

IDENTIFICATION

Species: *Theobroma cacao*

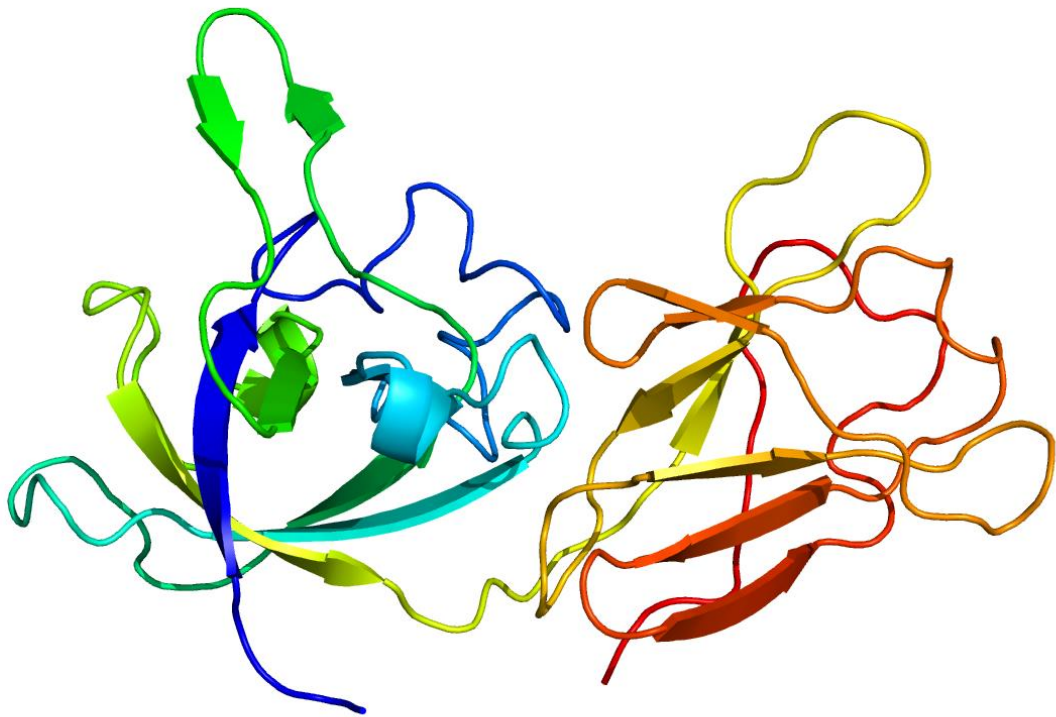
Locus: Thecc.02G333600

Gene Model: Thecc.02G333600.1.p

Description: TcEXPA-03

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

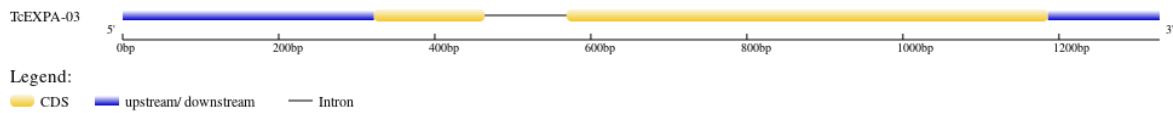
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

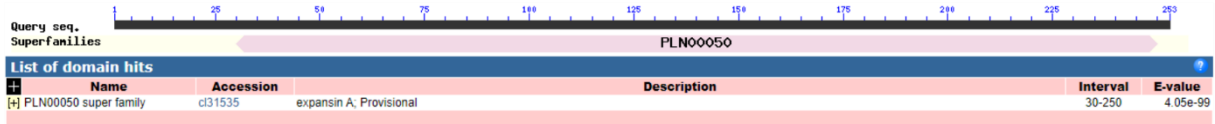
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-03

MAGGGSIRVIMLLGLCLVSRASLQAEQEWNVAHATFYGDMQGTETMMGACGYGN
LLEQYGLETTALSTALFNGLTCGACYEIRCYDSAQWCLNETVVVTATNFCPPNYR
KPEGNWCNPPRRHFDLSQPMFRRIAVYRAGIVPVIYKRVPCRKYGGLKIEIMGNPNWI
LVLVYNVGGAGHVINVKIKGSSTAWIQMSRNWQNWQTAAQVVGQSLSFQVTTSD
GKMVQSDDVAPGDWKFGGVYEGIQFPS*

CDS (coding sequence)

>TcEXPA-03

ATGGCTGGTGGGGGGTCTATTCGTGTGATAATGCTCTTGGGTTTGTGCTTGGTCAG
TAGGACTGCTAGCTTACAAGCAGAACAGGAATGGAATGTAGCACATGCAACTTTT
TATGGAGACATGCAAGGGACTGAAACAATGATGGGTGCTTGTGGCTATGGCAAT
CTACTAGAACAAGGTTATGGCCTTGAAACAACGGCCTTAAGCACAGCTCTCTTCA
ATGGTGGTCTAACCTGCGGAGCCTGCTATGAAATCCGGTGCTACGACTCCGCACA
GTGGTGCCTAAACGAGACCGTTCGTTGTAACCTGCCACAACCTTCTGCCCTCCCAAC
TACAGAAAACCTGAAGGAAATTGGTGCAATCCACCGAGGCGACACTTTGACCTAT
CCCAACCAATGTTCCGAAGAATTGCTGTTTACAGGGCTGGCATTGTTCCCGTCAT
ATATAAGCGCGTTCCATGCAGGAAATATGGTGGTCTCAAGATTGAGATTATGGGA
AATCCGAACTGGATACTGGTGCTAGTGTACAATGTGGGGGGGGCTGGTCATGTTA
TAAACGTTAAGATCAAGGGTTCAAGCACAGCATGGATACAAATGTCACGAACT
GGGGCCAGAACTGGCAGACTGCAGCTCAGGTTGTCGGGCAAAGCTTGTCCTTCCA
AGTGACTACTAGTGATGGTAAAATGGTGCAATCTGATGATGTTGCACCAGGTGAT
TGAAATTTGGTGGCGTTTATGAGGGAATCCAATCCCAAGTTAA

Nucleotide

>TcEXPA-03

AATTATTATTGTATTTACAGTACTAAGAAGTTGATTTCCCCCAGAGTTATCGGTTG
CAACAGCAGTAAATGGAAATCCGAAAAGATTATCTTTACATGTCCAACAATTCCA
AACACAGAATCCCAAACCTTATTAATCAACACAACTACACTTGCATGGCTTCCA
CTTTGATAACTTTCAATTAATTTTTATTTTTCCAATTTAAGATTCCCTCCCTCTTATC
ATTTACTCACATACCTGCCTATTTATACCTAGTACAGATTTCCAACCTTCCCAAG
CATTCAAGCATAAGAACAAGTAACAGAAAGCTGGAGAGGTTACATGGCTGGTGG

GGGGTCTATTCGTGTGATAATGCTCTTGGGTTTGTGCTTGGTCAGTAGGACTGCTA
GCTTACAAGCAGAACAGGAATGGAATGTAGCACATGCAACTTTTTATGGAGACAT
GCAAGGGACTGAAACAATGAGTAAGTTTTGTAATCCGTTGTCATTTCTGAATTTT
GATTCTAAGTATTCCTAGTCTAATACTACTAACATCATCCCTTGTATTTGTGGTAT
GCATGAGGATATAGTGGGTGCTTGTGGCTATGGCAATCTACTAGAACAAGGTTAT
GGCCTTGAAACAACGGCCTTAAGCACAGCTCTCTTCAATGGTGGTCTAACCTGCG
GAGCCTGCTATGAAATCCGGTGCTACGACTCCGCACAGTGGTGCCTAAACGAGAC
CGTCGTTGTAAGTCCACAACTTCTGCCCTCCCAACTACAGAAAACCTGAAGGA
AATTGGTGCAATCCACCGAGGCGACACTTTGACCTATCCCAACCAATGTTCCGAA
GAATTGCTGTTTACAGGGCTGGCATTGTTCCCGTCATATATAAGCGCGTTCATGC
AGGAAATATGGTGGTCTCAAGATTGAGATTATGGGAAATCCGAACTGGATACTG
GTGCTAGTGTACAATGTGGGGGGGGCTGGTCATGTTATAAACGTTAAGATCAAGG
GTTCAAGCACAGCATGGATACAAATGTCACGAAACTGGGGCCAGAACTGGCAGA
CTGCAGCTCAGGTTGTCGGGCAAAGCTTGTCTTCCAAGTGACTACTAGTGATGG
TAAAATGGTGCAATCTGATGATGTTGCACCAGGTGATTGGAAATTTGGTGGCGTT
TATGAGGGAATCCAATTCCCAAGTTAAAGGCAACGGTGTCTGCTTTTCTTAACTG
AATGCACATGAAGCTTGTTAACCTGTGTAATAGCATGAATAACTGTGATGACTCT
TGACTTCCAGGGTCAGAAATTGTTCAAACTCCACCTCTTAATAAATGAAAACCTT
GTCC