

RESEARCH ARTICLE

Cytotaxonomic characterization and estimation of migration patterns of onchocerciasis vectors (*Simulium damnosum sensu lato*) in northwestern Ethiopia based on RADSeq data

Shannon M. Hedtke¹, Rory J. Post^{2,3}, Sindew Mekasha Feleke⁴, Fikre Seife Gebretsadik⁵, Daniel A. Boakye⁶, Andreas Krueger⁷, Warwick N. Grant¹, Craig S. Wilding^{2*}

1 Department of Environment and Genetics, School of Agriculture, Biomedicine and Environment, La Trobe University, Bundoora, Victoria, Australia, **2** School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, United Kingdom, **3** Disease Control Department, London School of Hygiene and Tropical Medicine, London, United Kingdom, **4** Ethiopia Public Health Institute, Addis Ababa, Ethiopia, **5** Neglected Tropical Disease Prevention and Control Program, Federal Ministry of Health, Addis Ababa, Ethiopia, **6** Parasitology Department, Noguchi Memorial Institute for Medical Research, Accra, Ghana, **7** Military Hospital Hamburg, Department Tropical Medicine, Hamburg, Germany

* C.S.Wilding@ljmu.ac.uk



OPEN ACCESS

Citation: Hedtke SM, Post RJ, Feleke SM, Gebretsadik FS, Boakye DA, Krueger A, et al. (2024) Cytotaxonomic characterization and estimation of migration patterns of onchocerciasis vectors (*Simulium damnosum sensu lato*) in northwestern Ethiopia based on RADSeq data. *PLoS Negl Trop Dis* 18(1): e0011868. <https://doi.org/10.1371/journal.pntd.0011868>

Editor: James Cotton, University of Glasgow, UNITED KINGDOM

Received: June 8, 2023

Accepted: December 17, 2023

Published: January 4, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pntd.0011868>

Copyright: © 2024 Hedtke et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All RADSeq data are deposited in GenBank as a Sequence Read Archive

Abstract

Background

While much progress has been made in the control and elimination of onchocerciasis across Africa, the extent to which vector migration might confound progress towards elimination or result in re-establishment of endemism in areas where transmission has been eliminated remains unclear. In Northern Ethiopia, Metema and Metekel—two foci located near the Sudan border—exhibit continuing transmission. While progress towards elimination has been faster in Metema, there remains a problematic hotspot of transmission. Whether migration from Metekel contributes to this is currently unknown.

Methodology/Principle findings

To assess the role of vector migration from Metekel into Metema, we present a population genomics study of 151 adult female vectors using 47,638 RADseq markers and mtDNA *CoI* sequencing. From additional cytotaxonomy data we identified a new cytoform in Metema, closely related to *S. damnosum s.str.*, here called the Gondar form. RADseq data strongly indicate the existence of two distinctly differentiated clusters within *S. damnosum s.l.*: one genotypic cluster found only in Metema, and the second found predominantly in Metekel. Because blackflies from both clusters were found in sympatry (in all four collection sites in Metema), but hybrid genotypes were not detected, there may be reproductive barriers preventing interbreeding. The dominant genotype in Metema was not found in Metekel while the dominant genotype in Metekel was found in Metema, indicating that (at the time of sampling) migration is primarily unidirectional, with flies moving from Metekel to Metema. There

(Bioproject PRJNA819953). Unique Col sequences have been submitted to NCBI's GenBank database (<https://www.ncbi.nlm.nih.gov/>) with accession numbers OM866887-OM866988.

Funding: The genetic analysis work in this study was financially supported by the Global Challenge Research Fund (to CSW) and field collections were funded by the End Fund inc. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No authors received any salary from the funder of this study.

Competing interests: The authors have declared that no competing interests exist.

was strong differentiation between clusters but little genetic differentiation within clusters, suggesting migration and gene flow of flies within the same genetic cluster are sufficient to prevent genetic divergence between sites.

Conclusions/Significance

Our results confirm that Metekel and Metema represent different transmission foci, but also indicate a northward movement of vectors between foci that may have epidemiological importance, although its significance requires further study.

Author summary

Onchocerciasis is a severely debilitating disease caused by infection with a parasitic nematode worm (*Onchocerca volvulus*) transmitted from person to person by blood-sucking blackflies (Diptera: Simuliidae). Ethiopia has set a goal to eliminate the transmission of onchocerciasis through administration of the drug ivermectin, with 80% coverage, to all endemic areas. One potential challenge that might impede elimination in Ethiopia (and elsewhere in Africa) is the movement of vector blackflies carrying parasites between endemic areas, leading to failure to eliminate transmission. Cytogenetic data are the primary means for delimiting vector species, and we report the detection of a new cytotaxonomic form in northwestern Ethiopia, the “Gondar form” of *Simulium damnosum s.l.* Comparisons of nuclear genetic diversity of vectors in northwest Ethiopia suggests genetic differentiation consistent with reduction in mating between flies from two foci of ongoing transmission—Metema and Metekel. These results further suggest movement of blackflies is likely from Metekel, an area with ongoing transmission, into Metema, an area where persistent hotspots of transmission have been recently identified. Further research is warranted to determine whether the migration of blackflies between these areas that results in the observed genetic similarity is occurring at a timescale epidemiologically relevant to the persistence of these transmission hotspots.

Introduction

Onchocerciasis (river blindness) is a severely debilitating parasitic disease caused by infection with the filarial nematode *Onchocerca volvulus*. Symptoms include impaired vision (including blindness), skin disease (skin lesions, intolerable itching, and depigmentation leading to social exclusion), and occasional neurological-endocrine disorders (epilepsy and Nakalanga syndrome) [1–4]. Worldwide, approximately 244 million people are at risk of onchocerciasis, with >99% of affected individuals living in sub-Saharan Africa [5]. Annual, biannual, or even quarterly mass drug administration of ivermectin (MDA) was effective at eliminating transmission in most endemic areas in the Americas; MDA implemented using community-directed treatment with ivermectin (CDTI), initially by the African Programme for Onchocerciasis Control (APOC; 1995–2015), has been similarly effective at eliminating transmission in some parts of Africa [6,7]. Consequently, endemic countries in Africa with encouragement, co-ordination, and advice from the World Health Organization (WHO) have declared their intention to eliminate transmission [6], and in a significant number of endemic areas it has been possible to stop treatment because transmission has already been interrupted.

However, these early successes tended to have been in areas where onchocerciasis was endemic in isolated foci (of various sizes) such as those in Uganda, Sudan, and countries in the Americas [8]. Infective vector migration might interfere with progress towards elimination where endemic areas are geographically closer together or continuously distributed, as is the case for much of sub-Saharan Africa [9]. Long-distance migration by infective adult female blackflies of some vector species in West Africa brought significant numbers of parasites into areas where they had been controlled by larviciding during the WHO Onchocerciasis Control Programme (1975–2022) (before the widespread adoption of CDTI) and were considered a serious threat to the control of onchocerciasis [10–12]. Migration by infective vectors could similarly contribute to failure to achieve elimination of transmission by CDTI, or might result in re-establishment of endemism in areas where onchocerciasis was previously considered to have been eliminated ([13–15]; see also e.g., [16–18] with regards to mosquito vectors).

In Ethiopia, over 17 million people live in endemic areas or are at risk of infection [19]. In 2013, the Ethiopian Federal Ministry of Health launched the Ethiopian Onchocerciasis Elimination Program with the initial goal of interrupting transmission by 2020 [19]. The main control strategy is CDTI delivered twice per year, with a goal of maintaining treatment coverage at >80% across all known endemic districts. While there has been considerable success, transmission has not been eliminated throughout the country and onchocerciasis remains an issue in some regions. In the Amhara region of Ethiopia, approximately 2 million people are at risk of onchocerciasis, including in the Ethiopian part of the Metema-Galabat cross-border focus (which consists of the Metema sub-focus in Ethiopia and the Galabat sub-focus directly adjacent in Sudan). In the Metema-Galabat focus, onchocerciasis has been well controlled, although a transmission hotspot has been identified on the Wudi Gemzu River in Metema [20]. Transmission of the parasite also continues in the Metekel focus 20 km to the south [20]. While CDTI has been increased to three times per year in the Metema hotspot [20], one concern is that blackfly immigration from Metekel could be responsible for ongoing transmission in the hotspot and/or re-introduce transmission of onchocerciasis into Metema should CDTI be stopped. Whether adult flies from Metekel are migrating into the cleared areas of the Metema focus is unknown, but the distance between the two foci is certainly within the known distance for long-range wind-assisted migration [10,12,21,22]. The epidemiological risk of blackfly migration should be incorporated into decisions about whether to stop CDTI in these areas, but there are currently no routinely employed methods by which the pattern and extent of vector migration can be measured. Mapping and quantifying the movement of blackflies would be useful for determining optimal strategies for successful elimination.

The situation is considerably complicated by the multiplicity of vectors. In Ethiopia a number of *Simulium* species have been recorded as anthropophilic (S1 Table) and amongst these *Simulium damnosum s.l.*, the vector of most onchocerciasis in Africa (including Ethiopia), is not a single species but a species complex. The historical development of our understanding of the *S. damnosum* complex in Ethiopia has been previously described [23–26]. Studies of the polytene chromosomes found in the silk glands of larvae have revealed many structural mutations (mostly inversions) which, according to population cytogenetic analyses, indicate distinct species known as sibling species or cytospecies because they are morphologically more or less identical but reproductively isolated [27]. Because not all chromosomal rearrangements have been proven to lead to reproductively isolated species [27], we will use the more general term “cytoform” for blackflies that are structurally distinct. However, cytoforms commonly show important differences in their roles as vectors [28,29]. There are approximately 31 distinct cytoforms recognized within the *S. damnosum* complex across Africa, with four in Ethiopia: Kisiwani, *S. kaffaense*, *S. soderense*, and *S. kulfoense* [30]. Of these, cytoform *S. kaffaense* is currently the only proven vector in Ethiopia [23,24]. However, cytoforms have not yet been

identified in many onchocerciasis-endemic areas in Ethiopia, and it is likely that additional variation exists, especially in areas with distinctive ecology (such as Metema). One challenge is that blackflies only have suitable polytene chromosomes for identification in the silk glands of the larvae (homologous to the salivary glands of other dipteran larvae), which does not allow for the direct identification of biting adult females. Morphological or molecular traits associated with cytoform identity would be more convenient, but despite more than 40 years of research, these have proved generally difficult to find [28].

Population genetics may assist in the assessment of the role played by vector migration (if any) in ongoing transmission, because genetic sequence similarities and differences can be used to estimate the per-generation rate of movement of blackflies from one onchocerciasis endemic area to another. *Simulium damnosum s.l.* breeding sites are found in fast-moving waters and can be permanent, seasonal, or ephemeral in location and productivity. This demographic instability, combined with local breeding and short generation times, is expected to increase similarity within a transmission zone due to inbreeding, while also increasing differences from other locations due to stochastic changes in allele frequencies (i.e., genetic drift). This increases our ability to detect clusters of genetic diversity that may be relevant to how flies are moving across the geographic landscape, and to detect individual flies that are migrants. Quantification of the genetic similarities and differences in samples of blackflies that are competent vectors of *O. volvulus* may be used to determine whether ongoing or renewed transmission of the parasite is likely to be driven by movement of blackflies.

In theory, cytotaxonomic variation (in the form of inversion frequencies) may also give some qualitative indication of migration, because differences in inversion frequencies between localities indicate that blackflies from these localities are not mixing freely. However, in practice, the number of characters (inversions) available for analysis is too small for quantitative assessment, and the number of specimens from which cytological data is derived are low. Furthermore, because inversions cause physical linkage between genes, natural selection could disturb signals of migration by rapidly removing chromosomal variation. The use of genetic information less likely to be subject to selection would be more informative about gene flow between areas. DNA barcoding utilizing sequencing of the mitochondrial cytochrome oxidase I (*COI*) gene [31] has been applied widely for studies of population subdivision, particularly species delimitation, including in blackflies (e.g. [32,33]). However, phylogenetic analyses of blackflies suggest that nuclear sequence data tend to be more congruent with species delimitation based on morphology and cytotaxonomy, while the mitochondrial genome appears prone to introgression, leading to stronger geographic similarities across species boundaries [34, 35]. Despite its epidemiological importance, there remains no nuclear genome available for any member of the *S. damnosum* species complex. However, recently developed reduced-representation protocols such as restriction-site associated DNA sequencing (RADseq) [36] can be used to interrogate variation at thousands of loci without the need for pre-existing genome assemblies.

Here, we describe analyses of chromosomal, mitochondrial, and nuclear genetic variation in anthropophilic blackflies from the Metema and Metekel regions of Ethiopia. The objectives were to identify vector cytoforms in Metema, and to examine patterns of genetic variation to explore the potential role of vector movement in transmission of *O. volvulus* within and between the onchocerciasis-endemic areas of Metema and Metekel.

Materials and methods

Study area

The study area consisted of a number of sub-parallel river basins running northwesterly from the Ethiopian highlands to join the River Nile. The Metema-Galabat onchocerciasis focus is

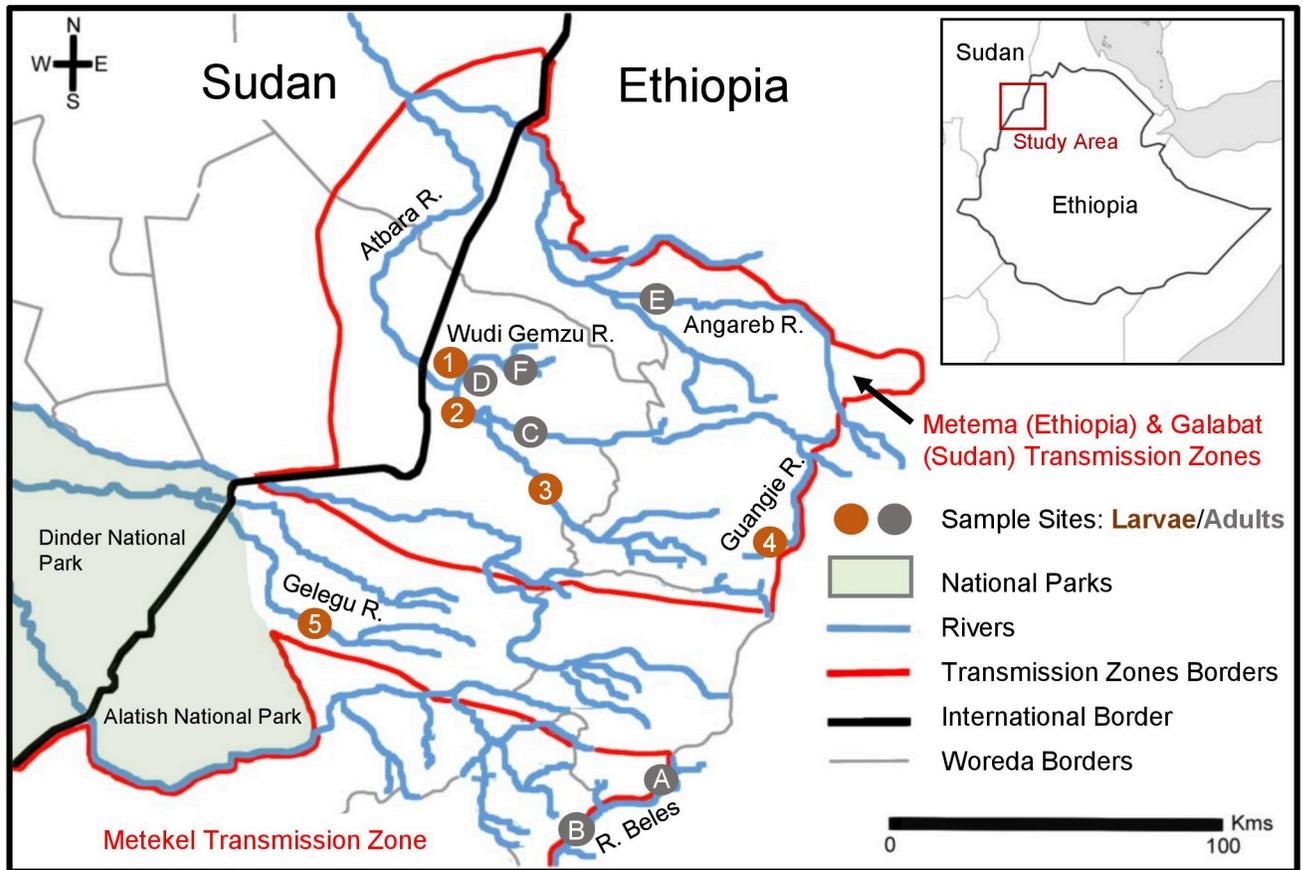


Fig 1. Map showing the Metema-Galabat and Metekel foci of onchocerciasis and collection sites sampled for *Simulium* in northwest Ethiopia. Transmission zone borders are outlined in red. Numbers/red circles (1–5) indicate larval samples collected in 2013 for cytotaxonomy and letters/grey circles indicate blood-seeking adult females collected in 2017–8 for DNA sequencing (see Tables 1 and S2). Basemap property of the Government of Ethiopia and supplied by the Ethiopian Ministry of Health.

<https://doi.org/10.1371/journal.pntd.0011868.g001>

centered on the Atbara river basin in Ethiopia and Sudan, but in Ethiopia it also includes some parts of the Angareb river basin which lies to the north of the Atbara. The two rivers flow NNW separately out of Ethiopia to join each other in Sudan and eventually to join the River Nile at Atbara town (north of Khartoum). South of the Metema sub-focus, an area free of onchocerciasis occupies the Shinfa river basin which flows NW into Sudan and joins the Blue Nile at Wad Madani. This area is referred to as a ‘buffer zone’ because it is situated between the Metema sub-focus to the north and the Metekel focus to the south in the Dinder River basin, which also flows NW into Sudan joining the Blue Nile upstream of Wad Madani (separately from the Shinfa). To the south of the Dinder river basin is the Blue Nile itself (Fig 1).

The altitude of the study area rises from 500 m near the border with Sudan to approximately 2000 m. The geology is largely Mesozoic volcanic basalts with many faults [37], and is thus suitable for the development of broken-water rapids and vector breeding sites. Meteorological data (www.ethiomet.gov.et –accessed 2013) indicates that in NW Ethiopia there is a single dry season October–April and a single rainy season May–September. The prevailing wind comes from the northeast during the dry season and from the south or southwest during the rainy season [38,39]. River discharge is undoubtedly seasonal, but there is little evidence about the extent to which rivers and streams actually cease to flow. Total rainfall is 800–1200 mm

Table 1. *Simulium* collection sites in northwestern Ethiopia. Larvae collected in 2013 (sites 1–5) were used for cytotaxonomy, and adult females (f) collected in 2017–18 (sites A–F) were used for DNA sequencing. See also Fig 1 and S2 Table.

Site	Transmission zone	Locality	River	Life stage	Number + <i>S. bovis</i>	Year of collection
1	Metema	Wudi Gemzu	Wudi Gemzu	larvae	10	2013
2	Metema	Guange N° 6	Kibe	larvae	3	2013
3	Metema	Meka	Meka	larvae	2	2013
4	Metema	Dubaba Kebele	Guangie	larvae	17	2013
5	Metema	Delegu Anana	Delegu Namu	larvae	4	2013
A	Metekel	Selassie Godiguadit	Tana Beles	adult f	33+0	2017
B	Metekel	Block 4	Abat Beles	adult f	33+0	2017
C	Metema	Asakefari	Asakefari	adult f	15+16	2018
D	Metema	Wudi Gemzu	Wudi Gemzu	adult f	29+2	2018
E	Metema	Kisha	Angareb	adult f	6+23	2017–18
F	Metema	Nega Wuha	Wudi Gemzu	adult f	32+0	2018

<https://doi.org/10.1371/journal.pntd.0011868.t001>

rainfall p.a., and the natural vegetation is a form of open savanna woodland known as ‘Combretum-Terminalia woodland and wooded grassland’ [40].

Cytotaxonomy

Larval specimens of the *S. damnosum* complex were collected in October 2013 (by DAB) from the Metema sub-focus of the Metema-Galabat onchocerciasis focus and from the buffer zone south of the Metema sub-focus, all in the North Gondar Zone of Amhara Region (Fig 1 and Tables 1 and S2). Larvae were hand-picked using forceps from vegetative and stony substrates found in white-water rapids and preserved in Carnoy’s fixative according to standard methods [41] and subsequently maintained at 4°C (except during transport at ambient temperature). Larvae of *S. damnosum s.l.* were identified according to Davies and Crosskey [41], and chromosome preparations were made from larval silk glands using standard methods and compared with the banding sequences of published chromosome maps [24,42–44].

Collection of anthropophilic adult female blackflies

Adult female appetitive blackflies were collected at various times during the period April 2017 to August 2018 from the Metema sub-focus (four sites) and the northern part of the Metekel focus (two sites), (Fig 1; Tables 1 and S2), using human landing catches between 07:00–18:00 by Ethiopia Ministry of Health and the Public Health Institute (EPHI) and staff using standard procedures [41]. Ethical approval was obtained from the Ethiopian Public Health Institute (EPHI) and local administrative offices prior to collection. All staff involved in collections received standard training, consented to undertake fly collections, and received ivermectin as part of the routine MDA. Flies were preserved in 95% ethanol in the field and subsequently maintained at 4°C (except during transport at ambient temperature). Preliminary morphological identification of flies was undertaken by EPHI in Ethiopia to remove any non-blackflies, and specimens were identified to morpho-species by RJP and CSW (according to [41,45,46]). Two species were identified among catches: *S. damnosum s.l.* and *Simulium bovis*.

DNA extraction from adult female blackflies

DNA was extracted from 192 specimens, including 41 *S. bovis* and 151 *S. damnosum s.l.*, using a GeneJet genomic DNA extraction kit (Thermo Fisher Scientific, Altrincham, UK) following the manufacturer’s instructions. DNA concentration was assessed using the Qubit dsDNA HS Assay Kit (Invitrogen, Thermo Fisher Scientific).

mtDNA barcoding

Partial mtDNA *CoI* fragments were amplified using the primers of Folmer *et al.* [47] at 0.2 μ M using 1 \times GoTaq HS mastermix (Promega Corp., Madison, Wisconsin, USA) with cycling conditions of 95°C for 3 min followed by 35 cycles of 95°C for 1 min, 40°C for 1 min, and 72°C for 1.5 mins, with a final extension step at 72°C for seven minutes. PCR products were purified using a GeneJet PCR purification kit (Thermo Fisher) and sequenced by GATC Biotech, Konstanz, Germany. Sequences were aligned using *ClustalW* v 2.0 [48]. A haplotype network of all available haplotypes was constructed using the TCS algorithm in *PopART* v 1.7 [49] and variability and dN/dS statistics calculated in *DnaSP* v 6 [50].

RADseq

A minimum of 100ng DNA was used for RADseq following digestion with *MsiI*. Libraries were sequenced on an Illumina NextSeq 500/550 by LGC Genomics GmbH (Berlin, Germany) using v2 chemistry with 150 bp paired-end reads. Reads were filtered for quality and adapter sequence and the *MsiI* restriction enzyme cut sites removed. Orthologs were identified and extracted using the program *Stacks* [51], with further filtering for quality [52] and parameter settings chosen using the r80 method [53] (S1 Text).

Population statistics and differentiation

Basic population genetic diversity statistics for each sampling location were estimated using *Stacks* [51]. Principal components (PCs) of genetic variation were plotted using the R library package *adegenet* v. 2.1.3 [54,55]. Linkage disequilibrium was estimated using r^2 for unphased genotypes, and sites with an estimated $r^2 > 0.1$ were pruned such that only the first site was retained using *vcftools* v. 0.1.16 [56]. Two alternative algorithmic approaches for identifying subpopulations of blackflies and computing the posterior probability of each blackfly's membership in each inferred subpopulation were used: (1) discriminant analysis of principle components (DAPC; [57]) using *adegenet* and (2) the program *structure* v.2.3.4 for Linux [58].

DAPC was used to infer the number of clusters or subpopulations suggested by the data using the Bayesian information criterion (BIC) and to assign individuals to those subpopulations. In a separate analysis, maximizing and plotting the differentiation among the six sampling locations was explored. *Structure* [58] was used to both define the number of subpopulations and to assign individuals to those subpopulations. For specifics on parameterization of these analyses, see S1 Text.

Results

Cytotaxonomy

Table 2 shows the inversions that are characteristic and diagnostic of the cytospecies and other cytoforms that are members of the *S. damnosum* subcomplex within the *S. damnosum* complex (i.e., they are closely related to *S. damnosum s.str.* [27]). Table 3 shows the status of inversions within the material examined from Metema. The frequencies of inversions in Metema, distribution between the sexes (potential sex-linkage), and photographs of selected chromosomal inversions are presented in S3–S6 Tables and S1–S6 Figs.

Only 36 specimens were analysed from Metema, but it was clear that all specimens belonged to the *S. damnosum* subcomplex because of their fixed inversions, especially 2L-C. However, beyond that, they did not correspond precisely to any known cytoform. Specimens homozygous for 2L-C.8 would normally be identified as *Simulium sirbanum*, while 2L-8 is

Table 2. Cytotaxonomic summary of the different members of the *Simulium damnosum* subcomplex (modified from [27]).

Cytoform	1S	1L	2S	2L	3S	3L
<i>S. damnosum</i> Nile form	<u>1.2</u> *.3*	<u>1.3</u>	St	<u>C</u>	St	<u>2</u>
<i>S. damnosum</i> Volta form	1.2.3	<u>1.3</u>	St	<u>C.8</u> *	St	<u>2.6</u> [◊]
<i>S. sirbanum</i> sirba form	<u>1.2.3</u> *	<u>1.3</u>	St	<u>C.8</u>	St	<u>2.6</u> [♦]
<i>S. sirbanum</i> sudanense form	<u>1.2.3</u> *	<u>1.3</u>	St	<u>C.8</u>	St	<u>2.6</u> [♦]
<i>S. sirbanum</i> type IV	<u>1.2.3</u>	<u>1.3</u>	St	<u>C.8</u>	St	<u>2</u> .(6?)
Hamadense	<u>1.2.3</u>	<u>1.3</u>	St	<u>C.8</u>	St	<u>2.6</u>

Notes: Fixed inversions are bold and underlined

*Sex-linked inversions

♦Inversion normally at high frequency

◊Inversion normally at low frequency. St denotes the standard chromosome sequence arbitrarily defined as that shown by *Simulium kilibanum* [34].

<https://doi.org/10.1371/journal.pntd.0011868.t002>

expected to be absent or sex-linked polymorphic in *S. damnosum s.str.* In the samples from North Gondar, it was found to be polymorphic but not sex-linked.

There are no chromosomal data available from other populations of the *S. damnosum* sub-complex in Sudan, South Sudan, or Ethiopia, except for Abu Hamed on the River Nile in Sudan [59]. The population examined from North Gondar is genetically differentiated from the Hamedense cytoform by eight inversions. Hamedense has no polymorphic inversions and nine fixed inversions (including 2L-8). Four of these fixed inversions are polymorphic in North Gondar, Ethiopia, with five additional polymorphic inversions unique to North Gondar. The populations in North Gondar zone and Abu Hamed are clearly different from each other, but there could still be some low-level genetic exchange.

Mitochondrial DNA barcoding

645 bp of *CoI* sequence was amplified successfully from 189 of the 192 samples. Unique sequences have been submitted to NCBI’s GenBank database (<https://www.ncbi.nlm.nih.gov/>) with accession numbers OM866887-OM866988. The haplotype network shows three main groupings of sequences with *S. damnosum s.l.* distinct from *S. bovis* (Fig 2). However, in addition to the main body of *S. damnosum s.l.* specimens, a small number ($n = 6$) appear genetically distinct from the other *S. damnosum s.l.* specimens, being separated by a minimum of 27 mutational steps. Although the closest match through a *BLAST* analysis [60] is in the *S. damnosum* species complex from Benin (Accession KF839980.1 [61]), due to the differences in sequence these are subsequently referred to here as “query”-*Simulium damnosum*.

Table 3. Cytotaxonomic summary of the *Simulium damnosum* subcomplex from the study area (Amhara Region, Ethiopia).

Site	Sample River	River Basin	n	1S	1L	2S	2L	3S	3L
1	Wudi Gemzu	Atbara-Angereb	10	<u>1.2.3</u>	<u>1.3.22</u> *.23*	St	<u>C.8.2b</u> *	St	<u>2.6</u>
2	Kibe	Atbara-Angereb	3	<u>1.2.3</u>	<u>1.3</u>	St	<u>C.8.2b</u> *	St	<u>2.6</u>
3	Meka	Atbara-Angereb	2	<u>1.2.3</u>	<u>1.3</u>	St	<u>C.8</u>	St	<u>2.6</u>
4	Guangie	Blue Nile-Rahad	17	<u>1.2.3</u>	<u>1.3.22</u> *.23*.24*	St	<u>C.8.2b</u> *.70*	St	<u>2.6</u>
5	Delegu Namu	Blue Nile-Dinder	4	<u>1.2.3</u>	<u>1.3</u>	St	<u>C.8</u>	St	<u>2.6</u>

Notes: Inversions which were only found in the homozygous state and are probably fixed inversions are shown in bold and underlined (although samples from Meka and Delegu Namu rivers were very small, and it is quite likely that inversion 1S-3 was polymorphic in these populations, as it is in the other three rivers); *New inversions unknown outside Ethiopia.

<https://doi.org/10.1371/journal.pntd.0011868.t003>

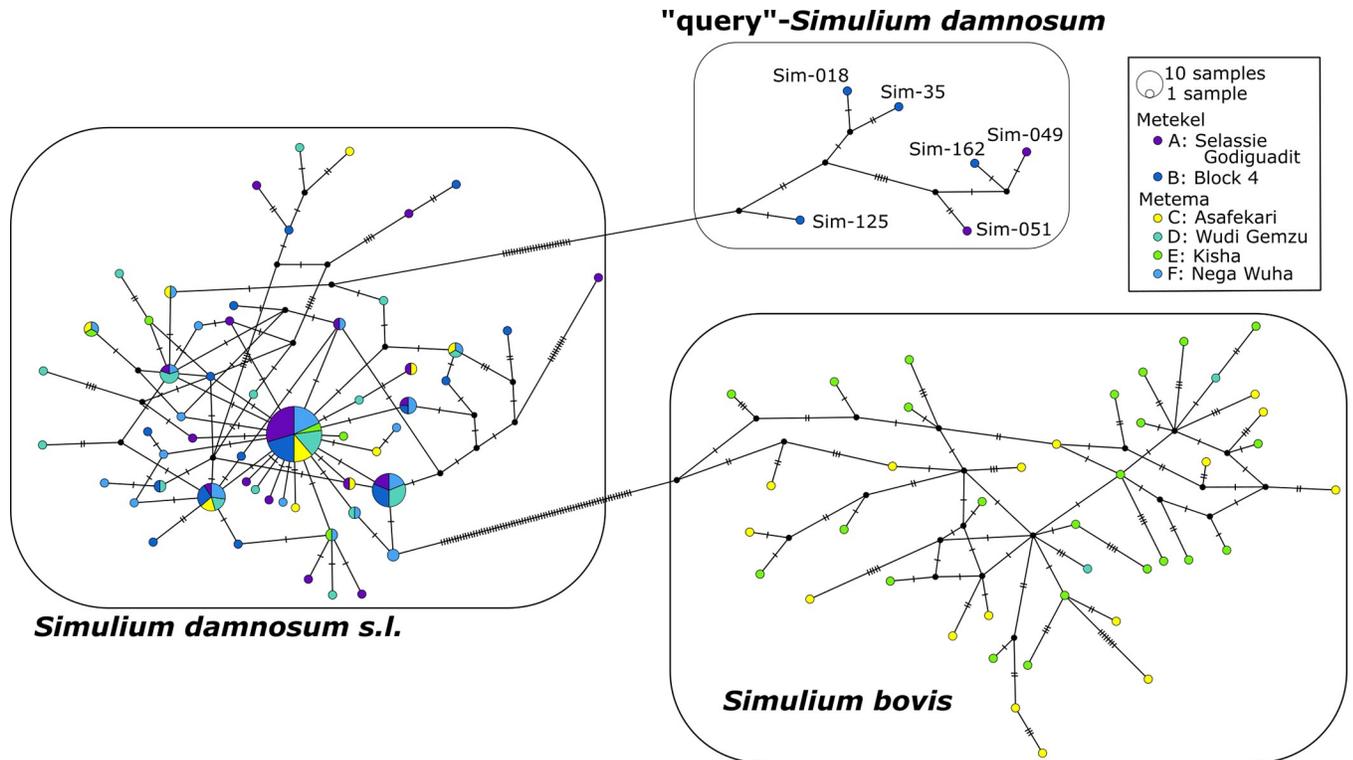


Fig 2. Mitochondrial *CoI* haplotype network of 189 *Simulium* from Metema and Metekel foci. Circles represent unique haplotypes; the size of the circles indicate the number of blackflies with that haplotype. Hatch marks along the lines connecting haplotypes indicate the number of sequence differences between them. Colors indicate the proportion of each haplotype found in each sampling location.

<https://doi.org/10.1371/journal.pntd.0011868.g002>

There were geographic differences in the distribution of *S. bovis*, *S. damnosum*, and “query”-*Simulium damnosum*. *Simulium bovis* was found only in 4 sites in Metema (C, D, E, F) and at sites D and F only at low frequency, while query-*damnosum* were found only at sites A and B in Metekel (Table 4).

Table 4. Numbers of *Simulium damnosum s.l.*, query-*S. damnosum s.l.*, and *S. bovis* as determined by mitochondrial DNA *CoI* analysis among 192 specimens extracted for molecular analysis. Numbers in parentheses represent samples morphologically identified as *S. damnosum s.l.* but which failed to amplify usable *CoI* barcodes.

Collection Site (River)		Species		
		<i>S. damnosum s.l.</i>	“Query” <i>S.damn</i> osum	<i>S.bovis</i>
A	Selassie Godiguadit (Tana Beles)	31	2	0
B	Block 4 (Abat Beles)	29	4	0
C	Asakefari (Asakefari)	15(+2)	0	16
D	Wudi Gemzu (Wudi Gemzu)	29(+1)	0	2
E	Kisha (Angareb)	6	0	23
F	Nega Wuha (Wudi Gemzu)	32	0	0

Haplotype diversity based on *CoI* was high for *S. damnosum* ($\pi_h = 0.885$), and for “query”-*Simulium damnosum* and *S. bovis*, each haplotype was unique ($\pi_h = 1$; S7 Table). All *S. damnosum s.l.* populations (with the exception of site E, where there were only 6 samples sequenced) displayed significantly negative Tajima’s D statistics (over all *S. damnosum s.l.* combined, Tajima’s D = -2.236; S7 Table), indicative of the action of either a recent selective sweep of the mitochondrial lineage or a recent population expansion.

<https://doi.org/10.1371/journal.pntd.0011868.t004>

RADseq Stacks parameter estimation

All RADseq data are deposited in GenBank as a Sequence Read Archive (Bioproject PRJNA819953). Here, we report RADSeq results only from anthropophilic *S. damnosum s.l.* samples ($N = 151$). Per-fly sequencing statistics (at $M = 1$) can be found in S8 Table. However, seventeen flies were removed from downstream analysis as their estimated average depth was lower than 10 when $M = 1$ or $M = 8$: two from Selassie Godiguadit (A: SIM-68, SIM-179), two from Block 4 (B: SIM-53, SIM-106), five from Asafekari, Metema (C: SIM-19, SIM-21, SIM-37, SIM-111, SIM-210), five from Wudi Gemzu, Metema (D: SIM-41, SIM-59, SIM-95, SIM-166, SIM-221), two from Nega Wuha, Metema (F: SIM-64, SIM-136), and one from Kisha (E: SIM-116). Thus, 134 *S. damnosum s.l.* remained for downstream analyses: 62 collected in the Metekel region (A: 31, B: 31) and 72 collected in the Metema focus (C: 12, D: 26, F: 29, E: 5).

RADseq population statistics. After filtering, 2,506,166 bp were sequenced across 14,469 orthologous loci, within which 47,638 single nucleotide sites were variable in the dataset. Estimates of genetic diversity were similar across all populations, although the number of polymorphic sites and number of private alleles varied with the number of samples in the population (S7 Table). After pruning for linkage disequilibrium, 23,860 single nucleotide sites were variable in the dataset.

Two approaches were used to explore how flies sampled clustered with regards to geography, and both were congruent despite their methodological differences. DAPC is a multivariate approach using PCA to maximize differences among groups while minimizing variance within groups [57]. The first PCs from PCA were plotted and show strong genetic differentiation between samples collected in Metekel and those collected in Metema (Fig 3A). The first ten PCs were used for the DAPC analysis using six groups as determined by cross-validation (S7 Fig). DAPC maximizes the differentiation observed in PCs and is consistent with the PCA plot, with most flies from Metema clustered together and independent of flies from Metekel (Fig 3B). However, nine flies collected from sampling locations within the Metema focus do not fall within this cluster, instead grouping with flies from Metekel (SIM-24, SIM-29, SIM-43, SIM-97, SIM-151, SIM-154, SIM-173, SIM-186, SIM-188). When *adegenet* was utilized for model-based K inference, the BIC supported a model of two clusters (S8 Fig). With two clusters, 90 PCs were determined to be optimal for DAPC using cross-validation (S9 Fig), and cluster assignment was consistent with the six-population DAPC with regards to identifying flies from Metema that were genetically more similar to flies from Metekel (Fig 3C). We do note that the six flies with strongly differentiated mtDNA haplotypes (in the group 'query-*S. damnosum*') are not differentiated from other flies on the basis of these nuclear markers (Figs 3A–3C).

The Bayesian model-based method implemented in the program *structure* estimates the cluster membership of each individual given a certain number of possible clusters, K [58]. The algorithm optimizes the groupings of those individuals to meet the assumptions of Hardy-Weinberg and linkage equilibrium. The results of population assignment did not change across all values of K tested (2 through 8) and were identical to the results from the DAPC analysis: nine flies collected in the Metema focus were inferred to be in the cluster which contained the flies from Metekel (Fig 3C).

Discussion

The data presented here are from two distinct studies carried out to examine the cyto- and molecular genetic characteristics of vectors in the Metema and Metekel foci of northwestern Ethiopia. The cytogenetic study, carried out in 2013, sampled a relatively small number of larval *S. damnosum s.l.* and aimed to place vectors in the Metema focus into the current

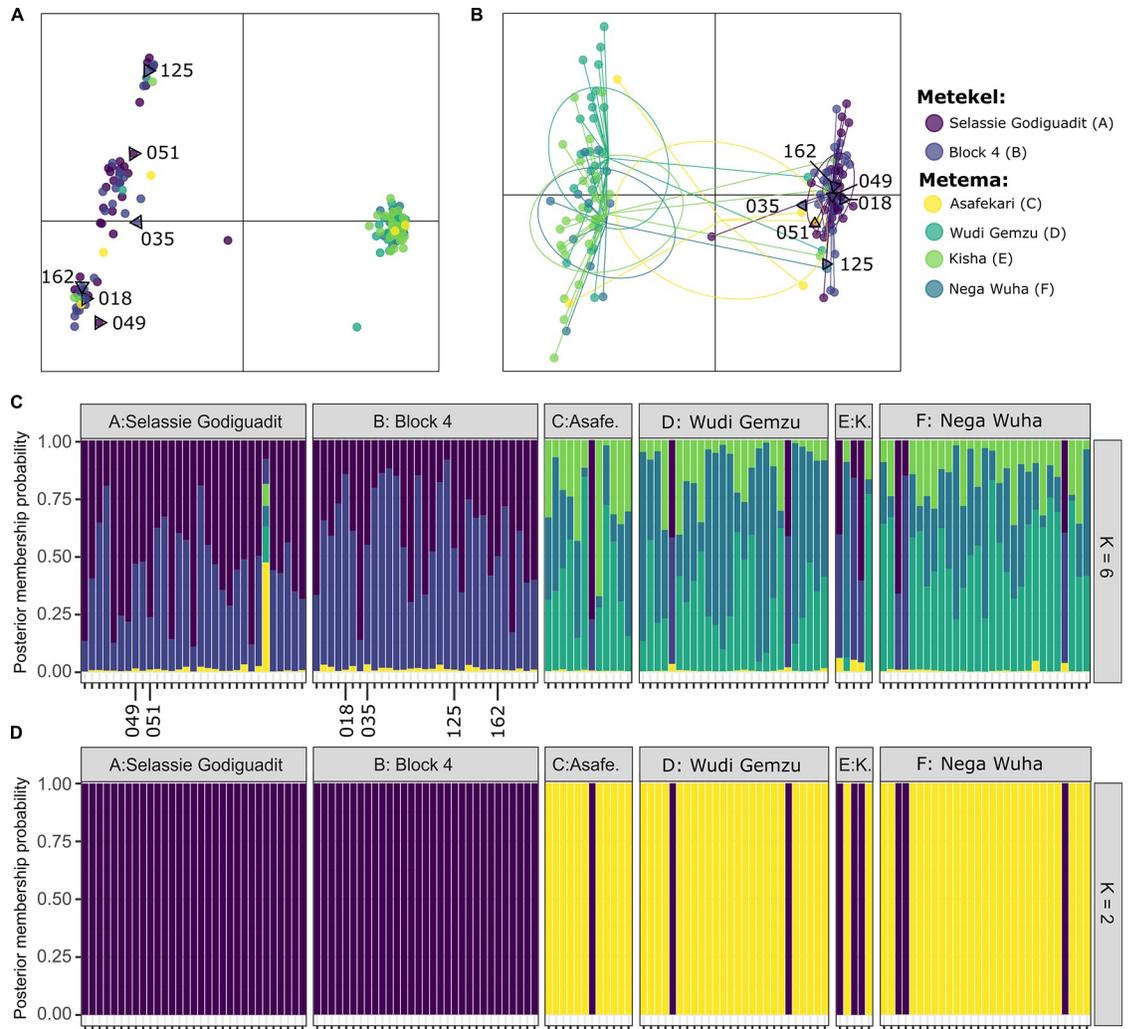


Fig 3. Population genetic structure of *Simulium damnosum* from Ethiopia genotyped at 23,860 variant sites. A) First (7.17% of total variance) and second (1.25% of variance) principal components (PCs) of variation in the genetic data (third and fourth PCs are presented in S10 Fig) and B) discriminant analysis of principal components (DAPC): each dot represents a genotyped fly; its color represents the sampling location. Lines connect to the center of the inertia ellipse for each sampling location. C) Genotype assignment of 134 flies based on DAPC for 6 sampling locations (K = 6) and D) based on assignment of each fly to one of two clusters (purple or yellow; K = 2). The results of *structure* with two to eight populations were identical to (D): all flies from Metekel were assigned to a cluster containing only Metekel flies, but 9 flies collected in Metema were inferred to be from this Metekel cluster. Numbered samples on A-C are the six “query”-*Simulium damnosum* specimens identified through mtDNA barcoding.

<https://doi.org/10.1371/journal.pntd.0011868.g003>

cytotaxonomic classification of *S. damnosum s.l.* DNA sequencing was performed using a larger, more geographically extensive sample of adult blackflies collected in 2018, and aimed to investigate the population structure of anthropophilic *Simulium* across the Metema and Metekel foci (without reference to their cytotaxonomic classification).

Cytotaxonomy and the Gondar form of the *S. damnosum* complex

The specimens examined cytotaxonomically from North Gondar are a previously unknown, genetically distinct cytoform within the *S. damnosum* subcomplex, and we have named this cytoform the ‘Gondar form’. It is unknown whether the Gondar form of the *S. damnosum* complex has a wider distribution in Ethiopia, or whether it extends into Sudan, because our

larval specimens were collected in 2013 and from the Metema sub-focus and the buffer zone immediately to the south only.

While we did not gather cytological data from samples collected in 2018, at least some of the adults collected in Metema may be the Gondar cytoform. Since no larvae were collected in Metekel, the adult blackflies collected there might be an as yet unidentified cytoform or could be a distinct lineage of the Gondar cytoform. Because rivers in northwestern Ethiopia are seasonal, breeding populations of *S. damnosum s.l.* are unstable, with vector migration between sites. It is therefore also possible that between 2013, when the blackflies used for cytotaxonomy were collected, and 2018, when collections for sequencing were done, the Gondar cytoform could have been replaced by some other cytoform(s) in the Metema sub-focus (e.g., fluctuations in cytoforms in Ghana [62]). This highlights one of the challenges in *Simulium* taxonomy as a whole: cytological data must be collected from larvae, whose adult characteristics may be unknown, while DNA is typically sequenced from adult flies, often collected because they are anthropophilic and thus potential disease vectors.

Two Groups Revealed by RADseq Analyses of Genomic DNA

Analysis of DNA sequence data (RADseq and mtDNA) was carried out on adult blackflies collected in 2018 and found two genetically distinct clusters of *S. damnosum s.l.*: one from Metema sub-focus and one from Metekel focus, with no evidence of hybridization between them (Fig 3C and 3D). Blackflies from both clusters are found in the Metema focus sampling sites, while members of only one cluster were found across Metekel focus sites. Within both the Metekel and Metema foci, there was little genetic differentiation among blackflies, suggesting that interbreeding of blackflies within the same genetic cluster is sufficient to prevent population structure among sampling locations within each focus. Furthermore, between locations, there was no evidence of differentiation between the blackflies of the same cluster collected in both Metema and Metekel. This suggests that there is or has been gene flow between the two foci, and that this gene flow is likely to be unidirectional from Metekel to Metema. There are two possible explanations for this gene flow: (1) all sampled blackflies are the same species, and those from the second group collected in Metema are very recent migrants from Metekel that have not yet interbred with the majority Metema genetic cluster (and/or sample sizes were too low to detect hybrids) or (2) both clusters are endemic to Metema but they represent species that do not successfully interbreed, and migration from Metekel to Metema homogenizes genetic variation in that species found in both locations. While we do not have cytogenetic data from blackflies from either genetic cluster, reproductive isolation has been shown to evolve in other dipterans as a result of chromosome inversions, and inversions have contributed to adaptive variation for traits such as preference for indoors versus outdoors, optimal climate, or insecticide resistance in mosquitoes (e.g., [63–65]).

Gene flow could occur as a result of either of two sorts of vector movements: long-distance migration (wind-assisted >100 km) and/or appetitive dispersal (active flight looking for a bloodmeal up to 20 km). The wind direction in this area is seasonal, from the southwest in the rainy season and from the northeast in the dry season. While northerly and southerly wind-assisted migration could both theoretically occur, ephemeral breeding sites affected by reduced rainfall would mean that far fewer flies would be available to be transported in the dry season than in the rainy season, when flies are abundant and winds are from the southwest. Such northerly migration could therefore result in long-distance migration from Metekel to Metema.

Appetitive dispersal does not occur over distances that would permit an individual blackfly to migrate through the buffer zone from Metekel to Metema. However, if suitable habitat is

continuous, regular small-scale migration could homogenize the blackflies in the two locations by gradual movement of genetic diversity across the landscape (i.e., gene flow through dispersal). This “stepping-stone” migration pattern would homogenize genetic diversity within a species, as might be expected to occur if the Metekel genetic cluster were a separate species with low differentiation across its sampled range.

The consequences of these two migration hypotheses on the epidemiology of onchocerciasis in northwestern Ethiopia are drastically different. If wind-dispersed in a single step, infective blackflies from an endemic area with ongoing transmission (Metekel) could cause hot-spots of transmission to arise in an area close to elimination (Metema-Galabat) despite continuing drug treatment in Metema. If dispersal through the buffer zone is occurring in small steps between breeding sites over several generations, then the risk of recrudescence due to infective blackflies is low and would not need to be considered for stop-treatment decisions. Sampling within the buffer zone and/or sampling of blackflies at multiple time points would be required to discriminate between these hypotheses.

Incongruent results from mitochondrial, nuclear, and cytotoxic data

While nuclear RADSeq markers effectively discriminate samples in our study, this is not the case for mitochondrial DNA. *CoI* haplotypes are shared across foci and provide no evidence of the discrimination of foci seen from multi-locus RADSeq markers. However, six flies from Metekel (two from Selassie Godiguadit [A] and four from Block 4 [B]) were genetically distinct based on mitochondrial genotyping and are clearly separated from the main body of *S. damnosum s.l.* samples in the mtDNA haplotype network (Fig 2). The high genetic distance between these six samples and all other *S. damnosum* is at odds with the nuclear DNA data (in which these six samples are indistinguishable from other Metekel flies). Previous phylogenetic analyses of blackflies from Africa have also found incongruence between trees built from mitochondrial data and those from both nuclear sequence data and cytotoxicity (e.g., [34,35]). When incongruence is observed, it can be attributed to biological processes—such as incomplete lineage sorting (when shared ancestral polymorphism is randomly lost in one population) or introgression (interspecific hybridization followed by unidirectional backcrossing)—or to systematic errors in phylogenetic reconstruction [66], or even to PCR/sequencing error. Because of the substantial number of (presumably) independent nuclear loci reported here, and the depth of differentiation between the two mitotypes, systematic error is not a convincing explanation for the incongruence. However, testing the hypothesis for mitochondrial introgression from an unsampled species is not possible because of the absence of available *CoI* barcodes for all *Simulium* species in this region of Africa.

Conclusions and significance

This study revealed that there was at least one new and unique cytotoxic form found within the *S. damnosum* complex in the study area and we have proposed to call it the Gondar form of the *S. damnosum* complex. The RADseq data indicated two genetically distinct clusters within the *S. damnosum* complex in the study area. It is possible, but unproven, that one or both of these genetic clusters is the novel cytotoxic form. Since these blackflies have not been characterized before, very little is known about whether they might exhibit traits which differ significantly from other known cytospecies for the transmission of onchocerciasis.

The RADseq analyses supports the hypothesis of northward geneflow from Metekel into Metema, but it remains unclear whether this represents long-distance movement of infective blackflies, which could cause a problem to onchocerciasis elimination efforts by bringing

parasites from Metekel into Metema, or gradual movement of genes north through the buffer zone, which would be less likely to introduce parasites. To quantify the risk of migration, it would be necessary to collect specimens regularly throughout the year and to demonstrate that at least some migrants were infective. However, this study has provided a proof of principle that nuclear genetic data (in this case partial sequencing using the RADseq approach) can reveal migration patterns in blackfly vectors of onchocerciasis.

Supporting information

S1 Text. Preparation and analysis of RADseq data.

(DOCX)

S1 Table. List of *Simulium* species recorded as anthropophilic in Ethiopia.

(DOCX)

S2 Table. Sample sites in northwestern Ethiopia where larvae or biting adult female *Simulium* sp. were collected.

(DOCX)

S3 Table. Inversion frequencies in the *Simulium damnosum* subcomplex from Ethiopia.

(DOCX)

S4 Table. Karyotype distribution of *Simulium damnosum* subcomplex from Ethiopia: Chromosome 2.

(DOCX)

S5 Table. Karyotype distribution of *S. damnosum* subcomplex from Ethiopia: Chromosome 1L.

(DOCX)

S6 Table. Karyotype distribution of *S. damnosum* subcomplex from Ethiopia: Chromosome 1S.

(DOCX)

S7 Table. Genetic diversity statistics across species and populations of *Simulium damnosum s.l.*, *S. bovis*, and the query-*damnosum* samples based on CoI DNA sequence data.

(DOCX)

S8 Table. Sequencing statistics for RADSeq data based on STACKS [1] when $M = 1$.

(DOCX)

S9 Table. Estimates of population genetic diversity of *Simulium damnosum sensu lato* samples collected from six locations in Ethiopia averaged across 47,638 variant sites.

(DOCX)

S10 Table. Estimates of pairwise genetic distance (below the diagonal) and F_{ST} (above the diagonal) for six locations in Ethiopia based on 23,860 variant sites of *Simulium damnosum s.l.* collected in 6 locations in Ethiopia.

(DOCX)

S1 Fig. Inversion 2L-2b marked on chromosome homozygous 2L-C.2b/C.2b.

(DOCX)

S2 Fig. Heterozygous inversion 2L-70.

(DOCX)

S3 Fig. Homozygote 2L-C.8.70/C.8.70 showing breakpoints of inversion 2L-70.
(DOCX)

S4 Fig. Breakpoints of Inversions 1L-22, 1L-23 & 1L-24 on chromosome 1L-1.3/1.3. Also showing heterozygous inversion 1S-2 on chromosome 1S-2.3/3.
(DOCX)

S5 Fig. Heterozygous inversions 1L-22 & 1L-23 on chromosome 1L-1.3/1.3.22.23.
(DOCX)

S6 Fig. Heterozygous inversions 1L-23 & 1L-24 on chromosome 1L-1.3/1.3.23.24.
(DOCX)

S7 Fig. Cross-validation for determining the optimal number of principle components to use in a discriminant analysis of principle components when $K = 6$, the number of sampling sites for *Simulium damnosum s.l.* flies collected for sequencing.
(DOCX)

S8 Fig. Bayesian Information Criterion (BIC) for the number of clusters in nuclear sequence data of *Simulium damnosum s.l.* from Ethiopia from $K = 1$ through $K = 40$.
(DOCX)

S9 Fig. Results of cross-validation using 100 replicates for determining the optimal number of principle components to use in a discriminant analysis of principle components when $K = 2$, the number of clusters inferred based on nuclear sequence data of *Simulium damnosum s.l.* flies collected in Ethiopia.
(DOCX)

S10 Fig. The third (0.89% of all variance) and fourth (0.87%) principal components from analysis of 23,860 genetic variants in linkage equilibrium of *Simulium damnosum s.l.* blackflies collected at 6 sites in Ethiopia.
(DOCX)

Acknowledgments

We are grateful to Mr. Kalkidan Mekete, Mr. Gemechu Tadesse, Mr. Mulugeta Ejigu and Mr. Kadu Meribo from EPHI and FMOH for conducting the field collections. We additionally thank Himal Shrestha and Emily Hendrickson at La Trobe University for useful discussion of the results and the WHO/APOC and The Carter Center staff who participated in supervision and training.

Author Contributions

Conceptualization: Rory J. Post, Craig S. Wilding.

Data curation: Shannon M. Hedtke, Rory J. Post, Daniel A. Boakye, Andreas Krueger, Craig S. Wilding.

Formal analysis: Shannon M. Hedtke, Rory J. Post, Daniel A. Boakye, Andreas Krueger, Craig S. Wilding.

Funding acquisition: Craig S. Wilding.

Investigation: Shannon M. Hedtke, Rory J. Post, Sindew Mekasha Feleke, Fikre Seife Gebretsadik, Daniel A. Boakye, Andreas Krueger, Craig S. Wilding.

Methodology: Shannon M. Hedtke, Rory J. Post, Craig S. Wilding.

Project administration: Rory J. Post, Warwick N. Grant, Craig S. Wilding.

Resources: Rory J. Post, Sindew Mekasha Feleke, Fikre Seife Gebretsadik, Daniel A. Boakye, Andreas Krueger.

Supervision: Rory J. Post, Warwick N. Grant.

Validation: Shannon M. Hedtke, Rory J. Post, Craig S. Wilding.

Visualization: Shannon M. Hedtke, Rory J. Post, Sindew Mekasha Feleke, Fikre Seife Gebretsadik, Warwick N. Grant, Craig S. Wilding.

Writing – original draft: Shannon M. Hedtke, Rory J. Post, Craig S. Wilding.

Writing – review & editing: Shannon M. Hedtke, Rory J. Post, Warwick N. Grant, Craig S. Wilding.

References

1. Chesnais CB, Nana-Djeunga HC, Njamnshi AK, Lenou-Nanga CG, Boule C, Bissek AZ, et al. The temporal relationship between onchocerciasis and epilepsy: a population-based cohort study. *Lancet Infect Dis*. 2018; 18(11):1278–86. [https://doi.org/10.1016/S1473-3099\(18\)30425-0](https://doi.org/10.1016/S1473-3099(18)30425-0) PMID: 30268645
2. Fodjo JNS, Mandro M, Mukendi D, Tepage F, Menon S, Nakato S, et al. Onchocerciasis-associated epilepsy in the Democratic Republic of Congo: clinical description and relationship with microfilarial density. *PLoS Negl Trop Dis*. 2019; 13(7):e0007300. <https://doi.org/10.1371/journal.pntd.0007300> PMID: 31314757
3. Mukendi D, Tepage F, Akonda I, Siewe JNF, Rotsaert A, Ndimun CN, et al. High prevalence of epilepsy in an onchocerciasis endemic health zone in the Democratic Republic of the Congo, despite 14 years of community-directed treatment with ivermectin: A mixed-method assessment. *Int J Infect Dis*. 2019; 79:187–94. <https://doi.org/10.1016/j.ijid.2018.10.021> PMID: 30711145
4. Colebunders R, Njamnshi AK, van Oijen M, Mukendi D, Kashama JM, Mandro M, et al. Onchocerciasis-associated epilepsy: From recent epidemiological and clinical findings to policy implications. *Epilepsia Open*. 2017; 2(2):145–52. <https://doi.org/10.1002/epi4.12054> PMID: 29588943
5. World Health Organization. Elimination of human onchocerciasis: progress report, 2021 –Élimination de l'onchocercose humaine: rapport de situation, 2021. *Weekly Epidemiological Record*. 2022; 96(46):591–8.
6. World Health Organization, African Programme for Onchocerciasis Control. Conceptual and operational framework of onchocerciasis elimination with ivermectin treatment. Geneva: African Programme for Onchocerciasis Control; 2010.
7. World Health Organization. Elimination of human onchocerciasis: progress report, 2020. *Weekly Epidemiological Record*. 2021; 46(96):557–67.
8. Lakwo T, Oguttu D, Ukety T, Post R, Bakajika D. Onchocerciasis elimination: progresses and challenges. *Res Rep Trop Med*. 2020; 11:81–95. <https://doi.org/10.2147/RRTM.S224364> PMID: 33117052
9. Zouré HG, Noma M, Tekle AH, Amazigo UV, Diggle PJ, Giorgi E, Remme JH. The geographic distribution of onchocerciasis in the 20 participating countries of the African Programme for Onchocerciasis Control: (2) pre-control endemicity levels and estimated number infected. *Parasit Vectors*. 2014; 7:326. <https://doi.org/10.1186/1756-3305-7-326> PMID: 25053392
10. Baker RH, Guillet P, Seketeli A, Poudiougo P, Boakye D, Wilson MD, et al. Progress in controlling the reinvasion of windborne vectors into the western area of the Onchocerciasis Control Programme in West Africa. *Philos Trans R Soc Lond B Biol Sci*. 1990; 328(1251):731–47. <https://doi.org/10.1098/rstb.1990.0141> PMID: 1976266
11. Boatman B. The Onchocerciasis Control Programme in West Africa (OCP). *Ann Trop Med Parasitol*. 2008; 102(Suppl 1):13–7. <https://doi.org/10.1179/136485908X337427> PMID: 18718148
12. Le Berre R, Garms R, Davies JB, Walsh JF, Philippon B, Johnson CG, et al. Displacements of *Simulium damnosum* and strategy of control against onchocerciasis. *Philos Trans R Soc Lond B Biol Sci*. 1979; 287(1022):277–88. <https://doi.org/10.1098/rstb.1979.0061> PMID: 43521
13. Bakajika D, Senyonjo L, Enyong P, Oye J, Biholong B, Elhassan E, et al. On-going transmission of human onchocerciasis in the Massangam health district in the West Region of Cameroon: Better

- understanding transmission dynamics to inform changes in programmatic interventions. *PLoS Negl Trop Dis*. 2018; 12(11):e0006904. <https://doi.org/10.1371/journal.pntd.0006904> PMID: 30427830
14. Hedtke SM, Kuesel AC, Crawford KE, Graves PM, Boussinesq M, Lau CL, et al. Genomic epidemiology in filarial nematodes: transforming the basis for elimination program decisions. *Front Genet*. 2020; 10:1282. Epub 2020/01/31. <https://doi.org/10.3389/fgene.2019.01282> PMID: 31998356
 15. Tekle AH, Zouré HG, Noma M, Boussinesq M, Coffeng LE, Stolk WA, Remme JH. Progress towards onchocerciasis elimination in the participating countries of the African Programme for Onchocerciasis Control: epidemiological evaluation results. *Infect Dis Poverty*. 2016; 5(1):66. <https://doi.org/10.1186/s40249-016-0160-7> PMID: 27349645
 16. Huestis DL, Dao A, Diallo M, Sanogo ZL, Samake D, Yaro AS, et al. Windborne long-distance migration of malaria mosquitoes in the Sahel. *Nature*. 2019; 574(7778):404–8. <https://doi.org/10.1038/s41586-019-1622-4> PMID: 31578527
 17. Midega JT, Smith DL, Olotu A, Mwangangi JM, Nzovu JG, Wambua J, et al. Wind direction and proximity to larval sites determines malaria risk in Kilifi District in Kenya. *Nat Commun*. 2012; 3:674. <https://doi.org/10.1038/ncomms1672> PMID: 22334077
 18. Atieli HE, Zhou G, Zhong D, Wang X, Lee MC, Yaro AS, et al. Wind-assisted high-altitude dispersal of mosquitoes and other insects in East Africa. *J Med Entomol*. 2023; 60(4):698–707. <https://doi.org/10.1093/jme/tjad033> PMID: 37094808
 19. Meribo K, Kebede B, Feleke SM, Mengistu B, Mulugeta A, Sileshi M, et al. Review of Ethiopian Onchocerciasis Elimination Programme. *Ethiop Med J*. 2017; 55(Suppl 1):55–63. PMID: 28878430
 20. Katarbarwa MN, Zarroug IMA, Negussu N, Aziz NM, Tadesse Z, Elmubark WA, et al. The Galabat-Metema cross-border onchocerciasis focus: The first coordinated interruption of onchocerciasis transmission in Africa. *PLoS Negl Trop Dis*. 2020; 14(2):e0007830. <https://doi.org/10.1371/journal.pntd.0007830> PMID: 32027648
 21. Garms R, Walsh JF, Davies JB. Studies on the reinvasion of the Onchocerciasis Control Programme in the Volta River Basin by *Simulium damnosum s.l.* with emphasis on the south-western areas. *Tropenmed Parasitol*. 1979; 30(3):345–62.
 22. Walsh JF, Davies JB, Garms R. Further studies on the reinvasion of the onchocerciasis control programme by *Simulium damnosum s.l.*: the effects of an extension of control activities into southern Ivory Coast during 1979. *Tropenmed Parasitol*. 1981; 32(4):269–73.
 23. Hadis M, Wilson MD, Cobblah M, Boakye DA. Cytotaxonomic description of *Simulium kaffaense*, a new member of the *S. damnosum* complex (Diptera: Simuliidae) from south-western Ethiopia. *Ann Trop Med Parasitol*. 2005; 99(3):267–91. <https://doi.org/10.1179/136485905X24210> PMID: 15829136
 24. Hadis M. **New members of *Simulium damnosum* species complex from Ethiopia**. Saarbrücken, Germany: VDM Verlag; 2010. xii+140 p.
 25. Hadis M, Wilson MD, Cobblah MA, Boakye DA. Cytotaxonomic description of *Simulium soderense* sp. nov. and a redescription of the 'Kulfo' form (Dunbar, 1969) of *Simulium damnosum s.l.* (Diptera: Simuliidae) as *Simulium kulfoense* sp. nov. in Ethiopia. *Int J Trop Insect Sci*. 2008; 28:3–11.
 26. Krüger A, Car M, Maegga BT. Descriptions of members of the *Simulium damnosum* complex (Diptera: Simuliidae) from southern Africa, Ethiopia and Tanzania. *Ann Trop Med Parasitol*. 2005; 99(3):293–306. <https://doi.org/10.1179/136485905X28009> PMID: 15829137
 27. Post RJ, Mustapha M, Krueger A. Taxonomy and inventory of the cytospecies and cytotypes of the *Simulium damnosum* complex (Diptera: Simuliidae) in relation to onchocerciasis. *Trop Med Int Health*. 2007; 12(11):1342–53. <https://doi.org/10.1111/j.1365-3156.2007.01921.x> PMID: 18045261
 28. Adler PH, Cheke RA, Post RJ. Evolution, epidemiology, and population genetics of black flies (Diptera: Simuliidae). *Infect Genet Evol*. 2010; 10:846–65. <https://doi.org/10.1016/j.meegid.2010.07.003> PMID: 20624485
 29. Cheke RA, Garms R. Indices of onchocerciasis transmission by different members of the *Simulium damnosum* complex conflict with the paradigm of forest and savanna parasite strains. *Acta Trop*. 2013; 125(1):43–52. <https://doi.org/10.1016/j.actatropica.2012.09.002> PMID: 22995985
 30. Adler PH. World blackflies (Diptera: Simuliidae): A comprehensive revision of the taxonomic and geographical inventory 2022. Available from: <https://biomia.sites.clemson.edu/pdfs/blackflyinventory.pdf>.
 31. Hebert PDN, Cywinska A, Ball SL, DeWaard JR. Biological identifications through DNA barcodes. *P Roy Soc B-Biol Sci*. 2003; 270(1512):313–21. <https://doi.org/10.1098/rspb.2002.2218> PMID: 12614582
 32. Ruiz-Arrondo I, Hernandez-Triana LM, Ignjatovic-Cupina A, Nikolova N, Garza-Hernandez JA, Rodriguez-Perez MA, et al. DNA barcoding of blackflies (Diptera: Simuliidae) as a tool for species identification and detection of hidden diversity in the eastern regions of Spain. *Parasit Vectors*. 2018; 11(1):463. <https://doi.org/10.1186/s13071-018-3046-7> PMID: 30103795

33. Pramual P, Adler PH. DNA barcoding of tropical black flies (Diptera: Simuliidae) of Thailand. *Mol Ecol Resour.* 2014; 14(2):262–71. <https://doi.org/10.1111/1755-0998.12174> PMID: 24112561
34. Krueger A, Hennings IC. Molecular phylogenetics of blackflies of the *Simulium damnosum* complex and cytophylogenetic implications. *Mol Phylogenet Evol.* 2006; 39(1):83–90. <https://doi.org/10.1016/j.ympev.2005.11.007> PMID: 16360322
35. Morales-Hojas R, Krueger A. The species delimitation problem in the *Simulium damnosum* complex, blackfly vectors of onchocerciasis. *Med Vet Entomol.* 2009; 23(3):257–68. <https://doi.org/10.1111/j.1365-2915.2009.00816.x> PMID: 19712156
36. Hohenlohe PA, Catchen J, Cresko WA. Population genomic analysis of model and nonmodel organisms using sequenced RAD tags. *Methods Mol Biol.* 2012; 888:235–60. https://doi.org/10.1007/978-1-61779-870-2_14 PMID: 22665285
37. Ethiopian Mapping Authority. National Atlas of Ethiopia. Addis Abeba: Ethiopian Mapping Authority; 1988.
38. Conway D. The climate and hydrology of the Upper Blue Nile River. *Geogr J.* 2006; 166:49–62.
39. Nicholson SE. A review of climate dynamics and climate variability in eastern Africa. In: Johnson TC, Odada AO, editors. *Limnology, Climatology and Paleoclimatology of the East African Lakes*. Boca Raton, Florida: CRC Press Taylor & Francis Group; 1996. p. 25–56.
40. Friis I, Demissew S, van Breugel P. Atlas of the potential vegetation of Ethiopia. Addis Ababa: Addis Ababa University Press & Shama Books; 2010. 307 p.
41. Davies JB, Crosskey RW, World Health Organization Division of Control of Tropical Diseases. *Simulium—vectors of onchocerciasis*. WHO/VBC/91.992 ed. Geneva: World Health Organization; 1991. 115 p.
42. Vajime CG, Dunbar RW. Chromosomal identification of eight species of the subgenus *Edwardsellum* near and including *Simulium* (*Edwardsellum*) *damnosum* Theobald (Diptera: Simuliidae). *Tropenmed Parasitol.* 1975; 26(1):111–38.
43. Boakye DA. A pictorial guide to the chromosomal identification of members of the *Simulium damnosum* Theobald complex in west Africa with particular reference to the Onchocerciasis Control Programme Area. *Trop Med Parasitol.* 1993; 44(3):223–44.
44. Krueger A. Guide to blackflies of the *Simulium damnosum* complex in eastern and southern Africa. *Med Vet Entomol.* 2006; 20:60–75.
45. Mehbrahtu Y, Abebe M, Mekuria Y. Blackflies (Diptera: Simuliidae) of Ethiopia: checklist and distribution. *SINET Ethiopian J Sci.* 1980; 3:1–20.
46. Freeman P, de Meillon B. Simuliidae of the Ethiopian region. London: British Museum (Natural History); 1953. vii+224 p.
47. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol.* 1994; 3(5):294–9. PMID: 7881515
48. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007; 23(21):2947–8. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
49. Leigh JW, Bryant D. POPART: full-feature software for haplotype network construction. *Methods Ecol Evol.* 2015; 6(9):1110–6. <https://doi.org/10.1111/2041-210X.12410>
50. Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol.* 2017; 34(12):3299–302. <https://doi.org/10.1093/molbev/msx248> PMID: 29029172
51. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. *Mol Ecol.* 2013; 22(11):3124–40. <https://doi.org/10.1111/mec.12354> PMID: 23701397
52. Rochette NC, Catchen JM. Deriving genotypes from RAD-seq short-read data using Stacks. *Nat Protoc.* 2017; 12(12):2640–59. <https://doi.org/10.1038/nprot.2017.123> PMID: 29189774
53. Paris JR, Stevens JR, Catchen JM. Lost in parameter space: a road map for STACKS. *Methods Ecol Evol.* 2017; 8(10):1360–73. <https://doi.org/10.1111/2041-210X.12775>
54. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics.* 2008; 24(11):1403–5. <https://doi.org/10.1093/bioinformatics/btn129> PMID: 18397895
55. Jombart T, Ahmed I. adegenet 1.3–1: new tools for the analysis of genome-wide SNP data. *Bioinformatics.* 2011; 27(21):3070–1. <https://doi.org/10.1093/bioinformatics/btr521> PMID: 21926124
56. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics.* 2011; 27(15):2156–8. <https://doi.org/10.1093/bioinformatics/btr330> PMID: 21653522

57. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 2010; 11:94. <https://doi.org/10.1186/1471-2156-11-94> PMID: 20950446
58. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000; 155(2):945–59. <https://doi.org/10.1093/genetics/155.2.945> PMID: 10835412
59. Higazi TB, Boakye DA, Wilson MD, Mahmoud BM, Baraka OZ, Mukhtar MM, Unnasch TR. Cytotaxonomic and molecular analysis of *Simulium* (*Edwardsellum*) *damnosum sensu lato* (Diptera: Simuliidae) from Abu Hamed, Sudan. *J Med Entomol.* 2000; 37(4):547–53. <https://doi.org/10.1603/0022-2585-37.4.547> PMID: 10916295
60. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990; 215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
61. Hernández-Triana LM, Prosser SW, Rodríguez-Perez MA, Chaverri LG, Hebert PDN, Gregory TR. Recovery of DNA barcodes from blackfly museum specimens (Diptera: Simuliidae) using primer sets that target a variety of sequence lengths. *Mol Ecol Resour.* 2014; 14(3):508–18. <https://doi.org/10.1111/1755-0998.12208> PMID: 24299419
62. Post RJ, Cheke RA, Boakye DA, Wilson MD, Osei-Atweneboana MY, Tetteh-Kumah A, et al. Stability and change in the distribution of cytospecies of the *Simulium damnosum* complex (Diptera: Simuliidae) in southern Ghana from 1971 to 2011. *Parasit Vectors.* 2013; 6:205. <https://doi.org/10.1186/1756-3305-6-205> PMID: 23849451
63. Cheng C, Tan JC, Hahn MW, Besansky NJ. Systems genetic analysis of inversion polymorphisms in the malaria mosquito *Anopheles gambiae*. *Proc Natl Acad Sci USA.* 2018; 115(30):E7005–E14. <https://doi.org/10.1073/pnas.1806760115> PMID: 29987007
64. Brooke BD, Hunt RH, Chandre F, Carnevale P, Coetzee M. Stable chromosomal inversion polymorphisms and insecticide resistance in the malaria vector mosquito *Anopheles gambiae* (Diptera: Culicidae). *J Med Entomol.* 2002; 39(4):568–73. <https://doi.org/10.1603/0022-2585-39.4.568> PMID: 12144286
65. Coluzzi M, Sabatini A, Petrarca V, Di Deco MA. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans R Soc Trop Med Hyg.* 1979; 73(5):483–97. [https://doi.org/10.1016/0035-9203\(79\)90036-1](https://doi.org/10.1016/0035-9203(79)90036-1) PMID: 394408
66. Heath TA, Hedtke SM, Hillis DM. Taxon sampling and the accuracy of phylogenetic analysis. *J Syst Evol.* 2008; 46(3):239–57.