

IDENTIFICATION

Species: *Theobroma cacao*

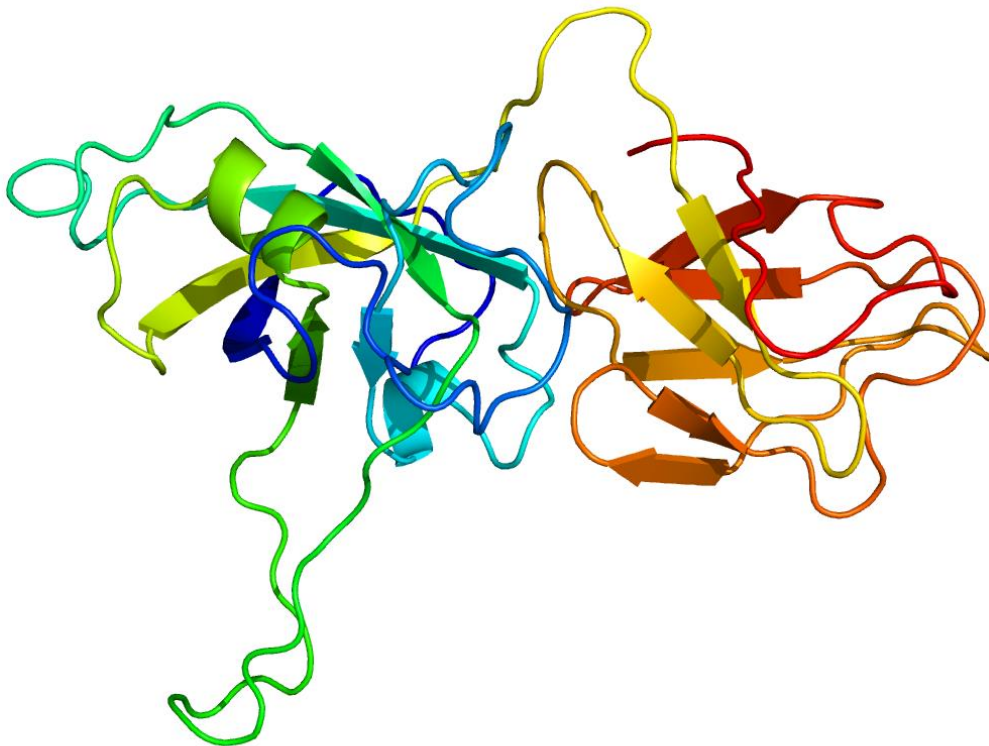
Locus: Thecc.05G038700

Gene Model: Thecc.05G038700.1.p

Description: TcEXPA-11

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

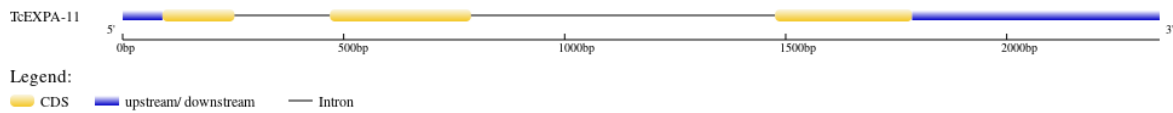
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

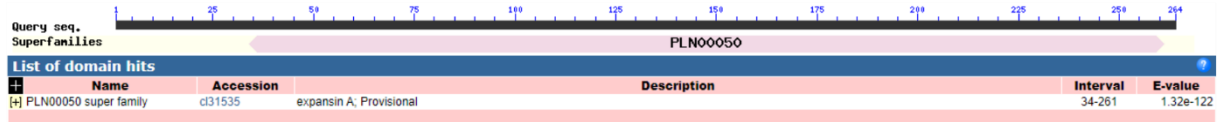
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-11

MAAAATSSLTVLCFASLLSFSVVVNARIPGVYTGGSWESAHATFYGGSDASGTMGG
ACGYGNLYSQGYGVNTAALSTALFNGLSCGACFEIKCANDPRWCHPGSPSIFVTAT
NFCPPNYALPSDNGGWCNPPRSHFDLAMPMLKIAEYRAGIVPVSFRRVPCRKQGGI
RFTINGFRYFNLVLVTNVAGAGDIVKVSVKGTGTGWLSMSRNWQNWQSNVAVLVG
QALSFRVTGSDKRTSTSWNVASANWQFGQFTGKNFRV*

CDS (coding sequence)

>TcEXPA-11

ATGGCTGCTGCTGCTACTTCTTCCCTGACTGTGCTGTGTTTTGCTTCATTGCTGTCC
TTTTCTGTTGTGGTGAATGCAAGGATCCCTGGAGTTTACACCGGTGGCTCCTGGG
AAAGTGCCCATGCTACCTTTTACGGTGGCAGTGATGCTTCTGGAACAATGGGCGG
TGCATGTGGGTACGGGAATCTCTACAGCCAAGGCTATGGCGTGAACACAGCAGCT
CTCAGCACAGCACTTTTCAACAATGGCCTTAGCTGTGGTGCTTGCTTTGAGATTAA
GTGTGCAAATGACCCAAGATGGTGTACCCTGGCAGCCCTTCCATTTTCGTCACA
GCCACCAACTTTTGCCCTCCTAACTATGCGCTCCCAAGTGACAATGGCGGCTGGT
GCAACCCTCCGCGCTCTCACTTCGACCTTGCCATGCCTATGTTCCCTAAGATCGCT
GAGTACCGTGCTGGTATTGTCCCCGTCTCTTTCCGTAGGGTGCCGTGCCGAAAGC
AAGGGGGAATCAGATTCACCATTAACGGATTCCGTTACTTCAACTTGGTTTTGGT
CACCAACGTCGCGGGTGCAGGGGATATCGTGAAGGTGAGCGTGAAGGGAACCAA
AACCGGCTGGTTGAGCATGAGCCGTAAC TGGGGTCAGAACTGGCAGTCCAACGC
CGTCTTGGTGGGCCAGGCGCTCTCCTTTAGAGTCACGGGCAGTGACAAGCGAACC
TCCACCTCGTGGAACGTGGCATCAGCTAATTGGCAGTTCGGCCAAACCTTCACGG
GAAAGAATTTCCGCGTTTGA

Nucleotide

>TcEXPA-11

ATTCGCCCCAGAATCAACACACCCAACATTTCTCTTTCCCTTTTCTCTTTCTTTCTT
TCTTCTTTGTTCTCTCTGAAAGAAGAAAACAGAATGGCTGCTGCTGCTACTTCTTC
CCTGACTGTGCTGTGTTTTGCTTCATTGCTGTCCTTTTCTGTTGTGGTGAATGCAAG
GATCCCTGGAGTTTACACCGGTGGCTCCTGGGAAAGTGCCCATGCTACCTTTTAC
GGTGGCAGTGATGCTTCTGGAACAATGGGTATTGGCTTTTAAGCATAATTTTGTGA

CAATTAGTACGAATTTTCGTCAATGAGGCCCTTGTTCTCATGTTCTGTTTTTGGAAAG
CATGAAAACCCCATTTCTTCAACATGCATTCACTAACAAAATCTCAAACCTGTAATA
TATATTTCTCTACATTCCCTAAAACCATTA AAAATGTGCCTCAAAAATTTGCTCGTT
TTGTTTTGTTTTTCTGGTTAGGCGGTGCATGTGGGTACGGGAATCTCTACAGCCAA
GGCTATGGCGTGAACACAGCAGCTCTCAGCACAGCACTTTTCAACAATGGCCTTA
GCTGTGGTGCTTGCTTTGAGATTAAGTGTGCAAATGACCCAAGATGGTGTCACCC
TGGCAGCCCTTCCATTTTCGTACAGCCACCAACTTTTGCCCTCCTAACTATGCGC
TCCAAGTGACAATGGCGGTGGTGCAACCCTCCGCGCTCTCACTTCGACCTTGC
CATGCCTATGTTCCCTAAGATCGCTGAGTACCGTGCTGGTATTGTCCCCGTCTCTT
TCCGTAGGTGAGTCCAAAATCTCTCTCCTTTGCCTCAGAATTTTCATGCCTTTTTGT
TTTCTTAAAGTACTAGTGTTCCTTAAGTTGGTATTTTCGTTACTTAGAGAAGTAAAT
GCGAGTGAAAGGCCAAAAGAATACTTGATAACAGTACATAAAATGTTGTGCATTTT
TGCTCAATTTTCTTAGGAACCAAACAGATTTTTTTTTTTGTTTTTAAATTTCTATTAAT
ATTAAATTTTCGTTTTTCTTCCCATCACTTTCTTGGGAAAGCAAAAACGGGGTTGGC
CTCAAAGGTATCGCCCTTTTGTGAGCTCATTTGGACGGCGTCCGTTTGAATAAATC
CAGATTTTTTAATGTCTTTTACCCAAGTTTAAACCAACTGTCTTTCGATCCGTGTTGT
ATGACCAAACACCCCTTGCCTTTTAAAGGCTGACCGTGCTTCCAACCCATTGATCT
CTCGAACCTCAAAGGGCATCAAAGTGAAAAATGTTGCAAGGGTATTATTGTAAT
TTAGTGTTTACTAAGATTTTCCGGGAAATCGGACAGCCCAAGGGACAAAATTCA
CGACATCCCACGCGTGGGTTCGGTCATTTTCATCAATGAGGAAAAAATGGGGAATG
ACACAATATTGAATTTCCATAATTACCCCCAACTCTGTTTTACGGAGTTTAATTAT
TATTATTATTATTATTTTTTACAGGGTGCCGTGCCGAAAGCAAGGGGGGAATCAGA
TTCACCATTAACGGATTCCGTTACTTCAACTTGGTTTTTGGTCCCAACGTCGCGGG
TGCAGGGGATATCGTGAAGGTGAGCGTGAAGGGAACCAAACCGGCTGGTTGAG
CATGAGCCGTAACCTGGGGTCAGAACTGGCAGTCCAACGCCGTCTTGGTGGGCCAG
GCGCTCTCCTTTAGAGTCACGGGCAGTGACAAGCGAACCTCCACCTCGTGGAACG
TGGCATCAGCTAATTGGCAGTTCGGCCAAACCTTCACGGGAAAGAATTTCCGCGT
TTGAAGACAACAGACAGAGACAGACAGTCACAGTCCCCCCTTCTCCGCCCTTT
TTCTTTTAATTTTTATTAATTAACCTTTTTTTTTATCTTTGTTTCCCGCCCTTCTATT
TTTCCTTTATTTTCCCGGAAAAATTTGAAGTGGTGGGTTGGGTGGGTAGAATCCG
TGAAAGTAACAGTGGGGTTAAAATGGGGTTTTTGGTAGGATTTACGGAAAAGAA
AGTGAAAAAAGTAAAAAGAAAAAGTGGTGATAGACTTAGGAAAGTGACTAAA
GTGAACATATGTGTGGTAACTTTTTTTTTTCTTCTTTTGGGGCCTAATATTGGGTCCA
TGTTTCTTTTCTTTGGTTTGTCAATTGTGCTGGTGGTAATAGTAAAAATCCACTGA
ATGGTAAAAAAGGGCTGTAAGTGAGGGGCTGAAGTGGCTGCAAAATC
CAAAGAAGCATGTAGCCCGCAGCTCTTGGTCATATAGTAATTTCCCATTTTACTAC
TATATATTA AAAATGTTAGAAGTATTTTACTTTTTGTCTTTAAATCAAATGGTGTTG
CTTTCTTCC