horticulturae

MDPI

Article

Genotyping of Autochthonous Rose Populations in the Netherlands for Effective Ex Situ Gene Conservation Management

Joukje Buiteveld [1,*], Alisia Smolka [2] and Marinus J. M. Smulders [2,*]

[1] Centre for Genetic Resources, Wageningen University & Research, P.O. Box 47, NL-6700 AA Wageningen, The Netherlands
[2] Plant Breeding, Wageningen University & Research, P.O. Box 386, NL-6700 AJ Wageningen, The Netherlands
* Correspondence: joukje.buiteveld@wur.nl (J.B.); rene.smulders@wur.nl (M.J.M.S.)

Abstract: Most wild rose species in the Netherlands belong to *Rosa* section *Caninae* (dogroses), with *Rosa arvensis* (section *Synstylae*) and *Rosa spinosissima* (section *Pimpinellifoliae*) as other indigenous species. All species are rare, often found in small populations or as scattered individuals, except for *Rosa canina* and *Rosa corymbifera*. Conservation strategies have been developed for these roses, with a focus on ex situ methods, including clonal archives and seed orchards, using vegetative propagation from the original shrubs. Efficient collection management aims at preservation of maximum genetic diversity with a minimum of duplicated genotypes. However, dogrose taxonomy is complex because of species hybridization, different ploidy levels, and their matroclinal inheritance due to Canina meiosis. They can also reproduce vegetatively through root suckers. In order to assess the genetic structure and the levels of genetic diversity and clonality within and among the wild rose populations in the Netherlands, we genotyped individuals in wild populations and accessions in the ex situ gene bank with 10 highly polymorphic microsatellite markers. The analysis revealed 337 distinct multilocus genotypes (MLGs) from 511 sampled individuals, with some MLGs shared across different species and sites. The genetic structure analysis showed distinct clusters separating non-dogrose species from the *Caninae* section. Geographic distribution of MLGs indicated both local and widespread occurrences. Redundancy analysis identified 152 distinct MLGs from 244 gene bank accessions, suggesting a 38% redundancy rate. Core collections were optimized to retain genetic diversity with minimal redundancy, selecting subsets of 20–40 individuals from different species groups. The study highlights the value of genetic characterization in guiding sampling strategies for dogroses. We propose a two-step approach that may be used to reveal clonality and redundancy and to optimize core collections of species that combine sexual and vegetative reproduction, to maximize genetic capture in ex situ gene banks.

Keywords: conservation genetics; rose; *Rosa*; dogrose; clonality; microsatellite marker; sampling strategy; gene bank

Citation: Buiteveld, J.; Smolka, A.; Smulders, M.J.M. Genotyping of Autochthonous Rose Populations in the Netherlands for Effective Ex Situ Gene Conservation Management. *Horticulturae* **2024**, *10*, 777. https://doi.org/10.3390/horticulturae10080777

Academic Editor: Jinzhi Zhang

Received: 14 June 2024
Revised: 15 July 2024
Accepted: 19 July 2024
Published: 23 July 2024

1. Introduction

In the Netherlands, the majority of wild rose taxa belong to the *Rosa* section *Caninae*, commonly referred to as dogroses (*Rosa* L. section *Caninae* DC.). Of the section *Synstylae*, only *Rosa arvensis* is an indigenous wild species in the Netherlands. The indigenous *Rosa spinosissima* belongs to the section *Pimpinellifoliae*. Many habitats for roses have disappeared in the past 75 years due to cutting of hedges and forest edges, land consolidation in agriculture, and grazing. Additionally, changes in forest management, such as abandoning the traditional coppice practice, which resulted in closure of forest canopies, contributed to the disappearance of light-demanding woody species such as roses [1]. As a result, all rose taxa are now considered rare, often found in small groups or scattered, except for *R. canina*,

R. corymbifera, and *R. spinosissima*, which are more common [1]. Nine rose species are threatened (i.e., they have been classified as vulnerable, endangered, or critically endangered) or near threatened according to the Red List of Vascular Plants of the Netherlands [2].

Conservation strategies for wild rose species in the Netherlands have focused on ex situ, in vivo measures such as establishing clonal archives and seed orchards. For this, cuttings of the original shrubs at different localities in the Netherlands are vegetatively propagated and maintained as clonal archives, or, when the number of accessions collected is greater than 30, are planted as a seed orchard. In the latter case, the collections are also registered as a seed source on the National List of Recommended Varieties and Provenances of Trees and Shrubs [3].

Maintaining living shrubs in field gene banks is a widely used method of ex situ conservation, but it is labour-intensive and expensive. When limited resources are available, efficient collection management is of utmost importance. This includes prioritizing which indigenous populations to preserve and to sample those individuals that are as representative as possible of the diversity present in the in situ populations. The ultimate goal is to establish collections that have the most genetic diversity possible within the fewest possible accessions. Molecular genetic diversity studies provide one of the best tools for making these decisions. Microsatellite markers, also known as Simple Sequence Repeat markers, are useful for informing conservation management, and they are particularly effective for identifying clonality. Esselink et al. [4] developed a set of highly polymorphic microsatellite markers in roses which was used to identify hybrid rose varieties [5] and garden rose varieties [6], but it is also suitable to characterize genetic diversity in various wild rose species [7,8].

Dogroses have a number of special features that affect the distribution of genetic variation within and between populations. Dogroses can hybridize naturally with each other [9–12]. This means that in mixed populations of different taxa, intra- and interspecific hybridization may occur. As a result, populations of dogroses often exhibit complex genetic structures with varying degrees of introgression and hybrid swarm formation. Rose species are preferentially outcrossing [13], but polyploid dogroses are self-compatible and can produce seeds by selfing, albeit with a lower seed production compared to outcrossing. The extent to which selfing or cross-pollination occurs in natural populations has not been determined [14]. Self-compatible species have significantly lower within-population variation and higher between-population differentiation [15]. Dogroses can show apomixis as detected in crosses involving *R. caesia*, *R. dumalis*, *R. rubiginosa*, *R. sherardii*, and *R. mollis* and verified by microsatellites [15,16]. Nybom et al. [16] estimated that a minority (5.5%) of the seedlings in their investigation were derived by apomixis, although selfing could not be ruled out. Like selfing, apomixis can lead to reduced genetic diversity within populations and enhanced genetic differentiation among different populations of dogroses. Moreover, dogroses exhibit matroclinal inheritance. This results from the so-called 'Canina meiosis', a modified form of meiosis in which only two sets of homologous chromosomes pair (seven bivalents), independent of the ploidy level (tetraploid, pentaploid, or hexaploid), while the remaining chromosomes occur as univalents [16]. Only the single set of chromosomes that form bivalents are transmitted by male and female gametes, whereas the remaining non-recombining univalents are exclusively transmitted through the egg cells. As a result of this unbalanced meiosis, offspring closely resemble their mother. Even if dogroses were highly outcrossing, this predominance of maternal inheritance of the chromosomes would result in a diversity pattern comparable to that of selfing or apomictic individuals. As they spread by seeds, some of these near-clones were found to occur over large distances across North-Western Europe [8]. Finally, besides sexually or apomictic seed production, dogroses are also able to propagate vegetatively by root suckers (e.g., *R. canina*). Hence, plants growing in a short distance of each other may also form clonal groups. Most population genetic studies on European dogroses have focused on validating the taxonomic systematics of these complex and polyploid taxa (phylogeny) and gaining insight into hybridization within and between subsections [10,12,17–19]. The results generally confirmed the current

taxonomic classification of subsections but found no support for several of the dogrose (micro)species [8,18,19].

This study aimed to investigate the genetic diversity of wild rose populations within and among populations in the Netherlands using a set of 10 microsatellite markers to support effective ex situ gene conservation management. The marker set developed by Esselink et al. [4] was chosen for this study as they demonstrated that fingerprinting with a subset of two or three microsatellite loci was already sufficient to distinguish the varieties tested in their study, which makes it a useful tool for clone detection and collection management. We genotyped rose individuals collected in the Dutch field gene bank, together with additional individuals from the wild populations from which they originated, to determine the occurrence of (near) clones in the collection and in situ populations. By conducting a clonal analysis at the group-level (e.g., *Rubigineae*), we were able to measure the amount of redundancy in the gene bank collections. With this information, we addressed the following questions: (1) How many clones occur within the same population and taxon or are shared between populations and taxa? (2) Which individuals from the wild populations are clones of individuals already present in the gene bank? (3) Where do these clones occur geographically, and does this coincide with the locations of the clones that are present in the gene bank? (4) How can this information guide additional sampling of the wild populations to maximize the genetic capture in the collection?

2. Materials and Methods

2.1. Plant Material and Sampling

We collected 511 samples of native wild roses (*Rosa* spp.) at 35 localities in the Netherlands (Table 1). Sampling occurred in the frame of the European Generose project (QLK5-CT-2002-01278). Most of the samples (394) in the dataset belong to the section *Canineae* (dogroses), but samples of *R. spinosissima* (58) of section *Pimpinellifoliae* and *R. arvensis* (59) of section *Synstylae* were also included. In total, the dataset contained 17 species and two unnamed taxa (*Caninae_spec* and *Rubigineae_spec*) for which the original species determination in the field was unclear. The identification of species followed Henker [20], which is similar to that used by De Riek et al. [19] and Reichel et al. [8]. Recently, the species nomenclature of dogroses in the Netherlands' Flora [21] has been changed to follow the wider recognition of hybrids as proposed by Bakker et al. [22]. These names are not used here, but they are shown in Table 2 for cross-reference.

Table 1. Samples and genotypic richness measures. Abbreviations: N—total number of samples per species; MLG—distinct multilocus genotypes; eMLG—expected number of MLGs at the lowest common sample size; SE—the standard error for the rarefaction analysis; H—Shannon's diversity index; G—genotypic richness; E—evenness; R—clonal richness.

Group	Species	N	MLG	eMLG	SE	H	G	E	R
arv	*R. arvensis*	59	52	9.87	0.35	3.96	50.45	0.96	0.88
can	*R. subcanina*	6	6	6.00	0.00	1.79	6.00	1.00	1.00
can	*R. subcollina*	11	7	6.55	0.50	1.77	4.84	0.79	0.60
can	*R. canina*	93	76	9.73	0.51	4.23	58.84	0.85	0.82
can	*R. corymbifera*	56	43	9.26	0.83	3.59	27.03	0.74	0.76
can	*R. balsamica*	44	30	8.74	1.00	3.18	17.93	0.74	0.67
can	*R. dumalis*	5	5	5.00	0.00	1.61	5.00	1.00	1.00
can	*R. pseudoscabriuscula*	1	1	1.00	0.00	0.00	1.00	NA	NA
can	*R. tomentosa*	57	39	9.04	0.91	3.46	23.72	0.74	0.68
can	*R. caesia*	3	3	3.00	0.00	1.10	3.00	1.00	1.00
can	*Caninae_spec.*	10	10	10.0	0.00	2.30	10.00	1.00	1.00
can	*R. sherardii*	10	2	2.00	0.00	0.50	1.47	0.72	0.11
rub	*R. gremlii*	35	10	4.45	1.13	1.55	3.07	0.55	0.26
rub	*R. rubiginosa*	34	18	6.58	1.29	2.30	5.16	0.46	0.52
rub	*R. elliptica*	2	2	2.00	0.00	0.69	2.00	1.00	1.00

Table 1. *Cont.*

Group	Species	N	MLG	eMLG	SE	H	G	E	R
rub	*R. micrantha*	15	8	6.00	0.92	1.80	4.79	0.75	0.50
rub	*R. agrestis*	10	3	3.00	0.00	0.64	1.52	0.58	0.22
rub	*Rubigineae_spec.*	2	1	1.00	0.00	0.00	1.00	NA	0.00
spi	*R. spinosissima*	58	36	8.98	0.89	3.39	24.03	0.81	0.61
	Total	511	352						

Table 2. Species classification used in this study, according to Henker et al. [20], with cross-reference to the species names used by Duistermaat [21].

Species Name in This Study	Taxon Name Used by Duistermaat [21]	(Sub)section	Group
R. arvensis	*R. arvensis*	*Synstylae*	arv
R. balsamica	*R. tomentella*	*Tomentellae*	can
R. caesia	*R. caesia*	*Caninae*	can
R. canina	*R. canina*	*Caninae*	can
Caninae_spec.	*n.d.*	*Caninae*	can
R. corymbifera	*R. corymbifera*	*Caninae*	can
R. dumalis	*R. vosagiaca*	*Caninae*	can
R. pseudoscabriuscula	*R. x suberectiformis*	*Vestitae*	can
R. sherardii	*R. sherardii*	*Vestitae*	can
R. subcanina	*R. x subcanina*	*Caninae*	can
R. subcollina	*R. x subcollina*	*Caninae*	can
R. tomentosa	*R. tomentosa*	*Vestitae*	can
R. agrestis	*R. agrestis*	*Rubigineae*	rub
R. elliptica	*R. elliptica*	*Rubigineae*	rub
R. gremlii	*R. x gremlii*	*Rubigineae*	rub
R. micrantha	*R. micrantha*	*Rubigineae*	rub
Rubigineae_spec.	*n.d.*	*Rubigineae*	rub
R. rubiginosa	*R. rubiginosa*	*Rubigineae*	rub
R. spinosissima	*R. spinosissima*	*Pimpinellifoliae*	spi

The intended number of individuals per species sampled at each site was five (for small populations) or 30 (for large populations). However, for the very rare species *R. elliptica*, *R. pseudoscabriuscula*, and *R. caesia*, only a few individuals could be sampled, while for *R. canina*, *R. arvensis*, and *R. spinossisima*, one to two populations were sampled more intensively with up to 45 individuals. In addition to tissue sampling for DNA analysis, cuttings were taken from some of the individuals for vegetative propagation to establish the gene bank collection. This ex situ collection included 244 individuals (hereafter called accessions).

2.2. Microsatellite Genotyping

Genomic DNA was extracted from silica-gel-dried leaf samples, quantified, and stored at −80 °C until use. For all samples, we performed multiplex PCRs to amplify allele combinations at 10 microsatellite loci, originally developed by Esselink et al. [4] (Table 3). PCR amplification was based on Esselink et al. [4] but modified as described by Reichel et al. [8] for the GE set. Fluorescent-labelled amplification products were separated and detected on an ABI Prism 3700 DNA Analyser (Applied Biosystems, Foster City, CA, USA). All samples were genotyped in accordance with reference alleles for each locus, as described in Esselink et al. [4], using Genotyper 3.5 NT (Applied Biosystems). For 409 wild accessions, the genotype data of 7 of the 10 microsatellite markers were also included in the analysis of genetic diversity across Central and North-Western Europe [8]. The genotype data for the gene bank accessions have not been analysed before.

Table 3. The 10 microsatellite markers used. The linkage groups are according to Spiller et al. [23] and Hibrand Saint-Oyant et al. [24]. Primer pairs were published in Mayland-Quellhorst and Wissemann [25]. All reverse primers are pigtailed with GTTT according to Brownstein et al. [26] to reduce stuttering.

Locus	Linkage Group	Allele Size Range [bp]	Repeat Motif	Forward Primer Fluorescently Labelled	Reverse Primer	Fluorescent Label Used
RhAB40	4	201–236	(TC)(AC)	AAT TTG TGT CAA TGC CAA ACA C	GTTTCTTGTCTCCAACCCATCGAGGTTTG	Pool 4-HEX
RhAB73	7	151–215	(CT)(CA)	GGT TAG ACG GGT GGA AGA AG	GTT TAC TGC CGA TAG AAG TAT TTC ATC A	Pool 1-NED
RhB303	2	83–151	(GA)	CAC TGC AAC AAC CCA ATA GC	GTT TCT TGT CTT CAG CTT AGA CTG TGC TG	Pool 2-HEX
RhD201	1	165–242	(TCT)	GGT ATG CAA ATA AGA GAT ACA GT	GTT TCT TCC TAA CAA ACC CAT TTT GAA AGG G	Pool 1-6-FAM
RhD221	4	163–233	(TCT)	CGT AAT TGC TGT GTG ACT GCT TT	GTTTCTTGCCGCTACGAGGAAAATCAA	Pool 3-HEX
RhE2b	6	151–195	(TGT)	CTT TGC ATC AGA ATC TGC TGC ATT	GTTTCTTGCAGACACAGTTCATTAAAGCAG	Pool 4-NED
RhEO506	2	186–343	(CAG)(CAA)(CAG)	GAA GCC TCA GCA GCA TCC TCA AAT	GTT TCT TCA GTG CCA ACA AGC CCA TTG G	Pool 2-6-FAM
RhO517	1	164–275	(GAC)	CGG CGA CGA ACA AAT CAG CAT ATC	GTT TCT TTG AAG AAC GAG GCG CAG CGT AA	Pool 2-NED
RhP50	3	225–406	(CGG)	TGA TGA AAT CAT CCG AGT GTC AG	GTT TCA CTT TCA TTG GAA TGC CAG AAT	Pool 3-6-FAM
RhP518	5	119–184	(CAT)CAAT(CAT)	TTC GAT CTC CAT CTG CAA GA	GTT TCT TCT TAT AAT CTA TTA CGA AGG CTG G	Pool 4-6-FAM

2.3. MLGs and Genetic Diversity

Allele dosage of polyploids can be assessed from microsatellite data but only with difficulty (as described for tetraploid rose samples by Esselink et al. [27]. More generally, "allelic phenotypes" [5] are used, hereafter called multilocus genotypes (MLGs). For this, every allele is recorded, but the dosage of each allele is ignored. Identical MLGs based on perfectly matching DNA profiles for all alleles of the 10 markers were considered to be clonal MLGs.

Multilocus genotypes (MLGs) were analysed with the R packages polysat [28] and poppr [29] to obtain indices of genotypic diversity, including genotypic richness (G), clonal richness as $R = (G - 1)/(N - 1)$, Shannon's diversity index (H), and Evenness (E) for each species in the dataset. Genetic structure and the genetic relationships among genotypes were analysed by principal coordinate analysis (PCoA) based on Bruvo distances, as described in Reichel et al. [8]. QGIS (v3.28) was used for maps and estimation of the geographical distance between individuals sharing an MLG.

2.4. Core-Set Analysis

The R package Core Hunter 3 [30] in R (https://cran.r-project.org/package=corehunter, accessed on 10 July 2022) was used to construct core sets. Two different distance functions were used, as proposed by Odong et al. [31]: average entry-to-nearest-entry distance (E-NE) and average accession-to-nearest-entry distance (A-NE). The E-NE algorithm aims to maximize the average distance between each selected individual and the nearest other selected item in the core set. It favours diverse core sets in which each individual is sufficiently different from the most similar other selected item. Higher values for E-NE indicate lower redundancy. A-NE minimizes the average distance between each individual (from the full dataset) and the closest selected item in the core set. It favours representative core sets in which all items from the original dataset are represented by similar individuals in the selected set. A precomputed distance matrix of Bruvo distances was applied for calculating the core sets.

Our objective for the core-set analysis was to optimize the set concerning redundancy and to investigate how the current gene bank collection for *R. arvensis*, the *Rubigineae,* and the *Canineae* group could be optimized by adding new individuals to the collection. Depending on the initial sample size of each taxonomic group, we evaluated a range of sample sizes for each of the four groups, from 10 to 200. Using the two optimization objectives, we assessed and compared the whole set and the core sets with decreasing sizes for the total number of alleles retained and for the 'min dist'. In the next step, we selected the 'best' core set, which should represent the genetic diversity in the whole sample set and have no redundant samples. If different individuals from the same MLG group were selected in different sets by Core Hunter, we randomly chose one of them to use across the whole list. When the MLG group included samples both in and not in the gene bank, we always chose one of the individuals (sample IDs) that were already present in the gene bank.

3. Results

3.1. Multilocus Genotypes (MLGs) in the Netherlands

Reichel et al. [8] have already investigated the level of clonality in *Rosa* species collected throughout Europe. As expected, they found that shared multilocus genotypes (MLGs) mostly belong to the same species. However, some MLGs were shared between samples classified as different species growing in the same site (especially species in the subsection *Rubigineae*). In a few cases, MLGs were shared between species from different subsections. Some MLGs were widespread (across more than 1000 km). A sizable portion (409 individuals) of the samples in our study were part of their investigation, for which we used three more microsatellite markers. Our results on the 511 samples with 10 markers are consistent with their findings on the existence of clonality within species and the spatial distribution of repeated MLGs and will be briefly described here.

In the dataset of 511 individuals from 17 rose species and two unidentified taxa, sampled from 35 sites in the Netherlands, we distinguished only 337 distinct MLGs (66%, Table 4). Of these MLGs, 59 occurred in more than one individual, and these are potential clonal genotypes. As indicated by the genotypic diversity metrics in Table 1, several species exhibited high values for clonal richness, but notably *R. agrestis*, *R. gremlii*, and *R. sherardii* had MLGs present in multiple samples, resulting in clonal richness values of 0.22, 0.26, and 0.11, respectively.

Table 4. Multilocus genotypes (MLGs) within gene bank and in additional samples in in situ populations (not in gene bank), per taxonomic group.

Group	# Sites (Populations)	#Samples			#MLGs		
		In Gene Bank	Not in Gene Bank (Only Additional)	Total	In Gene Bank	Not in Gene Bank (Only Additional)	Total
spi	4	0	58	**58**	0	36	**36**
arv	3	11	48	**59**	11	41	**52**
rub	15	82	16	**98**	32	3	**35**
can	31	151	145	**296**	109	105	**214**
Total		244	267	**511**	152	185	**337**

Although the majority of shared MLGs were within the same species, some were shared between species. Therefore, we have classified species at a higher taxonomic level and conducted a clonal analysis at the group level (group composition is indicated in Table 2).

In group "spi" (*R. spinossima*) and group "arv" (*R. arvensis*), 36 and 52 MLGs were detected, respectively (Table 4). We detected locally occurring repeated MLGs, but samples from these species did not share MLGs between sites, nor with other species.

The "rub" group consisted of species from the subsection *Rubigineae*. The 98 individuals in the "rub" group resulted in 35 different MLGs, of which six were repeated MLGs (Table 4). One of these (MLG3) was found in 38 samples and at seven sites and was shared between four species (*R. gremlii*, *R. rubiginosa*, *R. micrantha,* and *Rubigineae_spec*.) (Table 5). These occurrences are part of the widespread genotype MLG4 described in Reichel et al. [8].

Table 5. Multilocus genotypes (MLGs) shared between taxa and/or sites in the "rub" group (species from the subsection *Rubigineae*).

MLG	Shared between					
	R. gremlii	*R. micrantha*	*Rubigineae_spec.*	*R. rubiginosa*	*R. agrestis*	Sites
3	x	x	x	x		x
13	x			x		x
66					x	x
72	x	x		x		x
73	x			x		-
209		x				x

The group "can" was formed by species from the subsections *Tomentellae*, *Caninae*, and *Vestitae*. The 296 individuals in the "can" group consisted of 214 MLGs (Table 4). Most of the 36 repeated MLGs were found within a single species and at the same site. However, seven MLGs were shared between species (*R. subcollina*, *R. subcanina*, *R. dumalis*, *R. balsamica*, *R. corymbifera*, *R. tomentosa*, *R. pseudoscabriuscula*, and *Caninae_spec.*) (Table 6). The only *R. pseudoscabriuscula* individual sampled had an identical genotype to an *R. tomentosa* sample. The ten individuals of *R. sherardii*, which is limited to only one population in the Netherlands in the coastal dunes (POP5), were assigned to only two MLGs.

Table 6. Multilocus genotypes (MLGs) shared between taxa and/or sites within the "can" group (species from the subsections *Tomentellae*, *Caninae*, and *Vestitae*).

MLG	Shared between Taxa									Shared between/within Sites		Shared between Subsections
	R. subcollina	*R. dumalis*	*R. subcanina*	*R. canina*	*R. corymbifera*	*Caninae_spec.*	*R. balsamica*	*R. pseudoscabriuscula*	*R. tomentosa*	Between Sites	Within Sites	Between Subsections
23	x	x	x							-	x	
29								x	x	-	x	
108	x						x			-		x
110					x		x			x		x
126					x	x				-		
199					x		x			x		x
319							x			x		
325				x						x		
339							x		x	-		x
342				x						x		

3.2. Genetic Structure

A PCoA analysis was performed on the whole dataset to visualize the genetic structure and relationships among the samples. The first two coordinates of PCoA explained 59.7% of the variation (Figure 1A). As expected, three distinct clusters were formed, dividing samples of the non-dogrose species *R. arvensis* and *R. spinossisima* from each other and from the *Caninae* section. When performing a second PCoA analysis with only samples from the *Caninae* section, the "rub" group (subsection *Rubigineae*) was separated from the "can" group (species of subsections *Canineae*, *Vestitae*, and *Tomentellae*) along the first axis (Figure 1B). Within these two clusters, the species appear to be intermixed.

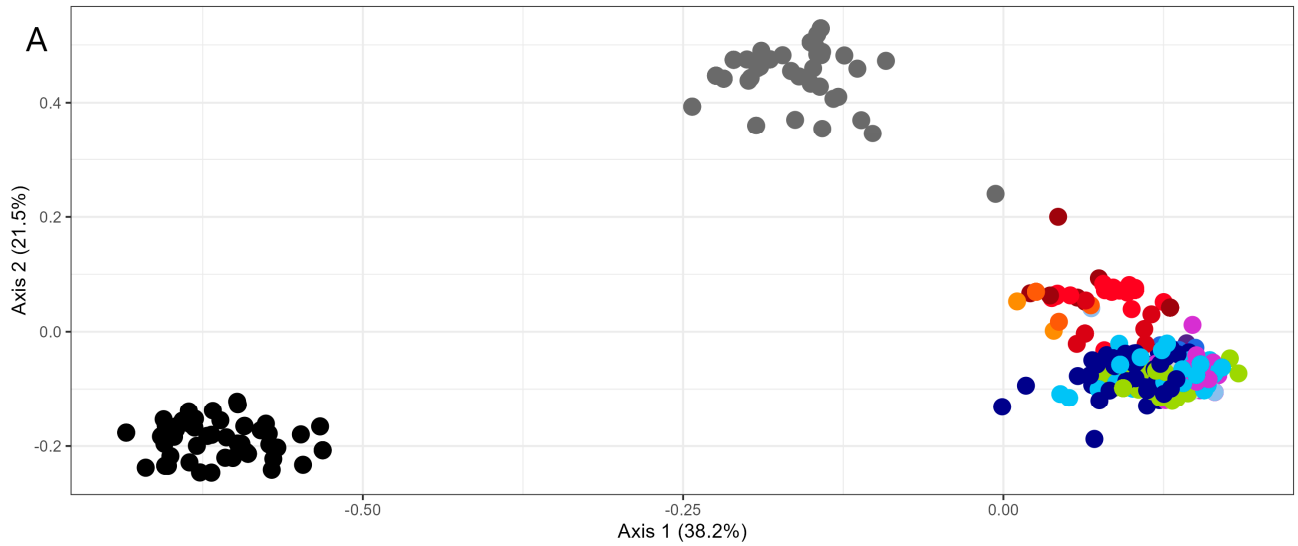

Figure 1. *Cont.*

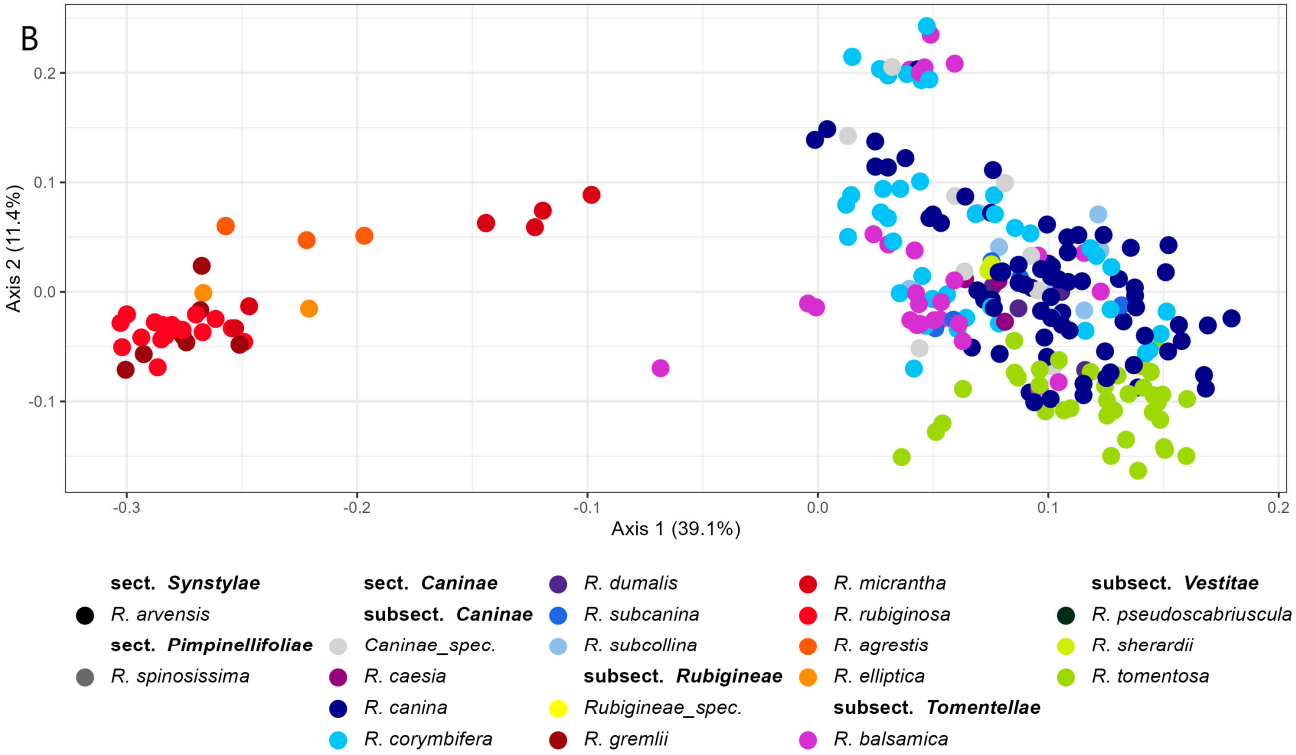

Figure 1. Principal coordinate analysis based on 10 microsatellite loci (**A**) of all rose samples and (**B**) of the samples of the "can" group only. Colours indicate species/subsections.

3.3. Geographic Origin of the MLGs

Most repeated MLGs were local, but some MLGs were widespread. Within the species complex comprising *R. micrantha*, *R. gremlii*, and *R. rubiginosa*, two genotypes (MLG3 and MLG13) were detected at several sites along the coast from the northern island of Schiermonikoog to the south of the Netherlands (Figure 2A), which is a distance of several hundred kilometres. Two other genotypes (MLG72 and MLG209) within this species complex were detected inland at two sites at approx. 10 km from each other. In *R. agrestis*, MLG66 was shared between two sites at a distance of 2.5 km. Within the "can" group, *R. canina* had two widespread MLGs (MLG342 and MLG325) with a maximum distance of 4.3 km. In *R. tomentosa*, the greatest distance measured between two individuals within the same MLG was 2.8 km. Four MLGs involving *R. balsamica* (MLG110, MLG199, MLG319, and MLG339) were found at sites with distances up to 277 km apart.

In all other cases, repeated MLGs were local and were mostly mingled with other genotypes, for instance at the site POP12 for *R. canina*. Also, identical genotypes within *R. tomentosa* populations were mostly mingled with other genotypes (Figure 2B shows POP26 as an example).

3.4. Multilocus Genotypes in the Gene Bank and Evaluation of Redundancy

Of the 511 genotyped individuals, 244 individuals (hereafter referred to as accessions) were from the gene bank (*R. spinossisima* is not included as it is widespread in the Netherlands). We measured the extent of redundancy in the entire collection and at the group-level by identifying accessions with identical MLGs. Combining all species, we identified 152 distinct MLGs from these 244 accessions, demonstrating a possible redundancy of 38% (Table 4). There were no duplicated accessions in the *R. arvensis* collection. The percentage of putatively duplicated accessions in the "rub" and "can" group was 61% and 28%, respectively (Table 4). The high redundancy in the "rub" group was mainly caused by 29 accessions, assigned to *R. gremlii*, *R. micrantha*, and *R. rubiginosa*, sharing the same widespread MLG3.

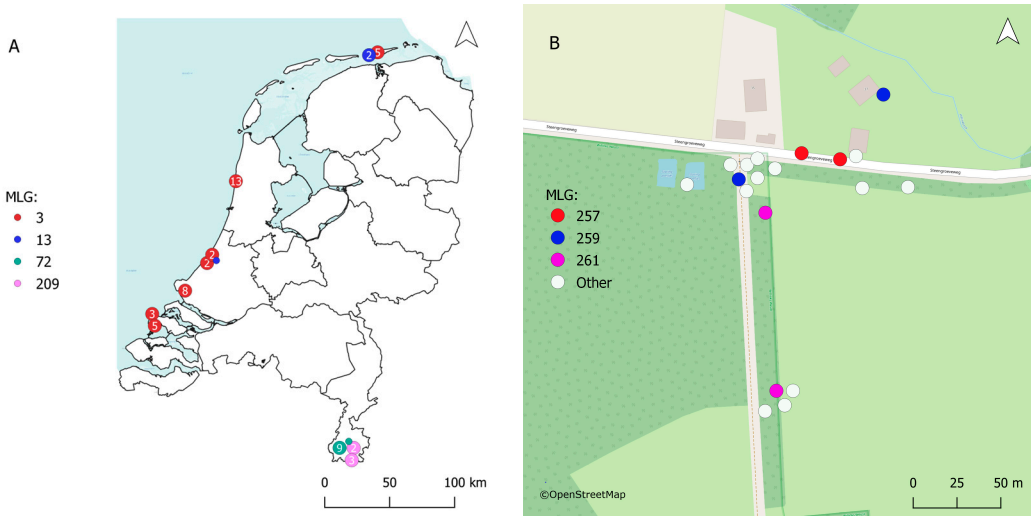

Figure 2. (**A**) Two repeated (clonal) multilocus genotypes (MLG3 and MLG13) in the species complex comprising *R. micrantha*, *R. gremlii*, and *R. rubiginosa*. The two genotypes occur at several sites (indicated by dots) along the coast from the northern island of Schiermonnikoog to the south of the Netherlands. The number in the dots indicates the number of individuals at the site. (**B**) Repeated (clonal) MLGs within *R. tomentosa* populations occur mingled with other genotypes (POP26 as example). Coloured dots indicate different repeated MLGs. White dots indicate unique genotypes.

We checked to what extent the other 267 individuals from the wild populations are from the repeated MLGs already present in the gene bank or represent new genotypes. For *R. arvensis* ("arv" group), most of the 48 additional samples represented new genotypes (41). In the "rub" group, the 16 additional samples resulted in three new MLGs (MLG46, MLG111, and MLG271). For the "can" group, 105 new genotypes were found in the 145 additional individuals.

3.5. Construction of Core Collections

For core collection construction using Core Hunter, the four groups ("arv", "spi", "rub", and "can") were analysed separately. The results regarding the two diversity parameters ('observed number of alleles' and 'min dist') were compared for the whole sample set (all) and the different subsets. Preliminary analysis showed that using the E-NE method better captured the number of alleles of the total sample set, but the 'min dist' values increased more strongly than with the A-NE method. As our goal for a core selection is maximization of the representativeness of the genetic diversity (conserving as much genotypic variation as possible in few accessions) instead of capturing the maximum number of alleles of these loci, we chose the A-NE method to further analyse the different subsets.

Depending on the redundancy detected in the four groups, smaller or larger subsets were chosen for comparison with the total sample set. For the "rub" group, which has a high redundancy (the whole set of 98 samples contains only 35 unique genotypes), reducing the sample set to 20 or 10 individuals did not drastically increase the value for 'min dist' (Table 7). For *R. spinossisma*, reducing the sample set to 10 samples gave a core with substantial larger 'min dist', and including more samples in the subset (20) had added value as it reduced the 'min dist'. For *R. arvensis*, similar results were obtained. In the "can" group, the whole set only contained 214 genotypes due to redundancy, while reducing the sample size from 180 down to 40 had only a slight effect on the genetic distance values.

Table 7. Poppr analysis of the different core sets with decreasing sample sizes and the whole sample set (all) for the diversity parameters number of alleles (#alleles) and genetic distance ('min dist'). The core sets were constructed with Core Hunter using the average accession-to-nearest-entry distance (A-NE) optimization method.

Group	Samples	#Alleles	min_dist
spi	58 (all)	64	0.010
spi	30	63	0.123
spi	20	62	0.123
spi	10	47	0.233
rub	98 (all)	76	0.017
rub	20	73	0.025
rub	10	67	0.050
can	296 (all)	131	0.009
can	200	131	0.013
can	180	130	0.019
can	160	128	0.038
can	140	127	0.038
can	120	127	0.038
can	100	123	0.048
can	80	120	0.048
can	60	116	0.048
can	40	104	0.048
can	20	88	0.120
can	10	75	0.199
arv	59 (all)	81	0.010
arv	50	81	0.021
arv	40	80	0.096
arv	30	70	0.096
arv	20	70	0.175
arv	10	64	0.201

Using these results for 'min dist', we selected the subsets spi20, rub20, arv30, and can40 to compare with the total sample set. This comparison aimed to determine from which populations individuals should be sampled for establishing a new collection in the case of *R. spinossisima* and where additional individuals in these selected subsets could be sampled to extend the gene bank collection. The spi20 core set comprised individuals from all four populations, with samples proportionally extracted based on the populations' sample sizes. Similarly, the optimized subset for *R. arvensis* included samples from all three populations and were also proportional to their sample sizes. This subset included 24 of the 41 additional sampled genotypes, along with half of the genotypes already present in the gene bank. The rub20 set included individuals sourced from 10 of the 15 populations sampled. Core Hunter selected two of the three additionally sampled individuals that were not yet present in the collection. All species were represented in the subset, including both individuals of *R. elliptica* and two of the 10 *R. agrestis* individuals. However, within the *R. micrantha–R. gremlii–R. rubiginosa* species complex, individuals of *R. gremlii* and *R. rubiginosa* were underrepresented relative to the number of individuals sampled due to high redundancy levels in these species. The can40 subset comprised 22 genotypes already present in the collection and 18 additional genotypes from 16 out of the 31 populations, typically one or two genotypes per population. The five *R. dumalis* individuals, currently not present in the gene bank, were not included in the core set.

4. Discussion

4.1. Clonality

The pattern of genetic diversity in rose species in relation to clonality within species (or groups of (micro)species) varied across species or species groups. The level of clonality

was particularly high in *Rubigineae*. The largest groups of identical multilocus genotypes (MLGs) were found in species belonging to this group such as *R. rubiginosa* and *R. gremlii*, but also in species belonging to the *Caninae* subsection. Some of them were geographically widespread, as was also observed by Reichel et al. [8]. Sharing of identical MLGs between species within subsections was explained by misidentification of species, while the sharing of MLGs between species of different subsections may be the result of too low resolution of the marker system [8]. Our investigation of clonality in the Dutch material aligns with the outcomes of this European study. Nonetheless, in certain species, *R. agrestis*, *R. gremlii*, and *R. sherardii*, the occurrence of clonality was higher compared to the observations at the European level. Notably, the sole population of *R. sherardii* within the Netherlands, situated within the coastal dunes, appeared to be largely clonal. Similarly, *R. agrestis*, which occurs in a few sites only, of which Bemelerberg (POP32) is the most important, demonstrated pronounced clonality.

Dogrose species can reproduce vegetatively by root suckers. In *R. sherardii*, this might explain the high clonality within the analysed population as this species is known to form rhizomatous thickets. However, alongside vegetative reproduction, there is concurrent hemisexual reproduction, selfing, and occasionally apomixis [15,32]. Nevertheless, discerning the specific impacts of selfing or apomixis from those attributed to Canina meiosis remains challenging [16]. Consequently, it remains uncertain what the predominant reproductive strategy within the examined dogrose populations is. In contrast, non-dogrose species such as *R. arvensis* and *R. spinossisima* did not exhibit widespread clonality or shared clones with other taxa, and the observed clonality in these species is consistent with vegetative reproduction. *R. arvensis* has the ability to root branches upon contact with soil [33], while *R. spinossisima* is known to form dense thickets through root suckers [25].

4.2. Redundancy in the Gene Bank

The patterns of redundancy in the collected gene bank accessions largely followed the pattern of clonality, with, among others, many identical MLGs in accessions of *R. rubiginosa* and *R. gremlii*. Genetic data can provide important insights in the occurrence of clonality in natural populations, causing redundancy in gene bank collections (e.g., in black poplar and cassava [34–36]), next to detecting redundancy by other causes including documentation and determination errors, as detected in a collection of old apple varieties [37,38] and cherry varieties [39]. However, genetic data also allow devising a strategy for improved sampling based on core collection concepts [40]. Various software programs exist to assist in building core collections. Here, we used Core Hunter to establish core sets for all four groups based on the entire sample set, which clearly demonstrated the potential for eliminating the duplicated MLGs and genetically close MLGs. By using a distance-based criterion, subsets representing the genotypic diversity without redundant individuals were efficiently constructed at the group level. However, disregarding the species identity resulted in certain species being underrepresented within these subsets. For example, *R. dumalis* had no representatives in the core set can40. Reducing redundancy at the group level, particularly in living collections, keeps them manageable. From a gene bank management perspective, we recommend conserving genotypes at the group level (a higher taxonomic level) rather than maintaining taxa separately, notably in genera in which there is uncertainty over species' identities and potential oversplitting. Nonetheless, recognizing that some clonal collections also function as seed orchards for seed and plant production for re-introduction into natural habitats, combining genotypes at the group level would lead to the harvesting of what technically should be called hybrid seeds. Given that it is standard practice to issue and certify seeds of pure species, harvesting from mixed-species collections is impractical. However, it can be argued that this should not pose a significant issue here, as the species/taxa, particular those belonging to the *R. micrantha*–*R. gremlii*–*R. rubiginosa* species complex and those in subsection *Caninae*, typically coexist in mixed populations in the Netherlands and may hybridize naturally [11,12,41].

Several general sampling guidelines for ex situ collections have been established [42], typically emphasizing sampling a greater number of populations, larger populations, and a geographically broad range. Recent studies indicated that the representation of in situ genetic diversity in ex situ collections can be enhanced by employing an optimized and species-specific sampling strategy [43–45]. It is generally acknowledged that genetic diversity increases with geographic and environmental distance, a factor that should be considered when prioritizing populations for ex situ sampling [46]. Therefore, it seems logical to maximize geographic sampling across the range of wild rose species in the Netherlands to capture the genetic diversity of these species in the ex situ collections. However, our findings suggest that broad geographic sampling is not the most efficient strategy for capturing genotypic diversity in the wild rose species analysed. For the non-dogrose species in our study ("arv" and "spi" group), we conclude that, in particular, clonality and its spatial structuring within populations must be considered. Specifically, *R. spinossisima* exhibits substantial clonality within populations. For this species, which predominantly inhabits coastal areas, extensive geographic sampling within these coastal populations is advisable for establishing ex situ collections. Conversely, for *R. arvensis*, which exhibited low clonality within populations and no redundancy in the collections, additional sampling efforts across all populations could significantly enhance the genotypic diversity of the collection. For the dogrose species, the widespread occurrence of clonal genotypes indicates that geographically broad sampling may not be the most efficient strategy. Specifically, within the *R. rubiginosa–R. gremlii–R. micrantha* species complex, there was extensive sharing of identical genotypes among the sampled populations. Consequently, future sampling at new sites in the Netherlands is unlikely to yield new genotypes. In the "can" group, additional sampling did result in a substantial number of new genotypes, which could enrich the genotypic diversity of the collection. However, there was no discernible pattern of clonality to guide sampling efforts. Although widespread MLGs occasionally occurred, most clonal MLGs groups were found within populations, and some populations exhibited extensive clonal groups. These results suggest that future sampling should particularly aim at targeting populations which were undersampled.

The widespread occurrence of clonal MLGs is likely attributable to human intervention, as suggested by Reichel et al. [8]. An alternative explanation could be the dispersal of seeds by birds, given that many rose species in the Netherlands are found in dune areas. Birds, which follow migration routes along the coast and use rose hips as a food source, may facilitate the spread of these species and clonal individuals [1].

5. Conclusions

Given the observation that clones are shared among species within the dogrose section, and that widespread clones exist, we propose a two-step sampling strategy to establish ex situ collections of roses, particularly living collections. In living collections, individuals are conserved ex situ via cuttings, with the objective of preserving as many distinct genotypes as possible. Establishing living collections is both labour-intensive and costly. Genetic data can assist in optimizing the composition of the collection by selecting genetically diverse individuals, which is especially important when space constraints limit the number of genotypes that can be included in a field collection. In a two-step approach, individuals from as many populations as possible would first be genotyped to assess genetic diversity and the presence of clones and then propagated. Based on this genetic data, decisions can be made to enhance representativeness either by sampling a larger area of the distribution or by re-sampling specific populations that are the most diverse. Furthermore, knowing that some propagation failures are inevitable, specific individuals can be re-sampled to ensure the success of the collection while other failures may be ignored if they are found be part of large MLGs.

This two-step approach may be broadly applicable while establishing ex situ collections of species which combine sexual and vegetative reproduction through various mechanisms, including woody species such as Populus nigra [34], P. tremuloides [47], and

willows through root suckers; figs and vines by areal or adventitious roots; and herbaceous species (including the strawberry, grasses and sedges [48], and lilies [49]) through runners, rhizomes, or adventitious bulbs.

Author Contributions: Conceptualization: J.B. and M.J.M.S.; methodology and analyses: J.B. and A.S.; writing—original draft preparation: J.B. and M.J.M.S.; writing—review and editing: all; visualization: J.B. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the Netherlands' Ministry of Agriculture, Fisheries, Food Security and Nature (KB project KB-34-013-003). Part of the genotypic data were produced in the European Framework V project Generose (QLK5-CT-2002-01278).

Data Availability Statement: Most data are available within the article. The data of the underlying genotypes are available on request from the authors.

Acknowledgments: The authors would like to thank Heleen de Jonge for help in creating the figures in QGIS.

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

References

1. Maes, B. *Inheemse Bomen en Struiken in Nederland en Vlaanderen: Herkenning, Verspreiding, Geschiedenis en Gebruik*; Uitgeverij Boom: Amsterdam, The Netherland, 2013.
2. Sparrius, L.; Odé, B.; Beringen, R. Basisrapport Rode Lijst Vaatplanten 2012 Volgens Nederlandse en IUCN-Criteria. FLORON Rapport 57. 2012. Available online: https://www.floron.nl/Portals/1/Downloads/Basisrapport%20Rode%20Lijst%20vaatplanten%202012.pdf (accessed on 13 June 2024).
3. Rassenlijst Bomen. Available online: https://www.rassenlijstbomen.nl/nl/home/soorten.htm (accessed on 13 June 2024).
4. Esselink, D.; Smulders, M.J.M.; Vosman, B. Identification of cut-rose (*Rosa hybrida*) and rootstock varieties using robust Sequence Tagged Microsatellite markers. *Theor. Appl. Genet.* **2003**, *106*, 277–286. [CrossRef] [PubMed]
5. Smulders, M.J.M.; Esselink, D.; Voorrips, R.E.; Vosman, B. Analysis of a database of DNA profiles of 734 Hybrid Tea Rose varieties. *Acta Hortic.* **2009**, *836*, 169–174. [CrossRef]
6. Vukosavljev, M.; Zhang, J.; Esselink, G.D.; Van 't Westende, W.P.C.; Cox, P.; Visser, R.G.F.; Arens, P.; Smulders, M.J.M. Genetic diversity and differentiation in roses: A garden rose perspective. *Sci. Hortic.* **2013**, *162*, 320–332. [CrossRef]
7. Samiei, L.; Naderi, R.; Khalighi, A.; Bushehri, A.-A.; Mozaffarian, V.; Esselink, G.D.; Smulders, M.J.M. In search of genetic diversity in Rosa foetida Herrmann in Iran. *Acta Hortic.* **2009**, *836*, 25–30. [CrossRef]
8. Reichel, K.; Herklotz, V.; Smolka, A.; Nybom, H.; Kellner, A.; De Riek, J.; Smulders, M.J.M.; Wissemann, V.; Ritz, C.M. Untangling the hedge: Genetic diversity in clonally and sexually transmitted genomes of European wild roses, *Rosa* L. *PLoS ONE* **2023**, *18*, e0292634. [CrossRef] [PubMed]
9. Ritz, C.M.; Wissemann, V. Microsatellite Analyses of Artificial and Spontaneous Dogrose Hybrids Reveal the Hybridogenic Origin of Rosa micrantha by the Contribution of Unreduced Gametes. *J. Hered.* **2011**, *102*, 217–227. [CrossRef]
10. Kellner, A.; Ritz, C.M.; Wissemann, V. Hybridization with invasive Rosa rugosa threatens the genetic integrity of native Rosa mollis. *Bot. J. Linn. Soc.* **2012**, *170*, 472–484. [CrossRef]
11. Herklotz, V.; Ritz, C.M. Spontane Hybridisierung von Hundsrosen (*Rosa* L. sect. *Caninae* (DC). Ser.) an einem natürlichen Vorkommen in der Oberlausitz (Sachsen, Deutschland). *Peckiana* **2014**, *9*, 119–131.
12. Herklotz, V.; Ritz, C.M. Multiple and asymmetrical origin of polyploid dogrose hybrids (*Rosa* L. sect. *Caninae* (DC) Ser.) involving unreduced gametes. *Ann. Bot.* **2017**, *120*, 209–220. [CrossRef]
13. Wissemann, V.; Hellwig, F.H. Reproduction and Hybridisation in the Genus *Rosa*, Section *Caninae* (Ser.) Rehd. *Bot. Acta* **1997**, *110*, 251–256. [CrossRef]
14. Werlemark, G.; Nybom, H. Dogroses: Botany, Horticulture, Genetics, and Breeding. *Hortic. Rev.* **2010**, *36*, 199–255. [CrossRef]
15. Nybom, H.; Esselink, G.D.; Werlemark, G.; Vosman, B. Microsatellite DNA marker inheritance indicates preferential pairing between two highly homologous genomes in polyploid and hemisexual dog-roses, *Rosa* L. *Sect. Caninae DC. Heredity* **2004**, *92*, 139–150. [CrossRef] [PubMed]
16. Nybom, H.; Esselink, G.D.; Werlemark, G.; Leus, L.; Vosman, B. Unique genomic configuration revealed by microsatellite DNA in polyploid dogroses, Rosa sect. *Caninae. J. Evol. Biol.* **2006**, *19*, 635–648. [CrossRef] [PubMed]
17. Jürgens, A.H.; Seitz, B.; Kowarik, I. Genetic differentiation of *Rosa canina* (L.) at regional and continental scales. *Plant Syst. Evol.* **2007**, *269*, 39–53. [CrossRef]
18. Koopman, W.J.M.; Wissemann, V.; de Cock, K.; van Huylenbroeck, J.; De Riek, J.; Sabatino, G.J.; Visser, D.; Vosman, B.; Ritz, C.M.; Maes, B.; et al. AFLP Markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). *Am. J. Bot.* **2008**, *95*, 353–366. [CrossRef] [PubMed]
19. De Riek, J.; De Cock, K.; Smulders, M.J.M.; Nybom, H. AFLP-based population structure analysis as a means to validate the complex taxonomy of dogroses (*Rosa* section *Caninae*). *Mol. Phylogenet. Evol.* **2013**, *67*, 547–559. [CrossRef] [PubMed]

20. Henker, H. Rosa. In *Gustav Hegi–Illustrierte Flora von Mitteleuropa. IV/2C*; Conert, H.J., Jaeger, E.J., Kadereit, J.W., Schultze-Motel, W., Wagenitz, G., Weber, H.E., Eds.; Parey: Berlin, Germnay, 2000; pp. 1–108.
21. Duistermaat, L. *Heukels' Flora van Nederland*; Noordhoff: Groningen, The Netherlands, 2020; ISBN 9789001589561.
22. Bakker, P.; Maes, B.; Maskew, R.; Stace, C. Dog-roses (*Rosa* sect. *Caninae*): Towards a consensus taxonomy. *Br. Ir. Bot.* **2019**, *1*, 7–19. [CrossRef]
23. Spiller, M.; Linde, M.; Hibrand-Saint Oyant, L.; Tsai, C.-J.; Byrne, D.H.; Smulders, M.J.M.; Foucher, F.; Debener, T. Towards a unified genetic map of diploid rose. *Theor. Appl. Genet.* **2011**, *122*, 489–500. [CrossRef] [PubMed]
24. Hibrand Saint-Oyant, L.; Ruttink, T.; Hamama, L.; Kirov, I.; Lakhwani, D.; Zhou, N.N.; Bourke, P.M.; Daccord, N.; Leus, L.; Schulz, D.; et al. A high-quality genome sequence of Rosa chinensis to elucidate ornamental traits. *Nat. Plants* **2018**, *4*, 473–484. [CrossRef]
25. Mayland-Quellhorst, E.; Wissemann, V. Population structure of Rosa spinosissima L. on the Frisian Islands and introgression from cultivated material. *Nord. J. Bot.* **2021**, *39*, e02941. [CrossRef]
26. Brownstein, M.L.; Carpten, J.D.; Smith, J.R. Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* **1996**, *20*, 1004–1010. [CrossRef] [PubMed]
27. Esselink, G.D.; Nybom, H.; Vosman, B. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting—Peak ratios) method. *Theor. Appl. Genet.* **2004**, *109*, 402–408. [CrossRef] [PubMed]
28. Clark, L.; Drauch Schreier, A. Resolving microsatellite genotype ambiguity in populations of allopolyploid and diploidized autopolyploid organisms using negative correlations between allelic variables. *Mol. Ecol. Resour.* **2017**, *17*, 1090–1103. [CrossRef] [PubMed]
29. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2014**, *2*, e281. [CrossRef] [PubMed]
30. De Beukelaer, H.; Davenport, G.F.; Fack, V. Core Hunter 3: Flexible core subset selection. *BMC Bioinform.* **2018**, *19*, 203. [CrossRef] [PubMed]
31. Odong, T.L.; Jansen, J.; Van Eeuwijk, F.A.; van Hintum, T.J. Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. *Theor. Appl. Genet.* **2013**, *126*, 289–305. [CrossRef] [PubMed]
32. Werlemark, G. Evidence of apomixis in hemisexual dogroses, Rosa section Caninae. *Sex. Plant Reprod.* **2000**, *12*, 353–359. [CrossRef]
33. De Cock, K. Genetic Diversity of Wild Roses (*Rosa* spp.) in Europe, with an in-Depth Morphological Study of Flemish Populations. Ph.D. Thesis, Ghent University, Faculty of Bioscience Engineering, Ghent, Belgium, 2008. Available online: http://hdl.handle.net/1854/LU-472039 (accessed on 13 June 2024).
34. Storme, V.; Vanden Broeck, A.; Ivens, B.; Halfmaerten, D.; Van Slycken, J.; Castiglione, S.; Grassi, F.; Fossati, T.; Cottrell, J.E.; Tabbener, H.E.; et al. Ex-situ conservation of Black poplar in Europe: Genetic diversity in nine gene bank collections and their value for nature development. *Theor. Appl. Genet.* **2004**, *108*, 969–981. [CrossRef] [PubMed]
35. Singh, N.; Wu, S.; Raupp, W.J.; Sehgal, S.; Arora, S.; Tiwari, V.; Vikram, P.; Singh, S.; Chhuneja, P.; Gill, B.S.; et al. Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. *Sci. Rep.* **2019**, *9*, 650. [CrossRef]
36. Carvajal-Yepes, M.; Ospina, J.A.; Aranzales, E.; Velez-Tobon, M.; Correa Abondano, M.; Manrique-Carpintero, N.C.; Wenzl, P. Identifying genetically redundant accessions in the world's largest cassava collection. *Front. Plant Sci.* **2024**, *14*, 1338377. [CrossRef]
37. Van Treuren, R.; van Hintum, T.J. Marker-Assisted Reduction of Redundancy in Germplasm Collections: Genetic and Economic Aspects. *Acta Hortic.* **2003**, *623*, 139–149. [CrossRef]
38. van Treuren, R.; Kemp, H.; Ernsting GJongejans, B.; Visser, L. Microsatellite genotyping of apple (*Malus* × *domestica* Borkh.) genetic resources in the Netherlands: Application in collection management and variety identification. *Genet. Resour. Crop Evol.* **2010**, *57*, 853–865. [CrossRef]
39. Barreneche, T.; Cárcamo de la Concepción, M.; Blouin-Delmas, M.; Ordidge, M.; Nybom, H.; Lacis, G.; Feldmane, D.; Sedlak, J.; Meland, M.; Kaldmäe, H.; et al. SSR-Based Analysis of Genetic Diversity and Structure of Sweet Cherry (*Prunus avium* L.) from 19 Countries in Europe. *Plants* **2021**, *10*, 1983. [CrossRef]
40. Van Hintum, T.H.J.L.; Brown, A.H.; Spillane, C.; Hodgkin, T. *Core Collections of Plant Genetic Resources*; IPGRI Technical Bulletin No. 3; International Plant Genetic Resources Institute: Rome, Italy, 2000.
41. Schanzer, I.A.; Kutlunina, N.A. Interspecific hybridization in wild roses (*Rosa* L. sect. Caninae DC.). *Biol. Bull. Russ. Acad. Sci.* **2010**, *37*, 480–488. [CrossRef]
42. Brown, A.H.D.; Marshall, D.R. A basic sampling strategy: Theory and practice. In *Collecting Plant Genetic Diversity: Technical Guidelines*; CAB International: Wallingford, UK, 1995; Chapter 5, pp. 75–91. Available online: http://cropgenebank.sgrp.cgiar.org/images/file/procedures/collecting1995/Chapter5.pdf (accessed on 13 June 2024).
43. Hoban, S.; Strand, A. Ex situ seed collections will benefit from considering spatial sampling design and species' reproductive biology. *Biol. Conserv.* **2015**, *187*, 182–191. [CrossRef]
44. Hoban, S.; Callicrate, T.; Clark, J.; Deans, S.; Dosmann, M.; Fant, J.; Gailing, O.; Havens, K.; Hipp, A.L.; Kadav, P.; et al. Taxonomic similarity does not predict necessary sample size for ex situ conservation: A comparison among five genera. *Proc. R. Soc. B* **2020**, *287*, 20200102. [CrossRef] [PubMed]

45. Rosenberger, K.; Schumacher, E.; Brown, A.; Hoban, S. Proportional sampling strategy often captures more genetic diversity when population sizes vary. *Biol. Conserv.* **2021**, *261*, 109261. [CrossRef]
46. Di Santo, L.N.; Hamilton, J.A. Using environmental and geographic data to optimize ex situ collections and preserve evolutionary potential. *Conserv. Biol.* **2020**, *35*, 733–744. [CrossRef] [PubMed]
47. DeWoody, J.; Rowe, C.A.; Hipkins, V.D.; Mock, K.E. "Pando" lives: Molecular genetic evidence of a giant aspen clone in central Utah. *West. N. Am. Nat.* **2008**, *68*, 493–497. [CrossRef]
48. Arens, P.; Bijlsma, R.-J.; van 't Westende, W.P.C.; van Os, B.; Smulders, M.J.M.; Vosman, B. Genetic Structure in Populations of an Ancient Woodland Sedge, Carex sylvatica Hudson, at a Regional and Local Scale. *Plant Biol.* **2005**, *7*, 387–396. [CrossRef]
49. Arens, P.; Grashof-Bokdam, C.J.; van der Sluis, T.; Smulders, M.J.M.; Vosman, B. Clonal diversity and genetic differentiation in Maianthemum bifolium between forest fragments of different age. *Plant Ecol.* **2005**, *179*, 169–180. [CrossRef]