

IDENTIFICATION

Species: *Theobroma cacao*

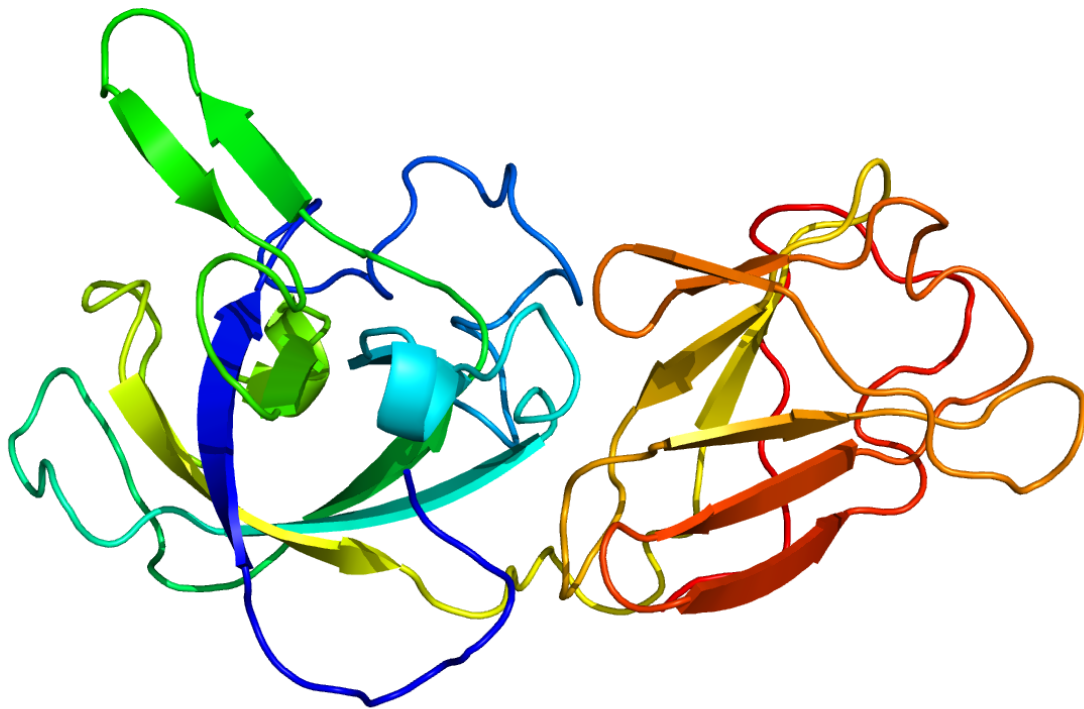
Locus: Thecc.08G049600

Gene Model: Thecc.08G049600.1.p

Description: TcEXPA-15

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

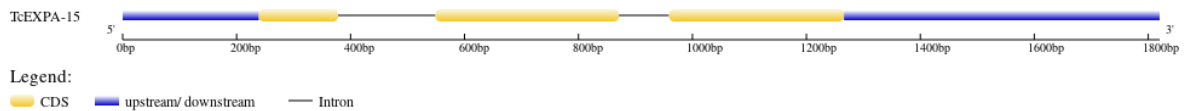
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

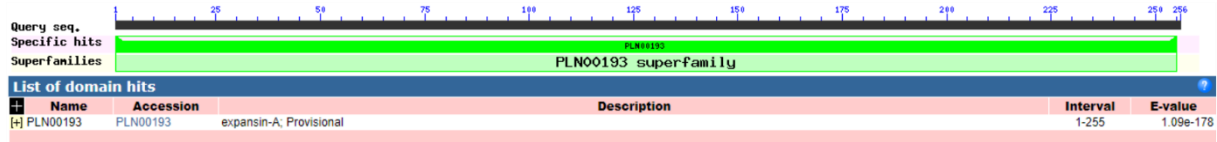
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-15

MAKIIFASAILLGLCYLATSVNAFSASGWTKAHATFYGGSDASGTMGGACGYGNLY
ATGYGTRTAALSTALFNDGASCGQCYKIMCDYQTDPRWCIKGASVTITATNFCPPNV
ALPNNAGGWCNPPLQHFDMAQPAWEKIGIYRGGIVPVIFQRPCKKHGGVRFTINGR
DYFELVLISNVGGAGSIQSVSIKSGKTGWMAMSRNWGANWQSNAYLNGQSLSFKVT
TTDGVTRLFPDVVPANWGFQSFSSNLQF*

CDS (coding sequence)

>TcEXPA-15

ATGGCTAAAATCATT TTTTGCATCAGCAATCTTGCTTGGACTCTGCTATCTTGCTAC
CAGCGTGAATGCCTTCTCGGCCTCTGGGTGGACTAAAGCTCATGCCACATTTTAT
GGAGGCAGCGATGCCTCGGGA ACTATGGGGGGGAGCTTGTGGATATGGAAACTTG
TATGCCACTGGCTATGGAACAAGAACTGCTGCTTTAAGTACTGCCTTGTTC AATG
ATGGAGCTTCATGTGGACAGTGCTACAAGATCATGTGTGATTATCAAACAGACCC
CAGATGGTGCATAAAAGGAGCATCTGTGACCATTACAGCTACAAACTTTTGCCCT
CCCAACGTTGCTCTTCCAAACAATGCCGGAGGCTGGTGCAACCCTCCTCTCCAGC
ATTTTGATATGGCTCAGCCTGCCTGGGAAAAGATCGGCATCTACAGAGGCGGAAT
TGTGCCTGTTATCTTCAAAGGGTTCCATGCAAGAAGCATGGTGGAGTAAGG TTC
ACCATCAACGGTAGAGACTACTTTGAGCTTGTATTGATCAGCAATGTGGGAGGAG
CAGGATCAATCCAATCTGTATCCATCAAAGGCTCTAAACTGGGTGGATGGCAAT
GTCAAGGA ACTGGGGAGCTAACTGGCAATCCAATGCCTATCTCAATGGCCAATCT
TTGTCGTTCAAGGTCACCACCACTGATGGAGTGACTCGACTATCCCTGACGTTGT
GCCAGCAAATTGGGGTTTCGGCCAGAGTTTCTCTAGCAATCTTCAATTCTAA

Nucleotide

>TcEXPA-15

CAGGGATAATTCAATTTTAAATGCTTAAAATAAATCGGTTTTCTCTCTTCTCTA
CTTGCAGGATCACATGCGAGGTTGTTGCATAGCGAGAGTTTGGTCCCTTATTTAA
AGAGCCAATTCTTCTGCTTGGAAATGCCAATAAATGTAAATACTACATTGATTC
TCAGTACTGCTCCTCATCCCGATCTTCTCTCTTTCTTCTCCACCAATTCGTTGTCT
TTTAAGCTACCGAGAAATGGCTAAAATCATT TTTTGCATCAGCAATCTTGCTTGG A

CTCTGCTATCTTGCTACCAGCGTGAATGCCTTCTCGGCCTCTGGGTGGACTAAAGC
TCATGCCACATTTTATGGAGGCAGCGATGCCTCGGGA ACTATGGGTTAGTAGCCA
CCCCCTTTTTCTTTTTCTTTATCTTTTTGTTTTGGATAAATGAAGTTTTATTCATCCA
AAGAAGAAGGAAGAGACTGAGAACCTGCAAGAAGCAGGAAACAAATACATTTCT
TCTGTAATTTTCGTCATCATTTGATATTGAATGATTTGTTGACAATAGGGGGAGCT
TGTGGATATGGAAACTTGTATGCCACTGGCTATGGAACAAGA ACTGCTGCTTTAA
GTA CTGCCTTGTTCAATGATGGAGCTTCATGTGGACAGTGCTACAAGATCATGTG
TGATTATCAAACAGACCCCAGATGGTGCATAAAAGGAGCATCTGTGACCATTACA
GCTACAAACTTTTGCCCTCCCAACGTTGCTCTTCCAAACAATGCCGGAGGCTGGT
GCAACCCTCCTCTCCAGCATTTTGATATGGCTCAGCCTGCCTGGGAAAAGATCGG
CATCTACAGAGGCGGAATTGTGCCTGTTATCTTCCAAAGGTAGCATCTCTATCTCA
GTCTCTTCCCTCTTTTTACTCTTAAAATCTTATATCCTTATGTCAACGTTTCGGTG
TTCTTACGATACAGGGTTCCATGCAAGAAGCATGGTGGAGTAAGGTTACCATCA
ACGGTAGAGACTACTTTGAGCTTGTATTGATCAGCAATGTGGGAGGAGCAGGATC
AATCCAATCTGTATCCATCAAAGGCTCTAAA ACTGGGTGGATGGCAATGTCAAGG
AACTGGGGAGCTAACTGGCAATCCAATGCCTATCTCAATGGCCAATCTTTGTCGT
TCAAGGTCACCACCACTGATGGAGTGA CTGACTATTCCCTGACGTTGTGCCAGC
AAATTGGGGTTTTCGGCCAGAGTTTCTCTAGCAATCTTCAATTCTAAGTTACTAAAA
GCTTTGGAGTTCAGGATTCTGGGTTTTTAGTTCCCGGCAGAGGGCGTGCTTATATTT
TACAAAAAAGCTGCCCGCCATACATTTCTTGGGATATACCAGATGAGATGGTTA
AACTCACACGGCTACTTTTCTGTTCTCTGGGTGGTGGTTCAGATGGAAGTGAATCC
TTGAACTCAAAGCCGCTACATTACTTTTTGTGCTTTCGATACAAGAAATTCGGACG
ATGTTTGGAATTATTCTAAGATCATATCAATAAGAAGCGCAAGAAAAGAAAAA
CTTGCAGCTTCAGCTTATGTAATTTTGTTCACAAATTTATATGTTTGTAATAG
GAGTTGTTGCTTGATCGTTCTACCCATTTTTCATTCCAATTGAATAAAGCAACCGA
CAAACCTTCTGTCATTTACCCCAACAAGGTTAAATTCTGTTCTGTCTCATAGATTG
AAAGCTCTGTTGTTTGGATGTAGACTATAGCAGAAGAGGCTACTAACTCATGGCC
TCTGCCTTTTTTAAGGGAATAAAAGAAACGAAACTTGAAAAGAC