agronomy

MDPI

Article

Assembly and Analysis of Plastomes for 15 Potato Cultivars Grown in Russia

Dmitry I. Karetnikov [1,2,3], Elena A. Salina [1,2,4], Alex V. Kochetov [1,2,3,4] and Dmitry A. Afonnikov [1,2,3,*]

[1] Federal Research Center Institute of Cytology and Genetics SB RAS, 630090 Novosibirsk, Russia
[2] Kurchatov Center for Genome Research, Institute of Cytology and Genetics SB RAS, 630090 Novosibirsk, Russia
[3] Faculty of Natural Sciences, Novosibirsk State University, 630090 Novosibirsk, Russia
[4] Faculty of Agronomy, Novosibirsk State Agrarian University, 630039 Novosibirsk, Russia
* Correspondence: ada@bionet.nsc.ru; Tel.: +7-(383)-363-49-63

Abstract: Chloroplasts are important organelles in a plant cell, having their own DNA (cpDNA), transmitted only through the female line, and performing the function of photosynthesis. The determination of chloroplast DNA is of interest in the study of the genetic diversity and phylogeny of potatoes, and of cytoplasmic sterility, as well as for applications in biotechnology and genetic engineering. Here, we reconstructed the complete plastomes of 15 *S. tuberosum* potato cultivars grown in Russia. Our analysis allowed us to determine the composition and location of genes for these plastid DNAs. It was shown that the plastid genome contains both highly and low-variable regions. The region at position 63,001–68,000 nt has the highest variability. We determined the types of cpDNA based on in silico approaches: 10 cultivars have cpDNA of the W-type and 5 cultivars have cpDNA of the T-type. The genetic diversity of the plastid DNA for these potato cultivars was analyzed alongside the previously reconstructed plastomes of South American accessions, European/North American commercial cultivars and potato cultivars bred in the Ural region. The results show that plastid DNAs of the same type form clusters by sequence similarity, in agreement with previous studies.

Keywords: potato; *Solanum tuberosum* L.; chloroplast DNA; cytoplasmic DNA type; genetic diversity; SNP; InDels

Citation: Karetnikov, D.I.; Salina, E.A.; Kochetov, A.V.; Afonnikov, D.A. Assembly and Analysis of Plastomes for 15 Potato Cultivars Grown in Russia. *Agronomy* **2023**, *13*, 1454. https://doi.org/10.3390/agronomy13061454

Academic Editor: Chenggen Chu

Received: 30 April 2023
Revised: 21 May 2023
Accepted: 23 May 2023
Published: 24 May 2023

1. Introduction

Chloroplasts are important organelles in a plant cell, having their own DNA (cpDNA), transmitted only through the female line, and performing the function of photosynthesis. Plastid DNA is circular, is between 120 and 220 kb in size, and encodes 120–130 genes [1,2]. The chloroplast genome typically comprises a pair of inverted repeat regions (IRA and IRB) separated by small (SSC) and large (LSC) single copy regions. Chloroplast genomes are mostly conserved and contain two main groups of genes: components for the photosynthetic machinery and genes required for the genetic system of plastids [1]. The peculiarities of plastid DNA inheritance and its conservation make it a marker for studying evolutionary relationships in plants [3,4]. The protein coding regions of cpDNA evolve slowly, which reduces their general utility for evolutionary and population genetic studies, especially at lower taxonomic levels. However, noncoding regions, including introns and intergenic spacers, demonstrate a high level of variations of various types (insertions, deletions, inversions, single nucleotide polymorphisms and variations in tandem repeat regions or simple sequence repeats (SSRs)). These types of cpDNA variations are widely used to study genetic diversity and structure, evolutionary history, and hybridization in native and agricultural species [5].

Chloroplast DNA in *Solanum tuberosum* L. and its relatives consists of ~155,000 bp and contains about 140 genes encoding proteins, tRNAs and rRNAs [6,7]. The study of chloro-

plast DNA is of interest for assessing the genetic diversity and phylogeny of potatoes [4,8,9], as well as for solving problems in biotechnology and genetic engineering [10–13].

The interaction of chloroplast DNA genes with nuclear DNA genes in some cases causes cytoplasmic male sterility in potato hybrids [10,11]. Cytoplasmic sterility depends on the type of chloroplast DNA, of which there are five basic ones [12]: A, C, S, W and T. It has been shown that all the major types of plastome resulted from point mutations in the course of evolution from the most primitive W-type; the T-type is also characterized by a 241-nucleotide deletion [13]. Three subtypes, W1, W2 and W3, are also known for the W-type. Cytoplasmic male sterility depends on both the chloroplast and the mitochondria genomes. In this regard, an extended classification of potato cytoplasmic DNA based on chloroplast and mitochondrial DNA types has been proposed [14–16]: A, M, P, W (W/α, W/β, W/γ), T (T/β) and D.

The determination of chloroplast DNA type is of interest not only in the study of cytoplasmic sterility, but also in relation to a number of agronomic traits. For example, a higher starch content was found in hybrids and breeding clones with the W/γ and W/α cytoplasm compared to potatoes with other cytoplasm types [16]. A study of the cytoplasmic DNA of more than one thousand European potato cultivars showed that the samples with the W/γ cytoplasm had a higher starch content and longer ripening times than those with the T-type, and statistically significant correlations were observed between the presence of the D-type and M-type cytoplasm and quantitative resistance to late blight [17]. Analysis of cytoplasmic DNA types in Indian potato cultivars showed that genotypes with D-type on average had higher levels of foliar resistance to late blight than genotypes with T-type cpDNA. They were also found to be related to a high total starch content and better processing attributes [18].

The experimental determination of cpDNA type is based on the restriction and RFLP analysis of SSR, STS and CAPS markers [15,16,19] and is widely used in potato research [14,16–18,20,21]. With the advent of high-throughput DNA sequencing, it became possible to use full-genome sequences of chloroplasts for their type determination based on in silico PCR and restriction digest reaction methods [7].

In the present work, we reconstructed the full plastome sequences of 15 potato cultivars grown in Russia using short-paired reads we obtained earlier [22]. We annotated the chloroplast genomes, identified single-nucleotide polymorphisms and InDels, and assessed the variability of the cpDNA. We identified cpDNA types based on in silico methods and evaluated the genetic diversity of plastid genomes for the potato cultivars grown in Russia and for previously reconstructed plastome sequences of South American potato accessions, European/North American commercial cultivars and potato cultivars bred in the Ural region of Russia.

2. Materials and Methods

2.1. Nucleotide Sequences

We used libraries of raw reads of 15 potato cultivars grown in Russia from previous work [22]: Golubizna, Krepysh, Nevsky, Meteor, Zhukovsky, Severnoe siyanie, Krasa Meshchery, Fritella, Udacha, Krasavchik, Grand, Sudarinya, Gusar, Nikulinsky and Symphonia. The dataset includes 14 cultivars developed by different Russian breeding programs and one Dutch cultivar (Symphonia). These varieties demonstrate very broad ecological plasticity [22]. Some cultivars (Nevsky, Zhukovsky, Udacha) have been approved for 12, 10 and 9 regions of the Russian Federation, respectively (relating to the contrasting light zones). Some cultivars (Krasavchik, Severnoe Siyanie and the Dutch variety Symphonia) are approved for only one region of the Russian Federation. Since the number of narrow adaptive varieties in the study was small, we included Symphonia in the analysis as well [22].

The SRA archives were downloaded from NCBI BioProject PRJNA933976.

The plastid genome sequence of *S. tuberosum* cultivar (cv.) Desiree (NCBI identifier DQ386163.2) was used as a reference.

To evaluate the genetic diversity of cultivated potato cpDNAs we extended our dataset of plastid genomes using the sequences of 13 South American potato accessions [7], 28 cultivars bred in the Ural region of Russia [23] and 6 commercial cultivars from Europe/North America [24]. Specimen names, abbreviations, species and sequence identifiers in the NCBI database are given in Table S1, Supplementary File.

2.2. Plastome Assembly and Annotation

The raw reads preprocessing was performed using fastp v. 0.20.1 [25]. Quality control and filtering was performed using FastQC v. 0.11.9 [26]. The resulting filtered reads were assembled using GetOrganelle v. 1.7.6.1 with the default seed library [27]. The resulting assemblies were annotated by GeSeq [28] using annotation of the *S. tuberosum* cv. Desiree plastome [6]. Predicted genes were manually corrected in the annotation of each cultivar. Inverted repeats and genes at joining sites were visualized using IRscope [29]. Microsatellites were identified using MISA v. 2.1 [30] with a minimum number of repeat units set as follows: 10 for mononucleotide repeats, 5 for dinucleotide repeats, 4 for trinucleotide repeats, and 3 for tetranucleotide, pentanucleotide and hexanucleotide repeats.

2.3. SNP and InDels Identification

Multiple sequence alignment of the plastid genomes was obtained using MAFFT v. 7 [31]. SNP detection was performed using SNP-sites v. 2.5.1 [32]. SNPs coordinates were determined using plastid genome of the cv. Desiree. Nucleotide diversity, π, along the genome was estimated based on multiple alignment using DNAsp v. 6 [33] with a sliding window size of 3000 bp and step of 1000 bp. Visualization of all SNPs and nucleotide diversity was performed using Circos v. 0.69.8 [34].

To identify insertions and deletions, each assembly was aligned by MAFFT with the *S. tuberosum* cv. Desiree plastid genome. We considered InDels with respect to plastomes of Russian cultivars: deletions in cpDNA sequences of Russian cultivars implies insertions in reference plastome and vice versa.

2.4. Cytoplasmic DNA Type Determination

To determine the type of cytoplasmic DNA, we used a method described previously [7]. Seven primer sets (NTCP6, NTCP7, NTCP8, NTCP9, NTCP12, NTCP14, NTCP18) [19] were used for the in silico PCR. We estimated product size using an online PCR and PCR-RFLP tool [35]. We also performed restriction endonuclease analysis using five restriction enzymes: *Bam*HI, *Hind*III, *Kpn*I, *Pvu*I and *Xho*I. The restriction fragments sizes were compared with previously obtained maps to determine the cpDNA types [12].

2.5. Phylogenetic Tree Reconstruction

To evaluate genetic diversity between plastomes of the Russian potato cultivars [22] and potato cultivars of the Ural selection [23], as well as South American accessions [7] and European/North American cultivars [24], we performed phylogenetic tree reconstruction using the IQ-TREE with automatic model selection [36] based on cpDNA multiple sequence alignment. The inverted repeat region (IRb) was removed in the multiple alignment to avoid duplication. The tree was visualized using iTOL v. 6 [37].

3. Results

3.1. Structural Characteristics of Plastid Genomes and Their Annotations

The plastid genomes of 15 potato cultivars grown in Russia were assembled into circular DNA. Their complete sizes as well as the sizes of LSC, SSC, IRb and IRa regions are shown in Table 1. The size of the plastid genomes for the potato cultivars grown in Russia varies from 155,296 bp to 155,565 bp. The largest differences in length are observed for the LSC region (from 85,737 bp to 86,006 bp); for other regions we observed differences of 1 nucleotide.

Table 1. Sizes of the plastid genomes, location and sizes for their LSC, SSC, IRb and IRa regions for 15 Russian potato cultivars.

Cultivar	cpDNA	LSC		IRb		SSC		IRa	
	size, bp	location	size, bp	location	size, bp	location	size, bp	location	size, bp
Fritella	155,565	1–86,006	86,006	86,007–111,599	25,593	111,598–129,972	18,375	129,973–155,565	25,593
Golubizna	155,296	1–85,737	85,737	85,738–111,330	25,593	111,329–129,703	18,375	129,704–155,296	25,593
Grand	155,549	1–85,991	85,991	85,992–111,583	25,592	111,582–129,957	18,376	129,958–155,549	25,592
Gusar	155,549	1–85,991	85,991	85,992–111,583	25,592	111,582–129,957	18,376	129,958–155,549	25592
Krasa	155,562	1–86,003	85,991	86,004–111,596	25,593	111,595–129,969	18,375	129,970–155,562	25,593
Meshchery									
Krasavchik	155,296	1–85,737	85,737	85,738–111,330	25593	111329–129703	18,375	129,704–155,296	25,593
Krepysh	155,562	1–86,003	86,003	86,004–111,596	25,593	111,595–129,969	18,375	129,970–155,562	25,593
Meteor	155,549	1–85,991	85991	85,992–111,583	25,592	111,582–129,957	18,376	129,958–155,549	25,592
Nevsky	155,565	1–86,006	86,006	86,007–111,599	25,593	111,598–129,972	18,375	129,973–155,565	25,593
Nikulinsky	155,296	1–85,737	85,737	85,738–111,330	25,593	111,329–129,703	18,375	129,704–155,296	25,593
Severnoe	155,296	1–85,737	85,737	85,738–111,330	25,593	111,329–129,703	18,375	129,704–155,296	25,593
siyanie									
Sudarinya	155,549	1–85,991	85,991	85,992–111,583	25,592	111582–129957	18,376	129,958–155,549	25,592
Symphonia	155,296	1–85,737	85,737	85,738–111,330	25,593	111,329–129,703	18,375	129,704–155,296	25,593
Udacha	155,565	1–86,006	86,006	86,007–111,599	25,593	111,598–129,972	18,375	129,973–155,565	25,593
Zhukovsky	155,562	1–86,003	86,003	86,004–111,596	25,593	111595–129969	18,375	129,970–155,562	25,593

The plastid genomes of the Russian potato cultivars contained 143 genes each, which is equal to the number of genes in the reference cv. Desiree plastome. We compared the location of genes relative to the boundaries between the LSC, SSC, IRb and IRa regions for all the plastomes of the Russian potato cultivars (Figure S1, Supplementary File). Regardless of the genome lengths, this location appeared to be identical.

In silico microsatellite analysis revealed the presence of mono- to pentanucleotide microsatellite repeats in the genomes (Table S2, Supplementary File). The number of di-, tri-, tetra- and pentanucleotide microsatellites was identical for all the potato cultivars grown in Russia (6, 2, 8 and 1, respectively). The only difference between the plastomes was observed for the monomer repeats, whose size varied from 36 to 39 bp. The microsatellite repeats of either nucleotide A (14 to 15 bp) or nucleotide T (22 to 24 bp) were found in the cpDNAs of the Russian potato cultivars (Table S3, Supplementary File).

3.2. Identification of SNPs, Insertions, and Deletions and Assessment of Nucleotide Diversity

Using multiple alignment of the plastome sequences of the potato cultivars grown in Russia and the reference genome cv. Desiree, we identified sets of SNPs (Table 2, Figure 1). A total of 128 unique polymorphisms were identified, of which 67 were transitions and 61 were transversions. In two cultivars (Severnoe siyanie and Symphonia), the plastome sequences were found to be identical to the plastome of cv. Desiree. The cvs. Golubizna, Krasavchik and Nikulinsky have one SNP each in the intergenic regions in comparison with the cv. Desiree plastome. In the remaining plastomes, 121 to 114 SNPs were identified. The number of SNPs located in the coding regions varies from 52 to 55 (slightly less than a half of the total number of SNPs).

The nucleotide diversity, π, calculated for the entire cpDNA was 0.0004. The calculation of π using sliding windows shows that the variability is not uniform along the sequence of the plastid genome. There are regions with diversity close to 0. Such regions ($0 < \pi < 0.0001$) include genes *rps2*, *rpoC2* (positions 16,094–22,093), *ndhJ*, *ndhK*, *ndhC*, *trnV*, *trnM*, *atpE*, *atpB* (positions 50,115–56,355), *rbcL*, *accD*, *psal* (positions 57,356–62,363), *rps12*, *clpP*, *psbB*, *psbT*, *psbN*, *psbH*, *petB* (positions 71,433–78,442), *ycf2* (positions 86,444–94,443), *ndhB*, *rps7*, *rps12*, *trnV*, *16S*, *trnI*, *trnA*, *23S*, *4.5S*, *5S*, *trnR*, *trnN* (positions 96,444–111,444), *ycf1*, *trnN*, *trnR*, *5S*, *4.5S*, *23S*, *trnA*, *trnI*, *16S*, *trnV*, *rps12*, *rps7*, *ndhB* (positions 128,463–145,463), *ycf2*, *trnI*, *rpl23*, *rpl2* and *rps19* (positions 146,464–155,656). At the same time, there are regions characterized by a high value of diversity, π. These are the region of genes *rpl32*, *sprA*, *trnL*, *ccsA*, *ndhD* and *psaC* (positions 113,445–119,462, $0.00079 < \pi < 0.00174$) and the region of genes *cemA*, *petA*, *psbJ*, *psbL*, *psbE* and *psbF* of the LSC site at positions 62,364–67,431, for which the π value peaks ($0.00287 < \pi < 0.00304$; Figure 1).

Table 2. Statistics of the SNPs identified in plastid genomes of the potato cultivars grown in Russia.

Cultivar	Number of SNP	Number of SNP in Genes	Genes with SNPs
Fritella	121	55	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, rps8, ycf1, ycf2, ycf3*
Golubizna	1	1	-
Grand	114	52	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl14, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*
Gusar	114	52	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl14, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*
Krasa Meshchery	121	54	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhD, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*
Krasavchik	1	1	-
Krepysh	121	54	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhD, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*
Meteor	114	52	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl14, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*
Nevsky	121	55	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, rps8, ycf1, ycf2, ycf3*
Nikulinsky	1	1	-
Severnoe siyanie	0	0	-
Sudarinya	114	52	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl14, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*
Symphonia	0	0	-
Udacha	121	55	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, rps8, ycf1, ycf2, ycf3*
Zhukovsky	121	54	*23S, atpA, ccsA, infA, matK, ndhA, ndhB, ndhD, ndhF, ndhH, ndhK, petD, psaA, psaB, psbA, rbcL, rpl16, rpl20, rpl22, rpoB, rpoC1, rpoC2, rps11, rps14, rps16, rps3, ycf1, ycf2, ycf3*

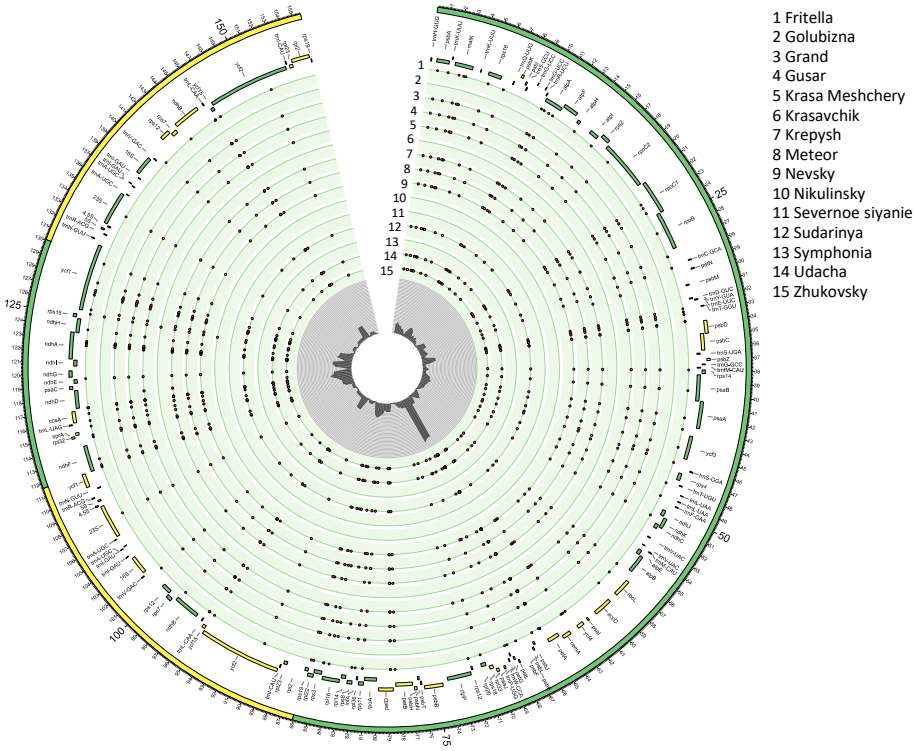

Figure 1. Circos plot showing SNP distribution in the plastid genomes of the 15 cultivars grown in Russia (light green tracks) and the cv. Desiree reference genome (white track with gene annotation).

On the outermost track, inverted repeats are shown in light green and single-copy regions are shown in green. The green bars in the annotation panel show genes transcribed clockwise; the yellow bars show genes transcribed counter-clockwise. The gray diagram in the center shows the distribution of nucleotide diversity π along the genome. The correspondence of the track number and the cultivar is shown in the top right corner.

We also assessed nucleotide diversity along the genomic sequence of the reference genome of cv. Desiree based on a comparison of 49 genomes of *S. tuberosum* accessions (15 potato cultivars from the current work, 1 South American accession TBR, 26 cultivars bred in the Ural region, 6 commercial cultivars from Europe/North America and cv. Desiree). The results (Figure S2, Supplementary File) demonstrated that the most variable region was the LSC segment at positions 63,001–68,000.

The results of the InDel identification and analysis are presented in Table 3. There are no InDels in the plastid genomes of cvs. Krasavchik, Golubizna, Krasa Meshchery, Severnoe siyanie and Symphonia, indicating high similarity with the plastid genome of the *S. tuberosum* cv. Desiree. This is consistent with the genomic size data from Table 1 and with the absence or low number of single-nucleotide substitutions for these cultivars (Table 2).

Table 3. InDel statistics for plastid genomes of the potato cultivars grown in Russia.

Cultivar	Insertions			Deletions			Genes with Insertions	Genes with Deletions
	Total Number	>1 bp	Maximal Size	Total Number	>1 bp	Maximal Size		
Fritella	19	12	241	13	9	18	-	*petB*
Golubizna	0	0	0	0	0	0	-	-
Grand	17	9	241	11	6	18	-	*petB*
Gusar	17	9	241	11	6	18	-	*petB*
Krasa Meshchery	20	12	241	13	9	18	-	*petB*
Krasavchik	0	0	0	0	0	0	-	-
Krepysh	20	12	241	13	9	18	-	*petB*
Meteor	17	9	241	11	6	18	-	*petB*
Nevsky	19	12	241	13	9	18	-	*petB*
Nikulinsky	0	0	0	0	0	0	-	-
Severnoe siyanie	0	0	0	0	0	0	-	-
Sudarinya	17	9	241	11	6	18	-	*petB*
Symphonia	0	0	0	0	0	0	-	-
Udacha	19	12	241	13	9	18	-	*petB*
Zhukovsky	20	12	241	13	9	18	-	*petB*

The genomes of the remaining cultivars have 17 to 20 insertions relative to the reference cpDNA, among which 9 to 12 insertions are larger than 1 nucleotide, and the maximum insertion size is 241 nucleotides. The number of deletions varies from 11 to 13, with 6 to 9 deletions being larger than 1 nucleotide. The maximum length of the deletions is 18 nucleotides for all varieties in which deletions are found. Interestingly, the number of deletions is less than that of insertions for all the genomes in which both types of mutations are present (Table 3).

Note that all the insertions we detected are located in the intergenic regions. Deletions are also predominantly located in the intergenic regions. Only one deletion 8 bp long is located in the second exon of the *petB* gene (position in the reference genome 77,435–78,076) (Figure S3, Supplementary File). This deletion affects the stop codon (TAG) of the gene. However, genomes with such a deletion do not have abnormalities in the open reading frame of the *petB* gene because the TAG codon is also located immediately after the deletion.

3.3. Identification of Chloroplast DNA Types

We determined the type of chloroplast DNA for potato varieties based on several criteria. First of all, using an in silico PCR reaction and several primers [19], we determined

the product sizes for each potato cultivar. This allowed us to determine unambiguously the type of cytoplasmic DNA for a number of varieties. The results are shown in Table 4.

Table 4. In silico PCR product sizes for a set of NTCP primers and the types of cpDNA for potato cultivars grown in Russia.

Cultivar	NTCP6	NTCP7	NTCP8	NTCP9	NTCP12	NTCP14	NTCP18	Type
Fritella	174	174	255	310	125	150	188	C/W
Golubizna	173	173	252	279	125	149	188	W/T
Grand	175	174	254	280	125	150	188	W
Gusar	175	174	254	280	125	150	188	W
Krasa Meshchery	174	174	254	310	125	151	188	A/W
Krasavchik	173	173	252	279	125	149	188	W/T
Krepysh	174	174	254	310	125	151	188	A/W
Meteor	175	174	254	280	125	150	188	W
Nevsky	174	174	255	310	125	150	188	W
Nikulinsky	173	173	252	279	125	149	188	W/T
Severnoe siyanie	173	173	252	279	125	149	188	W/T
Sudarinya	175	174	254	280	125	150	188	W
Symphonia	173	173	252	279	125	149	188	W/T
Udacha	174	174	255	310	125	150	188	C/W
Zhukovsky	174	174	254	310	125	151	188	A/W

For three sets of primers (NTCP8, NTCP9, NTCP12), we did not find the size of obtained products among the fragment lengths characteristic of the known types of chloroplast DNA [19]. Evaluation of the fragment lengths for other primers allowed us to attribute the chloroplast DNA of cvs. Grand, Gusar, Meteor, Nevsky and Sudarinya to type W unambiguously. For other cultivars, the data on fragment lengths (Table 4) do not allow us to determine the type of chloroplast DNA unambiguously. According to Table 4, cvs. Golubizna, Krasavchik, Nikulinsky, Severnoe siyanie and Symphonia have either the W-type or T-type of cytoplasmic DNA. Cultivars Krasa Meshchery, Krepysh and Zhukovsky have either W- or A-type cytoplasmic DNA. For the Fritella and Udacha cultivars, cytoplasmic DNA can be assigned to either type W or type C.

Additional information to identify the type of chloroplast DNA can be obtained by analysis of restriction enzyme sites [12]. We performed such an analysis in silico. The results are given in Table S3 (Supplementary File). The plastid genomes studied contain different numbers of restriction sites. For the *Bam*HI enzyme, 60 restriction sites were detected in the cvs. Golubizna, Krasavchik, Nikulinsky, Severnoe siyanie and Symphonia; 61 sites were detected for the other cultivars. Analysis of restriction fragment sizes allowed us to identify unambiguously type T cytoplasmic DNA for cvs. Golubizna, Krasavchik, Nikulinsky, Severnoe siyanie and Symphonia (Table S4, Supplementary File). Because the plastid genome sequences of cvs. Severnoe siyanie and Symphonia are identical to the reference sequence of the cv. Desiree, its cytoplasmic DNA is also of the T-type.

Of the remaining five cultivars, a 15.6 kb fragment was detected in the genome by restriction with the *Bam*HI enzyme (Table S4, Supplementary File), which is a W-type marker. In addition, a sequence comparison of the plastid genomes of these varieties demonstrates a 241-bp deletion that affects one of the *Bam*HI enzyme restriction sites. This is a major marker for W-type chloroplast DNA.

Thus, we identified 5 cultivars with T-type cpDNA (Golubizna, Krasavchik, Nikulinsky, Severnoe siyanie and Symphonia) and 10 cultivars with W-type cpDNA (Fritella, Grand, Gusar, Krasa Meshchery, Krepysh, Meteor, Nevsky, Sudarinya, Udacha and Zhukovsky) in a sample of potato cultivars grown in Russia.

3.4. Genetic Diversity of Potato Plastomes

To assess the genetic diversity of the plastid genomes of the potato cultivars from this work, we reconstructed a phylogenetic tree based on the multiple alignment of their

sequences with those of other potato specimens. Additional sequences include 28 cultivars bred in the Ural region [36], 6 commercial cultivars from Europe/North America [24], 13 South American accessions [7] and the *S. tuberosum* cv. Desiree [6]. The resulting tree is shown in Figure 2.

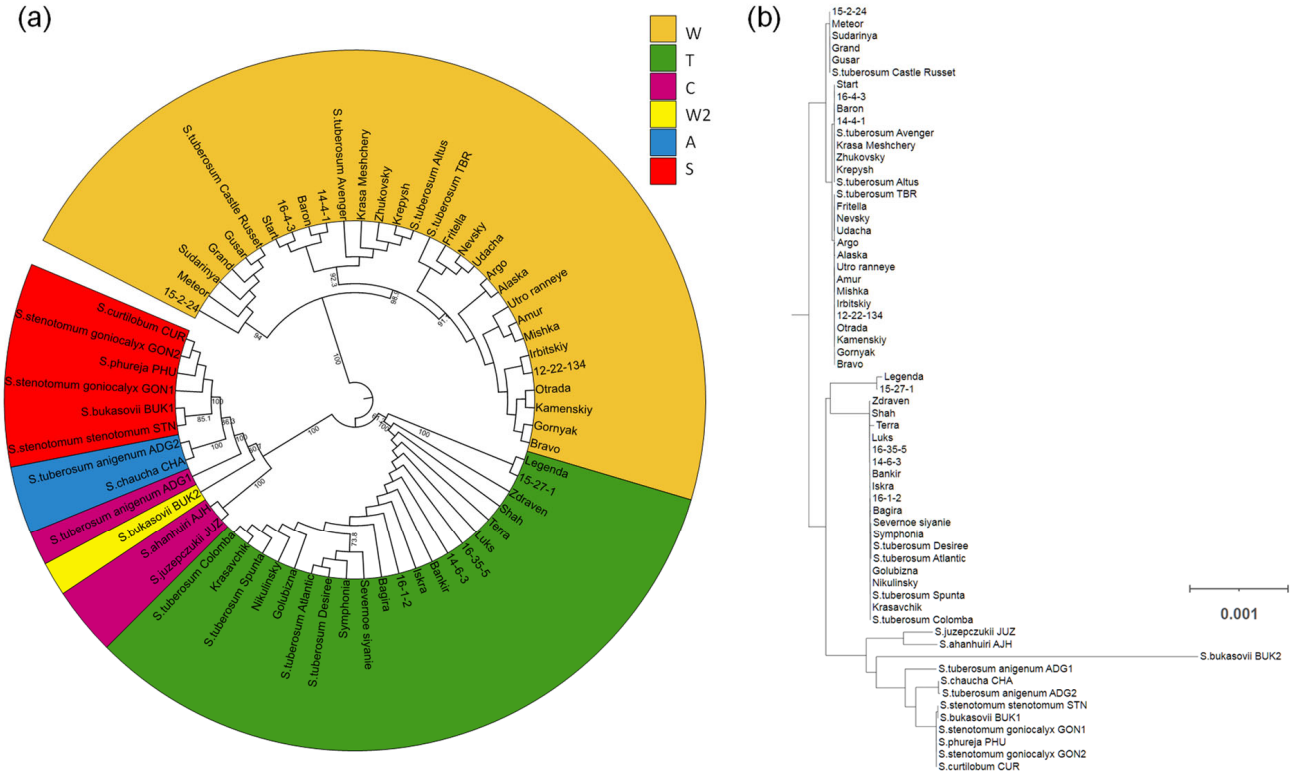

Figure 2. Phylogenetic tree of 63 potato plastomes: (**a**) Circular cladogram; (**b**) Phylogram. The background color of the specimens in the cladogram (**a**) corresponds to the type of the cpDNA (see legend in the upper right corner of the panel): sand (W-type), green (T-type), purple (C-type), yellow (W2-type), blue (A-type) and red (S-type).

In the cladogram (Figure 2a), the specimens studied are divided into three large clusters. The first one includes South American potato accessions (with the exception of TBR). They are represented by four types of cytoplasmic DNA: S, A, C and W2. The second cluster includes *S. tuberosum* potato cultivars with T-type cpDNA. This cluster includes the plastomes of *S. tuberosum* cv. Desiree, commercial cultivars Atlantic, Spunta and Colomba, 12 potato cultivars bred in the Ural region and 5 potato cultivars from the current work. The third cluster includes *S. tuberosum* specimens with W-type cpDNA. These are commercial cultivars Altus, Avenger and Castle Russet, *S. tuberosum* accession from South America (TBR), 16 cultivars bred in the Ural region and 10 potato cultivars from the present work. Thus, the results of the full sequence plastome similarity analysis demonstrate separation of the potato specimens into clusters with similar types of cpDNA.

The phylogram (Figure 2b) demonstrates the high degree of similarity between potato cultivars within the W and T clusters, in comparison with the plastomes of potato accessions from South America. These plastomes are of high variability; the *S. bukasivii* BUK2 accession demonstrates the highest difference from other specimens.

The cluster structure of potato samples with W-type chloroplast DNA (Figure 2a) is noteworthy. Two sub-clusters are well distinguished on the tree, the first of which includes potato cvs. Grand, Gusar, Meteor and Sudarinya from the present work. W/γ cytoplasmic DNA type was previously determined for these four cultivars using the experimental approach [22,38,39]. The second cluster includes cvs. Krasa Meshchery, Zhukovsky, Krepysh, Fritella, Nevsky and Udacha from the present work. These six

cultivars have type D cytoplasmic DNA according to experimental data [14,22,38,39]. Thus, cultivars that have the same type of cpDNA (W) but different types of cytoplasmic DNA (W/γ, D) [14] on this tree fall into two different sub-clusters.

4. Discussion

Chloroplast DNA for different potato accessions provides important information about their genetic diversity and cytoplasmic DNA type. It is of interest in the study of the genetic diversity and phylogeny of potatoes, and of cytoplasmic sterility, as well as for applications in biotechnology and genetic engineering. In the present work, we reconstructed and analyzed plastomes of 15 potato cultivars grown in Russia. The results demonstrate the conservation of plastome gene composition. Differences are observed in single nucleotide polymorphisms, microsatellite repeats and insertions/deletions. The number of SNPs ranged from 114 to 121 for our genomes. This is smaller in general than the SNP numbers for the South American accessions [7] and close for the cultivars bred in the Ural region [23] (note that in our work we compared our sequences with single reference plastome). Fewer than half of the SNPs are localized within the genes. This is in agreement with the data for the South American accessions [7]. We also identified a highly variable plastome region at positions 63,001–68,000, which is also characteristic of the South American accessions and the potato cultivars bred in the Ural region.

The results of the analysis of microsatellite repeats are also consistent with the data on South American potato accessions [7]: the number of repeats varies from 53 to 56 and among them mono-repeats prevail (~70%). In general, compared with the South American accessions, the plastomes of the potato cultivars grown in Russia are more homogeneous, which is generally consistent with our results on the whole genome analysis [22].

Using in silico methods, we have identified different types of cpDNA of our potato cultivars. Among the identified types of cytoplasmic DNA, there are 10 cultivars with W-type cpDNA and 5 cultivars with T-type cpDNA. The data we obtained are consistent with the results of the experimental determination of the cytoplasmic DNA type of Russian potato cultivars [14,38,39] for all the accessions except cv. Nikulinsky. Its cytoplasmic DNA type was recently determined as D [14], but in this work we showed that its DNA type is T (by the presence of a 241 bp deletion).

Reconstruction of the phylogenetic tree for the plastomes based on their multiple alignment showed their clustering by similarity in agreement with the cpDNA type. This is also in full agreement with previous studies on the reconstruction and analysis of potato plastid genomes [7,23]. However, the clustering of plastomes for the potato cultivars grown in Russia is dissimilar to the results of the assessment of their genetic similarity using protein sequences from the nuclear genome [22]. This may reflect the complex mechanisms of potato genome evolution during the breeding process.

It was shown that the type of cpDNA is related to a number of important agronomic traits [16–18]. Recently, Antonova et al. [40] evaluated the genetic diversity of potato cultivars bred in Russia and its neighboring countries using SSR-loci and markers associated with the resistance *R*-genes *H1*, *Gro1-4* and *Sen1*. They compared molecular screening data with the published results of the laboratory and field-tested resistance of potato varieties to *Globodera rostochiensis* (pathotype Ro1) and *Synchytrium endobioticum* (pathotype 1). They also clustered potato varieties according to the similarity of these markers. Six cultivars from our work were used in ref. [40]: Golubizna and Nikulinsky with T-type cpDNA; and Krepysh, Nevsky, Sudarinya and Udacha with W-type cpDNA. We did not find any similarity between the clustering of these cultivars according to the cpDNA and *R*-gene markers. Both T-type cultivars (Golubisna and Nikulinsky) appeared to be susceptible to *G. rostochiensis* pathotype Ro1. The W-type cultivars Krepysh and Sudarinya are resistant to *G. rostochiensis* pathotype Ro1; Nevsky and Udacha are susceptible. Thus, the relationship between the types of cpDNA and resistance to *G. rostochiensis* pathotype Ro1 is unlikely, according to these data. Further analysis is required to identify agronomical traits of Russian cultivars that may be related to the cpDNA type.

5. Conclusions

Based on the results of potato genome sequencing using short reads, we reconstructed the complete sequences of 15 plastids of *S. tuberosum* potato cultivars grown in Russia. Our analysis allowed us to determine the composition and location of genes for these plastid DNAs. It was shown that the plastid genome contains both highly and low-variable regions. The region at position 63,001–68,000 nt has the highest variability. We determined the types of cpDNA based on in silico methods: 10 cultivars have cpDNA of the W-type and 5 cultivars have cpDNA of the T-type. We evaluated the genetic diversity of plastid DNA for our potato cultivars and for the previously reconstructed plastomes of the South American accessions, European/North American commercial cultivars and potato cultivars of the Ural selection. The results show that plastid DNAs of the same type forms clusters by sequence similarity, in agreement with previous studies. Our analysis provides the basis for further genetic studies of Russian potato cultivars, including the study of the plastome structure's relationship with the agronomic characteristics of plants.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy13061454/s1, Supplementary file.pdf. References [7,23,24] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, E.A.S., A.V.K. and D.A.A.; methodology, D.I.K.; software, D.I.K.; validation, D.I.K. and D.A.A.; formal analysis, D.I.K. and D.A.A.; investigation, D.I.K. and D.A.A.; resources, E.A.S. and A.V.K.; data curation, D.I.K. and D.A.A.; writing—original draft preparation, D.I.K. and D.A.A.; writing—review and editing, E.A.S., A.V.K. and D.A.A.; visualization, D.I.K. and D.A.A.; supervision, E.A.S., A.V.K. and D.A.A.; project administration, E.A.S. and D.A.A.; funding acquisition, E.A.S. and A.V.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Kurchatov Genome Center of the Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, in agreement with the Ministry of Education and Science of the Russian Federation, no. 075-15-2019-1662.

Data Availability Statement: The assembled chloroplast genomes have been deposited to NCBI GenBank database (accessions OQ865325-OQ865339).

Acknowledgments: Authors are grateful to Ekaterina Sergeeva for valuable help and discussions. The data analysis was performed using computational resources of the Novosibirsk State University High-Performance Computing Center and the "Bioinformatics" Joint Computational Center supported by the budget project No. FWNR-2022-0020.

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

References

1. Rogalski, M.; do Nascimento Vieira, L.; Fraga, H.P.; Guerra, M.P. Plastid Genomics in Horticultural Species: Importance and Applications for Plant Population Genetics, Evolution, and Biotechnology. *Front. Plant Sci.* **2015**, *6*, 586. [CrossRef] [PubMed]
2. Wicke, S.; Schneeweiss, G.M.; Depamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [CrossRef] [PubMed]
3. Gitzendanner, M.A.; Soltis, P.S.; Wong, G.K.-S.; Ruhfel, B.R.; Soltis, D.E. Plastid Phylogenomic Analysis of Green Plants: A Billion Years of Evolutionary History. *Am. J. Bot.* **2018**, *105*, 291–301. [CrossRef] [PubMed]
4. Huang, B.; Ruess, H.; Liang, Q.; Colleoni, C.; Spooner, D.M. Analyses of 202 Plastid Genomes Elucidate the Phylogeny of Solanum Section Petota. *Sci. Rep.* **2019**, *9*, 4454. [CrossRef]
5. Wheeler, G.L.; Dorman, H.E.; Buchanan, A.; Challagundla, L.; Wallace, L.E. A Review of the Prevalence, Utility, and Caveats of Using Chloroplast Simple Sequence Repeats for Studies of Plant Biology. *Appl. Plant Sci.* **2014**, *2*, 1400059. [CrossRef]
6. Chung, H.-J.; Jung, J.D.; Park, H.-W.; Kim, J.-H.; Cha, H.W.; Min, S.R.; Jeong, W.-J.; Liu, J.R. The Complete Chloroplast Genome Sequences of Solanum Tuberosum and Comparative Analysis with Solanaceae Species Identified the Presence of a 241-Bp Deletion in Cultivated Potato Chloroplast DNA Sequence. *Plant Cell Rep.* **2006**, *25*, 1369–1379. [CrossRef]
7. Achakkagari, S.R.; Kyriakidou, M.; Tai, H.H.; Anglin, N.L.; Ellis, D.; Strömvik, M.V. Complete Plastome Assemblies from a Panel of 13 Diverse Potato Taxa. *PLoS ONE* **2020**, *15*, e0240124. [CrossRef]
8. Yan, L.; Zhu, Z.; Wang, P.; Fu, C.; Guan, X.; Kear, P.; Zhang, C.; Zhu, G. Comparative Analysis of 343 Plastid Genomes of *Solanum* Section *Petota*: Insights into Potato Diversity, Phylogeny, and Species Discrimination. *J Syst. Evol.* **2022**. [CrossRef]

9. Gavrilenko, T.; Antonova, O.; Shuvalova, A.; Krylova, E.; Alpatyeva, N.; Spooner, D.M.; Novikova, L. Genetic diversity and origin of cultivated potatoes based on plastid microsatellite polymorphism. *Genet. Resour. Crop. Evol.* **2013**, *60*, 1997–2015. [CrossRef]
10. Anisimova, I.N.; Gavrilenko, T.A. Cytoplasmic Male Sterility and Prospects for Its Utilization in Breeding, Genetic Studies and Seed Production of Potato. *Vavilov. J. Genet. Breed* **2017**, *21*, 83–95. [CrossRef]
11. Lössl, A.; Adler, N.; Horn, R.; Frei, U.; Wenzel, G. Chondriome-Type Characterization of Potato: Mt α, β, γ, δ, ε and Novel Plastid-Mitochondrial Configurations in Somatic Hybrids. *Theor. Appl. Genet.* **1999**, *98*, 1–10. [CrossRef]
12. Hosaka, K. Who Is the Mother of the Potato? Restriction Endonuclease Analysis of Chloroplast DNA of Cultivated Potatoes. *Theor. Appl. Genet.* **1986**, *72*, 606–618. [CrossRef] [PubMed]
13. Hosaka, K. Successive Domestication and Evolution of the Andean Potatoes as Revealed by Chloroplast DNA Restriction Endonuclease Analysis. *Theor. Appl. Genet.* **1995**, *90*, 356–363. [CrossRef]
14. Gavrilenko, T.A.; Klimenko, N.S.; Alpatieva, N.V.; Kostina, L.I.; Lebedeva, V.A.; Evdokimova, Z.Z.; Apalikova, O.V.; Novikova, L.Y.; Antonova, O.Y. Cytoplasmic Genetic Diversity of Potato Varieties Bred in Russia and FSU Countries. *Vestn. VOGiS* **2019**, *23*, 753–764. [CrossRef]
15. Hosaka, K.; Sanetomo, R. Development of a Rapid Identification Method for Potato Cytoplasm and Its Use for Evaluating Japanese Collections. *Theor. Appl. Genet.* **2012**, *125*, 1237–1251. [CrossRef]
16. Lössl, A.; Götz, M.; Braun, A.; Wenzel, G. Molecular Markers for Cytoplasm in Potato: Male Sterility and Contribution of Different Plastid-Mitochondrial Configurations to Starch Production. *Euphytica* **2000**, *116*, 221–230. [CrossRef]
17. Sanetomo, R.; Gebhardt, C. Cytoplasmic Genome Types of European Potatoes and Their Effects on Complex Agronomic Traits. *BMC Plant Biol.* **2015**, *15*, 162. [CrossRef]
18. Vanishree, G.; Patil, V.U.; Kaur, R.P.; Bhardwaj, V.; Chakrabarti, S.K.; Kumar, M. Cytoplasmic Types of Indian Potato Cultivars and Their Effect on Important Agronomic Traits. *Agric. Res.* **2022**, *11*, 390–397. [CrossRef]
19. Tiwari, J.K.; Devi, S.; Chandel, P.; Ali, N.; Bhardwaj, V.; Singh, B.P. Organelle Genome Analysis in Somatic Hybrids Between Solanum Tuberosum and S. Pinnatisectum Revealed Diverse Cytoplasm Type in Potato. *Agric. Res.* **2016**, *5*, 22–28. [CrossRef]
20. Martyrosyan, E.V.; Ryzhova, N.N.; Kochieva, E.Z. Polymorphism of Chloroplast Microsatellite DNA Loci in Russian Potato Cultivars. *Russ. J. Genet.* **2007**, *43*, 1325–1327. [CrossRef]
21. Smyda-Dajmund, P.; Śliwka, J.; Janiszewska, M.; Zimnoch-Guzowska, E. Cytoplasmic Diversity of Potato Relatives Preserved at Plant Breeding and Acclimatization Institute in Poland. *Mol. Biol. Rep.* **2020**, *47*, 3929–3935. [CrossRef] [PubMed]
22. Karetnikov, D.I.; Vasiliev, G.V.; Toshchakov, S.V.; Shmakov, N.A.; Genaev, M.A.; Nesterov, M.A.; Ibragimova, S.M.; Rybakov, D.A.; Gavrilenko, T.A.; Salina, E.A.; et al. Analysis of Genome Structure and Its Variations in Potato Cultivars Grown in Russia. *Int. J. Mol. Sci.* **2023**, *24*, 5713. [CrossRef] [PubMed]
23. Lihodeevskiy, G.A.; Shanina, E.P. Structural Variations in the Genome of Potato Varieties of the Ural Selection. *Agronomy* **2021**, *11*, 1703. [CrossRef]
24. Hoopes, G.; Meng, X.; Hamilton, J.P.; Achakkagari, S.R.; de Alves Freitas Guesdes, F.; Bolger, M.E.; Coombs, J.J.; Esselink, D.; Kaiser, N.R.; Kodde, L.; et al. Phased, Chromosome-Scale Genome Assemblies of Tetraploid Potato Reveal a Complex Genome, Transcriptome, and Predicted Proteome Landscape Underpinning Genetic Diversity. *Mol. Plant* **2022**, *15*, 520–536. [CrossRef]
25. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef]
26. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 1 February 2023).
27. Jin, J.-J.; Yu, W.-B.; Yang, J.-B.; Song, Y.; de Pamphilis, C.W.; Yi, T.-S.; Li, D.-Z. GetOrganelle: A Fast and Versatile Toolkit for Accurate de Novo Assembly of Organelle Genomes. *Genom. Biol.* **2020**, *21*, 241. [CrossRef]
28. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—Versatile and Accurate Annotation of Organelle Genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [CrossRef]
29. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An Online Program to Visualize the Junction Sites of Chloroplast Genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [CrossRef]
30. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-Web: A Web Server for Microsatellite Prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [CrossRef]
31. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]
32. Page, A.J.; Taylor, B.; Delaney, A.J.; Soares, J.; Seemann, T.; Keane, J.A.; Harris, S.R. SNP-Sites: Rapid Efficient Extraction of SNPs from Multi-FASTA Alignments. *Microb. Genom.* **2016**, *2*, e000056. [CrossRef] [PubMed]
33. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [CrossRef] [PubMed]
34. Krzywinski, M.; Schein, J.; Birol, İ.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An Information Aesthetic for Comparative Genomics. *Genom. Res.* **2009**, *19*, 1639–1645. [CrossRef] [PubMed]
35. San Millán, R.M.; Martínez-Ballesteros, I.; Rementeria, A.; Garaizar, J.; Bikandi, J. Online Exercise for the Design and Simulation of PCR and PCR-RFLP Experiments. *BMC Res. Notes* **2013**, *6*, 513. [CrossRef]
36. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef]

37. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]
38. Rybakov, D.A.; Antonova, O.Y.; Chukhina, I.G.; Fomina, N.A.; Klimenko, N.S.; Zheltova, V.V.; Meleshin, A.A.; Kochieva, E.Z.; Oves, E.V.; Apshev, K.K.; et al. Nomenclatural Standards and Genetic Passports of Potato Cultivars Bred in the A.G. Lorkh All-Russian Research Institute of Potato Farming. *Biotehnol. I Sel. Rastenij* **2021**, *3*, 5–52. [CrossRef]
39. Klimenko, N.S.; Gavrilenko, T.A.; Chukhina, I.G.; Gadzhiev, N.M.; Evdokimova, Z.Z.; Lebedeva, V.A. Nomenclatural Standards and Genetic Passports of Potato Cultivars Bred at the Leningrad Research Institute for Agriculture "Belogorka". *Biotehnol. I Sel. Rastenij* **2021**, *3*, 18–54. [CrossRef]
40. Antonova, O.Y.; Shvachko, N.A.; Novikova, L.Y.; Shuvalov, O.Y.; Kostina, L.I.; Klimenko, N.S.; Shuvalova, A.R.; Gavrilenko, T.A. Genetic Diversity of Potato Varieties Bred in Russia and Its Neighboring Countries Based on the Polymorphism of SSR-Loci and Markers Associated with Resistance *R*-genes. *Russ. J. Genet. Appl. Res.* **2017**, *7*, 489–500. [CrossRef]