

## IDENTIFICATION

**Species:** *Theobroma cacao*

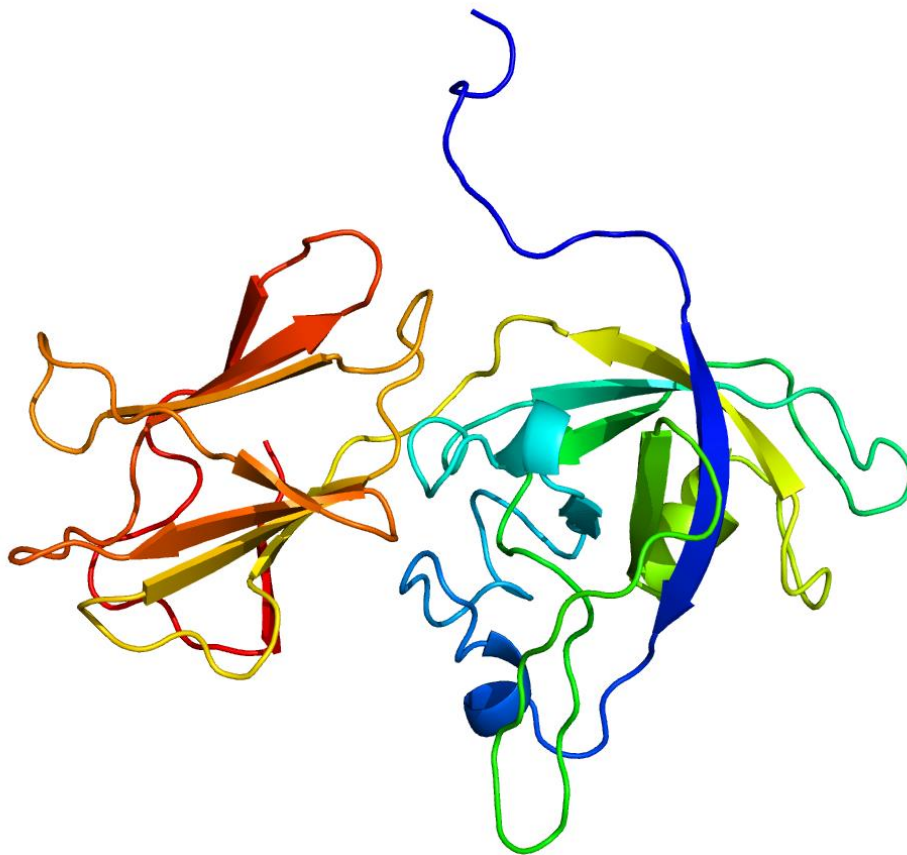
**Locus:** Thecc.04G065000

**Gene Model:** Thecc.04G065000.1.p

**Description:** TcEXPA-07

**Family:** Alpha Expansin

**3D structure:**



## GENOME DATABASES

Phytozome: [https://phytozome-next.jgi.doe.gov/info/Tcacao\\_v2\\_1](https://phytozome-next.jgi.doe.gov/info/Tcacao_v2_1)

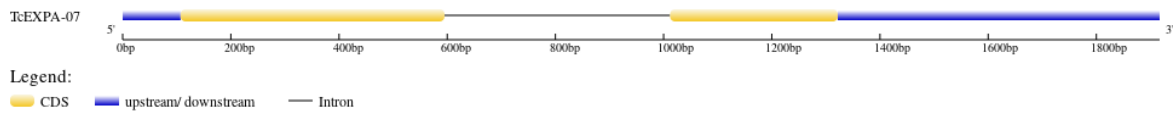
KEGG: <https://www.genome.jp/entry/gn:T02994>

## EXTERNAL RESOURCES

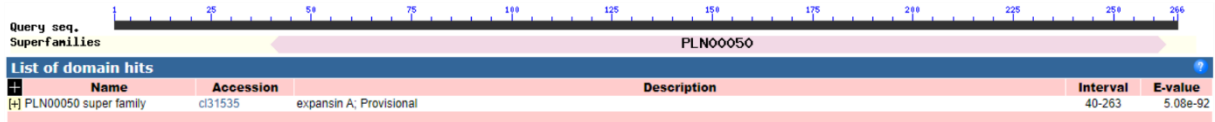
<https://www.cacaogenomedb.org/>

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

## GENE STRUCTURE



## DOMAIN ARCHITECTURE



## SEQUENCES

### Peptide

>TcEXPA-07

MSPPAKTEAPLLALLLLTLTSTATSHYTSSSPAPQSSQWHPARATYYAASDPRDAVG  
GACGYGDLVKAGYGMATVGLSESLFERGQICGACFELRCVDDLRCIPGTSIIVTAT  
NFCAPNYGFTAEGGGHCNPPNKHVFLPIEAFEKIAIWKAGNMPVQYRRIKCRKEGGV  
RFTVDGSGIFVSVLISNVAGAGDVVAVKIKGSKTGWLQMGRNWGQNWHINSDLKN  
QPLSFEVTTSDGLTLTSYNVAPKDOWNFGQTFEGKQFQS\*

### CDS (coding sequence)

>TcEXPA-07

ATGTCACCGCCAGCAAAGACTGAAGCTCCTCTCCTTGCACCTTCTGCTTCTTACACT  
AACCTCGACAGCCACTTCCCATTACACCTCTTCGTCCCCACCCGCGCCTCAGTCCT  
CCCAGTGGCATCCCGCACGCGCCACTTACTATGCAGCATCGGATCCTCGGGACGC  
GGTGGGTGGCGCGTGTGGGTATGGGGACCTGGTGAAGGCTGGCTACGGCATGGC  
CACGGTGGGTTTGAGCGAGTCCCTGTTTGAGCGTGGGCAGATCTGCGGCGCGTGC  
TTCGAGCTGAGGTGCGTGGATGATCTGCGCTGGTGCATACCCGGGACCTCTATCA  
TTGTGACAGCCACCAATTTCTGTGCTCCCAATTATGGGTTTACAGCGGAGGGTGG  
AGGGCATTGTAATCCTCCCAACAAGCACTTTGTGCTTCCCATTGAGGCTTTTGAGA  
AGATCGCCATTTGGAAGGCTGGCAACATGCCCGTTCAGTATCGAAGGATCAAGTG  
CAGAAAAGAAGGAGGTGTTTCGATTTACAGTTGATGGTTCAGGTATCTTCGTTTCA  
GTGCTCATCAGCAATGTTGCTGGTGTGAGATGTAGTTGCAGTTAAGATTAAGG  
GTTCAAAAAGTGGGTGGCTGCAAATGGGTTCGGAATTGGGGACAAAATTGGCATA  
TTAATTCTGATCTAAAGAATCAGCCTCTTTCGTTTGAGGTCACCACTAGTGATGGG  
CTTACGCTTACATCTTACAATGTTGCTCCTAAGGATTGGAACCTTGGGCAGACTTT  
TGAAGGAAAGCAATTCCAATCATAA

### Nucleotide

>TcEXPA-07

CTTAAAATGAATGAATGAACTTGATTGAAGACGCGGATTAACCTTTTCGCTTTCCA  
CTTTTGTCTCTCGTCTCTCTCAAGTGTAAGGAGTTGTTGTTTCAGGTTAAGAAATGT  
CACCGCCAGCAAAGACTGAAGCTCCTCTCCTTGCACCTTCTGCTTCTTACACTAACC  
TCGACAGCCACTTCCCATTACACCTCTTCGTCCCCACCCGCGCCTCAGTCCTCCCA  
GTGGCATCCCGCACGCGCCACTTACTATGCAGCATCGGATCCTCGGGACGCGGTG

GGTGGCGCGTGTGGGTATGGGGACCTGGTGAAGGCTGGCTACGGCATGGCCACG  
GTGGGTTTGAGCGAGTCCCTGTTTGAGCGTGGGCAGATCTGCGGGCGCGTGCTTCG  
AGCTGAGGTGCGTGGATGATCTGCGCTGGTGCATACCCGGGACCTCTATCATTGT  
GACAGCCACCAATTTCTGTGCTCCCAATTATGGGTTTACAGCGGAGGGTGGAGGG  
CATTGTAATCCTCCCAACAAGCACTTTGTGCTTCCCATTGAGGCTTTTGAGAAGAT  
CGCCATTTGGAAGGCTGGCAACATGCCCGTTCAGTATCGAAGGTACAAGCAGTTT  
TAAGTTGCTTTGCTTAACATTGTACTACTTGATATTA ACTTATCATTGGCAGCTAA  
GAATCGCTAGTAATTTGTTATATGTAGCTGCAGATATTTCCCTCGATTTTATACCTA  
TAAAGGAAAAATTGATGGACAAAAGCTAGGAATTGACTATCAATACGTTAATGC  
TTATTTGTATTAGTTAATGTTACCGTCATCAAGTCAGGTTAGGAAAATCTAGATTA  
GATCGGTTTTATTTGCATCAAATCAGGAGCCAAAGCTGGAATTTAGGCAAATGTT  
TATAGTGACACAAGGTTTCATGAACCCAAACAGCTGGGAATTAAGTTGGATTTATT  
GAGATAAACTGAAAGCTATGATCTTACGGACATAACTCCTAACAAGTATTTTCTC  
TTGCTTGGCTTAAACAGGATCAAGTGCAGAAAAGAAGGAGGTGTTTCGATTTACAG  
TTGATGGTTCAGGTATCTTCGTTTCAGTGCTCATCAGCAATGTTGCTGGTGCTGGA  
GATGTAGTTGCAGTTAAGATTAAGGGTTCAAAA ACTGGGTGGCTGCAAATGGGTC  
GGAATTGGGGACAAAATTGGCATATTAATTCTGATCTAAAGAATCAGCCTCTTTC  
GTTTGAGGTCACCACTAGTGATGGGCTTACGCTTACATCTTACAATGTTGCTCCTA  
AGGATTGGA ACTTTGGGCAGACTTTTGAAGGAAAGCAATCCAATCATAAGAGAT  
GTCTTTATATATACACAAGACTATTCTTTTTTCTCCTCTTTCTTTCTTCCCTTCTGTA  
ATCTTTTGTTTCTACTACTGATTGATTGGGTTGGAAAGTGTA AAAAGTGGTGCTTT  
TTGATTTAGGAGCAAGTAGTAAGCTGATATTCAAGTGATGACATTGTTGTGTGCA  
TAGCTCCATTGTCAAATTAGGATAGATATCAGGTTTGTCTCTTACACAATAATACC  
TAAAAATGTGAATTTAAAACAGATCCCTCCAATGGCATGGCCCAAGTGGTTTCAT  
TTAGGCCATCTCAAGGTAATAAGGCAAATGAATGTGTTTATTGATTGGAGCACAC  
ACAAGTAACAAGGCATTTGTAGGGGTGAACA ACTGCTACCACATGATTCTGGGCT  
TATGGTTACGGATTCCGGATGAAGCTAGCTATTCTTCTGCAATTGGTCGATGGTATA  
TCAACTAGCATGTACATTAACGAGAAGGAAGGTGGAAGCACAATATTGTTTCTAT  
TTTTCTTTTCTATTTTTTTTTTTCATTTTCGTAGGATGAAAATAAGAGAGAGAATGAA  
AATGGATAGATTTGCACTGTTTTCTTTCAAATT