

## IDENTIFICATION

**Species:** *Theobroma cacao*

**Locus:** Thecc.09G320300

**Gene Model:** Thecc.09G320300.1.p

**Description:** TcEXPA-16

**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)

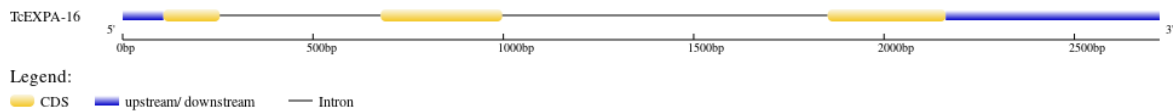
KEGG: <https://www.genome.jp/entry/gn:T02994>

## EXTERNAL RESOURCES

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

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

## GENE STRUCTURE



## DOMAIN ARCHITECTURE

Query seq. Superfamilies

PLN00050

Name	Accession	Description	Interval	E-value
PLN00050 super family	cl31535	expansin A; Provisional	29-256	8.30e-123

## SEQUENCES

### Peptide

>TcEXPA-16

MAVVSVVCVLLFISSMWMAEAKIPGVYSSGAWQNAHATFYGGSDASGTMGGACG  
YGNLYSQGYGVNTAALSTALFNGLSCGACFEIKCANDPKWCHSGGPSIFITATNFCP  
PNYALPNDNGGWCNPPRPHFDLAMPMLKIAEYRAGIVPVSYRRVPCRKRGGIRFTI  
NGFRYFNLVLISNVAGAGDIVKVSVKGSRTGWMSMSRNWQNWQSNVAVLVGQSL  
FRVTGGDRRTSTSWNIVPANWQFGQTFAGKNFRV\*

### CDS (coding sequence)

>TcEXPA-16

ATGGCTGTGGTGAGTGTAGTATGCGTTTTACTCTTTATCTCCTCAATGTGGATGGC  
TGAAGCAAAAATTCCAGGAGTCTACTCTGGTGGAGCATGGCAAAATGCCCATGCT  
ACCTTCTATGGAGGTTCTGATGCTTCTGGCACCATGGGAGGAGCTTGTGGCTATG  
GAAATTTATACAGCCAAGGGTATGGAGTTAACACGGCAGCATTGAGCACAGCAT  
TGTTCAACAATGGCTTAAGCTGCGGTGCTTGCTTTGAGATCAAGTGCGCAAATGA  
TCCAAAGTGGTGCCATTCAGGCGGCCCTTCCATCTTCATCACAGCAACCACTTCT  
GCCACCAAACATGCTCTTCCCAACGACAATGGTGGCTGGTGCAACCCTCCTCG  
CCCTCACTTTGACCTTGCCATGCCCATGTTCTTAAAATTGCAGAGTACCGTGCCG  
GTATCGTTCCCGTTTCTACCGCCGGGTGCCATGCCGAAAGAGGGGAGGGATCAG  
GTTACGATCAACGGCTTCCGTTACTTCAACTTGGTCTTGATCTCCAACGTTGCGG  
GTGCAGGGGATATCGTGAAGGTTAGCGTGAAAGGATCCAGGACCGGTTGGATGA  
GCATGAGCCGAAACTGGGGTCAAACTGGCAGTCAAACGCGGTTCTCGTCGGTC  
AGTCACTCTCCTTCAGGGTCACAGGCGGTGACAGGCGCACTTCCACCTCGTGGAA  
CATCGTCCCCGCCAATTGGCAGTTCGGCCAAACATTCGCTGGAAAGAATTCAG  
GTTTGA

### Nucleotide

>TcEXPA-16

ACCCCCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTGAAGAAAGAGTTGAGTTT  
TCAGGCTTTTTTTAAGTCAATAGTCTTTGAAGTAGTTGGAGAAGGGAAAAATGGC  
TGTGGTGAGTGTAGTATGCGTTTTACTCTTTATCTCCTCAATGTGGATGGCTGAAG  
CAAAAATTCCAGGAGTCTACTCTGGTGGAGCATGGCAAAATGCCCATGCTACCTT  
CTATGGAGGTTCTGATGCTTCTGGCACCATGGGTATGTACCTTCATCAAGGCCTAT

TTCTGCCTCATTTTCATCCCCATTTTCGGATTATGCCTCAAAGCTTCTTTCTTCCCC  
CCGCGCTCTCTTAATCTTCTTCTTTTCACAGTTATTTTTGACTACCCAGCTGCATT  
ATTTCTTACAGAATTTGCCATATCAGAAGTAGCATAAAAAGATAACATTTGGGC  
TTTCATGTCTGACAAATCCTCCCTAATTCTACAAAAAGAAAACCAACTCTCTAGC  
CTGCATTAAAGGCAAAGCTCTGAAATTTCCCTCATTGTCAAACCCTAACTCCTC  
ATGTCACATTTCTACTAAAAAACCACTTAAAACCGCCAACAAGTTAAGCACTAA  
CTCAACTGCATTACCTTAATTCCCCACTAATTATCTTTTTTTTTTCTCTCTCTAAT  
CCTTGCAGGAGGAGCTTGTGGCTATGGAAATTTATACAGCCAAGGGTATGGAGTT  
AACACGGCAGCATTGAGCACAGCATTGTTCAACAATGGCTTAAGCTGCGGTGCTT  
GCTTTGAGATCAAGTGCGCAAATGATCCAAAGTGGTGCCATTCAGGCGGCCCTTC  
CATCTTCATCACAGCAACCAACTTCTGCCACCAAACTATGCTCTTCCCAACGAC  
AATGGTGGCTGGTGCAACCCTCCTCGCCCTCACTTTGACCTTGCCATGCCCATGTT  
CCTTAAAATTGCAGAGTACCGTGCCGGTATCGTTCCCGTTTCCCTACCGCCGGTATG  
TGCTCTGTCTCCCGCCCTTTTCCGTACTCTGTTTCCCTTTTAATGCTACCATGTAGTG  
TGTTGAAGACAGGGCTAGACAGTCAAAGTGTGATTTTCATAAACTTTTAAGAGT  
CAAAATCAATATAATAATAATGCTAGAAATGATATGATCTAACTCTTTTCTTGCG  
GGGGTTAGGTA ACTCTATTTAATTTCAATTGGTTAGGGAATTGATCTCTATTACATG  
TGACTAATATTGTTAGATGTGGGTCAAGTTTCTGCATATGGGTTGGATTCTGGTGA  
GGTATGTGTCAGAGGAGTTACTATCTTGGGTCATTATGAGTTTTACGAAAATACC  
CATATTCATTTTCTCATTGCTCACCTACCCCAATCCTCTAGTCTAAGTTGTACCT  
AACTGACATTCAAAGCTTCGTTAAGAGCCGTTTGTGTGTTTTGCTCCATGTAGCT  
GATCTTAGCTTTGGCCAAAATATTTCTAGTCACTTTCCATCCTCTGTTCTCACTCGT  
GTTGTCTTTTTACGCTGTAGCTAAGGTACCAAATAGCCCAAAGCCAAAACAAAAC  
CAGACATTGGAATAACATTGTACGTTGTAGTAGGAATAGGTTGGGTTGTAAACCA  
ATTTCCATTAGTTCTACTTCTAGTAATGGACTAATCGTCAAACATCAAACCCAA  
GACACACACGACATCACAGTCCCCAAGGTCCTAAGACTCTTTTATTTTAATAATA  
ATAATAAAAAAAGGATTA AAAAGGCACTAAAACCCTAGCTGCTGTCTTGTGACCA  
AATCTTTGTTCCACGCCTAGAATCACGCTCCGCGTTTCATGCTAACATTTGCTTCG  
TCCTTAATTACGCAGGGTGCCATGCCGAAAGAGGGGAGGGATCAGGTTACGAT  
CAACGGCTTCCGTTACTTCAACTTGGTCTTGATCTCCAACGTTGCGGGTGCAGGG  
GATATCGTGAAGGTTAGCGTGAAAGGATCCAGGACCGGTTGGATGAGCATGAGC  
CGAAACTGGGGTCAAACCTGGCAGTCAAACGCGGTTCTCGTCCGGTCAGTCACTCT  
CCTTCAGGGTCACAGGCGGTGACAGGCGCACTTCCACCTCGTGGAACATCGTCCC  
CGCCAATTGGCAGTTCGGCCAACATTCGCTGGAAAGAATTTAGAGTTTGATTG  
TTCAAATTTGAAAACCAAACAACGAAAAAAGAAAAAACCAAACAAGGAACG  
AAACTATTAACCCTCCATAAGTTAAAAAACCAAAGTTAGCAGCAGCAAAAATGA  
AGTTTCCCGCCATAATCCATTTCCCATTTTACCGTCATTGGCGCCAACTTTTCTT  
CGTTAGGGTGATGTTTCTTTTGGACCAATTGGGTTTCGGGTTGGGGCTGCTGTCAC  
TCATTTAGGTTAAGGGGAAAGAAAAGTGAAGATTTTCATGGATTTTTCTCTCTCT  
TTTTGTTCTGACTTTTGGATTGGTGGGTTTGTTTTTATTGTGTGAAGGGCTTCTTT  
TTTTCATCTTTTTGTTGTTTTAGCGTCAAGAGAGGAAGGGAAAGAAGAAGTGTAG  
GGGAAGGACTGAGGCGGCTGCACACAAAAGGAAATGTAGCCCGCAGCCTAATA  
TAAATTTAGTGTCACTAGCTATAGTATGATTATATATATATATATTTGGGAAGTGT  
ATTGTTGTAACCTTATATATAATTGAATTGGAATTTTATATTTGGTTTGTCTCTT  
ATGC