

IDENTIFICATION

Species: *Theobroma cacao*

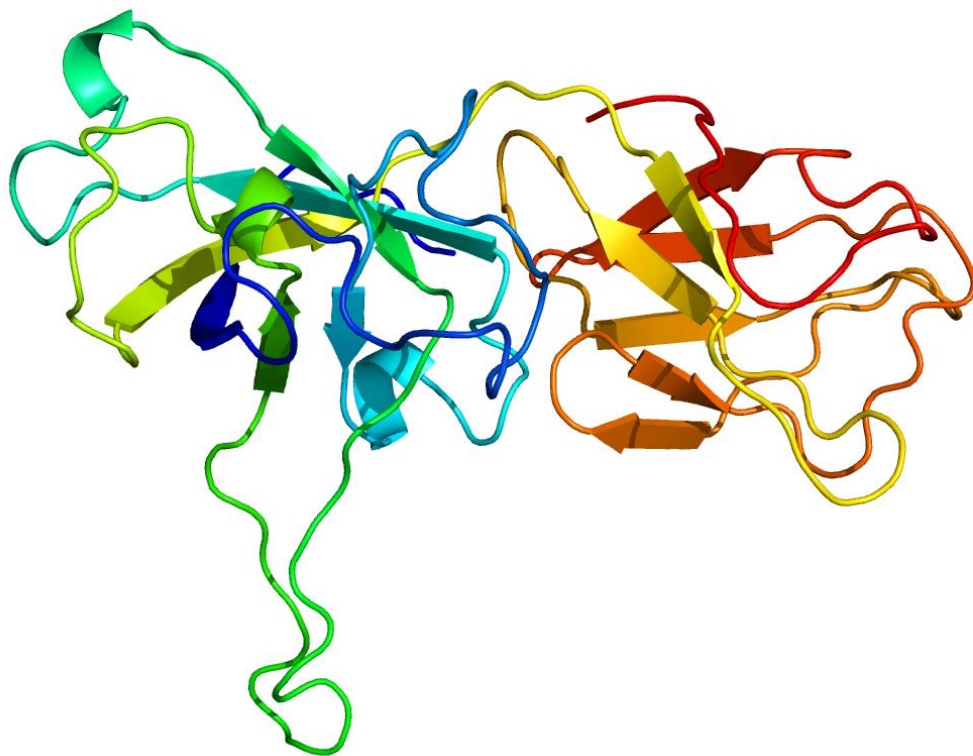
Locus: Thecc.02G147900

Gene Model: Thecc.02G147900.1.p

Description: TcEXPA-02

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

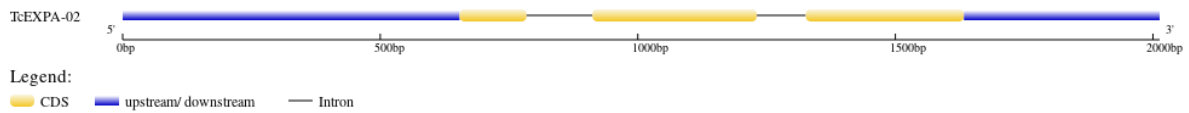
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

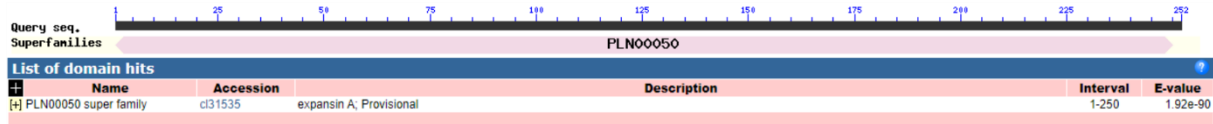
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-02

MGALEVTVLYLILLTTKVISKEE EWKSATATYTKETD GSIIRGGACGYGDIHKATYG
KYSAGLSTKLFNKGSTCGGCFELRCVDHILWCLQGSPSVILTATDFCPPNYGLSADYG
GWCNFPKEHFEMSEAAFAEIAERKADIVPVQYKRVKCERRGGIRFTMSGSSHFCQVLI
TNVGLDGEVVGVKVKGPKTGWIPMARNWQNWQCNINLIGQPLSFEVTASSGRTLA
SYSVAPANWQFGQTYEGKQFK*

CDS (coding sequence)

>TcEXPA-02

ATGGGTGCCCTGGAAGTTACAGTTCTCTATTTGATTCTACTGCTAACAACCAAGGT
CATTTC AAGGAAGAAGAATGGAAGTCAGCTACTGCAACATATACCAAAGAGAC
AGACGGATCAATCATCAGAGGAGGAGCTTGTGGTTATGGGGATATTCACAAGGC
GACCTATGGGAAGTACAGTGCTGGACTTAGTACCAAGTTATTCAACAAAGGGAGT
ACCTGTGGAGTTGCTTTGAGCTCAGATGTGTTGACCATATCTTGTGGTGTCTCCA
AGGGAGCCCCTCTGTCATCCTCACTGCCACAGATTTCTGTCCACCGAATTATGGG
CTTTCAGCAGATTACGGTGGCTGGTGCAATTTTCCAAGGAACACTTTGAGATGT
CTGAGGCAGCGTTTGCTGAAATTGCAGAGAGAAAAGCAGACATTGTGCCGGTTC
AGTATAAGAGGGTAAAGTGTGAAAGAAGGGGTGGGATAAGATTCACAATGAGTG
GGAGCTCTCACTTCTGTCAGGTCCTGATTACCAATGTGGGATTGGATGGTGAAGT
AGTAGGAGTGAAAGTGAAGGGCCCCAAAACAGGGTGGATACCTATGGCGAGAAA
CTGGGGACAAAAC TGGCAGTGCAATATCAACCTTATAGGACAGCCACTTTCTTTT
GAGGTA ACTGCCAGTAGCGGCCGGACACTTGCATCTTACAGTGTTGCTCCAGCAA
ATTGGCAGTTTGGTCAGACGTATGAAGGAAAACAATTCAAGTAA

Nucleotide

>TcEXPA-02

GCTAGTGAGCGCCTCCCAACAGTCAAAGCCTTTTTCCCTGTGAAGTCCCTCCCATC
TTTATCTTTCTCTCCTAAAAAGAAGAGCCAATGGATGTATCTTGACTTTTTTAAGT
ACTTTGCTAGTTGCTATAAACA AATCTAAAACCCTTCATAATTGTTTTAGGATTCT
TAATAGTCCTAGTTTATCTTTTTCTCAGACTTTCGCTTTTCCTGTCTGCCACCATTA
CAAAGCCTTAGACCCTTCTGGCCTTGCTCTCGTAATTATGTAGCCAAATGCAGGC
AGCACACTTTTTTCAGTCTCATTACCAACTAGAAAGAAGAAATGAAAAAAAAG

AAAGCATTCCCTTCTTTTCCGTTATAGACAAAGATACCCTTCATCAGGCTCTCGTCCG
AGTGACCCCAACAATATGAGCCAAGGAGCCCTATTTGTCTTTTGGCCTAAGTTCA
AACTTTGTA CTCTGAATGCGGTTTCCACCTAGATTACAAGTTAAGGACAGACCTA
GCTTAGTTTACCTCAACAAGGAGTTTGAGGTCCTTTCCCCAGGGAAAGGGGCTT
TTTTAACCTTCCTCTCACAACAATCCCCATAAAAAGGAAAGAAGAAAAAGAATC
ACATATAATCATTAATCACTACTCAAAGTTCTGAGAGGCCTCAAATGGGTGCCCT
GGAAGTTACAGTTCTCTATTTGATTCTACTGCTAACAACCAAGGTCATTTCCAAGG
AAGAAGAATGGAAGTCAGCTACTGCAACATATACCAAAGAGACAGACGGATCAA
TCATCAGAGGTA CTTTCTATTGCAACTTTTCGATTTAGAGTTTCAACCACAGAATGC
TCTTTTATATTAACAAGTATTTCAAATTGGTACAGAAA ACTAAAATTAATTTAAGC
TAACAGTACTGGCCTAATCTGGCAGGAGGAGCTTGTGGTTATGGGGATATTCACA
AGGCGACCTATGGGAAGTACAGTGCTGGACTTAGTACCAAGTTATTCAACAAAG
GGAGTACCTGTGGAGGTTGCTTTGAGCTCAGATGTGTTGACCATATCTTGTGGTGT
CTCCAAGGGAGCCCCCTCTGTCATCCTCACTGCCACAGATTTCTGTCCACCGAATTA
TGGGCTTTCAGCAGATTACGGTGGCTGGTGCAATTTTCCCAAGGAACACTTTGAG
ATGTCTGAGGCAGCGTTTGCTGAAATTGCAGAGAGAAAAGCAGACATTGTGCCG
GTTCAGTATAAGAGGTATGACATATCAA ACTATATACGTTTGGCAAGTTTTAATA
AAATTTTCATATATTTGTATCAAAGGAATCATAGTTTGTGCAAACATTGTGTAGG
GTAAAGTGTGAAAGAAGGGGTGGGATAAGATTCACAATGAGTGGGAGCTCTCAC
TTCTGTCAGGTCCTGATTACCAATGTGGGATTGGATGGTGAAGTAGTAGGAGTGA
AAGTGAAGGGCCCCAAAACAGGGTGGATACCTATGGCGAGAAACTGGGGACAAA
ACTGGCAGTGCAATATCAACCTTATAGGACAGCCACTTTCTTTTGAGGTA ACTGC
CAGTAGCGGCCGGACACTTGCATCTTACAGTGTTGCTCCAGCAAATTGGCAGTTT
GGTCAGACGTATGAAGGAAAACAATTCAAGTAACGAAACGATCTCTGAGATTTTT
ATATAGGGAAAGATGGCATAGATAATTTACCTCTCAGCCATTGCTGGAAACTTT
AAGATCTATATTGAAGAGGTGTACAAGTTCGCACTAACGAGATAGTGTGAAAGT
GAGAACATCTCCAGATTTAACATTGAAGCGGCATAATAGTCAAAAAGAGAAGAAA
AAGTGTAGAAAAAGGTTGAAGAAACAGGAATAGAAGGTCCCAGTTTTTGTCACTT
TTTTAATATTGTTGCTAGAGAGCAAAAATTTCTATGCAATGTATGATTATTTAAAA
GCAACGTTATATATAAATCTCACCTTGCATTCTCTGGTTTTGCTTTTGGGTGTATTT
CCAGATGCTGCGCCCGCAAAATATTTTG