

IDENTIFICATION

Species: *Theobroma cacao*

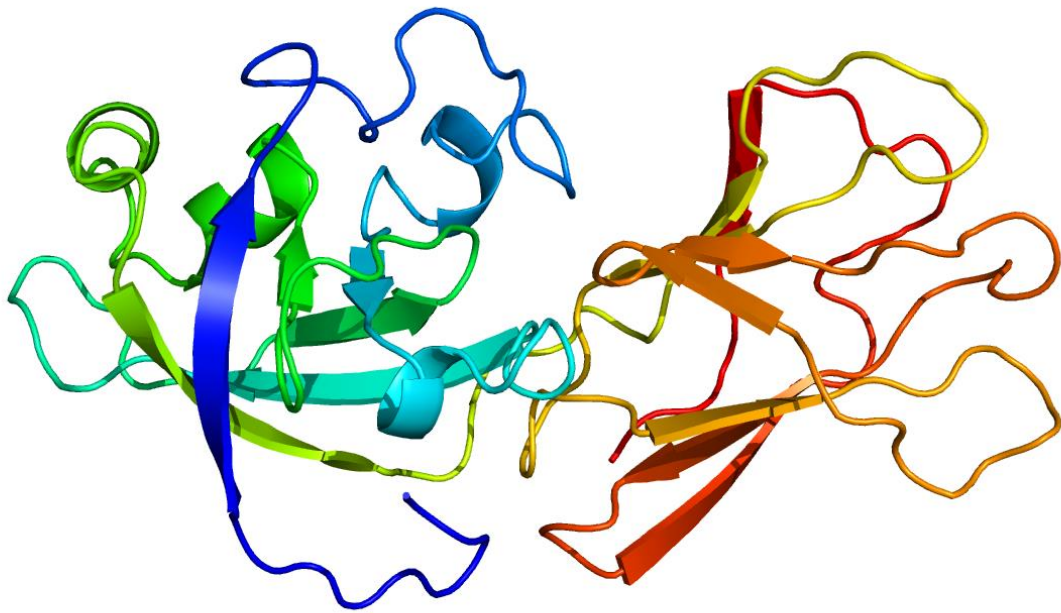
Locus: Thecc.03G241800

Gene Model: Thecc.03G241800

Description: TcEXLB-02

Family: Expansin Like Beta

3D structure:



GENOME DATABASES

Phytozome: https://phytozome-next.jgi.doe.gov/info/Tcacao_v2_1

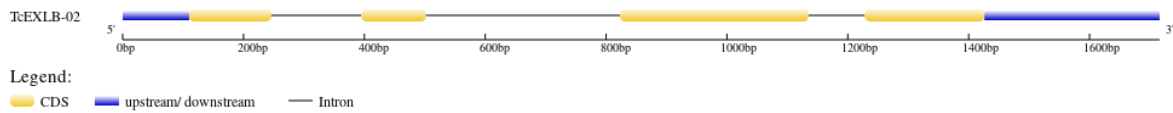
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

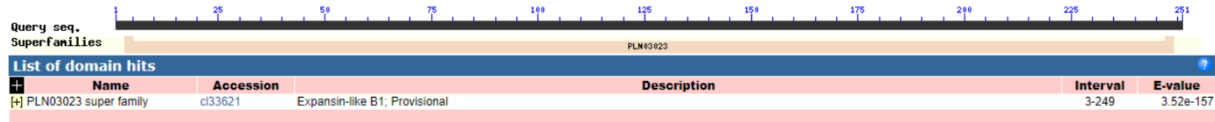
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXLB-02

MGFSLKFRYCLVSVMLLPALCYSQDYFVRSRATYYGSPDCLGTPSGACGFGEYGR
VNDANVAGVSRLYKNGTGCGACYQVRCTNPQLCDDNGVNIVVTDYGEDHTDFILS
PRAYTRMARSNTAAQLFAYGVVDVEYQRIPCGYGGKKLQFKVHEHSRYPYLAIVIL
YPAGRNEIQAI EIYREDCKQWIGMRRAYGAVYDMANPPQGSISLRFQVSGSAGLTWV
QAPNAIPSNWEAGVAYESDIQLE*

CDS (coding sequence)

>TcEXLB-02

ATGGGGTTCTCATTA AAAATTTCTGTTACTGCCTTGTGTCTGTCATGGTGCTCTTGCC
TGC ACTGTGTTACTCTCAGGACTATTTCTGTCGGTCCAGAGCAACATACTATGGC
AGTCCTGACTGCTTAGGGACTCCTAGTGGAGCTTGC GGGTTCGGAGAATACGGAA
GGACTGTCAATGATGCGAACGTGGCTGGGGTTTCCAGGCTCTACAAGAATGGAAC
TGGCTGCGGTGCTTGCTATCAGGTTAGGTGCACGAACCCTCAACTGTGTGATGAT
AATGGGGTGAACATTGTGGTGACTGACTATGGCGAAGGAGACCACACCGATTTC
ATCCTGAGCCCCCGTGCCTACACAAGGATGGCGCGATCAAACACTGCCGCGCAGC
TGTTTGCTTATGGCGTAGTTGACGTGGAATACCAGAGAATCCCCTGCGGCTACGG
TGGTAAAAGCTCCAGTTCAAGTTTCAATGAACACAGCAGATACCCAAGCTACTTG
GCCATCGTAATCTTGTACCCAGCTGGCAGAAACGAAATCCAAGCCATTGAAATTT
ATCGGGAGGATTGTAAGCAATGGATAGGAATGAGGAGGGCGTATGGGGCAGTTT
ACGACATGGCTAATCCGCCACAGGGTCTATCAGCTTGAGGTTCCAAGTGAGCGG
CAGCGCCGGTTTGACATGGGTGCAGGCACCAACGCTATCCCCAGCAATTGGGA
AGCTGGAGTTGCATATGAATCAGATATTCAGCTTGAGTAA

Nucleotide

>TcEXLB-02

ATCTCTCACAAGCTCATCTAACTTCTCTCCTCTCTATCTTCTAAGTTATCTTTCTCT
TTTCTTGGTTACGTTATTCTCTTTGGTACTGCTATAAACTTTGTAGCAGAAAGATG
GGTTCTCATTA AAAATTTCTGTTACTGCCTTGTGTCTGTCATGGTGCTCTTGCCCTGC
ACTGTGTTACTCTCAGGACTATTTCTGTCGGTCCAGAGCAACATACTATGGCAGT
CTGACTGCTTAGGGACTCCTAGTATGTACATTTCTCCTATGATACATGCATGTTT
ATTTATTTATTACAATTAAGTTCACAAGTGGATATTGACCGTGATTTTAAACCTT

CAAATTCACTTAAGGATGTTTCATCATGTCATACATGAGCTAACAGTTATGTGAAT
TTAGGTGGAGCTTGC GG GTTCGGAGAATACGGAAGGACTGTCAATGATGCGAAC
GTGGCTGGGGTTTCCAGGCTCTACAAGAATGGAAGTGGCTGCGGTGCTTGCTATC
AGGTAAAACCTCAGGCCCCCTAATCATTTCAGCTCTGAAAGCGACAGGTAAAGCA
AGGACGAAGCTACCAAAGTGATGAGCTCGGCAGAAATTTCCGCAAAAATAAACT
AATTTCCGCAATAACCTTTAGAATAGTATATATTTCTCACTTGCTAAAAGCATTTC
TAGTGTTCAACATGTGACCCTACTGCTTAAAATAATTAAGTTTAATTATTATTA
ATTGCTAATTTGGCACTAATCTATAAACCAAGTGGCTTCTTATTTTATGTTGCT
AGCTAATCCATGACAATTCCTTTTTTGGCTTAATTAACAAACTCAGGTTAGGTGCA
CGAACCTCAACTGTGTGATGATAATGGGGTGAACATTGTGGTGACTGACTATGG
CGAAGGAGACCACACCGATTTTCATCCTGAGCCCCCGTGCCTACACAAGGATGGCG
CGATCAAACACTGCCGCGCAGCTGTTTGGCTTATGGCGTAGTTGACGTGGAATACC
AGAGAATCCCCTGCGGCTACGGTGGTAAAAGCTCCAGTTCAAGGTTTCATGAACA
CAGCAGATACCCAAGCTACTTGGCCATCGTAATCTTGTACCCAGCTGGCAGAAAC
GAAATCCAAGCCATTGAAATTTATCGGGTAAATTTTACTTATTTATCACAACTTC
TGACCCCTTGAAGTACGGAAATCGTACATATTGTTCAATAATTATGCATTTGGTGT
GTGAGCAGGAGGATTGTAAGCAATGGATAGGAATGAGGAGGGCGTATGGGGCAG
TTTACGACATGGCTAATCCGCCACAGGGGTCTATCAGCTTGAGGTTCCAAGTGAG
CGGCAGCGCCGGTTTGACATGGGTGCAGGCACCAAACGCTATCCCCAGCAATTGG
GAAGCTGGAGTTGCATATGAATCAGATATTCAGCTTGAGTAAAGTAGCAAATTTT
AGAAGGAAAAAAATCAGACTACTTGTATTAATTCCTTAATTATGTAGTTGCTTTAT
GATGTCACTAGTAAGACAGCAAAGTGCTCGGAACCTTTGCCCTGTGCTCGATAAAA
TAAAAGTTGTTATATGTACTTTTGGCTTCATGAGTCAATAAAGTTACACCTTCAGA
GTATTTTTCCTTCCACTGAGTTTGTATCATGATGACGACGTCTTATAGTAATTGA
TAGTGTTTTGTTTTCACTACTTTTCGTAATCATTAATAAAAAATACTACATTC