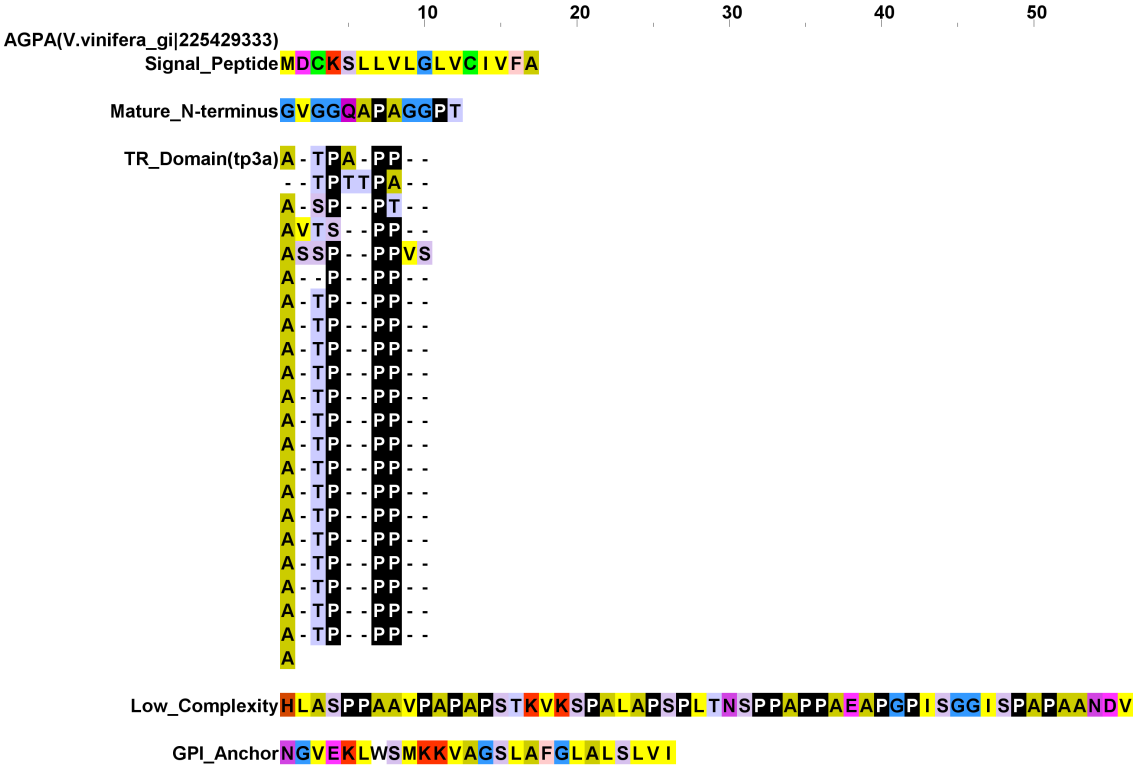
Dataset S1. TR architectures of representative proteins from 31 Pro-rich TRP classes.

A AGPA – TR-AGP type alpha.



B AGPB – TR-AGP type beta.

```
                                        10        20        30        40
AGPB(B.napus_TC77477)
          Signal_Peptide MASSSQAFLLLTLSMVLVHFSLA

      Mature_N-terminus QSPMMAPSGSMSMPPMPSGGS

       TR_Domain(mtp3) PMPMMTPP
                       PMPMMTPP
                       PMPMMTPP
                       PM-AMAPP
                       PMP-MTPP
                       PMP-M--A
                       PMPM-TP

     Low_Complexity SSSPMSPPTTMAPSPETVPDMASPPMMPGMDSSPSPGPMPPAMASPD

         GPI_Anchor SGAFNVRNDVVAISFLVAAHLLLV
```

C AGPC – TR-AGP type gamma.

```
                                        10        20        30        40        50
AGPC(T.aestivum_gi|62861391)
          Signal_Peptide MASPRGLAGLLCFALAAAAASA

Cu_bind_like_domain(PF02298) AQYRVGEQRGWSVPAAGAEPLNTWSARMRFIIGDQLLFVYPKDTDSVLLVDQAAYNACNT
                             TTYVSKFQGGSTVFTLDRSGPFFFISGNEASCKAEQKLIVVVLSVDHTPPGLLPPPP

     TR_Domain1(mtp2) PPSM-PPSSAPPMPSPMS
                      PPSMTPP-SAAPMPSPMS
                      PPSMTPP-SAAPMPSPMS
                      PPSMTPP-SAAPMPSPMS
                      PPSMTPP-SAAPMPSPMS
                      PPSMAPP-SSEPMPSPMT
                      P

                      AAAPGTTPVSPASPAGPAP

        TR_Domain2 SPGTPGGGS
                   SPGTPGSDT
                   S

                   SPPAPAADGANSTTPG

       GPI_Anchor SAAAPVTAGLIGTLAAGIGYAMLAI
```

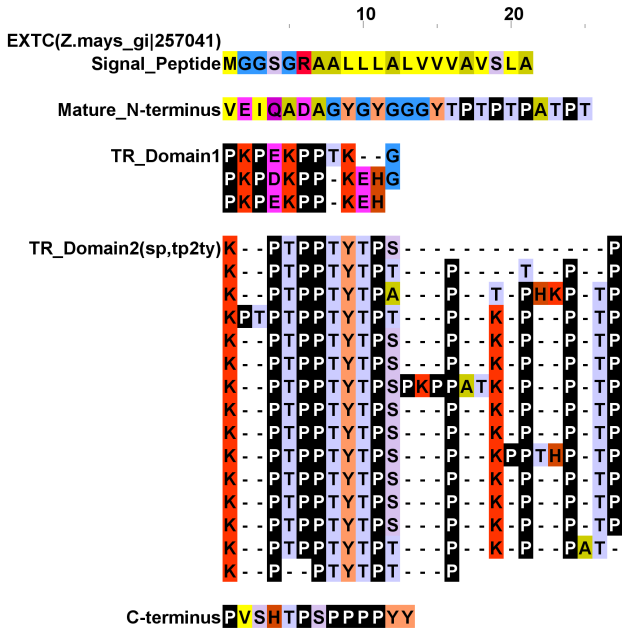
D EXTA – Extensin type alpha.

```
                                        10        20
EXTA(AtEXT6)
Signal_Peptide          MGIENCIVYVVVLSAIAATVTS

Mature_N-terminus       YPYSSPQTPSYNSPSYEHKGPKYAPHPK

TR_Domain(sp4+,yxy)     PYVKSSPPPQYYTPSPKVNYKSPPP
                        PYVYSSPPPPYYSPSPKVDYKSPPP
                        PYVYSSPPPPYYSPSPKVDYKSPPP
                        PYVYSSPPPPYYSPSPKVDYKSPPP
                        PYVYNSPPPPYYSPSPKVDYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYNSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVYYKSPPP
                        SYVYSSPPPPYYSPSPKVYYKSPPP
                        SYVYSSPPPPYYSPSPKVYYKSPPP
                        PYVYSSPPPPYYSPSPKVTYKSPPP
                        PYVY

C-terminus              KTPYY
```

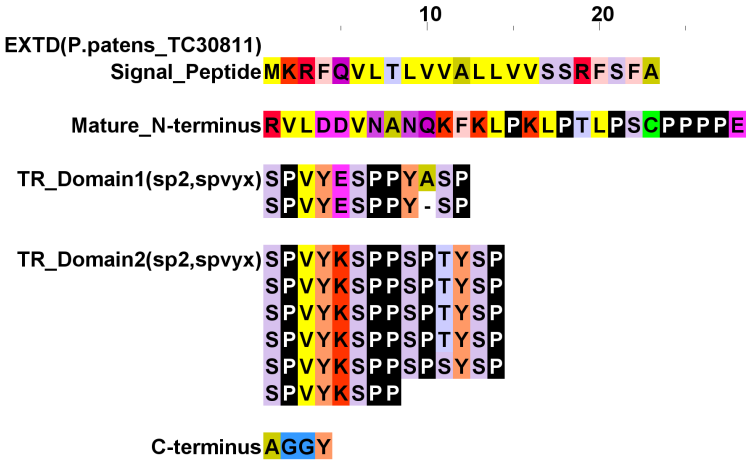
E EXTB – Extensin type beta.

```
                                        10        20        30        40        50
EXTB(S.tuberosum_gi|296401)
Signal_Peptide          MGKMASLVATLLVVLVSLSLA

Mature_N-terminus       SESSANYQYSSPPPPVHVYPSPPHHPVYKSPPPHHHHPVYKSPPPSEKPHYPPHTPVYKS
                        PPPHHHHPVYKSPPPPT

TR_Domain1(sp4+,vyk)    PVYKSPPPPKTPHY--PPHTPVYKSPPP-HHH-H
                        PVYKFPPPP-TPVYKSPP-PP--KD--PHYPPHT
                        PVYKSPPPP-TPVYKS---PP------P---P-T
                        PVYKSPPPP-TPVYKSPP-PPV-K---P-YHP-A
                        PVYKSPPPP-TPVYKS---PPV-K---P-YHP-A
                        PVYKSPPPP-TPIYKSPP-PPV-K---P-YHP

                        SPTPYHPK

TR_Domain2(sp4+,vyk)    PVYKSPPPPT-
                        PVYKS-PPPTH
                        YVYSSPPPP

C-terminus              YHY
```

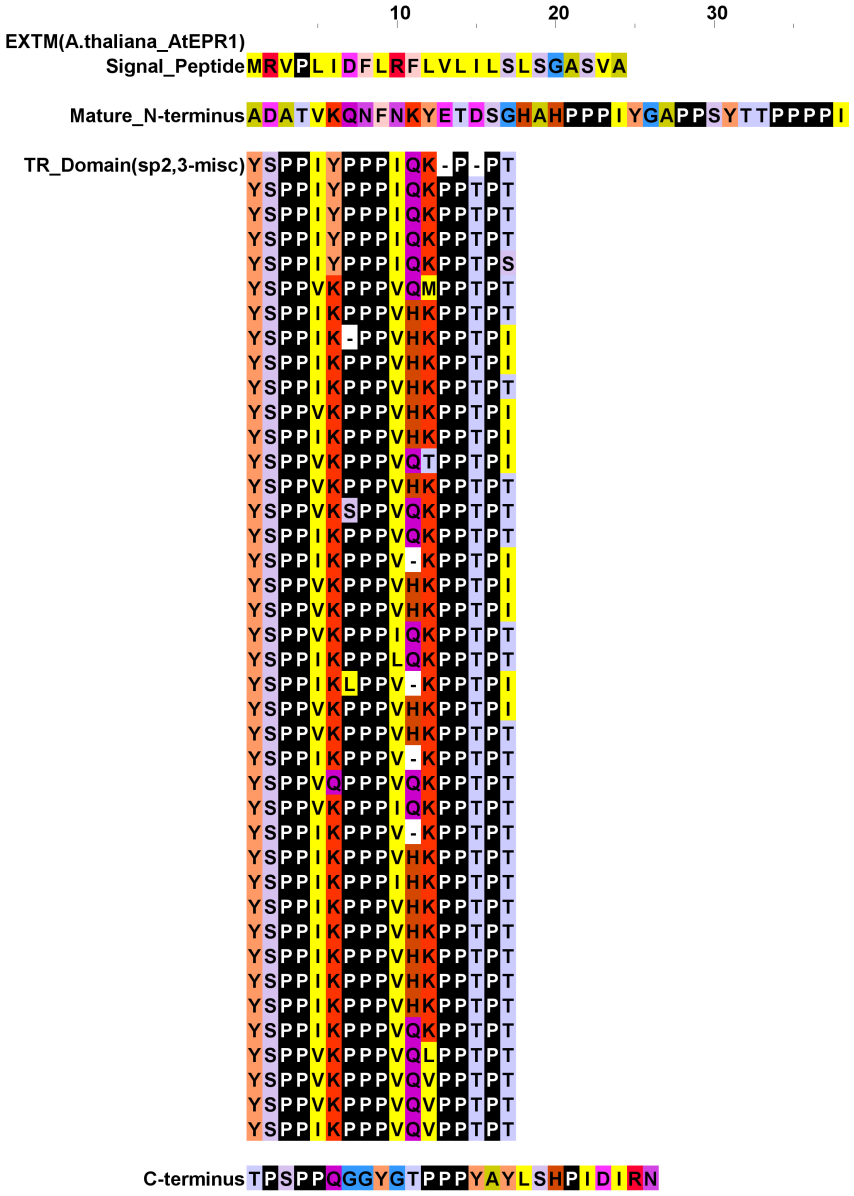
F EXTC – Extensin type gamma.



G EXTD – Extensin type delta.



H EXTM – Extensin type miscellaneous (S/T)$P_{2,3}$.



I HEXA – Hybrid Extensin type alpha.

```
                                   10        20        30        40        50
HEX(AtPEX4)
Signal_Peptide MPFYKQPWVFSKVFVLAMAKPPSFGCCFFLLFFSFLSSSFVSFA

LRR_Domain     LTDTEAAFIVQRQLLTLPDNGELPDDIEYEVDLKATFANTRLKRAYIALQAWKKAIFSDP
               FNTTGNWHGPHVCGYTGVVCAPALDDSDVTVVAGVDLNGADIAGHLPAELGLMTDVAMFH
               LNSNRFCGIIPKSFEKLKLMHEFDVSNNRFVGPFPNVVLSWPDVKYFDLRFNDFEGQVPP
               ELFKKELDAIFLNDNRFTSVIPESLGESPASVVTFANNKFTGCIPKSIGNMKNLNEIVFM
               DNDLGGCFPSEIGKLSNVTVFDASKNSFIGRLPTSFVGLTSVEEIDISGNKLTGLVPHNI
               CQLPNLVNLTYSYNYFSGQGGSCVPGGSRKEIALDDTRNCLASRPEQRSAQECAVVINRP
               VDCSKDKCAGGSSTPSKPS

TR_Domain1     PVHKPTPVPTT
               PVHKPTPVPTT
               PVQKPSPVPTT
               PVQKPSPVPTT
               PVHEPSPV

               LATPVDKPSPVPSRPVQKPQPPKESPQPDDPYDQSPVTKRR

TR_Domain2(sp4vh) S---PPPAPVN
               S---PPP-PVY
               SPPPPPP-PVH
               S---PPP-PVH
               SP--PPP-PVY
               SPPPPPP-PVH
               S---PPP-PVF
               S---PPP-PVY
               S---PPP-PVH
               S---PPP-PVH
               SP--PPPAPVH
               S---PPP-PVH
               SPP-PPP-PVY
               S---PPP-PVF
               S---PPP

               SQSPPVVYSPPPRPPKINSPPVQSPPPAPVEKKETPPAHAPAPSDDEFIIPPFIGHQ

C-terminus     YASPPPPMFAGY
```

J HLTA – Hybrid PRP/Lipid Transfer Protein type alpha.

```
                                        10        20        30        40        50
HLTA(L.tulipifera_TA1304_3415)
            Signal_Peptide MDSSKISALLLIFMLFVSSA

     Mature_Signal_Peptide SPILACTYCKNPPKHKSPK

          TR_Domain(p2vtl) GKPPKGPINLPPVI
                           VK-P--PITRPPVV
                           GK-P--PITLPPVI
                           GK-P--PVTLPPVI
                           GK-P--PITLPPVI
                           GK-P--PVTLPPVI
                           GK-P--PITLPPVI
                           GK-P--PITLPPVI
                           GK-P--PITLPPVI
                           GK-P--PITLPPVI
                           GK-P--AMTLPPVI
                           GK-P--PITLPPVI
                           GK-P--PWTLPPVI
                           GK-P--PITLPPVI
                           GK-P--PITLPPVI
                           GK-P--PITLPPVI
                           GK-P--PITLPPV

            LTP_Domain IGTPPTTKPPGCGCSPPPKEETCPIDTLKLGACVDLLGGLVHIGLGDPVVNTCCPLLQGL
                       VEIEAAVCLCTTIKLKLLNLKIFVPIALQLLATCGKTPPPGFTCSL
```

K HLTB – Hybrid PRP/Lipid Transfer Protein type beta.

```
                                        10        20        30        40        50
HLTB(M.sativa_PRP2)
            Signal_Peptide MANYALANVFILLLNLSTLLIVLA

      Mature_N-terminus CPYCPYPSPKPPTHHPPIVKPPVHKRRKYSPTPKPPVHKPPRYPPKPSPCPPPSST

            TR_Domain1 PKPPHV
                       PKPPHH
                       PKPP-V
                       VHPPHV
                       PKP

TR_Domain2(p2yxpkp2) P-VHPPYVPKPPIVKP-
                     PIVHPPYVPKPP-V-VK
                     P---PPYVPKPP---V-
                     --VRPPYVPKPP-V-V-
                     P-VTPPYVPKPP---V-
                     --VRPPYVPKPP-V-V-
                     P-VTPPYVPKPP-I-V

            TR_Domain3 KPPIVFPPHVPLPPVVPSPPPYV-PSPPIV
                       KPPIVFPPHVPLPPVVPVTPPYVQP-PPIV
                       TPP

            LTP_Domain TPTPPIVTPPVVSPPTPPSETPCPPPPLVPYPPTPPAQQTCSIDALKLGACVDVLGGLIH
                       IGIGGSAKQTCCPLLQGLVDLDAAICLCTTIRLKLLNINLVIPLALQVLIDCGKTPPEGF
                       KCPAY
```

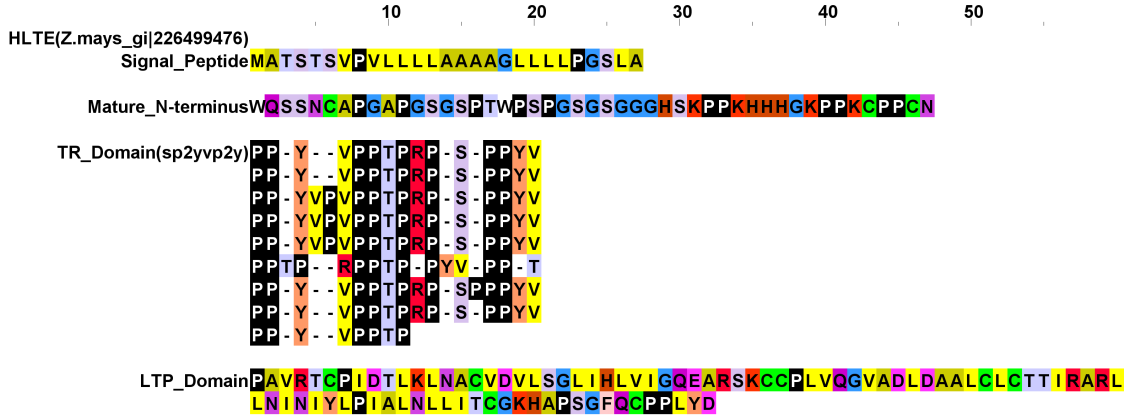
L HLTC – Hybrid PRP/Lipid Transfer Protein type gamma.

```
                                        10        20        30        40        50
HLTC(A.thaliana_gi|18396941)
            Signal_Peptide MDSSKLSSLSLCLFLICIIYLPQHSLA

        Mature_N-terminus CGSCNPRKGGKHSPKA

         TR_Domain(p2vtv) PKLPVPPVTV
                          PKLPVPPVTV
                          PKLPVPPVTV
                          PKLPVPPVTI
                          PKLPVPPVTV
                          PKLPVPPVTV
                          PKLPVPPVTV
                          PKLPVPPVTV
                          PKLPVPPVTV
                          PKLPVPPVTV
                          PKLP

            LTP_Domain    LPPISGLPIPPVVGPNLPLPPLPIVGPILPPGTTPPATGGKDCPPPPGSVKPPSGGGKAT
                          CPIDTLKLGACVDLLGGLVKIGLGDPAVNKCCPLLKGLVEVEAAACLCTTLKLKALDLNL
                          YVPVALQLLLTCGKNPPPGYTCSI
```

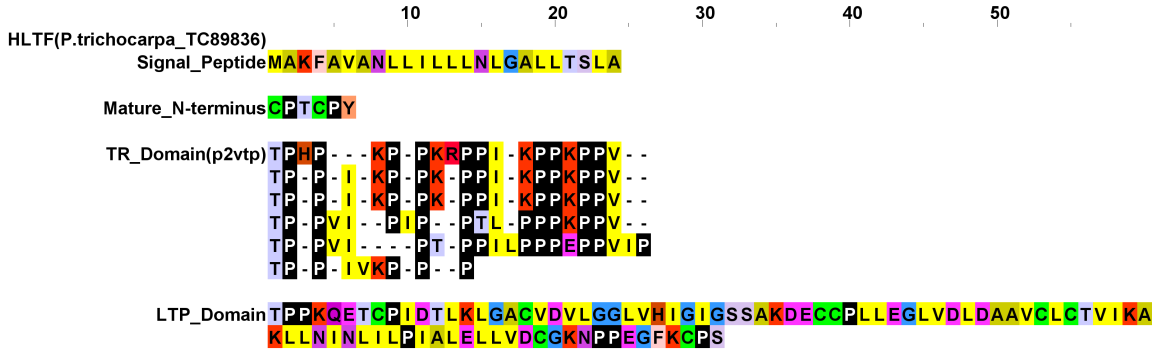
M HLTD – Hybrid PRP/Lipid Transfer Protein type delta.

```
                                        10        20        30        40        50
HLTD(A.thaliana_gi|42565109)
            Signal_Peptide MGSRSQNLSFLVLLLLGFVAVSYA

        Mature_N-terminus CDCTPPKPSPAPH

     TR_Domain1(p2tvk)    K--P-P-K-HP--V
                          K--P-P-K-PP-AV
                          K--P-P-K-PP-AV
                          KP-PTP-K-PP-TV
                          K--PHP-K-PP-TV
                          K--PHP-K-PP-TV
                          K--PHP-K-PP-TV
                          KP-PHP-K-PP-T-
                          KPHPHP-K-PP-IV
                          K

            TR_Domain2    PPTKPPPSTPK-
                          PPTKPPPSTPKP
                          PTTKPPPSTPKP
                          PHHKPPP-TPCP
                          PPT

            TR_Domain3    PTPTPPVVTP
                          PTPTPPVITP
                          PTPTPPVVTP
                          PTPTPPVITP
                          PTPTPPVITP
                          PTPTPPVVTP
                          PTPTPPVVTP
                          PTPTPP--TP

            LTP_Domain    IPETCPIDTLKLGACVDVLGGLIHIGLGKSHAKAECCPVLGGLLDLDAAVCLCTTIKLKL
                          LNIDLVLPIALELLLDCGKTPPSDFKCPA
```

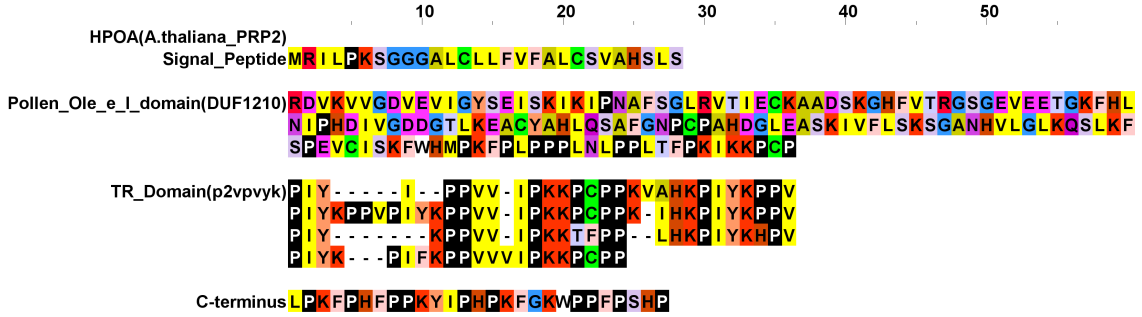
N HLTE – Hybrid PRP/Lipid Transfer Protein type epsilon.



O HLTF – Hybrid PRP/Lipid Transfer Protein type phi.



P HPOA – Hybrid PRP/Pollen Ole e I type alpha.



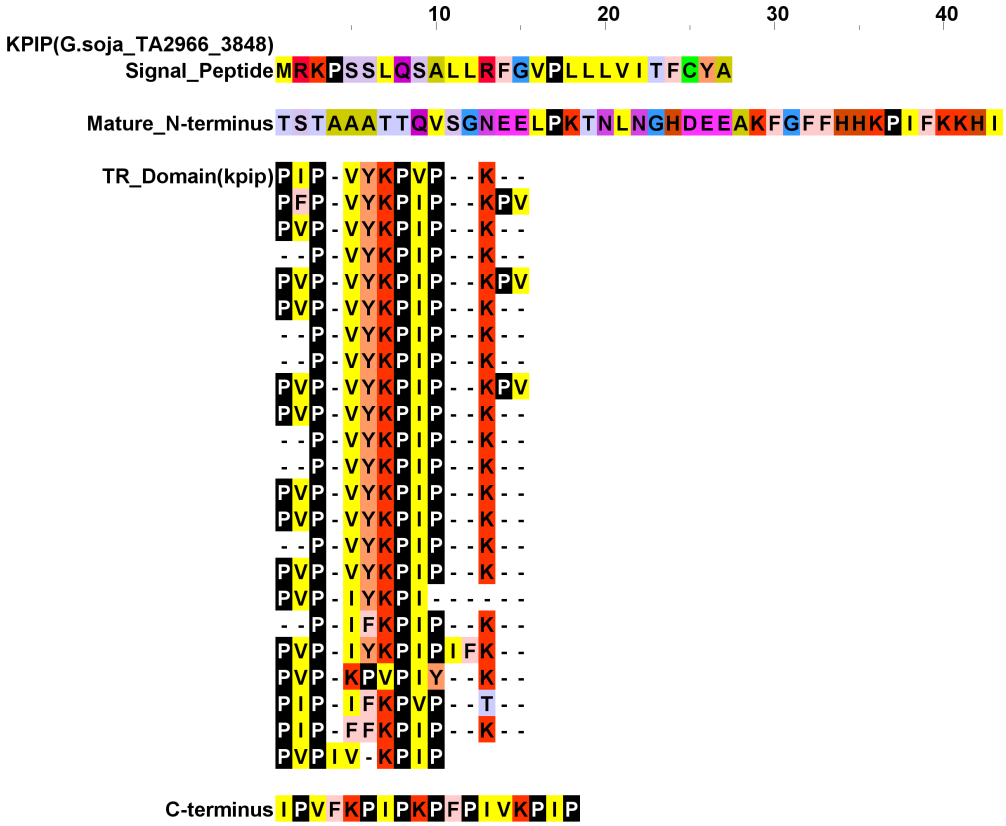
Q HPOB – Hybrid PRP/Pollen Ole e I type beta.

```
                                        10        20        30        40        50        60        70
HPOB(Z.mays_gi|162461641)
          Signal_Peptide MAADALRGLLPGVYAVLMVIAVTSA

Pollen_Ole_e_I_domain(DUF1210) ASSDASSVVVGLAKCADCTRKNLKAEAAFKGLEVAIKCKNSKGEYESKAMGKLDGFGAFS
                               IPLSTDLHGADCVAQLHSAAGNPCPGQEPSRIVPQSSKGNFVVVPGKTDYPSKECASATL
                               CGSIKKHLLDHFHKKPVPPKPKPE

TR_Domain(pepk-embed1) PKPE--LEPKPEPKPEPENHPPTPTYGSSTPIYHPP--SRQLFDKKHMF-DNFHKHHDYHHFFDHFHKKPVPPK
                       PKPEPKPEPKPKPQ--PEYHSPTPTYGSPTPIYHPPA--RHLFDKKP-LLDHFHKDHDYHHFLDHFHKKPVPQK
                       PK----LE--PKPQPKPEYHPPTPTYVSPTPIYHPPAETEMKYDKKH-F-DHFHKDHDYHHFFDHFHKKPVPPK
                       PKPE--PEPKPKPQPKPEYKPPMPTYGSPTPIYHPPAKKEVAFDKKH-FLDHFHKDHDYHHFFDHFHKKPVPPK
                       PKPE--PKPEPKPQPEPEYHPPTP

C-terminus TYGSPTPIYHPPAKH
```

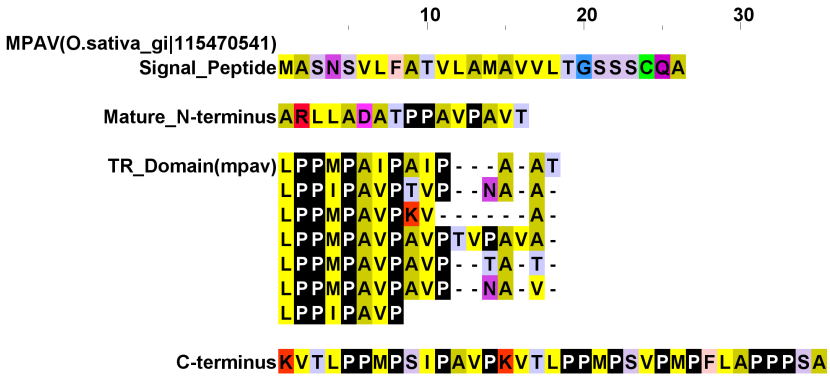
R HPOC – Hybrid PRP/Pollen Ole e I type gamma (AGP).

```
                                        10        20        30        40        50
HPOC(A.thaliana_AGP31)
          Signal_Peptide MGFIGKSVLVSLVALWCFTSSVFT

Mature_N-terminus EEVNHKTQTPSLAPAPAPYHHG

TR_Domain1 HHHPHP-P
           HHH-HPHP
           HPHPHP-P

           AKS

TR_Domain2(p2vk) PVKPPVKAPVSPPAKPPVKP
                 PVYPPTKAPVKPPTKPPVKP
                 PVSPPAKPPVKPPVYPPTKA
                 PVKPPTKPPVKPPVYPPTKA
                 PVKPPTKPPVKPPVYPPTKA
                 PVK

TR_Domain3(p2vk) PPTKPPVKPPVS
                 PPAKPPVKPPVY
                 PPTKAPVKPPVS
                 PPTKPPVTPPVY
                 PP

Pollen_Ole_e_I_domain(DUF1210) KFNRSLVAVRGTVYCKSCKYAAFNTLLGAKPIEGATVKLVCKSKKNITAETTTDKNGYFL
                               LLAPKTVTNFGFRGCRVYLVKSKDYKCSKVSKLFGGDVGAELKPEKKLGKSTVVVNKLVY
                               GLFNVGPFAFNPSCPK
```

S KPIP – KPIP domain protein.



T MPAV – MPAV domain protein.



U PEHK – PHEK domain protein.

```
                                        10        20        30        40        50
PEHK(V.vinifera_gi|225433073)
            Signal_Peptide MSPTYLLVVLLGLVVLTAPSLA

         Mature_N-terminus DYPTHPPF

       TR_Domain(p3ehk1) EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPLHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHKPPV
                         EKPPPEHK---
                         EKPLPEHKPP

            C-terminus TPVGKPPKGEKPPHCGHNPGHPPAENAEDSYKPPRKIKPPSTPAEKQPGPGKKLPTPPHK
                         PPHKPPTPTHPN
```

V PELPK – PELPK domain protein.

```
                                        10        20
PELPK(O.sativa_gi|115473185)
            Signal_Peptide MVFKKNAMSSSVLFLAALLLLSSSSMSSA

         Mature_N-terminus ARWLEEEY

       TR_Domain(pelpk2) PPHPT-VPELPKPEV
                         PPHPA-VPELPKHEE
                         PPHPV-VPELPKHEE
                         PPHPV-VPELPKPEL
                         PPHPV-VPELPKHEE
                         PPHPAVVPELPKHEE
                         PPHPAVVPEFPKHEE
                         PPHPA-VPELP-H--
                         ---PA-V---P---E
                         IPHPA-VPELPKHEE
                         PPHPV-VPELPK-PE
                         VPHAA-VPELPKPEL
                         PPHPA-VPELPKHEE
                         PPHPV-VPELPKHEE
                         PPHPV-VPELPKPEE
                         PHHPE-VPE

            C-terminus HEQPPKPESHYPEVPMAKP
```

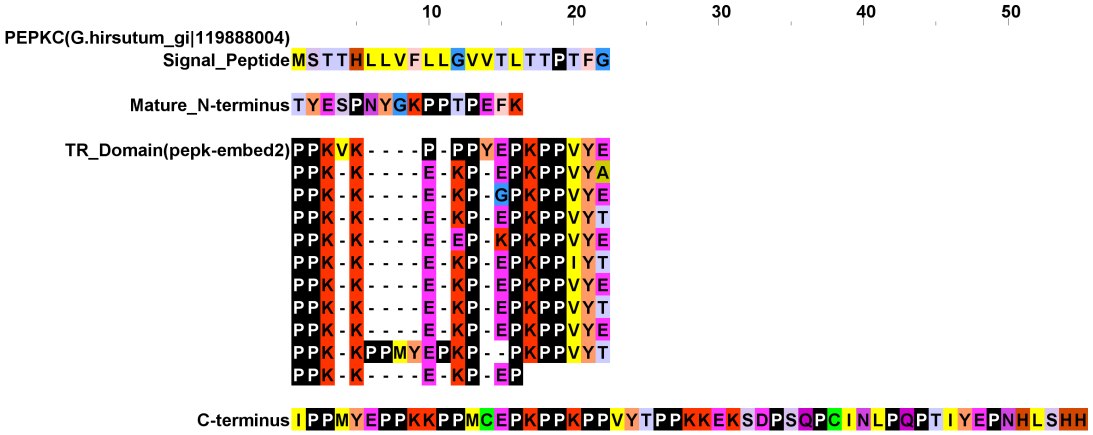
W PEPKA – PEPK type alpha.

```
                                        10        20
PEPKA(H.vulgare_TC165053)
          Signal_Peptide MARHSLLAVLLVGLLAASGFSLAAGAGRGLA

   Mature_N-terminus EKLPEPEPKPTPY

     TR_Domain(pepk) PE--PKPM--P-KPE--PMPK
                     PE--PTPM--P-KPE--PMPK
                     PE--PKPM--P-KPE--PMPK
                     PE--PKPE--P-KPE--PM--
                     ----PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-MPK--PEPK
                     PE--PKPE--PMKPEPKPMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPY--PMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPE--PMPK
                     PE--PKPE--P-KPE--PM-K
                     PE--PKPM--P-KPEPKPEPK
                     PEPMPKPE--P-KPE--PMPK
                     PE--PKPEPMP-KPE--PKPY
                     PM--PKPE--P-KPQ--PMPK
                     PE--PKPE--P-EPK--PEP

      C-terminus PPKGKPPMTDN
```

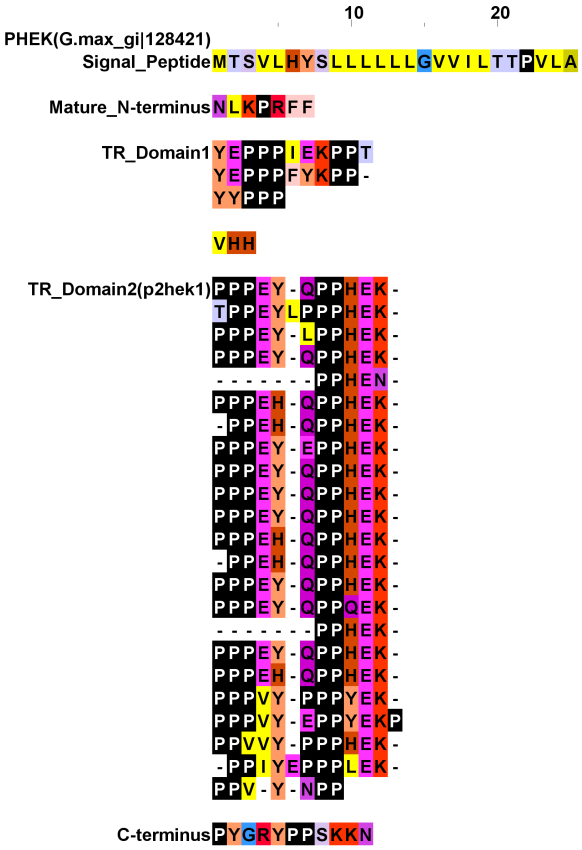
X PEPKB – PEPK type beta.

```
                                        10        20        30
PEPKB(O.sativa_gi|115471749)
          Signal_Peptide MRRSILSLCFHLALVIALAANVPDIANG

   Mature_N-terminus RVIEAKSDPKPADPKPKPDPTPKPQPETKPS

     TR_Domain1(pdpk) P-QPNPQPNPQPDPKPSP-QPD
                      P-KPTPQPEPKQDPQPNP-QPD
                      PKQ-SPQPDPKPTPQPNPKQ-D
                      P-QPNPQPDPKPTLQPNPKQ-D
                      P-QPNPQPNPKPTPQLDPKQ-D
                      P-QPNPQP

                      SPKAD

     TR_Domain2(pdpk) PKPNPKPKPQPEPSPNPKPEPKPEPKP-EP-SPN
                      PKPNPNPK--PEPQPDPKPEPKPQPEPSQPKLPP

      C-terminus LSPAIAIIVPGN
```

Y PEPKC – PEPK type gamma



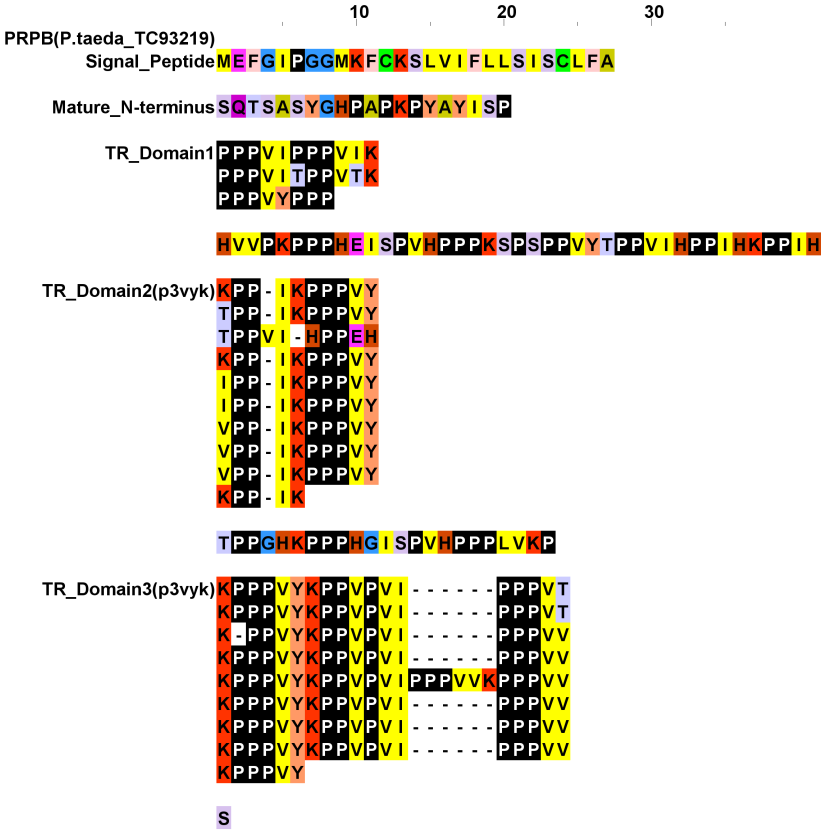
Z PHEK – PHEK domain protein.



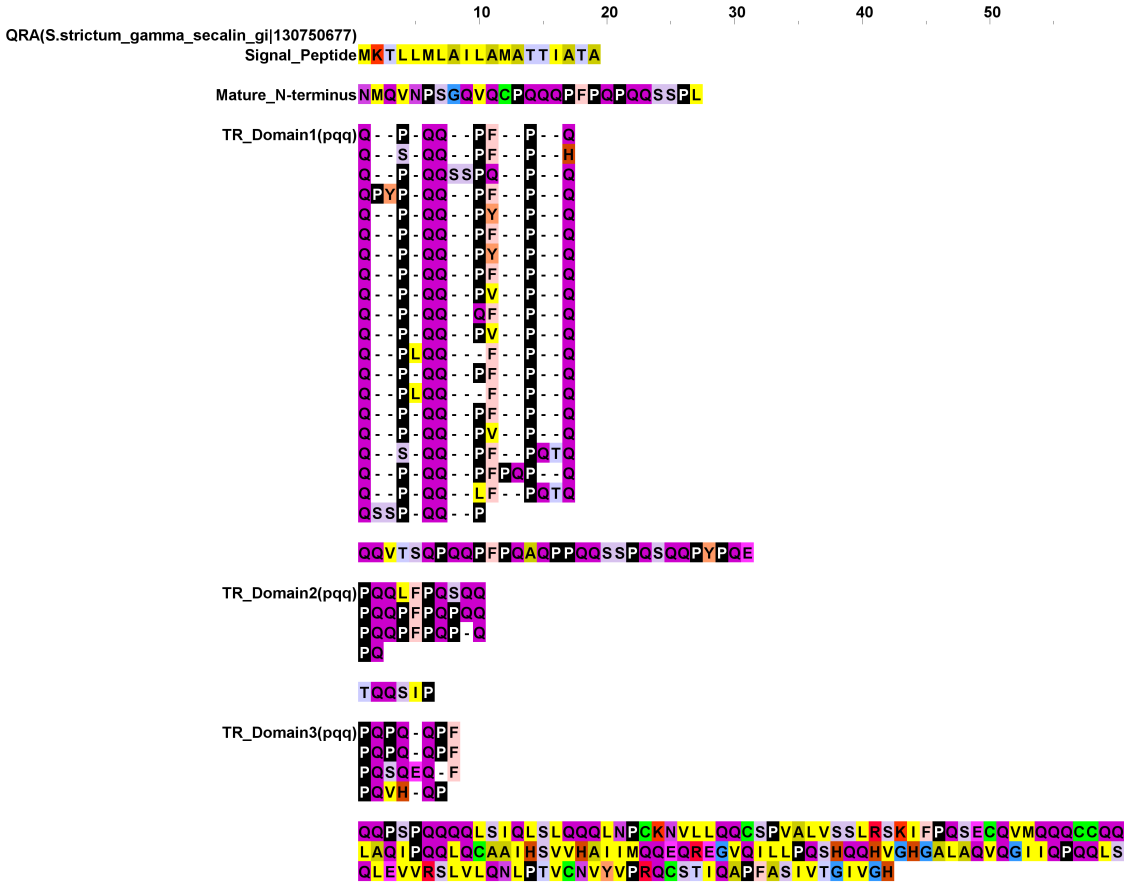
AA PRPA – Proline-Rich Protein type alpha.

```
                                        10        20
PRP(M.truncatula_gi|410107)
          Signal_Peptide MASSNFLVLLLFALFAIPRGLA

     Mature_N-terminus NYD

     TR_Domain(p2vyk) KPPVYQPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPIY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVEKPPVY
                      KPPVYKPPVY
                      KPPVEKPPVY
                      KPPVYKPPVE
                      KPPVYKPPVY
                      KPPVEKPPVY

          C-terminus GPPHHP
```

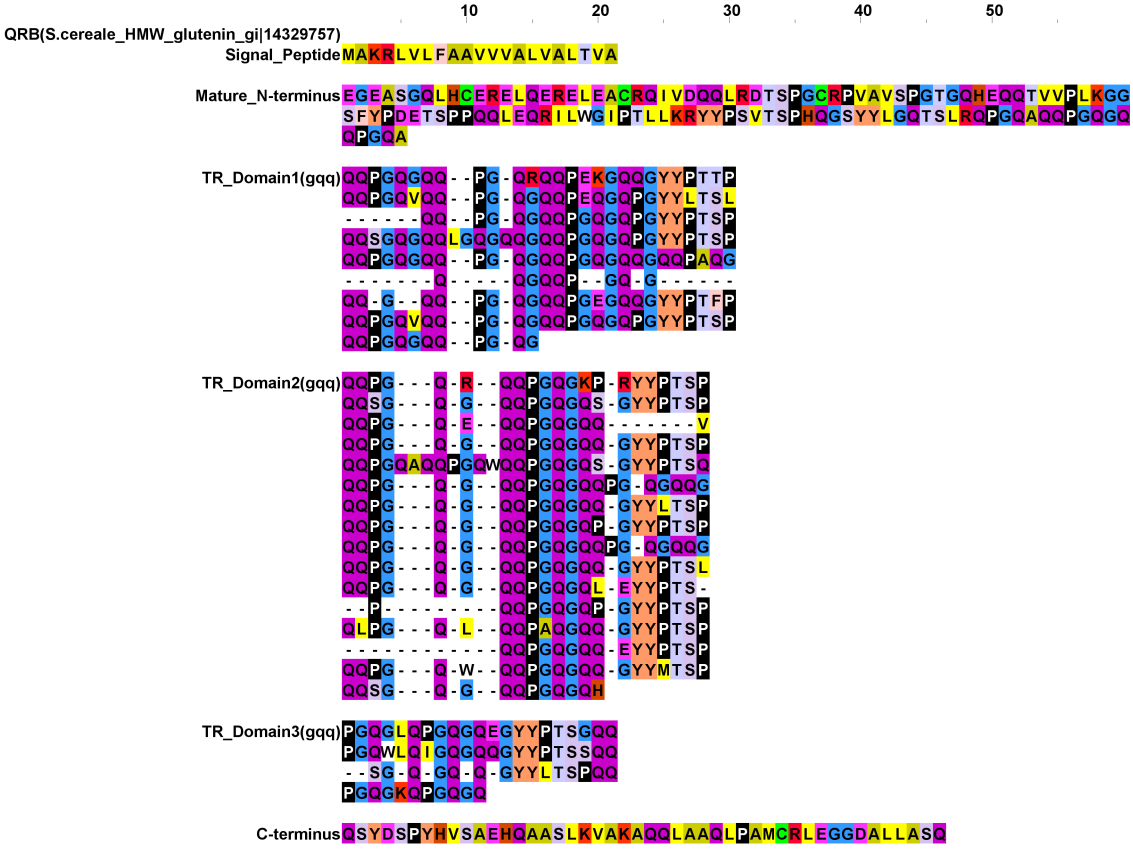
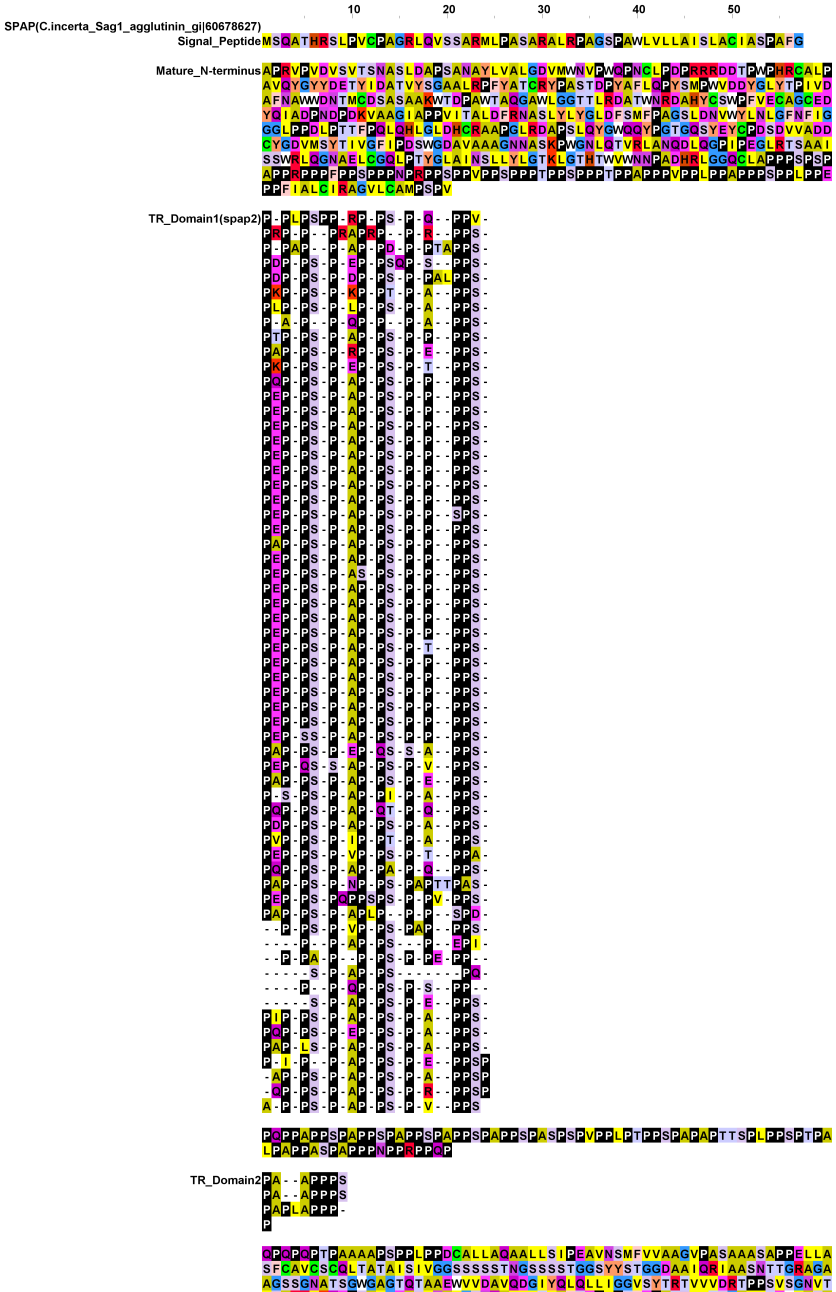
AB PRPB – Proline-Rich Protein type beta.



AC QRA – Gln-rich type alpha.



AD QRB – Gln-rich type beta.



AE SPAP – SPAP domain protein.