



OPEN

## Coupling genetic structure analysis and ecological-niche modeling in Kersting's groundnut in West Africa

Mariam Coulibaly<sup>1,2</sup>, Rodrigue Idohou<sup>3</sup>, Félicien Akohoue<sup>1</sup>, Andrew Townsend Peterson<sup>4</sup>, Mahamadou Sawadogo<sup>2</sup> & Enoch Gbenato Achigan-Dako<sup>1</sup>✉

Orphan legume crops play an important role in smallholder farmers' food systems. Though less documented, they have the potential to contribute to adequate nutrition in vulnerable communities. Unfortunately, data are scarce about the potential of those crops to withstand current and future climate variations. Using *Macrotyloma geocarpum* as an example, we used ecological niche modeling to explore the role of ecology on the current and future distributions of genetic populations of Kersting's groundnut. Our findings showed that: (1) the models had good predictive power, indicating that *M. geocarpum*'s distribution was correlated with both climatic and soil layers; (2) identity and similarity tests revealed that the two genetic groups have identical and similar environmental niches; (3) by integrating the genetic information in niche modeling, niches projections show divergence in the response of the species and genetic populations to ongoing climate change. This study highlights the importance of incorporating genetic data into Ecological Niche Modeling (ENM) approaches to obtain a finer information of species' future distribution, and explores the implications for agricultural adaptation, with a particular focus on identifying priority actions in orphan crops conservation and breeding.

Defining how and where plant varieties will adequately respond to environmental variations is a central topic in plant science research. This is more preoccupying for orphan crops that are largely grown in marginal areas and neglected in the mainstream research agenda<sup>1</sup>. Availability of the genetic resources of those crops is still a challenge for many genebanks at national, regional, and international levels. This situation jeopardizes the sustainable utilization of the plant genetic diversity that can be useful for current and future food systems and secured nutrition<sup>2-4</sup>. Such genetic resources are important for successful cultivars development and selection of economic and agronomic traits and could confer resilience to evolving climate.

In Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal and Baudet], a multipurpose staple orphan crop with high nutritional and economic values for smallholder farmers in West Africa<sup>5-9</sup>, the need to solve the ecological suitability of the extant genetic resources arose despite the significant achievements made recently on the germplasm collection, conservation, and characterization. Kersting's groundnut plays an important role in farming sustainability through its ability to fix atmospheric nitrogen in the soil and enhance soil fertility<sup>10</sup>. Furthermore, It serves in traditional medicine for local populations<sup>6,7,11</sup>. However, the production of Kersting's groundnut is declining rapidly and the genetic resources were rarely collected and safeguarded for the future generation. In addition, environmental stresses are among the main causes for declining Kersting's groundnut production from its cultivated areas<sup>5,12</sup>. Though Kersting's groundnut has relatively good adaptation to low-input conditions<sup>13,14</sup>, increased frequency of drought, intense precipitations, elevated temperatures, and increased salt and heavy metals in soils will often be accompanied by increased infestation by pests, and pathogens, are expected to limit the plant growth and productivity, and consequently the crop's yield and production<sup>15</sup>.

<sup>1</sup>Laboratory of Genetics, Biotechnology and Seed Science (GBioS), Faculty of Agronomic Sciences, University of Abomey-Calavi, 01 BP 526 Abomey-Calavi, Republic of Benin. <sup>2</sup>Laboratory of Biosciences, Faculty of Earth and Life Science, University of Ouaga I Pr. Joseph Ki-Zerbo, 03 BP 7021 Ouagadougou, Burkina Faso. <sup>3</sup>Laboratory of Biomathematics and Forest Estimates, Faculty of Agricultural Sciences, University of Abomey-Calavi, BP 1493 Abomey-Calavi, Republic of Benin. <sup>4</sup>Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA. ✉email: e.adako@gmail.com

Recent studies revealed a low variation within the species<sup>10,16</sup> that limits the extent of its genetic diversity and cultivated zones. Kersting's groundnut counts six landraces set mostly within three agroclimatic zones; Northern-Guinean (NG), Northern-Sudanian (NS), and Southern-Sudanian (SS) of Benin, Burkina Faso, Ghana, and Togo with the predominance of genetic resources and diversity in the Southern-Sudanian zone<sup>5,12</sup>. Overall, the area of cultivation and adaptation of landraces differ among agroclimatic regions. The Black landrace was largely collected in the Northern-Sudanian environmental conditions and was widely preferred, cultivated, and maintained by farmers<sup>12</sup>. The White landrace was widely grown in the Northern-Guinean transition zone of Benin<sup>5,6</sup> while less cultivated in Burkina Faso<sup>12</sup> and absent in other countries of West-Africa. The production of the Brown landrace was specifically limited to the Ghana farming system<sup>12</sup>.

Kersting's groundnut landraces are the direct results of farmer selection, cultivation, and maintenance over the centuries. This continual adaptation of the crop to smallholders farming conditions could continue to play a role in adapting production to climate change. Also, local adaptation of landraces could vary in their climatic response and requirement and therefore, may spread differentially under evolving environmental conditions<sup>17</sup>. To find an adequate preferendum where the species can thrive, it has become crucial to approximate the potential distribution of the crop and its genetic resources.

Unfortunately, with the rapid evolution in climate conditions and the further introduction and adoption of new cash crops with high economic importance, local seed systems alone will likely be insufficient to ensure the endurance of the crop genetic resources and diversity. In these conditions, applying ecological research is required to inform conservation and management decisions to mitigate a species genetic erosion<sup>18</sup>, as Kersting's groundnut at National and Regional levels. Ecological niche modeling (ENM) can identify the environmental parameters that can impact a species' distribution and project its potential distribution area onto new environmental surfaces to examine the effect of present or future environmental change<sup>19,20</sup>.

Several statistical and mechanistic techniques proved effective in quantifying niches and spatial distribution of natural and cultivated species<sup>21–25</sup>. The basic modeling framework of species distribution models (SDMs) in general has been criticized on a number of gaps, such as ignoring heterogeneity in population and genetic structure in different parts of a species geographical range<sup>26</sup>. However, many species are organized into differentiated genetic lineages across their geographical ranges<sup>27,28</sup> and populations differ in their adaptive potential to respond to environmental change<sup>29</sup>. Studies proved that incorporating molecular data into SDMs represents an important step forward for modeling the effects of climate change on species geographical ranges<sup>30–32</sup>.

In the case of Kersting's groundnut<sup>33</sup>, much uncertainty remains concerning the ability of the crop to withstand the changing climate, suggesting that there is a clear need to comprehensively analyze the response of the crop diversity under new environmental conditions of the coming decades.

The present study was undertaken to determine the environmental conditions in which KG will need to be cultivated in the coming decades and to utilize this information to prioritize genetic resource conservation and breeding efforts. Recent molecular studies involving 281 individuals from Benin and Togo identified two major genetic clusters of KG and these two groups were distributed across Southern-Sudanian and Northern-Guinean agroclimatic zones<sup>16</sup>. Kafoutchoni, et al.<sup>34</sup> also assessed the genetic structure of the species through the GBS approach with 217 individuals, and Discriminant analysis of principal components (DAPC) and found eight genetically distinct groups from five origins. In this context, the following questions are of high interest: do the agroclimatic niches of KG vary with climate changes? Would KG genetic groups differ in their ability to respond to present and future climatic scenarios?

This study examines the response of orphan crops to future climates by using genetic information and ecological niche modeling approach (gENMs) using KG as an example. Therefore, we combined KG population genomics data with ecological niche modeling: (1) to analyze the relationship between climate factors and species populations distribution in agroclimatic zones of Burkina Faso, Benin, Ghana, and Togo, and (2) to predict areas that would be suitable for the species and genetic populations under the future scenarios. We hypothesized that: (1) the future climates will impact the distribution of KG, and (2) genetically distinct populations of orphan crops would respond differently to climate change.

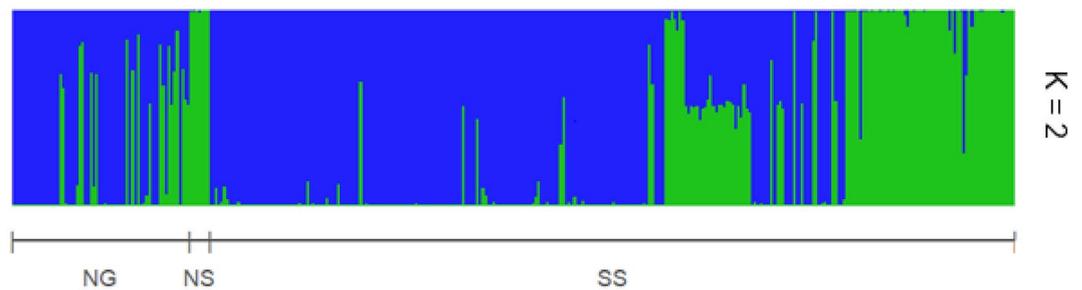
## Results

**Population structure and admixture.** Despite low levels of diversity ( $He=0.021$  and  $Ho=0.0053$ ,  $p\text{-value}\leq 0.001$ ), Kersting's groundnut populations remained genetically well differentiated. Admixture models with a putative number of tested genetic clusters (K) from one to five, showed that the most likely number of inferred members was 2 with  $\Delta K=101.917$  (Fig. S1: Evanno output plots; Table S1). The classification of the 361 accessions into populations based on the model-based structure (Fig. 1) showed that more than 65% of individuals belonged to the Pop1 (in blue colour,  $N=231$ ).

Table 1 showed the distribution of populations across agroclimatic zones. The majority of Kersting's groundnut individuals were collected in the Southern-Sudanian zone (< 80%); 83.117% and 75.385% of individuals of Pop1 and Pop2, respectively collected in this zone. All the accessions collected in the Northern-Sudanian zone belonged to the Pop2. In the Northern-Guinean zone, 19.481% of accessions belonged to Pop1 and 14.615% were included in Pop2.

The random distribution of landraces (based seed coat colour, Fig. 2) into the genetic populations indicated that Pop1 was mostly characterized by the White landrace (97.403% of individuals in Pop1). On the other hand, the Pop2 was composed of all the six landraces included in this study with a predominance of Black (43.846%) and Brown ones (30%).

The analysis of genetic distance revealed relatively strong genetic differentiation among the two distinct groups of KG with a pairwise  $F_{st}$  value of 0.583 (Table 2). The estimated values of gene diversity (Hs) showed an overall low genetic divergence between individuals within each population (0.017). Meanwhile, the heterozygosity



**Figure 1.** Structure diagram showing the delineation of Kersting's groundnut individuals into two genetic populations ( $K = 2$ ), Pop1 (Blue) and Pop2 (Green). Vertical lines represent individuals within populations and those with more than one color share genetic information with other populations. The horizontal line shows the distribution of the populations across agroclimatic zones (Northern-Guinean NG, Northern-Sudanian NS, and Southern-Sudanian SS).

Group by	Features of populations	Pop1		Pop2	
		N	Freq (%)	N	Freq (%)
Landraces seeds coat color	Black	12	5.195	49	37.692
	Brown greyed orange eye	0	0	39	30
	Red	2	0.866	17	13.077
	White	223	96.537	4	3.077
	White black eye	0	0	14	10.769
	White greyed orange eye	0	0	1	0.769
Agroecological zones	Northern-Guinean	2	19.48	19	14.615
	Northern-Sudanian	0	0	7	5.385
	Southern-Sudanian	192	83.117	98	75.385
Totals		237	100	124	100

**Table 1.** Number (N) and proportions (Freq) of accessions for each genetic Group per Landrace and agroclimatic zone.



**Figure 2.** Kersting's groundnut seed coat colors. (a) White mottled with greyed orange eye (b), Black, (c) White, (d) Red, (e) White mottled with black eye, (f) Brown mottled with greyed orange eye.

Source of variation	Df	SS	MS	Variation (%)	Pairwise Fst	He	Ho	Hs	P-value
Within pop	3	248.276	82.756	4.958		0.021	0.0053	0.017	<0.001
Between pop	1	2879.945	2879.945	47.895	0.583				<0.001
Pop1						0.005		0.004	<0.001
Pop2						0.029		0.029	<0.001

**Table 2.** Analysis of molecular variance (AMOVA), and genetic variation within and among populations. *DF* degree of freedom, *SS* sum of squares, *MS* expected mean squares, % percentage of the variance, *He* expected heterozygosity, *Ho* observed heterozygosity. *P*-value: significance tests after 1000 permutations.

analysis indicated that the genetic divergence between individuals in Pop2 was relatively greater comparing to Pop1.

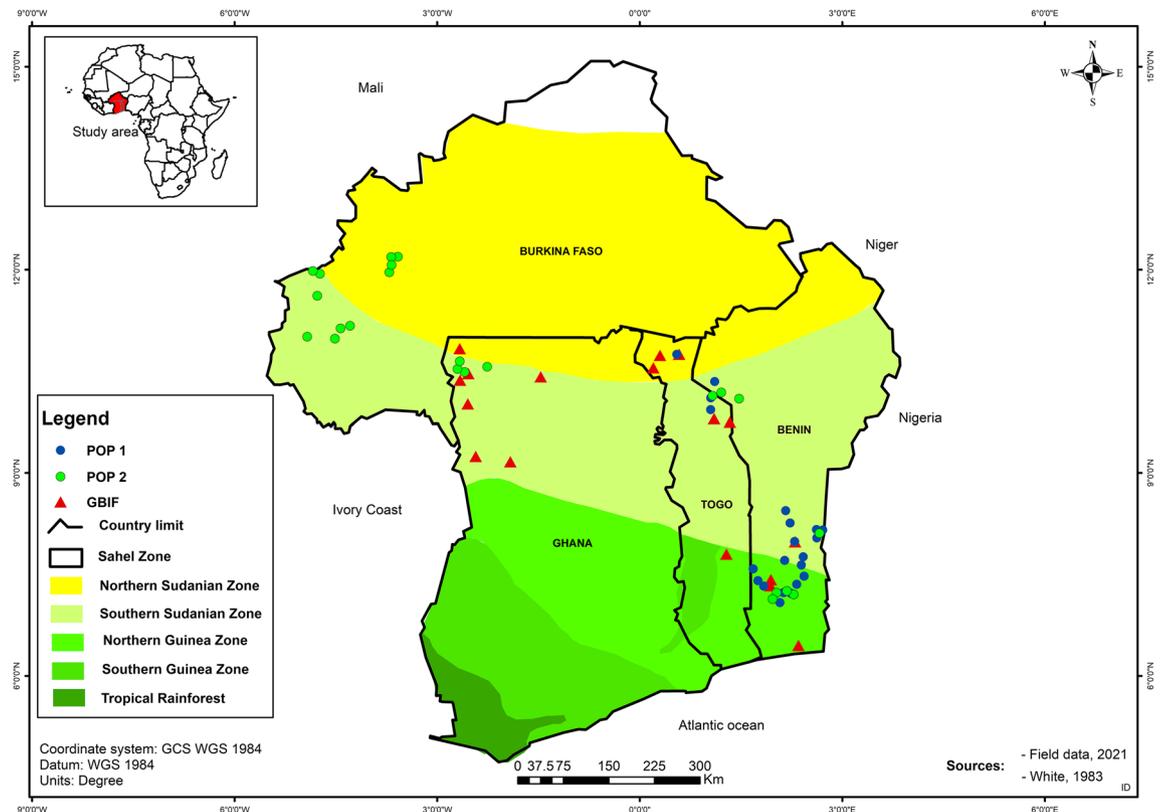
**Occurrences dataset construction.** Figure 3 shows Pop1 distributed across two agroclimatic zones, the Northern-Guinean and Southern-Sudanian zones whereas the Pop2 and Global Biodiversity Information Facility (GBIF) points were represented in all agroclimatic zones. At the country level, Pop2 was distributed across the four countries while Pop1 was located in Benin and Togo.

**Niche overlap and conservatism among *M. geocarpum* and genetic populations.** *Niche similarity between KG and Pop1.* The quantification of the niches of species (Fig. 4a) and genetic Pop1 (Fig. 4b) in environmental space indicated that more individuals were loaded in PC2 while low number on PC1. Furthermore, the principal Components Analysis (PCA) showed that the first two axes explained 68.82% of the environmental (E-space) variation (PC1 = 48.73% and PC2 = 20.09%) between the species and genetic Pop1. The difference between the E-space of the two groups is illustrated by Fig. 4c. Climate variables correlated with the two principal components were mean annual rainfall (bio12), rainfall driest month (bio14), bulk density in kg/cubic-meter for 10 cm depth (blt\_d2) and soil texture fraction clay at 10 cm depth (clyppt\_d2) (Fig. 4d: PC1 in horizontal axis and PC2 in vertical axis). The values of the Potential Niche Truncation Index (PNTI) were very low for the species (0.09, Fig. S2a) and Pop1 (0.03, Fig. S2b). These values are below the suggested range of values showing either moderate risk (0.15–0.3) or high risks (0.3) for which the observed niches do not represent the fundamental niches<sup>35</sup>. In addition, We found a non-significant niche equivalency test statistic ( $D = 0.311$ ,  $p = 0.990$ ), and significant background test statistic ( $p = 0.009$ , Fig. S2c–e).

*Niche similarity between KG and Pop2.* The quantified niches in environmental spaces of the species and Pop2 showed that the individuals of the species and Pop2 were mostly loaded on the PC1 (Fig. 5a,b). The difference between the E-space of the two groups is illustrated by Fig. 5c. The PCA analysis revealed that 66.55% of the variance (PC1 = 48.79% and PC2 = 17.76%) in environmental data input can be represented in a two dimensional E-space (Fig. 5d). Climate variables correlated with the first two axis included the max temperature warmest month (bio5), rainfall driest month (bio14), and bulk density in kg/cubic-meter for 10 cm depth (blt\_d2). For the species and genetic Pop2, the PNTI value was low and showed no niche truncation (Fig. S3a,b). Our results showed a non-significant niche equivalency test statistic ( $D = 0.342$ ,  $p = 0.871$ ) indicating identical environmental niches of *M. geocarpum* and Pop2 (Fig. S3c,d). We obtained a significant background test statistic ( $p = 0.014$ ) (Fig. S3e), which indicates that the species and genetic Pop2 niches were similar.

*Niche similarity between Pop1 and Pop2.* The quantification of the two genetic populations niches showed that the individuals were mainly loaded on PC1 axis (Fig. 6a,b) with Fig. 6c showing the difference between the E-space of the two groups. Moreover, when examining the results of PCA analysis, we found that the first two axes explained 67.79% of the (PC1 = 52.14% and PC2 = 15.65%) in environmental variations between the two genetic populations (Fig. 6d). The PCA correlation circle showed that max temperature warmest month (bio5), soil texture fraction clay at 10 cm depth (clyppt\_d2), bulk density in kg/cubic-meter for 10 cm depth (blt\_d2), and Soil texture fraction silt in percent for 10 cm depth (Sltppt\_d2) were the environmental variables correlated with the two principal components. The PNTI value ( $< 0.15$ ) indicated that the niches occupied by the two genetic populations represent the fundamental niches of these groups (Fig. S4a,b). We obtained a non-significant niche equivalency test statistic ( $D = 0.213$ ,  $p = 0.792$ ), and significant background test statistic ( $p = 0.015$ , Fig. S4c–e). These results indicates identical and similar environmental niches of the genetic populations.

**Ecological niche modeling.** *Climatic variables analogy.* Figure 7 shows the importance of the six selected environmental variables in terms of their contribution to each model. When examining the important variables associated with the occurrences data, the three most important variables for prediction models of *M. geocarpum* (Fig. 7a) and Pop2 (Fig. 7c) were the mean annual rainfall (bio12), rainfall driest month (bio14), and bulk density in kg/cubic-meter for 10 cm depth (blt\_d2). On the other hand, for Pop1 (Fig. 7b), soil texture fraction clay at 10 cm depth (clyppt\_d2), soil texture fraction silt in percent for 10 cm depth (Sltppt\_d2), and rainfall driest month (bio14) had the greatest contributions to the model. In contrast, clyppt\_d2 and Sltppt\_d2 were far less important for model 1 and model 3.



**Figure 3.** Map showing the distribution of Kersting's groundnut across agroecological zones of Burkina Faso, Benin, Ghana, and Togo: species collections points (circles) and the locations points downloaded from GBIF (triangles) used in the ecological niche modeling. Genetic groups are depicted by the blue (Pop1) and green (Pop2) colours. Software used: ArcGIS v. 10.7.1.

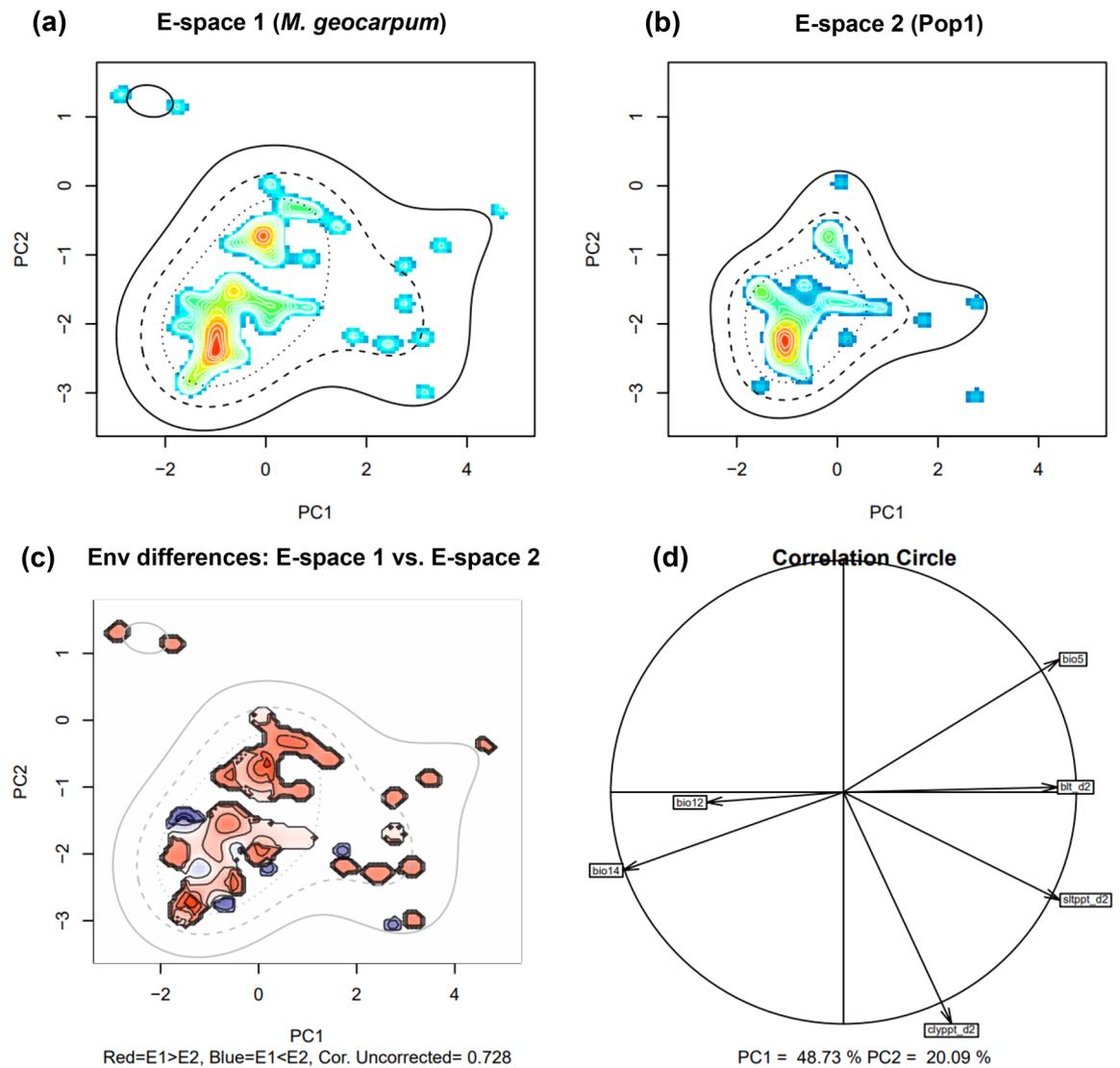
**Models' outputs and performance evaluation.** The results in Table 3 showed that the environmental predictive models displayed high predictive power for the three models. We found that the ENM for KG location points without genetic information was significant, with pROC scores of  $1.294 \pm 0.144$  (t-test,  $p < 0.001$ ). The Pop1 and Pop2 models had the highest pROC scores of  $1.755 \pm 0.115$  and  $1.525 \pm 0.110$ , respectively. These populations exhibited the best models performance with the lowest AICc scores (Pop1 = 571.670, and Pop2 = 696.185, Table 3). Additionally, these best-supported models had mean omission rates of 0.039 (Pop1) and 0.085 (Pop2) for the 10th percentile training presence. The AUC values were also high for the three niche models: *M. geocarpum* had a mean test AUC of  $0.916 \pm 0.027$ , the pop1,  $0.979 \pm 0.009$ , and Pop2,  $0.926 \pm 0.023$ , indicating that occurrences points were strongly differentiated from background locations, so model distributions were not random.

**Impact of climate change on Kersting's groundnut and genetic populations.** Based on the 10 percentile training thresholds (Table 3), the defined calibration areas (Fig. 8) revealed large areas with suitable conditions for the cultivation of KG and genetic populations in the present day (Table S2), except for Pop1. The ENMs predicted broader suitable areas for KG (21.578%) and Pop2 (23.696%), while suitable climatic conditions for Pop1 were restricted to a much smaller area (2.423%).

The Maxent model for the species predicted a large area of suitable conditions across the three agroclimatic zones, Southern Sudanian (SS), Northern Sudanian (NS), and Northern-Guinean (NG) (Fig. 8a1). The SS and NG zones were the areas forecasted to have high suitable climatic conditions for species production. For the Pop1, the areas predicted to have high likely cultivability conditions were concentrated in the NG zone of Benin, but very less and sparsely distributed in the SS zone (Fig. 8b1). The Pop2 was projected across the three studied agroclimatic zones of the four countries (Fig. 8c1).

Analyses based on projecting the final models into the future (2055) scenarios revealed varied patterns in KG and genetic populations cultivable areas (Fig. 8, Figs. S5, S6, and S7). Under the two future climatic scenarios RCP4.5 and RCP8.5, an increase in the species cultivable areas for about 91.137% and 16.567%, respectively were observed due to the decrease of the non-suitable areas (Fig. 8a2,a3, and Fig. S5). The expansion of this area was observed mainly in the NG zone further Southern of Benin, and Togo.. On the other hand, the NG zone of Southern and Central Benin and Togo, and the SS and NS zones of the four countries became climatically unsuitable to crop production.

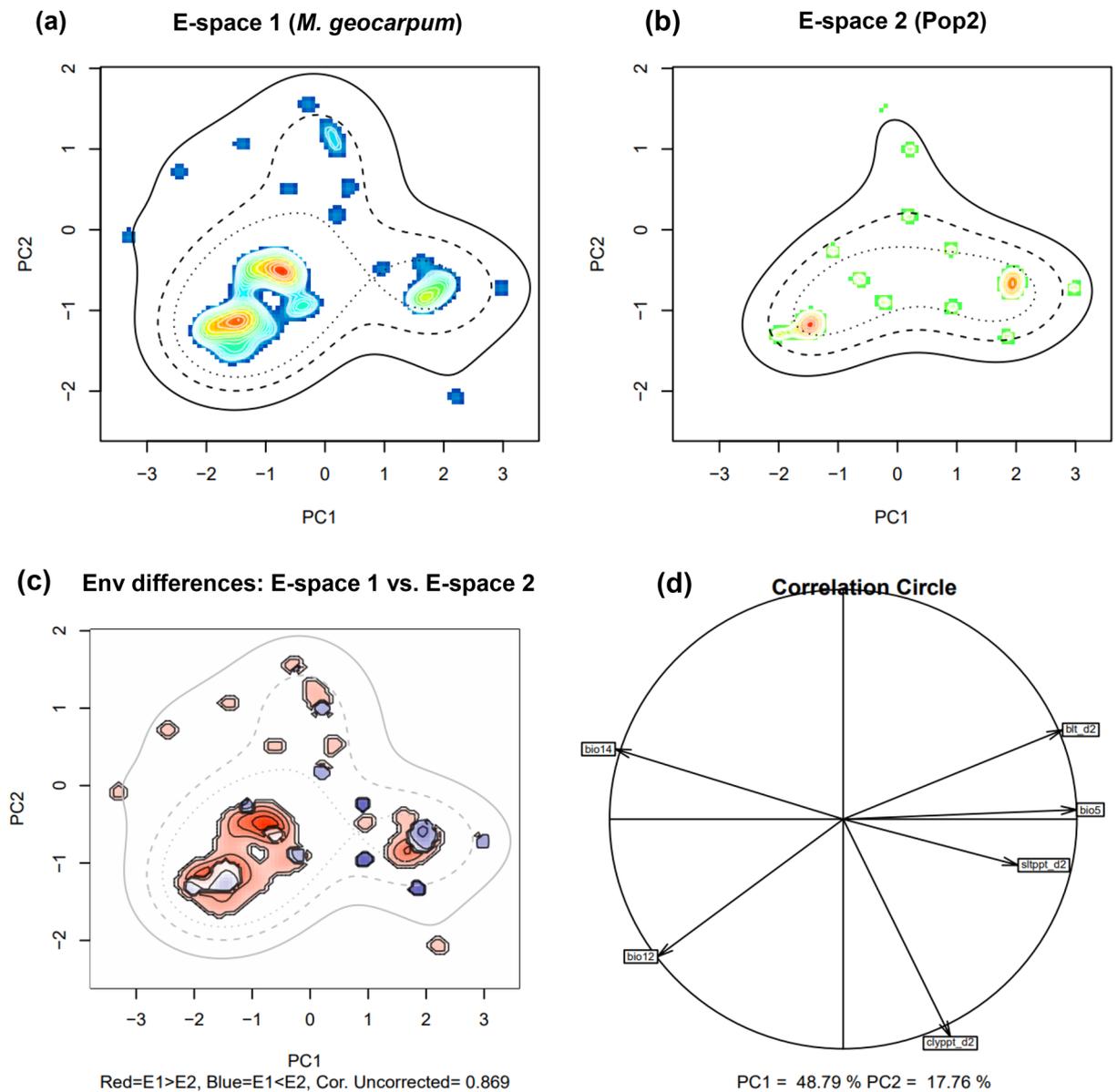
Important increases were observed in the potential cultivable areas of the Pop1 under future climatic scenarios (> 500% under RCP 4.5 and  $\approx 50\%$  under RCP 8.5) (Fig. 8b2,b3, and Fig. S6). This trend was observed across the



**Figure 4.** Niches of the two compared groups (*M. geocarpum* and genetic Pop1) in two dimensional E-space. Graphs (a) and (b) represent the niches of the species and Pop1, respectively along the first two axes of the PCA. The occurrences are represented by kernel density isopleths, red colour indicates high density and blue colour indicates low density. Solid and dotted contour lines illustrate 100% and 50% of the available background (environmental space). (c) is the difference in the E-space of the two groups, and Niche E-space Correlation Index (NECI). NECI determines if one should correct the occurrence densities of each group by the prevalence of their environments in their range for equivalency and background tests. For high NECI (>0.5) groups occupied niches are recommended to be corrected by the frequency of E-space in accessible environments to reduce the chances of committing type 1 errors, and (d) is the correlation circle based on the two principal components of the environmental input data. *bio5* max temperature warmest month, *bio12* mean annual rainfall, *bio14* rainfall driest month, *blt\_d2* bulk density in kg/cubic-meter for 10 cm depth, *clpypt\_d2* soil texture fraction clay at 10 cm depth, *Sltppt\_d2* soil texture fraction silt in percent for 10 cm depth.

three agroclimatic zones of the four countries under scenario RCP 4.5. However, under the scenario RCP 8.5, the cultivable areas decreased in the Northern and Southern Guinean zones of Ghana. There was also a slight decrease in the suitable areas of production in some areas of the Northern-Guinean and Southern-Sudanian zones Benin, Burkina Faso, and Togo. The Pop2 showed to be the more vulnerable to future scenarios as the suitable areas significantly decreased (65.758% under RCP4.5 and 76.941% under RCP8.5), while the unsuitable ranges increased (Fig. 8c2,c3, and Fig. S7). Meanwhile, this genetic group is projected to gain in very small suitable areas in the Northern Guinean zone further in the South of Benin, Togo, and Ghana under the severe scenario of RCP 8.5.

The MOP results indicated that the calibration environments were dissimilar from the extrapolative conditions in the two future scenarios of 2055 (Fig. 9). Figure 9 showed that the Sahel zone in Burkina Faso and the Northern-Sudanian climatic conditions of eastern Burkina Faso, and further North of Benin are unlikely to be suitable for the species cultivation in real life.

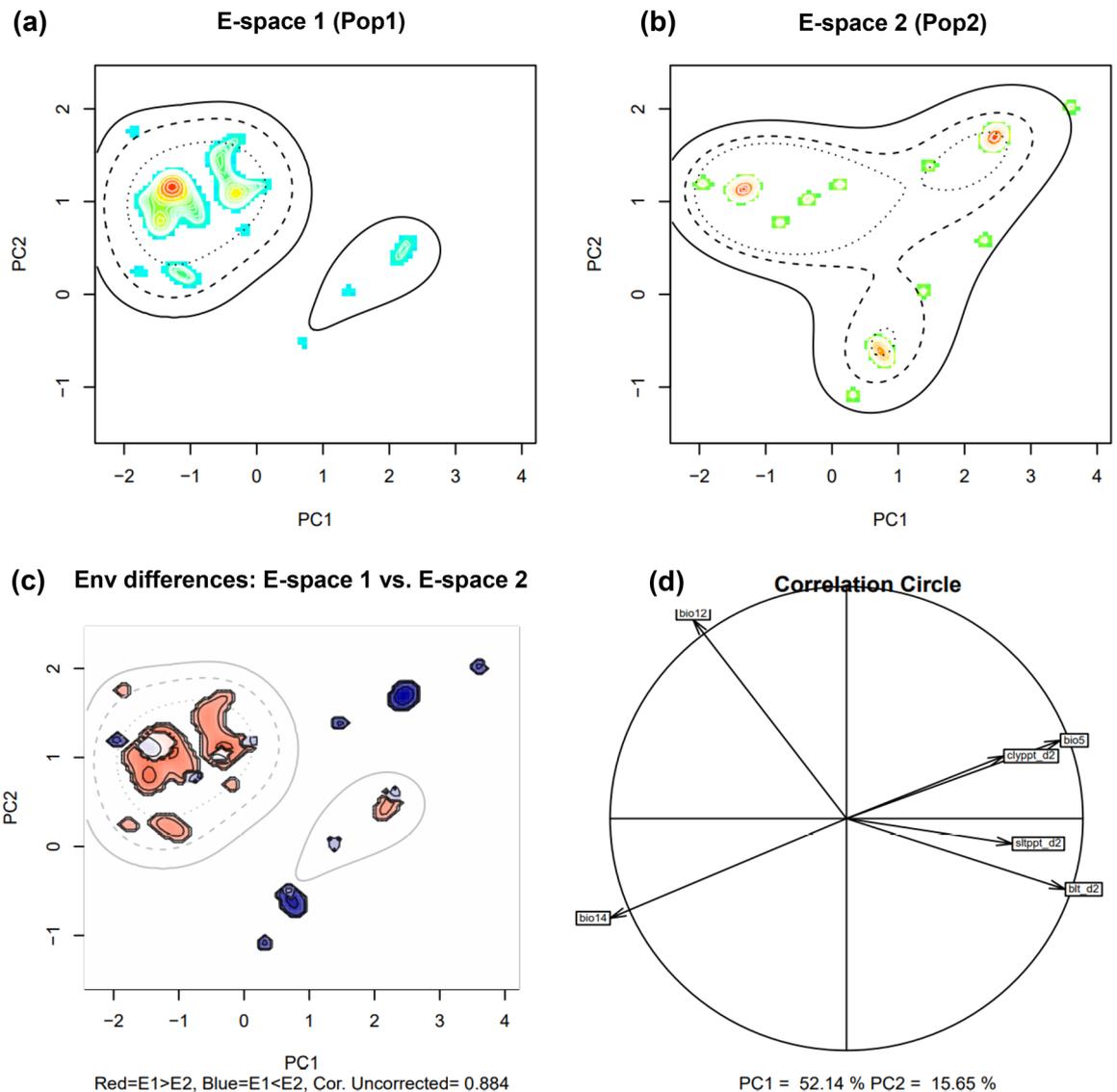


**Figure 5.** Niches of the two compared groups (*M. geocarpum* and genetic Pop2) in two dimensional E-space. Graphs (a) and (b) represent the niches of species and Pop2, respectively along the first two axes of the PCA. The occurrences are represented by kernel density isopleths, red colour indicates high density and blue colour indicates low density. Solid and dotted contour lines illustrate 100% and 50% of the available background (environmental space). (c): is the difference in the E-space of the two groups, and Niche E-space Correlation Index (NECI). NECI determines if one should correct the occurrence densities of each group by the prevalence of their environments in their range for equivalency and background tests. For high NECI (>0.5) groups occupied niches are recommended to be corrected by the frequency of E-space in accessible environments to reduce the chances of committing type 1 errors, and (d) is the correlation circle based on the two principal components of the environmental input data. *bio5* max temperature warmest month, *bio12* mean annual rainfall, *bio14* rainfall driest month, *blt\_d2* bulk density in kg/cubic-meter for 10 cm depth, *clpypt\_d2* soil texture fraction clay at 10 cm depth, *Sltppt\_d2* soil texture fraction silt in percent for 10 cm depth.

## Discussions

By incorporating genetic information in ENM, this study provided evidence that the two genetically distinct groups in Kersting's groundnut respond differently to climate changes though distributed in similar environmental niches, and suggested implications for sustainable management and breeding perspectives under current and future climate variations.

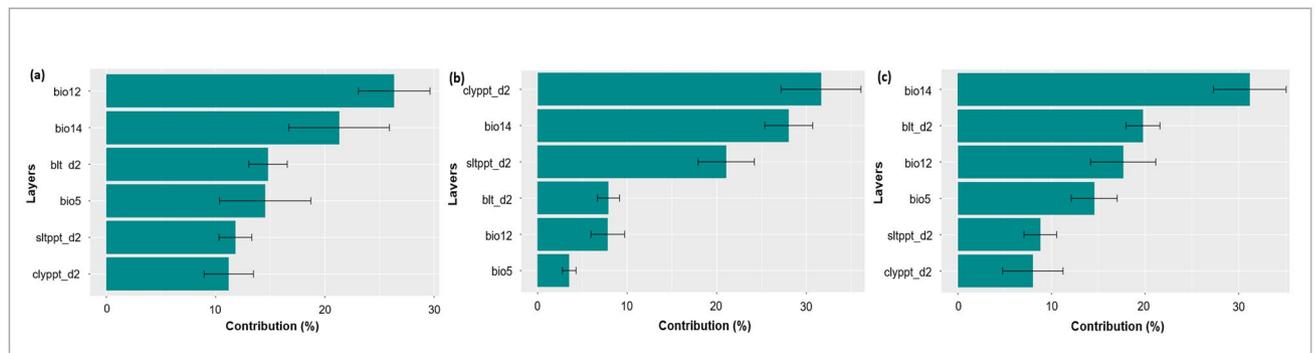
**Current potential distribution ranges of Kersting's groundnut and genetic groups.** We used two approaches to model KG potential environmental niches across agroclimatic zones of four West African



**Figure 6.** Niches of the two genetic populations in two dimensional E-space. Graphs (a) and (b) represent the niches of Pop1 and Pop2, respectively along the first two axes of the PCA. The occurrences are represented by kernel density isopleths, red colour indicates high density and blue colour indicates low density. Solid and dotted contour lines illustrate 100% and 50% of the available background (environmental space). (c): is the difference in the E-space of the two populations, and Niche E-space Correlation Index (NECI). NECI determines if one should correct the occurrence densities of each population by the prevalence of their environments in their range for equivalency and background tests. For high NECI ( $>0.5$ ) genetic populations occupied niches are recommended to be corrected by the frequency of E-space in accessible environments to reduce the chances of committing type 1 errors, and (d) is the correlation circle based on the two principal components of the environmental input data. *bio5* max temperature warmest month, *bio12* mean annual rainfall, *bio14* rainfall driest month, *blt\_d2* bulk density in kg/cubic-meter for 10 cm depth, *clyppt\_d2* soil texture fraction clay at 10 cm depth, *sltppt\_d2* soil texture fraction silt in percent for 10 cm depth.

countries: a traditional species ENM using species-wide cultivated range, and intraspecific models based on genetic information. We found that decomposing a species into intraspecific genetic groups increases our understanding of the potential distribution of KG, which is consistent with results from previous studies incorporating genetic information into ENM<sup>30,31,36,37</sup>.

**Environmental variables contribution.** Our approach allowed to identify six environmental variables correlated with the potential distribution ranges of the species and populations. The max temperature warmest month (*bio5*), mean annual rainfall (*bio12*), and rainfall driest month (*bio14*) were the climatic variables mostly involved in fitting the three models. The ecological weight that precipitations may have in KG's distribution is also supported by a recent ethnobotanical study where farmers identified drought and high moisture during the reproductive stage as the main factors hindering species production<sup>5,12</sup>. Tamini also showed that the different



**Figure 7.** Environmental contribution for the species (a) and genetic Pop1 (b) and Pop2 (c). *bio5* max temperature warmest month, *bio12* mean annual rainfall, *bio14* rainfall driest month, *blt\_d2* bulk density in kg/cubic-meter for 10 cm depth, *clyppt\_d2* soil texture fraction clay at 10 cm depth, *Sltppt\_d2* soil texture fraction silt in percent for 10 cm depth.

Species/population	Occ	pROC	AUC	Omission rate	AICc	TPLT
<i>M. geocarpum</i>	53	1.294 ± 0.144	0.916 ± 0.027	0.093	1362.357	0.313
Pop1	24	1.755 ± 0.115	0.979 ± 0.009	0.039	571.67	0.350
Pop2	26	1.525 ± 0.110	0.926 ± 0.023	0.085	696.185	0.386

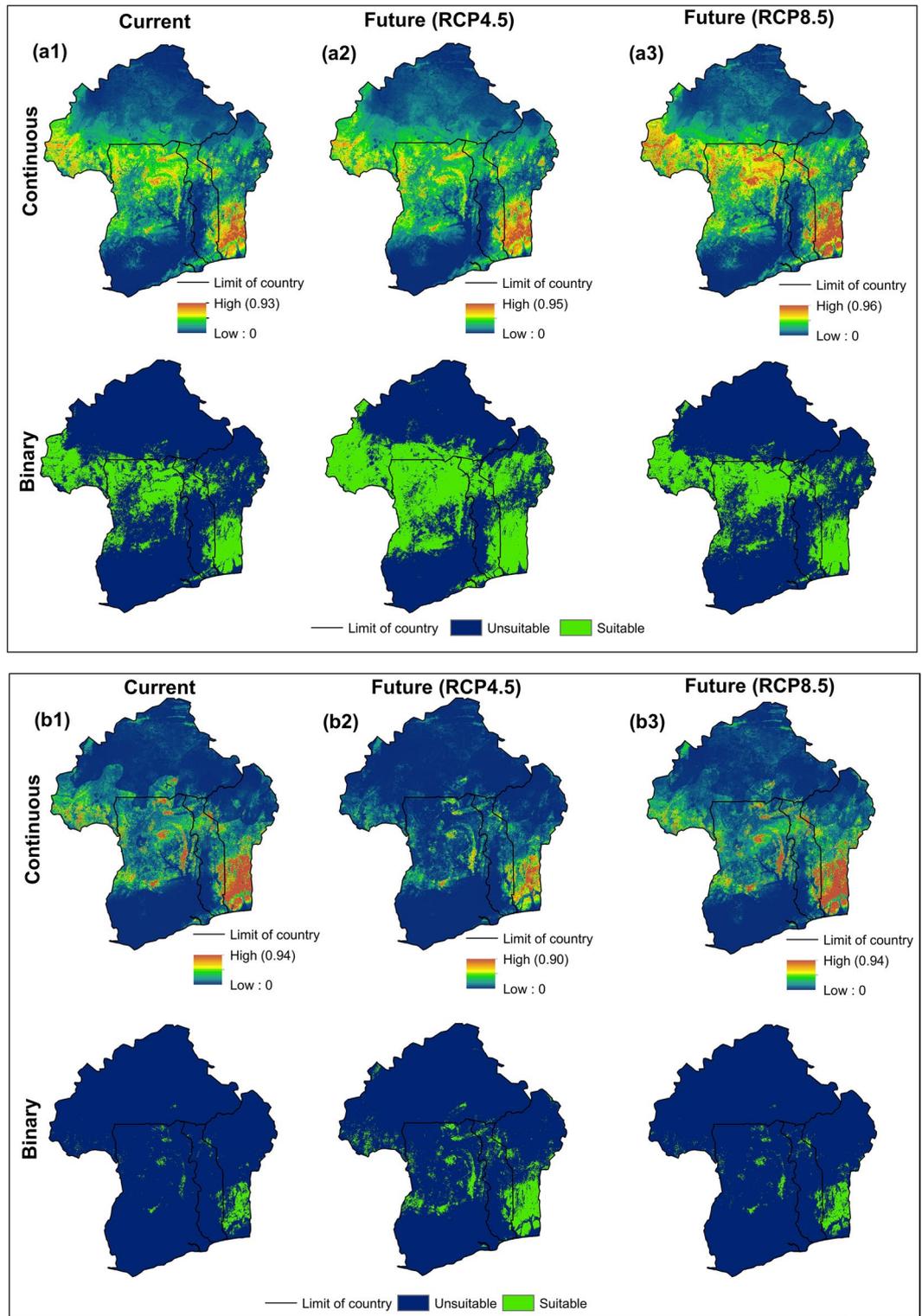
**Table 3.** Maxent results showing the parameters measuring the three models' performance and the partial receiver operating characteristic curve (pROC) for each model. *Occ* number of occurrences, *pROC* partial receiver operating characteristic curve, *AUC* area under the receiver operating characteristic curve, *AICc* corrected Akaike Information Criterion, *TPLT* 10 percentile Training Plogistic Threshold.

sowing dates and seasons with varied temperatures influenced KG growing cycle by delaying or accelerating the flowering time. To cope with these weather issues, farmers managed their cropping calendar by advancing (in June) or delaying (in August) the date of planting KG. The change in temperature and moisture levels are not only determined by climatic parameters but may also lead to a change in the absorption rate of water, fertilizers, and other minerals in soils, which determine yield output<sup>38,39</sup>. Therefore, the consideration of both climate and soil type was essential to better predict KG and genetic groups' distribution. The physical properties of soil such as bulk density of soil in kg / cubic-meter for 10 cm depth (*blt\_2*), soil texture fraction clay at 10 cm depth (*clyppt\_d2*), and soil texture fraction silt in percent for 10 cm depth (*Sltppt\_d2*) also play a relevant role in KG and populations distribution. Kouelo et al.<sup>40</sup> reported that the soil texture and preparation (type of tillage) influenced the crop productivity in Benin. Although applying ENM to intraspecific genetic groups allowed the detection of environmental variables, a multidisciplinary approach involving agronomists and biologists would allow a clear understanding of the weight that each of these components may have on crop growth and development.

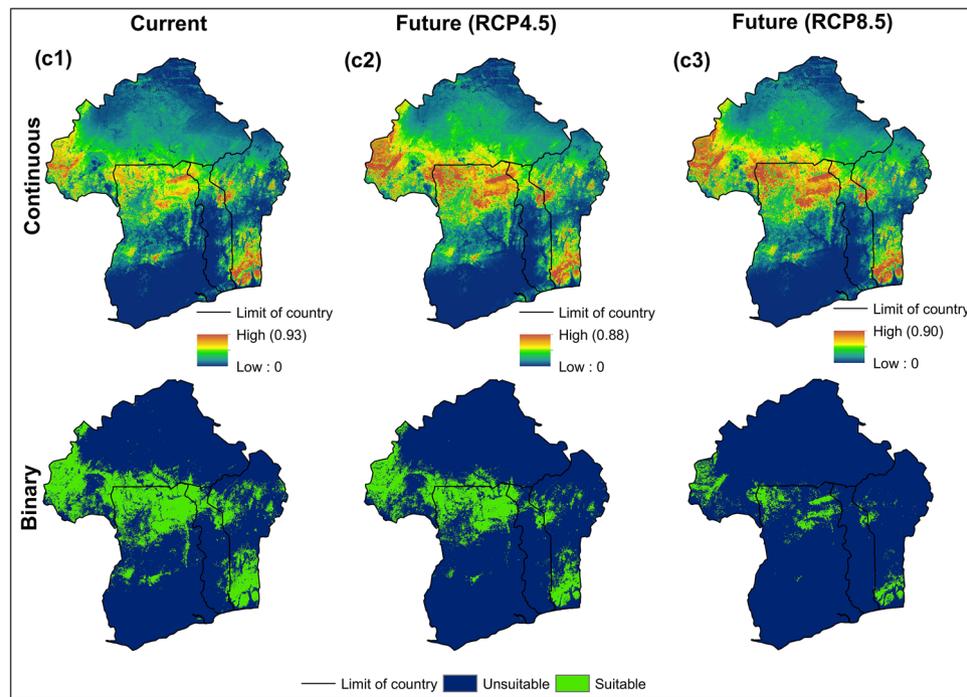
**Niche equivalency and similarity.** We found identical and similar environmental niches for the species and the two genetic populations, suggesting no variation in their environmental niches and adaptation to climate conditions. Our results did not support the view that a given species could be considered as an assemblage of genetic units differing in their spatial distribution<sup>31,36</sup>. Populations of KG are both cultivated in the Northern-Guinean and the Southern-Sudanian zones of Benin, characterized by bimodal (rainfall ≤ 1500 mm) and unimodal (rainfall ≤ 1100 mm) growing seasons, respectively. Although Pop 2 is grown also in the Northern-Sudanian zone, where rainfalls are relatively low (rainfall ≤ 900 mm), very few number of occurrences (7) from this region were used in the model. Therefore, they may not show variations in their environmental niches and adaptation to different climate conditions, as previously reported by Wellenreuther, et al.<sup>41</sup> who analyzed the ecological causes of the ranges limits and the coexistence of two congeneric damselflies (*Calopteryx splendens* and *C. virgo*).

In contrast, Maxent's outputs revealed that climatic niches of the two genetic populations of *M. geocarpum* are relatively different in the distribution areas. The species and genetic populations occupied the distribution ranges but do not respond similarly to the environmental conditions. Indeed, KG individuals displayed difference in their performance when evaluated under the same environmental conditions<sup>42–44</sup>.

**Predicting evolutionary change based on genetic variation.** Our results, found evidence of variability in the species response to future environmental conditions, confirming the assumption that changes in climate will influence the environmental niches of KG species and populatons. This corroborates many other studies which showed the impact of future climatic conditions on crops production, such as cereals, legumes, vegetables<sup>22–24,45,46</sup>.



**Figure 8.** Binary and continuous maps showing the spatial distribution of *M. geocarpum* (a1–a3) and genetic Pop1 (b1–b3) and Pop2 (c1–c3) in present and the future days. Each model was set to thresholds with the 10 percentile training presence values to produce continuous and binary raster maps for current and future scenarios in ArcGIS v. 10.7.1.

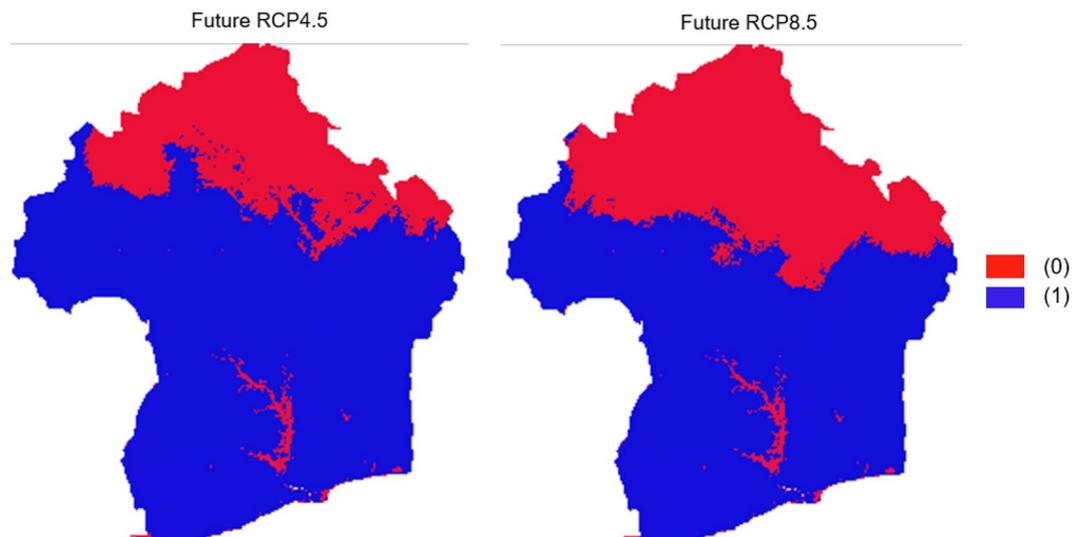


**Figure 8.** (continued)

We also combined genetic information to ENM to test the hypothesis that genetic divergent groups would respond differently to environmental change. The projection under future climatic scenarios RCP4.5 and RCP8.5 indicate that the response of Kersting's groundnut and genetic groups varied significantly across agroclimatic zones. Our results confirmed the idea that different genetic clusters potentially showed adaptive variation to different abiotic conditions within the geographical range of the species. Globally, although Central Benin and Northern Togo are presumed to be a centre of origin for the crop, we observed a loss in suitable areas for the species production. The same trend was observed in other countries where the crop has been grown for thousand of years. Similarly, Ikeda et al.<sup>31</sup> tested and demonstrated the hypothesis that species locally adapted to current environments are likely to become maladapted in the future. The same trends were reported on other cereals such as maize, wheat, sorghum, and barley which tend to decrease their area of suitability in their centres of origin<sup>47</sup>. Out of the 361 individuals included in this study, 101 belong to the genetic Pop2, and counts all the landraces grown—diversity based on seed coat color<sup>5,12</sup>. According to the conclusion of Vigouroux, et al.<sup>48</sup> who described the evolution and adaptation of pearl millet in West Africa, such diversity was supposed to confer adaptation to climate variation. The southern and Northern-Sudanian benefit of the lowest precipitations and are hence, the most prone to drought, widely reported as one of the main factors hindering KG production<sup>49</sup>. In the context of an inevitable agricultural reduction in these agroclimatic zones, the earlier the mitigation actions are taken, the more successful will be the collection and conservation efforts of KG's resources.

On the other hand, future environmental conditions remain, even become more favorable to the Pop1 cultivation (with an increase in suitable areas of 588.812% compared to the present day). However, under the severe environmental conditions (RCP8.5), the cultivable area of this genetic Pop1 is predicted to decline from RCP4.5 (−89,802.711 km<sup>2</sup>) in the Northern and Southern-Guinean zones of Ghana, Benin and Togo. With regards to these results, significant headway can be made by creating more favorable policy environments. Two strategies proposed by Ramirez-Cabral, Kumar and Taylor<sup>23</sup> can be used to mitigate loss in KG resources and diversity: first, the use of all the resources of Pop2 in regions where the stress of climate conditions become unsuitable for the species production. An example of this is the Southern-Sudanian zone of Central Benin, where the use of KG is more important than other regions. The second option was to address future loss of suitability by shifting Pop2's current cultivation areas to new regions, such as the Northern-Guinean zone of Southern Benin, Ghana and Togo where future climatic conditions are projected to become more cultivable for KG.

**Conservation implications.** Successful management of an endangered species often depends on the accurate identification of current and future cultivable areas. Thus, we combined genetic diversity and ecological niche modeling to understand the evolutionary dynamics in Kersting's groundnut species across its cultivated regions. Although the predictive models showed that future climatic conditions will be more favorable to KG production, the Sudanian and some areas of Northern-Guinean zones were identified as climatically unsuitable for the species. Notwithstanding the above results, this approach may provide a valuable tool for genetic resources managers for implementing collection and conservation strategies of this orphan legume, as sustained by Sohn et al.<sup>50</sup> in their study on the endemic bird in South America. In particular, our suggestion is to focus on the Pop2 with particularly high conservation value. This genetic group is represents the genetic diversity includ-



**Figure 9.** Output of the mobility-oriented parity analysis of KG from 'ntbox'; 0 (Red) represents strict extrapolation areas meaning complete dissimilarity between calibration environments and extrapolation regions in the future; 1 (Blue) area similar to areas where the models have been calibrated. Software used: ArcGIS v. 10.7.1.

ing all the coloured Landraces grown by farmers, which is unvaluable resources for the plant genetic improvement. These resources are still managed by farmers through an informal seed system. However, the local seed system alone will likely be insufficient to adapt the crop production to changing climate. In doing so, the formal seeds management efforts can improve the decision-making process that would greatly help in the collection and conservation of the genetic Pop1 at higher risk of extinction but also the Pop1 predicted to thrive with climate change. In general, African crop genetic resources conservation is unfortunately poorly supported by National Governments, and material from the regions is not fully represented in the important international genebanks which afford the baseline for sustained public breeding efforts<sup>45</sup>. In many areas where KG is produced, its production is declining gradually and consequently followed by the lost in genetic resources. Hence, many of these regions would stand to benefit from the genetic resources of other areas of cultivation, if these resources can be effectively managed and shared. Perhaps more importantly, the international cooperation on genetic resources conservation and use, between countries where KG is produced, will be crucial in adapting the crop to the imminent threats of climate change.

**Implications of ENM results for selection for adaptation.** Undertaking any global initiatives to overcome food insecurity challenges requires an assessment of staple crop productivity, quality, and environmentally suitable areas under climate change<sup>23</sup>. KG is a staple food crop and a source of proteins and nutrients for smallholder farmers in West Africa. In this research, we expected to identify the intraspecific genetic group that would be resilient under future climate scenarios, and which areas are concerned by changes in KG diversity. The model projections indicate that a shift in Kersting's groundnut productive areas is slightly and likely with a loss of cultivability of the Pop2 cultivation areas and an increase for Pop1. Although KG is predicted to gain in a suitable area in future environmental changes, its cultivation is, however, decreasing year to year from its cultivated areas<sup>5,51</sup>. Therefore, it will be crucial to adapt *M. geocarpum* species to the increasingly challenging environmental conditions through the development of new resilient cultivars that meet farmers' needs. The Pop2 comprises all Landraces (diversity, mainly the coloured landraces) of the species and has been reported having the best phenotypic performance<sup>42</sup>. The individuals in this genetic population can serve as a potential source for diversity on which breeding efforts could be based to confer best performance for plant growth and yield to Pop1, mainly composed of the white landrace mostly preferred, grown, and sold by farmers in Benin<sup>5,6</sup>. Similarly, the genetic Pop1 can serve as potential parents to confer resilience to changing climates and increase sustainability of genetic Pop2. In addition, advancements in molecular plant breeding would be an importance of paramount to increase the genetic gains and make more accurate the breeding process. In the particular case of KG, currently, available partial GBS data<sup>16,34</sup> would allow the analyses of intraspecific genetic clusters based on gene network variation for various important phenotypic traits (e.g.: grain yield, yield-related traits, flowering time). This novel approach would provide possibilities to assess the extent to which key functional genes and genetic variation may be threatened under future ecological conditions<sup>52</sup>. Plant breeders have widely and successfully implemented genome-wide breeding approach for the development of climate-resilient varieties<sup>53–55</sup>, through marker-assisted selection and genomic selection. Another approach is to increase variability within the species, particularly in Pop1, clustering mostly white seed coloured individuals, through mutation methods (using physical or chemical mutagens) combined with molecular markers (Targeted Induced Local Lesions in Genomes (TILLING)). Such techniques have been successfully used in breeding of many legume crops to enhance diversity and to develop mutant cultivars<sup>2,56,57</sup>.

## Conclusion and perspectives

This study assessed the potential impacts of climate variations on environments suitability for Kersting's groundnut cultivation, and consequently its distribution around four West-African countries. The use of Maxent' models and genetic information allowed a preliminary understanding of the stress factors influencing the climate suitability of the species and genetic populations under two future scenarios (2055, RCP4.5 and RCP8.5). The overall trend shown by our results indicates an increase in climate suitability for the species cultivation in the Northern-Guinean zone of Southern Benin and Togo. However, an important decrease was predicted in other agroclimatic areas, while genetic Pop1 production areas increase. Our findings illustrate also that projected areas of environments cultivability for *M. geocarpum* and the two populations are on different climate change trajectories. The projected distribution maps presented in this paper have been hence, analysed and used to identify strategic measures to manage the impacts of reduced climate cultivability while taking advantage of the opportunities in areas of improved suitability for Kersting's groundnut cultivation in the future. Our findings could be downscaled to a country level to assist national policymakers in developing strategic control initiatives to prevent the scarcity of this legume.

Although this study represents the first insight into examining the potential of Kersting's groundnut as a resilient crop under climate change, several limitations are to be noted to develop the right tools to reduce model uncertainty and make better predictions in future research. First, the global distribution of the species and its wild relatives is still incompletely documented online: collecting more and finer occurrence data, especially in regions where its production was previously reported would greatly help in refining or confirming our results. Second, our models identified rainfall, temperature, and soil variables that contributed significantly to their fits. However, socioeconomic factors such as the local market value of the species and cultural preferences must also be considered in the predictions. Combining also measures of key phenotypic traits and biotic stress factors (pests and diseases associated with KG production) in the modeling process would contribute to improving predictions of the impact of climate change on this legume crop. Third, the non-availability of whole genome-wide data in the case of Kersting's groundnut limited access to phylogenetic information and identification of key functional genes for various important phenotypic traits. That may provide means to assess the response of key functional genes under evolutionary climate change.

Nevertheless, using our approach, we identified species, genetic populations, and cultivable areas for further germplasm collecting to enhance available germplasm and better direct Kersting's groundnut breeding priorities in the future.

## Materials and methods

**Genotyping and genetic clustering.** For the present work, a total of 361 accessions of KG collected from Benin, Burkina Faso, Ghana and Togo were used. The DNA of each sample was extracted from young leaves of each accession, using the protocol of the Integrated genotyping service and support (IGSS) at the Biosciences Eastern and Central Africa (BecA: <http://hub.africabiosciences.org/activities/services>) located in Nairobi, Kenya. The quality of the DNA was confirmed by electrophoresis in 0.8% agarose, and the quantification was carried out using UVP BioDoc-It2 Imaging System. All of the samples were diluted to 50 ng/μl for the DArT genotyping platform. Genotyping was performed using the DArT-Seq™ platform at Diversity Arrays Technology<sup>58</sup>. The quality analysis of the genotypic data was performed using Illumina HiSeq 2500<sup>59</sup>. The SilicoDArt calling algorithms (DArTsoft14) were used to score DArTseq markers into a binary format (presence = 1 and absence = 0) for each sample genomic representation. A total of 2844 SNP markers were obtained and processed in TASSEL v5<sup>60</sup>, for quality check. SNPs were filtered with TASSEL v5 for further analysis using the proportion of missing data < 20%. A total of 2323 SNPs were retained and the Nipals model in kdcompute (<https://kdcompute.igss-africa.org>) was used for data imputation.

The program Structure 2.3.4<sup>61</sup> was used to assign individuals to different genetic clusters based on the admixture model. The population structure was performed based on the Bayesian clustering approach using the following settings: correlated allelic frequencies, burn-in period of 20,000 and 20,000 Markov Chain Monte Carlo (MCMC) interactions; and grouping (K) ranging from 1 to 5 in 10 independent runs. The results generated were used as input into the POPHELPER version 2.3.1, an R package<sup>62</sup>, to determine the most likely genetic clusters based on the Evanno method<sup>63</sup> and to generate the delta K ( $\Delta K$ ) graph and bar graph. The most likely number of K clusters was determined by its higher  $\Delta K$  value. After these searches, we found a likely number of K = 2.

The molecular variation (AMOVA) was determined using the R package adegenet<sup>64</sup> to test for statistically significant differences between the two genetic groups. The pairwise Fst comparisons between the two groups were calculated with 9999 permutations using package 'hierfstat' v 0.5–10 in R environment. General patterns of genetic diversity were also evaluated by calculating observed ( $H_o$ ) versus expected ( $H_e$ ) heterozygosity and gene diversity ( $H_s$ ) within each population. Based on the genetic analyses, we classified the two clusters as distinct genetic populations (Pop1 and Pop2) encompassed within three agroclimatic zones of Burkina Faso, Benin, Ghana, and Togo (northern-Sudanian, southern-Sudanian, and northern-Guinean). We then used these genetically defined populations and their GPS points available, to define populations' location points.

**Occurrence data.** The occurrence data for *M. geocarpum* was obtained from the Global Biodiversity Information Facility (GBIF, [www.gbif.org](http://www.gbif.org)), an online available database, and the crop self-collected material of the Laboratory of Genetics, Biotechnology, and Seeds Science (GBioS). As the different data sources and a large dataset (> 500 occurrence records) would likely carry elevated geographical or environmental space biases<sup>65,66</sup>, the number of records were decreased in Wallace package, an online workspace based on R interface<sup>67</sup> using four complementary approaches: (1) we first removed occurrences collected before 1986 to match with environmental layers and soil properties; (2) considerable ambiguity may exist in GBIF data over the identity of the

species due to synonymous names (*M. geocarpum* var. *geocarpum*, *M. geocarpum* var. *Tisserantii*; *Kerstingiella geocarpa*, *Kerstingiella tisserantii*). To avoid any confusion arising from this taxonomic ambiguity, we searched through the online databases using the following keywords: *Macrotyloma geocarpum*, *Kerstingiella geocarpa*, var. *geocarpa*, or var. *geocarpum*; orphan legumes. We then harmonized the GBIF database and discarded the reports on var. *tisserantii* and *Kerstingiella tisserantii*; (3) we used spatially filtering occurrences located  $\leq 10$  km from other occurrences using the *spThin*<sup>68</sup>, an R package; finally, (4) we manually checked isolated locations points in Africa (in ArcGIS ver. 10.7.1) and removed occurrences in areas where *M. geocarpum* is not generally grown.

The defined genetic clusters data with their geographic coordinates were also filtered separately to ensure the real distribution of each population within agroclimatic zones. After data cleaning steps, the final occurrence datasets contained: for the species, a dataset of 53 occurrences (which include 17 points retained from GBIF), used for modeling without genetic information. And two data files that comprised 24 locations points for Pop1, and 26 occurrences for Pop2, used in subsequent analyses integrating genetic information (see Supplemental Table S3).

**Environmental variables.** We used bioclimatic layers combined with soil properties to project current and future niches for the species and each genetic group. A total of 15 bioclimatic variables were downloaded from Africlim online regional climate models (RCMs) data portal (<https://webfiles.york.ac.uk/KITE/Africlim/>)<sup>69</sup>. Current and future variables averaged between the periods 1986–2015 (2000) and 2041–2060 (2055) were downloaded at a 30 arc-s ( $\sim 1$  km) spatial resolution. For future climatic conditions, predictions from the Ensemble model<sup>69</sup> were used. This model simulates changes based on a set of scenarios. The projections were run under Representative Concentration Pathway (RCP), RCP 4.5, and RCP 8.5 for the 2055 time horizon<sup>70</sup>. In all RCPs, the climatic conditions are extreme in RCP 8.5 scenarios compared to RCP 4.5. RCP 4.5 projects temperatures to rise above industrial levels by at least 1.5 °C in West Africa, with atmospheric CO<sub>2</sub> reaching 500 ppm while in RCP 8.5 projections, temperatures are predicted to rise by 2.8 °C and atmospheric CO<sub>2</sub> to be over 550 ppm<sup>71</sup>. These climate projections were statistically downscaled to match the bioclimatic variables using the delta method<sup>72</sup>.

Data related to soil characteristics were available in the World Soil Information (ISRIC) databases (Soil-property-maps-of-Africa-at-250-m-resolution) at 250 m resolution<sup>73</sup>. These spatial predictions of soil properties were generated based on two predictive approaches such as random forests and linear regression<sup>73</sup>. Soil characteristics included 30 variables related to the soil's physical, chemical, and nutritional properties. Soil data were then converted to 30 arcseconds using ArcGIS software v 10.7.1 to match with bioclimate layers. Finally, using shapefile boundaries of four West African countries (Benin, Burkina Faso, Ghana, and Togo) we cropped all variables to encompass the broad geographic regions that define Kersting's groundnut global distribution.

We discarded highly correlated variables ( $\geq 0.8$ ) by Pearson's rank correlation coefficients analysis using the R library *ntbox*<sup>74</sup> (see correlation matrix in Supplementary Table S4). A total of 18 variables were retained as less correlated. We then conducted a few preliminary model runs without genetic information (with species occurrences) in Maxent v. 3.4.4<sup>75</sup> to reduce the number of variables to be included in the prediction models<sup>75</sup>. For each run, we removed variables with the lowest contribution. Six bioclimatic variables including Max Temperature warmest month (bio5), Mean annual rainfall (bio12), Rainfall driest month (bio14), Bulk density in kg/cubic-meter for 10 cm depth (blt\_d2), soil texture fraction clay at 10 cm depth (clyppt\_d2), = Soil texture fraction silt in percent for 10 cm depth (Sltppt\_d2) were chosen based on their noncollinearity and contribution to the prediction models and were used in the final models of the species, and with genetic information (Pop1 and Pop2).

**Ecological predictive models' development and evaluation.** To calibrate our models, we employed the maximum entropy method<sup>75,76</sup> implemented in Maxent ver. 3.4.4<sup>75</sup>. The algorithm has been extensively tested and benchmarked<sup>77,78</sup>. Many studies have reported Maxent as one of the highest performing presence-background algorithms<sup>21,79</sup>. As the selection of sample points can influence model performance in Maxent<sup>80</sup>, we restricted the selection of background points using the regularization of 10,000 background points<sup>21</sup>. Models were trained with data from the present and projected in the future. Three Maxent models were generated with different occurrence datasets:

- model 1: all accessions location points together without genetic information were used in projecting the entire distribution of KG, distributed in four West African countries (Burkina Faso, Benin, Ghana, and Togo) where the distribution range of the species is characterized by three agroclimatic zones including the Northern-Sudanian, Southern-Sudanian, and Northern-Guinean.
- model 2 and model 3: were developed using occurrences of Pop1 and Pop2, respectively the genetically defined populations. Pop1 location points are covered by the agroclimatic zones of Southern-Sudanian, and Northern-Guinean of Southern and central Benin, and Northern Togo. On the other hand, the occupation range of Pop2 extends over the four countries and the three agroclimatic zones.

We used a Bootstrap method, by defining 75% of the data for model training and 25% for model testing for 10 iterations<sup>76</sup>. We kept the Auto Features box checked Linear, quadratic, product, hinge, and threshold functions of predictor variables were employed, and variable importance was assessed using a jackknife analysis<sup>79</sup>. For easy interpretation of the results of niche models, we specified the Maxent' output format to a logistic form.

The three models performance was tested based on the partial receiver operator characteristics (pROC) and the corrected Akaike information criterion (AICc) approaches. We used the outputs from Maxent to perform pROC in R package 'ENMGadgets' v 0.1.0.1<sup>81</sup>, by specifying 500 iterations with the omission threshold set at five percent. The statistical significance of pROC values was examined using t-test statistic. The AICc was estimated based on the number of parameters and likelihoods of continuous raw outputs, using R library ENMEval v

2.0<sup>82</sup>, following the method described by Warren and Seifert<sup>83</sup>. The lower is the value of AICc, more accurate is the predictive model<sup>84</sup>. Models omission rates calculated with the 10th percentile training presence threshold, and the area under the receiver operating characteristic curve (AUC), were also used as secondary criteria for testing performance of each model<sup>66,85</sup>. A model is considered as having a good fit when its AUC is closed to one (AUC ≥ 0.75)<sup>21</sup>. The final models were projected onto future climatic scenarios (RCP4.5 and RCP8.5) of the horizon 2055, using Maxent software with ten bootstrap iterations. Each model was set to thresholds with the 10 percentile training presence values to produce continuous and binary raster maps for current and future scenarios in ArcGIS v. 10.7.1.

Two different levels were, therefore, defined: unsuitable and suitable. Finally, we quantified in km<sup>2</sup> the surface occupied by each condition (suitable and unsuitable) using the continuous raw outputs from maxent in ArcGIS v. 10.7.1. We estimated the range size change under scenarios RCP 4.5 and RCP 8.5 of the period 2055 for the species and two genetic populations, following the model used by Hu et al.<sup>86</sup>:

$$\Delta(\%) = \frac{(FA - CA)}{CA} \times 100$$

where, FA corresponds to the future suitable areas (in km<sup>2</sup>) under a given future scenario; CA is the suitable areas in current conditions. Negative, null, and positive values represent range lost, stable, and gained, respectively. Furthermore, to visualize the potential changes of suitable areas for Kersting's groundnut production, we compared current and future distribution ranges of the crop and genetic populations using package "tmap" version 3.3-1<sup>87</sup> in R. Finally, to determine the zones where each model transfers would require extrapolation under future climatic conditions, we used the mobility-oriented parity (MOP) metric. The MOP analysis was performed with the R library 'ntbox' by setting for environmental distances to the nearest 10% of the reference region and 500 runs.

**Niche differentiation.** To understand whether differences between species and genetic populations emerge from true niche divergences, we performed Schoener's niche equivalency (identity) and similarity test (D), and Warren's niche background test (I) using the R package 'humboldt'<sup>35</sup>. Niche equivalency test is used to test out the null hypothesis that two species have identical environmental niches. The niche equivalency test examines the observed niche similarities between the predicting models of two species. The niche background test evaluates the power to detect differences between two ENMs. Compared to the existing tests, these methods improve the accuracy of niche similarity quantifications and corresponding statistical tests<sup>35</sup>. The Schoener's D statistic is used to determine how similar the occupied niches of two species are based on the original input occurrences. The observed values of D in both the niche identity test and the background similarity test were compared to the mean values of the randomized runs using t-test<sup>88</sup>. Hence, the environmental niches were considered significantly different if the observed values of niche overlap were less than 95% or 99% (alpha = 0.05 and 0.01, respectively) of the overlap values derived from the pseudoreplicates.

We visualized and analyzed E-space as two axes of a Principal Component Analysis (PCA) of input environmental variables throughout the entire study areas of the species and genetic populations. Broennimann et al.<sup>89</sup> used a kernel density function<sup>90</sup> to create a continuous E-space surface in a grid of 100 × 100 cells, estimating the occupied E-space of the focal species or its environment, respectively, using PC values from either the input occurrence localities or study region data.

Furthermore, The species fundamental niche from observed niche was analyzed with the package 'humboldt'<sup>35</sup>. The fundamental niche was characterised by truncating species occupied E-space by the available E-space in its environment. The larger the truncated proportion, the greater the risk that the occupied niche does not accurately reflect fundamental niche of the species. Hence, the Potential Niche Truncation Index (PNTI) was estimated to measure the amount of the observed E-space truncated by the available E-space. When the values of PNTI are ranged between 0.15 and 0.3, there is a moderate truncation risk while values > 0.3 explained a high risk of niche truncation due to limited available E-space<sup>35</sup>.

## Data availability

The molecular data underlying this article (DarTSeq data) will be submitted to NCBI.

Received: 29 July 2021; Accepted: 14 March 2022

Published online: 04 April 2022

## References

- Mabhaudhi, T. et al. Prospects of orphan crops in climate change. *Planta* **250**, 695–708. <https://doi.org/10.1007/s00425-019-03129-y> (2019).
- Singh, M., Bisht, I. S., Dutta, M., Springer. India, 221. <https://doi.org/10.1007/978-81-322-2023-7> (2014).
- Litrico, I. & Violle, C. Diversity in plant breeding: A new conceptual framework. *Trends Plant Sci.* **20**, 604–613. <https://doi.org/10.1016/j.tplants.2015.07.007> (2015).
- Govindaraj, M., Vetriventhan, M. & Srinivasan, M. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genet. Res. Intern.* **431–487**, 2015. <https://doi.org/10.1155/2015/431487> (2015).
- Akohoué, F., Sibiyi, J. & Achigan-Dako, E. G. On-farm practices, mapping, and uses of genetic resources of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal et Baudet] across ecological zones in Benin and Togo. *Genet. Resour. Crop. Evol.* **66**, 195–214. <https://doi.org/10.1007/s10722-018-0705-7> (2018).
- Assogba, P. et al. Indigenous knowledge and agro-morphological evaluation of the minor crop Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal et Baudet] cultivars of Benin. *Genet. Resour. Crop. Evol.* **63**, 513–529. <https://doi.org/10.1007/s10722-015-0268-9> (2015).
- Adu-Gyamfi, R., Fearon, J., Bayorbor, T. B., Dzomeku, I. K. & Avornyo, V. K. The Status of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Marechal and Baudet]. *Outlook Agric.* **40**, 259–262. <https://doi.org/10.5367/oa.2011.0050> (2011).

8. Obasi, M. O. & Agbatse, A. Evaluation of nutritive value and some functional properties of Kersting's groundnut seeds for optimum utilisation as a food and feed resource. *E. Afr. Agric. For. J.* **68**, 173–181. <https://doi.org/10.4314/eaaf.v68i4.1794> (2003).
9. Ajayi, O. B. & Oyeyayo, F. L. Potentials of Kerstingiella geocarpa as a health food. *J. Med. Food* **12**, 184–187. <https://doi.org/10.1089/jmf.2008.0100> (2009).
10. Mohammed, M., Jaiswal, S. K., Sowley, E. N. K., Ahiabor, B. D. K. & Dakora, F. D. Symbiotic N<sub>2</sub> fixation and grain yield of endangered Kersting's groundnut landraces in response to soil and plant associated bradyrhizobium inoculation to promote ecological resource-use efficiency. *Front. Microbiol.* **9**, 1–14. <https://doi.org/10.3389/fmicb.2018.02105> (2018).
11. Tamini, Z. Étude ethnobotanique de la Lentille de Terre [*Macrotyloma geocarpum* Maréchal & Baudet] au Burkina Faso. *J. Agric. Trad. Bot. Appl.* **37**, 187–199. <https://doi.org/10.3406/jatba.1995.3569> (1995).
12. Coulibaly, M., Agossou, C. O. A., Akohoué, F., Sawadogo, M. & Achigan-Dako, E. G. Farmers' preferences for genetic resources of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal and Baudet] in the production systems of Burkina Faso and Ghana. *Agronomy* **10**, 1–20. <https://doi.org/10.3390/agronomy10030371> (2020).
13. AchiganDako, E. G. & Vodouhe, S. R. *Macrotyloma geocarpum* (Harms) Marechal & Baudet. In *Plant Resources of Tropical Africa I: Cereals and Pulses* (ed. Brink, M. B. G.) 111–114 (Backhuys Publishers CTA, PROTA, 2006).
14. Mergeai, G. Influence des facteurs sociologiques sur la conservation des ressources phylogénétiques: Le cas de la lentille de terre [*Macrotyloma geocarpum* (Harms) Marechal et Baudet] au Togo. *Bull Rech Agron* **28**, 487–500 (1993).
15. Long, S. P., Marshall-Colon, A. & Zhu, X. G. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56–66. <https://doi.org/10.1016/j.cell.2015.03.019> (2015).
16. Akohoue, F., Achigan-Dako, E. G., Sneller, C., Van Deynze, A. & Sibiyi, J. Genetic diversity, SNP-trait associations and genomic selection accuracy in a west African collection of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Marechal & Baudet]. *PLoS ONE* **15**, 1–24. <https://doi.org/10.1371/journal.pone.0234769> (2020).
17. Schierenbeck, K. A. Population-level genetic variation and climate change in a biodiversity hotspot. *Ann. Bot.* **119**, 215–228. <https://doi.org/10.1093/aob/mcw214> (2017).
18. Araújo, M. B., Whittaker, R. J., Ladle, R. J. & Erhard, M. Reducing uncertainty in projections of extinction risk from climate change. *Glob. Ecol. Biogeogr.* **14**, 529–538. <https://doi.org/10.1111/j.1466-822X.2005.00182.x> (2005).
19. Araujo, M. B. & Peterson, A. T. Uses and misuses of bioclimatic envelope modeling. *Ecology* **93**, 1527–1539. <https://doi.org/10.1890/11-1930.1> (2012).
20. Martínez-Meyer, E. Climate change and biodiversity: Some considerations in forecasting shifts in species' potential distributions. *Biodiv. Inform.* **2**, 42–55. <https://doi.org/10.17161/bi.v2i0.8> (2005).
21. Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151 (2006).
22. Pironon, S. *et al.* Potential adaptive strategies for 29 sub-Saharan crops under future climate change. *Nat. Clim. Chang.* **9**, 758–763. <https://doi.org/10.1038/s41558-019-0585-7> (2019).
23. Ramirez-Cabral, N. Y. Z., Kumar, L. & Taylor, S. Crop niche modeling projects major shifts in common bean growing areas. *Agric. For. Meteorol.* **218–219**, 102–113. <https://doi.org/10.1016/j.agrformet.2015.12.002> (2016).
24. Syfert, M. M. *et al.* Crop wild relatives of the brinjal eggplant (*Solanum melongena*): Poorly represented in genebanks and many species at risk of extinction. *Am. J. Bot.* **103**, 1–17. <https://doi.org/10.3732/ajb.1500539> (2016).
25. Blonder, B. Hypervolume concepts in niche- and trait-based ecology. *Ecography* **41**, 1441–1455. <https://doi.org/10.1111/ecog.03187> (2018).
26. Hampe, A. & Petit, R. J. Conserving biodiversity under climate change: The rear edge matters. *Ecol. Lett.* **8**, 461–467. <https://doi.org/10.1111/j.1461-0248.2005.00739.x> (2005).
27. Leimu, R. & Fischer, M. A meta-analysis of local adaptation in plants. *PLoS ONE* **3**, 1–8. <https://doi.org/10.1371/journal.pone.0004010> (2008).
28. Hereford, J. A quantitative survey of local adaptation and fitness trade-offs. *Am. Naturalist.* **173**, 579–588. <https://doi.org/10.1086/597611> (2009).
29. Shaw, R. G. & Etterson, J. R. Rapid climate change and the rate of adaptation: Insight from experimental quantitative genetics. *New Phytol.* **195**, 752–765. <https://doi.org/10.1111/j.1469-8137.2012.04230.x> (2012).
30. Gotelli, N. J. & Stanton-Geddes, J. Climate change, genetic markers and species distribution modelling. *J. Biogeogr.* **42**, 1577–1585. <https://doi.org/10.1111/jbi.12562> (2015).
31. Ikeda, D. H. *et al.* Genetically informed ecological niche models improve climate change predictions. *Glob. Change Biol.* **23**, 164–176. <https://doi.org/10.1111/gcb.13470> (2016).
32. Alvarado-Serrano, D. F. & Knowles, L. L. Ecological niche models in phylogeographic studies: Applications, advances and precautions. *Mol. Ecol. Resour.* **14**, 233–248. <https://doi.org/10.1111/1755-0998.12184> (2014).
33. Thoen, M. P. *et al.* Genetic architecture of plant stress resistance: Multi-trait genome-wide association mapping. *New Phytol.* **213**, 1346–1362. <https://doi.org/10.1111/nph.14220> (2017).
34. Kafoutchoni, K. M., Agoyi, E. E., Agbahoungba, S., Assogbadjo, A. E. & Agbangla, C. Genetic diversity and population structure in a regional collection of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal & Baudet]. *Genet. Resour. Crop. Evol.* <https://doi.org/10.1007/s10722-021-01187-4> (2021).
35. Brown, J. L. & Carnaval, A. C. A tale of two niches: Methods, concepts, and evolution. *Front. Biogeogr.* <https://doi.org/10.21425/f5fbg44158> (2019).
36. Marcer, A., Mendez-Vigo, B., Alonso-Blanco, C. & Pico, F. X. Tackling intraspecific genetic structure in distribution models better reflects species geographical range. *Ecol. Evol.* **6**, 2084–2097. <https://doi.org/10.1002/ece3.2010> (2016).
37. Oney, B., Reineking, B., O'Neill, G. & Kreyling, J. Intraspecific variation buffers projected climate change impacts on *Pinus contorta*. *Ecol. Evol.* **3**, 437–449. <https://doi.org/10.1002/ece3.426> (2013).
38. Tamini, Z. *Etude ethnobotanique et analyses morphophysiologicals du développement de la lentille de terre* [*Macrotyloma geocarpum* (harms) Maréchal et Baudet] (Université de Ouagadougou, 1997).
39. Yohannes, H. A review on relationship between climate change and agriculture. *J. Earth Sci. Clim. Change* **7**, 1–8. <https://doi.org/10.4172/2157-7617.1000335> (2015).
40. Sileshi, G. *et al.* Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. *Field Crops Res.* **116**, 1–13. <https://doi.org/10.1016/j.fcr.2009.11.014> (2010).
41. Padi, F. K. & Ehlers, J. D. Effectiveness of early generation selection in cowpea for grain yield and agronomic characteristics in Semiarid West Africa. *Crop Sci.* **48**, 533–540. <https://doi.org/10.2135/cropsci2007.05.0265> (2008).
42. Kouelo, K. A. F. *et al.* Impact du travail du sol et de la fertilisation minérale sur la productivité de [*Macrotyloma geocarpum* (Harms) Maréchal et Baudet] au centre du Bénin. *J. Appl. Biosci.* **51**, 3625–3632 (2012).
43. Wellenreuther, M., Larson, K. W. & Svensson, E. I. Climatic niche divergence or conservatism? Environmental niches and range limits in ecologically similar damselfly. *Ecology* **93**, 1353–1366. <https://doi.org/10.1890/11-1181.1> (2012).
44. Akohoue, F., Achigan-Dako, E. G., Coulibaly, M. & Sibiyi, J. Correlations, path coefficient analysis and phenotypic diversity of a West African germplasm of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal & Baudet]. *Genet. Resour. Crop Evol.* **66**, 1825–1842. <https://doi.org/10.1007/s10722-019-00839-w> (2019).
45. Assogba, P. *et al.* Indigenous knowledge and agro-morphological evaluation of the minor crop Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal et Baudet] cultivars of Benin. *Genet. Resour. Crop Evol.* **63**, 513–529. <https://doi.org/10.1007/s10722-015-0268-9> (2015).

46. Adu-Gyamfi, R., Dzomeku, I. K. & Lardi, J. Evaluation of growth and yield potential of genotypes of Kersting's groundnut (*Macrotyloma geocarpum* Harms) in Northern Ghana. *Int. Res. J. Agric. Sci. Soil Sci.* **2**, 509–515 (2012).
47. Burke, M. B., Lobell, D. B. & Guarino, L. Shifts in African crop climates by 2050, and the implications for crop improvement and genetic resources conservation. *Glob. Environ. Change* **19**, 317–325. <https://doi.org/10.1016/j.gloenvcha.2009.04.003> (2009).
48. Ramirez-Cabral, N. Y. Z., Kumar, L. & Shabani, F. Global alterations in areas of suitability for maize production from climate change and using a mechanistic species distribution model (CLIMEX). *Sci. Rep.* **7**, 1–13. <https://doi.org/10.1038/s41598-017-05804-0> (2017).
49. Hancock, A. M. *et al.* Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* **334**, 83–86. <https://doi.org/10.1126/science.1209244> (2011).
50. Bellon, M. R. & van Etten, J. Climate change and on-farm conservation of crop landraces in centres of diversity. In *Plant Genetic Resources and Climate Change* Vol. 30 (eds Jackson, M. *et al.*) (CAB International, 2014).
51. Lane, A. & Jarvis, A. Changes in climate will modify the geography of crop suitability: Agricultural biodiversity can help with adaptation. *ISI J* **4**, 12 (2007).
52. Vigouroux, Y., Barnaud, A., Scarcelli, N. & Thuillet, A. C. Biodiversity, evolution and adaptation of cultivated crops. *C.R. Biol.* **334**, 450–457. <https://doi.org/10.1016/j.crvi.2011.03.003> (2011).
53. Coulibaly, M., Agossou, C. O. A., Akohoué, F., Sawadogo, M. & Achigan-Dako, E. G. Farmers' preferences for genetic resources of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal and Baudet] in the production systems of Burkina Faso and Ghana. *Agronomy* **10**, 371. <https://doi.org/10.3390/agronomy10030371> (2020).
54. Sohn, N., Fernandez, M. H., Papes, M. & Anciães, M. Ecological Niche modeling in practice: Flagship species and regional conservation planning. *Oecol. Aust.* **17**, 429–440. <https://doi.org/10.4257/oeco.2013.1703.11> (2013).
55. Amujoyegbe, B., Obisesan, I., Ajayi, A. & Aderanti, F. Disappearance of Kersting's groundnut [*Macrotyloma geocarpum* (harms) Maréchal et Baudet] in South-Western Nigeria: An indicator of genetic erosion. *Plant Gen Res News* **152**, 45–50 (2007).
56. Banta, J. A. *et al.* Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in *Arabidopsis thaliana*. *Ecol. Lett.* **15**, 769–777. <https://doi.org/10.1111/j.1461-0248.2012.01796.x> (2012).
57. Kumar, J., Choudhary, A. K., Gupta, D. S. & Kumar, S. Towards exploitation of adaptive traits for climate-resilient smart pulses. *Int. J. Mol. Sci.* **20**, 1–30. <https://doi.org/10.3390/ijms20122971> (2019).
58. Bohra, A., Mir, R. R., Jha, R., Maurya, A. K. & Varshney, R. K. Advances in genomics and molecular breeding for legume improvement. In *Advancement in Crop Improvement Techniques* (eds Bohra, A. *et al.*) 129–139 (Elsevier Inc, 2020).
59. Gobu, R. *et al.* Accelerated crop breeding towards development of climate resilient varieties. In *Climate Change and Indian Agriculture: Challenges and Adaptation Strategies* (eds Srinivasarao, C. *et al.*) 49–69 (ICAR-National Academy of Agricultural Research Management, 2020).
60. Aliyu, S., Massawe, F. & Mayes, S. Genetic diversity and population structure of Bambara groundnut [*Vigna subterranea* (L.) Verdc.]: Synopsis of the past two decades of analysis and implications for crop improvement programmes. *Genet. Resour. Crop. Evol.* **63**, 925–943. <https://doi.org/10.1007/s10722-016-0406-z> (2016).
61. Al-Khayri, J. M., Jain, S. M., Johnson, D. V. Springer Nature Switzerland AG. Switzerland. <https://doi.org/10.1007/978-3-030-23400-3> (2019).
62. Kilian, A. *et al.* Diversity arrays technology: a generic genome profiling technology on open platforms. *Methods Mol. Biol.* **888**, 67–89. [https://doi.org/10.1007/978-1-61779-870-2\\_5](https://doi.org/10.1007/978-1-61779-870-2_5) (2012).
63. Illumina, I. HiSeq<sup>®</sup>2500 Sequencing System: Unsurpassed power and efficiency for production scale sequencing. System Specification Sheet: Sequencing, 1–4. [https://www.illumina.com/documents/products/datasheets/datasheet\\_hiseq2500.pdf](https://www.illumina.com/documents/products/datasheets/datasheet_hiseq2500.pdf) (2015).
64. Buckler, E. *et al.* User Manual for TASSEL Trait Analysis by association, Evolution and Linkage Version 5.0. The Buckler Lab at Cornell University. 1–70. <https://www.maizogenetics.net/tassel> (2014).
65. Pritchard, J. K., Wen, X., Falush, D. Department of Human Genetics, University of Chicago. (2010).
66. Francis, R. M. pophelper: An R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* **17**, 27–32. <https://doi.org/10.1111/1755-0998.12509> (2016).
67. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software structure: A simulation study. *Mol. Ecol.* **14**, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> (2005).
68. Jombart, T. & Ahmed, I. adegenet version 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btr521> (2011).
69. Boria, R. A., Olson, L. E., Goodman, S. M. & Anderson, R. P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol. Model.* **275**, 73–77. <https://doi.org/10.1016/j.ecolmodel.2013.12.012> (2014).
70. Peterson, A. T. *et al.* *NicheBook* (Princeton University Press, 2011).
71. Kass, J. M., Pinilla-Buitrago, G. E., Vilela, B., Aiello-Lammens, M. E., Muscarella, R., Merow, C., Anderson, R. P., Wallace: A Modular Platform for Reproducible Modeling of Species Niches and Distributions. R package version 1.0.6.3. (2020).
72. Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. & Anderson, R. P. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545 (2015).
73. Platts, P. J., Omeny, P. A. & Marchant, R. AFRICLIM: High-resolution climate projections for ecological applications in Africa. *Afr. J. Ecol.* **53**, 103–108 (2014).
74. Meinshausen, M. *et al.* The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Clim. Change* **109**, 213–241. <https://doi.org/10.1007/s10584-011-0156-z> (2011).
75. IPCC, I. P. o. C. Cambridge University Press. <https://doi.org/10.1017/CBO9781107415324> (2013).
76. Ramirez-Villegas, J. & Jarvis, A. Downscaling global circulation model outputs: The delta method decision and policy analysis, Working Paper No. 1. *Policy Anal. Manag.* **1**, 1–18 (2010).
77. Hengl, T. *et al.* Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve current predictions. *PLoS ONE* **10**, 1–26. <https://doi.org/10.1371/journal.pone.0125814> (2015).
78. Osorio-Olvera, L. *et al.* ntbox: An R package with graphical user interface for modeling and evaluating multidimensional ecological niches. *Ecol. Evol.* **11**, 1199–1206. <https://doi.org/10.1111/2041-210X.13452> (2020).
79. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum Entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026> (2005).
80. Elith, J. *et al.* A statistical explanation of MaxEnt for ecologists. *Diversity Distrib.* **17**, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x> (2010).
81. Richards, C. L., Carstens, B. C. & Lacey Knowles, L. Distribution modelling and statistical phylogeography: An integrative framework for generating and testing alternative biogeographical hypotheses. *J. Biogeogr.* **34**, 1833–1845. <https://doi.org/10.1111/j.1365-2699.2007.01814.x> (2007).
82. Phillips, S. J. & Dudík, M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* **31**, 161–175. <https://doi.org/10.1111/j.2007.0906-7590.05203.x> (2008).
83. Merow, C., Smith, M. J. & Silander, J. A. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **36**, 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x> (2013).
84. Phillips, S. J. *et al.* Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecol. Appl.* **19**, 181–197 (2009).
85. Barve, N., Barve, V. ENMGadgets: Pre and post processing in ENM Workflow. R package version 0.1.0.1. (2019).

86. Kass, J. M. *et al.* ENMeval 20: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Ecol. Evolut.* <https://doi.org/10.1111/2041-210X.13628> (2021).
87. Warren, D. L. & Seifert, S. N. Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* **21**, 335–342. <https://doi.org/10.1890/10-1171.1> (2011).
88. Warren, D. L., Glor, R. E. & Turelli, M. Environmental Niche equivalency versus conservatism: Quantitative approaches to Niche evolution. *Evolution* **62**, 2868–2883. <https://doi.org/10.1111/j.1558-5646.2008.00482.x> (2008).
89. Broennimann, O. *et al.* Measuring ecological niche overlap from occurrence and spatial environmental data. *Glob. Ecol. Biogeogr.* **21**, 481–497. <https://doi.org/10.1111/j.1466-8238.2011.00698.x> (2012).
90. Benhamou, S. & Cornélis, D. Incorporating movement behavior and barriers to improve kernel home range space use estimates. *J. Wildl. Manag.* **74**, 1353–1360. <https://doi.org/10.2193/2009-441> (2010).

## Acknowledgements

This work was supported by the “Intra-Africa Academic Mobility Scheme” under the project Grant number 2016-2988 on “Enhancing training and research mobility for novel crops breeding in Africa (MoBreed)”, funded by the Education, Audiovisual and Culture Executive Agency (EACEA) of the European Commission. The authors gratefully acknowledge Biosciences eastern and central Africa—International Livestock Research Institute (BeCA-ILRI Hub), Nairobi, Kenya, for the crop genotyping and assistance for SNPs data treatment. We also highly appreciate the assistance of Mr Gafarou Agounde (from LABEF) and Mr Alban Zossou in the data preparation and data analysis.

## Author contributions

M.C.: conceptualization, investigation, data curation, methodology, molecular data analysis, writing—review and editing; E.G.A.D.: conceptualization, writing—review and editing, visualization, supervision, funding management, project administration; R.I.: software, data analysis, review and editing; F.A.: additional molecular data resources, review and editing; A.T.P., M.S.: review and editing, supervision.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-09153-5>.

**Correspondence** and requests for materials should be addressed to E.G.A.-D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022