Habitat segregation and ecological character displacement in cryptic African malaria mosquitoes

To cite this version:

HAL Id: hal-03002419
https://hal.umontpelliier.fr/hal-03002419
Submitted on 17 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Habitat segregation and ecological character displacement in cryptic African malaria mosquitoes

Billy Tene Fossog, Diego Ayala, Pelayo Acevedo, Pierre Kengne, Ignacio Ngomo Abeso Mebuy, Boris Makanga, Julie Magnus, Parfait Awono-Ambene, Flobert Njokou, Marco Pombi, Christophe Antonio-Nkondjio, Christophe Paupy, Nora J. Besansky and Carlo Costantini

1 Institut de Recherche pour le Développement (IRD), UMR MIVEGEC (UM1, UM2, CNRS 5290, IRD 224), Montpellier, France
2 Laboratoire de Recherche sur le Paludisme, Organisation de Coordination pour la lutte contre les Endémies en Afrique Centrale (OCEAC), Yaoundé, Cameroon
3 Department of Animal Biology, Faculty of Sciences, University of Yaoundé I, Yaoundé, Cameroon
4 Eck Institute for Global Health & Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA
5 Centre International de Recherches Médicales de Franceville (CIRMF), Franceville, Gabon
6 SaBio, Instituto de Investigación en Recursos Cinegéticos (IREC), CSIC-UCLM-JCCM, Ciudad Real, Spain
7 Faculty of Medicine, Universidad Nacional de Guinea Ecuatorial (UNGE), Bata, Equatorial Guinea
8 Institut de Recherche en Ecologie Tropicale (IRET), Libreville, Gabon
9 Sezione di Parasitologia, Dipartimento di Sanità Pubblica e Malattie Infettive, Università di Roma 'La Sapienza', Rome, Italy

Keywords
Anopheles coluzzii, Anopheles gambiae, cryptic species, ecological character displacement, ecological speciation, habitat segregation, malaria vector, molecular forms, niche partitioning, saltwater tolerance, spatial ecology, species distribution modelling, urban pollution.

Abstract
Understanding how divergent selection generates adaptive phenotypic and population diversification provides a mechanistic explanation of speciation in recently separated species pairs. Towards this goal, we sought ecological gradients of divergence between the cryptic malaria vectors Anopheles coluzzii and An. gambiae and then looked for a physiological trait that may underlie such divergence. Using a large set of occurrence records and eco-geographic information, we built a distribution model to predict the predominance of the two species across their range of sympatry. Our model predicts two novel gradients along which the species segregate: distance from the coastline and altitude. Anopheles coluzzii showed a ‘bimodal’ distribution, predominating in xeric West African savannas and along the western coastal fringe of Africa. To test whether differences in salinity tolerance underlie this habitat segregation, we assessed the acute dose–mortality response to salinity of thirty-two larval populations from Central Africa. In agreement with its coastal predominance, Anopheles coluzzii showed overall more tolerant than An. gambiae. Salinity tolerance of both species, however, converged in urban localities, presumably reflecting an adaptive response to osmotic stress from anthropogenic pollutants. When comparing degree of tolerance in conjunction with levels of syntopy, we found evidence of character displacement in this trait.

© 2014 The Authors. Evolutionary Applications published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
Introduction

One of the major interests of evolutionary ecology is understanding the ecological forces that are responsible for the evolution and coexistence of closely related sympatric species. Cryptic species are morphologically indistinguishable and often very closely related genetically and ecologically. Increasing application of DNA sequence analysis has accelerated discovery of cryptic species and revealed their abundance across taxonomic groups and biogeographic regions (Bickford et al. 2007; Pfenninger and Schwenk 2007). Despite theoretical developments, however, empirical genetic and ecological investigations have yet to provide a detailed understanding of the mechanisms responsible for the build-up of ecological and reproductive barriers during speciation (Sobel et al. 2010; Nosil 2012). To the extent that seemingly identical, recently diverged or emerging species exploit the same limiting resource, classical ecological theory predicts that competition will prevent their stable spatial coexistence. As such, the long-term sympatric coexistence of cryptic species implies the presence of intrinsic ecological differences that facilitate specialization and ecological niche partitioning, reducing the strength of interspecific competition (Szilagyi and Meszena 2009; Sobel et al. 2010). Uncovering the nature of these differences is a fundamental goal in studies of speciation, but also is critically important to biodiversity assessment, conservation efforts and control of economically and medically important insect pests (Coluzzi 1970; Ferguson et al. 2010).

Particularly well-known examples of cryptic diversity occur in the mosquito genus *Anopheles*, which contains the vectors of human malaria parasites (Coluzzi 1970). Nearly all *Anopheles* vector species are members of cryptic species complexes (Harbach 2013). By far the best studied is the group of Afrotropical species known as the *Anopheles gambiae* complex, containing the vectors responsible for the majority of malaria transmission in sub-Saharan Africa. All eight currently recognized species in this complex (Coluzzi et al. 2002; Coetsee et al. 2013) are morphologically indistinguishable, yet differ not only in terms of their contribution to malaria transmission but also in ecological characteristics, notably their trophic habits and aquatic larval habitats (Coluzzi et al. 1979; White et al. 2011b). Three species with nonoverlapping distributions are saltwater tolerant, breeding in brackish water along the coastline and associated islands of Western or Eastern Africa (*An. melas* and *An. merus*, respectively), or in geothermal springs of high mineral content in the Semliki forest of Uganda (*An. bwambae*). The distribution of each of the salt-tolerant species overlaps that of at least one of the obligate freshwater species in the complex, suggesting that tolerance to saltwater may have played a role in ecological diversification of this species group.

One of the most recent examples of cryptic diversification in the *An. gambiae* complex is occurring in the nominal species and most important malaria vector, *An. gambiae*, a freshwater taxon whose geographic range spans most of tropical Africa (Coetzee et al. 1993; Sinka et al. 2010). *Anopheles gambiae* carries extensive inversion polymorphisms on chromosome 2, which may contribute to ecological segregation of populations (Coluzzi 1982; Coluzzi et al. 2002; Manoukis et al. 2008; Cheng et al. 2012). However, genetic evidence independent of chromosomal inversions supported the subdivision of *An. gambiae* into two assortatively mating cryptic lineages provisionally referred to as the M and S molecular forms, whose distributions overlap in West and Central Africa. Premating reproductive isolation – although significant (Tripet et al. 2001; Diabate et al. 2006, 2009; Pennetier et al. 2010; Dabire et al. 2013; Sawadogo et al. 2014) – is incomplete, and there is no indication of intrinsic (genetic) postmating isolating barriers in the F1 (or F2) hybrids (Diabate et al. 2007; Hahn et al. 2012). Nevertheless, SNP genotyping at ~400 000 markers across the genome in paired M and S samples from natural populations in West and Central Africa revealed exclusive form-specific clustering, and the largely shared pattern of M-S differentiation across sampling locations suggested that each form is evolving collectively across this large study area regardless of localized M-S gene flow (Neafsey et al. 2010; Reidenbach et al. 2012). This and other evidence led to the recent naming of M and S as distinct species (Coetzee et al. 2013), hereafter referred to as *An. coluzzii* (formerly, M form) and *An. gambiae* (formerly, S form).

The persistence of sympatric and synchronous populations of *An. coluzzii* and *An. gambiae* despite imperfect premating barriers suggests that ecologically based barriers contribute to their isolation. In fact, there is mounting evidence that the ecology of the aquatic larvae plays a significant role. In the more arid savannas of West Africa (e.g. Burkina Faso and Mali) where the bulk of studies have been conducted, differences between *An. coluzzii* and *An. gambiae* in larval behaviours that affect the risk of predation may help explain why *An. coluzzii* is better adapted to live in ecologically more complex and stable habitats, where predators are more abundant and diverse compared to the ephemeral predator-barren larval habitats of *An. gambiae* (Diabate et al. 2005; Gimonneau et al. 2010, 2012a,b). These studies were spurred by the appreciation of landscape-level heterogeneities in the spatial distribution of the two species in West African savannas, where *An. coluzzii* prevails in more xeric areas characterized by larger, more permanent, aquatic habitats, generally of anthropogenic origin and often associated with irrigation (e.g. rice paddies) (Costantini et al. 2009; Gimonneau et al. 2012a).

However, detailed knowledge about the distribution and habitat segregation of *An. coluzzii* and *An. gambiae* remains
Habitat segregation in Anopheles gambiae

Tene Fossog et al.

fragmentary at a continental scale. Outside of Burkina Faso and Mali, only in Cameroon have the ecological factors underlying the divergence of An. coluzzii and An. gambiae been investigated (Lee et al. 2009a; Simard et al. 2009; Kamdem et al. 2012), and initial results suggest important and unanticipated regional differences. Contrary to the situation in West Africa, only An. gambiae occurs in the drier northern savannas of Cameroon (Simard et al. 2009). In southern Cameroon where An. coluzzii and An. gambiae co-occur, strong habitat segregation exists along gradients of urbanization in the interior of the country, with An. coluzzii prevailing in the core of densely urbanized settings in more stable but polluted larval habitats (Kamdem et al. 2012), in accordance with its higher tolerance of ammonia (Tene Fossog et al. 2013). On the other hand, at sites closest to the Cameroon coast, An. coluzzii prevails regardless of the degree of urbanization (Simard et al. 2009; Kamdem et al. 2012), consistent with the results of entomological surveys conducted piecemeal in individual African countries bordering the Gulf of Guinea (Awolola et al. 2005; Calzetta et al. 2008; Caputo et al. 2008; Ridl et al. 2008; Vezenehgo et al. 2009; de Souza et al. 2010; Djogbenou et al. 2011).

Species distribution models (SDMs) can be used in many aspects of fundamental and applied ecology and evolution (Peterson et al. 1999; Guisan and Zimmerman 2000). When SDMs are applied to study closely related species, these models can be helpful in exploring biogeographic relationships and niche evolution (Peterson et al. 1999; Losos et al. 2003). Comprehensive species distribution modelling of An. coluzzii and An. gambiae across their shared range in West and Central Africa has not been attempted before now, and there exist large gaps in our understanding of their biogeographic relationships, constraining the ability to target species-specific control to vectors known to differ in their degree of insecticide resistance and susceptibility to Plasmodium (Santolamazza et al. 2008; Ndath et al. 2011; White et al. 2011a; Dabire et al. 2012). In this study, we aimed to (i) build a model capable of predicting predominance of one or the other species across their sympatric range and (ii) identify ecological factors and physiological correlates that may have shaped the habitat segregation of An. coluzzii and An. gambiae. Using a large set of distribution records extracted across the study area from published and unpublished data, in conjunction with layers of geographic information that included climatic, spatial and anthropogenic variables, we developed an SDM mapping the relative probability of occurrence of the two species across Western and Central Africa. This model predicts significant habitat segregation between An. coluzzii and An. gambiae at this continental scale, with An. coluzzii showing a ‘bimodal’ distribution of relative abundance resulting from its predominance not only in its well-known xeric habitat in West African savannas, but also along the coastal fringe of West and Central Africa. As the variable ‘distance to coast’ was among the significant predictors of the relative abundance of the two species, we compared the salinity tolerance of coastal and inland populations to identify a physiological mechanism potentially underlying the coastal distribution of An. coluzzii. Based on assessments of the acute dose–mortality response after exposure to increasing concentrations of salt, we confirmed that An. coluzzii larvae were more tolerant than An. gambiae, suggesting a continuing role for salinity tolerance in ecological diversification of the An. gambiae complex. Moreover, we found that salinity tolerance increased in both species from their baseline level – corresponding to that of rural populations – whenever larvae were collected from aquatic urban habitats, suggesting a role for osmoregulation in local adaptation to anthropogenic pollutants. Finally, by matching degree of salinity tolerance with levels of syntopy, we found evidence of ecological character displacement between this cryptic species pair, indicating the occurrence of competition-driven divergent natural selection in this trait.

Materials and methods

Study area and occurrence records

Anopheles gambiae and An. coluzzii are broadly sympatric over much of West and Central Africa (della Torre et al. 2005), across a total surface of ~8 000 000 km² (Fig. 1A). Anopheles gambiae occurs allopatrically in East Africa, although there is a single record of An. coluzzii occurrence in Zimbabwe (Masendu et al. 2004). In the absence of further such records, we excluded East Africa. Employing an estimate of the eastern boundary of An. gambiae-An. coluzzii sympathy (see Fig. 1A), we based the analysis on this shared West and Central African range. The study area covers highly diverse habitats, including coastal mangroves, the Guineo-Congolese equatorial rainforest block, forest-savanna mosaic and different types of savanna characterized by increasing degrees of aridity culminating in xeric and predesertic steppes at higher latitudes.

Distribution data for An. gambiae and An. coluzzii were gathered from various published and unpublished sources reporting about their occurrence and abundance. The data set consisted of 1140 georeferenced localities across the study area sampled during 2001–2011 and included a total of 45 834 An. gambiae sensu lato (i.e. An. gambiae and An. coluzzii) that were molecularly identified to species using rDNA-based PCR assays (Scott et al. 1993; Fanello et al. 2002; Santolamazza et al. 2004) (Table 1, Fig. 1A).

Environmental predictors

Thirteen eco-geographic variables, classified into three factors, were considered as predictors: spatial (four variables),...
climate (six variables) and land cover (three variables) (Table 2). Spatial variables included geographic coordinates for latitude, longitude and their product, to describe the spatial structure of species occurrence (Legendre and Legendre 1998). Based on regional or country-wide entomological surveys indicating a predominance of An. coluzzii at sites nearest the coast in several West and Central African countries (Awolola et al. 2005; Calzetta et al. 2008; Caputo et al. 2008; Ridl et al. 2008; Simard et al. 2009; Vezenezho et al. 2009; de Souza et al. 2010; Djogbenou et al. 2011), we also included the variable ‘distance to coast’ as a predictor.

Climatic variables were annual estimates of mean temperature, mean temperature of the wettest quarter, mean precipitation, mean precipitation of the wettest quarter and length of the dry season (defined as the number of months when rainfall is below 150 mm). Elevation was
Table 2. Eco-geographic predictors used to model the relative probability of occurrence of An. gambiae and An. coluzzii.

<table>
<thead>
<tr>
<th>Class</th>
<th>Predictor</th>
<th>Description</th>
<th>Wald test value</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial</td>
<td>Lat</td>
<td>Latitude</td>
<td>20.70</td>
<td>1.29</td>
<td>±0.11</td>
</tr>
<tr>
<td></td>
<td>Lon</td>
<td>Longitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lat × Lon</td>
<td>Product of Latitude and Longitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coast</td>
<td>Distance to Coast (km)</td>
<td>8.40</td>
<td>1.36</td>
<td>±0.60</td>
</tr>
<tr>
<td>Climate</td>
<td>Temp Mean</td>
<td>Annual Mean Temperature (°C)</td>
<td>27.29</td>
<td>0.85</td>
<td>±0.06</td>
</tr>
<tr>
<td></td>
<td>Temp Quarter</td>
<td>Mean Temperature of Wettest Quarter (°C)</td>
<td>23.97</td>
<td>1.12</td>
<td>±0.04</td>
</tr>
<tr>
<td></td>
<td>Precipitation Mean</td>
<td>Mean Annual Precipitation (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Precipitation Quarter</td>
<td>Precipitation of Wettest Quarter (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dry Season</td>
<td>No. Months with Rainfall &lt; 150 mm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>Mean Elevation (m)</td>
<td>26.06</td>
<td>0.99</td>
<td>±0.01</td>
</tr>
<tr>
<td>Land cover</td>
<td>NVDI Mean</td>
<td>Annual NVDI</td>
<td>10.86</td>
<td>0.98</td>
<td>±0.01</td>
</tr>
<tr>
<td></td>
<td>NVDI Variation</td>
<td>Annual Variation in NVDI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Night Light Intensity</td>
<td>5.03</td>
<td>1.07</td>
<td>±0.06</td>
</tr>
</tbody>
</table>

Bold font indicates variables retained at $P < 0.05$ in the final minimal adequate model. 95% CI: 95% confidence interval of odds ratio.

Species distribution modelling

Statistical models to distinguish between populations of An. gambiae and An. coluzzii were developed. As a first step, we fitted logistic regressions to a subset of training localities where the two species occurred allopatrically ($n = 622$; Table 1) (Hosmer and Lemeshow 2000): for localities where only An. coluzzii or An. gambiae was recorded, the response variable was set to one or zero, respectively. To identify the subset of significant eco-geographic predictors explaining species habitat segregation, we used a forward–backward stepwise model selection procedure. Progressively more parsimonious models (those with fewer predictors) were selected if they decreased the Akaike information criterion [AIC; (Akaike 1974)], until the minimal adequate model—the one returning the lowest AIC—was identified (Acevedo et al. 2010). Its predictive performance was evaluated using the subset of localities where both species were recorded in sympathy ($n = 518$; Table 1) as an independent data set. For this purpose, we used calibration plots (Pearce and Ferrier 2000) in which the observed frequency of An. coluzzii (relative to An. gambiae) was regressed against the predicted probability of (relative) An. coluzzii occurrence (Zheng and Agresti 2000). Given complete coincidence between observations and predictions (i.e. perfect calibration), the regression line is expected to pass through the origin with a slope of one (Miller et al. 1991). The range of predicted probabilities of occurrence [0,1] was partitioned in ten equidistant classes of bins, and the mean predicted probability and corresponding mean observed frequency were calculated within each bin class. In estimating the calibration regression parameters, we excluded those bins where the number of localities was $n < 15$, the minimum sample size necessary to yield mean frequency estimates within ±15% of the true mean (Jovani and Tella 2006).
Mosquito larval sampling and salinity tolerance tests

Larval salinity tolerance was investigated in thirty-two populations from coastal and inland localities lying in the rainforest eco-climatic domain of southern Cameroon, Equatorial Guinea and Gabon, Central Africa (Fig. 1B; Table S1). About 170 water collections representing individual mosquito breeding sites were sampled, yielding ~15,800 An. gambiae s.l. larvae (i.e. a mix of An. coluzzii and An. gambiae in this study area) that were subsequently tested for salinity tolerance (Table S1).

Sampled localities were selected based on two criteria: (i) type of habitat with respect to degree of urbanization (i.e. rural versus urban) and (ii) ecological proximity to saltwater (i.e. the sea or river estuaries), so that both coastal and inland locations were included in the study. Based on previous results (Kamdem et al. 2012), urban locations were defined as those hosting ≥5000 inhabitants and urban infrastructures such as paved roads and sewage collection ditches; inland locations were defined as those lying ≥50 km from the coastline or ≥200 m above sea level. Sampled rural breeding sites included pools, rain puddles, ruts, ditches and stranded pirogues containing residual sea water diluted by freshwater. Urban larval habitats included rain puddles, pools and ditches, often located near dumping grounds; accordingly, many of these breeding sites showed signs of apparent organic pollution (filamentous algae, cyanobacterial mats, pungent odours, decomposing waste, etc.) in the slums of urban areas. Only breeding sites returning a sufficient number of late-instar larvae (cf. below) were selected for the tests. Because of the way these taxa partition their habitat in this eco-geo-graphic region, both species occurred in each locality either in allopatry, sympatry or parapatry according to geography and degree of urbanization (Fig. 2); when in sympatry, they often co-occurred in the same breeding sites in more or less balanced proportions. The number of breeding sites tested from each locality varied according to their local abundance (range: 1–27; median: 5; interquartile range: 2.25–7), as sampling effort (i.e. the number of man × hours devoted to larval collections in a single locality) was kept approximately constant across localities. Exceptions were represented by the major urban centres of this region (Bata, Bonaberi/Douala, Libreville, Malabo, Yaoundé – Figs 1B and 2), which were also part of a broader scale study on the urban ecology of malaria vectors (Kengne et al., unpublished data); accordingly, these towns were generally represented by a greater number of breeding sites due to their inclusion in multiple surveys (Table S1).

Figure 2 Relationship between the size of each locality shown in Fig. 1B, expressed as number of inhabitants and the relative abundance of An. coluzzii and An. gambiae. Ellipses identify urban centres with respect to their position relative to the Cameroon Volcanic line (CVL): westwards (red), eastwards (blue) or on the CVL (green). On the opposite sides of the CVL, urban centres are dominated by one or the other of the two species, while localities on the CVL identify a ‘contact zone’. Exceptions to this pattern are Bata (N°1) and Malabo (N°15) in Equatorial Guinea, Libreville (N°12) in Gabon and Tibati (N°25) in Cameroon. The latter is situated on the east of the CVL, but it lies at >800 m altitude in the forest/savanna mosaic, outside of the rainforest ecozone. In Bata and Libreville, the two species are in apparent parapatry (data not shown). Malabo is on the island of Bioko, on the CVL, under the influence of insular biogeography (e.g. Deitz et al. 2012).
Habitat segregation in *Anopheles gambiae*

To relate tolerance with the salinity of the natural breeding sites where each species occurred, larval sampling was accompanied by measurements of ions concentration in individual larval habitats using a portable field tester (Wagtech conductivity meter CyberScan CON11 C551-210), which normalizes readings of electrical water conductivity (EWC) at 25°C.

Mosquito larvae were dipped (Service 1993) from individual breeding sites and then brought to the laboratory or field facilities for salinity tolerance testing on the same day of collection. Before assaying, *An. gambiae* s.l. larvae were sorted out from other species of mosquitoes under a dissecting scope according to a morphological identification key (Gillies and Coetze 1987). *Anopheles coluzzii* and *An. gambiae* larvae are morphologically undistinguishable; so at this point of the experimental protocol, it was not feasible to distinguish which of the larvae were *An. gambiae* and which *An. coluzzii*. Except in cases of known habitat segregation (e.g. Kamdem et al. 2012), each batch of test mosquitoes (usually) contained a mix of both species in variable and *a priori* unknown proportions, and it was only through a *posteriori* DNA-based molecular identification of tested larvae that it was possible to count how many *An. coluzzii* or *An. gambiae* died or survived at each salinity concentration.

Larvae from each individual breeding site were kept and tested separately from those collected from other sites. The pool of larvae collected from a breeding site was randomly allocated to seven cups representing increasing test concentrations of commercial sodium chloride (refined large-grained table salt containing no additives) in distilled water. Six cups contained, respectively, 5%, 10%, 20%, 30%, 40% and 50% dilutions of a 35 g/L NaCl solution assumed to represent 100% sea water. The seventh cup contained only distilled water to act as a control (0% sea water). Occasionally, other concentrations were also tested (35%, 45%, cf. Table S1). Depending on the yield of individual field collections, which varied from one breeding site to another as a function of larval density, the number of larvae tested in each cup ranged from a minimum of 1 to a maximum of 25 (median: 14; interquartile range: 10–17). We defined each test cup as a single batch and the whole series of batches of increasing salinity concentrations from the same individual breeding site as an experimental replicate.

To minimize pupation during the assays, we selected late 3rd to early 4th instar larvae. Before transfer into the test solution, these were fed for 2 h in distilled water with baby-fish food to avoid cannibalism and standardize as much as feasible the physiological status of tested larvae, after which they were not fed. In the field, tests were run indoors at ambient temperature; in Yaoundé, tests were carried out in the laboratory, where temperature averaged 25°C. After 24 h exposure, all larvae that did not react with an escape response (swimming and/or diving) upon stimulation with the tip of a pipette were scored as dead.

**Molecular identification of *An. gambiae* s.l.**

At the end of the exposure interval, all larvae were killed and preserved in microtubes containing 70% ethanol to assess their taxonomic status. DNA of individual larvae extracted by a cetyl trimethyl ammonium bromide (CTAB) protocol (Morlais et al. 2004) was targeted for species identification based on a ribosomal DNA-based PCR-RFLP assay (Fanello et al. 2002).

**Statistical analysis of dose–mortality response to salinity**

To assess the dose–response and calculate the median lethal concentration (LC50) for different habitats (rural versus urban), locations (coastal versus inland) or species (*An. coluzzii* versus *An. gambiae*), we fitted binomial generalized linear mixed models (GLMM) with a logit link function using the lme4 package in R v.2.15.3. We fitted 'salinity concentration', 'habitat', 'location' and 'species' as fixed effects in the GLMMs; the random effects were 'salinity concentration' included as a linear model term grouped by 'locality' and 'breeding site', with 'breeding site' nested within 'locality'. In this way, a separate logistic regression with random intercept and slope was fitted to each individual breeding site to account for heterogeneities and non-independence of the dose–response among breeding sites from the same locality, as well as in the average dose–response among different localities. From the intercept and slope of the fixed effects, the median lethal concentration (LC50) of the main factors of interest (habitat, location and species) was calculated. In comparing the magnitude of fixed effects, we defined per cent relative difference between two values x and y as \( \frac{|x - y|}{\max(x,y)} \).

Inference on the size of fixed effects was based on an information-theoretic approach (Burnham and Anderson 2002, 2004): the set of competitive statistical models that were evaluated included those having the random effects and all possible combinations of fixed effects and their second- and third-order interactions conditional to marginality constraints of main effects. After having verified that the fitted models did not show overdispersion, these were sorted according to increasing values of the second-order bias correction Akaike information criterion (AICc), as appropriate for models in which the ratio of number of observations n to number of parameters K is \( n/K \leq 40 \). The \( \Delta \text{AIC}_c \) statistic was calculated as the AICc difference between a focal model and that having the minimum AICc. The best model has \( \Delta \text{AIC}_c = 0 \), but models considered to have substantial statistical support were identified by the
rule of thumb $\Delta AIC_c \leq 2$ and the analysis of Akaike weights (Burnham and Anderson 2002, 2004). Whenever there was evidence for more than a single best model, we calculated multimodel average parameters. We used full-model averages with shrinkage, which do not have a tendency to bias the value away from zero. As it is an unresolved issue how the variance of these estimates should be calculated, the standard errors and confidence intervals were returned for the subset-averaged (or ‘conditional’) coefficients. All calculations were carried out using functions available in the R package MuMIn v.1.9.0.

Results

Species distribution models

The eco-geographic predictors retained in the final minimal adequate model are summarized in Table 2. The stepwise procedure selected 7 of 13 variables representative of all three factors (i.e. spatial, climate and land cover), which explained 64% of the overall variance in the data (estimated from the Nagelkerke coefficient of determination $R^2$).

The minimal adequate model was projected across the study area, producing a map of predicted relative probability of occurrence of An. coluzzii relative to An. gambiae (Fig. 3A). One of the main features of this map is the ‘bimodal’ distribution of An. coluzzii in contrasting geographic domains: the highest probabilities of occurrence correspond to a wide belt covering the most xeric savannas found at higher latitudes in Western Africa – from the westernmost limit in Senegambia and eastwards to Chad – and along a narrow strip of coastline bordering the Gulf of Guinea, regardless of the precise environmental domain encountered along this strip. If it were not for a zone of contact between these two regions, corresponding to the coastal savanna areas at the westernmost limit in Senegambia, the range of An. coluzzii would approximate a disjunct distribution, particularly at its easternmost limits.

To provide better visual and quantitative assessment of the ‘bimodality’ in probability of occurrence of An. coluzzii, we plotted its relative frequency across all localities (i.e. training + validation data sets) as a function of distance from the coast (Fig. 3B). The highest relative frequencies occurred at the shortest and longest distances. In the coastal zone within 5 km of the sea, the average observed relative frequency of An. coluzzii was 0.62 $\pm$ 0.05, and beyond 900 km from the Gulf of Guinea the observed relative frequency was 0.88 $\pm$ 0.04. The minimum average frequency (0.06 $\pm$ 0.04) was found at an intermediate distance, 350–450 km from the coast.

Conversely, An. gambiae had the highest relative probability of occurrence across the central block of moist

Figure 3 (A) Relative probability of occurrence predicted for An. gambiae and An. coluzzii (dark blue: 1 An. coluzzii, and 0 An. gambiae; dark red: 1 An. gambiae, and 0 An. coluzzii) across their sympatric range. The inset shows an enlarged view of the area delimited by a black line in the map where high altitudes are encountered close to the coastline due to the presence of Mount Cameroon and the Cameroon Volcanic line; (B) Frequency of An. coluzzii relative to An. gambiae plotted as a function of distance from the coast. The red line depicts the second-order polynomial regression curve fitted to the observation records, represented by grey dots. Black circles and associated standard error bars illustrate the mean frequency of An. coluzzii for arbitrary classes of distance plotted for a visual assessment of goodness-of-fit.
savannas, in the forest/savanna mosaic, in most of the rain-forest area and in highlands. Although the dry savannas of West Africa share climatic similarities with the southern arid regions of Angola and Democratic Republic of Congo (DRC), An. gambiae had contrasting relative probabilities of occurrence in these two regions – low in western, but high in southern arid savannas. Along major lowland rivers such as the Niger and the Benoue/Donga in Nigeria, the Ogoue in Gabon and the Congo in the DRC, both species have similar (intermediate) relative probabilities of occurrence.

Calibration plots assessing the predictive performance of the final model showed that predictions closely fitted observations (coefficient of determination of the regression line: $R^2 = 0.99$; Figure S1). There was only a slight amount of calibration bias and spread, indicating that at locations where the predicted probability of occurrence was $>60\%$ the model slightly overestimated the frequency of An. coluzzii, and conversely, at locations below the 60% threshold, the model underestimated the relative frequency of An. coluzzii.

### Population patterns in salinity tolerance

During the salinity tolerance assays, occasionally we observed unusual mortality rates in the control batches. Thus, we investigated first how larval control mortality was partitioned across habitats, locations and populations by fitting binomial GLMMs only to the control batches. Even if average control mortalities were well below 5% (the arbitrary consensus threshold accepted to validate an acute toxicity bioassay), habitat and location of the test population had an impact on control mortality rates (Table 3). The relative importance of the fixed effects was ‘habitat’ = 0.66, ‘location’ = 0.46, and ‘habitat x location’ = 0.14. The full set of models was not well resolved (Table 3); hence, we calculated multimodel parameter estimates to fit average control mortalities. Mortality rankings were rural inland < rural coastal < urban coastal < urban inland (Table 4). Urban control mortalities were more than three-fold higher compared to rural localities, regardless of the location of the population (Table 4), potentially reflecting greater levels of stress due to increased exposure to xenobiotics resulting from anthropogenic pollution. To reduce possible biases in subsequent analyses, we excluded all replicates in which the control batch scored at least one death at the end of the 24-h exposure interval. Results from the statistical analysis of the full data set, that is without excluding replicates with mortality in the controls, were qualitatively similar to those presented below (data not shown).

After having verified with the full data set that ‘habitat’ (urban/rural), ‘location’ (inland/coastal), ‘salinity concentration’ and their second- and third-order interactions had significant impacts on larval mortality (Appendix S1), next

### Table 4. Ratio (above the main diagonal) and per cent relative difference (below the main diagonal) of multimodel control mortality estimates observed during the salinity tolerance assays.

<table>
<thead>
<tr>
<th></th>
<th>Rural</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coastal</td>
<td>Inland</td>
</tr>
<tr>
<td>Mean mortality</td>
<td>0.25%</td>
<td>0.05%</td>
</tr>
<tr>
<td>Rural</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coastal</td>
<td>–</td>
<td>5.4</td>
</tr>
<tr>
<td>Inland</td>
<td>81%</td>
<td>–</td>
</tr>
<tr>
<td>Urban</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coastal</td>
<td>66%</td>
<td>94%</td>
</tr>
<tr>
<td>Inland</td>
<td>78%</td>
<td>96%</td>
</tr>
</tbody>
</table>

### Table 3. Evaluation of competitive binomial generalized linear mixed models fitted to control batches of the salinity tolerance assays data set.

<table>
<thead>
<tr>
<th>Model No. (rank)</th>
<th>Intercept</th>
<th>Habitat</th>
<th>Location</th>
<th>Habitat x Location</th>
<th>d.f.</th>
<th>$\Delta$AIC&lt;sub&gt;C&lt;/sub&gt;</th>
<th>Akaike weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (1)</td>
<td>6.473</td>
<td>+</td>
<td></td>
<td></td>
<td>4</td>
<td>0</td>
<td>0.349</td>
</tr>
<tr>
<td>1 (2)</td>
<td>4.985</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1.19</td>
<td>0.193</td>
</tr>
<tr>
<td>4 (3)</td>
<td>6.641</td>
<td>+</td>
<td>+</td>
<td></td>
<td>5</td>
<td>1.35</td>
<td>0.178</td>
</tr>
<tr>
<td>3 (4)</td>
<td>5.431</td>
<td>+</td>
<td></td>
<td></td>
<td>4</td>
<td>1.78</td>
<td>0.143</td>
</tr>
<tr>
<td>8 (5)</td>
<td>5.949</td>
<td>+</td>
<td>+</td>
<td></td>
<td>6</td>
<td>1.87</td>
<td>0.137</td>
</tr>
</tbody>
</table>

Fixed effects included in the full model (Model No. 8) are the factors ‘habitat’ (two levels: rural versus urban), ‘location’ (two levels: coastal versus inland) and their interaction. Crosses (+) indicate the occurrence of a term in the model, while the intercept (logit scale) refers to the fitted value for baseline factor levels. Random effects of all models are breeding sites nested within localities. Models are sorted according to increasing values of the second-order bias correction Akaike information criterion ($\Delta$AIC<sub>C</sub>). The $\Delta$AIC<sub>C</sub> is the AIC<sub>C</sub> difference between a model and that having the minimum AIC<sub>C</sub>, so that the best model has $\Delta$AIC<sub>C</sub> = 0. Models whose $\Delta$AIC<sub>C</sub> ≤ 2 are considered to have substantial statistical support and those whose $\Delta$AIC<sub>C</sub> > 2 have less to no support. Akaike weights quantify the ‘strength of evidence’ for a given model, that is the probability that the model is the best one, conditional on both the data and full set of models considered.
we fitted binomial GLMMs to test whether salinity tolerance differed between An. coluzzii and An. gambiae. To do so, we had to take into account two constraints: (i) that results from different populations could not be pooled in a single analysis, as just mentioned above (Appendix S1); and (ii) that molecular identifications were carried out either at batch level for sympatric populations or at breeding site level for all other populations. In the first instance, specific dose–response curves could be estimated, whereas in the latter case, specific parameter estimates were obtained by carefully choosing the populations or breeding sites to be included in the analysis.

Overall, An. coluzzii was significantly (Table 5) more tolerant than An. gambiae to salinity, except when populations of the latter occurred – in fewer instances – in urban environments (Table 6). The level of tolerance within species, however, was not predicted by distance of the population from the coast or altitude, as the effect due to urbanization swamped that due to location (Table 6; Appendix S1).

### Salinity tolerance and ionic environment

Salinity tolerance may increase in urban habitats (regardless of inland or coastal location) due to higher content of ions derived from pollution by nitrogenous and phosphorous compounds (Antonio-Nkondjio et al. 2011; Tene Fossog et al. 2013). To investigate the relationship between the degree of salinity tolerance and the amount of ions available in the aquatic milieu of natural breeding sites, we plotted the LC\textsubscript{50} of individual populations as a function of the average electrical water conductivity (EWC) measured in those breeding sites from which test larvae were collected in each locality (Fig. 4). The general shape of the functional

### Table 5. Evaluation of competitive generalized linear mixed models fitted to the salinity tolerance assays data set pertaining to rural coastal populations of An. coluzzii and An. gambiae from Cameroon and Equatorial Guinea.

<table>
<thead>
<tr>
<th>Model No. (rank)</th>
<th>Intercept Salinity Species Salinity ( \times ) Species</th>
<th>d.f.</th>
<th>( \Delta \text{AIC}_c )</th>
<th>Akaike weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 (1)</td>
<td>-11.280 0.3617 +</td>
<td>9</td>
<td>0</td>
<td>0.748</td>
</tr>
<tr>
<td>8 (2)</td>
<td>-11.040 0.3538 +</td>
<td>10</td>
<td>2.24</td>
<td>0.244</td>
</tr>
<tr>
<td>2 (3)</td>
<td>-10.080 0.3442 +</td>
<td>8</td>
<td>9.12</td>
<td>0.008</td>
</tr>
<tr>
<td>1 (4)</td>
<td>1.179</td>
<td>7</td>
<td>25.26</td>
<td>0</td>
</tr>
<tr>
<td>3 (5)</td>
<td>-3.762</td>
<td>8</td>
<td>42.04</td>
<td>0</td>
</tr>
</tbody>
</table>

Parameter estimates of the model with best statistical support (No. 4) are shown in Table 6. For an explanation of statistics and symbols, cf. Table 3.

### Table 6. Dose–response parameters of salinity tolerance for several populations of Anopheles coluzzii and An. gambiae collected from Cameroon, Gabon and Equatorial Guinea (Central Africa).

<table>
<thead>
<tr>
<th>Species</th>
<th>Habitat</th>
<th>Location</th>
<th>Intercept (±SE)</th>
<th>Slope (±SE)</th>
<th>LC\textsubscript{50} (±SD)</th>
<th>Populations*</th>
</tr>
</thead>
<tbody>
<tr>
<td>An. gambiae</td>
<td>Rural</td>
<td>Coastal</td>
<td>-9.103 (±0.706*)</td>
<td>0.362 (±0.046)</td>
<td>25.2 (±2.8)</td>
<td>Bouandjo, Bwambe, Ebo, Eboudja II, Grand Batanga, Lolabe</td>
</tr>
<tr>
<td></td>
<td>Inland</td>
<td>-8.259 (±0.952)</td>
<td>0.350 (±0.053)</td>
<td>23.6 (±2.9)</td>
<td>Ongot, Tibati outskirts, Tongo-Gadima</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>-8.587 (±1.176)</td>
<td>0.279 (±0.035)</td>
<td>30.8 (±3.6)</td>
<td>(Bata), Campo, Libreville</td>
<td></td>
</tr>
<tr>
<td>An. coluzzii</td>
<td>Rural</td>
<td>Coastal</td>
<td>-11.280 (±1.826)</td>
<td>0.362 (±0.046)</td>
<td>31.2 (±2.8)</td>
<td>Bouandjo, Bwambe, CDC Plantations, Ebo, Eboudja II, Grand Batanga, Lolabe, Sacriba, Sipopo</td>
</tr>
<tr>
<td></td>
<td>Inland</td>
<td>-8.374 (±1.078)</td>
<td>0.235 (±0.029)</td>
<td>35.6 (±4.2)</td>
<td>Buea, Kumba, Manfe, Mbanga, Nguti, Tibati</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>-8.271 (±0.798)</td>
<td>0.274 (±0.021)</td>
<td>30.1 (±3.6)</td>
<td>Bata, Bonaberi, Campo, Douala, Kribi, Libreville, Malabo, Mbin, Nzou, Tiko</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inland</td>
<td>-8.833 (±0.673)</td>
<td>0.246 (±0.024)</td>
<td>35.9 (±4.1)</td>
<td>(Buea), Lour, Mbanga, Njome, Nkongsamba, Yaoundé</td>
<td></td>
</tr>
</tbody>
</table>

The median lethal concentration (LC\textsubscript{50}), intercept and slope refer to the logistic regression lines relating larval mortality to water salinity, expressed in per cent sea water. Dashes indicate lack of data due to the absence of populations from certain habitats (Kamdem et al. 2012). The ‘populations’ column shows the localities that were included in the analysis to calculate specific parameter estimates. When the same locality appears in multiple rows, specific parameters were estimated by taking into account results of molecular identifications at the breeding site level. * cf. Figures 1B, 2 and Table S1; populations in parentheses are not included in the analysis for lack of adequate samples for that specific population. * SE of the difference of the corresponding parameter estimate for An. coluzzii in the same Habitat/Location.
relationship between the two variates was explored by fitting a generalized additive model to the data. The resulting loess smoother (continuous solid black curve in Fig. 4) suggests that salinity tolerance of populations increased with the average conductivity of the breeding sites, and that the relationship plateaus, or perhaps humps down, at the highest EWC values that we measured in the field. To test the significance of this association, as well as whether the relationship is linear or asymptotic, we tested three competitive regression models to the data, whose parameters are presented in Table 7. The solid continuous black curve represents a loess smoother fitted to the data in order to extract the general shape of the functional relationship between the two variates. The blue asymptotic curve is the most parsimonious model based on the AIC. The other competitive models are shown as a dashed black (linear model) and horizontal grey (null model) lines.

![Figure 4](image-url)

**Figure 4** Functional relationship between electrical water conductivity and salinity tolerance in individual populations of *An. coluzzii* and *An. gambiae* from Central Africa. Trend lines show competitive regression models fitted to the data, whose parameters are presented in Table 7. The solid continuous black curve represents a loess smoother fitted to the data in order to extract the general shape of the functional relationship between the two variates. The blue asymptotic curve is the most parsimonious model based on the AIC. The other competitive models are shown as a dashed black (linear model) and horizontal grey (null model) lines.

Table 7. Parameter estimates of the competitive linear and nonlinear models shown in Figure 4.

<table>
<thead>
<tr>
<th>Model name (rank)</th>
<th>Model Equation</th>
<th>( a (\pm SE) )</th>
<th>( b (\pm SE) )</th>
<th>Asymptote</th>
<th>Residual d.f.</th>
<th>( AIC ) Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymptotic (1)</td>
<td>( y = ax/(1+bx) )</td>
<td>0.9621* (±0.4117)</td>
<td>0.0266* (±0.0127)</td>
<td>36.2</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>Linear (2)</td>
<td>( y = a + bx )</td>
<td>12.29 (±8.80)</td>
<td>7.78* (±3.47)</td>
<td>–</td>
<td>23</td>
<td>2.49</td>
</tr>
<tr>
<td>Null (3)</td>
<td>( y = a )</td>
<td>31.9*** (±1.1)</td>
<td>–</td>
<td>31.9</td>
<td>24</td>
<td>5.42</td>
</tr>
</tbody>
</table>

Asterisks denote levels of statistical significance for the null hypothesis that the parameter is equal to zero: * \( P < 0.05 \); ** * \( P < 0.001 \).

Table 7 shows the parameter estimates for the competitive linear and nonlinear models fitted to the data, whose parameters are presented in Table 7. The solid continuous black curve represents a loess smoother fitted to the data in order to extract the general shape of the functional relationship between the two variates. The blue asymptotic curve is the most parsimonious model based on the AIC. The other competitive models are shown as a dashed black (linear model) and horizontal grey (null model) lines.

Notably, mean conductivity values of the sampled larval habitats were higher in coastal compared to inland, as well as in urban compared to rural sites within each location (Figure S2). The difference is close to statistical significance for the comparison between locations (ANOVA, \( F_{1,157} = 3.23; P = 0.07 \)). However, the EWC data are distributed log-normally, and more revealing is the variability in EWC among larval habitats within each environment. Importantly, coastal breeding sites were significantly more variable in EWC than inland ones (Bartlett test of homogeneity of variances: \( K^2 = 68.04; \text{d.f.} = 1; P < 0.0001 \)), regardless of type of habitat (Figure S2).

**Discussion**

Phenotypic divergence can result from divergent or disruptive natural selection arising from differential resource exploitation, competitive interactions and/or ecological opportunity. These processes can be important causal drivers of divergence ultimately leading to speciation (Schluter 2000; Coyne and Orr 2004; Dieckmann et al. 2004). Traits evolving under the influence of divergent natural selection can provide fitness advantages under specific environmental conditions, leading to local adaptation of subpopulations. Adaptation to alternative environments therefore plays a crucial role in generating phenotypic and population divergence and provides a mechanistic ecological understanding of the process of speciation (Orr and Smith 1998; Turelli et al. 2001; Bolnick and Fitzpatrick 2007).

Recently diverged species, for which clues about past speciation history have not yet been completely erased, provide opportunities to test these ideas. The cryptic species pair *An. coluzzii* and *An. gambiae* investigated here represents a good model because their divergence is recent and reproductive isolation has been accomplished to variable degrees across their geographic range (e.g. Nwakanma et al. 2013). As their speciation history cannot be studied directly, it is crucial to understand which traits may be exposed to divergent natural selection and which ecological conditions promote diversification of these traits. Towards this aim, we sought major ecological gradients of diver-
Habitat segregation in *Anopheles gambiae*

Tene Fossog et al.

The unexpected and conspicuous second ‘mode’ of the predicted bimodal predominance of *An. coluzzii* is a ribbon of coastline bordering the Gulf of Guinea, largely captured by the spatial descriptor, ‘distance to coast’. This is visualized in the map of relative probability of occurrence (Fig. 3A) as a thin strip of *An. coluzzii* predominance along the whole coastline of West and Central Africa, expanding inland to a considerable extent in the forested region between Côte d’Ivoire and Ghana potentially due to the influence of altitude (see below). The steep drop in *An. coluzzii* relative frequency when moving away from the coastline is especially striking in light of the fact that the localities serving to fit the parabolic relationship between relative frequency and distance from the coast (Fig. 3B) do not fall along a linear transect, but instead are spread across a vast geographic area. This consideration underscores the homogeneity of the coastal distribution pattern, which may reflect the existence of a coastal population of *An. coluzzii* that is isolated to some extent from northern savanna populations.

The global pattern of relative probability of occurrence of *An. coluzzii* along the coastline appears to be modulated locally by environmental heterogeneities such as altitudinal gradients, along which *An. coluzzii* and *An. gambiae* segregate. This can be seen clearly considering the arc of ancient volcanoes emerging as oceanic islands in front of Mt. Cameroon (including Bioko Island) and present-day volcano Mt. Cameroon itself, whose slopes emerge from the coastline and rise to more than 4000 m above sea level. The data predict a reversal in the relative abundance of the two species along such altitudinal gradients near the coast: *An. coluzzii* predominates at sea level, but is replaced by *An. gambiae* above ~430 m (Fig. 3A). This predicted pattern conforms to the observed relative abundance of *An. coluzzii* and *An. gambiae* in the region of Mt. Cameroon and the backbone of the Cameroon Volcanic Line. *An. coluzzii* is predominant, as expected, in localities such as Tiko, Limbe or Idenau occurring along the coast at elevations <200 m (Bigoga et al. 2007; Lee et al. 2009a; Simard et al. 2009); conversely, its abundance decreases below the threshold of detection in mosquito samples above 800–1000 m (Tanga et al. 2010). At intermediate elevations, such as Buea, both species can be present in balanced proportions (Tanga et al. 2010; this study, Figs 1B, 2 and 3A). The association of coastal *An. coluzzii* populations with low altitude may explain the expanded patch of predicted suitable habitat between Ghana and Côte d’Ivoire (Fig. 3A), an area mostly lying below 200 m elevation. Conversely, the absence of *An. coluzzii* beyond 800–1000 m elevation may explain its surprising absence in xeric savannas at higher latitudes in the southern hemisphere (e.g. in central and south-eastern Angola), which are mainly highland plateau regions above 800 m elevation. At these higher altitudes, patches of *An. coluzzii* apparently occur only in strictly urban environments (e.g. Yaoundé...
Habitat segregation in Anopheles gambiae and Nkongsamba at 800–850 m in Cameroon or Huambo at 1700 m in Angola [Cuamba et al. 2006; Calzetta et al. 2008; Kamdem et al., 2012; this study].

Somewhat surprisingly, the major environmental predictor ‘degree of urbanization’, which was identified by species distribution models developed from regional and microgeographic surveys conducted in central and southern Cameroon (Kamdem et al. 2012), was not a significant predictor in our continental-scale SDM. This apparent contradiction may be explained by several factors, both technical and biological. First, the geographic scale of analysis differs [1 × 1 km in Kamdem et al. (2012) vs 5 × 5 km spatial units in this study]. Given that reversals in the relative abundance of the two species can happen along urban–rural transects as narrow as 6 km (Kamdem et al. 2012), the spatial resolution of the present study may have blurred the effect of urbanization on probability of occurrence. Second, the impact of urbanization on An. coluzzii prevalence has been extensively studied and validated only in populations occurring in the rainforest eco-climatic domain in Cameroon, Central Africa. The impact of this environmental variable may be smaller (or nil) in the xeric savannas in Western Africa. In fact, it is noteworthy that our samples from larval populations located in urban habitats west of the Cameroon Volcanic line were predominantly An. gambiae, not An. coluzzii (Figs 1B and 2). The Cameroon Volcanic line is a geographic barrier structuring populations of An. coluzzii (Pinto et al. 2013) and at least one other species of the An. gambiae complex, An. melas (Deitz et al. 2012). Third, as most of the largest cities of the rainforest ecozone in Western and Central Africa are coastal, the degree of urbanization is statistically confounded with the other spatial predictor used in our SDM, ‘distance to coast’.

The ecological significance of differences in salinity tolerance

In addition to comprehensive modelling of An. coluzzii versus An. gambiae distributions, a second major objective of this study was to identify physiological correlates that may have shaped the predicted distribution patterns. In particular, we aimed to test the hypothesis that An. coluzzii has greater tolerance to salinity than An. gambiae, a capacity that may contribute to its predicted coastal distribution along the entire Gulf of Guinea and its observed coastal distribution in southern Cameroon. Acute toxicity bioassays demonstrated moderate but statistically significant differences in larval salinity tolerance between the two species whenever they co-occurred in strict sympatry in rural habitats, as well as between rural and urban populations independent of distance to the coast. Both coastal and inland populations of An. coluzzii expressed higher tolerance to salinity than An. gambiae, except in instances when the latter species occurred in urban habitats. Although beyond the scope of the present study, further insights could be gained from investigations of chronic exposure to salinity and its impact on life history traits and fitness of natural populations.

Relative to the bona fide saltwater tolerant sibling species An. melas and An. merus that develop successfully in brackish water (Gillies and De Meillon 1968; Mosha and Mutero 2015), An. coluzzii even at its most salt-tolerant manifestation must be considered an obligate freshwater taxon, confined to water whose osmotic concentration is less than larval haemolymph (~300 mOsm/kg or ~30% sea water; Bradley 1987). Nevertheless, increased salinity tolerance in An. coluzzii may be ecologically relevant, as it may provide a mechanism to breed more successfully in marginal freshwater habitats prone to contamination by salts, thereby conferring an advantage against less tolerant competitors such as An. gambiae. For example, larval habitats exclusive to An. coluzzii during our surveys were fishermen pirogues stranded on the beach or emerging water table pools along the beach or river estuaries, which contained much greater amounts of salts. Given the predicted coastal distribution of An. coluzzii along the vast Gulf of Guinea, we hypothesize that higher salt tolerance may be a more global attribute across coastal An. coluzzii populations, and if so, that this difference may play a role in hitherto unexplained features of the geographic distribution of An. coluzzii and An. gambiae elsewhere in Central and West Africa. For example, An. coluzzii and An. gambiae have been reported to segregate along the river Gambia according to eco-geographic gradients (Caputo et al. 2008, 2014): the frequency of An. coluzzii decreased when moving away from the river banks and upstream, in accordance with a gradient in the degree of exposure to tidal brackish water. These same studies highlighted the association of An. coluzzii with coastal rice cultivation, where this species would be exposed to brackish water. In another example from coastal Benin, the saltwater tolerant species An. melas was found to co-occur in breeding sites of relatively high salinity together with a taxon that was almost certainly An. coluzzii (unknown at the time; Akogbeto 1995; Djogbenou et al. 2011). Together, these biogeographic records suggest that the greater degree of salinity tolerance measured in An. coluzzii from Central Africa may be more widely applicable geographically and combined with other evidence, prompt a more general working hypothesis: An. coluzzii is adapted to tolerate greater levels of environmental stressors in its larval habitats, not only in sites closely associated with the coast/sea, but also those associated with urban centres containing xenobiotics and ammonia (Kamdem et al. 2012; Tene Fossog et al. 2013; Cassone et al. 2014). Extending bioassays of
An. coluzzii and An. gambiae across their predicted geographic range will be a critical next step towards confirming this inference.

Rather paradoxically, urban coastal populations were consistently less tolerant than urban inland ones in both species. Only in the case of An. gambiae did rural populations conform to the expectation that inland populations would be less saline-tolerant than coastal ones. Moreover, An. gambiae exhibited a twofold more variable degree of tolerance than An. coluzzii across populations (Coefficient of Variation: 19.0% versus 9.5%, respectively). To shed more light on these results, it is informative to present salinity tolerance in the context of joint species distributions (Fig. 5). Considering only An. coluzzii, its level of salinity tolerance was similar across biotopes. However, the tolerance of An. gambiae varied substantially as a function of biotope and degree of sympathy with An. coluzzii. In particular, sympatric populations of the two species showed divergent levels of salinity tolerance (with An. coluzzii having the greater tolerance), but allopatric or parapatric populations of the two species in urban habitats had similar levels of tolerance, largely due to increased salinity tolerance by An. gambiae – one of the signatures of ecological character displacement.

Character displacement in salinity tolerance

Ecological character displacement occurs when competition imposes divergent selection on interacting species, causing divergence in traits associated with resource use (Martin and Pfennig 2011). This process can facilitate species coexistence, enhance phenotypic differences between sympatric species and promote or even initiate speciation and adaptive radiation (Rice and Pfennig 2010; Martin and Pfennig 2011).

The classic pattern to recognize the existence of character displacement is to observe the degree of phenotypic divergence in allopatric and sympatric populations of the same species pair: when sympatric trait values are more extreme than values occurring in allopatry, there is evidence for character displacement (Brown and Wilson 1956). However, this pattern can emerge also as a consequence of other processes. Schluter and McPhail (1992) have summarized the conditions under which character displacement evolves as an adaptive response to interspecific competition. Because of the involvement several different fields of study, it is generally difficult to test all these conditions in a single analysis, which explains why it has been historically difficult to prove the existence of character displacement (Schluter and McPhail 1993; Losos 2000). Clearly, here we have not tested the full range of conditions; therefore, further work is needed to cover the whole set. Most notably, we do not know whether differences in salinity tolerance have emerged in situ under coexistence of the two species, or rather as a result of differences evolved in allopatry followed by subsequent range expansion. In other words, it is difficult with our current data to assess the role played by the observed differences in salinity tolerance in the speciation history of An. coluzzii and An. gambiae. Experimental manipulations can provide a more stringent test of character displacement (e.g. Tyerman et al. 2008): we expect that if allopatric populations of An. coluzzii and An. gambiae are placed together, they should compete more strongly than sympatric populations.

Figure 5 Evidence for ecological character displacement in populations of An. gambiae (red) and An. coluzzii (blue) from Central Africa. The Gaussian curves show the probability density functions of tolerance thresholds, according to habitat, location and degree of syntopy. The dotted red curve in all panels depicts the tolerance of inland rural forest populations of An. gambiae, which are postulated to represent the ancestral state of the trait. Panel (A) includes localities where both species extensively share the same breeding sites, whereas in (B), both species share the same locality but not the same breeding sites due to spatial segregation (Libreville, Gabon). The dotted blue line refers to coastal urban populations of An. coluzzii in allopatry, where An. gambiae does not occur (Bonaberi/Douala, Malabo, Tiko). In (C), both species share the same habitat but not the same localities, each locality being exclusive of one of the two taxa; An. gambiae urban populations are represented by towns located in the savanna (Tibati, Cameroon) or in the forest westwards of the Cameroon Volcanic line (cf. Figs 1–2 and Table S1).
Habitat segregation in *Anopheles gambiae*

Linking distribution modelling with functional ecology: a theoretical framework for testing *An. coluzzii* versus *An. gambiae* interactions in the forest ecozone of Central Africa

The evidence that salinity tolerance changes in relation to the joint occurrence of *An. coluzzii* and *An. gambiae* in this geographic region prompts a hypothetical working model of species interactions, whose aim is to provide testable working hypotheses for future quantitative research (Fig. 6). We assume that larvae of *An. gambiae* and *An. coluzzii* respond differently to two environmental stressors, resource and osmotic stress, such that intersecting isoclines define four regions of population persistence in stressor state space (Fig. 6A). For reasons presented above, osmotic stress is assumed to increase in urban and coastal environments. Resource stress corresponds to the breeding opportunities and quality of larval habitats that each biotope provides. In the rainforest of Central Africa, the presence of *An. gambiae* s.l. is strictly dependent on the presence of human activities that clear the forest canopy and expose bare soil for larval habitats to occur; pristine rainforest is not a suitable habitat for the larvae of these mosquitoes. Thus, size of human settlements is taken as a proxy for available breeding opportunities, whereby small rural settlements are generally poorer than urban centres in breeding opportunities, and organic enrichment of larval habitats in denser settlements may increase the bioavailability of nutrients. Similarly, nutrients such as dissolved phosphate, nitrate and silicate can accumulate at coastal river mouths and alluvial plains through drainage and river transport and can lead to eutrophication of surface waters (e.g. Turner & Rabalais, 1994). Thus, we assume resource stress to increase along the urban < rural and coastal < inland environmental gradients (Fig. 6A).

Beyond the zero isoclines, population growth is negative, and stable population persistence is not possible except as sink populations. Region ‘C’ in Fig. 6A defines the set of values in which osmotic stress is too high for *An. gambiae* despite the abundance and quality of resources (i.e. lower resource stress); this is where *An. coluzzii* occurs alone. Coastal urban habitats are presumed to be mostly representative of this condition. Region ‘G’ defines areas where

![Figure 6](image.png)

**Figure 6** A theoretical framework of functional ecological interactions between *An. gambiae* and *An. coluzzii* in the forest ecozone of Central Africa. Panel (A) represents the zero isoclines of population growth in relation to different environmental stressors: blue (*An. coluzzii*) and red (*An. gambiae*) lines set the limits for combinations of osmotic and resource stress beyond which population growth becomes negative, leading to no stable persistence of either species. On the abscissa, osmotic stress is assumed to increase along the inland versus coastal, as well as the rural versus urban environmental gradients. On the ordinate, stress from lack of resources – particularly nutrients and suitable larval habitats – is supposed to increase in the reverse order along the same environmental gradients. The ‘N.A.’ region delimited in state space by a dotted curve is presumably not available in the study area (i.e. regions with abundant resources in presence of very low levels of osmotic stress). The ‘C’ and ‘G’ letters in state space along the same environmental gradients. The ‘C’ and ‘G’ letters in state space correspond to the breeding opportunities and quality of larval habitats that each biotope provides. In the rainforest of Central Africa, the presence of *An. gambiae* s.l. is strictly dependent on the presence of human activities that clear the forest canopy and expose bare soil for larval habitats to occur; pristine rainforest is not a suitable habitat for the larvae of these mosquitoes. Thus, size of human settlements is taken as a proxy for available breeding opportunities, whereby small rural settlements are generally poorer than urban centres in breeding opportunities, and organic enrichment of larval habitats in denser settlements may increase the bioavailability of nutrients. Similarly, nutrients such as dissolved phosphate, nitrate and silicate can accumulate at coastal river mouths and alluvial plains through drainage and river transport and can lead to eutrophication of surface waters (e.g. Turner & Rabalais, 1994). Thus, we assume resource stress to increase along the urban < rural and coastal < inland environmental gradients (Fig. 6A).

Beyond the zero isoclines, population growth is negative, and stable population persistence is not possible except as sink populations. Region ‘C’ in Fig. 6A defines the set of values in which osmotic stress is too high for *An. gambiae* despite the abundance and quality of resources (i.e. lower resource stress); this is where *An. coluzzii* occurs alone. Coastal urban habitats are presumed to be mostly representative of this condition. Region ‘G’ defines areas where

![Figure 6](image.png)

**Figure 6** A theoretical framework of functional ecological interactions between *An. gambiae* and *An. coluzzii* in the forest ecozone of Central Africa. Panel (A) represents the zero isoclines of population growth in relation to different environmental stressors: blue (*An. coluzzii*) and red (*An. gambiae*) lines set the limits for combinations of osmotic and resource stress beyond which population growth becomes negative, leading to no stable persistence of either species. On the abscissa, osmotic stress is assumed to increase along the inland versus coastal, as well as the rural versus urban environmental gradients. On the ordinate, stress from lack of resources – particularly nutrients and suitable larval habitats – is supposed to increase in the reverse order along the same environmental gradients. The ‘N.A.’ region delimited in state space by a dotted curve is presumably not available in the study area (i.e. regions with abundant resources in presence of very low levels of osmotic stress). The ‘C’ and ‘G’ letters in state space correspond to the breeding opportunities and quality of larval habitats that each biotope provides. In the rainforest of Central Africa, the presence of *An. gambiae* s.l. is strictly dependent on the presence of human activities that clear the forest canopy and expose bare soil for larval habitats to occur; pristine rainforest is not a suitable habitat for the larvae of these mosquitoes. Thus, size of human settlements is taken as a proxy for available breeding opportunities, whereby small rural settlements are generally poorer than urban centres in breeding opportunities, and organic enrichment of larval habitats in denser settlements may increase the bioavailability of nutrients. Similarly, nutrients such as dissolved phosphate, nitrate and silicate can accumulate at coastal river mouths and alluvial plains through drainage and river transport and can lead to eutrophication of surface waters (e.g. Turner & Rabalais, 1994). Thus, we assume resource stress to increase along the urban < rural and coastal < inland environmental gradients (Fig. 6A).

Beyond the zero isoclines, population growth is negative, and stable population persistence is not possible except as sink populations. Region ‘C’ in Fig. 6A defines the set of values in which osmotic stress is too high for *An. gambiae* despite the abundance and quality of resources (i.e. lower resource stress); this is where *An. coluzzii* occurs alone. Coastal urban habitats are presumed to be mostly representative of this condition. Region ‘G’ defines areas where
larval habitats are osmotically favourable but resources too few and poor for *An. coluzzii* to persist stably in the presence of *An. gambiae*; this region in state space is postulated to correspond to inland rural areas in the forest where *An. gambiae* occurs alone. Conversely, in the ‘C+G’ region, both environmental stressors are within the limits of positive population growth for both species; here, *An. coluzzii* and *An. gambiae* can compete for the same resources. In this region of state space, density dependence is likely to modulate the outcome of interspecific competition.

Because of little empirical knowledge, we consider here just the simple case of density-dependent competition modelled by the classic Lotka–Volterra equations (Fig. 6B). In rural coastal areas, where populations are smaller due to the lower carrying capacity of this biotope, and the environment is coarse-grained (*sensu* Rosenzweig, 1981) due to a broader spectrum of alternative – marginal – larval habitats (Figure S2), both species can occur in sympatry (stable equilibrium point at $K_0$). Larval habitat selection and character displacement in salinity tolerance contribute to the stable persistence of both species because they reduce interspecific competition; in particular, here *An. coluzzii* is prone to segregate in osmotically more stressful habitats than *An. gambiae*. As human settlements get larger the carrying capacity of the environment increases ($K_0 > K_0$); this can potentially shift the stable equilibrium point towards higher densities of *An. gambiae* (*e.g.* in the smaller urban habitats). Finally, in large and dense urban settlements, where the carrying capacity is at its greatest ($K_0 > K_0$), and the larval habitats are osmotically fine-grained (giving less opportunities for the stabilizing effect of habitat selection), the equilibrium point becomes unstable, so that competitive exclusion produces allopatry and competitive release in salinity tolerance (Figs 5C and 6B).

**Ecological divergence, phylogeography and speciation**

This study provides further evidence of ecological divergence and associated phenotypic differences between *An. coluzzii* and *An. gambiae*. Two novel major environmental gradients along which the species segregate (*i.e.* distance from the coastline and altitude) have been identified, and their contribution quantitatively integrated in an SDM providing a global map of the predicted relative frequency of *An. coluzzii* and *An. gambiae*. The ‘bimodality’ of the *An. coluzzii* distribution pattern suggests that its coastal and savanna populations may be evolving along independent trajectories. Previous population genetic studies have noted a clear subdivision within *An. coluzzii*, between populations derived from southern coastal Cameroon and populations derived from the West African savannas of Burkina Faso and Mali (Slotman et al. 2007; Lee et al. 2009a). This has led to the view that *An. coluzzii* is structured between West Africa and the Central African rainforest. In accordance with more extensive recent population genetic data (Pinto et al. 2013), our ecological modelling and physiological assays suggest an alternative view that explains the genetic structure equally well: subdivision of *An. coluzzii* is not necessarily between West and Central Africa, but rather between populations occupying the xeric savanna belt in the northern hemisphere, and populations occupying the ribbon of coastline along the Gulf of Guinea (Yawson et al. 2007). If this hypothesis proves correct, it suggests that the region of rampant and apparently recent hybridization between *An. coluzzii* and *An. gambiae* in The Gambia and Guinea Bissau (Caputo et al. 2008, 2011; Oliveira et al. 2008; Marsden et al. 2011; Nwakanma et al. 2013) may represent contact between savanna populations of *An. gambiae* and coastal populations of *An. coluzzii*, not between *An. gambiae* and the better studied northern savanna *An. coluzzii* populations. Only careful reconstruction of the phylogeographic relationships of these taxa across their entire distribution, information sorely lacking at present, will allow this situation to be further clarified.

**Conclusions**

Future applications of the quantitative predictive model of habitat segregation and qualitative theoretical framework of species interactions can help address longstanding questions, such as whether ‘mosquitoes select habitats’ or ‘habitats select mosquitoes’ (Harrison 1986; Nosil et al. 2005). Rather than behavioural preference, our data suggest that differences in physiological performance, modulated by environmental-dependent competition, may explain the observed patterns of habitat segregation.

As found in previous studies (Diabate et al. 2008; Lehmann and Diabate 2008; Gimonneau et al. 2010, 2012a,b; Tene Fossog et al. 2013), the ecology of larval stages appears to provide a fertile field of investigation to identify life history traits underlying divergence between *An. coluzzii* and *An. gambiae*. By integrating knowledge about spatially explicit ecological patterns as well as physiological and behavioural responses, we hope to gain a better understanding of the processes that are at the heart of *An. coluzzii* and *An. gambiae* evolution. Ultimately, the aim is to answer questions about how natural selection shapes phenotypic divergence, and what is its role in speciation in this model system. In light of their roles as malignant vectors of malaria, the relevance of evolutionary diversification in this system is amplified by its consequences for disease transmission.

**Acknowledgements**

Our gratitude goes to the inhabitants of the sampled locations and the administrative authorities of Cameroon,
Equatorial Guinea and Gabon for their collaboration and permission to carry out this study, as well as the staff of the French Embassy in Malabo, the OCEAC Malaria Research Laboratory in Yaoundé, the CIRMF in Franceville and Bouygues Contractors and Somague in Equatorial Guinea for their precious assistance during field work. We are especially grateful to Jean-Pierre Agbor, Serge Donfack, Sylvie Kemleu and Jean-Claude Toto for technical support and to Drs Xavier Garde, Richard Oslishy, Philippe Le Gall, Thierry Baldet, Isabelle Morlais for their help during the course of this work. Funding was provided by the Institut de Recherche pour le Développement, the National Institutes of Health (Projects R01 AI63508, and R01 AI76584 to NJB) and the Wellcome Trust through a fellowship (RYC-2012-11970). INAM was supported by a fellowship from the Universidad Nacional de Guinea Ecuatorial.

Data archiving
Data for this study are available at the Dryad Digital Repository: http://doi.org/10.5061/dryad.61m2s.

Literature cited


Habitat segregation in *Anopheles gambiae*

Tene Fosso et al.


Tene Fossog et al.

Habitat segregation in *Anopheles gambiae*


Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Calibration of the minimal adequate SDM mapped in Figure 3A.

**Figure S2.** Coefficients of Variation (CV) of electrical water conductivity (µS/cm) in *Anopheles gambiae* s.l. breeding sites from the forest ecozone in Central Africa.

**Table S1.** Geographical distribution of the breeding sites from which *Anopheles gambiae* and *An. coluzzii* larvae were collected for the salinity tolerance assays.

**Appendix S1.** Statistical analysis of the ecological factors associated with differences in larval salinity tolerance of *Anopheles gambiae* s.l. populations from Central Africa.