

Dataset S1: Recombinant fragments produced in this work from the selected gene products. In gray are highlighted the sequences amplified (from the corresponding genes) and produced (from the corresponding proteins).

spr0012. Cell division protein FtsH (EC:3.4.24.-)

aa sequence: 652

MKKQNNGLIKNPFLWLLFIFFLVTGFQYFYSGNNSGGSQQINYTELVQEITDGNVKELTY
QPNGSVIEVSGVYKNPKTSKEGTGIQFFTPSVTKVEKFTSTILPADTTVSELQKLATDHK
AEVTVKHESSSGIWINLLVSI VPF GILFFFLFSMMGMGGNGRNPMSFGRSKAKAANKE
DIKVRFSVDVAGAEEEKQELVEVVEFLKDPKRFTKLGARIPAGVLLLEGPPGTGKTLLAKAV
AGEAGVPPFFSISGSDVFVEMFVGVGASRVRS LFEDAKKAAPAIIFIDEIDAVGRQRGVGLG
GGNDEREQTLNQLLIEMDGFEGNEGIIVIAATNRSVDVLDPALLRPGRFDRKVLVGRPDVK
GREAILKVHAKNKPLAEDVDLKLVAQQT PGFV GADLENVLNEAALVAARRNKSIIDASDI
DEAEDRVIAGPSKKDKTVSQKERELVAYHEAGHTIVGLVLSNARVVHKVTIVPRGRAGGY
MIALPKEDQMLLSKEDMKEQLAGLMGGRVAEEIIFNVQTTGASNDFEQATQMARAMVTEY
GMSEKLGVPVQYEGNHAMLGAQSPQKSI SEQTAYEIDEEVRSLLNEARNKAAEIIQSNRET
HKLIAEALLKYETLDSTQIKALYETGKMPEAVEEESHALS YDEVKSKMNDK

Nt sequence: 1959

atgaaaaacaaaataatggtttaattaaaaatccttttctatggttattatcttt
ttccttgtgacaggattccagtatcttctatctggaataactcaggaggaagtcagcaa
atcaactatactgagttggtacaagaaattaccgatggtaatgtaaaagaattaacttac
caaccaaattggtagtgttatcgaagtttctggtgtctataaaaatcctaaaacaagtaaa
gaaggaacaggtattcagtttttcacgccatctgttactaaggtagagaaatctaccagc
actattcttctgcagatactaccgatcagaattgcaaaaacttgctactgaccataaa
gcagaagtaactgtaagcatgaaagttcaagtggtatatggattaatctactcgtatcc
attgtgccatttggaaattctattcttcttcttcttctctatgatgggaaatatgggagga
ggcaatggccgtaatccaatgagtttggacgtagtaaggctaaagcagcaataaagaa
gatattaaagtaagatcttcagatggtgctggagctgaggaagaaaaacaagaactagtt
gaagttggtgagttcttaaaagatccaaaacgattcacaaaacttgagcccgtattcca
gcaggtgttcttttgaggggacctccggggacaggtaagactttgcttgctaaggcagtc
gctggagaagcaggtgttccattctttagtatctcaggttctgactttgtagaaatgttt
gtcggagttggagctagtcgtgttcgctctctttttgaggatgccaaaaagcagcacca
gctatcatctttatcgatgaaattgatgctgttggacgtcaacgtggagtcggtctcggc
ggaggtaatgacgaacgtgaacaaacctgaaccaacttttgattgagatggatgggtttt
gagggaaatgaagggattatcgtcatcgtcgcacaaaccgttcagatgacttgatcct
gcccttttgcgtccaggacgttttgatagaaaagtattggttggccgtcctgatgttaaa
ggtcgtgaagcaatcttgaaagttcacgctaagaacaagccttttagcagaagatggtgat
ttgaaattagtggtcaacaaactccaggctttggttggtgctgatttagagaatgtcttg
aatgaagcagctttagttgctgctcgtcgcataaatcgataattgatgcttcagatatt
gatgaagcagaagatagagttattgctggaccttctaagaaagataagacagtttcacaa
aaagaacgagaattggttgcttaccatgaggcaggacataaccattggttggtctagtcctg
tcgaatgctcgcgttgccataaggttacaattgtaccacgcccgtgcaggcggatac
atgattgcacttctaaagaggatcaaatgcttctatctaaagaagatatgaaagagcaa

ttggctggcctaataatgggtggacgtgtagctgaagaaattatctttaatgtccaaactaca
 ggagcttcaaacgactttgaacaagcgacacaaatggcacgtgcaatggttacagagtac
 ggtatgagtgaaaaacttggcccagtacaatatgaaggaaaccatgctatgcttgggtgca
 cagagtcctcaaaaatcaatttcagaacaaacagcttatgaaattgatgaagaggttcgt
 tcattattaaatgaggcacgaaataaagctgctgaaattattcagtc aaatcgtgaaact
 cacaagttaattgcagaagcattattgaaatacgaaacattggatagtacacaaattaa
 gctctttacgaaacaggaaagatgcttgaagcagtagaagaggaatctcatgcactatcc
 tatgatgaagtaa **agtcaaaaatgaatgacgaaaaataa**

Protein	Forward (5'→3')	Reverse (5'→3')
spr0012	TAAT GAT ATC GAA GTT GTT GAG TTC TTA AAA GA	GCTGTA GCG GCC GCA TTT TTC GTC ATT CAT TTT TGA CT

1359 nt cloned

spr0561. Cell wall-associated serine proteinase precursor PrtA (EC:3.4.21.-)

aa sequence: 2144

MKKSTVLSLTAAVILAAAYAPNEVVLADTSSSEDALSISDKEKVVVDKETENKEKHKDIIH
 NAIETSKDTEEKKTTIIIEEKEVVSKNPVIDTKTSNEEAKIKEENSNQSQGDHTDSFVNKN
 TENPKKEDKVVIYIAEFKDKESGEKAIKGLSNLKNTKVLYTYDRIFNGSAIETTPDNLDKI
 KQIEGISSIERAQKVQPMNHRKEIGVEEAIDYLSKINAPFGKNFDGRGMVISNIDTGT
 DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKI PHAFNYNGGKITVEKYDDGRDY
 FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE
 DSIKHNVDVVS SVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA
 NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTFEFDKVNIGGESFKYRNIGAFFDKNKITT
 NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM
 VVNTVNYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE
 DFKDKLEQYYPIDMESFNPNVNGDEKEIDFKFAPDTEKELYKEDIIVPAGSTSWGPRI
 DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERP
 LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE
 VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD
 RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHGTFTIGANSSFDLNAVINVGEAK
 NKNKFVESFIHFESVEEMEALNSNGKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS
 RSKTLGGYDDD GKPKIPGTLNKGIGGEHGIDKFNPAQVIQNRKDKNNTSLDQNPPELFAFN
 NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE
 NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLVWGLKWDGLIYNPRGREENAPESKD
 NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK
 IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFEIANEVVYAGAALVNEDGEVEKNLEV
 TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIEVIALDGSSNFTKIHRIFADQADEK
 GMISYLVDPDKASKYEKLGEISEDKLNKAKSPEENTNNNQAKDEDSKPDEKSSVEGEA
 SLEINKTISTIREFENKDLKLIKFKFREVNDFTSETGKRIEYDYKYDDKGNIIAYDDG
 SALQYETEFKFEIKSKIYGVLSPSKDGHFELGKISNVSKNAKVYYGNSYKSIEIKATKY
 DSHSKTMIFDLYANINDIVIDGLAFAGDMRLFVKDDNQIKAETKIRMPEKNKETKAEYPYV
 SSYGNVIELGEGDLSKNKPDNLTKMESGKIYS DSEKQQYLLKDNIILRKGALKVTTYNP
 GKTDMLEGNVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGFN
 IIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLF GKDDKEYTGEDKSNVEAIKEDGSMFL
 IDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIYG
 DLHIDNTRDFNIKLVKDG DIMDWGMKDYKANGFPDKVTDMDGNVYLQGTGYSDLNAKAVG
 VHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNGFDGEIQEQHIYVNGKEY
 TSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEMS
 STIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEEK

KEEENKPTFDVSKKKDNPQVNHSQNLNESHKEDLQREDHSQKSDSTKDVATVLDKNNIS
SKSTTNNPNKLPKTGTASGAQTLAAGIMFIVGIFLGLKKNQD

Nt sequence: 6435

atgaaaaaaagcacagtattgtcactaactacagctgcagttatTTTTagcagcctatgcc
cctaatagaggtagtcttagca**gacacatctagctctgaagat**gctttaagcatctctgat
aaagaaaaagtagtagtagataaggaaacagaaaaataaagagaaacataaagatattcat
aatgctatagaaaacttcaaaggatactgaagaaaagaaaacaacaattattgaggaaaaa
gaagttgtagtaaaaaatcctgtaatagacactaaaactagcaatgaagaagcaaaaatc
aaagaagaaaattccaatcaatcccaaggagatcatacggactcatttTgtgaataaaaaac
acagaaaatcccaaaaaagaagataaaagtTgtctatattgctgaatttaagataaaagaa
tctggagaaaaagcaatcaagggactatcaaactctaagaatacaaaagTTTTatatact
tatgatagaattTTTaaacggtagtgccatagaacaactccagataactTggacaaaatt
aaacaaatagaaggattttcatcgattgaaagggcacaaaagtccaacccatgatgaat
catgccagaaaggaattggagttgaggaagctattgattacctaagTctatcaatgct
cggTTTTgggaaaaattTtgatggttagaggtatggTcatttcaaatatcgatactggaaca
gattatagacataaggctatgagaatcgatgatgatgccaagcctcaatgagattTaaa
aaagaagactTaaaaggaactgataaaaaattTggTtgagtgataaaaatccctcatgcg
ttcaattattataatggTggcaaaatcactgtagaaaaatTgatgatggaagggattat
TTTgaccacatgggatgcatattgcagggattctTgctggaaatgatactgacaagat
atcaaaaactTTaaacggcatagatggaattgcacctaatgcacaaattTctctTtaaaa
atgtattctgacgcaggatctgggtTgCGGTgatgaaacaatgtTtcatgctattgaa
gattctatcaaacacaacgtTgatgtTtTcggtatcatctggTTTTacaggaacaggt
ctTgtaggt**gagaaatattggcaagctattc**gggcattaagaaaagcaggcattccaatg
gtTgtcgctacgggtaactatgCGacttctgctTcaagTctTcatgggatttagtagca
aataatcatctgaaaatgaccgacactggaaatgtaacacgaactgcagcacatgaggat
gCGatagcggTcgctTctgctaaaaatcaaacagTtgagTtgataaagTtaacataggt
ggagaaagTTTTaaatacagaaatataggggcTTTTctgataagaataaaaatcacaaca
aatgaagatggaacaaaagctcctagTaaattaaaattTgtatataataggcaaggggcaa
gaccaagattTgataggtTtgatcttaggggcaaaattgcagTaatggatagaattTat
acaaggtTtaaaaaatgctTTTaaaaaagctatggataagggTgcacgcgccattatg
gtTgTaaatactgTaaattactacaatagagataattTggacagagctTccagctatggga
tatgaagcggatgaaggtactaaaagTcaagTgtTTTcaattTcaggagatgatggTgta
aagctatggaacatgattaatcctgataaaaaaactgaagTcaaaagaaataataaagaa
gattTTaaagataaattTggagcaatactatccaattgatatggaaagTTTTaattccaac
aaaccgaatgtaggtgacgaaaaagagattgactTTaagTTTgcacctgacacagacaaa
gaactctataaagaagatatcatcgTtccagcaggatccacatctTgggggccaagaata
gattTactTTTaaaaccgatgtTtccagcacctggTaaaaatattaaatctacgcttaat
gttattaatggcaaatcaacttatggctatTgtcaggaactagTatggcactccaatc
gtggcagctTctactgtTTTgattagaccgaaattaaaggaatgctTgaaagacctgta
ttgaaaaatctTaaaggagatgacaaaatagatctTacaagTctTcaaaaaattgCccta
caaaatactgctcgacctatgatggatgcaactTctTggaaagaaaaaagTcaatactTt
gcatcacctagacaacagggagcaggcctaattaatgtggccaatgctTtgagaaatgaa
gtTgtagcaactTTcaaaaactgattctaaaggtTtgTaaactcatatggTtccatt
tctctTaaagaaataaaaggtgataaaaaatactTTTacaatcaagctTcaatacatca
aacagacctTtgactTTTaaagTTTcagcatcagcgataactacagattctTtaactgac
agattaaaactTgatgaaacatataaagatgaaaaatctccagatggTaaagcaaatgtT
ccagaaattcaccagaaaaagTcaaaggagcaaatatcacattTgagcatggTactTtC
actataggcgcaaatTctagctTtgattTgaaTgCGgtataaaatgtTggagaggccaaa
aacaataaaatTgtagaatcattTattcattTtgagTcagTggaagaaatggaaagct
ctaaactccaacgggaagaaataaaactTccaactTctTgtcgatgctTcaatggga
ttTgctgggaattTggaaccacgaaccaatcctTgataaatgggctTgggaagaagggTca
agatcaaaaacactgggaggttatgatgatgatggTaaaccgaaaattccaggaacctta
aataagggaaatTgTgggagaacatggTatagataaaattTaatccagcaggagTtatacaa
aatagaaaagataaaaaatacaacatccctggatcaaaaatccagaattattTgctTtcaat
aaccaagggatcaacgctccatcatcaagTggtTctaagattgctaacattTatcctTta

gattcaa atggaa atcctca agatgctca acttgaa agaggatta acaccttctccactt
gtatta agaagtg cagaaga aggattg atttca atagtaa atacaa ataa agaggagaa
aatcaa agagact taaa agtcatttc gagagaa cacttt attag aggaattt taaattct
aaaag caatgat gcaa aggga atcaa atcatc taaactaaa agtttg ggggtg acttga ag
tgggat ggactcat ctata atcctag aggtag agaa gaaa atgc accagaa agta aggat
aatca agatcct gctacta agata agaggt caattt gaaccg attg cga aggtca atat
ttctata aattt aaatat agatta actaa agattac ccattggc aggtttc ctatattc ct
gtaaaa attgata acaccgccc taagattg tttcggttg attttt caaatcctg aaaaa
atta agttg attaca aaggata ctctat cata aggtaaa agatc agtata agaatgaa acg
ctattt gcgag agatcaaaa agaacatcctg aaaa atttg acgag attg cga acgaa gtt
tggatg ctggc gctctt gttaat gaagat ggag aggtg aaaaaa atctt gaagta
acttac gcaggtg aggggt caagga agaa atagaaa acttg acaa agacg gaa atacc att
tatgaa attaa aggtgc aggag attta agagg aaaa atcatt gaagtc attgc attagat
ggttct agca atttc acaa agattc ataga attaa atttg ctgatc aggtgat gaaa ag
gggatg atttctt attac ctatg atgcctg acaa agatgctt ctaa atatg aaaa actt
ggtgaa atttct gaagata aactcaaaa atgcaa agagccc agagg aaaa atacca ataat
aatca agcta aggat gaagatt caaaa ccagatg aaaaa agttc agttg agggg aga agct
agcctt gaaata aataaaa actattt caaca attag agagttt gaaaa taagac ctaa ag
aaactc attaaa agaa atttag agaa gttaat gatttt caagtg aaa actgg taagaga
atag agga atacg attata aatac gatgata aggg aaaa atattt gcttat gacgat ggt
agtgcctt caaatat gaaactg aaaa atttg acgaa ataaa atcaaaa atttat ggtggt
ctaagccc atcta agatgg acactttt gaaattctt ggaa agata agca atgtttctaaa
aatgcca aggtatattat ggaa atagctata aatcg atagaa atcaa agcg acta agtat
gattcac attcaaaa acgatgat atttg atttata cgc taatatta atgatattg tagat
ggattag cttttg ctgg agatag tagattattt gttaa agatg acaatca aataaaa agct
gaaacta agattaga atgcctg aaaaaa ataa aggaa actaa agcaga atatcc ctatgta
tcaagttat gggaa atgta atcga attagg agag aggtatctt caaaga acaaac cagac
aattta actaaa atgga atctggtaaa atctattctg attcag aaaa acaaca atatctg
ttaa aggata acatcatt ctaagaaa aggtatgc actaaa agtgactac ctata atcct
ggaaaa acggatag tttaga aggaa atggag tctatag caagga agatag caaaaa ta
caaa aggccaatc cta atcta agagtcctttc agaa acaaca atttatg ctgatag taga
aatgtt gaagatg gaaga agtacc caagctgt attaatgtc ggtttt ggacgg ctttaac
attata aggtatca agtgttac atttaaa atga acgata aagggg aagctatcg ataaa
gacggaa atcttgtg acagattcttctaaa acttg tattattt ggta aggatgata aagaa
tacttg gagagg ataa gttcca atgtaga agctataaaa agaatgg ctccatgttattt
attgata ccaa accagtaaacctt caatgg ataa gaactactt taatcc atctaa atct
aataaaa tttatg tacgaa atccaga attttattt aagagg taagatttctg ataa ggt
ggtttta actgg aattg agagttaat gaatc ggttg tagata attattt aatctac gga
gatttac acattgata acactag agattt taatatta agctga atgtta aagac ggtgac
atcatg gactgg gaa atgaa agactata aagcaaac ggtttcc agata aggtaac agat
atggatg gaa atg tttatctt caaactg gctatagc gatttga atgcta aagc agttgga
gtccactatc agttttt atagata atgttaaacct gaagtaaac attgatc ctaaggg a
aatactag tatcga atatgctgat gggaaa atctgtag tctt taacatca atgataaaa ga
aataatgg attcgatgg tgagatt caaga acaacatattt atgtaa atggaaa ga atat
acatcattt gatgat attaa caaata acagaca agacactaa acatta agattg ttgta
aaagatttt gcaagaa ataca accgtaaa agaatc attttta aataa agatac gggg agag
gta agtgaa ataaa cctcat agggta actgtg accattcaaaa tggaaa agaatg agt
tcaacgat agtgtc ggaaga agattttattttac ctgtttata aggggtga attag aaaaa
ggatacca atttgat ggttgg gaa atttctg gtttc gaaggta aaaaa agacgctgg ctat
gttatta atctat caaa agatac ttttataaaa cctgtatt caagaaa atag aggg agaaa
aagg agga agaga ataa accctactttt gatgtatc gaaaa agaa agata acccaca agta
aatcat agtca attaa atgaa agtcac agaaa agagg atttaca aagaga agatcattca
caaaa atctgatt caacta aggatgttac agctac agttcttg ataaa acaaatatc agt
agtaaat caactacta acaatc ctaata agttg ccaaaa actgga acagca agcgg agcc
cagacact attagctg ccgga ataatgtttat agtagga attttctt ggattgaa gaaa
aaaa atca agattaa

Protein	Forward (5'→3')	Reverse (5'→3')
spr0561	ATAT TCG CGA GAC ACA TCT AGC TCT GAA GAT	GATGAT GCG GCC GCG AAT AGC TTG CCA ATA TTT CTC

1090 nt cloned

spr0121. surface protein *pspA* precursor

aa sequence: 619

MNKKKMILTSLASVAILGAGFVASQPTVVRAEESPVASQSKAEKDYDAAKKDAKNAKKAV
EDAQKALDDAKAAQKKYDEDQKKTEEKAALKAASEEMDKAVAAVQQAYLAYQQATDKAA
KDAADKMIDEAKKREEEAKTKFNTVRAMVVPEPEQLAETKKKSEEAKQKAPELTKKLEEA
KAKLEEAEEKKATEAKQKVDAAEEVAPQAKIAELENQVHRLEQELKEIDESESEDYAKEGFR
APLQSKLDAKKAKLSKLEELSDKIDELDAEIAKLEDQLKAAEENNNVEDYFKEGLEKTIA
AKKAELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQPAPAPKPEKPAEQP
KPEKTDQQAEEEDYARRSEEEYNRLTQQQPKAIEKPAPAPKTGWKQENGMWYFYNTDGS
M
ATGWLQNNGSWYYLNSNGAMATGWLQYNGSWYYLNANGAMATGWAKVNGSWYYLNANGAM
ATGWLQYNGSWYYLNANGAMATGWAKVNGSWYYLNANGAMATGWLQYNGSWYYLNANGAM
ATGWAKVNGSWYYLNANGAMATGWVKDGDWYYLEASGAMKASQWFKVSDKWYYVNLGLG
LAVNTTVDGYKVNANGEWV

Nt sequence: 1860

atgaataagaaaaaatgattttaacaagtctagccagcgtcgctatcttaggggctggt
tttggtgcgtctcagcctactggtgtaagagcagaagaatctcccgtagccagtcagtct
aaagctgagaaagactatgatgcagcgaagaagatgctaagaatgcaaaaaagcagta
gaagatgctcaaaaggcttagatgatgcaaaagctgctcagaaaaatgatgacgaggat
cagaagaaaactgaggagaaagccgctagaaaaagcagcgtctgaagagatggataag
gca^{gtggcagcagttcaacaagcg}tatctagcctatcaacaagctacagacaaagccgca
aaagacgcagcagataagatgatagatgaagctaagaaacgcgaagaagaggcaaaaact
aaatthaataactggttcgagcaatggttagttcctgagccagagcagttggctgagactaag
aaaaatcagaagaagctaaacaaaaagcaccagaacttactaaaaaactagaagaagct
aaagcaaaattagaagaggctgagaaaaagctactgaagccaaacaaaaagtggatgct
gaagaagtcgctcctcaagctaaaatcgctgaattggaaaatcaagttcatagactagaa
caagagctcaaagagattgatgagctcgaatcagaagattatgctaaagaaggtttccgt
gctcctcttcaatctaaattggatgccaaaaagctaaactatcaaaacttgaagagtta
agtataagattgatgagtttagacgctgaaattgcaaaacttgaagatcaacttaaagct
gctgaagaaaacaataatgtagaagactactttaagaaggttttagagaaaactattgct
gctaaaaaagctgaattagaaaaaactgaagctgaccttaagaaagcagttaatgagcca
gaaaaaccagctccagctccagaaactccagccccagaagcaccagctgaacaacccaaa
ccagcgcggctcctcaaccagctcccgcacccaaaaccagagaagccagctgaacaacca
aaaccagaaaaaacagatgatcaacaagctgaagaagactatgctcgtagatcagaagaa
gaatataatcgcttgactcaacagcaaccgccaagctgaaaaaccagctcctgcacca
aaaacaggctggaacaagaaaacgggtatgtggtacttctacaatactgatggttcaatg
gcgacaggatggctccaaaacaacgggtcatggtactacctcaacagcaatgggtgctatg
gctacagggttggtccaatacaatgggtcatggtattacctcaacgctaacgggtgctatg
gcaacagggttgggctaaagtcaacgggtcatggtactacctcaacgctaactgggtgctatg
gctacagggttggctccaatacaacgggtcatggtattacctcaacgctaacgggctatg
gcaacagggttgggctaaagtcaacgggtcatggtactacctcaacgctaactgggtgctatg
gctacagggttgggctccaatacaacgggtcatggtactacctcaacgctaactgggtgctatg
gcaacagggttgggctaaagtcaacgggtcatggtactacctcaacgctaactgggtgctatg
gcaacagggttgggctaaagtcaacgggtcatggtactacctcaacgctaactgggtgctatg
aaagcaagccaatgggtcaaagtatcagataaatgggtactatgtcaatgggttaggtgcc
cctgtagtcaacacaactgtagatggctataaagtc^{aatgccaatggtgaatgggttaa}

Protein	Forward (5'→3')	Reverse (5'→3')
spr0121	TAGA GAT ATC GTG GCA GCA GTT CAA CAA GCG	GATACA GCG GCC GCA AAC CCA TTC ACC ATT CGC ATT G

1557 nt cloned

spr0328. Cell wall surface anchor family protein

aa sequence: 1767

MNKGLFEKRCKYSIRKFSLSGVASVMIGATFFGTSPVLADSVQSGSTANLPADLATALATA
 KENDGHDFEAPKVGEDQGSPEVTDGPKTEEELLALEKEKPAEEKPKEDKPAAAKPETPKT
 VTPewQTVekkeQQGTVTIREEKGVRYNQLSSTAQNNDNAGKPALFEKKGLTVDANGNATV
 DLTFKDDSEKGSRFGVFLFKDTKNNVFGYDKDGFWEYKSPTTSTWYRGSRVAAAPET
 GSTNRLSITLKS DGQLNASNNDVNLFDVTLPAAVNDHLKNEKKILLKAGSYDDERTVVS
 VKTDNQEGVKTEDTPAEKETGPEVDDSKVTYDTIQSKVLKAVIDQAFPRVKEYSLNGHTL
 PGQVQQFNQVFINNHRITPEVTYKKINETTAEYLMKLRDDAHLINAEMTVRLQVVDNQLH
 FDVTKIVNHNQVTPGQKIDDERKLLSSISFLGNALVSVSSDQTGAKFDGATMSNNTHVSG
 DDHIDVTNPMKDLAKGYMYGFVSTDKLAAGVWSNSQNSYGGGSNDWTRLTAYKETVGNAN
 YVGIHSSewQWEKAYKGI VFP EYTKELPSAKVVITEDANADKKVDWQDGAIA YRSIMN NP
 QGWKVKDITAYRIAMNFGSQAQNPFLMTLDGIKKINLHTDGLGQGVLLKGYGSEGHDSG
 HLN YADIGKRIGGVEDFKTLIEKAKKYGAHLGIHVNASETYPESKYFNEKILRKNPDGSY
 SYGWNWLDQGINIDAAYDLAHGRLARWEDLKKKLG DGLDFIYVDVWNGQSGDNGAWATH
 VLAKEINKQGWRF AIEWGHGGEYDSTFHHWAADLT YGGYTNKGINSAITR FIRNHQKDAW
 VGDYRSYGGAAANYPLLGGYSMKDFEGWQGRSDYNGYVTNLFAHDVMTKYFQHFTVSKWEN
 GTPVTMTDNGSTYKWTPEMRVELVDADNNKVVVTRKSN DVNSPQYRERTVTLNGRVIQDG
 SAYLTPWNWDANGKKLSTDK EKMYFNTQAGATTWTLPSDWAKSKVYLYKLTDQ GKTEEQ
 ELTVKDGKITL DLLANQP YVLYRSKQTNPEMSWSEG MHIYDQGFNSGTLKHW TISGDASK
 AEIVKSQGANM LRIQGNKEK VSLTQKLTGLKPN TKYAVYVGV DNRSNAKASITVNTG EK
 EVTTYTNKSLALNYVKAYAHNTRRN NATVDDTSYFQNM YAFFTTGSDVSNVTLTLSREAG
 DEATYFDEIRTFENSSMYGDKHDTGKGT FKFQDFENVAQGI FPFVVG GVEGVEDNRTHLS
 EKHPYTORGWNGK KVDVIEGNWSLKTNGLVSRRLVYQTI PQNFRFEAGKTYRVTFEY
 EAGSDNTYAFVVGKGEFQSGRRGTQASNLEMHEL PNTWTDSKKAKKATFLVTGAETGDTW
 VGIYSTGNASNTRGDSGGNANFRGYNDFMMDNLQIEEITLTGKMLTENALKNYLPTVAMT
 NYTKESMDALKEAVFNLSQADDDISVEEARAEIAKIEALKNALVQKKTALVADDFASLTA
 PAQAQEGLANAFDGNLSSLWHTSWGGGDV GKPATMVLKEATEITGLRYVPRGSGSNGNLR
 DVKLVVTDESGKEHTFTATDWPDNNKPKDIDFGKTIKAKKIVLTGT KTYGDGGDKYQSAA
 ELIFTRPQVAETPLDLSGYEAALAKAQKLT DKNQEEVASVQASMKYATDNHLLTERMVE
 YFADYLNQLKDSATKPDAPTVEKPEFKLSSVASDQ GKTPDYKQEIARPETPEQILPATGE
 SQFD TALFLASVSLALSALFVVKTKD

Nt sequence: 5304

atgaataaaggattatttgaaaaacggttgtaaataatagtagtattcggaaatcttcattaggt
 gttgcttctggtatgattggagctacattccttgggacaagtccgggtccttgcatagc
 gtgcagctctggttccacggcgaacttaccagctgatttagctactgctccttgcaacagca
 aaagagaatgatgggcatgattttgaagcgcctaaggtgggagaagaccaagggttctcca
 gaagttacagatggacctaagacagaagaagaactattagcacttgaaaaagaaaaaccg
 gctgaagaaaaacccaaaagaggataaacctgcagctgctaaacctgaaacacctaagacg
 gtaaccctgaaatggcaaacggtagagaaaaaagaacaacagggaaacagtcactatccga
 gaagaaaaagggtgtccgctacaaccaattatcctcaactgctcaaaatgataacgcaggt
 aaaccagccctgtttgaaaagaagggttgaccggttgatgccaatggaaatgcaactggt
 gatttaaccttcaaagatgattctgaaaagggcaaatcacgcttgggtgtcttcttgaaa
 tttaaagataccaagaataatgtttttgtcggttacgacaaggatggctggttctgggag
 tataaatctccaacaactagcacttggatagaggtagtcgtggtgctgctcctgaaaca
 ggatcaacaaaccgtctctctatcactctcaagtcagacggtcagctaaatgccagcaat

aacgatgtcaatctctttgacacagtgactctaccagctgcggtcaatgaccatcttaaa
aatgagaagaagattcttctcaaggcgggctcttatgacgatgagcgaacagttgttagc
gttaaaacggataaccaagaggggtaaaaaacagaggatacccctgctgaaaaagaaaca
ggctctgaagttgatgatagcaaggtgacttatgacacgattcagttcaaggttctcaa
gcagtgattgaccaagccttccctcgtgtcaaggaatacagcttgaatggacatactttg
ccaggacaggttcaaca**gttcaaccaagtctttatcaata**accaccgaatcaccctgaa
gtcacttataagaaaatcaatgagacaacagcagagtacttgatgaagcttcgogatgat
gctcacttaatcaatgcggaaatgacagtagccttgcaagttgtggacaatcaattgcac
tttgatgtgaccaagattgtcaaccacaatcaagtcactccaggtcaaaagattgatgac
gaaagaaaactactttcttctattagtttctcgcgcaatgctttagtctctgtttctagt
gatcaaaactggtgctaagttgatggggcaacatgtcaaacaatacgcagtcagcggga
gatgatcatalcagatgtaaccaatccaatgaaagatctagccaagggttacatgtatgga
ttgtttctacagataagccttgctgctgggtgttgaggtaactctcaaaacagctatggt
ggtggttcgaatgactggactcgtttgacagcctataaagaacagtcggaaatgccaac
tatgtaggaatccacagctctgaatggcaatgggaaaaagcttataagggcattgtttc
ccagaatacacgaaggaactccaagtgctaaggttggtatcactgaagatgccaatgca
gacaagaaagtcgattggcaggatgggtgccattgcttatcgtagcattatgaacaatcct
caaggttggaaaaaagtttaaggatatcacagcttaccgtatcgcgatgaactttggttct
caagcacaaaaccatttcttatgaccttggtatcaagaaaatcaatctccacaca
gatggtcttgggcaaggtgttctccttaaaggatatggtagcgaaggccatgactctggt
cacttgaactatgctgatattggttaagcgtatcgggtggtgctgaagacttcaagacccta
attgagaaggctaagaaaataggagctcatctaggtatccacgttaacgcttcagaaact
tatcctgagttcaataacttcaatgaaaaaatttctccgtaagaatccagatggaagctat
agctatgggttggaaactggctagatcaaggtatcaacattgatgctgcctatgacctagct
catgggtcgtttggcacggttgggaagatttgaagaaaaaacttgggtgacgggtctcgacttt
atctatgtggacggtttgggtaatgggtcaatcaggtgataacgggtgcctgggctaccac
gttcttgctaaagaaattaacaacaaggctggcgttttgcgatcagagtggggccaatggt
ggtgagtagcactctaccttccatcactgggcagctgacttgacctacgggtggctacacc
aataaaggtatcaacagtgccatcaccgctttatacgttaaccacaaaagatgcttgg
gtaggggactacagaagttatggtggtgacgccaactatccactgctaggtggctacagc
atgaaagactttgaaggctggcaaggaagaagtgactacaatggctatgtaactaactta
tttgcccatgacgtcatgaccaagtaacttccaacacttactgtaagtaaatgggaaaat
ggtacaccggtgactatgaccgataacggtagcacctataaatggactccagaaatgcga
gtggaattggttagatgctgacaataataaagtagttgtaactcgttaagtcaaatgatgct
aatagttccacaatatcgcgaacgtacagtaactctcaacggacgtgtcatccaagatggt
tcagcttacttgactccttggaaactgggatgcaaatggtaagaaactttctactgataag
gaaaagatgtactacttcaatacgcaggccggtgcaacaacttggacccttccaagcgat
tgggcaaagagcaaggtttacctttacaagtaactgaccaaggtaagacagaagagcaa
gaactaactgtaaaagatggtaaaattaccctagatcttctagcaaatcaaccatacgtt
ctctatcgttcgaaacaaaccaatcctgaaatgtcatggagtgaaggcatgcacatctat
gaccaaggatttaacagtggtaccttgaacattggaccatttcaggcgatgcttctaag
gcagaaattgtcaagttcaaggggcaacgatatgcttcgtattcaaggaacaaagaa
aaagttagttcactcagaaattaactggcttgaaccaaataccaagatgcccgtttat
gtcgggtgctgataaccgtagtaatgccaaaggcagcatcactgtaataactgggtgaaaa
gaagtgactacttataccaataagttctctcgcctcaactatgtaaaagcctatgcccac
aatacacgctcgtacaatgctacagttgacgatacaagttacttccaaaacatgtacgcc
ttctttacaactggatcggacgtatcaaatgttactctgacattgagtcgtgaagctggt
gatgaagcaacttactttgatgaaattcgtacctttgaaaacaattcaagcatgtacgga
gacaagcatgatacaggtaaaggcaccttcaagcaagactttgaaaatggtgctcagggt
atcttcccattttagtggtggtgctgaaggtgctgaagacaaccgactcacttgtct
gaaaaacacgatccatatacacaacgtgggtggaatggtaagaaagtcgatgatggtatc
gaaggaaattggtcactcaagacaaatggactagtgagccgtcgttaacttggtttaccaa
actattccgcaaaacttccgttttgaagcaggttaagacctaccgtgtaacctttgaatac
gaagcaggttcagacaatacctatgctttttagtgcgtaagggagaattccagtcaggt
cgtcgtgggtactcaagcaagcaacttggaaatgcatgaattgccaaatacttggacagat
tctaagaaagccaagaaggcaacttctcgtgacaggtgcagaaacaggggatacttgg

gtaggtatctactcaactggaatgcaagtaatactcgtggtgattctggtggaaatgcc
aacttccgtggttataacgacttcatgatggataatcttcaaatcgaagaattacccta
acaggaagatggtgacagaaaatgctctgaagaactacttgccaacggttgccatgact
aactacaccaagagtctatggatgctttgaaagaggcggctctttaacctcagtcaggcc
gatgatgatatcagtggtggaagaagcgcgtgcagagattgccaagattgaagccttgaag
aatgctttggttcagaagaaaacggctttggtagcagatgactttgcaagtcttacagct
cctgctcaggctcaagaaggtcttgcaaatgcctttgatggaaacttatctagtttatgg
catacatcatggggcgaggagatgtaggcaagcctgcaacctgggtcttgaaagaagca
actgaaatcactggacttcggttatggtccacgtggatcagggttcaaattggttaacttgcgt
gatgtgaaacttgttgacagatgagctggcaaggagcatacctttactgcaactgat
tggccagataacaataagccaaaagacattgattttggtaagacaattaaggctaagaaa
attgctcttacaggtactaagacttacggagatgggtggcgataaataccaatctgcagcg
gaactcatctttactcgtccacaggtagcagaaacacctcttgacttgtcaggctatgaa
gcagctttggctaaggctcagaaattaacagacaaagacaatcaagaggaagtagctagc
gttcaggcaagcatgaaatatgcgacggataacctctcttgacggaaagaatgggtggaa
tactttgcagattatctcaaccaattaaaagattctgctacgaaaccagatgctccaact
gtagagaaacctgagtttaacttagctctgtagcttccgatcaaggtaagacgccagat
tataagcaagaaatagctagaccagaaacacctgaacaaatcttgccagcaacagggtgag
agtcaatttgacacagccctcttcctagcaagtggttagcctagccctatctgctctctt
gtagtaaaaacgaagaaagactag

Protein	Forward (5'→3')	Reverse (5'→3')
spr0328	TTTA GAT ATC AGC GTG CAG TCT GGT TCC A	GATGAT GCG GCC GCT ATT GAT AAA GAC TTG GTT GAA C

1003 nt cloned

sp670_2141. TMP repeat family

aa sequence: 1217

MATLDELKVMIDAEIAPFRKKMKEVENQVKGTSQVKNATAKVREQSSSIGSAFGKLAKF
AGFAILGKLLLDVGMYSTQTALEVSAAMNQIKRQMGESSQSFLKWVNDNANAMNMGVGEA
TNYGAVYSNLFSGFIKDTNKLSAYTAKMLQTSAVVAEGSGRTITDVMERIRSGLLGNTEA
IEDLGINVNVAMIESTEAFKKFANGQSWQQLDYQTQQQIRLMAILEQATAKYGDTLNSNV
NGRISLFSKSLMKDAALNLGNSMLPIINAIMPVLSNFAMVLKNVTAKLAEFIALMFNKKAT
VKDGVGGAVGDMGNAMKDAAGGAGDLADAVDDAGDSAGGLADNLGDSAKNAKKAKELLG
LLGFDEINILQPKDDDDAGSGGGGKGGKGGGGGPFKDILPEVELTDMDNKFKSIFDG
LGDKLGLFDLFFKGFDAAFRPEGIKRIKTALDQIAKTMGEIATDPRVVFNAFNRMAEKIA
YALGQVTGSITITIGLGIGVFLAESIANGLGRQKERIIRALVALFDNVGNLSEAVGNIAQD
FSSAFYDVTITSTGAVRIGSAIVSTLLSLTSTIVEVGSKLAGSLFKGFQVVTTSAPKISS
VFQSLLDTVAPVFESIERSVNKFGDGLSRVYDEHVVPAINSIANAFNGLIDI IQILWENS
WQPFAEFLSGVFGVSIEGISDLLGGLLATLGLLADAIKLVADGFTVFSWCKENKEPIL
ALITTWQ TINFLSWAEQAGGLAGAFSLGSKVSLIVGGIKNLGLAIKALTFDKLVSFGET
IYLNTRYAKDFVNSGKTIAQLGKTALELGKSALAWTAHAAKMGLATAAEFAHSVAAGVA
TAATWAFNAALAVLTSPITWIIAIIAALIAIGVLLYQNWDTVVEFAKTAWQGLCDFISGI
CRAIGEFFSGLWTKLQEIFEPIGQWFGKFKQAWDAIVNIFSGIGEWFSGVFQAWDAIV
NIFTPIGSWFGQRWADVTSALANIGAWFTDIFQAWTGLTNIFSKLGLWFGGERWADVTSV
LANVSSWFGNMFSAVNAVKNAFSSIGGFFSGVWSTVQSIFVNAGQKVGSAVGGAFKSAV
NAVLGTIENVVNGFIGMINGVLGVVRNLPGLGWVGSVSTVSLPRLARGGIVDSPTIAMIG
EAGKEAVVPLENTGFIQTLGRVSSAVVNAMAGISPQGGFSSDGDIVIQIAGHEFGRVAI
QEINKEHERAGQTLTKI

Nt sequence: 3654

Atg**gcaactcttgatgaattgaaag**tcatgattgacgctgagatagcgcctttcaggaag
aagatgaaagaagtcgagaatcaggtcaaaggaacatctgaccaagtgaaaaatgctact
gccaaagttcgtgaacagtcgagctcaatcggtagtgcgtttgcaagctggctaagtct
gctggttttgcaatccttggttaagaaattacttgatggtgggatgtattcaacgcagacg
gctcttgaagtatcagcgtctatgaaccaaatacaagcgacagatgggcgagagttcgcaa
tctttcttaaaatgggttaacgataatgccaacgctatgaatatgggtgtgggtgaggct
accaactacggtgcagtcactcaaacttattttctggatttatcaaagataccaacaag
ctaagcgcctataccgctaagatggtgcagacatcggcagttggtgctgaaggttcaggg
cgcacgattacagacggttatggagcggattcgcctcaggttactagggaaacaccgaagcg
attgaggacctaggaatcaacgtcaacgtggctatgattgagtcactgaagcctttaag
aagttcgcaaacggacagagctggcaacagttggattaccaaacccagcaacaaatccgc
cttatggctattctggaacaggctacagccaagtatggggataccttgcctaattctgta
aatggctcgtatcagcctatttaagtcgctgatgaaggacgcagcattgaacctggtaac
tctatgttaccgattatcaatgccattatgcctgtcttgaactcttttgctatggctta
aagaacgttactgctaaactcgcctgagttatcgccttgatggttcaacaagaaagcaaca
gtgaaagatgggtgttggtggagcagttggagacatgggtaacgccatgaaggatgctgca
ggcggagcaggagaccttgctgacgcagtagacgacgctggagattcagcaggaggactt
gctgataatccttgagactcagccaaaaacgctaagaaagctgctaagagattgctaggt
cttttgggatttgatgagattaacatcttgcaaaaaccaaagacgcagatgcaggcgg
tctggaggaggtggcaaggtggttaaaggaaagggaggcgggtggcggacctttcaaagac
atcttgccagaagtcgagttgaccgacatggacaacaaatcaagagcatttttgatggt
cttgagataagctcaaaggtggtttgacctcttcaagaaaggttttgatgcagcattt
agaccagaaggtataaaacgcattaagactgccttagaccaaatagctaagacaatggga
gaaatcgccactgacccaaggttgtgaatgcctttaaccgaatggctgagaaaattgct
tatgcttttagggcaagtgcaggtcaataaccactatcgggctaggtatcgggtgttttc
cttgccgaaagtattgcaaatggccttggaaggcaaaaagaacgcattatcagggcgcta
gtcgcctttgtttgataatggtggtaacctttccgaggcagtaggaaacatagctcaggac
ttttctagtgccttctacgacgctcattacctcaactgggtgcgggttcgtagcggtagcgt
attgtgtcaactctggtgagttgacatctaccattggtgaagttggtagtaaatagca
ggaagttgtttaaggttttgaaaaagtcggtgtgacaagcgtcctaaaatttcatca
gtcttccaaagtttattagatactggtgcgctgtatgtgagagcattgaaaggtctggt
aacaatttggcgatggcttaagtcgtgtttatgatgaacatgtagtccttgcattaac
tctattgctaagtcttttaaggttaattgacattattcagattctctgggagaaattcc
tggcaacctttgctgagttttatcaggagatctcgggttagtattgaaggaatttca
gatttattaggaggtggccttttagccactttgggactattggcggatgctattaagtta
gtggcagatggtttcaccggtttttctgactggtgtaagaaaacaaagaacctatcttg
gctttgataacaacttgcaaacgattaatttcttatcatgggctgaacaagccggggga
cttgacaggagcattcagcttgtaggttagtaaggtctctttgattggtggagggattaag
aatctaggtcttgctattaaagcattgacatttgataagttggtcagttttgggtgaaaca
atctatttgaaacacttatatgcaaaagattttggtggtcaattcagggtaaaacaattgca
cagctaggaaaaactgctttagaacttggtaaatcagctctagcatggactgctcatgca
gcaaaaatgggattagcaaccgcgcggaatttgacacattctggtgcagcaggagtcgct
acagctgcaacatgggcttttaatgcagcgttagcagttttgacaagccaataacatgg
attattgcagcaatcgcagctttgattgctatcgggtgtcttgctctacaaaactgggac
actggtgttgagtttgctaaaactgcatggcaaggactatgtgattttattagtggtatt
tgctgagctattggcgaatttttcagtggtctatggacaaaactacaagaaatctttgag
ccgataggatcaatgggtttggcgagaagttccagcaagcatgggatgccattgtcaacatc
ttctctggatcggagagtggttctctggtgtattccaaggtgcatgggacgctatcggt
aatatcttcaactccaatcggctcatggtttggacaacgttgggcagatgtgactagtgc
ttggctaataatcggggcatggtttactgacatattccaaaagcatggactggtctaaca
aacatctttagcaaacctaggtttatggtttggcgagagatgggcagatgttacaagtggt
cttgcaaatgtatcttctgggtttgggaatatggttactagtgccttataatgcagtcag
aacgcgtttagttcaattgggtggcttcttcagcgggtgtatggtcaacggttcaaagcata
tttgcaatgctggacaaaaggttggttagcgtgtaggtggagcgtttaagagtgccggtc
aatgcgggttcttggaacgattgaaaatgtagtcaatggcttcatcggaatgattaatgga

gttttaggcggttgctcagaaacttacctggtctaggatggggttggtagtgtaagcacagtt
agcctccctcgtcttgcccgtggtggtatcgctcgatagccaacaatcgccatgattggt
gaagctggtaaagagggcggtcgtaccacttgaaaatacaggatttatccaaacacttga
cgagtagtcagcagtgcggtagtaaatacctggctggtattagtcacaaggtggtttt
tctagcgacggcgacatcggtatttcaaatacgcaggccatgagttcggacgggtagctatc
caagaaatcaacaaggaacatgaacgagcaggtcaaaccttgctcaagatttag

Protein	Forward (5'→3')	Reverse (5'→3')
sp670_2141	CCCG GAT ATC GCA ACT CTT GAT GAA TTG AAA G	AAAGCA GCG GCC GCG GTT AAT CTC ATC AAA TCC CAA AA

1100 nt cloned