Oxamniquine resistance alleles are widespread in Old World *Schistosoma mansoni* and predate drug deployment

Frédéric D. Chevalier[1*], Winka Le Clec'h[1], Marina McDew-White[1], Vinay Menon[1], Meghan A. Guzman[2], Stephen P. Holloway[3], Xiaohang Cao[3], Alexander B. Taylor[3,4], Safari Kinung'hi[5], Anouk N. Gouvras[6,7], Bonnie L. Webster[6,7], Joanne P. Webster[6,8], Aidan M. Emery[6,7], David Rollinson[6,7], Amadou Garba Djirmay[9,10], Khalid M. Al Mashikhi[11], Salem Al Yafae[11], Mohamed A. Idris[12], Hélène Moné[13], Gabriel Mouahid[13], P. John Hart[3,4], Philip T. LoVerde[2], Timothy JC. Anderson[1]*

[1] Texas Biomedical Research Institute, San Antonio, Texas, USA

[2] Departments of Pathology and [3]Biochemistry & Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

[4] X-ray Crystallography Core Laboratory, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

[5] National Institute for Medical Research, Mwanza, United Republic of Tanzania

[6] London Centre for Neglected Tropical Disease Research (LCNDTR), Imperial Collge, London, United Kingdom

[7] Wolfson Wellcome Biomedical Laboratories, Natural History Museum, London, United Kingdom

[8] Centre for Emerging, Endemic and Exotic Diseases (CEEED), Royal Veterinary College, University of London, United Kingdom

[9] Réseau International Schistosomiases Environnemental Aménagement et Lutte (RISEAL), Niamey, Niger

[10] World Health Organization, Geneva, Switzerland

25 [11] Directorate General of Health Services, Dhofar Governorate, Salalah, Sultanate of Oman

26 [12] Sultan Qaboos University, Muscat, Sultanate of Oman

27 [13] Host-Pathogen-Environment Interactions laboratory, University of Perpignan, France

28

29 *Corresponding authors: fcheval@txbiomed.org, tanderso@txbiomed.org

30

31 Short title: Standing variation for oxamniquine resistance in schistosomes

ABSTRACT

Do mutations required for adaptation occur *de novo*, or are they segregating within populations as standing genetic variation? This question is key to understanding adaptive change in nature, and has important practical consequences for the evolution of drug resistance. We provide evidence that alleles conferring resistance to oxamniquine (OXA), an antischistosomal drug, are widespread in natural parasite populations under minimal drug pressure and predate OXA deployment. OXA has been used since the 1970s to treat *Schistosoma mansoni* infections in the New World where *S. mansoni* established during the slave trade. Recessive loss-of-function mutations within a parasite sulfotransferase (SmSULT-OR) underlie resistance, and several verified resistance mutations, including a deletion (p.E142del), have been identified in the New World. Here we investigate sequence variation in *SmSULT-OR* in *S. mansoni* from the Old World, where OXA has seen minimal usage. We sequenced exomes of 204 *S. mansoni* parasites from West Africa, East Africa and the Middle East, and scored variants in *SmSULT-OR* and flanking regions. We identified 39 non-synonymous SNPs, 4 deletions, 1 duplication and 1 premature stop codon in the *SmSULT-OR* coding sequence, including one confirmed resistance deletion (p.E142del). We expressed recombinant proteins and used an *in vitro* OXA activation assay to functionally validate the OXA-resistance phenotype for four predicted OXA-resistance mutations. Three aspects of the data are of particular interest: (i) segregating OXA-resistance alleles are widespread in Old World populations (4.29 – 14.91% frequency), despite minimal OXA usage, (ii) two OXA-resistance mutations (p.W120R, p.N171IfsX28) are particularly common (>5%) in East African and Middle-Eastern populations, (iii) the p.E142del allele has identical flanking SNPs in both West Africa and Puerto Rico, suggesting that parasites bearing this allele colonized the New World during the slave trade and therefore predate OXA deployment. We conclude that standing variation for OXA resistance is widespread in *S. mansoni*.

AUTHOR SUMMARY

It has been argued that drug resistance is unlikely to spread rapidly in helminth parasites infecting humans. This is based, at least in part, on the premise that resistance mutations are rare or absent within populations prior to treatment, and take a long time to reach appreciable frequencies because helminth parasite generation time is long. This argument is critically dependent on the starting frequency of resistance alleles – if high levels of "standing variation" for resistance are present prior to deployment of treatment, resistance may spread rapidly. We examined frequencies of oxamniquine resistance alleles present in *Schistosoma mansoni* from Africa and the Middle East where oxamniquine has seen minimal use. We found that oxamniquine resistance alleles are widespread in the Old World, ranging from 4.29% in the Middle East to 14.91% in East African parasite populations. Furthermore, we show that resistance alleles from West African and the Caribbean schistosomes share a common origin, suggesting that these alleles travelled to the New World with *S. mansoni* during the transatlantic slave trade. Together, these results demonstrate extensive standing variation for oxamniquine resistance. Our results have important implications for both drug treatment policies and drug development efforts, and demonstrate the power of molecular surveillance approaches for guiding helminth control.

74 **INTRODUCTION**

75 The rate at which drug resistance alleles (or any other beneficial alleles) spread within
76 populations in response to a drug treatment (or any other selection pressure) is critically
77 dependent on the starting frequency of resistance alleles in the population when new drugs
78 are deployed (1,2). If no such alleles are present when a novel drug is introduced, then there is
79 a waiting time for resistance alleles to arise. Furthermore, because the starting allele frequency
80 will be $1/(2N_e)$, where N_e is the effective population size, the vast majority of resistance alleles
81 that arise will be lost due to genetic drift and fail to establish (2,3) (Fig. 1). The barrier to
82 establishment is particularly severe for recessive traits because these will be present in
83 heterozygotes at low frequency and therefore not exposed to selection. This effect, otherwise
84 known as "Haldanes's sieve" (4), may be particularly relevant for drug resistance evolution in
85 diploid pathogens, because resistance mutations in drug targets typically result in phenotypic
86 resistance only when two copies are present. However, if resistance alleles are already
87 segregating as standing variation in pathogen populations when drug treatment is initiated,
88 then there is no waiting time for mutations to arise. Furthermore, there is a high probability of
89 fixation and spread is rapid because many resistance alleles are present in homozygous
90 recessive genotypes and therefore exposed to selection. Hence, whether resistance alleles (or
91 alleles for other traits) arise *de novo* or are segregating as standing variation, is a central
92 question in evolutionary biology and public health, and key to predicting the effective "shelf
93 life" of new drug treatments (3,5).

94 There is growing evidence for resistance to anti-helminthics in several helminths infecting
95 humans, including *Onchocerca volvulus* (6), the filarial nematode causing river blindness, and in
96 soil transmitted helminths (hookworm, whipworm and *Ascaris* roundworms) (7) which
97 cumulatively infect over one billion people worldwide (8). However, perhaps the best
98 understood is oxamniquine resistance in schistosome blood flukes, for which the mechanism of
99 drug action and the genetic and molecular basis for drug resistance are now known.
100 Oxamniquine (OXA) kills *Schistosoma mansoni*, but is ineffective against two other major
101 schistosome species infecting humans (*S. haematobium* and *S. japonicum*) (9). OXA is a pro-
102 drug that is activated by a schistosome sulfotransferase (SmSULT-OR) encoded on chromosome

103 6 (10). Loss-of-function mutations within SmSULT-OR result in resistance: only parasites that
104 are homozygous for resistance alleles are phenotypically resistant (OXA-R). We initially
105 identified the locus and mutation, a single amino acid deletion (p.E142del), underlying OXA-R
106 using a genetic cross involving a resistant parasite line selected from a parasite isolate from a
107 Puerto Rican patient in 1971 (11). Additionally we identified a second mutation (p.C35R) in *S.*
108 *mansoni* obtained from an incurable Brazilian patient (12). These two mutations result in
109 disruption of the active site of the SmSULT-OR and inability to activate OXA (10). Both
110 p.E142del and p.C35R, and two additional mutations resulting in truncated proteins, were
111 subsequently identified in a field survey of SmSULT-OR genetic variation of parasites from a
112 single Brazilian village (13). The combined allele frequency of these four resistant alleles was
113 1.85%.

114 OXA was widely used in the New World, where only *S. mansoni* is present, during the 1970s
115 to early 2000's (14), so the OXA-R alleles we observed in Brazil (13) could conceivably have
116 resulted from drug selection. In contrast, OXA saw minimal usage in Africa where both *S.*
117 *haematobium* and *S. mansoni* are present. In South America, an estimated 6 million doses of
118 OXA were used to treat the 7 million people infected with *S. mansoni* prior to 1987 (15). In
119 contrast, 3 million doses had been used in Africa (Egypt, Ethiopia, Ivory Coast, Kenya, Malagasy,
120 Malawi, Rwanda-Burundi, South Africa, Sudan, Tanzania, Uganda, Zambia, Zaire) and the
121 Middle East (Arabian penensula) (15), where an estimated ~60 million people were infected
122 with *S. mansoni* (16). Hence, OXA selection was minimal in Africa compared with South
123 America. The central aim of this paper is to determine whether *de novo* mutation or standing
124 variation best explains OXA-R in *S. mansoni* populations. To do this we examined *SmSULT-OR*
125 genetic variation in Old World parasite populations where OXA treatment has been minimal.

126 **RESULTS**

127 **1. Samples**

128 We sequenced exomes from 92 miracidia from Tanzania (n=57), Niger (n=10), and Senegal
129 (n=25) from the SCAN collection at the Natural History Museum (17), with each miracidium
130 derived from a different patient. In addition, we included 112 samples collected from Oman:
131 these included 86 single worms and 26 pools of cercariae from single infected snails. The Omani
132 samples were derived from 11 human infections, 2 naturally infected rodents, and 26 naturally
133 infected snails (Table 1). We obtained an average of 15.7 ± 8 (mean ± standard deviation, s.d.)
134 million reads per library (supp. table 1). A very high proportion of these reads were mapped to
135 the reference genome (91.9 ± 6.29 (s.d.) % on average). We captured on average 99.08 ± 3.92
136 (s.d.) % of the bait regions. We obtained an average read depth of 64.40 ± 32.27 (s.d.) in the
137 bait regions. For *SmSULT-OR*, we obtained sequence across 95.06 ± 6.54 (s.d.) % of the coding
138 sequence with an average read depth of 43.09 ± 26.96 (s.d.).

139 We compared these Old World parasite samples with: (i) the OXA-R parasite from Puerto
140 Rico (HR) carrying the causative p.E142del mutation identified previously (10), for which whole
141 genome sequence is available, and (ii) published *SmSULT-OR* sequences from a single Brazilian
142 location (n=189) (13) (Table 1).

143

144 **2. New and known mutations identified in the *SmSULT-OR* gene**

145 We identified a total of 85 *SmSULT-OR* mutations across all three populations (West Africa, East
146 Africa and Middle East), including 76 coding mutations and 9 non-coding mutations (supp. table
147 2). Exon 1 carried 24 mutations. These included 13 non-synonymous single nucleotide
148 polymorphisms (SNPs), 8 synonymous SNPs, one duplication, one deletion, and one premature
149 stop codon. Exon 2 carried 52 mutations: 26 non-synonymous SNPs, 23 synonymous and three
150 deletions. There were no differences in the density of mutations in the two exons (Fisher's
151 exact test, d.f.=1, p=0.09). The overall transition/transversion ratio was 1.94. Forty-three
152 percent (37/85 mutations) showed an average allele frequency greater than 5%.

153 Among the 85 mutations, we have previously found three coding and one non-coding
154 mutations in the New World (Caribbean and Brazil): the substitution p.P67L (g.200C>T), the
155 deletion p.E142del (g.4348_4350delGAA), the substitution p.L256W (g.4691T>G) and the 3'
156 untranslated region (UTR) substitution g.4720C>T (10,13). Just one of these mutations
157 (p.E142del) encodes a confirmed OXA-R allele. Frequencies of all mutations in each population
158 are shown in supp. table 2.

159 We measured nucleotide diversity (π) in both New and Old World populations. Nucleotide
160 diversity was the highest in Oman ($\pi = 0.02424 \pm 2.4 \times 10^{-4}$) followed by East Africa ($\pi = 0.00195$
161 $\pm 2 \times 10^{-4}$), West Africa ($\pi = 0.00126 \pm 1.7 \times 10^{-4}$) and Brazil ($\pi = 0.00021 \pm 0.3 \times 10^{-4}$). If drug
162 pressure drives the changes at *SmSULT-OR*, we might expect to see signals of positive selection.
163 We examined sequence conservation by measuring proportions of non-synonymous difference
164 per non-synonymous site relative to synonymous changes per synonymous site (K_a/K_s) between
165 *S. mansoni* and a *S. rodhaini* outgroup. K_a/K_s ranged from 0.20 – 0.24 suggesting that changes at
166 non-synonymous sites are 4-5 times less abundant than at synonymous sites, consistent with
167 weak purifying selection. Similarly, the ratio of non-synonymous polymorphisms per non-
168 synonymous site relative to synonymous polymorphisms per synonymous site (pN/pS) ranged
169 from 0.111 to 1.36 and none of these values were significantly greater than unity (Table 2).
170 These tests did not provide evidence that *SmSULT-OR* is under positive selection.

171

172 **3. *In vitro* tests and *in silico* evaluation of impact of mutations**

173 We used three approaches to identify *SmSULT-OR* mutations that are likely to result in OXA-R.

174 a. Visual assessment

175 Some mutations have obvious deleterious effect on the protein by introducing premature stop
176 codons. This is the case for three mutations identified in *SmSULT-OR*. These mutations will
177 generate truncated proteins that do not contain the active site. Other mutations, such as
178 p.L179P, p.P225S or p.W120R, were either close enough to the OXA or the 3'-
179 phosphoadenosine-5'-phosphosulfate (PAPS) binding sites or showed potential steric hindrance

180 or structural deformation likely to have an effect on protein activity. Other mutations, such as
181 p.A74T, p.P106S or p.Q176R, were on the surface of the protein, distant to any active site, and
182 therefore unlikely to have an impact.

183 b. Functional assay of OXA binding

184 We used an *in vitro* OXA activation assay to functionally assay the impact of mutations on
185 SmSULT-OR activity for four newly discovered mutations and the two known OXA-R mutations
186 (Fig. 3). The known resistant mutations, p.C35R and p.E142del, showed from zero to very low
187 level of OXA activation, as expected (10). p.L179P showed a similarly low level of activation. The
188 mutations p.S160L and p.P225S showed intermediate OXA activation: we conservatively classed
189 these as OXA sensitive alleles. p.P106S showed similar activity to the wild type and p.S160L. We
190 were able to produce recombinant protein p.W120R but all our attempts to fold this protein
191 were unsuccessful. This mutation clearly has a dramatic impact on protein stability as predicted
192 by visual inspection and thermodynamic modelling results (below). We considered this
193 mutation to be OXA-R.

194 c. Thermodynamic modelling

195 We used thermodynamic modelling to evaluate the potential impact of substitutions on protein
196 stability (indels such as p.E142del cannot be modelled). The difference in free enthalpy ($\Delta\Delta G$)
197 ranged from -2.942 to 19.826 for the 37 mutations (Fig. 4). Among those, p.L179P and p.C35R
198 which have the greatest impact on OXA activation (Fig. 3) showed the second and third highest
199 $\Delta\Delta G$, respectively. p.W120R, for which we were unable to fold recombinant protein, showed
200 the highest $\Delta\Delta G$, consistent with a dramatic impact on stability. p.S160L and p.P225S showed
201 intermediate $\Delta\Delta G$ consistent with the results of the OXA activation assay.

202 In total, we identified 7 independent OXA-R mutations in Old and New World parasite
203 populations examined using these three approaches (Table 3, Fig. 2). These included 3 indels
204 and 4 amino acid mutations.

205

206 **4. Frequency of OXA-R alleles**

207 Having determined which of the mutations identified are likely to cause OXA-R, we measured
208 the frequency of resistant variants and the frequency of resistant parasites in each population
209 (Table 3, Fig. 5). To accurately identify resistant alleles and parasites, we phased our variant
210 calling data on the first 3 Mb of the chromosome 6 (corresponding to 30,812 variable sites).
211 West African samples carried only two resistant variants (p.S12X and p.E142del). East African
212 samples carried five resistant variants, two of them at a frequency over 0.05 (p.W120R and
213 p.N171IfsX28). The p.W120R mutation was also found at high frequency in Middle East (0.065).
214 We found the highest frequency of OXA-R alleles in East Africa (14.91%, 17/114), followed by
215 the Middle East (6.25%, 14/224), West Africa (4.29%, 3/70), and Brazil (1.85%, 7/378).

216 Most parasites carried OXA-R alleles in the heterozygous state and are predicted to be OXA
217 sensitive. Heterozygous OXA-sensitive parasites were found in West and East Africa (8.6% and
218 24.6%, respectively), Middle East (13.4%) and in Brazil (1.06%). Parasites that are homozygous
219 for OXA-R alleles at *SmSULT-OR* (but not necessarily carrying the same resistant variant) are
220 OXA-R. We found homozygous parasites predicted to be phenotypically OXA-R at a frequency of
221 5.26% (3/57) in East Africa and 1.06% (2/189) in South America only (Fig. 5).

222

223 **5. Haplotype analysis of *SmSULT-OR* and flanking regions**

224 We used our phased data to investigate the haplotypes surrounding our resistant variant
225 (p.E142del) in the Old and New World (Caribbean) (Table 1) and to investigate whether these
226 resistant alleles derived from a common ancestor. This is of interest because it would suggest
227 that these alleles predate the slave trade. We identified an identical 102.5 kb haplotype block
228 around p.E142del shared between the Caribbean sample and one of the West African samples
229 from Niger (Fig. 6A). This haplotype block contained 399 SNPs that varied in at least one of the
230 205 parasites examined (all samples except those from Brazil where haplotype data were not
231 available). We used a minimum spanning network to investigate haplotype relationships using
232 the 399 bi-allelic variants found in this 102.5 kb block across all phased haplotypes (Fig. 6B). The
233 haplotypes associated with the p.E142del OXA-R alleles found in both Puerto Rico and Niger

234 clustered together, consistent with a common origin for West African and Caribbean p.E142del
235 OXA-R mutations.

236 The p.E142del mutation was also identified in Brazil (13) but we sequenced the *SmSULT-OR*
237 exons only, rather than complete exomes, for these Brazilian samples preventing
238 reconstruction of extended haplotypes. This Brazilian p.E142del was also found associated with
239 two mutations (p.P67L and g.4720C>T) found in West Africa but without the p.L256W found in
240 samples from the Caribbean and Niger. Similarly, the p.E142del was also found in a Tanzanian
241 sample but this sample has a distinct haplotype from those found in West Africa (Fig. 6). These
242 results suggest either that there are several independent origins of East and West African
243 p.E142del mutations, or alternatively, that the p.E142del is extremely old and has distinctive
244 flanking haplotypes resulting from recombination and mutation.

245 **DISCUSSION**

246 **1. Evidence for standing genetic variation for OXA resistance**

247 Two lines of evidence support the view that standing variation is the source of OXA-R alleles in
248 schistosomes. First, we have indirect evidence: OXA-R alleles are geographically widespread in
249 Africa and the Middle East despite limited treatment with OXA (or hycanthone, its structural
250 analog; (18)) in these regions. Furthermore, that OXA-R alleles are found at high frequency in
251 East Africa, where there has been minimal use of OXA, while OXA-R alleles are found at low
252 frequency in Brazil, where OXA has been used extensively, is also consistent with the idea that
253 mutation and drift, rather than selection, explain the patterns of variation observed. The
254 second line of evidence is more direct: we showed that an OXA-R mutation, p.E142del, with an
255 identical 102.5 kb flanking haplotype, is sampled from both West African and Caribbean
256 schistosomes. This result strongly suggests that p.E142del alleles were present in West Africa
257 prior to the transatlantic slave trade (1501-1867) (19), and were transferred to the New World
258 on ships carrying West African slaves. Hence, OXA-R alleles were segregating within *S. mansoni*
259 populations at least 470 years before deployment of OXA in the 1970s (20).

260 Our molecular data are backed up by clinical observations from the early use of OXA.
261 Resistant parasites were detected in Brazil in the 1970s (21,22), before any mass drug
262 administration (14). Similar observations of *S. mansoni* infected patients resistant to OXA
263 treatment were made in East Africa (23). The existence of OXA-R parasites was confirmed
264 experimentally by infecting mice with parasites isolated from patients that were parasite
265 positive following drug treatment (21–23). Together these observations support the existence
266 of segregating OXA-R alleles in schistosome populations before OXA deployment for parasite
267 control.

268

269 **2. Differences in OXA treatment efficacy between East and West Africa**

270 Given the variation in OXA-R frequencies that we observe in different geographical samples, it is
271 interesting to examine the literature on clinical treatment efficacy of OXA for treating

272 schistosomiasis patients. Early trials of OXA to treat patients with intestinal schistosomiasis
273 resulted in multiple treatment failures in Egypt, East Africa and South Africa (23–27). As a
274 consequence, the WHO recommended use of higher doses of OXA in East Africa, compared
275 with West Africa (28). Human host metabolism did not explain this lack of efficacy because the
276 OXA availability in blood was the same or higher in the East African patient populations than in
277 West African or South American populations (29,30). Therefore, the presence of resistant
278 parasites was the most likely explanation. OXA-R parasites were identified in Kenya by treating
279 mice infected with parasites from patients showing poor treatment response (23). We observed
280 a 14.91% frequency of OXA-R alleles (and 5.26% frequency of homozygous OXA-R parasites) in
281 East Africa compared with 4.29% frequency of OXA-R alleles (and no homozygous OXA-R
282 parasites) in West Africa. Our results provide a molecular explanation for the poor treatment
283 response observed in East Africa, relative to other locations.

284

285 **3. Implications for development of OXA derivatives**

286 OXA kills *S. mansoni* but not *S. haematobium* or *S. japonicum*. OXA derivatives are currently
287 under development to obtain more potent molecules acting against all three schistosome
288 species. In fact, OXA derivatives that kill adult worms of the three main species of schistosomes
289 infecting humans *in vitro* have now been developed (31). The low frequency of homozygous
290 OXA-R parasites in *S. mansoni* suggests that OXA derivatives will result in imperfect cure rate
291 for this species, particularly in East Africa. However, the primary purpose of generating OXA
292 derivatives is to treat *S. haematobium* and *S. japonicum*, perhaps as a partner drug for
293 praziquantel (PZQ). Examining natural variation in homologous sulfotransferase genes from
294 these two species (*ShSULT-OR* and *SjSULT-OR*) to better understand the potential for resistance
295 evolution will be critical if OXA derivatives are to be developed to control these parasites.

296

297 **4. Implications of standing variation for drug resistance evolution**

298 Standing variation in drug resistance genes has a major impact on how fast resistance evolves
299 (2). If resistant alleles are already present in the population prior to treatment, drug resistance
300 has the potential to spread much more rapidly because there is no waiting time for a resistance
301 mutation to appear *de novo* (Fig. 7A) (3). An increasing number of examples support this view.
302 The nematode *Caenorhabditis elegans* carries a large diversity of *β-tubulin* alleles allowing
303 worms to be resistant to benzimidazoles (BZ) (32). In the filarial worm *Onchocerca volvulus*,
304 standing variation appears to be involved in resistance to ivermectin (6). Standing variation is
305 also involved in herbicide resistance in plants: resistance alleles from the weed *Alopecurus*
306 *myosuroides* were identified in plant collection almost 100 years before herbicides were used
307 (33). Antibiotic resistance provides particularly dramatic examples of standing variation:
308 plasmids encoding multidrug resistance were found in frozen bacteria isolated from the
309 permafrost demonstrating that antibiotic resistance plasmids were already present thousands
310 of years ago (34,35).

311 We do not want to give the impression that standing variation is the only source of variation
312 for evolution of drug resitance. In the case of other well studied helminth parasites, such as
313 *Haemonchus contortus* and *Teladorsagia circumcincta*, resistance alleles are consistent with
314 evolution through recurrent *de novo* mutations or very rare standing variants. While selection
315 of rare pre-existing BZ resistance alleles was shown in French farms (36), selection of *de novo*
316 resistance mutations, associated with different haplotypes, was shown to be common in UK
317 farms (37).

318 The fact that there is strong evidence of standing variation in OXA resistance but not in BZ
319 resistance may reflect the differences in the nature of the mutation underlying resistance. In
320 the case of OXA, resistance is due to loss-of-function mutations resulting in non functional
321 SmSULT-OR. As a consequence, mutations can occur in any position that disrupts substrate
322 binding or protein structure. Hence mutations are expected to arise frequently and will exist in
323 natural populations. In contrast, mutations underlying BZ resistance occur in specific position
324 within the *β-tubulin* gene. As consequence, such mutations occur much more rarely.
325 Furthermore because *β-tubulin* is an essential gene, mutations that disrupt protein structure

326 and result in non functional proteins are strongly selected against and cannot reach high
327 frequency.

328 Standing variation for OXA resistance alleles could arise by drift, or alternatively as a result
329 of selective forces other than OXA treatment. For example, Hahnel *et al*. (32) demonsrated BZ
330 resistance in natural population of *C. elegans* which are not targetted by BZ drugs and
331 speculated that natural compound with similar structure to benzimidazoles could have pre-
332 selected for benzimidazole resistance alleles. In the case of our OXA-R mutations in *SmSULT-OR*,
333 we do not know what selective forces might be involved (see also discussion in section 5).

334 Drug resistance in parasitic helminths infecting livestock or humans is now recognized as a
335 potential concern for control efforts, but was neglected for a long time using the argument that
336 the slow reproduction rate of worms would greatly delay resistance evolution (38). The
337 observation that resistance alleles exist as standing variation within multiple pathogen or
338 parasite populations in the absence of drug treatment suggests previous optimism about the
339 shelf life of drugs was poorly founded. If resistance alleles are already at relatively high
340 frequency in populations when new drugs are deployed, resistance may evolve extremely
341 rapidly (3,5).

342

343 **5. Prediction of the resistance spread**

344 We can use the frequency of resistance allele to model resistance spread in parasite
345 populations. In the case of OXA resistance, using an N_e estimate of 65,000 (39), OXA-R alleles
346 spread from 15% starting frequency to fixation in 170 generations in 100% of simulations when
347 selection is strong (s=0.1). Assuming a generation time between 3 months and 1 year, this is
348 equivalent to a period of 42.5 to 170 years (Fig. 7A). However, if OXA-R alleles were evolving *de*
349 *novo*, simulations indicate that fixation will occur rarely (0.04% of the simulations) and this
350 would take 1,197 generations (corresponding to a period of 299.25 to 1,197 years) from the
351 time a mutation arises. We can use the observed frequency of OXA-R alleles segregating in East
352 African populations to make predictions about the spread of phenotypically resistant parasites
353 (i.e. those homozygous for OXA-R alleles) (Fig 7B). If we use a frequency of 10% resistant

354 parasites to denote unacceptable efficacy (this threshold is used for antimalarial drugs (40))
355 then this threshold will be reached in 19 to 444 generations (4.75 to 444 years) using a range of
356 selection coefficients (from 0.01 to 0.2). The selection coefficients driving spread of drug
357 resistance alleles in schistosomes is not known, but the range we have used covers measured
358 selection coefficients driving drug resistance spread in another human parasite (*Plasmodium*
359 *falciparum*) (41,42). These simulations suggest that, while OXA-R allele frequencies are
360 relatively high in East Africa, OXA derivatives could still be extremely effective for schistosome
361 control in the short-medium term, particularly if deployed as a combination therapy, partnered
362 with a second drug with a different mode of action (e.g. PZQ).

363 Our simple models for predicting spread of OXA-R alleles under drug selection ignore fitness
364 costs of OXA-R alleles, which might limit the rate of spread. The fitness cost of inactive SmSULT-
365 OR in the absence of drug treatment still needs to be assessed accurately (13) but the
366 occurrence of natural homozygotes for defective alleles suggests limited costs. The exact role of
367 this enzyme in schistosome biology is not known. Sulfotransferases transfer the sulfo group
368 from a donor (usually PAPS) to the substrate, resulting in inactivating or solubilizing of the
369 substrate (43). Substrates can be endogenous signal molecules (hormones, neurotransmitters)
370 or exogenous chemicals. In the latter case, sulfotransferases play a role in detoxifying natural or
371 synthetic toxins (43), but sometimes they do the opposite by activating pro-drugs as observed
372 with the anti-tumor drug N-benzyl indolecarbinols (44) or OXA (10,45). Because the same
373 enzyme can sulfate both endogenous and exogenous compounds, it is often difficult to
374 determine their exact biological role. In schistosomes, we speculate that SmSULT-OR could
375 have roles in regulating endogenous molecules as well as detoxifying chemicals present in their
376 environment (*i.e.* in the feces, water, snail or human blood, etc.).

377

378 **6. Molecular markers for resistance surveillance**

379 Molecular markers of OXA-R now allow efficient monitoring of the distribution of OXA-R alleles
380 in schistosome populations. This approach is widely used for other parasites and pathogens
381 such as malaria parasite (46) and HIV (47). For schistosome, collections such as SCAN which

382 contain thousands of parasite samples from multiple locations (17) will be extremely valuable
383 for this work: regular sampling in different endemic regions will help to update the resistance
384 landscape. However, not all identified *SmSULT-OR* mutations will lead to resistance, so it is
385 critical that molecular screening is paired with functional evaluation or computational
386 prediction, as we have done here.

387 PZQ is the only drug currently available to treat schistosomiasis and effort in mass drug
388 administration has recently been expanded 10 fold (48), with a target of administering 250
389 million treatments per year, increasing selective pressure for PZQ resistance. We suspect that
390 standing variation for PZQ resistance is also likely, because this trait is easy to select in the
391 laboratory. This has been done by several groups using independent laboratory populations,
392 and resistance has spread within very few generations (49–54), strongly suggesting that
393 resistance alleles were already present in these laboratory populations. Furthermore, the
394 laboratory populations used for selection were isolated before PZQ was available (23,55) and so
395 were not exposed to drug selection. Molecular markers for monitoring the frequency of PZQ-R
396 in natural parasite populations would be extremely valuable, as demonstrated by our work on
397 OXA-R.

398 **MATERIALS AND METHODS**

399 **1. Ethics statement**

400 a. African samples

401 Samples from Senegal and Niger were collected as part of the EU-CONTRAST project (56), a
402 multidisciplinary alliance to optimize schistosomiasis control and transmission surveillance in
403 sub-Saharan Africa, as detailed in Webster *et al.* (57) with ethical approval granted by ethical
404 committees of the Ministry of Health (Dakar, Senegal), and the Niger National Ethical
405 Committee (Niamey, Niger) with additional ethical approval obtained from the St Mary's
406 Hospital Local Ethics Research Committee, R&D office (part of the Imperial College Research
407 Ethics Committee (ICREC; EC no. 03.36. R&D no. 03/SB/033E)), in combination with the ongoing
408 CONTRAST and Schistosomiasis Control Initiative (SCI) activities.

409 For the Tanzanian samples collected as part of the Schistosomiasis Consortium for
410 Operational Research and Evaluation (SCORE), ethical approvals were granted by the Imperial
411 College Research Ethics Committee and the SCI Ethical approval (EC no. 03.36. R&D no.
412 03/SB/033E); the National Institute for Medical Research (NIMR, reference no.
413 NIMR/HQ/R.8a/Vol. IX/1022); University of Georgia Institutional Review Boards, Athens, GA
414 (2011-10353-1).

415 Following routine procedures in the field, the objectives of the study were first explained to
416 the local village chiefs and political and religious authorities who gave their consent to conduct
417 the study. Written consent for the schoolchildren to participate in longitudinal monitoring of
418 the national control programme for schistosomiasis was given by head teachers, and/or village
419 chiefs where there were no schools, due to the fact that in African villages, written consent of
420 the child's guardian is often very difficult to obtain owing to the associated impoverished
421 conditions and often low literacy. Each individual child also gave verbal consent before
422 recruitment. Following sampling, a praziquantel treatment (40 $mg.kg^{-1}$) was offered to infected
423 participants.

424

425

426 b. Omani samples

427 We obtained ethical clearance from the Sultan Qaboos University and the Ministry of Health of
428 Oman to use positive stool samples for schistosomiasis collected by the Ministry of Health of
429 Oman during epidemiological screening. Ethical approval was given by the Medical Research
430 and Ethical Committee (MREC) of the Non-Communicable Disease Control Section of the
431 Directorate General of Health Affairs, Headquarters, Ministry of Health, Oman (no.
432 MH/DGHA/DSDC/NCD/R&S/167/01).

433 We obtained ethical approvals of animal studies in France from the French Ministère de
434 l'Éducation Nationale, de la Recherche et de la Technologie, from the French Ministère de
435 l'Agriculture et de la Pêche (agreement no. A 66040), and from the French Direction
436 Départementale de la Protection des Populations (no. C 66-136-01 with prefectoral order no.
437 2012-201-0008). Certificate for animal experimentation was given to H.M. (authorization no. C
438 66.11.01; articles no. R 214-87, R 214-122 and R 215-10). Housing, breeding and animal care
439 followed the guidelines of the French CNRS. The different protocols used in this study had been
440 approved by the French veterinary agency from the DRAAF Languedoc-Roussillon (Direction
441 Régionale de l'Alimentation, de l'Agriculture et de la Forêt), Montpellier, France (authorization
442 no. 007083).

443

444 **2. Sampling**

445 a. African samples

446 *S. mansoni* miracidia were collected from individual patients in three different West African
447 countries in 2007: 27 patients from two locations in Senegal, and 17 patients from two
448 locations in Niger (supp. table 3) (57,58). Additionally, *S. mansoni* miracidia were collected as
449 part of the SCORE program from 64 children in seven villages on the shores of Lake Victoria in
450 Tanzania (East Africa) in January 2012 (supp. table 3) (59).

451 Collection of miracidia was performed as previously described (57,60). Briefly, individual
452 stool samples from positive patients were homogenized through a mesh and washed through
453 with water. The content was then transferred in a Pitchford funnel assembly, washed with

454 additional water, and the filtered homogenate was drained into a Petri dish. The Petri dish
455 containing the homogenate was then left in bright ambient light (not direct sunlight) to allow
456 hatching of miracidia. Miracidia were visualized under a dissecting microscope and individually
457 captured in 3-5 µL of water using a micropipette. Miracidia were then pipetted individually onto
458 Whatman FTA cards for DNA preservation. Cards were allowed to dry for 1 hour and then
459 stored for future research (57,58,60).

460 b. Omani samples

461 The Omani schistosome samples were collected during 8 field trips, from 2001 to 2015, in
462 different areas in Dhofar, Oman and originated from either *Homo sapiens*, *Rattus rattus* or
463 *Biomphalaria pfeifferi* (supp. table 3): 11 patients from 4 localities, 2 rats from 2 localities, 26
464 snails from 3 localities.

465 We obtained schistosome adult worms from naturally infected rats and from laboratory
466 infected mice. The mice were infected with cercariae from naturally infected snails or from
467 laboratory snails infected with miracidia isolated from stool samples. We collected miracidia
468 from stool samples as described in Moné *et al.* (61). We exposed *Biomphalaria pfeifferi* snails to
469 miracidia as described in Mouahid *et al.* (62). We maintained laboratory and naturally infected
470 snails at constant temperature (26°C) and balanced photoperiod (12h light / 12h dark) and fed
471 them with fresh lettuce *ad libitum*. Prior to infection, we anaesthetized the mice by
472 intraperitoneal injection of an anesthetic solution (0.01 $mL.g^{-1}$ of mouse body weight). We
473 prepared the anaesthetic solution using 0.5 mL of Rompun (20 $mg.mL^{-1}$; Bayer) and 1 mL of
474 Imalgène (100 $mg.mL^{-1}$; Rhône Mérieux) diluted in 8.5 mL of autoclaved NaCl 8.5 ‰. Abdomens
475 of anesthetized mice were shaved and exposed during 1 hour to cercariae shed from laboratory
476 or naturally infected snails. Cercariae from naturally infected snails were kept in 95° ethanol
477 when the number of cercariae was too low for mouse infection. Laboratory infected mice or
478 naturally infected rats were euthanized by intraperitoneal injection of sodium pentobarbital
479 solution (1.1 mL diluted in 10 mL of ethanol 10%). We recovered adult worms by perfusion (63);
480 the worms were washed in NaCl 8.5 ‰, and single worms (female or male) were fixed in 95°

481 ethanol and preserved at -20°C. Worms and cercariae were washed in 1X TE buffer for 1 hour
482 prior to DNA extraction.

483

484 **3. DNA processing and library preparation**

485 We amplified DNA and sequenced exomes from single miracidia preserved on FTA cards (1
486 miracidium per patient) for African samples or from DNA extracted from worms or cercariae for
487 Omani samples. FTA-preserved samples were processed following our published protocol (64).
488 Briefly we punched a 2 mm disc containing the miracidium from the FTA card, washed it with
489 the FTA Purification Reagent (GE Healthcare Life Sciences), rinsed it twice with TE^{-1} buffer, and
490 finally dried it. DNA from single worms and cercariae were extracted using the DNeasy Blood
491 and Tissue kit (Qiagen) following the tissue protocol with an incubation time of 2h at 56°C and
492 an elution in 200 µL. Samples were quantified using the Qubit dsDNA HS assay kit (Invitrogen).

493 We then performed whole genome amplification (WGA) on each FTA punch or on 2 µL or 4
494 µL of DNA solution using the Illustra GenomiPhi V2 DNA Amplification kit (GE Healthcare Life
495 Sciences). Amplified DNA was purified with the SigmaSpin™ Sequencing reaction Clean-up
496 (Sigma-Aldrich), following the manufacturer protocol. We quantified purified samples using the
497 Qubit dsDNA BR assay (Invitrogen). DNA samples that failed WGA were cleaned up and
498 concentrated using the Genomic DNA Clean & Concentrator kit (Zymo Research) following the
499 manufacturer protocol and these samples with concentrated DNA were used to perform a new
500 WGA.

501 Because WGA reactions are non-specific, we performed qPCR on each sample from FTA
502 cards to quantify proportion of schistosome DNA (i.e., check for an excess of amplified DNA
503 from the environment (water, fecal matter, etc.)). We excluded samples that showed less than
504 100 schistosome genomes copies in 20 ng of DNA which were not suitable for subsequent
505 exome capture. We assessed genome copies by quantifyng the *S. mansoni α-tubulin* single copy
506 gene (64).

507 We prepared exome capture libraries on the selected African samples and on all the
508 successfully amplified Omani samples using the pre-capture pooling method of the SureSelect
509 XT^2 Target Enrichment System (Agilent) (64). Finally, we sequenced libraries on a HiSeq 2500
510 and data were demultiplexed using the Casava pipeline. Raw sequencing data are accessible
511 from the NCBI Sequence Read Archive under BioProject accession numbers PRJNA439266,
512 PRJNA560069, PRJNA560070.

513

514 **4. Alignment, variant calling and phasing**

515 We aligned sequences as described in Le Clec'h *et al.* (64). Briefly, we aligned data against the
516 v5 *S. mansoni* genome using BWA and SAMtools, realigned around indels using GATK, PCR
517 duplicates were marked using picard and Q-score were recalibrated using GATK.

518 To identify variants in the *SmSULT-OR* (Smp_089320) gene, we performed a variant calling
519 using FreeBayes (v1.1.0-46-g8d2b3a0) (65) in the first 3 Mb of the chromosome 6. We used a
520 minimum base quality (-q) of 20 and a minimum mapping quality (-m) of 30. We included
521 mutation sites identified previously (10,13) in order to genotype these sites specifically for
522 further comparison. We defined 4 populations: Caribbean (HR), West Africa (SN and NE), East
523 Africa (TZ) and Middle-East (OM). We excluded variants supported by less than 4 reads using
524 VCFtools (v0.1.14) (66). We then used vcf-subset from VCFtools to remove sites with reference
525 alleles only. We adjusted variant positions using vt normalize (v0.5772-60f436c3) (67). We
526 finally used Beagle (v4.1) (68) to first estimate genotypes and to then phase data. This approach
527 resolves haplotypes using statistical models rather than by direct linkage of SNPs on short
528 reads. Unphased variant calling data and scripts used to analyze the data are available on
529 Zenodo (DOI: 10.5281/zenodo.2850876 and 10.5281/zenodo.3370039, respectively).

530

531 **5. Genetics analysis**

532 We performed all the genetics analysis using phased exome data. We handled the VCF data in R
533 (R v3.5.0 (69)) using the vcfR package (v1.8.0.9) (70). We determined the longest haplotype

534 block from the samples carrying the p.E142del using a custom R code after filtering invariable
535 sites in these samples. We built the haplotype network using the R package poppr (v2.8.1)
536 (71,72) after selecting variants within the longest haplotype coordinates showing no more than
537 20% of missing data. We computed the frequency of the OXA-R allele using a custom R script
538 and functional annotation obtained through a custom bash script.

539 We generated haplotype sequences of the coding sequence only for the Old World samples
540 using the updated phased VCF data exported from R and the BCFtools (v1.2) (73). We
541 generated haplotypes of the South American samples from previous genotyping using a custom
542 bash script. We included the *SmSULT-OR* homologue sequence from *S. rodhaini* previously
543 identified (13). We aligned the haplotypes using Clustal Omega (v1.2.2) (74). We determined
544 nucleotide diversity, number of synonymous and non-synonymous sites using DnaSP software
545 (v6.12.01) (75). Populations were defined as described in the previous section. Scripts used to
546 analyze the data are available on Zenodo (DOI: 10.5281/zenodo.3370039).

547

548 **6. Recombinant SmSULT-OR protein production and OXA activation assay**

549 We produced recombinant SmSULT-OR proteins following Chevalier *et al.* (13). Briefly,
550 mutations were introduced in the cloned *SmSULT-OR* gene sequence (Smp_089320; GenBank
551 accession no. HE601629.1) and introduced into *Escherichia coli* for protein production. Proteins
552 were extracted from bacterial culture and purified on an affinity chromatography column. His-
553 tag was removed using Tobacco etch virus (TEV) protease, the solution was dialyzed overnight
554 and passed through affinity chromatography again to remove His-tag and TEV protease. The
555 sample was loaded onto a GE-pre-packed Q anion exchange column and eluted, and the pooled
556 fractions were dialyzed overnight. The protein was finally concentrated at 10 $mg.mL^{-1}$.

557 We tested the impact of a subset of the mutations observed on OXA activation using an *in*
558 *vitro* assay (13). We tested the ability of recombinant SmSULT-OR proteins to activate OXA in a
559 protease inhibitor cocktail with sheared *S. mansoni* gDNA as a final target for the tritiated OXA
560 pre-mixed with ATP, $MgCl_2$ and PAPS co-factor. After 2.5 h of incubation at 37°C, the reaction
561 was stopped and DNA was extracted three times. Radioactivity in the remaining aqueous phase

562 containing the DNA and in water (blank) was counted in a liquid scintillation spectrometer.
563 Blank values were subtracted from sample values. We performed three independent reactions
564 for each recombinant protein.

565

566 **7. Evaluation of impact of mutations on protein stability**

567 We first visually assessed the scored mutations on the SmSULT-OR protein structure (PDB code
568 4MUB (10)) using the mutagenesis function of PyMol software (v2.3.0; Schrödinger, LLC).

569 We then used the Rosetta package (v3.9) to test the impact of scored mutations on protein
570 stability. We used the ddg_monomer application to compute the difference in free enthalpy
571 ($\Delta\Delta G$) between the mutated (amino acid substitutions only) and the wild-type protein. The
572 higher the $\Delta\Delta G$, the more unstable is the mutated protein. For this, we prepared ligand files
573 (OAQ and A3P) by downloading SDF files containing all ligand structures from SmSULT-OR
574 crystals available from the Protein Data Bank website. We modified SDF files to add hydrogens
575 using Openbabel (v2.4.0) (76), and generated params files using molfile_to_params.py from
576 Rosetta. We used the structure of SmSULT-OR to generate a constraint file using the
577 minimize_with_cst application and the convert_to_cst_file.sh script. We removed ligand
578 information from the constraint file. For each mutation, we finally generated a resfile and ran
579 the ddg_monomer application with 50 iterations and options adapted from the protocol 16 of
580 Kellogg *et al.* (77). Scripts used to analyze the data are available on Zenodo (DOI:
581 10.5281/zenodo.3370039).

582

583 **8. Statistical analysis**

584 We conducted statistical analyzes using R (v3.5.0) (69). We performed simulations using a
585 modified version of the driftR simulator (78). The OXA activation assay data were tested for
586 normality using a Shapiro-Wilk and compared by a parametric ANOVA followed by a Tukey
587 post-hoc test.

588 **ACKNOWLEDGMENTS**

589 This study was supported by NIH grants [R01-AI097576 (T.J.C.A.), R21-AI096277 (T.J.C.A.), R01-
590 AI133749 (T.J.C.A.), R01-AI115691 and P50-AI098507 (P.T.L./P.J.H.)], World Health Organization
591 [HQNTD1206356 (P.T.L.)], by the National Center for Advancing Translational Sciences
592 (CTSA/IIMS award no. UL1-TR001120), the UTHSCSA Presidents Collaborative Research Fund
593 (P.T.L./P.J.H.) and the Robert A. Welch Foundation [AQ-1399 (P.J.H.)]. The molecular work at
594 Texas Biomed was conducted in facilities constructed with support from Research Facilities
595 Improvement Program Grant (C06-RR013556) from the National Center for Research Resources
596 (NIH). The AT&T Genomics Computing Center supercomputing facilities were supported by the
597 AT&T Foundation and the National Center for Research Resources Grant (S10-RR029392). The
598 X-ray Crystallography Core Laboratory is a part of the Institutional Research Cores at the
599 University of Texas Health Science Center at San Antonio supported by the Office of the Vice
600 President for Research and the Mays Cancer Center (NIH P30-CA054174). W.L was supported by
601 a Cowles fellowship from Texas Biomedical Research Institute.

602 CONTRAST was funded by the European Commission (FP6 STREP contract no. 032203).
603 SCORE (https://score.uga.edu/) activities were funded by the University of Georgia Research
604 Foundation Inc. (prime award no. 50816, sub awards RR374-053/4785416 and RR374-
605 053/4785426), which is funded by the Bill & Melinda Gates Foundation for SCORE projects.
606 SCAN is funded with support from the Wellcome Trust (grant no. 104958/Z/14/Z). For the
607 samples collected as part of the EU-CONTRAST project we would like to thank Dr. Oumar T.
608 Diaw and Dr. Moumoudane M. Seye (Institut Sénégalais de Recherches Agricoles, ISRA, route
609 des Hydrocarbures, Bel Air, Dakar, Senegal) for their coordination of the field work and the
610 collections in Senegal and Dr. Mariama Lamine for fieldwork coordination and the RISEAL team
611 for collections in Niger. Also, we would like to thank Dr. Fiona Allan and Ms. Muriel Rabone for
612 sample and data curation and preparation for material held in SCAN. For the collection of the
613 Tanzania SCORE samples we would also like to acknowledge the hard work of Ms. Teckla Angelo
614 for the fieldwork coordination, and Mr. Honest Nagai, Mr. Boniface Emmanuel, Mr. John
615 Igogote, Dr. Sarah Buddenborg and Mr. Reuben Jonathan for collections in Tanzania. Dr. David
616 Rollinson, Prof. Joanne Webster, Dr. Anouk Gouvras, Dr. Bonnie Webster and Dr. Aidan Emery

617 are members of the London Centre for Neglected Tropical Disease Research, a collaboration
618 between the London School of Hygiene & Tropical Medicine, the Natural History Museum, the
619 Royal Veterinary College, Imperial College London and the Sanger Institute.

620 All the biological material from Oman was obtained thanks to the financial supports from
621 the Ministry of Health in Oman, the Sultan Qaboos University (grant no. IG/MED/MICR/00/01),
622 the French Ministry of Foreign Affairs (French Embassy in Oman) (grants nos. 402419B,
623 402415K and 339660F), the CNRS-Sciences de la Vie (grants no. 01N92/0745/1 and
624 02N60/1340), the CNRS-Direction des Relations internationales (grants no. 01N92/0745 and
625 02N60/1340, and the PICS-CNRS no. 06249 : FRANC-INCENSE), the University of Perpignan and
626 the National Institute of Health (NIH) (grant no. R01-AI133749 T.J.C.A.; Subaward no. 53409
627 H.M.).

628 **REFERENCES**

629

630 1. Barrett RDH, Schluter D. Adaptation from standing genetic variation. Trends in ecology &
631 evolution. 2008 Jan;23(1):38–44.

632 2. Hermisson J, Pennings PS. Soft sweeps: molecular population genetics of adaptation from standing
633 genetic variation. Genetics. 2005 Apr;169(4):2335–52.

634 3. Messer PW, Petrov DA. Population genomics of rapid adaptation by soft selective sweeps. Trends
635 Ecol Evol. 2013 Nov;28(11):659–69.

636 4. Orr HA, Betancourt AJ. Haldane's sieve and adaptation from the standing genetic variation.
637 Genetics. 2001 Feb;157(2):875–84.

638 5. Hawkins NJ, Bass C, Dixon A, Neve P. The evolutionary origins of pesticide resistance. Biological
639 reviews of the Cambridge Philosophical Society. 2018 Jul;

640 6. Doyle SR, Bourguinat C, Nana-Djeunga HC, Kengne-Ouafo JA, Pion SDS, Bopda J, et al. Genome-
641 wide analysis of ivermectin response by *Onchocerca volvulus* reveals that genetic drift and soft
642 selective sweeps contribute to loss of drug sensitivity. PLoS neglected tropical diseases. 2017
643 Jul;11(7):e0005816.

644 7. Diawara A, Schwenkenbecher JM, Kaplan RM, Prichard RK. Molecular and biological diagnostic
645 tests for monitoring benzimidazole resistance in human soil-transmitted helminths. The American
646 journal of tropical medicine and hygiene. 2013 Jun;88(6):1052–61.

647 8. Pullan RL, Smith JL, Jasrasaria R, Brooker SJ. Global numbers of infection and disease burden of soil
648 transmitted helminth infections in 2010. Parasites & vectors. 2014 Jan;7:37.

649 9. Pica-Mattoccia L, Novi A, Cioli D. Enzymatic basis for the lack of oxamniquine activity in
650 *Schistosoma haematobium* infections. Parasitol Res. 1997;83(7):687–9.

651 10. Valentim CLL, Cioli D, Chevalier FD, Cao X, Taylor AB, Holloway SP, et al. Genetic and molecular
652 basis of drug resistance and species-specific drug action in Schistosome parasites. Science.
653 2013;342(6164):1385–9.

654 11. Rogers SH, Bueding E. Hycanthone resistance: development in *Schistosoma mansoni*. Science. 1971
655 Jun;172(3987):1057–8.

656 12. Pica-Mattoccia L, Dias LC, Cioli D. Genetic complementation analysis of two independently isolated
657 hycanthone-resistant strains of *Schistosoma mansoni*. Mem Inst Oswaldo Cruz. 1992;87 Suppl
658 4:211–4.

659 13. Chevalier FD, Le Clec'h W, Eng N, Rugel AR, Assis RR de, Oliveira G, et al. Independent origins of
660 loss-of-function mutations conferring oxamniquine resistance in a Brazilian schistosome
661 population. Int J Parasitol. 2016 Jun;46(27):417–24.

662 14. Coura JR, Amaral RS. Epidemiological and control aspects of schistosomiasis in Brazilian endemic
663 areas. Mem Inst Oswaldo Cruz. 2004;99(5 Suppl 1):13–9.

664 15. Foster R. A review of clinical experience with oxamniquine. Transactions of the Royal Society of
665 Tropical Medicine and Hygiene. 1987;81(1):55–9.

666 16. Hotez PJ, Bundy DA, Beegle K, Brooker S, Drake L, de Silva N, et al. Disease Control Priorities in
667 Developing Countries. In: Jamison D, Breman J, AR, M, *et al.*, editors. 2nd ed. Washington (DC):
668 Oxford University Press & World Bank; 2006. p. 467–82. Available from:
669 http://www.ncbi.nlm.nih.gov/books/NBK11748/

670 17. Emery AM, Allan FE, Rabone ME, Rollinson D. Schistosomiasis collection at NHM (SCAN). Parasit
671 Vectors. 2012;5:185.

672 18. Cioli D, Pica-Mattoccia L, Moroni R. *Schistosoma mansoni*: hycanthone/oxamniquine resistance is
673 controlled by a single autosomal recessive gene. Exp Parasitol. 1992 Dec;75(4):425–32.

674 19. Eltis D, Richardson D, Blight DW. Atlas of the Transatlantic Slave Trade [Internet]. Yale University
675 Press; 2010. Available from: http://www.jstor.org/stable/j.ctt5vm1s4

676 20. Katz N, Pellegrino J, Grinbaum E, Chaves A, Zicker F. Preliminary clinical trials with oxamniquine, a
677 new antischistosomal agent. Rev Inst Med Trop Sao Paulo. 1973;15(1):25–9.

678 21. Katz N, Dias EP, Araújo N, Souza CP. Estudo de uma cepa humana de *Schistosoma mansoni*
679 resistente a agentes esquistossomicidas. Rev Soc Bras Med Trop. 1973 Dec;7(6):381–7.

680 22. Dias, L. C. de S. AND Pedro, R. J. AND Rigo, E. AND Goto, M. M. F. AND Mafra, G. L. Linhagem
681 humana de *Schistosoma mansoni* resistente a esquistossomicidas. Revista de Saúde Pública. 1978
682 Mar;12:110–110.

683 23. Coles GC, Mutahi WT, Kinoti GK, Bruce JI, Katz N. Tolerance of Kenyan *Schistosoma mansoni* to
684 oxamniquine. Transactions of the Royal Society of Tropical Medicine and Hygiene. 1987;81(5):782–
685 5.

686 24. Gupta KK. Schistosoma mansoni treatment with oral oxamniquine in Zambia. East African medical
687 journal. 1984 Aug;61(8):641–4.

688 25. Omer AH. Oxamniquine for treating *Schistosoma mansoni* infection in Sudan. British medical
689 journal. 1978 Jul;2(6131):163–5.

690 26. Ongom VL, Wamboka GW, Kadil AUK. Oxamniquine (UK 4271): a potential antischistosomal drug in
691 the treatment of *Schistosoma mansoni* infections in Uganda. The East African medical journal.
692 1976 Aug;53:505.

693 27. Pitchford RJ, Lewis M. Oxamniquine in the treatment of various schistosome infections in South
694 Africa. S Afr Med J. 1978 Apr;53(17):677–80.

695 28. World Health Organization. WHO model prescribing information: drugs used in parasitic diseases
696 [Internet]. World Health Organization; 1995. Available from:
697 http://apps.who.int/medicinedocs/en/d/Jh2922e/

698 29. Daneshmend TK, Homeida MA. Oxamniquine pharmacokinetics in hepatosplenic schistosomiasis in
699 the Sudan. The Journal of antimicrobial chemotherapy. 1987 Jan;19(1):87–93.

700 30. Kokwaro GO, Taylor G. Oxamniquine pharmacokinetics in healthy Kenyan African volunteers. East
701 African medical journal. 1991 May;68(5):359–64.

702 31. Rugel A, Tarpley RS, Lopez A, Menard T, Guzman MA, Taylor AB, et al. Design, Synthesis, and
703 Characterization of Novel Small Molecules as Broad Range Antischistosomal Agents. ACS medicinal
704 chemistry letters. 2018 Oct;9(10):967–73.

705 32. Hahnel SR, Zdraljevic S, Rodriguez BC, Zhao Y, McGrath PT, Andersen EC. Extreme allelic
706 heterogeneity at a *Caenorhabditis elegans* beta-tubulin locus explains natural resistance to
707 benzimidazoles. PLoS pathogens. 2018 Oct;14(10):e1007226.

708 33. Délye C, Deulvot C, Chauvel B. DNA analysis of herbarium Specimens of the grass weed *Alopecurus*
709 *myosuroides* reveals herbicide resistance pre-dated herbicides. PloS one. 2013;8(10):e75117.

710 34. D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, et al. Antibiotic resistance is
711 ancient. Nature. 2011 Aug;477(7365):457–61.

712 35. Perry J, Waglechner N, Wright G. The Prehistory of Antibiotic Resistance. Cold Spring Harbor
713 perspectives in medicine. 2016 Jun;6(6).

714 36. Silvestre A, Humbert JF. Diversity of benzimidazole-resistance alleles in populations of small
715 ruminant parasites. Int J Parasitol. 2002 Jun 15;32(7):921–8.

716 37. Redman E, Whitelaw F, Tait A, Burgess C, Bartley Y, Skuce PJ, et al. The emergence of resistance to
717 the benzimidazole anthlemintics in parasitic nematodes of livestock is characterised by multiple
718 independent hard and soft selective sweeps. PLoS Negl Trop Dis. 2015 Feb;9(2):e0003494.

719 38. Geerts S, Gryseels B. Drug resistance in human helminths: current situation and lessons from
720 livestock. Clinical microbiology reviews. 2000 Apr;13(2):207–22.

721 39. Crellen T, Allan F, David S, Durrant C, Huckvale T, Holroyd N, et al. Whole genome resequencing of
722 the human parasite *Schistosoma mansoni* reveals population history and effects of selection.
723 Scientific Reports. 2016 Feb;6:20954.

724 40. Organization WH. Artemisinin resistance and artemisinin-based combination therapy efficacy:
725 status report [Internet]. Organization WH, editor. World Health Organization; 2018. Available
726 from: https://apps.who.int/iris/bitstream/handle/10665/274362/WHO-CDS-GMP-2018.18-eng.pdf

727 41. Nair S, Williams JT, Brockman A, Paiphun L, Mayxay M, Newton PN, et al. A selective sweep driven
728 by pyrimethamine treatment in southeast asian malaria parasites. Molecular biology and
729 evolution. 2003 Sep;20(9):1526–36.

730 42. Anderson TJC, Nair S, McDew-White M, Cheeseman IH, Nkhoma S, Bilgic F, et al. Population
731 Parameters Underlying an Ongoing Soft Sweep in Southeast Asian Malaria Parasites. Molecular
732 biology and evolution. 2017 Jan;34(1):131–44.

733 43. Negishi M, Pedersen LG, Petrotchenko E, Shevtsov S, Gorokhov A, Kakuta Y, et al. Structure and
734 function of sulfotransferases. Archives of biochemistry and biophysics. 2001 Jun;390(2):149–57.

735 44. Rothman DM, Gao X, George E, Rasmusson T, Bhatia D, Alimov I, et al. Metabolic Enzyme
736 Sulfotransferase 1A1 Is the Trigger for N-Benzyl Indole Carbinol Tumor Growth Suppression.
737 Chemistry & Biology. 2015 Sep;22(9):1228–37.

738 45. Pica-Mattoccia L, Carlini D, Guidi A, Cimica V, Vigorosi F, Cioli D. The schistosome enzyme that
739 activates oxamniquine has the characteristics of a sulfotransferase. Mem Inst Oswaldo Cruz. 2006
740 Sep;101 Suppl 1:307–12.

741 46. Nsanzabana C, Djalle D, Guérin PJ, Ménard D, González IJ. Tools for surveillance of anti-malarial
742 drug resistance: an assessment of the current landscape. Malaria journal. 2018 Feb;17(1):75.

743 47. Keating P, Pharris A, Leitmeyer K, De Angelis S, Wensing A, Amato-Gauci AJ, et al. Assessment of
744 HIV molecular surveillance capacity in the European Union, 2016. Eurosurveillance. 2017
745 Dec;22(49).

746 48. Fenwick A. Praziquantel: do we need another antischistosoma treatment? Future medicinal
747 chemistry. 2015;7(6):677–80.

748 49. Fallon PG, Doenhoff MJ. Drug-resistant schistosomiasis: resistance to praziquantel and
749 oxamniquine induced in Schistosoma mansoni in mice is drug specific. Am J Trop Med Hyg. 1994
750 Jul;51(1):83–8.

751 50. Couto FFB, Coelho PMZ, Araújo N, Kusel JR, Katz N, Jannotti-Passos LK, et al. *Schistosoma mansoni*:
752 a method for inducing resistance to praziquantel using infected *Biomphalaria glabrata* snails. Mem
753 Inst Oswaldo Cruz. 2011 Mar;106(2):153–7.

754 51. Lotfy WM, Hishmat MG, El Nashar AS, Abu El Einin HM. Evaluation of a method for induction of
755 praziquantel resistance in *Schistosoma mansoni*. Pharmaceutical biology. 2015 Aug;53(8):1214–9.

756 52. Pinto-Almeida A, Mendes T, Armada A, Belo S, Carrilho E, Viveiros M, et al. The Role of Efflux
757 Pumps in Schistosoma mansoni Praziquantel Resistant Phenotype. PloS one.
758 2015;10(10):e0140147.

759 53. Lamberton PHL, Faust CL, Webster JP. Praziquantel decreases fecundity in *Schistosoma mansoni*
760 adult worms that survive treatment: evidence from a laboratory life-history trade-offs selection
761 study. Infectious diseases of poverty. 2017 Jun;6(1):110.

762 54. Sanchez MC, Cupit PM, Bu L, Cunningham C. Transcriptomic analysis of reduced sensitivity to
763 praziquantel in *Schistosoma mansoni*. Molecular and biochemical parasitology. 2019 Mar;228:6–
764 15.

765 55. Lewis FA, Stirewalt MA, Souza CP, Gazzinelli G. Large-scale laboratory maintenance of *Schistosoma*
766 *mansoni*, with observations on three schistosome/snail host combinations. J Parasitol. 1986
767 Dec;72(6):813–29.

768 56. Utzinger J, Brattig NW, Kristensen TK. Schistosomiasis research in Africa: how the CONTRAST
769 alliance made it happen. Acta Tropica. 2013 Nov;128(2):182–95.

770 57. Webster BL, Webster JP, Gouvras AN, Garba A, Lamine MS, Diaw OT, et al. DNA "barcoding" of
771 *Schistosoma mansoni* across sub-Saharan Africa supports substantial within locality diversity and
772 geographical separation of genotypes. Acta Trop. 2013 Nov;128(2):250–60.

773 58. Gower CM, Gouvras AN, Lamberton PHL, Deol A, Shrivastava J, Mutombo PN, et al. Population
774 genetic structure of *Schistosoma mansoni* and *Schistosoma haematobium* from across six sub-
775 Saharan African countries: Implications for epidemiology, evolution and control. Acta Trop. 2013
776 Nov;128(2):261–74.

777 59. Ezeamama AE, He C-L, Shen Y, Yin X-P, Binder SC, Campbell CH, et al. Gaining and sustaining
778 schistosomiasis control: study protocol and baseline data prior to different treatment strategies in
779 five African countries. BMC Infect Dis. 2016 26;16:229.

780 60. Gower CM, Shrivastava J, Lamberton PHL, Rollinson D, Webster BL, Emery A, et al. Development
781 and application of an ethically and epidemiologically advantageous assay for the multi-locus
782 microsatellite analysis of *Schistosoma mansoni*. Parasitology. 2007 Apr;134(Pt 4):523–36.

783 61. Moné H, Holtfreter MC, Allienne J-F, Mintsa-Nguéma R, Ibikounlé M, Boissier J, et al. Introgressive
784 hybridizations of *Schistosoma haematobium* by *Schistosoma bovis* at the origin of the first case
785 report of schistosomiasis in Corsica (France, Europe). Parasitology research. 2015
786 Nov;114(11):4127–33.

787 62. Mouahid G, Mintsa Nguema R, Al Mashikhi KM, Al Yafae SA, Idris MA, Moné H. Host-parasite life-
788 histories of the diurnal vs. nocturnal chronotypes of *Schistosoma mansoni*: adaptive significance.
789 Tropical medicine & international health: TM & IH. 2019 Mar;

790 63. Duvall RH, DeWitt WB. An improved perfusion technique for recovering adult schistosomes from
791 laboratory animals. Am J Trop Med Hyg. 1967 Jul;16(4):483–6.

792 64. Le Clec'h W, Chevalier FD, McDew-White M, Allan F, Webster BL, Gouvras AN, et al. Whole
793 genome amplification and exome sequencing of archived schistosome miracidia. Parasitology.
794 2018 May;1–9.

795 65. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012 Jul 17;
796 Available from: http://arxiv.org/pdf/1207.3907v2

797 66. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and
798 VCFtools. Bioinformatics. 2011 Aug;27(15):2156–8.

799 67. Tan A, Abecasis GR, Kang HM. Unified representation of genetic variants. Bioinformatics (Oxford,
800 England). 2015 Jul;31(13):2202–4.

801 68. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for
802 whole-genome association studies by use of localized haplotype clustering. American journal of
803 human genetics. 2007 Nov;81(5):1084–97.

804 69. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria;
805 2015. Available from: http://www.R-project.org/

806 70. Knaus BJ, Grünwald NJ. VCFR: a package to manipulate and visualize variant call format data in R.
807 Molecular ecology resources. 2017 Jan;17(1):44–53.

808 71. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: an R package for genetic analysis of populations with
809 clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014;2:e281.

810 72. Kamvar ZN, Brooks JC, Grünwald NJ. Novel R tools for analysis of genome-wide population genetic
811 data with emphasis on clonality. Frontiers in genetics. 2015;6:208.

812 73. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and
813 population genetical parameter estimation from sequencing data. Bioinformatics. 2011 Nov
814 1;27(21):2987–93.

815 74. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-
816 quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011 Oct
817 11;7:539.

818 75. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al.
819 DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and
820 Evolution. 2017 Dec 1;34(12):3299–302.

821 76. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open
822 chemical toolbox. Journal of Cheminformatics. 2011;3(1):33.

823 77. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-
824 induced changes in protein structure and stability. Proteins. 2011 Mar;79(3):830–8.

825 78. Battey C. Cjbattey/Driftr: Driftr_V1.3 [Internet]. Zenodo; 2017 [cited 2019 May 13]. Available from:
826 https://zenodo.org/record/345172

827 79. Ogino S, Gulley ML, den Dunnen JT, Wilson RB, , Association for Molecular Patholpogy Training,
828 Committtee E. Standard mutation nomenclature in molecular diagnostics: practical and
829 educational challenges. J Mol Diagn. 2007 Feb;9(1):1–6.

830

831 **TABLES**

832

833 **Table 1 – Summary statistics for sequence variation in *SmSULT-OR*.**

Location	Sample size	Parasite stage sampled	Host	Sequence	Notes[b]
South America (Brazil)[a]	189	Miracidia	Human	Sanger sequence of *SmSULT-OR* exons	1-11 miracidia from each of 49 people from single village (13)
West Africa (Niger/Senegal)	25/10	Miracidia	Human	Exome sequence	1 miracidia from each person
East Africa (Tanzania)	57	Miracidia	Human	Exome sequence	1 miracidia from each person
Middle East (Oman)	112	Miracidia/adult worms/cercariae	Human/Rat/Snail	Exome sequence	1-53 parasite genotypes analysed per host from 11 people, 2 rats and 26 snails from 8 locations
Caribbean (Puerto Rico)[a]	1	Adult worms	Hamster	Genome sequence	OXA-resistant parasite used in a genetic cross (10)

834

835 [a] published data

836 [b] see supplementary table 2 for detailed information on sample locations

837
838
839
840 **Table 2 – Summary statistics for sequence variation in *SmSULT-OR*.**

Location	Nucleotide diversity ($\pi \pm$ standard deviation)	Dn	Pn	Ds	Ps	N	S	pN/pS	Ka/Ks
South America (Brazil)[a]	$2.1 \pm 0.3 \times 10^{-4}$	17	4	23	0	598.19	163.81	N/C	0.202
West Africa (Senegal / Niger)	$7.2 \pm 1.5 \times 10^{-4}$	19	9	23	2	600.93	164.07	1.229	0.225
East Africa (Tanzania)	$22.8 \pm 2.3 \times 10^{-4}$	18	19	23	9	593.80	162.20	0.577	0.214
Middle East (Oman)	$247.5 \pm 3.4 \times 10^{-4}$	10	15	11	36	598.80	160.20	0.111	0.243
Overall	$148.1 \pm$ N/C $\times 10^{-4}$	10	39	11	40	583.61	157.39	0.263	0.245

841
842 [a] Data from Brazil from a previous study (13) added for comparison.
843 Dn, Ds: Number of fixed non-synonymous and synonymous differences between species (*S. mansoni* and *S. rodhaini*)
844 Pn, Ps: Number of polymorphic non-synonymous and synonymous differences within species (*S. mansoni*)
845 N: Number of non-synonymous sites
846 S: Number of synonymous sites
847 pN/pS=(Pn/N)/(Ps/S)
848 Ka/Ks=(Dn/N)/(Ds/S)
849 N/C: not calculable

850 **Table 3 – OXA-R mutations scored in the *Schistosoma mansoni SmSULT-OR* gene and their frequency in the New and Old World.**

Nucleic acid mutation	Amino acid mutation	Mutation frequency			
		South America (Brazil)[a] n=189	West Africa (Senegal / Niger) n=25 / n=10	East Africa (Tanzania) n=57	Middle East (Oman) n=112
c.3dupGTTTATCCATAATG	p.I2VfsX3	-	-	0.0096	-
c.35C>A	p.S12X	-	0.014	-	-
c.103T>C	p.C35R	0.0027	-	-	-
c.358T>C	p.W120R	-	-	0.0510	0.065
c.424_426delGAA	p.E142del	0.008	0.035	0.0098	-
c.510delT	p.N171IfsX28	-	-	0.0625	-
c.536T>C	p.L179P	-	-	0.0089	-
Total frequency		0.0107	0.044	0.1418	0.065

851 The code used for nucleic acid mutations indicates the sequence type (c = coding), the position, and the mutation type (X>Y = substitution of X
852 by Y, insN = insertion of N, delN = deletion of N, dup = duplication). The code used for amino acid mutations indicates the sequence type (p =
853 protein), the reference amino acid, the position, and finally the alternative amino acid, and when frame shift (fs) occurs, the position of the stop
854 codon (X) after the mutation. For details about the nomenclature, see Ogino *et al.* (79).
855 [a] Data from Brazil were obtained in a previous study (13) and added for comparison.

856 **FIGURES**

857 **Fig. 1 – Drug resistance evolution from standing variation or *de novo* mutation**. Drug
858 resistance alleles spread rapidly and reach fixation with a high probability when drug resistance
859 alleles are already present as standing variation. In contrast, if resistance alleles are absent
860 when treatment is initiated, resistance mutation must arise *de novo*, so there is a waiting time
861 before the resistance alleles appear, and most resistance alleles are lost by genetic drift, so the
862 probability of establishment and fixation is low.

863

864 **Fig. 2 – Mapping of the resistance mutations on the gene sequence and structure of**
865 ***Schistosoma mansoni* SmSULT-OR sulfotransferase.** Exon 1 and exon 2 are represented in
866 orange and beige, respectively. Single nucleotide polymorphisms and insertion/deletion events
867 are represented in cyan and magenta, respectively. (A) Linear representation of the *SmSULT-OR*
868 gene showing the relative position of the mutations and their translation in amino acid
869 sequences. (B) Positions of mutations on the SmSULT-OR protein. Oxamniquine is represented
870 in yellow, 3'-phosphoadenosine-5'-phosphosulfate (PAPS) co-factor is represented in green.

871

872 **Fig. 3 – Enzymatic activity of recombinant *Schistosoma mansoni* SmSULT-OR sulfotransferase**
873 **expressed from different allelic variants.** This *in vitro* oxamniquine activation assay quantifies
874 DNA-oxamniquine complexes by scintillation (counts per minute). (A) Bars show the mean of
875 three replicates, while error bars are S.E.M. (B) The triangular matrix shows the p-values from
876 the pairwise comparisons (Tukey's HSD, $p < 0.05$). The color is proportional to the level of
877 significance, from white (not significant) to red (most significant). Enzyme carrying known loss-
878 of-function mutations, such as p.C35R or p.E142del, as well as a newly identified variant
879 (p.L179P) showed no or low oxamniquine activation, while two newly identified variants
880 (p.S160L and p.P225S) showed intermediate activation. The newly identified p.P106S did not
881 impair oxamniquine activation.

882

883 **Fig. 4 – *In silico* evaluation of mutations on protein stability.** We computed the difference in
884 free enthalpy ($\Delta\Delta G$) between the mutated and the wild type proteins. The higher the $\Delta\Delta G$, the
885 more unstable is the mutated protein. Only single amino acid changes within the resolved
886 crystal structure were examined. For completeness, we included mutations from our current
887 dataset and from previous studies (10,13). Grey bars correspond to known sensitive alleles. Red
888 bars correspond to validated resistant alleles. Grey labels correspond to mutations identified
889 previously from South America (13).

890

891 **Fig. 5 – World map showing proportion of resistance and sensitive alleles in South America**
892 **(Brazil), West Africa (Senegal and Niger), East Africa (Tanzania) and Middle East (Oman).**
893 Number of homozygous (hmz) and heterozygous (htz) parasites for resistant alleles, and total
894 number of parasites sampled (n) are shown below the pie charts. East Africa showed the

895 highest frequency of resistant alleles and resistant parasites. Data from Brazil from a previous
896 study (13) was added for comparison.

897

898 **Fig. 6 – Common origin of the p.E142del mutation in the Old and New World.** (A) Haplotype
899 variation in all the samples bearing p.E142del from Caribbean (HR9), Niger (NE) and Tanzania
900 (TZ) across a 102.5 kb region of chr. 6. Each row represents a chromosome, HR9 was used as
901 reference, blank squares reflect HR9 allele state, and black squares correspond to the
902 alternative allele. Relative bp position to p.E142del (0 bp) is shown on the x-axis. The first and
903 last variants showed on the block correspond to the break of the haplotype block. The
904 Caribbean sample (HR9) and a Nigerien sample (Sm.NE_Di158.1) share an identical haplotype
905 block of 102.5 kb. (B) Minimum spanning network of 410 haplotypes of the 102.5 kb region
906 previously identified. The network was built using 399 bi-allelic variants. Node size is
907 proportional to sample size (smallest node: n=1; biggest node n=50). Nodes with samples
908 carrying p.E142del are circled in blue. Caribbean and West African haplotypes carrying
909 p.E142del clustered together indicating a common origin. The p.E142del from East Africa has
910 different flanking haplotypes.

911

912 **Fig. 7 – Impact of starting allele frequency on OXA-R allele change.** (A) Monte Carlo simulation
913 over 1,500 generations using strong selection (selection coefficient (*s*) = 0.1) on a population
914 size (N) of 65,000 with a starting allele frequency (p(R)) of 0.15 for standing variation or 1/(2N)
915 for new mutation. These simulations underestimate time to fixation for true *de novo* mutation,
916 because we do not account for the waiting time for resistance to appear which is dependant on
917 the rate of drug resistance mutations and N_e (3). The starting frequency of 0.15 corresponds to
918 the frequencies of OXA-resistance alleles observed in Kenya. (B) Change in frequency of
919 resistant parasites (i.e. homozygotes for OXA-R alleles) from standing variation under a range of
920 selection coefficients. The dashed line corresponds to the two thresholds (10% and 20%) at
921 which treatment efficacy would be compromised. The numbers at the dashedlines correspond
922 to the parasite generations needed to cross these thresholds. These predictions are
923 deterministic, because we expect minimal stochastic variation when starting resistance allele
924 frequencies are high.

925 **SUPPLEMENTARY TABLES**

926

927 **Supp. table 1 – Library sequencing statistics.**

928

929 **Supp. table 2 – Mutations scored in samples from West Africa (Senegal and Niger), East Africa**
930 **(Tanzania) and Middle East (Oman).**

931

932 **Supp. table 3 – Sampling information for the schistosome material used.**

Standing variation

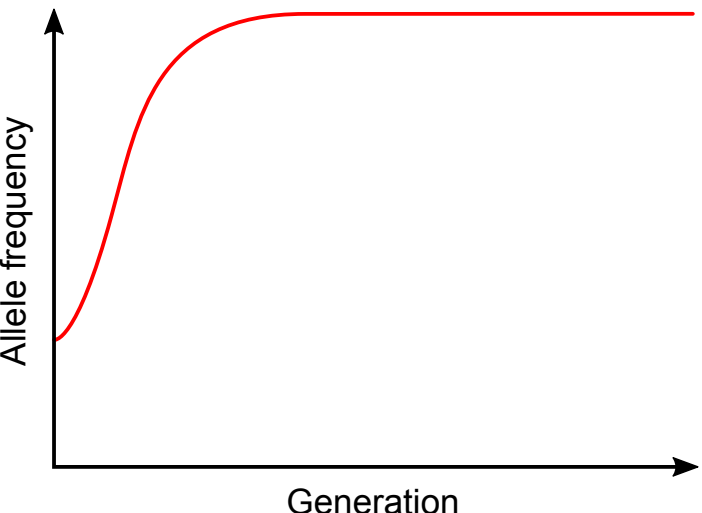
Mutation **present** before selection

Starting allele frequency = p(resistant allele)

→ **No** waiting time

Drug resistance evolves rapidly

High probability of fixation



De novo variation

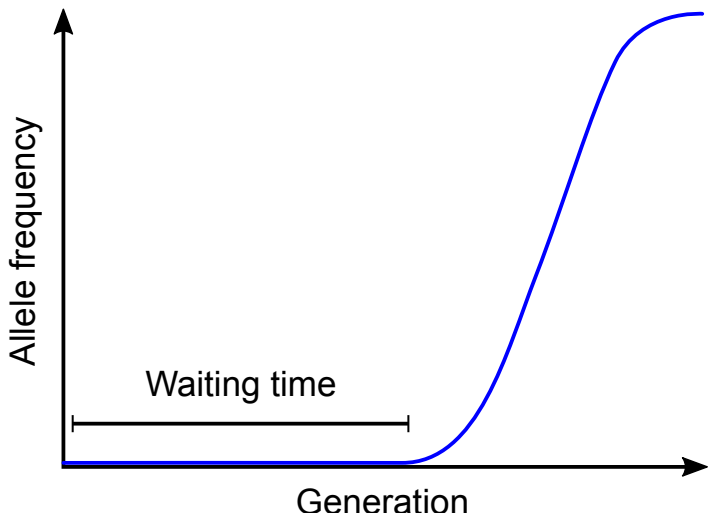
Mutation **not present** before selection

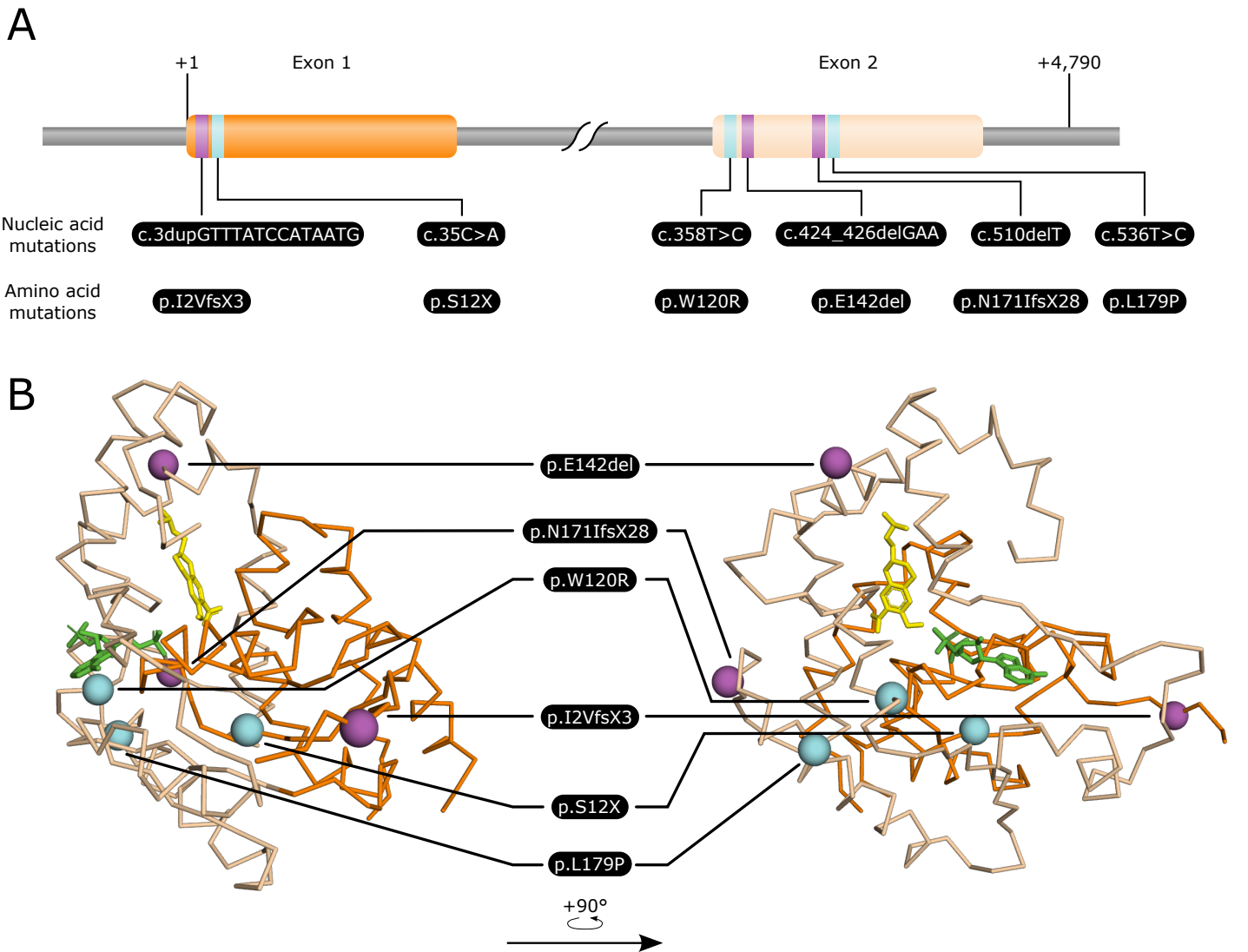
Starting allele frequency = $1/(2N_e)$

→ Waiting time

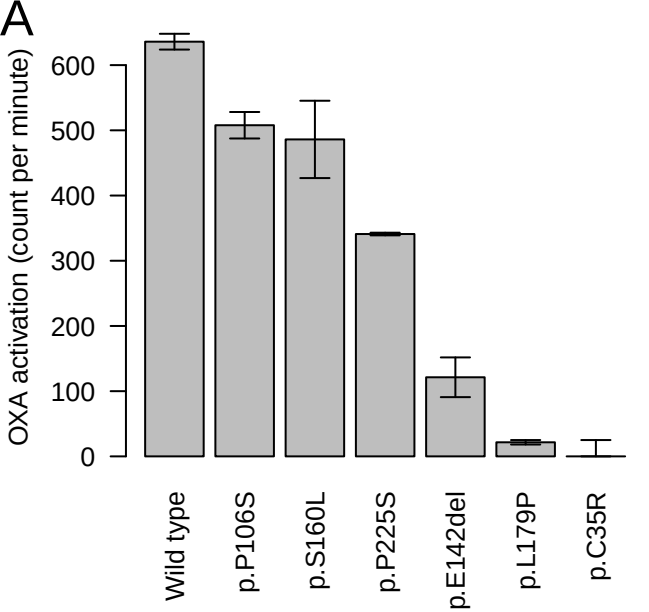
Drug resistance evolves slowly

Low probability of fixation



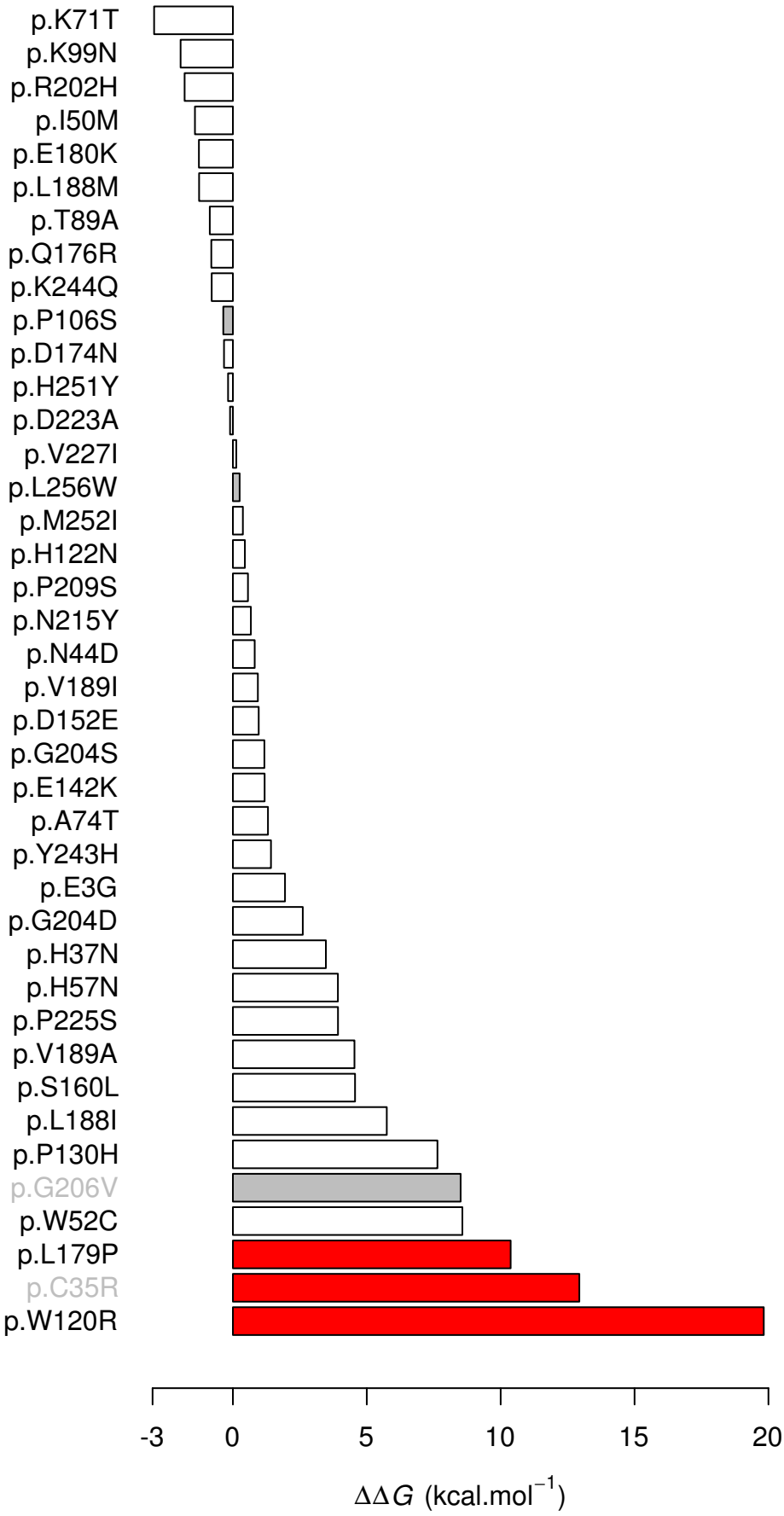

A
+1
Exon 1
Exon 2
+4,790
Nucleic acid mutations
c.3dupGTTTATCCATAATG
c.35C>A
c.358T>C
c.424_426delGAA
c.510delT
c.536T>C
Amino acid mutations
p.I2VfsX3
p.S12X
p.W120R
p.E142del
p.N171IfsX28
p.L179P
B
p.E142del
p.N171IfsX28
p.W120R
p.I2VfsX3
p.S12X
p.L179P
+90°

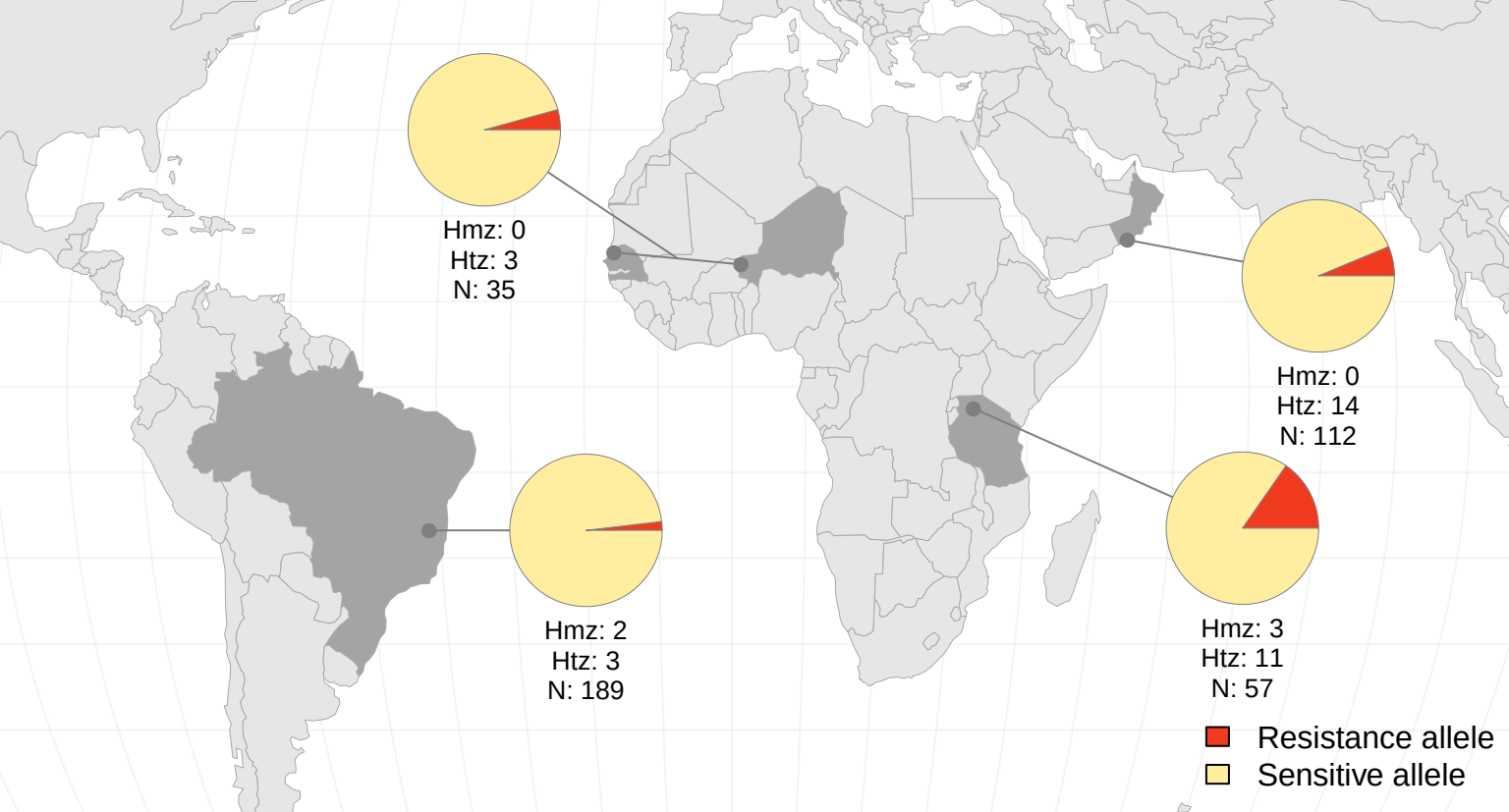
A

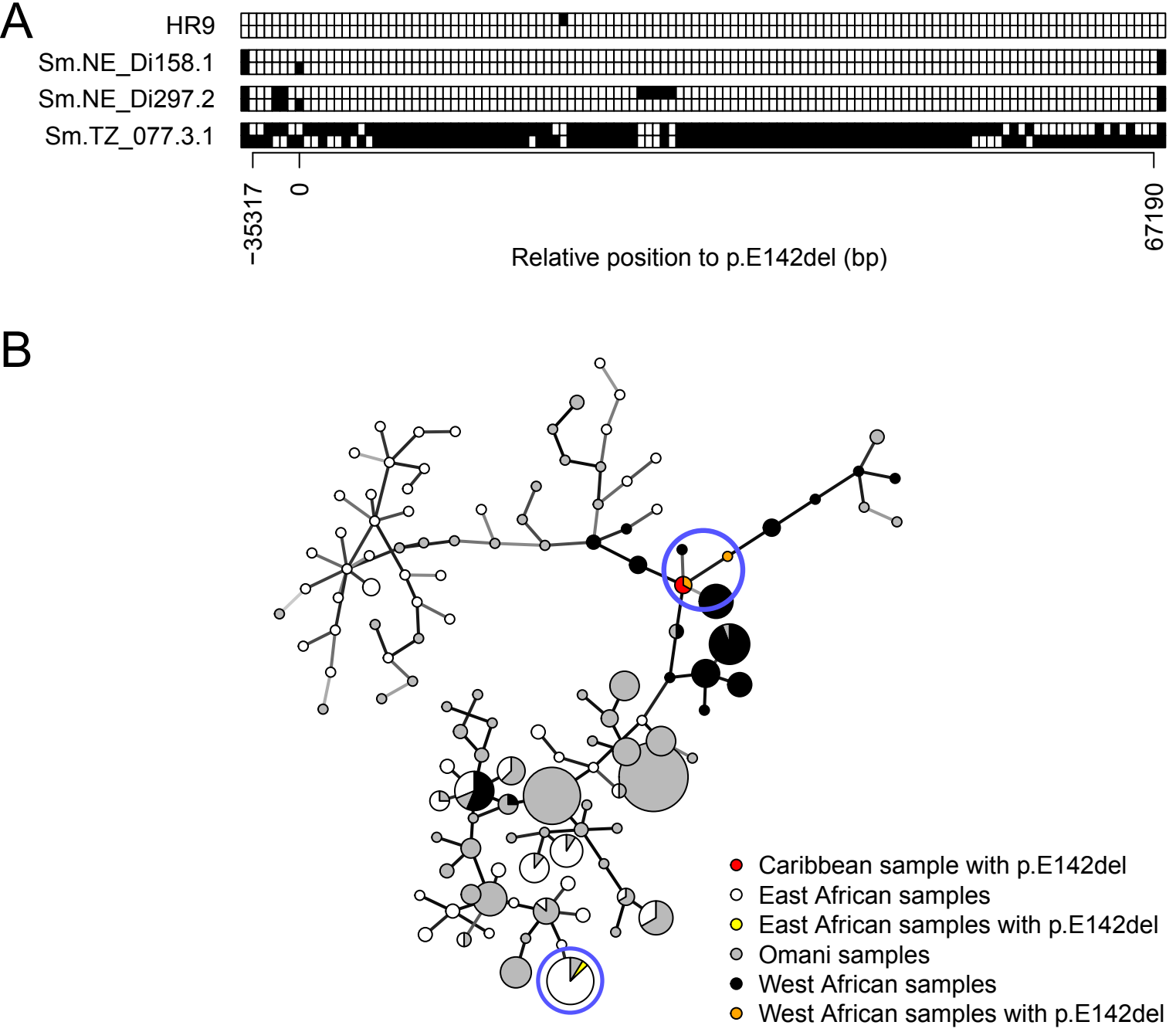


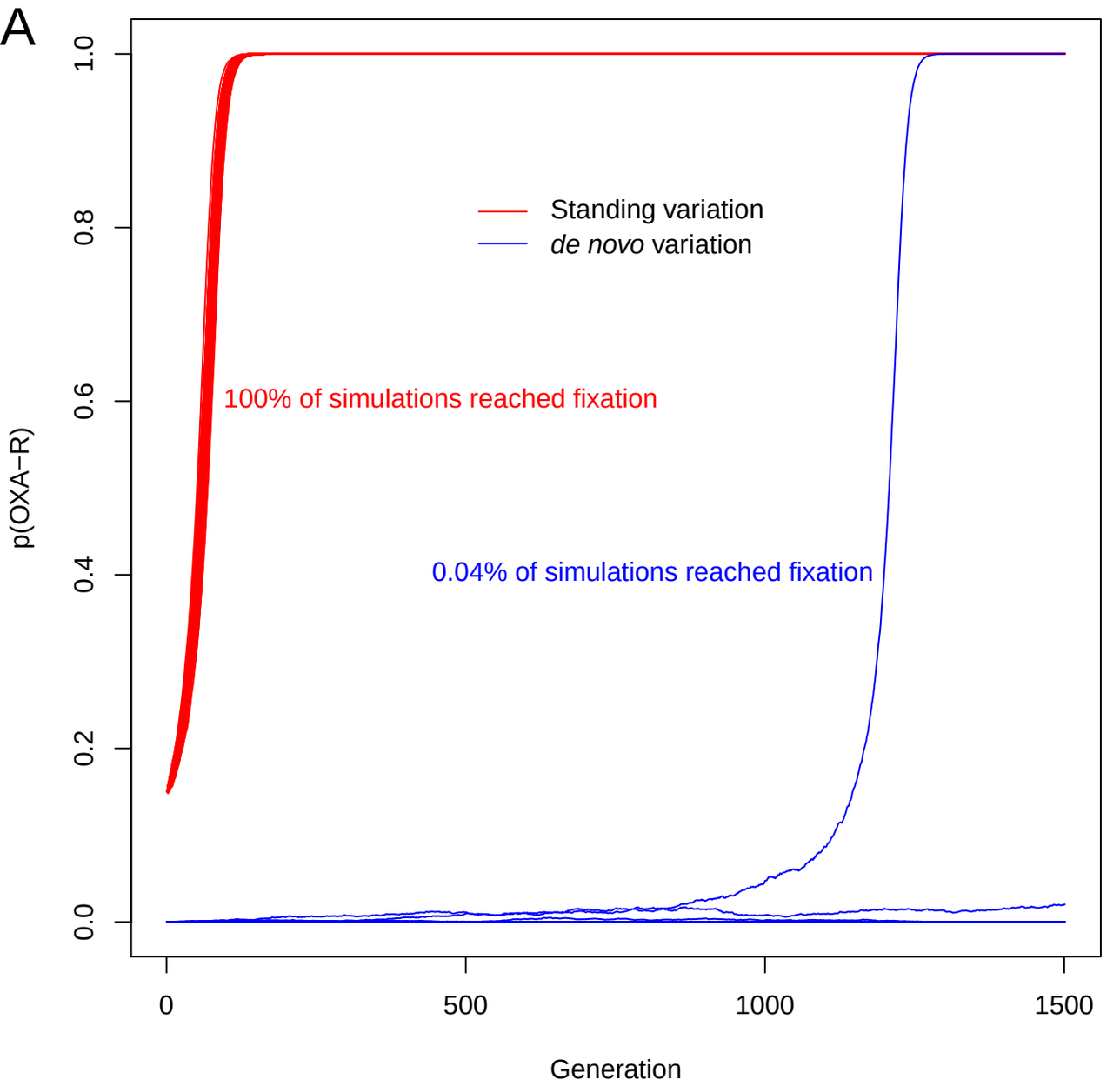
B

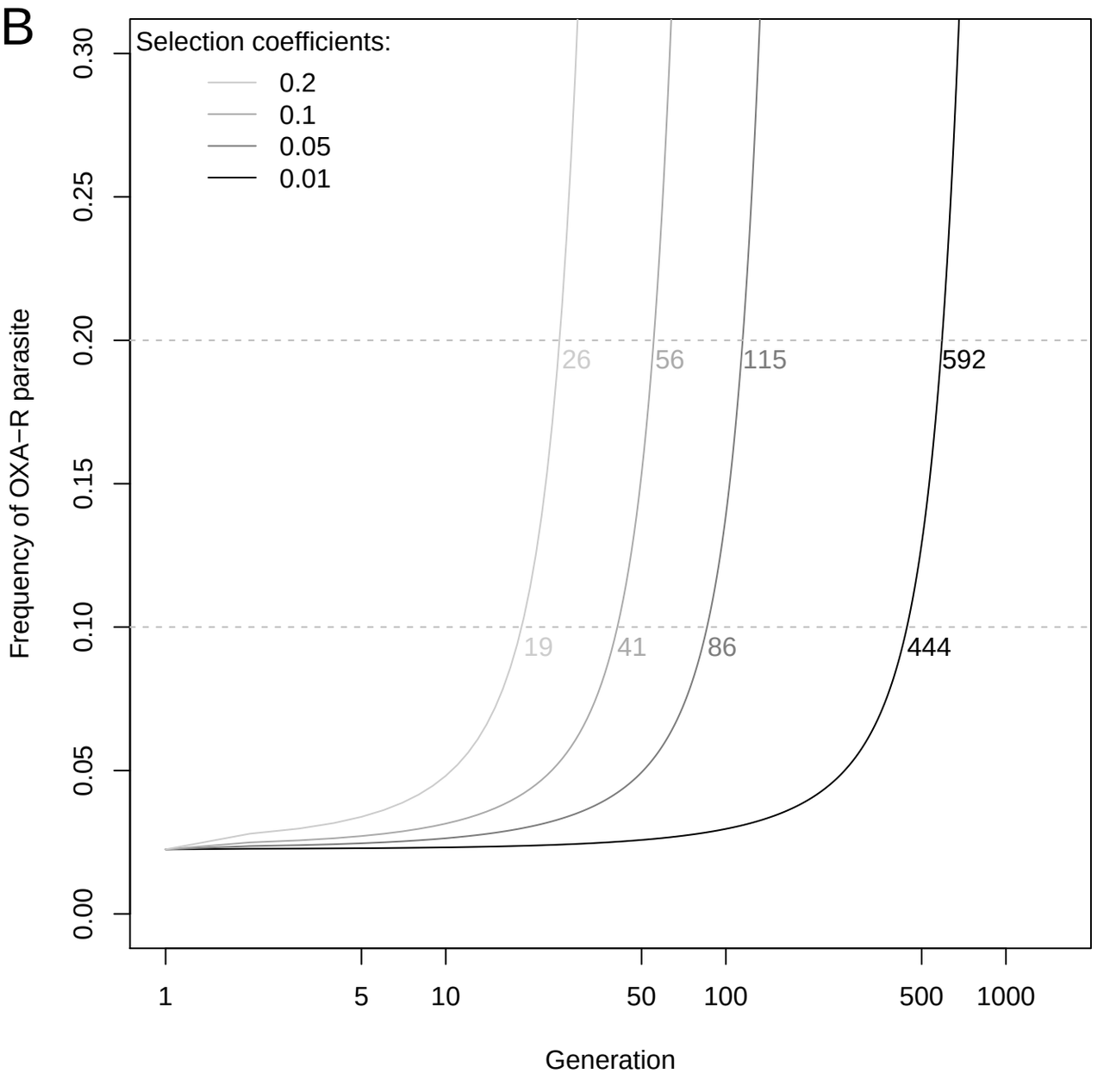
	Wild type	p.P106S	p.S160L	p.P225S	p.E142del	p.L179P
p.C35R	1.0e−09	1.5e−08	2.5e−08	1.2e−06	4.4e−03	3.2e−01
p.L179P	6.6e−09	1.4e−07	2.6e−07	2.4e−05	2.4e−01	
p.E142del	6.9e−08	2.5e−06	5.1e−06	1.2e−03		
p.P225S	5.9e−05	1.3e−02	3.5e−02			
p.S160L	2.8e−02	1.0e+00				
p.P106S	7.4e−02					


p.K71T
p.K99N
p.R202H
p.I50M
p.E180K
p.L188M
p.T89A
p.Q176R
p.K244Q
p.P106S
p.D174N
p.H251Y
p.D223A
p.V227I
p.L256W
p.M252I
p.H122N
p.P209S
p.N215Y
p.N44D
p.V189I
p.D152E
p.G204S
p.E142K
p.A74T
p.Y243H
p.E3G
p.G204D
p.H37N
p.H57N
p.P225S
p.V189A
p.S160L
p.L188I
p.P130H
p.G206V
p.W52C
p.L179P
p.C35R
p.W120R
-3
0
5
10
15
20
ΔΔG (kcal.mol^{-1})


Hmz: 0
Htz: 3
N: 35
Hmz: 0
Htz: 14
N: 112
Hmz: 2
Htz: 3
N: 189
Hmz: 3
Htz: 11
N: 57
Resistance allele
Sensitive allele


A
HR9
Sm.NE_Di158.1
Sm.NE_Di297.2
Sm.TZ_077.3.1
−35317
0
67190
Relative position to p.E142del (bp)
B
Caribbean sample with p.E142del
East African samples
East African samples with p.E142del
Omani samples
West African samples
West African samples with p.E142del


A
Standing variation
de novo variation
100% of simulations reached fixation
0.04% of simulations reached fixation
p(OXA−R)
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
Generation


B
Selection coefficients:
0.2
0.1
0.05
0.01
26
56
115
592
19
41
86
444
Frequency of OXA−R parasite
0.00
0.05
0.10
0.15
0.20
0.25
0.30
1
5
10
50
100
500
1000
Generation