TABLE S1 Sources of isolation and general genome features of *L. plantarum* strains. The previously determined genomes of WCFS1 (1, 2) and *L. plantarum* type strain DSM20246, *L. plantarum* subsp. *plantarum* DSM20174 and *L. plantarum* subsp. *argentoratensis* DSM16365 (3) are indicated by #.

Sources	**Accession No.**	**Strains**	**GC content (%)**	**Size (bp)**	**Number of Subsystems**	**Number of Coding Sequences**
Fermented Rice	LEBT00000000	*L. plantarum* K35	44.1	3,492,709	330	3345
	LEBW00000000	*L. plantarum* K36	44.5	3,211,684	332	3042
	LEBS00000000	*L. plantarum* T10	44.2	3,494,222	337	3387
Pla Pang Dang (Fermented fish)	LEBV00000000	*L. plantarum* P14	44.3	3,346,627	335	3257
	LEBR00000000	*L. plantarum* P22	44.3	3,303,556	327	3123
	LEBQ00000000	*L. plantarum* P26	44.4	3,275,958	326	3052
	LEBP00000000	*L. plantarum* P31	44.5	3,264,053	334	3095
	LEBO00000000	*L. plantarum* P42	44.5	3,343,255	334	3100
	LEBN00000000	*L. plantarum* P62	44.5	3,237,213	334	3089
	LEBM00000000	*L. plantarum* P67	44.4	3,238,093	326	3052
	LEBL000000000	*L. plantarum* P73	44.4	3,310,440	335	3146
	LEBK00000000	*L. plantarum* P76	44.3	3,343,255	328	3197
	LEBJ00000000	*L. plantarum* P86	44.4	3,237,213	323	3057
Fermented food (Pickled cabbage)	LEAW00000000	*L. plantarum* TISTR875	44.4	3,225,632	329	3055
Shrimp intestine	LEBI000000000	*L. plantarum* MHO2.4	44.5	3,207,275	329	3047
	LEBH00000000	*L. plantarum* MHO2.5	44.4	3,349,544	337	3195
	LEBG00000000	*L. plantarum* MHO2.9	44.3	3,338,588	333	3190
Marine fish	LEBF00000000	*L. plantarum* A3	44.2	3,394,509	336	3237
infant feces 8 months	LEBE00000000	*L. plantarum* I08	44.4	3,313,485	330	3170
infant feces 6 months	LEBU00000000	*L. plantarum* I61	44.4	3,314,417	330	3167
Infant feces 5 months (aerobic)	LEBD00000000	*L. plantarum* CIF17A2	45.1	3,105,497	330	2945
	LEBC00000000	*L. plantarum* CIF17A4	45.1	3,107,640	329	2932
	LEBB00000000	*L. plantarum* CIF17A5	45.1	3,106,602	329	2939
Infant feces 5 months (anaerobic)	LEBA00000000	*L. plantarum* CIF17AN2	45.1	3,128,930	330	2942
	LEAZ00000000	*L. plantarum* CIF17AN8	45.1	3,107,439	330	2942
Human	LEAV00000000	*L. plantarum* 299V	44.4	3,304,086	330	3134
Pickled cabbage	LEAY00000000	*L. plantarum* DSM2601	44.5	3,199,427	329	3020
	#	*L. plantarum* DSM20174	44.5	3,208,905	335	3060
Plant material	LEAX00000000	*L. plantarum* DSM20246	44.4	3,242,593	331	3063
Fermented cassava roots	#	*L. plantarum* DSM16365	45	3,201,989	334	2999
Human (oral cavity)	NC_004567.2	*L. plantarum* WCSF1	44.5	3,308,274	330	3013

Fig. S1 Pan and core genome of 28 *L. plantarum* strains isolated from various sources.

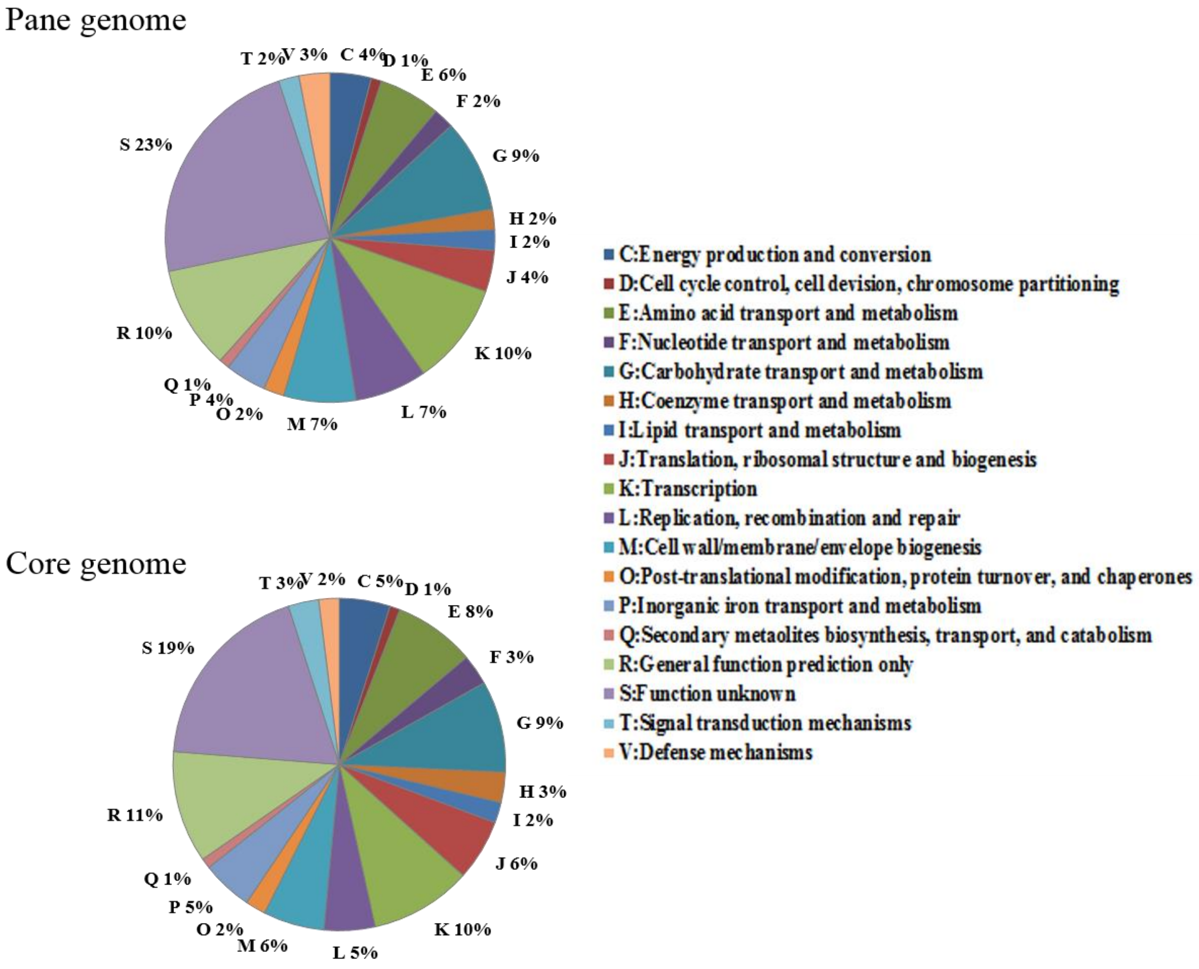


Fig. S2 Analysis of genome diversity based on carbohydrate gene presence and absence in *L. plantarum* isolated from various sources by using Mega 6 programme, dark orange colour: gene presence in the genome, white orange colour: gene absence in the genome. The PTS and other transport and metabolic genes for the following carbohydrates were identified: 1: Trehalose biosynthesis; 2: Sucrose utilization; 3: Fructooligosaccharide (FOS) and raffinose utilization; 4: Lactose utilization; 5: Moltose and moltodrextrin utilization; 6: Beta-glucoside metabolism; 7: Trehalose uptake and utilization; 8: Lactose and galactose uptake utilization; 9: Mannose metabolism; 10: D-Tagatose and galactitol utilization; 11: D-gulconate and ketogluconates metabolism; 12: Fructose utilization; 13: D-ribose utilization; 14: L-rhamnose utilization; 15: D-sorbitol (D-glucitol) and L-sorbose utilization; 16: L-arabinose utilization; 17: Glycerol and glycerol-3-phosphate uptake and utilization; 18: Mannitol utilization; 19: Inositol catabolism.

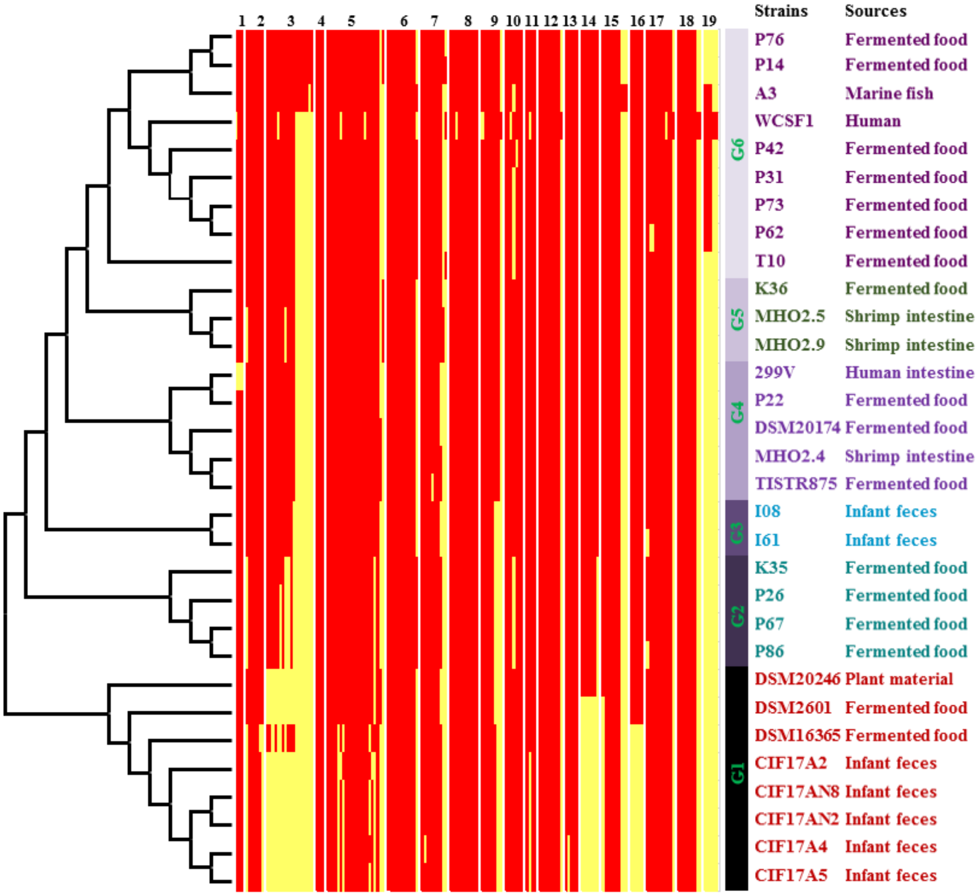


Fig. S3 Genes involved in carbohydrate metabolism in the RAST-annotated genomes of the *L. plantarum* strains.

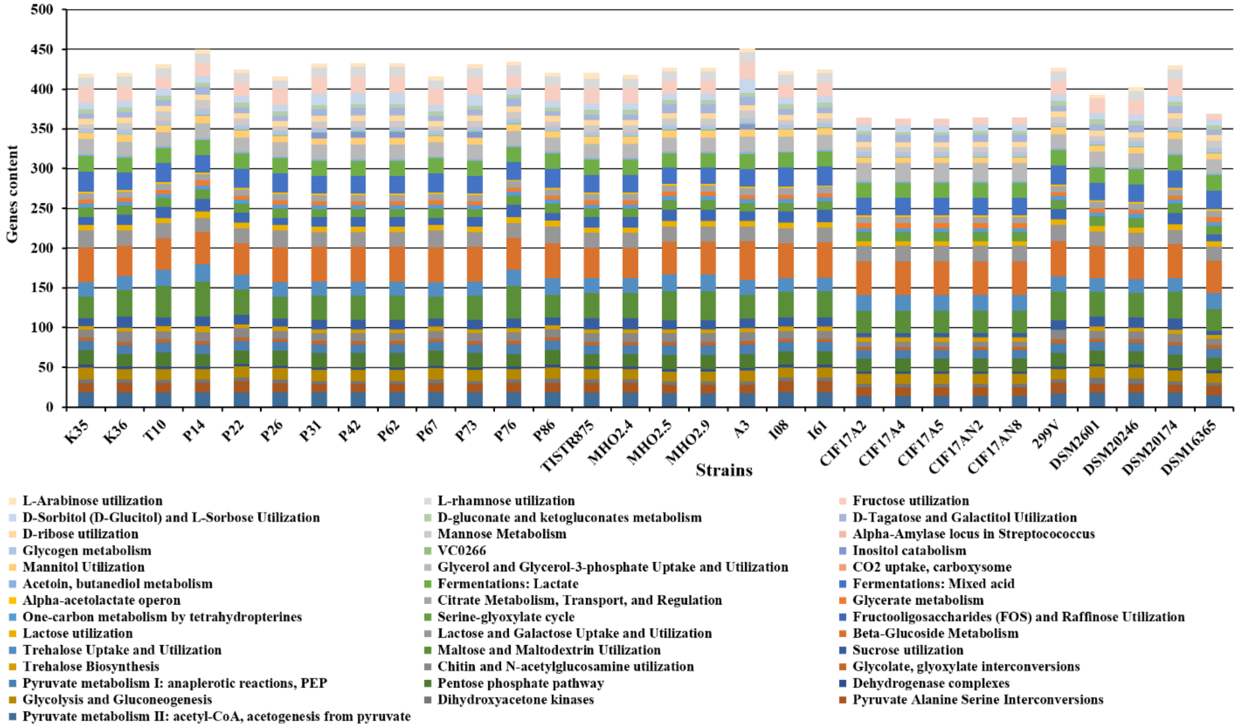
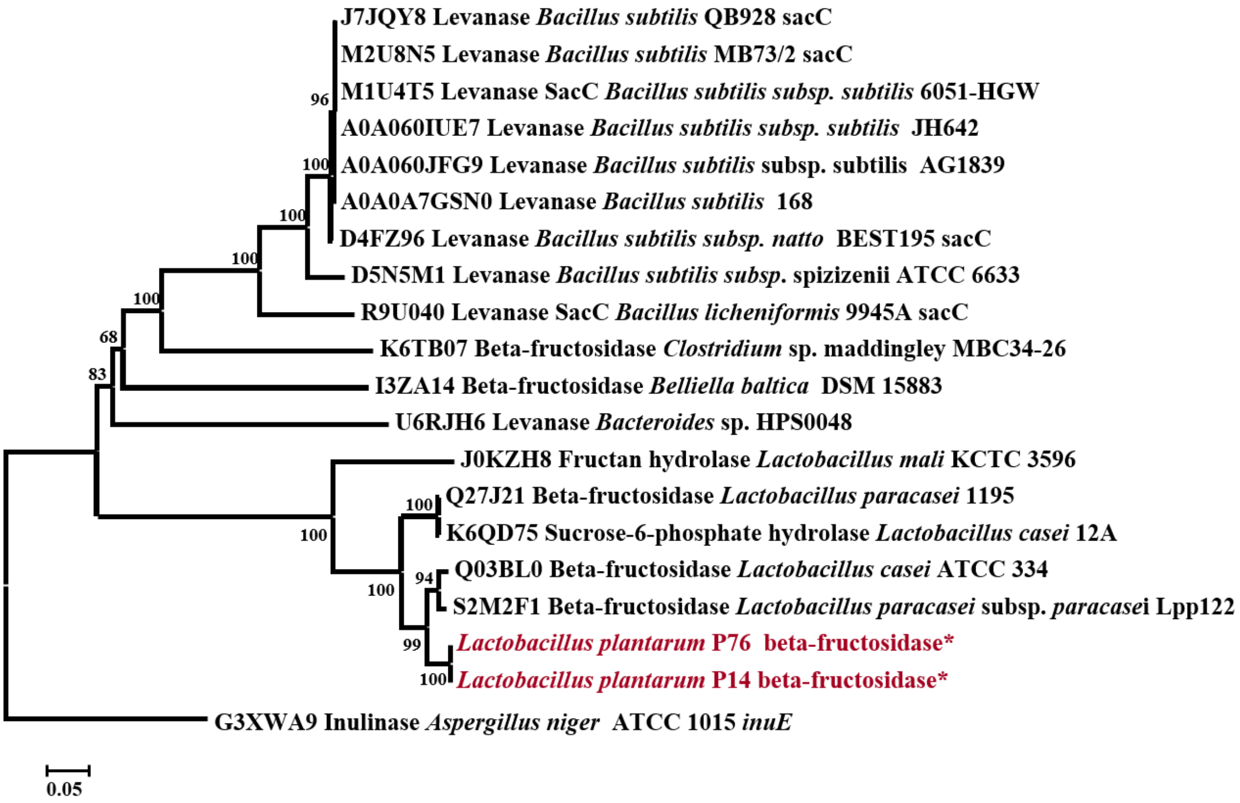


Fig. S4 Phylogenetic tree based on amino acid sequence of beta-fructosidase of *L. plantarum*. The tree was constructed by using MAGA 6 (NJ plot with 1000 bootstraps value, P-distance model).



References

1. **Sun Z, Harris HM, McCann A, Guo C, Argimón S, Zhang W, Yang X, Jeffery IB, Cooney JC, Kagawa TF, Liu W, Song Y, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O'Sullivan O, Ritari J, Douillard FP, Paul Ross R, Yang R, Briner AE, Felis GE, de Vos WM, Barrangou R, Klaenhammer TR, Caufield PW, Cui Y, Zhang H, O'Toole PW.** 2015. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. Nat Commun **6**:8322. doi:10.1038/ncomms9322

2. **Kleerebezem M, Boekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, Leer R, Tarchini R, Peters SA, Sandbrink HM, Fiers MW, Stiekema W, Lankhorst RM, Bron PA, Hoffer SM, Groot MN, Kerkhoven R, de Vries M, Ursing B, de Vos WM, Siezen RJ.** 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc Natl Acad Sci USA **100**:1990-1995. doi: 10.1073/pnas.0337704100

3. **Siezen R, Tzeneva VA, Castioni A, Wels M, Phan HTK, Rademaker JLW, Starrenburg MJC, Kleerebezem M, Molenaar D, van Hylckama Vlieg JET.** 2010. Phenotypic and genomic diversity of *Lactobacillus plantarum* strains isolated from various environmental niches. Environ Microbiol 12:758-773. doi: 10.1111/j.1462-2920.2009.02119.x.