

Supplementary Information

Comparative genome-based identification of a cell wall-anchored protein from *Lactobacillus plantarum* increases adhesion of *Lactococcus lactis* to human epithelial cells

Bo Zhang¹, Fanglei Zuo¹, Rui Yu¹, Zhu Zeng¹, Huiqin Ma², Shangwu Chen^{1*}

¹ *Key Laboratory of Functional Dairy, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, P. R. China*

² *College of Agriculture and Biotechnology, China Agricultural University, Beijing, China*

*Corresponding author. Mailing address for Shangwu Chen: College of Food Science and Nutritional Engineering, China Agricultural University, No. 17 Qinghua East Road, Hai-dian District, Beijing 100083, China. Phone: (86)1062738663. Fax: (86)1062738663. E-mail: swchen@cau.edu.cn.

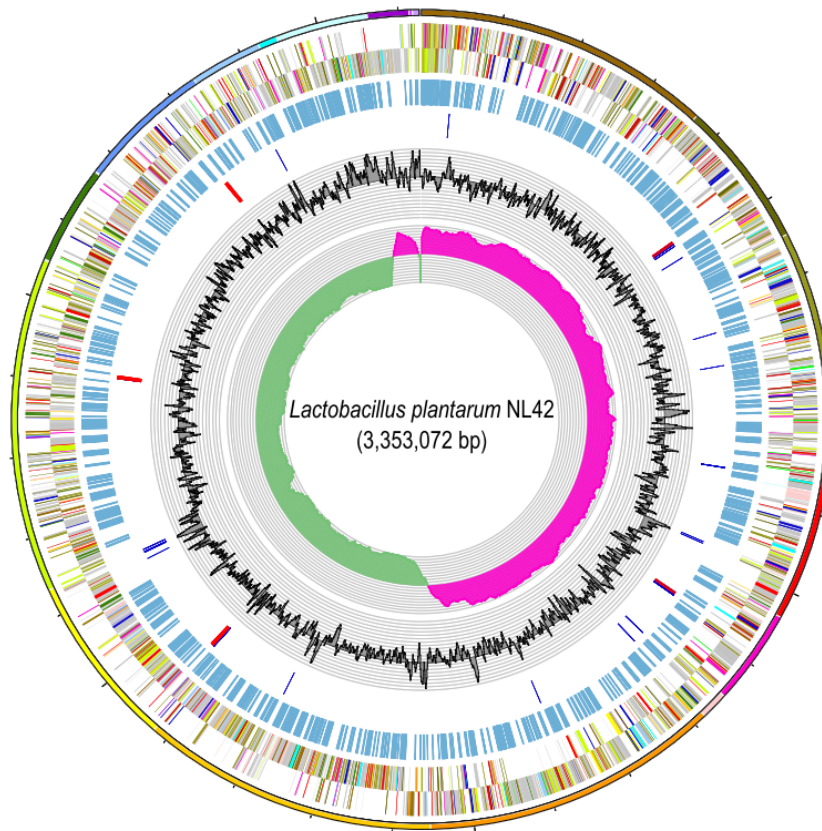


Fig. S1 Genome map of *Lt. plantarum* NL42. The circles from inside to outside indicate: GC skew (ring 1); GC content (ring 2); RNA genes (ring 3); KEGG-annotated enzymes (ring 4); COG-annotated coding sequences (rings 5 and 6) and genome sequence (ring 7, each tick indicates 100 kb). The order of the contigs was arranged according to *Lt. plantarum* WCFS1.

		Percent Identity																						
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
Divergence	1	■	99.9	99.9	99.9	99.9	100.0	100.0	100.0	99.9	99.5	100.0	99.9	99.9	99.9	99.9	99.9	99.5	99.9	100.0	100.0	99.9	1	JDM1
	2	0.1	■	100.0	100.0	100.0	99.9	99.9	99.9	100.0	99.6	99.9	99.9	99.9	99.9	100.0	100.0	99.4	99.8	99.9	99.9	100.0	2	Lp90
	3	0.1	0.0	■	100.0	100.0	99.9	99.9	99.9	100.0	99.6	99.9	99.9	99.9	99.9	100.0	100.0	99.4	99.8	99.9	99.9	100.0	3	LP91
	4	0.1	0.0	0.0	■	100.0	99.9	99.9	99.9	100.0	99.6	99.9	99.9	99.9	99.9	100.0	100.0	99.4	99.8	99.9	99.9	100.0	4	NC8
	5	0.1	0.0	0.0	0.0	■	99.9	99.9	99.9	100.0	99.6	99.9	99.9	99.9	99.9	100.0	100.0	99.4	99.8	99.9	99.9	100.0	5	NL42
	6	0.0	0.1	0.1	0.1	0.1	■	100.0	100.0	99.9	99.5	100.0	99.9	99.9	99.9	99.9	99.9	99.5	99.9	100.0	100.0	99.9	6	P8
	7	0.0	0.1	0.1	0.1	0.1	0.0	■	100.0	99.9	99.5	100.0	99.9	99.9	99.9	99.9	99.9	99.5	99.9	100.0	100.0	99.9	7	ST-III
	8	0.0	0.1	0.1	0.1	0.1	0.0	0.0	■	99.9	99.5	100.0	99.9	99.9	99.9	99.9	99.9	99.5	99.9	100.0	100.0	99.9	8	UCMA_3037
	9	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.1	■	99.6	99.9	99.9	99.9	99.9	100.0	100.0	99.4	99.8	99.9	99.9	100.0	9	WCFS1
	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	■	99.5	99.5	99.5	99.5	99.6	99.5	98.9	99.4	99.5	99.5	99.6	10	WHE_92
	11	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.1	■	99.9	99.9	99.9	99.9	99.9	99.5	99.9	100.0	100.0	99.9	11	wikim18
	12	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	■	99.9	99.8	99.9	99.9	99.5	99.8	99.9	99.9	99.9	12	ZJ316
	13	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	■	99.8	99.9	99.9	99.4	99.8	99.9	99.9	99.9	13	16
	14	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	■	99.9	99.9	99.4	99.7	99.9	99.9	99.9	14	19L3
	15	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	■	100.0	99.4	99.8	99.9	99.9	100.0	15	AG30
	16	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.0	■	98.5	99.8	99.9	99.9	100.0	16	ATCC_14917
	17	0.2	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.3	0.3	0.2	0.1	0.3	0.3	0.3	0.6	■	99.4	99.5	99.5	99.4	17	AY01
	18	0.1	0.2	0.2	0.2	0.2	0.1	0.1	0.1	0.2	0.3	0.1	0.2	0.2	0.3	0.2	0.2	0.3	■	99.9	99.9	99.8	18	EGD-AQ4
	19	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.2	0.1	■	100.0	99.9	19	EW_p
	20	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.0	■	99.9	20	PLA88
	21	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.3	0.2	0.1	0.1	■	21	JCM_1149

Fig. S2 Sequence distance of the 16S rRNA genes of *Lt. plantarum*. Strains used here are the same with those in Fig 1a. Strain names represented by the numbers 1 to 21 are listed on the right.

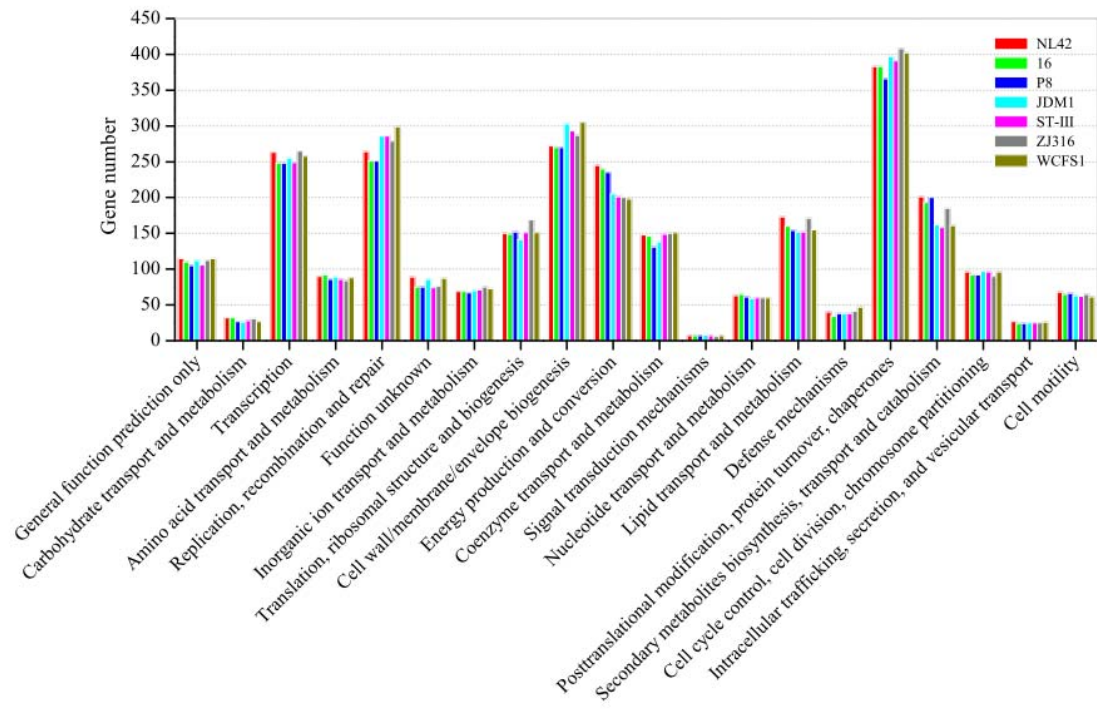


Fig. S3 Numbers of genes assigned to different COG functional categories among *Lt. plantarum* genomes.

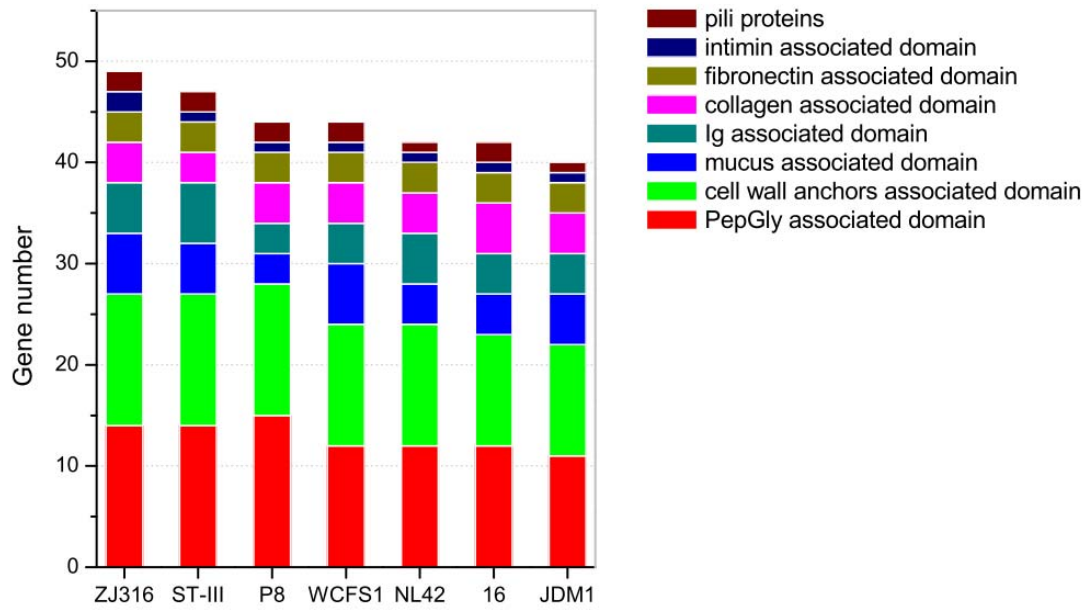


Fig. S4 Numbers of adhesion-associated and cell wall-anchored genes in different *Lt. plantarum* strains.

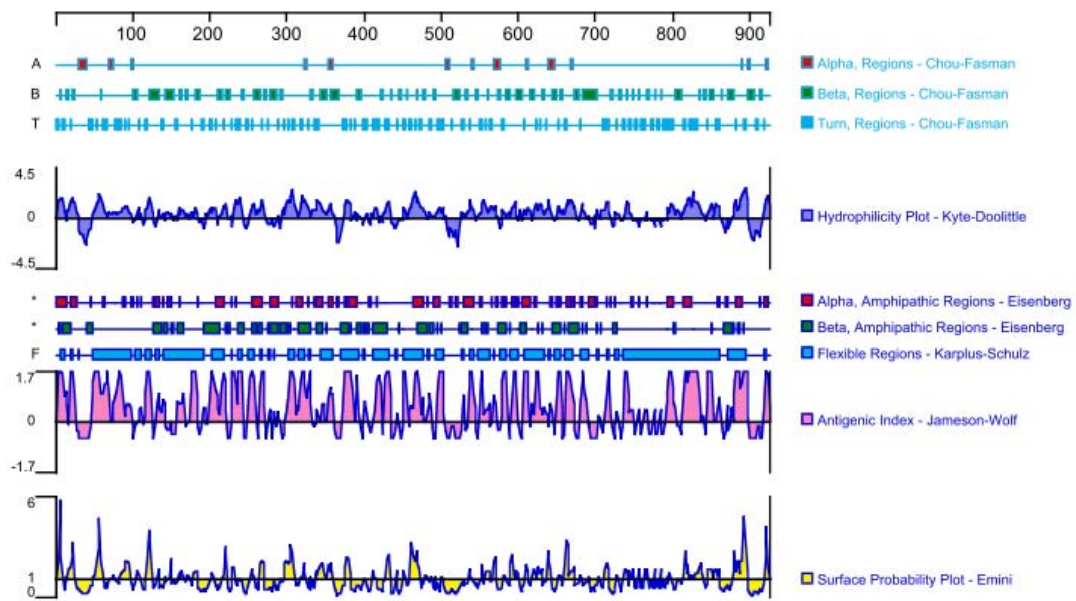


Fig. S5 Predicted secondary structure of CwaA. The analysis was performed using Protean (Lasergene package, DNASTAR, Madison, WI, USA).

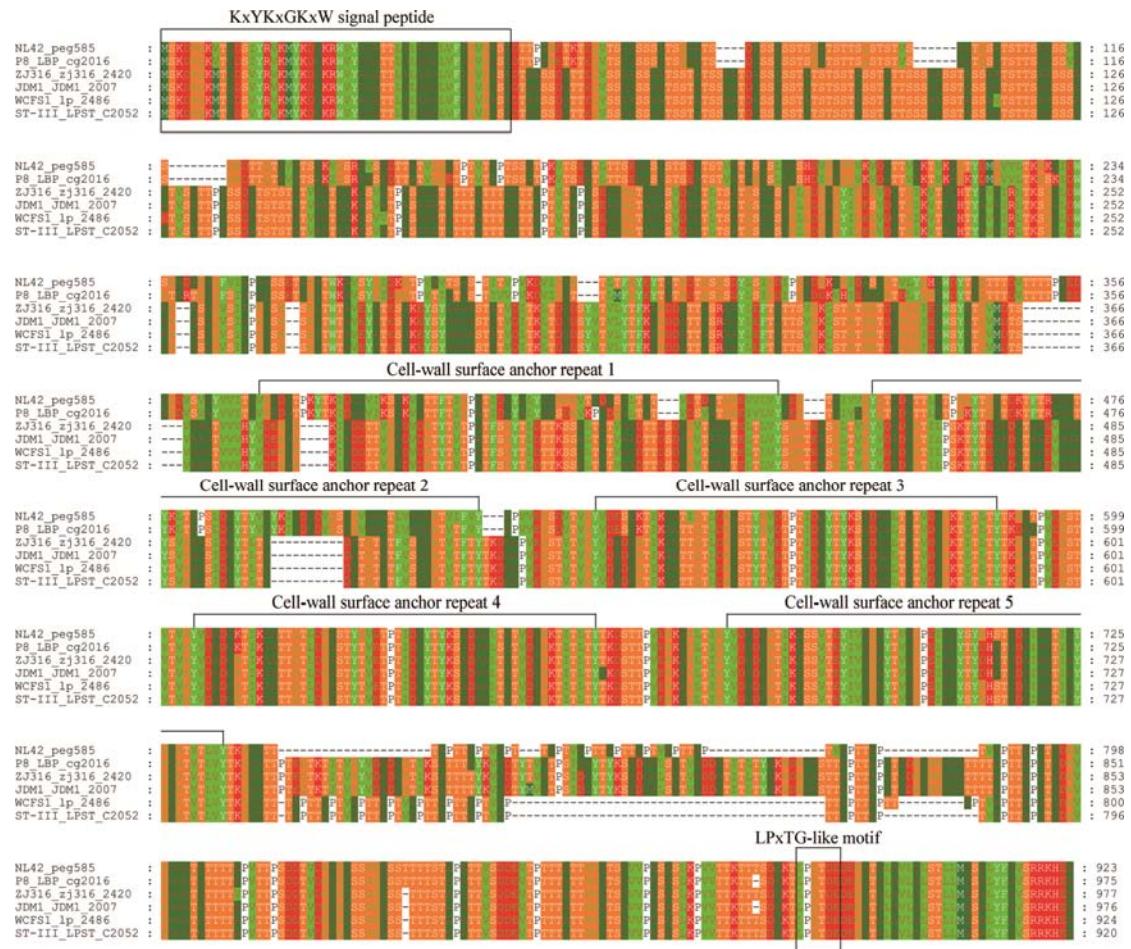


Fig. S6 Multiple sequence alignment of *Lt. plantarum* NL42 CwaA with its homologues in other *Lt. plantarum* strain complete genomes. Regions of signal peptide, cell wall surface anchor repeats and LPxTG-like motif are labeled. NL42_peg585 indicates CwaA.

1	11	21	31	41
MSKDNQKV T G	DSI Y RVKMYK	DGKR W VYAGA	T TLALAAGLV	FANVNASAD T
51	61	71	81	91
TPAN E TKTEQ	V TSGASS S AT	S QAAT S DASS	AS S T S QAT S T	T SQ S T S T V SA
101	111	121	131	141
A TQ S AT S T T S	A ASS V ASQ N D	T TQ T AVAT S A	K VQ S RA V SAD	T TAT V Q Q AT E
151	161	171	181	191
T V T APT S AT	P K V T S EAT L T	T S E AAS G S T S	E AT S T V AT G S	G SLAAS H D V N
201	211	221	231	241
A NN V K V D G T T	V A K T V K G Q T Y	D MQ V V L T K N K	A ID W S N G D A N	G Q F V I A P A G S
251	261	271	281	291
S E T ANG T W K A	V S Y Q I D K N T P	V A V T S G S S I T	V P V K D V I N A N	T L T F N Y E Y T A
301	311	321	331	341
T N D T N S N S D Y	L S I N D I P N A D	N K A N I D A S G N	T V L Y Q H G W D Y	T A Q T T T D V T T
351	361	371	381	391
T T P A E D A Q D V	S V A Y V V I T G V	N A D G T P K Y T K	I D G G V I K S G K	V G T T F T I N P A
401	411	421	431	441
T I D G Y A L Y G A	N N N V T G D A S G	L T G T V D T D A T	N N E V V L V Y A E	N T G I V V N Y Q I
451	461	471	481	491
A D G T L L A T P N	Q Y T Q G T D K T F	T R A G G T Y K L T	A P S L D G Y T V	G Y K I G D G D V Q
501	511	521	531	541
S C N V A AG T L V	A G A N T V I F V Y	A P V V E Q S D V T	V N Y V D E S G K T	I K A A T I Q T L D
551	561	571	581	591
N G S T Y N V D T P	T I D G Y T Y K S A	D A A L T G T V D G	N K T I T L T Y T K	N A T P V E Q S T V
601	611	621	631	641
T V N Y V D A D G K	T I K A A T T Q T L	D N G S T Y N V D T	P T I D G Y T Y K S	A D A A L T G T V D
651	661	671	681	691
G N K T I T L T Y T	K D S T T P V E N K	A N L T I N Y V D A	D G N T I K A S S V	T E Y I V G Q A Y T
701	711	721	731	741
V G Q P E I A G Y S	Y N H S T G D A I A	G T I G Y N G N T V	T L V Y T K N G G T	T T A P T T A P T V
751	761	771	781	791
A P T T A P T V A P	T T A P T T A P T V	A P T T A P T V A P	T T A P T V A P T T	A P G T G D N V N G
801	811	821	831	841
G G T G T T T T A P	V I T P S D D T V D	N G N G S S N N G S	S T I T T T S T A P A	T T V S D D E V T P
851	861	871	881	891
T T T A T T N N G T	S G V V P A S A S L	K E V V T T K T T T	S D A K T L E Q T D	E D E N G T A L A V
901	911	921		
L G L S T L L M G S	A L Y F G V S R R K	H E A		

Legend:

The conservation scale:



Variable | Average | Conserved

X - Insufficient data - the calculation for this site was performed on less than 10% of the sequences.

Fig. S7 Analysis of the evolutionary conservation of amino acids in CwaA using the web server ConSurf (<http://consurf.tau.ac.il/>).

Table S1 Bacterial strains and plasmids used in this study

Strains/plasmids	Relevant characteristics	Source or reference
<i>Lt. plantarum</i> NL42	Isolate from cheese	Zuo et al. 2014
<i>Lc. lactis</i> NZ9000	MG1363 <i>pepN::nisRK</i>	Kuipers et al. 1998
<i>Lc. lactis</i> NZ9000-401	<i>Lc. lactis</i> NZ9000 harboring empty vector pNZ401	This study
<i>Lc. lactis</i> NZ9000-401- <i>cwaA</i>	<i>Lc. lactis</i> NZ9000 harboring pNZ401- <i>cwaA</i>	This study
<i>E. coli</i> DH5 α	Cloning host	Takara
<i>E. coli</i> DH5 α -401	<i>E. coli</i> DH5 α harboring empty vector pNZ401	Laboratory stock
<i>E. coli</i> DH5 α -401- <i>cwaA</i>	<i>E. coli</i> DH5 α harboring pNZ401- <i>cwaA</i>	This study
<i>E. coli</i> H10407 (ETEC) ATCC 35401	Enterotoxigenic <i>E. coli</i> strain	ATCC ¹
<i>S. aureus</i> ATCC 25923	Quality control strain	CGMCC
pNZ401	Em ^r , inducible expression vector containing the <i>nisA</i> promoter, derivative of pNZ8048	Laboratory stock
pNZ401- <i>cwaA</i>	pNZ401 carrying <i>cwaA</i> gene	This study

¹<http://www.atcc.org/products/all/35401.aspx>