BMC Microbiology

Supplementary material

Functions of *prp* operon in carbon source metabolism and acetic acid resistance in acetic acid bacteria: a case study in *Acetobacter pasteurianus* CGMCC 1.41

Haoran Yang[1,2,3], Armin Ehrenreich[3]*, Wolfgang Liebl[3]*, Fusheng Chen[2]*

1. Agro-Product Processing Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China
2. Hubei International Scientific and Technological Cooperation Base of Traditional Fermented Foods, Huazhong Agricultural University, Wuhan, Hubei, China
3. Chair of Microbiology, Technical University of Munich, Freising, Germany

*Corresponding author: Armin Ehrenreich: aehrenr@tum.de;
Wolfgang Liebl: wliebl@tum.de;
Fusheng Chen: chenfs@mail.hzau.edu.cn

Table S1 Primers used for molecular cloning in markerless gene deletion

Primer name	Seq 5'to 3'*	Usage
pKOS6b_SLICE_fw	ACATGGCGATAGCTAGACTGG	Amplifying pKOS6b backbone for deletion vectors. 5' sequences should be added according to target genes
pKOS6b_SLICE_r	ACAACATACGAGCCGGAAG	
prpD-SL-FLA_fw	CTTCCGGCTCGTATGTTGTGTATAGGGCAACAGCAAG	Amplifying Flank A for Δ*prpDCBE'*, Δ*prpDCB* and Δ*prpD*
prpD-SL-FLA_r	CCTGTGCGTTATCGCCTGTTGTTTTTACCTCATGATCGG	Amplifying Flank A for Δ*prpD*
prpDCB-SL-FL-A_r	CGGGATAGAAAGAGAATGTTGTTTTTACCTCATGATCGG	Amplifying Flank A for Δ*prpDCB*
Op-SL-FL A_r	GTAAAACCTCGCTTTAAGGTGTTTTTACCTCATGATCGG	Amplifying Flank A for Δ*prpDCBE'*
prpC-SL-FL-A_fw	CTTCCGGCTCGTATGTTGTCAGCGGTTATCTATCAGG	Amplifying Flank A for Δ*prpC*
prpC-SL-FL-A_r	TGCCTACCAGATATGTCATTGTTCTATTCCTTCCTGTG	
prpD-SL-FLB_fw	CCGATCATGAGGTAAAAACAACAGGCGATAACGCACAGG	Amplifying Flank B for Δ*prpD*
prpD-SL-FLB_r	CCAGTCTAGCTATCGCCATGTCGCTCGGCAACATGATACAG	
prpB-SL-FL-A_fw	CTTCCGGCTCGTATGTTGTAGGTCGCGCATCTGCTTATC	Amplifying Flank A for Δ*prpB*
prpB-SL-FL-A_r	CGGGATAGAAAGAGAATGTTCATGCTGCATGCTCCAGG	
prpC-SL-FL-B_fw	CACAGGAAGGAATAGAACAATGACATATCTGGTAGGCA	Amplifying Flank B for Δ*prpC*
prpC-SL-FL-B_r	CCAGTCTAGCTATCGCCATGTACTGTGTAGGCAGGATAG	
prpE'-SL-FL A_fw	CTTCCGGCTCGTATGTTGTGCAGCATGACATATCTGG	Amplifying Flank A for Δ*prpE'*
prpE'-SL-FL A_r	GTAAAACCTCGCTTTAAGGTGTTCCGTCTCCCTTAC	
prpB-SL-FL B_fw	CCTGGAGCATGCAGCATGAACATTCTCTTTCTATCCCG	Amplifying Flank B for Δ*prpB*
prpDCB-SL-FL B_fw	CCGATCATGAGGTAAAAACAACATTCTCTTTCTATCCCG	Amplifying Flank B for Δ*prpDCB*
prpDCB-SL-FL-B_r	CCAGTCTAGCTATCGCCATGTTACACAACGGATGTATGTC	Amplifying Flank B for Δ*prpDCB* and Δ*prpB*
prpE'-SL-FL B_fw	GTAAGGGAGACGGAACACCTTAAAGCGAGGTTTTAC	Amplifying Flank B for Δ*prpE'*
Op-SL- FL B_fw	CCGATCATGAGGTAAAAACACCTTAAAGCGAGGTTTTAC	Amplifying Flank B for Δ*prpDCBE'*
Op-SL-FL B_r	CCAGTCTAGCTATCGCCATGTGGCTCATGAAATTGAAAG	Amplifying Flank B for Δ*prpDCBE'* and Δ*prpE'*

* Underlined part means the 5' extension for SLiCE

Table S2 Primers for verification processing markerless gene deletion

Primer name	Seq 5'to 3'	Usage
2-mccOpSeq-1_fw	CTGGCATCACTTCCCGTATC	Genomic specific forward check primer for entire *prp* operon, *prpDCB* and *prpD*
2-mccOpSeq-2_fw	TCGGCAGCACAATCGAGTTC	Sequencing primer for checking Δ*prpDCBE'*, Δ*prpDCB* and Δ*prpD*
2-mccOpSeq-3_fw	GGCCAAATGTGCGTGTG	Genomic specific forward check primer for *prpC*
2-mccOpSeq-4_fw	TGGCGCAAGATCCATAC	Sequencing primer for checking Δ*prpC* & Genomic specific forward check primer for *prpB*
prpE'_check_fw	CATTTACCGCCAGAGTAGTG	Genomic specific forward check primer for *prpE'*
prpD_check_r	TGGCGGATAAAGAAACACG	Genomic specific reverse check primer for *prpD*
delta_prpE'_check_fw	CCCGAAGCACTGATTTCTG	Sequencing primer for checking Δ*prpE'*
prpC_check_r	GCGCCGCTATCTATAATACG	Genomic specific reverse check primer for *prpC* & Sequencing primer for checking Δ*prpB*
prpB_check_r	CTGTGTATCGGCTATGACC	Genomic specific reverse check primer for *prpB* and *prpDCB*
delta_prpE'_check_r	GCCGTGTCATCAACGTAAG	Sequencing primer for checking Δ*prpE'*
2-mccOpSeq-10_r	GGTTGTCCGCCACCTTTAC	Genomic specific reverse check primer for *prpE'* and the entire *prp* operon
pKOS6b_check_fw	ATACGCAAACCGCCTCTC	pKOS6b specific checking primer pair
pKOS6b_check_r	AGGGCTTCCCAACCTTAC	

Table S3 Count of *acs* and *prpE'* genes in genera of AAB

Genera	Acs (PrpE') count	Genomes with Acs (PrpE') count	Genomes analyzed
Acetobacter	341 (134)	187 (132)	202
Ameyamaea	3 (0)	3 (0)	3
Brytella	1 (0)	1 (0)	1
Endobacter	3 (0)	3 (0)	3
Gluconacetobacter	11 (0)	10 (0)	30
Gluconobacter	113 (0)	113 (0)	114
Granulibacter	10 (0)	10 (0)	10
Komagataeibacter	7 (7)	7 (7)	60
Kozakia	5 (0)	5 (0)	5
Neoasaia	1 (0)	1 (0)	1
Nguyenibacter	1 (0)	1 (0)	1
Saccharibacter	5 (0)	5 (0)	5
Tanticharoenia	1 (0)	1 (0)	1
Other genera	0 (0)	0 (0)	75
Total	501 (141)	346 (139)	510

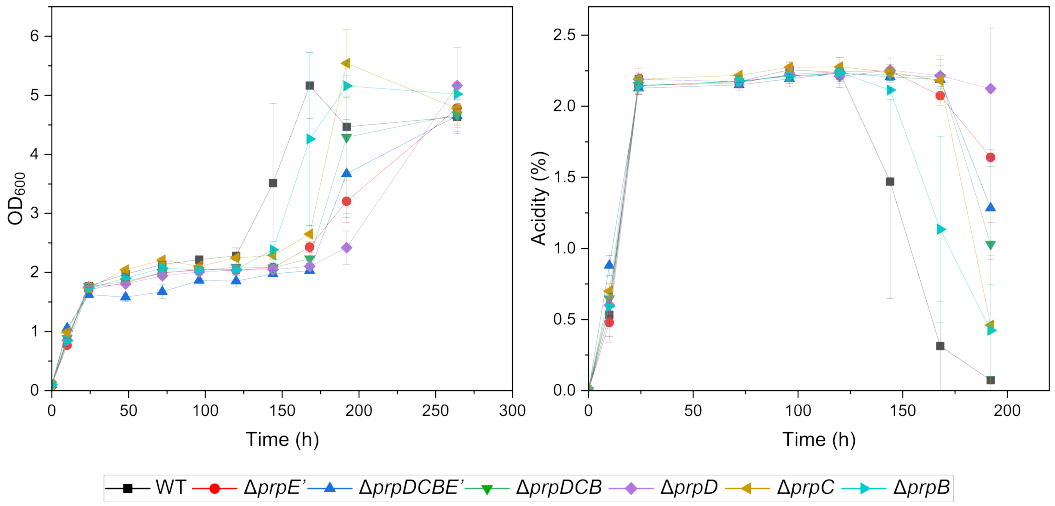

Fig. S1 Growth and acidity of all genotypes under the condition of 2% ethanol with 3 biological replicates

WT and all deletion mutants were inoculated into GYP liquid medium containing 2% ethanol with the initial OD_{600} value set as 0.1. Cultivation was carried out under 30°C and 200 r/min. Growth and acidity were measured regularly. The results of 3 biological replicates are shown in Fig. S1. After 24 h (1 d) of cultivation, all genotypes were able to complete the oxidation of 2% ethanol to the corresponding concentrations of acetic acid with growth entering the first stationary phase. From the 5th day (120 h) of the cultivation, the acetic acid overoxidation process and secondary growth of each strain could be observed successively. However, during this period, OD_{600} values as well as acidity had remarkable standard deviations, indicating that the time point at which the 3 samples from each strain initiated their acetic acid overoxidation, that is, the duration of the first stationary phase between the biological replicates of the same genotype, were largely different. As the existence of first stationary phase can be to some extent regarded as the adversity of acetic acid imposed on the cells, this uncertainty further reflects the unstable resistance against acetic acid of the strains.

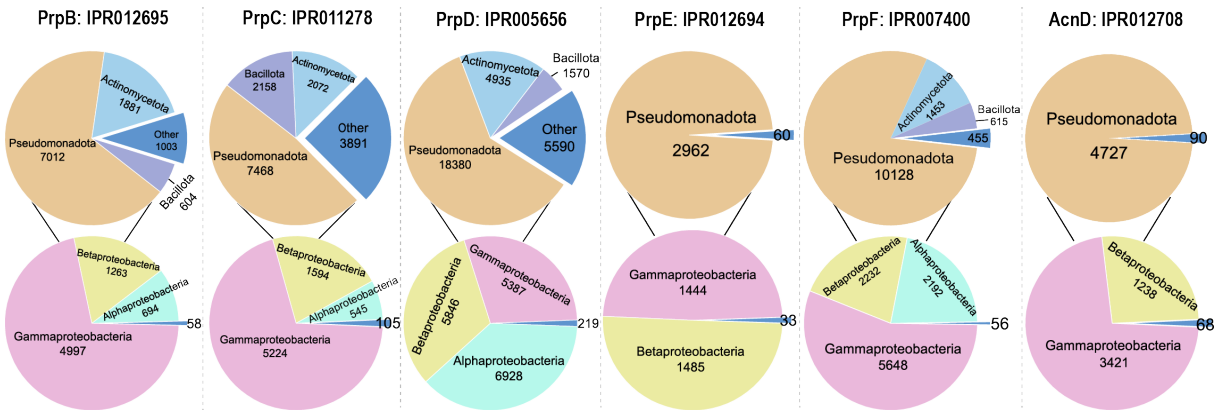

Fig.S2 Distribution of sequences translated from *prp* operon genes in bacteria that are included in the Uniprot database

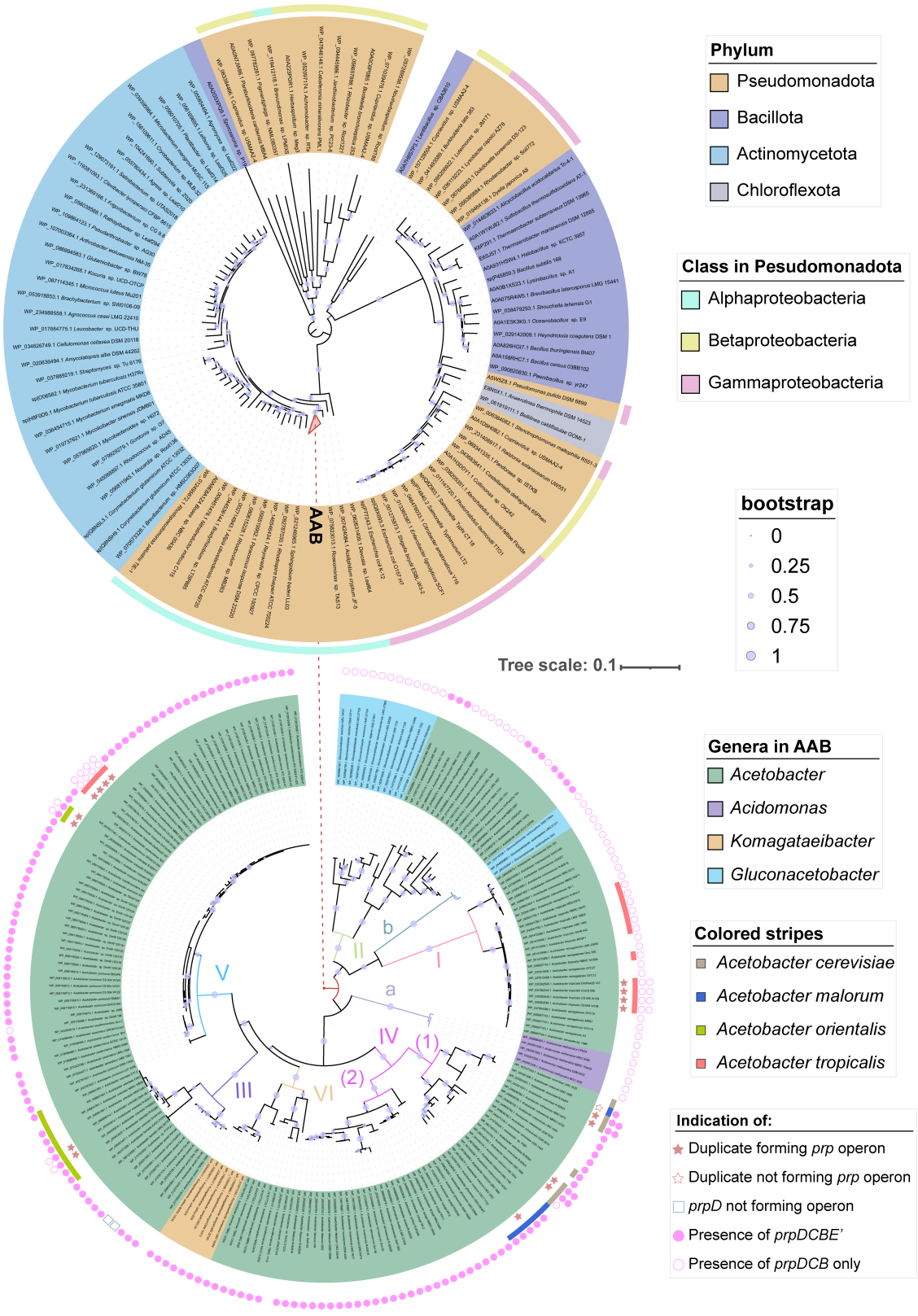

Fig. S3-A Phylogenic tree of PrpD from AAB and other bacteria

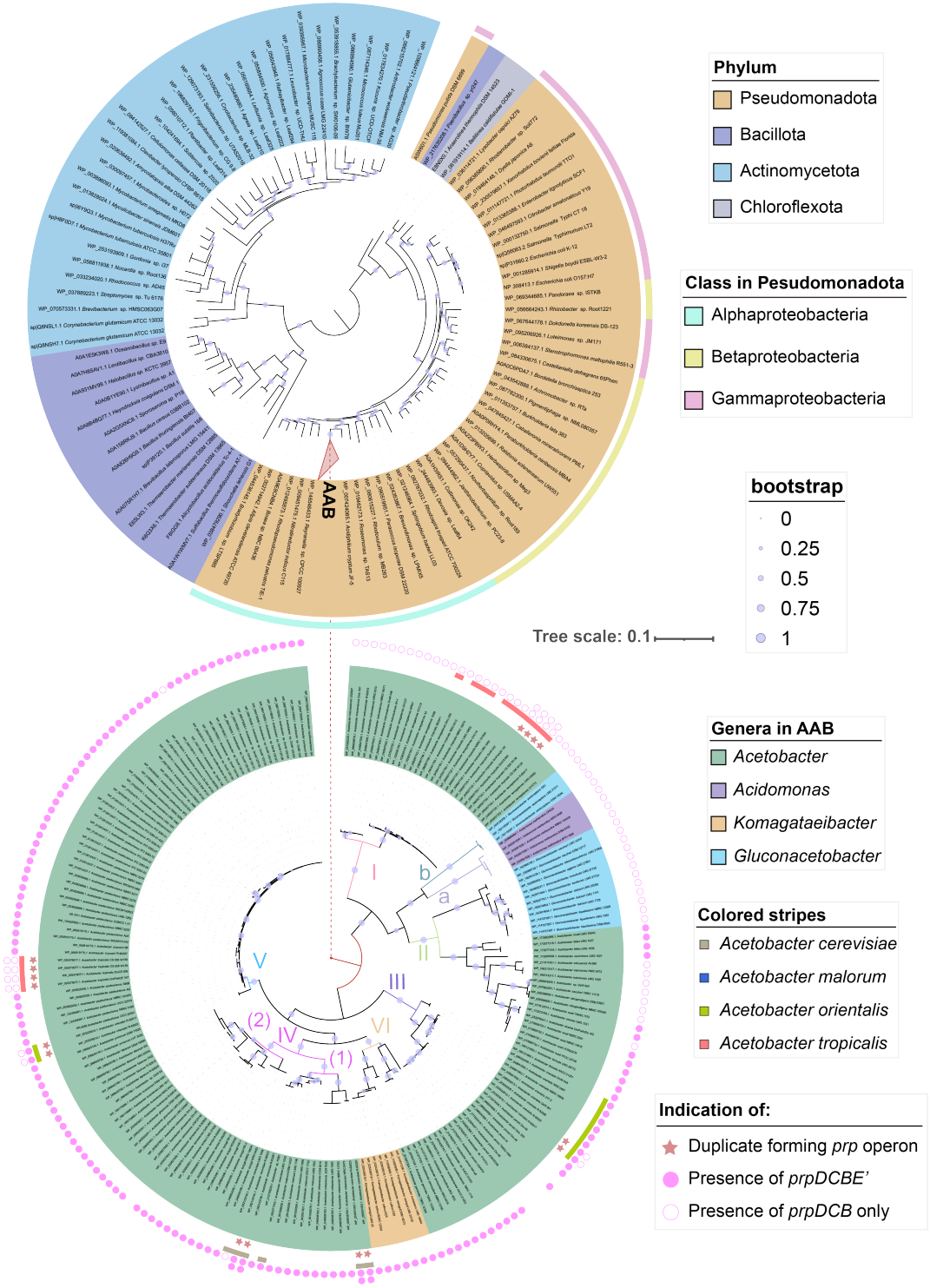

Fig. S3-B Phylogenic tree of PrpC from AAB and other bacteria

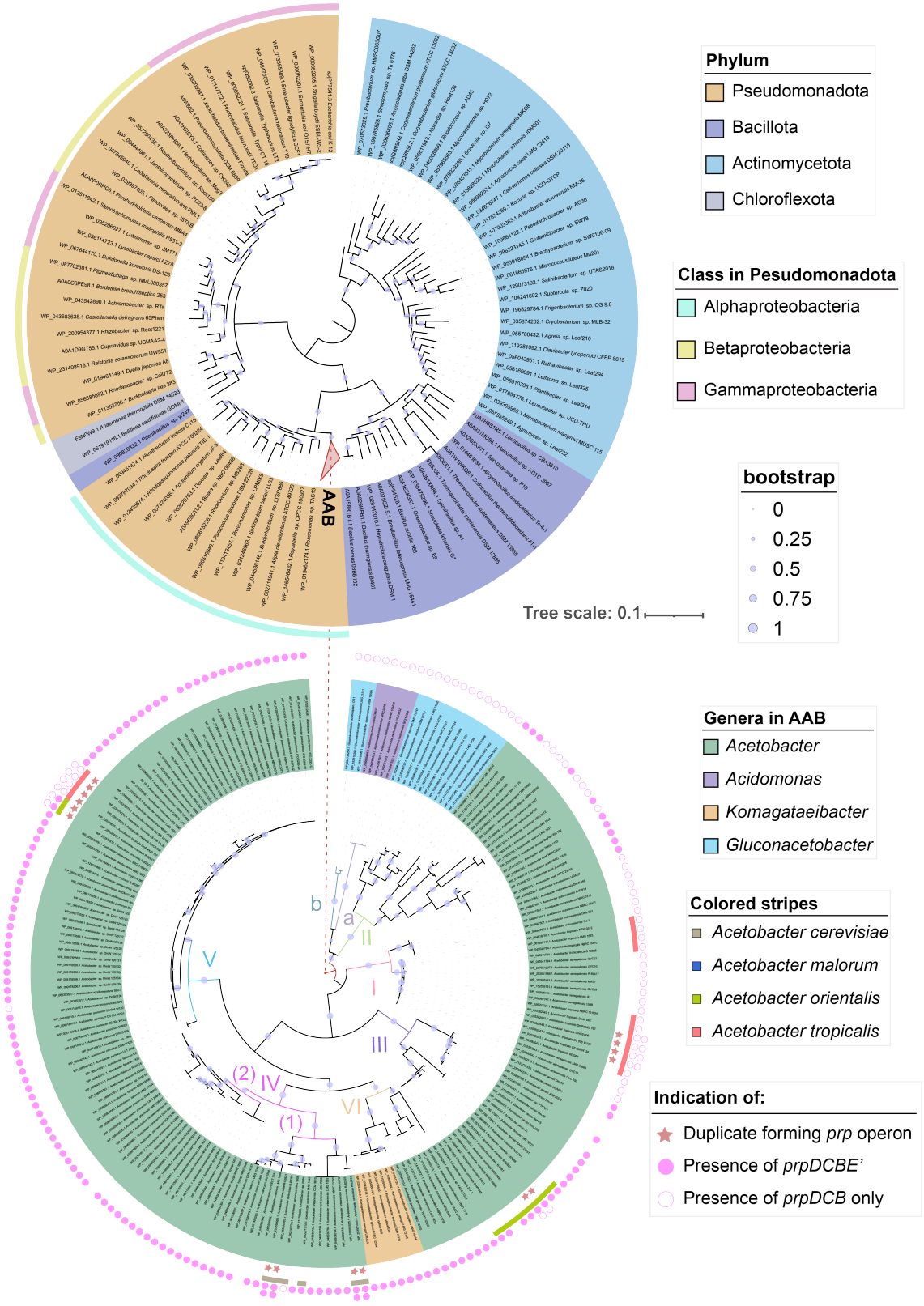

Fig. S3-C Phylogenic tree of PrpB from AAB and other bacteria

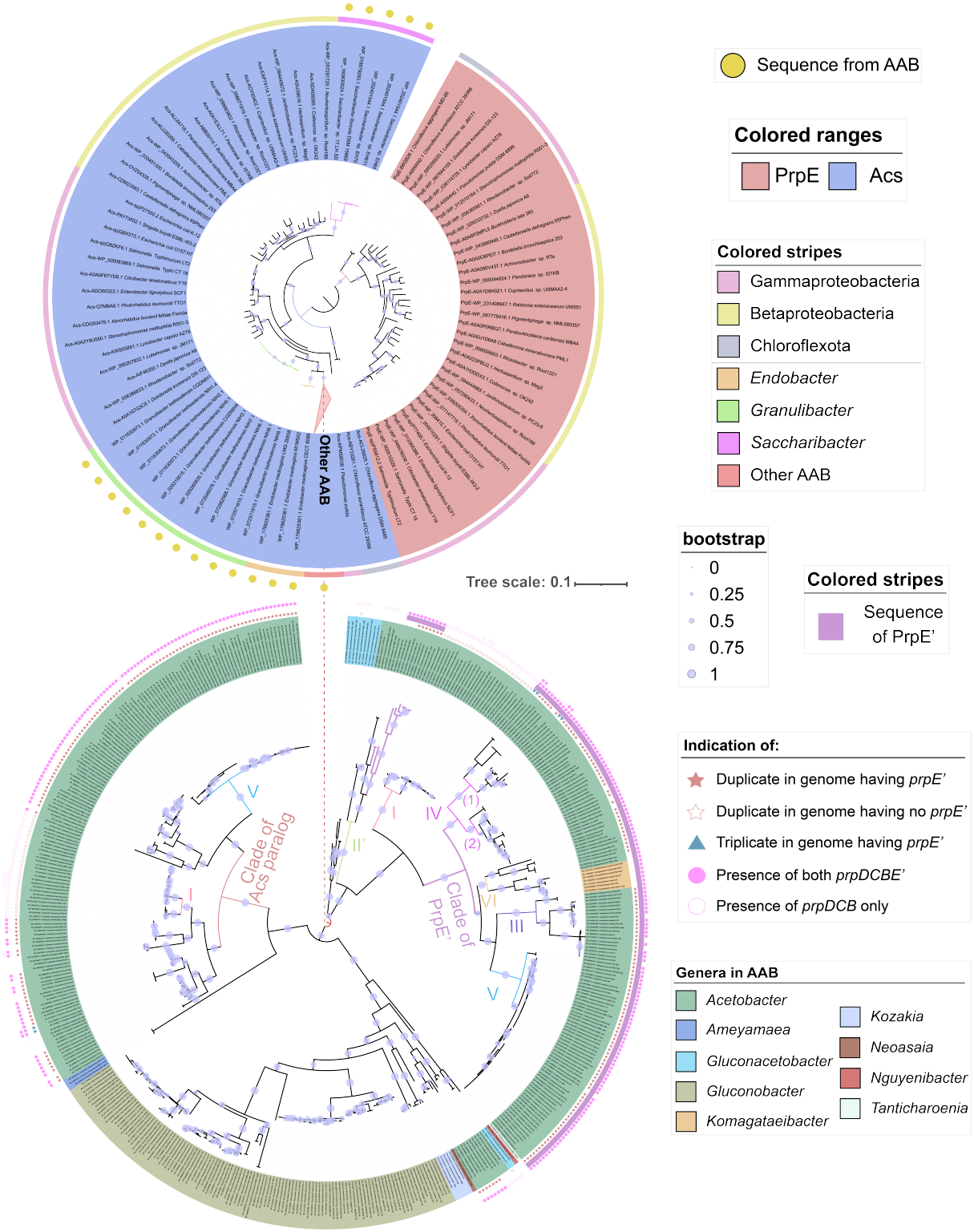

Fig. S3-D Phylogenic tree of PrpE and Acs from AAB and other bacteria

As *prp* operons in bacteria are predominantly found in members of the phyla Pseudomonadota, including the classes α-, β-, and γ-Proteobacteria, as well as in the phyla Actinomycetota and Bacillota (Fig. S2), and given that AAB also belong to the class α-Proteobacteria, the present study mainly focuses on exploring the phylogenetic relationships of *prp* operon–related genes in these bacterial groups. To this end, representative strains from the

aforementioned taxa that simultaneously harbor the three genes *prpB*, *prpC*, and *prpD* were selected, and the corresponding amino acid sequences were retrieved for phylogenetic tree construction. For strains belonging to the classes β- and γ-Proteobacteria, the presence of *prpE* and *acs* was additionally required. Moreover, two thermophilic anaerobic strains from the phylum Chloroflexota, *Anaerolinea thermophila* DSM 14523 and *Bellilinea caldifistulae* GOMI-1, were included as outgroups to support the phylogenetic analyses. However, as neither of these strains possesses *prpE*, the phylogenetic trees of PrpE and Acs instead employed sequences from *Chloroflexus aurantiacus* ATCC 29366 and *Chloroflexus aggregans* MD-66 that also belong to the phylum Chloroflexota as outgroups. In addition, given that AAB lack *prpF* and *acnD*, and that the *prp* operon structures in AAB are consistently organized as *prpDCB* or *prpDCBE'*, *prpF* and *acnD* were not included in the present analyses. Accordingly, the presentation and discussion of the results are organized following the gene order of the *prp* operon in AAB.

Amino acid sequences of PrpD, PrpC, PrpB, Acs and PrpE from bacteria strains chosen for the analysis were collected from Uniprot (https://www.uniprot.org/) or RefSeq (https://www.ncbi.nlm.nih.gov/refseq/) databases. Then, homologous sequences were aligned by Muscle algorithm with default parameters using MEGA 11 software. Based on the alignment results, maximum likelihood phylogenic trees were constructed by FastTree software with default parameters. Phylogenic trees were visualized and annotated by the online tool iTOLs (https://itol.embl.de/).

As shown in the above figures, almost all the sequences from AAB were gathered within a single clade in the corresponding phylogenic tress, except for that of PrpE and Acs, in which sequences from genus *Saccharibacter* were distinguished from other AAB. Meanwhile, genera *Granulibacter* as well as *Endobacter* obviously belong to the outgroup of AAB's Acs.

In a closer view for AAB's *prp* genes, it can be found that sequences of PrpB, PrpC, and PrpD from AAB can be clearly classified into three major distinct groups. *Acetobacter tropicalis*, *A. senegalensis* and *A. indonesiensis* formed a separate and distinct Clade I. Besides, sequences from *A. aceti* and *A. estunensis*, along with those from the majority of *Gluconacetobacter* strains, shared a common node and formed Clade II. All the sequences of other *Acetobacter* strains were developed from a single node, encompassing Clades III to VI, in which Clade V included all *A. pasterurianus* strains as well as 4 *A. tropicalis* and 2 *A. orientalis* strains possessing an additional *prp* operon. The sequences of *Acidomonas* and *Ga. azotocaptans* were found in the relatively small clades of a and b, respectively. Additionally, structures of Clades III to VI with respect to PrpE' were also present and shared a common node in the tree of Acs. On the other hand, many genomes within the Clade II have no *prpE'* or even *acs* gene, leaving an incomplete clade termed as Clade II'. Moreover, although none of

the *acs* gene in genomes of Clade I participates in the formation of *prp* operon, the two paralogous of their Acs still exhibited the Clade's structure.

Notably, the structures of Clades III to VI were found identical in both PrpC and PrpB tress, suggesting the coevolution of *prpC* and *prpB*. Given that all AAB's *prp* core genes were arranged in the order of *prpDCB*, it is implied that the relevant strains, including *A. pasteurianus* CGMCC 1.41, obtained *prpCB* simultaneously. On the contrary, differences exist in the same clades of PrpD tree compared to those in PrpB and PrpC counterparts, reflecting the independence of *prpD*'s evolution from that of *prpCB* in these genomes. Furthermore, *A. malorum* LMG 1699, *A. persici* Dm-46 and Dm-49 all harbor an addition *prpD* in the form of single gene (*A. persici* Dm-46 and Dm-49 are both supposed to have intact *prp* operons, but their operon-forming *prpD* genes are separated by scaffolds and therefore incomplete, which is why that the corresponding proteins were not shown in Fig. S2), while there are indeed strains possessing more than one *prp* operons. Besides, none of *prpC* or *prpB* were present as single genes. Thus, it could be presumed that new operons might be generating at the loci of the single *prpD*, which also suggested that this gene was initially acquired when forming an operon. Taken together, it is inferred that during the evolution of *prp* operon formation, genomes in Clades III to VI acquired *prpD* first, followed by the incorporation of *prpCB* to form *prpDCB*. Moreover, considering the absence of *prpE'* in many operons, it is plausible that this gene was integrated at a later stage during the evolution of the *prpDCBE'* operon.

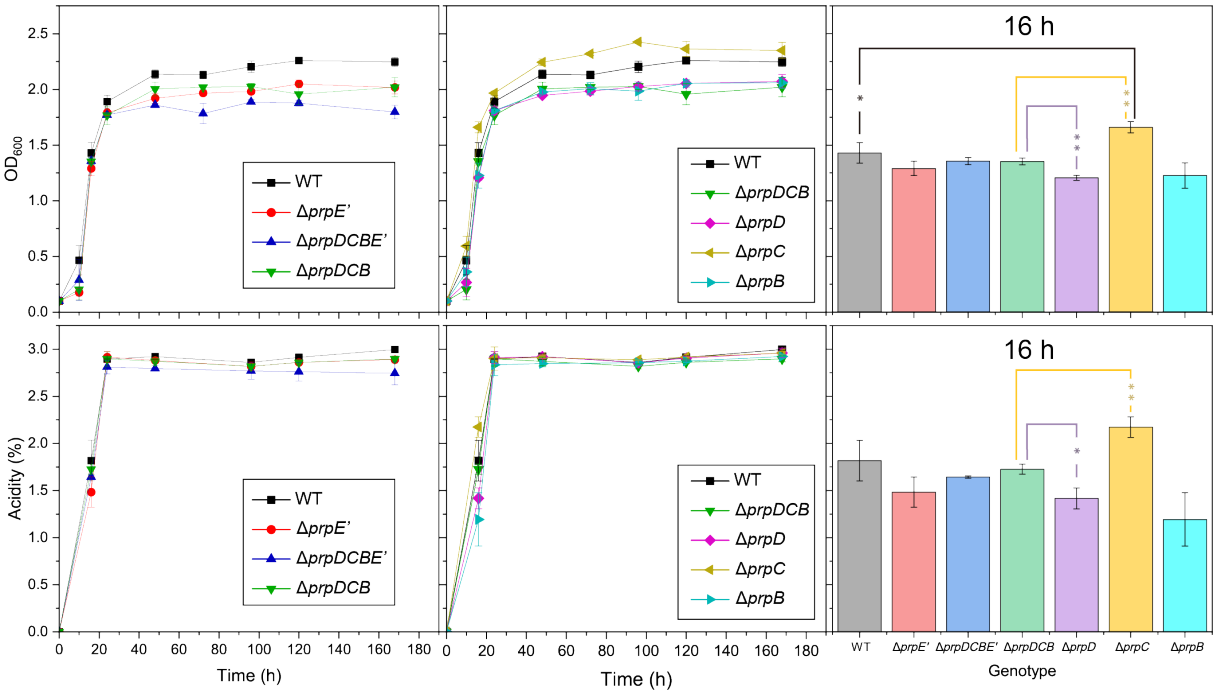

Fig. S4 Growth and acidity of all genotypes under the condition of 3% ethanol with 3 biological replicates

Data at 16 h were also present as the bar plots. Indication of significance: * ($p < 0.05$), ** ($p < 0.01$).

All genotypes finished the oxidation of 3% ethanol with acidity also reaching 3% after cultivating for 1 d (24 h). However, at 16 h of the cultivation, the content of acetic acid produced by each strain were different, indicating that their acid production rates varied. Among these, the acidity of Δ*prpC* was significantly higher than that of Δ*prpDCB*, of which the difference in the acidity compared to that of Δ*prpD* also showed significance. This indicates that the loss of Δ*prpD* and Δ*prpB* upon deleting *prpC* impaired acetic acid fermentation process, further reflecting the enhancing effect of PrpD or PrpB. Besides, as also discussed in the Discussion section, since PrpC starts the first 2-methylcitrate cycle (2-MCC) core reaction, the promoting effect from PrpD or PrpB is not related to 2-MCC. In regard to the growth, Δ*prpC* also demonstrated obvious higher OD_{600} values than WT throughout the cultivation, illustrating the inhibitory effect of PrpC.

In addition, it is worth noting that acetic acid overoxidation was not observed throughout the cultivation. As limit of ethanol concentration surely exists for its occurrence, 3% obviously exceeds this limit for *A. pasteurianus* CGMCC 1.41 under the experimental conditions of this study, which is why that acetic acid overoxidation assays were conducted with 2% ethanol.

Table S4 Mann Whitney U test results for acetic acid overoxidation data

Variable 1	Variable 2	*p*_value	Effect size	*p*_adjusted[1]	Significance[2]
		OD_{600} 4d			
WT	Δ*prpE′*	2.31E-04	0.847222	0.00162	**
Δ*prpE′*	Δ*prpDCBE′*	3.42E-02	0.444444	0.136843	ns
WT	Δ*prpDCB*	4.00E-04	0.8125	0.002398	**
Δ*prpE′*	Δ*prpDCB*	7.47E-01	-0.15278	0.747196	ns
Δ*prpC*	WT	9.20E-02	0.326389	0.275885	ns
Δ*prpC*	Δ*prpDCB*	4.42E-04	0.805556	0.002398	**
WT	Δ*prpB*	2.29E-05	0.986111	0.000183	***
Δ*prpDCB*	Δ*prpD*	1.77E-01	0.229167	0.353721	ns
		OD_{600} 5d			
WT	Δ*prpE′*	7.03E-06	0.930556	5.62E-05	****
Δ*prpE′*	Δ*prpDCBE′*	6.46E-03	0.604167	2.58E-02	*
WT	Δ*prpDCB*	9.98E-04	0.75	5.99E-03	**
Δ*prpE′*	Δ*prpDCB*	9.50E-01	-0.38889	9.50E-01	ns
Δ*prpC*	WT	7.44E-02	0.354167	1.67E-01	ns
Δ*prpC*	Δ*prpDCB*	6.72E-04	0.777778	4.70E-03	**
WT	Δ*prpB*	4.67E-03	0.631944	2.34E-02	*
Δ*prpDCB*	Δ*prpD*	5.57E-02	0.388889	1.67E-01	ns
		Acidity 5d			
WT	Δ*prpE′*	0.032161	-0.45139	0.257291	ns
Δ*prpE′*	Δ*prpDCBE′*	0.646395	0.083333	1	ns
WT	Δ*prpDCB*	0.124004	-0.28472	0.496015	ns
Δ*prpE′*	Δ*prpDCB*	0.864037	0.256944	1	ns
Δ*prpC*	WT	0.041565	-0.42361	0.290722	ns
Δ*prpC*	Δ*prpDCB*	0.041532	-0.42361	0.290722	ns
WT	Δ*prpB*	0.087333	-0.33333	0.436664	ns
Δ*prpDCB*	Δ*prpD*	0.162858	-0.24306	0.496015	ns
		OD_{600} 6d			
WT	Δ*prpE′*	2.96E-05	0.972222	0.000237	***
Δ*prpE′*	Δ*prpDCBE′*	5.25E-02	0.395833	0.157358	ns
WT	Δ*prpDCB*	1.47E-03	0.722222	0.010266	*
Δ*prpE′*	Δ*prpDCB*	9.97E-01	-0.65278	1	ns
Δ*prpC*	WT	6.03E-01	-0.05556	1	ns
Δ*prpC*	Δ*prpDCB*	6.48E-03	0.604167	0.038891	*
WT	Δ*prpB*	7.04E-03	0.597222	0.038891	*
Δ*prpDCB*	Δ*prpD*	8.90E-03	0.576389	0.038891	*
		Acidity 6d			
WT	Δ*prpE′*	0.00017	-0.86806	0.00136	**
Δ*prpE′*	Δ*prpDCBE′*	0.580329	0.041667	1	ns

Variable 1	Variable 2	*p*_value	Effect size	*p*_adjusted[1]	Significance[2]
WT	Δ*prpDCB*	0.010377	-0.5625	0.062264	ns
Δ*prpE′*	Δ*prpDCB*	0.947243	0.381944	1	ns
Δ*prpC*	WT	0.106748	-0.30556	0.320244	ns
Δ*prpC*	Δ*prpDCB*	0.041364	-0.42361	0.206821	ns
WT	Δ*prpB*	0.007048	-0.59722	0.049336	*
Δ*prpDCB*	Δ*prpD*	0.062805	-0.375	0.251221	ns
		OD_{600} 7d			
WT	Δ*prpE′*	1.81E-05	1	0.000145	***
Δ*prpE′*	Δ*prpDCBE′*	4.77E-01	0.020833	1	ns
WT	Δ*prpDCB*	2.26E-01	0.1875	1	ns
Δ*prpE′*	Δ*prpDCB*	9.90E-01	-0.55556	1	ns
Δ*prpC*	WT	9.03E-01	-0.30556	1	ns
Δ*prpC*	Δ*prpDCB*	5.35E-01	-0.01389	1	ns
WT	Δ*prpB*	7.99E-01	-0.19444	1	ns
Δ*prpDCB*	Δ*prpD*	1.82E-03	0.680556	0.012728	*
		Acidity 7d			
WT	Δ*prpE′*	1.77E-05	-1	0.000141	***
Δ*prpE′*	Δ*prpDCBE′*	7.91E-01	0.1875	1	ns
WT	Δ*prpDCB*	1.16E-03	-0.73611	0.008143	**
Δ*prpE′*	Δ*prpDCB*	9.81E-01	0.486111	1	ns
Δ*prpC*	WT	3.07E-01	-0.125	0.920215	ns
Δ*prpC*	Δ*prpDCB*	1.90E-02	-0.5	0.075886	ns
WT	Δ*prpB*	1.72E-03	-0.70833	0.010307	*
Δ*prpDCB*	Δ*prpD*	8.64E-03	-0.56944	0.043178	*

1: *p* values were Holm–Bonferroni adjusted.

2: indication of significance, ns: *p*_adjusted ≥ 0.05; *: *p*_adjusted < 0.05; **: *p*_adjusted < 0.01; ***: *p*_adjusted < 0.001; ****: *p*_adjusted < 0.0001.