Genetic Structure and Diversity of Rice From Certified Cultivars and Farmer’s Fields in Burkina Faso

Mariam BARRO
INERA: Institut de l’Environnement et de Recherches Agricoles

Abdourasmane KONATE
INERA: Institut de l’Environnement et de Recherches Agricoles

Issa WONNI
INERA: Institut de l’Environnement et de Recherches Agricoles

Abalo Itolou KASSANKOGNO
INERA: Institut de l’Environnement et de Recherches Agricoles

François SABOT
IRD: Institut de recherche pour le developpement

Laurence ALBAR
IRD: Institut de recherche pour le developpement

Irénee SOMDA
University Nazi Boni

Gilles BENA
IRD: Institut de recherche pour le developpement

Alain GHESQUIERE
IRD: Institut de recherche pour le developpement

Honoré KAM
INERA: Institut de l’Environnement et de Recherches Agricoles

Moussa SIE
INERA: Institut de l’Environnement et de Recherches Agricoles

Philippe CUBRY
IRD: Institut de recherche pour le developpement

Charlotte TOLLENAERE ( charlotte.tollenaire@ird.fr )
IRD: Institut de recherche pour le developpement  https://orcid.org/0000-0003-2045-2824

Short communication

Keywords: Rice, Genetic diversity, SNP, Burkina Faso

DOI: https://doi.org/10.21203/rs.3.rs-823882/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: In West Africa, two rice species are cultivated, the African rice *Oryza glaberrima* and the Asian rice *Oryza sativa*, the second one being much more frequent. Despite its importance for food security in West Africa, the genetic diversity of *Oryza sativa* in farmer's fields remains understudied in this region. Documenting the dynamics of diversity in the fields between landraces and improved cultivars is of importance to ensure rice cultivation adaptation to global change.

In this study, we genotyped 77 rice samples from Burkina Faso using the C6AIR SNP array from IRRI. Among the studied samples, 27 were certified cultivars and 50 were sampled in rice fields from three geographical zones in western Burkina Faso, each zone comprising one irrigated area and a neighboring rainfed lowland. Obtained data were analyzed in the frame of the worldwide rice genetic diversity by using the 3K genomes as reference.

Results: Most of the certified cultivars clustered with the *indica* genetic group, with a minority (26%) assigned to the *japonica* group. All except one of the rice samples from farmer's fields belonged to the *indica* group. The peculiar one unexpectedly clustered with the *Aus* genetic group and originated from Tengrela (rainfed lowland in Karfiguela zone). This site, already known to differ in terms of agronomic practices, showed the highest genetic diversity compared to the five other sites, as well as highest genetic differentiation. Obtained genetic data confirmed the high global frequency of one cultivar, in accordance with the data from farmer's interview. However, at the field level, genetic data rarely corresponded to the cultivar's names obtained from farmer's interviews.

Conclusions: Overall we found a high genetic diversity in the studied samples from Burkina Faso (certified and field's samples). We argue on the importance to document and preserve this agro-biodiversity as a prerequisite to face the current challenges of growing rice demand and global change. To this purpose, are required further scientific studies to refine our understanding of the dynamics of diversity in farmers’ fields, as well as a better knowledge of rice agro-biodiversity and awareness of its importance by rice farmers themselves.

Findings

Crop genetic diversity is a component of agro-biodiversity, with high value for nutrition and adaptation to biotic and abiotic stresses (Zimmerer et al. 2019), particularly in the context of global changes (Pironon et al. 2019). It contributes to render farming systems more stable, robust, and sustainable. On the other hand, the development, dissemination and adoption of improved cultivars is a pathway to increase crop productivity and align on market demands. A deep knowledge of certified cultivars at genomic level, as well as crop genetic diversity actually grown locally, are important information to take into account for plant diversity management and crop improvement.

Rice is rapidly becoming a staple food in the African diet. In West Africa, average annual production is 10.098.106 tons (milled equivalent, 2009–2019 period, Soullier et al. 2020), and average annual growth of production (2009–2019) is higher than 10% (Soullier et al. 2020). However, it still does not meet the demand, the imports representing 46% of rice consumption on average annually over the 2009–2019 period (Soullier et al. 2020), so that the efforts to increase rice production in a sustainable manner have to be pursued.

Two rice species are cultivated in West Africa: the African rice *Oryza glaberrima* Steud., and the Asian rice *Oryza sativa* L. Asian rice is nowadays by far the more cultivated in West African farming fields, with *O. glaberrima* being restricted to small areas, grown in places that are unsuitable for *O. sativa*, maintained mainly for social reasons (Linares 2002). African rice is known for its resistance to biotic and abiotic stresses and was consequently included to *O. sativa* genetic improvement programs, led by *AfricaRice* (*Africa Rice Center*) at the continental scale, to create
improved cultivars named NERICA (*New Rice for Africa*), and some of the cultivars named ARICA (*Advanced Rice for Africa*) (Sarla and Swamy 2005; Ndjiondjop et al. 2018a).

A small number of studies aimed at describing rice genetics in West Africa, some of them being limited in terms of number of genetic markers (< 100), (Chen et al. 2017; Luther et al. 2017; Yelome et al. 2018). However, Ndjiondjop et al. (2018a) used a panel of 27,560 SNPs to characterize 330 widely used rice genotypes in Africa, including NERICA and ARICA cultivars, and found two highly differentiated genetic groups corresponding to the different rice growing systems: lowland (primarily *indica* and lowland NERICAs) and upland (*japonica* and upland NERICAs). In addition, Diop et al. (2020) analyzed a set of 280 samples (landraces from Casamance/Senegal and other cultivars from West Africa in general) by genotyping-by-sequencing (GBS) approach. Analyzed samples belonged to four genetic subpopulations: *O. glaberrima*, *O. sativa japonica*, *O. sativa indica* Group 1, and *O. sativa indica* Group 2, with high genetic diversity within the majority group *O. sativa indica*.

Focusing on Burkina Faso in particular, research projects aiming to describe the genetic diversity from rice landraces began with a prospection by Moussa Sié in 1983–1984 all over the country (Sié et al. 1998; Sié et al. 1999). Then, a new sampling was performed in 2008 by Honoré Kam in 59 villages from four regions in western Burkina Faso, including 47 *O. glaberrima* samples among the 330 collected accessions, with a characterization of agromorphological diversity (Kam et al. 2017a) and genetic diversity using 23 microsatellites markers (Kam et al. 2017b). As expected, considering that upland ecology is not very common in Burkina Faso, only 11 accessions were assigned to *O. sativa japonica* genetic group; the *indica* group was consequently by far the most abundant and could be subdivided into three genetic groups (Kam et al. 2017b).

Considering the recent rise in areas cultivated in rice in Burkina Faso (three-fold between 2006 and 2016, FAO), and the intensification of rice cultivation (Demont et al. 2013), it is likely that the rice genetic diversity in the fields is rapidly evolving as well, rendering important to get a more recent picture of rice genetic diversity in this country.

Among the four rice production systems existing in West Africa (rainfed upland, rainfed lowland, irrigated and mangrove swamps, *AfricaRice* 2011; Katic et al. 2013), two are important in Burkina Faso. First, rainfed lowland rice represents the largest rice growing area in the country (67% between 1984 and 2009, MAHRH 2011). Second, irrigated rice represents less than 30% of harvest areas (FAO), but constitutes more than half of the total paddy rice produced (MAHRH 2011). Rice is grown in Burkina Faso over all 13 regions but most of the production comes from Hauts Bassins, Centre-Est, Boucle du Mouhoun and Cascades regions (DGPERS 2010; MAHRH 2011).

Because rice genomics is important for agrobiodiversity characterization in an objective of rice self-sufficiency (McCouch et al. 2013), it is timely to provide an update on the genetic characterization of rice cultivated in Burkina Faso. To this purpose, we selected two sets of samples for genotyping: cultivars certified in Burkina Faso or West Africa and samples from farmer's fields located in six study sites in western Burkina Faso, in order to address the following questions: (1) What is the genetic diversity of rice from Burkina Faso compared to worldwide rice diversity? (2) What is the genetic diversity and differentiation of rice actually grown by the smallholder rice farmers in irrigated and rainfed lowland sites in western Burkina Faso? (3) What is the correspondence between certified cultivars, genetic assignation of field samples and cultivar's names as known by the farmers?

We selected 27 certified cultivars from INERA Farako-Bâ, to represent the diversity of certified cultivars including some already adopted by the farmers in Burkina Faso, and recent elite varieties from African breeding programs. The information related to each of them appears in Table 1.
Table 1
List of registered cultivars from INERA/ Burkina Faso analyzed in this study.

| Cultivar   | Synonym    | Origin                | Introduction Date | Group           |
|------------|------------|-----------------------|-------------------|-----------------|
| FKR02      | GAMBIAKA   | Gambie                | 1970              | indica          |
| FKR04      | SINTANE DIOFOR | Senegal (Casamance) | 1960              | -               |
| FKR14      | 4418       | Inde                  | 1976              | indica          |
| FKR16      | 4456       | Inde                  | 1976              | indica          |
| FKR18      | SC 27      | Burkina Faso          | 1980              | indica          |
| FKR21      | ITA 257    | IITA/Nigéria          | 1987              | japonica        |
| FKR33      | 1195-5-2   | Burkina Faso          | 1982              | japonica        |
| FKR34      | RP 1125-1526-2 | Inde                | 1984              | indica          |
| FKR42      | IR64       | IRRI/Philippines      | 1989              | indica          |
| FKR45N     | WAB880-1-38 | AfricaRice            | 1999              | NERICA (upland) |
|            | WAB450-I-BL-1-736-HB | AfricaRice/INERA | -               | NERICA (upland) |
| FKR55      | WAB450-I-BL-1-736-HB | AfricaRice/INERA | -               | NERICA (upland) |
| FKR56N     | WSA 161-B-9-3 | AfricaRice            | 1999              | NERICA L (lowland) |
|            | WAB9984    | AfricaRice            | 2009              | japonica        |
| FKR60N     | WAS122-IDSA-1 | AfricaRice          | 1999              | NERICA L (lowland) |
| FKR61      | WAB C1 65  | AfricaRice            | 2009              | japonica        |
| FKR62N     | WAS 122-IDSA-1 | AfricaRice          | 1999              | NERICA L (lowland) |
| FKR64      | TS2        | Taiwan/DGPV/INERA     | -                | indica          |
| FKR66      | WAT1046-B43 | AfricaRice          | 2010              | indica          |
| FKR70      | IR 75-884-12-12 | AfricaRice       | 2009              | indica          |
| FKR74      | WAB 2094-WAC2-TGR2-B | AfricaRice/INERA | -               | ARICA (lowland) |
| FKR76      | F6-36      | DGPV/INERA            | 2012              | indica          |
| FKR78      | F6-41      | DGPV/INERA            | 2012              | indica          |

For each cultivar, are indicated potential synonym, the country or organism of origin, the date of introduction in Burkina Faso and the genetic group, based on a priori knowledge. Most of the cultivar's names begins with 'FKR' for 'Farako-Bâ Riz'.
| Cultivar | Synonym | Origin               | Introduction Date | Group                |
|----------|---------|----------------------|-------------------|----------------------|
| FKR80    | F6-49   | DGPV/INERA           | 2012              | indica               |
| FKR84    | ORYLUX 6 | AfricaRice/INERA     | -                 | indica               |
| NERICA4  | WAB450-I-B-P-91-HB | AfricaRice/INERA | -               | NERICA (upland)      |
| SAHEL177 | WAS 197-B-6-3-11 | AfricaRice/Saint-Louis | -   | indica               |
| SAHEL328 | WAS 197-B-4-1-5 | ISRA/AfricaRice      | -                 | indica               |

For each cultivar, are indicated potential synonym, the country or organism of origin, the date of introduction in Burkina Faso and the genetic group, based on a priori knowledge. Most of the cultivar's names begins with 'FKR' for 'Farako-Bâ Riz'.

In addition, we took advantage of a sampling previously performed in six sites in western Burkina Faso (Barro et al. 2021). These six sites are located in three geographical zones (Bama, Banzon and Karfiguela), each zone comprising one irrigated area and a neighboring rainfed lowland (Fig. 1a and https://dataverse.ird.fr/dataset.xhtml?persistentId=doi:10.23708/8FDWIE). The present study focused on the 50 fields studied in 2018 (7–11 fields per site, 8.3 on average, see Table 2). A map of the studied fields is presented in Fig. 1a. Each sample used in this study corresponded to one leaf per field, and was collected between September and December 2018. In 40 out of these 50 fields (80%), we interviewed the farmer and asked for the rice cultivars grown. We also performed farmer's interviews in the same sites in the two previous years (2016–2017) and subsequent year (2019) (Barro et al. 2021). A synthesis of obtained responses (annual frequencies of cultivars grown over the four-year period) are reported in Fig. 1b.

### Table 2

| Zone     | Site     | Rice production system | Latitude | Longitude | N samples | Gene Diversity |
|----------|----------|------------------------|----------|-----------|-----------|----------------|
| Bama     | Badala   | RL                     | 11.368   | -4.373    | 7         | 0.125 0.117    |
|          | Bama     | IR                     | 11.390   | -4.410    | 9         | 0.131           |
| Banzon   | Senzon   | RL                     | 11.288   | -4.829    | 7         | 0.113 0.124    |
|          | Banzon   | IR                     | 11.335   | -4.796    | 11        | 0.108           |
| Karfiguela | Tengrela | RL                     | 10.648   | -4.838    | 7         | 0.155 0.132    |
|          | Karfiguela | IR                    | 10.678   | -4.813    | 9         | 0.130           |
| Average  |          |                        |          |           |           | 0.131 0.124    |
| Global   |          |                        |          |           | 50        | 0.137           |

Wet lab work (both DNA extraction and SNP genotyping) was performed at the Genotyping Services Lab at the International Rice Research Institute (IRRI). DNA fingerprinting approach used the Illumina Infinium rice 6K chip (C6AIR) (Thomson 2014), a set of SNP designed to characterize the diversity within *Oryza sativa* species. This chip was already used in rice diversity studies, for example to characterize rice samples from Bangladesh (Kretzschmar et
Genotyping data table was provided by the genotyping platform and used for subsequent analysis (raw data available at https://github.com/Africrop/divrice_bf).

In order to place the rice diversity from Burkina Faso in the global context of Asian rice diversity, we downloaded the 29mio SNP datasets from the 3K genome data available at https://snp-seek.irri.org/. Obtained PLINK binary files were converted to VCF using PLINK software v1.9 (www.cog-genomics.org/plink/1.9/) (Chang et al. 2015), enabling the keep-allele-order option. SNPs positions corresponding to the C6AIR were then extracted from the VCF file using bcftools v1.9 (Danecek et al. 2021), then imported within the R software v4.1.0 (R Core Team 2018), as well as the genotyping table from accessions from Burkina Faso. Datasets from the chip genotyping and from the 3K genome data were merged prior to apply genomic filters. In order to keep best-quality SNPs, we applied the following filters SNPwise: less than 15% of missing data considering only the accessions from Burkina Faso, less than 10% missing data considering the whole dataset and an additional filter on heterozygosity which get rid of positions with more than 45% heterozygosity. We ended up with a final dataset including 5,247 SNPs.

We first made a Principal Component Analysis (PCA, Fig. 2) using LEA3.1 R package (Gain and François 2021), graphical display was made using ggplot2 R package v3.3.5 (Wickham 2016). Datasets were converted to the genind format from the adegenet R package v2.1.3 (Jombart 2008) and analyzed both with adegenet and hierfstat R package v0.5-7 (Goudet and Jombart 2020). A Discriminant Analyses on Principal Components (DAPC) was performed, using the 3K diversity groups as reference to assign accessions from Burkina Faso to these groups (Fig. S2). We then restrained our analysis on the accessions from this study (samples from Burkina Faso) and first computed a genetic tree within these samples (Fig. 3). Genetic distances between the accessions were computed using the dist.gene function and the resulting Neighbor-Joining tree was computed using the ape R package v5.5 (Paradis and Schliep 2019). Graphical representation was made using the “fan” option of the ggtree R package v3.1.2 (Yu 2020). A PCA was then computed considering only the field genotypes (with or without a peculiar accession, Fig. 3 and Fig. S1) with the dudi.pca function of ade4 R package v1.7-17 (Dray and Dufour 2007). Finally basic population genetics descriptive statistics (gene diversity and populations pairwise $F_{ST}$) considering different levels of hierarchy were computed using hierfstat (Table 2 and Table 3).

### Table 3

|               | Bama_RL | Bama_IR | Banzon_RL | Banzon_IR | Karfiguela_RL | Karfiguela_IR |
|---------------|---------|---------|-----------|-----------|---------------|---------------|
| Bama_RL       | NA      |         |           |           |               |               |
| Bama_IR       | 0.001848587 | NA      |           |           |               |               |
| Banzon_RL     | 0.012288212 | -0.019369126 | NA        |           |               |               |
| Banzon_IR     | 0.056308688 | -0.012856699 | -0.01936913 | NA       |               |               |
| Karfiguela_RL | 0.294933952 | 0.253058369 | 0.28337828 | 0.32875531 | 0.26147627   | NA            |
| Karfiguela_IR | 0.031039162 | 0.011808917 | 0.04078993 | 0.10737341 | 0.26147627   | NA            |

The PCA analysis including both this study’s samples and the 3K rice genomes (Fig. 2), as well as the DAPC analysis (Fig. S2), showed that the vast majority of the samples corresponds to *indica* group. This was expected as none of the samples came from upland growing systems, where japonica are generally found (Ndjiondjop et al. 2018a). Also,
Diop et al. (2020) found that *indica* was the most widely cultivated type of lowland rice in West Africa; with very few of their samples revealed as *japonica*.

On the other hand, seven reference cultivars (FKR45N, FKR59, FKR33, FKR21, FKR61, NERICA4, NERICA8, see Fig. 2, Fig. 3) were attributed to the *japonica* group. These cultivars were known to be japonica or Nerica upland (Table 1, Fig. 3) so that this result is congruent with expectations. The *japonica* group do not contain any of the analyzed samples from farmer's fields (Fig. 2, Fig. 3). The cultivar FKR04, a cultivar introduced from Casamance (Senegal) in 1960 belonged to admix (Table 1, Fig. 2, Fig. 3). Finally, one field sample, from the field labelled 'TG02', belong to *Aus* group (Fig. 2, Fig. 3). The *Aus* genetic group does not seem to be common in West Africa in general, as it does not appear in the two studies cited previously (Diop et al. 2020; Ndjiondjop et al. 2018a), whereas Sié et al. (1999) reported the presence of two *Aus* rice samples in a previous sampling performed in Burkina Faso.

Global gene diversity estimated in the 27 certified cultivars from Burkina Faso was 0.282. It represented various genetic groups (*indica, japonica* and admix, Fig. 2). Within-group diversity was also apparent as we noticed the certified cultivars from Burkina Faso are not so closed from each other's within *indica* and *japonica* diversity groups (Fig. 2). We argue however that the diversity used for the certified cultivars in Burkina Faso could still be enlarged by mobilizing more genetic diversity of the rice worldwide germplasm.

Global gene diversity estimated in the 50 analyzed field samples from Burkina Faso was 0.137, and within-site genetic diversity was the highest in Tengrela site (0.132; Table 2). On the other hand, Banzon irrigated site presented the lowest genetic diversity (0.108; Table 2). Tengrela was also involved in all highest pairwise genetic differentiation values (Table 3, Table S1). Highest between site genetic differentiation was between Tengrela rainfed lowland and Banzon irrigated perimeter ($F_{ST} = 0.328 [0.311–0.347]$). Such a specificity of this site of Tengrela, compared to others study sites, was also evidenced in a previous study on agricultural practices, showing that rice was mostly grown by women, for self-consumption only, with low frequency of chemical fertilization but often manure from household waste (Barro et al. 2021).

Tengrela was also the only site among the six where only landraces were grown (no use of certified cultivars, Fig. 1b, Barro et al. 2021) and where a sample was attributed to *Aus* group (Fig. 2, Fig. S1). The farmer named its cultivar ‘Samperema’. We note that the six other samples from Tengrela (whose names include ETP, Bedankaki, Bandakadi / Debale, Tchombiais) are also differentiated from the samples found in the five other sites (Fig. 3, and see the point aspect in Fig. 4 and Fig. S1). They could derive from hybridization between the locally grown landrace belonging to the *Aus* diversity group and introduced *indica* varieties, as suggested by their affinity in the PCA with TG02 (Fig. S1).

In terms of rice growing systems, we note that the samples from irrigated areas (in blue, Fig. 3) and rainfed lowlands (in orange, Fig. 3) do not specifically differ from each other, with the exception of Tengrela village, as previously mentioned (TG samples in Fig. 3). Rainfed lowlands considered as a whole (three sites) had a higher gene diversity (0.149) compared to irrigated areas (0.125). The genetic differentiation between all rainfed lowlands fields and all irrigated areas was estimated to $F_{ST} = 0.030 [0.027–0.034]$. This is likely due to the peculiarity of Tengrela site, as we note that the $F_{ST}$ obtained between rainfed lowlands and irrigated areas from the geographic zones of Banzon and Bama both do not differ from zero (Table 3 and Fig. S1).

In the phylogenetic tree (Fig. 3), as well as PCA analyses (Fig. 4 and Fig. S1), we also note that many field samples (from five sites: Senzon, Banzon, Karfiguela, Badala, Bama) are identical to each other's and to the reference cultivar FKR64 (commonly named *TS2* in Burkina Faso, see Table 1). This cultivar, originating from Taïwan (Table 1), was frequently mentioned by the farmers from all sites, except in the peculiar site of Tengrela (Fig. 1b). Consequently,
genetic data and farmer's responses were in agreement in that this rice cultivar is the most frequently grown in the studied sites from western Burkina Faso. However, at the field level, the genetic assignation of samples did not always correspond to the cultivar given names according to farmers’ interviews (see the color of labels in Fig. 4 and Fig. S1). This likely reflects the dynamics of rice genetic diversity in farmer's field and illustrates that genetic pool is not fixed but still evolving.

Our results offer a picture of rice genetic diversity in six sites from lowland and irrigated areas of western Burkina Faso. We confirmed that indica rice is by far the most frequently grown, but we identified a sample from the Aus genetic group. We globally did not notice major differences between lowland and irrigated areas, except for the lowland site of Tengrela, where rice is grown traditionally (low input) by women using traditional landraces. We confirmed the predominance of one certified cultivar in the five other sites.

Further research is required to encompass rice agrobiodiversity actually grown in this country and in West Africa in general. As a perspective of this study, we first propose to include more certified varieties and to extend the geographic areas to cover all important rice production areas from Burkina Faso. For instance, it would be interesting to study samples from other regions in Burkina Faso, such as the Boucle du Mouhoun, that was shown to be the most diversified in a previous study (Kam et al. 2017a). In addition, it could have been interesting to include rainfed upland rice fields, although this rice production system is minor in Burkina Faso (10% of the rice land area and 5% of national rice production, MAHRH 2011). Third, deciphering potential within-field rice genetic diversity is also an interesting research question for future work, that was not addressed in our study where only one plant sample from each field was analyzed (but see Gouda et al. 2020).

We also documented the genetic diversity of 27 certified cultivars, including indica, japonica and Nerica. This may offer the perspective (not straightforward though) to try to design easy-to-use genetic markers (see Ndjiondjop et al. 2018b) for markers discriminant for rice species) useful for quality control and seed certification. Finally, we argue for more studies combining rice genetic diversity with human and social science to understand further the rationale behind rice farmer's seed choice. This would allow understanding the apparent discrepancies between genetic assignations and naming of the cultivars in the fields, and more importantly getting useful information to design suitable strategies for rice genetic diversity management.

Indeed, while genetic improvement is very important to increase yield and fight poverty and food insecurity (Arouna et al. 2017), it is also critical to preserve agrobiodiversity and to include landraces, especially the preferred by farmers and consumers, in breeding programs and dissemination projects. This is especially pertinent in the context of global changes because farmer rice varieties in West Africa were shown to be robust / tolerant to sub-optimal conditions (Mokuwa et al. 2013).

**Abbreviations**

ARICA: Advanced Rice for Africa (https://www.africanrice.org/arica)

BL: Badala study site in Burkina Faso

BM: Bama study site in Burkina Faso

BZ: Banzon study site in Burkina Faso
Acknowledgements

This work was performed thanks to the facilities of the “International joint Laboratory LMI PathoBios: Observatory of plant pathogens in West Africa: biodiversity and biosafety” (www.pathobios.com; twitter.com/PathoBios). We are very grateful to Kader Abdoul Guigma, Yacouba Kone, Edouard Kabore, Daouda Hema, Traoré Momouni, Dabire Roméo, for their contributions to the fieldwork in Burkina Faso. We thank the rice farmers from Badala, Bama, Senzon, Banzon, Tengrela and Karfiguela for their kind collaboration. We are grateful to Martine Bangratz for testing some DNA extractions, and to Maria Ymber V. Reveche at IRRI Service Laboratories - Genotyping Services Lab (GSL) for her help in receiving the samples and genotyping. We thank Alexandre Soriano who inserted the data in the rice genome hub (https://rice-genome-hub.southgreen.fr/). We thank Eugénie Hebrard for helpful comments on earlier version of the manuscript.

Funding
This work was publicly funded by the CGIAR Research Program on Rice Agri-food Systems (RICE) and the ANR (the French National Research Agency) under «Investissements d’avenir» programme with the reference ANR-10-LABX-001-01 Labex Agro (RiPaBIOME project), coordinated by Agropolis Fondation under the frame of I-SITE MUSE (ANR-16-IDEX-006).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

Field data are available on IRD platform DataSuds: https://dataverse.ird.fr/dataset.xhtml?persistentId=doi:10.23708/8FDWIE

Genotyping data and R script used to analyze the data and perform the figures is available on github: https://github.com/Africrop/divrice_bf

Authors’ contributions

IW, AK, AG and CT coordinated the project, with advises from FS and LA. MB, AIK, IW and CT coordinated the sampling in farmer’s fields, and MB and AIK performed the sampling. IS and GB supervised the PhD work of MB, with the help of IW and CT. AK, IW, HK and MS chose and prepared the certified cultivars. PC, CT and MB analyzed the data. CT and PC wrote the manuscript and MB, IW, MS, GB, LA, FS edited previous versions of the manuscript. All authors reviewed the final manuscript.

Ethics approval and consent to participate

In every case, we obtained permission from the farmers to work in their fields, and the management of the entire project followed the guidelines of the Nagoya protocol regarding access and benefit sharing.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. AfricaRice ARC (2011) Boosting Africa’s Rice Sector: A research for development strategy 2011-2020. Cotonou, Benin

2. Arouna A, Lokossou JC, Wopereis MCS, Bruce-Oliver S, Roy-Macauley H (2017) Contribution of improved rice varieties to poverty reduction and food security in sub-Saharan Africa. Glob Food Sec 14:54-60. doi:https://doi.org/10.1016/j.gfs.2017.03.001
3. Barro M, Kassankogno AI, Wonni I, Sereme D, Somda I, Kaboré KH, Beno G, Brugidou C, Tharreau D, Tollenaere C (2021) Spatiotemporal Survey of Multiple Rice Diseases in Irrigated Areas Compared to Rainfed Lowlands in the Western Burkina Faso. Plant Dis. doi:10.1094/PDIS-03-21-0579-RE

4. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4 (1). doi:10.1186/s13742-015-0047-8

5. Chen C, He W, Nassirou TY, Nsabiyumva A, Dong X, Adedze YMN, Jin D (2017) Molecular characterization and genetic diversity of different genotypes of *Oryza sativa* and *Oryza glaberrima*. Electron J Biotechnol 30:48-57. doi:https://doi.org/10.1016/j.ejbt.2017.08.001

6. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. GigaScience 10 (2):giab008. doi:10.1093/gigascience/giab008

7. Demont M, Ndour M, Zossou E (2013) Can local African rice be competitive? An analysis of quality-based competitiveness through experimental auctions. Cah Agric 22 (5):345-352. doi:10.1684/agr.2013.0664

8. DGPER M (2010) Résultats définitifs de la campagne agricole et de la situation alimentaire et nutritionnelle 2009/2010. Burkina Faso

9. Diop B, Wang DR, Drame KN, Gracen V, Tongoona P, Dzidzienyo D, Naruye E, Greenberg AJ, Djiba S, Danquah EY, McCouch SR (2020) Bridging old and new: diversity and evaluation of high iron-associated stress response of rice cultivated in West Africa. J Exp Bot 71 (14):4188-4200. doi:10.1093/jxb/eraa182

10. Dray S, Dufour A-B (2007) The ade4 Package: Implementing the Duality Diagram for Ecologists. J Stat Softw; 1(4). doi:10.18637/jss.v022.i04

11. Food and Agriculture Organization of the United Nations, FAO. (1997). FAOSTAT statistical database http://www.fao.org/faostat/.

12. Gain C, François O (2021) LEA 3: Factor models in population genetics and ecological genomics with R. Mol Ecol Resour. doi:https://doi.org/10.1111/1755-0998.13366

13. Gouda AC, Ndjiondjop MN, Djedatin GL, Warburton ML, Gounoulou A, Kpeki SB, N'Diaye A, Semagn K (2020) Comparisons of sampling methods for assessing intra- and inter-accession genetic diversity in three rice species using genotyping by sequencing. Sci Rep 10 (1):13995. doi:10.1038/s41598-020-70842-0

14. Goudet J, Jombart T (2020) hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.5-7. https://CRAN.R-project.org/package=hierfstat.

15. Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24 (11):1403-1405. doi:10.1093/bioinformatics/btn129

16. Kam H, Ndjiondjop M, Sanou J, Sow M, Ouédraogo I, Laing M (2017a) Diversity assessment and geographical variability of rice landraces collected from Burkina Faso. Int J Dev Sustain 6:479-496

17. Kam H, Ndjiondjop M, Laing M, Ahmadi N (2017b) Molecular characterization and diversity analysis of Burkina Faso rice landraces using 23 microsatellite markers and establishment of a core collection. Int J Curr Res 9:56222-56232

18. Katic PG, Namara RE, Hope L, Owusu E, Fujii H (2013) Rice and irrigation in West Africa: Achieving food security with agricultural water management strategies. Water Resour Econ 1:75-92. doi:https://doi.org/10.1016/j.wre.2013.03.001

19. Kretzschmar T, Mbanjo EGN, Magalit GA, Dwiwanti MS, Habib MA, Diaz MG, Hernandez J, Huelgas Z, Malabayabas ML, Das SK, Yamano T (2018) DNA fingerprinting at farm level maps rice biodiversity across Bangladesh and reveals regional varietal preferences. Sci Rep 8 (1):14920. doi:10.1038/s41598-018-33080-z
20. Linares OF (2002) African rice (*Oryza glaberrima*): History and future potential. Proc Natl Acad Sci U.S.A. 99 (25):16360. doi:10.1073/pnas.252604599

21. Luther Z, Akromah R, Nyadanu D, Tokpah D, Page Z, Voor V, Kwaloe A (2017) Evaluation of genetic diversity in rice (*Oryza sativa* and *Oryza glaberrima*) germplasm from Liberia and Ghana using simple sequence repeat (SSR) markers. Afr J Biotechnol 16:1990-1996

22. MAHRH (2011) Burkina Faso _ National Rice Development Strategy (NRDS). https://riceforafrica.net/images/stories/PDF/burkina_faso_en.pdf

23. McCouch S, Wing R, Semon M, Venuprasad R, Atlin G, Sorrells M, Jannink J (2013) Making Rice Genomics Work for Africa. In: Wopereis M, David Johnson D, Ahmadi N, Tollens E, Jalloh A (eds) Realizing Africa's Rice Promise. CABI, pp 108-129

24. Mokuwa A, Nuijten E, Okry F, Teeken B, Maat H, Richards P, Struik PC (2013) Robustness and Strategies of Adaptation among Farmer Varieties of African Rice (*Oryza glaberrima*) and Asian Rice (*Oryza sativa*) across West Africa. PLoS One 8 (3):e34801. doi:10.1371/journal.pone.0034801

25. Ndjiondjop MN, Semagn K, Sow M, Manneh B, Gouda AC, Kpeki SB, Pegalepo E, Wambugu P, Sie M, Warburton ML (2018a) Assessment of Genetic Variation and Population Structure of Diverse Rice Genotypes Adapted to Lowland and Upland Ecologies in Africa Using SNPs. Front Plant Sci 9. doi:10.3389/fpls.2018.00446

26. Ndjiondjop MN, Semagn K, Zhang JW, Gouda AC, Kpeki SB, Goungoulou A, Wambugu P, Drame KN, Bimpong IK, Zhao DL (2018b) Development of species diagnostic SNP markers for quality control genotyping in four rice (*Oryza L.*) species. Mol Breed 38 (11). doi:10.1007/s11032-018-0885-z

27. Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35 (3):526-528. doi:10.1093/bioinformatics/bty633

28. Pironon S, Etherington TR, Borrell JS, Kühn N, Macias-Fauria M, Ondo I, Tovar C, Wilkin P, Willis KJ (2019) Potential adaptive strategies for 29 sub-Saharan crops under future climate change. Nat Clim Change 9 (10):758-763. doi:10.1038/s41558-018-0585-7

29. R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/.

30. Sarla N, Swamy BPM (2005) *Oryza glaberrima*: A source for the improvement of *Oryza sativa*. Curr Sci 89 (6):955-963

31. Sié M, Ghesquière A, Miezan K (1999) Structure génétique des variétés traditionnelles de riz (*Oryza sp.*) du Burkina Faso. Agron Afr 11 (2):57-71

32. Sié M, Zongo JD, Dona Dakouo D (1998) Prospection des cultivars traditionnels de riz du Burkina Faso. Revue CAMES Sci Med 00:21-27

33. Soullier G, Demont M, Arouna A, Lancon F, del Villar PM (2020) The state of rice value chain upgrading in West Africa. Glob Food Sec 25. doi:10.1016/j.gfs.2020.100365

34. Thomson MJ (2014) High-Throughput SNP Genotyping to Accelerate Crop Improvement. Plant Breed Biotechnol 2 (3):195-212. doi:10.9787/PBB.2014.2.3.195

35. Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Vol ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org . Springer-Verlag New York

36. Yelome OI, Audenaert K, Landschoot S, Dansi A, Vanhove W, Silue D, Van Damme P, Haesaert G (2018) Analysis of population structure and genetic diversity reveals gene flow and geographic patterns in cultivated rice (*O. sativa* and *O. glaberrima*) in West Africa. Euphytica 214 (11):215. doi:10.1007/s10681-018-2285-1
37. Yu G (2020) Using ggtree to Visualize Data on Tree-Like Structures. Current Protocols in Bioinformatics 69 (1):e96. doi:https://doi.org/10.1002/cpbi.96

38. Zimmerer KS, de Haan S, Jones AD, Creed-Kanashiro H, Tello M, Carrasco M, Meza K, Plasencia Amaya F, Cruz-Garcia GS, Tubbeh R, Jiménez Olivencia Y (2019) The biodiversity of food and agriculture (Agrobiodiversity) in the anthropocene: Research advances and conceptual framework. Anthropocene 25:100192. doi:https://doi.org/10.1016/j.ancene.2019.100192

**Figures**

*Figure 1*
Information on the studied sites: geographic location, rice growing systems and rice cultivars grown according to interviewed farmers 1a. Map of the rice sampling locations: Up right: location of Burkina Faso in Africa (from https://upload.wikimedia.org/wikipedia/commons/d/d0/Location_map_of_Burkina_Faso_in_Africa.svg). Up left: location of the three studied geographic zones in western Burkina Faso. The locations of all rice fields and irrigated areas from the Burkina Faso land occupation database (https://www.ignfi.fr/fr/portfolio-item/occupation-des-terres-burkina-fao/) appear in green on the map. Down: location of the studied fields within each of the three zone (Bama, Banzon and Karfiguela, satellite Google images). Irrigated locations (BZ: Banzon; KA: Karfiguela; BM: Bama) are represented with blue scares and surrounded in blue; while rainfed lowland (SZ: Senzon; TG: Tengrela, BL: Badala) places are represented similarly but in orange. 1b. Frequency of rice cultivars according to farmer’s interviews in the different study sites and the different years (2016-2019 dataset). Each plot corresponds to one site and each bar to a particular year and colors correspond to rice cultivars cited by the farmers.

Figure 2

Principal Component Analysis (PCA) based on SNP data for the rice reference data for worldwide diversity (3K genomes) and this study’s 77 samples from Burkina Faso. The reference samples represented as colored points according to the genetic group. The samples from Burkina Faso are shown in black, with certified cultivars represented with crosses and samples from farmer’s fields with squares. On the left are presented the first and second axis of the PCA while on the right are shown the first and third axis. Among the 77 analyzed samples, two (FKR04 and TG02) revealed very different and are labelled on the right-hand side of the figure.
Figure 3

Unrooted Neighbor-Joining tree based on distances computed from SNP data for the 77 samples from Burkina Faso. Samples from farmer’s fields begin with two letters reflecting the sampling site and appear in blue when sampled in irrigated areas (BZ: Banzon, KA: Karfiguela, BM: Bama) and in red/orange from rainfed lowlands (SZ: Senzon, TG: Tengrela, BL: Badala). On the other hand, registered cultivars appear in grey. Most of the samples (68/77=88.3%) are from indica genetic groups, surrounded in purple. The other genetic groups are indicated with colored shapes: blue for japonica, pink for Aus and grey for admix (see also Figure 2).
Figure 4

Principal Component Analysis (PCA) based on SNP data for the 49 samples from farmer’s fields in Burkina Faso assigned to the genetic group indica. The color of the labels indicate the cultivar as named by the farmer during the interview. The labels are the field’s identifier, with two letters reflecting the site: BL: Badala, BM: Bama, BZ: Banzon, KA: Karfiguela, SZ: Senzon TG: Tengrela. The aspect of the point indicates the sampling site: the shape corresponding the geographical zone and the color reflecting the rice growing system.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.docx
- FigureS1.tif
- FigureS2.tif