agriculture

MDPI

Article

Genomic Prediction for Inbred and Hybrid Polysomic Tetraploid Potato Offspring

Rodomiro Ortiz [1,*], Fredrik Reslow [1], Ramesh Vetukuri [1], M. Rosario García-Gil [2], Paulino Pérez-Rodríguez [3] and José Crossa [4]

[1] Department of Plant Breeding, Swedish University of Agricultural Sciences (SLU), P.O. Box 190, SE 23436 Lomma, Sweden; fredrik.reslow@slu.se (F.R.); ramesh.vetukuri@slu.se (R.V.)
[2] Umeå Plant Science Center, SLU Department of Forest Genetics and Plant Physiology, SE 90183 Umeå, Sweden; m.rosario.garcia@slu.se
[3] Colegio de Postgraduados (COLPOS), Montecillos 56230, Edo. de México, Mexico; perpdgo@gmail.com
[4] International Maize and Wheat Improvement Center (CIMMYT), Carretera México-Veracruz Km. 45, El Batán, Texcoco 56237, Edo. de México, Mexico; j.crossa@cgiar.org
* Correspondence: rodomiro.ortiz@slu.se

Abstract: Potato genetic improvement begins with crossing cultivars or breeding clones which often have complementary characteristics for producing heritable variation in segregating offspring, in which phenotypic selection is used thereafter across various vegetative generations (T_i). The aim of this research was to determine whether tetrasomic genomic best linear unbiased predictors (GBLUPs) may facilitate selecting for tuber yield across early T_i within and across breeding sites in inbred (S_1) and hybrid (F_1) tetraploid potato offspring. This research used 858 breeding clones for a T_1 trial at Umeå (Norrland, 63°49′30″ N 20°15′50″ E) in 2021, as well as 829 and 671 clones from the breeding population for T_2 trials during 2022 at Umeå and Helgegården (Skåne, 56°01′46″ N 14°09′24″ E), respectively, along with their parents (S_0) and check cultivars. The S_1 and F_1 were derived from selfing and crossing four S_0. The experimental layout was an augmented design of four-plant plots across testing sites, where breeding clones were non-replicated, and the parents and cultivars were placed in all blocks between the former. The genomic prediction abilities (r) for tuber weight per plant were 0.5944 and 0.6776 in T_2 at Helgegården and Umeå, respectively, when T_1 at Umeå was used as the training population. On average, r was larger in inbred than in hybrid offspring at both breeding sites. The r was also estimated using multi-environment data (involving at least one S_1 and one F_1) for T_2 performance at both breeding sites. The r was strongly influenced by the genotype in both S_1 and F_1 offspring irrespective of the breeding site.

Keywords: *Solanum tuberosum*; crossing; polyploidy; genomic estimated breeding values; linear models; Nordic latitude; selfing; tetrasomic inheritance

check for updates

Citation: Ortiz, R.; Reslow, F.; Vetukuri, R.; García-Gil, M.R.; Pérez-Rodríguez, P.; Crossa, J. Genomic Prediction for Inbred and Hybrid Polysomic Tetraploid Potato Offspring. *Agriculture* **2024**, *14*, 455. https://doi.org/10.3390/agriculture14030455

Academic Editor: John C. Snyder

Received: 7 January 2024
Revised: 3 March 2024
Accepted: 8 March 2024
Published: 11 March 2024

1. Introduction

A cycle of breeding tetraploid ($2n = 4x = 48$ chromosomes) potatoes (*Solanum tuberosum*) takes between one to one and a half decades from crossing to identifying superior germplasm for potential cultivar release. The main challenge in improving breeding efficiency in this tuber crop which shows tetrasomic inheritance remains, therefore, identifying promising breeding clones in early clonal generations (T_i). Reducing the breeding cycle increases genetic gains, which thus makes it a main priority for improving the genetic enhancement of the potato crop.

During the first year, a few dozens to several hundreds of F_1 full-sib family offspring ensue from crossing blocks of small or large potato breeding programs, respectively [1,2]. The segregated F_1 offspring (T_1) resulting from crossing two heterozygous potato parents are very heterogeneous and exhibit the maximum intrafamily genetic variance. Selecting efficiently at this stage may significantly improve genetic gains by increasing the intensity of selection due to the quality of the material that advances as T_2 to further stages of clonal selection in the target population of environments.

Visual selection of promising F_1 clones in early segregating generations is ineffective in tetrasomic potato breeding [3–5]. Tuber yield is, however, an important characteristic for visual preference scoring in T_1 and T_2 [6,7]. In this regard, family selection for tuber yield has been proposed for early (T_1 and T_2) clonal generations [8] since interfamily heritability appears to be higher than intrafamily heritability [9,10] as there is less of an environmental effect seen in the family means than within each family [4,11,12]. Unfortunately, family selection for tuber yield in the greenhouse does not seem feasible in pots whose diameter affects tuber characteristics [13]. Although, as indicated by Ticona-Benavente et al. [14], there is a lack of repeatability for tuber yield among individual clones across early (T_1–T_4) clonal generations, the mean ranking of a cross remains the same across sites and over years in potatoes [15], thus showing repeatability in the early generations. The tuber yield of F_1 clones in non-replicated T_1 plots cannot be used for selection, while family selection seems feasible particularly when, as noted by Brown and Caligari [16], considering both the family mean and its within-progeny standard deviation. Nonetheless, selection in T_1 may be useful for rejecting a family [17] or individual clones which show irregular tuber shapes or deep eyes [18], rather than for the selection of individual clones [19], which may randomly reduce the number of genotypes for further testing.

The use of best linear unbiased predictors (BLUPs) for between- or within-family selection seems to increase genetic gains for tuber yield in the potato crop [20]. BLUPs take into account pedigrees to utilize additive genetic variance [21]. Moreover, Slater et al. [22,23] demonstrated the advantage of BLUP-estimated breeding values (EBVs) for predicting the performance of low heritability traits in potatoes. These BLUPs provide EBVs for each family, individuals within each family and all cultivars in their pedigree [24]. Slater et al. [25] also proposed the use of selection based on genomic EBVs (GEBVs) for increasing genetic gains. In this regard, Ortiz et al. [26] have shown that multi-environment (ME) modeling increases GEBV prediction accuracy.

Jinks and Lawrence [27] presented a compelling challenge to the prevailing notion that heterozygotes are optimal for achieving high crop yields. They argued that, contrary to conventional wisdom, inbreeding plays a pivotal role in elevating the prevalence of beneficial alleles while simultaneously solidifying and eliminating homozygotes carrying harmful recessive variants. Moreover, Jinks [28,29] proposed that the manifestation of heterosis is not exclusively contingent on heterozygosity itself but is intricately linked to the genotype. This perspective opens up the possibility of fixing advantageous homozygous recombinants through inbreeding, thereby contributing to the enhanced performance of crops. In this regard, Ortiz et al. [30] recently pointed out that GEBVs may lead to eliminating undesired inbred or hybrid offspring in potato breeding. Hence, the objectives of our research, from which this manuscript ensued, were to assess the value of genomic best linear unbiased predictors (GBLUPs) based on the reaction norm model [31] for early generation selection for tuber yield in the potato crop, as well as the effect of inbreeding on genomic prediction ability (r) for this characteristic, thus emphasizing its critical role in enhancing selection accuracy and efficiency.

2. Materials and Methods

Our multi-site trials over two years used up to 858 hybrid (F_1) and inbred (S_1) clones resulting from the 2020 crossing block of Svenska potatisförädling, which is about 10% of the population size for year 2021 of this potato breeding program targeting Fennoscandia [2]. Four cultivars (S_0), namely Colleen (Ireland, 1991, Manna × Mizzen), Melody (The Netherlands, 2001, VE 74–75 × W 72–22–496), Queen Anne (Germany, 2012, SA 99–002–44 × Gala) and Rudolph (Great Britain, 2006, Chieftain × Stirlinh'g), were used as parents. Colleen's tubers are used for boiling and frying; while tubers of Melody are suitable for boiling, mashing and wedging. The tubers of Queen Anne are appropriate for frying, mashing and roasting, and those of Rudolph for baking, boiling and frying. True seeds from the crossing block were grown as seedlings in a greenhouse. The tuberlings (i.e., tubers derived from seedlings) for each offspring were grown in a field along with their parents for producing the next year's tubers for planting.

The tuber harvests of field trials included the four parents and 858 breeding clones at Umeå (63°49′30″ N 20°15′50″ E, Norlland, Sweden) in 2021, as well the parents and 829 clones at same site in 2022, and the parents and 671 clones at Helgegården (56°01′46″ N 14°09′24″ E, Skåne, Sweden) in 2022 (Table 1). The field layouts in each site were augmented designs with 4-plant plots, in which the four parents were replicated in each of the incomplete blocks. The spacing was 0.7 m between rows and 0.3 m spacing among plants within the plot. The tubers used for planting 'Rudolph' did not sprout well in 2021, and this cultivar was not included in further data analysis at Umeå in 2021 due to the number of unevenly lost plants in each of the nine blocks. The characteristic evaluated was tuber yield per plant.

Table 1. Total number of offspring after selfing (S_1) and crossing (F_1) used for trial plantings.

S_1 or F_1	Offspring Number
Colleen S_1 (C S_1)	162
Melody S_1 (M S_1)	94
Queen Anne S_1 (QA S_1)	177
Rudolph S_1 (R S_1)	173
Queen Anne × Colleen (QA × C)	36
Queen Anne × Melody (QA × M)	272

The effect of the block was accounted for when running the spatial analysis for each testing site by means of the AR1 × AR1 models [32], which adjusted the data by the row and column effects simultaneously [33,34]. Heritability and intra-class correlation estimates were based on the components of variance. The intra-class correlation (${I_R}^2$) was estimated as follows:

$${I_R}^2 = {\sigma^2}_A/[{\sigma^2}_A + {\sigma^2}_W]$$

where ${\sigma^2}_A$ and ${\sigma^2}_W$ are the variances among and within the six offspring, respectively. ${I_R}^2$ is the degree of resemblance of the among group to the total variance [35].

Targeting genotyping using a genotype-by-sequencing approach [36] was used for characterizing the S_0, S_1 and F_1 with 2340 single-nucleotide polymorphisms (SNPs), which were previously used in genomic prediction of breeding values for cultivars released in western Europe along with Svensk potatisförädling clones [25]. We removed markers whose minor allele frequency was smaller than 0.05; after that, 2194 SNPs were available for further analyses. Thirteen of these SNP—distributed across the 10 potato chromosomes—had highly significant marker trait association with tuber yield per plant in the Svensk potatisförädling population (unp. results). The SNPs had five different allelic stages, which ranged from 0 to 4. In this scale, 0 and 4 are the two homozygotes (OOOO or nulliplex and AAAA or quadriplex), while 1, 2 and 3 refer to simplex (AOOO), duplex (AAOO) or triplex (AAAO) genotypes [25]. We computed the genomic relationship matrix for a full tetrasomic polyploid model including non-additive effects using the method proposed by Slater et al. [25] as implemented in the R package AGHmatrix [37].

2.1. Statistical Models

We used the reaction norm model [30] as follows:

$$\boldsymbol{y} = \boldsymbol{Z}_E \boldsymbol{\beta}_E + \boldsymbol{Z}_g \boldsymbol{g} + \boldsymbol{u} + \boldsymbol{e}, \ldots \quad (1)$$

where $\boldsymbol{y}$ is the response vector (total tuber yield per plant); $\boldsymbol{Z}_E$ is the incidence matrix for environments (Umeå 2021, Umeå 2022 and Helgegården 2022); $\boldsymbol{\beta}_E$ is a vector of environmental effects, $\boldsymbol{\beta}_E \sim MN\left(0, \sigma_E^2\boldsymbol{I}\right)$, where σ_E^2 is the variance parameter associated with environment and $\boldsymbol{I}$ is the identity matrix and MN stands for multivariate normal distribution; $\boldsymbol{Z}_g$ is an incidence matrix that connects phenotypes with genotypes. $\boldsymbol{g}$ is a vector of random effects due to additive genotypic effects of genotypes; we assume $\boldsymbol{g} \sim MN\left(0, \sigma_g^2\boldsymbol{G}\right)$, where $\boldsymbol{G}$ is the genomic relationship matrix, σ_g^2 is the variance parameter associated with the genotypes. $\boldsymbol{u}$ is a random effect that takes into account the interaction between genotype and environments and $\boldsymbol{u} \sim MN\left(0, \sigma_{g\times E}^2\boldsymbol{Z}_g\boldsymbol{G}\boldsymbol{Z}_g^t\#\boldsymbol{Z}_E\boldsymbol{Z}_E^t\right)$, where # denotes the Haddamard product (cell by cell) between two matrices, and $\sigma_{g\times E}^2$ is the variance component associated with the interaction. Finally, $\boldsymbol{e} \sim NM\left(0,\ \sigma_e^2\boldsymbol{I}\right)$, where σ_e^2 is the variance parameter associated with the error. We assume that the random terms are independently distributed. We fitted model (1) using all available records for the three environments (Umeå 2021, Umeå 2022, Helgegården 2022) using the Bayesian generalized linear regression (BGLR) package [36] within the Bayesian framework and using default priors and hyperparameters provided by the software. Inferences were based on 15,000 MCMC iterations and obtained after discarding 15,000 iterations. Then, we obtained the posterior distributions for the variance parameters and the posterior means.

Cross-Validation

We investigated the prediction ability of the model considering the same breeding populations as in Ortiz et al. [30], i.e., four S_1 and their two related F_1 offspring. We extended the original prediction problems in Ortiz et al. [30] to multi-environment settings, so we considered the following problems: (1) prediction of the response variable in one environment using the data registered in another environment; (2) prediction among full sibs; and (3) prediction of non-related individuals and half sibs.

In prediction problem 1, we predicted the response variable in both Umeå 2022 and Helgegården 2022 independently using the phenotypic and genotypic information registered in Umeå 2021 as training data. We computed the Pearson's correlation between observed and predicted tuber yield for observations in Umeå 2022 and Helgegården 2022. For prediction problem 2, the training consisted of all phenotypic records for a given population evaluated in Umeå 2021 and 70% of the records of that population evaluated in Umeå 2022 or Helgegården 2022. The testing consisted of 30% of the records for a given population evaluated in Umeå 2022 or Helgegården 2022. We generated 50 partitions at random with training and testing sets defined as described. For each partition, we computed Pearson's correlation between observed and predicted tuber yield per plant. In the case of prediction problem 3, we predicted breeding populations of individuals evaluated in Umeå 2022 and Helgegården 2022 using records of non-related individuals with the non-breeding populations evaluated in Umeå 2021 and evaluations of the same population to be predicted from previous years.

2.2. Software

All computations were performed using the R statistical package [38]. Model (1) was fitted using the package BGLR [39] within the Bayesian framework and using default priors and hyperparameters provided by the software. Scripts are similar to those from Pérez-Rodríguez et al. [40]. Inferences were based on 15,000 MCMC iterations and obtained after discarding 15,000 iterations. The 15,000 iterations were subsequently thinned with a thinning parameter equal to 10 and posterior means of the parameters of interest were obtained. Convergence was visually inspected using the trace plot of variance parameters.

3. Results

Heritability estimates for total tuber yield were 0.70 for Umeå in 2021, 0.24 for Umeå in 2022 and 0.48 for Helgegården in 2022, while the intra-class correlation estimates for this characteristic were 0.50, 0.46 and 0.71 in each of these environments, respectively. There were significant differences ($p < 0.05$) in the tuber yields between S_0, S_1 and F_1 in each of the two testing sites over the two years (Table 2). The S_0 parents always had a greater tuber yield per plant than the inbred offspring (S_1). The average total tuber weight of the S_1 was, on average, less than half of the S_0 across sites over the two years, thus indicating a significant inbreeding depression for this tuber characteristic. The inbreeding depression was significantly higher in Umeå 2021 (tuber yield of the S_1 was below $\frac{1}{4}$ of the S_0) than in Umeå 2022 and Helgegarden 2022, where the S_1 tuber yield was 40 and 50% less than that of the S_0, respectively. The inbreeding depression for tuber yield was more noticeable in Queen Anne (36% of the S_0 was the tuber yield per plant in the S_1 across the testing environments) than in Colleen (47%), Melody (50%) and Rudolph (56%). The two F_1 were, on average, significantly lower ($p < 0.05$) than the parents in Umeå, which suggests a lack of average heterosis in the hybrid offspring in this site. Nonetheless, Queen Anne and Melody's F_1 offspring's total tuber yield was higher than that of the S_0 by 30% at Helgegården 2022, thus revealing transgressive segregation for this characteristic in the heterogeneous F_1 family at this testing site. The average total tuber yield of the F_1 was always significantly ($p < 0.05$) above that of the S_1 across testing sites over years, i.e., the F_1 tuber yield was 89% higher than that of the S_1, which was more noticeable at Umeå 2021, where the F_1 offspring's tuber yield was more than double that of the S_1. The S_0 always had the highest tuber yield per plant in Umeå (almost doubling in 2021 and about 11% higher in 2022) during the 90-day cropping season with about 14.5–ca. 21 h daylength, but was almost the same in Helgegården with a 120-day cropping season and a shorter daylength than in Umeå.

Table 2. Average tuber weight (kg plant^{-1}) of cultivars and their first inbred (S_1) and F_1 hybrid offspring at Umeå (Norlland, Sweden) and Helgegården (Skåne, Sweden) in 2021 (first clonal generation T_1 for S_1 and F_1) and 2022 (second clonal generation T_2 for S_1 and F_1).

Genotype	Umeå		Helgegården
	2021	2022	
Colleen	0.693	0.867	1.594
Colleen S_1	0.181	0.601	0.733
Melody	0.826	1.128	1.746
Melody S_1	0.195	0.739	1.040
Queen Anne	0.562	1.145	1.585
Queen Anne S_1	0.124	0.392	0.807
Queen Anne × Colleen	0.290	0.857	1.189
Queen Anne × Melody	0.386	0.981	2.058
Rudolph	N/A	1.100	1.472
Rudolph S_1	0.222	0.759	0.619
$LSD_{0.05}$	0.094	0.125	0.279

Figures S1 and S2 shows scatter plots of the observed vs. predicted total tuber yield per plant in Helgegården 2022 and Umeå 2022 (Problem 1). The Pearson's correlation coefficients between the observed and predicted values are shown in the respective legend. The correlation for Helgegården 2022 was 0.5944, whereas in Umeå 2022 the correlation was 0.6766. In the latter, the higher value could be attributed to the fact that both the T_1 and T_2 populations were grown in the same breeding site, i.e., Umeå. Tables 3 and 4 list the correlation between the observed and predicted tuber yield per plant in one environment using the data registered in another environment (Problem 2). The correlations are the average across 50 partitions and their corresponding standard deviation (SD) are included to assess the variability of the prediction ability of the model (Figures S3 and S4).

Table 3. Genomic prediction ability (r) for tuber yield (kg plant^{-1}) at Helgegården 2022 for full sibs.

Breeding T_2 population (30% of the records for a given population) in Helgegården 2022	(QA × C)	(QA × M)	(C S_1)	(M S_1)	(QA S_1)	(R S_1)
Training T_1 population in Umeå 2021 and 70% of that population evaluated in Helgegården 2022	(QA × C)	(QA × M)	(C S_1)	(M S_1)	(QA S_1)	(R S_1)
r	0.503	0.141	0.262	0.392	0.683	0.389
SD	0.210	0.129	0.113	0.162	0.087	0.122

Table 4. Genomic prediction ability (r) for tuber yield (kg plant^{-1}) at Umeå 2022 for full sibs.

Breeding T_2 population (30% of the records for a given population) in Umeå 2022	(QA × C)	(QA × M)	(C S_1)	(M S_1)	(QA S_1)	(R S_1)
Training T_1 population in Umeå 2021 and 70% of that population evaluated in Umeä 2022	(QA × C)	(QA × M)	(C S_1)	(M S_1)	(QA S_1)	(R S_1)
r	0.607	0.428	0.514	0.448	0.769	0.517
SD	0.23	0.116	0.092	0.172	0.058	0.109

The variance in the phenotypes was 0.3694 and the variance associated with the interaction was estimated as $\hat{\sigma}^2_{g \times E} = 0.1094$ (Figure 1). The interaction explains about 30% of the phenotypic variance. Hence, it was necessary to fit a model that includes the genotype-by-environment interaction. Tables 5 and 6 provide the prediction ability of model (1) for individuals evaluated in Helgegården 2022 and Umeå 2022, respectively, using records of non-related individuals evaluated in Umeå 2021 and of the same population to be predicted evaluated in 2022 (Problem 3). The training and training sets are clearly noted in each table. On average, the prediction ability for tuber yield per plant for each family, as computed by the Pearson's correlation between the observed and predicted values for this characteristic, was higher in Umeå 2022 (0.5360) than in Helgegården 2022 (0.3610) (Table S1), though the heritability and intraclass correlations were smaller in the former. The S_1 offspring, on average, had a higher prediction ability for tuber yield per plant than the F_1 hybrids at both testing sites (0.5651 vs. 0.5069 in Umeå 2022, and 0.4034 vs. 0.3187 in Helgegården 2022, respectively). Queen Anne S_1 had the highest prediction ability for tuber yield per plant in both testing sites, irrespective of which training population was used in the model, while among the two F_1 hybrids, Queen Anne × Colleen had, on average, the highest prediction ability in both sites (0.5770 in Umeå 2022 and 0.5256 in Helgegården 2022).

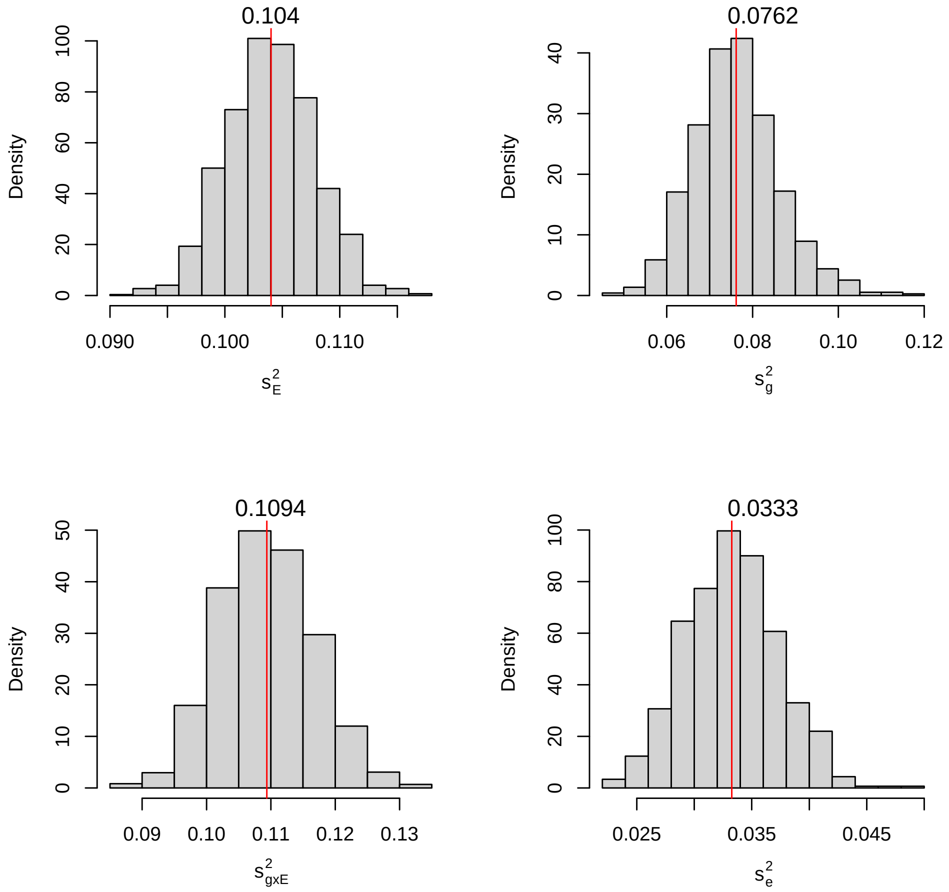

Figure 1. Histograms of posterior distributions of variance parameters in reaction norm model. The variance among environments, genotypes, genotype-by-environment interaction and residual are $\hat{\sigma}^2_E$, $\hat{\sigma}^2_g$, $\hat{\sigma}^2_{g\times E}$ and $\hat{\sigma}^2_e$, respectively.

Table 5. Genomic prediction ability (r) for tuber yield (kg plant^{-1}) at Helgegården 2022 (H2022).

Training Population	**Breeding T_2 Population**	**r**
U2021 C S_1 + U2021 QA × M + H2022 C S_1	H2022 QA × C	0.6183
U2021 M S_1 + U2021 QA × M + H2022 M S_1	H2022 QA × M	0.0994
U 2021 QA S_1 + U2021 QA × C + H2022 QA S_1	H2022 QA × C	0.5345
U2021 QA S_1 + U2021 QA × M+ H2022 QA S_1	H2022 QA × M	0.1089
U2021 QA × C + U2021 QA × M + H2022 QA × C	H2022 Q × M	0.1229
U2021 QA × C + U2021 QA × M + H2022 QA × M	H2022 QA × C	0.4650
U2021 R S_1 + U2021 Queen Anne × C + H2022 R S_1	H 2022 QA × C	0.5137
U2021 RS_1 + U2021 QA × M + H2022 R S_1	H2022 QA × M	0.1139
U2021 C S_1 + U2021 R S_1 + H2022 R S_1	H2022 C S_1	0.2636
U2021 M S_1 + Umeå 2021 R S_1 + H2022 R S_1	H2022 M S_1	0.4154
U2021 QAS_1 + U2021 R S_1 + H2022 R S_1	H2022 QA S_1	0.6570
U2021 C S_1 + U2021 M S_1 + H2022 C S_1	H2022 M S_1	0.1891
U2021 QA S_1 + U2021 C S_1 + H2022 C S_1	H2022 QA S_1	0.6342
U2021 C S_1 + U2021 M S_1 + H 2022 S_1	H2022 C S_1	0.1359
U2021 QA S_1 + U2021 M S_1 + H2022 M S_1	H2022 QA S_1	0.6656
U2021 QA S_1 + U2021 C S_1 + H2022 QA S_1	H2022 C S_1	0.2635
U2021 QA S_1 + U2021 M S_1 + H2022 QA S_1	H2022 M S_1	0.4059
U2021 M S_1 + U2021 QA × C + H2022 M S_1	H2022 QA × C	0.4965
U2021 C S_1 + U2021 QA × M + H2022 C S_1	H2022 QA × M	0.1137

U2021 = Umeå 2021; C = Colleen; M = × Melody; QA = Queen Anne; R = Rudolph; S_1 = selfing.

Table 6. Genomic prediction ability (r) for tuber yield (kg plant^{-1}) at Umeå 2022 (U2022).

Training Population	**Breeding T_2 Population**	**r**
U2021 C S_1 + U2021 QA × C + U2022 C S_1	U2022 QA × C	0.5870
U2021 M S_1 + U2021 QA × M + U2022 M S_1	U2022 QA × M	0.4363
U2021 QA S_1 + U2021 QA × C + U022 QA S_1	U2022 QA × C	0.5720
U2021 QA S_1 + U2021 QA × M + U2022 QA S_1	U2022 QA × M	0.4363
U2021 QA × C + U2021 QA × M + U2022 QA × C	U2022 QA × M	0.4358
U2021 QA × C + U2021 QA× M + U2022 QA × M	U2022 QA × C	0.5863
U2021 R S_1 + U2021 QA × C + U2022 R S_1	U2022 QA × C	0.5700
U2021 R S_1+ U2021 QA × M + U2022 R S_1	U2022 QA × M	0.4406
U2021 C S_1 + U2021 R S_1 + U2022 R S_1	U2022 C S_1	0.4944
U2021 M S_1 + U2021 R S_1 + U2022 R S_1	U2022 M S_1	0.4231
U2021 Q S_1 + U2021 R S_1 + U2022 R S_1	U2022 Q S_1	0.7815
U2021 C S_1 + U2021 M S_1 + U2022 C S_1	U2022 M S_1	0.4324
U2021 QA S_1 + U2021 C S_1 + U2022 C S_1	U2022 QA S_1	0.7528
U2021 C S_1 + U2021 M S_1 + U2022 M S_1	U2022 C S_1	0.4713
U2021 QA S_1 + U2021 M S_1 + U2022 M S_1	U2022 Q S_1	0.7577
U2021 QA S_1 + U2021 C S_1 + U2022 QA S_1	U2022 C S_1	0.5095
U2021 QA S_1 + U2021 M S_1+ U2022 QA S_1	U2022 M S_1	0.4629
U2021 M S_1 + U2021 QA × C + U2022 M S_1	U022 QA × C	0.5698
U2021 C S_1 + U2021 QA × M + U2022 CS_1	U2022 QA × M	0.4350

U2021 = Umeå 2021; C = Colleen; M = × Melody; QA = Queen Anne; R = Rudolph; S_1 = selfing.

4. Discussion

The results across the testing sites and over the years (Table 2) confirm that inbreeding significantly affects tuber yield, as noted before in this and other potato cultigen pools [30]. This inbreeding depression relates to the accumulation of deleterious recessive alleles due to vegetative propagation in this polyploid tuber crop with tetrasomic inheritance [41]. Polyploidization hid them and masked their harmful effects in the heterozygous tetrasomic potato. Inbreeding exposes these deleterious recessive alleles that reduce offspring fitness, i.e., survival, vigorous growth (influencing yield) and ability to reproduce [42]. Next-generation sequencing facilitates the identification of deleterious variants with major and minor harmful effects in the potato genomes [43], which may further lead to their purging through recombination and selection using large segregating populations. The significant differences in the inbreeding effect on tuber yield per plant among the four cultivars corroborate, as noted by Zhang et al. [44], that these variants are specific. The cultivars may differ in the inbreeding coefficient due to their co-ancestry, in which some breeding clones or cultivars appear more than once in the respective pedigree.

The range of the prediction ability for tuber yield per plant (0.5944–0.6776) when using the T_1 at Umeå 2021 as training population for both testing sets at Helgegården and Umeå in 2022 (Figures S1 and S2) was larger than previous estimates of prediction accuracy in early T_is in the same Nordic sites [45] or in other potato breeding populations [46–50]. Each of the above studies' results depend, of course, on their datasets, reference populations and their size, as well as the testing sites used, i.e., all of them affect the prediction accuracy of the different models. Nevertheless, this increase in r might ensue from using a full tetrasomic polyploid model including non-additive effects [24]. As indicated by Batista et al. [51], including allele dosage and dominance effects improve genomic prediction in polyploid species, particularly for those with a high frequency of different heterozygous types and high dominance degree. Likewise, the reaction norm model [31] could account for the improvement of the prediction ability for tuber yield per plant because the genetic and environmental gradients are defined as linear functions of SNPs and of environmental covariates, respectively.

The differences in the prediction ability between S_1 and their half sib F_1 offspring depend on the parents used (S_0) as well as the training population used for modeling GBLUPs (Tables 5 and 6). Hence, genomic predictions depend on the genetic background of both training and testing sets. The fact that the prediction ability, on average, was larger for inbred (S_1) than for hybrid (F_1) offspring suggests that GBLUPs are likely to be effective for purging deleterious alleles with harmful effects on tuber yield per plant, which are noticed after inbreeding. Likewise, the linkage disequilibrium between the tested SNPs and causal genes is kept more efficiently in S_1 than in F_1 offspring, thereby showing that decreasing the recombination rate may increase r. Although selfing is seldom used in potato breeding, inbreeding along with selection based on GBLUPs could both reveal deleterious recessive alleles and identify promising germplasm based on their EBVs for further crossing. In large potato breeding programs with hundreds of full sib F_1 family offspring, GBLUPs may also allow for a two-stage selection in early segregating generations, i.e., selecting the best among the F_1 full sib family offspring, and thereafter the best breeding clones within each family.

The emphasis on early generation selection, specifically in T_1, signifies the importance of identifying and harnessing heritable variations in the initial stages of potato breeding. This approach enables breeders to streamline the breeding process by focusing on traits relevant to tuber yield, contributing to developing improved cultivars more efficiently. The use of tetrasomic GBLUPs in this research introduces a genomic prediction method that considers the tetraploid nature of potatoes. This methodology holds promise in accurately predicting the tuber yield trait, as it accounts for the complexity arising from the tetraploid genome. This study explored how this method compares with traditional breeding approaches and its potential implications for enhancing breeding precision.

4.1. Inbred vs. Hybrid Offspring Performance

The observed larger average r values in inbred offspring compared to hybrid offspring raise intriguing questions about the genetic dynamics at play. This study revolves around the genetic basis of this disparity and how it might influence the crossbreeding strategy for potatoes. The results show that that some inbreds are more predictable in terms of tuber yield. However, results suggest a need for further investigation into hybrid vigor and its impact on yield prediction. Nevertheless, the inclusion of multi-environmental data in the analysis enhances the robustness of the findings of this study. Exploring how r may vary across different breeding sites provides valuable insights into the adaptability and stability of the selected breeding clones. Clearly the results of this research highlight the challenges posed by environmental variations and the need for cultivars (inbreds and hybrids) with broad adaptability.

4.2. Genotype Influence on Predictive Ability

The strong influence of genotype on the predictive ability (r) for both S_1 and F_1 offspring underscores the genetic control of tuber yield. The discussion could delve into the potential genetic markers associated with high-yielding traits and how this knowledge can be leveraged for marker-assisted selection, thus accelerating the breeding process.

In summary, the results are encouraging for using GBLUPs for T_1 selection for tuber yield per plant using a model based on a training population evaluated in four-plant plots using an augmented design. Such an approach, as envisaged by Bradshaw [52], would greatly impact potato breeding efficiency. Likewise, our findings confirm the in silico research by Wu et al. [53], who stated that GBLUPs used in successive early selection stages may increase short term genetic gains. Indeed, GBLUPs allows selecting clones with different inbreeding levels in early stages of potato breeding.

5. Conclusions

Considering the findings, future research should explore how these insights can be integrated into current potato breeding practices. Are there implications for the design of breeding programs, selection criteria, or the identification of parents? Moreover, how might this research contribute to addressing challenges such as disease resistance, climate resilience and market preferences for potato cultivars? In this regard, this research provides valuable insights into the ongoing efforts to enhance potato genetic improvement. The discussion can stimulate further exploration of the intricacies involved in early generation selection, genomic prediction methods and the genetic factors influencing tuber yield, ultimately informing more effective and sustainable potato breeding strategies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriculture14030455/s1, Figure S1 Genomic prediction using Umeå 2021 of inbred and hybrid performance as training population and Helgegården 2022 of inbred and hybrid performance as breeding population. r: prediction ability; Figure S2 Genomic prediction using Umeå 2021 of inbred and hybrid performance as training population and Umeå 2022 of inbred and hybrid performance as breeding population. r: prediction ability; Figure S3 Distribution of correlations for prediction problem described in Table 3; Figure S4 Distribution of correlations for prediction problem described in Table 4.

Author Contributions: Conceptualization, R.O., F.R. and M.R.G.-G., formal analysis, R.O., P.P.-R. and J.C.; investigation, R.O., F.R. and R.V.; resources, R.O. and M.R.G.-G.; data curation, R.O., F.R., J.C. and P.P.-R.; writing—original draft preparation, R.O.; writing—review and editing, R.O., R.V., M.R.G.-G., P.P.-R. and J.C.; visualization, P.P.-R. and J.C.; supervision, R.O.; project administration, R.O.; funding acquisition, R.O. and M.R.G.-G. All authors have read and agreed to the published version of the manuscript.

Funding: Swedish University of Agricultural Sciences (SLU) and the Swedish Research Council Formas for växtförädlingsområdet, SLU Breeding Network via the grant project “Genetic basis of inbreeding depression and heterosis”, and Formas through the grant project (2019–2022) “Genomisk prediktion i kombination med högkapacitetsfenotypning för att öka potatisens knölskörd i ett föränderligt klimat”.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: We thank Petra van Roggen and Yaseen Alnaasan (Intertek Agritech, Alnarp, Sweden) for faciltating the DNA extraction and further genotyping with Diversity Array Technology Pty Ltd. (ACT, Australia). We also acknowledge the field work for planting and managing the multi-site trials by Boel Sandström and other SLU staff in Umeå and to Hushallningssallskapet staff at both Helgegården and Mosslunda.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Eriksson, D.; Carlson-Nilsson, U.; Ortíz, R.; Andreasson, E. Overview and breeding strategies of table potato production in Sweden and the Fennoscandian Region. *Potato Res.* **2016**, *59*, 279–294. [CrossRef]
2. Ortiz, R.; Selga, C.; Reslow, F.; Carlson-Nilsson, U. Svensk potatisförädling: Breeding the new table and crisp potatoes. *Sveriges Utsädesförenings Tidskrif* **2020**, *2020*, 16–26.
3. Brown, J.; Caligari, P.D.S.; Mackay, G.R.; Swan, G.E.L. The efficiency of visual selection in early generations of a potato breeding programme. *Ann. Appl. Biol.* **1987**, *110*, 357–363. [CrossRef]
4. Gopal, J. Progeny selection for agronomic characters in early generations of potato breeding programme. *Theor. Appl. Genet.* **1997**, *95*, 307–311. [CrossRef]
5. Tarn, T.R.; Tai, G.C.C.; Jong, H.D.; Murphy, A.M.; Seabrook, J.E.A. Breeding potatoes for long-day, temperate climates. *Plant Breed. Rev.* **1992**, *9*, 219–332.
6. Brown, J.; Caligari, P.D.S. The efficiency of seedling selection for yield and yield components in a potato breeding programme. *Pflanzenzucht* **1986**, *96*, 53–62.
7. Neele, A.E.F.; Nab, H.J.; Louwes, K.K. Identification of superior parents in a potato breeding programme. *Theor. Appl. Genet.* **1991**, *82*, 264–272. [CrossRef] [PubMed]
8. Rodrigues de Figueiredo, I.C.; Pinto, C.A.B.P.; Ribeiro, G.H.M.R.; de Oliveira Lino, L.; Lyra, D.H.; Moreira, C.M. Efficiency of selection in early generations of potato families with a view toward heat tolerance. *Crop Breed. Appl. Biotechnol.* **2015**, *15*, 210–217. [CrossRef]
9. Bradshaw, J.E.; Dale, M.F.B.; Swan, G.E.L.; Todd, D.; Wilson, R.N. Early-generation selection between and within pair crosses in a potato (*Solanum tuberosum* subsp. *tuberosum*) breeding programme. *Theor. Appl. Genet.* **1998**, *97*, 1331–1339. [CrossRef]

10. Diniz, M.C.D.R.; Pinto, C.A.B.P.; Lambert, E.S. Sample size for family evaluation in potato breeding programs. *Ciência e Agrotecnologia* **2006**, *30*, 277–282. [CrossRef]
11. Simmonds, N.W. Family selection in plant breeding. *Euphytica* **1996**, *90*, 201–208. [CrossRef]
12. Simon, G.A.; Pinto, C.A.B.P.; Benites, F.R.G. Seleção de familias clonais de batata em diferentes ambientes. *Ciência e Agrotecnologia* **2009**, *33*, 164–169. [CrossRef]
13. Verissimo, M.A.A.; Pereira, A.S.; Silva, S.D.A.; Terres, L.R.; Ney, V.G.; Silva, G.O. Expressão de caracteres de tubérculos em função do tamanho de recipiente usado no cultivo de batata na geração de plântulas. *Rev. Ceres* **2012**, *59*, 787–793. [CrossRef]
14. Ticona-Benavente, C.A.; Pinto, C.A.B.P.; Rodrigues de Figueiredo, I.C.; Ribeiro, G.H.M.R. Repeatability of family means in early generations of potato under heat stress. *Crop Breed. Appl. Biotechnol.* **2011**, *11*, 330–337. [CrossRef]
15. Brown, J.; Caligari, P.D.S.; Mackay, G.R. The repeatability of progeny means in the early generations of a potato breeding programme. *Ann. Appl. Biol.* **1987**, *110*, 365–370. [CrossRef]
16. Brown, J.; Caligari, P.D.S. Cross prediction in a potato breeding programme by evaluation of parental material. *Theor. Appl. Genet.* **1989**, *77*, 246–252. [CrossRef]
17. Brown, J.; Caligari, P.D.S.; Mackay, G.R.; Swan, G.E.L. The efficiency of seedling selection by visual preference in a potato breeding programme. *J. Agric. Sci.* **1984**, *103*, 339–346. [CrossRef]
18. Swieżyński, K.M. Selection of individual tubers in potato breeding. *Theor. Appl. Genet.* **1978**, *53*, 71–80. [CrossRef]
19. Brown, J.; Caligari, P.D.S.; Dale, M.F.B.; Swan, G.E.L.; Mackay, G.R. The use of cross prediction methods in a practical potato breeding programme. *Theor. Appl. Genet.* **1988**, *76*, 33–38. [CrossRef] [PubMed]
20. Ticona-Benavente, C.A.; da Silva Filho, D.F. Comparison of BLUE and BLUP/REML in the selection of clones and families of potato (*Solanum tuberosum*). *Genet. Mol. Res.* **2015**, *14*, 18421–18430. [CrossRef] [PubMed]
21. Henderson, C.R. *Applications of Linear Models in Animal Breeding*; University of Guelph: Guelph, Canada, 1984.
22. Slater, A.T.; Wilson, G.M.; Cogan, N.O.; Forster, J.W.; Hayes, B.J. Improving the analysis of low heritability complex traits for enhanced genetic gain in potato. *Theor. Appl. Genet.* **2014**, *127*, 809–820. [CrossRef]
23. Slater, A.T.; Cogan, N.O.; Hayes, B.J.; Schultz, L.; Dale, M.F.; Bryan, G.J.; Forster, J.W. Improving breeding efficiency in potato using molecular and quantitative genetics. *Theor. Appl. Genet.* **2014**, *127*, 2279–2292. [CrossRef]
24. Slater, A.T.; Cogan, N.O.I.; Rodoni, B.C.; Hayes, B.J.; Forster, J.W. Improving the selection efficiency in potato breeding. *Acta Hort.* **2016**, *1127*, 237–242. [CrossRef]
25. Slater, A.T.; Cogan, N.O.; Forster, J.W.; Hayes, B.J.; Daetwyler, H.D. Improving genetic gain with genomic selection in autotetraploid potato. *Plant Genome* **2016**, *9*, 3. [CrossRef]
26. Ortiz, R.; Crossa, J.; Reslow, F.; Perez-Rodriguez, P.; Cuevas, J. Genome-based genotype environment prediction enhances potato (*Solanum tuberosum* L.) improvement using pseudo-diploid and polysomic tetraploid modeling. *Front. Plant Sci.* **2022**, *13*, 785196. [CrossRef]
27. Jinks, J.L.; Lawrence, M.J. *The Genetic Basis of Inbreeding Depression and Heterosis: Its Implications for Plant and Animal Breeding*; Palm Oil Research Institute of Malaysia: Kuala Lumpur, Malaysia, 1983.
28. Jinks, J.L. The genetic framework of plant breeding. *Phil. Trans. R. Soc. Lond. B* **1981**, *292*, 407–419.
29. Jinks, J.L. Biometrical genetics of heterosis. In *Heterosis*; Frankel, R., Ed.; Monograph Theoretical and Applied Genetics 6; Springer: Berlin, Germany, 1983; pp. 1–46.
30. Ortiz, R.; Reslow, F.; Vetukuri, R.; García-Gil, M.R.; Pérez-Rodríguez, P.; Crossa, J. Inbreeding effects on the performance and genomic prediction for polysomic tetraploid potato offspring grown at high Nordic latitudes. *Genes* **2024**, *14*, 1302. [CrossRef]
31. Jarquín, D.; Crossa, J.; Lacaze, X.; Cheyron, P.D.; Daucourt, J.; Lorgeou, J.; Piraux, F.; Guerreiro, L.; Pérez, P.; Calus, M.; et al. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* **2014**, *127*, 595–607. [CrossRef] [PubMed]
32. Burgueño, J.; Cadena, A.; Crossa, J.; Banziger, M.; Gilmour, A.R.; Cullis, B. *User's Guide for Spatial Analysis of Field Variety Trials Using ASREML*; Centro Internacional de Mejoramiento de Maíz y Trigo: Texcoco, Mexico, 2000.
33. Cullis, B.R.; Lill, W.J.; Fisher, J.A.; Read, B.J.; Gleeson, A.C. A new procedure for the analysis of early generation variety trials. *Appl. Stat.* **1989**, *38*, 361–375. [CrossRef]
34. Gleeson, A.C.; Cullis, B.R. Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. *Biometrics* **1987**, *43*, 277–288. [CrossRef]
35. Hill, J.; Becker, H.C.; Tigerstedt, P.M.A. *Quantitative and Ecological Aspects of Plant Breeding*; Springer: Dordrecht, The Netherlands, 1998.
36. Endelman, J.B.; Kante, M.; Lindqvist-Kreuze, H.; Kilian, A.; Shannon, L.M.; Caraza-Harter, M.V.; Vaillancourt, B.; Mailloux, K.; Hamilton, P.P.; Buell, C.R. Targeted genotyping-by-sequencing of potato and software for imputation. *bioRiv* **2024**. [CrossRef]
37. Amadeu, R.R.; Ferrão, L.F.V.; Oliveira, I.B.; Benevenuto, J.; Endelman, J.B.; Munoz, P.R. Impact of dominance effects on autotetraploid genomic prediction. *Crop Sci.* **2020**, *260*, 656–665. [CrossRef]
38. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022; Available online: https://www.R-project.org/ (accessed on 24 May 2023).
39. Perez, P.; de los Campos, G. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **2014**, *198*, 483–495. [CrossRef] [PubMed]
40. Pérez-Rodríguez, P.; Crossa, J.; Bondalapati, K.; De Meyer, G.; Pita, F.; Campos, G.d.l. A pedigree-based reaction norm model for prediction of cotton yield in multienvironment trials. *Crop Sci.* **2015**, *55*, 1143–1151. [CrossRef]

41. Lian, Q.; Tang, D.; Bai, Z.; Qi, J.; Lu, F.; Huang, S.; Zhang, C. Acquisition of deleterious mutations during potato polyploidization. *J. Integr. Plant Biol.* **2019**, *61*, 7–11. [CrossRef]
42. Charlesworth, D.; Willis, J. The genetics of inbreeding depression. *Nat. Rev. Genet.* **2009**, *10*, 783–796. [CrossRef]
43. Zhang, C.; Yang, Z.; Tang, D.; Zhu, Y.; Wang, P.; Li, D.; Zhu, G.; Xiong, X.; Shang, Y.; Li, C.; et al. Genome design of hybrid potato. *Cell* **2021**, *84*, 3873–3883. [CrossRef]
44. Zhang, C.; Wang, P.; Tang, D.; Yang, Z.; Lu, F.; Qi, J.; Tawari, N.R.; Shang, Y.; Li, C.; Huang, S. The genetic basis of inbreeding depression in potato. *Nat. Genet.* **2019**, *51*, 374–378. [CrossRef] [PubMed]
45. Selga, S.; Reslow, R.; Pérez-Rodríguez, P.; Ortiz, R. The power of genomic estimated breeding values for selection when using a finite population size in genetic improvement of tetraploid potato. *G3* **2022**, *12*, jkab362. [CrossRef] [PubMed]
46. Endelman, J.B.; Carley, C.A.S.; Bethke, P.C.; Coombs, J.J.; Clough, M.E.; da Silva, W.L.; De Jong, W.S.; Douches, D.S.; Frederick, C.M.; Haynes, K.G.; et al. Genetic variance partitioning and genome-wide prediction with allele dosage information in autotetraploid potato. *Genetics* **2018**, *209*, 77–87. [CrossRef]
47. Gemenet, D.C.; Lindqvist-Kreuze, H.; De Boeck, B.; da Silva Pereira, G.; Mollinari, M.; Zeng, Z.B.; Yencho, G.C.; Campos, H. Sequencing depth and genotype quality: Accuracy and breeding operation considerations for genomic selection applications in autopolyploid crops. *Theor. Appl. Genet.* **2020**, *133*, 3345–3363. [CrossRef]
48. Sood, S.; Lin, Z.; Caruana, B.; Slater, A.T.; Daetwyler, H.D. Making the most of all data: Combining non-genotyped and genotyped potato individuals with HBLUP. *Plant Genome* **2020**, *13*, e20056. [CrossRef] [PubMed]
49. Stich, B.; Van Inghelandt, D. Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. *Front. Plant Sci.* **2018**, *9*, 159. [CrossRef] [PubMed]
50. Wilson, S.; Zheng, C.; Maliepaard, C.; Mulder, H.A.; Visser, R.G.F.; van der Burgt, A.; van Eeuwijk, F. Understanding the effectiveness of genomic prediction in tetraploid potato. *Front. Plant Sci.* **2021**, *12*, 672417. [CrossRef] [PubMed]
51. Batista, L.G.; Mello, V.H.; Souza, A.P.; Margarido, G.R.A. Genomic prediction with allele dosage information in highly polyploid species. *Theor. Appl. Genet.* **2022**, *135*, 723–739. [CrossRef]
52. Bradshaw, J.E. Review and a nalysis of limitations in way to improve conventional potato breeding. *Potato Res.* **2017**, *60*, 171–193. [CrossRef]
53. Wu, P.Y.; Stich, B.; Renner, J.M.; Muders, K.; Prigge, V.; van Inghelandt, D. Optimal implementation of genomic selection in clone breeding programs–Exemplified in potato: I. Effect of selection strategy, implementation stage, and selection intensity on short-term genetic gain. *Plant Genome* **2023**, *16*, e20327. [CrossRef]