

IDENTIFICATION

Species: *Theobroma cacao*

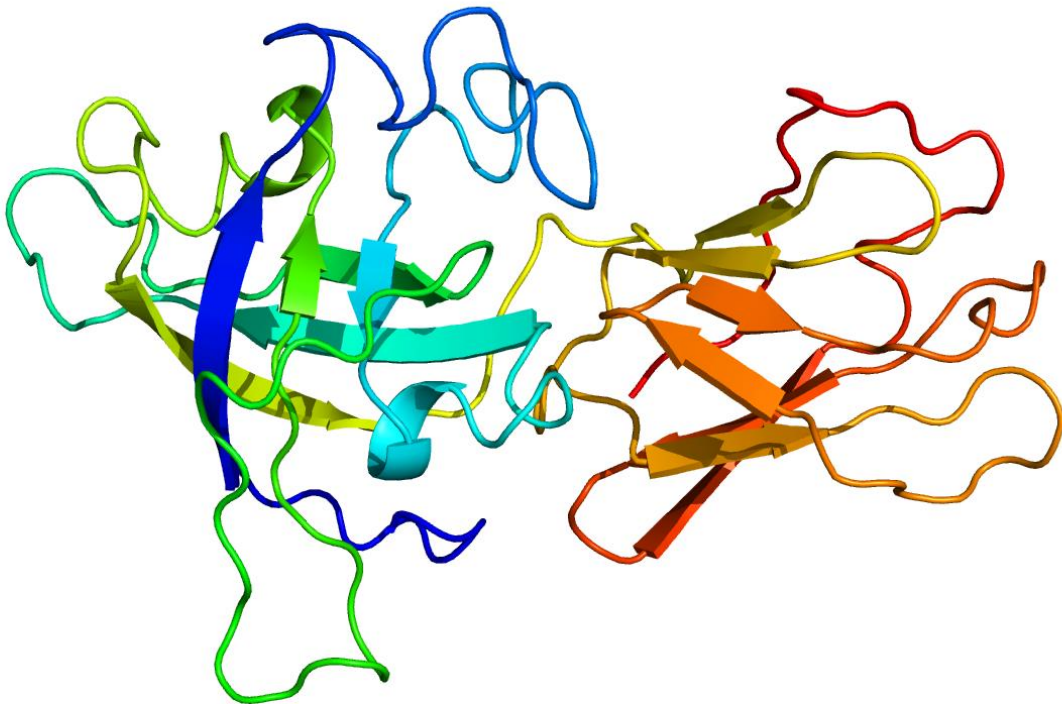
Locus: Thecc.04G093100

Gene Model: Thecc.04G093100.1.p

Description: TcEXPA-09

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

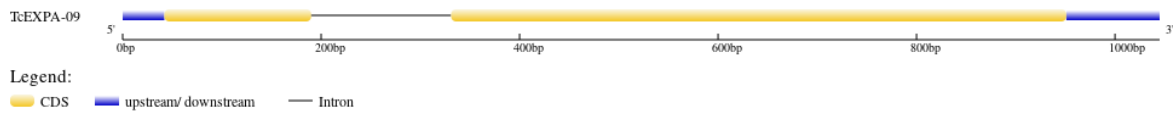
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

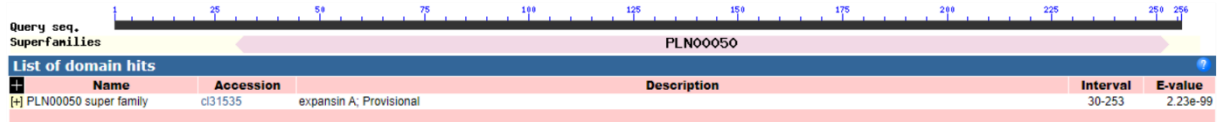
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-09

MAVLKFFFSFLLCMSCLAAAMGKWQRHIDTNWHDAAHATFYGDMSGGETLQGACG
YGDLFEQYGLETTALSTALFNNGLTGACFEIKCVNNPQWCYPNAGTIKVTATNFC
PPNYSKPDGNWCNPLKHFDSLMMKMF TKLAQYKAGIIPVQYRRVLCSKKGGVQFQV
KGNPYWTLVLLYNVGGAGDVKNVVKVGSSTGWIQMSRNWGVNWQTGTKLTGQSL
SFQVTTS DGKMLEFDNVAPANWLF DQIFDGKKNF*

CDS (coding sequence)

>TcEXPA-09

ATGGCTGTTCTTAAGTTTTTCTTCTCATTCTCTTGTGCATGTCCTGCTTGGCTGCA
GCTATGGGCAAGTGGCAAAGACATATTGATACCAATTGGCACGACGCCCATGCA
ACATTTTATGGTGATATGAGCGGTGGAGAAACCCTGCAGGGAGCTTGCGGATATG
GTGATCTCTTCGAGCAAGGGTACGGACTCGAGACAACAGCCCTTAGCACAGCTCT
ATTCAACAACGGACTCACCTGCGGTGCGTGT TTTGAAATCAAGTGTGTCAACAAC
CCTCAATGGTGTTATCCCAATGCTGGCACCATCAAAGTGACCGCAACCAATTTT
GCCCTCCTAATTACTCAAAGCCTGATGGCAACTGGTGCAATCCCCACTAAAACA
CTTTGATCTCTCAATGAAAATGTTCAACCAACTAGCGCAATACAAAGCTGGAATC
ATCCCCGTCCAGTACCGTCGCGTCTTGTGTTCCAAGAAAGGTGGGGTTCAGTTCC
AGGTCAAGGGAAACCCTTACTGGACTTTGGTGTGTTGTACAATGTTGGAGGTGC
CGGTGATGTTAAGAATGTTAAAGTCAAGGGCTCTAGCACTGGCTGGATACAGATG
AGCCGTA ACTGGGGAGTGAATTGGCAGACCGGAACAAAGTTAACAGGGCAAAGC
CTGTCATTCCAGGTGACTACCAGTGATGGCAAATGTTGGAGTTTGATAACGTGG
CACCGGCAATTGGCTGTTTGATCAAATTTTGGATGGGAAAAAGAACTTCTAG

Nucleotide

>TcEXPA-09

CAGACTTAGATTTAGTTTAATCTCTCTATATCGTTTATAGCAATGGCTGTTCTTAA
GTTTTTCTTCTCATTCTCTTGTGCATGTCCTGCTTGGCTGCAGCTATGGGCAAGTG
GCAAAGACATATTGATACCAATTGGCACGACGCCCATGCAACATTTTATGGTGAT
ATGAGCGGTGGAGAAACCCTGCGTAAGTATACCGTTTACATGTATGGATTTTGG
TGCTATTATTTCTGCGATGATATATCAAATCATCCACTTTTGGATGATTTGATCAT
AACACAATCGTTAAATAGTTAAGAGGTATTATACGTGATTATATATTGCAGAGGG

AGCTTGCGGATATGGTGATCTCTTCGAGCAAGGGTACGGACTCGAGACAACAGCC
CTTAGCACAGCTCTATTCAACAACGGACTCACCTGCGGTGCGTGTTTTGAAATCA
AGTGTGTCAACAACCCTCAATGGTGTTATCCCAATGCTGGCACCATCAAAGTGAC
CGCAACCAATTTTTGCCCTCCTAATTACTCAAAGCCTGATGGCAACTGGTGCAAT
CCCCACTAAAACACTTTGATCTCTCAATGAAAATGTTACACAAACTAGCGCAAT
ACAAAGCTGGAATCATCCCCGTCCAGTACCGTCGCGTCTTGTGTTCCAAGAAAGG
TGGGGTTCAGTTCCAGGTCAAGGGAAACCCTTACTGGACTTTGGTGTTGTTGTAC
AATGTTGGAGGTGCCGGTGATGTTAAGAATGTTAAAGTCAAGGGCTCTAGCACTG
GCTGGATACAGATGAGCCGTA ACTGGGGAGTGAATTGGCAGACCGGAACAAAGT
TAACAGGGCAAAGCCTGTCATTCCAGGTGACTACCAGTGATGGCAAATGTTGGA
GTTTGATAACGTGGCACCGGCCAATTGGCTGTTTGATCAAATTTTTGATGGGAAA
AAGA ACTTCTAGGAAAAGTTTCTTGAGCTTCTTTAGAATGTAGTGCTGCCTTTCTA
CATGTCGTTTTGGTGCAAGTTTGTGATGCAAATGGGTCTGATTGGGCTGT