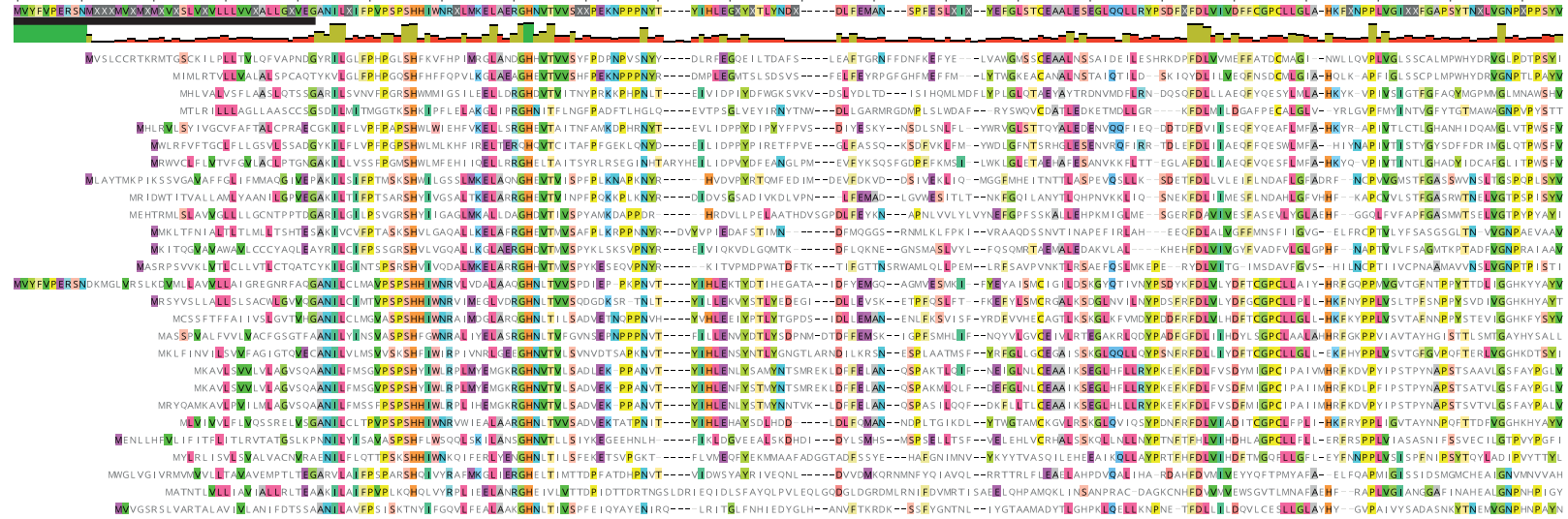


Consensus sequence signal peptide

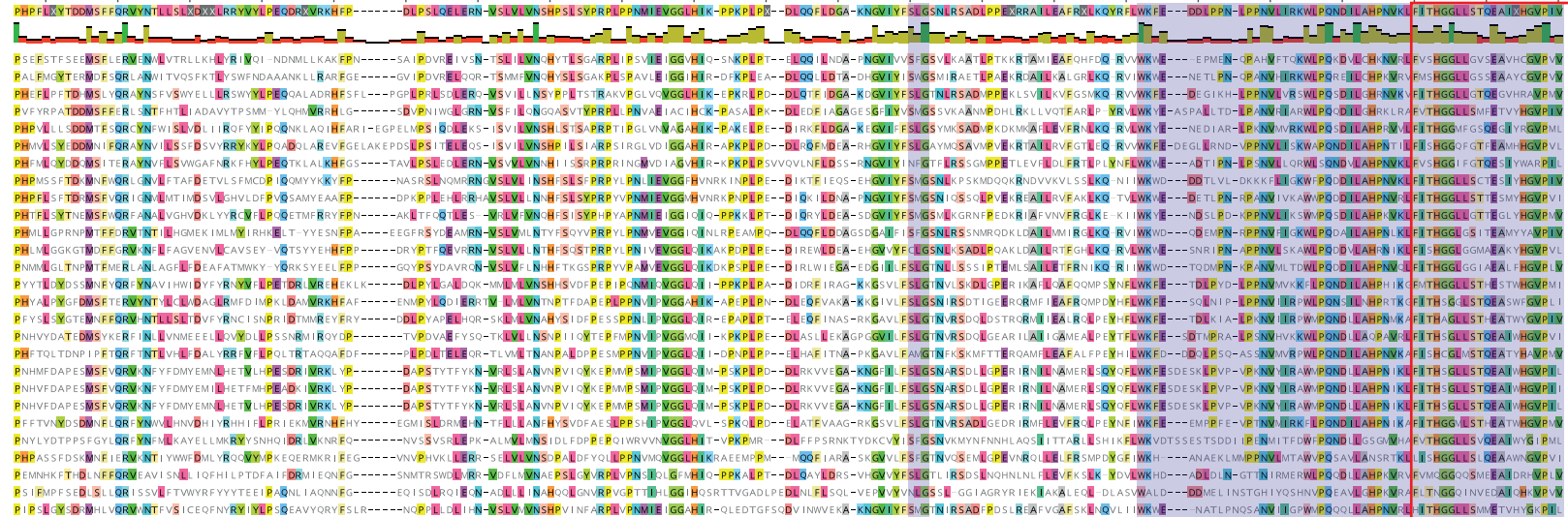
Consensus Identity

1. AFUGT36B3
2. AFUGT36C3
3. AFUGT49A4
4. AFUGT50B8
5. AFUGT301A3
6. AFUGT301C2
7. AFUGT301E3
8. AFUGT302A3
9. AFUGT302H3
10. AFUGT302J2
11. AFUGT306A3
12. AFUGT306C2
13. AFUGT306B2
14. AFUGT308A3
15. AFUGT308B3
16. AFUGT308C3
17. AFUGT308D2
18. AFUGT308F2
19. AFUGT308G2
20. AFUGT308G3
21. AFUGT308H2
22. AFUGT308I2
23. AFUGT309B2
24. AFUGT310B2
25. AFUGT313B2
26. AFUGT314A3
27. AFUGT315A3



Consensus Identity

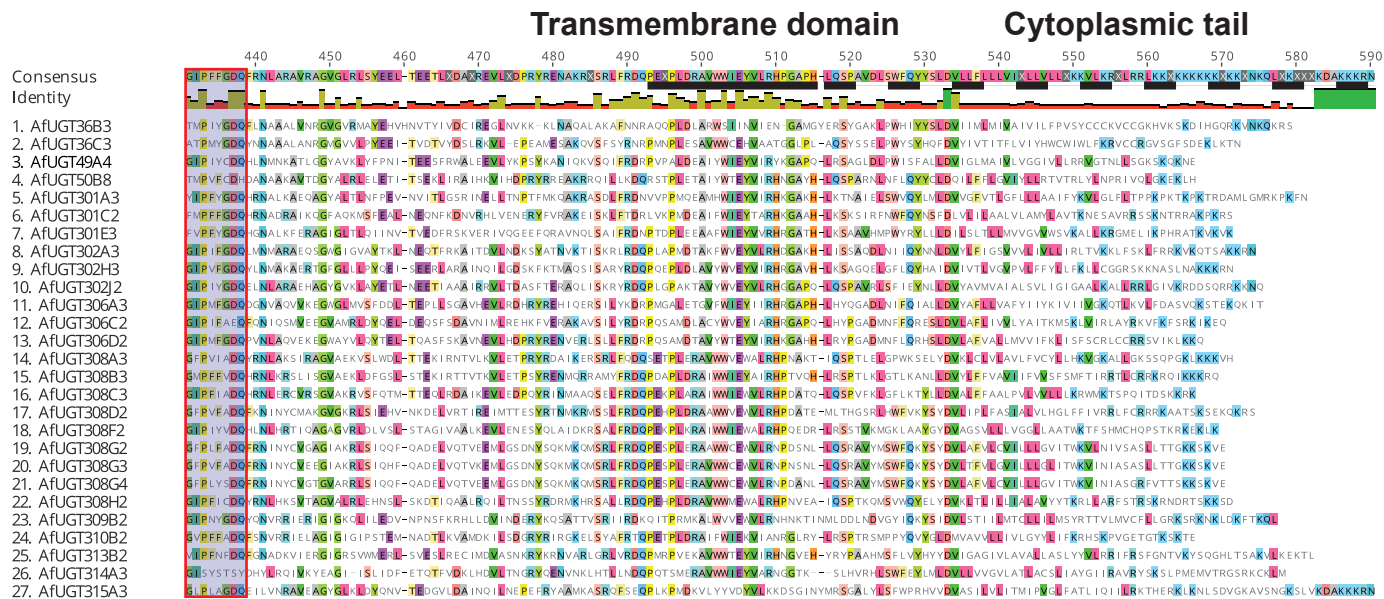
1. AFUGT36B3
2. AFUGT36C3
3. AFUGT49A4
4. AFUGT50B8
5. AFUGT301A3
6. AFUGT301C2
7. AFUGT301E3
8. AFUGT302A3
9. AFUGT302H3
10. AFUGT302J2
11. AFUGT306A3
12. AFUGT306C2
13. AFUGT306B2
14. AFUGT308A3
15. AFUGT308B3
16. AFUGT308C3
17. AFUGT308D2
18. AFUGT308F2
19. AFUGT308G2
20. AFUGT308G3
21. AFUGT308H2
22. AFUGT309B2
23. AFUGT310B2
24. AFUGT313B2
25. AFUGT314A3
27. AFUGT315A3



DBR1

DBR2

signature motif



signature motif

Fig. S1. Alignment of *An. funestus* UGTs amino acids. Alignment reveals that the C-terminal domain is more conserved than the highly variable N-terminal domain. The 29 amino acids sequence long signature motif on the C-terminal domain is highlighted by a red box. Conserved signature motifs including the DBR1 and the DBR2 motifs are highlighted in purple.

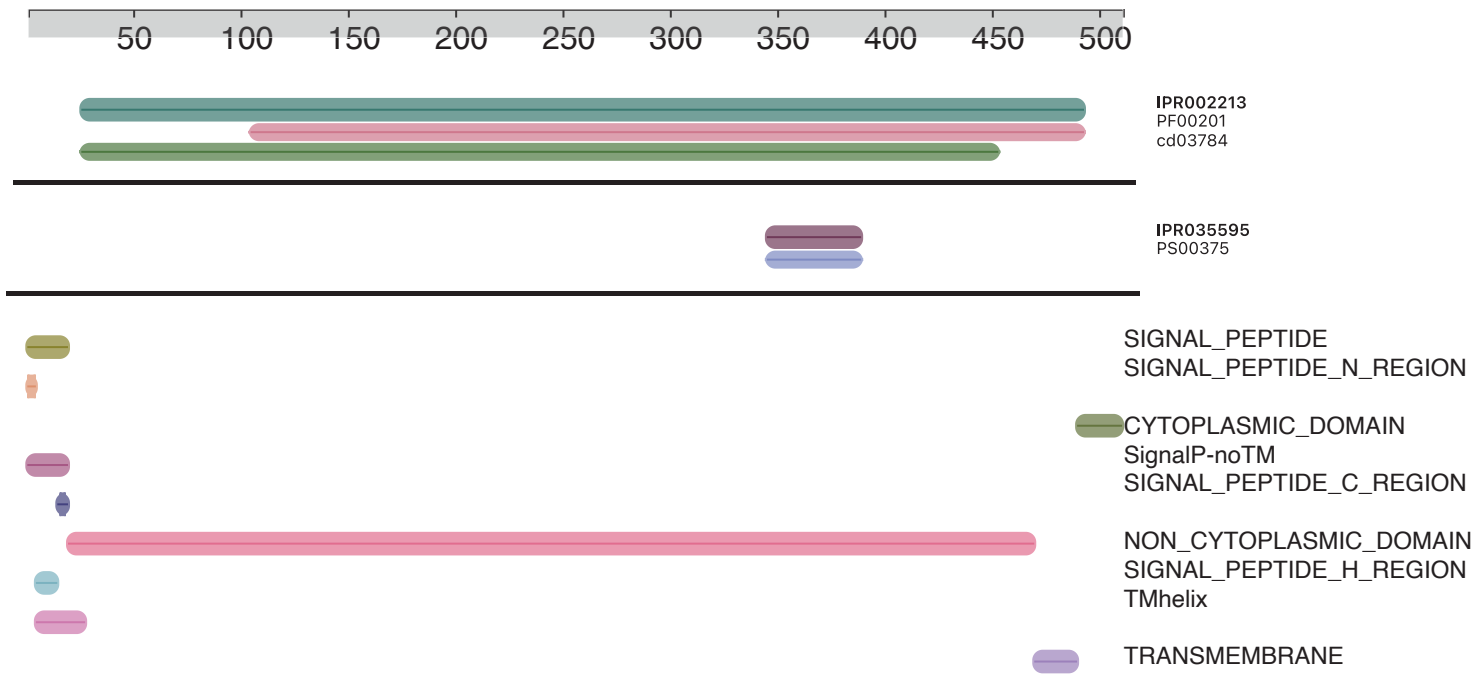
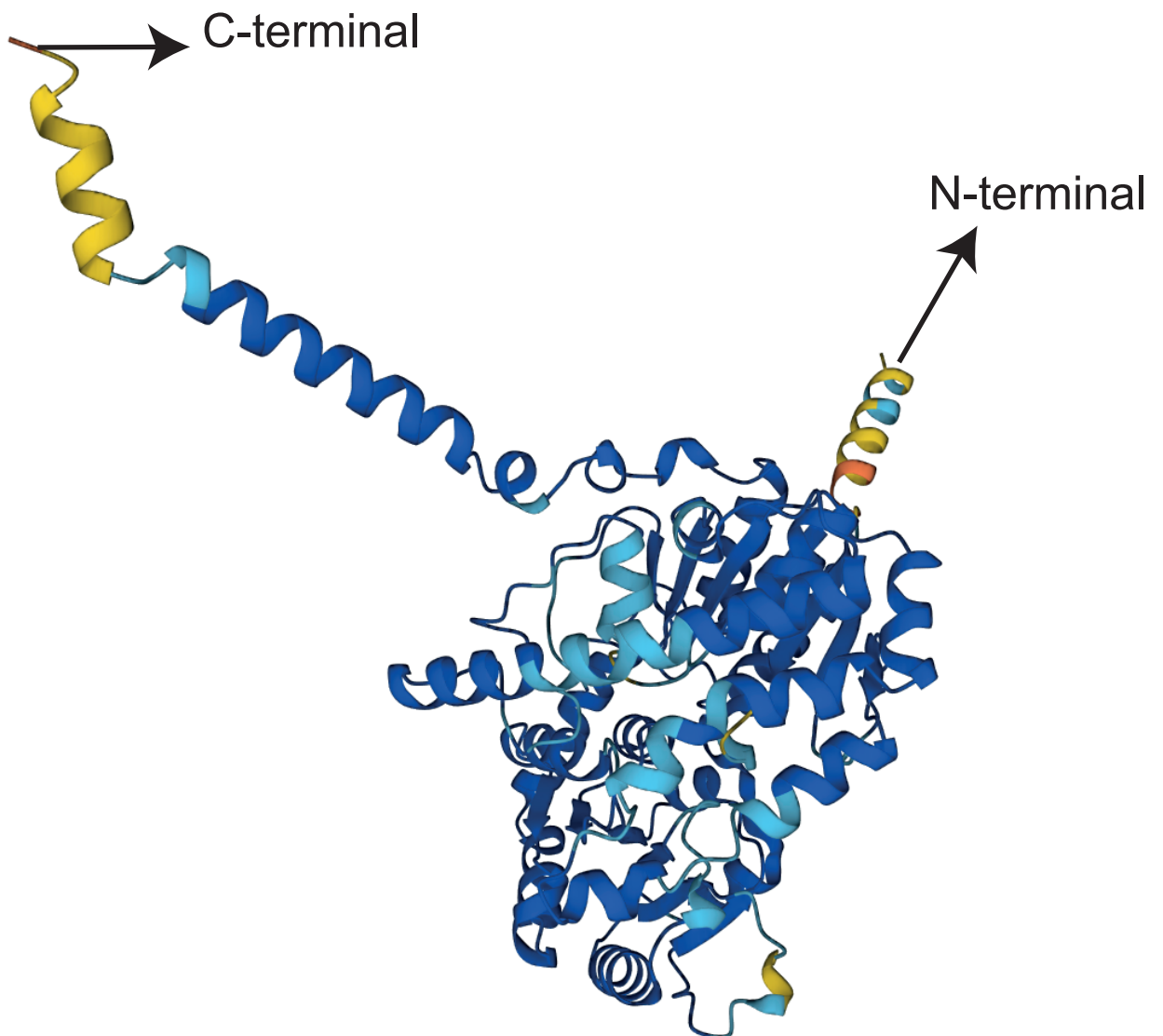
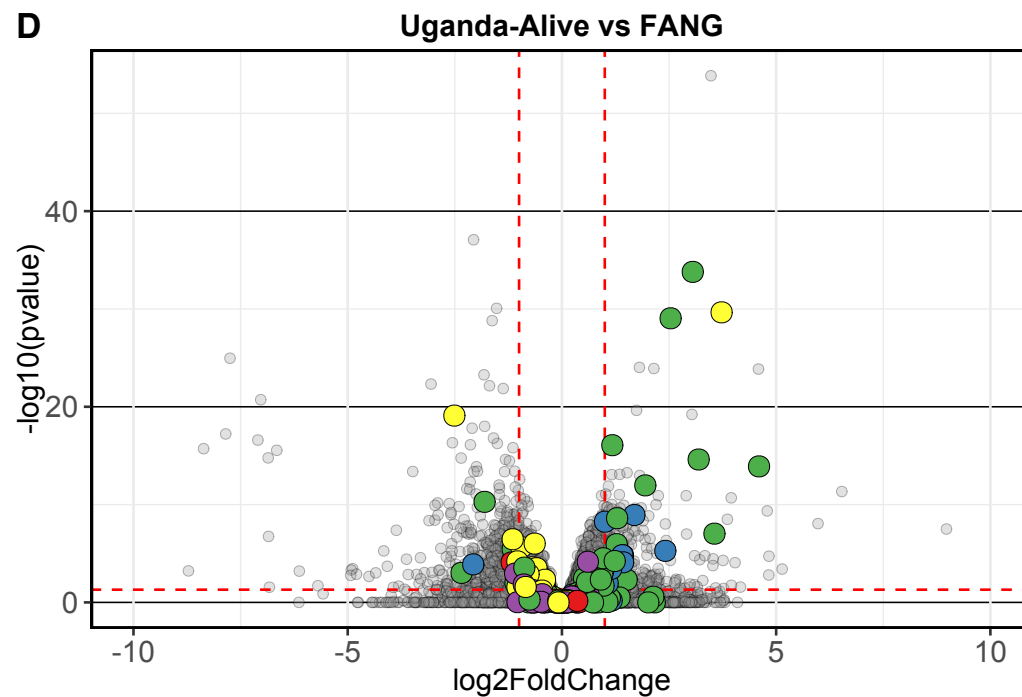
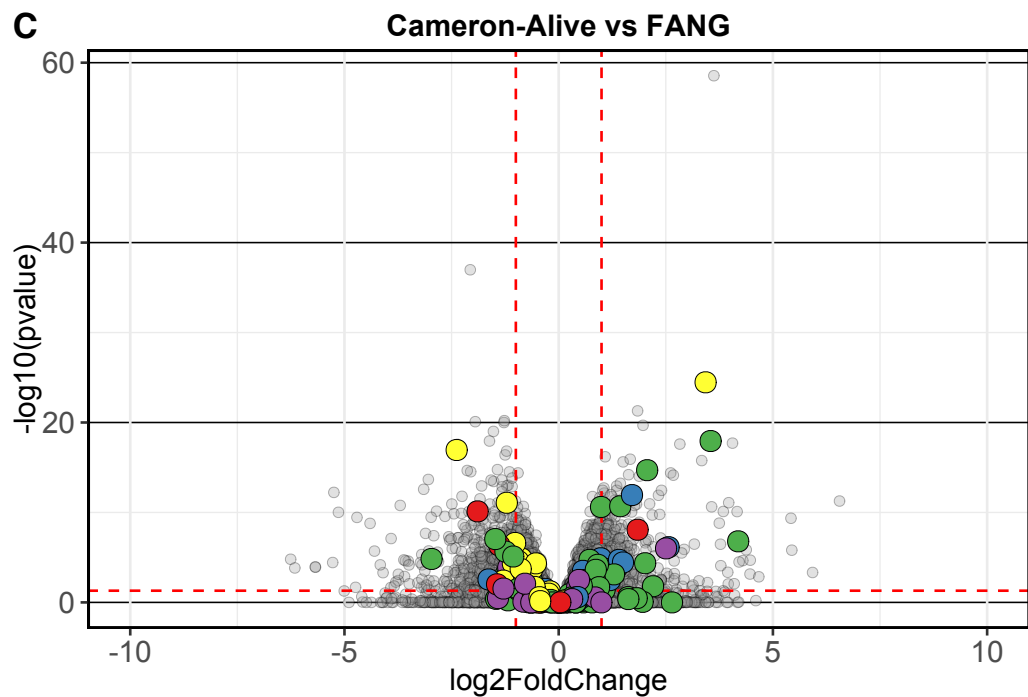
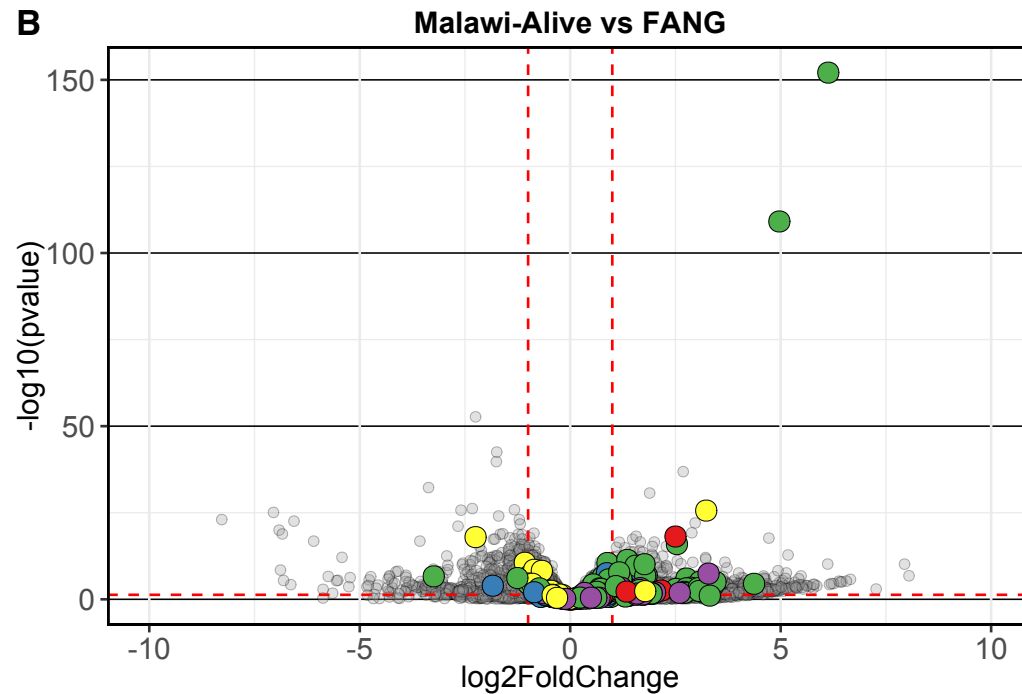
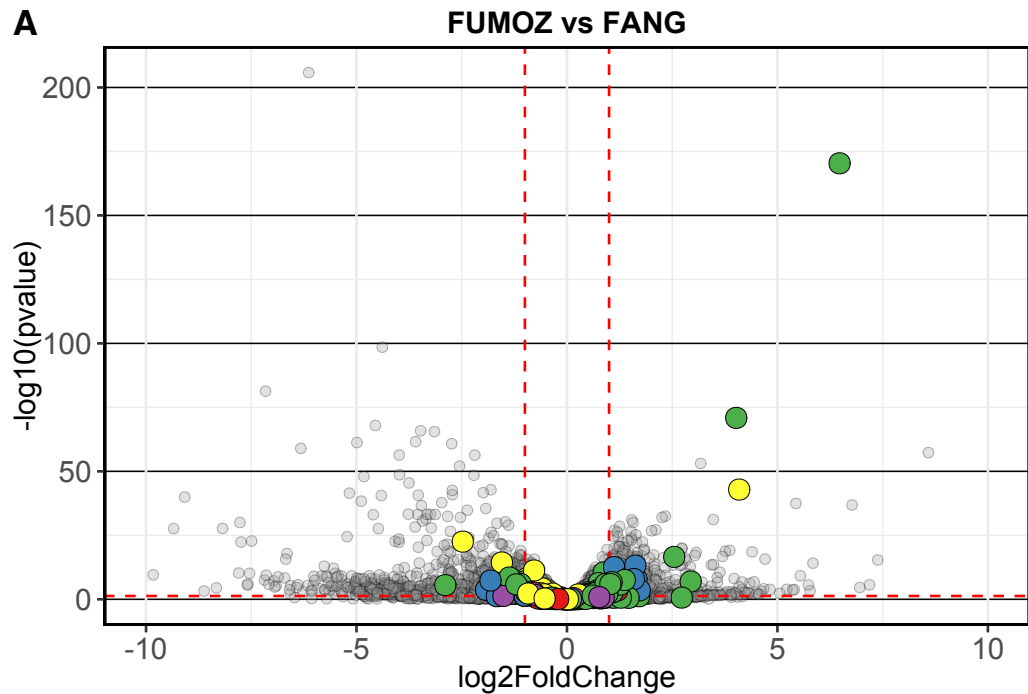
A**B**

Fig. S2. Protein function analysis of UGT50B8. *An funestus* UGT50B8 was selected for functional analysis because UGT50 is the only family that is found universally in all insect species. Protein functional analysis conducted by InterPro scan <http://www.ebi.ac.uk/interpro/search/sequence/> indicates that the protein pfam Id is PF00201 belonging to the UDP glycosyltransferases superfamily **(A)**. The functional analysis locates the UGTs conserved motif in the C- terminal domain. The majority of the protein is non-cytoplasmic resides in the endoplasmic reticulum (ER) while a small portion of the protein is cytoplasmic. The retention of the protein inside the ER is mediated by hydrophobic trans-membrane domain **(A)**. The protein structure was predicted using AlphaFold **(B)**. AlphaFold structures are colored using a per-residue confidence metric called pLDDT, which is scaled from 1 - 100. **Very high** (pLDDT > 90) **Confident** (90 > pLDDT > 70) **Low** (70 > pLDDT > 50) **Very low** (pLDDT < 50)

● ABCs ● COEs ● CYPs ● GSTs ● UGTs



● ABCs ● COEs ● CYPs ● GSTs ● UGTs

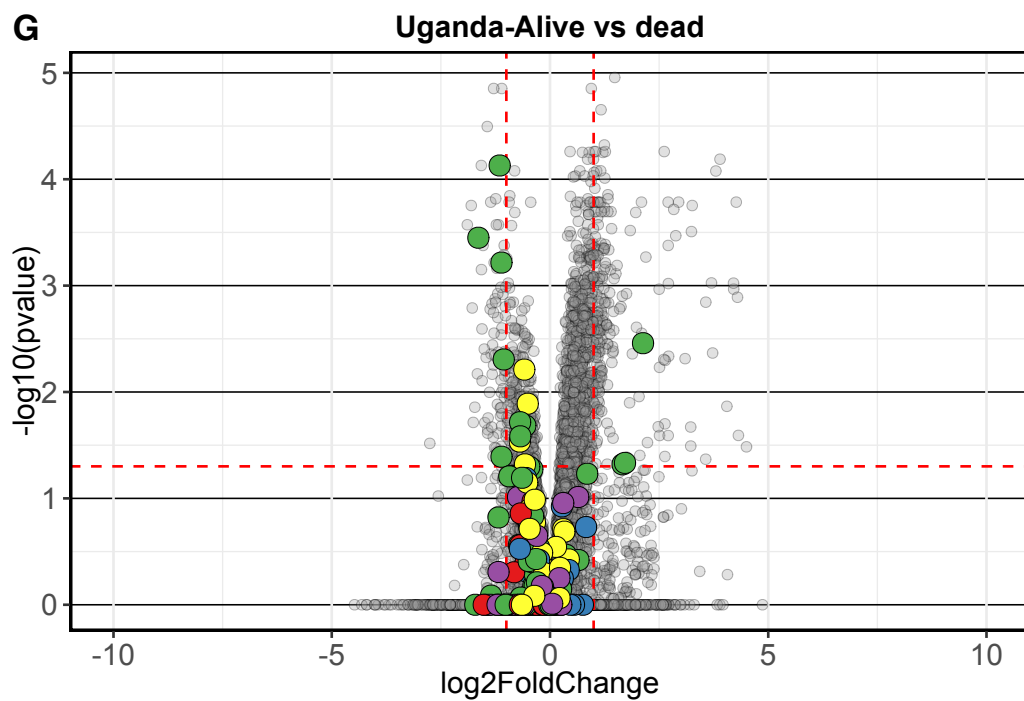
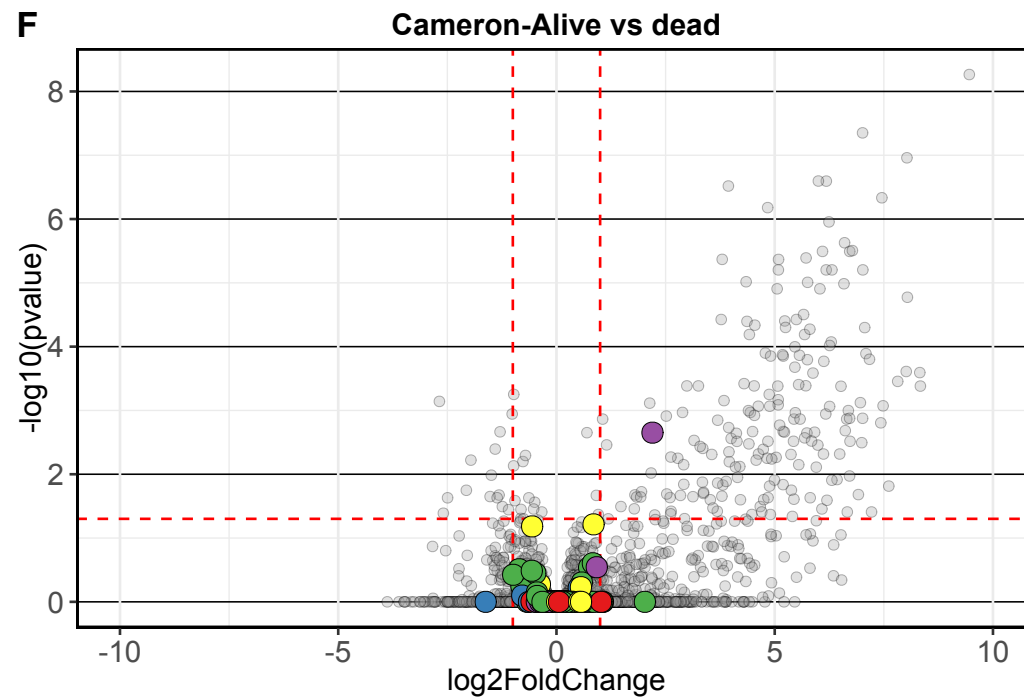
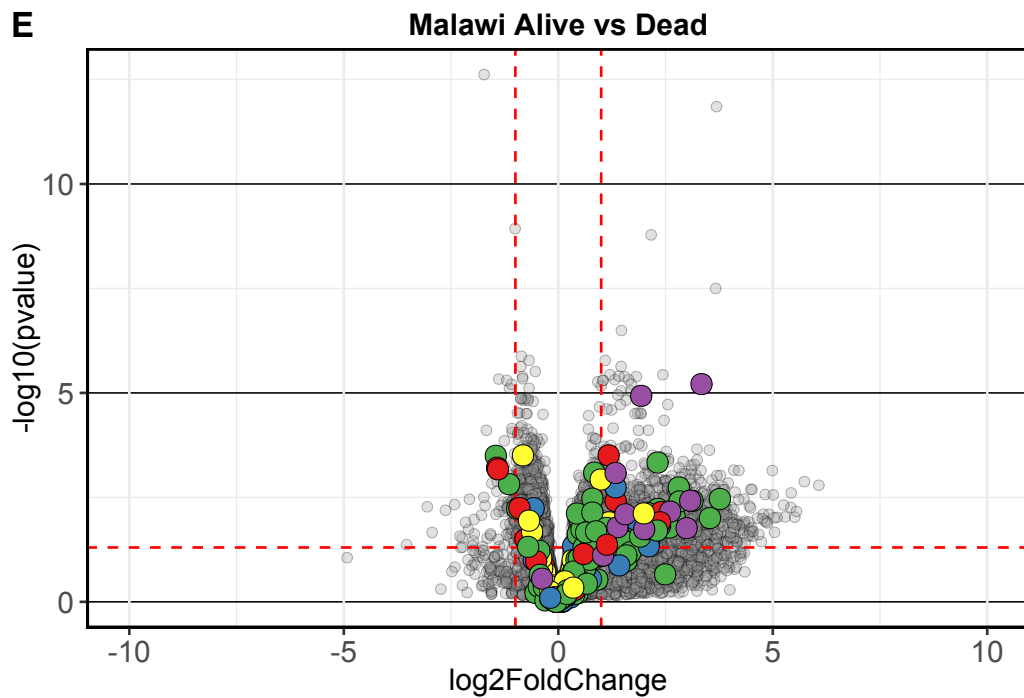
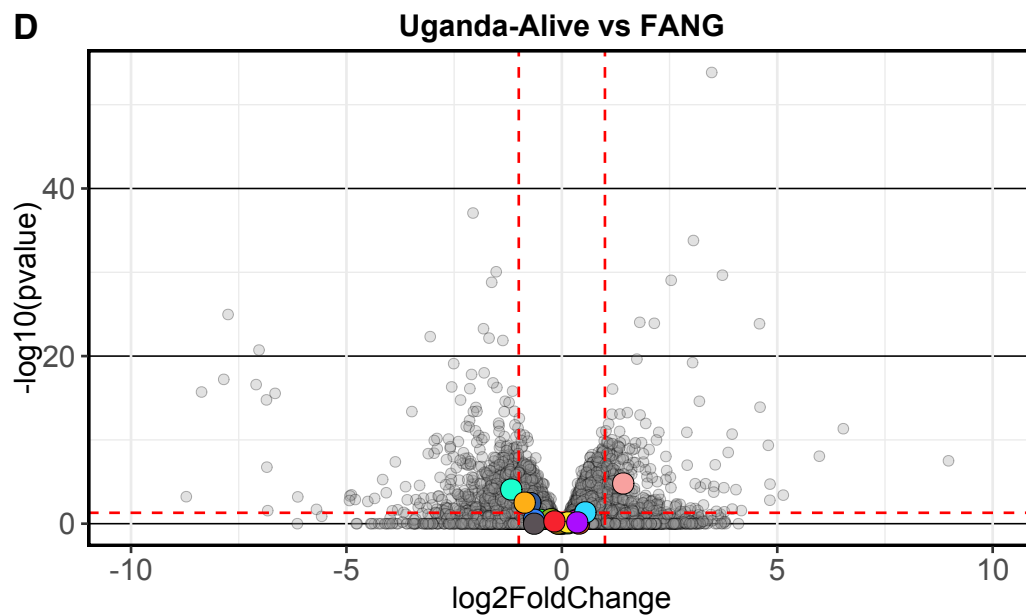
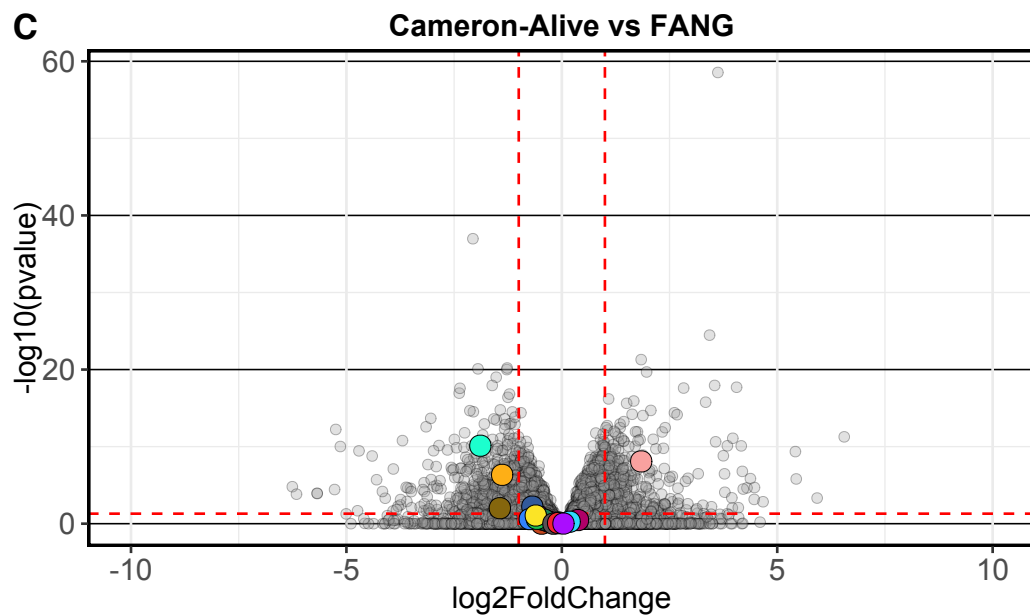
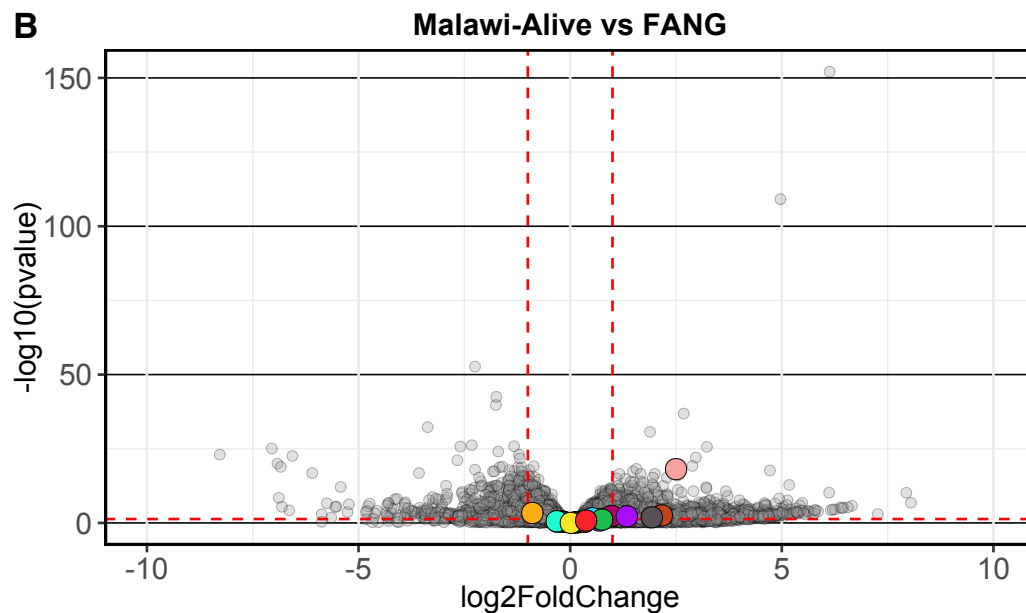
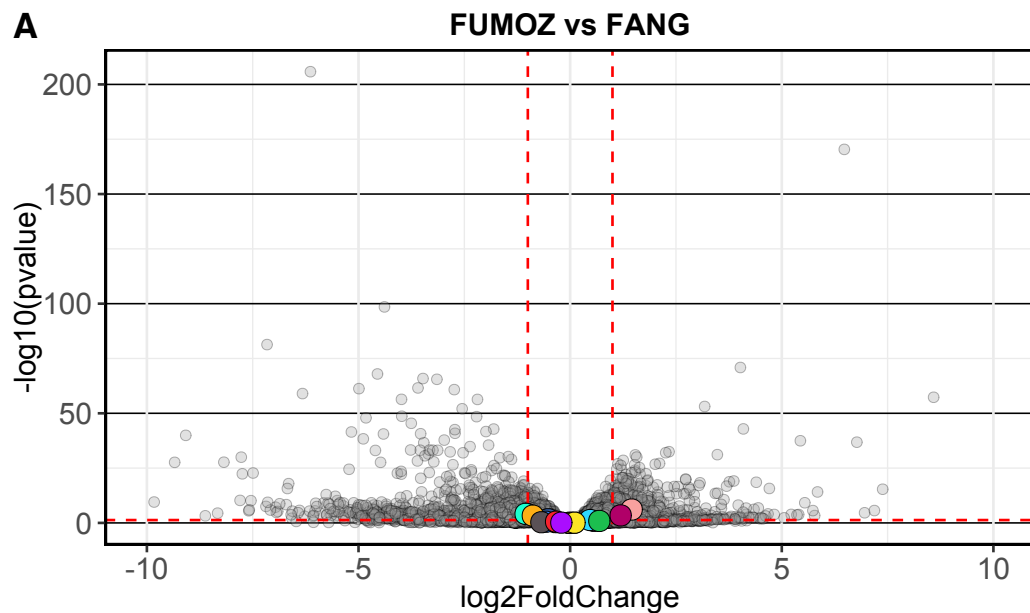


Fig. S3. Detoxification genes differential expression analysis in Malawi, Uganda, Cameroon and the resistant colony FUM0Z. The transcription profiles of the resistant colony FUM0Z, field-collected resistant mosquitoes from Malawi, Uganda and Cameroon was compared to the FANG transcription profile **(A-D)**. In addition transactional profiles of field-collected resistant population was compared to the transactional profile of a population of mosquitoes unexposed to permeation from the same same country **(E-G)**. Differentially regulated genes were defined as those with a corrected p-value threshold of < 0.05 and $\log_2(\text{fold change}) > 1$ indicated by the dotted red line in each plot. Differentially expressed detoxification genes were color coded according to the plot legend.



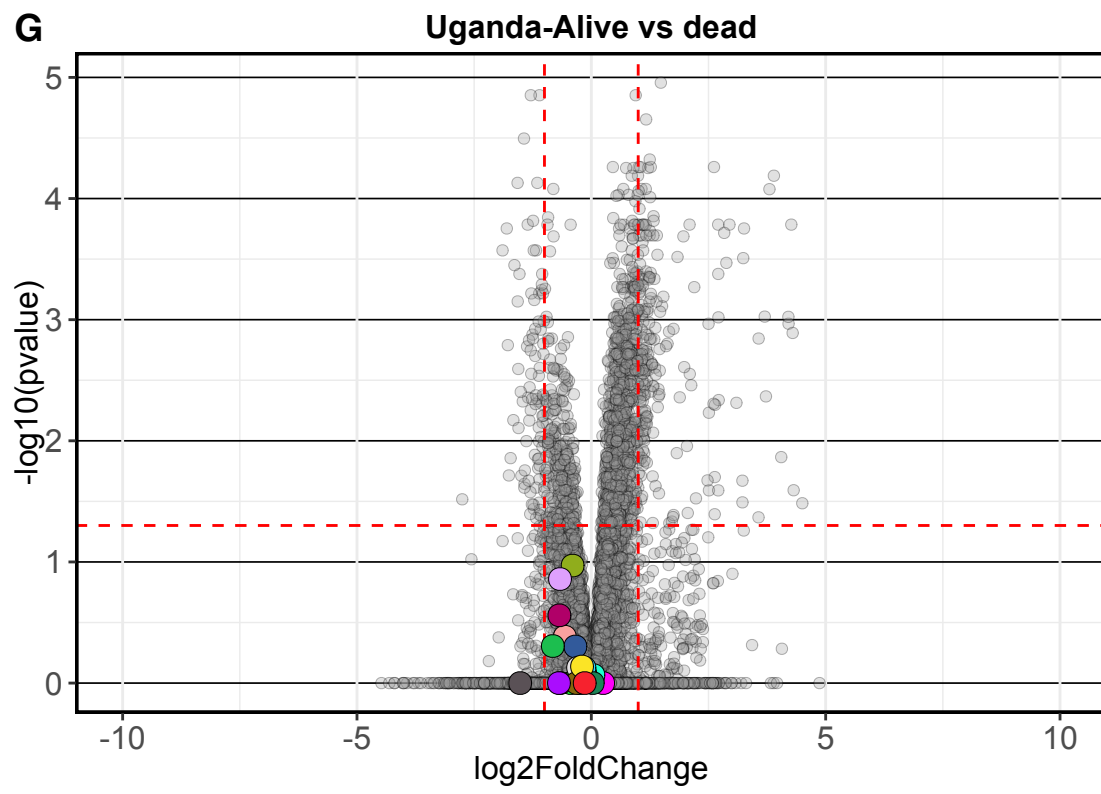
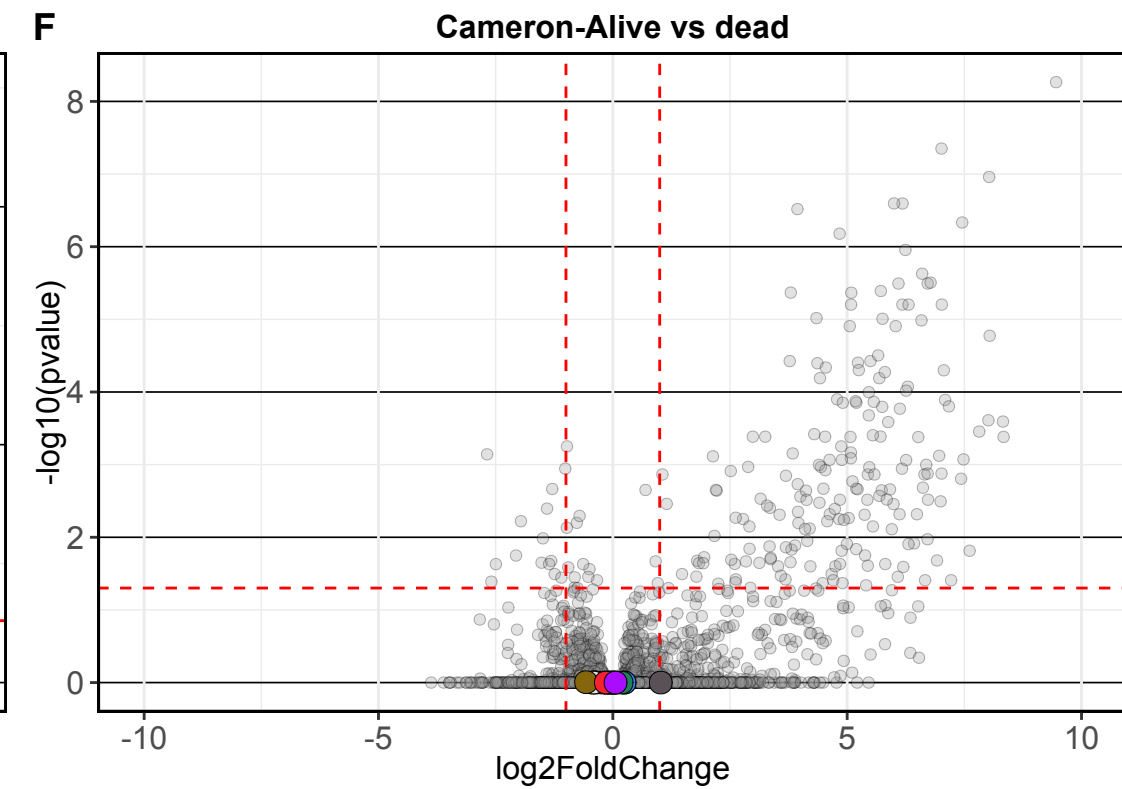
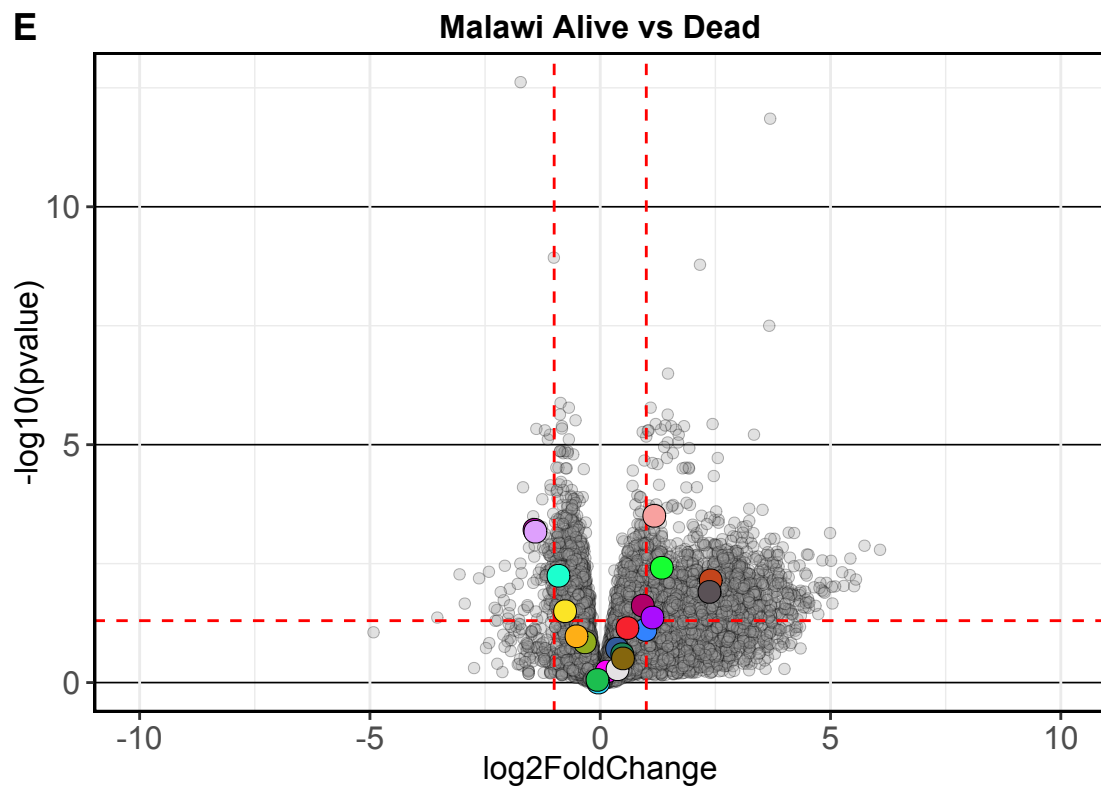


Fig. S4. UGTs differential expression in the FUMOS colony and in field-collected mosquitoes from Malawi, Uganda and Cameroon. The transcription profiles of the resistant colony FUMOS, field-collected resistant mosquitoes from Malawi, Uganda and Cameroon was compared to the FANG transcription profile (**A-D**). In addition transactional profiles of field-collected resistant population was compared to the transactional profile of a population of mosquitoes unexposed to permeation from the same same country (**E-G**). Differentially regulated genes were defined as those with a corrected p-value threshold of < 0.05 and $\log_2(\text{fold change}) > 1$ indicated by the dotted red line in each plot. Differentially expressed UGT genes were color coded according to the plot legend.

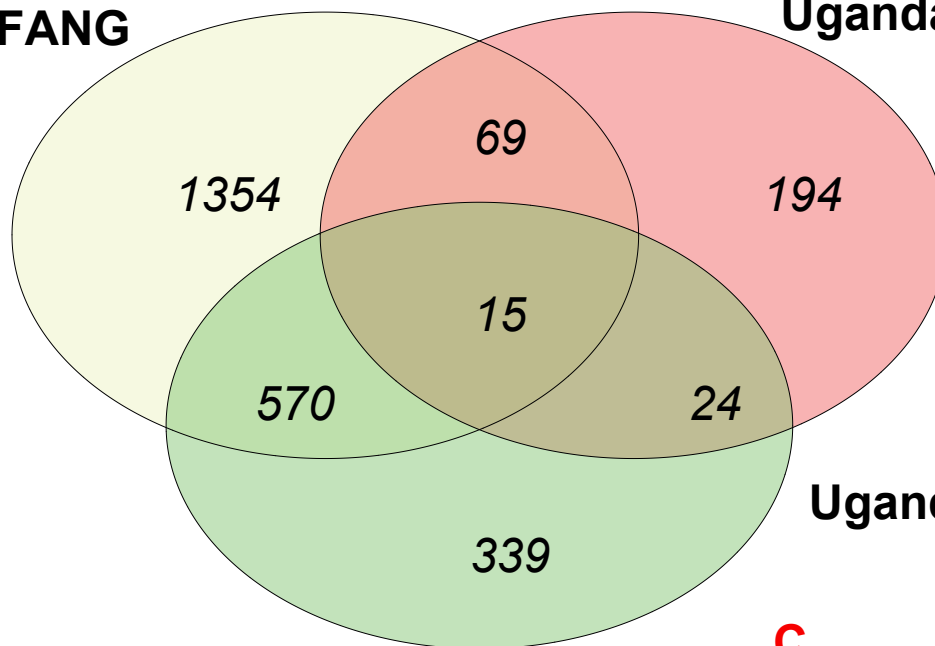
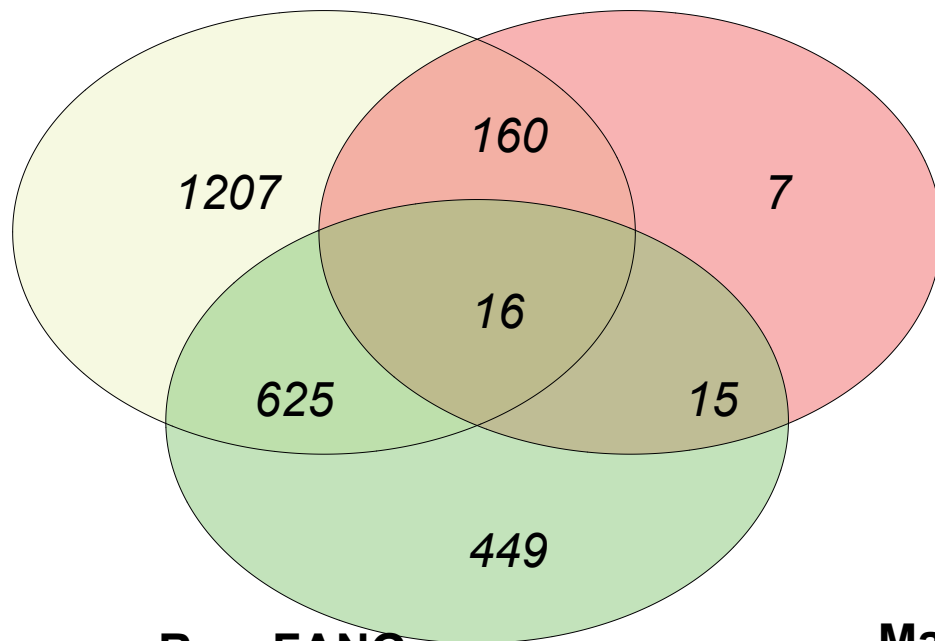
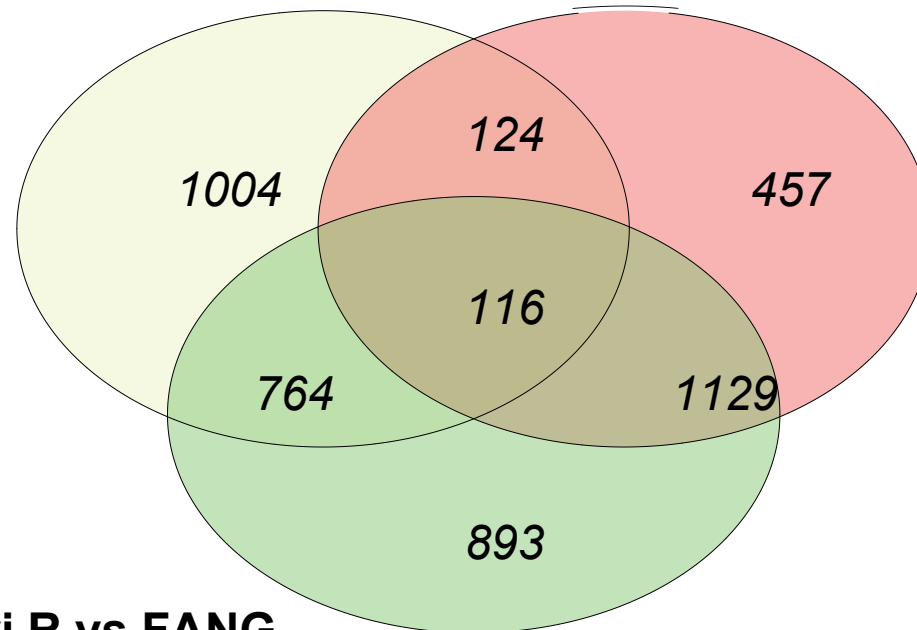
A**FUMOZ vs FANG****Uganda R vs UNEX****B****FUMOZ vs FANG****Cameroon R vs UNEX****C****FUMOZ vs FANG****Malawi R vs UNEX****Cameroon R vs FANG****Malawi R vs FANG**

Fig. S5. Venn diagrams illustrating the comparison of differentially expressed genes. The diagram displays the number of shared and unique differentially expressed genes between three different scenarios: 1) FUMOZ against FANG, 2) resistant field-collected population (R) against unexposed populations from the same country (UNEX), and 3) resistant field-collected population (R) against FANG. This comparison was conducted in Uganda (**A**), Cameroon (**B**), and Malawi (**C**).

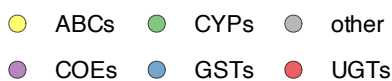
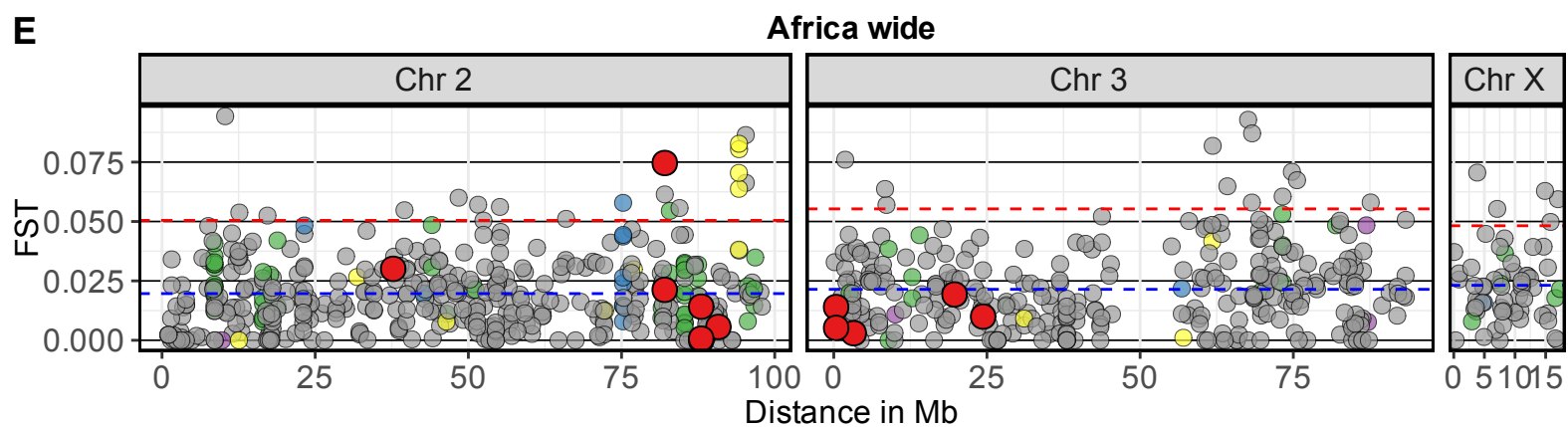
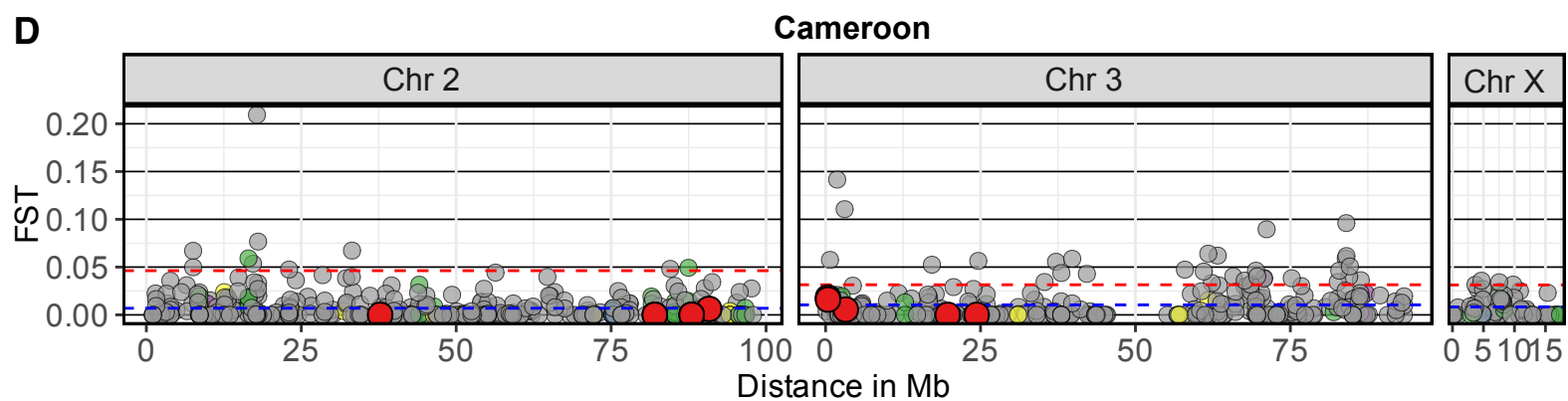
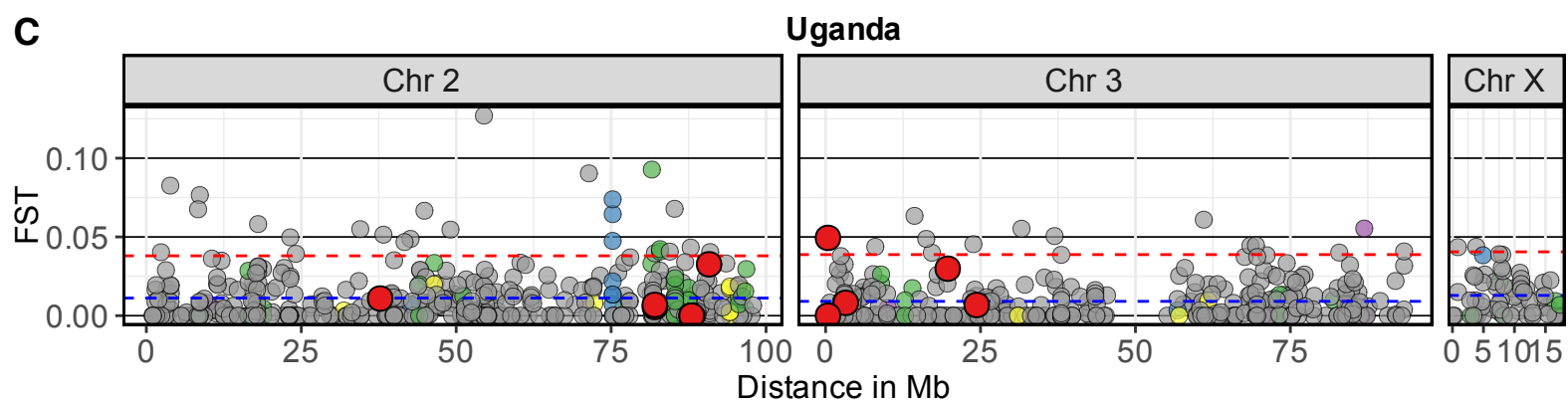
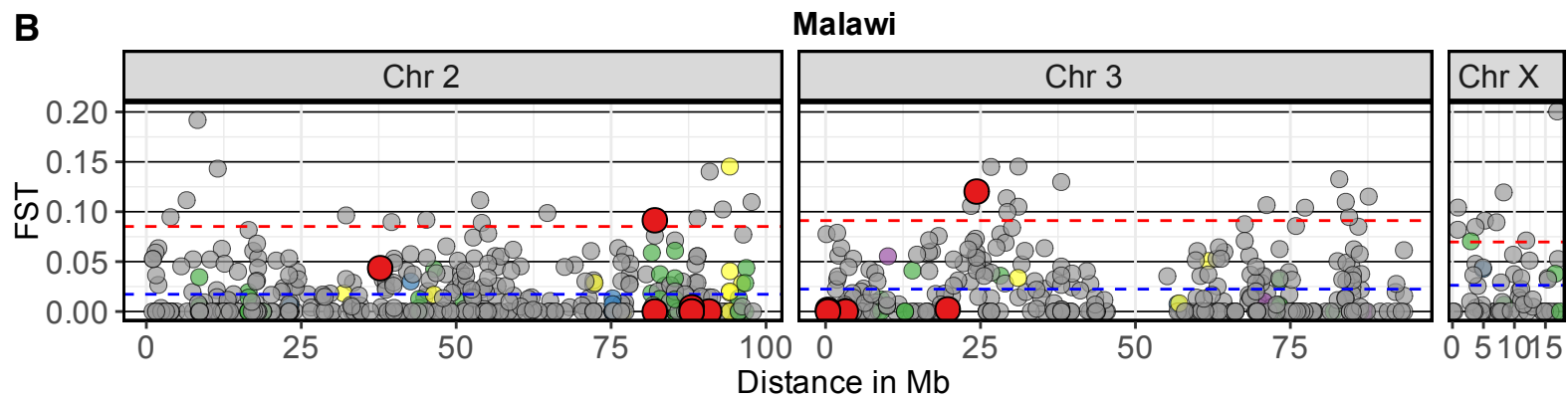
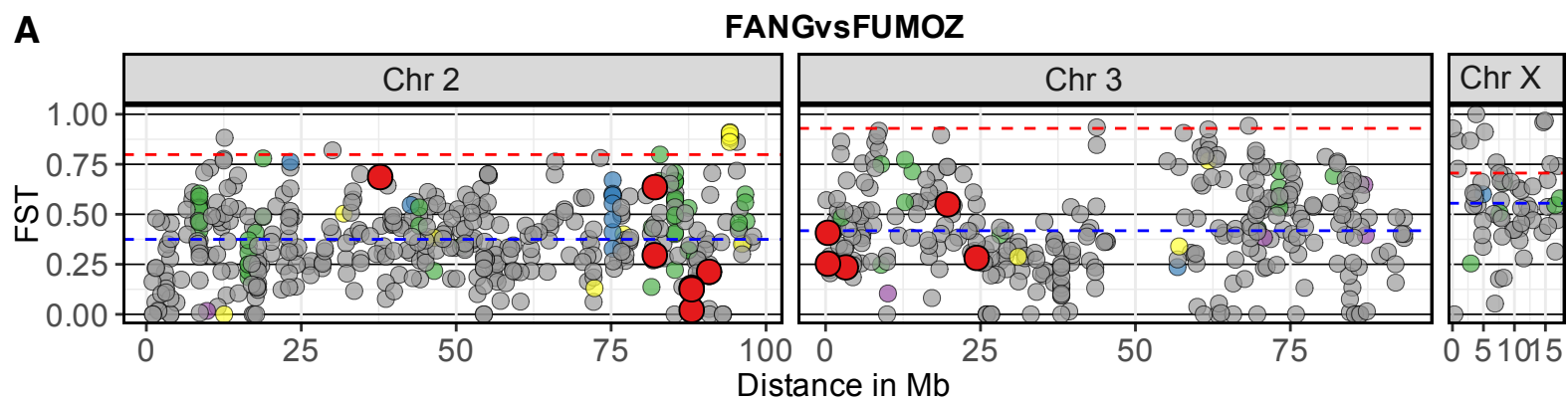
FANGvsFUMOZ

Fig. S6. Gene-wise F_{st} for all genes included in the targeted sequencing in all analyses between resistant and susceptible. The average and 0.95 quantiles of gene-wise F_{st} for each chromosome are represented by the blue horizontal line and the red horizontal line respectively. The genomic location of all genes in the targeted region (x-axis), represented by a circle, was plotted against the Gene-wise F_{st} (y-axis). Detoxification genes included by the targeted sequencing are color coded according to the plot legend.

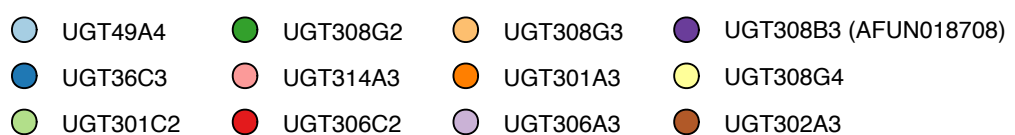
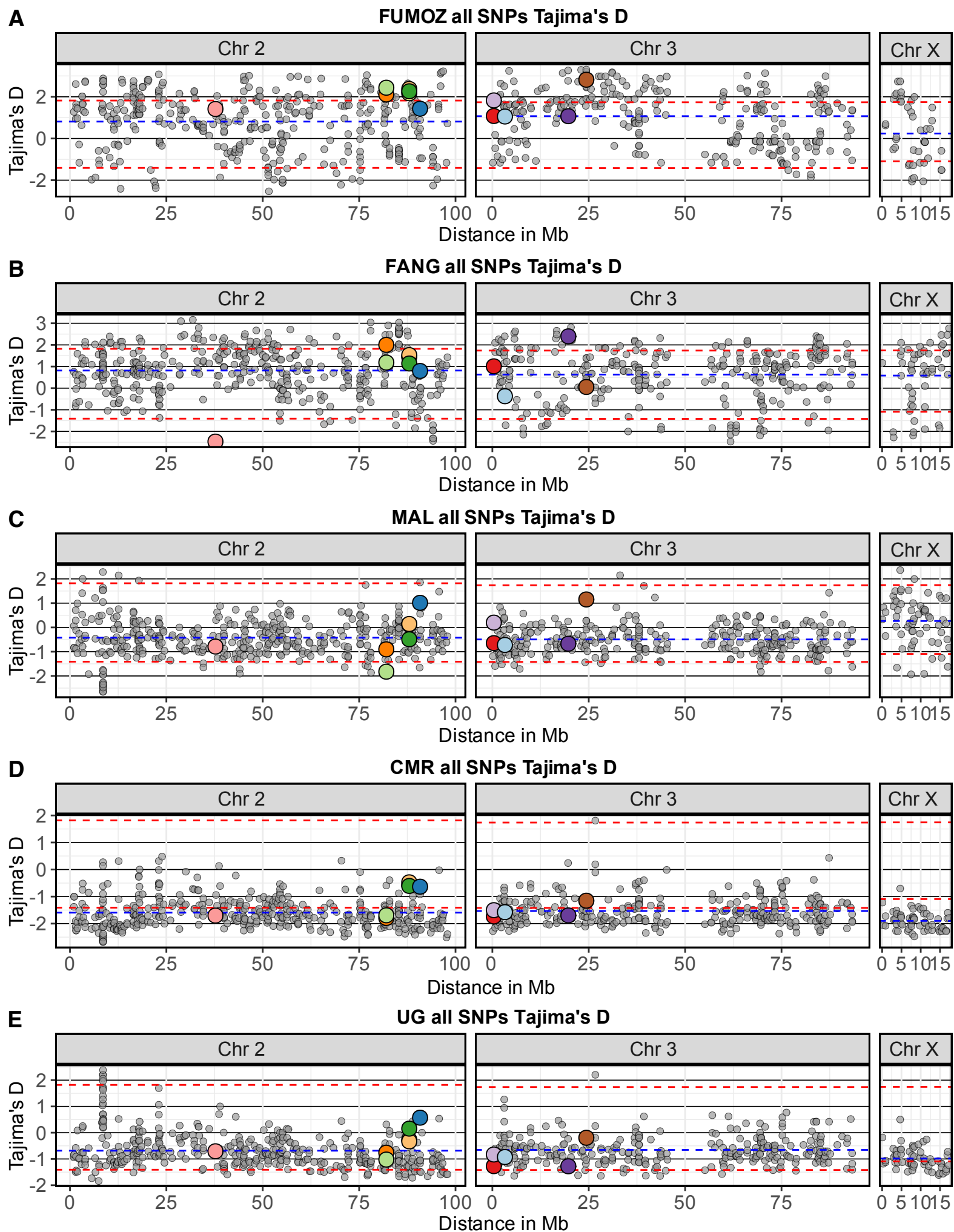


Fig. S7. Gene-wise Tajima's D for all genes included by the targeted enrichment sequencing. In each country plot, the average is represented by the blue horizontal line and the 0.5 quantiles and 0.95 quantiles are highlighted by the dotted red line. The genomic location of all genes in the targeted region (x-axis), represented by a circle, was plotted against the gene-wise Tajima's D (y-axis). UGT genes included in the targeted sequencing are color coded according to the plot legend.

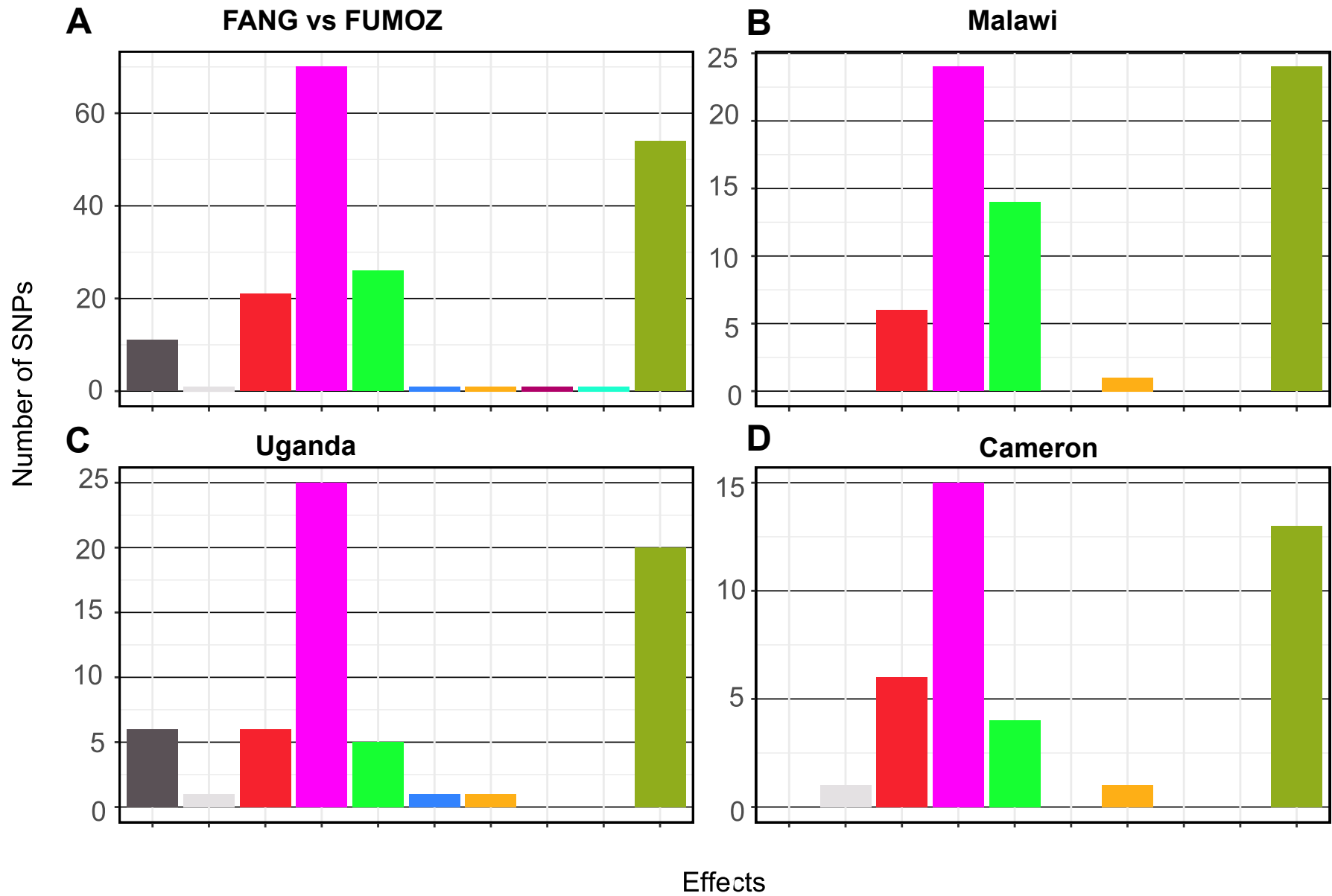
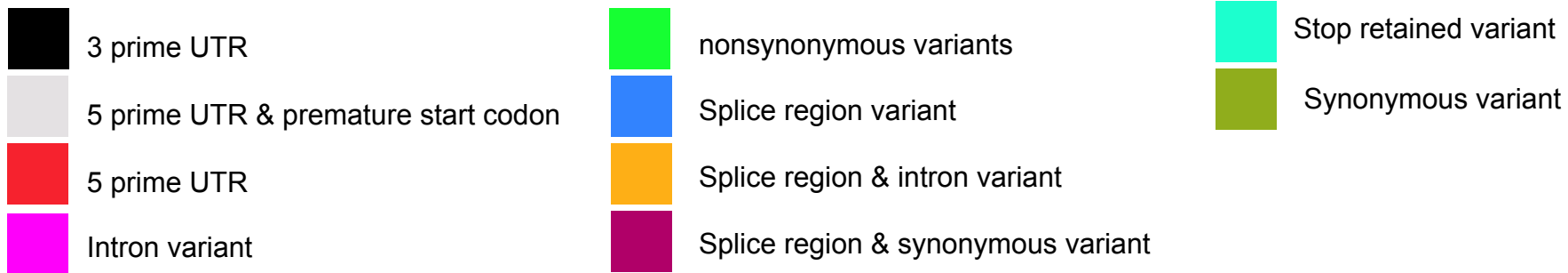
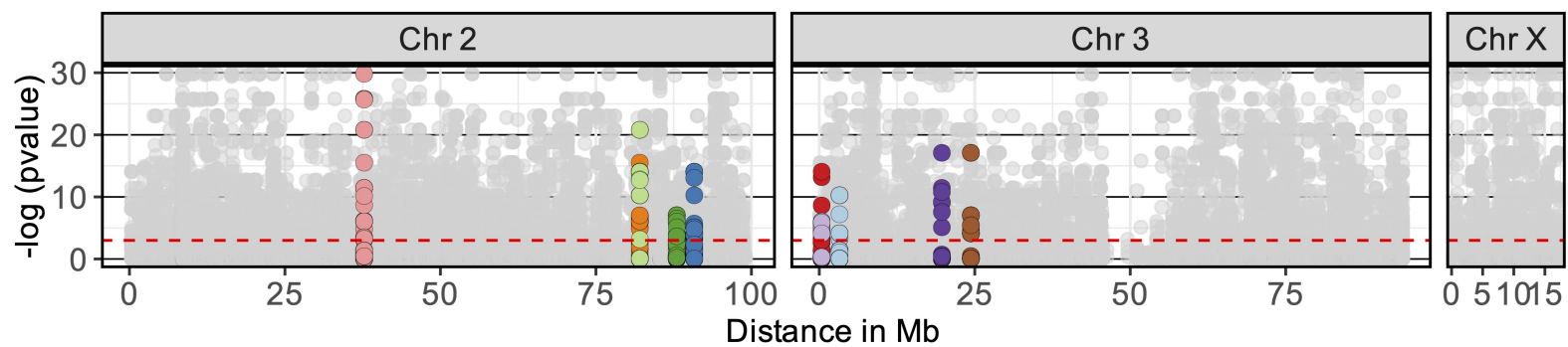
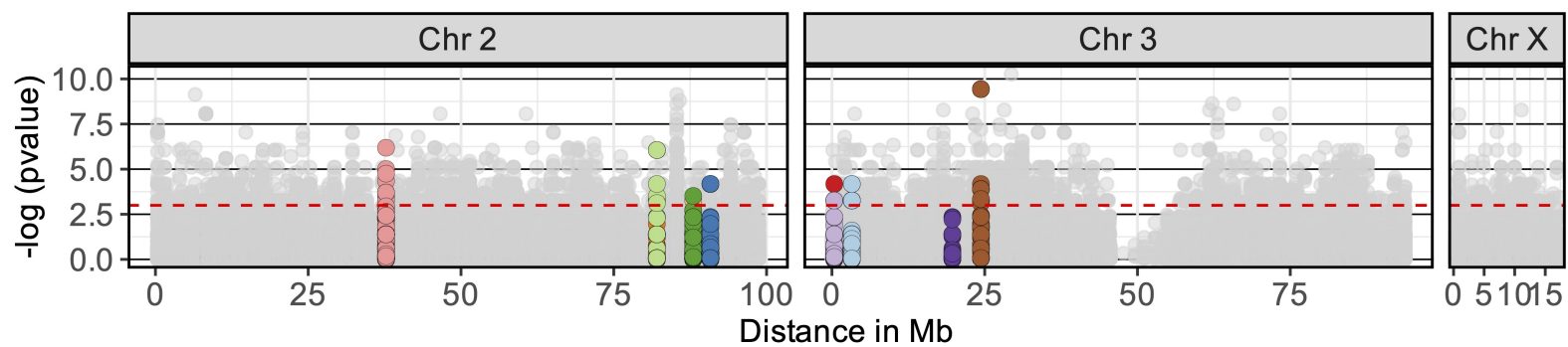


Fig. S8. Effect caused by UGT genes significantly differentiated SNPs. Most of the differentiated UGTs' SNPs in all analyses are intronic SNPs. We focused our study on UGT genes non-synonymous SNPs that are significantly differentiated. Details of the number of UGT genes SNPs in each comparison including nonsynonymous SNPs are included in (Table S7).

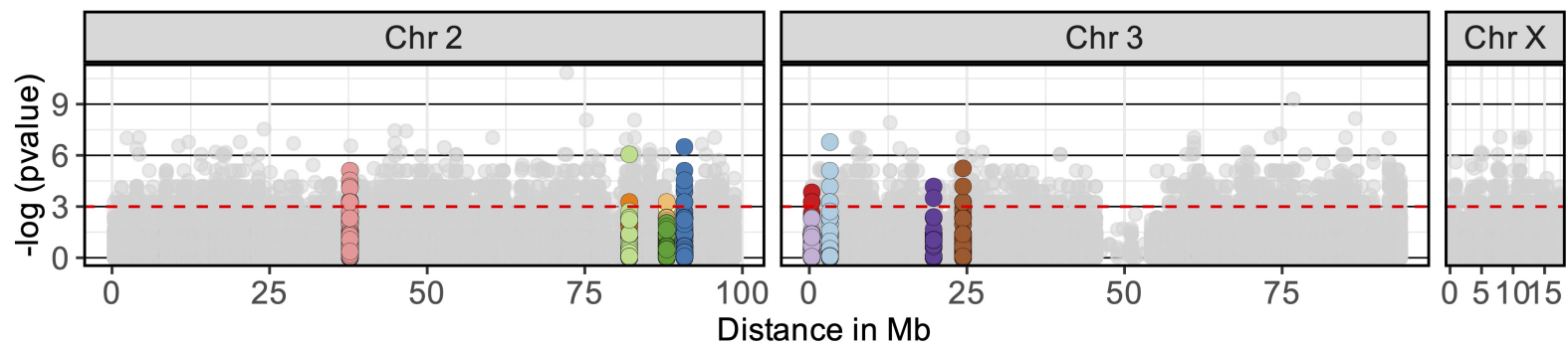
FANG and FUMOZ UGTs all SNPs



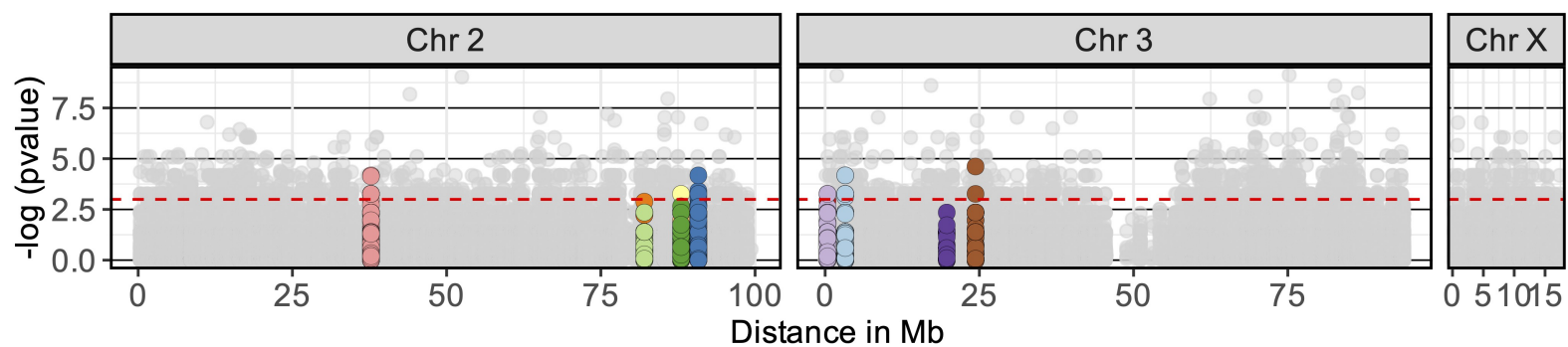
Malawi UGTs all SNPs



Uganda UGTs all SNPs



Cameron UGTs all SNPs



Africa_wide UGTs all SNPs

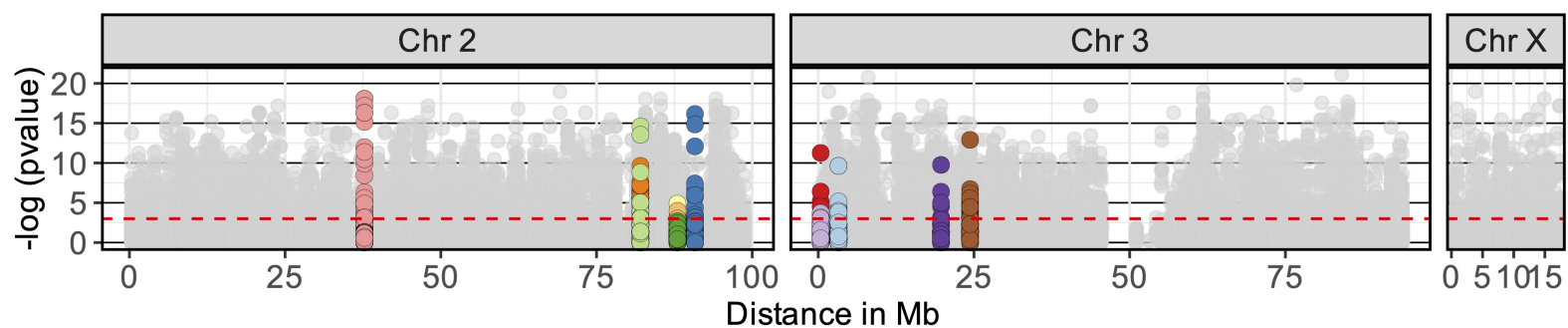
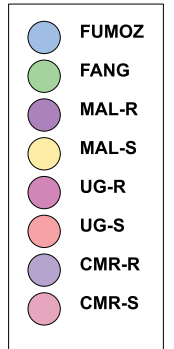
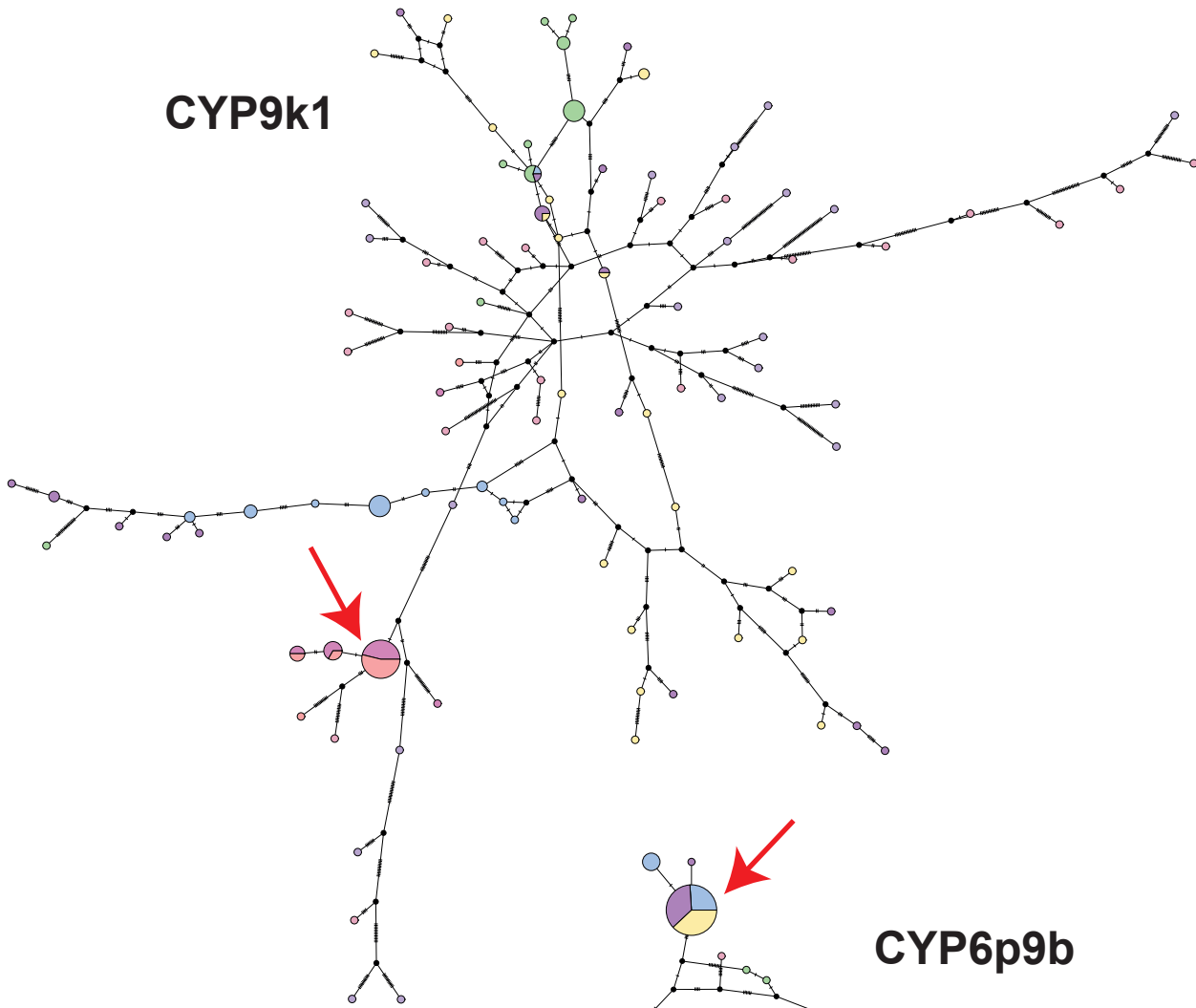
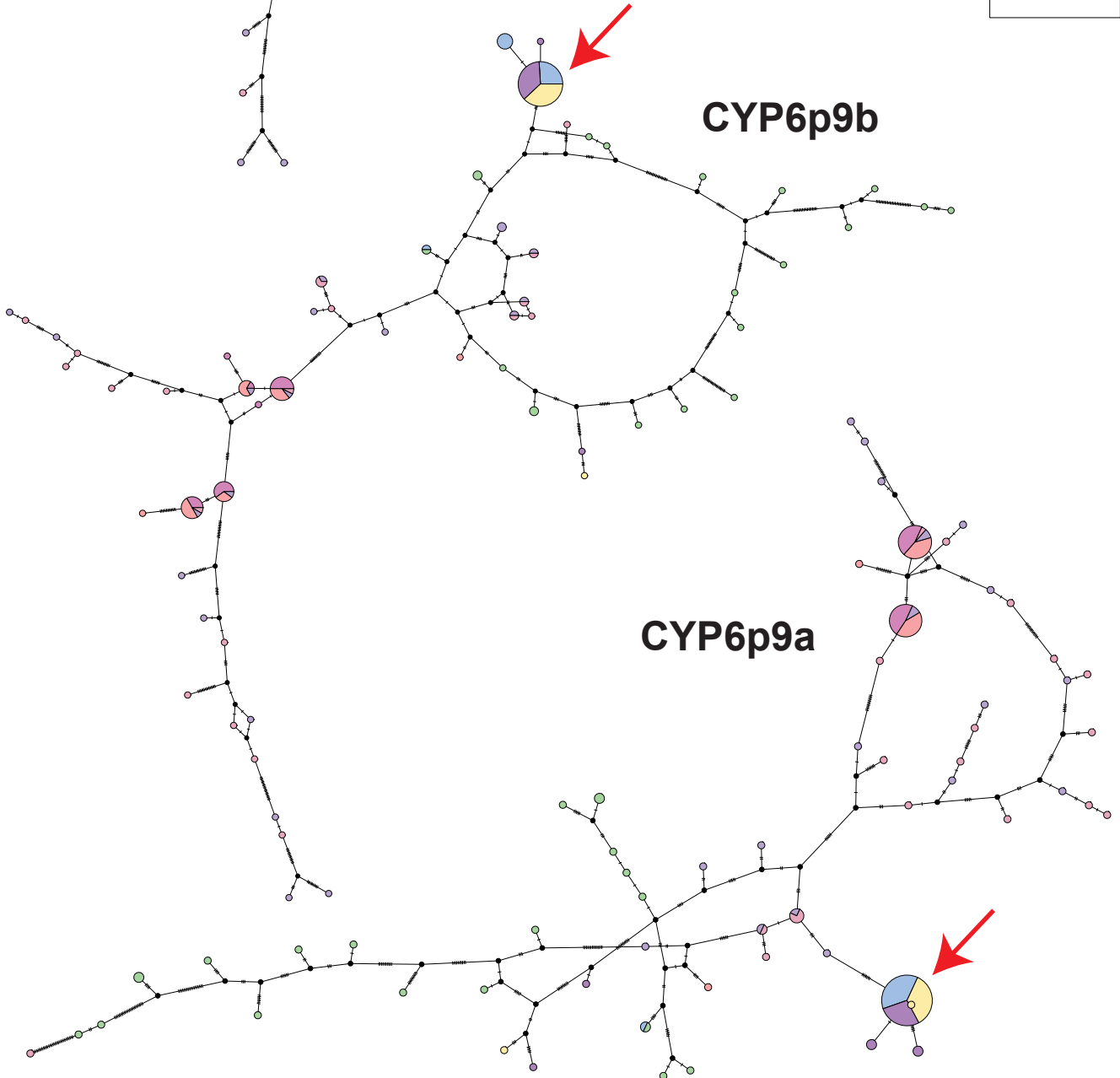


Fig. S9. UGT genes SNPs divergence between laboratory colonies and in each country. SNPs located within UGT gene are color-coded according to the plot legend. The blue dotted line indicates the average value in each comparison and the red dotted line indicates the significance level (p-value = 0.05).

CYP9k1



CYP6p9b



CYP6p9a

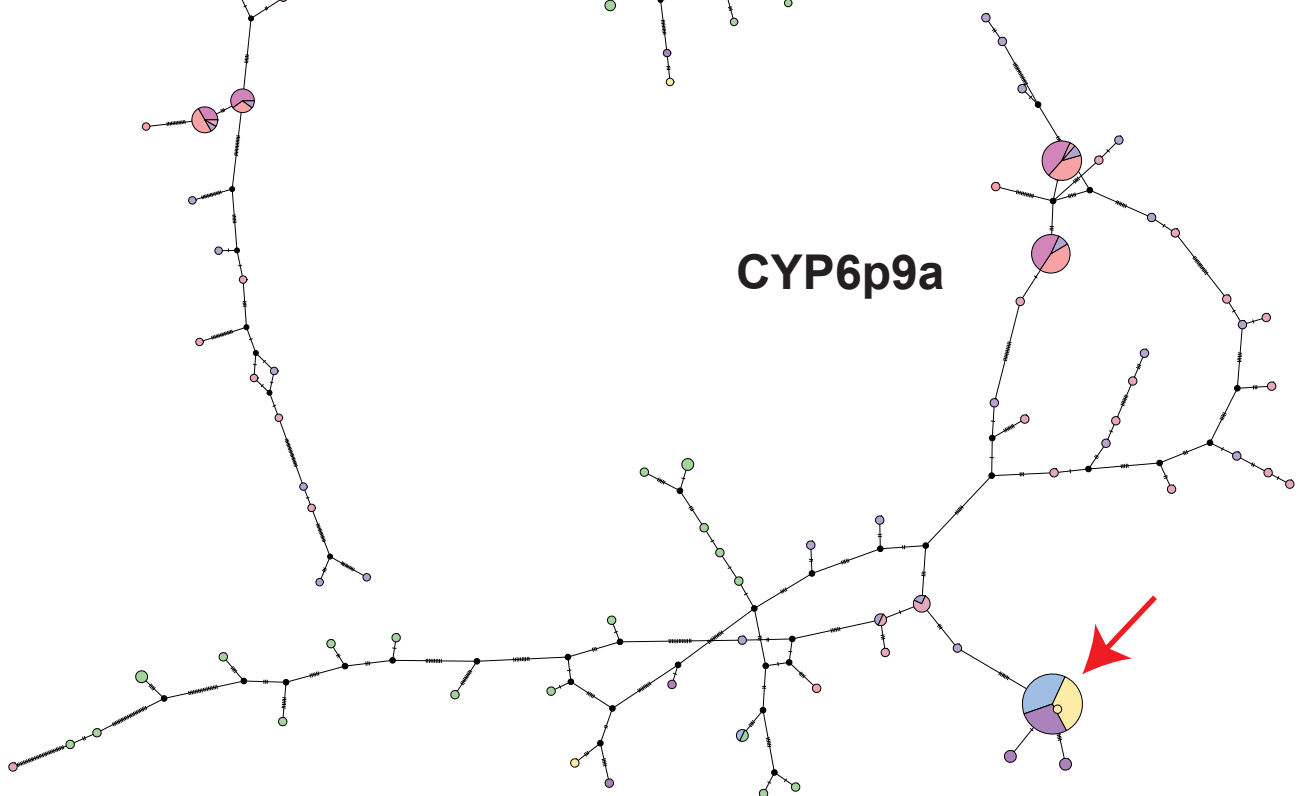


Fig. S10. Hapotype network for *cyp6p9a*, *cyp6p9b* and *cyp9k1*. TCS haplotype network for *cyp9k1* reveal a directions selection in Uganda supporting previous findings **(A)**. While TCS haplotype network for *cyp6p9a* and *b* support previous observation of directional selection in southern Africa represented by Malawi and the resistant FUMOS colony originally from Mozambique **(B and C)**. The red arrow highlights the dominant haplotype node.

1. AfUGT314A3
2. AfUGT308G4
3. AfUGT308G3
4. AfUGT308G2
5. AfUGT306C2
6. AfUGT306A3
7. AfUGT302A3
8. AfUGT301C2
9. AfUGT301A3
10. AfUGT49A4
11. AfUGT36C3

1 10 20 30 40 50 60 70 80 90
MATNTLVLLIAVIALRLTEAAKILAI FVPVLPKQHQLVYRPLIEELANRGHEIVLVTTDPIDTTDRITNGSLDRIEQIDLSFAYQLPVL EQLGQD
MRYQAMKAVLPVILMLAGVSSQAANILFMSSFPSPSHHIWLRPLIHEMGKRGHNVTVLSADVEKPPANVTYIHL ENLYSTMYNNTVKLDFFFELAN
MKAVLSVVLVLAGVSSQAANILFMSSGVPSPSHYIWLRLPMYEMGKRGHNVTVLSADVEKPPANVTYIHL ENFYSTMYNTSMREKLDFFELANQSP
MKAVLSVVLVLAGVSSQAANILFMSSGVPSPSHYIWLRLPMYEMGKRGHNVTVLSADLEKPPANVTYIHL ENLYSAMYNTSMREKLDFFELANQSP
MKITQGVAVAVAVLCCCYAQL EAYRILCIPSSGRSHVLVGGALLKGLAERGHVDTMVSPLYKLSKSVPNYREIVIQKVDLGMQTKDFLQKNEGN
MMKLTFNIALTLTLMLLTSHTESA KIVCVFPTASKSHVLGAQALLKELAFRGHEVTMVSAPFLKRPPNNYRDVYVPIEDAFSTIMNDFMQGGR
MLAYTMKP IKSSVGAVAFFGLIFMMAQGI VEPAKILSIFPTMSKSHWILGSSLMKELAQNGHEVTVISPFPLKNAPKNYRHVDVPYRTQMFEDI
MWLRFVFTGCLFLLGSVLSSADGYKILFLVFPFGPSHWLMLKHFIRELTERQHQVTCITAFPFGEKLQNYDEILIDPPYPIRETFFVEGLFASS
MHLRVL SYIVGCVFAFTALCPRAEC GKILFLVFPFAPSHWLWIEHFVKELLSRGHEVTAITNFAMKDPHRNYTEVLIDPPYDIPYYFPVSDIYE
MHLVALVSFLAASLQTSSGARILSVNVFPGRSHMMIGSILEELDRGHVTVITNYPRKKPHPNLTEIVIDPIYDFWGKSVKVDSDLYDLTDIS
MIMLRTVLLVALALSPCAQTYKVLGLFPHPGQSHFHFQPVKGLAEAGHEVTVVSHFPEKNPPNNYRDMPLEGMTSLSDSVSEFLFEYRPGFG

1. AfUGT314A3
2. AfUGT308G4
3. AfUGT308G3
4. AfUGT308G2
5. AfUGT306C2
6. AfUGT306A3
7. AfUGT302A3
8. AfUGT301C2
9. AfUGT301A3
10. AfUGT49A4
11. AfUGT36C3

100 110 120 130 140 150 160 170 180
GLDGRDMLRNI FDMVRTISA EELQHPAMQKLINSANPRACDAGKCNHFDVVMVEWVGVTLMNAFAEHFRAPLVGIANGGAFINAH EALGNPNHP
QSPASILQQFDKFLTLCEAAIKSEGLHLLRYPKEFKFDL FVSDFMIGPCIPAIMHRFKDVPYIPSTPYNAPSTSVTVLGSFAYPALVPHNV
AKMLQLFDEFGNLNCEAAIKSEGLHFLRYPKEFKFDL FVSDFMIGPCIPAIMHRFKDLPFIPSTPYNAPSTSATVLGSFAYPGLVPHNVFDA
AKTLQIFNEIGLNLCEAAIKSEGLHFLRYPKEFKFDL FVSDYMIIGPCIPAIMHRFKDVPYIPSTPYNAPSTSAAVLGSFAYPGLVPHNMFDA
SMASLVYLFQSQMRTAEMAL EDAKVLALKHEHFDLVI VGVFVADFLGLGPHFNAPT VVLSAGMTKPTAD FVGNPRAIAAVPHMLGGKGTMD
NMLKLFPKIVRAAQDSSNVTINAP E FIRLAHEEQFDLALVGGFMNSFIIGV GELFRCP TVLYFSASGSLTNVVGNP AEVAAVPHMLLGP RNPM
MDEVFDKVDSSIVEKLIQMGGMHEITNTT LASPEVQSLLKSDETFDLVLVLEIFLNDALGFADR FNCPVGMSTFGASSWVNSLTGSPQPLSY
QKSD FVKLFMYWDLGFNTSRHGLESENVRQFIRRTDLEFDL IAEQFFQESWLMFAHIYNAPIVTISTYGYSDFFDRIMGLQTPWSFVPHMVL S
SKYNSDL SNLFLYWRVGLSTTQYAL EDENVQQFIEQDDTDFDVI ISEQFYQEAFLMFAHKYRAPIVTLCTLGHANHIDQAMGLVTPWSFVPHPV
IHQMLMDFLYPLGLQTA EYAYTRDNVMDFLRNDQSQFDLLAEQFYQESYLMLAHKYKVP IVSIGTFGFAQYMGPMMLMNAWVSHVPHFELPFT
HFMEFFMLYTWGKEACANALNSTAIQTILDSKI QYDLILVEQFNSDCMLGIAHQKAPF IGLSSCPLMPWHYDRVGNPTLPAYVPALFMGYTER

1. AfUGT314A3
2. AfUGT308G4
3. AfUGT308G3
4. AfUGT308G2
5. AfUGT306C2
6. AfUGT306A3
7. AfUGT302A3
8. AfUGT301C2
9. AfUGT301A3
10. AfUGT49A4
11. AfUGT36C3

190 200 210 220 230 240 250 260 270 280
IGYPSIFMPF EDLSLLQRISSVLFTVWYRFYYYTEEIPAQNLIAQNNFGEQISDLRQIEQNADLLL INAHQQLGNVRPVGP TTIHLGGIHQSR
FDAPESMSFVQRVKNFYFDMYEMNLHETVLHPESDRIVRKLYPDAPSTYTFYKNVRLSLANVNPVIQYKEPMMPSMIPVGG LQIMP SKPLPDDL
PESMSFVQRVKNFYFDMYEMNLHETFMHP EADKIVRKLYPDAPSTYTFYKNVRLSLANVNPVIQYKEPMMPSMIPVGG LQIMP SKPLPDDL RKV
PESMSFVQRVKNFYFDMYEMNLHETVLHPESDRIVRKLYPDAPSTYTFYKNVRLSLANVNPVIQYKEPMMPSMIPVGG LQIMP SKPLPDDL RKV
FFGRVKNFLFAGVENVLCAVSEYVQTSYEHHP PDRYPTFQEVRRNVSLVLLNTHFSQSTPRPYLPNI VEVGG LQIKAKPDPLPEDIREWLDE
TFFDRVTNTILHGMEKIMLMYIRHKELTYYE SNFPAEEGFRSYDEAMRNVSLVMLNTYFSQYVPRPYLPNMVEVGGIQINLRPEAMPQDLQQFL
VPHMSSFTDKMNFWRQLGNVLF TAFDET VLSFMCDP IQQMYKKYFPNASRSLNQMRNGVSLV LINSHFSLSPRPYLPN LIEVGGFHVNRK
YEDDMNIFQRAYNVI LSSFD SVYRRYKYL PQADQLAREVFGELAKEPDSLPSITEL EQSISVILVNSHPILS IARPSIRGLVDIGGAHIRAPKP
LLLSDDMTF SQRCYNFWISLVDL IIRQFYIIPQKNLAQIH FARI EGPELMPSIQDLEKSI SVILVNSHL STSAPRPTIPGLVNVAGAHIKPAK
DHMSLYQRAYNSFVSWYELLLRSWY YLPEQQALADRHSFLPGPLPRLSDLERQVSVILLNSY PPLTSTRAKVPGLVQVGG LHIKEPKRLPDDL
MDFSQR LANWITVQSFKTLYSWFNDAAANKLLRARFGE GVI PDVRELQQR TSMMFVNQHYSLSGAKPLSPAVLEIGGIHIRD FKPLEADLQQLL

1. AfUGT314A3
2. AfUGT308G4
3. AfUGT308G3
4. AfUGT308G2
5. AfUGT306C2
6. AfUGT306A3
7. AfUGT302A3
8. AfUGT301C2
9. AfUGT301A3
10. AfUGT49A4
11. AfUGT36C3



1. AfUGT314A3
2. AfUGT308G4
3. AfUGT308G3
4. AfUGT308G2
5. AfUGT306C2
6. AfUGT306A3
7. AfUGT302A3
8. AfUGT301C2
9. AfUGT301A3
10. AfUGT49A4
11. AfUGT36C3



1. AfUGT314A3
2. AfUGT308G4
3. AfUGT308G3
4. AfUGT308G2
5. AfUGT306C2
6. AfUGT306A3
7. AfUGT302A3
8. AfUGT301C2
9. AfUGT301A3
10. AfUGT49A4
11. AfUGT36C3

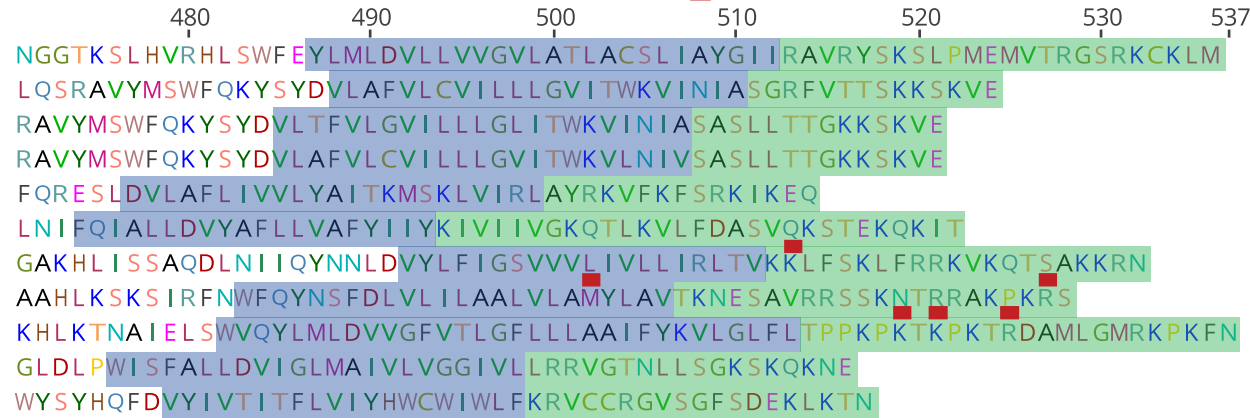


Fig. S11. nonsynonymous SNPs location in UGTs targeted by sequencing. Functional protein regions in non-aligned UGTs included in the targeted enrichment sequencing are highlighted in the figure legend. Nonsynonymous SNPs that are significantly differentiated between putatively susceptible and resistant from Malawi, Cameroon, and Uganda, and between the FANG and the FUMOZ colony are highlighted with a red color in the protein sequence. For further details on those SNPs see Table S8.