

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence	intracellular	extracellular	cytosol	nucleus	mitochondrion	endoplasmic reticulum	g a
Q02224	CENPE_HUMAN	Homo sapiens	Centromere-associated protein E	22.501419	T1928;S1933;T1951	S611;S2083;S2389;S2639;S2647;S2651	37217939;40136647;30379171	MAEEGAVAVCVRRVPLNSREESLGE TAQYVWKTDDNNVIYQVDSKSFNFD RVFHGNETTKNVEEIAAPIHDSAIQ GYNGTIFAYGQTASGKTYTMMGSED HLGVIPRAIHDFQKIKKFPDREFLLR VSYMEIYNETITDLCGTQKMKPLII REDVNRNVYVADLTEEVVYTSSEMAL KWITKGEKSRHYGETKMNQRSSRS HTIFRMILESREKGEPSNCEGSVKVS HLNLVDLAGSERAAQTGAAGVRLKE GCNINRSLFILGQVIKLLSDGQVGGF INRDSKLRILQNSLGGNAKTRIICT ITPVSFDETLTALQFASTAKYMKNT YVNEVSTDEALLKRYRKEIMDLKKQ LEEVSLETRAQAMEKDQLAQLLEEK DLLQKVQNEKIEINLTRLMLVTSSSLT LQQELKAKRRRVTWCLGKINKMK NSNYADQFNIPNITTKTHKLSINLL REIDESVCSSESDVFSNTLDTLSEIEW NPATKLLNQENIESELNSLRADYDN LVL DYEQLRTEKEEMELKKEKNDL DEFEALERKTKKQEMQLIHEISNL KNLVKHAEVYNQDLENELSSKVELL REKEDQIKKLEQYIDSQKLENIKMDL SYSLESIEDPKQMKQTLFDAETVALD AKRESAFLRSENLELKEKMKELATT YKQMENTIQLYQSLEAKKMQVD LEKELQSAFNEITKLTSLIDGKVPKD LLCNLELEGGITDLQKELNKEVEEN EALREEVILLSELKSLPSEVERLRKEI QDKSEELHIITSEKDKLFSEVVHKE RVQGLLEEIGKTKDDLATTQSNYKST DQEFONFKTLHMDFEQYKVMVLEE NERMNQEIVNLSKEAQKFDSSLGAL KTELSYKTQELQEKTRVQERLNEM EQLKEQLENRDSTLQTVEREKTLITE KLQQTLEEVKTLTQEKDDLKQLQES LQIERDQLKSDIHDVTNMINIDTQEQ LRNALESCLKQHQETINTLKSISEEV SRNLHMEENTGETKDEFQOKMVGI DKKQDLEAKNTQTLTADVKNIEIE QQRKIFSLIQEKNELQOMLESVIAEK EQLKTDLKENIEMTIENQEEELRLG DELKKQOEIVAEKKNHAIKKEGELS RTCDRLAEVEEKLEKESQQLQEKQ QLLNQVEEMSEMOKKINEIENLKN ELKNKELTLEHMETERLELAQKLNE NYEEVKSITKERKVLKELQKSFETER DHLRGYIREIATGLQTKELKIAHI HLKEHQETIDELRRSVSEKTAQIINT QDLEKSHTKLQEEIPVLHHEEQELLP NVKEVSETQETMNELELLTEQSTTK DSTTLARIEMERLRLNEKFQESQEEI KSLTKERDNLKTIKALEVKHDQLK EHIRETLAKIQESQSKOEQSLNMKE KDNETTIVSEMEQFKPKDSALLRIE IEMGLSKRLQESHDEMKSVAKEKD DLQRLQEVLSQESDQKNIKEIVAK HLETEEELKVAHCCLKEQEETINEL RVNLSEKETEISTIQKLEAINDKLO NKIQEIYEKEEQFNKQISEVQEKVN ELKQFKEHRKAKDSALQSIESKMLE LTNRLQESQEEIQIMIKEKEEMKRV QEALQIERDQLKENTKEIVAKMKES QEKEYQFLKMTAVNETQEKMCIEIE HLKEQFETQKLNLENIETENIRLTQI LHENLEEMRSVTKERDDLRSVEETL KVERDQLKENLRETITRDLEKQEEL KIVHMHLEHQETIDKLRGIVSEKT NEISNMQKDLHSNDALKAQDLKIQ EELRIAHMHLKEQQETIDKLRGIVSE KTDKLSNMQKDLSENSAKLQEKIQ ELKANEHQLITLKKDVNETQKKVSE MEQLKKQIKDQSLTSLKLEIENLNA QKLHENLEEMKSMKERDNLRRVE ETLKLERDQLKESLQETKARDLEIQ ELKTARMLSKHEKTVDKLREKISE KTIQISDIQKDLKSKDELQKKIQEL QKKELQLLRVKEDVNMSHKKINEM EQLKKQFEAQLNSMQSVRMDNFQL	True	False	4.732	4.291	False	False	F

TKKLHESLEEIRIVAKERDELRRIKES
LKMERDQFIATLREMIARDRQNHQV
KPEKRLLSDGQOHLTESLREKCSRI
KELLKRYSEMDDHYECLNRLSLDLE
KEIEFQKELSMRVKANLSL PYLQTK
HIEKLFTANQRCSMEFHRIMKCLKY
VLSYVTKIKEEQHESINKFEMDFIDE
VEKQKELLIKIQHLQDCDVPSREL
DLKLNQNMDLHIEEILKDFSESEFP
SIKTEFQVLSNRKEMTQFLEEVLN
TRFDIEKLNKIQKENDRICQVNNF
FNNRIAIMNESTEFEEERSATISKEW
EQDLKSLKEKNEKLFKNYQTLKTSL
ASGAQVNPITQDNKNPHVTSRATQL
TTEKIRELENSLHEAKESAMHESKI
IKMQKELEVTNDIIAKLQAKVHESN
KCLEKTETIQVLQDKVALGAKPYKE
EIEDLKMMLVKIDLEKMNAKEFEK
EISATKATVEYQKEVIRLLRENLR
QQAQDTSVISEHTDPQPSNKPLTCG
GGSGIVQNTKALILKSEHIRLEKEISK
LKQQNEQLIKQKNELLSNNQHLSN
EVKVTWKERTLKREAHKQVTCENSPK
SPKVTGTASKKKQITPSQCKERNLQ
DPVPKESPKSCFFDSRSKSLPSHPV
RYFDNSSLGLCFEVQNAGAESVDSQ
PGPWHASSGKDVPECKTQ