

	10	20	30	40	50	60	70	80	90	100	110	120	130	140
	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Xlr2	ATGGCAGCTGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCAGCGCAGCGAGACGCCGACACGGAGCACGCCCGACCGCGAAAAACGGACCCGTAGGGCTTGTGACAAGTGCAGCGCCTCGAGGA												
Xlr2-000202	ATGGCAGCTGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCAGCGCAGCGAGACGCCGACACGGAGCACGCCCGACCGCGAAAAACGGACCCGTAGGGCTTGTGACAAGTGCAGCGCCTCGAGGA												
TAS_375582	ATGGCAACCGAAGCCGACGGTGAACAGCGGGATCGCCCTCAGCCTCTGGTCGCGTCTGGCC	-ATG-ACCGCGGGCTCAACGCCGATGCGGAGCACGCCAGGCCTCGGAAGCGCACCCGAAGAGCGTGCGACAAGTGCAGCGCCTCGAGGA												
TA_229644	ATGGCTACCGAAGCCGACGGTGAAGAGCGGCATCGCCCTCAGCCTCTCGTCGCGTCTGGCC	-ATG-ACCGCGGGGGCCACGCCGATGCGGAGCACGCCCGGCCTCGGAAGCGCACCCGGAGAGCGTGCGACAAGTGCAGCGCCTCGAGGA												
TLONGI_14376	ATGGCAGCTGAAGCCGACGGTCAAGAGCGGCACCGCCCGAGCCTCTTGTCGCGCC	-GTTG-GCGCAGCGCAGCGAGACGCCGATACGGAGCACGCCCGGCCGAGGAAGCGGACCCGGAGGGCTTGCGACAAGTGCAGCGCCTCCAGGA												
TR_60282	ATGGCAGCTGAAGCCGACGGCCAAGAGCGGGATCGCCCGAGCCTCTTGTCGGCTC	-ATTG-GCGCAGCATAGCGAGACGCCGATACGGAGCACGCCCGGCCGAGGAAGCGGACCCGGAGGGCTTGCGACAAGTGCAGCGCCTCAAGGA												
TRICITRO_110	ATGGCAGCCGAAGCCGACGGTCAAGAGCGGCATCGCCGCGGTTCTCTTGTCGCGTC	-AGTG-ACGCAGCGCAGTGAGACGCCGACACGGAGCACGCCCGGCCGAGGAAGCGGACCCGGAGGGCTTGCGACAAGTGCAGCGCCTCGAGGA												
THZ_79686	ATGGCAGCCGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCACCCGACGCGAAACGCCAGACACGGAGCACGCCCGCCACGGAACGGACCCGTAGGGCTTGTGATAAAGTGCAGTGCCTCGAGGA												
TH_701679	ATGGCAGCCGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCACCCGACGCGAAACGCCAGACACGGAGCACGCCCGCCACGGAACGGACCCGTAGGGCTTGTGATAAAGTGCAGTGCCTCGAGGA												
A0A1T3C8Z2	ATGGCAGCCGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCACCCGACGCGAAACGCCAGACACGGAGCACGCCCGCCACGGAACGGACCCGTAGGGCTTGTGATAAAGTGCAGTGCCTCGAGGA												
TGAM_8287	ATGGCAACCGAAGCCGACGGTGAACAGCGGGATCGCCCTCAGCCTCTCGTCGCGTC	-TGAGCATG-ACCGCGGGTCGACACCCGATGCGGAGCACGCCCGGCCCGGAAGCGCACCCGGAGAGCGTGCGACAAGTGCAGCGCCTCGAGGA												
OPB37586	ATGGCAGCCGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCACCCGACGCGAAACGCCAGACACGGAGCACGCCCGCCACGGAACGGACCCGTAGGGCTTGTGATAAAGTGCAGTGCCTCGAGGA												
cDNA OPB3758	ATGGCAGCCGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCACCCGACGCGAAACGCCAGACACGGAGCACGCCCGCCACGGAACGGACCCGTAGGGCTTGTGATAAAGTGCAGTGCCTCGAGGA												
cDNA Xlr2*	ATGGCAGCTGAAGCCGACGGTCAAGAACGGGATCGCCCTGAGCCTCTCGTCGCGTC	-ATTG-ATGCAGCGCAGCGAGACGCCGACACGGAGCACGCCCGACCGCGAAAAACGGACCCGTAGGGCTTGTGACAAGTGCAGCGCCTCGAGGA												

	150	160	170	180	190	200	210	220	230			
	*	*	*	*	*	*	*	*	*			
Xlr2	CACGATGCGATGGAGAATG	TGAGT----	CT----	GATACTACAGCCAACCGGTGGCTC----	GTTATCGT-GATATCT-----	GAA-C-----	TG-----	ACTGA----	GATA----	T-----	C--TG--CGC	
Xlr2-000202	CACGATGCGATGGAGAATG	TGAGT----	CT----	GATACTACAGCCAGCCGGTGGCTC----	GTTATCGT-GATATCT-----	GAA-C-----	TG-----	ACTGA----	GATA----	T-----	C--TG--CGC	
TAS_375582	CTCGTTGTGATGGAGAATAGT	GAGT----	CT-----	CTCGTGCCAGGTGTTGAGTG----	GCGTTGGTTGATTGATGATTTCATGTGCAACC-----	TG-----	ACTGACT--	GATACAAAT-----	CA-TA--	CGT		
TA_229644	CTCGCTGTGATGGAGAATAGT	GAGT----	CTC----	TCTCTCTCTCATACGATCTCTC----	A-GATC-T-GAGATGT-----	TAC-CCATGCGGGCGTTGTG-----	ATTGA----	CTAA----	C-----	A--CG--CGT		
TLONGI_14376	CGCGCTGCGATGGCGAGTGGT	GAGTTTTGCT----	GGT-CCACGGCCTGCATGTCAGT-----	TGAT-GATGT-----	C-----	TA-----	ACCGA----	GACG----	T-----	G--TG--TGT		
TR_60282	CGCGGTGCGATGGCGAGTGGT	GAGT-----	T----	GGGA-----	CGGTTTCATG-----	GTT-TCTG-CATATCA-----	GTG-A-----	TGGTGTCTAACC GG----	GACA----	T-----	TTGTG--CGC	
TRICITRO_110	CGCGGTGCGATGGCGAGTGGT	GAGT-----	T-----	TGTGCCGTTCCACGGGTTCTGCAGATGAAGTCAT-GATGTCT-----	A-----	ACTGA----	GACA----	T-----	G--TG--CGC			
THZ_79686	CGCGATGCGATGGAGAATGGT	AAGT----	CT----	GAT-GTTGTGGC--TGG--AGCTC----	CTCATCGT----	TGAT-----	AT-C-----	TG-----	ACTGA----	GATT----	T-----	G--TG--CCC
TH_701679	CGCGATGCGATGGAGAATGGT	AAGT----	CT----	GAT-GTTGTGGC--TGG--AGCTC----	CTCATCGT----	TGAT-----	AT-C-----	TG-----	ACTGA----	GATT----	T-----	G--TG--CCC
A0A1T3C8Z2	CGCGATGCGATGGAGAATGGT	AAGT----	CT----	GATGCT--GC--GGCTGAAGCTC----	CTCATCGT----	TGAT-----	AT-C-----	TG-----	ACTGA----	GATT----	T-----	G--TG--CCC
TGAM_8287	CTCGCTGTGACGGAGAATAGT	GAGT----	CTCTCTGTCTCTTTCTCATCCCGAT--	CTC----	TAAGATCT-GAGATCT-----	TAT-C-----	CG-----	TATGGCGTTGATA----	TGATTGAC--	TGACACGC		
OPB37586	CGCGATGCGATGGAGAATGGT	AAGT----	CT----	GATGCT--GC--GGCTGAAGCTC----	CTCATCGT----	TGAT-----	AT-C-----	TG-----	ACTGA----	GATT----	T-----	G--TG--CCC
cDNA OPB3758	CGCGATGCGATGGAGAATG											
cDNA Xlr2*	CACGATGCGATGGAGAATG											

Intron 1

	240	250	260	270	280
	*	*	*	*	*
Xlr2	GTCTCC--AGTCCTTGTCGTCGTTGTGAAGTAAGC-----	T----	A-GCT-----	TCGATT-----	G-----GGGTC
Xlr2-000202	GTCTCC--AGTCCTTGTCGTCGTTGTGAAGTAAGC-----	T----	A-GCT-----	TCGATT-----	CCTATTCTG-AATTGGGGTC
TAS_375582	GTTAAA--AGTCCCTGTCGTCGTTGCGAAGTAAGCAAAGTCTAAATCCTAACACT----	A-TCTCTCACTACCATCTCCCGTGGGATCCTTCTATT-----	GTTCCCGGGGGC		
TA_229644	GTGAAT--AGTCCCTGTCGTCGTTGCGAAGTAAGCAAAGCCTAAATCC-----	T----	A-CCA-----	TCTCTTACTACCAGGAATCTCTCGTGAGATCCTTCAATCGTTCCCGG-----	GGCTC
TLONGI_14376	GT---C--AGTCCTTGTCGACGTTGTGAAGTAAGCAGCATG-----	T----	C-TCA-----	TTCATT-----	G-----TCTAC
TR_60282	GT---C--AGTCCTTGTCGTCGTTGCGAAGTAAGCACCATG-----	T----	C-TCT-----	TCCATT-----	T-----GTCCT
TRICITRO_110	G-CGTC--AGTCCTTGCCGTCGTTGTGAAGTAAGC-----	A-GCA-----	GCTCTT-----	G-----TCTCT	
THZ_79686	GCGTGC--AGTCCTTGTCGCCGTTGTGAAGTAAGC-----	TCAATAA-GCA-----	GCTGTT-----	T-----CTATT	
TH_701679	GCGTGC--AGTCCTTGTCGCCGTTGTGAAGTAAGC-----	TCAATAA-GCA-----	GCTGTT-----	T-----CTATT	
A0A1T3C8Z2	GCGTGC--AGTCCTTGTCGCCGTTGTGAAGTAAGC-----	T----	AAGCA-----	GCTGTT-----	T-----CTGTT
TGAM_8287	GTGTCAATAGTCCCTGTCGTCGTTGCGAAGTAAGCAAAGTCTAAATCC-----	T----	A-CCA-----	TCTCTACTACCAGGTATCTCGCGAGATCCTTCAATGGTTCCC--	G-----GGCTC
OPB37586	GCGTGC--AGTCCTTGTCGCCGTTGTGAAGTAAGC-----	T----	AAGCA-----	GCTGTT-----	T-----CTGTT
cDNA OPB3758	-----TCCTTGTCGCCGTTGTGAAG-----				
cDNA Xlr2*	-----TCCTTGTCGTCGTTGTGAAG-----				

Intron 1

Microexon

Intron 2

	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470
	*	*	*	*	*	*	*	*	*	*	*	*	*
Xlr2	G	A	C	T	G	C	T	G	A	T	G	T	C
Xlr2-000202	G	A	C	T	G	C	T	G	A	T	G	T	C
TAS_375582	A	A	C	T	G	C	C	G	T	A	C	G	A
TA_229644	A	A	C	T	G	C	T	G	A	T	G	T	C
TLONGI_14376	G	A	C	T	G	C	C	G	G	C	T	C	A
TR_60282	G	A	C	T	G	C	C	G	G	C	T	C	A
TRICITRO_110	G	A	C	T	G	A	T	G	C	T	G	C	A
THZ_79686	G	A	C	T	G	C	C	G	G	C	T	C	A
TH_701679	G	A	C	T	G	C	C	G	G	C	T	C	A
A0A1T3C8Z2	G	A	C	T	G	C	T	G	A	T	G	T	C
TGAM_8287	A	A	C	T	G	C	T	G	A	T	G	T	C
OPB37586	G	A	C	T	G	C	T	G	A	T	G	T	C
cDNA OPB3758	G	A	C	T	G	C	T	G	A	T	G	T	C
cDNA Xlr2*	G	A	C	T	G	C	T	G	A	T	G	T	C

Intron 3

	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570
	*	*	*	*	*	*	*	*	*	*
Xlr2	-	C	C	A	T	T	A	-	-	-
Xlr2-000202	-	C	C	A	T	T	A	-	-	-
TAS_375582	-	T	C	G	T	T	A	-	-	-
TA_229644	T	C	T	G	T	A	-	-	-	-
TLONGI_14376	-	G	A	C	T	G	A	-	-	-
TR_60282	-	A	T	T	G	T	A	-	-	-
TRICITRO_110	-	T	G	C	-	-	-	-	-	-
THZ_79686	-	C	C	A	T	T	A	-	-	-
TH_701679	-	C	C	A	T	T	A	-	-	-
A0A1T3C8Z2	-	C	C	A	T	T	A	-	-	-
TGAM_8287	-	T	C	A	A	C	T	-	-	-
OPB37586	-	C	C	A	T	T	A	-	-	-
cDNA OPB3758	-	-	-	-	-	-	-	-	-	-
cDNA Xlr2*	-	-	-	-	-	-	-	-	-	-

Intron 3

	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720
	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Xlr2	C	C	G	A	A	A	C	C	A	C	G	G	T	G	A
Xlr2-000202	C	C	G	A	A	A	C	C	A	C	G	G	T	G	A
TAS_375582	C	C	C	A	A	A	C	T	A	A	T	C	A	A	T
TA_229644	C	C	G	A	A	A	C	T	A	A	T	C	A	A	T
TLONGI_14376	C	C	A	A	A	C	A	C	G	G	T	G	A	A	T
TR_60282	C	C	A	A	A	C	A	C	G	G	T	G	A	A	T
TRICITRO_110	C	C	G	A	A	A	C	A	C	G	G	T	G	A	T
THZ_79686	C	C	A	A	A	C	A	C	G	G	T	G	A	A	T
TH_701679	C	C	A	A	A	C	A	C	G	G	T	G	A	A	T
A0A1T3C8Z2	C	C	A	A	A	C	A	C	G	G	T	G	A	A	T
TGAM_8287	C	C	G	A	A	A	C	T	A	A	T	C	A	A	T
OPB37586	C	C	A	A	A	C	T	A	A	T	C	A	A	T	C
cDNA OPB3758	C	C	A	A	A	C	T	A	A	T	C	A	A	T	C
cDNA Xlr2*	C	C	G	A	A	A	C	T	A	A	T	C	A	A	T

173017401750176017701780179018001810182018301840185018601870

*

*

*

*

*

*

*

*

*

*

*

*

*

*

Xlr2

CTCAATCGAGAGGATGAACAATGTGCCGGACCTGTGACACCATGTGCAAACCCCCTTTGCTCGTGCAAAAAGGACCAAGATGGCAGCATATACAATATGGAGTTACGGGAAGAACGCCGGAGAGTCTTTTGGCTGCTCTACGCCTTGGAT

Xlr2-000202

CTCAATCGAGAGGATGAGCAATGTGCCGGACCTGTGACGCCATGTGCAAACCCC-TTTGCTCGTGCAAAAAGGACCAAGATGGCAGCATATACAATATGGAGTTGCGGGAAGAACGCCGGAGAGTCTTTTGGCTGCTCTACGCCTTGGAT

TAS_375582

CTCAATCGAGAAGACGAGCAGTGTGCTGGACCTGTAACTCCATGCGCAAATCCCCTCTGCTCATGTGCGAAAAGGAGCAGGACGGCAGCTTATACAATCTCGAGCATCGAGAGGAACGCCGGAGAGTCTTTTGGCTTCTCTACGCTCTGGAC

TA_229644

CTCAATCGAGAAGACGAGCAGTGTGCCGGGCCCCGTAACTCCTTGCGCGAATCCCCTCTGCTCATGTGCGAAAAGGAGCAGGACGGCAGCTTGTACAATGTGGAGCATCGAGAGGAACGCAGGAGAGTCTTTTGGCTTCTGTACGCTCTGGAT

TLONGI_14376

CTCAACCGGGAGGATGAGCAGTGCGCCGGGCCCCGTAACGCCTTGCGCAAACCCCTCTCTGCTCATGTGCGTAAGGAGCAGGATGGCAGCCTGTATAACCTCGAGCTGCGAGAAGAGCGCCGGCGGGTCTTTTGGCTTCTCTATGCCCTGGAT

TR_60282

CTCAACCGAGAGGACGAGCAATGCGCCGGACCCGTGACGCCCTGCGCGAATCCTCTCTGCTCGTGTGCGAAAAGAGCAGGATGGTAGCTTGTACAATCTCGAGCTGCGAGAAGAGCGCCGGCGAGTCTTTTGGCTTCTCTATGCCTTGGAC

TRICITRO_110

CTGCAACCGGGAGGATGAACAGTGCGCCGGGCCCCGTAACGCCGTGCGCAAACCCCTCTGCTCATGTGCGAAAAGAGCAAGATGGTAGTCTGTATAATCTCGAGCTGCGAGAAGAGCGCCGAAGAGTCTTTTGGCTTCTCTATGCCCTGGAC

THZ_79686

CTACACCGAGAGGACGAGCAATGTGCCGGGCCTGTGACGCCGTGCGCAAACCCCCTTTGCTCGTGCAAAAAGACCAAGATGGCAGCATATATAACTTGGAGCTACGAGAAGAGCGTCGAAGAGTCTTTTGGCTTCTCTACGCTCTGGAT

TH_701679

CTACACCGAGAGGACGAGCAATGTGCCGGGCCTGTGACGCCGTGCGCAAACCCCCTTTGCTCGTGCAAAAAGACCAAGATGGCAGCATATATAACTTGGAGCTACGAGAAGAGCGTCGAAGAGTCTTTTGGCTTCTCTACGCTCTGGAT

A0A1T3C8Z2

CTACACCGAGAGGATGAGCAATGTGCCGGACCTGTGACCCCTTGCGCAAACCCCCTTTGCTCATGCAAAAAGGATCAAGATGGCAGCATATATAACTTGGAGCTACGAGAAGAGCGCCGAAGAGTCTTTTGGCTTTTATACGCCCTGGAT

TGAM_8287

CTCAATCGAGAAGACGAGCAATGCGCTGGACCCGTAACTCCTTGTGCAAATCCTCTTTGTTTCATGCCGAAAGGAGCAGGATGGCAGCTTGTACAATCTGGAGCATCGTGAGGAACGCAGGAGAGTCTTTTGGCTTCTGTACGCTCTGGAT

OPB37586

CTACACCGAGAGGATGAGCAATGTGCCGGACCTGTGACCCCTTGCGCAAACCCCCTTTGCTCATGCAAAAAGGATCAAGATGGCAGCATATATAACTTGGAGCTACGAGAAGAGCGCCGAAGAGTCTTTTGGCTTTTATACGCCCTGGAT

cDNA OPB3758

CTACACCGAGAGGATGAGCAATGTGCCGGACCTGTGACCCCTTGCGCAAACCCCCTTTGCTCATGCAAAAAGGATCAAGATGGCAGCATATATAACTTGGAGCTACGAGAAGAGCGCCGAAGAGTCTTTTGGCTTTTATACGCCCTGGAT

cDNA Xlr2*

CTCAATCGAGAGGATGAACAATGTGCCGGACCTGTGACACCATGTGCAAACCCCCTTTGCTCGTGCAAAAAGGACCAAGATGGCAGCATATACAATATGGAGTTACGGGAAGAACGCCGGAGAGTCTTTTGGCTGCTCTACGCCTTGGAT

18801890190019101920193019401950196019701980

*

*

*

*

*

*

*

*

*

*

*

Xlr2

CGCCATCTAGCCTTGTCTTTCAACTCGGCTCTAACGATACCCGATTCACTACTGCGAAGTTTATTGTTAGT--A--T--A--C--CA-A--T----T--CGTT-C--T-TGA-CTG-GCTT-T-----TGTT---CTAACTT-T--T--G-

Xlr2-000202

CGCCATCTAGCCTTGTCTTTCAACTCGGCTCTAACGATACCCGATTCACTACTGCGAAGTTTATTGTTAGT--A--T--A--C--CA-A--T----T--CGTT-C--T-TGA-CTG-GCTT-T-----TGTT---CTAACTT-T--T--G-

TAS_375582

AGGCATCTGGCTTTGTCTTTTCAACTCGGCTTTGACAATCCCCGATTCTTACTGTGAAGTTTATTGTAAGTGA--CTGA--G--TA-A--C----C--CATT-C--TGCTT-TTGTGCTTGA-----TTAA---CCAACCT-T--T--G-

TA_229644

CGGCATCTGGCCTTGTCTTTTCAACTCGGCGTTGACGATACCTGATTCTTACTGCGAAGTCTATTGTAAGTA-AC TT--AGGC--TA-A--GTAAT--TGTTAA--TATGT-TCT-GCTT-T-----CGTTAAACTAAC-C-C--T--G-

TLONGI_14376

CGGCATTTGGCCTTGTCTTTTCAACTCTGCTCTCACAATCCCTGACTCGTACTGTGAGTTTACTGTAAGT--A--T-----C--CA-A-----G--CCTG-C--T-TCA-GCA-GCTTCT-----TCTT---CACTCTT-T--CTAA-

TR_60282

CGGCACCTGGCCCTGTCTTTTCAACTCTGCCCTTACCATCCCTGACTCGTACTGCGAAGTCTACTGTAAGG--A--C--C--A--AA-A--C-----C--TGCC-C--C-CAATCTG-TTCA-T-----TGTC---CACCTT-TGCT--GA

TRICITRO_110

CGGCACCTGGCCTTGTCTTTTCAACTCTGCTCTCACAATTCCTGACTCGTACTGCGAAGTTTACTGTAAGT----T--T--C--CA-AACA----T--TGCT-C--T-AAC-CTGTTTCT-T-----TGATTA-CTCCCTTCT--A--A-

THZ_79686

CGCCATCTAGCCTTGTCTTTTCAACTCGGCTCTAACAATACCCGACTCTTATTGCGAAGTTTATTGTAAGT--A--T--A--C--TATA--TGA--TG-CCTC-CAGC-TTA-TTG-ACTG-G-----CCTTG--CTAA-TT-C--T--G-TH_701679

CGCCATCTAGCCTTGTCTTTTCAACTCGGCTCTAACAATACCCGACTCTTATTGCGAAGTTTATTGTAAGT--A--T--A--C--TATA--TGA--TG-CCTC-CAGC-TTA-TTG-ACTG-G-----CCTTG--CTAA-TT-C--T--G-

A0A1T3C8Z2

CGCCATCTAGCTTTGTCTTTTCAACTCGGCTCTAACAATACCTGACTCTTATTGTGAAGTTTATTGTAAGC--A--T--A--C--TA-T--G----T--TGTG-CC-T-TCA-CCC-TCCG-ACCAACTGTTG--CTAA-TT-C--T--A-

TGAM_8287

CGGCATCTGGCCTTGTCTTTTCAACTCGGCGCTGACAATACCTGATTCACTACTGCGAAGTCTATTGTAAGT--ACCT--AG-ACTGA-A--TG---ATCCGTT-C--T-G---CTT-TC-C-A-----TAAA---CTAAC-C-T--T--G-

OPB37586

CGCCATCTAGCTTTGTCTTTTCAACTCGGCTCTAACAATACCTGACTCTTATTGTGAAGTTTATTGTAAGC--A--T--A--C--TA-T--G----T--TGTG-CC-T-TCA-CCC-TCCG-ACCAACTGTTG--CTAA-TT-C--T--A-

cDNA OPB3758

CGCCATCTAGCTTTGTCTTTTCAACTCGGCTCTAACAATACCTGACTCTTATTGTGAAGTTTATT-----

cDNA Xlr2*

CGCCATCTAGCCTTGTCTTTTCAACTCGGCTCTAACGATACCCGATTCACTACTGCGAAGTTTATT-----

Intron 4

199020002010202020302040205020602070208020902100211021202130

*

*

*

*

*

*

*

*

*

*

*

*

*

*

Xlr2

-C--T--C--TCAGCTCCGCTACCTGAAGCTGTTTGGGAAAACCTTGATACAATCCCGCCAGCGAAATCCCGCTCGAAGGATCGGACCGCCAACTACCGCCTCTGGATCTGCCTTTTTTTGAATACTTTTTGCCCTCATGGCGATACT

Xlr2-000202

-C--T--C--TCAGCTCCGCTACCTGAAGCTGTTTGGGAAAACCTCGATACAATCCCGCCAGCGAAATCCCGCTCGAAGGATCGGACCGCCAAACACCGCCTCTGGATCTGCCTTTTTTTGAATACTTTTTGCCCTCATGGCGATACT

TAS_375582

-C--A--A--CCAGCCCCCTCTGCCTGAAGCTGTTTGGGAAAACCTCGACACAATTCCCTCCAAACGAGATTCCCGTTTCGAACAATTGGACCACCAACGACTGCCTCCGGATCTGCCTTTTTTCGAATACTTCTTACCCTCATGGCTATACT

TA_229644

-T--T--CA-TCAGCCCCCTACCTGAAGCTATTTGGGAGAATCTCGACACCATTCCTCCCAACGAGATGCCCGTTTCGACCATTTGGACCGCCAACACGGCCTCTGGATCTGCCTTTTTTCGAATACTTTTTTGCCACTCATGGCTATACT

TLONGI_14376

-CA-A--GCTTCAGCCCCGCTTCCTGAAGCCATATGGGAGAACCTCGACACGATCCCACCCAGCGAAATCCCGCTCGAAGAATAGGACCGCCAACGACTGCCTCTGGATCTGCCTTTTTTTGAGTACTTCTTACCTCTCATGGCGATACT

TR_60282

CCGGC--C--TCAGCACCTCTTCCCGAAGCCATCTGGGAGAACCTCGACACGATTTTCGCCAGTGAGATTCCCGCCCCGAAGAATCGGACCGCCCACGACTGCCTCTGGATCTGCATTTTTTCGAGTACTTTTTACCCTCATGGCGATACT

TRICITRO_110

-C--TGACT-TCAGCCCCCTCCCCGAAGCCATCTGGGAAAACCTCGACACGATCCCACCCAGCGAGATTCTTGCCCGAACAATCGGACCGCCGACGACTGCCTCTGGGTCTGCATTTTTTTGAGTACTTTCTACCCTCATGGCGATACT

THZ_79686

-C--T--C--TTAGCCCCGTTGCCTGAAGCTGTTTGGGAAAACCTCGATACAATCCCACCTAGCGAAATCCCGCCCCGAAGGATCGGACCGCCAACGACTGCCTCTGGGTCTGCCTTTTTTCGAATACTTTTTTGCCACTCATGGCCATACT

TH_701679

-C--T--C--TTAGCCCCGTTGCCTGAAGCTGTTTGGGAAAACCTCGATACAATCCCACCTAGCGAAATCCCGCCCCGAAGGATCGGACCGCCAACGACTGCCTCTGGGTCTGCCTTTTTTCGAATACTTTTTTGCCACTCATGGCCATACT

A0A1T3C8Z2

-C--T--C--TCAGCCCCGTTGCCTGAAGCTGTTTGGGAAAACCTCGATACAATCCCACCCAGCGAAATCCCGCTCGAAGGATCGGACCGCCAACGACTGCCTCTGGGTCTGCCTTTTTTCGAATACTTTTTTGCCGCTCATGGCGATACT

TGAM_8287

-T--T--CA-TCAGCCCCCTACCTGAAGCTGTTTGGGAAAATCTCGACACCATTCCTCCCAACGAGATGCCCGTTTCGAACCATTTGGACCACCAACAACAGCCTCTGGATCTGCCTTTTTTCGAATATTTCTTACCCTCATGGCTATACT

OPB37586

-C--T--C--TCAGCCCCGTTGCCTGAAGCTGTTTGGGAAAACCTCGATACAATCCCACCCAGCGAAATCCCGCTCGAAGGATCGGACCGCCAACGACTGCCTCTGGGTCTGCCTTTTTTCGAATACTTTTTTGCCGCTCATGGCGATACT

cDNA OPB3758

-----CCCCGTTGCCTGAAGCTGTTTGGGAAAACCTCGATACAATCCCACCCAGCGAAATCCCGCTCGAAGGATCGGACCGCCAACGACTGCCTCTGGGTCTGCCTTTTTTCGAATACTTTTTTGCCGCTCATGGCGATACT

cDNA Xlr2*

-----CTCCGCTACCTGAAGCTGTTTGGGAAAACCTTGATACAATCCCGCCAGCGAAATCCCGCTCGAAGGATCGGACCGCCAACTACCGCCTCTGGATCTGCCTTTTTTTGAATACTTTTTGCCCTCATGGCGATACT

Intron 4

	2140	2150	2160	2170	2180	2190	2200	2210	2220	2230	2240	2250	2260	2270	2280
	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Xlr2	GGGCGACATCATAGAAGTCC	ACCATAGGCGCCGTCATCCA	AGGCTTGGAGATCTCGACG	ATTACATTCTGTAGCTGTTA	TTCAGGGGGCTTCTTTGAC	CTACGAGATGAGTCTTGCCG	CGCTTTCGTCGGAGGGTA	ATGACAGCAACAT							
Xlr2-000202	GGGCGACATCATAGAAGTCC	ACCATAGGCGCCGTCATCCA	AGGCTTGGAGATCTCGACG	ATTACATTCTGTAGCTGTTA	TTCAGGGGGCTTCTTTGAC	CTACGAGATGAGTCTTGCCG	CGCTTTCGTCGGAGGGTA	ATGACAGCAACAT							
TAS_375582	AGGAGACATTATAGAAGTCC	ATCACAGACGCCGCCATCCA	AAGACTTGGAGATCTGGACG	ATTGCACTCCGTATCTATT	ATCCAAGGACTACTTACGAC	GTAACGAACTGAGTCTTGCCG	CGCTTTCATCAGAAGGAA	ATGACAACGGCAT							
TA_229644	TGGAGACATCATAGAGGTCC	ATCACAGGCGCCGTCATCCA	CGGCTCGGAGATTTAGACG	ATTACACTCCGTGTCTGT	CATCCAGGGGGCTTCTGACA	ACATACGAACGAGTCTCGCG	GGCACTCTCGTCAGAA	GGCAACGACAGTGGCAT							
TLONGI_14376	AGGCGACATCATCGAAGTCC	ATCATCGGCGCCGTCACCCG	AGACTCGGAGGTCTCGACG	ACTCACACTCCGTGGCCGT	CATTAGGGGGCTGCTTTGAC	CTACGAAATGAGTCTCGCCG	CGCTTTCGCCCGACGGC	GATGACAACGGCAC							
TR_60282	AGGCGACATTATAGAGGTCC	ATCATCGGCGCCGTCACCCG	AGACTCGGAGGCCTCGACG	ATTACACTCCGTGGCCGT	CGTTAGGGGGCTGCTTTG	GCCTACGAAATGAGTCTCG	CCGCGCTGTACCAGACG	GGCGACGACAACGGCAC							
TRICITRO_110	GGGCGACATCATAGAAGTTCA	TATCATCGGCGCCGTCATCCG	AGACTCGGAGGCCTCGACG	ATTACACTCTGTAGCCGTA	ATCCAGGGGACTGCTTTGAC	CTACGAGTTGAGTCTCGCCG	CGCTTTCACCAGACGGC	GACGACAACGGCAC							
THZ_79686	GGGCGACATTATCGAAGTCC	ACCACAGACGCCGTCATCCA	AAGACTGGCAGGTCTCGAC	GACTCGCACTCTGTAGCCG	TATCCAGGAGCTCCTTTCC	ACCTACGAGCTCAGTCTTG	CCGCTCTTTACCAGAGAG	CAATGATAACAACAC							
TH_701679	GGGCGACATTATCGAAGTCC	ACCACAGACGCCGTCATCCA	AAGACTGGCAGGTCTCGAC	GACTCGCACTCTGTAGCCG	TATCCAGGAGCTCCTTTCC	ACCTACGAGCTCAGTCTTG	CCGCTCTTTACCAGAGAG	CAATGATAACAACAC							
A0A1T3C8Z2	GGGTGACATTATAGAAGTCC	ACCACAGGCGCCGTCATCCA	AAGACTGGCAGGTCTCGAC	GACTCGCACTCTGTAGCCG	TATCCAGGAACTCCTTTCC	ACCTACGAGCTCAGTCTTG	CCGCTCTTTGCCAGAGAG	CAATGATAGCAGCAC							
TGAM_8287	AGGAGACATCATAGAGGTTC	ATCACAGGCGCCGTCACCC	ACGCCTTGGAGATTTAGACG	ATTACACTCCGTGTCTGT	CATCCAAGGGGCTTCTCACA	ACATACGAATTGAGTCTCG	CGGCACTCTCGTCAGA	AGGCAACGACAGCGGCAT							
OPB37586	GGGTGACATTATAGAAGTCC	ACCACAGGCGCCGTCATCCA	AAGACTGGCAGGTCTCGAC	GACTCGCACTCTGTAGCCG	TATCCAGGAACTCCTTTCC	ACCTACGAGCTCAGTCTTG	CCGCTCTTTGCCAGAGAG	CAATGATAGCAGCAC							
cDNA OPB3758	GGGTGACATTATAGAAGTCC	ACCACAGGCGCCGTCATCCA	AAGACTGGCAGGTCTCGAC	GACTCGCACTCTGTAGCCG	TATCCAGGAACTCCTTTCC	ACCTACGAGCTCAGTCTTG	CCGCTCTTTGCCAGAGAG	CAATGATAGCAGCAC							
cDNA Xlr2*	GGGCGACATCATAGAAGTCC	ACCATAGGCGCCGTCATCCA	AGGCTTGGAGATCTCGACG	ATTACATTCTGTAGCTGTTA	TTCAGGGGGCTTCTTTGAC	CTACGAGATGAGTCTTGCCG	CGCTTTCGTCGGAGGGTA	ATGACAGCAACAT							

	2290	2300	2310	2320	2330	2340	2350	2360	2370	2380	2390	2400	2410	2420
	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Xlr2	AATACCTCTCCCTATTCA	CACCCCCAAAGGCATTG	TGTGCTGGAATGCCTT	CAAGGC---CCT--C--	CAT-C-TCTCACCCGC	ACCCATCATCTGCAGCG	CCCATACAATGCGGGT	GATCCTTCCAAGATG	CGGGCTCGCAAAAGC	CTATAG				
Xlr2-000202	AATTCCCCTCCCTATTCA	CACCCCCAAAGGCATTG	TGTGCTGGAATGCCTT	CAAGGC---CCT--C--	CAT-C-TCTCACCCGC	ACCCATCATCTGCAGCG	CCCATACAATGCGGGC	GATCCTTCCAAGATG	CGGGCTCGCAAAAGC	CTATAG				
TAS_375582	AGTTTCCCTGCCTATCC	ACACTCCAAAGGGCATG	GTTGCCGGCATGCCGT	CGAGGTCATCCATGTCT	CATCCATCTCATCCAT	ACCCACCTTCTACAGCG	CCCATACAACGCGAG	GTGATCCCTCCAAG	ATGCGCCTAGCCAAGG	CCTATAG				
TA_229644	AGTTCCCCTGCCCATCC	ACACGCCAAGGGGCATG	GTCGCCGGCATGCCTT	CGCGGT---CAT--C--	CAT-G-TCCCATCCGT	ATCCGCCCTCCACAGCG	CCCATACAACGCAAG	CGATCCGTCCAAG	ATGCGCCTCGCCAAGG	CCTACAG				
TLONGI_14376	AATGCCTCTGCCGATTCA	CACACCCCAAGGGAATAG	CTTCTGGAATGCCTT	CCAGGC---CGT--C--	CAT-T-TCCCATCCAC	ACCCGTC---CGCAGCT	CCGTACAATGCAAGG	CGATCCCTCGAAG	ATGCGACTGGCAAAGG	CCTATAG				
TR_60282	CATGCCCTGCCTATTCA	CACACCCCAAGGGATCG	TTGCCGGAATGCCTT	CCAGGC---CAT--C--	CAT-T-TCCCATCCAC	ACCCGTC---CGCAGCC	CTTACAACGCAAGG	CGATCCCTCCAAG	ATGCGGCTGGCAAAGG	CCTATAG				
TRICITRO_110	CATGCCCTGCCTATTCA	CACACCCCAAGGGATG	CTTGCCGGAATGCCTT	CCAGGC---CAT--C--	CAT-T-TCCCATCCAC	ACCCGTC---CGCAGCC	CTTACAACGCAAGG	CGATCCCTCCAAG	ATGCGGCTGGCAAAGG	CCTATAG				
THZ_79686	CATGCCTTTGCCTATCC	ATACCCCAAGGGGCATG	CTCGCCGGAATGCCTT	CGAGGC---CAT--C--	CAT-C-TCTTATCAGC	ACCCGCCCTCTGCAGCT	CCATAACAACGCAAG	TGACCCTTCCAAG	ATGCGGCTAGCAAAGG	CCTATAG				
TH_701679	CATGCCTTTGCCTATCC	ATACCCCAAGGGGCATG	CTCGCCGGAATGCCTT	CGAGGC---CAT--C--	CAT-C-TCTTATCAGC	ACCCGCCCTCTGCAGCT	CCATAACAACGCAAG	TGACCCTTCCAAG	ATGCGGCTAGCAAAGG	CCTATAG				
A0A1T3C8Z2	CATGCCTTTGCCTATCC	ATACCCCAAGGGCATG	GTCGCCGGAATGCCTC	CGAGAC---CAT--C--	CAT-C-TCGCACCAGC	ACCCGCCCTCTGCAGCT	CCATAACAACGCAAG	CGACCCGTCCAAG	ATGCGGCTAGCAAAGG	CCTATAG				
TGAM_8287	CGTTCCCCTGCCCATCC	ACACGCCAAGGGGCATG	GTCGCCGGCATGCCTT	CGCGAT---CAT--C--	CAT-G-TCCCATCCGT	ATCCTCCATCTACAGCG	CCCATACAACGCGAG	TGATCCGTCCAAG	ATGCGCCTCGCCAAGG	CCTACAG				
OPB37586	CATGCCTTTGCCTATCC	ATACCCCAAGGGCATG	GTCGCCGGAATGCCTC	CGAGAC---CAT--C--	CAT-C-TCGCACCAGC	ACCCGCCCTCTGCAGCT	CCATAACAACGCAAG	CGACCCGTCCAAG	ATGCGGCTAGCAAAGG	CCTATAG				
cDNA OPB3758	CATGCCTTTGCCTATCC	ATACCCCAAGGGCATG	GTCGCCGGAATGCCTC	CGAGAC---CAT--C--	CAT-C-TCGCACCAGC	ACCCGCCCTCTGCAGCT	CCATAACAACGCAAG	CGACCCGTCCAAG	ATGCGGCTAGCAAAGG	CCTATAG				
cDNA Xlr2*	AATACCTCTCCCTATTCA	CACCCCCAAAGGCATTG	TGTGCTGGAATGCCTT	CAAGGC---CCT--C--	CAT-C-TCTCACCCGC	ACCCATCATCTGCAGCG	CCCATACAATGCGGGT	GATCCTTCCAAGATG	CGGGCTCGCAAAAGC	CTATAG				

	2430	2440	2450	2460	2470	2480	2490	2500	2510	2520	2530	2540	2550	2560	2570
	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Xlr2	TACGCACATTCTACACGT	GCTACATGTGCTTCTAC	ACGGCAAATGGGATGCC	CATATCCATGTTGGAC	GACGGCGACGACTGG	ATTACGTCCAAGAGATT	CAACGAGTGCGCGTCA	CATGCCATTTCCGCCT	CACAATCGGTCTCG	ACCAT					
Xlr2-000202	TACGCACATTCTACACGT	GCTACATGTGCTTCTAC	ACGGCAAATGGGATGCC	CATATCCATGTTGGAC	GACGGCGACGACTGG	ATTACGTCCAAGAGATT	CAACGAGTGCGCGTCA	CACGCCATTTCCGCCT	CACAATCGGTCTCG	ACCAT					
TAS_375582	CACGCATATCCTGCATGT	GCTGCATGTACTCCTAC	ACGGCAAATGGGATGCC	CATATCCATGTTGGAC	GATGGCGACGATTGG	ATCACGTCCAAGAGATT	CAACGAATGCGCATCA	CATGCCATTTCTGCTT	CGCAGTCTGTCTCG	ACCAT					
TA_229644	CACGCACATTCTCCATGT	GCTGCATGTGCTCCTCA	CGGCAAATGGGATGCC	CATATCCATGCTGGAC	GACGGCGATGACTGG	ATCACATCTAAAGATT	CAACGAGTGCGCGTCA	CACGCCATTTCTGCGT	CGCAGTCCGTTTCG	ACAAT					
TLONGI_14376	CACGCACATCCTGCACGT	GTTGCATGTACTGCTG	CACGGAAAATGGGATGCC	ATCTCCATGCTGGAC	GACGGCGACGATTGG	ATCACTTCCAAGAGGTT	CAACGAGTGCGCGTCA	CATGCCATCTCTGCGT	CGCAGTCTGTCTCG	ACTAT					
TR_60282	CACGCACATCTTGACGTG	CTGCACGTGCTGCCATG	GAAAAATGGGATGCC	ATCTCCATGCTGGAC	GACGGCGACGATTGG	ATCACGTCCAAGAGTT	CAACGAATGCGCGTCA	CATGCCATTTCTGCGT	CGCAGTCCGTCTCG	ACTAT					
TRICITRO_110	CACGCATATTCTGCACGT	GCTGCATGTCTTGCTAC	ACGGCAAGTGGGATGCC	ATCTCCATGCTGGAC	GACGGCGACGATTGG	ATCACGTCCAAGAGTT	TAATGAGTGCGCGTCA	CATGCCATTTCTGCGT	CGCAATCCGTCTCG	ACCAT					
THZ_79686	CACCCACATTCTACACGT	ATTGCATGTGCTTCTG	CACGGCAAATGGGATG	CTATATCTATGTTGG	ACGACGGCGACGATTGG	ATCACGTCCAAGAGATT	CAACGAGTGCGCGTCT	CACGCCATTTCTGCGT	CGCAATCGGTCTCG	ACCAT					
TH_701679	CACCCACATTCTACACGT	ATTGCATGTGCTTCTG	CACGGCAAATGGGATG	CTATATCTATGTTGG	ACGACGGCGACGATTGG	ATCACGTCCAAGAGATT	CAACGAGTGCGCGTCT	CACGCCATTTCTGCGT	CGCAATCGGTCTCG	ACCAT					
A0A1T3C8Z2	CGCTCACATTCTGCACGT	ATTGCATGTGCTTCTAC	ATGGCAAATGGGATG	CTATATCTATGTTGG	ACGACGGCGACGATTGG	ATCACGTCCAAGAGATT	CAACGAGTGCGCGTCT	CACGCCATTTCTGCAT	CACAATCCGTCTCG	ACCAT					
TGAM_8287	CACGCATATTCTCCATGT	GCTGCATGTACTCCTG	CACGGTAAATGGGATGCC	CATATCCATGCTGGAC	GATGGCGATGACTGG	ATCACATCCAAGAGATT	CAACGAGTGCGCGTCA	CATGCCATTTCTGCGT	CGCAGTCTGTTTCA	ACAAT					
OPB37586	CGCTCACATTCTGCACGT	ATTGCATGTGCTTCTAC	ATGGCAAATGGGATG	CTATATCTATGTTGG	ACGACGGCGACGATTGG	ATCACGTCCAAGAGATT	CAACGAGTGCGCGTCT	CACGCCATTTCTGCAT	CACAATCCGTCTCG	ACCAT					
cDNA OPB3758	CGCTCACATTCTGCACGT	ATTGCATGTGCTTCTAC	ATGGCAAATGGGATG	CTATATCTATGTTGG	ACGACGGCGACGATTGG	ATCACGTCCAAGAGATT	CAACGAGTGCGCGTCT	CACGCCATTTCTGCAT	CACAATCCGTCTCG	ACCAT					
cDNA Xlr2*	TACGCACATTCTACACGT	GCTACATGTGCTTCTAC	ACGGCAAATGGGATGCC	CATATCCATGTTGGAC	GACGGCGACGACTGG	ATTACGTCCAAGAGATT	CAACGAGTGCGCGTCA	CATGCCATTTCCGCCT	CACAATCGGTCTCG	ACCAT					

Figure S3. Alignment of Xlr2 encoding gene orthologs in diverse *Trichoderma* species. The alignment includes cDNA obtained from our *Trichoderma virens* Gv29.8 RNA-seq data (cDNA Xlr2*, Castañeda-Casasola, Lawry et al., in preparation), predicted mRNA sequences from *T. guizhouense* (cDNA OPB3758) and *T. parareesei* (OTA07803.1), as well as the extended 5’ region of various *Trichoderma* gene orthologs. The presence of the microexon sequence is observed in all Xlr2 orthologs from the *Trichoderma* species analyzed in this study. Introns are indicated by orange bars. The sequence Xlr2-000202 corresponds to the ortholog in *T. virens* FT-333, the gene reported in NCBI corresponds to TrVFT333_000202. It was obtained from a search using NCBI Blast. The orthologous gene was identified on chromosome 1 (<https://www.ncbi.nlm.nih.gov/nuccore/CP071115.1?from=816716&to=819677>). The correspondence names between this figure and figure 4B is as follow:

This Figure	Figure 4B
Xlr2	PP712105
Xlr2-000202	Xlr2-000202**
TAS_375582	TAS_375582
TA_229644	Triat2_229644
TLONGI_14376	Trilo3_1437635
TR_60282	TRQM6a_60282
TRICITRO_110	Trici4_1109904
THZ_79686	Triha1_79686
TH_701679	Trihar1_701679
A0A1T3C8Z2	A0A1T3C8Z2
TGAM_8287	Trigam1_8287
OPB37586	OPB37586
cDNA OPB3758	cDNA OPB3758
cDNA Xlr2*	cDNA PP712105

** this is only present in this table and corresponds to TrVFT333_000202 encoding gene see the figure legend.