LETTER

https://doi.org/10.1038/s41586-019-1622-4

Windborne long-distance migration of malaria mosquitoes in the Sahel

Diana L. Huestis[1], Adama Dao[2], Moussa Diallo[2], Zana L. Sanogo[2], Djibril Samake[2], Alpha S. Yaro[2,3], Yossi Ousman[2], Yvonne-Marie Linton[4,5], Asha Krishna[1], Laura Veru[1], Benjamin J. Krajacich[1], Roy Faiman[1], Jenna Florio[1], Jason W. Chapman[6,7,8], Don R. Reynolds[9,10], David Weetman[11], Reed Mitchell[4], Martin J. Donnelly[11], Elijah Talamas[12,13], Lourdes Chamorro[5,12], Ehud Strobach[14,15] & Tovi Lehmann[1]*

Over the past two decades efforts to control malaria have halved the number of cases globally, yet burdens remain high in much of Africa and the elimination of malaria has not been achieved even in areas where extreme reductions have been sustained, such as South Africa[1,2]. Studies seeking to understand the paradoxical persistence of malaria in areas in which surface water is absent for 3–8 months of the year have suggested that some species of *Anopheles* mosquito use long-distance migration[3]. Here we confirm this hypothesis through aerial sampling of mosquitoes at 40–290 m above ground level and provide—to our knowledge—the first evidence of windborne migration of African malaria vectors, and consequently of the pathogens that they transmit. Ten species, including the primary malaria vector *Anopheles coluzzii*, were identified among 235 anopheline mosquitoes that were captured during 617 nocturnal aerial collections in the Sahel of Mali. Notably, females accounted for more than 80% of all of the mosquitoes that we collected. Of these, 90% had taken a blood meal before their migration, which implies that pathogens are probably transported over long distances by migrating females. The likelihood of capturing *Anopheles* species increased with altitude (the height of the sampling panel above ground level) and during the wet seasons, but variation between years and localities was minimal. Simulated trajectories of mosquito flights indicated that there would be mean nightly displacements of up to 300 km for 9-h flight durations. Annually, the estimated numbers of mosquitoes at altitude that cross a 100-km line perpendicular to the prevailing wind direction included 81,000 *Anopheles gambiae* sensu stricto, 6 million *A. coluzzii* and 44 million *Anopheles squamosus*. These results provide compelling evidence that millions of malaria vectors that have previously fed on blood frequently migrate over hundreds of kilometres, and thus almost certainly spread malaria over these distances. The successful elimination of malaria may therefore depend on whether the sources of migrant vectors can be identified and controlled.

In Africa, malaria spans the humid equatorial forest to the semi-arid zones in the north and south. In regions in which surface water—which is essential for larval development—is absent during the 3–8-month dry season, mosquito densities and disease transmission drop markedly[3–8]. However, shortly after the first rains, populations of vectors surge[6], and transmission recommences. Recent studies suggest that in the Sahel *A. coluzzii* survives the long dry season by aestivation (a period of dormancy)[3,6,9,10], whereas *A. gambiae* sensu stricto (hereafter, *A. gambiae*) and *Anopheles arabiensis* re-establish populations by migration from distant locations, at which larval sites are perennial[3]. However, direct evidence, such as the capture of aestivating adults in their shelters or the recapture of marked mosquitoes at sites that are hundreds of kilometres from their release sites, remains limited.

The dispersal of mosquitoes (hereafter referred to as migration[11]) has been extensively studied because it directly affects disease transmission, the spread of adaptations (for example, resistance to insecticides) and strategies for controlling mosquitoes (such as insecticide barriers)[12,13]. Although tracking mosquitoes over large scales has seldom been attempted[12,13], the prevailing view is that the dispersal of malaria mosquitoes[12–15] does not exceed 5 km and long-range movements[16–19] represent 'accidental events' that are of minimal epidemiological importance[12]. Nonetheless, the prediction of long-distance migration of anopheline mosquitoes in the Sahel prompted us to question this view. To our knowledge, our study is the first to systematically sample insects that migrate at high altitudes over multiple seasons in Africa. We aimed to determine whether malaria vectors engage in wind-assisted movements and—if so—to assess the epidemiological relevance of this movement by addressing questions regarding the species involved, the frequency and heights of flights, how many mosquitoes migrate and how likely these mosquitoes are to carry *Plasmodium*. We then use simulations to estimate how far mosquitoes may have travelled, and from where.

During 617 nights on which aerial sampling took place, we caught 461,100 insects at heights that ranged between 40 and 290 m above ground level in four villages in the Sahel of Mali (Extended Data Fig. 1). These insects included 2,748 mosquitoes, of which 235 were anopheline mosquitoes (Table 1). These mosquitoes belonged to ten species: *A. coluzzii*, *A. gambiae*, *Anopheles pharoensis*, *Anopheles coustani*, *A. squamosus*, *A. rufipes* and *A. namibiensis*, as well as three distinct but currently undetermined *Anopheles* species (referred to here as *Anopheles* Mali species 1, *Anopheles* Mali species 2 and *Anopheles* species near (sp. nr) *concolor*) (Table 1). *A. coluzzii* and *A. gambiae* are the primary vectors of malaria in Africa, and *A. pharoensis*, *A. coustani*, *A. squamous* and *A. rufipes* are of secondary importance[20]. Mosquitoes were not among the 564 insects that were captured on 508 control nets (Table 1, Methods), which confirmed that these *Anopheles* mosquitoes were intercepted at altitude rather than near the ground during deployment. The maximum number of an anopheline species caught per night was five, which indicates that migration occurred over many nights. Consistent with Poisson distributions, the values of the variance:mean ratio were all near one (Table 1, Supplementary Discussion). Unless specified otherwise, the quantitative results presented here refer to the 5 most-abundant species, represented by more than 20 individuals (Table 1).

[1]Laboratory of Malaria and Vector Research, NIAID, NIH, Rockville, MD, USA. [2]Malaria Research and Training Center (MRTC), Faculty of Medicine, Pharmacy and Odonto-stomatology, University of Bamako, Bamako, Mali. [3]Faculte des Sciences et Techniques, Universite des Sciences des Techniques et des Technologies de Bamako (FSTUSTTB), Bamako, Mali. [4]Walter Reed Biosystematics Unit, Smithsonian Institution Museum Support Center, Suitland, MD, USA. [5]Department of Entomology, Smithsonian Institution, National Museum of Natural History, Washington, DC, USA. [6]Centre for Ecology and Conservation, University of Exeter, Penryn, UK. [7]College of Plant Protection, Nanjing Agricultural University, Nanjing, China. [8]Environment and Sustainability Institute, University of Exeter, Penryn, UK. [9]Natural Resources Institute, University of Greenwich, Chatham, UK. [10]Rothamsted Research, Harpenden, UK. [11]Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK. [12]Systematic Entomology Laboratory - ARS, USDA, Smithsonian Institution National Museum of Natural History, Washington, DC, USA. [13]Florida Department of Agriculture and Consumer Services, Department of Plant Industry, Gainesville, FL, USA. [14]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA. [15]Global Modeling and Assimilation Office, NASA GSFC, Greenbelt, MD, USA. *e-mail: tlehmann@nih.gov

Table 1 | Summary of mosquitoes collected in aerial samples in 2013–2015

	Standard panels[a]											Control panels[b]		
Taxa	Total captured	Mean panel density	Lower 95% CLM[c]	Upper 95% CLM[c]	Maxium per panel	Per cent nightly presence	Variance: mean ratio	Per cent female	Per cent post-blood feed[d]	Per cent infected[e]	Per cent anthropophily[g]	Total captured	Mean panel density	Maximum per panel
A. squamosus	100	0.053	0.042	0.063	3	11.02	1.37	76.0 (96)	93.2 (73)	0 (73)	41.1 (17)	0	0	0
A. pharoensis	40	0.021	0.015	0.028	2	6.00	1.08	82.5 (40)	100 (33)	0 (33)	33.3 (6)	0	0	0
A. coustani	30	0.016	0.01	0.022	2	4.38	1.05	88.9 (27)	87.5 (24)	0 (24)	14.3 (7)	0	0	0
A. rufipes	24	0.013	0.008	0.018	2	3.24	1.16	80 (20)	93.8 (16)	0 (16)	0 (4)	0	0	0
A. coluzzii	23	0.012	0.007	0.017	2	3.08	1.16	95.5 (22)	90.5 (21)	0 (21)	100 (1)	0	0	0
Anopheles. Mali species 1	2	0.001	0	0.003	1	0.32	1	100 (2)	100 (2)	0 (2)	ND	0	0	0
A. gambiae	1	0.0005	0	0.002	1	0.16	1	100 (1)	100 (1)	0 (1)	ND	0	0	0
Anopheles sp. nr *concolor*[f]	1	0.0005	0	0.002	1	0.16	1	0 (1)	NA	NA	NA	0	0	0
Anopheles Mali species 2	1	0.0005	0	0.002	1	0.16	1	100 (1)	100 (1)	0 (1)	ND	0	0	0
A. namibiensis	1	0.0005	0	0.002	1	0.16	1	100 (1)	100 (1)	0 (1)	ND	0	0	0
Anopheles unidentified	12	0.006	0.003	0.01	1	1.78	0.99	33.3 (6)	100 (2)	0 (2)	ND	0	0	0
Culicinae	2,340	1.236	1.185	1.286	22	58.19	4.83	86.4 (1,866)	96.7 (1,629)	ND	ND	0	0	0
Culicid unidentified	173	0.091	0.078	0.105	8	17.18	1.92	62.9 (116)	91.8 (73)	ND	ND	0	0	0
Total Culicidae	2,748	1.451	1.397	1.505	23	64.18	4.92	84.5 (1,876)	96.2 (1,804)	ND	ND	0	0	0
Total insects	461,100	243.58	242.88	244.29	2,601	100	314.75	ND	ND	NA	NA	564	1.110	31

Numbers in parentheses refer to the number of insects from which the percentages were calculated (that is, some of the total captured insects could not be assigned a sex or could not be tested). CLM, confidence limit of the mean; NA, not applicable; ND, not determined.

[a]Nightly aerial sampling using sticky nets (panels, usually three per balloon) were launched and retrieved at 17:00 and 07:00, respectively. Nets were raised to set altitudes between 40 and 290 m above ground (Methods). $n = 1{,}894$ panels.

[b]Control panels were raised to 40–120 m above ground level, and immediately retrieved during the launch and retrieval of the standard panels, to estimate the number of insects that were captured during the ascent and descent (Methods). $n = 508$ panels.

[c]These values were estimated using the normal approximation of the Poisson distribution. When a single mosquito per taxon was captured, the low negative values (< -0.0001) were rounded to zero.

[d]Only a few blood-fed and half-gravid females were pooled with gravid females to reflect those that were evidently exposed to at least one blood meal. Unfed mosquitoes consisted of the rest.

[e]Infection with human *Plasmodium* species was tested as described in the Methods.

[f]This species was identified on the basis of its genitalia (male).

[g]Identified using PCR (Methods), with additional confirmations by sequencing. The nonhuman hosts included cow, goat and—possibly—unknown rodents.

Females outnumbered males by more than 4:1 (Table 1). Critically, more than 90% of the anopheline females had taken a blood meal (2.9% were blood-fed, 87.5% were fully gravid and 0.7% were semi-gravid; blood feeding is required for the eggs of females to mature) before their high-altitude flights (Table 1), which suggests they were probably exposed to malaria and other pathogens. Although 31% of the blood meals came from humans, no *Plasmodium*-infected mosquitoes were detected among the 22 *A. gambiae* sensu lato (here represented by *A. gambiae* and *A. coluzzii*) that we tested, or the 174 secondary vectors (Table 1). Considering the typical rates of *Plasmodium* infections in primary (1–5%) and secondary (0.1–1%) vectors[5,21–23], our results probably reflect the small sample size: the likelihood for zero infected mosquitoes was over 30% and over 18% (assuming the highest rates in each range) in the primary and secondary vectors, respectively (Supplementary Discussion). Therefore, unless infection reduces migratory capacity or migrants are resistant to parasites (and there is currently no evidence for either), *Plasmodium* and other pathogens are almost certainly transported by windborne mosquitoes that may infect people post-migration.

The mosquitoes were intercepted between 40 and 290 m above ground level (Fig. 1a). The mean panel density of mosquitoes per altitude and the corresponding aerial density (the panel density divided by the volume of air sampled) increased with altitude and there was a significant effect across species on mean panel density ($P < 0.037$, $F_{1,24} = 4.9$) (Extended Data Fig. 2b), which suggests that the migration of anopheline mosquitoes also occurs more than 290 m above ground level. The similar distribution of species across years and locations (Extended Data Fig. 2c; non-significant effects of year and location shown in Extended Data Table 1), combined with the marked seasonality of the migration or high-altitude flight activity (the aerial captures of mosquitoes occurred between July and November, and peaked between August and October) (Fig. 1b, Extended Data Table 1), attest to the regularity of the migration of windborne *Anopheles* mosquitoes.

Using mean aerial densities and wind speeds at altitude (4.8 m s^{-1}) (Fig. 1c) and conservatively assuming that mosquitoes fly in a layer between 50 and 250 m above ground level, we estimated the nightly expected numbers of migrants crossing a 1-km line perpendicular to the wind direction. Nightly estimates ranged between 27 (for *A. gambiae*) and 3,719 (for *A. squamosus*) (Fig. 1d). When interpolated over a 100-km line that links our sampling sites (Extended Data Figs. 1a, 2c), annual migrations are estimated to exceed 80,000 *A. gambiae*, 6 million *A. coluzzii* and 44 million *A. squamosus* mosquitoes in that region alone (Fig. 1d). Thus, the migration of windborne mosquitoes in the Sahel occurs on a massive scale.

For each mosquito capture, we estimated the flight trajectories for 2- and 9-h-long flights using the hybrid single-particle Lagrangian integrated trajectory (HYSPLIT) model[24] with the most accurate assimilated meteorological data available (ERA5), and assuming that mosquitoes ascend by their own flight but are passively carried by the wind at altitude (Methods). The mean nightly displacements were 30 and 120 km (maxima of 70 and 295 km) for 2-h and 9-h flights, respectively (Fig. 2, Table 2). Notably, the maximal 9-h nightly flight displacements ranged between 257 and 295 km for all anopheline species with sample size of more than 20 insects (Table 2). These backwards trajectories exhibited a southwestern origin (Rayleigh test; a mean bearing of 212°, $r = 0.54$, $P < 0.0001$) (Table 2), which corresponds to the prevailing winds during peak migration (August to September) (Fig. 2). The trajectories of most species originated from a broad arc (over 90°) (Fig. 2), which suggests migrants emanated from multiple sites across a large region. Migration from this direction is consistent with the presence of high-density populations due to perennial larval sites and earlier population growth following the monsoon rains. The backwards trajectories with a strong northerly component, which were observed during the sparsely sampled period of October to December (Fig. 2), might indicate southward 'return flights' on the Harmattan winds that prevail during this season.

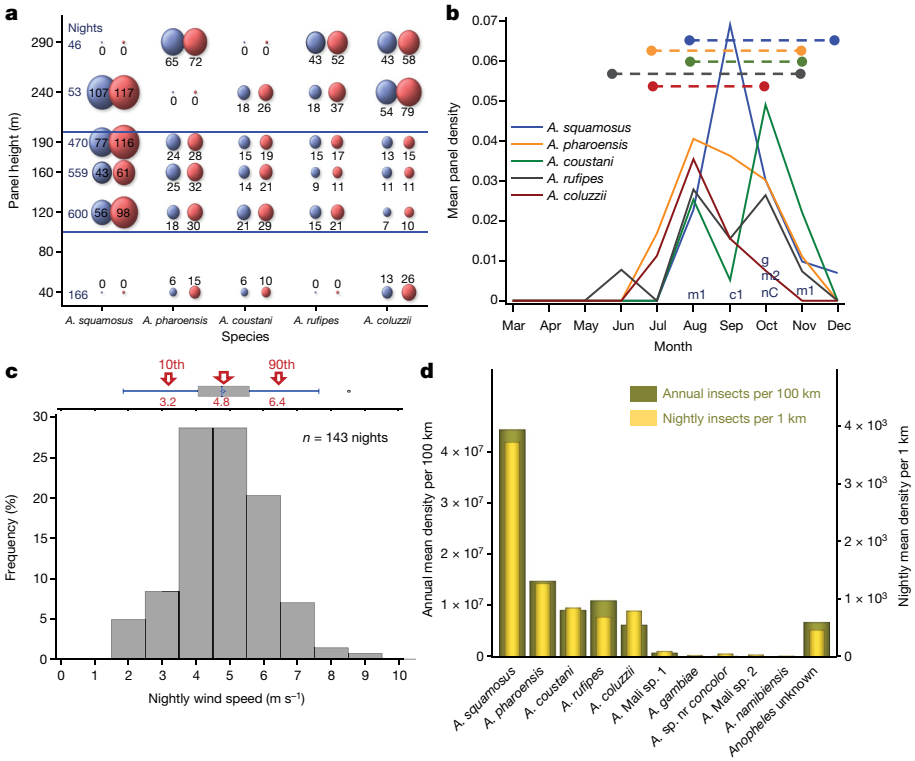


Fig. 1 | Flight altitude, seasonality, wind speed and abundance of migratory anopheline species. a, The relationship of altitude (height of the collecting panel) to panel density (blue) and aerial density (orange, mosquitoes per 10^6 m^3 of air), for the five most-common anopheline species (Table 1). Bubble size is proportional to mosquito density (the value shown in the bubble $\times$ 10^3); when the value is zero, only a dot is shown. The number of sampling nights for each of the collecting-panel heights is shown on the left. **b**, Monthly panel density (n = 1,894 panels) for the five most-common species (Table 1), overlaid by the length of migration period (dashed lines). Values for *A. squamosus* were divided by three to preserve the scale. The sampling month for species that were collected only once or twice is shown by letters: c1, *A. namibiensis*; g, *A. gambiae*; m1, *Anopheles* Mali species 1; m2, *Anopheles* Mali species 2; nC, *A.* sp. nr *concolor*. **c**, Distribution of the mean nightly wind speed at flight height on nights on which one or more anopheline mosquitoes were collected. Wind-speed data were taken from ERA5 database after matching the height of the collecting panel to the nearest vertical layer (Methods). The corresponding box and whisker plot (top) shows the median, mean, quartiles and extreme values overlaid by arrows indicating the mean and 10th and 90th percentiles (red). **d**, The number of mosquitoes per species that cross, at altitude (50–250 m above ground level), imaginary lines perpendicular to the prevailing wind direction. Migrants per night per 1 km (right y axis) are superimposed on the annual number of migrants per 100-km line (left y axis).

Contrary to the conventional view that dispersal of African anopheline mosquitoes occurs over distances[12,14,15,25] of less than 5 km, our results provide compelling evidence that primary and secondary malaria vectors regularly engage in high-altitude flights (or windborne) migrations that span tens to hundreds of kilometres per night. Because this migration includes large numbers of females that had taken at least one blood meal, it probably involves human *Plasmodium* (among other pathogens). Separate outbreaks of malaria in Egypt and Israel have previously been attributed[16] to *A. pharoensis* travelling more than 280 km. Assuming a conservative[22,26] 1% infection rate in migrating females of *A. coluzzii, A. gambiae*, *A. coustani* and *A. pharoensis* and 0.1% in the remaining anopheline mosquitoes (excluding the unknown *Anopheles* Mali species 1 and *Anopheles* Mali species 2) (Supplementary Discussion), more than 286,000 infected migrant mosquitoes are expected to annually cross, at altitude, a 100-km line perpendicular to the prevailing wind direction. Accordingly, *A. pharoensis*, *A. coustani* and *A. coluzzii* would contribute 41%, 25% and 17%, respectively, to malaria transmission by infected windborne mosquitoes. Although these estimates are coarse, this suggests that migratory secondary vectors could be a major source of infections, and that they should be included in studies of transmission and in control programs.

A. coluzzii was more common than *A. gambiae* among the migrants, which was contrary to initial predictions[3] that were based on data suggesting that *A. coluzzii* aestivates locally and therefore may not require migration to recolonize the Sahel. Indeed, migration occurs from the end of July to October—well after the surge of *A. coluzzii* populations in the Sahel following the first rain (May to June)[3,6]. Northward and southward oscillations of the intertropical convergence zone during the wet season continually reconfigure the better resource patches for mosquitoes, as the intensity of the rains shifts in location. Additionally, wet-season droughts endanger local mosquito populations every decade or two[27]. Thus, selection pressures to track freshwater resources by riding the winds that bring rain[28] may explain why residents of the Sahel, such as *Oedaleus senegalensis* grasshoppers and *A. coluzzii*, have a mixed strategy of migration[29] and local dormancy. *A. gambiae*, which presumably recolonizes the Sahel every wet season, is relatively rare in villages in the Sahel[3]—therefore, only one specimen of this species was captured by our nets. It may migrate on fewer nights and constitute a smaller fraction of windborne migrants (Supplementary Discussion).

In areas in which malaria is approaching elimination, cases of the disease that occur without travel history are presumed to represent indigenous transmission. We propose that a substantial fraction of such cases, especially those that occur within about 300 km of areas with high rates of malaria transmission, arise from the bites of exogenous, windborne infected mosquitoes. For example, northeastern South Africa has the highest incidence of persistent malaria in the country, with many cases not associated with human travel: these cases are concentrated in an arc that extends over about 150 km from the borders with Zimbabwe and Mozambique, where transmission is high. This area also includes the Kruger National Park, in which roads are scarce and vehicular transport of infected mosquitoes[30] may be hampered. Testing the correlation of infection events such as these with the corresponding prevailing wind

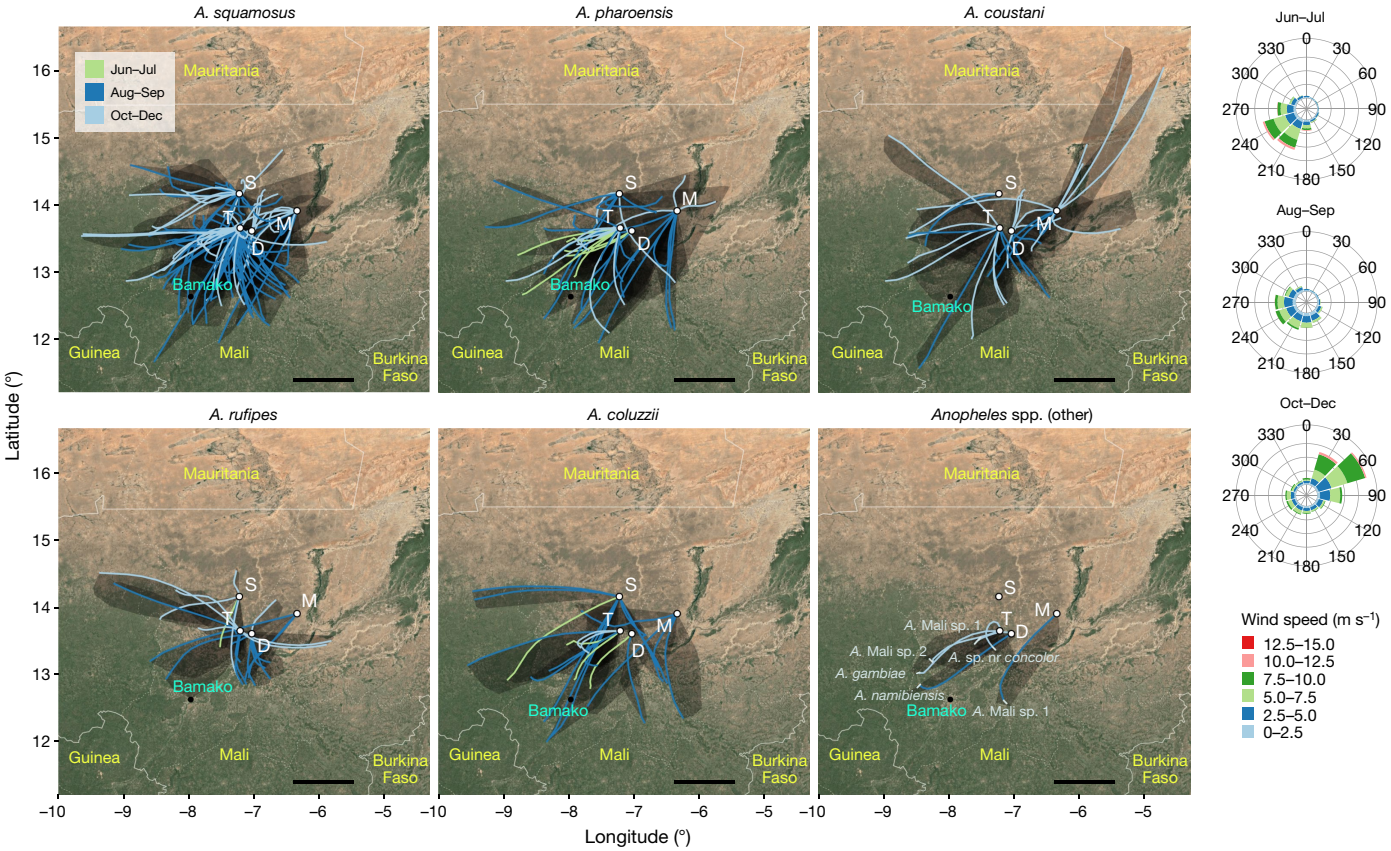


Fig. 2 | Backwards trajectories of the flights for each *Anopheles* capture event. Backward 9-h trajectories were estimated by HYSPLIT (Table 2), and are overlaid on a map showing parts of Mali and neighbouring countries. Map data are from Google, Landsat/Copernicus 2019. Each line represents one of four simulated trajectories of one (or more) mosquitoes intercepted at that location and on that night. The area encompassed by the four trajectories is shadowed. Migration season is shown by line colour. The *Anopheles* species is indicated above each panel. D, M, S and T are the balloon launch locations in the villages of Dallowere, Markabougou, Siguima and Thierola, respectively. Scale bars, 100 km. The seasonal-wind rose diagrams, reflecting wind conditions at 180 m above ground level averaged from 2013 to 2015, are shown at the right.

Table 2 | Summary of displacement distance (in a straight line) and source direction, based on 2-h and 9-h flight trajectories of mosquitoes

	Trajectories for 2-h flights				Trajectories for 9-h flights								
Taxa	Trajectories[a]	Displace mean	Displace 95% CLM	Displace min.–max.	Trajectories[a]	Displace mean	Displace 95% CLM	Displace min.–max.	Hourly disp. mean[c]	Actual hourly disp. mean[d]	Mean final bearing[e]	*R* (bearing)[f]	P_R
A. squamosus	1,100	27.7	27–29	2–68	400	109.1	103–115	4–265	13.3	12.1	213	0.516	0.0000
A. pharoensis	440	31.1	30–33	1–65	160	125.3	116–134	24–260	14.7	13.9	214	0.660	0.0000
A. coustani	330	28.5	27–30	2–60	120	125.8	114–138	16–295	14.5	14.0	199	0.270	0.0802
A. rufipes	264	26.1	24–28	2–70	96	109.2	97–121	24–257	12.5	12.1	199	0.454	0.0003
A. coluzzii	253	38.6	37–41	3–69	92	154.1	140–168	47–270	17.3	17.1	217	0.815	0.0000
Anopheles Mali species 1	22	20	14–26	6–52	8	94.3	52–136	51–172	10.2	10.5	223	0.947	0.0000
A. gambiae	11	33.5	ND[b]	ND[b]	4	131.1	ND[b]	ND[b]	15.9	14.6	254	ND[b]	ND[b]
A. sp. nr *concolor*	11	17.2	ND[b]	ND[b]	4	48.2	ND[b]	ND[b]	8.4	5.4	184	ND[b]	ND[b]
Anopheles Mali species 2	11	29.9	ND[b]	ND[b]	4	104.4	ND[b]	ND[b]	13.1	11.6	234	ND[b]	ND[b]
A. namibiensis	11	40.1	ND[b]	ND[b]	4	149.3	ND[b]	ND[b]	16.7	16.6	241	ND[b]	ND[b]
Overall	2,453	29.4	28.8–30.0	1–70.4	892	118.8	115–123	4–295	14.1	13.2	212	0.540	0.0000

Trajectories were produced using HYSPLIT; further information is provided in Methods and Fig. 2. Disp., displacements; displace mean, mean displacement; displace min.–max., minimum and maximum range of displacement.

[a]The number of unique nightly trajectories assumes all possible nightly interception times, given flight duration and a flight start and end of 18:00 and 06:00, respectively. Thus, for each night with a captured mosquito there were 11 unique 2-h flight trajectories and 4 unique 9-h flight trajectories.

[b]Not determined for species of which only a single specimen was captured.

[c]Hourly displacement between successive 1-h points along the 9-h trajectory.

[d]Effective hourly displacement computed as the quotient of the total 9-h trajectory displacement by 9.

[e]The mean bearing (angle) between the interception point (zero) and the final point of the 9-h trajectory computed from the north.

[f]A measure of angular dispersion that varies from 0 (uniform dispersion from all directions) to 1 (a single angle to which all points align).

direction will help to assess this hypothesis. If confirmed, the incorporation of disease-control efforts in source populations to minimize or block migration is likely to be an essential element in strategies to eliminate malaria.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1622-4.

Received: 15 February 2019; Accepted: 6 September 2019;
Published online: 02 October 2019

1. WHO. *World Malaria Report 2017*. (WHO, 2018).
2. Gething, P. W. et al. Mapping *Plasmodium falciparum* mortality in Africa between 1990 and 2015. *N. Engl. J. Med*. **375**, 2435–2445 (2016).
3. Dao, A. et al. Signatures of aestivation and migration in Sahelian malaria mosquito populations. *Nature* **516**, 387–390 (2014).
4. Fontenille, D. et al. High annual and seasonal variations in malaria transmission by anophelines and vector species composition in Dielmo, a holoendemic area in Senegal. *Am. J. Trop. Med. Hyg*. **56**, 247–253 (1997).
5. Fontenille, D. et al. Four years' entomological study of the transmission of seasonal malaria in Senegal and the bionomics of *Anopheles gambiae* and *A. arabiensis*. *Trans. R. Soc. Trop. Med. Hyg*. **91**, 647–652 (1997).
6. Lehmann, T. et al. Aestivation of the African malaria mosquito, *Anopheles gambiae* in the Sahel. *Am. J. Trop. Med. Hyg*. **83**, 601–606 (2010).
7. Simard, F., Lehmann, T., Lemasson, J. J., Diatta, M. & Fontenille, D. Persistence of *Anopheles arabiensis* during the severe dry season conditions in Senegal: an indirect approach using microsatellite loci. *Insect Mol. Biol*. **9**, 467–479 (2000).
8. Omer, S. M. & Cloudsley-Thompson, J. L. Dry season biology of *Anopheles gambiae* Giles in the Sudan. *Nature* **217**, 879–880 (1968).
9. Mamai, W. et al. Monitoring dry season persistence of *Anopheles gambiae* s.l. populations in a contained semi-field system in southwestern Burkina Faso, West Africa. *J. Med. Entomol*. **53**, 130–138 (2016).
10. Yaro, A. S. et al. Dry season reproductive depression *of Anopheles gambiae* in the Sahel. *J. Insect Physiol*. **58**, 1050–1059 (2012).
11. Chapman, J. W., Reynolds, D. R. & Wilson, K. Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecol. Lett*. **18**, 287–302 (2015).
12. Service, M. W. Mosquito (Diptera: Culicidae) dispersal—the long and short of it. *J. Med. Entomol*. **34**, 579–588 (1997).
13. Service, M. W. *Mosquito Ecology Field Sampling Methods* (Elsevier Applied Science, 1993).
14. Costantini, C. et al. Density, survival and dispersal of *Anopheles gambiae* complex mosquitoes in a West African Sudan savanna village. *Med. Vet. Entomol*. **10**, 203–219 (1996).
15. Touré, Y. T. et al. Mark-release-recapture experiments with *Anopheles gambiae* s.l. in Banambani Village, Mali, to determine population size and structure. *Med. Vet. Entomol*. **12**, 74–83 (1998).
16. Garrett-Jones, C. The possibility of active long-distance migrations by *Anopheles pharoensis* Theobald. *Bull. World Health Organ*. **27**, 299–302 (1962).
17. Sellers, R. F. Weather, host and vector—their interplay in the spread of insect-borne animal virus diseases. *J. Hyg. (Lond.)* **85**, 65–102 (1980).
18. Glick, P. A. *The Distribution of Insects, Spiders, and Mites in the Air.* Technical Bulletin No. 673 (US Department of Agriculture, 1939).
19. Reynolds, D. R. et al. Atmospheric transport of mosquitoes in northeast India. *Med. Vet. Entomol*. **10**, 185–186 (1996).
20. Kyalo, D. et al. A geo-coded inventory of anophelines in the Afrotropical Region south of the Sahara: 1898–2016. *Welcome Open Res*. **2**, 57 (2017).
21. Beier, J. C. et al. Characterization of malaria transmission by *Anopheles* (Diptera: Culicidae) in western Kenya in preparation for malaria vaccine trials. *J. Med. Entomol*. **27**, 570–577 (1990).
22. Antonio-Nkondjio, C. et al. Complexity of the malaria vectorial system in Cameroon: contribution of secondary vectors to malaria transmission. *J. Med. Entomol*. **43**, 1215–1221 (2006).
23. Toure, Y. T. et al. Perennial transmission of malaria by the *Anopheles gambiae* complex in a north Sudan Savanna area of Mali. *Med. Vet. Entomol*. **10**, 197–199 (1996).
24. Stein, A. F. et al. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Am. Meteorol. Soc*. **96**, 2059–2077 (2015).
25. Verdonschot, P. F. M. & Besse-Lototskaya, A. A. Flight distance of mosquitoes (Culicidae): a metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. *Limnologica* **45**, 69–79 (2014).
26. Hay, S. I., Rogers, D. J., Toomer, J. F. & Snow, R. W. Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review. *Trans. R. Soc. Trop. Med. Hyg*. **94**, 113–127 (2000).
27. Nicholson, S. E. The West African Sahel: a review of recent studies on the rainfall regime and its interannual variability. *ISRN Meteorol*. **2013**, 453521 (2013).
28. Wilson, K. in *Insect Migration: Tracking Resources through Space and Time* (eds Drake, V. A. & Gatehouse, A. G.) 243–264 (Cambridge Univ. Press, 1995).
29. Pedgley, D. E., Reynolds, D. R. & Tatchell, G. M. in *Insect Migration: Tracking Resources through Space and Time* (eds Drake, V. A. & Gatehouse, A. G.) 3–30 (Cambridge Univ. Press, 1995).
30. Frean, J., Brooke, B., Thomas, J. & Blumberg, L. Odyssean malaria outbreaks in Gauteng Province, South Africa, 2007–2013. *S. Afr. Med. J*. **104**, 335–338 (2014).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Study area. Aerial sampling stations were located in four villages in the Sahel in Mali (Extended Data Fig. 1): Thierola (13° 39′ 30.96″ N, 7° 12′ 52.92″ W) from March 2013 to November 2015, Siguima (14° 10′ 3.36″ N, 7° 13′ 40.44″ W) from March 2013 to October 2015; Markabougou (13° 54′ 51.84″ N, 6° 20′ 37.68″ W) from June 2013 to April 2015; and Dallowere (13° 36′ 56.88″ N, 7° 2′ 12.84″ W) from July 2015 to November 2015. This study area has previously been described in detail[3,6,9,11,31–33]. In brief, the region is rural, characterized by scattered villages with traditional mud-brick houses that are surrounded by fields. A single growing season (June to October) enables the farming of millet, sorghum, maize and peanuts, as well as subsistence vegetable gardens. Over 90% of the annual rains fall during this season (about 550 mm). Cattle, sheep and goats graze in the savannah that consists of grasses, shrubs and scattered trees. The rains form small puddles and larger seasonal ponds that usually are totally dry by the end of November. From November until May, rainfall is absent or negligible (total precipitation of less than 50 mm), and by December water is available only in deep wells.

Aerial sampling and specimen processing. Aerial sampling stations were placed about 0.5 km from the nearest house of the village in open areas away from large trees. The method of aerial collection of insects was adapted from a study on high-altitude mating flights in ants[34]. Rectangular 3 × 1-m nets (3 m^2), cut from a roll of tulle netting (mesh of 8 holes per square centimetre, with hole diameter of 1.2 mm), were sewn to form 4 narrow sleeves, 1 m apart, along the net (Extended Data Fig. 3). A 1-m carbon rod was inserted into each sleeve and glued to the net using Duco Cement Glue (Devcon) (Extended Data Fig. 3). Three nets were spread over each other on a clean large wooden table topped by 3.5 × 1.5-m plywood and coated with a thin film of insect glue (Tanglefoot, Tropical Formula, Contech Enterprises) by rolling a PVC pipe smeared with this glue over them, while applying moderate pressure downward. The pipe was held at each end (from each side of the long table) by two persons and repeatedly rolled (and smeared) until a uniform thin layer of glue coated the net but did not block its holes. After coating, the sticky nets were immediately rolled individually, and kept in two tightly secured plastic bags indoors, to avoid accidental contact with insects before setup.

Before the launch, polyurethane balloons (3 m in diameter; from Mobile Airship & Blimps, or Lighter than Air) were inflated to full capacity with balloon-grade helium (>98.5%) and topped up to ensure full capacity as needed, usually every 1–3 days based on the balloon condition (Extended Data Fig. 3). Typically, balloons were launched over about ten consecutive nights per month. The balloon was kept stationary at about 200 m above ground level by a cord (AmSteelBlue, synthetic rope sling, Southwest Ocean Services) secured to a 1-m^3 cement block inserted under the ground. The cord then passed through a horizontal, manually rotating drum made of a garden-hose reel that was used for reeling it in. A larger 3.3-m diameter balloon (Lighter than Air) was used between July and September 2015, and launched to about 300 m above ground level.

A team of five trained technicians operated each aerial-sampling station. During the launch of a balloon, one team member held the cord under the balloon with heavy-duty gloves and manually controlled its ascent and descent, another team member controlled the reel, and the other three team members added or removed the sticky nets to and from their specified positions on the cord. The nets were attached to Velcro panels that had previously been placed on the cord at desirable positions, and spaced to fit each of the matching Velcro pieces on the four carbon rods (Extended Data Fig. 3). A knot was made below the top-most Velcro panel and above the bottom-most Velcro panel, to ensure that the nets would remain stretched (rather than slip on the cord) even in strong winds. Additionally, the team secured the balloons over a 'landing patch' that was padded by tyres covered by a tarpaulin. The balloon was secured to the ground through its main cord by a central hook (at the middle of the landing patch), and by a large tarpaulin that covered it from the top, which was secured to the ground using 14 large stakes. Team members inspected the nets upon launch to verify that they were free of insects. Upon retrieval of the balloon, the team worked in reverse order and immediately rolled each sticky net and placed it into a clean labelled plastic bag, and inserted this into another bag; each of the bags was tightened with a cord until inspection.

Each balloon typically carried three sticky nets. Initially, the nets were suspended at 40, 120 and 160 m above ground level. From August 2013, the typical altitude was set to 90, 120 and 190 m above ground level. When the larger balloon was deployed in the Thierola station (August–September 2015), two additional nets were added at 240 and 290 m above ground level. Balloons were launched approximately 1 h before sunset (about 17:00) and retrieved 1 h after sunrise (about 07:30) the following morning. To control for insects trapped near the ground as the nets were raised and lowered, control nets were raised up to 40 m above ground level and immediately retrieved during the launch and retrieval operations; between September and November 2014, the control nets were raised to 120 m above ground level. The control nets spent 5 min in the air, or up to 10 min when raised to 120 m. Once retrieved, the control nets were processed in a manner identical to that of other nets. After retrieving the panels, inspection for insects was conducted between 09:00 and 11:30 in a dedicated clean area. The panel was stretched between two posts and scanned for mosquitoes, which were counted, removed using forceps and preserved in 80% ethanol before all other insects were similarly processed and placed in other tubes. Depending on their condition, the sticky panels were sometimes reused the subsequent night.

Species identification. Glue attached to the insects was washed off with 100% chloroform. The mosquitoes were gently agitated (for less than 30 s) to loosen them from one another. Individual mosquitoes were transferred into consecutive wells filled with 85% ethanol. Using a dissecting scope, the samples were morphologically sorted by mosquito subfamily (Anophelinae or Culicinae), and were tentatively identified to the *Anopheles* species or species-group level. All mosquitoes that were morphologically confirmed as *A. gambiae* sensu lato (and two insects that were identified on the basis of molecular barcode analysis) were identified to the species level on the basis of fragment-size differentiation after amplification of the nuclear ITS2 region and digestion of the product[35]. Validation was carried out in LSTM (laboratory of D.W.), in which each specimen was washed with 500 μl heptane followed by two further washes with ethanol. DNA was then extracted using the Nexttec (Biotechnologie) DNA isolation kit, according to the manufacturer's instructions. Species identification used a standard PCR method, including all primers[36], with products visualized on 2% agarose gel. Samples of *A. gambiae* sensu lato were further identified to species by short interspersed element insertion polymorphism[37]. In cases in which no species-specific bands were detected using the first method, an approximately 800-bp region of the mtDNA cytochrome oxidase I (COI) genes was amplified using the primers C1_J_2183 and TL2_N_3014[38]. PCR products were purified using the QIAquick PCR-Purification kit (Qiagen) and sequenced in both directions using the original PCR primers by MacroGen. Sequences were aligned using CodonCode Aligner (CodonCode) and compared to existing sequences in GenBank to identify species. All other *Anopheles* mosquitoes were identified by the retrospective correlation of DNA barcodes, with morphologically verified reference barcodes compiled by Walter Reed Biosystematics Unit and the Mosquito Barcoding Initiative in the laboratory of Y.-M.L. Head and thorax portions of all samples were separated from abdomens and used for DNA extraction using the Autogen automated DNA extraction protocol. mtDNA COI barcodes were amplified using the universal LCO1490 and HCO2198 barcoding primers[39], and amplified, cleaned and bi-directionally sequenced according to previously detailed conditions[40]. All DNA barcodes generated from this study are available under the project 'MALAN – windborne *Anopheles* migrants in Mali' on the Barcode of Life Database (www.boldsystems.org) and in GenBank under accession numbers MK585944–MK586043. *Plasmodium* infection status was tested for all available *Anopheles* on DNA extracts from the head and thorax portions ($n = 190$) and also from abdomens ($n = 156$) following previously described protocols[41–43]. Owing to the nature of the collections, all body parts were not available for each specimen, which accounts for any discrepancies in the numbers. Blood-meal identification was carried out following a published protocol[44].

Data analysis. Although aerial collections started in April 2012, protocol optimization and standardization took most of that year; the data included in the present analysis cover only the period March 2013–November 2015. Nights on which operations were interrupted by storms or strong winds (for example, the balloon was retrieved during darkness) were also excluded.

The total number of mosquitoes per panel represents 'panel density' of each species. Aerial density was estimated on the basis of the panel density of the species, and total air volume that passed through that net that night: that is, aerial density = panel density/volume of air sampled, and volume of air sampled = panel surface area × mean nightly wind speed × sampling duration. The panel surface area was 3 m^2. Wind-speed data were obtained from the atmospheric reanalyses of the global climate (ERA5). Hourly data were available at 31-km surface resolution, with multiple vertical levels that included ground, 2, 10, 32, 55, 85, 115, 180, 215, 255 and 300 m above ground level. Overnight records (18:00 through to 06:00) for the nearest grid centre were used to calculate the nightly direction and mean wind speed at each village (Siguima, Markabougou and Thierola). Dallowere, which is located 25 km south of Thierola, was included in the same grid cell as that of Thierola. The mean nightly wind speed at panel height was estimated on the basis of the nearest available altitude layer.

To evaluate clustering in mosquito panel density and the effects of season, panel height, year and locality, mixed linear models with either Poisson or negative-binomial error distributions were implemented by procGLIMMIX[45]. The clustering at the levels of the panel and night of sampling were evaluated as random effects, as was the case for the year of sampling and locality. These models accommodate counts as non-negative integer values. The ratio of the Pearson χ^2 to the degrees of freedom was used to assess the overall 'goodness of fit' of the model, with values of more than two indicating a poor fit. The significance of the scale

parameter that estimates k of the negative-binomial distribution was used to choose between Poisson and negative-binomial models. Sequential model fitting was used, starting with random factors before adding fixed effects. Lower Bayesian information criterion values, and the significance of the underlying factors, were also used to select the best-fitting model for each species.

The magnitude of migration of windborne migrants was expressed as the expected minimum number of migrants per species crossing an imaginary line of 1 km perpendicular to the wind direction at altitude. This commonly used measure of abundance assumes that the insects fly in a layer that is 1-km wide, and does not require knowledge of the distance or time the insects fly to or from the interception point[46–48]. We used the mean wind speed at altitude (4.8 m s^{-1}; Fig. 1c), and assumed that mosquitoes fly in a layer depth of 200 m between 50 and 250 m above ground level, which conservatively reflects the fact that mosquitoes were captured between 40 and 290 m above ground level (Fig. 1a and Extended Data Fig. 2b). Accordingly, this nightly migration intensity was computed as the product of the mean aerial density across the year (conservatively including periods during which no migrants were captured) by the volume of air passing over the reference line during the night. The corresponding annual index was estimated by multiplying the nightly index by the period of migration of windborne migrants, estimated from the difference between the first and last day and month that a species was captured over the three years. Species that were captured once were assumed to migrate during a single month. The annual number of migrants per species that cross a 100-km line was used because of the similar composition of species across our sampling sites, which spanned 100 km (Extended Data Fig. 1a).

Similar to most insects in their size range[47,49,50], the flight speed of mosquitoes[51,52] does not typically exceed 1 m s^{-1}. Because winds at panel altitude attain speeds considerably higher than the speed of the mosquito, flight direction and speed are governed by the wind[46,47] and flight trajectories can therefore be simulated on the basis of the prevailing winds during the night of capture at the relevant locations and altitudes, as has previously been done[53–55]. Accordingly, backward trajectories of mosquito flights were simulated using HYSPLIT[25] on the basis of ERA5 meteorological reanalysis data. Data that are available in ERA5 present the highest spatial and temporal resolution available for this region. Comparisons with data of lower spatial and temporal resolution that are available from the MERRA2 reanalysis data[56], and with the Global Data Assimilation System (available at 0.5° spatial resolution), showed good agreement in trajectory direction and overall distance (data not shown). Trajectories of each captured mosquito were simulated starting at its capture location and altitude, and all multiple interception (full) hours during the night of the collection. Because anopheline mosquitoes are nocturnal, we conservatively assumed that flights started at or after 18:00 and ended by 06:00 the following morning, and computed trajectories for every hour that allowed for a total of a 2-h or 9-h flight. For example, to complete a 9-h flight by 06:00, a mosquito could have started at 18:00, 19:00, 20:00 or 21:00. The total flight duration of tethered female *A. gambiae* sensu lato and *Anopheles atroparvus* reached or exceeded 10 h, with an average speed[51] of 1 km h^{-1} that is consistent with other studies[52,57,58]. Similarly, *Anopheles vagus* and *Anopheles hyrcanus*, caught 150 m above ground level after midnight over India, would have been migrating for more than 6 h, assuming they took off around dusk[20]. Thus, we conservatively assumed that anopheline mosquitoes at high altitude fly between 2 and 9 h per night, although longer durations are possible. Each trajectory consisted of the global positions of the mosquitoes at hourly intervals from the interception time. In addition to plotting trajectories[59–66], the linear distance from the interception site and the simulated position of the mosquito, and the azimuth (angle between the interception site and the simulated position of the mosquito from the north, projected on a plane) were computed for all trajectories. To evaluate the distance range and dominant directions of flight, the mean and 95% confidence interval of the distance and azimuth (as a circular statistic) were computed for the 2-h and 9-h flight trajectories. The dispersion of individual angles (azimuths) around the mean was measured by the mean circular resultant length r, which can vary from 0 to 1, with higher values indicating tighter clustering around the mean. Rayleigh's test was used to test whether there was a mean direction; no mean direction occurs when the angles form a uniform distribution over a circle[67].

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Data on anopheline capture, identification, sex and gonotrophic status are available from www.boldsystems.org (project code MALAN) and in GenBank (MK585944–MK586043, inclusive).

Code availability

SAS code used for statistical analyses (and data manipulations) and 9-h backward trajectories data for each mosquito-capture event (based on HYSPLIT) are available from the corresponding author upon request. The code for plotting trajectories is available at https://github.com/benkraj/anopheles-migration.

31. Lehmann, T. et al. Tracing the origin of the early wet-season *Anopheles coluzzii* in the Sahel. *Evol. Appl.* **10**, 704–717 (2017).
32. Lehmann, T. et al. Seasonal variation in spatial distributions of *Anopheles gambiae* in a Sahelian village: evidence for aestivation. *J. Med. Entomol.* **51**, 27–38 (2014).
33. Huestis, D. L. et al. Seasonal variation in metabolic rate, flight activity and body size of *Anopheles gambiae* in the Sahel. *J. Exp. Biol.* **215**, 2013–2021 (2012).
34. Fritz, G. N., Fritz, A. H. & Vander Meer, R. K. Sampling high-altitude and stratified mating flights of red imported fire ant. *J. Med. Entomol.* **48**, 508–512 (2011).
35. Fanello, C., Santolamazza, F. & della Torre, A. Simultaneous identification of species and molecular forms of the *Anopheles gambiae* complex by PCR-RFLP. *Med. Vet. Entomol.* **16**, 461–464 (2002).
36. Scott, J. A., Brogdon, W. G. & Collins, F. H. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. *Am. J. Trop. Med. Hyg.* **49**, 520–529 (1993).
37. Santolamazza, F. et al. Insertion polymorphisms of SINE200 retrotransposons within speciation islands of *Anopheles gambiae* molecular forms. *Malar. J.* **7**, 163 (2008).
38. Simon, C. et al. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* **87**, 651–701 (1994).
39. 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.* **3**, 294–299 (1994).
40. Linton, Y.-M. et al. Mosquitoes of eastern Amazonian Ecuador: biodiversity, bionomics and barcodes. *Mem. Inst. Oswaldo Cruz* **108**, 100–109 (2013).
41. Bass, C. et al. PCR-based detection of *Plasmodium* in *Anopheles* mosquitoes: a comparison of a new high-throughput assay with existing methods. *Malar. J.* **7**, 177 (2008).
42. Demas, A. et al. Applied genomics: data mining reveals species-specific malaria diagnostic targets more sensitive than 18S rRNA. *J. Clin. Microbiol.* **49**, 2411–2418 (2011).
43. Steenkeste, N. et al. Towards high-throughput molecular detection of *Plasmodium*: new approaches and molecular markers. *Malar. J.* **8**, 86 (2009).
44. Kent, R. J. & Norris, D. E. Identification of mammalian blood meals in mosquitoes by a multiplexed polymerase chain reaction targeting cytochrome *b*. *Am. J. Trop. Med. Hyg.* **73**, 336–342 (2005).
45. SAS. SAS for Windows version 9.3 https://www.sas.com/ (2011).
46. Hu, G. et al. Mass seasonal bioflows of high-flying insect migrants. *Science* **354**, 1584–1587 (2016).
47. Drake, V. A. & Reynolds, D. R. *Radar Entomology: Observing Insect Flight and Migration* (CABI International, 2012).
48. Reynolds, D., Chapman, J. & Stewart, A. Windborne migration of Auchenorrhyncha (Hemiptera) over Britain. *Eur. J. Entomol.* **114**, 554–564 (2017).
49. Taylor, L. R. Insect migration, flight periodicity and the boundary layer. *J. Anim. Ecol.* **43**, 225–238 (1974).
50. Chapman, J. W., Drake, V. A. & Reynolds, D. R. Recent insights from radar studies of insect flight. *Annu. Rev. Entomol.* **56**, 337–356 (2011).
51. Kaufmann, C. & Briegel, H. Flight performance of the malaria vectors *Anopheles gambiae* and *Anopheles atroparvus*. *J. Vector Ecol.* **29**, 140–153 (2004).
52. Snow, W. F. Field estimates of the flight speed of some West African mosquitoes. *Ann. Trop. Med. Parasitol.* **74**, 239–242 (1980).
53. Eagles, D., Walker, P. J., Zalucki, M. P. & Durr, P. A. Modelling spatio-temporal patterns of long-distance *Culicoides* dispersal into northern Australia. *Prev. Vet. Med.* **110**, 312–322 (2013).
54. Stefanescu, C., Alarcón, M. & Àvila, A. Migration of the painted lady butterfly, *Vanessa cardui*, to north-eastern Spain is aided by African wind currents. *J. Anim. Ecol.* **76**, 888–898 (2007).
55. Klausner, Z., Fattal, E. & Klement, E. Using synoptic systems' typical wind trajectories for the analysis of potential atmospheric long-distance dispersal of lumpy skin disease virus. *Transbound. Emerg. Dis.* **64**, 398–410 (2017).
56. Gelaro, R. et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
57. Pedgley, D. E. *Windborne Pests and Diseases: Meteorology of Airborne Organisms* (Ellis Horwood, 1982).
58. Gillies, M. T. & Wilkes, T. J. Field experiments with a wind tunnel on the flight speed of some west African mosquitoes (Diptera: Culicidae). *Bull. Entomol. Res.* **71**, 65 (1981).
59. Kahle, D. & Wickham, H. ggmap: spatial visualization with ggplot2. *R J.* **5**, 144–161 (2013).
60. Hijmans, R. J. geosphere: spherical trigonometry. R package version 1.5-10 https://cran.r-project.org/package=geosphere (2017).
61. Slowikowski, K. ggrepel: automatically position non-overlapping text labels with 'ggplot2'. R package version 0.8.1 https://cran.r-project.org/package=ggrepel (2018).
62. Santos Baquero, O. ggsn: north symbols and scale bars for maps created with 'ggplot2' or 'ggmap'. R package version 0.5.2 https://cran.r-project.org/package=ggsn (2019).
63. Arnold, J. B. ggthemes: extra themes, scales and geoms for 'ggplot2'. R package version 4.2.0 https://cran.r-project.org/package=ggthemes (2019).
64. Grolemund, G. & Wickham, H. Dates and times made easy with lubridate. *J. Stat. Softw.* **40**, 1–25 (2011).

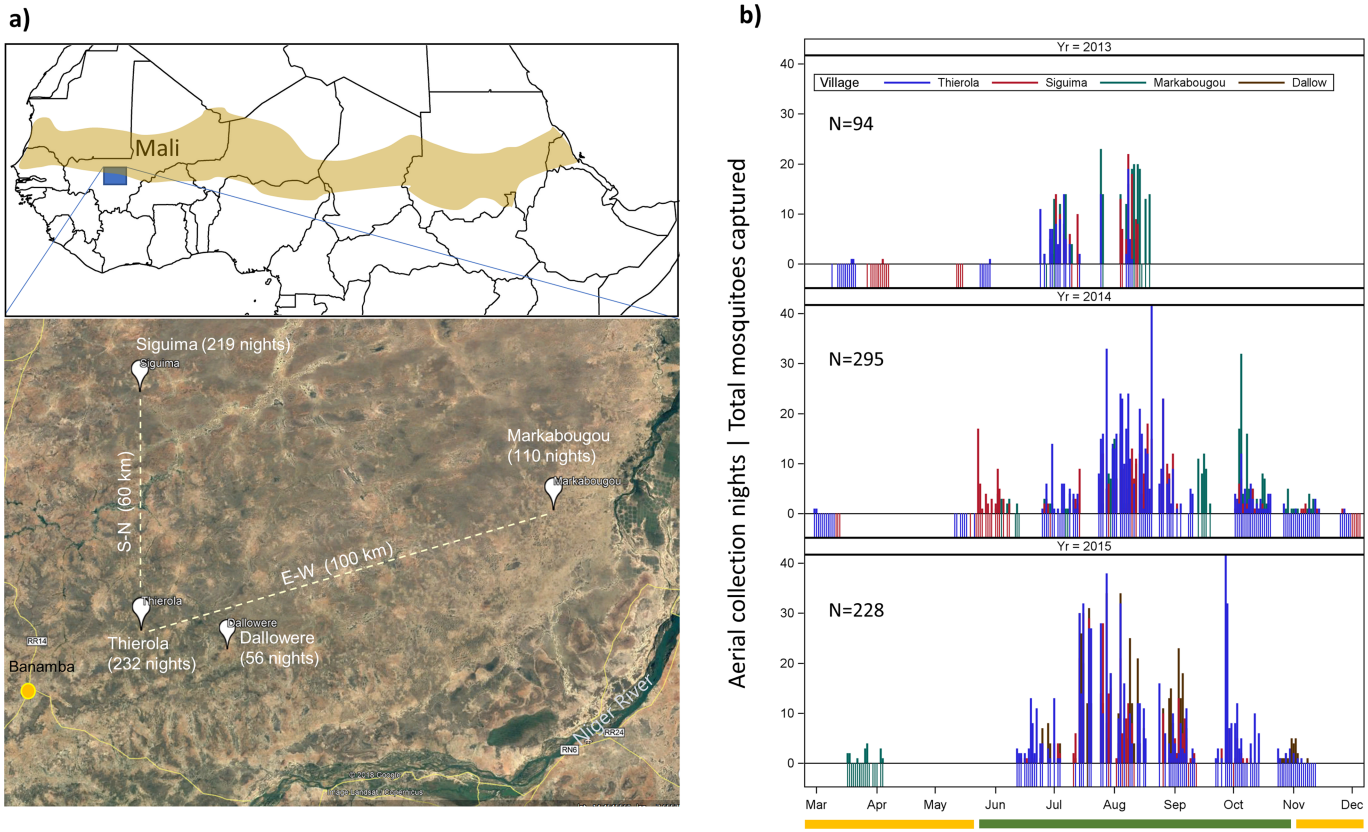
65. R Studio Team. RStudio: integrated development environment for R. R package version 3.6.0 https://www.rstudio.com/ (2015).
66. R Core Team. R: a language and environment for statistical computing. https://www.R-project.org/ (2016).
67. Fisher, N. I. *Statistical Analysis of Circular Data* (Cambridge Univ. Press, 1993).
68. Wickham, H. et al. ggplot2: create elegant data visualisations using the grammar of graphics. R package version 3.2.1 https://cran.r-project.org/web/packages/ggplot2/index.html (2019).

Acknowledgements We thank the residents of Thierola, Siguima, Markabougou and Dallowere for their consent to work near their homes, and for their wonderful assistance and hospitality; M. Keita, B. Coulibaly and O. Kone for their valuable technical assistance with field and laboratory operations; G. Fritz for consultation on the aerial sampling method using sticky panels; D. Sakai, S. F. Traore, J. Anderson, T. Wellems, M. Sullivan and S. Moretz for logistical support, F. Collins and N. Lobo for support to initiate the aerial-sampling project; J. M. C. Ribeiro and A. Molina-Cruz for reading earlier versions of this manuscript and providing us with helpful suggestions; and A. Crawford and F. (F.) Ngan for conversions of the MERRA2 and ERA5 data files to HYSPLIT format. This study was primarily supported the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health. Rothamsted Research received grant-aided support from the UK Biotechnology and Biological Sciences Research Council (BBSRC). Y.-M.L. and R.M. are supported by the US Army. Views expressed here are those of the authors, and in no way reflect the opinions of the US Army or the US Department of Defense. The USDA is an equal opportunity provider and employer. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

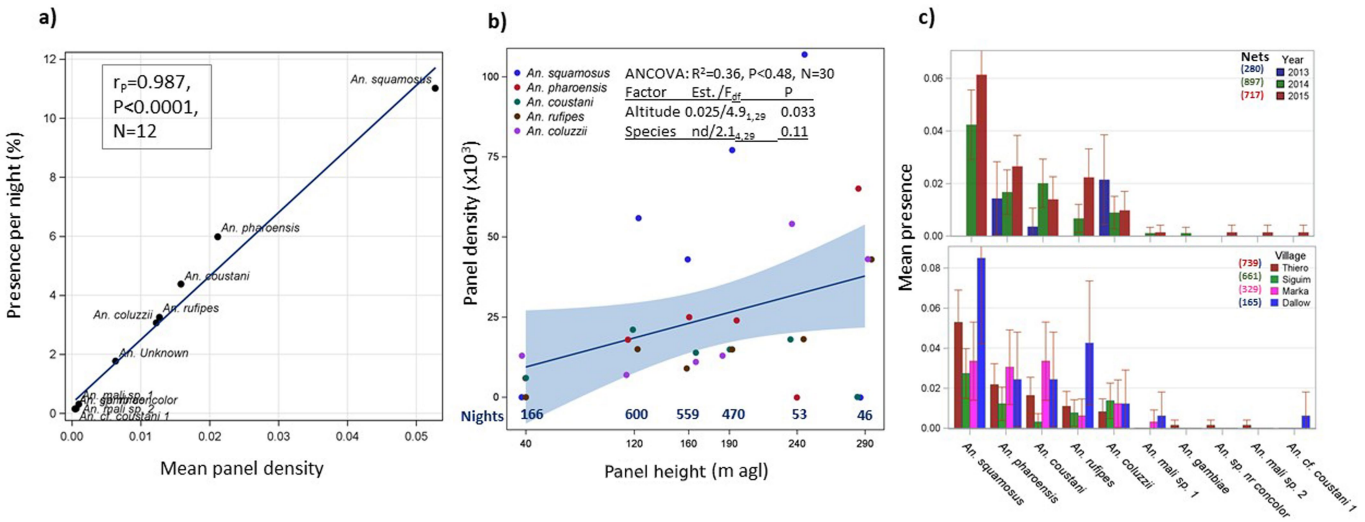
Author contributions The project was conceived by T.L. and D.L.H. Field methods and operations were designed by D.L.H. with input from D.R.R. and J.W.C. Fieldwork, protocol optimization, data acquisition and management, and initial processing of specimens, including tentative species identification, was performed by A.D., A.S.Y., M.D., D.S., Z.L.S. and Y.O. and subsequent processing by A.K., J.F. and L.V. with inputs from E.T. and L.C. Species identification and molecular analysis of specimens were conducted primarily by Y.-M.L., R.M., A.K. and B.J.K. with contributions by D.W., R.F. and M.J.D. Data analysis and HYSPLIT simulations were carried out by T.L. with inputs from all authors, especially R.F., B.J.K., D.R.R., J.W.C., E.S. and Y.-M.L. B.J.K. mapped simulated trajectories. The manuscript was drafted by T.L. and revised by all authors. Throughout the project, all authors have contributed key ideas that have shaped the work and the final paper.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1622-4.
Correspondence and requests for materials should be addressed to T.L.
Peer review information *Nature* thanks Nora Besansky, Simon Hay and Daniel Neafsay for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



Extended Data Fig. 1 | Study area and aerial sampling effort. a, Map of the study area, showing aerial-sampling villages, as well as the number of sampling nights per village. Schematic map of Africa, showing the Sahel region. The base map was generated using the ggplot2 package in R[68], under a GPL-2 license. **b**, Nightly sampling effort by year. The extension of the axes under zero indicates the sampling nights (by village), and the needles denote the total number of mosquitoes per night (regardless of the number of collecting panels per night). Dry and wet seasons are indicated by yellow and green, respectively, in the key under the *x* axis.



Extended Data Fig. 2 | Regularity of migratory flights, flight altitude and variability among years and localities in the aerial presence of species. a, Relationship between mosquito presence (fraction of positive nights) and the mean density of mosquitoes on collecting panels, to evaluate whether appearance can be accounted by overall abundance rather than by unique migratory nights. **b**, The relationship between the height of the collecting panel and mean density of mosquitoes per panel ($\times 10^3$; the regression line with shading denotes the 95% confidence interval of the mean), showing the mean density of mosquitoes per panel by species. The inset summarizes the covariance analysis that underlies this regression, which includes the species and height of the collecting panel. The number of nights per collecting-panel height is given in blue along the *x* axis (Fig. 1a). agl, above ground level. **c**, Variation in mosquito presence (fraction of positive nights) by species between years (top) and villages (bottom), with their 95% confidence interval. Sampling effort, expressed as the number of collecting panels per year or village, is shown adjacent to the key.

Extended Data Fig. 3 | Photograph showing a tethered sticky-panel setup and attachment. A sticky panel (3 × 1-m net) on a test helium balloon (of a lower volume and capacity), showing the attachment of the net covered with glue to the cord that tethers the balloon to the ground. The four carbon poles and Velcro attachment points are shown. A close-up image of the attachment of the panel to the cord, and an image of preparations to launch a standard 3-m balloon, are also shown.

Extended Data Table 1 | Variation in the rate of mosquito capture between years, localities and heights above ground

Dependent: Panel Density	Parameter	*A. squamosus*	*A. pharoensis*	*A. coustani*	*A. rufipes*	*A. coluzzii*
Random vars only: Poisson	Pearson χ2/df (BIC)	1.13 (793.5)	**1.04 (394.4)**	**0.90 (306.52)**	**1.11 (260.4)**	**1.16 (252.8)**
Random vars only: Negative Binomial	Pearson χ^2/df, Scale[a] (BIC)	**0.83, 5.98*** (756.2)**	0.97, 3.84^{ne} (391.4)	0.87, 2.09^{ns} (306.7)	0.99, 10.6^{ne} (254.5)	0.98, 15^{ns} (246.7)
	intercept[mean] (SD)	-4.06^{ns} (1.23)	-3.9** (0.226)	-4.4* (0.63)	-4.7*** (0)	-4.4** (0.23)
	Year (SD)	3.24^{ns} (4.36)	0^{ns} (0.06)	0.09^{ns} (0.31)	0.55^{ns} (0.56)	0^{ne}
	Locality[b] (SD)	0.075^{ns} (0.116)	0.04^{ns} (0.15)	0.73^{ns} (3.19)	0^{ne}	0^{ne}
Random vars only: Poisson	Night[c] (SD)	**4.02** (1.42)**	**1.78* (0.99)**	6.57^{ns} (7.3)	**29.0* (16.8)**	**32.0* (17.9)**
Random vars only: Neg. Bin.	Night[c] (SD), scale	3.9** (1.5), 0.74^{ns}	1.6^{ns} (1.1), 0.34^{ns}	0.5^{ne} (ne), 0^{ne}	30.1* (17.5), 0.7^{ns}	33.5* (18.7), 0.76^{ns}
Fixed and random: Poisson	Pearson χ^2/df (BIC)	0.37 (700)	0.6 (403)	0.2 (308)	0.09 (258)	0.08 (243)
	Night	1.4** (0.0)	0.78^{ns} (0.8)	1.9* (1.1)	14.0^{ns}(13.3)	21.9^{ns} (15.2)
	Period[d]	Aug-Oct*	Aug-Oct*	Aug-Oct^{ns}	Aug-Oct^{ns}	Aug-Oct***
	Panel height (m)	0.001***(0)	0.003***(0)	-0.007***(0)	0.001***(0)	0.014*(0.006)
Dependent: Aerial Density	Pearson χ^2/df (BIC)	0.42 (938)	0.41 (503)	0.2 (378)	0.1 (304)	0.09 (283)
Fixed and random: Poisson	Night	2.9*** (0.8)	2.6* (1.2)	5.2^{ns} (3.9)	26.8*(16.0)	31.5* (17.6)
	Period[d]	Aug-Oct^{ns}	Aug-Oct*	Aug-Oct^{ns}	Aug-Oct^{ns}	Aug-Oct***
	Panel height (m)	-0.003***(0)	-0.002***(0)	-0.008*(0.004)	-0.001***(0)	0.01*(0.005)

GLIMMIX models of random and fixed variables; total number of panels was 1,894. $^{***}P = 0.001$, $^{**}P = 0.01$, $^{*}P = 0.05$, NS, not significant ($P > 0.05$); NE, parameter could not be estimated.
[a]The negative-binomial-scale parameter estimates the *k* parameter of this distribution.
[b]When estimating the effect of locality (Extended Data Fig. 1a), only three locations were considered (after pooling Dallowere and Thierola, which are only 20 km apart (Methods)).
[c]The significance of clustering by night (across locations) estimated as the only random effect (using subject statement) after finding insignificant variance components of year and location.
[d]Periods were March to May, June to July, August to October and November to December. The period of the highest panel density is shown, with its statistical significance.
[e]The heights of the panels were 40, 120 (90–120), 160, 190 and 250 (220–290) m above ground level; some heights were combined into ranges, owing to small sample sizes (nights) for particular altitudes.

nature research

Corresponding author(s): Tovi Lehmann

Last updated by author(s): 2019/07/15

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for primary data collection. ERA5 weather data at altitude were used as part of the data interpretation as detailed in the Methods.
Data analysis	Data manipulation and analysis was carried out primarily by SAS 9.4. Flight trajectory simulations were carried out by HYSPLIT 4 using ERA5 weather data. ERA5 weather data were downloaded from the Climate Data Store (https://cds.climate.copernicus.eu/) using Python 3.6.0 scripts, and analyzed with R 3.5.2 (code will be available soon at https://github.com/benkraj/anopheles-migration). Additional R packages, tidyverse, lubridate, ggmap, ggsn, ggthemes, ggrepel, and geosphere were also used in plotting and analysis. All codes used in the analyses relevant to this paper are available as explained in the Data and Code Availability.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data availability statement is included in out paper: "Data on anopheline capture, identification, sex, and gonotrophic status are available from www.boldsystems.org (Project code: MALAN). Mithochondrial DNA barcode sequences are depsosited in MALAN and in Genbank: MK585944–MK586043). Additionally, 9-hour backward trajectories data for each mosquito capture event based on HYSPLIT are available from TL upon request." These data will be available to the public upon publication of this paper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Aerial sampling of insects between 40 and 290 m using nets attached to a stationary helium balloons was performed over 3 years (2013-2015) in four Sahelian villages of Mali. Anopheles mosquitoes collected were identified using morphological keys and molecular barcode analyses.
Research sample	Each balloon typically carried three nets. Initially, they were suspended at 40, 120, and 160 m agl, but from August 2013, the typical altitude was set to 90, 120, 190 m agl. When the larger balloon was deployed in the Thierola station (August–September 2015), two additional nets were added at 240 and 290 m agl. Nets tethered to the balloons were launched approximately 1 hour before sunset (~17:00) and retrieved 1 hour after sunrise (~07:30), the following morning. Each panel was stretched between two posts and scanned for mosquitoes, which were counted, removed using forceps, and preserved in 80% ethanol before all other insects were similarly processed and placed in other tubes. To control for insects trapped near the ground as the nets were raised and lowered, control nets were raised up to 40 m agl and immediately retrieved (between September and November 2014 the control nets were raised to 120 m agl) during the launch and retrieval operations. The control nets spent 5 minutes in the air (up to 10 minutes when raised to 120 m). Once retrieved they were processed as other nets.
Sampling strategy	We aimed to sample high altitude flying mosquitoes in the Sahel throughout the year for three years to assess regularity of high altitude flight patterns of the species involved. Four sites separated by 25-100 km were compared to assess similarity between locations. Field operations lasted typically two weeks/month (except January and February) to ensure successful sampling for at least 8 nights/month while deferring operation during storms or very windy nights due to major damage to equipment and supplies. However the availability of helium and security concerns, which were outside of our control has affected our operations (e.g., no helium was available for 5 months after August 2013 (despite early and full payment).
Data collection	Trained experienced field entomologists supervised, in person the field work in each sampling village. Data, consisting of anopheline and culicine mosquitoes per net per night were recorded on paper and transfered to excel sheets asap by the same team members.
Timing and spatial scale	Operations in different villages varied based on availability of helium and security concerns. In Thierola (13.6586, -7.2147) from March 2013 to November 2015, Siguima (14.1676, -7.2279) from March 2013 to October 2015; Markabougou (13.9144, -6.3438) from June 2013 to April 2015; and Dallowere (13.6158, -7.0369) from July 2015 to November 2015.
Data exclusions	Aerial sampling started in March 2012. However, protocol optimization took most of this year. Therefore, only data collected after standard operations protocols have been formalized were used. Additionally, sampling nights with strong winds and storms have been excluded because either balloon launch or retrieval were affected.
Reproducibility	To ensure reproducibility we only used data obtained after protocol optimization, ie., since March 2013 (ie., we excluded data collected in 2012 see above). To control for insects trapped near the ground as the nets were raised and lowered, control nets were raised up to 40 m agl and immediately retrieved (between September and November 2014 the control nets were raised to 120 m agl) during the launch and retrieval operations. The control nets spent 5 minutes in the air (up to 10 minutes when raised to 120 m). Once retrieved they were processed as other nets. Slight changes in panel altitude were introduced after noticing a trend of higher mosquito catch with altitude as described in Methods.
Randomization	NA. All data were used (after exclusion of data prior to 2013 and during storms and strong winds, as noted above)
Blinding	Blinding was not intentionally done. However, ultimate mosquito identification using molecular barcode analysis was carried out by collaborators not involved in the field operations.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Field operations were carried out in the Sahel of Mali. The region is rural, characterized by scattered villages with traditional mud-brick houses, surrounded by fields. A single growing season (June–October) allows the farming of millet, sorghum, maize, and peanuts, as well as subsistence vegetable gardens. Over 90% of the annual rains fall during this season (~550mm). Cattle, sheep, and goats graze in the savannah that consists of grasses, shrubs, and scattered trees. The rains form small puddles and larger seasonal ponds that usually are totally dry by the end of November. From November until May, rainfall is absent or negligible (total precipitation < 50mm), and by December through most of May, water is available only in deep wells. Temperatures peak (night high >30C) during the late dry season (March-May) and reach their minima (night near 15-20C) during the early dry season (December-February). Troughout the dry season (December-May) relative humidity is low (25-15%) with few short exceptions, whereas during the wet season (June-November) it is higher (60-90%) based on the monsoon rains. In the

	rest of the year, the temperatures are intermediate and depend on monsoon rains. No electricity or water pipes are available in the villages near our aerial sampling stations.
Location	Sampling stations were in Thierola (13.6586, -7.2147), Siguima (14.1676, -7.2279), Markabougou (13.9144, -6.3438), and Dallowere (13.6158, -7.0369).
Access and import/export	All the work was coordinated and carried out by the Malaria Research and Training Center of Mali (MRTCH, now the Mali ICEMR). Specimens were exchanges based on existing agreements for scientific cooperation between MRTC, NIH and participating collaborators. Dead and preserved insects were transfered between countries following existing procedures.
Disturbance	We believe that aerial sampling using small nets had minimal effect on diversity and biological systems being sampled given the fraction of the space actually sampled.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The study did not involve laboratory animals
Wild animals	Only anopheles mosquitoes are included in this paper. Other insects were also collected and be summarized separaetly.
Field-collected samples	Almost all insects collected by sticky nets were preserved immediately in 80% ethanol. They were kept in field conditions for several days before they were moved to a laboratory freezer in Bamako, Mali.
Ethics oversight	No ethical approval is required for sampling aerial flying mosquitoes.

Note that full information on the approval of the study protocol must also be provided in the manuscript.