Amino acid alignment of the pilus gene cluster
L.rhamnosus GG DNA-seq libraries and *L.rhamnosus* GG annotation

L. rhamnosus GG DNA-seq libraries

Initial strain

~ 50 generation

Product

CHCC
CHCC_Isolate
PF1
PF1_Isolate
PF2
PF2_Isolate
FD ON
FD

Sortase C
(359 aa)

Spa A
(334 aa)

Spa B
(241 aa)

Spa C
(895 aa)

Figure S1:
Multiple sequence alignment of amino acids derived from genomic sequences from all samples in Figure 1 for genes in the spaCBA-srtC1 cluster. Top alignment shows a representative zoom-in (from the Sortase C gene) of the multiple alignment, colored by amino acid. Whole gene alignments are shown for each gene as small version of the zoom-in. All samples are identical for all genes. Sample colored as in Figure 1.

A)
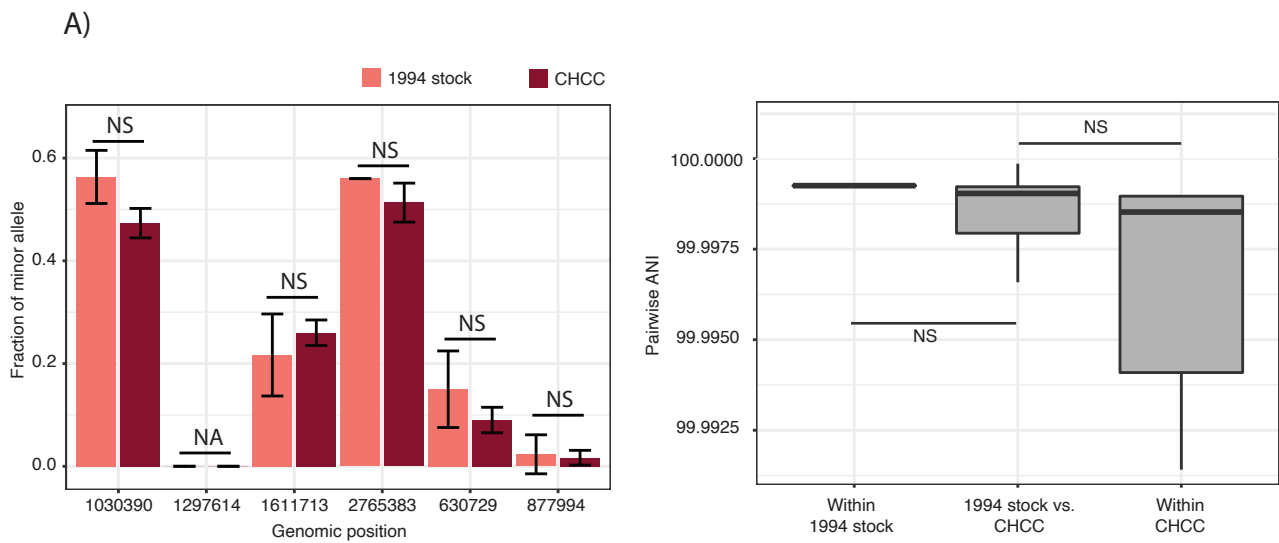
B)

Genomic Position	Amino Acid Position	Reference Codon or Nucleotide	Mutated Codon or Nucleotide	Reference Amino Acid	Mutated Amino Acid	Refseq Gene ID	Gene Product	Amino Acid Substitution Consequence
1030390	308	CAT	CAG	Histidine	Glutamine	WP_003564468.1	CamS family sex pheromone protein	Neutral
1297614	27	ACG	ATG	Threonine	Methionine	WP_005715754.1	5'-methylthioadenosine/ S-adenosylhomocysteine nucleosidase	Neutral
1611713	1827	CAG	CGG	Glutamine	Arginine	WP_014569703.1	Peptidase	Neutral
2765383	249	CAC	TAC	Histidine	Tyrosine	WP_014571681.1	6-phospho- alpha-glucosidase	Neutral
630729	227	CAG	CTG	Glutamine	Methionine	WP_014571154.1	Hypothetical protein	Neutral
877994	347	TAG	TAA	Stop codons	Stop codons	WP_014569342.1	Hypothetical protein	Unchanged Stop Codon

Figure S2:

A) Fraction of read support in the CHCC stock compared to the original stock from 1994, for the SNPs identified in Figure 4A.

Left panel: Y axis shows the fraction of reads supporting the alternative allele. Bars represent mean +/- 95% confidence intervals for the three samples of the original stock from 1994 and the three samples of CHCC from the same batch. NS (not significant) indicates P>0.05 (two-sided Mann-Whitney test) while NA (not applicable) means P was not possible to calculate (zero fraction in all samples in both groups).

Right panel: Analysis of genome-wide similarity of the three samples of the original stock from 1994 and the three samples of CHCC from the same batch. Y axis shows the distribution of pair-wise average nucleotide identity (ANI) scores for a given set of samples as boxplots (within the original stock from 1994, within the CHCC samples from the same batch, and between the two groups). NS (not significant) indicates P>0.05 (two-sided Mann-Whitney test).

B) Amino acid substitution consequence prediction of the SNPs identified in Figure 4A.

Prediction was performed by PROVEAN. Each row shows data for one genomic position.