

IDENTIFICATION

Species: *Theobroma cacao*

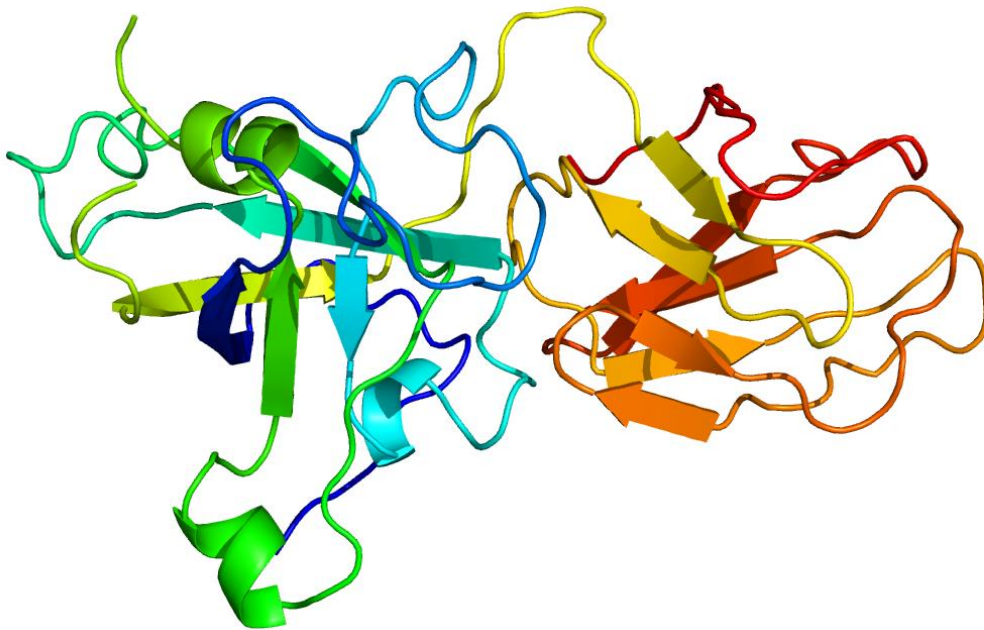
Locus: Thecc.07G082500

Gene Model: Thecc.07G082500.1.p

Description: TcEXPA-14

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

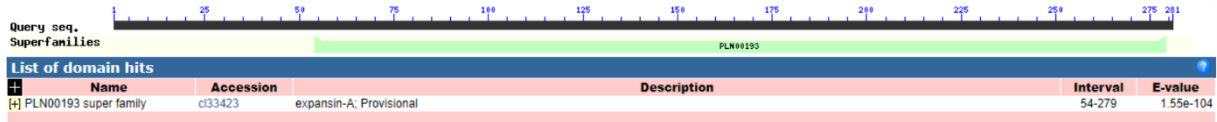
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-14

MEVLASRNHVCQSLLFICLERRRRMGPVIIPSVLLWHLFFFGSTILEGVNAQVNGWVN
AHATFYGADQSPFSLGGACGYDNTFHAGFGANTAALSAALFREGEACGACYQLRCD
YRADPKWCLRRASVTITATNFCPSNNNGGWCNPPHHHFDMSMPAFLRIAWQNEGI
VPVLYRRVPCRRRGGVRFLLKGGSSFNLMISNVGGSGNVKAAWIRGSRTTRTWLPM
HRNWGVNWQSSVDLRNQRLSFKLTLVDGKTLEFLNVVPSTWKFGQTFASPRQFY*

CDS (coding sequence)

>TcEXPA-14

ATGGAGGTTCTTGCTTCAAGGAATCATGTGTGTCAAAGTTTGTGTTTCATTTGTTT
GGAGAGAAGAAGAAGAATGGGGCCTGTCATCATTCCCTTCTGTTTTGCTTTGGCAC
TTGTTTTTCTTTGGCAGTACAATTCTTGAAGGAGTCAATGCTCAAGTTAATGGTTG
GGTTAATGCTCATGCAACTTTCTATGGGGCTGACCAAAGCCCCTTCTCTCTTGGAG
GAGCTTGTGGATATGATAACACCTTCCATGCTGGATTTGGAGCTAACACAGCCGC
GTTGAGCGCTGCACTTTTTAGAGAGGGTGAGGCTTGTGGTGCTTGTACCAACTA
AGGTGCGATTATAGGGCTGACCCCAAGTGGTGCCTCCGGCGTGCTAGTGTGACCA
TAACCGCCACCAACTTTTGCCCTTCCAACAACAATGGTGGGTGGTGCAATCCCCC
TCACCATCATTTTGACATGTCGATGCCTGCCTTCCTTCGCATTGCCTGGCAAGGCA
ACGAAGGCATTGTTCCGGTTCCTTATAGGAGGGTTCATGCAGGAGGAGAGGTGG
AGTCCGGTTCACCTTGAAAGGGCAATCAAGTTTCAACCTGGTGATGATCTCCAAC
GTTGGTGGCAGCGGCAATGTGAAGGCGGCATGGATAAGGGGCTCGAGAACAAGA
ACATGGCTACCTATGCATAGGAACTGGGGTGTCAATTGGCAAAGCAGCGTTGATC
TCAGGAACCAAAGACTGTCTTTTAAGCTGACTCTGGTCGATGGCAAACGTTGGA
GTTTCTCAATGTTGTTCTTCTACGTGGAAGTTTGGACAAACATTTGCTTCGCCAC
GACAGTTCTACTAG

Nucleotide

>TcEXPA-14

ACAGGACGTGCAAACCGCGTTATTTTTTCATATAGTTTCTTTCTTTTTTTTCTACCCC
ATTTTTAACCAATTCCTTTGAAGATAATGCAATTGAGAAAAAATTAATTAATTA
AATTCTAGATCACCGACAAACATCGTTTGGAGTCTATAAAATGGAGGTTCTTGCT
TCAAGGAATCATGTGTGTCAAAGTTTGTGTTTCATTTGTTTGGAGAGAAGAAGAA

GAATGGGGCCTGTCATCATTCCCTTCTGTTTTGCTTTGGCACTTGTTTTCTTTGGCA
GTACAATTCTTGAAGGAGTCAATGCTCAAGTTAATGGTTGGGTTAATGCTCATGC
AACTTTCTATGGGGCTGACCAAAGCCCCTTCTCTCTTGGTAAGCCACACTTTCTCT
TTAGTGTGCCCGCCTAATTATTAAGTGTTCGAATCATGCACACATTTGATTTGG
TTGCATGATCATGCATGCCATGAGCCTATAGGGTTTCTTCAACTTTGGTCTTTACA
ATTTATTTTATTCGATATTGATATGTTATTATAGGAGGAGCTTGTGGATATGATAA
CACCTTCCATGCTGGATTTGGAGCTAACACAGCCGCGTTGAGCGCTGCACTTTTTA
GAGAGGGTGAGGCTTGTGGTGCTTGCTACCAACTAAGGTGCGATTATAGGGCTGA
CCCCAAGTGGTGCCTCCGGCGTGCTAGTGTGACCATAACCGCCACCAACTTTTGC
CCTTCCAACAACAATGGTGGGTGGTGCAATCCCCCTCACCATCATTTTGACATGTC
GATGCCTGCCTTCCTTCGCATTGCCTGGCAAGGCAACGAAGGCATTGTTCCGGTT
CTTTATAGGAGGTAATTTTCAGTTTGTTTGACGAGTAGGAATTAATAGTTATTAAG
TCAGGTGATTAATAATCTCAAAGTTTAAGTTGATAAGAAAATGTCTGATATATT
ATATTAATAATTTAACAGTCTCTCACATGCAAATTAGATGAATGATTTATGAGTTT
TCACGTGAATTCAATTGCTTTGATACTATATTAATTAGGTAGTCCGTTTTTTAAA
AAAGTTAAGTAGATTAGAAGAACTCAACTTTATATTTTATTGTAACAATTGATTTT
ATTATGATGATAGGTGATTTCCCTAACTATGAATATATATATATATATATATATA
TATATATATATATATATATGTTACTAAAATATAAGAGATATTGCCAACAATTATGT
AAAATACCAATCATTTTCATGCAAATCTTACAAATTTACACTTTAATCAAGTGGTT
AATTTGATTATCATAATAAATGGGTGGTTTAGGATTCAATAGGGAAGTATATA
TTCTTTACATGACAAGTTTTTGTTCCTTTGTTTTGAAGGGTTCCATGCAGGAGGAG
AGGTGGAGTCCGGTTCACCTTGAAAGGGCAATCAAGTTTCAACCTGGTGATGATC
TCCAACGTTGGTGGCAGCGGCAATGTGAAGGCGGCATGGATAAGGGGCTCGAGA
ACAAGAACATGGCTACCTATGCATAGGAACTGGGGTGTCAATTGGCAAAGCAGC
GTTGATCTCAGGAACCAAAGACTGTCTTTTAAGCTGACTCTGGTCGATGGCAAAA
CGTTGGAGTTTCTCAATGTTGTTCTTCTACGTGGAAGTTTGGACAAACATTTGCT
TCGCCACGACAGTTCTACTAGCTGCCTAAAGAAATGTCCTAGAAACCATTACTTT
TTACATTGATGGTTTTAGGCTAAAAATCATAATGACTTAAAACATGTCATTGAAG
CTTGTCTTTATCTTATTGTACTTTAAACAAGCAGGCATGGCTCCAAAAAATTGGGAG
CAATATGCTTCATTTGTAATTTTTTTCTCCACTTTTCGTCATATATTGGGGGGAC
AGATTGTATTTTTTATTACCAAACATTCTAAATAAAAAAGATTAGCAAAGGGCTG
TTAAAAATAGTCCTTTAACTTCTTTATAAGCTCTCCAATCTTTGCTATATTCAACTC
AATTCAAAACCTTTTACCTTGAAATTCAACTCAAACATCAACCAGATCCCATTAC
TCAGAATCAAATCTTTAGCCTCTTCTTACCATAAAAAGGCCACTGGATAGCAC
TAGCAACTAGTGGTTAAGACAAAATATGTTATACCTTGAAAGAATCGTGTATTTT
AAATTTCTATCAATAAAGAGTTGTCAATAAAATTCTAATTA AAAATCAATTATGT
TGAGATAAAAACGA ACTTACCAATTGAATTTCTTATCTGTCCATGTATATGCTAG
ACTAGTAATTCTAACTTTCAGCTATGTTAATAATGAAAAACCCCTTTCACCAA
ATTTGATGAGTCTTACACCTCAAGATTTGATTGCATGCATTTTAATATTTGACTTA
ATGACAAGTATATGTATTTTTTACTCAATAAATCAAGTCTAAATCTTCTTCTCGAA
ATTCAAATATCTGGAAAACGACTTTAACATCGAGTAACACATTA ACTAAGGTCT
ACCGGATACCAATTGCTCAAAAAGACAAAAAAGAAAAA AAAAAAACCTTCCTAA
CCATAACTCTTTTAGATGTTTTAGGACAAATGAGATCTTTGAATTTTTATTTAAC
ACTGTTATGCTAGTGATTTATAGTTAAACATGCAGGCATGGCACCCGGAAGTGGG
AGCAATACTAGCATCCTTTTGGCAATAAAGACATATATAGAAATATAAATTGATT
GAAAGGTTTATAATTAGCAAGTAACTTGAATATCAAGG