OXFORD

Briefings in Functional Genomics, 17(3), 2018, 198–206

doi: 10.1093/bfgp/ely014
Advance Access Publication Date: 23 April 2018
Review paper

Role of genomics in promoting the utilization of plant genetic resources in genebanks

Peterson W. Wambugu, Marie-Noelle Ndjiondjop and Robert J. Henry

Corresponding author: Robert Henry, Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane, QLD 4072, Australia. Tel.: ±61733460551; Fax: ±61733460555; E-mail: robert.henry@uq.edu.au

Abstract

Global efforts have seen the world's plant genetic resources (PGRs) conserved in about 1625 germ plasm repositories. Utility of these resources is important in increasing the resilience and productivity of agricultural production systems. However, despite their importance, utility of these resources has been poor. This article reviews the real and potential application of the current advances in genomic technologies in improving the utilization of these resources. The actual and potential application of these genomic approaches in plant identification, phylogenetic analysis, analysing the genetic value of germ plasm, facilitating germ plasm selection in genebanks as well as instilling confidence in international germ plasm exchange system is discussed. We note that if genebanks are to benefit from this genomic revolution, there is need for fundamental changes in the way genebanks are managed, perceived, organized and funded. Increased collaboration between genebank managers and the user community is also recommended

Key words: plant genetic resources; utilization; genomics; DNA sequencing; genebank and genetic diversity

Introduction

Plant genetic resources (PGRs) form the natural variations that have supported human kind for several millennia. These resources are the basis for food security in addition to being sources of energy, animal feed, fibre as well as other ecosystem services. They are important in addressing the global challenges that are currently facing the human population, particularly the twin challenge of climate change and food scarcity. Owing to their great importance, effective conservation and sustainable utilization of these resources is critically important and has never been more urgent. As evidenced by the huge number of accessions that are conserved in genebanks for various species (Table 1), it is clear that enormous progress has been made in conserving germ plasm in seed banks (Table 1), and they remain under exploited because of a variety of factors. Promoting the sustainable utilization of biodiversity is a key goal of various global and regional efforts and initiatives as well as international agreements and treaties governing genetic resources.

In addition to these administrative, legal and political measures, which have been put in place, the use of scientific advances particularly genome sequencing has the potential to address some of the challenges that limit sustainable utilization of PGR. Over the past decade, there have been significant advances in DNA sequencing technologies, which are driving many areas of plant science. This has led to dramatic changes in read length, sequencing chemistry, instrumentation, throughput and cost. The current genomic revolution provides tools that help to cost effectively study genetic diversity, identify desirable genes and alleles as well as facilitating their transfer during crop improvement, thus reducing the time to deliver new varieties. Reference genome sequences for a large number of model and non-model species have been published and others continue to be released at a highly unprecedented rate.

Peterson Wambugu is a plant genetic resource scientist at the Genetic Resources Research Institute, Kenya Agricultural Livestock Research Organization (KALRO).
Marie-Noelle Ndjiondjop is a molecular geneticist at Africa Rice Center and is currently the head of the Genetic Resources Unit.
Robert Henry is a professor of Innovation in Agriculture, and director of the Queensland Alliance for Agriculture and Food Innovation at the University of Queensland.

Table 1: Top 10 crop collections held in the world's genebanks

Species	Common names	Number of accessions
Triticum aestivum	Wheat	856 168
Oryza sativa	Rice	773 948
Hordeum vulgare	Barley	466 531
Zea mays	Maize	327 932
Phaseolus vulgaris	Bean	261 963
Sorghum bicolor	Sorghum	235 688
Glycine max	Soybean	229 944
Avena sativa	Oat	130 653
Arachis hypogeae	Groundnut	128 435
Gossypium hirsutum	Cotton	104 780

These reference genome sequences provide perhaps the most important genome resource for promoting use of these species.

However, PGR conservation is lagging behind in embracing advances in molecular biology especially in genome sequencing compared with other areas of plant science [1]. These advances have the potential to aid in addressing fundamental biological questions and greatly impact many aspects of the conservation and utilization of PGR in genebanks (Figure 1). As observed by McCouch *et al.* [2], the application of these approaches, which have been referred to as next-generation genebanking [3], has capacity to dramatically transform previously dormant genebanks into research centres with robust research activities. The question that remains open for speculation is whether the current genomic revolution will translate into increased utilization of PGR from genebanks. This article reviews the real and potential application of genomics in various areas of germ plasm conservation and exchange as well as their potential for promoting germ plasm utility in genebanks.

Overview of sequencing technologies

An account of modern sequencing technologies begins in the 1970s when Sanger sequencing was introduced [4]. The Sanger protocols used went through many innovations, but early versions involved manual cloning of DNA fragments, radiolabelling, polyacrylamide electrophoresis and manual scoring of autoradiograms [3]. Though the Sanger method eventually produced longer read lengths (around 1000 bp) and had a low error rate, it was costly and laborious compared with the present technologies.

The next-generation sequencing (NGS) technologies emerged with improving performance over the past 10 years. For example, the early Illumina platforms generated only short reads of around 30 bp, but this increased to routine analysis of 150 bp. The numbers of reads generated by NGS has also continued to grow with the latest Illumina platform targeting 3 Tbp per run at full development. The increases in data volume have been associated with a continuing decline in the cost per bp with the cost of obtaining sufficient sequence data to cover even large plant genomes becoming much more affordable.

Long-read technology has also been developed and continues to improve with the latest PacBio platform, the sequel delivering increased volumes of data with read lengths >20 kbp. Recent improvements in the system produced by Oxford Nanopore have resulted in generation of long sequences of improving quality [5]. Illumina has also deployed the synthetic long-read technology, which is enabling the accurate and cost-effective assembly of complex and repetitive genomes [6]. These longer-read technologies make genome assembly much easier and have greatly improved the quality of many genome sequences.

The combination of technologies can deliver high-quality plant genomes at relative low cost [7]. It can be expected that the technologies will continue to improve such that obtaining whole-genome sequences of all samples in genebanks may be an option that becomes available to many collections in the near future.

Status of PGR conservation and utilization

Currently, 7.4 million accessions of the world's PGR are conserved in about 1625 genebanks spread globally [1]. Wheat has the highest number of conserved accessions followed by rice, barley and maize in a decreasing order (Table 1). While great efforts have been put in the collection of major crops, thereby resulting in tremendous success in their conservation, minor crops, crop wild relatives (CWRs) as well as neglected and underutilized crops remain grossly under-represented in genebanks [1, 8, 9].

Though it is difficult to accurately assess the extent of use of PGR, it is estimated that <1% of accessions conserved in various germ plasm repositories globally have been used in crop improvement [10]. Since 2006, Food and Agriculture Organization (FAO) [1] reported little change in the use of PGR in developing new varieties. The low use of PGR can be attributed to the lack of information on the potential value of conserved germ plasm, which is arguably one of the greatest challenges that faces genebanks [1, 11, 12]. The capacity to identify and transfer useful alleles to improved varieties has also been inadequate. The various challenges limiting germ plasm use from genebanks will be highlighted in this review, and the potential application of genomics (Figure 1) in addressing them will be discussed.

Plant identification, phylogenetic relationships and DNA barcoding

Proper conservation and effective utilization of germ plasm will require accurate plant identification and a clear biosystematics framework. However, this is usually constrained by the lack of taxonomic expertise [13]. Morphology-based plant identification, which is common in genebanks, increases chances of misidentification especially in case of morphologically similar and closely related species. Well-resolved phylogenetic relationships between cultivated species and their CWRs are vital in making germ plasm conservation management decisions. Additionally, they aid in gene discovery as well as defining strategies for gene transfer during crop improvement. A large proportion of accessions conserved in genebanks remain unidentified or identified up to genus level. DNA barcoding is an effective species identification tool, but there is no universally agreed locus for plant barcoding [14, 15]. Recently, the potential of whole chloroplast genome sequences as a universal barcode in plant identification as well as in resolving phylogenetic relationships has been demonstrated [16–20]. The ongoing advances in sequencing coupled with decreased sequencing costs as well as the high multiplexing capacity for chloroplast genomes will continue to make the whole plastid sequences a popular tool that may eventually replace Sanger-based DNA barcoding.

Owing to the challenges of plastid enrichment [21], sequencing of total DNA and then isolating chloroplast sequences is now the method of choice for most researchers. Chloroplast sequences can be assembled by reference guided assembly where reads are mapped to a reference [16] or by *de novo* assembly followed by selection of chloroplast contigs through homology

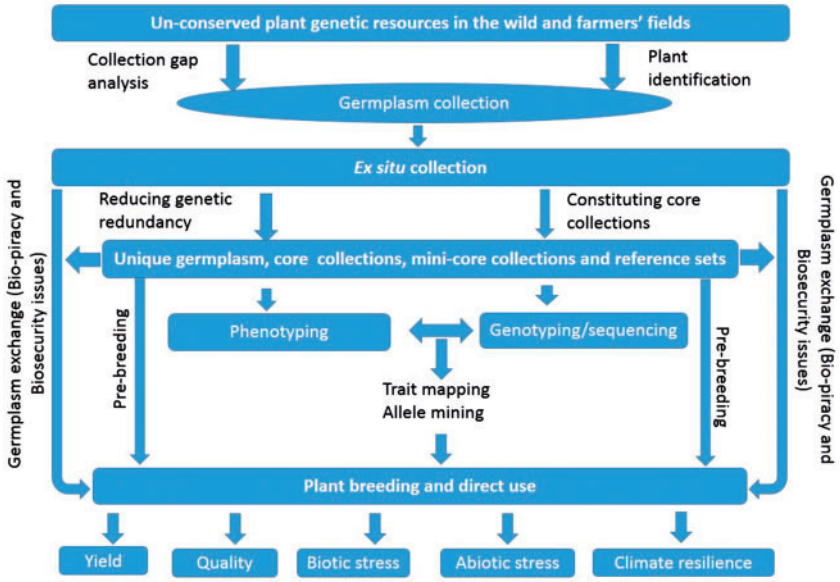

Figure 1. Schematic representation of the application of genomic technologies in germplasm utilization.

searches [21]. Recently, a robust approach using both of these approaches has been used to define genetic and evolutionary relationships between wild and cultivated species that constitute the primary gene pool for rice [22]. The use of whole chloroplast sequences eliminates the need to have a priori information on a locus of choice, a difficulty that acts as a major hindrance to single or multilocus studies. With a much longer sequence than most commonly used DNA barcodes, it has more variation that can help discriminate closely related species. Whole plastid sequences have also been used to identify novel genetic resources [23]. In addition to the use of chloroplast genomes, a set of well-selected informative Single Nucleotide Polymorphisms (SNPs) has been used to detect cases of species misidentification in genebanks [24], which is one of the factors that hinder utility of conserved germ plasm. The capacity for rapid and inexpensive analysis of complete plastid genomes as well as analysis of large numbers of nuclear loci is offering unprecedented opportunities in the field of plant systematics. Previously intractable phylogenetic relationships are now easily being resolved using genomic-based approaches. The routine sequencing of complete nuclear genomes might in the future make whole-genome sequences a tool for use in plant systematics. These advances are expanding the types of questions that genebank managers can ask in the area in plant systematics, thereby potentially addressing challenges that have always limited germ plasm utilization.

Germ plasm characterization

Trait mapping

Genetic diversity in *ex situ* collections have to a large extent been studied anonymously. A large majority of studies have mainly focused on diversity richness by reporting on the number of alleles detected in a collection. This has done little to encourage utility of conserved germ plasm; however, with current improvements in genomic approaches, there is increased impetus towards studying the functional genetic variation of genebank collections. The increased availability of high-quality reference sequences has opened almost unlimited possibilities in deciphering the molecular and genetic basis of biologically and economically important traits. Resequencing of several genotypes through whole-genome shotgun sequencing followed by mapping to the reference is currently the most popular approach for genetic analysis and marker discovery [25–30]. The limited availability of accurate phenotypic data now presents a challenge in studying the value of genetic resources by linking genotypes and phenotypes.

Owing to the large number of unstructured natural populations in genebanks, association mapping studies render themselves better suited for characterizing natural variation of genebank samples as opposed to QTL mapping. Association mapping offers the advantage of higher resolution than QTL mapping, as it helps sample as much natural variation at a particular loci as possible. It also helps in taking advantage of numerous historical recombination events in the study population. Genome-wide association studies (GWASs) of genebank material continue to gain popularity and have been conducted in rice [31, 32], barley [33, 34] alfalfa [35] and wheat [36, 37] among other crops. Genebank samples have wide ecogeographical origins from which they are collected and usually subjected to diverse local adaptations. Consequently, they are likely to show strong population structure [38], which may result in spurious associations [39]. Before association mapping, there is therefore need to subject the genotypes being studied to statistical analysis so as to remove population structure [40]. Whole-genome association studies have been used to construct haplotypes, which are useful in allele mining and elucidating the molecular basis of important traits [31, 41]. One limitation in the use of GWAS in genebanks is that they require a lot of genomic information about SNPs and may therefore not be suitable for species whose genomes are poorly studied and/or not well annotated [42]. A large proportion of genebank collections comprise minor, neglected and underutilized species, which are not well studied and lack enough genomic resources.

Despite the dramatic decrease in sequencing costs, whole-genome analysis of a large number of samples is not economically feasible for a large majority of genebanks and researchers [43]. As reviewed by Schlötterer *et al.* [44], the search for cost-effective approaches has led to the increased popularity of pool

sequencing (pool-seq) in genetic mapping and population genetic studies. The costs of library preparation still form a substantial cost of NGS [45, 46]. Bulk segregant analysis (BSA) [47] is one approach that helps to reduce library preparation costs, as it reduces the number of samples to sequence. BSA, also called DNA pooling [48], is a trait-based genetic mapping approach where DNA samples from individuals on the extreme ends of the phenotypic distribution are pooled. It has recently been demonstrated that in addition to the traditionally recommended procedure of bulking individuals from the extreme ends, bulks can also be formed by randomly choosing individuals from the population. This helps increase resolution and reduce false positives [49]. When coupled with whole-genome analysis, BSA presents a particularly powerful and cost-effective genetic mapping tool. NGS-aided BSA has largely been used on pedigree-based bi-parental populations where it has successfully been used to map various traits in several crops, among them tomato [50], sorghum [51], chick pea [52], cucumber [53], rice [54] and pigeon pea [55]. Recently, Yang *et al.* [49] demonstrated the potential of using NGS-aided BSA on natural populations as opposed to bi-parental populations. These authors suggested that this application is best suited for use in species with inadequate genomic resources. This makes it particularly useful for trait mapping in orphan crops whose germ plasm forms a substantial proportion of genebank collections. Trait mapping studies have led to the identification of numerous marker trait relationships which, as whole-genome sequences increase, have the capacity to ease the task of selecting the right germ plasm by genebank managers. Owing to the limited capacity that exists in genebanks, phenotyping and genomic analysis can be more efficiently undertaken in collaboration with the user community. The user community can provide phenotypic and genomic data as well as other important feedback on materials received from the genebank, thereby helping add value to the conserved germ plasm.

Source of materials for characterization on climate change adaptation

Genebanks act as an important source of genetic material for genomic studies in various areas of plant science, which would otherwise be difficult to conduct with experimental populations. For example, application of genomics as a tool for achieving greatly accelerated breeding for climate resilient crops requires a good understanding of the molecular and genetic basis of climate change adaptation. Conducting these kinds of studies requires long-term experiments or the availability of genetic resources, which have been collected from localities with varied climatic conditions and have thus been subjected to different climatic regimes over a long period of time. These long-term experiments are rare, and genebanks, most of which house collections largely comprising natural populations, which have been assembled over several decades, therefore provide a valuable resource for such genomic studies. Using genomics, it is possible to identify factors that lead to successful adaptation during periods of climate change [56]. By studying the genomes of populations collected in localities with contrasting environmental conditions, it is possible to identify genomic regions that are under selective pressure and these might suggest loci that may be important for adaptation to climate change [57]. When analysing adaptive changes that have taken place in finger millet samples collected over a period of 27 years, Vigouroux *et al.* [58] observed an increase in frequency of an early flowering allele at the PHYC locus. Similarly, analysis of samples of wild progenitors of barley and wheat collected over a period spanning about 28 years showed that their flowering time had shortened by an average of about 10 days [59]. Studying the transcriptome of populations growing along an environmental gradient may also reveal changes in gene expression of the same set of genes, thus potentially shedding light on the genetics of adaptation [60]. The drawback with this approach is that the number of upregulated or downregulated genes can be high [61] and may require significant additional efforts to narrow them down. Genomic information emanating from studies on climate change adaptation will support decision-making on what genetic resources to conserve in a genebank in future. Genomics will facilitate identification of novel alleles emerging because of climate change adaptation. These novel alleles, which give plants the adaptive capacity, should be prioritized for conservation, as they are important in developing climate resilient crops.

Constituting core collections

The concept of core collections was proposed by Frankel [62] and involves selecting a small subset of germ plasm that represents maximum proportion of genetic diversity present in the entire collection. Sequencing and high-throughput genotyping are providing efficient, reliable and cost-effective tools to establish core collections. These tools are for example enabling the identification of genetic redundancy in previously constituted core collections [34], thus demonstrating their capacity to unambiguously discriminate closely related accessions. With the current genomic revolution, this concept however appears to be getting redundant. It appears like the current trend is to extend the molecular characterization of genebank samples outside the limits of core collections by undertaking large-scale sequencing of genebank collections. For example, plans are underway to genotype the entire Oryza *glaberrima* collection conserved at AfricaRice genebank using SNP markers. As already stated, a large international initiative known as DivSeek aimed at sequencing all the accessions held in *ex situ* conservation is underway [63]. A total of 3000 rice genomes representing a broad spectrum of rice genetic diversity selected from various germ plasm collections have been sequenced and published [64]. Though it is expected that sequencing of these genomes will facilitate rice genomic analysis, there is no available evidence on whether these genomic resources have led to increased use of these genetic resources. The reduced interest by genebank managers to constitute core collections can probably be attributed to the capacity offered by sequencing and genotyping. These technologies are enabling the fast and cost-effective genomic analysis of a large number of samples. The availability of genomic data for a large number of genebank samples promotes greater use as compared with the case of core collections.

Pre-breeding and broadening genetic base

A significant proportion of genebank collections comprise wild species, which represent the primary, secondary and tertiary gene pool [65]. These CWRs have immense value in terms of the useful genes and alleles that have potential to improve the gene pool of crop species. However, some breeders are reluctant to use CWRs in their breeding programmes because of linkage drag. For example, for several decades, breeders in Japan have faced challenges in developing elite varieties with resistance against blast and also possessing good quality traits because of the co-introduction of desirable alleles for blast resistance and

the undesirable ones controlling poor grain quality [66, 67]. While such associations could be because of pleiotropy, they have in most cases been found to be because of tightly linked genes [67]. Breaking this linkage is usually costly and time-consuming. Most breeders therefore prefer to reuse their usually limited working collections, thereby leading to release of varieties with narrow genetic diversity. This narrow genetic base negatively affects the resilience, productivity and sustainability of agricultural production systems. Pre-breeding is therefore an important activity that helps to improve the genetic value, attractiveness as well as suitability of genebank materials to breeders.

Currently, there are predictive models that have been developed, which have the capacity to predict those SNP variations that are most likely to lead to deleterious phenotypic effects [68]. This means that materials with such SNP alleles can be eliminated from breeding programmes at an early stage. In addition to linkage drag, the reluctance by breeders to use wild species can also be attributed to their unwillingness to disrupt the favourable linkage blocks in their breeding materials that takes time to create. Using high-throughput sequencing and genotyping approaches, it is currently possible to obtain cross-specific sequence markers such as SNPs that can be used to saturate the genetic background of both parents [69]. Using these markers, it is possible to monitor the degree of introgression of specific alleles or genomic regions in the offspring [10, 70]. This monitoring ensures that the genome of the recurrent parent can be efficiently regained, and the tracking of both desired and undesired alien alleles ensures that only narrow segments of the wild species, preferably having only the desired allele, are introgressed [69]. To minimize linkage drag, it is recommended that the markers to monitor the introgression should be as close as possible to the desired genomic region [71]. The use of functional markers linked to the gene of interest is preferable as random markers may be located far away [72]. Deep sequencing of the genomic region controlling a particular trait can help identify the loci/alleles responsible for the undesirable trait and thus select recombinants lacking this undesirable allele [67, 73]. Genomics therefore plays an important role in the identification of both beneficial and deleterious alleles as well as facilitating the transfer of the beneficial ones during crop improvement. This minimizes the challenges associated with wild and unadapted materials, thereby enhancing their utility in crop improvement.

Selecting germ plasm from a genebank

Genebank managers occasionally receive germ plasm requests for accessions with specified phenotypic traits. With the huge numbers of accessions held in most of the genebanks, searching for an accession with certain specified traits has been likened to searching for the proverbial needle in haystack. Faced with this challenge, curators can effectively make use of genomic, passport as well as ecological data in trying to identify materials that are likely to possess the trait of interest. Though huge volumes of whole-genome sequence data continue flowing from sequencing machines, questions have been asked about how such data will be useful to a genebank manager in aiding germ plasm selection.

Homology and candidate gene analysis-based selection

With the growing volume of fully annotated genomes and knowledge of candidate genes, it is possible to use SNPs to narrow down the number of accessions likely to possess a particular trait by identifying and eliminating those that have no SNP variations in the gene of interest [68]. The improved understanding of the metabolic processes controlling various traits has led to increased availability of information on candidate genes. Information on whether the SNP is in the coding, non-coding or promoter region as well as whether it is synonymous or non-synonymous will help understand its functional consequences or effect on gene expression [74, 75]. SNP markers developed from expressed sequence tags, or from genes, referred as genic markers [76], are important from a genebank perspective, as they may help to assign available SNPs putative functions through homology inferred from a BLASTX analysis [77]. This in turn helps to further narrow down the number of accessions likely to possess the specific phenotype of interest. However, the absence of germ plasm evaluation data in most genebanks makes it difficult to confirm the veracity of the assigned putative function with the corresponding phenotype, hence limiting the application of these fundamental genomic resources. While access to genotypic data was previously a major challenge, the current advances in genomics have led to a dramatic shift that now makes access to high-quality phenotypic data the key bottleneck in functional genomics. Selection of accessions from genebanks is also benefiting from the increased number of functional markers, which are continually being developed though linkage as well as association studies. These genetic markers and candidate genes are a valuable resource that aids in the deployment of germ plasm for crop improvement.

Ecology and adaptive genomics-assisted selection

Genebank managers have sometimes used the Focussed Identification of Germplasm Strategy (FIGS) for targeted germ plasm selection. This approach is based on the idea that plants will develop adaptive traits based on the selection pressures that they have been subjected to in their particular growth environment. FIGS has successfully been used to identify new variation in various crops [78–80]. The use of FIGS will therefore facilitate identification of a manageable number of accessions likely to possess a particular trait, whose potential genetic value can then be analysed using genomic approaches. Whole chloroplast genome sequence data, for example, have shown that wild populations from drier and hotter environments are more genetically diverse than those from wet and cooler areas [81]. In the same vein, analysis of genomes of genebank samples obtained from different adaptation conditions will reveal different adaptive traits and identify genes/variations responsible for trait adaptation. In a study conducted by Parida and Mukerji [82], analysis of SNPs found in biotic and abiotic stress-responsive genes in a diverse set of rice genotypes revealed nine genotype-discriminating SNPs, which were found to have strong association with ecological adaptation. Information on expressed genes obtained through transcriptome sequencing is likely to be useful in adaptation and ecological genomics. However, as noted by Ganal and Wieseke [83], use of SNP identification in transcribed sequences may be challenging, as relatively fewer SNPs may be available because of selection pressure particularly in organisms with low polymorphism rates. Whole-genome analysis should help unravel the genetic basis of ecological adaptation. Analysis of the loci responsible for ecological adaptation can help identify beneficial alleles, which can be the basis for germ plasm selection from genebanks.

Instilling confidence in germ plasm exchange in genebanks

One of the key responsibilities of genebanks is the provision of good quality germ plasm. The free flow of germ plasm from genebanks has however, at times, been constrained by a variety of factors, thereby limiting utilization. These factors include strict biosecurity restrictions and the reluctance to share germ plasm because of fear of biopiracy. The lack of clear and conducive germ plasm access policies is also a factor that is responsible for reduced germ plasm exchange. While some of these challenges will largely require sociopolitical interventions, the use of genomics has potential to somewhat address them.

Fear of biopiracy

There are numerous reported cases where biodiversity or important traditional knowledge has been misappropriated with no benefits accruing to the providers/owners/custodians of such resources [84]. Similarly, intellectual property rights have been severally infringed on in case of germ plasm shared by breeders or genebanks. This misappropriation of genetic resources is against the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) as well as the Convention of Biological Diversity (CBD) and its implementing instruments, namely, the Nagoya Protocol, which aims at sharing the benefits arising from the utilization of genetic resources in a fair and equitable way. This has resulted in a climate of mistrust in international germ plasm exchange especially from national genebanks, thus leading to protectionist tendencies. Studies have shown that there is a palpable reluctance to share germ plasm by genebank managers especially in some national genebanks and plant breeders because of fear of biopiracy [85]. The proliferation of policies and regulations aimed at limiting germ plasm exchange will ultimately hurt bioprospecting.

In the face of these fears, there is need to instil confidence in the germ plasm exchange system by assuring germ plasm providers that the system has the capacity to identify and possibly prevent cases of biopiracy. DNA evidence is admissible in a court of law [86, 87]. Genomics therefore provides a method of legally protecting germ plasm and prosecuting cases of infringement of intellectual property rights or germ plasm sharing agreements. DNA fingerprinting using low- and medium-throughput markers has previously been used to show the geographic origin of a sample [88, 89]. It has also been used to identify varieties or cultivars as well as enhance patent protection [90, 91]. In cases of endemic species localized in a certain geographic region, prosecuting cases of biopiracy may be relatively easy. However, in other circumstances, courts of laws might require a genotyping system with higher precision and accuracy to enhance integrity and avoid miscarriage of justice. There is need for the fight against biopiracy to leverage the current advances in genomic technologies, which are driving contemporary forensic science. Targeted sequencing of a high number of informative loci increases resolution and should help discriminate even closely related accessions. These advances provide the capacity to track a distributed accession and link it to the original genebank sample. To protect plant breeder rights, SNPs are the basis of cost-effective, precise and more efficient intellectual property systems that have capacity for greater harmonization between various jurisdictions [92]. A carefully selected set of SNPs with good discriminating power has the capacity to determine the distinctness, uniformity and stability of a variety, a criterion that is important in awarding plant breeders rights.

Restrictive plant biosecurity requirements

In an attempt at containing the spread of pests and diseases, which may be costly to control, various countries have put in place extremely stringent quarantine regimes [1, 85]. Though quarantine-related issues are not the mandate of genebanks but rather of quarantine agencies, genebanks have the responsibility to ensure conservation and sharing of healthy materials. Ideally, seed health tests should be routine activities conducted on any germ plasm that is set to be shared. However, the capacity of many genebanks to ensure safe international exchange by effectively conducting such tests especially for vegetatively propagated species is limited [93]. With the ever-increasing incidences of pests and diseases, the need for reliable, robust and high-throughput quarantine methods has never been more important. Advances in NGS are set to revolutionize plant quarantine diagnostics with great success already recorded in the identification and characterization of plant pathogens [94]. The number of sequenced genomes of plant pathogens is increasing rapidly [95], and by using these sequences, it is now possible to more accurately distinguish between species as well as subspecies. The use of high-throughput whole-genome sequencing helps to develop high-resolution markers, which can discriminate closely related pathogens as well as clonal lineages [96, 97]. For example, using high-throughput whole-genome sequencing, it was possible to trace and explain the origin and emergence of the new maize streak virus in Africa [98] and the spread of tomato yellow leaf curl virus [99]. With quarantine restrictions increasing day by day, genebanks may be forced to invest in sequencing technologies for plant pathogen diagnostics. Such efforts and investments, which can be undertaken in collaboration with biosecurity agencies, will instil confidence in germ plasm exchange significantly. This might cause countries to relax their quarantine restrictions, thereby leading to increased flow of germ plasm between genebanks and various users.

Challenges

As already noted in this review, though we are on the threshold of a genomic revolution, which has capacity to transform genebanks, this is likely to be unattainable goal because of a variety of factors. The greatest challenges revolve around cost, funding, availability of genomic resources and technical and infrastructural capacity. A large majority of the genebanks, save for those in the CGIAR system and a few selected national genebanks in developed countries, have inadequate infrastructural and bioinformatics capacity. The high-performance computing resources required to store and analyse NGS data are beyond the financial capacity of a large majority of genebanks. Cloud computing is however becoming popular and provides a ray of hope for genebanks, as it is increasingly becoming possible to share computing resources between partner institutions [57]. Equally challenging is the lack of genomic resources for a majority of the minor, neglected and underutilized plant species, which constitute a large proportion of genebank samples. The African Orphan Crops Consortium is seeking to address this lack of genomic resources for orphan and underutilized crops by sequencing, assembling and annotating genomes of 101 species, which are important for food and nutritional security in Africa. Availability of these sequences will aid in gene discovery and allele mining in these species, which will be a major boost in efforts aimed improving food and nutritional security.

Future perspectives

In the rapidly changing field of genomics, there is no telling what the future holds for the application of genomics in genebanks. Though it has been suggested that the current genomic revolution has capacity to dramatically change genebank activities, this vision is likely to remain a mirage unless there is a paradigm shift in the way genebanks are perceived, organized, managed and funded. There is need to develop a critical mass of scientists trained in areas such as genomics, computational biology and population genetics to work in genebanks alongside other genebank staff. One of the ways that the genebanks can benefit from the current advances in genomic technologies is through greater collaboration with the user community. Some of the linkages and collaborations that have been established with the aim of supporting conservation, management and utilization of Plant Genetic Resources for Food and Agriculture (PGRFA) include DivSeek and Global Information System (GLIS). The assigning of Digital Object Identifiers (DOIs), which is a permanent identifier of PGRFA, makes it easier to share PGRFA information by easily and unambiguously referencing PGRFA samples across organizations. It will for example be easy to link research outputs with the PGRFA samples used to produce that output [100]. These linkages and collaborations will help genebanks leverage on the greater infrastructural, technical and financial capacity available to the user community, which currently, and in the foreseeable future, remains a great constraint in promoting effective conservation and sustainable use of PGR. Application of genomics in genebanks has potential of ultimately having an increasing impact in the development of more resilient varieties. This is likely to result in increased agricultural productivity, thereby having a positive impact on global food and nutritional security.

Key Points

- Genomics is finding wide application in the conservation and utilization of genetic resources and has the potential to revolutionize the way genebanks are managed.
- Utilization of conserved germ plasm has been poor, but current sequencing and genotyping technologies have potential to address the various challenges limiting germ plasm utility.
- Though most genebanks lack the capacity to access and use current genomic technologies, they can circumvent this challenge by ensuring greater collaboration with the user community.

Funding

This work was supported by Australia Awards through an ADS scholarship granted to the first author.

References

1. FAO. Second report on the world's plant genetic resources for food and agriculture. Rome, Italy: FAO, 2010, 299.
2. McCouch SR, McNally KL, Wang W, *et al*. Genomics of gene banks: a case study in rice. *Am J Bot* 2012;**99**(2):407–23.
3. Treuren R, Hintum TJL. Next-generation genebanking: plant genetic resources management and utilization in the sequencing era. *Plant Genet Resour* 2014;**12**:298–307.
4. Thudi M, Li Y, Jackson SA, *et al*. Current state-of-art of sequencing technologies for plant genomics research. *Brief Funct Genomics* 2012;**11**(1):3–11.
5. Jason DM, Aaron MW, Tam S, *et al*. Long-read genome sequencing identifies causal structural variation in a Mendelian disease. *Genet Med* 2018;**20**(1):159–63.
6. McCoy RC, Taylor RW, Blauwkamp TA, *et al*. Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements. *PLoS One* 2014;**9**(9):e106689.
7. Lee H, Gurtowski J, Yoo S, *et al*. Error correction and assembly complexity of single molecule sequencing reads. *bioRxiv* 2014.doi: 10.1101/006395.
8. Wambugu P, Furtado A, Waters D, *et al*. Conservation and utilization of African Oryza genetic resources. *Rice* 2013;**6**(1):29.
9. Bordoni P, Hodgkin T. Climate change and the conservation of plant genetic resources. *J Crop Improv* 2012;**26**(3):329–45.
10. Sharma S, Upadhyaya HD, Varshney RK, *et al*. Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Front Plant Sci* 2013;**4**:309.
11. Wambugu PW, Muthamia ZK. Second report on the state of plant genetic resources for food and agriculture in Kenya. Rome, Italy: FAO, 2009.
12. Khoury C, Laliberté B, Guarino L. Trends in ex situ conservation of plant genetic resources: a review of global crop and regional conservation strategies. *Genet Resour Crop Evol* 2010;**57**(4):625–39.
13. Chase MW, Fay MF. Barcoding of plants and fungi. *Science* 2009;**325**(5941):682–3.
14. CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci USA* 2009;**106**:12794–7.
15. Lahaye R, van der Bank M, Bogarin D, *et al*. DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA* 2008;**105**(8):2923–8.
16. Nock CJ, Waters DLE, Edwards MA, *et al*. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol J* 2011;**9**(3):328–33.
17. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol* 2009;**7**(1):84.
18. Yang JB, Tang M, Li HT, *et al*. Complete chloroplast genome of the genus Cymbidium: lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol Biol* 2013;**13**(1):84.
19. Erickson DL, Spouge J, Resch A, *et al*. DNA barcoding in land plants: developing standards to quantify and maximize success. *Taxon* 2008;**57**(4):1304–16.
20. Waters DLE, Nock CJ, Ishikawa R, *et al*. Chloroplast genome sequence confirms distinctness of Australian and Asian wild rice. *Ecol Evol* 2012;**2**(1):211–17.
21. McPherson H, van der Merwe M, Delaney SK, *et al*. Capturing chloroplast variation for molecular ecology studies: a simple next generation sequencing approach applied to a rainforest tree. *BMC Ecol* 2013;**13**(1):8.
22. Wambugu PW, Brozynska M, Furtado A, *et al*. Relationships of wild and domesticated rices (Oryza AA genome species) based upon whole chloroplast genome sequences. *Sci Rep* 2015;**5**(1):13957.
23. Brozynska M, Omar ES, Furtado A, *et al*. Chloroplast genome of novel rice germplasm identified in Northern Australia. *Trop Plant Biol* 2014;**7**(3–4):111–20.
24. Orjuela J, Sabot F, Chéron S, *et al*. An extensive analysis of the African rice genetic diversity through a global genotyping. *Theor Appl Genet* 2014;**127**(10):2211–23.

25. Li S, Xie K, Li W, *et al*. Re-sequencing and genetic variation identification of a rice line with ideal plant architecture. *Rice* 2012;**5**(1):18.
26. Huang X, Huang T, Dong G, *et al*. High-throughput genotyping by whole-genome resequencing. *Genome Res* 2009;**19**(6): 1068–76.
27. Arai-Kichise Y, Shiwa Y, Nagasaki H, *et al*. Discovery of genome-wide DNA polymorphisms in a landrace cultivar of japonica rice by whole-genome sequencing. *Plant Cell Physiol* 2011;**52**(2):274–82.
28. Xu X, Liu X, Ge S, *et al*. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat Biotechnol* 2012;**30**(1):105–11.
29. Wang MH, Yu Y, Haberer G, *et al*. The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nat Genet* 2014;**46**(9):982–8.
30. Li J, He W, Yang G, *et al*. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet* 2010;**42**(12):1053–9.
31. Lu Y, Li J, Buckler ES, *et al*. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 2010;**42**(11): 961–7.
32. Huang X, Zhao Y, Wei X, *et al*. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 2012;**44**(1):32–9.
33. Haseneyer G, Sharma R, Graner A, *et al*. Genome-wide association studies for Agronomical Traits in a world wide Spring Barley Collection. *BMC Plant Biol* 2012;**12**:16.
34. Muñoz-Amatriaín M, Hayes PM, Muehlbauer GJ, *et al*. The USDA barley core collection: genetic diversity, population structure, and potential for genome-wide association studies. *PLoS One* 2014;**9**(4):e94688.
35. Zhang T, Yu LX, Zheng P, *et al*. Identification of loci associated with drought resistance traits in heterozygous autotetraploid alfalfa (*Medicago sativa* L.) using genome-wide association studies with genotyping by sequencing. *PLoS One* 2015;**10**(9): e0138931.
36. Mora F, Castillo D, Lado B, *et al*. Genome-wide association mapping of agronomic traits and carbon isotope discrimination in a worldwide germplasm collection of spring wheat using SNP markers. *Mol Breed* 2015;**35**(2):69.
37. Tadesse W, Ogbonnaya FC, Jighly A, *et al*. Association mapping of resistance to yellow rust in winter wheat cultivars and elite genotypes. *Crop Sci* 2014;**54**(4):1905–616.
38. Yu J, Buckler ES. Genetic association mapping and genome organization of maize. *Curr Opin Biotechnol* 2006;**17**(2):155–60.
39. Flint-Garcia SA, Thornsberry JM, Buckler ES, IV. Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol* 2003;**54**: 357–74.
40. Utomo H,S, Wenefrida I, Linscombe S,D, Progression of DNA Marker and the next generation of crop development. In: Goyal A (ed). *Crop Plant*. Croatia: InTech, 2012.
41. Jia G, Huang X, Zhi H, *et al*. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat Genet* 2013;**45**:957–61.
42. Wu B, Han Z, Xing Y, Genome mapping, markers and QTLs. In: Zhang Q., Wing R (eds). *Genetics and Genomics of Rice*. NewYork, NY: Springer-Verlag New York, 2013.
43. Kilian B, Graner A. NGS technologies for analyzing germplasm diversity in genebanks. *Brief Funct Genomics* 2012;**11**(1): 38–50.
44. Schlötterer C, Tobler R, Kofler R, *et al*. Sequencing pools of individuals—mining genome-wide polymorphism data without big funding. *Nat RevGenet* 2014;**15**(11):749–63.
45. Harakalova M, Mokry M, Hrdlickova B, *et al*. Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective targeted next-generation sequencing. *Nat Protoc* 2011;**6**(12):1870–86.
46. Rohland N, Reich D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res* 2012;**22**(5):939–46.
47. Michelmore RW, Paran I, Kesseli RV. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci USA* 1991;**88**(21):9828–32.
48. Giovannoni J, Wing R, Ganal M, *et al*. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Res* 1991;**19**(23):6553–8.
49. Yang J, Jiang H, Yeh CT, *et al*. Extreme-phenotype genome-wide association study (XP-GWAS): a method for identifying trait-associated variants by sequencing pools of individuals selected from a diversity panel. *Plant J* 2015;**84**(3):587–96.
50. Illa-Berenguer E, Van Houten J, Huang Z, *et al*. Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. *Theor Appl Genet* 2015;**128**(7):1329–42.
51. Han Y, Lv P, Hou S, *et al*. Combining next generation sequencing with bulked segregant analysis to fine map a stem moisture locus in sorghum (*Sorghum bicolor* L. Moench). *PLoS One* 2015;**10**(5):e0127065.
52. Das S, Upadhyaya HD, Bajaj D, *et al*. Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. *DNA Res* 2015;**22**: 193–203.
53. Lu H, Lin T, Klein J, *et al*. QTL-seq identifies an early flowering QTL located near flowering locus T in cucumber. *Theor Appl Genet* 2014;**127**(7):1491–9.
54. Yang Z, Huang D, Tang W, *et al*. Mapping of quantitative trait loci underlying cold tolerance in rice seedlings via high-throughput sequencing of pooled extremes. *PLoS One* 2013;**8**(7):e68433.
55. Singh VK, Khan AW, Saxena RK, *et al*. Next-generation sequencing for identification of candidate genes for Fusarium wilt and sterility mosaic disease in pigeonpea (*Cajanus cajan*). *Plant Biotechnol J* 2015;**14**:1183–94.
56. Henry RJ. Genomics strategies for germplasm characterization and the development of climate resilient crops. *Front Plant Sci* 2014;**5**:68.
57. Stillman JH, Armstrong E. Genomics are transforming our understanding of responses to climate change. *BioScience* 2015;**65**(3):237–46.
58. Vigouroux Y, Ali A, Ndjeunga J, *et al*. Selection for earlier flowering crop associated with climatic variations in the Sahel. *PLoS One* 2011;**6**(5):e19563.
59. Nevo E. Evolution of wild barley and barley improvement. In: Zhang G, Li C, Liu X (eds). *Advance in Barley Sciences: Proceedings of 11th International Barley Genetics Symposium*. Dordrecht; Hangzhou; New York, NY: Springer Science + Business Media, 2012, 1–23.
60. Franks SJ, Hoffmann AA. Genetics of climate change adaptation. *Annu Rev Genet* 2012;**46**:185–208.
61. Urano K, Kurihara Y, Seki M, *et al*. 'Omics' analyses of regulatory networks in plant abiotic stress responses. *Curr Opin Plant Biol* 2010;**13**(2):132–8.
62. Frankel OH. Genetic perspectives of germplasm conservation. In: Arber WK, Llinensee K, Peacock WJ., *et al*. (eds). *Genetic Manipulation: Impact on Man and Society*. Cambridge: Cambridge University Press, 1984.

63. McCouch S, Baute GJ, Bradeen J, *et al*. Agriculture: feeding the future. *Nature* 2013;**499**(7456):23–4.
64. Li JY, Wang J, Zeigler RS. The 3, 000 rice genomes project: new opportunities and challenges for future rice research. *Gigascience* 2014;**3**:8.
65. Harlan JR, de Wet JMJ. Toward a rational classification of cultivated plants. *Taxon* 1971;**20**(4):509–17.
66. Saka N. A rice (*Oryza sativa* L.) breeding for field resistance to blast disease (Pyricularia oryzae) in Mountainous Region Agricultural Research Institute, Aichi Agricultural Research Center of Japan. *Plant Prod Sci* 2006;**9**:3–9.
67. Fukuoka S, Saka N, Koga H, *et al*. Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* 2009;**325**(5943):998–1001.
68. Xu Y. Plant genetic resources: management, evaluation and enhancement. *Mol Plant Breed* Wallingford, UK: CABI, 2010, 151–94.
69. Henry RJ, Rice N, Waters DLE, *et al*. Australian Oryza: utility and conservation. *Rice* 2010;**3**(4):235–41.
70. McNally KL, Bureau TE, Stokowski R, *et al*. Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc Natl Acad Sci USA* 2009;**106**(30):12273–8.
71. Hospital F. Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. *Genetics* 2001;**158**(3):1363–79.
72. Lau WCP, Rafii MY, Ismail MR, *et al*. Review of functional markers for improving cooking, eating and the nutritional qualities of rice. *Front Plant Sci* 2015;**6**:832.
73. Varshney RK, Terauchi R, McCouch SR. Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. *PLoS Biol* 2014;**12**(6):e1001883–8.
74. Hirakawa H, Shirasawa K, Ohyama A, *et al*. Genome-wide SNP genotyping to infer the effects on gene functions in tomato. *DNA Res* 2013;**20**(3):221–33.
75. Frousios K, Iliopoulos CS, Schlitt T, *et al*. Predicting the functional consequences of non-synonymous DNA sequence variants—evaluation of bioinformatics tools and development of a consensus strategy. *Genomics* 2013;**102**(4):223–8.
76. Varshney RK, Mahendar T, Aggarwal RK, *et al*. Genic Molecular Markers in Plants: Development and Applications. In: Varshney RK, Tuberosa R. (eds). *Genomics-Assisted Crop Improvement*. Dordrecht: Springer Netherlands, 2007, 13–29.
77. Varshney RK, Baum M, Guo P, *et al*. Features of SNP and SSR diversity in a set of ICARDA barley germplasm collection. *Mol Breed* 2010;**26**(2):229–42.
78. Bari A, Street K, Mackay M, *et al*. Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. *Genet Resour Crop Evol* 2012;**59**(7):1465–81.
79. Endresen DTF, Street K, Mackay M, *et al*. Sources of resistance to stem rust (Ug99) in bread wheat and durum wheat identified using focused identification of germplasm strategy. *Crop Sci* 2012;**52**(2):764–73.
80. Bhullar NK, Street K, Mackay M, *et al*. Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the "Pm3' resistance locus. *Proc Natl Acad Sci USA* 2009;**106**(23):9519–24.
81. Fitzgerald TL, Shapter FM, McDonald S, *et al*. Genome diversity in wild grasses under environmental stress. *Proc Natl Acad Sci USA* 2011;**108**(52):21140–5.
82. Parida SK, Mukerji M, Singh AK, *et al*. SNPs in stress-responsive rice genes: validation, genotyping, functional relevance and population structure. *BMC Genomics* 2012;**13**(1):426.
83. Ganal M, Wieseke R, Luerssen H, *et al*. High-throughput SNP profiling of genetic resources in crop plants using genotyping arrays. In: Tuberosa R, Graner A, Frison E (eds). *Genomics of Plant Genetic Resources: Volume 1. Managing, Sequencing and Mining Genetic Resources*. Dordrecht: Springer Netherlands, 2014, 113–29.
84. McGown J. *Out of Africa: Mysteries of Access and Benefit Sharing*. In: Burrows B (ed). Washington, USA: Edmonds Institute and the Africa Centre for Biosafety, 2006, 54.
85. Wambugu PW, Muthamia Z. Incentives and disincentives for Kenya's participation in the multilateral system of access and benefit sharing. In: Noriega IL, Halewood M, Lapeña I (eds). *The Multilateral System of Access and Benefit Sharing: Case Studies on Implementation in Kenya, Morocco, Philippines and Peru*. Rome, Italy: Bioversity International, 2012, 9–41.
86. Henry RJ. Evolution of DNA marker technology in plants. In: Robert H (ed). *Molecular Markers in Plants*. United Kingdom: Blackwell Publishing Ltd., 2012, 1–19.
87. Henry RJ. Next-generation sequencing for understanding and accelerating crop domestication. *Brief Funct Genomics* 2012;**11**(1):51–6.
88. Tnah LH, Lee SL, Ng KKS, *et al*. Forensic DNA profiling of tropical timber species in Peninsular Malaysia. *Forest Ecol Manag* 2010;**259**(8):1436–46.
89. Honjo M, Ueno S, Tsumura Y, *et al*. Tracing the origins of stocks of the endangered species Primula sieboldii using nuclear microsatellites and chloroplast DNA. *Conserv Genet* 2008;**9**(5):1139–47.
90. Bandelj D, Jakse J, Javornik B. DNA fingerprinting of olive varieties by microsatellite markers. *Food Technol and Biotech* 2002;**40**:185–90.
91. Brunings AM, Moyer C, Peres N, *et al*. Implementation of simple sequence repeat markers to genotype Florida strawberry varieties. *Euphytica* 2010;**173**(1):63–75.
92. Smith JSC, Jones ES, Nelson BK, *et al*. Genomic approaches and intellectual property protection for variety release: a perspective from the private sector. In: Tuberosa R, Graner A, Frison E (eds). *Genomics of Plant Genetic Resources Volume 1. Managing, Sequencing and Mining Genetic Resources*. London: Springer, 2014.
93. GCDT. A Global Conservation Strategy for Cassava (Manihot Esculenta) and Wild Manihot Species. Rome, Italy: Global Crop Diversity Trust, 2008.
94. Espindola A, Schneider W, Hoyt PR, *et al*. A new approach for detecting fungal and oomycete plant pathogens in next generation sequencing metagenome data utilising electronic probes. *Int J Data Min Bioinform* 2015;**12**:115–28.
95. Studholme DJ. Application of high-throughput genome sequencing to intrapathovar variation in Pseudomonas syringae. *Mol Plant Pathol* 2011;**12**(8):829–38.
96. Achtman M. Insights from genomic comparisons of genetically monomorphic bacterial pathogens. *Philos Trans R Soc Lond B Biol Sci* 2012;**367**(1590):860–7.
97. Lewis T, Loman NJ, Bingle L, *et al*. High-throughput whole-genome sequencing to dissect the epidemiology of Acinetobacter baumannii isolates from a hospital outbreak. *J Hosp Infect* 2010;**75**(1):37–41.
98. Harkins GW, Martin DP, Duffy S, *et al*. Dating the origins of the maize-adapted strain of maize streak virus, MSV-A. *J Gen Virol* 2009;**90**(Pt 12):3066–74.
99. Lefeuvre P, Martin DP, Harkins G, *et al*. The spread of tomato yellow leaf curl virus from the middle east to the world. *PLoS Pathog* 2010;**6**(10):e1001164.
100. FAO. Global information system on PGRFA. 2018. http://www.fao.org/plant-treaty/areas-of-work/global-information-system/en/ (11 January 2018, date last accessed).