1-LS* genes have not been detected on plasmids, except for *P. syringae* pv. glycinea PG4180, a strain carrying three almost identical *T1-LS* genes (two on chromosome and one on an indigenous plasmid) (Li and Ullrich, 2001).

On nucleotide level, the components of the *T2LS-SacC* operon encoded by various AAB (Fig. 5B) are next related to those of the corresponding operon present in at least 24 strains of the taxonomically distant genera *Burkholderia* and *Paraburkholderia* (β-proteobacteria, Burkholderiaceae) (Supplementary File S3), some of which are also described as plant-interactive diazotrophs (Supplementary File S2) (Estrada-De Los Santos *et al*., 2001).

In *Ga. diazotrophicus*, the expression of the *T2LS-SacC* (*lsdA-lsdB*) operon is finely regulated allowing that levan synthesis and degradation are not excluding but complementing habitat-related processes. The *T2-LS* (*lsdA*) gene is transcribed in a constitutive manner, while *SacC* (*lsdB*) transcription is induced by fructose at low concentrations and repressed by glucose (Menéndez *et al*., 2009). The attained ability to express the *T2-LS-SacC* operon must contribute to plant infection by *Ga. diazotrophicus* and likely other N_2-fixing, plant-interactive AAB, as they are inevitably faced with exogenous sucrose during colonization of their host plants. In this sense, adaptation to the more favoured endophytic

habitat may have functioned as the selective force driving retention of the horizontally acquired *T2LS-SacC* operon. All sequenced strains from plant-interactive species of *Burkholderia* (*vietnamiensis*) and *Paraburkholderia* (*graminis*, *phymatum*, *diazotrophica*, *silvatlantica*) contain a conserved type II secretion (Gsp) operon in the chromosome but with a gene organization (D,E,F,C,G,H, I,J,K,L,M,N) differing from that of *Gluconacetobacter* and *Asaia* species, and located distantly from the *T2LS-SacC* operon (Fig. 5B, Supplementary File S4). A *Gsp* operon next to the *T2-LS* gene is also absent in the sequenced genomes of the plant-associated diazotrophs *Beijerinckia indica* (Becking, 2006) and *Azotobacter chroococcum* P204 (Kumar *et al*., 2001) (Fig. 5B). The *T2-LS* and *SacC* genes are positioned in opposite transcriptional direction in *Beijerinckia indica* and *A. chroococcum*, while in the latter strain a GH 32 protein exhibiting low homology to SacC is located next to its T2-LS (Fig. 5B). The role of T2-LS in *Asaia*, which comprises strains found as commensals of malaria-transmitting *Anopheles* sp. (*Asaia species* SF 2.1, Fig. 5B) or spoilers of beverages (Favia *et al*., 2007; Kregiel *et al*., 2012), can actually not be interpreted due to limited knowledge about their specific natural habitats and to the non-conserved distribution of T2-LS among species of this AAB genus. The fact that at least some *Asaia* strains are described as N_2-fixing organisms (Samaddar *et al*., 2011) could, however, point at a similar adaptation to the plant environment. Noticeably, in *Zymomonas* (*Z*.) *mobilis*, which is also capable of N_2-fixation (Kremer *et al*., 2015), a two-gene operon (Fig. 1A/5A) encoding T1-LS proteins with sequence similarity above 68% resembles the organization of the *T2LS-SacC* operon of *Ga*. *diazotrophicus*. In *Z*. *mobilis*, the first *T1-LS* gene is fully responsible for levan production (Fig. 1A), while the second *T1-LS* gene encodes a sucrose-hydrolysing enzyme unable to synthesize levan (Gunasekaran *et al*., 1995). Taken into account that T1-LS evolved vertically within *Gluconobacter* (Matsutani *et al*., 2011) and *Zymomonas* (Fig. 1A) and both genera are supposed to have evolved from a common, α-proteobacteria-like ancestor due to their highly similar intracellular sugar metabolism (Swings and De Ley, 1977), the 'primitive' levansucrase-invertase organization found in facultative anaerobic *Z*. *mobilis* could be considered as an ancient operon version adapted for the colonization of sucrose-rich environments under more anoxic conditions. The close ecological and evolutionary connection between *Gluconobacter* and *Zymomonas* is substantiated by their reported co-occurrence in sucrose-rich environments like plant saps or honey (Ruiz-Argueso and Rodriguez-Navarro, 1975). However, as observed for *Gluconobacter* (Fig. 5A), related T1-LS genes are not organized in *LS-SacC* operons in plant-pathogenic *E*. *amylovora* and *P*. *syringae*, which also strictly encode T1-LS and commonly invade plants from oxic surfaces such as damaged leaves or fruits (Osman *et al*., 1986; Gross and Rudolph, 1987; Gross *et al*., 1992; Kasapis *et al*., 1994; Hettwer *et al*., 1995; Bereswill *et al*., 1997; Zhang and Geider, 1999; Du and Geider, 2002; Laue *et al*., 2006; Koczan *et al*., 2009). *Gluconobacter* is supposed to be the causative agent of the pineapple pink disease (Swings and De Ley, 1981). Recent studies confirmed the plant-pathogenicity of *Gluconobacter cerinus* CDF1 (He *et al*., 2017), which fastens the rottening of bananas while being beneficial for the development of its transmitting fruit fly *Bactrocera dorsalis*. Any specific involvements of T1-LS from *Gluconobacter* in plant pathogenicity or animal interactions, however, still need to be demonstrated. The functional role of LS in the interaction between LS expressing microbes and sucrose-feeding animal hosts has so far exclusively been shown for cariogenic *Streptococcus mutans* strains, which build up levan and dextran containing biofilms on human teeth (Mukasa and Slade, 1973).

The active expression and functional role of GH32 genes in the sucrose/fructan metabolism of AAB remains unclear

Our database search further revealed that many AAB genomes contain one or more genes encoding putative GH32 enzymes, most likely exo-β-fructosidases. GH32 genes are strictly present in the genera *Gluconobacter*, *Saccharibacter*, *Parasaccharibacter* and *Bombella* (Fig. 1B). On the contrary, other AAB contain GH32 encoding genes at non-conserved genomic positions, at which these genes are more frequently found in genomes of *Kozakia* and *Asaia* than in *Acetobacter* and *Komagataeibacter*, similarly to the distribution of LS (Fig. 1B). Noticeably, the bispecific AAB *Frateuria aurantia* (γ-proteobacteria, Xanthomonadaceae) harbours a closely related GH32 encoding sequence in its genome (Fig. 6) again indicating the ecological relatedness of certain AAB to γ-proteobacteria. Because of the strict abundance of putative GH32 proteins in sugar-tolerant *Gluconobacter* and *Saccharibacter* (Jojima *et al*., 2004), their expression under certain growth conditions seems likely, while their specific function as sucrose or fructan hydrolases still has to be proven.

In total, three clades are distinguishable in the phylogenetic tree of GH32 proteins from AAB (Fig. 6). Regardless their source, all enzymes encoded by the second gene of the horizontally acquired *T2LS-SacC* operon are grouped in clade 3, including SacC (LsdB) from *Ga*. *diazotrophicus* SRT4 that was experimentally proven to be an exolevanase (Menéndez *et al*., 2002, 2004). A predicted signal-peptide containing GH32 enzyme (GenBank: ACI52355) with presumably vertical evolution in *Ga*. *diazotrophicus* PAl5^T is positioned in clade 1 together, among others, with the unique GH32 protein encoded by T1-LS containing strains

of *Acetobacter aceti* and some species of *Gluconobacter* (*G. oxydans*, *G. roseus*, *G. thailandicus*). The functional expression of the intrinsic GH32 gene in *Ga. diazotrophicus* PAl5^T is uncertain, as the bacterium lost the ability to grow on sucrose as the sole carbon source after insertional disruption of the *T2LS-SacC* (*lsdA-lsdB*) operon (Hernández *et al*., 2000).

Concluding remarks

Our study reveals that AAB encode two clearly distinguishable LS types. The differences in these extracellularly active LS and their encoding genes can be considered as evolutionary driven adaptations, which are used by many α-, β- and γ-proteobacteria for the

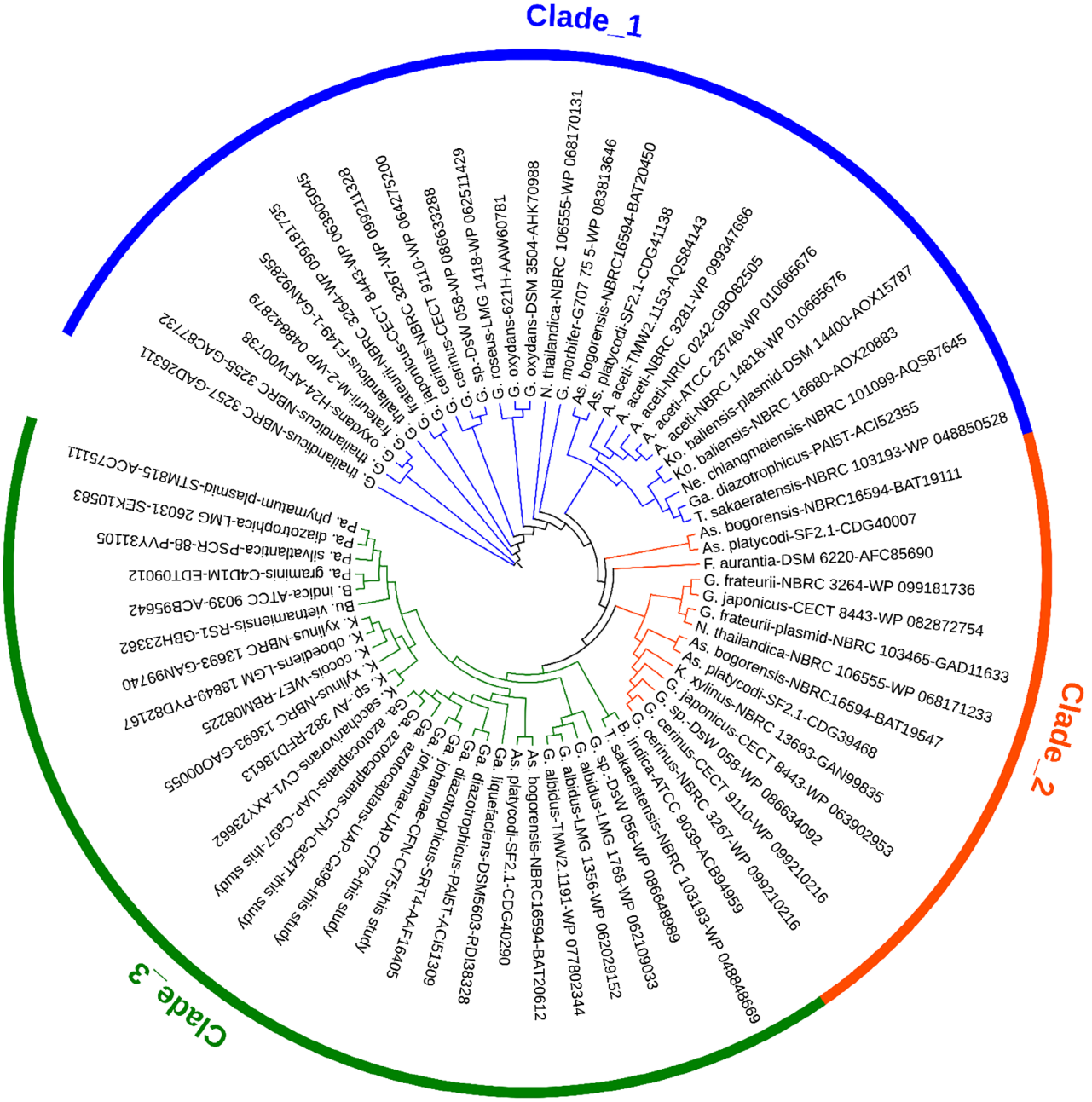

Fig. 6. Phylogenetic clustering of GH32 proteins identified in AAB carrying levansucrase genes. Entries are named as species, strain and sequence ID. Clades are shown with different branch colours and numbered in the outer ring. Clustal Omega (https://www.ebi.ac.uk/clustalo/) and Simple Phylogeny (https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/) were used for sequence alignment and tree generation respectively. Genera in the species names are abbreviated as follows: *A*, *Acetobacter*; *As*, *Asaia*; *B*, *Beijerinckia*; *Bo*; *Bombella*; *Bu*, *Burkholderia*; *E*, *Erwinia*; *F*, *Frateuria*; *Ga*, *Gluconacetobacter*; *G*, *Gluconobacter*; *H*, *Halomonas*; *K*, *Komagataeibacter*; *Ko*, *Kozakia*; *Ne*, *Neoasaia*; *N*, *Neokomagataea*; *Pa*, *Paraburkholderia*; *Ps*, *Pseudomonas*; *S*, *Saccharibacter*; *T*, *Tanticharoenia*; *Z*, *Zymomonas*. [Color figure can be viewed at wileyonlinelibrary.com]

occupation of specific sucrose containing habitats. Accordingly, T1-LS are secreted by proteobacteria known to be plant-pathogenic, while T2-LS genes are common in endophytic diazotrophs. LS could thus be regarded as useful ecological determinants for predictions of the role and host specificity of AAB and other (so far non-characterized) fructan producers in nature. This finding is a key step towards studying the underlying mechanisms of microbial biofilm formation via evolutionary adapted sucrases in plant environments and sucrose-feeding animal hosts.

Methods

Strains, media and cultivation conditions

Gluconobacter strains were cultivated aerobically at 30°C in/on sodium gluconate (NaG) media (20 g/l sodium gluconate, 3 g/l yeast extract, 2 g/l peptone, 3 g/l glycerol, 10 g/l mannitol, optional 80 g/l sucrose for levan production and 20 g/l agar). *Gluconobacter* strains marked by 'TMW' (Technical Microbiology Weihenstephan) were isolated on solid NaG media from water kefir (TMW 2.1191, TMW 2.767) or spoiled beer samples (TMW 2.1085, TMW 2.155, TMW 2.1555, TMW 2.1556, TMW 2.1557, TMW 2.1558, TMW 2.339). *Gluconacetobacter* strains were cultivated aerobically at 30°C in/on SYP medium modified by increasing the amount of yeast extract to 0.3% (w/v) (Caballero-Mellado and Martinez-Romero, 1994).

Genetic screening for AAB levansucrases and strain identifications

AAB strains marked by 'TMW' were identified via 16S rRNA gene amplification/sequencing (primers 07 Forward: AGA GTT TGA TCC TGG CTC AG + 1507 Reverse: TAC CTT GTT ACG ACT TCA C; Lane, 1991). *Gluconobacter* and *Gluconacetobacter* isolates from beer, kefir and plant were investigated regarding the presence of GH68 gene sequences. DNA was isolated from these strains following the instructions of the E.Z.N.A. Bacterial DNA Kit (Omega Biotek, Norcross) and used as template for PCR amplification of the respective (partial) target genes. Two primer sets were deduced from multiple nucleotide alignments of known levansucrase genes from AAB (Jakob *et al*., 2012b), targeting either against the T1- or T2-like levansucrase types (T1: GAT CCG ACR ACG ACV ATG CC (forward)/TAB GGR CCG AAA ATN CCS TT (reverse) and T2: TAT AAY GGN TGG GAD GTB AT (forward)/GGC ATG ACR TAR TGC GAR TA (reverse). PCR conditions were as follows: initial denaturation (3 min, 94°C); 35 cycles: denaturation (30 s, 94°C), annealing (1 min, 52°C), extension (1 min, 72°C), final elongation (10 min, 72°C). Reaction mixtures generally were prepared according to the suppliers' instructions (*Taq*-DNA-Polymerase-Kit, Qbiogene). Clear bands of the expected size (16S: ~1500 bp; T1: ~900 bp; T2: ~1000 bp) were cut out of the gels, following preparative DNA extraction (peqGOLD Gel extraction Kit, peqlab, Germany) and subsequent sequencing (GATC Biotech, Konstanz, Germany).

Data analysis and processing

Public available gene sequences were collected from NCBI either by manual search for the respective sequence types or by nucleotide and protein BLAST search (using complete GH 68/32 sequences from AAB as template) against deposited genomes of AAB or proteobacteria, which share similar nucleotide sequences coding for putative GH 68/32 proteins. Dendrograms were generated on the basis of multiple alignments using BioNumerics 6.50 software (Applied Maths, Belgium).

The SignalP 5.0 server (http://www.cbs.dtu.dk/services/SignalP) was used to predict the presence of signal peptides and the location of their cleavage sites in proteins. The molecular weight and isoelectric point of the predicted mature proteins were determined using PROTEIN CALCULATOR v3.4 (http://protcalc.sourceforge.net). Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) was used for multiple sequence alignments. Simple Phylogeny (https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/) was used for generating phylogenetic trees in a radial format. The available crystal structures of T1-LS from *Erwinia amylovora* (PDB code: 4D47) were used as template for molecular modelling of T1-LS from *Gluconobacter albidus* (Protein Id. WP_082780180). The 3D structure model was constructed using the SWISS-MODEL tools (Biasini *et al*., 2014).

Acknowledgements

Part of this work was supported by funds of the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office of Agriculture and Food (BLE) in Project 2816IP001.

References

Arrieta, J., Hernandez, L., Coego, A., Suárez, V., Balmori, E., Menéndez, C., *et al*. (1996) Molecular characterization of the levansucrase gene from the endophytic sugarcane bacterium *Acetobacter diazotrophicus* SRT4. *Microbiology* **142**: 1077–1085.

Arrieta, J.G., Sotolongo, M., Menéndez, C., Alfonso, D., Trujillo, L.E., Soto, M., *et al*. (2004) A type II protein secretory pathway required for levansucrase secretion by

Gluconacetobacter diazotrophicus. *J Bacteriol* **186**: 5031–5039.

Avigad, G. (1982) Sucrose and other disaccharides. In *Plant Carbohydrates I*. Berlin: Springer, pp. 217–347.

Barata, A., Malfeito-Ferreira, M., and Loureiro, V. (2012) The microbial ecology of wine grape berries. *Int J Food Microbiol* **153**: 243–259.

Becking, J.H. (2006) The genus *Beijerinckia*. In *The Prokaryotes: Volume 5: Proteobacteria: Alpha and Beta Subclasses*, pp. 151–162. New York, NY: Springer.

Bereswill, S., Jock, S., Aldridge, P., Janse, J., and Geider, K. (1997) Molecular characterization of natural *Erwinia amylovora* strains deficient in levan synthesis. *Physiol Mol Plant Pathol* **51**: 215–225.

Betancourt, L., Takao, T., Hernandez, L., Padron, G., and Shimonishi, Y. (1999) Structural characterization of *Acetobacter diazotropicus* levansucrase by matrix-assisted laser desorption/ionization mass spectrometry: identification of an N-terminal blocking group and a free-thiol cysteine residue. *J Mass Spectrom* **34**: 169–174.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., *et al*. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**: W252–W258.

Brandt, J.U., Born, F.-L., Jakob, F., and Vogel, R.F. (2017) Environmentally triggered genomic plasticity and capsular polysaccharide formation are involved in increased ethanol and acetic acid tolerance in *Kozakia baliensis* NBRC 16680. *BMC Microbiol* **17**: 172.

Brandt, J.U., Jakob, F., Behr, J., Geissler, A.J., and Vogel, R.F. (2016) Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*. *Microb Cell Fact* **15**: 170.

Brandt, J.U., Jakob, F., Wefers, D., Bunzel, M., and Vogel, R.F. (2018) Characterization of an acetan-like heteropolysaccharide produced by *Kozakia baliensis* NBRC 16680. *Int J Biol Macromol* **106**: 248–257.

Caballero-Mellado, J., and Martinez-Romero, E. (1994) Limited genetic diversity in the endophytic sugarcane bacterium *Acetobacter diazotrophicus*. *Appl Environ Microbiol* **60**: 1532–1537.

Cavalcante, V.A., and Dobereiner, J. (1988) A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane. *Plant Soil* **108**: 23–31.

Chouaia, B., Rossi, P., Epis, S., Mosca, M., Ricci, I., Damiani, C., *et al*. (2012) Delayed larval development in *Anopheles* mosquitoes deprived of *Asaia* bacterial symbionts. *BMC Microbiol* **12**: S2.

Corby-Harris, V., and Anderson, K.E. (2018) Draft genome sequences of four *Parasaccharibacter apium* strains isolated from honey bees. *Genome Announce* **6**: e00165.

Corby-Harris, V., Snyder, L.A., Schwan, M.R., Maes, P., McFrederick, Q.S., and Anderson, K.E. (2014) Origin and effect of alpha 2.2 Acetobacteraceae in honey bee larvae and description of *Parasaccharibacter apium* gen. nov., sp. nov. *Appl Environ Microbiol* **80**: 7460–7472.

Corby-Harris, V., Snyder, L., Meador, C., Naldo, R., Mott, B., and Anderson, K. (2016) *Parasaccharibacter apium*, gen. nov., sp. nov., improves honey bee (Hymenoptera: Apidae) resistance to Nosema. *J Econ Entomol* **109**: 537–543.

Crotti, E., Rizzi, A., Chouaia, B., Ricci, I., Favia, G., Alma, A., *et al*. (2010) Acetic acid bacteria, newly emerging symbionts of insects. *Appl Environ Microbiol* **76**: 6963–6970.

Damiani, C., Ricci, I., Crotti, E., Rossi, P., Rizzi, A., Scuppa, P., *et al*. (2010) Mosquito-bacteria symbiosis: the case of *Anopheles gambiae* and *Asaia*. *Microb Ecol* **60**: 644–654.

De Roos, J., and De Vuyst, L. (2018) Acetic acid bacteria in fermented foods and beverages. *Curr Opin Biotechnol* **49**: 115–119.

Deppenmeier, U., and Ehrenreich, A. (2009) Physiology of acetic acid bacteria in light of the genome sequence of *Gluconobacter oxydans*. *J Mol Microbiol Biotechnol* **16**: 69–80.

Deppenmeier, U., Hoffmeister, M., and Prust, C. (2002) Biochemistry and biotechnological applications of *Gluconobacter* strains. *Appl Microbiol Biotechnol* **60**: 233–242.

Dogsa, I., Brloznik, M., Stopar, D., and Mandic-Mulec, I. (2013) Exopolymer diversity and the role of levan in *Bacillus subtilis* biofilms. *PLoS One* **8**: e62044.

Du, Z., and Geider, K. (2002) Characterization of an activator gene upstream of lsc, involved in levan synthesis of *Erwinia amylovora*. *Physiol Mol Plant Pathol* **60**: 9–17.

Estrada-De Los Santos, P., Bustillos-Cristales, R.o., and Caballero-Mellado, J. (2001) *Burkholderia*, a genus rich in plant-associated nitrogen fixers with wide environmental and geographic distribution. *Appl Environ Microbiol* **67**: 2790–2798.

Favia, G., Ricci, I., Damiani, C., Raddadi, N., Crotti, E., Marzorati, M., *et al*. (2007) Bacteria of the genus *Asaia* stably associate with *Anopheles stephensi*, an Asian malarial mosquito vector. *Proc Natl Acad Sci U S A* **104**: 9047–9051.

Fels, L., Jakob, F., Vogel, R.F., and Wefers, D. (2018) Structural characterization of the exopolysaccharides from water kefir. *Carbohydr Polym* **189**: 296–303.

Fuentes-Ramírez, L.E., Bustillos-Cristales, R., Tapia-Hernández, A., Jiménez-Salgado, T., Wang, E.T., Martínez-Romero, E., and Caballero-Mellado, J. (2001) Novel nitrogen-fixing acetic acid bacteria, *Gluconacetobacter johannae* sp. nov. and *Gluconacetobacter azotocaptans* sp. nov., associated with coffee plants. *Int J Syst Evol Microbiol* **51**: 1305–1314.

Gross, M., and Rudolph, K. (1987) Demonstration of levan and alginate in bean plants (*Phaseolus vulgaris*) infected by *Pseudomonas syringae pv*. *phaseolicola*. *J Phytopathol* **120**: 9–19.

Gross, M., Geier, G., Rudolph, K., and Geider, K. (1992) Levan and levansucrase synthesized by the fireblight pathogen *Erwinia amylovora*. *Physiol Mol Plant Pathol* **40**: 371–381.

Gulitz, A., Stadie, J., Wenning, M., Ehrmann, M.A., and Vogel, R.F. (2011) The microbial diversity of water kefir. *Int J Food Microbiol* **151**: 284–288.

Gullo, M., and Giudici, P. (2008) Acetic acid bacteria in traditional balsamic vinegar: phenotypic traits relevant for starter cultures selection. *Int J Food Microbiol* **125**: 46–53.

Gunasekaran, P., Mukundan, G., Kannan, R., Velmurugan, S., Ait-Abdelkader, N., Alvarez-Macarie, E., and Baratti, J. (1995) The sacB and sacC genes encoding levansucrase and sucrase form a gene cluster in *Zymomonas mobilis*. *Biotechnol Lett* **17**: 635–642.

He, M., Jiang, J., and Cheng, D. (2017) The plant pathogen *Gluconobacter cerinus* strain CDF1 is beneficial to the fruit fly *Bactrocera dorsalis*. *AMB Express* **7**: 207.

Hernández, L., Arrieta, J., Betancourt, L., Falcón, V., Madrazo, J., Coego, A., and Menéndez, C. (1999) Levansucrase from *Acetobacter diazotrophicus* SRT4 is secreted via periplasm by a signal-peptide-dependent pathway. *Curr Microbiol* **39**: 146–152.

Hernández, L., Sotolongo, M., Rosabal, Y., Menéndez, C., Ramírez, R., Caballero-Mellado, J., and Arrieta, J. (2000) Structural levansucrase gene (lsdA) constitutes a functional locus conserved in the species *Gluconacetobacter diazotrophicus*. *Arch Microbiol* **174**: 120–124.

Hettwer, U., Gross, M., and Rudolph, K. (1995) Purification and characterization of an extracellular levansucrase from *Pseudomonas syringae pv*. *phaseolicola*. *J Bacteriol* **177**: 2834–2839.

Jakob, F. (2014) *Novel Fructans from Acetic Acid Bacteria*. Munich: Technische Universität München.

Jakob, F., Steger, S., and Vogel, R.F. (2012a) Influence of novel fructans produced by selected acetic acid bacteria on the volume and texture of wheat breads. *Eur Food Res Technol* **234**: 493–499.

Jakob, F., Meißner, D., and Vogel, R.F. (2012b) Comparison of novel GH 68 levansucrases of levan-overproducing *Gluconobacter* species. *Acetic Acid Bacteria* **1**: e2–e2.

Jakob, F., Pfaff, A., Novoa-Carballal, R., Rübsam, H., Becker, T., and Vogel, R.F. (2013) Structural analysis of fructans produced by acetic acid bacteria reveals a relation to hydrocolloid function. *Carbohydr Polym* **92**: 1234–1242.

Jimenez-Salgado, T., Fuentes-Ramirez, L.E., Tapia-Hernandez, A., Mascarua-Esparza, M.A., Martinez-Romero, E., and Caballero-Mellado, J. (1997) *Coffea arabica* L., a new host plant for *Acetobacter diazotrophicus*, and isolation of other nitrogen-fixing acetobacteria. *Appl Environ Microbiol* **63**: 3676–3683.

Jojima, Y., Mihara, Y., Suzuki, S., Yokozeki, K., Yamanaka, S., and Fudou, R. (2004) *Saccharibacter floricola* gen. nov., sp. nov., a novel osmophilic acetic acid bacterium isolated from pollen. *Int J Syst Evol Microbiol* **54**: 2263–2267.

Kasapis, S., Morris, E.R., Gross, M., and Rudolph, K. (1994) Solution properties of levan polysaccharide from *Pseudomonas syringae pv*. *phaseolicola*, and its possible primary role as a blocker of recognition during pathogenesis. *Carbohydr Polym* **23**: 55–64.

Kato, N., Mizuno, M., Nakai, Y., Nozaki, K., Suga, H., Kanda, T., *et al*. (2007) Structural analysis of the water-soluble carbohydrate from *Asaia bogorensis* by NMR spectroscopy. *J Appl Glycosci* **54**: 231–233.

Katzen, F., Ferreiro, D.U., Oddo, C.G., Ielmini, M.V., Becker, A., Pühler, A., and Ielpi, L. (1998) *Xanthomonas campestris pv*. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *J Bacteriol* **180**: 1607–1617.

Koczan, J.M., McGrath, M.J., Zhao, Y., and Sundin, G.W. (2009) Contribution of *Erwinia amylovora* exopolysaccharides amylovoran and Levan to biofilm formation: implications in pathogenicity. *Phytopathology* **99**: 1237–1244.

Komagata, K., Iino, T., and Yamada, Y. (2014) The family Acetobacteraceae. In *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, pp. 3–78. New York, NY: Springer.

Kregiel, D., Rygala, A., Libudzisz, Z., Walczak, P., and Oltuszak-Walczak, E. (2012) *Asaia lannensis*–the spoilage acetic acid bacteria isolated from strawberry-flavored bottled water in Poland. *Food Control* **26**: 147–150.

Kremer, T.A., LaSarre, B., Posto, A.L., and McKinlay, J.B. (2015) N2 gas is an effective fertilizer for bioethanol production by *Zymomonas mobilis*. *Proc Natl Acad Sci U S A* **112**: 2222–2226.

Kumar, V., Behl, R.K., and Narula, N. (2001) Establishment of phosphate-solubilizing strains of *Azotobacter chroococcum* in the rhizosphere and their effect on wheat cultivars under green house conditions. *Microbiol Res* **156**: 87–93.

Laue, H., Schenk, A., Li, H., Lambertsen, L., Neu, T.R., Molin, S., and Ullrich, M.S. (2006) Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. *Microbiology* **152**: 2909–2918.

Laureys, D., and De Vuyst, L. (2014) Microbial species diversity, community dynamics, and metabolite kinetics of water kefir fermentation. *Appl Environ Microbiol* **80**: 2564–2572.

Li, H., and Ullrich, M.S. (2001) Characterization and mutational analysis of three allelic lsc genes encoding Levansucrase in *Pseudomonas syringae*. *J Bacteriol* **183**: 3282–3292.

Li, L., Praet, J., Borremans, W., Nunes, O.C., Manaia, C.M., Cleenwerck, I., *et al*. (2015) *Bombella intestini* gen. nov., sp. nov., an acetic acid bacterium isolated from bumble bee crop. *Int J Syst Evol Microbiol* **65**: 267–273.

Lisdiyanti, P., Kawasaki, H., Widyastuti, Y., Saono, S., Seki, T., Yamada, Y., *et al*. (2002) *Kozakia baliensis* gen. nov., sp. nov., a novel acetic acid bacterium in the alpha-proteobacteria. *Int J Syst Evol Microbiol* **52**: 813–818.

Loganathan, P., and Nair, S. (2004) *Swaminathania salitolerans* gen. nov., sp. nov., a salt-tolerant, nitrogen-fixing and phosphate-solubilizing bacterium from wild rice (Porteresia coarctata Tateoka). *Int J Syst Evol Microbiol* **54**: 1185–1190.

Martínez-Fleites, C., Ortíz-Lombardía, M., Pons, T., Tarbouriech, N., Taylor, E.J., Arrieta, J.G., *et al*. (2005) Crystal structure of levansucrase from the Gram-negative bacterium *Gluconacetobacter diazotrophicus*. *Biochem J* **390**: 19–27.

Matsushita, K., Toyama, H., and Adachi, O. (1994) Respiratory chains and bioenergetics of acetic acid bacteria. In *Advances in Microbial Physiology*, pp. 247–301. Amsterdam: Elsevier.

Matsutani, M., Hirakawa, H., Yakushi, T., and Matsushita, K. (2011) Genome-wide phylogenetic analysis of *Gluconobacter*, *Acetobacter*, and *Gluconacetobacter*. *FEMS Microbiol Lett* **315**: 122–128.

Menéndez, C., Hernández, L., Banguela, A., and País, J. (2004) Functional production and secretion of the *Gluconacetobacter diazotrophicus* fructose-releasing exolevanase (LsdB) in *Pichia pastoris*. *Enzyme Microb Technol* **34**: 446–452.

Menéndez, C., Banguela, A., Caballero-Mellado, J., and Hernández, L. (2009) Transcriptional regulation and signal-peptide-dependent secretion of exolevanase (LsdB) in the endophyte *Gluconacetobacter diazotrophicus*. *Appl Environ Microbiol* **75**: 1782–1785.

Menéndez, C., Hernández, L., Selman, G., Mendoza, M.F., Hevia, P., Sotolongo, M., and Arrieta, J.G. (2002) Molecular

cloning and expression in *Escherichia coli* of an exo-levanase gene from the endophytic bacterium *Gluconacetobacter diazotrophicus* SRT4. *Curr Microbiol* **45**: 5–12.

Meneses, C.H., Rouws, L.F., Simões-Araújo, J.L., Vidal, M. S., and Baldani, J.I. (2011) Exopolysaccharide production is required for biofilm formation and plant colonization by the nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus*. *Mol Plant Microbe Interact* **24**: 1448–1458.

Meng, G., and Fütterer, K. (2003) Structural framework of fructosyl transfer in *Bacillus subtilis* levansucrase. *Nat Struct Mol Biol* **10**: 935–941.

Moore, J.E., McCalmont, M., Xu, J., Millar, B.C., and Heaney, N. (2002) *Asaia* sp., an unusual spoilage organism of fruit-flavored bottled water. *Appl Environ Microbiol* **68**: 4130–4131.

Mukasa, H., and Slade, H.D. (1973) Mechanism of adherence of *Streptococcus mutans t*o smooth surfaces I. Roles of insoluble dextran-levan synthetase enzymes and cell wall polysaccharide antigen in plaque formation. *Infect Immun* **8**: 555–562.

Mullins, E.A., Francois, J.A., and Kappock, T.J. (2008) A specialized citric acid cycle requiring succinyl-coenzyme A (CoA): acetate CoA-transferase (AarC) confers acetic acid resistance on the acidophile *Acetobacter aceti*. *J Bacteriol* **190**: 4933–4940.

Öner, E.T., Hernández, L., and Combie, J. (2016) Review of levan polysaccharide: from a century of past experiences to future prospects. *Biotechnol Adv* **34**: 827–844.

Osman, S., Fett, W., and Fishman, M. (1986) Exopolysaccharides of the phytopathogen *Pseudomonas syringae pv*. *glycinea*. *J Bacteriol* **166**: 66–71.

Pedraza, R.O. (2008) Recent advances in nitrogen-fixing acetic acid bacteria. *Int J Food Microbiol* **125**: 25–35.

Peters, B., Junker, A., Brauer, K., Mühlthaler, B., Kostner, D., Mientus, M., *et al*. (2013) Deletion of pyruvate decarboxylase by a new method for efficient markerless gene deletions in *Gluconobacter oxydans*. *Appl Microbiol Biotechnol* **97**: 2521–2530.

Pijning, T., Anwar, M.A., Böger, M., Dobruchowska, J.M., Leemhuis, H., Kralj, S., *et al*. (2011) Crystal structure of inulosucrase from *Lactobacillus*: insights into the substrate specificity and product specificity of GH68 fructansucrases. *J Mol Biol* **412**: 80–93.

Prechtl, R.M., Janssen, D., Behr, J., Ludwig, C., Küster, B., Vogel, R.F., and Jakob, F. (2018) Sucrose-induced proteomic response and carbohydrate utilization of *Lactobacillus sakei* TMW 1.411 during dextran formation. *Front Microbiol* **9**, 2796.

Prust, C., Hoffmeister, M., Liesegang, H., Wiezer, A., Fricke, W.F., Ehrenreich, A., *et al*. (2005) Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. *Nat Biotechnol* **23**: 195–200.

Ramírez-Bahena, M.H., Tejedor, C., Martín, I., Velázquez, E., and Peix, A. (2013) *Endobacter medicaginis* gen. nov., sp. nov., isolated from alfalfa nodules in an acidic soil. *Int J Syst Evol Microbiol* **63**: 1760–1765.

Raspor, P., and Goranovič, D. (2008) Biotechnological applications of acetic acid bacteria. *Crit Rev Biotechnol* **28**: 101–124.

Reid, S.J., and Abratt, V.R. (2005) Sucrose utilisation in bacteria: genetic organisation and regulation. *Appl Microbiol Biotechnol* **67**: 312–321.

Ruiz-Argueso, T., and Rodriguez-Navarro, A. (1975) Microbiology of ripening honey. *Appl Environ Microbiol* **30**: 893–896.

Samaddar, N., Paul, A., Chakravorty, S., Chakraborty, W., Mukherjee, J., Chowdhuri, D., and Gachhui, R. (2011) Nitrogen fixation in *Asaia* sp. (family Acetobacteraceae). *Curr Microbiol* **63**: 226–231.

Saravanan, V., Madhaiyan, M., Osborne, J., Thangaraju, M., and Sa, T. (2008) Ecological occurrence of *Gluconacetobacter diazotrophicus* and nitrogen-fixing Acetobacteraceae members: their possible role in plant growth promotion. *Microb Ecol* **55**: 130–140.

Semjonovs, P., Shakirova, L., Treimane, R., Shvirksts, K., Auzina, L., Cleenwerck, I., and Zikmanis, P. (2016) Production of extracellular fructans by *Gluconobacter nephelii* P1464. *Lett Appl Microbiol* **62**: 145–152.

Swings, J., and De Ley, J. (1977) The biology of *Zymomonas*. *Bacteriol Rev* **41**: 1.

Swings, J., and De Ley, J. (1981) The genera *Gluconobacter* and *Acetobacter*. In *The Prokaryotes*. Berlin: Springer, pp. 771–778.

Swings, J., Gillis, M., Kersters, K., De Vos, P., Gosselé, F., and De Ley, J. (1980) *Frateuria*, a new genus for "Acetobacter aurantius". *Int J Syst Evol Microbiol* **30**: 547–556.

Tajima, K., Uenishi, N., Fujiwara, M., Erata, T., Munekata, M., and Takai, M. (1997) The production of a new water-soluble polysaccharide by *Acetobacter xylinum* NCI 1005 and its structural analysis by NMR spectroscopy. *Carbohydr Res* **305**: 117–122.

Trujillo, L., Arrieta, J., Dafhnis, F., Garcıa, J., Valdes, J., Tambara, Y., *et al*. (2001) Fructo-oligosaccharides production by the *Gluconacetobacter diazotrophicus* levansucrase expressed in the methylotrophic yeast *Pichia pastoris*. *Enzyme Microb Technol* **28**: 139–144.

Ua-Arak, T., Jakob, F., and Vogel, R.F. (2017) Fermentation pH modulates the size distributions and functional properties of *Gluconobacter albidus* TMW 2.1191 levan. *Front Microbiol* **8**: 807.

Urakami, T., Tamaoka, J., Suzuki, K.-I., and Komagata, K. (1989) *Acidomonas* gen. nov., incorporating *Acetobacter methanolicus* as *Acidomonas methanolica* comb. nov. *Int J Syst Evol Microbiol* **39**: 50–55.

van Hijum, S.A., Kralj, S., Ozimek, L.K., Dijkhuizen, L., and van Geel-Schutten, I.G. (2006) Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol Mol Biol Rev* **70**: 157–176.

Vu, H.T.L., Yukphan, P., Chaipitakchonlatarn, W., Malimas, T., Muramatsu, Y., Bui, U.T.T., *et al*. (2013) *Nguyenibacter vanlangensis* gen. nov., sp. nov., an unusual acetic acid bacterium in the α-Proteobacteria. *J Gen Appl Microbiol* **59**: 153–166.

Wuerges, J., Caputi, L., Cianci, M., Boivin, S., Meijers, R., and Benini, S. (2015) The crystal structure of *Erwinia amylovora* levansucrase provides a snapshot of the products of sucrose hydrolysis trapped into the active site. *J Struct Biol* **191**: 290–298.

Yakushi, T., and Matsushita, K. (2010) Alcohol dehydrogenase of acetic acid bacteria: structure, mode of action, and applications in biotechnology. *Appl Microbiol Biotechnol* **86**: 1257–1265.

Yamada, Y. (2016) Systematics of acetic acid bacteria. In *Acetic Acid Bacteria*. Berlin: Springer, pp. 1–50.
Yamada, Y., and Yukphan, P. (2008) Genera and species in acetic acid bacteria. *Int J Food Microbiol* **125**: 15–24.
Yukphan, P., Malimas, T., Potacharoen, W., Tanasupawat, S., Tanticharoen, M., and Yamada, Y. (2005) *Neoasaia chiangmaiensis* gen. nov., sp. nov., a novel osmotolerant acetic acid bacterium in the α-Proteobacteria. *J Gen Appl Microbiol* **51**: 301–311.
Yukphan, P., Malimas, T., Muramatsu, Y., Potacharoen, W., Tanasupawat, S., Nakagawa, Y., *et al*. (2011) *Neokomagataea* gen. nov., with descriptions of *Neokomagataea thailandica* sp. nov. and *Neokomagataea tanensis* sp. nov., osmotolerant acetic acid bacteria of the α-Proteobacteria. *Biosci Biotechnol Biochem* **75**: 419–426.
Yukphan, P., Malimas, T., Muramatsu, Y., Takahashi, M., Kaneyasu, M., Tanasupawat, S., *et al*. (2008) *Tanticharoenia sakaeratensis* gen. nov., sp. nov., a new osmotolerant acetic acid bacterium in the α-Proteobacteria. *Biosci Biotechnol Biochem* **72**: 672–676.
Yukphan, P., Malimas, T., Muramatsu, Y., Takahashi, M., Kaneyasu, M., Potacharoen, W., *et al*. (2009) *Ameyamaea chiangmaiensis* gen. nov., sp. nov., an acetic acid bacterium in the α-Proteobacteria. *Biosci Biotechnol Biochem* **73**: 2156–2162.
Yun, J.-H., Lee, J.-Y., Hyun, D.-W., Jung, M.-J., and Bae, J.-W. (2017) *Bombella apis* sp. nov., an acetic acid bacterium isolated from the midgut of a honey bee. *Int J Syst Evol Microbiol* **67**: 2184–2188.
Zhang, Y., and Geider, K. (1999) Molecular analysis of the rlsA gene regulating levan production by the fireblight pathogen *Erwinia amylovora*. *Physiol Mol Plant Pathol* **54**: 187–201.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

File S1. Phylogenetic clustering of GH68 proteins from AAB and other proteobacteria. Levansucrase (LS) sequences were retrieved from the NCBI non-redundant protein database. Colours in entries (species name and sequence ID) represent the proteobacteria classes alpha (green), beta (blue) and gamma (yellow). T1-LSs and T2-LSs are clustered in two separate clades as defined in the outer ring. Clustal Omega (https://www.ebi.ac.uk/clustalo/) and Simple Phylogeny (https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/) were used for sequence alignment and radial tree generation, respectively.

File S2. Distinctive traits between T1- and T2-LSs from AAB and other proteobacteria. The SignalP 5.0 server (http://www.cbs.dtu.dk/services/SignalP) was used to predict the presence of signal peptides and the location of their cleavage sites in proteins. The molecular weight and isoelectric point of the predicted mature proteins were determined using PROTEIN CALCULATOR v3.4 (http://protcalc.sourceforge.net).

File S3. Phylogenetic trees calculated on the basis of partial T2-LS gene (1000 bp) (**A**) and SacC gene sequences (900 bp) (**B**) using the neighbour-joining method. Scale-bars indicate numbers of changes per nucleotide. Numbers following species names represent the accession numbers of the corresponding T2-LS and SacC genes as deposited in GenBank. The T1-LS and GH 32 encoding genes of *G. oxydans* 621 H were used as outgroups, respectively.

File S4. Schematic overview of the genetic organization of LS in AAB and related proteobacteria (supplement to Fig. 5).