

IDENTIFICATION

Species: *Theobroma cacao*

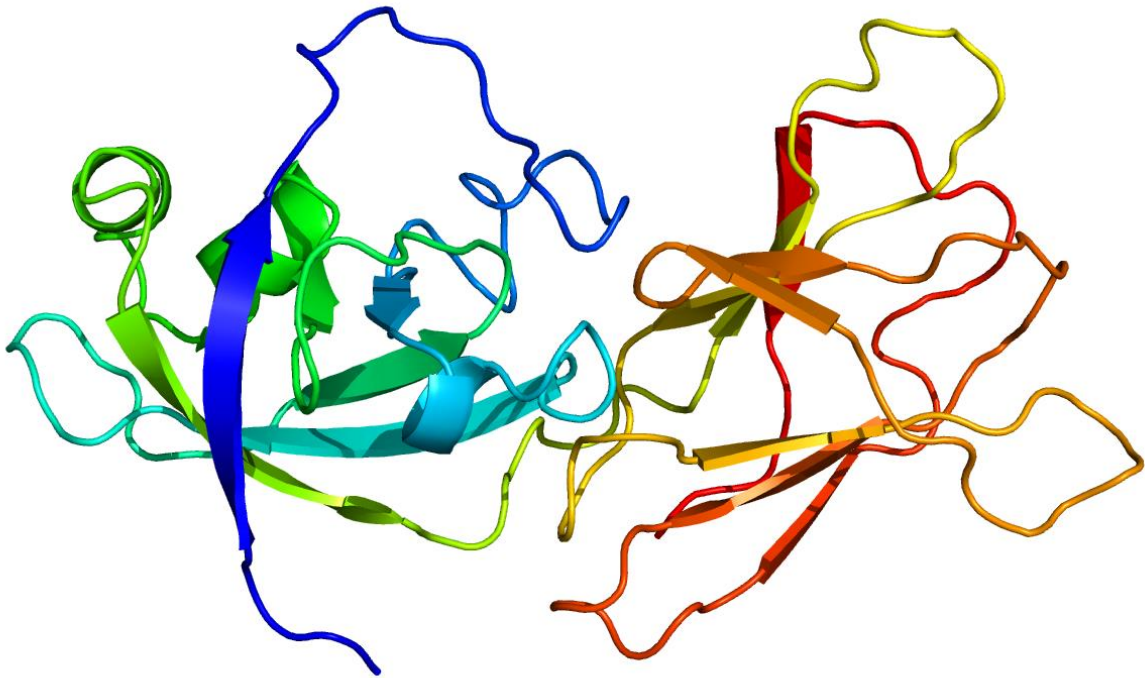
Locus: Thecc.02G049900

Gene Model: Thecc.02G049900.1.p

Description: TcEXLA-01

Family: Expansin Like Alpha

3D structure:



GENOME DATABASES

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

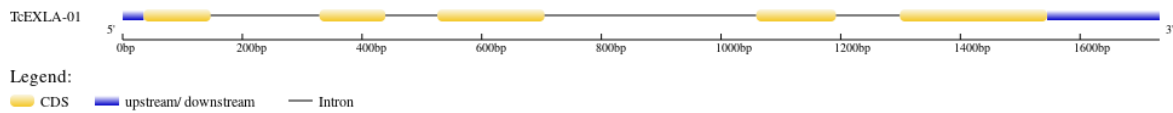
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

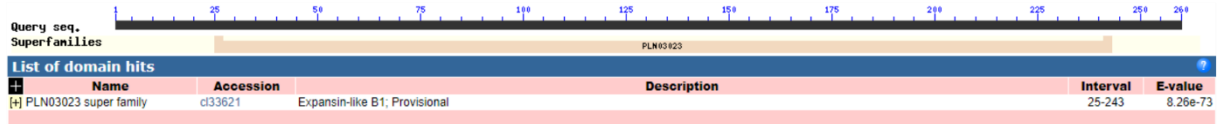
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXLA-01

MALFLCLLFVLVSSANACDRCVHQSKAAYFSKASALSSGACGYGSLALGLSGGHLA
AGVSSLYKDGAGCGACFQIRCINSTLSSKGTRVTLTDLNHNNTDFVLSSRAFMAM
ANQGMGQDILKLGIVDVEYKRIPCEYKNQNLAVRVEESSQKPNYLAIKLLFQGGQTEI
VGMDVAQVGSSNWNFMSRNYGAVWDTSRVPTGALQFRFVVTSGFDGKWIWAKSV
LPADWKTGVIYDSGVQITDIAKEGCSPCDDGSWR*

CDS (coding sequence)

>TcEXLA-01

ATGGCTCTCTTTCTTTGCTTGTTGTTGTCGTCCTCGTTTCTTCTGCTAATGCTTGTGAT
CGCTGCGTGCACCAGTCCAAGGCTGCTTATTTCTCCAAGCCTCTGCACTTTCATC
TGGAGCTTGTGGATATGGTTCCTTGGCATTAGGCTTAAGCGGTGGACACCTTGCA
GCTGGTGTGTCTTCTCTGTACAAAGATGGAGCTGGTTGTGGTGCCTGCTTTCAGAT
AAGGTGCATAAATTCAACTTTGTGCAGCAGTAAAGGGACTAGGGTGACACTAACT
GATCTCAATCACAAACAAATCAAACAGACTTTGTTCTTAGCAGCAGAGCTTTCATGG
CCATGGCTAACCAGGGCATGGGCCAAGACATTTTGAACTTGGGATTGTCGACGT
TGAATATAAAAGGATACCTTGTGAATACAAAACCAGAACTTAGCTGTTAGAGTT
GAAGAATCAAGCCAAAAACCAAATTAATTGCGATTAAATTGCTGTTCCAGGGTG
GTCAGACAGAGATAGTAGGCATGGATGTGGCTCAGGTTGGTTCTTCAAACCTGGAA
TTTCATGAGCAGGAATTATGGAGCAGTGTGGGACACAAGCAGAGTTCCAACCTGGT
GCCCTGCAATTTAGATTTGTGGTGACATCTGGGTTTGATGGCAAGTGGATTTGGG
CAAAGAGCGTACTTCCAGCTGACTGGAAAACCTGGAGTTATATATGATTCTGGTGT
TCAAATCACCGATATTGCAAAAAGAAGGTTGCTCTCCTTGTGATGATGGGAGTTGG
AGATGA

Nucleotide

>TcEXLA-01

AGTCACCTACATTTCTCAACTCAACTCAACGCGAAATGGCTCTCTTTCTTTGCTTG
TTGTTTCGTCCTCGTTTCTTCTGCTAATGCTTGTGATCGCTGCGTGCACCAGTCCA
GGCTGCTTATTTCTCCAAGCCTCTGCACTTTCATGTAAGCAGCCGCAAAGTTCCA
TCTTATTTTCATTTTCTGGTTTCTTTTTTTTTTCTATCGATTTGATTGATACCCA
GAGTAGTTTTCTTCTTGA AAAACGAAATAGGCAAAGACCCATATGATATTTCTG

TGAGAAATTTGGTGGGTAACATTTGTTTTGTGGCTTTGATATTTGCAGCTGGAGCT
TGTGGATATGGTTCCTTGGCATTAGGCTTAAGCGGTGGACACCTTGCAGCTGGTG
TGTCTTCTCTGTACAAAGATGGAGCTGGTTGTGGTGCCTGCTTTCAGGTCTGTAAT
TTTACAAAACAATTAATTCTGTTATCTGTACCATCCATAGTGGAAACTGAGTTTG
ATTATTAACATATTTTTTGGCAGATAAAGGTGCATAAATTCAACTTTGTGCAGCAGT
AAAGGGACTAGGGTGACACTAACTGATCTCAATCACAACAATCAAACAGACTTT
GTTCTTAGCAGCAGAGCTTTCATGGCCATGGCTAACCAGGGCATGGGCCAAGACA
TTTTGAAACTTGGGATTGTTCGACGTTGAATATAAAAGGTGGGCGTGATCAAACAT
TTAATCAAATTACTATAACCAAACAAAATAGTAATACTAGTGATGAAAATGAATGA
TTATGATATATTTTTTTGAAGTTTTCTCATCCGATGGTTTTTGTTACAACCACCACAT
AGAAAAGTCAGTCCCGGCGTGGACCAACTTATAATATTATGTGGCACCCCGGCA
GGCACATTAATTGCTATAACGGTCCAATATCTTCATTATCATATTCCTATAATTTA
CAATGCTGGGGTATTATCGTGATTTCGTGATTGGTTCAATAAATATTGTTTTAACTT
TGGTTAGCATTATATCATCAATACTACAATGCTGATCAACTTATTATGTGAAACA
GGATACCTTGTGAATACAAAACCAGAACTTAGCTGTTAGAGTTGAAGAATCAA
GCCAAAACCAAATTACTTGGCGATTAAATTGCTGTTCCAGGGTGGTCAGACAGA
GATAGTAGGCATGGATGTGGCTCAGGTCAGCTCAGATAAAGTACTCAAAGAGTG
CTAAATTTACATTATAGCTAAAACAATTTGTTTTGTATTTCATTACACTGATGTTCTG
ATATCGCTATGTATCATTCCAGGTTGGTTCTTCAAACCTGGAATTCATGAGCAGGA
ATTATGGAGCAGTGTGGGACACAAGCAGAGTTCCAACCTGGTGCCCTGCAATTTAG
ATTTGTGGTGACATCTGGGTTTGATGGCAAGTGGATTTGGGCAAAGAGCGTACTT
CCAGCTGACTGGAAAACCTGGAGTTATATATGATTCTGGTGTTCAAATCACCGATA
TTGCAAAAAGAAGGTTGCTCTCCTTGTGATGATGGGAGTTGGAGATGAGAGGGTCT
TGCTTTGCCAATCTTGTGTGAATACACATATTTATAGTGAATTTGAATAATTTAGA
AGATAAAAAAAAAAATGAATATATAGTATTTTTTTTCCTCGTTGTTATTTTCTAGGTT
TGCTCTTGAAGATGTCTAACATATTTTTTTTCCTACTCTACCATAATAAAATGCTC
CTTTGTTTCTTT