

Review

# Can Parentage Analysis Facilitate Breeding Activities in Root and Tuber Crops?

Prince Emmanuel Norman <sup>1,2,3,\*</sup> , Asrat Asfaw <sup>2,\*</sup> , Pangirayi Bernard Tongoona <sup>3</sup>,  
Agyemang Danquah <sup>3</sup> , Eric Yirenkyi Danquah <sup>3</sup>, David De Koeyer <sup>2</sup> and Robert Asiedu <sup>2</sup>

<sup>1</sup> Sierra Leone Agricultural Research Institute, Tower Hill, Freetown PMB 1313, Sierra Leone

<sup>2</sup> International Institute of Tropical Agriculture, Ibadan PMB 5320, Nigeria; d.dekoeyer@cgiar.org (D.D.K.); r.asiedu@cgiar.org (R.A.)

<sup>3</sup> West Africa Centre for Crop Improvement, College of Basic and Applied Sciences, University of Ghana, P.O. Box LG 30, Legon, Accra, Greater Accra, Ghana; ptongoona@wacci.edu.gh (P.B.T.); adanquah@wacci.edu.gh (A.D.); edanquah@wacci.edu.gh (E.Y.D.)

\* Correspondence: P.Norman@cgiar.org or pnorman@wacci.edu.gh (P.E.N.); a.amele@cgiar.org (A.A.); Tel.: +234-8126-217-609 or +232-7661-8454 (P.E.N.)

Received: 26 April 2018; Accepted: 14 June 2018; Published: 27 June 2018



**Abstract:** Controlled pollination in root and tuber crops is challenging. Complex ploidy, cross-incompatibility, erratic flowering patterns, outcrossing, etc., limit the efficiency of breeding progress in these crops. Half-sib breeding that involves random pollination among parents is a viable method to harness genetic gain in outcrossing crops that are problematic for performing planned and controlled pollination. The authenticity of resulting progenies from the half-sib breeding is essential to monitor the selection gain in the breeding program. Parentage analysis facilitated by molecular markers is among the available handy tools for crop breeders to maximize genetic gain in a breeding program. It can help to resolve the identity of half-sib progenies and reconstruct the pedigree in the outcrossing crops. This paper reviews the potential benefits of parentage analysis in breeding selected outcrossing root and tuber crops. It assesses how paternity analysis facilitates breeding activities and the ways it improves genetic gain in the root and tuber breeding programs. Conscious use of complementary techniques in the root and tuber breeding programs can increase the selection gain by reducing the long breeding cycle and cost, as well as reliable exploitation of the heritable variation in the desired direction.

**Keywords:** parentage assignment; DNA markers; pedigree; breeding efficiency; roots and tubers

## 1. Introduction

Root and tuber crops are important crops with increasing food, feed and industrial applications in Sub-Saharan Africa and many other regions of the world [1–3]. Root and tuber crops have tremendous potential to contribute to food, nutrition and income security of many families around the globe, but this has not yet been fully exploited. Variety development through breeding is among the action steps to unlock the potential of these crops for food, feed and industrial applications. Breeding root and tuber crops, however, presents special challenges and heavily relies on the traditional techniques of exploiting the existing variation. Irregularity in flowering time and flowering intensity, cross incompatibility, polyploidy and fertility are among the factors that add challenges to genetic improvement through breeding in these crops [4,5].

Traditional breeding in root and tuber crops often utilizes open pollination by wind or insects to generate sufficient families with unstructured pedigree and subsequent selection of progeny. This method is not efficient for high quality true seed production, identification of parents that

contributed to the progenies developed and estimation of reliable genetic parameters to harness diversity in the desired direction. Many root and tuber breeding programs have generated a significant number of varieties; but these varieties lack reliable and complete pedigree information to monitor the progress [5]. For instance, in potatoes (*Solanum tuberosum* L.), despite the availability of a fairly large amount of pedigree information on the web database, pedigree data of some genotypes are missing [5]. A similar situation also exists for cassava, yam and sweet potatoes. The lack of structured and complete pedigree information in root and tuber crops hinders development and efficient management of databases of crops. The lack of structured pedigree data also suggests that parental control is less exploited to harness selection gain in these crops.

The efficiency of breeders, geneticists, botanists and other specialties in plant science is enhanced through their ability to identify, distinguish and estimate the extent of genetic diversity and relatedness in breeding populations and the extent of parental control to guide the breeding progress. In traditional breeding using phenotypic traits alone, efficiency is diminished due to variations in environment and genotype, inadequate data collection and subjective treatment(s). The advent of molecular marker technology has complemented the traditional breeding technique in adequately identifying genotypes and estimating genetic parameters [6].

A record of reliable pedigrees and parental profiles is useful in modern plant breeding. Reliable pedigree information guides breeders in making prudent decisions on existing divergence in progeny, hybrid vigor and effects of inbreeding depression [7,8]. Parental profile information is useful to classify recombinant and parental genotypes for linkage analysis [9]. Such information is widely applicable for various studies including multiple quantitative traits loci (QTL) mapping [10], association mapping [11,12], resistance inheritance [13] and determination of genetic estimates, breeding values and relationships [14–16]. Demeke et al. [17] and Isenegger et al. [18] noted that clustering in full-sibs and half-sibs is an indication of a link between estimated relationships and known pedigrees.

Information on the application of parentage analysis in major root and tuber crops such as yam, cassava, sweet potato and potatoes has not been well reported. Moreover, unlike potatoes, which have the highest application of the technique, yam is among the crops with the lowest application (Table 1). In yams, parentage analysis has not been done on progenies derived from open pollination and polycross blocks. It was, therefore, imperative to catalogue existing literature on parentage analysis on the root and tuber crops to serve as a basis to improve the breeding strategy of the crops. Pedigree reconstruction is necessary for enhanced breeding and genetic studies. Such information will help with the estimation of reliable heritability and genetic correlation parameters in open pollinated plants to maximize genetic gain and design efficient breeding program. It also ascertains the genetic identity of mislabeled genotypes used in breeding programs. Thus, a good understanding of life history traits and pedigree reconstruction is necessary for the formulation of effective strategies for genetic conservation, management and utilization of genotypes in breeding and genetics programs. In this paper, we review the application and potential of parentage analysis to facilitate breeding activities in selected root and tuber crops.

## 2. Overview of Parentage Analysis in Breeding Programs

Parentage analysis was historically determined using the traditional morphological analysis technique [19]. This technique is limited and cumbersome, especially when candidate parents exhibit similar phenotypes and progenies are generated by open pollination or polycross [19,20]. Moreover, phenotypic parentage analysis is complicated where species used in crosses possess similar traits within each subgenus. These limitations of the phenotypic parentage analysis led to the advent of improved parentage analysis techniques to augment the traditional one.

Parentage analysis in natural populations started with chromosomal polymorphisms [21], followed by allozyme electrophoresis [22] and DNA molecular technology [20]. The DNA fingerprinting technology was originally developed to identify human remains in forensic research [23], but was later utilized for resolution of immigration [24] and paternity conflicts [25]. The technique

has also been useful to accurately resolve genetic relationships in plants [20], as well as other organisms. The efficiency of DNA fingerprinting-led parentage analysis was first discovered in birds in the 1980s, which led to a great paradigm shift in behavioral ecology [26]. Statistical techniques were simultaneously developed for parentage analysis using single-locus polymorphisms such as allozymes [27]. This simple statistical analysis was ineffective at determining parentage using multi-loci polymorphisms [28]. The limitations of the earlier parentage analysis techniques led to the discovery and development of microsatellites and other advanced molecular markers, as well as robust statistical packages [29]. These advances have made determination of genetic polymorphisms in populations and parentage analysis easier in the fields of breeding, genetics, evolution, behavioral ecology and molecular ecology [29,30].

**Table 1.** Inventory of databases and studies on parentage analysis in selected root and tuber crops.

Crop	Type	Database	Number of Studies	Reference
Yam ( <i>Dioscorea</i> spp.)	Tuber	Google Scholar, Research Gate, BioMed Central	12	[31–42]
Potato ( <i>Solanum tuberosum</i> L.)	Tuber	PubMed, Google Scholar Springer Link, Springer,	24	[5,15,19,43–63]
Cassava ( <i>Manihot esculentus</i> Crantz)	Root	Google Scholar, BioMed Central, Research Gate	18	[20,64–80]
Sweet potato ( <i>Ipomoea batatas</i> L.)	Root	Google Scholar, Springer Link	11	[81–91]

### 3. The Concept and Application of Parentage Analysis

Patterns of pollination may influence reproductive capacity of plants, hybridization within and among populations and habitat fragmentation [92]. Pollen dispersal determines the extent of gene flow between cultivated and wild crop species and contaminated seeds. In the past, several indirect techniques were used to trace the physical movement of pollen, including traps [93,94], dyes [95] and the paths of pollinators [92,96,97]. These techniques are limited regarding the actual patterns of fertilization and gene flow. The path of a pollen-forming seed in an open-pollinated environment may be successfully tracked using a direct technique like parentage analysis. Parentage analysis involves DNA profiling of progeny and potential candidate parents, comparing their alleles for identification and confirmation of existing relationships. In this technique, the DNA profiles of progeny and maternal and paternal parents are verified by assigning each allele of tested progeny to its parents. The maternal parent is known, because seeds are harvested from it. If the assigned alleles of progeny align more with a putative male parent than the remaining candidates, it implies a high probability of being the true pollen parent. Where alleles fail to be assigned to one of the putative parents, the relationship between progeny and parents is excluded [30]. The choice of any parentage analysis technique depends on adequate answers that address their corresponding empirical questions. Some of these questions include: (i) Which parentage analysis technique has a higher probability of success regarding the research objective of the study? (ii) Which sampling design is appropriate for a typical parentage analysis experiment? (iii) How are samples collected to ensure desired results from data subjected to an appropriate statistical method? (iv) Which molecular markers provide high levels of polymorphism per locus? The general roadmap on the application of the parentage analysis in natural and artificial population of organisms is elaborately illustrated in Jones et al. [29]. According to these authors, the type of parentage analysis to employ depends on the sampling techniques applied. When putative parents are mated to produce progenies in a breeding population, all parentage analysis techniques are possible. However, the absence of one or more putative parents in the mating population limits the number of applicable parentage analysis

options. For instance, parental reconstruction becomes the only applicable technique where several half- or full-sib groups are identifiable and sampled. For small family groups, parental reconstruction, or sibship reconstruction, or assessment for multiple mating within family groups is applicable. Multiple mating (mixed parentage) occurs where more than one putative male parent contributes pollen to fertilize the female stigmatic organ leading to fruit and seed development. Its influence on sample size depends on ploidy complexity, cross compatibility, pollen viability, flowering intensity and synchronization, as well as biotic and abiotic factors. In breeding programs where progenies are not collected in family groups, only the sibship reconstruction technique is applicable. Successful utilization of parental reconstruction or sibship reconstruction technique allows comparison among reconstructed parental genotypes for their mating patterns.

Spanoghe et al. [8] successfully applied parentage analysis in European potato cultivars for pedigree validation, as well as potential parental assignment. Accordingly, a root or tuber breeding program can make informed decision with the application of parentage analysis in instances where the pedigree is either in doubt or incomplete for the materials in the breeding program.

Parentage analysis is widely applicable in assessing reproductive success, mating patterns, kinship, fitness in natural populations and in developing highly polymorphic molecular markers of multilocus genotypes [98,99]. Parentage analysis is useful in multiple mating breeding systems where individuals of one sex mate with two or more partners of the opposite sex. It enhances investigation of the genetic mating system of organisms by evaluating the actual reproductive success of paternity and construction of pedigrees [99,100]. Further, it is used to confirm monogamy in some species [101] and also those that exhibit extra pair copulation [102]. Parentage analysis resolves difficulties in direct mating systems [103]. It also enhances heritability and genetic correlation estimates in open pollinated plants needed for evaluation of expected genetic gains and designing of breeding programs. The parentage analysis technique is a useful tool in the protection of Plant Breeders' Rights (PBR). The PBR comprise a globally-accredited system that gives breeders intellectual property rights for unique technology such as a new variety or varieties developed or adapted for commercial or various end uses [104]. The PBR accord breeders with end point royalties (EPR) utilized to support and sustain breeding activities. The use of DNA profiling resolves issues of genetic mixtures, mislabeling, mutations and outcrossing in natural populations that often confound efficient selection in breeding programs. The use of markers that are highly informative, high-throughput and reliably reproducible clearly and robustly determine genetic identity based on morphological attributes, aiding in the protection of PBR [104].

Despite its relevance and wide application, the success of parentage analysis depends on the sampling design used, choice of molecular markers, family structure species, pollen and seed dispersal [29,30,92,105]. Additionally, various unfavorable biological attributes of organisms and use of inadequate markers may also limit the potential of parentage analysis.

## 4. Factors Affecting Parentage Analysis

### 4.1. Sampling Techniques

Parentage analysis is often used in systems that allow the collection of candidate parents with the assumption of the presence of a sample of adult genotypes. An ideal parentage study involves large family sets of progeny obtained from the maternal parent and a complete set of potential candidate paternal parents in the breeding population [106]. This permits application of all parentage analysis methods with a high probability of success. However, the ideal situation is very difficult in natural populations. An excessively large number of genotyping tests for assaying many progenies per family may be needed depending on the goals of the study. Nonetheless, parentage inference has good prospects with exclusion and assignment techniques remaining useful options if candidate parents are present.

Sampling of specimens is the hardest aspect of parentage analysis; however, it is also the most important criterion required for good analysis. Care must be taken to maintain complete samples from the field. Knowledge of one of the parents of the offspring makes parentage analysis easier and more powerful than when both parents are unknown. Thus, use of adequate sampling design and collection of offspring from a known parent facilitates parentage analysis targeted at answering various questions of interest to the researcher.

#### 4.2. Molecular Marker Systems

Generally, molecular markers that identify heritable variations among genotypes are useful in parentage analysis. The identification of appropriate markers determines their successful application in parentage analysis [105]. A reasonable number of highly polymorphic molecular markers or a large number of the low to moderate polymorphic types is needed for successful parentage analysis [29].

Despite the various markers utilized in parentage studies, few useful ones have been earmarked such as microsatellites, single nucleotide polymorphisms (SNPs) and amplified fragment length polymorphisms (AFLPs). Of these, microsatellites are often utilized in parentage analysis due to their repeatability, high polymorphism, co-dominance and PCR-based attributes [105]. The hyper-variability and codominant inheritance attributes of microsatellite alleles aid molecular identification of organisms utilized in mating system experiments including parentage analysis [92]. Microsatellites are DNA repeats with 2–6 base pair sequences recurring at least 12 times. They are also referred to as short tandem repeats (STRs) [107] or simple sequence repeats (SSRs) [108]. Most microsatellites utilized in parentage analysis are di-, tri- or tetra-nucleotide repeats, with repeated motifs of two, three or four base pairs, respectively. Of the three, tetra-nucleotide microsatellites are preferred since different alleles are easily separated on a gel, compared to di-nucleotides, which exhibit less ‘stuttering’ with slightly larger or smaller bands than the true allele on the gel [29,109]. Stuttering is a process whereby stutter bands with different sizes from normal PCR products are produced in multiples of the short repeat unit (1–2 bp) [110,111]. Stuttering is caused by strand slippage, replication slippage and the structure of short tandem repeats (STRs) during polymerase chain reactions producing a highly polymorphic stutter product with the wrong number of repeats. The presence of stutter bands in SSRs makes scoring cumbersome, producing quasi-scoring in ladders lacking prominent bands [111]. Moreover, although multiplexing permits moderate-level throughput in microsatellites, these markers are poorly transferable across species, necessitating their investigation and optimization prior to use for parentage analysis in a new species [112].

A typical paternity analysis using microsatellites involves, firstly, sampling and genotyping parents to obtain microsatellite loci. Progeny obtained from known maternal parents are genotyped at these loci. A diploid progeny exhibits two alleles; of which, one allele is contributed by the seed parent and the other by the pollen parent. The maternal allele at each locus of the genotyped progeny is identifiable. The progenies are tested against a pool of pollen parents to determine their paternity. The exclusion probability of 97–99% determines paternity with a high degree of confidence. The exclusion probability is the probability of excluding a misrelated pollen parent as the father of a progeny, where knowledge of maternity exists. Allelic mismatch of pollen parents implies that pollens that formed the progeny were obtained outside the studied area. Genotyping of progeny families provides a good understanding of the distribution of pollination distances, directions and the emergence of successful reproductive paternity. In root and tuber crops, microsatellites have been utilized to determine the progeny-paternity relationship (Table 1). Microsatellites have been used for parentage analysis in yams [32–39], potatoes [44–46,48–50,52,53,55], cassava [63–73] and sweet potato [76–82,84,85].

Besides microsatellites, other useful markers for parentage analysis are SNPs, AFLPs [113,114] and diversity array technology (DArT) markers [6]. The AFLPs and SNPs have many loci, with each locus exhibiting two alleles per locus and low polymorphism. The alleles of SNPs are codominant, whilst those of AFLPs are dominant, implying that AFLPs disallow the separation of the heterozygote

from one of the homozygous groups. The application of AFLP markers in parentage analysis has only been reported in plants [6]. One of the merits of AFLP markers is the flexible use of its commercial kit for parentage analysis of many organisms with little need to develop new ones [29]. This contrasts SNP markers that are dependent on known DNA sequences with polymorphic nucleotide positions. The accessibility of SNP markers and the availability of hundreds of loci of model organisms are more easily obtainable compared to other marker systems. Moreover, SNPs are more easily assayed per locus compared to microsatellites, leading to the futuristic prediction of their preferential use for parentage analysis [115]. However, the low per-locus polymorphism of SNPs limits some data analysis techniques such as Sanger's method [116] and *in silico* SNP discovery using mining of SNPs within expressed sequence tag (EST) databases followed by PCR-based validation [117]. Single nucleotide polymorphism discovery is more difficult in crops with complex genomes than those with simple genomes due to their highly repetitive nature [118]. Despite the gene discovery technique, the low per-locus polymorphism in conserved genic regions, low-copy noncoding regions and intergenic spaces are independent of the detection of gene-based SNPs [119].

As molecular paternity research progressed, it became imperative to develop markers with wider genome coverage and higher level throughput for the increased resolution and speed required for various applications. Jaccoud et al. [120] suggested DArT as a promising alternative that satisfies the throughput, genome coverage and transferability criteria. DArT is a high-throughput complexity reduction technique involved in the hybridization of fluorescent DNA probes to targeted DNAs spotted on a microarray [121]. The DArT assays produce whole genome fingerprints used for scoring the presence or absence of DNA fragments in genomic sequences generated from genomic DNA samples through complexity reduction. The DArT genome profiles enable fast mapping of QTLs, accelerate the introgression of a selected genomic region into an elite genetic background and guide the assembly of many different genomic regions into improved varieties. The number of markers detectable by DArT depends mainly on the level of DNA sequence variability in analyzed samples and the complexity reduction technique used [122]. The merits of DArT include a high multiplexing level, simultaneous genotyping of several thousand loci per assay and its simultaneous screening of thousands of polymorphic loci lacking prior genome sequence information; additionally, DArT provides hundreds to tens of thousands of highly reliable markers [120], as well as good genome coverage information [123]. Moreover, sequences of DArT markers are easily accessible compared to AFLPs, making them a preferred technique for non-model species [124]. However, DArT is limited in some applications due to high technicality in preparing genomic representation for the target species, cloning, data management and analysis. Other demerits include its requirement of robust software (DArTsoft and DArTdb) for analysis and the dominance of DArT markers (present or absent) or differential intensity [125].

Prediction and selection of improved plants on a genome-wide scale requires large populations with dense molecular markers across the genome [126]. This is achieved through a technique known as genotype-by-sequencing (GBS). GBS is an enzyme-based complexity reduction technique that utilizes restriction endonucleases targeted at a small portion of the genome and DNA barcoded adapters for the production of multiplex libraries of samples used in next-generation sequencing (NGS). The GBS technique is noted for its robustness across many plant species and capability to produce 10,000s–100,000s of molecular markers [126,127]. It is flexible regarding species, populations and research objectives, indicating its suitability for plant genetics studies. Another merit of this technique is that both marker discovery and genotyping are completed at the same time. Moreover, GBS aids the exploration of new germplasm or species without prior discovery and characterization of polymorphisms. Reanalysis of raw sequences from GBS uncovers new polymorphisms, annotated genes, etc. [126].

Based on the above facts, it is clear that DArT and microsatellites are useful and preferred for parentage studies. This partly corroborates earlier prediction of the continued usefulness and preference of microsatellites for parentage analysis [29,128]. The AFLPs may be less popular and less

utilized for parentage determination, except for species lacking justification to develop microsatellites. With increasing availability of genomic data, the potential use of SNPs for exclusion or parentage assignment in model systems is increasing, though they might not be applicable in some molecular work. It is also probable that the use of tightly-linked loci in experiments with the linkage phase will increase the applicability of SNPs in parentage analysis [129]. The linked SNPs are dubbed super-loci, with the potential of many alleles under a sufficiently low rate of recombination that permits stable inheritance of haplotypes [29]. Similarly, it is also probable to increase the use of microsatellites in species with low polymorphism [29]. Besides the super-locus scenario, one should avoid loci that are in linkage disequilibrium to adhere to the independence among the loci statistical assumption of parentage analysis [30].

### 4.3. Analysis Method

Various analysis methods are available for the implementation of the parentage analysis in the breeding of crops whenever applicable for the program. However, which method is more appropriate to maximize the genetic gain in a breeding program is most important when it comes to the crop improvement. Jones et al. [9] provided the pros and cons of each of the techniques for parentage analysis. We briefly summarize the methods here with respect to their possible application in facilitating root and tuber crop breeding.

#### 4.3.1. Exclusion Technique

This technique involves comparison of the alleles of progenies with their putative parents followed by elimination of parents lacking matching alleles at all loci with tested progenies. The exclusion technique leverages the Mendelian inheritance principle that is typical in sexual diploid plants, as the individual progeny has at least one allele per locus the same as each of its parents. This technique is more feasible to dissect the parentage issues in a population with fewer progenies and candidate parents. It requires highly polymorphic genetic markers. Larger population size, genotyping error, mutation in progenies and complex ploidy are among the factors affecting its successful implementation.

#### 4.3.2. Categorical Allocation Technique

Categorical allocation involves assigning of progenies to putative non-excluded parent with the highest likelihood or posterior probability of consideration as the true parent. This technique is applicable for assignment of non-excluded putative parents that lack assignment by the complete parental exclusion technique. The technique is also useful in resolving scoring errors or mutations, as well as in estimating confidence.

Categorical allocation measures the logarithm of likelihood ratio (LOD score). It is the quotient of the likelihood of the putative parent(s) of a given progeny and the likelihood of unrelated individuals. The progenies are unassigned when the LOD score is negative or zero.

#### 4.3.3. Fractional Allocation Technique

This technique involves assigning a fraction or proportion of progenies to non-excluded putative parents based on their likelihoods of parentage. The fraction of progenies assigned to a putative parent is proportional to the likelihood of parenting them compared with non-excluded putative parents. This works on the assumption that the genotypes of all parents in the population, as well as one parent of the specified progeny are known.

The fractional allocation is estimated as the proportion of progeny ( $O = k$ ) assigned to putative male  $j$  ( $MP = j$ ) conditional on female  $i$  ( $FP = i$ ). This is summarized in the equation below.

$$\hat{F}_{ij} = \sum_k X_{ik} P(MP = j | FP = i, O = k)$$

where  $X_{ik}$  is the proportion of non-excluded female parents ( $i$ ) and progenies ( $k$ );  $P$  indicates proportion;  $MP$  and  $FP$  represent male parent and female parent, respectively.

#### 4.3.4. Full Probability Parentage Analysis Technique

The full probability technique involves simultaneous estimation of patterns of parentage and other desired population level variables using models. The technique offers better data mining through incorporation of available uncertainties in determining desired variables. The full probability models are suggested for the estimation of population-level variables of interest, particularly where parentage analysis methods show <95–100% confidence in assigning progeny to parents [29].

#### 4.3.5. Parental Reconstruction Technique

In this technique, the genotype of each progeny in full- or half-sib families is used for parental reconstruction. The parental reconstruction is applicable where at least one parent is shared among the full- or half-sib progeny arrays. Genotypes of unknown parents are determined using parental reconstruction, by matching alleles of progeny to a set of candidate parents and verifying those using assignment or exclusion techniques [130]. However, the parental reconstruction technique is not applicable in situations lacking large sets of related progeny. This technique often works on the assumption of a minimum number of parents, maximum likelihood or Bayesian approaches.

#### 4.3.6. Sibship Reconstruction Technique

This analysis technique involves the collection of large groups of full- or half-sib progenies that lack putative parents. The algorithms of this technique group progenies into unique clusters of full-sibs, half-sibs and unrelated progenies using patterns of relatedness or maximum likelihood techniques. Successful reconstruction of progenies is contingent upon adequate identification of clusters of half-sib or full-sib progenies.

All techniques of parentage analysis are possible where candidates are identifiable. In the absence of candidate parents, the options are fairly limited. For instance, for a breeding program in root and tuber crops employing a half-sib breeding method with an identifiable and comparatively large family size, the parental reconstruction technique is useful to establish the relationship. Several authors have also reported various techniques used to investigate proof of multiple mating in small family sets [131–134]. Sibship reconstruction is the only appropriate technique applicable where offspring are not collected per family.

#### 4.4. Genotyping Errors, Mutations and Null Alleles

The occurrence of genotyping errors, mutations and null alleles negatively affects parentage analysis even where appropriate sampling design and molecular markers are used. Genotyping errors and mutations are particularly the most critical problems that account for inconsistent results in parentage analysis. Genotyping errors arise from genotypic misreads, amplification failure or spurious production of false results. Mutations are allelic alterations in the progeny compared to those inherited from their parents. The existence of both genotyping errors and mutations produces incompatibilities between true candidate parents and their progeny. This necessitates use of good quality control measures that facilitate successful parentage analysis [135,136]. For instance, in the exclusion or parental reconstruction technique, parents are excluded or an additional parent is invoked provided the result is verifiable by one or more additional loci. This technique is probably excessively conservative, consequently producing several incorrect inclusions. It requires care for the assurance of sufficiently adequate power. Unlike the conservative technique, full probability models and sibship reconstruction methods of the categorical or fractional assignment build a model of error into likelihood estimates or posterior probabilities [137,138].

Lack of allelic amplification, or “null alleles,” is a problem in parentage analyses [139]. Null alleles account for mismatches between parents and progenies by presenting truly heterozygous genotypes

for the null allele as homozygous. These alleles are mostly improperly handled by parentage analysis programs. The detection of null alleles is often evident as a deviation from Hardy–Weinberg equilibrium at the null-bearing locus or as a non-Mendelian pattern of segregation in known sets of families [29,140]. Since null allelic loci are often detectable, they are apparently not a major threat to parentage analysis. A well-articulated study on the validity of null alleles is reported by Jones and Ardren [30].

This notwithstanding, Wang has developed a model to resolve issues of allelic dropouts, stochastic errors, genotyping error, mutations and null alleles [137]. Wang’s model works on a similar principle as the Kalinowski et al. model by fitting systematic allelic dropouts (null alleles) and microsatellite mutations and scoring errors (stochastic error category) [141,142]. The technique allows variation among loci for the rate of error. Wang’s model has been utilized for sibship reconstruction [137] and full probability parentage analysis [143]. The model is, however, still under examination to determine incremental accuracy in handling errors and null alleles for parentage analysis.

#### 4.5. Family Structure of Candidate Parents

The family structure sometimes constitutes candidate parents, their relatives or desired progeny. Besides parent-offspring relationships, most inference methods work on the assumption that candidate parents and progeny are unrelated, presenting difficulty for breeding populations with an existing family structure [144,145]. Several authors have noted the effects of the family structure of candidate parents in parentage analysis [29,146,147]. Parentage analysis often only produces the desired impact where very close relatives of progeny are added to candidate parents. In reality, non-exclusion of full progeny siblings exhibits higher likelihoods of parentage than the true parents [148]. The addition of many close progeny relatives to candidate parents necessitates the use of complete exclusion or parental reconstruction methods, due to their comparative insensitivity to family structure and reliable diagnosis of true patterns of parentage [30].

#### 4.6. Species Effect

Paternity assignment is also influenced by the ploidy level of the studied crop species. Paternity assignment or exclusion is more direct and simpler in diploids than polyploids since the developed progeny inherits one paternal allele from the putative male plant [92]. Genotypic scoring is also simpler in diploids than polyploids since each individual candidate exhibits one allele (homozygote type) or two alleles (heterozygote type) at each locus. Generally, about 30–50% of plant species are polyploids [143,149]. This presents an acute paternity assignment limitation especially in highly polyploidized genomes with complicated inheritance. However, the genomes of some polyploids are highly diploidized due to ancient polyploidy events, resulting in the formation of bivalent chromosomes at meiosis. This is typical of octoploid strawberry [150], as well as some genotypes of polyploid yams, cassava, potato, sweet potato, etc. [4]. In polyploids with polyploidized genomes, parentage assignment is complicated due to cumbersome scoring of alleles and inference of their polysomic nature of inheritance [151,152]. The parentage assignment complications in polyploids are nowadays resolved by the transformation of polyploid codominant genotypes into pseudodiploid-dominant genotypes [153].

#### 4.7. Pollen and Seed Dispersal Mechanism

Pollen- and seed-mediated gene flows are the two major determinants of genetic structure and diversity in plants [154]. Pollen-mediated gene flow contributes to plant genetic structure and diversity through immigration, hybridization and introgression, whereas seed-mediated gene flow functions in the colonization of new habitats. Determination of the extent of pollen- and seed-mediated gene flow is imperative for the prediction of its effects on the genetic structure of progenies generated in mating designs [155]. Gene flow by pollen in open pollination and polycross mating schemes has been studied in various plant species using molecular markers and paternity analysis [4,81,156]. These studies

describe the pathway of effective gene flow between the pollen donor and the pollen recipient plants, leading to botanic seed development. In root and tuber crops, however, a dearth of well-articulated information on the extent of pollen-mediated gene flow exists especially for the polycross or open pollination mating design. This information is vital since it represents the realized gene flows by pollen and seed [157]. Studies on open-pollinated or polycross-derived seeds permit determination of the rate of pollen immigration, the pattern and distance of pollen dispersal, paternal and maternal fertility levels and the relationship with the plant phenotype and genotype.

Dispersal of seeds results in seed or seedling mixtures where both parents are unknown. Such a situation makes parentage analysis difficult. Seeds or seedlings may be obtained under different scenarios such as from one parent known as a pollen or seed parent; two parents, known as pollen and seed parents; or no parents, known as immigrant seeds [92]. Most plant species are cosexual with either separate, monoecious male and female flowers or hermaphrodite flowers (i.e., possessing both male and female organs). Assignment of parentage of cosexual plants to either the seed- or pollen-parent using nuclear microsatellites is impossible. Dioecious plants lack such a problem since one of the parents is known [158–160].

Yam is a typical tuber crop that is mostly dioecious, with the staminate (androecious) and pistillate (gynoecious) organs borne on separate plants [161]. A few monoecious and trimonoecious types also exist [162]. Yams have varying flower, fruit and seed attributes depending on species. Detailed distinctions of these attributes have been reported by various authors [4,162–164]. The sticky nature of pollen promotes its pollination by insects [4]. Yam fruits are dry dehiscent trilobular capsules, each containing 2–6 seeds per capsule and measuring about 1–3 cm. Seeds of most species are winged to facilitate wind dispersal [4]. Seeds of some species such as *Dioscorea rotundata* are small, flat, light, lenticular and consist of a small embryo surrounded by a relatively large endosperm. At dehiscence, seeds are dispersed from dry locules. The shape and wing patterns of seeds depend on the species [163]. For instance, in *D. rotundata*, the botanic seed is completely encircled by the wing, whereas in *D. bulbifera* or other species, the wing is found on one or both sides of the seeds, respectively [163].

Cassava is a protogynous monoecious root crop with male and female flowers borne on the same plant. Cassava pollen is small and sticky, which favors pollination by insects such as wasps (mainly *Polistes* spp.) and honeybees (*Apis mellifera*). A comprehensive review of the reproductive biology of cassava has been reported [165].

Sweet potato (*Ipomoea batatas* (L.) Lam.) is an annual, autogamous and outcrossing root crop with both male and female organs borne in the same flower [85]. The pollen grains are spherical and mainly dispersed by bees for pollination. A comprehensive description of the floral biology of the crop is found at [166].

Potato (*Solanum tuberosum*) is a monoecious tuber crop with both male and female organs borne in the same flower. Natural pollen-mediated gene flow in potatoes is mainly by insects including bumblebees (*Bombus impatiens* Cresson and *B. terrestris* L.) [167] and pollen beetle species *Meligethes aeneus* Fabricius [168] with small pollen flow by wind [169]. Seed-mediated gene flow is noted to be facilitated by birds [170] and small mammals [171]. A comprehensive description of the floral biology of the crop is reported by Canadian Food Inspection Agency (CFIA) [170]. The flowers, fruits and seeds of yam and cassava are shown in Figures 1 and 2 (International Institute of Tropical Agriculture (IITA)/Prince E. Norman).



**Figure 1.** Floral, fruit and seed photos of yam. Photo source: International Institute of Tropical Agriculture (IITA)/Prince E. Norman.



**Figure 2.** Floral, fruit and seed photos of cassava. Photo source: IITA/Prince E. Norman.

Mitochondria or chloroplast organelle DNA has been used to distinguish the maternity and paternity of organism with uniparental inheritance [172,173]. However, both mtDNA and cpDNA lack sufficient variability to identify individuals. An improved technique utilized for identification of seed parents is seed tissue obtained by maternal inheritance such as endocarp [174,175], pericarp [176] and seed wing [177] tissue or megagametophyte tissue of conifers [178]. However, this method is limited to intact seeds, since seedlings generally lack maternal tissues.

Parentage analysis may exhibit any of the three outcomes including identification of (i) none of the parents, (ii) one of the parents or (iii) both male and female parents of the progenies [179]. The first scenario is typical of seed immigration, whereas the second represents pollen immigration. However, assuming that the closest parent in the second and third scenarios is the maternal parent, a conservative distance of seed dispersal occurs [179].

Based on the above information, it can be deduced that a good understanding of the dynamics of natural populations is obtained using seedling plants, since they give additional information such as the rate of seed immigration, the distance of seed dispersal and the pattern of seed dispersal. The determination of gene flow in seedlings also permits comparisons between the dispersal distances and patterns of pollen and seeds. The effects of biotic and abiotic factors on pollen and seed dispersal, as well as genetic factors on pollen and seed development may also limit the extent of variability expected in parentage analysis.

## 5. Potential of Parentage Analysis in Root and Tuber Breeding Programs

Population improvement is an important breeding strategy to harness genetic gain in root and tuber crops [4]. As breeding populations increase, issues of genealogy overlaps arise. Sometimes, when breeders release varieties, farmers give the varieties local names. Tracking of these varieties with time becomes difficult using phenotypic markers alone. Breeders often use routine molecular parentage analysis to resolve these issues and maintain useful allelic and gene diversity in their breeding populations. Maintenance of high genetic variability from generation to generation facilitates long-term sustainability of root and tuber breeding programs. Moreover, parentage analysis ascertains existing changes in allelic frequency of genotypes and how much is maintained in progenies.

Tracking pedigrees is crucial in root and tuber breeding programs. It provides the building blocks upon which heterotic groups are formed. Pedigrees are tracked by the name of the original development cross from which selection is made. Parentage analysis enhances root and tuber breeding activities through establishment of these heterotic patterns, which contribute to improving the heterosis in the new hybrids developed. This improvement leads to a gain in the selection of parents with desired complementary alleles for improvement of target traits.

The knowledge of the ploidy level of the parent, fertility-regulating mechanism of putative parents, genes controlling traits of interest, etc., also guides the breeder in planning cross combinations and the choice of breeding strategy to maximize genetic gain. Molecular parentage analysis could help breeders to make informed decisions in a more efficient, accurate, creative and rapid manner [85–87].

Parentage analysis has many potential applications in root and tuber breeding programs. Some of the applications include: (i) Increased parental control in the breeding program: Adequate identification of parents contributing alleles to the next generation improves the breeding efficiency. (ii) Reduced cost and time of maintaining parents in the breeding program: Adequate genetic profiling of parents and progenies guide the selection of genotypes with desired traits at an early stage of breeding, thereby saving time, cost and labor. (iii) Reliable estimation of paternal breeding values in the half-sib family: Adequate molecular analysis complements conventional breeding values in the half-sib family, especially where the expression of alleles is masked by epistatic, dominance and other genetic and environmental factors. (iv) Parentage analysis minimizes pollination and labeling errors in breeding program; the effects of pollination and mislabeling errors are so costly in a breeding program, reducing the validity of progeny developed. Parentage analysis reduces these effects by adequately tracing the identities of progenies to their parents. (v) Assess pollen donor rate and pollen movement in breeding blocks: Pollen movement and amount of pollen load on the female stigmatic surface are influenced by several factors, including spatial arrangement and distance between male and female plants, genetic, biotic and abiotic factors. Successful parentage analysis helps to ascertain pollen donor rate and dispersal. (vi) Estimate the level of inbreeding in the breeding programs: In monoecious and perfect flower genotypes of roots and tubers, parentage analysis helps to determine the extent of inbreeding in the progeny developed. (vii) Assess incompatibility in breeding programs: Good knowledge of ploidy status and other fertility-regulating mechanisms in parents guides parental selection. Thus, parentage analysis helps to uncover any genetic aberrations limiting breeding progress. (viii) Estimate of the genetic effects including general combining ability (GCA) and specific combining ability (SCA) in the breeding program: In half-sib breeding, genetic estimates are done by assumption. Such assumptions are less accurate and therefore misleading. However, parentage analysis contributes to generating reliable pedigree and progeny genetic data, which invariably aid reliable estimation of genetic effects. (ix) Assess the rate of pollen contamination in breeding blocks: Pollen contamination may occur in both controlled and artificial mating schemes through insects, wind, human and other agents of pollination. Routine parentage analysis helps to assess whether pollen used is pure and the progeny developed is true to type. (x) Identify parents that produce superior progenies for selection and high quality seed production.

Complete parentage assignment has higher prospects where maternity is known, progeny samples are collected in family groups and complete samples of putative male parents are represented. The probability of adequate parentage assignment decreases with decreasing completeness of the dataset. The use of higher resolving power molecular markers counterbalances reconstruction of less ideal samples. The number of loci and the locus heterozygosity determine the maximum number of samples needed for parentage assignment. A priori knowledge of kinship relations is, therefore, needed to guide the type of parentage assignment technique to implement in root and tuber population development.

Root and tuber improvement for desired traits is crucial to societies in Africa and elsewhere making their livelihoods directly or indirectly from these crops. The future prospects of genetic improvement of long-duration root and tuber crops should target a good understanding of SWOT

analysis of progeny and parent analyses using genomic markers, the development of a robust schematic pathway of parentage analysis in root and tuber crops as typified in aquatic animals [180], the reduction of the long breeding cycle for desired consumer and market traits and improved collaboration with relevant stakeholders in the crops' product value chains. Another prospect of enhancing parentage analysis relates to improvement in translational genetics, routine optimization of DNA sequencing protocols of recalcitrant root and tuber species, effective and efficient quality control and development of more user-friendly, but highly efficient parentage analysis software packages. Despite these prospects, more efforts are needed in root and tuber crops regarding the contribution of parentage analysis to improve estimates of genetic parameters, the extent of pollen and seed gene flows, the determination of the paternity of progenies harvested from known parents in random open pollination or polycross blocks especially in neglected root and tuber crops and the determination of the gap between the actual population and the expected population sizes.

Efficient reconstruction of genealogies of root and tuber crops also requires more efforts. Genealogy reconstruction may be limited by parent-offspring marker genotype data incompatibility using the paternity exclusion technique. Buteler [83] noted two limitations of paternity assignment in root and tuber crops using Mendelian segregation probabilities: ambiguous progeny genotypic profiles and statistical bias in favor of homozygotes. The ambiguity favors homozygous individuals with higher likelihood for the particular allele. However, the bias can be reduced by increasing the number of genetic markers.

## 6. Current Application of Parentage Analysis in Root and Tuber Crops

Molecular parentage analysis has been extensively utilized for the discovery of pedigree errors in breeding populations developed using natural and artificial mating schemes [30,59,181,182]. In potatoes, many researchers have noted conflicting level of accuracies in pedigree assignment. For instance, Douches et al. [183] discovered mismatches in parental assignment in their isozyme-based marker pedigree investigation. Similarly, using 17 SSR markers, [8] reported some discrepancies in the pedigree records of 577 potato varieties. The quest to resolve the discrepancies in pedigree data and understand the level of accuracy led to the development and utilization of other marker types. The development of an Infinium SNP array from sequence data of six potato varieties including Bintje, Kennebec, Premier Russet, Shepody, Snowden and Atlantic [184,185] has facilitated the genotyping of several elite North American potato germplasm. However, pedigree errors or mismatches were also observed in the SNP-based markers [59].

Parent-progeny trios' relationships, as well as the paternity of some varieties generated with bulk pollen were determined in a genome-wide SNP analysis in potato [59]. A trio consists of two parents and one progeny assessed at a time during parentage analysis. A total of 719 tetraploid potatoes were genotyped with the Infinium SNP array to produce 5063 high-quality markers. The curated information was used to verify pedigree records and establish parent-progeny relationships. Findings revealed pedigree errors in some trios. Of the 198 parent-progeny trios studied, 182 trios had accurate pedigree match (mean 0.02%, range 0–0.73%), while 16 trios exhibited pedigree error or mismatch (mean 16.4%, range 7–30%). Besides the 719 curated potato dataset, 24 genotypes were utilized to establish an alternative parentage hypothesis. Of the 24 genotypes, 17 pedigree modifications were suggested to have occurred. Of the 17 pedigree modifications, four were not due to pedigree error, but an unknown male, since the female parent was either open-pollinated by All Red or hybridized with bulked pollen of genotypes Allegany, AmaRosa and Purple Pelisse.

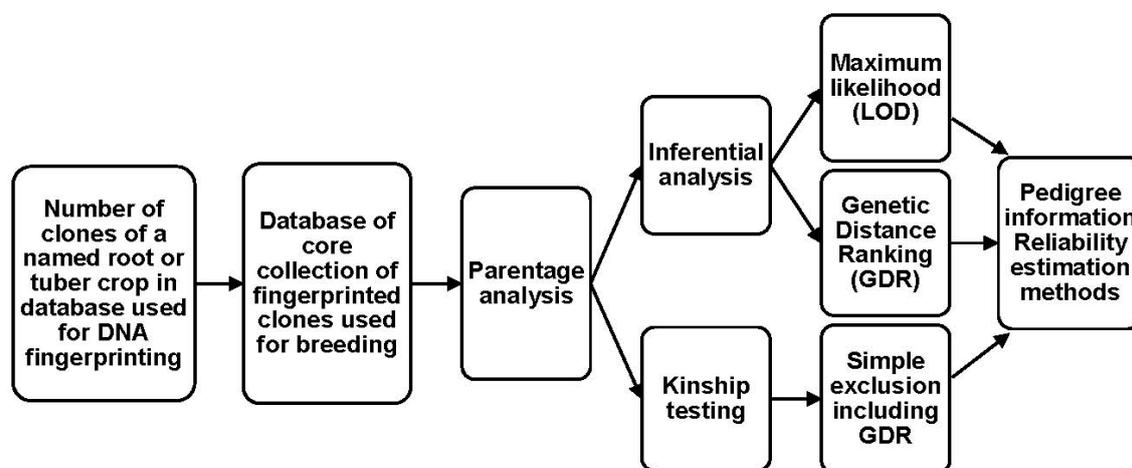
The exclusion analytical technique has been used to exclude candidates with a higher fraction of segregation mismatches at homozygous loci [30]. The same technique has been utilized in root and tuber crops. In potato, for instance, [8] identified segregation mismatches in the pedigree information using 17 SSR markers and the analytical exclusion technique. However, they noted that the markers used were unreliable. The segregation mismatch identified using at least 5000 SNP markers was very informative with unambiguous determination of the pedigree and level of accuracy [59]. Furthermore,

the authors remarked that as the inbreeding coefficient increases, the use of homozygous loci alone to check segregation mismatch becomes less effective [59]. In yams, little is known with regards to segregation mismatches in the pedigree information of existing germplasm. Currently, studies on parentage analysis in controlled (North Carolina 1) and polycross-derived progeny using DArT seq are being conducted at IITA. The findings of this study will be published as soon as the analysis and reports are completed.

Based on the above information, it is clear that parentage analysis is possible in roots and tuber crops. However, the existence of genetic bottlenecks with regards variable ploidy may contribute to segregation mismatches in pedigree records of genotypes of these crops.

## 7. Generic Framework of Parentage Analysis in Root and Tuber Crops

The flowchart in Figure 3 shows the experimental procedure of parentage analysis for a typical root or tuber crop. Accordingly, from many germplasm accessions in a root or tuber breeding program, molecular markers can be used to confirm phenotypic variations observed in-field. The database of the core collections of fingerprinted genotypes is useful in breeding. At this preliminary assessment stage, informative markers that contribute to most of the variability are identified and the genetic relationships or similarities of clones determined using various coefficients such as Jaccard, Dice, etc., and their rankings are based on their genetic distances. Putative parental candidates with desired complementary traits used in population development are subjected to parentage analysis to determine the paternity or maternity, or both, for the tested progenies. The choice of the type of parentage analysis depends on the sampling technique and other factors previously discussed. The pedigree information obtained could be subjected to inferential analysis using threshold levels for the genetic distance ranking (GDR) or logarithm of odds score (LOD); and simple exclusion for the kinship testing option.



**Figure 3.** A flowchart of the experimental procedure of parentage analysis of a typical root or tuber crop. LOD, logarithm of odds.

## 8. Conclusions

Successful application of DNA-informed breeding techniques in root and tuber crops in complementarity with conventional population development techniques, and other recent advances in genomics and phenomics tools can accelerate the genetic gain compared to using the conventional breeding technique alone. Utilization of molecular technique in parentage analysis and realization of its merits have contributed to a greater impact on root and tuber breeding programs. The effectiveness of the molecular techniques in parentage assignment contributes to resolving many genetic bottle necks associated with half-sib breeding using polycross and open pollination mating designs compared to the full-sib breeding using controlled pairwise matings. Moreover, successful use of parentage

analysis improves half-sib breeding efficiency by increasing selection gain or accuracy, leading to higher selection pressure (<3%), more accurate estimation of genetic parameters, sib screening, the protection of breeders' rights, the utilization of new mating designs, the selection of new traits, the identification of the genetic identity of candidate genotypes and the reduction of time and costs. Efficient parentage assignment also establishes the extent of pollen and seed gene flow necessary for prediction of its effects on the genetic structure of progenies. These merits make population improvement of organisms, including niche and neglected species, simple and flexible.

**Author Contributions:** P.E.N. and A.A. conceived of the idea and drafted the manuscript. All the authors contributed to writing the article, read and approved its submission.

**Funding:** The authors are grateful for the financial aid from the International Development Research Centre (IDRC) under the IDRC/CORAF-WECARD/IITA sub-grant agreement for developing capacity for the Agricultural Research in Sub-Saharan Africa (PJ-2126) research grant to pursue a Ph.D. in Plant Breeding at the West African Centre for Crop Improvement (WACCI), University of Ghana. The funding support of the Bill and Melinda Gates Foundation (BMGF) through the AfricaYam project (OPP1052998) for the field work is dully acknowledged.

**Acknowledgments:** We are also grateful for the institutional and technical support of the International Institute of Tropical Agriculture (IITA).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Akanbi, W.B.; Adeboye, C.O.; Togun, A.O.; Ogunride, J.O.; Adeyeye, S.A. Growth, herbage and seed yield and quality of *Telfairia occidentalis* as influenced by cassava peel compost and mineral fertilizer. *World J. Agric. Sci.* **2007**, *3*, 508–516.
2. Iyagba, A.G. A review on root and tuber crop production and their weed management among small scale farmers in Nigeria. *ARPN J. Agric. Biol. Sci.* **2010**, *5*, 52–58.
3. Okigbo, B.N. New crops for food industry: The roots and tuber in tropical Africa. In *New Crops for Food Industry*; Weekens, G.E., Haq, N., Day, P., Eds.; Chapman and Wall: London, UK, 1989; pp. 123–134.
4. Lebot, V. *Tropical Root and Tuber Crops Cassava, Sweet Potato, Yams and Aroids*; Crop Production Science in Horticulture Series 17; CABI Publishing: Wallingford, UK, 2008.
5. Berloo, R.; Hutten, R.C.B.; Eck, H.J.; Visser, R.G.F. An online potato pedigree database resource. *Potato Res.* **2007**, *50*, 45–57. [[CrossRef](#)]
6. Nybom, H.; Weising, K.; Rotter, B. DNA fingerprinting in botany: Past, present, future. *Investig. Genet.* **2014**, *5*, 1–35. [[CrossRef](#)] [[PubMed](#)]
7. Tarn, T.R.; Tai, G.C.C.; De Long, H.; Murphy, A.M.; Seabrook, J.E.A. Breeding potatoes for long-day, temperate climates. In *Plant Breeding Reviews*; Wiley: New York, NY, USA, 1992; pp. 217–332.
8. Spanoghe, M.; Marique, T.; Rivière, J.; Lanterbecq, D.; Gadenne, M. Investigation and development of potato parentage analysis methods using multiplexed SSR fingerprinting. *Potato Res.* **2015**, *58*, 43–65. [[CrossRef](#)]
9. Luo, Z.W.; Hackett, C.A.; Bradshaw, J.E.; McNicol, J.W.; Milbourne, D. Predicting parental genotypes and gene segregation for tetrasomic inheritance. *Theor. Appl. Genet.* **2000**, *100*, 1067–1073. [[CrossRef](#)]
10. Bink, M.P.; Uimari, P.; Sillanpaa, J.; Janss, G.; Jansen, C. Multiple QTL mapping in related plant populations via a pedigree-analysis approach. *Theor. Appl. Genet.* **2002**, *104*, 751–762. [[PubMed](#)]
11. D'hoop, B.B.; Paulo, M.J.; Kowitwanich, K.; Sengers, M.; Visser, R.G.F.; Van Eck, H.J.; van Eeuwijk, F.A. Population structure and linkage disequilibrium unravelled in tetraploid potato. *Theor. Appl. Genet.* **2010**, *121*, 1151–1170. [[CrossRef](#)] [[PubMed](#)]
12. Baldwin, S.J.; Dodds, K.G.; Auvray, B.; Genet, R.A.; Macknight, R.C.; Jacobs, J.M.E. Association mapping of cold-induced sweetening in potato using historical phenotypic data. *Ann. Appl. Biol.* **2011**, *158*, 248–256. [[CrossRef](#)]
13. Song, Y.S.; Hepting, L.; Schweizer, G.; Hartl, L.; Wenzel, G.; Schwarzfischer, A. Mapping of extreme resistance to PVY (Ry (sto)) on chromosome XII using anther-culture-derived primary dihaploid potato lines. *Theor. Appl. Genet.* **2005**, *111*, 879–887. [[CrossRef](#)] [[PubMed](#)]
14. Sun, G.; Wang-Pruski, G.; Mayich, M.; Jong, H. RAPD and pedigree-based genetic diversity estimates in cultivated diploid potato hybrids. *Theor. Appl. Genet.* **2003**, *107*, 110–115. [[CrossRef](#)] [[PubMed](#)]

15. Braun, A.; Wenzel, G. Molecular analysis of genetic variation in potato (*Solanum tuberosum* L.). I. German cultivars and advanced clones. *Potato Res.* **2004**, *47*, 81–92. [[CrossRef](#)]
16. Gjedrem, T. The first family-based breeding program in aquaculture. *Rev. Aquac.* **2010**, *2*, 2–15. [[CrossRef](#)]
17. Demeke, T.; Lynch, D.R.; Kawchuk, L.M.; Kozub, G.C.; Armstrong, J.D. Genetic diversity of potato determined by random amplified polymorphic DNA analysis. *Plant Cell Rep.* **1996**, *15*, 662–667. [[CrossRef](#)] [[PubMed](#)]
18. Isenegger, D.A.; Taylor, P.W.J.; Ford, R.; Franz, P.; McGregor, G.R.; Hutchinson, J.F. DNA fingerprinting and genetic relationships of potato cultivars (*Solanum tuberosum* L.) commercially grown in Australia. *Aust. J. Agric. Res.* **2001**, *52*, 911–918. [[CrossRef](#)]
19. Vincent, K.; Robert, K.; Morag, F.; Tadeo, K.; Yona, B.; Peter, V. Identification of F1 cassava (*Manihot esculenta* Crantz) progeny using microsatellite markers and capillary electrophoresis. *Am. J. Plant Sci.* **2014**, *5*, 119–125. [[CrossRef](#)]
20. Contreras, R.N.; Ranney, T.G.; Milla-Lewis, S.R.; Yencho, G.C. Investigating parentage and hybridity of three azaleodendrons using amplified fragment length polymorphism analysis. *Hortscience* **2007**, *42*, 740–743.
21. Levine, L.; Asmussen, M.; Olvera, O.; Powell, J.R.; De La Rosa, M.E.; Salceda, V.M.; Gaso, M.I.; Guzman, J.; Anderson, W.W. Population genetics of Mexican *Drosophila*. V. A high rate of multiple insemination in a natural population of *Drosophila pseudoobscura*. *Am. Nat.* **1980**, *116*, 493–503. [[CrossRef](#)]
22. Ellstrand, N.C. Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. *Am. Nat.* **1984**, *123*, 819–828. [[CrossRef](#)]
23. Gill, P.; Jeffreys, A.J.; Werrett, D.J. Forensic application of DNA ‘fingerprints’. *Nature* **1985**, *318*, 577–579. [[CrossRef](#)] [[PubMed](#)]
24. Jeffreys, A.J.; Brookfield, J.F.Y.; Semeonoff, R. Positive identification of an immigration test case using DNA fingerprints. *Nature* **1985**, *317*, 818–819. [[CrossRef](#)] [[PubMed](#)]
25. Jeffreys, A.J. Genetic fingerprinting. *Nat. Med.* **2005**, *11*, 1035–1039. [[CrossRef](#)] [[PubMed](#)]
26. Gibbs, H.L.; Patrick, J.; Weatherhead, P.J.; Boag, P.T.; White, B.N.; Tabak, L.M.; Hoysak, D.J. Realized reproductive success of polygynous red-winged blackbirds revealed by DNA markers. *Science* **1990**, *250*, 1394–1397. [[CrossRef](#)] [[PubMed](#)]
27. Meagher, T.R.; Thompson, E.A. The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. *Theor. Popul. Biol.* **1986**, *29*, 87–106. [[CrossRef](#)]
28. Pena, S.D.J.; Chakraborty, R. Paternity testing in the DNA era. *Trends Genet.* **1994**, *10*, 204–209. [[CrossRef](#)]
29. Jones, A.G.; Small, C.M.; Paczolt, K.A.; Ratterman, N.L. A practical guide to methods of parentage analysis. *Mol. Ecol. Res.* **2010**, *10*, 6–30. [[CrossRef](#)] [[PubMed](#)]
30. Jones, A.G.; Ardren, W.R. Methods of parentage analysis in natural populations. *Mol. Ecol.* **2003**, *12*, 2511–2523. [[CrossRef](#)] [[PubMed](#)]
31. Scarcelli, N.; Daïnou, O.; Agbangla, C.; Tostain, S.; Pham, J.L. Segregation patterns of isozyme loci and microsatellite markers show the diploidy of African yam *Dioscorea rotundata* ( $2n = 40$ ). *Theor. Appl. Genet.* **2005**, *111*, 226–232. [[CrossRef](#)] [[PubMed](#)]
32. Bousalem, M.; Viader, V.; Mariac, C.; Gomez, R.M.; Hochu, I.; Santoni, S.; David, J. Evidence of diploidy in the wild Amerindian yam, a putative progenitor of the endangered species *Dioscorea trifida* (Dioscoreaceae). *Genome* **2010**, *53*, 371–383. [[CrossRef](#)] [[PubMed](#)]
33. Gedil, M.; Sartie, A.M. Perspectives on molecular breeding of Africa’s main staple food crops—Cassava and yam. *Asp. Appl. Biol.* **2010**, *96*, 123–136.
34. Sartie, A.; Asiedu, R. Development of mapping populations for genetic analysis in yams (*Dioscorea rotundata* Poir. and *Dioscorea alata* L.). *Afr. J. Biotechnol.* **2011**, *10*, 3040–3050.
35. Nemorin, A.; David, J.; Maledon, E.; Nudol, E.; Dalon, J.; Arnau, G. Microsatellite and flow cytometry analysis to help understand the origin of *Dioscorea alata* polyploids. *Ann. Bot.* **2013**, *112*, 811–819. [[CrossRef](#)] [[PubMed](#)]
36. Arnau, G.; Maledon, E.; Bachand, I.; Abraham, K. Production of interploid hybrids and molecular markers heterozygosity determination using microsatellite markers in the greater yam, *D. alata*: Importance for the genetic improvement of the greater yam. In Proceedings of the 14th Triennial Symposium of the International Society for Tropical Roots Crops, Thiruvananthapuram, India, 20–26 November 2006.

37. Girma, G.; Hyma, K.E.; Asiedu, R.; Mitchell, S.E.; Gedil, M.; Spillane, C. Next-generation sequencing based genotyping, cytometry and phenotyping for understanding diversity and evolution of guinea yams. *Theor. Appl. Genet.* **2014**, *127*, 1783–1794. [[CrossRef](#)] [[PubMed](#)]
38. Arnau, G.; Bhattacharjee, R.; Shela, M.N.; Chair, H.; Malapa, R.; Lebot, V.; Abraham, K.; Perrier, X.; Petro, D.; Penet, L.; et al. Understanding the genetic diversity and population structure of yam (*Dioscorea alata* L.) using microsatellite markers. *PLoS ONE* **2017**, *12*, e0174150. [[CrossRef](#)] [[PubMed](#)]
39. Tamiru, M.; Natsume, S.; Takagi, H.; White, B.; Yaegashi, H.; Shimizu, M.; Yoshida, K.; Uemura, A.; Oikawa, K.; Abe, A.; et al. Genome sequencing of the staple food crop white Guinea yam enables the development of a molecular marker for sex determination. *BMC Biol.* **2017**, *15*, 86. [[CrossRef](#)] [[PubMed](#)]
40. Mignouna, H.D.; Mank, R.A.; Ellis, T.H.N.; Van den Bosch, N.; Asiedu, R.; Ng, S.Y.C.; Peleman, J. A genetic linkage map of Guinea yam (*Dioscorea rotundata* L.) based on AFLP markers. *Theor. Appl. Genet.* **2002**, *105*, 716–725. [[PubMed](#)]
41. Mignouna, H.D.; Mank, R.A.; Ellis, T.H.N.; Van den Bosch, N.; Asiedu, R.; Abang, M.M.; Peleman, J. A genetic linkage map of water yam (*Dioscorea alata* L.) based on AFLP markers and QTL analysis for anthracnose resistance. *Theor. Appl. Genet.* **2002**, *105*, 726–735. [[PubMed](#)]
42. Cao, J.; Jiang, D.; Zhao, Z.; Yuan, S.; Zhang, Y.; Zhang, T.; Zhong, W.; Yuan, Q.; Huang, L. Development of chloroplast genomic resources in Chinese Yam (*Dioscorea polystachya*). *BioMed Res. Int.* **2018**, *2018*, 1–11. [[CrossRef](#)] [[PubMed](#)]
43. Smyda-Dajmund, P.; Śliwka, J.; Wasilewicz-Flis, I.; Jakuczun, H.; Zimnoch-Guzowska, E. Genetic composition of interspecific potato somatic hybrids and autofused 4x plants evaluated by DArT and cytoplasmic DNA markers. *Plant Cell Rep.* **2016**, *35*, 1345–1358. [[CrossRef](#)] [[PubMed](#)]
44. Thieme, R.; Darsow, U.; Gavrilenko, T.; Dorokhov, D.; Tiemann, H. Production of somatic hybrids between *S. tuberosum* L. and late blight resistant Mexican wild potato species. *Euphytica* **1997**, *97*, 189–200. [[CrossRef](#)]
45. Bradshaw, J.E.; Hackett, C.A.; Meyer, R.C.; Milbourne, D.; McNicol, J.W.; Phillips, M.S.; Waugh, R. Identification of AFLP and SSR markers associated with quantitative resistance to *Globodera pallida* (Stone) in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*) with a view to marker-assisted selection. *Theor. Appl. Genet.* **1998**, *97*, 202–210. [[CrossRef](#)]
46. Ashkenazi, V.; Chani, E.; Lavi, U.; Levy, D.; Hillel, J.; Veilleux, R.E. Development of microsatellite markers in potato and their use in phylogenetic and fingerprinting analyses. *Genome* **2001**, *44*, 50–62. [[CrossRef](#)] [[PubMed](#)]
47. Simko, I.; Haynes, K.G.; Jones, R.W. Mining data from potato pedigrees: Tracking the origin of susceptibility and resistance to *Verticillium dahliae* in North American cultivars through molecular analysis. *Theor. Appl. Genet.* **2004**, *108*, 225–230. [[CrossRef](#)] [[PubMed](#)]
48. Reid, A.; Kerr, E.M. A rapid simple sequence repeat (SSR)-based identification method for potato cultivars. *Plant Genet. Res.* **2007**, *5*, 7–13. [[CrossRef](#)]
49. Bryan, G.; Lloyd, D.; Bradshaw, J. *Understanding and Improving Potato Flavour Characteristics*; Project Report; Agriculture and Horticulture Development Board (AHDB): Warwickshire, UK, 2008.
50. Côté, M.J.; Leduc, L.; Reid, A. Evaluation of simple sequence repeat (SSR) markers established in Europe as a method for the identification of potato varieties grown in Canada. *Am. J. Potato Res.* **2013**, *90*, 340–350. [[CrossRef](#)]
51. Hackett, C.A.; McLean, K.; Bryan, G.J. Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. *PLoS ONE* **2013**, *8*, e63939. [[CrossRef](#)] [[PubMed](#)]
52. Hirsch, C.N.; Hirsch, C.D.; Felcher, K.; Coombs, J.; Zarka, D.; Van Deynze, A.; De Jong, W.; Veilleux, R.E.; Jansky, S.; Bethke, P. Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *Genes Genomes Genet.* **2013**, *3*, 1003–1013.
53. Sharma, S.K.; Bolser, D.; de Boer, F.; Sønderkær, M.; Amoros, W.; Carboni, M.F.; Amoros, J.M.; de la Cruz, G.; Di Genova, A.; Douches, D.S.; et al. Construction of reference chromosome-scale pseudomolecules for potato: Integrating the potato genome with genetic and physical maps. *Genes Genomes Genet.* **2013**, *3*, 2031–2047. [[CrossRef](#)] [[PubMed](#)]
54. Sharma, R.; Bhardwaj, V.; Dalamu, D.; Kaushik, S.K.; Singh, B.P.; Sharma, S.; Umamaheshwari, R.; Baswaraj, R.; Kumar, V.; Gebhardt, C. Identification of elite potato genotypes possessing multiple disease resistance genes through molecular approaches. *Sci. Hortic.* **2014**, *179*, 204–211. [[CrossRef](#)]

55. Yada, B. Genetic Analysis of Agronomic Traits and Resistance to Sweetpotato Weevil and Sweet Potato Virus Disease in a Bi-Parental Sweetpotato Population. Ph.D. Thesis, North Carolina State University, Raleigh, NC, USA, 2014.
56. Goyer, A.; Hamlin, L.; Crosslin, J.M.; Buchanan, A.; Chang, J.H. RNA-seq analysis of resistant and susceptible potato varieties during the early stages of potato virus Y infection. *BMC Genom.* **2015**, *16*, 472. [[CrossRef](#)] [[PubMed](#)]
57. Vos, P.G.; Uitdewilligen, J.G.A.M.L.; Voorrips, R.E.; Visser, R.G.F.; van Eck, H.J. Development and analysis of a 20 K SNP array for potato (*Solanum tuberosum*): An insight into the breeding history. *Theor. Appl. Genet.* **2015**, *128*, 2387–2401. [[CrossRef](#)] [[PubMed](#)]
58. Chen, L. Potato Variety Identification with a Panel of SNP Markers. Master's Thesis, Wageningen University, Wageningen, The Netherlands, 1 March 2016.
59. Endelman, J.B.; Carley, C.A.S.; Douches, D.S.; Coombs, J.J.; Bizimungu, B.; De Jong, W.S.; Holm, K.G.; Holm, D.G.; Creighton Miller, J., Jr.; Novy, R.G.; et al. Pedigree reconstruction with genome-wide markers in potato. *Am. J. Potato Res.* **2017**, *94*, 184–190. [[CrossRef](#)]
60. Maras, M.; Sedlar, A.; Reid, A.; Božović, V.; Jovović, Z.; Meglič, V.; Peter Dolničar, P. Genetic diversity and redundancy among potato accessions in the Montenegrin collection as revealed by microsatellite markers. *Am. J. Potato Res.* **2017**, *94*, 306–313. [[CrossRef](#)]
61. Roger, A.C.; Jones, R.A.C.; Vincent, S.J. Strain-specific hypersensitive and extreme resistance phenotypes elicited by *Potato virus Y* among 39 potato cultivars released in three world regions over a 117-year period. *Plant Dis.* **2018**, *102*, 185–196.
62. Stich, B.; Inghelandt, D.V. Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. *Front. Plant Sci.* **2018**, *9*, 159. [[CrossRef](#)] [[PubMed](#)]
63. Deperi, S.I.; Tagliotti, M.E.; Bedogni, M.C.; Manrique-Carpintero, N.C.; Coombs, J.; Zhang, R.; Douches, M.; Huarte, M.A. Discriminant analysis of principal components and pedigree assessment of genetic diversity and population structure in a tetraploid potato panel using SNPs. *PLoS ONE* **2018**, *13*, e0194398. [[CrossRef](#)] [[PubMed](#)]
64. Da Silva, D.C.S.; Martins, M.L.L.; Santos, A.S.; da Silva Santos, V.; Alves, A.A.C.; da Silva Ledo, C.A. Obtaining hybrids of cultivars and wild subspecies of cassava. *Pesqui. Agropecu. Bras.* **2018**, *53*, 182–188. [[CrossRef](#)]
65. Chavarriaga-Aguirre, P.; Maya, M.M.; Bonierbale, M.W.; Kresovich, S.; Fregene, M.A.; Tohme, J.; Kochert, G. Microsatellites in cassava (*Manihot esculenta* Crantz): Discovery, inheritance and variability. *Theor. Appl. Genet.* **1998**, *97*, 493–501. [[CrossRef](#)]
66. Ceballos, H.; Kawuki, R.S.; Gracen, V.E.; Yencho, G.C.; Hershey, C.H. Conventional breeding, marker-assisted selection, genomic selection and inbreeding in clonally propagated crops: A case study for cassava. *Theor. Appl. Genet.* **2015**, *128*, 1647–1667. [[CrossRef](#)] [[PubMed](#)]
67. Da Costa, T.R.; Filho, P.S.V.; Gonçalves-Vidigal, M.C.; Galván, M.Z.; Lacanallo, G.F.; da Silva, L.I.; Kvitschal, M.V. Genetic diversity and population structure of sweet cassava using simple sequence repeat (SSR) molecular markers. *Afr. J. Biotechnol.* **2013**, *12*, 1040–1048.
68. De Oliveira, E.J.; de Resende, M.D.V.; da Silva Santos, V.; Ferreira, C.F.; Oliveira, G.A.F.; da Silva, M.S.; Oliveira, L.A.; Aguilar-Vildoso, C.I. Genome-wide selection in cassava. *Euphytica* **2012**, *187*, 263–276. [[CrossRef](#)]
69. Vieira, L.D.J.; Tavares Filho, L.F.D.Q.; Souza, F.V.D.; Alves, A.A.C.; De Oliveira, E.J. Development of interspecific hybrids of cassava and paternity analysis with molecular markers. *J. Agric. Sci.* **2013**, *151*, 849–861. [[CrossRef](#)]
70. Sakurai, T.; Plata, G.; Rodríguez-Zapata, F.; Seki, M.; Salcedo, A.; Toyoda, A.; Ishiwata, A.; Tohme, J.; Sakaki, Y.; Shinozaki, K.; et al. Sequencing analysis of 20,000 full-length cDNA clones from cassava reveals lineage specific expansions in gene families related to stress response. *BMC Plant Biol.* **2007**, *7*, 66. [[CrossRef](#)] [[PubMed](#)]
71. Mba, R.E.C.; Stephenson, P.; Edwards, K.; Melzer, S.; Mkumbira, J.; Gullberg, U.; Apel, K.; Gale, M.; Tohme, J.; Fregene, M. Simple Sequence Repeat (SSR) Markers Survey of the Cassava (*Manihot esculenta* Crantz) Genome: Towards an SSR-Based Molecular Genetic Map of Cassava. *Theor. Appl. Genet.* **2001**, *102*, 21–31. [[CrossRef](#)]

72. Mohan, C.; Shanmugasundaram, P.; Senthil, N. Identification of true hybrid progenies in cassava using simple sequence repeat (SSR) markers. *Bangladesh J. Bot.* **2013**, *42*, 155–159. [[CrossRef](#)]
73. Otti, G.; Ayodele, F.; Ikpan, A.; Melaku, G. Development of genomic tools for verification of hybrids and selfed progenies in cassava (*Manihot esculenta*). *Afr. J. Biotechnol.* **2011**, *10*, 17400–17408.
74. Olsen, K.M.; Schael, B.A. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5586–5591. [[CrossRef](#)] [[PubMed](#)]
75. Olsen, J.B.; Busack, C.; Britt, J.; Bentzen, P. The aunt and uncle effect: An empirical evaluation of the confounding influence of full sibs of parents on pedigree reconstruction. *J. Hered.* **2001**, *92*, 243–247. [[CrossRef](#)] [[PubMed](#)]
76. Olsen, K.M. SNPs, SSRs, and inferences on cassava's origin. *Plant Mol. Biol.* **2004**, *56*, 517–526. [[CrossRef](#)] [[PubMed](#)]
77. Ouborg, N.J.; Piquot, Y.; Van Groenendael, J.M. Estimating pollen flow using SSR markers and paternity exclusion: Accounting for mistyping. *Mol. Ecol.* **2005**, *14*, 3109–3121.
78. Roa, A.C.; Chavarriaga-Aguirre, P.; Duque, M.C.; Maya, M.M.; Bonierbale, M.W.; Iglesias, C.; Tohme, J. Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: Allelic polymorphism and degree of relationship. *Am. J. Bot.* **2000**, *87*, 1647–1655. [[CrossRef](#)] [[PubMed](#)]
79. Rabbi, I.; Hamblin, M.; Gedil, M.; Kulakow, P.; Ferguson, M.; Ikpan, A.S.; Ly, D.; Jannink, J.-L. Genetic mapping using genotyping-by-sequencing in the clonally propagated cassava. *Crop Sci.* **2014**, *54*, 1384–1396. [[CrossRef](#)]
80. Rabbi, I.Y.; Kulakow, P.A.; Manu-Aduening, J.A.; Dankyi, A.A.; Asibuo, J.Y.; Parkes, E.Y.; Abdoulaye, T.; Girma, G.; Gedil, M.A.; Ramu, P.; et al. Tracking crop varieties using genotyping by-sequencing markers: A case study using cassava (*Manihot esculenta* Crantz). *BMC Genet.* **2015**, *16*, 115. [[CrossRef](#)] [[PubMed](#)]
81. Jones, A. Cytological observations and fertility measurements of sweetpotato (*Ipomoea batatas* (L.) Lam.). *Proc. Am. Soc. Hortic. Sci.* **1965**, *86*, 527–537.
82. Jarret, R.L.; Gawel, N.; Whittensmore, A. Phylogenetic relationships of the sweet potato (*Ipomoea batatas* (L.) Lam.). *J. Am. Soc. Hortic. Sci.* **1992**, *117*, 633–637.
83. Buteler, M.I. Microsatellite-Based Paternity Analysis in Hexaploid Sweetpotato (*Ipomoea batatas* (L.) Lam.). Ph.D. Thesis, Louisiana State University, Baton Rouge, LA, USA, 1996.
84. Buteler, M.I.; Labonte, D.R.; Macchiavelli, R.E. Determining paternity in polyploids: Hexaploid simulation studies. *Euphytica* **1997**, *96*, 353–361. [[CrossRef](#)]
85. Buteler, M.I.; Jarret, R.L.; La Bonte, D.R. Sequence characterisation of microsatellites in diploid and polyploid *Ipomoea*. *Theor. Appl. Genet.* **1999**, *99*, 123–132. [[CrossRef](#)]
86. Buteler, M.I.; LaBonte, R.L.; Jarret, D.R.; Macchiavelli, R.E. Microsatellite-based paternity analysis in polyploid sweetpotato. *J. Am. Soc. Hortic. Sci.* **2002**, *127*, 392–396.
87. Lopez-Lavalle, L.A.B.; Orjeda, G. Occurrence and cytological mechanism of  $2n$  pollen formation in a tetraploid accession of *Ipomoea batatas* (sweet potato). *J. Hered.* **2002**, *93*, 185–192. [[CrossRef](#)]
88. Hwang, S.Y.; Tseng, Y.T.; Lo, H.F. Application of simple sequence repeats in determining the genetic relationships of cultivars used in sweet potato polycross breeding in Taiwan. *Sci. Hortic.* **2002**, *93*, 215–224. [[CrossRef](#)]
89. Hu, J.; Nakatani, M.; Mizuno, K.; Fujimura, T. Development and characterization of microsatellite markers in sweetpotato. *Breed. Sci.* **2004**, *54*, 177–188. [[CrossRef](#)]
90. Mwangi, R.O.M.; Andrade, M.I.; Carey, E.E.; Low, J.W.; Yencho, G.C.; Grüneberg, W.J. Sweetpotato (*Ipomoea batatas* L.). In *Genetic Improvement of Tropical Crops*; Springer: Cham, Switzerland, 2017.
91. Rosmayati, M.S.; Bakti, D. Identification and phylogenetic analysis of local yellow and orange sweet potatoes genotypes in Sumatera Utara. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *122*, 012048. [[CrossRef](#)]
92. Ashley, M.V. Plant parentage, pollination, and dispersal: How DNA microsatellites have altered the landscape. *Crit. Rev. Plant Sci.* **2010**, *29*, 148–161. [[CrossRef](#)]
93. Caron, G.E.; Leblanc, R. Pollen contamination in a small black spruce seedling seed orchard for three consecutive years. *For. Ecol. Manag.* **1992**, *53*, 245–261. [[CrossRef](#)]
94. Greenwood, M.S. Gene exchange in loblolly pine: The relation between pollination mechanism, female receptivity and pollen availability. *Am. J. Bot.* **1986**, *73*, 1443–1451. [[CrossRef](#)]

95. Linhart, Y.B.; Busby, W.H.; Beach, J.H.; Feinsinger, P. Forager behavior, pollen dispersal, and inbreeding in two species of hummingbird-pollinated plants. *Evolution* **1987**, *41*, 679–682. [[CrossRef](#)] [[PubMed](#)]
96. Parra, V.; Vargas, C.F.; Eguiarte, L.E. Reproductive biology, pollen and seed dispersal, and neighborhood size in the hummingbird-pollinated *Echeveria gibbiflora* (Crassulaceae). *Am. J. Bot.* **1993**, *80*, 153–159. [[CrossRef](#)]
97. Walther-Hellwig, K.; Frankl, R. Foraging distances of *Bombus muscorum*, *Bombus lapidarius*, and *Bombus terrestris* (Hymenoptera, Apidae). *J. Insect Behav.* **2000**, *13*, 239–246. [[CrossRef](#)]
98. Sunnucks, P. Efficient genetic markers for population biology. *Trends Ecol. Evol.* **2000**, *15*, 199–203. [[CrossRef](#)]
99. Cercueil, A.; Bellemain, E.; Manel, S. PARENTE: Computer program for parentage analysis. *J. Hered.* **2002**, *93*, 458–459. [[CrossRef](#)] [[PubMed](#)]
100. Coltman, D.W. Male reproductive success in a promiscuous mammal: Behavioural estimates compared with genetic paternity. *Mol. Ecol.* **1999**, *8*, 1199–1209. [[CrossRef](#)] [[PubMed](#)]
101. Brotherton, P.N.M.; Pemberton, J.; Komers, P.; Malarky, G. Genetic and behavioural evidence of monogamy in a mammal, Kirk's dik-dik (*Madoqua kirkii*). *Proc. R. Soc. Lond. B Biol. Sci.* **1997**, *264*, 675–681. [[CrossRef](#)] [[PubMed](#)]
102. Goossens, B.; Graziani, L.; Waits, L. Extra-pair paternity in the monogamous Alpine marmot revealed by nuclear DNA microsatellite analysis. *Behav. Ecol. Sociobiol.* **1998**, *43*, 281–288. [[CrossRef](#)]
103. Clapham, P.J.; Palsboll, P.J. Molecular analysis of paternity shows promiscuous mating in female humpback whales (*Megaptera novaeangliae* Borowski). *Proc. R. Soc. Lond. B Biol. Sci.* **1997**, *264*, 95–98. [[CrossRef](#)] [[PubMed](#)]
104. Howitt, C.A.; Miskelly, A. Identification of grain variety and quality type. In *Cereal Grains: Assessing and Managing Quality*, 2nd ed.; Wrigley, C., Batey, I., Miskelly, D., Eds.; Woodhead Publishing: Cambridge, UK, 2017; pp. 311–341.
105. Pemberton, J.M. Wild pedigrees: The way forward. *Proc. R. Soc. Lond. B Biol. Sci.* **2009**, *275*, 613–621. [[CrossRef](#)] [[PubMed](#)]
106. Teixeira, S.; Bernasconi, G. High prevalence of multiple paternity within fruits in natural populations of *Silene latifolia*, as revealed by microsatellite DNA analysis. *Mol. Ecol.* **2007**, *16*, 4370–4379. [[CrossRef](#)] [[PubMed](#)]
107. Craig, J.; Fowler, S.; Burgoyne, L.A.; Scott, A.C.; Harding, H.W.J. Repetitive deoxyribonucleic acid (DNA) and human genome variation: A concise review relevant to forensic biology. *J. Forensic Sci.* **1988**, *33*, 1111–1126.
108. Jacob, H.J.; Lindpaintner, K.; Lincoln, S.E.; Kusumi, K.; Bunker, R.K.; Mao, Y.P.; Ganten, D.; Dzau, V.J.; Lander, E.S. Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. *Cell* **1991**, *67*, 213–224. [[CrossRef](#)]
109. Walsh, P.S.; Fildes, N.J.; Reynolds, R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res.* **1996**, *24*, 2807–2812. [[CrossRef](#)] [[PubMed](#)]
110. Smulders, M.J.M.; Bredemeijer, G.; Rus-Kortekaas, W.; Arens, P.; Vosman, B. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theor. Appl. Genet.* **1997**, *94*, 264–272. [[CrossRef](#)]
111. Park, Y.-J.; Lee, J.K.; Kim, N.-S. Simple sequence repeat polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. *Molecules* **2009**, *14*, 4546–4569. [[CrossRef](#)] [[PubMed](#)]
112. Grattapaglia, D.; Kirst, M. *Eucalyptus* applied genomics: From gene sequences to breeding tools. *New Phytol.* **2008**, *179*, 911–929. [[CrossRef](#)] [[PubMed](#)]
113. Morin, P.A.; Luikart, G.; Wayne, R.K. SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* **2004**, *19*, 208–216. [[CrossRef](#)]
114. Meudt, H.M.; Clarke, A.C. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends Plant Sci.* **2007**, *12*, 106–117. [[CrossRef](#)] [[PubMed](#)]
115. Anderson, E.C.; Garza, J.C. The power of single-nucleotide polymorphisms for large-scale parentage inference. *Genetics* **2006**, *172*, 2567–2582. [[CrossRef](#)] [[PubMed](#)]
116. Wright, S.I.; Bi, I.V.; Schroeder, S.C.; Yamasaki, M.; Doebley, J.F.; McMullen, M.D.; Gaut, B.S. Evolution: The effects of artificial selection on the maize genome. *Science* **2005**, *308*, 1310–1314. [[CrossRef](#)] [[PubMed](#)]
117. Batley, J.; Barker, G.; O'Sullivan, H.; Edwards, K.J.; Edwards, D. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiol.* **2003**, *132*, 84–91. [[CrossRef](#)] [[PubMed](#)]

118. Meyers, B.C.; Tingey, S.V.; Morgante, M. Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. *Genome Res.* **2001**, *11*, 1660–1676. [[CrossRef](#)] [[PubMed](#)]
119. Mammadov, J.; Aggarwal, R.; Buyyarapu, R.; Kumpatla, S. SNP markers and their impact on plant breeding. *Int. J. Plant Genom.* **2012**, *2012*, 1–11. [[CrossRef](#)] [[PubMed](#)]
120. Jaccoud, D.; Peng, K.; Feinstein, D.; Kilian, A. Diversity arrays: A solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* **2001**, *29*, e25. [[CrossRef](#)] [[PubMed](#)]
121. Lezar, S.; Myburg, A.A.; Berger, D.K.; Wingfield, M.J.; Wingfield, B.D. Development and assessment of microarray-based DNA fingerprinting in *Eucalyptus grandis*. *Theor. Appl. Genet.* **2004**, *109*, 1329–1336. [[CrossRef](#)] [[PubMed](#)]
122. Kilian, A.; Huttner, E.; Wenzl, P.; Jaccoud, D.; Carling, J.; Caig, V.; Evers, M.; Heller-Uszynska, K.; Cayla, C.; Patarapuwadol, S.; et al. The fast and the cheap: SNP and DarT-based whole genome profiling for crop improvement. In Proceedings of the International Congress in the Wake of the Double Helix: From the Green Revolution to the Gene Revolution, Bologna, Italy, 27–31 May 2003; Tuberosa, R., Phillips, R.L., Gale, M., Eds.; Oxford University Press: Oxford, UK; pp. 443–461.
123. Akbari, M.; Wenzl, P.; Caig, V.; Carling, J.; Xia, L.; Yang, S.; Uszynski, G.; Mohler, V.; Lehmensiek, A.; Kuchel, H.; et al. Diversity arrays technology (DarT) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* **2006**, *113*, 1409–1420. [[CrossRef](#)] [[PubMed](#)]
124. James, K.E.; Schneider, H.; Ansell, S.W.; Evers, M.; Robba, L.; Uszynski, G.; Pedersen, N.; Newton, A.E.; Russell, S.J.; Vogel, J.C.; et al. Diversity arrays technology (DarT) for pan-genomic evolutionary studies of non-model organisms. *PLoS ONE* **2008**, *3*, e1682. [[CrossRef](#)] [[PubMed](#)]
125. Semagn, K.; Bjørnstad, Å.; Ndjiondjop, M.N. An overview of molecular marker methods for plants. *Afr. J. Biotechnol.* **2006**, *5*, 2540–2568.
126. Poland, J.A.; Rife, T.W. Genotyping-by-Sequencing for Plant Breeding and Genetics. *Plant Genome* **2012**, *5*, 92–102. [[CrossRef](#)]
127. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [[CrossRef](#)] [[PubMed](#)]
128. Glaubitz, J.C.; Rhodes, O.E.; DeWoody, J.A. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. *Mol. Ecol.* **2003**, *12*, 1039–1047. [[CrossRef](#)] [[PubMed](#)]
129. Jones, B.; Walsh, D.; Werner, L.; Fiumera, A. Using blocks of linked single nucleotide polymorphisms as highly polymorphic genetic markers for parentage analysis. *Mol. Ecol. Res.* **2009**, *9*, 487–497. [[CrossRef](#)] [[PubMed](#)]
130. Jones, A.G.; Arguello, J.R.; Arnold, S.J. Validation of Bateman's principles: A genetic study of mating patterns and sexual selection in newts. *Proc. R. Soc. Lond. B Biol. Sci.* **2002**, *269*, 2533–2539. [[CrossRef](#)] [[PubMed](#)]
131. DeWoody, J.A.; Walker, D.; Avise, J.C. Genetic parentage in large half-sib clutches: Theoretical estimates and empirical appraisals. *Genetics* **2000**, *154*, 1907–1912. [[PubMed](#)]
132. DeWoody, J.A.; DeWoody, Y.D.; Fiumera, A.C.; Avise, J.C. On the number of reproductives contributing to a half-sib progeny array. *Genet. Res.* **2000**, *75*, 95–105. [[CrossRef](#)] [[PubMed](#)]
133. Neff, B.D.; Pitcher, T.E.; Repka, J. A Bayesian model for assessing the frequency of multiple mating in nature. *J. Hered.* **2002**, *93*, 406–414. [[CrossRef](#)] [[PubMed](#)]
134. Sefc, K.M.; Koblmueller, S. Assessing parent numbers from offspring genotypes: The importance of marker polymorphism. *J. Hered.* **2009**, *100*, 197–205. [[CrossRef](#)] [[PubMed](#)]
135. Marshall, T.C.; Slate, J.; Kruuk, L.E.B.; Pemberton, J.M. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **1998**, *7*, 639–655. [[CrossRef](#)] [[PubMed](#)]
136. Hoffman, J.I.; Amos, W. Microsatellite genotyping errors: Detection approaches, common sources and consequences for paternal exclusion. *Mol. Ecol.* **2005**, *14*, 599–612. [[CrossRef](#)] [[PubMed](#)]
137. Wang, J.L. Sibship reconstruction from genetic data with typing errors. *Genetics* **2004**, *166*, 1963–1979. [[CrossRef](#)] [[PubMed](#)]
138. Koch, M.; Hadfield, J.D.; Sefc, K.M.; Sturmbauer, C. Pedigree reconstruction in wild fish populations. *Mol. Ecol.* **2008**, *17*, 4500–4511. [[CrossRef](#)] [[PubMed](#)]

139. Dakin, E.E.; Avise, J.C. Microsatellite null alleles in parentage analysis. *Heredity* **2004**, *93*, 504–509. [[CrossRef](#)] [[PubMed](#)]
140. Kalinowski, S.T.; Taper, M.L. Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conserv. Genet.* **2006**, *7*, 991–995. [[CrossRef](#)]
141. Kalinowski, S.T.; Taper, M.L.; Creel, S. Using DNA from noninvasive samples to identify individuals and census populations: An evidential approach tolerant of genotyping errors. *Conserv. Genet.* **2006**, *7*, 319–329. [[CrossRef](#)]
142. Kalinowski, S.T.; Wagner, A.P.; Taper, M.L. ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship. *Mol. Ecol. Notes* **2006**, *6*, 576–579. [[CrossRef](#)]
143. Grant, V. *Plant Speciation*, 1st ed.; Columbia University Press: New York, NY, USA, 1971.
144. Double, M.C.; Cockburn, A.; Barry, S.C.; Smouse, P.E. Exclusion probabilities for single-locus paternity analysis when related males compete for matings. *Mol. Ecol.* **1997**, *6*, 1155–1166. [[CrossRef](#)]
145. Olsen, K.M.; Schael, B.A. Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae) and its wild relatives: Further evidence for a southern Amazonian origin of domestication. *Am. J. Bot.* **2001**, *88*, 131–142. [[CrossRef](#)] [[PubMed](#)]
146. Nielsen, R.; Mattila, D.K.; Clapham, P.J.; Palsboll, P.J. Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. *Genetics* **2001**, *157*, 1673–1682. [[PubMed](#)]
147. Duchesne, P.; Meldgaard, T.; Berrebi, P. Parentage analysis with few contributing breeders: Validation and improvement. *J. Hered.* **2008**, *99*, 323–334. [[CrossRef](#)] [[PubMed](#)]
148. Thompson, E.A.; Meagher, T.R. Parental and sib likelihoods in genealogy reconstruction. *Biometrics* **1987**, *43*, 585–600. [[CrossRef](#)] [[PubMed](#)]
149. Stebbins, G.L. *Chromosomal Evolution in Higher Plants*; Addison-Wesley: Reading, MA, USA, 1971.
150. Ashley, M.V.; Wilk, J.A.; Styan, S.M.N.; Craft, K.J.; Jones, K.L.; Feldheim, K.A.; Lewers, K.S.; Ashman, T.L. High variability and disomic segregation of microsatellites in the octoploid *Fragaria virginiana* Mill. (Rosaceae). *Theor. Appl. Genet.* **2003**, *107*, 1201–1207. [[CrossRef](#)] [[PubMed](#)]
151. Hanson, T.; Brunfeld, S.; Finegan, B.; Waits, L. Conventional and genetic measures of seed dispersal for *Dipteryx panamensis* (Fabaceae) in continuous and fragmented Costa Rican rain forest. *J. Trop. Ecol.* **2007**, *23*, 635–642. [[CrossRef](#)]
152. Hanson, T.R.; Brunfeld, S.J.; Finegan, B.; Waits, L.P. Pollen dispersal and genetic structure of the tropical tree *Dipteryx panamensis* in a fragmented Costa Rican landscape. *Mol. Ecol.* **2008**, *17*, 2060–2073. [[CrossRef](#)] [[PubMed](#)]
153. Wang, J.; Scribner, K.T. Parentage and sibship inference from markers in polyploids. *Mol. Ecol. Resour.* **2014**, *14*, 541–553. [[CrossRef](#)] [[PubMed](#)]
154. Dick, C.W.; Hardy, O.J.; Jones, F.A.; Petit, R.J. Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Trop. Plant Biol.* **2008**, *1*, 20–33. [[CrossRef](#)]
155. Aguilar, R.; Quesada, M.; Ashworth, L.; Herrerias-Diego, Y.; Loco, J. Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Mol. Ecol.* **2008**, *17*, 5177–5188. [[CrossRef](#)] [[PubMed](#)]
156. Bacles, C.F.E.; Ennos, R.A. Paternity analysis of pollen mediated gene flow for *Fraxinus excelsior* L. in a chronically fragmented landscape. *Heredity* **2008**, *101*, 368–380. [[CrossRef](#)] [[PubMed](#)]
157. Burczyk, J.; DiFazio, S.P.; Adams, W.T. Gene flow in forest trees: How far do genes really travel? *For. Genet.* **2004**, *11*, 1–14.
158. Bittencourt, J.M.; Sebbenn, A.M. Patterns of pollen and seed dispersal in a small fragmented population of a wind pollinated *Araucaria angustifolia* in southern Brazil. *Heredity* **2007**, *99*, 580–591. [[CrossRef](#)] [[PubMed](#)]
159. Hardesty, B.D.; Hubbell, S.P.; Bermingham, E. Genetic evidence of frequent long-distance recruitment in a vertebrate dispersed tree. *Ecol. Lett.* **2006**, *9*, 516–525. [[CrossRef](#)] [[PubMed](#)]
160. Slavov, G.T.; Leonardi, S.; Burczyk, J.; Adams, W.T.; Strauss, S.H.; DiFazio, S.P. Extensive pollen flow in two ecologically contrasting populations of *Populus trichocarpa*. *Mol. Ecol.* **2009**, *18*, 357–373. [[CrossRef](#)] [[PubMed](#)]

161. Hamadina, E.I.; Craufurd, P.Q.; Asiedu, R. Flowering intensity in White yam (*Dioscorea rotundata*). *J. Agric. Sci.* **2009**, *147*, 469–477. [CrossRef]
162. Sadik, S.; Okereke, O.U. Flowering, pollen grain germination, fruiting, seed germination and seedling development of white yam, *Dioscorea rotundata* Poir. *Ann. Bot.* **1975**, *39*, 597–604. [CrossRef]
163. Orkwor, G.C.; Asiedu, R.; Ekanayake, I.J. *Food Yams: Advances in Research*; IITA: Ibadan, Nigeria, 2000.
164. Govaerts, R.; Wilkin, P.; Saunders, R.M.K. *World Checklist of Dioscoreales: Yams and Their Allies*; The Board of Trustees of the Royal Botanic Gardens: London, UK, 2007; pp. 1–65.
165. Halsey, M.K.; Olsen, K.M.; Taylor, H.J.; Chavarriaga-Aguirre, P. Reproductive biology of cassava (*Manihot esculenta* Crantz) and isolation of experimental field trials. *Crop Sci.* **2008**, *48*, 49–58. [CrossRef]
166. Taxonomy of Sweet Potato Batata (*Ipomoea batatas* L.) Sec. il Cronquist System. Available online: <http://sperimentazione.altavista.org/Sweetpotato.html> (accessed on 6 July 2015).
167. Scurrah, M.; Celis-Gamboa, C.; Chumbiauca, S.; Salas, A.; Visser, R.G.F. Hybridization between wild and cultivated potato species in the Peruvian Andes and biosafety implications for deployment of GM potatoes. *Euphytica* **2008**, *164*, 881–892. [CrossRef]
168. Petti, C.; Meade, C.; Downes, M.; Mullins, E. Facilitating co-existence by tracking gene dispersal in conventional potato systems with microsatellite markers. *Environ. Biosaf. Res.* **2007**, *6*, 223–236. [CrossRef] [PubMed]
169. White, J.W. Pollination of potatoes under natural conditions. *CIP Circ.* **1983**, *11*, 1–2.
170. Canadian Food Inspection Agency (CFIA). The Biology of *Solanum tuberosum* (L.) (Potatoes). Available online: <http://www.inspection.gc.ca/plants/plants-with-novel-traits/applicants/directive-94-08/biology-documents/solanum-tuberosum-l-/eng/1330982063974/1330982145930> (accessed on 21 April 2016).
171. Hawkes, J.G. The evolution of cultivated potatoes and their tuber-bearing wild relatives. *Die Kulturpflanze* **1988**, *36*, 189–208. [CrossRef]
172. Lian, C.; Goto, S.; Kubo, T.; Takahashi, Y.; Nakagawa, M.; Hogetsu, T. Nuclear and chloroplast microsatellite analysis of *Abies sachalinensis* regeneration on fallen logs in a subboreal forest in Hokkaido, Japan. *Mol. Ecol.* **2008**, *17*, 2948–2962. [CrossRef] [PubMed]
173. Ouborg, N.J.; Piquot, Y.; Van Groenendael, J.M. Population genetics, molecular markers and the study of dispersal in plants. *J. Ecol.* **1999**, *87*, 551–568. [CrossRef]
174. Garcia, C.; Jordano, P.; Godoy, J.A. Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: Patterns in distance and direction. *Mol. Ecol.* **2007**, *16*, 1947–1955. [CrossRef] [PubMed]
175. Terakawa, M.; Isagi, Y.; Matsui, K.; Yumoto, T. Microsatellite analysis of the maternal origin of *Myrica rubra* seeds in the feces of Japanese macaques. *Ecol. Res.* **2009**, *24*, 663–670. [CrossRef]
176. Grivet, D.; Smouse, P.E.; Sork, V.L. A novel approach to an old problem: Tracking dispersed seeds. *Mol. Ecol.* **2005**, *14*, 3585–3595. [CrossRef] [PubMed]
177. Jones, F.A.; Chen, J.; Weng, G.J.; Hubbell, S.P. A genetic evaluation of seed dispersal in the neotropical tree *Jacaranda copaia* (Bignoniaceae). *Am. Nat.* **2005**, *166*, 543–555. [CrossRef] [PubMed]
178. Iwaizumi, M.G.; Takahashi, M.; Watanabe, A.; Ubukata, M. Simultaneous evaluation of paternal and maternal immigrant gene flow and the implications for the overall genetic composition of *Pinus densiflora* dispersed seeds. *J. Hered.* **2009**, *101*, 144–153. [CrossRef] [PubMed]
179. Bacles, C.F.E.; Lowe, A.; Ennos, R. Effective seed dispersal across a fragmented landscape. *Science* **2006**, *311*, 628. [CrossRef] [PubMed]
180. Vandeputte, M.; Haffray, P. Parentage assignment with genomic markers: A major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. *Front. Genet.* **2014**, *5*, 432. [CrossRef] [PubMed]
181. Visscher, P.M.; Woolliams, J.A.; Smith, D.; Williams, J.L. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. *J. Dairy Sci.* **2002**, *85*, 2368–2375. [CrossRef]
182. Muñoz, P.R.; Resende, M.F.R., Jr.; Huber, D.A.; Quesada, T.; Resende, M.D.V.; Neale, D.B.; Wegrzyn, J.L.; Kirst, M.; Peter, G.F. Genomic relationship matrix for correcting pedigree errors in breeding populations: Impact on genetic parameters and genomic selection accuracy. *Crop Sci.* **2014**, *54*, 1115–1123. [CrossRef]
183. Douches, D.S.; Ludlam, K.; Freyre, R. Isozyme and plastid DNA assessment of pedigrees of nineteenth century potato cultivars. *Theor. Appl. Genet.* **1991**, *82*, 195–200. [CrossRef] [PubMed]

184. Hamilton, J.P.; Hansey, C.N.; Whitty, B.R.; Stoffel, K.; Massa, A.N.; Van Deynze, A.; De Jong, W.S.; Douches, D.S.; Buell, C.R. Single nucleotide polymorphism discovery in elite north American potato germplasm. *BMC Genom.* **2011**, *12*, 302. [[CrossRef](#)] [[PubMed](#)]
185. Felcher, K.J.; Coombs, J.J.; Massa, A.N.; Hansey, C.N.; Hamilton, J.P.; Veilleux, R.E.; Buell, C.R.; Douches, D.S. Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS ONE* **2012**, *7*, e36347. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).