A geo-coded inventory of anophelines in the Afrotropical Region south of the Sahara: 1898-2016 [version 1; referees: 3 approved]

David Kyalo¹, Punam Amratia¹, Clara W. Mundia¹, Charles M. Mbogo¹, Maureen Coetzee²,³, Robert W. Snow¹,⁴

¹Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya
²Centre for Emerging, Zoonotic & Parasitic Diseases, National Institute for Communicable Diseases, Johannesburg, South Africa
³Wits Research Institute for Malaria, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
⁴Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

Abstract

Background: Understanding the distribution of anopheline vectors of malaria is an important prelude to the design of national malaria control and elimination programmes. A single, geo-coded continental inventory of anophelines using all available published and unpublished data has not been undertaken since the 1960s.

Methods: We have searched African, European and World Health Organization archives to identify unpublished reports on anopheline surveys in 48 sub-Saharan Africa countries. This search was supplemented by identification of reports that formed part of post-graduate theses, conference abstracts, regional insecticide resistance databases and more traditional bibliographic searches of peer-reviewed literature. Finally, a check was made against two recent repositories of dominant malaria vector species locations (circa 2,500). Each report was used to extract information on the survey dates, circa village locations (geo-coded to provide a longitude and latitude), sampling methods, species identification methods and all anopheline species found present during the survey. Survey records were collapsed to a single site over time.

Results: The search strategy took years and resulted in 13,331 unique, geo-coded survey locations of anopheline vector occurrence between 1898 and 2016. A total of 12,204 (92%) sites reported the presence of 10 dominant vector species/sibling species; 4,473 (37%) of these sites were sampled since 2005. 4,442 (33%) sites reported at least one of 13 possible secondary vector species; 1,107 (25%) of these sites were sampled since 2005. Distributions of dominant and secondary vectors conform to previous descriptions of the ecological ranges of these vectors.

Conclusion: We have assembled the largest ever geo-coded database of anophelines in Africa, representing a legacy dataset for future updating and identification of knowledge gaps at national levels. The geo-coded database is available on Harvard Dataverse as a reference source for African national malaria control programmes planning their future control and elimination strategies.
Introduction

In 1939, Botha de Meillon stated that “Malaria in South Africa, as elsewhere in the world, is an entomological disease. Its epidemiology only becomes clear when knowledge of its entomology has been elucidated” (De Meillon & Gear, 1939). This sentiment has been variously accepted and neglected throughout the history of malaria control in Africa.

The first global inventory of the Genus Anopheles (Diptera: Culicidae) was published in 1901 and reproduced in 1903 and 1910 (Theobald, 1901). Sir Rickard Christophers updated this inventory in 1924 “as a necessary preliminary to studying the geographical distribution of species, has been published in the belief that, as a handy means of reference to known species with their correct names, it would be useful to medical men and others” (Christophers, 1924). In 1929, an assembly of reported locations of vectors from published and unpublished sources from the beginning of the 1900s was developed and presented as lists per country, which included location names and for the first time was shown on regional maps (Kumm, 1929). This was updated for the Africa region in 1938, providing bibliographic sources, locations, taxonomic keys for adult mosquito stages and more details on bionomics by the Natural History Museum, London (Evans, 1938), and repeated for larval stages in 1952 (Hopkins, 1952). During the Second World War, the US Sanitary Department developed a separate inventory (Ross & Roberts, 1943).

The most definitive catalogue of recorded anopheline species for the Afrotropical region was published in 1968 by Mick Gillies and Botha de Meillon, updating the earlier work (De Meillon, 1947) and capturing a wealth of published and unpublished observations from across the continent, linked to spatial grids of their distributions (Gillies & de Meillon, 1968). This geo-referenced catalogue was accompanied by comprehensive descriptions of the morphology, notes on the species role in malaria transmission, and bionomics (Gillies & de Meillon, 1968). Updated inventories of anophele distributions were published in 1972 (Fritz, 1972) and 1987 (Gillies & Coetzee, 1987). More localised inventories were published on the distribution of the Anopheles gambiae complex in southern Africa (Coetzee et al., 1993) and on insecticide resistance in southern African vectors (Coetzee et al. 1999).

During the Global Malaria Eradication Programme (GMEP) era, from 1955-69, descriptions of the anopheline vectors, shown as sub-national distributions, were regarded as important preludes to pre-eradication and attack phases of eradication. The surveillance of malaria vectors in sub-Saharan Africa (SSA) was limited to those countries where eradication was pursued (e.g. the Horn of Africa, southern Africa and Africa’s offshore island states). When the emphasis on malaria control shifted to presumptive treatment of fevers through primary health care, entomological reconnaissance for malaria control became a forgotten public health science during the 1970s and 1980s across much of SSA (Mnzava et al., 2014).

In 1996, the Mapping Malaria Risk in Africa (MARA) collaboration was launched (Coetzee et al., 2000; Lee Sueur et al., 1997; Snow et al., 1996; http://www.mara-database.org) to assemble, geo-code and map malaria parasite and vector surveys undertaken across Africa, south of the Sahara. The initiative focused only on documenting records of the sibling species of the An. gambiae complex and An. funestus s.l. malaria vectors from reports of surveys undertaken between 1920 and 2004. This was a milestone collaboration, managed by scientists across the Africa region, and started a renaissance in the assembly of empirical malaria information as geo-coded inventories. Other initiatives to create entomological and insecticide resistance data repositories followed, including the African Network for Vector Resistance (ANVR) (ANVR, 2005), the Malaria Atlas Project (MAP) (Sinka et al., 2010; http://www.map.ox.ac.uk; Wiebe et al., 2017), the Disease Vectors Database (Moffett et al., 2009), MosquitoMap [http://www.mosquitomap.org], Walter Reed Bioinformatics Unit’s systematic catalogue of Culicidae [http://www.vectormap.org; Foley et al., 2008; Foley et al., 2010], IRBase [http://www.irmapper.com; Dialynas et al., 2009; Knox et al., 2014], VectorBase [http://www.vectorbase.org], which focuses on descriptions of bionomics and gene libraries of many disease vectors, and the Vector-Borne Disease Network (VecNet) [http://www.vecnet.org].

The more recent on-line, global or regional, vector species location databases do not represent the entire historical reference material for any given country; they often focus only on peer-reviewed published sources without reference to unpublished reports from national control agencies and research partners, and they do not always cover the entire range of potential secondary vectors reported in countries. Here we present an assembly of geo-coded anopheline species data abstracted from survey reports published since 1900 across SSA and its offshore islands that have yet to eliminate malaria.

Methods

Data search

Methods used by us to identify sources of information have been opportunistic, cascaded approaches and began with bibliographies provided in earlier regional inventories published between 1929 – 1987 (Evans, 1938; Fritz, 1972; Gillies & de Meillon, 1968; Gillies & Coetzee, 1987; Hopkins, 1952; Kumm, 1929). Manual searches were undertaken at the archives and libraries of ex-colonial tropical medicine institutes to locate unpublished reports from malariologists working in Africa before countries achieved independence. We searched the archives of the World Health Organization (WHO) in Geneva and regional archives in Brazzaville and Cairo, and identified consultant’s trip reports and quarterly reports from malariologists working on behalf of the WHO from the 1950s through to the 1970s. National archives of the Ministry of Health offices were visited in Burkina Faso, Ghana, Kenya, Senegal, South Africa, Sudan, Tanzania and Uganda. Annual medical and sanitation department reports from 1919, produced mostly by the Anglophone pre-independent colonial
governments, were available at the library in the National Public Health Laboratories of the Ministry of Health, Nairobi, Kenya.

Post-graduate theses undertaken with entomological components were sourced from local university libraries in the faculties of zoology, medicine or related biological sciences in Kenya, Mali, Mozambique, Senegal, Sudan, Tanzania, Belgium, France and the UK. National and international malaria congresses and conference proceedings were also reviewed for abstracts that contained information on species identifications at specific localities. These sources of possible information are incomplete, and it is therefore to be expected that substantial data are available across university departments in Africa, not captured by us. Entomologists working across Africa, within research institutes or as part of National Malaria Control programmes, were also contacted directly for any unpublished survey reports, notably as part of more recent malaria vector surveillance and insecticide monitoring since the re-launch of indoor residual house-spraying in Africa.

Using more traditional methods of online electronic peer-review, published reports were searched using free text keywords “Anopheles” and “country-name” in March 2010 and repeated at least once per year before the search ended on 31st March 2017. Bibliographic on-line resources included PubMed, Google Scholar and the World Health Organization Library Database. Regional journals, including a large number of national medical, public health and parasitological journals, were not identified readily from the above sources, but titles and abstracts were available on African Journals Online (http://www.ajol.info). All publications were cross-referenced using the bibliographies for additional sources that may have been missed or that may correspond to unpublished or ‘grey’ literature, not controlled by commercial publishers. Finally, our database was compared with the 2,535 site location data reported by MARA/ARMA (Coetzee et al., 2000) and the 2,582 site locations recorded by MAP (Wiebe et al., 2017) to ensure these were captured as part of our search strategy.

Data abstraction

Data organisation: The basic principle of the database was to develop a site-specific inventory. As such, multiple reports from the same site were collapsed to a single entry, with all citations combined to that site. Invariably, multiple authors of published material report on the same surveys or aspects of entomological work from the same site across a period of several years, and in these instances the first and last survey years were retained. Some bibliographic sources cite previous vector descriptions from unpublished sources by other entomologists; these have been documented as op cit to the original author against the reference source. For older reports it was not always possible to define the year of sampling, and where dates of sampling were not provided we have presumed that it happened within the preceding five years of the publication date. Separate reports on the same location vary in the stages of vector sampled and the precision methods used to distinguish species and sibling species of complexes. In such cases, all sampling, vector stage and species identification methods were recorded across surveys, often increasing with time in species identification precision.

Sampling methods: For each record, we documented whether adults or larvae were sampled and a summary of methods used to sample vectors, for example animal bait catches, bed net traps, CDC light traps, human landing catches, human bait catches (where someone was protected by a double net), indoor resting searches, pyrethrum spray catches, exit traps, outdoor bait traps, Ifakara tent traps, Monks Wood traps, larval searches or larvae reared to adults. If there were no details available then “unknown” was recorded, which was often the case from national reviews of previously unpublished data. Insecticide resistance data often sample larvae at sites that were then reared to adults. In older literature, notably assemblies of unpublished data, it was not always possible to define whether adults or larvae were sampled and here we have defaulted to assuming both.

Species identification: Throughout the data assembly we have only recorded the reported presence of a species where this was described during a survey. Even if the report stated absence of sibling species/species we have not recorded absence. True absence is not possible to define within an ecological niche of vector species, as it depends critically on the intensity, duration and repeat sampling at any given site.

Dominant and potential secondary vectors

A perennial problem with assemblies of vector inventories over time are the ambiguities in taxonomy and nomenclature. These improve with time as part of detailed mosquito systematics, improvements in morphological keys and genetic techniques. Methods used to identify species were recorded for all surveys at each location as per morphological keys, cross-mating, Polymerase Chain Reaction (PCR), chromosome banding sequences, DNA probes or enzyme electrophoresis.

We have regarded as primary, dominant vectors within their ecological range as An. gambiae s.l. (An. gambiae s.s, An. coluzzii, An. arabiensis, An. melas, An. merus, An. bwambae); An. funestus s.s.; An. nili s.l., (An. nili s.s., An. carnevalei, An. ovengensis), An. moucheti s.l. (An. moucheti moucheti, An. moucheti nigeriensis); and An. mascarensis as a primary vector in Madagascar, Comoros and Mayotte (Fontenille & Campbell, 1992; Fontenille et al., 2003; Marrama et al., 1999).

For the Gambiæ complex, as much detail as possible from the reports was extracted and updated with new more specific chromosomal or cross-mating information from subsequent reports at the same location: An. gambiae s.l. (if only complex mentioned or An. costalis in very early reports), An. gambiae s.s. (Species A) when possible to differentiate from An. arabiensis (Species B) and saltwater breeding species (An. melas and An. merus), An. gambiae...
S form (when indicated as Savannah or Bamako or S forms) and An. coluzzii (when indicated as M form or Mopti form). The zoophilic An. quadriannulatus A and An. quadriannulatus B were described as sibling-species of the An. gambiae complex (previously species C) in the early 1980s, but not regarded as vectors of malaria within their geographic ranges of southern Africa and Ethiopia (Coetze, 1987; Coluzzi, 1984). An. quadriannulatus B from Ethiopia was later renamed An. amharicus Hunt, Wilkerson & Coetze sp. n. (Coetze et al., 2013; Hunt et al., 1998) while the name An. quadriannulatus was retained for the southern African species. An. quadriannulatus is recorded under other species. An. bwambae is a member of the Gambiæa complex involved in malaria transmission only within a very restricted geographical range in Uganda (Davidson & Hunt, 1973; White, 1985).

The An. funestus group has taxonomic complexity similar to that of the Gambiæa complex. The taxonomic classification and systematic of the An. funestus group has resulted in a reclassification of the group with An. funestus s.s., An. arunii, An. parentis, An. confusus, An. vaneedeni and An. funestus-like (described in Malawi (Spillings et al., 2009)) being grouped together as members of the “An. funestus subgroup”; An. rivularum, An. rivularum-like, An. brucei and An. fuscevenosus form their own subgroup; and An. leesoni has been grouped with the Asian An. minimus subgroup (Choi et al., 2012; Coetze & Koekemoer, 2013; Harbach, 2004; Spillings et al., 2009). Among the Funestus group, An. funestus s.s. is a significant vector in the transmission of malaria (Coetze & Koekemoer, 2013); An. rivularum has been recently implicated in transmission in Tanzania and might contribute as a secondary vector to transmission elsewhere (Kawada et al., 2012; Wilkes et al., 1996); and An. vaneedeni was implicated recently in residual transmission in South Africa (Burke et al., 2017). We have documented only the presence of sibling species within groups where these have been uniquely defined in the reports.

An. moucheti is an important vector in equatorial forests in Central and West Africa (Antonio-Nkondjio et al., 2002; Antonio-Nkondjio et al., 2008; Mattingly, 1949; Manga et al., 1995). This vector was originally divided into three morphological forms An. moucheti moucheti (type form), An. moucheti hervoetsi and An. moucheti nigeriensis (Brunhes et al., 1998). However, recent classifications recognize An. moucheti and An. hervoetsi as formal species, while An. moucheti nigeriensis is considered as a morphological subspecies of An. moucheti (Antonio-Nkondjio et al., 2002; Antonio-Nkondjio et al., 2008).

The An. nili complex currently comprises four formal species, An. nili s.s., An. somalicus, An. carnevalei and An. ovengensis (Awono-Ambene et al., 2004; Kengne et al., 2003). An. somalicus, has never been incriminated in human malaria transmission, however the three other members are highly anthropophilic and are important vectors of malaria within most of their geographical range (Carnevale & Zoulani, 1975; Mouchet et al., 2008).

The definition of secondary vectors is complex and often site/time specific (Afrane et al., 2016; De Meillon, 1951; Gillies & de Meillon, 1968; Holstein, 1951). Most non-dominant, potential vectors are exophilic (outdoor resting), exophagic (outdoor biting) and zoophilic (preference, but not exclusive, for non-human hosts). They only feature as possible vectors of malaria where they are abundant and have life-expectancies long enough to support transmission. The single most important characteristic is whether the vectors have been detected harbouring Plasmodium falciparum sporozoites in their salivary glands, suggesting they have acquired infections from human hosts and have survived long enough for a complete sporogonic cycle. We have proposed a revised list of potential secondary vectors based on historical and contemporary reports of detection of sporozoite infected adult sampled species (Table 1). Despite occasional reference to sporozoite positivity in several locations before 1950 (De Meillon, 1950), we have not treated An. austeni, An. brunnipes, An. christyi, An. hargreavesi, An. maculipalpis, An. theileri, An. pretoriensis and An. rhodesiensis as secondary vectors; however they are documented within the database alongside other anopheline species detected at sites across Africa. In addition, we have not included An. dhali as a secondary vector, since reports of contributions to transmission in the arid areas of the Horn of Africa are extremely localized (Lega et al., 1937; Rishikesh, 1961). Records under the “other anophelines” field were listed per the taxonomic older synonym or variety names, but subsequently corrected to nomenclature used today.

Survey geo-coding

Each survey location was attributed a decimal longitude and latitude. Where household level data were reported these were re-aggregated to village levels as a single entry. Where only mapped distributions of species locations were provided, without location names, these were overlaid in digital formats with Google Earth and location/site names extracted. Mosquito systems, however, often reported longitude and latitude of survey locations and detailed descriptions of survey sites, and the more recent use of Global Positioning Systems (GPS) during survey work has increased reporting of longitudes and latitudes. All reported coordinates were re-checked in Google Earth. For georeferencing survey locations, where longitudes and latitudes were not available in the original survey reports, we have used a variety of digital resources, amongst which the most useful were Microsoft Encarta Encyclopedia, Google Earth, GeoNames and OpenStreetData. Other sources of digital place name archives used are shown in Table 2 and used in combination with increasingly available national statistics bureau, ministry of health or ministry of education geo-coded place name databases.

Results

We identified 2,221 published and unpublished reports. The earliest survey report was of An. costalis (now known as An. gambiae s.l.) collected in Freetown and its environs in Sierra Leone in expeditions carried out by Christophers and Stephens...
Table 1. List of potential secondary vectors described as being sporozoite positive during field surveys in 20 countries since 1929.

<table>
<thead>
<tr>
<th>Secondary malaria vectors</th>
<th>Location/country where sporozoite positive samples found</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anopheles pharoensis</td>
<td>Benin, Burkina Faso, Chad, Cameroon, Ethiopia, Ghana, Kenya, Mali, Nigeria, Senegal, Tanzania, Uganda</td>
<td>Animut et al., 2013; Antonio-Nkondjio et al., 2006; Carrara et al., 1990; Cavalie &amp; Mouchet, 1961; De Meillon, 1950; Dery et al., 2010; Dia et al., 2008; Gari et al., 2016; Gibbins, 1932; Gillies, 1964; Hamon &amp; Mouchet, 1956; Kerah-Hinzoumbé et al., 2009; Kibret et al., 2014; Miles et al., 1983; Mukama &amp; Mwangi, 1989; Tchouassi et al., 2012; Wanji et al., 2003</td>
</tr>
<tr>
<td>Anopheles squamosus</td>
<td>Kenya, Mali, Tanzania, Zambia</td>
<td>De Meillon, 1947; Fornadel et al., 2011; Gillies, 1964; Hamon et al., 1956; Lobo et al., 2015; St Laurent et al., 2016</td>
</tr>
<tr>
<td>Anopheles wellcomei</td>
<td>Cameroon, Senegal</td>
<td>Antonio-Nkondjio et al., 2006; Dia et al., 2008; Wanji et al., 2003</td>
</tr>
<tr>
<td>Anopheles rufipes</td>
<td>Burkina Faso, Cameroon, Gambia, Ghana, Kenya, Mali, Nigeria, Senegal, Togo</td>
<td>Da et al., 2013; Dery et al., 2010; Dia et al., 2008; Gelfand, 1947; Hamon &amp; Rickenbach, 1955; Hamon et al., 1956; Holstein, 1960; St Laurent et al., 2016; Tabue et al., 2017; Tchouassi et al., 2012</td>
</tr>
<tr>
<td>Anopheles hancocki</td>
<td>Cameroon, Uganda</td>
<td>Antonio-Nkondjio et al., 2006; De Meillon, 1950; Gibbins, 1932; Wanji et al., 2003</td>
</tr>
<tr>
<td>Anopheles marshalli</td>
<td>Cameroon, Uganda</td>
<td>Antonio-Nkondjio et al., 2006; De Meillon, 1950; Gibbins, 1932; Wanji et al., 2003</td>
</tr>
<tr>
<td>Funestus subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anopheles leesoni</td>
<td>Kenya, South Africa, Tanzania</td>
<td>Conn, 2016; Garros et al., 2004; St Laurent et al., 2016</td>
</tr>
<tr>
<td>Anopheles parensis</td>
<td>Kenya, South Africa</td>
<td>Garros et al., 2004; Kamau et al., 2003</td>
</tr>
<tr>
<td>Anopheles vaneedeni</td>
<td>South Africa</td>
<td>Burke et al., 2017; Garros et al., 2004; Green &amp; Hunt, 1980</td>
</tr>
<tr>
<td>Anopheles rivulorum</td>
<td>Kenya, Tanzania, Zambia</td>
<td>Gillies, 1964; Lobo et al., 2015; Wilkes et al., 1996</td>
</tr>
<tr>
<td>Coustani group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anopheles coustani</td>
<td>Benin, Burkina Faso, Cameroon, DRC, Ethiopia, Ghana, Kenya, Madagascar, Mozambique, Tanzania, Zambia</td>
<td>Antonio-Nkondjio et al., 2006; Fornadel et al., 2011; Gillies, 1964; Govoetchan et al., 2014; Hamon &amp; Mouchet, 1961; Kibret et al., 2014; Lobo et al., 2015; Mendis et al., 2000; Mwangangi et al., 2013; Nepomichene et al., 2015; Sovi et al., 2013; Tchouassi et al., 2012; Vincke &amp; Jadin, 1946</td>
</tr>
<tr>
<td>Anopheles paludis</td>
<td>Cameroon, DRC</td>
<td>Antonio-Nkondjio et al., 2006; De Meillon, 1950; Karch &amp; Mouchet, 1992; Lips, 1961; Wanji et al., 2003</td>
</tr>
<tr>
<td>Anopheles ziemanni</td>
<td>Benin, Burkina Faso, Cameroon, Chad, Côte d'Ivoire, Ethiopia, Kenya, Tanzania</td>
<td>Animut et al., 2013; Gari et al., 2016; Githeko et al., 1990; Gillies &amp; de Meillon, 1968; Govoetchan et al., 2014; Hamon &amp; Mouchet, 1961; Kamau et al., 2006; Kerah-Hinzoumbé et al., 2009; Lobo et al., 2015; Robert et al., 1992; Sovi et al., 2013; Vincke &amp; Jadin, 1946; Wilkes et al., 1996</td>
</tr>
</tbody>
</table>

Notes:
1. An. squamosus in Madagascar was variously recorded as An. squamosus, An. squamosus var. cydippis [Nepomichene et al., 2015] and later unique reports of An. cydippis. The taxonomic keys reported in individual reports are vague and we have recorded An. squamosus presence where used in the species report, but as An. cydippis under other species where specified without reference to An. squamosus.
2. An. marshalli is a complex, although rarely differentiated during field surveys except when specified as An. keniensis. In Madagascar, early surveys often misclassified An. macarensis as An. marshalli.
3. The taxonomic name of An. mauritianus was sunk as a synonym of An. coustani in 1932. The varieties of te negerosus and ziemanni were raised to species status by Gillies & de Meillon (1968), while An. paludis was a recognised species by Evans (1938). Where early reports document An. mauritianus and the varieties, the current species names have been recorded here.
in 1898 and 1899 (Christophers & Stephens, 1900; Stephens & Christophers, 1900). Important sources of information have been national inventories developed by entomologists working in Africa, the earliest were those developed before the launch of the GMEP in Cape Verde, Democratic Republic of Congo (DRC), Eritrea, Ethiopia, Gabon, Kenya, Liberia, Mozambique, Nigeria, Rwanda and Burundi, South Africa, Sudan, Zambia and Zanzibar. During the preparation for the GMEP, countries often undertook national malaria reconnaissance surveys that included detailed descriptions of local ecologies and vector habitats, host infection studies and the presence of all anophelines by species and stages. After the Second World War, work of individual malariologists and entomologists began to be assembled into national inventories of anopheline distributions and in several instances mapped to provide a visual display of the recorded species distributions. These early national inventories, for the most part, were based on unpublished survey data from ministry of health reports spanning decades of entomological investigation. Some countries have recognised the importance of updating inventories of anopheline distributions, notably where regional compendia are incomplete for country purposes. Examples of these more contemporary national inventories were identified for Cote d’Ivoire, DRC, Kenya, Madagascar, Mali, Mauritania, Niger, Nigeria, Senegal, Somalia and Tanzania.

The final database contained 13,465 unique survey locations where anophelines were sampled between 1898 and 2016. We were able to geo-locate 13,331 (99%) using a variety of on-line digital gazetteers (Figure 1; Table 3). Sampled locations that have included surveys since 2005 cover 4,494 survey sites and highlight the paucity of available contemporary information on malaria vector species distributions in Congo, Central African Republic, Chad, Eritrea, Namibia, Sao Tome and Principe, South Sudan and Togo (Figure 1; Table 3).

### Dominant vectors

Within the Gambiae complex, *An. arabiensis* has a more extensive range (Figure 2), including the more arid areas south of the Sahara and the Horn of Africa. Conversely, *An. gambiae s.s.* and *An. coluzzii* have been described more frequently in West Africa compared to Central and East Africa. The saltwater breading *An. melas* has a restricted range along the West African coast; however the ranges of *An. merus* in East and Southern Africa show an extended, inland geographic range (Figure 2) (Bushrod, 1981;
Coetzee et al., 1993; Cuamba & Mendis, 2009; Kigadye et al., 2010). An. funestus s.s., the dominant malaria vector within its group, has only recently been described uniquely through advances in molecular techniques, its distribution being within the range of more ubiquitous data on An. funestus s.l. (Figure 3). An. moucheti s.l. is located principally in central Africa, but described further west across Nigeria and recorded as far west as Sierra Leone and Guinea (Figure 4). An. nili s.l. has a more extended range (Figure 4) covering areas occupied by other dominant vectors. While An. mascarensis, also a dominant vector within its range, is constrained to Madagascar, Comoros and Mayotte (Figure 4).

Secondary vectors
Fewer site locations reported possible secondary vectors compared to reporting of dominant vectors. We identified survey reports of potential secondary vectors (Table 1) at 4,442 site locations sampled since 1901, and only 1,106 site locations where the sampling date included 2005–2016. Among the possible secondary vectors of the An. funestus group both An. rivulorum and An. leesoni showed an extensive range across SSA (Figure 5A), there were fewer reports of An. parensis and An. vaneedeni presence; however these were largely at sites located in Eastern and Southern Africa (Figure 5A). Reports of the presence of members of the An. coustani group were ubiquitous and extensive in their range (Figure 5A), while sibling species of this group, An. ziemanni and An. paludis, were more frequently reported within a constrained range across the central belt of Africa (Figure 5A).

An. rufipes, An. pharoensis, An. squamosus and An. marshallii all have a cosmopolitan range across Africa (Figure 5B). An. hancocki and An. wellcomei have been largely reported from the central African belt (Figure 5B).
Table 3. Numbers of geo-coded site locations documenting dominant vector species (DVS) and potential secondary vector species (SVS) by country and lists of all anophelines identified during entomological surveys in the current database and supplemented by anopheline descriptions provided in (Gillies & de Meillon, 1968; White, 1980).

<table>
<thead>
<tr>
<th>Country</th>
<th>DVS survey locations (date range) [No. locations since 2005]</th>
<th>SVS survey locations 1900–2016 [No. locations since 2005]</th>
<th>Lists of anophelines described</th>
</tr>
</thead>
</table>

Wellcome Open Research 2017, 2:57 Last updated: 11 SEP 2017
<table>
<thead>
<tr>
<th>Country</th>
<th>DVS survey locations (date range) [No. locations since 2005]</th>
<th>SVS survey locations 1900–2016 [No. locations since 2005]</th>
<th>Lists of anophelines described</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Lists of anophelines described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>DVS survey locations (date range) [No. locations since 2005]</td>
<td>SVS survey locations 1900–2016 [No. locations since 2005]</td>
<td>Lists of anophelines described</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------------------------</td>
<td>-----------------------------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Country</td>
<td>DVS survey locations (date range) [No. locations since 2005]</td>
<td>SVS survey locations 1900–2016 [No. locations since 2005]</td>
<td>Lists of anophelines described</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Country</td>
<td>DVS survey locations (date range) [No. locations since 2005]</td>
<td>SVS survey locations 1900–2016 [No. locations since 2005]</td>
<td>Lists of anophelines described</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>--------------------------------</td>
</tr>
</tbody>
</table>
Figure 2. Distribution of sampled sites reporting presence of *Anopheles gambiae* complex (11,494), and sibling species, regarded as dominant vectors of malaria within their ecological range. *An. gambiae* s.s (3,988 locations) is shown separately to allow for reporting that did not distinguish, for example, between *An. coluzzii/M* forms and *An. gambiae S* forms. *An. gambiae* s.s (S form) 1,574 locations, *An. coluzzii* 1,331 locations, *An. arabiensis* = 3,635 locations, *An. melas* 538 locations, *An. merus* 253 locations and *An. bwambae* 26 locations.

Figure 3. Distribution of 5,052 sampled locations reporting the presence of *Anopheles funestus* s.l. (green) or where survey reports specified *An. funestus* s.s. (Blue) (n=727).
Figure 4. Distribution of sampled locations reporting the presence of dominant vectors *Anopheles nili* s.l., *An. moucheti* and *An. mascarensis*. (A) *An. nili* s.l. (*n* = 822 locations); (B) *An. moucheti* s.l. (*n* = 499 locations), the two sites in Namibia were noted as unusual for this far south (De Meillon, 1951) and (C) *An. mascarensis* (*n* = 483 locations).
Figure 5. Distribution of sampled locations of potential secondary malaria vectors (Table 1). (A) Anopheles funestus group: An. rivulorum (n = 244), An. leesoni (n = 187), An. parensis (n = 41) and An. vaneedeni (n = 34); the An. coustani group (n = 2,689) and where specified sibling species An. ziemanni (n = 662) and An. paludis (n = 375). (B) An. squamosus (n= 1,294), An. pharoensis (n = 1,889), An. rufipes (n = 998), An. marshallii (n = 442), An. wellcomei (n = 272) and An. hancocki (n = 297).
Discussion
The geo-coded inventory of anopheline species in SSA covers over 13,000 locations and represents the most spatially comprehensive description of dominant and secondary malaria vectors in Africa to-date. Of the sites sampled since the turn of the last century, 4,494 (33.7%) have been sampled at least once since 2005 (Figure 1). The use of historical archive and unpublished material has been fundamental in expanding our understanding of the spatial ranges of primary and secondary malaria vectors in Africa. Our assembly process highlights the significance of ensuring unpublished materials are sourced at country levels. A wealth of unpublished information existed 50 years ago and this is equally true today. We have not visited every country in Africa nor visited every national malaria control programme and university archive on the continent. The data on anopheline vectors could be improved substantially at country levels.

We have not attempted to model the likely presence of dominant or secondary vectors where currently no data exist. This has been attempted by others (Foley et al., 2010; Levine et al., 2004; Lunde et al., 2013; Moffett et al., 2007; Sinka et al., 2010; Wiebe et al., 2017) and used to project the likely ranges of dominant vectors under different climate (Drake & Beier, 2014) and intervention (Sinka et al., 2016) scenarios. It is our belief that modelling cannot replace actual field survey data. In addition, we cannot reliably document the temporal changes in vector species compositions. The precision in species identification has changed with time, making it hard to compare pre-1990 dominant species complex composition with current compositions, and relative abundance is often not reported. Such time-series analysis is better described at small, site specific locations of repeat, standardised sampling rather than from large summary database repositories as presented here.

Recent modelling exercises to predict niches, using climate and ecological predictors of vector species presence (Foley et al., 2010; Levine et al., 2004; Lunde et al., 2013; Moffett et al., 2007; Sinka et al., 2010; Wiebe et al., 2017), largely confirm the natural geographic ranges described during early cartographies of species distributions in Africa (Gillies & de Meillon, 1968). Likewise, we show the constrained, and sympatric, ranges of An. gambiæ and An. funestus groups across the west, central and southern African belts where these highly efficient vector groups continue to contribute to some of the highest P. falciparum transmission rates in Africa (Noor et al., 2014). This ecological range for the most significant of malaria vectors is unlikely to change dramatically with more data or more elaborate modelling. However, far less is known empirically about the possible secondary vectors of malaria in Africa (Table 1).

Secondary vectors of malaria might become increasingly important as indoor-centric vector control efforts change the landscape of dominant vector compositions, resulting in residual transmission being maintained by outdoor biting and resting vectors (Afrane et al., 2016). The description of the complete range of anophelines at sampled locations is often incomplete, largely because sampling strategies focus on dominant indoor biting and resting vectors. Therefore, in the absence of outdoor adult mosquito sampling, possible secondary vectors are not documented (Stevenson & Norris, 2017). It is important to note that these predominantly zoophilic mosquitoes do at times feed on humans and can be found resting indoors. However, more information is required beyond simply defining their ecological range to be able to interpret the importance of these potential vectors of residual malaria transmission in Africa. Abundance is a key feature of a vector’s ability to transmit malaria, and often these data are not readily available from reports of secondary vectors. Some anophelines, however, have other roles in public health, notably their ability to transmit other parasites and viral infections and contributing to the transmission of: filariasis (An. pharoensis) (Gillies & de Meillon, 1968), Rift Valley fever (An. coustani, An. squamosus) (Tantely et al., 2013) (An. pharoensis, An. rufipes, An. coustani) Zika virus (An coustani s.s., An paludis) and Chikungunya virus (An. rufipes, An. coustani) (Tantely et al., 2016).

The most recent geo-coded inventory of anophelines in Africa focused only on dominant vectors, published and contemporary data sources (Wiebe et al., 2017). We have used unpublished national reports dating back to the period before the GMEP, which have provided a rich source of additional information related to dominant and secondary malaria vectors, spanning over 100 years. This geo-coded repository of data is provided on Harvard Dataverse (Snow, 2017), and original source materials have been provided to the Global Malaria Programme of the WHO. These data are therefore available to every national malaria agency responsible for the future control or elimination of malaria across SSA. In addition, the data are also available to national academic counterparts to malaria programmes in SSA interested in vector species niche mapping. We imagine that a first step is that the maps of sampled presence of dominant vectors, their sibling species and potential secondary vectors will be used to highlight where information within national borders is currently absent. Linking mapped species distributions to information on insecticide resistance provides a layered, information platform to manage insecticide use (IRBase; Coleman et al., 2017).

Using information to plan an effective control programme is crucial, this was recognized over 50 years ago in SSA, but conspicuous by its absence at the launch of the recent Roll Back Malaria initiative. As malaria programmes need to become more selective, nuanced and focused in the application of interventions, data platforms are an essential part of the preparation stage. There is a growth in sub-national targeting of resources to meet epidemiological needs, driven largely by variations in empirical data on malaria infection prevalence or routine clinical data. We would encourage countries to build their knowledge base on malaria vector species compositions as part of a broader epidemiological profile within their national borders.

Data availability
The Vectors Database, which includes all the data that support the findings of this study, are available from the KEMRI Wellcome Trust Research Programme’s Population Health Dataverse, http://dx.doi.org/10.7910/DVN/NQ6CUN (Snow, 2017), under a CC-BY 4.0 license.
Competing interests
No competing interests were disclosed.

Grant information
This work was supported by the Wellcome Trust [079080] and [103602] to RWS as part his Principal Fellowship support since 2007, which supported PA, CWM and DK, and [077092] and [203077] are acknowledged by DK, CMM and RWS (Kenya Major Overseas Programme); RWS is also grateful to the UK’s Department for International Development for their continued support to a project Strengthening the Use of Data for Malaria Decision Making in Africa first funded and piloted in 2013 [203155], which provided funding for DK; MC is funded by the South African DST/NRF Research Chairs Initiative.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
The following individuals from the Kenya Medical Research Institute-Wellcome Trust Collaborative Programme provided technical support to locate published materials and geo-code survey sites Caroline Kabaria, Robi Okara, Philip Mbithi, Caroline Tago, Viola Kurui, Betsy Makena, Joe Maina and Winnie Musivo. The authors are grateful to Philip Bejon for reviewing an earlier version of the manuscript. Librarians, archivists and entomologists from across Africa have generously produced unpublished reports as follows:


References
null


Reference Source


Rousseau Djouaka
AgroEcoHealth Platform, International Institute of Tropical Agriculture (IITA), Cotonou, Benin

This submitted manuscript is very relevant for malariologists including policy/decision makers working on malaria vector control. The authors have generated a very impressive set of data from 1898 to 2016 to map and report Anopheles distribution in Africa. This review is very unique in the sense that: (i) it builds on data of more than 118 years old, (ii) it includes both published and unpublished references collected from several libraries and institution archives, (iii) could serve as strong starting point for other research works such as: (a) Mapping and identifying some key historical events which occurred on the dynamics/distribution of malaria vectors in Africa over the past 100 years, (b) generate information on climate changes/variations (when these data are coupled with climatic parameters) and analyse how changing climate conditions have affected the distribution of malaria vectors over the past 100 years in Africa.

As complementary information, I will make 2 suggestions:

- Authors should be more precise on the duration it took to generate such volume of information (...See Abstract Results: “It took several years ... How many years?”

- Authors should clearly describe (under the methodology section) the selection/rejection criteria used for unpublished data and how these data were cleaned before been incorporated among the set of published data.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes
Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.

**Referee Expertise:** Malaria Entomologist, vector-borne disease control.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

**Referee Report 11 August 2017**

doi:10.21956/wellcomeopenres.13191.r24527

Michael J. Bangs
International SOS Public Health & Malaria Control, Kuala Kencana, Indonesia

Kyaloo et al. provide an exceptional paper and excellent summarization (detailed, yet concise) with important historical references to produce a timely resource covering approximately 118 years of observations on anopheline mosquitoes and their respective distributions in sub-Saharan Africa. A number of painstaking years in production, the scope of the coverage is impressive. Data covers not only published accounts but also the partially hidden ‘grey’ literature and miscellaneous reports lurking in the dusty recesses of government archives and elsewhere. As the authors point out, what is presented as inventory is actual recorded collections of specimens, not just predicted distributions (albeit useful in its own right). Both approaches have merit and inherent limitations as to what the data can and cannot provide or be extrapolated upon.

Although the systematics of the anopheline mosquitoes has resulted in revised taxonomy and nomenclature over the many decades (precision in species identification, expanding groups and complexes, refined phylogenetic relationships, for example), one overriding theme to these assembled works has been a greater understanding of the mosquitoes’ complex role in malaria transmission in specific areas. Where competent vectors exist (primary or secondary), perennial or seasonal, the specter of transmission risk persists.

The compiled inventory is a prime example of Science as a series of building blocks, with history serving as its foundation (with the occasional evolving paradigm shift). Never lose sight of the past – it can be invaluable no matter the age. Historical references provide both a temporal-spatial ‘baseline’ and confirmation of vector presence that can too easily become ‘lost’, thereby denying a larger audience of invaluable data on specific locations that might not be revisited for a long time to come. Moreover, this article serves as a template to emulate in other malaria endemic regions of the world – for those so daring.

This publication provides a solid foundation from which to continue to add ‘points’ to the African map as older information continues to surface and new data is gathered. Despite the monumental effort to assemble and display vector distributions, the authors acknowledge there remain documents yet to recover. To point, what is not clear in the presentation is how interested individuals and institutions might be able to contribute additional entries to the vector databases (Harvard Dataverse, KEMRI Vectors Database) from information still ‘hidden’ and yet to be captured.
Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.

**Referee Expertise:** Public health entomology, mosquito bionomics, vector-borne disease control, infectious disease epidemiology.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

Josiane Etang
Organisation for Coordination in the Fight Against Endemic Diseases in Central Africa (OCEAC), Yaoundé, Cameroon

The subject is of relevance to malaria entomology and will be of significant interest to those in the field of malaria control or elimination in Sub-Saharan Africa. The team of authors has a significant track record in this area, having previously carried out major surveys of malaria vectors in Africa. The paper provides an up-to-date catalogue of recorded anopheline species for the Afrotropical region, based on data extracted from peer-reviewed published sources and unpublished research reports, which allow to gather a wide-range of information on primary and secondary malaria vectors in Africa.

However, following are two issues to address:

1. It appears that important sources of information have been national inventories developed by entomologists working in Africa, which sounds very interesting. Furthermore, the use of historical archive and unpublished material has been fundamental in expanding the understanding of the spatial ranges of primary and secondary malaria vectors. However, the authors have not provided a detailed description of data sources to highlight the contribution of unpublished material in the
inventory of anopheles that is being published. It would be interesting to build a table or a figure on breakdown of published and unpublished data.

2. The authors made a huge effort to gather recent data including those from islands; recent data on malaria vectors in the island of Manoka in Cameroon (Mbida et al, 2017, Etang et al. 2017…) may also be included.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
No

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Discuss this Article

Version 1

Author Response 08 Sep 2017

Robert Snow, Centre for Tropical Medicine and Global Health, University of Oxford, UK

We are grateful to the three reviewers for their observations regarding the assembly of anopheline mosquito distributions in Africa. As highlighted by Mike Bangs the added value of the current anopheline site description inventory is that we have consulted a multitude of unpublished sources of information, that date back to the halcyon days of malariology of the 1950s and 1960s. These resources have been neglected in previous assemblies of mosquito data and enrich what has previously been published.

In addition, we have provided digital copies of original published and unpublished reports to the Global Malaria Programme, World Health Organization, so that these might be shared with countries in the Africa region. This valuable resource is the legacy of information from malaria vector surveys undertaken over the
past 100+ years.

To answer Rousseau Djouaka’s question, the work has spanned many years, in different guises. 21 years ago, we initiated a Pan-African collaboration known as the Mapping Malaria Risk in Africa (MARA/ARMA) project, in 2005 the Malaria Atlas Project (MAP) was launched in Nairobi and more recently we have begun collaborations around building layers of malaria epidemiological data to support national malaria control programmes in Africa (INFORM & LINK). A history of these data assembly initiatives can be found at [http://kemri-wellcome.org/programme/population-health/](http://kemri-wellcome.org/programme/population-health/). We began to systematically review, extract data and geo-code all assembled published and unpublished information identified by us since 1996 in March 2010.

Josiane Etang’s comment regarding sources of information, these are available alongside each record in the main database available at Snow RW (2017). A geo-coded inventory of anophelines in the Afrotropical Region south of the Sahara: 1898-2016, doi:10.7910/DVN/NQ6CUN, Harvard Dataverse, V1, under a CC-BY 4.0 license.

We will have missed unpublished reports, still buried in academic and ministry of health archives across Africa. We might also have missed several peer-reviewed reports. When data searches are opportunistic, cascaded and dependent upon access to historical archives, it is always hard to know what has been captured and what remains to be assembled. What we do hope however, is that national malaria control programmes across Africa and their national academic partners would use what we have started to resurrect the need to maintain, map and improve mapped representations of mosquito vector inventories.

Mike Bangs asks how will this inventory be updated. Short term grants help to initiate data assembly resources, but rarely are these funded beyond 3-5 years and often funding is awarded to research programmes based outside of the malaria endemic world.

We would argue that any basic information platform on malaria at country-levels must include an understanding of the vectors; only through an understanding of the complex epidemiology of malaria, past and present, will future disease control and elimination be successful; and using data to inform malaria must become the norm and driven by partnerships between academics and ministries of health at country level.

As such we believe that the inventories developed by us should serve as the first step in establishing national geo-coded databases of mosquito vectors and that these should ultimately be linked to many other malaria information layers through government led initiatives at country levels. This may take time, and it may require WHO technical guidance.

The authors

**Competing Interests:** We have no competing interests

-----

Reader Comment 27 Jul 2017

**Peter Macharia**, KEMRI Wellcome Trust Research programme, Kenya

A legacy data set that will inform National Malaria Control Programmes in Africa on appropriate control interventions targeting for years to come.
**Competing Interests:** Researcher at KEMRI - Wellcome Trust