

TGAGTTTTTCATACTTCAGAACAGTGTGTAATATCCCAAATTTTTATACTATTTTTATAA 79114
AATTAATTTTTGTTATAATAAGATTTTTATATATATATAAATTCTGAAAAATAAAGTAAAAATC 79054
ATAATAAATCAGATTTTTATGAACTGTTTTATGTATTAAATCAGATTTTTTTTTTAAAA 78994
GAAATTTTCGAAATCAAAATTTTTGAAATTAATCAGATTTTTATCTTTTGAAGAAAAAAA 78934
AATTGGAATTTGATTTTCAGATTTTTCGTCCAAAACGAAAAATAGAGAGAAAAGAAAATT 78874
CTGAAATTATAATTTGAGTTGGTTTTGGAAAAGGATTAGATTTTTGTAAATACTTATCTT 78814
TTAGTGAAACCTAGATTTACATATATATATATATACCCCCAAAACACCAAAAAAATTCT 78754
CACGTAATACACTTCTTCATCTTTTTGGTAAGTTTTAAATTTTCAGATCTAAAATCACCA 78694
TTGTTGTGTTTGAAGGTTTCAACAAAAAAAAAAAAAACTTTTTAAATCCGATGACTTGGT 78634
CGGAGTCCGGCGTCGGTTTTCTTCTTCTTCTTCTTCTGTTTTCTTCTTCTTCTTCTT 78574
TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 78514
GCGTGCGATGCTTCACAGTACCTCCAAGATTTGATGTTTTTAAAATAAAAAAAGAAAAA 78454
AAAAAAGAAAAGGAATTACATGTGCATTAAGTCCACGAGAGATTAGACTTACTGTAA 78394
CATGAAAGTTGAAAAAAATTTCAAGAAAAATTTCCCTTTCAATCTGTGACGACGCTGTC 78334
ACGTTTTTTTTCAAGTAAAAGCTTTTTAATTTTTTTTTCTTTTCAACCTTGGTTATTTAA 78274
CTTTGTCAACTATTAAGTATTAACCTTGTAAGCTTTGATTAATTTTTCTTTCGGAATATT 78214
ATATTAAGTGACTAACATTAGGTAATATTATTTGTTGTAGACGGAACTTCGTAGAAGGC 78154
GATTTTTAGTTGATATTGCTATACTTGGGACGCGTTTGAGGTAATTCACAATGTCCATCA 78094
ACAAC TAGAATCCCGTTAGGAATTGTATGTGCTATGTGTAATGCATGTGGTTATGTGATT 78034
CTATATTTATGTTAGATGTATATTATGTGACAGGAATCGTATAGTACTTCTATGACTAA 77974
AATTATTATTATTATTACTATAGTACTATTATCCCTGCGTATAATATATATATGTATCT 77914
ACTGGTATTGTTGTTATGAATTTGGATTTATAATGTACCTAATTATAACCATCGTATTTAA 77854
ATTATGCTAATATTGACAATGTACTTAATGGGTATAACTAGTGTGTTAATGATGTAGCGA 77794
ACGGTATTATAAGTTGACATATATTGTAACCTCTATGAGGCTCTAATGATGGATATATTGG 77734
TTACTACGAATTATGGACGTTGATTAGTATTAAGTGGCTAAGTTGTGAAAAATATTATTT 77674
GAAC TAAAGATGTTCCCTGCTAATGTAAATGTGATGTTTGATGTGTCACAAC TTTTAAAA 77614
TCTATTAATCACGTAAAGTGAAATTGAGGGAATTATCCTGTGGATTTGGATCCTCCATA 77554
GGTGATGAACAGTACTTGATTTTATTATGATACAACCTTTTATTGTCTTCTCCTAATAC 77494
TATTGGTGCGCATTGCGGATACCCATTAGATTAGTGAGGGGTGTGCACACTAGGGACGCA 77434
CTGCAAAACGATATTGGCCATTGCTCTTAATTTATTGGGTGACGACCCATGTTGAAGTAG 77374
GTGGTTACTGGGATTAGTCCCGCCTACTGACGTTTCGATTCCCTCTAATAATTAAATGT 77314
GACGTTGTGTGCATCATATAAACTTGGAGTATATAAATTGATTAAGTGTGTTATGTACTT 77254
GTTTGTATTACGTTATACTTGTCTGCATTGTATTATATTTGTATGTATGTTATAAAAAATT 77194
ATTTTTTAAAATACAAC TATTATTATTATTTTACGGGATGGAGTTGTAATTACTTAGC 77134
TTTCGCTAATTTTTGTGTTTTTGTTTTTCTTTGCTCTTTTCTTATTTATATTGTGCAGGT 77074
TGGTGAAAGGGACTACGTTGCAGGAATGAGTGAATTTATAGTTTTTAAGTTCACCTAGGT 77014
TAATAAACACTACAGTACTGCCATTTGAGATTGTTTAGTTTTGTTTTGGTATAAAGTAGA 76954
CAACTCAGTTTGGTTTATTTTGATATTTGGTTTGTGAGTAAAAATTTTTATTTCTAATAA 76894
TTAGCCTTGCATTTATTTTGAACAATAGATGCGGCGGTAACACCCTAAAGGTTTGGTCAG 76834
ATATTTTTAATGAAAAGTGAATATTTTTATAAAGTTTTAGAACCAAAGTTTTGGGGTGTTAC 76774

AAATTGGTATTCAGAGCTTAGGTTATAATAAATAAATATAAATTAGGAGTTAATAAAATT 76714
TGTTAAATAATTAAGAGTCTAAGATTAATTAAGTAATAAGCATTAAAAATAAATTGTGAG 76654
AGTGGGATCGAAAAGGGTTTATGTGCCTTAATGTGAAAGCACTTATTTTCAGGTGCTTATT 76594
AATTGTTTCTTTTCGTTCTTGTATGTTTAGATGTTGTAAGTACAAGAAAATGTTGTGCTC 76534
GGATGACCGACCATAGGATTTAGAGATTTATATTAAGGTCCGTATACGAAGTTATTTAGT 76474
AGTATTGAAAATTCTCTAAGTTCCTATTTATGTGTTTTAATTATGAATCTTATTACTTGA 76414
CTATTTGAGTTGATTAGAGTTTAAGTTCTTGTATAATTATTGCTTACTTGTCAAATTGA 76354
TGTTATGTGGAAGCATGCAAAGTTGTCTTGAAATGTATGTTATGATGAAATATAGTTTG 76294
GTTGAGTCATGAAATTTTCTTATGATGAAAGTTTGTATTGAGTAAATAAAACGATGTGCA 76234

AGACCATGTATGGGTGAATAGATGTAGTAATACCAATGCTGATGTGCAGGAATTGTATGC 76174
M L M C R N C M Q

AAACCTATCAAGTATACACTAGATCTCGTAGGAGGACTTTTGAACAAATGGCTGAAACCC 76114
T Y Q V Y T R S R R R T F E Q M A E T P

CTGAACAACACTACTTGAGAGATTGAGATCTCTTGAACAACCTTTTCGCAACGTATGGGTTTAG 76054
E Q L L E R L R S L E Q L S Q R M G L V

TGTTACAAAACCAATTAGGAAATAATGGTGGAGAGGACCCACAAGCTGCTATGGCAAAGA 75994
L Q N Q L G N N G G E D P Q A A M A K K

AGTTAGCAACCCTTAAACCTCCAATCTTTGTAGGAAAGGAAGACCCCTTACTCTTAGAGA 75934
L A T L K P P I F V G K E D P L L L E N

ACTGGCTAAGAGACTTTGATAAGTTATTTCACTGCTACTGGGACACCTGAAGCTCAAAAAG 75874
W L R D F D K L F T A T G T P E A Q K V

TAGACCAAGCTACCTTTTATCTGAGGGAGGATGCAGACACTTGGTGGGAGAGTCAAGGAC 75814
D Q A T F Y L R E D A D T W W E S Q G P

Retrotransposon gag protein (pfam accession PF03732)

CTATTGTTAGAGCTCAGGAAAACCTTTAATTGGAATGCTTTTAAGGTTGCTATTAAGGATA 75754
I V R A Q E N F N W N A F K V A I K D R

GATTTTTCCCTGAACATATTAGGAGGCAGAAATACAATGAGTTCACTAGATTTAACCAGG 75694
F F P E H I R R Q K Y N E F T R F N Q G

GAGGTACTATGTCTGTGCAAGAGTATGCCAAAAGTTCAATGAGTTAGCTAGATTTTGCC 75634
G T M S V Q E Y A Q K F N E L A R F C P

CTAATGTTGTGCCAGATGAGAGAGCTAAGGCTCAAAAGTTTGAGGATGGTTTAGCATTTA 75574

N V V P D E R A K A Q K F E D G L A F R

GAATTCAGACCAGACTTGGGGGAGCAACTTCTGCAACTTTTCAGGAAGCTTATGCTAAGG 75514

I Q T R L G G A T S A T F Q E A Y A K A

CTTCTAATATTGAGAGGATTTTGAGGCGTGAAGAGGAAGTTATGGGGAGGAATAAGAGAA 75454

S N I E R I L R R E E E V M G R N K R K

AAGACCCACCTAGCAACCAAAATGACCATGGAAATGACAAGAAACCTCGATATGGGGGTA 75394

D P P S N Q N D H G N D K K P R Y G G N

ACAATAATAATGGGGGCAATAATCACACTAATGGTGGTGGTAATTATCAAGGGAATCGTA 75334

N N N G G N N H T N G G G N Y Q G N R S

GCAACTACCAAGGTCAGGGGAGATCAAACCAGCAAGGATCCCGTACCCAGAACCCTACTT 75274

N Y Q G Q G R S N Q Q G S R T Q N P T C

GTAGAAAGTGTAACAAAAGCCACCCAGGATTTACCTGTCAAGGAGACCCAATAACTTGTT 75214

R K C N K S H P G F T C Q G D P I T C Y

Zinc knuckle (pfam accession PF00098)

ATGCTTGTGGAGAGAAAGGGCATAAGGCTAATCAGTGTCCCAAGCGTCAGAATAATGGAC 75154

A C G E K G H K A N Q C P K R Q N N G Q

AAAATGGAAACAATGGGGGAAATAGGAATGGTCATGGGCCTAATCAGAACCAGAATAACA 75094

N G N N G G N R N G H G P N Q N Q N N N

ATAACCGTCCCTACAACAACAACAACCTCTCAAGGTCAAACCTTCGAATGCTCAAGGGGGGA 75034

N R P Y N N N N S Q G Q T S N A Q G G N

ACAATACTCAGCATAATGGTCAGAATAACAATCGAGCAAATGGAGGAAACAACAATCAGA 74974

N T Q H N G Q N N N R A N G G N N N Q N

ACGGCAATGGAAATGGTGCTCGAGGCAACAATGGAAGAATCTATGTTATGAACCAGAATG 74914

G N G N G A R G N N G R I Y V M N Q N E

Retroviral aspartyl protease (pfam accession PF08284)

AAGCAGACACCAACGCCAATGTTGTGACGGGTACTTTCCCTCGTAAACTCTAACCTGCTT 74854

A D T N A N V V T G T F L V N S N P A Y

ACTTGCTTTTTGATTCTGGGGCGTCTCATTCTTTTCATAGCTAGTTCATTTGTTGAAAAGT 74794
 L L F D S G A S H S F I A S S F V E K L

TAGGTCTAAAACCTCAATCTTGTGTCAAACCTTTCATTACAATACCTTCAGGAGAAGTAG 74734
 G L K P S I L C Q T F I T I P S G E V V

TTCCTTGTAGTTCTCTATACCAAGACATACCCATTACCATATTAGGATCTGATTTGCCGG 74674
 P C S S L Y Q D I P I T I L G S D L P A

CTGATCTTATTCAGTTTGACCTACCCGACTTTGATGTAATATTGGGAATGGATTGGCTTG 74614
 D L I Q F D L P D F D V I L G M D W L A

CTAAGTATAGAGCTAGGATAGAGTGTCACTCAAAGGTGTCTCTTAGGGGGCCAAAGG 74554
 K Y R A R I E C H T Q K V S L R G P K G

GAAATAGAATATCCTATCAAGGAATTGTTTCTAAACCTGGAGTCAGTATTGTGTCAGCCA 74494
 N R I S Y Q G I V S K P G V S I V S A M

TGTCAATCAAACCTATATTAGGAAGGGCTACCCCATATACTTGTGCCATGTGAAGGATG 74434
 S F K T Y I R K G Y P I Y L C H V K D V

TGAGTGTGGAGGATGGAGAGATATCTCAAATACCTGTGGTGAGTGAGTTCCAAGATGTTT 74374
 S V E D G E I S Q I P V V S E F Q D V F

TTCCAGAAGAAATTCCAGGGATGCCGCCAGTGAGAGAAATGGATTTTAAGATTGACCTAG 74314
 P E E I P G M P P V R E M D F K I D L V

TGCCTGGAACCTGGAGCTATTTCTAAGGCACCATATAGGATGGCACCTGCAGAGATGCAAG 74254
 P G T G A I S K A P Y R M A P A E M Q E

AGTTGAAAGTGCAATTGGAGGAATTATTGGAGAAAGGGTACATTAGGCCAAGTGTTCAC 74194
 L K V Q L E E L L E K G Y I R P S V S P

CTTGGGGAGCACCAAGTGTATTTGTTTCGAAAGAAGGATGGAACCTTGAGGTTGTGTATTG 74134
 W G A P V L F V R K K D G T L R L C I D

RNA-dependent DNA polymerase (pfam accession PF00078)

ATTACAGAGAGTTGAATAATGTCACAATAAAGAATAAGTACCCATTGCCTAGGATTGAGG 74074
 Y R E L N N V T I K N K Y P L P R I E D

ATTTATTTGATCAACTTAAGGGTGCTGGAATTTTCTCTAAGATTGATTTGAGGTCTGGGT 74014

L F D Q L K G A G I F S K I D L R S G Y

ATCACCAATTGAGAATTTTCGGAGGAAGATATACCAAAAAACAGCTTTTTCGTACGAGGTATG 73954

H Q L R I S E E D I P K T A F R T R Y G

GGCATTATGAGTTCACAGTGATGCCATTTGGACTTACTAATGCACCTGCAGCATTATGG 73894

H Y E F T V M P F G L T N A P A A F M D

ATCTTATGAATAGAACATTTTCAGCCGTATTTAGATAGATTTGTGGTGGTGTTCATAGATG 73834

L M N R T F Q P Y L D R F V V V F I D D

ATATATTGGTGTATTCGAAGGATAAAGAAGAGCATGAAGGTCATTTAAGGAAAGTTTTGG 73774

I L V Y S K D K E E H E G H L R K V L E

AGATACTTCGAGAGAAAAGGTTGTATGCTAAGTTATCAAAATGTGAGTTTTGGCTTGAGA 73714

I L R E K R L Y A K L S K C E F W L E K

AAGTTGCATTTTTAGGTCATGTGATTTTCGAAGGAAGGTGTTGCTGTAGATCCATCAAAGA 73654

V A F L G H V I S K E G V A V D P S K I

TACAAGCAGTAACAGAATGGGTGAGACCTAGTAATGTGACTGAGATTAGAAGTTTCTTAG 73594

Q A V T E W V R P S N V T E I R S F L G

GACTTGCTGGCTACTATAGGAGGTTTGTGCAAGATTTCTCAAAAAGTAGCTCAACCTTTGA 73534

L A G Y Y R R F V Q D F S K V A Q P L T

CAAATTTGATGAAGAAAACAACCTCGATTTTCAGTGGGATGAGAGGTGTGAGAAAGCTTTTC 73474

N L M K K T T R F Q W D E R C E K A F Q

AGGAATTGAAGCAAAGACTTACTTCAGCACCAGTTTTGACATTACCATCTGGATTAGAAG 73414

E L K Q R L T S A P V L T L P S G L E G

GTTTTGAGGTGTATAGTGACGCTTCTAAGAATGGGTTAGGATGTGTATTGATGCAACATA 73354

F E V Y S D A S K N G L G C V L M Q H S

GTAAGGTGGTAGCATATGCTTCGAGACAACCTTAAGCCTTATGAACAGAATTACCCTACTC 73294

K V V A Y A S R Q L K P Y E Q N Y P T H

ATGATTTAGAGTTAGCTGCTGTAGTATTCGCATTGAAAATTTGGAGGCATTATTTGTATG 73234
 D L E L A A V V F A L K I W R H Y L Y G

GTGTGTCATGTAAGATTTTCACTGATCATAAAAAGTCTGAAATATATATTTACTCAGAAGG 73174
 V S C K I F T D H K S L K Y I F T Q K E

AGTTGAACATGAGACAGAGGAGATGGCTTGAACCTTATTAAGGATTATGATTTAGAGATTT 73114
 L N M R Q R R W L E L I K D Y D L E I L

TGTATCATGAGGGTAAAGCGAATAAAGTTGCTGATGCATTGAGTAGGAAGACTAGTCATT 73054
 Y H E G K A N K V A D A L S R K T S H S

CGATGAACATGATGGTGTATCTGAGAGATTGTGTGAAGATTTTCAGGAGCATGAGTTTAG 72994
 M N M M V L S E R L C E D F R S M S L E

AAGTCATGGAGCAAGGGCAAGTGAAGCTCAATTGAATGCACTATGCGTGCAACCCACCT 72934
 V M E Q G Q V E A Q L N A L C V Q P T L

TATTCGATGAGATTCGAGAGAAGCAAAGTAGTGATGAGTGGATGGTGAAGATAAAGAAAA 72874
 F D E I R E K Q S S D E W M V K I K K M

TGAAAGAAGATGGAGTTGTCATCGAGTTTGACATTGATGAAAATGGTGTGTGAAGTACA 72814
 K E D G V V I E F D I D E N G V V K Y K

AGGGAAGATGGTGTGTTCCCTAAGGATGAGGAGTTAAAAAGAAAGATTTTGAAGAAGCTC 72754
 G R W C V P K D E E L K R K I L E E A H

ATAATACTCCATATTCTGTGCATCCTGGAGGAGATAAACTTTATAAGGATTTGAAGCAGC 72694
 N T P Y S V H P G G D K L Y K D L K Q H

ATTTTTGGTGGAAAAACATGAAACGTGAAGTGGCAGAGTTTGTGCAAAGTGTGTTGACGT 72634
 F W W K N M K R E V A E F V A K C L T C

GTCAGAAAGTGAAGATTCAGCATATGAGACCTGGTGAATGATGCAACCTTTAGAAGTGC 72574
 Q K V K I Q H M R P G G M M Q P L E V P

Integrase core domain (pfam accession PF00665)

CGAGTTGAAATGGGAGTCTATTTCAATGGATTTTGTGATGGGATTACCACTTACTAAGT 72514
S W K W E S I S M D F V M G L P L T K S

CAGCTAAGAATGCCATATGGGTTATAGTGGATCGATTGACAAAGTCGGCCAGATTTATAG 72454

A K N A I W V I V D R L T K S A R F I A

CAATGAAGGATACATGGAGTATGCAACAGTTGGCTAGTGCATATGTGCGAGAGGTTGTTA 72394

M K D T W S M Q Q L A S A Y V R E V V R

GACTGCATGGAATACCAAAGGATATCGTTTTAGATAGAGACTCGAGATTTTTGTCCAAGT 72334

L H G I P K D I V S D R D S R F L S K F

TTTGGGGGAGGTTACAACAAGCCTTTGGGACATTGCTCAAATTTAGTACAGCTTTCCACC 72274

W G R L Q Q A F G T L L K F S T A F H P

CTGCAACAGATGGACAGACAGAGAGAACAATTCAAACATTGGAGGATATGTTGAGAGCAT 72214

A T D G Q T E R T I Q T L E D M L R A C

GTGTGATAGACTTTGGAGGATCTTGGGATGATTATTTGCCAACTATAGAGTTTTCGTATA 72154

V I D F G G S W D D Y L P T I E F S Y N

ACAACAGTTATCACTCAAGCATAAAGATGGCACCGTATGAAGCATTGTATGGGCGAAAAT 72094

N S Y H S S I K M A P Y E A L Y G R K C

GTAGGAGTCCTTTGTGTTGGAGTGACATAAGTGAGACGATGACTTTAGGGCCTGAGATGA 72034

R S P L C W S D I S E T M T L G P E M I

TTGAAGAAACAACGAAACAAGTTAGGCTTATTCAGGAGCACATGAGGGCAGCTCAAGATA 71974

E E T T K Q V R L I Q E H M R A A Q D R

GACAAAAGGCTTACGCAGATCAGAATAGAAGGGAGATGGAATTTGAGGTTGGGGAGAAGG 71914

Q K A Y A D Q N R R E M E F E V G E K A

CTTTGCTAAAAGTGTACCAACAAGGGGGTCATGAGATTTGGTAGGAAAGGAAAGTTGA 71854

L L K V S P T K G V M R F G R K G K L S

GTCCACGTTACATTGGACCATATGAGATCTTGAACGAATTGGGAAAGTAGCCTATAGAT 71794

P R Y I G P Y E I L E R I G K V A Y R L

TAGCCTTACCAATGGAGTTAGCTAATGTCCATAACGTCTTTCATGTGTCTCAACTTCGAA 71734

A L P M E L A N V H N V F H V S Q L R K

AATATGTCCATGATCCTACCCATATCATTCAACCTGAAACCATTGAACTAGATGAAACCT 71674
Y V H D P T H I I Q P E T I E L D E T L

TATCCTTTGAGCAACGCCAGTTAGGATTCTTGATACCAAAACGAGAAGTACCCGGAACA 71614
S F E Q R P V R I L D T K T R S T R N K
Chromo domain (pfam accession PF00385)

AGGCGGTAAAACCTAGTCAAGGTGTTATGGTCAAGTCAAACCTTCTGAAGAGGCTACTTGGG 71554
A V K L V K V L W S S Q T S E E A T W E

AAGCCGAAGATGATATGAAAAACCGATATCCCGAACCTTCCAGCAGGTACGCTTGAGTT 71494
A E D D M K N R Y P E L S Q Q V R L S F

TCGGGGACGAAACTCTTTAAGGGGGGTAGAATGTGATACTAACTTTTTGTTTGTATATTA 71434
G D E T L *

GTAGCGAGCGATAACGTTAAAGTTCGAGGACGAACTTTCTTTTAAGGGAGAGTAGATGTA 71374
ATATCCCAAATTTTTATAACTATTTTTATAAAATTAATTTTTGTTATAATAAGATTTTTATAT 71314
ATATATAATTCTGAAAAATAAAGTAAAATCATAATAAATCAGATTTTTATGAAACTGTTT 71254
TTATGTATTAATCAGATTTTTTTTTTAAAGAAAATTTGAAATCAAAATTTTGAAATTA 71194
ATCAGATTTTTATCTTTTGAAGAAAAAAAAAATTGGAATTTGATTTTCAGATTTTTCGTC 71134
CAAAACGAAAAATAGAGAGAAAAAGAAAATTTCTGAAATTTATAATTTGAGTTGGTTTTGGAA 71074
AAGGATTAGATTTTTGTAAATACTTATCTTTTAGTGAAACCCTAGATTTACATATATATA 71014
TATATACCCCCAAAACACCAAAAAAATTTCTCACGTAATACACTTTCTTCATCTTTTTGGT 70954
AAGTTTTAAATTTTCAGATCTAAAATCACCATTGTTGTGTTTGAAGGTTTCAACAAAAAAA 70894
AAAAAACTTTTTAAATCCGATGACTTGGTCGGAGTCCGGCGTCGGTTTTCTTCTTCTCT 70834
TCTTCTTCTGTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 70774
TCCTTCTTCTTCTTCTTATTTTCAATTTGGCTTTGCGTGCGATGCTTTCACAGTACCTCCAAGAT 70714
TTGATGTTTTTAAAATAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAGGAATTACATGTGCATT 70654
AAAGTCCACGAGAGATTAGACTTACTGTAACATGAAAGTTGAAAAAAAAAATTTCAAGAAAA 70594
AATTTCTTTCAATCTGTGACGACGCTGTACGTTTTTTTTTCAAGTAAAAGCTTTTTAAT 70534
TTTTTTTTCTTTTCAACCTTGGTTATTTAACTTTGTCAACTATTAAGTATTAACCTTGTA 70474
AGCTTTGATTAATTTTTCTTTCGGAATATTATATTAAGTGACTAACATTAGGTAATATTA 70414
TTTGTTGTAGACGGAACTTCGTAGAAGGCGATTTTTTAGTTGATATTGCTATACTTGGGA 70354
CGCGTTGAGGTAATTCACAATGTCCATCAACAAGTAAATCCCCTTAGGAATTGTATGT 70294
GCTATGTGTAATGCATGTGGTTATGTGATTCTATATTTATGTTAGATGTATATTATGTGA 70234
CAGGAATCGTATAGTACTTTCTATGACTAAAATTTATTATTATTATTACTATAGTACT 70174
ATTATCCCTGCGTATAATATATATGTATCTACTGGTATTGTTGTTATGAATTTGGATTTA 70114
TAATGTACCTAATTATAACCATCGTATTTAAATTTATGCTAATATTGACAATGTACTTAATG 70054
GGTATAACTAGTGTGTTAATGATGTAGCGAACGGTATTATAAGTTGACATATATTGTAAC 69994
TCTATGAGGCTCTAATGATGGATATATTGGTTACTACGAATTATGGACGTTGATTAGTAT 69934

TAAGTGGCTAAGTTGTGAAAAATATTATTTGAACTAAAGATGTTCCCTGCTAATGTTAATG 69874
TGATGTTTGATGTGTCACAACCTTTTAAAAATCTATTAATCACGTAAAGTGGAAATTGAGG 69814
GAATTATCCTGTGGATTTGGATCCTCCATAGGTGATGAACAGTACTTGATTTTATTATGA 69754
TACAACCTTTTATTGTCTTCCTCCTAATACTATTGGTGCGCATTGCGGATACCCATTAGA 69694

FIGURE S4.– Nucleotide sequence of *bvgypsy-1*. Numbers of nucleotide residues are coordinated with those of BAC clone 33E19. The open reading frame is indicated with its putative translation product. The 1684-bp repeated sequences are underlined. Putative domains are boxed and shown with pfam descriptions.