communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-08227-0

Cowpea (*Vigna unguiculata* L. Walp.) landraces in Mozambique and neighbouring Southern African countries harbour genetic loci with potential for climate adaptation

Check for updates

Mercy Wairimu Macharia[1], **Leonardo Caproni**[1], **Marta Solemanegy**[1,2], **Robel Takele**[1],
Rogerio Marcos Chiulele[3], **Paulino Munisse**[2], **Manuel Amane**[2], **Mario Enrico Pè**[1], **Roberto Buizza**[1] &
Matteo Dell'Acqua[1] ✉

Large untapped potential for local climate adaptation is present in plant genetic resources maintained by smallholder farmers in Southern Africa. Cowpea (*Vigna unguiculata* L.Walp.), a staple crop for local farmers in the arid and semiarid regions of Southern Africa, is mostly unexplored by modern breeding approaches. In this paper, we assemble and characterize a collection of 389 cowpea accessions from Southern Africa, mainly landraces, including the entire ex-situ collection from the national Mozambique genebank. We use a genotyping-by-sequencing approach to describe the diversity in the collection and we concurrently characterize historical and projected climate at sampling points of landraces in search of genomic signatures of local adaptation. Our results show unique cowpea diversity in Mozambique, which can be partially put in relation with bioclimatic variation. The genotype-environment association approaches, Latent Factor Mixed Models (LFMM) and partial Redundancy Analysis (pRDA) allowed us to identify 36 genomic loci potentially involved in local climate adaptation. This included a region on chromosome 7 tagging four candidate genes linked to flowering and including a homolog of *GIGANTEA*, a gene regulating flowering time in response to day length and temperature in Arabidopsis. Finally, we estimated cowpea landraces adaptation to projected climate in the region, highlighting regions of maladaptation in southern Tanzania and Zimbabwe. Our results show that genetic resources maintained by farmers in Southern Africa bear traits for local climate adaptation and may contribute to enhancing the adaptability of cowpea to a shifting climate.

Smallholder farming systems of sub-Saharan Africa (SSA) contribute to up to eighty percent of the agricultural production on the continent[1]. They are characterized by low input agriculture as a mean of sustenance of local livelihoods, often in off-grid communities, and have a crucial role in the resilience of small-scale and large-scale socioeconomic systems in the whole African continent[2]. These farming systems are markedly exposed to the impacts of the climate crisis, which results in lower and more unsteady harvests[3,4]. The 6th Assessment Report of Working Group II of the Intergovernmental Panel on Climate Change (IPCC) published in 2022, reported Africa as a hotspot of anthropogenic climate change, with areas in which trends in mean maximum and minimum temperatures range between +2 °C and +3 °C per century[5]. As climate change accelerates, this region is likely to experience increasingly extreme levels of warming. This, combined with the rapidly growing population in Sub-Saharan Africa (SSA), continues to widen the gap between food

---

[1]Institute of Plant Sciences, Scuola Superiore Sant'Anna, Pisa, Italy. [2]Instituto de Investigação Agrária de Moçambique, Maputo, Mozambique. [3]Centre of Excellence in Agri-Food Systems and Nutrition, UEM, Maputo, Mozambique. ✉e-mail: matteo.dellacqua@santannapisa.it

production and demand. Therefore, it is imperative to enhance agricultural resilience across the continent.

Traditional SSA crops such as cowpea (*Vigna unguiculata* (L.) Walp.) hold significant promise for addressing climate adaptation challenges in smallholder farming systems[6]. Cowpea is valued for its drought tolerance, early maturity, and nitrogen-fixing ability, making it a valuable resource for food security and sustainable farming in limiting conditions. Due to these characteristics, it is often favored by SSA farmers over more productive crops[7]. Cowpea cultivation in Africa dates as far back as 1500 BCE[8] and has been since maintained and selected by local farmers in thousands of landraces distributed across the continent and adapted to different growing conditions and farmer cultural preferences[9]. Today, more than 90% of cowpea global production is concentrated in smallholder systems of SSA, particularly in West Africa, and is characterized by a productivity of only around 0.6 tonnes per hectare (t/ha)[10,11]. Increased incidence of pests and diseases[12,13], declining soil fertility, and use of low-yielding seed varieties challenge cowpea productivity in the continent. The climate crisis is expected to have a further adverse effect on cowpea productivity exacerbated by irregular and erratic rainfall patterns—often linked to drought and aridity[14]. There is the need to tap into the undisclosed potential of cowpea genetic resources to boost its capacity for adaptation and productivity in the continent[6].

Traditional agrobiodiversity maintained by local farmers is a trove of allelic variation contributing to adaptation, a promising genetic pool for selecting and developing new varieties[15]. Accessing this diversity is facilitated through ex-situ collections which maintain landraces along with their passport data detailing their geographic origin, environment, and farming practices. When combined with genomic tools, these collections can be effectively explored to uncover their adaptive potential and accelerate the integration of their genetic diversity into breeding programs[16]. Concurrently, climate data can be interrogated pivoting sampling information of the ex-situ accessions, opening the possibility to identify genetic factors that can explain climate adaptation of landraces[17]. Combined biogenetic and climate data analysis can be then exploited to identify the best adaptation strategies for projected climate scenarios[18,19], shedding light into to future trends and vulnerabilities[20] and giving the opportunity to effectively screen plant genetic resources collections for alleles with potential contribution in enhancing climate change resilience.

In this study, we aimed at tapping into the untapped gene pool of Southern Africa cowpea to look for genomic potential for climate adaptation. The Mozambican allele pool is particularly interesting due to its high genetic diversity and its distinction from germplasm sourced from other geographical regions around the world[15,21]. We assembled a collection of 389 cowpea landrace accessions, representing the whole diversity of cowpea in Mozambique and a sample of diversity of neighbouring countries and 28 improved varieties for comparison, and applied genomics to derive molecular markers. We leveraged landrace sampling information to identify genetic regions potentially associated with temperature and precipitation factors, providing insights into the molecular basis of cowpea adaptation to current and future climate conditions.

Results

Genetic diversity and population structure of cowpea genetic resources

We assembled a cowpea collection of 389 genotypes, mostly landraces and 28 improved varieties, representative of Southern Africa countries and maintained ex-situ in African genebanks (Supplementary Data 1; Fig. S1). Samples were genotyped using a double-digestion Restriction site Associated DNA markers (ddRAD) approach producing more than three billion paired end reads (Supplementary Data 2). Reads were aligned to the cowpea reference genome ASM411807v2[22], yielding 202,946 genome-wide Single Nucleotide Polymorphisms (SNPs) that were further reduced to an operational, high-quality set of 18,969 SNPs with minor allele frequency (MAF) above 5%. To conduct diversity and genetic structure analyses, the SNP set was further reduced by dropping markers in Linkage Disequilibrium (LD), resulting in a set of 2242 LD-pruned SNPs. The LD-pruned SNPs were used to compute a Neighbour-Joining (NJ) tree of the accessions included in the study (Fig. 1a). The NJ tree showed a distinct group of cowpea accessions for the most part consisting of landraces from Mozambique, while the rest of the collection did not show obvious geographical grouping. A Discriminant Analysis of Principal Components (DAPC) suggested the existence of three genetic clusters best describing the distribution of allelic frequencies within the collection (Fig. 1b). These clusters can be best depicted in a Principal Component Analysis (PCA) computed on SNP data and explaining ~36% of the genotypic variance in the dataset in its first two dimensions (Fig. 1b). Cluster 3 consisted majorly of Mozambique landraces while clusters 1 and 2 consisted of a mixture of accessions from South Africa, Malawi, Tanzania, Zimbabwe, Mozambique and improved varieties. Observed heterozygosity was generally low with values ranging from 0.02 to 0.05 (Supplementary Data 3), and accordingly, F_{IS} values were generally high, ranging from 0.61 in cluster 3 to 0.91 in cluster 1 (Supplementary Data 3). The overall nucleotide diversity (π) for the entire dataset was 0.23. Clusters 1 and 2 exhibited higher π, 0.28 and 0.24, respectively, indicating substantial genetic diversity within each group. Cluster 3, predominantly composed of Mozambique samples, had the lowest nucleotide diversity ($\pi = 0.12$), suggesting lower within-group genetic variation compared to Clusters 1 and 2. Nonetheless, we identified rare alleles (with a frequency below 5% in each cluster group) at rates of 2.3% in Cluster 1, 6.5% in Cluster 2, and 14.6% in Cluster 3. Bayesian structure analysis revealed that the most probable number of genetic clusters in the cowpea core collection was two ($K = 2$) (Fig. S2). One cluster predominantly consisted of Mozambique landraces, while the second cluster included mostly improved varieties and accessions from other Southern African countries (Fig. 1c).

Cowpea diversity on the landscape is partially explained by climate

Cowpea landraces included in this study originated from several Agroecological Zones (AEZs), the majority in low land semiarid, and low land subhumid zones (Figure S3: Supplementary Data 4). We explored climate diversity of sampling points within these AEZ's using historical bioclimatic variables corresponding to accessions' sampling points. For this purpose, we selected noncollinear bioclimatic variables using a Variance Inflation Factor ($VIF < 5$) (Supplementary Data 5) to ensure the absence of multicollinearity. Pairwise correlation analysis confirmed that the selected variables were non-redundant ($r < 0.6$) (Figure S4). They includes bio2 (mean diurnal range), bio3 (isothermality), bio5 (maximum temperature of warmest month), bio12 (annual precipitation), bio14 (precipitation of driest month) and bio18 (precipitation of warmest quarter). Overall, we found large climatic diversity at the cowpea points of origin (Fig. 2), maximum temperature of warmest month (bio5) ranged more than ten degrees, from 24.8 °C to 36.3 °C (Supplementary Data 6). We observed annual precipitation (bio12) values between 381 to 1553 mm, typical of semi-arid and subhumid AEZs, respectively (Supplementary Data 6). In a multivariate analysis of the bioclimatic variables, the first principal component (Dim 1) accounted for 36% of the total variance while the second principal component (Dim 2) contributed an additional 22.6% of variance, capturing most of the remaining variability (Fig. S5). The bioclimatic variables most associated with Dim 1 were bio12 (Annual precipitation) and bio2 (Mean diurnal range of temperature), with loading values of 0.56 and 0.49, respectively (Supplementary Data 7). For Dim 2, the primary contributors were bio18 (Precipitation of the coldest quarter) and bio3 (Isothermality), with loadings of 0.57 and 0.47, respectively. The variance partitioning conducted using partial Redundancy Analysis (pRDA) indicated that in total, climate, geography, and genetic structure collectively accounted for 15% of the variance in the collection (Table 1). Notably, climate alone contributed to 5% of the explained variation. The remaining variation could not be effectively explained by any covariate in the model, in accordance with a fairly admixed diversity in the landraces collection.

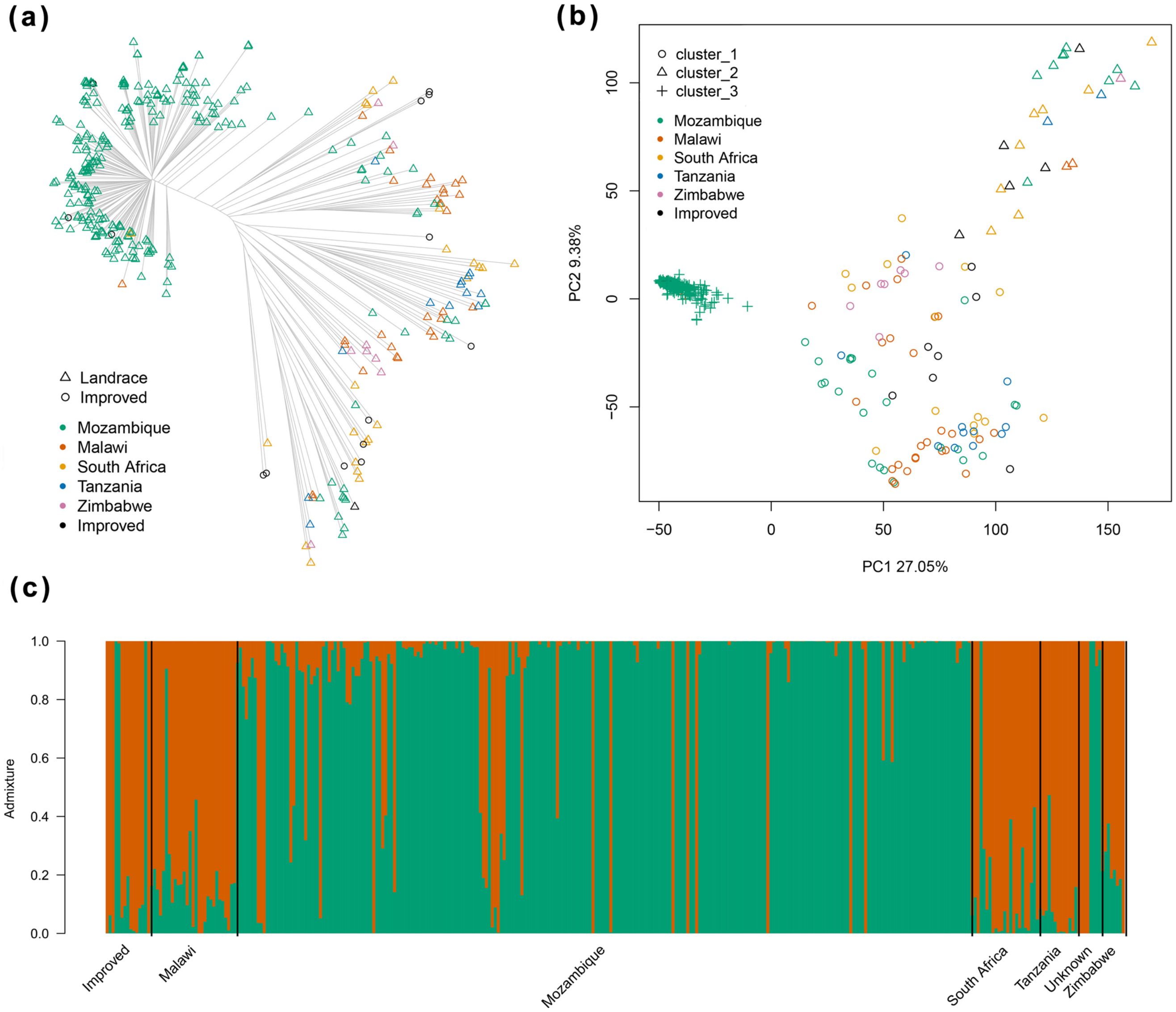

Fig. 1 | Genetic diversity and population structure of 389 cowpea landrace accessions and 28 improved varieties. a Neighbour-joining tree derived from Single Nucleotide Polymorphism (SNP) data. **b** Principal Component Analysis (PCA) derived from SNP data. Different colours represent the regions of origin, as shown in the legend (**a**, **b**), with symbols indicating type of material in (**a**) and Discriminant Analysis of Principal Components (DAPC) clusters in (**b**). **c** Structure analysis of the core collection of cowpea ($K = 2$). The bar plot shows the genetic ancestry of each accession, with the two colors representing two distinct genetic clusters. The accessions are ordered by country of origin, with black lines demarcating groups by origin. Country names are displayed on the x-axis.

In the pRDA, the significance of RDA1 to RDA4 was confirmed through ANOVA ($p < 0.05$), with RDA1 and RDA2 being highly significant ($p < 0.001$), RDA3 significant ($p < 0.01$) and RDA4 marginally significant ($p = 0.016$) (Supplementary Data 8).

The potential impact of climate change on the cultivation of cowpea in Southern Africa

Next, we evaluated the potential impact of the local climate on the extant genetic makeup of cowpea accessions, giving us indications of how the different cowpea varieties could adapt to the projected climate scenarios in Southern Africa. A Gradient Forest (GF) model was created utilizing the set of non-collinear historical bioclimatic variables and MEMs (Moran's Eigenvector Maps) as predictors, with 2242 LD-pruned SNPs tested as response variables. Bioclimatic variables represent climatic differences, MEM represent spatial structure. The allelic turnover in the cowpea collection was best predicted by MEM5 (Fig. 3a); among the bioclimatic variables, bio3 (isothermality) had the highest importance in terms of contribution to prediction accuracy, followed by precipitation of warmest quarter (bio18), and maximum temperature of warmest month (bio5) (Fig. 3a). We utilized these bioclimatic predictors to estimate the genomic variation influenced by climate across the landscape (Fig. 3b, c). Out of all the SNPs tested as response variables, approximately 35% ($n = 787$) were successfully predicted by the GF model ($R^2 > 0$) (Supplementary Data 9). We then modelled the expected genomic composition under the SSP 585 (2041riance indicates the overall c2060) climatic scenario, representing a significant increase in global temperatures, based on landscape-genome relations modelled by the GF. We used it to derive cowpea genomic offset in future scenarios as the discrepancy between the current genetic variation and the genetic variation that would be needed under a changed climate for the tested scenario. Overall, we observed moderate offset, with some areas of predicted higher vulnerability in the southeastern and northeastern parts of Zimbabwe, as well as in the southern region of Tanzania (Fig. 3d). The lowland sub-humid coastal regions of Mozambique showed little to no offset of cowpea diversity.

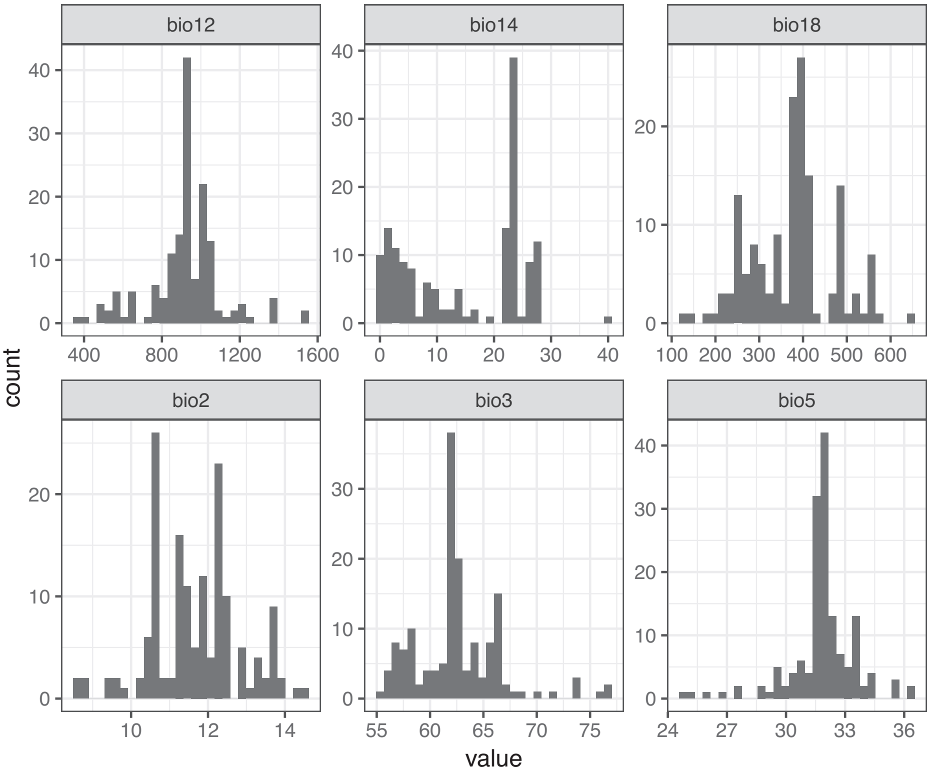

Fig. 2 | Distribution of bioclimatic variables at the original sampling locations of the accessions. Histograms showing frequency of selected bioclimatic variables including bio2 (mean diurnal range, °C), bio3 (isothermality, %), bio5 (maximum temperature of the warmest month, °C), bio12 (annual precipitation, mm), bio14 (precipitation of the driest month, mm), and bio18 (precipitation of the warmest quarter, mm).

Table 1 | A partial redundancy analysis (pRDA) showing the impact of climate, geography, and genetic composition on the detected genetic diversity of cowpea landraces

Partial RDA model	Inertia	R^2	adj R^2	*p*(>F)	Proportion of explainable variance	Proportion of Total Variance
Full model: G ~ clim. + geog. + struct.	137.2	0.15	0.08	0.001 ***	1.00	0.15
Pure climate: G ~ clim. | (geog. + struct)	49.6	0.05	0.02	0.001 ***	0.36	0.05
Pure structure: G ~ struct. | (clim. + geog.)	17.9	0.02	0.00	0.001 ***	0.13	0.02
Pure geography: G ~ geog. | (clim. + struct.)	24.1	0.03	0.02	0.001 ***	0.18	0.03
Confounded climate/structure/geography	45.6					0.05
Total unexplained	771.9					0.85
Total inertia	909.1					1.00
***$P \leq 0.001$						

The proportion of explainable variance indicates the overall constrained variation explained by the complete model, while "inertia" denotes variance.

Genomic loci with adaptive potential are available in Southern African cowpea landraces

We first searched for adaptive loci using a Latent Factor Mixed Model (LFMM) approach (Supplementary Data 10), testing associations between the full set of markers ($n = 18{,}969$) and the six non-collinear bioclimatic variables. With this analysis we identified 11 unique Marker-environment associations (MEAs), in bio2, bio3, bio12, and bio18 (Figs. 4, S6; Supplementary Data 11).

Two common MEAs were identified within a shared LD block on chromosome 4 at 41.3 Mb, associated with bio12 and bio18. There were two MEAs located within an LD block on chromosome 11, spanning positions 1.73 Mb and 1.75 Mb associated with bio2, bio12, and bio18. On chromosome 7, a MEA at 36.7 Mb was common for bio3 and bio18, which ranked as the most important climatic predictors towards the GF model (Fig. 4).

Additionally, we searched for putative adaptive loci using an approach based on pRDA, specifically implemented with the function in R rdadapt()[23]. This was done by employing a model that controls for genetic structure, with the LD-pruned set of markers as response variables (Figs. 5a, S5). This analysis revealed 21 putative adaptive loci distributed across nine chromosomes (Fig. 5b; Supplementary Data 12).

We then compared the SNPs predicted by the GF model with the LFMM MEAs and with the pRDA SNPs to prioritize loci relevant for different dimensions of environmental adaptation. We found 16 loci overlapping between the pRDA and GF datasets (Fig. 5c; Supplementary Data 13). One MEA locus overlapped with a locus reported by the GF dataset but not by RDA and was located on chromosome 4 (Supplementary Data 11). Only one LD block harboring two loci was common across all three datasets, located on chromosome 7 at 36.7 Mb and 37.3 Mb (Supplementary Data 11; Supplementary Data 13). Subsequently, we scanned the genes within the overlapping loci, harbored by LD blocks adjusted both above and below the LD decay distance and identified the annotated protein-coding gene transcripts (Supplementary Data 14; Supplementary Data 15). We identified two LD blocks: one on chromosome 3 harboring the locus at 0.62 Mb, and another on chromosome 7, harboring loci 36.7 Mb and 37.3 Mb which contain five genes with suggestive roles in environmental regulation that may confer climate adaptive potential (Supplementary Data 16). These genes include *GIGANTEA* (Vigun07g213100.1) which was identified within this LD block harbouring two outlier SNP loci on chromosome 7: Vu07_33160392 and Vu07_33797176 (Supplementary Data 16). Three more genes—*GRAS* (Vigun07g206200.1), *No apical*

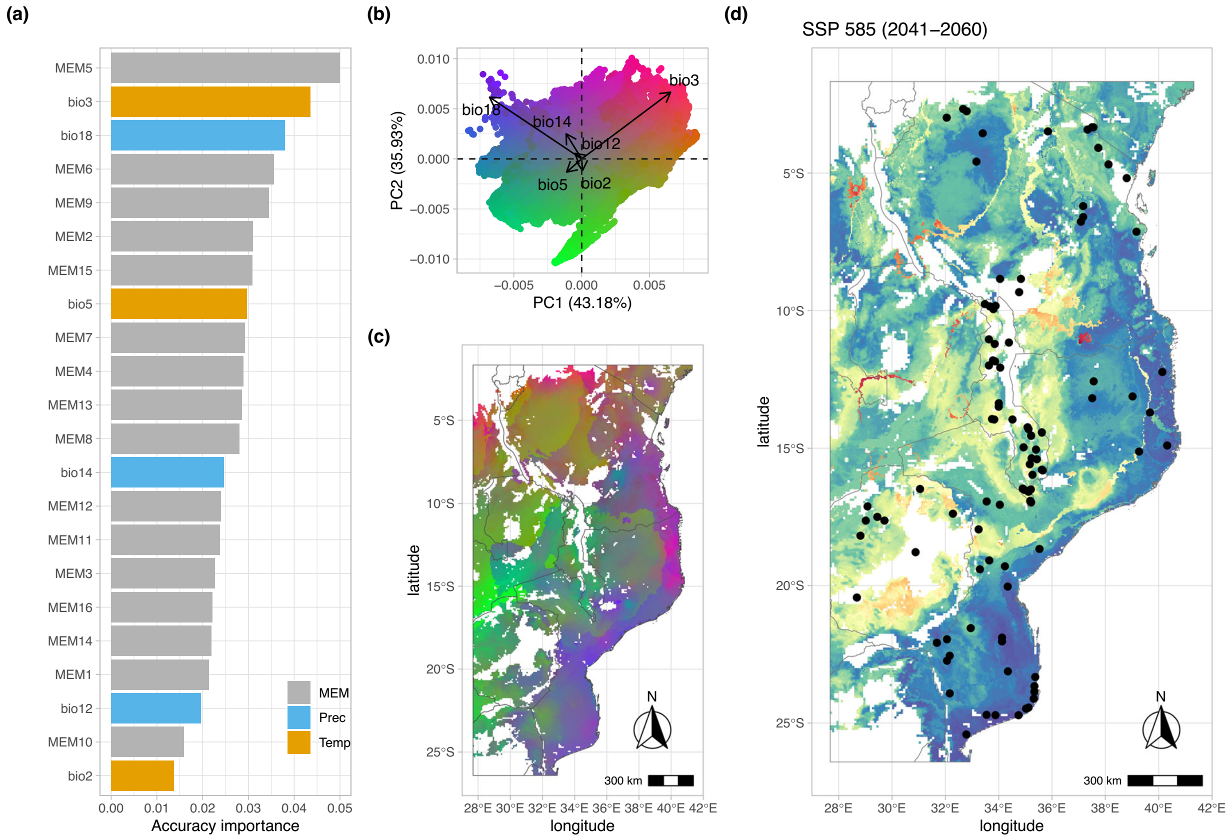

Fig. 3 | The relationship between bioclimatic and spatial diversity and genomic variation of cowpea landraces in Southern Africa. a The accuracy importance of bioclimatic and spatial variables is ranked based on the Gradient Forest (GF) model. **b** A biplot of the biological space displays Principal Components of the transformed bioclimatic predictors using an RGB colour palette relative to its first three dimensions. **c** GF-transformed bioclimatic variables across landscape (Southern Africa). Colours correspond to those on bioclimatic-genetic space in (**b**). **d** Genomic offset based on the Shared Socioeconomic Pathway (SSP) 585 (2041–2060) climate scenario. The colour scale illustrates the extent of the discrepancy between present and anticipated shifts in allele frequencies driven by climate change; the black dots represent cowpea sampling sites of origin of georeferenced cowpea accessions.

meristem (*NAM*) (Vigun07g207300.1), and *somatic embryogenesis receptor kinase 1* (*SERK1*) (Vigun07g209900.1)—were identified within 378 kb, 291 kb, and 8 kb, respectively, in the vicinity of the outlier SNP Vu07_33160392 on the same LD block. On chromosome 3, the gene *Dehydrin* (Vigun03g008500.1) was found within the LD block harbouring outlier SNP at locus Vu03_628006.

Discussion

Climate change challenges plant breeding to rapidly identify solutions to enhance crop adaptation to rapid fluctuations in cultivation conditions and resource availability. Genetic traits in landraces, selected and preserved across diverse agroecological zones, offer allelic variations that can be harnessed to accelerate the development of new, improved varieties with reduced input requirements and greater yield stability under local conditions[24]. Within Southern Africa, including Mozambique, arid and semi-arid areas cover more than one third of the arable land and sustainable utilization of these agroecological zones remains a priority[25]. Cowpea farmers in these areas often prefer landraces to improved varieties and select them based on adaptability to prevailing environmental stresses including drought[26]. Hence the potential value of characterizing and utilizing existing genetic diversity of cowpea landraces to develop resilient crop varieties that can thrive in these challenging environments, ensuring food security and sustainable agricultural practices in the face of climate change.

In this study, we characterized the genetic diversity of Southern African cowpea landraces to gain insights into the variation present in local farmer fields and maintained ex-situ in SSA. The snapshot of current cowpea variation is arguably the result of a combination of evolutionary mechanisms, among which selection exerted by local pedoclimatic conditions and farmers' choice of materials[27,28]. Analysis of population structure identified three primary genetic clusters (Fig. 1b). Despite obtaining accessions from five countries across Southern Africa, only a subset of Mozambican landraces displayed clear and distinct geographical clustering (cluster 3). The observed clustering pattern reflects differences likely resulting from different recent evolutionary dynamics, including adaptation driven by farmer choice and local environment. The distinct grouping of Mozambican landraces (Cluster 3) suggests separation of the allele pool, while the admixture in Clusters 1 and 2, which include accessions from multiple countries and improved varieties, indicates historical gene flow and shared ancestry. Previous work[29] noted that cowpea landraces from Southern Africa exhibited higher variability among subpopulations within the region compared to other parts of the world, possibly due to limited germplasm exchange or a greater diversity of local environmental conditions. Our data support the idea of historical gene flow as evidenced by the lack of more distinct subpopulations and the minimal presence of private alleles, even in the existing ones (Supplementary Data 3). Such finding may be conflated by the overrepresentation of samples from Mozambique in the collection, although an analysis restricted to Mozambican samples suggests that this has a limited effect on the reported genetic structure (Fig. S7).

Cluster 3, characterized by most Mozambique landrace accessions, presented a π of 0.12 (Supplementary Data 3), meaning that, on average, there is 1 different nucleotide for every ~8 nucleotide sites. This low π value suggests that the genetic variation within this cluster is relatively limited, perhaps highlighting the need for a revised collection strategy to capture a more comprehensive representation of the genetic diversity present in

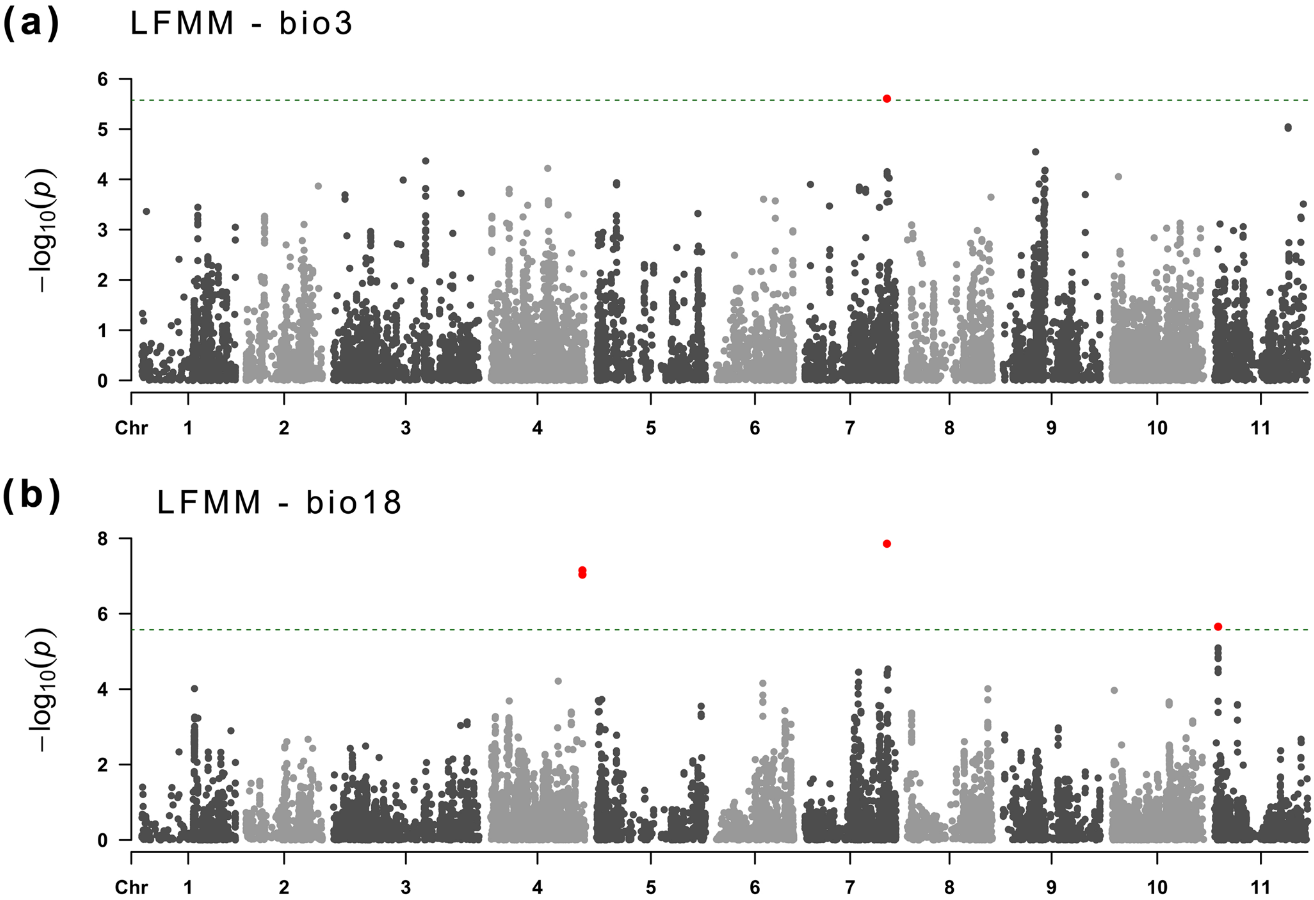

Fig. 4 | Genome-wide Associations with climate variables. Latent Factor Mixed Models (LFMM) Mahattan plots of: **a** isothermality (bio3) and **b** precipitation of the warmest quarter (bio18). In both panels, the x-axis represents the genomic coordinates of SNPs across the chromosomes, while the y-axis shows the $-\log10(p)$ from LFMM analysis, indicating the strength of association between each SNP and the respective bioclimatic variable. The dashed green horizontal lines mark the significance threshold, determined using Bonferroni correction with an alpha level of 0.05. SNPs exceeding this threshold are considered significantly associated with the respective bioclimatic variable and are highlighted in red.

natural populations within Mozambique[15]. On the other hand, Clusters 1 and 2 consisted of a mix of accessions from different countries and include improved varieties. Improved varieties had higher nucleotide diversity ($\pi = 0.33$) compared to landraces ($\pi = 0.20$), likely due to their origin from a broader genetic base (Supplementary Data 1; Supplementary Data 3), combining diverse genetic backgrounds. Nonetheless, landraces exhibited relatively higher observed heterozygosity (Supplementary Data 3), possibly resulting from their admixed origin in farmer fields.

Cowpea landraces were sourced from heterogenous AEZs (Fig. S3) from a large, multi-country area. We performed an admixture analysis to assess the genetic structure of our samples, which suggested no significant genetic differentiation based on the AEZs (Fig. 1c; Fig. S2). The structure algorithm primarily detects the largest change in model likelihood between consecutive K values and may overlook genetic distinctions within groups[30], contributing to the identification of broader divisions in the dataset, such as the $K = 2$ clustering observed here. It's also likely that cowpea accessions largely belong to the same genetic pool because of homogenizing effect of seed exchange amongst farmers. Farmers in Mozambique prioritize drought tolerance as a key trait when selecting cowpea varieties[31], and similar constraints may have contributed to shaping a relatively narrow genetic base across AEZs.

We characterized local climate conditions by considering 30 years of historical climate data, Worldclim Version 2.1[32]. We focused on 6 bioclimatic variables including bio2, bio3, bio5, bio12, bio14 and bio18 (Supplementary Data 5). These selected variables are relevant to cowpea cultivation, as they directly influence critical environmental factors. Temperature variables like bio2, bio3 and bio5 are crucial for cowpea adaptation because they directly influence key physiological processes such as germination, flowering, and seed development[33,34]. Cowpea can tolerate high-temperature environments; however, its productivity can be significantly compromised by heat stress, particularly in relation to the timing and duration of high temperatures[35,36]. In cowpea, night temperatures exceeding 16 °C can lead to a substantial decrease or complete loss of yield[37,38], a phenomenon attributed to the impairment of pollen viability and anther maturation[39]. Drought stress also impacts the growth and development of cowpea with higher vulnerability reported during the flowering and seedling stages[40,41]. The precipitation variables bio12, bio14 and bio18 are critical for cowpea adaptation as they influence not only water availability but also the timing and distribution of precipitation, which directly affect growth, development, and yield potential.

We found significant climate variability in Southern African regions where cowpea is traditionally grown, a prerequisite for identification of potential alleles relevant for local adaptation. We used this diversity to estimate genomic offset for cowpea in future climate scenarios via a GF model (Fig. 3)[42]. We found a moderate offset in most cowpea cultivation areas of Southern Africa. In a few areas where genomic offset was high, there is a likelihood of substantial maladaptation of current cowpea genotypes to future climate. This highlights the importance of supplementing existing diversity through expanding germplasm collection efforts across Southern Africa and/or breeding programs. Adaptive loci may already be found in today's landraces. We thus took a locus-specific approach chasing adaptive loci through pRDA, LFMM, and GF SNPs. Among the adaptive loci

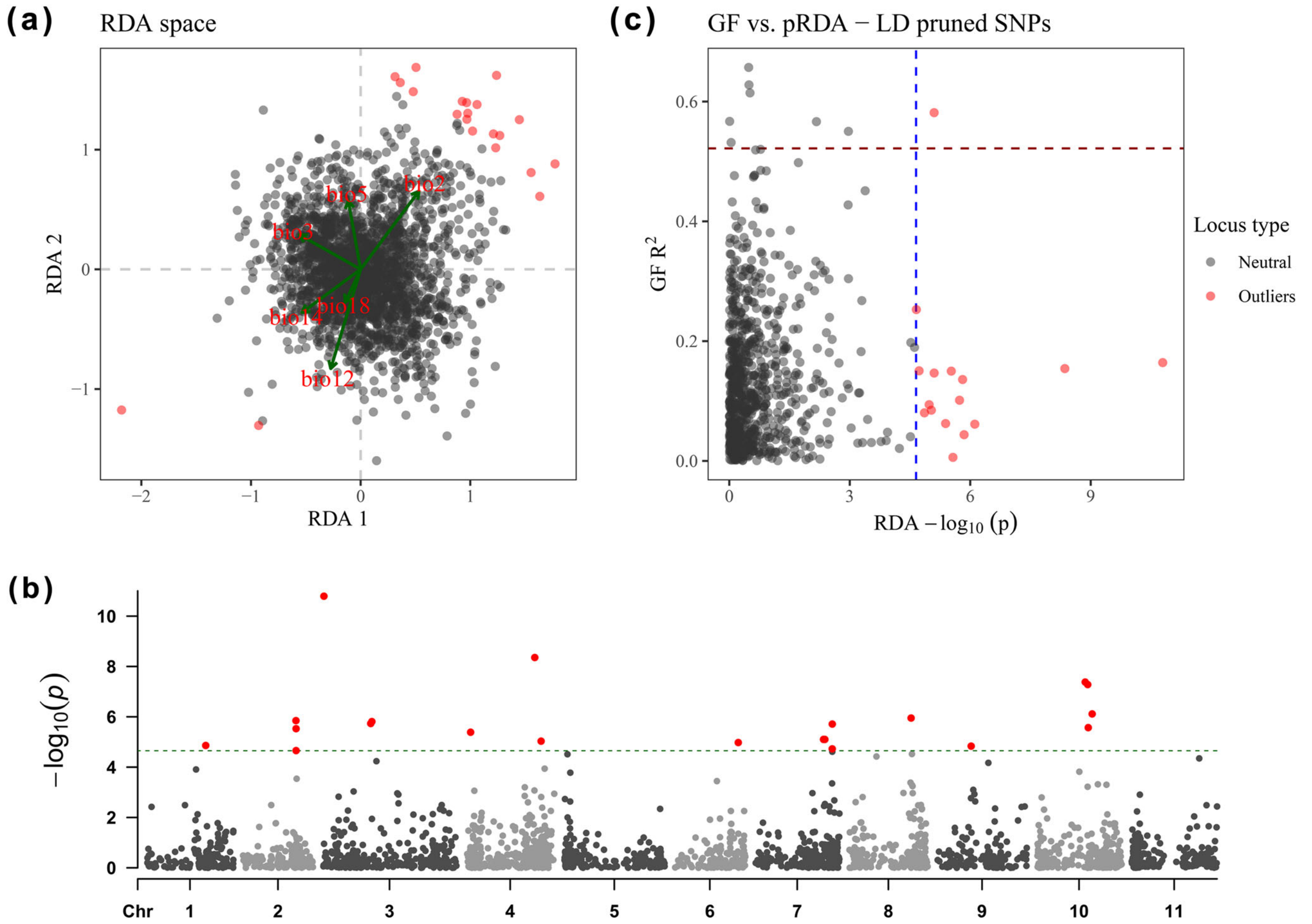

Fig. 5 | Schematic representation of adaptive loci identified through partial Redundancy Analysis (pRDA) and predicted single nucleotide polymorphisms (SNPs) from Gradient Forest (GF) analysis. a Projection of SNPs and selected bioclimatic variables along the first two RDA axes. Outlier SNPs are highlighted in red. **b** Manhattan plot corresponding to pRDA. The plot reports individual SNPs across all chromosomes (x-axis) and -log10(*p*) value of each SNP association (y-axis). The horizontal lines represent a threshold for a nominal *p*-value of 0.05. SNPs above the threshold are highlighted in red. **c** Visualization depicting the overlap between LD-Pruned Single Nucleotide Polymorphisms (SNPs) in the pRDA space (-log10(*p*)) and the R^2 values from the GF analysis. Dashed lines indicate the thresholds in both dimensions. Outlier SNPs from the pRDA are highlighted in red.

identified through pRDA, sixteen overlapped with SNPs highlighted by the GF (Supplementary Data 13). One locus overlapped the GF and LFMM datasets (Supplementary Data 11), and a single LD block was targeted by the three methods. The convergence of different methods on specific loci provides support for signals of adaptation. Still, the overall overlap remains limited, reflecting differences in statistical principles, model assumptions, and sensitivity to genetic–environment associations[43]. pRDA is a constrained ordination method that detects multivariate associations between allele frequencies and environmental variables, maximizing explained genetic variation while explicitly correcting for population structure[44]. LFMM, in contrast, is a univariate mixed model that estimates the relationship between allele frequencies and environmental values one by one[45]. Unlike both these association-based methods, GF is a machine-learning approach that models non-linear genotype-environment relationships without explicitly testing for direct associations but identifies important SNPs based on their contribution to explaining environmental variation and provides a cumulative importance score for each locus[46]. These differences highlight the importance of integrating multiple approaches to capture a more comprehensive set of adaptive loci and reduce potential biases inherent in any single method.

LD decay across the chromosomes was observed to occur within ~391–1746 kilobases, providing a moderate resolution framework for mapping adaptive loci (Supplementary Data 14). In accordance with this, candidate genes were found in association with loci prioritized for local adaptation (Supplementary Data 16). We identified four genes in an LD block on chromosome 7. This block was consistently detected by three statistical methods—pRDA, LFMM, and GF—highlighting the robustness of the findings. These genes are associated with flowering, a biological process strongly associated with local climate adaptation[47,48]. Within this LD block, the ortholog of *GIGANTEA* is suggestive of physiological roles already observed in Arabidopsis and including positive regulation of long-day photoperiodism, flowering time regulation, temperature response and circadian rhythm[49,50]. *GIGANTEA* has also been linked to local adaptation in Arabidopsis[51,52]. In cowpea, *GIGANTEA* has already been identified as a candidate gene for flowering time with a phenotyping-based experiment[53]; it is notable that in our study, this association emerges by connecting genomic data with climate data alone. The other genes identified in this LD block include the *GRAS domain family (GRAS)*, a transcription factor associated with flower development and circadian rhythm regulation in plants[54] and the *No apical meristem (NAM)* transcription activation factor, associated with regulation of flower development and abiotic stress[55,56]. The *somatic embryogenesis receptor kinase 1 (SERK1)* is associated with floral organ abscission, a process tightly controlled by environmental cues[57]. We also identified *Dehydrin gene* on another LD block targeted by multiple methods, a gene associated with response to cold and desiccation in plants[58,59]. The genes highlighted may contribute to the adaptation of cowpea to local environments; we expect that other loci identified in this study, although not associated with clear candidate genes, also contribute significantly to local adaptation, particularly in relation to drought and heat tolerance, which are crucial for maintaining productivity under changing climatic conditions. Further validation of these loci through functional genomics and field trials is necessary to confirm their roles in environmental stress tolerance. Integrating these loci into breeding programs could enhance the resilience of cowpea to environmental stresses. The efficacy of such adaptive loci has been showcased in rice, particularly in conferring traits like cold tolerance[60] and salinity tolerance[61].

Conclusion

The effect of climate change on agriculture does not occur in isolation, and therefore, it's crucial to look at climate adaptation in a multifaceted way. Here, by integrating genomics and climate modelling we have identified genomic loci with potential for local adaptation to environmental conditions. The identification of these loci contributes to understanding the polygenic nature of climate adaptation, highlighting key processes such as flowering, circadian rhythm regulation, and temperature stress response. These processes are influenced by environmental stresses and could facilitate genome-assisted selection for improving cowpea's tolerance to drought, heat, and cold stresses.

Although our study focuses on Southern African region, the genetic mechanisms of adaptation to climate stressors are likely to be consistent across regions with similar environmental conditions. The insights gained from identifying these genomic loci can inform crop improvement strategies even in regions with different environmental challenges, making our findings applicable not only across Africa but also globally in the pursuit of resilient agricultural systems. Most farmers currently have a high incentive to abandon crop landraces for high yielding varieties and thus, landraces may fall out of cultivation. By uncovering the genetic basis of local adaptation in cowpea, we can better prioritize conservation initiatives, complement existing breeding strategies, and promote sustainable management practices in agriculture amidst the challenges posed by climate change.

Methods

Plant material and DNA extraction

A total of 417 accessions, designated as the cowpea collection, were used in this study. The collection consisted of landraces from Malawi (41), Mozambique (283), South Africa (34), Tanzania (19), Zimbabwe (12) and improved varieties (28) (Fig. S1; Supplementary Data 1). The collection was sourced from the International Institute of Tropical Agriculture (IITA) in Nigeria, the national gene bank at the Agricultural Research Institute of Mozambique (IIAM), and the gene bank at the Eduardo Mondlane University in Mozambique. Seed accessions from gene banks are often acquired from heterogeneous sources characterizing farmer fields and associated habitats[62]. Since these materials may not always be genetically pure, nor seed quantities provided by gene banks sufficient to support downstream characterization, we conducted purification and multiplication fields during the growing seasons of 2020/2021 and 2021/2022 at Umbeluzi Research Station (26°03′S, 32°23′E, 12 m a.s.l.) and Nhacoongo Research Station (24°19'49.00" S, 35°12'55.00" E, 68 m a.s.l.), Mozambique. Umbeluzi Research Station is in the Agroecological zone 1 which is a semi-arid to dry climate, with a mean annual precipitation of 679 mm, average temperature ranging from 23 °C to 26 °C during the rainy season. Nhacoongo Station is situated in Agroecological Zone 2 in the southern region of Mozambique. The rainy season at this site typically begins in October and ends in March, with an average annual precipitation of 1000–1200 mm, peaking in February and March. The average annual temperature ranges from 18 °C to 33 °C.

Each cowpea accession was planted in a 5 m x 1 m plot, with 1.6 m spacing between plants and 1 m between rows. Crop management followed the guidelines provided by IITA. Purity was maintained by removing off-type plants during growth, and after harvest, seed sorting was carried out to separate healthy, well-formed seeds from any non-standard or damaged seeds.

DNA extraction was conducted at IIAM molecular laboratory. Briefly, five seeds per accession were germinated in individual pots at the IIAM greenhouse for two weeks. Total genomic DNA was isolated using the GenElute™ Plant Genomic DNA kit following the manufacturer's protocol. Genomic DNA quality was evaluated in 1% agarose gel and quantified using a Nanodrop spectrophotometer. Samples were shipped to IGA Technology Services (Udine, Italy) for library construction and sequencing.

DNA sequencing and genotyping

A custom ddRAD protocol from IGA Technology Services was used for sequencing with slight adjustments based on Peterson's double-digest restriction-site associated DNA preparation[63]. The target DNA was digested using the restriction enzymes *SphI* and *MboI* to achieve an optimal fragment distribution. The quality and quantity of the resulting libraries were assessed using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and a Bioanalyzer DNA assay (Agilent Technologies, Santa Clara, CA). The libraries were then sequenced in paired-end mode with 150 cycles on a NovaSeq 6000 instrument following Illumina's instructions (San Diego, CA). Raw reads were demultiplexed using the process_radtags utility included in Stacks v2.61[64]. Reads were mapped against the *Vigna unguiculata* reference genome[22] ASM411807v2 available at (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_004118075.2/) using BWA-MEM algorithm[65] at default parameters and selection of unique reads with a mapping quality greater than 4. From the aligned reads, detection of covered loci was read using the gstacks program included in Stacks v2.61[64]. Filtering of detected loci was done with option ($-R = 0.75$) retaining only loci that have a 75% representation of all the samples. Heterozygosity was maintained at a maximum of 80%. HaplotypeCaller[66]. Genome Analyzer Tool Kit package version 4.2.0.0 was used for variant calling and was run in GVCF mode (https://github.com/broadinstitute/gatk/releases), with suggested parameters. Only high quality (QUAL > 30.0) biallelic SNPs were retained. Raw sequencing reads can be found at NCBI under BioProject PRJNA951214. For diversity analyses, additional filtering was applied to maintain a dataset with a minor allele frequency (MAF) > 0.05, ensuring that SNP markers and individuals with no more than 10% missing data were retained. Heterozygosity was limited to a maximum of 10%.

This set of markers underwent linkage disequilibrium (LD) pruning with a threshold of $r^2 = 0.5$. LD pruning reduces redundancy, ensuring that only independent genetic variants are considered in downstream analyses thus improving the accuracy and reliability of results[67]. This step was carried out using the *--indep-pairwise* function implemented in PLINKv1.9[68] with a window size of 50 SNPs and a step size equal to 5 SNPs. These parameters ensured the retention of minimally correlated variants, which served as for diversity analyses, genetic structure analysis, GF modelling and pRDA. To contextualize associations and outlier loci within locus-specific linkage disequilibrium (LD), we identified haplotype blocks using PLINK v1.9 with default parameters. These LD-blocks were then extended in both directions based on chromosome-specific LD decay distances and were considered to assort independently.

Genetic diversity analysis

Genetic diversity analyses were performed on a set of quality filtered SNPs using R[69]. A discriminant analysis of the principal components (DAPC) implemented in R/adegenet[70] was used to identify the presence of putative genetic groups within the core collection. A neighbor-joining phylogeny was produced with R/adegenet and a principal component analysis (PCA) was performed on the SNPs dataset to qualitatively describe genetic relatedness among the genotypes. Summary statistics including observed and expected heterozygosity, allelic richness, fixation indices (F_{IS} and F_{ST}), pi (π) and rare alleles were computed in R/adegenet to further describe the core collection.

Structure 2.3.4[71] was used to assign individuals to genetic clusters by determining the number of clusters that best fit the data. The analysis was conducted using the admixture model with 10,000 burn-in iterations followed by 100,000 MCMC repetitions. The number of genetic groups (K) was tested from $K = 1$ to $K = 10$, with 10 replicates for each value of K. The output data were then analysed in Structure Harvester[72] to calculate ad hoc statistics and identify the most likely number of clusters based on the Evanno method[73].

Acquisition of spatial and climatic data

GPS sampling locations of cowpea landrace original collection sites were obtained from available genebank passport data. South African accessions ($n = 34$) have been left out of these analyses as passport data was unavailable. We selected a subset of 153 cowpea genotypes for use in climate analyses based on a combination of factors including the availability of high-quality passport data, which was essential for accurate geographic and ecological representation. The selected genotypes represent the full spectrum of Agro-Ecological Zones (AEZs) across Southern Africa, with a focus on regions with varied climatic conditions, soil types, and farming systems (Fig. S3). Sampling points were projected onto the covered geographical area (i.e. Tanzania, Mozambique, Malawi and Zimbabwe) using R/raster[74] and altitudes of sampling points were derived using the CGIAR SRTM database at 90 m resolution[75]. Agroecological Zones (AEZs) were derived from the Food and Agricultural Organization (FAO) GAEZ v4 database. Historical climate data was derived from WorldClim version 2.1. The data is available at 2.5 min (~21 km^2 at the equator) horizontal resolution[32]. Similarly, CMIP6 downscaled future climate projections were obtained from WorldClim. The downscaling and calibration (bias correction) were done with WorldClim v2.1 as baseline climate. We used the available 13 models (Supplementary Data 17) to prepare multi-model ensembles (MME) for future projected climates with the same spatial resolution as historical data for Shared Socioeconomic Pathway (Supplementary Data 18). Using historical and future climate data, 19 biologically meaningful indicators, called bioclimatic variables, were derived using R/dismo[76]. These 19 indicators provide valuable information for assessing the suitability of specific regions for agricultural activities and crop production. Bioclimatic variables from bio1 to bio11 refer to temperature while bio12 to bio19 describe to rainfall distribution and patterns (Supplementary Data 19). To minimize collinearity among bioclimatic variables, we performed Variant Inflation Factor analysis with the *VIF()* function implemented in R/Biodiversity[77]. This procedure retained variables with a variance inflation factor below five for downstream analysis (Supplementary Data 5).

Gradient forest analysis to estimate cowpea vulnerability to future climates

A GF machine-learning regression tree-based algorithm modelled the influence of environmental factors on the genetic diversity of cowpea across its cultivation regions in Southern Africa. This analysis was implemented in *R/gradientForest*[78]. The GF analysis focused on AEZs in which representative cowpea accessions were collected. GF measures the accuracy of each predictor in estimating the genetic composition across the climatic landscape. The GF model was trained using the set of non-collinear bioclimatic variables and Moran's Eigenvector Maps (MEM) derived in *R/adespatial*[79] from the original accessions sampling points. The LD pruned SNPs were used as response variables. Using the above-described set predictors, a forest of 500 trees was built to predict the allelic state of each SNP that served as response variable. The GF model was then employed to predict and assess the difference in genetic composition between historical and future climate conditions, following Fitzpatrick and Keller's method[80]. The mismatch between the genetic composition we observe using historical bioclimatic variables as predictors and the expected genetic composition at future climate scenarios and horizons with an approach derived from Caproni et al.[81] (2023). The genomic offset of Southern Africa cowpea was estimated as Euclidean distance between the allelic turnover predicted using historical and projected climate datasets.

Characterizing the drivers of cowpea adaptation

In this study, we employed LFMM to analyze genotype-environment associations. We used the full set of high-quality SNPs, which allowed for a broader exploration of potential adaptation signals. We focused on bioclimatic variables with a Variance Inflation Factor (VIF) below 5 to minimize multicollinearity and ensure the reliability of the results. For this analysis, we used a procedure implemented in the R package LEA. In the first step of the analyses, we performed inference of individual admixture coefficients using sparse nonnegative matrix factorization[71,82,83] and used an entropy criterion that evaluates the quality of fit of the statistical model to select the number of ancestral populations that best explains the genotypic data. The number of ancestral clusters, determining the number of latent factors, was interpreted considering the outcome of the analyses carried out in admixture, and was used to correct for confounding effects due to population structure. Before testing association, we performed genotype imputation, with a LEA built-in function that carries out imputation-based ancestry coefficients.

The genome-wide association was tested using the *lfmm()* function[45], a Bayesian approach that uses a Monte-Carlo Markov Chain algorithm. After correction for confounding effects, the level of association between the tested bioclimatic variables and the molecular markers is measured and mapped to the corresponding genomic position and represented as a Manhattan plot.

Marker-Environment Associations were also studied by means of pRDA using functions developed in R/vegan[84]. This analysis was conducted using LD-pruned genetic markers to minimize redundancy to improve the accuracy and reliability of our results[23]. The environmental diversity was described as non-collinear bioclimatic variables, geography by means GPS coordinates while genetic structure using the first three principal components of the genetic PCA as in Caproni et al.[81] (2023). The proportion of genetic variance which could be attributed to each set of explanatory variables was estimated in a step-wise manner[23]. We used RDA analysis to identify genomic loci underlying signatures of adaptation. To the purpose we used the R function *rdadapt()*[23]. The adaptive loci were identified based on their position along a Mahalanobis distance D distribution, corrected for the inflation factor to derive p-values using a chi-squared distribution with two degrees of freedom[43]. Finally, a Bonferroni threshold at a nominal p-value significance level of 0.05 was used for outlier identification.

Detection of candidate genes

To identify gene functions potentially associated with climatic adaptation -related SNPs, annotations were searched in the *Vigna unguiculata* genome (ASM411807v2) from the genome browser: https://phytozome-next.jgi.doe.gov/info/Vunguiculata_v1_2[85]. The LD decay was estimated for each of the chromosomes using a threshold of $r^2 = 0.22$ which represents half of the maximum interpolated r-squared value (around 0.45 across all chromosomes)[86] (Supplementary Data 14). Gene models and associated protein coding genes were searched in regions within LD blocks, then extended in both directions based on chromosome-specific LD decay distances. The genes located in the candidate regions were considered candidate genes.

Statistics and reproducibility

We applied multiple statistical tests to ensure the robustness of our findings. To minimize genetic linkage bias, we applied linkage disequilibrium (LD) pruning using PLINK v1.9 with an r^2 threshold of 0.5. The resulting set of SNPs was used for Marker-environment association (MEA) analyses using pRDA and for the GF modelling approach. Adaptive loci were indeed identified using multiple approaches. LFMM was also used with the full set of 18,969 markers. With this approach, we detected significant associations between SNPs and bioclimatic variables using a Bonferroni correction to control for false positives, with an alpha of 0.05. The pRDA significance was assessed with the same approach using an alpha of 0.05. The GF model predicted SNP turnover along environmental gradients, identifying 787 SNPs with predictive power ($r^2 > 0$) and selecting those above the 99th percentile of R^2. Reproducibility was ensured by using publicly available datasets, and analysis scripts have been provided for transparency.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data generated or analyzed during this study are included in the published article and its supplementary materials. Source data for Fig. 1 can be found in Supplementary Data 1, for Fig. 4 in Supplementary Data 10, and for Fig. 5 in Supplementary Data 12 and Supplementary Data 13. Raw reads are available at the National Center for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov) under BioProject PRJNA951214.

Code availability

Scripts used to analyse the data are available at https://github.com/mdellh2o/cowpeaMOZ and archived on Zenodo https://doi.org/10.5281/zenodo.15384460.

Received: 19 September 2024; Accepted: 14 May 2025;

References

1. Lowder, S. K., Sánchez, M. V. & Bertini, R. Which farms feed the world and has farmland become more concentrated? *World Dev.* **142**, 105455 (2021).
2. Giller, K. E. et al. Small farms and development in sub-Saharan Africa: farming for food, for income or for lack of better options?. *Food Sec.* **13**, 1431–1454 (2021).
3. Mekonnen, T. W., Gerrano, A. S., Mbuma, N. W. & Labuschagne, M. T. Breeding of vegetable cowpea for nutrition and climate resilience in sub-Saharan Africa: Progress, opportunities, and challenges. *Plants* **11**, 1583 (2022).
4. Omotoso, A. B. et al. Climate change and variability in sub-Saharan Africa: a systematic review of trends and impacts on agriculture. *J. Clean. Prod.* **414**, 137487 (2023).
5. IPCC. Climate change 2022: impacts, adaptation, and vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press*: https://doi.org/10.1017/9781009325844 (2022).
6. Gerrano, A. S., Lubinga, M. H. & Bairu, M. W. Genetic resources management, seed production constraints, and trade performance of orphan crops in Southern Africa: a case of Cowpea. *S. Afr. J. Bot.* **146**, 340–347 (2022).
7. Mohammed, S. B. et al. Appraisal of cowpea cropping systems and farmers' perceptions of production constraints and preferences in the dry savannah areas of Nigeria. *CABI Agric. Biosci.* **2**, 46 (2021).
8. Flight, C. The Kintampo culture and its place in the economic prehistory of West Africa in *Origins of African Plant Domestication* 211–221 (Mouton, 1976).
9. Mwale, S. E. et al. Response of cowpea genotypes to drought stress in Uganda. *Am. J. Plant Sci.* **8**, 720–733 (2017).
10. FAO. FAOSTAT Statistical Database. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/ (2022).
11. Bolarinwa, K. A. et al. An investigation of cowpea production constraints and preferences among smallholder farmers in Nigeria. *GeoJournal* **87**, 2993–3005 (2022).
12. Oyerinde, A. A., Chuwang, P. Z. & Oyerinde, G. T. Evaluation of the effects of climate change on increased incidence of cowpea pests in Nigeria. *J. Plant Prot. Sci.* **5**, 10–16 (2013).
13. Afutu, E. et al. Cowpea scab disease (*Sphaceloma sp.*) in Uganda. *Crop Prot.* **92**, 213–220 (2017).
14. Omomowo, O. I. & Babalola, O. O. Constraints and prospects of improving cowpea productivity to ensure food, nutritional security, and environmental sustainability. *Front. Plant Sci.* **12**, 751731 (2021).
15. Gomes, A. M. F. et al. Diversity of cowpea *Vigna unguiculata* (L.) Walp. landraces in Mozambique: new opportunities for crop improvement and future breeding programs. *Agronomy* **11**, 991 (2021).
16. Sansaloni, C. et al. Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nat. Commun.* **11**, 18404 (2020).
17. Menamo, T. et al. Genetic diversity of Ethiopian sorghum reveals signatures of climatic adaptation. *Theor. Appl. Genet.* **134**, 731–742 (2021).
18. Rhoné, B. et al. Pearl millet genomic vulnerability to climate change in West Africa highlights the need for regional collaboration. *Nat. Commun.* **11**, 5274 (2020).
19. Yuan, S. et al. Genomic vulnerability to climate change in *Quercus acutissima*, a dominant tree species in East Asian deciduous forests. *Mol. Ecol.* **32**, 1639–1655 (2023).
20. Jägermeyr, J. et al. Climate impacts on global agriculture emerge earlier in the new generation of climate and crop models. *Nat. Food* **2**, 873–885 (2021).
21. Guimarães, J. B. et al. Genetic diversity and population structure of cowpea (*Vigna unguiculata* (L.) Walp.) landraces from Portugal and Mozambique. *Plants* **12**, 846 (2023).
22. Lonardi, S. et al. The genome of cowpea (*Vigna unguiculata* [L.] Walp. *Plant J.* **98**, 767–782 (2019).
23. Capblancq, T. & Forester, B. R. Redundancy analysis: a Swiss Army Knife for landscape genomics. *Methods Ecol. Evol.* **12**, 2298–2309 (2021).
24. Mercer, K. L. & Perales, H. R. Evolutionary response of landraces to climate change in centers of crop diversity. *Evol. Appl.* **3**, 480–493 (2010).
25. Spear, D. D. et al. Vulnerability and adaptation to climate change in semi-arid areas in Southern Africa. *Technical Report*; https://doi.org/10.13140/RG.2.2.32046.92488 (2018).
26. Kuruma, R. W., Sheunda, P. & Kahwaga, C. M. Yield stability and farmer preference of cowpea (*Vigna unguiculata*) lines in semi-arid Eastern Kenya. *Afr. Focus* **32**, 65–82 (2019).
27. Woldeyohannes, A. B. et al. Data-driven, participatory characterization of farmer varieties discloses teff breeding potential under current and future climates. *eLife* **11**, e80009 (2022).
28. Burton, G. et al. Climate change, food systems, and adaptive livelihoods in rural Africa: towards resilient, inclusive agricultural systems. *Farming Syst.* **3**, 149–164 (2021).
29. Huynh, B.-L. et al. Gene pools and the genetic architecture of domesticated cowpea. *Plant Genome* **6**, plantgenome2013.03.0005 (2013).
30. Younger, J. L., Clucas, G. V. & Kao, D. The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. *Mol. Ecol.* **26**, 3883–3897 (2017).
31. Chiulele, R. M. et al. Assessment of farmers' perceptions and preferences of cowpea in Mozambique. *Afr. Crop Sci. Conf. Proc.* **10**, 311–318 (2011).
32. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
33. Ismail, A. M. & Hall, A. E. Positive and potential negative effects of heat-tolerance genes in cowpea. *Crop Sci.* **38**, 381–390 (1998).
34. Barros, J. R. A. et al. Selection of cowpea cultivars for high temperature tolerance: physiological, biochemical and yield aspects. *Physiol. Mol. Biol. Plants* **27**, 29–38 (2021).
35. Craufurd, P. Q., Bojang, M., Wheeler, T. R. & Summerfield, R. J. Heat tolerance in cowpea: effect of timing and duration of heat stress. *Ann. Appl. Biol.* **133**, 257–267 (1998).
36. Mohammed, S. B., Ongom, P. O., Togola, A. & Boukar, O. Enhancing cowpea tolerance to elevated temperature: achievements, challenges and future directions. *Agronomy* **14**, 513 (2024).

37. Nielsen, C. L. & Hall, A. E. Responses of cowpea in the field to high night air temperature during flowering. II. Plant responses. *Field Crops Res.* **10**, 181–196 (1985).
38. Hall, A. E. Breeding for adaptation to drought and heat in cowpea. *Eur. J. Agron.* **21**, 447–454 (2004).
39. Barros, J. R. A. et al. Pollen viability, and the photosynthetic and enzymatic responses of cowpea (*Vigna unguiculata* (L.) Walp., *Fabaceae*) in the face of rising air temperature: a problem for food safety. *Agronomy* **14**, 463 (2024).
40. Ravelombola, W. et al. Evaluation of cowpea for drought tolerance at seedling stage. *Euphytica* **216**, 123 (2020).
41. Ngompe Deffo, T. et al. Identifying critical growth stage and resilient genotypes in cowpea under drought stress contributes to enhancing crop tolerance for improvement and adaptation in Cameroon. *PLoS One* **19**, e0304674 (2024).
42. Dauphin, B. et al. Re-thinking the environment in landscape genomics. *Trends Ecol. Evol.* **38**, 261–274 (2023).
43. Forester, B. R., Lasky, J. R., Wagner, H. H. & Urban, D. L. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Mol. Ecol.* **27**, 2215–2233 (2018).
44. Capblancq, T., Luu, K., Blum, M. G. B. & Bazin, E. Evaluation of redundancy analysis to identify signatures of local adaptation. *Mol. Ecol. Resour.* **18**, 1223–1233 (2018).
45. Frichot, E., Schoville, S. D., Bouchard, G. & François, O. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* **30**, 1687–1699 (2013).
46. Fitzpatrick, M. C., Keller, S. R. & Lotterhos, K. E. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Mol. Ecol. Resour.* **21**, 2749–2765 (2021).
47. Ågren, J., Oakley, C. G., Lundemo, S. & Schemske, D. W. Adaptive divergence in flowering time among natural populations of *Arabidopsis thaliana*: Estimates of selection and QTL mapping. *Evolution* **71**, 550–564 (2017).
48. Price, N., Lopez, L., Platts, A. E. & Lasky, J. R. In the presence of population structure: From genomics to candidate genes underlying local adaptation. *Ecol. Evol.* **10**, 1889–1904 (2020).
49. Mishra, P. & Panigrahi, K. C. *GIGANTEA*–an emerging story. *Front. Plant Sci.* **6**, 8 (2015).
50. Park, Y.-J. et al. *GIGANTEA* shapes the photoperiodic rhythms of thermomorphogenic growth in arabidopsis. *Mol. Plant.* **13**, 459–470 (2020).
51. Keller, S. R., Levsen, N., Olson, M. S. & Tiffin, P. Local adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera* L. *Mol. Biol. Evol.* **29**, 3143–3152 (2012).
52. de Montaigu, A. & Coupland, G. The timing of GIGANTEA expression during day/night cycles varies with the geographical origin of Arabidopsis accessions. *Plant Signal Behav.* **12**, e1342026 (2017).
53. Paudel, D. et al. Genome-wide association study reveals candidate genes for flowering time in cowpea (*Vigna unguiculata* [L.] Walp.). *Front. Genet.* **12**, 667038 (2021).
54. Ma, J. et al. GRAS transcription factors mediate flowering through signaling pathways of gibberellin and circadian rhythm in *Lonicera japonica* Thunb. *Plant Gene* **28**, 100340 (2021).
55. Cheng, X. et al. *NO APICAL MERISTEM* (*MtNAM*) regulates floral organ identity and lateral organ separation in Medicago truncatula. *N. Phytol.* **195**, 71–84 (2012).
56. Shao, H., Wang, H. & Tang, X. NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Front. Plant Sci.* **6**, 156056 (2015).
57. Lewis, M. W. et al. The SERK1 receptor-like kinase regulates organ separation in Arabidopsis flowers. *Plant J.* **62**, 817–828 (2010).
58. Tiwari, P. & Chakrabarty, D. Dehydrin in the past four decades: from chaperones to transcription co-regulators in regulating abiotic stress response. *Curr. Res. Biotechnol.* **3**, 249–259 (2021).
59. Sun, Z. et al. Plant dehydrins: expression, regulatory networks, and protective roles in plants challenged by abiotic stress. *Int. J. Mol. Sci*. **22** https://doi.org/10.3390/ijms222312619 (2021).
60. Guo, H. et al. Differentiation, evolution and utilization of natural alleles for cold adaptability at the reproductive stage in rice. *Plant Biotechnol. J.* **18**, 2491–2503 (2020).
61. Haque, T. et al. Salt tolerance QTLs of an endemic rice landrace, Horkuch at seedling and reproductive stages. *Sci. Rep.* **12**, 17306 (2022).
62. Hammer, K., Arrowsmith, N. & Gladis, T. Agrobiodiversity with emphasis on plant genetic resources. *Naturwissenschaften* **90**, 241–250 (2003).
63. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**, e37135 (2012).
64. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).
65. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
66. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*. **201178** https://doi.org/10.1101/201178 (2018).
67. Aguirre, N. C. et al. Comparison of ddRADseq and EUChip60K SNP genotyping systems for population genetics and genomic selection in *Eucalyptus dunnii* (Maiden). *Front. Genet.* **15**, 1361418 (2024).
68. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
69. R Core Team. R: The R Project for Statistical Computing.CRAN. https://www.r-project.org/ (accessed 10 Jul 2023).
70. Jombart, T., Devillard, S. & Balloux, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94 (2010).
71. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
72. Earl, D. A. & vonHoldt, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
73. 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 (2005).
74. Hijmans, R. J. raster: Geographic data analysis and modeling. R package version 2.8 (2018).
75. Reuter, H. I., Nelson, A. & Jarvis, A. An evaluation of void-filling interpolation methods for SRTM data. *Int. J. Geogr. Inf. Sci.* **21**, 983–1008 (2007).
76. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species distribution modeling. *R. package version* **1**, 1–4 (2017).
77. Kindt, R. & Coe, R. Tree diversity analysis: a manual and software for common statistical methods for ecological and biodiversity studies. *World Agroforestry Centre* (2005).
78. Ellis, N., Smith, S. J. & Pitcher, C. R. Gradient forests: calculating importance gradients on physical predictors. *Ecology* **93**, 156–168 (2012).
79. Dray, S. et al. Adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.3-23, https://CRAN.R-project.org/package=adespatial (2023).
80. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic

landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).

81. Caproni, L. et al. The genomic and bioclimatic characterization of Ethiopian barley (*Hordeum vulgare* L.) unveils challenges and opportunities to adapt to a changing climate. *Glob. Chang. Biol.* **29**, 2335–2350 (2023).
82. Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G. & François, O. Fast and efficient estimation of individual ancestry coefficients. *Genetics* **196**, 973–983 (2014).
83. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
84. Oksanen, J. et al. vegan: Community Ecology Package. R package version 2.3-2. https://CRAN.R-project.org/package=vegan (2015).
85. Goodstein, D. M. et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186 (2012).
86. Hill, W. G. & Weir, B. S. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* **33**, 54–78 (1988).

Acknowledgements

We thank the IIAM technical team, Frederico Madabula, Fidélio Gonçalves, Nilso da Silva, Camilo Bruno and Agnaldo Alicete who assisted with green house experiments and DNA extraction. We also thank the IITA genebank, Nigeria, for providing additional germplasm. We extend our thanks to PLAN International for assisting the activites in Mozambique. This work was financially supported by the European Commission's Horizon 2020 Research and Innovation Programme through the FOCUS-Africa project (grant agreement n. 869575). Additional funding was provided by the TALISMAN project, supported by the Italian Ministry of Foreign Affairs (MAECI) and by the Patent21 and the PRISMA project, funded by the Italian Ministry of University and Research (MUR), under the National Plan for Research (PNR) within the framework of the PNRR 2021–2026. We also acknowledge the support of the Doctoral School of Agrobiodiversity at Scuola Superiore Sant'Anna, Pisa, Italy.

Author contributions

M.D. designed the study and supervised research. M.S. contributed to the assembly of the collection and in the generation of genomic data. R.M.C. provided support in germplasm acquisition and analysis. M.E.P., R.B., P.M., and M.A. contributed to experimental design and interpretation of results. M.W.M. managed molecular experiments and data analysis with contributions from L.C. L.C. and R.T. contributed with analysis of climatic data. M.W.M. drafted the manuscript, interpreted results, and produced figures with support of L.C. All authors have read and approved the manuscript.

Competing interests

The authors declare no competing interests. Matteo Dell'Acqua is an Editorial Board Member for Communications Biology, but was not involved in the editorial review of, nor the decision to publish this article.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-08227-0.

Correspondence and requests for materials should be addressed to Matteo Dell'Acqua.

Peer review information *Communications Biology* thanks Celestina Jochua, Maria Sanchez and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Jorge Duitama and David Favero. A peer review file is available.