

Supplementary Material

The Histone Chaperone Network is highly conserved in *Physarum polycephalum*

Axel Poulet^{1*}, Ellyn Rousselot^{2*}, Stéphane Téletchéa², Céline Noirot³, Yannick Jacob¹, Josien van Wolfswinkel¹, Christophe Thiriet⁴ and Céline Duc^{2#}

¹Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences, Yale University, New Haven, CT, 06511, USA

²Nantes Université, Faculté des Sciences et Techniques, CNRS, US2B, UMR 6286, F-44000 Nantes, France

³INRAE, UR 875 Unité de Mathématique et Informatique Appliquées, Genotoul Bioinfo Auzeville, 31326 Castanet-Tolosan, France.

⁴Université Rennes 1, CNRS, IGDR (Institut de génétique et développement de Rennes) - UMR 6290, 35043 Rennes, France

*Contributed equally.

#Corresponding author : celine.duc@univ-nantes.fr

1. Analysis of chaperone transcripts

For each chaperone transcript, we provided below the sequence. The transcript name corresponds to the transcript ID followed by the protein ID. The coding sequence is displayed in upper case in bold italics while 5'UTR and 3'UTR are in lower case.

>Phypoly_transcript_17984_PpASF1

tgtcgtgttcagtgagg**ATGGGTTGTGTCAACGTCACTAACATCCAGATACTGGATAATCCAGCCCCTTTCATGA**
ATCCCTTGCAATTTGAAATCACGTACGAGTGCATCACAGACTTACAAGAAGATCTTGAGTGGAAGGTTGTGTATG
TTGGAAGCGCAGAAAGTGAAGCACATGACCAAGTGTGGATTCTATTCTTTTGGGACCTGTACGTGTAGGAGTGT
CAAATTTTGTTTTACAAGTTGACCCTCCAGATTCTTCTAGAATTCCAAAAAAGACTTACTGGGGGTTACTGCGA
TATTACTAATGGGACTGTACAAATCACGGGACTTTATACGGGTGGGCTATTTTGTGAACAACGAGTATGCGGAAG
AGGAGCTAAGACTCAATCCTCCAGAGGGAGAGGTAGACTTAAACAAAAGTGGTACGAAATATACTCGCAGAAAAGC
CCCGCGTAACCCGTTTCCCCATCCCATGGGACCAAGAAACCATCATTCCCCCTGCAAAACACAGAAAATTCCTCCT
CCCTCCCCCGCATCACAACCGAAACAGCCTTGATGCCCGCTCAAAACACAGCATGACGGAAGAGATGATGCCTA
TGGAAGTCATGTCTACTTCAACTTAAGgcggtcaaaattgtaaaataccataaggttgaagcacaagagctgggta
gaacttggttaaaattacccaaatacgaaaatacgcaaaaaattataggcggatttggcctagaaaaactggataa
actagttatccgtgcatttctaggttttgacgaaa

>Phypoly_transcript_01999_PpMCM2

ttgcgacgcgcaagaattggcgccaaacctacaccacacaagacctgccatacaaaacatccccctacttccctac
caacacaaaccaccttcccatcctactcagtc**ATGAGTGATACCAACAGCAACCACAACTCTGACGAGGAGGGCG**
AGGAGGAGTATGACCCCTCCATCCCGAAGGACAAGGAGGCAGAGGAGGATGTGGAGTCCGAGACCTGTCCGATG
GCAACCAGGAGGCTATGGATGATATTTTGGATGGTGAGGAGAATGTAGTTTATGACCTGAAGAATTGGAGGCTA
TCAACCAGGGCGAGGATTTGTACAATGGTGATTTCGAACAGGATTACGAGTATGACCCCAACTTGGATCGATATG
ATGAAAAAGATTTGGATGAAAACGAATATGCTGAAATGGGTCCAGAGGAACGAGCAGCAGCAGAAGAAGCTATTA
AGAAACGGAATCGTAAAGTGCAAAAAGCAAAAAGCTGGAAGACTTCCTAGTGCTCTACTCTATGATACCGACGAAG
AAGAAGATGCCAAGCCCGAATACAAACCGAGGAGGAAAAGGCCACGAGTCAACGACCTAATGTTGGAATGTGT
TTGTTGAAGAAGAGGATACAAGAGATTTGGGTGATTTGGATGTTAATTTGGAAGATGTAAAGGTCCCTTGCGCG
AATGGATTAGCTTAGATCCCACAAAGAGGACAATTATGCGTCGATTGAGGAAATTCCTAAAAACATATGGTCCTG
CCAAAGCCCCCTTTTATCAACAGGCCATTCAAACAATGTGTGCGGCAAATCAAGAAAGTTTGTGGTGAACCACA
ATCACCTGACACAGACGGAGCCCTTGATAGCCATTTGGACGGCCGACGCGCTACTGAAATGATCCAAATTTTTG
ATGAGGTTGCCATGAACTCGTTTTGGAATGTTTCCCAATTACGAAAACATATGTCCAGAAGTACACGTTCTGTA
TAACCAACTTGCCGGTGAGTGATTGCTGCGCGAGTTGCGTCAAATTCATTTAAATGTTTGGTAAAGGTTTCTG
GTGTGGTCACTCGCAGAACATCCGTTTTCCCTCAACTTAAATATGTAAATTCGACTGCTTGAAATGTTCAACGA
TCCTTGGGCCATATTACCAGGACGGACAAGCGGAGATCAAGGTGCGAATATGCCCAATTGTCAAAGTAAGGGCC
CGTTCGCCATCAATGCGGAGCAAACTGTGTATCAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAGATCAAAAGAGGTGATACTTTTGTGGGACCTAATCGACAGCGCGCTCTGGAG
AGGAAGTAGAAGTAACAGGCATTTACCGCCACAACCTTCGACGAGCACTTGAACAAAACACACGGATTTCCGGTTT
TCGCGACTATCATCGAGGCAATTATATATCAAAAAAGACGACTTGCTAGCGTCGTTCCGCCTTACTGAGGCCG
ATGAGAGAGAGATTAGGCGACTCGCCAGAGACGAGCGTATAGGAGAGAAAAATTGTACAAAGTATTGCTCCTTCTA
TTTATGGGCATGATGACATAAAATTAGCACTCGCGTTATCCCTATTTCGGAGGACAATCAAAAAACATTGGTGACA
AACATAGGATTGCTGGGGACATCAACATCCTCTTGCTGGGCGATCCTGGGGTCGCCAAAAAGTCAGTTTTTGAAGT
ACATTGAGAAAACCTGCCACAGGGCTGTGTACACTACCGGTGAGGGTGCCAGTGCGGTAGGTTTGCAGCTGCCG
TGAGGATTGATCCCGTGACTCGTGAGTGGGTTTTAGAGGGGGGTGCGCTTGTGCTTGCCGATAGAGGAGTGTGCC
TTATCGACGAGTTGCATAAAATGAATGATAAGGATCGTACTTCCATCCACGAAGCTATGGAACAGCAATCCATTT
CTATATCAAAAGCAGGAATCGTAACGACACTAATGGCGCGCTGCGCAGTGATCGCAGCAGCCAATCCCATAAAAG
GACGATACGACATCAGCATTCCCTTCTCGCAAAACGTAGAGCTCACTGACCCCATCCTTTCCCGTTTTGATATTA
TGTGTGTTGTTAGGGATATTGTGGATCCAGTAAAAAGACCAGCTCCTTTTCGCGATTGCTGCTGCAGAGCCACATCC
GCTCCACCCGATCGTGAAAGCGCAGGCGTCTTTAGCGAACGCCCCAAAAAACCTTTTCCGCCCTCCCCCACCT
CTTCCGCCTCTTTTCGATGCACCCCTCAGGACCCATGGAGGATGGATCACAATCCCAGTCGCAATCTCAATCGCAAG
ACCTCCCCCTCCCAACCAGATCCACCTATACCGCCAACCACATGCCCGAAAAAGACCAACAAAAACATTTTATCTC
AAGAACTCTTACGGAATACATTCTATACAGCAAAACAAAAAGTGTTCCTCAAGGTCTTCAGGAATTGATAAGGATA
AAATCAGCAGCCTGTACGCAGACTTGCGTCGCGAGAGCAGCGCTCCTGGGTCTGTCCCGATGACCGTGCGTCATC

TTGAAAGCATCATTGGAATGGCTGAAGCGCATGCAAGAATGCACTTGAGAGAACAAGTGAGAGAGGATGACATAA
ATGTCGCGATCCGCGTGCTGCTAGAAACGTTTCTGCGCAAAAGCTTAATGTGGCTAAAACGATGCGTAAAT
CCTTCCAGAAATACATTACTTACAAGAAGGACTACAACCGATTGCTTTTCCACATACTGACAATGGAAGTGC
AACAAATGAACTACCTCTCCGTGAAGCACCAAATCGACAACCCGGAGCGCGTGGAAGTGATGTGCCAAGACTTTG
AAACAAAAGCGCGCAACTTAATATTACAGATTTTTATCCGTTCTACAACCTCAGATATGTTCCACAGTAATTTCA
TTTTAGATAGAAGGGAAGAGAAAATCATCAAAACATTTTCATAAacatcaatacgttttcagtacgtaattgattc
cttgtaaaaatattattaaatttttghtaat

>Phypoly_transcript_03991_PpHSP90

aaactacctgaaaaacttgacatcactccacaaaaaacaggcacgcaggacctgagagtcttctgctggttcttctc
aaacacattgcatcgctctgaagtagcttcgatctaccaaaaccaacttatatacacatacacacaaccATGTCTGAAG
TCGAAACATACACGTTCCAGGCTGAGATCAACCGCTTATGAGCTTGATCATCAACACTTTCTACTCCAACAAAG
AGATTTTCCTTCGTGAACCTCATTCCAATGCATCTGATGCTTTGGACAAGATCCGCTATGAGTCCCTTACTGACA
AGAGCAAATTGGAGGGTGCTCCCGAGCTCTTCATCCACATTGTCCCGACAAGGAAAACAAGACACTCACCTCA
TCGATACCGGTATCGGTATGACCAAGGCAGACCTCATCAACAACCTTGGTACAATTGCACGCTCTGGAACCAAGA
ACTTCATGGAGCACTTGCAATCCGGATCAGCTGATATTAGCATGATTGGACAATTCCGTGTGGGTTTCTACTCTG
CATACTTGGTCGCGGACAAGGTTACCGTGACCAGCAAACACAACGATGACGAACAATACATTTGGGAGTCCAGCG
CCGGTGGTGAGTTCTTCATCCGCCGCGACACCGGAGAACAGCTCGGCCGTGGTACCAAGATCATCTCCACCTCA
AGGACGACCAACTCGAGTACCTCGAGGAGAAGAGAATCAAGGACTTAGTCAAGAAGCACTCCGAGTTCATCCAGT
ACCCCATCTCTCTCTGGGAGACCAAGGAGACTGAGAAGGAGGTCTCTGACGATGAGGAAGAGGTCAAGGAGGGAG
AGAAGCCCACCATCGAGGAAGTTGATGAGGACAAAGAGGCCGACAAGAAGAAAAAGAAGGTCAAGGAAGTCA
CACACGAGTGGGTCTCTCTCAACAAGCAAAAGCCCATCTGGACACGCAACCCCGCCGACATCACCAAGGAGGAGT
ACGCCGCTTTCTACAAGAGCATCTCCAACGACTGGGAGGAGCACCTCGCTGTCAAGCACTTCTCCGTTGAGGGTC
AACTCGAGTTCGCGCCCATCTCTTCGTGCCCAAGCGCGCCCCCTTCGACATGTTTCGAGTCCCGCAAGAAGCTCA
ACAACATCAAGCTCTACGTGCGTGCGGTGTTTCATCATGGACGACTGCAAGGACTTGATCCCCGAGTGGCTCAACT
TCATCAAGGGTATCGTCTGACTCTGAGGATCTCCCCCTCAACATTTCCCGTGAGACCCTCCAACAGAAACAAGATCC
TCAAGGTCAATCAAGAAGAACTTGGTCAAGAAGTGATCGAGCTCTTCTTCGAGATCGCCGAGAACGCTGAGGACT
TCAAGAAGTTCTACGAGGCCTTCGGCAAGAACCTCAAGCTCGGCATCCACGAGGACAGCCAGAACC CGCCCAAGC
TCGCCGAGCTCTCCGCTACAACCTCCACCAAGTCCGGAGACGAGCCACCAGCCTCAAGGACTACATCACCCGCA
TGCCCGAGTCCGAGGAAGGACATCTACTACATCACCGGTGAGAGCAAGAAGGCCGTGAGAACTCCCCCTTCTCTCG
AGGCGCTCAAGAAGAAGGGATACGAGGTGCTCTTCTCGTGACCCCATCGACGAGTACTGCGTGACGCAACTCA
AGGAGTTTCGAGGGCAAGAAGCTCGTGTGCATCACCAAGGAGGGCTTGAAAGCTCGATGAGACCGAGGAGGAGAAGA
AGAAGGCCGAGGAAGAGAAGGCCAAGAACGAGCACCTCCTGAAGCTCGTCAAGGACGTCTCGCGGACAAGGTCTG
AGAAGGTGCTGTTGTCCAACCGCATCGTCTGACTCCCCATGCGTCTTGGTCACTGGAGAGTACGGATGGTCCGCTA
ACATGGAGAGAATCATGAAGGCACAAGCCCTCCGCGACTCCAGCATGAGCACTTACATGCTCTCCAAGAAGACCC
TCGAGCTCAACCCCAACCAACCCCATCGTTCATCGAGCTCCGCAAGCGCACTGACGCCGACAAATCCGATAAGACCG
TGAAGGACCTCGTCTGGCTGCTTTTCGAGACCGCCCTCCTTTCTCCGTTTCTCCCTGGAAGAGCCCAACTCTT
TCGCCAACCGCATTCACCGCATGATCAAACCTCGGACTCAACGTTGATGACACAACATCCGAAGAGCCCAACCG
AAGACCTCCCACCCCTTGAAGGAGAGGCCGAGGCTTCCAAGATGGAGGAAGTGGACTAAaaaaacgcgatcctct
ttgtaacgatagcaaaaaaattataaaaaatgatgaaggaagag

>Phypoly_transcript_04582_PpHSC70

gggggtggtacggttacccctggtcttacgcaagacttgagatcaatcagttcgtacacacaATGAGCAAGACAG
CACCCGCTATTGGAATCGACTTGGAACCACTACTCCTGTGTGGGTGTGTGGCAAAATGATAGAGTGGAAATCA
TTGCCAACGACCAGGGTAACCGTACCACCCCTCCTATGTTGCCCTTCACTGACCAAGAGCGTCTTATTGGTGATG
CCGCAAGAACAAGTGGCCCTCAACCCCGAGAACACAGTGTTTCGATGCTAAGCGTCTCATTGGTTCGAGGTTCT
CCGACCCCGTCTGCCAAGACGACATGAAGCACTGGCCCTTCAAGGTCAATTGCCAAGGAGGGAGACAAGCCCCACA
TCAGAGTCACATTCAAGGGAGAGCAGAAGGACTTCTCCCCGAAGAAATCAGTGCGATGGTGTGATTAAGATGA
AGGAGACCGCCGAGGCTTTCTCGGAGCCACAGTCACCAACGCTGTGATCACAGTGCCCGCTATTTCATGACA
GCCAGCGCCAGGCCACCAAGGATGCTGGAACCATTTCTGGTTTGAATGTGCTCCGTATCATTAACGAGCCCACTG
CTGCCGCTATTGCCTATGGTTTTGGACAAGACTGCCGAGGCCGAGAAGAAGCTGCTCATCTTCGATTTGGGAGGTG
GTACATTTGATGTGTCCCTCCTTACCATCGAGAGCGGTGTGTTTGGAGTCAAGTCCACTGGTGGAACACCCATC
TTGGAGGTGAGGACTTCGATAACCGCCTCGTGAACCACTTCATGGAGGAGTTCAAGCGCAAGTTCAAGAAGGACA
TCACCGGAAACAGAGAGCCATCCGCCGTCTCAGAACCGCCTGCGAGCGTGCGAAGAGAACCTTGTCTTCCAGCA
CCCAGGCCTCCATCGAAATCGACTCCCTCCACGAGGGTGTGACTTCTATACCAACATCACCAAGAGCCAGGTTCCG
AGGAGCTCTGCATTGACTTGTTCGCTGGCTGCTTGGAGCCCGTGAGAAAGGTGCTCCGCGACGCCAAGCTCAGCA
AGAGCCAAGTCCACGAGATTGTGTTGGTGGTGGATCTACCCGTATCCCCAAGGTCCAGCAACTCTTGCAAGACT

TCTTCAATGGAAGGAGCTCAACAAGTCCATCAACCCCGATGAGGCCGTGCTTACGGAGCTGCCGTGCAAGCTG
 CCATTCTACCGGAGAGGGAGGCAAGAAGGTCCAGGATCTCTTGCTTCTCGATGTGACCCCCCTCACCCCTTGGTA
 TCGAGACTGCCGGAGGAGTGATGACTTCCCTCATCTCCCGTAACTCCACCATTCCCTTGCAAGAAGGCCAGGTGT
 TCTCCACCTACTCCGACAACCAACCTGGTGTGCTCATCCAAGTATATGAGGGTGAGCGTTCCATGACCCGCGACA
 ACAACTTGCTCGGAAAGTTTCGAGCTCTCCGGAATCCCCCTGCGCCCCGTGGTGTTCCTCAAATCGAAGTGACCT
 TCGATATCGATGCCAACGGTATCCTCAACGTGTCTGCTGAGGACAAGTCCACCGGAAAGAAGGACAAGATCACCA
 TCACCAACGAGAAGGGACGTCTCAGCAAGGACCAGATCGAGAAGATGGTCAAGGAGGCCGAGAACTACAAGAAGC
 AAGACGAGGAGGCCAAGGACAGAGTTGAGTCCAAGAACAAGCTCGAGAACTACGCCTACTCCGTGCGCAACACCA
 CCAAAGAGGACAACCTCAAGGACAAGCTCAGCGCCGAGGACAAGGCTGCCATCGACAAGGCGGTTGATGAGACCA
 TCAAGTGGCTCGACGCCAACCAGAGCGCCGCCAAGGAGGAGTTGACCACAAGACGAGCGAGCTCGAGAAGGTCTG
 TGAACCCCATCATGACCAAGCTCTACCAATCCGCCGGCGGAGCTCCCGGAGGAGCGCCGGAGGTTTCCCCGGAG
 GTGCTGGCGGATTCCCCGGAGGTGCCGGAGGTTTCCAGGCGGAGGTGCCCCAGGAGGCGGTGCTTCCAGCGCTG
 GTGCCGGCCCAACCGTTGAGGAAGTCGACTAAAtcgagactcgaaattttgaaaaataaaacccttgtaaaata
 gcaaaacaaaaaagagaatagtcaacaaaaaaataaaacccttgtaaaat

>Phypoly_transcript_01375_PpIPO4

cgcacacacgagcaccgcgttttcgtttgcctctgtgctgccATGGGGGTACTGTGGAGCAATTGACACACTTATG
 TGAGCAGGCCCTCCATCCCAACACAGAGATTGTTTCGTGCTGCCGTACCCAGCTCAATGCCGCGCTCACCGACCC
 TCAATCAACCATGCTTTTCTTGGCTATTGCAACAATCCCCCATACACAGTAAAGCAATTTGCAGCAGTGGT
 GATGAGGAAGCGAGTGGTGGGCCACTACTACAAGCTGCCTCAAAATGAGAGAGGGGTGTTAAAGGCCTCCCTCCT
 TTTGATGCTGCATTCTGACACAAAAGTCGAGGTTATTGTAAGGAGGGCAGTGGCAGACGTTGTGGTGACACTTGG
 TAAAGCTATTGTGCCAAGAGGTGACTGGAATGTTCTGTTACCCACTATCCTCAATGCTGCTTCTAGTGCTGTTTT
 GATTGAGGACAGAGAGATAGCAATGTATTTGTTAGCAGCAGTAATAGACTCCCTTGGGGACCAGATGAGAGAACA
 CTATGCATCATTGCACACCCTCCTGGCTGCCTGCTTACGAGACAGCAACCCCTTCTATCCGTTTATTGGCTCTCCA
 CACCACAAGTGCTTTAATTCCATGGCTGGTAACCAAAGATGATATAGCACTATATGGATCTCTAATACCATTGGT
 GGTGATGTTGTAAATATTTTGTACAAATGAAATGGATGATGATGCTATTGAGTCTTTCTCAATTTTAGAGGA
 TCTTGTAGAGACACCCATACCTGTGATTACAAATTCATACCAGCTTTGGTGAACTCATGCTTGAAATATGTGC
 TAACCATCAAGTTGATATTAGTGTCCGACAAAGTGCTCTGTCAATTTGTCCAACCTTGTGTGCGGGTGAAACCTAA
 GATGGTTGTGAAGAACAATTTGGTCACTCCAACACTCGCAGCAGCAATTTGCCATGTGTGCAGAGGGAGAGGGTAA
 CGATGAGGCTCAACAGGATCTCTATGAACAAGAGGAGACACTGTCCCCATACATCTTTGGCTCACAAGTGATTGA
 TACTGTGGCACTCTCACTACCTGTGAACCAGGTGATAGGCCCTTCGCTCGCACTAGTCGCCCCAATATGGCAGGCC
 TGAGATGCACTTCCGCCAGCGTCGTGCTGCCATTACTGCAATTACTGTGTTGACCGAAGGCTGTGCGGAGTCAAT
 CAAGGACACTCTGGCTAAGGTGCTACGTACGTGGTGGGGGCACTAGTAGATCCCTCGTTTCTGGTGCGCCAGGC
 AGCATGCTTGGCCATAGGGCAATGTGCAGATAACTTACAACCTGAAATCATTGACTACCATAATGTAGTGGTTCC
 TGCTTTATTGCAAGGATTACGAGATCCCATGGAATCGGTTGTGCTCCGCTGCTGCTACGCCCTCGAGGCGTTCTG
 CGAGTCCATGGGCGACAAATTTGTGCCATATCTAGCCACGTTTCATGGAATACATGGGCGAGGTCATGCGCACGAA
 GGGCATAGAGACGCAGGAGGTGGCCGTCAAGTGCGATTTCGTCCGCGACTGTGCGCCGCGGTCTCTGTTCTGTTACC
 CTATTTTCGGTGCAAGTTTACGAGCTGATGCGAACTATGATGATGGCTACCGGCTTAGGTGCTACGGAGCCAGAAAA
 GATCATAGTGGTGCGAGCCAGGGCGACGGAGTGCATAGGGTTGTGCGCTCTGGCAGTGGGAGAGCCGGCGAAAGG
 GATAGTGCATGACGCGATGTCTCTTGCAATTGGCAGGTCTGGAACCTTGGTTTTTCGGAATTGAATGAGTTTACCTT
 TGGATTTTTTTGCAACTTTTTCGAATTGCTAGGCGCCGATTTTGTGAACTACTTACCAGTTGTGATGCCGCACAT
 ATGGAAGCCTGCGAATCTTCGGATGGAATATTCCCTACAGCAAAAACCAACTAGAAGTAGAGGGCGATGACGA
 AGAGCCTCTTGATGTTAGCGTGCGCACTTCCTTCCTTTACGAAAAAGTGCTGCTTGTGAGGCTATCGGGTCTTT
 TGCAATACATGTCAAGGTTGGCTTTATACCATTGTAGACCAGTCAATAGCAGTATTGAATCCTTTGCTTAATTA
 TTTCCACGAGACTGTTAGAGCTGCTGCCATAGTTTCTATCCAACAGATCGCGGTTCTATTACACAGCTTTTCC
 TCCACCTACGCCTTTTGTACCTGGCGATTTGACCATCCCTCTGTGACGCAGAGCAAGGAGTACAGCACACGAGT
 GATGCCCGTATTAATCTCGGCTATGCGTGATAAATCGTCTGATGTAGTTGTGCGCGCCAGTGACTCGATTGGTGA
 AATTGCTCAGCTGTTAGGTCCTGTTGCTGTGCGCGAATTCATGCCAGCAATTGTAGAAGAATTTTCATCCATTTT
 GCAACAAAAGTCTCCTTGCCAAATCCATTTTACTGAAGGTGATGAACTGGATGATGAAGGCAATTCAGAGGATGA
 AGAGGGAGGTCCGGCAGAAGAAGAAGAAGAGGATGACGATGCACGCTTCTTACACTACTCGATAATGTGACCGA
 GTGTATGGTGGATCTTGCACATGCTCTTGGGTACAGTTTACCACAGCCATGCAAGCATTCTTCCACAACATCAT
 GAAATACGCACGCAACAGTGCACATCACTCACTGCGTGCTACTTCCATCGGTGCAATGGCTGAGACCATCAAAGG
 CGCACATACCCTCCCTCCCCAGACCTCACCTCTCAATTGTTCCAGCTTGGGCTAAATGGTATGACCAACCGGCA
 AGCTGACCCAGATATTAGCGCACGGCTGCGTTTTTGTGTGGTGTGTGATTGAGATGGGCGGAAACAACATGGA
 AACGCAGATTACAAATGCACTCCAAGCCCTGCACCCCTGTTTGGCAGTCCAGATCCGACTGTGATTGATAATGC
 ATGTGGTGCAGTAGCCGAATTATACTCACATTCCCCGATTACCCACATGTGCCCATTTGAATTCATCCTTCCCAC
 CCTTCTCCCTGCACTCCCCCTCAGAAAAGACCACCAAGAAGATGAACCTGTGTGGAAATGCGTTGCGGAGCTAAC

ACGGTCTGCAAAAGTCAATGCACAGGGGGAAACGACAGTTGCCCTCCTAGAGACAATACTCCGTGGATTGGAAAA
TCCCCACTACAACCCCAACTCAAAGATTTACTGAAGCATGCACTTGGTGGATGGTTCCCCACATTGCAAGCACT
CTCCAGTTCTACCCCTTCCTTAGTCCCTCTGTTGCACACTGTCACCAATAAAAAACCTGTAAattacaacccatct
cctaatatcccaata

>Phypoly_transcript_00177_PpSPT6

ctcctgcgaagcacacataaaaagctggttgtttgtgatttcaccagcattcagttgctgagcatcgctggttgaca
caagacATGTCTAGCAATGAATATGTGGAAGAGGAGCAAGACACCCCTTGTAGTGCCAGAAGGTGAGGCCGCCGCG
CCAGAGGGAGCGGAAGAAGGATATCAGGAGGGACAAGCGGAAGGTCAAGAGGAAACACAAGAGGGGAGGTGGCa
gagcaggaggaaggcgacagaagaggaaggcgaaggggaatcaagagaaaagaagaagaagaaaaagaagaaaaag
aGGAGCGCTTTTATAGAAGATGAGGCGATACTCAGCGGAAGTGAGGTGACAGCAGCGAAGAAGAGGACGGAGAC
TCGCAGAACGAGTACGAGTACGACGACTTCACGGTGCCACCGGGATGgagaggaggaggggagagaggaggaggga
ggaggagagaggaggaggaggaggaagaagaagaggagggaatgaggagagactggaggaggaTGACTATGATTTA
GTCAGAGAGAACACAGGACAAAACATAGCCAGGAGGAAACAGGATTCTGCAGGGGCAGGAGGGGGGAACAAAAAC
TCAAAGAACTAAAGAGGCTGAAGCAGATGGGAAACAGGACAGACTTGTACAGTAATATCCAGAGCAGGCTTTTT
GGTGATGATGATGACGAAGCGGGGCCATCAGAAGCCATCACCCACGAGGAACCGCGATCGGATGACGAACAGAT
GACGAGGACGAGGACGATTTTCATCGTGGGCGCGGACGGGAAGCCCATCAAGAGGTGGAACAGAGGAGGTACAGG
GGAGGAGAGGACAGGGGAGCCCTGCAGGACGCAGAAGACATTTTTGGCAGTGTGATTTTTTCAAACAAGCATCA
GGCCGGCGTTCCTCGTCTCGGGGCGATGAGTTTCGACGACGATGGTGACTATGTGACGGGCACGCAGGGCCCTGAGC
GAGGATGAGCTGGACGCAGAGATAGAGGAGGGGGACGAGAAGCAAAGGACCCCCCTCGCGGACATCTACGAGCAC
TACGAGCCGAGCCTGCTCGAGCAGAAGCACTTCACCTCGGAGGACGAGAAAAATTCGCCAGTTGGACGTTCCCGAG
CGATTGCAGATGCGAAAAGGCACAGCCTACGACGAGAAGCTCGTGCCCGAAGAGGCCGAGTGGAATTTACGCGC
GCCTTCGCGCGCACGCACGCGACCGGATAAGGCAATCATCATCCAGCGCATCATCCACGTCTTGCGCTTCCTGCAA
GAACAGTTGCTGGAGGTCCCCTTTATTAGCACATACAGAAAAGACTATTACCAACCCCTTAAAAATATCTGAACTT
TGGGACATTTATGACTGGGACGAGAAATGGGCTCATTTACAAGCGCGAAAAGAATAACATTATGCAAATGTGCAGG
ATGCTCGAGGGCAACCCCAACACGGCGACGAGCTGCGCGAGTACATGGAGCTCCTGCGCGACGGGCGACGGAG
GAGGACCTGAGCGACCTGCAGGACTACTTCCAGATGTTCTACGCGGACGAGGTGCGCGAGAGCGCCGCGCACGCC
GAATCGGGAAAGGTCAAGCGCGCGCTGAGGAGGGACTTGTACACTACGGCGAAGCGCAACGGGTTGGGCGACTTT
GCAAAGAAATTTCGGAATAACCCCCAAAGAATTTCGGAGACAATCTTATGGACAACACACACACACCTATA
GACAACGAAGAAGACCCCCGAGGCACTTGCGCAAACTACTTGTGCAAAACATTCCCCGACACAGAGTCCGTGCTG
AAAGCGGCGCGAGTCATGTGTGCGCAGGAAATCGCATTCGACATCCACGTGCGCCAGAGTGTGCGCTCAGTATAC
GTGAAACGCGCCATTCTTTCCACCGTGCCAACGCTAAAAGGCAAAAAGGACATTGATGCATTTTCATCCATATATG
ATCATAAAACAAATGGACAAGCCGGCCATGCAGATGGAGGACTTCCAAATGTTACTCGTACTCAAAGCAGAGAAG
GAAGGTTTTATCACAACAAAAATACACACGGACGAGAAAGACGTGCTCCAGGAAATGGACACCTTATATACAAGC
GACGGAACGAGTGCTACCGCAGAAGCGTGGAACAACACCGCCGCGAAATCCTCGCCCTCACGCTGCGGAACCTAC
CTCTTCCCTCTCTTCGCGCGCGAGCTACGCGCGAACTCACGCAGGAAGCGCAGAACTTCGTGCGGTACGACGTG
TCGCGCCGTCTTGAGCAGCAGCTTCTCGTGCGCCGTGGGCTCCTGAGGAGGACGACTCGACCGACGCAGAGGAA
GAGGATGACACGCCGCGCAGGATGCGCATCATCGCAACGTGTGCGGGGGGCTCGGGGAACCCACGATGGCTGTG
GTTTTGGATGACGAGGGAGAGGTTGTAGAGTACCTTAAGCTCAGCTTTTTGAACGAAAGGCGTACAAATGCGGAC
GGACGAGATAGCACCAAGACCTAAAAACCTCGAAAAATTTGTACAGGACCACCAACCTCGAGCTATCATTGTC
GCTGCAGAGAACCTCGAGTGCAGGCGATACTACGAAGATATTATGGCAATTGTCCAAAACTGAGGGGGTCTCGC
GAATTGCGCAAGCAGTGCAAGGTGATGTATGGTGACGCGGAGCTTGCGCGTGCAATTAAGCAATCCCATCGTGGA
ATTTCAGAATTCCCTGAGTATCCTCCACTTTTGCGCCAAGCTATTTGTTAGGACGCAAATGTTGGATCCAGTA
ACAGAATTTGCTAATTCTTGCGATGACGCTAACGGTCACTTTGTGTTACGCATGCACGATTTGCAAGACATGGTG
CCCCAAGACAGGCTCGCGAAGTCAATCCACCGCGTGTTCATCAACGTGCTGAACGCAGTGGGAGCCGACGTGAAC
CGCATGCTCACGCACAAGTACGCGAACGCATCGCTCCAGTACGTTTCCGGTCTGGGCGTGCGCAAGGCCGAGTAC
CTATTGCAAGCTATTTTCAGAAAGGGTCGATCGCTCGCGTCCAGAGACGAGCTCACAAGTTTCGTGGACAAGTGC
GTTTTTCATGAACTGCGCTGGATTCAATAGGATTTCGCGACAAGCATTCTACAACGAAAACGTAGAACTTGAGCCA
CTCGACGATACACGTATACACCCGGAGGACTACAGGCTCGCGCGCAAGATGGCTGCCGATGCGCTGGATCGCGAG
GGCGAGGATGCAGTTTTACAAGAGTGCAATTTACGCGATTATGGAAAAGCCGAAGAAGTTGGATGATATTGATTTG
GATGCTTTTGCTACAATGTTGGAAAGCAGTGGAGAGATCAAAAAGATGACACTGTACGACATCAAGAAAGAGCTC
ACACATCAATTGCGCGACATCCGCACTAAGTACACTGACCCTACTAACGTTGATATCTTCCACTGGCTCACGGGA
GAGACCGACAACCTCGCTGGCGGGCGGAATGATCATCGCGCTCGCGCTTTCGCGCAATCGGAACGTTCCGTTTTG
TGCCGGCTAGAGAGTGGGTTGATCGGATTCATCAATTCGAAAATGTGAGCGATGACAGAGAGGCCGCGTCCAGT
TCTACCATTGCTGTGGGGCAAACCGTTACGCGCGTGTCTCAGCATTGACAAGGACAAGTTTTCTCTAAGGCTC
TCGTGTGCGGCTCGGATCTGGACAACCCGAAGTGGGAGTACGACAAGTACGAAGAATTGAAAAAATTGAACAA

TTCTTGGTTATCGAGTTGGAGCCCAACAAGGACGAGCCCAAGAAAGAGGCCGTGCGCAGGACCACGTGGAACGTG
CACCACCCGCTCTTCCAGGACACGGATTTTCATCGGGGCCGAGAAGTTTTTGGCCGACAAACCGATCGGGGAAGTT
GTAATCCGGCCCTCGTCCAAGGCCCCCGACCACCTCACCTTTTCTTGGAAAGTTTTTTGACGGCGTGTACATTAC
ACTGATGTCCAAGTGAGGGACAAGAAAAGTGCGAGGAGTTTGGGATCACTTTTCATTGGAGAAAATAAATATGAT
GACTTGGACGAAATTGTGGCGAGATACATGGACCCGATAAACTTTTTTGTACTCGAGATGATTAACACCAGAAAG
TTCAAGCGATCTTCCAGCAGGAAGTTGATCAATTGGTCACGTTGAGAAAGGAGCAAGCCCTACATCGATCCCC
TACTACATCACAGCCCCTAAGGAGCACCTTGGACGCTTCGAAATCACTTACATGCCCTCAACGCGCCTTAAGCAC
GAGTTCATCACGGTGAACCCAGAAGGCTTCTCGTTCAGGAAACAGCCCTTCAAGAGCGTGGAGCAGCTGATCAAG
TGGTTCAAGTTACACTACAGGGATCCTATTCTCCACCAAAGAAAGCCGACGAGAAACCCCCACGCCAGCCCCA
ACACCCGCGCAACTGCAGCAGCTGCAGCAAATGCAAACACAGCAGCGCATGTCGACGCCGAGATCAACGCAACA
CCCGCGCACCTTGGCGCAGCGACGCCCCGCGCACCCCGGCAACGCAACACCCGCGCACCCGGGCAACGCGACGCCC
GCGCACCGAGGGCAGCGGATGGGCGAACGCCGCGCAGGGCGGGAACAACCTTCCCGGGGATGCAGGCCATGCAGCAA
AACCTACCTGCCACGTCACTTTTCTTATGGGGCGGGTGGGACCAAAGCGGAACGAGCACGCCGGGGCACTCG
CAGGGAGGGGCAGGAGGGTGAATAACAATAACACAAACAATTACGGAAATAACAATAACGCGTTTCGCTAACCCA
ACGAATGACGTCACGATGGGATGGGAGGACAACATAACAACAATAACAACAATTCCAACAATTACGGAAATAGC
AATTCCAATTACGGACAGAACGACAGTTACGGGAATAACCGGGACGATCGCTCGCGAAGGGGACAATAATTCCAAT
AATTACGGAAATAACAGAGATAATTATGGGAATAATCGGGGGAATAGTTACAGAGATAACAATAACTACAACAGT
AACAGAGATAGCAGTAGGTACGGCGGGGGAGGGTAtgggggaggaggaagagacaactacaggagtggaggagga
ggagggggaggga tacggaggaggaagcggaggggggatacggaggagATAGAAACAGAGACAGCAGGGGAGGAAGA
GATAACTACAACAATAACAGGAGAGACGACAGGGATTACAAcaggaggagcgggggaggagggggagggggatac
ggaggagacaggaacagggaTAGTCGAGATAa tagaggggagagacggaggggagaggggtgggacgatagagataag
gggagtggaggaggaggaggagggggagggaagaagaagcagagggggagagGGTTTCGAAATATCGGCGCAGCAG
GTCTTCCACGGACTCCTCAGCTTTCCCTCGTACCGTTCTGGAAGAGTCTTTTGTAAaaaaagttgggagctgaag
 ttggcctatgtaggtgtcatttggcaaggagctgagcacttcagaggtttcttcgtcataaaactgatgttgaag
 gaaagcatttttgtgagttgttgggaccgaagaaaagccctcaaacactaaaatgggttccaagaaatttcttag
 aggtacctcttcgaggatatctgtgcgaacaactcgtgttggaaatatttttctcattttaatactgcggtc
 cttagaacagcgataaaactggcaaatagcaaccaaacgctcttccgggttcagttcctccgtatcctctacaat
 gagcaatatatttataatagaagggatacatggaacattgagatcaatatatttttgcagttgcaaagtgtcctac
 tgaaaagtggacgagtggttctgcgcgagaagtgagtgcggaaggaatcgtagagattgcaattcttaattgttaaa
 ggtgtggatttgcctcctagtagcgttcggtcgcggtgggggctcctcaccttcgactacgttatcaaaatattg
 gttcgatatgccttcatttatttcttggccttttgccttgaaatgttggtagcaaaggagctacgtccaccatatt
 ggtgtagatctagagaatgaacgtgagaaaagtactgggttaaaagaaccaggagcttaattgatcgtgaatgtgc
 aaacaaatgtggtattaaggtctattttaagtatagtgagaaaaatgtggcgctaaacatatttgcacgtgtg
 gaaagaagaaat

>Phypoly_transcript_03499_PpAPLF

ataatccggtacaaagaacgctctggatggcaacaactcctacacagtagattatgcacccacagcaaaagctg
 gatgcaaattttcacaatgcaaaacaaaaatagaaaagggatcttgcgcattggaaaatcatatccctctgata
 ggtttactgatggagggctctggaaatgaatattatcatcccaagtgtggtttgaggtcaaaagaagacaaaaga
 aaactacaaagaaagtagagagcgaagatgatttacagggattcgacgttatttcgtcgacccgatcaagaactta
 ttctggatcttattaaacataggaatgatccccgcgtacactctccaaagaaaaaagcaaaaaaagacgcaggag
 gcgaggggggagggggaacgccccaaaaagaacaccaaacagacctacatcgacagcgacgaggaaaaaagcacgc
 ccaaaaaaggaaatttctaataatagcagcaccagtaacggaaataacaagaataaggacaacggcaaaaaacaaag
 ggaagcaaatttccggggtcggatggggaggaatcggaggacctggaaaaggacgcgtgggctgctgtggcgctcaa
 acgggagcaggggtggccagagtcctaaatagcacagatatcagacaagtttcagttgtgatactgcctcagcct
 cgagcagggccataaaagcgtgctgcccgcgaagatgatcgcgcttttgaggatgaatttggaggagaccoggatg
 catatctacaatgggttggcaggaactggagtgctatggatcagcacggattttgatgatgcaggctatacatctc
 caggcttgtggaccgatacgtatcgagatcgtgcccagattttcatgcaaaaaccaaagaagggccttgtgg
 tggctatgtggggacctgaacaaaagcgtttccgaaatgagataaagaagcactaccaagacaccatggggacag
 tcaagtaccgagcgctcgagggtcccacgccccctaaactcagccttcaactcgaaagacgagctctcgaaaatcg
 acgacgcacgcaccttattgagc**ATGCACGCGCTGCACTGGTGGCCGTTTCGAGGGCGAGGTCGCGCAGAACCAGG**
AAAAGAAGGAGAAAGAGAAAGAGAAGAAGAAGGGGGTAGTGTGGGGGAGGGAGGGGCGAGGTGGGTGCTGGAGT
GCAGCGAGGAGGGGGACATCAAGGTGGAAGTATGGGCGGTCTACGGAAGTACGAGCGGGGGTTAGTGAACAACA
AGAAGATCAGTCGAAGCAGTGTGTCTGTGGAGACGGTGGATGACGGCGTTGAACTCACTCCCCCTTGGTGTGAATC
CCATGACAATGTGCGCGCGCAGGAAGCGAAGACGTGGAGATCCTCAGCCGAAACGAGAAGTATACATTGCACGAGG
GCGATTTCGTTACCTTATGCGGCGAGCAGTACCGATTTCGTGCTCCAGCGGCCAACGGTCAACCCAGACTACAACGA
GCCAAACTACAAGTCAAAAAGATGGCGAAAAATCGGCGAAAACCCGCTCGCGCCAAGCTTCGAAAAATCGAGGGCG
CGGAAGGCGAGGGCGAGAAAACGAAGAGGGACGGGGAGGCAGATGAGATAAATGAGGGAGATAAAGGGGAGGAGG

AAGGAAGGGAGAAGCAACGAGGGGAGAAGGGAGAGGAAGGGGAGAAGGGAGAGGAGGGATCACCCGAAAAGAAAAG
GAGTGAAACGCAAGGCGAGCGATATCGCCACGACTACTACCACAACCTCCTCGGCCAAAAAAACCCCGCAAGGAT
CCGCAGAAAGCGATTTCGGCTACCGTTCCGTTTAGCGCGGGCTACGATTCCCCAGAGGAGAGCAAGGCAAAAAGCA
AGAAGAAGCCCGACGTCAAAACGGAGGAGGCTGATGACGTAATCGATGACGAAGAGCTCGCGCGCAAACTACAAG
AGGAGCTGAACAAGGAAGACGAGGCGCCAAACCGAAAAACCCGAAATCGAAGAACAGCACCCCCGAAAAAGACG
CAGAGATCGCGAGGAAGTTACAGGAGAAGTACAACAAGGAGAGCGGGGGAGGAGGGGAAGAAGAAGAGGAAGCAGG
AGGACGACGAGGACGATGGCGAGGTGAACTGGGCGCGAAAGTGAAAAAGCAAAAGAAAGGGAAGAAGGGGGGGA
GCGACGACGAGGATGAGGACTGGAACGACGTGGGAGGAGGAGGGGACGAGGACAGCGACGAGGACTACCAAGAGG
AGGAGAGCGACAGCGAAAAAGAACAAGAAAAATAAGAAAAAGGGGGCGAGTCCCAAGAAACCACCATGCAAGTACG
GCAAGTCGTGCTACCGCACCAACCCAGACCACCTCCGGGATTCTCACACCCCAATAAGttatgaacgtatccg
tttatgtaattgtacttaatttttatagttatatgcacctc

>Phypoly_transcript_27931_PpPolE3

Aaattcgaaaaatacgagagctcgaaaattttaaaattttcaaaaaaggcgcggttcttgcgcgggaaacgcaggaaa
gagacaaggATGAGCGGAGACAAGACAGAACCAAAAGAAGAGAAAGAGAAATCAGGAGAGGACCTGGAATTACCT
CGAGCAGTAGTGCAAAGACTAATTAATATTTCATTACCAGACAATGTACATGTACAACAAGAGGCTAAACTAGCT
ATTGCACACGCGGGGAAAGTTTTTATCAATTATTTAACAGCTTGTGCGAATGATTACTGCATACAAGGAACGAGA
TCTACCATTTCAGCTAAAGATGTATTAAATGCATGTGACGAGCTGGAATGCCTGAAATCACTAAAAAATTA
GAAGTTTTTGAAGCATACAAAAAGAGCAATCAGAGAAAAAGAAAAAGGAAAGACAAAGAAAAAGACGAGGGT
AAGTCCGAGGATTAAactttgttatactatgtacaatatataaaatcatttattcaaaaaaaaaaaaaaaaaagagcaa
tcagagaaaaa

>Phypoly_transcript_00085_PpSWR1

ctgaattacgcgaggagcaacatccatagtgttcacgcatctctatcaacatgatgagcatcttctgagccttt
ggcatggaatggagatgccacgtgATGGTGCAGAAGGGAAGGGAAGCAGCAGAGGTGGAGAGCAGGGGGAATAA
AAGACAGAAGGGGGTTGCAGAGCCCCCATTCATGATACACGCAATGCCAACACTAAGAGACCATCTCGCAC TTC
TGAGCAAGCCTCCCCAATCCCCAAGAAGAGAGTGCCTGGGTCTGGTGTGCCTGAATCACCATCAGTAGAAGCAAT
CGATGTGAGCCAAGAAAGTGAAGTTGCACCTAAGGCTGGAGTAGAGACAGACATTAACATTCCAGCCATTTCGGGA
GTATGAAGTAGAGAAAAGGCTAGAATTGATAAAACAACCTCACAGAACAGAAAAGAAATGAACTCAAAGAACTTGC
ATATCTGGAAGCGGGGAACAACCTTGATGGACTATGAAGCCGGCTTCACAAAGGGTAAGATAATGACCGCAAAGGA
GGAGCCAGGGGAGAATGTGGCAGTAGGTGCAGAAGTCCCCAACAAAGAGGGTCTGGACGTCTTCCAAAGGATCC
CAACTCTCAATCAAGGAGGAGATCAACAGTTGGCCAGTCTCCTGTTCAAGGTGAATCACTACAAGCCCAATCACC
AATTCCACAAACTCCCTCAGCGCCACAACTCCAGTACAACCCAGTCAACCTACTCCACAACACCCACCCAACC
TACAGAGCAGATCACAGCACAGAAGCTAAGCCAGCGGTAGTATTTACCAGCCGCTCCAAGTGCCACTTCCATC
CATCTTTACCGTCCCTCCACCTACTTCGGAACAAGACCGTGAGCGCGCACTTAAAGACGCAGCGATCATACGCCG
CGTAGGCGAGCTTCAGCGCACGGGCATGTGGTCCCTCAAGCGTATTCTTAAGGCTGTAGAGCCGCCGCGCGCCAA
GGTCCACTGGGATCACGTGCTCGCGGAGATGGGGTGGCTGGCGAACGATTTCAAGCAGGAGCGCAAGTGGAAGAT
GGCTCTTGCGAAGAAAGTGTCTAAACAAGTCATGAAGTACCACCAGATGCAGGAGACGCGGAGCAGCGCAAGCG
CAAGGAAGAGGAACACCAGATCCGCAAGATCGCGAGTGGCATTGCTCGTGAAGTGAAGAAATTTTGGCAGAAAAAT
CGAGAAGGTGGTCATGTTCAAGGAACAGACCAAAATTGGACGAGAAGAAGAAAGAAAGTGTAGACAAGCACCTGGA
TTTTCTCTGGGACAAACAGAGAAGTACTTCGAGATGATCGCCAACGACCTCAAGCAACCCGAAGGAGGAGCTCA
GAAGGACGCACCAACGACTGAGCAATCCACAGAAGAGAAGGCGGACGGAGAACCAGCCGACGAGAAGCCACTCGT
CCGAAGCCACACGCAAGAGCTGCTGATGGGGGCTGAGGAAACGGGAGGAGAAGAGGGAGCTGAAGATGACGCCGA
ATTGCTTCCACCTGAGGAAACTGTAGATGAGGTGCAGCAGCAGGAGACGATGGAGGAGGAGGAGACCAAGGAGGA
CTCCGCTGACCTCCCCAAGGGTGAGGAGGAGCTCTCCAGCTCCAGCGGAGAGCGAGATGCCACTAGAGGAGAT
CATCAAACAATTCAAGGAACAGGGCGGCGGGGAGGGTGAGCAAGATGACGCCGAGGATGCAGACGCGGATGAAGA
CGACGAAGGGGAGGACGACGATGGGAAAACGGGTCAAGAGCGTATCAGCGAAGCAGCAGCCGACGCAACGCAGC
TCAGCCTACGGGATTTACTGTCCACCACCAAGGTCAAGACGAAAGTCCCTTTTTTACTCCACGGAACCTTGCG
CGAGTATCAACACATTGGTCTCGACTGGATCGTAACCATGTATGAGAAGCGCTTGAACGGAATCTTGGCTGATGA
AATGGGTTTTGGGCAAGACCATCATGACAATCGCCATGCTTGCCCATCTCGCCTGCGAGAAAGGAATCTGGGGTCC
TCATCTAATCGTGGTTCCCAAGCGTGATGCTGAACTGGGAAATGGAATTCAAAAAGTGGTCCCCCGCGTTCAA
GATCCTCACCTACTTCGGAACGCAAAAGGAGCGCAAGCTCAAGCGCCAGGGCTGGTTCGAAGACGGAACCTCTTCCA
TGTTTTGCATCACCTCGTACAAGGTGGTTGTGACAGGATCAACAGGCCTTCCGCCGCAAGAAATGGAAGTATTTTCAT
CCTGGACGAGGCGCACAAATCAAGAACTTCCGGTCCCAGCGCTGGCAAGTCTTCTCAACTTCAACTCCAAGCG
CCGCCTGCTTCTTACAGGAACCCCACTCCAGAACAGTCTGATGGAGCTCTGGTCCCTCATGCATTTCTCATGCC
ACATATGTTCCAGTCCCAAGGAATTCAAAGACTGGTTCTCCAATCCCGTCACCAATATGATAGAAGGGACTGA
GACCCTCAACGAAGACCTCATCACGAGACTCCACGCCGTTCTTCGTCCCTTCTTCTACGTGCGATGAAGAAAGA
CGTAGAGAAGCAGTTGCCTAACAAGTACCACCAGCTCGTTACCTGCCAGCTCTCGAAACGACAACGATTTCTGTA

CGAGGAGTTTCATGTCTAGCGCCGTTACGCAAGACACCCCTCAAATCAGGCAACTTTTTGGGGATCATCAACGTTTT
GATGCAACTCCGAAAGGTGTGCAATCACCCCGACCTCTTTGAGGGTCGGCCCATCATCTCTCCCTTTGATATGGA
CCCGATTGAGTTCTACACCAGCTCAGCAGTGTGCAGAGCCCTGGAAGAGGACCCATTCAAGTCCGTCAACCTCCC
TTTCTCAACCTCTCTCTTGCCTCCACCGAACTCGAAGTTGGCCACATACAGGCACTGCGCACGTACAGTTACG
CACCCCGTCACCCAAAATCGTCGAGCTTTCTTCCACTCCCTCTTCTAAATCCACCCCTCTTACAACATTTCGCCCC
GAGCACCGTCGCTCCTTCCATGTTAGACGAGCTGCGTTCCGCACGTGAAGCGCGCATGCACCAAGAAGCGCTGGA
CAACCTCAAGCATGCTGCGTATATCAACGAATTTAGGTGCGCTCGCAAGCCGCTCTATGGATGGGAGCTACATCG
CTCTGTGCGCATCCACCGCCAGTCCATGACGTTTATACCTACGCCAAGACCCCAAGCACTTCCTAGACTACCC
TTCCACCCCTAAGTACTAGTTAGTGCCTTCTGAAAGCTCGCGCCGGAACCTCAACCTACCTCATCGCCCCATT
TGTATGCATCATCCAAAAGCCCGTTCTCTACCCCCACGACTCGCCTGCCACATCCCAACCCCTCCACCGTCTCT
CGAAACCAAGCGATCAGAACTCGAGTTGGTTTCAGAACTATCGCCTACAGTGGATATTTTCCGTCCCGCGTTCTGT
GCGTTCCAGGTCTACTTCCAGACAAGAGCCTCGTGCAGTACGACTGCGGTAAGCTGCAAGAGCTAGCTGTCTCT
CCTGCGCACGCTGAAGAGCGGCGGACACCGCGCGCTCATTTTTCACGCAGATGACGCGCATGCTCGATATCCTCGA
GGCCTTCTGAATCTCCACGGGCACACGTACCTGCGCCTGGATGGCTCCACGAAGGTGCGAGAAGAGACAACAGCT
GACGGAGCGTTTTCAATACGGATCCCAAAGTGTTTTGTTTTTCATCTTGTCCACGAGGAGCGGTGGTGTGCGAATTAA
CTTGACTGGTGTGACTCAGTTATCTTCTACGATAGCGATTGGAATCTGCTATGGATGCACAAGCGCAAGACCG
CTGTACCCGCATAGGACAAACGCGAGACGTGCACATCTACCGCCTTGTACAGAATTCACCATCGAGGAAAACAT
TTTGAAGAAGGCAAACCAAAGCGGCATTTGGACGACATGGTAATCAAAGGTGGTGGCTTTACTACCGATTTCTT
CAAAAATGTAGACGTCAGAGATTTGTTTGAGTCAACCAAGGACCAAAATTGACCGAGAGAGCGCACCAATTCAGCT
GAGAGAAAGCAAGGCAATTTTCGAGCAGGACTGGGAGCGTGCAGCGGCTAGTGTGAGGATGAAACCGATGCAAT
AGCCCTACGGCAAGCACAAGCCGAACAAGATTTTGATAACGAAGAGTTTACCGAAGAGGCTGGGCAACCTAGCGA
ATCTCAGCCCAACGGAACGGACGCCCAACGGCAGCCGACCAAGACATGTCATTTCTTGACGACAAGAAGACGAATA
CGAGCTCCAGGCCACGAATGGCAAGTTGGCTAGCAGTGCCAGTGGATTTCGAGTTCGAAGATAGCCTACTCCCCAT
CCAACGTTATGCGGTACATTTCTTGGAGCCGTGAACGCAACAACGAGCGGAGAGCCAACCGAAAATGTTACTGA
GGAAGTGAGCACCCGAGAGCAAGAATGGGAGCTGGAGAATCAAGCGCATGAAGGAAGAGGAAGAAGAGCGCAT
GTCAGTAAGAAGATCAGCTTTTTTACGAGGTTTTCTGAAATGCGGCCGAGTTCCGGGCTAAGCTGCTCGCCGA
AATGGAAGTGACTGCCACCCAAACCTGACTCGAGACCTGTACATCCCCCTACGATGGGGACTACCTCTC
CCAATTTCCCCCGTTTCCACCGCCGTTTCGAGGACGACGATACTCTCCCAAAAACGAACCCCGAGAGAAGCGATC
CCCGCCCCCTAGCCCTCCGGGAGGAGAAGTGGGGGATCAGGAGGCATACACCACTCGCCCGCAAAAAGAGAGAAAG
GAAGAAGCATGCCAGACTGATCATGTTCTCAGTCTCCTCCCTCCCTCCCATGCCAACTAAAAGGAGGAAGGGG
CAGGAAAGCTAATGGCGGTCTTTTCCCAAGTAGGATTGTACCGAAAACGGCCAGAGACGGGAAGAAAAATCGCACC
GTCAGCCCAACAGGACCTCTACCCCTGGTCCCCCGAGGAGGACGAGGCACTGACCAACGCCACCCGGACCTACGG
GCTCAACTGGGATCTGGCCTGCGAGATGCTCCTCTCAACCCCGCCCTACCTCGCGGAAGTTTCATGCGCACGGC
CACTGCGTGCGCGGAACGGTATCGCACGGTTCTGCAGCCGCGCGCAGCGGGCGAGGCGGATAAGGACCGCAAGTC
TCCTCCCTCGTCTCCCTCATGTGCGCATCCCAAGCGCAGCCACCGCACTTTGGTATCGTGGATGTGTTTCGCAG
TATGGTTAACCAGAAGAAGGAGATCCCTGGCATGATAGGGAGTGTTCGCGCAGACGGACCAATCTGTGCCACCACA
TCCTTCCCATATGCAGGCTGCCATGCACGCTGGAATCACGTGAGCAAAAGCCACTTACCCCCCGCGATGTTCTCGC
TCTCATTTTCACAGCGCGTCTGTGTTGACAATGCCACAAAAGGAAGTCCCGATCCTAACAACCAACTGGCTACCAAGCT
TGACCCCAAAATATGCGGCTTCGCGCAGGGGCTCGGCTCGAGCATGGCGCTCGCGATCGCAGTGGCGAATCACGC
CGCGGCATCTGCGGCAGCAGCTCAACCGGAATCCAGCCAAATTCATGGTGCAGGGCAAGGACAACCCGGCGT
GCCTGCCAACACCGCCGCGCCAGCCGCGACGTTTACGGTGCCTGGACCCGGGGTCTGTCGGGGGCGTTCCGGCCGC
CAACAGCATCATCTCAACCCCAACGCGAGCAGCAACGCGCAGACCTCGCGCAAGCTCGCGCAGATCAACCAAG
CATGCTCCCCACCGCCCGAATGTACCATCCCCCGCGCAACCTCCCCGGGGGTGTCCCCACGGTTAACCGCGC
GCCTGCGGTAGTTGTAGCAATCCCGAACAACCCAGCTAGTGTGCGGGTCCCCCGCTCTAATCCGAATAATACCAT
TAACCCGAGGGCAACCCCATCGCACCTACTGTAAATATGACTGTCCACACCCGTTCTTCTTCAACCAACCC
ACAACACCCGAGCAGGTTCCAAACCTCACCGTCCAAGAGATCGAAGAAAAACGCAGAAAAATGCCGCAAAATGG
GGCTGCAGGTGGCAACAAACGCCAACGATCGAAACAAAACGCCGCTCCGCCCCGTCGCCGAACAAACACCCGCACA
ACCCACACAATCAGTCCAACCAACCCACACAATCAGCCACGCTACAGCAAAATCAGATGCACCAACTACAACA
AATACAACACCAGTTACAACAGCAACACATGCAACACCAGCAACAACCCCACTTAGCGGGCCAGATTCCACAAGT
GACACAGCCACATTTGCTCACACAACCAACCCACCCATGTGCACCAATACTCCACCCAAACCAACCCACACA
ATCCACAGTTCTGGTCAACCAATCTTTCAGGCACGAGTACACCACTACACAACCATAACAACCTACTCCCAC
GTATACAACAGGGCCTACGTACGGTACACCAACCACTACACCACTCCACAACAGGGATTACGTACGCTACAGC
TTACACACCCCCACTTATGCCAACCTACCTCCACACCTAGCACCTACACAACAGCCACCCCGGCACTACCTA
CCCAGCAACTACCACCTTCAACAACACGACGACACCCGGAACGTATACTACCGCGTACCCCGGCACCACGCCCGC
GGCAACGTACGCAACGATGCCAGCAACTACTCCAGCTGGTGCACGTACACGACCGCTGCCACGTACAGCAACGTA
CGCGGCACCCACAGCCGCGGACCATTCACCACAGCACCAGCGGGCTCCACCTTTCAGGGGCTCTCTGCGACGTA
CATCTCCGAGGAAATCTGCGGGCACGTTTGCCACACCTGCCTCGACCCGTACACGAACCTCGGCTGTGGGATA
TTCGCCGAGGGTATGGTTATCGGAGTTCCGCAACAGCCAACAATGATAGGAACAATACCTCCAAACGCTCTCAA
ACCCCGCGCAGGGCAACCAACACGGGCGCATCTCGCCCCGGGGCACTGCGCAGTCTTCCCCGGGCACTGCGCA
AAACGCGGAGCTGCTCGCAATGATGTGCAGTGCAGTTCCCTGCGTCCACGAACGCAGTGCAGGGCGTGCTAAT

GCGCAAAGACTTGACTGAGGAGCAGCGCCTGGAAGCCGTGCGCAAGATTATGAGCGAGCACACTCCCATGTGAtc
ttgcttctcgtgttattgtatgttaaagtagtgtacactttgtgaa

>Phypoly_transcript_09236_PpSWC2

attgagatttaaacatgtctaccagaagtttatatatacttacatttgaatattaaggagggggaatattacacca
gaaattgaaaaatcatcaaacagcttagttcttcagaaaatattgagatttacaa**ATGTCTACCAGAAAGTTTACC**
ATCCCGCTCCACAAGGGGCAAAAGGATGAACGCGCTCGTTGGGGATGATGCGCTTGTAGATGAAGAGTTTTGGAA
AATGGATATGTTTAAGGAAGATGAAGAGGATGATGAATATAAATCTGAAGCAGAGGAGGAAGATATCATCGATGC
CGATTTTCGACAATCCAGAGGAGGCCGAGTcagaagaagaagttgaagttgccaagagaaaaaagaaaaagaaaag
gatggcctatgtagatccaaaagctaaaaaagggccaaaaCGACGGAAAGATCCAAGATAAAAGCGAAAGCACTCGA
CGCAGAACTACTAAAGGGGCCGAATCGGACGTGGAAGTGATCATGGACGAAGAACCATAATGAGGAAAAAGGCGA
GGCGCCAAAAGCCGAGCCCAAAAAGGAAGGGGAGGCAGACGAAAAAGACGCCCAAGCCCTAGAAAATCCCGGA
GGGCGAGGAAGGGGCTGATGCTGCGTCACCCCCCTAGCACCCCCGGCTCAGTCCCTAACACACCTAAATCCACGCG
CAAGTCCACGCGCGCAGTAGCTATTCAAAGGGCCGAGCAGCGAGAAAAAGGAACGAGAAGAAACAGAGAAAAAGAG
GTCACAACGGTCAAGGAAAAGATCTCCTGTGGTCTGTCGGCGCATTACACAGGAAGAAGTACTAGAAAGCAAAA
AATGACCGAGATTTATAATACAGAGTCTCTAAATATATTCTGTCTATGGAAGAAGAAAAGAAGATAAAGGA
AGCACCCACAATCGTAGGTCCAGCAATAAAATATTACTCAAAAGACGGCAAAACGTGCGTGTGTTTTGACGAGTT
GCCCCGACGTTTTGGCTTCAGCCCCGCCACCCCCATCCCGCGAAATTTGCCCAATTACTGGCCTTCCAGCGAA
GTACAAGGATCCCAAAACAGGAATTTGTTACGCCAATTTGGACGCGTTCAAAACACTTCGCACCAAGCTACTACA
ACAAGAAGAAGCAGAAACCAATTTGAAAATTTGCCAGTTGACCGCCTTGTGCAAGAATCTAAACGAAAAAAGA
TCAATTACAAAACACGCCCACGGGAGTTTTGGTGTGAaggaaactttttattagatttgggtgcatggcaggtaggt
taggctcggctaagctaggctagtctagtctaaattttgcagagggacaaagataaagacggaagagagtaaggc
atcttggcaaaggcttgggtggatttggaaatgtaccaaggtctttgtattgtattttatgtaaaatagttcacac
tt

>Phypoly_transcript_00559_PpSPT16

acagcgcaaggaaacctgttaccaagacc**ATGGCTGACAAGTCTGCTGAAGCTTCTGCAGGTTCTAAAAATACTA**
TTAAGCTAGACAGTAAGTCTTTTTGGAGAAGAGTAAAAGCCTTATATGATTCTTGGCAGAAAAAGATGATTATT
TATGGAAGGGACCGGATGCATTGGTTATAGCACAAGGAACACCAGATGATGATAATCCGTATAAAAAAACAAATG
CTTTGCAATATTGGCTTTTGGGATATGAGTTTGGCGATACCATCATTGTATTTTGGCAGTCAAAATATATATGTAC
TCATTTCTCCGAAAAAATTGGATTGCTTGAAGTATTGAAGGAAGAAAAATGCGCCATACGTGTTGTGTTTTTGC
AGCTTTCCAAAGCTGAAAAACAACAAAACCTTGAATACTTTATTCGAGGCTGTCCACAAGAGCAAAAGTGCGAAAA
GCATTGGAGTGTTGCCAAAAGAGAAAAATGCAAGGAGCTTTGGGGTCACTTTGGGAGGAAGCGCTTGCGAAATCGG
GGCTGACACTAGTGGATGTGGCACCAGGATACTCTGAAATGTTGTCTGTAAAGGATACTGTTGAAATGAACTCT
TAAAAACAGCTGCAGAACTACTTCCAAAGTGATGTCAAATCTTAAATCCAGCTGTTGAACTGTGATTGATG
AAGGAACACAGAAATCACACACTGCCCTAGCAGAAGAAATGGAAGAAGTATTGCTCAACCTGAAAAAATTAAT
CTAAGTTGGTAGGTGAGCACGTGCAATCATGCTACATTCCAATCATTCAAAGCGGGGCAAAATATGACCTGAAAT
TAAGTGCAACTAACGACAACGAGAATTTACACTTTGGCACTATCGTGTGTGCTCTTGGTGCGAGATATAAAGGAT
ATTGTTCCAATGCAGCGCGCACTTTTTTTATAGATCCGAAAAAGGAAGTAGAGATCAATTATAGTTTTCTTCTTG
AAGTACACAATATAGTAGTGAAATCCATCAAAACAGGTGTAAAAATATCGGTGATAATGGACAAAGCGACCAAAAT
ATATTGCCGACAAGCGCCCTGAGTTATCCAGCTCTTTCTTAAAGTGCTGCGGTTATGGAATAGGAATTGAGTTCC
AAGAAACGTACCTGATGCTCAGCAGCAAGAACACACGCACCATCAAACCTGGCATGGCTTTTAACATCATGGTAG
GGTTTCAAAATCTCGAACTTAAGGAAGAGCCTAAAGACCCCAAGATAAAGTGATTCTATGTTGTTGGCAGACA
CAGTGATTGTTAGTGAAACCGGAGAGGCGGAGGTTTTAACTAATTCTGTAAGCAAAAGATTGCGTGATGTCTCTT
ATTTCAATTGAGGAGCTAGACAACGAAGAAAAGGCCTTGCCGACTACTACCAAGAAATCCTCTGCCACTGTGAAGA
GCGAGAAAGCCGAGAAGCTTGATGACGCTTTCAAGAGTGCCACGTTTGACGGACCACGCTCTCGAGACCAGAAGG
GCAAATCAAAAGGCTCGGAGGATAAGATAAAGAAAACATCAAGAAAGAACTAGCAGAGAAGTTGGCACGGGATCGAC
GGAAAGGTGCAGACGAAGATAGTACCTCGGCTGGCCCTTCCACCGTCAAAACCTCCTCAAAAAGATATTGTGGCTT
ACCCTGACACATCTGCATACCTCACGAAGCTCACAGAAACAAGATTTTGGTTGACCACAAGGCCGAAGCTATTTC
TTTTGCCCATCTATGGGTATCTAGTCCCCTTCCACATTTTCGATCGTGAAAAATGCAAGCAAAACAGACGACTATC
TCAGAAATCAACTTCGTAACACCCACCCAGGTCAATTGCTCAATCACAGGGCCAAGGAAAGGAATTGCTATATATTA
GGGAGGTGACATACAGGATTCAAGACCAAGAAACCTTACAAATTCCTTGCTGCTCATCAAGCAACTGAGGAAGC
AAGTGAGCGAGCGCAGCAGAGCAGCAATGTCATGGCCTGGTCTCCAGGACAACTTATTGTTAGTAAAA
ATCCAGGAATCCCAAGCTCATGGGCCTCTTTATTCGTCCTAATATCAGTGCCACAGGACAGGTGGGACACTTG
AAGCACATACAAACGGCTTCCGTTTTTCTACTGCAAGGGTACTACTGTTGATATTTTCAAGAAATATCAAAC
ACGCATTTTTTCAGCCATGTGACCATGAACAAATTTGTTTTAGTGCACTTCCATCTTCACAACTATATCATGATCG
GCAAAAAGAAGGCTAAGGATGTTTCAGTTCTGCACAGAAGTGATGGAGGGGTCATCACACTTAGATGGAAATCAC
GTTTCATCATCACGTGATGAGATTGAGGATGAACAAGCTGAGGCTGAAATGCGAAATAAATCAACAGGGATTTC
AGACTTTTGTCAAGCGTGTTGAAGAGGTTTCTGCCCAATCTGGCACTGAAATAGAGTTTGATATTCCGTTCCGTG

AGCTAGGATTCTTTGGAGTGCCCTCCAGGAGTAATGTGTTCCCTTCAACCCCTCTGTGCATTGCTTGGTGAGTCTGA
TTGAGCCCCCTTTCTTTGTACTGACACTCAAAGAAATTGAAATTTGCAACTTTGAGCGTGTATCGTTTGGGTTGA
GGAACCTTTGATATGGTGTGTTTGTGTTCAAGGACTACGAACGACCAGTAACCAAAGTTAACAATATCCCCGTCGATT
CCCTAGAGGCCATCAAGGAATGGCTTGATGAATGTAAACATAAAATTTCTACGAGGGCCCCACAAATCTGAACTGGA
CCCACATCATGGGCACTGTAAGGCAGGACGTGGATCAATTCTTTGAAGATGGTGGCTGGAGCTTCTTGGAGGACA
AGGACTCCGATGCAGGCGATTTCGTCACTCTGAAGATGGTTCCGAGTTTGAACCTGAAAGCAGCGGCAGTGGTGAAG
AAGAAGAGGCCAGTGAGGATGATTACTCATCAGATTCCGATGCCTCAGGATCGGATTACAGCTCTGCTTCAGATG
AGGAATCAGAAGAAGAGGGATTGGATTGGGAAGAGCTTGAGCGCAAAGCTGCTGCTGAGGATAGAGCAAAGAAAT
ACGATGAAGATGAGCCTCCTAAGCGAAAGAGAGCTGGAGGAGCGTACTCGGACGACGAGGACGATCACCCCAAAT
CCAAAAAACAGCAGTATCGAAGGGAGATAGTTTTGTCCAGAAATGGAAGCAGTGAGTGGTGCCAGAGACATCACT
AAgtacccaggactggatattgttcttaacattcacatTTTTtatttctagccacccccccagggggttggtgtg
gaggggtgggccccccactgaccagcacaaatttgtacatttgcaccaggcaaatggataatttcttgagctg
aatcatttggagtgctgtgtgggctttattcataagcagtttggggggcatttgatagtagtaccactgtattcctt
gtttgattatgatttcaagtgtcttgggtgtgttagaggggtgaatttgcTTTTgtgtagcagtttttttatcat
caagcttatgtgaatctgtgctgggcatgtgtgtacttttatcagcaacttttgctaaagcattagctggtgtgtg
gcacaaaaaatccatgcactgcactatattcacattcttgggacaataggccaggcttgacatgcttaagtactg
aagcacacatataaattggcatcaaaggggaagtgttaaaaaacacatcaaacaatcgtgccacacaagtaacat
catctaagttgtgtgcaaaccaagttaagacccaactaagcgcaaaaaaggggtgcacatgtgattcaacaatga
aatgggaaatttttggctcaaataattgataataatggcacaaggtgatccaagtatgctataataataatccaagt
taggtcgttaaggcatctctaattgtgaaaaagggacagtttttctaacattgcaaatgccaatttctcatcagtta
ccattaagaaaacactagcaatgtcatgatatccctgaaagtatttcaaggtggggtgctcagaaagaacagcca
aaataatatttgataactctttgcgcttctgagcaagtagcattggatcacccctctgtgaacttccacagtgatc
tctccacatctctctccacttgatgcgctctgggtgttcttcttctgtattaattcagtgatgcacgaccatcag
caggaatgcactcacactgctggatgccaaagtaagatgggcccaggcccttgcctgattgcattgttcaccagcc
caccctctgttagtgcatattcaatcagtgctgcccgtgacacaggggggtgatctaacaataaacaatgcatt
gggctttggaatctttggaacccttcatctattgcgcggtgtattgcggcaagaagaagaggagaatgtagttgg
tggggcaagaaagtttgacctgcgtgtccttcggtgataataaaaaaaaaaagtgctggtgtgttt

>Phypoly_transcript_05191_PpSSRP1

gtctttcttgtatcttgcctttcacacaccaacatacatatgattcatttttaattttgttttttaaaatattact
aatgatatacatgttgtattatTTTTTcaaagtgcacatatttttgtgaataaaagtatgtttgtgtt
gttttgtgtgaaatttgatgtgaatgtgagggcatgaacatgatggaaatgagaggtgcagctgtacaagagcga
aaagaaaatgacgcaagaaatccattgcgtgggcagctttttattcgcaggtaggtctttaaaaatttagcgcaa
caggttcccgcgcaagttgttcttccaa**ATGGCCACAACAGCACATCAATTCAACAACATTACCCCTTGGGGGTA**
GGGTTAGTAGTACAAAAGGTGTTCTTAAGGTGAACAGCACCATGTTCTTGTGGCGGACAGAAGCTGGAAGAGTGA
TATCGGTACTTCCAACCTGACCTATCACAAGCTTTTTTGGACACGTATTTACGGAAGAGTATTTACAGCTCAAATTGC
AGTTAAACATTGAAATTTCTATCAAATTTGATGGATTTAGAGAGTCCGATTACGATGTTTTGAAGACATTTTTGT
CTGAAAATTTTCATGTTGAACTTACTACTGTACAATTAGCTACAAAAGGACTGAATTGGGGGGAATACGAATTAG
CTGGACCTGTGATCACTCTCGCAGTTGACGGAAGCACGCTTTTGAAATTCCTGTTTCTGAAATTTCTGGTGCTG
TAATGCAACCTACTCAAAAAATGAAGTTGCAATTGAATTTTCATCAAGATTCCACATTGGAAGATGAGGACGAAT
CCCTTGTTGATATTCTGTTTCTTTATCCCCAATCGATCCACAGAAGAAGGAAGCAAGACATCAGCACAAATATTAC
ACAAAGGACTGTTAGACGTAATTGGTGGTGGCATTCAAGGAAAAGATCGCATTGCAACTCTTAACAAAGTACCCA
TACTTACTCCAAGAGGTGTTATGATATTGAACTTTATCCCACTTTTCATGAAATTCGTGGCAAAACACATGATT
ACAAAGTCACTTATGAATCCGTATCTCGTCTATTCCAGCTTCCAAAGCCTGACCAACATCATATATTCTTTATTG
TTACACTAGATCCCCCAATTGACAGAAGGAACAATGCGGTATCCTCACCTTGTCATGCACTTCAAAATGGACGAGA
AGTCTCCTCCAGTACCTATCACACTGCCTGCAGATCTCAAAGAAAAAGAACAAATTAAGAGCTTGCCAGAGGAAT
TCAAGGATGCAGATGCTTATGAAGTGGTTAACAAGTTATTCCGTGCTTTTACACAGCGCAAAATTACCACTCCAG
GCTCCTTCAAGTCTCATTCTGGCGCTTCTGCTGTCAAATGCGCCTTCAAGGCTAATGATGGATTCTTTTTCTTC
TCGAACGAAGCTTTTTCTTTGTACACAAACCTCCTCTGCATATGCGATTTGAAGAAGTCTCCAGTGTGGAGTTTG
CACGTGTGGCATCAAGCACTAATGCACAGGCCAACCGTACGTTTCGATCTTGTGATTAGGCTAAGTGACTCAACTG
TGCACCAATTCATGTCATTCAAAGAAATGAATACGGGTCTCTTTCAATTTTATTCAATCAAAGAAGCTTCGCA
TAGAAAATCCTGATGAAGCAACTGAGGAATTGGCAGGGTTGGCAGGAGTGATCGGAGCAGATGATGATGAAGAAT
CTGAGGAGGATGAGGACTTTGTTGATGAGGGTGAAGGAAGATGAAGAGTATGACTCTAATGAATCAGAAAGCG
GTGAGGAGGCAGAAGAAAAGCCTAAGAAGGACAAGCACGAAGGTAAGCGCAAAAAATCAGACGGAGAAGAGAAAC
CCAAGAAGAAAAAGAAGGCAGCGGAAGAAGAGGAAGACTAGgggtgtgatcattttgttttaataactttaactgta
tgtctattttccactaagctgtaaaatagtagtcttttacggggaataaattctcggttataatg

>Phypoly_transcript_13763_PpSET

AGCGAGAAAAGCGCGAGAAAAAACCATGTATAATAAGCACGACATTGACCCGGATTCTTTGGACAGTTCTGCGA
AGGCGATCTTGACAAATCTTTCCGTTTCAGCCTCTCGTTGATTGCCCCGATTGATTAGATTGGAAGCCTCCCGCAG
ATCCAGAAGAACAAGTGCCGATTGCGGAAGTTCAAAACAAGTAGTTCCAGATAGTTTAAATGCCTGCTTTGATTT
TGTACATTACACCCATCCAATGTCTCTACACAAAATTATTGAAGGTTTCCACTCTGAAAATACGCAAGTTTCCA
AGCGCCAACTAGAGCTAAAAATACGAGAAATTGCCGTGAAAGTCAAAAAATCTTCGTGGCAGGTCCGAGAGGAAG
TCGTCAAATCCTATGACGTCACACTACCCGAAATAATCGAGAGCAAACCCGTTCCGCCCCAAAACCCCCCAGAAAA
ACATGACAGATTTCTTTGCCAAAGCCGCGCGCCAAAAATCAACCACCAAATTCAAACTCAAATCAAATTCAA
ATTCAAACCCCAATTCAAATTCAGCACACCCACCAAAAAAACCCCAACCAAAAAACCCCCACCAAAAAACCTCC
CCACACCCTCTTCCACACCCTCCCTTCCGAAAACAGCACAAAAAACGCACCTACCGAAAACGGATGTCACACGA
CCCCAAAAGGAATATTATGGATATGTTTCAGAAACATCAAAAGACCCAGGGTTAGgttccactgcgtctttttg
tttgtgtaattattttgaataatttcaaaaaataaaaatatcttgcgaaagt

>Phypoly_transcript_09859_PpCAF1B

attaatttgaatttttagcggaagcaccagctacaatcttcttgtgcctggttgtttagtttatgctcccttagt
tgcccttctttatgtcctagagttgtcttctgtgcaccataaaATGAAGGCGGACACTCTTATGGTCCTATGGCAC
AGCAAGGAACCTGTCTTCTGTGGACTTTCACTCTTCTGGTAGACTTGCTTCAAGTGGTGCAGATAATGATATA
AAGATATGGAATAATTGGAACCGTGGGATGGAAGTGTTCACCTTGATTTCTTGTCTAGCCTCACGAGACACACC
AAAGCAGTGAATGTGGTTAGATTTTACCTGATGGAACCTTCTTGCCACAGGATCAGATGATGGATACATATAT
ATTTGGAAGTTAAATACATCTGTGCCAGTATTACCCGATTCTATTGACAAAAGAGTGTGGTCCGTGGCTTATACC
TTGAGAGGAAAAGCACAAGATGTGTATGATTTGGCTTGGTCCCCAGACAGCTCGCAGATCATTGTAGGTAGCACC
GACAATAGTGCCTCTGTTTGGAGTGTGAAAGGAGCGCAATGCAAAACACCCTTTGAACACCACAACCATTACGTA
CAAGGAGTAGCATGGGATCCTTTGGGAATATACCTTTTGACGTTTCAGCAGCGATCGCACGCTCCGCGCCTACACG
CGAAAGCTCCCAAAGAACAAGAAGTCCACAAAGCCCTTCTTCCCCCTACCCCTTCTTCCGCTCCGCTCT
ACCTCTTCTCTCTCTCCGACCCCCCTTCTCCCTGTCCCGGACCCCCCTTCTTCCCCCTCTCCCCCGTCCCCGCC
CCAGCGCAAAACCCGCCCCGCCACGCCCAGCAAAAAACGGGTGAAAAGAAAAGTGGATCGTCGCGCAGCTCGTC
AACAAGCGCGAGTTCCCGGGGACACAAGGGCTGCTCGAAAAAGAGAAAAGAAAAGGAAATCCGGTAAA
GGAAAAGAGAAAAGAAAAGAAAAGGACTGTCTACAACACAAGCTATTTTTTGACGAAAGTGTAAATACATTTTTT
AGGAGGCTCACATGGTCCCTGAGGGCTCCCTTCTGTTTCTTCTACTGGAATTCACAAAGCCTCCCTGCTCT
CCTCCTACCAATGTATGTTATGCATTTTCTAGGAATAATCTTAACAGGCCGTTATTTTACATACCAAGCGGGAAA
CCGGTTATAGCCGTGCGATGTAGCCCAATCCTGTACAACTGTGGAACGGCAAGCACCTTGGCAAACCAACCC
CTCTCTACCTCTCCACCCCCCTCCACACCCTCCTCCGCACCCTCCCCACCTCCTCTTCTACCTCCCCACTCCC
GCCTCCTCCACCCCTGTTTCTACCGATGCACCCTCCGCCCCCTACCTCCACACCCTTGGCACTATCCAGATACC
CCGATGAGTGACTCCCCACCAACACCGCCCCCTACCCCCACCCCCACCCCTGCTGCCCCCTCTCCCCATCTCCCC
CCATCTCCCTCCTCCCCCTCCCTCGTTTCGTGAAGCTCGATTACCGAATTGTGTTTGGCATCGCGACGGTGGATTG
GTGGTGGTTTACGACACGCAACACTTACACCCCTCGCGTTTATGGGAGACATGCATTACGCGAATATGACTGAT
GTAACCTTGGTCGGCTGACGGTTCGTCCCTGATCATGTCTCGACGGATGGCTATTGCTCGTATGCCACCTTCGCT
CCTGGCGAGCTTGGCACTCCCTGGAAGAAAATGAGATACCAAGTTGTATGAAAGAGGCCATTTTACAAAAGACAG
CAGAAAGCAGAGGAAGAGGAAGAGGAGCCAGCCATCATGCCTCCGCTTCTCAGGAACCAATGGAGAGCCATCC
GCGTCCGCGTCCGCGTCTGCAGTCGCATCTGCGCCTGCCGCCACGTCCGAAGATACGGCTGATGCAGGCTCGCTT
TTTTTTGGGAAGAGAAAAGAGATTCCCCAAATACTACCCAAGACCCCTCGCTCGCCAGCGCAAAACCTGCCGAA
AAGAAGCCTCGCAGATTGGCTCCTACGTTTATTTCCGCTCTTCCCAATAAGTAAgtaaaatgttaaagcggttg
ttttaaggttaataaatgcacttaataactaaa
aaaaaaaaaaaaaaaaataaaaaagggtgatggggagggggtgggggtgggaaagttt

>Phypoly_transcript_09477_PpCAF1C

cccacaacaagcaataaacgaggcgtgcgcagctttcccttcccgcgcatttgttgcaagATGGCCAAAGCTACT
GAGGAGGAGATACCAGAGGCAGAGGAAAAAATTATAAATGAGGAGTACAAAAATATGGAAGAAAGAACGCACCATTC
TTGTACGATTTGGTGATGACACACGCTCTTGAATGGCCAAAGTTTAACTGTTTCATTGGTTGCCAGACAAGAAAATA
CAATCCAACAAAGAGTTTTCTGTGCAACGACTCGTTTTTGGGGACACATACATCTGATGCAGAACAGAATTACCTC
ATGATAGCCAAAGTTACCTCCCATTTGGATGACGCAAAATATTGACCCACGTAAAATATGATGACCAGAAAGGAGAG
GTCGGTGGATTTGGTGCAGTGAGCGAGAAGATAGAGATTGAACAACGTATCAACCACAGTGGTGAGGTGAATAGA
GCTCGTGCTATGCCTCAAAACCCCTACTATCATTGCTACTAAGACGGTTACTTCTGATGTTTTTCTCTTTGATTAC
ACTAGACACCCCTCAAAGCCATCAGCAGATGGAAGTGAACCCCAATGCCATTTTGAAGGGACACACAAAAGAA
GGGTATGGTATTGACTGGAATCCTTCTCGCGAGGGCCATCTCCTTTTCGGCTTCTGATGATAAACTATTTGTTTG
TGGGACATACAGATGCCAAAGCAAATACCACTATAGAGGCCAAAGCAAATCTTTAGGGGCCACAAAGAGTTGTT

GAGGATGTTGCCTGGCATTGCAGCCATGCAGACTATTTTGCTTCTGCTGGCGACGACAAAAAATTGATGATATGG
 GATGTCAGGACAAATGCAGATCCACTGCAAGTAGAAGCACATGCTGGAGAAGTAACTGCGTGGCTTTTAAACCCG
 TTTTGTGAGTTTCTTTTAGCTACTGGTTCTGCTGACAAGACTGTCGCCCTGTGGGATATGCGCAACCTCGATGCC
 AAAGTGCATTCTTTGTCTCACATACTGATGAGGTATTCCAAGTGAATGGTCAACACACCAGAAACAGTACTG
 GCCAGCTGTGGCTCGGATAGGCGAGTGAATGTGTGGGATCTATCTCGCATTGGTGTGTAACAAACAGCAGAGGAC
 GCAGAGGATGGTCCCCCTGAGTTGCTCTTCATCCATGGAGGGCACACCTCAAAGATCTCTGACTTCTCTTGGAAC
 TACAATGAGCCATGGGTGCTTGCATCTGTGGCAGAGGACAACATCATTCAAGTGTGGCAAATGGCAGAAAAATATT
 TACAATGATGATGACACAGATGAAGCCCAGCCCTCCAATGAGGGGgatttggagtgatgtgtgggttaatcacc
 tacttttcttttctctctttgtacacatagattaaattatatgttctttctgt

>Phypoly_transcript_01462_HIRA

ggtttagagttaattgtatttttttaattttagtagcacaggaatataaaaaatgataagcagtttacatctctatttga
 agtagcaaaagaacagctgcaatcttgtacacaatttgtgtgcaagtagataaacactatccaaccatcctggat
 tcttgaatttcttcatcttctgttattgttgtatgaatgcttggacacagtgtgtagaaaagagctttcactggc
 agaacgagcaatggacggttaacttgagacaaccagattgtgtaacacttgggtgtgggtgATGCCCCGACAAAGGA
 AATAAATGCAATCCTTCACAGCCACGTGCTTGCGCGAAATTTGAAGATCAGCAGAACACAAGTTATAATGAACAA
 ATGAGGATTTTAAAGCCTGATTGGGTACCCACACAGGATACCCAATATTCTCGGTAGACATTACCCAGATGGT
 AGTCGATTTGCTACAGGAGGCGGAGACAACAAAGTAATCATATGGGCTTCTGCACCTGTACAATATGAGGAACAA
 GAGGCCAACAAATTCATGCCACGAAAGCTTGACAGAGCTCACCTACCACATGTCTCTATTAGCTGTGTTAGGTGG
 TCACACAACGGGCAATTTTATAGCATCCGGCTCAGATGATGCACTAGTGATGCTGTGGAACTCTCTTCAGGTCCT
 GCGTCAAAAAGTTTTGGTTCAGATGAGCTCACTCAAGAAAAGTGGCGCTGCACCGCTACTCTTCGTGGACATTCTG
 AAGGACATTACTGATGTTGCATGGTCTCCTGATGACAGTCAGCTCGCTTCGTGTAGTGTGGATAACGAAATTTTG
 GTCTGGGATACTGCTAAAAATTCTGTAGTAGCAAAATTGAGTGGGCACACTGGATTTGTAAAAGGAGTGGCATGG
 GACCCCATGGGGCGATACATTGCTTCTCAGTCAGAGGATAAGACAATGATAGTATGGCGGTGCTCCGACTGGAAA
 ATCGAAACAAAATCGAAGAACCATTGTCCGGTCTGCTGCCACTACTTTTACAGAAGACCAAGCTGGTCCCCA
 GATGGGCAATTCATAAGCTCTGCACACGCGTTAGCGGTTCAAACCATGTAGCCATGATCACAGCACGTGACACG
 TGGCGGTCTGTGTGTGACTTTGTAGGCCACAAATTACCTATTGTGGTAACTCGGTTTAACTCCTGTTATATTGCA
 GACACAGTCGCAATAATTATGGAGTGTGCCTAATAGGAGGGCAAGACAATACCATTCTGTGTGGACAACCATA
 TCATCTCGTCTCTTCTGCTCTTAATTCTTGTGTTTAAACAGAGTGTGCTTGATGCATCATGGGCAAGGATGGA
 TTCAGCTTTATGTGTTGCTCTACAGATGGGACAGTAGTATGGGTAGAATTTGCAAAAAGGAGAATTAGGAACTGCT
 CTTTCACCAAAAGAAATGGATAAAGTGCTCAGGAGCATCTATGGTGACCATATCCTCTCTCGCCCCAAGAGCACA
 GTAACAGTGGCAGAGGATCCCTCTCAGCTCTACCTCCCAAAGACCACTGCACCAATGAAACCAGTAGCAGCACAG
 CCACCTCCTGTGCCTACACAGCAAGTAGAAACCAGATTAGCCAATGGCAAAAAGACGCATACAACCCCAATTTCTT
 GGCAACATTGCTTCTACCACTGCAACCACTCCCTTCATCCCCCCCCCTACCCCTTCTCCTCCTCTTCTCCTCCTC
 TCTACCCCTTTCCCTCTTCTTCTTCTCGTCTTCTCTTTTCTGCTCCCTCTGCCTTCTCCTCTTCTCCTCCTCCT
 ACCCCCTGCTTACTGCTGCACAAGTGAATGGATTGCAGGTGCTGCCAACAGCTTGTCTATTAATGGCTTGATT
 ACGTCATCATCTGTGAACAGCTCTGTGATGTCATCAGCTACTACAAATCACTCAACCACATCATCAGCTGTACCT
 GTGTTTCATCCCTCCACCCCGCTCCCTGGTTCAAATCAACTTCACCTGTTGCTGTTGCGCACACAATCTCTCTC
 CCAGATGTTGGAAAGAGGAAAGAAAGTGATGATGTAGTAGAAGTGCCACCACCAACAAAAAGCACAAAAAGAT
 GAAGAAAGTGATAAAGAGAAGGGAAAAAGAGGCAGAAAAAGGGGACAAAAACAAAGACATAGAGGCAGACCACACA
 AAAGAGAAAGAGAGGGAAAAAGGAAAGAGATAAGGATAGGGAGAAGGAAAGGAGGAGGCGTGAGAAAAGAAAAAGAG
 AAGGAAAGGTAAGAAAGAGAAAGAAAGGAGCGGGAAGCAACAAAGAGAGCACAGGAGGGACAAAAAGAGCGA
 CACATGTTGTCAAGGCTGCACATTGCACATGCACAACATGCTCCAAATTTATGCAAACTTATTTGGGAACCCACA
 GGACTACAGTTAGAAACAAAAACAGAAGGAATGGGAACACAGGTGAGCTGCCGCAACACAAAAATTTCCACAGCAG
 TTGTGGCGCTCATATTCTTGGAAAAGCTGTTTTGGTGTGTGGAATCAATCTTTTCGGCAGTGGCTATGGAA
 GATTCTTGTGTATATATTTTTCAGCAGTTGGCAGATTGATATTTCTCCTATTCTATGAGTGTCTCCACTTTCA
 TTTTATAGAGGCTAATAAAAACTTCTACTTGGCTGCTATTCTTCTGACTGTCTACTGCATGTGTGGAAATATAGCA
 CAGCAATGTGCTACTATAACAGCACAACCAGTTAATATTTCTCCTCCTTCCAATCTACACATTTCCCATTTGGTG
 CTCAGTGATAATGGCCAACCTGTAATAATTTGCTCTTCACATACATATGCATTTACACTGAGATGAAAGTGTGG
 ATGAAAATCGACTCAGTGTCTTCCCTTCCACAGGTATAACAGATGGATTTTATCAAAATTAATGGCTACAACA
 GAAAGTGGAAATAGAAATATTGCCACAATGACACAAGAAAGCGATGAAATGCTTGGTGGTTTACTGAGCATGGAA
 TCACATCTTGCTGCTAGTTTAGTTGTTGGCTCTAGTGTAGAATACAAATATTGGCTGCTTACATATACCCGCTAT
 CTTTCTCAGAAGCTGATGTCAACCGCTCACAGAATCTGCAATCTCTTAGGTCTTCCACCGGAAAAACT
 GGCCCTGCCCCACCCCTGCACAGGGTAAAACTGGGAACCTTCCATATTAGGCATGGCAAAACGCGAGCTGCTG
 CGTGAGCTGTTGCTATTATGTCTACGAACCGCAGCCTGCAACGTTTAGTCGCGCAGTTCCGTGAAGGCTCGAG
 AGTATATCTAAATAAatgaagtaaccacaaatc

>Phypoly_transcript_03291_PpUBN

>Phypoly_transcript_00351_PpCABIN1

TGCATTTCAGGCAGAAGAAATACCGATGGCCATGGAATGCTTGAAGGATGCAAGGAGCTCTTGGCCGATAGCGGT
ATTCCCATCATTTGGGTGCGCACATTGTACGCACAACCGCAATATTACTGCACAGGGAATCGAGCGAAAACTCCAG
GAAATTTTCAGGACAGAACTATAACGACCGCGGCAGACCACTTCATGCGCGAGAACTACGAGCAAGTCGTGCGCA
TCGCTCGAGCCTATCGCGCGTGGCTACCCCGAATCGGCGATAGAACTGACAGCAGGCAAAAAGGACATAACAAGC
GCCAAACGAAATAGACTTAAATCTTCGATTTATTACACCAGGCGTATTTGAACTGAAGGATTTGAAGAACGCA
TTCTGGTGTGCTGCGAGCGGATGCTCGCGGAATCACTGCCCCACTACAACATGCCCGAAATGCACGCCAACTTCCAA
TCTCTGCTGGATTGTATTATAGCCACTGAGAAGCAAGCGCAGACATTAGGAGTCTCTCTGTGACCCCTGAACCTC
CAAAAACATAAGAAAGCTATTATGGTCTTGGTAAAGAAATCAACTGTCCCTATGCTCGTGATAAACGGCGTGATG
GTGTTTTACGCAATTGCATCCGAAAACATGACCCCGACGCAAGAGTGGTCCCTTATGCACCTAACCCACCACGAA
CTTTTCAGGAAGAAGGCTTGTATTGCAAGATGGCAAATTTCTTCGCCGTTTTATCGCGCTGCTCATCCCGCTG
CACTTCCCGGAGTCGGACGAGCCAATGAGCACGGAAGCTTGGAACTCCTTGCCGAATACCGAATCGACATAGCA
ACGGAATTGGCTCAATGCATTTGGTGTCTTTACGGCATAAAAACCTCCACCCAAAACATAGACGAGCACGCGCACAAA
ACGCAACAAAACACCGCGAAGCGCGAAGAGCTCATGCTGATTTTCATCCTCGCAAGACGTATTTATCAGGACAAC
AAACAAAAGCTTCCCAAGGAGACGAAGGATGCACTTTCCAAGCTCTACTCCTCCTTCAAAGATCCCCCGGAAAGA
GTCACCCCTTTAGTAAAGGAAATCGAAAGATATCTGATCGATACCGAAGCCCGCCTCTCCTTGCAAAAAATAAAC
GAGCTCAAAGCCGAGCATTCCGCGAAGGCCAACAGTTACCGTGCGGTGCTCGGAAAATGGACCGAAGACTCGGAA
CCGGAGACTGAGCAGTTTGACCTGGCTTCTAATCCCGATTGCGATTTTGTATATAGAGATTTGTATTATATGATG
GCACTTACCTACCCAGACTCGGACTCAGCCTATGAAGAATTGATTAAGAAAGATTTATTCTGCAACCCGTGGCAG
ATTGAATCGTGGGTTGCTCTTGGGGTGATGCACCTACGCAACCTAGTGCTGCTCACCAGAGAGAAAGCGTTTTTC
TTCCCTAATCAGCCTCCGTTGGTGGTTTTCCGTGGAAAATTTGAATCGGTTGTGCGATAAATGCCCTCTCACAAGCA
CTAGAGCTGGATGGCGAAAGAGAGGACCTGTGGAAGGAACTCGGCCTCCTGCTCTACCAACAACACAAGCACATC
CTCCTCCCGTCTTAccccctttcttccccctctccctcttttaccctcttctcctcctcgccctccccctccttcc
tctcttctctctctctctcggaACCTAACACCTCGCCCAATGACTCAATTACCGCATCGAACGCCGAACCCGAAAAA
GACACGAGCGACGATTTCCAAAAGAGGAAAGAACTCGTCGCGCGCGTTTACGAGCGCGCTGCGTCTGTGCTTCAA
CGCGCGTCAACGCTCAATCCCAAAGAAATGGTTATATCCACAGATGCTAGGCAAAATAAGCGAAAAATGCGGGAAA
CCCCTCAAGGAATATTAAACATTATACCGTTCTGGATCGGAAATGCGGAAGAAAGAGGAACCTGAACCACTGTAC
CGCCTGCACGCTAGCAGCTCAAACCTTTTACTACGAACCTCTACACTTAGTGAAGCTAAGATGGAGGAATTAGAA
CAATTTAACTTCGCACCTATTGAAAATCCAAGTGAGCTTTTCGCATGAACAGAGGTACAAAAAACTTGTCACTGAT
TGTCTGGCCGCAATGGAAGAATGCAAAAACGACTACAAATGGTTCCACAAATCCGTTTACATGGTGTGCATACACT
TTGAAATATGGGCCAGAACAGACGTGAAAACGGCGCAGAAAGCTCTCAGCTTCCCTTTTCAAGCCCATGAGCCAG
AATTTCTGGTACATTTGGAAGAAAGAAAGGATCAGAAATTCACCGGGGAAAAATGGACCAGTACAGAATAAAGTAC
ACAAAGTTATTGCTTGTCTGCTGGAGGACAATGTAGACAAAAGCGCACTGGAAAACCTGCTTGTCTAAAAGTGAAG
AGAGACCCCAAGTGCGTGGAAGTGGGCTTGGTTGAGGATACTTTTGTGCTTCTGTGCGGTCCAAGCTATTGTA
ATCCAGCAACTATCAGAAACCTTAAACAGCAGCGTCCCTCCCCCTCCCTCACAAGACTCTTTGCCCAACCCC
TCCGAACCTCTGTCAATCCAACCTTCTCTTCCGAAACCTCACCCACCTCTACTGCTCAAACCTCTACCGCTATT
CCGGCAACCGAAGACTTTACCTCTTCTTCTGTCTCTTCTTCTACCCAGCTACTTCTACTCCCCAACCTCTTGCT
TCTGCTATCCCTCCTTCTGCCTCCTCTACCTCTTCAACTATCCCTTCTGCATCTCCTTCTATCCCTCCCTCTACG
ACCTCCTCTTCCATCATCTCGCCTCACCTGGAATCGCTCGAAAAAATTTTACTGAACGCGTGGGAGAGCCACAG
GAAGCCAATACAAAAAATCGCGCGATGCAGGACCTCGCTAAccaagttctagaacaagccttttaatttatatat
ggcatacggccaagaccaaataatgccaactgacgaaccaactgatatcctacaagcggttctcagttttgoga
aagaagtttgccaagggaagaagacgaaaaagaaacagactctaaagatgcaaacaagaaagacggagccgcacc
acgaagaagaaggaaagtgaagaaaaagagaaacctacaaatgaaccaccacaaagcgaagaaatggaaattga
ttcttcacaaaatgcgcctcccggtgtagatacatTTTTTAAAAAACCGTgcccgttattataaaaaaagccata
aaagttcgaattttacatcaccg

> S3S101K5_5354_LULU_Contig588_N11D4_PpNASP

gcgcgcaaaacttacaaagaATGAGCACAAACAAGAGAAGATGGTCAACCTCTAGAGGGTGGATCAAACACAACCAC
TACCACCACAACACTAGCACAACTACCTCTCCCCAAGAGACCGCTGTAGAGGCTGGCGCCTCCTCTTCAGAGCC
CAAAGAATAACCCAGAAAGAGGCGAGCGAAGAAATATGAAGAAGCAAAGGCTTTGTTTAATGAAGACAAGTACGC
AGAAGCTTCTGATTTATTTGCGGATGTGTTGACCCACAAGGTTGCTGTCCATGGCGAGCTAGGGCTAGAGTGCGC
TGAGTGCTACTTCTTGTATGGAAAGTGTCTATTCTTGCAAGGCAAAAAGCAGAAAAGCTCCGTGTTCCGTGCCACAGC
ACAACAAAAAATAAATGAGGATGACGAAGACAGCAGCGATGAAGAAGACGTACCACCAGTTGGTGAGGGAGGAGA
GAAAACCGAAGGAAAAGAGCTAGAGGAGAAACAGAGGATGTAGAGAATACACAAGTCGCGTGGGAGACCCCTCGA
AGTAGCCCGACTAATTTACGAGAAGCACCTGAAAAAGCGGCCGAGTTGTCTGACACCCGATATGCATTGGCAGA
GATCTTTTTTGGAAACAGATAATTGCAATGACGCCATCGAAGAGTTCAAAAAATGTGTGTCTTTACGCGAATCGAT
CTTTGTTTCTAATGACAGAAGGATTGCTGAAGCTCACTATCACCTTGGACTCGCGTATTTTTTCGACGGACAGACC
AGCAGACGCACACTTCCACTACTCGAAAGCCGAGACGTGCTCGATTGCAAACTGAAGTCACTTCTCCTTGGCAG
TGCAGAAATCGAGGACATAAAATCCATTTTGGAAAGACTTGGCGGCTAAAAATGGAAGAAGTAGATCCAGCCATTT
GGCCAAAGAGAAAGAAGCCACCGAAAAATGCAATAAATGGCACGACAAAATGGAAAGGGCAAGCACCCGCTCCTC
ATCCTCATCTCCGCTCCCCCTCCTCATTTCCCAAAGAAGAGACCGGCTTCGCGGCCCCCATCTTCGACAAGAC

ACAACCAGTCACGAGGTTTGGTGTGGTTGGTAGCTCAGTAAAACGTGCTACACCTGTCACCGCCTCCTCTTCCTC
CTCTTCCTCTTCGTCCTCCTCTTCTTCTTCCACTCCCTCAGCCTTCCAAATCGCTTCCGCTTCTCTTCATCCTC
ATCCTCATCTCATCCACAACCCCATCTCCTGCACCACGACCGGCCCTATCATCTCCGGAGAAAAACGAAAACT
GGAGGACATAATGGGAGGAGAAGAGGGATTGAGAAGCGAACGAAAAACGCAGGACACAGTATGTGAAACCAGAAC
ATGAataaacctgcttcgtagctaaaccaaaccggtacataagcgcaagtacttccagga

>Phypoly_transcript_05212_PpXRCC1

gttgcgcgaggacgcatgggcttttacactaatcatgcaccaaacatgaatcattttctccttttagttgATGGC
AGAAACCTGCTGTGAGTTCAAAGTGATATCAGTCTCCTCCGAAGATGCGGAGCAGAAAGCAAGCAACATTGCAGT
GCAAGATTTGTATGCGACATGGCAAACAGAACCTGGAGAAGAGAAAGCAAGTATATTGTTGGAACCTACCAGCTGA
AACAACTTGCATACTATCGAAATTGTGAATGCTGGAAGTACAAATGTAGAGATATACGTGACAACTCGGAAAA
CCAGCAGGATGAAGAAATCTTATTACCAACTGCTATCATGATGACTCTGGGCGAAGTGAATAGCAATACTAACAC
CAATCGGTTACGATTTTTTTGGACCTGACAAGCTCACAAAAGGAATATCGGATAAGGCGTGGAAATACATAAAGAT
CGTATGCATATCATTGACACTTCTAGGAAGCCCTTTGGCTTAGCTCGCATTTCGCCCTTTTTCTCCGCTCCTTC
TGACGAATACATTGCTTCTCTATTGCCGACAACCGCTCTTCTATGGTCTCCGTTCCCTCTTCTCAACCTACCAC
CTCAACTTATGTACCTCAAAGTATGCTTCAAGGAGCAAGTCTTTTGGGGGCCACCTCCTGATACCACAATAGC
CCAGGACACAAACAACAAAAGACTCCAAACCATACAAACACAGCCAACAAATCCAACAATAATAATCGTACAAA
CGACACAAACAACCTACCGGCCAAGTCAGCAGCAAAGCGCAAGAGGGTCATCATGAATGACGATTATGACGATAC
AGAAGGCGATTGAGCGAAAGACCAAAACCCACGATCCCTCACTTCTTACACATTTCCTTCATGTGAAGAAGAG
CGTACCAGTAGCAAATAGGTGCAAATCGCTTGTATTTACCTCATTTAACAAAATATTGAGCGGCGTTTCGTTTGT
AGTTTCCGGCATCGCTAACCCAGAGCGACAAGAGCTGCGAGATAAGGCCATAGAAATGGGAGCCCAGTATTACGC
AGATTGGATCGAATCTGCTACTCATCTTATTTGCCCATTTACCAACACTCCTAAGTTTATTTTGGCCCCAAAAGGC
AGGTGGTACCATTGTCACTCCAGCATGGATTTTAGACTGCTATAAAAAACAAAGTGCGCCTTCCAGAAACAGAATA
CCTATTAGCGGAGGTGAAGAAGATCGACGACGCGGGGATGAAAAGCACCATGGGCAACATGCGCGATTGGCTCGC
TACATTTAGACCGAAGTTTGACAAACCCGCAAATAATAGTACCAATACTACCCAATCCCAACCTCAACCAAACT
AACAAAATCCCTAACATTTGTTAAACACGCCCGAATCACCAAGGACTCACAAACAGAAAGCTGACGTCGATGATAC
CTCTTCCGTCAGCACTGAACCGATGAGCGACGACGAAAAACAAATAATGACAAAGAACCCCCAACCAAAATGTA
CAAAAAAGATCGGTGGTGAAGTAGCGCCCTTACAGACACCAAGGAGGATATGGGACCAACTCAAGATAACGTTTA
CTTCTCGACAAGCGAAGACGACACATATCGACAACAAAGGACAATACGAAAGGACTTCCAAGTGAACAAACC
CAAGAAATCCGGCGTTTTTTGACATCTCAGCACCTTTGCCGAATATTTCAAGAACATCAGGGTTGTATTATATGG
CAACTTCGACCCCGCTTTTGAACCACTTCTGTACCGATACGTGACTGCTTATTCTGGAGAAATAGAGGATAAGCC
GACTAGCAAGACAACCTCACATTATAACCAACCAATTGTGGGATAAAGAAATTGCAACAACCTGCGAGACGAATATCC
CTCAGCATGTATAGTAAGACCCAGTTGGGTGTTTGTATTGCGACAACGAAACAAAGAAAGGTGGAAACCAATTTGCA
CAGTACACAAACGGTTTAAcaaaacttgatctcttttattttaatgtaaaatatttctaaggc

>Phypoly_transcript_12278_PpNHP6A

tatcttcccgaacacatttctactctgacttgaaagttgacatgtgtataaagtagtttcgaaacaaagaattgta
gcgaatcttcaaagatatatcaactcgagaattgtagcgaaagggacatacaacgagtattctgcgaaaatgcaaa
tgagttagcgagaacttctatcttctttttcaaacattgttaaaagcctgcagccggtctttcgagttctattgttct
ttctgcttttattctgctataATGGGAAGGTGAATCTCAAGCTCCTGAGGATATGGTTAAGGAGGAAGCAAG
GACACTGGTAAGAAGAAGAGGAGGCAAGAAAGCGGTTAAGCAACTAGTGAAAGCCGAAAGCTCCAGCCAAAGAGAC
AAGAAGAAGGTTGACGCCAAGGAGGATAAGAAAAAGAAAGACGCCAAGGATGACAAGCAGAAAGAAAAAGAGAG
GAGGCGGAGAGCGAGGAGCAAGAAAGCGACAAGGAGGAGGAACAGGAGGAGGAAGACGGCGCGGAGACCAAGGCT
AAAGACAAGACCACGGCCAAAGGGAAGTAGGAGCAAAGAAGGCGACTACCAAGGAGAGCAAGCCCCAAGAAAGAG
AAGGCCGGGAAAGACAAGAAAGACAAGAAGGACAAGACGGAAGAAAGGTAAGGATCCCAACGCCCCCAAGAGG
CCCCAGACAGCGTACCTTATCTTCGCAAATGAGCAACGTGAAACCATCAAGAAGGACAACCCAGCTTTTCTCCG
CCAGACATACTGAAGGAAGCTGGAGCACGCTGGAAGCGCTCACCGACAAGGAAAGAGGAAATACCAGGATAAG
AATGAAATAGAGAAGCACAGATATAAGAAAGAAATGGAGGCCTATAATGCGAATGGTGGCGAGAAAGACACTGAT
AAGAAGGGTGCGAAGAGGAAGAAGGGCGAAGAGGAGGAAGAAGAGGACGAGGAAACGCCAAGCCGAACAAAGAG
GCAGCCGCAGCAGACAAAGGCAAGGCAAAACCGAAGGGTCTCCTTCTCGCAAAAGCAGGGGCGAAAAAAAGTAG
agtccacgaccgaagagtgtaaggatgtaatttgccaccgacacctctgacgaactaatacgcgtacgtgtacatact
atacaataaaaaccagacttatattt

>Phypoly_transcript_14313_PpNHP6B

aaacgtcgtcaaagagccggcatgcgagcaaccctgcccgtgagcagcccgctATGGATACTCCCAAAGAGTGCCC
TGCCAAGGAGGACGCCCTGCTGAGAAGACCGAGCAGCCTGCGCCGGAGAAGAAGGAGGCCGAGAAGGAGGAAGG
GAAGGAAGGCAAGACGAGGAGAAGGACGTCAAAGTTAGCGAGAGCGAGGAGGAGACCCCTCAAAACAAGAACCCAA

2. Detailed analysis of the chaperone structure

2.1. The H3/H4 chaperones involved before chromatin incorporation

PpHSP90 harbors 4 domains: (i) the N-terminal ATP-binding domain with high conservation, (ii) the glutamic acid-rich (20% in the PpHSP90 linker) and charged linker that is less conserved, the highly conserved (iii) middle and (iv) C-terminal domains (Figure 3A). **HSC70** interacts with H3 histones right after their ribosome exit as well as with H4 and HSP90 [5,6]. As other HSC70s, PpHSC70 presents the 3 general structural domains (Figure 3B): (i) a NBD (N-terminal Nucleotide Binding) domain containing an ATPase domain, (ii) a SBD (Substrate-Binding Domain) domain divided in SBD β and SBD α , the latter acting as a flexible lid and (iii) a flexible linker connecting NBD and SBD [96]. Overall, the HSC70 is well conserved across species, SBD α being more divergent (Figure S6F). In PpHSC70, we noticed the presence of the EEVD motif (Figure S3F, red box) involved co-chaperone binding [96].

The **NASP** proteins frequently present two isoforms corresponding to splice variants but *P. polycephalum* only present one splice variant and only one NASP orthologue is present in studied organisms excepted in *C. elegans* which contain two paralogues. NASP is a protein that bound H3/H4 constituting a reservoir of soluble H3 and H4. The PpNASP protein presents 4 TPR and one SHNi-TPR (Sim3-Hif1-NASP interrupted TPR) repeats as well as a "E/D-rich" region (Figure 3C). TPRs are involved in H1 and H3/H4 binding [25,26]. Moreover, PpNASPs presents an "E/D-rich" region that interrupts the TPR2 and is involved in H1 binding [25], this region being larger in human (Figure 3C) and mice NASPs. Besides, PpNASP presents a well conserved SHNi-TPR (Sim3-Hif1-NASP interrupted TPR).

The nuclear translocation protein **IPO4** is associated with the cytosolic multiprotein complex involved in H3.1/H4 dimer nuclear import [7]. Another model has been proposed in which H3 and H4 are bound as monomers to IPO4 in the cytosol before being translocated to the nucleus and being addressed in the nucleus as dimers by the NASP, ASF1, p46^{RbAp46} and HAT1 complex [97]. Excepted *Drosophila* that presents tandem duplicate genes giving rise to Artemis and Apollo proteins [98], *Physarum* and studied organisms present one single IPO4. PpIPO4 contains an importin- β N-terminal domain and two HEAT repeats (Figure 3D). Acetylation of K5 and K14 residues of the histone H4 increased affinity with IPO4 [99] and facilitates histone nuclear import, notably in *P. polycephalum* [8] and in yeast [100].

The **ASF1** is highly conserved from yeast to mammals. Two isoforms exist in mammals and plants but *P. polycephalum* contains only one isoform similarly to yeast or *Drosophila*. Alignment of ASF1 proteins confirmed that the N-terminal is highly conserved (50-60% of identity regarding the first 155 amino acids) while the C-terminal is rather divergent and smaller than in yeast similarly to other

metazoans (Figure S6). Moreover, the C-terminal contains several acidic amino acids - mainly glutamate - as it was reported in other organisms [101] but to a smaller extent.

2.2. The H3/H4 chaperones involved in chromatin incorporation

The CAF-1 complex comprising 3 subunits, named PpCAF1A, PpCAF1B and PpCAF1C in *Physarum*. **PpCAF1A** harbors 4 domains: (i) the A domain involved in ScCac3 interaction and flanked by (ii) the “E/D-rich” domain enriched in glutamic acid (E) and aspartic acid (D) involved in H3/H4 interaction, (iii) the KER domain rich in lysine (K), glutamic acid (E) and arginine (R) involved in ScCac2 interaction and (iv) a PIP2 domain (Figure 4A). Besides these domains, Hs-p150 also presents PEST, PIP1 and C-terminal domains as well as the HP1 motif. The PEST domain - rich in proline (P), glutamic acid (E), serine (S) and threonine (T) [102]- is absent in *P. polycephalum*, yeast and plant proteins (Figure 4A). The CAF-1 large subunit associates *via* the PIP domain (PCNA-Interacting Peptide) with PCNA (Proliferating Cell Nuclear Antigen) which triggers CAF-1 recruitment to the replicating sites [102]. While animal CAF1A proteins display two PIPs (PIP1 in the N-terminal region and PIP2 in the KER domain, proteins from the other analyzed organisms only present PIP2 (Figure 4A). Besides, Hs-p150 and ScCac1, respectively harbor a p60-binding and a WHD (Winged Helix domain) involved in DNA binding without sequence specificity [103,104] in their C-termini (Figure 4A). **PpCAF1B** harbors several domains: (i) WD40 repeats involved in the interaction between the middle subunit of CAF-1 and the H3/H4 histones and ASF1 [103] and in the interaction with the large subunit of CAF-1 [105], (ii) a PEST domain comprising a B-like similar domain to the HIRA B domain and (iii) a PIP domain. WD40 repeats are clustered at the N-terminal except one in the C-terminal region for PpCAF1B, ScCac2 and AtFAS2 (Figure 4B). Most animal CAF1B proteins contain two B-like domains (B-like 1 and B-like 2) located in the middle and C-terminal regions of the PEST domain for Hs-p60 while *Physarum*, yeast and plants only display B-like 2 (Figure 4B). **PpCAF1C** harbors 2 domains: (i) two H4-interacting regions located at the N- and C-termini and (ii) several WD40 repeats in the internal region (Figure 4C).

The HIR complex comprising 3 subunits, named PpHIRA, PpUBN and PpCABIN in *Physarum*. **PpHIRA** presents three domains: (i) the N-terminal region composed of several WD40 repeats involved in the UBN1 interaction [106], (ii) the B domain involved in ASF1 interaction and (iii) the HIRA domain involved in CABIN1 binding [33] and HIRA homooligomerization [107] (Figure 4D). The B-domain and the HIRA domain are well conserved (Figure S8B-C). One should note that HIRA protein has two orthologues in yeast, ScHIR1 and ScHIR2 respectively homologous to the N- and C-terminal regions of HsHIRA (Figure 4D). Besides, HsHIRA was also reported to bind H2B and H4 [108]. **PpUBN** presents three domains: (i) a N-terminal domain named NHRD (N-terminal to HRD -

Hpc2p-Related Domain), the (ii) HRD and (iii) ubinuclein middle domains (Figure 4E). UBN proteins present similar domain organization across eukaryotes, except for ScHpc2; the human UBNs display longer C-termini (Figure 4E). The NHRD domain mediates HsHIRA interaction [106,109] through 5 key residues (R46/E53/K74 conserved in most species but E61/F62 only conserved in animals; Figure S8D, in blue). NHRD and HRD domains are located in the N-terminal region of UBN proteins excepted for ScHpc2 in which they localize in the C-terminal (Figure 4E). The HRD domain confers the H3.3/H4 specificity to the HIR complex whereas the ubinuclein middle domain is able to bind both H3.1/H4 and H3.3/4 [32,110]. HsUBN1 also presents a DNA-binding domain which participates in HIR-mediated histone deposition in chromatin [111]. We did not find this lysine-rich domain in PpUBN. As other UBN, PpUBN presents a ubinuclein middle domain which is absent from ScHpc2 and DdUBN (Figure 4E). **PpCABIN1** presents several TPR and TPR-like repeats and HsCABIN1 a MEF2-binding domain (Figure 4F). These TPRs are sufficient for the HsCABIN1-HsHIRA interaction [18]. This MEF2-binding domain was retrieved only in animal CABIN1 (CePICD-1 excepted). The TPRs are sufficient for the HsCABIN1-HsHIRA interaction [33] and HsCABIN1 and HsUBN1 can interact together [109].

2.3. The H2A/H2B chaperones

Proteins of the **NAP** family have various functions beyond their roles in chromatin assembly. ScNAP1 participate to the H2A.Z/H2B variant exchange [15] and AtNRP proteins counteract SWR-C activity in Arabidopsis [111]. More generally, NAP proteins are able to bind H2A/H2B and H3/H4 histones (reviewed in [13]). **PpSET** and **PpNAP1L1** harbor a N-terminal tail of variable length and an acidic C-terminal tail as most NAP proteins. Sequences participating to the NAP hydrophobic core and protein assembly [13] are highly conserved in both *Physarum* NAPs (Figure S9A-B). Finally, the NES and its key residues as well as the NLS [112,35] are conserved only in PpNAP1L1 (Figure S9H-I, Tables S5 and S6). The **FACT** complex is composed of two subunits SPT16 and SSRP1 (or ScPob3) and contributes to chromatin dynamics during transcription, DNA repair and replication. This heterodimer deposits H2A/H2B dimers, H2A.X/H2B dimers and H3/H4 tetramers into chromatin but not H2A.Z/H2B [19]. It also disrupts interactions between core histones as well as DNA/histones interaction, notably during replication. It is thus able to bind either nucleosomes or free histones. **PpSPT16** present four domains: the (i) N-terminal, (ii) dimerization, (iii) middle and (iv) C-terminal domains (Figure 5B), based on the homology with HsSPT16 [113]. N-terminal domains are conserved across species (Figure S10A) and present an aminopeptidase-like “pita bread” fold [114]. Moreover, they are binding modules for H3/H4 in fission yeast [115] and human [116]. The SPT16 dimerization domain enables formation of the FACT complex through an interaction with the N-

terminal/dimerization domain of SSRP1 [117] and presents some similarities across studied species (Figure S10B). The middle domain presents a double pleckstrin homology structure in yeast that is able to bind H3/H4 [39] during nucleosome dynamics operated by FACT and it is conserved across studied species (Figure S10C). Besides, the H2A/H2B minimal binding domain defined in ScSPT16 [39] is poorly conserved across species (Figure S10D, green square) but its key residues are either conserved or substituted by an amino acid of similar properties in animals and *Physarum* (Figure S10D). Finally, the C-terminal domain is an intrinsically disordered region involved in mononucleosome binding and weakening of interactions between H2A/H2B dimers and H3/H4 tetramers [118]. **PpPob3** is the orthologue of ScPob3, the second yeast FACT subunit. SSRP1 proteins display 5 domains: (i) a N-terminal/dimerization domain with a single pleckstrin homology structure [113] involved in dimerization with SPT16 [117], (ii) a middle domain with a double pleckstrin homology structure [113] involved in DNA binding [119], (iii) an HMG (high-mobility group) domain flanked by two intrinsically disordered regions, the (iv) IDD (Intrinsically Disordered) and (v) C-terminal regions (Figure 5C). However, ScPob3 and PpPob3 presents only three domains (Figure 5C) similarly to *T. thermophila* and *D. discoideum* [37] orthologues. The N-terminal region of the HMG domain is involved in DNA binding in human [120] and yeast [121].

2.4. The DNA replication-associated histone chaperones

MCM2 is a component of the DNA replicative helicase MCM2-7 complex. This complex is loaded on double-stranded DNA during late M and G1 phases and participates to DNA unwinding during replication. Moreover, MCM2 can associate with H3/H4-ASF1 during replication [12] and with the histones-FACT complex during transcription [40]. Besides, MCM2 can bind all H3 isoforms (H3.1, H3.3 and CENP-A) [41]. We analyzed the amino acid composition of PpMCM2. The PpMCM2 protein displays an acidic tail (33% aspartate and glutamate in the first 186 amino acids) as the other eukaryotic MCM2 orthologues (~30% for the human, yeast or Arabidopsis MCM2; [40]). PpMCM2 displays five domains: (i) a HBD (histone binding domain), (ii) MCM N-terminal, (iii) OB(oligonucleotide/oligosaccharide)-fold, (iv) MCM and (v) AAA+ that confers the ATPase activity (Figure 5E). The DNA polymerase ϵ also named **PolE** is responsible for DNA synthesis on the leading strand during replication. PolE3 and PolE4 are its small subunits. HsPolE3 and HsPolE4 promote H3/H4 dimers and tetramers deposition into chromatin during replication [122]. ScDPB4, the yeast PolE3 orthologue, interacts with the histone chaperones ScHIR1 and ScNAP1 and is involved in heterochromatin silencing [123]. PpPolE3 presents a histone-like fold (Figure 5F) with 4 helices similar to H2A and H2B ones and an acidic C-terminal (Figure S13A).

2.5. The chaperones associated with histone recycling and exchange

SPT6 is present as one isoform in most organisms such as *Physarum* but *Arabidopsis* presents two isoforms, AtSPT6 and AtSPT6L. The former presents a very low expression level in most tissues while the later displays a wide expression [124] and its chromatin localization was studied [125]. For this reason, we only included AtSPT6L in our study. PpSPT6 presents 8 domains. Interproscan did not retrieve the HtH DNA-binding domain in the *Arabidopsis* protein (Figure 6A, [125]). The N-terminal region of SPT6 is involved in histone interaction. The core region is composed of: (i) 3 DNA-binding domains HtH (Helix-turn-Helix), HhH (Helix-hairpin-Helix) and HHH domain 9, (ii) the YqgF domain from the ribonuclease family and (iii) the S1 RNA-binding domain (Figure 6A). The C-terminal region of SPT6 is involved in the binding to the Ser2-phosphorylated RNA Polymerase II thanks to the SH2 (Src-homology 2) domain. Moreover, AtSPT6L presents a plant-specific WG/GW domain required for Argonaute interaction and small RNA-mediated gene silencing [126] also present in PpSPT6 (Figure 6A). To perform chromatin remodeling, substitution of nucleosomal canonical histones by histone variants, the so-called “histone replacement”, takes place. The H2A.Z/H2B variant exchange is performed by the **SWR-C** complex [15,14,16-18] which destabilize the chromatin to allow the ATP-dependent substitution of H2A/H2B dimers by H2A.Z/H2B dimers. **PpSWR1** is the *Physarum* orthologue of ScSWR1, DmDomino [127] or AtPIE1 (Photoperiod-Independent Early flowering 1, [128]). To note, two orthologues ScSWR1 have been identified in human, HsSRCAP (Snf2-Related CBP Activator Protein; [129]) and Hs-p400 (400KDa protein, [130]) as well as in mouse, fish and xenopus. Several members of the SWI2/SNF2 protein family were first identified from genetic studies conducted in *S. cerevisiae* (SWI refers to yeast mating type *switching*, while SNF is an abbreviation for Sucrose Non Fermenting). These proteins behave as chromatin remodelers and display ATPase and Helicase activities. **PpSWR1** presents 4 characteristic domains (Figure 6B): (i) a HSA (Helicase/SANT-associated) domain conserved in studied orthologues; an ATPase domain split in two segments (ii) SNF2 and (iii) Helicase domains, both being highly conserved and responsible for SWC2 binding [131,45]; (iv) a well-conserved SANT (Swi3/Ada2/N-Cor/ TFIIB) domain in the C terminal. The HSA domain is located at the N-termini of most proteins excepted for p400 ones which harbor a longer N terminal (Figure 6B). The AtPIE1 N-terminal region binds to H2A.Z but not canonical H2A [18] and the N-terminal region of ScSWR1 is responsible for H2A.Z binding [45] but residues identified in yeast for their role in H2A.Z binding [46] are not conserved across studied species. Besides, some SWR1 orthologues present a SANT domain while SCRAP, yeast, *C. elegans*, *T. thermophila*, *A. trichopoda* and maize proteins do not (Figure 6B and S15D). Instead, SRAP proteins have A/T-hook motifs and harbor a long spacer between the SNF2 and Helicase domains (Figure 6B; [132]). Moreover, p400, DmDomino

and DdSWR1 proteins present polyglutamines in their C-termini downstream of the SANT domain (Figure 6B) while PpSWR1, AtPIE1 and PpaPIE1 do not. **PpSWC2** is the *Physarum* orthologue of ScSWC2 and of animal YL1 proteins, which are present in both SRCAP and p400 complexes [16]. The ScSWC2 N-terminal region binds H2A.Z [57] and is responsible for DNA recognition [133]. Similar to ScSWC2, PpSWC2 contains ~40% of charged residues. PpSWC2 presents N- and C-terminal regions characteristic of SWC2 proteins as well as the Z domain (Figure 6C). **APLF** belongs to the Aprataxin and PNK (polynucleotide kinase) family which are DNA-break response factors. This family is characterized by the presence of a FHA (ForkHead-Associated) domain [134] which interacts with the DNA-damage repair facilitators XRCC1 (X-Ray Cross Complementing group 1) and XRCC4. PpAPLF presents a single PBZ located at the very end of the protein while HSAPLF has tandem PBZ domains (Figure 6D).

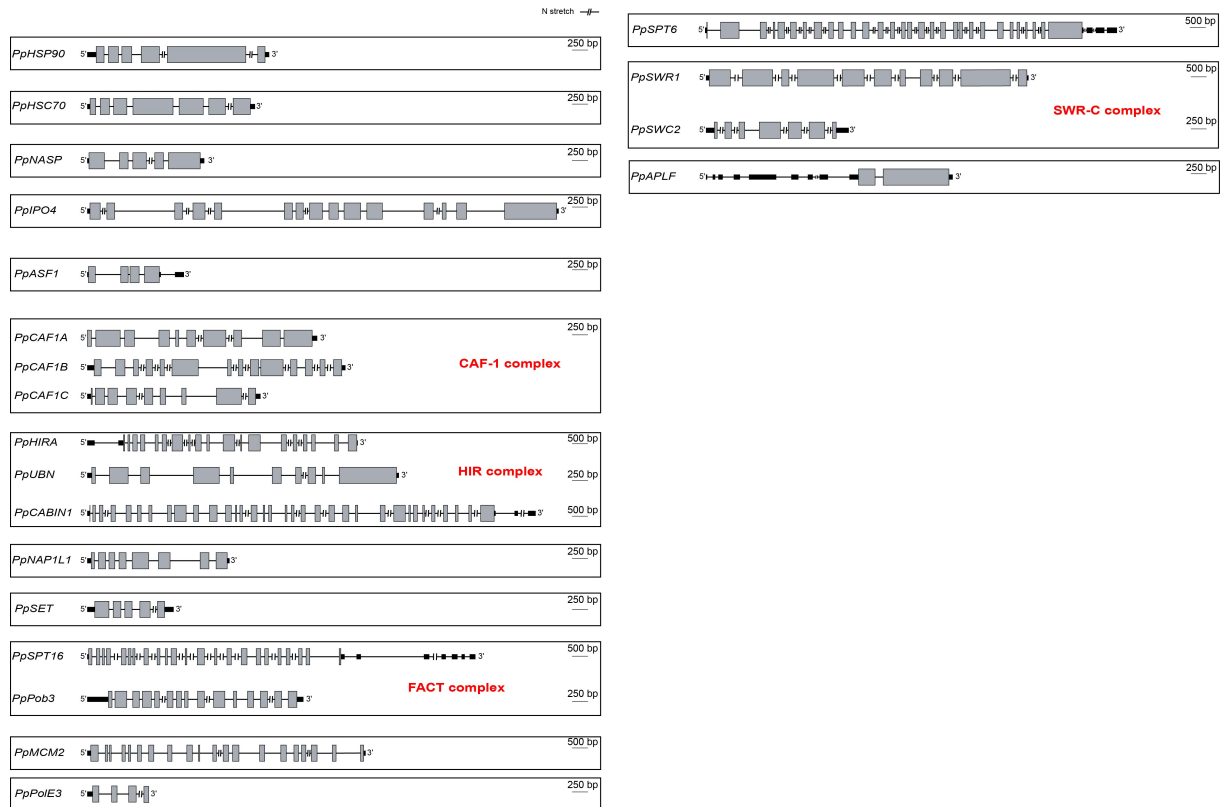


Figure S1. Structure of the various *Physarum* histone chaperone genes. Exons, grey rectangles; untranslated regions, black rectangles. Presence of N stretches are indicated by an interrupted line as indicated on the top right of the figure.

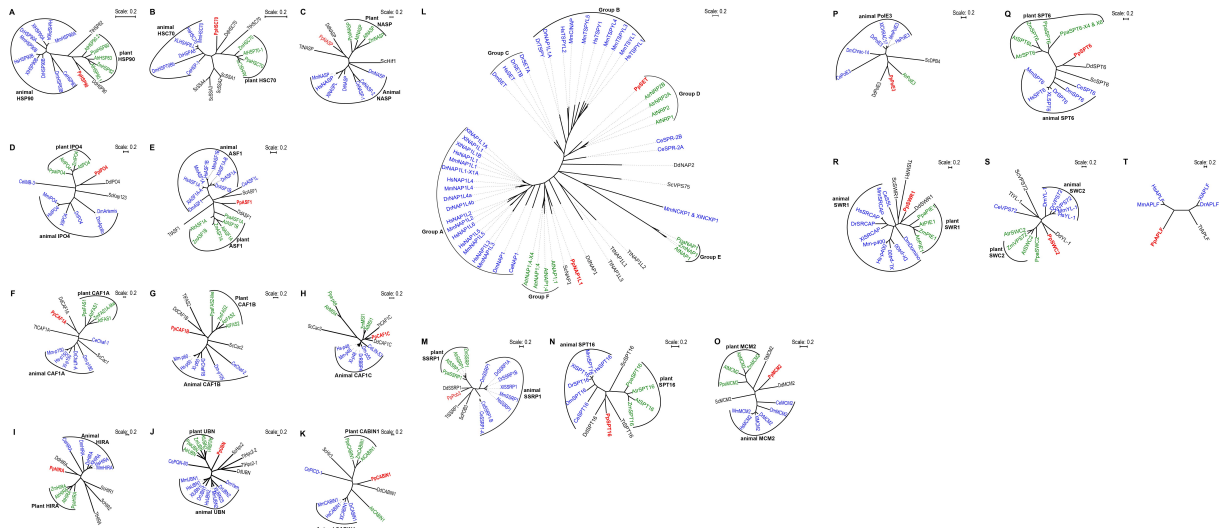


Figure S2. Phylogenetic trees of *Physarum* histone chaperones. Phylogenetic tree of the chaperones HSP90 (A), HSC70 (B), NASP (C), IPO4 (D), ASF1 (E), CAF1A (F), CAF1B (G), CAF1C (H), HIRA (I), UBN (J), CABIN1 (K), of the NAP family (L), SSRP1 (M), SPT16 (N), MCM2 (O), PoIE3 (P), SPT6 (Q), SWR1 (R), SWC2 (S) and APLF (T) from *P. polycephalum* and various eukaryotes. The NAP family (L) is very large and displays a complicated phylogenetic history. Given the high number of NAP proteins in the selected organisms, we condensed highly similar copies in each of the species (between 92 to 99% of identity over the whole protein). All the animal NAP protein cluster together into a group that we named group A, the ScVPS75 protein being found in this group. The animal TSPY proteins cluster into a group that we named group B. The SET chaperones constitute a single clade that we name group C, excepted for PpSET which belongs to the same clade than the plant NRP proteins that are closely related to HsSET. The plant kingdom presents three groups of NAP proteins: (i) group D comprising the NRP proteins and PpSET, this *Physarum* protein presenting a longer branch suggesting a lower homology with the plant NRPs; (ii) group E comprising three proteins and arising from the same clade than the mouse and *Xenopus* NCPK1 but the long branches suggested a lower homology between them; (iii) group F comprising four plant proteins and PpNAP1L1 presenting a longer branch suggesting a lower homology. The three *T. thermophila* NAP proteins group together in a single clade as well as the *C. elegans* SPR proteins. We found two NAP protein in *D. discoideum* and both do not belong to any other clade and are not evolutionary related to each other. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

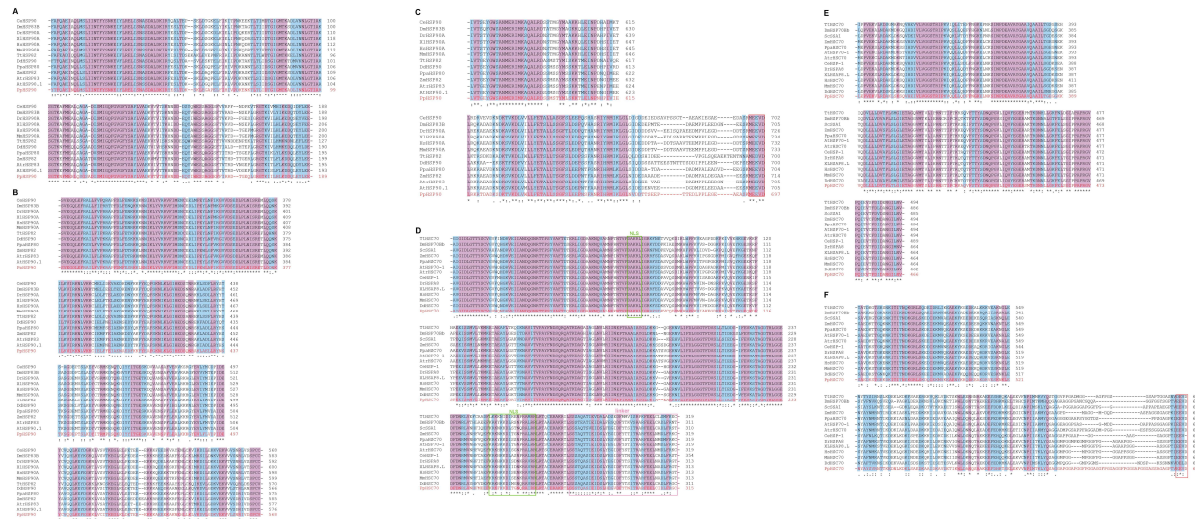


Figure S3. Protein alignments for the *Physarum* HSP90 and HSC70 heat shock proteins. **a to c.** Alignments correspond to the N-terminal (A), middle (B) and C-terminal (C) domains for the HSP90 proteins. The C-terminal region of PpHSP90 contains a strictly conserved MEEVD motif responsible for interaction with TPR-containing proteins [135] (C) which is displayed by a red box. Alignments are performed with HSP90 proteins used in the phylogenetic tree reported in Figure S2A, only one HSP90 protein per organism being conserved for an easier reading. **D to F.** Alignments correspond to the NBD and linker domains (D) and to the SDB β (E) and SDB α (F) domains for the HSC70 proteins. Two NLS (Nuclear Localization Signal) were reported in HsHSP70 [136], the first only being strictly conserved while the second one is less: both NLS signals are indicated by a green rectangle and the linker by a pink rectangle (D). The HSC70 EEVD motif involved co-chaperone binding is displayed by a red rectangle (F). Alignments are performed with HSC70 proteins used in the phylogenetic tree reported in Figure S2B, only ScSSA1 being used among yeast homologues for an easier reading. The *Physarum* proteins appears in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

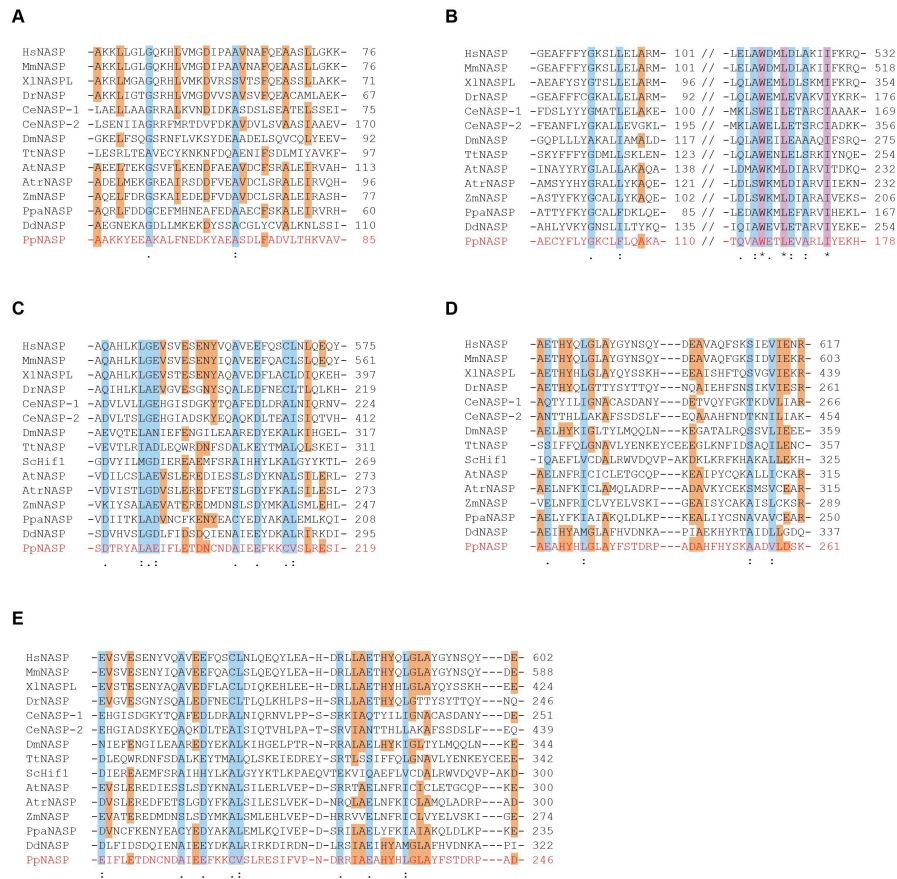


Figure S4. Protein alignments for the *Physarum* NASP protein. Alignments correspond to the TPR1 (A), TPR2 (B), TPR3 (C), TPR4 (D) and SHNi-TPR (E) repeats for the NASP proteins. Several TPR residues were previously described to be almost unchanged through evolution [137] which we could also observe. Alignments are performed with NASP proteins used in the phylogenetic tree reported in Figure 2C. ScHif1 was removed from TPR1 and TPR2 for an easier reading. For the interrupted TPR2, the TPR2-1 and TPR2-2 alignments are displayed together and separated by a double slash. The *Physarum* protein appears in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Almost conserved residues are labelled by an orange rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*; Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

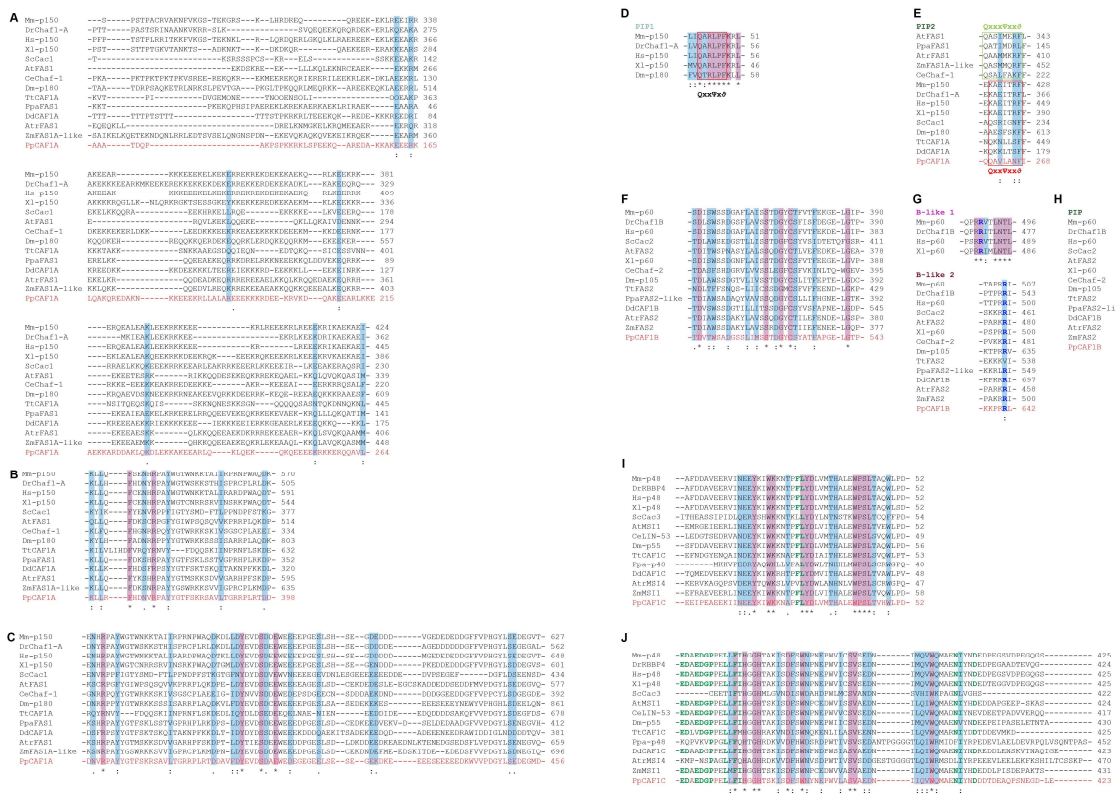


Figure S7. Protein alignments for the *Physarum* CAF-1 complex. **A** to **E**. CAF1A Alignments correspond to the CAF1 domain (**A**) involved in the interaction with ScCac2, the A domain (**B**) involved in the interaction with ScCac3, the E/D rich region (**C**) involved in the interaction with H3/H4 [139], the PIP1 (**D**) and PIP2 (**E**) boxes. Alignments are performed with CAF1A proteins used in the phylogenetic tree reported in Figure S2F. **F** to **H**. CAF1B alignments correspond to the PEST domain (**F**) which is strongly conserved for the studied organisms, the B-like1 (top) and B-like 2 (bottom) domains with the key residues for CAF1B binding to ASF1 depicted in blue (**G**), the PIP box (**H**). Residues R483 and R499 from Hs-p60, respectively located in the B-like 1 and 2 domains, are involved in the interaction between Hs-p60 and ASF1 [9]: these residues are displayed in blue in **g** and both are highly conserved. The PIP domain of PpCAF1B is more degenerated in sequence compared to the ones in CAF1A proteins (**G**). Alignments are performed with CAF1B proteins used in the phylogenetic tree reported in Figure S2G. **I-J**. CAF1C alignments for the CAF1C proteins correspond to the N- (**I**) and C-terminal H4-interacting regions (**J**) with the key residues for CAF1C binding to H4 depicted in green. Based on the crystal structure, several residues were previously identified for their role in the interaction with the histone H4 (F29/L30 and F367/I368 in **i** and I407 of the Hs-p48 hydrophobic patch in **j**; E356/ D357/ E359/ D360/ G361/ P362/ L365/ N406/I407/ D410 in **j** that form hydrogen bonds with specific residues of H4). These residues are displayed in green (**I** and **J**) and they all are strictly conserved in PpCAF1C. Alignments are performed with CAF1C proteins used in the phylogenetic tree reported in Figure S2H. The *Physarum* proteins appear in red. Positions refer to the mature protein

with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

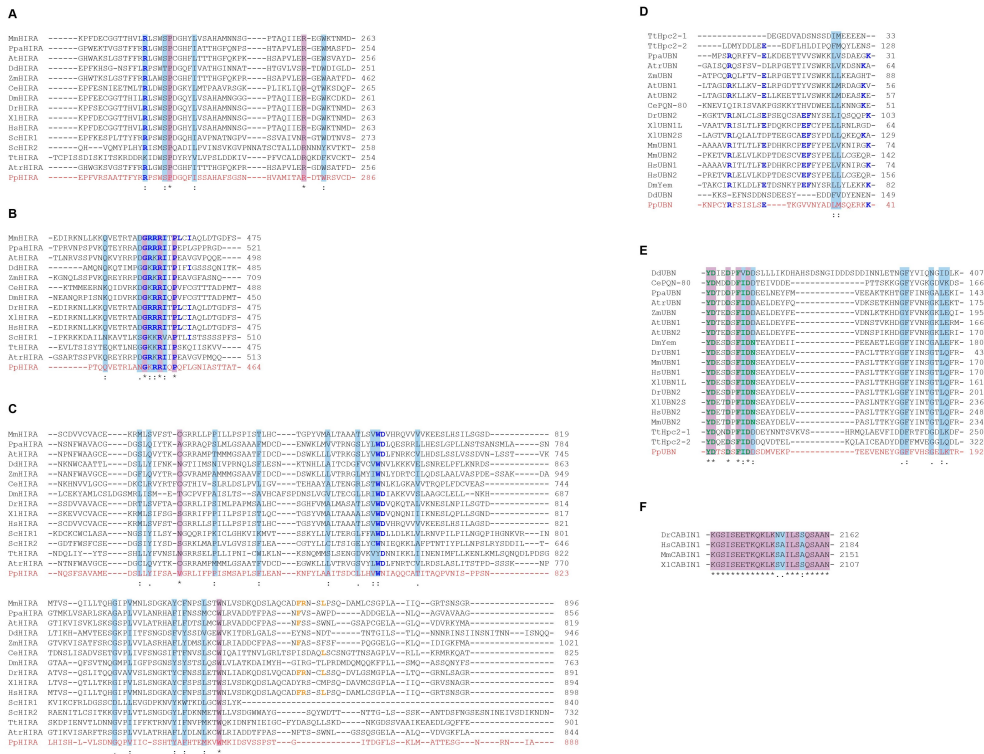


Figure S8. Protein alignments for the *Physarum* HIR complex. **A.** Alignment correspond to the N-terminal domain for the HIRA proteins. The R227-HsHIRA residue is involved in HsHIRA-HsUBN1 interaction [106]. This residue is displayed in blue and highly conserved. **B.** Alignment correspond to the B domain for the HIRA proteins. The GRRRIxPLxI (with x is any amino acid) motif is involved in ASF1 binding [9] and corresponding residues are in blue. This motif is highly conserved except the last isoleucine. **C.** Alignment correspond to the HIRA domain for the HIRA proteins. The W799/D800-HsHIRA residues reported for their putative role in homooligomerization [107] are depicted in blue and those for CABIN1 binding (with the N-terminal region of CABIN1 containing TPR repeats, [33]) in orange (F870/R871/L874-HsHIRA). Only alignment of the beginning of the HIRA domain is displayed since the end is less conserved. Alignments are performed with HIRA proteins used in the phylogenetic tree reported in Figure S2I. **D.** Alignment correspond to the NHRD domain for the UBN proteins. Several residues (R46/E53/E61/F62/K74) were identified in the HsUBN1 NHRD domain for their role in the interaction with HsHIRA [109]. These residues are depicted in blue. **E.** Alignment correspond to the HRD domain for the UBN proteins. The HRD domain confers a specificity for H3.3/H4 binding and key residues were identified (Y132/D133, D136, F138/I139/D140/N141) for their role in histone binding [32]. These 7 residues are displayed in green. Alignments are performed with CAF1B proteins used in the phylogenetic tree reported in Figure S2J. DrUBN1 was removed from the NHRD alignment since the sequence available in databases is truncated and does not have such a domain. ScHpc2 was removed from both alignments since both domains are localized in the C-terminal. **F.** Alignment correspond to the MEF2-binding domain retrieved only in animal CABIN1.

Alignments are performed with animal CABIN1 proteins used in the phylogenetic tree reported in Figure S2K. The *Physarum* proteins appear in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

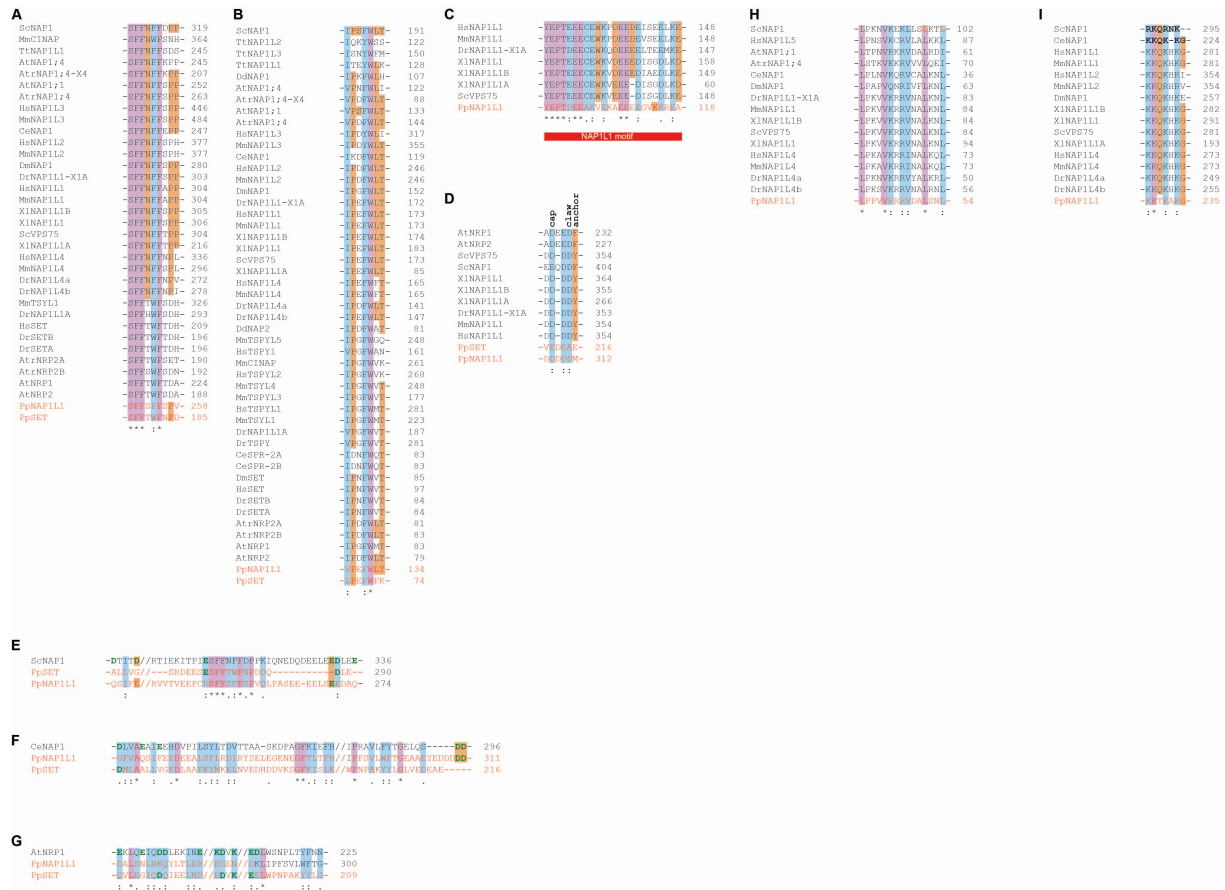
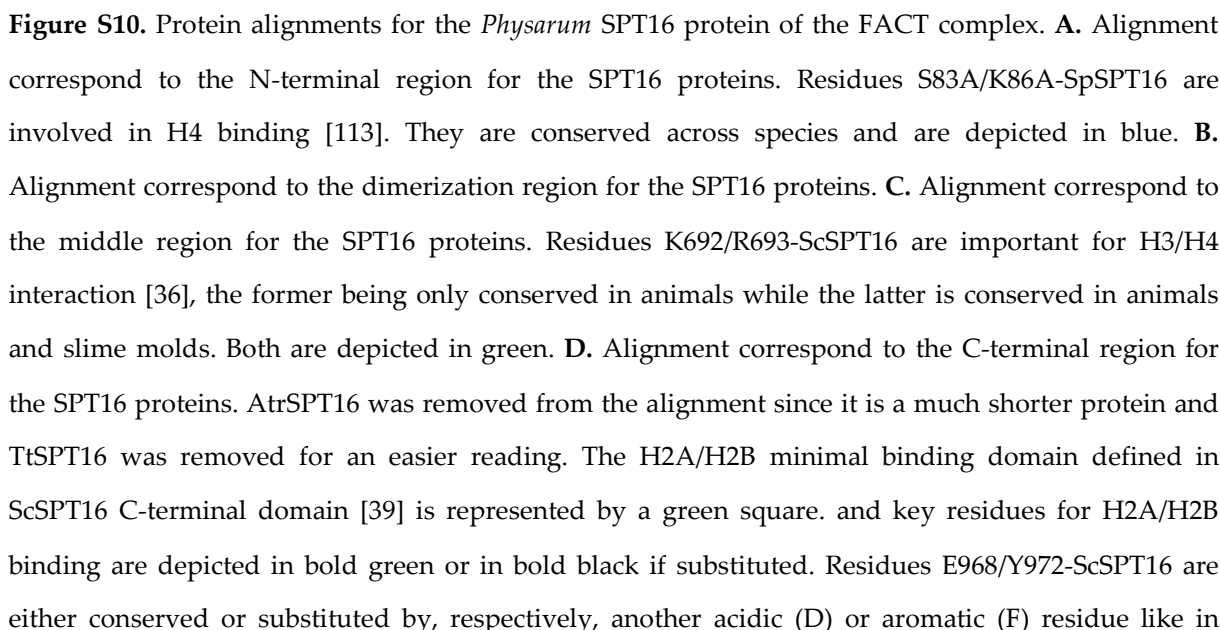


Figure S9. Protein alignments for the *Physarum* proteins of the NAP family. **A-B.** Alignments with the SFFNFFDPP and IPSFWLT sequences of ScNAP1 that participate to the NAP hydrophobic core and protein assembly [13]. Both sequences they are highly conserved among NAP family members, especially in PpSET and PpNAP1L1. **C.** Alignment for the NAP1L histone-binding motif [49]. It was retrieved in PpNAP1L1 but not in SET proteins since this motif is known to be absent from SET proteins. **D.** Alignment for the Cap-Claw-Anchor (cap residue: E/D/S; claw residue: D/E; anchor residue: Y/F; [34]) motif located in the C-terminal acidic tail of some NAP protein. This motif is also important for H2/H2B binding but only the cap and claw residues are conserved in PpNAP1L1 and PpSET. **E to G.** Alignments of the PpNAP1L1 and PpSET proteins with the histone binding modules of NAP proteins from yeast (**E**), *C. elegans* (**F**) and *Arabidopsis* (**G**). Mutation of residues D201/ D205/ E310/ E332/ D333/ E336-ScNAP1 (**E**), D125/ E129/ E132/ D295/ D296 of CeNAP1 (**F**) and K115/ D116/ K118/ E213/ D214 of AtNRP1 (**G**) reduce H2A/H2B binding [34,35]. Those residues involved are depicted in green and only few residues are conserved in *Physarum* NAP proteins which suggests PpNAP1L1 and PpSET present distinct surfaces for H2A/H2B binding. **H.** Alignment correspond to the NES signal for NAP proteins that present such signals. NES enables cytoplasmic localization [112]. Mutations of L99/L102-ScNAP1 prevent its cytoplasmic localization [112]: these two residues are depicted in orange and are conserved in PpNAP1L1 and in some other NAP proteins but not in PpSET. Those findings were further confirmed by predictions tools (Tables S5 and S6 for NSL and

NES). **I**. Alignment correspond to the NLS signals for NAP proteins that present such signals. AtNRP1 was removed from the alignment since its NLS signal is too divergent (**I**). In bold, are presented the NLS signals described in [35]. Alignments are performed with NAP proteins used in the phylogenetic tree reported in Figure S2L. The *Physarum* proteins appear in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Almost conserved residues are labelled by an orange rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.



PpSPT16. They are depicted in green if conserved or in bold black if substituted to an amino acid of similar properties. Alignments are performed with SPT16 proteins used in the phylogenetic tree reported in Figure S2M. The *Physarum* protein appears in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

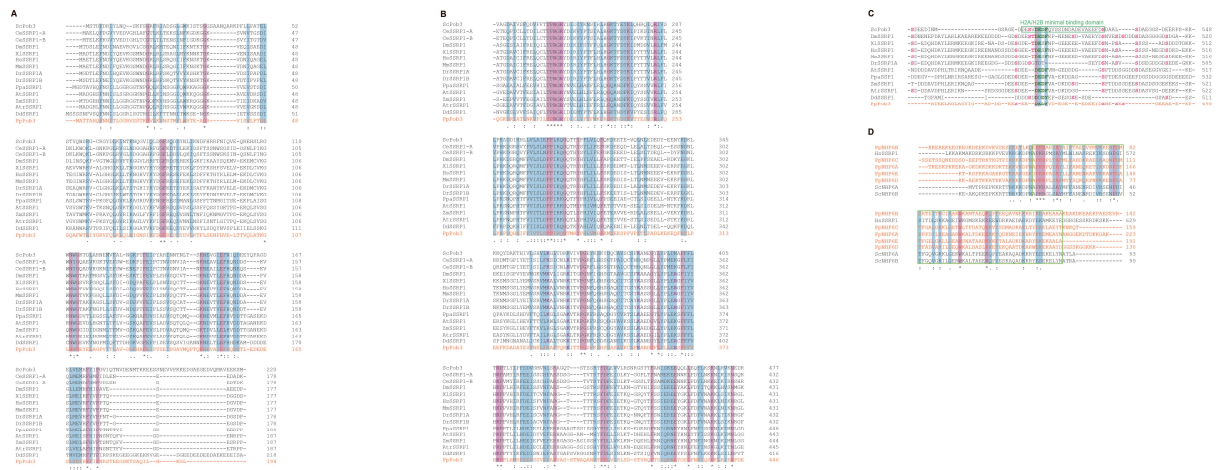


Figure S11. Protein alignments for the *Physarum* Pob3 and NHP6 proteins of the FACT complex. **A.** Alignment correspond to the N-terminal and dimerization regions for the Pob3/SSRP1 proteins. **B.** Alignment correspond to the middle region for the Pob3/SSRP1 proteins. Residues R211/R213/R241/K319/K325/K346/R357/K364-HsSPT16 involved in DNA binding [119] are well conserved and depicted in purple. **C.** Alignment correspond to the IDD region for the Pob3/SSRP1 proteins. The H2A/H2B minimal binding domain defined in ScPob3 IDD region [39] is represented by a green square. It is poorly conserved across species. Key residues of the H2B-binding motif (D/E)xxQ (where Q is F or Y, and x is any residue) [39] are depicted in green and those for nucleosomal DNA binding in pink. The H2B-binding motif is strongly conserved in SPT16 proteins, especially in PpSPT16 and in DdSSRP1 [37]. Serine phosphorylation in DmSSRP1 C-terminal region inhibits nucleosomal DNA binding. These residues are depicted in pink and conserved but to a smaller extent in PpPob3. **D.** Alignment correspond to the HMG region for the Pob3/SSRP1 proteins. The HMG box is depicted by a light green box. The yeast and *Physarum* Pob3 proteins lack the HMG box present in HsSSRP1. The HMG box is located in the NHP6 proteins of yeast and *Physarum*. Alignment is performed with NHP6 proteins reported in Table S2. Alignments are performed with Pob3/SSRP1 proteins used in the phylogenetic tree reported in Figure S2N. TtSSRP1 was removed from the alignment for an easier reading in Figure S11A-C. The alignment in c was performed according to [140]. The *Physarum* proteins appear in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

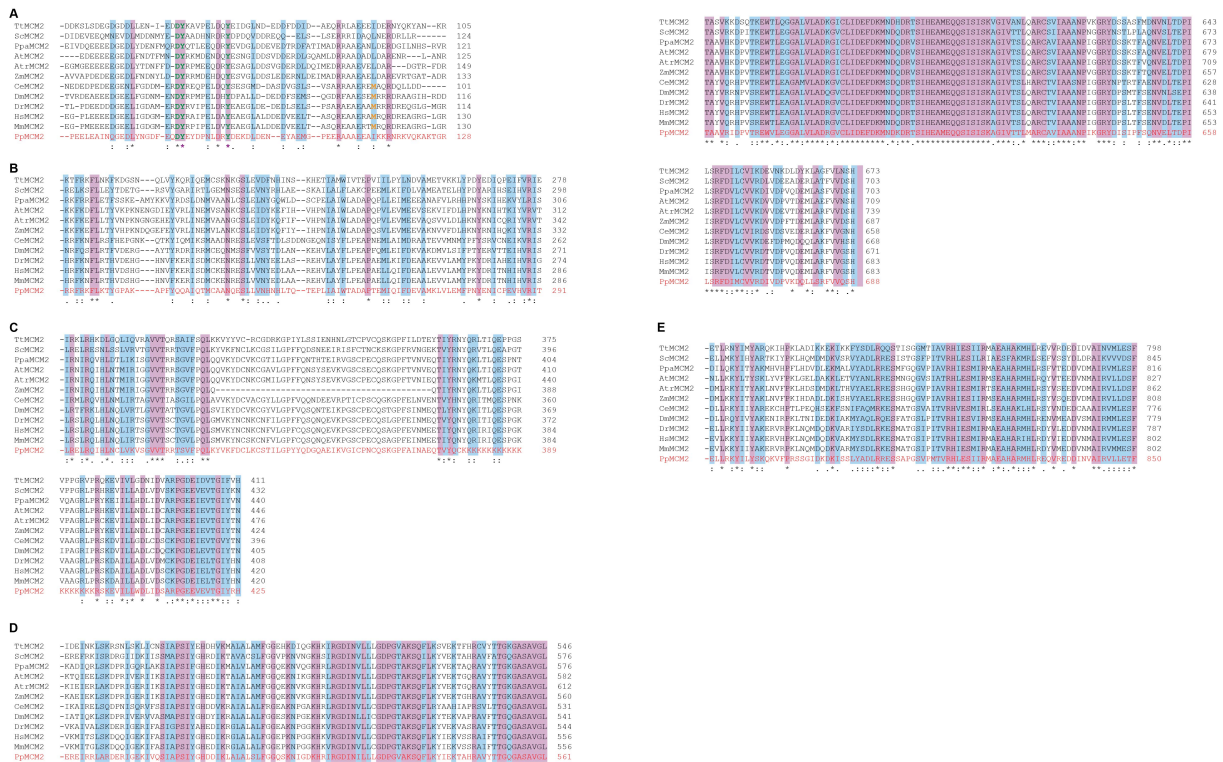


Figure S12. Protein alignments for the *Physarum* MCM2 protein. Alignments correspond to the HBD (A), MCM N-terminal (B), OB-fold (C), MCM (D) and AAA-lid (E) domains for the MCM2 proteins. The D80/Y81/Y90-HsMCM2 residues are involved in H3/H4 binding; they are conserved in all analyzed species including *Physarum* and are displayed in green (A). The Y81/Y90-HsMCM2 residues are also involved in the MCM2 association with the histones-FACt complex [12]. Both are conserved in all analyzed species which is represented by a purple asterisk (A). Alignments are performed with MCM2 proteins used in the phylogenetic tree reported in Figure S2O. The *Physarum* protein appears in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*; Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

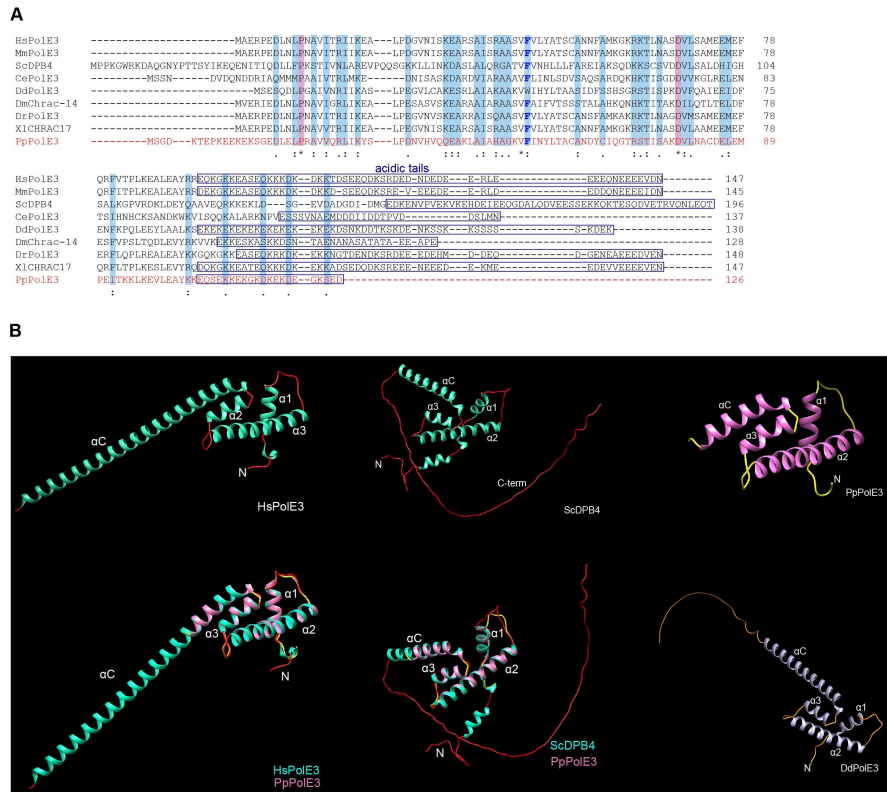


Figure S13. Analysis of the PolE3 chaperone. **A.** Alignment correspond to the PolE3 proteins. The C-terminal region of PolE3 is required for interaction with H3/H4 and the F44 residue involved in the interaction between HsPolE3 and HsPolE4, HsPolE1, HsPolE2 [122] is conserved in PpPolE3 and displayed in blue. Acidic tails are indicated by blue rectangles. Alignment is performed with PolE3 proteins used in the phylogenetic tree reported in Figure S2P. The *Physarum* proteins appear in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*. **B.** Ribbon representation of HsPolE3, ScDPB4, PpPolE3 and DdPolE3 proteins. The α C helix characteristic of H2B family is much shorter in PpPolE3 than in HsPolE3 and that HsPolE3 presents a long α C helix while ScDPB4 presents a long C-terminal tail. Representation of each chaperone is displayed on the upper panel, excepted for DdPolE3 presented in the right lower panel. Superimpositions of HsPolE3 and ScDPB4 with PpPolE3 are presented in the lower panel. The various helices (α 1, α 2, α 3, α C) and the N-terminal tail (N) are indicated. Models were generated with Phyre2 for PpPolE3 and AlphaFold for HsPolE3, ScDPB4 and DdPolE3. Superimpositions were done with the Chimera software.

A

PpaSPT6	-YSVGS-YGYSA-----GSRPSESAS-----	1483
AtrSPT6	-GSSSGDG GW RGRSSFDRLRS--TPGSRTGRHGNNARDGHPGVPRPYG---QG-RGRGRNDEGN	1529
AtsPT6L	-HSSSGGS GW SSSQSE GW KGNSDRSGSGRGGEYRNGGGRDGHPSGAPRPYGGRRGR-RGRGRDDMNS	1502
ZmSPT6	- W SGG-----GGDANG GW RGDSNDR-----DRPFSGRSGGRFDSRNSSGGRGRG-RGRGRGN---	1509
PpSPT6	-MST-----PQINATPAHPGAATPAHPGNATPAHPGNATPAHQSG GW ANAAQGGNNFPGM	1518
.		
PpaSPT6	GFDSRNSQ W GDSR--AGG GW GSS-----S GW SR--GNS GW DKAV-----VR	1522
AtrSPT6	DSGYGSSK W SGAN-DSGG GW SGSKDSGG GW SGSKDTGG GW SGSKDTGGG W SGSKE	1588
AtsPT6L	DRQDNGND W GNNDTGTAD GW GN-----GGG GW SES-----	1535
ZmSPT6	-----FGDD--DNAG GW SGG-----GGG--	1527
PpSPT6	QAM--QQNLPATSSLSL W GWDSGTST---PGHSQGGAG W NNNNNTNNGNNNAFAN	1572
***. . *		
PpaSPT6	PSG-----DN-----RGTSN W GPEPSPRNNW	1544
AtrSPT6	SGGG W SGLSGHKSSESFSRDDLQAGG W SGSAG-----GGRNSD W GGGGGGGG	1638
AtsPT6L	-----AGKK-----TG-----GGSTG W SGSESNGKS-	1557
ZmSPT6	-----NSGRWTDNIGT-GG-	1540
PpSPT6	-----P---TNDVT W GWEDNNNNNNNNNNYGNSSNSNYGQNDYGNRR	1612
. . . .		
PpaSPT6	DKNGGAGRSAWDNSQPGAPSQWDNVT-----	1570
AtrSPT6	-SGGKSSSDWVGSGSGRNSE W SGESGQNSDVGVDAGSGRNND W GGGGSGSGRN-ND W	1696
AtsPT6L	-DG-----AGS W SGS-----	1567
ZmSPT6	-SG-----GSS W GGGGD-----	1553
PpSPT6	DDR-----SRDN-----NSNNYGNRD---NYGNRGNSYRDNNNNYNSNRDSSRY	1655
.		
PpaSPT6	-----NSEGGAAGD W AAVPGATPT-----SGS W DAVP--	1599
AtrSPT6	GGGGSVAGGKGPNSD W GGGGGGGGSGGGGKGPNSD W GGGGGGGGGKGPNS W GGGGG-	1755
AtsPT6L	-----GGGGSG W GNDSGGKKSSD---GGFGSGSGGGSD W GNES--	1605
ZmSPT6	-----DSNRGGGGG W GSAGGGAD-----NAGGNG W AGAAA--	1585
PpSPT6	GGGYGGGGG-DN-YRSGGGGGGGGGG-----GGYGDRNR	1692
* . . . * . *		
PpaSPT6	---GAT---P-----TSGS W MPGGHAGPRGF---	1620
AtrSPT6	-GSGGGE---KGPDGD W G--GGSGGGGKGPDGD W GGDS W NSGKKGGSSGAWPES	1807
AtsPT6L	---GGKK---SSADGG----- W GS-----ESG	1621
ZmSPT6	---GGSD---DPGLGS-----A-----KKV	1599
PpSPT6	DSRGGRDNNNNRRDRDYNNRSGGGGGGGGGGDRNRD-----SRD	1733
*		
PpaSPT6	-----	1620
AtrSPT6	GGGNSGGG W GNASTTTSATNVGGG W A-----	1834
AtsPT6L	GKSDGEG W GNPESSRKSDGGGG W -----	1647
ZmSPT6	VPAQDGGT W GS-----EGGG W -----	1617
PpSPT6	NRGRDGG W DDRDKGSGGGGGGGGRRSGGEGFEISAQQVFHGLLSFPLVPFWKSL	1791

B

ScSPT6	-SDKIDEMYDIFGDG-HDYDWALEIEN---EELN---GNDNN-EAEEEEIDEET--GA-----IKSTKKKISLQDIYDLEDLKKNLMTEG	314
CeSPT6	-EGAEDDARDVFGVEDFNLDE-----FYDDDDGEDGLEDEEEEIIEDDGEGGEIKIRKKDTTK---KSTLLESIEPSEIDRGFLLPG	292
DmSPT6	-DASLQEGQDIFGV-DFDYDD-----FSKYEEDDYEDDS-EGDEYDEDLGVGDDTRVKKKKKALKKKVVKKTIFDIYEPSELKRGHFTDM	263
DrSPT6	-DAALQEAQEIFGG-DFDFAE-----FDTEAYDHAEE-----EE-EEDQDDESWDPRPKQTKRRVSRRSIFEIYEPSELESSHMTDQ	280
XLSP6	-DAALQEAQEIFGV-DFDYDE-----FEKYDEDEDEM-E-EYD-YE-DDEGDAETVRPKKTAKKRVSRRSIFEMYEPSELESSHLTDQ	283
HsSPT6	-DAALQEAQEIFGV-DFDYDE-----FEKYNEYDEELEEE-EYE-YE-DDEAGEIRVRPKKTTKKRVSRRSIFEMYEPSELESSHLTDQ	273
MmSPT6	-DAALQEAQEIFGV-DFDYDE-----FEKYNEYDEELEEE-DYE-YE-DDETEGEIRVRPKKTTKKRVSRRSIFEMYEPSELESSHLTDQ	281
PpaSPT6	-SIALQEAQEIFGDVTSLEERRRRERDTLAGVDDEDRVDD-----DD---DDGYGR-----SRKPPSQKLEKQFEPSSLLEKFLTER	364
AtrSPT6	-SSALQEAHEIFGDVDELLRFRKLGHGKAGDGFDDVD-----DS-----VDGPKRLEDEFEPSEILEEKYMTKE	286
AtsPT6L	-ITAMRDANEIFGDVDELLTIRKKGLA-----S-----NQRMERRLEDEFEPSEILEEKYMTGN	286
ZmSPT6	-SSALQEAQDIFGDVDELLALRKQELEREL-----NS-----GELRGARLEDEFEPSEILEEKYMTSK	338
PpSPT6	-RGALQDAEDT EG SVDFFKQASGRSRPRGDEFDDGDYDGTQGLSELDLDAEIEEGD-----EKQKDLPLAD Y EHYEPSSL EQ KHFTSE	275
: : **		

Figure S14. Protein alignments for the *Physarum* SPT6 protein. **A.** Alignment correspond to the WG/GW domain for the SPT6 proteins. The WG/GW domain was described as a plant-specific WG/GW domain required for Argonaute interaction and small RNA-mediated gene silencing; this domain is surprisingly also present in PpSPT6 (f, WG/GW in purple). **B.** Alignment correspond to the H3-binding region for the SPT6 proteins. The H3-binding region corresponds to the region 239-314AA of ScSPT6 [44] which is located between the acidic N-terminal domain responsible for Iws1 (Interact-With-SPT6) binding and the HtH domain. The F249-ScSPT6 residue involved in nucleosome binding is depicted in green (B). Alignments are performed with SPT6 proteins used in the phylogenetic tree

reported in Figure 2B. The *Physarum* proteins appears in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

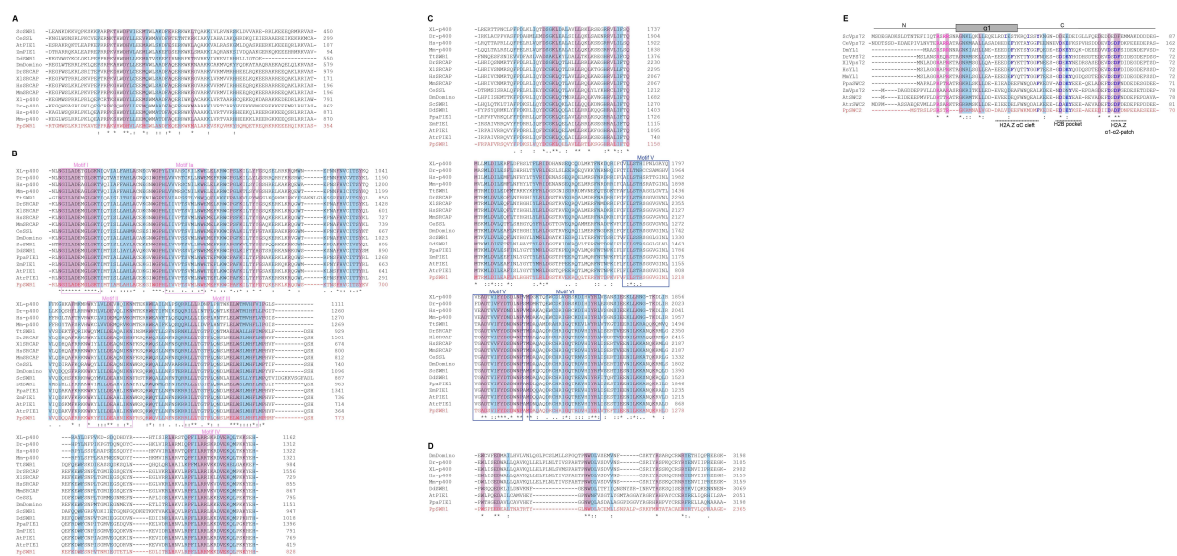


Figure S15. Protein alignments for the *Physarum* SWR-C complex. **A to D.** Alignments correspond to the HSA (**A**), SNF2 (**B**), Helicase (**C**) and SANT (**D**) domains for the SWR1 proteins. No HSA domain was found neither in AtrSWR1 nor in TtSWR1 with Interproscan. Five of the seven motifs (I and Ia to IV) that constitute the ATPase domain are located in the SNF2 region (**B**) and two (V and VI) in the Helicase domain (**C**). Alignments are performed with SWR1 proteins used in the phylogenetic tree reported in Figure S2Q. **E.** The helix $\alpha 1$ present in the Z domain is indicated as well as the N-terminal (N) and C-terminal regions. Dashed lines indicate regions which interact with the H2A.Z–H2B dimer (namely H2A.Z α C cleft, H2B pocket and H2A.Z $\alpha 1$ – $\alpha 2$ -patch). The Z domains (Figure 6C) responsible for H2A.Z binding and selectivity as well as DmYL1 key residues [47] are widely conserved (**E**, in blue). The RxxR motif (x for any residue) anchors ScSWC2 to the nucleosome surface [48] and its arginines are displayed in pink. Alignments are performed with SWC2 proteins used in the phylogenetic tree reported in Figure S2R, TtYL1 and DdYL1 were removed from alignment due to lower conservation. The *Physarum* proteins appear in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

A

```

TtAPLF -LKEHILGRRGDLKVD---NPKVSGQHCVLKYDYAQKKAYIIDVSSNGTS-----LFNKKLEKNKEVELENGDLVNLQDKS----- 100
DrAPLF -NSECTVGRKKDCDLSPANKLVSGNHCKITHDQNSGKQVWLEDMSNGTV-----INMSKVVKQTHLLQNGDVIYFYVRKNEPEQNIAYVYQ- 114
XlAPLF -KKEWTIGRKKACDLSPGKLVSGEHCKITVNEESGEVSLDSTNGTV-----INKLKVIRKQTYPLKNGDVIYVYVRKNEPEQ----- 110
HsAPLF -PGETVIGRGLLGIT---DKRVSRRAHILEVAGGQLRIKP--IHTNPCFYQSSEKSQLPLKPNLWCYLNPGDSFSLVDKYI-----FRIL- 101
MmAPLF -PGQTVIGRGLLGIT---DKRVSRRAHILEVVDSQLRIKP--IHRNPCFYQSSEKSQLHPMETQVWSQLHPGDSFSLLLDKYA-----FRVF- 98
PpAPLF -GGPTELGRGL---VN---NKKLSRKQCVVETVDDGVELTP--LGVNPMTMSRAGEDVEILSRNEKYTLHEGDSFTLCGEQYR-----FVLQ- 133
      : **      :      :      : * . . . :      .      *      :      :      * . ** . . :

```

B

```

                                HBD
                                cap-anchor  NAP1L1 motif  KR-motif
HsAPLF -DEDN---DNVGQPN---YDLND--SF--LDDEEDYEPT-DESDSWEPK---EDEEKDVEELLK-EAKRFMKRK- 512
MmAPLF -DEDD---DDVGQPSD-----DEDEEDYEPT-DESDSWHPGK---DDEEQEDVDELLK-EAKSSLHLKH- 491
XlXRCC1 -DEDTEDENTRIVPSKANH-----KPKQEDYDASTDEE---ATGERQQHEDDSGEDTEDEL-RFQEEKQGKK- 521
DrXRCC1 -DESEETPKAKTPEKKPVTPKKQPVVEEEEEEEYGGSTDEE---GPG---DDGSGMDTEDELN-RVEMASRKKK- 490
PpXRCC1 --SDDEKTNDKEPTKMYKK-DSVVSSALDTETLMGPTODN-VYFSTS---EDDTHIDKQGTIRKDFQINPKPK- 530
      . .      *      . : : : * :      . : * : : : : *
                                cap-anchor  NAP1L1 motif  KR-motif
                                HBD

```

Figure S16. Analysis of the APLF chaperone. **A.** Alignment correspond to the FHA domain for the APLF proteins. In response to DNA damage, APLF proteins can be phosphorylated. The putative phosphate-binding residues are displayed in light blue. Basic residues supposed to confer specificity for binding CK2-phosphorylated motifs are underlined, CK2 being a serine/threonine protein kinase. The HsAPLF S116 which is phosphorylated in response to DNA damage [141] is conserved in PpAPLF and depicted in blue and underlined. **B.** Alignment correspond to the acidic domain for the APLF proteins. The acidic tail of HsAPLF is integrated in XRCC1 for *Xenopus*, zebrafish and *Physarum*. The acidic domains present the NAP1L motif [49] of the NAP family (red rectangle). They also harbor the H2A-H2B binding cap-anchor motif of HsAnp32E, ScSPT16 and DmYL1 chaperones (purple rectangle) as well as the KR-motif (KR, lysine/arginine, blue rectangle) [50]. The NAP1L motif and the cap-anchor motif constitute the HBD domain of AFLF (orange line). Residues Y462/F468/Y476/E477/W485-HsAPLF are involved in histone binding [49,50,142]. They are depicted in green and are not conserved in most analyzed acidic domains. Alignments are performed with APLF proteins used in the phylogenetic tree reported in Figure S2S. HBD, histone binding domain; KR, lysine/arginine. The *Physarum* proteins appear in red. Positions refer to the mature protein with the initial methionine. Asterisk indicates a fully conserved residue labelled with a pale pink rectangle. Colon and dot indicate conservation between residues of, respectively, strongly and weak similar properties and are labelled with a pale blue rectangle. Atr: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans*; Dd: *Dictyostellium discoideum*; Dr: *Danio rerio*; Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Pp: *Physarum polycephalum*; Ppa: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae*; Tt: *Tetrahymena thermophila*; Zm: *Zea mays*.

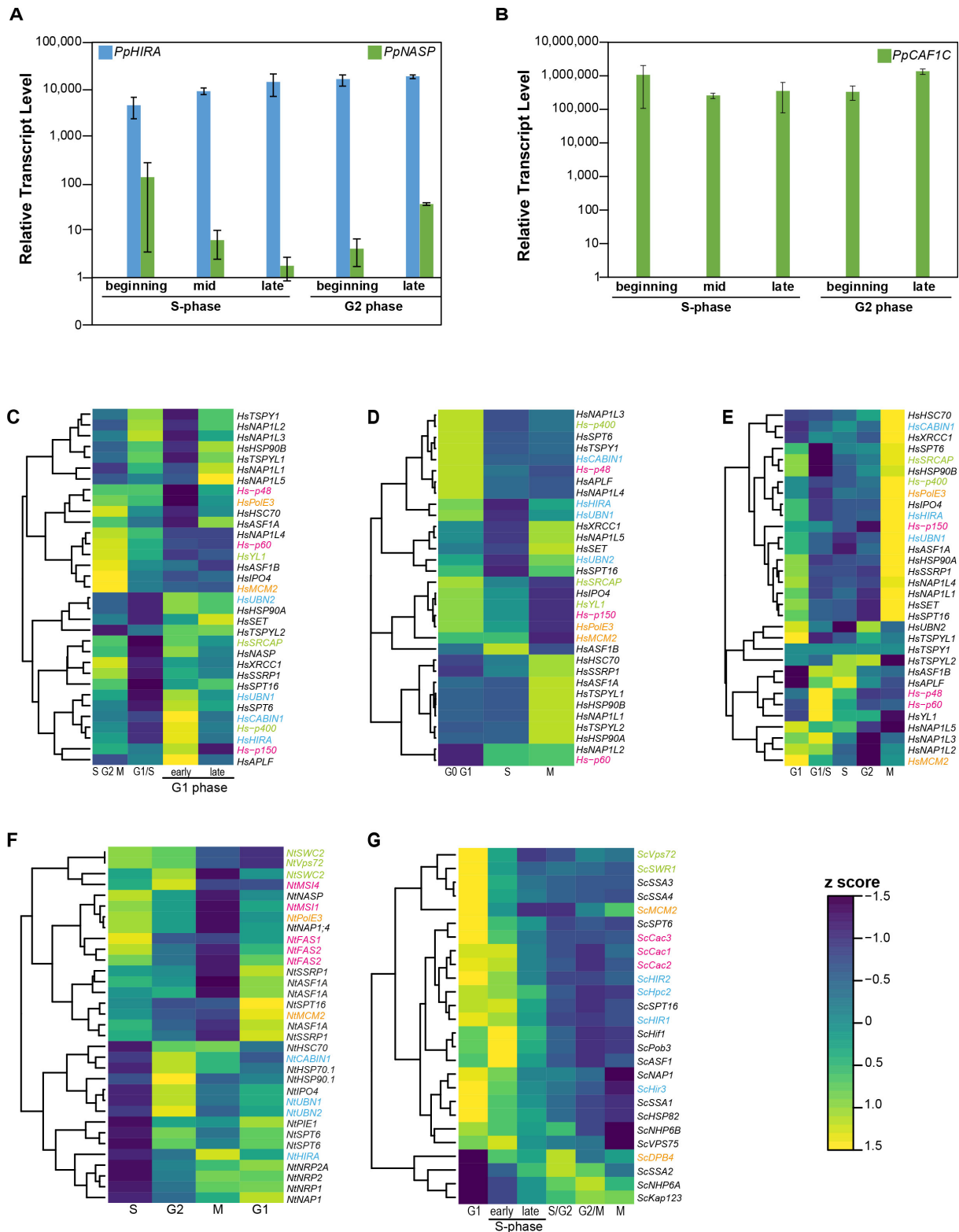


Figure S17. Expression analysis histone chaperone-coding genes. **A-B.** Mean relative transcript level for *Physarum PpHIRA*, *PpNASP* (**A**) and *PpCAF1C* (**B**) genes measured by qRT-PCR. Means are displayed with a logarithmic scale and obtained from three biological replicates consisting of independent mitotically synchronous plasmodia cultures harvested at different times of the cell cycle. Samples used for RNA-Seq and qRT-PCR are independent. **C-G.** Heat maps displaying the RNA-Seq expression of histone chaperones at various stages of the cell cycle in human (HUVEC, **C**; MCF-7, **D**;

U2OS, E), tobacco TN-90 (F) and yeast (G) cells. Cell cycle stages are indicated at the bottom of the map. Each row corresponds to a transcript listed on the right and each column to a cell cycle time point. The color bar at the bottom right depicts the scale for z-score, with blue representing lowest expression and yellow representing the highest. The CAF-1 complex is associated with H3.1/H4 incorporation during DNA replication and its subunits are displayed in pink. The HIR complex is associated with H3 variant incorporation throughout the cell cycle and its subunits are displayed in blue. The SWR-C complex is involved in H2A.Z/H2B replacement and its subunits are displayed in green. MCM2 and PolE3 are histone chaperones associated with replication and are displayed in orange.

Table S1: List of histone chaperones identified in *Physarum polycephalum*.

Physarum Histone chaperone	Physarum Transcript ID	Scaffold	Number of exons/introns	Predicted Protein Molecular Weight (KDa)	Human homologue(s)	% of Disordered residues (RAPID)
PpASF1	Phypoly_transcript_17984	Scaffold927	4/4	22.56	Anti-silencing function 1A/B (ASF1A/B)	27.23%
PpMCM2	Phypoly_transcript_01999	Scaffold1687	20/19	108.7	Minichromosome maintenance protein 2 (MCM2)	25.29%
PpHSP90	Phypoly_transcript_03991	Scaffold842 & Scaffold258	6/5	80.31	Heat shock protein 90A/B (HSP90A and HSP90B)	21.09%
PpHSC70	Phypoly_transcript_04582	Scaffold831	15/14	91.91	Heat shock cognate 70 (HSC70: HSPA8)	10.99%
PpIPO4	Phypoly_transcript_01375	Scaffold213	7/6	117.65	Importin 4 (IPO4)	11.3%
PpSPT6	Phypoly_transcript_00177	Scaffold740 & Scaffold436	30/32	202.76	Suppressor of Ty 6 (SPT6)	33.22%
PpAPLF	Phypoly_transcript_03499	Scaffold4622 & Scaffold945	2/9	48.09	APLF aprataxin and PNK-like factor	78.67%
PpPOLE3	Phypoly_transcript_27931	Scaffold54	4/3	14.22	PolE3	58.73%
PpSWR1	Phypoly_transcript_00085	Scaffold2256	11/10	261.27	SRCAP, E1A-binding protein p400 (hDomino)	36.0%
PpSWC2	Phypoly_transcript_09236	Scaffold551	7/6	41.56	Protein YL1 (YL1), vacuolar protein sorting 72 homologue (VPS72)	64.67%
PpSPT16	Phypoly_transcript_00559	Scaffold433 & Scaffold1547	23/28	119.68	Suppressor of Ty 16 (SPT16)	25.3%
PpSSRP1	Phypoly_transcript_05191	Scaffold957 & Scaffold3182	15/14	59.88	Structure-specific recognition protein 1 (SSRP1)	28.6%
PpSET	Phypoly_transcript_13763	Scaffold146	5/4	28.52	Patient SE translocation (SET)	45.36%
PpNAP1L1	Phypoly_transcript_11436	Scaffold35	8/7	42.35	Nucleosome assembly protein 1-like (NAP1L1-6)	51.22%
PpCAF1C	Phypoly_transcript_09477	Scaffold1106	ONLY RBAP48	47.76	Retinoblastoma associated protein 46 (RBAP46)	18.44%
PpCAF1A	Phypoly_transcript_03998	Scaffold5423 & Scaffold4498	10/9	83.98	Chromatin assembly factor 1, subunit A (CAF1A or p150)	34.31%
PpCAF1B	Phypoly_transcript_09859	Scaffold6006 & Scaffold8861 & Scaffold10170	14/13	70.28	Chromatin assembly factor 1, subunit B (CAF1B or p60)	61.05%
PpHIRA	Phypoly_transcript_01462	Scaffold428	19/19	110.63	Histone regulation A (HIRA)	25.7%
PpUBN	Phypoly_transcript_03291	Scaffold395	10/9	87.31	Ubinuclein 1 (UBN1) and Ubinuclein 1 (UBN2)	68.04%
PpCABIN1	Phypoly_transcript_00351	Scaffold5	34/35	176.49	Calcineurin-binding protein cabin-1 (CABIN1)	23.1%
PpNASP	S3S101K5_5354_LULU & Contig588_N11D4	Scaffold42	5/4	47.52	Nuclear autoantigenic sperm protein (NASP)	45.72%

Table S2: List of protein sequences used for phylogenetic trees. NA, no homolog identified by a blast search in this organism.

RefSeq ID	Tree ID	Organism	Chaperone	Protein Sequence
NP_176846.1	AtASF1A	A. thaliana	ASF1	MSAIKITNVAVLHNP APFVSPFQFEISYECL NSLKDDLEWKLIYVG SAEETYDQLLESVLV GPVNVGNVRFVQA DPPDPKIQEEDIIGV TVLLTCSYMGQEFIR VGYVNNDEDEQLK EEPPTKVLIDKVQRNI LSDKPRVTKFPIDFHP EEEQTAATAAPPEQS DEQQPNVNGEAQVL PDQSVEPKPEES
NP_198627.1	AtASF1B	A. thaliana	ASF1	MSSINITNVTVDNP APFVNPFFQFEISYECL TSLKDDLEWKLIYVGS AEDETYDQVLESVLV GPVNVGNVRFVLA DSPDPLKIREEDIIGVT VLLLTCSYMDQEFIRV GYVNNDYDDEQLRE EPPTKVLIDKVQRNLI TDKPRVTKFPIHFHPE NEQTLGDGPAPTEPF ADSVVNGEAPVLEEQ PQKLQIEQFDDSDV NGEAIALLDQPNLQ ET
XP_006843956.1	AtrASF1A	A. trichopoda	ASF1	MSAVNITNVTVDNP AAFLNPFQFEISYECLL PLKDDLEWKLIYVGS AEDETYDQLLESVLV PINVGNVRFVQADP PDPTKIPDEDIIGVTVL LLTCSYMGQEFIRVGY YVNNDYDDEQLREEP PPRVLIDRVQRNIALD KPRVTKFPIHFHPE DSTDAPPSPSPSEN PIEEPQTEPSARLDT
NP_492567.3	CeASF1L	C. elegans	ASF1	MASRVNIVQVQILDN PAMFVDKFKMEITFE VFEHLPHDLEWELVY VSGTSRDFDQVLDL ALVGPIPEGRHKVVF DAEHDPDISKIPVEDIV GVSVLLLRCKYNDQE FINMGWVAVANEYTD EELKENPPAKPLIEKLS RKIETEDLRVTTFPIR WTDEDPAEPPVDEE ANKVFDEDDLMPH DDGQDDDEEEDDD ETGPNTTEVDLNEF NERMANAHDGTEQ KNGEESMEHDGASG DVEMGDKH
XP_638098.1	DdASF1	D. discoideum	ASF1	MSHIRLTQVLHNNP SSFNSPFIIDISFECVS PIKEDLEWKVIYVGS DNEKNDQVLDLILLG PVAVGQNGQVFEVD PPDANKIPKDLLGV TVVFLICAYKGEDFIR VGYVNNDYFEQELK DNPPETPDLSKIQRN VMDDKPVVTGFPIQ WN
NP_001189131.1	DmASF1	D. melanogaster	ASF1	MAKVHITNVVVDN PSSFFNPFQFELTFECI EELKEDLEWKMIYVG SAESEHHDQVLDTIYV GPVPEGRHIFVQAD PPDVSKIPEPDAVGV TIVLLTCSYRGQEFVR VGYVNNDYADPEM RENPTKPLFEKLRN ILASKPRVTRFKINWD YGHINGNGNGVENG HQDEMATDGPSTSE AASAVIHPEDDNL MPMENGKALNENS NSLAMEC

NP_996946.1	DrASF1A	D. rerio	ASF1	MAKVQVLNVAVLDN PSPFGNPFQFEITFEC MEDLPEDLEWKIIYV GSAESEEYDQILDVSL VGVPVAGRHMVFVQ ADAPNCSLIPETDAV GVTVVLTCTYRGQEF IRIGYVNNNEYTDTEL RENPPPKPNYGQLQR NILASNPRVTRFHIN WEGCAEKMEDESENV DPAPNAMLPPSCTP GKAPLLGLVPDMSM DCL
NP_999922.1	DrASF1B	D. rerio	ASF1	MAKVQVLNVAVLDN PSPFGNPFQFEITFEC MEDLPEDLEWKIIYV GSAESEEYDQILDVSL VGVPVAGRHMVFVQ ADAPNTGLIPESDAV GVTVVLTCTYRGQEF IRIGYVNNNEYDPEL RENPPVKPDYTLQQR NILASNPRVTRFHIN WEATMDKMEDESEN VDAPNAMLPTCM PGKAPPLGLMPDMS MDCL
NP_054753.1	HsASF1A	H. sapiens	ASF1	MAKVQVNNVVLD NPSPFYNPFQFEITFE CIEDLSEDLWKIIYVG SAESEEYDQVLDVSLV GPVPAGRHMVFVQA DAPNGLIPDADAVG VTVVLTCTYRGQEFI RVGYVNNNEYTETEL RENPPVKPDFSKLQR NILASNPRVTRFHIN WEDNTEKLEDAESSN PNLQSLSTDALPSAS KGWSTSENSLNVML ESHMDCM
NP_060624.1	HsASF1B	H. sapiens	ASF1	MAKVSVLNVAVLEN PSPFHSPPRFEISFEC EALADDLEWKIIYVGS AESEEFQILDVSLVG PVPAGRHMVFVQAD APNPSLIPETDAVGT VVLITCTYHGQEFIRV GYVNNNEYNPELRE NPPMKPDFSQLQRNI LASNPRVTRFHINWD NNMDRLEAIETQDPS LGCGLPLNCTPIKGLG LPGCIPGLLPENSMD CI
NP_079817.1	MmASF1A	M. musculus	ASF1	MAKVQVNNVVLD NPSPFYNPFQFEITFE CIEDLSEDLWKIIYVG SAESEEYDQVLDVSLV GPVPAGRHMVFVQA DAPNAGLIPDADAV GVTVVLTCTYRGQEF IRVGYVNNNEYTETEL RENPPVKPDFSKLQR NILASNPRVTRFHIN WEDNTEKLEDAESSN PNLQSLSTDALPSAS KGWSTSENSLNVML ESHMDCM
NP_077146.1	MmASF1B	M. musculus	ASF1	MAKVSVLNVAVLEN PSPFHSPPRFEISFEC EALSDDLWKIIYVGS AESEEFQILDVSLVG PVPAGRHMVFVQAD APNPSLIPETDAVGT VVLITCTYHGQEFIRV GYVNNNEYPPELRE NPPPKPDFSQLQRNI LASNPRVTRFHINWD NNPDSLEAIENQDPN VDFSLSLCTPVKSLG LPSCIPGLLPENSMD I
XP_001768290.1	PpaASF1A	P. patens	ASF1	MTAVNVTNVTVLDN PSMFMNPFQFEISYE CLCPLQDDLEWKLIYV GSAEDEKYDQVLESV LVGPVNIGNYRFVQ ADPPELSKIPEEDTIG VTVLLTCSYQGQEFI RVGYVNSNDYVEAL REEPPARVLIEKVQRN ILADKPRVTKFPIVFN SPAPVLNHIESERITD SYISEPAAISTLHTYEN VSDVHSTPPAAVDIM QTDENIVHAASLDT PMDLAPADEVRVDL RTPQTLQEV

-	PpASF1	P. polyceph	ASF1	MGCNVNTNIQILDNP APFMNPLQFEITYECI TDLQEDLEWKVVVV GSAESEAHQVLDLSIL LGPVRVGVSNFVLQV DPPDSSRIPKQDLLGV TAILLMGLYKSRDFIR VGYFVNNEYAEELRL NPPPEGEVDLTKVVRN ILAEKPRVTRFPIPWD QETIIPPANTENSSSLP AITTETSLDAASNHS MTEEMMPMEVMST ST
NP_012420.1	ScASF1	S. cerevisiae	ASF1	MSIVSLLGIKVLNNPA KFTDPYEFITFECLLES LKHDLEWKLTYVGSS RSLDHDQELDSILVGP VPVGVNKFVFSADPP SAELIPASELVSVTVIL LSCSYDGREFVRVGYV VNNEYDEELRENPP AKVQVDHIVRNILAE KPRVTRFNIVWDNE NEGDLYPEQPGVD DEEEEDDEEEDDED DEDDDDQEDGEG EAEAAAAEEEEEEKT EDNETNLEEEEDDIEN SDGDEEEGEEVGSV DKNEDGNOKRRKIE GGSTDIESTPKDAARS TN
XP_001033143.3	TtASF1	T. thermophi	ASF1	MALVNIQNIQFHNN PCPFLSPFKLDVTFECI KPIPPDIEWQLIYIGS AKDEKYDQLDKFSI GSLDQGVQLQFTIETN PPDHTKIPNKDDLGL VTAILTVSYHNQEFF RVGYVYNQYNDQE LIINDPPQLIDKVERSI LDKQPRITHFNIKWG TEDENKETDPNTLAIL QQQLAQNGTIPNQL MQEMQQVNSQMQ SFLDPFSKSSGILQEL TTQGQTNSTNSFMF GQGLDIPQNNAFQP TNVFSSNPY
NP_001086449.1	XIASF1A-A	X. laevis	ASF1	MAKVQVNNVVVLD NPSPFYNPFQFEITFE CIEDLSEDLWKIIVVG SAESEYDQVLDSVLV GPVPAGRHMVFVQA DAPNPGIIPDADAVG VTVVLTCTYRGQEFI RVGYVNNNEYTETEL RENPPVKPDFSKLQR NILASNPRVTRFHIN WEENTEKLDDSNPH MHPVLSIEARPSASK GWPMSENSLNVMLE SHMDCM
NP_001080310.1	XIASF1A-B	X. laevis	ASF1	MAKVQILNMVVLDN PCPFHNPQFEITFEC IEDLPDDLEWKIIVVG SAESEYDQTLDSVLV GPVPAGRHMVFVQA DAPNCSLIPESDAVG VTVVLTCTYRGQEFI RVGYVNNNEYSDPPEL RENPPKHAHFGQLQR NILASNPRVTRFHIN WECTSEAKMEDIENV DPASNTMLPPNCAP SKGLAAALNTLPENS MDCM
NP_001150107.2	ZmASF1A	Z. mays	ASF1	MSAVNITNVAVLN PTAFLNPFQFEISYECL VPLDDDLWKLIYVG SAEDENYDQQLSVL VGPVNVGTYRFVLQA DPPDPSKIREEDIIGVT VLLTCSYMGQEFMR VGYVNNNDYDDEQL REEPPAKVLIDRVQR NILADKPRVTKFPINF HPEPSTGPGQQQE PQTTSPEHTGNGEA NGSKPEADQ

XP_008656015.1	ZmASF1B	Z. mays	ASF1	MSAVNITNVAVLN PTAFINPFQFEISYEC VPLDDLEWKLIYVG SAEDENYDQQLSVL VGPVNVGTYRFVLQA DPPDPISKIREEDIIGVT VLLTCSYMGQEFMR VGYYVNNDYGDQEL REEPPAKVLIDRVQR NILADKPRVTKFPINF HPEPSTGTGRQQQQ QEPQTASPEKHAGSG EGNGSKPEAD
NP_175112.2	AtMCM2	A. thaliana	MCM2	SPSSAGFNTDQLPIST SQNSENFSDEEEAAV DTQVIRDEPDEAEDE EEEEGEDLFNDTFMN DYRKMDENDQYESN GIDDSVDDERDLGOA MLDRAAADADLDAR ENRLANRKLPHLLHD NDSDDWNYRPSKRS RTTVPFRNGGDPD GNPPSPGVSQPDIS MTDQTDYQDEDD NDDEAEFEMYRIQGT LREWVMRDEVRRFIA KKFKDFLLTYVKPNE NGDIEYVRLINEMVS ANKCSLEIDYKEFIHV HPNIAIWLADAPQPV LEVMEEVSEKVIDLH PNYKNIHTKIYVRTN LPVNDQIRNIRQIHLN TMIRIGGVVTRRSV FPQLQKVKYDCNKC GAVLGPFQNSYSEV KVGSCSECQSKGPFT VNVEQTIYRNYQKLT QESPGTVPAGRLPRH KEVILLNDLIDCARPG
XP_020523093.1	AtrMCM2	A. trichopoda	MCM2	GARSPNPSGNGNPR SMDSDDPSTPASPTS AGFSTDRLPNTSRTS QSLSDAAVDPAIIRD EEEEGMGEEEEEGE DLYTDNFFDDYRPM EQDRYESAGLDDSVG DERDLQIMEDRRA AEVELDARDGTRFDR KLPRLLHDQDTDDI DYRPKRSRVYRPR GGMSYEDVDGATPIS PGGGGSHGGYHGGE PPTTDQTEDEDEDD AEYGIYRVQGTLEW VTRDEVRRVIKKFKE FLLTYVNPKNNGEH EYVRLINEMVAANGC SLEIDYKQFIFVHPNIA IWLADAPQVLEVM EEVAQSVVLDLHKNY KNICQRIYVRTNLPI YDQIRNIRQIHLNTMI RIGGVVTRRSVFPQ LQQVKYDCNCKGMIL GPPFQNSYSEVKVGS CPECQSKGPTVNIE QTIYRNYQKMTLQES
NP_001022416.1	CeMCM2	C. elegans	MCM2	QPLPIADDDDDVD GIDEMFNNEDEPE DEEENLFGDDMER DYREQPELDQYSESG MDDASDVGLSVSA RRAAEREMAQRDQL LDDALMYEDGDSEE VDTRRRGRGRGRG DAADDDSVPMEEEDI PVDILENIRGRTIRDH VSDEAVAKEIERFKN FLRSFHEPGNKQTKYI QMIKSMAADNRESL EVSFTDLSDDNGEQN ISYFLPEAPNEMLAIM DRAATEVVMNMYPF YSRVCNEIKVRISQLP VEEDIRMLRQVHLN MLURTAGVVTIASGIL PQLAVVKYDCVACGY LLGPFVQQNDEEVRP TICPSCQKGPPFELNV ENTVYHNYQRITMQ ESPNKVAAGRLPRSK DVILLGDLCDSCKPGD EIEVTGVYTNNFDGSL NYKQGFVFNLIHA NHITNKDKMASDQL

				NNNNNNNNKRKT KTPSSVNNHNTKINE DKDLYNFIDKIRDKTD KSHLEDEYDEDDNTK DGKFNKVEDEIKNDI DSEEDDDIIEDNEHL YDGEDSENLSSEGE GDDLLEDLDQDQDEE EVLGDEDNHYEMEG LDNETYEDIDKDKKTE VDSILTKRDREEQKTR QATNRHQYRISSFLD ENDENMVDEDTENN KLEKRNKLIKRRRE TIRQQNDNIKKDYDG SVPSFKDGFVNSERYF EDGDGDDDDDDDDF DIGDNGQVNVLEEPD KDKCLREYVCMGSPR NEISKQFAKLLFEKEK DGFNPIYQESIQKMC AANKESLMINFTHLC ASTVFGVWVAEIPTE MIEIFDEVALKVVLRIY PNYRNIVKSIHVRIITH LPICESLRDIRQSNLN KLTQVGGVITRRSNVY PQLKHVKYDCVKCKT
XP_637579.1	DdMCM2	D. discoideus	MCM2	
				AAERRDLRAAMTSPV GDPEFENEDEILGD QTVRDEAEEDGEEL FGDNMENDYRPMPE LDHYDPALLDDEDDF SEMSQGDRAAESE MRRRDRAAGIHRDD RDLGFGQSDDEDDV GPRAKRAGEKAAV GEVEDTEMVESIENLE DTKGHSTKEWVSML GPRTEIANRFQSFLRT FVDERGAYTYRDRIRR MCEQNMSFVVSYT DLANKEHVLAYFLPE APFQMLEIFDKVAKD MVLISIFTYERVITTEI HVRISELPLIELRTFR KLHLNLQVRLGVVT ATTGVLPQLSVIKYDC VKCGYVLGPFVQSQN TEIKPGSCPECQSTGP FSINMEQTLRYNYQKI TLQESPGRIIPAGRI SKDVILLADLCDQCKP GDELEVTGIYTNNDY GSLNTDQGFVPVATV IIANHVVKDSKQVV
NP_477121.1	DmMCM2	D. melanogaster	MCM2	
				TRGSRRGDLTSSPGR DLPPFEDESEGLLGD LPDEEDDDGEELIGD AMERDYRVIPELDY EAEGLDEDEDLSELS SARAAEAAMRRRD REQQLGMGRIGRGLL YDSEEDDKRPTKRQ RVLAERAAEGGAME GEDEEMIESIENLED MKGHTVREWVSMA APRLEIYHRFKNFLRT HVDEHGHNVFKERIS DMCKENKESLLVNYE ELASREHVLAYFLPEA PAEMLKIFDEAAKEV VLAMYPKYDRIAHEI HVRIGNLPLVEELRSL RQLHLNLQIRTSGVV TSCTGVLPQLGMVKY NCNKC�FILGPFQFS QNQEVKPGSCPECQ SLGPFEINMEQTVYQ NYQRITIQESPGKVA AGRLPRSKDAILLADL VDMCKPGDEIELTGIY HNNDYGSNLNANGF PVFATVILANHIARKD
NP_775364.2	DrMCM2	D. rerio	MCM2	

NP_004517.2	HsMCM2	H. sapiens	MCM2	QRRRGNDPLTSSPGR SSRRTDALTSPPGRDL PPFEDESEGLLGTGEP LEEEDGEEELIGDGM ERDYRAIPELDAYEAE GLALDDEDVEELTAS QREAAERAMRQRDR EAGRGLGRMRRLLY DSDEEDEERPARKRR QVERATEDGEEDEE MIESIENLEDKGHVS REWVSMAGPRLEIH HRFKNFLRTHVDSHG HNVFKERISDMCKEN RESLVVNYEDLAARE HVLAYFLPEAPAEELLQ IFDEAALEVVLAMYP KYDRITNHHVRIHSL PLVEELRSLRQLHLNQ LIRTSGVVTSCTGVLP QLSMVKYNCNKCNE VLGPFQSQNQEVKP GSCPECQSQAGPFEVN MEETIQNYQRIRIQE SPGKVAAGRLPRSKD AILLADLVDSCKPGDE IELTGIYHNNYDGSLN TANGFPVFATVILAN
NP_032590.2	MmMCM2	M. musculus	MCM2	QRRRISDPLTSSPGRS SRRADALTSPPGRDL PPFEDESEGLLGTGEP MEEEDGEEELIGDGM ERDYRPIPELDVYAE GLALDDEDVEELTAS QREAAERTMRQRDR EAGRGLGRMRRLLY DSSEEDEERPARKRR HVERATEDGEEDEE MIESIENLEDKGHVS REWVSMAGPRLEIH HRFKNFLRTHVDSHG HNVFKERISDMCKEN RESLVVNYEDLAARE HVLAYFLPEAPAEELLQ IFDEAALEVVLAMYP KYDRITNHHVRIHSL PLVEELRSLRQLHLNQ LIRTSGVVTSCTGVLP QLSMVKYNCNKCNE LGPFQSQNQEVKP GSCPECQSQAGPFEIN MEETIQNYQRIRIQE SPGKVAAGRLPRSKD AILLADLVDSCKPGDE IELTGIYHNNYDGSLN TANGFPVFATIIILANH
XP_024380390.1	PpaMCM2	P. patens	MCM2	GDFDTRPLSSTGG GRRGSMGGSEFDSE NEAVEDRNLFIDEIIV DQEEEDGEDLYDEN FMQRDYQTLEEQR YEVDGLDDEVEDTRD FATIMADRRAAEANL DERDGLNHSRVRKL PTMLQEYDDDENEEY QPARRRMRLHEPVA SPVMSMSEHDETGD NEEEDDEEDDGGTD FLNVQGSIAEWVA RDDVRRFVQRKFRRF LETSSKEAMYKKVYR DSLDNMVAANLCSLE LNYGQWLDSCPELAI WLADAPQPLLEIMEE EANAFVLRHHHPNYSK IHEKVYLRISNPLEDK IRNIRQVHLDTLIKISG VVTRSGVFPQLQQV KYDCVKCGTILPFFQ NTHTEIRVGSCPECQ SRGPFTVNVEQTIYR NYQKLTQESPNTVQ AGRLPYKEIILLHDLI DVARPGEEIEVTGIYV

-	PpMCM2	P. polyceph	MCM2	EEEYDPSIPKDKEAE DVESETLSDGNQEA MDDILDGEENVVYDP EELEAINQGEDLYNG DFEQDYEYDPNLDY DEKDLDENEYAEMG PEERAAAAEAIKKRN RKVQKAKTGRLPSAL LYDTDEEDAKPEYKP RRRKRPVNDPNVGN VFVEEDTRDLGOLD VNLEDVKGPLREWIS LDPTKRTIMRRFRKFL KTYGPAKAPFYQQA QTMCAANQESLLVN HNHLTQTEPLIAIWT ADAPTEMIQIFDEVA MKLVLEMFPNYENIC PEVHVRTNLPVSDSL RELRQIHLNCLVKVSG VVTRRTSVFPQLKYVK FDCLKCSTILGPYYQD GQAEIKVGICPNCQS KGPFAINAEQTVYQC KKKKKKKKKKKKKKK KKKRSKEVILLWDLID SARPGEEVEVTGIYRH NFDEHLNKTGHFPVF
NP_009530.1	ScMCM2	S. cerevisiae	MCM2	DSENELPPSSPQQHF RGGMNPVSSPIGSPD MINPEGDDNEVDV PDIDEVEEQMNEVDL MDDNMVEDYAADH NRDROYDPDQVDDRE QQELSLSERRRIDAQL NERDRLRNVAIIDD EDEEQEGAAQLDEM GLPVQRRRRRRQYED LENSDDLLSDMDID PLREELTLESLSNVKA NSYSEWITQPNVSRTI ARELKSFLLEYDETG RSVYGARIRTLGEMN SESLEVNRYRHLAESKA ILALFLAKCPEEMLKIF DLVAMEATELHYPDY ARIHSEIHVRISDFPTI YSLRELRESNSSLVRV TGVVTRRTGVFPQLK YVKFNCLKCGSILGPF FQDSNEEIRISFCTNC KSKGPFRVNGEKTYY RNYQRVTLQEAPGTV PPGRLPRHREVILLAD LVDVSKPGEEVEVTGI YKNNYDGNLNAKNG
XP_001009217.2	TtMCM2	T. thermophi	MCM2	ADNDSASVSQEEAIE ENQJVEEDDKSLSDE GDGDDLLNIEDDYK AVPELDQYEIDGLND EDDFDDIDAEQRRLA EEEIDERNYQKYANK RRIPTALREFDLGEED EEDLNVQLQRRRQFY EAHNLYESEDEQME EEEENKDEREYDKYLD REQARGKLAVWIKKE STVRYIRKTRKFLNKF KDGSNQLVYKQRIQE MCSKNKGSLEVDNF HINSKHETIAMWIVT EPVILPYLNDVAMET VKKLYPDYEDIQPEIF VRIENYTILDIRKLRH KDLGQLIQVRAVVTQ RSAIFSQKKVYYVCR CGDRKGPYLSSENH NLGTCPCVQSKGPFIL DTEYTIYRNYQRLTIQ EPPGSVPGRVPRQK EVIVLGDNI DVARPG DEIDVTGIFVHRYDYA LNVKHGFPVFSTMIE ANYIKRVREGDQSSL

NP_001080759.1	XIMCM2	X. laevis	MCM2	AGSRRDALTSSPGRD LPPFEDESEGMFGDG VVPEEEEEGEELIGDA MERDYRPISELDREYEV EGLDDEEDVEDLTAS QREAAEQSMRMRD REMGRELGRMRRL LYDSDEEEDRPARKR RMAERAAEGAPEED EEMIESIENLEDMKG HTVREWVSMAATRL EIYHRFKNFLRTHVDE HGHNVFKEKISDMCK ENKESLPVNYEDLAA REHVLAYFLPEAPAE MLKIFDEAAKEVVLV MYPKYDRIAREIHVRI SHLPLVEELRSLRQLH LNQLIRTSGVVTCCTG VLPQLSMVKYCNCK CNFILGPFFQSQNQE VRPGSCPECQSFQPF EINMEETVYQNYQRI TIQESPGKVAAGRLP RSKDAILLADLVDSCK PGDEIELTGIYHNNYD GSLNTANGFPVFATV ILANHITKKDDKVAV
PWZ26529	ZmMCM2	Z. mays	MCM2	PGFSTDRLPPNNTTSR GATDPSSYSDDGGEA EVDPHVLPDDDDPA AVVAPDEDEEGEDLF NDNYLDDYRRMDEH DQYESVGLDSDLEDE RNLDEIMADRRAAEA ELDAREVRTGATADR KLPRMLHDQDTDDD TNFRPRKRHRASFRQ PSGPRTPRSSDDGDGA ATPSSPGRSHRGMSS GGDVPMTDQTDOD PYEDEFDEEDEMNM YRVQGTLEWVTRD EVRRFIKKKFEFLTY VHPKNDQGEFEYVRL INEMVLANKCSLEIDY KQFIYIHPNIAIWLAD APQSVLEVMEEVAK NVVFDLHKNYRNIHQ KIYVRISNLPVVDQIR NIRQIHLNTMIRIGGV VTRRSQVFPQLQQT YRNYQKLTQESPGIV PAGRLPRYKEVILLND LIDCARPGEEIEVTGIY TNNFDLSLNTKNGFP
NP_194494.2	AtIPO4	A thaliana	IPO4	DNDARRQAEQIKRL AKDPQVVPALVQHL RTAKTPNVRQLAAVL LRKRITGHWAKLSPQ LKQHVKQSLIESITVE NSPPVRRASANVVS VAKYAVPAGWPD LTFLLQCSQSAQEDH REVALIFSSLTETIGN TFRPYFADLQALLKC MQDESSSRVRVAALK AVGSFLEFTNDGDEV VKFRDFIPSILDVSRKC IASGEEDVAILAFEIFD ELIESPAPLLGDSVKAI VQFSLEVSCNQNL STRHQAIQIVSWLAK YKYNSLKKHKLVIPIQ VMCPLLAESDQEDD DDLAPDRASAEVIDTL AMNLPKHVFLPVLEF ASVHCQSTNLKFREE SVTALGVISEGCFDL MKEKLDTVLNIVLGA LRDPELVVRGAASFAI GQFAEHLQPEILSHY QSVLPCLLIAEDTSEE VKEKSHYALAAFCEN

XP_006838378.1	AtriIPO4	A. trichopoda	IPO4	DNDARRQAEQJKRL AKDPHVIPALLQHMR TAKTANVRQLSAVLL RKKITGHWMKLSPQ VKDSVKSALMESITVE HSPPVRRASANVVS VAKYAVPAGEWPELL PFLFQCSQSPQEDHR EVALILFSSLTETIGEM FRPHLATLQSLFLKSL QDETSTRVRIAALKAV GSFVEFVHVRDEVV MVRELIPILNVSRQC LANGDEEVAVIAFEIF DELVESAPLLGSFVV SIVQFSLEVSSSENLET NTRHQAIQISWLAK YKPKSLVKHKLVPIL QVMCPLLAESHGD DDDDDLADRAAAE VIDTMAVNLPRKLVF PPVLEFASIGYQNP PKYREASVMALGVVS EGCFELMKNKLEDVL RIVLEALKDPEQLVRG AASFALGQFAEHLQP EIVTYYESVLPCLNSI GDASEEVEQKSYAL
NP_490715.1	CelMB-3	C. elegans	IPO4	ADNDIRKQAEVYEKI EGPTKVAALFEAYTH HTNNSDARSTVLVFL RRVLARDWDIAIWET LNEENKQRIKAVLE MIVHETELSIKKIADL ISEIASNLIDDSGDMT WGGVLELMDHCLKS EDLTGNYIALILRGCP IIFGNRLAHFLPTLKV VLEKCMATPDLQIKA TAVRAVIAFAVDNDE EKDVRLMTSLVPSNV LQVCNETSDEDDSDG PLGEFAELASSLPKCL NTHMSQVLQVTLAIA GNKEKNEMVRQNAI EVICSYMESAPKGLKK YAPGALGPILETLISC MTMDDDLNEWL NEIEEEDYEDIPIAES AIDRVACCINGKVML PVFLPLVEKLLTSEDW KMKHAALRAFSAVG EGCQRSMEPHIEQIM AHITNYVNDAPRV QYACNAIGQMSSD FAPTLQKKCHAAVIP
XP_646718.1	DdiIPO4	D. discoideum	IPO4	RALLQPNNDIIKAAT DKLNKLLKSEYSLYLF HLEESSPYDEIKQLSA VLLRQKLVAHWTKFS VESRKYIKDSILKLVIS QPSQLVRRSISEVIIIA RLEVATGTWGDLPFF LLQLSSSPDITVRQIQI HILDLSISNVDFLKYF PQLPTVLSQAVIDPQL SVRALAVKAIGSSIIYA VQTDSKLPFIDLPTT LQVIKQCIENEMEDD VISAFEIFNDLVESPPYS SIKAHIPLIVNFSIEIVK QPEIDNSIKTIALEFLE TCIKYQPKILKNSNLL NPILEILFKILTFESDSSI DENDYEYNILQSASV AIKECGKSYSSKLIYPI LPTLKQFLESENVNVR NAVMMVIIQQLSYGCIE TMKDDLDNIIQFVLR GLKDQEKVRSACV TIGKLSQTLTPEIYKYT NQVFPLLFQQLSDPD DQFILRCFALENFLL NLDSQELIPLPNVMD

				ERIRESTAKMLKAYEN PDSLLVLTQIVMSDRP VQERQVAAVLLKRRV KKLRHWQLVPAEHQ AAIKSNMLQVLIQV EKTVMGTVAFIIGSLV RHEEDKQNSWREEIL KFIYERCSPPIESER GSSIFSSLMDAAPDQ FSNHTDTIFPLLAGILV TAEANGNMATPTVH NMLAGTCVLLPFISG HSDAEQIMVKAVPLI LKALAAFAEKGNSNE FMGAFDIIDSMAEYV PHLLTGNVKLLLEFCL MIARNKQFDASIRVQ VLTQVGLVRLKKII MKQKLLQPTLSVLF VICQDDLKEDDDYFSS ESLNSPSNAAQTL LMALHMVDPKFIPPL LDLLEPALQSPVLR RSSFCMGVIAEGCSE AIGKKYLEVMLNIIKA GVLDVSMFVRTAAFF ALGQFSEFLQPTICKF APQILPVLFQDYNQLV
NP_730184.1	DmArtemis	D. melanogaster	IPO4	
				ERIRESTAKMLKAYEN PDSLLVLTQIVMSDRP VQERQVAAVLLKRRV KKLRHWQLVPAEHQ AAIKSNMLQVLIQV EKTVMGTVAFIIGSLV RHEEGEQNSWREEIL KFIYERCSPPIESER GSSIFSSLMDAAPDQ FSDHTDTMFPLLAGIL VTAEANGNMATPTV HNMLTGSCFLLPFVS GHSNAEQIVVKAVPL ILKALAAFEKGYSEIF MGAFDIIDSMAEHVP HLLTGNVKLLLEFCLM IARNKQFDASIRVQV LTFVGLVRLKKIIM KQKLLQPTLSVLFVI CQDDLEEGDDYFSS ESLNSPSNAAQTL MALHMVDPKFIPPL DLLEPALQSPVLR SSFCMGVIAEGCSEA IGKKYLEVMLNIIKAG VLDVSMFVRTAAFFA LGQFSEFLQPTICKFA PQILPVLFQDYNQLV
NP_730182.1	DmApollo	D. melanogaster	IPO4	
				NAVIQQTAEKQAF KDPALPILCAVMTG SQNPQVLPFIPAFIKF SLFFNVKLLCVLLQVR QSAVMMLRMVRK QWKKISPDHRESLKA VVLQAFQTEHTVR HSLQSLSAVLKHETP DRWPALLELNQSTK SNNPQDRQVGLLLS KVVGSNPEPFKPHYK QLLQLFGTVLQDLNN PTALYICITLTAITAY TGTEEMNLMRSLIPK LLVALKHILQADQDQ ASEAMEVDFELMESE VSIVVPHIAEIVRFCL ISADASLSDSLVRKAL SCIAVLIRLKNKAVLK HKLLQPIQVVFPIISA APPPGEEDPEDEEND TGGDSENPKHFAVQ VIDTMLHMPPEKLF NQLLPFTQACLSSEN PYERKGGMLCMAVL AEGCADHIRTMLSS MLQTVCRSLSDNNQ VVRSAALFALGQFSE
F1QIP3	DrIPO4	D. rerio	IPO4	

NP_078934.3	HslIPO4	H. sapiens	IPO4	DTERIRRATEQLQIVL RAPAAPALCDLLAS AADPQIRQFAAVLTR RRLNTRWRLAAEQ RESKSLILTALQRETE HCVSLSLAQLSATIFR KEGLEAWPQLQLLQ HSTHSPHSPEMGL LLLSVVVTSRPEAFQP HHRELLRLNETLGEV GSPGLLFYSLRTLTTM APYLSTEDVPLARML VPKLIMAMQTLPIDE AKACEALEALDELLES EVPVITPYLSEVLTFCCL EVARNVALGNAIRIRI LCCLTLFLVKVSKALLK NRLLPPLLHLTFPIVA AEPPPGQLDPEDQDS EEEELEELMGETPKH FAVQVVDMLALHLP PEKLCQQLMPMLEEA LRSESPYQRKAGLLVL AVLSDGAGDHIRQRL LPPLLQIVCKLEDPS QVVRNAALFALGQFS ENLQPHISSYSREVM PLLLAYLKSVP LGHTH
NP_077229.4	MmlIPO4	M. musculus	IPO4	DTERIRRATEQLQITL RDPAAPALFDLLATA TDSQIRQFAAVLTRR RLNNRWRLAPEQR ESLSLVLTAQKETV HSVSVSLAQLSATIFR KEGLQGWQPQFMNLL QHSTHSSHSPKEVG LLLSVVVSSQPEAFH AHQHELLQLNETLS DVSFPGVLFYSLRTL AIARYVRPDDVSLAR MLVPKVVTALRTLPL DEVKACEALEALDEM LETELPIINPHLSEVLT FCLEVAKNVALGEPL RVRVLCCLTLFLVKVKS KALLKNRLVPLLHAL FPLMAAEPPMGQLD PEDQSDDDDDLEIGL MGETPKHFAVQVVD MLALHLPPEKLCPHV MPMLEEALRSEDPYQ RKAGFLVLAVLSDGA GDHIRQRLLYPLLQIV CKGLDDPSQIVRNAAL LFALGQFSENLPQPHIS SYSEEVMPLLLSYKLS
XP_024374755.1	PpalIPO4	P. patens	IPO4	DNASRKQAEQIRRL SKDPLLVPALLHHVRC SPYPEVRQLAAVLLRK KITGHWMLQSAEMR NNVKSTLLESITLENSP PVRRGSADVVSVAK HAVPAGIEWPELLPFL HQCSQSAQEDHREV ALILFSSLTETIGDLLRP HFATLQSVFITGLND QQSNRVRVAALKAV GALVGVIQSEQVVM MFRELIPILNVSRCL ANGDEDVAILAFEIFD ELIESAAPVLGPTIPVI VQFALEVCSNKHLEA NTRHQAIQIIFWLAK YKPKTLVKHKMVTPIIL SVICPILAEPESTRTHED DIACERAAA EVLDTM ATSLPKKHVFPVPLHF ATSNFHNPDPNYRD AAVMSLGVISEGCYE AMKSRLEDVLSLVLEA LKDKQAVRGAASFA LGQFAEHLQPEISEHY ERVLPICIFAVLSDAVP DVQEKAFYALAAFCE

-	PpIPO4	P. polyceph	IPO4	ALHPNTEIVRAAVTQ LNAALDTPQSTMFL ALLQQSPHPTHVKQFA AVVMRKRVVGHYYK LPQNERGVLKASLL MLHSDTKVEVIVRRA VADVVTGKAIVPR GDWNVLLPTILNAAS SAVLIEDREIAMYLLA AVIDSLGDQMRHY ASLHTLLAACLRDSNP SIRSLALHTTSALIPWL VTKDDIALYGSILPLVV DVVKYFVQNEMDDD AIESFSILEDLVETPIPV ITKFIPALVNLMEICA NHQVDISVRQSALSF VQLVVRVKPKMVVK NKLVTPTLAAAFAMC AEGEGNDEAQQDLY EQEETLSPYIFGSQVI DTVLSLPVNQVIGP SLALVAQYGRPEMHF RORRAAITAITLVEG CAESIKDTLAKVVQYV VGALVDPFSLVRQAA CLAIGQCADNLQPEII DYHNVVVPALLQGLR
NP_011035.1	ScKap123	S. cerevisiae	IPO4	AITSGVGLKEATKTLQ TQFYTQPTTLPALIHIL QNGSDDSLKQLAGV EARKLVSKHWNAIDE STRASIKTSLQTAFSE PKENVVRHSNARVIASI GTEELDGNKWPDLV PNLIQTASGEDVQTR QTAIFILFSLLEDFTSSL SGHIDDFLALFSQTIN DPSSLEIRLSAQALN HVSALIEEQETINPVQ AQKFAASIPSVVNVL DAVIKADDTMNAKLI FNCLNDFLLDSQLTG NFIVDLIKLSQJAVNS EIDEDVRVFALQFISS LSYRKSKVSQSKLGPE ITVAALKVACEIDVD DELNNEDETGENEEN TPSSSAIRLLAFASSEL PPSQVASVIVEHIPA MLQSANVFERRAILL AISVAVTGSPDYILSQ FDKIIPATINGLKDTEP IVKLAALKCIHQLTDD LQDEVAKFHIEEYLPJII DIIDSAKNIVIYNYATV
NA	NA	T. thermophi	IPO4	
XP_018114829.1	XIIPO4	X. laevis	IPO4	NAVIQQATAQLKEAF KDPQIIPALFEILRGSQ ELQIRQFAAVLLRRRL NKHWWKAVQPEQQH NLKIVLESIQREPEHK VRYALQQLIAVILKNE RLEHWPEFIKVLQLS HSDVPDQKQVILVL WCSLHLKASLFQPHV HDLGLFKQTLSDLN NGPLIYYTVQSLTCILP YIVGNETNLLRPFIPKI LAAIRQLILVNQVQA CEAMEFFDVLMEDE VPVIVHYIADTVQFCL EMAVNTSLSDELRVK ALSCIMCLIKKSKSIIK QKLLPQILNLFPIIMC AAPPAGEMDKEDQE EEDDDIEDSVETPKEY AMQVIDMLALHLPP EKLKELSPLMEPCLLS SNPYPRKAGLMCLAV LSEGCDFICDKHLQP MLSLVCQSLSDDNQV VRNAALYALGQFSEH LQPDITNYSYDTVLP LEYFSRVDPSTAHLT

				DNDARRQAEQJRRRL ARDPQVVPALVHHL RTAKTPNVRLAAVL LRKKITSHWPKLPADS KASKLQALIDSITLDHS HPVRRASANVVSIAK YAIPAGEWPELLPFLF QCSQSPQEDHREVAL ILFSSLTETIGATFQSH LNNLQPILLKCLQDET SSRVRIAALKAVGSFI EYVNDGGDIVKMFRL DFVPSILNISRQCLAN GEEDVASIAFEIFDELI ESPAPLLGDSVRSIVQ FSLEVSANQDLEINIR QQAIQIISWLVKFKAS FLKHHKLVVPIQVM CPLLTETANEDESDL AADRSAAEVIDTMAI NLPRHVLAPVLEFASL SFHHINPKYREAAVTS LGVISEGCCHEHLKDKL EDCLKIVLEALKDQEQ MVRGAASFALGQFA EHLQPEILSHYASVLP CILNALEDPSDEVKEK SYALAAFCEDMGED
ONL96235.1	ZmlIPO4	Z. mays	IPO4	
				GVPPKAGNTYSIDVK NFISRARALYEHWKK HSADLWGSADALAIA TPPASDDLRLKSSAL NIWLLGYEFPDTIMV FTKKQHFLCSRNKAS LLEVVKKPAHDELKLD VIMHVKPKGDDGTG LMDAIFRAIRDLRG DGNDSQVVGHIARE APEGKLLTWTTERLK NANFQFVDITGGLSD LFAVKDDTEVMSVKK AAYLAYSVMKNVVV PNLESAIDEEKDVTSH ALMDLTEKAILEPTKA SVKLKPENVDICYPPIF QSGGKFDLKPSAASN DELLTYDPASIIICAVG ARYNSYCSNVARTYLI DATSLQSKAYEVLLKA HEAAIDALRSGRKINT VYQAALSVVEKNAPE FVDKLTKSAGTGIGLE FRESGLNINAKNDKV LRPKMAFNVS LGFQ NLECESESRSKKNKFS LLLADTVLVTQKPEL
NP_001329034.1	AtSPT16	A thaliana	SPT16	
W1NNZ7				NASKSSSYQIDLDTFSS KRLKAFYSSWEENRD SLWASADAIAVATPP PSDDLRYLKSSALNIW LLGYEFPETIMVFTSK QIHFLCSQKKANLLET LGKSAKSALNLEVLH VKAKNEDGASQMEEI FNAIKSHAESKVVV GHIAREAAEGKLLT WREKLELNFQLGDVT NGFSELFVAVKDENEIT NVKKASFSLASVLKNF VVPKLEVIIDEKKVT HSSLMDTTEKAILDP AKSKVKLKAENVDICY PPIFQSGGQFDLRPN SSSNDDHLYDATSVI ICAIGSRYSYCSNVA RTYLIDADDIQNRAY NVLFKAHEAAIGSLKP GKKIGEAYKAAMEVV EREAPEFASNLTKSAG TGIGLEFQESGLRLNA KNDRLVKAGMIFNVVS LGFQNLKAPTNNPKT ESFSLLLADTVIVREN VPEVATSTSSKAFKDV
	AtrSPT16	A. trichopoda	SPT16	

				RAERLYEHWEKGAD GLDSIKSLAFVYGETD NPYTKTSALFTWLF HEIADTVLLKKDHIYI LGSNRKVEFFGVS DNQSSGKVPTVSTLL RDKTDKDAGNFEKLI DHIKSAGGDVGNFVK EKFSSEFVSSWNKALE EGGVNKNDDTLAFTH LFAVKDDKEMDLIRK SAQATTASWTAARA RYVEIIDQEKVRHVS LSNEFAAFMKDSKVQ QALAKYEADTCYDPI VMSGGNYSFKWNHE SSESHLSQFGTIITSF GARLSEYCTNLTTRTM LIFPSSELETAYEAILA AELAVIAALPGAKLS DVYKIGIDLTLEKSPKL AETLNKKELGFATGIE FRESRLAISAKCDEVV KAGMVFIVYIGVDSIP NKNKGEKGKPAIAI SDTILVKEEGDNEILTE KAKSRLKSNVIKFKEE QENREAEKDNDQKK
NP_492821.1	CeSPT16	C. elegans	SPT16	
				AVPTGPREATLDAGN FCKRVKILYDSWNSD SNLWKSANSLVLALG QPNESNPKVTSLSQ TWLFGVELKDTIIVFLE KEIVIVSTSKINLFQK LSETEQVKTELSSIKFN FLTIDKSDKNKSNFEK LIGEATKAGSNIGVIIK ETYIGDLALQWEAAL NECPLTKVDITPALSS CLLVKDLQEQKNIITS AKITSKVLKSHILPKIET IIDKGERQTHNQLAD YAADIFESPEKISSKLT VEHVDYSYVPIIQSGG IYDLRASASSDDNPLH FGTIIVSCGARYKNYC SNIARTYIIDPTSDQK KNYAILLNVQSNVIKA IKPDVTFSSLYEKAIQT IKESSKPELVDFHFPKN VGYGIGIEFQESLAVL NATNSRTLKAGMTL NIACGFQKISNPEGK DEKSKTYLSLISDTVLL NDEGKVEVLTDVGKK ASDVVYMLGGEDDD
XP_640196.1	DdSPT16	D. discoideus	SPT16	
				KRLYTEWRAPSIGHD DALRNLDCIMSIVGV EEDVMYSKSMALQL WLLGYELTDTSVFCS DAVYFLTSKKKIEFLK QTQNITEEGFPEINLL VRDRTDKDQGNFEKL IKALQNSKKGKRLGV FAKDAYPGEFSEAWK KSLTASKFEHVDISTII AYLMCPKDESEINNIR KASLVSMDFINKYLK DEIMDIIDSDRKVKH NKLSDGCEAAIGEEK YTSGLDPRLLDMAYP PIIQSGGAYSLKFS ADKNPLHFGVIVCSL GARYKSYCSNISRTFL VNPTEAMQENYTFLLV SVQEEILKLLVPGTKLC DVYEKTLDFVKKEKPS MVDNLPKSGGFAMG LEFRENSIVIGPKCQA LLKKNMVFNLHVGIS NLTNPEATDKEGKNY ALFIGDTVLVGEQSPA SVMTPSKKKIKNVGIF IKDDSDDEEDVDDKKT
NP_476610.2	DmSPT16	D. melanogaster	SPT16	

NP_001091053.1	DrSPT16	D. rerio	SPT16	RLYGNWKKGEDEFG KVDAIVSVGVDEEIV YAKSTALQTWLFGYE LTDTIMVFCEKIIFLA SKKKVEFLKQVAVTK GNENANGVPPITLLV REKNESNKVNFDKMI EAIRGSKEGKTGVGFI KDKFPGEYMKSWSD TITAEGLQKVDISTVV AYTMAVKEDGELAL MKKAASITTDVFSKFF KERVMEIVDADEKVK HSRLAESVEKAIEDRK FLGGVDPSTVEMCYP PIIQSGGNYSLKFSVV SDKNHMHFGAITCA MGIRYSYCSNLVRTL MVDPSQEMQDNYN FLLQVEEELLKEMKH GVKLSEAYNAVMEFV KKEKPDVLSKLTKNLG FAMGIEFREGSLVIN QKNQFKLKRGMVFSI SLGFADMINKEGKKE EQKKYALFIGDTIQIN EEDQATVLTTPVKKIK NVGIFLKNDDDEDE
NP_009123.1	HsSPT16	H. sapiens	SPT16	RLYSNWRKGEDEYA NVDAIVSVGVDEEI VYAKSTALQTWLFGY ELDTIMVFCDDKIIIF MASKKKVEFLKQJAN TKGNENANGAPAITL LIREKNESNKSSFDKM IEAIESKNGKKIGVFS KDKFPGEFMKSWND CLNKEGFDKIDISAVV AYTIAVKEDGELNLM KKAASITSEVFNKFFK ERVMEIVDADEKVRH SKLAESVEKAIEKKYL AGADPSTVEMCYPPII QSGGNYNLKFSVSD KNHMHFGAITCAMG IRFKSYCSNLVRTLMV DPSQEVQENYNFLQ LQEELLKLRHGVKIC DVYNAMDVVKKQK PELLNKITKNLGFGM GIEFREGSLVINSKNQ YKLKKGMMVFSINLGFS DLTNKEGKKPEEKTY ALFIGDTVLVDEDEGP ATVLTSVKKVKKNVGI FLKNEDEEEEEKDE
NP_291096.2	MmSPT16	M. musculus	SPT16	RLYSNWRKGEDEYAS IDAIVSVGVDEEIVY AKSTALQTWLFGYEL TDTIMVFCDDKIIIFM ASKKKVEFLKQJANTK GNENANGAPAITLLV REKNESNKSSFDKMI DAIKESKSGKKIGVFS KDKFPGEFMKSWSD CLNKEGFDKVDISAV VAYTIAVKEDGELNL MKKAASITSEVFNKFF KERVMEIVDADEKVR HSKLAESVEKAIEKK YLAGADPSTVEMCYP PIIQSGGNYNLKFSVV SDKNHMHFGAITCA MGIRFKSYCSNLVRTL MVDPTQEVQENYNF LLQLQEELLKLRHGV KICDVYNSVMDVVKK QKPELLNKITKNLGFG MGIEFREGSLVINSKN QYKLKKGMMVFSINLG FSDLTNKEGKKPEEKT YALFIGDTVLVDEDEGP ATILTSVKKVKKNVGI FLKNEDDEEEEEKDE

				QINLELYGKRLLNN KWKEHKEMWGGGA DAIAVVTPPASEDLRY LKSTALHIWLLGYEFP ETVMVFMFGALHFV CSSKKAHLEELQKSS KMLTGVDIHHMKER KQDGSVQMNNSVLDA VKGFSKGKTPTVGVL SREATEGSVMEKWA ECLEASGAATVDVSG GFSEIFAVKDEEEISNI KNASHLSAAILKSFVV PKLEVIIDEEKVTHFE LMEQTENVITNPGKY VKLKAEDVDICYPPVF QSGGVFDLKPASVSN EEPLYDTLGVILCAIG ARFRSYCSNVARSIMI DADKTQEKAYKILLKA HEAAIAALRPGNAM SVAYKAAYAVVESGG PEFLPYFTKNAGTGIG IEFRESGLTLNAKNER VIRPGMAFNVSLGFH NLTTESNPKSKTFSLL LADTAIVVEKGPPPEVP TLKCSKTYTDIAYSFKD
XP_001783705.1	PpaSPT16	P. patens	SPT16	KLDKSFWRVRKALY DSWQKKDDYLWKG DALVIAQGTDDNDP YKKTNALQYWLLGYE FADTIIVFCESNIYVLT SPKKIGLLEVLKEENA PYVLCFLQLSKAENN QNLNTLFEAVHKS GKSGVLPEKMQGA LGSWEALAKSGLTL VDVAPGYSEMLSVKD TVEMKLLQAAETTS KVMSNFLIPAVETVID EGTQKSHALAEEME EVFAQPEKINSKLVGE HVESYPIIQSGGKY DLKLSATNDNENLHF GTIVCALGARYGYCS NAARTFFIDPKKEVEI NYRFLLEVHNIVVKS KPGVKISVIMDKATY IADKRPELSSFLKCC GYGIGIEFQETYLMLS SKNTRTIKPGMAFNI MVGFQNLKEEKP PKNKVYSMLLADTVI VSETGEAEVLTVSVSK RFGDVSFYIEELDNEE
-	PpSPT16	P. polycephala	SPT16	LLYSKYNEFEKSPNSL LFVLGSSNAENPYQK TTILHNWLLSYEFPAT LIALVPGKVIITSSAKA KHLQKAIIDFKDPESK ITLELWQRNNKEPEL NKKLFDDVIALINSAG KTVGIPEKDSYQKGF MTEWNPVWEAAVK ENEFNVIDISGLSKV WEVKDVNEQAFLSV SSKGSDKFMLLSNE MVRVDEELKITNAK LSDKIENKIDDVKFLK QLSPDLSALCPPNYKF NFDLLDWTSYPIIQSG KKFDLRVSARSTNDQ LYNGGICILASGIRYN NYCSNITRTFLIDPSEE MANNYDFLLTLQKEI VTNILKPGRTPEVYE SVIEIEKTKPELVNPF TKNIGSLGLEFRDSN FILNVKNDYRKIQRG DCFNISFGFNLLKDS QSANNYALQLADTV QIPLDETEPPRFLTNY TKAKSQISFYFNNEE
NP_011308.1	ScSPT16	S. cerevisiae	SPT16	

				QLLSVWNKQPQYAN IDAFIINKGKQGGN KIKTSAISMWYFGFD FIDTILLITKKTFAIIGG NKKINMLKSVQEHAE AKEYNLVFIKEDQAN NSNQLQQLFEILDKD LNKSSFNIGTLAKEQQ VGPFMTEDSFIDK NQYKFADCSFVQD CLSVKDQNEISYIGKA AKVSVYLESKLIKEIETI IEDEGKTHSQIATMI EGUENELKKISEEI GGESDNLDAVPIV QSGGKYDLKPNAQS NEDILSYDTIIVSVGTK YMEYHANIVRTLFDIP TNDQKKIQRVYELQ NQIAVQLKPGIKLKT YENAVNFINEKVPQL KDKIPANFGFGIGLEF RESNLYINAKNEKEVE EGMVFNVVVGFDNL QSEKEKAYAIQSDTV AIRKQNTPNAMVTFK VSKKYEDISYSIQDEG QDEEQEEEEDDLEKE
XP_001018189.2	TtSPT16	T. thermophilus	SPT16	
				RFFGSWKKGDDEFA NVDAIVSVGVDEEI VYAKSTALQTWLFGY ELTDTIMVFCEKILF MASKKKVEFLKQJAN TKGNNANGTPAITL LVREKQNESNKGNF KMIEAIKVSCKGRIG VFIKDFPFGDFMKS YDILNKESFEKVDISAS VAYTIAVKEEGELNL MKKAASITSDVFSKFF KDRVMEIVDADEKVR HGKLAESVEKAIEDKK YLGGTDPSTIEMCYP IIQSGGNYNLKFSVVS DKNHMHFGAITCAL GIRKSYCSNLVRTLM VDPTQEMQENYNFL LQLQEELKELKHGAK ICDAYQVIMDQVKK QKPDLMKITKTGFG AMGIEFREGSLVINN KNQYKLLKGMVFSV HGLAELNNKMGKK PEEKTYALFVGDTVLV NEEGAATVLTNVKKK VKNVGIFLKKEDDEEE
NP_001084166.1	XISPT16	X. laevis	SPT16	
				YAINIENFSKRLKVY DHWKEHKSDDLWGSS DAIAIATPPPSDDLRY LKSSALDIWLLGYEFP ETIIVFMHKQIHLVSS QKKGNLIGTLKKAAN EAVGVDIVLHVKT SDGADLMDDIVHAA RNQSKDKPVVGHIA KEAPEGKLLTWWIKL SSGSLRLVDVTNGFS ELFAVKDTTEITCVKK AAYLTSSVLKNFVIPKL EKVIDEEKVSHSSLM DDAEKAILDPLKVKVK LKPDNVVICYPVVFQ SGGKFDLPGASSND EYLYYDSAIICAGSK YSSYCSNVARTYLIDA TPTQNKAYETLRKAH EAAIQQVKPGNQMS AVYQAAVAVIERDAP ELLPNLTKSAGTGIGL EFRESGLNLNAKNDR RIKKGMVFNVSGLGH NIQAETTSEKTKQFSL LLADTVLVNERGHEIL TAPCSKAFKDVAYSF
NP_001105557.1	ZmSPT16	Z. mays	SPT16	

				GGKNPGLLKINSGGI QWKKQGGGKAVEV DRSDIVSVSWTKVTK SNQLGVKTKDGLYYK FVGFRDQDVPSSLSSFF QSSYGKTPDEKQLSV SGRNWGEVDLHGNT LTLVLGSKQAFEVSLA DVSQTQLQGKNDVT LEFHVDDTAGANEKD SLMEISFHIPNSNTQF VGDENRPPSQVFND TIVAMADVSPGVEDA VUTFESIAILTPRGY NVELHLSFLRLQGQA NDFKIQYSSVVRLL PKSNQPHTFVVISLDP PIRKQGTMYPHIVM QFETDTVSELSISD ELMNTKFKDKLERSY KGLIHEVFTTVLRWLS GAKITPKGFRSSQD GFAVKSSLKAEDGVL YPLEKGGFFLPKPPTLI LHDEIDYVEFERHAA GGANMHYFDLLIRLK TDHEHLFRNIQRNEY HNLYTFISSKGLKIMN
NP_189515.1	AtSSRP1	A thaliana	SSRP1	
				RGGTNPQGLRIHSGG IVWRKQGGGKVVEV GKSDLVGSWMKVP KSYQLGVRIKAGLVYK FIGFREQDVNNLNSFI SNTLGITPQEKQLSVS GRNFGIEELNGNMLT FLVGSKQAFEVSLAD VSQTQLQGKNDVSLE FHVDDTTGANEKDSL VELAFHIPNSNTTFVG DETRPPAQVFRDKIM SMADVGPSGEEAVV TFDGIAVLTPRGYTV ELHISFFRLLGQANDF KIQYSSVVRFLVLPKF NQPHTFVIITLDPPIR KGQTLPHIVLQFETE YVVESTLSISEDLLSTK YKDRLEASYGGLYDV FTAILRGLSGAKLTRP GKFRSCQDGYAVKSS LKAEDGILYPLEKSSFF LPKPPTLHDEIECLE FERHGAGGSSISSHYF DLLVRLKNEQEHLFR NIQRNEYHNLFEFINS KGLKITNLGETQATG
XP_006842123.1	AtrSSRP1	A. trichopoda	SSRP1	
				GHLAFGTLKLTKESLN FKGDKGKSVNVVTS DIDKLKWKLGKNP GLRVGLNDGGAHRF GGFKDLDLEKIQSFTS SNWSQSIDQSNLFK GWNYYQAEVKGKTV EFSWEDKPIFEIPCTN VSNVIANKNNEAVLEF HQNDDSKVQLMEM RFHMPIDLENEEDAD KVEEFKAVLAYAGLE AETEQPICLLTDILCTT PRGRYDIKVYPTSIAL HGKTYDYKIPKINRL FLVPHKDGRRHVYFVL SLNPPIRQGQTRYSYL IIEFFGKDEEQDLELAL TDEQLESSNGNLRRD MTGPİYETISILFKSIC NLKITVPGRFLGSSGT PAIQCTHRQNPGLLY PMEKGFLLHKKPAMY IRFEISSCHFARSDSG TVTRTFDFEIDLKYG PLTFNAMEKEENNKL FDYLNKKNIKIRNSQR VENTVADSSDDEIDP
NP_498633.1	CeSSRP1-A	C. elegans	SSRP1	

NP_491688.1	CeSSRP1-B	C. elegans	SSRP1	HLTCTGLTLTENSINFI GDKGGKSVITGTDV DKLKWQKLGNKPGL RVGLSDGGAHRFGG FLDDDLQKISFTSSN WSKSINQSNLFINGW NYGQADVKGKNEFS WENEPIFEIPCTNVSN VIANKNEAILEFHQNE QSKVQLMEMRFHM PVDLENEEDTDKVEE FKKAVLAYAGLEAETE QPICLLTDILCTTPRGR YDIKVYPTSIALHGKT YDYKIPVTINRFLVLP HKDGRQVYFVLSLNP PIRQQQTHYSYLIFE GKDEEEDLESLTDEQ LDYFNGNLQREMTG PIYETISILFKSICNLKV TVPGRFLGSSGTPAIQ CTHRQNLGLLYPMEK GFLFIQKPMYIRFEEI SSCHFARSDSGTVTR TFDIEDLKTGSSLTFS AMDKEENNKLFDYL NKKEIKIRNSHRIDNK SAGYSSDEDDIDDPY
XP_635805.1	DdSSRP1	D. discoideu	SSRP1	LGGRISGTRGILKFTT NNITWKESENGKIETV SSSDIKRANWARVTP RIFQLILSIKGGATVKF DGFKEQDYEVVRKYL SDQYNVSPLEIIELSSK GCNWGEVKVNGPMP IQFTTDHGKVGFEFPI SEVSQSVIGANNKNE LTLEFHHDKAMDDE DETMVEMRFFTPIRP SKEGEEGGKEKKVGE DGEDEDEDEDAEKE EEISALEQFQQTIMN KSDMVSNVSKSLVVF SAIQFLTPRGRIDIEM YPTFLKLHGKTHDYKV PYESISRLQFFRPDQ KHIFFISLDPPIRQGO TKYAHLVIQFQAEENI HLELNLDELQKQFK DQLSPIMNGNANALI CKILKALTGKKITIPGN FQSDSGANSIKCSLKA NEGLYPLERCFFFVH KPPTYIKFEDISNIEFA RYGAPSVRGGSNRTF DLSINLKNSTSIQFVNI
NP_523830.2	DmSSRP1	D. melanog	SSRP1	GVLCSGRLKMTQNI IFKNTKTGKVEQISAE DIDLINSQKFVGTWG LRVFTKGGVLHRTG FRDSEHEKLGKFIKAA YSQEMVEKEMCVKG WNWGTARFMGSVL SFDKESKTIFEVPLSHV SQCVTGKNEVTLEFH QNDDAPVGLLEMRF HIPAVESAEDPVDKF HQNVMSKASVISASG ESIAIFREIQLTPRGRY DIKIFSTFFQLHGKTF DYKIPMDSVLRFLML PHKDSRQMFVLSLD PPIKQGQTRYHYLVLL FAPDEETTIELPSEAE LRDKYEGKLEKESGP VYEVMGKVMKVLIG RKITGPGNFIGHSGT AAVGCSFKAAGYLY PLERGFYIHKPPLHIR FEEISSVNFARSGGST RSFDFEVTLKNGTVHI FSSIEKEEYAKLFDYIT QKKLHVSNMGKDKS GYKDVDFGSDNEN

				GSWNDGRLRFSKQT VVYKSHKTGKVDSP APDLSEAWRRVCL GHGLKATSTGHIYKY DGFKETDYEKISAFFK ANYKVELEEDKMCVK GWNWGTAKFAGSLL SFDVSDSPVFEIPLSSV SQCATGKNEVTVEFH QNDDAEVSLMEVRF YVPPNTGDDGSDPVE AFAQNILSKADVIQA TGDAVCIFKELQCLTP RGRYDIRIYPTFLHLH GKTFDYKIPYTTVLRLF LLPHKDQRQMFFVIS LDPPIKQGQTRYHFLI LLFSKDEDISLALNMS EDEVEKRYEGKLSKN MSGPLYEIVSRVMKA LVNRKITVPGNFQGH SGSQCITCAYKASSGL LYPLERGFYVHKPPV HLRFEEISCNVFARGT TTTRSDFEIEKQNN QFTFSNIREEYGKLF DFVNAKLTIKNRGF KEGMKGAEYDSDSD
A0A2R8QK93	DrSSRP1A	D. rerio	SSRP1	
				GSWNDGRLRFSKQT VVYKNSKTGKVDTP VPELTQAQWRRVCL GHGKLTSTGHIYK YDGFKDADLEKISEYF KDNVYKELTEKDMCV KGWNWGTAKFNGP LLSFDVNDSPTEIPLS SVSQCTTGKNEVTVE FHQNDDEVSLEMEV RFYVPPTTGDEGSDP VEAFAQNVLSKADVI QATGDAVCIFRELQC LTPRGRYDIRIYPTFLH LHGKTFDYKIPYTTVL RLFLPHKDQRQMFF VISLDPPIKQGQTRYH FLILLFSKEETISLTLM NEDEVERRFEGKLNK NMSGSLYEMVSRVM KALVNRKITVPGNFQ GHSGAQCTCSYKAS SGLLYPLERGFYVHK PPVHLRFEEIACVNFA RGTTTTRSDFEIEK QGNQYTFSSIEREEYG KLDFDFVNAKLSIKNR GFKEGMKGNDMDMY
A0A2R8QBX6	DRSSRP1B	D. rerio	SSRP1	
				GSMNDGRLRLSRQGI IFKNSKTGKVDNIQA GELTEGIWRRVALGH GLKLLTKNGHVYKYD GFRSEFEKLSDFEKT HYRLELMKEDLCVKG WNNWGTVKFGGQLLS FDIGDQPVFEIPLSNV SQCTTGKNEVTLEFH QNDDAEVSLMEVRF YVPPTQEDGVDPVEA FAQNVLSKADVIQAT GDAICIFRELQCLTPR GRYDIRIYPTFLHLHG KTFDYKIPYTTVLRLFL LPHKDQRQMFFVISL DPPIKQGQTRYHFLIL LFSKDEDISLTLMNE EEVEKRFEGRLTKNM SGSLYEMVSRVMKAL VNRKITVPGNFQGH GAQCITCSYKASSGLL YPLERGFYVHKPPVH IRFDEISFVNFARGTT TTTRSDFEIEKQGTQ YTFSSIEREEYKLFDF VNAKLNKIKRGLKE GMNPSYDEYADSDE
NP_003137.1	HsSSRP1	H. sapiens	SSRP1	

				GSMNDGRLRLSRQGI IFKNSKTGKVDNIQA GELTEGIWRRVALGH GLKLLTKNGHVYKYD GFRESEFEKLSDFFKT HYRLELMEKDLCKVG WNWGTVKFGGQLLS FDIGDQPVFEIPLSNV SQCTTGKNEVTLEFH QNDDAEVSLMEVRF YVPTQEDGVDPVEA FAQNVLSKADVIQAT GDAICIFRELQCLTPR GRYDIRIYPTFLHLHG KTFDYKIPYTTVLRFL LPHKDQRMFFVISL DPPIKQGQTRYHFLIL LFSKDEDISLTNNMNE EEVEKRFEGRLTKNM SGSLYEMVSRVMKAL VNRKITVPGNFQGH GAQCITCSYKASSGLL YPLERGFYVHKPPVH IRFDEISFVNFARGTT TTRSFDFEIEKQGTQ YTFSSIEREEYKLFDF VNAKKLNKRNGLKE GINPGYDDYADSD
NP_001129553.1	MmSSRP1	M. musculus	SSRP1	GRGGTNPQGLKINS GFQWRKAGGKVVVE VSTSDIASLWTRVPK GFQLGVKLKAGRDVK FNGFRESDMANLSF FTTNMGITSEEKPLAI SGRNWGDVTLEGSSL AFSVGGKHAFDVSA DVSQTQMQGKNEV MLEFHVDDTTGASEK DTLMELSFHIPTNTT YIGDEERPPAQIFREKI LSMGDVGPSGAEAIA LFEEVHILTPRGYNV ELHASFLRLQGQAND FKIQYSSVRLFILPKS NQPHTFVVITLDPPIR KGQTFYPHIVLQFPSE EIAECTLSIGEELLNTK YKDRLQPAYKDLSDH VFTQILKGLSGAKVTR PGKFRSAQDGYCVRT SLKAEEGTLYPLEKSFF FLPKPPTLULHDEIEYLE FERHGAAGTSSISSHY FDLIRLKSEQEHQFR NIQRNEYHNLNFI KNLKIMNLGDAQGTS
A0A2K1K1Z7	PpaSSRP1	P. patens	SSRP1	GRVSSTKGVKVNST MFLWRTEAGRVISVL PTDLSQAFWTRIYGR VFQLKLQNLIGNSIKF DGFRESDDVLKTFLS ENFHVELTTVQLATK GLNWGEYELAGPVIT LAVDGKHAFIPVSEI SGAVMQPTQKNEVA IEFHQDSTLEDEDESL VDIRFFIPNRSTEEGS KTSAILHKGLLDVIG GGIQKDRATLNKV PILTPRGYDIELYPTF MKFRGKTHDYKVITYE SVSRLQLPKPDQHII FFIVTLDPPIRQGTMR YPHLVMHFKMDEKS PPVPITLPADLKEKEQ FKSLPEEFKDADAYEV VNKLFRFTQRIKTP GSFKSHSGASAVKCA FKANDGFLFLERSFF FVHKPPLHMRFEVS SVEFARVASSTNAQA NRTFDLVIRLSDSTVH QFTGIQRNEYGSLFN FIQSKLRIENPDEAT
-	PpPO3	P. polycephala	SSRP1	

NP_013642.1	ScPOB3	S. cerevisiae	SSRP1	SGRFRIADSGLGWKIS TSGGSAANQARKPFL LPATELSTVQWSRGC RGYDLKINTKNQGV QLDGFSDQDDYNLKN DFHRRFNQVEQREH SLRGWNNWGKTDLAR NEMVFALNGKPTFEI PYARINNTNLTSKNE VGIEFNQDEEYQPA GDELVEMRFVPGVI QTNVDENMTKKEESS NEVVPKEDGAEGED VQMAVEEKSMAEAF YEELKEADIGEAGD AIVSFQDVFFTTPRGR YDIDIYKNSIRLRGKTY EYKLQHRQIQRVSLP KADDIHLLVLAIEPP LRQGQTTYPFLVLQF QKDEETEVLNLEDE DYEENYKDLKKQYD AKTHVLSHLKGLTD RRVIVPGEYKSKYDQ CAVSCSFKANEGYLY PLDNAFFFLTKPTLYIP FSDVSMVNISRAGQT STSSRTFDEVLVLRSN
XP_001014803.2	TtSSRP1	T. thermophilus	SSRP1	VTFFEDQLVLSHNKR RLCKLSLKKFTNSTVN KTDIVIDLNTVDLQDD EDQLCEMRLFIPQQQ DAQMKPEDDGESEE KEGEDSVGGYADQL NSEIITKAKIGQYSGQ SIVKFEDISLLVPRGKY QLDMYKTVRFHGSS FNVIVEYSNVKGFLLP QPDEVHVAFLGLD QPLKIGNTVHSYIVM QFKKEQKANKVNI PEEKEDKLDLDEEY DGFLYIAGQLFKTLC NNVQIIMPAGFQSSD KQNCCLKCTLKTHQGL LYPMRKSIFIKPVIH IQISDIQKVEFNVRGN ATLNKLFVVKVFTKTT TPQFFGFERKELDVLL EYFKSNKITYDDTN QGATFDDDEEFTDSI SEDEEEGKRAQRKSA KRANKAVKGQSGPLPL DDDEDEEDEFDADE FESEDEEEEEEDEDED EDFEQ
NP_001084164.1	XlSSRP1	X. laevis	SSRP1	GSMNDGRLRLSRAG LMYKNNKTGKVENIS AADIAEVVWRRVAL GHGIKLLTNGGHVYK YDGFRETEYDKLFDYF KSHFSVELVEKDLCKV GWNWGSVRFGGQL LSFDIGDQPAFELPLS NVSQCTTGKNEVTLE FHQNDDEVSLEIR FYVPPTQDDGGSVE AFAQNVLSKADVIA TGDAVCIFRELQCLTP RGRYDIRIYPTFLHLH GKTFDYKIPYTTVLRLF LLPHKDQRQMFFVIS LDPPIKQGQTRYHFLI LLFSKDEDMTLTLNM SEEEVERRFEGKLKKS MSGCLYEMVSRVMK ALVNRKITVPGNFLG HSGSQCITCSYKASSG LLYPLERGIYVHKPP VHIRFDEITCVNFARG TTTTRSFDEIETKQG SQYTFSSIEREEYGKLF DFVNAKLSIKNRGLK EGMKPAYDDYADSD

Q9LEF5	ZmSSRP1	Z. mays	SSRP1	RGGTNPQGQFKVHSG GLAWKRQGGGKTIET DKADVTAVTWMKVP RAY QLGVRIKAGLFYRFIG FREQDVSNLTNFIQK NMGVTPDEKQLSVS GQNWGGIDIDGNM LT FMVGSKQAFEVSLPD VAQTQMQGKTDVLL ELHVDDTTGANEDKS LMDLSFHVPTSNTQF V GDESPPPAHILWETIL KFADVGSSEEPVVTFE GIAILTPRGYSVELHL SFLRLQGQAND FKIQYSSIVRLFLPKS NNPHTFVVITLDPPIR KGQTLYPHIVIQFETE AVVERDLALSK ELLVEKYKDRLEESYK GLIHEVFTKVLRLSG AKVTRPGSFRSCQDG YAVKSSLKAEDGL LYPLEKGFFFLPKPPTL ILHEIEFVEFERHGA
NP_015377.1	ScNHP6A	S. cerevisiae	NHP6	MVTPREPKRTRKK KDPNAPKRALSAYMF FANENRDIVRSENPI TFGQVGGKLGKWK ALTPEEKQPYEAKAQ ADKKRYESEKELYNAT LA
NP_009647.2	ScNHP6B	S. cerevisiae	NHP6	MAATKEAKQPEPK KRTTRRKDPNAPKR GLSAYMFFANENRDI VRSENPDVTFGQVG RILGERWKALTAEK QPYESKAQADKKRYE SEKELYNATRA
-	PpNHP6A	P. polycephala	NHP6	MEGESQAPEDMVNE EEVKDTGKKKGKKA VKQLVKPAKAKEDK KKVDKEDKKKDAK DDKQKKKEEEAESEE QESDKKEEQEEDGA ETKAKDKTTAKGKVG AKKATTESKPKKEKA GKDKDKDKTEKKV KDPNAPKRQATAYLIF ANEQRETIKKDNPSFS PPDILKEAGARWKAL TDKEKRYQDKNEIEK HRYKEMEAYNANG GEKDTDKKGAKRKKG EEEEEEDEENAKPNK KAAAADKGKGPKG LLAKAGAKKK
-	PpNHP6B	P. polycephala	NHP6	MDTPKECPAKEDAP AEKTEQPAPEKKEAE KEEGKEGKDEEKDVK VSESEETLKQEPKRA RSAYAIFTAEQVPMV KKEHPDMSRTEITCI AAKWKAMTAEQKQ PFEEQAVNEKKRIKEA KEAAAEAKDKEAEK PAESEVRYSLRNSTVI KDGEELVNTKKRALE TEEDSANKKKCVTSEE
-	PpNHP6C	P. polycephala	NHP6	MITLHECLKTLLADTE NVTLDTEAELLYNFV SFISTRGTGKNKENTY PTPISSETSSQNEDDS DEEPTRTKTGTPIKRK KDPNAPKRPLTAFML YAHDKRAEIKAQNP MTFGALGKLGKGTW LTFDATQKKVYSOMA DKAKLTYEKELAEYTK NKQT
-	PpNHP6D	P. polycephala	NHP6	MSSDEESASSASEVEE VEEKKRGPKKGAKK AEKTKKATKKVAKKK KDPNAPKRPLSAFMY YSQQRPNVKKENPS ATFADVARIERGEQWK ALDASLKEYEKQAA TDKARWEKEKAAYD GGSSKGKKKAAAGS DDEEEESGSDE

-	PpNHP6E	P. polycepha	NHP6	MVRFPSYFGPQDFEF QKGRARPYFKRKRL LSRCPLVFQTRTHS QHSQEHKMVAKKEK KEAVEGERKSSRTSK QVERLSVSPPTILKKK ATKSTTASKTSPKK AGAKKGTKKAKGEK GPKRPASAYLLFANE NRAKVKKANPDATF GELGKLLGEAWSNAT AADKKKFTALADKDK ARYEKEKAALA
NP_001031485.1	AtNAP1	A thaliana	NAP family	MSPTSVRSREWEGPS RWTEYLGPEMAASV SSTRSSKQIDGHVGG STKALNIQWVVMIE VADGLMAKMYRLN QILEYDPVGHVFSEA FWKAGVFPNHPRICT LLSKKFPEHFSKLQLE RIDKFSLDSLHDGAEL HLQSLEPWQLLLDL MAFREQALRULDLSS TVITLLPHQNSLILHAF MDLFCAFVRVNLFAE KIPRKMLLQVYNLLH ALSRNDRDCDFYHRL VQFIDSYPPLKGLQE DLNFVSPRIGEVLEAV GPSIFLSADTRKLRNE GFLSPYHPRFPDILT SAHPMRAQDLANVT SYREWVLLGYLVCPD ELLRVTSIDIALVVLKE NLVVTLFRDEVSLYQ MVCEKEFGIGISFASA DSINLTMQYILLHEDY QLYVLPVLESKKMA KSGRTKQKEADLEYS VAKQVEKMISEVHEQ
NP_194341.1	AtNAP1;1	A thaliana	NAP family	MSNDKDSFNVSOLD AALKDEDAGLVNAL KNKLQNLQAGQSDVL ENLTPNVRKRVDALR DIQSQHDELEAKFREE RAILEAKYQTLYQPLY VKRYEIVNGTTEVELA PEDDTKVDQGEETA EEKGVPSFWLTALKN NDVISEEVTREDEGAL KYLKDIKWCKIEEPKG FKLEFFFDTPNYFKNT VLTKSYHMIDEPELL EKAMGTEIDWYPGK CLTKILKPKKKGSK NTKPITKLEDCESFFN FFSPPEVPDEDEDIDE ERAEDLQNLMEQDY DIGSTIREKIIPRAVSW FTGEAMEAEDEIDD DEEDDIDEDEDEEED DEEDDDDEDEEESKT KKKPSIGNKKGGRSQJ VGEGKQDERPPECK QQ
NP_001319544.1	AtNAP1;4	A thaliana	NAP family	MSNEENIKSDNKS SSDLPTIPALDIGAEEC DLAELKNLTKRPF VKKLSPKVTRKVLFLK DIQVTHDELEEF KSALEATYDNLKPLF AKRYEIVNGVVEAEA EKEGVPNFWLIAMKT NEMLANEITERDEAA LKYLKDIRSCRVEDTS RNFKLEFLFDSNLYFK NSVLKTYHVNDEG PVLEKVIKTDIEWFPG KCLTHKVVVKKTKK GPKKVNNIPMTKTEN CESFFNFKPPEIPEID EVDDYDDFDITMTEE LQNLMDQDYDIAVTI RDKLIPHAVSWFTGE ALVDEDDSDNDND DNDEKSD

				MVADKSKKSKIEEG EEENLEQIDAEVLV KLQEIQDDLEKINE SDEVLEVEQKYNVIR PVYDKRNEVIQSI WMTAFLSHPALGDL TEEDQKIFKYLNSLE EDAKDVKSGYSITFH TSNPFEDAKLTKTFT FLEEGTTKITATPIKW KEGKGLPNGVNHDD KKGKRALPEESFFT WFTDAQHKEDAGDE IHDEVADIKEDLWSN PLTYFNDADEEDFD GDDDGDEEGEDDD DEEEEDGEE
NP_001077822.1	AtNRP1	A thaliana	NAP family	
				MVTDKSKKAKTEEN VEQIDAEVLVLSIE IQDDLEKINEKASDEV LEVEQKYNVIRKPVYD KRNEIITIPDFWLTA FLSHPALGELLTEEDQ KIFKYLSSLDVEDAKD VKSGYSITFSFNP FEDGKLTKTFTFLEEG TTKITATPIKWKEGKG LANGVNHKNGNKR ALPEESFFTWFSDAQ HKEDVEDEMQDEVA DIIKEDLWPNPLTYFN NDADEEDFGDDDG DEEEKEGDSDEDDDE EDEVGEE
NP_564063.1	AtNRP2	A thaliana	NAP family	
				MVEKGKSKAQPVE DDVSDHIDGELVLSIE KLQEIQDELEKVN SDKVLEVEQKYNEIRR PVYGRRNEIQTIPDF WLTAFLSHPALCDLLS EEDQKIFKYLHSLDVE DSKDIKSGYTITFNFN TNPYFEDAKLIKLTFS DDGSTKVTGTTIKWK EGMDVSNGINHGKE GKKRPFSEGSFFTWF SETDQQKDFTEGLQD EVAEIIKEDLWPNPLK YFNNEAEGDFDEDD DDDENEDDEGESEDS SDEEGEDDS
XP_006836211.1	AtrNRP2A	A. trichopoda	NAP family	
				MVLSKGGKAKVSAPE EEGLEHIDGELVLSIE RLQEIQDELEKVN SDKVLEVEQKYNEIRR PVYNKRKEIQTIPDF WLTAQSHPALGELL TEEDEKVFYQLQSLDV QDFKDVKSGYSIAFN FSTNPYFENTKLTKTIS FSDEGTSKITGTTIKW KDGMEPMNGVDHK KKGKRAFTEESFES WFSDNRQNGGLSES EDESIVCDEVAEIIKED LWPNPLKYFNNDAD EEEMDEEEEEDEDD DALGKNENGSDDDD SDDDEQDGDDEDEA
XP_006827206.1	AtrNRP2B	A. trichopoda	NAP family	
				MLDLPVYILFHKSLPR WFMSEYDALEAKFFE ERAALEAKYQKQYEP LYTKRFEIVNGVVEVD SVKNDTTEETAADDN AKGVPDFWLTAMKT NEVLAGEIYERDEGAL KYLKDIKWCKIEDPKG FKLEFFFDTNPFKNS VLTKTYHVEEDDEPLL EKAIGTEIEWYPGKCL TQKVLKKPKKGVKN AKPITKTEPCESFFNFF KPPQVPEDDADIDED TAEELQNLMEQDYDI GTTIRDKIIPHAVSWF TGEAIDGEEFEGETG DEEDDDDDDEDGDD DEDEEDDDVEDEDD DDEDEEEGSKKSG VVRLKVRLLLSFSLFPP SRMQFTDIRLFFGV QKAGSRQNGDQGD RPAECKQQ
XP_020520079.1	AtrNAP1;4-X4	A. trichopoda	NAP family	

XP_011620892.1	AtrNAP1;4	A. trichopoda	NAP family	MSKQNDQLSPYNM ADLGSALPGAAALSA EDRAGLVNALKNKLQ HLAGQHSDVLESLS KVKKRVVVLQEIQSQ HDELEAKFFEERAAL AKYQKLYDPLYTKRYE IVNGIVEVDSVENEEA SEKTPRDDTDNNEEK GVPDFWLTAMKTNE VLSGEIYERDEEALKYL RDIKWCRIDNPKGFK LEFFFDPNPYFKNSVL TKTYHMIDDEPILER AIGTDIEWYPGKSLT QKVLKKPKKGSKNA KPIKTEDCESFFNFFS PPQVPEDDEDIDEET AEQLQNLMEQDYDI GSTIRDKIIPHAVSWF TGEAVQGDEFEDMD EEDDEDGDEGDGDD EEEEEEDEDEDED DEEEERGSRKKPVA AGHKIAGGVVGE QGGERPPECKQQ
NP_501422.1	CeNAP1	C. elegans	NAP family	MADQEHIDAGLLSTN FDMIQALPLNVKQRV CALKNLQMKTIQIES DFYKRVHELEIEFEGK FKSTFDQRKAIVAGE VEPTKEQIDTPILEGLE GDQLAELYKAAEADP SAKGKDFWLTALRT HDLVAEAEIEHDVPIL SYLTDVTTAASKDPA GFKIEFHATNPYFKN QVLTKTLLGFDPAE APLQFDGPHVIRAVG DTIEWEDGKNVTKKA VKKKQKKGANAGKFL TKTVKADSFNFFPP KSKDERNEDEDEQA EEFLELDYEMGQAIR DTIIPRAVLFTYGELQS DDMFDFPGEDGDDV SDFSDEA
NP_001021332.1	CeSPR-2A	C. elegans	NAP family	MSEEPAAKRMKNAD EFPNLRSEVAGPVKE KLIELDAIQHQLDVM SENAAEVLKVEQIFN QKRLPIYEKRKLTTKI DNFWQTAFLNHHLL STAIPPEEQEDLLAAL DLEVQEFEDLRSGFKII MTFDPNEYFTNEVIT KSYHLQSESPSTEITEI EWKENKKPPFPEDG DSAHTFLEWLTYAAL PDSDEIAEVIKDDLTV NPLQYYVMPDMQEV EEDDIEDFLNEERGVD ENGQRIIPRNISDSLK VDQDESADGQEGED EEEEDMGEEEDGVE EEAEGEEEEEGAIEE EGGDENVEAHVAVN PENTAE
NP_001021333.1	CeSPR-2B	C. elegans	NAP family	MSEEPAAKRMKNAD EFPNLRSEVAGPVKE KLIELDAIQHQLDVM SENAAEVLKVEQIFN QKRLPIYEKRKLTTKI DNFWQTAFLNHHLL STAIPPEEQEDLLAAL DLEVQEFEDLRSGFKII MTFDPNEYFTNEVIT KSYHLQSESPSTEITEI EWKENKKPPFPEDG DSAHTFLEWLTYAAL PDSDEIAEVIKDDLTV NPLQYYVMPDMQEV EEDDIEDFLNEERGVD ENGQRIIPRNISDSLK VDQDESADGQEGED EEEEDMGEEEDGVE EEAEGEEEEEGAIEE EGGDENVEAHVAVN PENTAE

NP_958876.1	DrSETB	D. rerio	NAP family	MSASAAKVSKEQN SNHDGADETSEKEQ QEAIHIDEVQNEIDR LNEQASEEILKVEQKY NKLRQPFQKRSELIA KIPNFWVTTFVNHPQ VSALLGEEDEEALHYL TRVEVTEFEDIKSGYRI DFYFDENPYFENKVLS KEFHNLNESGDPSSKST EIKWKAGKDLTKRTG QTPNKAGKKRQHEE PESFFTWFTHSDAG ADELGEVIKDDIWPN PLQYYLVPDMDEEEG EGEEEDDEEEEGLED IDEEGEDDDGEEDEE EDEGEDGEDDGEDD
NP_958883.1	DrSETA	D. rerio	NAP family	MSASAAKVSKEKELNS NHDGGDETSEKEQQ EAIEHIDEVQNEIDRL NEQASEEILKVEQKY NKLRQPFQKRSELIA KIPNFWVTTFVNHPQ VSALLGEEDEEALHYL TRVEVTEFEDIKSGYRI DFYFDENMYFENKVL SKEIHLNESGDPTSKS TEIKWKPGKDLTSRSS QTQSKAGKKRQHEE PESFFTWFTHDADSG ADELGEVIKDDIWPN PLQYYLVPDMDEEEG EGEDEDDEEGLDDI DEEGDDDGEEEDDD DGEDDEGD
AAH52128.1	DrNAP1L1A	D. rerio	NAP family	HKALASLTGGDADD DCKEMPCSSKKASRE KSRDKSNRPGCSQSD RRTEAAAADSSSL CEDQEDPSFVDDDEE EEEDDSLPGSSSTAS SSVASDNEDNEDGEC AIVSVKMAPEVRQSV ALLAQVQMRLDALEK KGARLHQRLEMKLSR QRRPHLDQRSITQA VPGFWVTALLNHPH LSAHIDETDEDAISY MTNLEVESFKNNKLG YRICFHFRNPFFQN KMIVKELHLMGGG PVSFSNPILWHRGQN LVGSGEPRTSQGVY QSFFHWFSDHSNPG RDDIAQLREDLYRNP LRYLTPLWEPRENG SSAPKPGSSNRDEC VIISDSDEEEQNSQS LGQEDDEDDQGGV SEENERDEGDSCEE RDEEEVDVEEVESH DSGGCEAREQEEEDI DVDEEKS
NP_956349.1	DrTSPY	D. rerio	NAP family	RKRSASPSGECSTQL KQAKVSVEEQKNP QEDAEPKVGNGEPK DGAEPSEGAKEEAKI EDEVAAQESTGQAN VSGGESARQPSDSAA IAAAEALASLTGGDA DDCKEMPCSSKKAS REKSRDKSNRPGCSQ SDRRTAAAAADSSSL LLCEDQEDPSFVDD EEEEEDSLPGSSST ASSVASDNEDNEDG ECAIVSVKMAPEVRQ SVALLAQVQMRLDAL EKKGARLHQRLEMKL SRQRRPHLDQRSIT QAVPGFWVTALLNH PHLSAHIDETDEDAIS YMTNLEIESFKNNKL GYRICFHFRNPFFQ NKMIVKELHLMGGG SPVSFSNPILWHRGQ NLVSGEPRTSQGV YQSFFHWFSDHSNP GRDDIAQLREDLYRN PLRYLTPLWEPREN GSSAPKPGSSNRDE

				MDKKNEKGARGKG GTKPDKPEAFLQVER VPYFNLLPKNVKRRV YALKRLQLQSGNIEAK FYEEVHELERYAGLY QPILDKRREIVTGAVE PTDEECEWHSKDKEED ELADDLQKKALEEK QADSASAVDPKGIPD FWLTIFKHVDMLGE MLQEHDPEILKHLQD VTVKFSEPGQPMST LEFHFEPNSYFSNTVL TKVYKMKSEPDAAADP FSFEGPEIVDCEGCKI DWHKGKDVTVKIVK KKQKHKGRTIRTVS KEIPQDSFFNFFNPVK ASPDGMDEDLDFTLA TDFEIGHFRERVIPR AVLYFTGEALEDDES EEEELEEGEEEDLDEE GEEEEEGDFDPTA
NP_001082816.1	DrNAP1L4a	D. rerio	NAP family	
				MEADNEAKGVSDQ AMQNPQALALQDK LDSVSHSASIMDSLPK SVKRRVNALRNQLQVD STHIEAKFYKEVHELE RKYSALYQPLFDKRRK VVSGEVEPTDEECEW QSDHEDEAALAECLK KKAIEEETEDANEK PKGIPEFWLTIFRSVD MLSDMLQEHDPEILK HLHDIKVFSGPDQP MSFTLEFHFEPNEYFT NTVLTKVYKMKSEPD ADDPFSFEGPEIIDCE GCEIDWQKGKDVTV KTIKKQKHKGRTVS RVITKQVPNDSFFNFF NPIKVSPDKELDEDESE YTLATDFEIGHFRERI VPRAVLYFTGEALED DESFEEDLEDEGEED QDDEEEEEEEEGDS KTGDPQPAECRQQ
NP_001007454.1	DrNAP1L4b	D. rerio	NAP family	
				MDKKNEKGARGKG GTKPDKPEAFLQVER VPNFNLLPKNVKRRV YALKRLQLQSGNIEAK FYEEVHELERYAGLY QPILDKRREIVTGAVE PTDEECEWHSKDKEED ELADDLQKKALEEK QADSASAVDPKGIPD FWLTIFKHVDMLGE MLQEHDPEILKHLQD VTVKFSEPGQPMST LEFHFEPNSYFSNTVL TKVYKMKSEPDAAADP FSFEGPEIVDCEGCKI DWHKGKDVTVKIVK KKQKHKGRTIRTVS KEIPQDSFFNFFNPVK ASPDGMDEDLDFTLA TDFEIGHFRERVIPR AVLYFTGEALEDDES EEEELEEGEEEDLDEE GEEEEEGDFDPTKDP PPAECKQQ
XP_005170732.1	DrNAP1L4-X1	D. rerio	NAP family	

				MADIDNKEQSELDQ DLDDVEEVEEEETGEE TKLKARQLTVQMMQ NPQLAALQERLDGL VETPTGYIESLPRVVK RRVNALKNLQVKCA QIEAKFYEEVHDLERK YAVLYQPLFDKRFEL NAIYEPTEECEWKP DEEDEISEELKEKAKIE DEKKDEEKEDPKGIPE FWLTVFKNVDLLSDM VQEHDEPILKHLKDIK VKFSDAGQPMSEVLE FHFEPNEYFTNEVLTK TYRMRSEPDSDPFS FDGPEIMGCTGCQID WKKGKNVTLKTIKKK QKHKGRTVRTVTKT VSNDSFFNFAPPEV PESGDLDDAEAILA ADFEIGHFLRERIIPRS VLYFTGEAIEDDDDDY DEEGEEADEEGEEEG DEENDPDYDPKKDQ
NP_001294853.1	HsNAP1L1	H. sapiens	NAP family	NPAECKQQ
				EEAGNQIMVEGLGE HLERGEDAAAGLGD DGKCGEEAAAGLGEE GENGEDTAAGSGED GKKGGDTDEDSEAD RPKGLIGYVLDTFVE SLPVKKYRVLALKKL QTRAAANLESKFLREF HDIERKFAEMYQPLL EKRRQIINAIYEPTEEE CEYKSDSEDCDDEEM CHEEMYGNEEGMV HEYVEDDDGYEDYYY DYAVEEEEEEEEDDI EATGEENKEEDPKGI PDFWLTVLKNVDTLT PLIKKYDEPILKLLTDIK VKLSDPGEPLSFTLEF HFKPNEYFKNELLTKT YVLKSLAYYPHPYR GTAIEYSTGCEIDWN EGKNVTLKTIKKQK HRIWGTIRTVTEDFPK DSFFNFFSPHGITSNG RDGNDLGLGHNLRT YIIPRSVLFFSGDALES QQEGVVREVNDIY
NP_068798.1	HsNAP1L2	H. sapiens	NAP family	DKIIYDNWMAAIEEV
				HGVAEEEMASSTSDS GEESDSSSSSSTSDSS SSSSTSGSSSGSGSSS SSGSTSSRSRLYRKRR VPEPSRRARRAPLGT NFVDRLPQAVNRNV QALRNIQDECDKVD LFLKAIDLERKYAEL NKPLYDRRFQIINAEY EPTEECEWNSDEEE FSSDEEVQDNTPSEM PPLEGEENPKENPE VKAEEKVPKEIPEVK DEEEKVPKEIPEVKA EKADSKDCMEATPEV KEDPKEVPQVKADDK EQPKATEAKARAAR ETHKRVPEERLQDSV DLKRARKGPKKREDP KGIPDYWLIVLKNVD KLGPMIQKYDEPILKF LSDVSLKFSKPGQPV YTFEFHLPNPYFRNE VLVKTYIIKAKPDHND PFFSWGWEIEDCKG CKIDWRRGKDVTTT TQSRRTATGEIEIQPR VVPNASFFNFSPPEI
NP_004529.2	HsNAP1L3	H. sapiens	NAP family	

NP_005960.1	HsNAP1L4	H. sapiens	NAP family	MADHSFSDGVPSDS VEAAKNASNTEKLT QVMQNPRVLAALQE RLDNVPHTPSSYIETL PKAVKRINALKQLQ VRCAHIEAKFYEEVH DLERKYAALYQPLFDK RREFITGDVEPTDAES EWHSENEEEKLAGD MKSKVVVTEKAAATA EEDPKGIPEFWTIF RNVDMSELVQEQYDE PILKHLQDIKVKFSDP GQPMSEVLEHFHFPN DYFTNSVLTKYKMK SEPDKADPFSEFGEPI VDCDGTIDWKKGK NVTVTIKKQKHKG RGTVRTITKQVPNESF FNFFNPLKASGDGES LDESEFTLASDFEIG HFFRERIIVPRAVLYFT GEAIEDDDNFEEGEE GEEEELEGDEEGEDE DDAEINPKV
NP_715638.1	HsNAP1L5	H. sapiens	NAP family	MADSENQGPAPESQ AAAAEAAAEVMA EGGAQGGDCDSAAG DPDSAAGQMAEEPQ TPAENAPKPNDFIES LPNSVKCRVLALKKLQ KRCDKIEAKDKEFQ ALEKKYNDIYKPLLAKI QELTGEMEGCAWTL EGEEEEEEEDDEEE GEDEEEEEAAEAAA GAKHDDAHAEMPD DAKK
A6NFF2	HsNAP1L6	H. sapiens	NAP family	MMEGLGEHSTAGE MGPLLGAVAATASP QSLMEYSSDADFIESL PLVVKYRVYTLKKLQA KCAVLEAKYLFHFHSV ERKFATYIGPLLEKRR QITNALYEPTKEECER
NP_001071165.2	HsTSPY1	H. sapiens	NAP family	MRPEGSLTYRVPERL RQGFCEGVGRAAQAL VCASAKEGTAFRMEA VQEGAAGVESEQAA LGEEAVLLDDIMAEV EVVAEEEEGLVERREEA QRAQQAVPGPGPM TPESAPELLAVQVEL EPVNAQARKAFSRQ REKMERRRPHLDRR GAVIQSVPGFWANVI ANHQPMSALITDEDE DMSYMSVLEVGEEK HPVHLCKIMLFFRSN PYFQNKVITKEYLVNI TEYRASHSTIEWYPD YEVEAYRRRHHNSSL NFFNWFSHDHNFAGS NKIAEILCKDLWRNPL QYYKRMKPPEGTTET SGDSQLLS
NP_003300.1	HsTSPYL1	H. sapiens	NAP family	HSIIISDQVPSDQDAH QYLRLRDQSEATQV MAEPGEGGSETVALP PPPPSEEGVPQDAA GRGGTPQIRVVGGR GHVAIKAGQEEGQP PAEGLAAASVVMMA DRSLKKGQVQGEKAL EICGAQRSASELTAG AEAEAEVKTGKCAT VSAAVAERESAEVVK EGLAEKEVMEEQME VEEQPPEGEEIEVAEE DRLEEEAREEEGPWP LHEALRMDPLEAIQLE LDTVNAQADRAFQQ LEHKFGRMRRHLYER RNYIIQNIPGFWMATA FRNHPQLSAMIRGQ DAEMLRYITNLEVVEL RHPRTGCKFKFFFR NPYFRNKLIVKEYEVR SSGRVVSLSPIIWRR GHEPQSFIRRNQDLIC SFTWFSHDHSLPESDK IAEIIKEDLWPNPLQY YLLREGVRRARRRPLR EPVEIPRPFQGSQ

				LSSSESQQRDP PPPPP PPPLRLPLPPQQR RLQEETAAQVLAD MRGVGLGPALPPP PYVILEEGIRAYFTLG AECPGWDSTIESGYG EAPPTESLEALPTPE ASGGSLEIDFQVVQS SSFSGGGALETCSAV GWAPQRLVDPKSKE EAIIVEDEDEDERES MRSSRRRRRRRRRK QRKVKRESRERNAER MESILQALEDIQLDLE AVNIKAGKAFRLKRK FIQMRRPFLERRDLII QHIFGFVVKAFNLH PRISILINRRDEDFRYL TNLQVQDLRHISMGY KMKLYFQTNPYFTN MVIVKEFQRNRSGRRL VSHSTPIRWHRGQEP QARRHGNQDASHSF FSWFSNHSLPEADRI AEIKNDLWVNPRLYY LRERGSRIKRRKQEM KKRKTGRCEVVIME DAPDYYAVEDIFSEIS
NP_071400.1	HsTSPYL2	H. sapiens	NAP family	
				MADIDNKEQSELDQ DLEDVEVEEEETGEE TKIKARQLTVQMMQ NPQLAALQERLDGL VDTPTGYIESLPKVVK RRVNALKNLQVKCA QIEAKFYEEVHDLERK YAVLYQPLFDKRFELI NAIYEPTEECEWKP DEEDEVSEELKEKAKI EDEKKDEEKEDPKGIP EFWLTVFNVDLLSD MVQEHDEPILKHLKD IKVKFSDAGQPMSEV LEFHFEPNDFYTNVNL TKTYRMRSEPDSDP FSFDGPEIMGCTGCQ IDWKKGNVTLKTIK KKQKHKGRTVRTVT KTVSNDFFNFAPPE VPENGLDDDAEAIL AADFEIGHFLRERIIPR SVLYFTGEAIEDDDDD YDEEGEEADEEGEEE GDEENDPDYDPKKD QNPAECKQK
NP_056596.1	MmNAP1L1	M. musculus	NAP family	
				EELGSQVMAEGPGES QDRSEGVSIEPGDGG QHGEETVAAGVGEE GKGEEAAAGSGEDA GKCGGTDESDSDRP KGLIGYLLDTDFVESLP VKVKCRVLALKKLQT RAAHLESKFLREFHDI ERKFAEMYQPLLEKR RQJINAVYEPTEEECE YKSDCEDYFEEEMDE EETNGNEDGMVHE YVDEDDGYEDCYDY DDEEEEEEDDSAGA TGGEENVNEEDPKGIP DFWLTVLKNVEALTP MIKKYDEPILKLLTDIK VKLSDPGEPISFTLEF HFKPNYFKNELLTKT YVLKSLACYDPHPYR GTAIEYATGCDIDWN EGKNVTLRTIKKKQR HRVWGTVRTVTEDF PKDSFFNFSPHGSL NGGDENDDFLLGHN LRTYIIPRSLVFFSGDA LESQQEGVVREVND E IYDKIYDDWMAAIEE
NP_032697.2	MmNAP1L2	M. musculus	NAP family	

				HGVAEEAMASTACD SGDESDSNSSSSTNSC SSSGSSSSGSSSSSSSS SSSSSSSSSSSSGSSGS SSNGSHLNRKKRVPE PSRRAQRRPSGKLFL DKLPQAVRNRVQAL RNIQNECDKVDTLFL RAIHDLERKYAELNKP LYDKRFQIINAEYEPT EEEECEWNSEEFSGD EEMQDDTPNEMPPL EGEEEEECNEKAEVK EEGTHVPEEVPEAKV EEEEAPKETPEVKTEE KDIPKEGAEKAEQEE SSKEIPEVKGEKADS TDCIDIAPEEKEDVKE VTQANTENKDQPTTEE FTPRAPAREAQKRVP ETRPEEGVNIKRAK GKPKKEDPKGIPDYW LTVLKNVDKLGPMIQ KCDEPLKFLSDVSLKF SNPGQPIGYTFEFHFL PNPYFRNELLMKTYII RSKPDHYDPFFAWG WEIEECKGCKIDWRR
NP_620081.1	MmNAP1L3	M. musculus	NAP family	
				MAENSLSDGGPADS VEAAKNASNTEKLT QVMQNPQVLAALQE RLDNVSHTPSSYIETL PKAVKRRINALKQLQ VRCAHIEAKFYEEVH DLERKYAALYQPLFDK RREFITGDVEPTDAES AWHSENEEEDKLAG DMKNKVIVIAEKAAAT VEELNPKGIPEFWFTI FRNVDMSELVQEYD EPILKHLQDIKVKFSD PGQPMSFVLEFHFEP NDYFTNPVLTKTYKM KSEPKADPFSFEGPE IVDCDGTIDWKKGK NVTVKTIKKQKHKG RGTVRTITKQVPNESF FNFFSPLKASGDGESL DEDEFTLASDFEIGH FFRERIVPRAVLYFTG EAIEDDDNFEEGEEG EEEELEGDEEDED DADVNPKV
NP_001272418.1	MmNAP1L4	M. musculus	NAP family	
				MADPEKQGPAAESRA EDEVMEGAQGGEDA ATGDSAAAPAAEPP QAPAEANAPKPKDF MESLPNSVKCRVLAL KKLQKRCDKIEAKFDK EFQALEKKYNDIYKPL LAKIQELTGEMEGCA WTLLEGEDDEDEED DEEEEEEEAAGAT GGPNFAKK
NP_067407.1	MmNAP1L5	M. musculus	NAP family	
				MSSPERDEGTPVPDS RGHCDADTVSGTPD RRPLLGEKAVTGEG RAGIVGSPAPRDVEG LVPQIRVAAARQGES PPSVRGPAAAVFVTP KYVEKAQETRGAESQ ARDVKTEPGTVAAAA EKSEVATPGSEEVME VEQKPAGEEMEMLE ASGGVREAPPEAGP WHLGIDLRNPLEAI QLELDTVNAQADRAF QHLEQKFGRMRRHY LERRNYIIQNIPGF MTAFRNHPQLSAMI RGRDAEMLRVTSLE VKELRHPKTGCKKFF FRRNPYFRNKLIVKEY EVRSSGRVVSLSPII WRRGHEPQSFIRRN QDLICFFTFWSDHSL PESDRIAEIHKEDLWP NPLQYYLCREGIRRP RRPIREPVEIPRPF QSG
NP_033459.1	MmTSYL1	M. musculus	NAP family	

NP_084112.1	MmCINAP/M	M. musculus	NAP family	SSSEPRQRDLPPPPPP PLQLRLPPPQQRRPR PQEETEAQVLADM RGVGPTLPPPLPYVIL EEGGIRAYFTLSAESP GWDHAMESGFGEA PSTGIMETLPSSEISG GSLAIDFQVAEPSSLG EKALETCSLGGWGPQ MLVGPKRKEEAIIVE DEDEDDKESVRRRQR RRRRRRKQRKAKESR ERSAQRMESILQALES IQMDLEAVNIKAGKA FLRLKRKFIQMRRPFL ERRDLIQHIPGFVWK AFLNHPRISILINQRD RDIFRYLTNLQVQDLR HISMGYKMKLYFQTN PYFTNMVIVKEFQRN RSGRLVSHSTPIRWH RGQEPQAYNRRSHD TRESFFNWFSNHSLP EADRIAIEIKNDLVV NPVRYMRRGGYRS SRKKQHGERAKNQ YEMVIMEDAHDHYAI EDILSDISEIDEITDNET
NP_941019.2	MmTSPYL3	M. musculus	NAP family	MAWEGIGVLEATMC LPPIQERDAHPARPR VPGAELRPCNPAAS PATAPPSSSENTGENS ASRVNLNDGCEKSGG NGVGTKKKTVEVTTE EDSVAAEPAEVGEK LEWVVAEQESLRPLD LRALIVDPLEAIQWEL EAMSAQADGAHLQL VRRFGRMRRLLHARR SFIIQNIPGFWVTAF NHPQLSAMISPRDED MLGYLMNLEVREL RH SRTGCKFKFLFESNPY FRNEVIVKEYECRASG GVVSIATRILWHRGQ EPPALVHRNRDAVRS FFSWFSQHSLEADR VAQIIKDDLWPNPLQ YYLLGQRPYRARRSLA RWSAEARPRPYGFQ SG
NP_084479.1	MmTSYL4	M. musculus	NAP family	MNGVEGNNELSLAN TTTPSHASEDLDLKQ DQGLQEETDTVREM EAAGEAGADGGASP DSEHCPELCFRVAE NSCAAAARGLEDAPS PSKGGDAPSAPVAA DDSSKNGCQLEGPHS PAKPKALEACGAVGL GSQQMPGPKKTKEM TTTKCAISVATGKEGE AGAAMQEKKGLQKE KKVAGGKKEETRPRA PKINCMDSLAIDQEL SNVNAQADRAFLQLE RKFGMRRLHMQRR SFIIQNIPGFWVTAFR NHPQLSPMISGQDE DMMRYMINLEVEEL KQPRVGCKKFIFQS NPYFRNEGLVKEYER RSSGRVVSLSPIRW HRGQEPQAHHRNR EGNTIPSFFNWFS DH SLLEFDRIAIEIKGELW SNPLQYYLMGDGPR RGVRVPPRPQPVESPR SFRFQSG

NP_001078890.1	MmtTSPYL5	M. musculus	NAP family	MSGRSRGRKSSRAK GRGKGRARARVRAA AEDAWHDEKPPQSP RLGEDSAAAQVQAG AAQGGAEPAELREEA ACRLPLDCGLALRAR AADERGLAAPDPDLE RARSLAERLTSDFV GTVGALAKLRGSR GNRRVPGRKAPDTRS ATGRGPQATVSGKPK MASAGLCAAAPVGE EKKMTEKHAGAGSP ATVGSMDTLETVQLK LETMNAQADRAYLRL SRKFGQLRLHHLERR NLIQSIPIGFWGQAF QNHPLQLSAFLNTKDK EVLSYLNRLVEEELGL ARLGYKIKFYFGRNPY FQNKVLIKEYGCGPS GQVVSRSAPIQWLP GHDLSLSKENPENN GSFFGWFSNHSSIES DKIVEINEDLWPNPL QYYLISEEARGEKKE ERPGPAKLSAPAVR QPN
NP_012974.1	ScNAP1	S. cerevisiae	NAP family	MSDPRTKPKSSMQI DNAPTPHNTPASVLN PSYLKNGNPVRAQA QEQQDKIGTINEEDIL ANQPLLLQSIQDRLG SLVGQDSGYVGGLPK NVKEKLLSLTLQSELF EVEKEFQVEMFELEN KFLQKYKPIWEQRSRI ISGQEQPKPEQJAKG QEIVESLNETELLVDE EEKAQNDSEEEQVKG IPSFWLTALENLPIVC DTITDRDAEVLEYLQD IGLEYLTDGRPGFKLLF RFDSSANPFFTNDILC KTYFYQKELGYSGDFI YDHAEGCEISWKDN AHNVTVDLEMRKQR NKTTKQVRTIEKITPIE SFFNFFDPPKIQNED QDEELEEDLEERLALD YSIGEQLKDKLIPRAV DWFTGAALFEFEED EEEADEDEDEEDDD HGLEDGDSAEQ DDFAGRPEQAPECK QS
NP_014153.1	ScVP575	S. cerevisiae	NAP family	MMSDQENENEHAK AFLGLAKCEEVDAIE REVELYRLNKMMPVY EKRDAYIDEIAEFWKI VLSQHVSFANYIRAS DFKYIDTIDKIKVEWL ALESEMYDTRDFSITF HFHGIEGDFKEQQVT KVFQIKKGKDDQEDG ILTSEPVPIEWPQSYD SINPDLIKDKRSPEGK KKYRQGMKTIFGWF RWTGLKPGKEPHG DSLASFSEEIYPFCVK YYAEAQRDLEDEEGE SGLSADGDSEDDDGS LGEVDLPLSDEEPSSK KRKV

				MGKGSTFGGPPSAN AAAEHDEEELD LAEL QDALGSPHELPPVV RRRVDALSNLHKQYL TLQKQYHAEVRELER KYQKLYDPVFDKRAN IVSGSYEPTDEEAKVE KAEEEEGVKKREAQP GDADIKGVPEFWLTV LQNNGFVAQSIFEED EEALSFLRDIRYSELEG ENEGFTLTFFHAENPF FSNTTSLKTYHMEQD ELMGELQYDHATGTK IEWKAGKNLTVKQTK KKTAKAGKKPARVVT VEEPCDSFFSFFSPVQ LPASEEEELSEEDAQN MEMDFEIGCIKDKLI PFSVLWFTGEAAEYE DDDDDMFGDEEGD DDEEDDEEDDDEEEE EEAPKRGKRGQGGG PGHSFAPPPGAKPGA
-	PpNAP1L1	P. polyceph	NAP family	GAPGQQQPECKQQ
				MAAPKTRKSTEPED TVDDEDIQVLEGIQD QIEELNDKATEILEVE RKYNRLRKPMDKRA EAIKLPFWFKVFTS QDMLAALLVGEDLA AFKYMKELNVEDHD DVKSGFKISLKFDPNP FFKNELLSKEFRYDKE GNLTVIPTKIEWKDG KDLTKKSQSDTKGK RSRDEESESFTWFN PDDQDLAEIIEEL WPNPAKYLLGLVEDE AEEEGDDDEGEDD GEDGEGEGEGEAE EDQ
-	PpSET	P. polyceph	NAP family	
				MAPKRSPLPPQKK KPRPPPALGPEETSAS AGLPKKGEKEQQAIE EHIDEVQNEIDRLNE QASEEILKVEQKYNKL RQPFQKRSELIAKIP NFWVTTFVNHPQVS ALLGEEDEALHYLTR VEVTEFEDIKSGYRIDF YFDENPYFENKVLKE FHLNESGDPSSKSTEI KWKSGKDLTKRSSQT QNKASRRQHHEEPES FFTWFTHSDAGAD ELGEVIKDIWPNPL QYYLVPDMDDDEEGE GEEDDDDEEEGLE DIDEEGDEDEGEDE DDDEEGEEDEGEDD
NP_001116293.1	HsSET	H. sapiens	NAP family	D
				MSENNEIEMELPFDS SAITSTEVLDRVNALL TLQDTQNELTEQMEK EILEIEKKYKKFQPLA EKRFEIVSGKVEPTKE DQCKKAPIQVENLKS VPTDKGIPKFWLHVL QNTEVKDIIIECDIEAL EYLVDIKIVQVGDAQ DYSLDFHFSENPFFTN TVISKTVKLEEDNELN EIVSTPINWKDGKNF TVQSKKKTVKSPTK GKAATTTSTTVQEVV PCFFSTFVSPNQDPTS DEEAEIMYIQYQIA KLKDIVIPEAVNFFLG RASDAEENDYDFGED FEDEEGEDDEEDDE EEQTIKKPSGKGAQ PQQPQDCKQQ
XP_645850.1	DdNAP1	D. discoideum	NAP family	

XP_642786.1	DdNAP2	D. discoideur	NAP family	MSKHQKKEVEFEK DDPKAVYLELFDKTE LKSLSDEIEKLTGEKDL ELMKCDKDIYKSAEPL YRERRALLEIPDFWA TIFTNLFKLGFVDEL MCVDDFFVEDTETEF RLFIDFRENDIISNKQ VIITVTPTTSEELSQE VKISTTPIELKQNKET EKKNKDSKKSTDDED EEDDDDEISGFLKLL NPTKDIATFIVGSVW SNPIDSFNCDDDEDY DGLIASDSEGEEDK
NP_650438.2	DmSET	D. melanogas	NAP family	MSSVPKRAKLDGAP ADGNTSAAAGNNEE ESEALEQIDACQNEID ALNEKASEILKVEQK YNKLRKPCYKRSELV KRIPNFWVTSFINHP QVSGILDEEEECLEHA LNKLEVEEFEDIKSGY RINFHFDENPYFENK VLTKEFHLSAAASE NGDWPASTSTPIKW KEGKNLLKLLTKPYG NKKRRNSEYKTFDFW FSDNTDPVNDEIAELI KDDLWPNPLQYYLVP DIEVEPEEEDNEDN DEEAFDEDEDGEDGE GEEEEDEDDK
NP_001246485.1	DmNAP1	D. melanogas	NAP family	MDAPAEGHVDPESC NEIEDEKSGSDCQSM PAYMNSVMRRQYLQ QMVKMLPAPVQNRI VFLKNLQQLHNLIEA QFFEDVYKLEQYQV QYQPLFDKRREIEGK VDPAEKPKQWKEPES STDNEADAEHFREAL SSLKSIPKDAKGIPGF WLTVFRNTAIMSEM VQPHDEPAIRKLIDISI KYDNGHSYTLFHFHD KNEYFSNSVLTQYVL KSTVDPNDPFAFEGP EYKCTGCTINWEKK MNLTKTIRKKQKHK ERGAVRTIVKQVPTD SFFNFFSPPEVPSDQE EVDSDSQILATDFEI GHFLRARIIPKAVLYY TGDIVDEDEDEDEEE YDENEEDEYDDDDAP PPKGPKSAGIKKQSP NDPCPNQ
NP_058661.1	MmNCKP1	M. musculus	NAP family	KLTLNDRGVGMLTR LYNIKKACGDPKAKPS YLIDKNLESVAVKFIVRK FPAVETRNNNQQLA QLQKEKSEILKNLALY YFTFVDVMEFKDHVC DLLNTIDVCQVFFDIT VNFDLTKNYLDLTVTY TTLMILLSRIERKAIIG LYNYAHMETHGASD REYPRLGQMIVDYEN PLKKMMEEFVPHSKS LSDALISLQMVYPRR NLSADQWRNAQLLS LISAPSTMNPAQSD TMPCEYLSLDAEK WIIFGFILCHGMLNTE ATALNLWKLALQSSS CLSFRDEVFHIHKAA EDLFVNIRGYNKRIND IRECKEAAVSHAGSM HRERRKFLRSALKELA TVLSQDPGLGPKALF VFMALSFARDEIWL RHADNMPPKSADDF IDKHIAELIFYMEELRA HVRKYGPVMQRYV QYLSGFDVAVVLNVL

				LKRNDASPIRRGPT EEAWNADDTGTRST DSPMWRRVSMSTN RSQPNGKNVQKTRQ IEWVIHLQNVAQGL MTKLHLHQILGPPE PGSNMQYPESFWKT GIIPDMPKLCQHVAR KFPEHPAKIQLDKVD KAGVDLLHEHAGRYV ATLEPWIMVLEDLMT FREQLRVILDLSSTVV TLMPPNQNPILLQVF MDLFCCFVRVNLLAD KVPRKMLQMHNLV HTIMKSGRDYGAYHR MVQFVEAYDPPLKGL HDDLNFVSPRIGEILD AVGPTVLLGSDFQRL RSEGYSFHPRYPEK LTNSAHPARAQDLV NIEAYREWALIGYLC PTELLRPGAVDIALVL LKESLLPLFRDEYVLL HEEYQQVLPRIAESK KLAKAGRAKGGDAD VEYNLAKQVEKQICD AQDTAINNADVHR
XP_024361543.1	PpaNAP1	P. patens	NAP family	
				MSSETAAKHDHQC EDHHDHEIDIDQQ IEESIKELALGEKIRAV AINHHLAEKKALDEEL QKKIQALTFEYEQKS MPIYAKTQEIIVQGRV PTEELKDLEKYLKEEE KTQVEGAKKEEPIEY WLKAMKSNVDLAM EIKEQDETALKTLTKIE YVLEDTKKFHIVFTFG PNDYFTNTELKKTVEL DENEKPVKTGTPIE WKEGNTTVKITKKT QKNKKTGVKRVVERE TKIESFFNFSDSAPT DAPEEEDKEDEQA MDGDRINIDFDIARSL IDEVIPYSLEYFLGIKT GGEDHDEIDMDDL DEEEIAQLQKEYAAKT GKKGPAAGGANDKK DCKQQ
XP_001022132.2	TtNAP1L1	T. thermophi	NAP family	
				MRSKKNIVEEDYISFH KDLSNKMSEKILHFEP KFRIKLCAMLYHHKQ KQNLDKDLHNELLKL KRNYEKKTYEIIKQM NSLIKEGTIILNDFSDI SDHFTMPGDPNYLEP LAYPKIQKYWSSCIQ NCTILQHQFTKDDQD IAQCLETIESREETKTT KSFTLSFQFSQDQKYF TNSFVQVTFYLDISVN AYRQESNAEFKIKTG NKETATTGIFLQLFQN VNLEEGLNKISDAHD LCTKLGRSYEIALFKD ELIPYSLEYLNAARTPN LKQEMLEKQLMEQY QEKKKIKSFVQQQQ EEADEELADEQYLED NQCKQQ
XP_001024488.1	TtNAP1L2	T. thermophi	NAP family	

				MDEELKAVITIKDEYQ LVKQKMEQQIQQL QNEGYQLIHLYLK DDLEQENKIYQSKVK EILEKYTLGQQIFQS VNQIFLKNKNYFKFIN AIQQQDIISGRKFEQ FSKEETADLDYFNIT NHNDQRAKNERFAS HEVLPISNYWFMAFK NFDLVKPYVKEKDDQ ILKHLHNIFVNIQDKS SIELKFEFKENEFQ QYLVKVFFLNDNMEP ERSEGTQJEWYEGKD VCRKTVKKSQKNRKT GAIRVITHEVEDEEFF NFFNSLNLEEIKLMG QKKNSLIQQKNYID RMNIDYDIARAIIDEL LPYSLEYLDVRRLLSK QAEVQDLNFLKEINE SVSDSDSDGGDNQ KNDNSPTKQSQNNK GKKDIKDTSKDKSGV QDNTNECKQY
XP_012655909.1	TtNAP1L3	T. thermophi	NAP family	KLTLNDRGVGMLTR LYNIKKACGDPKAKPS YLIDKNLESVAKFIVRK FPAVETRNNNQQLA QLQKEKSEILKNLALY YFTFVDVMEFKDHVC ELLNTIDVCQVLDIT VNFDLTKNYLDLIVTY TTLMIMLSRIEERKAI GLYNYSHEMTHGAS DREYPRLGQMIVDYE NPLKMMEEFVPHS KSLSDALISLQMVYPR RNLSAEQWRNAQLL SLISAPSTMLNPAQS DTMPCEYSLDTMEK WIIIGFILCHGILNTD ATSLNLWKLALQSSS CICLYRDEVLFHFKAA EDLFVNIRGYNKRIND IRECKENAVSHAGAT HRERRKFLRSALKELA TVLSQPGLLGPKALF VFMALSFARDEIILL RHADNIPKIIADDFM DKHIAELIFYMEELRA HVRKYGPVMQRYV QYLSGFDVVLNELV
XP_012655909.1	XINCKP1	X. laevis	NAP family	MCTDRSKFYEEVHEL ERKYAALYQPLFDRS DIINATYEPTEECEW KVEEEDISGDLKEKAK LEEKKDEEKEDPKGI PEFWLTVFKNVDLLS DMLQEHDEPILKHLK DIKVKFSDAQPMSE TLEFYFEPNEFFTNEV LTKTYKMRSEPDSD PFSFDGPEIMGCTGC LIDWKKGKNVLTCTIK KKQKHKGRGTVRTVT KTVPNDSFFNFTTPE VPENGELDDAEAIL TADFEIGHFLRERIIPR SVLYFTGEAIEDDDDD YDEEGEEADDEEGEE EADEDNDPDYEPKKG QNPAECKQ
NP_001087969.1	XINAP1L1A	X. laevis	NAP family	

				MANIDNKEQTELDQ QDMEDVEDVEEET GEEANSKARQLTAQ MMQNPQVLAALQE RLDDLVTPTGYIESL PKVVKRRVNALKNLQ VKCAQIEAKFYEEVHE LERKYAALYQPFKEKR SDIINASYEPTEECE WKVDEEDDIAEDLKE KAKLEEEKKDEEKDP KGIPEFWLTVFKNVD LLSDMVQEHDPEILK HLKDIKVFSEAGQP MNFMLEHFHFEPEFF TNELLTKTYKMRSEP DESDPFSFDGPEIMG CTGCLIDWKKGNVT LKTIKKKQKHKGRT VRTVTKTVPNDSFFN FFSPPEVPENGELDD DAEAILTADFEIGHFL RERIIPRSVLVFTGEAI EDDDDDYDEEGEEAD DEEGEEADEHDHDPD FDPKKAQNPAECKQ Q
NP_001080547.1	XINAP1L1B	X. laevis	NAP family	
				MANIDKTKKKNAALS KEQTELDQQDMEDV EDIEEEEAGEDANSK ARQLTAQMMQNPQ VLAALQERLDDLVT PTGYIESLPKVVKRRV NALKNLQVKCAQIEA KFYEEVHELERYAAL YQPLFKRSIIINATY EPTEECEWKVDEEE DISGDLKDKAKLEEK KDEEKEDPKGIPEFW LTVFKNVDLLSDMVQ EHDEPILKHLKDIKVK FSEAGQPMFTLEFH FEPNDDFTNEVLTKTY KMRSEPDESDFPSFD GPEIMGCTGCLIDWK KGKNVTLKTIKKKQK HKGRGTVRTVTKTVP NDSFFNFSPPEVPE NGELDDDAEAILTAD FEIGHFLRERIIPRSVL YFTGEAIEDDDDDYD EEGEEADDEEGEEA DEDNDPDYEPKKDQ NPAECKQQ
NP_001001249.1	XINAP1L1	X. laevis	NAP family	
				MSRWSEYLSVEEPIPS AMATWRNMSADAP QGSSAGSHKHLQME PVVQLSKVAEGLLAK MYRLNSILDYDPNPT HTFSDAFWKAGVFP NFPKICITMSKKFPEH PNKLQLERVDFGLD ALNENAEGYMHNLE QWILLLLDLAFREQA LRLLDLSSTVITLPH QNSLILHAFMDLFCF VRVNLFSDKMPRKMI LQVYNILHIMLKGGR DCEFYHRLVQFVDSY DPPIKGLHEDLNFVSP RIGEVLEAVGPIIFLST DTKKLRNEGFLSPFHP RYPDILTNSAHPMRA QDLANVTSYREWVLL GYLVCPDELLRVTSID VAMVVLKENLILPLFR DEVILLHENYQLYVLP KVLESKRMAKSGRTK QKEADLEYNVAKQVE RMLTEVHEQALVSCD AMHRERRILLKQEI G RMVLFTDQPSLLAP
NP_001335312.1	ZmNAP1	Z. mays	NAP family	

				EDDDGEPVHGDPAE HDENDDDEDDDDVG NEYENDGFIVNDEDE EEEEEEDEERKDSDEE RQKKKKRKKKDEGL DEDDYLLQDNNVKF KKRQYKRLKKAQREQ GNGQGESSDDEFDS RGGTRRSAEDKIKDR LFDDVDVDDPPDDV GDEEDLVVEEDVVG EDEMAFIVDEDEH GPPKRGNSKKKKYRQ GSDITAMRDANEIFG DVDELLTIRKKGLASN QRMERRLEDEFPTV LSEKYMGTGNDDEIRQ LDIPERMQISEESTGS PPVDEISIEEESNWIY AQLASQLRESOGTFD GRGFSVKNKDDIAKFL ELHHVQKLEIPFIAMY RKEQCRSLDTGDFD GANQGGKPKETKWHK VFWMIHDLKKWLL LRKRKMALHGYTKR YEESSRRVYDETRNLN NQYLFESVIKSLKVAE
NP_001321759.1	AtSPT6L	A thaliana	SPT6	EEVDRIVQDNEATNV DGGDDDEEEEEEGP DEYEKDGFIQDVED EQEEEEERKDSDE DLKAKKKRKRKRETE NYLEDDEYELLQEA NVTGFRPKPGSKKF KRLKAGRDTQDEER TRFSDDEEIEEGHSG RTAEKLRSLFGDDE GVPPEDISEDEQQEG EDDEDIGDEMRDF IVDEEVDENGLPVR RRKPKKKVLRQAPGV SSSALQEAHEIFGDVD ELLFRKLGHGKAGD GFDGVDSDVGKPK RLEDEFEPSEEEKYM TEKDRIREVDRPERI QLSEETGPVPTDEESI TKEAEWYISQLVGS TGSPVGFQHIMRLD KDHITAEIANVIGML HVQKDFIPFIAMRYK ELCLDLRDPPEEVLS NEERDKSDKAPPLRW HKVLWAIHNFDRKW LLLQRRKKALAVSYK
XP_020528358.1	AtrSPT6	A. trichopoda	SPT6	GHSDDDEEPQSKMK MAKEKSKRKKMVA SSDEDEDDDDDEEN RKEMQGFIADDDE EEDAKSEKSESRHSG EDELDEDLINENY DIRETKQNRVQLGD SSDEDEPIRRPNHED DLLSERGSDDGDRR KDRGRGDRGGYGE SERSEDFIEDGDGA PRRHRKRHRGDENLP EGAEDDARDVFGVE DFNLDEFYDDDDGED GLEDEEEIEEDGEG GEIKIRRKDTTKSTL LESIEPSEIDRGFLPG DKKIATEDLPERFQLR RTPVTEADDELESE ALWIIKYAFEETVTN QADLDQDDKLDLIM NLDPSVYEDRKKAVIK SIKKVLQFIRVRSNSFE PTFIGYRKEDIDNLLT MNNLWRVYDFDEK WCHLSEKKNKYDLM RRMREYQELSDDLTA KRRPISDADLMDTKY
NP_497969.1	CeSPT6	C. elegans	SPT6	

				NIHPKKKFKNLNKNK DDESELEEYHDSSEE DENEPNQYQYDDFV VGGDDDDERRRKKKKE KRKKSSNSDNEDDN EEEEEEDRGGGGSG SGGSGSGGRNSSP KRTNNHLDEDIDGDE DGDDGSDGSESESES EEDEEVAQARRERRR KRREERRRREKGGLO RLKLRASNKDLSKE DDVEDGYNGEQEDD GFVVDHSGKPKRSH ERERMSYESSNRSRI MSDIFDDDEEGYGS GEEEEDSGINKTEAD RAKDRLAIREQEYPS LLEEKHFTDADEEIRN KNVPERLQSRKGTQY AGEMACEEEAEWIYE VAFEQRDFQQQQQR NPNTPEVRDAKAITAI FKILQFIQRDLLEIPFIY TYEKDIEVPYFTLQEL WNIFDLDEKWAHM KVNKRNLQIGSNNQ QLEPYKSVLLESSEES
XP_645136.1	DdSPT6	D. discoideu	SPT6	ELDVNERKRLKLLKA AVSDSSEEEEDDEERL REELKDLDNDPIEED DSGGYSDGVSOGK KRKKHEDDDLDRLE DDDYDLIEENLGKVK ERRKRFRRLRIHDNE SDGEEQHVEGLVRE QIAEQLFDENDESIGH RSERSHREADDYDDV DTESDADDFIVDDNG RPIAEKKKKRRPIFTD ASLQEGQDIFGVDFD YDDFSKYEEDDYEDD SEGDEYEDLGVGDD TRVKKKALKKKVVK KTIFDIYEPSELKRGHF TDMDNIEIRKTDIPER MQLREVPTVPVPEG DELLEAEWIKYAF KHTVSEKQEPESREK MRKPPTTVNKKIKQL EFIRNQLEVPPIAFY RKEYVKPELNDDLW KVVYDGIWCQLNER KRKLKLVFEKMRQFQ LDTLCADTDQVPVDD VRLLDSDFERLADVQ
NP_001284936.1	DmSPT6	D. melanog	SPT6	EEKDLKPKKTQRFME EEEEEEEENTEDQDE HGNLRGLDDDDVEE EEEEEGEPPAGEDS DSGEEVRHRRKRFS DDYLDLDDDLIEENL GVKVRRKKKYSRVK TMDDEGDDDEKDL IADEIFTGDDGEGE VEDGEAVDTLHPRDD EEEEEDESDIDDIV DDDGQPIKKKGKKF SGYDAALQEAQEIF GGDFFAEFDTEAYD HAEIEEEQDDSW DRPKKQTKRRVSRRSI FEIYEPSESSHMTD QDNEIRSTDMPERFQ LRAIPVKPAEDELEE EAEWIYRNAFSTPTIS MQESTDYLDRTTTN FSRKGPSTIAKKEALN FMRNQHFVPPPIAFY RKEYVEPELNNDLW KVVQWDEKWTQLK TRKQNLTRLFQRMQ SYQFEQISADPKPLA DSTRPLDTADMERLK
NP_660094.1	DrSPT6	D. rerio	SPT6	

NP_001307684.1	HsSPT6	H. sapiens	SPT6	NDEGEVVRVTKKFFV EEEEDEEEEEENLD DQDEQGNLKGFINDD DDEDEGEDEGSDSDS GSEDDVGHKKRRKRT SFDDRLEDDDFDLIEE NLGVKVKRGQKYRR VKKMSDDEDDDEEEY GKEEHEKEAIAEEIFQ DGEGEGQEAAMEAP MAPPEEEEEDEESD IDDFIVDDDGQPLKK PKWRKKLPGYTDAAAL QEAQEIFGVDFDYDE FEKYNEYDEEEEEYE YEDDEAEGEIRVRPKK TTKKRVSRRSIFEMYE PSELESSHLTDQDNEI RATDLPERFQLRSIPV KGAEDDELEEEADWI YRNAFATPTISLQESC DYLDRGQPASSFSRK GPSTIQKIKEALGFMR NQHFVFPFIAFYRKEY VEPELHINDLWRVWV QWDEKWTQLRIRKE NLTRLFEKMQAYQYE QISADPKPLADGIR
NP_033323.2	MmSPT6	M. musculus	SPT6	NHEGEVVRVTKKFFV EEEEDEEEEEENLD DQDERGNLKDFINDD DDEEEGEDEGSDSG DSEDDVGHKKRRKPS FDDRLEDDDFDLIEEN LGVKVKRGQKYRRVK KMSDDEDDDEEEYG KEEHEKEAIAAGEIFQD EEGEEGQEAVEAPM APPDEEEEDDEESDID DFIVDDDGQPLKKPK WRKKLPGYTDAAALQE AQEIFGVDFDYDEFEK YNEYDELEEDYED DETEGEIRVRPKKTTK KRVSRRSIFEMYEPSE LESSHLTDQDNEIRAT DLPERFQLRSIPVAAA EDDELEEEADWIYRN AFATPTISLQDSCDYL DRGQPTSSFSRKGPS TVQKIKEALGFMRNQ HFEVFPFIAFYRKEYVE PELHINDLWRVWQ WDEKWTQLRIRKEN LTRLFEKMQAYQYEQ ISADPKPLADGIRAL
XP_024376317.1	PpaSPT6	P. patens	SPT6	GQDEYEKDGFIWDEE EEEEEEPEEPQESSD EDSKAKKKKKRGE ETYELDEDDYDLQEA NVTGFHRPPKADGKT KFKRLKKGGIGEGEK SGGIGLSDDDEEDG GGRRGRRTAEELKRS LFGEDDGAAEDVVE EEELDEEEVDEDEM ADFIVDEEEEVDEH GEPKHKRKKKKVKR KDFRQAPGINSIALQE AQEIFGVDVTSLEERR RRERDTLAGVDDED VDDDDDDGYGRSRK PPSQKLEKQFEPSSL EEKFLTERDDRLRETD VPERLQILEEAVGALP TSDSAIRQAAEWIYD RAFGLAVPVRPEFR HLSRDRDLISRQIANV LHLIHDDKFEIPFIAM YRREECFDLLKEASDD YEDEERPLVRTYAAL WAVQQWDKKWLLL QRRKTALQAAEKRI PSDIRDDPEKDELVDK

				VDEENELEDETVGVE DRSFFKYGHLQKRRR RKVQSKSTVFQKGAQ WYKGSRPSVGSRRK RDRNSSAVINEESDD SDENDGNDQETLAR KFLEKQFEPSPLEAHN VIARDRLRSTDIPERI QVLEEVAGPLPRSIVE LRRASEWIFDRAFGH LCLYSRSEFQHLNLM DKYEVVKQJANVLHL VHNEKLEIPFIAMYRR EECLDLLKELTDEVSH EDRRTIYRFEALWAV HEWNKKWLILQRRK SDLQDSYEMRIPIGIIA IPEKDNLFKRILSLAD AESEQAVGDCCEAQF NLHFPTVEVDMVDR GFKRPKCRSSYCAH DAGLGSITKYFGLTSE AFGDNLKAGHKMNE ILDEAFLPDDIASKFC NESGPTTQFRDPASIL RGARHMLAVEISTEP SVRECVRMFYSKNAT VNTRPTAGMEIIDA
XP_024375737.1	PpaSPT6-X4	P. patens	SPT6	VNTRPTAGMEIIDA EVSHEDRRTIYRFEAL WAVHEWNKKWLIL QRRKSDLQDSYEMRI PIGIIAIEKDNLFKRIL RSLADAESQAVGDC EAQFNLHFPTVEVD MVDRGFKRPKCRSSY SCAHDAGLGSITKYF GLTSEAFGDNLKAGH KMNEILDEAFLPDDIA SKFCNESGPTTQFRD PASILRGARHMLAVEI STEPVRECVRMFY KNATVNTRPTAGM EIIDAFHQFEVGLIK DKQISAFDDEQWLLI HKAEEKLIDVTVGVS KESIDTFLSECEPLFSS DGSTCTSQHWNEQR KQILQEATVTILLSKM EKEARMNLVTNAKR VVVAQCVLQLWNKV SIGPYDPKGSGRKC RNEEGSALRVLACCW GPASKGTTFVMLDA DGEILNLTGYLSTK VFTPEQAQRKENDQ SRLLQVLIDFKPHIVL
XP_024375742.1	PpaSPT6-X6	P. patens	SPT6	VPEGEAAPEGAEEG YQEGQAEGQEETQE GEVAEQEEGAEEEGE GESREKKKKKKKKRS AFIEDEAILSGSEVDS EEEDGDSQNEYEYDD FTVPTGMEEEGREGG GGGERRRRKKKKRR HEETLEEDDYDLVRE NTGQNIARRKQDSA GAGGGNKNKKLKRRL KQMGNRTDLYSNIQS RLFDDDDDEAGPSEA ITHEEPRSDDELDDDED EDDFIVGADGKPIKRS KQRRYRGGEDRGAL QDAEDIFGSVDFFKQ ASGRRSRPRGDEFDD DGDYVDGTQGLSEDE LDAEIEEGDEKQKDPL ADIEHYEYPSLLEQKH FTSEDEKIRQLDVPER LQMRKGTAYDEKLVP EEAEWIFTRAFARTH ATDKAIIQRIIHVLRFL QEQLLEVPFISTYRKD YYQPLKISELWDIYD WDEKWAHLQARKN
-	PpSPT6	P. polycephala	SPT6	

				EIVNDNDETKAPSEEE EGEDVFDSSSEEDID EDEDEARKVQEGFIV NDDDENEDPGTSISK KRRKHRREREEDDR LSEDDDLLMENAGV ERTKASSSSGKFKRLK RVGDEGNAAESES NVAASRQDSTSKLED FFSEDEEEESGLRNG RNNEYGRDEEDHEN RNRTADKGGILDELD DFIEDDEFSDDEDETR QRRIQEKLLREQSIK QPTQJITGLSSDKIDE MYDIFGDGHDYDWA LEIENEELENGNDNN EAEIEEEIDEETGAIKST KKKISLQDIYDLEDLKK NLMTEGDMKIRKTDI PERYQELRAGITDYG NMSSSEDQELERNWI AEKISVDKNFDANYD LTEFKEAIGNAIFITK ENLEVFFIAYRRNYIS SREKDGFLLTEDDLW DIVSLDIEFHSLVNKK DYVQRFYAELHIDDPI
NP_011631.1	ScSPT6	S. cerevisiae	SPT6	RNIGDEEEEDYQGOV EEEEEGNNQVEEDDE EEEDDDQDDYQQD GFVVGDSSEEEIEEEE DEDAEERRRRKKEKK RQKKLEKQNQHRIL KKGPKKRELSEDEME NNEIQDLEEHEDSEE GIPDRKRNRQQRSTG NDGHQGDKMDLED QGNEDYLEDVGDYG AKYGKTAFEELFMD ESGEEEQNENEDDEN REVDDYIDVQQLFEP DDLKKRFERDDDKKIK EEDIPERLQIRMQGR KAPDPEELMEETKWI SEKIKTIKDYKGSKLLN DINFHKKIFSFLQSLHL DKEEVMYIYTYKNEF HPQFDLEDLWRLYDL DGEWAQFNQKNRI LLQIQQLRRELENTPL LIQNNIIDFEQVRKVE EFFHKAIDTQSLKFIKE YFDYLFKIYPQKEHQ RLKIRKDRKTRMINTY IQCKVHKLVTSLTLP
XP_001030352.2	TtSPT6	T. thermophilus	SPT6	WAIAAAGSDVTPACS AFYPPASALDLLHF WTKLFGKEYGFRRRQ LGSALFTSERDSPRQ SRERLCSLIAMDFVES EAESEGEYEDGNNTA ERSNKKFVEDDDEEE EEENPEDQDELGNLK GFINDDDDEDEEEEE DEGGSGSDEDVGH KKRKH5FDDRLEDDDD FDLIEENLGKVKRQK FRRVRKMSDDEDEEE EDFGKEEHEKEAIAEE IFQDGEDES RPDHVD QPVAGPEEEEEEEDE ESDIDDFIVDDGQP LKKPKWRKKLPGYTD AALQEAQEIFGVDFD YDEFKYDEDDDEEME EYDYEDDEGDAETRV RPKKTAKKRVSRRSIF EMYEPSELESSHLTDQ DNEIRTTMPERFQL RTISVKSAEDDELEEE AEWIYRNAFATPTISL QETNDYLERGQVGS NFNRRKGPSTIQKIREA
XP_041439251.1	XISPT6	X. laevis	SPT6	

				ADEEDPRPSRRNRDD MDERDDDDDDDED GQDEFEKDDFIVDDE EEEEVEEEEEQKSDD ERRRKKKKKKRESDD FMLDEDDYMLLQDN NITGISRPKPGNKFKR LKKAGRESEMDEHSG FSDGDGTGKKRSGKE RVEYSLFGDHDQAAP FEEDFEDDQQAAGEGE EEVGDEDEDEMADFIV EEDEIDGNGQVVRK KFKKKVPRQAAGVSS SALQEAQDIFGDVDE LLALRKQELERREALNS GELRGARLEDEFEPFI LAEKYMTSKDEQIKE NDVPERMQLSEELTG YPPTDKTTIEESSWI HSHLTGDGFLSFFGN EHINKDIDQKDIVNV VTMLHANKFEIPFIA MYRKESCPTLLSFDS EDNEDIESDARKLRW HKLLWAVQALDRKW LLLQKRKVALQLYK RFDDEKRRIDDVTRQ
PWZ19394.1	ZmSPT6	Z. mays	SPT6	
NA		A. thaliana	APLF	
NA		A. trichopoda	APLF	
NA		C. elegans	APLF	
NA		D. discoideum	APLF	
NA		D. melanogaster	APLF	
				VKVDASPGSEIVLINS ECTVGRKKDCDLSFP ANKLVSGNHCKITHD QNSGKVWLEDNSTN GTVINMSKVVKQT HLLQNGDVIVFVYRK NEPEQNIAYVYQSITP QESASHDVEDAGREE DSDLTETESAPVPEP VIVKPLPQSGHEDPQ PSTSSSLHFYNNPLS TCSDVSARKNPVSS AVCKGDSSTSGSPAQ TRLKWTCWTDGEPE EEMQKRRTDRDD PGFGSAHSDASADIP LRGASGKEKTEGATT DKMEESLTCICQDLL YDCISVQPCMHTEFCA ACYSGWMERSFCP TCRCPPERIRKNHILN NLVEAYLLQHPEKCR TEDDLRSMADARNKIT QDMLQPKVERFSFDE EASSDYLFELSDNDS ISDMSQPYMMCRQC PGYRKELSSALWICES AQSESIAKTAGDGPS
NP_001093485.1	DrAPLF	D. rerio	APLF	
				RVALAPGETVIGRGP LLGITDKRVSRRHAILE VAGGQLRIKPIHTNP CFYQSSEKSQLPLKP NLWCYLNPGDSFSL VDKYIFRILSIPSEVEM QCTLRNSQVLDEDNI LNETPKSPVINLPHET TGASQLEGSTEIAKTQ MTPTNSVSFLGENRD CNKQQPILAERKRILP TWMLEHLSQNLNLS VPAISGGNVIQSGGK EEICKDKSQLNTTQ GRRQLISSGSENTSA EQDTGEECKNTDQEE STISSKEMPQSFAITL SNTEMNNIKTNAQR NKLPIELGKVSXHKI ATKRTPHKEDEAMSC SENCSSAQGDSLQDE SQGSHSESSNPSNP ETLHAKATDSVLQGS EGNKVKRTSCMYGA NCYRKNPVHFQHF HPGDSQYGGVQJVG QDETDDRPECYPGPS CYRKNPQHKEYRHN
NP_775816.1	HsAPLF	H. sapiens	APLF	

				VPVGPQGTIVIGRPL LGITDKRVSRRHAILE VVDSQLRIKPIHRNPC FYQSSEKQHSMPMET QVWSQLHPGDSFSL LDKYAFRVFSAESEVE MECTLRNSQMLDED DILSEMQKSPVVNLP DKTTGASQLQGSPFIT KTKCPTIDPMSSSGEC RAFSEHQRPRTQRKR ILPAWMLAESLSDQS LSTPAEGGDKDVIQR SGKAGTCEDRTPGNT SWHGKKRLSPSGNSK SVSAEQDPGKKCRKA DQEGPGVSSENVPE SSSNIVKDPDVIDVT NKQKGILIEELGEVS KHKAATKPTTNEEGE SCARVQSKSPPEKSQ GCHPESSAPSSPDAL HTDTADPVLGCSEES KVRRTACMYGANCY RRNPLHFQHFSPHG DSDYGEVHGTDEGVI GDRPECPYGASCYRK NPQHKMEYRHSALP
NP_001163960.1	MmAPLF	M. musculus	APLF	
NA		P. patens	APLF	
				AQNQEKKEKEKKK GGSVGEGGARWVLE CSEEGDIKVELMGGP TELGRGLVNNKISRK QCVVETVDDGVELTP LGVNPMTMSRAGSE DVEILSRNEKYTLHEG DSFTLCGEQYRFVLQ RPTVTQTTSQTTSQ KDGEKSAKPASRQAS KIEGAEGEGEKTRD GEADINEGDKGEEE GREKQRGEKGEEGK GEEGSPKKGVKRKA SDIATTTTTSSAKKTP QGSAESDSATVPFSA GYDSPEESKAKSKKKP DVKTEEADDVIDDEEL ARKLQEELNKEDEAP KPKTPKSKNSTPEKD AEIARKLQEKYNKESG GGGKKRKQEDDED DGEVNWAAKVKKQK KGKKGGSDDEDEDW NDVGGGGDESDDED YQEEESDSEKNKKNK KKGASPKKPPCKYVK SCYRTNPDHLRDFSH
-	PpAPLF	P. polycephala	APLF	
NA		S. cerevisiae	APLF	
				KGGLVQETFVLNQKE HILGRRGDLKVDNPK VSGQHCVLKYDYAQK KAYIIDVSSNGTSLFN KKLEKNKEVELENGDL VNLLQDKSQWIGYIF KLVDNVDSFNKDQQ DTATKNNTDQKQLE QSLEQYKQNEKEQQE QNEQIKKMQNELEER LKKVKEDDEHFEKDQ TCVVCIDLYNPYLMIT PCLHNYCCDCMCELL KNKDIACPQCCEKPI VQKNYQLNNLIEAFIK RNPDKKWQEDVIKK KNESNLLNKDFLDQI NEKLSNKTKKVGIYKR QRSYSDSESSDRNNN RNYSEDEEEFDEDEE EEEYQYNNNQPFQ FGYNYNFIMPAHLNR CVECQNARQDDNFK CSPYQQHIKCINCDK MMPQRNDNNLYPQ NCTICERSFCNLYFKQ DCSSKYNSSSHHLRL FQDIPIPQVLNNNVLS
XP_001009833.2	TtAPLF	T. thermophilus	APLF	

				RLLGAETDSSSEFLYK KEWTIGRKKACDLSF PGNKLVSGEHKITV NEESGEVSLDSTN GTVINKLVIRKQTY LKNQDVIVVYRKNE PEQNVAYLYKSLNQ QDSLHDPADTSGSEE AETQLSSQDDQLSY EEPQPTSTSSLFSTPT TSAIPGVQLESAEKS ESLGGHSSTSDASPA RASIPKSNLSTQEQGS LGPPKKRIRTEDHWT TNKNFVPASCPIGAS DESKTPSMKPKDMEE TLTICQELLHDCVSL QPCMHTFCAACYSG WMERSSLCPTCRCPV ERICKNHILNNLVEAY LIQHPEKCRSEEDRCS MDARNKITQDMLQP KVRRSFSDEEGSSEDL LELSDVDSESDISQP YTVCRQCPGFVRHS MQPPPPPPPSDTETS RTQGDAPSTSTNFPT ATQEYVCPSHGSHVI
NP_001089240.1	XIAPLF	X. laevis	APLF	
NA		Z. mays	APLF	
				DAEQKASNIQVQDLY ATWQTEPGEEKASILL ELPAETTLHTIEIVNA GSTNVEIYVTNSENQ QDEIILLPTAIMMTL GEVNSNTNTNRLRFF GPDKLTGKISDKAWK YIKIVCISFDTSRKPG LARIRLFSAPSDEYIA SSIADNRSSMVSVPS QPTTSTYVPQKYASR SKSFAGPPPTTIAHD TNNKTPNHTNTAN KSNNNNRTNDTNNS PAKSAKRKRVMND DYDDTEGDSAKDQN PRSLTSLHISPHVKKS VPVANRSKSLVFTSFN KILSGVSFVVGIANP ERQELRDKAIEMGAQ YYADWIESATHLICPF TNTPKFILAQKAGGTI VTPAWILDCYKNKVR LPETEYLLAEVKKIDD AGMKSTMGMNMRD WLATFRPKFDKPAN NSTNTTQSQPQPKLT KSLTFVNNARITKDSQ
-	PpXRCC1	P. polyceph	XRCC1	
				DSTHCAENLLKADTY RKWRAAKAGEKTISV VLQLEKEEQIHSVDIG NDGSFVVEVLVGSSA GGAGEQDYEVLLVTS SFMSPSERSGSGNPN RVRMFGPDKLVRAA AEKRWDRVKIVCSQP YSKDSFGLSFVRFHS PPDKDEAEAPSQKVT VTKLGQFRVKEEDES ANSLRPGALFFSRINK TSPVTASDPAGPSYA AATLQASSAASSASP VSRAIGSTSKPQESPK GKRKLDLNQEEKTP SKPPAQLSPSVKRPK LPAPTRTPATAPVPA RAQGAVTGKPRGEG TEPRRPRAGPEELGKI LQGVVVVLSGFQNP RSELRDKALELGA PDWTRDSTHLICAF NTPKYSQVLGLGGRI VRKEWVLDCHRMRR RLPSQRYLMAGPGSS SEEDASHSGSGDE APKLQKQPKTKTKP
NP_006288.2	HsXRCC1	H. sapiens	XRCC1	

XP_041427601.1	XIXRCC1	X. laevis	XRCC1	DTHTTAENLLKADTY RKWKAARPGEKQSV ILQFEKEEQIHSIDIGN EGSAFVEVLVGRSTSI SEQEVEVLGMSSFM SPSESKNESNLNRTR MFGPDKLVKGAAEK KWDRVKIMCTQPYT KNLAYGISFIRLNSPPE DGSPTAPSPKVTKLG QFMVKEEENSPPSM RPGSLFFNRTSKPQV TPPKTPPATQSYAAA ALQGTAEISSSTEKQI KTPPSNNITKEPSSSK RKFEFNKEPSSHSYAVK KPDVKESPSSTKAKTS SQPAPKKPKVETPPT KKPSPSEKPNQKKSP PSSQPMDLGRILQGT VFVLSGFQNPFRSDL RDKALEMGAKYRPD WTPDSTHLICAFANT PKFSQVKAAGGIIVRK EWWLDCYKRRQLP QKQYLLGTAESSDEE DDSEEDPPKAPQHK PHPDPSRTNHNKKPSP
NP_001003988.2	DrXRCC1	D. rerio	XRCC1	DNTHKADNLLSDTY RKWKAARPGEKQVS VILQLEKEEQIHSIDIG NEGSAFIEVLVGHSTS VKDQDFEVLVTSFF MSPTESRNGTNTNR VRFFGPGQLVKAQS QEKWDRVKIVCSQPY SKTIAYGISFIKLHSP DGNEALVSTPPKLTKL GQFRVKDESPSAGSN VQPGALFFSRDTSK SSTGLKVSPQNERLSY AAAALKTEGSSSTHTP ASDTQQAPVKRKEFE SKERLEASGPPAKKPS SSASAETTPKAKARSS MGTSPPSRASPAQKS SDKRESPKTPESKPK PKPKPKSSESVPFNRI LEGVVVLSGFQNP RADLRDKALALGARY RPDWTPDATHLCAF ANTPKYSQVKAAGGI IVRKEWVMDCHKNK QKISCKRYLMDGAES SSESEAEQDESEETP KAKTPEKKPVTPKKQ
NA		T. thermophi	XRCC1	
NP_001326839.1	AtPIE1	A. thaliana	SWR1	ASKSGKSPDNESRA KRQKTLEAPKEPRRP KTHWDHVLEEMAW LSKDFESERKWKLAQ AKKVALRASKGMLD QASREERKLEEEQRL RKVALNISKDMKKF WMKVLYKHQLVRNE KKKKAMDQLEFLLG QTERYSTMLAENLVE PYKQGQNTPSKPLTI ESKSDEERAQIPPEI NSSAGLESGPELDED YDLKSEDETEDEDTI EEDEKHFTKRERQEEL EALQNEVDLPVEELLR RYTSGRVSRVTSVKD ENEDNLTSVSRVTSP VKDENQDNLASVGQ DHGEDKNLAASEET EGNPSVRRSNDYGH LAISETHSHDLEPGM TTASVSKRKEDHTYDF NDEQEDVDFVLANG EEKDDEATLAVEEELA KADNEDHVEEIALLO KESEMPIEVLARYKE DFGGKDISEDESSSF

XP_020520860.1	AtrPIE1	A. trichopoda	SWR1	KDDETTLAEEEEFAKS EANDPSDEIKLLQQES EMPLEELLAMYKMD TCKDEDDLSASECAST SSEEQVDHSGNKEVK GEDSGPDEDRDGLTV LPDSTEIESFSPLKHS VMQKLKGLKLENRTTE GRESENIADAAAAA RSAQPTGNTFLTTKV RTKFPLLKHPLREYQ HIGLDWLVTMYEKRL NGILADEMGLGKTIM TISLLAHLACEKGIWG PHLUVPTSVMLNWE TEFLKWCAPFKILTYF GTAKERKNKRQGWL KPNSEFHVCTTYRLVI QDAKVFKRKKWKYLI LDEAHLKNNWKSQR WQTLNLFNSKRRILLT GTPQLNDLMELWSL MHFLMPHVQSHQE FKDWFSNPISGMVE GQDRVNKEVVDRLH NVLRPFILRLKRDVE KQLPKKYEHVYICRLS RRQRNLYEDFIASSET RTSSRSVADDQPSTSS AVAPPPSPAIETDED AVVEEKKKKKTSDDL EITPRTPVDRRIPYICS ILLTENRSIRKLVLS GPVRQEDHEEQIARA QRIQPVDQIQRVEQ IILNGSVEDILKDPFPA VMADLTKEPPPTAP PPPIQKTMQPIEVKIE DSEGSNTAQPSVLP5 CGGGETNVERAAKRE AHVLARIAELRKNGL WSNSRLPKCVEPERN KTHWDYLLLEVVKWM AVDFRTETNTRKRIAK VIAHAIAKQHRDKQI EIERAAEREIKEKRKM CAGIAKMVRDFWSS TDKVVDIRAKEVLESR LRKARNKHLMFVIGQ VDEMSNIVQEGLVSS SKSPSIASDRDDKDEE FKAPGSDESDEQTI ANAEKSQKKEDVRQE VDALQNEATVDMDD FLYTLPEYLKAYGLT QEDLEEMKREKLEEQ TDESMNIDDNGGDQ IHNAISPTLSEDTVTS NTNPTTTTTPRQTRS SLSLSPSIDNTIALAT ATNSSKKRSRSNSTSK NKSNTGSSSSTNPSP MGSPTLSTSTINLLNT TTTTTTSTIKIELNL QKNKDEFIDINNKRNI LIKELLQQQDNNN NNNNNNNNNNNN NNNNNNNNNNND DIKQQQLSLLLDQ DKLQLINGKVEEFCNS NGFSFYQIFIDNLSSL TSKSLQPTTSSKIPTI PVEKPIIEKPTLPERPK SERLKRQQSLLNSTDS KKLSTSTTTTTTTPT TTTATPPPPSSSSTS TSTSTGTPIGGINTP TNTPSTNNNNNK GICNIKDTPIPFSLAE ETLTDKSTPIKTSQQP QKKQRRESSSTQLNK QITEIQQQQQPIEEI TIPIEQQQMQLLAQV NPQDFKPIPLPPPSAL
NP_001255180.1	CeSSL	C. elegans	SWR1	
XP_647182.1	DdSWR1	D. discoideum	SWR1	

				SPAPPAVPDRVTPHS TEISVAPANSTSTTVR AAGSVGAALPATRH HQHIATQVKGIASSSS KQKQLASQQLPVPL SPLPQQQQQTAEAT AAAAAPAHSNVSVSS STIEASVLPQAKRQR LDDNEDRTSAASIVG PAESSNIVSLLPASV ASSSEVGGLSSTALQD LNALKKRILQQLQIL RNLKERHLENVSEYFY LQNGGSMMDYPAW RKKTPTPQFISYSNAN RIDQLIHEDKPSTAA AAAAAQNQKYTTQQT DSVESLVSIGTGAT KGAPLDGNISNSTVK TNTQSQVPSKIGSFTE STPAATESNSSTVPG TATSGAATSTSATSAE ASGNVLAVEAEIKIPA VGATPVAISTKLPAAV VQLTQGGGTPLLPCN TSAGSTALRRPQGQN NASSGSAASGGGG SLTPTPLYTGNGPAAL
NP_524833.2	DmDomino	D. melanogaster	SWR1	QHHGDTGGVSAPFQ TAKEPLPPETIKLDYK NVSGKATSGESPSAP KTEVTPSVKGAVKST DKDMEHPKPTPLFET NQTAASSVLQKQSTT NSATQLCKFDSKETM EDVTRDLLKTSKEDDI SETLSAGGTQKGISG NSICSGYSSKQGGQH ITPNRGNTPNDSKD GSCYLTECSQTQTS TQKTSTLPKTPGIPTL VENIQCVQELSSQYL HINPKMVTLSKETSTR GDLYCNVAAKEVLCN VSVDEPKLSGHGTCT DTAHVSIMSERSQV MCGGSSDVILKQND ANQIEISMSSNHQQ DGSILLTHAENCVTSE GLESDQPNLSDCLVQ QLQESIPDHLKLSQIQ ICLNNTPDIVNEEERQ EMNAEALSAVEVKDL IDVGILMQPDMEDT WLVVSEIDGQPEHAE MEVSDESLVQCPESEQ
XP_005156368.1	DrSRCAP	D. rerio	SWR1	LQRSKSFTGSEAEQ QQQQSQQPQQQQ QQPVPTQQQSANA NLPQSPVTSFAPSASP SAPQSPNYQIIMSR PVTGQNMNITLQNV GPMVTGNQQTILTS PLQSPASPGFQHQ QQWRFEFGSSSVIQV TSPLQTMQPSQPT QHSPVPLQAVPRPN APGSALGVCQSPTR FVEANIMVRQINLSS PPGSGHFVYQDGAG LATGTVHLTSPGTPG AVRERRLSQPHSQPG SSTVHHKALPSQVPT LGSPGPITTTNLPPQI SSIIQGQLARPTMFEK TSQSVVGVANSAAT FGVSTIPPSSPSRSN SSQGLSSQSLTPSM KKMQPKKLEEIAPSN PEVAQLRKQCLEHHG KKMEGLEVFKEYLIE LFFLQHLQGNMMDY IAFKKKPCVPLFTYLR QNDLDLEDEEEEEEQ
XP_009302483.1	Dr-p400	D. rerio	SWR1	

NP_006653.2	HsSRCAP	H. sapiens	SWR1	QTQMVSDGMTGSN PVSPASSSSPASSGAG GISPQHIAQDSSLG PPGPPDGATVPLEGF SLSQAADLANKGPK WEKSHAEIAEQAKHE AEIETRIAE LRKEGFW SLKRLPKVPEPPRPGK HWDYLCEEMQWLS ADFAQERRWKRGVA RKVVRMVIRHHEEQ RQKEERARREEQAKL RRIASTMAKDVRQF WSNVEKVVQFKQQS RLEEKRKALDLHDF IVGQTEKYSDLLSQSL NQPLTSSKAGSSPCL GSSSAASSPPPASRL DDEGDGFQEQEED EDEETIEVEEQQEG NDAEAQRREIELLRRE GELPLEELLRLPPQLL EGPSSPSQTPSSHDS DTRDGPEEGAEIEPP QVLEIKPPPSAVTQR NKQPWHPDEDEEF TANEEAEDEEDTIAA EEQLEGEVDHAMELS
NP_056224.3	Hs-p400	H. sapiens	SWR1	LQSRACPGSEGEQ PAHPNPPSPAAPFA PSASPSAPQSPSYQIQ QLMNRS PATGQNVN ITLQSVGPVVGNNQ QITLAPLPSPSPSGF QFSAQPRRFEHGSPS YIQVTPSLQQVQTQ SPTQSPSPGQALQ NVRAGAPGPGGLCS SSPTGGFVDASVLR QISLSPSSGGHVFQD GSGLTQIAQGAQVQL QHPGTPITVRERRPS QPHTQSGGTIHLGP QSPAAAGGAGLQPL ASPSHITTANLPPQIS SIHQGLVQQQVLQ GPPPLRPLGFERTPG VLLPGAGGAAGFGM TSPPPPTSPSRTAVPP GLSSLTSLVGNMG KKVPKLEEIPASPE MAQMRKQCLDYHY QEMQALKEVFKEYLIE LFFLQHFQGNMMDF LAFKKKHYPALQAYLR QNDLDIEEEEEEEEE
NP_001291195.1	MmSRCAP	M. musculus	SWR1	QTQMVSDGMTGSN PVSPASSSSPDSSGA GGISPQHIAQDSSLD GPPGPQDGTTPLEG FSLSHAADLVNRGQK WEKSHAEIAEQAKHE AEIETRIAE LRKEGFW SLKRLPKVPEPPRPGK HWDYLCEEMQWLS ADFAQERRWKRGVA RKVVRMVIRHHEEQ RQKEERARREEQAKL RRIASTMAKDVRQF WSNVEKVVQFKQQS RLEEKRKALDLHDF IVGQTEKYSDLLSQSL NQPPASSKAGSSPCL GSSSAASSPPPVSRRL DDEGDGFQEQEED DDEETIEVEEQQEGN DAETQRREIELLRHEG ELPLEELLRLPPQLLG GPFSPSQTPSHSDT QDGPENIEEESQD LEVHPSSAVTQCNC QRWHPDEDEEFT NEDEAEDEEDTIAAEE QLEGEVDHAMELSEL

				LQRRSR5FTGSEEEQPA HPNLPPSPAAPFAPS ASPSAPQSPGYQIQQ LMSRSPVAGQNVNIT LQNVGPVVGNNQOI TLAPLPLNPTSPGFQ FGAQQRFFEHGSPSY IQVTSPPMSQQVQTQ SPTQSPSPGPGQTLQN VRAGAPGPGLGICSN SPTGGFVDASVLRQ ISLSPSSGGHFVFQEA PGLTQMAQGAQVQ LQHS GAPITVRERRLS QPHAQSGGTIHHLG POSPAAGGTGLQPL ASPNHITTASLPPQIS SIQQQLIQQQQVL QGQPMNRS LGFERT PGVLLPGVGGPSAFG MTSPPPPTSPSRTTM PPGLSSVPLTSMGSS GMKKVPKKLEEIPPA SQEMAQMRKCLDY HYKEMEALKEVFKEYL IELFFLQHLQGNMM DFLAFKKHYAPLQA YLRQNDLDIEEEEE
NP_083613.2	Mm-p400	M. musculus	SWR1	GRRNRAKPEPSSVES GPPLVKDVRGLTPVK KVSPPTSDQAVDF VAQVKKPHAGRGRP RKVVGVPPKQGVKSLS DVVTKKMPATRSSA GKPVAFVADGREVA KDSISIEETVPEGSSLK TLVRSQDKVPVRVRI VGPFSSNGGSKHGRD ESFGSLSSGSQKSRG QIVRQEAQRAADDET SDGLQPNQRKSDGL AARDEKISCRREIERS IHEKKVAEEQVIDKRE EGQTVPEKKRTAHFH YGRKDAVEQVIDKRN QWLPVTERKMGGKT HHDIHETGEQVIHKR GEGRYIFEMSAGGKS QRRIDGEHSTRGKH PSASEGTLASRRRFM ALRRRQREDLLSRA DEILLKQEFLEANKDK KVQPKSKKPARPKTH WDFVIEEMTWLAKD FERERKWKLTQAKKI VLRVNRSKLDVVARE
XP_024369996.1	PpaPIE1	P. patens	SWR1	GNKRQKGVAEPPHID TRNANTKRPSRTSEQ ASPIPKRVRGSGVPE SPSVEAIDVSQSEVA PKAGVETDINIPAI YEVEKRLELIKQLTEQ KENELKELAYLEAGN NLMDYEAGFTGKI MTAKEEPPGENVAVG AEVPNKKGPGRPAK DPNSQSRRRSTVGQS PVQGESLQAQSPIPQ TPSAPQTPVQPSQPT PQTPTQPTQJTAQE AKPAVVFHQPLQVPL PSIFTVPPPTSEQDRE RALKDAAIIRRVGELQ RTGMWSLKRIPKAVE PPRAKVHWDHVLAE MGWLANDFKQERK WKMALAKKVSQV MKYHQMQUETREQR KRKEEEHQIRKIASGI AREVKKFWQIEKV MFKEQTKLDEKKKEV LDKHLDFLVGQTEKYF EMIANDLKQPEGGA QKDAPTTEQSTEEKA
-	PpSWR1	P. polycephala	SWR1	

				GGEQDLADLKFRYDL LTNELFHLREFVSLVD YDPTHFNDSSEFQKFL RETHLSLEERGEKFTD DVAKKGTNGDLTRRR RNLRTSTVVSSETTNE KKGDIELKLESIAPLVR NKCEELKYKLSDHSN RKSIVPQKRPIQHLKK REAAKSLFKSERKEN PLPLHEHIAEERYDHI AKVEEPSEAFKICPS DDSSFENTSEHYSDN FYFTTSSEEDIKKRG RKKKKPRIKLVVHPPK QTITNPLHVVKPGYES LHEYIASFKSLEDDLTL EEYNKYIDEQRLLSR LKKGIENGALKYDKET DSLQPTSKEIKTIITYK PDPISYFYKQDQLQIH TDHLINQGIHMSKLF RSSTKARIARAKKVSQ MIEQHFKHVAGAE RKAKEERHHKSLARF AVQAVKKRWNNMAE KAYRILRKDEEEQLKRI EGKQHLKMLEKSTQ YFFDCLFVSFRKAWS CVNFIYICYQIYYGAE GNLFSYTKIFRRKK QIGSQLKKIQEQVQK QMNSTEHDGQGLNS SIRNYKNNSDISYEQN STNLSKNQDSYKDSIK KILETMKTESEKQKK YDENLRKIDQWRQE MKINFEISDDEMND DSYNFDFPENQLRQ SDLANQKYNQLIEEQI DKLLQSSIIDKELDIVK QRIHNAHSSSTFLESK KAFLKNKRKNLFKNY VKTVNQNNQNAEKI QNWQNGNNNLSTG VSSTSSQVSSSNVQT LEQIJKESMKLQSSKN SKLIEHQLMRLTQAIE SPRHDDFIENHIFKKF KVTKDKLQTRTQQVI GYKPSDFKIKRYRFN PONMPFCGSNFDEA KKALVPQYHAVYND QSDPFNRLLSKPIIPRI GVNKKQLQMKPKK NVNEGQPEPTTSPL KMTGSNPVSPGSSAE LSPFDLSGEMGDTAC GSPTYDGMGDSQQD NSGGQKWDGQSEIA EQAKHEAEIENRIAE RKEGFWTPRRLSKVP EPTRPKVLWDYLCEE MQWLSADFAQERR WKRGVARKLVRMVV RHHEELKQKEERARR EEQAKLRRATTIAKE VRQFWSNVEKVQF KQQRLEEKKKALD LQLDFIVGQTEKYSDL LSQSLNETLLPVSKSS SCIGSSQGGSLRTSPT PSVHHNYDGDPLPH DEEDDEETIEIERQD GNDAESQRLEIELKK ESELPLEELLES LGPNF LQDVPDQDMQESD SPLLPQVFEEEDGEFT VNEEEGEDEETIEAE EALEGQQDHAEELSE LAKEGELSLEELLQKYS MVKTDEEIFDPSEES QQSDQLSDSSHCHENS DSDSEGVFLVKPED
NP_010621.1	ScSWR1	S. cerevisiae	SWR1	
XP_001018403.2	TtSWR1	T. thermophilus	SWR1	
XP_018092196.1	XISRCA	X. laevis	SWR1	

				PSSGNDGDQQRQK TTPPSPAQSFTPTVSS PAPQSPSYQIQHLIM SRNSMPGQNVNITL QNVGPVVPQSQQITL TSLPITSPASPGFQFN TQPRRFEHGSPSYIQ VTSVVSQQVQTQNS TQPNTPVPIQALQGV TGTSSASLGMCSQSP TRGFVDAGLLVRQLS LGPSNTGTFVIEGGS GITQIAQSAQVHLSP GTPTVPAHNLLHSSS PTVGPVHQFGSQNS VTGSVNVQSPSTPSQ LTTNLSQQLSSIIQGG LIQQQQQQVLHGG QVSRTFSDGTSSGM IAGVTGTSAFGITSAA TPTSPSRVTGPQARS SLPLTTTLGTTLRQA RKLEIIPATQEDALL RKLCFDYHHQMQA LKETYTEYLIEFLQH CQGNMMDFLPFKKK RNIIYQFLRQNDLD MEVDEEEADKNSRA
XP_018117870.1	XL-p400	X. laevis	SWR1	RARRQKALEAPREPR RPKVHWDHVLAEML WLAKFDSERKWKLS MAKKIAQRANKSIVD QAAKGERKQKEEHR MRKVAVNISKDVKKF WIKIEKLASSFLSSNVI VNCOTTLDISFLKKT YFATSLSSNMVGKQ VVYKHQLELEERKKK ALDKQLDFLLGQTER YSTMLAENLVDMPYS QNLETETSQINQPSH QEVVAEENLNATIPD DLDNMEVDDDYGSS LDEEPEDEHTIDEDE AQJTEAERNEELAAL AEADLPDLVILKMYT TKVTPNQANGCDHE LAHSSSDEGNSSEEV DDGHSYAEFVKKNH VHLLLSYMWVKVMEI FLLSMIRHVLVVVVY GFIYSLSHEDDYIAAE EVKDEATLSEEEQLA KKEDPDHLEIKLLQK ESEIPLEELLARYREDG CADHETAELNSPHF
AQK73705.1	ZmPIE1	Z. mays	SWR1	MEIDEEPMVFLDRT TRATRGKRMTKLLDD EVEEDEQFWNQEAL KEEEHDEYEAEREV ADEFDSDFNDEPEP DAVAVNEKLRDLPK KRLIYPGKTASKKKK KTKVVSQLEIYPGDEK PGEELGNKEQEEKEE NEAQEDMEGEKVIRK STRTSVVVRQAERDA LRAAIQATTKPIQRKK VGEEKRMTQEEMLLE AAQTEIMNLRNLERV LAREEEVKKAIVHKA VYKGPQIRYHSKOGC NYLEFCNGASFNSLS TKSVYPPEKAVCVITG LPAKYRDPKTGLPYAT RDAFKAIRERFMDEH DGLRKKMEMGDLFD TLVAKGFATKQKRTKI PKSNKSFSLRSSARFL SSESEEESEEDSD
NP_181212.2	AtSWC2	A. thaliana	SWC2	

				MDPMASAEQEV LLPRSARKTRGRMS KLLDEEIEDELFWNQ DALKEEELDAEYEEA EVADEFDSDFDEDEP EPDDEEAKNDGDERE RVKKKLIFPGKKAVKK KSKKKVLAAPKAENS SETPEPMQLEHQDIP DDLEGERAVRKSTRT SVIVRQAERDAIKAAL QATMRPVKRKKEGE EKRMTOEEMLEAA QTEIVNLNLERVLAR EEEVKKRAVVHKAVY NGPQIRFLSRNGKFTL EFNNGATFESAIAATP ASYAEPAVCAITSMP AKYRDPKTGLLYATKE AFKIIRERFYKEVSERE KERPDMGSLFALISGE GFSMKRRKSCMENR NQHLDYRRQYIF
XP_011620654.1	AtrSWC2	A. trichopoda	SWC2	
				MAPRRSSRAVKNK KESREGSSNDEHSNE SSTRSSKSRDGSHHN EEEGETSKASDPQINE ESEDETSNASDEQEE NDDTSSDEDAEPIVL NVTERARRATAGNK MAALLASADHEDEFY KTAYGGFEENNEVDK EFKSPVHSDDDDEVDS DFDKPEEEDEPASGG EEDGRPRRKRKFNE PKRGMTADDILAKNK KWAMARLAGNIVAA NSVDDKTQAAMLKE AEKTEKMNIESLKYYE AFELERKKREKNTVR VFPPGPREQKMTES GTTITLSEIKTFKCERP RERNLCAVTGRPARY LDPVTRLPYSTAYAFK VIRDYHKKHLSIRGN DEVTTYLKSALKALPSP QISPRVPLTTGPITSG PLVASSAGLTVMAT TSSAPKTTTH
NP_492505.1	CeVP572	C. elegans	SWC2	
				LSTADEELWQQLKE SDNEKEDPDYIVNSD EEEEEDIVSDAASD EEEEEGGGDDDETETK STKKSCKSKYKDPKSK KSISRRKKSSSDTDSK STMETELEKHLDNNT STTTTNDNTPNKRKR KEPSPPKTPNSEPTSP STRSSQSSAIEQLQ KSQPQTLTKNKKLV DDDDEEDESTTTTT TTATTSRRSTRNSES IKTTETTTTRKGGKNS KSDKNDEISSTPSTSI DGEGLSKLDISTYGL TQEELLECKETEINT ESLNHLQEEEDKKK VFHPKKALTGPRRIYR STPEQTTITFTDSLIPY CLTNQPQINNITSSG NEKSKKVRASRKSLD PPPIVDITSTTTTTTA ATTATINPTTATVTT ETNPSTTSTQEPIVEQ KDKGDTNRTENSENE IKKNIELCVITGLPAKY IDPESKKPFANLEAYKI
XP_629104.1	DdYL1	D. discoideum	SWC2	

NP_001285823.1	DmYL1	D. melanogaster	SWC2	MAASRSRRNNAGNK IAHLLNEEEEDDFYKT SYGGFQEEDEEDKEYE QKDEEEEDVVDSDFSI DENDEPVSDQEEAPE KKRKRGVVNTKAYKE TKPAVKKETKATPAL HKKRPGGGVTKRPRR PRFTVLDSGRKSIRTS TAIKTQATKIRLKELD DARKRKKKVRVEDY MPTQEELLEAKITEE ENTKSLEKFQKMELE KKKSRTPTKRTFSGPTI RYHSLTMPAMRKPT RGANPAVDKDLAG KCERTFVTIENDFNCK VFQSLFRHKAPPKAS NGICPITRLPARYFDPI TQQPYYSIQAFKILRE AAYMQLQEQGGGSE QPELAKWLEWRKLV KENRLKASAAASKNG DN
NP_001076413.1	DrVps72	D. rerio	SWC2	MSLANSREQRSTAG NRMSKLLDAEEDEF YKTTYGGFNDESGDD EYHGDHSDTEDEVDS DFDIDEGDEPDSGQE EDAPKRKSRVVTKAY KEPLKVSKPKVKRVSE ELKPERPRVERRTVRD ELQDLGDIRKSVRKST SEHTRKTNERLQERQ QEAPRRKRGAQSERV LTQDELLDEAKLTAE NLRSLENYERLEADKK KQVHKRRFEGPMIR YHSVLMPLLPDTHLK EENVDEGLDQDTP QATPTSSSTQGAGS LCSRTYITFSDDEAFSS AFPSAARCTPTHVPQ EVCPTVTHKALYRDP VTDIPYANARAFRIIR EAYQKYIAAHGFPNA SGSFSANTDASDSPA NLKSARPKAVLKQSA AAT
NP_001258016.1	HsYL1	H. sapiens	SWC2	MSLAGGRAPRKTAG NRLSGLLEAEEDEFY QTTYGGFTEESGDDE YQGDQSDTEDEVDS DFDIDEGDEPSSDGE AEEPRRRRVVTKAY KEPLKSLRPRKVNTPA GSSQKAREEKALLPLE LQDDGSDSRKSMRQ STAETHRTQTLRVQE RQGQSRRRKGPHCE RPLTQEELLREAKITEE LNLSLETYERLEADK KKQVHKRKCPCPIIT YHSVTVPLVGEPGPK EENVDIEGLDPAPSVS ALTPHAGTGPVNPPA RCSRTFITFSDATFE EWFPPQGRPPKVPVR EVCPTVTHKALYRDP VTDIPYATARAFKIIRE AYKKYITAHGLPPTAS ALGPGPPPPPEPLPGS GPRALRQKIVIK MSLAGGRAPRKTAG NRLSGLLEAEEDEFY QTTYGGFTEESGDDE YQGDQSDTEDEVDS DFDIDEGDEPSSDGE AEEPRRRRVVTKAY KEPLKSLRPRKVSTPA SSSQKAREEKTLLPLEL QDDGSDSRKSMRQS TAEHTRQTLRVQER QGQSRRRKGPHCER PLTQEELLREAKITEEL NLRSLETYERLEADKK KQVHKRKCPCPIITY HSVTVPLVGEPGPK ENVDVEGLDPAPTAS ALAPHAGTGTGAAA ATPPAHCSTRTITFSD DATFEWFPQGRPP KVPVREVCPTVTHRPA LYRDPVTDIPYATARA FKIIREAYKKYITAHGL PPTASALGPGPPPE PLPGSGPRALRQKIVI K
NP_033362.2	MmYL1	M. musculus	SWC2	

XP_024394407.1	PpaSWC2	P. patens	SWC2	RLSKLLDEEIEADEEF WNQDAFKEDVADDE YEEEQELADEFDSDF NDDEESGDDEEVEEK ERRPKKKQLPPGGKK PLKKGPSGSTKKKGV SFLEAAAAADAGET PLPTLKTPTQATRPSP EKEDGETEEGEKLLRK STRTSVIVQAEREA QRAIQQALPKIVKKK REGEDRKMTQEEML LEAAQTEIQNRQSLET LLAREEEVKRKAIINKE VYSGPLIRFYSKEGIN VLEFVKMPEVPEVIN AKAPPTCTFTFRNPKQ PVCVVTGLLAKYRDP KSGFPYATKEAFIIRE RQRKGEDISRKDSLEK VQRKRTKVEREKFAP GTRTGAVKIKYKKPD KPVETKFQNSPKPQS QPSVENRASGGAGP SGATSSGLTLDKPE NSNKVGEFEDKLDHT STPAPAPAPLPTPMV EANEADFMLEDLGV5
-	PpSWC2	P. polyceph	SWC2	MSTRSLPSRSTRGKR MNALVGDDALVDEE FWKMDMFKEDEED DEYKSEAEEDDIIDAD FDNPEEAEESEVEVA KEKKEKKRMAYVDPK AKKGPKRRKIQDKAK ALDAELLKGPESDVEV IIMDEEPNEEKGEAPK AEPKKGRGRRKKTTP KALEIPEGEGADAAS PPSTPGSVNTPKSTR KSTRAVAIQRAEQRE KEREETEKKRSQRSRK RSPVVRRITQEELLE EAKMTEIYNTESLKYIL SMEEKKKIKEAPTIV GPAIKYYSKDGTVCVC FDELDPVLASAPPTPP SREICPITGLPAKYKDP KTGICYANLDAFKTLR TKLLQQEEAETNLKIA QLTALLQESKRKKDQ LQNTPTGVLV
NP_010773.4	ScVps72	S. cerevisiae	SWC2	FIIQTRSRRSNAGNKL QKLLLEQLRDIESTKR QISSYKNGNDDEEDEI GLLFQEDDEDEDFEM MAKDDDDDEEEKED ETQSIRKEPSQASSEQ AADDLMFSSSESED5 SNENDEDAEEKEIRR QELLSRKKRNKRLQK GPVVIKKQKPKPSG EAIIPRSHHTHEQLNA ETLLLNTRRTSKRSSV MENTMKVYEKLSKAE KKRKIIQERIRKHKEQ ESQHMLTQEERLRIA KETEKLNILSLDKFKEQ EVWKKENRLALQKR QKQKFQPNETILQFLS TAWLMTPEMELEDR KYWQEQNLNRDKKK KKYPRPKKNLNLGK QDASDDKKRESEESIK NDGDVNSLGENSSSV HNQKRIETSTNDTV EGESSPDAAVSRVNS DELKPTALPDVTLDAI ANKQSTVDEAPNSQ PQKNIITNEQKITNVG

				KNEEIYNSDQDDQD QTPONKSHSNNSNQ VGAQSKKGAVQDKS AKNSKVTASKGNQKK KSQEVDSKAPKKGKA ISKGKTNNKKGDKGG AASKGGKKKVQKG KKGEKAPVKGSA TKK KGGAQDKKKADNSA LKKTAKISQGA EIIED EKEEVFNDESATPP QYSQTNATSYQSDNI SKLKKKPAKASNKK TQQTAAASKLKTQKE EQGEKSSSLKAFSSIIK DKKNGKNVPFQPT NKESEEIEIEPETDEQE SDPNQLKKLGKSDL NTEHQNGGNQVKD DKKGQSESLKNRSK VLVDSKKPNSKLYKN DNGHEQSTTSKGSKV SAIKNIRKDSLDMELD NTLENEEDEDILSGS EDSENSQNSDNITN KKQLDKKKSDLSLSSKL KNKADESEDDGEGED
XP_001019354.1	TtYL1	T. thermophilus	SWC2	NEEDEVDDEDEEE MNLADGRAPRKTAG NRMSGLLQAEEDDF YKTTYGGFNESGDE EYNEDRSASEDEVDS DFDIDEGDEPTSDHE EDEPKKKRRVVTKAY KEPIQLLKPKPKKPEA PPNTAAKS RPEKPQE PPDDTVDSRKQMRQ STTEHTRQTLRVKER QIQSKKKKGPHLDRP LTQEELLEAKITEINI RSLNRYERLEADRKK QVHKRRRCAGPTIRY HSMVMPLITELRMKE ENVDVGLDHEQTD RTHAGKCSRSFITFSD DETFFRFPKSRGKF SVRDVCPVTHKPALY RDPITDIPYNSKAFKI IRDAYKKYITTHGLPN AAMATTMGPSADA AQRNTRQIIIKQSVPSA
NP_001085907.1	XlVps72	X. laevis	SWC2	MDAGDDEPPVLLDR AARASRGKRITKLLLED EVEQDEAFWNQEAL KDEENDDNVVEEPDA GDEFDSDFGEDESET DDEPEKEVRERLPIKK RLMFPGKTLRKTNVK KKKATPKPEDGTKAD KSADKPSSTQADLP DELEAEKAIRKSSRTS VIVRQAEREAIRAEKE ATAKPIKRKKEGEEKR MTQEEMLLEAAETEI MNMNRLERVLAREE EVKKKAVVHKDITYEG PTIRFFSRDGESRLEFI NGASFGESELCCTSSPY PEKPVCVVTGLPAKY RDPKTGLPYATMEAF KIIRSFLEKEADRKR NMSNMGELFESITG GHLMPKKRRVEIRSP NMLGGSRHGGRTW RIPAFDMVDED
ACG47775.1	ZmVps72	Z. mays	SWC2	MEVDLPRSVKRMV KEKLSYLSKVDDPSK DVNISNEALSASFESA RIFIHYLSATANDICQ ESKRQTINAEDVFKAL EEIDFPFTEPLRVSL DFRKKKAFKSKETN KKRKSEEPESNVIGS EGKGEETEEGGHTK EDMDEQEEDCGDEG KEDEEE
XP_006855071.1	AtrPolE3	A. trichopoda	PolE3	MSSNDVDQNDRIA QMMMPAAIVTRLMI KEDNISASKDARDVIA RAAAVFNLNSDVSA QSARDQKHKTISGDD VVKGLRELENTSIHN HCKSANDKWKVISQ QKALARKNPVESSSV NAEMDDDIIDTPVD DSL MN
NP_498421.1	CePolE3	C. elegans	PolE3	

XP_645298.1	DdPolE3	D. discoideum	PolE3	MSESQDLPGAIVNRH KASLPEGVLCAKESRL AIAKAAKVWIHYLTA ASIDFSSHSGRSTISPK DVFAQIEEIDFENFKP QLEEYLAALKSEKEKE KEKEKEKEKEKEKD SNKDDTKSKDENKSS KKSSSSSKDEK
NP_476646.2	DmChrac-14	D. melanogaster	PolE3	MVERIEDLNLPNAV GRLIKEALPESASVSK EARRAIAAASVFAIF VTSSSTALAHKQNHK TITAKDILQTLTLDPE SFVPSLTQDLEVVRKV VKEKKESKASKDSNT AENANASATAAEEA PE
NP_957095.1	DrPolE3	D. rerio	PolE3	MAERPEDLNLPNAV TRIIKEALPEGVNVSK EARRAISQAASVFVLY ATSCANSFAMKAKRK TLNAGDVMSAMEE MEFERFLQPLREALE AYKKGQKGKKEASEQ KRDKEKKNGTDEND KSRDEEEDEHMDDE QDGENEAEEEDVEN
NP_001265184.1	HsPolE3	H. sapiens	PolE3	MAERPEDLNLPNAV TRIIKEALPDGVNISKE ARSAISRAASVFLYA TSCANNFAMKGKRR TLNASDVLAMEEME FQRFTPLKEALEAYR REQKGKKEASEQKKK DKDKKTDSEEQDKSR DEDNDEEERLEEEE QNEEEVDN
NP_067473.2	MmPolE3	M. musculus	PolE3	MAERPEDLNLPNAV TRIIKEALPDGVNISKE ARSAISRAASVFLYA TSCANNFAMKGKRR TLNASDVLAMEEME FQRFTPLKEALEAYR RDEKGKKEASEQKKK DKDKKTDSEEQDKSR VEEDEERLDEDDQN EEEEIDN
-	PpPolE3	P. polycephalus	PolE3	MSGDKTEPKKEEKS GEDLELPRAVVQRLIK YSLPDNVHVQAEAKL AIAHAGKVFINYLTAC ANDYCIQGTRSTISAK DVLNACDELEMPFIT KKLKEVLEAYKKEQSE KKEKGKDEKDEGKS ED
NP_010406.3	ScDPB4	S. cerevisiae	PolE3	MPPKGWRKDAQGN YPTTSYIKEQENITIQQ LLFPKSTIVNLAREVP QQSGKKLLINKDASL ALQRGATVFVNHLL FAREIAKSQDKKSCSV DDVLSALDHIGHSAL KGPVRDKLDEYQAAV EQRKKEKLDGGEVDA DGDIDMGEDKENVP VEKVKEHDEIEEQGD ALQDVEESSEKKQKT ESQDVETRVQNLEQT
XP_018084282.1	XICHRAC17	X. laevis	PolE3	MAERPEDLNLPNAV VTRIIKEALPEGVNISK EARSASRAASVFLY ATSCANNFAMKGKR KTLNASDVLAMEE MEFQRFPLTKESLEV YRQDQKGKKEATEQ KKDKKEKKADSEDD KSREENEEDEKME EDEVVVEEEVEN

OA093127.1	AthHSP90.1	A thaliana	HSP90	FQAEINQLLSIJINTFY SNKEIFLRELISNSSDA LDKIRFESLTDKSKLD GQPELFIRLVDPKAN KTLSIIDSGIGMTKAD LVNNLGTIARSGTKEF MEALQAGADVSMIG QFGVGFYSAYLVAEK VVVTTKHNDDEQYV WESQAGGSFTVTRD VDGEPLGRGTKISLFL KDDQLEYLEERRLKDL VKKHSEFISYPIYLWTE KTTKEISDDEDEDEP KKENEGEVEEVDEEK EKDGKKKKKIKEVSHE WELINKQKPIWLKRP EETKEEYAAFYKSLTN DWEDHLAVKHFSVE GQLEFKALFVPRAP FDLFDTRKKLNLIKLY VRRVFIMDNCEELIPE YLSFVKGVDSDDLPL NISRETLQQNLIKVI RKNLVKKCIEMFNEIA ENKEDYTKFYEAFSKN LKLGIHEDSQNRGKIA DLLRYHSTKSGDEMT
OA094685.1	AthHSP90.2	A thaliana	HSP90	QLLSIJINTFYSNKEIFL RELISNSSDALDKIRFE SLTDKSKLDGQPELFI HIIPDKTNNLTIIIDSG IGMTKADLVNNLGTI ARSGTKEFMEALAAG ADVSMIGQFGVGFYS AYLVADKVVVTTKHN DDEQYVWESQAGGS FTVTRDTSGETLGRGT KMLVLYKEDQMEYLE ERRKDLVKKHSEFISY PISLWIEKTIEKISDD EEEEEKKDEGKVEEV DEEKEKEKKKKIKE VSHWDLVNNKQPI WMRKPEEINKEEYAA FYKLSNDWEEHLAV KHFSVEGQLEFKALF VPKRAPFDLFDTKKK PNNIKLYVRRVFIMD NCEDIPEYLGFKGIV DSEDPLNISRETLQQ NKILKIRKNLVKKCL ELFFEIAENKEDYNKF YEAFSKNLKLGIHEDS QNRTKIAELLRYHSTK SGDELTSKDYVTRM
XP_006828794.1	AtrHSP83	A. trichopoda	HSP90	FAFQAEINQLLSIJINT FYSNKEIFLRELISNSS DALDKIRFESLTDKSKL DAQPELFIRLVDPKV NKTLSIIDSGIGMTKA DLVNNLGTIARSGTK EFMEALQAGADVSM IGQFGVGFYSAYLVAE KVIVTTKHNDDEQYV WESQAGGSFTVTRD VNGEPLGRGTKITLFL KEDQDYLEERRLKD LVKKHSEFISYPIYLWT EKTTEKESDDEDEEV DKEKKEGDVEEVDD EGKDKKKKKMKEVS HEWTLVNNKQPIWM RKPPEITKEEYASFYKS LTNDWEDHLALKHFS VEGQLEFKALFLPKR APFDLFDTRKKMNNI KLYVRRVFIMDNCEE LIPEYLGFKGVVDS DLPLNISREMLQQNK ILKVIRKNLVKKCIEMF FEIAENKEDYNKFYEA FSKNIKLGIHEDSQNR AKLADLLRYHSTKSGD

				QLMSLIINTFYSNKEIY LRELISNASDALDKIRY QALTEPSELDTGKELFI KITPNKEEKTLTIMDT GIGMTKADLVNNLG TIAKSGTKAFMEALQ AGADISMIGQFGVGF YSAFLVADKVVVTSK NNDDDSYQWESSAG GSFVVRPFNDPEVTR GTKIVMHKEDQIDFL EERKIKEIVKHSQFIG YPIKLVVEKEREKEVE DEEAVEAKDEEKEG EVENVADDADKKTK KIKEKYFEDELNKT PIWTRNPDDISNEEY AEFYKSLNDWEDHL AVKHFSVEGQLEFRA LLFVPRAPFDLFENK KSKNSIKLYVRRVFIM ENCEELMPEYLNFIG VVDSDELPLNISREML QQSKILKIRKNLVKK CMELIDEVAEDKDNF KKFYEQFGKNLKLGIH EDSTNRKKLSDFLRY TSAGDEPTSLKEYVSR
NP_506626.1	CeHSP90	C. elegans	HSP90	
				NQLMSLIINTFYSNKE VFLRELISNASDALDKI RYQSLTDASVLESKTE LEIKIIPDKTAKTLTID SGIGMTKTDMMVKNL GTIARSGTKNFMEQL QSGAADISMIGQFGV GFYSAYLVADTVIVHS KNNDEQYVWESSA GGEFTIALDHTEPLGR GTKIVLHMKEDQLDY LDETKIKNLVKKHSEFI QYPISLLTIKEKEVDEE TTAKEGEEESTDAKIE EIEEEKEKKVKVQEK EWDVLNKTPLWTR NPSDVTKEEYNSFYKS ISNDWEEPLAVKHFS VEGQLEFKALFVPPK APFDLFESKKANNIK LYVRRVFIMDNCAHII PEYLNFRGIVDSDEL PLNISRETLQQNKILT VIRKNLVKKCIELEFNEI AENSEDYKKFYAFSK NLKLGVEDSQNREK FADLLRYQTSKSGDEL VTLKEYVGRMKEGQ
XP_647482.1	DdHSP90	D. discoideum	HSP90	
				QLMSLIINTFYSNKEIF LRELISNASDALDKIRY ESLTDPKSLDSGKELYI KLIPNKTAGTLTIIDTG IGMTKSDLVNNLGTI AKSGTKAFMEALQA GADISMIGQFGVGFY SAYLVADKVTVTSKN NDDEQYVWESSAGG SFTVRADNSEPLGRG TKIVLYIKEDQTDYLEE SKIKEIVNHSQFIGYP IKLLVEKEREKEVSDD EADDEKKEGDEKKE METDEPKIEDVGEDE DADKKDKDAKKKTI KEKYTEDELNKTPI WTRNPDDISQEEYGE FYKSLTNDWEDHLAV KHFSVEGQLEFRALLF IPRRTPPDLFENQKKR NNKLYVRRVFIMDN CEDLIPEYLNFMKGV VVDSDELPLNISREML QQNKVLKIRKNLVK KTMELIEELTEDKENY KKFYDQFSKNLKLGV HEDSNRRAKLADFLR
NP_001261362.1	DmHSP83B	D. melanogaster	HSP90	

				EVETFAFQAEIAQLM SLIINTFYSNKEIFLREL ISNSSDALDKIRYESLT DPSKLDSCDKLIELIP DQKERTLTIIDTGIGM TKADLINNLGTIAKSG TKAFMEALQAGADIS MIGQFGVGFYSAYLV AEKVTVITKHNDDEQ YIWESAAGGSFTVKP DFGESIGRGTKVILHL KEDQSEYVEEKRIKEV VKKHSQFIGYPITLYIE KQREKEVDLEEGERQ EEEEEVAAGEDKDKPKI EDLGADEDEDSKDGK NKRKKVKKEYIDAQ ELNKTPIWTRNPDD ITNEEYGEFYKSLSD WEDHLAVKHFVVEG QLEFRALLFVPRRAAF DLFENKKRRNNIKLYV RRVFIMDNCEELIPEY LNFIKGVVDSDELPLN ISREMLQQSKILKVIR KNLVKKCLDLFTELAE DKDNYKKYEQFSKN IKLGIHEDSQNRKLS
NP_571403.1	DrHSP90A	D. rerio	HSP90	
				AFQAEIAQLMSLIINT FYSNKEIFLRELVSNAS DALDKIRYESLTDPTK LDSGDKLIDIIIPNVQ ERTLTJDTGIGMTKA DUINNLGTIAKSGTKA FMEALQAGADISMIG QFGVGFYSAYLVAEK VTVITKHNDDEQYA WESSAGGSFTVKVD HGEPIGRGTKVILHLK EDQTEYIEEKRVKEVV KKHHSQFIGYPITLYVEK ERDKEISDDAEKEKA EKEKEEGEDKPKIE DVGSDDEEDTKDKDK KKKKKIKKYIDQEEL NKTPIWTRNPDDIS NEEYGEFYKSLTNDW EDHLAVKHFVVEGQL EFRALLFIPRRAPDFL ENKKKKNNIKLYVRR VFIMDNCEELIPEYLN FIRGVVDSDELPLNIS REMLQQSKILKVIRKN IVKKCLELFAELAEK DNYKKFYDAFSKNLKL GIHEDSQNRKLSSELL
NP_571385.2	DrHSP90B	D. rerio	HSP90	
				EEVETFAFQAEIAQL MSLIINTFYSNKEIFLR ELISNSSDALDKIRYES LTDPSKLDGKELHIN LIPNKQDRTLTIIVDTG IGMTKADLINNLGTIA KSGTKAFMEALQAG ADISMIGQFGVGFYS AYLVAEKVTVITKH DDEQYAWESSAGGS FTVRTDTGEPMGRG TKVILHLKEDQTEYLE ERRIKEIVKKHSQFIGY PITLFVEKERDKEVSD DEAEKEKDEKEKEKE EKESEDKPEIEDVGSD EEEEEKDGDKKKKKKI KEKYIDQEELNKTPI WTRNPDDITNEEYGE FYKSLTNDWEDHLAV KHFSVEGQLEFRALLF VPRRAPDFLFENRKK KNNIKLYVRRVFIMD NCEELIPEYLNFIIRGV VDSDELPLNISREML QQSKILKVIRKNLVKK CLELFTELAEDKENYK KFYEQFSKNIKLIGIHE
NP_001017963.2	HsHSP90A	H. sapiens	HSP90	

NP_001258898.1	HsHSP90B	H. sapiens	HSP90	AFQAEIAQLMSLIINT FYSNKEIFLRELISNAS DALDKIRYESLTDPSK LDSGKELKIDIIPNPQE RTLTLVDTGIGMTKA DLINNLGTIAKSGTKA FMEALQAGADISMIG QFGVGFYSAYLVAEK VVVITKHNDDEQYA WESSAGGSFTVRAD HGEPIGRGTKVILHLK EDQTEYLEERRVKEVV KKHSQFIGYPITLYLEK EREKEISDDEAEKEEG EKEEEDKDEEKPKE DVGSDEEDDSGKDKK KKTKKIKEKYIDQEEL NKTKEIPIWTRNPDDIT QEEYGEFYKSLTNDW EDHLAVKHFSVEGQL EFRALLFIPRRAPDFLF ENKKKKNNIKLYVRR VFIMDSCELIPEYLN FIRGVVDSDELPLNIS REMLQQSKILKVRKN IVKKCLELFSLEADKE NYKKFYEAFAFSKNLKL IHEDSTNRRRLSELLR
NP_034610.1	MmHSP90A	M. musculus	HSP90	EEVETFAFQAEIAQL MSLIINTFYSNKEIFLR ELISNSSDALDKIRYES LTDPSKLDSGKELHIN LIPSKQDRTLITVDGI GMTKADLINNLGTIA KSGTKAFMEALQAG ADISMIGQFGVGFYS AYLVAEKVTITKH DDEQYAWESSAGGS FTVRTDTGEPMGRG TKVILHLKEDQTEYLE ERRIKEIVKKHSQFIGY PITLFVEKERDKEVSD DEAEKEKEKEKEKEKE EKESDDKPEIEDVGSD EEEEKKDGDKKKKKK KIKEKYIDQEELNKT PIWTRNPDDITNEEY GEFYKSLTNDWEEHL AVKHFSVEGQLEFRA LLFVPRRAPDFLFENR KKKNIKLYVRRVFIM DNCEELIPEYLN FIRGVVDSDELPLNIS REMLQQSKILKVRKN IVKKCLELFSLEADKE NYKKFYEAFAFSKNLKL IHEDSTNRRRLSELLR
NP_032328.2	MmHSP90B	M. musculus	HSP90	AFQAEIAQLMSLIINT FYSNKEIFLRELISNAS DALDKIRYESLTDPSK LDSGKELKIDIIPNPQE RTLTLVDTGIGMTKA DLINNLGTIAKSGTKA FMEALQAGADISMIG QFGVGFYSAYLVAEK VVVITKHNDDEQYA WESSAGGSFTVRAD HGEPIGRGTKVILHLK EDQTEYLEERRVKEVV KKHSQFIGYPITLYLEK EREKEISDDEAEKEEG EKEEEDKDEEKPKE DVGSDEEDDSGKDKK KKTKKIKEKYIDQEEL NKTKEIPIWTRNPDDIT QEEYGEFYKSLTNDW EDHLAVKHFSVEGQL EFRALLFIPRRAPDFLF ENKKKKNNIKLYVRR VFIMDSCELIPEYLN FIRGVVDSDELPLNIS REMLQQSKILKVRKN IVKKCLELFSLEADKE NYKKFYEAFAFSKNLKL IHEDSTNRRRLSELLR

				AFQAEINQLLSLIINTF YSNKEIFLRELISNSSD ALDKIRFESLTDKSKLD GQPELFIHIVPDKAN NTLSIISDSIGMTKAD MVNNLGTIARSGTKE FMEALSAGADVSMI GQFGVGFYSAYLVAE KVVVTSKHNDDQYI WESQAGGSFTITRDT SGEPLGRGTHIKLYK EDQLEYLEERRKDLV KKHSEFISYPISLWTEK TTEKEVSDDDEDDK KDEEGKIEEVDEEKEK DKKKKKVKEISREWTL INKQKPIWMRKPED VTKEEYAAFYKSLTND WEEHLAVKHFSVEG QLEFKSVLFVPKRAPF DLFDSRKKQNNIKLYV RRVFIMDNCEELIPEY LGFVKGVVDSDELPL NISRETLQQSKILKVIR KNLVKKCMEMFSEIA ENKEDYQKFYEAFSK NLKLGIHEDSQNRSL ADLLRYHSTKSGDEM
XP_024395876.1	PpatHSP80	P. patens	HSP90	
				LMSLIINTFYSNKEIFL RELISNASDALDKIRYE SLTDKSKLEGAPELFIH IVPDKENKTLTLDGTI GMTKADLINNLGTIA RSGTKNFMHQLSGS ADISMIGQFGVGFYS AYLVADKVTVTSKHN DDEQYIWESSAGGEF FIRRDTEQLGRGTKI ILHLKDDQLEYLEEKRI KDLVKKHSEFIQYPISL WETKETEKEVSDDDE EVKEGEKPTIEEVDED KEADKKKKKKVKEVT HEWVLLNKQKPIWT RNPADITKEEYAAFYK SISNDWEEHLAVKH SVEGQLEFRAILFVPK RAPFDMFESRKKLNN IKLYVRRVFIMDDCKD LIPEWLNFIKIVDSE DLPLNISRETLQQNKI LKVIKKNLVKKCIELFF EIAENAEDFKKFYAF GKNLKLGIHEDSQNR AKLAELLRYNSTKSGD EPTSLKDYITRMPESEQ
-	PpHSP90	P. polycephala	HSP90	
				MSLIINTVYSNKEIFLR ELISNASDALDKIRYKS LSDPKQLETEPDLFIRI TPKPEQKVLEIRDSGI GMTKAELINNLGTIA KSGTKAFMEALSAGA DVSMIGQFGVGFYSL FLVADRVQVISKSND DEQYIWESNAGGSFT VTLDEVNERIGRGTIL RLFLKDDQLEYLEEKRI KEVIKRHSEFVAYPIQ LVVTKEVEKEVPIPEE EKKDEEKDEEKKDE DDKKPKLEEVDDEEEK KPKTKKVKEEVQEIEE LNKTKPLWTRNPSDI TQEEYNFYKSISND WEDPLYVKHFSVEG QLEFRAILFIPKRAPFD LFESKKKKNNIKLYVR RVFITDEAEDLIPEWL SFVKGVDSEDLPLNL SREMLQQNKIMKVIR KNIVKKLIEAFNEIAED SEQFEKFYSAFSKNIK GVHEDTQNRALAK LLRYNSTKSVDELSTLT
NP_015084.1	SchHSP82	S. cerevisiae	HSP90	

				INTFYSNKEIFLRELIS NASDALDKIRYISITDS EKAKLEVEPNFRIRIIP DKANNTLTLWDTGIG MTKKELINNLTGIAKS GTKAFMEALSSGADI SMIGQFGVGFYSAYL VAKVEVISKSNDDES QWRWESSAGGTFTV VNDDENPEKLTGRGK IILHMKNDNLEFLEER RIKDLIKKHSEFIAFPIE LQVEKTEEKETDEED EEKEKEDKEKTDEPEI KEETEKDKKKKKVKV VHTEFEEQNKNKPL WMRKPEITKEEYVN FYKSLTNDWEEHQA VKQFSVEGQLEFRAIL FIPKRAPFDLFETKKK KNNIKLYVRRVFIMD DCEELIPEYLNFIKGVV DSEDLPLNISREFLQH NKILKVIKKNIVKKCLD LIQEVADNEEDFKKFY EQFGKNLKLGIHEDS ANREKLSSFLRYHSSK SGEELTLTKDYVSRM
AF151114_1	TtHSP82	T. thermophi	HSP90	
				QDQQMEEDVETFAF QAEIAQLMSLIINTFY SNKEIFLRELISNSSDA LDKIRYESLTDPKLDOS GKELKIELIPNKQDRS LTIIDTGIGMTKADU NNLTGIAKSGTKAFM EALQAGADISMIGQF GVGFYSAYLVAEKT VITHIDDEQYAWES SAGGSFTVRVDNSEP LGRGTKVILHLKEDQS EYFEEKRIKEIVKKHSQ FIGYPITLFVEKERDKE ISDDEAEKEKEKKDE PKDEEKPEIEDVGSDD EEDKKEGDKKKKKIK EKYIDQEELNKTPIW TRNPDDITNEEYGEFY KSLTNDWEDHLAVK HFSVEGQLEFRALLFV PRRAPFDLFENRKKK NNIKLYVRRVFIMDN CDELIPEYLNFMRGV VDSDELPLNISREML QQSKILKVRKNLVKK CLELFTELSEDKENYK MFYEHFSKNIKLGIHE
XP_018086052.1	XIHSP90A	X. laevis	HSP90	
				AFQAEIAQLMSLIINT FYSNKEIFLRELISNAS DALDKIRYESLTDPK LDSGKDLKIDIIPNRQ ERTLTVIDTGIGMTKA DLINNLTGIAKSGTKA FMEALQAGADISMIG QFGVGFYSAYLVAEK VVVITHKNDDEQYV WESSAGGSFTVKIDH GEPGRGTKVILHLKE DQTEYLEEKRVKETVK KHSQFIGYPITLYLEKE REKEISDDEAEKEKE EEKKEEGENEKPKIE DVGSDDEEGKDKKK KTKKIKEKYIDQEELN KTKPIWTRNPDDITQ EEYGEFYKSLTNDWE DHLAVKHFSVEGQLE FRALLFIPRRAPFDLFE NKKKKNNIKLYVRRV FIMSDCELIPEYLNFI RGVVDSDELPLNISRE MLQQSKILKVRKNIV KKCLELFSELAEDKEN YKKFYEAFSKNLKLGI HEDSTNRKRLSSELLRY
XP_041418547.1	XIHSP90B	X. laevis	HSP90	

NP_001135416.3	ZmHSP82	Z. mays	HSP90	TETFAFQAEINQLLSLI INTFYSNKEIFLRELIS NASDALDKIRFESLTD KSKLDAQPELFIRLVP DKASKTLSIIDSGVGM TKSDLVNNLGTIARSG TKEFMEALAAAGATDV SMIGQFGVGFYSAYL VADRVMTTKHNDD EQYVWESQAGGSFT VTHDTTGERLGRGTK ITLFLKDDQLEYLEERR LKDLVKKHSEFISYPIY LWTEKTTKEISDDEE EDNKEEEGDVEEVD DEDKADKSKKKKK VKEVSHWVQINKQ KPIWLKPEEITREEY ASFYKSLTNDWEDHL AVKHFSVEGQLEFKAI LFVPRRAPFDLFDTRK KLNNIKLYRRVFIMD NCEELIPEWLGFKG VVDSDDLPLNISRETL QQNKILKIRKNLVKK CIEMFIEAENKDDYA KFYDAFSKNIKLGIHE DSQNRAKLADLLRYH
NP_195870.1	AtHSP70-1	A thaliana	HSC70	TTYSCGVWQHDRV EIIANDQGNRTTPSYV AFTDSERLIGDAAKN QVAMNPVNTVFDKAK RLIGRRFSDSSVQSD MKLWPFKIQAGPAD KPMIYVEYKGEEKEFA AEEISSMVLKMKREIA EAYLGVTIKNAVVTVP AYFNDSQRQATKDA GVIALNVMRIINEPT AAAIAYGDKKATSV GEKNVLIFDLGGGT DVSLLTIEEGIFEVKAT AGDTHLGGEDFDNR MVNHFVQEFKRRSK KDITGNPRALRRLRTS CERAKRTLSSTAQTTI EIDSLYEGIDFYSTITR ARFEELNMDLFRKC MEPVEKCLRDAKMD KSTVHDVVLVGGSTR IPKVQQLQDFFNGK ELCKSINPDEAVAYG AAVQGAILSSEGNEK VQDLLLDVTPLSLGL ETAGGVMTTLIPRNT TIPTKKEQVFSTYSDN
XP_006833219.1	AtrHSC70	A. trichopoda	HSC70	TTYSCGVWQHDRV EIIANDQGNRTTPSYV AFTDTERLIGDAAKN QVAMNPVNTVFDKAK RLIGRRFSDASVQSDI KLWPFKVIPGDKP MIVVQYKGEDKQFA AEEISSMVLKMKREI AEAYLGSVKNNAVVT VPAYFNDSQRQATK DAGVIAGLNVMRIIN EPTAAAIAYGDKKAT SVGEKNVLIFDLGGG TFDVSLLTIEEGIFEVK STAGDTHLGGEDFDN RMVNHVQEFKRRKH KKDISGNPRALRRLRT SCERAKRTLSSTAQTT IEIDSLYEGIDFYSTITR ARFEELNMDLFRKC MEPVEKCLRDAKMD KNTIHDVVLVGGSTRI PRVQQLQDFFNGKE LCKSINPDEAVAYGA AVQAAILSSEGNEKV QDLLLDVTPLSLGL TAGGVMTVLIIPRNTT IPTKKEQVFSTYSDNQ

				SCVGVFMHGKVEIIA NDQGNRTTPSYVAF DTERLIGDAAKNQVA MNPHTVFDARLI GRKFDDPAVQSDMK HWPFKVISAEGAKPK VQVEYKGENKIFTPEE ISSMVLKMKETAEEA FLGTTVKDAVTVPA YFNDSQRQATKDAG AIAGLNVLRIINEPTA AAIAYGLDKKGHGER NVLIFDLGGGTFDVS LTIEDGIFEVKSTAGD THLGGEDFDNRMVN HFCAEFKRKHKKDLA SNPRALRRRTACER AKRTLSSSSQASIEIDS LFEGIDFYTNITRARFE ELCADLFRSTMDPVE KSLRDAKMDKSQVH DIVLVGGSTRIPKVQK LLSDFSGKELNKSINP DEAVAYGAAVQAAIL SGDKSEAVQDLLLLD VAPLSGIETAGGVM TALIKRNTTIPKTAQ TFTTYSDNQPGVLIQ
NP_503068.1	CeHSP-1	C. elegans	HSC70	VWQNDRVEIIANDQ GNRTTPSYVAFDTE RLIGDAAKNQVAMN PTNTVFDARLIGRKF SDKEVQSDMKHWPF KVIPKDGDKPHIQVEF KGETKVFSPEISSMV LLKMKETAAYLGKTI NNAVITVPAYFNDSQ RQATKDAGTISKLVN QRIINEPTAAIAYGL EKKGSGEKNLIFDLG GGTFDVSLLTIEDGVF EVKATAGDTHLGGED FDNRLVSHFVDEFKR KHKDIMGNQRAVR RLRTACERAKRTLSS AQASIEIDSLFEGIDFY TSITRARFEELCADLFR GCLDPVEKVLKDSKL DKKSIHEIVLVGGSTRI PKVQQLLEFFNGKE LNKSINPDEAVAYGA AVQAAILSNEGGAKV ADLLLLDVAPLSMGL ETAGGVMTTIPRNT TIPCKKTQTFSTYSDN QPGVLIQVYEGERAM
XP_646617.1	DdHSC70	D. discoideum	HSC70	GVYQHGKVEIIANDQ GNRTTPSYVAFDSE RLIGDPAKNQVAMN PRNTVFDARLIGRKY DDPKIAEDMKHWPF KVVSDGGKPKIGVEY KGESKRFAPEEISSMV LTKMKETAAYLGESI TDAVITVPAYFNDSQ RQATKDAGHIAGLN VLRIINEPTAAALAYG LDKNLNGERNVLIFDL GGGTFDVSILTIDEGS LFEVRSTAGDTHLGG EDFDNRLVTHLAEFF KRKYKDLRSNPRAL RRLRTAAERAKRTLSS STEATIEIDALFEGQD FYTKVSRARFEELCAD LFRNTLQPVKALND AKMDKGQIHIVLVG GSTRIKQVQSLLEFF HGKNLNLINPDEAV AYGAAVQAAILSGDQ SGKIQDVLLVDVAPLS LGIETAGGVMTKLIER NCRIPCKQTKTFSTYS DNQPGVSIQVYEGER
AAG26909.1	DmHSP70Bb	D. melanogaster	HSC70	

				SCVGVFQHGKVEIA NDQGNRTTPSYVAF DTERLIGDAAKNQVA MNPTNTVFDKRLIG RRFDDGVVQSDMKH WPFNVINDNSRPKV QVEYKGESKSFYPEE SSMVLTKMKEIAEAY LGKTVSNAVITVPAYF NDSQRQATKDAGTIS GLNVLRIINEPTAAAI AYGLDKKVGAEARNVL IFDLGGGTFDVSILTIE DGIFEVKSTAGDTHL GGEDFDNRMVNHFI TEFKRKHKKDISDNKR AVRRLRTACERAKRT LSSSTQASIEIDSLYEGI DFYTSITRARFEELNA DLFRGTLDPVEKALR DAKMDKAQIHDIIVLV GGSTRIPKIQLLQDY FNGKELNKSINPDEA VAYGAAVQAAILSGD KSENVQDLLLLDVTP SLGIETAGGVMTVLK RNTTIPTKQTQFTTY SDNQPGVLIQVYEGE
NP_001103873.1	DrHSPA8	D. rerio	HSC70	
				SCVGVFQHGKVEIA NDQGNRTTPSYVAF DTERLIGDAAKNQVA MNPTNTVFDKRLIG RRFDDAVVQSDMKH WPFMVVNDAGRPK VQVEYKGETKSFYPEE VSSMVLTKMKEIAEA YLGKTVTNNAVTVPA YFNDSQRQATKDAG TIAGLNVLRINEPTA AAIAYGLDKKVGAE NVLIFDLGGGTFDVS LTIEDGIFEVKSTAGD THLGGEDFDNRMVN HFIAEFKRKHKKDISE NKRAVRRLRTACERA KRTLSSSTQASIEIDSL YEGIDFYTSITRARFEE LNADLFRGTLDPVEK ALRDAKLDSQIHDIIV LVGGSTRIPKIQLLQ DFFNGKELNKSINPD EAVAYGAAVQAAILS GDKSENVQDLLLLDV TPLSLGIETAGGVMT VLIKRNTTIPTKQTQT FTTYSNQPGVLIQV
NP_006588.1	HsHSC70	H. sapiens	HSC70	
				SCVGVFQHGKVEIA NDQGNRTTPSYVAF DTERLIGDAAKNQVA MNPTNTVFDKRLIG RRFDDAVVQSDMKH WPFMVVNDAGRPK VQVEYKGETKSFYPEE VSSMVLTKMKEIAEA YLGKTVTNNAVTVPA YFNDSQRQATKDAG TIAGLNVLRINEPTA AAIAYGLDKKVGAE NVLIFDLGGGTFDVS LTIEDGIFEVKSTAGD THLGGEDFDNRMVN HFIAEFKRKHKKDISE NKRAVRRLRTACERA KRTLSSSTQASIEIDSL YEGIDFYTSITRARFEE LNADLFRGTLDPVEK ALRDAKLDSQIHDIIV LVGGSTRIPKIQLLQ DFFNGKELNKSINPD EAVAYGAAVQAAILS GDKSENVQDLLLLDV TPLSLGIETAGGVMT VLIKRNTTIPTKQTQT FTTYSNQPGVLIQV
NP_112442.2	MmHSC70	M. musculus	HSC70	

XP_024379975.1	PpatHSC70	P. patens	HSC70	TTYSCGVWQHDRV EIIANDQGNRTTPSYV AFTDTERLIGDAAKN QVAMNPTNTVFDKAK RLIGRRFSDASVQSD MKLWPFKITPGPEGK PMISVQYKGEEKTFA AEEVSSMVLKMKKEIA EAYLGSTVKNAVTV PAYFNDSQRQATKD AGVIAGLNVLRINEP TAAAIAYGLDKKATSV GEKNVLIFDLGGGTF DVSILTIEEGIFEVKAT AGDTHLGGEDFDNR MVNHFVQEFKRKYK KDISSNPALRRLRTA AERAKRTLSSTAQTIT EIDSLYEGVDFYSTITR ARFEELNMDMFRKC MEPVEKSLRDAKMD KSSIHVDVVLGGGSTRI PKVQQLQDFFNGKE LCKSINPDEAVAYGA AVQAAILSGEGNEKV QDLLLLDVTPLSLGL TAGGVMTVLIARNTT IPTKKEQVFSTYSDNQ
-	PpHSC70	P. polycephala	HSC70	SCVGVWQNDRVEIIA NDQGNRTTPSYVAFT DQERLIGDAAKNQV ALNPENTVFDKRLI GRRFSDPVCQDDMK HWPFKVIAKEGDKPH IRVTFKGEQKDFSPEE ISAMVLKMKETAFAF LGATVTNAVITVPAYF NDSQRQATKDAGTIS GLNVLRINEPTAAAI AYGLDKTAEAEKNVLI FDLGGGTFDVSLLTIE SGVFEVKSTGGNTHL GGEDFDNRLVNHFM EEFKRKFKDITGNQR AIRRLRTACERAKRTL SSSTQASIEIDSLHEGV DFYTNITRARFEELCID LFRGCLEPVEKVLRLDA KLSKSQVHEIVLVGGS TRIPKVQQLQDFFN GKELNKSINPDEAVA YGAAVQAAILTGEGG KKVQDLLLLDVPTLTL GIETAGGVMTSUISR NSTIPCKKAQVFSTYS DNQPGVLQVYEGER
NP_009396.2	ScSSA1	S. cerevisiae	HSC70	VAHFANDRVDIIAND QGNRTTPSFVAFTDT ERLIGDAAKNQAAAM NPSNTVFDKRLIGR NFNDPVEVQADMKHF PFKLIDVDGKPKIQVE FKGETKNFTPEQISS MVLGKMKETAESYLG AKVNDNAVTVPAYF NDSQRQATKDAGTIA GLNVLRINEPTAAAI AYGLDKKGKEEHLIF DLGGGTFDVSLLSIED GIFEVKATAGDTHLG GEDFDNRLVNHFIQE FKRNKKDLSTNQRA LRLRLRTACERAKRTL SSAQTSVEIDSLFEGID FYTSITRARFEELCADL FRSTLDPVEKVLRLDAK LDKSQVDEIVLVGGST RIPKVQKLVTDYFNG KEPNRSINPDEAVAY GAAVQAAILTGDESS KTQDLLLLDVAPLSLG IETAGGVMTKLIPRNS TIPTKKSEIFSTYADN QPGVLQVFEGERAK

				VAHFSNDRVDIAND QGNRTTPSVFGTDT ERLIGDAAKNQAAM NPANTVFDKRLIGR NFNDPEVQGDMMHF PFKLDVDGKPKIQVE FKGETKNFTPEQISS MVLGKMKETAESYLG AKVNDVAVTVPAYF NDSQRQATKDAGTIA GLNVLRIINEPTAAAI AYGLDKKGKEEHLIF DLGGGTFDVSLLSIED GIFEVKATAGDTHLG GEDFDNRLVNHFIQE FKRKNKKDLSTNQRA LRLRTACERAKRTL SSAQTSEIDSLFEGID FYTSITRARFEELCADL FRSTLDPVEKVLDAK LDKSQVDEIVLVGGST RIPKVQLVTDYFNG KEPNRSINPDEAVAY GAAVQAAILTGDESS KTQDLLLLDVAPLSLG IETAGGVMTKLIPRNS TIPTKKSEVFSTYADN QPGVLIQVFEGERA
NP_013076.1	ScSSA2	S. cerevisiae	HSC70	VAHFSNDRVEIAND QGNRTTPSVYAFDT ERLIGDAAKNQAAM PHNTVFDKRLIGRK FDDPEVTDDAKHFPF KVISRDGKPVVQVEY KGETKTFTPEEISSMV LSKMKETAENYLGTT VNDVAVTVPAYFNDS QRQATKDAGTIAGM NVLRIINEPTAAAIAY GLDKKGRAEHNVLIF DLGGGTFDVSLLSIDE GVFEVKATAGDTHLG GEDFDNRLVNHILATE FKRKTKKDISNNQRS LRLRTAAERAKRALSS SSQTSIEIDSLFEGMD FYTSITRARFEELCAD LFRSTLEPVEKVLKDS KLDKSQIDEIVLVGGS TRIPKIQKLVSDFFNG KEPNRSINPDEAVAY GAAVQAAILTGQDST KTQDLLLLDVAPLSLG IETAGGIMTKLIPRNS TIPTKKSETFSTYADN QPGVLIQVFEGERT
NP_009478.1	ScSSA3	S. cerevisiae	HSC70	VAHFANDRVEIAND QGNRTTPSVYAFDT ERLIGDAAKNQAAM NPHNTVFDKRLIGR KFDDPEVTNDAKHYP FKVIDGGKPVVQVE YKGETKTFTPEEISSMI LTKMKETAENFLGTE VKDAVTVPAYFNDS QRQATKDAGTIAGLN VLRINEPTAAAIAYGL DKKSQKEHNVLIFDL GGGTFDVSLLSIDEV FEVKATAGDTHLGGE DFDSRLVNFLAEFKR KNKKDLTTNQRSLRR LRTAAERAKRTLSSSA QTSIEIDSLFEGIDFY SITRARFEELCADLFRS TLEPVEKVLADSKLDK SQIDEIVLVGGSTRIPK VQKLVSDFFNGKEPN RSINPDEAVAYGAAV QAAILTGQDQSTTQD LLLLDVAPLSLGIETAG GIMTKLIPRNSTIPTK KSEVFSTYADNQPGV LIQVFEGERTRTKDN
NP_011029.3	ScSSA4	S. cerevisiae	HSC70	

XP_001012263.1	TtHSC70	T. thermophi	HSC70	LGTTYSCVGVFINDRV EIIANDQGNRTTPSYV AFTETERLIGDAAKN QVARNPTNTVFDAK RLIGRKFNETVVQKDI KLWPFKVEAGPDDK PKIVVKHKGEVKKFH AEEISSMVLVKMREIA EAFLTKQIKNAVITVP AYFNDSQRQATKDA GAIALNLVRIINEPT AAAIAYGLDKKGQGE KNLVIFDLGGGTFDVS LLTLDDGIFEVKATAG DTHLGGEDFDNKLVE FCAADFLKKKNIDIRE NPRAMRRRLTQCER AKRILSSAQATIEVD ALADSEDFMMVISRP KFEELCLSMFKECIPP VEKVLKDSGMAKNQ VHEVVLVGGSTRIPK VIQLITEFFNGKEPNR SINPDEAVAYGAAIQ AAILTGSSEHIQDVL LLDVTPLSMGIETAG QVMTVLIPRNTTIPTK KSQVFTTYADNQPQGV
NP_001079632.1	XLHSAP8.L	X. laevis	HSC70	SCVGVFQHGKVEIIA NDQGNRTTPSYVAFT DTERLIGDAAKNQVA MNPTNTVFDAKRLIG RRFEDAVVQSDMKH WPFTVVSDDGGRPKV QVEYKAETKSFPPEEV SSMVLTKMKEIAEAY LGKTVTNVAVTVPAY FNDSQRQATKDAGTI SGLNLVRIINEPTAAAI AYGLDKKVGVERNVL IFDLGGGTFDVSILTIE DGIFEVKSTAGDTHL GGEDFDNRMVNHFV AEFKRKHKKDIIDNKR AVRRRLTACERAKRT LSSSTQASIEISLYEGI DFYTSITRARFEELNA DLFRGTLDPVEKSLRD AKLDSQIHDIVLVGG STRIPKIQLLQDFFN GKELNKSINPDEAVA YGAAVQAAILSGDKS ENVQDLLLDVTPLSL GIETAGGVMITVUKR NTTIPTKQTFTTYS DNQPGVLQVYEGER
NP_001148198.1	ZmHSC70	Z. mays	HSC70	GTTYSVGVWQHDR VEIANDQGNRTTPSY VAFTDSERLVGDAAK NQVAMNPINTVFDA KRLIGRRFSDASVQSD AKLWPFKVIPGSGDK PMIGVQFRGEEKQFS AEEISSMILNKMKETA EAYLGTTIKNAVITVP AYFNDSQRQATKDA GVISGLNVMRIINEPT AAAIAYGLDKKSSSD GEKNVLIFDLGGGTF DVSLTIEEGIFEVKAT AGDTHLGGEDFDNR LVNHFVQEFKRKNKK DITGNPRALRRLRTAC ERAKRTLSSTAQTITIEI DSLYEGIDFYTTITRAR FEELNMDLFRKCMEP VEKCLRDAKMDKSSV HDVVVLVGGSTRIPRV QQLLQDFNKGELCK SINPDEAVAYGAAVQ AAILTGEKNEKVQDL LLLDVSPSLGLETAG GVMITVLIPRNTTIPTK KEQVFSTYSDNQPQGV

				TMIEPKKLNKRKREPT AIENLTSEEKESQISSL NLEMKGFLDYFREY MDKSKRTDLFSGFSE CSSLNSMVALLMEE MSLPLSKLVDEIYLK EKTESVTMVAVKS VSVGQVRVSYGVLNV DADVLEDDSESCLWC WETRDLMPSVVRG VLKLRTCRKKIHERIT AVSAMLAAQREETE KLWRSDLSKAAEKL KILSEVDIRSFMNDM MQKNSSEMAEKDSK REEKLKLEKNRCE AEKEKKRMRQVLKE KLQKEKEQLLQKAI VDENNKEEETESRK RIKKQQDESEKEQKR REKEQAEKKQLQVQ KQASIMERFLKKS SLTQPKLPSSSEVTA LSCTKHENEIGKVQ AIDNAFSTTCEATVD DIRREHFASWRQLGH LLSSSKHWGMRRQ PKSELFPKLKLSTNSG
NP_176725.1	AtFAS1	A thaliana	CAF1A	
				RSPKTRMIEHENPRV SPVAVMAIESHGQP MEKSGVSCDMDCRG LVSPIQSQDLPKRKR RRSLEGNMSQEEKEL KIAAFREEIQGLRYY DEFMSKTDLGVDLN GSMDSVIACIEESEL PLKKLVDSVLKSLKR GGEGANITLLSVQNSI HRVGIRNKYGIENPN ADLMEDESKPCLWC WETRELKLLPKAHHG ILKIRRTGRRRIRERIA AVSEMISSLSAQENE NDYKAKIMEASGKLR KILNLERIRDLVSNLLQ KIEAGREVIGEEQKL LDRELKNMGKELKRQ MEEAEREERQREKEL KRQQDEADKEQRRR EREEAELKKQLKRQQ EDAAEQRRHEKEEA ESKKQLKKQEEAEK DKRRREKEEAERKKQ LSVQKQAAMMKRFF KSNSNVNISPTKVSSP RSPHKNEQISNAVILA
XP_006858774.1	AtrFAS1	A. trichopoda	CAF1A	
				LIETNENSCDPNQKG VKRVAQTPPEKDAKK VKLSRDSSNQSVIELS DTSPQKETSQQSPKA PKTPKTPKTPKVSREE REKQKREKMEEREKQ RIERERILEEKRLEKDK LAEEKRLDKEKKEKER LDKKLEEDKKKEEKRK EAEKKKKDEEEKMK KDEERNKRKKEEEKK EAKRREDEEKKEAKR KEEEAIEERKRQSAL FAKFFSKVEKKVAP QKESSNWLPFERKD GMSLAKIISRDPLPDD ADIFTQHEEINSIASFI ESAAKIVPIESSKARN MKAKLFQFHGNRRP QYYGTWRKSKIVSG SCPLAEIEGIDYNVVS DDEWEDEPSDGEEC NSDDDAEKDNDDDD GGEDDGFVPPCYL SDGEGDEDSTSDNDI AGDKKKEQPKRITID SDDDENSTDAERKA RLAQRAEDWAKRTG
NP_492440.1	CeChaf-1	C. elegans	CAF1A	

				IKDTEMKEETPNELPK PIPTSTTTTTTTTTTP TSTTTTTSTPKRRRLT EEEEQKREDEKKKREE DRIKREEDKKKDEEK KKLTEEREIEKKRDEE KKKLLLEEKINRKKREE EAEKKRIEKKKKEEIE KKKEEIERKKGEAERL KEEEQKKKLTSSFFS PIEAPIKPKSNSFIQP IELPSNTQLYQYQLPT YSNSYEVFKESISSN DTIKLETNQKKRDYFS KLNRGTOHGELGQR NHRFKLKIDSLPGAIM HHDLVSRLSALKLLKF HDSFRPSYYGTFSKTS KQJTAKNPFKDLTID YDYDSSDEWEEKDD DQAEKITSADEKEED QDADEEENEEDRAW IDDIGLNDDDDTQVD NSTNISVNERVFSRK NKKRFEKKPIILEPY FNQYPTIINLVHNK NNNNNNNNNNNN NNNNNNNNNNNN
XP_637076.1	DdCAF1A	D. discoideu	CAF1A	
				RSQLDGLFRHRDEEA DDEVEEEQNIVRRK PPNRANLYSELSSGR RKPLKMQRDVKLQR RTKDEEDDDVQVID YLSAPAGLIEVEQPKQ LTRMKAKYLFHADNR RPPYYGTWRKSSSIS ARRPLAQDKVLDYE VDSDCWEEEEPGES LSASEKEKESEEESE EEYNEWYVPHGHS DEELQNDGDGMEDG HTREAQKAKLQVLQ QEFAQEMKKQTKKIK ARLLGPVWLDENG KSELFPAIFAHTIDMY ACWQVEPLSEPPPE PERQDQTPQVPVLQ LDDRMLQQLVRLTH GNRNSKTFLINEYLEY LKTQATAEANQTTL SKAVLREFDELASW KTVELSTPEAAAASA NSAKKGRKPKKRLC WVVASDMLKQLP DLQLQNQWNYTLTP KVSDGPDAQQEQS
NP_572495.1	Dm-p180	D. melanog	CAF1A	
				TPRRDMDCVGKANT NKKLVQARLPFKRLN PEPKECNEPKRTKGP VAPKCEPSDQENDQ DSSSISHGPA LVNG RGPLDCFMSRRKRSP LRSAPeATIDLTEDSN DSAKQQPAPPIAATC PLSEEKTKTSEGTTPT IPLTEETEKDEAEDV DALPLLDITQSDTEE EEEEEEEEEQQAE VSHGNESVLSTGSTSS ASVIASSPEPSKSAPT TPASTSRINAANKVK RRSLKSVQEEKQQR QRDEKERLKQEAkaa KEKKKEEARKMKEEK EREKKKEKDEKERR EKKERDEKEKADKLK AKEEQRMKIEAKLE EKRRKEEKRLKEEKD RIKAEKAEITRFLQKP KTQLAPKTLASACGK FAPFEIKAHMSLAPLT RVQCEDSVLEDLDY LAQPDSTLNLKDWIT GHKPRSSGPTRRHS
NP_001038478.2	DrChaf1-A	D. rerio	CAF1A	

NP_005474.2	Hs-p150	H. sapiens	CAF1A	AATAMDCCKDRPAFP VKKLIQARLPFKRLNL VPKGKADDMDDQ GTSVQSKSPDLEASLD TLENNCHVGSDIDFR PKLVNGKGPLDNFLR NRIETSIGQSTVIIDLT EDSNEQPDLSVDHNL LNSEASPSREAINQ REDTGDQQGLLKAIQ NDKLAFPGETLSDIPC KTEEEGVGCGGAGR RGDSQECSPRSCPILT SGPRMCPRKEQDSW SEAGILFKGVPMV VLQDILAVRPPQIKSL PATPQGKNMTPESE VLESFPEEDSVLSHSSL SSPSSTSSPEGPPAPP KQHSSTSPFPTSTPLR RITKKFVKGSTEKNKL RLQRDQERLGKQLKL RAEREKEKLKEEAKR AKEEAKKKKEEEKELK EKERREKREKDEKEKA EKQRLKEERRKERQE ALEAKLEEKRRKKEEK RLREEEKRIKAEAEIT
NP_038761.1	Mm-p150	M. musculus	CAF1A	DCKDRPGFPVKRLIQ ARLPFKRLNLVPEKVV EEDTSPKAAVESKVP DLQLSLGTGFESQCHT GSHVGLSTKLVGQGQ PIDSFLRATIKPVPSVV IIDLTENCSDIPDSPEG HSELSPDTAGVTTV EGAAKQEQHSAALC LLETSPDITCHMEEEP GSPGDPKRTGDCQA GSLQSCPELTGSRCT PTKELSSWSKAGDLLF IEKVPVVVLEDILATKP SIASLPMMSLDRSVT SESEILESCPEDDSILS HSSSTNSSPTSSPEGP STPPEHRGGRSSPSTP ACRVAKNFVKGSTEK GRSKLHRDREQREE KEKLREEIRRAKEEAR KKKEEEKELKEKERRE KREKDEKEKAEKQRL KEEKRRKERQALEAKL EEKRKKEEKRLREEE KRLREEEKRIKAEAEI TRFFQKPKTPQAPKT LAGSCGKFAPFEIKEH
XP_024369555.1	PpaFAS1	P. patens	CAF1A	PAEREKLRKREKAERKA EKLRIRAEKEAARAER ELKKQQEEAEKEQRR KDRELKKNNEEEKKQ LDDVEKEQRRKEKEAI EAEKELKRKERLRRQ QEEVEKEQKRKEKEV AEKKRQLLLQKQATI MDRLFRRKESITPVA HEKQSPAQGVSSPSP AHQTPEGISTSRTEPS TVEMVMRLDEGLKV PCTLTEEEILRGHVSS WQEKWKRRSLPFR RWGVVRMAPKVTVVR ELRLQGASASFPRRVE SAGNAVSSSTLEAQA PSLKRSREDFESIRDC KVEDDCILLNEDPIDS PPRKKQCNTRWKLL QFDKSHRPAYYGTF KLSSTVGPRHPLRKDP SLDYEVDSDDEEWEE DPGESLSDCEDKEED VEKVDSELEDAADG FVVPDGYLSENEG LEETDVEDTEKAAEVK APCESGTGAECMVS

-	PpCAF1A	P. polyceph	CAF1A	VFMDAWLTKRVNT PQNDVNNAPAPITD NMDIDTTPKQEDTKI KSPSKTKSPSKPPPPD LQNCVLLSPCPTFFP TQNTPPPHNIPNG NTLQDKPILVDP5VN QSHAPNPAVDAAAT DQPAKPSPKRRLSP EEKQAREDAKAKEE ERKLQAKQREDAKNK KEEEKRLLALAREEEK KKRDEEKRVKDQAKE EARLKKEAEKKARDD AKLQKDLEKKAKEEA RLQKLEKQKQEEEE KRKKERQQAVALANFI KVAPVEAPRPQKVD PDLLIQPFVCGSEVF PPSYLTRQSPGFD DAIAHKDGGITIEKLQ VEFREHMHGKRNRK RDKKRAEKLQGLPP ALSNMKLLRFHDNVR PAYYGTFSKRSVLTG RRPLRTDDAVFDYEV DSDAEWEDEGEHEEL SESEGEKDKEEESEE
NP_015343.1	ScCac1	S. cerevisiae	CAF1A	KKGILSFFQNTTTVKS NKFLTKEKDVITLDDP KEDVSGPMIETVKQE TMKSINKECADEMKT TPKKANAEDKLLCYK NSPIQSTKYDRNTNK QVPNGNIIAIEKSR SSPCSKRELSSKKKEEA KREKELKKQRAEEK HRKELLRQEEKKKEL KVEEERQRRRELKKQ KEEEKRRKEEARLEAK RRKEEERLKEEEEIRLK EEAKERAQSRIGNFFK KLSDSNTPVVEKSDYE KFFLPFYAKDGVRRVS NKWKLTKVELEGSKR KIDDELNSKDKTSSD DLLNWLQSRRLPRGH KIKRKAVDVLQQMPL KEKTDELQSLAQV PHKYIKFYENVRPPFI GTYSMDFTLPPNDPF STKGTGFNYDYSDV EWWNEEEGEVDNLE SGEEEEEDDEDVPS EGEFDFGLDSENSDL DGLPCAARKFVGPLIP
XP_001021065.2	TtCAF1A	T. thermophi	CAF1A	NSROMNIECQGDSP AQTLNQSQNNVQQS AVEVNSQNVQENIE QAKVLNNQNGQS NISEQYWQSKKNGFI QELNKLQIDLSEQI MFAMPQEILSQKIQD YFSAINDSSTEQIAD LQRKLIKESCLWIVEG KNCNDFDLIKLCTVLQ EYTKQCLAQISDISQK DFDFYRDIIGEWAEK KNYGLKLTQDIQTL KQQNEGNLSDQQVS ALKIGLHEIWEVKEIQ CLSNNTKIEELDYIRKK RRQHQALLRGLEKLIK AIEKGKGEDEAFSAY QKTMNPPKMKVQ TKINGKSN5QKQNNQ VDKSIEDNLKKMKIDS GKVTPIDVGEMQNE TNQQENSQLIQQA PKKKTAKKSNKKA GSDQQIEQNEDTQK QSICSSVNNNSITQE QSKQISSKKQNESK KSGNQQQQASANT

NP_001082096.1	XI-p150	X. laevis	CAF1A	STKSNTKKMVQARLP FKRLNPVPKDEGCLEE KKVRLTKNVSPQKML HSLNSSMEDMENDC DMETETVPIPKAVNG KGPLDNYIRKAPKVS HAPSITITDLTEESNISI SNDCLPLNGESETHLA NGTLALEESTPNPLS AKEECTVSLNKTVEN THFSELKSDQLHQAA ATSTSASNFSPERVVK EDCNSSADDDASVS SSSSPVLSLSPDAQTG SQFRNRSSPSTTTPT GKVTANKTSADKNKT KDKDKQRQAEKEERE RAKKEARSAKKKRQ GLLKNLQKRKGTSES SGKEYKKEKEREKKE KAEKMKLKEEKREK LEALEAKQEEKRKDE EKRQKEEKRQKEEE KRLKEEERVKAEKAE ITRFFQPKTPQAPKT FSRSCGFAPFEIKKG MALAPLCRIDFEPEAS EELDRFLQEQNSKIYF
XP_008656983.1	ZmFAS1A-like	Z. mays	CAF1A	EAPDRNDAEPRRDM GQSQMQVGGPVVL NRSaelGSGDFMAID DAPVQAPSGQPAAA KHQSPAALTDTIMEV QKQLKRRASNNPVI AAADKDALAAGCRQ ELEGLFQFYKEVSDRK VQLDGGNLSVNGMV GCLLESSLGLTKLVD GIYEKLGLDGVSVAS VRGSVLLVGQRMMY GKSSLDADVLEDDTE SALWCWETRDCLKM PVKARSVLSTRRSVRK KIHERTAIYSTLSVLE NPGVYVQVNDLKKA SLRLNKSINLEGIRSM VERVAQKNLERGVKV AGSAIKELQETEKND QNLRRLEDTSVSELQ NGNSPDNEKEVQKA QKQVEKEIKRQKEEE ARMKKLQKKQQEDV LREQKRREKEEAEEK KQKQKQEEELKEQK RREKEEAEMKKQHKK QQEEAEKEQKRLKE
Q9SXY1	AtFAS2	A thaliana	CAF1B	PVLTVDFHPISGLLAT AGADYDIKLWLINSG QAEKKVPSVSYQSSLT YHGCANTIRFSPSG ELLASGADGGELFIW KLHPSETNQSWKVH KLSLFHRKDVLDLQW SPDDAYLISGSVDNSC IIWVDVNGSVHQILD AHCHYVQGVAWDPL AKYVASLSSDRTCRIY ANKPQTKSGVEKMI NYVCQHVIKADQQ RGDETKTIKTLFHDE TLPFFRRLSWSPDGS FLIPAGSFKVSPTSEA VNATYVFSRKDLSP ALQLPGASKPVVVVR FCPVAFKLRGSSSEEG FFKLPYRLVFAIATLNS VYIYDTECVAPIAVLA GLHYAAITDITWSPN ASYLALSSQDGYCTLV EFEDKELGEAVSISVG KKPVDGEEKKHDLK GDELMTETTPDESKK QAELEQNEESKQPLP SKITTDGKEKEHIMQ

XP_020531827.1	AtrFAS2	A. trichopoda	CAF1B	QPVLTLDFHSLTNLLA TGGADKDKLWALV QSEAQKKAPTACYQT SLPPHHSTAVNVLRFS PSGEQLASGADGGEL AIWKMQSDDRHTW KVLKTLSFHRKDVLDL QWSTDGAFLLSGSVD NSCIIWDVNGGAVH QILDAHLHYVQGVA WDPSGQYLASLSDDR TCRIYSRKPQKNKAK GQENLNYYVCQHVV KGEFQNLGDTKLSSA RHHLFHDETLPSFFRR LAWSPDGSFLLAPAG IDKLSSASPTCNTAYV FSRNDLARPALQLPG ASKPIVAVRFCPISTL RGSNSAGFFKLPYRIIF AVATLNSLYIYDTSV APVAILAGLHYAAITD IAWSSDAKYLAUSSQ DGYCTILEFENDELGS PLSLTAATEAKIVAKT SEESPVPKTEHMQVD ENRVTIAVEEKARIQI SKVEDNEPKEKNVEIK
NP_490902.2	CeChaf-2	C. elegans	CAF1B	GLLSVDLHHETRNKG YRLATASVQKEVRVW EFEFEMGLDPKTQEN KPQLTVGFLANLVFH NHAINQVKFSPSKEH ELLASGDCEGRITWK LSDQPVPVPPQDEMP SNKENWIRYKVLNH NSDVNALCWDPSGT QLASVNDHTLAVHD ALTGKRFLVASNFRSP NGVCWDPGSGKIAT MSPDRKMDLIDAVR GARLKHFSALLPSM TIPANGDLHLETKIH KLFHDDQLFSQRAL GFSPNGEFIAAPCAH LELGSSDLYGTYFFRR EDLGVKEAPYTFYPAP RPTFLVKFSPVTFSLP STKENHLGLPYRLLWI ALNKDAIFYDSQHSY PVAVVDNIHLNALTD ASFSDHGRVLVSSLE GFCFVKINLTQWGE VMTIIVPCGSPNLIE EKKQKKRKSTVAVVIE KEKEKEEDAASPMR
XP_646299.1	DdCAF1B	D. discoideur	CAF1B	PIYSADFPTSNNKCT TGFONEIKIWGYTKN KDGHLSVEFLSSLAKH VKPVNIARFSPGGNL LASGSDDGSIWVRL NQNAIPPSDSSMKEI WSIVTVLRVTTDVVD LSWSCDGGQLSTVST DNSVSIWSPLSKTHH QLITEHSYVQGVSW DPLNDFMITQSSDGT CRIYRNEKKKKIITQP SSCPSSTPVSEKDKD SEDTDDNDNSNSND SNNNNTNSNSTGNN VVNSATIEDVNTNSR KKLNLIIISNVLSRRSYN NTSIPNEDIAEIEELQ QQQQQLDEAKIKEES LVSHRMFYDERASTF FRRPSWSPDGSIFITP TGKFRDPSPTSKYQST YIFSRHIRDRPLIPLS NNPTVVVKFNPIIFKL RPDNNTNNNNNNNN DDDSNNDNNNNNN NNNNNNNNNNNNT NNDNDNDNDTSDSDS

NP_610589.2	Dm-p105	D. melanogaster	CAF1B	VLSVDIQQNGLGLRS PTICRLASGGS DAHVL IWYVNRSDDAEGVD VELAADLSRHQRAVN AVRWSPNGELLASG DDES VFIWKQKAD HEVVNIVDADGCSEQ DKEVWLT LKVLRGHR EDIYDLSWAPNSQFL VSGSV DNTAMLW DV HSGKSLALDDHKGY VQGVAWDPCNQYIA TMSTD RQMRIFDAN TKRVLHRVSKCVLPV KEDHEMHGKSMRLY QDGT LQTFFRRLCFT PDGKLLTPSGITDYD GVV KPINTSYGFSRH DLSKPAFVLPFPKEYA VAVRCS PVLRLRPY NAEKNP IISLPYRMI YAVATKN AVFFYDTQ QPVPFAIVSNIHYSRL TDLAWSSDGTVLIVSS TDGYCSLTFEPTELG DCYEDMETVLSVVLK SSENATVLKKRQKLR KVS LDEPRKPLQE KSK
NP_001315058.1	DrChaf1B	D. rerio	CAF1B	PVYSLDFQQSGDGKT QRLATAGVDTTVRM WRVDKGP DGKAVVE FLSNLARHTKAVNVV RFSPTAEVLASGGDD AAILLWKLNDNKEPE QTPTFQEEEDAQLNK ESWSVVKTLRGHIED VYDISWTS DGNFMA SGSV DNTAIMWDVN KGQKMCIFNDHKS Y VQGVAWDPLGQYIS TLSCDRVMRVYSAH NRKKAYSISKMTSSAT ADGEVKNYRMFHDD SMRSFFRRLTFTP DGS FLAPAGC VEA GENV TNTTYVFSRKSFKRPI AHLPCPSKATLAVRC SPVYFELRTKRAEDDS LKP VANTFNL PYRLVF AVASEDSIFFYDTQQT LPFGYVSNIIHYHTLSD LSWSRDGSFLAVSST DGYCSFISFDEGELGT PLKEKPPLEMVTPSST NEKKGKRALANGRTA SPVPRQANTPAAQD
NP_005432.1	Hs-p60	H. sapiens	CAF1B	VYSLDFQHG TAGRIH RLASAGVD TNVRIWK VEKGP DGKAI VEF LSN LARHTKAVNVVRFSP TGEILASGGDDAVILL WKVNDNKEPEQIAF QDEDEAQLNKENWT VVKTLRGHLEDVYDIC WATDGNLMASASV DNTAIIW DVSKGQKI SIFNEHKS YVQGV TW DPLGQYVATLSCDRV LRVYSIQKRVA FNVS KMLSGIGAEGEARSY RMFHDDSMKSFRR LSFTPDGSL LTPAGC VESGENVMNTTYVFS RKNLKRPIAHLPCPGK ATLAVRCCPVYFELRP VVETGVELMSLPYRL VF AVASEDSVLLYDT QQSF PFGYVSNIIHYH TLSDISWSSDGAFLAI SSTDGYCSFVTFEKDE LGIPLKEKPV LNM RTP DTAKKTKSQTHRGS S PGPRPVEGTPASRTQ DPSSPGTTPPQARQA

				VYSLDFQHGATWKIH RLASAGVDTAVRIWK LERGPDGKAIVEFSN LARHTKAVNVVRFSP TGEILASGGDDAVILL WKMNDSKEPEQJAF QDEEEAQLNKENWT VVKTLRGHLEDVYDIC WATDGNLMTSASVD NTVIIWVDSKGQKISI FNEHKSIVQGVTDW PLGQYIATLSCDRVLR IYNTQKKRVAFNISK MLSGQGPEGEARSFR MFHDDSMKSFRRILS FTPDGSLLLTPAGCM ESGENVTNTTYVFSR KHLKRPIAHLPCPGK ATLAVRCCPVYFELRP VAETEKASEEPSPELV NLPYRMVFAVASEDS VLLYDTQQSFPGYVS NIHYHTLSDISWSSD GAFLAISSTDGYCTFV TFEKGELGIPLKEKPV SIRTPDTAKKAKNQT HQGSSPGSRSEVETP SNRTQDPSPSCTTPS
NP_082359.1	Mm-p60	M. musculus	CAF1B	
				PVLSLDFHPSGGLLAT AGADHDIKIWSVHQ KESGSPTVEFEADLCY HSKAVNVLRFSPSGQ LLASGGDGGELIWKP IERADKKPSWKVLKTL QLHVRDVLDLAWSP DSLMLMSGVDNQC MIWDVVSGKVLQVL NDHQHFVQGVAVWD PAGEFTSISSDRTCRI YSRQSVPSKSKRKLAV VQNMFTCKQVLAKS DALPSKTNITIDADGR VAKASQLFHDETLPSF FRRLAWSPDSSFLVV PSGLHRIAHEAPSCN VSFIFSRDLRSPCIHL PAPTKPVAVRFCPI FSLRENETEKSSSGF ALPYRLVFAVATLDSL YVYDTQRSEPIVIFAGI HYAAITDIAWSADAQ YVAVSSQDGYCSLLIF HGNELGTKLSSGDVP ANIAKFLPENLLGKPQ ESLPNSTELSQVSEPV GLMNEEQSSPIVICPR
XP_024390997.1	PpaFAS2-like	P. patens	CAF1B	
				PVFSVDFHSSGRLASS GADNDIKIWKIGKRA DGSVSLDFLSLTRHT KAVNVVRFSPDGNFL ATGSDDGYIYIWKLN TSVPVLPDSIDKEVW SVAYTLRGKAQDVYD LAWSPDSSQIIVGSTD NSASVWSVKGQCK HHFEHHNHVYQGVA WDPLGIYLLTFSSDRT LRAYTRKLPKNKKSHK ALPPPTPSSASASTS SSSSDPLLVPDPLSS PSSVPVAPAQNPPAA TPSKKRKVKWIVAQ LVNKNREFGTQGLLE KEKEKEEESGKGKEK EKEKDCLQHKLFLEDES VNTFFRRLTWSPEGS LLFLPTGIHKASPASP PTNVCYAFSRNNLNR PLFYIPSGKPVIAVRCS PILYKLWNGKHPAQ KPLSTSPTPTSSAPS PTSSSTSPTPASSTPV TDAPSAPTSTPPATIP DTPMSDSPNTATPT
-	PpCAF1B	P. polycephala	CAF1B	

NP_013605.1	ScCac2	S. cerevisiae	CAF1B	PVYSLTFQKNSANDK LFTAGGDNKVRIWKL NRDENGQNGGVRKI ESLDFLGSALTHHEQAI NVIRFNSKGDVLASA GDDGGVLLWKQEDP NTQQESVVRPFGMD AETSEADENKEKVVV WKRLRGSGGATAAA EYDLAWSPDNRNIV VACMDNSIRLFDVGA GMLVCGQSDHGHYV QGVAVWDPLNQFILS QSADRSLHVGIVLS AGVVTGLKRSKIACA ELPCPGDVLRTNYLHH NETLPSFFRRCISPC GGLVVIPSGVYKVAG DEVANCVVYTRSGIL NSAGGVKNRPAIRIP SLKKPALMAAFSPVF YETCQKSVLKLPLYLV FAIATTNEVLVYDQDV LEPLCVVGNHYSPT DLAWSEDGSTLIISST DGFCSYVSIDTETQFG SRIEPPAMHAEPLDT DESAAAKNQREAG
XP_001020614.1	TtFAS2	T. thermophi	CAF1B	QRIVTLDVHFPFINLLV TGGSDDEEVYEGEDLEF EEEIGYKLVQINENY TPGQGGDAKPVTFIR ALKKHNSNPVNCVRF PNGQYLASASDDHKKI VIWHVDKGFQNIISTN QKIFNLVKNILEGHN REVCDLRWFNDSTHL ISGGMDYRAYIWNV KEGVIKQTVGAHKSY VQGVAVDPKMKFCL TLGNDRTVKVWRKLK SNNKKNIYEYIPST MKRLPFADVNIEEEM LSDDDEKSSITGNQM NGANGAVNQAMTY GMFLSERQLNTFVKR PDWSPDGSFLLPAA IVQEKRSKIEMCVYL YRRNVLNKPSLIINTN NKPAICTRFQKLFKK KEENQFSMVDIPYV FAISTIDNVMIYSTASL SPLAVVGNHIFALIND LTFFSNQSLIICSSDG MCSFVFEEENDLGKP LNSEEIQDEEIRNVN
NP_001090195.1	Xl-p60	X. laevis	CAF1B	VYSLDFQHGNKFNRR LASAGVDTAVRLWKL ERSPEGKAVVEFLASL ARHTKAVNVVRFSPN GEILASGGDDAAILL WKLNETKECEPAPFQ EQDEFELNKENWTVF KTLRGHLEDVYDICW TQDSNFMVSASVDN TAITWVDVTKGQKLSIF NEHKSYYQGVTVWDP LGQYIATLSCDRIMRV YKTETKRVAYNVSKM TSSPGPDGEAKSFRM YHDDSMKSFFRRLTF TPDGSLLTPAGCVEA GETVINTTYVFGKKNL RRPIAHLPCPTKATLA VRCCPVYFELRPTMK DDSGESRPQGLITQP YRMVFAVASEDAVL YDTQQLFPFGVVSNN HYHTLSDISWSSDGN FLAISSTDGYCSFVTFE EGELGVPLKEKPVALI PKTPATERKIKKAQSK KVSSPGSRQELTPSNK SMEHSAPSTPMNSR

NP_001132835.1	ZmFAS2	Z. mays	CAF1B	QPVLSLDFHPVSRRL ATAGADHDVFKIWEV ASDGKLPATATFKTGLV PNTTAHSSAVNVLRF SPSGEYLASGADGGG ITLWKLHPADDGEA WKIHKTLFFHHKDVL DLQWSHDSTFLVSAS VDNTCIIWEASKGTV QQKLEGLHYVQGV AWDPLGQYIASLSSD RTCKIYAKKPQGSK NAERLNFVCQHTLVK VEYQNHDESKPPTKS HLFHDETLPSFFRRLA WSPDGSFLVLPAGLS KHSSEVINTAYVMSR RDLSPRSIQLPGASKA IVAVRFCPLFKPRGS NSDGFCLKPYRVVFA VATLNSLYVYDTEVP PILVHAGLHYAAITDI AWSSDAKYLAUSSRD GYCTIIIEFENEELGQP HILPGSKEVANGNLT PETKKPVCADSMKVD NSANKLKMEASPA VGVRAPLLPTENIKRT
NP_200631.1	AtMSI1	A thaliana	CAF1C	LINEEYKIWKNTPEL YDLVITHALEWPSLTV EWLPDREEPSGKDYS VQKMILGTHTSESEP NYLMILAQVQLPLDDT ESEARQYDDDRSEFG GFGCATGKVQIIQQI NHDGEVNRARYMP QNPFIATKTVNAEVY VFDYSKHPSKPPLDG ACNPDILKRGHSEEG YGLSWSKFKQGHLLS GSDDAQICLWDINAT PKNKSLDAQQIFKAH EGVVEDVAWHLRHE YLFGSVGGDDQYLLIW DLRSPSASKPVQSVV AHSMEVNCLAFNPF NEWVATGSTDKTV KLFDLRKLSTALHTFD SHKEEVFQVGWNP NETILASCCLGRRLM VWDLRIDEEQTVED AEDGPPELLFIHGGH TSKISDFSWNPCEDW VISSVAEDNLIQIWQ MAENIYHDEDDAPG EPPSKAS
XP_006826731.1	AtrMSI4	A. trichopoda	CAF1C	RYTQWKSLLVPVLYD WLANHNLVWPSLSC RWGPQLEQATYKNR QRLYLSEQTDGSPVN TLVIANCEVVKPRVA AAEHISQFNEEARSPP VKKFKTIIHPGEVNRI RELQNSKIVATHTDS PDVLIWDVEAQPNR HAILGASDSRPDLILT GHKDNAEFALAMCP TDPYVLSGGKDKSVV LWSIEDHISAAAAEA PSTKSPGVPSSGSKS NTKTVGPSDKILDSPA IGPRGIYKGHSDTVED VQFCPSAQEFCSVG DSDCLLWDARIGVD PAIKVEKAHNADLHC VDWNPHDENLITGS ADNSVRMFDRRLT SNGIGSPVYKFEGHK AAVLCVQWSPDKAS VFGSSAEDGYLNVW DHEKVSKKENVGNK MPNSPAGLFFQHAG HRDKVVDHWNVSD PWTIVSVSDGESTG

NP_492552.1	CeLiN-53	C. elegans	CAF1C	MATLEDGTSEDRVA NDEYKIWKKNTPFLY DLVMTHALEWPSLSV QWLPDVAKDNSDHT IHRILGHTSDEQNH LLSKICMPTDDAQFD ASRYDTERSEYGGFG AVNGKVEPDIRINHE GEVNRARYMPQKSN IIATKSPHADVYIFDYL KHSAPRDNFTFNPLI RLKGHTKEGYGLSW NPNKEGLILSASDDQ TVCHWDINANQNVA GELQAKDVFGHESV VEDVAWHVLHDGVF GSVGDDKKLIWDVR TSTPGHCIDAHSAEV NCLAFNPYSEFILATG SADKTVALWDLRNLRL MKLHSFESHREIFQ VQWSPHNETILASSG TDKRLHVWDLKIGE DQSAEDAEDGPPELL FIHGGHTAKISDFSW NPNPWPVVCSVSED NILQVWQMAADNIYN EVDEETPADVVERQQ
XP_640105.1	DdCAF1C	D. discoideur	CAF1C	EEYKIWKRHPTFLYD MVITHALEWPSLTVA WLPVKTSPNKPYSI EKVILGHTSDEEQNY LMVAKVHLPVDEASI ESLKYDDTKGEVGGI GNVSEKIEIQINHE GEVNRARVMPQNHS IIATKTSSEVYIFDTT KHPLPTPDGKCSPN LKLTHGKKEGYGISW NPRKEGHLSCSDDQ SICMWDISAASKSDS TLDALNIYNGHTSIVE DVAWHYIHDTFGGSV GDDKKLMIWDRTRG TKPIHVVEAHNSEVN CLSFPNPFCEFLVATGS TDKTVLWDMRNLG NRLHSLISHTDEVFQV QFSPHNETVLASCGS DRRVNVWDLRIGEE QNNEDAADGPPELLF IHGGHTSKISDFSWN PNDPWSIASVAEDNI LQIWQMAENIYNDR EDDLENSKVTNAQIG E
NP_524354.1	Dm-p55	D. melanogas	CAF1C	AVEERVINEEYKIWKK NTPFLYDLVMTHALE WPSLTAQWLPDVTK QDQKDYSHRLILGT HTSDEQNHLLIASVQ LPSEDAQFDGSHYDN EKGEFGFGSVCGKI EIEIKINHEGEVNRAR YMPQNACVIATKTPS SDVLVFDYTKHPSKPE PSGECQPDRLRLRGHQ KEGYGLSWNPNLNG YLLSASDDHTICLWDI NATPKHRVIDAKNIF TGHTAVVEDVAWHL LHESLFGSVADDQKL MIWDRNNTSKPS HTVDAHTAEVNCLSF NPYSEFILATGSADKT VALWDLRNLKLLHS FESHKDEIFQVQWSP HNETILASSGDRRLH VWDLKIGEEQSTED AEDGPPELLFIHGGH TAKISDFSWNPNEP WIICSVSEDNIMQV WQMAENVYNDEEP EIPASELETNTA

NP_997760.1	DrRBBP4	D. rerio	CAF1C	RVINEEYKIWKNTPF LYDLVMTHALEWPSL TAQWLPDVTRPEGK DFSVHRLVLGHTSD EQNHLVIASVQLPND DAQFDASHYDSEKGE FGFGSVSGKIEIEIKI NHEGEVNRARYMPQ NPCIIATKTPTSDDL FDYTKHPSKPDPSGE CTPDLRLRGHQKEGY GLSWNPNLRSCLLSA SDDHTICLWDISTVPK EGKIVDAKTIFTGHTA VVEDVSWHLLHESLF GSVADDQKLMWDT RSNNTSKPSHAVDAH TAEVNCLSFNPYSEFIL ATGSADKTVALWDL RNLKLLHSFESHKDE IFQVQWSPHNETILA SSGTDRLNVWDL SK IGEEQSPEDAEDGPP ELLFIHGGHTAKISDF SWNPNEPWVICSVS EDNIMQVWQMAEN IYNDEDPEGAADTEV QG
NP_001128727.1	Hs-p48	H. sapiens	CAF1C	RVINEEYKIWKNTPF LYDLVMTHALEWPSL TAQWLPDVTRPEGK DFSIHRLVLGHTSDE QNHHLVIASVQLPNDD AQFDASHYDSEKGEF GGFGSVSGKIEIEIKIN HEGEVNRARYMPQN PCIIATKTPTSDDL YTKHPSKPDPSGECN PDLRLRGHQKEGYGL SWNPNLSGHLASD DHTICLWDISAVPKE GKVVDKTIFTGHTA VVEDVSWHLLHESLF GSVADDQKLMWDT RSNNTSKPSHSVDAH TAEVNCLSFNPYSEFIL ATGSADKTVALWDL RNLKLLHSFESHKDE IFQVQWSPHNETILA SSGTDRLNVWDL SK IGEEQSPEDAEDGPP ELLFIHGGHTAKISDF SWNPNEPWVICSVS EDNIMQVWQMAEN IYNDEDPEGSVDPEG QGS
NP_033056.2	Mm-p48	M. musculus	CAF1C	RVINEEYKIWKNTPF LYDLVMTHALEWPSL TAQWLPDVTRPEGK DFSIHRLVLGHTSDE QNHHLVIASVQLPNDD AQFDASHYDSEKGEF GGFGSVSGKIEIEIKIN HEGEVNRARYMPQN PCIIATKTPTSDDL YTKHPSKPDPSGECN PDLRLRGHQKEGYGL SWNPNLSGHLASD DHTICLWDISAVPKE GKVVDKTIFTGHTA VVEDVSWHLLHESLF GSVADDQKLMWDT RSNNTSKPSHSVDAH TAEVNCLSFNPYSEFIL ATGSADKTVALWDL RNLKLLHSFESHKDE IFQVQWSPHNETILA SSGTDRLNVWDL SK IGEEQSPEDAEDGPP ELLFIHGGHTAKISDF SWNPNEPWVICSVS EDNIMQVWQMAEN IYNDEDPEGSVDPEG QGS

				LVPALYDWLTNHHL MWPSLSRWGPQLE AGSSKTRQLLFYSERT DGECPNTIVVAQCDI MKPRTAAAEQISQAF IPPCCMSQRPEFKEG GKSPHLKKLTIHPG EVNRIREIPQNSNILA THTDSPKVLWNTKT QPNRATTSAASESKP DLVLIGHTDNAEFAL NVSRTAPYVISGGKTP KSITPTAAGSKQSGTA GGAADTTNVYTRGIF KGHTDVTEDVQFRPS SMNEFCSVGDDSCLL LWDARTGYQPISKVV KAHNADLHCVDWN AHDENULTGSADNS VRLFDHRKILARGQAI PVEQFEGHSAAVLCV QWCPDRASVFGSCA EDGLLNVDYKVG KALDTTNLKQPVKVP PGLFFQHTGHRDKVV DFHWDSDPWTIVS VSEDANTPGGGTTL QIWRMIDFIYRPEDE
XP_001776478.1	Ppa-p48	P. patens	CAF1C	NEEYKIWKKNAPFLY DLVMTHALEWPSLTV HWLPDKKIQSNEFS VQRLVLGHTSDAEQ NYLMIKVVHPLDDA NIDPRKYDDQKGEVG GFGAVSEKIEIEQRIN HSGEVNRARAMPQ NPTIATKTVSDVFLF DYTRHSPKSPADGKC NPNAILKGHTKEGYGI DWNPSREGHLLSASD DKTICLWDIHDAKAN TTIEAKQIFRGHKEVV EDVAWHCSHADYFA SAGDDKKLMIWDVR TNADPLQVEAHAGE VNCVAFNPFCEFLA TGSADKTVALWDMR NLDAKLHSFVSHTDE VFQVQWSPHHETVL ASCGSDRRVNVWDL SRIGDEQTAEDAEDG PPELLFIHGGHTSKIS DFSWNYNPWWLAS VAEDNIIQVWQMAE NIYNDDDTDEAQPSN EGDLE
-	PpCAF1C	P. polycephala	CAF1C	MNQCAKDITHEASSI PIDLQERYSHWKKNT KLLYDYLTNSTKW SLTCQFFPDLDTTSDE HRILLSFTSSQKPEDE TIYISKISTLGHKWSSL NNFDMDEMEFKPEN STRFPKHLVNDISIFF PNGEENRARYLPQN PDIIAGASSDGAIIYFD RTKHGSTRIRQSKISH PFETKLFSGHGVQDV EAMDTSSADINEATS LAWNLQQEALLSSH SNGQVQVWDIKQYS HENPIIDLPLVSINS GTAVNDVTWMPHT DSLFAACTEGNAVSL DLRTKKEKLSNREK HDGGVNSCRFNKYN SLILASADSNRNLN WDIRNMNKSPIATM EHGTSVSTLEWSPNF DTVLATAGQEDGLVK LWDTSCETIFTHGG HMLGVNDISWDAH DPWLMCSVANDNS VHIWKPAGNLVGHS
NP_009754.1	ScCac3	S. cerevisiae	CAF1C	

XP_001026968.2	TtCAF1C	T. thermophi	CAF1C	ENQAINDEYKIWKKN APFLYDIAITHELEWP SLSVQWLPTKDIPQE SDYAIHKLLGTHTS QDKDYLLIAKVRLPLE ETATDISEYQNAKE VGQTGLSAGENRIEIE TKILHDGEINRARYM PKYNIATKVTNGEI HVFDTQHTTTPQN DQVRPQLRLVGHSAE GYGISWNPKKQGYIV SGGYDKKICIWNVEA ASQLNSSISPLHDIEF HKSCVEDVAWHQIN PDIFGSVSDRRTVAI WDMRQKSNAGLINP THCTQAHTGDIYCLD FNPFNEYLFITGSEDK NIGFWDMRNTSKRL HTFVGHTDQVLRCE WSPFNVGVFSSCSAD RRVIVWDISKCGQE MKNEDLVDGPPELLF MHGGHRAKVNDIS WNQKENLILASVEEN NILQVWQMARNIYD DTDDEV MKD
NP_001083811.1	Xl-p48	X. laevis	CAF1C	RVINEEYKIWKNTPF LYDLVMTHALEWPSL TAQWLSVTRPDGK DFSIRHLVLGHTSDE QNHLLIASVQLPNDD AQFDASHYSEKGEF GGFGSVSGKIEIEKIT HDGEVNRARYMPQ NPCIATKTPTSDDL FDYTKHPKPDPSGE CNPNLRLRGHKEGY GLSWNPNLSGNLLSA SDDHTICLWDISAVP KEGKVVDAKTIFTGH TAVVEDVSWHLLHES LFGSVADDQKLMIW DTRSNTSKPSHSVD AHTAEVNCLSFNPYS EFILATGSADKTVAL WDLRNLKLLHSFES HKDEIFQVQWSPHN ETILASSGDRRLNV WDLKIGEEQSPEDA EDGPPELLFIHGGHT AKISDFSWNPNEPW VICSVSEDNIMQVW QMAENIYNDEDETEG GVDPEGQGS
NP_001105556.1	ZmMSI1	Z. mays	CAF1C	RAEVEERLINEEYKIW KKNTPFYDLVITHAL EWPSTLVQWLPDRT EPPGKDHVSQKMILG THTSDNEPNYMLA QVQLPLDDAEADAR HYDDDHADIGGFGA ASGKVIVQQINHD GEVNRARYMPQNSFI IATKTVAEVVVDYS KHPSKPLDGACNPD LRLKGHNSEGYGLSW SIFKEGHLLSGDDAQ ICLWDIKANSRNSL DALQIFKHHDGVVED VAWHLRHEYLFGSV GDDYHLLIWDLRSPA PTKPVQSVVAHQGE VNCLAFNPFNEWVV ATGSTDKTVKFLDRK IDTSLHTFDCHKEEVF QVGWSPKNETVLAS CCLGRRLMVWDLRSI DQEQTPEADGPPPE LMFIHGGHTSKISDFS WNPCEWVVASVA EDNILQIWQMAENIY HDEDDLPISEPAKTS

NP_001319681.1	AthIRA	A thaliana	HIRA	<p> QIFSIDVQPNGERFAT GGGDHKVRIWNMK SVDKDLQNIDTKERLL ATLRDHFGSVNCVR WAKNSRYVASGSDD QVIQIHERKPGSGTTE FGSGEAPDVENWKA VMTLRGHTADVVDL NWSPPDDSLASGSL DNTVHIWNMRTGM CTTVLRGHLSLVKGV WDPIGSFIASQSDDK TVIIWRTSDWGMMAH RTDGHWAKSLGSTFF RRLGWSPCGHFLT HGFQKPKHSAPVLER GEWSVAYDFLGHSA PIIVVRFNHSMFRIIP SSTHETKQVGWSNG TSKSGEKDLQSYNVIA MGSQDRTITVWTTG SARPLFVAKHFFGQS VVDLSWSPDGYSLFA CSLDGTVAMIHFDPK ELGVRLTDELDELKK SRYGDVVRGRQANLV ESPAQLLLETASTKQA GSKRAASDVQQNQV </p>
XP_006851504.1	AtrHIRA	A. trichopoda	HIRA	<p> QIFSIDIQPGGLRFAT AGGDHKVRIWNMKS VGRETEYDQTNQTL ATLRDHFGSVNCVR WAKHGRYVASGSDD QVILVHERKPGSGTTE FGSGEPDPVENWKV VMTLRGHTADVVDL NWSPPDDLILASCSLD NTVHIWNMSNGICT AVLRGHSSSLVKGIW DPIGSFIASQSDDKTV IIWRTNDWSQEHRT GHWGKSVGSTFFRRL GWSPCGHFITTHGF QKPRHSAPVLERGD WSATDFDLGHNAPV VAVKFNHSMFRKNL ANGADLKEFQVET VGWSNGATKSVPRE SMPYNVIAIGSQDCTI TVWTTASPRPLFAK HFFTQSVVDLSWSPD GYALFACSLDGTVAT FYFEVKELGHRISDAE LDELKRSRYGDVRGR QANLAESPAQLLLEV ASAKQLPTKRGLSNLS </p>
NP_498101.2	CeHIRA	C. elegans	HIRA	<p> DTGSILAIDCHPSGKK FITCGQKARTSNGLV VVWNAEPVLDKKKA SNENVPKLLFQVESQ SQSNSCRWSPDGKR FAFGSDDSSVSVWEY VGLINMSGITGGAQ NVERYKECCVLRGHS MEVLTVIEWSPNGKY LASGSIIDYRIIYNARK LPDRITVLDIQLPVK GLSWDPIGKYLASLE GDKKLRFWATDSWQ CVKSVTEPFESNIEET MLTRLDWSPDGKYL MTPAAVRSKGPLIKLI QRQTWKSQDFLAGH HKGTTTCVRAMPRIE ANLKNKGRMQLTCA AVGSRDKSISIVVFP GTLKPLFVINNFNHT VMDFAWCGRNLLAC SQDGTVKVIHLSSEVI GEMISNEAMSDLCY QIYSIRPPRYELTDKEE DESQDSFNLSDLSSSA NNASFVTCPEDVLIKR KKLVAAQQSSDIQLT </p>

				SIYSIDIHPDGRVAT GGGDAKIKIWSMAPI SLLAEEDAGIPKLLCS IENAHFHSVNSVKWS KDGKYLASGDDKLC MIWGLSNNSLLKN TTENWVCVATLRGH ASDISEVSWSPDNKYI ATCSFDKSIWIWETNK FQMVSKLEEKGFKV GLTWDPLGRYLASQS EDKSIIWRTSDWVV ETIVTEPFKHSGNSFF LRPSWTPDGQFIVAT HGINNATHTGVLSR TDWDIGLDLVGHRK AVVVSRCSSKIYKDFK SRDQKFCILLGGQDS TSLWSSSSPSRLMVT RSLFDQCIQDISWCS DGYSFVACSTDGTVG YISLDPEEIGGSPIGPE EKQFFKNYYGDAVTI DKEGNVIYGAAGGNL SSAANSUSHAILPENP DQLAMEESNANGVG SGVSGGVSGGGGS GSGINNQGNGENHS
XP_643117.1	DdHIRA	D. discoideur	HIRA	
				QIFSVDIHKDCTKFAT GGQGSDCGRVVIWN LLPVLSDKAEFDADVP KMLCQMDQHLACV NCVRWSQNGQNL SGSDDKLMIWKRSA GSSGVFGTGGMQKN HESWKCFTLRGHD GDVLDLAWSPNDVY LASCSIDNTVWIWDAQ AFPHSVATLKHTGL VKGVSWDPLGRFLAS QSDDRSIKIWNMTN WSLSHTITEFEECGG TTHILRLSWSPDGQYL VSAHAMNGGGPTA QIEREGWKCDKDFV GHRKAVTCVRFHNSI LSRQENDGSPSKPLQ YCCLAVGSRDRSLV WMTALQRPMVVIHE LFNASILDLTWGPQE CLLMACSDVGSIACL KFTEEELGKAISEEQ NAIRKMYGKNYVNG LGKSAPVLEHPQRL PQGDQPTKFPLSNNN EANQRPISKQTETRTK
NP_572401.2	DmHIRA	D. melanog	HIRA	
				PIFSVDIHPDGTKFAT GGQGEDSGKVVIWN MAPVLRREEDEKNENI PKLLCQMDNHLACV NCVRWSNGLYLAS GGDDKLMVMVWKRA AFIGPSTVFGSSSKLA NVEQWRCVTILRNH TGDVMDVAWSPHD VWLASCVDNTIVIW NARKFPEIVMTLKGH TGLVKGLTWDVPVGKY IASQADDHSLKVVWRT MDWQLETNITKPFSE CGGTTTHVRLSWSPD GQYLVSAHAMNNSG PTAQIIERDGVWKTN MDFVGHKAVTVVK FNPKIFKKKQKNGST PKPSCPYPCCAVGSK DRSLSVWLTSLKRPLV VIHDLFDKSIMDITWT LNLGLLLVCSMDGT AFLDFSQDELGDPLN EEKNAIHQNIYKSL AITMESQLSTTIENPE MLKYQQERQGNQN ANVAQSGSPENQTP
XP_696478.2	DrHIRA	D. rerio	HIRA	

NP_003316.3	HsHIRA	H. sapiens	HIRA	<p>PIFSVDIHPDGTKFAT GGQQQDSGKVIW/ NMSPVLQEDDEKDE NIPKMLCQMDNHLA CVNCVWRWSNGMYL ASGGDDKLIMVWKR ATYIGPSTVFGSSGKL ANVEQWRCVSILRN HSGDVMDDVAWSPH DAWLASCSDNTVVI WNAVKFPEILATLRG HSGLVKGLTWDPVG KYIASQADDRSLKVV RTLDWQLETSITKPF ECGTTHVLRLSWSP DGHYLVSAHAMNNS GPTAQIEREGWKTN MDFVGHRKAVTVVK FNPKIFKKQKNGSS AKPSCPYPCCAVGSK DRSLVWLTCLKRPLV VIHELFDKSIDISWT LNLGLILVCSMDGSV AFLDFSQDELGDPLSE EEKSRIHQSTYGKSLAI MTEAQLSTAVIENPE MLKYQRRQQQQQL DQKSAATREMGSAT</p>
NP_034565.2	MmHIRA	M. musculus	HIRA	<p>PIFSVDIHPDGTKFAT GGQQQDSGKVIW/ NMSPVLQEDDEKDE NIPKMLCQMDNHLA CVNCVWRWSNGMYL ASGGDDKLIMVWKR ATYIGPSTVFGSSGKL ANVEQWRCVSILRSH SGDVMDDVAWSPHD AWLASCSDNTVVI WNAVKFPEILATLRG HSGLVKGLTWDPVG KYIASQADDRSLKVV RTLDWQLETSITKPF ECGTTHVLRLSWSP DGHYLVSAHAMNNS GPTAQIEREGWKTN MDFVGHRKAVTVVK FNPKIFKKQKNGSS KPSCPYPCCAVGSKD RSLVWLTCLKRPLV IHHELFDKSIDISWTL NGLGLILVCSMDGSVA FLDFSQDELGDPLSEE EKSRIHQSTYGKSLAI MTEAQLSTAVIENPE MLKYQRRQQQQQL DQKNATTRETSASS</p>
XP_001762381.1	PpaHIRA	P. patens	HIRA	<p>QIFSVDTQPNGLRFA TAGGDHKVRIWNM KPLAEKETDSDTKLLA TLRDHFGSVNCVRW AKCGQKIASGSDDQV VYIHEKRPGSGTTEFG SGEPPDVENWKVLLT LRGHSADVVDLGWS PDDTQLASCSDNSIR IWQSTGALLSVLTG HQSLVKGLTWDPIGS FLATQSDDKSVIWRW SNWSMVKKVEGPW EKTVGSTFFRRLGWS PCGHFIATTHGFQNP SHTAPVLERGEWMA SDFDVGHNAPVVAV/ RFNHSMFRKVKTIAL SGNGAILGAANGTSS KTKEPYPYNVIAIGSQ DCCISVWTTGSPRPV FIGKHFFQQSVVDLS WSPDGYTLFCCSLDG SVASFQFERKELGEKI SDAEMEETFKSSRYGD LRVRQATVAESPAQL MLEAAAAAKWGS NASNANTIDSTKAVS</p>

-	PpHIRA	P. polycephala	HIRA	ACAKFEDQQNTSYNE QMRILKPDWVTHTG YPIFSVDIHDPGSRFA TGGGDNKVIIWASAP VQYEEQEQANNSCPRK LAELTYHMSPISCVR WSHNGQFLASGSDD ALVMLWLKSSGPASK SFGSDELTQENWRCT ATLRGHSKDITDVAW SPDDSQLASCSDNEI LVWDTAKNSVVAKLS GHTGFVKGVAWDP MGRYIASQSEDKTMI VWRCSDWKIETKIEE PFVRSAAATFYRRPS WSPDQGFISSAHAFS GSNHVAMITARDTW RSVCDVFGHKLPIVVT RFNPVIFADTAGNNY GVCLIGQDNTISVW TTISSRPLALNSCFKQ SVLDASWAKDGFSE MCCSTDGTVVWVEF AKGELGTALSPKEMD KVLRSIYGDHILSRPKS TVTVAEADPSQLYLPKT TAPMKPVAAQPPPV
NP_009545.2	ScHIR1	S. cerevisiae	HIRA	SRKYEIVTVDSHDK RLATGGLDGKIRIWSI DSILRCMELESLTPEIP LPQDLQMPLCSMSR HTGSITCVKFSPDGKY LASGSDRILLI WALD EEQSSQPAFGSEHER EHWTVRKRRLVAHDN DIQDICWAPDSSILVT VGLDRSVIVWNGSTF EKLKRFVDVHQS LVKG VVFDPANKYFATTSD DRTMKIFRYHKTGDIS FTIEHIITEFPKESPLTT YFRRPSWSPDGGQHIA VPNATNGPVSSVAIV NRGTWDTNVS LIGH DAPTEVARFNPRLFE RNAGVKQKDDDDPE NALVGQNDKVVHFF DKNIDSVVATAGQDK SLAVWSTSRPRPILVA FDIANKSITDMSWNP DGSLLFVASLDSITLF KFENNELGKPIPLEKN MEQLYRYGVDKDSL FPESINQLLLEDQTKS FKHTKISTSKLGENHP
NP_014681.1	ScHIR2	S. cerevisiae	HIRA	NALAALGPYIILAGSG GHVMAWRQQQLVD TAFDRVMIKDLKPEV SFQVDQDTTGDIFFIT GDLETLYIGSEHRLW GYSGLWLCRDTNNINS VEKMNSKLLFECKSPS TITDVKYDINLGILFVL LSNENKILLFRHKTFD KLSEITIDKASKPITGII DPTGQTFTVMTSDRS ILVYQINKTGTHKLIN KLTQHVMYPLHYRI SMSPQADILPVINSV KGVPPNNATCTALLD RNNNYKVTKTLVTPS SNGCRVLVYSPAFYEK PNLKKGTSTRYNLIAT SGSTDGTILVWNTKR MKPLFNALQVSSTAI NDMSWSQDGFTLFA ISNDATLYTFAFQEKD LGVALPQTEIKSLQEV NKKLPKLEELAEQIP KSPENIKLEESASAA PIPNDIGRSAVGKKPT KKKTANNQTNIGIKTI QSTSMFNTPSYTPV

XP_001014644.1	TtHIRA	T. thermophi	HIRA	EEIISINFQNGNVL VFCSGQDNDQKIVK TTINQDYMNQNVTN KVDFQPLNIRSVQSS NLIKWSPNDQFLAAA FEKLGWGYANGDIY KLKDFKDELEIAAIT WAPDSSRIASGGLDG KIVVRSVDIAKNYEKI KELSHDNKILGLVWD TYDKYLAALYSDNNVI IWKCNSEKSFIVSLT CPISSDISKITSKRDDR KIDWSPDYRYVLVPSL DDKIVPFVICALDRQK DFKVCKTFMGPFSSI NCVKFNPNLYENKES VINVFALGDNDGNISI WGIGDGYQAKKPFFL FKSHPNGNELIEDIC WNAEGNMLIATTLK KYVFLALFGETVFGSK LSPQQQDFLEGKYG TLRTVAATNMFKVVF KATQEEEEEQKNKE KQLFEAKGDNKQIDSI LGRGNNGKTATQF GSSNSTNQADQNQ
Q8QFR2	XIHIRA	X. laevis	HIRA	PIFSVDIHPDGTKFAT GGQGQDSGKVVIV NMPPMLKEEDEKNE NIPKMLCQMDNHIA CVNVCVRWSNNGAYL ASGGDDKLIMVWKR SGYIGPSTVFGSSSKL ANVEQWRCLSLRSH SGDVMMDVAWSPHD AWLASCSDNTVVI WNALKFPEISTLRGH SGLVKGLTWDVPVGKY IASQADDSIKVWRT MDWQLETSITKPFDE CGGTTHLRLSWSPD GHYLSAHAMNNSG PTAQJIERDGWKTN MDFVGHKAVTVVK FNPKIFKKQKNGSST KTSCPYCCAVGSKD RSLSWLTLCKRPLV IHELFDKSIMDISWTL NGLGILVCSMDGSVA YLDIFSQDELGDPLSEE EKNNIHQSTYKSLAI TTEPQLPNTVIENPE MLKFQQRQQLQQD GEHMQAQMEAPI
A0A3L6FHG3	ZmHIRA	Z. mays	HIRA	FSIDIQTGGLRFATGG GDQKIWASVLPAGR PVPKKERVSGGMKH GGEKQVCGAQLGFR PCFMAHLAPPPQLLA PSSPQPLPALASAAAS NAASVRLAKAARLVR LAKAARLRLAKAAR PLQLCHRRPRLTAGS PPTRTMGASCFTITT PLPPACTSAVSPPLPL LFYRQTPATKIILSLEH PATTRPPVPSLLQAPK RLQKKALDSQSEGML ACCVISLLRKMVRIV SMESVHKDNANND NQRLLATLRDHFGSV NCVRWAKHGRYLAS GSDDQVILHERKAGS GTSEFGSGEPPDAEN WKVIMTLRGHTADV VDLSWSPDDSTLASG SLDNTIHIWNMNGI CTAVLRGHTSLVKGV TWDPIGSFIASQSDD KTVMIWRTSDWSLA HKTEGHWTSLGSTF FRRLAWSPCGHFITT

				ELLRASPKVLTAGDRK LLKVELRPGDTTYVS WKKLMRDAGKVNL SASVPDPPPNANPNL EFRIAPGHPVEIETNE QPHSNRFNAVIEKIER LYKGNDS SDGEELDG APDDDEYDTEDSFID DAELDEYFEVDNSTV KHDGFYVNRGKLER MEPSTTSNQQPKKR RRKDSAKPCRDADV SDKHTKLSITARKKDQ STAPGSWKTQESPLP SGAQDANTSVP added VKHSDRANHQS added TSHKSRETGSSSALHQ KYSNKSLLHQSTSLG KSPPNVFAEVTVVRQ KENNGMHQLANVT GSRQSSQASKDGSN VKKTSILEKAIKIRELEK VVVESRPPAITENQE ADTSSQAVKRRLPRD VKLKLAKVARIAS QGKSTELINRLMSIV GHLIQLRSLKRNLMIM IDMGDSATREKDFR
NP_173580.2	AtUBN1	A thaliana	UBN	
				NESCKISSEILTAGDRK LLKVELLKEETTLVSW KKLMDEASKENGGLF VSAPERLLNANPNLE FRLAPGAQTENEMV NQPHPNRLNSVIAKI ERLYMGKDGSDGEEL DGAPDDDDYDTEDSF IDDAELDEYFEVDNSP IKHDGFFVNRGKLERI EPSATSNQQPKKRR RKESAKPCGDVVDVS RKRAKMAKTAGGKD QSASPGPSSKKISNDS KTVQDSFSPKQNG NDSLVLNVKHTDKA NHQPMNATSPKSKA AGSSGPLHPKCSSKS VHEQSNPPGKSRPN VSAKSAVVRQVNN GMPDLDIATESKTSIQ ISKKSNGRPRKYSTL EKAIRNLEKLVAESRP PAATENQDADISSQA VKRGLPGDVKLHLAK VARIAYASQGEISGELI NRLMGIVGHLIQIRSL KRNLMIMDSIVTANR
NP_177855.4	AtUBN2	A thaliana	UBN	
				SVSPSGAVASSLYDP KGAISQRQSFVSLR PGETTIVSWKRLVKD SNKAIRNSASVEPPV GAHPALESRIAPEGH ARHELEDVPPSNRF SAVIQIERLYKGRESS DEEDLDIPDDQYD TDDSFIDDAELDEVFQ VDKSETKHNGFFVNR GKLEKTNDPISSPVHA PKKRRRDLNNAISE NAAENLPKRHLNVG GVRMKAARNAPLV GNKLSVPSQAKTSV GGELGHDIKSSQNQL IACVETSKQPLDSS MKQENSISTKLPNKDI SMEDKDGSQKPGLL PSGEGASKLTVNKEH AHPIDRDLKGNPV QVESQHKKALKDAK GLLPSSKIRPKEGVGS SDMRDTHILASKSPL QKMKLPSLASKEGAA VRPKGTMLEKIRELE NKVAELRPPAMDLO EVDASAQGIKRRLPQ
XP_006847173.1	AtrUBN	A. trichopoda	UBN	

NP_001255177.1	CePQN-80	C. elegans	UBN	SAKNEVIQIRISVAKP GSKKYTHVDWEELLK NNGKERDEDDLRRFY DEDTLFMAKKLGETR SKSGKKLRVNLQDQ HFNRKCGYDMDDDF IDDEIVDDEPTTSKK GGFYVGKGDVKDSN DEDEEEEEETDVEESI QEVPKVKIRKIKEKS GEKSSKSSDSSSSSS DSEDAAGPPRMTGA PPSIRMVSSSSSLKKL PAPRKPSPPPAVEPT RSVIEIEKPEKTEKPAE KRKTAATTSALEPAEK KKKPDSSSSSEIILDD NAPAACKQPESAQK GSTTIVTMSAKKICKQ QEIEIAKQKPSSSSVTP SKPAAPTASTPKKEEI GTLETFLKNIFANLDA KGKLYQEKKLKTFDTE TLKEITTFIDVMKHHK TEKVNMPLIYSLAKS FGMTTQEVVMKQVE QEISSKATTPKPSTN AVLSTISSQKSLEKNA
XP_645639.1	DdUBN	D. discoideur	UBN	NKISNGNSFFSPNLFT FSSSKTLTSRILNFKN NNNNNNNNNNNNN NNNNNNNNNNNNN NNNNNNNNNNNNN NNNNNNDYNDSSNE EENEYSIIKRKKKIKIAP TISKLLSIKKSEFNDD NSDEESYEDDFVDYE NENENENENENEYK NENENEDNKNEDNQ DNDLAPSSSKLFRFLC TPKIQTNIKINNNKNK NKNKGNDSTKKVD SEINEKEEDSNAELKS KPKAIRIEVDLKSNGIII NFQQLVSNESNPLIS NNNNNNNKSLISQT NNFVKSFKNENISNIT FNERKNVLGMVLIKI GAGAKNNSSTRVEQ KVEKEEERGEEEDFHY YLDSGDSTISEYDIE DPFVDDSLLIKDHAAH SDSNGIDDDSDINN LETNGFYVIQNGIDLK FNNNNNNNNNNKNN NNNSGGANNNNNN
NP_476893.1	DmYem	D. melanogas	UBN	QGDAGFSRFGSNILE PDAASAAPDNSSKPT TKTAKCIRIKDLFETD SNKYPEFNYSRLLYLE KKKTKKLKQVSTTNG SASTDPFADNDDV ARIVKELEAKYGNSYA TGRGSKKDDYRDIG MGYDESDFSIDNTEA YDEIIPAEATLEGGFY INCGALEFKNLTKKS TTRTDIIKMPERSRK RMVSSSSSSSSSSGD DDENDDGNNNEEDDE SDEDDSEENDES EDDSESELEDEDSAA TAKSSSKYKDNHQAK RAKVIVTGKSKPSSSS LTSGKKPPTKPITSSS SNSPRPSTVEISDTED GQDPITQTPSSQLQS LPQSQAQAQALKKV VKTTTVKMDLAKARD SFLKSQSGTAAVKGV GNGELKCVSTDVSSS DSSDMESEHGRADR QAGQHKGKGQENLR TADTLPTTLADIVT

				EAYDELVPACLTTRYG GFYINSGTLQFRPASD EGENENDFEDSGLKP KKRKLQGGKDKKIMK KKRDEYGLAKKDEEEF TKSTPPDKSSAKKKKN KKPLSIDKMLKFFHKE KLQQMKEFNSREKEL HSLNAIHAGPAIQAEI QEAHVSSDPLFSLIGS ASADDLLQAVKAAEQ DFDLDDLLEGPQNIC PPGLEEVNGEALVVS VAEKPPALLPDGLPP MLEQHIKDISQAANK IEGQNKMEILSSELNS VLLDEVNSKQLGGK VRSRIFSYLASQLSCSK GTLVKRAKKLH5QQQ DEQMQLLKKLENAV ARSMPEQITRFQNH QAHSEARAAKLEAEK ERAIDGSDEEEERSG KRVFGPRKRFRWNEE IRELLCEIVNVKMMIY DSESPSCSSLEEYLKAF LEADVKSLLWPKGWM QSRILLIETRAKHGHIT
A0A0J9YIW0	DrUBN1	D. rerio	UBN	
				AGGPGADNRRRRLE EDGAAEMSFDRGR MKMMGTGVTTAGD RGGUGKRDSEEPKG KTVRLNLCLSEPEQC SAEFNYSLIQSQQPK KNPPCSKSLDPNDP FNDDEKERLQVEALA KKFESKYGNAGKKRR KDRMQDLIDIGFGYD ESDPFIDNSEAYDELV PASLTTKLGGFYINTG TLQFRAASESEGEDK VKDEKPRIRDGEEPL MKRRKKDVGIEEKK PRKNKVAQGVAGL SLHRPEKKRKLKLMK DSLCLAAMLRRFTRE KEEIRKQDISSHLGLT SAPNSNNLLQNSHLH IGHANHTDLSADLT ADPAMMSLLSSANE SELQDIMRDLDFSSLD AAPQTPPSCRENGQ VGLGSSLGNKIGVGG LNRGQAGVVSPFTLP NGLPAPLIKRIEDLRA ASRQFDQEGRKFFFT
NP_001073638.1	DrUBN2	D. rerio	UBN	
				LNPFLKKSKEEAG AGEQHQCCEPAAAA VRITLTLFEPDHKRC EFFYPVLVKNIRGKVK GLQPGDKKKDLSDPF NDEEKERHKVEALAR KFEKYGGKKRRKDRI QDLIDMGYGVDES FIDNSEAYDELVPASL TTKYGGFYINSGTLQF RQASEEDDFIKEKKK KSPKKRKLKEGGEKIK KKKKDDTYDKEKKSK KSKFSKAGFTALNASK EKKKKKYSGALSVKE MLKKFQKEKEAQKKR EEHKPVAVPSAEAQ GLRELEGASDPLSLF GSTSDNDLLQAATA MDSLTDLDLEHLLSES PEGSPFRDMDDGSD SLGVGLDQEFQRPSS LPEGLPAPLEKRVKEL AQAAARAAGESRQK FFTQDINGILLDIEAQ TRELSSQVRSGVYAYL ASFLPCS KDALLKRAR KLHLYEQGGRLKEPL
NP_001072982.1	HsUBN1	H. sapiens	UBN	

NP_775840.3	HsUBN2	H. sapiens	UBN	RREAEPGPEREPEYP REPPRLEPQPYREPA RAEPPAPREPAPRSD AQPPSREKPLPQREV SRAEPPMSLQREPPR PEPPPPFPPLPLQPPP PRESASRAEQPPRPP RETVRLELVKDPDTE SCVEFSYPPELLCGEQ RKKLHTEDPFNDEH QEROEVEMLAKKFE MKYGGKPRKHKRDR LQDLIDIGFYDETD FIDNSEAYDELVPASL TTKYGGFYINTGTLOF ROASDTEEDDITDNQ KHKPPKVPKIKEDDIE MKKRRKRKEEGEKEKK PRKKVPKQLGVVALN SHKSEKKKKRYKDSLS LAAMIRKFQKEKDAL KKESNPKVPVTLSTPS LNKPPCAAAALGND VPDLNLSSGDPDLPIF VSTNEHELFEAENA LEMLDDDFDRLDLD ASDGSPLSESGGENG TTTQPTYTSQVMPKV
NP_080942.1	MmUBN1	M. musculus	UBN	LNPAFLKKSKEEVG GTEQHQCCEPAAAA VRITLTLFEPDHKRC EFFYPVLVKNIRGKVK GLHPGDKKKDVLDPF NDEKERHKVEALAR KFEEKYGGKRRKDRI QDLIDMGYGYDESOS FIDNSEAYDELVPASL TTKYGGFYINSGTLOF ROASEEDDFIKEKKK KSPKKRKLKEGGEKIK KKKKDDTYDKEKKSK KSKFSKAGFTALNASK EKKKKKYSGSLSVRE MLKKFQKEKEAQKKR EEHKPVAVSSIEAQ GLRELEGTSDPLLSF GSTSDNDLLQAATA MDSLTDLDLEQLSES PEGSPFRDMDGSD SLGVGLDQEFQRPSS FPEGLPIPLEKRVKELA QAARAAEGESKQKFF TQDINGILLDIEVQTR ELTSQIRSGVFAYLAS FLPCS KDALVKRARKL HLYEQGGRLKEPLQK
NP_796159.3	MmUBN2	M. musculus	UBN	RREADFAGAEREPPR LEPQPYREPARAEPA PRADAQPPARDKLP QREVSRAEPPMALQ REPPRPEPPPPPLPLQ TPPPRESASRAEPPPR PPKETVRLELVKDPDTE DESCVEFSYPPELLCG EQRKLVHTEDPFTD EHKERQEVEMLAKKFE EMKYGGKARKHRKD RLQDLIDIGFYDETD PFIDNSEAYDELVPAS LTTKYGGFYINTGTLO FRQASDTEEDDFTDN QKHKPPKVPKIKEDDI EVKKRRKRKEEGEKEKK PRKKVPKQLGVVALN SHKSEKKKKRYKDSLS LAAMIRKFQKEKDAL KKESTPKVPVTPSSSS LPKPPCVTTALGDDIP DLGLNSADPDLPIFVS TNEHELFEAENALE MLDDDFDRLLDATS DGSPLSESGGENGNT THPTFPSQVVPKVVP TLPEGLPVLLEKRIEDL

XP_024392537.1	PpaUBN	P. patens	UBN	TTVSVWKKLVSDAEG KLVSARVSDAPAGAN PALEARIAPEKGAGS VTGNDDLPQVPNR FSSVIERIERLYKGGDS DDGGAEDTPDEDQY DTDDPFIDDEELNEVF MVVEAKTKHTGFFIN RGALEKISVSPSPPAV QPKKRRRRESKKDKK DEKEVKRRMKGRIM QGSRQFPSSHNFNA GGETQVSNHTLNKA QSSDPPEISSDSDAFR QGSAGPSRVSYKQEK LDFADVDNNGGVVE RLLETTAMLNTESVF GGGRNIQKSEYRYGS SFEEKWKQAAREPK GVMKVSESLMKMI QQSGKEDGAKDSOG GAGHAVAGYTKSLSP GPKEPSGRRGWPK GTVMERAIQDLIKMV ALLCPPSPEGEEQAL SQPGAKAKRLPREVK TRLAKVARLAQAKQG KVPDELVERLHGILGH
-	PpUBN	P. polycephala	UBN	CYRFSISLSETKGVVN YADLMSQERKKRHLL SNPPSDPSTNPITTTT TSEDTPKTPPLNGSV DEASIDTAGDESVEA ADESVEVEDEGNAS ASKDVKKSSPKRSMP VLPKKTTNRLVQIID KIEKMYTTGGDSFMN GYDTSDFSIDDSDMV EKPTEEVENEYGGFFV HSGELKTRKKKRVSDF QEDSPDKNKSRTKK LKIEDEKGSPEKDKK NDENGKKSDDKKPKD KKKEETPKDKEKKKED STKDKEKKKEDPGKE KEKKKEEKEETKDSEE EKAALRQKLLQTEEKI EQAIKELETEAKKKIVP GQTKRIPAEALAPQLS NLYEALQQTANGFIT DRLLDRLVEFLPFKKS TIKMHKKSKHKAHV DTMEVELKRLEDAYD TEYNLLRAAVFNTMA EEQKKQQNPPPTD PPADPSNPAEPKYA
NP_009774.4	ScHpc2	S. cerevisiae	UBN	KQTKSSGKMQTQTD TNAEVLNTDNSIKKET GSDSEDLFNKFSNKK TNRKIPNIAEELAKNR NYVKGASPSPIISGSS STSPSGPSSSTNPM GIPTNRFNKNTEVELY QHSPSPVMTTNKTD TEEKRQNNRNMNDNK NTPERGSSFAAKQL KISSLLTISSNEDSKTL HINDTNGNKNSNAA SNNIPSAYAEHTEG NSIESLIKPPSSPRNKS LTPKVILPTQNMGTI AKDPHLGDNTPGILIA KTSSPVNLDVESTAQ SLGKFNKSTNSLKAAL TKAPAEKVSLSRSISS VTNSDSNISSSKKPTS EKAKSSSASAILPKPT TTKTSKKAASNSSDST RKKNASNKTTSAIKKE SNAGSKLNTVKKENS SLSSIKATEKEKDKGG NSTEAKNSTSNVRKE PTAKSPKRLVAAPTYS PPKILQTAETKAKEPSI

XP_001017549.4	TtHpc2-1	T. thermophi	UBN	<p>PNFGFCQDKQIFQH DLNENNINQQHQKIL SKIEEKPNIKQNKHILE RMHDEQSHDEEHLE RGYEWIQSTNEEDFQ IMKNVCYLEQFLDMY DDLEEDFLHLDIPQF MQYLESESESESIQS DEEEMEDQFEDNDLI QNVNTHRYSESRLLN TNLKNEQGVKKTDEF QKFDQPANNAESTM NAEAVVKNNKIITNM KNKGKNQMSQNSID SQREDDIYSQINQKG KAKRYQSSALLAQE QFIRISKYGDIALTSNI ARKNPKSDDFYDQED SFIDDDQVDTELKQL AICEADYDDFFMVEG GLQDLVKSSEFRERLK DIQEYRDTKKIENHLK TDKKEKQPNPKIKISS ENSKNLSFNEKKQQG LNLLMDAEDEEDDK QJKKKKNAAKNKKQ FDLLHLKQPTFQEEM LQNAQNTKIQPAKT</p>
XP_001010706.2	TtHpc2-2	T. thermophi	UBN	<p>GEDVADSNSSDIMEE EENEVDKPSIAKFSNL NNNNNNKLDQVNG NKNSVNGSADKKKK KQRSTNKVSIEELTKQ IKQNQSKRVNGQED GVVIKANFDWVNYN TTEDKTLNLYSDFLKL NGENNLDYFKLDLVE FEDVNIVQHHSQVSS ITDDLKRLRYGQJEN IQRIAKKSGAISKGEG TNVNFREDELYDQND PFIDDEYNNTSVKV SHRMQLAEVFIDDFR TFDGLDKDFIKSEIYK NRVEIMHNLEKKNEI EEETENFKKKKKPN PPSNPAISSNNLSSN LANNSLTAALNSATA TLNNKIQQKATADVE KDLNQVESVKTQEEK KFKKDNVNNTTSLSS GOQLQNAQNIQAA GLNQNLQSVSGSVN AQSLPQAVQASNLLL GADLSTSLSQLLSQ AQSSNNNNLLQNA</p>
XP_018092158.1	XIUBN1L	X. laevis	UBN	<p>AVKRQRSEEQESETS VAATVRISLTLFEPDQ KRCPEFCYPELLRNLR GDGKGLHGDKSQK KVSVSIEEEAEKKEV ANIAKRFEKYGNKK RRDRKQDLIDMGY GYDESDFIDNSEAYD ELVPASLTTKHGGFYI NSGTLQFRQASESED EVVKEKKKSPKMM KDRGEVKKKKRAEE KKCKKNKPKPGFTA LNGTKDKKKKSAIF KMLARFEQEKNAQK NIITPPSPTTLKPSSIS QQPILSQPTQEAEP APDAAVSTHVSELLK AVSAIDSVREKQMMEN IHNGPPEKKHGVQG GSTSEEAIKTPASLPD GLPHALEKRIKELTKA VKASEGEKKGLFSQE MNSALLDIYLSRELSS SLRLAVFTHLASVLP GKDTLVKWATRLYLQ KQGGRLREPLRLKE AVARTMPEQITKYHE</p>

				KPRETGPVLSVEQQN QQTPPPLQPPAQR PQPRAAHQVMFQPR PPQALGHRLQREPS SPALEAGICHVLDKP QGDDTHQLAGTVRL QLALTDPTTEGCAEFS YPDLLKEQKAGRALI PVTRKETAGLSKPPAS NDPFNDDERERQVE NLAKKFEAKYGGKSH HHRKDRVQDLIDIGY GYDETDPFIDNSEAY DELVPASLNTKYGGF YINTGTLQFRQASDSE NEDFVENKKHKTSKV SKIIDGERSLKKRRK EENTEKEPKMKMAA KQJGVMAFNSHKPD KKKKKLYKDSVALAA MLRKQKEKDAIHKK VNKPNPPTVTVVNTA PKLVTAAASNDLSDLGL GADPVLAIFGSEREL MQEAESALEMMLSDF DFDKLLDSASNDSPA SGGENSRVGQSLNPS EVQLPKQVPLPEGL
XP_018116614.1	XIUBN2S	X. laevis	UBN	
				AAAAAARVPGAVAA APGVTSAPSALAPTQ PPATAAAVSEAATPC QRQLFTVELRPGETTI VSWKLLKEAGHTAA APPVAVEPAFAAHA GPSGPAHPAENDPK DPTQPNRNFNAVIEKIE RLYMGKHSSDEEDLN DVPDDDYDTEDSFI DDAELDEYFEVDNLK TKHDGYFVNKGKLEQ IEPGTAANVAPKKRR RKDASSYLESNHLGP IDYFDIGDVPEKSSAR GTVQTGKQLVSSNVS SYGQYHEDNRVVKN KISGPGGAPKRSSEF SGDAAARAKIHKDVS HAPLELRDMEKHKA AALPVDYAHKSKISSET FDYAYPAYRDKGTSG QLDFQQRKISRENOG PSNRTYRKEKHGTNE YPGIAMATAVYSTQT MHPVVGREGSGTKP KGTRLERAIIRDQKIV AEYRPPTIDINEVDPN
XP_020405992.1	ZmUBN	Z. mays	UBN	
				WEPLAPSKAEQEFHL SQTYHDGLKLQAKD YDKARELLESILKDPIT NSKVETIANDNHLHH LRFALKNLATVFLEL GSSHYENALNCYLQA IDLDAKDSVLWNHLG TLSCSMGLLSIRWAF EQGLLCSPNNWNC MEKLEVLIAVGDEV CLSVANLILRHWP SRALHVKKCIEDTDS APFAPKGIDKLEPQH VRLKFLGKRKVSMDN QDMDATSKKLKRV QFKLPEASWVALLNIL IGIVHPSRETGVISADI PITIELSLSTEAVMQG MKKKDHCVDSDSSN VSVKDCNIERESGGS VKEKEPVFSEHPQE RRSTRLERLRNQKPEK EGLEFDNSKDPSSDIL QYLEKFLKRGFDRES AGSFCNEESDPISEHA VVSNFVKENLENYGA YHMGHLLLEYIANKC EHVLSRETALKILELEK
NP_195008.6	AtCABIN1	A thaliana	CABIN1	

				QWETLAPTKEAQEFH LTQEYHEGLQLQAK EYSKAQRNFEAVLRD PLMSNAQVDGNTRD GHLLQLRFLALKNLAT VFLEQGSMMHYESAIN CYLQAVEIDTKDSVV WNQLGTLSCYIGLLSI ARWAFEQGLSCSPN NWNCMEKLEVLIAI GDEISCLSVANLILRN WPSHARALLVKTVE ESERIPFAPRGIDRLEP KHVRLKFPNKRKRSIDE PYDGADIKKRNQTV ELLLPASLSSLVDAIL KILRPLSAKGSSENSPIN EKQSRSAEFPSRNNE LHGKGETDDLNMKG RYTDMRMRINVSSES ENVGDPFIGKGMGM LLAGNQSSSTAENPE KIFHGKEKEGSDDEV HPQERRSTRLERLSR KPEKEELDYPLSKNRA EAAFQILKPFIVVGLE KKDYNYNQNTNTVLL MDTLEFGVNDISRF
W1P7Q2	AtrCABIN1	A. trichopoda	CABIN1	
				LNEDESSDGENSVII LDSOTEPEDGAGGEE DEDNGEDSEEEDEED SDEEEEDSDSSVDOT ESVNSVWSNVDEEN GEEITTEAAEKIKKF EKAIFSADNDSETGR RILEKLLNDIAIATYNQ ENPDFELDEPPRLAK MLQVFIASHRNLAKE CEKEVENADEKEKEQ LKSEEIQLCQVLAYE PNSSDIWLDVALKSIE FGDLNFSKYAFKRCE LKESLESHATLLYLTCD YNESLMIMKQFQEQ NDVLNDKLKYLKHKIR STNQYYKLSCDRIFEE DEVYADVQSVEKRKI QDFDERIEALRSRVV GKIEQRDVFDRDEEE KCSVIPRIEADQDLQ TVCTIFCDLFDRIHAY SHCRHQLIEFTEWES RRDHLDLVDTVEHLV DVVECVEDISQKVEEL ENLKRGAAGRKKQ QQKDEKEYLASWR
NP_001299995.1	CePICD-1	C. elegans	CABIN1	
				YSEYESEYDDTDG DDDNNNNNNNNNNI NNNNNNDKFDNMSI EAQETYNMKLYSNAL KHQQNNNDYLOAQQL YQELLQSPMIERINPQ SKKSILSSQMATLKYS TLKNMANINEAIDNK ELSKLSTQAIIDLGR DVSLWYSIGKISMELE KWNISRLALERAQIS PTHWLSLEKLIELLYVI GDDNSCKNIEQRILQ EQPSNLKAIKLLIQR SNQSNYNEINEILNNI LISNQNSKSYIQLIKK RDRIQQQQQQQQQ QQQQQQQQQQQQ QQQQQQQMIIGKQ PQLTNELDKPVEYIL ENKNWSSLLDLNSIY NEKKESTSSIFSPIEI KLHPSLLGIQLVLDG NEDDVDEMEEEEE VDITPINTVDNKKQ PIAIEITDTSITTATVTA TTSSTNLESSIRPATPV TPTPTINTSNTTTTTT
XP_647263.1	DdCABIN1	D. discoideus	CABIN1	
NA		D. melanogaster	CABIN1	

EOYA53	DrCABIN1	D. rerio	CABIN1	EHEDAFRSTKPGQTK EAQEAFAFYHKAL DLQKHDFDESAY HELLKTPLLKEAVASE DEKVGKHPGLMLKY STYKNLASLAVLKDDL DTAMDFYVQAVMLD STDVNMWYKMGKL ALRKVSMPARHAFE VGLHCNPEHWPCLD SLITVLALSDYSCCLY YICKALEKDSGYSKGR VLKEKIFQEQPCLRRD SMKMFSLDISLQYV DVEEEAHSSIVEEAL LRRQROAKLTRHPVA DLQLVQPIKYFTWKS VGESFLAMYKHQNA CLVPRPDFGRIDLS MYRDPDCLLQMPISA PVPAAAAAPSDPPQ PTVPISDPPAPVVLPP APSSPSASLQLYDTVI WYSRVNGALDDKVK KGTKRKRVEDTVETA KRRSARVRNTCKCKE EKIDFQELLKFLPSRL KKFEDDEDEESLSNM
NP_001186210.1	HsCABIN1	H. sapiens	CABIN1	HEGSFKSHKTQTKEA QEAFAFYHKALDL QKHDRFEESAKAYHE LLEASLLREAVSSGDE KEGLKHPGULKYSTY KNLAQLAAQREDLET AMEFYLEAVMLDSTD VNLWYKIGHVALRLI RIPLARHAFEEGLRCN PDHWPCLDNLITVLY TLDYTTCLYFICKALE KDCRYSGGLVLKEKIF EEQPCLRKDSLRFML KCDMSIHDVSVSAAE TQAIVDEALGLRKKR QALIVREKEPDLKLVQ PIPFFTWKCLGESLLA MYNHLTTCEPPRPSL GKRIDLSDYQDPSQP LESSMVVTPVNVQIP STVSTNPVAVAEPPV VSYTSVATTSFPLHSP GLLETGAPVGDISGG DKSKGVKRRKISEES GETAKRRSARVRNTK CKKEEKVDFQELLMK FLPSRLRKLDPEEEDD SFNNYEVQSEAKLESF
NP_766137.2	MmCABIN1	M. musculus	CABIN1	HEGSFKSHKIQTKEA QEAFAFYHKALDL QKHDRFEESAKAYHE LLEARLLREAVSSGDE KEGLKHPGULKYSTY KNLAQLAAQREDLET AMEFYLEAVMLDSTD VNLWYKIGHVALRLI RLPLARHAFEEGLRC NPDHWPCLDNLITVL YTLSDYTTCLYFICKAL EKDCRYSGGLVLKEKI FEEQPCLRKDSLRFML LKCDMSIHDVSVSAA ETQAIVDEALGLRKKR QALIVREKEPDLKLVQ PIPFFTWKCLGESLLA MYSHLTTCEPPRPSL GKRIDLSDYQDPSQLL ASSVVVTPVSVVQPS PICTNPVAVAEPPVLS YTSVTAASFPLHSPSL LDTGTPMGDVSGGD KSKKGVRKKTVEES GETAKRRSARVRNTK CKKEEKVDFQGLLVK FLPSRLRKLDPEEEDD PFNNYEVQAEAKLES

XP_024373289.1	PpaCABIN1	P. patens	CABIN1	FSEPLAPTEAQESRL TQMYGEALSHLQHG QJEKAQSLFQSIQDP ISIKAQVEPVSSANT MLQLRYSALKNLAET FSMQGSGHHVEAVD CFLQAAIDGKDVVL WNRLGTLSCALGNL NVARRAFEEGLRCSP RHWSCMEKLVIELIA VGDESACLSVAKRLLK LSPSHPRALQIQHAIE QGVSMRPFDPGPKSG HETSSVLKIGPRGFDL LQPEYFGLSFDNKRKL ENAYETVQHQQKRK VHSIDVHLPEASWFA LVNAVTDVLKGGRR WKDDPEPLNTQIVVS PGGNQSEQQKAASA GISAGLVNATVKFSV HKSHPIIVDLNSEYKE QVAGEEAEIISEDCLN SSSTKTSEEVSRGLG VAPVLAEPRESASE NRAEGHKDVENASV HQKQGGTGEKQEAS KLPDLMKPSDPTTEW
-	PpCABIN1	P. polycephala	CABIN1	KKHKHTPTKGDALEA ENFALYEKALRHQYA GEYQEAELKYSILAS PIVNEFVEVSKRDDLY YTLPVVTLRYLAYKNF ASLCETTSRKTEAMEY YAKALDVDRDWWV WYHMAVCMMSVERL QLARFALEQGLKASP DHWLSIDKLAEVFIL GDYGACMNVDRIIL ELEPGFSKAILLTKM MSENPDKFKVEAFL WEAGDLGDCNIDDL VSARSKRQKVADFDH DTVEEPPREFVLKEKT WSSFADLLYDIYKEM TSEKGSKFLNDPIYIT DPAPLSPPGSPAASA LNDSFPSLNFTTPSSG TDLNSGSGTASSGT NQPGLNESATNLNTS ATNLNTSATNLNTST SNLNTSATGVNSGET EAPNVESGRKSTRGR KPAADSATRQSARVK EIKHDKISSDLDEYLN KLLNTQEESDETTASS
NP_012674.1	ScHir3	S. cerevisiae	CABIN1	QYEAEEHSRELQIEQS FNILQDALIDLKKNDF EKSDSKFQELFQIDVV KPDWRGMYRNSPT LDNLRYLRYNRGMY YHLYLENNYERLNSQ ELVNCILKAVENLVESI QHSDADFAVTDLLAR IFKSFNSVKLERLISEYE FTKQENLSLLGRHRK FLLNDLTLMMNNYV ELTNKLLVPNLSDNTI FERVHLEKYKDIKPEPL AFGPILSRISEMKKQD EEIMKKLDVFNVTLN EESWDEVAKALKNLL PSVKTSSLIGRNMDP YNEIEEPIEAVKFELSE AINNTPSLDRESEKQE EEQDNESVRADDKSG NLAPSDIQTNEEARP NKRTDEHIDSTKPLQ RSSKRFRKEREQENSKE LVMDVHKRFFGEFNT LLSYIHILPFCDFDTFA SKFIIGSSDKQPEKFIP YTDLYECLKSWSSRYT DIFNQNDYLSSGSNE
NA		T. thermophila	CABIN1	

				EHEETFKNHKTQTKE AQEAEAFALYHKALD LQKLDHFDESSQAYH QLETRLLEAIPAND EKEGLKHPGLMLKYS TYKNLAQLASQRDDL ETAMEFYIEAVMLDS TDVNLWYKIGCLALR LVRIPLARHAFEEGLQ CNPDPHWPCLDNLITV LYTISDYTTCLYFICKAL EKDCRYSKGLVLKEKI FEEQPCLRRDSFRMF LKCNMSEIHEVDVSDE EARTIVEEAMDRLRR ROALTVRKAEPDLKL VQPIPIYFTWKCLGESL LNMYRHLISCEPPRPS LGKRIDLTEYKDPAVL FHPPPVSPDTIIVDP LMISTPPTPVSVIEMP SVLPTATPPASSSNVI DTSAVQADSSTGDKT KKGVKRKKISEDAGET AKRRSARVRNTRCKK EERVDFQELLKFLPS RLRKSDEEEDDPFCS FETQSEMKEHYAN
XP_018117354.1	XICABIN1	X. laevis	CABIN1	
				WEPLAPTEAQEFAL SQKYHEGLKLQQKD YAKARELLEDVLKDPL ISNVQADNIGSDQHL LQLRFLTLKNLASVFL QQGFEFYDNALHCYL QAVELDSNDSVWWN HLGTLSCMGLLSVSR WAFEQGLVCSPNN WNCMEKLEVLISIRD EVACLSVANLILNSW PSHHRALHVKKTIEC AEPVPFAPRGIDILEP KHAKLIFSNKRKSVDD EICQEIGTKSKQSAT LQLNEAKWLALLDGI LSLISANTEKAGEYNC SNTVDRCSISEGSAKG LAYNIIDVVISTDLKS VETAGGNGNDSHHG EILPYDCKATVKEKDG NSDREHPHERSTRL ERLRSRKSGBKDENG DGKDISHAMTQFLES FILKGPPSPVEKAYCSG NADASNPDTHNYTS NEEANDLKQFLCKISK NFGPHHIGYMLLEE
A0A317YDT9	ZmCABIN1	Z. mays	CABIN1	
				MVEESASASEASVIQ TLTEPATEIAQTLEPN LASIEATVESVQGGT ESTCNDANNNNNA DSAAATEVCDEEREKTL EFAEELTEKGSVFLKE NDFAEAVDCFSRALEI RVAHYGELDAECINA YYRYGLALLAKAQAE ADPLGNMPKKEGEV QQESSNGESLAPSVV SGDPERQGSSSGQEG SGGKDQGEDGEDCQ DDDLSDADGDADED ESDLDMAWKMLDIA RVITDKQSTETMEKV DILCSLAEVSLERDIE SSLSDYKNALSILERLV EPDSRRTAELNFRICI CLETGCQPKEAIPYQ KALLICKARMERLSNE IKGASGSATSSVSEI DEGIQSSNVPIYDKS ASDKEVEIGDLAELAE DLEKKAT
NP_974699.1	AtNASP	A thaliana	NASP	

XP_006856185.1	AtrNASP	A. trichopoda	NASP	AVVVEMAAEPSASV DGALSKGMDLGSQV GEASGSQGIEGSSEK ALEFADELMEKGREA IRSDDFVEAVDCLSRA LEIRVQHYGELAPKCA MSYYHYGRALLYKAQ EESDPLNVSVTMAKK AASSSGPGKMDIDLS KGKGVANGEASTAS DGLKDSSNKGKESEG SNGNEENDESSEDD GDGVEDDQGVDTGE EEEESDLDSWKMLD IARVIEKNPVDTLKVV DVISTLGDVSLEREDF ETSLGDYFKALSILESL VEKDNRLQALFNFKI CLAMQLADRPADAV KYCEKMSVCEARLQ RLTSEVSESTAKTTSV DNESFSVALEEDITKS SEGNDSTSSVQEKV SEIETLKMILLVDLKEKL EDLQQMMSSLTNPLI QATEKLSTIAEKAKE DSSMSSRMGTETCG FDSPKLTITTTTSTKE
NP_496380.1	CeNASP-1	C. elegans	NASP	MDTENIADASDIRVK DASGDSDEKNGTTT EEETVEQKEKRLAELL AAGRRLKVNIDDKA SDSLSEATELSSEIYGE NHENTFDSLYYYGMA TLELAKEESQLLKGPG EKESGDEEQAGNSDD KTDEENGETEKEDEGE ESGEEEDDDDTMKL SWEILETARCIAAAKI EALAEQSGISAIEEW NLKLADVLVLLGEHGI SDGKYTQAFEDLDRA LNIQRNVLPSSRKIA QTYLIGNACASDANY DETVQYFGKTKDVLIA RQTELKHELARGVDD KEKKSEFENELKELEE MMPGVEEMIADAV HSAAQVEETKKAICA QFEGFTQVLAKLPQE AGDQKEANDISSLVR RPAKRAVDAPTDNQ AVKKEKEEGTTSI
NP_506298.1	CeNASP-2	C. elegans	NASP	ETIAKNVDSTGEEKVA MTNVAEETTPTSDSA TTNEKTASENSPTEA AASDVKDVAEASTGA QETNTSSETPAVVED EKASMEVSIGTSIEDS ETSQKEIEGELVVPEN PVEEISPEQKRKSLSE NIIAGRRFMRTDVFD KAVDVLVSAASIAAE VFGDAHEDTFEANFL YKGALLEVGKLEDRLV ANALTDVPKMAEGE EEVDQDGVENPEDVP QDERAEIKQVEEAL GVASEEPETVADEAV KTEQKEAEEDSVEKD VENSDEQNQEVEEVV ENEATVDPTEDVEME GVEEVNDEKEVGQT AEVDGEAEEDDEAED DEADSMKLAWELLET SRCIADKKAASLAAES TVDGEAIKMWKLNL ADVLTSLSGEHGIADSK YEQAQKDLTEAISQIT VHLPATSRVIANTHHL LAKAFSSDSLFEQAAA

				ATTTTGTKKDEKSKVE IGEAGEGTFDEYEE GEEEGEEGEEEEEE EEEKDKETIIKSAEGN EKAGDLLMKEKDYSS AGLYCVALKNLSSY GALSPKVAHLYVKYG NSLILTYKQESENLM NLLGSKIKEIAEKGED STTSSTPPQTTTTTT TTTTTTTTTPTEDA PDKNASDKKNSKKEE EESEDISTEESKPVNK LNTKENEDDEATPIQI AWEVLETARVIYEKEK DKQIELSDVHVSLGDL FIDSDQIENAIIEYDK ALRIRKDIRDNDLRSI AEIHYAMGLAFHVD NKAPIAEKHRTAIDL LGDQLNKIQQLNGE TDTDKLSLADEIQEL KNIIVDLQGRLEVP IDTSNEPVPKELQQD EQQPVFSTASQNP KPGTPVNTFGVVGKS ARRSITPQTNNSSSS TPSTPNKFTPIETTTT
XP_629898.1	DdNASP	D. discoideu	NASP	
				VSSPSKTVAPEPVA DTPDNAPAVSTEGS GKAEQERAEKILKKE LFSQGSRNFLVKSDE AADELQVCQLYEEV YGELADELGQPLLYA KALIAMALDENKVID VPDEAADDDEDVD DDEESAEDGAAKKE EKKDTKEAANGASS NGKELDTKEGSDEA DSTGEAEQAQSDK SKKVPTGVDEVSSN GGGGAAVNDDERPS TSNGEVTASCSNGAA PAVEEPEEEEGVSGS LQLAWEILEAAQIFS RQGLSGLPYLAEVQT ELANIEFFENGILEAAR EDYEKALKIHGELPTR NRRALAEHYKIGLTY LMQQLNKEGATLR QSSVJEEIEAIEKGD EPSERDRNNMLDLEE TKQEILAKIQEIEEMQ AQTIAEVRAALDSYIK PMSSGDAAAASSSSS SSANGAASSSSSSSKG
NP_001246993.1	DmNASP	D. melanoga	NASP	
				MPEETGATSTAERM EEKPCSSSTGDSSVDV AEEAKKLIGTGSRLV MGDVVSASVVFQEA CAMLAEKYGDTADE CGEAFFFCGKALLELA RMENTVLGNALEGV PEESSEEGEKQDDSKI ESADNLDGDDDGDE DEDDDDAEGDAKD KESEEDVGNLQLAW EMLEVAKVIYKRKDD KEDQLMAAQIHLKLA EVGVESGNYSQALED FNECLTLQKHLPSHS RLLAETHYQLGTTYSY TTQYNQAIEHFNSIK VIESRLAMLQEVIDKA EGEDSAKEEGEFEEEL KQLLPEIKEKIEDAKES QRTAAAASEAIHQTL AGASTSAFPTENGG PSSSTASQ
E9QE69	DrNASP	D. rerio	NASP	

NP_001182122.1	HsNASP	H. sapiens	NASP	LVSADKIEDVPAPSTS ADKVESLDVDSEAKK LLGLGQKHLVMGDIP AAVNAFQEAASLLGK KYGETANECGEAFF YGKSLLELARMENG LGNALEGVHVEEEG EKTEDESLVENNDNI DEEAREELREQVYDA MGEKEEAKKTEDKSL AKPETDKEQDSEMEK GGREDMDISKSAAEP QEKVDLTLDWLTETS EEAKGGAAPEGPNE AEVTSKGPEQEVPA EEEKSVSGTDVQECC REKGGQEKQGEVIVS IEEKPKVSEEQPVVT LEKQGTAVEVEAESL DPTVKPVDVGGDEPE EKVVTSENEAGKAVL EQLVGQEVPPAEESP EVTTEAAEASAVEAG SEVSEKPGQEAAPVLP KDGAVNGPSVVGDDQ TPIEPQTSIERLTETKD GSGLEEKVRAKLVPS QEETKLSVEESEAAAG
NP_001074944.1	MmNASP	M. musculus	NASP	VSADKIEDAPAPSTS DKMESLDVDSEAKKL LGLGQKHLVMGDIPA AVNAFQEAASLLGKK YGETANECGEAFFY GKSLLELARMENGVL GNALEGVHVEEEEGE KTEDESLVENNDNVD EEAREELREQVYDAM GEKEAKKAEKSLTKP ETDKQESEVEKGGRR EDMDISEPEEKLQETV EPTSKQLTESSEEAKE AAIPGLNEDEVASGK TEQESLCTEKGKSISG AYVQNKEFRETVEEG EEIISLEKKPKETSEDQ PIRAAEKQGTLMKV EIEAIDPQVKSADVG GEEPQDQVATSESEL GKAVLMELSGQDVE ASPVVAAEAGAEVSE KPGQEITVIPNNGPV VGQSTVGDQTPSEP QTSARLTETKDGSSV EEVKAELVPEQEEAM LPVEESEAGDGVET KVAQKATEKAPEDKF
A0A2K1IFP3	PpaNASP	P. patens	NASP	TQGSAAVDLEEAQRL FDDGCEFMHNEAFE DAAECFSKALEIRVRH FGELAPECATTYFKYG CALFDKLQEEADPLG ETVNPNPALPSSHPE GSSGKEPEVTGKDPV EEAGGDDDDGDEDDV MGEGEEGKEEEEGDL EDAWKMLEFARVIHE KLSRTIEAVDIITKLA DVNCFKENYEACYED YAKALEMLKQIVEPD SRILAEYFKIAIAKQL DLKPKEALIYCSNAVA VCEARVQRLKHEINN VHTEGVDALGTSEAA SVVDVEHSEPQTNGS MKATEHDNDVINP PTEEKSAESKVNEIK DIEDLLVDLKDKEQEL KEMASAPTLVEQLQA ANPEAAASMKQLFSL AAGSQASDGGSSSS TADGIKTSNGFDQPT LANFTSSTAVVTDLG VVGRGIKRAEPVAVV PAGIVNSSQPSKKRSF

-	PpNASP	P. polyceph	NASP	NTTTTTTTTSTTSPO ETAVEAGASSEPKEL TPEEAAKYEEAKLF NEDKYAEASDLFADV LTHKVAVHGLGLEC AECYFLYGKCLFLQAK AESSVFGPTAQKIN EDDEDSSDEEDVPPV GEGGEKTEGKELEEK PEDVENTQVAVETLE VARLIVEKHPEKAAEL SDTRYALAEIFLETDN CNDIAIEEFKCCVSLRE SIFVPNDRRIAEAHYH LGLAYFSTDRPADAH FHYSKAADVLDKSLKS LPPGSAEIEDIKSILED LAAKMEEVDPSLLAK EKEATENAINGTNG KGKAPVSSSSSAPPS SFPKEETGFAAPIFDK TQPVTRFGVVGSSVK RATPVTASSSSSSSS SSSSTPSAFQJASASS SSSSSSTTPSPAPRP ALSSSGEKRKLEDIMG GEEGFERTKTQDTV CETRT
NP_013078.1	ScHif1	S. cerevisiae	NASP	MKLRAEDVLANGTSR HKVQIDMERQVQIA KDLLAQKKFLEAAKR CQQTLDLSPKDGLLP DPELFTFAQAVNYM EVQNSGNLFGDALLA GDDGSGSESESEPES DVSNGEEGNENGQT EIPNSRMFQDQEEE DLTGDDVDSGSDSDSG EGSEEEENVEKEER LALHELANSFPAHEH DDEIEDVSLRKSFG HIYFENDLYENALDLL AQALMLLGRPTADG QSLTENSRLRIGDVVIL MGDIEREAEAMFSRAI HHYKALGYKTLKPA EQVTEKVIQAEFLVCD ALRWVDQVPAKDKL KRFXHAKALLEKMT TRPKDSELQARLAQ IQDDIDEVQENQQH GSKRPLSQPTTSIGFP ALEKPLGDFNDLSQL VKKKPRRH
XP_001030823.3	TtNASP	T. thermophi	NASP	EITKEQKILDKEENG QNNLNDEGSRNQEE AEEITKILKESSKVSE KYLESRLEAVECYKN KNFDQAENIFSDLMI YAVKFYKSETHVKMS KYFFFGDMLLSKLEN NPDVFGDKAQKKLD NQTTVEVEKNPNQM ASEEASEQKMVAQK EQKEEEDNEESQS QGEEFIVQLDTVIFKD ENVEIRDETAPQNN EENQTEENNLPIGKV TMEKSDPIDDLQLA WENLESRKIYNQELL QLQDSQEDRKEKQI QVKKQIVEVTLRIADL EQWRDNFSDALKEYT MALQLSKEIEDREYSR TLSSIFFQLGNAVLYE NKEYCEEEGLKNFDS AQILENCLSKIAIPPQ ELPSFKILKKDDIKVD HLKHSFLDDDESKEIK DTLKVLYEKIEDTLIQR DERDKYKAYQELMKK QSQNQNQFAQPSQ

				DTATAPSTSADKADG MDIDTEAKRLMGAG QRHLVMKDVRSVTS FQEASSLLAKYGETA DECAEAFYSYGTSLLE LARMENGVLGNALE GMPEDDEEGEKEEDA NIPADNLDEKEREQL RDQVYDAMAEERA PKETSESEALGKPKDE SKPMDTDEQNKPVE DKMKNGNMETGKV TDGLKLESVNRDVSIM DESEKGEVPESKNAP TPGEHTEQKEKSEPKE SKAPEKKDTKDTANQ SPDSTEVAEERAHTD TAEKMDSEASAKTEE GASKTPTGNEASKSE DPEKMGEEEEEGEDS EESEENESTEEKETED EDVGNLQLAWEMLD LSKMIFKRQKSREAQ LKAAQAHKLGEVST ESENYAQAVEDFLAC LDIQKEHLEEHDRLLA ETHYHLGLAYQYSSK HEEAISHFTQSVGVIE
NP_001089907.1	XINASPL	X. laevis	NASP	
				PQAPPPPDTEPTAA PGEAEDEPQTLERAQ ELFDRGSKAIEDEDV DAVDCLSRALAIRASH YGEIAPCASTYFKYG CALLYKAQEEEDVLG NVPKSLPNEDSVKST ATKDDSGTSKVSMTD AEDAASSEKVDAAEG HNSNGKNQENGNG EVEKDDDDVDDEK MGDEEDNDLDSWK MLDIARAIVEKSQDN TMEKVKIYSALAEVAT EREDMDNSLSQYMK ALSMLEHLVEPDHRR VVLENFRLCLVYELVS KIGEAIYCAKAIKSLCK SRIQSLKSSKDALLAGI DGDASAAEAEAGGSEK SAIEKELEQLTSILPDL EKKLEDLSEANPSAD MDENVKAIVSRVTD VMPKAASFSSQMA TSSNGFDSSVMSTAE ASNELASSRLLNLLLR NLHLIHHLNKVTAAS TLRLFLQRRLETSQYLS
A0A804MWY6	ZmNASP	Z. mays	NASP	

Table S4: List of primers

Primer ID	Sequence	Gene
HIRA_F	GTGAATGGATTGCAGGTGCC	HIRA
HIRA_R	GTGGGAGGGATGAACACAGG	HIRA
NASP_F	GTACCACCAGTTGGTGAGGG	NASP
NASP_R	CCGCTTTTTCAGGGTGCTTC	NASP
CAF1C_F2	ATTTTGCTTCTGCTGGCGAC	CAF1C
CAF1C_R2	AAACGGGTAAAAGCCACGC	CAF1C
19S_QF	CAACGGGTACAGAGGATAAGGG	19S
19S_QR	ATTCATATGTCCCCTAACGCCC	19S

Table S5: Prediction of NLS signals in chaperone proteins.

Physarum Histone chaperone	Probability	Start	End	Signal Sequence
PpASF1				This protein has no NLS predicted by NLStradamus
PpMCM2	0.792	115	125	KKRNRKVQAK
	0.744	148	154	KPRRRP
	1.000	378	399	KKKKKKKKKKKKKKKKRSK
PpHSP90	0.879	244	251	KKKKKKVK
PpHSC70				This protein has no NLS predicted by NLStradamus
PpIPO4				This protein has no NLS predicted by NLStradamus
PpSPT6	0.998	63	74	REKKKKKKKKR
	0.996	122	136	GGGGERRRRKKRR
	0.706	171	179	NKNSKLLR
	0.643	240	245	KRSKQR
PpAPLF	0.837	25	35	KKEKEKEKKKG
	0.773	205	228	RGEKGEEGEGEESPEKKGVRK
	0.911	309	347	KPKTPKSKNSTPEKDAEIAKLQEKYNKESGGGGKKKRK
	0.889	362	372	KVKKQKKGKKG
	0.961	405	419	KNKKNKKKGASPKP
PpPolE3	0.743	114	124	EKKEKGDK
PpSWR1	0.808	1567	1575	REKRERKKH
	0.817	1590	1618	MPTKGGRGRKANGGLFPSRIVPKTARDGK
	0.779	1980	1992	KPSPSKRSKNAE
PpSWC2	0.992	79	107	KEKKEKKRMAYVDPKAKKGPKRRKIQDKA
	0.990	137	155	APKAEPKKGRRRKKTPK
	0.893	200	219	REKEREETEKKRSQRSRKRS
PpSPT16	0.744	494	509	QKGSKGSSEDKIKKHQ
PpPob3	0.996	503	528	KPKKDKHEGKRKKSDGEEKPKKKKA
PpSET				This protein has no NLS predicted by NLStradamus
PpNAP1L1	0.964	231	245	KQTKKTKAKGKKPA
	0.685	345	353	KRGGRQGG
PpCAF1A	0.994	143	264	KPSPKKRKLSPEEKQAREDAKKAKEEERKLQAKQREDAKNKKEEEKRLLALAREEEKKRDEEKRVKDQAKEEARLKKEAEKKARDDAKLQKDLEKKAKEEARLQKLQEKQKQEEEEERKKE
	0.982	345	359	GKRNKRDKKRAEK
	0.850	499	513	HMERDGTKREKREKK
PpCAF1B	0.677	194	201	LPKNKKSH
	0.704	283	296	EKEKEESGKGKEKE
PpCAF1C				This protein has no NLS predicted by NLStradamus
PpHIRA	0.999	599	669	TKKHKKDEESDKEKGKEAEKGDKNKDIEADHTKEKEREKERDKREKERRRREKEKEKEKVRKEKEKERE
PpUBN	0.717	190	194	RKKKR
	0.999	204	283	DNKSKRTKKLIEDEKGSPEKDKKNDENGKSKDKKPKDKKKEETPKDKEKKEDSTKDEKKKEDPGKEKEKKKEEK
	0.605	525	525	K
	0.981	615	644	KIKKEKAKNGEKAAGPGGAKPAKVYSRKK
PpCABIN1				This protein has no NLS predicted by NLStradamus
PpNASP				This protein has no NLS predicted by NLStradamus

Table S6: Prediction of NES signals in chaperone proteins.

Physarum Histone chaperone	Score	Start	End	Signal Sequence	number of matches for the NES
PpCabin1	100.0	766	778	QT-L-GVS-L-ST-L-N-L-Q	1
PpIPO4	75.0	566	581	DA-M-SLA-L-AG-L-E-L-GFSE	1
PpSPT6	73.6	1216	1230	EE-L-KK-L-NKF-L-V-I-ELE	4
PpCAF1A	70.2	522	534	SS-A-KAI-L-TN-L-S-I-O	2
PpSWR1	64.9	1764	1779	DV-L-AL-I-SQR-V-V-L-TMPQ	1
PpSPT6	61.1	634	647	Q-M-ED-F-QML-L-V-L-KAE	1
PpSPT6	61.1	959	972	S-L-QY-V-SG-L-G-V-RKAE	1
PpSPT16	56.2	839	854	SL-I-EPP-F-FV-L-T-L-KEIE	2
PpSPT16	56.2	106	121	EN-A-PYV-L-CF-L-Q-L-SKAE	1
PpMCM2	56.2	557	568	S-A-VG-L-TAA-V-R-I-D	1
PpSPT6	56.2	1308	1319	S-A-RS-L-GS-L-F-I-GE	1
PpCABIN1	51.8	157	170	S-I-DK-L-AEV-L-F-I-LGD	1
PpPob3	51.8	84	97	DV-L-KTF-L-SEN-F-H-V-E	1
PpNASP	51.8	204	217	DA-I-EE-F-KKC-V-S-L-RE	1
PpMCM2	49.2	262	274	Q-I-FDE-V-AMK-L-V-L-E	1
PpSPT6	48.7	474	489	ED-L-SD-L-QDY-F-Q-M-FYAD	2
PpPob3	46.4	114	125	EL-A-GP-V-IT-L-A-V-D	1
PpMCM2	43.9	878	891	QL-L-FH-I-LT-M-E-V-REO	2
PpSPT16	41.2	324	336	S-I-KPG-V-KIS-V-I-M-D	1
PpMCM2	39.8	656	670	DP-I-LSR-F-DI-M-C-V-VRD	2
PpCAF1A	20.3	6	18	S-I-YLC-L-VWA-V-F-M-D	1
PpASF1	NA				
PpSET	NA				
PpNAP1L1	NA				
PpSWC2	NA				
PpPolE3	NA				
PpHSP90	NA				
PpHSC70	NA				
PpCAF1B	NA				
PpCAF1C	NA				
PpHIRA	NA				
PpUBN	NA				

Table S2: RNA-Seq data for histone chaperones during the cell cycle. Normalized read count are presented in TPM.

Physarum Encoded Protein	Physarum Transcript ID	Average early S-phase	SEM early S-phase	Average mid S-phase	SEM mid S-phase	Average late S-phase	SEM late S-phase	Average early G2-phase	SEM early G2 phase	Average late G2-phase	SEM late G2-phase
PpASF1	c_ppoly_21148	1685.64018535042	382.571302686029	1059.9304195573	492.595273451016	362.802070527273	108.002266976866	524.77419380426	121.605051084244	1935.94237626234	220.851430113619
PpMCM2	c_ppoly_24879	348.316651403373	107.612867406748	30.736854935386	6.7519544495917	18.0262054632708	9.33210563139798	19.6441524563809	8.78750018426791	320.298470991307	83.2392746707242
PpIPO4	c_ppoly_21477	500.0753411145	155.059505888504	285.169396424205	125.5775857642	300.045891422379	44.5007710355702	440.72913492516	226.920583929418	320.56752056917	99.5996768964357
PpPT6	c_ppoly_23850	946.221323399772	267.782426259667	801.75039129217	196.592902112129	796.128154000321	225.869751198728	883.769437899125	418.910375801823	535.910510796423	108.739600667991
PpPolE3	c_ppoly_20662	498.946034384142	84.3251390005563	369.130590586745	20.427037370197	38.8726987573958	19.8434389071508	60.3746550698003	16.3747911981762	730.295813248133	202.149890161534
PpHSP90	c_ppoly_11600	20007.6443158073	6426.45272292963	9734.37029857941	5156.02179431029	15770.1416854257	265.79873656357	8178.16622784559	1159.11881355825	15918.9721285622	4950.54779397259
PpHSC70	c_ppoly_14645	23225.7456694507	3247.06191059978	9646.97444768135	3701.48661572195	24989.3498155015	9841.83672566814	17255.8499772107	5068.55676580514	18211.4644589064	3051.81467142549
PpNASP	c_ppoly_25875	1629.35403343165	620.414116060328	363.899319613087	209.842458800485	100.358580632022	49.2779416872384	135.875297632927	63.9403874024608	922.890744802152	440.673737748532
PpCAF1A	c_ppoly_04687	597.254804471717	120.092250699239	62.1585607690401	23.4137215227569	26.8581363237992	13.1076862635872	28.3681013175196	14.1973006626964	434.341742433927	99.4410586715044
PpCAF1B	c_ppoly_07527	750.710058213592	233.916619770675	183.331866595398	79.089799099446	14.2463273177883	7.9155809639024	6.39873554427367	2.74463936330242	493.114229648283	97.6520207782227
PpCAF1C	c_ppoly_21398	2501.2459346085	578.127883384376	3371.9451619237	764.048482179825	1943.08803111066	509.35511092029	2459.11637656382	271.929299332971	2379.00764440829	288.673270974413
PpHIRA	c_ppoly_12971	989.598820352507	140.79766480616	1160.8239286344	197.737924512506	1212.72394731905	452.371827102978	1134.73843959412	334.63691041208	760.968307356968	70.766415233273
PpUBN	c_ppoly_06650	378.743171268866	88.0311047036758	688.542276271353	26.190024394418	682.568343394393	262.45494895959	509.773405608509	70.615779913131	539.487296205266	90.4127318057242
PpCABIN1	c_ppoly_25516	305.317386411327	108.982428985004	123.793553723011	34.3114249875473	87.3060935803543	30.8516180968159	113.957483964267	47.5692809749332	182.024985513339	30.267088246206
PpPLF	c_ppoly_14659	602.253242194539	117.392752510644	696.635847592961	119.153243588837	567.470471090983	141.273674480576	594.889695389578	192.973403903031	463.49685999437	45.6443812599786
PpSWR1	c_ppoly_23444	646.256991190687	44.8386018649976	524.264004191157	114.675703587593	750.759297095279	200.595591582832	809.632893598996	283.378733605977	445.385395046608	51.3320997499832
PpSWC2	c_ppoly_14213	76.3530595685763	12.1138778936789	287.056516156046	60.1224577135304	192.637043757318	24.4457821070988	121.082884673759	7.89472148236264	104.872094197058	8.93252644801636
PpSPT16	c_ppoly_05072	1636.36073180243	303.322016727763	1199.89815004542	384.411533528417	599.80583751488	39.3030725590596	882.275887448549	206.48198026475	1607.57888712835	223.137688211474
PpPob3	c_ppoly_22049	2229.8086235541	570.071074636041	834.499671535581	296.684402950269	485.228384979042	144.516443107806	908.926892518441	279.746447641578	2779.57096039765	400.016961261955
PpNAP1L1	c_ppoly_06693	13756.1314103008	3384.84587293737	15303.6934508884	3841.67850719629	13997.2210675457	2383.26933024534	13826.715003288	3013.02502051686	10769.7059192618	2239.71095674852
PpSET	c_ppoly_06196	3864.36061664834	508.320802019727	7191.29184354547	2647.07569765998	4959.35118588884	732.62386714741	4791.89302691023	777.030852241146	4052.06069090031	209.40628677767
PpHT01	c_ppoly_11752	224.696253946371	64.0543175847619	82.1480278004567	36.0232191739293	56.8407614086324	17.1144801265771	31.3177687480039	14.1154140043007	201.232068992229	82.272118714642
PpHTA2	c_ppoly_09597	915.261558275856	157.750403876033	597.704488201613	74.6491583767923	540.232443693337	71.0450192193106	478.278797518563	94.3203004094633	919.533399670732	293.589766273955
PpHTA3	c_ppoly_08402	12768.2104766242	1764.19512439209	19699.028596941	8527.599947596	6298.59967750296	1310.2571828244	5418.08853376133	756.639194212458	13212.4754814818	1176.4678460523
PpHTB1	c_ppoly_07519	516.711905099623	73.1012567019167	286.04909909138	59.1602429755268	311.385778068313	141.223766772719	247.940835274094	72.7312655722548	437.695433961764	144.971408575862
PpHTB2	c_ppoly_07520	19591.0487732844	3852.84466629448	19477.6081457671	7330.41880876511	5953.18640798085	1305.43618961695	6995.68294236743	947.448208448614	29290.2210913313	9024.79107813862
PpHTT1	c_ppoly_23074	362.936493621585	30.7899100607491	184.837175371919	54.7652150451114	175.238136352701	72.839740023364	78.117185948318	8.82195351238481	343.537584809706	114.774352425865
PpHTT2	c_ppoly_07595	245.457643975657	71.5039049378998	228.797500426871	64.5931817164329	170.908285886344	106.117662637738	509.168664755438	289.786224425149	552.733119862111	246.6473716141
PpHTT3	c_ppoly_07042	813.285604922061	93.9522087593013	354.217927444129	51.6391828541076	311.668949854076	106.117662637738	509.168664755438	289.786224425149	552.733119862111	246.6473716141
PpHTT4	c_ppoly_00698	9895.6108639726	724.633057838997	19329.1179871322	5883.7362724994	12788.8950614719	1092.30683795137	11143.0920774536	466.57600829916	9981.80779712691	319.802030472577
PpHTT5	c_ppoly_10421	29.3192111194444	8.99358097544886	176.432671524867	37.6843272724607	84.3202148809206	9.232765909262	34.7932332606828	9.39613097875041	41.1153001275873	7.74354993753454
PpHTF1	c_ppoly_07791	3084.03601807698	576.640520811421	3191.90476212654	212.242124852856	3038.563595261	1064.27838964211	2143.55897320622	553.386093720452	4777.70786964036	1177.45821282105
PpHTF2	c_ppoly_15298	771.495274942455	189.771049071911	574.531486893002	50.5536608956485	548.0011158296472	39.0745859034198	389.985749592906	47.1831650477286	844.051363434518	374.735789162973

Table S8: RNA-Seq file details.

Sample Name	cell cycle phase	Sequence 17	Sequence i5	R1 file Name	R2 file Name
S4	early S-phase	TCCGGAGA	TATAGCCT	S4_TCCGGAGA-AGGCTATA-AHFMH7DRXY_L001_R1.fastq	S4_TCCGGAGA-AGGCTATA-AHFMH7DRXY_L001_R2.fastq
S13	early S-phase	TCCGGAGA	ATAGAGGC	S13_TCCGGAGA-GCCTCTAT-AHFMH7DRXY_L001_R1.fastq	S13_TCCGGAGA-GCCTCTAT-AHFMH7DRXY_L001_R2.fastq
S24	early S-phase	TCCGGAGA	CCTATCCT	S24_TCCGGAGA-AGGATAGG-AHFMH7DRXY_L001_R1.fastq	S24_TCCGGAGA-AGGATAGG-AHFMH7DRXY_L001_R2.fastq
S6	mid S-phase	CGCTCATT	TATAGCCT	S6_CGCTCATT-AGGCTATA-AHFMH7DRXY_L001_R1.fastq	S6_CGCTCATT-AGGCTATA-AHFMH7DRXY_L001_R2.fastq
S16	mid S-phase	CGCTCATT	ATAGAGGC	S16_CGCTCATT-GCCTCTAT-AHFMH7DRXY_L001_R1.fastq	S16_CGCTCATT-GCCTCTAT-AHFMH7DRXY_L001_R2.fastq
S7	late S-phase	GAGATTCC	TATAGCCT	S7_GAGATTCC-AGGCTATA-AHFMH7DRXY_L001_R1.fastq	S7_GAGATTCC-AGGCTATA-AHFMH7DRXY_L001_R2.fastq
S18	late S-phase	GAGATTCC	ATAGAGGC	S18_GAGATTCC-GCCTCTAT-AHFMH7DRXY_L001_R1.fastq	S18_GAGATTCC-GCCTCTAT-AHFMH7DRXY_L001_R2.fastq
S28	late S-phase	GAGATTCC	CCTATCCT	S28_GAGATTCC-AGGATAGG-AHFMH7DRXY_L001_R1.fastq	S28_GAGATTCC-AGGATAGG-AHFMH7DRXY_L001_R2.fastq
S10	early G2 phase	ATTCAGAA	TATAGCCT	S10_ATTCAGAA-AGGCTATA-AHFMH7DRXY_L001_R1.fastq	S10_ATTCAGAA-AGGCTATA-AHFMH7DRXY_L001_R2.fastq
S20	early G2 phase	ATTCAGAA	ATAGAGGC	S20_ATTCAGAA-GCCTCTAT-AHFMH7DRXY_L001_R1.fastq	S20_ATTCAGAA-GCCTCTAT-AHFMH7DRXY_L001_R2.fastq
S29	early G2 phase	ATTCAGAA	CCTATCCT	S29_ATTCAGAA-AGGATAGG-AHFMH7DRXY_L001_R1.fastq	S29_ATTCAGAA-AGGATAGG-AHFMH7DRXY_L001_R2.fastq
S2	late phase	ATTACTCG	TATAGCCT	S2_ATTACTCG-AGGCTATA-AHFMH7DRXY_L001_R1.fastq	S2_ATTACTCG-AGGCTATA-AHFMH7DRXY_L001_R2.fastq
S12	late phase	ATTACTCG	ATAGAGGC	S12_ATTACTCG-GCCTCTAT-AHFMH7DRXY_L001_R1.fastq	S12_ATTACTCG-GCCTCTAT-AHFMH7DRXY_L001_R2.fastq
S22	late phase	ATTACTCG	CCTATCCT	S22_ATTACTCG-AGGATAGG-AHFMH7DRXY_L001_R1.fastq	S22_ATTACTCG-AGGATAGG-AHFMH7DRXY_L001_R2.fastq