

L1.LtrB-WT

<u>Gene name</u>	<u>5' flanking</u>		<u>additional nts</u>	<u>3' flanking</u>	
	EBS2 <u>GUGUA</u>	EBS1 <u>GUGUUG</u>		EBS2 <u>GUGUA</u>	EBS1 <u>GUGUUG</u>
5'nucleotidase	(2)	ACA CACUUC CACAAA	/AAAGAUCAGAUUAG-253-CUU CACAUU UACUGG /	UGAGGAACUUCGGCA	
a-acetolactate decarboxylase	(1)	CAA UACAUU UAGCAC	/ACUCUCUAGCGGCUU-26--AAG CACUAA CACACG /	GCAAAGUCGGUAUCG	
<u>Ribosomal protein L13/S9</u>	(1)	ACC AACAUAC CACACC	/ACAUGCUGAUACAGG-431-GCAG CACAU CGUCUU /	GUUAUCAACCAACCA	
<u>Ribosomal protein L13/S9</u>	(1)	ACA CACAUA UACAGC	/UCAAAAACCCUGAAGU-355-UAC ACGUGA CGCCCG /	UAUGGUUGAACGUAA	
<u>Ribosomal protein L13/S9</u>	(2)	ACC AACAUAC CACACC	/ACAUGCUGAUACAGG-303-ACA UAUUGG CACAAG /	UACAAUUGCCGGCA	
L-lactate dehydrogenase	(1)	UGC AGCUUA CUCAAU	/CAUCGCUAAAAAAGG-170-UUC ACAUU CAUUGA /	ACG AUGCUGAAAUGC	
NADH dehydrogenase	(1)	AAC AGCACAC CGCAGC	/AGCGGCAAAUUAUCU-563-CUCA UAAAG CACAUC /	UAUGGGGUCUCUCCA	
ABC transporter permease	(1)	AUG CACUUU CCCAAU	/UUUCAUCUGUCCGUA-150-UUU CACCUU CUUCCA /	ACAGCCUUAGUACAU	
Glutamine synthetase	(1)	CAU CACAAA CCCAAC	/AGUUAAACUCAUCAA-386-CAAG CUAUG CACAAU /	UUGUAUCGCAAUGGG	
Glyceraldehyde dehydrogenase	(1)	AAU GUCAUA CACUUC	/AAACCUUGUUC-----GUA CACUUG CAUACU /	UCGCUAAAAUCGCUA	
Elongation factor Tu	(1)	ACU CCGAAAC CGCGAC	/ACUGACAAACCACUC-362-CUAC UUC CACAAC /	UGAGGUUACUGGUUC	
Ribosomal protein L19	(1)	ACU GAUAU CCUGAC	/UUUCGUCCUGGUGAC-153-GUU CACACU CCACGU /	GUUGAAAAAAUCGAA	
<i>Ribosomal protein L11</i>	(4)	AUU CACAUU CAUCAC	/AAAAACUCCUCCAGC-233-CCU CACCUA CUUAGG /	UAGGCUUUUGGACCUCU	
N-acetylglucosaminidase	(1)	GAC AACAAAC CACAAG	/UAAUACCGUUCUACAG-173-GCA CAUCAG CAUCAU /	CUGGUAGUUACACAA	
Aldehyde dehydrogenase	(1)	CCA CACGUC CACAAA	/GCUUUCAUUGUUGCC-254-UUC CGAUGA CGCAAG /	UUUGAAAGAGAUUCUU	
Phosphotransferase	(1)	UAU UGUGUA CACAAC	/UUUCAUGAGAUUCUUG-94--UAU CACAUU CCUUUG /	UAUGUGGUUGCUUAA	
Ferredoxin oxidoreductase	(1)	GAC CCUUUA UACAAC	/AAAUGAACAGGAACA-107-AAG CACUUU CAGAAA /	UGUCGGAUGAUUUAC	
ABC transporter	(1)	UGA ACAUAU UUGCGAU	/UCUAGGCCGUUCAGG-136-UAU CUCUUU UGCAACA /	ACGGUUCGUUCUAAU	
g-aminobutyrate permease	(1)	ACU CAUUAU CGCAAC	/UUAAUGAUGAUAAAA-56--UUC ACAUCA CUAUUG /	AUUUUUUUUACUCCA	
<u>Uncharacterized prot 1-2</u>	(1)	UAG ACCUUC CACGAU	/UUUAUGAUUACCUAG-214-AAC CACAGA CACAAU /	UAAAAUUUUGCUACA	
Thioredoxin reductase	(1)	AAU CAUUAU CGCGAC	/UGGAGCUAAUCACCG-185-AAU UACGUG CACAAG /	AAAUUUUUCAACAAA	
Arsenate reductase	(1)	CAC AAUCUA UACAGC	/ACCCUCUUGUACAAG-185-UUU CACUUU CACAAG /	CAAUUAAAAUCAUUU	
NADH dehydrogenase	(1)	AAC AGCACAC CACAGC	/AGCGGCAAAUUAUCU-173-CAG CACAUU UAUUUCU /	GGCGUUUAGCCUUGG	
Transcriptional regulator	(1)	UAA AACGCAC CACGAC	/AACAAUGGUUUCAAU-173-ACA GCUUUG CAACUG /	UAAAAUUUUGAUGGUU	
Uncharacterized protein	(1)	AAA CACAAG UCUACU	/GGAAAAUUAUCAAA-284-UGA CUCAGG CACCAA /	AAAAGAAAGCAGCAA	
Ser/Thr protein kinase	(1)	AAC AAUACA UCAACC	/AAAAUGCUCUCUACAG-51--UGG CUCAU UACAAG /	GAAUUUUUACCCGUG	
a-acetolactate decarboxylase	(1)	CAA UACAUU UACGAC	/ACUCUCUAGCGGCUU-50--GUA UCGGUA CGCUUG /	AUACGGCAAUUGGCG	
Ribosomal protein S14	(1)	AAU ACAUAU AAUGGCU	/AAGAAAUCUAUGGUU-64--GUC CACAUU CAGUUU /	ACCGCAAUUUAAAC	
N-acetylglucosaminidase	(1)	UAA CUCAAA CUCGAC	/UUCUUCUAAUCUCAA-50--AAA UAUCU GGCAGCC /	CAAUUGCUUCAUCA	

Figure S1. mRNA fragments identified at the splice junction of L1.LtrB-WT circles. Additional nts are shown along with their flanking sequences (5' flanking) (3' flanking), their origin (Gene name) and frequency of identification between parentheses. The junctions between the additional nts and their flanking regions (/) as well as the IBS1- (yellow) and IBS2- (green) like sequences are denoted. The bolded nts represent residues from the IBS1- and IBS2-like sequences that can potentially base pair with the intron's EBS1 and EBS2 sequences specified above. Sequences spanning two genes and including a short intergenic region are underlined while the mRNA sequence including 3' untranslated residues is italicized.