



Non-volatile bioactive compounds and best linear unbiased prediction (BLUP)-based genetic estimates in wild African coffee accessions: the case of *Coffea racemosa* and *Coffea zanguebariae*

Willian dos Santos Gomes · Niquisse José Alberto · Cristhiane Altoé Filete · Joana Scarparo Novello · Gracieli Lorenzoni Marotto · Emanuele Catarina da Silva Oliveira · Taís Rizzo Moreira · Aldemar Polonini Moreli · José C. Ramalho · Fábio Luiz Partelli

Received: 23 February 2026 / Accepted: 6 April 2026 / Published online: 17 April 2026
© The Author(s) 2026

Abstract The sustainability of global coffee production might be threatened by climate change and associated disease outbreaks, prompting the interest in underutilized wild *Coffea* species as reservoirs of genetic and chemical diversity. This study aimed to characterize the non-volatile chemical profiles and estimate genetic values for caffeine, trigonelline, and chlorogenic acid (CGA) of beans from 24 *Coffea racemosa* and 21 *C. zanguebariae* wild accessions from distinct Mozambican regions. Fruits were collected at full maturity, dried under standard conditions to stabilize the material, and analyzed by high-performance liquid chromatography (HPLC). Significant inter- and intra-specific variations were detected for all compounds. *C. racemosa* exhibited low mean caffeine content (7.3 mg g⁻¹ DW) and high

CGA levels (24.4 mg g⁻¹ DW), showing potential for naturally low-caffeine beverages. In sharp contrast, *C. zanguebariae* presented higher caffeine content (21.9 mg g⁻¹ DW) and intermediate CGA levels. A principal component analysis revealed three distinct chemical clusters per species, driven predominantly by caffeine and trigonelline. Additionally, in *C. racemosa* a Best Linear Unbiased Prediction (BLUP) analysis showed a large range of values from -2.00 to 1.94, pointing to genetic variability for CGAs, and considerable potential for selection, in contrast with *C. zanguebariae* that displayed narrower BLUP ranges for all compounds. These findings highlighted

W. dos Santos Gomes · J. S. Novello · G. L. Marotto · E. C. da Silva Oliveira · A. P. Moreli
Coffee Design Group, Federal Institute of Espírito Santo, Campus Venda Nova do Imigrante, Av. Elizabeth Minete Perim, S/N, Bairro São Rafael, Venda Nova do Imigrante, Espírito Santo 29375-000, Brazil

N. J. Alberto
Genetic Breeding Program, Federal University of Espírito Santo/UFES, Campus Alegre, Bairro Guararema, S/N, Alegre, Espírito Santo CEP 29500-000, Brazil

C. A. Filete
Department of Chemistry, Federal University of Espírito Santo/UFES, Campus Goiabeiras, Avenida Fernando Ferrari, 514, Vitória, Espírito Santo CEP 29075-910, Brazil

T. R. Moreira
Laboratório de Geoprocessamento, Sensoriamento Remoto e Análises de Dados de Biodiversidade (GEODABI), Instituto Nacional da Mata Atlântica (INMA), Rua Bernadino Monteiro (Rodovia Josil Espíndula Agostini), nº 261, Jardim da Montanha, Santa Teresa, Espírito Santo CEP 29650-0000, Brazil

J. C. Ramalho
Associate Laboratory TERRA, Forest Research Center (CEF), School of Agriculture (ISA), University of Lisbon (ULisboa), Quinta Do Marquês, Av. República, 2784-505, Oeiras and Tapada da Ajuda, 1349-017 Lisbon, Portugal

F. L. Partelli (✉)
Department of Agrarian and Biological Sciences (DCAB), Federal University of Espírito Santo/UFES, Rod. BR 101 Norte, Km 60, Bairro Litorâneo, São Mateus, Espírito Santo CEP 29932-540, Brazil
e-mail: partelli@yahoo.com.br

distinct biochemical profiles and genetic potentials of these underutilised wild *Coffea* species, emphasizing *C. racemosa* as a promising genetic resource for the development of differentiated, bioactive-rich coffee products and providing novel data for future breeding and conservation strategies.

Keywords Caffeine · CGAs · Genetic diversity · Trigonelline · Underused wild *Coffea* species potential

Introduction

Wild species of the *Coffea* genus represent an under-exploited genetic resource with potential to diversify and strengthen the resilience of global coffee cultivation, as well as to be introduced in the coffee value chain to be used as new distinct/specialty products. These species have also been described as niche coffee types with potential commercial value, particularly *Coffea racemosa*, which has attracted interest for differentiated beverages, specialty markets, and the diversification of coffee production systems (Davis et al. 2021).

Among wild underused species, *Coffea racemosa* Lour. and *C. zanguebariae* Lour., native to eastern and southern Africa, stand out for their characteristics to face climate change impacts regarding to biotic and abiotic stresses, such as high tolerance to thermal and water stress (Guerreiro-Filho 1992; Davis et al. 2021), while exhibiting naturally low caffeine content relative to cultivated *Coffea* species, and distinct reproductive cycles (Davis et al., 2019). Therefore, these species, which are naturally restricted to specific regions of Mozambique, Tanzania, and South Africa, might offer promising adaptive and phytochemical attributes, with potential applications in both conventional breeding programs and the production of differentiated coffees aimed at niche markets and potential interest for functional or differentiated coffee products.

Wild *Coffea* species have long been considered important genetic resources for coffee breeding programs, including the use of *Coffea racemosa* in interspecific hybridization with *Coffea arabica* aiming to improve tolerance to abiotic stress and to broaden the genetic base of cultivated coffee (Guerreiro-Filho 1992; Ky et al. 2001; Davis et al. 2021).

Coffea racemosa, endemic to a narrow coastal strip between northern KwaZulu-Natal in South Africa and to southern Inhambane province in Mozambique, but was recently shown a wider distribution in central and southern Mozambique (Tapaça et al. 2023). This species exhibits valuable agronomic traits such as high tolerance to thermal and drought stress and naturally low caffeine levels. *Coffea zanguebariae*, is more associated with Cabo Delgado (in the North). Although less documented, seemed to share somewhat similar attributes, including resistance to adverse environmental conditions and a rapid fruiting cycle; however, its chemical composition remains poorly understood, limiting the assessment of its potential for new agronomic and industrial applications (Davis et al. 2021; Tapaça et al. 2023; Navarini et al. 2024). Despite this potential, scientific knowledge regarding the chemical characteristics associated with the beverage quality of these species remains incipient, particularly in relation to non-volatile compounds such as caffeine, trigonelline, and chlorogenic acids are important constituents related to beverage quality, although coffee sensory characteristics result from the interaction of multiple chemical classes, including volatile compounds, sugars, lipids, and roasting-derived products (Lyrio et al. 2025; Ribeiro et al. 2014; Hameed et al., 2018), which are impacted by climatic changes and global warming (Ramalho et al. 2018).

Previous research demonstrated substantial diversity in caffeine and chlorogenic acid contents among wild *Coffea* species, highlighting their importance for both ecological adaptation and beverage quality (Campa et al. 2005; Ky et al. 2001). Still, recent advances have been made in the distribution, as well as genetic, morphological, agronomic and climate profiles characterization of *C. racemosa* and *C. zanguebariae* (Davis et al. 2021; Tapaça et al. 2023). In addition, recent metabolomic research expanded the chemical characterization of wild *Coffea* species (Lyrio et al. 2025), information regarding intra-specific variability among natural populations and predicted genetic values for key bioactive compounds remains scarce, particularly for accessions collected in situ under distinct environmental conditions.

Given the growing demand for genetically diverse materials adapted to climate change scenarios and with differentiated chemical profiles, the systematic evaluation of these species becomes strategically

important. Also, in addition to quantifying the main bioactive compounds, the application of predictive methodologies such as Best Linear Unbiased Prediction (BLUP) will constitute a step forward since it provide estimates of genotypic values and identify superior accessions for different agronomic and sensory purposes (Piepho et al. 2008). In this context, the present study aimed to provide detailed information on the intra- and inter-specific chemical diversity of key bioactive compounds related to coffee beverage quality (namely caffeine, trigonelline, and chlorogenic acid [CGAs]) in *Coffea racemosa* and *C. zanguebariae* accessions from Mozambique. Additionally, the study included the estimation of average content ranges for each compound across accessions. These findings contribute to expanding the phytochemical understanding of these species and support the identification of genotypes with desirable biochemical profiles, which could be strategically integrated into breeding programs or cultivation systems aiming to enhance coffee quality and resilience to environmental challenges.

Material and methods

Characterization of plant material and experimental design

In situ samples were collected in 2023 and 2024, in different regions of Mozambique, including plant populations occurring under natural or naturalized conditions. Collections of *Coffea zanguebariae* Lour. were carried out in northeastern Mozambique, specifically on Ibo Island (Cabo Delgado Province), where the species is currently cultivated or naturalized, although this region may not fully correspond to its original natural distribution (Davis et al. 2021). Samples of *Coffea racemosa* Lour. were obtained in the southern region of the country, including the districts of Chidenguele (Gaza Province) and Murumbene, Maxixe, Inharrime, Panda, and Zavala (Inhambane Province), which are within the known distribution range of this species.

Collections of *Coffea zanguebariae* Lour. were carried out in northeastern Mozambique, specifically on Ibo Island, Cabo Delgado Province, while samples of *C. racemosa* Lour. were obtained in the southern region of the country, covering the districts

of Chidenguele (Gaza Province) and Murumbene, Maxixe, Inharrime, Panda, and Zavala (Inhambane Province).

At each location, fruit samples were collected alongside records of geographic coordinates (Fig. 1). The *C. racemosa* and *C. zanguebariae* samples were obtained in December 2023, and February 2024, respectively, all of fruits at maturation the stage.

The sampling design included 24 accessions of *C. racemosa* and 21 accessions of *C. zanguebariae*. For each selected plant, between 150 and 200 g of ripe fruits were harvested. The samples were packed in properly labelled paper bags and subjected to sun drying, followed by oven drying with forced air circulation at 65 °C until a constant weight was achieved. The drying temperature was selected to ensure uniform moisture removal and sample stabilization prior to chemical analysis, following procedures commonly used in research of green coffee composition. The same drying protocol was applied to all samples to avoid differential degradation among accessions. Subsequently, the fruits were manually hulled, separating the beans from the husks.

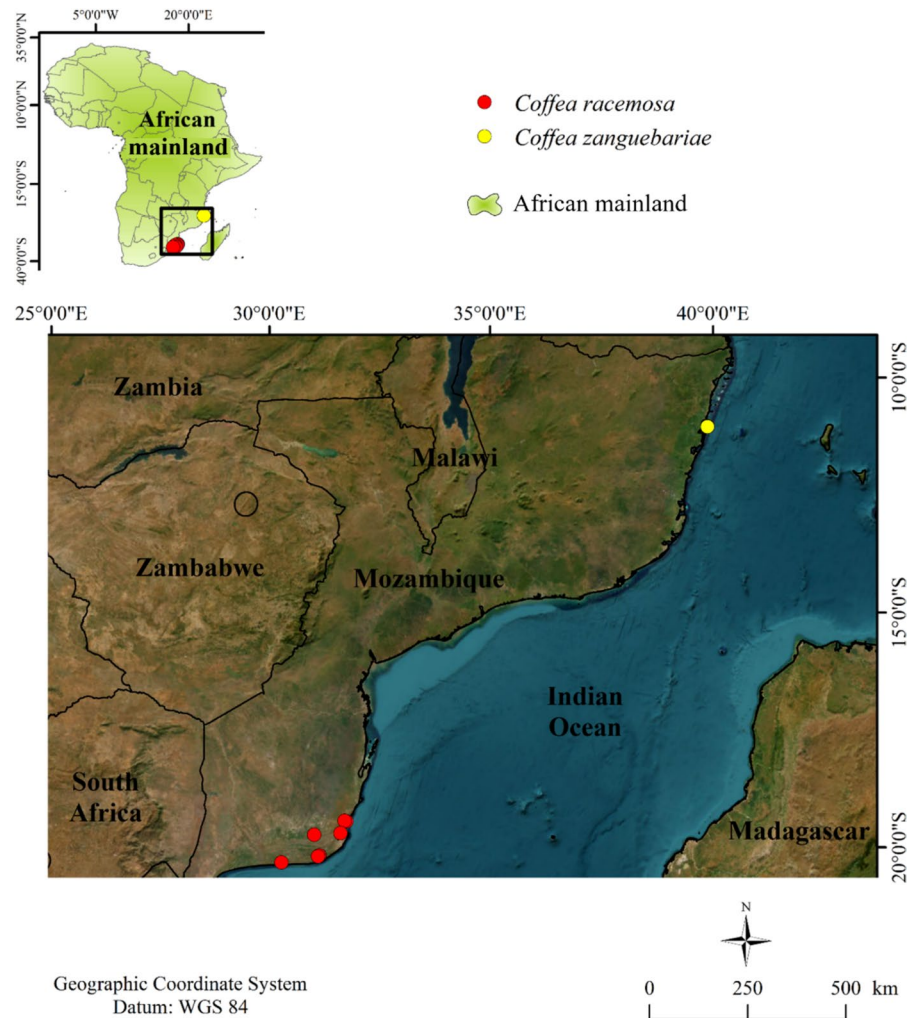
Determination of beans CGAs, caffeine, and trigonelline by high-performance liquid chromatography (HPLC)

Extractions were performed according to the method described by Trugo and Macrae (1984). A total of 500 mg of ground coffee green beans was weighed into an Erlenmeyer flask and dispersed in 100 mL of ultrapure water heated to 80 °C. This solution was stirred for 15 min in a shaker, and samples were then microfiltered (0.40 µm cellulose acetate filter) into a 100 mL volumetric flask, and the volume was made up with Milli-Q water.

A total of 500 mg of ground green coffee beans was weighed and transferred to an Erlenmeyer flask, then dispersed in 100 mL of ultrapure water heated to 80 °C. The solution was stirred for 15 min on a shaker, followed by microfiltration (40 µm cellulose acetate filter). The filtrate was transferred to a volumetric flask, and the volume was made up to exactly 100 mL with Milli-Q water.

An aliquot of the filtered solution (1 mL) was subsequently filtered again through a 0.45 µm MCE—Cellulose Esters membrane filter and transferred to a clean vial, followed by injection into

Fig. 1 Locations of the collection sites for *Coffea racemosa* and *C. zanguebariae* accessions



the liquid chromatography (HPLC) system from Shimadzu Corp., Japan, equipped with a UV detector set at 272 nm and a refractive index detector. Chromatographic separation was achieved using a C₁₈ chromatographic column (Aminex HPX-87H) (300 cm×7.8 mm) maintained at 40 °C, with a mobile phase consisting of acetic acid and methanol at 5 mM, delivered at a flow rate of 1.0 mL min⁻¹. Compounds were identified by comparing their retention times with those of analytical standards and quantified using calibration curves.

Quantification was performed using external calibration curves prepared with analytical standards. Each sample was analyzed in repeated measurements to ensure analytical consistency.

The analytical procedure followed previously described methods for the determination of caffeine, trigonelline, and chlorogenic acids in green coffee beans (Bicho et al. 2013; Ribeiro et al. 2014).

Statistical analysis

The data regarding the concentrations of CGAs, caffeine, and trigonelline were standardized using the z-score method to equalize the variable scales. Subsequently, a Hierarchical Cluster Analysis (HCA) was performed, employing Euclidean distance as the dissimilarity measure and the Ward's linkage method (Ward.D2) for cluster formation.

To enhance the visualization and interpretation of the clusters, a Principal Component Analysis

(PCA) was conducted based on the Euclidean distance matrix. All statistical analyses were performed in the R environment, using the *vegan*, *cluster*, *factoextra*, *tidyverse*, and *ggplot2* packages.

Cluster validation was performed separately for each species using three complementary approaches. Model-based clustering was conducted using the MCLUST algorithm implemented in the R package *mclust*, assuming three groups based on PCA structure. Cluster stability was evaluated by bootstrap resampling (500 iterations) combined with k-means clustering, using silhouette values as stability metrics.

Differences among clusters were tested using PERMANOVA based on Euclidean distance matrices with 999 permutations using the *vegan* package.

All analyses were performed in R (R Core Team 2024).

In addition, aiming to estimate the predicted genetic values for the contents of caffeine, trigonelline, and CGAs among the genotypes, while minimizing the influence of environmental factors, a Best Linear Unbiased Prediction (BLUP) analysis was carried out on samples of *Coffea racemosa* and *Coffea zanguebariae*.

The data obtained were organized in a spreadsheet and imported into the R software (R Core Team 2024). The categorical variables corresponding to genotypes and replications were converted into factors to ensure their proper recognition within the statistical models.

For data modelling, mixed linear models were fitted using restricted maximum likelihood (REML) via the *lmer* function from the *lme4* package. The model adopted for each chemical variable evaluated was as follows:

$$Y_{ij} = \mu + G_i + R_j + \epsilon_{ij}$$

where Y_{ij} represents the observed value of the response variable, μ is the overall mean, G_i is the random effect of the i -th genotype, R_j is the random effect of the j -th replication, and ϵ_{ij} is the random error term, assumed to be normally distributed and homoscedastic.

In this model, the replication term refers to repeated analytical measurements performed for each accession during the chemical analysis, corresponding to independent laboratory determinations used to estimate experimental error.

The model fit quality was assessed using the marginal and conditional coefficients of determination (R_m^2 , R_c^2), calculated with the *r2_nakagawa* function from the *performance* package.

To obtain the estimated predicted values of the genotypic effects, the Best Linear Unbiased Predictions (BLUPs) were extracted using the *ranef* function from the *lme4* package. These values were organized and displayed in scatter plots, constructed with the assistance of the *ggplot2* package, to facilitate the visualization and interpretation of the relative performance among the genotypes.

It is important to emphasize that the accessions were sampled in situ from distinct natural populations occurring under different environmental conditions. Therefore, environmental heterogeneity was not experimentally controlled through a common garden trial. In this context, the BLUP estimates obtained in this study represent predicted genotypic effects under the specific environmental conditions sampled and should be interpreted as exploratory indicators of genetic differentiation rather than definitive breeding values.

Results

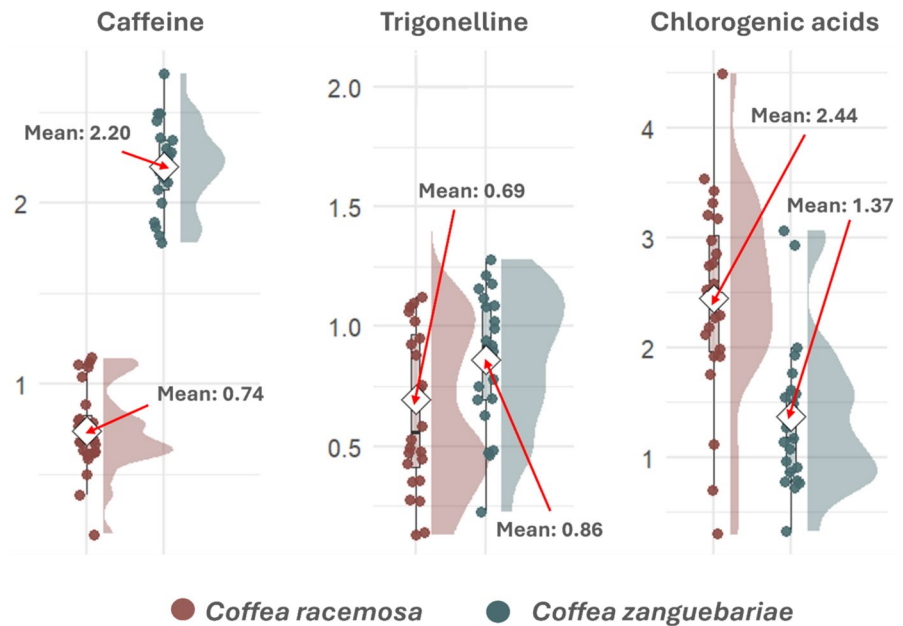
Quantification of bioactive compounds

The characterization of bioactive compounds in the 24 accessions of *Coffea racemosa* revealed mean values of 24.4 mg g⁻¹ DW for CGAs, 7.4 mg g⁻¹ DW for caffeine, and 6.9 mg g⁻¹ DW for trigonelline. However, for these compounds showed a wide range of values among accessions, as evidenced by the dispersion of individual values represented in Fig. 2, with genotypes exhibiting values both above and below the respective means.

For the 24 accessions of *Coffea zanguebariae*, for which reports on this secondary metabolite composition remain scarce in the literature, the present study established mean values of 13.7 mg g⁻¹ DW for CGAs, 22 mg g⁻¹ DW for caffeine, and 8.6 mg g⁻¹ DW for trigonelline.

The chemical diversity observed among the accessions in both species allowed the identification of distinct multivariate profiles, which were visualized using Principal Component Analysis (PCA) (Fig. 3).

Fig. 2 Distribution of caffeine, trigonelline, and chlorogenic acid contents in accessions of *Coffea zanguebariae* ($n=21$) and *Coffea racemosa* ($n=24$). Data expressed in g per 100 g



The cluster structure observed in the PCA was statistically validated for each species using model-based clustering (MCLUST), bootstrap resampling, and PERMANOVA based on Euclidean distance matrices.

For *Coffea racemosa*, the first two principal components (PC1 and PC2) explained 60.95% of the total data variability (Fig. 3A). The distribution of accessions suggested the formation of three clusters, which was supported by MCLUST analysis (silhouette=0.20). Bootstrap resampling indicated good cluster stability (0.42 ± 0.07), and PERMANOVA detected significant multivariate differences among clusters ($F=5.11$, $p=0.003$).

Cluster 1, represented by accessions 1, 7, 11, 13, and 16, was located in the positive region of PC1 and was associated with higher concentrations of trigonelline and caffeine, according to the direction of the variable vectors. Cluster 2, comprising the majority of accessions, was concentrated near the origin and in the negative region of PC1, being associated with higher chlorogenic acid contents. Cluster 3, composed of accessions 14, 19, and 23, was positioned in the lower left quadrant, grouping accessions with relatively lower concentrations of the three compounds analyzed.

For *Coffea zanguebariae*, the first two principal components explained 78.1% of the total variance

(Fig. 3B). The distribution of accessions also suggested the presence of three groups, which was supported by MCLUST (silhouette=0.32), bootstrap resampling (0.38 ± 0.07), and PERMANOVA ($F=7.47$, $p=0.001$), indicating significant multivariate differences among clusters based on the composition of caffeine, trigonelline, and chlorogenic acids.

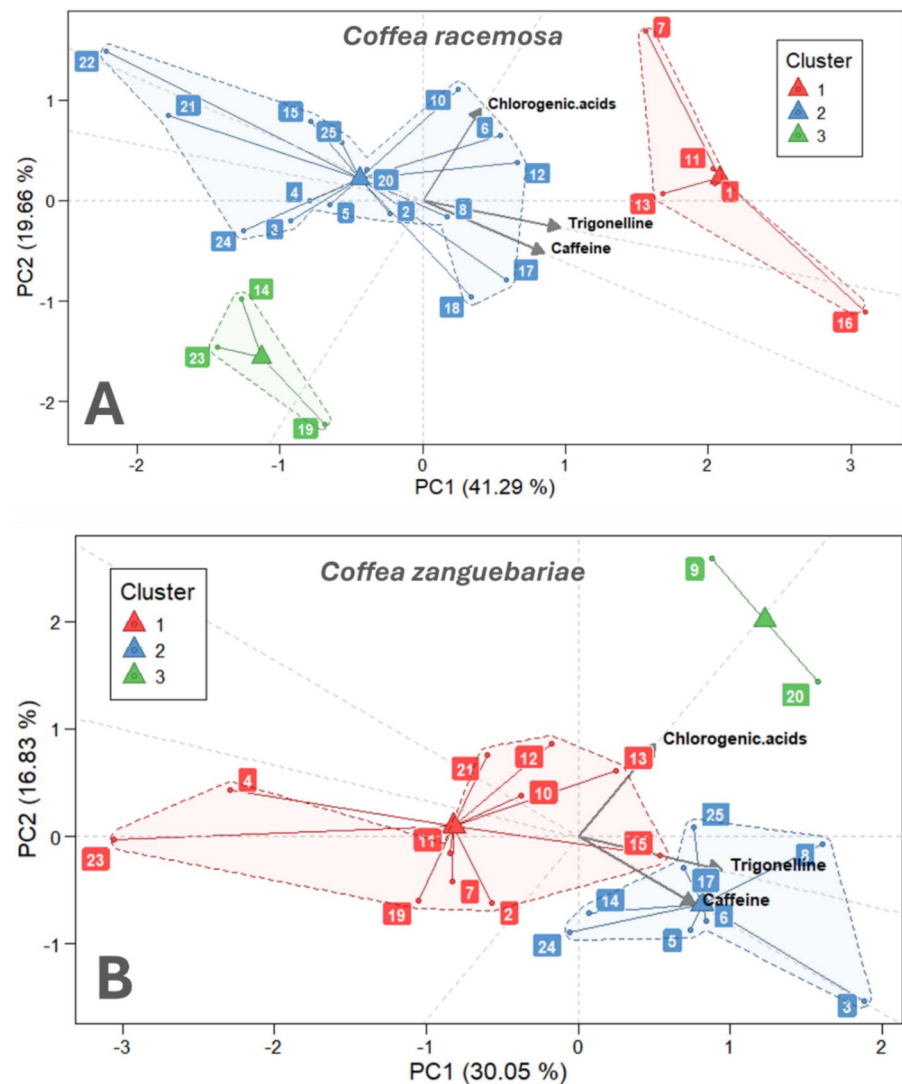
The contribution of variables to cluster separation also differed between the species. In *C. racemosa*, caffeine and trigonelline were the compounds that most explained the variability in component PC1, while CGAs contributed mainly to component PC2. In *C. zanguebariae*, caffeine and trigonelline also had the greatest influence on PC1, albeit with a lower explained variance, and CGAs exerted influence on both PC1 and PC2.

Best linear unbiased prediction

To complement, Best Linear Unbiased Prediction (BLUP) values were estimated for caffeine, trigonelline, and chlorogenic acid contents in the different genotypes of both species (Fig. 4).

The *C. racemosa* samples showed a greater range in the predicted values for chlorogenic acid contents, with genotypes varying from -2.00 to 1.94 suggesting greater putative genetic differentiation

Fig. 3 Principal Component Analysis (PCA) for genotype discrimination of *Coffea racemosa* (A) and *C. zanguebariae* (B) based on the composition of caffeine, trigonelline, and chlorogenic acids (CGAs)



among accessions for this compound under the sampled conditions for this compound in the species (Fig. 4A). Caffeine and trigonelline contents showed more modest variations, yet sufficient for identifying distinct, with BLUP values ranging approximately from -0.5 to 0.5 for caffeine and from -0.5 to 1.36 for trigonelline.

In *C. zanguebariae* the BLUP values for the compounds showed a narrower range compared to *C. racemosa*, especially for caffeine and trigonelline, whose values ranged approximately from -0.5 to 0.5 (Fig. 4B). CGAs also exhibited lower variability, with values distributed between -1.0

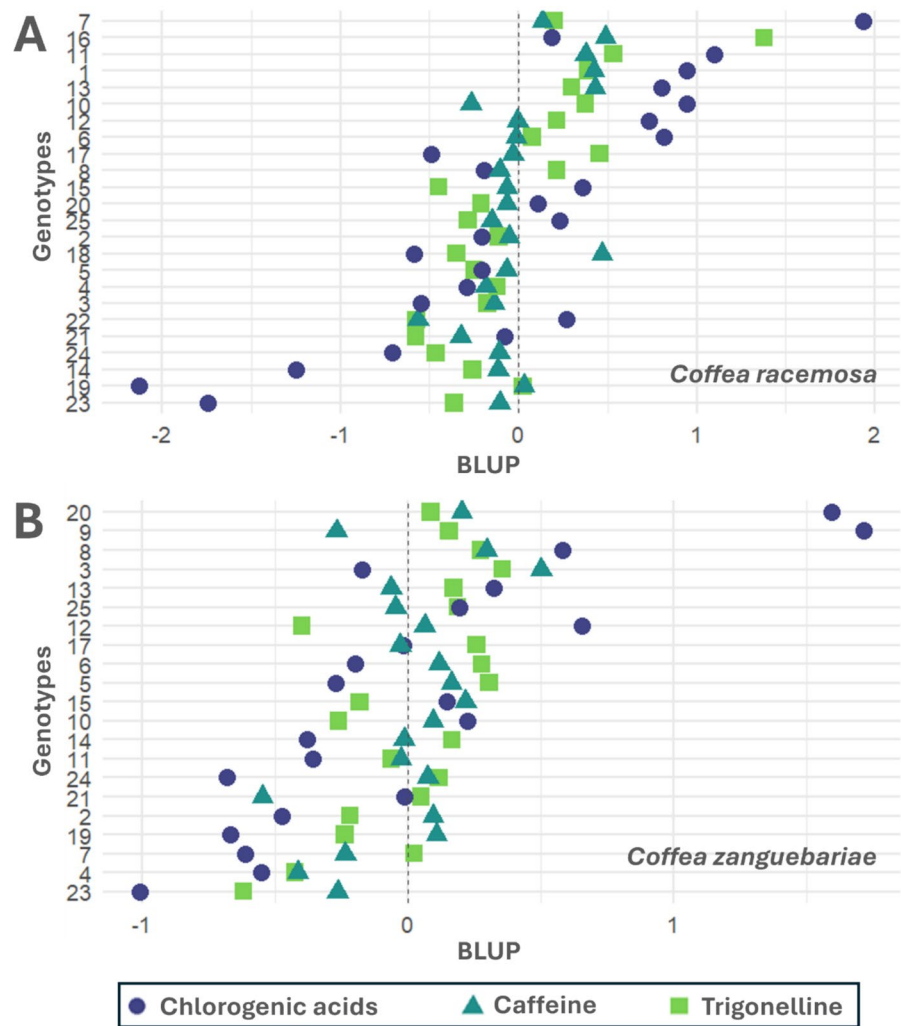
and 1.7 . Despite this, it was possible to identify genotypes with distinct performance for each compound, indicating relative performance differences among accessions within the environmental contexts evaluated.

Discussion

Quantification of bioactive compounds

The quantification of bioactive compounds in accessions of *Coffea racemosa* and *C. zanguebariae*

Fig. 4 Best linear unbiased prediction (BLUP) estimates for caffeine, trigonelline, and chlorogenic acid contents in different accessions of *Coffea racemosa* (A) and *C. zanguebariae* (B)



revealed contrasting chemical profiles between the species. In *C. racemosa*, the average content of the non-phenolic caffeine ($0.73 \text{ g } 100 \text{ g}^{-1}$) was below the values usually found in *C. arabica*, (usually in the range of 8 to about $16 \text{ mg g}^{-1} \text{ DW}$) and even lower than in *C. canephora* that frequently shows values between 16 and 22 or even $28 \text{ mg g}^{-1} \text{ DW}$ (Bicho et al. 2013; Oteef et al. 2024), although in green coffee the levels of caffeine might vary widely, mostly due to inter-specific differences, as pointed by several authors (Bicho et al. 2013). This low concentration suggests that *C. racemosa* may represent a promising genetic resource for the development of naturally low-caffeine coffee products and extracts, responding to the growing demand for beverages with reduced stimulant content, without the need for

chemical decaffeination (Hamon et al. 2025). However, it should be noted that coffee quality and sensory characteristics depend on the interaction of multiple chemical classes, including volatile compounds, sugars, lipids, and roasting-derived products.

Conversely, the chlorogenic acid content observed in *C. racemosa* ($24.4 \text{ mg g}^{-1} \text{ DW}$) was higher than in *C. zanguebariae*, indicating potential as a source of antioxidant-related compounds, although these values remain lower than those typically reported for total CGAs in cultivated species such as *C. arabica* (Bicho et al. 2013). The influence of CGAs on beverage quality is complex and depends on their interaction with other chemical constituents and processing conditions.

Meanwhile, *C. zanguebariae* exhibited an average caffeine content of 21.9 mg g⁻¹ DW, thus falling within the range usually found for *C. canephora* cultivars, whereas the CGA content (CGAs 13.6 mg g⁻¹ DW) is even lower to that found in *C. racemosa*, denoting a suitability for differentiated uses, particularly in formulations where the stimulant contribution associated with caffeine is valued.

Regarding trigonelline content, a known precursor of volatile compounds that contribute to the aroma and taste after roasting (Malta and Chagas 2009; Barbosa et al. 2012), the close values of 6.9 mg g⁻¹ and (*C. racemosa*) 8.6 mg g⁻¹ (*Coffea zanguebariae*) are somewhat below those previously observed in other *C. arabica* and *C. canephora* cultivars (Bicho et al. 2013), although this compound can be found in a wider range of values between 6–13.8 mg g⁻¹ (Ramalho et al. 2018), and values close to 6 mg g⁻¹ are common in coffee blends (Ribeiro et al. 2014). In fact, *Coffea zanguebariae* (sometimes mislead classified as *Coffea ibo* or even *C. racemosa*) from Cabo Delgado, Northeast Mozambique (as were our samples), have been long time noticed for its bean aromatic quality, as a new species for use in beverage (Vasconcellos 1906).

In the PCA, the formation of three well-defined clusters for both species demonstrates not only genotypic variability in caffeine, trigonelline, and chlorogenic acid contents, but also distinct patterns of association among these compounds, which may have relevant eco-physiological, chemical, and sensory implications. In *C. racemosa*, the greater variance explained by the first principal component (PC1) is strongly associated with the concentrations of caffeine and trigonelline, consistent with the metabolic dynamics of these alkaloids in *Coffea* spp., whose biosynthesis shares precursors in the shikimic acid pathway and is subject to genetic and environmental control (Ashihara and Crozier 2001). The segregation of Cluster 1, characterized by higher levels of these compounds in certain genotypes, is particularly relevant given that *C. racemosa* is generally recognized for its low caffeine content compared to the cultivated species *C. arabica* and *C. canephora* (Lyrio et al. 2025). The identification of accessions naturally richer in caffeine and trigonelline within this species suggests the existence of useful genetic variability, especially for breeding programs aimed at modulating chemical profiles for sensory or bioactive attributes.

Cluster 2, which encompassed the majority of the analyzed accessions, exhibited a profile characterized by higher levels of CGAs. These compounds play important ecological roles in plants, functioning as antioxidants, photoprotective agents, and mediators of defense against herbivores and pathogens (Clifford 1999). In *Coffea* species generally, variability in the contents of these compounds is associated with both genetic and environmental factors, with direct implications for the sensory profile of the beverage and its antioxidant capacity (Cheng et al. 2016). Considering that the individuals were under natural conditions and subjected to greater environmental challenges, the predominance of CGAs may support the hypothesis that *C. racemosa* tends to accumulate relatively higher proportions of these phenolics compared to alkaloids, a pattern previously observed in other wild African species (Anthony et al. 1993).

In *C. zanguebariae*, a distinct cluster configuration was observed, with lower variance explained by the first two principal components, indicating a more dispersed chemical structure. Despite the similarity to *C. racemosa* regarding the influence of caffeine and trigonelline on PC1, the formation of an exclusive cluster (Cluster 3) associated with accessions exhibiting elevated chlorogenic acid concentrations is of particular interest. This pattern, not observed in *C. racemosa*, may suggest specific genetic or biochemical mechanisms for the accumulation of these compounds in certain *C. zanguebariae* genotypes. Research in *C. canephora*, a species known for its high chlorogenic acid content, indicates that differential expression of genes in the phenylpropanoid pathway is strongly associated with intraspecific variability of these compounds (Francisco et al. 2024).

Best linear unbiased prediction

The analysis of predicted values (BLUPs) for bioactive compounds in *Coffea racemosa* and *C. zanguebariae* highlights genetic variability and improvement potential, particularly in *C. racemosa*. The wide range of BLUP values for CGAs suggests substantial genetic variability, which is important for selection efforts. This variability is supported by findings that genotype significantly influences the content of bioactive compounds such as CGAs, which are key

determinants of antioxidant capacity in coffee species (Rodrigues et al. 2015).

In contrast, *C. zanguebariae* exhibited a narrower range of BLUP values, indicating more restricted genetic variability for CGAs. This limited variability may constrain the genetic improvement potential of this species. The importance of CGAs as a selection criterion in coffee breeding is underscored by their significant quantitative differences across various genotypes, including *C. canephora* and *C. arabica*, as demonstrated through systematic research employing HPLC and multivariate analysis (Guerrero et al. 2001).

The lower variability observed in caffeine and trigonelline levels in *C. racemosa* and *C. zanguebariae* compared to CGAs may be attributed to biochemical and genetic factors. Caffeine and trigonelline, both methylxanthines, are synthesized through specific metabolic pathways that may exhibit tighter regulatory control in wild species, potentially due to evolutionary adaptations for survival in their native environments (Cardozo et al. 2010; Spinoso-Castillo et al. 2021). In contrast, CGAs, which perform various ecological functions (Kundu and Vadassery 2019), may display greater variability as they are influenced by environmental factors and genetic diversity, as observed in cultivated species such as *C. arabica* and *C. canephora* (Anthony et al. 1993; Ky et al. 2001).

This restriction in the BLUP range for caffeine and trigonelline may hinder genetic improvement strategies aimed at enhancing these alkaloids, as it limits the selection of diverse genetic material for breeding programs. Furthermore, the phenotypic stability of these compounds in wild species may reflect adaptive traits that confer resilience against environmental stresses, suggesting a complex interaction among genetic, biochemical, and ecological factors in the evolution of coffee (Francisco et al. 2024).

A limitation of the present study is that the accessions were collected from distinct natural habitats rather than evaluated under a common garden experimental design. Consequently, it was not possible to fully disentangle genetic effects from environmental influences or to estimate genotype $\times$ environment interactions. Although the mixed model approach minimizes part of the unexplained variance by treating genotype as a random

effect, the predicted values should be interpreted cautiously, as they may still partially reflect environmental heterogeneity inherent to in situ sampling. Therefore, the BLUP estimates reported here should be regarded as preliminary indicators of genetic differentiation. Future research should validate these findings through multi-environment trials or common garden experiments to confirm the stability and magnitude of genotypic effects.

Conclusion

The characterization of caffeine, trigonelline, and chlorogenic acid contents in *Coffea racemosa* and *C. zanguebariae* revealed high intraspecific variability, highlighting the potential of these species for the improvement of bioactive profiles.

Coffea racemosa exhibited a substantially lower average caffeine content compared to *C. arabica* and *C. canephora*, reinforcing its applicability for naturally low-caffeine coffees and functional products.

Coffea zanguebariae showed an average caffeine content similar to levels recorded in *C. canephora*, with intermediate chlorogenic acid contents, indicating potential for products where stimulant contribution is valued.

BLUP estimates suggested greater relative differentiation among *C. racemosa* accessions for CGAs under the sampled environmental conditions, indicating promising variability that warrants further validation under controlled experimental trials.

Future research should expand the characterization to minor compounds, integrate genomic analyses, and assess the chemical and functional stability of genotypes across different environments.

However, confirmation of these patterns under common garden or multi-environment trials is necessary to fully validate their genetic basis and breeding applicability.

Author contribution W.S.G. wrote the main manuscript text, performed data analysis, and interpreted the results. N.J.A. conducted the field collections and contributed to sample acquisition. C.A.F., J.S.N., G.L.M., and E.C.S.O. carried out the chemical laboratory analyses. T.R.M. assisted in data analysis and interpretation. A.P.M. and J.C.R. contributed to critical revision and improvement of the manuscript. F.L.P. supervised the study, contributed to the research design, and provided overall guidance. All authors reviewed and approved the final manuscript.

Funding The Article Processing Charge (APC) for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614). This work received funding support from the Espírito Santo Research and Innovation Support Foundation—FAPES (Proc. 2022-WTZQP and Proc. 2024-9H43M for FLP) and the National Council for Scientific and Technological Development—CNPq (Proc. 309535/2021-2 for FLP), United Nations Industrial Development Organization (UNIDO), OIKOS Institute, Lúrio University, Association of Coffee Producers of Ibo—Cabo Delgado, Coffee Producers of Inhambane and Gaza, Institute of Agricultural Research of Mozambique, the Agrarian Institute of Panda—Inhambane. Support by Fundação para a Ciência e a Tecnologia, I.P. (FCT), Portugal, through the projects UID/00239/2025 (<https://doi.org/10.54499/UID/00239/2025>) and UID/PRR/00239/2025 (<https://doi.org/10.54499/UID/PRR/00239/2025>), both of the Forest Research Centre), and the LA/P/0092/2020 (<https://doi.org/10.54499/LA/P/0092/2020>) of the Associate Laboratory TERRA, to J.C.R. is also greatly acknowledged.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interest.

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/>.

References

- Anthony F, Clifford MN, Noirot M (1993) Biochemical diversity in the genus *Coffea* L.: CGAs, caffeine and mozambioside contents. *Genet Resour Crop Evol* 40(2):61–70. <https://doi.org/10.1007/BF00052636>
- Ashihara H, Crozier A (2001) Caffeine: a well known but little mentioned compound in plant science. *Trends Plant Sci* 6(9):407–413. [https://doi.org/10.1016/S1360-1385\(01\)02055-6](https://doi.org/10.1016/S1360-1385(01)02055-6)
- Barbosa JN, Borém FM, Cirillo MA, Malta MR, Alvarenga AA, Alves HMR (2012) Coffee quality and its interactions with environmental factors in Minas Gerais, Brazil. *J Agric Sci* 4:181–190. <https://doi.org/10.5539/jas.v4n5p181>
- Bicho NC, Leitão AE, Ramalho JC, Alvarenga NB, Lidon FC (2013) Identification of chemical clusters discriminators of arabica and robusta green coffee. *Int J Food Prop* 16:895–904. <https://doi.org/10.1080/10942912.2011.573114>
- Campa C, Doulbeau S, Dussert S, Hamon S, Noirot M (2005) Qualitative relationship between caffeine and chlorogenic acid contents among wild species. *Food Chem* 93(1):135–139. <https://doi.org/10.1016/j.foodchem.2004.10.015>
- Cardozo EL Jr, Donaduzzi CM, Ferrarese-Filho O, Friedrich JC, Gonela A, Sturion JA (2010) Quantitative genetic analysis of methylxanthines and phenolic compounds in mate progenies. *Pesq Agrop Brasileira* 45(2):171–177. <https://doi.org/10.1590/S0100-204X2010000200008>
- Cheng B, Furtado A, Smyth HE, Henry RJ (2016) Influence of genotype and environment on coffee quality. *Trends Food Sci Technol* 57:20–30. <https://doi.org/10.1016/j.tifs.2016.09.003>
- Clifford MN (1999) CGAs and other cinnamates—nature, occurrence and dietary burden. *J Sci Food Agric* 79(3):362–372. [https://doi.org/10.1002/\(SICI\)1097-0010\(19990301\)79:3<362::AID-JSFA256%3e3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0010(19990301)79:3<362::AID-JSFA256%3e3.0.CO;2-D)
- Davis AP, Gargiulo R, Almeida INdasm, Caravela MI, Denison C, Moat J (2021) Hot coffee: the identity, climate profiles, agronomy, and beverage characteristics of *Coffea racemosa* and *C. zanguebariae*. *Front Sustain Food Syst*. <https://doi.org/10.3389/fsufs.2021.740137>
- Davis Helen, Chadburn Justin, Moat Robert, O'Sullivan Serene, Hargreaves Eimear, Nic Lughadha (2019) High extinction risk for wild coffee species and implications for coffee sector sustainability *Science Advances* 5(1) <https://doi.org/10.1126/sciadv.aav3473>
- Francisco JS, de Almeida IF, Portela CdaS, Alves EA, Rocha RB, Mori ALB, Benassi MdeT (2024) Teores de trigonelina, cafeína e ácidos clorogênicos em cafês *Coffea canephora* híbridos intervarietais naturais cultivados na Amazonia Ocidental. *Observatório Economia Latinoam* 22(7):e6020. <https://doi.org/10.55905/oelv22n7-295>
- Guerreiro-Filho O (1992) *Coffea racemosa* Lour. Une revue. *Café Cacao Thé* 36:171–186
- Guerrero G, Suárez M, Moreno G (2001) CGAs as a potential criterion in coffee genotype selections. *J Agric Food Chem* 49(5):2454–2458. <https://doi.org/10.1021/jf001286u>
- Hamon P, Raharimala NE, Akaffou S, Couturon E, Crouzillat D, Hamon S, Dupeyron M, Guyot R (2025) Caffeine-free species in the genus *Coffea*. In: *Coffee in health and disease prevention*, Elsevier, pp 59–67. <https://doi.org/10.1016/B978-0-443-13868-3.00004-1>
- Hameed Syed Ammar, Hussain Muhammad Umair, Ijaz Samee, Ullah Imran, Pasha Hafiz Ansar Rasul, Suleria (2018) Farm to Consumer: Factors Affecting the Organoleptic Characteristics of Coffee. II: Postharvest Processing Factors Abstract Comprehensive Reviews in Food Science and Food Safety 17(5):1184–1237. <https://doi.org/10.1111/1541-4337.12365>
- H. P., Piepho J., Möhring A. E., Melchinger A., Büchse (2008) BLUP for phenotypic selection in plant breeding and

- variety testing *Euphytica* 161(1-2):209–228. <https://doi.org/10.1007/s10681-007-9449-8>
- Kundu A, Vadassery J (2019) Chlorogenic acid-mediated chemical defence of plants against insect herbivores. *Plant Biol* 21(2):185–189. <https://doi.org/10.1111/plb.12947>
- Ky C-L, Louarn J, Dussert S, Guyot B, Hamon S, Noirot M (2001) Caffeine, trigonelline, CGAs and sucrose diversity in wild *Coffea arabica* L. and *C. canephora* P. accessions. *Food Chem* 75(2):223–230. [https://doi.org/10.1016/S0308-8146\(01\)00204-7](https://doi.org/10.1016/S0308-8146(01)00204-7)
- Lyrio MVV, Alberto NJ, Debona DG, Frinhani RQ, Ramalho JC, Pereira LL, de Castro EVR, Partelli FL, Romão W (2025) Comprehensive chemical profiling of wild *Coffea racemosa*, *C. zanguebariae*, *C. arabica*, and *C. canephora*: a metabolomic approach using LC-MSn and multivariate analysis. *Food Chem* 481:144062. <https://doi.org/10.1016/j.foodchem.2025.144062>
- Malta MR, Chagas SJR (2009) Evaluation of non-volatile compounds in different cultivars of coffee cultivated in southern Minas Gerais. *Acta Sci-Agron* 31:57–61
- Navarini L, Scaglione D, Del Terra L, Scalabrin S, Mavuke L, Turello L, Nguenha R, Luongo G (2024) Mozambican *Coffea* accessions from Ibo and Quirimba Islands: identification and geographical distribution. *AoB Plants* 16(2):plae004. <https://doi.org/10.1093/aobpla/plae004>
- Oteef MDY, Khemira H, Tounekti T, Mahdhi M, AlFaifi Z, Sharma M, Afzal M, AlSolami W (2024) Biochemical composition of the beans of local coffee cultivars from southwestern Saudi Arabia. *J Food Compos Anal* 129:106089. <https://doi.org/10.1016/j.jfca.2024.106089>
- Ramalho JC, Pais IP, Leitão AE, Guerra M, Reboredo FH, Máguas CM, Carvalho ML, Scotti-Campos P, Ribeiro-Barros AI, Lidon FJC, DaMatta FM (2018) Can elevated air [CO₂] conditions mitigate the predicted warming impact on the quality of coffee bean? *Front Plant Sci* 9:287. <https://doi.org/10.3389/fpls.2018.00287>
- Ribeiro VS, Leitão AE, Ramalho JC, Lidon FC (2014) Chemical characterization and antioxidant properties of a new coffee blend with cocoa, coffee silverskin and green coffee minimally processed. *Food Res Int* 61:39–47. <https://doi.org/10.1016/j.foodres.2014.05.003>
- Rodrigues NP, Salva TdeJG, Bragagnolo N (2015) Influence of coffee genotype on bioactive compounds and the in vitro capacity to scavenge reactive oxygen and nitrogen species. *J Agric Food Chem* 63(19):4815–4826. <https://doi.org/10.1021/acs.jafc.5b00530>
- Spinosa-Castillo JL, Escamilla-Prado E, Aguilar-Rincón VH, Corona-Torres T, García-de los Santos G, Morales-Ramos V (2021) Quantitative comparison of three main metabolites in leaves of *Coffea* accessions by UPLC-MS/MS. *Eur Food Res Technol* 247(2):375–384. <https://doi.org/10.1007/s00217-020-03631-8>
- Tapaca IdaPE, Mavuke L, Corti R, Pedrazzani S, Maquia ISA, Tongai C, Partelli FL, Ramalho JC, Marques I, Ribeiro-Barros AI (2023) Genomic evaluation of *Coffea arabica* and its wild relative *Coffea racemosa* in Mozambique: Settling resilience keys for the coffee crop in the context of climate change. *Plants* 12(10):2044. <https://doi.org/10.3390/plants12102044>
- Team, R. C. (2024) R internals.
- Vasconcellos E (1906) Catálogo da Exposição Colonial—Algodão, Borracha, Cacau e Café. In: Sociedade de Geographia de Lisboa, pp 138

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