

S2 Fig. Alignments of protein sequences of pan-edited dual-coding genes in *L. tarentolae*, *L. amazonensis*, *P. serpens*, and *Perkinsella CCAP1560/4* with *T. brucei* and *T. vivax* sequences. A: CR3, B: CR4, C: ND3, D: ND7 5' Editing Domain E: ND9, F: RPS12, G: ND8 (Nondual-coding). Absent sequences were unavailable. All ORF alignments show published protein sequences. All ARF alignments show +1 or +2 (ND7 only) reading frame translations of the full length mRNA sequences. In the ARF alignments of CR3, ND7 and RPS12, translations were made using the alternative *T. brucei* mRNA sequences shown in Fig 1. Two alignments of CR3 ARFs are presented to display the *P. serpens* +2 reading frame which has no stop codons. *L. tarentolae* CR4 published protein sequence shows limited homology in the C terminus to all other CR4 protein sequences. The edited mRNA has two editing sites where 13 U residues are inserted. If the second of these insertion sites is shortened to 12 U residues, the translation of this mRNA has much better homology to other CR4 proteins. Alignments with translations of the two different sequences (13U and 12U) are both shown, with the location of the altered site highlighted in red. While ND8 does not appear to be dual-coding, this alignment was included as well, for comparison of the conservation of a nondual-coding gene with that of the dual-coding genes. It should be noted that ND8 is the only nondual-coding gene that is pan-edited in *L. tarentolae*, *L. amazonensis*, and *P. serpens*, and ND7, A6 and COIII are only partially edited in these species. █=Termination codon

A. CR3

CR3 ORF Alignment

```
CR3ORF T. brucei      --MFDCLVLL--FFYCLFVHFFCFLFVCDLFLCLLFSFCFLLDFCFLFNMGLLLCLFFFFFI
CR3ORF T. vivax      --MFDCLVLL--FFLLLFVHFFCFLFICDLFLCLLFVFCFLFVDFCFLFNMGLLLCLFFFFFV
CR3ORF L. amazonensis --MFDFVIIMFL--FMSFVHFFCFLFIVDLLFCLMFFVFFLYDFCFVCNLGFCCCLFFFFFL
CR3ORF P. serpens    IFLFDFVLFLVLFLFFVHFFCFLFIIDLFCFLLLLFFLVDFCFCCCFGFVSCFLFFFV
                        : ** : : : : : ***** : ** : : : . . . ***** : * : *** : ** :

CR3ORF T. brucei      LSFDMLLSFLLLYISFRY█
CR3ORF T. vivax      LSFDMLLSFLLLYISFRY█
CR3ORF L. amazonensis LSIDMILSFILLYVVSFRY█
CR3ORF P. serpens    FHFDMVLSFILFVSFRY█
                        : : *** : *** : * : : * : * : *
```

CR3 ARF Alignment with *P. serpens* +1 RF

```
CR3ARF T. brucei      -----RNINMCMYKNNVYVVVLFWFWLYIFFVFYLFVICFYVCYLVFVYFVIFVF
CR3ARF T. vivax      -----CLI----V█FCCFFYCCLYIFFVFCLFVICFYVCCLFVYLWIFVF
CR3ARF P. serpens +1 █IVYNI█KHNLIFCL█----LFCF█CYFYCFLCFFVFYLLLCFVVFYFCFF█CLIFVF
CR3ARF L. amazonensis --K█NNMY█V█IYICLI----SLL█CFCLWVLYIFFVFYLLLCYFVWCFLFFFYMFVLCV
                        * : * . * * * * * : : * : * . * * :

CR3ARF T. brucei      YLIWVYCCVYFFFLFYHLICCYHFYCYI█VFVIRLKKYANNFC-
CR3ARF T. vivax      CLIWVCCCVYFFFLFYHLICCYRFCYI█VFVI-----
CR3ARF P. serpens +1 VVVLVLWVYFYFLFFILWFLSYLYL█VSVIKSI█KHKLIS-
CR3ARF L. amazonensis CVI█VFVVVCFFFFCYPLIWFCRLFYMYMLVSDIKI█LLFL█K
                        : : * * * : : * * * : * * *
```

CR3 ARF Alignment with *P. serpens* +2 RF

```
CR3ARF T. brucei      -----RNINMCMYKNNVYVVVLFWFWLYIFFVFYLFVICFYVCYLVFVYFVIFVYL
CR3ARF T. vivax      -----CLI----V█FCCFFYCCLYIFFVFCLFVICFYVCCLFVYLWIFVFCV
CR3ARF P. serpens +2 -----SKCIYKNIIFYI--FVWF--FVFSVIFIVFCAFFLFFIYYWFVLLFF
CR3ARF L. amazonensis K█NNMY█V█IYICLI----SLL█CFCLWVLYIFFVFYLLLCYFVWCFLFFFYMFVLCV
                        * : * : * * * : : * : : * : : : .

CR3ARF T. brucei      IWVYCCVYFFFLFYHLICCYHFYCYI█-----VFVIRLKKYA--NNFC-----
CR3ARF T. vivax      IWVCCCVYFFFLFYHLICCYRFCYI█-----VFVI-----
CR3ARF P. serpens +2 IIVFFSVWFL--FLLLFWFCFLFIFFCFSFWYGFIHFHIIICKFPLLKAFKNISL█V
CR3ARF L. amazonensis I█VFVVVCFFFFCYPLIWFCRLFYMYML-----VSDIKI█LL--FL█K-----
                        * * * * : * : : : : * :
```

CR4 ORF Alignment with *L. tarentolae* 13U translation (Published Sequence)

CR4ORF T. brucei	RFLF!
CR4ORF T. vivax	RFVF!
CR4ORF L. tarentolae 13U	KF---
CR4ORF L. amazonensis	RFVF!
	. *

```
CR4ORF T. brucei      FFMLFFVLGGFFVFVFFFCLCLFLFVVVVILLVWLLLLLVFYRFIYMRFLF-----
CR4ORF T. vivax       FFMLFFVLGGFFVFVFFFCLCLFFVVIIVLLVWLWLLLLLVFYRFIYMRFVF-----
CR4ORF L. tarentolae 12U LFMLFFVFCGGLLFVMFILFFISFFVLIVLVCWLLFIIFVRFFCMTFCILILIGLFINL
CR4ORF L. amazonensis LFMLFFVFCGGLLFVMFILFFVSFFVLIVLLFCWMLFIIFVRFICMRFVF-----
***** : * : * : * : * : * : * : * : * : * : *
```

CR4 ARF Alignment with *L. tarentolae* 13U translation (Published Sequence)

```
CR4ARF T. brucei      -----IYCYLCVFIIILF[F]FWLCIFFFFIWCVCVLCITVYGIFYCCFVFCFCLFVWVCFI
CR4ARF T. vivax       -----FFCLLCIFFFFIWCVCVLCIVCGIFCCCFCCFFCFLCVWVCFI
CR4ARF L. tarentolae 13U QIH[NTYMYNCKSVV-VFGFFMY-----YFVC---CIFYCFCCFCLFDLCVMVYFI
CR4ARF L. amazonensis -----DIKNIK[VI-FYYFYIFL---FTFVGWFLCIVVGIFWYFYYFYCYL[FTVYFI
                                   :      :      .:*      **      *.* :      *. * **

CR4ARF T. brucei      CFVIVVCFFLLFWVWVIFYWCFCDCIVYFFCDVIIILFIIFFFYFVLCFLYCCFYLVCLVFFC-
CR4ARF T. vivax       CFVIVVCFFFLFWVWVIFC[FIDCIVCFCDVIIILFIIFFCFVLCFLFCCFFLVCLVFFC-
CR4ARF L. tarentolae 13U YIAFVCLFVLVCYVVILLCCFDCIVFFLFTVIIIFFFLFIFWVFIFLLLFWFVFGFL[FFF
CR4ARF L. amazonensis CIALFVCLFYVCYVVILLCCFDCIVFFLFIDIILFFFFFFYGLCLFF[LFYLDLVFYFF
                                   :...  .. : :***:  :**** *:  ***:*:* : : : :  * :  ...

CR4ARF T. brucei      IFCCVLCYFLIYFLCCFLFWVFLFLFFFFVYVCFYLWLLLF[FGCCCYLCIG-----
CR4ARF T. vivax       IFCYVLCYFLICFLCCFLYWVFLFLFFFFVYVCFL[LLLFY[FGCCCYLCIG-----
CR4ARF L. tarentolae 13U [YVCVLIFYF-----EFLFMLFFVFCGFLFVFMFILFFISF
CR4ARF L. amazonensis F[FCVWFFILNFCLCYFLFFADFCLWCLFYFLCHFLF[LFYYFVGCYLYLYFV-----
                                   :  *  : :      :*  *: :  *  : : :

CR4ARF T. brucei      --LFICVFYFR[LWYWFYKMFF-----
CR4ARF T. vivax       --LFICVLCF-----
CR4ARF L. tarentolae 13U FVLIVLVFCWLLFIFV-FRFFCMTCILILIGLF[INL
CR4ARF L. amazonensis --LFVCVLCFK---VG-YEFIFI-----
                                   *:: *:: :
```

CR4 ARF Alignment with *L. tarentolae* 12U translation (Hypothetical Sequence)

```
CR4ARF T. brucei      -----IYCYLCVFIIILF[F]FWLCIFFFFIWCVCVLCITVYGIFYCCFVFCFCLFVWVCFI
CR4ARF T. vivax       -----FFCLLCIFFFFIWCVCVLCIVCGIFCCCFCCFFCFLCVWVCFI
CR4ARF L. tarentolae 12U QIH[NTYMYNCKSVV-VFGFFMY-----YFVC---CIFYCFCCFCLFDLCVMVYFI
CR4ARF L. amazonensis -----DIKNIK[VI-FYYFYIFL---FTFVGWFLCIVVGIFWYFYYFYCYL[FTVYFI
                                   :      :      .:*      **      *.* :      *. * **

CR4ARF T. brucei      CFVIVVCFFLLFWVWVIFYWCFCDCIVYFFCDVIIILFIIFFFYFVLCFLYCCFYLVCLVFFCI
CR4ARF T. vivax       CFVIVVCFFFLFWVWVIFC[FIDCIVCFCDVIIILFIIFFCFVLCFLFCCFFLVCLVFFCI
CR4ARF L. tarentolae 12U YIAFVCLFVLVCYVVILLCCFDCIVFFLFTVIIIFFFLFIFWVFIFLLLFWFVFGFL[FFF
CR4ARF L. amazonensis CIALFVCLFYVCYVVILLCCFDCIVFFLFIDIILFFFFFFYGLCLFF[LFYLDLVFYFF
                                   :...  .. : :***:  :**** *:  ***:*:* : : : :  * :  ...

CR4ARF T. brucei      FCCVL-CYFLIYFLCCFLFWVFLFLFFFFVYVCFYLWLLLF[FGCCCYLCIGLFICVF
CR4ARF T. vivax       FCYVL-CYFLICFLCCFLYWVFLFLFFFFVYVCFL[LLLFY[FGCCCYLCIGLFICVL
CR4ARF L. tarentolae 12U [MFVFWFFILNFCLCCFLFFVDFCCLWCLFYFLYHFLF[LCWCFCVGCYLYLCFDFVWRF
CR4ARF L. amazonensis F[FCVWFFILNFCLCYFLFFADFCLWCLFYFLCHFLF[LFYYFVGCYLYLYFVLFVCVL
                                   *  .  ::*  ** **::.. *  *: :* : : *  * :  .**  ** : : : :

CR4ARF T. brucei      YFR[LWYWFYKMFF
CR4ARF T. vivax       CF-----
CR4ARF L. tarentolae 12U VF[F]LDYFKI---
CR4ARF L. amazonensis CFKVGYEFIFI---
                                   *
```

C. ND3

ND3 ORF Alignment

```
ND3ORF T. brucei      --LLSLFLYLFLILWFIIILFIGFFLCFLCFLHFFDCFRCCLWFSCGLYDMNSRLVFYTL
ND3ORF T. vivax       --MLSIFLYLVLCCLWFIILLGGFFLCFLCFLHFFDCFRCCLWFSCGLYDMNSRLVFYTL
ND3ORF L. tarentolae  -----FGRREKVLHFFDCFRCCLWFSCGLYDMNSRFVYVSI
ND3ORF L. amazonensis MFFCSFYFNFVLIFCMLILSIGVLFYIFMFLHFFDCFRCCLWFSCGLYDMNSRLCYIFI
                        :          :*****: : :

ND3ORF T. brucei      DLCFVSCFFVLLNSIICVLLFVFVIVLFYFCYGFLFLWFLFFVVCIGFVWYFWDHVYLC
ND3ORF T. vivax       DLCFVSCFFVLLNSVICVLLFVFVLVLFYFCYGFLFLWFLFFVLLLGFWYFWDHVY--
ND3ORF L. tarentolae  DLCAVLLCFVMFYSIIGLILFLIVVVLVFMCKL-FFVWFCFVFL- ----FWSIV---
ND3ORF L. amazonensis DLCAILLCFVLFYSNFGLIIFVLVLLVFMCKL-FFVWFLLLFFM- ----LWLMYLI-
                        ****.  * **: * : ::*:*:*:*:*: * :*: * :... :*

ND3ORF T. brucei      GVILFCLWCFLLYTTYINKYIK-----
ND3ORF T. vivax       -----
ND3ORF L. tarentolae  FNIWFCV---LFVFIFLLILINSVFSYLILK
ND3ORF L. amazonensis FDSVYCL---FLFFFY-----
```

ND3 ARF Alignment

```
ND3ARF T. brucei      -SKNPRLFTLVCHYFYICF-W-YCGLLFYL VFFYVFYVFCYIFLIVFVVVCGFRVVC
ND3ARF T. vivax       -----CYQFFCIWF-C-VCGLLFYCVFFYVFYVFCYIFLIVFVVYDFHVVC
ND3ARF L. tarentolae  HSKNSSI-Y-----NLYRSIGISLGGGKKCCIFLIVFVVYDFRVVC
ND3ARF L. amazonensis TQKNSRFKFKICFFVHFILILCFVFCWFLVLYYFIFLCFCIFLIVFVVYDFHVVC
                        :          ***** .*:****

ND3ARF T. brucei      IWIHVWCFIHWIYVLLVVCFLYCIPLEVFCCFLWCCFIFVMVFCFCGFCFLLYVLDL
ND3ARF T. vivax       IWIHVWCFIHWIYVLLVVCFLCYWILLFVFCYLYLCWYCFIFVMVYFCGFCFLCCY-GL
ND3ARF L. tarentolae  TWIHDLFMFLIYVSQFYVLLCFIPLLVFYFLLWCCILCVNC-FLCGFVLYFCYFGV
ND3ARF L. amazonensis TWIHVYVIFLIYVSQFYCVLYCFIPILVWLFYLVCCYILCVNY-FLCDFYCYFLCYGW
                        ***  ::  ***  .  . *  * ::*  :  :  *  :*. *  .

ND3ARF T. brucei      CGIFGITYI--CVVIFYFVYDVFCIIHIILINI---
ND3ARF T. vivax       CGIFEIMYI-----
ND3ARF L. tarentolae  L--YLIFDSVYCLFLFFF-YFELIVYSLIYNNKI
ND3ARF L. amazonensis CIYLI LCTV-CFC-FFF-INLINSVK--LSLDNN--
                        :  *
```

D. ND7 5' Editing Domain

ND7 5' Editing Domain ORF Alignment

```

ND7ORF T. brucei      -----MLFLVVFLHLRYRFTFGPQHPAAHGVLCCLLYFCGEFIVYIDCIIGYLHR
ND7ORF T. vivax       -----MFIFVVVFLHLRYRFTFGPQHPAAHGVLCCLLYFCGEFIVYIDCIIGYLHR
ND7ORF L. tarentolae  ILFSRLHDNYILYLLIYVFLHLRYRFTFGPQHPAAHGVLCCLLYLSGEFITYIDVIIGYLHR
ND7ORF P. serpens     -----IIFFIFVFLHLRYRFTFGPQHPAAHGVLCCLLYFSGEYITYIDVIIGYLHR
                        :.:*****:.*:*:*

```

```

ND70RF T. brucei           GTEKLCE
ND70RF T. vivax            GTEKLCE
ND70RF L. tarentolae       GTEKLCE
ND70RF P. serpens          GTEKLCE
                          *****

```

ND7 5' Editing Domain ARF Alignment

ND7ARF T. brucei -IQKNMNTTWYSIIIVIFGSFFTFVSFYIWSTASRSTWCFMLFI VFLWWIYCLYWLYYRLFA
ND7ARF T. vivax -----VFCCSFFAFVFSFYIWPTASRSTWCMLFI VFLWWIYCLYWLYYRLFA
ND7ARF L. tarentolae LNFTIPTRILYFIFINCFITLVLYFRTPASSSPWRIMLFIISFWRIYNVYRCNYWVFT
ND7ARF P. serpens -----SNNFYFFYFCFFAFVSVFWASTSRSTWCVMLEVI FRVVNIYWRNRLLT

 :*:* ** * *: : *: * * :::

```

ND7ARF T. brucei          SWYRKVMWMI!
ND7ARF T. vivax          SRYRKVMWV!
ND7ARF L. tarentolae     SRYRKVMWV!
ND7ARF P. serpens       PWNKGKIVWI!
                                *:~::~*

```

E. ND9

ND9 ORF Alignment

```
ND9ORF T. brucei      MCIFLCLFRLCFCLILFYCLCCRWCFCVCFVDCSFLFFYCFVSFFLFYCMFLFFNLWFLFL
ND9ORF T. vivax       MGLLLCLFRLCFCLILFYCLCCRWCFCVCFVDCSFLFFYCFVSFFLFYCMFLFFNLWFLFL
ND9ORF L. tarentolae  MFLFLIMFRCVFVLLLFCLCCRWFVLCFVDCSFVFFYLFCFFLFFVMFLFFNLWFFLL
ND9ORF L. amazonensis MFLFLIMFRCVFVLCFFCLCCRWFVLCFVDCSFVFFYLFCFFLFFVMFLFFNLWFFLL
* :.* :.* * * **.****** *.*****.* **.* :. *****.*

```

```
ND9ORF T. brucei      YCCDLLLIDFCGFCFCRFMLLYILFCLFLCFLCFLVLCCLFVFFGLCFSFSCFCYAFLLL
ND9ORF T. vivax       YCCDLLLIDFCGFCFCRFMLLYILFCLFLCFLCFLVLCCLFVFFGLCFSFSCFCYAFLLL
ND9ORF L. tarentolae  YCLDLFCIDFCGFCFVRFILYVLFCLLLCFRVSEVFLICFFLFFGLVFSLFFCSYALCIF
ND9ORF L. amazonensis YCLDLFCIDFCGFCFVRVLLYVLFCLILCFRVSEVFLICFFLFFGLVFSLFFCSYALCIF
** **.* ***** **.*.*:*****.*.*.*** *.*:***** **.* :. :. :

```

```
ND9ORF T. brucei      ERECFDLFGFYFVGNDILHRLFVDWFFVGGFLLKCYPLFGLFVLLFCVLVEEIVCTFTML
ND9ORF T. vivax       ERECFDLFGFYFVGNDVLHRLFVDWFFVGGFLLKCYPLFGLFVLLFCVLNEMICTFTMV
ND9ORF L. tarentolae  EREVFDLFGFVFCGNDCLHRYVDWFFVGGFLLCKVYPLFGLFMLNFCMLCEDIVVIATSC
ND9ORF L. amazonensis ERECFDLFGFVFCGNDCLHRYVDWFFVGGFLLCKVYPLFGLFVLNFCMLVEEIVYATCC
*** ***** * *** **.*:***** * *****.* **.* :. :. : *

```

```
ND9ORF T. brucei      FLLLHTNFYLYHYFI-----
ND9ORF T. vivax       FLLLHTNFYLYHFFV-----
ND9ORF L. tarentolae  FVLCFSNFAT-----
ND9ORF L. amazonensis FVLVFPILHLFNLIYDNADLNIN
*: * . : :

```

ND9 ARF Alignment

```
ND9ARF T. brucei      NINLIFFINIILCVYFYVYVYVVFYFIVCVVDGVLFVLLIVVFCFFIVLLVFFCFIV
ND9ARF T. vivax       -----WVCCCVYFVCVFWFCFIVCVVGGVLFVLLIVVFCFFIVLLVFFCFIV
ND9ARF L. tarentolae  -RYIIYLLSIIICYFWLFCVVCLCCCYFFVCVVDGFFVLLIVVLFFFICLCVFFYFLW
ND9ARF L. amazonensis IINCLYNLIKLCFYFLCFVVYLYVYVFFVYVVGFEFFVLLIVVLFFFICLCVFFYFLW
. :.* : *.* **.* :.*****: *** * *** *:

```

```
ND9ARF T. brucei      CFYFLICDFCFYIVVICYWLIFVVVFVFLCCCFYFVCFCVFCVLCVVCVCLYFLDCVL
ND9ARF T. vivax       CFYFLIYGFCFCIVVIYCWLIFVVVFVFLCYIIFCFVCFCVFCVFLCVVCFCFLDCVL
ND9ARF L. tarentolae  CFYFLIYDFFYCIVFYFVIFAVFVLFVLEWYMFCFVYYYVFEVLYWVFVFCFLVWFL
ND9ARF L. amazonensis CFCFLIYDFFYCIVWICFVIFAVFVLFVLEWYMFCFVFCVFEVLYWVFVFFYFLVWFL
** *** .* : ** * **.***.*. * :.* **.* : ** **.* :.* : ** .*

```

```
ND9ARF T. brucei      VLVVFVMRFCCWNANVLICLVFILLVMIFYIVYLLIGFLLVFFCWSVIHYLVCLYCYFVC
ND9ARF T. vivax       VLVVFVMRFYC NVSVLICLVFILLVMMFYTVYLLIDFLLVFFCWSVIHYLVCLYCYFVC
ND9ARF L. tarentolae  VYFFVVMRYVFLNVKFLICLVFFVVMIVYIVFMLIDFLLVFFCCKFIHCLVCLCWIFVC
ND9ARF L. amazonensis VYFFVVMRYVFLSENVLICLVFFVVMIVYIVFMLIDFLLVFFYVKFIHCLVYLCWIFVC
* ...**.*: . ..*****:.*:***.* :.* **.* **.* **

```

```
ND9ARF T. brucei      WWRRYVRLQCYFCCCIPIFICIILFNITVVILNFSL----
ND9ARF T. vivax       LMRWYVRLQWFCCCIPIFICIFLF-----
ND9ARF L. tarentolae  YVKILLWLPLVVLCCVPILQYSFYFIFV--MNLSIKLIY
ND9ARF L. amazonensis WLKRLFLMLHVVLCLFFQFCIYLI FMIM---LTTLTKF-
*: * : : : *

```

F. RPS12

RPS12 ORF Alignment

```
RPS12ORF T. brucei -----MWFLYGCLRFVLFVLCYYMSPRLPSSGNRRVLYAVFYLYNFVWMLRCFF
RPS12ORF T. vivax -----MWFLYGCLRFVLFVLCYYMSPRLPSSGNRRVLYAVFYLYNFVWLLRCFF
RPS12ORF L. tarentolae -----MRVLFLYGLCVRFYFCLVLYLSPRLPSSGNRRCLYAICYMFNILWFFCVF-
RPS12ORF L. amazonensis TFLNLIYFVRVLYLYGLCVRFLEFLCLVLYLSPRLPSSGNRRCLYAISIMFNILWYFLVF-
RPS12ORF P. serpens -----MFFVRSYCLYGFCVRFCFVFLCIYVSPRLPSSGNRRVYVVCFNLYSFVIYCFLEFG
RPS12ORF Perkinsela -----MLFGFLVRYGFIEFFFFVSPRLPSSGNRFCYELDMRFFVFCYDFVLLG
                        *: * : . : ::***** : . :
```

```
RPS12ORF T. brucei CC-FIGLVMSLFIIEGGGF---VDLP-GVKYYTRIVS---
RPS12ORF T. vivax CCVFFGLHLSLFIIEGGGF---VDLP-GIKYYTRMFIN--
RPS12ORF L. tarentolae CCVCF-LNHLLFIVEGGGF---IDLPGVKYFSRFFLNA-
RPS12ORF L. amazonensis CCFVF-VIFQLFIVEGGGF---IDLPGVKYFSRFCNVS-
RPS12ORF P. serpens CCVICYSQSFYFLCEGGGF---VDLP-CIKLYVRVPIA-
RPS12ORF Perkinsela FSV---LLSSLVFYEGFGFWLFMDVPFGLYYFSRG-----
                        . .: ** ** :*: * : : *
```

RPS12 ARF Alignment

```
RPS12ARF T. brucei NTLLITNISK-----YILFFLRMWFCMVVVYVLFYLFYVIIWVRDCPVPTDVYCMP
RPS12ARF T. vivax -----CDFCMVVVCVLFCLFYVIIWVPDCPVVIDVCCMP
RPS12ARF L. tarentolae ---NTYRPI-----IIFILCVYYFCMVYVFVFYIFVWFYIIVHDYLVFVIDVVMQ
RPS12ARF L. amazonensis HK-YLFRPF-----IIFILFVYICMVYVFVFYFYVWFYIIVHDYQAPVIDVVMQ
RPS12ARF P. serpens ----LKPFIILSYFYLYLCFLFVVIVYMFVYVVFVLYFYVYMLVPVYPVQVIVVFMLE
RPS12ARF Perkinsela -----CCLVFWFVMVLSFFFLALVCPVLVIGFVMSW
                        : * * : . : . . * .
```

```
RPS12ARF T. brucei YF--IYIILFGCCVVFFVVL-LVILCHYLLRVVVLLIYPVLSIIHVLVRFYKICF--
RPS12ARF T. vivax CF--ICIILCGCCVVFFVVCFLVCICRYLLRV-VLLIYPVLSIIHVCLLI-----
RPS12ARF L. tarentolae YV--ICLIFYDFFVFFV-FVFWIIC-CLLKVVLVLLICQELSIHFVFFWMRKQIVIIKI
RPS12ARF L. amazonensis LV--LCLIFYDIFWFFAV-LFLWFFS-CLLKVVLVLLICQELSIFRVFVMCRKFNYLYFY
RPS12ARF P. serpens VL--ICIVLLFIVFYLVVVLVFLRVFIFVVRVVLLIYHVLSYMSVCQPKIIAS---
RPS12ARF Perkinsela IWGFFLFVMI LCCWVFPCCQVWFFMKVLV-----FGCLWMYRLDCIIFP
                        : ::: : . . :
```

```
RPS12ARF T. brucei ----
RPS12ARF T. vivax ----
RPS12ARF L. tarentolae ILFR
RPS12ARF L. amazonensis KN--
RPS12ARF P. serpens ----
RPS12ARF Perkinsela VV--
```

G. ND8 (Nondual-coding)

ND8 ORF Alignment

```
ND8ORF T. brucei      MFFFDLFFFFVCFYMCFVCCVTICLPIELTIVSLLVRGNHFLRFYWCGLERCIACRLCD
ND8ORF T. vivax       MFFFDLFFFFVCFYMCFVCCVTICLPIELTFCSSLTRGNHFLRFYWCGLERCIACRLCD
ND8ORF L. tarentolae  MFVYDFCFSSFVCFYMCFLCCVTILVLPLELTIVSICVRGNHFLRFYWCGLERCIACRLCD
ND8ORF L. amazonensis MFCYDFVFSFFVCFYMCFLCCVTILPCEITIVSICARGHHFLRFYWCGLERCIACRLCD
ND8ORF P. serpens     MFFDFFFCFFCFVYMCFCFCCVTIVVPCEVSLCSFLVRGTHFLRFYWCGLERCIACRMCD
** : ** * * * . ***** : * * : : : * : . ** ***** : **
```

```
ND8ORF T. brucei      LICPSLALDVRVGWSFGGHRFADWFTLSYRRCIYCGFCMHVCPTDAITHSLFVMCFCCLA
ND8ORF T. vivax       LICPSLALDVRVGWSFGGHRFADWFTLSYRRCIYCGFCMHVCPTDAITHSLFVMCFCCLA
ND8ORF L. tarentolae  FICPSLALDVRVRSLSGYSDFVFNISYRRCIYCGFCMHVCPTDAITHSCFLLFCCCIA
ND8ORF L. amazonensis FICPSLAIDVRCIRSLCGYRSDLFYISYRRCIYCGFCMHVCPTDAITHSCFLLFCCCIA
ND8ORF P. serpens     YICPSVAIDVRGCVSLIGHRAHLFFISYRRCIYCGFCMHVCPTDAITHSFVVLFSVLLS
***** : ** * : * : : . * : ***** : : : : :
```

```
ND8ORF T. brucei      MYLLAPKFLLFGCCFMLDFDYLCFV
ND8ORF T. vivax       MYLLAPKFLLFGCCFMLDFDYLCFV
ND8ORF L. tarentolae  MYLCAPKFVLFGCCFMLDFDYLCFV
ND8ORF L. amazonensis MYLCAPKFVLFGCCFMLDFDYLCFV
ND8ORF P. serpens     SYLVAPKFILFGCCFMVDFDLFCF
** ***** : ***** : ** : **
```

ND8 ARF Alignment

```
ND8RF2 T. brucei      NLIILSFGWLLFFYFFFI FVCFELIFCFFFLFVFICVLFVVLFFVYPLS PLLVYWFVVTI
ND8RF2 T. vivax       -----CFELIFCFFFLFVFICVLFVVLFPVYPLN HFVVCWPVVTI
ND8RF2 L. tarentolae  -----KHI CIRLKECLFMIFVFLFLFVFICVFYVVLWFYHWSWPLLVFVVTI
ND8RF2 L. amazonensis -----NIIRSILIKICFVMILFFFLFVFICVFYVVLWFYHVRPLLVFVVLVII
ND8RF2 P. serpens     -----I CVNVIKCFELIFFFVVFVLFICVFVVLPLLHFVKYHCVVFWFVVLII
* : : : : * : : : : ***** : ** . : : * ** *
```

```
ND8RF2 T. brucei      FCFVFIGVV SVVLLVVYVI FALV HWMFVLGGVLVVIVLRLDLHWVIVVVFIVVFVCMF
ND8RF2 T. vivax       FCFVIDVV NVVLLVVYVIWFVPV HMFVLGGVLVVIVLQIDLHWVTDVVFIVVFVCMF
ND8RF2 L. tarentolae  FCFVIDVV NVVLPVAVYVILYAQV LMFVVLEVYVVIGFPMCLILVIVVVFIVVFVCMF
ND8RF2 L. amazonensis FCFVIDVV NGVLPVAVYVILYALVWPLMFVVLEVYVVIVIPYFILVIVVVFIVVFVCMF
ND8RF2 P. serpens     FYVFIGVV NVVLLVVCV IIFVLVLP MFVVVLVWLVLVHLHICFLLVIDVVFIVVFVCMF
* ** . *** . ** . * * : . * : : : * : ** : : : * *****
```

```
ND8RF2 T. brucei      ARQMPLRIHCLLCVFVV PCIYWRPSFYCLVVVLCYLIFICVLCSYLFWVI-----YCV
ND8RF2 T. vivax       VQQMPLHIHCLLCVFVV PCIYWRPSFCYLVVVLCCLIFICVLC-----
ND8RF2 L. tarentolae  VQQTPLRIHVFCYFVVVLPICIYAHNLNLFYLVVVLCYLIFICVLFSLVYLEKY-IWLII
ND8RF2 L. amazonensis VPPMLLRIHVFCYFVVVLPICIYAHNLNLFYLVVVLCCLIFICVLFNL FEYFFYIFVVCV
ND8RF2 P. serpens     VQPMQLPIHLLFYLVCYVPIWLHPNLFCLVVVLWFLIYFCVFVSCIV FI ----LFVV
. * ** : . . * * : : : : ***** ** : ** :
```

```
ND8RF2 T. brucei      YDLKKFTVKLNFD
ND8RF2 T. vivax       -----
ND8RF2 L. tarentolae  INF-----
ND8RF2 L. amazonensis KNLVNV-----
ND8RF2 P. serpens     K K R T K Y --
```