

## **Additional file 10**

### **Protein-level alignments of intron gain retrogenes (“Sbjct”) and their parents (“Query”) by GeneWise [S1].**

Each alignment shows the protein sequence of the parental gene on the first line, followed by a line indicating the similarity level of the match followed by a line representing the translation of the DNA sequence of the retrogene. If the intervening sequence of the retrogene is long, it is indicated in the square brackets. If the intervening sequence of the retrogene is short and not divisible by 3, it is indicated by '!'. (for specifics, see the description of the output format of GeneWise (<http://www.ebi.ac.uk/Tools/Wise2/>)). Red downward arrows indicate the exon-exon borders of parental genes and the blue upward arrows and connecting lines mark the new intron in the retrogene. In *XXyac-R12DG2.2*, the retrogene is also annotated to be embedded near the 3' end of an ncRNA gene candidate (LOC100190939, Additional file 12). In *CSMD3*, the green upward arrows and connecting lines mark the intragenic region in the retroposed sequence. The retrogene is an alternative translation start point (Additional file 12). In *WBP2NL*, the retrogene is located in its first intron, but could be transcribed at sometimes (Additional file 1 and 2). Besides, the intronic region at *XXyac-R12DG2.2* and *AC019016.1* could be spliced in two patterns.

## References

- S1. Birney E, Clamp M, Durbin R: **GeneWise and Genomewise**. *Genome Res* 2004, **14(5)**:988-995.

**TMEM14D**

Query 1 MEKPLFPLVPLHWFGFGYTALVVS<sup>↓</sup>GGIVGYVKT<sup>↓</sup>GSVPSLAAGLLFGSLA  
MEKPLFPLVPLHWFGFGYTALVVS<sup>↓</sup>GGIVGYVKTG PSLAAGLLFGSLA  
Sbjct 2 MEKPLFPLVPLHWFGFGYTALVVS<sup>↓</sup>GGIVGYVKTGRAPSLAAGLLFGSLA

Query 50 GLGAYQLYQDPRNVWGFLA<sup>↓</sup>ATSVT<sup>↓</sup>FGVMGMRSY<sup>↓</sup>YYGK<sup>↓</sup>FMPVGLIAGAS  
G+GAYQLYQDPRNVW FLA<sup>↓</sup>ATSVT<sup>↓</sup>FG+MGMRSY<sup>↓</sup>YYGK<sup>↓</sup>FMPVGLIAGAS  
Sbjct 149 GVGAYQLYQDPRNVWDFLA<sup>↓</sup>ATSVT<sup>↓</sup>FGIMGMRSY<sup>↓</sup>YYGK<sup>↓</sup>FMPVGLIAGAS

Query 99 LLMAAKVGVRMLMTSD  
LLMAAKVGVRMLMTSD  
Sbjct 296 LLMAAKVGVRMLMTSD

**RPS3AP5**

Query 1 MAVGKNKRLTKGGKKGAKK<sup>↓</sup>VVDPFSK<sup>↓</sup>KDWYDVKAPAMFNIRNIGKTLV  
M V KNKRLTKGGKKGAKK<sup>↓</sup>VVDPFSK<sup>↓</sup>KDWYDV APAMFNIRNIGKTLV  
Sbjct 2 MVVDKNKRLTKGGKKGAKK<sup>↓</sup>VVDPFSK<sup>↓</sup>KDWYDVNAPAMFNIRNIGKTLV

Query 50 TRTQGT<sup>↓</sup>KIASDGLKGRVFEVSLADLQ<sup>↓</sup>NDEVAFRKFKLIT<sup>↓</sup>EDVQGNCLT  
TRTQGT<sup>↓</sup>KIASDG RVFEVSLADLQ<sup>↓</sup>NDEVAFRKFKLIT<sup>↓</sup>EDVQGNCLT  
Sbjct 149 TRTQGT<sup>↓</sup>KIASDG - - -RVFEVSLADLQ<sup>↓</sup>NDEVAFRKFKLIT<sup>↓</sup>EDVQGNCLT

Query 99 NFHGMDL<sup>↓</sup>TRDKMCSMVKKWQ<sup>↓</sup>TMIEAHVDVKT<sup>↓</sup>TDGYLLR<sup>↓</sup>LFCVGF<sup>↓</sup>TKKRN  
NFHG+DLT DKMCSMVKKWQ<sup>↓</sup>TMIEAHVDVKT<sup>↓</sup>TDGYLLR<sup>↓</sup>LFCVGF<sup>↓</sup>TKKRN  
Sbjct 287 NFHGVDL<sup>↓</sup>TSDKMCSMVKKWQ<sup>↓</sup>TMIEAHVDVKT<sup>↓</sup>TDGYLLR<sup>↓</sup>LFCVGF<sup>↓</sup>TKKRN

Query 148 NQIRKTSYAQHQQVR - - - - - QIRKKMMEIMTREVQ<sup>↓</sup>TNDLKEVV  
NQIRKTSYAQHQQV QIRKKMMEIMTREVQ<sup>↓</sup>TNDLKEVV  
Sbjct 434 NQIRKTSYAQHQQV!KSGRPLMLSTS<sup>↑</sup>QIRKKMMEIMTREVQ<sup>↓</sup>TNDLKEVV

Query 186 NKLI<sup>↓</sup>PDSIGK<sup>↓</sup>DIKACQSIYPLHDV<sup>↓</sup>FVRKVKMLKKPKFELGKLMELHGE  
NKLI<sup>↓</sup>PDSIGK<sup>↓</sup>+EKACQSIYPLHDV<sup>↓</sup>FVRKVKMLKKPKFELGKLMELHGE  
Sbjct 583 NKLI<sup>↓</sup>PDSIGK<sup>↓</sup>DVEKACQSIYPLHDV<sup>↓</sup>FVRKVKMLKKPKFELGKLMELHGE

Query 235 GSSSGKATGDETGAKVERADGYEPPVQESV  
G SSGKATGDETG KVERADGYE PVQESV  
Sbjct 730 GCSSSGKATGDETGVKVERADGYELPVQESV

**XXyac-R12DG2.2**

Query 1 MARGGRGRRLLGLALGLLLALV<sup>↓</sup>LAPRVLRAKPTVRKERVV<sup>↓</sup>RPDSELGERP  
MARGGR LGLALGLLLALV<sup>↓</sup>LAPRVLRAKPTVRKERVV<sup>↓</sup>RPDSELGERP  
Sbjct 2 MARGGR - - -LGLALGLLLALV<sup>↓</sup>LAPRVLRAKPTVRKERVV<sup>↓</sup>RPDSELGERP

Query 50 PEDNQSFQYDHEAFLGKEDSKTFDQLTPDESKERLGKIVDRIDNDGDGF  
Sbjct 140 PEDNQSFQYDHEAFLGKEDSK FD+LTPDESKERLGKIVDRIDNDG GF

Query 99 VTTEELKTWIKRVQKRYIFDNVAKVWKDYDRDKDDKISWEEYKQATYGY  
Sbjct 287 VTTEELKTWIKRVQKRYIFDNVAKVWKDYDRDKDD ISWEEYKQATYGY

Query 148 YLGNPAEFHDSDDHHTFKKMLPRDERRFKAADLNGDLTATREEFTAFLH  
Sbjct 434 YGNPAEFHDSDDHHTFKKMLPRDERRFKAADLNGDLTAT EEF TAFLH

Query 197 PEEFEHMKEIVVLETTLEDIDKNGDGFVDQDEYIADMFSHEENGPEPDWV  
Sbjct 583 PEEFEHMKEIVALETLEDIDKNGDGFVDQDEYIADMFSHEENGPEPDWV

Query 246 LSEREQFNEFRDLNKGKLDKDEIRHWILPQDYDHAQAEARHLVYESDK  
Sbjct 730 LSEREQFNEFRDLNKGKLDKDEIRHWILPQDYDHAQAEARHLVYESDK

Query 295 NKDEKLTKEEILENWNMFVGSQATNYGEDLTKNHDEL  
Sbjct 877 NKDEKLTKEEILENWNMFVGSQATNYGEDLTKNHDEL

### HSP90B2P

Query 8 GLCCVLLTFGSVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAI  
Sbjct 24 GLCCVLLTFGSVRA DEVDV+GTVEEDLGKS+GSR DDEVVQREEEAI

Query 57 QLDGLNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELI  
Sbjct 171 QLDLNASQIRELREKLEKFAFQAEVNRMMKLIINSLY+NKEIFLRELI

Query 106 SNASDALDKIRLISLTDENALSGNEELTVKIKCDKEKNLLHVTDGTGVGM  
Sbjct 318 !---DALVKIRLISLTDENALSGNEELTVKIKCDKKNLLHVTDGTGVGM

Query 155 TREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSELIGQFGVGFYSAF  
Sbjct 455 TREELVKNLGTIAKS T+EFLNKMTEAQEDGQSTSELIGQFGV FYS F

Query 204 LVADKVIIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLVLK  
Sbjct 602 LVADKVIIVTSKHNNDTQ+IWES+SNE+SVIADPRGNTLGRGTTITLVLK

Query 253 EEASDYLELDTIKNLVKKYSQFINFPIYVWSSKTET↓VEEPMEEEEAAKE  
 E+A DYL LDTIKNL+ KYSQFINFPIYVWSSKTETV+EP +EE AAKE  
 Sbjct 749 EKAFDYLALDTIKNLI!KYSQFINFPIYVWSSKTETVKEP!KEEAAAKE

Query 302 EKEESDDEAAVEEEEEEEKPKTKKVEKTVWDWELMNDIKPIWQRPSKEV  
 EKEES DEAAVEEEEE++KPKTKKVEKTV +WELMNDIKPIWQRPS EV  
 Sbjct 894 EKEESYDEAAVEEEEEEKQKPKTKKVEKTVCEWELMNDIKPIWQRPS!EV

Query 351 EEDEYKAFYKSFESKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFD  
 E DEYKAF KSFSKESDDP+A IHFTAEGEVTFKSILFVPT PRGLFD  
 Sbjct 1040 E-DEYKAFCKSFESKESDDPVACIHFTAEGEVTFKSILFVPTFVPRGLFD

Query 400 EYGSKKSDYIKLYVRRVFIITDDFHDMMPKYLNLFVKGVVDSDDLPLNVS  
 EYGSKKSDYIKLYVR VFITDDF D MPK LNFVKGVVDS L LNVS  
 Sbjct 1184 EYGSKKSDYIKLYVRCVFIITDDFRDTPKLNLFVKGVVDSGGLSLNVS

Query 449 ETLQQHKLKLVIRKKLVRKTLDMIKKIADDKYNDTFWKEFGTNIKLGVI  
 ETLQQHKLKLVIRKKL V KTLDMIKKIAD+KYNDTFWKEFGTNIKLGVI  
 Sbjct 1331 ETLQQHKLKLVIRKKL VHKTLDMIKKIADEKYNDTFWKEFGTNIKLGVI

Query 498 EDHSNRTRLAKLLRFQSSHPTDITSLDQYVERMKEKQDKIYFMAGSSR  
 EDHSNRT LAKLLRFQSSHHP DITSL Q VERMKEKQDKI +MAGSSR  
 Sbjct 1478 EDHSNRTCLAKLLRFQSSHHPADITSLHQDVERMKEKQDKICLMAGSSR

Query 547 KEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEG  
 KEAESSPFV RLLK+GYEVIYLTEPV EYCIQALPEFDGKRFQNVAKEG  
 Sbjct 1625 KEAESSPFVXRLLRGVEVIYLTEPVVEYCIQALPEFDGKRFQNVAKEG

Query 596 VKFDESEKTKESREAVEKEFEPELLNWMKDKALKDKIEKAVVSQRLTESP  
 VKFD+SEKTKES EAVEKEFEPL NW+KDKA+KDKIEKA+VSQ LTES  
 Sbjct 1772 VKFDDSEKTKESHEAVEKEFEPLPNWVKDKAIKDKIEKAMVSQCLTESL

Query 645 CALVASQYGWSGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPL  
 CALVASQYGWSGNMERIMKAQAYQTGK ISTNY+AS+KKTFEINPRHPL  
 Sbjct 1919 CALVASQYGWSGNMERIMKAQAYQTGKGIISTNYHASRKKTFEINPRHPL

Query 694 IRDMLRRIKEDEDDKTVLIDLAVVLFETAT--LRSGYLLPDTKAYGDRIE  
 IRDMLRRIKEDEDDKTVLIDLAVVLFETA +RSG LLPDTKAYGDRIE  
 Sbjct 2066 IRDMLRRIKEDEDDKTVLIDLAVVLFETA!AVIRSGCLLPDTKAYGDRIE

Query 741 RMLRSLNIDPDAKVEEPEEPEETAEDTTEDTEQDEDEEMDVGTDDEE  
 RM LSLN PDAKVEE P+EEPEETAED EQD+D+EMDVGTDDEE  
 Sbjct 2215 RMPCLSLNT!PDAKVEE-PDEEPEETAEDK---EQDKDKEMDVGTDDEE

Query 790 -EETAKESTAEKDEL  
 E TAKESTAEKDEL  
 Sbjct 2345 KQETAKESTAEKDEL

**HSP90AA4P**

Query 1 MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELI  
Sbjct 6 MPEETQTQDQP EE EVETF FQAEIAQLMSL+INTFY NKEIFLRELI

Query 50 SNSSD<sup>↓</sup>DALDKIRYESLTDPSKLD<sup>↓</sup>SGKELHINLIPNKQDR<sup>↓</sup>TLTIVDTGIGM  
SN SDALDKI ESLTDPSKLD<sup>↓</sup>SGKE HI+LIPNKQDR<sup>↓</sup>TLTIVDTGIGM  
Sbjct 153 SNLSDALDKILXESLTDPSKLD<sup>↓</sup>SGKEPHISLIPNKQDR<sup>↓</sup>TLTIVDTGIGM

Query 99 TKADLNNLGTIAKSGTKAFMEALQAGADISMIGQFVGFYSA<sup>↓</sup>YLVAEK  
TKADLNNLGTI KS TK FME LQAGADISMIGQF VGFYSA<sup>↓</sup>VAEK  
Sbjct 300 TKADLNNLGTITKSETKVFMEVLQAGADISMIGQFSVGFYSA<sup>↓</sup>SVAEK

Query 148 VTVITKHNDDEQYAWESSAGGSFTVVRTDTG<sup>↓</sup>EPMGRGTKVILHLKEDQTE  
VTVITKH<sup>↓</sup>N+DEQYAWESS GSFTVVRTDTGEP+G GTKVIL LKEDQTE  
Sbjct 447 VTVITKHNNDEQYAWESSLRGSFTVVRTDTGEP<sup>↑</sup>IGHGTKVILPLKEDQTE

Query 197 YLEERRIKEIVKKHSQFIGYPITLFV<sup>↓</sup>EKERDKEVSDDEAEKEDKKEEK  
YLEER+I EIVKKHS FIGYP+TLFVEK+ DKEVS+DEAEKEDKKEEK  
Sbjct 594 YLEERKINEIVKKHS!FIGYPVTLFVEKKHDKEVSNDEAEKEDKKEEK

---

Query 246 EKEEKESEDKPEIEDVGSDEEEE-KKDGD- - - -K<sup>↓</sup>KKKKKIKEYIDQEE  
EKEEKESEDKPEIEDVGSDEEEE KKDGD KKKKKKIKEYIDQ E  
Sbjct 739 EKEEKESEDKPEIEDVGSDEEEEXKKDGD<sup>↓</sup>KKKKKKKKKKKIKEYIDQGE

---

Query 290 LNKT<sup>↓</sup>KPIWTRNPDDITNEEYGEFYKSLTNDWEDHLAVKH<sup>↓</sup>FSVEGQLEFR  
LN TKPI TR PDD+TN EY EFYKSLT +WED+LAVKH<sup>↓</sup>FSVEGQLEFR  
Sbjct 886 LN!TKPICTRY<sup>↑</sup>PDDFTN<sup>↑</sup>XEYREFYKSLTINWEDYLAVKH<sup>↓</sup>FSVEGQLEFR

---

Query 339 ALLFVPRRAPFDL<sup>↓</sup>FENRKKKNNIKLYVRRVFIMDNCEELIPEYLN<sup>↓</sup>FIRG  
A L<sup>↓</sup>FVPR APF+L+E RKKKN IKL RR +IMDNCEELIPEYLN<sup>↓</sup>FIRG  
Sbjct 1032 AFLFVPRLAPFELLETRKKKNNIKLSARRDLIMDNCEELIPEYLN<sup>↓</sup>FIRG

Query 388 VVDS<sup>↓</sup>EDLPLNISREMLQOSKILKVIRK<sup>↓</sup>NLVKKCLELFT<sup>↓</sup>ELAEDKENYKK  
VVDS<sup>↓</sup>EDLPLNI REMLQ SKILKVIR NLVKKCLELFT<sup>↓</sup>EL E KE Y K  
Sbjct 1179 VVDS<sup>↓</sup>EDLPLNIFREMLQXSKILKVIRNNLVKKCLELFT<sup>↓</sup>ELVEGKERYXK

---

Query 437 FYEQFSKNIKLGIHEDSQNRKKLSELLRY<sup>↓</sup>YTSASGDEMVS<sup>↓</sup>LKD<sup>↓</sup>YCTR<sup>↓</sup>MK  
FYEQFSKNIKLGIH DSQN+KKLSELLRY<sup>↓</sup>YTSASGDEMVS<sup>↓</sup>LKD<sup>↓</sup>YCTR<sup>↓</sup>MK  
Sbjct 1326 FYEPFSKNIKLGIHGDSQNQKKLSELLRY<sup>↓</sup>YTYASGDEMVS<sup>↓</sup>LQD<sup>↓</sup>YCTR<sup>↓</sup>MK

---

Query 486 ENQKH<sup>↓</sup>IYYITGETKDQVANS<sup>↓</sup>AFVERLRKHGLEVIYMI<sup>↓</sup>EPIDEYCVQQLK  
ENQKH<sup>↓</sup>IYYITGETKDQVANS +V+RL KHGLEVIY IEPIDEYCVQQLK  
Sbjct 1473 ENQKH<sup>↓</sup>IYYITGETKDQVANSTIVQRLWKHGLEVIY<sup>↓</sup>TI<sup>↓</sup>EPIDEYCVQQLK

---

Query 535 EFEGKTLVSVTKEGLELPEDEEEKKKQEE- - - - -K<sup>↓</sup>TKFENLCKIMK  
EFEGKTLVSVTKE LELPEDEEEKKKQEE KK K<sup>↓</sup>FENLCKI+K  
Sbjct 1620 EFEGKTLVSVTKEDLELPEDEEEKKKQEEGKQTKQKK!<sup>↓</sup>KFENLCKI<sup>↓</sup>VK

Query 577 DILEKKVEKVVVSNRLVTSPCCIVTSTYGWTANMERIMKAQALRDNSTM  
 DILEK VEKVVVSN+LVTSPCCIV+STYGWTANMERIMKAQALRDNST  
 Sbjct 1769 DILEK!VEKVVVSNQLVTSPCCIVSSTYGWTANMERIMKAQALRDNSTT

---

Query 626 GYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDLVILLYETALLSS  
 GYMAAKKHLEINPDHS+I+TLRQKAE DKNDKSVKDLVILLYETALLSS  
 Sbjct 1914 GYMAAKKHLEINPDHSFIDTLRQKAEADKNDKSVKDLVILLYETALLSS

Query 675 GFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPLGDD  
 F LE+PQTHANRIYRM KLGLG DEDDPTADDTSAAVTEEMPPLGDD  
 Sbjct 2061 DFGLEGPQTHANRIYRMNKLGLGTDEDDPTADDTSAAVTEEMPPLGDD

Query 724 DTSRMEE  
 DTSRME+  
 Sbjct 2208 DTSRMEK

### HSP90AA5P

Query 192 EDQTEYLEERRIKEIVKKHSQFIGYPITLFFVEKERDKEVSDDEAEKED  
 E + EYLEERRIKEIVKKHSQFIGYPITLFFVEK+R+K+VSD EAE+KED  
 Sbjct 22 ERRPEYLEERRIKEIVKKHSQFIGYPITLFFVEKKRNKQVSDAEAEKED

Query 241 KEEEEKEKEEKESEDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEYIDQEE  
 K K++ KES DKPEIEDVGSDEEEEEKD DKKKKK KEYIDQE  
 Sbjct 169 ---KRKKKESNDKPEIEDVGSDEEEEEKDADKKKKKS-KEYIDQE-

Query 290 LNKTTP IWTRNPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLE  
 LNKTTP IWTRNPD ITNEEYGEF++SLTN+WEDHLAVKHFSVEGQLE  
 Sbjct 298 LNKTTP IWTRNPDAITNEEYGEFHQSLTNNWEDHLAVKHFSVEGQLE

Query 337 RALLFVPRRAPFDL FENRKKKNN  
 RA LFV R APFDL FENRKKKNN  
 Sbjct 439 <2-----[441 : 465]-2> RA-LFVSRCAPFDL FENRKKKNN

Query 361 IKLYVRRVFIMDNCEELIPEYLNFI RGVVDS EDLPLNISREMLQQSKIL  
 I+LYV RVFIM NCEELIPEYLNFI RGV+DSEDLPLNI EMLQQSKIL  
 Sbjct 533 IQLYVHRVFIMGNCEELIPEYLNFI RGVMDSEDLPLNIFSEMLQQSKIL

Query 410 KVIRKNLVKKCLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKK  
 KVIRKNLVK LELF ELAEDK+NYKKF+E++S+NI L + DSQN+KK  
 Sbjct 680 KVIRKNLVKTYLELFAELAEDKKNYKKFHEKLS ENINL!EYIDSQNRKK

Query 459 LSELLRYYTSASGDEMVS LKDYCTR MKENQKHIIYYITGETKDQVANS AF  
 LSELL+ TSAS DEMV L + RMK NQKHIIYYITGETKDQVANS AF  
 Sbjct 825 LSELLKXCTSASRDEM VYLIVF-SRMKANQKHIIYYITGETKDQVANS AF

Query 508 VERLRKHGLEVIYMIIEPIDEYCVQQLKEFEGKTLVSVTKEGLELPEDEE  
VE L+KHGLEVIYMIIE ID+YCVQQLKE+E KT+VSV KEGLELPEDEE  
Sbjct 969 VECLQKHGLEVIYMIELIDKYCVQQLKELESKTVVSVAKEGLELPEDEE

Query 557 EKKKQEEKKTGFENLCKIMKDILEKKVEKVVVSNRLVTSPCCIVTSTYG  
EKKKQEEKKTGFENLCKIMKD+LEKKV+KVVVSN VTSPCCIVTSTYG  
Sbjct 1116 EKKKQEEKKTGFENLCKIMKDMLEKKVKKVVVSN!-VTSPCCIVTSTYG

Query 606 WTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSI IETLRQKAEADK  
WTANMERIMKAQA DNSTMGY+AAKKHLEINPDHS IETLRQKAEADK  
Sbjct 1259 WTANMERIMKAQARXDNSTMGYVAAKKHLEINPDHS! IETLRQKAEADK

Query 655 NDKSVKDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGGLGIDEDDP  
N+KSVKDLVILLY TAL S GFS+EDPQ H N+IYRMIKLGGLG+DE DP  
Sbjct 1405 NDKSVKDLVILLYXTALPSPGFSMEDPQRHTNKIYRMIKLGGLGVDEYDP

Query 704 TADDTSAAVTEEMPPLGDDDTSRMEEV  
TA+D +AA+T+EMPPL G DDTSRMEEV  
Sbjct 1552 TANDINAAITKEMPPLRGGDDDTSRMEEV

**CSMD3**

Query 1 MGVDIRHNKDRKVRKPEKKSQDIYLRLLV-KLYRFLARRTNSTFNQVVL  
MGVD RH KDRKVRKPEKKSQDIYLR L V + ++ TNSTFN+VVL  
Sbjct 1 MGVDTRH-KDRKVRKPEKKSQDIYLR L SVGQAVQVSSKGTNSTFNQVVL

Query 49 KRLFMSRTNRPPPLSLSRMIRKMKLPGRENKTAVVVGTITDDVRVQEVPK  
KRLFMSR NRPPPL LSRMI KMKLPGRENKTAVVVG ITDDVRVQEVPK  
Sbjct 145 KRLFMSRNNRPPPLPLSRMIWKMKLPRENKTAVVVGAITDDVRVQEVPK

Query 98 LKVCALRVTSRARSILR-----AGGKILTFDQLALDSPKG  
LKVCALR+TSRARS I A GK LTF QLALDSPKG  
Sbjct 292 LKVCALRMTSRARSHIFXGVGWGGGAERSGARGKSLTFRQLALDSPKG

Query 134 CGTVLLSGPRKGREVYRHF GKAPGTPHSHTKPYVRSKGRKFERARGRA  
CG V LSGP KGR VY P + KPY+RS GRKFERAR RRA  
Sbjct 439 CGVVRLSGPPKGRGVY-----APET!CKPYIRSNRGRKFERARDRA

Query 183 SRGYKN  
SRGYKN  
Sbjct 560 SRGYKN

**WBP2NL**

Query 1 MTDAAVSFAKDFLAGGVAAAISK TAVAPIERVKLLLVQHASKQITADK  
+T+AAVSFAKD LAGGVAAAISK AVAPIERVKLLLVQHASKQITADK  
Sbjct 1 ITEAAVSFAKDSL AGGVAAAISKMAVAPIERVKLLLVQHASKQITADK

Query 50 QYKGIIDCVVRIPKEQGVLSFWRGNLANVIRYFPTQALNFAFKDKYKQI  
QYKGI DC+V IPKE GVLS W GNLA+VIRYFP QALNF FKDKYKQI  
Sbjct 148 QYKGIIDCVVRIPKEQGVLSFWRGNLANVIRYFPTQALNFAFKDKYKQI

Query 99 FLGGVVDKRTQFWLYFAGNLAGS GAAGATSLCFVYPLDFARTRLAADV GK  
FL GVDKR+QFW YFAGNLAGS GA GATSLCFVYPLDFA TR+AADV GK  
Sbjct 293 FLDGVDKRSQFWRYFAGNLAGS GATGATSLCFVYPLDFAHTRVAADV GK

Query 48 AGAEREFRGLGDCLVKIYKSDGIKGLYQGFNVSVQGI I IYRAA---YFG  
AGAERE+RG GDCLVKIYKSDGIK LYQGFNVSVQGI I YRAA YFG  
Sbjct 440 AGAERELRGFGDCLVKIYKSDGIKSLYQGFNVSVQGI I -YRAACC!YFG

Query 194 IYDTAKGMLPDPKNTHIVISWMI AQTVTAVAGLTSY PFDTVRRRMMMQS  
YDTAKGMLPDPKNTHIVISWMI AQTVTAVAGLTSY F+T+ MMMQS  
Sbjct 582 NYDTAKGMLPDPKNTHIVISWMI AQTVTAVAGLTSYSFNTICCYMMMQS

Query 243 GRKGT DIMYTGTDCWRKIARDEGGKAFFKGAWSNVL RGMGGAFVVLVLY  
G KGT DIMYTGT DCWRKI DEG KA FKGAW S+VLRGMGGAFVVLVLY  
Sbjct 728 GCKGT DIMYTGTDCWRKI THDEGSKAVFKGAW SSVLRGMGGAFVVLVLY

Query 292 DEIKKY  
DE KKY  
Sbjct 875 DETKKY

**AC019016.1**

Query 9 AEFIVGGKYKLV RKIGSGSFGDIYLAINITNGEEVAVKLESQKARHPQL  
AEFIVG KYKL+ KI SFGDI LAI+ITNGEEV VK ESQ ARHPQL  
Sbjct 393 AEFIVGRKYKLLWKIWCSSFGDICLAISITNGEEVVVKPESQ!ARHPQL

Query 58 LYESKLYKILQGGVGIPHIRWYGQEKDYNV LVM DLLGPSLEDLNFNFC SR  
LY+S+LYK+LQGGVGIPHIRW GQE+DYNV LVM D G L NFNFC SR  
Sbjct 538 LYKSRLYKMLQGGVGIPHIRWCGQEEDYNV LVM D FSG! -LASKTNFNFC SR

Query 107 RFTMKT V LMLADQ MISRIEYVHTKNFIHRDIKPDN FLMGIGRHCNKLF L  
RFTMKT V L+LADQ MISRIE+VHTKNFIHR IKP NFLMGIGRHC+KLF  
Sbjct 681 RFTMKT V LILADQ MISRIEYVHTKNFIHRAIKPGN FLMGIGRHC SKLFR

Query 156 IDFGLAKKYRDN RTRQHIPYREDKNLTGTARYASIN AHLGIEQSRRDD M  
IDFGLAK Y DNRTRQHIPYREDKNLTGT A YAS AHL IEQSRRDD M  
Sbjct 828 IDFGLAKTYGDN RTRQHIPYREDKNLTGTALYAS -HAHLSIEQSRRDD M

Query 205 ESLGYVLMYFNRTS---LPWQGLKAATKKQKYEKISEKKMSTPVEVLCK  
ESL YVLMYFNRTS LPWQGLKAA +KQKYEKISEKKMSTPV CK  
Sbjct 972 ESLEYVLMYFNRTSQP!LPWQGLKAAAEKQKYEKISEKKMSTPV--FCK

Query 251 ↓ GFPAEFAMYLNYCRGLRFEEAPDYMYLRQLFRILFRITLNHQYDYTFDWT  
GF AEFA+YLN CRGL +E+APDYMYL QLF ILFRITL HQYDYTFDWT  
Sbjct 1112 GFAAEFAIYLNPCRGLHLEKAPDYMYLTQLFCILFRITLIHQYDYTFDWT ↑

Query 300 MLKQKAAQQAASSSGGQQAQTPTGKQTDKTKSNMKGF  
LKQKAAQQAASSSG GQQAQTPTGKQ DKTK NMKGF  
Sbjct 1259 KLKQKAAQQAASSSG!GQQAQTPTGKQPDKTKGNMKGF