

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
F4J8G7	VIR_ARATH	Arabidopsis thaliana	Protein virilizer homolog	13.157209	NaN	NaN	28154133	MVRSEPCVLFAQTFVHPQLDEYVDE VIFAEPVIITACEFLEQNASSSSQAVS LVGATSPPSFALEVFVRCCEGESKFKR LCNPFLYTPSAPYPLEVEAVVTNHLV VRGSYRSLSLIVYGNIVKDLGQYNIIL EGRSVTDIVSSTEGNLEDLPLVLHSV NRTIEECLSSLDIVSLPLAAVDLPVEV KRLQLLLKIFDKLATNDVVNKFVDT VVSGVSSYVTDNVDFFLKNKNCSAV TSSLDSGLFHDIVDRVKEDILDNEI QESDVALGLFSFLESETYLATSQQLV VMLSPYIQFERDSLCTVLPKLSKGKA TLLGLSLAFLLCSGREGCLQFVNSG GMDQLVYLFVGHGQNSTTITLLLLG VVEQATRHSVGCCEGLGWWPREDG SIPSGKSEGYCLLLKLLMQKPCHEIA SLAIYILRRLRIYEVISRYEFAVLSALE GLNSHGAATHNLNMLSDAKSLOLQ KLQNLMSKLSGVEDPSPSAYAERSL VSDHSEGWLSYKATSKLTSSWTCPF YSSGIDSHILALLKERGFPLSAALLS MPELHSHKVDIMDVFTDIAMFIGNII LSFMFSRTGLSFLHHPELTATIIQS LKGSDLNKEECVPLHYASILISKGF TCSLLEIGINLEMHLRVVSAVDRLLK SIQQTEEFWLWELRDVSRSDCGR EALLTLGVFPEALAVLIEALHSKDM EPAVENSGISPLNLAICHSAAEIFEVI VSDSTASCLHAWIEHAPVLHKALHT LSPGGSNRKDAPSRLKWKIDAGVYV HKHGVGGLLRYAAVLASGGDAQLSS SSILALDLTPAENGAGESTNVSEMN VLDNLGKVIFEKSFEGVNLSDSSISQ LTTALRILALISDNSTVAAALYDEGAV TVVYAILVNCSEFMFERSNNIYDYLVD DDHGCSISDFLSERNREQSLVDLLI PSLVLLISVLQRLQGTKEQYRNTKLM KALLRLHREVSPKLAACAADLSSHYP DSALGFGAVCHLIVSALVCWPVYGW IPGLFHTLLSGVQTSSVPALGPKETC SFLCILSDILPEEGVWFVKSGMPLLS GLRKLAVGTLMGPOKEKQINWYLEP GPLEKLINHLTPNLDKIAKIIQHHAV SALVVIQDMLRVFIVRIACQVEHAS ILLRPIFSSIRDGILDQSSTRDTEAYM VYRYLNFLASLLEHHPHAKGLLLEGI VQLLVEVLERCYDATYPSENRVLEY GIVSASSVIQWCIPAFRSISLLCDSQV PLLCFQKHELLASLSAKDCALIFPFV

LKFCQVLPVGNELLSCLGAFKDLSS
CGEGQDGLVSLLFHLFSGTEESVSE
RWCDTNSLSLDQLDMKKNPPFLSC
WIKLLNSINSKDGSLSSLAMKAVNVL
SVGSIRLCLDGKSLDSKKVAALKSLF
GLPSEFSGTDTFREENIGLIEQMVTL
LSSMTSGSDSSATAEMKPYLHEASQ
SLLSLLKDGNIIDDIISCKGVFVSPGN
LDMDDLVSRIEDDLYQRGLEDKF
WWECPETLPERLPQSSLPKRKLPT
LESSRRAKGENSSVDIPTQNSIQRG
MGSVSLPPAPTRRDAFRQRKPNTSR
PPSMHVDDYVARERSVDTAGNSNAI
TISRAGSSSGRPPSIHVDEFMARQRE
RGQNPSTIVVGEAVVQVKNPTPARD
TEKVAGKPKQFKADPDDDLQIDIVF
DGEECEGPDDKLPFLQPDENLMQP
APVMVEQNSPHSIVEETESDANGSS
QFSHMGTVPVASNVDENAQSEFSSRI
SVSRPEMSLIREPSISSDRKFVEQAD
EAKKMAPLKSAGISESGFIPAYHMP
GSSGQNSIDPRVGPQGFYSKSGQQH
TGHIGHGFSGRGVYEQKVMNPQPPL
PLVPPPSVSPVIPHSSDSLNSQSSPFI
SHGTQSSGGPTRLMPPLPSAIPQYSS
NPYASLPPRTSTVQSFQYNHAGVGT
TEQQQSGPTIDHQSGNLSVTGMTSY
PPPNLMPSHNFSRPSLVPFYGPN
SHQGGDKPQTMLLVPSIPQSLNTQS
IPQLPSMQLSQLQRPMQPPQHVRPP
IQISQPSEQGVSMQNPFQIPMHQM
QLMQQTQVQPYHPPQQQEISQVQ
QQQQHHAVQGQQGAGTSQQQESG
MSLHDYFKSPEAIQALLSDRDKLCQ
LLEQHPKLMQMLQEKLGL