Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

WITHDRAWN: Molecular and cytogenetic analyses in Geranium macrorrhizum L. wild Italian plants

Irene Cardinali
University of Perugia: Universita degli Studi di Perugia
Marilena Ceccarelli
marilena.ceccarelli@unipg.it
University of Perugia: Universita degli Studi di Perugia https://orcid.org/0000-0003-4154-1543

Research Article

Keywords: chromosome number, DNA barcoding, FISH, Geranium macrorrhizum L., phylogenetic analysis

Posted Date: November 10th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3499810/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

EDITORIAL NOTE:

The full text of this preprint has been withdrawn by the authors while they make corrections to the work. Therefore, the authors do not wish this work to be cited as a reference. Questions should be directed to the corresponding author.

Abstract

Geranium macrorrhizum L. is a herbaceous species native to southern Europe, and introduced in central Europe and North America. It is widely distributed also in Italy, up to Campania region as the southern limit. In this study, molecular and cytogenetic analyses were carried out on 22 wild plants, collected in central and southern Italy, compared to five cultivated plants, with the main purpose to identify those living near the Marmore waterfalls in central Italy, recently described as the new species *G. lucarinii*. Four barcoding markers (*rbcL*, *matK*, *trnH-psbA* intergenic spacer, and internal transcribed spacer region, ITS) were sequenced and their variability among the plants was evaluated. Chromosome numbers were determined and 45S rDNA was physically mapped by fluorescence *in situ* hybridization (FISH) in both wild and cultivated plants. Moreover, genomic affinity between wild and cultivated plants was evaluated by genomic *in situ* hybridization (GISH). Our results confirmed that all the plants, including Marmore population, belong to *G. macrorrhizum*. Barcoding analyses showed a close similarity among the wild plants, as well as a differentiation, although not significant, between the wild plants on one hand and the cultivated plants on the other. Integrated studies focusing on morphological, genetic and ecological characterization of a larger number of wild populations will allow for knowing the extent of the variability within the species.

INTRODUCTION

Geranium macrorrhizum L. is a herbaceous perennial species with a robust and more or less horizontal rhizome. The aerial stem is erect, quite long, naked up to the inflorescence with zygomorphic flowers, pink or purplish in colour. Basal leaves are in persistent rosette, whereas cauline leaves are opposite, five-lobed (palmate), with glandular and eglandular hairs (Aedo 2017; Yeo 2004). An essential oil containing the monoterpenoids geraniol and beta-citronellol, and several sesquiterpenes including germacrone, are extracted from the aerial parts of the plant to be used in aromatherapy and phytotherapy (Harborne and Williams, 2002; Stoeva, 2002). However, the species is mainly cultivated as an ornamental plant, with cultivars selected for flower colours from white through pink to magenta (Yeo, 2004). Regarding the chromosome number, two cytotypes corresponding to two ploidy level, 2n = 2x = 46 and 2n = 4x = 92, were observed (Baltisberger and Baltisberger 1995; Baltisberger 1991; Gauger 1937; Petrova and Stanimirova 2002; Strid and Franzén 1981; Tan et al. 2011; van Loon 1984).

G. macrorrhizum belongs to the section *Unguiculata* of subgenus *Geranium*, one of the four subgenera of the genus (Aedo 2018). At present, section *Unguiculata* includes only one other species, *G. dalmaticum*, with 2n = 2x = 46 (Aedo 2017; Baltisberger 1984; Yeo 2004). A close morphological similarity exists between the two species, so that in the past *G. dalmaticum* had been considered a subspecies or a variety of *G. macrorrhizum* (cf. Aedo 2017). A third species, *G. kikianum*, endemic to the Taigetos Mountains (Peloponnese, Greece), tetraploid as *G. macrorrhizum*, was ascribed to this section (Tan et al. 2011), but considered a synonym of *G. macrorrhizum* after an in-depth revision (Aedo 2017). *G. dalmaticum* is endemic to a few localities in southern Croatia, Montenegro and northern Albania (Aedo 2017), whereas

G. macrorrhizum has a wider range of distribution, being native to southern Europe, from south-eastern France to Krym, and introduced in central Europe (Aedo 2017) and North America (Hawke 2004).

In Italy, *G. macrorrhizum* is present in most of the northern and central Regions, with few exceptions (Toscana, Marche), whereas it has been reported in only two southern Regions (Molise and Campania; cf. Wagensommer and Venanzoni, 2021). In Umbria (central Italy), the species was reported for the first time in 1837. A small population of *G. macrorrhizum* was observed near the Marmore waterfalls (Sanguinetti and Sebastiani 1837), reported again in 1869 (Fiorini Mazzanti) and then no more until 2016, when it was accidentally rediscovered during field research (Venanzoni 2017). Morphological analyses carried out on both living plants and herbarium samples highlighted that this population (hereinafter called Marmore population) diverged from *G. macrorrhizum* for three traits related to indumentum, leaves and calyx (Wagensommer and Venanzoni 2021). Notable differences also exist in the ecological niche (altitude and vegetation context). Indeed, almost all the Apennine and peninsular stations of *G. macrorrhizum* are located in the mountain belt on calcareous rocks and also in beech forests, or in orophilic screes, at over 1000 m a.s.l., whereas Marmore population lives on calcareous rocky slopes, at 190–250 m a.s.l., in a Mediterranean vegetation context (Wagensommer and Venanzoni 2021). On these bases, Marmore population was elevated to the rank of species, called *G. lucarinii* Venanzoni & Wagens. (Wagensommer and Venanzoni 2021). The new species has a restricted and punctiform distribution, being represented by only that small population. The growth environment is anthropically disturbed, therefore *G. lucarinii* has been considered critically endangered, according to the IUCN (2019) criteria (Wagensommer and Venanzoni 2021).

Accurate species delimitation is fundamental to biology. It has implications not only for a reliable evaluation of biodiversity but also in the use of the organisms at many levels, even for their conservation (Garnett and Christidis 2017; Li et al. 2019). It is now widely accepted that alpha taxonomy, based on morphological characters, is no longer sufficient to guarantee a reliable species description and delimitation. Rather, the integrated use of molecular, karyological, ecological data, along with quantitative morphology, is strongly recommended (De Queiroz 2007; Dejaco et al. 2016; Tiburtini et al. 2022 and references therein).

With the aim of testing the identity of the plants collected at Marmore waterfalls and to shed light on the relationships between the putative new species *G. lucarinii* and *G. macrorrhizum*, DNA barcoding was carried out on Marmore plants, compared to *G. macrorrhizum* wild plants from central and southern Italy and cultivated plants. In addition to the core barcode markers *rbcL* and *matK*, two supplementary markers, nuclear ITS and plastid intergenic spacer *trnH-psbA*, were used (CBOL Plant Working Group 2009; Hollingsworth et al. 2011). The latter was never been sequenced in *G. macrorrhizum* before, but it was used anyway because the high variability makes it a particularly suitable marker for discriminating between closely related species (Federici et al. 2013; Hollingsworth et al. 2011; Kress et al. 2007). In addition, cytogenetic analysis was undertaken in order to know the chromosome number of the plants studied, and to characterize their chromosome complement by means of fluorescence *in situ* hybridization (FISH) of 45S rDNA. Finally, cross GISH (genomic in situ hybridization) experiments were

carried out to assess the genomic affinity between Marmore plants and the cultivated ones. This method is commonly applied to reveal genomic similarity between closely related species, based on the homology of the repetitive DNA sequences (Falistocco 2019 and references therein).

MATERIALS AND METHODS

Plant Material

Plants were collected in field inspections in central and southern Italy, at Marmore waterfalls and Felitto (Campania Region), respectively (cf. Wagensommer and Venanzoni 2021). Ten plants per population were sampled. Being the species rhizomatous, care was taken to collect plants at a suitable distance from each other. The plants from Felitto had been described as *G. macrorrhizum* (Del Guacchio 2002; Salerno 2004; Wagensommer and Venanzoni 2021). Two plants of *G. macrorrhizum*, previously collected at National Park of Abruzzo, Lazio and Molise (NPALM, central Italy) and then transferred to the Botanical Garden of Camerino University (central Italy), were obtained from this institution. Other *G. macrorrhizum* plants were obtained from Botanical or public gardens (Table 1). All the plants were cultivated *ex situ* at the Department of Chemistry, Biology and Biotechnology of Perugia University and used for molecular and cytogenetic analyses.

Table 1
List of the plants studied, their status and geographical provenance

Sample	Status	Provenance
Plant 1	cultivated	Bologna (Emilia-Romagna, northern Italy), Botanical Garden
Plant 2	wild	National Park of Abruzzo, Lazio and Molise (central Italy)
Plant 3	wild	National Park of Abruzzo, Lazio and Molise (central Italy)
Plants from 4a to 4l (population 4)	wild	Felitto (southern Campania, Italy)
Plants from 5a to 5n (population 5)	wild	Cascata delle Marmore (Marmore waterfalls, Umbria, central Italy)
Plant 7	cultivated	Mount Terminillo (Latium, central Italy), public garden
Plant 9	cultivated	Varsavia (Poland), public garden
Plant 10	cultivated	Vipiteno (Trentino Alto-Adige, northern Italy), public garden
Plant 11	cultivated	Campotosto (Abruzzo, central Italy), public garden

DNA extraction, amplification and sequencing

Total genomic DNA was isolated from fresh leaves, using the DNeasy Plant Mini kit (Qiagen, Germany) according to the manufacturer's instructions.

Three plastid markers (*rbcL*, *matK* and *trnH-psbA* intergenic spacer) and the nuclear ITS region (ITS1-5.8S-ITS2), were amplified in a 25 µl volume reaction containing 20 ng of DNA template, 1 µl of each primer (10 pmol/µl) and 0.5 units of MyTaq HS polymerase (Bioline). Amplifications were performed on a thermal cycler 2720 (Applied Biosystems, Foster City, CA, USA). The primer pairs and cycling conditions are listed in Table S1. Two *matK* primer pairs were used. First, the 390F + 1326R pair (Cuènoud et al. 2002) failed amplification. The primer pair 3F + 1R KIM (Costion et al. 2011) produced multiple sequences. In order to obtain a single amplicon, the 1R KIM sequence was modified according to the complementary region on the *G. macrorrhizum matK s*equences found in GenBank database. Amplified products were purified using the ExoSAP-IT® *Express* reagent (Thermo Fisher Scientific Inc.). Sequencing in both directions was performed by Eurofins Genomics service (Germany). Primers used for sequencing were the same as those for amplifications. Electropherograms quality was visually inspected. Sequences were manually edited and aligned using the ClustalW algorithm implemented in BioEdit 7.1.7 (Hall 1999) with the default values. The sequences were compared with those available in GenBank (cf. Table S2) through a BLASTn search (Zhang et al. 2000).

Newly determined sequences were deposited in GenBank (accession numbers OK299101 for *rbcL*, OM417815 for *trnH-psbA*, from OR656480 to OR656483 for ITS, OR668227 for *matK*). Only differing sequences for each locus have been deposited.

Molecular analyses

The identification of variable and parsimony informative sites was carried out using MEGA 11 software (Tamura et al. 2021). MUSCLE (Edgar 2004) was used to align the sequences with the outgroup ones in MEGA 11. Genetic relationships among samples were inferred using both Neighbor-Joining (NJ; Saitou and Nei 1987) and Maximum Likelihood (ML) methods. In NJ analyses, the genetic distances were computed using Kimura 2-parameter (K2P) substitution model (Kimura 1980) for each locus, and were given as units of the number of base substitutions per site. All ambiguous positions were removed for each sequence pair. Bootstrap analysis was done using 1,000 replicates (Felsenstein 1985). For ML analyses, the Tamura-Nei model was used (Tamura and Nei 1993). Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances, and then selecting the topology with superior log likelihood value. Branch lengths measure the number of substitutions per site. Bootstrap analysis was done using 100 replicates. Three *Geranium* species whose ITS sequence showed the highest identity percentage (97.75–97.27%) with GenBank *G. macrorrhizum* sequence DQ525073 were added as outgroup (*G. dalmaticum* DQ525072, *G. lasiopus* KX421242, *G. glaberrimum* KX421239). Two species for which both *rbcL* and *matK* sequences from the same origin were available, were chosen as outgroups (*G. lucidum* MK542503 and JN896161, *G. robertianum* KP963378 and KY687141). Moreover, all the *rbcL* and *matK G. macrorrhizum* sequences available in GenBank were aligned to obtain those dendrograms (see accession numbers in dendrograms). NJ and ML trees were constructed both for each marker and for concatenated markers. In

concatenated trees, sequences of *G. macrorrhizum* and outgroup species were obtained by the sum of three marker sequences probably deriving from different individuals.

Genetic relationships among plants were also investigated through a median-joining network of haplotypes obtained by analyzing the concatenated sequences. The network was constructed with the Network software v.10.2 (www.fluxus-engineering.com) by using the reduced median algorithm (ρ = 2). The term haplotype here used indicates the list of mutations found in the examined sequences in each sample, arbitrarily numbered for the analyses.

Cytogenetic analyses

Root apices were treated with ice cold water for 24 hours at 4°C, then transferred in 8-OH-quinoline (Sigma) 0.02 M for 5 hours at room temperature, and fixed in ethanol-acetic acid 3:1 (v/v). Fixed root tips were washed in an aqueous solution of 6 mM sodium citrate plus 4 mM citric acid, digested with a mixture of 10% pectinase (Sigma), 8% cellulose (Calbiochem), and 2% macerozyme (Serva) in citrate buffer pH 4.6 for 1 h at 37°C, and squashed under a coverslip in a drop of 60% acetic acid. After removing coverslips by the solid CO_2 method, slides were air-dried and used for FISH or GISH experiments. FISH was performed as described in Mascagni et al. (2022). The wheat probe pTa71, containing 18S-5.8S-26S rDNA (Gerlach and Bedbrook 1979) was used. The DNA of nuclei and chromosomes was denatured in a thermal cycler for 6 min at 70°C and the preparations were incubated overnight at 37°C with 2 ng/µl of heat-denatured DNA probe which had been labeled with digoxigenin-11-dUTP (Roche) or biotin-16-dUTP (Roche) by nick-translation. Detection of the digoxigenin or biotin at the hybridization sites was carried out using anti-digoxigenin conjugated with FITC (Roche) or streptavidin conjugated with Cy3 (Sigma), respectively. The preparations were then counterstained with a 2% DAPI (4,6-diamidino-2-phenylindole) solution in McIlvaine buffer pH 7, mounted in antifade solution (AFI; Citifluor), and analyzed with a fluorescence microscope (DMRB, Leica). Images were captured with a digital camera (ILCE-7, Sony) and optimized using Adobe Photoshop 5.0.

The same slides were used for chromosome counting. At least ten DAPI-stained metaphases per plant were analyzed.

For GISH experiments (self-GISH and cross GISH), total genomic DNAs extracted from leaves were used as probes after labelling with biotin-11-dUTP by nick translation following the producer’s protocol (BioNick Labeling System, Invitrogen). GISH procedure was similar to FISH protocol with the exception of the probe concentration, that was 5 ng/µl. These experiments were replicated three times.

RESULTS

Molecular analyses

Sequences of different length were obtained for each marker (Table 2). The *rbcL* sequences used for subsequent analyses were longer than the gene fragment, 599 bp in length, considered the barcode

region by Hollingsworth et al. (2011). *matk* locus proved to be difficult to amplify. The mainly used primer combinations failed in amplification, therefore a specific reverse primer was designed and used in combination with primer 3F KIM (Table S1). Also in this case, the primer pair used produced sequences slightly differing in length from the gene barcode region (Hollingsworth et al. 2011). BLASTn analysis showed that all the *rbcL* and *matK* sequences, either from wild or cultivated plants, were identical to those of *G. macrorrhizum*. Indeed, the identity percentage was in the range 99.86–100.00 and 99.87–100.00 for *rbcL* and *matK* sequences, respectively (Table S2). The range was slightly wider for ITS sequences (98.55–100.00; Table S2). Regarding the intergenic spacer *trnH-psbA*, sequences 339–346 bp in length were obtained. They resulted monomorphic in nucleotide composition and showed the highest identity percentage (95%) with *G. maderense* sequence (Table S2). Their alignment with plastome of *G. maderense* showed that the intergenic spacer *trnH-psbA* in the analysed plants is 288 bp long and has a G + C content of 36.1%.

Table 2
Range of the amplicon sequences for each marker and some characteristics of the fragments aligned for NJ and ML trees

	rbcL	*matK*	ITS	*trnH-psbA*
Sequences length (bp)	692–703	822–837	677–709	339–346
Aligned length for NJ and ML trees (bp)	692	787	623	not aligned
No. of variable sites among aligned sequences	1	1	4	0
No. of parsimony informative sites	1	1	4	0

The network analysis, constructed with the reduced median method, was applied to analyse the genetic relationships among the plants. To this purpose, sequences were aligned and trimmed for each marker, and concatenated sequences were used. Six haplotypes were found, four within the wild plants, and two within the cultivated ones (Fig. 1a). Populations 4 (Felitto, Campania) was homogeneous such as population 5 (Marmore). The two plants from NPALM (plants 2 and 3) showed unique haplotypes, as well as the cultivated plant 10. All the other cultivated samples shared the same haplotype. Within *rbcL* sequences, only a single nucleotide polymorphism was found, consisting in a transversion A/C (nucleotide position 472 in Table S3) which distinguishes the plants from Marmore and NPALM from all the others. Only one SNP, a transition C/T (nucleotide position 149 in Table S3), was also observed among *matK* sequences, distinguishing all the wild plants from the cultivated ones. Instead, thirteen polymorphic sites were observed among ITS sequences. All of them were located in the sequenced portions of the intergenic spacers ITS1 and ITS2. Four polymorphic sites distinguished cultivated plants from the wild ones. Several polymorphic sites resulted heteroplasmic nucleotides in the *G. macrorrhizum* sequence available in GenBank (DQ525073) and only in those of cultivated plants. Heteroplasmy was showed at two sites in wild population 4 and one site in plant 2 from NPALM (Table S3).

Bearing in mind the need to clarify the taxonomical placement of Marmore population, the genetic relationships among plants were more investigated by Neighbor-Joining and ML methods, including sequences of *G. macrorrhizum* and outgroup species retrieved from GenBank. Some characteristics of the trimmed aligned fragments are shown in Table 2. The only polymorphism detected among *rbcL* sequences (see above) was responsible of a weakly supported but clear differentiation among wild plants (Fig. S1). Instead, all the wild plants were included in the main branch of *matK* tree, harbouring a weakly supported sub-cluster including the cultivated ones and two samples of *G. macrorrhizum* from GenBank (Fig. S2). The concatenated *rbcL* + *matK* tree, based on a total of 1,479 bp, showed a highly supported main branch harbouring all the analysed plants. The cultivated plants closely clustered with *G. macrorrhizum* samples, whereas a further differentiation emerged among the wild ones (Fig. S3). The ITS tree harboured two main clades, the first including all the wild plants, the second comprising the cultivated plants (Fig. S4). Some variability can be observed within each cluster. The two plants from NPALM formed a sub-cluster, whereas population 4 was slightly differentiated from population 5. Among the cultivated plants, sample 10 turned out to be more similar to *G. dalmaticum* than to *G. macrorrhizum*. The concatenated tree *rbcL* + *matK* + ITS, for a total of 2,101 bp (Fig. 1b) included only *G. robertianum* as outgroup because the sequences of all the three markers were available in GenBank only for this species. The tree highlighted the differences between the two groups of cultivated and wild plants, already observed in Fig. 1a. The wild plants were in turn grouped into three sub-cluster corresponding to their geographical provenance. ML trees showed the same NJ topology (not shown).

Cytogenetic analysis

Chromosome counts on the DAPI stained metaphases showed that the somatic chromosome number in all the wild plants, including those from Marmore, was 2n = 92, whereas in the cultivated plants it was 2n = 46, with the exception of plant 10, showing 2n = 92.

Due to the small chromosome size, it was difficult to arrange the karyotype. In order to establish at least the number of chromosome pairs carrying ribosomal DNA, FISH was carried out using pTa71 as a probe. Eight hybridization signals related to 45S rDNA were counted on metaphase plates of cultivated plants (Fig. 2a-b), whereas a maximum of 16 signals were observed on metaphase plates of wild plants, comprising those collected at Marmore waterfalls (Fig. 2c-d).

To evaluate the genome affinity between Marmore plants and *G. macrorrhizum*, GISH experiments were carried out by probing the genomic DNA of Marmore plants on chromosomes of cultivated plants and *vice versa* (Fig. 3). Preliminary experiments in which the labelled DNA of Marmore plants or cultivated plants was hybridized to its own chromosomes (self-GISH) were performed to better evaluate, by comparison, the results of cross GISH. Thus, after self-GISH, fluorescent signals, although of different intensity, were observed on each chromosome of the complement in both wild and cultivated plants (Fig. 3a-b). Low intensity signals were easily recognized at the centromeric and pericentromeric region, showing a hybridization pattern typical of satellite DNA. The same hybridization pattern was observed after cross GISH (Fig. 3c-d).

DISCUSSION

DNA barcoding has been developed as rapid and reliable method to identify a species (CBOL Plant Working Group 2009; Hebert et al. 2003). Since then, it has been largely used in basic and applied biodiversity research, to discriminate between morphologically similar taxa, reducing the number of misidentifications, for cultivated flora analyses, or to solve the doubtfully status of some alien species (De Castro et al. 2020 and references therein; Koblmüller 2023). In this study, DNA barcoding was applied to test the identity of plants living near the Marmore waterfalls in central Italy, considered a new species, *G. lucarinii*, morphologically very similar to *G. macrorrhizum* (Wagensommer and Venanzoni 2021). Plastomes of different *Geranium* species showed large structural variation and high rates and patterns of nucleotide substitutions (Park et al. 2017), therefore the use of plastid barcode markers appeared particularly appropriate for studying relationships between *G. lucarinii* and *G. macrorrhizum*.

All the sequences of markers *rbcL*, *matK* and ITS showed identity percentages equal or close to 100% with those of *G. macrorrhizum*. The greatest variability was observed among ITS sequences, as expected due to its nature of bi-parentally inherited marker. The minimum value of the range (98.55%) was higher than the identity percentage (97.76%) between ITS of *G. macrorrhizum* DQ525073 and that of the closely related species *G. dalmaticum* DQ525072, confirming that our ITS sequences correspond to *G. macrorrhizum*.

trnH-psbA intergenic spacer did not contribute to species identification because GenBank database was missing of the reference sequence for *G. macrorrhizum*. However, it proved likewise useful. Indeed, any *trnH-psbA* sequence variation was observed among the plants studied. This sequence monomorphism, unusual for the marker, supports the fact that all the analysed plants seem to belong to the same species.

The distribution of genetic variability in barcoding sequences suggested some differentiation within the plants studied (Fig. 1). The cultivated plants closely clustered with *G. macrorrhizum*, whereas wild plants were clearly grouped into three sub-clusters corresponding to their geographical origin. The plants from National Park of Abruzzo, Lazio and Molise were more closely related to Marmore population than to population 4 from Campania (Felitto). Interestingly, the two samples from the National Park are the same plants used to morphologically compare plants from Marmore, later considered a new species (Wagensommer and Venanzoni 2021). Despite this clustering, it is clear that genetic variation between cultivated and wild plants do not support the existence of two different species.

This finding is also confirmed by the cytogenetic analyses. Two cytotypes, diploid and tetraploid, were detected in this study. Ninety-two chromosomes, corresponding to the tetraploid level, were counted in Marmore plants, as well as in the other wild plants studied, *versus* the 46 chromosomes counted in almost the cultivated plants, with the exception of plant 10. However, the different chromosome number cannot be considered a discriminating factor, because the existence of diploid and tetraploid plants with n = 23 has long been known in *G. macrorrhizum* (see Introduction). Recently, plants with a genome size corresponding to the hexaploid level were found in Croatia (Ćavar-Zeljković et al. 2020). The occurrence

of different ploidy levels in the same species, due to endopolyploidy, is not exclusive to *G. macrorrhizum*. Rather, it is common to many taxa of the *Geranium* genus (Petrova and Stanimirova 2003).

Our FISH analyses confirmed that wild plants have a doubled chromosome number compared to the cultivated ones. The number of 45S rDNA signals was in agreement with the ploidy level, unlike what occurs in many species in which a reduction in the number of ribosomal DNA sites per monoploid genome is observed following polyploidization (Garcia et al. 2017). FISH also showed that the number of chromosome pairs carrying ribosomal DNA is higher than that previously observed in karyotype analyses carried out in different *G. macrorrhizum* Bulgarian populations with 2n = 46 (Petrova and Stanimirova 2002). Indeed, eight hybridization signals, corresponding to four chromosome pairs, were observed in our cultivated plants with 46 chromosomes, whereas only two or three nucleolar chromosome pairs were found by Petrova and Stanimirova (2002).

The genomic affinity between Marmore plants and *G. macrorrhizum* was cytologically investigate by GISH. Repetitive DNA sequences (satellite DNA) are mainly involved in the hybridization reaction. The method provides a powerful tool to study their distribution pattern along chromosomes, especially in species for which there is a lack of genome information (Falistocco 2023; Zhang et al. 2015). Since most satellite DNA sequences are fast evolving in structure, redundancy and localization even within the same species (Garrido-Ramos 2017; Thakur et al. 2021), their detection through GISH could give information about the relationship between related species. The comparison of hybridization patterns after self-GISH and cross GISH in our material showed homology of the repeated sequences between Marmore plants and the cultivated ones.

On account of these results, the taxonomical treatment of the plants from Marmore waterfalls as a new species appears questionable. *G. lucarinii* Venanzoni & Wagens. is described as morphologically closely related to *G. macrorrhizum*, differing in some features regarding especially leaves and calyx, despite the lack of statistical analyses. Further traits distinguishing *G. lucarinii* from *G. macrorrhizum* were the flowering period and the habitat, in term of vegetation context and altitudinal range of distribution. The latter (190–250 m a.s.l.) is partly overlapping with that of *G. macrorrhizum* (50-2800 m a.s.l.). In addition, one of the most southern *G. macrorrhizum* Italian stations, Felitto (population 4 in this study), is also located at low altitude (200–290 m a.s.l.; Del Guacchio 2002; Salerno 2004). A comparative climatic niche analysis could help to clarify the differences in flowering period and morphology observed between Marmore plants and *G. macrorrhizum*, which could be due to adaptation to environmental factors (Li et al. 2019).

Thus, at present, our molecular and cytogenetic data support the presence of only species *G. macrorrhizum* L. in central and southern Italy.

This study is the first report of molecular and cytogenetic characterization of *G. macrorrhizum* Italian populations. The topology of concatenated tree (Fig. 1b) suggests that *G. macrorrhizum* wild populations in central and southern Italy form a genetically fairly homogeneous group, well separated from the cultivated plants. The origin of cultivated plants here studied is not well known, just as there is not

enough information on the status, if cultivated or wild, of *G. macrorrhizum* plants whose sequences were retrieved from GenBank. Beyond this, it is significant that the cultivated plants cluster together and with known *G. macrorrhizum*, whether it be cultivated or not, whereas the wild plants form a distinct cluster. Work is in progress to deepen morphological and genetic studies, extending them to a greater number of wild populations, to estimate the degree of the variability within the species. For the same purpose, the role played by the geographical distribution of the populations, their spatial isolation and consequent gene flow, as well as ecological specialization, will be evaluated. Such integrative approach is fundamental to define different aspects of the speciation process and to delimit evolutionary distinct lineages (De Queiroz 2007; Dejaco et al. 2016). At the same way, it could be interesting to extend the study to the entire section *Unguiculata*, with the aim to better define the circumspection of the species *G. macrorrhizum*, its relationships with *G. dalmaticum*, and to solve the controversial case of *G. kikianum* (Aedo 2017; Wagensommer and Venanzoni 2021).

Declarations

ACKNOWLEDGMENTS

We are sincerely grateful to Prof. Hovirag Lancioni for her expertise, suggestions and contribute offered to this article.

COMPETING INTERESTS

The authors have no competing interests to declare that are relevant to the content of this article.

AUTHORS CONTRIBUTIONS

MC conceived the study, performed the molecular and cytogenetic analyses, wrote the draft of the manuscript. IC performed phylogenetic analysis and critically revised the manuscript. The authors read and approved the final manuscript.

References

1. Aedo C (2017) Taxonomic Revision of Geranium Sect. Ruberta and Unguiculata (Geraniaceae). Ann Mo Bot Gard 102:409–465. https://doi.org/10.3417/D-16-00016A
2. Aedo C (2018) Ajustes nomenclaturales en Geranium L. (Geraniaceae). Acta Bot Malacitana 42:297–298. https://doi.org/10.24310/abm.v42i2.2750
3. Baltisberger M (1984) Zytologische Untersuchungen an einigen Pflanzen aus Albanien. Ber Geobot Inst ETH, Stiftung Rubel, Zürich 51:63-77
4. Baltisberger M (1991) Cytological investigations of some Greek plants. Fl. Medit. 1:157-173
5. Baltisberger M, Baltisberger E (1995) Cytological data of Albanian plants. Candollea 50:457-493. https://doi.org/10.5169/SEALS-879478

6. Ćavar Zeljković S, Siljak-Yakovlev S, Tan K, Maksimović M (2020) Chemical composition and antioxidant activity of Geranium macrorrhizum in relation to ploidy level and environmental conditions. Plant Syst Evol 306:18. https://doi.org/10.1007/s00606-020-01649-9
7. CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106:12794–12797. https://doi.org/10.1073/pnas.0905845106
8. Costion C, Ford A, Cross H, Crayn D, Harrington M, Lowe A (2011) Plant DNA barcodes can accurately estimate species richness in poorly known floras. PLoS ONE 6:e26841. doi:10.1371/journal.pone.0026841
9. Cuènod P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW (2002) Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB and matK DNA sequences. Am J Bot 89:132-144
10. De Castro O, Del Guacchio E, Di Iorio E, Di Maio A, Di Martino L, Bartolucci F, Conti F (2020) Barcoding helps threatened species: the case of Iris marsica (Iridaceae) from the protected areas of the Abruzzo (Central Italy). Plant Biosyst 154:961–972. https://doi.org/10.1080/11263504.2020.1762786
11. De Queiroz K (2007) Species Concepts and Species Delimitation. Syst Biol 56:879–886. https://doi.org/10.1080/10635150701701083
12. Dejaco T, Gassner M, Arthofer W, Schlick-Steiner BC, Steiner FM (2016) Taxonomist's Nightmare … Evolutionist's Delight : An Integrative Approach Resolves Species Limits in Jumping Bristletails Despite Widespread Hybridization and Parthenogenesis. Syst Biol 65:947–974. https://doi.org/10.1093/sysbio/syw003
13. Del Guacchio E (2002) Note floristiche per la Campania. Delpinoa 44:75-80
14. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797
15. Falistocco E (2019) Chromosome investigations in annual Medicago species (Fabaceae) with emphasis on the origin of the polyploid Medicago rugosa and M. scutellata. Plant Biosyst 153:235–241. https://doi.org/10.1080/11263504.2018.1462864
16. Falistocco E (2023) Comparative cytogenetic study of Plantago species. Plant Biosyst 157:1078–1084. https://doi.org/10.1080/11263504.2023.2242335
17. Federici S, Galimberti A, Bartolucci F, Bruni I, De Mattia F, Cortis P, Labra M (2013) DNA barcoding to analyse taxonomically complex groups in plants: the case of Thymus (Lamiaceae): DNA Barcoding Of Genus Thymus. Bot J Linn Soc 171:687–699. https://doi.org/10.1111/boj.12034
18. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791
19. Fiorini Mazzanti E (1869) Cenno sulla vegetazione della caduta delle Marmore in una rapida escursione di luglio. Atti dell'Accademia Pontificia de' nuovi Lincei 22:143-144
20. Garcia S, Kovařík A, Leitch AR, Garnatje T (2017) Cytogenetic features of rRNA genes across land plants: analysis of the Plant rDNA database. Plant J 89:1020–1030. https://doi.org/10.1111/tpj.13442

21. Garnett ST, Christidis L (2017) Taxonomy anarchy hampers conservation. Nature 546:25–27. https://doi.org/10.1038/546025a
22. Garrido-Ramos M (2017) Satellite DNA: An Evolving Topic. Genes 8:230. https://doi.org/10.3390/genes8090230
23. Gauger W (1937) Ergebnisse einer zytologischen Untersuchung der Familie der Geraniaceae, I. - Pianta 26:529-531
24. Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucl Acids Res 7:1869–1885. https://doi.org/10.1093/nar/7.7.1869
25. Hall T (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Sympos Ser 41:95-98
26. Harborne JB, Williams CA (2002) Phytochemistry of the genus Geranium. In: Lis-Balchin M (ed) Geranium and Pelargonium. The genera Geranium and Pelargonium. Taylor & Francis, London and New York, pp 20-29
27. Hawke R (2004) Plant Evaluation Notes. Hardy Geraniums for Northern Gardens. Chicago Botanic Garden 22
28. Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313–321. https://doi.org/10.1098/rspb.2002.2218
29. Hollingsworth PM, Graham SW, Little DP (2011) Choosing and Using a Plant DNA Barcode. PLoS ONE 6:e19254. https://doi.org/10.1371/journal.pone.0019254
30. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120. https://doi.org/10.1007/BF01731581
31. Koblmüller S (2023) DNA Barcodes for Evolution and Biodiversity. Diversity 15:1003. https://doi.org/10.3390/d15091003
32. Kress WJ, Erickson DL (2007) A Two-Locus Global DNA Barcode for Land Plants: The Coding rbcL Gene Complements the Non-Coding trnH-psbA Spacer Region. PLoS ONE 2:e508. https://doi.org/10.1371/journal.pone.0000508
33. IUCN (2019) Guidelines for using the IUCN Red List categories and criteria. Version 14. Prepared by the Standards and Petitions Committee, Cambridge UK. https://www.iucnredlist.org/resources/redlistguidelines. Accessed 18 January 2021
34. Li Y, Wen J, Ren Y, Zhang J (2019) From seven to three: Integrative species delimitation supports major reduction in species number in Rhodiola section Trifida (Crassulaceae) on the Qinghai-Tibetan Plateau. Taxon 68:268–279. https://doi.org/10.1002/tax.12052
35. Mascagni F, Barghini E, Ceccarelli M, Baldoni L, Trapero C, Díez CM, Natali L, Cavallini A, Giordani T (2022) The Singular Evolution of Olea Genome Structure. Front Plant Sci 13:869048. https://doi.org/10.3389/fpls.2022.869048

36. Park S, Ruhlman TA, Weng M-L, Hajrah NH, Sabir JSM, Jansen RK (2017) Contrasting Patterns of Nucleotide Substitution Rates Provide Insight into Dynamic Evolution of Plastid and Mitochondrial Genomes of Geranium. Genome Biol Evol 9:1766–1780. https://doi.org/10.1093/gbe/evx124
37. Petrova A, Stanimirova P (2002) Reports 1288-1294 In: Kamari G, Blanché C and Garbari F (eds) Mediterranean chromosome numbers reports - 12. Fl Medit 12: 443-486
38. Petrova A, Stanimirova P (2003) Karyological study of some Geranium (Geraniaceae) species growing in Bulgaria. Bocconea 16: 675–682
39. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. https://doi.org/10.1093/oxfordjournals.molbev.a040454
40. Salerno G (2004) Segnalazioni floristiche Italiane: 1127. Informatore Botanico Italiano 36:89
41. Sanguinetti P, Sebastiani A (1837) Centuriae tres Prodromo florae Romanae addendae. Ex Typographia Contedini, Romae
42. Stoeva T (2002) Cultivation and harvesting of Geranium macrorrhizum and Geranium sanguineum for medicinal use in Bulgaria. In Lis-Balchin M (ed) Geranium and Pelargonium. The genera Geranium and Pelargonium. Taylor & Francis, London and New York, pp 30–35
43. Strid A, Franzén R (1981) In Löve A (ed) IOPB chromosome number reports LXXIII. Taxon 30:829-861
44. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. https://doi.org/10.1093/oxfordjournals.molbev.a040023
45. Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120
46. Tan K, Siljak-Yakovlev S, Vold G (2011) Geranium kikianum sp. nov. (Geraniaceae) from the southern Peloponnese, Greece. Nord J Bot 29:1–5. https://doi.org/10.1111/j.1756-1051.2010.01028.x
47. Thakur J, Packiaraj J, Henikoff S (2021) Sequence, Chromatin and Evolution of Satellite DNA. Int J Mol Sci 22:4309. https://doi.org/10.3390/ijms22094309
48. Tiburtini M, Astuti G, Bartolucci F, Casazza G, Varaldo L, De Luca D, Bottigliero MV, Bacchetta G, Porceddu M, Domina G, Orsenigo S, Peruzzi L (2022) Integrative Taxonomy of Armeria arenaria (Plumbaginaceae), with a Special Focus on the Putative Subspecies Endemic to the Apennines. Biology 11:1060. https://doi.org/10.3390/biology11071060
49. Van Loon JC (1984) Chromosome numbers in Geranium from Europe. I. The perennial species. Proc. Kon. Ned. Akad. Wetensch. Ser. C. 87:263-277
50. Venanzoni R (2017) Il geranio odoroso delle Marmore. In: Venanzoni, R. (Ed.) Il Giardino Botanico della Cascata delle Marmore – Biodiversità al lavoro. Comune di Terni, Fondazione Cassa di Risparmio di Terni e Narni
51. Wagensommer RP, Venanzoni R (2021) Geranium lucarinii sp. nov. and re-evaluation of G. kikianum (Geraniaceae). Phytotaxa 489:252–262. https://doi.org/10.11646/phytotaxa.489.3.2

52. Yeo PF (2004) The morphology and affinities of Geranium sections Lucida and Unguiculata. Bot J Linn Soc 144:409–429. https://doi.org/10.1111/j.1095-8339.2003.00258.x
53. Zhang Y, Cheng C, Li J, Yang S, Wang Y, Li Z, Chen J, Lou Q (2015) Chromosomal structures and repetitive sequences divergence in Cucumis species revealed by comparative cytogenetic mapping. BMC Genomics 16:730. https://doi.org/10.1186/s12864-015-1877-6
54. Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7: 203-214

Figures

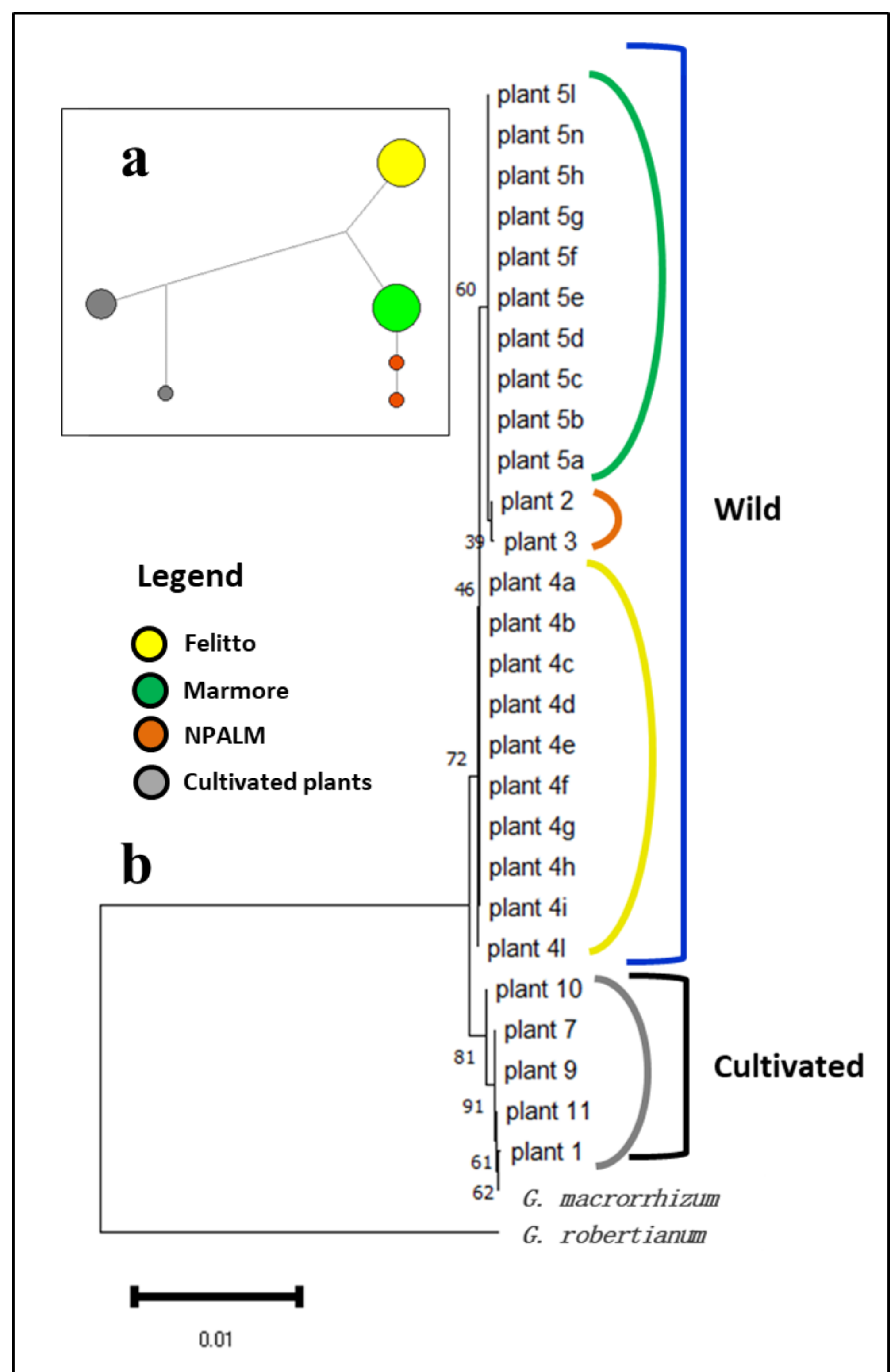

Figure 1

Legend not included with this version.

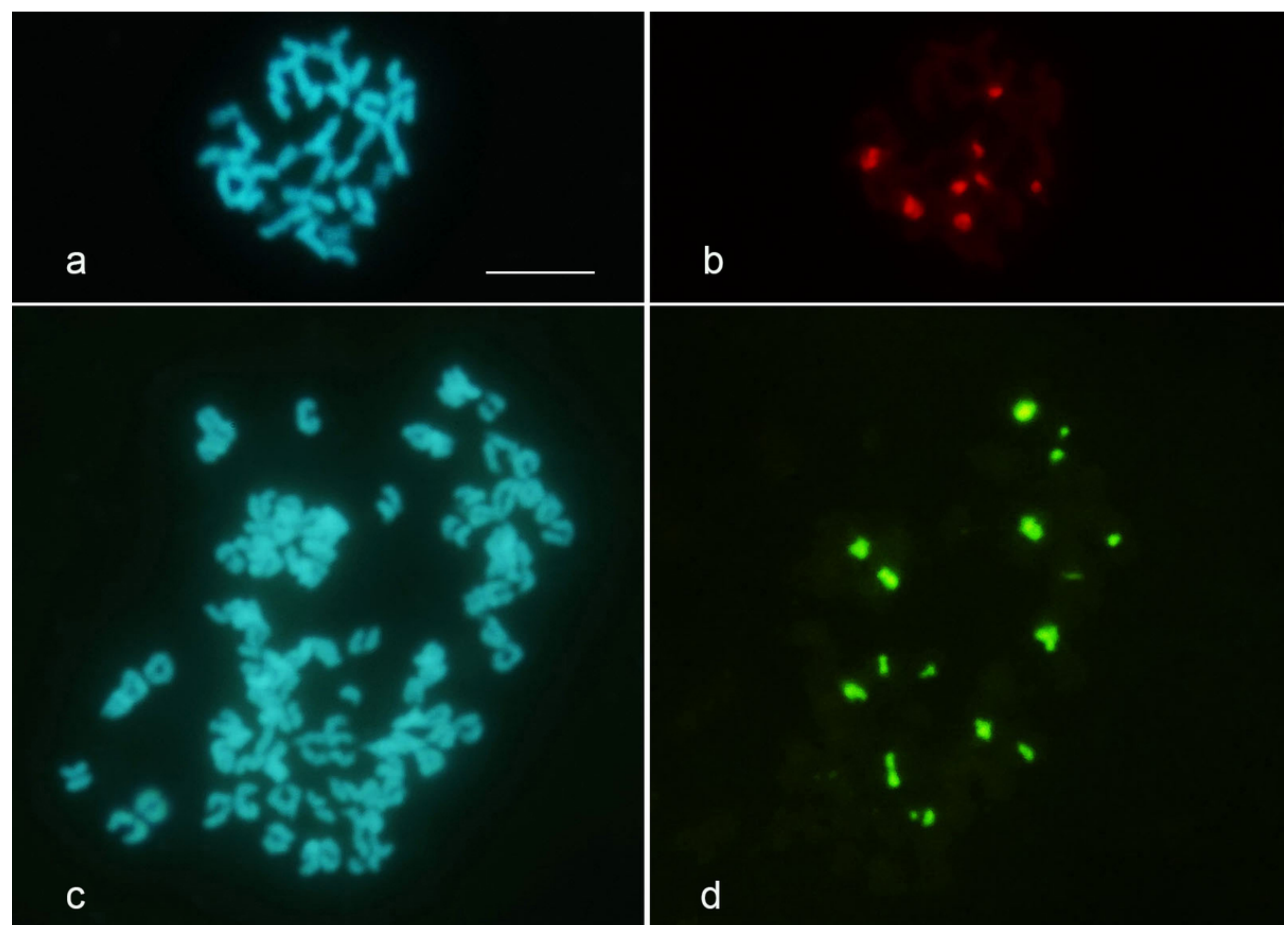

Figure 2

Legend not included with this version.

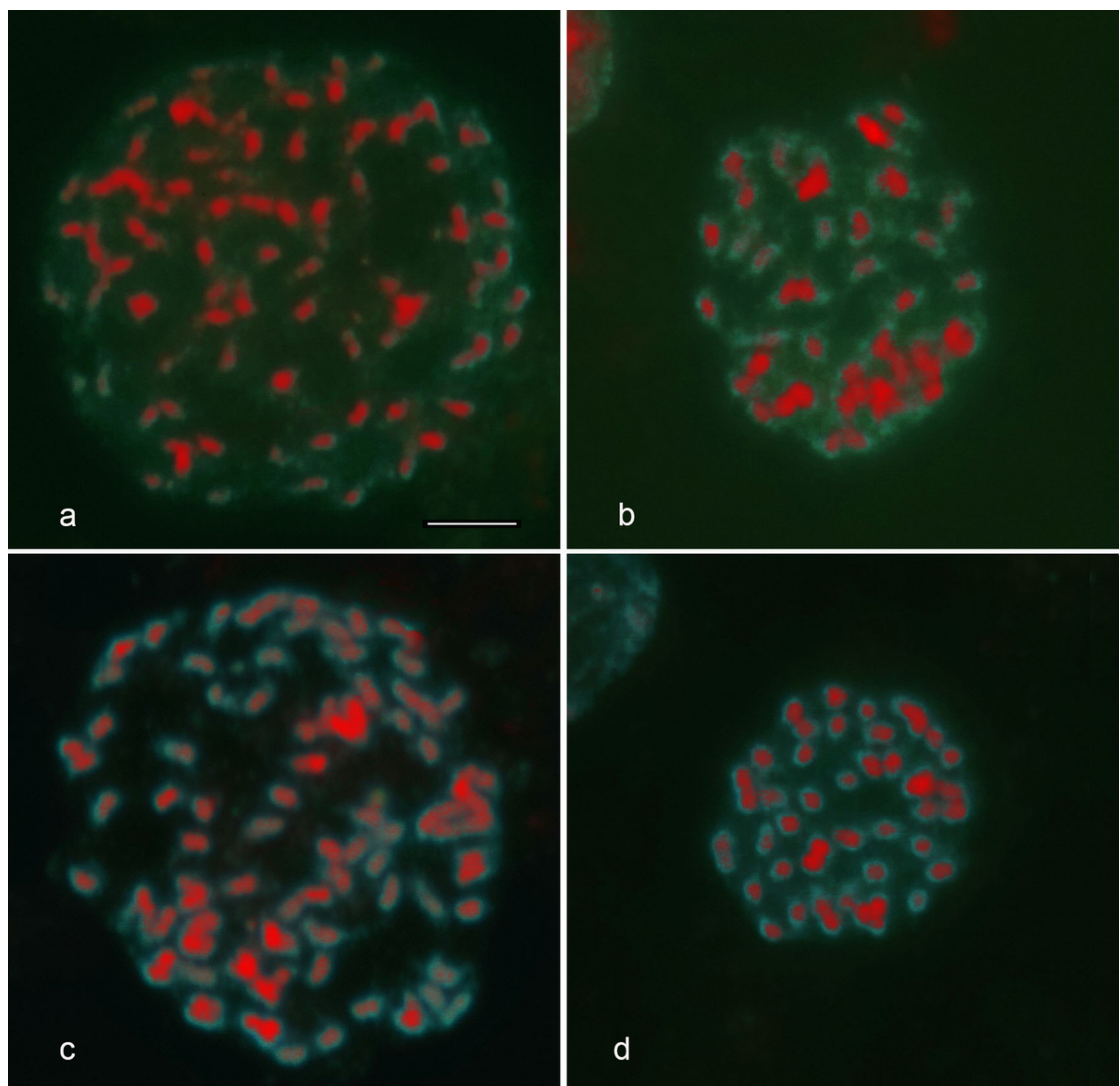

Figure 3

Legend not included with this version.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.pdf
- SupplementaryTables.pdf