SUPPLEMENTARY TABLES

Predicting N-terminal Myristoylation Sites in Plant Proteins

Sheila Podell and Michael Gribskov

San Diego Supercomputer Center and Department of Biology, University of California San Diego, 9500 Gilman Drive,
La Jolla, CA 92093-0537

Contents

Supplementary Table 1. Plant proteins used in myristoylation classified positive set.

Sequence (Classified Positives)	ID	Description	Species	Basis	Reference
Kinases - Calcium Dependent					
MGNTCVGPSRNGFLQSVSAAMWRPR	gi|15237928|ref|NP_197815.1|	CPK 01	Arabidopsis thaliana	homology + memb localiz	First 7 residues identical to Curcubita pepo cpCPK1. Damman et al 2003
MGNACVGPNISGNGFLQTVTAAMWR	gi|15227920|ref|NP_179379.1|	CPK 02	Arabidopsis thaliana	direct + peptide	Lu & Hrabak 2002, BGM
MGHRHSKSKSSDPPPSSSSSSSGNV	gi|15234656|ref|NP_193925.1|	CPK 03	Arabidopsis thaliana	peptide	BGM, Qi
MGNSCRGSFKDKLDEGDNNKPEDYS	gi|15236560|ref|NP_194096.1|	CPK 05	Arabidopsis thaliana	direct	Hrabak et al Plant Phys 132:666-680 2003
MGNSCRGSFKDKIYEGNHSRPEENS	gi|30679935|ref|NP_565411.2|	CPK 06	Arabidopsis thaliana	homology	First 7 residues identical to CPK5
MGNCCGNPSSATNQSKQGKPKNKNN	gi|18416872|ref|NP_568281.1|	CPK 07	Arabidopsis thaliana	homology + memb localiz	Damman et al 2003
MGNCCASPGSETGSKKGKPKIKSNP	gi|30687323|ref|NP_850853.1|	CPK 08	Arabidopsis thaliana	homology + memb localiz	Damman et al 2003
MGNCFAKNHGLMKPQQNGNTTRSVE	gi|15230321|ref|NP_188555.1|	CPK 09	Arabidopsis thaliana	homology + memb localiz	Damman et al 2003
MGNCNACVRPDSKESKPSSKPKKPN	gi|28416563|gb|AAO42812.1|	CPK 10	Arabidopsis thaliana	homology	First 7 reisdues almost identical to CPK30 & CPK33
MGNCCGTAGSLIQDKQKKGFKLPNP	gi|15227001|ref|NP_180459.1|	CPK 14	Arabidopsis thaliana	homology	First 7 residues identical to CPK32
MGCFSSKHRNTESDIINGSVQSSIP	gi|15234441|ref|NP_192383.1|	CPK 15	Arabidopsis thaliana	homology + memb localiz	First 7 residues identical to CPK 21
MGLCFSSAAKSSGHNRSSRNPHPHP	gi|15227525|ref|NP_181133.1|	CPK 16	Arabidopsis thaliana	homology + memb localiz	Damman et al 2003
MGNCCSHGRDSADNGDALENGASAS	gi|15239716|ref|NP_197437.1|	CPK 17	Arabidopsis thaliana	homology	First 7 residues identical to CPK34. (CPK17)
MGLCFSSPKATRRGTGSRNPNPDSP	gi|15234174|ref|NP_195066.1|	CPK 18	Arabidopsis thaliana	homology	First 7 residues identical to CPK16

Sequence (Classified Positives)	ID	Description	Species	Basis	Reference
MGNTCVGPNLNPNGFLQSVSAAVWR	gi|15224978|ref|NP_181425.1|	CPK 20	Arabidopsis thaliana	homology	First 7 residues identical to Curcubita pepo cpCPK1
MGCFSSKHRKTQNDGGEKSIPINPV	gi|15234432|ref|NP_192380.1|	CPK 21	Arabidopsis thaliana	homology	Damman et al 2003
MGNCCGSKPLTASDIVSDQKQETIL	gi|15234430|ref|NP_192379.1|	CPK 22	Arabidopsis thaliana	homology	First 7 residues almost identical to CPK32
MGCFSSKHRKTQNDGGGERSIPIIP	gi|15234435|ref|NP_192381.1|	CPK 23	Arabidopsis thaliana	homology	First 7 residues identical to CPK 21
MGSCVSSPLKGSPFGKRPVRRRHSS	gi|15225092|ref|NP_180708.1|	CPK 24	Arabidopsis thaliana	homology	First 7 residues almost identical to StCPK1
MGCFSSKELQQSKRTILEKPLVDIT	gi|15234291|ref|NP_195331.1|	CPK 27	Arabidopsis thaliana	homology	First 6 residues identical to CPK 21
MGVCFSAIRVTGASSSRRSSQTKSK	gi|34098881|gb|AAQ56823.1|	CPK 28	Arabidopsis thaliana	homology + memb localiz	Damman et al 2003
MGNCIACVKFDPDNSKPNQKKKPPR	gi|30699042|ref|NP_177612.2|	CPK 30	Arabidopsis thaliana	homology	First 7 residues almost identical to CPK33 & CPK10
MGCYSSKNLKQSKRTILEKPFVDIG	gi|22328350|ref|NP_680596.1|	CPK 31	Arabidopsis thaliana	homology	First 7 residues identical to membrane localized CPK 21,23 (CPK 31)
MGNCLAKKYGLVMKPQQNGERSVEI	gi|15223629|ref|NP_175485.1|	CPK 33	Arabidopsis thaliana	homology	First 7 almost identical to CPK30 & CPK10
MGNCCSHGRDSDDNKEEPRPENGGG	gi|15238353|ref|NP_196107.1|	CPK 34	Arabidopsis thaliana	homology	First 7 residues identical to CPK17
MGGCFSKKEYQADGNGYRSAPTAYH	gi|7434365|pir||T14335	CDPK	Daucus carota	peptide + homology	Qi. First 7 residues identical to L.esculentum CAC87494.1
MGGCFSKKYRQEGANGGYRATRRN	gi|3283996|gb|AAC25423.1|	CDPK (NtCDPK1)	Nicotiana tabacum	peptide + homology + memb localiz	Qi, First 7 residues identical to L.esculentum CAC87494.1. Yoon et al 1999
MGGCFSKKYTQQDANGHRAGRRVNQ	gi|19171502|emb|CAC87494.1|	CDPK (LeCDPK1)	Lycopersicon esculentum	direct	Rutschmann et al. 2002
MGICASKGKPNNANNGHHGSGSGGV	gi|14029712|gb|AAK52801.1|AF363784_1	CDPK	Lycopersicon esculentum	homology	First 7 residues identical to L. esculentum LeCDPK1
MGICASKNKATQPEDNGYTPANGVG	gi|4336426|gb|AAD17800.1|	CDPK	Mesembryanthemum crystallinum	homology	First 7 residues identical to L. esculentum LeCDPK1

Sequence (Classified Positives)	**ID**	**Description**	**Species**	**Basis**	**Reference**
MGNRTSRHKHGEKHRSEAGSHQTHI	gi|7434369|pir||T10938	CDPK	Ipomoea batatas	homology + memb localiz	Ohto et al. 1995
MGNSCVGAAGYFQGFTSVIALGGRS	gi|2315983|gb|AAB70706.1|	CDPK	Tortula ruralis	homology	First 7 residues almost identical to Curcubita pepo cpCPK1
MGNTCVGPSISKNGFFESVSAAMWR	gi|7434364|pir||T09940	CDPK	Cucurbita pepo	direct	Ellard-Ivey et al
MGNTCVGPSISKNGIFQSVSAAMWR	gi|16215471|emb|CAC82999.1|	CDPK (NtCDPK3)	Nicotiana tabacum	homology	First 7 residues identical to Curcubita pepo cpCPK1
MGNTCVGPSISKNGLFQSVSAAMWR	gi|16215467|emb|CAC82998.1|	CDPK (NtCDPK2)	Nicotiana tabacum	homology	First 7 residues identical to Curcubita pepo cpCPK1
MGNTCVGPSITMNGFFQSVSTALWK	gi|7428007|pir||T02784	CDPK	Zea mays	homology	First 7 residues identical to Curcubita pepo cpCPK1
MGNTCVGPSSAADRHGFFHSVSLAV	gi|33146743|dbj|BAC79646.1|	CDPK	Oryza sativa	homology	First 7 residues identical to Curcubita pepo cpCPK1
MGSCCSRATSPDSGRGGANGYGYSH	gi|23616997|dbj|BAC20693.1|	CDPK (OSCPK2)	Oryza sativa	direct	Martin & Busconi 2000.
MGVCLSKSKPAESKSDGHYRSGGSD	gi|10568116|gb|AAD28192.2|	CDPK (StCDPK1)	Solanum tuberosum	direct	Raices et al 2001
MGHCYSRNISTVDDDDEIPSATAQL	gi|15228912|ref|NP_188935.1|	CRK6	Arabidopsis thaliana	peptide	BGM
Kinases - non CDPK					
MGCVLCKESTGDKRKHNNPDEPPPA	gi|15241289|ref|NP_199899.1|	KINASE - cyclin	Arabidopsis thaliana	peptide	BGM
MGSKYSKATNSINDASNLSYGIPFE	gi|8547239|gb|AAF76314.1|	KINASE - Fen (Pto)	Lycopersicon esculentum	peptide	Qi .
MGSKYSKATNSINDASNSSYRVPFE	gi|16326617|gb|AAL17825.1|	KINASE - Fen (Pto)	Solanum nigrum	peptide + homology	Qi. First 14 residues identical in Capsicum annuum, Lycopersicon pimpinellifolium,Lycopersicon hirsutum
MGSKYSKATNSISDASNSRYGVPFE	gi|8547235|gb|AAF76310.1|AF220603_2	Kinase - Fen(Pto) LescPth3	Lycopersicon esculentum	peptide	Qi.
Phosphatases					
MGSGASKNTEEDDDGSNGGGGQLYV	gi|15222304|ref|NP_172191.1|	fructose-2,6-bisphosphatase	Arabidopsis thaliana	peptide	BGM

Sequence (Classified Positives)	**ID**	**Description**	**Species**	**Basis**	**Reference**
MGLKLSRGPVKEKSPLEFTRVHILT	gi|30693398|ref|NP_198756.2|	PTEN-related protein	Arabidopsis thaliana	peptide	BGM
Calcium Binding proteins					
MGCSVSKKKKKNAMRPPGYEDPELL	gi|3309575|gb|AAC26110.1|	SOS3	Arabidopsis thaliana	direct + peptide	Ishitani etal, BGM
MGGSSSKESPRGGGSGRRYERSVSG	gi|18408863|ref|NP_564907.1|	COPINE	Arabidopsis thaliana	peptide	BGM
Transription Factors					
MGCAQSKIENEEAVTRCKERKQLMK	gi|25405541|pir||D96563	bZIP	Arabidopsis thaliana	peptide	BGM
MGCTASKVEQEDTVRRCKERRRHMK	gi|13124871|gb|AAK01315.2|	bZIP	Oryza sativa	peptide	BGM
MGCTASKLDSEDAVRRCKERRRLMK	gi|15217620|ref|NP_171713.1|	proline-rich	Arabidopsis thaliana	peptide	BGM
Peroxidases					
MGGALSTVFGSGEDATAAGTESEPS	gi|15242525|ref|NP_198811.1|	thioredoxin	Arabidopsis thaliana	peptide	BGM
GTP-binding proteins					
MGAFMSRFWFMMFPAKEYKIVVVGL	gi|15228464|ref|NP_186962.1|	ARF	Arabidopsis thaliana	peptide	BGM
MGARFSRIAKRFLPKSKVRILMVGL	gi|15226086|ref|NP_179133.1|	ARF	Arabidopsis thaliana	peptide	BGM
MGILFTRMFSSVFGNKEARILVLGL	gi|30682545|ref|NP_850057.1|	ARF	Arabidopsis thaliana,Brassica rapa	peptide	BGM
MGLNFTKLFSRLFAKKEMRILMVGL	gi|15241463|ref|NP_196971.1|	ARF	Arabidopsis thaliana	peptide	BGM
MGLSFAKLFSRLFAKKEMRILMVGL	gi|30698721|ref|NP_850975.1|	ARF	Arabidopsis thaliana	peptide + homology	BGM, First 25 residues identical in Gossypium hirsutum
MGLSFGKLFSRLFAKKEMRILMVGL	gi|1703374|sp|P51821|ARF_CHLRE	ARF	Arabidopsis thaliana	peptide + homology	BGM, First 25 residues identical in Capsicum annuum
MGLSFTKLFSRLFAKKEMRILMVGL	gi|11131023|sp|O48649|ARF_SALBA	ARF	Daucus carota	homology	First 25 residues identical in Vigna unguiculata,Catharanthus roseus,Populus tremuloides,Salix bakko
MGLSFTKLLGRLFSKKEMRILMVGL	gi|29124981|gb|AAO63780.1|	ARF	Populus tremuloides	homology	First 25 residues identical in Salix bakko

Sequence (Classified Positives)	**ID**	**Description**	**Species**	**Basis**	**Reference**
MGLTFTKLFSRLFAKKEMRILMVGL	gi|32129330|gb|AAP73857.1|	ARF	Oryza sativa	homology	First 25 residues identical in Hordeum vulgare,Nicotiana tabacum, Triticum aestivum, Zea mays,Solanum tuberosum
MGLTISKLFSRLFAKKEMRILMVGL	gi|1703376|sp|P51824|ARF_SOLTU	ARF	Solanum tuberosum	homology	First 7 residues almost identical to AAP73857.1
MGQAFRKLFDTFFGNQEMRVVMLGL	gi|15237901|ref|NP_197208.1|	ARF	Arabidopsis thaliana	peptide	BGM
MGQAFRKLFDAFFGNKEMRVVMLGL	gi|31433708|gb|AAP55187.1|	ARF	Oryza sativa	peptide	BGM
MGQTFRKLFDTFFGNQEMRVVMLGL	gi|28827546|gb|AAO50617.1|	ARF	Arabidopsis thaliana	peptide	BGM
MGTTLGKPFAGFFHQEEARIVLFGL	gi|18378965|ref|NP_563652.1|	ARF	Arabidopsis thaliana	peptide	BGM
MGLLCSRSRHHTEDTDENTQAAEIE	gi|15225278|ref|NP_180198.1|	G protein alpha subunit (GPA1)	Arabidopsis thaliana	peptide	BGM
MGCASSLPDRNSGTLSGLSNSENAV	gi|18410144|ref|NP_567008.1|	Rab GTPase (Ara6)	Arabidopsis thaliana	direct + peptide	Ueda et al., BGM
Disease Resistance					
MGGCFSVSLPCDQVVSQFSQLLCVR	gi|15221252|ref|NP_172686.1|	CC-NBS-LRR	Arabidopsis thaliana	homology	First 6 residues identical to L.esculentum CAC87494.1
MGGCVSVQVSCDQLLNHLGRCFCRK	gi|15221280|ref|NP_172693.1|	CC-NBS-LRR	Arabidopsis thaliana	peptide	BGM
MGGCVSVSISCDQLTKNVCSCLNRN	gi|30697771|ref|NP_201107.2|	CC-NBS-LRR	Arabidopsis thaliana	peptide	BGM
MGGCVSVSLSCDREVNQFSQWLCVS	gi|15221250|ref|NP_172685.1|	CC-NBS-LRR - RFL1	Arabidopsis thaliana	peptide	BGM
MGISFSIPFDPCVNKVSQWLDMKGS	gi|15221520|ref|NP_176451.1|	CC-NBS-LRR	Arabidopsis thaliana	peptide	BGM
MGSVMSLGCSKRKATNQDVDSESRK	gi|30692151|ref|NP_190034.2|	TIR-NBS-LRR	Arabidopsis thaliana	peptide	BGM
MGSVMSLGCSKRKATSQDVDSESRK	gi|11357250|pir||T47442	dis_resist	Arabidopsis thaliana	peptide	BGM
Misc					
MGAAGSKLEKALGDQFPEGERYFGF	gi|18420557|ref|NP_568074.1|	ubiquitin-specific protease	Arabidopsis thaliana	peptide	BGM
MGANHSREDLELSDSESESEYGSES	gi|7488984|pir||T07737	Dem protein	Lycopersicon esculentum	peptide	Qi
MGASHSHEDLEICSSDEDEYEEYEE	gi|15234174|ref|NP_195066.1|	Dem protein	Arabidopsis thaliana	peptide	BGM
MGTSQSREDRISESDTESDSDYYEE	gi|15229785|ref|NP_190622.1|	Dem protein	Arabidopsis thaliana	peptide	BGM

Sequence (Classified Positives)	**ID**	**Description**	**Species**	**Basis**	**Reference**
MGLGASVLTKTLTFAAAFSAFGDPI	gi|13877587|gb|AAK43871.1|AF370494_1	oxoglutarate-malate translocator	Arabidopsis thaliana	peptide	BGM

Supplementary Table 2 - Plant proteins used in myristoylation classified negative set

Sequence (Classified Negatives)	ID	Function	Species	Basis for neg
MGFSFSSNRFGYLTVTFLLVISCLL	gi|15238663|ref|NP_197287.1|	cation exchanger, putative (CAX8)	Arabidopsis thaliana	BGM negative peptide GFSFSSNR
MGKKKSDESAATTKVKPSGKDASKD	gi|18410084|ref|NP_567001.1|	ABC transporter family protein	Arabidopsis thaliana	BGM negative peptide GKKKSDES
MGLLEAFLNWLRSLFFKQEMELSLI	gi|22331720|ref|NP_190556.2|	ADP-ribosylation factor, putative	Arabidopsis thaliana	BGM negative peptide GLLEAFLN
MGLLSNRIDRSSLKPGDHIYSWRTA	gi|18415178|ref|NP_568167.1|	expressed protein	Arabidopsis thaliana	BGM negative peptide GLLSNRID
MGRSSSKKKKKRGGSGRRGQLKDHG	gi|30694992|ref|NP_191500.2|	protein kinase like	Arabidopsis thaliana	BGM negative peptide GRSSSKKK
MGSDGDKKKKFIVAGSVSGFLVIMV	gi|22330893|ref|NP_187339.2|	pectin methylesterase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	BGM negative peptide GSDGDKKK
MGSEYKHILKSLCLSHGWSYAVFWR	gi|30699554|ref|NP_849917.1|	hypothetical protein [Arabidopsis thaliana]	Arabidopsis thaliana	BGM negative peptide GSEYKHIL
MGVKSVRFSIWFLFVVTDLVFCRTL	gi|18412743|ref|NP_565236.1|	glycosyltransferase family 47]	Arabidopsis thaliana	BGM negative peptide GVKSVRFS
MGAFRFHQYQVVGRALPTEKDVQPK	gi|15226755|ref|NP_180995.1|	60S ribosomal protein L18A (RPL18aB) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGAFRFHQYQVVGRGLPTPTDEHPK	gi|21362867|sp|Q943F3|RL1X_ORYSA	60S ribosomal protein L18a	Oryza sativa	Function annotation
MGAGRKLKQLRITQRWADKHYKKSN	gi|6682246|gb|AAF23298.1|AC016661_23	putative 40S ribosomal protein S23 [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGAMHSRGKGISSSALPYKRTPPTW	gi|419802|pir||S30146	ribosomal protein S13, cytosolic - maize	Zea Mays	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGAPCTWTIGWREFVRRACDAREEG	gi|5257276|dbj|BAA81775.1|	Similar to Mycobacterium tuberculosis cosmid SCY13E10. (Z95324) [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGAPKQKWTSEEEDALRAGVRKHGA	gi|18463961|gb|AAL73044.1|AF461815_1	histone H1-like protein [Zea mays]	Zea mays	Function annotation
MGAPKQKWTSEEEEALRRGVLKHGP	gi|20804653|dbj|BAB92342.1|	histone H1-like protein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGARDEARPPGKRRQNPFPRANLSL	gi|12580855|emb|CAC27136.1|	40S ribosomal protein S2 [Picea abies]	Picea abies	Function annotation
MGAYKYIEELWKKKQSDVMRFLQRV	gi|6093872|sp|O65082|R15B_PICMA	60S RIBOSOMAL PROTEIN L15-2	Picea mariana	Function annotation
MGAYKYIEELWRRKQSDVLRFLLRV	gi|25295572|pir||JC7521	ribosomal protein L15, large subunit - Tortula ruralis	Tortula ruralis	Function annotation
MGAYKYVAEMERRKQSDVLRFLLRV	gi|6094016|sp|O82712|RL15_QUESU	60S RIBOSOMAL PROTEIN L15	Pinus taeda	Function annotation
MGAYKYVSELWRKKQSDVMRFLQRV	gi|15236049|ref|NP_193470.1|	60S ribosomal protein L15 (RPL15B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGAYKYVSELWRRKQSDVMRFVQRV	gi|28875969|gb|AAO59978.1|	ribosomal protein L15 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGAYTYVSELWRKKQSDVMRFLQRV	gi|6094014|sp|O82528|RL15_PETHY	60S RIBOSOMAL PROTEIN L15	Petunia hybrida	Function annotation
MGCEGDVATLDVCKGKQQIENRWMA	gi|15290088|dbj|BAB63781.1|	putative RNA helicase RH2 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGDARDNEAYEEELLDYEEEDEKVP	gi|18416493|ref|NP_568245.1|	DEAD/DEAH box helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGDFDGEQKELIKKLVNFRMIDGKR	gi|23495376|dbj|BAC19857.1|	Ribosomal protein S7 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGDGEIGCHAQKFFTDIFTTNGIQV	gi|15230721|ref|NP_190141.1|	reverse transcriptase (RNA-dependent DNA polymerase) family protein [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGDIGDLLSEDDEEEDQDYDFLKKK	gi|15242854|ref|NP_201168.1|	DEAD box RNA helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGDIVRQILGWFPGHPASGNGATLS	gi|20521321|dbj|BAB91835.1|	putative ribosomal RNA apurinic site specific lyase [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGDKRKKTFMFIRLVSAAGTGFFYV	gi|15239562|ref|NP_197380.1|	ribosomal protein L33 - like [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGDSGDPFLRNPNAAVQARAKVQNR	gi|30693345|ref|NP_850676.1|	U1 small nuclear ribonucleoprotein 70 (U1-70k) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGDVEPEVAAAGVPKKRTFKKFAFK	gi|2129721|pir||S71259	ribosomal protein S15, cytosolic - Arabidopsis thaliana	Arabidopsis thaliana	Function annotation
MGDYGSGMMRGNPDGGGMQSRTKGQ	gi|31429881|gb|AAP51870.1|	putative U1 small nuclear ribonucleoprotein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGEAEVKDNEVYEEDLVDYEEEVEN	gi|20521462|dbj|BAB91970.1|	putative DEAD BOX RNA helicase [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGEDFQHILRLLNTNVDGKQKIMFA	gi|25553581|dbj|BAC24846.1|	putative ribosomal protein S18 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGEEVEERREAMGEEVEVESEAPTT	gi|22775636|dbj|BAC15490.1|	putative ATP-dependent helicase pitchoune [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGEGGVVVPESVLKKQKRNEEWALV	gi|445613|prf||1909359B	ribosomal protein L7	Solanum tuberosum	Function annotation
MGEIAVGLNKGHQVTKKAGTPRPSR	gi|25295570|pir||JC7579	ribosomal protein L36 - green alga (Enteromorpha prolifera)	Enteromorpha prolifera	Function annotation
MGESRVSSAGERDATMVDIGVKSRP	gi|22331612|ref|NP_680097.1|	reverse transcriptase (RNA-dependent DNA polymerase) family protein [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGETRRHSVDVPITRTLVALRRVRS	gi|15233988|ref|NP_194210.1|	DNA polymerase III like protein [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGFFKKKEKKAEFDVFRLYGLHHPE	gi|11497566|ref|NP_054974.1|	ribosomal protein L22 [Spinacia oleracea]	Spinacia oleracea	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGFKGNRFTWRRGLVESTFVAKRLD	gi|15234647|ref|NP_193294.1|	reverse transcriptase (RNA-dependent DNA polymerase) family protein [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGFKRFVEIGRVALVNYGEDYGKLV	gi|15225356|ref|NP_179635.1|	60S ribosomal protein L14 (RPL14A) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGFKRYVEIGRVALVNYGEDHGKLV	gi|15236981|ref|NP_194439.1|	60S ribosomal protein L14 (RPL14B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGFVKVVKNKQYFKRYQVKFKRRRE	gi|7440750|pir||T12615	ribosomal protein L5 - common sunflower	Helianthus annuus	Function annotation
MGGCDKQGFPMKQGVLTPGRVRLLL	gi|1480022|dbj|BAA11393.1|	putative ribosomal protein [Brassica rapa]	Brassica rapa	Function annotation
MGGESPEDYLKILSLWIRLRHIPVN	gi|22331551|ref|NP_683665.1|	reverse transcriptase (RNA-dependent DNA polymerase) family protein [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGGFRFHQYQVVGRALPTENDEHPK	gi|15231888|ref|NP_188078.1|	60S ribosomal protein L18A (RPL18aC) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGGFVKTQKTNAYHKRFQVKFKRRR	gi|22093592|dbj|BAC06888.1|	putative 60S ribosomal protein L5 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGGFVKTQKTNAYYKRFQVKFKRRR	gi|542157|pir||S39486	ribosomal protein L5 - rice	Oryza sativa	Function annotation
MGGLDGEQKLLIKKLVNFRMKEGKR	gi|13449397|ref|NP_085579.1|	ribosomal protein S7 [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGGLDSEQKQLIKKLVNFHMKEGKR	gi|9838445|ref|NP_064059.1|	ribosomal protein S7 [Beta vulgaris]	Beta vulgaris	Function annotation
MGGPVIERRGGSIGWVLRQNGTWWP	gi|32470660|gb|AAP45186.1|	putative micronuclear linker histone polyprotein [Solanum bulbocastanum]	Solanum bulbocastanum	Function annotation
MGGVGKTKRMRRLKRKRRKMRQRSK	gi|7441210|pir||T06233	ribosomal protein L41 - soybean (fragment)	Glycine max	Function annotation
MGGYFICGGMERLVRILILQKRNYP	gi|22748410|gb|AAN05375.1|	putative RNA polymerase [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGHSNVWNSHPKKYGPGSRLCRVCG	gi|15229844|ref|NP_189987.1|	40S ribosomal protein S29 (RPS29B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGHSNVWNSHPKNYGPGSRVCRVCG	gi|31745220|gb|AAP68880.1|	putative ribosomal protein S29 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGIDLIAGGKSKKTKRTAPKSDDVY	gi|15241061|ref|NP_198137.1|	60S ribosomal protein L18 (RPL18C) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGIDLKAGGKSKKTKRTAPKSNDIY	gi|28798549|gb|AAO46881.1|	60S ribosomal protein [Medicago sativa]	Medicago sativa	Function annotation
MGIDLVAGGRNKKTKRTAPRSDDVY	gi|14595709|gb|AAK70908.1|AC087551_7	cytoplasmic ribosomal protein L18 [Oryza sativa]	Oryza sativa	Function annotation
MGIDLVAGGRNKKTKRTAPRSEDVY	gi|34393858|dbj|BAC83538.1|	putative cytoplasmic ribosomal protein L18 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGIKKGEKVDYSLNQSGDLVEASPY	gi|24476045|gb|AAN62787.1|	Putative DEAD/DEAH box RNA helicase protein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGIRFLQAYTPGTRNRSVSDFSELT	gi|28269737|tpg|DAA00915.1|	TPA: ribosomal protein L2 [Chlamydomonas reinhardtii]	Chlamydomonas reinhardtii	Function annotation
MGIRFYRAHTPGTRNRSVSDFHEIT	gi|5880705|gb|AAD54798.1|AF137379_21	ribosomal protein L2 [Nephroselmis olivacea]	Nephroselmis olivacea	Function annotation
MGIRLYKAYTPGTRNRSISDFKELK	gi|22711952|ref|NP_683843.1|	ribosomal protein L2 [Chaetosphaeridium globosum]	Chaetosphaeridium globosum	Function annotation
MGIRLYKAYTPGTRNRSVLEFNDIT	gi|11466368|ref|NP_038371.1|	ribosomal protein L2 [Mesostigma viride]	Mesostigma viride	Function annotation
MGIRSYRAYTPGTRNRSVSEFTEIN	gi|14548219|sp|O98452|RK2_SPIMX	Chloroplast 50S ribosomal protein L2	Spirogyra maxima	Function annotation
MGISRDSIHKRRATGGKQKMWRKKR	gi|30697195|ref|NP_200732.2|	40S ribosomal protein S8 (RPS8B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGISRDSIHKRRATGGKQKQWRKKR	gi|15241316|ref|NP_197529.1|	40S ribosomal protein S8 (RPS8A) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGISRDSMHKRRATGGKKKAWRKKR	gi|17367479|sp|O81361|RS8_PRUAR	40S ribosomal protein S8	Prunus armeniaca	Function annotation
MGISSLFPVQVAVWHETIGPGGFER	gi|25307616|pir||E85175	ATP-dependent RNA helicase like protein [imported] - Arabidopsis thaliana	Arabidopsis thaliana	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGITTRARGKRRKLQGHQTPPPPPP	gi|22831229|dbj|BAC16087.1|	putative ribosomal RNA apurinic site specific lyase [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGKDTIADIITCIRNADMNRKGTVR	gi|11497562|ref|NP_054970.1|	ribosomal protein S8 [Spinacia oleracea]	Spinacia oleracea	Function annotation
MGKDTIADIITSIRNADMNKKGTVR	gi|6831644|sp|Q42362|RR8_MUSAC	Chloroplast 30S ribosomal protein S8	Musa acuminata	Function annotation
MGKDTIADIITSIRNADMNRKGTVR	gi|7525068|ref|NP_051093.1|	ribosomal protein S8 [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKDTIADIITSIRNADMNRKRTVQ	gi|13518473|ref|NP_084832.1|	ribosomal protein S8 [Lotus corniculatus var. japonicus]	Lotus corniculatus	Function annotation
MGKDTIADLLTSIRNADMNKKGIVR	gi|22296315|dbj|BAC10086.1|	Chloroplast 50S ribosomal protein S8 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGKDTIADLLTSIRNADMNKKGTVR	gi|14017607|ref|NP_114293.1|	ribosomal protein S8 [Triticum aestivum]	Triticum aestivum	Function annotation
MGKDTIANIITYIRNADMNKKGMVQ	gi|24212272|sp|P59033|RR8_PHAAN	Chloroplast 30S ribosomal protein S8	Phaseolus angularis	Function annotation
MGKFMKPGKVVLVLAGRYSGRKAVI	gi|445615|prf||1909362A	ribosomal protein L27	Pisum sativum	Function annotation
MGKGTGSFGKRRNKSHTLCVRCGRR	gi|18401007|ref|NP_566535.1|	60S ribosomal protein L37 (RPL37C) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKREIHFTPVGRQVQRVLEYPLRL	gi|28973767|gb|AAO64199.1|	putative transformer-SR ribonucleoprotein [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKRKQRGGGGGGGGAATEGGAATG	gi|29150367|gb|AAO72376.1|	putative RNA binding ribonucleoprotein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGKTHGMGAARKLKSHRRTQRWADK	gi|6716785|gb|AAF26742.1|AF220539_1	40s ribosomal protein S23 [Euphorbia esula]	Euphorbia esula	Function annotation
MGKTIQVFGFPNGVSAEEVKKFLER	gi|15223906|ref|NP_172932.1|	RNA-dependent RNA polymerase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKTIQVFGFPYLLSAEAVKKFLEN	gi|4138282|emb|CAA09697.1|	RNA-directed RNA polymerase [Nicotiana tabacum]	Nicotiana tabacum	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGKTIQVFGFPYLLSAEVVKSFLEK	gi|7489070|pir||T30819	RNA-directed RNA polymerase (EC 2.7.7.48) - tomato	Lycopersicon esculentum	Function annotation
MGKTKLKPVEDVNSEVVDEVEKAEE	gi|15236218|ref|NP_195217.1|	DEAD/DEAH box helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKTRGMGAARKLKNHRRRQRWADK	gi|1362041|pir||S56673	ribosomal protein S23.e, cytosolic (clone RJ3) - garden strawberry	Fragaria ananassa	Function annotation
MGKTRGMGAGRKLKQLRITQRWADK	gi|18398604|ref|NP_566351.1|	40S ribosomal protein S23 (RPS23A) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKTRGMGAGRKLKRLRINQRWADK	gi|15242574|ref|NP_195916.1|	40S ribosomal protein S23 (RPS23B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKTRGMGAGRKLKTHRRNQRWADK	gi|21671348|dbj|BAC02684.1|	40S ribosomal protein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGKVHGSLARAGKVRGQTPKVAKQD	gi|18399137|ref|NP_565458.1|	40S ribosomal protein S30 (RPS30A) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGKVHGSLARAGKVRGQTPKVAKQE	gi|27527721|emb|CAC86461.1|	putative 40S ribosomal protein S30 [Polytomella sp. Pringsheim 198.80]	Polytomella sp.	Function annotation
MGKVSKADKKIAYDQKLCQLLDDYN	gi|22759727|dbj|BAC10912.1|	putative 60S acidic ribosomal protein P0 [Zinnia elegans]	Zinnia elegans	Function annotation
MGKYKAPGPDGFQPVFYQDCWDIVG	gi|15226038|ref|NP_178766.1|	reverse transcriptase (RNA-dependent DNA polymerase), putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGLDQEDLDLTNDDHFIDKEKLSAP	gi|30694705|ref|NP_199327.2|	DNA-directed RNA polymerase subunit [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGLFLLYSKLQSEAVREAITTIKGK	gi|25294932|pir||A84674	60S ribosomal protein L10A [imported] - Arabidopsis thaliana	Arabidopsis thaliana	Function annotation
MGLVFRPTIHRPRSTAVIRRGMKFI	gi|9049466|dbj|BAA99431.1|	putative 60S acidic ribosomal protein P2A [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGNALPLTDMPLGTAIHNIEITRGR	gi|34394232|dbj|BAC84684.1|	ribosomal protein L2, chloroplast [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGNDERKRPTKKMKYGGKDDQKMKN	gi|18422089|ref|NP_568591.1|	TFIIH basal transcription factor complex helicase XPB subunit, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGNDIIANMITAIRNANLGRAETVE	gi|28202208|ref|NP_777449.1|	ribosomal protein S8 [Anthoceros formosae]	Anthoceros formosae	Function annotation
MGNDTIANLITSIRNADMVEKGTVR	gi|29565645|ref|NP_817224.1|	ribosomal protein S8 [Pinus koraiensis]	Pinus koraiensis	Function annotation
MGNDTIANMITSIRNANLEKTKTVQ	gi|34501440|ref|NP_904227.1|	ribosomal protein S8 [Physcomitrella patens subsp. patens]	Physcomitrella patens	Function annotation
MGNDTIANMITSIRNANLGKIKTVQ	gi|11466739|ref|NP_039335.1|	ribosomal protein S8 [Marchantia polymorpha]	Marchantia polymorpha	Function annotation
MGNDTITNLITSIRNADMVEKGTVR	gi|7524688|ref|NP_042442.1|	ribosomal protein S8 [Pinus thunbergii]	Pinus thunbergii	Function annotation
MGNGERGRPNKKMKYGGKDDQKMKN	gi|18422091|ref|NP_568592.1|	TFIIH basal transcription factor complex helicase XPB subunit, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGNSKKMLGDEKEGTSAIPGFNQIQ	gi|2196464|emb|CAA74024.1|	DNA-dependent RNA polymerase subunit beta [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGNVIKSQVGAQALGKNVRVSVTKM	gi|30352072|ref|NP_848099.1|	ribosomal protein L22 [Adiantum capillus-veneris]	Adiantum capillus-veneris	Function annotation
MGPLRPRLRSRRDRHGVVRRRLPRR	gi|34393244|dbj|BAC83094.1|	putative 60S acidic ribosomal protein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGPPDVVPSEHVEFDFSDVFGPSPT	gi|7649389|emb|CAB89082.1|	S6 ribosomal protein kinase [Asparagus officinalis]	Asparagus officinalis	Function annotation
MGPSLQLQFHYEDILRQELLLKLNY	gi|31322691|gb|AAP33162.1|	ribosomal protein L5 [Cucumis sativus]	Cucumis sativus	Function annotation
MGQKIHPLGFRLGITQQHRSTWFAP	gi|5880703|gb|AAD54796.1|AF137379_19	ribosomal protein S3 [Nephroselmis olivacea]	Nephroselmis olivacea	Function annotation
MGQKIHPLGFRLGTTQNHCSYWFAQ	gi|18860350|ref|NP_569667.1|	ribosomal protein S3 [Psilotum nudum]	Psilotum nudum	Function annotation
MGQKIHPLGFRLGVTQEHLSNWFAR	gi|27734479|sp|O98455|RR3_SPIMX	Chloroplast 30S ribosomal protein S3	Spirogyra maxima	Function annotation
MGQKIHPLGFRLGVTQEHRAHWFAK	gi|11466365|ref|NP_038368.1|	ribosomal protein S3 [Mesostigma viride]	Mesostigma viride	Function annotation
MGQKINPLGFRLGENQSHHSLWFAQ	gi|32480881|ref|NP_862792.1|	ribosomal protein S3 [Calycanthus floridus var. glaucus]	Calycanthus floridus	Function annotation

<table>
<tr><th>Sequence (Classified Negatives)</th><th>ID</th><th>Function</th><th>Species</th><th>Basis for neg</th></tr>
<tr><td>MGQKINPLGFRLGITQNHRSYWFAN</td><td>gi|11466742|ref|NP_039338.1|</td><td>ribosomal protein S3 [Marchantia polymorpha]</td><td>Marchantia polymorpha</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQGHHSLWFSQ</td><td>gi|28261755|ref|NP_783269.1|</td><td>ribosomal protein S3 [Atropa belladonna]</td><td>Atropa belladonna</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQKHHSFWFAQ</td><td>gi|14017610|ref|NP_114296.1|</td><td>ribosomal protein S3 [Triticum aestivum]</td><td>Triticum aestivum</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQNHHSFWFAQ</td><td>gi|11467230|ref|NP_043062.1|</td><td>ribosomal protein S3 [Zea mays]</td><td>Zea mays</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQSHDSIWFAQ</td><td>gi|13518476|ref|NP_084835.1|</td><td>ribosomal protein S3 [Lotus corniculatus var. japonicus]</td><td>Lotus corniculatus</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQSHHSFWFAQ</td><td>gi|11466969|ref|NP_054390.1|</td><td>ribosomal protein S3 [Epifagus virginiana]</td><td>Epifagus virginiana</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQSHHSLWFAQ</td><td>gi|7525071|ref|NP_051096.1|</td><td>ribosomal protein S3 [Arabidopsis thaliana]</td><td>Arabidopsis thaliana</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQSHRSLWFAQ</td><td>gi|34500952|ref|NP_904137.1|</td><td>ribosomal protein S3 [Amborella trichopoda]</td><td>Amborella trichopoda</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGTTQSHYSLWFSQ</td><td>gi|11497565|ref|NP_054973.1|</td><td>ribosomal protein S3 [Spinacia oleracea]</td><td>Spinacia oleracea</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGVTQNHCSHWFAK</td><td>gi|28202211|ref|NP_777452.1|</td><td>ribosomal protein S3 [Anthoceros formosae]</td><td>Anthoceros formosae</td><td>Function annotation</td></tr>
<tr><td>MGQKINPLGFRLGVTQNHHSYWFAQ</td><td>gi|34501443|ref|NP_904230.1|</td><td>ribosomal protein S3 [Physcomitrella patens subsp. patens]</td><td>Physcomitrella patens</td><td>Function annotation</td></tr>
<tr><td>MGQKSNPQALRTAAPYQQHESCWYH</td><td>gi|17222576|gb|AAL36749.1|AF353999_29</td><td>ribosomal protein S3 [Mesostigma viride]</td><td>Mesostigma viride</td><td>Function annotation</td></tr>
<tr><td>MGQKTNPISLRLQNVNRNFDSCWYS</td><td>gi|11497487|ref|NP_042277.1|</td><td>ribosomal protein S3 [Prototheca wickerhamii]</td><td>Prototheca wickerhamii</td><td>Function annotation</td></tr>
<tr><td>MGQKVHPFGFRVGITKKHQSQWFAK</td><td>gi|632229|pir||S40460</td><td>ribosomal protein S3 - Chlamydomonas frankii chloroplast (fragment)</td><td>Chlamydomonas frankii</td><td>Function annotation</td></tr>
<tr><td>MGQKVHPIGFRLGITQKHRSYWCTT</td><td>gi|7524928|ref|NP_045930.1|</td><td>ribosomal protein S3 [Chlorella vulgaris]</td><td>Chlorella vulgaris</td><td>Function annotation</td></tr>
<tr><td>MGQKVHPLGFRVGITKKHQSQWFAR</td><td>gi|28269769|tpg|DAA00947.1|</td><td>TPA: ribosomal protein S3 [Chlamydomonas reinhardtii]</td><td>Chlamydomonas reinhardtii</td><td>Function annotation</td></tr>
<tr><td>MGQKVHPSGFRVGITKKHQSQWFAR</td><td>gi|1173249|sp|P46307|RR3_CHLEU</td><td>Chloroplast 30S ribosomal protein S3</td><td>Chlamydomonas eugametos</td><td>Function annotation</td></tr>
</table>

Sequence (Classified Negatives)	ID	Function	Species	Basis for neg
MGQKVNPISLRLPINRSVDCLWFSD	gi|6066169|gb|AAF03187.1|AF110138_19	ribosomal protein S3 [Nephroselmis olivacea]	Nephroselmis olivacea	Function annotation
MGQKVNPLGFRLGVTQEHHSYWFAT	gi|22711972|ref|NP_683840.1|	ribosomal protein S3 [Chaetosphaeridium globosum]	Chaetosphaeridium globosum	Function annotation
MGQYTINLRESYPVWKIGSKIDSTK	gi|7388083|sp|O98462|RPOA_SPIMX	DNA-directed RNA polymerase alpha chain (PEP) (Plastid-encoded RNA polymerase alpha subunit) (RNA polymerase alpha subunit)	Spirogyra maxima	Function annotation
MGRDTIADIITSIRNVDMNRKGTVR	gi|13518369|ref|NP_084728.1|	ribosomal protein S8 [Oenothera elata subsp. hookeri]	Oenothera elata	Function annotation
MGRDTIADILTSIRNANMDKKGTVR	gi|32480878|ref|NP_862789.1|	ribosomal protein S8 [Calycanthus floridus var. glaucus]	Calycanthus floridus	Function annotation
MGRDTIAEIITSIRNADMDRKRVVR	gi|28261752|ref|NP_783266.1|	ribosomal protein S8 [Atropa belladonna]	Atropa belladonna	Function annotation
MGRDTIANIITCIRNADMDKKRMVR	gi|34500949|ref|NP_904134.1|	ribosomal protein S8 [Amborella trichopoda]	Amborella trichopoda	Function annotation
MGRDTILEIINSIRNADRGRKRVVR	gi|11466967|ref|NP_054388.1|	ribosomal protein S8 [Epifagus virginiana]	Epifagus virginiana	Function annotation
MGRGKTPNLGEFRVATDQSASRKIS	gi|26399124|sp|P59152|RR4_WOOUN	Chloroplast 30S ribosomal protein S4	Woodwardia unigemmata	Function annotation
MGRKIHPLGFRLGVSXKHYSYRFAQ	gi|30352071|ref|NP_848098.1|	ribosomal protein S3 [Adiantum capillus-veneris]	Adiantum capillus-veneris	Function annotation
MGRLAIRPNRSRLTWLRENKNTHPI	gi|19571093|dbj|BAB86518.1|	ribosomal RNA apurinic site specific lyase-like protein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGRLHSKGKGISSSALPYRRTAPSW	gi|480095|pir||S36423	ribosomal protein S13, cytosolic - garden pea	Pisum sativum	Function annotation
MGRLIRAQRKGAGSVFKSHTHHRKG	gi|7440614|pir||S22641	ribosomal protein L2, cytosolic - common tobacco	Nicotiana tabacum	Function annotation
MGRMHSRGKGISASALPYKRSSPSW	gi|18411716|ref|NP_567104.1|	40S ribosomal protein S13 (RPS13A) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRMHSRGKGISASALPYKRTPPSW	gi|8131699|dbj|BAA96366.1|	cytoplasmic ribosomal protein S13 [Panax ginseng]	Panax ginseng	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGRMHSSGKGMSCSVLPYRRAAPAW	gi|34393549|dbj|BAC83147.1|	putative 40S RIBOSOMAL PROTEIN S13 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGRMKKIFFNQKSDLNSMKLFFCNR	gi|11466380|ref|NP_038383.1|	RNA polymerase beta" subunit [Mesostigma viride]	Mesostigma viride	Function annotation
MGRRILNDALRTIVNAEKRGKASVE	gi|15224834|ref|NP_179562.1|	40S ribosomal protein S15A (RPS15aB) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRRILNDALRTIVNAERRGKASVE	gi|15233571|ref|NP_194672.1|	40S ribosomal protein S15A (RPS15aE) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRRPARCNRQIKNKPYPKSRYCRG	gi|1076752|pir||S49596	ribosomal protein L10.e, cytosolic - rice	Oryza sativa	Function annotation
MGRRPARCYRQIKGKPYPKSRYCRG	gi|30683726|ref|NP_563945.2|	60S ribosomal protein L10 (RPL10A)/Wilm's tumor suppressor protein-related [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRRPARCYRQIKNKPYPKSRFCRG	gi|18203270|sp|Q9M5M7|RL10_EUPES	60S ribosomal protein L10	Euphorbia esula	Function annotation
MGRRPARCYRQIKNKPYPKSRYCRG	gi|6093994|sp|O22431|RL10_PINTA	60S ribosomal protein L10 (Wilm's tumor suppressor homolog)	Pinus taeda	Function annotation
MGRSMLPEQQEDVSRKSKKEKKSKK	gi|34394349|dbj|BAC84904.1|	putative RNA helicase [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGRVIRAHVEGAGSVFKSHTHHRKG	gi|12054507|emb|CAC20221.1|	ribosomal protein L2 [Glycine max]	Glycine max	Function annotation
MGRVIRAQRKGAAGSVFKSHTHHRK	gi|15230421|ref|NP_190687.1|	60S ribosomal protein L8 (RPL8B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRVIRAQRKGAGSVFKSHTHHRKG	gi|15234298|ref|NP_195336.1|	60S ribosomal protein L8 (RPL8C) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRVIRTQRKGRGSIFTSNNTHRKG	gi|6440928|dbj|BAA78597.1|	60S ribosomal protein L2 [Chlamydomonas sp. HS-5]	Chlamydomonas sp.	Function annotation
MGRVRTKTVKKSSRQVIEKYYSRMT	gi|30680641|ref|NP_850765.1|	40S ribosomal protein S17 (RPS17D) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGRVRTKTVKKSSRQVIERYYSKMT	gi|20141760|sp|P49215|RS17_LYCES	40S ribosomal protein S17	Lycopersicon esculentum	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGRVRTKTVKKTSRQVIEKYYSRMT	gi|31432043|gb|AAP53735.1|	contains similarity to 40S ribosomal protein S17 [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGSDTEAEKSIQKEKKKYAITLAPI	gi|15241537|ref|NP_196435.1|	ribosomal protein L7Ae family [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGSEGNMKKSVVTQVSIGGFGESTT	gi|15229153|ref|NP_190519.1|	RNA-dependent RNA polymerase SDE1 [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGSLVQGDDFQHILRLLNTNVDGKN	gi|1350947|sp|P49202|RS18_CHLRE	40S ribosomal protein S18	Chlamydomonas reinhardtii	Function annotation
MGSNDLNTWVSDKLMVLLGFSQTAV	gi|15226961|ref|NP_181077.1|	RNA helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGSQLSNAIDGKDNTKTIGSADPDQ	gi|22328663|ref|NP_680697.1|	RNA helicase -related [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGSRGRMLFDLNELPTEAEEEEAAV	gi|31429906|gb|AAP51895.1|	putative DNA2-NAM7 helicase family protein [Oryza sativa (japonica cultivar-group)]	Oryza sativa	Function annotation
MGSRVPIETIEEDGEFDWEAAVKEI	gi|15219634|ref|NP_174785.1|	DEAD/DEAH box helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGSVKRKSVEESSDSAPPQKVQRED	gi|15218905|ref|NP_176185.1|	ATP-dependent RNA helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGTERKRKVSLFDVMEDPSLSSKNT	gi|15226549|ref|NP_182247.1|	RNA helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGTNEVTRIVTDEEKREAKNFNIFD	gi|15219773|ref|NP_176261.1|	RNA polymerase subunit [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGTVLKVVDNSGAKKVMCIQALKGK	gi|9757736|dbj|BAB08261.1|	50S ribosomal protein L14 [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGVDPFKTTETLEADKETNGGVPVK	gi|15239967|ref|NP_196805.1|	RNA helicase, putative [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGVFSFVCKSKGGEWTAKQHEGDLE	gi|15242089|ref|NP_200539.1|	60S acidic ribosomal protein P3 (RPP3B) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGVFTFVCKNGGGAWSAKQHEGELE	gi|15236029|ref|NP_194319.1|	60S acidic ribosomal protein P3 (RPP3A) [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGVLYRNSFMGYSSAHPHQHFAGEL	gi|3901091|emb|CAA06965.1|	arbuscular mycorrhiza protein [Pisum sativum]	Pisum sativum	Function annotation
MGVSKTGAFRLGLYGNLNVQSSVQE	gi|19912795|dbj|BAB40805.2|	PolI-like DNA polymerase [Oryza sativa]	Oryza sativa	Function annotation

Sequence (Classified Negatives)	**ID**	**Function**	**Species**	**Basis for neg**
MGVSLRHLSPSSFWVSRRPRVSSSI	gi|30685835|ref|NP_188690.2|	DNA polymerase -related [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGVSLTSHNNPLLRHLSPSSSWVSR	gi|25405411|pir||D96545	probable DNA polymerase A family protein [imported] - Arabidopsis thaliana	Arabidopsis thaliana	Function annotation
MGVYTFVCRNNGGEWTAKQHSGEIE	gi|7489709|pir||T02037	acidic ribosomal protein P3a - maize	Zea mays	Function annotation
MGWKAAEKLIRHWKILRGDNVMIIR	gi|22327012|ref|NP_680212.1|	50S ribosomal protein L24, chloroplast precursor -related [Arabidopsis thaliana]	Arabidopsis thaliana	Function annotation
MGYVKVVKTSPYFSRYQVKYRRRRQ	gi|7440744|pir||T08009	probable ribosomal protein L5 - green alga (Dunaliella salina)	Dunaliella salina	Function annotation

Supplementary Table 3. Plant proteins used to supplement training set, to build a second generation of HMM models.

Sequence (Supplementary positive)	**HMM$_{80}$ score**	**ID**	**Description**
MGGCASRPKESDMQNEEGSVPNKPV	17.2	gi|15229163|ref|NP_190523.1|	expressed protein [Arabidopsis thaliana]
MGLCTSKPNSSNSDQTPARNSPLPA	16.1	gi|15229785|ref|NP_190622.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]
MGACFSSHTATAAADGGSGKRQQRK	16.0	gi|6689920|gb|AAF23900.1|AF194413_1	calcium-dependent protein kinase [Oryza sativa]
MGLSFGKLFSKLFAKKEMRILMVGL	15.9	gi|15228723|ref|NP_191788.1|	ADP-ribosylation factor [Arabidopsis thaliana]
MGLTFTKLFSRLFSKKEMRILMVGL	15.4	gi|7242921|dbj|BAA92519.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGLAFGKLFSRLFAKKEMRILMVGL	15.3	gi|27261078|dbj|BAC45192.1|	ADP-ribosylation factor 1 [Oryza sativa (japonica cultivar-group)]
MGGCVSTPKSCVGAKLRSSKRRKSR	15.3	gi|15218522|ref|NP_172512.1|	expressed protein [Arabidopsis thaliana]
MGVCFSKASKPSGNGYQPPPPQPYK	14.9	gi|28866604|emb|CAD70165.1|	calcium-dependent protein kinase [Spirodela punctata]
MGNCCVTPPQTGSPLKNKKNKPNPF	14.6	gi|2665890|gb|AAB88537.1|	calcium-dependent protein kinase [Fragaria x ananassa]
MGSCVSKGKGDDDSVHNVEFSGGNV	14.5	gi|15231958|ref|NP_187483.1|	thioredoxin family [Arabidopsis thaliana]
MGGCCCCSSSRRADVDNGPAYYYYP	14.5	gi|18422087|ref|NP_568590.1|	zinc finger (C3HC4-type RING finger) protein family [Arabidopsis thaliana]
MGQCCSRAPAPDSGRGGTNGYGYSN	14.2	gi|7428011|pir||T02993	calcium-dependent protein kinase (EC 2.7.1.-) 9 - maize
MGNCCVTPEGSGRGRKKQQQEQKQK	14.0	gi|34393291|dbj|BAC83205.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGCVCGKPSAIEDSKDSPRERFSSK	13.9	gi|15219169|ref|NP_175713.1|	cell division-related protein -related [Arabidopsis thaliana]
MGACFSVAISCDQAVNNLTSCLSRN	13.8	gi|15239105|ref|NP_196159.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]
MGLRFTKALSRLFGKKEMRILMVGL	13.7	gi|1703374|sp|P51821|ARF1_CHLRE	ADP-ribosylation factor 1
MGGCTSKPSTSSGRPNPFAPGNDYP	13.7	gi|15222781|ref|NP_175381.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGICVSKPSPEPDLHNHHTSIPVND	13.5	gi|1706130|sp|P53681|CRK_DAUCA	CDPK-related protein kinase (PK421)
MGGCVSTPKSCVGGKIRSSKRRKTR	13.5	gi|18406410|ref|NP_564750.1|	expressed protein CW14 [Arabidopsis thaliana]
MGSKYSKATNSINDASNLSYGVPFE	13.5	gi|8547231|gb|AAF76307.1|	Fen kinase [Lycopersicon pimpinellifolium]
MGCFASRPNDTGGNRRKPTSIGDVS	13.4	gi|15219459|ref|NP_177482.1|	expressed protein [Arabidopsis thaliana]
MGCSHSKLDDEEAVQICKDRKRFIK	13.4	gi|15224715|ref|NP_179499.1|	hypothetical protein [Arabidopsis thaliana]
MGACASKPKESDIVEGSVSTENAVV	13.3	gi|18417019|ref|NP_567777.1|	expressed protein [Arabidopsis thaliana]
MGGCHAKPLTHEEADGACSPPRERP	13.1	gi|33146627|dbj|BAC79915.1|	putative CDPK-related protein kinase [Oryza sativa (japonica cultivar-group)]
MGNVISGGSRPENHRDRTSPPYPNP	13.0	gi|15231808|ref|NP_190909.1|	expressed protein [Arabidopsis thaliana]
MGCAASKLDNEDAVRRCKDRRRLMK	12.9	gi|30695172|ref|NP_191591.2|	bZIP protein [Arabidopsis thaliana]
MGGCVSSQRKLSNKLQQKKQKRGRS	12.8	gi|15223005|ref|NP_172850.1|	expressed protein [Arabidopsis thaliana]
MGLCSSSSARRDAGTPGGGNGAGNK	12.7	gi|12592069|gb|AAF23901.2|AF194414_1	calcium-dependent protein kinase [Oryza sativa]
MGCFASTPKDSGGNRRKPTSIGEVS	12.5	gi|15220908|ref|NP_173236.1|	hypothetical protein [Arabidopsis thaliana]
MGCVSSKILSKSGSFQEKVVSHGFK	12.4	gi|5257280|dbj|BAA81779.1|	Similar to neurofilament triplet M protein. (P07197) [Oryza sativa (japonica cultivar-group)]
MGSCLAKPAESAEQKKASSVPSPPP	12.3	gi|33591148|gb|AAQ23078.1|	gliding motility related CaM kinase [Chlamydomonas reinhardtii]
MGHCCSKGVTADNDGHVVSVVDGNS	12.2	gi|16904222|gb|AAL30818.1|AF435450_1	calcium/calmodulin-dependent protein kinase CaMK1 [Nicotiana tabacum]
MGTNYSKPTTSINDASNLRNRVPFE	12.1	gi|13021867|gb|AAK11569.1|AF318493_1	Pto-like protein kinase D [Lycopersicon hirsutum]
MGCCQSRIDSKEIVSRCKARKRYLK	12.1	gi|15230441|ref|NP_190697.1|	proline-rich protein family [Arabidopsis thaliana]
MGTNYSKPTTSINDASNLSNRVPFE	12.1	gi|8547236|gb|AAF76311.1|AF220603_3	LescPth2 [Lycopersicon esculentum]
MGGCVSVSLSCDRVVNQFSQWLCVS	11.8	gi|22497333|gb|AAL65635.1|	RFL1 [Arabidopsis thaliana]
MGSLCSRNKHYSQADDEENTQTAEI	11.7	gi|3913725|sp|P93564|GBA1_SOLTU	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGCSHSKFDDDEAVQICKDRKRFIK	11.7	gi|15234640|ref|NP_194742.1|	expressed protein [Arabidopsis thaliana]
MGSCVGKERSDEEDKIDFKGGNVHV	11.5	gi|20804820|dbj|BAB92503.1|	thioredoxin-like protein [Oryza sativa (japonica cultivar-group)]
MGTALGKPFAGFFHQEESRIVLFGL	11.5	gi|15217757|ref|NP_171745.1|	ADP-ribosylation factor, putative [Arabidopsis thaliana]
MGNCVGGHSDRKNALNPNEPAATNV	11.3	gi|5162878|dbj|BAA81749.1|	calcium-dependent protein kinase [Marchantia polymorpha]
MGCCLSKKPSPSLPSSVKPSDPIKP	11.3	gi|15218781|ref|NP_176748.1|	hypothetical protein [Arabidopsis thaliana]
MGSKYSKATNSINDALNPSYLVPFE	11.2	gi|32330888|gb|AAP79929.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGARCSKFSFCLFPSHFKSASVLES	11.1	gi|15232406|ref|NP_190971.1|	protein kinase family [Arabidopsis thaliana]
MGLLCSKNRRYNDADTEENTQTAEI	11.1	gi|1346103|sp|P49082|GBA1_LOTJA	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGCVHGRPSTSSPAAANASRRRDHP	11.1	gi|23495876|dbj|BAC20085.1|	putative CRK1 protein(cdc2-related kinase 1) [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSINDALNSSYLVPFE	11.1	gi|32330884|gb|AAP79927.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGACASRPKDLDRDLAPAPVPAEES	11.0	gi|5360192|gb|AAD42902.1|AF161330_1	unknown [Solanum chacoense]
MGNTCVGPSISKIGLFQSVSAAMWR	10.9	gi|16215475|emb|CAC83000.1|	calcium-dependent protein kinase 2 [Nicotiana benthamiana]
MGCAQSRVDNEEAVARCKERRNVIK	10.9	gi|15239473|ref|NP_197941.1|	expressed protein [Arabidopsis thaliana]
MGNCCVSSGTTDRENENVSDKNTTI	10.9	gi|18407049|ref|NP_564773.1|	expressed protein [Arabidopsis thaliana]
MGGCTSKPSSSVKPNPYAPKDAVLQ	10.9	gi|15230288|ref|NP_188541.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]
MGCGVSRPETDEGRQEGSRLGSIGK	10.9	gi|11994158|dbj|BAB01187.1|	gene_id:MQP17.4~unknown protein [Arabidopsis thaliana]
MGASCSSPFPPPPSPAESNAPSASA	10.9	gi|20146453|dbj|BAB89233.1|	pectate lyase-like protein [Oryza sativa (japonica cultivar-group)]
MGASGSKLEKALGDQFPEGERYFGL	10.8	gi|21740732|emb|CAD40853.1|	OSJNBa0086B14.25 [Oryza sativa (japonica cultivar-group)]

Sequence (Supplementary positive)	**HMM$_{80}$ score**	**ID**	**Description**
MGNCFSQSTQTRGSKRNQKPDPLDE	10.8	gi|28866606|emb|CAD70166.1|	putative calcium-dependent protein kinase [Spirodela punctata]
MGCLLSKDKDSEQEHNNGGYRYAES	10.7	gi|13561063|emb|CAA65500.1|	protein kinase [Medicago sativa]
MGNCCVSSGTTDRENENISDKNTTI	10.6	gi|21554122|gb|AAM63202.1|	unknown [Arabidopsis thaliana]
MGQCCSKGVSGENGGSVVAIGDGNS	10.5	gi|19547871|gb|AAL87457.1|	serine/threonine protein kinase pk23 [Lycopersicon esculentum]
MGACLSSSSVIDAKDQKPSSAYVIS	10.5	gi|13699079|dbj|BAB41200.1|	TMV response-related gene product [Nicotiana tabacum]
MGSKYSKATNSINDALSSSYLVPFE	10.5	gi|32330886|gb|AAP79928.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGGCVSTHSKAIRPRRKGRRRSSKH	10.4	gi|18406090|ref|NP_566845.1|	expressed protein [Arabidopsis thaliana]
MGLVCSRNRRYRDSDPEENAQAAEI	10.4	gi|3913709|sp|O04279|GBA2_PEA	Guanine nucleotide-binding protein alpha-2 subunit (GP-alpha-2)
MGNKTSRKSKEKSGRNYTTELRSYE	10.4	gi|15235333|ref|NP_195159.1|	conserved hypothetical common family [Arabidopsis thaliana]
MGCCFSVIDDDGYEHDDAVSGSVNE	10.4	gi|32488315|emb|CAE02892.1|	OSJNBa0015K02.9 [Oryza sativa (japonica cultivar-group)]
MGCSSSVPARSTGGLNNISNDNSAT	10.3	gi|5931625|dbj|BAA84717.1|	rab5B [Oryza sativa (japonica cultivar-group)]
MGLSFTKLLGRLFSEKEMRILMVGL	10.3	gi|29124979|gb|AAO63779.1|	ADP-ribosylation factor 1 [Populus tremuloides]
MGSCFSSEGGNESRKESAERPQITP	10.3	gi|24796805|gb|AAN64481.1|	putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGCCISKCSPKSKDFKEAEEPHEK	10.3	gi|15219009|ref|NP_173570.1|	hypothetical protein [Arabidopsis thaliana]
MGCAASRIDNEEKVLVCRQRKRLMK	10.3	gi|30686165|ref|NP_181014.2|	proline-rich protein family [Arabidopsis thaliana]
MGCFHSKAAKEFRGHEDPVKLASET	10.2	gi|18414931|ref|NP_567533.1|	calcineurin B-like protein 1 (CBL1) [Arabidopsis thaliana]
MGCNCTKGTRPDNDNVDNSNSIVSN	10.1	gi|15217565|ref|NP_172431.1|	protein kinase -related [Arabidopsis thaliana]
MGSKYSKATNSINDALSSGYLVPFE	10.1	gi|32330890|gb|AAP79930.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGCFHSTAAREFPDHENPVKLASET	10.1	gi|15238015|ref|NP_199521.1|	calcineurin B-like protein 9 (CBL9) [Arabidopsis thaliana]

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGLLCSKHQHSTKPDAENAQATGIE	10.0	gi|7441569|pir||T09152	GTP-binding regulatory protein alpha chain - spinach
MGGCASKPKESDIPETEKTVVESKN	10.0	gi|18423589|ref|NP_568803.1|	expressed protein [Arabidopsis thaliana]
MGNCVARSGTAVDAGGDGGEDGKRR	10.0	gi|33146759|dbj|BAC79670.1|	putative protein phosphatase 2C [Oryza sativa (japonica cultivar-group)]
MGCVFGKESLAPEGRERGGRREKEN	10.0	gi|7688002|emb|CAB89665.1|	CRK1 protein [Beta vulgaris subsp. vulgaris]
MGGCISVSLSCDRVVNQISQCLCVS	9.9	gi|22497401|gb|AAL65637.1|	RFL1 [Arabidopsis lyrata]
MGGCVSSNLLTPTDDSSFSHLTSSS	9.9	gi|15223291|ref|NP_174553.1|	expressed protein [Arabidopsis thaliana]
MGNTCRGSFKGNIFQGYSQPEESST	9.8	gi|20453015|gb|AAL68972.1|	calmodulin-like-domain protein kinase CPK2 [Cucurbita maxima]
MGSSASKTASSVSSASSSSAFSPPP	9.8	gi|30694571|ref|NP_567043.2|	expressed protein [Arabidopsis thaliana]
MGGGNSKEESSSPSSSSWASHQSYP	9.7	gi|15232157|ref|NP_186814.1|	copine-related [Arabidopsis thaliana]
MGCGSSLPDRDSRPFGRDNPENGGG	9.7	gi|1370180|emb|CAA98167.1|	RAB5B [Lotus corniculatus var. japonicus]
MGCCFSAGKVTAPPVAAAQIPPPLE	9.7	gi|34146808|gb|AAQ62412.1|	At4g28180 [Arabidopsis thaliana]
MGGVCSCVFKDDDKKKKLRSNDDDK	9.7	gi|18397646|ref|NP_564362.1|	expressed protein [Arabidopsis thaliana]
MGNAGSSPEQAAGNNKEDAEARRQP	9.6	gi|18844859|dbj|BAB85328.1|	OJ1656_A11.17 [Oryza sativa (japonica cultivar-group)]
MGNCLSASTLPTEPNAPKSVEEGVK	9.6	gi|22327062|ref|NP_197926.2|	protein kinase family [Arabidopsis thaliana]
MGNNFSVESPSLAPFLCGKRKYLYN	9.6	gi|15238051|ref|NP_199537.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]
MGTCFSSSTKSTAEISPFDLVVKPP	9.6	gi|18410034|ref|NP_566997.1|	expressed protein [Arabidopsis thaliana]
MGSCFSSRVKADIFHNGKSSDLYGL	9.5	gi|15225520|ref|NP_181496.1|	protein kinase, putative [Arabidopsis thaliana]
MGCISSKNVSCLTDQGDSPLPEPGL	9.4	gi|22327464|ref|NP_198758.2|	cdc2-related protein [Arabidopsis thaliana]
MGCCYSLSSTVDPVQDHTTDASSEP	9.4	gi|15227366|ref|NP_179301.1|	protein kinase family [Arabidopsis thaliana]
MGSCFSSSKVSGSNSNTPSTNNTAT	9.4	gi|3779218|gb|AAC78558.1|	protein kinase CPK1 [Solanum tuberosum]

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGNCFGSEEAEAAATVRAPAQGHHG	9.3	gi|32488580|emb|CAE03088.1|	OSJNBa0017B10.3 [Oryza sativa (japonica cultivar-group)]
MGCFNSKHSSPAGPPRPVRRRLDTN	9.2	gi|22330784|ref|NP_683519.1|	cdc2-related protein [Arabidopsis thaliana]
MGLVCSRNRRYRDSDPGENAQAAEI	9.2	gi|22476942|gb|AAM97352.1|	G protein alpha II subunit [Pisum sativum]
MGLCHSKIDKTTRKETGATSTATTT	9.2	gi|15219170|ref|NP_173072.1|	protein phosphatase 2C -related [Arabidopsis thaliana]
MGCISSKSVSCLTDQGDSPLPEPGL	9.2	gi|18032144|gb|AAL56635.1|AF120153_1	cyclin-dependent kinase CDC2C [Arabidopsis thaliana]
MGLLCSRNRRYNDADAEENAQAAEI	9.0	gi|3334195|sp|Q40224|GBA1_LUPLU	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGAGVSGLFGLGGDEGETSAAVGGA	9.0	gi|31433129|gb|AAP54682.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGAASSRLEKALGEQFPEGERYFGL	9.0	gi|24756890|gb|AAN64154.1|	Putative ubiquitin-specific protease 3 [Oryza sativa (japonica cultivar-group)]
MGSAMSLGCSKRKATNQDVDSESRK	8.9	gi|25453543|pir||T52347	disease resistance protein RPP1-WsB [imported] - Arabidopsis thaliana (fragment)
MGLLCSRNRRYNDADAEENAQTAEI	8.9	gi|3913724|sp|P93163|GBA2_SOYBN	Guanine nucleotide-binding protein alpha-2 subunit (GP-alpha-2)
MGSGFSSLLPCFNQGHRNRRRHSSA	8.8	gi|15226962|ref|NP_181078.1|	expressed protein [Arabidopsis thaliana]
MGNCCSDVASGAGATAGVGGSGSSA	8.7	gi|34365733|gb|AAQ65178.1|	At5g61900 [Arabidopsis thaliana]
MGLLCSKSNRYNDAKAEENAQTAEI	8.7	gi|3913708|sp|O04278|GBA1_PEA	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGCICSKGAAEDEEGVVYHREKANE	8.6	gi|15217643|ref|NP_174637.1|	protein kinase -related [Arabidopsis thaliana]
MGNAAAVPSSCASADAAPSTTPTIK	8.5	gi|13872953|dbj|BAB44058.1|	P0030H07.22 [Oryza sativa (japonica cultivar-group)]
MGLCYSVDRTTGKEPGEASSTATTA	8.4	gi|22330744|ref|NP_178081.2|	protein phosphatase-2C -related [Arabidopsis thaliana]
MGGCVSTSSKSTCSSWSNGEKPVRR	8.4	gi|22328287|ref|NP_680572.1|	protein phosphatase 2C (PP2C) -related [Arabidopsis thaliana]
MGSVMSLSCSKRKATSQDVECSSES	8.4	gi|15219186|ref|NP_173623.1|	Kelch repeat containing F-box protein family [Arabidopsis thaliana]

Sequence (Supplementary positive)	**HMM$_{80}$ score**	**ID**	**Description**
MGAAVAGEPRHNNANGGHREREGIK	8.4	gi|31430970|gb|AAP52815.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGQCYGKARGASSRADHDADPSGAG	8.3	gi|7446432|pir||T02033	calcium/calmodulin-dependent protein kinase homolog - maize
MGNSLRCISQEQDPNQKKPSSVVNG	8.3	gi|15222256|ref|NP_177084.1|	hydrolase, alpha/beta fold family [Arabidopsis thaliana]
MGCVSSNLLNHDEDFSQIGGGSSAF	8.3	gi|15235179|ref|NP_192801.1|	hypothetical protein [Arabidopsis thaliana]
MGNACRGSFRAKTFQGSYPQNHSTS	8.2	gi|31616519|gb|AAP55748.1|	calcium-dependent protein kinase 3 [Capsicum annuum]
MGACLTLSFSCDEVVNQISQGLCIN	8.2	gi|15221277|ref|NP_172692.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]
MGGAFSTSKPKPAAGEEGGESAVVA	8.1	gi|34394553|dbj|BAC83857.1|	putative thioredoxin [Oryza sativa (japonica cultivar-group)]
MGCVLGRPGSSGSVSGSRDEVSTRI	8.1	gi|15229881|ref|NP_187156.1|	cyclin-dependent protein kinase -related [Arabidopsis thaliana]
MGCCGSSLRVGSHAPEKPPRRARPP	8.1	gi|31433335|gb|AAP54864.1|	protein kinase-like protein [Oryza sativa (japonica cultivar-group)]
MGGALSIVFGSGEDATAAGTESSEP	8.0	gi|1388078|gb|AAC49353.1|	thioredoxin h
MGLCFSKKQQVRRRREEQQQPPCHE	8.0	gi|32483146|emb|CAE02138.1|	OSJNBa0074L08.6 [Oryza sativa (japonica cultivar-group)]
MGVCCSKGTGIIVEHGADDGNECGD	7.9	gi|30679372|ref|NP_195790.3|	protein phosphatase 2C (PP2C), putative [Arabidopsis thaliana]
MGCLRSLVRRKQFDSSNGKAETHHH	7.8	gi|30693001|ref|NP_198510.2|	cationic amino acid transporter -related protein [Arabidopsis thaliana]
MGCICSKGVRTNDDYIETNHVSIGK	7.8	gi|15223647|ref|NP_176083.1|	CRK1 protein -related [Arabidopsis thaliana]
MGSCLSSSGGGGSRRSLHGSPHVPG	7.8	gi|15239351|ref|NP_198474.1|	protein phosphatase 2C (PP2C), putative [Arabidopsis thaliana]
MGISFSIPFDPCVNKVSQWLDMKVS	7.7	gi|15221747|ref|NP_176525.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]
MGGCCSSSRKSHLVGTPVYYYCPES	7.7	gi|18411296|ref|NP_567171.1|	zinc finger (C3HC4-type RING finger) protein family [Arabidopsis thaliana]
MGNSFTCISHEQEQRPKKSSGGGGN	7.7	gi|15222699|ref|NP_173960.1|	hydrolase, alpha/beta fold family [Arabidopsis thaliana]

Sequence (Supplementary positive)	**HMM$_{80}$ score**	**ID**	**Description**
MGSKYSKATNSISDASNSFESYRFP	7.6	gi|8547238|gb|AAF76313.1|	Pto kinase [Lycopersicon esculentum]
MGFCFCLSSGGSTDKSQIYEITDYG	7.6	gi|15233997|ref|NP_195021.1|	protein phosphatase 2C (PP2C), putative [Arabidopsis thaliana]
MGSCLSAESRSPRPGSPCSPAFSVR	7.6	gi|15232961|ref|NP_186924.1|	protein phosphatase 2C (PP2C) [Arabidopsis thaliana]
MGGALSTVFGSGEDAAAAGTESSEP	7.5	gi|21595244|gb|AAM66084.1|	thioredoxin [Arabidopsis thaliana]
MGISFSNNNRRRDNNNRRHLHHYPP	7.5	gi|18397483|ref|NP_566274.1|	RING zinc finger protein -related [Arabidopsis thaliana]
MGCCLSKNPPQASSSSLIDGSCSVK	7.4	gi|30693012|ref|NP_198517.2|	serine-rich protein -related [Arabidopsis thaliana]
MGLSHARSGKRCNDFKVGGELSCRG	7.3	gi|13873011|dbj|BAB44115.1|	B1085F09.11 [Oryza sativa (japonica cultivar-group)]
MGNANAREDGAAVDGDGDGEVSGRR	7.3	gi|5702015|emb|CAB52141.1|	GAL83 protein [Solanum tuberosum]
MGQCYGKVNQSKQNGEEEANTATYV	7.3	gi|18087555|gb|AAL58909.1|AF462819_1	At2g46700/T3A4.8 [Arabidopsis thaliana]
MGCAGSTQSQADGSVKKIRKPKPWK	7.3	gi|18404455|ref|NP_564630.1|	expressed protein [Arabidopsis thaliana]
MGGCAGKVRRDDEEKLDFKGGNVHI	7.3	gi|24637227|gb|AAN63617.1|AF435816_1	thioredoxin h-like protein [Zea mays]
MGCCQSLFSGDNPLGKDGVQPQPLS	7.3	gi|18418600|ref|NP_567980.1|	protein kinase family [Arabidopsis thaliana]
MGRSSSKKKKKRGGSGRRGQLKDHG	7.2	gi|30694992|ref|NP_191500.2|	protein kinase like [Arabidopsis thaliana]
MGNLCSLFTPPKPVKKRKPITKRQS	7.2	gi|15230077|ref|NP_189622.1|	hydrolase, alpha/beta fold family [Arabidopsis thaliana]
MGNRTSRHHRAAPEQPPPQPKPKPQ	7.2	gi|12313676|dbj|BAB21081.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGNCLSTSDPSHEDVSKKLPKALPV	7.2	gi|30689674|ref|NP_850414.1|	expressed protein [Arabidopsis thaliana]
MGQCYGKVNQSKQNGEEEANTTTYV	7.1	gi|15226426|ref|NP_182193.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]
MGNCCGTAGSLAQNDNKPKKGRKKQ	7.0	gi|30694663|ref|NP_191312.2|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]
MGCAQSNEDGEGPVARCRERKHLLR	6.9	gi|23237834|dbj|BAC16409.1|	putative bZIP protein [Oryza sativa (japonica cultivar-group)]
MGLCHGKPIEQQSKNLPISNEIEET	6.9	gi|15229002|ref|NP_191235.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]

Sequence (Supplementary positive)	HMM_{80} score	ID	Description
MGCAGSTPKTDDNSKKLKKPKPWKH	6.9	gi|15289832|dbj|BAB63530.1|	P0435H01.26 [Oryza sativa (japonica cultivar-group)]
MGASGSKPEGAGPGPALPTPSPAPP	6.8	gi|9022433|gb|AAF82382.1|AF280543_1	putative copper chaperone Cox17 [Chlamydomonas reinhardtii]
MGACASRPKDLDKDLAPAPVPAEDA	6.8	gi|6019472|gb|AAD02545.2|	PGPS/D3 [Petunia x hybrida]
MGNYVSSALSKTSSSSSSAAKVILP	6.8	gi|15237951|ref|NP_197237.1|	expressed protein [Arabidopsis thaliana]
MGTSVSKPVSDDKKKGTSVSKPVKN	6.5	gi|15217704|ref|NP_174650.1|	AIG1-related protein [Arabidopsis thaliana]
MGCGGSKEAVATGNTVGSTGGRSSR	6.5	gi|20160760|dbj|BAB89701.1|	P0504E02.1 [Oryza sativa (japonica cultivar-group)]
MGGKSSKSKKNVEFGSPSTPVQIKI	6.5	gi|15224198|ref|NP_179453.1|	At14a, putative [Arabidopsis thaliana]
MGNNCVGPSAAGQNGFFANVALWRP	6.4	gi|27819504|gb|AAO24908.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGNCCGWGARIKDGSPQPGASGMFS	6.4	gi|27545044|gb|AAO18450.1|	putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGLLCSRNKGYNQADDEENTQTADI	6.3	gi|18369796|dbj|BAB84093.1|	heterotrimeric GTP binding protein alpha subunit [Nicotiana tabacum]
MGCTSSKLDDLPAVALCRERCAFLE	6.3	gi|15236929|ref|NP_195250.1|	expressed protein [Arabidopsis thaliana]
MGSLCCVAAKSDRSNSTSGDFSFGL	6.3	gi|26451018|dbj|BAC42615.1|	putative mandelonitrile lyase [Arabidopsis thaliana]
MGNANGREDGAIPDAGDPSGREPHA	6.3	gi|32364486|gb|AAO61676.1|	AKIN beta1 [Medicago truncatula]
MGVSASCLWGGSESRGNQNGSAAVT	6.3	gi|29893573|gb|AAP06827.1|	putative receptor ser/thr protein [Oryza sativa (japonica cultivar-group)]
MGSCISLQISCDQVLTRAYSCFFSL	6.3	gi|15236915|ref|NP_192816.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]
MGLSYSSNPTPLDNDQKKPSPATAV	6.3	gi|8885571|dbj|BAA97501.1|	Dof6 zinc finger protein-like [Arabidopsis thaliana]
MGANISGGSPEFDRNDDVYSRKLRL	6.2	gi|30695337|ref|NP_567108.2|	F-box protein (lectin-related) [Arabidopsis thaliana]
MGNKDSKGSSSSTNPKEEIAAVPQP	6.2	gi|15239322|ref|NP_196220.1|	hypothetical protein [Arabidopsis thaliana]
MGNTCRGSFKGNIFQGYSQPEDSST	6.2	gi|20453013|gb|AAL68971.1|	phloem calmodulin-like-domain protein kinase PCPK1 [Cucurbita maxima]

Sequence (Supplementary positive)	**HMM$_{80}$ score**	**ID**	**Description**
MGSSFRSSSSRNEHEDGGDEAEHAL	6.2	gi|15228112|ref|NP_181265.1|	ABC transporter family protein [Arabidopsis thaliana]
MGSFCSKSLGINFGSEYSGSSVADD	6.2	gi|30682523|ref|NP_680154.2|	expressed protein [Arabidopsis thaliana]
MGCASSKHRKRCLHCRRGYSPPVDV	6.1	gi|15228503|ref|NP_189527.1|	expressed protein [Arabidopsis thaliana]
MGICHGKPVEQQSKSLPVSGETNEA	6.1	gi|15226841|ref|NP_181647.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]
MGSSVSKTASSASASSITSILNSPP	6.1	gi|15227326|ref|NP_181664.1|	expressed protein [Arabidopsis thaliana]
MGCFLSKPAGAGPLPPNDAAALPAD	6.1	gi|31433613|gb|AAP55105.1|	putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGCGLSREKDAGGGPRRRPGSVGDV	6.0	gi|32487665|emb|CAE05655.1|	OSJNBa0038O10.21 [Oryza sativa (japonica cultivar-group)]
MGMAAGRSKPMAASSSERGRRWSAR	6.0	gi|18087678|gb|AAL58970.1|AC091811_19	hypothetical protein [Oryza sativa]
MGCTSSKQAKANVVADVYKPPPSSF	6.0	gi|15217659|ref|NP_176631.1|	peptide transporter -related [Arabidopsis thaliana]
MGLCHGKPSQIPEPEAEEAAAAAGV	6.0	gi|23237898|dbj|BAC16472.1|	putative calcium/calmodulin-dependent protein kinase CaMK [Oryza sativa (japonica cultivar-group)]
MGLSYSSNPTQLDNDQKKPSPATAA	6.0	gi|21554185|gb|AAM63264.1|	zinc finger protein-like [Arabidopsis thaliana]
MGGLCSKGSAVDKSPSDTTLGPDRV	6.0	gi|20146508|dbj|BAB89287.1|	P0491F11.23 [Oryza sativa (japonica cultivar-group)]
MGVSFSCPFAEQDDVEAALDSVTVK	6.0	gi|15231294|ref|NP_187969.1|	hypothetical protein [Arabidopsis thaliana]
MGSSFSASFTNSTTAAAVPPPSPPS	6.0	gi|15221835|ref|NP_173303.1|	expressed protein [Arabidopsis thaliana]
MGAGTSKNSDSTSHGDGDGEERDEY	5.9	gi|7484665|pir||T08994	6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2, 6-bisphosphate 2-phosphatase (EC 3.1.3.46) - spinach
MGNSVRSNLRDIRGRRSMDPRMWHK	5.9	gi|18395895|ref|NP_565317.1|	expressed protein [Arabidopsis thaliana]
MGNCVALEISCDQTLNHACGCLFGD	5.9	gi|15218365|ref|NP_173041.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGNCFGFSAKVGNRESPYRGSSRIS	5.8	gi|25287598|pir||G86396	protein T7N9.2 [imported] - Arabidopsis thaliana
MGNRARRRPGGSGDEENEQRKVGGS	5.8	gi|34393262|dbj|BAC83132.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGCEVSKLSALCCVSESGRSNPDVT	5.8	gi|15237604|ref|NP_198942.1|	protein kinase family [Arabidopsis thaliana]
MGNCCPGSGDAEPASSDASTGNGSS	5.8	gi|20160680|dbj|BAB89623.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGSACSRKRGQLLVDEEDLYSARFS	5.8	gi|32488227|emb|CAE02935.1|	OSJNBa0014K14.7 [Oryza sativa (japonica cultivar-group)]
MGCISSKLLPPGPGDARGGSGARAT	5.8	gi|21741722|emb|CAD40645.1|	OSJNBa0073L04.18 [Oryza sativa (japonica cultivar-group)]
MGNCCRSPAAAAREDVKSSHFPASA	5.8	gi|20805246|dbj|BAB92912.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGCWYSRIDREEIVSRCKARKRYMK	5.8	gi|22858664|gb|AAN05792.1|	unknown [Gossypium hirsutum]
MGNGVASFSGCCAGTTAGEISGRYV	5.7	gi|15232663|ref|NP_187551.1|	hypothetical protein [Arabidopsis thaliana]
MGGCVGKGRSIVEEKLDFKGGNVHV	5.7	gi|27543511|gb|AAO16555.1|	thioredoxin h [Leymus chinensis]
MGNQTSKKSQETSAKSVHYTTELRS	5.7	gi|15235330|ref|NP_195158.1|	hypothetical protein common family [Arabidopsis thaliana]
MGHSPSRHNACGGGGGDGESPPSPL	5.7	gi|20146228|dbj|BAB89010.1|	putative phosphatidylserine decarboxylase [Oryza sativa (japonica cultivar-group)]
MGCKGSKLEDQEAVALCRGRAELLA	5.6	gi|32483081|emb|CAE01923.1|	OSJNBb0078D11.6 [Oryza sativa (japonica cultivar-group)]
MGCFLACFGDRRRRRQPRRQSPARS	5.6	gi|32488453|emb|CAE03386.1|	OSJNBa0004N05.10 [Oryza sativa (japonica cultivar-group)]
MGSAMSLSCSKRKATSQDVDSESCK	5.6	gi|30692215|ref|NP_850655.1|	disease resistance protein RPP1-WsB-like (TIR-NBS-LRR class), putative [Arabidopsis thaliana]
MGVNASEPKRTRSRPSQSQGRHPST	5.5	gi|11357615|pir||T47772	hypothetical protein F24I3.200 - Arabidopsis thaliana
MGITSSTDPKTETFSDPKSQNNEEG	5.5	gi|21592562|gb|AAM64511.1|	unknown [Arabidopsis thaliana]
MGNRLCCHDAADDEKPAAAAVSGGG	5.5	gi|20503002|gb|AAM22711.1|AC098693_16	Putative protein kinase [Oryza sativa (japonica cultivar-group)]

Sequence (Supplementary positive)	HMM_{80} score	ID	Description
MGSVCCVAVKDRKVPPSGGPASSSV	5.4	gi|15222225|ref|NP_177673.1|	expressed protein [Arabidopsis thaliana]
MGCLSVVGASPCVSSSALSSPTSRL	5.4	gi|27807824|dbj|BAC55280.1|	nucleoside diphosphate kinase [Nicotiana tabacum]
MGASSSTDDKESSEKREIESLAAST	5.4	gi|15241814|ref|NP_198775.1|	expressed protein [Arabidopsis thaliana]
MGCASSAPDRTVSRGSSSLNSANEG	5.4	gi|7438400|pir||T12437	small GTP-binding protein - common ice plant
MGCFQSKSRRPFPGHEDPVLLASQT	5.4	gi|27650414|emb|CAD33259.1|	calcineurin B [Crocus sativus]
MGSCVSTTRRRRRSRKLSVAARKFR	5.4	gi|14488359|gb|AAK63926.1|AC084282_7	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGISLSKRRRDNNNNHHHPHHNPP	5.3	gi|15239642|ref|NP_197409.1|	expressed protein [Arabidopsis thaliana]
MGNGIGKLSKCLTGGAGRNKKPELS	5.2	gi|15226985|ref|NP_180455.1|	expressed protein [Arabidopsis thaliana]
MGSCVSSSQSKSSSASDSVVKLASF	5.2	gi|30678442|ref|NP_186835.2|	expressed protein [Arabidopsis thaliana]
MGTCMSTHSRRLRPRRKGRRRFSKN	5.2	gi|15241746|ref|NP_198759.1|	hypothetical protein [Arabidopsis thaliana]
MGGCAAQAARADAAATGQSAARERA	5.1	gi|31126743|gb|AAP44664.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGLTSSKGGRCRQVGPTGQRLKGGR	5.1	gi|34395360|dbj|BAC84430.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGLCWGSPSDSPPTTTPSSTGNISS	5.0	gi|18398350|ref|NP_565408.1|	protein kinase -related [Arabidopsis thaliana]
MGSSSSSPKPDNVQEAEKNEFASLS	5.0	gi|21593253|gb|AAM65202.1|	transcriptional regulator, putative [Arabidopsis thaliana]
MGNLHGIHRSHHGGSNFPGEAPNSP	5.0	gi|15232712|ref|NP_187565.1|	polysaccharide lyase family 1 (pectate lyase) [Arabidopsis thaliana]
MGGCTSKQERRGVRSVKETSKDQSR	5.0	gi|15229545|ref|NP_189035.1|	expressed protein [Arabidopsis thaliana]
MGCLHSKTANLPSSDDPSAPNKPES	5.0	gi|15237465|ref|NP_199469.1|	protein kinase family [Arabidopsis thaliana]
MGCCCCLPSIPESSRTIDEHLPLSR	4.9	gi|18416148|ref|NP_567682.1|	zinc finger (C3HC4-type RING finger) protein family [Arabidopsis thaliana]
MGAGMSSGRRPSFYGGGGGGGVGRP	4.9	gi|15408827|dbj|BAB64219.1|	P0581F09.16 [Oryza sativa (japonica cultivar-group)]

<table>
<tr><th>Sequence (Supplementary positive)</th><th>HMM$_{80}$ score</th><th>ID</th><th>Description</th></tr>
<tr><td>MGSCASVHKDLGFPKKLFLASSPTK</td><td>4.9</td><td>gi|19386733|dbj|BAB86115.1|</td><td>P0678F11.22 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGCLLGCFGRRKNRRRQRRRESFQP</td><td>4.9</td><td>gi|15240079|ref|NP_199217.1|</td><td>expressed protein [Arabidopsis thaliana]</td></tr>
<tr><td>MGHSARKKKKKKGGGGRKAAKDHGG</td><td>4.9</td><td>gi|32480225|emb|CAE02048.1|</td><td>OJ990528_30.6 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGNAIRLLRKCLNSHGVSASSGGVS</td><td>4.8</td><td>gi|18410440|ref|NP_567036.1|</td><td></td></tr>
<tr><td>MGCVASKNAVSVTPAADSSGALRER</td><td>4.8</td><td>gi|33146814|dbj|BAC79804.1|</td><td>putative cyclin-dependent kinase CDC2C [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGITSSTDPKTETFSDPKSQNNEEA</td><td>4.8</td><td>gi|30683589|ref|NP_849659.1|</td><td>expressed protein [Arabidopsis thaliana]</td></tr>
<tr><td>MGCASSKNRNRCRNCKGGLSPVIVP</td><td>4.7</td><td>gi|22327469|ref|NP_680368.1|</td><td>expressed protein [Arabidopsis thaliana]</td></tr>
<tr><td>MGNCSPSPRRRGRPTSPSGSPPLDR</td><td>4.7</td><td>gi|20521256|dbj|BAB91772.1|</td><td>P0679C12.24 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGSSQSSQLLDEEEEEEDEAESEGE</td><td>4.7</td><td>gi|18391446|ref|NP_563918.1|</td><td>expressed protein [Arabidopsis thaliana]</td></tr>
<tr><td>MGGCVGKDRSIVEDKLDFKGGNVHV</td><td>4.7</td><td>gi|11362710|pir||T50864</td><td>thioredoxin-like protein [imported] - Hordeum bulbosum</td></tr>
<tr><td>MGNGNSREAKESRRSRLRHKLQKFR</td><td>4.7</td><td>gi|15242057|ref|NP_200529.1|</td><td>phosphatidylserine decarboxylase, putative [Arabidopsis thaliana]</td></tr>
<tr><td>MGHTSSGNEVPLKYSSGKAFPLGVS</td><td>4.6</td><td>gi|29126649|gb|AAO17049.2|</td><td>isoamylase-type starch debranching enzyme ISO3 [Zea mays]</td></tr>
<tr><td>MGGLCSRSSSVNNAPGGTFAHVNGH</td><td>4.6</td><td>gi|15218580|ref|NP_174692.1|</td><td>expressed protein [Arabidopsis thaliana]</td></tr>
<tr><td>MGACNSCEATAVAAVNGRSASGEAT</td><td>4.6</td><td>gi|31126784|gb|AAP44703.1|</td><td>hypothetical protein [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGCVSSCFGVDDFEDYPNPSSSVNR</td><td>4.6</td><td>gi|15242393|ref|NP_197083.1|</td><td>zinc finger (C3HC4-type RING finger) protein family [Arabidopsis thaliana]</td></tr>
<tr><td>MGACISFFSSSSPSKTAATTNNHSN</td><td>4.5</td><td>gi|25287611|pir||C85420</td><td>protein kinase-like protein [imported] - Arabidopsis thaliana</td></tr>
<tr><td>MGCSFSGLNALYDAVNGGGDVWINE</td><td>4.4</td><td>gi|30682270|ref|NP_850792.1|</td><td>serine/threonine protein kinase, putative [Arabidopsis thaliana]</td></tr>
<tr><td>MGCFMGCFGLSSNKKRRNSIRKILP</td><td>4.4</td><td>gi|30685680|ref|NP_180898.2|</td><td>expressed protein [Arabidopsis thaliana]</td></tr>
<tr><td>MGPGASKETPTQQQKPAESVSTPPE</td><td>4.4</td><td>gi|20161674|dbj|BAB90592.1|</td><td>P0697D09.21 [Oryza sativa (japonica</td></tr>
</table>

Sequence (Supplementary positive)	HMM_{80} score	ID	Description
			cultivar-group)]
MGICFSAEDQHYQFSQQQNYPKKTT	4.4	gi|22325532|ref|NP_178731.2|	protein kinase, putative [Arabidopsis thaliana]
MGGCLSGDVKGGKQAIGGVQQRPTS	4.4	gi|15223946|ref|NP_172362.1|	copine, putative [Arabidopsis thaliana]
MGISSSSGSFRFSDNDHALGNSIEI	4.4	gi|25411036|pir||E84452	hypothetical protein At2g03810 [imported] - Arabidopsis thaliana
MGLTSSLRFHRQNNKTFLGIFMILV	4.4	gi|15227374|ref|NP_181682.1|	cytokinin oxidase family [Arabidopsis thaliana]
MGSNASVKDILGELQKGSYISLLSK	4.4	gi|10998336|gb|AAG25896.1|AF170086_1	silverleaf whitefly-induced protein 1; M20B peptidase [Cucurbita pepo]
MGAGASKDVDGGSLGSQEREVSDGG	4.3	gi|14149116|dbj|BAB55655.1|	fructose-6-phosphate 2-kinase/fructose-2,6-bisphosphatase [Bruguiera gymnorrhiza]
MGCFSCCGADDVGKKKKRDDPYVPI	4.3	gi|29027802|dbj|BAC65877.1|	OSJNBa0011P19.1 [Oryza sativa (japonica cultivar-group)]
MGGCVGKGRRHIEEDKLDFKGGNVH	4.3	gi|24637229|gb|AAN63618.1|AF435817_1	thioredoxin h-like protein [Oryza sativa]
MGCGGSKVDDQPLVILCRERKQLIK	4.2	gi|15219083|ref|NP_173593.1|	bzip-related transcription factor [Arabidopsis thaliana]
MGGRGGGSTDYDDKVGRKRKKEKER	4.2	gi|15222574|ref|NP_174492.1|	expressed protein [Arabidopsis thaliana]
MGGSSGGGVSYRSGGRSDVELEDYE	4.2	gi|25295800|pir||T52046	potassium channel protein SKOR [validated] - Arabidopsis thaliana
MGCFGCCGGGEDFRRVSETGPKPVH	4.1	gi|22331138|ref|NP_188367.2|	serine/threonine protein kinase, putative [Arabidopsis thaliana]
MGNACCVATRDKMVVPNTSAVENFQ	4.1	gi|30681426|ref|NP_850014.1|	expressed protein [Arabidopsis thaliana]
MGSSTSGNCSSVSTTGLANSGSESD	4.1	gi|15222223|ref|NP_177672.1|	bZIP transcription factor ATB2, putative [Arabidopsis thaliana]
MGICLSAQIKAVSPGASPKYMSSEA	4.1	gi|15227001|ref|NP_180459.1|	protein kinase (APK1b) [Arabidopsis thaliana]
MGCVFGRPSSSPPAAAARKGRRKKR	4.1	gi|15528679|dbj|BAB64745.1|	putative CRK1 protein [Oryza sativa (japonica cultivar-group)]
MGTAVAAACAASGDSGDSGPRGDRS	4.0	gi|31415905|gb|AAP50926.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]

Sequence (Supplementary positive)	**HMM$_{80}$ score**	**ID**	**Description**
MGHCFSLPSSQSE-IHEDNEHGDGN	4.0	gi|15239565|ref|NP_197973.1|	protein phosphatase 2C (PP2C), putative [Arabidopsis thaliana]
MGGSSGGGVSYRSGGESDVELEDYE	4.0	gi|15232991|ref|NP_186934.1|	
MGTASSKSNINLSQVLNTEEPYSSV	3.9	gi|15235775|ref|NP_194002.1|	cytochrome P450 family [Arabidopsis thaliana]
MGCSTSKLDDLPAVALCRDRCSFLE	3.9	gi|6598363|gb|AAF18601.1|AC002354_13	hypothetical protein [Arabidopsis thaliana]
MGSFHRRTFSYDKLPNEPIRLSVLK	3.9	gi|15239396|ref|NP_197916.1|	expressed protein [Arabidopsis thaliana]
MGNLGVREAGRMAVLSRRWRRLPGL	3.8	gi|34393595|dbj|BAC83248.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGNACCVAARDKMVVPNSSAGENLQ	3.8	gi|30692196|ref|NP_849522.1|	expressed protein [Arabidopsis thaliana]
MGPVFSSLMYEEEDQGNDSELGDIH	3.7	gi|32488631|emb|CAE03424.1|	OSJNBa0032F06.7 [Oryza sativa (japonica cultivar-group)]
MGGVLSSVLGGGGGDEPLAGNESES	3.7	gi|25990392|gb|AAN76509.1|AF352030_1	thioredoxin h [Brassica rapa]
MGSGCASSPPTSPQPRHGRLVTVLS	3.6	gi|14346205|emb|CAC41149.1|	unnamed protein product [Triticum aestivum]
MGNCCGTDGFKHERPLPAHPNRQGE	3.6	gi|23268465|gb|AAN11310.1|	calmodulin domain protein kinase 1 [Ceratopteris richardii]
MGLFTVTKKATTPFDGQKPGTSGLR	3.5	gi|12585309|sp|P93804|PGM1_MAIZE	Phosphoglucomutase, cytoplasmic 1 (Glucose phosphomutase 1) (PGM 1)
MGACASASRPPPSPRHGCSPPTPT	3.5	gi|29467524|dbj|BAC66713.1|	P0582D05.14 [Oryza sativa (japonica cultivar-group)]
MGRSRGVLSSGDDDTGHRSKRRRVS	3.5	gi|21898562|gb|AAM77037.1|	histone acetyltransferase complex component [Zea mays]
MGCMFSHLAAKFAFFPPSPPTYHLT	3.5	gi|15222598|ref|NP_174498.1|	expressed protein [Arabidopsis thaliana]
MGGFLVLNSSNNVVRTVPSKKRKHP	3.5	gi|15241355|ref|NP_197544.1|	hypothetical protein [Arabidopsis thaliana]
MGNTQGKGDGAMTGRAKAITGGAVK	3.4	gi|11230987|dbj|BAB18105.1|	cyclic nucreotide dependent protein kinase II [Chlamydomonas reinhardtii]
MGCCQSSFLKPSSLHDKKITSDDLS	3.4	gi|15223700|ref|NP_175512.1|	protein kinase - related [Arabidopsis thaliana]
MGGCVGKDRGIVEDKLDFKGGNVHV	3.3	gi|11362711|pir||T50862	thioredoxin-like protein [imported] - Phalaris coerulescens

Sequence (Supplementary positive)	HMM_{80} score	ID	Description
MGLRLAGGEVAVGSSGRGGGGRWHG	3.3	gi|31415932|gb|AAP50953.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGAAVALHFATSCALNHYTINPRVV	3.3	gi|15219001|ref|NP_175655.1|	hypothetical protein [Arabidopsis thaliana]
MGIFTVRSAYKLASKAELCSENSGE	3.2	gi|28376719|gb|AAO41149.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGGRCVGAAVKCGEGAAGGAEEGVG	3.2	gi|31431866|gb|AAP53578.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGPKVAGPCPADGNGKSSSSPLLSA	3.2	gi|21952827|dbj|BAC06243.1|	P0018C10.45 [Oryza sativa (japonica cultivar-group)]
MGLSARHGSLGGFYPGFIRRGSCGA	3.2	gi|18844884|dbj|BAB85353.1|	P0519D04.12 [Oryza sativa (japonica cultivar-group)]
MGNGGGGNKKGRRDSPSPPLETPPR	3.2	gi|31432873|gb|AAP54449.1|	putative lanthionine synthetase [Oryza sativa (japonica cultivar-group)]
MGHKQSVDSRGKGRKVPGSSSMVQK	3.2	gi|15226911|ref|NP_178332.1|	F-box protein (SKP1 interacting partner 3-related) [Arabidopsis thaliana]
MGNGNSTETKESRRSKMRKKIQNFR	3.1	gi|18416656|ref|NP_567736.1|	phosphatidylserine decarboxylase, putative [Arabidopsis thaliana]
MGATCVHGFRPDHNLQDLRAKIFAL	3.1	gi|25406936|pir||E86202	hypothetical protein [imported] - Arabidopsis thaliana
MGNSFGCSASGERLVSAARDGDFVE	3.1	gi|15240746|ref|NP_196344.1|	expressed protein [Arabidopsis thaliana]
MGGRGAGGEAHDDPFRSRSSSFGAV	3.1	gi|15128233|dbj|BAB62561.1|	OSJNBa0089K24.19 [Oryza sativa (japonica cultivar-group)]
MGGKLCLVAARSAHPSGRSAAVWPV	3.1	gi|11875149|dbj|BAB19362.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGSSCSRSHSLSEAETTKNAKSADI	3.1	gi|540533|dbj|BAA07405.1|	G protein alpha subunit [Oryza sativa (japonica cultivar-group)]
MGGFAASNIVSESGEKASNLVYRGR	3.0	gi|34394682|dbj|BAC83988.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNCVGASNRSTSRENKGNHRVSRR	3.0	gi|31415931|gb|AAP50952.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGGAVARAPGNQKQKSDLAGSQQRR	3.0	gi|22535687|dbj|BAC10861.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGSSRSEGSSSSSARQPYGSPIPYR	2.9	gi|31432406|gb|AAP54043.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGGVTSSTSSSSPATAAGGSSYSAE	2.9	gi|20160988|dbj|BAB89922.1|	putative hypersensitive-induced response protein [Oryza sativa (japonica cultivar-group)]
MG-CFSCFDSSDDEKLNPVDESNHG	2.9	gi|15240038|ref|NP_196820.1|	protein kinase family [Arabidopsis thaliana]
MGLCVSVNRNEYVSSSTTAKIVTIN	2.8	gi|18411155|ref|NP_565136.1|	expressed protein [Arabidopsis thaliana]
MGCFGRTPKSNKRSDTKTTKNNDFT	2.8	gi|15226901|ref|NP_180426.1|	protein kinase family [Arabidopsis thaliana]
MGNAIGKASNDIGGFIGNIFTAPLK	2.8	gi|25373345|pir||A96819	hypothetical protein F9K20.1 [imported] - Arabidopsis thaliana
MGASTSTSNPSLEAQEQETLASASL	2.8	gi|32489048|emb|CAE04845.1|	OSJNBa0084K01.17 [Oryza sativa (japonica cultivar-group)]
MGCTNSTEARREMVLADAGRRATRS	2.7	gi|32490043|emb|CAE05962.1|	OSJNBa0063C18.3 [Oryza sativa (japonica cultivar-group)]
MGLFTVTKKATTPFEGQKPGTSGLR	2.7	gi|12585310|sp|P93805|PGM2_MAIZE	Phosphoglucomutase, cytoplasmic 2 (Glucose phosphomutase 2) (PGM 2)
MGNCWGAKISSESPCRSASSPSGGT	2.7	gi|14719339|gb|AAK73157.1|AC079022_30	putative protein kinase [Oryza sativa]
MGNCICVTEKTTTSWSGDDNGSYNK	2.7	gi|15234649|ref|NP_193923.1|	hypothetical protein [Arabidopsis thaliana]
MGQCYGKAGGASSRRADHDDAVAPP	2.7	gi|7446431|pir||T03023	calcium-dependent protein kinase-related protein kinase - maize
MGEAASTEEGDADAIKEDEAASTEE	2.7	gi|14587269|dbj|BAB61187.1|	OSJNBb0032H19.20 [Oryza sativa (japonica cultivar-group)]
MGGALSSVLGGDAAAAAASADDSSS	2.6	gi|33621080|gb|AAQ23133.1|	thioredoxin H2 [Ipomoea batatas]
MGSFFSTMFTPPPAADDGGDSRVVA	2.6	gi|28209505|gb|AAO37523.1|	putative thioredoxin [Oryza sativa (japonica cultivar-group)]
MGNFVCIEISGDQMLDRIIRCLCGK	2.6	gi|15219879|ref|NP_176314.1|	disease resistance protein (CC-NBS-LRR class), putative [Arabidopsis thaliana]
MGCCAS---NTAGGSKANRISRWRS	2.6	gi|22326653|ref|NP_196408.2|	leucine rich repeat protein family [Arabidopsis thaliana]
MGAASSSIVRSEPFAGKLCGLEDVP	2.6	gi|15242282|ref|NP_200025.1|	F-box protein (SKP1 interacting partner 3-related) [Arabidopsis thaliana]
MGGCVGKDRSIVEEKLDFKGGNVHV	2.5	gi|24637237|gb|AAN63622.1|AF438359_1	thioredoxin [Triticum aestivum]

Sequence (Supplementary positive)	**HMM_{80} score**	**ID**	**Description**
MGNCLDSSAKVDNSNHSPHANSASS	2.5	gi|15223723|ref|NP_172889.1|	protein kinase (APK2a) [Arabidopsis thaliana]
MGNCAIKPKVLKDSDEDLVPVERET	2.5	gi|15237758|ref|NP_201274.1|	expressed protein [Arabidopsis thaliana]
MGSLMSGWDSRVRDPKSVRRCKSLT	2.5	gi|18394756|ref|NP_564090.1|	expressed protein [Arabidopsis thaliana]
MGCGGSKVDNQPIVILCRERKELLK	2.4	gi|15223973|ref|NP_177874.1|	expressed protein [Arabidopsis thaliana]
MGNTSSCAPLIISTNSSSGVVKILA	2.4	gi|15221054|ref|NP_173265.1|	hypothetical protein [Arabidopsis thaliana]
MGNCMGSSKCASASAPLESKVTRSS	2.4	gi|33146495|dbj|BAC79604.1|	putative serine/threonine-specific protein kinase [Oryza sativa (japonica cultivar-group)]
MGIYLSTPKTDKFSEDGENHKLRYG	2.4	gi|15236548|ref|NP_194914.1|	protein phosphatase 2C (PP2C), putative [Arabidopsis thaliana]
MGACISFFSSSSPSKTGLHSHATTN	2.4	gi|30690422|ref|NP_195285.3|	protein kinase family [Arabidopsis thaliana]
MGNACGGALRSKHQHSFKHAASSQR	2.4	gi|7428005|pir||T03263	calcium-dependent protein kinase (EC 2.7.1.-) 7 - maize
MGGRVARDCSESTRNGGFWGVMVVD	2.4	gi|24059969|dbj|BAC21431.1|	hypothetical protein~predicted by FGENESH etc. [Oryza sativa (japonica cultivar-group)]
MGGFCCCLCTDDFEEYAHPNNPIYR	2.4	gi|20805222|dbj|BAB92889.1|	OSJNBa0093F16.9 [Oryza sativa (japonica cultivar-group)]
MGNVCVHMVNNCVDTKSNSWVRPT	2.4	gi|15227525|ref|NP_181133.1|	calcium-dependent protein kinase, putative (CDPK) [Arabidopsis thaliana]
MGNGLMSKKKKKKNNKYKKKRPNVS	2.4	gi|15224411|ref|NP_178561.1|	F-box protein family [Arabidopsis thaliana]
MGQGQSGGLPKQGPGGERNKGDGQK	2.4	gi|18479040|gb|AAL73395.1|AF432345_1	26S proteasome regulatory subunit IV [Tortula ruralis]
MGCFGCSKKSSKRSETNKDTVINRK	2.3	gi|15223024|ref|NP_172265.1|	protein kinase family [Arabidopsis thaliana]
MGIAFSKNWWSRFSGSSPLSVMVCE	2.3	gi|34394463|dbj|BAC83676.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGGRRRGASSPALGRGRGRRRVAGV	2.3	gi|22324481|dbj|BAC10395.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGCTHTRHSRKRYQRRVTTRRAEHL	2.3	gi|8467997|dbj|BAA96598.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]

Sequence (Supplementary positive)	HMM$_{80}$ score	ID	Description
MGLCHGKSAAVLEPTVEEEEEGATR	2.3	gi|14133874|gb|AAK54157.1|AF368282_1	CaMK1 [Oryza sativa]
MGTTSSTSSPLFSVFCSPTPLENGR	2.3	gi|29244700|gb|AAO73292.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGGYQAVLDPDNQKPSSKNPGYFLS	2.2	gi|28411895|dbj|BAC57341.1|	leaf senescence related protein-like protein [Oryza sativa (japonica cultivar-group)]
MGCFGPSKASRTRNNEHDKITRQNP	2.2	gi|15233098|ref|NP_188797.1|	expressed protein [Arabidopsis thaliana]
MGLCSPSPSSHDCSSSSPRNSEKKQ	2.2	gi|25372915|pir||E86343	T22I11.12 protein - Arabidopsis thaliana
MGITHANSITTEPSPSPRPPLLRLL	2.2	gi|24059896|dbj|BAC21362.1|	aminoacyl peptidase-like [Oryza sativa (japonica cultivar-group)]
MGCIISKKKSPKRNPPWKETLEKRS	2.1	gi|15241455|ref|NP_199242.1|	cyclin-dependent protein kinase-related protein [Arabidopsis thaliana]
MGGRVSKAVACCCCRSQHHGVVVES	2.1	gi|32488705|emb|CAE03448.1|	OSJNBa0088H09.6 [Oryza sativa (japonica cultivar-group)]
MGSSHAGGEMCMEGGGGGRARRAGG	2.1	gi|7340912|dbj|BAA92984.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGCTTSHDAFAAARARARASSSSLE	2.1	gi|20146230|dbj|BAB89012.1|	bzip-like transcription factor-like [Oryza sativa (japonica cultivar-group)]
MGQGPSGGLHRQGDRKPDGGEKKEK	2.1	gi|21593177|gb|AAM65126.1|	26S proteasome subunit 4 [Arabidopsis thaliana]
MGNSCSSVDESAAAIATLAESAAAG	2.1	gi|33146803|dbj|BAC79751.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGSMYRHLRDRSSRRRRFSPAWRQE	2.1	gi|15451581|gb|AAK98705.1|AC069158_17	Hypothetical protein [Oryza sativa]
MGVLFSCPVDDETAAVEDAAPVAEQ	2.1	gi|22296385|dbj|BAC10154.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGQSMSCGSRPEHGIFASVQCGDII	2.1	gi|30684032|ref|NP_180450.2|	RING zinc finger ankyrin protein -related [Arabidopsis thaliana]
MGGSSSSSPTSSSSSPAKRGKNKNG	2.1	gi|7485535|pir||T04515	hypothetical protein F16A16.50 - Arabidopsis thaliana

Supplementary Table 4. *Arabidopsis thaliana* proteins predicted to contain N-terminal myristoylation sites.

Sequence (Arabidopsis predict myristoylated)	HM_{266} Score	ID	function	train	description
MGLNFTKLFSRLFAKKEMRILMVGL	4.1	At5g14670.1	ARF	yes	ADP-ribosylation factor ARF1, putative similar to ADP-ribosylation factor DcARF1 (GI:965483) [Daucu
MGLSFGKLFSRLFAKKEMRILMVGL	7.4	At2g47170.1	ARF	yes	ADP-ribosylation factor 1 (ARF1) identical to ADP-ribosylation factor ARF1({Arabidopsis thaliana} (S
MGLSFAKLFSRLFAKKEMRILMVGL	6.6	At1g23490.1	ARF	yes	ADP-ribosylation factor 1 (ARF1), putative identical to SP:Q9SRC3 ADP-ribosylation factor 1-like [Ar
MGLSFAKLFSRLFAKKEMRILMVGL	6.6	At1g70490.2	ARF	(yes)	ADP-ribosylation factor 1 (ARF1), putative nearly identical to ADP-ribosylation factor 1 GB:P36397 [
MGILFTRMFSSVFGNKEARILVLGL	1.7	At2g24765.1	ARF	yes	ADP-ribosylation factor 3 (ARF3) identical to GP:453191 ADP-ribosylation factor 3 {Arabidopsis thali
MGLSFGKLFSKLFAKKEMRILMVGL	7.2	At3g62290.1	ARF	no	ADP-ribosylation factor identical to GP:166586 ADP-ribosylation factor {Arabidopsis thaliana}; ADP-r
MGARFSRIAKRFLPKSKVRILMVGL	7.1	At2g15310.1	ARF	yes	ADP-ribosylation factor, putative similar to ADP-ribosylation factor (GI:861205) [Chlamydomonas rein
MGQAFRKLFDTFFGNQEMRVVMLGL	4.6	At5g17060.1	ARF	yes	ADP-ribosylation factor, putative similar to ADP-ribosylation factor 1; ARF 1 (GP:385340) {Drosophil
MGQTFRKLFDTFFGNQEMRVVMLGL	3.9	At3g03120.1	ARF	yes	ADP-ribosylation factor, putative similar to ADP-ribosylation factor 1; ARF 1 (GP:385340) {Drosophil
MGTTLGKPFAGFFHQEEARIVLFGL	4.3	At1g02440.1	ARF	yes	ADP-ribosylation factor, putative similar to ADP-ribosylation factor GB:AAA32729 GI:166586 from (Ara
MGAFMSRFWFMMFPAKEYKIVVVGL	1.0	At3g22950.1	ARF	yes	ADP-ribosylation factor, putative similar to ADP-ribosylation factor GB:P91924 [Dugesia japonica]
MGLSFAKLFSRLFAKKEMRILMVGL	6.6	At1g10630.1	ARF	(yes)	ADP-ribosylation factor, putative similar to ADP-ribosylation factor GI:166586 from [Arabidopsis tha
MGTALGKPFAGFFHQEESRIVLFGL	5.5	At1g02430.1	ARF	no	ADP-ribosylation factor, putative similar to ADP-ribosylation factor GI:166586 from [Arabidopsis tha
MGCAQSKIENEEAVTRCKERKQLMK	9.7	At1g52320	bZIP	yes	gi|25405541|pir||D96563

Sequence (Arabidopsis predict myristoylated)	HM_{266} Score	ID	function	train	description
MGCTASKLDSEDAVRRCKERRRLMK	14.1	At1g02110.1	bZIP	yes	proline-rich protein family contains proline-rich domain, INTERPRO:IPR000694
MGCAASKLDNEDAVRRCKDRRRLMK	13.1	At3g60320.1	bZIPprotein	no	bZIP protein
MGGGKSNSSSPDSREVREREGIPEL	2.3	At1g21740.1	bZIPprotein	no	bzip-related transcription factor similar to GB:AAB81673, EST gb|T20649 comes from this gene
MGAIYARQLKEDTSPVTDSSRKYSR	3.8	At4g30130.1	bZIPprotein	no	expressed protein bZIP protein - Arabidopsis thaliana, PID:g600855
MGCTSSKLDDLPAVALCRERCAFLE	6.9	At4g35240.1	bZIPprotein	no	expressed protein bZIP protein - Arabidopsis thaliana,PID:g600855
MGGAVTKSETLQKEWVPETKLEAKI	1.9	At4g26470.1	CALB	no	calcium-binding EF-hand family protein low similarity to SP|P06787 Calmodulin {Saccharomyces cerevis
MGCSVSKKKKKNAMRPPGYEDPELL	10.5	At5g24270.1	CALB(SOS3)	yes	calcineurin B-like protein, putative similar to calcineurin B-like protein 8 (GI:15866276) [Arabidop
MGCFHSKAAKEFRGHEDPVKLASET	13.3	At4g17615.1	CALB1	no	calcineurin B-like protein 1 (CBL1) identical to calcineurin B-like protein 1 (GI:3309082) [Arabidop
MGCVCSKQLEGRRQEDISLLASQT	2.2	At4g01420.1	CALB5	no	calcineurin B-like protein 5 (CBL5) identical to calcineurin B-like protein 5 (GI:9965366) [Arabidop
MGCFHSTAAREFPDHENPVKLASET	13.0	At5g47100.1	CALB9	no	calcineurin B-like protein 9 (CBL9) identical to calcineurin B-like protein 9 (GI:5866279) and calci
MGNCCSDVASGAGATAGVGGSGSSA	11.1	At5g61900.1	COPINE	no	copine BONZAI1 (BON1) nearly identical to BONZAI1 [Arabidopsis thaliana] GI:15487382; contains Pfam
MGGCLSGDVKGGKQAIGGVQQRPTS	6.7	At1g08860.1	COPINE	no	copine, putative Similar to BONZAI1 [Arabidopsis thaliana] GI:15487382; contains Pfam profile PF0016
MGGGNSKEESSSPSSSSWASHQSYP	10.7	At3g01650.1	COPINE	no	copine-related low similarity to SP|Q99829 Copine I {Homo sapiens}
MGGSSSKESPRGGGSGRRYERSVSG	12.7	At1g67800.1	COPINE	yes	copine-related low similarity to SP|Q99829 Copine I {Homo sapiens}
MGTGNSKENWRQSSFRSTSASSASP	0.6	At5g14420.2	COPINE	no	copine-related low similarity to SP|Q99829 Copine I {Homo sapiens}
MGNTCVGPSRNGFLQSVSAAMWRPR	4.0	At5g04870.1	CPK01	yes	calcium-dependent protein kinase (CDPK)(AK1) identical to calcium-dependent protein kinase, isoform

Sequence (Arabidopsis predict myristoylated)	**HM$_{266}$ Score**	**ID**	**function**	**train**	**description**
MGNACVGPNISGNGFLQTVTAAMWR	4.8	At3g10660.1	CPK02	yes	calcium-dependent protein kinase (CDPK)(CPK2) identical to calcium-dependent protein kinase isoform
MGHRHSKSKSSDPPPSSSSSSSGNV	13.3	At4g23650.1	CPK03	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [March
MGNSCRGSFKDKLDEGDNNKPEDYS	13.9	At4g35310.1	CPK05	yes	calcium-dependent protein kinase, putative (CDPK) similar to calmodulin-domain protein kinase CDPK i
MGNSCRGSFKDKIYEGNHSRPEENS	10.6	At2g17290.1	CPK06	yes	calcium-dependent protein kinase (CDPK)(CPK6) identical to calmodulin-domain protein kinase CDPK iso
MGNCCGNPSSATNQSKQGKPKNKNN	16.8	At5g12480.1	CPK07	yes	calcium-dependent protein kinase 7 (CDPK)(CPK7) identical to calmodulin-domain protein kinase CDPK i
MGNCCASPGSETGSKKGKPKIKSNP	16.2	At5g19450.1	CPK08	yes	calcium-dependent protein kinase 19 (CDPK19) identical to calcium-dependent protein kinase [Arabidop
MGNCFAKNHGLMKPQQNGNTTRSVE	12.3	At3g20410.1	CPK09	yes	calcium-dependent protein kinase 9 (CDPK)(CPK9) identical to calmodulin-domain protein kinase CDPK i
MGNCNACVRPDSKESKPSSKPKKPN	12.4	At1g18890	CPK10	yes	gi|28416563|gb|AAO42812.1|
MGNCCRSPAAVAREDVKSNYSGHDH	11.1	At3g51850.1	CPK13	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Arabi
MGNCCGTAGSLIQDKQKKGFKLPNP	15.0	At2g41860.1	CPK14	yes	calcium-dependent protein kinase, putative (CDPK) similar to calmodulin-domain protein kinase CDPK i
MGCFSSKHRNTESDIINGSVQSSIP	16.3	At4g21940.1	CPK15	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Nicot
MGLCFSSAAKSSGHNRSSRNPHPHP	19.2	At2g17890.1	CPK16	yes	
MGNCCSHGRDSADNGDALENGASAS	16.7	At5g12180.1	CPK17	yes	calcium-dependent protein kinase, putative (CDPK)
MGLCFSSPKATRRGTGSRNPNPDSP	17.3	At4g36070.1	CPK18	yes	Calcium-dependent serine/threonine protein kinase calcium-dependent protein kinase isoform AK1, Arab
MGNTCVGPNLNPNGFLQSVSAAVWR	4.8	At2g38910.1	CPK20	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase, isofo
MGCFSSKHRKTQNDGGEKSIPINPV	19.2	At4g04720.1	CPK21	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase(CDPK)
MGNCCGSKPLTASDIVSDQKQETIL	10.4	At4g04710.1	CPK22	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Nicot

Sequence (Arabidopsis predict myristoylated)	**HM_{266} Score**	**ID**	**function**	**train**	**description**
MGCFSSKHRKTQNDGGGERSIPIIP	17.7	At4g04740.1	CPK23	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Lycop
MGSCVSSPLKGSPFGKRPVRRRHSS	15.1	At2g31500.1	CPK24	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Arabi
MGNVCVHMVNNCVDTKSNSWVRPTD	3.4	At2g35890.1	CPK25	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase, isofo
MGCFSSKELQQSKRTILEKPLVDIT	6.5	At4g04700.1	CPK27	yes	protein kinase, putative similar to calcium-dependent protein kinase [Nicotiana tabacum] gi|3283996|
MGVCFSAIRVTGASSSRRSSQTKSK	12.5	At5g66210.2	CPK28	yes	calcium-dependent protein kinase
MGNCIACVKFDPDNSKPNQKKKPPR	11.1	At1g74740.1	CPK30	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Arabi
MGCYSSKNLKQSKRTILEKPFVDIG	4.8	At4g04695.1	CPK31	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [Lycop
MGNCCGTAGSLAQNDNKPKKGRKKQ	12.1	At3g57530.1	CPK32	no	calcium-dependent protein kinase, putative (CDPK) similar to calmodulin-domain protein kinase CDPK i
MGNCLAKKYGLVMKPQQNGERSVEI	9.2	At1g50700.1	CPK33	yes	calcium-dependent protein kinase, putative (CDPK) similar to calmodulin-domain protein kinase CDPK i
MGNCCSHGRDSDDNKEEPRPENGGG	20.7	At5g19360.1	CPK34	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium-dependent protein kinase [March
MGICHGKPVEQQSKSLPVSGETNEA	8.2	At2g41140.1	CRK1	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGGCTSKPSSSVKPNPYAPKDAVLQ	15.6	At3g19100.1	CRK2	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGQCYGKVNQSKQNGEEEANTTTYV	9.2	At2g46700.1	CRK3	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGHCYSRNISAVEDDEIPTGNGEV	7.7	At5g24430	CRK4	no	CDPK-related protein kinase isoform 4 (ppc:4.2.1)
MGLCTSKPNSSNSDQTPARNSPLPA	17.0	At3g50530.1	CRK5	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGHCYSRNISTVDDDDEIPSATAQL	13.4	At3g49370.1	CRK6	yes	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGLCHGKPIEQQSKNLPISNEIEET	7.7	At3g56760.1	CRK7	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGGCTSKPSTSSGRPNPFAPGNDYP	18.8	At1g49580.1	CRK8	no	calcium-dependent protein kinase, putative (CDPK) similar to calcium/calmodulin-dependent protein ki
MGTSVSKPVSDDKKKGTSVSKPVKN	11.4	At1g33880.1	disease	no	AIG1-related protein similar to AIG1 protein GB:P54120 [Arabidopsis thaliana], NTGP4 GB:AAD09518 [Ni

Sequence (Arabidopsis predict myristoylated)	**HM$_{266}$ Score**	**ID**	**function**	**train**	**description**
MGACFSVAISCDQAVNNLTSCLSRN	16.1	At5g05400.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGACLTLSFSCDEVVNQISQGLCIN	11.3	At1g12280.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGGCVSVQVSCDQLLNHLGRCFCRK	13.4	At1g12290.1	disease	yes	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGGCVSVSISCDQLTKNVCSCLNRN	11.4	At5g63020.1	disease	yes	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGGCVSVSLSCDREVNQFSQWLCVS	14.2	At1g12210.1	disease	yes	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGISFSIPFDPCVNKVSQWLDMKGS	4.0	At1g62630.1	disease	yes	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGISFSIPFDPCVNKVSQWLDMKVS	3.3	At1g63350.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGISFSIPFDPCVNKVSQWLDMKVS	3.3	At1g63360.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGNCVALEISCDQTLNHACGCLFGD	7.6	At1g15890.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGNNFSVESPSLAPFLCGKRKYLYN	9.7	At5g47260.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGSCFSLQVSDQTLNRIFNCLIGKS	4.3	At1g61180.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti
MGSCISLQISCDQVLTRAYSCFFSL	9.2	At4g10780.1	disease	no	disease resistance protein (CC-NBS-LRR class), putative domain signature CC-NBS-LRR exists, suggesti

Sequence (Arabidopsis predict myristoylated)	HM_{266} Score	ID	function	train	description
MGCCFSVQFSFDDQTLVRIFNFLCG	1.5	At1g61300.1	disease	no	disease resistance protein (NBS-LRR class), putative domain signature NBS-LRR exists, suggestive of
MGSVMSLGCSKRKATNQDVDSESRK	5.8	At3g44480.1	disease	yes	disease resistance protein (TIR-NBS-LRR class), putative domain signature TIR-NBS-LRR exists, sugges
MGSAMSLSCSKRKATSQDVDSESCK	5.4	At3g44630.3	disease	no	disease resistance protein RPP1-WsB-like (TIR-NBS-LRR class), putative domain signature TIR-NBS-LRR
MGGCFSVSLPCDQVVSQFSQLLCVR	16.4	At1g12220.1	disease	yes	disease resistance protein RPS5 (resistance to Pseudomonas syringae protein 5)(CC-NBS-LRR class) dom
MGSVMSLGCSKRKATSQDVDSESRK	5.7	At3g44670	disease	yes	gi|11357250|pir||T47442
MGLLCSRSRHHTEDTDENTQAAEIE	11.6	At2g26300.1	GTPB	yes	G protein alpha subunit 1 (GPA1) identical to GB:M32887
MGCASSLPDRNSGTLSGLSNSENAV	9.6	At3g54840.1	GTPB	yes	GTP-binding protein (Ara6)
MGCVFGREAATTTTAEAKQAKSSKA	7.4	At1g54610.1	KINASE	no	CRK1 protein -related similar to CRK1 protein GI:7671528 from [Beta vulgaris] - cdc Related 4.5.2
MGCICSKGVRTNDDYIETNHVSIGK	15.9	At1g57700.1	KINASE	no	CRK1 protein -related similar to CRK1 protein GI:7671528 from [Beta vulgaris]putative protein kinase (ppc:4.5.2)
MGCGKSKHDVVTGNTTRIKKPSEAE	3.1	At4g15650.1	KINASE	no	kinase like protein
MGCCQSSFLKPSSLHDKKITSDDLS	10.0	At1g50990.1	KINASE	no	protein kinase - related low similarity to SP|Q06548|APKA_ARATH Protein kinase APK1A Arabidopsis tha
MGNCL-KPLKEQPPSASPKPLTIPS	10.0	At5g11410.1	KINASE	no	protein kinase family contains eukaryotic protein kinase domain, INTERPRO:IPR000719
MGACISFFSSSSPSKTGLHSHATTN	4.1	At4g35600.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGARCSKFSFCLFPSHFKSASVLES	8.9	At3g54030.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGCCQSLFSGDNPLGKDGVQPQPLS	13.8	At4g35230.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGCEVSKLSALCCVSESGRSNPDVT	2.7	At5g41260.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGCFGCSKKSSKRSETNKDTVINRK	6.5	At1g07870.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069

Sequence (Arabidopsis predict myristoylated)	HM$_{266}$ Score	ID	function	train	description
MGCFGRTPKSNKRSDTKTTKNNDFT	5.8	At2g28590.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGCLHSKTANLPSSDDPSAPNKPES	8.8	At5g46570.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGCVNSRHRPFRRKSTTLKESSEEK	1.0	At1g03740.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGGCCFSTPLDNNQQPPQQPQIPPGE	3.3	At2g17170.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGNCLSASTLPTEPNAPKSVEEGVK	12.6	At5g25440.1	KINASE	no	protein kinase family contains protein kinase domain, Pfam:PF00069
MGLPHSGSVSPLSPLGEACSRRDLT	0.9	At4g00710.1	KINASE	no	protein kinase family low similarity to protein kinase [Arabidopsis thaliana] GI:2852449; contains P
MGRSSSKKKKKRGGSGRRGQLKDHG	6.7	At3g59410.1	KINASE	no	protein kinase like GCN2 - Saccharomyces cerevisiae, EMBL:M27082
MGLCWGSPSDSPPTTTPSSTGNISS	6.1	At2g17220.1	KINASE	no	protein kinase -related contains a protein kinase domain profile (PDOC00100)
MGCVNSKQTVSVTPAIDHSGVFRDN	7.7	At1g18670.1	KINASE	no	protein kinase -related similar to CRK1 protein GI:7671528 from [Beta vulgaris] putatuve 4.5.2
MGCVSSKQTVSVTPAIDHSGVFKDN	9.8	At1g74330.1	KINASE	no	protein kinase -related similar to p58 protein kinase GB:AAB59449 (Homo sapiens); contains Pfam prof
MGCICATARSPSAAVTDKDLLDSSK	8.2	At1g71530.1	KINASE	no	protein kinase -related similar to protein kinase GB:4502731 from (Homo sapiens)
MGICFSAEDQHYQFSQQQNYPKKTT	4.8	At2g07180.1	KINASE	no	protein kinase, putative contains protein kinase domain, Pfam:PF00069
MGNCWCRFEPLNHRVSANAKSESPK	4.3	At5g56460.1	KINASE	no	protein kinase, putative contains protein kinase domain, Pfam:PF00069
MGNCFESCSKRNDNEDDNLDSSVNSK	5.2	At1g74490.1	KINASE	no	protein kinase, putative similar to protein kinase [Arabidopsis thaliana] gi|2852449|dbj|BAA24695; c
MGSCFSSRVKADIFHNGKSSDLYGL	12.1	At2g39660.1	KINASE	no	protein kinase, putative similar to protein kinase gi|166809|gb|AAA18853
MGCSFSGLNALYDAVNGGGDVWINE	2.9	At5g08160.2	KINASE	no	serine/threonine protein kinase, putative identical to serine/threonine protein kinase [Arabidopsis
MGICLSAQVKAESSGASTKYDAKDI	6.9	At1g07570.1	KINASE - APK1	no	protein kinase (APK1a) identical to Protein kinase APK1A from [Arabidopsis thaliana] SWISS-PROT:Q065

Sequence (Arabidopsis predict myristoylated)	HM$_{266}$ Score	ID	function	train	description
MGICLSAQIKAVSPGASPKYMSSEA	5.1	At2g28930.1	KINASE - APK1	no	protein kinase (APK1b) identical to protein kinase APK1B [Arabidopsis thaliana] SWISS-PROT:P46573
MGNCLDSSAKVDNSNHSPHANSASSG	1.0	At1g14370.1	KINASE - APK1	no	protein kinase (APK2a) identical to protein kinase APK2a GI:2852447 from [Arabidopsis thaliana]
MGNCLDSSAKVDSSSHSPHANSASLS	1.1	At2g02800.1	KINASE - APK1	no	protein kinase (APK2b) identical to protein kinase APK2b [Arabidopsis thaliana] gi|2852449|dbj|BAA24
MGCCYSLSSTVDPVQDHTTDASSEP	18.5	At2g17090.1	KINASE - APK1	no	protein kinase family similar to Arabidopsis thaliana APK1A [SP|Q06548], APK1B [SP|P46573]; contains
MGGCFSNRIKTDIASSTWLSSKFLS	5.2	At5g02290.2	KINASE - APK1	no	protein kinase, putative similar to protein kinase APK1A [Arabidopsis thaliana] SWISS-PROT:Q06548
MGCISSKNVSCLTDQGDSPLPEPGL	15.1	At5g39420.1	KINASE - CDC	no	cdc2-related protein
MGCFNSKHSSPAGPPRPVRRRLDTN	17.1	At3g01085.1	KINASE - CDC	no	cdc2-related protein temporary automated functional assignment
MGCNCTKGTRPDNDNVDNSNSIVSN	14.5	At1g09600.1	KINASE - CDC	no	protein kinase -related Similar to cdc2 protein kinases
MGCICSKGAAEDEEGVVYHREKANE	11.2	At1g33770.1	KINASE - CDC	no	protein kinase -related similar to cdc2-related protein kinase GB:AAA58424 GI:180492 from [Homo sapi
MGNANGKDEDAAAGSGGADVTSSSA	4.6	At5g21170.1	KINASE - cyclin	no	AMPKBI (5'-AMP-activated protein kinase, beta subunit, complex-interacting region) domain family con
MGCVLGRPGSSGSVSGSRDEVSTRI	10.8	At3g05050.1	KINASE - cyclin	no	cyclin-dependent protein kinase -related similar to cyclin-dependent kinase GB:CAA65979 from [Medica
MGCVLCKESTGDKRKHNNPDEPPPA	14.0	At5g50860.1	KINASE - cyclin	yes	cyclin-dependent protein kinase-related
MGCIISKKKSPKRNPPWKETLEKRS	2.9	At5g44290.1	KINASE - cyclin	no	cyclin-dependent protein kinase-related protein
MGCFGCCGGGEDFRRVSETGPKPVH	6.9	At3g17410.1	KINASE- Pto	no	serine/threonine protein kinase, putative similar to Pto kinase interactor 1 GB:AAC61805 from [Lycop
MGSSFRSSSSRNEHEDGGDEAEHAL	9.2	At2g37280.1	MISC	no	ABC transporter family protein similar to PDR5-like ABC transporter GI:1514643 from [Spirodela polyr
MGGAGGTVTATSPGTSSAENNTWSS	3.4	At1g49720.1	MISC	no	abscisic acid responsive elements-binding factor identical to abscisic acid responsive elements-bind

Sequence (Arabidopsis predict myristoylated)	HM$_{266}$ Score	ID	function	train	description
MGNASAGERARNTGVQANELVESAL	3.3	At3g06810.1	MISC	no	acetyl-coA dehydrogenase -related similar to acetyl-coenzyme A dehydrogenase GB:NP_031408 from [Mus
MGMNASLHTTMSDPSNLNHHSIGQA	0.6	At3g23280.2	MISC	no	auxin-regulated protein contains Pfam profile: PF00023 ankyrin repeat
MGIVTSFAFPEASSSGDNSKVLPFI	0.6	At1g64170.1	MISC	no	cation/hydrogen exchanger, putative (CHX16) monovalent cation:proton antiporter family 2 (CPA2) memb
MGCLRSLVRRKQFDSSNGKAETHHH	7.5	At5g36940.1	MISC	no	cationic amino acid transporter -related protein cationic amino acid transporter-1, rat, PIR:T46813
MGTASSKSNINLSQVLNTEEPYSSV	2.7	At4g22710.1	MISC	no	cytochrome P450 family contains Pfam profile: PF00067 cytochrome p450
MGASHSHEDLEICSSDEDEYEEYEE	6.4	At4g33400.1	MISC	yes	Dem -related protein Dem (defective embryo and meristems) protein -Lycopersicon esculentum, PID:e321
MGTSQSREDRISESDTESDSDYYEE	6.7	At3g19240.1	MISC	yes	dem-related protein similar to dem GB:CAA73973 from [Lycopersicon esculentum] (Plant Cell (1998) 10
MGNSFGCSASGERLVSAARDGDFVE	7.9	At5g07270.1	MISC	no	expressed protein ankyrins - different species
MGNVTSNVAAKFAFFPPEPATYGVT	0.8	At4g31020.1	MISC	no	expressed protein BEM46 PROTEIN, Schizosaccharomyces pombe, PID:D1022245 (alpha beta hydrolase)
MGNSLRCISQEQDPNQKKPSSVVNG	8.9	At1g69240.1	MISC	no	hydrolase, alpha/beta fold family low similarity to SP|Q40708 PIR7A protein {Oryza sativa}, polyneur
MGNSFTCISHEQEQRPKKSSGGGGN	8.6	At1g26360.1	MISC	no	hydrolase, alpha/beta fold family similar to SP|Q40708 PIR7A protein {Oryza sativa}, ethylene-induce
MGNLCSLFTPPKPVKKRKPITKRQS	8.4	At3g29770.1	MISC	no	hydrolase, alpha/beta fold family similar to SP|Q40708 PIR7A protein {Oryza sativa}, polyneuridine a
MGCAGAAATDSPIQMNGLPISIYAP	2.0	At3g16150.1	MISC	no	L-asparaginase -related similar to L-ASPARAGINASE GB:P30364from [Lupinus angustifolius]
MGTKVTVISTSERKRDEAVTRLGAD	0.9	At4g37980.1	MISC	no	mannitol dehydrogenase (ELI3-1), putative identical to GI:16267

Sequence (Arabidopsis predict myristoylated)	**HM_{266} Score**	**ID**	**function**	**train**	**description**
MGKSGGRKKKSGGSNSNSSQVNSSE	0.9	At1g62390.1	MISC	no	octicosapeptide/Phox/Bem1p (PB1) domain-/tetratricopeptide repeat (TPR)-containing protein contains
MGCTSSKQAKANVVADVYKPPPSSF	11.6	At1g64500.1	MISC	no	peptide transporter -related
MGNGNSREAKESRRSRLRHKLQKFR	5.9	At5g57190.1	MISC	no	phosphatidylserine decarboxylase, putative similar to SP|P53037 Phosphatidylserine decarboxylase pro
MGNGNSTETKESRRSKMRKKIQNFR	0.9	At4g25970.1	MISC	no	phosphatidylserine decarboxylase, putative similar to SP|P53037 Phosphatidylserine decarboxylase pro
MGSVASLALAIVVKGKDSSKLWGLV	1.5	At4g18130.1	MISC	no	phytochrome E (PHYE) identical to SP|P42498 Phytochrome E {Arabidopsis thaliana}
MGNLHGIHRSHHGGSNFPGEAPNSP	5.2	At3g09540.1	MISC	no	polysaccharide lyase family 1 (pectate lyase) simliar to style development-specific protein 9612 SP:
MGLCFSSTYATVDVMPHEVDTEVPD	5.2	At3g49500.1	MISC	no	RNA-dependent RNA polymerase SDE1 RNA-directed RNA polymerase
MGHATSLSHFLILSSSRFSRLGSLT	0.7	At2g25840.1	MISC	no	trytophanyl-tRNA synthetase -related
MGNTRRSGGGGGRNNKNKKRSRKGS	1.6	At4g34910.1	NUCLEAR	no	DEAD/DEAH box helicase, putative
MGHFASNCEGKPKKRAGESDEKGDG	1.3	At1g75660.1	NUCLEAR	no	Dhp1-related protein similar to Dhp1 protein GB:BAA04601 GI:496393 from [Schizosaccharomyces pombe] Exoribonuclease
MGSHSSSQLSSELVEEGSERAIVSP	2.7	At4g25540.1	NUCLEAR	no	DNA mismatch repair protein MSH3 identical to SP|O65607 DNA mismatch repair protein MSH3 (AtMsh3) {A
MGSSFSASFTNSTTAAAVPPPSPPS	8.7	At1g18680.1	NUCLEAR	no	expressed protein contains similarity to DNA dependent reverse transcriptase GI:2920563 from [Spragu
MGSGESGSFSPSPPQKQLKIENDML	1.8	At3g01410.1	NUCLEAR	no	RNase H domain-containing protein low similarity to GAG-POL precursor [Oryza sativa (japonica cultiv
MGASASVPERSVHQFTVKDSSGKDL	5.3	At2g48150.1	PEROX	no	glutathione peroxidase, putative
MGASSSSSVSEKSIHQFTVKDSSGK	1.2	At3g63080.1	PEROX	no	glutathione peroxidase, putative phospholipid-hydroperoxide glutathione peroxidase, spinach, PIR:JC5
MGCYVVGSAGSKEKVDLLKTKFGFD	1.6	At5g16980.1	PEROX	no	NADP-dependent oxidoreductase, putative strong similarity to probable NADP-dependent oxidoreductase
MGSCVSKGKGDDDSVHNVEFSGGNV	16.9	At3g08710.1	PEROX	no	thioredoxin family similar to thioredoxin H-type GB:P29448 SP|P29448 [Arabidopsis thaliana], Thiored

Sequence (Arabidopsis predict myristoylated)	**HM$_{266}$ Score**	**ID**	**function**	**train**	**description**
MGGALSTVFGSGEDATAAGTESEPS	9.2	At5g39950.1	PEROX	yes	thioredoxin H-type 2 (TRX-H-2) identical to SP|Q38879 Thioredoxin H-type 2 (TRX-H-2) {Arabidopsis th
MGANVSTPDQRFQVTHFRSTKPWTP	8.9	At1g69880.1	PEROX	no	thioredoxin, putative similar to SP|Q38879 Thioredoxin H-type 2 (TRX-H-2) {Arabidopsis thaliana}; co
MGSGASKNTEEDDDGSNGGGGQLYV	12.8	At1g07110.1	PHOS	yes	fructose-2,6-bisphosphatase -related similar to fructose-2,6-bisphosphatase GI:8572069 from [Arabido
MGGCVSTSSKSTCSSWSNGEKPVRR	14.6	At4g03415.1	PHOS	no	protein phosphatase 2C (PP2C) -related temporary automated functional assignment
MGSCLSAESRSPRPGSPCSPAFSVR	13.5	At3g02750.1	PHOS	no	protein phosphatase 2C (PP2C) similar to putative protein phosphatase-2C (PP2C) GB:AAC36699 [Mesembr
MGSCLSSSGGGGSRRSLHGSPHVPG	9.1	At5g36250.1	PHOS	no	protein phosphatase 2C (PP2C), putative
MGTYLSSPKTEKLSEDGENDKLRFG	4.2	At2g25070.1	PHOS	no	protein phosphatase 2C (PP2C), putative
MGFCFCLSSGGSTDKSQIYEITDYG	8.5	At4g32950.1	PHOS	no	protein phosphatase 2C (PP2C), putative phosphoprotein phosphatase, Arabidopsis thaliana, PIR2:S5545
MGYLNSVLSSSSQVHSDDGPVSGGG	2.9	At4g31750.1	PHOS	no	protein phosphatase 2C (PP2C), putative protein phosphatase 2C, Medicago sativa, PID:g2582800
MGIYLSTPKTDKFSEDGENHKLRYG	4.3	At4g31860.1	PHOS	no	protein phosphatase 2C (PP2C), putative protein phosphatase 2C, Schizosaccharomyces pombe, PIR2:S542
MGVCCSKGTGIIVEHGADDGNECGD	12.2	At5g01700.1	PHOS	no	protein phosphatase 2C (PP2C), putative protein phosphatase type 2C - Saccharomyces cerevisiae, EMBL
MGHCFSLPSSQSEIHEDNEHGDGNV	8.8	At5g26010.1	PHOS	no	protein phosphatase 2C (PP2C), putative protein phosphatase-2C, Mesembryanthemum crystallinum, AF075
MGLCHSKIDKTTRKETGATSTATTT	12.4	At1g16220.1	PHOS	no	protein phosphatase 2C -related similar to GB:AAC36699
MGLCYSVDRTTGKEPGEASSTATTA	11.0	At1g79630.1	PHOS	no	protein phosphatase-2C -related similar to protein phosphatase-2C GI:3643088 from [Mesembryanthemum
MGLKLSRGPVKEKSPLEFTRVHILT	2.3	At5g39400.1	PHOS	yes	PTEN -related protein PTEN/MMAC1, Rattus norvegicus, EMBL:AF017185

Sequence (Arabidopsis predict myristoylated)	HM$_{266}$ Score	ID	function	train	description
MGNGFSYGSSSSSFGNEERNDGSAS	9.2	At5g56110.1	TRANSCRIPT	no	myb family transcription factor contains PFAM profile: Myb DNA binding domain PF00249
MGGSGSSSSSISNLAENKSSGSLLL	7.9	At5g23000.1	TRANSCRIPT	no	myb family transcription factor contains PFAM profile: myb DNA binding domain PF00249;
MGRCGRSNDGVIGGVRPYVRSPVPR	2.3	At1g14600.1	TRANSCRIPT	no	myb family transcription factor contains Pfam profile: PF00249 myb-like DNA-binding domain
MGFDFSTSKSKAKRQKRIEVRFASP	1.1	At1g68150.1	TRANSCRIPT	no	WRKY family transcription factor similar to DNA-binding protein ABF2 GI:1159879 from [Avena fatua]
MGQGPSGGLNRQGDRKPDGGEKKEK	0.7	At2g20140.1	UBQ	no	26S proteasome subunit 4
MGANISGGSPEFDRNDDVYSRKLRL	6.4	At3g61060.1	UBQ	no	F-box protein (lectin-related) low similarity to PP2 lectin polypeptide [Cucurbita maxima] GI:410437
MGAASSSIVRSEPFAGKLCGLEDVP	6.0	At5g52120.1	UBQ	no	F-box protein (SKP1 interacting partner 3-related) contains Pfam profile PF00646: F-box domain
MGVAHSDLHNDLSSSSCFGDRNILK	3.0	At1g12710.1	UBQ	no	F-box protein (SKP1 interacting partner 3-related) contains Pfam profile PF00646: F-box domain
MGHKQSVDSRGKGRKVPGSSSMVQK	2.3	At2g02250.1	UBQ	no	F-box protein (SKP1 interacting partner 3-related) contains similarity to SKP1 interacting partner 3
MGSSLCVAVRKKEKQKQKKTVVFLP	2.0	At3g52320.1	UBQ	no	F-box protein family contains F-box domain Pfam:PF00646
MGSVMSLSCSKRKATSQDVECSSES	5.4	At1g22040.1	UBQ	no	Kelch repeat containing F-box protein family contains Pfam profiles PF01344: Kelch motif, PF00646: F
MGAAGSKLEKALGDQFPEGERYFGF	7.1	At4g39910.1	UBQ	yes	ubiquitin-specific protease 3 (UBP3) identical to GI:2347098
MGAAGSKLEKALGDQFPEGERYFGF	7.1	At2g22310.1	UBQ	(yes)	ubiquitin-specific protease 4 (UBP4) identical to GI:2347100
MGNAIRLLRKCLNSHGVSASSGGVS	4.5	At3g56170.1	UNKN	no	
MGGKSSKSKKNVEFGSPSTPVQIKI	9.4	At2g18630.1	UNKN	no	At14a, putative similar to At14a, GI:11994571 and GI:11994573 [Arabidopsis thaliana]; unusual splice
MGCVCGKPSAIEDSKDSPRERFSSK	18.1	At1g53050.1	UNKN	no	cell division-related protein -related similar to GB:CAB89490 from [Beta vulgaris]
MGNKTSRKSKEKSGRNYTTELRSYE	9.5	At4g34330.1	UNKN	no	conserved hypothetical common family with hypothetical protein F10M10_90, Arabidopsis thaliana; simi
MGACASKPKESDIVEGSVSTENAVV	17.3	At4g27580.1	UNKN	no	expressed protein
MGASNVRSKSRKLVSPEQVGNNRSK	4.4	At5g11530.1	UNKN	no	expressed protein

Sequence (Arabidopsis predict myristoylated)	HM_{266} Score	ID	function	train	description
MGASSSTDDKESSEKREIESLAAST	5.0	At5g39590.1	UNKN	no	expressed protein
MGCAGSTQSQADGSVKKIRKPKPWK	14.2	At1g53400.1	UNKN	no	expressed protein
MGCAQSTIAVVGRKNKKRIIQESVV	2.1	At5g14020.1	UNKN	no	expressed protein
MGCASSKHRKRCLHCRRGYSPPVDV	16.3	At3g28850.1	UNKN	no	expressed protein
MGCASSKNRNRCRNCKGGLSPVIVP	9.7	At5g39865.1	UNKN	no	expressed protein
MGCFASRPNDTGGNRRKPTSIGDVS	13.3	At1g73390.1	UNKN	no	expressed protein
MGCFMGCFGLSSNKKRRNSIRKILP	6.4	At2g33400.1	UNKN	no	expressed protein
MGCGGSRLGGVATTRAEEGGVVPLP	5.7	At5g50830.1	UNKN	no	expressed protein
MGGCASKPKESDIPETEKTVVESKN	18.3	At5g54095.1	UNKN	no	expressed protein
MGGCASRPKESDMQNEEGSVPNKPV	18.2	At3g49540.1	UNKN	no	expressed protein
MGGCTSKQERRGVRSVKETSKDQSR	9.3	At3g23930.1	UNKN	no	expressed protein
MGGCVSSQRKLSNKLQQKKQKRGRS	12.6	At1g13970.1	UNKN	no	expressed protein
MGGCVSTHSKAIRPRRKGRRRSSKH	18.1	At3g29180.1	UNKN	no	expressed protein
MGGLCSRSSSVNNAPGGTFAHVNGH	6.4	At1g34320.1	UNKN	no	expressed protein
MGGRGGGSTDYDDKVGRKRKKEKER	4.1	At1g32130.1	UNKN	no	expressed protein
MGGSSSSSPTSSSSSPAKRGKNKNG	5.1	At4g28840.1	UNKN	no	expressed protein
MGGVTSSIAAKFAFFPPTPPSYEVI	1.8	At1g66900.1	UNKN	no	expressed protein
MGISGSKRVTTILSNSPEFDSACDS	1.6	At2g43540.1	UNKN	no	expressed protein
MGITSSTDPKTETFSDPKSQNNEEA	3.4	At1g14060.1	UNKN	no	expressed protein
MGLCVSVNRNEYVSSSTTAKIVTIN	1.3	At1g76600.1	UNKN	no	expressed protein
MGLFSSSSGDLRYKFQVCHENQVKS	1.0	At2g24240.1	UNKN	no	expressed protein
MGNACCVAARDKMVVPNSSAGENLQ	5.5	At4g39140.2	UNKN	no	expressed protein
MGNACCVATRDKMVVPNTSAVENFQ	5.1	At2g21500.2	UNKN	no	expressed protein
MGNALVKKEPPPPVVLVPPLFDYPP	3.8	At3g27930.1	UNKN	no	expressed protein
MGNCALKPKVLSETGAPAPEELKDS	1.9	At4g39320.1	UNKN	no	expressed protein
MGNCCVSSGTTDRENENVSDKNTTI	17.0	At1g61170.1	UNKN	no	expressed protein
MGNCLRHESEMHWAGEDWDEFITED	1.9	At3g20340.1	UNKN	no	expressed protein
MGNCLSTSDPSHEDVSKKLPKALPV	13.0	At2g44260.2	UNKN	no	expressed protein
MGNFLRRSCAVHPASGDGGRGRHVD	2.1	At3g61930.1	UNKN	no	expressed protein
MGNGIGKLSKCLTGGAGRNKKPELS	8.5	At2g28890.1	UNKN	no	expressed protein
MGNGTSRVVGCFVPSNDKNGVDLEF	6.9	At2g46920.2	UNKN	no	expressed protein
MGNGVTKLSICFTGGGGERLRPKDI	7.5	At1g07630.1	UNKN	no	expressed protein
MGNLFVKKPQITEVDRAILSLKTQR	0.8	At5g63880.1	UNKN	no	expressed protein
MGNNSSAQSSTPSLPIDSTFNLPSP	5.3	At2g44230.1	UNKN	no	expressed protein
MGNSVRSNLRDIRGRRSMDPRMWHK	7.2	At2g04900.1	UNKN	no	expressed protein

Sequence (Arabidopsis predict myristoylated)	HM$_{266}$ Score	ID	function	train	description
MGNYVSCALNKTSSSPLAKVILPDG	1.4	At3g03280.1	UNKN	no	expressed protein
MGNYVSSALSKTSSSSSSAAKVILP	7.9	At5g17350.1	UNKN	no	expressed protein
MGRLSSSMKKIEKGLKNGNKYQDED	1.3	At2g45740.2	UNKN	no	expressed protein
MGSAGVASSSSDVAISALREKHEKE	4.5	At4g14950.1	UNKN	no	expressed protein
MGSCVSSSQSKSSSASDSVVKLASF	8.9	At3g01860.1	UNKN	no	expressed protein
MGSCVVILSDGEVPVKADASTIAIS	2.4	At2g22400.1	UNKN	no	expressed protein
MGSFCSKSLGINFGSEYSGSSVADD	11.6	At5g08660.1	UNKN	no	expressed protein
MGSFHRRTFSYDKLPNEPIRLSVLK	2.6	At5g25340.1	UNKN	no	expressed protein
MGSLMSGWDSRVRDPKSVRRCKSLT	1.0	At1g19530.1	UNKN	no	expressed protein
MGSRLAESSTPMKACEEVGRPSFTS	1.0	At1g06660.1	UNKN	no	expressed protein
MGSSQSSQLLDEEEEEEDEAESEGE	3.8	At1g12850.1	UNKN	no	expressed protein
MGSSVSKTASSASASSITSILNSPP	8.9	At2g41330.1	UNKN	no	expressed protein
MGSVCCVAAKDRNVPSGAIDNSVCS	4.0	At1g19680.1	UNKN	no	expressed protein
MGSVCCVAVKDRKVPPSGGPASSSV	6.6	At1g75400.1	UNKN	no	expressed protein
MGTCFSSSTKSTAEISPFDLVVKPP	13.6	At3g54290.1	UNKN	no	expressed protein
MGWNFSKAPDNSKAYQNLPADEAAI	2.1	At3g22435.1	UNKN	no	expressed protein
MGSGFSSLLPCFNQGHRNRRRHSSA	10.2	At2g35350.1	UNKN	no	expressed protein and genefinder
MGGVTSSVAAKFAFFPPSPSSYKLV	3.1	At5g14390.1	UNKN	no	expressed protein CGI-67, Homo sapiens, EMBL:AF151825
MGISLSKRRRDNNNNHHHPHHNPPY	7.9	At5g19080.1	UNKN	no	expressed protein Common family members At5g03200, At3g09770 [Arabidopsis thaliana]
MGGVPSTPRKTGGDDVSVAEYLIAT	1.0	At1g04200.1	UNKN	no	expressed protein Contains similarity to gb|Z69902 from C. elegans
MGFTVVSRSGSPTRRENQKRRVCFR	1.3	At1g79950.1	UNKN	no	expressed protein contains similarity to helicase-like protein NHL GI:6969265 from (Homo sapiens)
MGGCVSTPKSCVGGKIRSSKRRKTR	17.3	At1g59650.1	UNKN	no	expressed protein CW14 annotation temporarily based on supporting cDNA gi|6520232|dbj|AB028233.1|
MGGCVGCYREHRSTAASLKDPPSNS	2.3	At1g15350.2	UNKN	no	expressed protein ESTs gb|R65145, gb|N96612 and gb|R90096 come from this gene
MGGVCSCVFKDDDKKKKLRSNDDDK	14.4	At1g30755.1	UNKN	no	expressed protein ESTs gb|T43746 and gb|AA721821 come from this gene.
MGSSASKTASSVSSASSSSAFSPPP	8.2	At3g57070.1	UNKN	no	expressed protein hypothetical protein At2g41330 - Arabidopsis thaliana, EMBL:AC005662
MGNVISGGSRPENHRDRTSPPYPNP	15.0	At3g53410.1	UNKN	no	expressed protein KIAA0544 protein, Homo sapiens, EMBL:AB011116

Sequence (Arabidopsis predict myristoylated)	**HM_{266} Score**	**ID**	**function**	**train**	**description**
MGSLSSEEDDEVSSERCGSYSPSAD	3.4	At4g38430.1	UNKN	no	expressed protein other predicted proteins Arabidopsis thaliana
MGGSASLQSDSMEAAAAADSSIPRD	4.5	At3g62970.1	UNKN	no	expressed protein PGP237-11, Petunia x hybrida, EMBL:AF049930
MGGSTSKDERNSSKRRIKVKANEQR	1.7	At4g13540.1	UNKN	no	expressed protein predicted protein, Arabidopsis thaliana
MGLFSSSSGELRYKFQVSHDNQVKS	0.8	At4g30940.1	UNKN	no	expressed protein predicted protein, Caenorhabditis elegans, PID:G1707206
MGGVTSSIAAKFAFFPPSPPSYGFV	3.8	At5g38220.1	UNKN	no	expressed protein predicted proteins, Arabidopsis thaliana
MGGSYSGSASMAFEYHETRAPVTES	9.6	At4g02880.1	UNKN	no	expressed protein similar to A. thaliana hypothetical protein F15K9.11, GenBank accession number AC0
MGGCVSTPKSCVGAKLRSSKRRKSR	16.6	At1g10410.1	UNKN	no	expressed protein similar to ESTs gb|N96021 and gb|N96863
MGCGGSKVDNQPIVILCRERKELLK	5.6	At1g77500.1	UNKN	no	expressed protein similar to GB:AAD41429
MGCMFSHLAAKFAFFPPSPPTYHLT	1.6	At1g32190.1	UNKN	no	expressed protein similar to hypothetical protein GB:AAD18105 GI:4337191 from [Arabidopsis thaliana]
MGGCVSSNLLTPTDDSSFSHLTSSS	14.1	At1g32760.1	UNKN	no	expressed protein similar to hypothetical protein GB:AAF25972 GI:6714276 from [Arabidopsis thaliana]
MGNLFCCVLVKQSDVAVKERFGKFQ	0.6	At3g01290.1	UNKN	no	expressed protein similar to hypothetical protein GB:CAA10289 (Cicer arietinum)
MGGVTSSVAAKFAFFPPSPPSYKVV	3.3	At3g01690.1	UNKN	no	expressed protein similar to putative protein GB:CAA22987 [Arabidopsis thaliana]
MGRRSSGGRSAPRPRPAAARSPAPQ	0.8	At5g64400.1	UNKN	no	expressed protein similar to unknown protein (emb|CAB89373.1)
MGCLLGCFGRRKNRRRQRRRESFQP	7.2	At5g44040.1	UNKN	no	expressed protein similar to unknown protein (gb|AAD10667.1)
MGRASSSLSPGHRHFAGQANELITE	0.7	At5g47380.1	UNKN	no	expressed protein similar to unknown protein (pir||T02421)
MGACASRESLRSDSAKLILLDGTLQ	4.5	At5g66580.1	UNKN	no	expressed protein similar to unknown protein (pir||T08401)
MGLFLVLFPDHNNNDSSSYYSPAKT	1.1	At5g52500.1	UNKN	no	expressed protein strong similarity to unknown protein (emb|CAB68146.1)

Sequence (Arabidopsis predict myristoylated)	HM_{266} Score	ID	function	train	description
MGACASRESRRTETAKLILPDGTLQ	4.7	At3g50800.1	UNKN	no	expressed protein unknown protein F27L4.13 Arabidopsis thaliana chromosome II BAC F27L4, PID:g315261
MGCAQSRVDNEEAVARCKERRNVIK	11.1	At5g25590.1	UNKN	no	expressed protein various predicted proteins, Arabidopsis thaliana
MGAAVALHFATSCALNHYTINPRVV	1.3	At1g52460.1	UNKN	no	hypothetical protein
MGASCSAPVSSTIKKNWRDHPPSS	5.1	At1g69650.1	UNKN	no	hypothetical protein
MGASDSTLLGSQENRGDVITTISHR	0.8	At5g66290.1	UNKN	no	hypothetical protein
MGCCFSAGKVTAPPVAAAQIPPPLE	14.7	At4g28180.1	UNKN	no	hypothetical protein
MGCCISKCSPKSKDFKEAEEPHEKY	15.7	At1g21510.1	UNKN	no	hypothetical protein
MGCCLSKKPSPSLPSSVKPSDPIKP	17.2	At1g65710.1	UNKN	no	hypothetical protein
MGCFASTPKDSGGNRRKPTSIGEVS	13.3	At1g17940.1	UNKN	no	hypothetical protein
MGCLFGCFRAKDDESTTDSVSQAKH	6.7	At2g30820.1	UNKN	no	hypothetical protein
MGCSPSKLDGLPAVSLCRDRCSSLE	3.4	At1g20530.1	UNKN	no	hypothetical protein
MGCSSSRTIAEGKKEKIRRPKTWKH	4.9	At1g16960.1	UNKN	no	hypothetical protein
MGGGGGSVSNANGGVPASSRKVIQD	1.2	At4g18150.1	UNKN	no	hypothetical protein
MGIVASKGDSNTPLLNLCKERKELI	3.0	At5g54480.1	UNKN	no	hypothetical protein
MGNCICVTEKTTTSWSGDDNGSYNK	8.2	At4g21920.1	UNKN	no	hypothetical protein
MGNCLRHDNGVARKEKDDLDPEPLV	1.3	At3g21680.1	UNKN	no	hypothetical protein
MGNSLGTKKTATIMNINGESFKLKT	1.3	At1g71015.1	UNKN	no	hypothetical protein
MGNTSSCAPLIISTNSSSGVVKILA	2.3	At1g18290.1	UNKN	no	hypothetical protein
MGNVTSNMAAKFAFFPPPPTYDVGK	2.1	At2g24320.1	UNKN	no	hypothetical protein
MGSIYARQVREDPNWSGSGSRRYSG	6.6	At2g19090.1	UNKN	no	hypothetical protein
MGTCMSTHSRRLRPRRKGRRRFSKN	13.0	At5g39430.1	UNKN	no	hypothetical protein
MGCVSSNLLNHDEDFSQIGGGSSAF	8.5	At4g10630.1	UNKN	no	hypothetical protein Arabidopsis thaliana chromosome II BAC F13H10.12,PID:g3894191
MGNQTSKKSQETSAKSVHYTTELRS	4.0	At4g34320.1	UNKN	no	hypothetical protein common family with F10M10_100, Arabidopsis thaliana; similar to At14a, GI:11994
MGASGRWIKALVGFTKSDKSRSSKK	1.2	At3g22190.1	UNKN	no	hypothetical protein contains PF00612 IQ:IQ calmodulin-binding motif
MGAVTSSMAAKFAFFPPNPPSYGVE	0.9	At3g30380.1	UNKN	no	hypothetical protein contains Pfam profile: PF00561 alpha/beta hydrolase fold
MGIISSRILKPEYRSQYNNQIPMQE	1.6	At5g02210.1	UNKN	no	hypothetical protein hypothetical protein F7A19.14 - Arabidopsis thaliana, EMBL:EMBL:AC007576

Sequence (Arabidopsis predict myristoylated)	**HM_{266} Score**	**ID**	**function**	**train**	**description**
MGGFLVLNSSNNVVRTVPSKKRKHP	5.1	At5g20440.1	UNKN	no	hypothetical protein probable mitosis and maintenance of ploidy protein, Schizosaccharomyces pombe,
MGVSFSCPFAEQDDVEAALDSVTVK	8.0	At3g13600.1	UNKN	no	hypothetical protein similar to hypothetical protein GB:CAB68186 GI:6735365 from [Arabidopsis thalia
MGNKDSKGSSSSTNPKEEIAAVPQP	11.1	At5g06010.1	UNKN	no	hypothetical protein similar to unknown protein (gb|AAF04428.1)
MGNGVASFSGCCAGTTAGEISGRYV	3.9	At3g09400.1	UNKN	no	hypothetical protein similar to unknown protein GB:AAC79593 [Arabidopsis thaliana]
MGNLFVKKPKITEVDRAILSLKTQR	1.8	At5g09260.1	UNKN	no	hypothetical protein SNF7 protein - Saccharomyces cerevisiae, PIR:S52590
MGCCASNTAGGSKANRISRWRSTGI	8.0	At5g07910.1	UNKN	no	leucine rich repeat protein family contains leucine rich repeat (LRR) domains, Pfam:PF00560
MGGLISLQSLDMSSNSLSGPLPKSL	1.7	At4g20940.1	UNKN	no	leucine rich repeat protein family contains leucine rich-repeat (LRR) domains Pfam:PF00560, INTERPRO
MGCCQSRIDSKEIVSRCKARKRYLK	15.6	At3g51290.1	UNKN	no	proline-rich protein family
MGCAASRIDNEEKVLVCRQRKRLMK	9.0	At2g34670.1	UNKN	no	proline-rich protein family contains proline-rich region, INTERPRO:IPR000694
MGCCLSKNPPQASSSSLIDGSCSVK	14.0	At5g37010.1	UNKN	no	serine-rich protein -related hypothetical protein F1E22.8 - Arabidopsis thaliana, EMBL:AC007234
MGGSSGGGVSYRSGGESDVELEDYE	2.9	At3g02850.1	UNKN	no	
MGCCCCLPSIPESSRTIDEHLPLSR	8.9	At4g23450.1	UNKN-ZINC	no	C3HC4-type zinc finger protein family contains Pfam profile: PF00097 zinc finger, C3HC4 type
MGCVSSCFGVDDFEDYPNPSSSVNR	8.1	At5g15790.1	UNKN-ZINC	no	C3HC4-type zinc finger protein family contains Pfam profile: PF00097 zinc finger, C3HC4 type (RING f
MGGCCCCSSSRRADVDNGPAYYYYP	17.5	At5g41350.1	UNKN-ZINC	no	C3HC4-type zinc finger protein family contains Pfam profile: PF00097 zinc finger, C3HC4 type (RING f
MGGCCSSSRKSHLVGTPVYYYCPES	15.9	At4g00335.1	UNKN-ZINC	no	C3HC4-type zinc finger protein family contains Pfam profile: PF00097 zinc finger, C3HC4 type (RING f
MGSGRSKSCLGSSAMAREASDVAQP	0.8	At5g45290.1	UNKN-ZINC	no	C3HC4-type zinc finger protein family contains Pfam profile: PF00097 zinc finger, C3HC4 type (RING f

Sequence (Arabidopsis predict myristoylated)	HM_{266} Score	ID	function	train	description
MGCVSSCFRVEDIDEYMNPNSSVYR	4.4	At3g02290.1	UNKN-ZINC	no	C3HC4-type zinc finger protein family contains zinc finger motif, C3HC4 type (RING finger)
MGLNHGVGLNNNNNNGGFNGISTGG	1.8	At5g62940.1	UNKN-ZINC	no	Dof zinc finger protein Dof zinc finger protein, Oryza sativa, EMBL:AB028129
MGNISSSGGEGRRRRRRNHTAAPPP	5.7	At3g09770.2	UNKN-ZINC	no	putative RING zinc finger protein annotation temporarily based on supporting cDNA gi|18377643|gb|AY0
MGQSMSCGSRPEHGIFASVQCGDII	5.2	At2g28840.1	UNKN-ZINC	no	RING zinc finger ankyrin protein -related
MGISFSNNNRRRDNNNRRHLHHYPP	7.7	At3g06140.1	UNKN-ZINC	no	RING zinc finger protein -related contains Pfam profile: PF00097 Zinc finger, C3HC4 type (RING finge
MGFLSSSTTRKEAKTTRPNKAHPST	0.7	At2g41835.1	UNKN-ZINC	no	zinc finger protein (PMZ) -related temporary automated functional assignment

Supplementary Table 5. Non-Arabidopsis plant proteins predicted to contain N-terminal myristoylation sites.

Sequence (non AT predict myrist)	**HMM$_{266}$ Score**	**ID**	**Description**
MGLCFSSAAKSSGHNRSSRNPHPHP	19.2	gi|15227920|ref|NP_179379.1|	
MGVCFSKASKPSGNGYQPPPPQPYK	18.3	gi|28866604|emb|CAD70165.1|	calcium-dependent protein kinase [Spirodela punctata]
MGHCCSKGVTADNDGHVVSVVDGNS	17.8	gi|16904222|gb|AAL30818.1|AF435450_1	calcium/calmodulin-dependent protein kinase CaMK1 [Nicotiana tabacum]
MGCVHGRPSTSSPAAANASRRRDHP	17.5	gi|23495876|dbj|BAC20085.1|	putative CRK1 protein(cdc2-related kinase 1) [Oryza sativa (japonica cultivar-group)]
MGNCCVTPEGSGRGRKKQQQEQKQK	17.0	gi|34393291|dbj|BAC83205.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGSCCSRATSPDSGRGGANGYGYSH	16.8	gi|23616997|dbj|BAC20693.1|	CDP2_ORYSA Calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGACASRPKDLDRDLAPAPVPAEES	16.0	gi|5360192|gb|AAD42902.1|AF161330_1	unknown [Solanum chacoense]
MGCCFSVIDDDGYEHDDAVSGSVNE	15.9	gi|32488315|emb|CAE02892.1|	OSJNBa0015K02.9 [Oryza sativa (japonica cultivar-group)]
MGGCFSKKYTQQDANGHRAGRRVNQ	15.4	gi|19171502|emb|CAC87494.1|	calcium-dependent protein kinase [Lycopersicon esculentum]
MGACASRPKDLDKDLAPAPVPAEDA	15.3	gi|6019472|gb|AAD02545.2|	PGPS/D3 [Petunia x hybrida]
MGACFSSHTATAAADGGSGKRQQRK	15.2	gi|6689920|gb|AAF23900.1|AF194413_1	calcium-dependent protein kinase [Oryza sativa]
MGCSSSVPARSTGGLNNISNDNSAT	15.0	gi|5931625|dbj|BAA84717.1|	rab5B [Oryza sativa (japonica cultivar-group)]
MGSCLAKPAESAEQKKASSVPSPPP	15.0	gi|33591148|gb|AAQ23078.1|	gliding motility related CaM kinase [Chlamydomonas reinhardtii]
MGQCCSRAPAPDSGRGGTNGYGYSN	15.0	gi|7428011|pir||T02993	calcium-dependent protein kinase (EC 2.7.1.-) 9 - maize
MGNCCVTPPQTGSPLKNKKNKPNPF	14.9	gi|2665890|gb|AAB88537.1|	calcium-dependent protein kinase [Fragaria x ananassa]
MGGCHAKPLTHEEADGACSPPRERP	14.9	gi|33146627|dbj|BAC79915.1|	putative CDPK-related protein kinase [Oryza sativa (japonica cultivar-group)]
MGICASKGKPNNANNGHHGSGSGGV	14.9	gi|14029712|gb|AAK52801.1|AF363784_1	calcium-dependent protein kinase CDPK1 [Lycopersicon esculentum]
MGCVFGRPSSSPPAAAARKGRRKKR	14.8	gi|15528679|dbj|BAB64745.1|	putative CRK1 protein [Oryza sativa (japonica cultivar-group)]
MGASCSSPFPPPPSPAESNAPSASA	14.6	gi|20146453|dbj|BAB89233.1|	pectate lyase-like protein [Oryza sativa (japonica cultivar-group)]
MGICVSKPSPEPDLHNHHTSIPVND	14.5	gi|1706130|sp|P53681|CRK_DAUCA	CDPK-related protein kinase (PK421)

Sequence (non AT predict myrist)	**HMM$_{266}$ Score**	**ID**	**Description**
MGVCLSKSKPAESKSDGHYRSGGSD	14.3	gi|10568116|gb|AAD28192.2|	calcium-dependent protein kinase [Solanum tuberosum]
MGSCVGKERSDEEDKIDFKGGNVHV	14.3	gi|20804820|dbj|BAB92503.1|	thioredoxin-like protein [Oryza sativa (japonica cultivar-group)]
MGNCFSQSTQTRGSKRNQKPDPLDE	14.3	gi|28866606|emb|CAD70166.1|	putative calcium-dependent protein kinase [Spirodela punctata]
MGCGLSREKDAGGGPRRRPGSVGDV	14.1	gi|32487665|emb|CAE05655.1|	OSJNBa0038O10.21x [Oryza sativa (japonica cultivar-group)]
MGSCFSSEGGNESRKESAERPQITP	14.0	gi|24796805|gb|AAN64481.1|	putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGGCFSKKYRQEGANGGYRATRRNA	13.9	gi|7428009|pir||T01989	calcium-dependent protein kinase (EC 2.7.1.-) 1 - common tobacco
MGACLSSSSVIDAKDQKPSSAYVIS	13.9	gi|13699079|dbj|BAB41200.1|	TMV response-related gene product [Nicotiana tabacum]
MGCCGSSLRVGSHAPEKPPRRARPP	13.9	gi|31433335|gb|AAP54864.1|	protein kinase-like protein [Oryza sativa (japonica cultivar-group)]
MGCVFGKESLAPEGRERGGRREKEN	13.9	gi|7688002|emb|CAB89665.1|	CRK1 protein [Beta vulgaris subsp. vulgaris]
MGLCSSSSARRDAGTPGGGNGAGNK	13.8	gi|12592069|gb|AAF23901.2|AF194414_1	calcium-dependent protein kinase [Oryza sativa]
MGICASKNKATQPEDNGYTPANGVG	13.8	gi|4336426|gb|AAD17800.1|	
MGGCFSKKEYQADGNGYRSAPTAYH	13.6	gi|7434365|pir||T14335	protein kinase, calcium-dependent (EC 2.7.1.-) - carrot
MGCLLSKDKDSEQEHNNGGYRYAES	13.6	gi|13561063|emb|CAA65500.1|	protein kinase [Medicago sativa]
MGTNYSKPTTSINDASNLSNRVPFE	13.6	gi|8547236|gb|AAF76311.1|AF220603_3	LescPth2 [Lycopersicon esculentum]
MGCAGSTPKTDDNSKKLKKPKPWKH	13.6	gi|15289832|dbj|BAB63530.1|	P0435H01.26 [Oryza sativa (japonica cultivar-group)]
MGLCHGKPSQIPEPEAEEAAAAAGV	13.5	gi|23237898|dbj|BAC16472.1|	putative calcium/calmodulin-dependent protein kinase CaMK [Oryza sativa (japonica cultivar-group)]
MGSCFSSSKVSGSNSNTPSTNNTAT	13.4	gi|3779218|gb|AAC78558.1|	protein kinase CPK1 [Solanum tuberosum]
MGQCCSKGVSGENGGSVVAIGDGNS	13.3	gi|19547871|gb|AAL87457.1|	serine/threonine protein kinase pk23 [Lycopersicon esculentum]
MGCTASKVEQEDTVRRCKERRRHMK	13.2	gi|31433523|gb|AAP55028.1|	putative leucine zipper protein [Oryza sativa (japonica cultivar-group)]
MGTNYSKPTTSINDASNLRNRVPFE	13.1	gi|13021867|gb|AAK11569.1|AF318493_1	Pto-like protein kinase D [Lycopersicon hirsutum]
MGNCVGGHSDRKNALNPNEPAATNV	12.9	gi|5162878|dbj|BAA81749.1|	calcium-dependent protein kinase [Marchantia polymorpha]

Sequence (non AT predict myrist)	**HMM$_{266}$ Score**	**ID**	**Description**
MGCVSSKILSKSGSFQEKVVSHGFK	12.8	gi|5257280|dbj|BAA81779.1|	Similar to neurofilament triplet M protein. (P07197) [Oryza sativa (japonica cultivar-group)]
MGGLCSKGSAVDKSPSDTTLGPDRV	12.8	gi|20146508|dbj|BAB89287.1|	P0491F11.23 [Oryza sativa (japonica cultivar-group)]
MGCGSSLPDRDSRPFGRDNPENGGG	12.8	gi|1370180|emb|CAA98167.1|	RAB5B [Lotus corniculatus var. japonicus]
MGNCVGASNRSTSRENKGNHRVSRR	12.7	gi|31415931|gb|AAP50952.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNCCRSPAAAAREDVKSSHFPASA	12.7	gi|20805246|dbj|BAB92912.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGLCFSKKQQVRRRREEQQQPPCHE	12.5	gi|32483146|emb|CAE02138.1|	OSJNBa0074L08.6 [Oryza sativa (japonica cultivar-group)]
MGGAFSTSKPKPAAGEEGGESAVVA	12.3	gi|34394553|dbj|BAC83857.1|	putative thioredoxin [Oryza sativa (japonica cultivar-group)]
MGNCVARSGTAVDAGGDGGEDGKRR	12.3	gi|33146759|dbj|BAC79670.1|	putative protein phosphatase 2C [Oryza sativa (japonica cultivar-group)]
MGCFQSKSRRPFPGHEDPVLLASQT	12.3	gi|27650414|emb|CAD33259.1|	calcineurin B [Crocus sativus]
MGCFLSKPAGAGPLPPNDAAALPAD	12.3	gi|31433613|gb|AAP55105.1|	putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSINDASNSSYRVPFE	12.2	gi|16326617|gb|AAL17825.1|	Pto-like protein [Solanum nigrum]
MGNAGSSPEQAAGNNKEDAEARRQP	12.2	gi|18844859|dbj|BAB85328.1|	OJ1656_A11.17 [Oryza sativa (japonica cultivar-group)]
MGQCYGKARGASSRADHDADPSGAG	12.0	gi|7446432|pir||T02033	calcium/calmodulin-dependent protein kinase homolog - maize
MGNANGREDGAIPDAGDPSGREPHA	11.7	gi|32364486|gb|AAO61676.1|	AKIN beta1 [Medicago truncatula]
MGSSCSRPHSVNEAEAAGNTRSADI	11.5	gi|11245655|gb|AAF71788.2|AF267485_1	G protein alpha subunit [Hordeum vulgare]
MGLLCSKHQHSTKPDAENAQATGIE	11.4	gi|7441569|pir||T09152	GTP-binding regulatory protein alpha chain - spinach
MGGCAGKVRRDDEEKLDFKGGNVHI	11.4	gi|24637227|gb|AAN63617.1|AF435816_1	thioredoxin h-like protein [Zea mays]
MGSSCSRPHSVNEAEAADNTRSADI	11.4	gi|22336032|dbj|BAC10501.1|	G protein alpha subunit [Triticum aestivum]
MGLLCSKNRRYNDADTEENTQTAEI	11.3	gi|1346103|sp|P49082|GBA1_LOTJA	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGLLCSRNKGYNQADDEENTQTADI	11.1	gi|18369796|dbj|BAB84093.1|	heterotrimeric GTP binding protein alpha subunit [Nicotiana tabacum]
MGNCCGWGARIKDGSPQPGASGMFS	11.1	gi|27545044|gb|AAO18450.1|	putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGCVASKNAVSVTPAADSSGALRER	11.1	gi|33146814|dbj|BAC79804.1|	putative cyclin-dependent kinase CDC2C [Oryza sativa (japonica cultivar-group)]

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGASGSKPEGAGPGPALPTPSPAPP	11.0	gi|9022433|gb|AAF82382.1|AF280543_1	putative copper chaperone Cox17 [Chlamydomonas reinhardtii]
MGSLCSRNKHYSQADDEENTQTAEI	11.0	gi|3913725|sp|P93564|GBA1_SOLTU	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGNACRGSFRAKTFQGSYPQNHSTS	10.9	gi|31616519|gb|AAP55748.1|	calcium-dependent protein kinase 3 [Capsicum annuum]
MGNRTSRHKHGEKHRSEAGSHQTHI	10.9	gi|7434369|pir||T10938	calcium-dependent protein kinase (EC 2.7.1.-) - sweet potato
MGSKYSKATNSISDASNSRYGVPFE	10.8	gi|8547235|gb|AAF76310.1|AF220603_2	LescPth3 [Lycopersicon esculentum]
MGNCFGSEEAEAAATVRAPAQGHHG	10.7	gi|32488580|emb|CAE03088.1|	OSJNBa0017B10.3 [Oryza sativa (japonica cultivar-group)]
MGAGTSKNSDSTSHGDGDGEERDEY	10.6	gi|7484665|pir||T08994	6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2, 6-bisphosphate 2-phosphatase (EC 3.1.3.46) - spinach
MGNAAAVPSSCASADAAPSTTPTIK	10.6	gi|13872953|dbj|BAB44058.1|	P0030H07.22 [Oryza sativa (japonica cultivar-group)]
MGLLCSKSNRYNDAKAEENAQTAEI	10.5	gi|3913708|sp|O04278|GBA1_PEA	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGSKYSKATNSINDASNLSYGVPFE	10.4	gi|8547231|gb|AAF76307.1|	Fen kinase [Lycopersicon pimpinellifolium]
MGGCVGKGRSIVEEKLDFKGGNVHV	10.3	gi|27543511|gb|AAO16555.1|	thioredoxin h [Leymus chinensis]
MGACNSCEATAVAAVNGRSASGEAT	10.2	gi|31126784|gb|AAP44703.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSINDASNLSYGIPFE	10.1	gi|8547239|gb|AAF76314.1|	Fen kinase [Lycopersicon esculentum]
MGLLCSRNRRYNDADAEENAQTAEI	10.1	gi|3913724|sp|P93163|GBA2_SOYBN	Guanine nucleotide-binding protein alpha-2 subunit (GP-alpha-2)
MGLLCSRNRRYNDADAEENAQAAEI	9.9	gi|3334195|sp|Q40224|GBA1_LUPLU	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)
MGLVCSRNRRYRDSDPEENAQAAEI	9.9	gi|3913709|sp|O04279|GBA2_PEA	Guanine nucleotide-binding protein alpha-2 subunit (GP-alpha-2)
MGCGGSKEAVATGNTVGSTGGRSSR	9.8	gi|20160760|dbj|BAB89701.1|	P0504E02.1 [Oryza sativa (japonica cultivar-group)]
MGANHSREDLELSDSESESEYGSES	9.8	gi|7488984|pir||T07737	dem protein - tomato
MGNANAREDGAAVDGDGDGEVSGRR	9.5	gi|5702015|emb|CAB52141.1|	GAL83 protein [Solanum tuberosum]
MGGCVGKDRSIVEDKLDFKGGNVHV	9.5	gi|11362710|pir||T50864	thioredoxin-like protein [imported] - Hordeum bulbosum
MGQCYGKAGGASSRRADHDDAVAPP	9.3	gi|7446431|pir||T03023	calcium-dependent protein kinase-related protein kinase - maize

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGNTCRGSFKGNIFQGYSQPEDSST	9.3	gi|20453013|gb|AAL68971.1|	phloem calmodulin-like-domain protein kinase PCPK1 [Cucurbita maxima]
MGNTCVGPSSAADRHGFFHSVSLAV	9.2	gi|33146743|dbj|BAC79646.1|	putative calcium-dependent protein kinase 2 [Oryza sativa (japonica cultivar-group)]
MGAGVSGLFGLGGDEGETSAAVGGA	9.2	gi|31433129|gb|AAP54682.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGCGQSKIDQEEAVCRCRDRRKLMA	9.2	gi|15408647|dbj|BAB64061.1|	putative bZIP (leucine zipper) protein [Oryza sativa (japonica cultivar-group)]
MGNTCRGSFKGNIFQGYSQPEESST	9.1	gi|20453015|gb|AAL68972.1|	calmodulin-like-domain protein kinase CPK2 [Cucurbita maxima]
MGCTNSTEARREMVLADAGRRATRS	9.0	gi|32490043|emb|CAE05962.1|	OSJNBa0063C18.3 [Oryza sativa (japonica cultivar-group)]
MGLVCSRNRRYRDSDPGENAQAAEI	9.0	gi|22476942|gb|AAM97352.1|	G protein alpha II subunit [Pisum sativum]
MGCISSKLLPPGPGDARGGSGARAT	9.0	gi|21741722|emb|CAD40645.1|	OSJNBa0073L04.18 [Oryza sativa (japonica cultivar-group)]
MGGGQSRSPRDGSGHGRYGHSPSFQ	9.0	gi|20804929|dbj|BAB92608.1|	P0456E05.7 [Oryza sativa (japonica cultivar-group)]
MGGCVGKDRGIVEDKLDFKGGNVHV	8.8	gi|11362711|pir||T50862	thioredoxin-like protein [imported] - Phalaris coerulescens
MGCAQSNEDGEGPVARCRERKHLLR	8.7	gi|23237834|dbj|BAC16409.1|	putative bZIP protein [Oryza sativa (japonica cultivar-group)]
MGSSCSRSHSLSEAETTKNAKSADI	8.7	gi|540533|dbj|BAA07405.1|	G protein alpha subunit [Oryza sativa (japonica cultivar-group)]
MGASGSKLEKALGDQFPEGERYFGL	8.6	gi|21740732|emb|CAD40853.1|	OSJNBa0086B14.25 [Oryza sativa (japonica cultivar-group)]
MGNRTSRHHRAAPEQPPPQPKPKPQ	8.5	gi|12313676|dbj|BAB21081.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGNCSPSPRRRGRPTSPSGSPPLDR	8.4	gi|20521256|dbj|BAB91772.1|	P0679C12.24 [Oryza sativa (japonica cultivar-group)]
MGALLSSPNSKNQPWEHGEASKADS	8.4	gi|31431565|gb|AAP53323.1|	putative kelch-containing protein [Oryza sativa (japonica cultivar-group)]
MGSGCASSPPTSPQPRHGRLVTVLS	8.4	gi|14346205|emb|CAC41149.1|	unnamed protein product [Triticum aestivum]
MGCFSCCGADDVGKKKKRDDPYVPI	8.3	gi|29027802|dbj|BAC65877.1|	OSJNBa0011P19.1 [Oryza sativa (japonica cultivar-group)]
MGNNCVGPSAAGQNGFFANVALWRP	8.3	gi|27819504|gb|AAO24908.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSINDALNPSYLVPFE	8.3	gi|32330888|gb|AAP79929.1|	Pto-like serine/threonine kinase [Capsicum annuum]

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGGCVGKGRGVVEEKLDFKGGNVHV	8.2	gi|24637225|gb|AAN63616.1|AF435815_1	thioredoxin h-like protein [Hordeum vulgare subsp. vulgare]
MGGALSIVFGSGEDATAAGTESSEP	8.2	gi|1388078|gb|AAC49353.1|	thioredoxin h
MGSCASVHKDLGFPKKLFLASSPTK	8.2	gi|19386733|dbj|BAB86115.1|	P0678F11.22 [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSINDALNSSYLVPFE	8.2	gi|32330884|gb|AAP79927.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGLCHGKSAAVLEPTVEEEEEGATR	8.0	gi|14133874|gb|AAK54157.1|AF368282_1	CaMK1 [Oryza sativa]
MGGALSSVLGGDAAAAAASADDSSS	7.8	gi|33621080|gb|AAQ23133.1|	thioredoxin H2 [Ipomoea batatas]
MGSFFSTMFTPPPAADDGGDSRVVA	7.8	gi|28209505|gb|AAO37523.1|	putative thioredoxin [Oryza sativa (japonica cultivar-group)]
MGGCVGKDRSIVEEKLDFKGGNVHV	7.8	gi|24637237|gb|AAN63622.1|AF438359_1	thioredoxin [Triticum aestivum]
MGNACGGALRSKHQHSFKHAASSQR	7.5	gi|7428005|pir||T03263	calcium-dependent protein kinase (EC 2.7.1.-) 7 - maize
MGSCVSTTRRRRRSRKLSVAARKFR	7.5	gi|14488359|gb|AAK63926.1|AC084282_7	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSINDALSSSYLVPFE	7.4	gi|32330886|gb|AAP79928.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGAAVAGEPRHNNANGGHREREGIK	7.4	gi|31430970|gb|AAP52815.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGASASSSAGRSKPPPPPPCSSSGV	7.4	gi|31432227|gb|AAP53889.1|	putative speckle-type protein [Oryza sativa (japonica cultivar-group)]
MGGVTSSTSSSSPATAAGGSSYSAE	7.2	gi|20160988|dbj|BAB89922.1|	putative hypersensitive-induced response protein [Oryza sativa (japonica cultivar-group)]
MGNSCSSVDESAAAIATLAESAAAG	7.1	gi|33146803|dbj|BAC79751.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGMAAGRSKPMAASSSERGRRWSAR	7.0	gi|18087678|gb|AAL58970.1|AC091811_19	hypothetical protein [Oryza sativa]
MGSKYSKATNSINDALSSGYLVPFE	7.0	gi|32330890|gb|AAP79930.1|	Pto-like serine/threonine kinase [Capsicum annuum]
MGQGQSGGLPKQGPGGERNKGDGQK	7.0	gi|18479040|gb|AAL73395.1|AF432345_1	26S proteasome regulatory subunit IV [Tortula ruralis]
MGCWYSRIDREEIVSRCKARKRYMK	7.0	gi|22858664|gb|AAN05792.1|	unknown [Gossypium hirsutum]
MGLAFGKLFSRLFAKKEMRILMVGL	6.9	gi|27261078|dbj|BAC45192.1|	ADP-ribosylation factor 1 [Oryza sativa (japonica cultivar-group)]
MGCLSVVGASPCVSSSALSSPTSRL	6.7	gi|27807824|dbj|BAC55280.1|	nucleoside diphosphate kinase [Nicotiana tabacum]
MGNGGGGNKKGRRDSPSPPLETPPR	6.7	gi|31432873|gb|AAP54449.1|	putative lanthionine synthetase [Oryza sativa (japonica cultivar-group)]
MGNTCVGPSISKNGFFESVSAAMWR	6.6	gi|7434364|pir||T09940	calcium-dependent protein kinase (EC 2.7.1.-) CDPK - pumpkin

Sequence (non AT predict myrist)	HMM_{266} Score	ID	Description
MGLSHARSGKRCNDFKVGGELSCRG	6.6	gi|13873011|dbj|BAB44115.1|	B1085F09.11 [Oryza sativa (japonica cultivar-group)]
MGACASASRPPPSPRHGCSPPTPTM	6.5	gi|29467524|dbj|BAC66713.1|	P0582D05.14 [Oryza sativa (japonica cultivar-group)]
MGSACSRKRGQLLVDEEDLYSARFS	6.5	gi|32488227|emb|CAE02935.1|	OSJNBa0014K14.7 [Oryza sativa (japonica cultivar-group)]
MGGVLSSVLGGGGGDEPLAGNESES	6.5	gi|25990392|gb|AAN76509.1|AF352030_1	thioredoxin h [Brassica rapa]
MGVLFSCPVDDETAAVEDAAPVAEQ	6.5	gi|22296385|dbj|BAC10154.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGLTSSKGGRCRQVGPTGQRLKGGR	6.4	gi|34395360|dbj|BAC84430.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGCTHTRHSRKRYQRRVTTRRAEHL	6.4	gi|8467997|dbj|BAA96598.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGAASSRLEKALGEQFPEGERYFGL	6.4	gi|24756890|gb|AAN64154.1|	Putative ubiquitin-specific protease 3 [Oryza sativa (japonica cultivar-group)]
MGASTSTSNPSLEAQEQETLASASL	6.4	gi|32489048|emb|CAE04845.1|	OSJNBa0084K01.17 [Oryza sativa (japonica cultivar-group)]
MGLCLSSGVAAAMAEAGMAASTAMV	6.3	gi|34394505|dbj|BAC83793.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGSKYSKATNSISDASNSFESYRFP	6.2	gi|8547238|gb|AAF76313.1|	Pto kinase [Lycopersicon esculentum]
MGAVVSHASAAAPPPPPTGYAPFVD	6.2	gi|6539563|dbj|BAA88180.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGCFLACFGDRRRRRQPRRQSPARS	6.2	gi|32488453|emb|CAE03386.1|	OSJNBa0004N05.10 [Oryza sativa (japonica cultivar-group)]
MGCASSAPDRTVSRGSSSLNSANEG	6.2	gi|7438400|pir||T12437	small GTP-binding protein - common ice plant
MGVSASCLWGGSESRGNQNGSAAVT	6.2	gi|29893573|gb|AAP06827.1|	putative receptor ser/thr protein [Oryza sativa (japonica cultivar-group)]
MGCKGSKLEDQEAVALCRGRAELLA	6.1	gi|32483081|emb|CAE01923.1|	OSJNBb0078D11.6 [Oryza sativa (japonica cultivar-group)]
MGNSLGCSASGERLVSAARDGDAVE	6.0	gi|31433119|gb|AAP54672.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNTCVGPSISKNGLFQSVSAAMWR	5.9	gi|16215467|emb|CAC82998.1|	calcium-dependent protein kinase 2 [Nicotiana tabacum]
MGLTISKLFSRLFAKKEMRILMVGL	5.9	gi|1703376|sp|P51824|ARF1_SOLTU	ADP-ribosylation factor 1
MGNLVSSIATSSLKRRNSILREPLF	5.9	gi|17222567|gb|AAL36740.1|AF353999_20	apocytochrome b [Mesostigma viride]

Sequence (non AT predict myrist)	**HMM$_{266}$ Score**	**ID**	**Description**
MGNTCVGPSISKNGIFQSVSAAMWR	5.8	gi|16215471|emb|CAC82999.1|	calcium-dependent protein kinase 3 [Nicotiana tabacum]
MGQCYARNVPVDGEGGGGVVTTTTT	5.8	gi|31433008|gb|AAP54572.1|	putative kinase [Oryza sativa (japonica cultivar-group)]
MGGCAAQAARADAAATGQSAARERA	5.7	gi|31126743|gb|AAP44664.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGCRMRSGSGLQDPFEGVREAGVDG	5.7	gi|20303607|gb|AAM19034.1|AC084748_24	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNGGSTAHRPRVRVHRNSGGDDEA	5.6	gi|10800066|dbj|BAB16486.1|	P0665D10.11 [Oryza sativa (japonica cultivar-group)]
MGLSFTKLFSRLFAKKEMRILMVGL	5.6	gi|1703375|sp|P51822|ARF1_DAUCA	ADP-ribosylation factor 1
MGCAGSKDVTVADVYRPPPTSVSLF	5.5	gi|7339706|dbj|BAA92911.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGHSARKKKKKKGGGGRKAAKDHGG	5.5	gi|32480225|emb|CAE02048.1|	OJ990528_30.6 [Oryza sativa (japonica cultivar-group)]
MGNTCVGPSITMNGFFQSVSTALWK	5.4	gi|7428007|pir||T02784	calcium-dependent protein kinase (EC 2.7.1.-) - maize (strain W64A)
MGSSRSEGSSSSSARQPYGSPIPYR	5.3	gi|31432406|gb|AAP54043.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNCCPGSGDAEPASSDASTGNGSS	5.3	gi|20160680|dbj|BAB89623.1|	putative calcium-dependent protein kinase [Oryza sativa (japonica cultivar-group)]
MGSCGAGAADSEWPDGVTGRDAEVG	5.3	gi|21328092|dbj|BAC00676.1|	B1144D11.23 [Oryza sativa (japonica cultivar-group)]
MGNACGGSLRSKYLSFKQTASQRHD	5.2	gi|32492297|emb|CAE04157.1|	OSJNBa0088A01.28 [Oryza sativa (japonica cultivar-group)]
MGQAFRKLFDAFFGNKEMRVVMLGL	5.2	gi|31433708|gb|AAP55187.1|	putative ADP-ribosylation factor [Oryza sativa (japonica cultivar-group)]
MGGCVGKGRRHIEEDKLDFKGGNVH	5.2	gi|24637229|gb|AAN63618.1|AF435817_1	thioredoxin h-like protein [Oryza sativa]
MGCFGRRPPSSSPPHMERREILVKA	5.2	gi|18461223|dbj|BAB84420.1|	P0690B02.8 [Oryza sativa (japonica cultivar-group)]
MGNRLCCHDAADDEKPAAAAVSGGG	5.2	gi|20503002|gb|AAM22711.1|AC098693_16	Putative protein kinase [Oryza sativa (japonica cultivar-group)]
MGNTCVGPSISKIGLFQSVSAAMWR	5.1	gi|16215475|emb|CAC83000.1|	calcium-dependent protein kinase 2 [Nicotiana benthamiana]
MGNSLGGKKKIAKVMKVDGTTFKLK	5.1	gi|16265861|gb|AAL16667.1|AF419849_1	unknown [Musa acuminata]
MGLVCSRSRRFREAHAEENAQDAEI	4.9	gi|1346105|sp|P49084|GBA1_SOYBN	Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1)

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGLTFTKLFSRLFSKKEMRILMVGL	4.6	gi|7242921|dbj|BAA92519.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGGGGGGSDGDDERRTVDPARRTYN	4.6	gi|23616960|dbj|BAC20663.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNAIRLLRKCLNSHGVSASSGGVS	4.5	gi|18410440|ref|NP_567036.1|	
MGAACGSAGPAHPACGSVAALTTPA	4.4	gi|31712062|gb|AAP68368.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNRARRRPGGSGDEENEQRKVGGS	4.3	gi|34393262|dbj|BAC83132.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGSAGALLSHSPPGLGGFPPRHHHH	4.3	gi|15408654|dbj|BAB64068.1|	OsTATC [Oryza sativa (japonica cultivar-group)]
MGLTFTKLFSRLFAKKEMRILMVGL	4.3	gi|1703380|sp|P51823|ARF_ORYSA	ADP-ribosylation factor
MGAGASKDVDGGSLGSQEREVSDGG	4.2	gi|14149116|dbj|BAB55655.1|	fructose-6-phosphate 2-kinase/fructose-2,6-bisphosphatase [Bruguiera gymnorrhiza]
MGNCWGAKISSESPCRSASSPSGGT	4.2	gi|14719339|gb|AAK73157.1|AC079022_30	putative protein kinase [Oryza sativa]
MGNTQGKGDGAMTGRAKAITGGAVK	4.1	gi|11230987|dbj|BAB18105.1|	cyclic nucreotide dependent protein kinase II [Chlamydomonas reinhardtii]
MGLLFGRCRVLQTGLSASTQSNSRP	4.0	gi|6063535|dbj|BAA85395.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGLITSGPLVPLSVGARSSPRWLSG	4.0	gi|22296391|dbj|BAC10160.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGLGASLGAGHAAAGVRRARRVVGG	4.0	gi|15623967|dbj|BAB68022.1|	P0455H0310.11 [Oryza sativa (japonica cultivar-group)]
MGAGHRGGSGELREGAGENFGELPP	3.9	gi|28564575|dbj|BAC57684.1|	putative receptor-like protein kinase [Oryza sativa (japonica cultivar-group)]
MGASASSSGERRRRCGSAAIVADTV	3.9	gi|31432236|gb|AAP53891.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGGSGSSGPFSGSTLKTSSYRSSPI	3.8	gi|20161116|dbj|BAB90045.1|	OSJNBb0049O23.2 [Oryza sativa (japonica cultivar-group)]
MGTRSSGSSSSAQRRRSDLPLIKCP	3.6	gi|27817838|dbj|BAC55606.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGTAVAAACAASGDSGDSGPRGDRS	3.6	gi|31415905|gb|AAP50926.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGAGMSSGRRPSFYGGGGGGGVGRP	3.5	gi|15408827|dbj|BAB64219.1|	P0581F09.16 [Oryza sativa (japonica cultivar-group)]
MGLSFTKLLGRLFSKKEMRILMVGL	3.5	gi|11131023|sp|O48649|ARF_SALBA	ADP-ribosylation factor 1

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGNSCVGAAGYFQGFTSVIALGGRS	3.4	gi|2315983|gb|AAB70706.1|	calmodulin-like domain protein kinase [Tortula ruralis]
MGLSFTKLLGRLFSEKEMRILMVGL	3.4	gi|29124979|gb|AAO63779.1|	ADP-ribosylation factor 1 [Populus tremuloides]
MGGAAAAHATTGDGKRRRERRQRTA	3.4	gi|32487740|emb|CAE05490.1|	OSJNBa0022H21.10 [Oryza sativa (japonica cultivar-group)]
MGHGVSCARTGDEHDFFRAAQLGDL	3.3	gi|14279688|gb|AAK58690.1|AF272860_1	receptor-like kinase Xa21-binding protein 3 [Oryza sativa]
MGQCAMSHTRSTTAGAAGGKPPIAP	3.3	gi|31432198|gb|AAP53860.1|	putative speckle-type protein [Oryza sativa (japonica cultivar-group)]
MGCILGKLAAAPGSSLFFPAAAAAS	3.2	gi|20805217|dbj|BAB92884.1|	putative CRK1 protein [Oryza sativa (japonica cultivar-group)]
MGARCSKLSVCWWPPHFKSPLLENG	3.2	gi|15042834|gb|AAK82457.1|AC091247_24	putative protein kinase [Oryza sativa]
MGNCFTKTYEIPITSGTMRRPASTA	3.2	gi|21741166|emb|CAD41068.1|	OSJNBa0084K11.12 [Oryza sativa (japonica cultivar-group)]
MGHASTDPSASDHDAPARRRRHRRR	3.2	gi|20804479|dbj|BAB92175.1|	OSJNBa0016I09.27 [Oryza sativa (japonica cultivar-group)]
MGAAVAGEPRHKYANGGHREKEREA	3.2	gi|32489252|emb|CAE03984.1|	OSJNBa0033H08.16 [Oryza sativa (japonica cultivar-group)]
MGLRFTKALSRLFGKKEMRILMVGL	3.1	gi|1703374|sp|P51821|ARF1_CHLRE	ADP-ribosylation factor 1
MGGGGGGGRGGGDHHSPTSGKTPRG	3.1	gi|20804522|dbj|BAB92216.1|	B1015E06.17 [Oryza sativa (japonica cultivar-group)]
MGCCSCFGFLRKPRVSVSRPRDADG	3.1	gi|13324784|gb|AAK18832.1|AC082645_2	putative protein kinase [Oryza sativa]
MGNGGGASGSRGCGLRQQRQRLSVQ	3.1	gi|32488775|emb|CAE04328.1|	OSJNBb0016D16.19 [Oryza sativa (japonica cultivar-group)]
MGNLGVREAGRMAVLSRRWRRLPGL	3.0	gi|34393595|dbj|BAC83248.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGLFTVTKKATTPFDGQKPGTSGLR	3.0	gi|12585309|sp|P93804|PGM1_MAIZE	Phosphoglucomutase, cytoplasmic 1 (Glucose phosphomutase 1) (PGM 1)
MGRGATSSSSSSFESSNYPSVSTKS	3.0	gi|20269055|emb|CAC84708.1|	aux/IAA protein [Populus tremula x Populus tremuloides]
MGVLTGRPRRPPFAGSAGGANDFLA	3.0	gi|20805277|dbj|BAB92943.1|	OJ1460_H08.9 [Oryza sativa (japonica cultivar-group)]
MGACTSKPSNFSVDDITVAGDGAIF	2.9	gi|16904226|gb|AAL30820.1|AF435452_1	calcium/calmodulin-dependent protein kinase CaMK3 [Nicotiana tabacum]
MGGKLCLVAARSAHPSGRSAAVWPV	2.9	gi|11875149|dbj|BAB19362.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]

<table>
<tr><th>Sequence (non AT predict myrist)</th><th>HMM$_{266}$ Score</th><th>ID</th><th>Description</th></tr>
<tr><td>MGGSSGGGVSYRSGGESDVELEDYE</td><td>2.9</td><td>gi|15232991|ref|NP_186934.1|</td><td></td></tr>
<tr><td>MGLSIALPGVSLPPLQPIKDPAVPE</td><td>2.8</td><td>gi|32489554|emb|CAE03795.1|</td><td>OSJNBa0063G07.19 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGCTTSRQARHDLRHCPSPLALPR</td><td>2.8</td><td>gi|19071642|gb|AAL84309.1|AC073556_26</td><td>hypothetical protein [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGLFTVTKKATTPFEGQKPGTSGLR</td><td>2.8</td><td>gi|12585310|sp|P93805|PGM2_MAIZE</td><td>Phosphoglucomutase, cytoplasmic 2 (Glucose phosphomutase 2) (PGM 2)</td></tr>
<tr><td>MGCTTSHDAFAAARARARASSSSLE</td><td>2.8</td><td>gi|20146230|dbj|BAB89012.1|</td><td>bzip-like transcription factor-like [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGRSRGVLSSGDDDTGHRSKRRRVS</td><td>2.8</td><td>gi|21898562|gb|AAM77037.1|</td><td>histone acetyltransferase complex component [Zea mays]</td></tr>
<tr><td>MGGFCCCLCTDDFEEYAHPNNPIYR</td><td>2.8</td><td>gi|20805222|dbj|BAB92889.1|</td><td>OSJNBa0093F16.9 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGCVCGRAAAVDDGRCGAAAEAAV</td><td>2.7</td><td>gi|20804521|dbj|BAB92215.1|</td><td>putative CRK1 protein [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGGRRRGASSPALGRGRGRRRVAGV</td><td>2.7</td><td>gi|22324481|dbj|BAC10395.1|</td><td>hypothetical protein [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGCAQGKCCVPRRQRGRGGGGAVGG</td><td>2.7</td><td>gi|23307591|dbj|BAC16725.1|</td><td>putative protein phosphatase 2C (PP2C) [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGNSYSNVGSPTAAGYVQAPELPLH</td><td>2.6</td><td>gi|31415968|gb|AAP50988.1|</td><td>hypothetical protein [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGVARVCLDRRSLAADRRGGGSPWP</td><td>2.6</td><td>gi|23617043|dbj|BAC20733.1|</td><td>hypothetical protein [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGSFASPLDAAPAPPTAKPSPPPAP</td><td>2.6</td><td>gi|21742766|emb|CAD40522.1|</td><td>OSJNBa0023J03.9 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGHSPSRHNACGGGGGDGESPPSPL</td><td>2.5</td><td>gi|20146228|dbj|BAB89010.1|</td><td>putative phosphatidylserine decarboxylase [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGCCGGMSSTSRAPRGIREETLLRV</td><td>2.4</td><td>gi|20160545|dbj|BAB89494.1|</td><td>P0022F10.10 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGQYSSTHCGGGPKLRHRNLHPSPL</td><td>2.4</td><td>gi|30013671|gb|AAP03878.1|</td><td>Avr9/Cf-9 rapidly elicited protein 189 [Nicotiana tabacum]</td></tr>
<tr><td>MGNCCGTDGFKHERPLPAHPNRQGE</td><td>2.3</td><td>gi|23268465|gb|AAN11310.1|</td><td>calmodulin domain protein kinase 1 [Ceratopteris richardii]</td></tr>
<tr><td>MGGRGAGGEAHDDPFRSRSSSFGAV</td><td>2.3</td><td>gi|15128233|dbj|BAB62561.1|</td><td>OSJNBa0089K24.19 [Oryza sativa (japonica cultivar-group)]</td></tr>
<tr><td>MGSRRAGSSSSPLFWPAPPSRAADP</td><td>2.3</td><td>gi|82363|pir||S04860</td><td>alpha-amylase/subtilisin inhibitor precursor - barley</td></tr>
<tr><td>MGARASAFLRLEINRQPQRLKRAQP</td><td>2.3</td><td>gi|20804644|dbj|BAB92333.1|</td><td>P0510F09.5 [Oryza sativa (japonica cultivar-group)]</td></tr>
</table>

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGNASGRLDDIADAEMDDGGGGGNR	2.2	gi|22324436|dbj|BAC10353.1|	AKIN beta1-like protein [Oryza sativa (japonica cultivar-group)]
MGNIASSGGASPRPPPPHLEAYRHG	2.2	gi|14587218|dbj|BAB61152.1|	B1045D11.20 [Oryza sativa (japonica cultivar-group)]
MGNMGSSGGHRRRNNGHGRHHHHGQ	2.1	gi|31430853|gb|AAP52712.1|	putative hydroxyproline-rich glycoprotein [Oryza sativa (japonica cultivar-group)]
MGGFAASNIVSESGEKASNLVYRGR	2.1	gi|34394682|dbj|BAC83988.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGPGASKETPTQQQKPAESVSTPPE	2.1	gi|20161674|dbj|BAB90592.1|	P0697D09.21 [Oryza sativa (japonica cultivar-group)]
MGNCAVTQHAVSWADDGEWELPESG	2.0	gi|20146273|dbj|BAB89055.1|	P0684C02.11 [Oryza sativa (japonica cultivar-group)]
MGSGCSSPSISLTTIATSHFQSQES	1.9	gi|20138145|sp|O22494|GSHB_LYCES	Glutathione synthetase, chloroplast precursor (Glutathione synthase) (GSH synthetase) (GSH-S)
MGLRLAGGEVAVGSSGRGGGGRWHG	1.9	gi|31415932|gb|AAP50953.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGQCAMSPARSGTTAGSGGGPTPSS	1.9	gi|31432194|gb|AAP53856.1|	putative speckle-type protein [Oryza sativa (japonica cultivar-group)]
MGSILSSLIGGEPAAALESSDHAVK	1.8	gi|27466896|gb|AAO12855.1|	thioredoxin h [Pisum sativum]
MGCGESKHAVATENATIPKNKRSLS	1.8	gi|1216214|emb|CAA65228.1|	glutamic acid-rich protein [Solanum berthaultii]
MGGRCVGAAVKCGEGAAGGAEEGVG	1.7	gi|31431866|gb|AAP53578.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGGGGGSSRSSKVYHERQRLQFCLL	1.6	gi|29893581|gb|AAP06835.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGCGESKHAVATANATIPKNKRSLS	1.6	gi|21435776|gb|AAM53961.1|AF515615_1	TSB [Lycopersicon esculentum]
MGSSSATSNRLLLSKIATSEGHGEN	1.5	gi|8570043|dbj|BAA96743.1|	1-aminocyclopropane-1-carboxylate synthase [Prunus persica]
MGTAGGCERAPRGRGSPAAAPARGG	1.4	gi|20804917|dbj|BAB92597.1|	B1070A12.21 [Oryza sativa (japonica cultivar-group)]
MGARRRRAWRSGPAAWRARGGGGGV	1.3	gi|22535613|dbj|BAC10787.1|	B1096D03.22 [Oryza sativa (japonica cultivar-group)]
MGSRTTRGFGVEDDKGAERGMPSFN	1.3	gi|15148924|gb|AAK84889.1|AF402608_1	TGA-type basic leucine zipper protein TGA2.1 [Phaseolus vulgaris]
MGSCGGAARAAGSAVTVGKLVKLLL	1.3	gi|18071350|gb|AAL58209.1|AC090882_12	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGLLLVLNLRSATSLSGSGGLIAVH	1.2	gi|31431566|gb|AAP53324.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]

Sequence (non AT predict myrist)	**HMM$_{266}$ Score**	**ID**	**Description**
MGGRRRRGGGEKTPYVEESDNAVRP	1.2	gi|14090218|dbj|BAB55479.1|	P0494A10.19 [Oryza sativa (japonica cultivar-group)]
MGHGLSCSRDTDEYDLFRAAQLGDI	1.2	gi|15290134|dbj|BAB63825.1|	putative receptor-like kinase Xa21-binding protein 3 [Oryza sativa (japonica cultivar-group)]
MGVLFSCPVDDYDALEESAAAAAAA	1.1	gi|31249708|gb|AAP46201.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGNCNCLERARAKVTAWGAEEGDGG	1.1	gi|27261040|dbj|BAC45156.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGNNCFSSSKVSGSNSNTPSTTATA	1.1	gi|16904224|gb|AAL30819.1|AF435451_1	calcium/calmodulin-dependent protein kinase CaMK2 [Nicotiana tabacum]
MGITHANSITTEPSPSPRPPLLRLL	1.1	gi|24059896|dbj|BAC21362.1|	aminoacyl peptidase-like [Oryza sativa (japonica cultivar-group)]
MGGAVARAPGNQKQKSDLAGSQQRR	1.0	gi|22535687|dbj|BAC10861.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
MGASTSTEQKVSIEQREAEALAAST	1.0	gi|27450531|gb|AAO14626.1|AF467900_3	hypothetical protein [Prunus persica]
MGMAASRRWAGGKRGWRRSDRGELL	0.9	gi|14090213|dbj|BAB55474.1|	P0494A10.14 [Oryza sativa (japonica cultivar-group)]
MGLLSSLRPHRRGALSGGGGGGGGQ	0.9	gi|32482882|emb|CAE02255.1|	OSJNBb0032E06.14 [Oryza sativa (japonica cultivar-group)]
MGFAVVRTNREFVRPSAATPPSSGE	0.9	gi|8468020|dbj|BAA96620.1|	unnamed protein product [Oryza sativa (japonica cultivar-group)]
MGIAFSKNWWSRFSGSSPLSVMVCE	0.8	gi|34394463|dbj|BAC83676.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
MGNCMGSSKCASASAPLESKVTRSS	0.8	gi|33146495|dbj|BAC79604.1|	putative serine/threonine-specific protein kinase [Oryza sativa (japonica cultivar-group)]
MGYLSSVIPTDGSPVSGGGLSQNGK	0.8	gi|32488314|emb|CAE02891.1|	OSJNBa0015K02.8 [Oryza sativa (japonica cultivar-group)]
MGGSSAPPSSTSSHNNSTFSPDFPS	0.7	gi|7488838|pir||T06536	protein-tyrosine-phosphatase (EC 3.1.3.48) - garden pea
MGASASVLSLPTAAAIPGAATAIAG	0.7	gi|32488872|emb|CAE03544.1|	OSJNBa0060D06.10 [Oryza sativa (japonica cultivar-group)]
MGGRVSKAVACCCCRSQHHGVVVES	0.7	gi|32488705|emb|CAE03448.1|	OSJNBa0088H09.6 [Oryza sativa (japonica cultivar-group)]
MGSSYGTVDVEAIREAQRAQGLATI	0.7	gi|1403038|emb|CAA62683.1|	chalcone sythase [Gerbera hybrid cv. 'Terra Regina']
MGGGGGSGSSSISPLVSFVLGAAMA	0.7	gi|28269397|gb|AAO37940.1|	putative regulatory protein [Oryza sativa (japonica cultivar-group)]
MGSALSGSGSGSLLVLSVSVVGILF	0.7	gi|28071301|dbj|BAC55990.1|	P0705A05.3 [Oryza sativa (japonica cultivar-group)]

Sequence (non AT predict myrist)	**HMM_{266} Score**	**ID**	**Description**
MGEAASTEEGDADAIKEDEAASTEE	0.6	gi|14587269|dbj|BAB61187.1|	OSJNBb0032H19.20 [Oryza sativa (japonica cultivar-group)]

Supplementary Table 6. Predicted internal myristoylation sites in *Arabidopsis thaliana.*

Sequence (predicted myrist - internal)	HMM_{266} Score	ID	Start pos	Description
MGSIYARQVREDPNWSGSGSRRYSG	19.4	At2g19090.1	744	hypothetical protein
MGAIYARQLKEDTSPVTDSSRKYSR	16.3	At4g30130.1	631	expressed protein bZIP protein - Arabidopsis thaliana, PID:g600855
MGGSSSKESPRGGGSGRRYERSVSG	12.7	At1g67800.2	21	copine-related low similarity to SP|Q99829 Copine I {Homo sapiens}
MGSKFSKNGKTAKRSISVEEENLVL	12.6	At2g35160.1	195	SET-domain transcriptional regulator family identical to SUVH5 [Arabidopsis thal
MGAAFSTQESDGSSTSENNRAYSSY	10.2	At5g07060.1	316	RRM-containing RNA-binding protein, putative
MGSSSSSPKPDNVQEAEKNEFASLS	9.8	At1g53670.1	50	transcriptional regulator -related found within GB:AE000666 from [Methanobacteri
MGGSYSGSASMAFEYHETRAPVTES	9.6	At4g02880.1	85	expressed protein similar to A. thaliana hypothetical protein F15K9.11, GenBank
MGNGFSYGSSSSSFGNEERNDGSAS	9.2	At5g56110.1	214	myb family transcription factor contains PFAM profile: Myb DNA binding domain PF
MGNGAAKVRGSLERKKSKSFKEGDS	9.1	At3g12710.1	36	expressed protein contains similarity to 3-methyladenine-DNA glycosidase I GB:P0
MGSSISTSKTSVWPPEDPSAKIDEA	9.1	At5g61300.1	93	hypothetical protein predicted protein, Arabidopsis thaliana
MGLSYSSNPTPLDNDQKKPSPATAV	9.0	At5g60200.1	21	Dof zinc finger protein similar to dof6 zinc finger protein GI:5689615 from [Ara
MGGGKSNSSSPDSREVREREGIPEL	8.8	At1g21740.1	321	bzip-related transcription factor similar to GB:AAB81673, EST gb|T20649 comes fr
MGMGMGGPYGSQDPNDPFNQPPSPP	8.4	At3g07560.1	168	glycine-rich protein
MGGSGSSSSSISNLAENKSSGSLLL	7.9	At5g23000.1	295	myb family transcription factor contains PFAM profile: myb DNA binding domain PF
MGLCLVRASAETGEEESDDQNKPER	7.9	At4g21445.1	47	receptor-interacting protein (PRIP) -related annotation temporarily based on sup
MGGSSSGNNTLSGPFGNSGVNWGAP	7.8	At3g07810.1	340	heterogeneous nuclear ribonucleoprotein (hnRNP), putative contains Pfam profile:
MGGSSSGNNTLSGPFGNSGVNWGAP	7.8	At3g07810.2	340	heterogeneous nuclear ribonucleoprotein (hnRNP), putative contains Pfam profile:
MGGACSRKRDQQVEDILNRGVSGKY	7.3	At1g15740.1	8	leucine rich repeat protein-related
MGLCCSDKLRMNRGAPPQDVVTAFV	7.1	At5g58690.1	10	phosphoinositide-specific phospholipase C, putative similar to phosphoinositide-
MGGLVSNNSNNSDHNKNCNKGKGKR	7.1	At1g64000.1	71	WRKY family transcription factor similar to WRKY DNA binding protein GB:CAB97004

Sequence (predicted myrist - internal)	HMM_{266} Score	ID	Start pos	Description
MGSGSSVVLDSENGVEAESRKLPSS	7.0	At1g13260.1	36	AP2 domain transcription factor, putative (RAV1) identical to RAV1 GI:3868857 f
MGNYFAVPPGVRVMPQYPGSPLASP	7.0	At3g20250.1	531	pumilio-family RNA-binding protein, putative contains Pfam profile: PF00806 Pumi
MGSNSSTEFPASSTQNSLTRKAGGR	7.0	At5g41150.1	693	repair endonuclease (gb|AAF01274.1)
MGNFCSRPCNLSFGSYSSLREQFGK	6.7	At5g22110.1	318	DNA polymerase epsilon subunit B-related temporary automated functional assignme
MGGSYSRNATNETANVVQQGQDKLK	6.7	At5g53150.1	295	DnaJ protein family low similarity to AHM1 [Triticum aestivum] GI:6691467; conta
MGLSASRATSQFCGRRANKNQAPEG	6.7	At3g42060.1	154	myosin heavy chain-related
MGVNASEPKRTRSRPSQSQGRHPST	6.7	At3g57120.1	23	protein kinase family contains eukaryotic protein kinase domain, INTERPRO:IPR000
MGSGSSSLFGNGNNKREIEHENGIT	6.5	At3g58110.1	527	expressed protein putative protein At2g42370 - Arabidopsis thaliana, TREMBLNEW:
MGLCNSENSSPVSVSVPSEVPATAE	6.4	At1g64380.1	280	AP2 domain transcription factor RAP2, putative contains Pfam profile: PF00847 AP
MGGTVSTSNQQVVPGQQNQFTGIPP	6.3	At1g21630.1	142	calcium-binding EF-hand family protein contains INTERPRO:IPR002048 calcium-bindi
MGLLVSRKDHPGHHQAPYNSNYCVV	6.2	At1g62190.1	220	hypothetical protein similar to unknown protein GI:3445206 from [Arabidopsis tha
MGGVNSKKPQDFGSGVQDPEKNKLF	6.1	At5g05160.1	307	leucine-rich repeat transmembrane protein kinase, putative
MGNCYSLQKDHETALKNFLRAVQLN	6.0	At2g20000.1	524	cell division cycle (CDC) protein - related low similarity to SP|P30260|CC27_HUM
MGTGTSSSWRSQQSQNSYYSHQENE	6.0	At4g02560.1	864	LUMINIDEPENDENS protein
MGSSVSKVVMSDKSKQVESSTDSDS	5.9	At2g34310.1	150	expressed protein
MGSSVSKVVMSDKSKQVESSTDSDS	5.9	At2g34310.2	150	expressed protein
MGSSVSKVVMSDKSKQVESSTDSDS	5.9	At2g34310.3	150	expressed protein
MGSGVSRALSYGFEVGDLVWGKVKS	5.9	At5g27650.1	159	PWWP domain protein hypothetical protein F22F7.12 - Arabidopsis thaliana, EMBL:A
MGQNYRGPSLPPPNMSQNYEGPPPP	5.8	At1g53260.1	297	hypothetical protein low similarity to SP|Q38732 DAG protein, chloroplast precur
MGNLFCSEYRIQRRFDLKGSSHGRS	5.8	At1g21980.1	519	phosphatidylinositol-4-phosphate 5-kinase -related similar to phosphatidylinosit
MGGSSSGSPSYGQRSNNSSKMALKE	5.7	At5g50380.1	567	exocyst subunit EXO70 family contains Pfam domain PF03081: Exo70 exocyst complex

Sequence (predicted myrist - internal)	**HMM_{266} Score**	**ID**	**Start pos**	**Description**
MGCSASLISVDVARDLLQVHPNSNA	5.5	At1g25450.1	218	very-long-chain fatty acid condensing enzyme (CUT1), putative nearly identical t
MGCCCGGTELIAENGTEEFSYGGCR	5.4	At2g16910.1	39	bHLH protein
MGCKVSWTENSVTVTGPPRDAFGMR	5.2	At2g45300.1	369	3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate
MGCKVSWTENSVTVTGPSRDAFGMR	5.2	At1g48860.1	370	3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate
MGSFLAGSTEAPGAYEYRNGRRVKK	5.2	At1g16350.1	376	inosine-5'-monophosphate dehydrogenase, putative strong similarity to SP|P47996
MGLCFSSTYATVDVMPHEVDTEVPD	5.2	At3g49500.1	552	RNA-dependent RNA polymerase SDE1 RNA-directed RNA polymerase
MGGSSRRESETESSLSLDTRNPPRR	5.2	At3g55005.1	150	tonneau 1b (TON1b) annotation temporarily based on supporting cDNA gi|11494366|g
MGGSSSSSPTSSSSSPAKRGKNKNG	5.1	At4g28840.1	11	expressed protein
MGVFLSNSSTAESPQKSLDDNVATG	5.1	At4g30790.1	778	expressed protein
MGAACVRVANELGKGDADAVRFSIK	5.1	At5g17700.1	314	MATE efflux protein family similar to ripening regulated protein DDTFR18 [Lycope
MGNLISVLVHCDERGSEVILGDDSH	5.0	At3g04520.1	68	L-allo-threonine aldolase -related similar to L-ALLO-THREONINE ALDOLASE GB:O0705
MGHSFRVPTSQALKTRTRRRSTAGP	5.0	At4g14210.1	30	phytoene dehydrogenase precursor [phytoene desaturase] (PDS) identical to GI:289
MGISVSSGLRLQIPEESEDQEFVEL	4.9	At2g36360.1	369	Kelch repeat-containing protein low similarity to rngB protein, Dictyostelium di
MGQSTSKFRRSKTTFTSPVLPNLRE	4.8	At4g07400.1	38	F-box protein family (FBL8) (FBL24) contains similarity to SKP1 interacting part
MGNSAGSLFRDKETQKRAGEIHAKE	4.7	At5g01030.1	197	expressed protein predicted proteins, Arabidopsis thaliana
MGNSAGSLFRDKETQKRAGEIHAKE	4.7	At5g01030.2	197	expressed protein predicted proteins, Arabidopsis thaliana
MGCCGSKMGKRGFSDRMVSLHNLVS	4.7	At1g03590.1	8	protein phosphatase 2C (PP2C) similar to GB:AAB97706
MGSRLGKPPSSMESSPPFIWSGLSS	4.6	At5g16280.1	1180	expressed protein KIAA1012 protein - Homo sapiens, EMBL:AB023229
MGCGFSGQIPESIGSLEQLVTLSLN	4.6	At5g49780.1	198	leucine-rich repeat transmembrane protein kinase, putative
MGSSSRTVSGSEIVRKAKQPPVLHG	4.6	At3g09600.1	188	myb family transcription factor contains Pfam profile: PF00249 myb-like DNA-bind
MGSRNSKALSYGFEVGDMVWGKVKS	4.6	At3g05430.1	121	PWWP domain protein contains Pfam profile:PF00855 PWWP domain
MGLSFSVPEGEKLPSSLLNIFKELH	4.6	At3g18630.1	182	uracil-DNA glycosylase -related similar to uracil-DNA glycosylase GB:P39615 from

Sequence (predicted myrist - internal)	**HMM$_{266}$ Score**	**ID**	**Start pos**	**Description**
MGSGTSVVLDSENGVEVEVEAESRK	4.5	At3g25730.1	32	AP2 domain transcription factor, putative contains Pfam profile: PF00847 AP2 dom
MGCSASLISVDIVKNIFKSYANKLA	4.5	At2g28630.1	166	beta-ketoacyl-CoA synthase family
MGNNLGKKESDREEKGSSIFDDLIP	4.4	At4g12770.1	169	auxilin-related low similarity to SP|Q27974 Auxilin {Bos taurus}
MGSKDSSEFFHDPSMKSSNAGQVRK	4.4	At1g71260.1	126	expressed protein
MGASNVRSKSRKLVSPEQVGNNRSK	4.4	At5g11530.1	161	expressed protein
MGAVISTQEAGGSSTENNGASSSSY	4.4	At1g07360.1	360	RRM-containing RNA-binding protein, putative
MGGQTSSSSSQGFVARNEKDCLDRS	4.3	At5g20200.1	264	expressed protein predicted proteins - Arabidopsis thaliana
MGSTITRNKGKFSRSSQKKKTQKPK	4.3	At3g14980.1	558	PHD finger transcription factor, putative contains Pfam profile: PF00628 PHD-fin
MGSSRGSPPRSNIEDFHGRSRMLDN	4.2	At1g70180.1	315	expressed protein contains similarity to arginine/serine-rich protein GI:6601502
MGSSRGSPPRSNIEDFHGRSRMLDN	4.2	At1g70180.2	315	expressed protein contains similarity to arginine/serine-rich protein GI:6601502
MGSSATTSIKEDVVSQVLRRDKPGR	4.1	At3g42230.1	167	hypothetical protein various predicted transposase like proteins
MGLVFSSNQDAATATATAAFSTSEA	4.1	At4g16110.1	538	response regulator protein family similar to ARR2 protein GI:4210451 from [Arabi
MGSGGSSVVLDSENGVETESRKLPS	4.0	At1g25560.1	45	AP2 domain transcription factor, putative similar to DNA-binding protein RAV2 GI
MGSRGSRSRAIDGEDEYYSLILNFK	4.0	At4g39960.1	188	DnaJ protein, putative similar to SP|Q9S5A3 Chaperone protein dnaJ {Listeria mon
MGSKISTGCKSLSGTSVASPVVAGI	4.0	At5g19660.1	394	subtilisin-related protein subtilisin/kexin isozyme SKI-1 precursor - Rattus nor
MGLGFSSNCSSDWNQTILQEDLNSS	3.9	At1g61660.1	91	expressed protein
MGLGFSSNCSSDWNQTILQEDLNSS	3.9	At1g61660.2	91	expressed protein
MGNSSSSTWKEIGFWPTHRFKESFG	3.9	At4g10210.1	277	hypothetical protein IB1C3-1 protein, Arabidopsis thaliana, AJ011845 contains Pf
MGAIRSRSRTSLSSTTSLPLHQTRP	3.8	At3g47820.1	135	armadillo repeat containing protein
MGGAFARGYVQQAVSDTDSRRCHPL	3.8	At1g01440.1	250	hypothetical protein similar to hypothetical protein GB:CAB80918 GI:7267606 from
MGALMASSVPARGRRRESNNKKKDV	3.8	At4g18770.1	342	myb family transcription factor identical to transcription factor (MYB98) GI:153
MGNLFCSEYRIHKRFDLKGSSHGRT	3.8	At2g26420.1	493	phosphatidylinositol-4-phosphate 5-kinase -related
MGNTVVRARREARFYKRIAIDYVYA	3.8	At1g31120.1	743	potassium transporter, putative (KUP10/HAK10/KT10/POT10) similar to potassium tr
MGISIAKHRLEILKLARRDRKNSPP	3.7	At1g15760.1	47	hypothetical protein preedicted by genemark.hmm

Sequence (predicted myrist - internal)	HMM_{266} Score	ID	Start pos	Description
MGSFLAGSTEAPGGYEYTNGKRIKK	3.7	At1g79470.1	377	inosine-5'-monophosphate dehydrogenase identical to inosine-5'-monophosphate deh
MGISFAKDTDCTVRAYSDSDYGGCK	3.7	At1g34090.1	477	polyprotein -related similar to polyprotein GI:4996361 from [Arabidopsis thalian
MGAVSSSKESGSSTSDNRGASSSSY	3.7	At2g29580.1	366	RRM-containing RNA-binding protein, putative
MGQSSSNEPEVEKKREPEPEPEPEV	3.7	At1g12270.1	207	stress inducible protein (sti), putative similar to sti (stress inducible protei
MGSSSASHSSTASLEAGVASLEIPG	3.7	At1g04140.1	466	transducin / WD-40 repeat protein family contains 4 WD-40 repeats (PF00400); sim
MGSSSASHSSTASLEAGVASLEIPG	3.7	At1g04140.2	466	transducin / WD-40 repeat protein family contains 4 WD-40 repeats (PF00400); sim
MGSYFRKMSKPEWVNALQRLRTSKL	3.6	At3g04210.1	444	disease resistance protein (TIR-NBS class), putative domain signature TIR-NBS ex
MGGVISTIGSGESESVVDRGVEVAV	3.6	At3g30170.1	6	hypothetical protein similar to mutator-like transposase GB:AAD19776 GI:4388821
MGLACASPIPQRRPNIKEALQVLER	3.6	At5g67280.1	719	leucine-rich repeat transmembrane protein kinase, putative
MGNLFCSEYRIQRRFDLKGSSHGRY	3.6	At1g77740.1	521	phosphatidylinositol-4-phosphate-5-kinase -related similar to phosphatidylinosit
MGNTVVRARREARFYKKIAIDYVYA	3.6	At2g35060.1	739	potassium transporter, putative (KUP11/HAK11/KT11/POT11) similar to potassium tr
MGGGDSLEAGTSTDQGAFGSLGYCT	3.5	At4g00840.1	25	DHHC-type zinc finger domain-containing protein contains Pfam profile PF01529: D
MGNFSSTLQALHLRKNHLSGVFPEN	3.5	At2g15040.1	659	disease resistance protein -related
MGLSATKNAEIPVGNSPEESRNLLN	3.5	At1g65070.1	71	DNA mismatch repair MutS family contains Pfam profile PF00488: MutS domain V
MGVRSSTGSLRSRHVSRDSGDYLID	3.5	At5g15680.1	1491	expressed protein
MGGAGGTVTATSPGTSSAENNTWSS	3.4	At1g49720.1	267	abscisic acid responsive elements-binding factor identical to abscisic acid resp
MGNGHVSGPSASVFFQPKFGVSSPS	3.4	At1g01500.1	163	expressed protein
MGNINSNSFQHPCVSKAQHYSAVEG	3.4	At5g40120.1	281	MADS-box protein predicted proteins, Arabidopsis thaliana
MGNLFCTGHSIHRRFDLKGSSHGRL	3.4	At3g07960.1	496	phosphatidylinositol-4-phosphate 5-kinase -related similar to phosphatidylinosit
MGNLFCSEYSIHRRFDLKGSSLGRT	3.4	At3g56960.1	557	phosphatidylinositol-4-phosphate 5-kinase-related protein phosphatidylinositol-4
MGGGSSHGGGQQYVPFATSSGSLRV	3.4	At4g11850.1	13	phospholipase D-gamma, putative similar to phospholipase D-gamma-2 GI:5052357 fr

Sequence (predicted myrist - internal)	HMM$_{266}$ Score	ID	Start pos	Description
MGNASAGERARNTGVQANELVESAL	3.3	At3g06810.1	342	acetyl-coA dehydrogenase -related similar to acetyl-coenzyme A dehydrogenase GB:
MGSSVCTSLVEVGELGRLSASIKKL	3.3	At1g01820.1	131	expressed protein
MGQLFSASRQTLIVRAQDVEQIPDI	3.3	At4g11130.1	518	RNA-directed RNA polymerase -related RNA-directed RNA polymerase - Nicotiana tab
MGAAAVLMSSNDHDRDNAKYELLHV	3.2	At5g49070.1	238	beta-ketoacyl-CoA synthase family similar to very-long-chain fatty acid condensi
MGCSVTGVKPASGDSPLQLGPKEED	3.2	At3g63410.1	269	chloroplast inner envelope membrane protein, putative similar to SP|P23525 37 kD
MGAGVSGEPTAEENVQVASSMSNQE	3.2	At3g17440.1	116	expressed protein contains Pfam profile: PF00190 11S plant seed storage protein
MGATSSTSPTTTITTDQSAYLQSFA	3.2	At1g03840.1	387	expressed protein contains Zinc finger,C2H2 type,domain
MGCNYSKKGSEKLSCNMVSPCPGSF	3.2	At3g54530.1	23	hypothetical protein
MGVVGSSLVKEFLDDGNFVIKNTWD	3.2	At2g01780.1	85	S-locus glycoprotein, putative similar to S-locus glycoprotein from [Brassica ol
MGSGGSSVVLDPENGLETESRKLPS	3.1	At1g68840.1	38	AP2 domain protein RAP2.8 (RAV2) identical to RAV2 GI:3868859 from [Arabidopsis
MGSGTGLGRSGATAMRPPPNPMTGS	3.1	At2g43160.1	774	clathrin binding protein (epsin) -related
MGSGTGLGRSGATAMRPPPNPMTGS	3.1	At2g43160.2	774	clathrin binding protein (epsin) -related
MGLVFGREYELESNNNLESECLKGL	3.1	At1g01190.1	223	cytochrome P450, putative similar to cytochrome P450 SP:O48927 from [Glycine max
MGNVHSGKRAMVISSLKKANISASR	3.1	At4g31115.1	40	expressed protein supported by full length cDNA gi:21436034 from [Arabidopsis th
MGSSLRRHATLLWPDDHAEFCNVME	3.1	At4g21690.1	135	gibberellin 3 beta-hydroxylase family similar to gibberellin 3 beta-hydroxylase
MGQTSVCEWRQFLNANQSSIAGEKV	3.1	At5g59660.1	779	leucine rich repeat protein kinase, putative similar to light repressible recept
MGSSSRRDSETESSSSLESRNPPRR	3.0	At3g55000.1	150	expressed protein
MGLLVGDSVKEAYARAKSAKPPTNP	3.0	At5g11000.1	144	expressed protein predicted proteins, Arabidopsis thaliana
MGASSSSHRISTRHRWTPTSTQLQI	3.0	At1g20700.1	81	homeodomain protein PALE-2, putative strong similarity to homeodomain protein PA
MGRRSSKIAGRKGAQDSKKAKLYCR	2.9	At2g25830.1	84	expressed protein
MGFLRSKSNGGSTSRSRSRSRGRSR	2.9	At5g03110.1	22	expressed protein various predicted proteins, Arabidopsis thaliana
MGASHGWGFFCDRTDRSVRISDIFN	2.9	At1g05550.1	196	hypothetical protein
MGNACSLFSAMPLKSPASWNILIGG	2.9	At4g22760.1	486	selenium-binding protein-related temporary automated functional assignment

Sequence (predicted myrist - internal)	HMM_{266} Score	ID	Start pos	Description
MGSSSSSTLTEHNSLTGAKSGLGPV	2.9	At3g05710.1	24	syntaxin of plants SYP43 similar to putative syntaxin protein (GB:CAB52175) [Ara
MGSSSSSTLTEHNSLTGAKSGLGPV	2.9	At3g05710.2	24	syntaxin of plants SYP43 similar to putative syntaxin protein (GB:CAB52175) [Ara
MGLCSPSPSSHDCSSSSPRNSEKKQ	2.8	At1g21060.1	2	expressed protein
MGAVVGGPNDNDVYGDERTDYQHAE	2.8	At2g44560.1	450	glycosyl hydrolase family 9
MGNSGSGGPPPPPPPMPLANGATPP	2.8	At1g61080.1	582	proline-rich protein family
MGAGLGKECALLTDGRFSGGSHGFV	2.7	At3g23940.1	506	dihydroxyacid dehydratase -related similar to dihydroxyacid dehydratase GB:CAA60
MGSHSSSQLSSELVEEGSERAIVSP	2.7	At4g25540.1	489	DNA mismatch repair protein MSH3 identical to SP|O65607 DNA mismatch repair prot
MGSGVSSSTAMAGSSSASASASSSS	2.7	At5g43410.1	77	ethylene response factor, putative contains AP2 DNA-binding domain
MGFCFGGGRVVDVLATDESGYFSTG	2.7	At1g35420.1	198	expressed protein
MGNTVVKARTGSWLPKKIAIDYVYA	2.7	At4g19960.1	789	potassium transporter, putative (KUP9/HAK9/KT9/POT9) similar to potassium transp
MGCSASVISIDIVKNIFKTYKNKLA	2.6	At1g07720.1	166	beta-ketoacyl-CoA synthase family similar to GB:AAC99312 from [Arabidopsis thali
MGHALSLAKDELYDCHELAKKFRAI	2.6	At4g38270.1	278	glycosyltransferase family 8 contains Pfam profile: PF01501 glycosyl transferase
MGQYFGRSSFQPQLIGNDNPDGVAG	2.6	At5g13250.1	234	hypothetical protein
MGLSSATGKGKASQNESTTSCMFFH	2.6	At3g42100.1	126	hypothetical protein very low similarity to SP|Q9UUA2 DNA repair and recombinati
MGCYVVGSAGSNEKVDLLKNKFGFD	2.6	At3g03080.1	184	NADP-dependent oxidoreductase (P2), putative similar to probable NADP-dependent
MGAGRGRGRGGGADGGAPGKQPSGA	2.6	At1g50300.1	173	RRM-containing RNA-binding protein, putative contains similarity to RNA-binding
MGNSSSSSSQNISTLYSLLHGSSSN	2.6	At3g02150.2	287	TCP family transcription factor TFPD identical to GP:6681577 TFPD {Arabidopsis t
MGCSAGLISVDLARDLLQVHPNSNA	2.6	At1g68530.1	223	very-long-chain fatty acid condensing enzyme (CUT1) identical to very-long-chain
MGCSAGLISVDLARDLLQVHPNSNA	2.6	At1g68530.2	223	very-long-chain fatty acid condensing enzyme (CUT1) identical to very-long-chain
MGGAAVLLSNRSSDRSRSKYQLIHT	2.5	At5g43760.1	298	beta-ketoacyl-CoA synthase, putative similar to beta-ketoacyl-CoA synthase [Simm
MGCLTSTKRLALCLQPPLDSCPIGV	2.5	At5g56440.1	266	F-box protein family contains F-box domain Pfam:PF00646

Sequence (predicted myrist - internal)	**HMM$_{266}$ Score**	**ID**	**Start pos**	**Description**
MGSSFSFGFGAGSGQTTKTRHKPRL	2.4	At2g41520.1	41	DnaJ domain-containing protein contains Pfam profiles PF00226: DnaJ domain, PF00
MGSNSSSSCLTGEAKKTNSQIIKKG	2.4	At1g75770.1	89	expressed protein
MGNSTSLRGSTSHSPRDTPVFINSG	2.4	At4g13800.1	299	expressed protein
MGSCVVILSDGEVPVKADASTIAIS	2.4	At2g22400.1	733	expressed protein
MGCLWSKLVKLKTSKEFKVVSDSYK	2.4	At1g42150.1	213	hypothetical protein similar to putative transposon protein GI:7267374 from [Ara
MGVFVSCSAGNGGPDPISLTNVSP	2.4	At5g51750.1	319	subtilisin-like serine protease similar to subtilisin-like protease GI:3687307 f
MGLSGSGSYYIHRGLSGSGPPTFHG	2.2	At2g45850.1	7	AT-hook DNA-binding protein -related
MGLSGSGSYYIHRGLSGSGPPTFHG	2.2	At2g45850.2	7	AT-hook DNA-binding protein -related
MGIGSTRNTQDQGSNSVSVSGASPG	2.2	At1g72390.1	81	expressed protein
MGISIAKHRLEILKLARRDRKPSPL	2.2	At1g80520.1	45	hypothetical protein
-GSSFSKEKGSVPNGTNPSTRRKLI	2.2	At4g19180.1	88	hypothetical protein apyrase (EC 3.6.1.5), potato, PIR2:JC4616
MGNINSNSFQNPCVSNTQHYSAVEE	2.2	At5g41200.1	282	MADS-box protein
MGNSASGRGSISMPTRDTPVFTNSG	2.2	At3g23870.1	299	purine permease - related low similarity to purine permease [Arabidopsis thalian
MGESSSKEPEINLNKNEGGLKKNAS	2.2	At5g60010.1	23	respiratory burst oxidase (NADPH oxidase) family similar to respiratory burst ox
MGNSSARGHQCVYEPRNPLDMSHRS	2.1	At1g68270.1	246	AMP-dependent synthetase and ligase family similar to AMP-binding protein GI:190
MGSTSSNNNNGSNTNNNNNASSILR	2.1	At2g02070.1	447	C2H2-type zinc finger protein -related
MGGLTASSPFSAGVLPETVSSRGST	2.1	At1g45207.1	292	conserved hypothetical protein contains similarity to hypothetical proteins from
MGCYVVGSAGSTEKVDLLKTKFGFD	2.1	At5g17000.1	179	NADP-dependent oxidoreductase, putative strong similarity to probable NADP-depen
MGLCISWNGAATNGPIFAEIVPERA	2.1	At5g10190.1	332	transporter - related low similarity to spinster membrane proteins from [Drosoph
MGLLFSSLSDAQDRGGGRVGWLTE	2.0	At5g30341.1	23	hypothetical protein
MGCAGAAATDSPIQMNGLPISIYAP	2.0	At3g16150.1	169	L-asparaginase -related similar to L-ASPARAGINASE GB:P30364from [Lupinus angusti
MGNLGSLGLDSLSSLNTLRAEEKRD	1.9	At5g44800.1	1882	chromodomain-helicase-DNA-binding (CHD) protein family similar to chromatin remo
MGNNSAKLGNGRTPPHCWTFFSTAA	1.9	At3g04650.1	314	expressed protein
MGRSITNSNRSSLDLGRPSSSSSSS	1.9	At4g19490.1	8	expressed protein probable membrane protein YDR027c, Saccharomyces cerevisiae, P

Sequence (predicted myrist - internal)	**HMM_{266} Score**	**ID**	**Start pos**	**Description**
MGRSITNSNRSSLDLGRPSSSSSSS	1.9	At4g19490.2	8	expressed protein probable membrane protein YDR027c, Saccharomyces cerevisiae, P
-GCSFSSGSEHTGDGIEDYEDDDSP	1.9	At3g07350.1	23	expressed protein similar to hypothetical protein GB:AAC17612 [Arabidopsis thali
MGNLQSQSGKTGSGIVKITPQYPLY	1.9	At2g23520.1	366	hypothetical protein
MGCVTSVTYSVLINGQHFGHITPER	1.9	At2g16680.1	559	reverse transcriptase (RNA-dependent DNA polymerase), putative similar to revers
MGLNHGVGLNNNNNNGGFNGISTGG	1.8	At5g62940.1	243	Dof zinc finger protein Dof zinc finger protein, Oryza sativa, EMBL:AB028129
MGSAFAAELGTMQVSEQTDTLRVLG	1.8	At1g19800.1	194	expressed protein
MGAATSLYSATCFALGKYGNGNPYP	1.8	At5g20060.1	128	expressed protein acyl-protein thioesterase, Homo sapiens, EMBL:AF098668
MGKCSASKSKEVVEKECLGRNYCSI	1.8	At4g38590.1	718	glycosyl hydrolase family 35 (beta-galactosidase) similar to beta-galactosidase
MGIVVSELLSQLQDWNNRKQFTSSS	1.8	At3g17205.1	425	hypothetical protein
MGTCPSVLSNEAISTEQASHDSIEL	1.8	At3g42420.1	298	hypothetical protein various predicted proteins, including predicted Helicases.
MGCYVVGSAGSKEKVDLLKNKFGFD	1.8	At5g16960.1	180	NADP-dependent oxidoreductase, putative similar to probable NADP-dependent oxido
MGSGESGSFSPSPPQKQLKIENDML	1.8	At3g01410.1	112	RNase H domain-containing protein low similarity to GAG-POL precursor [Oryza sat
MGSLTRASPDSQPVAREGRISSSLQ	1.8	At5g01770.1	756	transducin / WD-40 repeat protein family similar to WD-repeat protein mip1 (SP:P
MGTFSVTGSSGGIVNHQDGASSEWF	1.7	At2g07671.1	90	
MGSSLRRKKEDDWESILHRQENSLD	1.7	At1g63730.1	390	disease resistance protein (TIR-NBS-LRR class), putative domain signature TIR-NB
MGGAFARGYVKKAVSDTGVRRCHPL	1.7	At4g01090.1	289	expressed protein
MGTASSGSEISSNKGSTRKGLRGGS	1.7	At4g00450.1	1822	expressed protein
MGCSAGLISIDLANNLLKANPNSYA	1.7	At1g01120.1	256	fatty acid elongase 3-ketoacyl-CoA synthase 1 (KCS1) identical to GB:AAC99312 GI
MGMGYSGGSIPPQEETNTNNVGFVS	1.7	At1g68210.1	491	hypothetical protein
MGGLISLQSLDMSSNSLSGPLPKSL	1.7	At4g20940.1	146	leucine rich repeat protein family contains leucine rich-repeat (LRR) domains Pf
MGMVSAKPNASDKVVPAAAEMSSSR	1.7	At4g37210.1	394	tetratricopeptide repeat (TPR)-containing protein low similarity to SP|Q02508 Pr

Sequence (predicted myrist - internal)	**HMM$_{266}$ Score**	**ID**	**Start pos**	**Description**
MGACDCIITKAGPGTIAESLIRSLP	1.6	At5g20410.1	348	1,2-diacylglycerol 3-beta-galactosyltransferase (UDP-galactose:diacylglycerol ga
MGCYVVGSAGSKEKVDLLKTKFGFD	1.6	At5g16990.1	177	allyl alcohol dehydrogenase, putative similar to allyl alcohol dehydrogenase fro
MGQACSIVLRGASHHVLDEAERSLH	1.6	At5g20890.1	362	chaperonin, putative similar to SWISS-PROT:P78371- T-complex protein 1, beta sub
MGNTRRSGGGGGRNNKNKKRSRKGS	1.6	At4g34910.1	569	DEAD/DEAH box helicase, putative
MGGALGNSRSSMFKRQTKGNQIVQQ	1.6	At1g54060.1	200	expressed protein
MGCSLAAVTEVSFNLGGLSGAMISN	1.6	At5g17630.1	238	glucose-6-phosphate/phosphate translocator, putative similar to glucose-6-phosph
MGSTQRVLSFSDEKKKKKRKVKGEF	1.6	At2g17680.1	68	hypothetical protein
MGCYVVGSAGSKEKVDLLKTKFGFD	1.6	At5g16970.1	179	NADP-dependent oxidoreductase (P1), putative identical to probable NADP-dependen
MGCYVVGSAGSKEKVDLLKTKFGFD	1.6	At5g16980.1	73	NADP-dependent oxidoreductase, putative strong similarity to probable NADP-depen
MGGTTRGSSSSSGDDSNVFQMIFGS	1.6	At3g54920.1	450	polysaccharide lyase family 1 (pectate lyase) similar to pectate lyase 2 GP:6606
MGRRYSGSFGDLFSWPDLGGIGSSG	1.5	At1g68440.1	154	expressed protein
MGFSTGKRKSRDEEEDRVSFFASEF	1.5	At4g33160.1	102	F-box protein family (FBX13) contains similarity to fimbriata GI:547307 from [An
MGGYMSSKKTMEINPENSIMDELRK	1.5	At5g56010.1	595	heat shock protein, putative strong similarity to SP|P55737 Heat shock protein 8
MGGSMRQNKRRRILEKPTRRELSSS	1.5	At5g55820.1	929	hypothetical protein
MGLGASVLTKTLTFAAAFSAFGDPI	1.5	At5g12860.1	145	oxoglutarate/malate translocator, putative similar to 2-oxoglutarate/malate tran
MGSVASLALAIVVKGKDSSKLWGLV	1.5	At4g18130.1	330	phytochrome E (PHYE) identical to SP|P42498 Phytochrome E {Arabidopsis thaliana}
MGMVFGKIAVETPKYTVTKSGDGYE	1.4	At2g37970.1	11	expressed protein
MGYFSSGGNGYHRNHQNKSQRINGK	1.4	At4g32620.1	1411	expressed protein predicted protein T10M13.8, Arabidopsis thaliana
MGLLVSRRQHAEHHRAPYNNNYCIV	1.4	At4g27030.1	249	hypothetical protein gene F19K23.12 of BAC F19K23 from Arabidopsis thaliana chro
MGGRGGREAIYSSSTSSAATSSSSV	1.4	At4g37750.1	30	ovule development protein aintegumenta (ANT) identical to ovule development prot
MGLSTRKLEDQTLPVIDSSQSKTSF	1.4	At2g28880.1	614	para-aminobenzoate (PABA) synthase, component I family similar to PABA synthase

Sequence (predicted myrist - internal)	HMM$_{266}$ Score	ID	Start pos	Description
-GLFSSKKGKSKKNNSNPPPPPPPP	1.4	At1g72790.1	378	proline-rich protein family contains proline-rich extensin domains, INTERPRO:IPR
MGTRVGRDGRRGGSRPHKRHRFTRT	1.4	At1g05910.1	265	tat-binding protein -related similar to GB:CAA56963 from [Saccharomyces cerevisi
MGNKSSQRSGSTRRSRAAEVHNLSE	1.3	At3g59060.1	241	bHLH protein family contains Pfam profile: PF00010 helix-loop-helix DNA-binding
MGNKSSQRSGSTRRSRAAEVHNLSE	1.3	At3g59060.2	241	bHLH protein family contains Pfam profile: PF00010 helix-loop-helix DNA-binding
MGTPFGGPSTSAQNPTGAPANKDRN	1.3	At3g20800.1	10	cell differentiation protein -related similar to cell differentiation protein, R
MGHFASNCEGKPKKRAGESDEKGDG	1.3	At1g75660.1	270	Dhp1-related protein similar to Dhp1 protein GB:BAA04601 GI:496393 from [Schizos
MGRLSSSMKKIEKGLKNGNKYQDED	1.3	At2g45740.1	146	expressed protein
MGRLSSSMKKIEKGLKNGNKYQDED	1.3	At2g45740.2	146	expressed protein
MGLNCRCLELEERVLKGEERYTHLE	1.3	At5g53220.1	6	hypothetical protein
MGAAVALHFATSCALNHYTINPRVV	1.3	At1g52440.1	75	hypothetical protein
MGLACVQLFKIDEAGELFEEARGIL	1.3	At4g10840.1	480	kinesin light chain - related low similarity to kinesin light chain [Plectonema
MGGYCSKGKINKALRLYHEMTGKGI	1.3	At5g59900.1	479	pentatricopeptide (PPR) repeat-containing protein contains Pfam profile PF01535:
MNCASSSPSQSSPEEESVSLEDSDV	1.3	At1g07430.1	67	protein phosphatase 2C (PP2C), putative similar to GB:CAB90633 from [Fagus sylva
MGRSFSEESGEAERVRGLVTELDGL	1.2	At3g20110.1	196	cytochrome P450 family similar to Cytochrome P450 93A1 (SP:Q42798) {Glycine max
MGSTRSNSSTAPSFFAGPTMTSSSA	1.2	At5g03150.1	381	expressed protein zinc finger protein, Arabidopsis thaliana, EMBL:AL132953
MGVGDSSGSPKPIRIVVVGEKGSGK	1.2	At3g05310.1	3	GTP-binding protein - related low similarity to rac 1 protein [Physcomitrella pa
MGNCASSPRYENFMTPTIPSSKRL	1.2	At5g06100.1	209	myb family transcription factor contains Pfam profile: PF00249 myb DNA-binding d
MGNCASSPRYENFMTPTIPSSKRL	1.2	At5g06100.2	209	myb family transcription factor contains Pfam profile: PF00249 myb DNA-binding d
MGRSFRAGWGPNGVLFHTGKPICSS	1.2	At1g80680.1	370	nucleoporin PRECOZ (PRE) annotation temporarily based on supporting cDNA gi|2265
MGLSDSDFETEDDESGDDDSEDTGE	1.2	At3g49140.1	932	pentatricopeptide (PPR) repeat-containing protein contains Pfam profile PF01535:

Sequence (predicted myrist - internal)	**HMM$_{266}$ Score**	**ID**	**Start pos**	**Description**
MGNLFCSKYSVHRRFDLKGSSLGRT	1.2	At2g41210.1	550	phosphatidylinositol-4-phosphate 5-kinase -related
MGGFLTNSYSPKKFRSRHKKLCFGP	1.2	At3g03600.1	103	ribosomal protein S2p family similar to putative ribosomal protein S2 GB:CAA7422
MGAYHGKFSFDAFSHKKAVLYRSLF	1.1	At1g44170.1	423	aldehyde dehydrogenase, putative (ALDH) similar to aldehyde dehydrogenase ALDH [
MGAYHGKFSFDAFSHKKAVLYRSLF	1.1	At1g44170.2	423	aldehyde dehydrogenase, putative (ALDH) similar to aldehyde dehydrogenase ALDH [
MGSGSTVPGDPAGARLKPGSGYGYG	1.1	At5g19620.1	674	expressed protein outer membrane protein (IAP75) - Pisum sativum, L36858
MGAAMSLFRRRFNRDPSKYPNQRIT	1.1	At5g28235.1	415	Ulp1 protease family contains Pfam profile PF02902: Ulp1 protease family, C-term
MGAAMSLFRRRFNRDPSKYPNQRIA	1.1	At5g45570.1	692	Ulp1 protease family contains Pfam profile PF02902: Ulp1 protease family, C-term
MGAAMSLFRRRFNRDPSKYPNQRIA	1.1	At4g08430.1	579	Ulp1 protease family similar to At5g45570, At5g28235; contains Pfam profile PF02
MGRSCSEENGEAERVRGLVTKTDAL	1.0	At2g27010.1	174	cytochrome P450 family similar to Cytochrome P450 93A1 (SP:Q42798) {Glycine max
MGLSLGDGRSEADFIKEIVKEVERV	1.0	At1g47370.1	150	disease resistance protein (TIR class), putative domain signature TIR exists, su
MGNGHKRGRSEDSSSNNVSIGVETD	1.0	At3g13030.1	19	expressed protein
MGLFSSSSGDLRYKFQVCHENQVKS	1.0	At2g24240.1	188	expressed protein
MGCKARHAFKISRRVFELIRSEGSL	1.0	At5g60760.1	48	expressed protein various predicted proteins, Arabidopsis thaliana
MGTIASGCPETFAEVEGSSGRGGGG	1.0	At1g03890.1	105	globulin (seed storage protein) family similar to Arabidopsis thaliana 12S seed
MGNTITGITLGLGPDGELKYPSHQH	1.0	At5g18670.1	239	glycosyl hydrolase family 14 (beta-amylase) almost identical to beta-amylase BMY
-GCFSSSPFHGTTSSEHPPHSTPTV	1.0	At5g51920.1	17	hypothetical protein strong similarity to unknown protein (pir‖T05129)
MGIGAAGSTRISNELGAGNPEVARL	1.0	At2g04090.1	308	MATE efflux protein family similar to ripening regulated protein DDTFR18 [Lycope
MGGVSAAVSKTAAAPIERVKLLIQN	1.0	At3g08580.1	86	mitochondrial ADP,ATP carrier protein 1 identical to SWISS-PROT:P31167 ADP,ATP c
MGGVSAAVSKTAAAPIERVKLLIQN	1.0	At3g08580.2	86	mitochondrial ADP,ATP carrier protein 1 identical to SWISS-PROT:P31167 ADP,ATP c
MGGVSAAVSKTAAAPIERVKLLIQN	1.0	At5g13490.1	90	mitochondrial ADP,ATP carrier protein 2 identical to SWISS-PROT:P40941 ADP,ATP c

Sequence (predicted myrist - internal)	**HMM$_{266}$ Score**	**ID**	**Start pos**	**Description**
MGGVSAAVSKTAAAPIERVKLLIQN	1.0	At4g28390.1	85	mitochondrial ADP,ATP carrier protein, putative similar to mitochondrial ADP,ATP
MGTFHGKPSGCLLFELSHMLRKNTN	0.9	At3g25100.1	216	Cdc45-related protein similar to Cdc45 GB:AAC67520 [Xenopus laevis] (EMBO J. 17,
MGGVHASMVGSLHPTPVATLPFLSP	0.9	At3g12150.1	214	expressed protein
-GGSVSKSRLSDASKGENQNARKSR	0.9	At3g11590.1	598	expressed protein
MGIAFVKFLYDKDNKVVVGDDHGSK	0.9	At1g60110.1	484	jacalin lectin family similar to myrosinase-binding protein homolog [Arabidopsis
MGTKVTVISTSERKRDEAVTRLGAD	0.9	At4g37980.1	202	mannitol dehydrogenase (ELI3-1), putative identical to GI:16267
MGSLFAVLFVRNIPSISSGAAFFGL	0.9	At1g77860.1	171	membrane protein, Rhomboid family contains PFAM domain PF01694, Rhomboid family
MGHSYVRAKSGSSVMKKIAINFGYD	0.9	At5g14880.1	728	potassium transporter, putative (KUP8/HAK8/KT8/POT8) similar to potassium transp
MGLPHSGSVSPLSPLGEACSRRDLT	0.9	At4g00710.1	324	protein kinase family low similarity to protein kinase [Arabidopsis thaliana] GI
MGINLSSNTDSVLRCYSDSDWAGCK	0.9	At4g10690.1	1261	retrotransposon like protein copia-like retrotransposon Hopscotch polyprotein -
MGAASSDGSTDNDHVELLSSLLQRS	0.9	At1g20150.1	37	subtilisin-like serine protease similar to subtilisin-type protease precursor GI
MGLAISWVEKAALPEERRQGILRRL	0.8	At3g10572.1	173	3-phosphoinositide-dependent protein kinase-1, putative annotation temporarily b
MGLIHVRRISANEANGNLTPSMEVK	0.8	At3g05850.1	596	expressed protein
MGLFSSSSGELRYKFQVSHDNQVKS	0.8	At4g30940.1	186	expressed protein predicted protein, Caenorhabditis elegans, PID:G1707206
MGVNLRSLKDGFIDKGNSRVHLMAN	0.8	At3g42330.1	78	hypothetical protein
MGSVTSNDPSALFRGLASSSNSSSV	0.7	At2g02080.1	388	C2H2-type zinc finger protein -related
MGSGTGLGRAGAGAMRPPTNSMVGS	0.7	At3g59290.1	888	epsin-related protein Af10-protein - Avena fatua, EMBL:U80041
MGMVGSGQKLPMPPTTGRWNRSDPN	0.7	At2g33420.1	987	expressed protein
MGTVFTRRFEADRVGSQVPVKLMHL	0.7	At5g58510.1	175	expressed protein KIAA0066, Homo sapiens, EMBL:HSORFKG1O
MGRASSSLSPGHRHFAGQANELITE	0.7	At5g47380.1	117	expressed protein similar to unknown protein (pir||T02421)
MGSASARSFRISATASLPVLHRYKA	0.7	At5g67470.1	858	formin homology 2 (FH2) domain-containing protein contains formin homology 2 dom
MGNKMSVLAGDFLLSRACGALAALK	0.7	At2g34630.1	99	geranyl diphosphate synthase (GPPS)(dimethylallyltransferase) identical to GI:11
MGISLVRIDYAPHGQNPPHTHPRGS	0.7	At5g39130.1	92	germin-like protein (AtGER2) identical to germin-like protein subfamily 1 member

Sequence (predicted myrist - internal)	HMM_{266} Score	ID	Start pos	Description
MGISLVRIDYAPHGQNPPHTHPRGS	0.7	At5g39190.1	92	germin-like protein (AtGER2) identical to germin-like protein subfamily 1 member
MGISLVRIDYAPHGQNPPHTHPRGS	0.7	At5g39160.1	92	germin-like protein (GLP2a) (GLP5A) identical to germin-like protein subfamily 1
MGATVGGDGDTAVVAEQNRQLKGEI	0.7	At1g62990.1	9	homeodomain-containing protein HD1 -related similar to GB:Z29073 from [Brassica
MGIKGSIEQSTGDNGSEDKSIIHIE	0.7	At3g07530.1	637	hypothetical protein
MGIAFVKFLYDKDNKFVVGDDHGSK	0.7	At1g60130.1	486	jacalin lectin family similar to myrosinase-binding protein homolog [Arabidopsis
MGSVIGKSGKVINLIRQETRARIKV	0.7	At5g46190.1	55	KH domain protein strong similarity to unknown protein (pir||T04533)
MGSVNVNSASSASSCGSTSSSVMNK	0.7	At1g67890.1	446	protein kinase -related C-terminal region similar to MAP3K delta-1 protein kinas
MGSGYGVAAAAAYGGGKSHGNGNNG	0.7	At1g17640.1	323	RNA recognition motif (RRM) - containing protein contains InterPro entry IPR0005
MGFLSSSTTRKEAKTTRPNKAHPST	0.7	At2g41835.1	146	zinc finger protein (PMZ) -related temporary automated functional assignment
MGSNSTAELEAAGGDGGGNNGTHMP	0.6	At2g01420.1	422	auxin transport protein, putative similar to auxin transport protein PIN7[Arabid
MGSNSTAELEAAGGDGGGNNGTHMP	0.6	At2g01420.2	426	auxin transport protein, putative similar to auxin transport protein PIN7[Arabid
MGMNASLHTTMSDPSNLNHHSIGQA	0.6	At3g23280.2	275	auxin-regulated protein contains Pfam profile: PF00023 ankyrin repeat
MGMNASLHTTMSDPSNLNHHSIGQA	0.6	At3g23280.1	299	auxin-regulated protein contains Pfam profile: PF00023 ankyrin repeat
MGIVTSFAFPEASSSGDNSKVLPFI	0.6	At1g64170.1	183	cation/hydrogen exchanger, putative (CHX16) monovalent cation:proton antiporter
MGGGGGGGYNPFHGGGGGGQQYTFH	0.6	At5g03160.1	441	expressed protein P58 protein, Bos primigenius taurus, PIR:A56534
MGNVHGRVEPKNGSSQALLIGSHMD	0.6	At4g20070.1	145	hydantoin utilization protein-related contains similarity to hydantoin utilizati
MGISSSWSVADVLRSHVRRRHRTAT	0.6	At4g13470.1	87	hypothetical protein various predicted proteins, Arabidopsis thaliana
MGCSYRGLFRVGVVAVVCVMSVCSL	0.6	At5g24590.1	25	No apical meristem (NAM) protein family contains Pfam PF02365: No apical meriste
MGCSYRGLFRVGVVAVVCVMSVCSL	0.6	At5g24590.2	425	No apical meristem (NAM) protein family contains Pfam PF02365: No apical meriste
MGNSQICPTGTEKDCNGTQPKPMSI	0.6	At3g25560.1	204	somatic embryogenesis receptor-related kinase similar to somatic embryogenesis r