

SEQUENCE LISTING.txt

```

<110> AB Enzymes Oy
<120> A novel fungal protease and use thereof
<130> A9273PC
<150> FI 20095497
<151> 2009-04-30
<160> 15
<170> PatentIn version 3.5

<210> 1
<211> 20
<212> PRT
<213> Fusarium equiseti

<220>
<221> MISC_FEATURE
<223> Sequence of an aminoterminal peptide #3792 from Fusarium
equiseti RF6318 protease.

<220>
<221> VARIANT
<222> (18)..(18)
<223> /replace= "Ser"

<220>
<221> VARIANT
<222> (18)..(18)
<223> /replace= "Thr"

<220>
<221> VARIANT
<222> (18)..(18)
<223> /replace= "Arg"

<400> 1
Ala Leu Thr Thr Gln Ser Asn Ala Pro Trp Gly Leu Ala Ala Ile Ser
1 5 10 15

Arg Cys Thr Pro
20

<210> 2
<211> 10
<212> PRT
<213> Fusarium equiseti

<220>
<221> MISC_FEATURE
<223> Sequence of a tryptic peptide 1246.673 from Fusarium equiseti
RF6318 protease.

<400> 2
Thr Val Ala Ala Ala Asp Ser Ser Trp Arg
1 5 10

```

SEQUENCE LISTING.txt

<210> 3
 <211> 7
 <212> PRT
 <213> Fusarium equiseti

<220>
 <221> MISC_FEATURE
 <223> Sequence of a tryptic peptide 3341.633 from Fusarium equiseti RF6318 protease.

<220>
 <221> misc_feature
 <222> (1)..(1)
 <223> Xaa can be any naturally occurring amino acid

<400> 3

Xaa Thr Tyr Gly Val Ala Lys
 1 5

<210> 4
 <211> 14
 <212> PRT
 <213> Fusarium equiseti

<220>
 <221> MISC_FEATURE
 <223> Sequence of a tryptic peptide 1503.799 from Fusarium equiseti RF6318 protease

<220>
 <221> VARIANT
 <222> (3)..(3)
 <223> /replace= "Ile"

<400> 4

Glu Ala Leu Thr Val Gly Ala Thr Thr Ser Ala Asp Ala Lys
 1 5 10

<210> 5
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> The sequence of the oligonucleotide primer PR087 derived from the aminoterminal peptide SEQ ID NO:1.

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> r is a or g

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (9)..(9)

SEQUENCE LISTING.txt

```

<223> y is t or c

<220>
<221> misc_feature
<222> (12)..(12)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 5
cartcnaayg cnccntgggg                                     20

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> The sequence of the oligonucleotide primer PRO88 derived from the
aminoterminal peptide SEQ ID NO:1.

<220>
<221> misc_feature
<222> (3)..(3)
<223> r is a or g

<220>
<221> misc_feature
<222> (6)..(6)
<223> y is t or c

<220>
<221> misc_feature
<222> (9)..(9)
<223> y is t or c

<220>
<221> misc_feature
<222> (12)..(12)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 6
caragyaayg cnccntgggg                                     20

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> The sequence of the oligonucleotide primer PRO89 derived from
peptide SEQ ID NO:4.

<220>
<221> misc_feature
<222> (3)..(3)

```

<223> r is a or g
 <220>
 <221> misc_feature
 <222> (6)..(6)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (18)..(18)
 <223> n is a, c, g, or t

<400> 7
 gcrtcngcng angtngtngc

20

<210> 8
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> The sequence of the oligonucleotide primer PRO90 derived from peptide SEQ ID NO:4.

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> r is a or g

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> r is a or g

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature

SEQUENCE LISTING.txt

<222> (18)..(18)
 <223> n is a, c, g, or t

<400> 8
 gcrtcngcrc tngtngtngc 20

<210> 9
 <211> 866
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> The sequence of the PCR fragment obtained using the primers PRO88 (SEQ ID NO:6) and PRO89 (SEQ ID NO:7) and *Fusarium equiseti* RF6318 genomic DNA as a template.

<400> 9
 aaagcaacgc accgtgggggt cttgctgcca tctcccgccg aacccccggt ggcagcacct 60
 acacctacga caccactgcc ggtgccggtta cttacggtta cgtcggttgac tctggtatca 120
 acaccgcca cactgacttt ggcggccgtg cttctctcggttacaacgct gctggtggcg 180
 cccacactga tacccttggc cacggtaccc acgttgctgg taccattgcc tccaacacct 240
 acggtgttgc caagcgtgta agtacaatca taccacacat gagctacaac atgatctgaa 300
 ctttatactt actattatta ggccaacgtc atctctgtca aggttttcgt cggtaaccaa 360
 gcttctacct ctgttatcct tgctggtttc aactgggctg tcaacgacat cacctccaga 420
 accgtgctag cccgctctgt catcaacatg tctctcggtg gtccctcttc tcagacctgg 480
 gctactgcca tcaacgctgc ctacagccaa ggtgtcctct ccggttgttg tgccggtaac 540
 ggtgattcca acggtcgtcc tctcccgcc tctggccagt ctctgcca cgttcccaac 600
 gctatcaccg ttgctgccgc cgactccagc tggcgaactg cctctttcac caactacggt 660
 cctgaggtcg atgtcttcgg tcctgggtgtc aacatccagt ccacctggta cacctccaac 720
 agcgctacca acaccatcag cggtacctcc atggcttgcc ctacggttg tggtcttgct 780
 ctctacctcc aggctctcga gaacctcaat acccctgctg ccgtcaccaa ccgcatcaag 840
 tctcttgcaa ctacctccgc tgacgc 866

<210> 10
 <211> 1303
 <212> DNA
 <213> *Fusarium equiseti*

<220>
 <221> misc_feature
 <223> The nucleotide sequence of the full-length *Fusarium equiseti* RF6318 protease gene (Fe prts8A) .

<400> 10
 atgactagct tccgccgtat cgctcttggc cttgcagctc tgctgccgc agtcctcgcc 60
 gctcccaccg agaagcgaca ggagctcact gccgcgcctg acaagtacat catcacctc 120
 aagcccgagg ctgctgaggc caaggtcgag gctcacatgg cctgggttac cgacgtccac 180
 cgccgcagcc tcggcaagcg tgacacttcc ggtgttgaga agaagttcaa catcagcagc 240

SEQUENCE LISTING.txt

```

tggaacgcct actctggcga gttcgacgat gctaccattg ctgagatcaa gaagagcccc      300
gaggttgcct tcgtcgagcc cgactacatt gtcaccctcg actacaaggt tgagcctctc      360
tctgaccgtg ctctgaccac tcagagcaac gtccttggg gtcttgctgc catctcccgc      420
cgaacccccg gtggcagcac ctacacctac gacaccactg ccggtgccgg tacttacggt      480
tacgtcgttg actctggtat caacaccgcc cacactgact ttggcgggcg tgcttctctc      540
ggttacaacg ctgctggtgg cgcccacact gatacccttg gccacggtac ccacgttgct      600
ggtaccattg cctccaacac ctacggtggt gccaaagcgtg taagtacaat cataccccac      660
atgagctaca acatgatctg aactttatac ttactattat taggccaacg tcattctctgt      720
caaggttttc gtcggtaacc aagcttctac ctctgttatc cttgctgggt tcaactgggc      780
tgtcaacgac atcacctcca agaaccgtgc tagccgctct gtcataca tgtctctcgg      840
tggtccctct tctcagacct gggctactgc catcaacgt gcctacagcc aaggtgtcct      900
ctccgttggt gctgccggta acggtgattc caacggtcgt cctctccccg cctctggcca      960
gtctcctgcc aacgttcca acgctatcac cgttgctgcc gccgactcca gctggcgaac     1020
tgcctctttc accaactacg gtcctgaggt cgatgtcttc ggtcctgggt tcaacatcca     1080
gtccacctgg tacacctcca acagcgctac caacaccatc agcgggtacct ccatggcttg     1140
ccctcacgtt gctggctctg ctctctacct ccaggctctc gagaacctca ataccctgc      1200
tgccgtcacc aaccgcatca agtctcttgc cactaccggc cgcatcactg gcagcctcag     1260
cggcagcccc aacgcatgg ctttcaacgg cgctactgct taa                          1303

```

<210> 11
 <211> 412
 <212> PRT
 <213> *Fusarium equiseti*

<220>
 <221> MISC_FEATURE
 <223> The deduced amino acid sequence of the full-length *Fusarium equiseti* RF6318 protease (Fe_RF6318) including amino acids from Met1 to Ala412.

<400> 11

Met Thr Ser Phe Arg Arg Ile Ala Leu Gly Leu Ala Ala Leu Leu Pro
 1 5 10 15

Ala Val Leu Ala Ala Pro Thr Glu Lys Arg Gln Glu Leu Thr Ala Ala
 20 25 30

Pro Asp Lys Tyr Ile Ile Thr Leu Lys Pro Glu Ala Ala Glu Ala Lys
 35 40 45

Val Glu Ala His Met Ala Trp Val Thr Asp Val His Arg Arg Ser Leu
 50 55 60

SEQUENCE LISTING.txt

Gly Lys Arg Asp Thr Ser Gly Val Glu Lys Lys Phe Asn Ile Ser Ser
65 70 75 80

Trp Asn Ala Tyr Ser Gly Glu Phe Asp Asp Ala Thr Ile Ala Glu Ile
85 90 95

Lys Lys Ser Pro Glu Val Ala Phe Val Glu Pro Asp Tyr Ile Val Thr
100 105 110

Leu Asp Tyr Lys Val Glu Pro Leu Ser Asp Arg Ala Leu Thr Thr Gln
115 120 125

Ser Asn Ala Pro Trp Gly Leu Ala Ala Ile Ser Arg Arg Thr Pro Gly
130 135 140

Gly Ser Thr Tyr Thr Tyr Asp Thr Thr Ala Gly Ala Gly Thr Tyr Gly
145 150 155 160

Tyr Val Val Asp Ser Gly Ile Asn Thr Ala His Thr Asp Phe Gly Gly
165 170 175

Arg Ala Ser Leu Gly Tyr Asn Ala Ala Gly Gly Ala His Thr Asp Thr
180 185 190

Leu Gly His Gly Thr His Val Ala Gly Thr Ile Ala Ser Asn Thr Tyr
195 200 205

Gly Val Ala Lys Arg Ala Asn Val Ile Ser Val Lys Val Phe Val Gly
210 215 220

Asn Gln Ala Ser Thr Ser Val Ile Leu Ala Gly Phe Asn Trp Ala Val
225 230 235 240

Asn Asp Ile Thr Ser Lys Asn Arg Ala Ser Arg Ser Val Ile Asn Met
245 250 255

Ser Leu Gly Gly Pro Ser Ser Gln Thr Trp Ala Thr Ala Ile Asn Ala
260 265 270

Ala Tyr Ser Gln Gly Val Leu Ser Val Val Ala Ala Gly Asn Gly Asp
275 280 285

Ser Asn Gly Arg Pro Leu Pro Ala Ser Gly Gln Ser Pro Ala Asn Val
290 295 300

Pro Asn Ala Ile Thr Val Ala Ala Ala Asp Ser Ser Trp Arg Thr Ala
305 310 315 320

Ser Phe Thr Asn Tyr Gly Pro Glu Val Asp Val Phe Gly Pro Gly Val
325 330 335

SEQUENCE LISTING.txt

Asn Ile Gln Ser Thr Trp Tyr Thr Ser Asn Ser Ala Thr Asn Thr Ile
340 345 350

Ser Gly Thr Ser Met Ala Cys Pro His Val Ala Gly Leu Ala Leu Tyr
355 360 365

Leu Gln Ala Leu Glu Asn Leu Asn Thr Pro Ala Ala Val Thr Asn Arg
370 375 380

Ile Lys Ser Leu Ala Thr Thr Gly Arg Ile Thr Gly Ser Leu Ser Gly
385 390 395 400

Ser Pro Asn Ala Met Ala Phe Asn Gly Ala Thr Ala
405 410

<210> 12
<211> 1243
<212> DNA
<213> Fusarium equiseti

<220>
<221> misc_feature
<223> The nucleotide sequence encoding the amino acid sequence of the
proenzyme form of Fusarium equiseti RF6318 protease.

<400> 12
gctcccaccg agaagcgaca ggagctcact gccgcgcctg acaagtacat catcacccctc 60
aagcccgagg ctgctgaggc caaggctcag gctcacatgg cctggggttac cgacgtccac 120
cgccgcagcc tcggcaagcg tgacacttcc ggtgttgaga agaagttcaa catcagcagc 180
tggaacgcct actctggcga gttcgacgat gctaccattg ctgagatcaa gaagagcccc 240
gaggttgcct tcgtcgagcc cgactacatt gtcaccctcg actacaaggt tgagcctctc 300
tctgaccgtg ctctgaccac tcagagcaac gtccttggg gtcttgctgc catctcccgc 360
cgaacccccg gtggcagcac ctacacctac gacaccactg ccggtgccg tacttacggt 420
tacgtcgttg actctggtat caacaccgcc cacactgact ttggcggccg tgcttctctc 480
ggttacaacg ctgctggtgg cgccacact gatacccttg gccacggtac ccacgttgct 540
ggtaccattg cctccaacac ctacggtgtt gccaaagcgtg taagtacaat cataccccac 600
atgagctaca acatgatctg aactttatac ttactattat taggccaacg tcatctctgt 660
caaggttttc gtcggttaacc aagcttctac ctctgttatc cttgctgggt tcaactgggc 720
tgtcaacgac atcacctcca agaaccgtgc tagccgctct gtcataca tgtctctcgg 780
tggtccctct tctcagacct gggctactgc catcaacgct gcctacagcc aagggtgtcct 840
ctccgttggt gctgccggtg acggtgattc caacggctcg cctctccccg cctctggcca 900
gtctcctgcc aacgttccca acgctatcac cgttgctgcc gccgactcca gctggcgaac 960
tgctctttc accaactacg gtcctgaggt cgatgtcttc ggtcctggtg tcaacatcca 1020
gtccacctgg tacacctcca acagcgctac caacaccatc agcgggtac ccatggcttg 1080

SEQUENCE LISTING.txt

ccctcacgtt gctgggtcttg ctctctacct ccaggctctc gagaacctca ataccctgc 1140
tgccgtcacc aaccgcatca agtctcttgc cactaccggc cgcatactg gcagcctcag 1200
cggcagcccc aacgccatgg ctttcaacgg cgctactgct taa 1243

<210> 13
<211> 392
<212> PRT
<213> Fusarium equiseti

<220>
<221> MISC_FEATURE
<223> The amino acid sequence of the proenzyme form of Fusarium
equiseti RF6318 protease including amino acids Ala21 to Ala 412
of the full length protease.

<400> 13

Ala Pro Thr Glu Lys Arg Gln Glu Leu Thr Ala Ala Pro Asp Lys Tyr
1 5 10 15

Ile Ile Thr Leu Lys Pro Glu Ala Ala Glu Ala Lys Val Glu Ala His
20 25 30

Met Ala Trp Val Thr Asp Val His Arg Arg Ser Leu Gly Lys Arg Asp
35 40 45

Thr Ser Gly Val Glu Lys Lys Phe Asn Ile Ser Ser Trp Asn Ala Tyr
50 55 60

Ser Gly Glu Phe Asp Asp Ala Thr Ile Ala Glu Ile Lys Lys Ser Pro
65 70 75 80

Glu Val Ala Phe Val Glu Pro Asp Tyr Ile Val Thr Leu Asp Tyr Lys
85 90 95

Val Glu Pro Leu Ser Asp Arg Ala Leu Thr Thr Gln Ser Asn Ala Pro
100 105 110

Trp Gly Leu Ala Ala Ile Ser Arg Arg Thr Pro Gly Gly Ser Thr Tyr
115 120 125

Thr Tyr Asp Thr Thr Ala Gly Ala Gly Thr Tyr Gly Tyr Val Val Asp
130 135 140

Ser Gly Ile Asn Thr Ala His Thr Asp Phe Gly Gly Arg Ala Ser Leu
145 150 155 160

Gly Tyr Asn Ala Ala Gly Gly Ala His Thr Asp Thr Leu Gly His Gly
165 170 175

Thr His Val Ala Gly Thr Ile Ala Ser Asn Thr Tyr Gly Val Ala Lys
180 185 190

SEQUENCE LISTING.txt

Arg Ala Asn Val Ile Ser Val Lys Val Phe Val Gly Asn Gln Ala Ser
195 200 205

Thr Ser Val Ile Leu Ala Gly Phe Asn Trp Ala Val Asn Asp Ile Thr
210 215 220

Ser Lys Asn Arg Ala Ser Arg Ser Val Ile Asn Met Ser Leu Gly Gly
225 230 235 240

Pro Ser Ser Gln Thr Trp Ala Thr Ala Ile Asn Ala Ala Tyr Ser Gln
245 250 255

Gly Val Leu Ser Val Val Ala Ala Gly Asn Gly Asp Ser Asn Gly Arg
260 265 270

Pro Leu Pro Ala Ser Gly Gln Ser Pro Ala Asn Val Pro Asn Ala Ile
275 280 285

Thr Val Ala Ala Ala Asp Ser Ser Trp Arg Thr Ala Ser Phe Thr Asn
290 295 300

Tyr Gly Pro Glu Val Asp Val Phe Gly Pro Gly Val Asn Ile Gln Ser
305 310 315 320

Thr Trp Tyr Thr Ser Asn Ser Ala Thr Asn Thr Ile Ser Gly Thr Ser
325 330 335

Met Ala Cys Pro His Val Ala Gly Leu Ala Leu Tyr Leu Gln Ala Leu
340 345 350

Glu Asn Leu Asn Thr Pro Ala Ala Val Thr Asn Arg Ile Lys Ser Leu
355 360 365

Ala Thr Thr Gly Arg Ile Thr Gly Ser Leu Ser Gly Ser Pro Asn Ala
370 375 380

Met Ala Phe Asn Gly Ala Thr Ala
385 390

<210> 14
<211> 934
<212> DNA
<213> Fusarium equiseti

<220>
<221> misc_feature
<223> The nucleotide sequence encoding the amino acid sequence of the
mature form of Fusarium equiseti RF6318 protease.

<400> 14
gctctgacca ctgagagcaa cgctccttgg ggtcttgctg ccatctcccg ccgaaccccc 60

ggtggcagca cctacaccta cgacaccact gccggtgccg gtacttacgg ttacgtcggt 120

SEQUENCE LISTING.txt

gactctggta tcaacaccgc ccacactgac tttggcggcc gtgcttctct cggttacaac 180
gctgctgggtg gcgcccacac tgataccctt ggccacggta cccacgttgc tggtagcatt 240
gcctccaaca cctacgggtg tgccaagcgt gtaagtacaa tcatacccca catgagctac 300
aacatgatct gaactttata cttactatta ttaggccaac gtcattctctg tcaagggttt 360
cgtcggtaac caagcttcta cctctgttat ccttgctggg ttcaactggg ctgtcaacga 420
catcacctcc aagaaccgtg ctagccgctc tgtcatcaac atgtctctcg gtggtccctc 480
ttctcagacc tgggctactg ccatcaacgc tgcctacagc caagggtgtcc tctccgttgt 540
tgctgccggt aacgggtgatt ccaacggctg tcctctcccc gcctctggcc agtctctctg 600
caacgttccc aacgctatca ccgttgctgc cgccgactcc agctggcgaa ctgcctcttt 660
caccaactac ggtcctgagg tcgatgtctt cggtcctggg gtcaacatcc agtccacctg 720
gtacacctcc aacagcgcta ccaacaccat cagcgggtacc tccatggctt gccctcacgt 780
tgctgggtctt gctctctacc tccaggctct cgagaacctc aatacccctg ctgccgtcac 840
caaccgcatc aagtctcttg ccactaccgg ccgcatcact ggcagcctca gcggcagccc 900
caacgccatg gctttcaacg gcgctactgc ttaa 934

<210> 15
<211> 289
<212> PRT
<213> *Fusarium equiseti*

<220>
<221> MISC_FEATURE
<223> The amino acid sequence of the mature form of *Fusarium equiseti* RF6318 protease including amino acids Ala124 to Ala412 of the full length enzyme.

<400> 15

Ala Leu Thr Thr Gln Ser Asn Ala Pro Trp Gly Leu Ala Ala Ile Ser
1 5 10 15

Arg Arg Thr Pro Gly Gly Ser Thr Tyr Thr Tyr Asp Thr Thr Ala Gly
20 25 30

Ala Gly Thr Tyr Gly Tyr Val Val Asp Ser Gly Ile Asn Thr Ala His
35 40 45

Thr Asp Phe Gly Gly Arg Ala Ser Leu Gly Tyr Asn Ala Ala Gly Gly
50 55 60

Ala His Thr Asp Thr Leu Gly His Gly Thr His Val Ala Gly Thr Ile
65 70 75 80

Ala Ser Asn Thr Tyr Gly Val Ala Lys Arg Ala Asn Val Ile Ser Val
85 90 95

SEQUENCE LISTING.txt

Lys Val Phe Val Gly Asn Gln Ala Ser Thr Ser Val Ile Leu Ala Gly
100 105 110

Phe Asn Trp Ala Val Asn Asp Ile Thr Ser Lys Asn Arg Ala Ser Arg
115 120 125

Ser Val Ile Asn Met Ser Leu Gly Gly Pro Ser Ser Gln Thr Trp Ala
130 135 140

Thr Ala Ile Asn Ala Ala Tyr Ser Gln Gly Val Leu Ser Val Val Ala
145 150 155 160

Ala Gly Asn Gly Asp Ser Asn Gly Arg Pro Leu Pro Ala Ser Gly Gln
165 170 175

Ser Pro Ala Asn Val Pro Asn Ala Ile Thr Val Ala Ala Ala Asp Ser
180 185 190

Ser Trp Arg Thr Ala Ser Phe Thr Asn Tyr Gly Pro Glu Val Asp Val
195 200 205

Phe Gly Pro Gly Val Asn Ile Gln Ser Thr Trp Tyr Thr Ser Asn Ser
210 215 220

Ala Thr Asn Thr Ile Ser Gly Thr Ser Met Ala Cys Pro His Val Ala
225 230 235 240

Gly Leu Ala Leu Tyr Leu Gln Ala Leu Glu Asn Leu Asn Thr Pro Ala
245 250 255

Ala Val Thr Asn Arg Ile Lys Ser Leu Ala Thr Thr Gly Arg Ile Thr
260 265 270

Gly Ser Leu Ser Gly Ser Pro Asn Ala Met Ala Phe Asn Gly Ala Thr
275 280 285

Ala