

IDENTIFICATION

Species: *Theobroma cacao*

Locus: Thecc.09G363200

Gene Model: Thecc.09G363200.1.p

Description: TcEXPA-17

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

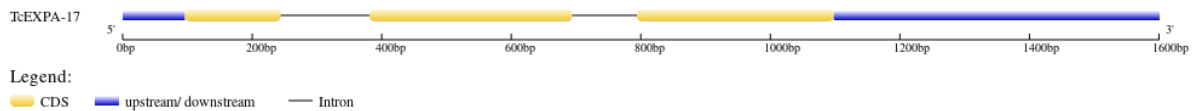
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

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

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

GENE STRUCTURE



DOMAIN ARCHITECTURE

Query seq. Superfamilies

PLN00050

Name	Accession	Description	Interval	E-value
PLN00050 super family	c131535	expansin A; Provisional	22-254	4.42e-129

SEQUENCES

Peptide

>TcEXPA-17

MIAQTFSSAMLLSIFLFTSLNLHGITGDYGGWQNAHATFYGGGDASGTMGGACGYG
NLYSQGYGTKAALSTALFNGLSCGSCYEMRCNDDPKWCLSGTITVTATNFCPPNF
ALSNDNGGWCNPLQHFDLAEPFLQIAQYRAGIVPVVFRRVPCVKKGGIRFTINGHS
YFNLVLITNVGGEGDVHAVSIKGSKTGWQPMSRNWQNWQSN SYLNGQGLSFQVT
TSDGQTVTSYNVVPADWQFGQTFEGGQF*

CDS (coding sequence)

>TcEXPA-17

ATGATTGCTCAGACATTTTCCTCAGCTATGCTTCTCTCTATTTTCCTGTTTACTAGC
TTAAATTTGCATGGAATTACTGGAGACTATGGCGGCTGGCAGAATGCGCATGCTA
CTTTCTATGGCGGTGGTGATGCCTCTGGCACAATGGGAGGAGCGTGTGGCTACGG
GAACTTGTACAGTCAAGGGTATGGAACAAAGACTGCAGCACTAAGCACAGCCCT
GTTCAACAACGGCCTCAGCTGTGGATCTTGCTATGAGATGCGATGCAATGATGAT
CCCAAATGGTGCCTTTCTGGCACCATCACTGTCACTGCTACAAACTTCTGCCCCCC
TAACTTTGCTTTGTCTAATGACAATGGCGGGTGGTGCAATCCTCCTCTCCAGCATT
TTGATTTGGCAGAGCCCGCCTTCTGCAAATTGCCCAATACCGAGCTGGAATCGT
ACCTGTTGTCTTTAGAAGGGTGCCATGTGTGAAGAAAGGAGGAATAAGATTCACC
ATTAATGGACACTCTTACTTCAACCTGGTCTTGATCACCAACGTTGGGGGTGAGG
GAGATGTTTCATGCAGTGTCCATCAAAGGGTCCAAAACAGGGTGGCAACCAATGT
CTAGAAACTGGGGCCGAACTGGCAGAGCAACTCCTACCTCAATGGCCAAGGCC
TGTCTTTCCAAGTCACCACCAGCGATGGTCAAACCTGTGACAAGCTACAACGTGGT
GCCTGCTGATTGGCAGTTTGGTCAGACTTTTGAAGGAGGTCAATTTTAG

Nucleotide

>TcEXPA-17

ACACACAACACAAAGCAAATCATATTCCCCACCTGTCCGTCATCCAACAGCTCAG
ATCCTGATCTTTTTCTGTAGCATTATTAATTATTAGTTAAAATGATTGCTCAGACA
TTTTCTCAGCTATGCTTCTCTCTATTTTCCTGTTTACTAGCTTAAATTTGCATGGA
ATTACTGGAGACTATGGCGGCTGGCAGAATGCGCATGCTACTTTCTATGGCGGTG
GTGATGCCTCTGGCACAATGGGTGAGTGCTGCACTACAACCTCAACTTCTGACCCT
TACCTCTTTCTGTCTTTAGAGTCTTCAACATTTAATCTTTAATTTGCCTGAGAATT

GGTGGTAATTTGTGAGAGGTTTTGTTATGTGAACATTGATTTCCAGGAGGAGCGT
GTGGCTACGGGAACTTGTACAGTCAAGGGTATGGAACAAAGACTGCAGCACTAA
GCACAGCCCTGTTCAACAACGGCCTCAGCTGTGGATCTTGCTATGAGATGCGATG
CAATGATGATCCCAAATGGTGCCTTTCTGGCACCATCACTGTCACTGCTACAAAC
TTCTGCCCCCTAACTTTGCTTTGTCTAATGACAATGGCGGGTGGTGAATCCTCC
TCTCCAGCATTTTGATTTGGCAGAGCCCGCCTTCCTGCAAATTGCCCAATACCGAG
CTGGAATCGTACCTGTTGTCTTTAGAAGGTAAGGGCATCATTAAAACATAACACC
AATGCTTGACATTGAACAATCACGACTGTCTTAATTTAGATTCTGATTGATGTACT
TTGCGTCGGTTTGGCAGGGTGCCATGTGTGAAGAAAGGAGGAATAAGATTACACC
ATTAATGGACACTCTTACTTCAACCTGGTCTTGATCACCAACGTTGGGGGTGAGG
GAGATGTTTCATGCAGTGTCCATCAAAGGGTCCAAAACAGGGTGGCAACCAATGT
CTAGAAACTGGGGCCAGAACTGGCAGAGCAACTCCTACCTCAATGGCCAAGGCC
TGTCTTTCCAAGTCACCACCAGCGATGGTCAAACCTGTGACAAGCTACAACGTGGT
GCCTGCTGATTGGCAGTTTGGTCAGACTTTTGAAGGAGGTCAATTTTAGAATTTGT
CTTCACTCTCCAATCCCAAGATAGATTAAGCTTATGTAACCTTGCATTGTAACAGTG
TTGTGATTTGTAAGAGTCTGGAAAATGCTGAGGTGACTAAAAATGCGCCCGCTAG
AGGCTAGACCTTTCCACAATTGTGGCACTCATAATCAAAGTTGACGAACCTATTC
TCCGTTTTCAACTTTTACCATGGAAATTTGGGCTACTAGAACAGGTGGGGCCATG
GGATTCACCCAGAGCCCAAAGGTTCAACACTTTTTGCAATGATTCTGAATCCAGA
GCCGGGGCCTCTCTGCATCGATAAAACACTAAAATCATACTAAACGAATCAAAAT
CAATCATAGGCTTTTATGTGGCATTACTCACAATCTTGAGGCCGAAAAGCAGAGG
CTCTAGTTCAATTTATGCTGAAGCTTTTTGGTCTTAAGTAGCAGGGCCCTTCGTTG
TTTTTCAACTACTTTTTGGGTAATGGAACCTGTGTGGAATGCCGTGGTAATATT