

UC Irvine

UC Irvine Previously Published Works

Title

Intraspecific Transcriptome Variation and Sex-Biased Expression in *Anopheles Arabiensis*
(vol 13, evab199, 2021)

Permalink

<https://escholarship.org/uc/item/09d4z2qz>

Journal

GENOME BIOLOGY AND EVOLUTION, 14(2)

ISSN

1759-6653

Authors

Jayaswal, Vivek
Ndo, Cyrille
Ma, Hsiu-Ching
[et al.](#)

Publication Date

2022-02-04

DOI

10.1093/gbe/evab285

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License,
available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Intraspecific transcriptome variation and sex-biased expression in *Anopheles arabiensis*

Vivek Jayaswal ^{1,†}, Cyrille Ndo ^{2,†,‡}, Hsiu-Ching Ma ³, Bryan Clifton ³, Marco Pombi ⁴, Kevin Cabrera ³,
Anna Couhet ⁵, Karine Mouline ⁵, Abdoulaye Diabaté ², Roch Dabiré ², Diego Ayala ⁵, José M. Ranz ^{3,*}

¹ School of Mathematics and Statistics, The University of Sydney, Sydney, NSW, Australia

² Institut de Recherche en Sciences de la Santé, Bobo-Dioulasso, Burkina Faso

³ Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, California, USA

⁴ Dipartimento di Sanità Pubblica e Malattie Infettive, University of Rome "Sapienza", Rome, Italy

⁵ Unité Mixte de Recherche Maladies Infectieuses et Vecteurs: Ecologie, Génétique, Evolution et Contrôle,
University of Montpellier, CNRS, IRD, Montpellier, France

[‡] Current address: Centre for Research in Infectious Diseases, Wellcome Trust, Yaoundé, Cameroon

^{*} Corresponding author: José M. Ranz (jranz@uci.edu)

[†] These authors contributed equally

Running Title: patterns of transcriptome variation in anophelines

ABSTRACT

The magnitude and functional patterns of intraspecific transcriptional variation in the anophelines, including those of sex-biased genes underlying sex-specific traits relevant for malaria transmission, remain understudied. As a result, how changes in expression levels drive adaptation in these species is poorly understood. We sequenced the female, male, and larval transcriptomes of three populations of *Anopheles arabiensis* from Burkina Faso. One-third of the genes were differentially expressed between populations, often involving insecticide resistance-related genes in a sample type-specific manner, and with the females showing the largest number of differentially expressed genes. At the genomic level, the *X* chromosome appears depleted of differentially expressed genes compared to the autosomes, chromosomes harbouring inversions do not exhibit evidence for enrichment of such genes, and genes that are top contributors to functional enrichment patterns of population differentiation tend to be clustered in the genome. Further, the magnitude of variation for the sex expression ratio across populations did not substantially differ between male- and female-biased genes, except for some populations in which male-limited expressed genes showed more variation than their female counterparts. In fact, female-biased genes exhibited a larger level of interpopulation variation than male-biased genes, both when assayed in males and females. Beyond uncovering extensive adaptive potential of transcriptional variation in *An. arabiensis*, our findings suggest that the evolutionary rate of changes in expression levels on the *X* chromosome exceeds that on the autosomes, while pointing to female-biased genes as the most variable component of the *An. arabiensis* transcriptome.

KEYWORDS

Transcriptome variation, functional diversification, sex-biased gene expression, faster-*X* effect, *Anopheles arabiensis*

46
47
48
49
50
51
52
53
54
55
56
57
58

SIGNIFICANCE

Transcriptome evolution in the anophelines has been primarily studied at the interspecific level, paying little attention to the patterns of interpopulation differentiation. As a result, our understanding about intraspecific variation in expression levels, particularly at different life stages, in relation to chromosome organization, and in relation to sex-biased expression, remains limited. We present a portrait of the transcriptome variation in *An. arabiensis*, an increasingly concerning malaria vector. Our results uncover the patterns of functional differentiation underlying expression variation at different life stages, expose a non-random distribution of gene expression polymorphism at the chromosome and fine scales, and reveal substantial differences in the patterns of interpopulation expression polymorphism between female- and male-biased genes.

INTRODUCTION

In Africa, three members of the species complex *Anopheles gambiae sensu lato* (s.l.), *An. gambiae* s.s., *An. coluzzii*, and *An. arabiensis*, are considered major vectors of transmission of the malaria parasite *Plasmodium* (Barron et al. 2019; White et al. 2011). As disease vectors, these species impose a substantial burden on human health with ~215 million cases and ~384,000 casualties just in 2019 (WHO 2020). Malaria transmission is shaped by multiple genetic and ecological factors, including: the immune response of these vectors against the *Plasmodium* parasite, which in turn affects vector reproduction and longevity (Mitchell et al. 2015); standing genetic variation in insecticide resistance (Lucas et al. 2019; Main et al. 2016) and differential survival during the larval and adult stages (Roux and Robert 2019); and access to feeding resources (Barredo and DeGennaro 2020).

Organismal traits relevant for understanding malaria transmission in the anophelines, such as reproduction and longevity, are interrelated through a partially shared genetic basis (Lawniczak et al. 2007; Mitchell and Catteruccia 2017; Rono et al. 2010). Large titers of the hormone 20-hydroxecdysone (20E) generated by the male accessory glands are transferred to the female during mating, inducing oogenesis, egg laying, and preventing further mating (Baldini et al. 2013; Gabrieli et al. 2014). This monandrous behaviour enhances female longevity (Dao et al. 2010), favouring indirectly the completion of the *Plasmodium* sporogony, i.e. the parasite multiplication within the mosquito. This lipid-transporter mediated male-enhanced oogenesis reduces the mosquito immune response against *Plasmodium*, further facilitating its development in the female (Rono et al. 2010). Equally important, in the anthropophilic anopheline species, blood-feeding triggers two temporally and physiologically tied process: oogenesis and *Plasmodium* sporogony (Mitchell and Catteruccia 2017). Further, longevity in the anophelines exhibits a complex interplay with insecticide resistance (Rivero et al. 2010), while survival is strongly associated with desiccation resistance (Ayala et al. 2019), in both cases affecting malaria transmission. All these organismal traits are susceptible to both inter- and intra-specific adaptive evolution as abundant naturally occurring genetic variation exists (Ayala et al. 2019; Main et al. 2016; Neafsey et al. 2010). Nevertheless, the functional consequences of this intraspecific

genetic variation in relation to organismal traits that impact malaria transmission remain poorly characterized in the anophelines.

An important aspect of the anopheline biology is its unique sex-specific component as females are not only mainly monandrous (Tripet et al. 2003; Yuval 2006), but are also the only blood-feeding sex (Barredo and DeGennaro 2020; Carnevale and Robert 2009; Yuval 2006). This could impact the rate of evolution of the above indicated organismal traits as they could be targeted by sex-specific natural and sexual selection (Darwin 1871; Ellegren and Parsch 2007). For example, in *An. arabiensis*, all the 20E content generated by the male accessory glands is transferred to the female during copulation while only half of the total 20E is transferred by other anophelines males (Mitchell et al. 2015). The genomic signature of this sex-specific fast evolution has been explored at the interspecific level, finding that the magnitude of divergence in expression levels, the rate of evolution of protein-coding genes, and the rate of gene turnover, are substantially higher for genes with female-biased expression than for male- and non-sex-biased genes (Papa et al. 2017). These patterns are in stark contrast to previous findings in other Diptera and animal lineages (Grath and Parsch 2016), being this difference interpreted as the outcome of an enhanced selection intensity operating on females coupled with ameliorated male-male competition at the post-copulatory level as a result of female monandry. Whether similar patterns are reproduced at the intraspecific level has not been explored.

An. arabiensis is the most common malaria vector in the driest parts of Africa (Main et al. 2016; White et al. 2011), and it is expanding its range due to climate change (Bimbile Somda et al. 2018; Dabire et al. 2014). In this species, our understanding of patterns of naturally occurring genetic variation is primarily confined to SNP and chromosomal inversion surveys (Ayala et al. 2017; Fontaine et al. 2015; Main et al. 2016), while functional genomic information derives primarily from interspecific studies based on a single laboratory strain from Dongola, Sudan (Izquierdo et al. 2019; Neafsey et al. 2015; Papa et al. 2017). To our knowledge, only one study has tried to decipher the genome-wide transcriptional patterns that arises from interpopulation genetic variation (Simma et al. 2019), and thus far no study has tried to investigate the functional patterns associated with this genetic variation while taking into account sex-specific aspects of the anopheline biology. As a result, we lack a basic understanding about the magnitude of this subset of

functional variation upon which selective mechanisms can operate and how it might be linked to the evolution of key traits that contribute to the role of *An. arabiensis* as a malaria vector and its ability to thrive in the most arid areas of Africa.

Using RNA-sequencing, we dissected the patterns of transcriptome variation in females, males, and larvae across three populations from Burkina Faso exposed to different types of savanna climate. Notably, in Burkina Faso, one of the countries most affected by malaria burden and deaths (WHO 2020), insecticide resistance in *An. arabiensis* has been increasingly documented (Jones et al. 2012a). This resistance can arise in different ways (Feyereisen et al. 2015; Ranson et al. 2002), in some cases in association with gene expression changes, typically in the form of overexpression of detoxifying enzyme-encoding genes leading to metabolic resistance (Ibrahim et al. 2016), and expression changes in genes relevant for cuticle composition leading to penetration resistance (Jones et al. 2013). Here, we aimed at: *i*) determining the magnitude and patterns of interpopulation variation in transcript abundance for *An. arabiensis* in Western Africa; *ii*) uncovering the functional attributes of differentially-expressed genes most strongly impacted by population differentiation, particularly in relation to metabolic and penetrance insecticide resistance; *iii*) testing for the existence of hot spots of transcriptional differentiation in the *An. arabiensis* genome, *e.g.* in regards to inversion polymorphism, which could favour the accumulation of transcriptional differences in particular chromosomes or genomic regions (Marques-Bonet et al. 2004); and *iv*) characterizing how genes with different sex biases might differ in their level of expression polymorphism such that we can determine the extent to which interspecific patterns of divergence (Papa et al. 2017) are reproduced at the intraspecific level and explore how gene expression variation is impacted by sex in the anophelines. Through this population framework we have generated a comprehensive portrait of the patterns of interpopulation functional variation in *An. arabiensis*, which should accelerate our understanding about the potential for adaptation and fast evolution in this vector species (Lai et al. 2019; Powell 2018), while providing a more nuanced picture of anopheline transcriptome evolution.

RESULTS

Field work and colony construction

We collected gravid females from rural areas of Nanoro, Bobo, and Kodené not overly exposed to insecticides. These three sites are located across two different ecological and climate zones of Burkina Faso (West Africa; fig. 1; Material and Methods). Subsequently, we constructed colonies by first molecularly excluding the presence of *An. coluzzi* and *An. gambiae* (supplementary fig. S1), and then by enriching through particular mating schemes for the presence of the most frequent chromosomal inversions according to cytological squashes, *i.e.* the 2Rb in all the colonies and the 2Ra and 3Ra in Kodené and Nanoro, respectively (supplementary fig. S2 and table S1; Material and Methods). This approach resulted in three stable colonies, one per collection site, virtually fixed for particular combinations of chromosomal inversions (supplementary table S1). Hereafter, we will refer to these colonies as populations.

Sequencing results

To ameliorate any unintended effect from long-term laboratory adaptation (Bonizzoni et al. 2012), RNA-seq data were generated shortly after colony construction. We sequenced in triplicates the transcriptome of L4 larvae, naïve males, and virgin nonblood-fed females. In the case of the larvae, their sex was unknown, nevertheless, the probability of including skewed biological replicates due to sex bias was low and therefore the impact in downstream analysis negligible (supplementary note S1). In the case of adult individuals, the choice of using only virgins was to minimize allometric differences between the sexes (Baker and Russell 2011; Jiang et al. 2015). Approximately 523 million sequencing reads (58.2 ± 10.58 ; mean $\pm$ SD per sample type and population) were used in subsequent downstream analyses. From those, $88 \pm 15.6\%$ per sample type and population mapped to the genome assembly of *An. arabiensis*. Mapping efficiency did not differ by collection site (Wilcoxon rank-sum test, $P = 0.701$), and was strongly correlated between pairs of biological replicates for the sample type and collection site (Pearson's $r = 0.929-0.997$). Upon discarding any batch effect (supplementary fig. S3), multidimensional scaling analysis additionally showed a perfect clustering of

the biological replicates per population for each of the sequenced sample types, confirming the reliability and reproducibility of the sequencing results (supplementary fig. S4).

Interpopulation transcriptional variation

The total number of genes with non-zero read counts in at least one sample was 13,657 (99.8%). However, a gene was considered to be expressed only if there was evidence of expression (≥ 1 cpm) in ≥ 2 biological replicates of at least one of the particular sample type by population combinations, *i.e.* a biological group, surveyed. Under these criteria, 12,575 genes (91.9% of all the genes surveyed) were considered in further analyses (supplementary table S2). First, we performed three pairwise population contrasts per sample type (supplementary table S3). At a 1% FDR, fold-change (FC) ≥ 2 , a minimum of 1,150 (9.1%) and a maximum of 2,304 (18.3%) genes showed differential expression in larvae and females, respectively. In total, across sample types, 4,035 genes were differentially expressed between populations, *i.e.* 32.1% of all annotated genes, of which, 76.2% showed sample type-specific expression (supplementary fig. S5 and table S4).

We next examined the commonalities and differences in expression differentiation across population contrasts, which were reproducible at different minimum expression thresholds (fig. 2A; supplementary figs. S6-S7). Females exhibit the largest magnitude of interpopulation differentiation, which is clearly linked to the Kodení population. This population shows the largest number of differentially expressed genes (1,813) in relation to Bobo. The differentiation of the Kodení female transcriptome from those from Bobo and Nanoro involves 1,063 common genes, which represent more than 50% of the subtotal in each pairwise contrast (58.6% and 68.3% for the Kodení-Bobo contrast and the Kodení-Nanoro contrasts, respectively) (fig. 2B). For the male and larval transcriptomes, the contrasts involving Nanoro entail the most differentially expressed genes. According to these levels of expression polymorphism, for males and larvae but not for females, the mode of interpopulation differentiation in mRNA abundance seems to mirror more closely the ecotype zone to which the collection site belongs (fig. 2C). Lastly, we examined whether there was any global bias in the directionality of the differential expression, *i.e.* towards over- or underexpression, across population contrasts and sample types. We found evidence of such bias at $P_{\text{adj}} < 0.01$ in 6 out of the 9 contrasts (supplementary fig. S7), although no consistency in the directionality of the trend was found for the

same population contrast across sample types. In sum, we documented an overall lack of consistency in magnitude, mode, and directionality in the patterns of differentiation in mRNA abundance across sample types and population contrasts.

Nonuniform interpopulation functional diversification across sample types

To understand the patterns of functional differentiation associated with changes in gene expression across populations and sample types, we tested for enrichment in Gene Ontology (GO) terms and pathways among the differentially expressed genes in each population contrast (supplementary table S6; supplementary note S2). For females, the populations from Bobo and Kodené showed the largest level of differentiation across functional rubrics (supplementary table S7). In contrast, for males, Nanoro exhibited the most pronounced pattern of functional differentiation, both in relation to Kodené and Bobo. For larvae, no consistent pattern of functional differentiation emerged across population contrasts. Overall, we found very limited overlap in trends of functional differentiation both among sample types within the same population and among populations for the same sample type (fig. 3A). These patterns were essentially reproduced when performing a complementary analysis focused on genes found overexpressed in one population in relation to the other two (supplementary table S5), and when considering only the 161 genes that contribute to the statistical enrichment across multiple functional terms, hereafter *top contributors* (fig. 3B; supplementary table S8 and note S3; Material and Methods).

When specific GO terms and pathways were considered, several related to protein metabolism stood out as being enriched across two or more population contrasts both in males and larva (supplementary fig. S8-S9). Also in males, enrichment for pathways and GO terms related to multiple detoxification processes underlying metabolic and penetration resistance was very common across population contrasts. For example, we found multiple genes encoding cytochrome P450s such as AARA015789 (*CYP6P4*), which has been shown to underlie *An. arabiensis* resistance against several pyrethroids in Central Africa (Ibrahim et al. 2016), and glutathione S-transferases, such as AARA015764 (*Gstd3*), which has been shown to underlie resistance of *An. arabiensis* to DDT in Burkina Faso (Jones et al. 2012b). We also detected genes encoding proteins relevant for cuticle composition such as AARA003376. This gene is involved in the

detoxification of bisphenol, a compound known for interfering in the tanning and sclerotization reaction in mosquito (Sacher 1971), and also in that of repellents such as nicotine and limonene (Huong et al. 2019). Although not part of enriched functional terms, we also found genes recently proposed to be candidates of insecticide resistance in *An. gambiae* such as AARA016114 (*hexamerin*) and AARA002269 (*α-crystallin*) (Ingham et al. 2018). Notably as well, 31.4% (126/401) of the differentially expressed genes at 1% FDR and FC>2 between Ethiopian resistant populations of *An. Arabiensis* and other populations not resistant DDT and the pyrethroid deltamethrin were also found to be polymorphic in expression across Burkina Faso populations (Simmam et al. 2019). In good agreement with the general patterns described above, most of these differentially expressed genes were not consistent across sample types (supplementary fig. S10).

Transcriptional variation across the genome

For each sample type, we sought to establish whether genes polymorphic in expression were randomly distributed across the five chromosomal elements of *An. arabiensis*: the X chromosome, and the left and right arms of the autosomes 2 and 3 (2L and 2R; 3L and 3R) (fig. 4A). We found a significant deficit of differentially expressed genes on the X chromosome across sample types, being more robust in larvae as such deficit was reproduced across all population contrasts (supplementary table S9); for females, two population contrasts, and for males, one population contrast showed this reduction. Additionally, in the case of females, we also detected an enrichment for differentially expressed genes on chromosome 3R (supplementary table S9).

Alternative chromosomal arrangements generated by inversions are known to facilitate functional divergence (Marques-Bonet et al. 2004). The populations of *An. arabiensis* analyzed carry the inversions 2Ra and 3Ra (supplementary table S1), allowing us to investigate any possible transcriptional effect of such inversions. Although we could not test precisely whether expression polymorphism is associated with the indicated inversions, as their breakpoints have not been mapped molecularly, none of our analyses identified the chromosomes carrying the indicated inversions as a preferential source of interpopulation polymorphism in gene expression. This result differs markedly from what has been reported in other Diptera

(Fuller et al. 2016; Lavington and Kern 2017). Our results provide strong evidence for a chromosome effect on the genome-wide expression polymorphism documented here, but not one mediated by chromosome inversions.

Expression polymorphism can also be non-randomly distributed at a finer scale, and not only at the chromosome-wide level. We tested this possibility across scaffolds, adjusting for their size in number of included genes. Only for females, in the case of the contrast between Kodení and Nanoro, is there evidence of enrichment in the 1,254 kb long scaffold KB704718, which is located on the chromosome 2L. Out of the 57 expressed genes harboured by this scaffold, 16 (30.4%) are differentially expressed in the indicated contrast, although there is no consistency in the population directionality of the expression changes. Therefore, we discarded the possibility of local chromatin-level regulatory differentiation in the genomic region represented by the scaffold KB704718.

Subsequently, we examined how the 165 *top contributor* genes were organized in the genome (fig. 4B). Of them, 154 genes were mapped to chromosomes, showing a small but significant underrepresentation on the X and 3R (two-sided Fisher's exact test, FDR-corrected; X, $P_{\text{adj}} = 0.0132$; 3R, $P_{\text{adj}} = 0.0175$) and a larger deficit on the 3L (two-sided Fisher's exact test, FDR-corrected, $P_{\text{adj}} = 9.53 \times 10^{-5}$). Notably, 60 of these *top contributors* (*i.e.* 36.4%) form part of 22 clusters, defined as groups of at least two differentially expressed genes located less than 10 kb away (supplementary table S8). Collectively, these findings support the conclusion that polymorphic genes underlying multiple functionally enriched trends of population differentiation are not randomly distributed across the *An. arabiensis* genome, but rather they tend to form clusters.

Sex-biased gene expression

We examined the expression ratio between the sexes, *i.e.* $\log_2(\text{male/female})$, from a population perspective. We considered increasing expression differences between the sexes in our gene categorization: moderate sex bias ($2 \leq \text{FC} < 4$); strong sex bias ($\text{FC} \geq 4$); and sex-limited expression (supplementary table S10; Material and Methods). We discovered overt interpopulation differences in the distribution of sex-biased expression

(fig. 5A). These differences are well reflected in the disparity among the overall proportion of genes showing sex bias in gene expression across populations (table 1; three-sample test for equality of proportions without continuity correction, $X^2=1303$, d.f.=2, $P_{adj}=8.8\times 10^{-16}$), which is also apparent when examined in a pairwise fashion (supplementary table S11). Thus, while 41.0% of the genes showed some degree of significant sex bias in Kodení, only 20.4% and 27.8% showed such pattern in Bobo and Nanoro, respectively. These percentages are similar to those in another *An. arabiensis* strain from Dongola, Sudan, in which 34.6% (3,911/11,300) of the genes were also found to be sex-biased in expression (Papa et al. 2017). In addition, the proportion of female-biased genes was greater than that of male-biased genes (three-sample test for equality of proportions without continuity correction, $X^2=33.81$, d.f.=2, $P_{adj}=7.36\times 10^{-8}$), which is significant in two of the three population contrasts (supplementary table S11). This overrepresentation of female-biased genes has also been reported for the Dongola strain, with 2,025 genes being female-biased and 1,886 genes being male-biased (Papa et al. 2017). Overall, Eastern and Western African populations of *An. arabiensis* show strong commonalities relative to sex-biased gene expression.

Given the known plasticity of the transcriptome and sex-biased expression (Grath and Parsch 2016), we inspected the degree of consistency of different expression trends across our populations (fig. 5B; supplementary table S12). We found that 48.8% of the expressed genes consistently show no sex bias, while the rest show evidence of significant sex bias in at least one of the populations. Within this last group of genes, 25.9% are female-biased in at least one of the populations and unbiased in the rest, 23.3% are male-biased in at least one of the populations and unbiased in the rest, and 2% show a mixed-bias pattern (*i.e.* reversal of sex bias across populations). When requiring consistency in the trend, *i.e.* the same sex bias pattern across the three populations, the fraction of genes showing such biases was reduced to 8.2% and 4.3%, for female- and male-biased expression, respectively. This lack of consistency affects even well-known fitness-increasing genes such as AARA017140 or MISO (Mating-induced stimulator of oogenesis; Mitchell et al. 2015), which is categorized as moderately female-biased in Bobo and Kodení but non-sex-biased in Nanoro.

At the interpopulation level, Kodení showed the most differentiated pattern of categorization of sex-biased expression relative to the other two populations (fig. 5C). This interpopulation lability was further underscored when we parsed our categorization against that in the Dongola population (Papa et al. 2017) (supplementary fig. S11, table S13, and note S4). Further, this lability in the categorization of sex-biased genes is even reflected in the differential enrichment for functional rubrics of the different sex-biased gene categories across Burkina Faso populations. Only in the case of genes with sex-limited expression, were more than 50% of the enriched functional rubrics in each population shared by at least two of the surveyed populations (supplementary table S14).

The interspecific comparison of the sex-biased expression across four *Anopheles* species, including *An. arabiensis*, revealed that male-biased genes showed a higher magnitude of sex bias compared to female-biased genes, while female-biased genes exhibited a higher rate of interspecific differentiation compared to male-biased genes (Papa et al. 2017). We examined the reproducibility of both patterns in the Burkina Faso populations. In the case of the first pattern, we found no differences in the magnitude of sex bias between the male- and female-biased gene categories when a combination of $P < 0.01$ and effect size (*i.e.* $\log_2$ FC) ≥ 1 was considered (supplementary fig. S12 and table S15). Subsequently, we tested for an increased expression variation of female-biased relative to male-biased genes. For that, we calculated the standard deviation of the $\log_2$ sex expression ratio (SD, Material and Methods), a common metric (Papa et al. 2017; Zhang et al. 2007). At a $P_{\text{adj}} < 0.01$, and an effect size > 0.3 to capture truly meaningful biological effects (Cohen 1992), only two out of the 12 contrasts showed differences in the SD of the sex expression between the sexes (Fig. 6; supplementary table S15). The differences found were consistent with higher expression variation for male-biased genes in Bobo and Nanoro, being primarily associated with the category of sex-limited expression. Repeating the analysis with only genes showing consistently the same pattern of sex bias across populations did not change the result (supplementary table S16). Collectively, these results indicate a lack of reproducibility of the patterns previously documented at the interspecific level in our Burkina Faso populations.

We also investigated whether there is a positive association between the magnitude of sex-biased expression and the degree of variation in mRNA abundance in Burkina Faso populations. Compared to non-sex-biased genes, we found evidence for a higher magnitude of variation for both sex-limited and strongly sex-biased genes, but not for moderately-sex-biased genes (supplementary fig. S13 and table S17). We saw no obvious difference between male- and female-biased genes in terms of magnitude of variation relative to non-sex-biased genes. Further, in other Diptera, male-biased genes are more polymorphic than female- or non-sex-biased genes, and generally any given gene category based on its sex bias (*i.e.* female-, male-, or non-sex-biased) shows more expression variability when assayed in males compared to females (Hutter et al. 2008; Meiklejohn et al. 2003; Muller et al. 2011). Contrary to this expectation, when examining the degree of interpopulation variation of different sex-biased gene categories within the same sex, we detected an overall higher level of expression polymorphism for female- compared to male-biased genes. This was obvious in males, particularly for strongly sex-biased genes and most comparisons including the non-sex-biased category (supplementary table S18). The same trend was reproduced in females, but in some cases the effect size was not as large as in males. Based on this trend, we predicted that the strongly sex-biased gene categories should display larger interpopulation variation when expressed in males compared to when expressed in females. Upon adjusting for population-level differences, we found support to our prediction (supplementary fig. S13 and table S19). Altogether, the patterns of sex-biased expression described here in *An. arabiensis* share limited similarity with those previously reported in *D. melanogaster* (Hutter et al. 2008; Meiklejohn et al. 2003; Muller et al. 2011).

DISCUSSION

Patterns of interpopulation variation in gene expression

Genome-wide analyses of molecular variation in *An. arabiensis* and other anopheline species have primarily focused on sequence evolution of protein-coding genes (Fontaine et al. 2015; Neafsey et al. 2010; Neafsey et al. 2015). In contrast, equivalent analysis on expression variability have been scarce (Bonizzoni

et al. 2012) and largely limited to the response to insecticides (Bonizzoni et al. 2015; Ingham et al. 2018; Simma et al. 2019). As the anophelines are subject to strong selection pressures including climate change, habitat disturbance, and human-mediated interventions to control malaria, transcriptome studies are arguably essential to understand the adaptive response of these species.

Our portrait of the expression polymorphism in *An. arabiensis* has revealed remarkable levels of mRNA variation among females, males, and larvae from three populations of Burkina Faso located within ~350 km², with larvae showing the lowest level of expression polymorphism. This is largely expected due to several factors. The first is the enhanced selection intensity that operates on this developmental stage, as it requires an aquatic environment and is highly susceptible to predation, which results in higher mortality rates compared to adulthood (Awono-Ambene and Robert 1999; Carnevale and Robert 2009). The second factor is the greater developmental constraints associated with the insect larval stage, constraints that decrease as development progresses (Artieri and Singh 2010). And the third is the reduced power to detect interpopulation expression differences when sequencing a mix of male and female larvae as a gene can be differentially expressed between two populations only in one sex, or in both sexes but in opposite directions, *i.e.* overexpressed in one sex and underexpressed in the other.

Notably, we observed discordant patterns of transcriptome variation among populations across sample types. Using the fraction of differentially expressed genes between two populations as a distance metric, the Nanoro population, which is the only one of the three located at the Sudano-Sahelian ecotype zone, shows the most variable transcriptome for males and larvae, whereas the Kodené population does so for the female transcriptome. Particularly noticeable is the relatively low number of differentially expressed genes between the females of Bobo and Nanoro, a pattern for which its biological underpinnings are not apparent at this time and should be further explored. Overall, these differential patterns of expression polymorphism could be explained by asymmetrical functional constraints and selective pressures among sample types, which underscores the need of performing expression analyses with different sample types to obtain a more accurate view of the genome-wide patterns of population differentiation in the anophelines.

From a functional perspective, the patterns of transcriptome differentiation, based on the number of GO terms and pathways that showed significant enrichment among differentially expressed genes, only mirror the ecotype location of the three analyzed populations in the case of the males but not in the case of females and larvae. This is particularly evident for the females from Bobo and Koden, which exhibit the most marked patterns of functional differentiation despite belonging to the same ecotype zone. Overall, the disparate patterns of functional differentiation shown by different sample types across populations could be explained by uneven effects of local variation in abiotic and biotic factors although experiments directly testing this possibility remain to be done. We also documented very little overlap among enriched functional terms across sample types, which, to some extent, is unsurprising given the different ontogenetic constraints and selection pressures to which larvae, males, and females are exposed.

Expression polymorphism is abundant among insecticide resistance genes

We detected that many polymorphic genes are associated with functional terms relevant for insecticide resistance. For example, 251 (~2/3) of the differentially expressed genes in females in response to DDT and the pyrethroid deltamethrin in Ethiopian populations (Simm et al. 2019) were found polymorphic across our Burkina Faso populations in not only females (126 or 31.4%) but also males (162 or 40.4%) and larvae (91 or 22.7%). Although tentatively some of these differentiated genes could be used as markers for tracking resistance in Western and Eastern African populations of *An. arabiensis*, analyses of the genome-wide transcriptional response to insecticides are yet to be done in Burkina Faso populations. Furthermore, we note that among these genes, we have identified not only those typically known to mediate metabolic resistance in anopheline species but also others previously overlooked but recently proposed to contribute to alternative pathways and biological mechanisms of resistance (Ingham et al. 2018). An overall concerning aspect is the abundant naturally occurring variation associated with metabolic and penetration resistance across sample types, which might hinder long-term vector control strategies that rely primarily on insecticide-based interventions.

The X chromosome exhibits reduced levels of transcriptional variation relative to the autosomes

Genes showing interpopulation expression polymorphism are non-randomly distributed between the *X* chromosome and the autosomes of *An. arabiensis*. The *X* chromosome appeared significantly depleted of these genes across all sample types investigated, particularly in larvae, which should be studied more precisely in this developmental stage by analyzing each sex separately. This decreased level of expression polymorphism of *X*-linked genes is reminiscent of observations at the intraspecific level in embryos and adult males of different *D. melanogaster* inbred strains (Hutter et al. 2008; Kayserili et al. 2012), potentially reflecting a faster-*X* effect, *i.e.* a rate of evolution on the *X* that exceeds that on the autosomes. This effect can result from two non-mutually exclusive mechanisms. First, it could result from a more efficient fixation of beneficial recessive mutations via positive selection when hemizygous (Charlesworth et al. 1987). Second, it could reflect a higher fixation of dominant slightly deleterious mutations through genetic drift due to a reduced effective population size of the *X* chromosome relative to the autosomes (Charlesworth et al. 1987; Vicoso and Charlesworth 2009) –the *X* chromosome but not the autosomes is in lower number in males relative to females, *i.e.* $N_{eX}=3/4N_{eA}$. Upward alterations of the relationship between the N_e of the *X* and autosomes due to intense male-male competition as in *D. melanogaster*, *i.e.* $N_{eX}>3/4N_{eA}$ (Laporte and Charlesworth 2002; Vicoso and Charlesworth 2009), would not apply to the same extent to *An. arabiensis* as females are essentially monandrous (Carnevale and Robert 2009; Tripet et al. 2003; Yuval 2006).

A faster-*X* effect on expression has been often interpreted as a result of an adaptive scenario driven by *X*-linked recessive *cis*-regulatory mutations (Coolon et al. 2015; Kayserili et al. 2012; Llopart 2018, 2012; Meisel et al. 2012), particularly in association with male-biased genes. Male-biased, as opposed to female-biased genes, are preferentially located on the autosomes of *Drosophila* species, are the gene category showing the largest divergence between species, and, in some reports, have the highest levels of gene expression polymorphism (Hutter et al. 2008; Meiklejohn et al. 2003; Ranz et al. 2003; Zhang et al. 2007). As reported at the species level in the anophelines (Baker and Russell 2011; Jiang et al. 2015; Magnusson et al. 2012; Papa et al. 2017), and here for the Burkina Faso populations of *An. arabiensis*, female-biased genes are also located in excess on the *X* chromosome while male-biased genes are preferentially located

on the autosomes (supplementary fig. S15). However, it is not clear whether male-biased genes are driving a faster-*X* effect in *An. arabiensis*, as the strongest evidence for depletion of polymorphic genes is found in larvae—a developmental stage more moderately impacted by sexual dimorphism—, followed by females, with males showing the weakest evidence for depletion of *X*-linked polymorphic genes. Additionally, some categories of female-, but not male-biased genes, are in fact the most polymorphic in expression in *An. arabiensis* (supplementary Table S18). Further, adopting the rationale that a high correlation between male and female gene expression reflects strong selective constraints, *i.e.* weak relaxed selection (Meisel et al. 2012), we found that the *X* chromosome only exhibits higher levels of selective constraints relative to the 3R arm (supplementary figure S16). Therefore, a faster-*X* effect in *An. arabiensis* seems, if anything, more influenced by non-adaptive than adaptive causes, showing substantial general differences with previous findings in *Drosophila*.

Genomic hot spots of transcriptome variation

Inversion polymorphism might favour the accumulation of transcriptional differences in particular chromosomes or genomic regions (Marques-Bonet et al. 2004). The lack of information about the location of the breakpoints of the inversions 2Ra and 3Ra precluded a precise assessment of whether expression differences were particularly associated with these inversions or whether there was any position effect at their breakpoints. Nevertheless, we did not find that the chromosomes carrying such inversions, the 2R and 3R, were particularly enriched in differentially expressed genes as it has been found in other chromosome inversion polymorphism systems in *Drosophila* species (Fuller et al. 2016; Lavington and Kern 2017). One possible explanation is that these inversions had arisen very recently and therefore not enough time has elapsed as for detecting the expected signature of expression differentiation between populations in association with such inversions. Additional work deciphering the evolutionary origin of these inversions will be necessary to test this possibility.

Further, we found evidence of a non-random distribution of the top contributing genes to functional enrichment patterns of population differentiation across the genome. This unequal distribution is not only

happening at the whole-chromosome scale, but more importantly, it is consistent with an incipient emergence of hot spots of transcriptome differentiation, possibly denoting effects of population divergence due to genetic changes affecting *cis* regulatory sequences or local chromatin properties.

The lability of sex-biased expression requires a population framework

Our analysis of sex-biased gene expression stresses the need for adopting a population framework relative to the categorization of genes according to their sex bias trend. We found that 2,227 (68%) of the female-biased genes and 2,381 (81%) of the male-biased genes in at least one population showed transitions between non-sex bias and sex bias in gene expression –or vice versa– across Burkina Faso populations. This intraspecific lability of sex-biased expression is even more evident when we compared our sets of genes consistently showing male or female bias in relation to a population from Dongola, Sudan (Papa et al. 2017). We found that 0.63% (2/318) and 14.29% (60/420) of the genes in these two categories respectively, which also show sex-biased expression in the Dongola population, reverse the direction of their bias between studies. Similarly, when we considered jointly genes that show consistent and non-consistent sex-bias, we detected 248 genes exhibiting reversals of the sex bias across Burkina Faso populations, a number significantly higher than the only seven equivalent cases documented between the Dongola population of *An. arabiensis* and another of *An. gambiae* (two-tailed Fisher's exact test; $P=1.18 \times 10^{-18}$) (Papa et al. 2017). The lability of sex-biased gene expression featured by the Burkina Faso populations is not surprising as different populations might be exposed not only to different abiotic factors but also to very different regimes of sexual selection, which has been proven to be a main factor in the evolution of sex-biased expression in experimentally evolved populations of *Drosophila* species over very short timespans (Hollis et al. 2014).

Patterns of sex-biased transcriptome evolution

A previous comparative analysis of four *Anopheles* species, including *An. arabiensis*, found evidence that female-biased genes were more rapidly evolving interspecifically than male-biased genes, in sharp contrast to former findings in other eukaryotes (Khaitovich et al. 2005; Ranz et al. 2003; Voolstra et al. 2007). This peculiarity of the anophelines was interpreted to be the consequence of two factors. First, females in these

species are under strong selection pressures, as only this sex blood-feeds, which is essential for oogenesis (Attardo et al. 2003; Mitchell and Catteruccia 2017). Second, anopheline males would be exposed to decreased post-copulatory sexual selection compared to other Diptera, as females typically mate only once during their lifetime (Carnevale and Robert 2009; Tripet et al. 2003; Yuval 2006), which is due to a male-induced mating plug and the subsequent refractory response (Mitchell et al. 2015; Rogers et al. 2009). Nevertheless, an increased rate of evolution at the interspecific level could also result from nonadaptive processes. Finding evidence of genes with female-biased expression being more influenced by positive selection compared to male-biased genes will ultimately require determining whether a significantly higher proportion of between-species divergence to within-species polymorphism exists that is associated with female-biased genes (Harrison et al. 2012; Meiklejohn et al. 2003; Nuzhdin et al. 2004; Yeh et al. 2014).

At the intraspecific level in *An. arabiensis*, we have found that overall female-biased genes show similar magnitude of sex-expression ratio variation to that of male-biased genes, although male-limited expressed genes in Bobo and Nanoro showed a significantly higher expression variation than their female-limited counterparts. This overall lack of differences between female- and male-biased genes at the intraspecific level does not refute nor confirm whether intense positive selection is the primary factor driving the fast evolution of female-biased genes in anopheline species. Nevertheless, several observations are not in line with the factors considered as evidence for the preferential action of positive selection on the female transcriptome, which call for additional studies to more reliably discard relaxed selection as a relevant contributing factor to the fast evolution of female- relative to male-biased genes at the interspecific level. First, we have found unambiguous evidence that female-biased genes are more variable than male-biased genes at the interpopulation level when assayed both in males and females, which is compatible with female-biased genes being subject to less stringent selective constraints. Note that this last observation does not contradict the notion reported in other Diptera that gene expression is more variable in males than in females (Meiklejohn et al. 2003; Muller et al. 2011), due to an increased susceptibility to genetic or environmental modifiers (Jiang et al. 2010). In fact, we do observe a larger magnitude of interpopulation differentiation for the same gene category, regardless of its sex bias, when expressed in males compared to in females.

Additionally, there is evidence for accelerated evolution of male-related traits at the post-copulatory level, which can have adaptive potential. In *An. arabiensis*, the entire 20E content generated by the male accessory glands is transferred to the female during copulation, while other anophelines only transfer half (Mitchell et al. 2015). As this 20E content elicits vital physiological responses in the females, impacting their remating ability (Baldini et al. 2013; Gabrieli et al. 2014), and ultimately their vectorial capacity (Rono et al. 2010), male-male competition could still take place, however, not by minimizing the direct fertilization success of the rival sperm but by impacting indirectly some sperm traits such as longevity and motility, or by modulating the reproductive fitness of the female, in part through its endocrine system (Rogers et al. 2008). In fact, we found significant differences in the magnitude of sex bias among genes expressed in the female spermathecae (AARA002778 and AARA017412), whose orthologs in *An. gambiae* (an ABC transporter and a peroxidase encoding genes) regulate the lipid reserves for sperm storage and acquisition of sperm motility respectively (Aitken and Baker 2006; Rogers et al. 2008). Our results add to the increasingly accepted view that the strength of sexual selection operating on males might be higher than previously recognized (Cator et al. 2021).

CONCLUSIONS

Intraspecific expression surveys in the anophelines are fundamental to understand the potential of these species for adaptation and trait diversification, which is crucial for vector control. These studies should interrogate the transcriptome response to environmental cues such as insecticides and increasing temperatures while considering sex-specific aspects of the anopheline biology. Equally important, these studies should also address what genetic changes are responsible for differential expression. While expression differences between the sexes primarily reflect their distinct reproductive systems (Assis et al. 2012; Grath and Parsch 2016; Jiang et al. 2015; Papa et al. 2017), intraspecific expression differences for the same sample type under controlled environmental conditions can result from *cis* and *trans* regulatory changes as well as from copy number variation (Cardoso-Moreira et al. 2016; Faucon et al. 2015; Itokawa et al. 2010; Meiklejohn et al. 2014). Therefore, genome sequencing of these, or similar, populations should help gauge the relative importance of these different types of changes in their contribution to the

transcriptome diversification of *An. arabiensis*, and consequently to the expansion of its geographic range while transmitting malaria across Africa.

ACKNOWLEDGMENTS

We are grateful to Elizabeth Hemming-Schroeder, Anthony James, Colin Meiklejohn, Igor Sharakhov, and Guiyun Yan for providing useful comments on early versions of the manuscript, and to the staff at the IRSS for technical assistance at different stages in establishing and maintaining the mosquito populations. This work was supported by an NIH award (R21AI096084) and UCI Intramural funds to J.M.R.

MATERIAL & METHODS

Mosquito collection and identification. Mosquitoes were collected from July to November 2012. No specific permissions were required for the locations where the collection was performed. Blood-fed and unfed resting females were collected using manual aspirators in human dwellings. Subsequently, *An. gambiae s.l.* mosquitoes were sorted from other species using morphological identification criteria (Gillies and Coetzee 1987; Gillies and De Meillon 1968). Once in the laboratory, unfed mosquitoes were provided a blood meal on a rabbit. Field collected or laboratory blood-fed females were kept in cages for 4 days until they became gravid and ready to lay eggs. Then, females were individually placed in plastic cups containing approximately 25ml of water for oviposition. Once having laid eggs, they were killed and genomic DNA extracted from legs, followed by species identification by PCR (Santolamazza et al. 2008) as the member species from the *An. gambiae s.l.* are morphologically indistinguishable. [In Burkina Faso, three members of this complex are present: *An. gambiae*, *An. coluzzii*, and *An. arabiensis* (Costantini et al. 2009)] The PCR is based on the amplification of a ~200 nt long Short Interspersed Element (SINE 200) locus on the X chromosome (locus S200 X 6.1.) using the flanking primers SINE 200X6.1F (5'-TCGCCTTAGACCTTGGGTTA-3') and SINE 200X6.1R (5'-CGCTTC AAGAATTCGAGATAC-3'). The reaction mixture contained 2µl of 10x PCR buffer (Promega, Madison WI, USA), 1 mM MgCl₂, 0.2 mM of

each dNTP, 0.3 μ M of each primer (Eurogentec, Angers, France), 0.35 unit of Taq DNA Polymerase (Promega) and 2 μ L of template DNA, in a total volume of 20 μ L. Amplification runs were performed under the following conditions: an initial denaturation step at 94°C for 10min followed by 35 cycles of 30s at 94°C, 30s at 54°C, 60s at 72°C, and a final elongation step of 10min at 72°C. The amplification products were run in 2% agarose gels (supplementary fig. S1). Progeny of all mosquitoes molecularly identified as *An. arabiensis* were pooled according to the study sites and reared under common insectary conditions. Larval offspring were reared at 30°C and fed with Baby Tetramine fish food. Adults were maintained at 28°C and 80-85% relative humidity with *ad libitum* supply of 10% sucrose solution.

Chromosome karyotyping. Ovaries were extracted from 20-22h post blood feeding females under a stereoscope (Leica Microsystem GmbH, Wetzlar, Germany) and preserved in Carnoy's fixative solution (three parts of absolute ethanol and one part of glacial acetic acid). After one night at room temperature, ovaries were kept at -20°C until chromosomal analysis was performed. Ovaries were first removed from the Carnoy's fixative solution and placed in a drop of 50% propionic acid on a microscope slide. Then, follicles were carefully separated from each other using fine needles under a stereoscope. After 3-5 minutes, a cover slip was placed over the preparation and a gentle pressure applied by tapping with a pencil to squash the follicles allowing polytene chromosomes to spread out. The slides were examined under a contrast phase microscope (Motic BA410, Motic China group co., Ltd., Xiamen China) and chromosome arrangements identified using a reference of the cytogenetic map for *An. arabiensis* (Coluzzi et al. 2002).

Construction of isofemale lines. Naïve males and virgin females were obtained by sexing pupae prior to emergence (Jones 1957), and placing them into separate cages. Individuals were supplied with 10% sucrose during the first two days to allow complete reproductive maturation. On the third day after emergence, females were blood-fed and subsequently forced to mate with naïve males (one male for one female). Four days after insemination, females were provided a blood meal and subsequently allowed to lay eggs as described above. Isofemale progenies were then carefully reared separately until F4-F6.

Construction of homokaryotypic colonies. Subsets of mosquitoes were blood-fed and allowed to lay eggs individually. Once having laid eggs, mosquitoes were given a second blood meal and their ovaries

were extracted 20-22h later for karyotyping. Specimens carrying the inversion 2Ra, 2Rb or 3Ra were then identified and progeny of females sharing identical chromosome arrangement were pooled together for mating. After mating and blood feeding, F1 mosquitoes were placed in individual cups for oviposition, re-blood fed, and their ovaries extracted for karyotyping. As previously, F2 mosquitoes were sorted by their karyotype. The process was repeated until each inversion was completely fixed in the colony as tested by further karyotyping. The resulting colonies were maintained in the insectary of the Institute de Recherche en Sciences de la Santé (IRSS) at Bobo Dioulasso, Burkina Faso, under controlled conditions of 28°C, 80-85% relative humidity and a 12 h / 12 h light–dark cycle without crepuscular transitions. Larvae were fed using TetraMin baby fish food while adults were provided *ad libitum* 10% sucrose solution. Females were allowed to feed on a rabbit for production of eggs, following the approved UCI IACUC protocol number 2012-3035.

RNA extraction and sequencing. Stage 4 larvae and 3- to 5-day-old naïve males and virgin nonblood-fed females from homokaryotypic colonies were frozen in TRIzol (Invitrogen) for subsequent total RNA extraction. Specimens were macerated using motorized pestles and the resulting total RNA samples were purified with RNeasy Mini kit (Qiagen). Concentration, quality, and integrity of the RNA samples were assessed using a Qubit 2.0 Fluorometer, a NanoDrop 8000 Spectrophotometer, and with the RNA 6000 Pico kit in a BioAnalyzer 2100 (Agilent Technologies), respectively. Unstranded (female) and stranded (male and larvae), non-polyA, ribodepleted libraries were prepared using the TruSeq Total RNA (Illumina) and the Ribo-Zero™ (Epicentre) kits. cDNAs were sequenced using 100-bp paired end reads on an Illumina HiSeq 2500. Size selection was performed using 2% agarose gels on the Sage Sciences Blue Pippin, resulting in a library insert size of ~480 nt. Three biological replicates were sequenced per biological group, *i.e.* a particular sample type by population combination, being 27 the total number of libraries (*i.e.* 3 samples × 3 populations × 3 replicates). The set of nine libraries corresponding to a particular sample type were sequenced jointly in two lanes.

Transcriptome analyses. Quality control and processing of RNA-seq outputs were performed using the pipeline HTStream (<https://github.com/s4hts/HTStream>), including removal of known *Anopheles* rRNA-related sequences as present in NCBI, PCR duplicates, adapter sequences, reads shorter than 50 nt, and

591 filtered for low quality bases using a sliding window approach that required a window size of 10 nt and a
592 minimum quality Qscore of 20. Reads were mapped using the splice-aware STAR Aligner v2.7.3a (Dobin et
593 al. 2013) to the AaraD1 (release 47; April 2020) genome assembly, which was softmasked using
594 RepeatMasker, Dust, and TRF. The AaraD1.11 gene annotation was used. Both annotation and genome
595 assembly were downloaded from VectorBase (as of July 31 2020). Scaffold location at the chromosomal
596 level was taken as reported (Waterhouse et al. 2020). SAM alignment files were sorted and converted into
597 BAM files using SAMtools v1.10 (Li et al. 2009) and subsequently into count matrices based on the gene
598 annotation used.

599 Differential expression analysis was performed on those genes deemed as expressed, which implied the
600 detection of at least 1 count per million (cpm) in at least two of the three biological replicates of at least one
601 biological group. Firstly, read counts across sample types for the genes deemed as expressed were scaled
602 using TMM normalization (Robinson and Oshlack 2010) and log2 transformed. Secondly, precision weights
603 were estimated using the *voom* method (Law et al. 2014) and provided as input for linear models. For each
604 sample type, three linear models were considered corresponding to the three pairwise comparisons
605 between populations. Mathematically, each model was of the form $y_{ij} = \alpha + \beta_i + \epsilon_{ij}$ where, y_{ij} denoted the
606 normalized and log2 transformed gene expression in the j^{th} replicate from the i^{th} population, α denoted the
607 baseline expression, β_i denoted the average increase (or decrease) in expression with respect to baseline
608 in replicates from the i^{th} population, and $\epsilon_{ij} \sim N(0, \sigma^2)$ denoted an independent and identically distributed
609 error term with mean 0 and variance σ^2 . The above model was implemented in *limma* (Ritchie et al. 2015)
610 to test the null hypothesis that $\beta_1 = \beta_2$, *i.e.* for a given sample type, the average expression of a gene was
611 the same in the two populations labelled 1 and 2. We rejected the null hypothesis and considered a gene to
612 be differentially expressed if the false discovery rate (FDR) (Benjamini and Hochberg 1995) adjusted p -
613 value was < 0.01 and $FC \geq 2$, *i.e.* more strict criteria than those conventionally used to ensure that the same
614 genes will be identified as differentially expressed even in the case that the number of biological replicates
615 per biological group was increased. Lastly, to determine whether the differentially expressed genes in a

616 given contrast, *i.e.* between two populations of the same sample type were preferentially over- or under-
617 expressed, a goodness-of-fit test was performed under the null hypothesis that $p_1 = p_2$, where p_1 and
618 p_2 denote the proportion of differentially expressed genes that are over- and under-expressed, respectively.
619 We rejected the null hypothesis if the FDR-adjusted p -value < 0.01 .

620 **Functional enrichment analysis.** Lists of differentially expressed genes were used to identify enriched
621 pathways using VectorBase's Metabolic Pathway Enrichment tool (Giraldo-Calderon et al. 2015), and
622 enriched categories of Gene Ontology (GO) terms using g:Profiler (Raudvere et al. 2019), in both cases as
623 of Oct 2020. For any relevant set of genes, a particular pathway was considered to be enriched for
624 differentially expressed genes if the FDR adjusted p -value was < 0.01 . In the case of the GO terms, the
625 g:SCS multiple test correction was applied also at a p -value < 0.01 . This more conservative correction does
626 not assume independence among the multiple tests that have to be done as it considers the dependencies
627 of GO terms within the GO term hierarchy.

628 For the identification of *top contributor* genes in terms of biological relevance, *i.e.* those genes that
629 contribute to the statistical enrichment across multiple functional terms, we determined whether the joint
630 probability of the gene being differentially expressed and an associated pathway (or GO term) being
631 enriched was greater than or equal to an arbitrary threshold value Q_{thresh} . For a given sample type, the joint
632 probability was defined as $Q(g, w) = P(g) \times P(w|g)$. Here, g and w denote a differentially expressed gene and
633 an enriched pathway, respectively. Since there are only three pairwise comparisons of interest, $P(g) = N/3$,
634 where $0 \leq N \leq 3$ and $P(w|g) = M/N$, where $0 \leq M \leq N$. We set Q_{thresh} to 0.33 or 0.66 depending on whether the
635 gene was differentially expressed in only one of the pairwise comparisons or more. If a gene is differentially
636 expressed in only one comparison, then the maximum permissible value for $Q(g, w) = \frac{1}{3} \times \frac{1}{1} = 0.33$. In
637 contrast, if a gene is differentially expressed in ≥ 2 comparisons, then $Q(g, w) = 0.66$, which ensures that the
638 pathway is enriched in more than 50% of such cases. For each gene examined, we counted X , *i.e.* the

639 number of pathways (or GO terms) where $Q(g, w) \geq Q_{\text{thresh}}$ and if $\frac{X}{Y} > 0.25$, where Y denotes the total
640 number of enriched pathways (or GO terms), then the gene was deemed a *top contributor* gene. The
641 threshold of 0.25 was selected based on the empirical observation that, for most pairwise comparisons, it
642 resulted in a small number of genes (< 30) being included in the list of *top contributors*, thereby facilitating
643 the downstream examination of their properties.

644 **Sex-biased expression analysis.** Population-specific linear models, similar to those described above,
645 were developed to identify differentially expressed genes ($FC \geq 2$ and FDR-adjusted p -value < 0.01)
646 between males and females. The differentially expressed genes were divided into six groups based on the
647 type (overexpressed in male or female) and magnitude of sex bias. Unlike genes showing sex-limited
648 expression, those preferentially expressed in one of the sexes were further subdivided according to the fold-
649 change ($2 \leq \text{fold-change} < 4$, moderate sex bias; and $\text{fold-change} \geq 4$, strong sex bias). A fold-change
650 threshold of 4 was selected based on the empirical observation that, in each of the three populations, the
651 third quartile of fold-changes was approximately 4 implying that only 25% of the genes had fold-change
652 exceeding this threshold.

653 Subsequently, a Wilcoxon rank-sum test was performed to determine whether male- and female-biased
654 genes showed differences in expression variability. First, the six groups were mapped to three categories:
655 exclusive (expressed only in male or female); non-exclusive with strong bias (overexpressed in male or
656 female with $\text{fold-change} \geq 4$); and non-exclusive with moderate bias (overexpressed in male or female and
657 $2 \leq \text{fold-change} < 4$). Second, for each differentially expressed gene, the standard deviation for the expression
658 ratio between the sexes, *i.e.* the standard deviation for $\log_2(\text{male/female})$, was estimated using the empirical
659 Bayes method in *limma*. Third, for each of the three categories, a two-sided Wilcoxon rank-sum test was
660 performed to test the null hypothesis (H_0) that the standard deviation was the same in male- and female-
661 biased genes. Fourth, the U-statistic for Wilcoxon rank-sum test was converted into rank-biserial correlation
662 (Kerby 2014) to estimate the “effect size” or $|f - s|$ where f and s denote the probability of a male- or
663 female-biased gene, respectively, having a higher standard deviation. Lastly, for a given category, we

664 rejected H_0 if the FDR-adjusted p -value (Wilcoxon rank-sum test) < 0.01 and effect size ≥ 0.3 . For the effect
665 size, the cut-off value used was in accordance with reported guidelines (Cohen 1992) for considering an
666 effect size to be biologically meaningful rather than just statistically significant, as the latter can simply result
667 from an increase in the power of the test associated with a large sample size. It should be noted that the use
668 of the effect size for Wilcoxon rank-sum test is equivalent to the use of a fold-change threshold in gene
669 expression analysis in addition to the p -value for identifying genes with non-trivial changes.

670

671 **DATA AVAILABILITY**

672 All raw sequencing data were deposited as part of the NCBI BioProject PRJNA707074.

673

FIGURE CAPTIONS

Fig. 1. Collection sites of gravid females *An. arabiensis* in Burkina Faso. Nanoro is located in the Sudano-Sahelian zone, which extends throughout much of the central part of the country. In this zone, vegetation is constituted of an arid savannah and the climate is characterized by a short rainy season from July to September. Bobo Dioulasso, or Bobo, and Kodené are situated in the Sudano-Guinean zone at the Southwest of the country. In this zone, the type of vegetation is humid savannah, and the rainy season extends from June to October.

Fig. 2. Magnitude and patterns of interpopulation differentiation of the transcriptome of *An. arabiensis* across sample types. (A) Number of differentially expressed genes at a 1% FDR and fold-change ≥ 2 for females, males, and larvae. The total number of unique differentially expressed genes across contrasts for each sample type is indicated in parenthesis. (B) Venn diagrams showing the degree of overlap among genes differentially expressed across the three contrasts performed for each sample type. (C) Hierarchical clustering of differentially expressed genes for different sample types. The genes included correspond to those in the union set of the Venn diagrams in (B). B, Bobo; K, Kodené; and N, Nanoro. Unlike Nanoro, Bobo and Kodené are associated with the same ecotype zone.

Fig. 3. Commonalities in functional enrichment patterns across different functional rubrics are limited among sample types. (A) Degree of overlap among sets of differentially expressed genes in at least one population contrast at a 1% FDR across the three types of samples assayed for GO terms and pathways. GO term and pathway enrichment was assessed using g:Profiler (Raudvere et al. 2019) and VectorBase (Giraldo-Calderon et al. 2015), respectively. (B) Degree of overlap among sets of differentially expressed genes tagged as *top contributors* to patterns of functional enrichment for GO terms (without distinguishing among Biological Process, Molecular Function, and Cellular Component) and pathways.

Fig. 4. Differentially expressed genes are non-randomly distributed across the *An. arabiensis* chromosomal element. A) Per sample type. (B) Main contributors to functional enrichment patterns. Top, different chromosomes, with L and R denoting the left and right arms of the metacentric chromosomes 2 and

3, respectively. Only genes with $FC \geq 2$ in expression at a 1% FDR in at least one of the population contrasts performed are considered. Each of these genes is indicated with a vertical solid line. Significant enrichment ($\uparrow$) or depletion ($\downarrow$) at 1% FDR-corrected p -value (*) or at 5% FDR-corrected p -value (•) is shown.

Fig. 5. Abundant and inconsistent sex-biased gene expression across populations of *An.*

***arabiensis*.** (A) The violin plots show the proportion and range of fold-change values of differentially expressed genes between males and females in the three surveyed populations. (B) For sex-biased gene expression, female (red) or male (blue) bias might not be found in all three populations (*no consistently* category). In fact, the percentage of genes showing the same sex bias across populations (*consistently* category) is only 12.54% (4.32% + 8.22%). (C) Sankey charts showing the lability in the sex-biased expression between populations. The seven delineated categories in table 1 are color-coded. Bobo and Nanoro are more closely related to each other based on the patterns of sex-biased expression than either of them to Kodení. See supplementary table S12 for the precise number of genes associated with different pairwise combinations of categories based on the magnitude and type of sex-biased expression between the two populations involved in each case. F, female-biased; M, male-biased; No DE, no differentially expressed between the sexes, *i.e.* no sex-biased.

Fig. 6. Magnitude of variation in sex expression ratio of sex-biased genes. The violin plots show the proportion and range of the variation in sex expression ratio across the three populations of *An. arabiensis* investigated. For each population, all sex-biased genes, or subsets of them based on the magnitude of the sex bias, are plotted side by side. The magnitude of variation in sex expression ratio was calculated as the standard deviation (SD) of $\log_2(\text{male/female})$ expression. Only in two out of 12 contrasts is there evidence of significantly higher variation between the sexes, in all cases male-biased genes show significantly higher variation than female-biased genes with an effect size >0.3 and statistical significance (***) at $P_{\text{adj}} < 0.001$ (supplementary table S15).

722 TABLES

723

Table 1. Population patterns of sex-biased gene expression

Population	Non-Sex-Biased	Female-Biased				Male-biased				Sex-Biased
		Moderate	Strong	Limited	Subtotal	Moderate	Strong	Limited	Subtotal	
Bobo	10006	676	262	587	1525	572	199	273	1044	2569
Kodeni	7422	1465	662	594	2721	1403	475	554	2432	5153
Nanoro	9073	911	395	690	1996	935	264	307	1506	3502

* Fold-change differences: Moderate ($2 \geq FC \leq 4$); Strong ($FC > 4$); Limited (expression evidence in only one sex). See supplementary table S10.

725

REFERENCES

- Aitken RJ, Baker MA 2006. Oxidative stress, sperm survival and fertility control. *Mol Cell Endocrinol* 250: 66-69. doi: 10.1016/j.mce.2005.12.026
- Artieri CG, Singh RS 2010. Demystifying phenotypes: The comparative genomics of evo-devo. *Fly (Austin)* 4: 18-20.
- Assis R, Zhou Q, Bachtrog D 2012. Sex-biased transcriptome evolution in *Drosophila*. *Genome biology and evolution* 4: 1189-1200. doi: 10.1093/gbe/evs093
- Attardo GM, Higgs S, Klingler KA, Vanlandingham DL, Raikhel AS 2003. RNA interference-mediated knockdown of a GATA factor reveals a link to anautogeny in the mosquito *Aedes aegypti*. *Proceedings of the National Academy of Sciences of the United States of America* 100: 13374-13379. doi: 10.1073/pnas.2235649100
- Awono-Ambene HP, Robert V 1999. Survival and emergence of immature *Anopheles arabiensis* mosquitoes in market-gardener wells in Dakar, Senegal. *Parasite* 6: 179-184. doi: 10.1051/parasite/1999062179
- Ayala D, et al. 2017. Chromosome inversions and ecological plasticity in the main African malaria mosquitoes. *Evolution* 71: 686-701. doi: 10.1111/evo.13176
- Ayala D, et al. 2019. Association mapping desiccation resistance within chromosomal inversions in the African malaria vector *Anopheles gambiae*. *Mol Ecol* 28: 1333-1342. doi: 10.1111/mec.14880
- Baker DA, Russell S 2011. Role of testis-specific gene expression in sex-chromosome evolution of *Anopheles gambiae*. *Genetics* 189: 1117-1120. doi: 10.1534/genetics.111.133157
- Baldini F, et al. 2013. The interaction between a sexually transferred steroid hormone and a female protein regulates oogenesis in the malaria mosquito *Anopheles gambiae*. *PLoS biology* 11: e1001695. doi: 10.1371/journal.pbio.1001695
- Barredo E, DeGennaro M 2020. Not Just from Blood: Mosquito Nutrient Acquisition from Nectar Sources. *Trends Parasitol* 36: 473-484. doi: 10.1016/j.pt.2020.02.003
- Barron MG, et al. 2019. A new species in the major malaria vector complex sheds light on reticulated species evolution. *Sci Rep* 9: 14753. doi: 10.1038/s41598-019-49065-5
- Benjamini Y, Hochberg Y 1995. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 57: 289-300.
- Bimbile Somda NS, et al. 2018. Ecology of reproduction of *Anopheles arabiensis* in an urban area of Bobo-Dioulasso, Burkina Faso (West Africa): Monthly swarming and mating frequency and their relation to environmental factors. *PLoS One* 13: e0205966. doi: 10.1371/journal.pone.0205966
- Bonizzoni M, et al. 2012. Strain Variation in the Transcriptome of the Dengue Fever Vector, *Aedes aegypti*. *G3 (Bethesda, Md)* 2: 103-114. doi: 10.1534/g3.111.001107
- Bonizzoni M, et al. 2015. RNA-seq analyses of changes in the *Anopheles gambiae* transcriptome associated with resistance to pyrethroids in Kenya: identification of candidate-resistance genes and candidate-resistance SNPs. *Parasit Vectors* 8: 474. doi: 10.1186/s13071-015-1083-z

763 Cardoso-Moreira M, et al. 2016. Evidence for the fixation of gene duplications by positive selection in
764 *Drosophila*. *Genome research* 26: 787-798. doi: 10.1101/gr.199323.115

765 Carnevale P, Robert V. 2009. Les Anophèles: Biologie, transmission du Plasmodium et lutte
766 antivectorielle. Marseille: IRD Éditions.

767 Cator LJ, Wyer CAS, Harrington LC 2021. Mosquito Sexual Selection and Reproductive Control
768 Programs. *Trends Parasitol* 37: 330-339. doi: 10.1016/j.pt.2020.11.009

769 Charlesworth B, Coyne JA, Barton NH 1987. The relative rates of evolution of sex chromosomes and
770 autosomes. *Am Nat* 130: 113-146.

771 Cohen J 1992. Statistical power analysis. *Current Directions in Psychological Science* 1: 98-101.

772 Coluzzi M, Sabatini A, della Torre A, Di Deco MA, Petrarca V 2002. A polytene chromosome analysis of
773 the *Anopheles gambiae* species complex. *Science* 298: 1415-1418.

774 Coolon JD, et al. 2015. Molecular Mechanisms and Evolutionary Processes Contributing to Accelerated
775 Divergence of Gene Expression on the *Drosophila* X Chromosome. *Mol Biol Evol* 32: 2605-2615. doi:
776 10.1093/molbev/msv135

777 Costantini C, et al. 2009. Living at the edge: biogeographic patterns of habitat segregation conform to
778 speciation by niche expansion in *Anopheles gambiae*. *BMC Ecol* 9: 16. doi: 10.1186/1472-6785-9-16

779 Dabire KR, et al. 2014. Occurrence of natural *Anopheles arabiensis* swarms in an urban area of Bobo-
780 Dioulasso City, Burkina Faso, West Africa. *Acta Trop* 132 Suppl: S35-41. doi:
781 10.1016/j.actatropica.2013.12.012

782 Dao A, et al. 2010. Reproduction-longevity trade-off in *Anopheles gambiae* (Diptera: Culicidae). *J Med*
783 *Entomol* 47: 769-777. doi: 10.1603/me10052

784 Darwin C. 1871. The descent of man and selection in relation to sex. London: John Murray.

785 Dobin A, et al. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* (Oxford, England) 29: 15-
786 21. doi: 10.1093/bioinformatics/bts635

787 Ellegren H, Parsch J 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev*
788 *Genet* 8: 689-698.

789 Faucon F, et al. 2015. Identifying genomic changes associated with insecticide resistance in the dengue
790 mosquito *Aedes aegypti* by deep targeted sequencing. *Genome research* 25: 1347-1359. doi:
791 10.1101/gr.189225.115

792 Feyereisen R, Dermauw W, Van Leeuwen T 2015. Genotype to phenotype, the molecular and
793 physiological dimensions of resistance in arthropods. *Pestic Biochem Physiol* 121: 61-77. doi:
794 10.1016/j.pestbp.2015.01.004

795 Fontaine MC, et al. 2015. Mosquito genomics. Extensive introgression in a malaria vector species
796 complex revealed by phylogenomics. *Science* 347: 1258524. doi: 10.1126/science.1258524

797 Fuller ZL, Haynes GD, Richards S, Schaeffer SW 2016. Genomics of Natural Populations: How
798 Differentially Expressed Genes Shape the Evolution of Chromosomal Inversions in *Drosophila*
799 *pseudoobscura*. *Genetics* 204: 287-301. doi: 10.1534/genetics.116.191429

800 Gabrieli P, et al. 2014. Sexual transfer of the steroid hormone 20E induces the postmating switch in
801 *Anopheles gambiae*. Proceedings of the National Academy of Sciences of the United States of America
802 111: 16353-16358. doi: 10.1073/pnas.1410488111

803 Gillies MT, Coetzee M. 1987. A supplement to the Anophelinae of Africa south of the Sahara (Afrotropical
804 Region): Publications of the South African Institute for Medical Research.

805 Gillies MT, De Meillon B. 1968. The Anophelinae of Africa south of the Sahara (Ethiopian
806 Zoogeographical Region): Publications of the South African Institute for Medical Research.

807 Giraldo-Calderon GI, et al. 2015. VectorBase: an updated bioinformatics resource for invertebrate vectors
808 and other organisms related with human diseases. Nucleic Acids Res 43: D707-713. doi:
809 10.1093/nar/gku1117

810 Grath S, Parsch J 2016. Sex-Biased Gene Expression. Annu Rev Genet 50: 29-44. doi: 10.1146/annurev-
811 genet-120215-035429

812 Harrison PW, Wright AE, Mank JE 2012. The evolution of gene expression and the transcriptome-
813 phenotype relationship. Semin Cell Dev Biol.

814 Hollis B, Houle D, Yan Z, Kawecki TJ, Keller L 2014. Evolution under monogamy feminizes gene
815 expression in *Drosophila melanogaster*. Nat Commun 5: 3482. doi: 10.1038/ncomms4482

816 Huong LT, et al. 2019. Chemical Compositions and Mosquito Larvicidal Activities of Essential Oils from
817 Piper Species Growing Wild in Central Vietnam. Molecules 24. doi: 10.3390/molecules24213871

818 Hutter S, Saminadin-Peter SS, Stephan W, Parsch J 2008. Gene expression variation in African and
819 European populations of *Drosophila melanogaster*. Genome biology 9: R12. doi: 10.1186/gb-2008-9-1-
820 r12

821 Ibrahim SS, Riveron JM, Stott R, Irving H, Wondji CS 2016. The cytochrome P450 CYP6P4 is responsible
822 for the high pyrethroid resistance in knockdown resistance-free *Anopheles arabiensis*. Insect Biochem
823 Mol Biol 68: 23-32. doi: 10.1016/j.ibmb.2015.10.015

824 Ingham VA, Wagstaff S, Ranson H 2018. Transcriptomic meta-signatures identified in *Anopheles*
825 *gambiae* populations reveal previously undetected insecticide resistance mechanisms. Nat Commun 9:
826 5282. doi: 10.1038/s41467-018-07615-x

827 Itokawa K, et al. 2010. Genomic structures of Cyp9m10 in pyrethroid resistant and susceptible strains of
828 *Culex quinquefasciatus*. Insect Biochem Mol Biol 40: 631-640. doi: 10.1016/j.ibmb.2010.06.001

829 Izquierdo A, et al. 2019. Evolution of gene expression levels in the male reproductive organs of
830 *Anopheles* mosquitoes. Life Sci Alliance 2: e201800191. doi: 10.26508/lsa.201800191

831 Jiang PP, Hartl DL, Lemos B 2010. Y not a dead end: epistatic interactions between Y-linked regulatory
832 polymorphisms and genetic background affect global gene expression in *Drosophila melanogaster*.
833 Genetics 186: 109-118. doi: 10.1534/genetics.110.118109

834 Jiang X, Biedler JK, Qi Y, Hall AB, Tu Z 2015. Complete Dosage Compensation in *Anopheles stephensi*
835 and the Evolution of Sex-Biased Genes in Mosquitoes. Genome biology and evolution 7: 1914-1924. doi:
836 10.1093/gbe/evv115

837 Jones CM, et al. 2013. The dynamics of pyrethroid resistance in *Anopheles arabiensis* from Zanzibar and
838 an assessment of the underlying genetic basis. Parasit Vectors 6: 343. doi: 10.1186/1756-3305-6-343

839 Jones CM, et al. 2012a. Additional selection for insecticide resistance in urban malaria vectors: DDT
840 resistance in *Anopheles arabiensis* from Bobo-Dioulasso, Burkina Faso. PLoS One 7: e45995. doi:
841 10.1371/journal.pone.0045995

842 Jones FC, et al. 2012b. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:
843 55-61. doi: 10.1038/nature10944

844 Jones JC 1957. A Simple Method for Sexing Living *Anopheles* Larvae (Diptera, Culicidae). Annals of the
845 Entomological Society of America 50: 104-106. doi: 10.1093/aesa/50.1.104

846 Kayserili MA, Gerrard DT, Tomancak P, Kalinka AT 2012. An excess of gene expression divergence on
847 the X chromosome in *Drosophila* embryos: implications for the faster-X hypothesis. PLoS Genet 8:
848 e1003200. doi: 10.1371/journal.pgen.1003200

849 Kerby DS 2014. The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation.
850 Comprehensive Psychology 3. doi: 10.2466/11.lt.3.1

851 Khaitovich P, et al. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and
852 chimpanzees. Science 309: 1850-1854. doi: 10.1126/science.1108296

853 Lai YT, et al. 2019. Standing genetic variation as the predominant source for adaptation of a songbird.
854 Proceedings of the National Academy of Sciences of the United States of America 116: 2152-2157. doi:
855 10.1073/pnas.1813597116

856 Laporte V, Charlesworth B 2002. Effective population size and population subdivision in demographically
857 structured populations. Genetics 162: 501-519.

858 Lavington E, Kern AD 2017. The Effect of Common Inversion Polymorphisms In(2L)t and In(3R)Mo on
859 Patterns of Transcriptional Variation in *Drosophila melanogaster*. G3 (Bethesda, Md) 7: 3659-3668. doi:
860 10.1534/g3.117.1133

861 Law CW, Chen Y, Shi W, Smyth GK 2014. voom: Precision weights unlock linear model analysis tools for
862 RNA-seq read counts. Genome biology 15: R29. doi: 10.1186/gb-2014-15-2-r29

863 Lawniczak MK, et al. 2007. Mating and immunity in invertebrates. Trends Ecol Evol 22: 48-55. doi:
864 10.1016/j.tree.2006.09.012

865 Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England)
866 25: 2078-2079. doi: 10.1093/bioinformatics/btp352

867 Llopart A 2018. Faster-X evolution of gene expression is driven by recessive adaptive cis-regulatory
868 variation in *Drosophila*. Mol Ecol 27: 3811-3821. doi: 10.1111/mec.14708

869 Llopart A 2012. The rapid evolution of X-linked male-biased gene expression and the large-X effect in
870 *Drosophila yakuba*, *D. santomea*, and their hybrids. Mol Biol Evol 29: 3873-3886. doi:
871 10.1093/molbev/mss190

872 Lucas ER, et al. 2019. Whole-genome sequencing reveals high complexity of copy number variation at
873 insecticide resistance loci in malaria mosquitoes. Genome research 29: 1250-1261. doi:
874 10.1101/gr.245795.118

875 Magnusson K, et al. 2012. Demasculinization of the *Anopheles gambiae* X chromosome. BMC Evol Biol
876 12: 69. doi: 10.1186/1471-2148-12-69

877 Main BJ, et al. 2016. The Genetic Basis of Host Preference and Resting Behavior in the Major African
878 Malaria Vector, *Anopheles arabiensis*. PLoS Genet 12: e1006303. doi: 10.1371/journal.pgen.1006303

879 Marques-Bonet T, et al. 2004. Chromosomal rearrangements and the genomic distribution of gene-
880 expression divergence in humans and chimpanzees. Trends in Genetics 20: 524-529. doi:
881 10.1016/j.tig.2004.08.009

882 Meiklejohn CD, Coolon JD, Hartl DL, Wittkopp PJ 2014. The roles of cis- and trans-regulation in the
883 evolution of regulatory incompatibilities and sexually dimorphic gene expression. Genome research 24:
884 84-95. doi: 10.1101/gr.156414.113

885 Meiklejohn CD, Parsch J, Ranz JM, Hartl DL 2003. Rapid evolution of male-biased gene expression in
886 *Drosophila*. Proceedings of the National Academy of Sciences of the United States of America 100: 9894-
887 9899.

888 Meisel RP, Malone JH, Clark AG 2012. Faster-X evolution of gene expression in *Drosophila*. PLoS Genet
889 8: e1003013. doi: 10.1371/journal.pgen.1003013

890 Mitchell SN, Catteruccia F 2017. Anopheline Reproductive Biology: Impacts on Vectorial Capacity and
891 Potential Avenues for Malaria Control. Cold Spring Harb Perspect Med 7. doi:
892 10.1101/cshperspect.a025593

893 Mitchell SN, et al. 2015. Mosquito biology. Evolution of sexual traits influencing vectorial capacity in
894 anopheline mosquitoes. Science 347: 985-988. doi: 10.1126/science.1259435

895 Muller L, et al. 2011. Population transcriptomics of *Drosophila melanogaster* females. BMC genomics 12.
896 doi: Artn 8110.1186/1471-2164-12-81

897 Neafsey DE, et al. 2010. SNP genotyping defines complex gene-flow boundaries among African malaria
898 vector mosquitoes. Science 330: 514-517. doi: 10.1126/science.1193036

899 Neafsey DE, et al. 2015. Mosquito genomics. Highly evolvable malaria vectors: the genomes of 16
900 *Anopheles* mosquitoes. Science 347: 1258522. doi: 10.1126/science.1258522

901 Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM 2004. Common pattern of evolution of gene
902 expression level and protein sequence in *Drosophila*. Mol Biol Evol 21: 1308-1317. doi:
903 10.1093/molbev/msh128

904 Papa F, et al. 2017. Rapid evolution of female-biased genes among four species of *Anopheles* malaria
905 mosquitoes. Genome research 27: 1536-1548. doi: 10.1101/gr.217216.116

906 Powell JR 2018. Genetic Variation in Insect Vectors: Death of Typology? Insects 9. doi:
907 10.3390/insects9040139

908 Ranson H, et al. 2002. Evolution of supergene families associated with insecticide resistance. Science
909 298: 179-181. doi: 10.1126/science.1076781

910 Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL 2003. Sex-dependent gene expression and
911 evolution of the *Drosophila* transcriptome. Science 300: 1742-1745.

912 Raudvere U, et al. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of
913 gene lists (2019 update). Nucleic Acids Res 47: W191-W198. doi: 10.1093/nar/gkz369

914 Ritchie ME, et al. 2015. limma powers differential expression analyses for RNA-sequencing and
915 microarray studies. *Nucleic Acids Res* 43: e47. doi: 10.1093/nar/gkv007

916 Rivero A, Vezilier J, Weill M, Read AF, Gandon S 2010. Insecticide control of vector-borne diseases:
917 when is insecticide resistance a problem? *PLoS Pathog* 6: e1001000. doi: 10.1371/journal.ppat.1001000

918 Robinson MD, Oshlack A 2010. A scaling normalization method for differential expression analysis of
919 RNA-seq data. *Genome biology* 11: R25. doi: 10.1186/gb-2010-11-3-r25

920 Rogers DW, et al. 2009. Transglutaminase-mediated semen coagulation controls sperm storage in the
921 malaria mosquito. *PLoS biology* 7: e1000272. doi: 10.1371/journal.pbio.1000272

922 Rogers DW, et al. 2008. Molecular and cellular components of the mating machinery in *Anopheles*
923 *gambiae* females. *Proceedings of the National Academy of Sciences of the United States of America* 105:
924 19390-19395. doi: 10.1073/pnas.0809723105

925 Rono MK, Whitten MM, Oulad-Abdelghani M, Levashina EA, Marois E 2010. The major yolk protein
926 vitellogenin interferes with the anti-plasmodium response in the malaria mosquito *Anopheles gambiae*.
927 *PLoS biology* 8: e1000434. doi: 10.1371/journal.pbio.1000434

928 Roux O, Robert V 2019. Larval predation in malaria vectors and its potential implication in malaria
929 transmission: an overlooked ecosystem service? *Parasit Vectors* 12: 217. doi: 10.1186/s13071-019-3479-
930 7

931 Sacher RM 1971. A mosquito larvicide with favorable environmental properties. *Mosquito News* 31: 513-
932 516.

933 Santolamazza F, et al. 2008. Insertion polymorphisms of SINE200 retrotransposons within speciation
934 islands of *Anopheles gambiae* molecular forms. *Malar J* 7: 163. doi: 10.1186/1475-2875-7-163

935 Simma EA, et al. 2019. Genome-wide gene expression profiling reveals that cuticle alterations and P450
936 detoxification are associated with deltamethrin and DDT resistance in *Anopheles arabiensis* populations
937 from Ethiopia. *Pest Manag Sci* 75: 1808-1818. doi: 10.1002/ps.5374

938 Tripet F, Toure YT, Dolo G, Lanzaro GC 2003. Frequency of multiple inseminations in field-collected
939 *Anopheles gambiae* females revealed by DNA analysis of transferred sperm. *Am J Trop Med Hyg* 68: 1-5.

940 Vicoso B, Charlesworth B 2009. Effective population size and the faster-X effect: an extended model.
941 *Evolution* 63: 2413-2426. doi: 10.1111/j.1558-5646.2009.00719.x

942 Voolstra C, Tautz D, Farbrother P, Eichinger L, Harr B 2007. Contrasting evolution of expression
943 differences in the testis between species and subspecies of the house mouse. *Genome research* 17: 42-
944 49. doi: 10.1101/gr.5683806

945 Waterhouse RM, et al. 2020. Evolutionary superscaffolding and chromosome anchoring to improve
946 *Anopheles* genome assemblies. *BMC Biol* 18: 1. doi: 10.1186/s12915-019-0728-3

947 White BJ, Collins FH, Besansky NJ 2011. Evolution of *Anopheles gambiae* in Relation to Humans and
948 Malaria. *Annual Review of Ecology, Evolution, and Systematics*, Vol 42 42: 111-132. doi:
949 10.1146/annurev-ecolsys-102710-145028

950 WHO. 2020. World Malaria Report 2020. In. Geneva.

951 Yeh SD, et al. 2014. Functional divergence of the miRNA transcriptome at the onset of *Drosophila*
952 metamorphosis. *Mol Biol Evol* 31: 2557-2572. doi: 10.1093/molbev/msu195

953 Yuval B 2006. Mating systems of blood-feeding flies. *Annual review of entomology* 51: 413-440.

954 Zhang Y, Sturgill D, Parisi M, Kumar S, Oliver B 2007. Constraint and turnover in sex-biased gene
955 expression in the genus *Drosophila*. *Nature* 450: 233-237.

956