

## **Supplemental material**

### **Enantioselective regulation of lactate racemization by LarR**

#### **in *Lactobacillus plantarum***

Benoît Desguin, Phillipe Goffin, Nordine Bakouche, Aurélie Diman, Eric Viaene, Damien Dandoy, Laetitia Fontaine, Bernard Hallet, and Pascal Hols

**Fig. S1.** Phylogeny of LarR homologs.

**Fig. S2.** Sequence of the *L. plantarum larR-larA* intergenic region.

**Fig. S3.** rLarR purification.

**Fig. S4.** Comparison between the 20 *larA-larR* intergenic regions.

**Fig. S5.** Mutagenesis of the Lar-box.

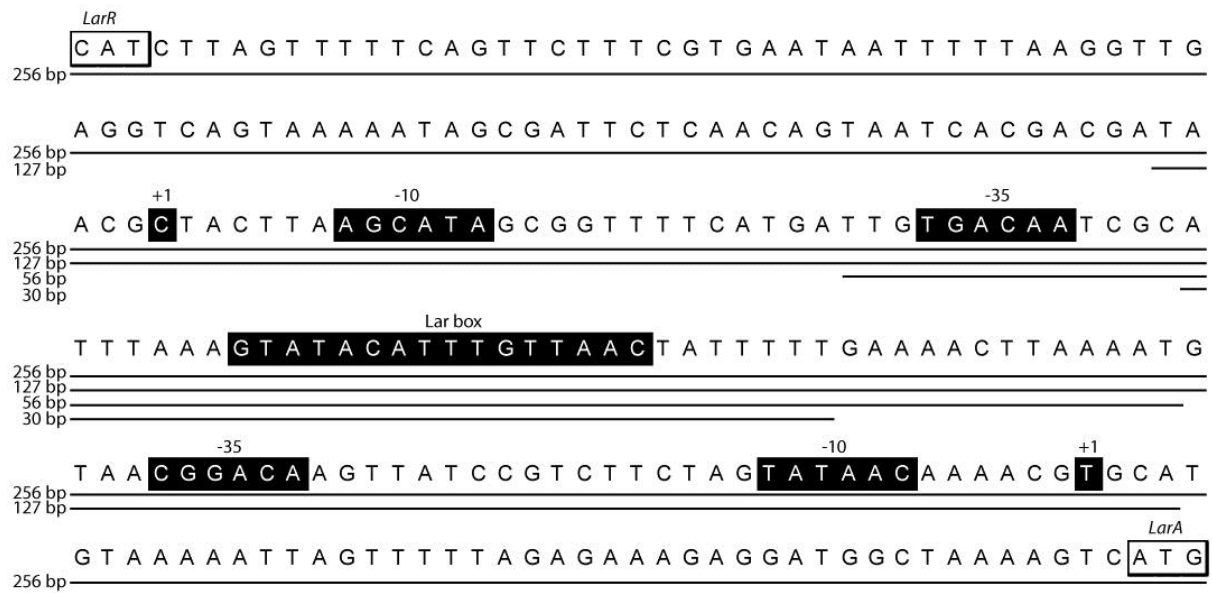
**Table S1.** Primers used in this study.

**Table S2.** MALDI-TOF analysis of the two forms of rLarR.

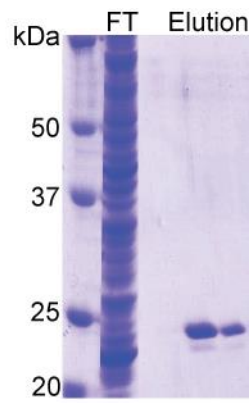
**Table S3.** Bioinformatic analysis of LarR homologues.



**Fig. S1. Phylogeny of LarR homologs.** The 55 Lar homologs used to construct the circular phylogram are reported in Table S3. The proteins identified as known members of the PrfA group (PrfA, PrfA-like, RcfB, Srv, Ers and Fnr<sub>bac</sub>) or suggested as LarR regulators based on their genetic context (LarR) are indicated between parentheses. Numbers in between parentheses correspond to the several LarR paralogs found in the same species as reported in Table S3. Crp is used as an outgroup to root the tree. The evolutionary history was inferred by using the Maximum Likelihood method based on the Le\_Gascuel\_2008 model (parameters: model LG+G) (1). The tree with the highest log likelihood is shown. The trees are drawn to scale, with branch lengths measured in the number of substitutions per site (scale bar, 1 substitution/site). All positions containing gaps and missing data were eliminated. Evolutionary analyses and multiple alignments using the ClustalW algorithm were conducted with MEGA6 (2).



**Fig. S2. Sequence of the *L. plantarum* *larR-larA* intergenic region.** The promoter -35 and -10 boxes, the Lar box and transcriptional starts (+1) are highlighted in black. The *larR* and *larA* start codons are boxed. The positions of the three [<sup>32</sup>P] radiolabelled probes used in this study are drawn as lines below the sequence. The +1 transcription start of the P<sub>larA</sub> has been determined by primer extension. The +1 transcription start of P<sub>larR</sub> has been deduced from the positions of the predicted -10 and -35 boxes.

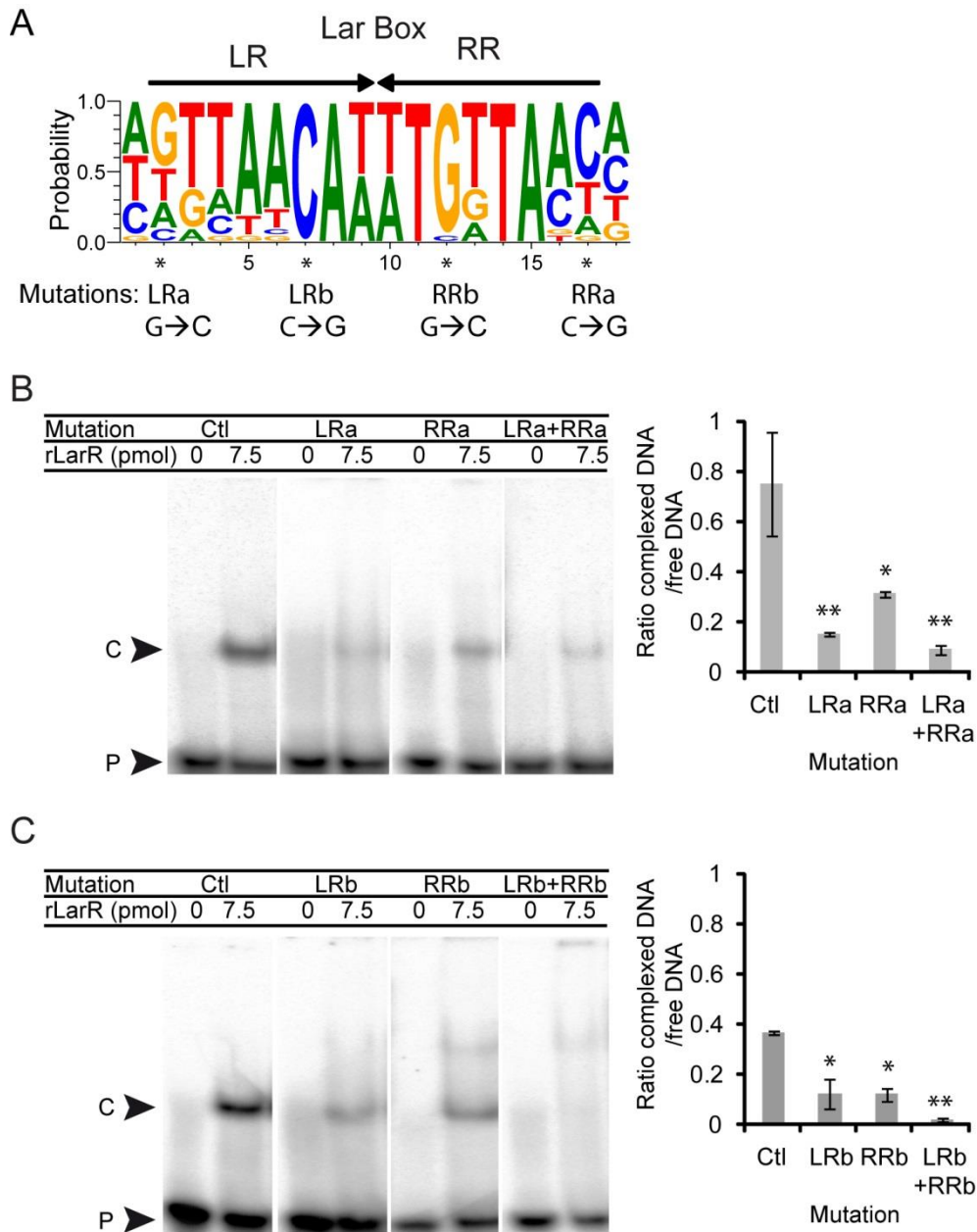


**Fig. S3. rLarR purification.** SDS-PAGE gel of the flow through (FT) and the elution fractions from rLarR purification by affinity chromatography on a Strep-Tactin column.

	bp from the center of the Lar box						
species	<i>larR</i> side	-50.5	-40.5	-30.5	-20.5	-10.5	-0.5 0.5
<i>Cl. genomosp.</i>		CATGGGAAATCTT	TCTCTCAGTCATA	GTTTTCGCGCACATGCGTT	ATTAACAAATGTTAGC		
<i>En. faecalis</i>		TGACTAAGTTTCTGTTTGATT	CGTTAACTATTATAAACACTCTCACCAACCTTTT	GTTAGCAAATGTTACC			
<i>Eu. halii</i>		CATTCTTTTTCACCCCTTTCTTCATAAAAT	AGTATATCATAT	TTTAACAGATGTT	GGTAACATTTGATAAC		
<i>G. adjacens</i>		CTCCTCTAAGACAATTTCTTATACACACT	ATTATCAT	CTTACCACGAAAATATAC	GTTAACAAATGTTACT		
<i>G. elegans</i>		AAAAAATTCACCTTGCAATTT	CTTAACATT	ATTATCCT	TTGATAAA	AATATTTTT	GTTAACATATGTTAAC
<i>L. coryniformis</i>		TACAATTACTTAATGTATCATAGT	CTTTTAGTAAA	AATTAGCGACAGCATT	TAAAGTATACATTTGTTAAC		
<i>L. fermentum</i>		CATACCTATTTGA	GTTTAGCACACTTCCAGTCAT	TGTGAAGTTAATTTTTCTTAT	GTAACAAATCATACG		
<i>L. brevis</i>		AAAATTAATTATC	GTTAGCAAAGTCCTAT	CCTTAGCGAAAAGGATACCCCTTCGC	CTTAACATTTGGTAAT		
<i>L. mali</i>		ATTATGTAAAACCTTATTAGTT	GATTTAAATAATA	CTCCTAATTTTATCAT	ATCATCTTATCAATTGGTAAT		
<i>L. pentosus</i>		ACAATGACGCTTCTTAAGCATAGCGGTTTTCGTGAT	TGTGACAA	TCGCATTTAAAGTATACATTTGTTAAC			
<i>L. plantarum</i>		ACGATAACGCTACTTAAGCATAGCG	GTTTTCATGAT	TGTGACAA	TCGCATTTAAAGTATACATTTGTTAAC		
<i>St. vestibularis</i>		ATCTTAATAACACTAATAATA	CACCTATACTACCAATTCCTATATCATCCT	TTGTTAAC	ATCAAATGATACT		
<i>L. sakei</i>		TAATCATTTATACTTATTATT	CGATAACTTAAGTCTTAAGTGTA	CGCTATTTTAATTGAACATTTGTTAAC			
<i>Sp. inulinus</i>		TGTTATCTATATTCT	ATTTTAACGCCATTT	CAGCTAATATATATCACAAATAAAGGAAACATTTGTTAAC			
<i>Eu. biforme</i>		CATATTCGCTCTCCTTTACCA	GTTTTGGAT	TATAACAA	TACTCCAAATTTTTCGAAACAACAAATGTTAAC		
<i>Pe. clausenii</i>		ATATAATCCCCCTGCACTTAT	GTGTTTTATTATCTTTAAATTCCGATACGCGCACT	GTAACAAATGTTACA			
<i>Pe. pentosaceus</i>		TAAAAGTAAATAAT	ATGTACAGACATCT	ATTATCAT	AAGAAAAAATTGTGAGAAGT	GTAACAAATGTTATA	
<i>L. curvatus</i>		TACTAATCATTAAATAACTTTA	ATCTAAATTATA	CGTCTTGAATACTATAAAGTT	GTTATCATTGTTACA		
<i>Co. intestinalis</i>		CGCCCCTTTCAGATTTGTTAAG	CCGACCACTGTATACACCTAGCGAAACGTTGAGTTGACATTTGGTAAC				
<i>Co. tanakaei</i>		CCTCCCCTACCTGACCTCTGGAA	AGCGCAACTCTATCGGACTGTGAAGCCGAAGC	TGCACCAATGGTAAC			
<b>consensus</b>						T	GTTAACATTTGGTAAC AA

species	bp from the center of the Lar box						<i>larA</i> side
	-0.5	0.5	10.5	20.5	30.5	40.5	
<i>Cl. genomosp.</i>	ATTAACAAATGTTAGC	ATTTTTT	TGAGCTAACCGGCAATTATTTTTTC	CGT	GCTAAATTT	TAGGTC	CAAGTTGCGA
<i>En. faecalis</i>	GTTAGCAAATGTTACC	CACAATAA	GGACTGTGTGATTCCTAAATTT	CATT	TATACTTAGTCT	CGT	AAAATAT
<i>Eu. halii</i>	GGTAACATTTGATAAC	ATTTTTT	TATTATTATCT	TTT	ACTAACACTTTTT	ATTAATATGAAAT	GTTGTTTTC
<i>G. adjacens</i>	GTTAACAAATGTTACT	AACCTTT	TGAGACTTCTGCTGTGCGTTATTT	GAGTATACT	TATAAT	CGACAA	ACAA
<i>G. elegans</i>	GTTAACATATGTTAAC	ACCATTT	TGATTTACATT	CAAAA	GAAAAAATAA	TATACTATATT	GTTAGCTAT
<i>L. coryniformis</i>	GTATACATTTGTTAAC	CTACTTT	GAAAATGCTAAA	CGTGGC	AGACA	AGCCAAAA	AAGATGCGTATAACTG
<i>L. fermentum</i>	TGTAACAAATCATACG	CTTATTTT	TGACGCTGAAA	GCAACGCA	ACAATGAGAGTATTTT	TGAAA	ACTGAGA
<i>L. brevis</i>	CTTAACATTTGGTAAT	GACACTGGCT	ATTGATAT	CGATTATTC	ACTTTCATAGT	GAGGAT	TGTCAATGAA
<i>L. mali</i>	CTTATCAATTGGTAAT	GTTTAA	AGAA	CATACACTTAAATA	ATGTTTT	AAACAATACTTAT	TGACAAGAAA
<i>L. pentosus</i>	GTATACATTTGTTAAC	TATTTT	TGAAA	ACTTGAAATGTAGC	GGACA	ACCGTCGCGACTACTAGTATAACAA	
<b><i>L. plantarum</i></b>	GTATACATTTGTTAAC	TATTTT	TGAAA	CTTAAATGTAAC	GGACA	AGTTATCCGTCTTCTAGTATAACAA	
<i>St. vestibularis</i>	AACATCAAATGATACT	TATGTAT	TTTAAA	AAAAGACTTATAA	ACTGGT	GTAAGTATTAAT	TAATAAAGGAG
<i>L. sakei</i>	TTGAACATTTGTTAAC	ATTTT	TGAGC	TTTCTTTTGTAAC	CGTTTTAATGTTGTCTATAA	ATTATACCGCAT	CC
<i>Sp. inulinus</i>	GGAAACATTTGTTAAC	ATTTTT	TGATGATAGC	AAACGGAA	CCGACA	ACCTTTTTGTTTTGATT	TATAACAG
<i>Eu. biforme</i>	AACAACAAATGTTAAC	ATAAAA	ATGAG	GTTTCATAATTATTTA	ATTCGGATTATATT	AATTCGGT	AAAAAA
<i>Pe. clausenii</i>	TGTAACAAATGTTACA	CCTAATT	GTTGTATCCA	CGACCACTTATCTTTCCTAACAT	GTTGTAT	AGCGCCAA	
<i>Pe. pentosaceus</i>	TGTAACAAATGTTATA	CCCAAAT	GTGATAC	CGGTCCG	ATTTTACATAGCCTAATATATT	GTCTAGC	GCCAA
<i>L. curvatus</i>	GTTATCATTTGTTACA	CCTAAA	AAGTAATA	ATGAAA	TGTTGTT	CATAAAT	TATACCGCATCCTTTTCAGAAA
<i>Co. intestinalis</i>	GTTGACATTTGGTAAC	TACATTC	GACATAACGCGCAC	CCTTTCAA	ACCCCATAGTT	ACCCCG	GCTTAAAC
<i>Co. tanakaei</i>	TGCACCAAATGGTAAC	AAACAA	AGGT	CAGACCGCCCGATGCCATCACAAC	CATAGT	TACCCCT	CGTCCGTA
<b>consensus</b>	<b>GTTAACA</b>	<b>TTTGGTAA</b>	<b>TTTG</b>				
	<b>AA</b>						

**Fig. S4. Comparison between the 20 *larA-larR* intergenic regions.** The suspected +1 transcription starts, the promoter -10 boxes, -35 boxes and the Lar boxes are highlighted in red, yellow, orange, and green, respectively. The identified half-Lar boxes are boxed in black. The consensus sequence is shown below the sequences.



**Fig. S5. Mutagenesis of the Lar-box.** (A) Display of mutations LLa, LRb, RRa, and RRb performed in the left and right inverted repeats (LR and RR) of the Lar-box (inverted arrows). The mutations are indicated underneath the logo. (B and C) Left panels: EMSA of 56-bp [<sup>32</sup>P] radiolabeled probes containing the native Lar-box or mutated Lar-boxes in absence or presence of purified rLarR. Ctl, control with the native Lar-box. The rLarR-DNA complex (C) and unbound probe (P) are indicated by arrow heads. Right panels: mean values of ratios of complexed DNA/free DNA of EMSA as in left panels. Data are from 3 replicates from one significant experiment. The error bar represents the standard deviation. Significance based on Student *t*-test; \* *p*-value < 0.05, \*\* *p*-value < 0.01. Mutated Lar-boxes are compared to the native Lar-box (Ctl).

**Table S1. Primers used in this study.**

Primer name	Sequence <sup>a,b</sup>	Mutant strain or plasmid
<b>Primers used for construction of the deletion vector</b>		
LP101A1	5'-CGGGGTACCAATCTTCAAACGGCGTGGTTAGACG-3'	
LP103A4	5'-GCTCTAGAGATGTTAGTTACCTCTTAGATCAAATTGC-3'	
LP105B4	5'-GGACTAGTCTTACCAGCACCGTAGATGACCTCC-3'	
LP103B4	5'-GCTCTAGAAAGTGTTCGGCTAAACCATGATAGG-3'	pGIR002
LOX66A5-2	5'-GCTCTAGATACCGTTCGTATAATGTATGCTATACGAAGTTATTGCAGTTTAAATTTCG GTCCTCG-3'	
LOX71A5-2	5'-CTCTAGATACCGTTCGTATAGCATAACATTATACGAAGTTATAAGTACAGTCGGCATT ATTC-3'	
<b>Primers used for the validation of the deletion</b>		
LP096UP1	5'-CGGGGTACCTAGCTGAGATCAATGCGAACTTGG-3'	LR0002
LP105B1	5'-GCTCTAGAAAAGCGTCTAAATAAAGTGTTC-3'	
<b>Primers used for the construction of overexpression and purification vectors</b>		
LP103OXA1	5'-CATCCCATGGCAAGCTGGAGTCACCCGCAGTTCGAGAAAGGTGCTGG TATGGTATTAAGTATTTGAATAT-3'	pGIR991
LP103OXB1	5'-GCGGAATTCAAAACGGCGACGGTCCAGACGGGAG-3'	
Lp1031	5'-GGGGTACCTAAAACGGCGACGGTCCAGACG-3'	pGIR090
Lp1032	5'-CATGCCATGGTATTAAGTATTTGAATATTTACT-3'	
<b>Primers used for the construction of controlled gene expression vectors</b>		
PSIP103-104A1	5'-CGGGATCCCATCTTAGTTTTTCAGTTCTTTTCG-3'	pGIR003
PSIP103-104B2	5'-CATGCCATGGCTTTTAGCCATCCTTCTCTA-3'	
<b>Primers used for the sequencing of vectors</b>		
P32B1	5'-CGCGGATCCGAATGCATTCTGCTGAAACGATTGCCATTTTC-3'	PGIR002
CATA1	5'-AAAACTGCAGATGAAGAAAGCAGACAAGTAAGC-3'	
UP_PNZ8048	5'-TAGATACAATGATTTTCGTTTCG-3'	pGIR090/091
CAT1	5'-AATTGTCAGATAGGCCTAATGACTGG-3'	
GUSA2	5'-TTCCATGATTTCTTTAACTATGC-3'	pGIR003
GUSB2	5'-TTTTGGCTATCAATCAAAGCAACACG-3'	
<b>Primers used for the amplification of P<sub>32</sub> radiolabeled and unlabeled probes</b>		
LarBox_A1	5'-CATCTTAGTTTTTCAGTTCTTTTCG-3'	Lar box 256 bp
LarBox_B1	5'-CATGACTTTTAGCCATCCTCTTTC-3'	
LarBox_A2	5'-TAACGCTACTTAAGCATAGCGGTTTTTC-3'	Lar box 127 bp
LarBox_B2	5'-TGCACGTTTTGTTATACTAGAAGAC-3'	
Lar-A_A	5'-CGATCGATATGAAAAAGGCCTACGCACGCGAGGGTC-3'	Unspecific DNA 137 bp
Lar-B_B	5'-GCATCGATTTGCCGCCACTTGTGTAATATTTTC-3'	
<b>Primers used for primer extension</b>		
LP104PE1	5'-TTGTTATGATATGTTGCAGCTTGC-3'	
<b>Primers used directly as P<sub>32</sub> radiolabeled probes</b>		
LarBox_lp1	5'-ATTTAAAGTATACATTTGTTAACTATTTTT-3'	<i>L. plantarum</i> Lar box
LarBox_lp1	5'-TAAATTTCAATGTAACAATTGATAAAAA-3'	
LarBox_cons1	5'-ATTTAAAGTAACTTTGTTAACTATTTTT-3'	Consensus Lar box
LarBox_cons1	5'-TAAATTTCAATGTAACAATTGATAAAAA-3'	
LarBox_rand1	5'-ATTTAAAGTTCTGATTTATAAACTATTTTT-3'	Random Lar box
LarBox_rand2	5'-TAAATTTCAAGACTAAATATTTGATAAAAA-3'	
LarBox3_A	5'-TTGTGACAATCGCATTTAAAGTATACATTTGTTAACTATTTTTGAAAACCTAAAAAT-3'	Wild type
LarBox3_B	5'-ATTTAAAGTTTTCAAAAATAGTTAAACAATGTATACTTTAAATGCGATTGTCACAA-3'	
LarBox3'3_A	5'-TTGTGACAATCGCATTTAAAGTATACATTTGTTAACTATTTTTGAAAACCTAAAAAT-3'	Mutation LRA
LarBox3'3_B	5'-ATTTAAAGTTTTCAAAAATAGTTAAACAATGTATACTTTAAATGCGATTGTCACAA-3'	
LarBox3'4_A	5'-TTGTGACAATCGCATTTAAAGTATACATTTGTTAAAGTATTTTTGAAAACCTAAAAAT-3'	Mutation RRA
LarBox3'4_B	5'-ATTTAAAGTTTTCAAAAATACTTAAACAATGTATACTTTAAATGCGATTGTCACAA-3'	
LarBox3''_A	5'-TTGTGACAATCGCATTTAAAGTATACATTTGTTAAAGTATTTTTGAAAACCTAAAAAT-3'	Mutation LRA+RRA
LarBox3''_B	5'-ATTTAAAGTTTTCAAAAATACTTAAACAATGTATACTTTAAATGCGATTGTCACAA-3'	
LarBox3'5_A	5'-TTGTGACAATCGCATTTAAAGTATAGATTTGTTAACTATTTTTGAAAACCTAAAAAT-3'	Mutation LRb



LarBox3'5_B	5'-ATTTTAAGTTTTCAAAAATAGTTAACAAT <u>CTATACTTTAAATGCGATTGTCACAA</u> -3'	
LarBox3'6_A	5'-TTGTGACAATCGCATTAAAGTATACATTT <u>CTTA</u> ACTATTTTTGAAAACCTAAAAT-3'	Mutation RRb
LarBox3'6_B	5'-ATTTTAAGTTTTCAAAAATAGTTAAGAAATGTATACTTTAAATGCGATTGTCACAA-3'	
LarBox3'7_A	5'-TTGTGACAATCGCATTAAAGTATAGATTT <u>CTTA</u> ACTATTTTTGAAAACCTAAAAT-3'	Mutation
LarBox3'7_B	5'-ATTTTAAGTTTTCAAAAATAGTTAAGAAAT <u>CTATACTTTAAATGCGATTGTCACAA</u> -3'	LRb+RRb
<b>Primers used for QuickChange mutagenesis</b>		
QC_LarBox_A3	5'-ATCGCATTAAAGTATACATTT <u>CTTA</u> ACTATTTTTGAAAAC-3'	pGIR003B (mutation RRb)
QC_LarBox_B3	5'-TAGTTAAGAAATGTATACTTTAAATGCGATTGTCACAATCATG-3'	
QC_LarBox_A2	5'- <b>CATTTGTTAACT</b> ATTTTTCAAAAACCTAAAATGTAACGGAC-3'	pGIR003C (mutation RRc)
QC_LarBox_B2	5'-TTTAAGTTTTGAAAATAGTTAACAATGTATACTTTAAATGCG-3'	
QC_LarBox_A4	5'- <b>AACT</b> ATTTTTCAAAAACCTAAAATCTAACGGACAAGTTATCC-3'	pGIR003D (mutation RRc+RRd)
QC_LarBox_B4	5'-GTCCGTTAGATTTTAAGTTTTGAAAATAGTTAACAATG-3'	

<sup>a</sup> Restriction sites and mutations introduced in the primers are underlined

<sup>b</sup> Lar box sequences are in bold

**Table S2. MALDI-TOF analysis of the two forms of rLarR.**

<b>1<sup>st</sup> band</b>		<b>Match to: LarR Score: 470 Expect: 1.6e-045</b>					
<b>Start - End</b>	<b>Observed</b>	<b>Mr(expt)</b>	<b>Mr(calc)</b>	<b>Delta</b>	<b>Miss</b>	<b>Sequence</b>	
34 - 51	2209.0051	2207.9978	2200.1003	-0.1025	0	R.HTYLTYHGLAEHYTYVLK.D	(Ion score 98)
57 - 65	1015.5026	1014.4953	1014.5458	-0.0505	0	K.NSILQDGR.E	(Ion score 61)
66 - 82	1994.0045	1992.9972	1993.0884	-0.0911	0	R.EYNLSYIAKPDVISLLR.D	(No match)
88 - 96	1063.4653	1062.458	1062.5094	-0.0514	0	R.STDQPFNVR.I	(Ion score 44)
97 - 109	1619.6981	1618.6908	1618.7627	-0.0719	0	R.IESEYATFYQVNR.V	(Ion score 96)
115 - 127	1554.7046	1553.6973	1553.7725	-0.0752	0	K.YVNSTPELQNYVK.N	(Ion score 81)
133 - 140	972.5323	971.525	971.5763	-0.0513	1	K.KLSENILR.L	(No match)
134 - 140	844.4393	843.432	843.4814	-0.0494	0	K.LSENILR.L	(Ion score 30)
167 - 192	2853.303	2852.2957	2852.4377	-0.142	1	R.KVNEGILIDFVVTNDDIAGFCGISSR.S	Propionamide (C)
168 - 192	2725.2021	2724.1948	2724.3428	-0.148	0	K.VNEGILIDFVVTNDDIAGFCGISSR.S	Propionamide (C)

<b>2<sup>nd</sup> band</b>		<b>Match to: LarR Score: 280 Expect: 1.5e-026</b>					
<b>Start - End</b>	<b>Observed</b>	<b>Mr(expt)</b>	<b>Mr(calc)</b>	<b>Delta</b>	<b>Miss</b>	<b>Sequence</b>	
34 - 51	2209.1116	2200.1043	2200.1003	0.004	0	R.HTYLTYHGLAEHYTYVLK.D	(No match)
57 - 65	1015.5601	1014.5528	2208.1003	0.007	0	K.NSILQDGR.E	(Ion score 53)
66 - 82	1994.0621	1993.0548	1014.5458	-0.0335	0	R.EYNLSYIAKPDVISLLR.D	(No match)
88 - 96	1063.5229	1062.5156	1993.0884	0.0062	0	R.STDQPFNVR.I	(Ion score 46)
97 - 109	1619.7723	1618.765	1062.5094	0.0023	0	R.IESEYATFYQVNR.V	(Ion score 89)
115 - 127	1554.7827	1553.7754	1618.7627	0.0029	0	K.YVNSTPELQNYVK.N	(Ion score 50)
134 - 140	844.4944	843.4871	843.4814	0.0057	0	K.LSENILR.L	(No match)
144 - 150	807.4059	806.3986	806.4142	-0.0156	1	R.MVMNGKK.G	(No match)

The MASCOT search engine was used to identify the proteins after Maldi-TOF analysis.

**Table S3. Bioinformatic analysis of LarR homologues.**

Species	N <sup>oa</sup>	Protein <sup>b</sup>	Accession	Annotation	LarA <sup>c</sup>
<i>Bacillus</i> sp. 1NLA3E		Fnr <sub>Bac</sub>	ZP_09599669.1	cyclic nucleotide-binding protein	
<i>Clostridiales</i> genomosp. BVAB3 str. UPII9-5		LarR	YP_003475460.1	Crp/Fnr family transcriptional regulator	YP_003475461.1
<i>Collinsella intestinalis</i> DSM 13280		LarR	ZP_04446027.1	hypothetical protein COLINT_02751	ZP_04446028.1
<i>Collinsella tanakaei</i> YIT 12063					
	1	LarR	ZP_08853122.1	hypothetical protein HMPREF9452_00991	ZP_08853123.1
	2		ZP_08853117.1	hypothetical protein HMPREF9452_00986	
<i>Enterococcus casseliflavus</i> ATCC 12755			ZP_08145504.1	crp/Fnr family transcriptional regulator	
<i>Enterococcus faecalis</i> T3					
	1	LarR	ZP_05503397.1	transcription regulator	ZP_05503398.1
	2	Ers	ZP_05502499.1	transcriptional regulator	
<i>Enterococcus faecium</i> E0045			ZP_19440287.1	CarD family transcriptional regulator	
<i>Enterococcus gallinarum</i> EG2			ZP_05648526.1	transcriptional regulator	
<i>Enterococcus hirae</i> ATCC 9790			YP_006487412.1	Crp/Fnr family transcriptional regulator	
<i>Enterococcus italicus</i> DSM 15952			ZP_07896383.1	crp/Fnr family transcriptional regulator	
<i>Enterococcus</i> sp. 7L76			CBL31034.1	cAMP-binding proteins-catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	
<i>Eubacterium bifforme</i> DSM 3989					
	1	LarR	ZP_03487930.1	hypothetical protein EUBIFOR_00495	ZP_03487929.1
	2		ZP_03489375.1	hypothetical protein EUBIFOR_01964	
	3		ZP_03487649.1	hypothetical protein EUBIFOR_00208	
	4		ZP_03487783.1	hypothetical protein EUBIFOR_00348	
<i>Eubacterium hallii</i> DSM 3353		LarR	ZP_03715095.1	hypothetical protein EUBHAL_00131	ZP_03715096.1
<i>Granulicatella adiacens</i> ATCC 49175		LarR	ZP_05738512.1	Crp/Fnr family transcriptional regulator	ZP_05738513.1
<i>Granulicatella elegans</i> ATCC 700633		LarR	ZP_05852068.1	transcriptional regulator, Crp/Fnr family	ZP_05852067.1
<i>Lactobacillus brevis</i> ATCC 367		LarR	YP_795942.1	cAMP-binding protein	YP_795943.1
<i>Lactobacillus coryniformis</i> subsp. <i>coryniformis</i> KCTC 3167		LarR	ZP_08478609.1	transcription regulator	ZP_08478608.1
<i>Lactobacillus curvatus</i> CRL 705		LarR	ZP_09041129.1	bacterial regulatory s, crp family protein	ZP_09041121.1
<i>Lactobacillus fermentum</i> IFO 3956		LarR	YP_001844550.1	transcriptional regulator	YP_001844549.1
<i>Lactobacillus mali</i> KCTC 3596 = DSM 20444		LarR	ZP_09448267.1	cAMP-binding protein	ZP_09448266.1
<i>Lactobacillus pentosus</i> MP-10		LarR	CCB81962.1	transcription regulator	CCB81961.1
<i>Lactobacillus plantarum</i> WCFS1		LarR	YP_004888173.1	Crp/FNR family transcriptional regulator	YP_004888174.1
<i>Lactobacillus sakei</i> subsp. <i>sakei</i> 23K		LarR	YP_395397.1	unnamed protein product	YP_395398.1
<i>Lactococcus garvieae</i> ATCC 49156					
	1		YP_004779753.1	transcription regulator	

	2	YP_004778376.1	transcription regulator	
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363				
	RcfB	YP_001033749.1	rcfB gene product	
<i>Lactococcus lactis</i> subsp. <i>lactis</i> KF147				
		YP_003354892.1	Crp family transcriptional regulator	
<i>Listeria grayi</i> DSM 20601				
	1	ZP_07052624.1	crp/Fnr family transcriptional regulator	
	2	ZP_07052673.1	global nitrogen regulator NtcA	
<i>Listeria innocua</i> Clip11262				
		NP_470555.1	hypothetical protein lin1218	
<i>Listeria ivanovii</i> subsp. <i>ivanovii</i> PAM 55				
		YP_004854974.1	putative regulator of the Fnr CRP family	
<i>Listeria marthii</i> FSL S4-120				
		ZP_07870549.1	Crp/FNR family transcriptional regulator	
<i>Listeria monocytogenes</i> EGD-e				
	1 PrfA	NP_463731.1	listeriolysin positive regulatory protein	
	2 PrfA like	NP_464776.1	hypothetical protein lmo1251	
<i>Listeria seeligeri</i> FSL N1-067				
		ZP_16495800.1	Crp/FNR family transcriptional regulator	
<i>Listeria welshimeri</i> serovar 6b str. SLCC5334				
		YP_849464.1	Crp/FNR family transcriptional regulator	
<i>Listeriaceae bacterium</i> TTU M1-001				
		ZP_09893295.1	Crp/FNR family transcriptional regulator	
<i>Pediococcus clausenii</i> ATCC BAA-344				
	LarR	YP_005005673.1	unnamed protein product	YP_005005672.1
<i>Pediococcus pentosaceus</i> ATCC 25745				
	LarR	YP_805158.1	cAMP-binding protein	YP_805157.1
<i>Pseudomonas putida</i> KT2440				
	Crp	NP_742590.1	cAMP-regulatory protein	
<i>Sporolactobacillus inulinus</i> CASD				
	LarR	ZP_09714632.1	transcription regulator	ZP_09714633.1
<i>Streptococcus agalactiae</i> ATCC 13813				
		ZP_08649995.1	crp/Fnr family transcriptional regulator	
<i>Streptococcus equi</i> subsp. <i>zooepidemicus</i>				
		YP_002745168.1	Crp family regulatory protein	
<i>Streptococcus ictaluri</i> 707-05				
		ZP_09124891.1	transcriptional regulator, Crp/Fnr family	
<i>Streptococcus pyogenes</i> M1 GAS				
	Srv	NP_269852.1	transcription regulator Srv	
<i>Streptococcus urinalis</i> 2285-97				
		ZP_09136776.1	transcriptional regulator, Crp/Fnr family	
<i>Streptococcus vestibularis</i> F0396				
	LarR	ZP_07723811.1	transcriptional regulator, Crp/Fnr family	ZP_07724144.1
<i>Weissella koreensis</i> KACC 15510				
	1	YP_004725607.1	unnamed protein product	
	2	YP_004725591.1	unnamed protein product	
	3	YP_004726285.1	unnamed protein product	

<sup>a</sup> The LarR homologs are numbered when several paralogs are found in the same species. These numbers are reported in Fig .S1

<sup>b</sup> The identification of LarR proteins is based on the synteny with a *larA* gene

<sup>c</sup> When present, the NCBI accession numbers of *L. plantarum* LarA homologs are indicated

## References

1. **Le SQ, Gascuel O.** 2008. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**:1307-1320.
2. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**:2725-2729.