

1 accaatgcgctgaacgtccccatcttacctaagagagaaccccctcctacgactggaagtgcctccgaact
71 gatatatgacaatatctacttttggaatcagctgctcgggGAGAGAGACTAAGACGGATAACGCGTCATCTC
141 GCCTTCCCAAATTTTACCCCTCGCTAGACCTGGTCCAAAACCTCCATCTGGAGCCGGCAGAAGAGAACG
211 ATGTTTAACTCGGTCAACCTGGGCAACTTCTGCTCGCCGTCGGCAAGGAGAGGGGGCGCCGATTT
M F N S V N L G N F C S P S R K E R G A D F
277 GGCGAGCGAGGGAGCTGCGCCTCCAACCTCTATCTGCCAGTTGCACTTACTACGTGCCCGAGTT
G E R G S C A S N L Y L P S C T Y Y V P E F
343 TCCACGGTCTCCTCCTTCCCTGCCCCAGCCCCCTCTCGTCAGATCTCCTATCCCTACCCGGCCCAA
S T V S S F L P Q A P S R Q I S Y P Y P A Q
409 GTGCCCCCGGTCCGAGAGGTCTCCTACGGCCTGGAGCCCTCCGGCAAGTGGCACCCCGGAACAG
V P P V R E V S Y G L E P S G K W H P R N S
475 TACTCCTCCTGCTATGCGGCGGGCAGGACTCATGCACCGGGAGTGCCTGCCTCCTTCCACCGTC
Y S S C Y A A A D E L M H R E C L P P S T V
541 ACCGAGATCCTCATGAAAAACGAAGGCTCCTACGGCGGCCACCACCACCCAGCGCCCCGACGCC
T E I L M K N E G S Y G G H H H P S A P H A
607 GCCCCGCGGGCTTCTACTCCTCAGTCAACAAGAACAGTGTCTGCCCAAGCCTTTGACCGCTTC
A P A G F Y S S V N K N S V L P Q A F D R F
673 TTTGACAACGCCTACTGCGGCGGCAGCGACCCGCCCGGAGCCCCCTGCCCGGGCAAGGGTGAG
F D N A Y C G G S D P P A E P P C P G K G E
739 GCCAAAGGGGAGCCTGAGGCGCCCCGGCCTCGGGACTGGCGTCCCGGGCTGAGGCGGGGGCCGAG
A K G E P E A P P A S G L A S R A E A G A E
805 GCCGAGGCCGAAGAGGAGAACACGAATCCAGCTCGTCCGGTTCAGCCACTCGGCGGCCAAGGAG
A E A E E E N T N P S S S G S A H S A A K E
871 CCGGCCAAGGGAGCCGCCCAAGtaggtagcagcggcgggaagcagggcgggaggaggaggagg
P A K G A A P N
937 ctagagaggggagggcagcagggtaggggggagggcagggggagcggggacggcctcgtgttttgggtca
1007 gtccaattttatgtggagttttataagcattcaaaggattttatataacctaacaagccctctctccca
1077 acgactcgaactttttacgatggagaaggggtggggaggggagggaaaagggctctttggaaaagccgggtg
1147 aacccccctctcggttatctcctctcgggtctgtgaaatttttaaaagcggccaccatcggcgggcaatataa
1217 ctttgatcgtgaacttagaggagcatttaaggaaaggtggggagccccggcgggggttgggggtggga
1287 gggaggggaggggtggagggggagaaaggggagggcaagggaaagacaggggagaaactccgagaggggga
1357 gtgggtggaagagagagggcattggagcagagctggagcccggggagtcagccagggccgctctgcttctt
1427 ttaaaaaaactattttttgaaatgttcaaacctctgtgttggcgggtccctgagcaggggctcggagcgacc
1497 cgaggggtcagcggcgacagggggtggggggcagggcgggctcagggctcccctctgctccagggcctggggc
1567 ttggcgtttcccctggggcccgggcctccggcggcggcctgcccaggccgggggaagggagcggcccgccg
1637 ggcgtccccggcgccagggcccccgggaagcggctgtcgggtgtctctggtgctctctctctctctctct
1707 ccccgcatcccagtgactctcggcggatcgcactctcttggcgcctgtctctcggcgggtccccccacattc
1777 tgggggtggggagtggggggtgggttttagagccaccgggagggcagaggtgcgcggtgagcggggcggg
1847 gggccttggggacgcacttggcccgggctggcgcgcatctcccgaagaccggccacagccagcggccctc
1917 gccggaggggagggggcagggcccccgggtggccgggagcgtgggacgcgcttgcccggcaggtcggg
1987 ggtcggcggctcggcccagcggcggcttggcggagggagaccggcggcctggtgggagggctgcccggc
2057 ggtggggccggcagagggggcccgcgagcgtcggcgggggggctgggggtcggcaggggtctcaagctgtctct

2127 cccctctcctcgcacttgccccctcccctcccctccagACGCCCCCGCACCCGCAAGAAGCGC
A P R T R K K R
2192 TGCCCTTATTTGAAATTCAGATCCGGAACTGGAGCGAGAGTTTTTTTTTCAACGTGTATATCAAC
C P Y L K F Q I R E L E R E F F F N V Y I N
2258 AAAGAGAAGCGGTCAGCTGTTTCGGATGCTGAACCTGACGGACCGACAAGTGA AAAATTTGGTTT
K E K R V Q L F R M L N L T D R Q V K N W F
2324 CAGAACAGAAGGATGAAAGAAAAAATAAGCAGAGACCGGCTGCAGTATTTCTCGGGAAATCCT
Q N R R M K E K K I S R D R L Q Y F S G N P
2390 TTGCTGTAACCGCAGACTGGGCCCTTTTGGAGGGGAGGGGGAAAGGGGGAAATTATTTTATTTTAT
L L * 304
2456 TTTTATTTTTTATTTTCTAACTCGCCTACTTTCCGCGGGTGGAAAACCTGGACTGTGGCCAGGGCTGGCCC
2526 CCACTGCCGTTTCCCGCACGTCTTTCTGGACAACCTCCTTCACTTGGTCTGACTccatgtgggacaggg
2596 acagggcctggaatgggggatgaaggaaagtgtctgacccacggcagtggaacccgttggcgcgagggc
2666 caagactctttatataaaagaaaacacacctcggcagccatgtcttgctgctcgaattgggtgggggaggg
2736 gcgagagtagtgggcgcctgaactgaacaactccttgaactaaaagccttccttgcccaagtgaacaagt
2806 ctgctaggacaccgggtctgagggggagacttctgggctccccatcccataccccatgctccatg
2876 gtagggggagcgtaaaatgtgaccccgcccttggctgtgaaatttctgttcccttgacctggtgttaggtgtg
2946 caaactcgtgtcctacctcgtctacgcccagggccaccaccggcctagctgttctccccttgaaggtgt
3016 agaaccctctttagaattttcttaactggtgctcagaggtttcctaactcttttctagcccgcccccac
3086 ttcaatgaatattatagctatagctatgcatgtttgtaacctcttttttttggttggttgaaggaaaaaa
3156 taaaaaagattagggtggaagggcaataggggagatgggattggggcaccctcccatagctgctggcctgaa
3226 ttttttgaataaaatccacaagcatttctttctacccctgaggggggtgtgggtaggactggcctggaat
3296 gagattggaatcaccatctccatgtgcgtgagtggtgtacatgtgcaatccccatccggctcgtccc
3366 gtcccagagggattgctgtgactttttttttttttttttttttgtggcaataaagataacttcatt 3430

SUPPLEMENTARY FIG. S3. The nucleotide sequence of *B. bubalis* *Hoc11* gene and its *in silico* translation. Exons (black) and UTRs (purple) are shown in upper-case letters. The intron (blue) and 5' upstream and 3' downstream regions (green) are in lower-case letters. The deduced amino acid sequence of 304 residues is shown under the nucleotide sequence in a single-letter code. The triangles indicate exon-intron boundaries.