plants

MDPI

Article

Paternity Assignment in White Guinea Yam (*Dioscorea Rotundata*) Half-Sib Progenies from Polycross Mating Design Using SNP Markers

Prince E. Norman [1,2,3,*], Agre A. Paterne [2], Agyemang Danquah [3], Pangirayi B. Tongoona [3], Eric Y. Danquah [3], David De Koeyer [4], Ugochukwu N. Ikeogu [5], Robert Asiedu [2] and Asrat Asfaw [2]

1 Sierra Leone Agricultural Research Institute, Tower Hill, Freetown PMB 1313, Sierra Leone
2 International Institute of Tropical Agriculture, Ibadan PMB 5320, Nigeria; p.agre@cgiar.org (A.A.P.); r.asiedu@cgiar.org (R.A.); a.amele@cgiar.org (A.A.)
3 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; adanquah@wacci.ug.edu.gh (A.D.); tongoona@ukzn.ac.za (P.B.T.); edanquah@wacci.ug.edu.gh (E.Y.D.)
4 Fredericton Research and Development Centre, Agriculture and Agri-Food Canada, P.O. Box 20280, Fredericton, NB E3B 4Z7, Canada; david.dekoeyer@canada.ca
5 Integrative Plant Breeding, Cornell University, Ithaca, New York, NY 14850, USA; uni3@cornell.edu
* Correspondence: pnorman@wacci.ug.edu.gh; Tel.: +232-76-618-454

Received: 2 March 2020; Accepted: 16 April 2020; Published: 19 April 2020

Abstract: White Guinea yam is mostly a dioecious outcrossing crop with male and female flowers produced on distinct plants. Fertile parents produce high fruit set in an open pollination polycross block, which is a cost-effective and convenient way of generating variability in yam breeding. However, the pollen parent of progeny from polycross mating is usually unknown. This study aimed to determine paternity in white Guinea yam half-sib progenies from polycross mating design. A total of 394 half-sib progenies from random open pollination involving nine female and three male parents was genotyped with 6602 SNP markers from DArTSeq platform to recover full pedigree. A higher proportion of expected heterozygosity, allelic richness, and evenness were observed in the half-sib progenies. A complete pedigree was established for all progenies from two families (TDr1685 and TDr1688) with 100% accuracy, while in the remaining families, paternity was assigned successfully only for 56 to 98% of the progenies. Our results indicated unequal paternal contribution under natural open pollination in yam, suggesting unequal pollen migrations or gene flow among the crossing parents. A total of 3.8% of progenies lacking paternal identity due to foreign pollen contamination outside the polycross block was observed. This study established the efficient determination of parental reconstruction and allelic contributions in the white Guinea yam half-sib progenies generated from open pollination polycross using SNP markers. Findings are useful for parental reconstruction, accurate dissection of the genetic effects, and selection in white Guinea yam breeding program utilizing polycross mating design.

Keywords: *Dioscorea*; parentage analysis; pedigree reconstruction; hybrid progeny; DArT markers

1. Introduction

Yam belong to the genus *Dioscorea* of the family Dioscoreaceae, order Dioscoreales. *Dioscorea* is the largest genus of Dioscoreaceae comprising about 614 species [1]. Yam species are categorized into different sections. The five most important sections are *Enantiophyllum*, *Lasiophyton*, *Combilium*, *Opsophyton* and *Macrogynodium* [2]. *Dioscorea rotundata* Poir, *D. alata* L., *D. cayenensis* Lam, *D. opposite*

Thumb and *D. japonica* Thunb belong to *Enantiophyllum* and are distinguished by clockwise twining on support [2]. White Guinea yam (*Dioscorea rotundata* Poir) is native to West Africa [3]. White Guinea yam possesses an allotetraploid structure with disomic inheritance [4,5]. Allotetraploids originate from the merging of two species' genomes of two sets of homologous chromosomes, with each chromosome pairing only with its homologous form during meiosis [6] to form bivalents [7] and disomic inheritance. This leads to the transmission of heterozygous gametes to the progeny [8].

White Guinea yam is a highly prized African food crop cultivated for its underground starchy tubers. It is a food of choice in West Africa, where it has a deep-rooted connection with the social and cultural facets of the people [9]. White Guinea yam holds tremendous potential as a source of food and income for more than 300 million people [10] and as a raw material for industrial applications. However, the crop's potential is yet not fully exploited.

Breeding is among the tools to release the yam potential as food source and industrial raw material. Yam breeding utilizes both sexual and asexual reproduction where an artificial hand or natural open pollination applied for the generation of full-sib or half-sib progenies segregating for desired traits and subsequent selection of their clonal derivatives for performance reproducibility and stability over seasons and locations [11]. Yam species often produce male and female flowers on separate plants. Hence, it requires the establishment of separate crossing blocks of male and female parents with multiple planting dates to enhance the synchronization of flowering for hybridization. However, executing planned and controlled crosses are often constrained by the tiny size of flowers, the availability of fertile female and male plants, and the short pollination window due to partial asynchrony [12]. Half-sib breeding, which entails the random open pollination among desirable parents, is a viable alternative to the full-sib breeding which utilizes planned and controlled artificial pollination [13]. Open pollinations involving fertile parents in polycross blocks and trial fields produce a high fruit-set, and hence a cost-effective and convenient way to produce large numbers of seedlings for selection [12].

Open pollination produces half-sib progenies with unstructured pedigree. DNA profiling of progeny and possible parents and comparing their alleles for determination and validation of existing relationships [14] is a viable tool to elucidate the identity of half-sib progenies and reconstruct the pedigree in the outcrossing crops [15]. The technique works on the principle that progeny constitutes allelic contributions from maternal and paternal parents [16]. Such an analysis can provide a reliable estimation of paternal breeding values in the half-sib family and reduces crossing and labeling errors associated with conventional hand pollination. Zoundjihékpon et al. [17] performed the first parentage analysis in cultivated yam, applying six isozyme markers. They validated the progenies of crosses involving well known genitors (one male and three females). Sartie and Asiedu [18] employed nine SSR markers to determine the success of the hybridization of seven *D. rotundata* and *D. alata* mapping populations. The progenies of each of the mapping populations showed combinations of their parental alleles, indicating the success of hybridization.

Paternity testing improves selection gains by increasing parental control in the selection gain equation [19]. Accurate determination of parentage and pedigree relationships helps in determining trait inheritability and ascertaining the genetic progress [20]. Despite the enormous merits of parentage analysis using DNA markers, the technique has not been fully utilized to exploit the potential of open pollination in polycross blocks for yam breeding. The objective of this study was to assess the potential of parentage analysis in determining paternity in the white Guinea yam half-sib progenies generated from open pollination polycross blocks.

2. Results

2.1. Heterozygosity and Genetic Diversity

The genetic parameter estimates of half-sib progenies are presented in Table 1. The mean proportion of heterozygosity ranged from 0.272–0.328; the minimum proportion of heterozygosity ranged from 0.116–0.294; and the maximum proportion of heterozygosity ranged from 0.360–0.531. The genetic diversity differed among the polycross-derived families. Family TDr1686 had the widest

(0.116–0.531) proportion of heterozygosity compared to family TDr1691 (0.216–0.360), which was among families with the narrowest gene diversity. The average distance between the genetic values and the mean (standard deviation) was highest (0.065) for family TDr1686, whereas family TDr1688 exhibited the lowest genetic value deviation from the mean, 0.02. The average minor allele frequency ranged from the highest of 0.238 in family TDr1689 to the lowest, 0.185 in family TDr1690. The average major allele frequency ranged from the highest of 0.813 in family TDr1690 to the lowest, 0.762 in family TDr1689.

Table 1. Summary statistics: minor allele frequency, major allele frequency, proportion of heterozygosity, and QC = quality check of genetic (6602 SNPs, 394 individuals) of segregating progenies of white yam.

Family	Sample	Sample after QC	Hybrid	Proportion of Heterozygous					Minor Allele Frequency	Major Allele Frequency
				Mean	Min	Max	σ^2	σ		
TDr1685	50	50	50	0.276	0.193	0.486	0.002	0.043	0.205	0.794
TDr1686	50	50	37	0.276	0.116	0.531	0.004	0.065	0.211	0.788
TDr1687	50	46	46	0.291	0.229	0.397	0.001	0.030	0.209	0.791
TDr1688	50	50	50	0.327	0.294	0.407	0.000	0.020	0.233	0.766
TDr1689	50	28	28	0.328	0.161	0.517	0.002	0.048	0.238	0.762
TDr1690	45	44	44	0.272	0.255	0.437	0.001	0.029	0.186	0.813
TDr1691	49	48	48	0.281	0.216	0.360	0.001	0.029	0.204	0.795
TDr1692	50	50	49	0.293	0.251	0.460	0.001	0.038	0.209	0.791

σ^2 = variance, and σ = standard deviation.

2.2. Parental Reconstruction and Gametic Composition in Half-Sib Progenies

Generally, there were unequal proportions of half-sib progenies confirmed as true hybrids among crossing parents (Table 2). Of the 394 half-sib progenies sampled, 352 (96.2%) were with fully recovered pollen parent identity, whereas 3.8% lacked paternal identity (Table 2). The pollen parent contribution to progenies varied among male parents utilized in the crosses. The paternal contribution to the progenies was highest for male parent TDr9501932 (65.63%), followed by TDr9902607 (24.43%) and lowest for TDr8902789 (9.94%). Families TDr1685 and TDr1688 had 100% of progenies with fully recovered pedigree, whereas family TDr1689 had the lowest proportion of offspring (56%) with successfully assigned paternity. Of the 50 progenies assessed in family TDr1685; 52, 46 and 2% were contributions from male parents TDr9501932, TDr9902607 and TDr9902789, respectively (Table 2, Figures 1 and 2).

Table 2. The number and proportions of progenies with recovered paternal identity.

Female	Family	Male			Progeny with Fully Recovered Pedigree	Percent of Progenies with Fully Recovered Pedigree
		TDr9501932	TDr9902789	TDr9902607		
TDr9700793	TDr1685	26 (52%)	1 (2.0%)	23 (46%)	50	100
TDr8902157	TDr1686	12 (24%)	0 (0%)	25 (50%)	37	74
TDr8902475	TDr1687	46 (92%)	0 (0%)	0 (0%)	46	92
TDr9700632	TDr1688	23 (46%)	1 (2%)	26 (52%)	50	100
TDr9700205	TDr1689	18 (36%)	6 (12%)	4 (8%)	28	56
TDr9519158	TDr1690	39 (86.7%)	0 (0%)	5 (11.1%)	44	97.8
TDr9518988	TDr1691	45 (91.8%)	0 (0%)	3 (6.1%)	48	97.9
Ojuiyawo	TDr1692	22 (44.0%)	27 (54%)	0 (0%)	49	98
Total progenies		231	35	86	352	
% contribution		65.63	9.94	24.43	100	96.2

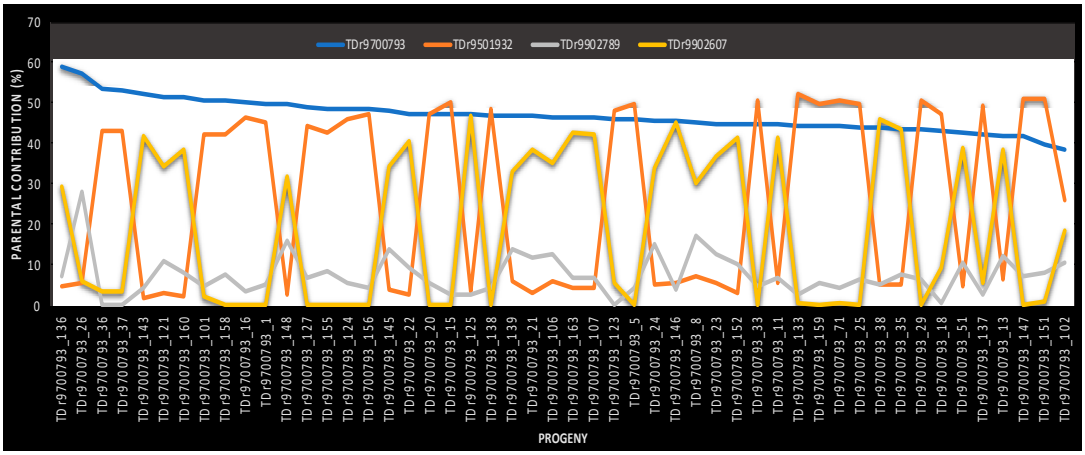

Figure 1. Pedigree reconstruction of polycross family TDr1685 based on genetic contribution of shared parental alleles in progeny**.** TDr9501932, TDr9902789 and TDr9902607 = male parents; TDr9700793 = female parent.

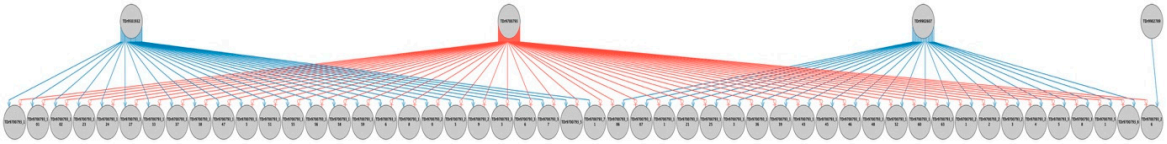

Figure 2. Pedigree reconstruction of polycross family TDr1685 based on Helium pedigree visualization**.** TDr9501932, TDr9902789 and TDr9902607 = male parents (blue); TDr9700793 = female parent (red).

The pedigree recovery of families TDr1686, TDr1687, TDr1688, TDr1689, TDr1690 and TDr1691 are displayed in Figures 2–8. Paternal parents TDr9501932 and TDr9902607 contributed 24.0 and 50.0% to family TDr1686; 46.0 and 54.0% to family TDr1688; 86.7 and 11.1% to family TDr1690; 91.8 and 6.1% to family TDr1691 with no contributions from paternal parent TDr9902789 (Table 2). Family TDr1692 had the second-highest proportion (98%) of progenies with fully recovered paternal identity with male parents TDr9501932 and TDr9902789 contributing 44 and 54%, respectively (Table 2 and Supplementary Material Table S1).

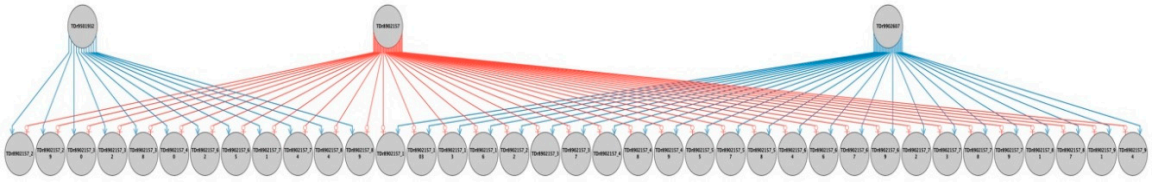

Figure 3. Pedigree reconstruction of polycross family TDr1686 based on Helium pedigree visualization**.** TDr9501932 and TDr9902607 = male parents (blue); TDr8902157 = female parent (red).

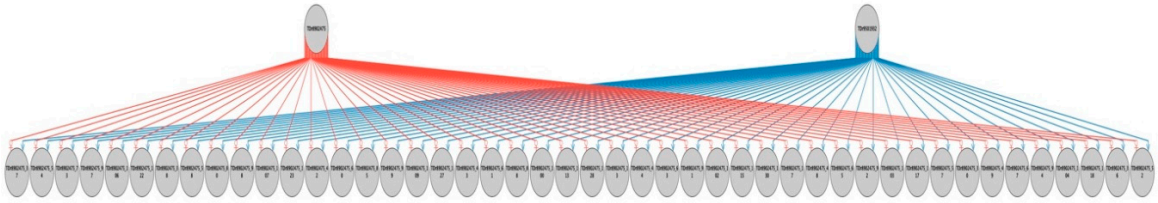

Figure 4. Pedigree reconstruction of polycross family TDr1687 based on Helium pedigree visualization. TDr9501932 = male parent (blue); TDr8902475 = female parent (red).

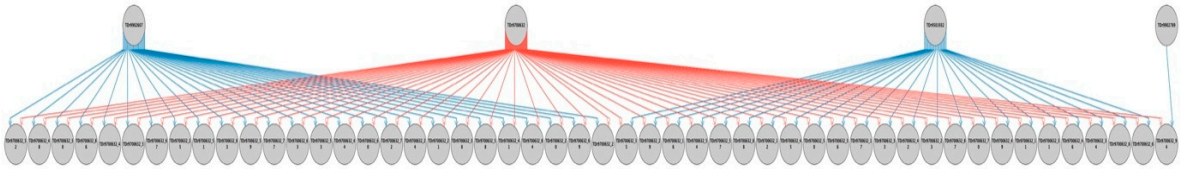

Figure 5. Pedigree reconstruction of polycross family TDr1688 based on Helium pedigree visualization**.** TDr9902607, TDr9501932 and TDr9902789 = male parents (blue); TDr9700632 = female parent (red).

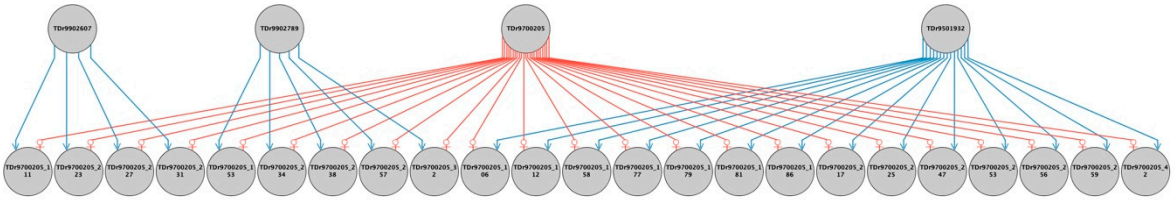

Figure 6. Pedigree reconstruction of polycross family TDr1689 based on Helium pedigree visualization**.** TDr9501932, TDr9902789 and TDr9902607 = male parents (blue); TDr9700205 = female parent (red).

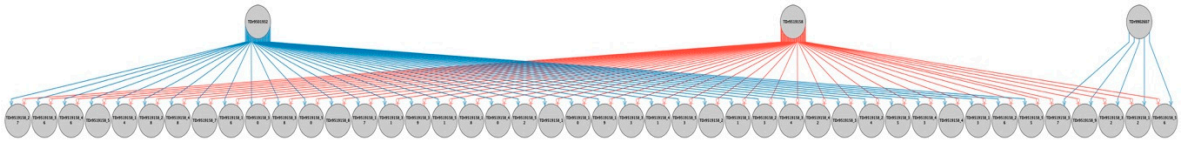

Figure 7. Pedigree reconstruction of polycross family TDr1690 based on Helium pedigree visualization**.** TDr9501932 and TDr9902607 = male parents (blue); TDr9519158 = female parent (red).

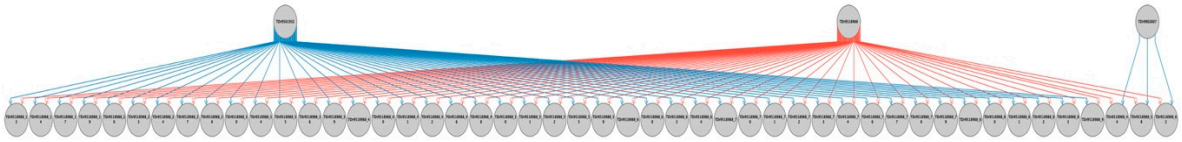

Figure 8. Pedigree reconstruction of polycross family TDr1691 based on Helium pedigree visualization**.** TDr9501932 and TDr9902607 = male parents (blue); TDr9518988 = female parent (red).

We also checked maternal identity in the progenies and found a mismatch in one out of the eight maternal parents that successfully produced offspring from open pollination polycross block. Our analysis revealed that maternal parent 'Ojuiyawo' had a low allelic contribution to the progenies despite the highest putative male contributors assigned to the progenies in family TDr1692 (Supplementary Material Table S1).

The proportion of gametes in the studied materials are presented in Table 3. Overall, the eight polycross families exhibited a higher proportion of nonmissing gametes. Family TDr1686 had the highest proportion of nonmissing gametes (94.9%) compared to family TDr1690 (92.6%), which had the lowest. The missing gametes ranged from 5.1 to 7.4% with family TDr1686 (7.4%) exhibiting the highest missing gametes, while family TDr1690 (5.1%) had the lowest.

Table 3. Gamete frequency of eight polycross-derived families of yam.

Family	Number of Nonmissing Gametes	Proportion of Nonmissing Gametes	Number of Missing Gametes	Proportion of Missing Gametes
TDr1685	669930	0.940	42870	0.060
TDr1686	660336	0.926	52680	0.074
TDr1687	672606	0.943	40410	0.057
TDr1688	676128	0.948	36888	0.052
TDr1689	681738	0.939	44482	0.061
TDr1690	601600	0.949	32192	0.051
TDr1691	643928	0.938	42680	0.062
TDr1692	674222	0.946	38794	0.054

2.3. Numbers and Patterns of SNP Mutations in Half-Sib Progenies of White Yam

Transition SNPs were higher in number and frequency in all progeny families studied (Table 4). The C to T transition accounted for the highest frequency ranging from 30.87 (family TDr1685) to 32.62% (family TDr1687), while the G to C transversion occurred at the lowest frequency ranging from 6.48 (family TDr1691) to 7.47% (family TDr1689). The Tv to Ts ratio ranged from 1:0.60 (family TDr1687) to 1:0.67 (family TDr1685) with a mean of 1:0.64.

Table 4. Percentage of transition and transversion SNPs across the *Dioscorea rotundata* genome.

Family	SNP Type						Tv to Ts Ratio
	Transitions (Ts)		Transversions (Tv)				
	AG	CT	AT	AC	GT	GC	
TDr1685	805 (29.10)	854 (30.87)	394 (14.24)	257 (9.29)	261 (9.44)	195 (7.05)	1.67
TDr1686	822 (28.20)	938 (32.18)	413 (14.17)	251 (8.61)	280 (9.61)	211 (7.24)	1.66
TDr1687	785 (30.05)	852 (32.62)	344 (13.17)	227 (8.69)	230 (8.81)	174 (6.66)	1.60
TDr1688	829 (28.57)	924 (31.84)	416 (14.33)	259 (8.92)	266 (9.17)	208 (7.17)	1.66
TDr1689	869 (27.86)	1005 (32.22)	419 (13.43)	292 (9.36)	301 (9.65)	233 (7.47)	1.66
TDr1690	756 (29.62)	820 (32.13)	349 (13.68)	223 (8.74)	228 (8.93)	176 (6.90)	1.62
TDr1691	755 (29.13)	841 (32.45)	353 (13.62)	237 (9.14)	238 (9.18)	168 (6.48)	1.62
TDr1692	831 (29.37)	888 (31.39)	403 (14.25)	256 (9.05)	256 (9.05)	195 (6.89)	1.65
Mean	806.5 (28.99)	890.25 (31.96)	386.38 (13.86)	250.25 (8.98)	257.5 (9.23)	195.0 (6.98)	1.64

Values in brackets are percentages.

3. Discussion

3.1. Allelic Diversity and Parental Reconstruction

Yam breeding is a two-step process that involves both sexual and asexual reproduction. Sexual reproduction generates recombinant progenies either through artificial hand or natural open pollination. Generating a reasonable number of recombinant progenies for valid selection from potentially interesting parental sources with the controlled cross is a challenging task in yam due to its tiny sized flowers and partial asynchrony [12,13]. However, random open pollination involving fertile parents in polycross blocks and trial fields often produces high fruit-set, and hence, is a cost-effective and convenient way to produce large numbers of seedlings for selection [12,13]. This study assessed the potential of parentage analysis in maximizing heterozygosity in the progenies from an open pollination mating design involving potential parents with a new source of genes. Natural open pollination in white yam generated half-sib families with a variable proportion of heterozygosity. Progenies of half-sib family TDr1686 exhibited the widest (0.116–0.531) heterozygosity and highest standard deviation compared to the remaining half-sib families. Families with higher average distance between the genetic values and the mean (standard deviation) implies that their data points are spread over a large range of values, whereas those with low standard deviation had data points that are close to the mean. The high heterozygosity in some of the families indicated the high genetic

variability for a response to within-family selection, whereas the within-family insufficiency in heterozygosity implied the low genetic diversity to respond to within-family selection. Shete et al. [21] also noted the influence of heterozygosity and polymorphic information content on the genetic variability and response to selection. The allelic diversity and genetic differences have been reported to be influenced by the genome transmission mode and number of successful hybridizations [22]. The allelic richness and the initial allelic composition determine the selection and the response to selection, respectively [23,24]. The high allelic richness also indicates higher accessibility of a more substantial fraction of the genotypic space by fewer mutational events [25]. The allelic richness and heterozygosity of a breeding population may be recovered by gene flow induced by migrants carrying lost alleles [26]. Thus, open pollination might facilitate the introduction of new alleles from diverse origins to broaden the genetic base in yam breeding programs.

Sexual fertility and flowering synchrony are essential for success with open natural pollination in yam. Our results showed variations in the proportion of gametes among families (Table 3). Generally, a low proportion of sterile gametes were produced among the parental clones. The observed variation in the proportion of gametes was possibly due to genotype composition, cross-compatibility, and allelic diversity among the parental genotypes. The pedigree reconstitution in studied materials showed that the male and female parents differed widely in terms of their parental ability to produce successful offspring (Table 2). More specifically, a low number of half-sib progenies (only 2–12%) was derived from the pollen parent TDr9902789 compared to the others. This could be attributed to pollen incompatibility, bad pollen shedding, or low pollen viability. Variation in flowering times between the male and female parents could not be the possible cause. Sequential planting (three planting dates in 10 days interval) was employed to facilitate flowering synchronization among the parents. As a result, all the female parents except TDr08-21-3 (Ekpe II) successfully produced progenies. Female parent TDr08-21-3 (Ekpe II) did not flower during the season, which resulted in no fruit and seed sets. Being shy to flower is a common phenomenon in yam [13]. Conversely, the maternal parent 'Ojuiyawo' had a low allelic contribution to the progenies despite the highest putative male contributors assigned to the progenies in family TDr1692. The maternal mismatch in family TDr1692 was highly probable to the mishandled maternal identity in field or laboratory during the DNA extraction. Labeling error is the potential cause of mismatch of parents and progenies in the breeding program [27]. This signifies the importance of breeding trial information management for tracking materials using barcodes in the field and during laboratory analysis to avoid errors.

More interestingly, the paternal contribution under natural open pollination was unequal, suggesting unequal pollen migrations or gene flow among the crossing parents. The causes of the varying parental ability and their unequal contribution to progenies have also been noted in other species. In *Hibiscus moscheutos* [28], *Betula pendula* [29] and *Picea abies* [30], unequal paternal contribution was found to influence the pollen-tube growth rate. Spira et al. [31] also observed that the time of arrival of pollen on the stigmatic surface influences paternal contribution in *Hibiscus moscheutos*. Paternal contributions were affected by pollen–pollen interactions in artificial crossing experiments using pollen mixtures in *Pinus sylvestris* [32]. Moreover, paternal contributions were noted to be affected by genetic incompatibility between male and female gametophytes in *Pseudotsuga menziesii* [33] and pollen germination rate in polycrosses of *Cryptomeria japonica* [34]. Detailed studies on these aspects form part of future studies in yam. We also found a low proportion (3.8%) of progenies lacking paternal identity, possibly due to low foreign pollen contamination outside the polycross block. Our polycross block was established in an isolation field with no yam plants in approximately 500 m surroundings. Yam pollen is sticky and the chance for wind pollination is minimal. Flower thrips (*Thysanoptera*) usually play key role with natural open pollination in yam [35]. However, insects belong to *Coleoptera*, *Diptera*, *Hymenoptera* and *Hemiptera* reported visiting the yam field during the flowering period presumably contribute to natural open pollination [36].

The average minor allele frequency of the risk allele tested for all families was greater than 10% indicating its significance in detecting genetic effects in the studied populations. The findings are also in concurrence with Ardlie et al. [37], who noted that loci with high minor allele frequency have a

higher power to detect weak genetic effects compared with those with lower minor allele frequency values.

Reliable pedigree information is useful for the accurate dissection of the genetic effects, and thereby maximize selection gain in crop breeding. Molecular markers have been effectively applied in many crops for the reconstruction of pedigree and understand the footprint of breeding history of contemporary germplasms [14,38]. Our result with the application of SNP markers from the DArT genotyping platform successfully recovered paternity in half-sib progenies with known maternity in *D. rotundata*. This suggests the enormous potential of parentage analysis for selection, pedigree identification, and accurate dissection of the genetic effects that could contribute to accurate prediction of the extent of gene flow in heterogeneous populations of yam, as well as other root and tuber crops. Similar analysis could also be applied to other crops too. Telfer et al. [15] reported a high throughput and reproducibility qualities of SNPs in paternity determination and pedigree reconstruction in *Eucalyptus nitens*. Sartie and Asiedu [18] reported the efficiency of molecular markers in confirming the true hybridity of progenies from *D. rotundata* and *D. alata* controlled pairwise crosses. Tamiru et al. [39] confirmed parental line-specific heterozygous SNP markers on 150 F1 individuals obtained from a bi-parental cross in white Guinea yam. They also identified parental-specific region of the genome linked with flower sex determination in white Guinea yam using whole-genome resequencing of bulked DNA samples from female and male F1 progenies, that confirmed that all the progenies were from a specific bi-parental cross. Although genetic variation in breeding populations are well known to be mainly created through mutation, population gene flow, sexual and asexual recombination, the genetic diversity in the current study may have been more attributable to sexual recombination and gene flow.

Parental reconstruction has also been successfully done in other root and tuber crops such as cassava (*Manihot esculenta* Crantz) [40–42], potato (*Solanum tuberosum* L.) [43,44] and sweetpotato (*Ipomoea batatas* L.) [45]. In these crops, molecular parentage analysis facilitated the discovery of pedigree errors or mismatches in populations improved using natural and artificial mating schemes [44,46,47]. A detailed review on parentage analysis of root and tuber crops has been reported by Norman et al. [27].

3.2. Numbers and Patterns of SNP Mutations in Half-Sib Progenies of White Yam

The numbers and patterns of SNP mutations in this study indicate a bias in chloroplast genome evolution in half-sib progenies of white yam. The diversity and patterns of SNP mutations in the studied half-sib progeny families were possibly due to the function of genes as previously suggested by Cao et al. [48]. The mean Tv to Ts ratio in the eight progeny families was 1:0.64 indicating a bias in favor of transition. This finding is similar with the Tv to Ts ratio of 1:0.60 reported for *D. polystachya* plastid genomes [48]. Small mutations at high frequency have also been reported in plants [49]. However, the frequency of these mutations is suggested to be depleting in coding regions of plant genome compared to those in the noncoding regions [49]. The scope of the present study did not cover comparison of mutational events in the coding and noncoding regions of the white yam genome. This aspect may form part of future genomic research in this plant organism.

4. Materials and Methods

4.1. Experimental Materials

The experimental materials were botanic seeds generated using polycross mating designs. The polycross design was established at the International Institute of Tropical Agriculture, Ibadan, Nigeria, using 12 genotypes of *D. rotundata* comprising nine females and three males with desired complementary traits for fresh tuber, dry matter, tuber shape, earliness and tolerance to yam mosaic virus [13]. The mating scheme utilized targeted crossing of three female parents to a male parent (3:1) producing nine cross combinations. Staggered planting (three plants at 10 days interval) applied to synchronize flowering among the parents. However, progenies of eight families were successfully generated and utilized in this study except female parent, TDr08-21-3 (Ekpe II) that did not flower to

produce viable seed. The botanic seeds generated from the eight families were processed from trilobated fruits harvested from maternal parents in late February 2017.

Prior to establishing the prenursery and nursery, cocopeat and topsoil samples were collected and analyzed at IITA soil analytical lab using standard procedures described by the International Soil Reference and Information Centre (ISRIC) and the FAO [50]. The chemical properties of the cocopeat medium used in the prenursery include 4.61 pH, 0.133% N, 1.401% OC, 0.057% P, 1.618% Ca, 0.199% Mg, 0.380% K, 0.339% Na, 0.008% Zn, 0.002% Cu, 0.019% Mn and 0.006% Fe. The chemical properties of the topsoil medium used in the nursery included 6.08 pH, 0.078% N, 1.489% OC, 0.0004% P, 0.050% Ca, 0.019% Mg, 0.007% K, 0.003% Na, 0.001% Zn, 0.002% Cu, 0.043% Mn and 0.015% Fe.

The prenursery was laid out in a completely randomized design with two replicates in the glasshouse. The cocopeat was put in perforated polyethylene seedling trays and slightly soaked. Different botanic seeds were sown in 5 cm holes created in the growth medium in early March 2017. The trays were well labeled at planting and irrigated every other day until 4 weeks after sowing (WAS). The growing seedlings were fertigated to field capacity every other day for 4 WAS; and fumigated with cypermethrin at the rate of 15 mL per 1 L water at 6 WAS prior to transplanting in the seedling nursery at 8 WAS.

The nursery was laid out in a randomized complete block design in the screen house. Seedlings were transplanted in early June in blocks of 20 plots per block, with inter- and intra-spatial block spaces of 0.5 m apart. The seedlings were transplanted in holes created in the crest of the sterilized topsoil and irrigated to field capacity every three days until 6 months after transplanting (MAT).

4.2. Genotyping

Approximately 1 g young, healthy and fully expanded leaves were collected per genotype. Approximately 54 samples comprising 50 progenies, three putative male and one maternal-bearing fruit parents were collected per family. Genomic DNA was extracted from representative samples using modified CTAB protocol with slight modification [51]. The genomic DNA quality and concentration were done using agarose gel and nanodrop as described by Aljanabi and Martinez [52]. Concentrated DNA of 50 μL from each sample was sent to Diversity Array Technology (DArT) Pty Ltd., Canberra, Australia for sequencing.

4.3. Analysis

The raw HapMap file from DArTseq genotyping platform (Australia) was first converted to a variant call format (VCF) file using perl programming language and TASSEL v.5.2.43 [53,54]. The output was scored for presence or absence of DNA fragments in genomic sequences generated from genomic DNA samples. A total of 20,000 SNP derived DArTseq markers were identified from the raw data and subjected to filtering using VCFtools [55] at the following conditions: MAF (0.05), no missing data, depth (>5), genotype quality (GQ = 20), maximum and minimum allele = 2, reproducibility = 1 and no indels. Noninformative markers or missing data (in clone and SNP) in the raw data of the VCF file were filtered out. Of the 20000 SNPs subjected to filtering, highly informative 6602 were retained for different analyses. Dosage format was generated for the 6602 informative SNPs using plink and the recodeA function alongside with the double-id [56].

Various population genetic analyses were conducted to explore the genetic properties of the markers. The number of alleles and allele frequencies for the selected SNPs were estimated using the VCFtools [55]. The proportion of heterozygous genotypes also referred to as the gene diversity (GD) of a locus under Hardy-Weinberg equilibrium [57], the minor allele frequency, major allele frequency and the proportions of nonmissing and missing gametes were assessed using TASSEL software version 5.2.51 [53] and the plink command–freq and –hardy [56].

The Penalized and doMC libraries were used to estimate the contribution of the male and female parents in the offspring. The value of each parent was estimated using the multinomial model of the penalized function formula as described by McIlhagga [58].

The log-likelihood of parentage of the observations studied was calculated as:

$$L = \sum_{i=1}^{n} \sum_{j=1}^{q} \mathrm{y_{i,j}} \log \mathrm{p_{i,j}}. \quad (1)$$

where the (*i*, *j*)th element $\mathrm{y_{i,j}}$ is the number of times category *j* occurred in observation *i*. Each element $\mathrm{y_{i,j}}$ has a certain probability $\mathrm{p_{i,j}}$ of occurrence.

$$\mathrm{P}_j = \begin{bmatrix} p_{1,j} \\ p_{2,j} \\ \vdots \\ p_{\mathrm{n},j} \end{bmatrix} \quad (2)$$

The observation vectors $\mathrm{y_1}$, $\mathrm{y_2}$, ..., $\mathrm{y_q}$ were stacked into a single vector:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\mathrm{q}} \end{bmatrix} \quad (3)$$

The probability vectors p_1, p_2, ..., p_{q} were used to calculate a single vector:

$$p = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{\mathrm{q}} \end{bmatrix} \quad (4)$$

The probability vector (p_{k}) of parental contribution to progeny was estimated as:

$$p_{\mathrm{k}} = \frac{\exp\left(\eta_{\mathrm{k,k}}\right)}{\sum_{j=1}^{q} \exp\left(\eta_{\mathrm{k,j}}\right)} \quad (5)$$

where each vector $\eta_{\mathrm{k,k}} = \widetilde{\mathrm{X}}\beta_{\mathrm{j}}$ and the exponentiation and division are both element-by-element; $\widetilde{\mathrm{X}}$ is a matrix of covariates. The values obtained were then expressed as percentage. The coefficient vectors β_{j} was stacked to form a single coefficient matrix.

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{\mathrm{q}} \end{bmatrix} \quad (6)$$

The gradient (intercept) of *L* with respect to β was estimated using the chain rule as:

$$\frac{dL}{d\beta} = \frac{dp}{d\beta} \frac{dL}{dp} \quad (7)$$

The gradient dL/dp is simply y/p, where the division is element-by-element. The Jacobian $dp/d\beta$ is

$$\frac{dp}{d\beta} = \begin{bmatrix} dp_1/d\beta_1 & dp_2/d\beta_1 & \cdots & dp_{\mathrm{q}}/d\beta_1 \\ dp_1/d\beta_2 & dp_2/d\beta_2 & \cdots & dp_{\mathrm{q}}/d\beta_2 \\ \vdots & \vdots & \ddots & \vdots \\ dp_1/d\beta_{\mathrm{q}} & dp_2/d\beta_{\mathrm{q}} & \cdots & dp_{\mathrm{q}}/d\beta_{\mathrm{q}} \end{bmatrix} \quad (8)$$

Each of the submatrices $dp_{\mathrm{k}}/d\beta_{\mathrm{j}}$ is a Jacobian matrix represented as:

$$\frac{dp_{\mathrm{k}}}{d\beta_{\mathrm{j}}} = \frac{dp_{\mathrm{k},1}}{d\beta_{\mathrm{j}}}\frac{dp_{\mathrm{k}}}{d\eta_{\mathrm{k},1}} + \frac{dp_{\mathrm{k},2}}{d\beta_{\mathrm{j}}}\frac{dp_{\mathrm{k}}}{d\eta_{\mathrm{k},2}} + \cdots + \frac{dp_{\mathrm{k},1}}{d\beta_{\mathrm{j}}}\frac{dp_{\mathrm{k}}}{d\eta_{\mathrm{k,q}}} = X^{\mathrm{T}}\frac{dp_{\mathrm{k}}}{d\eta_{\mathrm{k,j}}} \quad (9)$$

where $dp_{\mathrm{k}}/d\eta_{\mathrm{k,j}}$ is a diagonal matrix of derivatives.

The values generated for each progeny were ranked and the true hybrids were identified based on the estimated male and female coefficients. The profile values of progeny samples of each family and maternal parent were surrounded by the three putative male parents and analyzed for possible alignments. Helium pedigree visualization software was used for pedigree transmission patterns, and visualization of the various polycross families generated as described by Shaw et al. [59] and Shaw [60]. Figure 9 illustrates the process employed for parentage analysis in this paper.

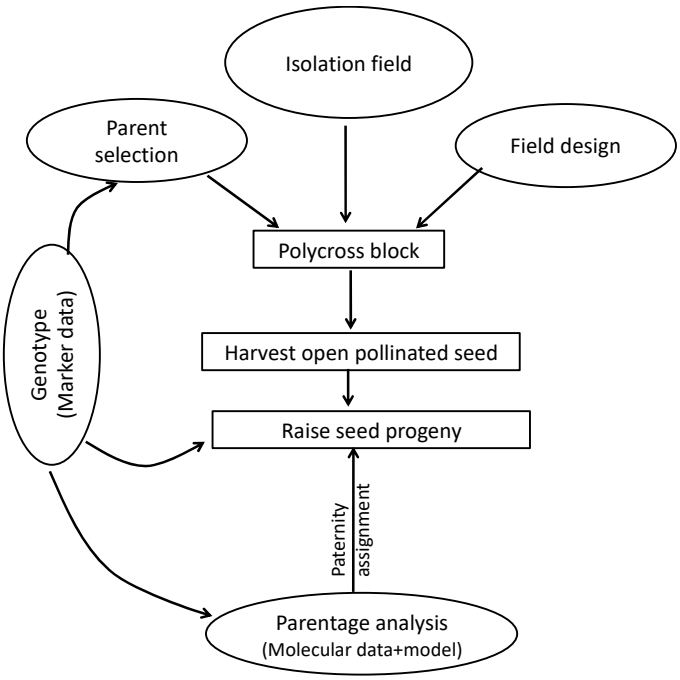

Figure 9. A flow diagram showing the experimental procedure of parentage analysis in white yam.

5. Conclusions

In the current study, the identification of parentage relationships in the half-sib progenies of white Guinea yam using SNP markers aids the accurate traceability of genotypes with unknown identity, and determination of genetic diversity across families and generations that could be exploited for breeding.

Supplementary Materials: The following materials are available online at www.mdpi.com/2223-7747/9/4/527/s1, Table S1. Contributions of the female (Ojuiyawo**)** and putative male (TDr9501932, TDr9902789 and TDr9902607) parents to progeny of family TDr1692

Author Contributions: P.E.N. and A.A. designed the experiment. P.E.N. executed the experiment. A.A., P.B.T., E.Y.D. and A.D. supervised the work. P.E.N., A.A., A.A.P. and U.N.I. performed the data analysis. P.E.N. and A.A. wrote the manuscript. All the authors contributed to writing the article, read and approved its submission. All authors have read and agreed to the published version of the manuscript.

Funding: The work was financially supported by Bill and Melinda Gates Foundation (BMGF) through AfricaYam project (OPP1052998) to IITA and the International Development Research Centre (IDRC) under the IDRC/CORAF-WECARD/IITA sub-grant agreement for developing capacity for Agricultural Research in Sub-Saharan Africa (PJ-2126) to PEN (PhD student) at the West African Centre for Crop Improvement (WACCI), University of Ghana.

Acknowledgments: The authors are thankful to IITA Bioscience Centre, the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria, for facilitation with DNA extraction and analysis. Technical support from entire yam breeding team, IITA, Ibadan, Nigeria, is dully acknowledged.

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

Abbreviations

CTAB	cetyl trimethyl ammonium bromide
DArT	diversity array technology
DNA	deoxyribonucleic acid
IITA	international institute of tropical agriculture
ISRIC	international soil reference and information centre
MAT	months after transplanting
SNP	single nucleotide polymorphism
VCF	variant call format
WAS	weeks after planting

References

1. 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: Kew, London, UK, 2007; pp. 1–65.
2. Bai, K.V.; Ekanayake, I.J. Taxonomy, morphology and floral biology. In *Food Yams: Advances in Research*; Orkwor, G.C., Asiedu, R., Ekanayake, I.J., Eds.; NRCRI and IITA: Ibadan, Nigeria, 1998; pp. 13–38.
3. Lebot, V. *Tropical Root and Tuber Crops: Cassava, Sweet Potato, Yams and Aroids*; CABI: Wallingford, UK, 2009; 413p.
4. Mignouna, H.D.; Abang, M.M.; Onasanya, A.; Agindotan, B.; Asiedu, R. Identification and potential use of RAPD markers linked to *Yam mosaic virus* resistance in white yam (*Dioscorea rotundata* Poir). *Ann. Appl. Biol.* **2002**, *140*,163–169.
5. 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. Gen*. **2005**, *111*, 226–232.
6. Sybenga, J. Cytogenetics in Plant Breeding. In *Monograph on Theoretical and Applied Genetics 17*; Springer Science and Business Media; Technology and Engineering; Springer: Berlin, Germany, 2012; p. 469.
7. Stebbins, G. Types of polyploids: Their classification and significance. *Adv. Genet.* **1947**, *1*, 403–429.
8. Stift, M.; Berenos, C.; Kuperus, P.; van Tienderen, P.H. Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: A general procedure applied to *Rorippa* (yellow cress) microsatellite data. *Genetics* **2008**, *179*, 2113–2123.
9. Obidiegwu, J.E.; Akpabio, E.M. The geography of yam cultivation in southern Nigeria: Exploring its social meanings and cultural functions. *J. Ethn. Foods* **2017**, *4*, 28–35.
10. Alabi, T.R.; Adebola, P.O.; Asfaw, A.; De Koeyer, D.; Lopez-Montes, A.; Asiedu, R. Spatial multivariate cluster analysis for defining target population of environments in West Africa for yam breeding. *Int. J. Appl. Geo. Res*. **2019**, *10*, 1–30.
11. Darkwa, K.; Olasanmi, B.; Asiedu, R.; Asfaw, A. Review of empirical and emerging methods and tools for yam (*Dioscorea* spp.) improvement: Status and prospects. *Plant Breed.* **2019**, 1–24.
12. Lebot, V.; Abraham, K.; Kaoh, J.; Rogers, C.; Molisalé, T. Development of anthracnose resistant hybrids of the Greater yam (*Dioscorea alata* L.) and interspecific hybrids with *D. nummularia* Lam. *Gen. Res. Crop Evol*. **2019**, *66*, 871–883.
13. Norman, P.E.; Asfaw, A.; Tongoona, P.B.; Danquah, A.; Danquah, E.Y.; De Koeyer, D.; Asiedu, R. Pollination success in some white yam genotypes under polycross and nested mating designs. *Int. J. Biol. Sci. Appl*. **2018**, *5*, 19–28.
14. 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.
15. Telfer, E.J.; Stovold, G.T.; Li, Y.; Silva-Junior, O.B.; Grattapaglia, D.G.; Dungey, H.S. Parentage reconstruction in *Eucalyptus nitens* using SNPs and microsatellite markers: A comparative analysis of marker data power and robustness. *PLoS ONE* **2015**, *10*, e0130601.

16. Herbinger, C.M.; Doyle, R.W.; Pitman, E.R.; Paquet, D.; Mesa, K.A.; Morris, D.B.; Wright, J.M.; Cook, D. DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. *Aquaculture* **1995**, *137*, 245–256.
17. Zoundjihékpon, J.; Hamon, S.; Tio-Touré, B.; Hamon, P. First controlled progenies checked by isozyme markers in cultivated yams Dioscorea cayenensis-rotundata. *Theor. Appl. Genet*. **1994**, *88*, 1011–1016.
18. Sartie, A.; Asiedu, R. Development of mapping populations for genetic analysis in yams (*Dioscorea rotundata* Poir and *Dioscorea alata* L). *Afr. J. Biotech*. **2011**, *10*, 3040–3050.
19. Mammadov, J.; Aggarwal, R.; Buyyarapu, R.; Kumpatla, S. SNP markers and their impact on plant breeding. *Int. J. Plant Genom.* **2012**, *2012*, 1–11.
20. Nybom, H.; Weising, K.; Rotter, B. DNA fingerprinting in botany: Past, present, future. *Investig. Genet*. **2014**, *5*, 1–35.
21. Shete, S.; Tiwari, H.; Elston, R.C. On estimating the heterozygosity and polymorphism information content value. *Theor. Popul. Biol*. **2000**, *57*, 265–271.
22. Pruvost, N.B.M.; Hoffmann, A.; Reyer, H.U. Gamete production patterns, ploidy, and population genetics reveal evolutionary significant units in hybrid water frogs (Pelophylax exculentus). *Ecol. Evol*. **2013**, *3*, 2933–2948.
23. Petit, R.J.; El Mousadik, A.; Pons, O. Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol*. **1998**, *12*, 844–855.
24. Zeng, Z.B.; Cockerham, C.C. Long-term response to artificial selection with multiple alleles—study by simulations. *Theor. Pop. Biol*. **1990**, *37*, 254–272.
25. Wagner, A. Robustness and evolvability: A paradox resolved. *Proc. R. Soc. B Biol. Sci*. **2008**, *275*, 91–100.
26. Greenbaum, G.; Templeton, A.R.; Zarmi, Y.; Bar-David, S. Allelic richness following population founding events–A stochastic modeling framework incorporating gene flow and genetic drift. *PLoS ONE* **2014**, *9*, e115203.
27. Norman, P.E.; Asfaw, A.; Tongoona, P.B.; Danquah, A.; Danquah, E.Y.; Koeyer, D.D.; Asiedu, R. Can parentage analysis facilitate breeding activities in root and tuber crops? *Agric. J*. **2018**, *8*, 1-24.
28. Snow, A.A.; Spira, T.P. Pollen-tube competition and male fitness in *Hibiscus moscheutos*. *Evolution* **1996**, *50*, 1866–1870.
29. Pasonen, H.L.; Pulkkinen, P.; Käpylä, M.; Blom, A. Pollen-tube growth rate and seed-siring success among *Betula pendula* clones. *New Phytol*. **1999**, *143*, 243–251.
30. Aronen, T.; Nikkanen, T.; Harju, A.; Tiimonen, H.; Häggman, H. Pollen competition and seed-siring success in *Picea abies*. *Theor. Appl. Genet*. **2002**, *104*, 638–642.
31. Spira, T.; Snow, A.A.; Puterbaugh, M.N. The timing and effectiveness of sequential pollinations in Hibiscus moscheutos. *Oecologia* **1996**, *105*, 230–235.
32. Parantainen, A.; Pasonen, H.L. Pollen-pollen interactions in Pinus sylvestris. *Scand. J. For. Res*. **2004**, *19*, 199–205.
33. Nakamura, R.R.; Wheeler, N.C. Pollen competition and paternal success in Douglas-fir. *Evolution* **1992**, *46*, 846–851.
34. Moriguchi, Y.; Ishiduka, D.; Kaneko, T.; Itoo, S.; Taira, H.; Tsumura, Y. The Contribution of pollen germination rates to uneven paternity among polycrosses of *Cryptomeria japonica*. *Silvae Genet.* **2009**, *58*, 139–144.
35. Mignouna, H.D.; Abang, M.M.; Asiedu, R.; Geeta, R. Producing yam (Dioscorea) seeds through controlled crosses. *Cold Spring Harbor Protoc.* **2009**, prot5327, doi:10.1101/pdb.prot5327.
36. Segnou, M.; Fatokun, C.A., Akoroda, M.O.; Hahn, S. Studies on the reproductive biology of white yam (*Dioscorea rotundata* Poir.). *Euphytica* **1992**, *64*, 197–203, doi:10.1007/BF00046049.
37. Ardlie, K.G.; Lunetta, K.L.; Seielstad, M. Testing for population subdivision and association in four case-control studies. *Amer. J. Hum. Genet*. **2002**, *71*, 304–311.
38. Vos, P.G.; Uitdewilligen, J.G.; Voorrips, R.E.; Visser, R.G.; van Eck, H.J. Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): An insight into the breeding history. *Theor. Appl. Genet*. **2015**, *128*, 2387–2401, doi:10.1007/s00122-015-2593-y.
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.

40. 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.
41. 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.
42. 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.
43. Smyda-Dajmund, P.; Sliwka, 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.
44. 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.
45. 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.
46. Jones, A.G.; Ardren, W.R. Methods of parentage analysis in natural populations. *Mol. Ecol*. **2003**, *12*, 2511–2523.
47. 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.
48. 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*, 6293847.
49. Exposito-Alonso, M.; Becker, C.; Schuenemann, V.J.; Reiter, E.; Setzer, C.; Slovak, R.; Brachi, B.; Hagmann, J.; Grimm, D.G.; Chen, J.; et al. The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet*. **2018**, *14*, e1007155.
50. International Soil Reference and Information Centre and the Food and Agriculture Organization (ISRIC/FAO). *Procedures for Soil Analysis*, 6th ed.; Technical Paper 9; FAO: Roma, Italy, 2002; p. 120.
51. Dellaporta, S.L.; Wood, J.; Hicks, J.B. A plant DNA mini preparation Version II. *Plant Mol. Biol. Rep*. **1983**, *1*, 19–21.
52. Aljanabi, S.M.; Martinez, I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res*. **1997**, *25*, 4692–4693.
53. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**, *23*, 2633–2635.
54. 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.
55. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158.
56. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet*. **2007**, *81*, 559–575.
57. Nei, M. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 3321–3323.
58. McIlhagga, W. Penalized: A MATLAB Toolbox for fitting generalized linear models with penalties. *J. Stat. Soft* **2016**, *72*, 1–21.
59. Shaw, P.D.; Graham, M.; Kennedy, J.; Milne, I.; Marshall, D.F. Helium: Visualization of large scale plant pedigrees. *BMC Bioinform.* **2014**, *15*, 259.
60. Shaw, P.D. Visualizing Genetic Transmission Patterns in Plant Pedigrees. Ph.D. Thesis, Edinburgh Napier University, Edinburgh, Scotland, 2016; p. 221.