

Text S1 - Curated sequences of PGTG peptides using a comparative approach as described in Table S2.

>PGTG_06567.2
MLHPESYTTVPSASISQAVLKSVDIIIVGAGIAGSALAFALTDPSSRNRTGKSPSVLLIERDLRQPDRIVGEELLQPGGCLAVKRLGLRD
CLDEIEAVEVNGYGVYWGTDASQITQLALPYPPESVPMWAKDGLWNGKSPKQAPRQGRSFHHGRFVQRLRWKAHSRPTVTVLQATV
TDLIRCPKTDHVIIGVTVKSAEEETVSFFAPINFIMDGCFSKFRRIIAPDGFQKQPTVRSHFVGLLLOTPAPFFDCIPLPGHGHVILRKKDP
AADQPVDGELGVGPVLVYQIGTGETRMLVDVPGAKVPSISNGSLHSYLERQVGPILPRSLETFHATLDSNDFADRLRVMPNSYLPH
RQEGHGGVILMGDSMNRHPLTGGGMTVALLDVEIISNLLGLDLDFFEDWKAIEDRLAIWQQRKSTSTCINVLAQALYSLFGAEDDNL
ILKEGCFKYFELGGKRVSDPISLLSALIPSPLLLFFHYHFFSVAFYAIWIFTQQNGVQINKMFQILWTACVTILPVLWAE*

>PGTG_15040.2
MNNRSVSGPRLPPGSRPKAIHIKQKIPSSSLAAITRRRLTLLTIFSAFIGVYRWADAIKSRWYIFDPSSELHKITLEAIERYPNSTSAVI
HHIVQSFESQPTLKPYYITANPFPDPSQAVSAYPSEWVFNAGGAMGAMYIIHASITEYLIVFGTPIGTEGHTGRHTADDYFNILAGQQ
TAFAGSLEAEVYRKGSVHLLRRGQVKYKMPDGGCWALELAQGWIPMLPFGGLADSVFSTVDLITTYHTIRITAREIIRNLLAGKI*

>PGTG_07202.2
SSLIDPLIEFIGSFSTFNQILYIFLLSITSIIISINIFNLAI PKDPTTTPVVHFLFPFIGSAVSYGIDPYAFLESCRKKGYNVFTFVLL
NKKVTVALGLEGNALILNGKLSQVNAEEAYTAL
GAP IN GENOMIC DNA SEQUENCING
PKTQQAVKDCFKVASEITICTASATLQGPVEVREGLNKSFANLYHDLDDGGFTPLHFAPNLPLPSYRRDRRAQVAMRNFMNIIQKRRED
NREGQLGDMIDSLQGGQTYKDGRLTDKEIAHIMIALLMAGQHTSAATGSWLLHLASRPDIVAELRQEIEVFGKPGQTDKDELDPDL
ERVQSPMLLACIKEVRLRHPPHISIMRKKVSPITVPRTLASHNEDTPYIIPSSNFVLAAPGASQIDPAIWSSPHEFEPSSRWLKLTSFPK
AGGGETQEEMVDYGFGMISGANSFPFLPFGAGRHRICIGEFAYLQLSLTGLATVIRNCELELVSNQFPKPDYTTMLVCPKPRDVKFTRR
NTHS*

> PGTG_06567.2
MRIHCIGRGSIGSLLCFHLQSIPTITLLRSRQAQHRRSIPTLSIQLEQQDRTRTATGFTYEFNLKQQQPIESLIVTTKAPHVLESLO
RVRHRLSANSTVLLLHNGLGVVEELIETCFQEPSSRPTFVLATTSHGYYRIDKGLPGTQAGSHGRFCHAGLDIRLGLVLPNTTIRNCL
RLRGHSNQPSNDDSQLQDDNPLLNPHSRTPVLEELHPDIEPETRSLHYTLSSLLNPLMVKELNWKWPMGDFQTSALIKLTVNAAI
NPISALLETRNEALYRESSFESLCRQVCQEASAVFAAQAGQFPRPHHLSAPNLQRVVNDIVLATRANISSMCSDIRTLATNRI SPHKT
LSKANLNRIAASQAPIKIPNYQSLISGQEEKSVKETSSTEIDYINGYICRLGSGFNVDTPLNQSLSDLIKLSVAIKRAQVLPKLQVRNR
RLKINRPEDNPATPDQLDLFEKPENNPGSHQLDLFEKPENNHLADNQASVVVDKS*

> PGTG_11438.2 - PGTG_11439.2
MDIILDADQLIILDCWRSRIFPIQLNDLKRPNPSLPLASFSPRITVTQNPQANISYSSVFDYFTCGSQAFQYQADRKLISITEV
VNSNLTNFGAHQRCLSRFRDHIIFRQSVSIFVIFIMGIFSMYFFLSLSYFFIFDHRLKHKHPKFLNQIKLEILCALKAIGPIDILA
LPLHLAQVRGHSKLYTHVSDIKGVSGFRFLFKPILDFFKLTKEFQDDDSWHQLRATQPTSRLSYPMLDNVKKRISLALRKTFFHNFGGG
WFYFIFSYFLFLYI IHRIEHHPVLYKRIHKTHHKWVITPFFASYAFHPVDGFLQSMPIYHIFVIFPFFHRFLYLFALFVTVVWTILIHDS
ELIVGHRLEDYINGPHTHTLHLLYFNCFNGQYFTWTDKLFSTYRNPEVDDKSTLDLAQQKSTIRKNKEMVLEKDYQ

>PGTG_17808.2
MKQVIDSVIGTGGPNPPRRDEKSDAQYSEKSGAYQASNHQNLIPGKFNPDADPHTHYEFGGPWGSLGLMILFPCLMYFFIICLWCYDG
KLSRPDSLHPTETIRWSSEFWQLIKLHTRPTWSATYLMGLLIHQVALAWYMPGVVQEGPLIPSLNGKLSYCNALCSWYATLATV
LHSVLRVRLGDVFDQLGHLMTIATIFGFAISLFYVLPILQGGAVRMSGNHIDFFMGAALSPKIGHIDVKLFAEVRIIPVWLLFIIAV
AGSVKQYETIGVYTPNSLFMVGTYGLYINACARKEECIPLTWMAYEKWGWLLSFWNFAGVAFYCHSVIYITNQPPSKYHFSTWSYVA
LYLTYTAAYVFDTSNSQKARFKMENDSKSQITSRIGYFPPQLSYGLTKLPKVLVVGNGDRKLLIDGWAVCRHPNYTADFIQALCWAAC
SGTGLIPIFYFAFFLVMILHRCTRNFERCSSKYGKSWDEYCSLVPYFSFIGVI

> PGTG_08907.2
MSRVALLLSDFVGLLSSKFLLESFERNPFLDRLVVFCLSWPIAREIRLRLEQRDLRARASSRGALLAPLVSSPLPFGLSVLLTRLRMI
HSGSPGDVILHFNRSRVPKPIHPDQPTKVFRRSRVMGVETIWTIDHDDAKYFLSTGFPNFGKSPFLKAGFRRLLDGDFASDQRLGAWHR
SLTRPHFVRERIAADVAMEHSHRVATWLSTQTDLGKSVDIQDIFARYTLVTGTQHLFGRCVDSLNDLIHDIRIQTGPNADFAQNFVAA
QHWAIINSLLLHPLLISLGFRIADRATEEVQVVDTLVQDASLSLASQIKKNNESDSDQAEAGGTASENLDDHLLTSGCSKELVRQECL
NILLAARDTTASLLSICIYELARDSPRKTAMWRKLDKDEVERLGGSDIVTLDQVRELKYLRAVLNESLRLHPPVWANTRHAFEDDVLPSG
VFVPAGTDCRFFIREFQRNPEVWGMGDAEEFDPDRWIDSRKALQVKDPFSPQFPFAGPRI CLGQQFALAEASMMIRVIEGFEGVDLDDL
DGPVGAEPVAVVLSFRGGLKVRFKK

>PGTG_12360.2
MTNNPKPSWAGPLGSDTPSRWHLQVPVDEDAVWVYDSSPDRNLGAQSFQSKYWSRHSKSPALPDPEGDPLQAAKNGFEFYKQLQMPD
GHWSGEFSGFLPLTPGMVIACYITKTPLAEEVKIELARRFANDQRQGNVRDRGWGLHTAGKSTVFGTVLNYVACRLGIDAEHTMMVR
ARATLHALGGATGPTWGVVWALLGVYDWEVNPVPELWLLPEILPLHPWRWVHRSRQVYLPISYLCGKRLQAQSDPTLDSLRKELY
TQPYESIDWPCRNLIAKEDLYSPRHPIANGLFWILGYWEKICPSSIRNLGLNRAHELCKMEDENTDFNDLAPVNKVLNLIVCWDRYGF
ESDELRRHQKLLKNFMWMDKMGMSATNGSGLWDLGFIITQALVESGLAKTEEPSTQDSVIRALQWIDRCQILENPKHFKSGYRHSKG
AWPFSTKSQSYTVSDCTAEALKSVLCLQEELSYPKLSKERLCLAVDVILSLQNPNGGFASVELIRGPSWLEYLSPAIEVFGKTMIEVN
YPECTTACLTAMSLFSQYYPDYRAPEISRARQAAIKFISHAQRDDGSWYGSWGVCFYATMFALESLSLNNETYKNSLLVKKACRFLLD
RQMDGGWGESFKSCEQGVYIHHQTSQVFQTAWAVLALLAAKYPEPEPIQRACRLIISRQTADGQWLDGAEIEGVFNKTTSVTYPHYKFA
WSISALGKAHKRFPDVQW

>PGTG_02889.2
MAFQPKKIIVIGGEGFLGHNLVQTLHRTYPDSTISSLDLTKRFPDCKDEEPRKEQQQAENQETNNTHQFIQADLTSLDSLLEAFQQTETP

ELVFHTASPWGSSSEICEKVNIQGTLLNTIAACLKFGVQRLVYTSSAGVVFNGNDLINVDERLPVPKIGCDPYNTSKARAEAVLEANG
KDSLLTCAIRPAGIFGPGDRQAIPLGLIEVLKTRKHGVQIGGNNTLFDWYVDNVVHAHILAAEKLDREVPLGEFSTLSLSPISKTVERRN
LLTSGSKAEEDSSMGGTDEVSLVDLSTSSKDAESYNPTGWIQGHGEMVDQAVPAKRHRWDQWAPISTQINYRDDQVRVAGEAFFVTGG
EPVFFWDFARAVWHEYAAHSPQAKALNLDPKPRFTIVIPTFLALFLASLAQLFAKLTNSTTLFTPEKVRYSASKYHNIEKARVVLGYE
PLIGIHEGIQKAVQWYISNESLTQSSQTSDDKID

> PGTG_19538.2

METKNPCSTGGGLHPPDPTMKIPRLYQAEELLEAAKRNIIIRADTGTGKTFVALSLITWIAAQSPANHDDHRIQAFLAPTRPLAHQQAQAY
IQKHCTLRVKAYTGDLPQLWNIKWHSELNEVDVIVSTAQVSLLI FDEAHHCRKNHIYNQIMRSHYHRLAKDPTVRLPKILGLTASPI
WNYKDLERADSDIKSLQSAALAAQIYEVKTHTEVDCQHNFKNPKVVFYFEPSEFEKNSHPPWDQINQLLSLHASPKMIAAMESVSLELG
TYAHS LAVLDWLKSLLTVGASNQAMPGRLLDPNKQKQIREVIQELEELVNIDDI PETQFSSKVAVLNKILVSYKEKDNHDFLCIVFVE
RRQHAQLLPFLLERNAQLKGFRLPTALTGHAGNVNDLIGIKMSRTQNKAVAKFRGTGEYNLT IATSVAEEGLDFRSCRVVIRFDLITTT
WKGYIQSRGRARARESDYIVMLPNGTNNKYLEFTGKEEQKAAALYNRPEDELIEEGEEWTPQLICQLAGGKESILTYSAATSLNDVVC
QLIPPDEFPLPVVAPQYEITWLGDNFQCQVTLPPMAALHPSQRTFTGLAMATKKDAKRSAAFEACKVLRDLVNLQHFPLPQREGKSAQIC
DADGREIEATPLSDQVEAII PNVYGDFRSTETIWLHKFSFPDDSPDGFSTMGLLCARHLTVPDGLQFDHYTDSRPLPITIEQSKRIQW
GQDDAPTNLQRLETFSRVVMQAAINRKAYEGKLYFLVAPLLRDTCEIDWNLVDTPMIPLSDTADSLRYQNTIAPIRHLHYRIFDTCEPA
GDI SEASPQSV PACPSMRDFCKKISKFHNLGHFYKVYDLKKEEQFQGELVYLETTFHVLNNLSKSESTVVQPHRILLPLKLCGTHIP
RSMWKVFSYLP SLTRLLHDSLQATTLFKRLDFPTISLHGIQALT PPGGGVPWDYQTLLETLGDAFLKLATS VHVYLSHLKKGEGDMSHV
RSRSDVNAYLRRKAIQANLPASILSQRFRTRDRFRDPQTEDGKELPNGNFSRKI PKRVLSDVVEALLGAGFLTGGIELGLKIGTALDLCF
GGTAPWSERPVNI GLESI THDALEPSTLLKQALEEKIGYVFKEKLLVQALTHRANSFMTN CYEREWLGDVIDMWIVEHAYKRFD
HATAEELTLARAKVVSNGSLGFLALKKLGLQEIMHASENFEQACTEAIEAIKPFKIEEYFSTIDNLFVVFDPPKIINDVLEAIVGAV
FVDSGFNLQSAYRTLDIIFEDVIPGLSRLVARDPLSTMLRLRDQYQCAELRRISEPNPNGETKDPISVKVCRIELHGQEIASGRHKSSA
SVAEQRASLEALKVLQEPPESEPTSHSVWSTCQCKTLAVAATTLASSNLAASKV