

IDENTIFICATION

Species: *Theobroma cacao*

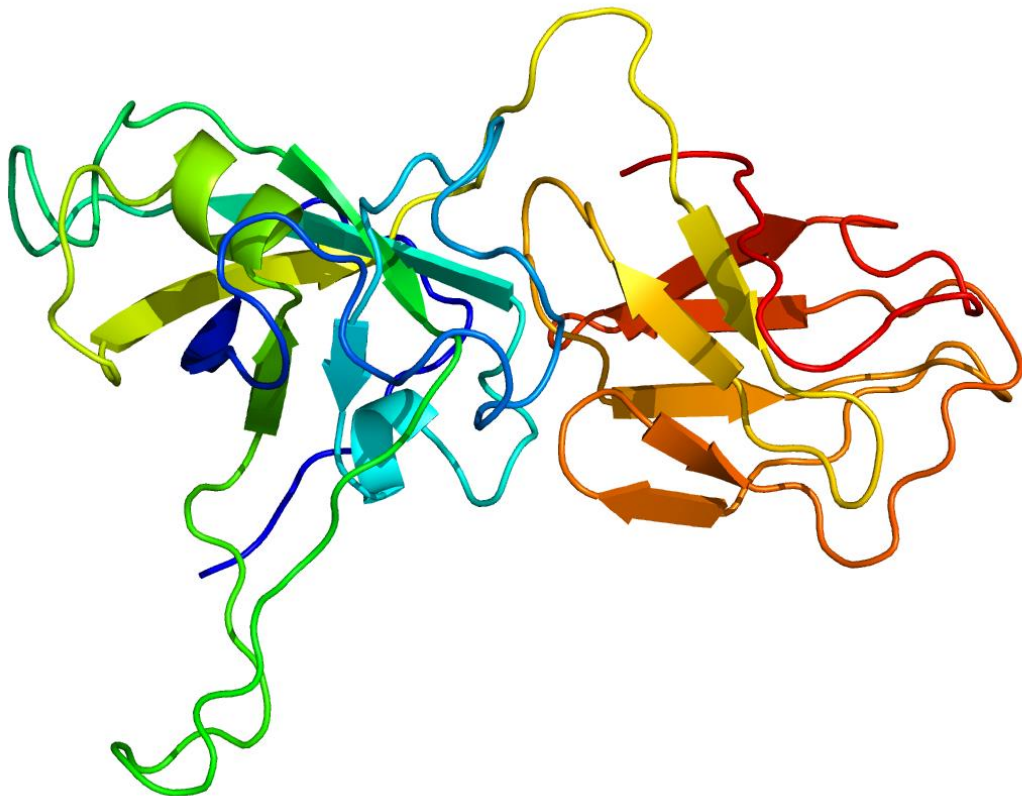
Locus: Thecc.10G036000

Gene Model: Thecc.10G036000.1.p

Description: TcEXPA-18

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

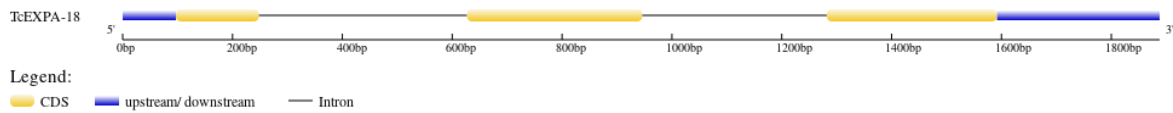
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

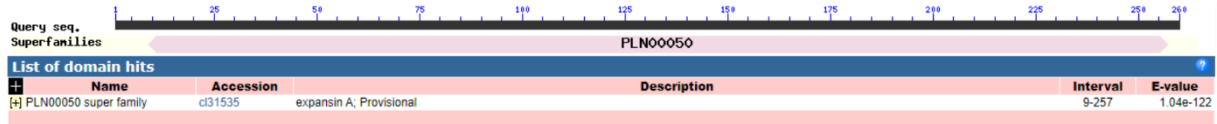
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-18

MAIGSVVACLVALNLILVAVEARIPGVYSGAEWQTAHATFYGGSDASGTMGGACGY
GNLYSQGYGVNTAALSTALFNGLSCGACFEIKCANEPQWCHAGSPSIFVTATNFCPP
NYALPSDNGGWCNPPRPHFDLAMPMFLKIAEYRAGIVPVS YRRVPCRKQGGIRFTIN
GFRYFNLVLITNVAGAGDIVKASVKGSKTGWMSLSRNWQNWQSNVAVLVGQSLSF
RVTGSDRRSTSTSWNIVPSNWQFGQFTFTGKNFRV*

CDS (coding sequence)

>TcEXPA-18

ATGGCAATTGGGAGTGTTGTTGCTTGTGGTTGCTTTGAATCTCATTGTTGGTAGC
TGTTGAAGCTAGAATTCCTGGTGTACTCAGGGGCTGAATGGCAGACTGCCCAT
GCTACTTTCTATGGTGGCAGTGATGCTTCTGGGACAATGGGAGGAGCTTGTGGCT
ATGGGAACCTATACAGCCAAGGCTACGGAGTTAACACGGCAGCTTTGAGCACTGC
CCTTTTCAACAATGGCCTCAGCTGCGGTGCCTGTTTCGAGATCAAATGCGCAAAT
GAACCACAATGGTGCCATGCTGGTAGCCCTTCAATTTTCGTGACTGCAACCAACT
TTTGCCCTCCAAATTATGCTCTTCTAGTGACAATGGTGGCTGGTGCACCCTCCT
CGCCCTCACTTTGATCTCGCCATGCCCATGTTCCCTAAGATTGCTGAGTACCGTGC
TGGCATTGTCCCTGTCTCTTATCGCAGGGTGCCATGCCGGAACAAGGAGGAATC
AGGTTACAATCAATGGTTTCCGTTACTTCAACTTGGTCCTAATCACCACGTGGC
AGGCGCTGGTGACATCGTGAAGGCAAGCGTGAAGGGTTCGAAAACAGGGTGGAT
GAGCTTGAGCCGAAACTGGGGCCAGAACTGGCAATCCAACGCCGTTCTGGTCGG
CCAATCTCTCTTTTCAGGGTCACCGGCAGCGACCGCAGAACCTCTACCTCATGG
AACATTGTCCCTTCCAATTGGCAGTTTGGCCAAACTTTCACAGGCCAAAATTTCA
GGGTCTAA

Nucleotide

>TcEXPA-18

ACACTCTACCTTCTCTTCAAAACCCTTCAAACCCACACCTTTAGCTTCTCCCTC
CCTCCACTTTCTCTGTCTGTTTAGGACAGTCAAAAAGAAAAATGGCAATTGGGAG
TGTTGTTGCTTGTGGTTGCTTTGAATCTCATTGTTGGTAGCTGTTGAAGCTAGAA
TTCCTGGTGTACTCAGGGGCTGAATGGCAGACTGCCCATGCTACTTTCTATGGT
GGCAGTGATGCTTCTGGGACAATGGGTAATCTTCTACTGCATTCAAACCTTT

AGAAAATGAAAAAATAAAAAAAAATTCTTTCTTTCACCACCTTGAGTTGTCAATT
TGAAGACTTAACCATGTTCAATCTTTTTGTCCCTATCATTGTTTGTAGACTTCAA
CTCTGCCTCCCCTGCTGCTTTTCCTGGATATCATAGCTAAAACAGACGTGTGATGG
CCAGTTTTTTTCATTGGAAAAGGCCTTTTAGTCAATGTTAATAAGCACTCGTGAAC
GTTTACTTTCTGGAAATCCAAATGACTTCCAAGAGAATGCTTTTACAACCAATTCT
AATTCTCAAACCAGGAAGATTTTCGACTTTTTATCTGACATTTTCTTTTCCACTTC
GAATAATCTTCAGGAGGAGCTTGTGGCTATGGGAACTTATACAGCCAAGGCTACG
GAGTTAACACGGCAGCTTTGAGCACTGCCCTTTTCAACAATGGCCTCAGCTGCGG
TGCCTGTTTCGAGATCAAATGCGCAAATGAACCACAATGGTGCCATGCTGGTAGC
CCTTCAATTTTCGTGACTGCAACCAACTTTTGCCCTCCAAATTATGCTCTTCCTAG
TGACAATGGTGGCTGGTGCAACCCTCCTCGCCCTCACTTTGATCTCGCCATGCCCA
TGTTCCTTAAGATTGCTGAGTACCGTGCTGGCATTGTCCCTGTCTCTTATCGCAGG
TAAAAAACAAAACCTTCCGTAGTTCCTTCGTCAACCAGTTTAGCATGATTGAGCT
CCTCATTTTAACTTTTCAGGGCAATTTTTATCATTTTTCACTTGATATATTCATAAA
AATTTGACATATTTTCGTGTTAGTTGAGGCAATTTTCCTCTTTAATTCAAGACCAT
CCATTGCTTTTTGGCCCTACATTCTACTCTTCAGAATATGAAACCAGTTTGAGGTT
TAACCCACACTTTTTCCC GCGTAACTCACTATTCCTGAGAAACGAACAGAACCCA
AGAAACCCAATCTTTTACTTTTCTTCAAGCTGATTGAAAAATATTTCTTCGCCAG
GGTGCCATGCCGGAAACAAGGAGGAATCAGGTTCAACAATCAATGGTTTCCGTTAC
TTCAACTTGGTCCTAATCACCAACGTGGCAGGCGCTGGTGACATCGTGAAGGCAA
GCGTGAAGGGTTCGAAAACAGGGTGGATGAGCTTGAGCCGAAACTGGGGCCAGA
ACTGGCAATCCAACGCCGTTCTGGTCGGCCAATCTCTCTCTTTCAGGGTCACCGGC
AGCGACCGCAGAACCTCTACCTCATGGAACATTGTCCCTTCCAATTGGCAGTTTG
GCCAAACTTTCACAGGCAAAAATTTCAAGGTCTAATCTAATCTTAATGAAGGGTT
TCTTTCCTTTTTTGGGGGTTTTAATATCTTTTAGTTTTATTTTCTGGAAAATGAGG
GAAAGTGTGTGGTGAAGTTGCTTTGACTTTTGAGGCTGGGGGAAGTGGCTTTCAT
TTTATAGTTTTTACTGTTTTTCTATTTTTGGGTAAGAGGGTAAAGGTAAAGAAGTG
AGATGTAATGGGGGCTGAAGTGGCTTGAAAGAAGTTGTAGCCCGCAGCTTTGAT
ACTCTTTTTTCTAATATATATATATAATAATTAATGAAAGAAGTGTTTTAT