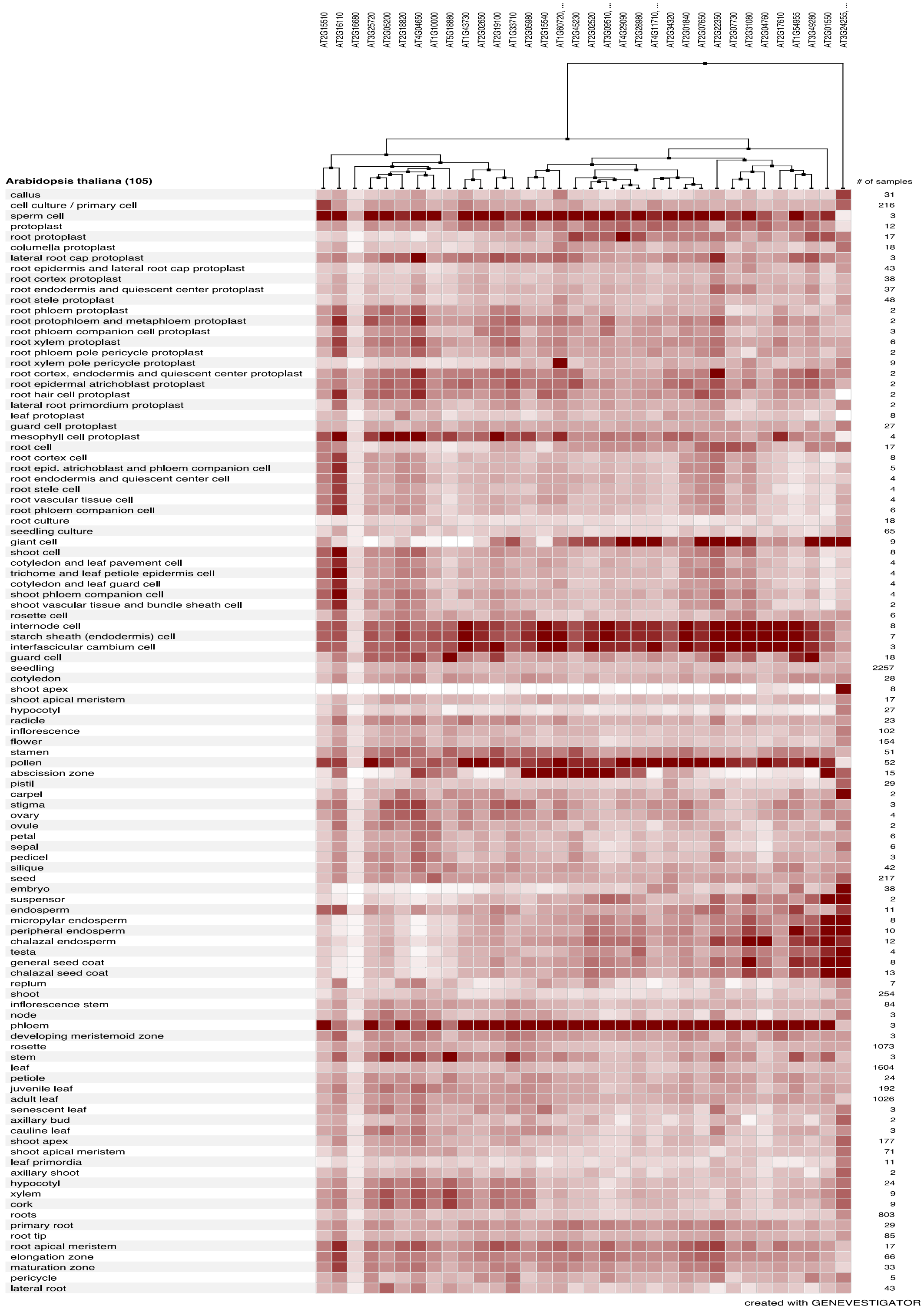
Supplementary materials to Galvan-Gordillo et al., Biologia 71, 11, 2016.

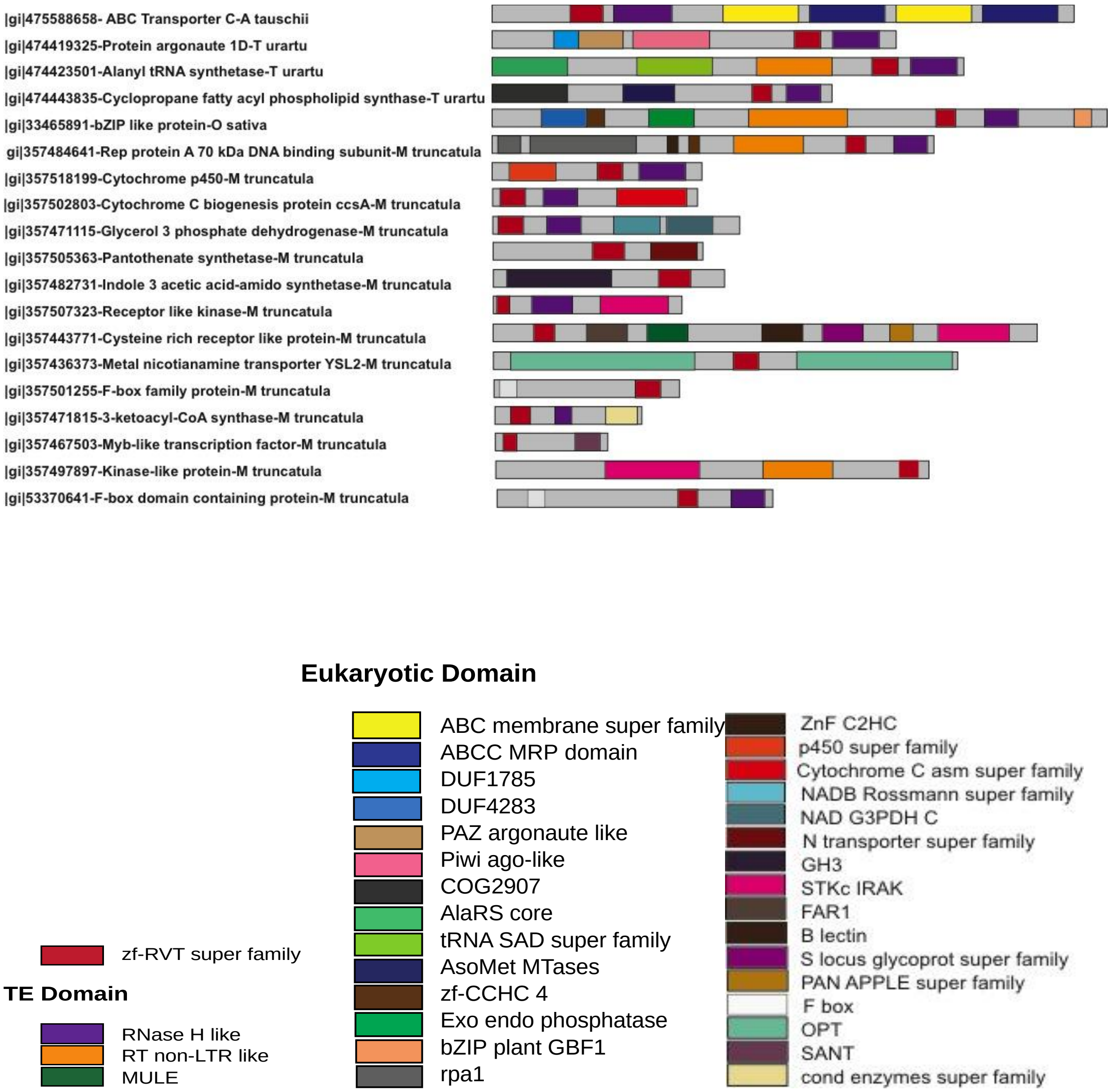
Supplementary Table 1. GO annotations of putative endogenized genes from *Arabidopsis*. Expression pattern, subcellular localization and biological roles of the putative endogenized Arabidopsis RVT-Znf genes are shown, based on TAIR (https://www.arabidopsis.org/tools/bulk/go/) and Genevestigator databases (https://genevestigator.com/gv/).

Putative Endogenous Gene	**Subcellular Location**	**Expressed in**	**Expressed During**	**Involved in**
At1g43730				
At4g04650	Stromule			Response to cold
At2g16110		hypocotyl, petiole, shoot apex		
At2g02520	Mitochondrion	Collective leaf structure, stamen, vascular leaf	LP0.2 two leaves visible stage, petal differentiation and expansion stage	
At1g33710				
At1g60720				
At5g18880	Nucleus			
At3g24255	Nucleus	Guard cell		
At4g11710				
At4g29090	Mitochondrion			
At1g10000		cauline leaf, collective leaf structure, cotyledon, flower, guard cell, hypocotyl,inflorescence meristem, leaf apex, petal, petiole, plant embryo, plant sperm cell,pollen, seed, sepal, shoot apex, shoot system, stem, vascular leaf	LP.02 two leaves visible stage, LP.04 four leaves visible stage, LP.06 six leaves visible stage, LP.10 ten leaves visible stage, LP.12 twelve leaves visible stage, M germinated pollen stage, flowering stage, mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, vascular leaf senescent stage	
At3g25270	Nucleus			
At3g09510	Mitochondrion			
At2g02650	Mitochondrion	Sperm cell		
At5g36905				
At2g22350				

Supplementary Fig. 1. 105 anatomical parts from 34 Arabidopsis RVT Znf genes were analyzed using the hierarchical clustering tool, using the Pearson correlation, from the GENEVESTIGATOR microarray dataset.

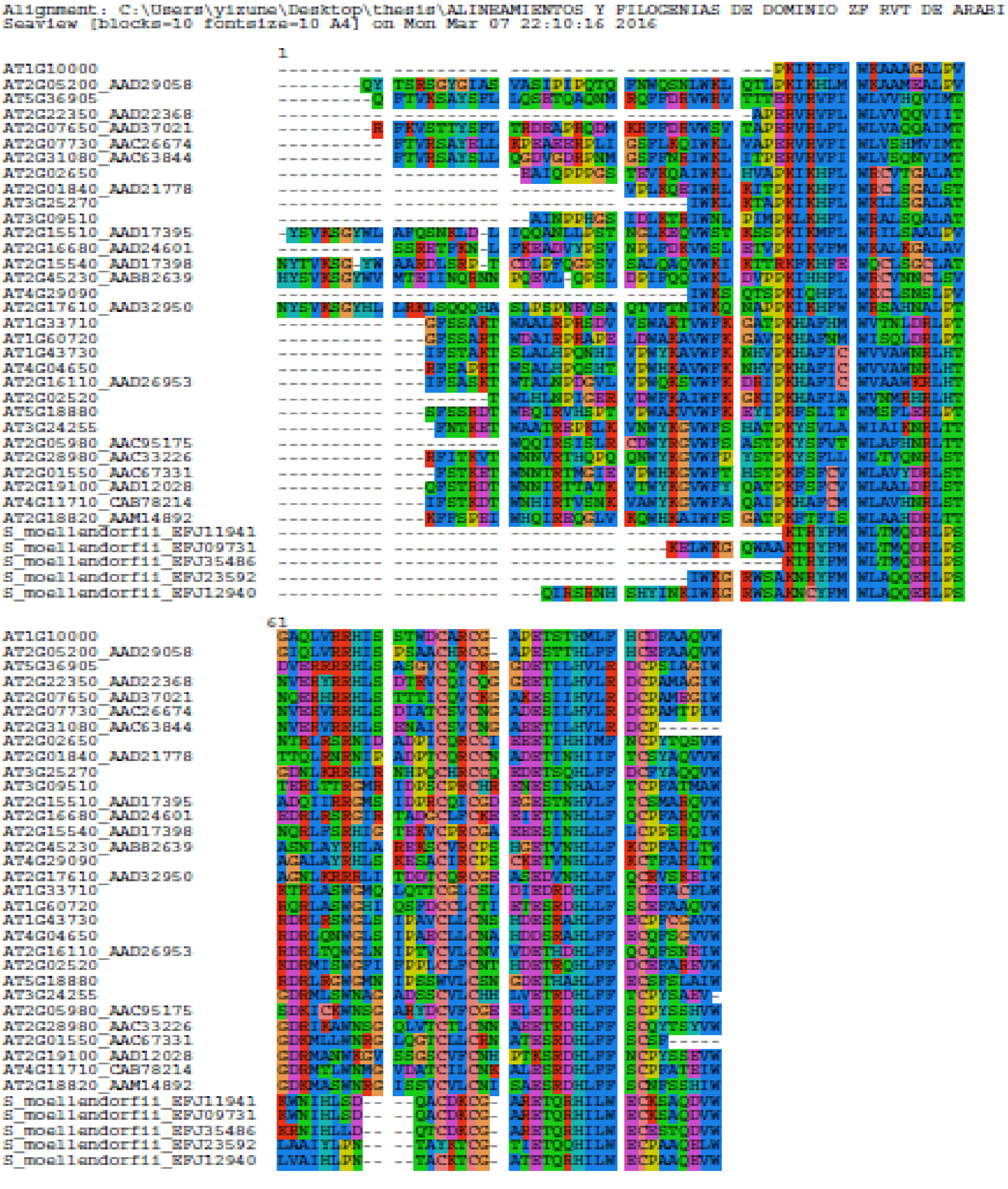


Supplementary Fig. 2. Schematic representation of additional domains in zf-RVTs from plant species. Genes containing RVT-Znf domain were obtained with the NCBI search and analysis engine using the BLAST tool against the Viridiplantae database. At4g29090 was used as query, and domain structures were obtained using the conserved domain database from NCBI.



Supplementary Fig. 3. Alignment of the RVT-Znf domain containing genes from *Arabidopsis*.

The total of 36 amino acid sequences analyzed were obtained as described in materials and methods, aligned using ClustalX2, manually edited using Seaview.

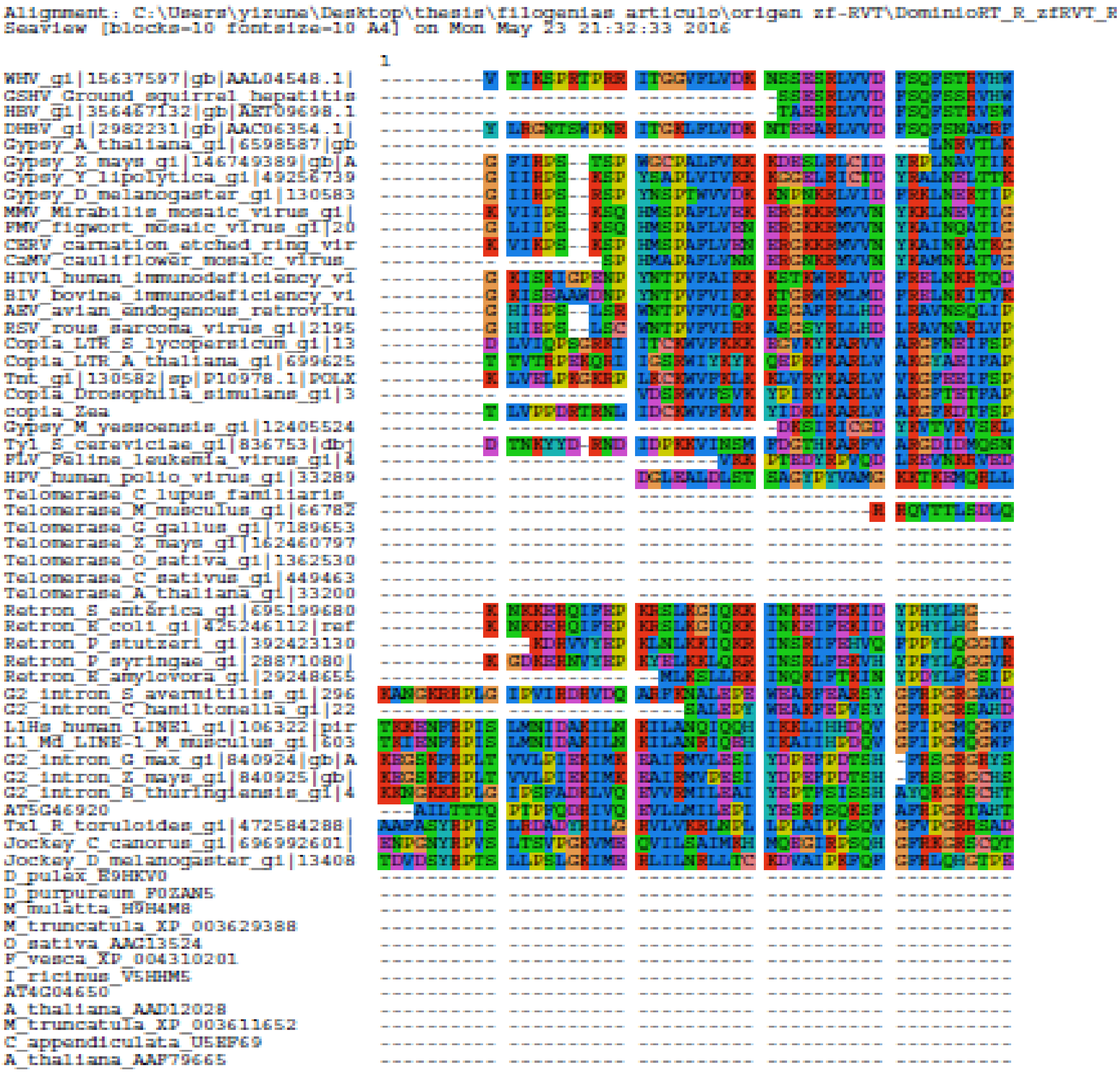


Alignment: C:\Users\yizuna\Desktop\thesis\ALINEAMIENTOS Y FILOGENIAS DE DOMINIO ZF RVT DE ARABI
Seaview [blocks=10 fontsize=10 A4] on Mon Mar 07 22:10:16 2016

Supplementary Fig 4-8. Alignment of the origin of the RVT-Znf domain.

The total of 60 amino acid sequences analyzed were retrieved from the NCBI and Pfam databases, aligned using ClustalX2 and manually edited using Seaview.

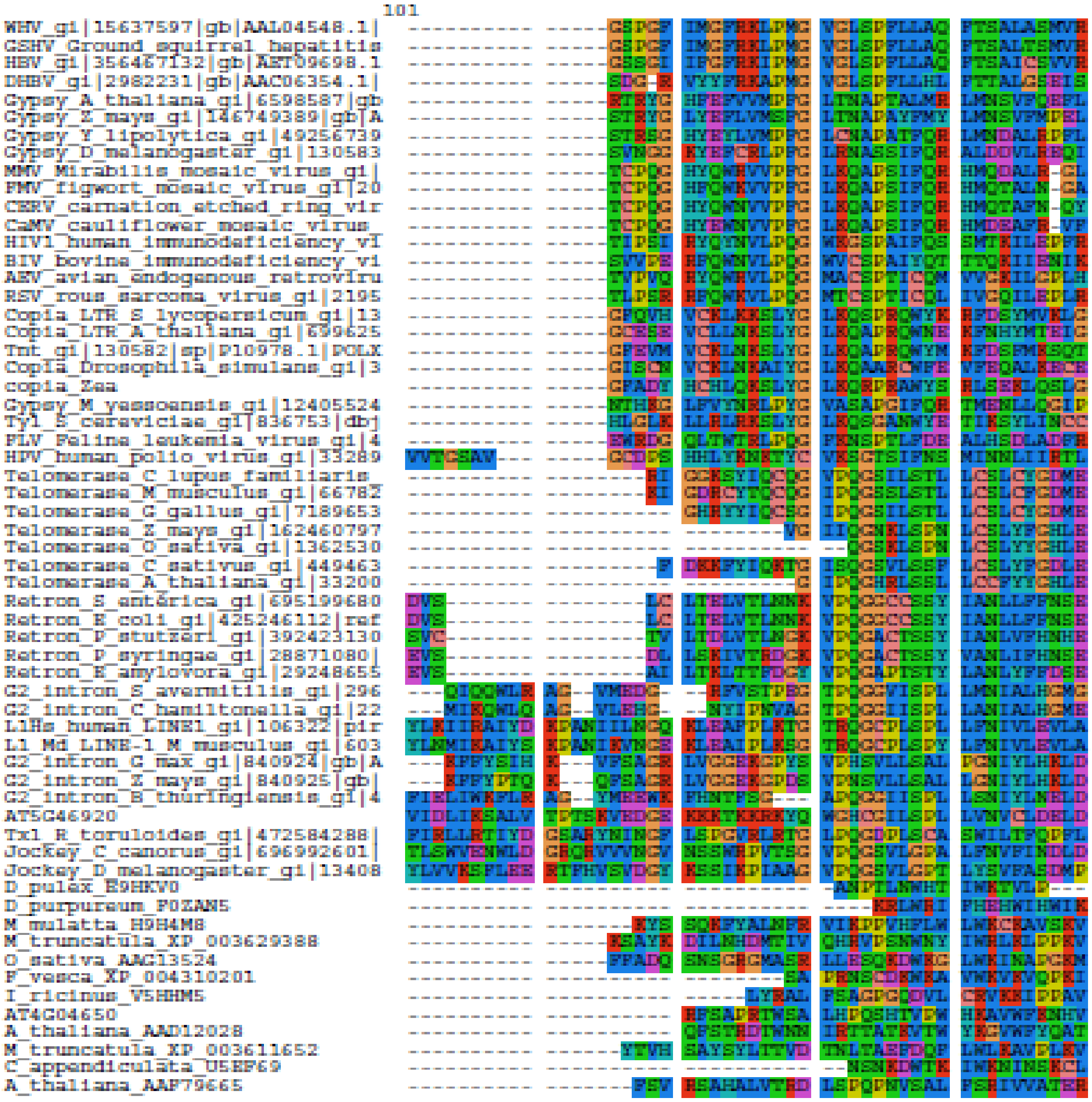
Suppl. Fig. 4.



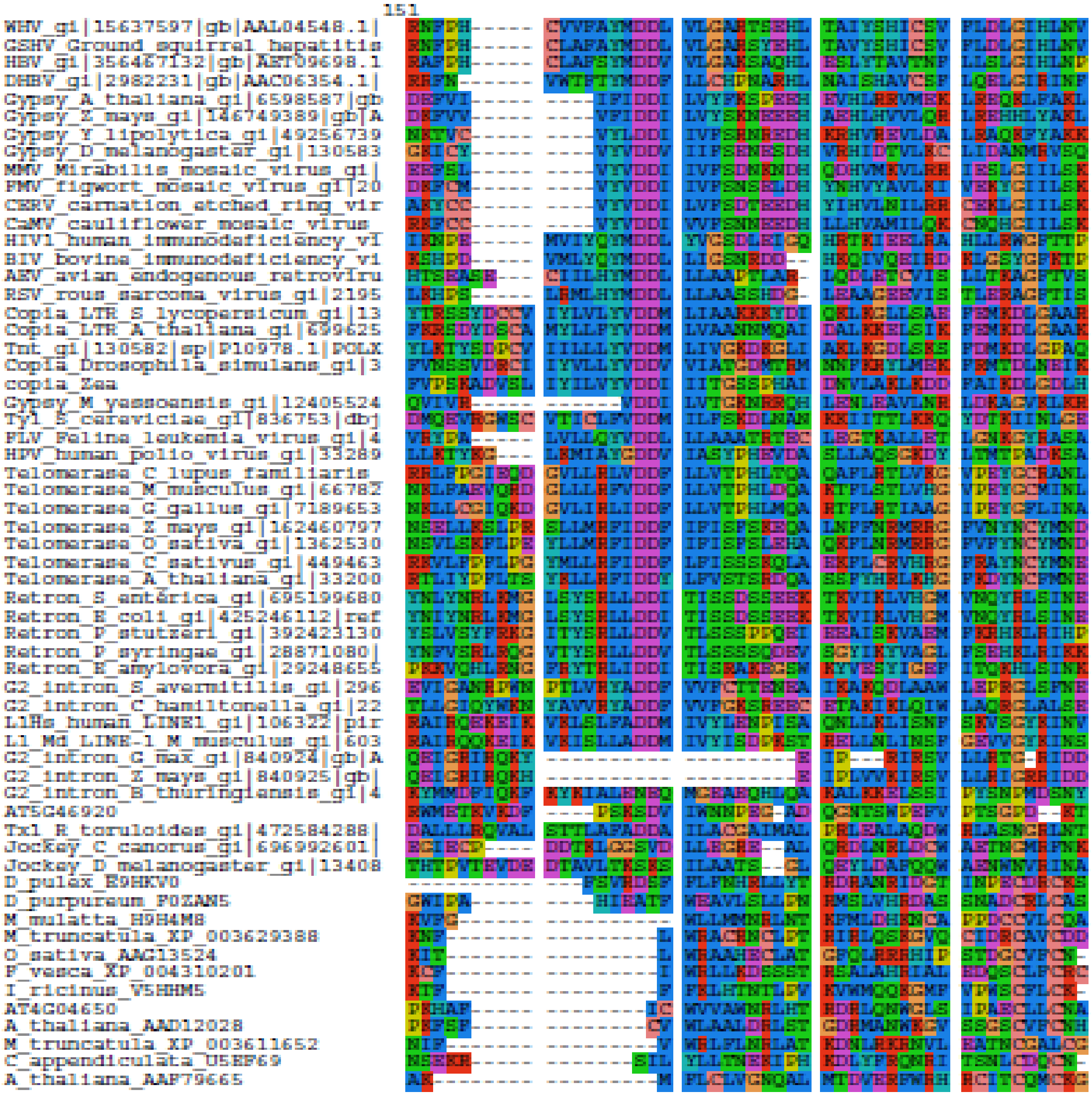
```
Alignment: C:\Users\yizune\Desktop\thesis\filogenias articulo\origen zf-RVT\DominioRT_R_zfRVT_R
Seaview [blocks=10 fontsize=10 A4] on Mon May 23 21:32:33 2016

                                   1
WHV_gi|15637597|gb|AAL04548.1|     ---------V TIKSPRTPRR ITGGVFLVDK NSSRSRLVVD FSQFSTRVHW
GSHV_Ground_squirrel_hepatitis     ---------- ---------- ---------- -SSRSRLVVD FSQFSSRVHW
HBV_gi|356467132|gb|AET09698.1     ---------- ---------- ---------- -TARSRLVVD FSQFSTRVSW
DHBV_gi|2982231|gb|AAC06354.1|     ---------Y LRGNTSWPNR ITGKLFLVDK NTREARLVVD FSQFSNAMRF
Gypsy_A_thaliana_gi|6598587|gb     ---------- ---------- ---------- ---------- ---LNRVTLK
Gypsy_Z_mays_gi|146749389|gb|A     ---------G FIRPS--TSP WGCPALFVKK KDRSLRLCID YRPLNAVTIK
Gypsy_Y_lipolytica_gi|49256739     ---------G IIRPS--RSP YSAPLVIVKK KDGELRICTD YRALNELTTK
Gypsy_D_melanogaster_gi|130583     ---------G IIRPS--RSP YNSPTWVVDK KNPNKRLVID FRKLNEKTIP
MMV_Mirabilis_mosaic_virus_gi|     ---------K VIIPS--KSQ HMSPAFLVEK ERGKKRMVVN YKKLNEVTIG
FMV_figwort_mosaic_virus_gi|20     ---------G LIIPS--KSQ HMSPAFLVEN ERGKKRMVVN YKAINQATIG
CERV_carnation_etched_ring_vir     ---------K VIKPS--KSP HMSPAFLVEN ERGKKRMVVN YKAINKATKG
CaMV_cauliflower_mosaic_virus_     ---------- --------SP HMAPAFLVNN ERGNKRMVVN YKAMNKATVG
HIV1_human_immunodeficiency_vi     ---------G KISKIGPENP YNTPVFAIKK KSTKWRKLVD FRELNKRTQD
BIV_bovine_immunodeficiency_vi     ---------G KISRAAWDNP YNTPVFVIKK KTGRWRMLMD FRELNKITVK
AEV_avian_endogenous_retroviru     ---------G HIRPS---LSR WNTPVFVIQK KSGAFRLLHD LRAVNSQLIP
RSV_rous_sarcoma_virus_gi|2195     ---------G HIRPS---LSC WNTPVFVIRK ASGSYRLLHD LRAVNAKLVP
Copia_LTR_S_lycopersicum_gi|13     ---------D LVIQPSGRKI ITCKWVFKKK RGVKYKARVV ARGFNEIFSP
Copia_LTR_A_thaliana_gi|699625     ---------T TVTRPEKQRI IGSRWIYKYK QEPRFKARLV AKGYASIPAP
Tnt_gi|130582|sp|P10978.1|POLX     ---------K LVELPKGKRP LGCKWVFKLK KLVRYKARLV VKGFEEIFSP
Copia_Drosophila_simulans_gi|3     ---------- ---------- VDSKWVFSVK YPIRYKARLV ARGFTETPAP
copia_Zea                          ---------T LVPPDRTRNL IDCKWVFKVK YIDRLKARLV AKGFKDTFSP
Gypsy_M_yessoensis_gi|12405524     ---------- ---------- ---------- -DKSIRICGD YKVTVKVSKL
Ty1_S_cereviciae_gi|836753|dbj     ---------D TNKYYD-RND IDPKKVINSM FDGTHKARFV ARGDIDMQSN
FLV_Feline_leukemia_virus_gi|4     ---------- ---------- -------VKK PTRDYRPVQD LREVNKRVED
HPV_human_polio_virus_gi|33289     ---------- ---------- DGLEALDLST SAGYPYVAMG KKTKEMQRLL
Telomerase_C_lupus_familiaris      ---------- ---------- ---------- ---------- ----------
Telomerase_M_musculus_gi|66782     ---------- ---------- ---------- ---------R RQVTTLSDLQ
Telomerase_G_gallus_gi|7189653     ---------- ---------- ---------- ---------- ----------
Telomerase_Z_mays_gi|162460797     ---------- ---------- ---------- ---------- ----------
Telomerase_O_sativa_gi|1362530     ---------- ---------- ---------- ---------- ----------
Telomerase_C_sativus_gi|449463     ---------- ---------- ---------- ---------- ----------
Telomerase_A_thaliana_gi|33200     ---------- ---------- ---------- ---------- ----------
Retron_S_enterica_gi|695199680     ---------K NKKERQIFEP KRSLKGIQKK INKEIFEKID YPHYLHG---
Retron_E_coli_gi|425246112|ref     ---------K NKKERQIFEP KRSLKGIQKK INKEIFEKID YPHYLHG---
Retron_P_stutzeri_gi|392423130     ---------- --KDRVVYEP KLNLKKIQKR INSRIFEHVQ FPPYLQGGIK
Retron_P_syringae_gi|28871080|     ---------K GDKERNVYEP KYRLKKLQKR INSRLFERVH YPPYLQGGVR
Retron_E_amylovora_gi|29248655     ---------- ---------- --MLKSLLRK INQKIFTKIN YPDYLFGSIP
G2_intron_S_avermitilis_gi|296     KANGKRRPLG IPVIRDRVDQ ARFKNALEPE WEARFEARSY GFRPGRGAWD
G2_intron_C_hamiltonella_gi|22     ---------- ---------- ----SALEPY WEARFEPVSY GFRPGRSAHD
L1Hs_human_LINE1_gi|106322|pir     TKKENFRPIS LMNIDAKILN KILANQIQQH IKKLIHHDQV GFIPGMQGWF
L1_Md_LINE-1_M_musculus_gi|603     TKIENFRPIS LMNIDAKILN KILANRIQEH IKAIIHPDQV GFIPGMQGWF
G2_intron_G_max_gi|840924|gb|A     KEGSKFRPLT VVLPIEKIMK EAIRMVLESI YDPEFPDTSH -FRSGRGRYS
G2_intron_Z_mays_gi|840925|gb|     KEGSKFRPLT VVLPIEKIMK EAIRMVPESI YDPEFPDTSH -FRSGRGCHS
G2_intron_B_thuringiensis_gi|4     KKNGKKRPLG IPSPADKLVQ EVVRMILEAI YEPTPSISSH AYQKGRSCHT
AT5G46920                          ---AILTTIQ PTPPQDRIVQ EVLLMILEPI YESRFSQKSF AFRPGRTAHT
Tx1_R_toruloides_gi|472584288|     AAPASYRPIS LRDADYRILG RVLVKRLNPL LPLAIPLSQV GFVPGRRSAD
Jockey_C_canorus_gi|696992601|     ENPGNYRPVS LTSVPGKVME QVILSAIMKH MQRGIRPSQH GFRKGRSCQT
Jockey_D_melanogaster_gi|13408     TDVDSYRPTS LLPSLGKIME RLILNRLLTC KDVAIPKFQF GFRLQHGTPE
D_pulex_E9HKV0                     ---------- ---------- ---------- ---------- ----------
D_purpureum_F0ZAN5                 ---------- ---------- ---------- ---------- ----------
M_mulatta_H9H4M8                   ---------- ---------- ---------- ---------- ----------
M_truncatula_XP_003629388          ---------- ---------- ---------- ---------- ----------
O_sativa_AAG13524                  ---------- ---------- ---------- ---------- ----------
F_vesca_XP_004310201               ---------- ---------- ---------- ---------- ----------
I_ricinus_V5HHM5                   ---------- ---------- ---------- ---------- ----------
AT4G04650                          ---------- ---------- ---------- ---------- ----------
A_thaliana_AAD12028                ---------- ---------- ---------- ---------- ----------
M_truncatula_XP_003611652          ---------- ---------- ---------- ---------- ----------
C_appendiculata_U5EF69             ---------- ---------- ---------- ---------- ----------
A_thaliana_AAP79665                ---------- ---------- ---------- ---------- ----------
```

Suppl. Fig. 5.


101
WHV_gi|15637597|gb|AAL04548.1|
GSHV_Ground_squirrel_hepatitis
HBV_gi|356467132|gb|AET09698.1
DHBV_gi|2982231|gb|AAC06354.1|
Gypsy_A_thaliana_gi|6598587|gb
Gypsy_Z_mays_gi|146749389|gb|A
Gypsy_Y_lipolytica_gi|49256739
Gypsy_D_melanogaster_gi|130583
MMV_Mirabilis_mosaic_virus_gi|
FMV_figwort_mosaic_virus_gi|20
CERV_carnation_etched_ring_vir
CaMV_cauliflower_mosaic_virus_
HIV1_human_immunodeficiency_vi
BIV_bovine_immunodeficiency_vi
AEV_avian_endogenous_retroviru
RSV_rous_sarcoma_virus_gi|2195
Copia_LTR_S_lycopersicum_gi|13
Copia_LTR_A_thaliana_gi|699625
Tnt_gi|130582|sp|P10978.1|POLX
Copia_Drosophila_simulans_gi|3
copia_Zea
Gypsy_M_yessoensis_gi|12405524
Ty1_S_cerevisiae_gi|836753|dbj
FLV_Feline_leukemia_virus_gi|4
HPV_human_polio_virus_gi|33289
Telomerase_C_lupus_familiaris
Telomerase_M_musculus_gi|66782
Telomerase_G_gallus_gi|7189653
Telomerase_Z_mays_gi|162460797
Telomerase_O_sativa_gi|1362530
Telomerase_C_sativus_gi|449463
Telomerase_A_thaliana_gi|33200
Retron_S_enterica_gi|695199680
Retron_E_coli_gi|425246112|ref
Retron_P_stutzeri_gi|392423130
Retron_P_syringae_gi|28871080|
Retron_E_amylovora_gi|29248655
G2_intron_S_avermitilis_gi|296
G2_intron_C_hamiltonella_gi|22
L1Hs_human_LINE1_gi|106322|pir
L1_Md_LINE-1_M_musculus_gi|603
G2_intron_G_max_gi|840924|gb|A
G2_intron_Z_mays_gi|840925|gb|
G2_intron_B_thuringiensis_gi|4
AT5G46920
Tx1_R_toruloides_gi|472584288|
Jockey_C_canorus_gi|696992601|
Jockey_D_melanogaster_gi|13408
D_pulex_E9HKV0
D_purpureum_F0ZAN5
M_mulatta_H9H4M8
M_truncatula_XP_003629388
O_sativa_AAG13524
F_vesca_XP_004310201
I_ricinus_V5HHM5
AT4G04650
A_thaliana_AAD12028
M_truncatula_XP_003611652
C_appendiculata_U5EF69
A_thaliana_AAF79665

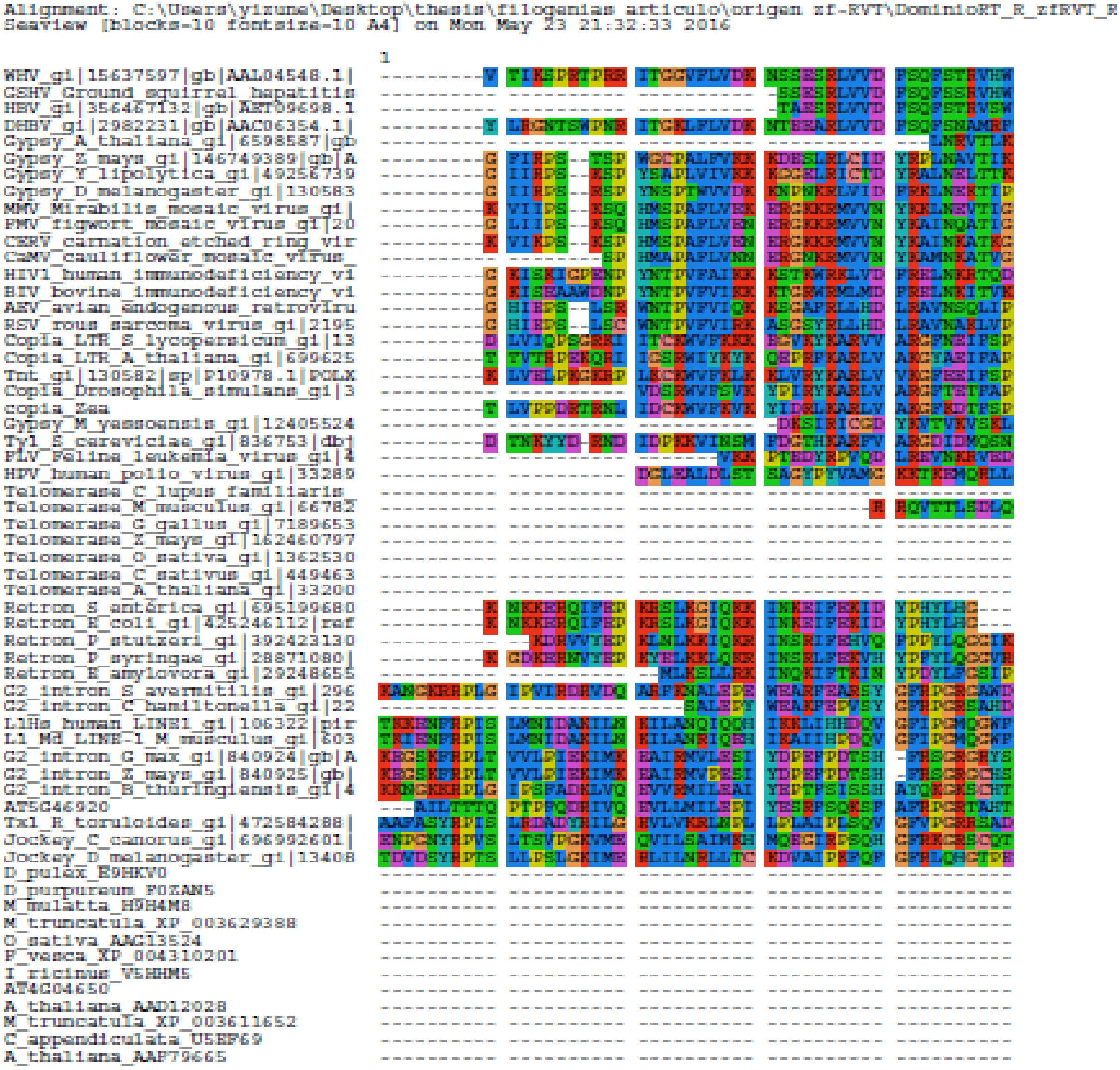
Suppl. Fig. 6.


151
WHV_gi|15637597|gb|AAL04548.1|
GSHV_Ground_squirrel_hepatitis
HBV_gi|356467132|gb|AET09698.1
DHBV_gi|2982231|gb|AAC06354.1|
Gypsy_A_thaliana_gi|6598587|gb
Gypsy_Z_mays_gi|146749389|gb|A
Gypsy_Y_lipolytica_gi|49256739
Gypsy_D_melanogaster_gi|130583
MMV_Mirabilis_mosaic_virus_gi|
FMV_figwort_mosaic_virus_gi|20
CERV_carnation_etched_ring_vir
CaMV_cauliflower_mosaic_virus
HIV1_human_immunodeficiency_vi
BIV_bovine_immunodeficiency_vi
AEV_avian_endogenous_retroviru
RSV_rous_sarcoma_virus_gi|2195
Copia_LTR_S_lycopersicum_gi|13
Copia_LTR_A_thaliana_gi|699625
Tnt_gi|130582|sp|P10978.1|POLX
Copia_Drosophila_simulans_gi|3
copia_Zea
Gypsy_M_yessoensis_gi|12405524
Ty1_S_cerevisiae_gi|836753|dbj
FLV_Feline_leukemia_virus_gi|4
HPV_human_polio_virus_gi|33289
Telomerase_C_lupus_familiaris
Telomerase_M_musculus_gi|66782
Telomerase_G_gallus_gi|7189653
Telomerase_Z_mays_gi|162460797
Telomerase_O_sativa_gi|1362530
Telomerase_C_sativus_gi|449463
Telomerase_A_thaliana_gi|33200
Retron_S_enterica_gi|695199680
Retron_E_coli_gi|425246112|ref
Retron_P_stutzeri_gi|392423130
Retron_P_syringae_gi|28871080|
Retron_E_amylovora_gi|29248655
G2_intron_S_avermitilis_gi|296
G2_intron_C_hamiltonella_gi|22
L1Hs_human_LINE1_gi|106322|pir
L1_Md_LINE-1_M_musculus_gi|603
G2_intron_G_max_gi|840924|gb|A
G2_intron_Z_mays_gi|840925|gb|
G2_intron_B_thuringiensis_gi|4
AT5G46920
Tx1_R_toruloides_gi|472584288|
Jockey_C_canorus_gi|696992601|
Jockey_D_melanogaster_gi|13408
D_pulex_E9HKV0
D_purpureum_F0ZAN5
M_mulatta_H9H4M8
M_truncatula_XP_003629388
O_sativa_AAG13524
F_vesca_XP_004310201
I_ricinus_V5HHM5
AT4G04650
A_thaliana_AAD12028
M_truncatula_XP_003611652
C_appendiculata_U5EF69
A_thaliana_AAF79665

Suppl. Fig. 7.


201
WHV_gi|15637597|gb|AAL04548.1|
GSHV_Ground_squirrel_hepatitis
HBV_gi|356467132|gb|AET09698.1
DHBV_gi|2982231|gb|AAC06354.1|
Gypsy_A_thaliana_gi|6598587|gb
Gypsy_Z_mays_gi|146749389|gb|A
Gypsy_Y_lipolytica_gi|49256739
Gypsy_D_melanogaster_gi|130583
MMV_Mirabilis_mosaic_virus_gi|
FMV_figwort_mosaic_virus_gi|20
CERV_carnation_etched_ring_vir
CaMV_cauliflower_mosaic_virus_
HIV1_human_immunodeficiency_vi
BIV_bovine_immunodeficiency_vi
AEV_avian_endogenous_retroviru
RSV_rous_sarcoma_virus_gi|2195
Copia_LTR_S_lycopersicum_gi|13
Copia_LTR_A_thaliana_gi|699625
Tnt_gi|130582|sp|P10978.1|POLX
Copia_Drosophila_simulans_gi|3
copia_Zea
Gypsy_M_yessoensis_gi|12405524
Ty1_S_cerevisiae_gi|836753|dbj
FLV_Feline_leukemia_virus_gi|4
HPV_human_polio_virus_gi|33289
Telomerase_C_lupus_familiaris
Telomerase_M_musculus_gi|66782
Telomerase_G_gallus_gi|7189653
Telomerase_Z_mays_gi|162460797
Telomerase_O_sativa_gi|1362530
Telomerase_C_sativus_gi|449463
Telomerase_A_thaliana_gi|33200
Retron_S_enterica_gi|695199680
Retron_E_coli_gi|425246112|ref
Retron_P_stutzeri_gi|392423130
Retron_P_syringae_gi|28871080|
Retron_E_amylovora_gi|29248655
G2_intron_S_avermitilis_gi|296
G2_intron_C_hamiltonella_gi|22
L1Hs_human_LINE1_gi|106322|pir
L1_Md_LINE-1_M_musculus_gi|603
G2_intron_G_max_gi|840924|gb|A
G2_intron_Z_mays_gi|840925|gb|
G2_intron_B_thuringiensis_gi|4
AT5G46920
Tx1_R_toruloides_gi|472584288|
Jockey_C_canorus_gi|696992601|
Jockey_D_melanogaster_gi|13408
D_pulex_E9HKV0
D_purpureum_F0ZAN5
M_mulatta_H9H4M8
M_truncatula_XP_003629388
O_sativa_AAG13524
F_vesca_XP_004310201
I_ricinus_V5HHM5
AT4G04650
A_thaliana_AAD12028
M_truncatula_XP_003611652
C_appendiculata_U5EF69
A_thaliana_AAF79665

Suppl. Fig. 8.



```
Alignment: C:\Users\yizune\Desktop\thesis\filogenias articulo\origen zf-RVT\DominioRT_R_zfRVT_R
Seaview [blocks=10 fontsize=10 A4] on Mon May 23 21:32:33 2016

                                1
WHV_gi|15637597|gb|AAL04548.1|  ---------V TIKSPRTPRR ITGGVFLVDK NSSESRLVVD FSQFSTRVHW
GSHV_Ground_squirrel_hepatitis  ---------- ---------- ---------- -SSESRLVVD FSQFSSRVHW
HBV_gi|356467132|gb|AET09698.1  ---------- ---------- ---------- -TAESRLVVD FSQFSTRVSW
DHBV_gi|2982231|gb|AAC06354.1|  ---------Y LRGNTSWPNR ITGRLFLVDK NTEEARLVVD FSQFSNAMRF
Gypsy_A_thaliana_gi|6598587|gb  ---------- ---------- ---------- ---------- ---LNRVTLK
Gypsy_Z_mays_gi|146749389|gb|A  ---------G FIRPS--TSP WGCPALFVKK KDESLRLCID YRPLNAVTIK
Gypsy_Y_lipolytica_gi|49256739  ---------G IIRPS--KSP YSAPLVIVKK KGGELRICTD YRALNELTTK
Gypsy_D_melanogaster_gi|130583  ---------G IIRPS--RSP YNSPIWVVDK KNPNKRLVID FRKLNEKTIP
MMV_Mirabilis_mosaic_virus_gi|  ---------K VIIPS--KSQ HMSPAFLVEK ERGKKRMVVN YKKLNEVTIG
FMV_figwort_mosaic_virus_gi|20  ---------G LIIPS--KSQ HMSPAFLVEN ERGKKRMVVN YKAINQATIG
CERV_carnation_etched_ring_vir  ---------K VIKPS--KSP HMSPAFLVEN ERGKKRMVVN YKAINKATKG
CaMV_cauliflower_mosaic_virus_  ---------- --------SP HMAPAFLVNN ERGNKRMVVN YKAMNKATVG
HIV1_human_immunodeficiency_vi  ---------G KISKIGPENP YNTPVFAIKK KSTKWRKLVD FRELNKRTQD
BIV_bovine_immunodeficiency_vi  ---------G KISEAAWDNP YNTPVFVIKK KTGKWRMLMD FRELNKITVK
AEV_avian_endogenous_retroviru  ---------G HIEPS--LSR WNTPVFVIQK KSGAFRLLHD LRAVNSQLIP
RSV_rous_sarcoma_virus_gi|2195  ---------G HIEPS--LSC WNTPVFVIRK ASGSYRLLHD LRAVNAKLVP
Copia_LTR_S_lycopersicum_gi|13  ---------D LVIQPSGRKI ITCKWVFKKK ECVKYKARVV ARGFNEIFSP
Copia_LTR_A_thaliana_gi|699625  ---------T TVTRPEKQRI IGSRWIYKYK QEPRFKARLV AKGYAEIFAP
Tnt_gi|130582|sp|P10978.1|POLX  ---------K LVELPKGKRP LKCKWVFKLK KLVRYKARLV VKGFEEIFSP
Copia_Drosophila_simulans_gi|3  ---------- ---------- VDSKWVFSVK YPIRYKARLV ARGFTETFAP
copia_Zea                       ---------T LVPPDRTRNL IDCKWVFKVK YIDRLKARLV ARGFKDTFSP
Gypsy_M_yessoensis_gi|12405524  ---------- ---------- ---------- -DKSIRICGD YKVTVKVSKL
Ty1_S_cerevisiae_gi|836753|dbj|  ---------D TNKYYD-RND IDPKKVINSM FDGTHKARFV ARGDIDMQSN
FLV_Feline_leukemia_virus_gi|4  ---------- ---------- -------VKK PTEDYRPVQD LREVNKRVED
HPV_human_polio_virus_gi|33289  ---------- ---------- DGLEALDLST SAGYPYVAMG KKTKEMQRLL
Telomerase_C_lupus_familiaris_  ---------- ---------- ---------- ---------- ----------
Telomerase_M_musculus_gi|66782  ---------- ---------- ---------- ---------R RQVTTLSDLQ
Telomerase_G_gallus_gi|7189653  ---------- ---------- ---------- ---------- ----------
Telomerase_Z_mays_gi|162460797  ---------- ---------- ---------- ---------- ----------
Telomerase_O_sativa_gi|1362530  ---------- ---------- ---------- ---------- ----------
Telomerase_C_sativus_gi|449463  ---------- ---------- ---------- ---------- ----------
Telomerase_A_thaliana_gi|33200  ---------- ---------- ---------- ---------- ----------
Retron_S_entérica_gi|695199680  ---------K NKKERQIFEP KRSLKGIQKK INKEIFEKID YPHYLHG---
Retron_E_coli_gi|425246112|ref  ---------K NKKERQIFEP KRSLKGIQKK INKEIFEKID YPHYLHG---
Retron_P_stutzeri_gi|392423130  ---------- --KDRVVYEP KLNLKKIQKR INSRIFEHVQ FPPYLQGGIK
Retron_P_syringae_gi|28871080|  ---------K GDKERMVYEP KYELKKLQKR INSRLFEKVH YPPYLQGGVR
Retron_E_amylovora_gi|29248655  ---------- ---------- --MLKSLLRK INQKIFTKIN YPDYLPGSIP
G2_intron_S_avermitilis_gi|296  RANGKRRPLG IPVIRDRVDQ ARPKNALEPE WEARPEARSY GPRPGRGAWD
G2_intron_C_hamiltonella_gi|22  ---------- ---------- ----SALEDY WEAKPEDVSY GPRPGRSAHD
L1Hs_human_LINE1_gi|106322|pir  TKRENFRPIS LMNIDAKILN KILANQIQQH IKKLIHHDQV GFIPGMQGWF
L1_Md_LINE-1_M_musculus_gi|603  TKIENFRPIS LMNIDAKILN KILANRIQEH IKAIIHPDQV GFIPGMQGWF
G2_intron_G_max_gi|840924|gb|A  KEGSKFRPLT VVLPIEKIMK EAIRMVLESI YDPEFPDTSH -FRSGRGRYS
G2_intron_Z_mays_gi|840925|gb|  KEGSKFRPLT VVLPIEKIMK EAIRMVPESI YDPEFPDTSH -FRSGRGCHS
G2_intron_B_thuringiensis_gi|4  KKNGKKRPLG IPSFADKLVQ EVVRMILEAI YEPTFSIESH AYQKGKSCHT
AT5G46920                       ---AILTTTQ PTPPQDRIVQ EVLLMILEPI YEERFSQKSF AFRPGRTAHT
Tx1_R_toruloides_gi|472584288|  AAFASYRPIS LRDADYRILG RVLVKRLNPL LPLAIPLSQV GFVPGRRSAD
Jockey_C_canorus_gi|696992601|  ENPGNYRPVS LTSVPGKVME QVILSAIMKH MQEGIRPSQH GFRKGRSCQT
Jockey_D_melanogaster_gi|13408  TDVDSYRPTS LLPSLGKIME RLIILNRLTC KDVAIPKFQP GFRLQHGTPE
D_pulex_E9HKV0                  ---------- ---------- ---------- ---------- ----------
D_purpureum_F0ZAN5              ---------- ---------- ---------- ---------- ----------
M_mulatta_H9H4M8                ---------- ---------- ---------- ---------- ----------
M_truncatula_XP_003629388       ---------- ---------- ---------- ---------- ----------
O_sativa_AAG13524               ---------- ---------- ---------- ---------- ----------
F_vesca_XP_004310201            ---------- ---------- ---------- ---------- ----------
I_ricinus_V5HHM5                ---------- ---------- ---------- ---------- ----------
AT4G04650                       ---------- ---------- ---------- ---------- ----------
A_thaliana_AAD12028             ---------- ---------- ---------- ---------- ----------
M_truncatula_XP_003611652       ---------- ---------- ---------- ---------- ----------
C_appendiculata_U5EF69          ---------- ---------- ---------- ---------- ----------
A_thaliana_AAP79665             ---------- ---------- ---------- ---------- ----------
```