

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q3TLH4	PRC2C_MOUSE	Mus musculus	Protein PRC2C	48.301093	S878;S882;S1918;S1929;T1932;S1933;T2143;S2148;T2163;T2173;T2181;S2190;T2194;T2195;S2197;S2421;S2429;S2431;S2585;S2632;S2637	S102;S189;S192;S336;S501;S755;S761;S777;S853;S904;S1215;S1219;S1221;S1222;S1236;T1238;T1240;S1516;T1917;S1935;S1938;S1965;S2057;S2095;S2212;T2625;T2634;S2638;S2646	21540332;22645316;22517741;36852467;33300544;36288343;30059200;28528544;35822049;26192747;34678516	MSEKSGQSTKAKDGKKYATLSLFNT YKGSLETQKTTVAARHGLQSLGKV GISRRMPPPANLPSLKAENKGNPN VNIIVPKDGTGWASKQEQHHEEKAPE VSPAQPKPGVAAPPEVAPAPKSWAS NKQGGGQDGIQVNSQFQQEFP AAGDQEKKEKEANDENYGPGPSLRP PNVACWRDGGKSAGSPSSDQDEKQ LGQDESTAITSEQNDILKVVEKRIAC GPPQAKLNGQQPALASQYRAMMPP YMFQQYPRMAYPPLHGPMPRFPPSL EANKSLRGRGPPPSWASEPERPSILS ASELKELDKFDNLDAEADGEGWAGA QMEVDYTEQLNFSDDDEQGSTSPK ESSEEDQTAKTPESTENRKEVDEAS STKSSSQIPAPPPVTKSPYKGGPPFN QERGPSSHLPPPKLLAQHPPPPD RQIPGRQGPPPSKPPVPDNDIWKQ RRKQQSEISAVERARKRREEERR MEEQRKAACAEKQLKQDEKLGIEK QPSPEELREREREKERERELEKEKE RELEKEQEKQREMERARQKEKELE QQREKEQELQRLREQEKEGEPKEQ EKEEKVEPQEPVVEPATENQESENN CKKEEPIFTRQDSNRSEKETTQV QEAEPESGAQPRPGYFKQFQKSLPP RFQRQQEQMKQQQWQQQQQQQQ QGVLQTVPSQPSNGSVPPPHRPL YQPMQPHPQHLASMGFDRPRLMM QSYM DPRMISGRPAMDIPPIHPGMI PPKPLIRRDQMEGSPNSSESFEHIAR SARDHGISLSEPRMMWGSDPYHAE PQQAATPKSAEETGDARPETAMDHE HMTAAYPVEHSQLETHSKTDVARD TETEEQKFLSRSLDVKPRHVDNT QSACFDVIDQKSLPTSAEERISALE QPARKRSVSHGNSHAQNAEEQRNE PSVSIPKVINRCMDSKETVEKPEKP RKDGF LRSSEGPKEKVKSKSETR WGPRPSSNRREEGNDRPVRRSGPIK KPVLRDMKEEREQRKEKEGEKLEKV TEKVVKAEKPEKDLPLPLPPAPAQ PQPQLVSPVQPEPEKPPSTETSTL TQKPSQDEKPLEPVGVSQVEPVVKT VNQQSVAAPTvkEEKPPEKVINKDV GIERSRPDSRLAVKDDSSLPTRTYWK EARDRDWFPDQYRGRGRGEYYSR GRSYRGSYGGRRGGRGHTREYPQY RDNKPRTEHVPSGPLRQRESETRS ESSDFEVVPKRRRQRGSETD TDSEV HESASDKDSVSKGKLPKREERPENK KPVKQSSFKPENHVRIDNRPLEKP YIREDKSKPGFLPKGEPTRRGRGGT FRGGDRDPGGRPSRPATLRRPAYRD NQWNTRQAEPKPEDGEP RRHEQ FMPIPADKRPPKFERKFDPARERPR RQRPTRPPRODKPPRFRRRLREREA SKTSEVLVPSNGTANNVQEPVNP ADISGNKTPDLNQNSSDQANEW ETASESSDFNERRERDEKKNADKSS QAVVKAGESVLPKREIAKRSFSSQR PGVDRQNRNGNPPKSGRNFSGP

RNERRNGPPSKGGKRGPFDDQASG
TAGADPVSGNSAHHQEGVPNGAGQ
KNSKDAAGKKREDTKPGPKPKKEKV
DALSQFDLNNYASVVIIDHPEVTVI
EDPQSNLNDGTFTEVVSKKQKRL
QDEERRKKEEQVVQVWSKKNIGEK
GRSQTSLPPRFACKQATGTQQIQAP
PSAPVPVSSAPGLTAPAAAAPASTP
APVPILASATALVPVSTPAPVLTSCPA
PVPTSASAPVPASTSSPVTASSSSQP
SVPAPTPVLASASTTVSVPILTSASIPI
LASALAPATVSSPTPVVSATAVPSIST
PAVPASAPTASVPLAPASAASTVPPP
ASTVQTQTQTQTHKPVQSPPLPPSAPS
SKQPPPSIRLPSAQTSNGTDFVAAGK
SMPTSQSHGSLTAELWDSKVAAPAV
LNDISKKLGPIPPQPPSVSAWNKPL
TSFGSATSSSEGTRNGQESGVEIGIDT
IQFGAPASNGNENEVVPVLSEKATD
KVPEPKEQRQKQPRAGPIKAQKLPD
LSLVENKEHKPGPIGKERSLKNRKV
KDAQQVEPEGQEKPSPAVVRSTDPE
TAKETKAVSEMSAEIGAMISVSSAEY
GSDAKESVTDYTTSSSLPNTVATN
NAKMEDTLVNNVPLPNTLPLPKRET
IQSSSLTSVPPTTFLTKMESARK
AWENSPNLREKQSPVTSTAPPVSG
VSSSASGPSTANYSSFFSSASMPQIPV
ASVTPASLSGAGTYTTSSLSTKSTTT
SDPPNICKVKPQQLQTSSLPSASHFS
QLSCMPSLIAQQQSPQVYVSQSAA
AQIPAFYMDTSHLFNTQHARLAPPS
LAQQQGFQPGLSQPTSVOQIPIPIYA
PLQGQHQAQLSLGAGPAVSAQELF
SSSIQPYRSQAFMQSSLSQPSVVLV
GTAIHNFPAVQHQLAKAQSGLAFO
QTSNPQIPILYDHLGQASGLGSSQ
LIDTHLLQARANLTQASNLYSGOVQ
QPGQTNFYNTAQSPSALQQVTVPLP
ASQLSLTNFGSTGQPLIALPQTLQPQ
LQHTTPQAQAQSLSRPAQVSQPFRG
LIPAGTQHSMMATTKMSEMELKA
FGSGIDIKPGTPPIGGRSTTPTSSPFR
ATSTSPNSQSSKMNSVYQKQFQSA
PATVRMAQPFPAQFAPQILSQPNLV
PPLVRAPHTNTFPAPVQRPPMALAS
QMPPPLTTGLMSHARLPHVARGPC
GSLSGVRGNQAQAALKAEQDLKAK
QRAEVLQSTQRFFSEQQNKQIGGK
TQRVSDTNSPETLSDPPGTCPEKV
EEKPPPAPTITTKPVRTGPIKQAIKT
EETKS