



# African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops

Prasad S. Hendre<sup>1</sup> · Samuel Muthemba<sup>1</sup> · Robert Kariba<sup>1</sup> · Alice Muchugi<sup>1</sup> · Yuan Fu<sup>2,3</sup> · Yue Chang<sup>2,3</sup> · Bo Song<sup>2,3</sup> · Huan Liu<sup>2,3</sup> · Min Liu<sup>2,3</sup> · Xuezhong Liao<sup>2,3</sup> · Sunil Kumar Sahu<sup>2,3</sup> · Sibow Wang<sup>2,3</sup> · Linzhou Li<sup>2,3</sup> · Haorong Lu<sup>2,3</sup> · Shufeng Peng<sup>2,3</sup> · Shifeng Cheng<sup>2,3</sup> · Xun Xu<sup>2,3</sup> · Huanming Yang<sup>2,3</sup> · Jian Wang<sup>2,3</sup> · Xin Liu<sup>2,3,4</sup> · Anthony Simons<sup>1</sup> · Howard-Yana Shapiro<sup>1,5</sup> · Rita H. Mumm<sup>6</sup> · Allen Van Deynze<sup>5</sup> · Ramni Jamnadas<sup>1</sup>

Received: 2 October 2018 / Accepted: 1 April 2019 / Published online: 9 May 2019  
© The Author(s) 2019

## Abstract

**Main conclusion** The African Orphan Crops Consortium (AOCC) successfully initiated the ambitious genome sequencing project of 101 African orphan crops/trees with 6 genomes sequenced, 6 near completion, and 20 currently in progress.

Addressing stunting, malnutrition, and hidden hunger through nutritious, economic, and resilient agri-food system is one of the major agricultural challenges of this century. As sub-Saharan Africa harbors a large portion of the severely malnourished population, the African Orphan Crops Consortium (AOCC) was established in 2011 with an aim to reduce stunting and malnutrition by providing nutritional security through improving locally adapted nutritious, but neglected, under-researched or orphan African food crops. Foods from these indigenous or naturalized crops and trees are rich in minerals, vitamins, and antioxidant, and are an integral part of the dietary portfolio and cultural, social, and economic milieu of African farmers. Through stakeholder consultations supported by the African Union, 101 African orphan and under-researched crop species were prioritized to mainstream into African agri-food systems. The AOCC, through a network of international–regional–public–private partnerships and collaborations, is generating genomic resources of three types, i.e., reference genome sequence, transcriptome sequence, and re-sequencing 100 accessions/species, using next-generation sequencing (NGS) technology. Furthermore, the University of California Davis African Plant Breeding Academy under the AOCC banner is training 150 lead African scientists to breed high yielding, nutritious, and climate-resilient (biotic and abiotic stress tolerant) crop varieties that meet African farmer and consumer needs. To date, one or more forms of sequence data have been produced for 60 crops. Reference genome sequences for six species have already been published, 6 are almost near completion, and 19 are in progress.

**Keywords** Orphan crops · Nutrition · Hunger · Genomics · Next-generation sequencing · Crop breeding

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00425-019-03156-9>) contains supplementary material, which is available to authorized users.

✉ Prasad S. Hendre  
P.Hendre@cgiar.org

<sup>1</sup> African Orphan Crops Consortium, World Agroforestry Centre (ICRAF), Nairobi, Kenya

<sup>2</sup> BGI-Shenzhen, Shenzhen 518083, China

<sup>3</sup> China National GeneBank, BGI-Shenzhen, Shenzhen 518120, China

<sup>4</sup> BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China

<sup>5</sup> University of California, 1 Shields Ave, Davis, CA 95616, USA

<sup>6</sup> University of Illinois, Urbana, IL 61801, USA

## Introduction

The world is facing an enormous challenge to feed a growing population with huge economic disparity and uneven development across the globe. Among all the continents, Africa is predicted to contribute to half of the world's population growth by 2050 (FAO 2017). Nigeria will be the third most populous country by 2050 and, along with the Democratic Republic of Congo, Ethiopia, Tanzania, and Uganda, will contribute to most of the world's population growth (United Nations 2017). At the same time, climate change scenarios also threaten current production levels with anticipated decline in the production of staple crops (Dinesh et al. 2015). All these factors together are going to aggravate the burgeoning problem of nutritional insufficiency in the African population. At present, among the top 36 countries with a high hidden hunger index (HHI, a combination of iron deficiency, zinc deficiency, and vitamin A deficiency), 31 are from sub-Saharan Africa (SSA) and they harbor 91% of the world's HHI-affected preschool children (Muthayya et al. 2013). Among the adults, every fourth African above the age of 15 years has suffered some form of hunger (Food and Agriculture Organization 2017). The adulthood diseases called as non-communicable diseases (NCDs), the majority of which are exacerbated by malnutrition, will be the major cause of deaths in SSA by 2030 (Food and Agriculture Organization 2017). Most of these countries are ill-prepared to tackle these problems, as 33 of the 44 resource-poor, least developed countries are in the African continent (Food and Agriculture Organization 2017). Thus, the economic disparity, which is closely linked with development and health indices, has a huge impact on the present and future human capital, which is a prime driver of a country's growth and development. These problems should be urgently tackled to fulfill the ambitious sustainable development goals (SDG) of 'zero hunger' (SDG 2) and 'good health and well-being' (SDG 3) by 2030, as agreed by the United Nations General Assembly (United Nations 2015). It is a well-accepted fact that access to good quality, diverse, healthy, and nutritious food has an ability to immensely improve the quality of life at all the stages (Micronutrient Initiative 2009).

The Green Revolution that substantially increased agricultural production in the last century was mainly based on improved grain yield and focused primarily on wheat and rice. While effective in saving lives, it also resulted in monoculture, decreased dietary diversity, and destruction of natural habitat/forest, (Pingali 2012). On similar lines, the African countries have traditionally been positively over-incentivizing production and consumption of maize, which has reached a point where there is an unbalanced

ratio of calorie intake and other dietary components such as minerals, vitamins, and proteins (Madzivhandila et al. 2016). To overcome the devastating effects of hunger as well as malnutrition and hidden hunger, solutions beyond field-based productivity measurements and calorie sufficiency must be achieved. Traditional African crops can provide the balance needed in African agri-food system through a systems and landscape approach, which is productive, as well as nutrient dense, diverse, resilient to climate change, and culturally linked with food habits of the local communities.

## Africa and the African Orphan Crops Consortium (AOCC)

Africa is bestowed with immense diversity of traditional food crops which are being consumed by local communities since time immemorial. These crops are distributed discretely across the farming landscapes, but most often with low yields due to low or almost no investment (monetary, human resource, organizational, policy, government, crop improvement, etc.) into research, development, marketing, and any other portfolio. Thus, these neglected, under-researched, "orphan" crops have not benefitted from modern scientific and technological advancements. Nevertheless, due to the potential they hold to address food, nutritional, and economic security of the developing and undeveloped parts of the world, they are appropriately referred to as crops for the future (Baldermann et al. 2016). Most African orphan crops have been consumed by local communities without extensive selection and domestication, thereby maintaining genetic diversity for stress and disease tolerance, adaptability, and production traits as well as nutrition. They are rich in micro- and macronutrients and can adapt to changing climatic conditions as well as diseases and pests (Baldermann et al. 2016). Thus, these crops have immense scope for innovation in research, capacity building, social empowerment, and food value chains (i.e., production, processing, consumption, marketing, and product development), but need a nonconventional approach which is open, inclusive, and welcoming to involvement and investment from public, private, national, and international partners. Understanding this as an important African agenda, the African Orphan Crop Consortium (AOCC) was established with the full support of the African Union in 2011, mandated to work on 101 selected crops originated or naturalized in Africa (<http://africanorphancrops.org>) by investing in training, products, tools, services, practices, and processes to mainstream them into the African agri-food system.

## Selection of AOCC crops

The list of 101 species of African orphan crops (<http://africanorphanocrops.org/meet-the-crops>) important to African agriculture and agroforestry systems was drawn from an Africa-centric survey conducted by African Union's (AU) New Partnership for African Development (NEPAD) with participation from African agricultural scientists, sociologists, anthropologists, nutritionists, policy makers, farmers, government representatives, universities, and various other stakeholders (Hendre and Van Deynze 2015). Three primary selection criteria used for prioritization were (i) being rich in micro- and macronutrient contents, (ii) relevant to Africa, and (iii) need for developing breeding resources. The list was judiciously spread across woody trees/shrubs (50), non-woody annuals (33), non-woody fruits (3), one palm, two each of woody climber fruits, perennial vegetables, perennial rhizomes, and one each from a perennial root, non-tree fruit, and perennial succulent species. It comprises 28 orders and 45 families. Table 1 gives information about the stage of resource generation for 60 species and areas in Africa, where they are grown or found. In addition, Supplementary Table 1 lists the nutritional importance of some of the AOCC trees and crops. The list of crops is generally fixed, but is open to evolve with time and need.

## Application of genomics

Genomics technology has been positively contributing to producing quality seed stocks with desirable traits by providing tools and methods to have varieties with high productivity, high nutrients, and tolerance to various biotic and abiotic stresses (Kole et al. 2015). Genomic resources for a crop include the genome sequence, annotation of gene functions and extant diversity in the gene pool. All these together help in developing single nucleotide polymorphism (SNP) and other variant panels, which can be used to associate and introgress traits of economic and agronomic importance. Until now, 236 plant genomes have been sequenced using next-generation sequencing technologies (Chen et al. 2018) of varying quality. It can also be seen that some of the crops, which were considered as orphan crops in the last century, have now entered the next era of breeding and improvement after their genomes were sequenced and molecular breeding tools have been adopted for many of them, e.g., sorghum (Paterson et al. 2009; Fernandes et al. 2018), pearl millet (Varshney et al. 2017; Liang et al. 2018), peanut (Bertioli et al. 2016; Varshney 2016; Janila et al. 2016), chickpea (Varshney et al. 2013, Li et al. 2018), foxtail millet (Zhang et al. 2012), tef (Cannarozzi et al. 2014), finger millet (Hittalmani et al. 2017), cowpea (Boukar et al. 2018), bitter melon (Urasaki et al. 2017), and cucurbits (Zheng et al. 2019). The AOCC-mandated genomes, if already sequenced

and published, will be assessed for their quality and if needed will be corroborated using complementary sequencing technologies.

A need to attract funds by building investment cases to develop research and development programs for these orphan crops was aptly explained by Dawson et al. (2018). This report emphasized the fact that those crops which received some investments in their research and development programs had an increased yield. Thus, inclusion of genomics tools in designing breeding programs such as quantitative trait mapping, genome-wide association mapping (GWAS), and genomic selection (GS) has a huge potential to contribute by reducing the length of varietal improvement program up to a third of the existing traditional pipelines, which do not use trait–marker associations (Hickey et al. 2017).

With the publication of the first set of five genomes (Chang et al. 2018a, b, c, d, e, f), we have published the standard data generation, analysis, and annotation pipeline which is being used by AOCC in the sequencing projects. But as explained elsewhere, these pipelines will evolve to suit technological and analytical updates. As some genomes are complex and with subtle or major genomic duplications, they pose significant challenges for genome assembly and calls for technological upgradation to long-read sequencing technologies, which the consortium has already taken a note of. A general roadmap to breed orphan crops using genomic tools is also well laid out by Sogbohossou et al. (2018) using *Gynandropsis gynandra* (cleome) as an example. This signifies and justifies the initial investment necessary to generate genomic resources to aid in modernizing breeding programs. Using a similar model to develop improved varieties/clones but encouraging diversified farming landscape will limit some drawbacks associated with green revolution technologies such as monoculture, high input, and high emission practices. Publication of the first set of five orphan crop genomes was considered as an important step to boost research and improvement for these species (Tena 2019).

## Partnership

The consortium is an uncommon partnership between public, private, academic institutes, universities, and non-government, international, and development organizations (Hendre and Van Deynze 2015). The consortium has a healthy mix of partners who bring firsthand scientific knowledge (BGI, Shenzhen, China; University of California, Davis, USA (UCD); World Agroforestry (ICRAF), Nairobi, Kenya; Agriculture Research Council of South Africa (ARC), Pretoria, South Africa; Wageningen University and Research, Wageningen, Netherlands; Ghent University, Ghent, Belgium; CyVerse, Tucson, USA); offer analytical and computational capacity (BGI, Shenzhen, China; Benson Hill Biosystems, St. Louis, USA;

**Table 1** Progress of African Orphan Crop Consortium (AOCC) for three major activities in genomics workflow

| Sl# | Species name (growth habit/type of crop) | Common name and distribution in Africa  | Whole genome sequencing status (WGS) | Transcriptome sequencing status | Re-sequencing status |
|-----|--|---|--------------------------------------|---------------------------------|----------------------|
| 1   | <i>Faidherbia albida</i> (T-Fr)          | Apple ring acacia, distributed throughout the dry zones of Africa, predominantly in the east and southern Africa (Orwa et al. 2009), is an important agroforestry species for land restoration (PROTA 2018)                                       | Published, Chang et al. (2018a, b)   | RNA sequenced                   | In the pipeline      |
| 2   | <i>Vigna subterranea</i> (A-OS)          | Bambara groundnut originated in western Africa and grown for seeds in sub-Saharan semi-arid tropical regions of Nigeria, Cameroon and eastern Africa (Tanzania, Uganda) (PROTA 2018)  | Published, Chang et al. (2018a, c)   | RNA sequenced                   | Partially done       |
| 3   | <i>Lablab purpureus</i> (A-LS)           | Hyacinth bean is consumed as a vegetable in tropical Africa (Nigeria and Kenya, Rwanda, and Ethiopia) (PROTA 2018)  | Published, Chang et al. (2018a, d)   | RNA sequenced                   | Partially done       |
| 4   | <i>Sclerocarya birrea</i> (T-Fr)         | Marula tree occurs throughout sub-Saharan Africa except the humid zones. A wine called Amarula is commercially made from fruits in southern Africa (PROTA 2018)   | Published, Chang et al. (2018a, e)   | RNA sequenced                   | In the pipeline      |
| 5   | <i>Moringa oleifera</i> (T-LFV)          | Drumstick, Indian in origin but now naturalized in Africa and found throughout tropics and subtropics of Africa (PROTA 2018), is used as a leafy and fruit vegetable  | Published, Chang et al. (2018a, f)   | RNA sequenced                   | Partially done       |
| 6   | <i>Solanum aethiopicum</i> (A-LV)        | African eggplant is found across the tropical sub-Saharan Africa, especially western, central, and some parts of eastern Africa (Uganda) Schippers (2000; Maundu et al. 2009), consumed as a leafy vegetable                                      | Published, Song et al. (2019)        | RNA sequenced                   | Yet to begin         |
| 7   | <i>Digitaria exilis</i> (A-G)            | Fonio, consumed in the form of grains, is cultivated in western Africa since ancient time and now is grown scattered across the same region (PROTA 2018)  | Final assembly                       | Yet to begin                    | Yet to begin         |
| 8   | <i>Eleusine coracana</i> (A-G)           | Finger millet, an African-origin cereal, is found to be growing in all the semi-arid and arid tropics of Africa in eastern, western, and southern regions (USDA 2019). Grains are consumed in the form of flour or fermented to make edible drink | Final assembly                       | RNA sequenced                   | Partially done       |
| 9   | <i>Gynandropsis gynandra</i> (A-LV)      | Spider plant is generally found distributed all over the Africa near agricultural land and human settlements, but grown in drier zones of southern Africa in South Africa, Zimbabwe etc. (PROTA 2018), and is consumed as a leafy vegetable       | Final assembly                       | In the pipeline                 | Partially done       |

**Table 1** (continued)

| Sl# | Species name (growth habit/type of crop) | Common name and distribution in Africa   | Whole genome sequencing status (WGS) | Transcriptome sequencing status | Re-sequencing status |
|-----|--|--|--------------------------------------|---------------------------------|----------------------|
| 10  | <i>Amona cherimola</i> (T-Fr)            | Cherimoya, a fruit tree, also referred to as custard apple is grown in tropical Africa specifically in western Africa (Algeria, Egypt, Libya, Somalia, and South Africa) (USDA 2019)   | Final assembly                       | In the pipeline                 | Yet to begin         |
| 11  | <i>Artocarpus heterophyllus</i> (T-Fr)   | Jackfruit is originally from Southeast Asia (India), and now naturalized in Africa and found to be growing in western and south-western Africa (Orwa et al. 2009)  | Final assembly                       | In the pipeline                 | Yet to begin         |
| 12  | <i>Artocarpus altilis</i> (T-Fr)         | Bread fruit is a naturalized widely found across Africa, mainly in tropical sub-Saharan Africa (CABI 2018)   | Final assembly                       | In the pipeline                 | Yet to begin         |
| 13  | <i>Celosia argentea</i> (A-LV)           | Cockscomb is used as a leafy vegetable and medicinal plant in Kenya, Congo, Benin, and Ethiopia, grown on scattered plots in home gardens and priced during dry seasons (PROTA 2018)   | Data generation on                   | In the pipeline                 | Yet to begin         |
| 14  | <i>Phaseolus vulgaris</i> (A-LV, FV)     | Common bean is widely grown all over Africa with the main countries being Burundi, Kenya, Rwanda, DR Congo, Tanzania, Uganda, and Sudan (Wortmann et al. 1998). The green pods and young or dried seeds are consumed as food                                   | Data generation on                   | Yet to begin                    | In the pipeline      |
| 15  | <i>Solanum nigrum</i> (A-LV)             | Black or African nightshade is consumed as fruits or leaves. There is no exact information about its distribution (PROTA 2018), but found in tropical sub-Saharan regions of Africa (USDA 2019)  | Data generation on                   | In the pipeline                 | Yet to begin         |
| 16  | <i>Adansonia digitata</i> (T-LV, F)      | Baobab occurs naturally all over the sub-Saharan Africa. Fruits and leaves are consumed in various parts of Africa. Except for a few countries in central Africa (Rwanda, Burundi, Djibouti, and Uganda), it is found all over sub-Saharan Africa (PROTA 2018) | Data generation on                   | Tissue collected                | In the pipeline      |
| 17  | <i>Anacardium occidentale</i> (T-Fr)     | Cashew nut is found and grown in coastal and humid countries in eastern Africa (Kenya, Mozambique, Gambia, Sudan, Tanzania, Uganda) (Orwa et al. 2009)   | Data generation on                   | In the pipeline                 | Yet to begin         |
| 18  | <i>Doryyalis caffra</i> (T-Fr)           | Kei apple or wild apricot is a fruit tree found in southern Africa (South Africa, Botswana, Namibia, Lesotho, Malawi, Mozambique, South Africa, Swaziland, Zimbabwe) and introduced in Ethiopia, Kenya, Sudan, Tanzania, and Uganda (Orwa et al. 2009)         | Data generation on                   | In the pipeline                 | Yet to begin         |

Table 1 (continued)

| Sl# | Species name (growth habit/type of crop) | Common name and distribution in Africa   | Whole genome sequencing status (WGS) | Transcriptome sequencing status | Re-sequencing status |
|-----|--|--|--------------------------------------|---------------------------------|----------------------|
| 19  | <i>Parinari curatellifolia</i> (T-Fr)    | Mboba plum, found in almost all the sub-Saharan Africa (Orwa et al. 2009), is grown for fruits   | Data generation on                   | Yet to begin                    | In the pipeline      |
| 20  | <i>Parkia biglobosa</i> (T-Fr)           | Monkey cutlass tree or African locust bean produces edible fruits and seeds the tree is a native of western Africa and found distributed across central and eastern Africa (Orwa et al. 2009; PROTA 2018)  | Data generation on                   | In the pipeline                 | Yet to begin         |
| 21  | <i>Saba senegalensis</i> (WCL-Fr)        | Saba produces edible fruits and is found in western Africa and eastern-most Africa (Orwa et al. 2009)  | Data generation on                   | Yet to begin                    | Yet to begin         |
| 22  | <i>Uapaca kirkiana</i> (T-Fr)            | Wild loquat is a native fruit species of southern Africa (Angola, DR Congo, Malawi, Mozambique, Tanzania, Zambia, and Zimbabwe (Orwa et al. 2009))   | Data generation on                   | In the pipeline                 | In the pipeline      |
| 23  | <i>Vitellaria paradoxa</i> (T-Fr, OS)    | Shea butter tree is a native of western Africa, but found across sub-Saharan belt till Sudan and Ethiopia (Orwa et al. 2009). Kernels, seed oil, and pulp are consumed by locals and seed oil is widely used in cosmetic and confectionary industry                  | Data generation on                   | In the pipeline                 | In the pipeline      |
| 24  | <i>Casimiroa edulis</i> (T-Fr)           | White sapote is a fruit tree widely found in Tanzania and South Africa (Orwa et al. 2009)  | Data generation on                   | In the pipeline                 | Yet to begin         |
| 25  | <i>Dacryodes edulis</i> (T-Fr)           | African plum or Safou is a native fruit species of western Africa, but also distributed in central Africa (Orwa et al. 2009). Fruits are eaten raw or after boiling or roasting  | Data generation on                   | In the pipeline                 | Yet to begin         |
| 26  | <i>Ziziphus mauritiana</i> (T-Fr)        | Jujube is a fruit species naturalized in Africa which is grown in western, central, and eastern Africa (Orwa et al. 2009)  | Data generation on                   | In the pipeline                 | Yet to begin         |
| 27  | <i>Tamarindus indica</i> (T-Fr)          | Tamarind, found across sub-Saharan western, central, and eastern Africa (Orwa et al. 2009), produces edible sour fruit which is used as a part of culinary preparations or consumed as juice   | Data generation on                   | Tissue collected                | Yet to begin         |
| 28  | <i>Syzgium guineense</i> (T)             | Water berry is a fruit tree found in its native range of eastern and southern Africa (Orwa et al. 2009)  | Data generation on                   | Tissue collected                | Yet to begin         |
| 29  | <i>Irvingia gabonensis</i> (T)           | African bush mango is a native fruit species of western Africa and is found distributed across central and eastern sub-Saharan Africa till Uganda and Sudan (Orwa et al. 2009). Fruits can be eaten raw, but some tree parts are also used in the local dye industry | Data generation on                   | Yet to begin                    | Yet to begin         |

**Table 1** (continued)

| Sl# | Species name (growth habit/type of crop)   | Common name and distribution in Africa  | Whole genome sequencing status (WGS) | Transcriptome sequencing status | Re-sequencing status |
|-----|--|---|--------------------------------------|---------------------------------|----------------------|
| 30  | <i>Ricinodendron heudelotii</i> (T-Fr, Sp) | Groundnut tree is a native tree of western Africa and found distributed across central and eastern sub-Saharan Africa (Orwa et al. 2009). The fruits are edible and after drying are used a spice   | Data generation on                   | Tissue collected                | Yet to begin         |
| 31  | <i>Detarium microcarpum</i> (T-Fr)         | Sweet datock occurs in the drier regions of west and central Africa (PROTA 2018). The fruit can be eaten raw or fruit pulp can be dried and stored  | Data generation on                   | Yet to begin                    | Yet to begin         |
| 32  | <i>Garcinia livingstonei</i> (T-Fr)        | Wild mangosteen is a fruit tree found across southern and eastern Africa (Orwa et al. 2009)   | Data generation on                   | Yet to begin                    | Yet to begin         |
| 33  | <i>Dioscorea alata</i> (Tdd) (PT-T)        | Water yam is an introduced tuberous root vegetable crop naturalized in Africa. It is primarily grown predominantly in western Africa (Obidiegwu et al. 2009)  | Data generation on                   | By collaborator                 | In the pipeline      |
| 34  | <i>Macropylioma geocarpum</i> (A-PS)       | Kersting's groundnut is a pulse crop with similar growth habit as peanut. The seeds are consumed after cooking or roasting and can be made into flour for various uses. It is suggested to have originated from Western Africa and grown as a minor crop in West African Savanna zone, Cameroon, Tchad, and also reported to be grown in Mauritius, Tanzania, and Fizi (Ayenan and Ezin 2016) | Yet to begin                         | Yet to begin                    | In the pipeline      |
| 35  | <i>Strychnos cocculoides</i> (T-Fr)        | Monkey orange, a native fruit tree of southern Africa, is also found in Botswana, Kenya, South Africa, Tanzania, Uganda, Zambia, and Zimbabwe (Orwa et al. 2009)  | Yet to begin                         | In the pipeline                 | In the pipeline      |
| 36  | <i>Strychnos spinosa</i> (T-Fr)            | Kaffir orange is mainly found in eastern Africa (Kenya Ethiopia, Sudan, Tanzania, Uganda) and Mali, Zambia, and Madagascar (Orwa et al. 2009), which is eaten as raw fruit  | Yet to begin                         | Tissue collected                | Yet to begin         |
| 37  | <i>Vitex doniana</i> (T-Fr)                | Black plum is predominantly distributed across east and south Africa (Orwa et al. 2009). The fruits are consumed raw or after boiling   | Yet to begin                         | Tissue collected                | Yet to begin         |
| 38  | <i>Allanblackia stuhlmannii</i> (T-Fr, OS) | Allanblackia is a native tree of Tanzania and found only in the humid tropical regions of the country (Orwa et al. 2009). The edible seed oil is in high demand locally and internationally and is used in cosmetics and confectionary industry   | Yet to begin                         | Tissue collected                | Yet to begin         |

Table 1 (continued)

| Sl# | Species name (growth habit/type of crop) | Common name and distribution in Africa   | Whole genome sequencing status (WGS)  | Transcriptome sequencing status | Re-sequencing status |
|-----|--|--|---|---------------------------------|----------------------|
| 39  | <i>Amona reticulata</i> (T-Fr)           | Bullock's heart, a naturalized fruit species in Africa from central America, is predominantly grown in South Africa and sporadically in other tropical African countries (Orwa et al. 2009)                                    | Yet to begin  | Tissue collected                | Yet to begin         |
| 40  | <i>Vangueria madagascariensis</i> (T-Fr) | Wild medlar, a fruit tree, is found in east, southern, and southwest Africa (Orwa et al. 2009)   | Yet to begin  | In the pipeline                 | Yet to begin         |
| 41  | <i>Passiflora edulis</i> (WCL-Fr)        | Passion fruit is naturalized fruit species found in sub-Saharan Africa and in grown in eastern and southern African countries (USDA 2019)  | Yet to begin  | In the pipeline                 | Yet to begin         |
| 42  | <i>Brassica carinata</i> (A-LV)          | Ethiopian kale is cultivated in Ethiopia since thousands of years, and also grown in east and southern Africa to some extent (PROTA 2018)  | By collaborator   |                                 | In the pipeline      |
| 44  | <i>Dioscorea rotundata</i> (PT-T)        | White yam probably originated in west Africa, where it is extensively cultivated (Aighevi et al. 2015)   | By collaborator   |                                 | In the pipeline      |
| 45  | <i>Lens culinaris</i> (A-LS)             | Lentils are grown in tropical parts of Africa, predominantly in eastern and southern Africa (PROTA 2018)   | By collaborator   |                                 | In the pipeline      |
| 46  | <i>Musa acuminata</i> (NWF-Fr)           | <i>Musa acuminata</i> is sequenced as one of the parents of widely cultivated triploid banana.   | Other group   |                                 | In the pipeline      |
| 47  | <i>Allium cepa</i> (A-RV)                | Onion is an introduced root (bulb) vegetable species to Africa and is grown across tropical west, east, and southern Africa (PROTA 2018)   | By collaborator   |                                 | Yet to begin         |
| 48  | <i>Amaranthus tricolor</i> (A-G)         | Amaranth species consumed as leafy vegetable, is introduced from tropical Asia, but is now naturalized in Africa and grown in west, east and southern Africa (PROTA 2018)  | By collaborator   |                                 | Yet to begin         |
| 49  | <i>Carica papaya</i> (P-Fr)              | Called as Papaya or paw paw in Africa, it is grown for its fruit in east and southern Africa as fruit crop (da Silva et al. 2007)  | By collaborator   |                                 | Yet to begin         |
| 50  | <i>Citrullus lanatus</i> (A-Fr; OS)      | Referred to as egusi melon or west African melon, the fruit species originated from western Kalahari region of Namibia and Botswana. Apart from southern Africa, now it is also found across Mediterranean Africa (PROTA 2018) | By collaborator   |                                 | Yet to begin         |
| 51  | <i>Cucurbita maxima</i> (A-FV)           | Pumpkin has been reported from many countries in tropical Africa and probably occurs in all the countries (PROTA 2018), where it is grown as a fruit vegetable   | Done by CucGenDB; <a href="http://cucurbitgenomics.org">http://cucurbitgenomics.org</a> (Zheng et al. 2019) |                                 | Yet to begin         |



**Table 1** (continued)

| Sl# | Species name (growth habit/type of crop)                   | Common name and distribution in Africa   | Whole genome sequencing status (WGS) | Transcriptome sequencing status | Re-sequencing status |
|-----|--|--|--------------------------------------|---------------------------------|----------------------|
| 52  | <i>Eragrostis tef</i> (A-G)                                | Tef originated in northern Ethiopia. It is cultivated for its grains and cultivation is confined mainly to Eritrea and Ethiopia (PROTA 2018)   | By collaborator                      |                                 | Yet to begin         |
| 53  | <i>Ipomoea batatas</i> (PR-RV)                             | Sweet potato is an introduced root tuber species to Africa, but now it is grown extensively in all the tropical Africa primarily in eastern and western Africa (Glato et al. 2017)                         | By collaborator                      |                                 | Yet to begin         |
| 54  | <i>Musa acuminata</i> AAA Group ( <i>Matoke</i> ) (NWF-Fr) | East African highland bananas or matoke originated in African great lakes region, is eaten as fruit vegetable in Uganda, Tanzania, Burundi, Democratic Republic of Congo, Kenya, and Rwanda (PROMUSA 2017) | By collaborator                      |                                 | Yet to begin         |
| 55  | <i>Momordica charantia</i> (A-FV)                          | Bitter melon, a fruit vegetable probably domesticated in India and southern China, occurs almost throughout tropical Africa and occasionally cultivated in east Africa (PROTA 2018)                        | By collaborator                      |                                 | Yet to begin         |
| 56  | <i>Mangifera indica</i> (T-Fr)                             | Mango is a naturalized fruit tree in Africa and is grown in tropical parts of Western, and eastern Africa (Orwa et al. 2009)   | By collaborator                      |                                 | Yet to begin         |
| 57  | <i>Macadamia ternifolia</i> (T-Fr)                         | Macadamia nut is an introduced nut tree species grown in Ethiopia, Kenya, Malawi, South Africa, Tanzania, Zimbabwe (Orwa et al. 2009)  | By collaborator                      |                                 | Yet to begin         |
| 58  | <i>Persea americana</i> (T-Fr)                             | Avocado is an introduced fruit tree in Africa, which is grown almost all over sub-Saharan Africa (Orwa et al. 2009)  | By collaborator                      |                                 | Yet to begin         |
| 59  | <i>Psidium guajava</i> (T-Fr)                              | Guava, a naturalized fruit tree grown in Ethiopia, South Africa, Uganda, Togo, Senegal, and Nigeria (Orwa et al. 2009)   | Other group                          |                                 | Yet to begin         |
| 60  | <i>Morus alba</i> (T-LV, FV)                               | Mulberry is grown in eastern and southern Africa (Orwa et al. 2009), where the fruit is consumed fresh or as juice   | Other group                          |                                 | Yet to begin         |

Growth habit depicted as: A annual, T tree, PT perennial tuber, WCL woody climber or liana, NWF non-woody fruit, PR perennial root. Economic part depicted as: Fr fruit, FV fruit vegetable, LV leafy vegetable, RV root vegetable, G grains, T tuber, OS oil seed, Sp spice, PS pulse seed. As soon as the DNA/RNA enters into the specific activity, it gets logged as “in the pipeline”. For 18 species, the WGS and transcriptome sequencing is being done either by one of the AOCC’s collaborators or other groups. The information about “Other group” and “By collaborator” is still being worked out and thus is not final

CyVerse, Tucson, USA; Ghent University, Ghent, Belgium); provide sequencing, genotyping and other ancillary technologies, supplies, data logistics, and infrastructure (Illumina Inc., San Diego, USA; LGC Genomics, Hoddesdon, UK; Google Inc., Mountain View, USA; Ghent University, Ghent, Belgium; KeyGene Inc., Rockville, USA; Oxford Nanopore Technologies, London, UK; Thermo Fisher Scientific, Waltham, USA); share experience of industrial food processing, technological up-scaling, marketing, and building seed value chains (Mars Incorporated., McLean, USA; Corteva Agriscience, Johnston, USA; Benson Hill Biosystems, St. Louis, USA); work on African development questions, advocacy, and policy framework (World Wildlife Fund (WWF), Gland, Switzerland; New Partnership for African Development (NEPAD), Midrand, South Africa; African Alliance for a Green Revolution for Africa (AGRA), Nairobi, Kenya; Food and Agriculture Organization (FAO), Rome, Italy; United Nation's International Children's Emergency Fund (UNICEF), Nairobi, Kenya); build training and capacity; provide funding support; and include ground-level organizations working on agricultural questions, germplasm repositories/gene banks, crop/tree improvement and breeding, socioeconomic context, and translational research (UC Davis African Plant Breeding Academy (AfPBA); Bioscience eastern central Africa-International Livestock Research Institute (BecA-ILRI), Nairobi, Kenya; Mars Incorporated, ICRAF, UNICEF, Integrated Breeding Platform, El Batan, Mexico). The partners and their contribution to the consortium are summarized in Supplementary Fig. 1 and described in Supplementary Table 2. The AOCC has grown from 5 founding members, NEPAD, WWF, Mars Incorporated, ICRAF, and UCD, to 24 core members involved in strategizing and developing AOCC's roadmap and they bring onboard a substantial and broad-based expertise and complementing skill sets cutting across all the crops. The sequencing work is coordinated from the genomics laboratory of AOCC located at ICRAF, Nairobi, Kenya. Initially, four Ion Protons donated by Thermo Fisher Scientific in 2014 were used for re-sequencing, which was subsequently upgraded in 2017 to HiSeq4000 donated by Illumina Inc. An open network of more than 25 institutes and organizations primarily handling crop-specific portfolios also supports ground activities related with field experimentation, crop breeding, and translational research (<http://africanorphancrops.org/partners-and-networks/>). Most importantly, the AOCC is currently made up of over 116 African plant breeders from 28 countries, who are trained in the AFPBA and leaders in their institutions.

## Timelines

The timelines to achieve the sequence targets are dependent largely on funding support, technology upgrades, building national crop breeders' networks, successful advocacy, and policy initiatives. The whole genome sequencing is anticipated to be completed by 2020 and re-sequencing by 2022.

## Major constraints

Working on orphan crops in the context of Africa poses significant challenges such as availability of reliable cultivation and production data, nutritional quality data, information on breeding technologies, and data on local and regional markets and value chains (Dawson et al. 2018). The AOCC will gather this information by working with local and regional partners and by encouraging the administrative bodies to collect these data through advocacy and policy measures. Apart from these important considerations, funding support is a major decisive factor. The plant breeders trained under the AfPBA as well as other collaborators, and networking partners are encouraged to support sequencing and re-sequencing activities through bilateral projects. Efforts are always on to garner support from national, international, and private funding bodies.

## Important achievements of the consortium

The consortium has been successful in gathering world class expertise from diverse and unusual partners and collaborators across the research, science, technology, social, commercial, and development sectors of the agri-food system under a single net. A genomics laboratory was established at ICRAF, which also acts as a secretariat of the consortium. This laboratory hosts a HiSeq 4000 sequencing system donated by Illumina Inc. and all the necessary instrumentation and workflows. Shared responsibilities of all the partners are listed in Supplementary Table 2 and described in Supplementary Fig. 1. Various partners are involved in building investment cases, budget, and projections of deliverables with timelines.

## Present status and progress

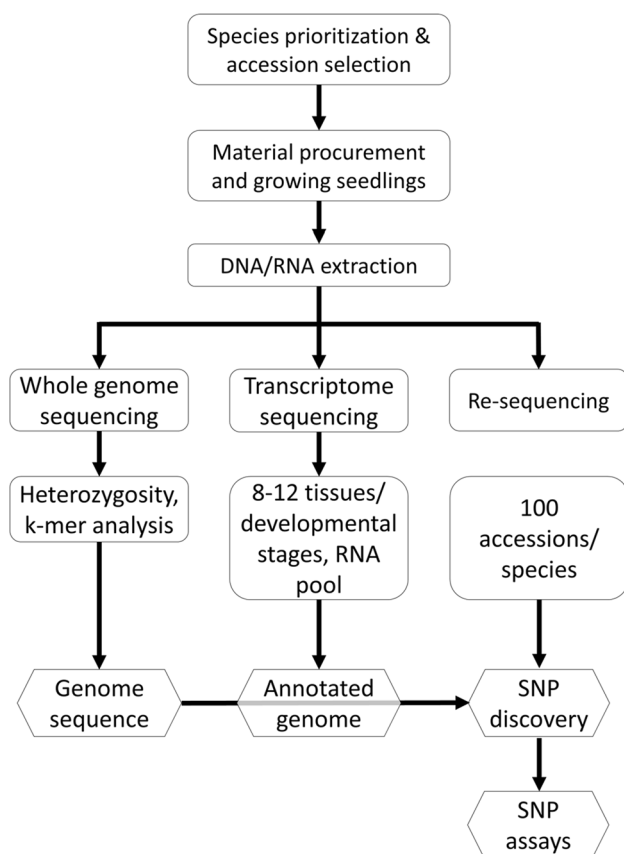
As an immediate goal, AOCC started developing genomics resources—reference genome sequencing transcriptome sequencing and re-sequencing of 100 accessions/species for these 101 crops. Extraction of nucleic acids posed peculiar challenges, as many of the species contain large quantities of mucilage, phenolics, and secondary metabolites and the processes were modified to suit the quality requirements. DNA extraction primarily used extensively modified manual

CTAB-based methods, whereas RNA extraction was done using any suitable kit with certain modifications if required. Reference genome sequencing is attempted using short-read NGS technology from Illumina, and BGI-Seq, but now has been expanded to long-read and scaffolding technologies. Transcriptome sequencing is carried out using Illumina's short-read technology, but now also extended to Oxford NanoPore. The re-sequencing was initially done over Ion Proton machines (generously donated by the then Life Technologies, Carlsbad, CA, USA now Thermo Fisher Scientific Inc.), but now shifted to Illumina's short-read NGS technology on HiSeq 4000 (generously donated by Illumina Inc.). Standard protocols recommended by the respective supplier are used for all these activities unless some tweaking is required to suit genome complexity or any unexplained problems.

The workflow of the three activities, provided in Fig. 1, runs in parallel to each other. The first step was species prioritization and is explained above in "Selection of AOCC crops". Material procurement involved either getting DNA,

clones, or seeds from the respective collaborator(s) and growing them in a nursery. The accessions used for reference genome sequencing were selected by the researchers/breeders working in the respective crops/trees to be physically accessible with minimal passport data. The first step for whole genome sequencing (WGS) is heterozygosity survey and k-mer analysis over a shallow sequencing data (5–10 Gb). This analysis helps to understand genome complexity to decide the course of sequencing and analysis (amount of data needed, analysis pipeline, etc.). The second activity is generation of transcriptome data by sequencing an RNA pool of 8–12 tissues/developmental stages. The third activity of re-sequencing is done by sequencing 8–15 times depth/accession for 100 selected accessions/species. The panel for re-sequencing is selected to represent genetic, trait, and/or geographic diversity present within Africa or diversity important for trait improvement. The selection of material is primarily done by the breeders or groups working on these crops with inputs from the AOCC team. Standard processes were followed for genome assembly, transcriptome assembly (Chang et al. 2018a), and then mapping of re-sequenced reads. All the three activities happen independent of each other and get merged when individual results are available.

The present status of AOCC for the three major activity workflows is illustrated in Table 1. Six reference genome sequences have been already published (*Lablab purpureus*, *Vigna subterranea*, *Faidherbia albida*, *Sclerocarya birrea*, and *Moringa oleifera* (Chang et al. 2018a), *Solanum aethiopicum* (Song et al. 2019). The curated whole genome sequences are now freely available in the GigaScience database (GigaDB) (Chang et al. 2018b, c, d, e, f, g; Tena 2019; Song et al. 2019). Six other genomes are in the final stages of assembly (*Eleusine coracana*, *Digitaria exilis*, *Gynandropsis gynandra*, *Annona cherimola*, *Artocarpus heterophyllus*, and *Artocarpus altilis*). Twenty-one species are in the pipeline for WGS, 17 under transcriptome sequencing, and 15 under re-sequencing workflow. Re-sequencing has been partially completed for five species (*Eleusine coracana*, *Vigna subterranea*, *Faidherbia albida*, *Moringa oleifera*, *Gynandropsis gynandra*). As soon as DNA/RNA gets logged into the workflow, the status is shown as "in the pipeline". A total of 46 species have been touched upon by AOCC where at least one of the three activities has been initiated. In addition, 19 species have been sequenced by our collaborators or other partners.



**Fig. 1** Workflow used by African Orphan Crops Consortium (AOCC) to generate genomics resources for 101 African orphan crops. The three parallel workflows are independent of each other and each gets initiated as soon as material is available; however, to develop SNP panels, information from all the three sources are taken into account

## African Plant Breeding Academy (AfPBA)

UC Davis African Plant Breeding Academy (AfPBA, [http://pba.ucdavis.edu/PBA\\_in\\_Africa/](http://pba.ucdavis.edu/PBA_in_Africa/)) is a capacity building arm of AOCC, which plans to train 150 mid-career African plant

breeders by 2021. This training is designed to enable African plant breeders in the use of genomics-based data in their breeding programs with skills and tools to increase the efficiency of cultivar development programs and faster release of improved varieties to the farmers. It aims to empower African plant breeders to put the DNA sequence information on African orphan crops into action in developing new, improved varieties that meet farmer, consumer, and processor needs. To date, 80 scientists have completed the intensive 6-week training, and 36 are participating in the current class. Overall, these scientists represent 28 countries across the African continent; they work with over 105 crop species including 55 orphan crops. Most are mid-career scientists employed in national agricultural research programs; over 80% are PhD scientists, 33% are women.

Although plant breeding is a long-term effort, the outcomes of the training have already been realized (Sogbohossou et al. 2018). Collaborations have been established among the cohorts and these collaborations have been successful in securing substantial funding for research and graduate student training. A community of practice has been developed among the AfPBA graduates, which has given rise to the African Association of Plant Breeders, intended to extend continuing professional development beyond the formal training of the AfPBA. The AfPBA is an outlet for translating the resource of sequence data into crop varieties that can provide the basis of food and nutritional security for Africa.

## Open data access policy

Open access and democratization of sequence information and other genomics data is considered as an important part of the process of finding new innovative solutions for emerging scientific, social, and political challenges (Pauwels 2017). This involves putting the data in public domain and making the translational technologies approachable and affordable to the general users with a minimal skill set. This is considered as an important commitment of AOCC toward the African agricultural research landscape. All the data and accessions created by the AOCC will be made publicly available through gene banks, partners, partners websites, and publications such as GigaScience, NCBI, and/or CNGB Nucleotide Sequence Archive (CNSA: <https://db.cngb.org/cnsa>) databases.

## Conclusion

The AOCC was established with a goal of supporting food, nutrition, and income generation capacity of the African population, mainly the smallholder farmers, by providing diversified options through locally available underutilized

crops/trees on farming landscape driven by nutrient-sensitive food systems approach. The AOCC is committed to using high-end technology-driven solutions to mainstream 101 under-researched, underinvested African crops, referred to as orphan or neglected crops by using next-generation technologies such as genomics and genomics-assisted breeding. In the current phase, AOCC is sequencing and generating genomics resources for 101 target African orphan crops using next-generation sequencing technologies through a vast network of core partners and a network of collaborators. The list of these crops was drawn considering African needs to support food and nutrition targets as well as economic empowerment of smallholder farmers. The AOCC has made a reasonable progress with regard to generating genomic resources for 60 species. To date, the AOCC, through the AfPBA, has empowered 116 of Africa's top plant breeders to expedite development of improved varieties of African orphan crops and other food plants. The AOCC solutions will increase dietary diversity, create new avenues of income generation, and increase the health and standard of living of the African population, including that of smallholder farmers.

**Author contribution statement** ICRAF hosts the AOCC genomics lab and UC Davis hosts AfPBA; ICRAF and UCD team were involved in overall workflow management. The authors shared responsibility as follows—PSH wrote the manuscript with relevant inputs from other authors; PSH, AM, AVD, RJ: supervision, coordination, management of genomics laboratory workflows, and logistics; AM: germplasm acquisition and logistics; PSH, RK, SM: DNA/RNA extractions, logistics, and re-sequencing workflow. BGI team was involved in planning and execution of WGS pipelines as follows—YF, BS, YC: designing and implementing genome surveys and assembly; YC, YF, ML, XL, SW, LL: genome annotations; HL, SP: NGS libraries and primary data generation; BS, SKS: manuscript revision; HL, SC, XX, HY, JW, XL: initiation, supervision, and management of overall WGS pipeline; RM: Director, AfPBA; RM and AVD: core instructor, AfPBA; H-YS, TS, AVD, RJ: conceptualization and establishment of a functional AOCC consortium, collaborations, and partnerships; all the authors read and approved the final manuscript.

**Acknowledgements** The authors acknowledge contributions from Drs Sean Mayes, Wai Kuan Ho, and Presidor Kendabie from Nottingham University and Crops for Future (CFF), Malaysia/UK toward sequencing and publishing bambara groundnut genome (Chang et al. 2018a, c); Dr Damaris Odeny, International Centre for Arid and Semi-arid Tropics, ICRISAT, Nairobi, Kenya, and Drs Elizabeth Balyejusa Kizito, Pamela Nahamya Kabod, and Sandra Ndagire Kamenya all from Uganda Christian University, Mukono, Uganda, for sequencing and publishing African eggplant genome (Song et al. 2019). Other collaborators and partners provided in the text and in the supplementary

material are duly acknowledged. Donations of HiSeq 4000 by Illumina Inc., San Diego, USA, and Ion Protons and Ion Chefs by Thermo Fisher Scientific, Waltham, USA, are also duly acknowledged. The initial genome sequencing of the first five crops (Chang et al. 2018a, b, c, d, e, f) was supported by the Shenzhen Municipal Government of China, (nos. JCYJ20150831201643396 and JCYJ20150529150409546), the State Key Laboratory of Agricultural Genomics (no. 2011DQ782025), and Guangdong Provincial Key Laboratory of Genome Read and Write (no. 2017B030301011). Sequencing of African eggplant was supported by the National Natural Science Foundation of China (no. 31601042), the Science, Technology and Innovation Commission of Shenzhen Municipality under grant, (nos. JCYJ20151015162041454 and JCYJ20160331150739027), as well as the funding from Guangdong Provincial Key Laboratory of Genome Read and Write (no. 2017B030301011).

## Compliance with ethical standards

**Conflict of interest** All the authors declare no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Aighewi BA, Asiedu R, Maroya N, Balogun M (2015) Improved propagation methods to raise the productivity of yam (*Dioscorea rotundata* Poir.). *Food Secur* 7:823–834
- Ayenán MAT, Ezin VA (2016) Potential of Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal & Baudet] and prospects for its promotion. *Agric Food Secur* 5:10. <https://doi.org/10.1186/s40066-016-0058-4>
- Baldermann S, Blagojević L, Frede K, Klopsch R, Neugart S, Neumann A, Ngwene B, Norkoweit J, Schröter D, Schröter A, Schweigert FJ, Wiesner M, Schreiner M (2016) Are neglected plants the food for the future? *Crit Rev Plant Sci* 35:106–119. <https://doi.org/10.1080/07352689.2016.1201399>
- Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, Liu X, Gao D, Clevenger J, Dash S, Ren L, Moretzsohn MC, Shirasawa K, Huang W, Vidigal B, Abernathy B, Chu Y, Niederhuth CE, Umale P, Araújo ACG, Kozik A, Do Kim K, Burow MD, Varshney RK, Wang X, Zhang X, Barkley N, Guimarães PM, Isobe S, Guo B, Liao B, Stalker HT, Schmitz RJ, Scheffler BE, Leal-Bertioli SCM, Xun X, Jackson SA, Michelmore R, Ozias-Akins P (2016) The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat Genet* 48:438–446. <https://doi.org/10.1038/ng.3517>
- Boukar O, Belko N, Chamarthi S, Togola A, Batiéno J, Owusu E, Haruna M, Diallo S, Umar ML, Olufajo O, Fatokun C (2018) Cowpea (*Vigna unguiculata*): genetics, genomics and breeding. *Plant Breed*. <https://doi.org/10.1111/pbr.12589>
- CABI (2018) <https://www.cabi.org/isc/datasheet/1822>. Accessed 30 Jan 2019
- Cannarozzi G, Plaza-Wüthrich S, Esfeld K, Larti S, Wilson YS, Girma D, de Castro E, Chanyalew S, Blösch R, Farinelli L, Lyons E, Schneider M, Falquet L, Kuhlmeier C, Assefa K, Tadele Z (2014) Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (*Eragrostis tef*). *BMC Genom* 15:581. <https://doi.org/10.1186/1471-2164-15-581>
- Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro H, Xu X, Yang H, Wang J, Liu X (2018a) The draft genomes of five agriculturally important African orphan crops. *GigaScience* 8:giy152. <https://doi.org/10.1093/gigascience/giy152>
- Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro H, Xu X, Yang H, Wang J, Liu X (2018b) Genomic data of the Bambara Groundnut (*Vigna subterranæa*). *GigaScience Database*. <https://doi.org/10.5524/101055>
- Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro H, Xu X, Yang H, Wang J, Liu X (2018c) Genomic data of the Apple-Ring Acacia (*Faidherbia albida*). *GigaScience Database*. <https://doi.org/10.5524/101054>
- Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro H, Xu X, Yang H, Wang J, Liu X (2018d) Genomic data of the Hyacinth Bean (*Lablab purpureus*). *GigaScience Database*. <https://doi.org/10.5524/101056>
- Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro H, Xu X, Yang H, Wang J, Liu X (2018e) Genomic data of Marula (*Sclerocarya birrea*). *GigaScience Database*. <https://doi.org/10.5524/101057>
- Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro H, Xu X, Yang H, Wang J, Liu X (2018f) Genomic data of the Horseradish Tree (*Moringa oleifera*). *GigaScience Database*. <https://doi.org/10.5524/101058>
- Chen F, Dong W, Zhang J, Guo X, Chen J, Wang Z, Lin Z, Tang H, Zhang L (2018) The sequenced angiosperm genomes and genome databases. *Front Plant Sci* 9:418. <https://doi.org/10.3389/fpls.2018.00418>
- Dawson IK, Hendre P, Powell W, Sila D, McMullin S, Simons T, Revoredo-Giha C, Odeny DA, Barnes AP, Graudal L, Watson CA, Hoad S, Burnett F, Muchugi A, Roshetko JM, Hale IL, Van Deynze A, Mayes S, Kindt R, Prabhu R, Cheng S, Xu X, Guarino L, Shapiro H, Jamnadass RR (2018) Supporting human nutrition in Africa through the integration of new and orphan crops into food systems: placing the work of the African Orphan Crops Consortium in context. ICRAF Working Paper no 276. Nairobi, World Agroforestry Centre. <http://dx.doi.org/10.5716/WP18003.PDF>
- Dinesh D, Bett B, Boone R, Grace D, Kinyangi J, Lindahl J, Mohan CV, Ramirez-Villegas J, Robinson R, Rosenstock T, Smith J, Thornton P (2015) Impact of climate change on African agriculture: focus on pests and diseases. Findings from CGIAR research program on climate change, agriculture and food security (CCAFS) submissions to the United Nations Framework Convention on Climate Change Subsidiary Body for Scientific and Technological Advice (UNFCCC SBSTA), info note submitted to CGIAR research program on CCAFS, Copenhagen, Denmark. <http://www.ccafs.cgiar.org>. Accessed 20 Sept 2018
- Fernandes SB, Dias KOG, Ferreira DF, Brown PJ (2018) Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *Theor Appl Genet* 131:747–755. <https://doi.org/10.1007/s00122-017-3033-y>

- Food and Agriculture Organization (2017) Regional overview of food security and nutrition in Africa 2016, the challenges of building resilience to shocks and stresses, Accra, <http://www.fao.org/3/a-i6813e.pdf>. Accessed 20 Sept 2018
- Glato K, Aidam A, Kane NA, Bassirou D, Couderc M, Zekraoui L, Scarcell N, Barnaud A, Vigouroux Y (2017) Structure of sweet potato (*Ipomoea batatas*) diversity in West Africa covaries with a climatic gradient. PLoS One 12:e0177697. <https://doi.org/10.1371/journal.pone.0177697>
- Hendre PS, Van Deynze A (2015) How a consortium is changing Africa's food systems? <http://www.africanorphancrops.org/category/resources/blog/page/3/>. Accessed 20 Sept 2018
- Hickey JM, Chiurugwi T, Mackey I, Implementing genomic selection in CGIAR breeding programs workshop participants et al (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nat Genet 49:1297–1302. <https://doi.org/10.1038/ng.3920>
- Hittalmani S, Mahesh HB, Shirke MD, Biradar H, Uday G, Aruna YR, Lohithaswa HC, Mohanrao A (2017) Genome and transcriptome sequence of finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. BMC Genom 18:465. <https://doi.org/10.1186/s12864-017-3850-z>
- Janila P, Variath MT, Pandey MK, Desmae H, Motagi B, Okori P, Manohar SS, Rathnakumar AL, Radhakrishnan T, Liao B, Varshney RK (2016) Genomic tools in groundnut breeding program: status and perspectives. Front Plant Sci 7:289. <https://doi.org/10.3389/fpls.2016.00289>
- Kole C, Muthamilarasan M, Henry R, Edwards D, Sharma R, Abberton M, Bentley J, Blakeney M, Bryant J, Cai H, Cakir M, Cseke LJ, Cockram J, de Oliveira AC, De Pace C, Dempewolf H, Ellison S, Gepts P, Greenland A, Hall A, Hori K, Hughes S, Humphreys MW, Iorizzo M, Ismail AM, Marshall A, Mayes S, Nguyen HT, Ogonnaya FC, Ortiz R, Paterson AH, Simon PW, Tohme J, Tuberosa R, Valliyodan B, Varshney RK, Wullschlegel SD, Yano M, Prasad M (2015) Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. Front Plant Sci 6:563. <https://doi.org/10.3389/fpls.2015.00563>
- Li Y, Ruperao P, Batley J, Edwards D, Khan T, Colmer TD, Pang J, Siddique KHM, Sutton T (2018) Investigating drought tolerance in chickpea using genome-wide association mapping and genomic selection based on whole-genome resequencing data. Front Plant Sci 9:190. <https://doi.org/10.3389/fpls.2018.00190>
- Liang Z, Gupta SK, Yeh C-T, Zhang Y, Ngu DW, Kumar H, Patil HT, Mungra KD, Yadav DV, Rathore A, Srivastava RK, Gupta R, Yang J, Varshney RK, Schnable PS, Schnable JC (2018) Phenotypic data from inbred parents can improve genomic prediction in pearl millet hybrids. G3 8:2513–2522. <https://doi.org/10.1534/g3.118.200242>
- Madzivhandila T, Sibanda S, Yamdjeu AW, Moalosi K, Gwelo FA (2016) Achieving food security and nutrition. In: AGRA, D Sumba, J Niuguna, A Marie, J Njuguna, J Njau A Thiuta (eds) Africa agriculture status report 2016-progress towards agricultural transformation in Africa. Alliance for a Green Revolution in Africa (AGRA), Nairobi, Kenya, pp 235–250. <https://agra.org/aasr2016/public/assr.pdf>
- Maundu P, Achigan-Dako E, Morimoto Y (2009) Biodiversity of African vegetables. In: Shackleton CM, Pasquini MW, Drescher AW (eds) African indigenous vegetables in urban agriculture. Earthscan, London, pp 65–104
- Micronutrient Initiative (2009) Investing in the future: a united call to action on vitamin and mineral deficiencies. Global report 2009. [http://www.unitedcalltoaction.org/documents/Investing\\_in\\_the\\_future.pdf](http://www.unitedcalltoaction.org/documents/Investing_in_the_future.pdf). Accessed 20 Sept 2018
- Muthayya S, Rah JH, Sugimoto JD, Roos FF, Kraemer K, Black RE (2013) The global hidden hunger indices and maps: an advocacy tool for action. PLoS One 8:e67860. <https://doi.org/10.1371/journal.pone.0067860>
- Obidiegwu JE, Asiedu R, Ene-Obong EE, Muoneke CO, Kolesnikova-Allen M (2009) Genetic characterization of some water yam (*Dioscorea alata* L.) accessions in West Africa with simple sequence repeats. J Food Agric Env 7:634–638
- Orwa C, Mutua A, Kindt R, Jamnadass R, Anthony S (2009) Agroforestry Database: a tree reference and selection guide version 4.0. World Agroforestry Centre, Kenya. <http://www.worldagroforestry.org/treedb/>. Accessed 20 Sept 2018
- Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Ohtillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman Ware D, Westhoff P, Mayer KFX, Messing J, Rokhsar DS (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457:551–556. <https://doi.org/10.1038/nature07723>
- Pauwels E (2017) The new bio-citizen: how the democratization of genomics will transform our lives from epidemics management to the internet of living things. Wilson briefs. [https://www.wilsoncenter.org/sites/default/files/new\\_bio\\_citizen\\_democratization\\_of\\_genomics\\_transform\\_our\\_lives.pdf](https://www.wilsoncenter.org/sites/default/files/new_bio_citizen_democratization_of_genomics_transform_our_lives.pdf). Accessed 20 Sept 2018
- Pingali PL (2012) Green Revolution: impacts, limits, and the path ahead. Proc Natl Acad Sci USA 109:12302–12308. <https://doi.org/10.1073/pnas.0912953109>
- PROMUSA (2017) Musapedia, the banana knowledge compendium. <http://www.promusa.org/East+African+highland+banana+subgroup>. Accessed 31 Jan 2019
- PROTA (2018) PlantUse English. [https://uses.plantnet-project.org/e/index.php?title=Main\\_Page&oldid=328573](https://uses.plantnet-project.org/e/index.php?title=Main_Page&oldid=328573). Accessed 30, 31 Jan 2019
- Schippers RR (2000) African indigenous vegetables: an overview of the cultivated species. University of Greenwich, Natural Resources Institute/ACP-EU Technical Centre for Agricultural and Rural Cooperation, London, United Kingdom, pp 214
- Sogbohossou EOD, Achigan-Dako EG, Maundu P, Solberg S, Degue-non EMS, Mumm RH, Hale I, Van Deynze A, Schranz ME (2018) A roadmap for breeding orphan leafy vegetable species: a case study of *Gynandropsis gynandra* (Cleomaceae). Hort Res 5:2. <https://doi.org/10.1038/s41438-017-0001-2>
- Song B, Song Y, Fu Y, Kizito EB, Kabod PN, Liu H, Kamenya SD, Muthemba S, Kariba R, Li X, Wang S, Cheng S, Muchugi A, Jamnadass R, Shapiro H-Y, Van Deynze A, Yang H, Wang J, Xu X, Odeny DA, Liu X (2019) Draft genome sequence of the *Solanum aethiopicum* provides insights into disease resistance, drought tolerance and the evolution of the genome. BioRxiv. <https://doi.org/10.1101/532077>
- Teixwira da Silva JA, Rashid Z, Nhut DT, Sivakumar D, Gera A, Souza MT Jr, Tennant PF (2007) Papaya (*Carica papaya* L.) biology and biotechnology. Tree For Sci Biotechnol 1:47–73
- Tena G (2019) Sequencing forgotten crops. Nat Plants 5:5. <https://doi.org/10.1038/s41477-018-0354-z>
- United Nations, Department of Economic and Social Affairs, Population Division (2017) World population prospects: The 2017 revision, key findings and advance tables. Working Paper no. ESA/P/WP/248. [https://esa.un.org/unpd/wpp/Publications/Files/WPP2017\\_KeyFindings.pdf](https://esa.un.org/unpd/wpp/Publications/Files/WPP2017_KeyFindings.pdf). Accessed 20 Sept 2018
- United Nations General Assembly (2015) Transforming our world: the 2030 Agenda for Sustainable Development A/RES/70/1, <https://www.refworld.org/docid/57b6e3e44.html>. Accessed 14 Feb 2019

- Urasaki N, Takagi H, Natsume S, Uemura A, Taniai N, Miyagi N, Fukushima M, Suzuki S, Tarora K, Tamaki M, Sakamoto M, Terauchi R, Matsumura H (2017) Draft genome sequence of bitter melon (*Momordica charantia*), a vegetable and medicinal plant in tropical and subtropical regions. *DNA Res* 24:51–58. <https://doi.org/10.1093/dnares/dsw047>
- USDA, Agricultural Research Service, National Plant Germplasm System (2019) Germplasm resources information network (GRIN-Taxonomy), National germplasm resources laboratory, Beltsville, Maryland. <https://npgsweb.ars-grin.gov/gringlobal/search.aspx>. Accessed 30 Jan 2019
- Varshney RK (2016) Exciting journey of 10 years from genomes to fields and markets: some success stories of genomics-assisted breeding in chickpea, pigeonpea and groundnut. *Plant Sci* 242:98–107. <https://doi.org/10.1016/j.plantsci.2015.09.009>
- Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, Cannon S, Baek J, Rosen BD, Tar'an B, Millan T, Zhang X, Ramsay LD, Iwata A, Wang Y, Nelson W, Farmer AD, Gaur PM, Soderlund C, Penmetza RV, Xu C, Bharti AK, He W, Winter P, Zhao S, Hane JK, Carrasquilla-Garcia N, Condie JA, Upadhyaya HD, Luo M-C, Thudi M, Gowda CLL, Singh NP, Lichtenzveig J, Gali KK, Rubio J, Nadarajan N, Dolezel J, Bansal KC, Xu X, Edwards D, Zhang G, Kahl G, Gil J, Singh KB, Datta SK, Jackson SA, Wang J, Cook DR (2013) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat Biotechnol* 31:240–246. <https://doi.org/10.1038/nbt.2491>
- Varshney RK, Shi C, Thudi M, Mariac C, Wallace J, Qi P, Zhang H, Zhao Y, Wang X, Rathore A, Srivastava RK, Chitkineni A, Fan G, Bajaj P, Punnuri S, Gupta SK, Wang H, Jian Y, Couderc M, Katta MAVSK, Apudel DR, Mungra KD, Chen W, Harris-Shultz KR, Garg V, Desai N, Doddamani D, Kane NA, Conner JA, Ghatak A, Chaturvedi P, Subramaniam S, Yadav OP, Berthouly-Salazar C, Hamidou F, Wang J, Liang X, Clotault J, Upadhaya HD, Cubry P, Rhone B, Gueye MC, Sunkar R, Dupuy C, Sparvoli F, Cheng S, Mahala RS, Singh B, Yadav RS, Lyons E, Datta SK, Hash CT, Devos KM, Buckler E, Bennetzen JL, Paterson AH, Ozias-Akins P, Grando S, Wang J, Mohapatra T, Weckwerth W, Reif JC, Liu X, Vigouroux Y, Xu X (2017) Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat Biotechnol* 35:99–976. <https://doi.org/10.1038/nbt.3943>
- Wortmann CS, Kirkby RA, Eledu CA, Allen DJ (1998) Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 131 p. (CIAT publication no. 297). <https://hdl.handle.net/10568/54312>. Accessed 20 Sept 2018
- Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, Xie M, Zeng P, Yue Z, Wang W, Tao Y, Bian C, Han C, Xia Q, Peng X, Cao R, Yang X, Zhan D, Hu J, Zhang Y, Li H, Li H, Li N, Wang J, Wang C, Wang R, Guo T, Cai Y, Liu C, Xiang H, Shi Q, Huang P, Chen Q, Li Y, Wang J, Zhao Z, Wang J (2012) Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat Biotechnol* 30: 549–554. <https://doi.org/10.1038/nbt.2195>
- Zheng Y, Wu S, Bai Y, Sun H, Jiao C, Guo S, Zhao K, Blanca J, Zhang Z, Huang S, Xu Y, Weng Y, Mazourek M, Reddy UK, Ando K, McCreight JD, Schaffer AA, Burger J, Tadmor Y, Katzir N, Tang X, Liu Y, Giovannoni JJ, Ling KS, Wechter WP, Levi A, Garcia-Mas J, Grumet R, Fei Z (2019) Cucurbit Genomics Database (CuGenDB): a central portal for comparative and functional genomics of cucurbit crops. *Nucleic Acids Res* 8:D1128–D2236. <https://doi.org/10.1093/nar/gky944>

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