

Supplementary material for

Nuclear Pore Complex Components in the Malaria Parasite *Plasmodium*

berghei

Jessica Kehrer, Claudia Kuss, Amparo Andres-Pons, Anna Reustle, Noa Dahan-Pasternak, Damien Devos, Mikhail Kudryashev, Martin Beck, Gunnar R. Mair and Friedrich Frischknecht

Contents:

- Supplementary Fig. S1
- Supplementary Fig. S2
- Supplementary Fig. S3
- Supplementary Fig. S4
- Supplementary Fig. S5
- Supplementary Fig. S6
- Supplementary Fig. S7
- Supplementary Fig. S8
- Supplementary Fig. S9
- Supplementary Fig. S10
- Supplementary Fig. S11
- Supplementary Fig. S12
- Supplementary Fig. S13
- Supplementary Fig. S14
- Supplementary Fig. S15
- Supplementary Fig. S16
- Supplementary Fig. S17

Plasmodium berghei NUP637

P. falciparum syntenic ortholog PF3D7_0609000

>PBANKA_0107600 | Plasmodium berghei ANKA | conserved Plasmodium protein, unknown function | protein | length=5440

MNLLIKTMNISEGNDKLVGGGKVDSEIRKENNKKQKKEHNTSFYFVEVYHKLKIYYEYSIN
EEERKNKCFLNKNPLTEITKLQDELVSFFDGNIELFSNPFYFLNVEEDNNINKNDLKNY
LGTKYPNLLNIMEDVIKCSLI INRVVYVCDLFDNDVYKNI EKNNIHTEISKTIYINEVLK
ECIEKENLIIIMIFLEILYI INRNEKIGITLDNDNKSMKHEYFSEKEDNYTDNELKNEEV
KSSSDNSSLTNNSLASLNI SKKKKKKNINNKINGDKFSYVVLVNL TENDFVENIFLEID
KIIKFLIKFDGICDKAEYAFNYKTHKINILFNLLDLLFLYFSKFHANINNISSIFKSI NH
LVSTKYFDFISTDYVVKELDNLLFLESQKKNKKIIKYIKNIKNSKLSIQVDDYKLDINI
SDSGLNKFDFISSEGLNSSFYKSLNFEPSTIFPKINQSLSEGITNFQGNVSNPISIGNT
NSIPNIIITEPENMSNINQKTQNLVDVDTLNKFSNNIVTNNQINYSNFHSEINNLRGKN
SGNTMINPQNKCGIKNNIFNKCLDCTGYWYSCNCSLNNFESTLTLSTQISLLLLILCLHP
NIDKYVYERKNKININSSSPDGINISVQSSGKENINELLELKKKIDFENFQTPTIWKY
NCMSTEKDYKHLMGKNNYIISENYISENYSYYLFFFSKI TEAKKKKKYKNI EEEKPVEEF
WNDKIFQDASNDGITGSCSYDSNIMLKFVIGLFINSKEDILSSIFDDNIFKVLNRFILK
KISSYTLIGNLIFHLFHSIFLSAFIKNGRITDVTWNTFIDSHIKRDFNLKRKKKIDNKDAN
IFKDTITSSNNNTASFIQDNAQTRQNFQDGINKINISQHRMGSDSLLNRTIQNPSDLYMH
QNSYRNMHNNIYIYKXKGEYIIDLAFLLCINYPRLITKYVCLKNI INRYHSRIME
FSEIDDTFYHHNLEQIHEIDNSGSRNGVNTIPTRNSKNEHDKYSQLKENYNNINEERM
IMKYRSLKLLAEEMKICIDFYSDIFVQILDFAVLCNNIYDLKIIENVVMCFKPLTA
NLNIFNLIKLLFNQFTCALNGKFNFSKITNKQHIYKTRGNTKSGVVLNGFNII DDET
SNDITNNGFIDSINIDSTNNGQVNVKNSRDLYRNVFCGRSVNIISSNQATTSSNNVLSFS
SYKDGNGKNGEYEEYENGNYSYDGNDRYDDNDDDDYCSNAYFTYINYPKLNLEKLF
EQNNYLKEILKNNVMYELEKNNAYLRNLIKENYLVEKHIINDKKIFKNYASNLDAYNKLL
SYENDYTASARDANKILQDIEQSSNKKLFLNFKGVTSLFCQAQVNPISYVGGQLQNNLDNN
SLQLQSSMNSFNLEKKNICINYLADNKVAENLLRENNILKYEYKYNMIDLFPKNKT
EENQLDNIKSIMLFEYNVNIIFSIEKRFKILKTNFDGSYNNKEAIKFFFFEDNKEAINFFK
DKKNEPPIFKDDSSSLNENIMNRAKVIKRELEDDSKSYIDARYLLNII FEKRFYPYNC
IKMLEKCLLFLSSVNNKLCYMPLNKNLLVDYIILEQYENASGNTASSTSMDIANMHLQ
ENNVNKLFPDDDEKNPFRIRNKKYNNMISLIDTQNYINTTQFKNVFIHNLNMHSNKKSK
DKHKIVKHXYTLHLLDMLYEIVKTKGNSSLDLLNLKCLSLLEILCSSFVKNSSSALSVLCI
ITNIFPDLFLDVVKQFKDQNDQLKKKLDLFDKEASIEBPKYKIEQLTNNLSVPKINLKV
IIIFMKCVNPLIYIIGIKRKFNMSLTKIWHFFNII EKIKEIDAKMKSKPMNSNSSTKTQ
NCAFFDIFYNFDDFDEIDNEQNSEDAENCQNRNNGQEFNILNMGKSNIGNKICIESD
QEKNGKNIKNNLEQMMNNSTIGNNIKEINSNDFEDNENNDIYEKCLKKLMIISDDKNV
CKHLIVLTNYIVQNI FYNL INNFILKKIRDQKYEAEFDL DENIKLEIFTNKQENS MFKRR
YNYIYINKSKNIESQIITMQNDVNINTSNIIYQIEQMLRQNNITFEPYINSQKMFDAKSSN
LGGNLSASIPNTSLNPFPTNFCIGTSTSQLGTSPIALGNSTMSLFPKQDNKGTNFSMFMN
SSNTQTNII LSGTSGNTLNNSTSSLSTSNFNNTSNFNLNPSQTNTTSHSIMDSQRTMFASTSR
LLSPSQETLGLNSSSNNTIRDASNSLNFAPISSNNNIFKPQTGLNMFSTNSGTNTSSI
FSTPQSNTTGLFCGSSSLGNNSNINDSRSSIIFCSNMQTNKNNPITSSFCGSMNLNQTSSST
NTGTNSLGLSGGFCGLGSRPGMNTSNLLLNMDSKNSQLGGTLGASTSTSLFCANRNSNT
LTNNTIFSSNTLNSLNIISQNNSTSNIFKNKPLGTSNTISNNNSLFSPTASTSLFCAN
NTTSQISSFCANTTSNMTQTTSLSLFNNSFCSQQNNITSSNNLGLSSNNLALTGQGMNQ
SSLFNNSTGFSSNLMGNKGNSSYLFCNETNQVMNNAALFSTGNIYSHLGNHNLQNI
LNREGINNINKLEYYEDLKIKTNVLNAEIKYKMKELKKNENLQKEKRIQLIKEAKLRY
MKSXSDINFKKELIILIVLVFMYFKLILKGLPLSTKNNKLNVTMNERFDYLCDFRPTM
YLFQILGESNRFMNTFFNIIFLDNII LHEENKNEKII SKYKNSDYITIIYKKTFLFRT
NIEEEDDENKISDQPKDFKNIKNNLFTTLTPETYKKKMEYIICSYSKESYNNILPHKIK
SLGLSILKILFERDVLFIHMYNQWNEKILYEKINKFHLNFNSINNNETSMNMMNLNVN
DSISESLNNNNNIITSNFNDINSRGGEQIKEGLEAGASTGGDGKEINFTESLCTP
HNFLFKNININSSTYLLILLKNFRFTSEINKIIIFILQIIIRDTKTTINILKRDAESF
NYLKYALKNIFIYNLNQKFNNSYIISNREIKMQNILNFIIDIPLLYLLKKNKIKTKDS
VISTRKDRREFYERNKNSISGNEVLNTSNWLSYINKHNSGNMEMDEIDLIRVEKNKLES
ENSNMVLKTKKDNLYLFEIEEGNSANSYSTYNSYSSDISDNEYKINNLLKNKGGHILGN
DNYSLRNEMILKYKFNKLNKSKRNKSKYVYDYLSDIDLIDTISKSKICIFYDNKEKDEEN
DNITFYNLSSENINDNYSKNDYNEIKSINYNVTFPCSSYNFLP INQFYTEKYDFIESEL
VYISDLKLYFIKSKFRHSYFTKLLNIEKVSATSENDKYKTNTSENINDDINYEIKENNL
SFSNLRSGFDDSSLTSNKIIINENDI IKYQNCBEKDIKFRTNKLESENEQNDVKNIV
EDEIKVDFYFNKLNKNLNFNPFNIFYDNKVDKMKYIEYIKKIAEKEQKENDGKYKNILN
PINNLIKTKQNEIEKQLSIELQKKNIIYINNEYCEWIELSHLLNVDIDYNEIRGGLNKNNS
VLNYSRMPITKLIIFYEVI TRKFLFLNNSDSLIESCVLSLLGLSFTPSKQMDKKNIIT
NIDESLEENCFNLNLIKNVIVNFVNFNYEHVSETIKVGYNNNIAEIGEQRNDRKRLITN
EKLNEFEFKDDLFIHQDEEIIILGNYINKADNENNFNNENLYEMGSGLNNDNSKIVFL
NKKRKYKKNIFLNNNSNLKQYNTFEKNIYFIKSLNIIYMLSKNKKIKEYIINLINTMW
HNKYNVFYYITQNVNIEKLDTEKIIFFKISLHILNFIPIILDHILINIDKHVEKFINLS
FPNDKEDNNSIQIFESINNAATNNELGNKNNNIGPQKYNCFHI INKENFNEKKNLFEFL
NPIFTFDNAQNF INHYINIVKNYNDNYLKHYYHYLYFNYNFQKIYNSNLEHIHYIK
SNKLCNDINEKCFNTKSRINEYKKNVHLLYTNI INVYVNFCKILNSYKVKSHVNVNYTF
NDGKYVENYENNNIDENVKNKILKPINIENENEFNFENLNEIIGNKSYLLKFLDDYYINE
NISKEEKKHEIKENVYIYNSINYLNIYKNWLLVYKTYIEKITYLIDLILKKNNTIKKK
IYYLKNFYFIVNNIQRHLKSVTSLIKSSYYINFNIQITPILFIVYLFITIFFLSHNNKFA
FKYLSKINKGFHNLFPKGNMLDFKNTEINIDGEDDTNNDDDSGSEILKHNNKNEINKKIK

FYDEINGSKQKESDAHLCDNDEFKIDHLKYNNKLMSSVFSKQILTKEFQNNIINIPKA
YNQDSSNIINNFDDKTRTRYNNNTSSSIHDDGIDTISNLQKGELYKSVFFINNNYNNNN
NSNNVGSKNLDDNKLYYDIHSEYYKLIILKSNEF**FG**DLQLLIILITKKHPVLNKYTIYI
YECLYTLNIIYHYSYCFAEKEYNSKNNKMFISLIKRIDHNILNEFISTLFTDSYKNFFT
NNNSTIPIYFYNNIYKFIDYNTIEIFMIKNREELTDNASILESKNRVSMHNLIDSDSDIG
NYFLNKEQMYMKEKKKVHFADSEPTKIVNEKVDVKMSEEDNGVINLNFKNFENKIHLN
SRSLNLYEEIYCSNFLYNKNDLNLLSINLLIFLLTRIGIYEIDIKIDENNVKNLIKGNLF
IFNKNSGNENENKILHTCFLLSALCDTSYINEFISSEDLFSLFLKSNIFFNNLYEYKFNYL
SNSISFLTPTNFLENKNIYIPLPMLIISYITVILKVIDKKGLQLFNIIGKWITTNITIF
TNHLIVNNYLMKSYEIVHEINNFTTYNHNLVYKCLCYASTCNFTREPSNNTSNSSNKPLN
FVHESKTDVKTNKENDINDNTKYTENLINHSNVKLVNLDIISSTYLLRLYKNYLLYIH
QNYLNLKCLKIVENINKKNNKSKKIKQNLTYMEKKLINKKGNIGEIDSSKNINRYINKF
NDENNIIPNKIPIQKENYFRNHFPNYSEYESNEHNSDYDESESDNEEEEEEEEEEEEE
KNIREMDSIKNNFSKLINSVYNEKNSQFIIGNYNNKNVLSNMNTI**FG**NSNSNNGFIENN
YFYIDKGLCFEFNEISLNICNIKEHTYILQYVKDLLNLSIFVIYDISWKDENELVTTD
IDSHKETRGNHVNVKSDNNIDENNYNFDTKEMRRKNIRTHRILQNIIPKYSIKFHCINL
CLDIEYDIGIYDEYVQYKTKYEEYIKIVLKVFMNTSFYYINI IKYENKHLKSVSVAM
KCLTNIIFLLTNLIRKEKTTDEDKHHALINKFLYDKNGFKKNAHEITTELFNCEDNIF
TDKCSLDHVKKFLENTEIDIHVLKILTCIIYFTFAMLNNQKVNKINIENINLNHINKI
NLKKIANTYLENNNILKVFTFILQRSTQIYYLLYNYAMQS

Figure S1. FG repeat *Plasmodium berghei* protein (sequence from release 28 of plasmodb.org). FG di-amino-acids are highlighted in blue.

Plasmodium berghei NUP221

P. falciparum syntenic ortholog PF3D7_0905100

>PBANKA_0416300 | Plasmodium berghei ANKA | nucleoporin NUP100/NSP100, putative (NUP100) | protein | length=2004
MYMQNSSSNVFGSYNNQGNLNKGLFGNAGNPSSLTNTNTSNLFGGSKPQQTNMMVNVK
NLFSGMTTTSGLDGNKSIYDSMNSQSNLGNKNLFGASGVNNSMQSGGGTVRNNLFMNTN
NQNNLNMKNLFDGSGNLNNAQGTNLSNKGLFGGLQNNQSSSGGNLFGNLSSQTNPGGMY
SGLSPGANQNKGTGLFGTSGNTNQVSSSTSLFGGTSTMGQNKLGVLFGNLQSTDPNAQNS
GNTTNSLFGGMSGNQLKPMASAGSNIFGGMSNTTTNTGTSNLFGNNNTASGMNQSKPGVF
FGNLQGSNQGTGSSVFGNSSNNTSEMSQNKSGVFFGNLQGSNQGTGNSMFGGMSGGMN
QNKPGLLGNLQSPSQGTTGTTNSSI FGSSVSGMNQNRGTTLFGSMSNTGASTTNTTNN
LFGGTSTMGNKLGAFGNLQSTDPNAQNSGNTTNSLFGGMPGNQLKPMASAGSNLFGGMS
TNTTNTGTSNLFGNNNTASGMNQSKPSSNI FGSLSSTTQPTGATSSSSSTGNI FGSTQ
TANQGMGSNVFGSSGMNQKFGNI FGGTSAHQSSSTFGSTGHGTSGI SPTSALGSG
TNTSTGLGGGLGISGGIGGTGTNTGMNINLNTTGTGTTGISTISGSSNLFGGLSSTS
NTNTMSLGDKNLFGSKPGFSLGTATNISGQSSTNLGLGTMNLVGLLSPSLNTTSSGLG
TNISSSTGTSNLLTGLNGGIGSNSTLGVGATTSPITIGISSSSSSNILGSNNALNNTSL
SGNLLSNGSGMNQGTGISFLGDKNLIGESRGNMQNNLLTVDGKDKSNDLLTNKLNINKESS
ENLTKTKFDDDLVLSKKKENIFNKAYDDNDISNDEINLKNQDGTFFYNYINLI INDNI
NDIQKTTFLNDHDSLMWDFKSNIFKNFVDYLVNFKQKNQKVIKSNVLDLDQHEFQI
LIWLYNINSYKIDFIPFKSFYLLSNTNLNYSKMFISNSDKSILPRNFVEINDHYMNY
KSYFLKSQINKHHDNYSTLKDGI GNSNGVSGNSISIGMNSVLDIDS IYKQILKDIL
NSLLMNLSITKENIKKKEEYENGDPNNKNSANCGSYIYENLLINYLFGTLQHLHDK
FYKLEISNYLSKDLNQIDEYFVDTKSNEIFSYCYDNFYKNEMDKENCTIHFFFYFVNIMF
RCGNYIGLIQI IKNDEFIKNLNIDSYLNDFIILLIRILLIYNQNNESIFENMKIDIDC
SDIFKKKIHMNSLINFFHSILYLCSSNIYAYDLLCILFSDIFYKKGMYTNRKKEIEMKN
IFYVYQKKRNSNYVDNNSDRNSSNEDNYDSGGDNNNIERGSNSINNNKGGKGGSERKG
FFLDVFTNMLHKKSKKDEDEPEFETMEINEEIEKLNRENMTSNTMELGITNKYSYNF
EYCNIEYSIWIELSLFLSKNFYLYGSKFQENFDIILDTNLSYNYDDNNAFNLDTKHRRD
MINSKIEELFISISSCIINENREYFVCSKLRADDVINNFIVDGKISEKVSQVFKVLR
TNFLLYIKLFFYLLVGNIVYTVGFLSCISNNIQRILLVLTIFLHKNNVFNENSKLNKI
KTLGFLKLDNNTILYNSKINDNVPAGSDNMNLLLSMNNTDIPFDYFLLRNNINIL
LKASYLLNLKTNICIKLLKSIVQENQGILLNESIIGHINNDGNIFYGKLHDFLFLFKNKV
KSRNVKCFFLMHYVLYKKIYNSNILRKIRKIDIDNEYHYKCRNFQVQKNINTLNKIS
LDNSIINVHSSILYKISQFYSILAYFANQRKHYIVSFICYIILNDEQSAINSLRLIYNDE
LIYYFNNKEKEYIRKCAFRFYHLAKNMWPSNVKLESIEQKSYLILSILFMKKMFEEA
FAIFSLALIPDNMLEKLSIADYYNIDLSSNFLVTLRELCKKNIKISELVDQSTIRSVIKF
LLPIKDKLDAQVAESINYLGSLSF

Figure S2. FG repeat *Plasmodium berghei* protein (sequence from release 28 of plasmodb.org). FG di-amino-acids are highlighted in blue. Two predicted transmembrane domains are displayed in red font. An internal repeat extending over the FG region is underlined.

***Plasmodium berghei* NUP138**

***P. falciparum* syntenic ortholog PF3D7_0903500**

>PBANKA_0417900 | *Plasmodium berghei* ANKA | conserved *Plasmodium* protein, unknown function | protein | length=1237

MNNNGGNANNPNMWGMPNNNLGNNSF**FG**TNMSQNLNENNNNSLNNQTNMTGNNM**FGA**
SVNNQNNNIFENNMNSQNNVVKSDNSI**FG**TPSNDLNKSNNTSL**FG**SLSSNTSVNNNTTNA
FNANLNFNAAKKDVPYPGFTRSLGNSPNNSSNIFKQSTLGSSIALGSQLNEDRGSESGGLF
SKEQLEAAKQIFANSSSGNLSFNNKTNNKLTFSGGFSSSSSAFSKNNINPFQSMAMQA
TNQNDKSSLMNNNMNKL**FG**DTNNNNINGGGKIN**FG**MPTVNNLTSPNKLNEIGPGGFSQ
AIKSFNPNTNSLFSSTNTGSTASTSFNNNDSSQKSLFSN**FG**GGFNFAKTNTNDFKKN
TTENATNNNTNTSFSFLPSFNGSKNNDNSFSLNSFNATNNEKSSGSLFSFNNTNTAS
PDFSLFSNNNINKTTTPDNSINSSINKDAENKAIEKSNQNI GESKDSGITSVVSEFKAN
DASGTDKSI FSKETEAEKDEKAGDKNEQNDNDKEKDETKETNVTLEIKKDENDSTKNHD
EKLIDHKS ESKENQNNENEKDDKNDKKYETSDNAKDNKDGTTNNDKMSL**FG**KS LTFNSSF
FKTSTANLTDKKS NLESSSSLNNDTENKKEKTGFF**FG**NDDKNENTNTSTESKWD**FG**SK**FG**N
NNLNFNKEESESFSFKETSTTDKKN ESTNLKDNKKEEDTKPSVSTESKDKNKEATATPKK
SIFN**FG**IKSNFLNQSSKDSKDDNINESNKEKDNEKPTKGAETDETADGTDHNTTGVVY
KLDSDVGDKKKIWKLSFTKKKQDSTDAQKELKENELKNNLLNKGPSLQPSDITLGKG
SL**FG**SFTKDNEKKNEDTKNTKSDHFSNINS**FG**KNNTSISFNQSSL**FG**KPDEKGNLTFSSS
NNTFQFFNSSQNDKKKTEFQTKTQNIPEIKNNQNADEDDVNNVSNLINFISLEDRKKNVY
INNNAQCDDENHEKTYISSNLKSNNNNETSHRSQTANFQLKETKGTLRMDNDDYADFND
IDFINGQQNLEINMEHQRL ENDRKSVENNIKTNECFLKKNLDQEMAVDVINNLSFVKK
KINFMYCSNEILDIYNKVSHYEKMYALISEDQIKIEKKQESLEKRLRLIQSEQCDMLSL
LNELDNENSLTFLKVLNQNLNKDSDSINNKNLYLDIDKFEKLADK IENLEELIDS IHNTS
KHDIVNDLVNKC YTNEINCEQIEKQLNGYSHEL RNMK

Figure S3. FG repeat *Plasmodium berghei* protein (sequence from release 28 of plasmodb.org). FG di-amino-acids are highlighted in blue. An internal repeat at the C-terminus is underlined

***Plasmodium berghei* NUP205**

***P. falciparum* syntenic ortholog PF3D7_1364200**

>PBANKA_1140100 | *Plasmodium berghei* ANKA | conserved *Plasmodium* protein, unknown function | protein | length=1828
MDVEDNSMDFYSMSNYEDIENNNKTKLNDKEQGGIISNKDRPISTNGITDLSCPLEKSCES
FVTVSDAYLSSSSLYKSNFYDDYEDNDAYNINNEVSTLRYSFSSKNSSSDISNDSFSDQY
ETPYINKNNNINVKNCFNKNSHITYQDENNSNDNKICEGKSLNIENNLKNIKISMLKKT
QDISAFSKRENKTNESDDIYSNRGSNDLDNSTHIENAKNGNKKTDVSEENDSNMGGKQKQ
NEHSKSDQEDKNSKSENSDDLYSVISGPNYDEIFLRTQKDDYMNNDIKLKRNI FDE
NQNSGHTKNIENEDDKNFDDTKTKYTTSTNFIKIKFSDQLESIKKKKNDGNVFRSYDS
SQIEGGEDEKTIKEKKEKKKQKKNLNDKYNVQKDYEFKFLVRSTINMCHPKIKNIDENED
ILFNMDKISKIIVIENNNKDAYKTDSENDDCDTNNSLYDRVDDILDDRKYIREKRDKY
FSEVKNGLNIGTNNNGDGIRIKNPSDKNGDHLNSVIGTIKNDAINNCLKTENNKNEEQ
YIKTNMANSVYYLNHDMIEINENVLKLYSGMCNYVERNCTKNIVRIVPHLNNIYDNKKL
CYIKDSPSFSEIYKVIPEWKDSFELLNKAQTDTEDMKNILKRIENIKNDLNLKDLVD
DKKEFEKIKFETIQHESMI SNFEKKNKTYIKGVKLDNMSFKLKKLHKINTNNLNFNIT
ENVNRLKLYVDNAMTDYSQFLNEEKSIEILRFQQLDNTFNILNTDDEYSFLMNLKDEL
KEMICEKKEQLQNIIRRTMLYKAFDTTSFKDQITNTLNKQTLINQIKYNLQCNVLLFD
MYKKFSKKNKIKIIPENYVKDYTNEVIKKSIIKTFNDTKKNVSKYFNCVHFTRNRPIT
HPFFIHFCLEKFLYHFIFNVLPLSEENQKFMQYCSIPGFNSFLVFDKVLKHLKKEEQ
FPSTDLTYDFSSAYDRIGKTDYNDTSLINQYMQWNELNPLYLLALRNASEFCFPSSNDI
NDNNDYTSYTLMDSKKNVFDLINHIPDSIGMSNAATAAAAIASSSGICNNINPSLGF
SSFDNMKNLSTFSLNNQMSNNTNTMFNITGVGTNISGLNSGNSFASNNLNFANNSGS
NLFGSTFNNSNTNTMRNNSYSGTTNNLLSGGNSQSMFNTNFPKSSSTSIFGNASSTS
SFMGSGINNTPNVSVGGSTSLFGGSANLGNMMLTNNRNPLENNSAVGIFGSSNYSTSAN
NNTLTVSTPFGTTNNDLLSGGNSQSMFNTNFPKSFASSIFGNTSSTSNFMGSGINNTPN
VSVGGSTSIFGGSTNLGNILNTNNSNINLNIPNTPTMGNQNSLLINNNATGNNNSGL
FNKSTLGSSSLFNSSMNNANSQFNRTTLNTASSLAFANSITPNNNILLGSNVGMNNTS
SIFSGLNNKTAMLSNNMFGNNLNNTNSGLNTGLFSLSSDTNKIGNNNNSNMSNSLF
TNSSGMSRINTNNNNNNNSLFSNVNNGASSNKNMVNTS SFNRGINMGSNNTISSYDLG
GNTNFGSFMGGNTSGNNSTFSGNNYGGLFGGNAQQGASSIFNNNNTSMFSNTRDTN
IFNNSNNSPSKNMFNNRMSYNSTNITSNNSSTFGNKTGFGSNNLNQFSTTLNNSNTS
GNNSNTSGNNSMFLSNNNMNRKSSFPPLSNNNYMSSSNLNNNSLFQKNTFSSTQNNRD
SLSGFNTLNNSSSNFGNKFQQNVGNTFGSTFGNTSNNMFSNNSSSFGANTNTGLGA
NNNSVFQQTNSNFQYNFKPQNNSSLSR

Figure S4. FG repeat *Plasmodium berghei* protein (sequence from release 28 of plasmodb.org). FG di-amino-acids are highlighted in blue. Two internal repeats—one at the N-terminus, one closer to the C-terminus—are underlined.

Plasmodium berghei NUP313

P. falciparum syntenic ortholog PF3D7_1446500

>PBANKA_1310200 | Plasmodium berghei ANKA | conserved Plasmodium protein, unknown function | protein | length=2715
MINNKRKRAYIQNERHENSNESESIVKLENEKGHDKDIIYDQIYSIPKKNKPRCLVLSNS
KDNLIANFTNLLKIDNYGVEFKTKWYVHLFRICLFSERLYGPTQTGSGKNKNEKGNEEN
INKLILEYVNDVDELLKNVQKILLRKIDCENLYYQKILVIVCLLFDYFQNKKYHNIKTVKHL
EGEEDTKNSIENFETKNGSDIHINDNSKKIKSKKRYLETTKKQIYDDSGDDTTNKDKL
SNYSENNKLLGSELKLNNEKNKFKPIDIINNQNYYIIEFKNILEYNEKKKFNLEIFYHI
TNLIHKEYYSQYAHYNIPIWFLNYINNLEIKKSINNINSKLEMFINLHPKIFNLSHA
KKRKRKNYQHNNI IHNDINPTNPNNSHYFEGGIVDYPYIKNRFKQLGGGFSSFTNPINS
STPSNI INNANNNNNNNNNF FGVNNNSTNTNFNGFKSTNTDTSTIGNTFGSTFKGFGQSN
ASI FGNNNTISNTVTKNIETGSTPFTSTNTNTSL FGNSINSTSTNTNTNSIPKFPPLF
PPNSSGSTNNMNNINMNSFNSQLNKPLFANITPIGNMMPNLQKNVNL FGSINNNAPG
PNNNSINTTLDNNNNNNNNASIF FGFDKNNQNTSIF FGNPSLNNTGTNTNFISASNNNATN
NIFSNKDSQKLLFQPTQIENLENNNNNSNNNNLNTTTSNGNNTAL FGQKASGTGFN
STINGNINKDLLLNKAGFENDMLLQKTNMFSTKPADSINFMNNNNMINTNTNIISS
NNNAVSSFNVTGNSNDILGGVSSNLNGQNTNPNLNFNTDNGSKELFKDLKITNEPTS
KFTVKKPLFSRRPVKSNLLISNSITSENNIFSSTYSNKDSLLNNQSKPVKVI FGQNVNT
STPI TNENTTAI FGGSSNTTNNPNNSTYTNTMQIPSI STKEIKPITSTMCKGNI FNQ
ENNKNI FSNNTKLPFNASNNNTT FGSTNCNI FGCKSEISNIFSTSANNTISISDTKTNP
DANANTNVNDNLI GTNTNSV FGNTSIF FGQNEINNDQNNFQGINFPVKLLTLEERKKK
TAKNMLSLSFSTNTNNSIQTQNSITLNDTSLINKNNNDTKKTNI FGQFKSTQSDINTKT
DTSIF FGGFNKSNI FDI GKSQENFEKKFQ FGNNKPSGNNTNI IRSIPNDTL FGSKNEDKQ
KTS DNTPKT FGTNLF FGFNFTNKEKDESKNEVKNVKNIEIKNEVNDKQVKNVNDKQV
KNEVNDKQVKNVNDKQVKNVNDKQVKNVNDKQVKNVNDKQVKNVNDKQVKNVNDKQV
NENKIDITKEPTIF FGCSILENQIDEANKSASKSFI FGLGIKEKDEKKTSLTSDINK
DDNALTTPK FGFTNNITTVSNLNNIFNLGNKKPDDNDNNGNAKSSIFIERNEKEDGN
GAGLFASTTNNKQKDNANLFSKAEKDTNISGINRFSSTNLVGNKNKESDILTINYDN
IISRKRKRDVLDYKDEIKNMNFSNLNENNDNNLRNNETNHTNLSQINNNMFLNEENK
NNEGLFLKSKIGNTLNHTKHLNTIKFDNIPSKDKLSASFSENNNTNNDADQTDNRNIF
NINMLISDQITKENELTNTITSTLANNESLNKSQLNFLSKNTGKNNNQYEADELGKK
KTENQNNHHYQQNESNFICIPDGNKSKLAEKQNDFYEKMRINEYLMNNNYNSDENKNN
NLEKKNITVIDKSMNRKISYNKHSFDSIDVDLDRDVIDINELRQELFLCALNFHLHYW
AEKVLIFYPHNEKINKYCTSIKIKNKFDETYDFNLFDDIHSISRKLLKRLDKNSCIEYES
VLFLSADNIQILKNAKLSLFDVFNIIYHFWFKKNSDLKKNFQNFLYTNKIYPSYLNYYMS
LYKNSKTSNNKFTDLKGNRTKIVKGLTPQNKLTNEETQKHSSPALSFYEKFSQFSYSKQ
NSKSDLHIPNDESSEKNSSTSEHNNSVIDNEQSESYSTNDENTDENSENSTSESQENSQ
IYRNFENSDLNENRIKKNKKTITENNLAEIFNSDSSDASNVLNNYDDEPFNMINEKY
IYLEKEEKRRKKLFLPSHIKIKKLSPECILELELILKDNVNLVKKYKLFKKEKYEYLYV
NLIAMLKHYNYFSTYKSEDINVENGNNNFANIIPSNVSKDKMDKNHRQYINNDKGSNLK
RISNLETNCTNIKNDKINILKIKCDKNLLYYINLQKNI AHLNLYKEIEI ACTYLNHIK
NNLFVIFQIPILLFNVLYDKLIKCYNDFSIMDNYFNNSLLYKDFRYEYILKTKIAYMLLN
KFNEQDDIENIKFAIYFDNRRVKAHKNNIYNDYIFHKNLINSYIKVHENIVPSEWIEAK
YAKNGELSDINDSEYLDYYSFFFKDISKYIKLSIIDKIINSNIFSLRKENYTI CQSSDR
ENTNKKYSNNGTNSGNNSNFPNNNTNENNETSQNDIEKKKKFACYFMYMVRRS IMDDV
VNIYSSKGEYFYKLLATLVNNNFTKDYEKSHKHFHMF IQVANINIFMKNVFFEYINFF
KKKELYNI EKIKNEDEIMDNLVNVKVNGRNFFVINTLAEIATKIEKNI VKDVNIHILI
NLRNI IDTFYLFKYVFNRRFSLSLIENLYKNVLFNFDINSFFSDDVNFYKPEQIKDL
YTHIRHKFTSQNVIN

Figure S5. FG repeat *Plasmodium berghei* protein (sequence from release 28 of plasmodb.org). FG di-amino-acids are highlighted in blue; two internal repeats positioned at the C-terminus are underlined.

Toxoplasma gondii NUP145

>TGGT1_203780 | *Toxoplasma gondii* GT1 | hypothetical protein | protein | length=1475
MAFPSPFFSGAVPSAAGGPERRPNTQRPIHRADPYAASLLGRLASKLASSVSSFFSSAPA
RKSPSVGVSTSSSHAVRLAPVVRGSAPSPSLWGASELSADGEATGREVTVVRQPTSSRQSL
AEVLNPSLWPRDNSFHSSALPRQGEPACLHPGSDPAAQSHDVLTLSTSGAVHYLQSSSP
PDMSDRAGTQWRRLDAVPASTSVSPNGDMQMGWQSQVADQTTFFVPRAAPLYGDGVRGLQV
HFPAAGDPLSNLEASASRLTQQWLLQQGATGPPVMEGTVIPTPAHTGYPPQSIERPVT
VSSRAFSGYAIPEEPLAQPALRPTFNGHTRVAPAASHSLFPTEERVASSLLAAGPIRRR
GSEARATPASGPPVPLRLHPRRDRSAAAGSLRQEHLTVDQLRQLYLRCPQRWRVRL
RELFRQEVESAGAVVVDALSGAPQDRATGHKSGQGSQVDTEGVVSVQDDLRRMDDSAG
RGTSSGPDTSAGSASCDKSEKESAASNGPQGSASGVPEGMKAGEDRVTNQGSTEGTSPVK
GGLFGSLGQMDTKTSFATERRGIILETSKELPATGFCEVDAAGTTPQSSLFAGDRRLGD
SGAQLFGGTAARAEDGLSAGGVTTNSKEIAREPPAPFSFPAVGGGAGGTAPLFSSGKAA
DQDTASQGRQDSQMGIKASGLGAASVSDQAYATLHASTLLGASGPAPGSIHQSSSLSP
ASTTMAGDKSEKQGSASGGSGLFGATKADSGGKGDINASVSQSTAETPALSFKPGAVSLFL
GGSSSLTGGSNLFGKTPDQDAVSNDSQTLPTFCRPAASSLVSGSASETVTRGASLFGGA
TFSAVSQPPSTNSKLSSRQREPAEKPEESSSEEGITVSDGAEVGETASSSSNTASDRDAK
PQDRSAPGGQVTQSEASAVPWQNVGKACLQVEDDGFAPDADDADADEKETGSAPGGGP
AASRHTFATPVSGGSLFSSSGTESKPATTPSSLFGTYSSMASTAASSSSGAGGGLFGGV
TTSNLFGKSSSSTGILGTGTQSQTVSSSAPPFLVFCGGGPAKTAGEAGKQTSGLSVT
STASGPGSGLFGNSVSTAAPLAGGTAAAASSAEKDKPSTVSLFCFKDSASATDSRVGGT
APAPKTSGLCLFVFGSTVGSQTACASSGGSVKRGREGCDIGAPASKSVFGTTSQAST
TGGLFGVSSAPSASLFGGNTSSGTGGSSSNLFFGLSTNGANDAATKEAGPAALSSKG
ASPFGTQSSTPVFCGGTTATGSSSSLSVFCASKADGNASNPFQFKVPGGSLGAGTGS
LFCNGGTTSSGTSLFGASLGAPSGGSGSLFGAAAGGTATGGSESQRPGAAPVVSSPFG
SQSASGGSSGGASVFCSAVASRPLTFCATGIGGNQSTGLGAGSAAGSTGFCATGA
GGLGQSGEDGNSLFCPQTGGPVIRRPRLTIKRTTK

Figure S6. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). 29 FG di-amino-acids are highlighted in yellow.

Toxoplasma gondii NUP302

>TGGT1_259640 | Toxoplasma gondii GT1 | nucleoporin autopeptidase | protein | length=2894

MFSSNTSSSLTGGGSLFGSSGGGGGLFGGAGTQQQQGSLFGGTTGGFMSQPQQQTQGRFLF
GNAGTTGMSLFGQQQTTQPQSGGLFGSSTTTNSGLFGSAPQQQTGGLFGSSGLSQPQQTP
SGGGLFGSTNTSTLGSSTITGGGLFGQQPQQTGGLFGSSTGTGTGAATGTGALGQQQSGS
LFGSSGFGTNTGGFGTSTLTAGQQTSGGLFGSQPQGGGLFGSNTSSFGAAGTSGTSAFG
TSTFGGGAATGTVLGGAGSLQQQPPFQPHKTEEGLLMSICFENLAEVSDDERWRFYQQRG
GGAMGASNTPGTGLFGQQTTQPPSGGLFGNSAGTTTSGGLFGSTPASTTGTLLGSTPQT
NQQQSGGLFGTSTTGGFGGGLFGSSATTQPSGQQQTGGLFGNTGGSTTTPGGGLFGSSTL
GNTTGTGLFGSTTTPQQTQSGGLFGQQQTAGSGGLFGSSSLTGTATNTGGGLFGSSTTPS
TGLFGSTTQQCPGATTTTGGGLFGSTTTSATTTGQATTGTGLFGGLSSGTGATQGGGL
FGSGGTATTGATGGGLFGNAQQSNLGGGLFGSSGGLKLTTPGTTTTPGATGTTGGLF
GNTTTTGASGTSLFGSSSLTSGGGLFGSSAAKPGETGGLFGSSGTSSTSGSLFGNTTGT
AGTGTLLGAAPASGSGTSGGGLFGSSTTGSSTGLFGSGIFGSAQTIQQNSAAGGLF
GSTSAGGVGGSTLGGQANQTSLFGAIGSLSGCAQGPASAAAADAYGLASLLGGHV
EVKLTLSARPSESSAQESAQIFPPRMQLLDPCCGTALGDTGLLGSVSSVWRSAGAPARLC
RGRFRPLPGMEGGLSGSYLSSGASNASFALPEVYIQKAFGTSRTASPLEDESKGGQANA
LRLLPDGKAFSAAPWASTTLERSAVEALLMQRRRKPGRLLPDQVSPWSSTALQFMRSNPF
YREQLAEVQASKDTRMQRDPTLPTSQSAFAAAAGVKETTLASEEVSAPASVSCPAVPGV
LPASAAVPEKRGWTANVEVSPSAKKLAGLIGSPIETFLATPAGSPGTSPNGSRTGPAE
ECTRWHGAPPASGSVAAGLESVAACLPAPEDLRPLVLRPDYETVPSIEVLGTMTEQKLSR
VQDFSITRRGYGSIWPGYTDLRGINLDEAVKIEKLEVTYVYNEAPPVGLNKRAVITL
KNCKPRSVKYLDTLTIQRPEDEAYVQDKQCQYVTKVRRYTERMGAKFVLDLNLATGEWTFE
VEHFSYTYRFLDEDDENDEELETQLRQKALLPSTSAHQPTLESPVKVLPAAHDSFPASPEN
RTRDFGEAPERITHLFFPERPANGYSSERSFLGVKQVQLKAGFAEADRETAYLGGVVAE
NLVNPFEFLKNVSLAESVNRHCQQRWRPLYVSGEAPSRGLKRVSYNSGKTVEAVHDLKEKA
LPGGERMVEATGVSGRMWFEPQVNGVEASHRGVQRFQDEKTEDKGVQDTERRDGGAAASL
RLEDGTGGSVRAVPPGRPCEDTRSSASFPHLQRRHCPYPVRCAPVISRDGLMALPLLS
LPVSLLEGITTASLPGSLVQLAHLSPLLREEMPVQNGGIISSVSGLEGGAGLEAPVQPS
PMWTVGEVPEDESVSREEEKEQALSMAAQRVYVQPLREAKGAPSLTAASLPADVSPALVA
ASRLRPFQADLATSAGAPYTNQSPTEQLSLVRDGCRRVRLTGASAMPLVLSAFLEEVVKEN
GESELQSREEDSQSESSARADATQEFPLALTPAPEVAEMGKEGRRGRRSKGEAGSSGDS
NRRRFRSQSEWLLTELCRQESHAKSRRSKVKTNLNAPSSGLVQRLRLRLLAFFEQ
QTRVYSGCMSPTSSVSVGNKCSLPLSSQPFQSAVSSAHPPDVPEASYLAPYMLQTWQ
LLVALMLSSPEQEEESVFTAPSGLFSQORDISPEAILESQRQSRLLLEWLLRESGREVTALLQ
RAAALRASPPFSTSEGLSLDRGESLCGATELCRQRGAAQNPAITRLGANALFRLAGTAGQ
RALQIAMAADHDEERKLLAVFHLAAGQLYDAVELLNSRPGEPYPHLALCLAHVQQQ
VGREFLYNNLFRATAPSFLLTPPPGIARLYRLLTPCSTRASPSPGSASVKAKEGKRRSI
SPGVVEPAEASRCASPSKRHAGATDRNSGAKHETHALSLHASWKSGETTEANSVSGMGL
EHFVSWRHQLTASLVFSSASPLEDPRVLSTGAVVSSGNAEKKVDSVDEGEKAAACTQED
TVTQTKGEGSAADVDFQIRTELTPPPQAADAPRELRRALLQFEHRLRFRENGDASQSG
VEFAVPASPAPLYRQDEDEDGETQKDRMHDGLQGATSNIFDLQYGLLRMHAGLAQPSLSIF
DPSHTPYGLDFFFQAWTAGVVTLLHRGSARARVRLKQTESEGTEAAEAIDAEDGLEQEE
ARQLHRLTVAFAAELECPLGCPWACAALLFTPFGRALQGMRALIGRHAAEFTCTGVGP
RQSGKESQGLQAFPEKAAFEERRREILRVLLEEVGVAQWWLDEADGLYALSQRKFMQAAF
LFYWAYLRRLRPLVRGFLGSPVAAPLREELNLNWSPTGAMERHLLRHAGRALLQCLPGFL
LAVLLQQIQHVEKETRVSRRRALEKCELTGSVTSGTTRALSDEL RAYERQERREEPTE
VCDEMGSMPSVLILSTEAKMRLLLQILEAIRVKLAGGLPDLSDANGTDGSRKGVKEDIL
ASCVVGGECRSKAFVVDLARLACMEKALRWLLKKQMRKQKQHNRTQRTSPGSARVPSD
AKLSGATAKSSLEEDGDRVDADFSAAALLREVEKQTKHVEAMSDANRGGFFPTDISPDEPAF
WIALKTALLEERRE

Figure S7. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). 45 FG di-amino-acids are highlighted in yellow. The autoproteolytically conserved HFS domain (part of the larger Nucleoporin2 domain) found also in yeast Nup145 or human Nup98 is highlighted in green. At the C-terminus of the protein (amino acids 2309-2571) lies a Nup96 or Nup145C domain.

Toxoplasma gondii NUP69

```
>TGGT1_273850 | Toxoplasma gondii GT1 | hypothetical protein | protein | length=723
MFGNTATTSSAASGGGGGGLFGSTSSPFASSSSLFGSSGTSKPGGLFGAAPAGSGTAT
PSLFGGTTGGASLFGAASTSSVSSGQTLSSPSTGSGLFGSASAAPSLFGSTVPASTASA
QAASTGTASGSSALFGALAKPAGADATGAPAGGLFGSSTGAQGTGAAAATTAPAGTG
LFGATSSASQPSTAGTGLFGAASSASQPSTAGTGLFGATSSASQPSTAGTGLFGAASSAS
QPSTAGTGLFGATSSSSQPSTAGTGLFGAASSSSQPSTAGTGLFGAASSASQPSTAGTGL
FGAAAGSSQPSTAGTGLFGAASSASQPSTAGTGLFGAASSASQPSTAGTGLFGAKPAADG
AGAHTAEKQATAKTEAPAASAAPAEGTKATGGATGLLGSSSLFGDTKAASAASSTGATT
SLFGAAAAAAGGATGTPAGDSASSSSVSAPFGAGLGASSTANSGSSGNAAGGLFGAA
AAGSAKAEKLGASGPAGATASGATTAETLPPPPQVALETQHERVEEVLAKEKRLQR
RVRRFNEVAEEVGSVEKAMIEESKKLHALREEQIKIEKRQTYICDFIDGLERQQRDLLTL
LASVEASVLRQIPQDNGDPTGAAGGDALAQRVQREWEAESGFHSSSEELLSRRLRNIDE
QLNDVGLALSEATERFQPGPLGTVAQVVGIIHQAAALQASWRQASELQQRMDALQRLTSDAK
HGE
```

Figure S8A. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). 25 FG di-amino-acids are highlighted in yellow. An internal repeat region (see below) is underlined; at the C-terminus lies an NSP1_C domain (in bold) which can bind Nup57 and Nup82 in yeast (<https://www.ncbi.nlm.nih.gov/pubmed/11689687>).

```
SSASQPSTAGTGLFGAA
SSASQPSTAGTGLFGAT
SSASQPSTAGTGLFGAA
SSASQPSTAGTGLFGAT
SSSSQPSTAGTGLFGAA
SSSSQPSTAGTGLFGAA
SSASQPSTAGTGLFGAA
AGSSQPSTAGTGLFGAA
SSASQPSTAGTGLFGAA
SSASQPSTAGTGLFGAK
```



Figure S8B. The internal repeat region of TGGT1_273850 is composed of 10 near-perfect 17 amino acid repeats of the format SSASQPSTAGTGLFGAA; a weblogo of the repeats is shown next to it.

Toxoplasma gondii NUP206

>TGGT1_305790 | *Toxoplasma gondii* GT1 | hypothetical protein | protein | length=1918
MASDKKDARSKSSKSEKVAKSKKSSSRSGSGKSSSKDGGKSAKSASKACGSTSEVKPC
EEQVSIPTQVGEPTDASLAPNTNPPPEQTFAVSVPEVPTWEAQVPAEEEDVEASVNASA
EPVVAAEPDMPHPAASAAAADENPLVQTEQTPVAGERKTDANNLAPPPARQPDSPMPSPT
RGMGPRGRLLAGSSSPRAAFGASQPPPPSFGVPPPPPPVGGVALLDQGTQKKGEHSRRHLN
ARLMDGDTYETERVGRIQFVSMARTCIGLHDESGAPARSNCPPCSYLHDVVMVCRPPG
GTLPRDWVDYQDMHDELVLKQIREWRRKEAEKADTVVNAEVAENVSIFPCKSFLARNFA
ETLPEGVRLLPSRLQQLNRNRRRRLRHEKTGTEDLEGEISQDEEDVAAEGGRVCAKPRG
SPQGAAEPTKSFADKHIKETPGADSRSATGEPETPEKAPAVEDNAQPASSEDAKAGAVE
TPADLAYEEDLTFGLNGGKVVAVINLCMTERYYSQPTLRRDGEIAYWIKVDGSGDIPDDK
VFCLFFRVIAFLAHKYSKLLPFSPESFATGWYSCGRDKMQEETVTVEDTPTGHAVCCDKF
TIVVHCTHGVNRTGLFVSLLLATLNFNCSAEFAVKAYEAKRGAPLSKEVFNWIRQCKKAG
IPERFRSALMPPEVSSSLFTFVQKLLLEGGDLAGEEELTLPETLTNALDGERKEQNGRQ
ERLEAHVERMRQLRAALKAKKAEKGEKTHEGQEAERDDQAEGEKAEKAEKAEESKDE
QAASAAEDSEEGQDAEETKENEELSRERALEVVKLPARLPADGIVLFGPIQSSLLAPEEL
LVSLSSQESAKISGFEIMTGDDLHGQPHYPQLLTLRTRRRETPEKGRRGDGNKRRRG
AKGKAGAAEETQGTGQQDGGPPSKEEMKGENEENEEACEVEKKEASVKSEKEAAAAAS
DEDSKGHVSEDGTEHDTPAKKRRLEAKLEDLEEAVKDEEDATAEMAEEKVEETQGEFTA
EQSAPAVPEPPNIDELSIPELHTKGQGLYSQKQIAYRLIQFREEDPLRLHYIVKVQCADI
ATFEWLLSHSFGMRMQMIQGYPSLYKLLYLRGKRRDKELFARAQEAQAENAARQRNR
NVAAPKKGFRTGARPGAPGPFMGSRPGVAPPPPPGIRPAGAVLPPPPAPFGPPVTPGF
APPFLPIPGMAPRGVPCAPGVSQPPYPPGASPLFCPPPPGAPPAPFGEGRRDGTPT
PGPPGPFGRLYEPRSPPPGAEPFRMYDDYVSKSAGEMGYRGYESQDRDFYGREPMKS
PQYPRGPGGETGRRGPGCSWPGGPEGDREGYHNQASSPTERGPYMRGPGPAGFGGPE
GMRDGLTRHGFGEEKRPFGPCGGFGFEGTRGGRGPDGFGSGARGCNESSGGRGPF
RSRPPGDYEGRGRDDFDRGYGPNRGGMYSNTGGTQGFEGPGFRGRGFGFSDSPGAGG
DRHEPGRVRFSSDRFGDRVSSDRYGDRGDMWRDGRDQDEFRRFRDGSFRNAARDNKN
FSNFGYGNQNPAEFGGPMGERGGDFRARMESGSGVGGGEDRGSRRISGWSSGFEDSRGN
SQFGRFSGDRSGDRFADSGERSRFGDKEQAGGFGRRQGRPNVSRDFRPPFGGSGSG
FADGSRDDGSRVPSLPGPPQWTIQKGPEAGGPPSGAFGAGHGQMGDAPGVEADKFRGSA
PPGIPTVESDSEDKSNLSASSVMQPQSFGIGQESSAQNAFAQFAKQLASGLQDGGQGG
NSMAATMAMMGASGSSQAMQQQLVALLSQQQMYGNSNLQADMQQLMMYAQYYGCLLNG
MSEQQAAAVVQTQQLLQQQQQLVNLQQLQSIQQQQQLQLSLAGLMQTGTPKPKDKV

Figure S9. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). 19 FG di-amino-acids are highlighted in yellow.

Toxoplasma gondii NUP67

```
>TGGT1_306560 | Toxoplasma gondii GT1 | hypothetical protein | protein | length=698
MFSSATASGTSGLFGSGTGGGLFGGSAATPAGASATAGTGLFGSAQATSGGLFGAASGG
AAGAAGTSRPGGGLFCGGTTTATGAGGTSLFGAAQKPKATGATTGLFCGGTAPAGTAGGG
GLFGNSQAAPATGGLFGAQAQSTGTTGGLFGAQAQNTGAAGGLFGAQAQNTGAASAG
AAVLPREPTVANIDRVVPGVLEKFKKIEEKMREEEKVMNELQAATRKMRESFSGFSSLS
THQSRVSWALHHVLLKREANLAPVVQRDKQLALQVEQLHQLEQNVTAANASLARACS
AGGSRDAVYVPLQVPCPLSTPLYQQLLQEVWSVSGKLQQLQEQVKKLLRLETERAARSSC
SYTVDGVS LAGGYSESGVDFHSEVQMIKEVLESQREVLLATAQKALELRENTRRMQDYLE
RRGVKIPVFPEDQEEVAVNAMAVQATQTTGSLFCSTGPTSSPLGGIFGPLSSQTPAAGG
LFGTVAGASATPATGTGGGLFGGSAATPAGASATAGTGLFGSAQATSGGLFGAASGGAAG
AAGTSRPGGGLFCGGTTTATGAGGTSLFGAAQKPKATGATTGLFCGGTAPAGTAGGGGLF
GNSQAAPATGGLFGAQAQSTGATGGLFGAQAQNTGAAGGLFGAQAQNTGATGGLFG
AQAQNTGTAGGLFCGGQPSGTSGLFCNTGAGLFGAK
```

Figure S10. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). 29 FG di-amino-acids (11 at the N-terminus; 16 at the C-terminus of the protein) are highlighted in yellow.

***Toxoplasma gondii* NUP43**

```
>TGGT1_310610 | Toxoplasma gondii GT1 | hypothetical protein | protein | length=435  
MAAWPSSSFISSSWGGGAGTGGAVSFASLAASSSPGGTAPGGNPVGLGGMNPFSPGQAA  
FGSGAGPSNTLSTTQPNAFSGQSSSVLGASLTASTSPFGASSSLFSSASSSFCASTLFG  
SASSSSSSPFGSSTLFGTSSSSPFGVSSSSSTSSLFGAPSSSSSPFGVSSSSSTSSLF  
GAPSSSSSPFGVSSSSSTSSLFGAPSSSSSPFGVSSSLFGSSSSPFGTAGGTAPAAA  
LEGGQAKSEGTKFIDAWKNLEAAKGGIWPFSFVAAEAECLFVGLDISPEHRVLYFYQ  
RPQSEWRALGEQIVATQLQHLRQFHALVGEKAAADRLSFEAAPPVVEFLRLVGGSLPQT  
SFSVPQETLQTPSLGALGAPPAMNKAEINQEVSNLESLGTPGARDET DARGAAFAAPQF  
EPGKIPDVPPPRELC
```

Figure S11. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). 15 FG di-amino-acids are highlighted in yellow.

Toxoplasma gondii NUP593

>TGGT1_313430 | Toxoplasma gondii GT1 | hypothetical protein | protein | length=5639

MSLFGGGGSSGTAPGSKSLFGATTSGTSSGLLGASSSSSSSLFGGGSSSLFGNSSSSIF
GAGSTAKPAQTPGLFGSGATSSSFASPPSLFGSAGATEKPAASGGLFGSTGANSSSSLF
SGGGAASAATSGGGLFGSAASGSASSSSSSSSSLFGGGGAANPAASGGLFAGTGASSS
SSSSSSSLFGSGSSAEKPTASGGLFASPGANASSSLFSGGSTATGTSGGGLFGNAASANS
SSSSSSSSSSSSSSSLFGSGGATETPATSGGLFGSKGANSSSSSLFSGAGTASTGTSKG
GLFGNATSASSSSSTSSASASSSSSSSSSSSLFGSAGATEKPAASGGLFGSTGANSSSSLF
SGGGAASAATSGGGLFGSAASGSASSSSSSSSSLFGGGGAANPAASGGLFAGTGASSS
SSSSSSSLFGSGSSAEKPTASGGLFGAPATASSLLPGAAAASSETSGSSIFGGAASASSSS
TSSASTSSASSSSASSSSSSSLFGGEGTAGAVPAKPSLPGSSSLFGSGQSTKAAE
APQSSLFSSSRLASASSAGASASSTASAPAVASSMAALAAGSSPEIRKCLVSYMHLKSV
LISLSLSECPAQVEELFSSPQLLQRFASLLSPFDPVDRALPQNI SADLHAAISAVRKQF
AQRARRRSASDATDAARDRLVPEAMIYEALSVYLLRKFNKQQQLLSPRFSAPFLQEA
DPPDPVMTREVEDILVDEQVDFWCLAFIFRTAASVAPSSSRTQAPKQRTPATLSDSSDK
SLSGQKSENDRDLASAGPQRDAAKTSSLEQACFAFACCLVREKNIIEGLFRTYDELRRFD
GLERSGRKRSALTEDTVKREARASRSPALQVVAEEPETVKALAAATLRIQKALLHCFAAF
YRARNVLEACAKEKPVSAATLHRLASMTDDSFVGAALASSSEDRLLHSSRRLRAGQE
VADASALLVAAYPALFRGALGHEQEAQETLAQLGAADHFAALHRGVMTPLLNAFK
QRLEREQSDWDLALSSIDPAGEPSQSVDSGDEAESRGPYSAEYPDALYEQFI AHPV
VPIVLF AHATALALLSSEDKDASSAGRSSKEDCSLSAAEVD AFLQPLLLCGSPSFAAALE
ALVVVRDGLRQAGSAAAADP LLSMQREILLEFVHLLHLLPHHFQGFELVHAVSGLV
DSPGVAEDLWLRFSRSGAFLFSPVAPDSAGPGASTGKAAALARVEPSCSQCGCWGDFV
SSLQRETRLTQVRSRSLASPPRYGGLYLVLDLLRCVSPFGLHLLHLLQVLSALVPRSPA
SRVVEKASEAEADAGDDEEVVDATSLAAFLAFLTEPLPGILLPPLPDLFLPADAGDER
ALRGETAENALPRPGPSPLRRPYISVGPVLELCLAVMLGFQEPVDDKREVGQGREAF
FDASAVSGGAEVWSSLLPRGVS AEVQGLPVWRLTNAGATACATTAALRQRFELAEGRSD
SRNPTASLLANDSRPPCLDGEGWGVSTDGARAQPALRVLEPGRDNATSRGRLLSGVSA
SWGPEQSDFFPAGESSGDASAFMLLHALRVRLGGRETECEETGRSGKGEEREHNVELPCL
SLLRTAWFVFDAGLQFLTQTGKPLPDAALKSFIAASALFARLAAWPPAMRLVAELVIAEE
VYNVGGPAAAFVSNAAATAASSNVREILALFPRKAKASKTKGDSACTGEDDKAAGADAGE
SEASGDSAAAVDAFLASSNRSSAAVAPFALMPLRFLALACLAAQQQLRCLPEALTGRCC
LTPCSLQNIHFVFRLLWPVSELGPSSSSSLSPARECRSLFSRFSSPDEGRLETPLASRLP
LVLLQPQTSELLTVDLASPEAVHTYQQLMLLISVGSRSRSEASSLRRSLFRPFSLSASP
SSGAFFPPEFFSLECFCEFFSAPFLSVSEGVWGDPTGAGDCGARGGTAGLLSPFLRRL
LHLLGEAMEEKS SVGAYPV LKSTLLHLLNLLQCLPPAALTLSWQQHAALS AATVSESARE
LLLRRERERRHEREESDPSLEGEAEHPMFRAFLAPFRDLARMQFGAPDADVCGDGLLC
PPDLAREETANKGGSLEESYWVLFPTPASRPANRRAGLFRGDREVI AVEPEEEAGTESRE
RTAAETGADLTF LAECAESQRLRVLFERRQLTQLEEQRDTFGAICEFLFSSVWTRLASCS
FHFAHRAEAVVLESLAATTSTLSLQGANRLPGAGEETPSVGPQGHRRPDSRTRKSRGD
RQRDRDGRERWRQCEALRGSEVDDERIGDLQRDADAPGTIKATLAPDGGAWVFTHAGEA
RAPHSRFERERDNLQILFWDKQGVFAKRAATLAVYGHVELTNRLLQLSSSNFLSSLA
SLLRCNILLELCPYPARQALPTEGSPFLFAFVNSLLSMASLDFHSLSSVAWITPLAAT
MLGTPPVQARRLLSADTTTLLPAPCCSNARGESTARGDSTAPGCRCACAFCSASSF
SFSSFPVAVGSSSGLFGGVQAAPTFLRETALGLLPRVSLPAGPSETPYSPVSSLGRWS
AHWTGRALVENVPLCLSASGLSAVSTFFAFRAAPSRLLDTPVALLPAEAAALCVSAVTSP
NTDRSQNILCLVAQAFHLFLATAVSP TGAALGACEAATSSLSPPGTSSACDASLLLPQ
RSPDGE SLGRKRLSLDQSTSCDSAPGLLSRLQRLQWLKLA FSPASTLPEDLGISSPL
WPDVDPDLLPSHPGDVLAALQE QETFLLSQGAATVRVSPAARRPVWRRSSGTTARAPANF
VQSVFALLEASSRLKVVTASLLNLLTVASLGLQARRRKGSETDPRGADCFPAGLGEAVA
PGRAHVSSSPFVAVCLTFAPKEVQTVLEREDARGQVEKASEKKEVEKAASRWRKRRLRFV
DLLFPGLSKGLRDLAYKEPMKEYRLLHLLQLLHLLSEPSDANCAPLALHDDLIFLGGG
GGFLDGEKEEPLLWAVDFAQQIILLDVLDTAAASSLGDAGEKSPGSRTQLRLASLGD
KLLRLGLLLEQLCF SAGA QAVHLQLLQQT VLGVRRLSRGRRLLHTRGAGVDDAFVALT
GGKGGGRVEKSVGWKLLLRVAHSLLATWQQVDWTRARLWLAATACTPGVPESEDSGRV
SEGHE TLLPLASVETPALDSGGS AVF SRLISSAGFDVPLYASLGASTATS GVHTPAFSRR
PSGNLRASVSGRDSAKAKSAAGGDEAGSGAGSASRSSLFPRASGDGDKGPVFSFGFDE
EEEELG FVASLAAIELCGAFQALLRILVSMHALPLAVRVQAKASEAGKPKHKSLSISDLF
VDHAPSLSDYREFFSLFFFLSSDPTLLAAFFPAQRKTHAPHTPSLGLSPLPCTQRSLLRL
LLWLRSRISPAYLHQAPRPPASFSLVPRPGDAAACVVAASSPSPVSHPTQTASMG
GRGGHCSRRTAAASASLESFSDFWTRRRLLRAAVEEGVGG EACAFLES LAATSATSETF
LAAERGRARNAPGSFSP LLSPLVRACGLVSSPEQHGPGLFDRRSLVLLATHVTLQAF
PLCRPGVVSETERERLPSEWRQCGEDERSVYLHPHERVENLRPRRRRGGAGGASRETDP
GELRRGDSFRSKTRREGNSSVEASFSGCLLGEAWIAGAFACRVGVGSLADAQLSLL
AAFFQLLCVSRERGVFLYLYLPKKEGNSSVQNALEAGEEGRS AEPKESREASFASSPLS
DVQPFPQA AVFPVLSLVEICEREKSLSSASCGLFRRSEFRLLVALALQVLTCTNWT TDV
RGAAPP SRLLLRASQASLHAAEKSLSSSAFASRRASLSAPPLSPLLRHPEELDDFSQRAS
LTPSRRGSEVVVSLCPAVASGAEGPSMDTSRSDRSRSPSPAALSFGCEQSYFLSRFS
AKIQPLADATGSAGRAAFLSPLWQDEALKGAAGESAHPPLAFLAKPLFAAGVRQPCIPSA
EILDNLERAALAEASECDAEDANEVDTRASGQKDEDEEKKKAASLRRLRASELLFAV
QKGLLEVVREKVKYQKTCGEVGEVGFSERFMERKREDQKPRKQVHLPAFASSLSFFPNAPSG
ADVADPGVHLAGPGTARRS VFSPPRYPPVSDLQLSALLLASSAI SSETRQEKTKRTASSGS
PRPRHTEHPGPDAGAPLAREKALHAI CRRIQQATRAAGHQLPEADLLD LLAALYGFTV
FSTVWSFLSLSGADGALPSLQSSRNSDAREETCP SPLVTFWAEALESQLAELLDQTETV
WGRSAGASPSVSESAPAFFSECFAFSDSSAVSEETRIATLLAKTPGPVCLVRVSAALLRLS

LPRLFLNALVSLLDASVAAVHARRREGDRHEEKKPAEPGVATPRLHLPLAASRRASEDGL
ASWRRASRAEGRDASSLSTETDADPTLEPFDPLVACALQRALERLPLRPRRAVSVFFSHS
KPPGKAPDVFEDDEAASDALEGDPLSSPPAPLRAFQLPLVRFSPLAGSLSAQIQQHIA
TLNSLFPGVAPVVTRELWASQHLLLRQPVARCISLSPLDAFGDRAASAGSGPESQRDEDS
LFSPKTSLLLHDASFLVEEVALRTLGSYVAFFALLQSGVTPRRPGGAAALFSRAAAAAA
PLLDAGDGEGFFVELHTCRNSRGARQGAGLLEAPPTSERERCTDSIDSIANSHSDACRSS
LFSLAKSVSASHPSLPGALLDADFFISLLTLAPFQSFLHPLRTPAFPLPRSPQKDVAPT
VSWLPAYEALVTSSQDASCVGKKIRRSPSHLLFCRLLGLLALGLRQLRPSSGGCASPSPGF
IPGDSGVECGTQLTTEAAGRDRERARTAGLIRSLQLLDGRMQFVLGPQGFAQTPQLATLE
EAALYAQLLTLPPWRSMEIEQQQLVWRWFSLLRDFALSVVHGLLSLVDGTCVLRPVSL
LECLACQPPPTASSPRFSAFCASPDDPEGRGESRERRAEHRQRRGFWDGGFAEASLGEDT
GESRRVIPSTFDQRGIYMLQICNHTVSASFSGFLLPLSHTLPPPPFFLPGCACSSS
FARFSSASSRWCRVRVAVSPKSYCALITSLFSSPRDEETRLASPSRAGRQREGNGRKDK
PEEEKRGAALVREREADKAEGPQGEQVLERRRHLLRDADVESEIPPPEREAPTPRGGK
CAKRNAARERRDFLPTADEVAQTLAAFVEIARSLAKALEELAREDRTPPLAVLTAGDRR
FLPLSLFLRCEGPMNRVSTLCSSPRPSQRQTSGEMSLLASGWARSPTPGSVSPSHSSVC
GAGRASSVVSMTAAGFLSEETPRELARRDVLKDTGLKDGGLGRDGASVVSTRRGPSTEL
QGGCASFSIVSAGRQRAEQRRSLQRRRRSSPTVTACRVDRVTPRGCVASLLVSKENEQRC
RGRDEGLVAPSLLGIEEDEELEVCPDLLLQDALWELLRDTLEKAMAAGTQIVNAFVAQ
NILRPYADAAKQFLLLLIHGAVSGHPSMSSLPSPPSFLPAGAAHFLRATETKGEETGKE
IFSGASSRNFYATLLEYLLKRYAVEEKIFEAAQGVGFPLLDLRGKRRKAEGADSDVES

Figure S12. FG repeat *Toxoplasma gondii* protein (sequence from release 29 of eupath.org). FG di-amino-acids are highlighted in yellow.

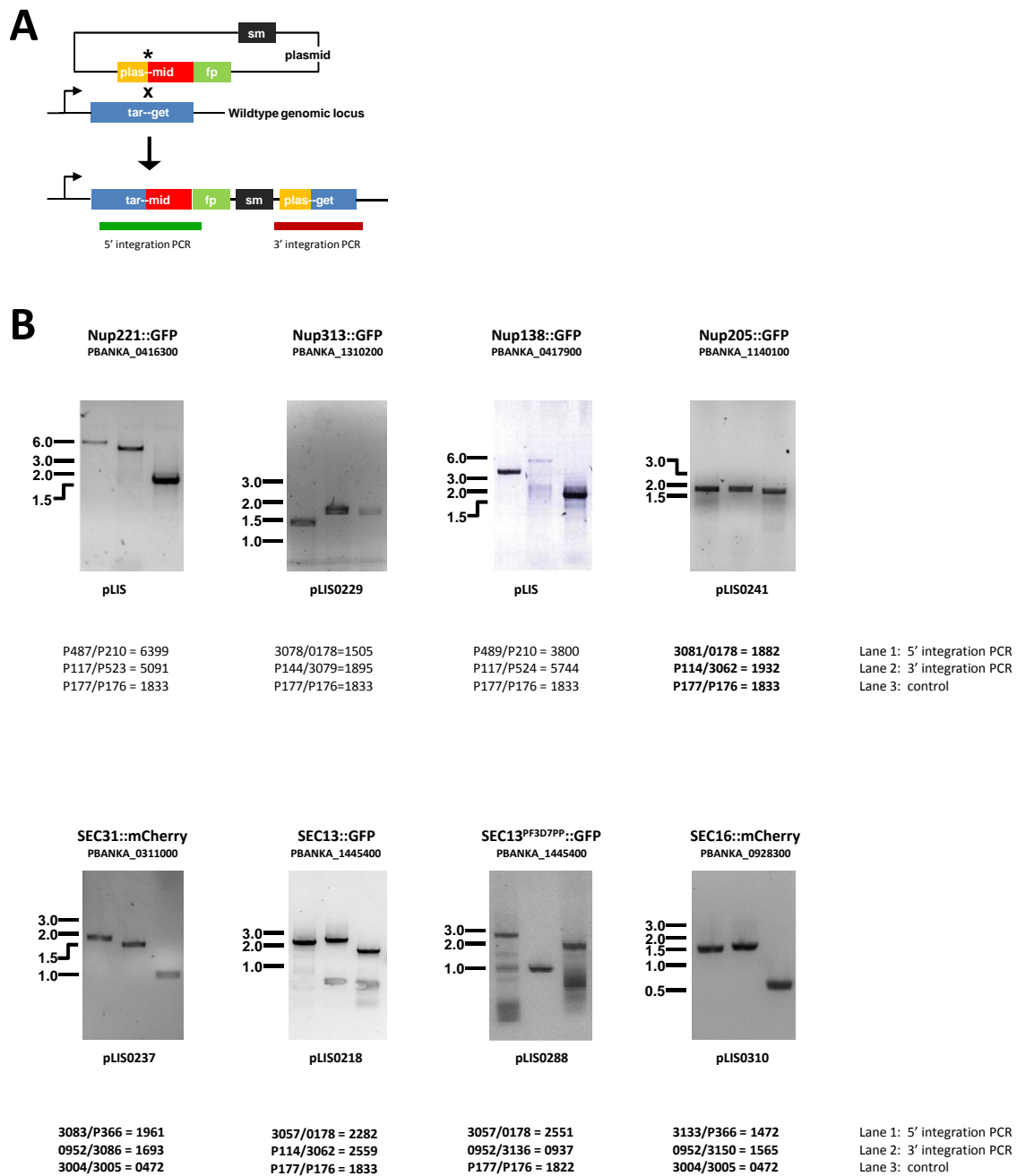


Figure S13. Polymerase Chain Reaction (PCR) genotyping of transgenic parasites. A Schematic overview of the *in situ* tagging strategy and indication of mutant-specific PCRs. **B** As shown for the top left panel: indicated are the tagged protein (Nup221::GFP) above the agarose gel separation of the PCR products with the size marker in kb; below the gel image are primer combinations and expected sizes of each of the three PCR reactions (5' integration, 3' integration and control)

PBANKA_031100 1 MALKSINISGNFDWCPFEYKNYLFCFNHNLNLYSNNGLNLYIYLLDIN
PF3D7_0214100 1 MALKSINISGNFEWCPFEYKNYLFCFNHNLNLYSNNGLNLYIYLLDIN

PBANKA_031100 51 LNSDIRSLDIVHKLNFEEALSRE-----NNKSNKTSSENEY
PF3D7_0214100 51 LNSEIRNLEIVNKYNFEDALKYDNDVIKGGKKNKNNKNNHNNNSVNEY

PBANKA_031100 86 VTSFEWLNCCNNEVESNNENELNKGIIIGGLTNGNITLLNAQNLFDTNCVN
PF3D7_0214100 101 VTCFEWNNNSNNFVDINNNNEELSKGIIIGGLTNGDITLLNAKNLFEINRN-

PBANKA_031100 136 YDNFILSQSNIEHSSINCLEYNNRHKNNLIATGGNDGQLFITDIENIFSP
PF3D7_0214100 150 YDNFILSKTNIHDNGINCLEYNNRHKNNLIATGGNDGQLFITDIENLYSPT

PBANKA_031100 186 SYDPYLDKNNLQKITCLNWNKKVSHILATSSNNGNTIWDLKIKKSAVSF
PF3D7_0214100 200 SYDPYLDKNNLQKITCLNWNKKVSHILATSSNNGNTIWDLKIKKSAVSF

PBANKA_031100 236 RDPHSRTKTSSLCWLENNOPTQTLISYDDDKNPCLQLWDLRNSNYPIKEII
PF3D7_0214100 250 RDPHSRTKTSSLSWLSNNOPTQTLISYDDDKNPCLQLWDLRNSNYPIKEII

PBANKA_031100 286 GHSKGINNICFSSIDSNLLSSGKDVTKCWYLNNSNFDIYNEVNNSGNNI
PF3D7_0214100 300 GHSKGINNICFSPIDSNLLSSGKDVTKCWYLDNNSNFDIFNEVNNSANNI

PBANKA_031100 336 YSKWSPPIPDIFASSTNMDTIQINSINNGSKMSTKYIPTFYKKEAGICIG
PF3D7_0214100 350 YSKWSPYIPDIFASSTNMDTIQINSINNGSKMSTKYIPTFYKKEAGICIG

PBANKA_031100 386 FGGKICLFDNFKSDNGSNLSASESASNKNNANQNLONIADEASMQINKNRG
PF3D7_0214100 400 FGGKICLFDNSTNNMSNVNMMNVMNMMNINISFNNDNSCDEYDSNKGK-

PBANKA_031100 436 LQNEGGNSNPNVLIKCHYPTVDLIEADKFEKYIACGKYQEFCEKNI
PF3D7_0214100 449 ----NKSTQKRELKHYHYPTDMLIEADNFEKYITSGNYREFCEKIN

PBANKA_031100 486 KCDDYHEKLTWQILQLLCTSQKEEIVKHIGYDMNEINQKIVEFIGEESGF
PF3D7_0214100 495 KCDDDEHEKLTWQILQLLCTSQKRGDIVKYLGHDIINNIVDKIMQIIGKQPGF

PBANKA_031100 536 IFKKYRCETNENMNNGCNITSNLEIDSDMNNKHDNTRMNNLNNGENISG
PF3D7_0214100 545 IFKTLIDEKENNNNNNNNNSTNQMYQNDVLLHNDPILMNNYLLKDNMNP

PBANKA_031100 586 YSNYGN-----MHGEFVTSQMENEPNFNEPFDLDPEKFFR
PF3D7_0214100 595 IMLNNNNNNINNRTGTNVMSYNGQNLGDTNHEENFNFNFDLDPEKFFR

PBANKA_031100 623 ELGKTEIEKKQENETFKEDKNGKEDLKS SVNVNVDNNTKDMISTIKQNP
PF3D7_0214100 645 ELGKTENEKIKQNEEDISGNDEHLLNSIKGKENTKNKKSGLGTDDNN

PBANKA_031100 673 EIGFGEINEDSKPN-----SNKQNSNNWNTGIESIIEKCVLIGNIE
PF3D7_0214100 695 DNGDHKNKNEGSNINGEHVSEHILNEKNNTNWNLGIETIIEKCVLIGNIE

PBANKA_031100 714 AAVELCLYQNRMADALLSSFGGENLWHKTKNIYIKKQONDSFLRNINYVL
PF3D7_0214100 745 TAVELCLHKNRMADALLSSFGGQLWHKTKNIYIKKQONDFLKNINYVL

PBANKA_031100 764 DDKLEYLVKTIIDLSSWDEALSILCTYAINNPNFNLCETLAKRLQNEKFD
PF3D7_0214100 795 DDKLENLINNVDLNSWDEALSILCTYAINNPNFNSLCEMLAKRLQNEKFD

PBANKA_031100 814 VRSASICYLCASNFPETVEIWDSPSPKSTLLNALQDIVEKTIIVLKMVIK
PF3D7_0214100 845 IRAASICYLCASNFPETVEIWNMPSPKSTLNLVQLDIVEKTIIVLKMVIK

PBANKA_031100 864 YNKYNSIMNQINQYAEILLANSGRLLKAAMTFLSFTENDNAIENLTLRDRI
PF3D7_0214100 895 YENFNSTIMNQISQYAEILLANSGRLLKAAMTFLCLIQHDSIESLILRDRI

PBANKA_031100 914 RNSAAHIMPPHFKPPSPFOYFDIKPFG---LSQKYNNNNITNTSITHGI

```

PF3D7_0214100 945 YNSANHVLCCQIKPPI SPFQIVDIKPSPNVYQNNMYNNNNNNNNININSS
PBANKA_031100 961 YNKSQIHPNKSITIS-----I
PF3D7_0214100 995 SNNNNNNNNKVLSSMHHMQQFNQCNVNKMYSSTSNLIINNMTMNSNFKS

PBANKA_031100 976 VPPLPSEMHKQHGPSTPMVDFSHNSPTKFNQOMISGSD-----ISR
PF3D7_0214100 1045 VILPPPMNNTQMNSTSSTIQPPSVPTKFFHTQIINNMMNSRSSIATTK

PBANKA_031100 1019 LPFQKDFMGSINSTSNRGYPTSNFNINVKYPSNGMGSIIYIPPPNIPSHST
PF3D7_0214100 1095 NYPTSNLNSVITPSMNNMNTNLSHGNNVTPPYMSQTNVAVPMNNNNNNN

PBANKA_031100 1069 TPINASSPPHTSVQNLAHLSSHSLNKLEHENKIQGNNMFESQSYTINVK
PF3D7_0214100 1145 NTMNPVYPSLPKFPNYNLNSQVQONSIIPEKQLTSPVFSNSYGNINKTH

PBANKA_031100 1119 AMONSAPPSPNKIPNISRSSFVSQNNIPPPFSQRNRHTSVSTSMGN---
PF3D7_0214100 1195 TTNNAVPPPNVTSSVTPPMPSQNLNTRSSFADIQNVVSPRNKNQSI

PBANKA_031100 1166 PIQONIPPIQESQLNKRESIDSQVYNPVSPSII SPISIIQQTQQONS---
PF3D7_0214100 1245 SSTANLNYQHNDQFNKRECMEQPVYPMTNQSSVFSMNTMQKKNVPGGFQ

PBANKA_031100 1213 -FTPPQFYSQKVGFTPPSGIKTTSPVAGAMSI TPGMPVWPPIPTTTQQL
PF3D7_0214100 1295 DNTSQMNYGMQPTGSPPPSSLSSTSPVAGALTVTPGMPVWPPIPTTTQQL

PBANKA_031100 1262 GSTTQSTANENKKIQTVTKEQNGVLMNRSNIENVKRVISNLLNMYTSQEL
PF3D7_0214100 1345 GSTTQSTANENKKIQTATKEQNGVLMNRNHIENIKKTI SNLLNIYTSQES

PBANKA_031100 1312 VKKKADDISVKVHELFDKIDNGAFNEQINTIITINLANSINENDFKTANKN
PF3D7_0214100 1395 VKKKADDVSSKVYELFEKIDCGAFNEQINDSLLNLVNCINANDFKTINKI

PBANKA_031100 1362 LMEVSRNLWDGNNKAWIMGLKCIIPKC
PF3D7_0214100 1445 IVDL SRNLWDGNNKAWIMGVKELIIPKC

```

Polyproline-rich domain

Figure S14. ClustalW alignment of *P. berghei* and *P. falciparum* Sec31. Boxshading provided by http://www.ch.embnet.org/software/BOX_form.html. Note the high sequence conservation apart from a polyproline-rich domain close to the carboxyterminus

Toxoplasma gondii Sec13

```
>TGGT1_201700 | Toxoplasma gondii GT1 | WD domain, G-beta repeat-containing protein |  
protein | length=654  
MAAPATLSTFETSHAGCLHSVEFDFFATRLATASSDRITIRLWLSLSTPEASTHAGEVAPKT  
ATFLQELRGHEGPVWQVRWAHPSFGNLLASCGYDRRIIVWRQSAAAVPQGPQTRFAPTQS  
LFTPVYTNEEDHTASVNSIAFCPHEFGLHLAAGSSDGSVSVLSLSGDPGAPGAQAQLFWSR  
KAFAAHFNGVNSVAWAFMPAASQAGTPALMLLATGGCDSQVRIWGLDPNSQEWQQLHQLT  
DADPHTDWWRDVAFQPASASSLLSSSRLASCSEDTVKLWVGEASTPSANPSATSITW  
SLLQTLRLHAPVWRVSWSVSCTILSVACGEKDVCLFRETVAGHWKVSRLGPDLSLQVA  
PRPLPPAAAFGVPPTPGGQTPQGFQLQTQAPQLQTQAPQLQTQAPQLQPPAPQLQPHGS  
AAPLGAYPPSHPPSLSSSPPTHPAHGASHPLSSFFSSHPSLQNPAPGGLSATPPSTAA  
TPRPLGPAAGQPPQGSPTPGVAFPAFGAPAYPGTPASAGLYGPPTPGAPGGAQSYPPPAF  
AAPYPQGSAPFPAVQFAQTSLGGQQA3SPASFFPAGTAGSARPKPPAFYNI337GAF337EAG337PTG  
AAP337AAQ337PPFSGGQT337GAA337GLQ337PT337GF337PR337PP337MAA337G337AQ337Y337AP337RA337PM337YA337Y337KGN
```

Figure S15. Sec13 protein from *Toxoplasma gondii* (sequence from release 29 of eupath.org). 6 WD domains are present starting at amino acid position 3 up to 337 (underlined). The unusual C-terminal extension contains 81 proline residues highlighted in blue.

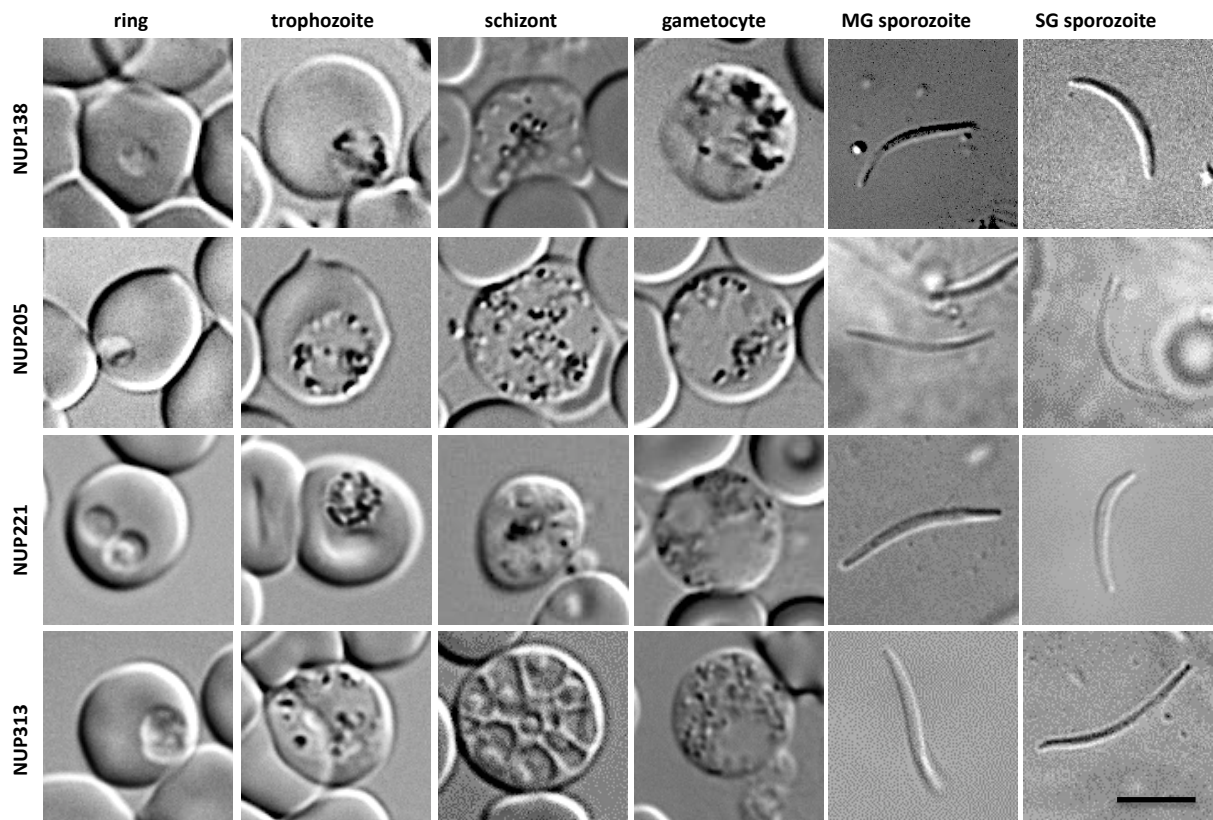


Figure S16: DIC Images to Figure 1B. Images taken with a Zeiss Axiovert microscope 63x. Scale bar 5 μ m.

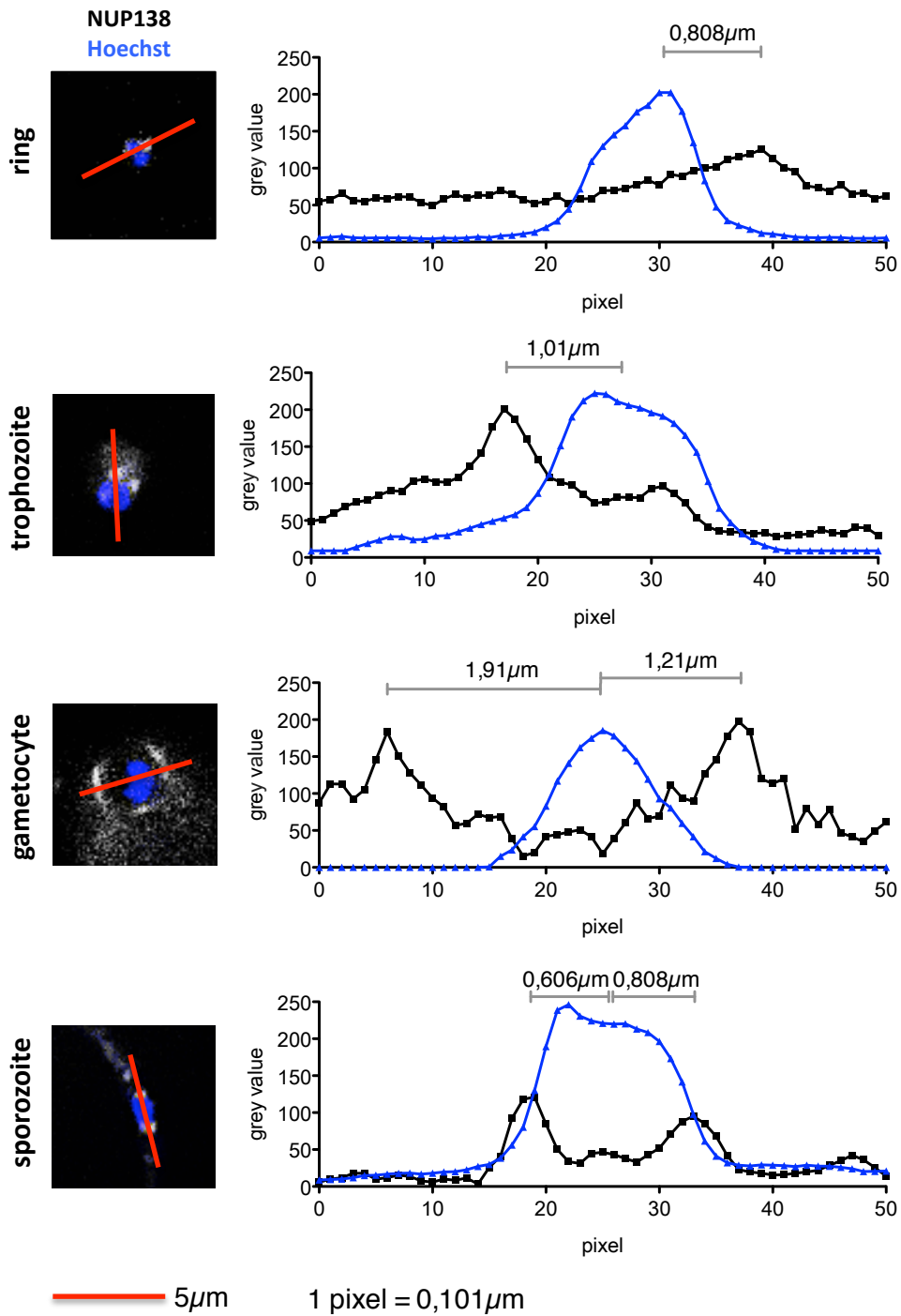


Figure S17: Position of the nuclear pore complex containing NUP138 relative to the DNA staining during life cycle progression as analyzed by line plots. Distances between signal peaks are indicated. Images taken with a Zeiss Axiovert Microscope 63x.