

An integrative taxonomic revision of the *Chaerophyllum hirsutum* complex (Apiaceae) using morphological and molecular markers

Thomas Reinhart^{1,*}, Lucile Guillon^{2,*}, Thomas Begoc^{1,*}, Pauline Chapotin², Jean-Pierre Reduron³, Armando Espinosa Prieto¹, Laurent Hardion¹

¹ University of Strasbourg, Laboratoire Image Ville Environnement, UMR 7362 CNRS, ENGEES, Strasbourg, France

² University of Strasbourg, Faculty of Life Sciences, Strasbourg, France

³ Via Apia, Mulhouse, France

* These authors contributed equally to this work.

Corresponding author: Laurent Hardion (hardion@unistra.fr)

Academic editor: Lorenzo Lazzaro ♦ Received 7 April 2024 ♦ Accepted 20 August 2024 ♦ Published 23 October 2024

Abstract

Background and aims – *Chaerophyllum hirsutum* represents a complex of taxa with varying treatments and ranks across floras. Using both morphometric and molecular markers, we assessed the robustness of *C. hirsutum*, *C. elegans*, *C. villarsii*, and *C. villarsii* var. *cicutariiforme*.

Material and methods – Ten morphometric variables and two ratios were calculated. Based on the sequencing of six plastomes, the *rps16* intron was selected as the more variable region and sequenced on a broader sampling. Additionally, we also sequenced the nrDNA internal transcribed spacer 2 (ITS2) using Illumina technology to obtain intra-individual allelic diversity.

Key results – Morphologically, the most easily differentiated taxon was *C. elegans*, especially using the number of subterminal umbels. The distinction between *C. hirsutum* and *C. villarsii* was rather clinal, but is mainly based on the degree of carpophore division. Finally, *C. villarsii* var. *cicutariiforme* was less easily distinguishable from the three others, but partly using the carpophore length and the total length of basal leaf blade. The cpDNA *rps16* clearly distinguished *C. elegans* from the three other taxa of the complex, which rather showed a geographical pattern of cpDNA diversity. The nrDNA ITS2 partially distinguished *C. villarsii* from the other taxa, without distinction of *C. elegans*.

Conclusions – The present study supports the species differentiation of *C. elegans* based on both morphology and chloroplast genome. Furthermore, *C. villarsii* var. *villarsii* and *C. villarsii* var. *cicutariiforme* could potentially be recognized as distinct varieties within *C. hirsutum*. This will need to be confirmed by future studies using a larger sampling size and more comprehensive markers, covering a broader portion of the nuclear genome.

Keywords

cpDNA *rps16* intron, internal transcribed spacer 2, plant systematics, plastome, morphometry

INTRODUCTION

Chaerophyllum L., tribe Scandiceae Spreng., subtribe Scandicinae Tausch, is the most widespread genus of the Apiaceae Lindl. (Spalik et al. 2010), comprising 40–50 species primarily native to the northern hemisphere,

but also occurring in the southern hemisphere since the incorporation of the genus *Oreomyrrhis* Endl. (Chung 2007). Within the *Chaerophyllum* genus, the *Chaerophyllum hirsutum* L. complex, characterised by its ciliate petals, poses a taxonomic challenge in terms of both morphological differentiation and evolutionary history.

Copyright Thomas Reinhart, Lucile Guillon, Thomas Begoc, Pauline Chapotin, Jean-Pierre Reduron, Armando Espinosa Prieto, Laurent Hardion. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Plant Ecology and Evolution is published by Meise Botanic Garden and the Royal Botanical Society of Belgium.

Its main taxa include *C. hirsutum* L., *C. elegans* Gaudin, and *C. villarsii* W.D.J.Koch, with various taxonomic treatments across floras regarding rank, synonymy, and additional infraspecific taxa. While several infraspecific taxa have been overlooked, the detailed monograph of the Apiaceae from France, *Ombellifères de France* (Reduron 2007), mentioned the following ones: *C. hirsutum* var. *glabrum* (Lam.) Briq. without bristles, *C. hirsutum* var. *roseum* W.D.J.Koch with pink flowers, and *C. hirsutum* var. *calabricum* (Guss. ex DC.) Pöhl. with larger leaves, as well as *C. villarsii* var. *glabrum* Beauverd without bristles, *C. villarsii* var. *alpestre* Jord. with taller fruits, and *C. villarsii* var. *cicutariiforme* Beauverd with intermediate forms. In the present study, we focused on the taxa considered in the French flora, *Flora Gallica* (Tison and de Foucault 2014): (i) *C. hirsutum*, (ii) *C. elegans*, (iii) *C. villarsii* var. *villarsii*, and (iv) *C. villarsii* var. *cicutariiforme*.

Morphologically, floras differentiate *C. hirsutum* from the other two species by its carpophore, which is less than 50% divided (vs more than 50%) and flame-shaped (flattened with a flame-like appearance). On the other hand, *C. elegans* is mainly distinguished from the other two by its opposite to whorled inflorescences. *Chaerophyllum villarsii* var. *cicutariiforme* also blurs these distinctions with its *hirsutum*-like leaf shape and ecology, and *villarsii*-like carpophore. Ecologically, *C. hirsutum* is a species found in mountain-lowland forests

and humid environments, *C. villarsii* is more associated with highland mesophile meadows, and *C. elegans* with highland humid meadows. Geographically, the three species exhibit a nested distribution, with *C. hirsutum* common in Western and Central European Mountains (Fig. 1), *C. villarsii* abundant mainly along the northern part of the Alps, and *C. elegans* rare in the northern and southwestern Alps. Additionally, these taxa are all perennial hemicryptophytes, mainly insect-pollinated, epizoochore, and diploids with $2n = 22$ chromosomes.

Phylogenetically, the *C. hirsutum* complex represents the second basal divergence of the genus (Piwczyński et al. 2015). To date, the only molecular study addressing the taxonomic delineation of this group focused on the entire genus using nuclear ribosomal DNA internal transcribed spacers (ITS). The authors observed slight divergence between *C. villarsii* and *C. hirsutum*, and a lack of divergence for *C. elegans* and *C. hirsutum* var. *calabricum* both assigned to *C. hirsutum* (Piwczyński et al. 2015). The authors suggested further molecular investigations based on a broader sampling.

The aim of this study is to test the robustness of taxa within the *C. hirsutum* complex using both morphometric and molecular markers, as a standard in integrative taxonomy (Hardion et al. 2020; Castro et al. 2023). The genetic markers target the nrDNA ITS2, which has been previously studied (Piwczyński et al. 2015), as

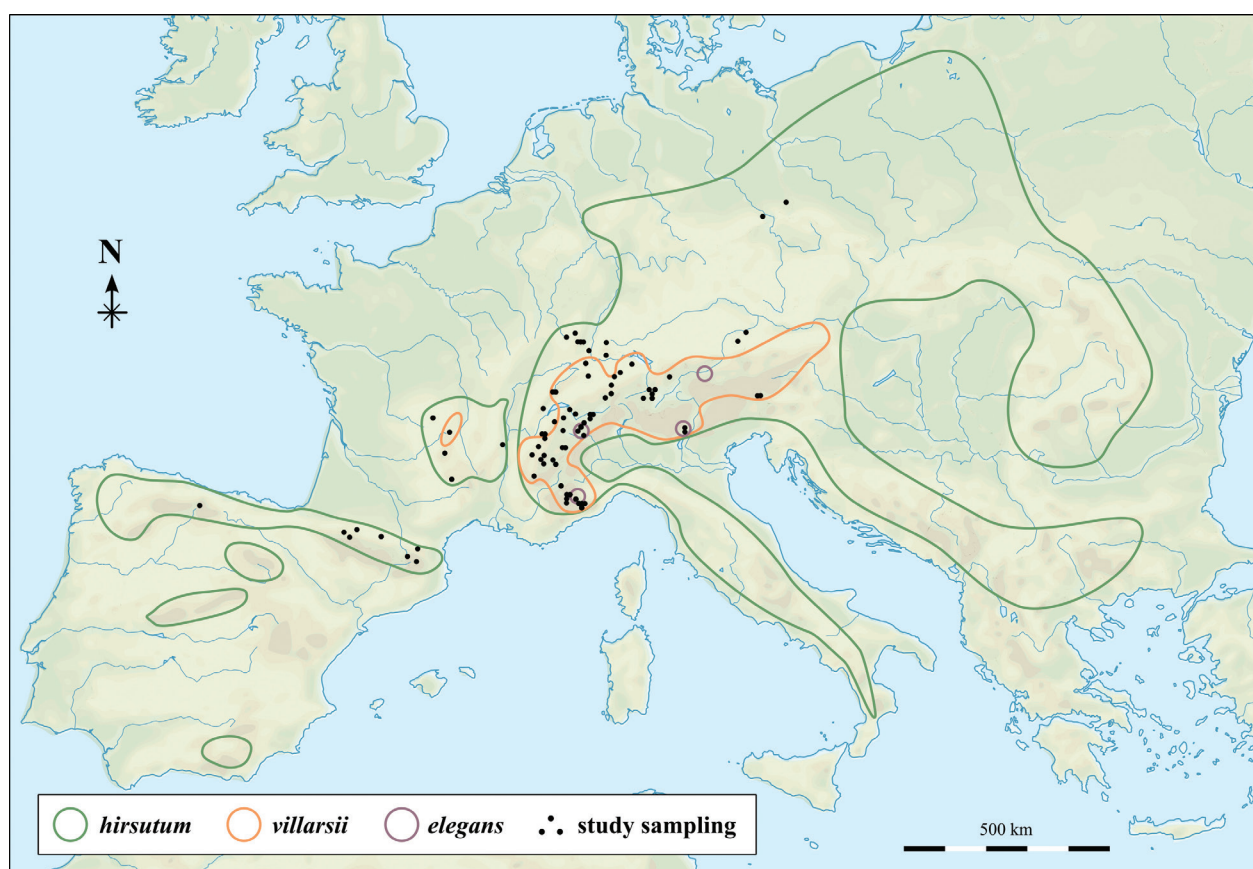


Figure 1. Geographic distribution of taxa within the *C. hirsutum* complex (distribution areas estimated from Meusel et al. 1965, and www.gbif.org).

well as a newly investigated chloroplast DNA (cpDNA) locus selected for its SNP variation based on plastome sequencing of a subset of the sample (Dodsworth 2015).

MATERIAL AND METHODS

Plant sampling

The plant material primarily originates from herbarium collections at the University of Strasbourg (STR), the National Museum of Natural History in Paris (P), and the University of Basel (BAS), as well as from the personal collection of Jean-Pierre Reduron (Suppl. material 1). We gathered 83 specimens of *C. hirsutum* (including five of *C. hirsutum* var. *calabricum*), 36 of *C. elegans*, 59 of *C. villarsii* var. *villarsii*, and 21 of *C. villarsii* var. *cicutariiforme*. However, due to variations in collecting dates, DNA preservation, and the availability of plant parts on specimens (such as basal or stem leaves, inflorescences, or infructescence), morphometric and molecular analyses are conducted using partially overlapping specimen lists. Respectively for morphometry, cpDNA *rps16* and nrDNA ITS2, we analysed 72, 13, and 15 specimens of *C. hirsutum*,

25, 8, and 13 of *C. elegans*, 50, 10, and 12 of *C. villarsii* var. *villarsii*, and 16, 6, and 6 of *C. villarsii* var. *cicutariiforme*.

Morphometric analysis

Ten quantitative variables were measured three times per specimen, whenever possible, and averaged (Fig. 2). These morphometric characters were selected based on a compilation of determination keys for the *C. hirsutum* complex in European Floras (referenced in Reduron 2007). Boxplots and adjusted R^2 were generated on the entire dataset using the `boxplot()` and `lm()` functions from the `stats` and `graphics` base packages in R v.4.3.1 (R Core Team 2024). The dataset contained numerous missing values (NA) due to the absence of various plant parts (e.g. carpophore) on herbarium specimens. Prior to further analyses, samples with more than 50% of missing values were excluded to reduce their occurrence to less than 25% across the entire dataset. Then, a Bayesian principal component analysis (bPCA) was conducted to provide a descriptive analysis of this reduced dataset despite some remaining missing values. This was achieved thanks to the “missing value estimation” parameter included in the `bpca()` function from the R package `pcaMethods` v.1.92 (Oba et al. 2003; Stacklies et al. 2007).

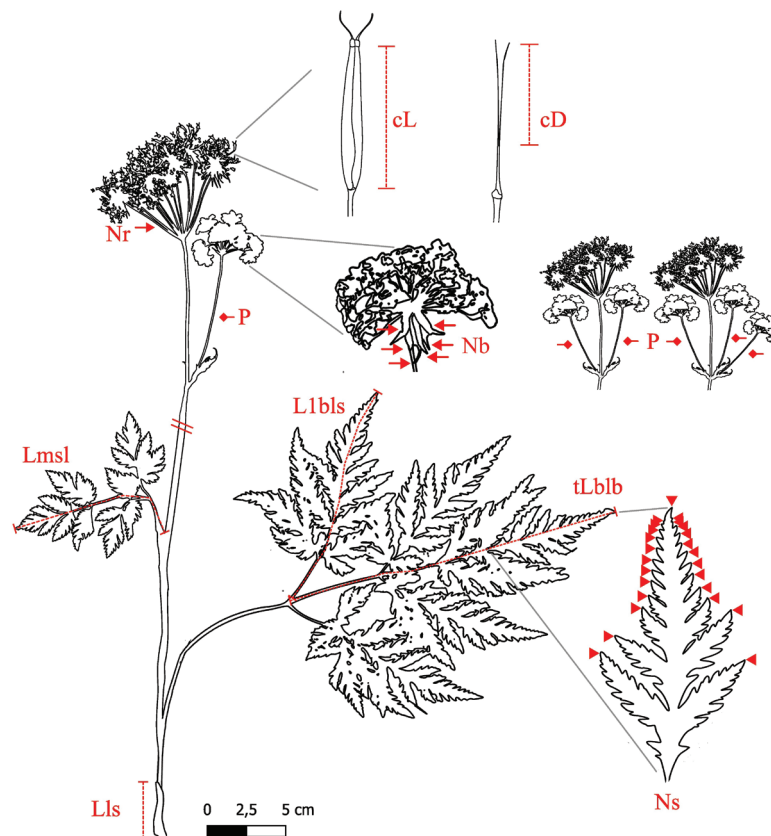


Figure 2. Description of the 10 morphometric quantitative variables measured: length of median stem leaf (Lmsl), length of leaf sheath (Lls), length of first basal leaf segment (L1bls), total length of basal leaf blade (tLblb), number of segments within the last-order division on the basal leaf (Ns), carpophore length (cL), length of carpophore division (cD), number of subterminal umbels (P), number of rays (Nr), and number of involucre bracteoles (Nb). Two ratios were also calculated, $tLblb / L1bls$ and cD / cL .

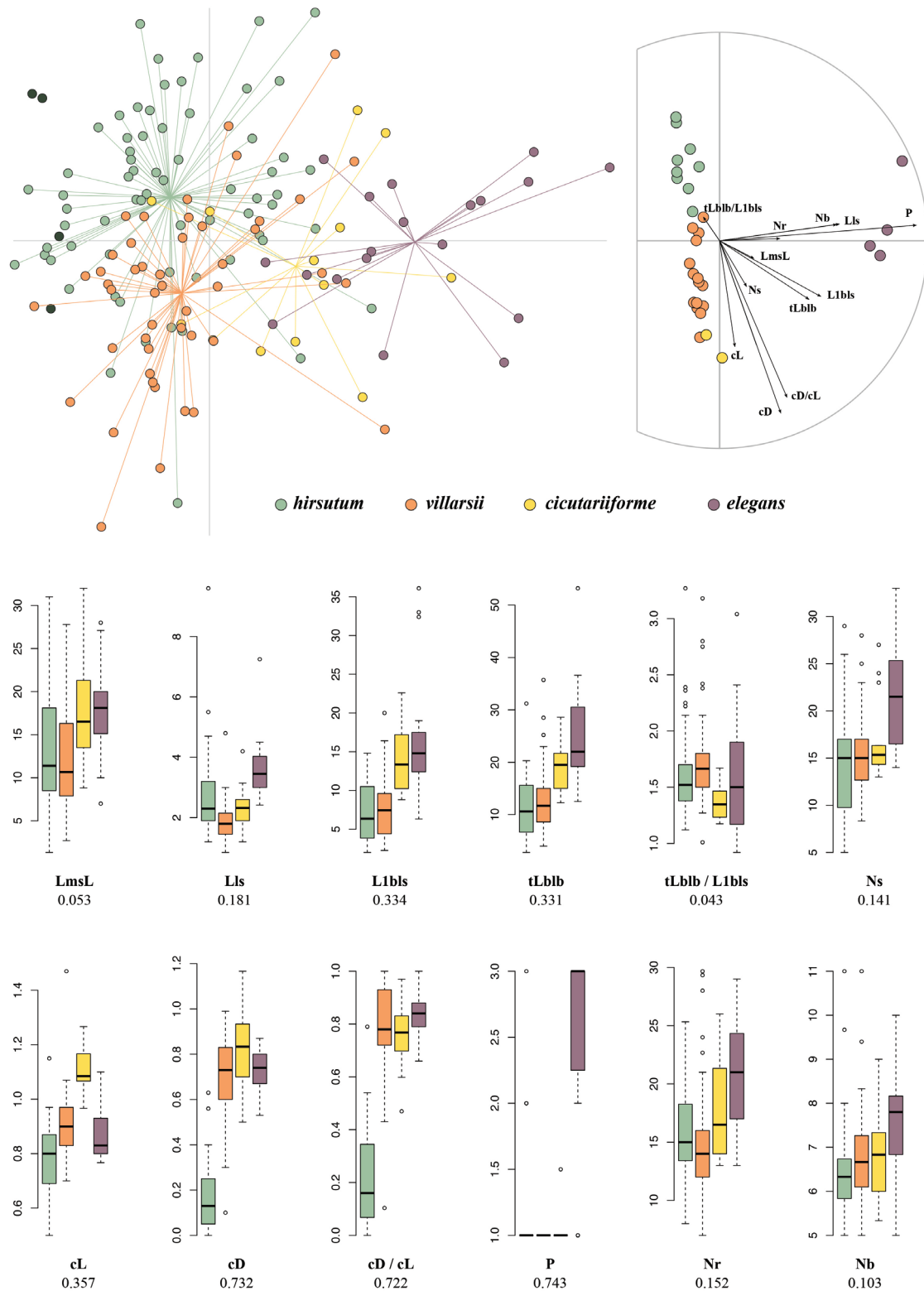


Figure 3. Top left, Bayesian principal component analysis (bPCA) with missing value estimation on samples with less than 50% of missing values, resulting to a dataset with less than 25% of NA ($n = 149$). Two principal components were imposed, with a R^2 (the importance of component) of 0.337 for the first (horizontal) axis and 0.135 for the second (vertical) axis. Top right, discriminant analysis (DA) on samples without NA values ($n = 30$), with two of the three axes obtained representing inertia of 0.988 and 0.764 (and the third, not represented here, 0.473); Bottom, boxplots for morphometrical values per taxa. LmsL, length of median stem leaf (cm); Lls, length of leaf sheath (cm); L1bls, length of first basal leaf segment (cm); tLblb, total length of basal leaf blade (cm); cL, carpophore length (cm); cD, length of carpophore division (cm); Ns, number of segments within the last-order division on the basal leaf; P, number of subterminal umbels; Nr, number of umbel rays; Nb, number of involucellar bracteoles. Dark green dots represent samples of *C. hirsutum* var. *calabricum* considered as *C. hirsutum* var. *hirsutum* samples in the present study.

Molecular markers

The DNA extractions followed a CTAB protocol (Doyle and Doyle 1987) on 40 mg of mechanically ground dried leaves. The DNA concentrations were measured using the QuBit dsDNA HS Assay kit (Invitrogen by Thermo Fisher Scientific) and diluted to approximately 10 ng/mL. The nrDNA ITS2 spacer was amplified using plant-specific primers forward u3 and reverse p4 described in Cheng et al. (2016). The PCRs were performed in a final volume of 25 μ L following the instruction of GoTaq G2 Master Mix (Promega, Madison, USA), with a thermocycler program of 94°C for 5 min, 30 cycles of 94°C for 45 s, 50°C for 45 s, 72°C for 45 s, and a final extension of 94°C for 5 min. A next generation sequencing method was chosen to include the infra-individual genetic variation of nrDNA (e.g. to detect putative hybrids). The PCR products were pooled thanks to 5' primer libraries and sequenced on an Illumina MiSeq platform using 2 \times 300 bp paired-end reads generating ca 60,000 paired-end reads. For each amplicon, we retained the ribotypes represented by at least 20% of the reads obtained for the amplicon. The ITS2 ribotype network was manually constructed considering indels as a fifth state.

To select the most resolutive cpDNA region, the plastomes of six samples (CE08.02; CH23.11; CV24.15; CV22.11; CH13.01; CE001), two per species, were sequenced in genome skimming using an Illumina system generating 5 M reads per sample in 2 \times 150 bp on a size selection of DNA fragments around 450 bp. The mapping of read pairs on a reference plastome was performed with the 'BBMap' algorithm in Geneious Prime 2023 (Dotmatics, Boston, USA) based on the reference plastome of *Anthriscus sylvestris* (L.) Hoffm. (MT561042). The plastome alignment was generated in MAFFT v.7.490 (Katoh and Standley 2013), and we then manually selected the *rps16* intron for its short size and its higher number of SNPs through the alignment. In addition, this marker was previously used to resolve the phylogeny of Apiaceae, but not on *Chaerophyllum* samples (Downie et al. 2000, 2001). We manually designed and tested four primer pairs in vitro, and the best PCR products were obtained using the forward 'cpChaeroF4' [5' TTTTCTCCTCGTACGGCTCG 3'] and the reverse 'cpChaeroR4' [5' ATGAAGGTGCTCTTGACCCG 3']. The PCRs were performed in a final volume of 25 μ L following the instruction of GoTaq G2 Master Mix, with thermocycler program of 94°C for 5 min, 30 cycles of 94°C for 30 s, 52°C for 30 s, 72°C for 30 s, with a final extension of 72°C for 5 min. The PCR products were sequenced on a Sanger system using the reverse primer.

Phylogenetic reconstructions were generated using MUSCLE algorithm for sequences alignment (Edgar 2004) and the PAUP algorithm for maximum parsimony using heuristic search and 1,000 bootstrapping replications, as implemented in Geneious Prime 2023. For the cpDNA *rps16* phylogeny, *Myrrhis odorata* (L.) Scop. was used as an outgroup because it is placed in the tribe

Scandiceae Spreng. and subtribe Scandicinae Tausch, as *Chaerophyllum*. The haplotype networks were generated using functions from the R package pegas v.1.3 (Paradis 2010).

RESULTS

Morphological distinction

The four taxa tested in this study (*C. hirsutum*, *C. elegans*, *C. villarsii*, and *C. villarsii* var. *cicutariiforme*) could be partially differentiated based on bPCA even with 25% missing values (Fig. 3). The most easily differentiated taxon was *C. elegans*, on the first axis of the bPCA and the DA. The distinction between *C. hirsutum* and *C. villarsii* is rather clinal on the second axes of the two analyses. Finally, *C. villarsii* var. *cicutariiforme* is less easily distinguishable, with often intermediate values between the three other taxa.

Each of the 12 morphometric variables exhibit overlapping values among the four taxa (Fig. 4), except the number of subterminal umbels (P, $R^2 = 0.743$) which clearly distinguishes *C. elegans*, in addition to some other variables such as length of leaf sheath (Ls) and the total length of basal leaf blade (tLbLb). Then, the second-best discriminatory variable was the carpophore division (cD, $R^2 = 0.732$), which clearly distinguishes *C. hirsutum*. The ratio between the carpophore division and carpophore length (cD/cL) is also a discriminatory variable, but it does not provide more information than cD. Finally,

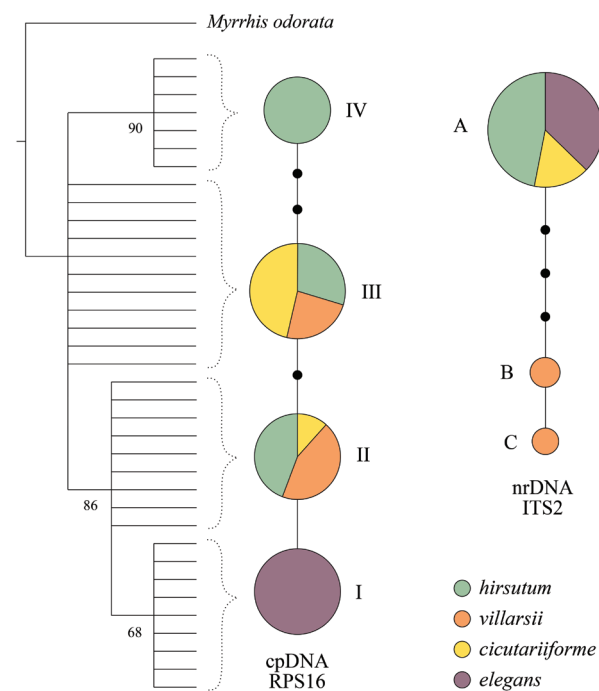


Figure 4. Left, maximum parsimony bootstrap tree (with bootstrap values) and haplotype network on cpDNA *rps16* intron ($n = 36$); right, haplotype network on nrDNA ITS2 ($n = 47$).

the carpophore length (cL, $R^2 = 0.357$), the length of first basal leaf segment (L1bls, $R^2 = 0.334$) and the total length of basal leaf blade (tLblb, $R^2 = 0.331$) partially distinguish *C. villarsii* var. *cicutariiforme* from the other taxa despite overlaps. The other variables show relatively poor resolutive power ($R^2 < 0.2$).

Phylogenetic analyses

The 43 samples successfully sequenced for the nrDNA ITS2 generated four ribotypes, divided into two groups separated by three substitutions. The ribotype A was found in 34 samples, representing *C. hirsutum*, *C. elegans*, and *C. villarsii* var. *cicutariiforme*. In contrast, the other group contains the ribotypes B and C associated with nine samples (one sample only had B) and mainly with *C. villarsii* ($n = 10$). One sample presented the association of ribotypes A and C (CV_20-08), and another showed the combination of the three ribotypes (CV_24-15). Without morphological data for these specimens, we chose to remove them from the analysis.

The 36 samples sequenced for the cpDNA *rps16* generated four haplotypes, without clear correspondence with the four ribotypes previously presented. Taxonomically, only *C. elegans* represented a monophyletic group with the clade I. The clade IV represented the first divergence of the group, and gathered only *C. hirsutum* samples, but only those from western Europe excluding the Alpes (Vosges, Pyrenees, and Massif Central in France). More than half of the sampling was gathered in the haplotypes II and III, combining samples of *C. elegans*, *C. hirsutum*, and *C. villarsii* var. *cicutariiforme*. Only one sample (CV_016), initially identified as *C. villarsii* var. *villarsii*, shared the haplotype I (*C. elegans*) and the ribotypes BC (*C. villarsii*), and this sample was morphologically more related to *C. villarsii* var. *villarsii* than to *C. villarsii* var. *cicutariiforme*.

DISCUSSION

The present study provides insights into the differentiation of taxa within the *C. hirsutum* complex, which vary from species to varieties. The common bias linked to the analysis of herbarium specimens, such as incomplete material, was predominant. Only a minority of specimens could inform all morphometric variables, leading to a significant amount of missing data. However, the present study has identified the most reliable characteristics for distinguishing taxa within this complex, thereby encouraging further investigations into this intricate group. The taxonomic identity of only three samples was not supported by a consensus of molecular markers and could correspond to methodological errors or putative hybrids (CV016, CV20-08, CV24-15; see Suppl. material 1).

Chaerophyllum elegans is the most differentiated taxon based on our genetic and morphometric data.

Its verticillate umbels combined with its larger leaves (sheaths and limbs) provide reliable information for its determination, with weak overlaps with other taxa. This taxon is also the only one to be monophyletic through the cpDNA *rps16*. This species was previously described as a variety and a subspecies of *C. hirsutum*, but several authors have proposed the species status based on its morphological differentiation, and its restricted and ecologically characteristic distribution. Indeed, this species has been documented across approximately ten locations as delineated by Wörz (1988), mainly in alpine and subalpine zones. Relating to phytogeographical observations from previous authors, Wörz (1988) hypothesised the preglacial origin of this species with some other ones, which could have persisted in alpine ice-free areas during the last glaciation. This theory offers a plausible explanation for the limited distribution of *C. elegans*, suggesting that its stations were established before this period, and then partitioned. Our phylogenetic hypothesis supports the monophyly of *C. elegans*, and the paraphyly of the other taxa. This phylogenetic pattern rather supports the membership of *C. elegans* in the progeny of the common ancestor of the *C. hirsutum* complex. As described from a neo-endemic origin, *C. elegans* may be the result of a peripatric speciation in an isolated alpine niche during the last glaciation. This recent divergence from a large entity is also supported by the incomplete segregation of nrDNA ITS2 ribotypes among the taxa.

The distinction of *C. villarsii* from *C. hirsutum* is less clear, and the two taxa seem to form a morphological cline. Their only distinctive criterion is the longer division of the carpophore for *C. villarsii*, which is relatively challenging to observe in the field. In addition, Beauverd (1902) described the compressed form (“flammuliform”) of the carpophore of *C. hirsutum*, a character indeed observed but not measured during our investigations. Some Floras mentioned the shorter stem leaf sheath of *C. villarsii* (e.g. (0.4–)0.6–0.8(–1.2) mm in Reduron 2007) as a distinctive character from other taxa. Our estimated size was 1.77 ± 0.78 mm, but only measured on the basal leaves. Another frequently mentioned characteristic in Floras is the equilateral triangular leaf shape of *C. hirsutum*, attributed to the basal segments nearly as wide as the rest of the leaf, contrasting with the narrower leaves (with smaller basal segments) of *C. villarsii*. However, our measurements indicate a similar length-to-width ratio between the two taxa. We also observed that the shape of the leaf segment tips was thinner in *C. villarsii* and sharper in *C. hirsutum* (Suppl. material 2), but this character was technically difficult to measure. The interesting but incomplete distinction of the individuals of *C. villarsii* from the remaining sampling based on nrDNA ITS2 was shown thanks to the use of Illumina technology for amplicon sequencing. This difference might not have been noticed with Sanger sequencing, due to the bias of nucleotide ambiguities. This incomplete differentiation of ITS2 ribotypes between taxa could be the result of an

incomplete lineage sorting, or a relatively ancient genetic differentiation (based on the four SNPs between A and B-C) followed by more recent events of hybridization. Inversely, the cpDNA *rps16* placed *C. villarsii* nested within the paraphyly of *C. hirsutum*. As for geographical distributions, *C. hirsutum* is the more extended taxon in the phylogeny of the taxonomic complex, with a specific clade (IV) showing a coherent geographical distribution in Western European massifs (see also spatial distribution of genetic data in Suppl. material 3). The gradient differentiation between *C. hirsutum* and *C. villarsii* is also found in their ecology, from a hygrophile and montane ecology for *C. hirsutum* to a more mesophile and subalpine ecology for *C. villarsii* (Reduron 2007). Wörz (1988) also considered their divergence as a more recent post-glacial. The differentiation between *C. hirsutum* and *C. villarsii* may be likely attributed to sympatric processes influenced by a divergence on ecological niches.

Traditionally attached to *C. villarsii*, *C. villarsii* var. *cicutariiforme* blurs the clinal distinction between *C. hirsutum* and *C. villarsii*. As represented in the morphometric bPCA, this taxon is historically regarded as intermediate between *C. hirsutum* and *C. villarsii* var. *villarsii*, displaying leaves like the former and fruit like the latter. The carpophore of *C. villarsii* var. *cicutariiforme* resembles that of *C. villarsii* var. *villarsii* (and *C. elegans*) in our morphometric dataset, while the leaf dimensions of *C. villarsii* var. *cicutariiforme* are rather closer to those of *C. elegans*. Regarding genetic data, the cpDNA haplotypes of *C. villarsii* var. *cicutariiforme* mainly correspond to the haplotypes of *C. villarsii* var. *villarsii* (and *C. hirsutum*), but the nrDNA ITS2 ribotypes are rather the same as *C. hirsutum* and *C. elegans*. Ecologically, some authors mentioned that this taxon is rather like the temperate *C. hirsutum*, which it seems to replace in the Mediterranean South Alpes (Reduron 2007). Our analyses did not clearly distinguish *C. villarsii* var. *cicutariiforme* from the rest of the complex, yet it underscores the need for further investigation, given its unique combination of characteristics.

Our sampling was partly sufficient to test the robustness of the four previous taxa, but the *C. hirsutum* complex also includes other putative entities. First, we did not analyse samples of the Italian *C. magellense* Ten., also considered as a subspecies of *C. hirsutum*. The only available DNA sequence for this species (GenBank accession [KJ956537](#), herbarium specimen E00040962; Piwczyński et al. 2015) targeted the nrDNA ITS. Including it in our ITS2 dataset connected it to ribotype A by three nucleotide differences (data not shown), partially confirming the genetic differentiation of *C. magellense* from the rest of the complex. Our morphometric analysis also included four specimens identified as *C. hirsutum* var. *calabricum*, an Apennine and Alpine taxon initially distinguished by De Candolle based on its indumentum and its poorly incised leaves. These samples represent the lowest values for every morphometric variable, explaining its position near *C. hirsutum*. Unfortunately, we did not obtain DNA

sequences for these samples. This taxon is represented by only one DNA sequence in international DNA databases (GenBank accession [KJ956578](#), herbarium specimen E00065526; Piwczyński et al. 2015), here again targeting the nrDNA ITS, and this sequence is identical to ribotype A in our ITS2 dataset. Assigned to *C. villarsii*, *C. villarsii* var. *alpestre* described based on its longer achenes and styles was absent from our sampling. Finally, we did not consider the infraspecific taxa *C. hirsutum* var. *roseum*, *C. villarsii* var. *glabrum*, and *C. hirsutum* var. *glabrum* based on their dark-pink flower and their lack of pubescence, respectively. In absence of biogeographic or even populational structuring of these characters (Reduron 2007), we hypothesised that their variations are plastic or not phylogenetically structured.

CONCLUSION

The present study provides further evidence to support taxa within and beyond *C. hirsutum*. Both morphometric analysis and the cpDNA genome clearly distinguish *C. elegans* from *C. hirsutum*. The differentiation between *C. hirsutum* and *C. villarsii* seems to correspond more to a morphological continuum, although the nrDNA ITS2 also partially distinguishes them. Based on our results, *C. villarsii* var. *villarsii* and *C. villarsii* var. *cicutariiforme* could be considered as varieties within *C. hirsutum*. This rank choice should be tested once again using more resolving markers covering a broader part of the nuclear genome. In addition, regarding the large infraspecific polymorphism of leaf characteristics described in Floras, and the importance of ecological distinction for these close taxa, it would be interesting to cultivate them in different conditions to observe the robustness of their morphological distinction, or their high plasticity along the temperate-alpine gradient.

ACKNOWLEDGEMENTS

The authors thank herbarium curators and botanical gardens, conservatories and societies that provided us with plant distribution data or material: Marion Martinez Martin, Céline Froissart, and Frédéric Tournay from the Herbarium and the Botanical Garden of the University of Strasbourg (STR), Aurélie Grall and Jurriaan de Vos from the Basel Herbarium (BAS), Pete Lowry from the Herbarium of Museum National d'Histoire Naturelle of Paris (P), Jérôme Hog and Julie Nguéfack from the Conservatoire Botanique d'Alsace-Lorraine, Emmanuelle Lehin from the Conservatoire Botanique National de Franche-Comté, Jean-Michel Genis from the Conservatoire Botanique National Alpin, Valerio Lazzeri from the Societa Botanica Italiana, Alfred Mayer from Rome University, Yvona Asbäck from Biodiversität Austria, and Jens Wesenberg from the Herbarium Senckenbergianum Frankfurt und Görlitz. We also thank

Pieter Verschuren for her help during wet lab work. This work was initiated and supported as a supervised project 'VegeLAB' of the master's track 'Plantes, Environnement, et Génie Ecologique' at the Faculty of Life Sciences of the University of Strasbourg.

REFERENCES

- Beauverd G (1902) Notes sur les caractères extérieurs du *Chaerophyllum hirsutum* L. Bulletin de l'Herbier Boissier 2ème série 8: 709–732.
- Castro S, Muratet A, Szczepaniak M, Nguéack J, Hardion L (2023) RAD sequencing, morphometry and synecology clarify the taxonomy of the *Melica ciliata* (Poaceae) complex in France and Poland. Journal of Systematics and Evolution 61: 764–775. <https://doi.org/10.1111/jse.12940>
- Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S (2016) Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Molecular Ecology Resources 16: 138–149. <https://doi.org/10.1111/1755-0998.12438>
- Chung K-F (2007) Inclusion of the South Pacific alpine genus *Oreomyrrhis* (Apiaceae) in *Chaerophyllum* based on nuclear and chloroplast DNA sequences. Systematic Botany 32: 671–681. <https://doi.org/10.1600/036364407782250517>
- Dodsworth S (2015) Genome skimming for next-generation biodiversity analysis. Trends in Plant Science 20: 525–527. <https://doi.org/10.1016/j.tplants.2015.06.012>
- Downie SR, Katz-Downie DS, Spalik K (2000) A phylogeny of Apiaceae tribe Scandiceae: evidence from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 87: 76–95. <https://doi.org/10.2307/2656687>
- Downie SR, Plunkett GM, Watson MF, Spalik K, Katz-Downie DS, Valiejo-Roman CM, Terentjeva EI, Troitsky AV, Lee B-Y, Lahham J, El-Oqlah A (2001) Tribes and clades within Apiaceae subfamily Apioideae: the contribution of molecular data. Edinburgh Journal of Botany 58: 301–330. <https://doi.org/10.1017/S0960428601000658>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11–15.
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32: 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Hardion L, Perrier A, Martinez M, Navrot N, Gaquerel E, Tournay F, Nguéack J, Combroux I (2020) Integrative revision of *Dianthus superbus* subspecies reveals different degrees of differentiation, from plasticity to species distinction. Systematics and Biodiversity 18: 255–268. <https://doi.org/10.1080/14772000.2020.1737979>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. <https://doi.org/10.1093/molbev/mst010>
- Meusel H, Jäger EJ, Weinert E (1965) Vergleichende Chorologie der Zentraleuropäischen Flora. Verlag Gustav Fischer, Jena, vol. 1: 1–333, vol. 2: 1–688.
- Oba S, Sato MA, Takemasa I, Monden M, Matsubara KI, Ishii S (2003) A Bayesian missing value estimation method for gene expression profile data. Bioinformatics 19: 2088–2096. <https://doi.org/10.1093/bioinformatics/btg287>
- Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26: 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Piwczyński M, Puchałka R, Spalik K (2015) The infrageneric taxonomy of *Chaerophyllum* (Apiaceae) revisited: new evidence from nuclear ribosomal DNA ITS sequences and fruit anatomy. Botanical Journal of the Linnean Society 178: 298–313. <https://doi.org/10.1111/boj.12282>
- R Core Team (2024) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> [accessed 16.09.2024]
- Reduron JP (2007) Ombellifères de France: Monographie des Ombellifères (Apiaceae) et Plantes Alliées, Indigènes, Naturalisées, Spontanées, Adventices ou Cultivées de la Flore Française. Société botanique du Centre-Ouest, Nercillac, 1–564.
- Spalik K, Piwczynski M, Danderson CA, Kurzyńska-Młynik R, Bone TS, Downie SR (2010) Amphitropic amphiantarctic disjunctions in Apiaceae subfamily Apioideae. Journal of Biogeography 37: 1977–1994. <https://doi.org/10.1111/j.1365-2699.2010.02334.x>
- Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007) pcaMethods—a bioconductor package providing PCA methods for incomplete data. Bioinformatics 23: 1164–1167. <https://doi.org/10.1093/bioinformatics/btm069>
- Tison JM, de Foucault B (2014) Flora Gallica: flore de France. Biotope, Mèze, 1–1196.
- Wörz A (1988) Investigations of distribution and phytosociological characteristics of *Chaerophyllum elegans* Gaudin (Apiaceae). Botanica Helvetica 98: 87–96. <https://doi.org/10.5169/SEALS-68573>

SUPPLEMENTARY MATERIALS

Supplementary material 1

Sampling information, morphological data, and DNA sequence accessions.

<https://doi.org/10.5091/plecevo.124907.suppl1>

Supplementary material 2

Photographs of the basal leaves of *Chaerophyllum villarsii* var. *villarsii* (left; A. Binz 1508, BAS-BU) and *Chaerophyllum hirsutum* (right; A. Thellung 1011, BAS-BU).

<https://doi.org/10.5091/plecevo.124907.suppl2>

Supplementary material 3

Geographic distribution of ribotypes, haplotypes, and taxa (map from www.openstreetmap.fr).

<https://doi.org/10.5091/plecevo.124907.suppl3>