

CTCATGACGTCAAGATCTCAAGTCTATAAAAAAAATTGTTTTAAATAAGTCAACCTTTG 11926
 TGCTGTACGCTTCTTATTTTGGAGTCTACCTTTTGGATTATTTCTGATTGAGTTAAGCTTG 11866
 TATGTATGTTCTCTTCTATTGAGTTTTAATTTATTTATGACTTGTTTAGTAGGTTACTTA 11806
 CTTACTTATGATACGTATTACAATGTCACTCTCGTCTGTTTGAGAAGAAATGACATTGTA 11746
 ACACACATCATAAGTAGTCATTGCATTTGTAATAGCAATCTTAGGTAGAGAGAGAATGCC 11686

 TAGAGAGAGAGAGAGAGAAAACTCTGGAGCGAGCGAAGAGAAGGAGAATGGACAATGGT 11626
 M V
 >ORF-A
 AAGAAGGAGACACCCCCAAGCCAGTAAACCACAACCTAGAGCCTTGAGAACAGCCTTCAT 11566
 R R R H P Q A S K P Q P R A L R T A F I
 RNA recognition motif (pfam accession PF00076)
 AGATTTCTTCTCTCCCAATATTGATACCCAAACAATCCACAACATATTCAGTAGATATGG 11506
D F L P P N I D T Q T I H N I F S R Y G
 TGATCTGGAGGACTTAGTGATACCAGCAAAACTCCGGAAAAACTGTGGGCACAAATACGC 11446
D L E D L V I P A K L R K N C G H K Y A
 ATTCATTAATTTTTTCTCCATGAATGCTTTACTCAATGCGATTAAGCAGGAGAATGGAAG 11386
F I K F F S M N A L L N A I K Q E N G R
 AAGAATGGGAAATTTTTTGGATGCGAGTTAACCTGCAAAATATGACAAACAAGACCCTCC 11326
R M G N F L M R V N P A K Y D K Q D P P
 CCATAAAAACCACTTTCCAAATCCTAAACCAAAATCACAGACAGCCTCAAAAAAACCCGGT 11266
 H K N H F P N P K P N H R Q P Q K N P V
 ACAATATCATCCAGCTTGGAGAGACCACCGATCGTATAAGGATGTCTCGAACCCAAACCA 11206
 Q Y H P A W R D H R S Y K D V S N P N Q
 AATACCAATCCACACTGATGTTCCACCAATCAATCCCTCAACCAAACCTAATACCCGGAA 11146
 I P I H T D V P P I N P S T K P N T R K
 ACCACCTCATCAAACGAATTTATCCTCTTACCTATAGAATCAATCATCCCTAACCAAAT 11086
 P P H Q T N L S S S P I E S I I P N Q I
 CCTTGAACCTCTCAGTACTGACATTGTGAAAGAAATGACAAAGCACCGTAGGATGAGTTC 11026
 L E P L S T D I V K E M T K H R R M S S

TAGGGTCCTTGGGGAAGACACGGAGAGAATAAGGGACCAAGTGGAACTTGTGGAAGTAGA 10966
R V L G E D T E R I R D Q V E L V E L E

GGGCGATCAGATTCTTGCCATCTCAGGGGAGAAAAATGAAGAGATCCTGGAGTTACTGGA 10906
G D Q I L A I S G E K N E E I L E L L E

AAGAAGCGTTATAGCAGTCGCAAACCTTTCATCTCCATCCAAGATTATCCATGAGCATAT 10846
R S V I A V A N S S S P S K I I H E H I

CTTGGCGGAAGGGGTTAACTATCTGAAGATTAAACCCCTTGGGGGAATGCTTCATCTTAT 10786
L A E G V N Y L K I K P L G G M L H L I

CCAGTTCAATTCGGTTGAAGAAAAGGATGACATGATAAAAAGCAAATGGCTTGAACGATG 10726
Q F N S V E E K D D M I K S K W L E R W

GTTCTGGAGCTAAGGGATGTGAATAACGCTAGCACGGCATTATGGAGGGAGATGTGGAT 10666
F L E L R D V N N A S T A L W R E M W I

CACAATTTATGGAGTTCCATTGATCGCATGGAGTTATGAAAATTTTCAGAAAATTGGTTG 10606
T I Y G V P L I A W S Y E N F Q K I G C

TATATTCGGGAGAGTGCTATCGGTGGAATATTCTCGCATGGATTACGCCAGAGTTCAGTT 10546
I F G R V L S V E Y S R M D Y A R V Q L

AATCACAGATTGTCTCTTCAAAGTCAATAACCCCATAGTTTTTTTACGTGGAAGATAAACC 10486
I T D C L F K V N N P I V F Y V E D K P

GTTTAAGATTTTTGTTACAGAAGACTTTGGTCTTGGTCCAAATCATGATCCTCCTGCAAG 10426
F K I F V T E D F G L G P N H D P P A S

TAAAGGTATGCCAAATCCCCTCTTCCATAGATTAGATTCTGATAACTCGAATTCGGAATC 10366
K G M P N P L F H R L D S D N S N S E S

CTCTGATAAAGATCCATTGGATGATGATGATCGTGATAGTGACGACTGGGATCCTCCGGG 10306
S D K D P L D D D D R D S D D W D P P G

AGGGGAAAGGTCACCCCAAAAACCCCTCCCAAACCTCCGAGTTCAATGCATCGGGAAATAC 10246
G E R S P Q K P L P N S E F N A S G N T

TCCAGGAATCTCGGACATTGGAGTTAATAATGAGACAATCATTACTTCGCCTACCAAAGC 10186
 P G I S D I G V N N E T I I T S P T K A

TAGTGCCAAGGTCTCTCCCAATGCAAAAACAAAACCTCCCATATACCCTAAACCTCCCAA 10126
 S A K V S P N A K Q K P P I Y P K P P K

AACTCAACTGAACTTTAATACCCCCACCACGTTCCCCAAGTCTGCTTTGCATTGGGAACTT 10066
 T Q L N F N T P P R S P S L L C I G N L

AAATCAACAAAAGTCTCCTCCCAACCACTTGAAGTCCAAAAAGCCCCACCTTCACCATC 10006
 N Q Q K S S S Q P L E L Q K A P P S P S

GAAAACCTTACCCTTCCCTCCAACAACGAAACTGGGCTCACCTTTTAGCCCTGATCCAAC 9946
 K T L P F P P T T K L G S P F S P D P T

CTTTAAATATAATAATCCCCCATCTCCCAAAATAATATAATCAGCCCAATAAGCCCATT 9886
 F K Y N N P P I S Q N N I I S P I S P L

GGTCCCCAACCTGCCCAAAATACACAAAACCTCCCTAGTTCTACAAGTCGAAACTCTCC 9826
 V P K P A Q N T Q N S P S S T S R N S P

TTTAAAGCCCAGCCTCAATGACCAAAGCTTTCCTTACTACAATCCTCTGATCCACACTGA 9766
 L K P S L N D Q S F P Y Y N P L I H T D

TAATTCCTTTGGCCCGCTACTAAGGAAAGCCCAATCAAAATCCCAAACTAAGACACTCTC 9706
 N S F G P L L R K A Q S K S Q T K T L S

ATCCTCTCCTTCGACGTCCAGCCCTTCTATCCCCCGGTTTTGAAGACTTCCTTCCTCC 9646
 S S P S T S S P S I P P G F E D F L P P

CCCTCTGAAAGCCCATCATGAAAAAAGGAGATTACAAAACGACTGAAGAAAAATAAAGC 9586
 P L K A H H E K R R L Q K R L K K N K A

CAAAAACCGCCTCTCCTCCTCCTCCTCCAATCCCCACCTCTTCCTCCCTCTCCCTCCCC 9526
 K N R L S S S S S N P P P L P P S P S P

AAACCCGAAAACATCTCATGAGAACACTGCCTCGGAAATTATTGAATTAGGCTTGCAACT 9466
 N P K T S H E N T A S E I I E L G L Q L

AGGAATGAAATTCAATGGTGAACATCAGATCTACAAGACAAAATTGTTGGAATTTTGTGTC 9406
 G M K F N G E L S D L Q D K I V G I L S

ACGCCAGGAGCAGGACTGGCTTTCCAATGTATAAGTACATCTTATACTCTCAATAAATTG 9346
 R Q E Q D W L S N V *

TTCCATGTTAATCTCGTGAATGTCAGGGGCTCGGAGCATGGCCTAAAAGAAATGTTCT 9286
 CAAAAAGTTACTACTCCTTCATGACCCCATGATAGTATTCATCCAAGAATCCAACTGGA 9226
 ATGTATTCCTTCTAAATTGCAAAAATCAATTTGGTGTGATGATGACCTCAGCCTCTGTAT 9166
 CAGTCCATCAAACGGATCCTCTGGAGGATTAATCTCCCTATGGAGACCCTCAAATTTCA 9106
 TCTGGTTTCCAGTAGAATCGAATCACAATGGATCGCAATGGAAGGAATGGTGGTGAGGGA 9046
 M E G M V V R E
 >ORF-B

AAATTTTCAATGCCTTCTCATAAATATTTATAACTCCTGTGATGCTTCGACTAGATCAGA 8986
 N F Q C L L I N I Y N S C D A S T R S D
 Endonuclease/exonuclease/phosphatase family (pfam accession PF03372)
 CACATGGAACCATATAGAGGATTTTTGCAGAACTCACACTTACCTCTTCTAATAGCGGG 8926
T W N H I E D F C R N S H L P L L I A G

GGATTTCAATGAGGTACTATCTTCCCAAGATCGAGGCAGCCGGATAATAGATGAACTAG 8866
D F N E V L S S Q D R G S R I I D E T S

TGCCGGAAAATTCAGGCAATTCATAACCAACCTTCATCTTACTGAAATCACACCCTCCAA 8806
A G K F R Q F I T N L H L T E I T P S N

TGGATACTTTACATGGTTTtaggggtcaatcaaaatcaaagctggatagaattcttgcca 8746
G Y F T W F R G Q S K S K L D R I L V Q

ACCAGATTGGATTCTGAAATTCTCATTCCCTTAATGCCTCCATCCTCAAAGGAGTATCTC 8686
P D W I L K F S F L N A S I L K R S I S

GGATCATTGCCCTCTTGTACTGAAGTCGCAATCTAAGGACCGGGGACCGAAACCTTTCAG 8626
D H C P L V L K S Q S K D R G P K P F R

ATTTCTTGACATGTGGCTCACCCACAAGGATTGCCTGATCCTTACTAGGAAAGTATGGGA 8566
 F L D M W L T H K D C L I L T R K V W E

AGATTGGAAGGATTACAATTTAGAGAAGTTTAAAGCTGTCAGAAAAGAGTTGAAAGT 8506
 D S K G F T I S E K F K A V R K E L K V

ATGGAACCAGTCAAAAATTTGGGAATCTAGAAACCAATATCTCTCAATTAGAAGACGAAAT 8446
 W N Q S K F G N L E T N I S Q L E D E I

TCACAAATGGGATACTGTTGCCAACACGAGAAACCTATCGGTTGATGAACTGAGTCTCAG 8386
 H K W D T V A N T R N L S V D E L S L R

ATCCAAGGCCCAACTGGATTTATGGGATTGGATCAAACGCAAGGAAATTCCTGGGCCCA 8326
 S K A Q L D L W D W I K R K E I H W A Q

GAACTCTCGTATAAGTTGGTTGAAGTGTGGGGATAAGAACTCAAAGTTCTTCCATGCCTA 8266
 N S R I S W L K C G D K N S K F F H A Y

TGCATCGATTAGAAGACGGAAGAATAATATCTCTTCCATCACGATCGATGGTGAGACCGT 8206
 A S I R R R K N N I S S I T I D G E T V

CTGTGACCCGGAAAAAATCAAAGCCGAAGCCTCACTCTATTTCCAAAATCTGTTCTCAGA 8146
 C D P E K I K A E A S L Y F Q N L F S E

AGAAACCTTTTCCAGACCAACTTTCTTGAACCTAGCCTTCAAAAAACTCTCATCAATACA 8086
 E T F S R P T F L N L A F K K L S S I Q

ATCCTCGGACCTCACCAAACCTTTCTCACACTCTGAAATAGAAAAAGCAGTAGCATCATG 8026
 S S D L T K P F S H S E I E K A V A S C

TAGCCCTTCAAAATCCCCTGGCCCGGATGGTTTCAATTTTAACTTCATAAAGTCTTCCTG 7966
 S P S K S P G P D G F N F N F I K S S W

GGCAATCATCAAAGAAGACATTTTCTCACTTGTCAATGAATTCTGGCAGTCTGGAACACT 7906
 A I I K E D I F S L V N E F W Q S G T L

ACCAAGGGGTAGTAATGTAGCGTTCATAGCGCTGATCGCCAAGGTGGAAGCCCCCTCAAA 7846
 P R G S N V A F I A L I A K V E A P S N

CTTCAAGGACTTCCGACCCATCAGTATGGTCGGTAGCCTTTACAAGATAATTGCGAAGTT 7786
 F K D F R P I S M V G S L Y K I I A K L

RNA-dependent DNA polymerase (pfam accession PF00078)

GCTTTCCTTCAGGCTGAAAAATGTTATGAACGATCTTATTGGGCCCAACAATCTTCTTT 7726
 L S F R L K N V M N D L I G P Q Q S S F

TATTGAGGGGCGCCAGATCTTGGATAGTGTTTTAATCACTGGCGAGTTATTGGACTCATA 7666
I E G R Q I L D S V L I T G E L L D S Y

CAAAAGTTCCAAGATGGGGGCGAGTAATGTTAAAACTGGACTTCCACAAGGCCTTTGACAG 7606
K S S K M G A V M L K L D F H K A F D S

TGTTTCCTGGTCTTTCTTGGATTGGACCATGGATCAAATGGGCTTCCCATTAACATGGCG 7546
V S W S F L D W T M D Q M G F P L T W R

AAAATGGATCTCCTCCTGTGTCTCATCTGCAGCCGCATCTGTCTCCTAAATGGCTCTCC 7486
K W I S S C V S S A A A S V L L N G S P

TTCGACTCCGTTCAAGCTCCAGAGGGGCTCCGTCAAGGAGACCCTCTCTCTCCCTTTCT 7426
S T P F K L Q R G L R Q G D P L S P F L

CTTTGTGTTAGCAGCGGAAGTTTTGAATCTCATGATCAGAAAAGCCACAGAATTGAATAA 7366
F V L A A E V L N L M I R K A T E L N K

ATGGTCAGGTATTGCTATTTGTAAATCGGGTCTATTCTAACTCATCTTCAATTTGCAGA 7306
W S G I A I C K S G P I L T H L Q F A D

TGATACGATAGTATTCTCAACTCCGGATTTGAAGGCGCTCAATAACATCCATAAAACTCT 7246
D T I V F S T P D L K A L N N I H K T L

CATCCTGTTCCAGCTATCCTCAGGCTTGAGATCAACTTCCACAAAAGTGAGATCCTTGG 7186
I L F Q L S S G L Q I N F H K S E I L G

AATCAATACTCCTCAATCTTGGCTTAAAGAAGCGGCAAGGCAATTATTTTGCAGAGTTGG 7126
I N T P Q S W L K E A A R Q L F C R V G

TAATTTCCCGATCACCTACCTGGGCCTTCCAATAGGTGGCAGTTCCGCGAGATTAGCAAC 7066
N F P I T Y L G L P I G G S S A R L A T

ATGGGAACCTCTCTTGGAGAGAATGAGGAAGAAATTGGCCACATGGAAAGAGAAATTACT 7006
W E P L L E R M R K K L A T W K E K L L

CTCGATTGGTGAAGACTCACTTTACTAAAAGCCTCACTCTCGAACCTGCCAATCTATTT 6946
S I G G R L T L L K A S L S N L P I Y F

CATGTCCTTATATCCTATGCCACAAGGAGTTATAGAAAAAATTAATAAAATTCAGAGAAG 6886
 M S L Y P M P Q G V I E K I N K I Q R S

CTTTCTTTGGAGTGGTGGTATGGATAAAAAGGGCTCTATCTATGGTGAAGTGGGAATATGT 6826
 F L W S G G M D K R A L S M V K W E Y V

CCAGCTTCCAAAAGCGTTGGGAGGCTTAAATGTGAGTAACCTTCTGATTAGAAATTTGGG 6766
 Q L P K A L G G L N V S N L L I R N L G

GCTCCTTTGTAAGTGGGTGTGGAGGTATTTTTTCAGAACCAGATTCGCTATGGAGACTATC 6706
 L L C K W V W R Y F S E P D S L W R L S

AATTAAAGCCAAATATAAATACCAGGCGCAAATGAATATGGCTGACATTGCTCCAATAAG 6646
 I K A K Y K Y Q A Q M N M A D I A P I R

ATCAGGTGGTCCCTTGGAGACATCTTTGCAACCATCTCCTAAAACACCAAGCAACAAATGA 6586
 S G G P W R H L C N H L L K H Q A T N E

ACTTCTGAAACAAGGTACCAGGAAAAGAATAGGGAATGGGGAGAATACCTTATTTTGGCA 6526
 L L K Q G T R K R I G N G E N T L F W H

TGACTCTTGGCTGGGCAATTTGCCTCTGAAATTAACCTTCCCAAGACTCTTCCTAATCTC 6466
 D S W L G N L P L K L T F P R L F L I S

AGTTTTACCCATGGCTTCAGTAGCGGAGATGGGTTCTTGGGTGAATTTGGAATGGAAATG 6406
 V L P M A S V A E M G S W V N L E W K W

GAATTTGCCATGGTCCAGAGAATTCAGAAAGAGAGACCGCATTGAATGGGAGCAGCTCCA 6346
 N L P W S R E F R K R D R I E W E Q L Q

ACCTTCCCTCCAGCAAATCTCAGTCCGCCTCAATGAATCAGATGAGTTAATATGGAACCTT 6286
 P S L Q Q I S V R L N E S D E L I W N F

TAGTATGGCTGGTAATTTCTCAGTTCGCTCCTTCTATGAAGAACTTCACAAGCGCTCGAA 6226
 S M A G N F S V R S F Y E E L H K R S K

GCCCTGTCTAGAAAATCTCCCTCAAAGATATGGAAAGGACTTGTTCCCTTCCGAATAGA 6166
 P C L E N L P Q K I W K G L V P F R I E

AATCTTCACTTGGTTATCAGTGCTAGAGAGAATCAATACTAAGAAGAACTAGCTTCTCT 6106
I F T W L S V L E R I N T K K K L A S L

GAACATTATCCCACCCGCTGAGGTGGGTTGCTCATTATGTAGTTTGGAGCCTGAGGATAT 6046
N I I P P A E V G C S L C S L E P E D I

TTCGCACCTCTTTTTGTTTTGCCCTTCTCAATGGAGATTTGGGCTTGGTGGTGGGACCT 5986
S H L F L F C P F S M E I W A W W W D L

TTGGAACCTATCTTGGGTATGGCCAAAATCTCTAAATCTTGCCCTCTCTCAATGGAATTG 5926
W N L S W V W P K S L N L A L S Q W N C

CCCAAGGAAGGAAAAATTATTCAAAAAATCTGGCTGGCAGCATTTCATTGTGATTATCTG 5866
P R K E K L F K K I W L A A F I V I I W

GTCAATCTGGAGAGAACGCAATGAGAGAATTTTCAATAAGAAAGAATCATCAGTTTCAGA 5806
S I W R E R N E R I F N K K E S S V S E

AATCAAAAACCTCATTCTTGTCCGTTTATGTTGGTGGATGAAGCCTTGGAACCTCTCCTT 5746
I K N L I L V R L C W W M K P W N L S F

CCCGTACACAATTGAAGAAGTCATCAGAATCCCACAATGTCTCTTATGGGGTAGCGCTGT 5686
P Y T I E E V I R I P Q C L L W G S A V

GCCTCGAAGAAGTAAACCTCCCATCTCCCCCTCTAATTCAGCTCAGATCTAACCCCC 5626
P R R S K T S H L P P L I Q L R S N P P

TGACCCTTGTCTCAAGTGGATGGTGGGTTTCACCCCGTTCTCGCCAAAAGAAGGTGCTAG 5566
D P C L K W M V G F T P F S P K E G A R

AGCAGGAGGCATTTTTGGAGGCTTCCTCAGAGATGAAGTGGGTGTGATCTTATGCTCCTT 5506
A G G I F G G F L R D E V G V I L C S F

CTCCTGCCCTTTTCCGCCAATGGGTATTAATGAAGTTGCAGTGATTGCAATTCACCGAGC 5446
S C P F P P M G I N E V A V I A I H R A

TCTGCAAATCTCTCTCAGTGTGAAAATCTAAAAGACCGAGAAATCTCAATTTTCTCTGA 5386
L Q I S L S V Q N L K D R E I S I F S E

ATCCAGCCAAGCTATCAGTTGGTGCCTCAATCTCTCATCTGGTCCGACTAATCTCTCCTT 5326
 S S Q A I S W C L N L S S G P T N L S F

CCTGTTGAACTTCATCAGATCTACATGCAAAAAGCTCCCTCTTCTGAAGTTTGATTATCT 5266
 L L N F I R S T C K K L P L L K F D Y L

CTCAAGCTGCTCAAGTCAAGTAAAACAGAATGCAATTGGAGAAATCTATGTTTTCTCAGA 5206
 S S C S S Q V K Q N A I G E I Y V F S D

TGTAGTTAGATGGAAAAAGTCTCCCATTTAATTTTGAACCGACCCCCTCCCATGGTATGA 5146
 V V R W K K S P I *

TGTAAGTGGCAGACGAAAGTACCCTTTTAAAAGGATTGTGAAATAAAATGATAAAAAAAAA 5086
 AAAAAACACACATCATAAGTAACCGATTAAACAAGTCACAAATAGGTTAAAACTCAATAGA 5026
AGAAAAACAGACATACAAACCTAACTCAACCAAGGATAATCAAAAAGGTAGATTTAAAATGA 4966
GAAAACGCGCAAAAATTGACTTATTTAAAACAATTTTCCAAAAAAAAAATGATATAAAC 4906

FIGURE S3.– Nucleotide sequence of the *BNR-like* element identified in this study. Numbers of nucleotide residues are coordinated with those of BAC clone 33E19. Each pair of palindromic sequences is indicated by single, double, or dotted lines. Putative domains are boxed and shown with pfam descriptions.