Signatures of aetivation and migration in Sahelian malaria mosquito populations

A.Dao\textsuperscript{1}, A.S.Yaro\textsuperscript{1}, M.Diallo\textsuperscript{1}, S.Timbiné\textsuperscript{1}, D.L.Huestis\textsuperscript{2}, Y.Kassogué\textsuperscript{1}, A.I.Traoré\textsuperscript{1}, Z.L.Sanogò\textsuperscript{1}, D.Samaké\textsuperscript{1} & T.Lehmann\textsuperscript{2}

During the long Sahelian dry season, mosquito vectors of malaria are expected to perish when no larval sites are available; yet, days after the first rains, mosquitoes reappear in large numbers. How these vectors persist over the 3–6-month long dry season has not been resolved, despite extensive research for over a century\textsuperscript{1–12}. Hypotheses for vector persistence include dry-season diapause (aetivation) and long-distance migration (LDM); both are facets of vector biology that have been highly controversial owing to lack of concrete evidence. Here we show that certain species persist by a form of aetivation, while others engage in LDM. Using time-series analyses, the seasonal cycles of Anopheles coluzzii, Anopheles gambiae sensu stricto (s.s.), and Anopheles arabiensis were estimated, and their effects were found to be significant, stable and highly species-specific. Contrary to all expectations, the most complex dynamics occurred during the dry season, when the density of A. coluzzii fluctuated markedly, peaking when migration would seem highly unlikely, whereas A. gambiae s.s. was undetected. The population growth of A. coluzzii followed the first rains closely, consistent with aetivation, whereas the growth phase of both A. gambiae s.s. and A. arabiensis lagged by two months. Such a delay is incompatible with local persistence, but fits LDM. Surviving the long dry season \textit{in situ} allows A. coluzzii to predominate and form the primary force of malaria transmission. Our results reveal profound ecological divergence between A. coluzzii and A. gambiae s.s., whose standing as distinct species has been challenged, and suggest that climate is one of the selective pressures that led to their speciation. Incorporating vector dormancy and LDM is key to predicting shifts in the range of malaria due to global climate change\textsuperscript{3}, and to the elimination of malaria from Africa.

Over half a million malarial deaths still occur annually, mostly in sub-Saharan Africa\textsuperscript{4}. Transmitted by Anopheles gambiae s.s., A. coluzzii (previously known as the A. gambiae S and M molecular forms\textsuperscript{5}), A. arabiensis and A. funestus, malaria is widespread, including in dry savannas and semi-arid areas. Persistence of malaria in areas where the surface waters required for larval development are absent for several months a year\textsuperscript{6–12} has been the subject of much interest, as it has long been recognized that, during the dry season, reproductively quiescent adult mosquitoes are ideally suited for vector control\textsuperscript{13–15}. Recent findings suggested that aetivation is used by A. coluzzii to persist throughout the dry season\textsuperscript{16–18}; yet, more definitive evidence is required to fully resolve this question.

Data from a five-year study of Sahelian A. coluzzii, A. gambiae s.s. and A. arabiensis population densities at an unparalleled resolution were subjected to time-series analyses to isolate the seasonal components, assess their magnitude, and determine if they were stable or time-varying (Methods). This statistical framework allowed identification of salient elements of the seasonal cycle of each species, providing unique ecological signatures, which were then deciphered to determine if populations endured the dry season locally or if populations recolonized the area by migration.

From September 2008 to August 2013, a total of 40,195 A. gambiae sensu lato (s.l.) (28,547 females and 11,648 males) were collected in the Sahelian village of Thierola, Mali, during 511 collection days (Fig. 1 and Extended Data Fig. 1; Extended Data Table 1, Supplementary Information). The complexity of the population dynamics of A. gambiae s.l. was epitomized by dramatic fluctuations during the dry season (Extended Data Figs 2 and 3). Putative seasonal elements were visually identified (Methods and Extended Data Table 2), providing a descriptive framework and expectations, to aid the interpretation of the statistical results. Briefly, the population growth phase (June–August) started \textasciitilde 3 weeks after the first rain, resulting in the wet-season peak (September–October). Density declined as larval sites dried (November), reaching its dry-season minima in February–March. Surprisingly, density started rising halfway into the dry season (March) and culminated in a dramatic dry-season peak lasting \textasciitilde 7 days, returning to the typical low density weeks later (April–May), and ending with the first rain surge, 3–7 days after the first rains (Extended Data Fig. 3 and Extended Data Table 2).

\textsuperscript{1}International Center for Excellence in Research (ICER), University of Sciences, Techniques and Technologies, BP 1805, Bamako, Mali. \textsuperscript{2}Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland 20852, USA.
Table 1 | Unobserved component time-series (final) models of the population dynamics for each taxon (Methods)

| Taxon* | Parameter | Variance (stochastic)$^\dagger$ | P (variance)$^\ddagger$ | Deterministic effects estimated$^\ddagger$ | P (effect)$|$
|-------|-----------|--------------------------|-----------------|---------------------------------|-------|
| **A. gambiae s.l.** | Levels $| E$ var | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| R$^2$ = 0.76 | Levels $| E$ var | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| AIC = 793.5 | Cycle$| DampF$ | 0.7 | 0.0001 | 0.99 |
| BIC = 808.1 | Cycle$| period$ | 5,758.7 | 0.99 | 0.0001 |
| | Cycle$| E$ var | 0.56 | 0.0001 | 0.99 |
| | Irreg.$| E$ var | 0.0112 | 0.85 | NA | 0.95 |
| **A. coluzzii** | Seasons | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| R$^2$ = 0.72 | Levels $| E$ var | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| AIC = 844 | Cycle$| DampF$ | 0.7 | 0.0001 | 0.99 |
| BIC = 858 | Cycle$| period$ | 47.1 | 0.5 | 0.0001 |
| | Cycle$| E$ var | 0.65 | 0.0001 | 0.99 |
| | Irreg.$| E$ var | 0.011 | 0.92 | NA | 0.95 |
| **A. gambiae s.s.** | Seasons | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| R$^2$ = 0.89 | Levels $| E$ var | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| AIC = 515 | Cycle$| DampF$ | 0.76 | 0.0001 | 0.99 |
| BIC = 545 | Cycle$| period$ | 16.1 | 0.0001 | 0.99 |
| | Cycle$| E$ var | 0.14 | 0.0017 | 0.21 |
| | Cycle2$| DampF$ | 1 | 0.0001 | 0.21 |
| | Cycle2$| period$ | 41.1 | 0.0001 | 0.21 |
| | Cycle2$| E$ var | 0.0001 | 0.54 | 0.0003 |
| | Irreg.$| E$ var | 0.042 | 0.28 | 0.0003 |
| | Irreg.$| AR(1)$$^\#$ | 0.94 | 0.0001 | 0.99 |
| **A. arabiensis** | Seasons | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| R$^2$ = 0.77 | Levels $| E$ var | 0 (fixed) | NA | See Fig. 2 | 0.0001 |
| AIC = 742 | Cycle$| DampF$ | 0.69 | 0.0001 | 0.99 |
| BIC = 757 | Cycle$| period$ | 11,966 | 0.99 | 0.0001 |
| | Cycle$| E$ var | 0.48 | 0.0001 | 0.99 |
| | Irreg.$| E$ var | 0.00001 | 0.99 | NA | 0.99 |

*All models (species) include 362 observations (5-day means from 22 September 2008 and 1 September 2013, based on all A. gambiae s.l. and those genotyped, see Methods and Extended Data Fig. 1).

$^\dagger$Stochastic variance and test of significance (P variance) indicate whether the parameter is time varying.

$^\ddagger$Effect size and test of significance (P effect) measure the overall deterministic effects.

$^\#$Seasonal component was modelled by 73 dummy variables. Individual effect of each of these parameters and 95% CI are shown in Fig. 2 (see text and Methods).

Time-series analysis of the log-transformed density (Extended Data Fig. 2), using an unobserved components model (Methods), was fitted for A. gambiae s.l. (Table 1). The model selected had a fixed level (equivalent to intercept) and no slope (trend), reflecting a stable mosquito density over the study. An additional non-seasonal cycle with a long period was also included (Methods and Supplementary Information).

The variance of the seasonal component was insignificant, indicating it was not time-varying; thus, it was modelled as a fixed component, simplifying its interpretation. The seasonal component of A. gambiae s.l. population dynamics was highly significant ($P < 0.0001$, Table 1). The estimated seasonal variation (Fig. 2a) revealed a large gap between the 95% confidence intervals (CIs) of the wet-season peak and that of the mid-dry-season low; thus, these elements and the decline between them are statistically well-supported. Likewise, large gaps were found between the 95% CIs of the mid-dry-season low and the late-dry-season peak, between this peak, the end-dry-season low, and the following wet-season peak, indicating that these elements (and the transitional phases connecting them) were statistically supported. Other putative elements (Extended Data Table 2) had insufficient statistical support.

The putative elements of each species’ seasonal cycle were identified (Extended Data Table 2). The seasonal component of all species was fixed (its variance was insignificant) and was highly significant ($P < 0.0001$, Table 1). The time-series model selected for A. coluzzii was structurally similar to that of A. gambiae s.l. (Table 1). On the basis of their 95% CIs, one wet-season peak and two dry-season peaks, which were observed in all years (Fig. 1), were statistically supported (Fig. 2b). The early wet-season decline of A. coluzzii produced the pre-dry-season trough in mid-November, before the last larval site dried, which was followed by an early dry-season peak in late December (Fig. 2b). Subsequently, its seasonal component was virtually identical to that of A. gambiae s.l. (Fig. 2), consistent with its predominance in species composition (Extended Data Fig. 1c). The model for A. gambiae s.s. included two non-seasonal cycles as well as an autoregressive (lag 1) error (Table 1). Only a single peak (wet season) and a long dry-season trough were statistically discerned in A. gambiae s.s. (Fig. 2c). The model for A. arabiensis was structurally similar to that of A. coluzzii (Table 1). A single wet-season peak and the long dry-season trough were supported (Fig. 2d), whereas changes during the dry season were not distinguished from noise.

The species-specific signatures manifested by their population dynamics provide compelling evidence that A. coluzzii persists locally in the Sahel during the dry season, whereas A. gambiae s.s. recolonizes via LDM after the first rains; the evidence is less clear for A. arabiensis. First, A. coluzzii was present throughout the dry season (albeit in small numbers), whereas A. gambiae s.s. was undetected from January to May (Fig. 2 and Extended Data Fig. 4), consistent with previous studies$^{14,19-27}$. Second, the density of A. coluzzii rose dramatically (ten- to ninety-fold from their preceding phase) twice during the dry season (Fig. 2). Since these peaks preceded the first rain by at least six weeks, any potential migrant mosquitoes would likely perish before reproductive opportunities were available, given the absence of surface waters in the area (ruling out dry-season reproduction). Third, the most crucial evidence relates to the period when population growth starts with respect to the first rain. The onset of population growth can be defined as the first time when the lower 95% CI of the seasonal component is greater than the upper 95% CI during the preceding dry season’s low phase (red arrows, Fig. 2). This phase started in June for A. coluzzii but in August for A. gambiae s.s. and A. arabiensis (Fig. 2). A delay of six to eight weeks in the onset of population growth for the latter two species corroborated our previous results in two other Sahelian villages, 10–25 km away from Thierola$^{14}$. Commencing population growth shortly after the first rain fits well with...
During the dry season, no surface waters were available near Thierola for at least a 30 km radius, and in the distant localities where surface waters did exist, overall density was very low and *A. gambiae s.s.* was not detected until the wet season\(^1\). The nearest high-density source is in the Niono rice cultivation area (~150 km ENE of Thierola), but it consists exclusively of *A. coluzzii*. Therefore, LDM spanning hundreds of kilometres is necessary to explain the re-colonization of *A. gambiae s.s.* Alternative explanations, including desiccation-tolerant dormant eggs, larvae or pupae, as well as larval growth in deep, underground water sources, should not be altogether dismissed, despite being contradictory to available knowledge.

Population dynamics of *A. arabiensis* exhibited mixed signatures. Statistically, it is similar to *A. gambiae s.s.*, and the long delay in population growth after the rains (Figs 1, 2 and Extended Data Fig. 4) indicates that it too persists by LDM. Yet, throughout the dry season, sporadic individuals were found every year, as opposed to zero *A. gambiae s.s.* Possibly, the dominant strategy of *A. coluzzii* is expressed in a small fraction of *A. arabiensis*, consistent with previous reports of local persistence of *A. arabiensis* by aestivation in the East African Sahel\(^1\) and perhaps in other parts of the West African Sahel\(^10,27\). Alternatively, the occasional *A. coluzzii* recovered in the dry season could represent backcrossed hybrids between *A. coluzzii* and *A. arabiensis*.

These results provide fresh insights that dramatically change our understanding of the ecology of African malaria vectors and resolve the ‘dry-season paradox’. Ignoring the pervasive effects of dormancy and LDM limits our understanding of malaria transmission and its response to control and elimination strategies. Dormancy shapes vector composition in the Sahel, where *A. coluzzii* comprised 75% of the overall indoor vector density with its wet-season peak being at least twice as high and broad than that of either *A. gambiae s.s.* or *A. arabiensis* (Fig. 1, Supplementary Information and refs 14, 16). Arguably, dormancy underlies the heavy burden of malaria transmission in such areas by exponential amplification of human-vector cycles that culminate in intense late-wet-season transmission. Although *A. gambiae s.s.* and *A. arabiensis* predominate during the end of the wet season (October–November), we doubt that they alone can sustain the high rate of malaria transmission had *A. coluzzii* not amplified infections from June to September. Therefore, vector-control strategies that eliminate *A. coluzzii* alone may cut peak malaria transmission to very low levels. Targeting *A. coluzzii* while in its hidden shelters during the dry season is probably the most efficient control strategy, if these sites are found, but the indoor population during the early wet season and the late-dry-season peak also presents promising targets. Thus, a single residual spraying indoors in the late dry-season (for example, March) that is effective for 4 months may achieve dramatic reduction in malaria transmission in the following wet season\(^1\). Moreover, the spread of introduced genes by genetically modified mosquitoes may be hindered or aided by dormancy and LDM, as would other forms of malaria control and elimination campaigns.

Divergent strategies of persistence through the dry season were revealed by species-specific seasonal dynamics: local persistence of *A. coluzzii*, as opposed to annual recolonization by LDM for *A. gambiae s.s.* and *A. arabiensis*. They signify a multitude of behavioural, physiological and molecular divergence processes and thus probably represent the most striking phenotypic differences between the species found so far\(^28-29\), lending support for the elevated taxonomic status of the molecular forms to species\(^9\). Consistent with previous interpretations\(^30\), the adaptation to exploit arid environment such as the Sahel via aestivation may represent the central dimension in the adaptive divergence between the species. The implications of these differences for understanding speciation
and for explaining their geographical range\textsuperscript{11,12} are just beginning to be appreciated.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 July; accepted 22 October 2014.

Published online 26 November 2014.

1. Donnelly, M. J., Simard, F. & Lehmann, T. Evolutionary studies of malaria vectors. Trends Parasitol. \textbf{18}, 75–80 (2002).
2. Omer, S. M. \& Cloudsley-Thompson, J. L. Dry season biology of Anopheles gambiae Giles in the Sudan. \textit{Nature} \textbf{217}, 879–880 (1968).
3. Holstein, M. H. Biology of Anopheles gambiae. \textit{Research in French West Africa} (World Health Organization, 1954).
4. Siraj, A. S. \textit{et al.} Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. \textit{Science} \textbf{343}, 1154–1158 (2014).
5. World Health Organization. \textit{World Malaria Report 2013} (World Health Organization, 2013).
6. Coetzee, M. \textit{et al.} Anopheles coluzzii and Anopheles arnhamicus, new members of the Anopheles gambiae complex. \textit{Zootaxa} \textbf{3619}, 246–274 (2013).
7. Omer, S. M. \& Cloudsley-Thompson, J. L. Survival of female Anopheles gambiae Giles through a 9-month dry season in Sudan. \textit{Bull. World Health Organ.} \textbf{42}, 319–330 (1970).
8. Taylor, C. E., Toure, Y. T., Coluzzi, M. \& Petrarca, V. Effective population size and persistence of Anopheles arabiensis during the dry season in west Africa. \textit{Med. Vet. Entomol.} \textbf{7}, 351–357 (1993).
9. Toure, Y. T. \textit{et al.} Ecological genetic studies in the chromosomal form Mopti of Anopheles gambiae s.s. in Mali, West Africa. \textit{Genetica} \textbf{94}, 213–223 (1994).
10. 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. \textit{Insect Mol. Biol.} \textbf{9}, 467–479 (2000).
11. Coetzee, M., Craig, M. \& \& Sueur, D. Distribution of African malaria mosquitoes belonging to the Anopheles gambiae complex. \textit{Parasitol. Today} \textbf{16}, 74–77 (2000).
12. della Torre, A., Tu, Z. \& Petrarca, V. On the distribution and genetic differentiation of Anopheles gambiae s.s. molecular forms. \textit{Insect Biochem. Mol. Biol.} \textbf{35}, 755–769 (2005).
13. Sogoba, N. \textit{et al.} Monitoring of larval habitats and mosquito densities in the Sudan savanna of Mali: implications for malaria vector control. \textit{Am. J. Trop. Med. Hyg.} \textbf{77}, 82–88 (2007).
14. Adamou, A. \textit{et al.} The contribution of aestivating mosquitoes to the persistence of Anopheles gambiae in the Sahel. \textit{Malar. J.} \textbf{10}, 151 (2011).
15. Huey, D. L. \textit{et al.} Seasonal variation in metabolic rate, flight activity and body size of Anopheles gambiae in the Sahel. \textit{J. Exp. Biol.} \textbf{215}, 2013–2021 (2012).
16. Lehmann, T. \textit{et al.} Aestivation of the African malaria mosquito, Anopheles gambiae in the Sahel. \textit{Am. J. Trop. Med. Hyg.} \textbf{83}, 601–606 (2010).
17. Yaro, A. S. \textit{et al.} Dry season reproductive depression of Anopheles gambiae in the Sahel. \textit{J. Insect Physiol.} \textbf{58}, 1050–1059 (2012).
18. Lehmann, T. \textit{et al.} Seasonal variation in spatial distributions of Anopheles gambiae in a Sahelian village: evidence for aestivation. \textit{J. Med. Entomol.} \textbf{51}, 27–38 (2014).
19. Coluzzi, M., Petrarca, V. \& Di Deco, M. A. Chromosomal inversion intergradation and incipient speciation in Anopheles gambiae. \textit{Boll. Zool.} \textbf{52}, 45–63 (1985).
20. Coluzzi, M., Sabatini, A., Petrarca, V. \& Di Deco, M. A. Chromosomal differentiation and adaptation to human environments in the Anopheles gambiae complex. \textit{Trans. R. Soc. Trop. Med. Hyg.} \textbf{73}, 483–497 (1979).
21. Sogoba, N. \textit{et al.} Spatial distribution of the chromosomal forms of Anopheles gambiae in Mali. \textit{Malar. J.} \textbf{7}, 205 (2008).
22. Toure, Y. T. \textit{et al.} Perennial transmission of malaria by the Anopheles gambiae complex in a north Sudan Savanna area of Mali. \textit{Med. Vet. Entomol.} \textbf{10}, 197–199 (1996).
23. Denlinger, D. L. Dormancy in tropical insects. \textit{Annu. Rev. Entomol.} \textbf{31}, 239–264 (1986).
24. Denlinger, D. L. \& Armbruster, P. A. Mosquito diapause. \textit{Annu. Rev. Entomol.} \textbf{59}, 73–93 (2014).
25. Tauber, M. J., Tauber, C. A. \& Masaki, S. \textit{Seasonal Adaptations of Insects} (Oxford Univ. Press, 1986).
26. Sogoba, N. \textit{et al.} Malaria transmission dynamics in Niono, Mali: the effect of the irrigation systems. \textit{Acta Trop.} \textbf{101}, 232–240 (2007).
27. Lernassorn, J. J. \textit{et al.} Comparison of behavior and vector efficiency of Anopheles gambiae and \textit{An. arabiensis} (Diptera:Culicidae) in Barkedji, a Sahelian area of Senegal. \textit{J. Med. Entomol.} \textbf{34}, 396–403 (1997).
28. Huestis, D. L. \& Lehmann, T. \textit{Ecophysiology of Anopheles gambiae s.s.: persistence in the Sahel. Infect. Genet. Evol.} \texttt{http://dx.doi.org/10.1016/j.meegid.2014.05.027} (14 June 2014).
29. Lehmann, T. \& Diabate, A. The molecular forms of \textit{Anopheles gambiae}: a phenotypic perspective. \textit{Infect. Genet. Evol.} \textbf{8}, 737–746 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgments We thank the residents of Thierola for their hospitality and assistance with mosquito collections: J. Ribeiro, T. Wellens, P. McQueen, R. Faiman and G. Wasserberg for their comments on previous versions of this manuscript; and C. Traoré, R. Sakai, R. Gwadz and T. Wellems for logistical support. This study was supported by the Tamaki Foundation and by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Author Contributions T.L. conceived the study and together with A.D. and A.S.Y. assisted with mosquito collections; J. Ribeiro, T. Wellems, P. McQueen, R. Faiman and G. Wasserberg for their comments on previous versions of this manuscript; and C. Traoré, R. Sakai, R. Gwadz and T. Wellems for logistical support. This study was supported by the Tamaki Foundation and by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.L. (tlehmann@niaid.nih.gov).
METHODS
The study was performed between September 2008 and August 2013 in Thierola (13.6583° N, 7.2155° W), a small rural village in the Malian Sahel. The village populations, ethnico-ecological composition, agricultural activities and house structure were described previously18. During the wet season, the rains fill two large ponds and many small puddles near the village. The small puddles require frequent rains, as they dry within a week without additional rain. The last rain typically falls in October and usually all surface water dries by December. From November until May, rainfall is altogether absent or negligible (total precipitation < 30 mm). In the course of this study, dry-season ‘mango’ rains (<20 mm) fell in the area in March 2009, but no rain fell during the dry season in any of the subsequent years, at least over a 30 km radius. During the dry season, water is only available in four deep wells (~25 m deep). Seepage of water around wells and troughs for animals was monitored every day, but no mosquito larvae were found in these small puddles, which typically dry up every evening. A few trees may be irrigated by bucket every several days, but all water dries within hours. Annual precipitation is approximately 500 mm (543 mm in Segou, which lies 30 km south and 100 km east of Thierola). For this study, the dry season refers to December–May and the wet season to July–October; the transition periods (June and November) are marked by climatic irregularity (surface water may or may not be available). In this paper, a year is defined as the period spanning from the first rains (1 July) to the end of the following dry season (30 June). On-the-ground searches for surface waters during the dry season were conducted every year, in consultation with herders and hunters, and a detailed examination of the satellite photographs available in Google Earth was also performed. Tree holes containing water that last until January (and rarely into February) were also monitored, but no anopheline larvae were found. Except after the mango rains of March 200916, no surface waters have been found during the dry season in a distance up to 30 km around Thierola.

Mosquito collection. Live collections using mouth aspirators inside all houses (n = 120), were conducted throughout the study period as described previously18. The houses searched). Mean daily density per house was transformed to stabilize the variance (including female and male mosquitoes) in each day by the number of houses sampled (missing) value differed from the corresponding 10 days mean by more than 30% of that mean, it was replaced by the latter. If no 10-day mean density was available for that period, the same was carried out with the corresponding global 10-day mean density (across the five years). Less than 5% of values required such substitutions. The fit between the equidistant time series and the observed daily mean density is depicted in Extended Data Fig. 2. Statistical analyses were performed on the equidistant log-density time series.

Mosquitoes that were morphologically identified as members of A. gambiae s.l. complex were subjected to molecular identification to determine their species15. To compare adult and larval composition, larval collections using dippers were conducted between adult and larval composition, larval collections using dippers were conducted during the wet season of 2009 and 2010 from multiple larval sites and multiple po-
sitions in each site.

Data analysis. The Anopheles gambiae s.l. indoor collection records, consisting of 511 collection days in all (~120) accessible houses from September 2008 through August 2013, were used to produce mean daily density per house (dividing the total collected (including female and male mosquitoes) in each day by the number of houses sampled). Mean daily density per house was transformed to stabilize the variance into natural log density as follows: ln(density) = ln(mean density + 0.9)(no. of houses sampled). Although statistically equivalent to a transformation using 1.0, 0.9 was used to signify that the drop of density from 1 to zero, which probably reflects, biologically, “more” than a change from 1.1 to 0.1.

To provide a descriptive framework as part of the data exploration, putative elements of the annual (seasonal) cycle were visually identified if they appeared in two or more years, after scanning graphs using linear and logarithmic scales (Fig. 1, Ex-
tended Data Fig. 2 and Extended Data Table 1) using two 15-day-long frames that were shifted horizontally along the figures. The time series of daily density were converted into an equidistant, 5-day interval time series consisting of 362 intervals using Proc Expand30, after fitting a continuous curve to the data by connecting successive straight-line segments between non-missing input values using the ‘join’ method. This procedure interpolated missing values in the time series. The 5-day interval estimates produced showed the best fit to the observed data based on visual inspection and the sum of the difference between the observed and expected values, when compared with 10-day, 14-day, and 1 month intervals. Moreover, if the interpo-
lated (missing) value differed from the corresponding 10 days mean by more than 30% of that mean, it was replaced by the latter. If no 10-day mean density was available, the same was carried out with the corresponding global 10-day mean density (across the five years). Less than 5% of values required such substitutions. The fit between the equidistant time series and the observed daily mean density is depicted in Extended Data Fig. 2. Statistical analyses were performed on the equidistant log-density time series.

30. SAS (Sas Institute, Cary, North Carolina, 2011).
31. 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).
32. Harvey, A. C. Forecasting, Structural Time Series Models and the Kalman Filter (Cambridge Univ. Press, 1989).

©2014 Macmillan Publishers Limited. All rights reserved
Extended Data Figure 1 | Sex-ratio, density and composition of *Anopheles gambiae* s.l. a–c, Overall monthly means of the proportion of *A. gambiae* s.l. females (a), house density (b), and species composition (c). \( N_m, N_d \) and \( N_g \) denote sample size of *A. gambiae* s.l., the number of collection days, and the number genotyped to species, respectively (Methods). Whiskers in box-whisker plots extend to the extreme values up to 1.5 the distance between the twenty-fifth and seventy-fifth percentiles. In a, blue triangles represent means that are significantly lower than the red triangles (based on the sequential Bonferroni test; see Extended Data Table 1) and the horizontal line represents 1:1 sex ratio.
Extended Data Figure 2 | Population dynamics of *Anopheles gambiae* s.l.
House density over time in linear (top) and natural logarithm (bottom) scale to evaluate systematic change over time. Circles denote observed daily mean density and the black lines show the interpolated series of mean house density over 5-day intervals. Grey lines depict interpolation during the longest time without field samples (December 2008 to April 2009). First rain events are shown by green lines (dates listed above). Sample sizes of mosquitoes and collection days are the same as in Extended Data Fig. 1b.
Extended Data Figure 3 | Population dynamics of *Anopheles gambiae* s.l. across years. Observed daily mean density (circles) is shown against 5-day means (line) on linear and log scales from July to June of every year to assess similarity among years. Green arrows mark the first rain and the tan background denotes the dry season. Shading during 2008 indicates a gap in sampling (December–March) when imputed values were used (Methods). Sample sizes ($N_m$ and $N_d$) are explained in Fig. 1.
Extended Data Figure 4 | Level-adjusted seasonal component of the density of *A. coluzzii*, *A. gambiae s.s.* and *A. arabiensis*. The level, seasonals, and their 95% CI (bands) were estimated using unobserved component time series models (Table 1; Methods). Arrows denote the start of the population growth (text and Fig. 2). Time is shown starting from May to maximize the comparability between the species. In the colour ruler on the x axis, yellow and green denote dry and wet seasons, respectively, and orange and light-green denote transition periods. Sample sizes are given in Extended Data Fig. 1.
Extended Data Table 1 | Annual and monthly variation in the indoor sex ratio (proportion of females)

| Source    | df | F/Z  | P     |
|-----------|----|------|-------|
| Year      | -- | 0.34 | 0.37  |
| Month     | 11/33 | 5.35 | 0.0001|
| AR(1)     | -- | 0.39 | 0.69  |
| Residual  | -- | 4.12 | 0.0001|
| AIC−2ResLL | 85/59 |      |       |

Month was treated as fixed effect and year as random effect in Proc Mixed™. No indication for serial correlation was detected.
### Extended Data Table 2 | Putative elements of the seasonal cycles of the members of *A. gambiae s.l.*

| Taxon | Year | Wet season (WS) | Dry season (DS) | WS first-rain surge |
|-------|------|----------------|----------------|---------------------|
|       |      | normal growth  | WS peak         | late-WS low         | rapid decline       | slow decline       | mid DS low         | DS rise            | late-DS peaks      | late-DS decline    | end-DS low         |                  |
|       |      |                |                 |                     |                     |                     |                   |                     |                    |                     |                    |                  |
|       |      | June - Aug     | Sep-Oct          | Oct                  | Nov                 | Dec - Feb           | Feb - Mar          | Mar                 | Apr                 | Apr                 | May               | May - Jun         |
| *A. gambiae s.l.* | 2008-09 | nd | Y | Y | nd | nd | nd | nd | nd | nd | nd | Y | Y |
|       | 2009-10 | Y | Y | Y | Y | nd | nd | nd | nd | nd | nd | nd | nd |
|       | 2010-11 | nd | Y | Y | nd | nd | nd | nd | nd | nd | nd | nd | Y |
|       | 2011-12 | Y | Y | Y | Y | nd | nd | nd | nd | nd | nd | Y | Y |
|       | 2012-13 | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |

| Taxon | Year | normal growth | WS peak | WS decline | DS desertion | onset |
|-------|------|---------------|---------|------------|-------------|-------|
|       |      | Aug - Sep     | Sep-Nov | Nov-Dec    | Jan - May   | Jun - Jul|
| *A. coluzzii* | 2008-09 | nd | Y | nd | nd | Y |
|       | 2009-10 | Y | Y | Y | nd | Y |
|       | 2010-11 | nd | Y | Y | nd | Y |
|       | 2011-12 | Y | Y | Y | nd | N |
|       | 2012-13 | Y | Y | Y | nd | N |

| Taxon | Year | normal growth | WS peak | WS decline | DS sporadic appearance |
|-------|------|---------------|---------|------------|------------------------|
|       |      | Aug - Sep     | Oct     | Nov-Dec    | Jan - Jul              |
| *A. artemisii* | 2008-09 | nd | Y | Y | nd |
|       | 2009-10 | Y | Y | Y | Y |
|       | 2010-11 | nd | Y | Y | Y |
|       | 2011-12 | Y | Y | Y | N |
|       | 2012-13 | Y | Y | Y | Y |

Elements were identified based on visual examination of their population dynamics over the five years of the study (Methods, Fig. 1 and Extended Data Fig. 3). 'Y', 'N', and 'nd', denote periods when a putative element was visible, invisible, or was not determined (because of insufficient data), respectively.

a Each taxon has its own 'element header' with the typical month in which an element was observed. Species-specific elements are shown in colour. Underlined elements were supported statistically (text, Methods and Fig. 2).