

IDENTIFICATION

Species: *Theobroma cacao*

Locus: Thecc.05G072200

Gene Model: Thecc.05G072200.1.p

Description: TcEXPA-12

Family: Alpha Expansin

3D structure:



GENOME DATABASES

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

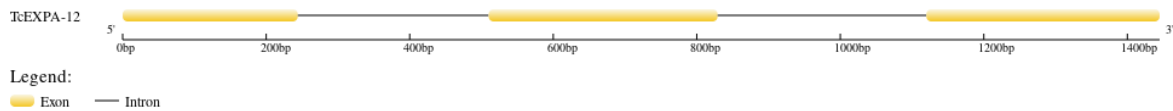
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

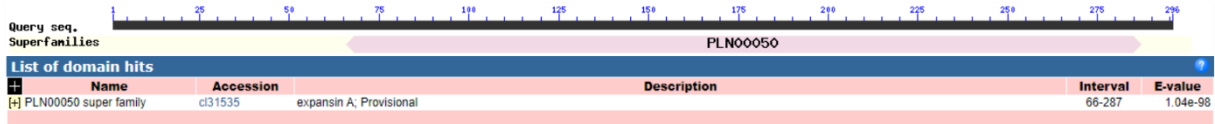
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPA-12

MASPSKAFSATLLLLLAVIIMLLLPECHADHPVKKTSSIIHHVRRVGNPVVSLKHKGR
PPFKPGPWKQGHATFYEGGSGTFGGACGYEDVVKMGYGLNTAALSAAFNNNGQKC
GACFEIKCVDSSEWCKPGQPSLIVTATDSCPPNYDLPSDNGGWCNPPRAHFDLAKPAF
GQIAEYQGGIVPVQYRRVPCKKQGGIRFTVTGNPYFNLVSVWNVGGAGDVISVHVL
KGDNKQTWIPLKRNWGQKWETDAKLVGEDLSFRVRTSDGKTSTSWHVAPKNWQF
DQTFGLGSKNFRGLP*

CDS (coding sequence)

>TcEXPA-12

ATGGCATCGCCTTCCAAGGCTTTCTCTGCAACAACATTGTTGCTTTTGGCTGTTAT
TATTATGTTATTATTGCCTGAATGCCATGCAGACCATCCTGTCAAAAAGACAAGC
TCGATTCACCATGTTCGTGAGAAGAGTTGGCAATCCTGTAGTTTCTCTTAAGCATA
AGGGCCGACCACCTTTCAAGCCTGGCCCATGGAAGCAGGGTCATGCCACCTTCTA
CGAGGGTGGCTCAGGGACATTTGGGGGAGCTTGTGGCTACGAGGATGTAGTGAA
AATGGGTTACGGCCTAAACACAGCGGCATTGAGCGCAGCTTTGTTCAACAACGGG
CAGAAATGCGGCGCTTGCTTTGAGATCAAATGTGTAGATTCTCTGAATGGTGCA
AACCAGGGCAGCCTTCTTTGATCGTGACCGCCACAGACAGTTGTCCACCTAATTA
CGATCTTCTAGCGACAATGGAGGATGGTGCAACCCACCCCGCGCGCATTTCGAC
TTAGCCAAGCCTGCTTTCGGTCAAATTGCTGAGTATCAAGGTGGCATTGTCCCTGT
CCAATACCGCAGAGTACCGTGCAAAAAGCAAGGAGGTATTCGTTTTACGGTCACA
GGCAACCCTTACTTCAACCTGGTGTCAAGTATGGAACGTGGGAGGAGCCGGGGAT
GTTATAAGCGTGCATGTACTGAAGGGTGATAACAAGCAGACATGGATAACCATTG
AAGAGAAACTGGGGTCAGAAATGGGAGACCGACGCCAAGTTAGTCGGGGAAGA
CCTGTCCTTCCGAGTTAGAACAAGCGATGGCAAACTTCAACCTCATGGCATGTC
GCTCCCAAGAACTGGCAATTTGATCAGACTTTTTTGGGTAGTAAAACTTCCGAG
GATTGCCTTGA

Nucleotide

>TcEXPA-12

ATGGCATCGCCTTCCAAGGCTTTCTCTGCAACAACATTGTTGCTTTTGGCTGTTAT
TATTATGTTATTATTGCCTGAATGCCATGCAGACCATCCTGTCAAAAAGACAAGC

TCGATTCACCATGTCGTGAGAAGAGTTGGCAATCCTGTAGTTTCTCTTAAGCATA
AGGGCCGACCACCTTTCAAGCCTGGCCCATGGAAGCAGGGTCATGCCACCTTCTA
CGAGGGTGGCTCAGGGACATTTGGTATGAACTATGGTTACCTTTACGAAGCTTTG
ACATTCTTGAAAAGCCGTCAAAGCTTCCCACTCTTCGTTTTGAAGGGTGATTAATC
TCTTGTTAATCCTAAGATTATATCAAAAAACAAAAGATTACAAGGAAAGCGAA
GATATCTATTCATTTAATCTGTGAAAAGTTATAAGTCCAGGATTCAATTCCATTCT
TGGCTCAAAGACATATTCATGGAGAAGTTATATAACTAGTCGTCGTTCCTTTCCTA
TGTTGAATGTAGGGGGAGCTTGTGGCTACGAGGATGTAGTGAAAATGGGTTACG
GCCTAAACACAGCGGCATTGAGCGCAGCTTTGTTCAACAACGGGCAGAAATGCG
GCGCTTGCTTTGAGATCAAATGTGTAGATTCTCTGAATGGTGCAAACCAGGGCA
GCCTTCTTTGATCGTGACCGCCACAGACAGTTGTCCACCTAATTACGATCTTCCTA
GCGACAATGGAGGATGGTGCAACCCACCCCGCGCGCATTTCGACTTAGCCAAGCC
TGCTTTCGGTCAAATTGCTGAGTATCAAGGTGGCATTGTCCCTGTCCAATACCGCA
GGTTTGCTTCCTTTTTTCTAATTTTCTTCCTTGCACTCTCAAGTCTAAAGTCGATGA
TTGAAGAACACAATTGAGAATAAAAATAAAATTTATATAATTTTCTCTTTAATAA
GTTATATTATTAATGCTTGATTTTGATATTTTATAATCAAATAGTATTTTCGACAA
CTAACACATAGCTTAACATTA AAAACGTA ACTCATGAAAAAATATCAAATTGTA
ATAAAAAAATTACATCCGAAAGAAACATGATCAAAATTCTTAAAAGATTAAAAT
TGAGTTATATTGCAGAGTACCGTGCAAAAAGCAAGGAGGTATTCGTTTTACGGTC
ACAGGCAACCCTTACTTCAACCTGGTGTGTCAGTATGGAACGTGGGAGGAGCCGGG
GATGTTATAAGCGTGCATGTACTGAAGGGTGATAACAAGCAGACATGGATACCA
TTGAAGAGAACTGGGGTCAGAAATGGGAGACCGACGCCAAGTTAGTCGGGGAA
GACCTGTCCTTCCGAGTTAGAACAAGCGATGGCAAACTTCAACCTCATGGCATG
TCGCTCCCAAGAACTGGCAATTTGATCAGACTTTTTTGGGTAGTAAAACTTCCG
AGGATTGCCTTGA