https://doi.org/10.17221/103/2025-CJGPB

Evaluating the discriminatory ability and informativeness of DArTseq markers in a comprehensive set of contemporary European potato varieties

Petr Sedlák[1], Vladimíra Sedláková[1], Jakub Vašek[1], Ervín Hausvater[2], Daniela Čílová[1], Martina Melounová[1], Jiří Ptáček[3], Petr Doležal[2]*

[1]Department of Genetics and Breeding, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czech Republic
[2]Department of Potato Protection, Potato Research Institute Havlíčkův Brod, Ltd., Havlíčkův Brod, Czech Republic
[3]Department of Genetic Resources, Potato Research Institute Havlíčkův Brod, Ltd., Havlíčkův Brod, Czech Republic

**Corresponding author: sedlak@af.czu.cz*

Citation: Sedlák P., Sedláková V., Vašek J., Hausvater E., Čílová D., Melounová M., Ptáček J., Doležal P. (2026): Evaluating the discriminatory ability and informativeness of DArTseq markers in a comprehensive set of contemporary European potato varieties. Czech J. Genet. Plant Breed., 62: 64–75.

Abstract: High-throughput molecular technology DArTseq generates markers for potential use in molecular breeding of crops. Using DArTseq, we analysed a comprehensive set of 333 European potato varieties reflecting the outcomes of long-term breeding history and representing a potential germplasm for future breeding of potatoes in the Central European region. The varieties were classified according to four factors: region of origin, breeder, earliness and utilisation mode, that may potentially reflect their genetic structure, and for which complete data were publicly available. The DArTseq analysis was performed by the service centre, the Diversity Array Technology (University of Canberra), which generated approximately 38 000 silicoDArT and 64 000 SNP (single nucleotide polymorphism) polymorphic markers. The discriminatory ability of the markers in relation to the factors was confirmed using neighbour-joining and principal coordinate analysis (PCoA), while the informativeness was assessed using the discriminant analysis of principal components (DAPC). The analyses identified the 50 SNPs most strongly associated with each factor, along with their highly probable chromosomal localisation. Herein presented research contributes to the evaluation of potato genetic resources by adding the novel molecular data of active germplasm and implies their future utilisation in genome wide association studies and marker assisted selection.

Keywords: diversity array technology; genetic diversity; potato genetic resources; SNP markers; tetraploid potato

Genetic diversity of crops is necessary for successful breeding (Swarup et al. 2021) and keeping sustainability of agriculture in changing environments. Potato breeding is challenging due to environmental changes highlighting higher resistance and tolerance to biotic and abiotic stress, respectively (Tiwari et al.

Supported by the Ministry of Agriculture of the Czech Republic within the framework of The National Agency for Agricultural Research, Project No. QK22010073 and Internal Grant Agency of Czech University of Life Sciences Prague (SGS 512 SV25-11-21360).

https://doi.org/10.17221/103/2025-CJGPB

2022). Commercial varieties represent a considerably adapted active reservoir of reliable genes for the future. Restricted varieties and economically unimportant lines use to be frequently preserved by breeders or in systems of germplasm conservation. In Czechia, such deposit is represented by gene banks framed in the National Programme on Conservation and Utilization of Genetic Resources Important for Food and Agriculture (NPGR, https://www.np-genetickezdroje.cz/?lang=en). Gene banks are responsible for conservation, evaluation, and distribution of germplasm, and inform about its morphological and ecological values.

Molecular breeding utilises phenotype data supported with valuable molecular data reporting causal genes or their markers. To study the genetic diversity of potato and to produce the needed molecular data, variety methods have been used. An approach detecting polymorphisms in cytoplasmic DNA indicated systematic reclassification of Solanaceae (Spooner et al. 1991). Methods Restriction fragment length polymorphism (RFLP) and amplified fragment length polymorphism (AFLP) initiated genetic mapping and detection of loci perspective for marker assisted selection (MAS) (Meksem et al. 1995). Later, the detection of single nucleotide polymorphisms (SNPs) using microarrays (Massa et al. 2018; Caraza-Harter and Endelman 2022; Pandey et al. 2023) and more recently PotatoMASH (multi-allele scanning haplotags) (Vexler et al. 2024) were extensively used for identification of genes and quantitative trait loci (QTL). Subsequently, next generation sequencing led to assembly potato genome (last version in 2025), which accelerated the identification of genes, QTL and the development of MAS (Godec et al. 2025). However, the direct genome resequencing is suboptimal for MAS of potato, because the genome is too large and only its minority actively contributes to individual breeding value. The diversity array technology (DArT) then represents a compromised approach, which reduces the genome complexity and increases the density of markers effective for MAS (Jaccoud et al. 2001). The DArTseq combines the reduction of genome complexity with selection of target fragments and their sequencing (Edet et al. 2018), which enables to detect two kinds of genetic markers at once: dominant silicoDArTs, represented by the presence or absence of target genomic fragments, and SNPs in these fragments. The potential of DArT to assess the genetic diversity and to identify markers associated with traits was demonstrated in the following studies. The method was designed and verified for *Oryza sativa* L. (Jaccoud et al. 2001) and subsequently used in other crops: *Hordeum vulgare* L. (Ovesná et al. 2013), *Secale cereale* L. (Milczarski et al. 2011), *Brassica napus* L. (Raman et al. 2012), × *Triticosecale* Wittm. (Badea et al. 2011), *Daucus carota* L. (Grzebelus et al. 2014), and *Fragaria* × *ananassa* Rozier (Sánchez-Sevilla et al. 2015). In potato, the DArT was used to detect resistance genes to late blight (Śliwka et al. 2012), to create a structural map of *Solanum bulbocastanum* Dun. (Iorizzo et al. 2014), to evaluate a composition of somatic hybrids (Smyda-Dajmund et al. 2016) and to evaluate diversity in small collections of tetraploid potato genetic resources (Rungis et al. 2017). The method and its current applications are comprehensively reviewed by Dereis and Freisia (2023). Larger DArTseq-based studies on diversity in larger sets of contemporary tetraploid potato varieties are, however, missing.

The objectives of this research were to assemble a collection of European potato varieties relevant to the environmental conditions of Czechia, analyse them using the DArTseq method, and assess the reliability of the obtained markers for future implementation in potato breeding programs. We hypothesised that differences in a region of variety origin, effect of the breeder, earliness and utilisation mode are reflected in diversity of DArTseq markers.

MATERIAL AND METHODS

Plant material. From 2010 to 2023, tissue samples of 333 advanced potato cultivars were obtained from long-term comparative experiments of Department of Potato Protection (Potato Research Institute in Havlíčkův Brod Ltd., PRI) in Valečov (49.6426781°N, 15.4958431°E) and from collections of *in vitro* gene bank PRI in the frame of the NPGR. The sample set represented various European regions, breeders, classes of earliness, utilisation modes, profiles of resistance against pathogens and diseases, and lineages. An individual dataset including the mentioned characteristics was prepared for each variety based on searching the publicly available data from the GrinCzech database (https://grinczech.carc.cz/gringlobal/search.aspx), the Europotato database (https://www.europotato.org), Potato pedigree database (https://www.plantbreeding.wur.nl/PotatoPedigree/index.html), documents of breeders and administrations in the field of the registration of cultivars. Anonymised concept of the dataset concerning to the evaluated characteristics is presented in Table S1 in Electronic Supplementary Material (ESM). Figure 1 demonstrates the composition of the collection.

https://doi.org/10.17221/103/2025-CJGPB

DNA extractions and DArT analysis. The youngest and healthy leaves were collected from field experimental plots in July, prior the flowering to avoid contamination by pollen, and immediately processed in the laboratory. DNA was extracted from 200 mg of fresh leaf tissue or an equivalent amount of *in vitro* tissue using the cetyltrimethylammonium bromide (CTAB) method (Doyle & Doyle 1990). Purified pellets were eluted in 50 µL of AE buffer (Qiagen, Germany), quantified using a nanospectrophotometer (Implen, Germany). DNA integrity was evaluated by electrophoresis in a 0.8% agarose gel. DNA samples were sent to Diversity Array Technology (University of Canberra, Australia) for analysis using the Potato DArTseq 1.0 service with 2.5 million reads. The resulting silicoDArT and SNP markers were delivered as a single-row *.csv format and mapped to four reference genomes to ensure the reliable mapping: *Solanum phureja* Juz. & Bukasov DM v6.1 (Pham et al. 2020), Potato v9 (DArT.com), Solynthus ASM1507626v1 (van Lieshout et al. 2020), and *S. chacoense* Bitter M6 v41 (Leisner et al. 2018).

DArT data processing and analysis. SilicoDArT and SNP data were analysed using MS Excel (Figure 1) and R Studio 4.5.1 (R Core Team 2025) by means package dartR (Gruber et al. 2018; Mijangos et al. 2022). Prior the analysis, all monomorphic loci were removed. SNP data were filtered by call rate (threshold = 0.93), followed by removal of loci with minor allele frequency < 0.01 and secondary loci. Individuals with an individual call rate < 85% were then excluded. Genetic diversity was assessed in relation to region of origin, breeder,

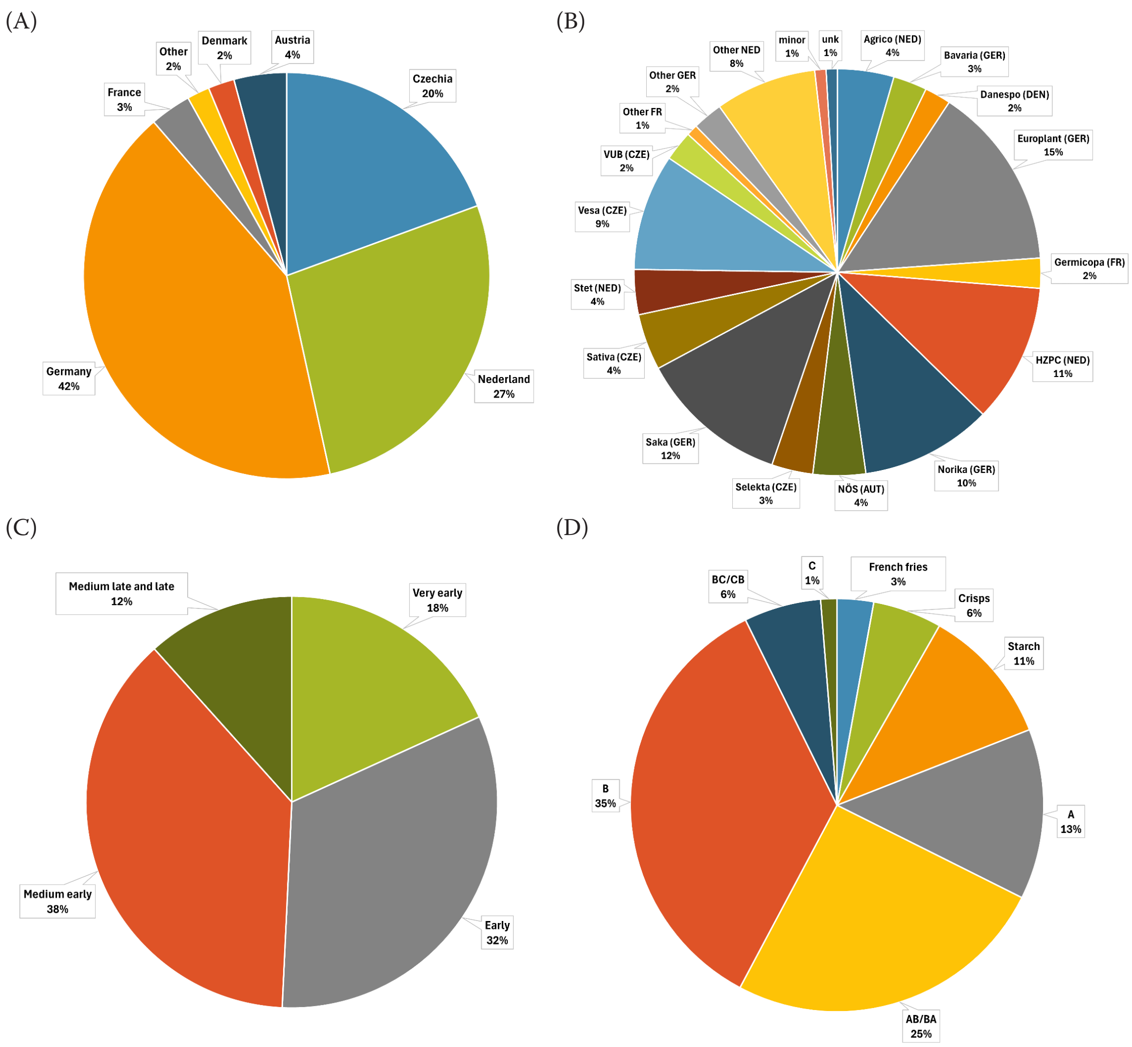

Figure 1. Structure of the collection of varieties; classification factors: region (A), breeder (B), earliness (C), utilisation mode (D)

https://doi.org/10.17221/103/2025-CJGPB

earliness, and mode of utilisation, using the analysis of molecular variance (AMOVA) with 1 000 bootstraps, neighbour-joining based on Nei's similarity coefficient and principal coordinate analysis (PCoA) based on Euclidean distances. The discriminant analysis of principal components (DAPC) was used for searching the SNPs explaining a genetic structure. The quality of the most strongly associated SNPs was evaluated based on standard population characteristics, as summarised in Table S4 in ESM. Figures 3, 4, and 6 were produced using the packages ggplot2 (Wickham 2016), ggsci (Xiao 2025), ggtree (Yu et al. 2017) and patchwork (Pedersen 2025).

RESULTS

Germplasm collection. The sample set consisted of 333 varieties produced by 42 breeders in nine countries (regions), mostly in the last 30 years; however, up to ten varieties represented older germplasm no longer actively used (i.e. Hindenburg, Rita, Clivia, Apta, Orlík, etc.). Relevant descriptive data were available for almost all varieties except for three genotypes. The varieties mostly originated from the Netherlands, Germany and Czechia. The genotypes were sorted to four earliness groups: very early, early, middle-early; the fourth group combined the middle-late and late varieties due to their small subgroup size. The utilisation mode comprised ten classes of table quality and suitability for processing (A, B, C, AB, BA, BC, CB, French fries, crisps and starch potato). The detailed structure of the collection concerning the mentioned factors is demonstrated in Figure 1, where the varieties with missing descriptors were classified as unknown (unk). If the number of varieties produced by a regional breeder was < 5, such varieties were classified other (Figure 1B).

Evaluation of genetic diversity. DArTseq identified 58 031 DNA fragments (silicoDArT markers) with more than 84 921 SNPs. The results of SNPs mapping to four reference genomes are presented in Table 1, while the best results were obtained using DM v6.1 with 16 483 thousand unmapped markers. Totally 58% of the SNP markers were mapped on the same chromosome in all reference genomes. The highest correlation between mappings was observed in the comparison of DM v6.1 and V9 ($r = 0.99$; $P < 0.001$). All these results are presented in Figure 2. The filtering of datasets on monomorphic markers removed 19 736 silicoDArTs and 20 694 SNPs. The final datasets comprised 38 295 polymorphic silicoDArTs with 2.2% of missing data and 4 243 primary SNPs with 2.42% of missing data.

The filtration on the individual call rate of SNPs retained 304 varieties. Higher, however insignificantly, allelic richness of SNPs and rate of loci deviated from Hardy-Weinberg equilibrium (unpresented result) was observed rather in the more numerous clusters of varieties originated from Germany, Nederland and Czechia. These observations were objectively confirmed by one-way AMOVA, which proved low, however, significant genetic differentiation ($\phi < 5\%$, $P < 0.001$) between groups (Table 2). The coefficients of variation (*CV*) ranged from 16 to 78% depending on the factor, while the highest *CV* was observed for earliness (> 70%) and the lowest for breeder (< 20%),

Table 1. Distribution of 84 921 single nucleotide polymorphism (SNP) markers mapped to four potato reference genomes

Chromosome	DM v6.1	M6 v41	Solynthus	Potato v9
1	8 854	7 509	6 588	8 624
2	6 836	5 861	6 045	6 770
3	6 586	5 250	6 401	6 062
4	6 002	4 102	5 931	5 761
5	4 609	3 982	4 699	4 479
6	5 467	4 524	5 934	5 354
7	5 135	4 150	4 533	4 867
8	5 199	2 992	7 606	4 832
9	5 337	4 021	5 417	4 985
10	4 142	2 939	3 901	3 904
11	5 269	4 457	5 529	4 931
12	4 848	4 465	4 928	4 885
Unmapped	16 637	30 669	17 409	17 657

https://doi.org/10.17221/103/2025-CJGPB

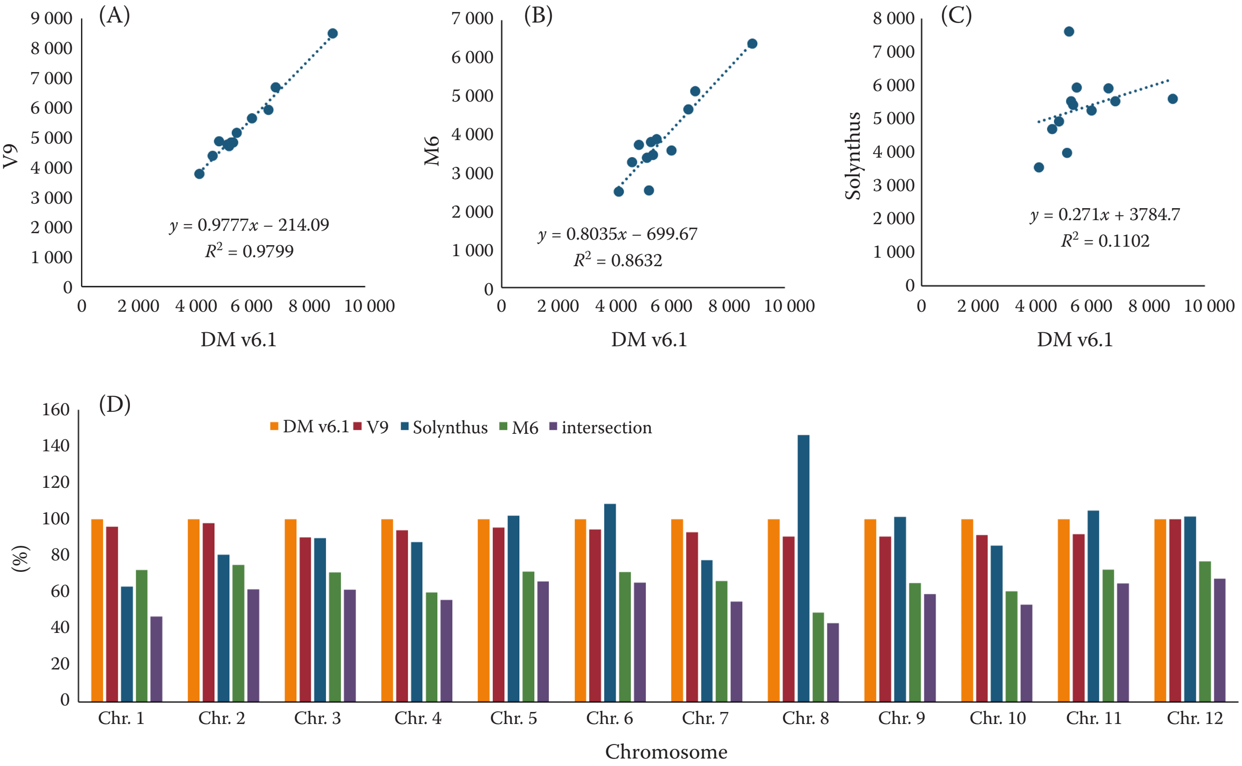

Figure 2. Analysis of single nucleotide polymorphisms (SNPs) mapping on chromosomes of reference genomes; panels (A), (B) and (C) show the correlations between mapping of DM v6.1 to other reference genomes; panel (D) presents the percentage of SNPs that are identically mapped to chromosomes relative to DM v6.1

which was associating with the number of groups within the factor. The differences in *CV*s between silicoDArTs and SNPs, tested using paired t-test, were found insignificant ($P = 0.056$). The results of AMOVA for SNPs were confirmed using neighbour-joining and PCoA (Figures 3 and 4, respectively).

The neighbour-joining based on SNPs suggested clear linkages of groups for all factors (Figure 3). Two dominant clusters of germplasm were indicated for breeder and region, while the first cluster included varieties from Czechia and Germany, and the second one from Nederland and Austria. Varieties from

Table 2. Results of analysis of molecular variance (AMOVA)

Method/factor	df	SSD	MSD	σ²	*P*	ϕ	Coefficient of variance
SilicoDArTs							
Region	6	0.064	0.011	0.0001	< 0.001	0.016	39.5
Breeder	17	0.176	0.010	0.0002	< 0.001	0.034	18.0
Earliness	3	0.036	0.012	0.0001	< 0.001	0.010	78.8
Utilisation mode	7	0.120	0.013	0.0002	< 0.001	0.035	30.5
SNPs							
Region	6	0.139	0.023	0.0003	< 0.001	0.026	35.9
Breeder	17	0.316	0.019	0.0004	< 0.001	0.028	16.4
Earliness	3	0.068	0.023	0.0001	< 0.001	0.011	71.2
Utilisation mode	7	0.247	0.027	0.0005	< 0.001	0.041	27.7

Degrees of freedom (df) corresponded to the number of groups under the factors presented in Figure 3; SSD – sum of squared deviancies; MSD – mean of square differences; σ² – estimated variance component; *P* – *P*-value; ϕ – phi value

https://doi.org/10.17221/103/2025-CJGPB

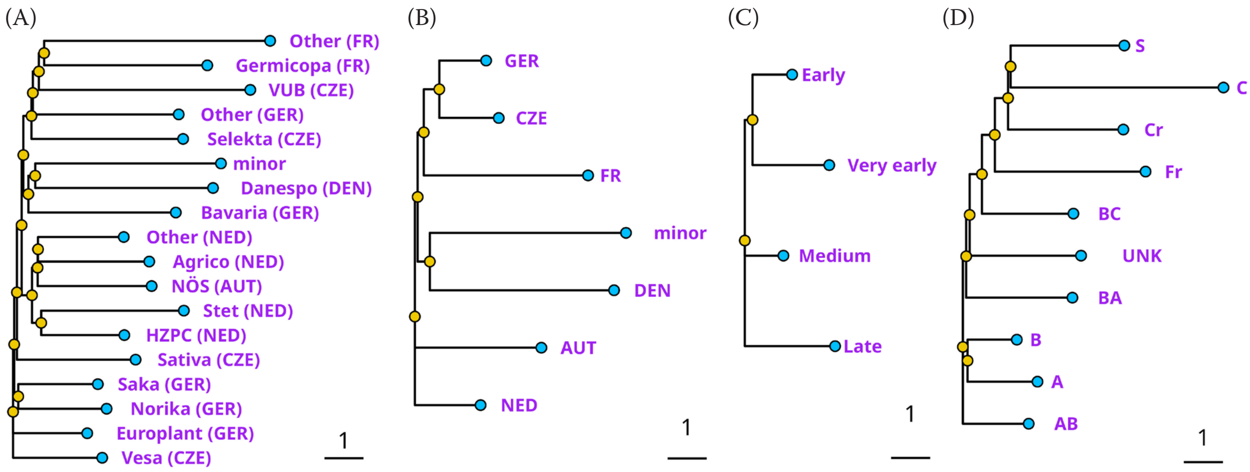

Figure 3. Neighbour-joining based on Nei's coefficients similarity of single nucleotide polymorphisms (SNPs) for the factors: breeder (A), region (B), earliness (C), utilisation mode (D)

Abbreviated mode of utilisation: Cr – crisps, Fr – French fries, S – starch, A–C – the table quality, UNK – unknown utilisation mode; regions: AUT – Austria, CZE – Czechia, DEN – Denmark, FR – France, GER – Germany, NED – Nederlands

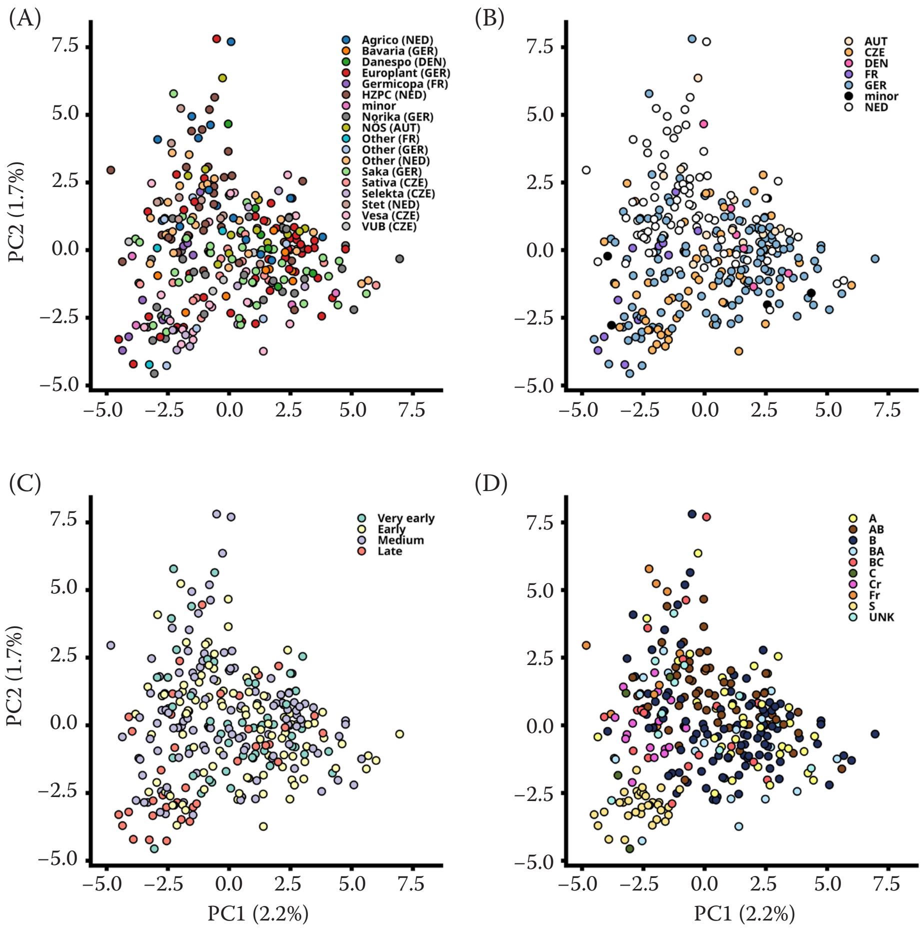

Figure 4. Results of principal coordinate analysis (PCoA) for factors: breeder (A), region (B), earliness (C), utilisation mode (D)

Abbreviated utilisation mode: Cr – crisps, Fr – French fries, S – starch, A–C – the table quality, UNK – unknown utilisation mode; regions: AUT – Austria, CZE – Czechia, DEN – Denmark, FR – France, GER – Germany, NED –Nederlands

https://doi.org/10.17221/103/2025-CJGPB

France and Denmark were placed to other relatively independent clusters. The varieties from Denmark were grouped together the artificial group (minor) shared together with unique genotypes from USA, Hungary, Poland, and genotypes of unknown origin. The analogous structure was observed for breeder. The earliness resulted in three major clusters, while the very early and early genotypes were clustered together far away from the medium and late varieties. The utilisation mode formed a sequentially descending structure from high table quality (A, AB, B) through the varieties intended for processing on crisps and French fries to the high-quality starch varieties. The varieties of poor table quality (BC and C) were grouped together those for industrial processing.

The PCoA for SNP diversity indicated several tens of components having relatively minor effects on diversity across the germplasm. However, based on SNPs, the PCoA (Figure 4) showed a structure for region, earliness and utilisation mode. No structure was found for breeder. The PCoA for regions indicated distinctness of varieties originating from Nederland, and the linkage of Czech and German germplasm. This agreed with the results of neighbour-joining. The dominant cluster of late varieties (Figure 4C) matched with a group of starch varieties (Figure 4D). Varieties for processing on crisps were clustered separately from starch varieties, and only small diversification was observed under the effect of earliness and within varieties of high table quality.

The SNP loci responsible for the clustering of varieties according to classification factors were identified using the DAPC, which detected approximately 300 principal components (PC). After selection of

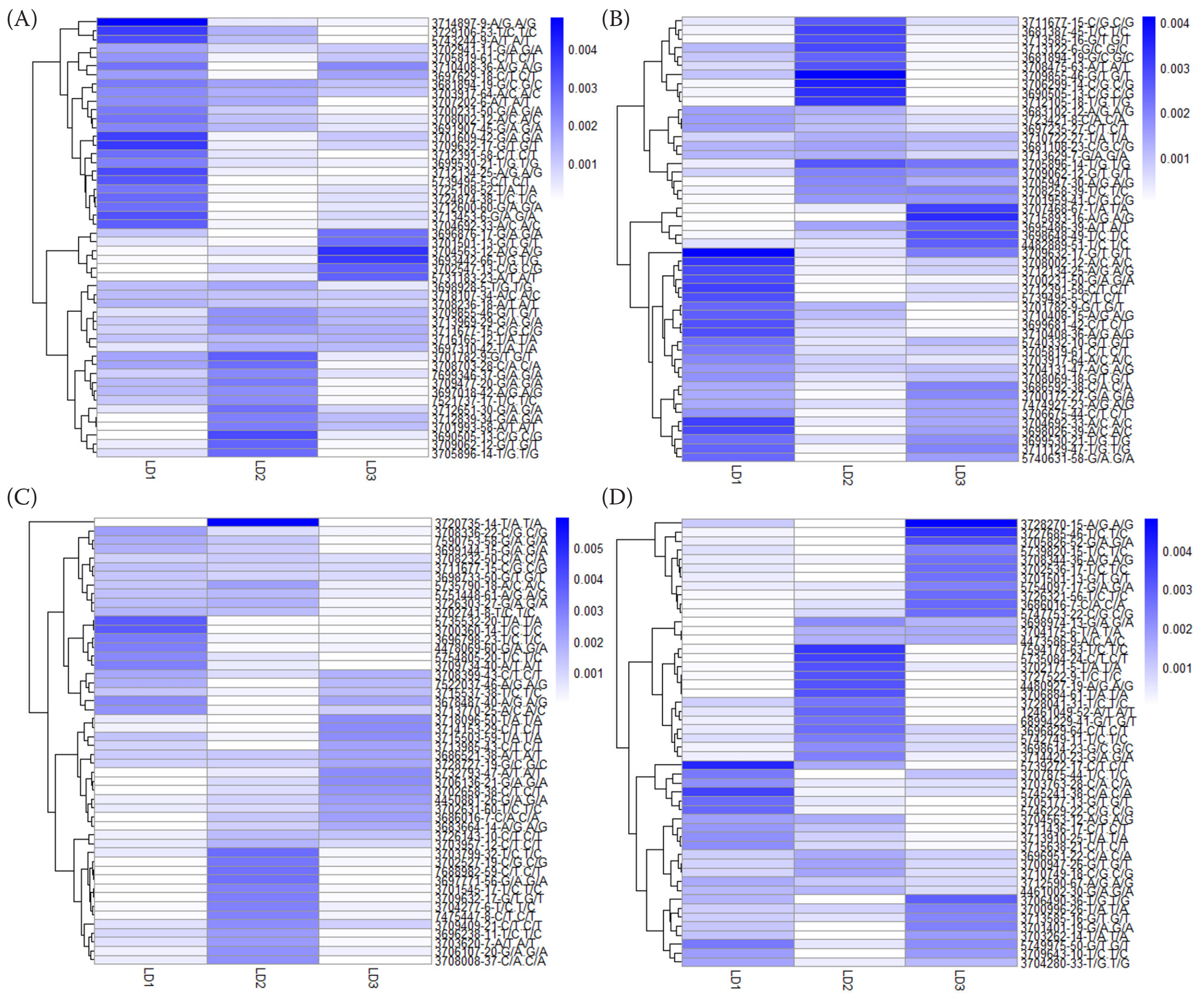

Figure 5. Single nucleotide polymorphism (SNP) markers most strongly associated with the most influential linear discriminants identified by discriminant analysis of principal component (DAPC) for breeder (A), region (B), earliness (C) and utilisation mode (D)

https://doi.org/10.17221/103/2025-CJGPB

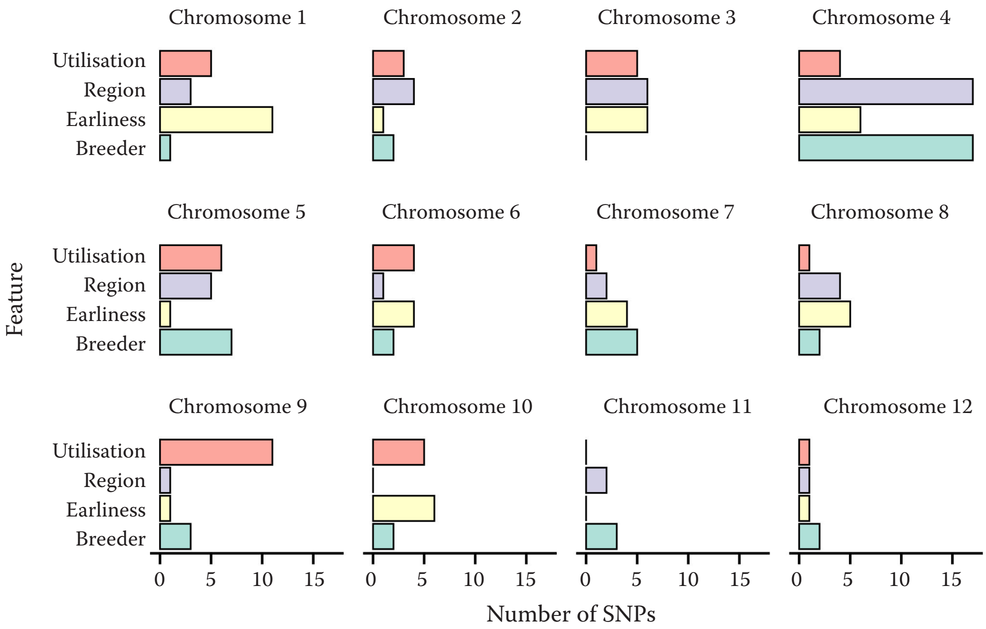

Figure 6. Distribution of the most influencing markers across genome
SNP – single nucleotide polymorphism

80 the most effective PCs, up to twelve linear discriminant functions (LD) were detected depending on the factor. The contribution of the 50 most associated SNPs, detected for three the most effective LDs, is presented in Figure 5. Of 200 markers analysed, 84% were consistently positioned by all reference genomes; 8% of markers showed incongruence caused by Solynthus, and the remaining 8% could not be mapped. The chromosomal locations of the SNPs are provided in Tables S2 and S3 in ESM and summarised in Figure 6. Based on the evaluation of predicted SNP positions, the LDs were attributed to the effects of different chromosomes.

Considerable number of SNPs (56%) associated with the regional structure were mapped to chromosomes 3 (12%), 4 (34%), and 5 (10%). The breeder effect was related to another set of SNPs located on chromosomes 4 (34%), and 5 (14%); however, eight SNPs overlapped with those influencing the regional structure. SNPs linked to the earliness were located on chromosomes 1 (22%), 3 (12%), 4 (12%), and 10 (12%), whereas those associated with the utilisation mode were positioned on chromosomes 1 (10%), 3 (10%), 5 (6%), 9 (22%), and 10 (10%). The remaining SNPs, except for four unmapped per factor, were relatively evenly distributed across the other chromosomes. The unmapped markers were found strongly associated with the first two, the most influential linear discriminant functions LD1 and LD2. This result is shown in Table S3 in ESM, in which the SNPs with the highest discriminant power and their positioning on chromosomes are listed. The results of SNPs quality evaluation are summarised in Table S4 in ESM. The call rate (CR ≈ 94%), reproducibility (RepAvg ≈ 95%) and polymorphic information content (PIC ≈ 0.30) indicate reliability and good biological informativeness of the markers. The average read depth of (AvgCountRef +AvgCountSnp > 80 per marker) confirms robust calling. Genotype frequencies demonstrate clear polymorphism and low monomorphism. Overall, these parameters support the informativeness of markers and validate the DAPC results.

DISCUSSION

The presented research conceptually belongs to the global systematic characterisation of germplasm. It is needed for the sustainable potato breeding depending on the long-term conservation, phenotyping and molecular genotyping (Sanchez et al. 2023). Molecular breeding yields form recent knowledge of the potato genome and various GWAS strategies have been utilised to identify QTLs responsible for tens of potato characteristics, as referenced more specifically below. We completed, and, using the DArTseq markers, characterised a relatively comprehensive set of recent European varieties representing the most important

https://doi.org/10.17221/103/2025-CJGPB

breeding programs and environmental adaptations for the Central-European region. A similar study on diversity was recently conducted to analyse the diversity of a broad collection of European diploid potato lines using PotatoMASH, while subsequent GWAS identified 19 new QTLs for nine traits based on 2,730 SNPs (Vexler et al. 2024). We finally analysed our sample set using almost twice-higher number of SNPs; however, to ensure better coverage of chromosomes and retaining diversity potential, we did not filter the markers on linkage disequilibrium.

The contemporary varieties represent valuable germplasm expected to serve as parental material in future tetraploid potato breeding programs. The expectation is reflecting in pedigrees related to our sample set (Table S1 in ESM), where commercial varieties participated on 220 from 252 available pedigrees; the most frequently used varieties Marabel, Laura and Carrera respectively contributed on pedigrees of 11, 9 and 7 co-analysed varieties. Although pedigree information was publicly available for only two-thirds of varieties, this still represent a substantial individual contribution of commercial varieties to overall diversity of European varieties. This may partly explain the regionally defined genetic differentiation detected by AMOVA and PCoA (Figure 4).

Discriminatory power and informativeness of DARTseq SNPs. To evaluate the discriminatory power of the DArTseq markers, we selected classification factors region, breeder, earliness, and utilisation mode, because they offered a relatively coherent and less fragmented structure, and their complete data were publicly available. The discriminatory power was confirmed using the neighbour-joining, clearly revealing interconnections within all factors (Figure 3), and PCoA indicating genetic differentiation according to region and utilisation mode. The regional effects on the diversity can be explained by differences in the evaluated germplasms, selection effects of the environment and the breeder/customer preferences on cultivars. This could be supported by strong relatedness between cultivars originated from the geographically near regions of Germany and Czechia. In contrast, the Czech varieties differed from those originated in the climatically different Netherlands. No clear structure was observed for the highly fragmented factor breeder (Figure 4A), nor for the least fragmented earliness (Figure 4C). This is partly confirmed by the results of AMOVA; significant differences under effect of breeder were identified, however the sigma and phi values are very low, which indicates weak influence of the breeder on the genetic structure. In contrast, significant effect of breeder was detected in previous study of European diploid germplasm (Vexler et al. 2024), probably due to the use of highly specific germplasm. We expected differences due to the breeder effects, as we hypothesised an influence of distinct germplasms. Markers strongly discriminating the varieties under the factor breeder were found on chromosomes 4, 5, 7 and 12. The trends in pedigrees, however, were difficult to evaluate, as some varieties occur frequently in various pedigrees provided by various breeders. Moreover, a significant number of pedigrees consist of anonymous genotypes or have not been provided by breeders in databases. The absence of a clear structure for earliness may be specifically attributed to the high number of low-efficiency QTL for the trait, combined with an excessively high number of SNPs, likely reducing the accuracy of genetic effect estimates.

The informativeness of SNP markers evaluated by DAPC are well reflected in the earliness and the utilisation mode. QTLs for these traits are known, so that we could use them as a scale. The earliness is a complex trait comprising the vine maturity, skin set, and optimal date of harvesting. The SNPs with the highest discrimination power, predicted by DAPC within our study, were mapped to chromosomes 1, 3, 5, 6, and 10. In previous studies, the QTLs associating with the earliness have been mapped to chromosome 5 (Li et al. 2018), the QTL associated with vine maturity to chromosomes 1, and 5 (Park et al. 2023), and a minor QTL for skin set to chromosome 9 (Caraza-Harter & Endelman 2022), which relatively agrees with our results. SNPs strongly associated with utilisation mode were herein confirmed on chromosomes that, according to previous research, harbour QTLs associated with tuber table and processing quality. The utilisation mode is complex characteristics involved by content and quality of starch, as well as the content of reducing sugars. The QTLs are known across potato genome. Previous study in tetraploid germplasm identified SNPs associated with 117 genes distributed equally across all chromosomes (Schönhals et al. 2017). Figure 6 and Tables S2 and S3 in ESM show that the SNPs with highest discriminatory power for the utilisation mode are predominantly located on the chromosome 9, which aligns with previous QTL mapping for cooking-type using RFLP (Kloosterman et al. 2010) and its recent confirmation using PotatoMASH (Vexler et al. 2024). Most of other

https://doi.org/10.17221/103/2025-CJGPB

herein detected SNPs associated with utilisation mode were equally distributed across chromosomes 1, 3, 4, 5, 6, and 10. This is also relatively consistent with previous research by Schäfer-Pregl et al. (1998), who mapped 18 loci involved in the starch content across all chromosomes. More recently, Werij et al. (2012) mapped QTLs for starch and crisps quality to eight chromosomes, and Śliwka et al. (2016) identified twelve additional QTLs for starch content using DArT markers on seven chromosomes. The crisp quality, directly influenced by glucose content and fry colour, associates with QTLs previously mapped to chromosomes 1, 3, 4, 5, 6, 7, 10 and 11 based on analyses of very closely specialised American germplasm (Massa et al. 2018; Pandey et al. 2023). In contrast to this, we did not detect any SNP on chromosome 11; this could be explained by genetic distances between European and American germplasm and by lower proportions of specialised varieties in our sample set.

The most effective SNPs detected by DAPC were chromosomally localised based on *in silico* mapping provided by Diversity Array Technology in Canberra. The presented results of mapping were reliable as we had at disposal four chromosomally assembled genomes, and most of the SNPs with the highest discriminatory power were mapped unambiguously on the same chromosome using at least two references. This approach resulted in sixteen unmapped SNPs per 200; however, some incongruences between reference genomes, attributed to Solynthus, were also observed. The overall successfulness of SNP mapping differed depending on used reference genome (Table 1), however best results were reached using the DM v6.1 (*S. phureja*) and M6 (*S. chacoense*), mapping 80% and 64% of SNPs respectively on any chromosome. This agreed with comparative study of coverage between references by Kyriakidou et al. (2020) who proved high homology of DM v6.1 and M6, and for 30% higher coverage of DM v6.1 with other genomes compared with M6.

CONCLUSION

The research brought considerable information background for the implementation of DArTseq markers to potato breeding programs. The herein presented results document the discriminatory power and informativeness of the acquired DArTseq markers, which suggests their immediate use for GWAS and the potential to accelerate the development and integration of new markers for MAS. This, however, accents a need for completing inclusions of missing phenotype data on resistance to biotic and abiotic stressors, and other breeding values, to the database of analysed cultivars, which requires next research.

Acknowledgements. We thank Dr. Andrzej Kilian (University of Canberra) for his valuable advice and expert assistance during the preparation and implementation of the DArT analysis.

REFERENCES

Badea A., Eudes F., Salmon D., Tuvesson S., Vrolijk A., Larsson C.-T., Caig V., Huttner E., Kilian A., Laroche A. (2011): Development and assessment of DArT markers in triticale. Theoretical and Applied Genetics, 122: 1547–1560.

Caraza-Harter M.V., Endelman J.B. (2022): The genetic architectures of vine and skin maturity in tetraploid potato. Theoretical and Applied Genetics, 135: 2943–2951.

Deres D., Feyissa T. (2023): Concepts and applications of diversity array technology (DArT) markers for crop improvement. Journal of Crop Improvement, 37: 913–933.

Doyle J.J., Doyle J.L. (1990): Isolation of plant DNA from fresh tissue. Focus, 12: 13–15.

Edet O.U., Gorafi Y.S.A., Nasuda S., Tsujimoto H. (2018): DArTseq-based analysis of genomic relationships among species of tribe *Triticeae*. Scientific Reports, 8: 16397.

Godec T., Beier S., Rodriguez-Granados N.Y., Sasidharan R., Abdelhakim L., Teige M., Usadel B., Gruden K., Petek M. (2025): Haplotype-resolved genome assembly of the tetraploid potato cultivar Désirée. Scientific Data, 12: 1044.

Gruber B., Unmack P.J., Berry O.F., Georges A. (2018): dartR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Molecular Ecology Resources, 18: 691–699.

Grzebelus D., Iorizzo M., Senalik D., Ellison S., Cavagnaro P., Macko-Podgorni A., Heller-Uszynska K., Kilian A., Nothnagel T., Allender C., Simon P.W., Baranski R. (2014): Diversity, genetic mapping, and signatures of domestication in the carrot (*Daucus carota* L.) genome, as revealed by Diversity Arrays Technology (DArT) markers. Molecular Breeding, 33: 625–637.

Iorizzo M., Gao L., Mann H., Traini A., Chiusano M.L., Kilian A., Aversano R., Carputo D., Bradeen J.M. (2014): A DArT marker-based linkage map for wild potato *Solanum bulbocastanum* facilitates structural comparisons between *Solanum* A and B genomes. BMC Genetics, 15: 123.

https://doi.org/10.17221/103/2025-CJGPB

Jaccoud D., Peng K., Feinstein D., Kilian A. (2001): Diversity arrays: A solid state technology for sequence information independent genotyping. Nucleic Acids Research, 29: E25.

Kloosterman B., Oortwijn M., de Willigen J., America T., de Vos R., Visser R.G., Bachem C.W. (2010): From QTL to candidate gene: Genetical genomics of simple and complex traits in potato using a pooling strategy. BMC Genomics, 11: 158.

Kyriakidou M., Achakkagari S.R., Gálvez López J.H., Zhu X., Tang C.Y., Tai H.H., Anglin N.L., Ellis D., Strömvik M.V. (2020): Structural genome analysis in cultivated potato taxa. Theoretical and Applied Genetics, 133: 951–966.

Leisner C.P., Hamilton J.P., Crisovan E., Manrique-Carpintero N.C., Marand A.P., Newton L., Pham G.M., Jiang J., Douches D.S., Jansky S.H., Buell C.R. (2018): Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. The Plant Journal, 94: 562–570.

Li X., Xu J., Duan S., Zhang J., Bian C., Hu J., Li G., Jin L. (2018): Mapping and QTL analysis of early-maturity traits in tetraploid potato (*Solanum tuberosum* L.). International Journal of Molecular Sciences, 19: 3065.

Massa A.N., Manrique-Carpintero N.C., Coombs J., Haynes K.G., Bethke P.C., Brandt T.L., Gupta S.K., Yencho G.C., Novy R.G., Douches D.S. (2018): Linkage analysis and QTL mapping in a tetraploid russet mapping population of potato. BMC Genetics, 19: 87.

Meksem K., Leister D., Peleman J., Zabeau M., Salamini F., Gebhardt C. (1995): A high-resolution map of the vicinity of the *R1* locus on chromosome V of potato based on RFLP and AFLP markers. Molecular Genetics and Genomics, 249: 74–81.

Mijangos J.L., Gruber B., Berry O., Pacioni C., Georges A. (2022): dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. Methods in Ecology and Evolution, 13: 2150–2158.

Milczarski P., Bolibok-Brągoszewska H., Myśków B., Stojałowski S., Heller-Uszyńska K., Góralska M., Brągoszewski P., Uszyński G., Kilian A., Rakoczy-Trojanowska M. (2011): A high density consensus map of rye (*Secale cereale* L.) based on DArT markers. PLoS ONE, 6: e28495.

Ovesná J., Kučera L., Vaculová K., Milotová J., Snape J., Wenzl P., Huttner E., Kilian A., Martelli G., Milella L. (2013): Analysis of the genetic structure of a barley collection using DNA diversity array technology (DArT). Plant Molecular Biology Report, 31: 280–288.

Pandey J., Scheuring D.C., Koym J.W., Endelman J.B., Vales M.I. (2023): Genomic selection and genome-wide association studies in tetraploid chipping potatoes. Plant Genome, 16: e20297.

Park J., Sathuvalli V., Yilma S., Whitworth J., Novy R.G. (2023). Identification of QTL associated with plant vine characteristics and infection response to late blight, early blight, and *Verticillium* wilt in a tetraploid potato population derived from late blight-resistant Palisade Russet. Frontiers in Plant Science, 14: 1222596.

Pedersen T. (2025): Patchwork: The composer of plots. R package version 1.3.2.9000. Available at https://patchwork.data-imaginist.com.

Pham G.M., Hamilton J.P., Wood J.C., Burke J.T., Zhao H., Vaillancourt B., Ou S., Jiang J., Buell C.R. (2020): Construction of a chromosome-scale long-read reference genome assembly for potato. GigaScience 9: giaa100.

R Core Team (2025). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available from: https://www.R-project.org

Raman H., Raman R., Nelson M.N., Aslam M.N., Rajasekaran R., Wratten N., Cowling W.A., Kilian A., Sharpe A.G., Schondelmaier J. (2012): Diversity array technology markers: genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). DNA Research, 19: 51–65.

Rungis D.E., Voronova A., Kokina A., Veinberga I., Skrabule I., Rostoks N. (2017): Assessment of genetic diversity and relatedness in the Latvian potato genetic resources collection by DArT genotyping. Plant Genetic Resources, 15: 72–78.

Sanchez D., Sadoun S.B., Mary-Huard T., Allier A., Moreau L., Charcosset A. (2023): Improving the use of plant genetic resources to sustain breeding programs' efficiency. Proceedings of the National Academy of Sciences, 120: e2205780119.

Sánchez-Sevilla J.F., Horvath A., Botella M.A., Gaston A., Folta K., Kilian A., Denoyes B., Amaya I. (2015): Diversity arrays technology (DArT) marker platforms for diversity analysis and linkage mapping in a complex crop, the octoploid cultivated strawberry (*Fragaria* × *ananassa*). PLoS ONE, 10: e0144960.

Schäfer-Pregl R., Ritter E., Concilio L., Hesselbach J., Lovatti L., Walkemeier B., Thelen H., Salamini F., Gebhardt C. (1998): Analysis of quantitative trait loci (QTLs) and quantitative trait alleles (QTAs) for potato tuber yield and starch content. Theoretical and Applied Genetics, 97: 834–846.

Schönhals E.M., Ding J., Ritter E., Paulo M.J., Cara N., Tacke E., Hofferbert H.-R., Lübeck J., Strahwald J., Gebhardt C. (2017): Physical mapping of QTL for tuber yield, starch content and starch yield in tetraploid potato (*Solanum tuberosum* L.) by means of genome wide genotyping

https://doi.org/10.17221/103/2025-CJGPB

by sequencing and the 8.3 K SolCAP SNP array. BMC Genomics, 18: 642.

Śliwka J., Jakuczun H., Chmielarz M., Hara-Skrzypiec A., Tomczyńska I., Kilian A., Zimnoch-Guzowska E. (2012): A resistance gene against potato late blight originating from *Solanum* × *michoacanum* maps to potato chromosome VII. Theoretical and Applied Genetics, 124: 397–406.

Śliwka J., Sołtys-Kalina D., Szajko K., Wasilewicz-Flis I., Strzelczyk-Żyta D., Zimnoch-Guzowska E., Jakuczun H., Marczewski W. (2016): Mapping of quantitative trait loci for tuber starch and leaf sucrose contents in diploid potato. Theoretical and Applied Genetics, 129: 131–140.

Smyda-Dajmund P., Śliwka J., Wasilewicz-Flis I., Jakuczun H., Zimnoch-Guzowska E. (2016): Genetic composition of interspecific potato somatic hybrids and autofused 4x plants evaluated by DArT and cytoplasmic DNA markers. Plant Cell Reports, 35: 1345–1358.

Spooner D.M., Sytsma K.J., Conti E. (1991): Chloroplast DNA evidence for genome differentiation in wild potatoes (*Solanum* Sect. Petota: *Solanaceae*). American Journal of Botany, 78: 1354–1366.

Swarup S., Cargill E.J., Crosby K., Flagel L., Kniskern J., Glenn K.C. (2021): Genetic diversity is indispensable for plant breeding to improve crops. Crop Science, 61: 839–852.

Tiwari J.K., Buckseth T., Zinta R., Bhatia N., Dalamu D., Naik S., Poonia A.K., Kardile H.B., Challam C., Singh R.K., Luthra S.K., Kumar V., Kumar M. (2022): Germplasm, breeding, and genomics in potato improvement of biotic and abiotic stresses tolerance. Frontiers in Plant Science, 13: 805671.

van Lieshout N., van der Burgt A., de Vries M.E., Ter Maat M., Eickholt D., Esselink D., van Kaauwen M.P.W., Kodde L.P., Visser R.G.F., Lindhout P., Finkers R. (2020): Solyntus, the new highly contiguous reference genome for potato (*Solanum tuberosum*). G3 Bethesda, 10: 3489–3495.

Vexler L., Leyva-Perez M. de la O., Konkolewska A., Clot C.R., Byrne S., Griffin D., Ruttink T., Hutten R.C.B., Engelen C., Visser R.G.F., Prigge V., Wagener S., Lairy-Joly G., Driesprong J.-D., Riis Sundmark E.H., Rookmaker A.N.O., van Eck H.J., Milbourne D. (2024): QTL discovery for agronomic and quality traits in diploid potato clones using PotatoMASH amplicon sequencing. G3 Genes/Genomes/Genetics, 14: jkae164.

Werij J.S., Furrer H., van Eck H.J., Visser R.G.F., Bachem C.W.B. (2012): A limited set of starch related genes explain several interrelated traits in potato. Euphytica, 186: 501–516.

Wickham H. (2016): ggplot2: Elegant Graphics for Data Analysis. New York, Springer-Verlag.

Xiao N. (2025): ggsci: Scientific journal and sci-fi themed color palettes for ‘ggplot2’. R package version 4.0.0. Available at https://CRAN.R-project.org/package=ggsci.

Yu G., Smith D., Zhu H., Guan Y., Lam T.T. (2017): ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution, 8: 28–36.

Received: October 21, 2025
Accepted: January 5, 2026
Published online: January 20, 2026