

IDENTIFICATION

Species: *Theobroma cacao*

Locus: Thecc.03G151400

Gene Model: Thecc.03G151400.1.p

Description: TcEXPB-04

Family: Beta Expansin

3D structure:



GENOME DATABASES

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

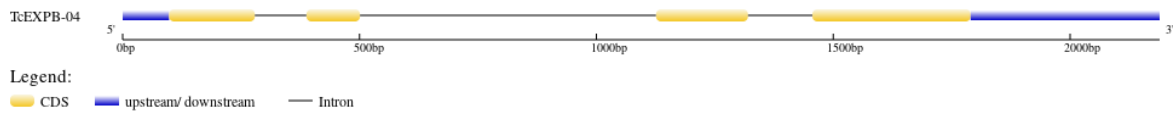
KEGG: <https://www.genome.jp/entry/gn:T02994>

EXTERNAL RESOURCES

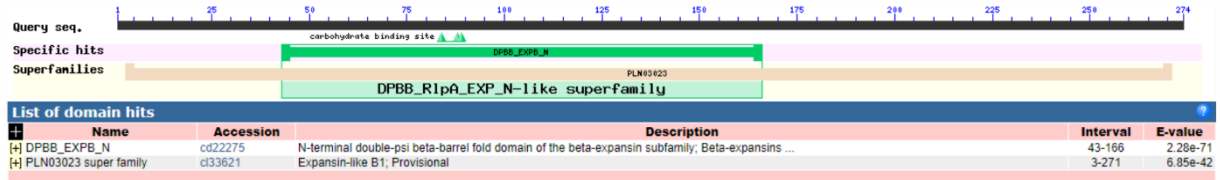
<https://www.cacaogenomedb.org/>

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

GENE STRUCTURE



DOMAIN ARCHITECTURE



SEQUENCES

Peptide

>TcEXPB-04

MAPLSYICLPLAYIIVLVFFCHFELCVSFKSKQFNLSTIEIHWSSAGATWYGSPDGAGS
DGGSCGYGNAVSLPPFSSMVTGIGPSLYKSGKECGACYQVKCTKHPFCSGKPVRVVI
TDFCPGGLCASDSAHFDLSGTAFGAMAIAKEDKLRDAGVLEIRYARVACDYSGKTI
VFHVDMGNSNSNYFAVLVEFEEGDGDLAGVYLKETSSGSGDWRAMQQSWGAVWKL
DAGSELHPPFSIRLKSQLSGQTLVAQNVIPQGWKPGATYRSLVNYL*

CDS (coding sequence)

>TcEXPB-04

ATGGCCCCTCTTTCTTATATTTGTTTACCTTTAGCATATATAATAGTATTAGTCTTC
TTTTGCCACTTTGAGCTGTGTGTTTCTTCAAGTCCAAACAATTCAACTTATCAAC
CATTGAAATCCATTGGTCATCCGCCGGAGCCACTTGGTATGGCAGCCCTGACGGC
GCCGGAAGTGATGGAGGTTCTTGTGGGTATGGTAATGCAGTGTCACTGCCTCCTT
TCTCTTCCATGGTAACTGGAATCGGCCATCTCTCTATAAATCAGGCAAAGAATG
TGGAGCCTGTTATCAGGTAAGTGACCAAACATCCCTTTTGTTCAGGGAAACCG
GTCAGGGTGGTGATAACAGATTTTGGCCCTGGAGGCCTTTGTGCATCCGACTCTG
CACATTTGATCTGAGTGGAACAGCATTTGGTGCCATGGCAATTGCCGGTAAAGA
AGATAAACTTCGAGACGCTGGAGTATTGGAAATTCGATACGCAAGAGTTGCCTGT
GACTATTCAGGAAAAACAATAGTCTTCCACGTTGACATGGGGTCAAACCTCTAACT
ACTTTGCAGTGCTGGTTGAGTTTGAAGAGGGAGATGGTGATCTTGCCGGTGTCTA
TCTGAAGGAGACGTCAAGTGGGTCGGGTGACTGGCGAGCCATGCAACAATCATG
GGGTGCAGTGTGGAAGCTTGATGCTGGGTCAGAATTACACCCTCCATTTTCAATC
CGATTAATAATCTCAATTATCAGGCCAAACCCTGGTTGCCCAAACGTGATCCCAC
AAGGCTGGAAGCCTGGAGCCACGTATAGATCCTTAGTTAATTACCTGTAA

Nucleotide

>TcEXPB-04

TATATCTAAGTAGCCAATAGCAATACGTTTAAAGAGTTTTTTTTCCAAGCCCCTTAAA
TTATTGTAGTGCCTCTATCTAGTGCCTAGTGAAAGTGTAACCTATGGCCCCTCTTTC
TTATATTTGTTTACCTTTAGCATATATAATAGTATTAGTCTTCTTTTGGCACTTGA

GCTGTGTGTTTCCTTCAAGTCCAAACAATTCAACTTATCAACCATTGAAATCCATT
GGTCATCCGCCGGAGCCACTTGGTATGGCAGCCCTGACGGCGCCGGAAGTGATG
GTAATTAATGCAGAGAGACTTGGAAGGATGGGGTGGGTTGGGGGGGATATATAT
TAAAGCATGACTAGGCTGACTAATGAATATCGATCTATGCGTGTCATTTATGCAG
GAGGTTCTTGTGGGTATGGTAATGCAGTGTCACTGCCTCCTTTCTCTTCCATGGTA
ACTGGAATCGGCCCATCTCTCTATAAATCAGGCAAAGAATGTGGAGCCTGTTATC
AGGTAAGAACCAGCTGCAGCTTTAGAGCTAACGCCTATTGTGCAACTTGCCTGAA
TCATGCACATCAACATTGTTCAACTTGACTTCCTAGTTCCAAGTACTTGGATATGA
CACGTTTAAAACCTTCTCATTTCCTTTAACAATAGAAATTTCTAAGTTTAACTA
TTCCTTCAGTTTAAATTTATAACAAGTTTTCAAAAAGTGCATTTAGATTTTCGAGTAC
AAGTAGTGCCATTATTAGTGGAAGTTTGAAGTCAAGAGTCAAGACTTAGCACAT
GCCATACTTAAAAAAGCACAGTACAGTATGCCACTGAGTTATACATCTCATCCT
GGCCATCCATTTCAATAGTAGCAGTCACTCGTAGCTTTTGCTTTTGTTACTACTTTT
CCTTCTTCTTGGCAGACAATGGGATGAGATAGTTCCATATGTACAAAATTTTCAG
GAGACAGAGACGGTCCCAGAAGTGTCAATTTTTCATTGCCTTCCTTGTTCTTTTCAT
TCTTCGACTTGTTGAATTGCATATTCTGATCTAATTACTTTTTGCTTCTCGATCATG
CTTGTTAACAAAGGCTGACCTGATCAGTCTGTTGTTTGTATAGTAATTAATGTG
TCTTTTGCTATAGGTAAAGTGCACCAAACATCCCTTTTGTTTCAGGGAAACCGGTC
AGGGTGGTGATAACAGATTTTGGCCCTGGAGGCCTTTGTGCATCCGACTCTGCAC
ATTTTCGATCTGAGTGGAACAGCATTGTTGGTCCATGGCAATTGCCGGTAAAGAAGA
TAAACTTCGAGACGCTGGAGTATTGGAAATTCGATACGCAAGGTAATAATCCTTT
TTCTAATCAATGTTATGAAAGACTGAAAAACAATAAAAAACTTATGGTCATTTT
CAAATCATGGAAAATATACAAATTATGAGGAACAGTCCATCATTTTTTTCTTGTTT
GGCGGATTGCAGAGTTGCCTGTGACTATTCAGGAAAAACAATAGTCTTCCACGTT
GACATGGGGTCAAACCTAACTACTTTGCAGTGTCTGGTTGAGTTTGAAGAGGGAG
ATGGTGATCTTGCCGGTGTCTATCTGAAGGAGACGTCAAGTGGGTCCGGTGACTG
GCGAGCCATGCAACAATCATGGGGTGCAGTGTGGAAGCTTGATGCTGGGTGAGA
ATTACACCCTCCATTTTCAATCCGATTAATAATCTCAATTATCAGGCCAAACCCTGG
TTGCCCAAACGTGATCCCAAGGCTGGAAGCCTGGAGCCACGTATAGATCCTT
AGTTAATTACCTGTAAGCTTGGGCAGTGCTCCCAAGGCCATGCATCAAACAGGGA
GTAAAGTTTAAACCCTGGAATTGTGGAAAACCTCTCCATCTTATAAGCTGAAGC
GAGACTACATGACCAAATGAGCAACGCTTAATCATGGACAATTACGTATGTTAAT
TAGCAGTTCAACTTCGGTCTAATTAGTGATAATGTCTAATTAGTGCCTTATAATA
TTGTGCATGAGTTGATTTTCATCACGTGCTCAAGATGGATGTAATTAGACATCGTTT
TCCGTTTTGCTATTTTTCAATTTTCAACACTAGTCCCTCCCTAATCTCTAGAAATGAA
AAATGGTAATGGGAAAACGGGCTAATTTTAGGTTTTCAAAGTCCCATCCTTTA
AGAAGGCACAATGAAGAAATGAGTAA