IDENTIFICATION

Species: *Theobroma cacao*

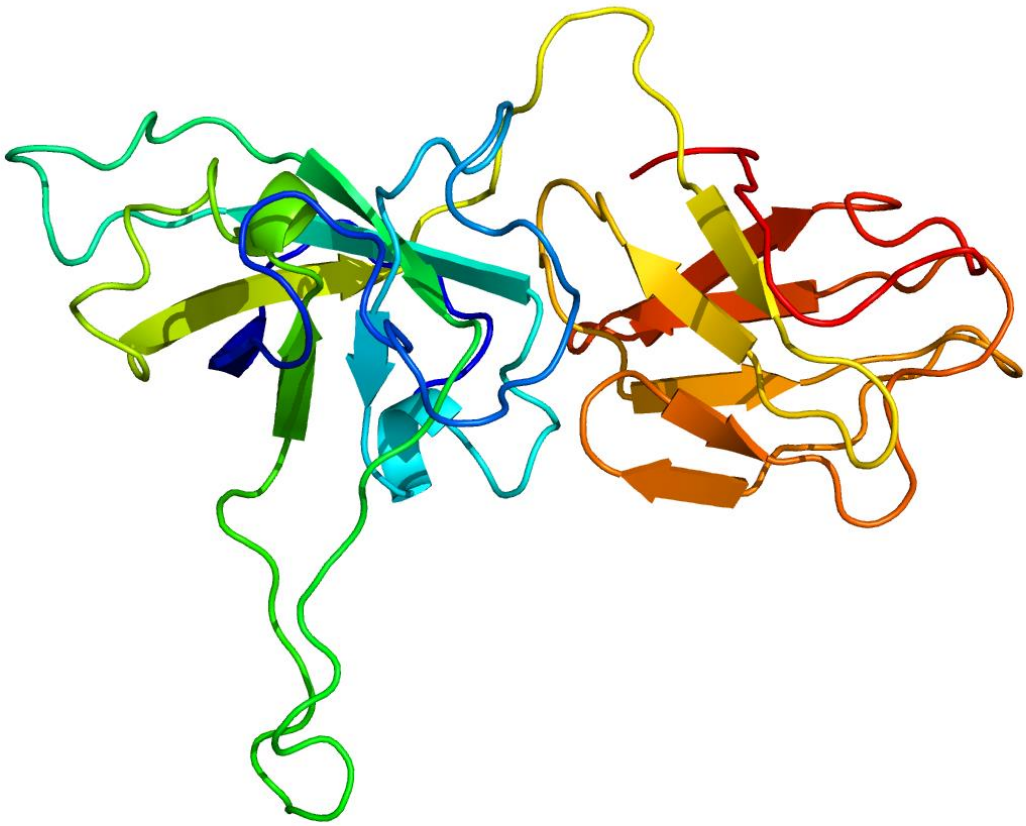
Locus: Thecc.04G152900

Gene Model: Thecc.04G152900.1.p

Description: TcEXPA-10

Family: Alpha Expansin

3D structure:



GENOME DATABASES

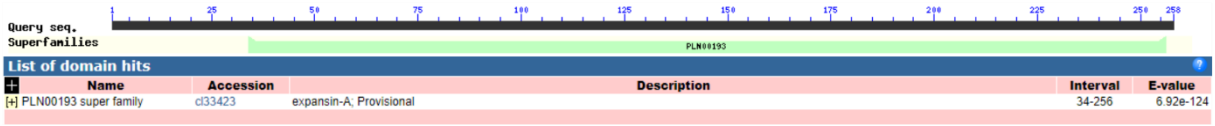
Phytozome: https://phytozome-next.jgi.doe.gov/info/Tcacao_v2_1
KEGG: https://www.genome.jp/entry/gn:T02994

EXTERNAL RESOURCES

https://www.cacaogenomedb.org/
https://cocoa-genome-hub.southgreen.fr/node/4

GENE STRUCTURE

DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-10

MATTSMQITLSFLFLFCLLGICSCFHGSNRDDGGWQTAHATFYGGADATGTMGGAC
GYGNLYSQGYGTTTAALSTALFNNGLTCGACYELRCNDDPKWCISRTITVTATNFCP
PNYALSSDNGGWCNPPREHFDLAEPAFLQIAEYRAGIVPVLFRRVSCVKKGGIRFTLN
GHSYFNLVLITNVGGAGDITSASIKGSKTGWLPMSRNWGQNWQNNAYLNGQSLSFK
LTASDGRTITKYNVVPAGWQFGQTYEGDQF*

CDS (coding sequence)

>TcEXPA-10

ATGGCCACCACATCAATGCAAATTACATTGTCATTTCTCTTCTTATTCTGTCTGCT
CGGTATCTGCAGCTGCTTCCATGGCTCTAACAGAGATGATGGAGGTTGGCAAACT
GCCCATGCCACATTCTATGGTGGTGCTGACGCTACCGGCACAATGGGGGGAGCAT
GTGGTTATGGAAACTTGTATAGCCAGGGGTATGGAACGACCACCGCAGCTTTGAG
CACTGCACTTTTCAACAATGGCTTGACCTGTGGTGCCTGCTACGAGCTGCGGTGC
AATGACGACCCGAAATGGTGCATTTCTCGAACAATAACCGTGACTGCCACCAACT
TTTGTCCACCTAATTATGCTTTATCTAGTGACAATGGTGGGTGGTGCAATCCCCCT
CGAGAACACTTCGATTTGGCAGAACCTGCTTTCTTGCAGATTGCAGAATATCGAG
CTGGTATTGTTCCTGTTCTCTTCAGAAGGGTTTCTTGCGTGAAGAAAGGAGGCATC
AGATTCACCCTCAATGGCCATTCTTACTTCAACCTGGTCTTGATAACAAATGTGGG
AGGTGCAGGGGACATAACTTCCGCCTCCATCAAGGGGTCCAAAACAGGGTGGCT
ACCCATGTCAAGAAATTGGGGTCAAAACTGGCAGAACAATGCCTACCTTAATGGC
CAAAGCCTCTCTTTCAAACTGACTGCCAGCGATGGCAGGACTATCACCAAGTACA
ACGTAGTGCCCGCTGGCTGGCAGTTTGGACAAACTTATGAAGGAGATCAATTCTA
G

Nucleotide

>TcEXPA-10

CTCAAAGCCTTGCTCCAAGCAATAAAGTCACCCCTATTGTTTGAGCTTGCTCTCAC
TCAAATGGCCACCACATCAATGCAAATTACATTGTCATTTCTCTTCTTATTCTGTC
TGCTCGGTATCTGCAGCTGCTTCCATGGCTCTAACAGAGATGATGGAGGTTGGCA
AACTGCCCATGCCACATTCTATGGTGGTGCTGACGCTACCGGCACAATGGGTGAG
TTTCAGACTTGAGAAACACATAAATGTCGTTACTTGAATTTTTCTGATCTTTGATG

GTCTGTATTTCTGGGATATTTGTGATGGTTTTTTTAGGGGGAGCATGTGGTTATGG
AAACTTGTATAGCCAGGGGTATGGAACGACCACCGCAGCTTTGAGCACTGCACTT
TTCAACAATGGCTTGACCTGTGGTGCCTGCTACGAGCTGCGGTGCAATGACGACC
CGAAATGGTGCATTTCTCGAACAATAACCGTGACTGCCACCAACTTTTGTCCACC
TAATTATGCTTTATCTAGTGACAATGGTGGGTGGTGCAATCCCCCTCGAGAACAC
TTCGATTTGGCAGAACCTGCTTTCTTGCAGATTGCAGAATATCGAGCTGGTATTGT
TCCTGTTCTCTTCAGAAGGTAGTGCTAATCCTCTCCCAAATTTTACTTCTATTAAT
AGCACCAACTGTCACAAAGTAATTAAATATGGCATGTTACTCTCACCGAACTATA
CCTAAAACCAAGCAACTACCCTTCCTTTTCCGTAAGTTTTAAACCACAACTCGATT
GAAATCACACTCTATCTTCACCTGTTTTCATTATGAAAGATCGTTCCTAATAGACC
AAACTCCTTGTTTTAACTTGCAGGGTTTCTTGCGTGAAGAAAGGAGGCATCAGAT
TCACCCTCAATGGCCATTCTTACTTCAACCTGGTCTTGATAACAAATGTGGGAGGT
GCAGGGGACATAACTTCCGCCTCCATCAAGGGGTCCAAAACAGGGTGGCTACCC
ATGTCAAGAAATTGGGGTCAAAACTGGCAGAACAATGCCTACCTTAATGGCCAA
AGCCTCTCTTTCAAACTGACTGCCAGCGATGGCAGGACTATCACCAAGTACAACG
TAGTGCCCGCTGGCTGGCAGTTTGGACAAACTTATGAAGGAGATCAATTCTAGAG
AAATTGTTTCTATATATATATATATA