PERSPECTIVE

Plant genomics in Africa: present and prospects

Hassan Ghazal[1,2], Yagoub Adam[3] (iD), Abdellah Idrissi Azami[2], Sofia Sehli[2], Hannah N. Nyarko[4], Bouchra Chaouni[5], Grace Olasehinde[3,6], Itunuoluwa Isewon[3,7], Marion Adebiyi[3,8], Olayinka Ajani[3,9], Enock Matovu[10], Olawole Obembe[3,6], Yvonne Ajamma[3], Gaston Kuzamunu[11,12,13], Samson Pandam Salifu[4], Jonathan Kayondo[14], Alia Benkahla[15] and Ezekiel Adebiyi[3,7,16,*] (iD)

[1]National Center for Scientific and Technical Research, Rabat, Morocco,
[2]Mohammed VI University of Health Sciences, Casablanca, Morocco,
[3]Covenant University Bioinformatics Research (CUBRe), Covenant University, Ogun State, Km 10 Idiroko Road, P.M.B. 1023, Nigeria,
[4]Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana,
[5]Laboratory of Plant and Microbial Biotechnology, Biodiversity and Environment, Faculty of Sciences, University Mohammed V, Rabat, Morocco,
[6]Department of Biological Sciences, Covenant University, Ogun State, Km 10 Idiroko Road, P.M.B. 1023, Ota, Nigeria,
[7]Department of Computer and Information Sciences, Covenant University, Ogun State, Km 10 Idiroko Road, P.M.B. 1023, Ota, Nigeria,
[8]Department of Computer Science, Landmark University, Kwara-State, Omu-Aran, Nigeria,
[9]Department of Chemistry, Covenant University, Ogun State, Km 10 Idiroko Road, P.M.B. 1023, Ota, Nigeria,
[10]College of Veterinary Medicine, Animal Resources and Biosecurity, Makerere University, P. O. Box 7062, Kampala, Uganda,
[11]African Institute for Mathematical Sciences, Cape Town 7945, South Africa,
[12]Department of Pathology, Division of Human Genetics, University of Cape Town, IDM, Cape Town, South Africa,
[13]Department of Integrative Biomedical Sciences, Computational Biology Division, University of Cape Town, Observatory 7925, South Africa,
[14]Uganda Virus Research Institute (UVRI), Uganda Research Unit on AIDS, Entebbe, Uganda,
[15]Bioinformatics and Biostatistics Laboratory (LR16IPT09), Pasteur Institute of Tunis, Tunis, Tunisia, and
[16]Division of Applied Bioinformatics, German Cancer Research Center (DKFZ), G200, Im Neuenheimer Feld 280, Heidelberg 69120, Germany

SUMMARY

Plants are the world's most consumed goods. They are of high economic value and bring many health benefits. In most countries in Africa, the supply and quality of food will rise to meet the growing population's increasing demand. Genomics and other biotechnology tools offer the opportunity to improve subsistence crops and medicinal herbs in the continent. Significant advances have been made in plant genomics, which have enhanced our knowledge of the molecular processes underlying both plant quality and yield. The sequencing of complex genomes of African plant species, facilitated by the continuously evolving next-generation sequencing technologies and advanced bioinformatics approaches, has provided new opportunities for crop improvement. This review summarizes the achievements of genome sequencing projects of endemic African plants in the last two decades. We also present perspectives and challenges for future plant genomic studies that will accelerate important plant breeding programs for African communities. These challenges include a lack of basic facilities, a lack of sequencing and bioinformatics facilities, and a lack of skills to design genomics studies. However, it is imperative to state that African countries have become key players in the plant genome revolution and genome derived-biotechnology. Therefore, African governments should invest in public plant genomics research and applications, establish bioinformatics platforms and training programs, and stimulate university and industry partnerships to fully deploy plant genomics, particularly in the fields of agriculture and medicine.

INTRODUCTION

Currently, the world is facing several challenges, including public health and food security for an ever-increasing population of approximately 8 billion people worldwide, with over 600 million living in Africa (Conway, 2012; Philp, 2018). In Africa, agriculture is one of the most prominent economic sectors, and about 70% of the continent's land is used for agriculture-related activities (Nhemachena et al., 2018). Agriculture can contribute significantly to achieving the continent's developmental goals, which include eradication of poverty and hunger, enhancing trade and investment, sustainable management of resources and environment, employment, and human security and prosperity. Although the continent possesses about 60% of the world's arable land, Africa accounts for just 4% of global agricultural output (Nhemachena et al., 2018). Researchers blame the low productivity on erratic weather conditions, drought, poor seed quality, and obsolete farming practices.

Regarding public health, the incidence and prevalence of non-communicable and infectious diseases have increased. Most of these diseases lack approved and effective treatments, while others are still being treated by decades-old protocols with many limitations. It is interesting to note that Africa hosts a large variety of plants, with about 60% of them being indigenous, and most have been reported to be used for the treatment of diverse diseases (Tuasha et al., 2018). The use of medicinal plants as therapeutics is a common practice. Ethnobotanical surveys have reported some plant species employed in the treatment of many diseases such as malaria, typhoid, tuberculosis, rheumatism, and asthma (Lawal et al., 2020).

BOX 1 . Summary of the main points

1 Plant genomics and its applications
2 The progress in sequencing and analysis of African plant genomes
3 Application of bioinformatics tools to study plant genomes in Africa

---

BOX 2 . Open question

1 What is the current state of the art of crop and medicinal plant genomics in Africa?

Harnessing the full therapeutic potential of plants may contribute to bridging the gap between disease outcomes and therapy management, thereby improving health in the African population. Genome projects on medicinal plants could advance their use as a source of natural drugs coupled with the development and production of pharmaceutical agents. The search for bioactive compounds from medicinal plants is not a new research effort in Africa. However, most of the previous research works are either national- or regional-based. Furthermore, bioinformatics technology has not been applied extensively in any of the previous efforts. Therefore, most of the approaches have only been wet lab-based, yet investigating wet lab validations of obtained promising active compounds, the experimental methods applied have often led to high attrition rates in the final active compounds that are validated.

The field of genomics has opened up a whole new plethora of opportunities and avenues, which can help achieve enhanced productivity and preserve biodiversity under the present-day challenges. The recent advances in genomics offer great potential in crop improvement and herb biodiversity programs. To this end, various crop improvement programs have been initiated (Aglawe et al., 2018; Butt et al., 2020; Moose and Mumm, 2008; Parry et al., 2009). Over the past few decades, genomics-based approaches have been extensively used to understand and dissect the genetic basis of crop improvement. The H3Africa Health Genomics Consortium undertakes efforts to integrate genomics knowledge and expertise, which have been acquired over the last decade, into African conventional crop improvement and biodiversity protection programs (H3Africa consortium et al., 2014). In all African countries, the supply and quality of food must rise to meet the demand. Genomics-driven biotechnology will generate prosperity for producers and consumers by lowering costs and increasing plant yield and quality. Furthermore, famine and malnutrition in Africa may be alleviated by applying genomics or other biotechnology tools to improving subsistence crops and investigating the potential of medicinal plants (for more detail, refer to http://africanorphancrops.org).

Modern genomics and genetic engineering technologies have emerged as powerful tools that can provide roadmaps to sustainable agriculture. These have been successfully utilized and have led to significant achievements like the development of high-yielding varieties, new hybrids, higher-quality products, and transgenic crops. Genetic engineering techniques to develop genetically modified (GM) crops have garnered much attention worldwide. Interestingly, most countries are unwilling to use GM

seeds, but cross-bred seeds can provide a more workable solution. Therefore, it is necessary to pay attention to the other non-controversial, non-invasive, promising applications of genomic tools in African agriculture, plant biodiversity, and pharmaceutical and herbal industries.

Furthermore, the major bottlenecks affecting the productivity of the so-called orphan (neglected or underutilized) crops in Africa include little or no selection of improved genetic traits and extreme environmental conditions. Nevertheless, some orphan crops have recently been brought to the attention of the global scientific community, where advanced research and development projects have been initiated (Chiurugwi *et al.*, 2019; Epping and Laibach, 2020; Hendre *et al.*, 2019; Jamnadass *et al.*, 2020; Kamei *et al.*, 2016; Kumar and Bhalothia, 2020; Tadele, 2018; Tadele, 2019; Tadele and Bartels, 2019). These initiatives, introducing a range of genetic and genomic methods, addressed significant constraints affecting orphan crop production and/or nutritional efficiency. There is an urgent need for concerted efforts to advance research and development of both major and orphan crops using plant genomics to achieve food security and ultimately improve the livelihood of the population.

Plants possess mitochondrial, chloroplast, and nuclear genomes. The nuclear genome, which is the largest and most complex, contains repetitive sequences and retrovirus-like retrotransposons. Plant genomics studies have elucidated the genetic composition, structure, organization, function, diversity, and interactions embedded in plant genome sequences, and high-throughput technologies, tools, and methodologies have been developed to study the plant genome at the molecular, chromosomal, biochemical, and phenotypic levels. Jackson *et al.* (2006) pointed out how to choose the plant species to sequence and the sequencing technologies to be used. Genome assembly of sequenced plants aims to generate the complete ordered genomic sequence of a plant. Many bioinformatics algorithms for *de novo* assembly of sequencing reads are available. However, genome assembly is challenging in the plant genome due to repeats in the genome sequence and/or polyploidy (Pirita *et al.*, 2019). In general, it is highly recommended to use plant-specific genome assemblers that consider the complexity of the plant genome (Michael and VanBuren, 2019; Schatz *et al.*, 2012). Understanding of the plant genetic structure and its genetic organization enhances our comprehension of plant system biology beyond the genomic level (Campos-de Quiroz, 2002). The bioinformatics approaches to investigate the plant genetic structure and its genetic organization include annotating functional and non-functional elements, studying genome regions' epigenetic status, identifying repetitive sequences, and studying the dynamic nature of the genome.

The term pan-genome was introduced in 2005 by Tettelin *et al.* to refer to the number of genes shared by taxonomically related species (Tettelin *et al.*, 2005). The pan-genome concept includes both function-based and structure-based analyses (Tranchant-Dubreuil *et al.*, 2019). The function-based pan-genome definition considers extending the shared genes into gene families. However, the structure-based definition of a pan-genome considers only the non-redundant set of sequences (about 100 base pairs) among the collection of individual genomes (Tranchant-Dubreuil *et al.*, 2019). The study of the plant pan-genome could result in the creation of a pan-reference genome, which could be used to improve all sequence-based bioinformatics analyses, such as the accuracy of the mapping process (Bayer *et al.*, 2020). The bioinformaticians' roles in the pan-genome study involve genome-wide association study (GWAS) data analysis, transcriptomics, proteomics, and constructing and mining the genomic networks. Nowadays, the pan-genome has been constructed for several plant species, including Asian rice (*Oryza sativa*) and African rice (*Oryza glaberrima*) (Tranchant-Dubreuil *et al.*, 2019). Nonetheless, the study of plant pan-genomes still faces some challenges, such as storing the huge genomes of most plants.

Since the first plant genome sequence was released in 2000 (Arabidopsis Genome Initiative, 2000), plant functional genomics have entered the high-throughput stage. Understanding of the genomic functions helps in establishing links between phenotypes and their underlying genotypes. Key contributions of bioinformaticians to the study of plant genomic functions include providing and maintaining databases (Ong *et al.*, 2016). The bioinformatics databases for plant genomics functions are snowballing rapidly. Examples of public databases for querying plant genomes include KEGG (Tokimatsu *et al.*, 2011), PLAZA (Van Bel *et al.*, 2018), Phytozome (Goodstein *et al.*, 2012), GreenPhylDB (Rouard *et al.*, 2010), and PlantsDB (Nussbaumer *et al.*, 2013). None of these plant genomics databases is hosted in Africa.

Arabidopsis thaliana is commonly utilized as a model for researching multiple facets of plant biology. Its small genome (125 Mb) was chosen as the subject of the first plant genome sequencing project in 2001, which was a boost for plant genomics research. Following the sequencing of the Arabidopsis genome, hundreds of plant genomes have been sequenced, and presently a deluge of sequence data is available. The sequences of model plants and economically useful plants help to understand the plants' genome architectures and identify biotechnologically important genes.

Worldwide, plant genomic studies have had a revolutionary impact on agriculture and herbal medicine. Major success stories of plant genomic studies include those facilitated by the National Plant Genome Initiative established in 1998 by the US Government, which sought to comprehend the structures and roles of genes in crops that are

essential to agriculture, environmental conservation, health, and energy. Significant work done includes the development of a new efficient method for sorghum (*Sorghum bicolor*) biofortification (Ashokkumar *et al.*, 2020; Chandra *et al.*, 2020; Cruet-Burgos *et al.*, 2020; Siwela *et al.*, 2020) and the development of innovative technology to enable high-throughput image capture and automated visual analysis of traits within mutant and naturally varying plant populations (Krishnamurthy *et al.*, 2019; Li *et al.*, 2020).

Furthermore, the plant genomics market, which was USD 7.3 billion in 2019 and is projected to grow to USD 11.7 billion by 2025 (https://icrowdnewswire.com/; https://www.marketsandmarkets.com/), continues to increase. This growth is likely to be driven by the research attention drawn to the application of genomics in plant breeding and conservation of genetic resources worldwide. Until recently, most plant genomics research had been secluded to the industrialized world, although tropical Africa alone houses about 50 times more native plant species than countries in the temperate zone (Slik *et al.*, 2015). Plant genomics research is currently gaining ground in Africa, and the continent is poised to become a hub of plant genomic research. In this review, we present the current status of plant genomics in Africa and explore all the opportunities plant genomics might offer to benefit optimally from the rich flora in the continent.

PROGRESS IN AFRICAN PLANT GENOME SEQUENCING AND ANALYSIS

Plant genomics in Africa is developing, and most plant genomes in Africa were sequenced in the last decade, with a clear acceleration in the recent 5 years thanks to the African Orphan Crops Consortium (http://africanorphancrops.org). For example, Chang *et al.* (2018) reported the draft genome sequences of five orphan crops, which are local plants that are either underused or neglected but are important to agriculture and agroforestry. They listed some of the potential applications of the released genomes, such as the improvement of crop breeds (Chang *et al.*, 2018). To date, 60 species have undergone initial sequencing, and many plant genomes, including eight genome assemblies from the African Orphan Crops Consortium, have been published and six others are nearing release, with 20 currently underway (Hendre *et al.*, 2019). Below we summarize the achievements of 'African' plant genome sequencing projects, which we define as 'plants that grow in Africa and that are sequenced by African-based groups or by groups involving African-based partners', by the end of 2020 (Table 1). Results and statistics of these plant genome sequencing, assembly, annotation, and transcriptome analyses are summarized in Table 2.

This century might see largely completed sequences for several branches of practically all angiosperm clades, including main crops growing in Africa. These sequences can offer a strong structure for connecting genome-level events to all facets of morphological and physiological variation in crops worldwide. Tables 1 and 2 illustrate interesting features about some of the most common plants in Africa. For instance, the sizes of the African plant genomes sequenced so far vary from 278 Mb to approximately 1.5 Gb and are said to be influenced by climatic changes (Veselý *et al.*, 2020). The olive tree has by far the largest genome (about 1.5 Gb) and the largest number of genes (50 684 genes), which mainly encode proteins; only 1982 RNA genes have been identified. Making comparisons, *O. glaberrima*, *Artocarpus altilis*, and *Solanum aethiopicum* have genome sizes of 316, 833, and 1020 Mb, respectively, yet they encode nearly the same number of conventional genes (33 164, 33 986, and 34 906, respectively). On the other hand, the genome of *Sclerocarya birrea* (331 Mb) is slightly larger than that of *Moringa oleifera* (278 Mb). Interestingly, they have about the same number of genes (18 937 and 18 451, respectively), which is almost half of the number of genes of *O. glaberrima*, with a similar genome size (316 Mb). *Moringa oleifera* has the lowest number of genes (18 451) and the highest numbers of rRNA (8406) and tRNA (1241) genes.

The described genomes of these African plant species will serve as a valuable complementary resource for non-model food crops, especially legumes. They will be of use to both agroforestry and evolutionary science. Improvement of these plants using genomics-assisted techniques and methods could help deliver food security to millions of people. The obtained genome data could help define and characterize agronomically significant genes and characterize their modes of operation, allowing genomics-based, evolutionary studies and breeding strategies to evolve quicker, in order to establish more oriented and predictable crop improvement programs.

One of the first 'African' plant genomes to be sequenced was *Olea Europaea*, by the International Olive Genome Consortium, which includes two African-based teams, from Morocco and South Africa. The *O. Europaea* genome offers significant insights into oil biosynthesis evolution and constitutes a valuable tool for the genomics of oil crops (Unver *et al.*, 2017). Whole-genome sequencing and gene expression studies provide insight into the nature of oil biosynthesis and assist future research aimed at further increasing olive oil production.

Faidherbia albida is known for its traditional use to treat conditions such as fever, diarrhea, hives, vomiting, cough, rheumatism, or even hemorrhage (Tijani *et al.*, 2008). The study of its genome promises to provide more insights into its anti-inflammatory properties (Chang *et al.*, 2018).

The genome of *Vigna subterranea*, or Bambara groundnut, has also been sequenced (Chang *et al.*, 2018). The plant originates from West Africa and is cultivated in sub-Saharan regions, particularly in Nigeria (Borget, 1992;

Table 1 List of 13 sequenced '**African**' plant genomes. 'African' here refers to plants that are grown in Africa and have been sequenced and analyzed by African-based groups or by groups involving African-based partners

	Plant species	Tax ID	Contributing African institution(s)	Publication
1	*Olea europea*	158386	• University Mohammed Premier, Oujda, Morocco • Faculty of Sciences, University Mohamed V, Rabat, Morocco • Genomics Research Institute, University of Pretoria, Pretoria, South Africa	Unver *et al.* (2017)
2	*Argania spinosa*	85884	• University Mohammed Premier, Oujda Morocco • National Institute of Agricultural Research (INRA), Rabat, Morocco • Faculty of Sciences, University Ibn Zohr, Agadir, Morocco • Faculty of Sciences, University Mohamed V, Rabat, Morocco • Genomics Research Institute, University of Pretoria, Pretoria, South Africa	Khayi *et al.* (2018)
3	*Faidherbia albida*	138055	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa	Chang *et al.* (2018)
4	*Vigna subterranea*	115715	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa	Chang *et al.* (2018)
5	*Lablab purpureus*	35936	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa	Chang *et al.* (2018)
6	*Sclerocarya birrea*	289766	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa	Chang *et al.* (2018)
7	*Moringa oleifera*	3735	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa	Chang *et al.* (2018)
8	*Artocarpus heterophyllus*	3489	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Center for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology; University of Pretoria, Pretoria, South Africa	Kumar and Bhalothia (2020)
9	*Artocarpus altilis*	194251	• African Orphan Crops Consortium, World Agroforestry Centre (ICRAF), Kenya • Center for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology; University of Pretoria, Pretoria, South Africa	Kumar and Bhalothia (2020)
10	*Solanum aethiopicum*	205524	• African Orphan Crops Consortium, World Agroforestry Centre, Kenya • Biosciences Eastern and Central Africa, - International Livestock Research Institute, Nairobi, Kenya • International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)-Eastern and Southern Africa, Nairobi, Kenya	Song *et al.* (2019)
11	*Oryza glaberrima*	4538	• Africa Rice Center, Cotonou, Benin	Wang *et al.* (2014)
12	*Dioscorea cayenensis* subsp. *rotundata*	55577	• International Institute of Tropical Agriculture, Ibadan, Nigeria	Tamiru *et al.* (2017)
13	*Digitaria exilis*	1010633	• Laboratory of Botany, University of Dakar, Dakar, Senegal • Senegalese Agricultural Research Institute, Dakar, Senegal • Laboratoire Mixte International LAPSE, Dakar, Senegal	Abrouk *et al.* (2020)

Table 2 Statistics and methods of genome sequencing, assembly, and annotation of 'African' plants

	1	2	3	4	5	6	7	8	9	10	11	12	13
Species	*Olea europaea*	*Argania spinosa*	*Faidherbia albida*	*Vigna subterranea*	*Lablab purpureus*	*Sclerocarya birrea*	*Moringa oleifera*	*Artocarpus heterophyllus*	*Artocarpus altilis*	*Solanum aethiopicum*	*Oryza glaberrima*	*Dioscorea cayenensis subsp. rotundata*	*Digitaria exilis*
Chromosome number (2n)	46	20	26	22	22	28	28	56	56	48	12	40	36
Genome size (Mb)	1480	615	661	550	385	331	278	982	833	1020	316	570	893
Genome sequencing													
DNA organ source	Leaves	Leaves	Leaves	Leaves	Leaves	Leaves	Leaves	Leaves	Leaves	Leaves	–	Leaves	Leaves
Sequencing platform	Illumina HiSeq 2000	Illumina HiSeq Xten/PacBio RSII	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina HiSeq 2000	Roche/454 GS FLX	Illumina HiSeq 2500	Illumina HiSeq 2500
Number of reads	710 M	957.5 M/6.7 M	–	–	–	–	–	2.736 G	2.727 G	242.61 G	3485 MTP BAC	–	–
Average read size	150 bp	150 bp/6000 bp	100 bp	100 bp	100 bp	100 bp	100 bp	100 bp	100 bp	100 bp	–	100 bp	–
Transcriptome sequencing													
RNA organ(s) source	Leaves, stem, pedicel, fruits	NA	Seeds, leaves, stem	Leaves, stem	Leaves, petal, sepal, stem	Leaves	Seeds, leaves, flowers, stem, bark	Leaves, stem	Leaves, roots, seedling, stem	Seedling	NA	Leaves, stem, flowers, tuber	Leaves, panicle, seeds
Sequencing platform	Illumina HiSeq 2000	NA	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina MiSeq	NA	Illumina HiSeq 2500	Illumina HiSeq 2500
Number of reads (M)	–	NA	360	360	360	360	360	394	223	–	NA	519	–
Read size (bp)	–	NA	100	100	100	100	100	100	100	–	NA	100	–
Assembly													
Assembly method	de novo single library assembly using SoapDeNovo	hybrid de novo assembly using Masurca	de novo single library assembly using SoapDeNovo2	de novo single library assembly using SoapDeNovo2	de novo single library assembly using SoapDeNovo2	de novo single library assembly using SoapDeNovo2	de novo single library assembly using SoapDeNovo2	de novo single library assembly using Platanus	de novo single library assembly using Platanus	de novo single library assembly using Platanus	Poolwise assembly of the MTP BAC, then mapping them to each other to construct chromosomes	de novo single library assembly using ALLPATHS-LG, and scaffolding using SSPACE	de novo single library assembly using DeNovo MAGIC3
N50 (kb)	364.6	49.916	692.039	640.666	621.373	335.449	957.246	547.861	1536	516.1	217	2120	83.702
Number of contigs/scaffolds	42 843	75 327	51 470	65 586	118 976	40 280	22 329	108 267	98 152	162 187	5309	4704	18 026
Genome size (Mb)	1480	615	661	550	385	331	278	982	833	1020	316	570	893
GC content	38.8%	33%	–	–	–	–	–	32.9%	32.3%	33.13%	–	35.83%	–

(continued)

Table 2. (continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13
Annotation method	Homology-based method and de novo method using RNA-seq libraries	NA	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries	AB initio gene prediction and homology-based annotation	AB initio gene prediction and homology-based annotation using RNA-seq libraries	Homology-based method and de novo method using RNA-seq libraries
Total number of predicted genes	50 684	NA	28 979	31 707	20 946	18 937	18 451	35 845	33 986	34 906	33 164	26 198	44 542
tRNA	798	NA	458	756	611	564	1241	713	689	960	701	–	–
rRNA	773	NA	1008	1080	633	313	8406	466	2706	1185	–	–	–
miRNA	411	NA	126	102	109	106	111	159	168	128	–	–	–

'NA', not applied (is not done yet); "–", not mentioned (the piece of information was not mentioned in their original paper); M, million; G, Billion.

Linnemann, 1992). This plant is drought-tolerant and has significant nutritional value (Chang *et al.*, 2018; Gbaguidi *et al.*, 2018). The study of this plant's genome will help to establish efficient breeding strategies to improve its decreasing yield.

Lablab purpureus, or Dolichos bean, is a member of the Fabaceae family, is one of the most ancient (>3500 years) domesticated legume species (Maass *et al.*, 2005) native to Africa (Maass *et al.*, 2010), and is cultivated throughout the tropics for food (Ojo *et al.*, 2013). It is often grown as forage for livestock and used as a medicinal plant (Al-Snafi, 2017). *Lablab purpureus* has a good nitrogen-fixing ability, and it is highly adaptable to diverse environmental conditions (Robotham and Chapman, 2017). Based on reports regarding the limited yields of this crop, it is critical to identify and characterize the agronomically important genes; hence, the *L. purpureus* genome is studied (Chang *et al.*, 2018). On the other hand, marula (*S. birrea*) is a popular wild tree found in many African countries. Its leaves, stem, bark, roots, and fruits are used for food and traditional medicine (Mariod and Abdelwahab, 2012). The *S. birrea* fruit is not only renowned for its abundance in ascorbic acid (Mariod and Abdelwahab, 2012), it also contains a hard-brown seed enclosing a soft white kernel rich in oil and protein (Mariod and Abdelwahab, 2012). The draft genome of this species will serve as a valuable complementary tool for non-model food crops to elucidate the medical and socio-economical importance of marula seeds, fruits, and oil and to discover genes involved in the biosynthesis of its metabolites (Chang *et al.*, 2018). Horseradish tree (*M. oleifera*), a member of the Moringaceae family, is a drought-tolerant plant distributed throughout tropical and subtropical countries including those in Africa (Rivas *et al.*, 2013). This plant covers the major agroecological regions in Nigeria (Busani *et al.*, 2011) and has important nutritional and medicinal values (Busani *et al.*, 2011). The seeds are used to extract oil, which can also be a good source of biodiesel (Ciolkosz, 2015; Rashid *et al.*, 2008). The draft genome of this species (Chang *et al.*, 2018) can be exploited to enhance biodiesel production. The use of biodiesel as fuel increases energy security, improves air quality in the environment, and provides safety benefits (Rashid *et al.*, 2008).

Two 'African' members of the Moraceae family, *Artocarpus heterophyllus* (Lam.) and *A. altilis*, have seen their genomes being recently sequenced (Kumar and Bhalothia, 2020). *Artocarpus heterophyllus*, also known as jackfruit (Prakash *et al.*, 2009), is native to the Western Ghats of India and Malaysia, but it is also found in Central and Eastern Africa (Prakash *et al.*, 2009). Flakes of ripe *A. heterophyllus* fruits are high in nutritive value (Elevitch and Manner, 2006). The plant *A. altilis*, also known as breadfruit, has been a reliable staple food in the Pacific region for more than 3000 years, with hundreds of named

cultivars (Ragone, 2018). Breadfruit is grown and used in many tropical regions, including Africa, with most trees originating from just a few Polynesian cultivars that were introduced in the late 1700s (Ragone, 2018). *Artocarpus altilis* trees generate a variety of nutritious fruits and are easy to cultivate. Breadfruit is rich in vitamin C and carbohydrates, and may be used to support cardiovascular health, aid digestion, and treat dandruff (Ragone, 2018). The study of the genomes of these two Moraceae species will help to elucidate the mechanisms underlying the high energy density in their fruits and the genes involved in sugar and starch metabolism.

The African eggplant (*S. aethiopicum*), an indigenous non-tuberous crop that mainly grows in tropical regions in Africa (Sunseri *et al.*, 2010), is a member of the Solanaceae family. By dint of its morphological hypervariability (Adeniji *et al.*, 2012; Plazas *et al.*, 2014), the *S. aethiopicum* is classified into four groups: Gilo, Shum, Kumba, and Aculeatum. The creation of a reference genome sequence will help identify disease resistance genes and discover the expanded gene family responsible for bacterial spot resistance in eggplant (Song *et al.*, 2019). Two members of the Poaceae family have undergone genomic studies. One is *O. glaberrima*, the first ever 'African' plant genome to be sequenced (Wang *et al.*, 2014). It is an African rice species that was independently domesticated from the wild progenitor *Oryza barthii* approximately 3000 years ago (Sweeney and McCouch., 2007), 6000–7000 years after the domestication of Asian rice (*O. sativa*) (Portères, 1976). *Oryza glaberrima* (African rice) is an essential food crop in several sub-Saharan countries. The study of the genome of African rice and the population genomics of rice helps to discover the evolution of such species in Africa and how to maintain their sustainability. The second 'African' Poaceae family member whose genome has been sequenced is fonio millet (*Digitaria exilis*), an African orphan cereal crop that has great potential for dryland agriculture. Abrouk *et al.* (2020) established high-quality genomic resources to improve fonio by molecular breeding. This study has provided new insights into the genetic diversity, population structure, and domestication of fonio millet (Abrouk *et al.*, 2020).

White guinea yam (*Dioscorea rotundata*), a member of the Dioscoreaceae family, is a tuber-bearing crop. It is a very popular species in West and Central Africa, which are considered as the main regions for yam production worldwide (Tamiru *et al.*, 2017). *Dioscorea rotundata* represents a major source of food and income in these regions as well as an integral part of the socio-cultural life of the natives. Genomic analysis of orphan crops such as yam promotes efforts to improve food security and the sustainability of tropical agriculture (Tamiru *et al.*, 2017). Finally, the genome of the argan tree (*Argania spinosa*), a member of the African Sapotaceae family which serves essential socio-

economic and ecological roles in various regions, such as the arid zone in the Southwestern region of the North African country Morocco, has also been sequenced (Khayi *et al.*, 2018). This tree is used as food, as animal feed, and for cosmetic and pharmaceutical purposes. This African endemic genetic asset has experienced serious threats from climate change, environmental factors, and human overexploitation in recent decades. Study of the genetics and genomics of the argan tree will help overcome these challenges (Ghazal *et al.*, 2021). The elucidation of the genomic basis of the production of important metabolites by *A. spinosa*, including fatty acids, tocopherols, squalene, sterols, and phenolic compounds, strongly stimulated molecular breeding studies in the argan oil industry. Genomic studies of the argan tree could facilitate the discovery of the biosynthesis pathways of oil and other important metabolites, which will help ameliorate the quality and quantity of these metabolites.

OTHER PLANT GENOME-RELATED PROJECTS

Genome-wide single-nucleotide polymorphism analysis of the Ethiopian *Capsicum*

Ethiopian *Capsicum*, a genus in the Solanaceae family, is commonly known as chili pepper. *Capsicum* species are major crop plants and are almost globally distributed (Panda *et al.*, 2004). Chili pepper fruits are not only used as spices and vegetables but also for medicinal purposes due to their richness in vitamins A and C (van Zonneveld *et al.*, 2015). Chili pepper fruits are also used as natural coloring agents, cosmetics, and active ingredients in host defense repellents. The genus includes 27 species, of which five are known to be domesticated (Ince *et al.*, 2010). The five cultivated *Capsicum* species, namely *C. annuum* L., *C. chinense* Jacq., *C. frutescens* L., *C. baccatum* L., and *C. pubescens* Ruiz & Pav., are among the most economically important vegetables worldwide (González-Pérez *et al.*, 2014). The characterization of genetic diversity in different *Capsicum* species growing in different regions has shown that the ecological distribution significantly influences plant genetic diversity (Lee *et al.*, 2016). Therefore, analysis of genetic variation in Ethiopian *Capsicum* species involving many cultivars will help breeders utilize the germplasm collection to improve existing commercial cultivars (Solomon *et al.*, 2019).

To capture genetic diversity, a collection of 142 *Capsicum* genotypes was established from different geographic areas of Ethiopia. The high-resolution melting curves and the morphological characteristics demonstrated that the collection included one *Capsicum baccatum*, nine *Capsicum frutescens*, and 132 *Capsicum annuum* accessions. Measurement of plant growth parameters revealed variation in germplasm parameters such as plant height, stem thickness, internode length,

the number of side branches, fruit width, and fruit length. Broad-sense heritability for fruit weight was maximum, followed by leaf length and width. Genotyping by sequencing (GBS) (Solomon *et al.*, 2019) was performed to identify single-nucleotide polymorphisms (SNPs) in the panel of 142 *Capsicum* germplasms, and 2 831 791 genome-wide SNP markers were identified. Among these, 53 284 high-quality SNPs were selected and used to estimate the level of genetic diversity, the population structure, and phylogenetic relationships. From model-based ancestry analysis, the phylogenetic tree, and principal coordinate analysis (PCoA), two distinct genetic populations were identified: one composed of 132 *C. Annuum* accessions and one consisting of nine *C. frutescens* accessions. GWAS (Solomon *et al.*, 2019) analysis identified 509 SNP markers that were significantly correlated with fruit, stem, and leaf traits.

3000-Year-old Egyptian emmer wheat (*Triticum turgidum* subsp. *dicoccum*)

Tetraploid emmer wheat (*Triticum turgidum* subsp. *dicoccum*) (NCBI tax ID: 49225) is a progenitor of the world's most widely grown crop, hexaploid bread wheat (*Triticum aestivum*). Emmer wheat was one of the first cereals domesticated in the old world as it was cultivated around 9700 BC in the Levant (Arranz-Otaegui *et al.*, 2016) and subsequently in Northern Africa. Radiocarbon dating showed that the entire genome sequence of the emmer wheat specimen from the Egyptian Museum could be traced back to the New Kingdom between 1130 and 1000 BC. The aforementioned wheat genome seems to be unusual, carrying haplotypes that are absent in modern emmer (Fuller and Lucas, 2014). Therefore, it is important to study the wheat genome to comprehend the history and diversity of ancient cereals, as ancient DNA sequences can reveal dispersal and domestication histories.

Since 1921, emmer wheat samples have been excavated from the archaeological site of the Hememiah North Spur in Egypt (Brunton and Caton-Tompson, 1928). Sequencing was performed using an Illumina NextSeq 500 sequencer, resulting in 861 million reads with a 75 bp length (Scott *et al.*, 2019). Next, sequences were aligned to the Zavitan v.2 reference genome (Avni *et al.*, 2017) for emmer wheat using BWA aligner (Li and Durbin, 2009). SNPs and genotypes were called using the Genome Analysis Tool kit GATK described by McKenna *et al.* (2010).

Based on the 99 078 SNPs, several methods were used to map genome-wide similarity between the Egyptian wheat and modern ones. This confirmed that Egyptian wheat is genetically closest to the domesticated ones, specifically to the domesticated Indian Ocean subgroup. Not only the identity but also the state of these SNPs confirmed that Egyptian wheat is most concordant (86.4–87.4%) with the Indian Ocean subgroup compared with the other domesticated ones (mean, 81.7%; SD, 1.8%) or wild ones (mean, 79.1%; SD, 1.6%). Phylogenetic analysis showed that Egyptian wheat branches are closest to (but not classified in) the Indian Ocean clade. Moreover, the admixture (Patterson *et al.*, 2012) source-population inference showed that Egyptian wheat shares most ancestry with the Indian Ocean subgroup (Scott *et al.*, 2019).

Egyptian wheat has several specific haplotypes that were described using 50 SNP sliding windows (McKenna *et al.*, 2010). Such unusual haplotypes may reflect missing alleles or incomplete sampling of modern emmer (in the Indian Ocean subgroup, none of the sequenced modern emmers is present, and none is from Africa). Emmer cultivation in the Nile Valley has all but vanished after the Roman period; therefore, it is likely that much of the genetic richness of ancient Egyptian emmer has been lost (Scott *et al.*, 2019).

APPLICATIONS OF PLANT GENOMICS IN AFRICA

Plant genetics resources and future trends in plant genomics in Africa

The strength of Africa lies in its natural resources, including genetic resources (Kent *et al.*, 2003; Vicente and Schlebusch, 2020). The latter is the foundation for advancement in agriculture, environment, and the medical field, especially in the areas of drug discovery and disease treatment. Plant genomic studies have buttressed the fact that many plants and food crops originate from different parts of Africa, including West Africa's Niger River basin, North of the Niger River, a part of the Western Sahara Desert that today includes Northern Mali and Mauritania, and East, Central, and Southern Africa. Examples of such crops are African rice (*O. glaberrima*), African yam (*D. rotundata*), pearl millet (*Cenchrus americanus*), and wild millet (*Pennisetum glaucum monodii*) (Pennisi, 2019). An early cradle of agriculture existed around West Africa's Niger River basin, where the first plant genomics studies started appearing. Many traditional food crops on the continent began with a cereal called pearl millet, which was the African version of rice (Pennisi, 2019).

African cereals have been part of the Cereal Genome Sequencing Efforts on Crop Improvement (Laura, 2017). Resources are available for rice, sorghum, millet, and maize (*Zea mays*). The use of emerging high-end genomic technologies can be expanded from crop plants to traditional medicinal plants. The use of high-end genomic technologies expedites medicinal plant breeding and transforms them into living factories of medicinal compounds (Da-Cheng and Pei-Gen, 2015). The utility of molecular phylogeny and phylogenomics in predicting chemodiversity and bioprospecting is vital to the discovery and development of natural product-based drugs (Peterson *et al.*, 2019).

The role of plant genomics in African agriculture

Agriculture is currently facing the 'perfect storm' of climate change, increasing fertilizer prices, and increasing food demand from a growing and wealthier human population (Abberton *et al.*, 2016). These indicators may lead to a global food shortage unless crop production is increased in terms of productivity and resilience. Agriculture intensification has centered on increasing productivity under improved conditions with important agricultural inputs (Abberton *et al.*, 2016). In addition, intensive cultivation of a limited number of crops has significantly limited the number of plant species on which humans rely. That alone needs a new agricultural paradigm, which decreases dependence on high inputs, increases crop diversity, and provides stability and resilience to the environment (Abberton *et al.*, 2016). Genomics brings great opportunities to boost crop yield, quality, and output stability via advanced breeding techniques. Genomics also offers the potential to improve major crop tolerance to climate change, enhance productivity, and select minor crops to diversify Africa's food supply (Abberton *et al.*, 2016). It is necessary to have a deep understanding of genomic-assisted breeding for most essential staples feeding the planet and how to use and adapt these genomic tools to improve the production of large and small crops with desirable features that enhance adaptation to or mitigate the effects of climate change (Abberton *et al.*, 2016).

Comprehension of genetic variation in a population at the level of the DNA sequence enables the identification of agronomically important genes and associated molecular genetic markers while offering a way of selecting certain genes in breeding programs (Abberton *et al.*, 2016). The key driver of the current genomic revolution is the development of next-generation sequencing (NGS) technology (Abberton *et al.*, 2016). This technology revolutionizes the production of crops as quickly as it has revolutionized medicine. This allows the sequencing of different crop genomes and promotes the correlation of genomic diversity with agronomic characteristics, thereby providing the basis for genomic-assisted breeding. Sequencing technology continues to develop and all the major crops will likely have benefited from sequence-based genomic advancement within a few years (Abberton *et al.*, 2016).

Genome editing technology has great potential to revolutionize agricultural productivity on the African continent, particularly in sub-Saharan Africa, as Komen *et al.* (2020) pointed out that CRISPR-Cas9 is already being used in Africa to improve primary staple foods including wheat, cassava (*Manihot esculenta*), and banana (*Musa* spp.). Indeed, the research results are looking propitious (Komen *et al.*, 2020). In addition, international agricultural science centers are now implementing genome editing in their research and development programs in partnership with national science organizations within Africa (Komen *et al.*, 2020). Research in a variety of African countries is currently in progress. For example, CRISPR-Cas9 technology is used by the Kenya Agriculture and Livestock Research Organization and two other international organizations to develop maize germplasm further until it is immune to maize lethal necrosis (Boddupalli *et al.*, 2020).

The role of plant genomics in herbal medicine in Africa

Herbal medicines have been widely used for more than 5000 years in traditional medicine and ethnomedicine worldwide (Tu, 2011). The World Health Organization has identified 21 000 medicinal crops, out of which 2500 are located in Africa (Modak *et al.*, 2007). Of the African and Indian population, 90% and 70% depends on traditional medicine, respectively. In China, traditional medicine accounts for around 40% of all healthcare delivered (WHO, 2005).

With the fast advances in high-throughput sequencing technologies, a new herbal research discipline, called 'herbal genomics' (herb-genomics), has been established, which combines herbal and genomic research and uses this information together with transcriptomic, proteomic, and metabolomic data to predict the secondary metabolic pathways in herbs. Therefore, it provides a general picture of the genetic background of traditional herbs, enabling researchers to investigate the mechanisms to secure the sourcing of medicinal plants and their active compounds in the future. Furthermore, herb-genomics helps prevent and treat human diseases from an omics perspective (Chakraborty, 2018; Hu *et al.*, 2019). Molecular techniques (such as the random amplified polymorphic DNA method, inter-simple sequence repeat-PCR, amplified fragment-length polymorphism analysis, microsatellite markers, SNPs, expressed sequence tags, microchip-based genomic profiling, and DNA barcoding) have been well proven to be powerful tools for authentication of medicinal plants based on phylogenic variation signatures on chloroplast and nuclear DNA (Gantait *et al.*, 2014). In recent years, over 7200 plant chloroplast genomic sequences have been published, and over 20% of these plants are believed to have medicinal properties (Hu *et al.*, 2019).

A study conducted by Su *et al.* (2008) reported the development of strain-specific sequence-characterized amplified region markers for strain detection and authentication of *Ganoderma lucidum* (Su *et al.*, 2008). Balasubramani *et al.* (2011) have developed DNA markers from genomic DNA by amplifying and sequencing whole internal transcribed spacer regions (ITS1, 5.8S rRNA, and ITS2). In addition, the use of universal primers has been proven to be effective in authenticating *Berberis* species. These primers are functional as a molecular pharmacognostic, which is used in quality assessment of raw drugs (Balasubramani *et al.*, 2011; Su *et al.*, 2008). Herb-genomics research has

obtained great results, even though it is still in its early developmental stage (Hu *et al.*, 2019). Advanced bioinformatics tools and omics approaches will allow herbgenomics to open a new field for fundamental and applied research into natural drugs (Hu *et al.*, 2019). Moreover, genome sequencing projects may be initiated for medicinal plants. To this end, there is a major need to develop a sequence database with all the bioinformatics tools for all plants and make them available for public use by different nations (Gantait *et al.*, 2014).

To date, just a little part of the plant metabolism has been explored to produce new medicines and other products. Accordingly, there is interest in more herbal genomics studies and the developments in other omics knowledge that may enable the rapid discovery of previously identified metabolic pathways and enzymes (Chakraborty, 2018). Unfortunately, till date, very little has been done in herbal genomics in Africa. There is still a huge gap to fill and plenty of benefits to gain in launching plant medicinal genomics for Africa soon.

The role of plant genomics in novel drug discovery in Africa

Novel drug discovery is often proceeded by screening of bioactive molecules in natural resources such as plants and fungi (Albarano *et al.*, 2020), where metabolites and gene products are screened for their therapeutic activities. However, the advent of genomic technologies has led to significant progress in novel drug discovery. The advance of drug discovery based on plant genomics is expanding into two directions, which are (i) wet-lab experiments and (ii) *in silico* analyses. Wet lab-driven drug discovery includes plant genome sequencing and chemogenomic analysis; *in silico* driven drug discovery includes virtual screening via molecular simulation experiments, identification of novel pathways for new drug discovery, and functional genomics analysis (Chakraborty, 2018). Several studies have been performed in several African countries using *in silico* and modeling tools and methods for plant-based drug discovery.

CONCLUSION AND THE WAY FORWARD

The African continent is currently facing various challenges, including global health and food security for an ever-increasing population. Furthermore, approximately 20% to 30% of the earth's surface is arable land, causing persistent reductions in productivity. As a result, millions may be facing protein deficiency, with three-quarters of the world population getting their daily protein dose from plants. Plant genomics may be able to provide the scientific advances the African societies need for crop improvement. NGS of agricultural crops will also allow for the development of more productive and sustainable practices so that Africa can meet the challenges of feeding and

healing a growing population in a changing environment. The current generation of DNA sequencing technologies is making genome sequencing a reality without requiring generous funding sources. This momentous leap brings exciting opportunities to the botanical and plant communities, also in Africa. Whole-genome sequences may contain thousands of nuclear markers for phylogenetic and population-level analyses, allowing genome-wide investigations of fundamental evolutionary and ecological concerns. Additionally, the creation of a pan-genome, recording the genomic richness of ecotypes, regional isolates, and associated organisms (Golicz *et al.*, 2016), would make it possible for comparative methods and association analyses to classify the genetic components of such characteristics and adaptations. The possibilities run the gamut from systematics, ecology, and evolution to molecular genetics.

On the other hand, the incidence and prevalence of non-communicable and infectious diseases have risen, affecting public health. Besides, the existence of new pathogenic species exacerbates the incidence and prevalence of infectious diseases. Most of these diseases lack appropriate therapies, whereas others are still being handled through decades-old procedures with several limitations. However, the production of novel drugs for certain diseases remains a major challenge. Notably, the pharmaceutical industry has entered an 'innovation crisis' and has been producing at a constant pace for nearly 60 years. Natural medicinal plant extracts have provided the basis for several Western medicines and herbal medicines, a cornerstone of healthcare in the developing world. Many ancient civilizations, including African, Arab, Persian, Indian, and Chinese civilizations, have popularized herbal remedies in developing countries. Chinese and Indian research institutes have initiated multiple genome projects on medicinal plants of Chinese or Indian origin to further explore their economic benefits. A primary propeller has been the effective synthesis of semi-synthetic artemisinin and its entry into commercial production for malaria treatment (Tiwari *et al.*, 2019). This insinuates that harnessing the plant's maximum therapeutic potential will help fill the gap between disease outcomes and therapy management, hence enhancing the health outcomes of the global population. Searching for bioactive compounds from medicinal plants is not a new African research initiative. Even so, in previous studies, two key components were lacking in our view: (i) Most of the studies have only been national or regional (e.g., Bouyahya *et al.*, 2020; Kamau *et al.*, 2016; Okot *et al.*, 2020; Oladeji *et al.*, 2020; Patrick *et al.*, 2015; Spiegler., 2020; Twilley *et al.*, 2020). To achieve the same effect and coverage as the traditional Chinese medicine effort, a continental coverage is needed. (ii) Bioinformatics or computational biology was not extensively applied in any past works. Hence, most of the strategies were based on the

wet lab. However, inspections of wet lab validations of obtained promising active compounds revealed that the experimental methods sometimes resulted in higher levels of defections in the final validated active compounds.

The global plant community is getting more convinced than ever that this is the perfect time to be involved in plant genomics research, especially in Africa, for crop improvement and optimization of medicinal plant use. The sequencing of African plant genomes presents a conceptual framework for the establishment of a regional and integrative omics strategy for comprehensive crop and medicinal plant characterization. Plant genomics research will concentrate on predictive tools and functional analysis to annotate the genomes and decipher metabolic pathways through the study of metabolomes and analysis of the biosynthesis of key metabolites as well as the plant microbiome. To this end, collaboration through a pan-African plant genome (PAPG) consortium would be one of the most suitable and efficient ways to organize the broad interests of the plant genomics community for the precious African endemic and orphan crops and plants.

The lack of large-scale plant genomics studies in Africa is the result of many deep-seated issues. Some of these issues include the lack of awareness of the pertinence of genomics in improving breeding techniques and agriculture, the lack of African scientists with genomic science experience, the lack of research infrastructure, the minimal availability of computational expertise and tools, the lack of sufficient funding for research activities from African governments, and the involvement of many African researchers in partnership research at not more than the stage of sample collection. Overcoming these limitations will, in part, depend on African scientists themselves acquiring the expertise and facilities necessary to lead high-quality plant genomics research aimed at solving plant and agriculture problems relevant to the African population and to become internationally competitive in plant genomic science and its applications.

Various initiatives have been established at the regional and continental scales, such as H3Africa and the African Orphan Crops Consortium to address these issues. The NIH-funded H3Africa initiative for a sustainable African bioinformatics network (h3abionet.org) has been instrumental in building bioinformatics facilities and skills in many African countries. H3Africa stands for Human Heredity and Health in Africa (h3africa.org). H3ABioNet projects have demonstrated the value of bioinformatics in the sense of health issues. Moreover, the has NIH sponsored many bioinformatics training programs in Africa. However, these measures have not been applied systematically to the plant-related biological field. To overcome this issue, we propose establishing a Pan African Medicinal Plant Bioinformatics Network (AMPBioNet) to support efforts to expand the reach of H3ABioNet and spread experience

with plants across the continent. A primary emphasis on medicinal plants would be an ideal missing link between the human/health and plant/agriculture fields with the ambition of extending the existing tools and learned methods and strategies to other plants such as major cultivars, endemic species, orphan crops, and tropical trees. The Pan African Medicinal Plant Bioinformatics Network (AMPBioNet) will do so by (i) pinpointing already sequenced plants, and if not yet sequenced, sequencing targeted plants based on their pharmacological properties for generating assembled and annotated plant genomes; (ii) developing both analytical and statistical models for screening such genomes to elucidate their pharmacological properties and completely characterize them; and (iii) build an interdisciplinary and multidisciplinary network of African researchers designing and implementing models using biocomputing tools and facilitating translation into novel plant-based health interventions.

Accordingly, the Pan-African medicinal plant bioinformatics network AMPBioNet will examine traditional African medicines for novel compounds to treat most common or re-emerging diseases in Africa, such as malaria, leishmaniasis, trypanosomiasis, tuberculosis, and colon and prostate cancers, through network pharmacology and genomics approaches. As mentioned above, although many genomes of African plants have been sequenced, most funding is external. Thus, the research is outsourced and training is mostly given by foreign experts. This often raises sustainability issues and brain retention concerns. From this viewpoint, H3Africa for human and health is a very convenient model for large-scale actions to follow in plant sciences and agriculture. The already launched genomics revolution for human health in Africa through the pan-African H3Africa project could support agriculture and speed up the discovery of new plant genomics methods provided the lessons learned from H3Africa are well implemented from a similar plant, physiology, pathology, and progeny (P4Africa) perspective.

The availability of affordable NGS technologies has boosted research on plant genomics worldwide. Africa is on the right track to benefit optimally from this situation. In recent years we have seen the emergence of a dozen sequenced African plant genomes thanks to the contributions of African scientists. The NGS technologies enabled sequencing, assembly, and annotation of key African plant genomes, thereby generating an unprecedented volume of information for local African researchers and breeders. Using this deluge of plant genomics data within the continent places an unprecedented challenge upon the few local bioinformaticians to deal with the volume, speed, and variety of the data generated by local academia and to improve storage, sharing, and analysis capacities. Plant bioinformatics is expected to bring new insights into the genetics of the crops, associated with phenotyping,

enabling the optimal exploration of plant genetic resources for a range of significant end-user characteristics to promote more uses to improve food production, nutritional safety, and plant biodiversity. The efforts of the African scientific community and their international colleagues will not in themselves be sufficient. National governments and regional political and economic organizations must support sustained funding of all research fields, including the development of plant genomics and associated research infrastructure.

With respect to data mining and sharing, the H3Africa consortium can be inspiring as it has developed an approach that attempts to balance (i) the protection of the ability of African scientists to be the first to analyze and publish their research findings given their limited resources and capacities to deal with data as quickly as scientists in developed countries and (ii) the benefit from global accessibility to H3Africa data and biospecimens. To meet these not fully compatible goals, the H3Africa consortium has decided that data will initially be made accessible to consortium members through H3ABioNet before they are sent to the European Genome Phenome Database, through which they will be publicly released (through an 'Independent Data and Biospecimen Access Committee'). As is typical in genomics, there will be a short lag (1 year) between submission and publication of the data. This is marginally longer than the norm (6 to 9 months) to provide resource-challenged African investigators a bit more time to analyze and submit their manuscripts for peer review.

Similar considerations were given to the development of the policy for the release of biospecimens collected in Africa. Biospecimens will indeed be stored within an African biorepository (with copies elsewhere on the continent) and circulated globally for further study. However, data and biospecimen sharing do raise the often contentious issues of ownership and commercialization rights. The prospective P4Africa consortium will address this issue while embracing an ethos that promotes research for the global common good. Resources generated by P4Africa will be expected to be useful in future plant genomic research in Africa and globally.

In conclusion, famine, malnutrition, and the health burden in the African continent may be alleviated by applying genomics or other biotechnology tools to improve subsistence crops and plants of nutritional and medicinal value. The role of the public sector could be great, but government policies on these issues are still unclear. Through the African Union and other affiliated institutions, African nations should invest in public plant genomics research with suitable data deposition and sharing policies. It is imperative that publicly funded plant genomics databases be equally accessible to everyone in the continent. Funds should be increased for improving crop germplasm in the continent. Indeed, agricultural plant genomics studies should be publicly funded as plants' DNA sequences are necessary for the continued progress to understand crop and plant ecology and biodiversity. Furthermore, plant genomics research is naturally organized in a team-based manner. This would constitute another tool to enhance South–South partnership along with the traditional South–North schema. Finally, ethical and social implications of plant genomics studies, genome editing, and breeding should be considered, and patent laws affecting plant genomics must be clarified to evaluate and improve the perception and acceptance by the African population, in order to ensure fair benefit sharing and biodiversity conservation. Policymakers have to be informed and well introduced to this new but quickly accelerating field. Undoubtedly, the future of plant genomics in Africa looks bright for both consumers and producers of food and herbal medicines. Still, without proper measures and targeted actions, the poorest nations of the continent might miss the plant genomics revolution, which will also allow for the development of more productive and sustainable practices. Together, we can meet the challenges of increasing demands for feeding and healing of a fast-growing population in a changing environment. If the dearth of plant genomics research involving Africans persists, the potential nutrition, health, and economic benefits emanating from genomic science may once again elude the entire continent.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

We declare that we have no conflict of interest.

REFERENCES

Abberton, M., Batley, J., Bentley, A., Bryant, J., Cai, H., Cockram, J. *et al.* (2016) Global agricultural intensification during climate change: a role for genomics. *Plant Biotechnology Journal*, **14**, 1095–1098.

Abrouk, M., Ahmed, H.I., Cubry, P., Simonikova, D., Cauet, S., Pailles, Y. *et al.* (2020) Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate. *Nature Communications*, **11**, 4488.

Adeniji, O., Kusolwa, P. & Reuben, S. (2012) Genetic diversity among accessions of *Solanum aethiopicum* L. groups based on morpho-agronomic traits. *Plant Genetic Resources*, **10**(3), 177–185.

Aglawe, S.B., Barbadikar, K.M., Mangrauthia, S.K. & Madhav, M.S. (2018) New breeding technique "genome editing" for crop improvement: applications, potentials and challenges. *3 Biotech*, **8**(8), 336.

Albarano, L., Esposito, R., Ruocco, N. & Costantini, M. (2020) Genome mining as new challenge in natural products discovery. *Marine Drugs*, **18**(4), 199.

Al-Snafi, A.E. (2017) The pharmacology and medical importance of Dolichos lablab (*Lablab purpureus*) – a review. *IOSR Journal of Pharmacy*, **07**, 22–30.

Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**(6814), 796–815.

Arranz-Otaegui, A., Colledge, S., Zapata, L., Teira-Mayolini, L.C. & Ibáñez, J.J. (2016) Regional diversity on the timing for the initial appearance of cereal cultivation and domestication in southwest Asia. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 14001–14006.

Ashokkumar, K., Govindaraj, M., Karthikeyan, A., Shobhana, V.G. & Warkentin, T.D. (2020) Genomics-integrated breeding for carotenoids and folates in staple cereal grains to reduce malnutrition. *Frontiers in Genetics*, **11**, 414.

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S.O., Gundlach, H. *et al.* (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, **357**, 93–97.

Balasubramani, S.P., Goraya, G.S. & Venkatasubramanian, P. (2011) Development of ITS sequence-based markers to distinguish *Berberis aristata* DC. from *B. lycium* Royle and *B. asiatica* Roxb. *3 Biotech*, **1**, 11–19.

Bayer, P.E., Golicz, A.A., Scheben, A., Batley, J. & Edwards, D. (2020) Plant pan-genomes are the new reference. *Nature Plants*, **6**(8), 914–920.

Boddupalli, P., Suresh, L.m., Mwatuni, F., Beyene, Y., Makumbi, D., Gowda, M. *et al.* (2020) Maize lethal necrosis (MLN): efforts toward containing the spread and impact of a devastating transboundary disease in sub-Saharan Africa. *Virus Research*, **282**, 19794.

Borget, M. (1992) Food legumes. In: Borget, M., Wibberley, E.J., Tindall, H.D., Wilding, J. & Angladette, A. (Eds.) *The Tropical Agriculturalist*. London: Macmillan Education, pp. 1–120.

Bouyahya, A., Chamkhi, I., Guaouguaou, F.-E., Benali, T., Balahbib, A., El Omari, N. *et al.* (2020) Ethnomedicinal use, phytochemistry, pharmacology, and food benefits of *Thymus capitatus*. *Journal of Ethnopharmacology*, **259**, 112925.

Brunton, G. & Caton-Thompson, G. (1928) *Te Badarian Civilization and Predynastic Remains Near Badari*. British School of Archaeology in Egypt.

Busani, M., Patrick, J.M., Arnold, H. & Voster, M. (2011) Nutritional characterization of Moringa (*Moringa oleifera* Lam.) leaves. African. *Journal of Biotechnology*, **10**(60), 12925–12933.

Butt, H., Zaidi, S.S., Hassan, N. & Mahfouz, M. (2020) CRISPR-based directed evolution for crop improvement. *Trends in Biotechnology*, **38**(3), 236–240.

Campos-de Quiroz, Q.H. (2002) Plant genomics: an overview. *Biological Research*, **35**(3–4), 385–399.

Chakraborty, P. (2018) Herbal genomics as tools for dissecting new metabolic pathways of unexplored medicinal plants and drug discovery. *Biochimie Open*, **6**, 9–16.

Chandra, A.K., Pandey, D., Tiwari, A., Sharma, D., Agarwal, A., Sood, S. *et al.* (2020) An omics study of iron and zinc homeostasis in finger millet: biofortified foods for micronutrient deficiency in an era of climate change? *OMICS: A Journal of Integrative Biology*, **24**(12), 688–705.

Chang, Y., Liu, H., Liu, M., Liao, X., Sahu, S.K., Fu, Y. *et al.* (2018) The draft genomes of five agriculturally important African orphan crops. *Gigascience*, **8**(3), 1–16.

Chiurugwi, T., Kemp, S., Powell, W. & Hickey, L.T. (2019) Speed breeding orphan crops. *TAG. Theoretical and Applied Genetics*, **132**(3), 607–616.

Ciolkosz, D. (2015) Chapter 34 - What's so different about biodiesel fuel? In: Dahiya, A. (Ed.) *Bioenergy*. Cambridge, MA: Academic Press, pp. 551–555. https://doi.org/10.1016/B978-0-12-407909-0.00034-1.

Conway, G. (2012) *One Billion Hungry: Can We Feed the World?* Ithaca, NY: Cornell University Press.

Cruet-Burgos, C., Cox, S., Ioerger, B.P., Perumal, R., Hu, Z., Herald, T.J. *et al.* (2020) Advancing provitamin A biofortification in sorghum: genome-wide association studies of grain carotenoids in global germplasm. *The Plant Genome*, **13**(1), e20013.

Da-Cheng, H. & Pei-Gen, X. (2015) Genomics and evolution in traditional medicinal plants: road to a healthier life. *Evolutionary Bioinformatics Online*, **11**(11), 197.

Elevitch, C.R. & Manner, H.I. (2006) Artocarpus heterophyllus (jackfruit), species profiles for Pacific Island agroforestry, **10**, 1–25. Available at: http://www.traditionaltree.org.

Epping, J. & Laibach, N. (2020) An underutilized orphan tuber crop – Chinese yam: a review. *Planta*, **252**(4), 1–19.

Fuller, D.Q. & Lucas, L. (2014) Wheats: origins and development. In: Smith, C. (Ed.) *Encyclopedia of Global Archaeology*. New York: Springer. https://doi.org/10.1007/978-1-4419-0465-2_2192.

Gantait, S., Debnath, S. & Nasim Ali, M. (2014) Genomic profile of the plants with pharmaceutical value. *3 Biotech*, **4**, 563–578.

Gbaguidi, A.A., Dansi, A. & Dossou-Aminon, I. (2018) Agromorphological diversity of local Bambara groundnut (*Vigna subterranea* (L.) Verdc.) collected in Benin. *Genetic Resources and Crop Evolution*, **65**(4), 1159–1171.

Ghazal, H. (2021) Argania genetics and genomics. In: Tombuloglu, H., Unver, T., Tombuloglu, G. & Hakeem, K.R. (Eds.) *Oil Crop Genomics*. Springer. https://doi.org/10.1007/978-3-030-70420-9.

Golicz, A.A., Batley, J. & Edwards, D. (2016) Towards plant pangenomics. *Plant Biotechnology Journal*, **4**, 1099–1105.

González-Pérez, S., Garcés-Claver, A., Mallor, C., Sáenz de Miera, L.E., Fayos, O., Pomar, F. *et al.* (2014) New insights into *Capsicum* spp relatedness and the diversification process of *Capsicum annuum* in Spain (analysis of a Capsicum collection from Spain). *PLoS ONE*, **9**(12), 1–23.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, **40**(D1), D1178–D1186.

H3Africa Consortium, Rotimi, C., Abayomi, A., Abimiku, A.L., Adabayeri, V.M., Adebamowo, C. *et al.* (2014). Research capacity. Enabling the genomic revolution in Africa. *Science*, **344**(6190), 1346–1348.

Hendre, P.S., Muthemba, S., Kariba, R., Muchugi, A., Fu, Y., Chang, Y. *et al.* (2019) African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops. *Planta*, **250**, 989–1003.

Hu, H., Shen, X., Liao, B., Luo, L., Xu, J. & Chen, S. (2019) Herbgenomics: A stepping stone for research into herbal medicine. *Science China Life Sciences*, **62**(7), 913–920.

Ince, A.G., Karaca, M. & Onus, A. (2010) Genetic relationships within and between *Capsicum* species. *Biochemical Genetics*, **48**(1–2), 83–95.

Jackson, S., Rounsley, S. & Purugganan, M. (2006) Comparative sequencing of plant genomes: choices to make. *The Plant Cell*, **18**(5), 1100–1104.

Jamnadass, R., Mumm, R.H., Hale, I., Hendre, P., Muchugi, A., Dawson, I.K. *et al.* (2020) Enhancing African orphan crops with genomics. *Nature Genetics*, **52**, 356–360.

Kamau, L.N., Mbaabu, M.P., Mbaria, J.M., Karuri, G.P. & Kiama, S.G. (2016) Knowledge and demand for medicinal plants used in the treatment and management of diabetes in Nyeri County, Kenya. *Journal of Ethnopharmacology*, **189**, 218–229.

Kamei, C.L.A., Severing, E.I., Dechesne, A., Furrer, H., Dolstra, O. & Trindade, L.M. (2016) Orphan crops browser: a bridge between model and orphan crops. *Molecular Breeding*, **36**, 9.

Kent, N., Robert, L., Carl, B., Susan, B. & Sarah, K. (2003) *African Perspectives on Genetic Materials*. Washington, DC: Environmental Law Institute. ISBN 1-58576-068-4.

Khayi, S., Azza, N.E., Gaboun, F., Pirro, S., Badad, O., Claros, M.G. *et al.* (2018) First draft genome assembly of the Argane tree (*Argania spinosa*) [version 1; peer review: 1 approved, 1 approved with reservations]. *F1000Research*, **7**, 1310.

Komen, J., Tripathi, L., Mkoko, B., Ofosu, D.O., Oloka, H. & Wangari, D. (2020) Biosafety regulatory reviews and leeway to operate: case studies from Sub-Sahara Africa. *Frontiers in Plant Science*, **2020**(11), 130.

Krishnamurthy, P., Fujisawa, Y., Takahashi, Y., Abe, H., Yamane, K., Mukaiyama, K. *et al.* (2019) High-throughput screening and characterization of a high-density soybean mutant library elucidate the biosynthesis pathway of triterpenoid saponins. *Plant and Cell Physiology*, **60**(5), 1082–1097.

Kumar, B. & Bhalothia, P. (2020) Orphan crops for future food security. *Journal of Biosciences*, **45**, 131.

Laura, R.B. (2017) Genome-wide diversity in the levant reveals recent structuring by culture. *Nature Communications*, **8**(1), 1–11.

Lawal, I.O., Olufade, I.I., Rafiu, B.O. & Aremu, A.O. (2020) Ethnobotanical survey of plants used for treating cough associated with respiratory conditions in Ede South Local Government Area of Osun State, Nigeria. *Plants*, **9**(5), 647.

Lee, H.Y., Ro, N.Y., Jeong, H.J., Kwon, J.K., Jo, J., Ha, Y. *et al.* (2016). Genetic diversity and population structure analysis to construct a core collection from a large *Capsicum germplasm*. *BMC Genetics*, **17**(1), 142.

Li, B., Cockerton, H.M., Johnson, A.W., Karlström, A., Stavridou, E., Deakin, G. *et al.* (2020) Defining strawberry shape uniformity using 3D imaging and genetic mapping. *Horticulture Research*, **7**, 115.

Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Linnemann, A. (1992). *Bambara Groundnut (Vigna subterranea) Literature. A Revised and Updated Bibliography*. Wageningen: Department of Tropical Crop Science, WAU.

Maass, B.L., Jamnadass, R.H., Hanson, J. & Pengelly, B.C. (2005) Determining sources of diversity in cultivated and wild Lablab purpureus related to provenance of germplasm by using amplified fragment length polymorphism. *Genetic Resources and Crop Evolution*, **52**(6), 683–695.

Maass, B., Knox, M. & Venkatesha, S. (2010) Lablab purpureus – A crop lost for Africa? *Tropical Plant Biology*, **3**(3), 123–135.

Mariod, A.A. & Abdelwahab, S.I. (2012) *Sclerocarya birrea* (Marula), An African tree of nutritional and medicinal uses: a review. *Food Reviews International*, **28**(4), 375–388.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A. *et al.* (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297–1303.

Michael, T.P. & VanBuren, R. (2019) Building near-complete plant genomes. *Current Opinion in Plant Biology*, **54**, 26–33.

Modak, M., Dixit, P., Londhe, J., Ghaskadbi, S. & Devasagayam, T.P. (2007) Devasagayam Indian herbs and herbal drugs used for the treatment of diabetes. *Journal of Clinical Biochemistry and Nutrition*, **40**, 163–173.

Moose, S.P. & Mumm, R.H. (2008) Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology*, **147**(3), 969–977.

Nhemachena, C., Matchaya, G., Nhemachena, C.R., Karuaihe, S., Muchara, B. & Nhlengethwa, S. (2018) Measuring baseline agriculture-related sustainable development goals index for Southern Africa. *Sustainability*, **10**, 849.

Nussbaumer, T., Martis, M.M., Roessner, S.K., Pfeifer, M., Bader, K.C., Sharma, S. *et al.* (2013) MIPS PlantsDB: a database framework for comparative plant genome research. *Nucleic Acids Research*, **41**(D1), D1144–D1151.

Ojo, V.O.A., Dele, P.A., Amole, T.A., Anele, U.Y., Adeoye, S.A., Hassan, O.A. *et al.* (2013) Effect of Intercropping *Panicum maximum* var. ntchisi and lablab purpureus on the growth, herbage yield and chemical composition of *Panicum maximum* var. Ntchisi at different harvesting times. Pakistan. *Journal of Biological Sciences*, **6**(22), 1605–1608.

Okot, D.F., Anywar, G., Namukobe, J. & Byamukama, R. (2020) Medicinal plants species used by herbalists in the treatment of snakebite envenomation in Uganda. *Tropical Medicine and Health*, **48**, 44.

Oladeji, O.S., Oluyori, A.P., Bankole, D.T. & Afolabi, T.Y. (2020) Natural products as sources of antimalarial drugs: ethnobotanical and ethnopharmacological studies. *Scientifica (Cairo)*, **2020**, 7076139.

Ong, Q., Nguyen, P., Thao, N.P. & Le, L. (2016) Bioinformatics approach in plant genomic research. *Current Genomics*, **17**(4), 368–378.

Panda, R.C., Kumar, O.A. & Rao, K. (2004) Cytogenetic studies of some F1 hybrids between wild and cultivated taxa of *Capsicum* L. *Cytologia*, **69**(2), 203–208.

Parry, M.A., Madgwick, P.J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M. *et al.* (2009) Mutation discovery for crop improvement. *Journal of Experimental Botany*, **60**(10), 2817–2825.

Patrick, V.T.F., Alexander, K.N., Regina, A.-O., Lauve, R.T.Y., Mark, O. & Fabrice, F.B. (2015) Update on medicinal plants with potency on *Mycobacterium ulcerans*. *BioMed Research International*, **2015**, 1–16. http://doi.org/10.1155/2015/917086.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y. *et al.* (2012) Ancient admixture in human history. *Genetics*, **192**, 1065–1093.

Pennisi, E. (2019) Plant genomics unearths Africa's 'fertile crescent. *Science*, **364**(6439), 422–423.

Peterson, W.W., Marie-Noelle, N. & Robert, H. (2019) Advances in molecular genetics and genomics of African Rice (*Oryza glaberrima* Steud). *Plants (Basel)*, **8**(10), 376.

Philp, J. (2018) The bioeconomy, the challenge of the century for policy makers. *New Biotechnology*, **40**(Pt A), 11–19.

Pirita, P., George, K., Elena, L.-G., Michael, G., Darren, H., David, B. *et al.* (2019) A critical comparison of technologies for a plant genome sequencing project. *GigaScience*, **8**(3), 1–12.

Plazas, M., Andújar, I., Vilanova, S., Gramazio, P., Herraiz, F.J. & Prohens, J. (2014) Conventional and phenomics characterization provides insight into the diversity and relationships of hypervariable scarlet (*Solanum aethiopicum* L.) and gboma (*S. macrocarpon* L.) eggplant complexes. *Frontiers in Plant Science*, **5**, 318.

Portères, R. (1976) *Origins of African Plant Domestication*. Harlan, J.R. (Ed.). Hawthorne, NY: Mouton Publishers, pp. 409–452.

Prakash, O., Kumar, R., Mishra, A. & Gupta, R. (2009) *Artocarpus heterophyllus* (Jackfruit): an overview. *Pharmacognosy Reviews*, **3**(6), 353–358.

Ragone, D. (2018) Breadfruit – *Artocarpus altilis* (Parkinson) Fosberg. In: Rodrigues, S., de Oliveira Silva, E. & Sousa de Brito, E. (Eds.) *Exotic Fruits*. Cambridge, MA: Academic Press, pp. 53–60.

Rashid, U., Anwar, F., Moser, B.R. & Knothe, G. (2008) *Moringa oleifera* oil: a possible source of biodiesel. *Bioresource Technology*, **99**(17), 8175–8179.

Rivas, R., Oliveira, M.T. & Santos, M.G. (2013) Three cycles of water deficit from seed to young plants of *Moringa oleifera* woody species improves stress tolerance. *Plant Physiology and Biochemistry*, **1**(63), 200–208.

Robotham, O. & Chapman, M. (2017) Population genetic analysis of hyacinth bean (*Lablab purpureus* (L.) Sweet, Leguminosae) indicates an East African origin and variation in drought tolerance. *Genetic Resources and Crop Evolution*, **64**(1), 139–148.

Rouard, M., Guignon, V., Aluome, C., Laporte, M.-A., Droc, G., Walde, C. *et al.* (2010) GreenPhylDB v2.0: comparative and functional genomics in plants. *Nucleic Acids Research*, **39**, D1095–D1102.

Schatz, M.C., Witkowski, J. & McCombie, W.R. (2012) Current challenges in de novo plant genome sequencing and assembly. *Genome Biology*, **13**, 243.

Scott, M.F., Botigué, L.R., Brace, S., Stevens, C.J., Mullin, V.E., Stevenson, A. *et al.* (2019) A 3,000-year-old Egyptian emmer wheat genome reveals dispersal and domestication history. *Nature Plants*, **5**(11), 1120–1128.

Siwela, M., Pillay, K., Govender, L., Lottering, S., Mudau, F.N., Modi, A.T. *et al.* (2020) Biofortified crops for combating hidden hunger in South Africa: availability, acceptability, micronutrient retention and bioavailability. *Foods*, **9**(6), 815.

Slik, J.W.F., Víctor, A.-R., Shin-Ichiro, A., Patricia, A.-L., Luciana, F.A., Peter, A. *et al.* (2015) An estimate of the number of tropical tree species. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(24), 7472–7477.

Solomon, A.M., Han, K., Lee, J.-H., Lee, H.-Y., Jang, S. & Kang, B.-C. (2019) Genetic diversity and population structure of Ethiopian *Capsicum* germplasms. *PLoS ONE*, **14**(5), e0216886.

Song, B., Song, Y., Fu, Y., Balyejusa, K.E., Ndagire, K.S., Nahamya, K.P. *et al.* (2019) Draft genome sequence of *Solanum aethiopicum* provides insights into disease resistance, drought tolerance, and the evolution of the genome. *GigaScience*, **8**(10), 1–16.

Spiegler, V. (2020) Anthelmintic A-type procyanidins and further characterization of the phenolic composition of a root extract from *Paullinia pinnata*. *Molecules*, **25**(10), 2287.

Su, H., Wang, L., Ge, Y., Feng, E., Sun, J. & Liu, L. (2008) Development of strain-specific SCAR markers for authentication of *Ganoderma lucidum*. *World Journal of Microbiology & Biotechnology*, **24**, 1223–1226.

Sunseri, F., Polignano, G.B., Alba, V., Lotti, C., Bisignano, V., Mennella, G. *et al.* (2010) Genetic diversity and characterization of African eggplant germplasm collection. *African Journal of Plant Science*, **2010**(4), 231–241.

Sweeney, M. & McCouch, S. (2007) The complex history of the domestication of rice. *Annals of Botany*, **100**, 951–957.

Tadele, Z. (2018) African Orphan crops under abiotic stresses: challenges and opportunities. *Scientifica (Cairo)*, **2018**, 1–19.

Tadele, Z. (2019) Orphan crops: their importance and the urgency of improvement. *Planta*, **250**(3), 677–694.

Tadele, Z. & Bartels, D. (2019) Promoting orphan crops research and development. *Planta*, **250**(3), 675–676.

Tamiru, M., Natsume, S., Takagi, H., White, B., Yaegashi, H., Shimizu, M. *et al.* (2017) Genome sequencing of the staple food crop white Guinea yam enables the development of a molecular marker for sex determination. *BMC Biology*, **15**(1), 86.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L. *et al.* (2005) Genome analysis of multiple pathogenic isolates of

Streptococcus agalactiae: implications for the microbial "pan-genome". *Proceedings of the National Academy of Sciences of the United States of America*, **102**(39), 13950–13955.

Tijani, A.Y., Uguru, M.O. & Salawu, O.A. (2008) Anti-pyretic, anti-inflammatory and anti-diarrhoeal properties of Faidherbia albida in rats. *African Journal of Biotechnology. Academic Journals (Kenya)*, **7**(6), 696–700. Available at: http://www.academicjournals.org/AJB.

Tiwari, M.K., Yadav, D.K. & Chaudhary, S. (2019) (2019) Recent developments in natural product inspired synthetic 1,2,4- trioxolanes (ozonides): an unusual entry into antimalarial chemotherapy. *Current Topics in Medicinal Chemistry*, **19**(10), 831–846.

Tokimatsu, T., Kotera, M., Goto, S. & Kanehisa, M. (2011) *Protein Function Prediction for Omics Era*. Springer. KEGG and GenomeNet resources for predicting protein function from omics data including KEGG PLANT resource. pp. 271–288.

Tranchant-Dubreuil, C., Rouard, M. & Sabot, F. (2019) Plant pangenome: Impacts on phenotypes and evolution. In: Roberts, J.A. (Ed.) *Annual Plant Reviews Online*. John Wiley & Sons, pp. 1–25. https://doi.org/10.1002/9781119312994.apr0664.

Tu, Y. (2011) The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nature Medicine*, **17**(10), 1217–1220. https://doi.org/10.1038/nm.2471.

Tuasha, N., Petros, B. & Asfaw, Z. (2018) Medicinal plants used by traditional healers to treat malignancies and other human ailments in Dalle District, Sidama Zone, Ethiopia. *Journal of Ethnobiology and Ethnomedicine*, **14**(1), 15.

Twilley, D., Rademan, S. & Lall, N. (2020) A review on traditionally used South African medicinal plants, their secondary metabolites and their potential development into anticancer agents. *Journal of Ethnopharmacology*, **261**, 113101.

Unver, T., Wu, Z., Sterck, L., Turktas, M., Lohaus, R., Li, Z. et al. (2017) Genome of wild olive and the evolution of oil biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, E9413–E9422.

Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A. & Van de Peer, Y. et al. (2018) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Research*, **46**(D1), D1190–D1196.

van Zonneveld, M., Ramirez, M., Williams, D.E., Petz, M., Meckelmann, S., Avila, T. et al. (2015) Screening genetic resources of Capsicum peppers in their primary center of diversity in Bolivia and Peru. *PLoS ONE*, **10**(9), 1–23.

Veselý, P., Šmarda, P., Bureš, P., Stirton, C., Muthama Muasya, A., Mucina, L. et al. (2020) Environmental pressures on stomatal size may drive plant genome size evolution: evidence from a natural experiment with cape geophytes. *Annals of Botany*, **126**(2), 323–330.

Vicente, M. & Schlebusch, C.M. (2020) African population history: an ancient DNA perspective. *Current Opinion in Genetics & Development*, **62**, 8–15.

Wang, M., Yu, Y., Haberer, G., Marri, P.R., Fan, C., Goicoechea, J.L. et al. (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nature Genetics*, **46**, 982–988.

World Health Organization (WHO). (2005) *National Policy on Traditional Medicine and Regulation of Herbal Medicines*. Geneva: Report of WHO Global Survey.