

IDENTIFICATION

Species: *Theobroma cacao*

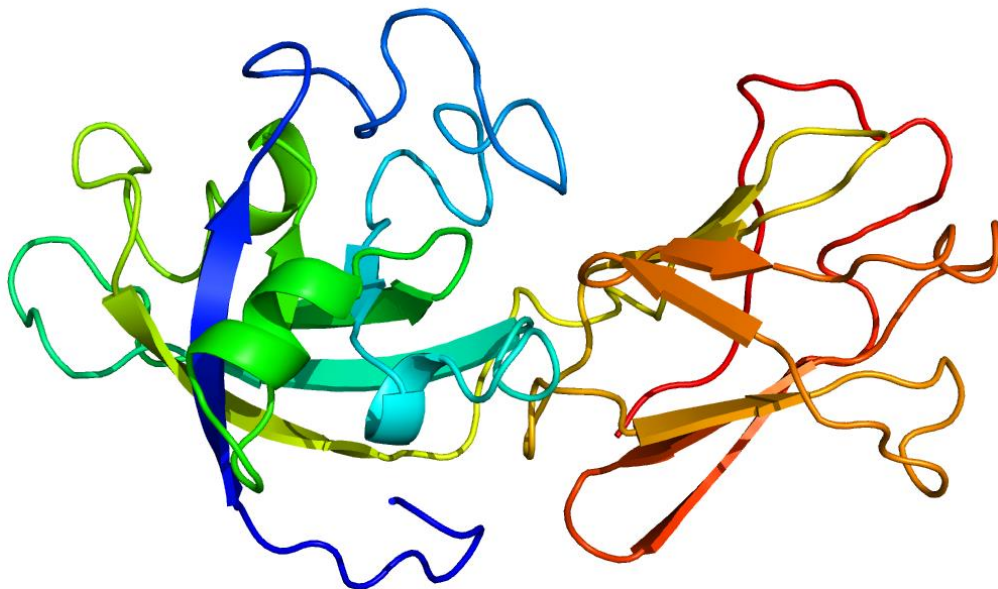
Locus: Thecc.04G084700

Gene Model: Thecc.04G084700.1.p

Description: TcEXPA-08

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

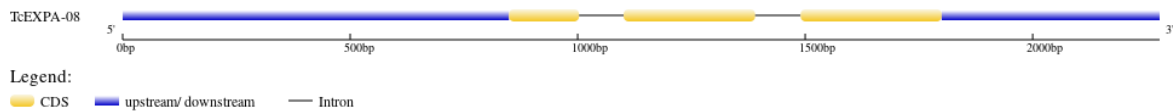
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

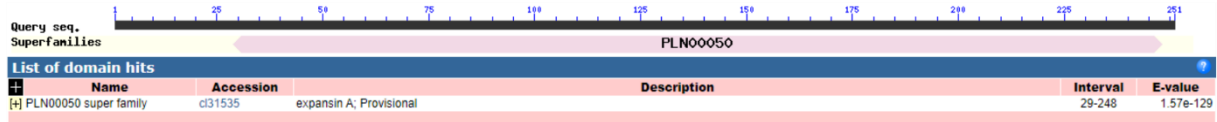
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-08

MGLLGFLLVLLSVLNFVHGFPRNHAPHGYLGPWINAHATFYGGGDASGTMGGAC
GYGNLYSQYGINTAALSTALFDNGLSCGACFELRCVNDPQWCLPGSIVVTATNFCP
PGGWCDPPNHHFDLSQPIFQHIAQYRAGIVPVIYRRVRCRRSSGGIRFTINGHSYFNLVL
ITNVGGAGDVHSVSIKGSRTRWQPMSRNWQNWQSN SYLNGQSL SFLVTTSDGRSV
VAYNVAPAGWSFGQTYAGGQFRF*

CDS (coding sequence)

>TcEXPA-08

ATGGGGCTTTTGGGGTTTCTTCTAGTTGTATTGCTCTCAGTCCTCAACTTTGTTCAC
GGTTTTCCAAGGAATCATGCCCTCATGGATACCTTGGACCCTGGATCAACGCTC
ATGCCACTTTCTATGGAGGTGGTGTGCTTCTGGAACAATGGGTGGAGCTTGTGG
CTATGGAAATCTATATAGCCAAGGTTATGGTATTAACACAGCAGCATTGAGCACT
GCATTGTTTGACAATGGCTTGAGCTGTGGCGCCTGCTTTGAGCTCAGGTGTGTCA
ATGACCCACAATGGTGCTTGCCCGGCTCCATTGTGGTTCACAGCCACCAATTTCTG
CCCTCCAGGAGGCTGGTGTGACCCTCCTAATCACCCTTCGATCTCTCTCAGCCTA
TTTTCCAACACATTGCCCAATACAGAGCTGGTATTGTTCCAGTAATTTACAGAAG
GGTGAGGTGCAGGAGAAGTGGGGGTATCAGGTTCAACAATTAATGGACATTCCTA
CTTCAATTTAGTACTGATCACCAATGTTGGAGGTGCTGGTGTGTTTCTGTAT
CCATTAAGGGTTCAAGGACCAGATGGCAGCCAATGTCAAGAACTGGGGCCAGA
ACTGGCAGAGCAATTCTTATCTTAATGGTCAAAGCCTCTCTTTTCTTGTTACAACC
AGTGATGGCCGCAGTGTGTCGCTTACAATGTTGCTCCAGCCGGTTGGTCTTTGG
ACAGACTTACGCTGGAGGGCAGTTTAGGTTCTAA

Nucleotide

>TcEXPA-08

CAAAGGTGTTTTAACTTTAAAAGTGTTTCATGACTTGGTGACGGAGATATAACTTA
ATTGATAGCCACATCTCTCCTTCTTCAAAGAAGATTCAAAGTTCAAATTCTAGCGT
GGTTTACGTGCAAAGTGAGAAAATTAGTTAATTTGGAAACCAGTAGTTTGTAT
GAATGAACATGTAGTGAAATGATGTGGAAACTTTTAAACAAAGGATTCGAAACAA
GTTGAGAGTGCAGAGTCAAAGGCCCTAAAATCTTCGCAGGAGCAAAGTTTCACT
CCTAGATGAAGCAATTCAGTTTTTTGTCATCCAATATTGTTTCGCTGCACGCTTG

TGATCTAATTAAGGATTTCAAGCAGAGAGCCAAATTCACCTACACCCCCTTTGG
ATTCATACCTTGTTTGTCCCCCTATGGCTGTAAACACCCCTCAAACCACGCGGTT
TGAGCCTTGAAAACCGCTTTTTCTAGCCCCATAAAAACCTGCATTTGTTTCCCCAAA
ATCAAGAGTCCCCAAAAGCCACTACGCAGTTGCTTGGTCCCAACAGCTCAATTGT
TGTCAGGTTGTCAGTCAGGTAGGATGTTGACTGCCCTCTATTTATCCCTTCAAAC
ACCCCCTCTTCATTTCTACAAACAACCTCCTTTCAAGGTTCAATTTACTAGCCCCAAT
CTTCTTTTTCTTCCTCTTCCAGGGTTTTTTTTTTTAAAACCTTTAAAACAGCATAATG
AAAAAAAAAAAAAAAAACAAAGGCTTTCTAGTAGTACTTAAATATGCTTGTACTCTT
CGTTAGTCTTAGTACAATTTTTTACATAATGTTGAGGGGTTTATGGAATATGTTT
TTTCATATGCAGGGCAATGGGGCTTTTGGGGTTTCTTCTAGTTGTATTGCTCTCAG
TCCTCAACTTTGTTACGGTTTTTCCAAGGAATCATGCCCTCATGGATACCTTGGA
CCCTGGATCAACGCTCATGCCACTTTCTATGGAGGTGGTGATGCTTCTGGAACAA
TGGGTACATTCCACATCACATTTTTTGTATTCTTGTGAGTGGCGTGTCTTCTTG
GTAATCTACTTTACTTAATGCATTACATTCTGGTATTGCTATAGGTGGAGCTTGTG
GCTATGGAAATCTATATAGCCAAGGTTATGGTATTAACACAGCAGCATTGAGCAC
TGCATTGTTTGACAATGGCTTGAGCTGTGGCGCCTGCTTTGAGCTCAGGTGTGTCA
ATGACCCACAATGGTGCTTGCCCGGCTCCATTGTGGTCACAGCCACCAATTTCTG
CCCTCCAGGAGGCTGGTGTGACCCTCCTAATCACCCTTCGATCTCTCTCAGCCTA
TTTTCCAACACATTGCCCAATACAGAGCTGGTATTGTTCCAGTAATTTACAGAAG
GTAACACAATTAATAATTTTAGCAAAAAGTTTCCACATTTTCTCTTATATTTCTA
ATGTGCGTAATAAGGTTACATATATAAAAATTTTCTGTTTTGCAGGGGTGAGGTGCA
GGAGAAGTGGGGGTATCAGGTTACAATTAATGGACATTCCTACTTCAATTTAGT
ACTGATCACCAATGTTGGAGGTGCTGGTGTGATGTTTCTGTATCCATTAAGGGTT
CAAGGACCAGATGGCAGCCAATGTCAAGAACTGGGGCCAGAACTGGCAGAGCA
ATTCTTATCTTAATGGTCAAAGCCTCTCTTTTCTTGTTACAACCAGTGATGGCCGC
AGTGTGTCGCTTACAATGTTGCTCCAGCCGGTTGGTCCTTTGGACAGACTTACGC
TGGAGGGCAGTTTAGGTTCTAAGTTTCAACAACCCCTAAGCATACTATAACTGCC
GTTGCTATAGTTAAGATATTTTGTCTGAACAAAACTAATGTGAGTTACAAAAG
AACTATATAGTAGAATAGTATGAGGGTGAATCTTGGGTCAAGATTGATTTTTA
AAGAGCCCTTTCCTGTGTAGCATTATTGTGAGTTTTTGCAGTTCTATTGGCTTTGA
ATTGAAGCCTGGGGGTAAGTGTGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGA
GAGAGGTACAGTATTTAGAGGCATATTATTAGTGAGATTCTGCCTAAAGGGCATT
CTTACTATCCAGCTACTGCTACATCAAATGATGAATAGCAGGTGTTTTTTCAGGCC
AAAAGGCACTTTTGCTTGCTTGTGAGAAGCCACTGTCTTTCTTGTTGTGCACTTTG
AACAATTATCCAAAATTAAGATTTCAGCAATTTGTGAGGACAGAGAGATCTCA
ACTTTT