

IDENTIFICATION

Species: *Theobroma cacao*

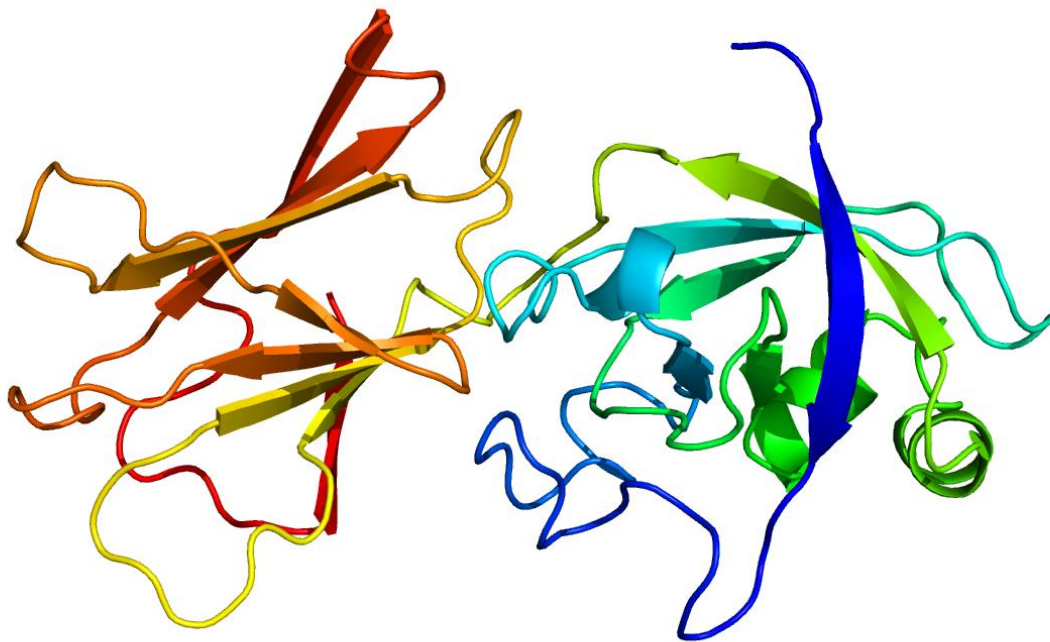
Locus: Thecc.03G242000

Gene Model: Thecc.03G242000.1.p

Description: TcEXLB-03

Family: Expansin Like Beta

3D structure:



GENOME DATABASES

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

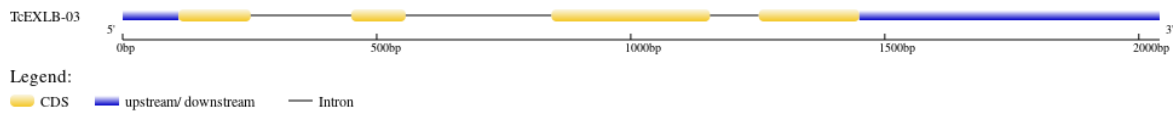
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

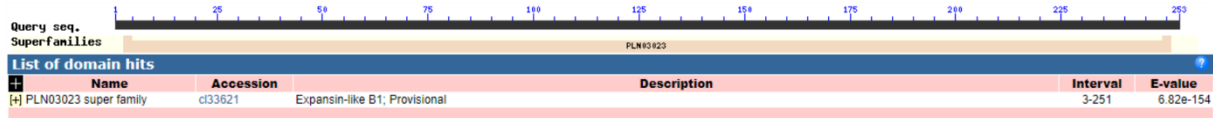
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXLB-03

MGFSVRFRCCLVSVMLLPALCLCYSDYFVRSRATYYGSPDCLGTPSGACGFGEYGR
RTVNDANVAGVSRLYKNGAGCGACYQVRCTNPQLCDGNGVNVVVTDYGEGDHTD
FILSPRAYTRMARSNTAAQLFAYGVVDVEFQRIPCRYGGHKLQFKVQEQRYPYLA
VILYQTGKNDIQAVDIWREDCKQWIGMRRAYGAVFDMASPPQGA VSLRLQVSGSAG
LTWVQAPNVIPSDWKAGVAYESDIQLE*

CDS (coding sequence)

>TcEXLB-03

ATGGGGTTCTCAGTAAGATTTTCGTTGTTGCCTTGTGTCTGTCATGGTGCTCTTGCC
TGC ACTGTGTCTGTGTTACTCTCAGGACTATTTTCGTCCGGTCCAGAGCAACTTACT
ATGGCAGTCCTGACTGCTTAGGGACTCCAAGTGGAGCTTGCGGGTTTGGAGAATA
CGGCAGGACTGTCAATGATGCGAACGTGGCTGGGGTTTCCAGGCTCTACAAGAAT
GGAGCTGGCTGCGGTGCTTGCTATCAGGTTAGGTGCACGAACCCTCAACTGTGTG
ATGGTAATGGGGTGAACGTTGTGGTGACTGACTATGGCGAAGGAGACCACACAG
ATTCATCCTGAGTCCCCGTGCCTACACAAGGATGGCACGATCAAACACCGCCGC
GCAGCTGTTTGCTTATGGCGTAGTTGACGTGGAATTCCAGAGAATCCCCTGCCGC
TACGGTGGTCACAAGCTCCAGTTCAAGGTTCAAGAACAAAGCAGATACCCAAGC
TACTTGCCATTGTAATCTTGTACCAAACACTGGCAAAAACGATATCCAAGCCGTTG
ATATTTGGCGGGAGGATTGTAAGCAATGGATAGGAATGAGGAGGGCGTATGGGG
CAGTTTTTCGACATGGCTAGTCCGCCACAGGGGGCTGTCAGCTTGAGGTTGCAAGT
GAGTGGCAGCGCGGGCTTGACATGGGTGCAGGCACCAAACGTTATCCCAGCGA
TTGAAAGCTGGAGTTGCGTATGAATCAGATATTCAGCTTGAGTAA

Nucleotide

>TcEXLB-03

ATCTCTCACAAAGCTCATCCAACCTTCTCCCCTCTCCATTTGCTAGGTTATCTTTCTCT
TTTCTTGGTTACATTATTCTCTTTGGTACTGCTAAAAGCTTGGTAGCAGAAACATG
GGGTTCTCAGTAAGATTTTCGTTGTTGCCTTGTGTCTGTCATGGTGCTCTTGCCCTGC
ACTGTGTCTGTGTTACTCTCAGGACTATTTTCGTCCGGTCCAGAGCAACTTACTATG
GCAGTCCTGACTGCTTAGGGACTCCAAGTATGTACATTTCTCCTATGATACGTGCA
TGTTTATTTATTTAGTAGTATTTAGTAGTAATATTAAGTGTGCATCGTCACGATTTTT

AAAATTAGCAAGTGGATATACTTATTCTTGAGGTCCTTGACTGTGATTTTAAATCT
TACTGCCTATGCAATGCTCATATGTCATACATGAGCTAACAATCATGTGAATTTA
GGTGGAGCTTGCGGGTTTGGAGAATACGGCAGGACTGTCAATGATGCGAACGTG
GCTGGGGTTTCCAGGCTCTACAAGAATGGAGCTGGCTGCGGTGCTTGCTATCAGG
TAAACTTCAAGCCCCTTAATTATTTCAACTCTGAAAGCGATCACAGGTAAAGCA
AGGAGTTATGGGTTCGGCAGAACTTCCGCAAAATCAAACACTTTTCCACTCAAC
CTTTAGAATAGTATAGTGTTTCTAGCGTTAAACATGTGCCCTACCGCTTAAACT
AATTAAGTTTTCATTATCATTAGTTGCTTTGGCACTGATCTATAAACACAAGTT
GCTTCTTATTTTATGTTGCTAGCTAATCCATGACGCTTGCTTTTGCTTAATTAACA
AACTCAGGTTAGGTGCACGAACCCTCAACTGTGTGATGGTAATGGGGTGAACGTT
GTGGTGACTGACTATGGCGAAGGAGACCACACAGATTTTCATCCTGAGTCCCCGTG
CCTACACAAGGATGGCACGATCAAACACCGCCGCGCAGCTGTTTGCTTATGGCGT
AGTTGACGTGGAATTCCAGAGAATCCCCTGCCGCTACGGTGGTCACAAGCTCCAG
TTCAAGGTTCAAGAACAAGCAGATACCCAAGCTACTTGGCCATTGTAATCTTGT
ACCAAACCTGGCAAAAACGATATCCAAGCCGTTGATATTTGGCGGGTAAATTTTA
CTTATTTATCACAACTTTTTACTCCTTCTTGAAGTACCGAATCGTACATATTGTT
ACATTATTATGCGTTTGGTGTGTGAGCAGGAGGATTGTAAGCAATGGATAGGAAT
GAGGAGGGCGTATGGGGCAGTTTTTCGACATGGCTAGTCCGCCACAGGGGGCTGT
CAGCTTGAGGTTGCAAGTGAGTGGCAGCGCGGGCTTGACATGGGTGCAGGCACC
AAACGTTATCCCCAGCGATTGGAAAGCTGGAGTTGCGTATGAATCAGATATTCAG
CTTGAGTAAAGTGCGTCAGGCTTGCACTACTAGAAACAAATAAGTAAGAAATTTT
AAAACGCAAAAATCAGACTACTTGTATTAAGTCCTTAATTACATGGTTGCTTTAT
GATGTCACTAGGAAGACAGCAAAGTGCTTAAACTTCGCCTTGCGCTCGATAAAA
TAAAGAGTTGGTTATGTGTACCTTTGCTTTTTTTTTTAAACATTAATCATCATAGACT
TTCATTCAAACCTGGGATGAACCTAGTAAAAAAGGAACAGAACGGTCAAGAGTT
AAAAGAGAACCAAGTTACACCATAAAGCAACAAAGAATTGCACCCAGCCCCAA
TCAATTTTCATCCTAACCAAGAAAGAACTAGCAAAGGCATCAGTAAGATCAGG
AAACAGGACCAAAGACACTGTTCCCTATTCTTGCTGTTACAAAATCCATTCATA
CCAGTCTATCCATCACACTATTACCATCAGACCTGATACAGAACATCAAAGGTA
CCTGAAGCATGAGATTGTCGCATATATCAGCAATCTTATCGAAGGACTTGGAAGC
CAATGCTTCATTCAGAACACTAAGATCCATCTCTTCTTATCTTCTGCTTTT