

## IDENTIFICATION

**Species:** *Theobroma cacao*

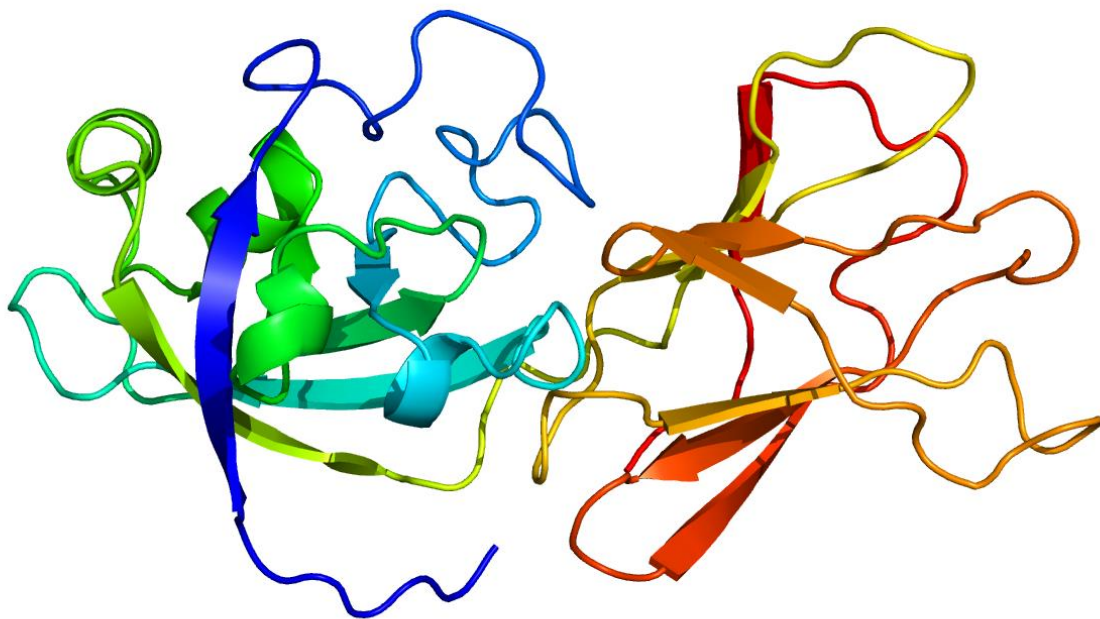
**Locus:** Thecc.01G384900

**Gene Model:** Thecc.01G384900.1.p

**Description:** TcEXPB-01

**Family:** Beta 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

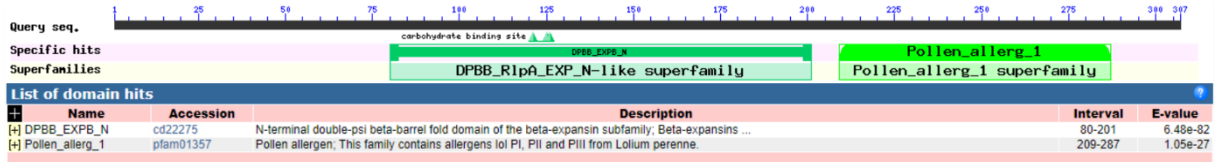
<https://www.cacaogenomedb.org/>

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

## GENE STRUCTURE



## DOMAIN ARCHITECTURE



## SEQUENCES

### Peptide

>TcEXPB-01

MEGSSGFTRSIGWGMLNGINHHQRQNQRSLIEVFMAFALRNMLSFFTFAATFSLLF  
NSCYCFYPKLLNASLAASESDWSPAGATWYGSPTGAGSDGGACGYGGAVEQAPFSA  
MVSAGGPSLFKSGKGCACGYEVQCTTNSACSENPTVTVITDECPGCVSESVHFDLSG  
TAFGAMAKSGQADQLRNAGVLQIQYRRVQCNYPGTSIAFKVDAGSNPYFATLVIY  
QNGDGDALVELKQALNSDPWLPMQQSWGAVWKLDSGSTLRPPFSIRLTSLSGQT  
LVATGVIPAGWQPGQTYTSVVF\*

### CDS (coding sequence)

>TcEXPB-01

ATGGAAGGATCCTCAGGGTTTACTCGGTCCATCGGATGGGGAATGTAAACGGCA  
TTAACCACCATCAACGGCAAATCAACGAACACTTTCTTTGATCGAAGTGTTCAT  
GGCTTTCGCTCTTCGAAACATGCTCTCTTTTTTCACTTTTTCGCGCCACATTTTCTCT  
GTTATTCAACTCCTGTTACTGCTTCTACCCTAAGCTCCTTAATGCCTCCTTGGCAG  
CATCTGAGTCCGATTGGTCCCCAGCCGGAGCAACGTGGTATGGCAGCCCCACGGG  
GGCTGGAAGTGATGGTGGAGCTTGTGGTTATGGGGGTGCTGTGGAGCAAGCCCC  
ATTTTCGGCGATGGTGTTCGGCTGGGGGCCCTTCCCTATTCAAATCTGGCAAAGGA  
TGTGGAGCCTGTTACGAGGTCCAGTGCACGACAAATTCAGCGTGCTCGGAAAATC  
CTGTGACTGTAGTTATCACCGATGAATGCCCTGGCTGTGTCTCGGAGTCCGTTTCAT  
TTTGACTTAAGTGGTACTGCTTTTGGTGCCATGGCAAAGTCTGGCCAGGCTGATC  
AACTCCGTAACGCTGGAGTCTTGCAGATTCAATATAGAAGAGTGCAATGTAACATA  
TCCCGGGACGTCCATTGCCTTCAAAGTTGATGCCGGCTCCAACCCTTACTATTTTCG  
CCACCCTAGTTATATACCAAATGGAGATGGTGACCTTGCCTTAGTTGAACTCAA  
ACAGGCGCTTAACCTTGACCCCTGGCTTCCCATGCAACAATCCTGGGGTGCAGTC  
TGAAACTAGACAGCGGCTCGACGTTGCCGGCTCCGTTCTCCATCAGGCTGACAT  
CACTACCTCTGGCCAGACTCTTGTGGCAACCGGTGTCATTCCTGCTGGTTGGCAG  
CCCGGACAAACCTATACATCAGTCGTTAATTTTAA

## Nucleotide

>TcEXPB-01

GGTGTAAAAAAAAAAAAAGCATTAAATCGCAATCATACAATCTCATCAATCTTATT  
TATTAATAAACGATTTGAATATTACAGAGCCCAAATCCCCGCCAAATATTACGTG  
CTGTAGTTTTTAAACGGAAATGGAAGGATCCTCAGGGTTTACTCGGTCCATCGGA  
TGGGGAATGTTAAACGGCATTAAACCACCATCAACGGCAAATCAACGAACACGT  
ATCCCTTGAATTTGGGGCAAGAAACATACGAAAATATTAATTTGGAGCAGTCGA  
TGACGTGGGCCCCGATGTTAACCTTCTACCGGGGGCGGGTTCGGAGACATAACCGC  
TTTGTGGTCGATTGCAAACGTAGACTGTCACAAACAGACAAAAGAGAGATGTTG  
GTAGTTAGGTTTCGTCTGAAAACGTACGTCTGAGTCAGAGAAACAGAGATAAAA  
AGAGCAAATCAGATCTTGCGTTTTGTTTTTTTCTTCTCTCCCGGCTCCGGCAACGA  
AGCCTGGAAACAAGCAATTTGACTGATGTAACCATGCAAGAGTATTACTGTGATT  
TGGACAGAACA AAAAGTTCTTTTATTTTCAACTGTGTGTGCATGGTGATTGAAATTA  
TTAGGCCATGGAACGTTTGGTTGATTCTCAAATGAATTGTATTTTTGTTCAATTAT  
AGAGAAGTAGACGGATTTTCTGATTCTCTGACAAATTTTGCCTGGACAACAGGGA  
GTTGATTTTTCTGTTCTTGATTCTGGTTTCTGGTCTCTCTGGGGTTTTGTGTTGTGT  
TACTCTTTGGCATTTTGTA CTTCGATTTCAATCAATTTAATTATGTGCTTACGAAT  
GTAATTAGACTTTTCTAGCTGATGTCATAGAAATGTATGGGAATTTAGCCAAATT  
CTAACTTAAAGTCTATTAAGTAGTATAGGTTATTAGGATTTTCGTATCTCAAGGAT  
CTTACTCATTTCGCTTGTTCCTGAAGAAAATAATGCATAAATCCTTCTAAAAACGT  
GCACGTGTCCTTCACTAAGTGGTTGCATTAGAACCACATGCACTAGCCTTTGGCTC  
TATAAGTACCTCGTCGTTGATCTGCTAAACCCATCAGCAAACAGAGCTTAAACT  
TTGCATTTAGCTAGCTTTGCTAATTCCTTTCTACTAGTAGTTTCTTTGATCGAAGTG  
TTCATGGCTTTTCGCTCTTCGAAACATGCTCTCTTTTTTCACTTTTTCGGGCCACATTT  
TCTCTGTTATTCAACTCCTGTTACTGCTTCTACCCTAAGCTCCTTAATGCCTCCTTG  
GCAGCATCTGAGTCCGATTGGTCCCCAGCCGGAGCAACGTGGTATGGCAGCCCCA  
CGGGGGCTGGAAGTGATGGTAACGATGAGCATGGAGAATGATTGAGACTATTTA  
CACATAGAAAATTGTTAATTTTATGAAATTTAACCCTTTTTTTTTTTTGTATCTAG  
GTGGAGCTTGTGGTTATGGGGGTGCTGTGGAGCAAGCCCCATTTTCGGCGATGGT  
GTCGGCTGGGGGCCCTTCCCTATTCAAATCTGGCAAAGGATGTGGAGCCTGTTAC  
GAGGTGACATACAATTTTATACATATATATTCATCAGTTTGTATCATGCCTCATT  
CCCATGTCCAAAACCTCCATAACTTTTCTGTCAAGTTTGAAAGATTGATAAACTTCC  
TTTGTGTTTATGATGATTCTCCATATACATATACTTCTTTTTTCTACTCTCGCATG  
CAAAGGGTACCTTTCAATGTTGTTTAAACATGCATAATATGGTTTAAATTCAGGTCC  
AGTGCACGACAAATTCAGCGTGCTCGGAAAATCCTGTGACTGTAGTTATCACCGA  
TGAATGCCCTGGCTGTGTCTCGGAGTCCGTTCAATTTGACTTAAGTGGTACTGCTT  
TTGGTGCCATGGCAAAGTCTGGCCAGGCTGATCAACTCCGTAACGCTGGAGTCTT  
GCAGATTCAATATAGAAGGTAAGATTCTACGTACGATTAGAAAGAGAGGATGAG  
AACAGAGAAGTAGCATGACAATATTCTTGTACAATTTTTCCATGAAGAATTTAAC  
ATATCCTCAGCCTGCAAAAAAAGGTGAAGAAGAATTGACTTTGTAAATTGCACTA  
TAGGTGGTATTTACTCTTGGGAATTGACTTTGTAAACCAAGCCTACTGCCTATGG  
GATTAGATTATTCTTCAAATAATTTTAAACCTTCATTTTCGCTGTCAAATTTCTAAC  
GTTTTCTAACCCTCATTTCATTGACGAAAATTTGGCTTACAGAGTGCAATGTAA  
CTATCCCGGGACGTCCATTGCCTTCAAAGTTGATGCCGGCTCCAACCCTTACTATT  
TCGCCACCCTAGTTATATACCAAATGGAGATGGTGACCTTGCCTTAGTTGAACT  
CAAACAGGCGCTTAACTCTGACCCCTGGCTTCCCATGCAACAATCCTGGGGTGCA

GTCTGGAAACTAGACAGCGGCTCGACGTTGCGGCCTCCGTTCTCCATCAGGCTGA  
CATCACTCACCTCTGGCCAGACTCTTGTGGCAACCGGTGTCATTCCTGCTGGTTGG  
CAGCCCGGACAAACCTATACATCAGTCGTTAATTTTTAATGCCTAGTAAGGCACC  
CTAAACGTAGCGATCTCTATACGATATAATGTATTCGATTCTGTACCACGTTTATA  
ACTGTCCTGTTAGTCAACCTGAAGATATTCTTCCATGTCTTAATTGCTTGATTATG  
TATGTTTATGGGAGAATTATCAATAAAAATACAGTTTATAGC