

## IDENTIFICATION

**Species:** *Theobroma cacao*

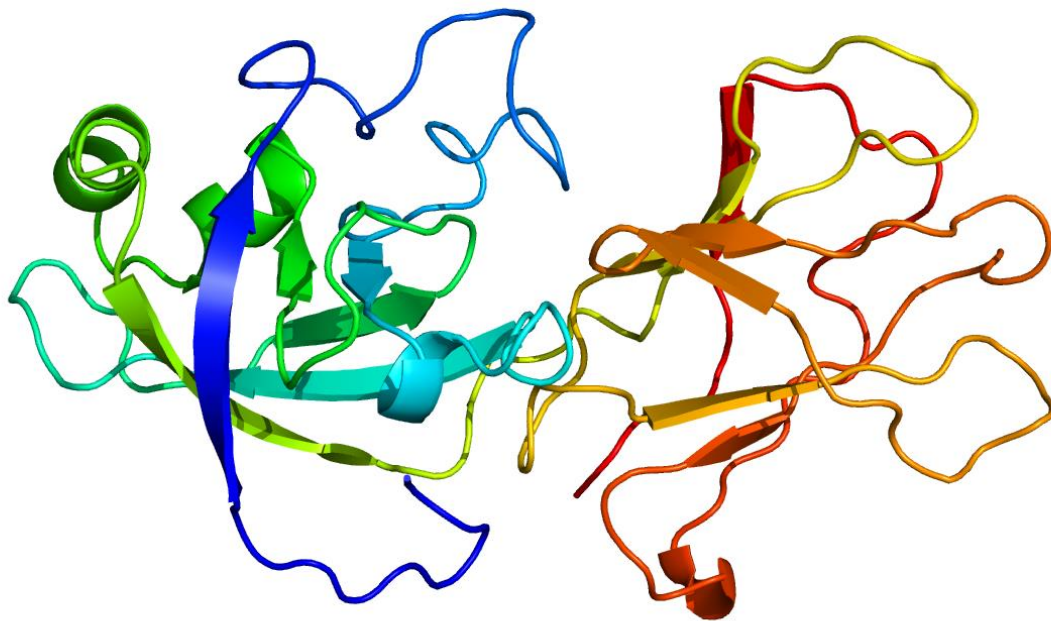
**Locus:** Thecc.03G234700

**Gene Model:** Thecc.03G234700.1.p

**Description:** TcEXLB-01

**Family:** Expansin Like Beta

**3D structure:**



## GENOME DATABASES

Phytozome: [https://phytozome-next.jgi.doe.gov/info/Tcacao\\_v2\\_1](https://phytozome-next.jgi.doe.gov/info/Tcacao_v2_1)

KEGG: <https://www.genome.jp/entry/gn:T02994>

## EXTERNAL RESOURCES

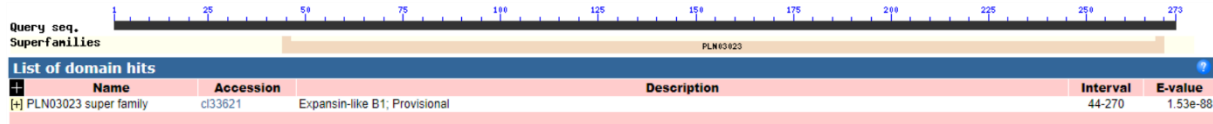
<https://www.cacaogenomedb.org/>

<https://cocoa-genome-hub.southgreen.fr/node/4>

## GENE STRUCTURE



## DOMAIN ARCHITECTURE



## SEQUENCES

### Peptide

>TcEXLB-01

MPLSSLISLSLLIESSMALS LQYLLILFAALLLTQSLAEAATCSNYFTHSRAAYYPNSDE  
HGTDV GACGFGSFGATINGGDVSTASELYRNGV GCGACYQVRCTNSNYCSDDGVTV  
VITDQ GSGPNTDFILSRRAFGRMAQT TDAASLLAQGVVDIEYRRVSCSYPNKNITIKI  
DENSNYPHYLAFVLWYQQGDKDITAVQLCETKNFVCKLLDRSYGAVWTTTSPPTGP  
LSLRMLLSGEDGDES WFPINNIPENWKGGETYDTGVQINI\*

### CDS (coding sequence)

>TcEXLB-01

ATGCCCTTATCTTCTTTGATTTTCATTGTCTTTGCTCATTGAATCCTCTATGGCTCTT  
TCTCTTCAATACCTTTTGATCCTTTTCGCAGCTTTGCTTCTCACACAAAGTCTGGCA  
GAGGCAGCAACATGCAGCAACTATTTACCCATTACGAGCAGCCTACTATCCAA  
ACTCTGATGAACATGGAACAGATGTAGGTGCATGCGGTTTTGGCTCCTTTGGAGC  
ACAATAAATGGTGGAGATGTATCCACAGCATCTGAACTCTACCGTAATGGTGTT  
GGTTGTGGTGCTTGCTACCAGGTGAGGTGCACCAACAGCAACTATTGCTCAGATG  
ACGGTGTGACAGTGGTTATAACTGATCAAGGCTCAGGTCCCAATACCGACTTTAT  
CCTGAGCAGACGAGCCTTTGGTAGGATGGCTCAGACCACAGATGCAGCTGCATCT  
CTATTAGCCCAAGGAGTTGTTGACATTGAATATAGGCGTGTTTCATGCAGCTACC  
CCAACAAAATATCACTATCAAGATTGATGAGAACAGCAATTACCCTCATTATTT  
GGCATTTGTTCTTTGGTATCAACAAGGTGACAAGGATATCACAGCAGTGCAGCTA  
TGTGAGACTAAAACCTTTGTGTGCAAGCTATTGGATAGGAGTTACGGAGCAGTAT  
GGACAACCACTTCAACCCCAACCGGTCCATTGTCTTTAAGGATGCTATTGAGTGG  
AGAGGATGGAGATGAGTCATGGTTTGTTCCTATTAATAACATAACCAGAAAATTGG  
AAAGGTGGAGAAACATATGACACAGGAGTTCAAATTAATATCTAG

### Nucleotide

>TcEXLB-01

AATGACGATATCGGTAGCAGA AACTTTGAAGATTGCAAACAGAACGAAATCTTCAT  
CTTCAAATGGCAGTTTGGGAGAACATTAATTAATAAACA ACTACAATAAAGTC  
TAGCTGCATGACTATTGTCAAACACTCGTAGGAGATTTCAATAAGGGAAGGAA  
AAGTGCTTCAGAATTTGAAAGTCAAGAGGGATGTTCCACATAGTTCCAAAACAAG  
TACTTCTAAGTCTTAAAAGATAGCTGGGAAATACAATCCTAACAGAATTCCACTC

AACGTTTATTTACAGAATGTCAAAGTTGGCAAGGTTTACCAGACTAGGATATTTT  
GGCTGCTTGCTTCACTATAAATACCAGCTGCAGACCCCGTTATTCATGCCCTTAT  
CTTCTTTGATTTCAATTGTCTTTGCTCATTGAATCCTCTATGGCTCTTTCTTTCAAT  
ACCTTTTGATCCTTTTCGCAGCTTTGCTTCTCACACAAAGTCTGGCAGAGGCAGCA  
ACATGCAGCAACTATTTACCCATTCACGAGCAGCCTACTATCCAAACTCTGATG  
AACATGGAACAGATGGTAAGGGAATACATACAAGGAATGCGTACCATGTATTGT  
GCATAATACCAACACAAATGCTCACAAATTTTCTTAACTGAATTCCAGTAGGTGC  
ATGCGGTTTTGGCTCCTTTGGAGCAACAATAAATGGTGGAGATGTATCCACAGCA  
TCTGAACTCTACCGTAATGGTGTGGTTGTGGTGCTTGCTACCAGGTATTAAGA  
AAGATGTGCCCTTGCACATGTGACGAAATTTAAGTATAGATGCCCTTTTACATGC  
TTAAATAATATTAACATGCTCTTTTATGATGCAAATCGCACTTTCAATTTACAT  
TGAACCAATTCTTACTTTTGCAAGGTGAGGTGCACCAACAGCAACTATTGCTCAG  
ATGACGGTGTGACAGTGGTTATAACTGATCAAGGCTCAGGTCCCAATACCGACTT  
TATCCTGAGCAGACGAGCCTTTGGTAGGATGGCTCAGACCACAGATGCAGCTGCA  
TCTCTATTAGCCCAAGGAGTTGTTGACATTGAATATAGGCGGTAAGATTTCTCTAC  
ATTGAGGTGCATTTAAGAAAGAAGCAACTTCAGGAAAATGCTGTACTTCCCCT  
TTTTATTCTCCTTAGTAGCAAGAAACCAGAAGTACACCACATAATGATAGAACTA  
ACTCATGGATCTCTTACACAATTCCTCTCTTTATACAGTGTTCATGCAGCTACCC  
CAACAAAATATCACTATCAAGATTGATGAGAACAGCAATTACCCTCATTATTG  
GCATTTGTTCTTTGGTATCAACAAGGTGACAAGGATATCACAGCAGTGCAGCTAT  
GTGAGGTATTATGCTGTTCCGAAAATAACTTTTTGTTTAGTCTTCTCAACTGGTA  
TGGACCAACTAAGGAGTTGTAAAGGGAGGATTAATTTGTTGGCCAAATCTACAAG  
AGACAGTACAATACTTATTAACGGGTTCCCTTCTGATATGTGCAGACTAAAACTT  
TGTGTGCAAGCTATTGGATAGGAGTTACGGAGCAGTATGGACAACCACTTCACCC  
CCAACCGGTCCATTGTCTTTAAGGATGCTATTGAGTGGAGAGGATGGAGATGAGT  
CATGGTTTGTTCCTATTAATAACATAACCAGAAAATTGGAAAGGTGGAGAAACATA  
TGACACAGGAGTTCAAATTAATATCTAGAAATATGCACCAGAGTCATTCATTTCT  
TTGTTTTAGCAATACCAATAAAAGAATTAAGTCCCTAGCTTTCTTAAAAGATGTGT  
TCTTTCATTATGTTTAATTGGTTAGCATATATTTTATTATTTTGTGTCAAGGAAGG  
GATGACACACTTTACTAAGAGATATTCTTATGTTTTCTTTGTAATATTGACTAATG  
AAATTTTATGTTTCTTAAATAAAGAATTAATTTCTCAAAGCAGTCATCTACAAA  
ACAGCACTAACCATATGGTCAAGGAAAAGACCTTCACTTCTACTTCTAAGCATAG  
TTCTCAAGTTTTATTGTTACTCCTGATTTTCATCAACAAACATAGTATCCCATTACTT  
AACCATACCAACAAATTTTCATAATTTTCATGACAATTC AACATACCCTGATTGTT  
CAGGCAGCAAATCCAGCAGTGGTTCTTAAAGACCTAATTACAAAAATGACCTCCT  
ATTCATTCTCATAATGCTTCTCCATTAGTTACCAATATTAATATTAGCAAAAATA  
CAGTGCCTTATCATGATGCTATTCAGAAGTGCAACAAATATCAGCTTATGTCAAC  
ATGTATATTCAGAATAGATTTCCAAACATTAACCTACAAGGAAGAACTTTAATAA  
TCATCAAAAGAGGGCAAGGGGGGAAAAAAAAAAATTGAAGGTTCAAAGTTAAGTAT  
ACTTACAAATAAAGGCAATGTAAGATGATACTGTTTGGTGATAATCTCCACTTGA  
AGCAGGTGCAGAAAGAAGGGAAGTTCTCCAAGTCTCGAATAGCATAGCA  
AAA