



HAL
open science

Evolution of the distribution area of the Mediterranean *Nigella damascena* and a likely multiple molecular origin of its perianth dimorphism

Florian Jabbour, Pierre-Emmanuel Du Pasquier, Léa Chazalviel, Martine Le Guilloux, Natalia Condee Silva, Yves Deveaux, Domenica Manicacci, Pierre Galipot, Andreas Heiss, Catherine Damerval

► To cite this version:

Florian Jabbour, Pierre-Emmanuel Du Pasquier, Léa Chazalviel, Martine Le Guilloux, Natalia Condee Silva, et al.. Evolution of the distribution area of the Mediterranean *Nigella damascena* and a likely multiple molecular origin of its perianth dimorphism. *Flora*, 2021, pp.151735. 10.1016/j.flora.2020.151735 . hal-03017899v1

HAL Id: hal-03017899

<https://hal.science/hal-03017899v1>

Submitted on 21 Nov 2020 (v1), last revised 25 Aug 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evolution of the distribution area of the Mediterranean *Nigella damascena* and a likely multiple molecular origin of its perianth dimorphism

Florian Jabbour^{a,§,*}, Pierre-Emmanuel Du Pasquier^{b,§}, Léa Chazalviel^a, Martine Le Guilloux^c, Natalia Conde e Silva^c, Yves Deveaux^c, Domenica Manicacci^c, Pierre Galipot^{a,c}, Andreas G. Heiss^d, Catherine Damerval^{c,*}

^a *Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP39, 57 rue Cuvier, 75005 Paris France*

^b *Laboratoire de Génétique évolutive, Université de Neuchâtel, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland*

^c *Université Paris-Saclay, INRAE, CNRS, AgroParisTech, Génétique Quantitative et Evolution-Le Moulon, 91190 Gif-sur-Yvette, France*

^d *Austrian Archaeological Institute (ÖAI), Austrian Academy of Sciences (ÖAW), 1190 Vienna, Austria*

Running title: Geographic and molecular origin of perianth dimorphism in *Nigella damascena*

§*Co-first authorship.*

**Corresponding authors.*

E-mail addresses: fjabbour@mnhn.fr (F. Jabbour) and catherine.damerval@universite-paris-saclay.fr (C. Damerval)

ABSTRACT

Nigella damascena (Ranunculaceae) is a common Mediterranean species traditionally used as condiment and healing herb. Its perianth dimorphism has been known since 1591. In all modern accessions examined to date, a single allele of the *NdAP3-3* gene characterized by a MITE insertion has been found to be responsible for the mutant morph, lacking a corolla. Based on herbarium collections, we aimed at (i) reconstructing the evolution of the worldwide distribution area of both morphs of *N. damascena*, and (ii) assessing the diversity in the molecular origin of the mutant allele.

We compiled a database of more than 3000 collections of *N. damascena* across four centuries. When possible, specimens were georeferenced, and distribution maps were built for the wild-type (WT) and mutant morphs. Additionally, we genotyped the *NdAP3-3* locus, based on DNA extracted from herbarium specimens collected from various localities and at different times.

We observe a consistently larger and denser distribution area of WT plants in Europe and North Africa compared with the mutant morph. The relative proportions of mutant and WT morph plants collected in the last couple of centuries are roughly similar. Interestingly in herbarium specimens, *NdAP3-3* alleles without a MITE insertion can be found in plants exhibiting the mutant morph.

The diachronic study of the distribution area of *N. damascena* and a literature survey showed that populations exhibiting the mutant phenotype have been present since the end of the 16th century, though not frequently, in the broad Mediterranean Basin. Our genotyping results indicate that the origin of the mutant morph may be multiple and suggest that the *NdAP3-3* allele with the MITE insertion typically found in all modern accessions to date, is only a sample of the variability susceptible to disrupt petal formation and development in *Nigella damascena*.

Keywords: floral dimorphism, genotyping, herbarium collections, museum time series, nectariferous petal, *NdAP3-3*

1. Introduction

Nigella damascena L. (Ranunculaceae) is a common eurymediterranean weed species (Zohary, 1983) and a popular ornamental plant known as love-in-a-mist. It is herbaceous, annual and diploid $2n = 12$ chromosomes (Ghosh and Datta, 2006). Its current distribution was likely shaped by ancient trade routes allowing synanthropic long-distance dispersal of its seeds (Heiss et al., 2011). Although less commonly used than *N. sativa*, its traditional uses by human populations as condiment and healing herb throughout history have already been well documented, in particular for the Mediterranean (Heiss et al., 2012/2013). In the past few decades, its pharmaceutical properties have been actively investigated (e.g. Agradi et al., 2002; Fico et al., 2004; Toma et al., 2015). Archeological finds of *N. damascena* date back to the Late Bronze Age (1410–920 BC) in central Europe. They suggest that the species was introduced to the Alps by foreign miners in the course of ore exploitation (Heiss and Oeggel, 2005).

Two floral morphs have been known in this species since 1591 (Clusius, 1601). The wild-type flower has a bipartite perianth composed of c. 5 petaloid sepals and a series of c. 8 pouch-shaped nectariferous petals. Stamens are numerous, surrounding the gynoecium constituted of most generally 5 fused carpels. The alternative morph differs mainly in the perianth that is unipartite, consisting of numerous petaloid tepals (Fig. 1; Deroin et al., 2015; Jabbour et al., 2015). Although wild-type nectariferous flowers can be pollinated by insects, self-fertilization is possible because as the flower matures, the curvature of the styles results in stigmatic surfaces coming into contact with the dehiscent stamens. The morph lacking petals appears mainly self-pollinating (D. Manicacci and B. Gonçalves, personal observations). The monogenic control of the dimorphism was first established by Toxopeus (1927) based on segregation analyses, and

its molecular basis was unraveled almost a century later (Gonçalves et al., 2013; Zhang et al., 2013). The shift to a unipartite perianth is caused by a mutation of an unknown nature entailing the absence of expression of a master gene (*NdAP3-3*) directing petal formation. We will adopt the terminology of Gonçalves et al. (2013) to designate the wild-type (homozygous dominant PP or heterozygous Pp) and mutant (homozygous recessive pp) floral morphs, called [P] and [T], respectively. The [T] morph with its conspicuous flowers harboring supernumerary petaloid organs has drawn attention of horticulturalists for long, which has likely played an important role in its long-term maintenance, whereas the self-maintenance of the [T] morph in natural populations can be questioned.

Herbarium collections are repositories of past and current plant biodiversity. The millions of specimens stored in the thousands of herbaria worldwide provide a diachronic view of species distribution and of species phenotype variation. An increasing number of studies rely on these natural history resources, addressing questions ranging from genetic population history to the effects of environmental changes (Besnard et al., 2018).

Knowledge about the spatial distribution and the frequency of occurrence through time of [P] and [T] individuals of *N. damascena* would help evolutionary botanists investigate the origin and ecological grounds for the maintenance of the [T] morph in the wild. To get insight into these topics, we built a quasi-exhaustive database of the world collection of *N. damascena* herbarium specimens. A preliminary study focusing on specimens collected in mainland France showed that the collection localities of the mutant morph were less numerous (one-sixth) but more scattered than those of the wild-type morph, and suggested that the presence of [T] plants found in natural populations in the vicinity of towns and villages are feral individuals escaped from gardens (Damerval et al., 2018).

In this paper, we reconstructed the evolution of the worldwide distribution area of *N. damascena*, tracking the distribution of both morphs. Additionally, we assessed the diversity of

the mutant allele at the *NdAP3-3* locus in the [T] individuals, based on ancient DNA extracted from herbarium specimens collected from various localities around the Mediterranean Basin and at different times.

2. Material and methods

2.1. Data collection and database construction

We built a database gathering the information found on specimens of *N. damascena* from 42 herbaria, located in the Mediterranean Basin or known to house large amounts of species growing in areas with a Mediterranean climate (Supplementary data Table S1). The codes of the 42 herbaria are: AD, ATHU, B, BC, BCN, BEI, BM, BUC, CAI, CANB, CNHM, COI, E, FI, FR, G, GAZI, HO, HUH, HUJ, ISTF, K, LINN, LISU, LP, MA, MEL, MELU, MPU, NE, NHMS, NSW, NY, P, PAL, S, SIENA, TO, UPS, VAL, W, ZA; they follow Thiers (2020). Online herbarium databases (when existing and available; e.g. P Herbarium of the Muséum national d'Histoire naturelle, Paris: <https://science.mnhn.fr/institution/mnhn/collection/p/item/search>, Kew Herbarium Catalogue: <http://apps.kew.org/herbcat/navigator.do>) were checked to get an insight about the collection and exploit the available images of herbarium specimens. The herbaria sent specimens on loan to P, where they were studied and photographed for the record. Alternatively, the herbaria sent images of their collection of *N. damascena*. We had access to all or part of the collection of *N. damascena* from the 42 herbaria that answered positively to our request; and although our database is not exhaustive, it is the most complete we could possibly build. All the material required for the constitution of the database (herbarium specimens, online images, and images sent by the curators) was gathered until summer 2018. New records or very recently collected specimens are hence not included in our database.

The database was built and handled in Excel 14.3.6 (Supplementary data Table S2). Its structure and the protocol to select the herbarium specimens for georeferencing and building the temporal sequence of distribution maps were described by Damerval et al. (2018) who then focused on mainland France. Entries corresponding to duplicates (same collection, i.e. identical collector and collection number) were discarded in the analyses.

We also checked whether *N. damascena* is mentioned in the floras of the Mediterranean and Irano-Turanian regions (cited in the Results and Discussion sections), and we gathered the ecological data when they were reported.

2.2. Diachronic study of the world distribution of *Nigella damascena*

Two datasets were extracted from the database:

(i) The georeferenced dataset, consisting in 2009 specimens from which the floral morph could be determined. We classified the entries into five time intervals ([1800–1849], [1850–1899], [1900–1949], [1950–1999], [2000–2018]), and two categories ([P] and [T] morph). Maps were built using QGIS 2.18.15 (2018). The geodetic coordinate system we used is EPSG: 4326, WGS 84. For ease of viewing, we decided to represent on the distribution maps the density of collection per standard area rather than the whole set of individual collection localities. We used the Vector Grid tool in QGIS to add a layer of 50 km x 50 km squares on the maps. The number of collections per square was calculated using ftools, and organized into four classes: [1 to 3 collections], [4 to 10], [11 to 30] and [more than 31].

(ii) The [P]/[T] dataset, consisting in 2763 specimens (2382 [P], 336 [T] and 45 [P]/[T] belonging to the same collection). This dataset was built to compare the number of collected specimens from each morph among 50-year periods starting from the year 1800.

2.3. Genotyping herbarium specimens at the NdAP3-3 locus

Leaves were sampled from plants from 12 herbarium specimens (8 [T] and 7 [P] individuals; Table 1). Stringent precautions were taken to prevent possible contamination by recent DNA: reagents and equipment were UV-treated, and experiments were done in a running laminar flow hood. An extraction control (without tissue sample) was run along the herbarium samples. Dried leaves samples (8–15 mg) were crushed to fine powder using the Qiagen TissueLyserII (5x20 sec at 23 Hertz). The CTAB procedure for ancient DNA described in Gutaker *et al.* (2017) was then used, and extracted DNA was checked for amount and quality with a Nanodrop 2000 Spectrophotometer (Thermo Scientific). The extraction was repeated twice for four samples. DNA was also extracted from leaves of the three genotypes (PP, Pp and pp) of two modern accessions maintained in the lab and used for comparison.

In all modern accessions with [T] morph flowers, a MITE insertion was observed in the second intron of the *NdAP3-3* gene. Two primers flanking the MITE insertion point (U7, R9) and a supplementary forward primer within the MITE sequence (U9) were designed and checked for specificity with recent DNA of homozygous PP and pp, and heterozygous Pp genotypes (Supplementary data Table S3, Supplementary Fig. 1). Fragments of 94 bp and 344 bp were expected to be amplified for the P and p allele respectively using U7-R9. Because ancient DNA may be degraded, the 344 bp fragment might not be amplified in the p allele. To overcome this difficulty, we use the U9-R9 which should amplify a fragment of 97 bp specific to the p allele with a MITE insertion. PCR amplification conditions were 95°C 3 min, then 40 cycles of 95°C 30 sec, 55°C 30 sec, 72°C 1 min, followed by 7 min final elongation at 72°C (20 µL mix with Promega buffer 1X, MgCl₂ 1.5mM, dNTP 200 µM, Forward and Reverse primers 0.5µM each, GoTaq 0.2 µL). Two DNA concentrations for each sample were generally used for amplification, and the PCR were repeated at least twice for all samples and primer pairs. The extraction control as well as the standard PCR control (without DNA) were run in parallel with the herbarium samples, to check for absence of contamination. Direct sequencing of the 344 bp

amplified fragment was done for a modern [T] morph accession (Supplementary data Alignment). Small size PCR products obtained using U7-R9 and U9-R9 primer pairs were cloned in pGEM-T Easy vector (Promega). One to two clones were sequenced for nine specimens (four and three herbarium specimens from the [P] and [T] morphs, respectively, and two modern accessions from the [P] and [T] morphs) (Supplementary data Alignment). Sequencing was performed by Eurofins Genomics (Cologne, Germany).

3. Results

3.1. A database compiling more than 3000 worldwide collections across four centuries for *Nigella damascena*

The whole dataset consists of 3294 collections. The oldest specimen included in the database dates back to 1600–1625 (*J. Burser s.n.*, [P] and [T]-morph plants, both cultivated in Leipzig, Germany), and the most recently collected were *Jabbour et al. 1–9* ([P] and [T], all cultivated), collected by F. Jabbour, M. Le Guilloux and C. Damerval in 2015. In the database, the oldest and most recent records that were georeferenced (i.e. collected in the wild) are *leg. ign. s.n.* (1793, Italy, TO) and *Lecchi et al. s.n.* (16 May 2014, Italy, FI).

The three herbaria that contributed most in terms of number of collections were P (613 specimens), FI (390) and BM (230). The three most active collectors of *N. damascena* specimens were Carlo Pietro Stefano Sommier (1848–1922; Italian; 28 collections), Eugène Bourgeau (1813–1877; French; 19), and Elisée Reverchon (1835–1914; French; 13). Information about these collectors are available at kiki.huh.harvard.edu/databases/botanist_index.html.

Figure 2 shows the likely total distribution area of *Nigella damascena* reconstructed based on 2009 georeferenced herbarium specimens collected in the wild from 1793 onwards, no matter

the floral morph, and no matter the floral morph could be identified at all. It is broadly comprised between 27°N and 55°N, and between 18°W and 50°E. *Nigella damascena* plants have rarely been collected from the wild outside the Mediterranean Basin *sensu lato* (including northern European regions) (not shown on Fig. 2). Plants from the [P] morph were collected in Jamaica (1 collection) and on Saint Helena (1). Few [T] plants were collected in South-East Australia (5), in the Azores Archipelago (1), Cuba (1), in the State of New York (1) and in Colorado (1). Both morphs were found together in South Africa (1).

In the Mediterranean Basin, the countries that contributed most to our database are France (463 [P] specimens, 93 [T] specimens, and 15 [P]/[T] specimens belonging to the same collection), Italy (469, 57, 7), Spain (512, 20, 3), Greece (174, 7, 3), Portugal (139, 11, 0), Croatia (93, 14, 3), Turkey (110, 3, 1), and Algeria (99, 1, 2). The collection effort of both morphs was the highest between 1850 and 1899 (700 specimens, Fig. 3), before decreasing in the early 20th century (565) and increasing in the late 20th century (616). The proportions of wild [P] and [T]-morph specimens collected during each time slice are similar, and the proportions of [P] specimens collected are five (before 1849) to more than ten-fold (1950–1999) the proportions of [T] specimens (Fig. 3).

3.2. Distribution and density of collections of [P] and [T] morphs in the Mediterranean Basin during the last couple of centuries are different

The evolution of the distribution area of both floral morphs of *N. damascena* was reconstructed by periods of 50 years beginning in 1800, focusing on the Mediterranean Basin. The distributions of both morphs do not change much over time, taking into account the variation in collection effort. However, they differ between morphs (Fig. 4). We describe here below our results following an anticlockwise direction around the Mediterranean Basin, from Morocco to

Portugal. In western North Africa, [P] plants have been collected since the first half of the 19th century in Algeria and since the second half of the 19th century in Morocco and Tunisia. Collections of [T] plants were made later (starting from the second half of the 19th century only in Algeria (2 collections) and from the first half of the 20th century in Morocco (a single collection)), and they were less numerous (quasi-anecdotal) and involved less localities. Plants from the [P] morph only were collected in Libya, in the end of the 20th century, and no record of *N. damascena* was ever made in Egypt where the species is not reported (Täckholm, 1956). There, collections were seldom and involved very few localities. *Nigella damascena* was collected in the Middle East only once, in Jerusalem ([P] morph) in 1953, although *Flora Palaestina* (Zohary, 1966) does not mention the species. The species is reported as absent from Lebanon and Syria (Mouterde, 1970). The easternmost limits of the distribution area of both morphs are located in SW Russia (between the Caspian Sea and the Black Sea) and in NW Iran. In each of these areas, collections were made in the second half of the 19th century and in the first half of the 20th century, respectively. In the literature, *N. damascena* was reported from Azerbaijan (Riedl, 1992) and from Pakistan (Riedl and Nasir, 1991).

During the 19th century, collections and collection localities of the [P] morph were more numerous than previously, especially in S Europe, with some northward expansion of the distribution area of the morph. While collection localities got more gathered on the southern coasts of Europe, they covered most of the general distribution area (Fig. 2) of the species at the end of the 20th century.

In the Balkan Peninsula (including Crete), the northwestern part of Turkey and Cyprus, the number of [P] morph collections increased greatly from the end of the 19th century to the end of the 20th century. Relatively fewer [T] morph collections have been made in these countries (Greece: 174 [P], 7 [T], 3 [P]/[T]; Turkey: 110 [P], 3 [T], 1 [P]/[T]; Cyprus: 39 [P], 2 [T]).

Plants from the [T] morph were mostly collected in Italy and France during the 19th century. The countries where most [T] morph plants were collected are France (93), Italy (57), Germany (32), Spain (20), Croatia (14) and Portugal (11). Additionally, more [T] than [P] plants were collected from Germany (32 [T]/ 28 [P]), Austria (8/1), Australia (7/0), and USA (5/1). The distribution area of the [T] morph expanded to northern Spain in the first half of the 20th century. In the following 50-year time period, collections got rarer and more scattered. Some were recorded in southern Spain. Plants exhibiting the mutant morph have very scarcely been collected on the North African coast during the last couple of centuries.

While collections of [T] plants from islands were rare during the whole study period, [P] morph plants were heavily collected from islands in the Mediterranean area. [P] plants have been collected in Madeira and in the Canary Islands, although not recently. [T] plants were collected in the Canary Islands in the first half of the 19th century. The number of collections of [P] morph plants from Mediterranean islands constantly increased during the last two centuries, except in Corsica where very few collections were recorded in the 20th century and in Sardinia where no collections were made in the first half of the 20th century. During the 20th century, the density of collections of [P] morph plants from the Aegean Islands and Crete was comparable to that in northern Italy, southern France and the southern half of Spain.

During the last couple of centuries, [T] morph plants have been less collected, their collection localities have been more scattered and located more northerly distributed than [P] morph plants.

The ecological data relative to *N. damascena* provided by herbarium specimens and those from literature (Supplementary data Floras) are either similar or complementary. Overall, both morphs of *N. damascena* occur as weeds and garden escapes, and the species' habitat corresponds mainly to harvests, cultivated fields, fallow lands, phrygana, maquis, garrigue, open woodlands, roadsides, and secondarily to ruderal habitats. To a lesser extent, *Nigella*

damascena is also reported from rocky, stony and grassy areas (rarely from cliffs, in Algeria), on calcareous soil and, in Greece, on soils rich in serpentine, dolomite, or metamorphic schists.

3.3. Genotyping herbarium specimens at the *NdAP3-3* locus

All [T] morph plants of modern accessions examined to date are characterized by a single MITE insertion in the second intron of the *NdAP3-3* gene (Gonçalves et al., 2013; Zhang et al., 2013). While this insertion cosegregates with the mutation responsible for the floral dimorphism, it still remains unclear whether it is the cause of the absence of expression of the *NdAP3-3* gene entailing the mutant [T] phenotype, or whether it is in strong linkage disequilibrium with the causal mutation. In any case, its presence/absence constitutes a reliable marker of the mutant/wild-type allele found in present-day populations. Using this marker, we wanted to assess whether the same mutant allele could be found in ancient populations, or if other alleles entailing the mutant phenotype may exist. We designed two pairs of primers able to amplify short fragments typical of ancient low-quality DNA in the wild-type and mutant alleles and checked their specificity on modern accessions (Supplementary Fig. 1). U7-R9 is flanking the MITE insertion point and is expected to amplify two fragments in a heterozygous genotype: the 344 bp corresponds to the allele with the insertion, and the 94 bp to the allele without the insertion. The upper primer of the second pair (U9) was designed within the MITE inserted sequence, and amplified a 97 bp fragment in combination with the R9 primer, that can be observed only in homozygous [T] and heterozygous [P] morph plants (Supplementary Fig. 1). The identity of the amplified fragments was confirmed by sequencing (Supplementary data Alignment).

Twelve herbarium specimens were selected, of which three consisted in plants of the two floral morphs. The collection dates ranged from 1847 to 1993, and the localities spanned France, Italy, Spain, Czech Republic, Germany, Algeria, Tunisia and Canada (Table 1). Because PCR

amplifications from the oldest specimen (P02391601, [T] morph) were not reliable on two different DNA extracts, this individual was finally discarded from the analyses. Based on the profile of PCR amplification using both pairs of primers, all seven [P] morph specimens were found homozygous for an allele without insertion (Fig. 5, Table 1, Supplementary Figs. 2 & 3, Supplementary data Alignment). Among the seven [T] morphs, two were found homozygous for an allele with the MITE insertion (P02570110, P02391524), four were heterozygous for the presence/absence of the MITE insertion (P03175016, P02559213, P02376183, P03174950) and one appeared homozygous for an allele without insertion (P02391571) (Fig. 5, Table 1, Supplementary Figs. 2 & 3). The 344 bp fragment encompassing the full insertion was not consistently amplified using U7-R9 primer pair but the U9-R9 primer pair retrieved a low size fragment of the expected size (Supplementary Figs. 2 & 3). Sequencing the low size amplified fragments in P03175016 and P02376183 confirmed their heterozygosity: absence of the MITE insertion for the product obtained by U7-R9 amplification and presence of the insertion for the product obtained by U9-R9 amplification (Supplementary data Alignment). In samples where no MITE insertion was revealed (all [P] morphs and the P02391571 [T] morph), we cannot exclude that a polymorphism in the binding region of the U9 primer results in the absence of amplification. In any case, because we have taken drastic precautions to avoid contamination by recent DNA, we have run negative controls for extraction and PCR in parallel with the herbarium samples (Fig. 5) and we have sequenced small size fragments obtained with both primer pairs, we came confident that our results reveal that *NdAP3-3* alleles without a MITE insertion can be found in [T] morph plants without petals. Our results suggest that in these ancient populations, mutant allele(s) without MITE insertion is almost as frequent as the mutant allele with MITE insertion typical of modern accessions. We cannot exclude that some [P] morph plants are in fact heterozygous (wild-type and mutant without MITE alleles): this could be the case of P03174950 and P02391571, where both morphs are encountered. However, the

frequency of mutant allele remains very low in natural populations, and [P] morph plants can be homozygous for a wild-type allele.

4. Discussion

4.1. Origin and expansion of *N. damascena*'s distribution area

We observe a consistently larger distribution area of plants from the [P] morph in Europe and North Africa compared with the [T] morph, especially in coastal areas characterized by a Mediterranean climate (Kottek et al., 2006). More specifically, plants from the [T] morph are very scarce on the North African coast. In contrast, they are more commonly collected North to the 45th parallel than plants from the [P] morph. For that matter, the *Atlas Flora Europaeae* (Jalas and Suominen, 1988) qualifies the species' occurrence as "only as an ephemeral escape from cultivation". The Mediterranean Basin hotspots *sensu* Médail and Quézel (1997) are included within the distribution area of *N. damascena* (Fig. 2), except the southern part of the Atlas Mountains and the eastern parts of Anatolia. The Syria-Lebanon-Israel hotspot is not represented in the distribution area of *N. damascena* (one could have expected the opposite, based on the species epithet *damascena*). Collections were made outside the broad Mediterranean Basin, and these sampled populations could be naturalized in regions with Mediterranean climate such as the western coast of North and South America, South Africa, and South-East Australia. Collections of likely naturalized plants in northern Europe and eastern North America show the tolerance of *N. damascena* to different climates. Whether this geographical expansion is natural or human-mediated is difficult to assess. *Nigella damascena* is a short-lived annual species frequently occurring on disturbed soils. However, *N. damascena* is also reported in several floras of the Mediterranean Basin (Hayek, 1927; Strid and Tan, 2002; Jeanmonod and Gamisans, 2007; Tison and de Foucault, 2014) and in herbarium collections on stony-rocky slopes or rocky grassland from low elevation to 1000 m or to cliffs, which could

be considered as a primary habitat. Interestingly, *N. damascena* seems to be rarely found in cultivated fields in Greece, where it occurs preferably on various substrates. A hypothesis would be that the species may have originated in Greece (or more largely from the Balkans) and/or western Turkey (the center of diversity of the genus (calculated from Davis, 1965; Strid and Tan, 2002)). It can nowadays be observed as a rapid colonizer of fallow land around the Mediterranean (Heiss et al., 2012/2013), and be considered a cereal weed (Saatkamp et al., 2011), possibly having colonized the Mediterranean Basin westward following the expansion of agrarian communities. Also, long-distance transport along trade and migration routes has to be considered for *N. damascena* (Heiss and Oeggl, 2005), a phenomenon which is already well-known for other periods and taxa (e.g. Bakels and Jacomet, 2003; Jacomet, 2014; Heiss et al., 2017). An alternative hypothesis would suggest that *N. damascena* already had a eurymediterranean distribution before the Neolithic, and then colonized and thrived in the first cultivated lands all around the Mediterranean Basin. We are convinced that the occurrences of the species outside the Mediterranean region are due to human activities, be they intentional or unintentional.

4.2. The evolution of Nigella damascena collection effort worldwide

Herbarium collection records should be analyzed with caution, as collections of natural history are intrinsically biased (see the discussion in Damerval et al., 2018). Not all biodiversity is captured, and the representativeness of a collection depends, among other factors, on the spatiotemporal uniformity of the collection effort (Auderset Joye et al., 2002; Daru et al., 2018; Besnard et al., 2018). For a given species, a way to partly circumvent collection bias, and to get a representative measure of the actual bias in the intraspecific diversity, is to consider the most exhaustive set of herbarium collections from the highest number of herbaria, and to consult

numerous authoritative floristic treatments from the wildland areas of the species' reconstructed distribution area. This is what we did for *N. damascena*.

The collection effort concerning *N. damascena* has not been uniform during the last two centuries. The decrease in specimen number collected in the first half of the 20th century could be due to the impact of the two world wars on the activity of the scientists (Le Bras et al., 2015). The fact that *N. damascena* has often been collected next to areas with high human population density could be due to the local presence of many collectors and to the naturalization of populations after escaping from gardens. Whilst this explanation could stand for the southern coast of France or the region of Valencia in Spain, we do not think it is valid for the whole Mediterranean Basin, as *N. damascena* has not been heavily collected from many densely inhabited areas (e.g. Middle East). Thus, we believe that the distribution of the species shown in Fig. 2 represents its likely total distribution area in the broad Mediterranean Basin.

The morph [T] is less frequent than the [P] morph over collections from the last couple of centuries suggesting that populations homogeneous for the [T] morph are less common in the wild than populations homogeneous for the [P] morph. This should not be surprising because the mutant allele is recessive, and even though [T] morph plants can self-fertilize, they do not compete favorably with allogamous [P] morphs in the wild (D. Manicacci, personal observations).

4.3. Contrasted distribution areas of [P] and [T] morph plants and long-term maintenance of populations with [T] morph plants in the wild

The temporal sequence of the distribution area of [T] morph populations in the last couple of centuries (Fig. 4) showed that populations exhibiting the mutant phenotype have always been present (and collected) in the Mediterranean Basin. Assessing whether populations with [T] morph plants always grow in the vicinity of wild-type populations, or are repeatedly found year

after year, is not possible based on the data at hand. Indeed, and to our knowledge, no collector included in our database planned a systematic and long-term survey focusing on *N. damascena*. Populations of the two morphs are not equally distributed. Relatively more [T] specimens have been collected in northern Europe and there is no reason to think that the collectors mostly living or botanizing in this part of the world were more sensitive to unusual phenotypes. A possible explanation is that in these regions where *N. damascena* is not spontaneous, most populations come from horticultural sources which are generally biased towards [T] morph flowers. Populations with [T] morph plants may only form ephemeral feral populations. Alternatively, they can get naturalized in the vicinity of towns and villages and from there colonize new areas. As the mutant phenotype is common in cultivation in gardens (alone or together with the wild-type morph), at least since the end of the 16th century, we assume that gardens would be effective sources of pollen and seeds for regularly starting new populations of [T] morph plants in the wild.

Although we acknowledge that the wild-type of *N. damascena* produces flowers with a bipartite perianth, it is to be highlighted that the type specimen of the species name is from the [T] morph (*LINN 700.1*, kept at the Linnean Herbarium in London). This lectotype was selected among the specimens cited by Linnaeus (1753) as belonging to his own original collections and was designated by Zohary (1983) in his taxonomic revision of the genus.

4.4. Genotyping herbarium samples suggests that the origin of the [T] morph is multiple

Our genotyping results indicate that the MITE insertion is not systematically associated with the [T] morph in herbarium samples. The sequence alignment of the short size PCR fragments shows a deletion of 3 nucleotides at the MITE insertion site, when compared to the wild type allele. However, this deletion does not correspond to a footprint of an excision of the MITE (Supplementary data Alignment), suggesting that this region may be polymorphic. Our results

therefore strongly suggest that the *NdAP3-3* allele with the MITE insertion typically found in all modern accessions to date, is only a sample of the variability susceptible to disrupt petal formation and development in *Nigella damascena*.

Knowing the major role of *AP3-3* genes in the formation and development of the petal in *N. damascena* but also in other Ranunculaceae species (Rasmussen et al., 2009, Zhang et al., 2013), three hypotheses can be put forward to interpret our results: (i) the MITE inserted secondarily in an already inactivated *NdAP3-3* allele that we also found in our ancient DNA analysis. Under this hypothesis, the MITE insertion is not responsible for *NdAP3-3* inactivation but in linkage disequilibrium with the causal mutation; (ii) inactive alleles of *NdAP3-3* can be produced not only by the MITE insertion observed in modern accessions (resulting in abnormal splicing and aberrant mRNA production, see Gonçalves et al., 2013), but also possibly by a MITE insertion in another region of the *NdAP3-3* locus, or by any coding sequence mutations resulting in inactive proteins, or mutations in *cis*-regulatory regions entailing the absence of gene expression; (iii) the absence of petals in some ancient *N. damascena* populations could be due to either the loss of function of a downstream target gene of *NdAP3-3* or alternatively a co-factor or a trans-acting regulatory factor of *NdAP3-3*. In any case, the first horticulturalists probably retained by chance the MITE inserted allele, which was later disseminated in horticultural varieties.

Note about the specific epithet damascena

The specific epithet *damascena* refers to Damascus, the capital city of Syria, and points at a possible Middle Eastern origin of the species. According to the International Plants Names Index (2020), 29 species from 18 families of angiosperms have *damascena* as a species epithet, as the botanists naming them should have believed these species were of eastern Mediterranean origin. As far as *N. damascena* is concerned, the reference to Damascus is far older than the

protologue provided by Linnaeus (1753; Heiss et al., 2012/2013) and we think it only indicates that the first people who coined the species name referred to its likely origin, as for many condiments and spices. The lack of collections in the Middle East does not necessarily mean that *N. damascena* has not been growing there during the last couple of centuries. However, we are quite confident in the absence of *N. damascena* from Syria, as we gathered data from Middle Eastern herbaria (Beirut and Jerusalem herbaria) and herbaria traditionally keeping extensive and old Middle Eastern collections (e.g. Geneva and Paris herbaria). A way to determine the geographical origin of the species would be to combine a phylogeographical study with palaeo- and archaeobotanical investigations.

Funding: The work of LC and PG at the Muséum national d'Histoire naturelle was supported by the Institute of Systematics, Evolution and Biodiversity. PG's PhD work is supported by the Ecole Normale Supérieure de Lyon. CD, MLG, YD and NCS benefit from the support of Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS, LabEx SPS and ANR-17-EUR-0007, EUR SPS-GSR).

CRedit author statement

Florian Jabbour: Project administration, Conceptualization, Methodology, Resources, Supervision, Investigation, Formal analysis, Visualization, Writing- Original draft preparation, Funding acquisition. **Pierre-Emmanuel Du Pasquier:** Investigation, Formal analysis, Visualization, Writing- Reviewing and Editing. **Léa Chazalviel:** Investigation, Formal analysis, Visualization. **Martine Le Guilloux:** Investigation, Validation. **Natalia Conde e Silva:** Formal analysis, Writing- Reviewing and Editing. **Yves Deveaux:** Formal analysis, Writing- Reviewing and Editing. **Domenica Manicacci:** Investigation, Formal analysis, Writing- Reviewing and Editing. **Pierre Galipot:** Writing- Reviewing and Editing.

Andreas G. Heiss: Writing- Reviewing and Editing. **Catherine Damerval:** Project administration, Conceptualization, Methodology, Resources, Supervision, Investigation, Formal analysis, Visualization, Writing- Original draft preparation, Funding acquisition.

Acknowledgements

We thank the curators and staff members of the Herbaria AD, ATHU, B, BC, BCN, BEI, BM, BUC, CAI, CANB, CNHM, COI, E, FI, FR, G, GAZI, HO, HUH, HUI, ISTF, K, LINN, LISU, LP, MA, MEL, MELU, MPU, NE, NHMS, NSW, NY, P, PAL, S, SIENA, TO, UPS, VAL, W, ZA for sending herbarium specimens on loan and/or providing digital images. The MNHN (National Museum of Natural History, Paris) gives access to the collections in the framework of the RECOLNAT national Research Infrastructure. We thank Alexandra Zisi and Jelena Aleksić for translations from Greek and Russian to English, respectively. We also thank two anonymous reviewers for their critical and helpful comments. Authors' contributions: FJ, NCS, DM, and CD designed the research; FJ, LC, DM, PEDP and CD generated the herbarium specimens' collection dataset and FJ, LC, PEDP and CD analyzed the corresponding results; MLG, NCS, YD and CD conducted the molecular study and FJ, MLG, NCS, YD and CD analyzed the corresponding results. FJ and CD wrote the paper and all authors contributed to the final writing of the manuscript and gave approval for publication.

References

- Agradi, E., Fico, G., Cillo, F., Francisci, C., Tome, F., 2002. Estrogenic activity of *Nigella damascena* extracts, evaluated using a recombinant yeast screen. *Phytother. Res.* 16, 414–416.
- Andersson, S., 1997. Genetic constraints on phenotypic evolution in *Nigella* (Ranunculaceae). *Biol. J. Linn. Soc.* 62, 519–532.

- Auderset Joye, D., Castella, E., Lachavanne, J.B., 2002. Occurrence of Characeae in Switzerland over the last two centuries (1800–2000). *Aquatic Bot.* 72: 369–385.
- Bakels, C.C., Jacomet, S., 2003. Access to luxury foods in Central Europe during the Roman period: the archaeobotanical evidence. *World Archaeol.* 34, 542–557.
- Besnard, G., Gaudeul, M., Lavergne, S., Muller, S., Rouhan, G., Sukhorukov, A.P., Vanderpoorten, A., Jabbour, F., 2018. Herbarium-based science in the twenty-first century. *Bot. Lett.* 165, 323–327.
- Bittkau, C., Comes, H.P., 2005. Evolutionary processes in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. *Mol. Ecol.* 14, 4065–4083.
- Bittkau, C., Comes, H.P., 2009. Molecular inference of a Late Pleistocene diversification shift in *Nigella* s. lat. (Ranunculaceae) resulting from increased speciation in the Aegean archipelago. *J. Biogeogr.* 36, 1346–1360.
- Clusius, C., 1601. *Melanthium pleno flore*. In: *Rariorum Plantarum Historia*. Ex officina Plantiniana Apud Ioannem Moretum, Antwerp, pp. 207–208.
- Comes, H.P., Tribsch, A., Bittkau, C., 2008. Plant speciation in continental island floras as exemplified by *Nigella* in the Aegean Archipelago. *Phil. Trans. R. Soc. B* 363, 3083–3096.
- Damerval, C., Ben Othman, W., Manicacci, D., Jabbour, F., 2018. Distribution area of the two floral morphs of *Nigella damascena* L. (Ranunculaceae): a diachronic study using herbarium specimens collected in France. *Bot. Lett.* 165, 396–403.
- Daru, B.H., Park, D.S., Primack, R.B., Willis, C.G., Barrinton, D.S., Whitfeld, T.J.S., Seidlerr, T.G., Sweeney, P.W., Foster, D.R., Ellison, A.M., Davis, C.C., 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytol.* 217, 939–955.

- Davis, P.H., 1965. Flora of Turkey and the East Aegean Islands. Edinburgh University Press, Edinburgh.
- Deroin, T., Damerval, C., Le Guilloux, M., Jabbour, F., 2015. Floral vascular patterns of the double-flowered and wild-type morphs of *Nigella damascena* L. (Ranunculaceae). *Modern Phytomorphol.* 7, 13–20.
- Fico, G., Panizzi, L., Flamini, G., Braca, A., Morelli, I., Tomè, F., Cioni, P.L., 2004. Biological screening of *Nigella damascena* for antimicrobial and molluscicidal activities. *Phyther. Res.* 18, 468–470.
- Ghosh, A., Datta, A.K., 2006. Karyotyping of *Nigella sativa* L. (black cumin) and *Nigella damascena* L. (love-in-a-mist) by image analyzing system. *Cytologia* 71, 1–4.
- Gonçalves, B., Nougé, O., Jabbour, F., Ridet, C., Morin, H., Laufs, P., Manicacci, D., Damerval, C., 2013. An *APETALA3* homolog controls both petal identity and floral meristem patterning in *Nigella damascena* L. (Ranunculaceae). *Plant J.* 76, 223–235.
- Gutaker, R.M., Reiter, E., Furtwängler, A., Schuenemann, V.J., Burbano, H.A., 2017. Extraction of ultrashort DNA molecules from herbarium specimens. *BioTechniques* 62, 76–79.
- Hayek, A., 1927. *Prodromus Florae Peninsulae Balcanicae*. Dahlem-Berlin.
- Heiss, A.G., Antolín, F., Bleicher, N., Harb, C., Jacomet, S., Kühn, M., Marinova, E., Stika, H.-P., Valamoti, S.M., 2017. State of the (t)art. Analytical approaches in the investigation of components and production traits of archaeological bread-like objects, applied to two finds from the Neolithic lakeshore settlement Parkhaus Opéra (Zürich, Switzerland). *PLoS One* 12, e0182401.
- Heiss, A.G., Kropf, M., Sontag, S., Weber, A., 2011. Seed morphology of *Nigella* s.l. (Ranunculaceae): identification, diagnostic traits, and their potential phylogenetic relevance. *Int. J. Plant Sci.* 172, 267–284.

- Heiss, A.G., Oeggl, K., 2005. The oldest evidence of *Nigella damascena* L. (Ranunculaceae) and its possible introduction to central Europe. *Veget. Hist. Archaeobot.* 14, 562–570.
- Heiss, A.G., Stika, H.P., De Zorzi, N., Jursa, M., 2012/2013. *Nigella* in the mirror of time: A brief attempt to draw a genus' ethnohistorical portrait. *Offa* 69/70, 147–169.
- The International Plant Names Index, 2020. Published on the Internet <http://www.ipni.org>.
- Jabbour, F., Udron, M., Le Guilloux, M., Gonçalves, B., Manicacci, D., Nadot, S., Damerval, C., 2015. Flower development schedule and *AGAMOUS*-like gene expression patterns in two morphs of *Nigella damascena* (Ranunculaceae) differing in floral architecture. *Bot. J. Linn. Soc.* 178, 608–619.
- Jacomet, S., 2014. Crop Diversity in Southwestern Central Europe from the Neolithic Onwards, in: Chevalier, A., Marinova, E., Peña-Chocarro, L. (Eds.), *Plants and People: Choices and Diversity through Time*. Oxbow Books, Oxford, pp. 82–95.
- Jalas, J., Suominen, J., 1988. *Atlas Florae Europaeae: Distribution of Vascular Plants in Europe* (Vol. 3). Cambridge University Press, Cambridge.
- Jeanmonod, D., Gamisans, J., 2007. *Flora corsica*. Edisud, Aix-en-Provence.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World Map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* 15, 259–263.
- Le Bras, G., Pignal, M., Jeanson, M.L., Muller, S., Aupic, C., Carré, B., Flament, G., Gaudeul, M., Gonçalves, C., Invernón, V.R., Jabbour, F., Lerat, E., Lowry, P.P., Offroy, B., Pérez Pimparé, E., Poncy, O., Rouhan, G., Haevermans, T., 2017. The French Muséum national d'histoire naturelle vascular plant herbarium collection dataset. *Sci. Data* 4, 170016.
- Linnaeus, C., 1753. *Nigella*. *Species Plantarum*, t. I. Laurentius Salvius, Stockholm, p. 534.
- Médail, F., Quézel, P., 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Ann. Missouri Bot. Gard.* 84, 112–127.

- Mouterde, P., 1970. Nouvelle flore du Liban et de la Syrie. Tome II. Texte. Dar-el-Machreq, Beyrouth.
- Penn, M.G., Cafferty, S., Carine, M., 2018. Mapping the history of botanical collectors: spatial patterns, diversity, and uniqueness through time. *Syst. Biodiv.* 16, 1–13.
- Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H., Ferguson, C.J., 2004. Implications of the decline in plant collecting for systematic and floristic research. *Syst. Bot.* 29, 216–220.
- QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.org>.
- Rasmussen, D.A., Kramer, E.M., Zimmer, E.A., 2009. One size fits all? Molecular evidence for a commonly inherited petal identity program in Ranunculales. *Am. J. Bot.* 96, 96–109.
- Renner, S.S., Rockinger, A., 2016. Is plant collecting in Germany coming to an end? *Willdenowia* 46, 93–97.
- Riedl, H., 1992. *Nigella* L. In: Iranshahr, M., Rechinger, K.H., Riedl, H. (Eds.), *Flora Iranica*, Vol. 171. Akademische Druck- und Verlagsanstalt, Vienna, pp. 24–34.
- Riedl, H., Nasir, Y.J., 1991. *Flora of Pakistan*. No. 193. Ranunculaceae. University of Karachi, National Herbarium, Pakistan Agricultural Research Council, Islamabad.
- Saatkamp, A., Affre, L., Dutoit, T., Poschlod, P., 2011. Germination traits explain soil seed persistence across species: the case of Mediterranean annual plants in cereal fields. *Ann. Bot.* 107, 415–426.
- Sieniawska, E., Sawicki, R., Golus, J., Swatko-Ossor, M., Ginalska, G., Skalicka-Wozniak, K., 2018. *Nigella damascena* L. essential oil—a valuable source of β -elemene for antimicrobial testing. *Molecules* 23, 256.
- Strid, A., 1969. Evolutionary trends in the breeding system of *Nigella* (Ranunculaceae). *Bot. Not.* 122, 380–397.

- Strid, A., Tan, K., 2002. Flora Hellenica Volume 2. Koeltz Scientific Books, Koenigstein.
- Täckholm, V., 1956. Students' flora of Egypt. Anglo-Egyptian Bookshop, Cairo.
- Tanikawa, N., Honma, K., Tatsuzawa, F., 2019. Flavonoids of the rose-pink, blue, and white flowers of *Nigella damascena* L. (Ranunculaceae). *Sci. Hortic.* 257, 108609.
- Thiers, B., 2020 and continuously updated. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. <http://sweetgum.nybg.org/science/ih/>.
- Tison, J.-M., de Foucault, B., 2014. Flora Gallica. Flore de France. Société Botanique de France. Biotope Editions, Mèze.
- Toma, C.C., Olah, N.K., Vlase, L., Mogoşan, C., Mocan, A., 2015. Comparative studies on polyphenolic composition, antioxidant and diuretic effects of *Nigella sativa* L. (black cumin) and *Nigella damascena* L. (lady-in-a-mist) seeds. *Molecules* 20, 9560–9574.
- Toxopeus, H.J., 1927. Erblichkeitsuntersuchungen an *Nigella damascena* L. *Genetica* 9, 341–440.
- Wang, P., Liao, H., Zhang, W., Yu, X., Zhang, R., Shan, H., Duan, X., Yao, X., Kong, H., 2015. Flexibility in the structure of spiral flowers and its underlying mechanisms. *Nature Plants* 2, 1–10.
- Yao, X., Zhang, W., Duan, X., Yuan, Y., Zhang, R., Shan, H., Kong, H., 2019. The making of elaborate petals in *Nigella* through developmental repatterning. *New Phytol.* 223, 385–396.
- Zhang, R., Guo, C., Zhang, W., Wang, P., Li, L., Duan, X., Du, Q., Zhao, L., Shan, H., Hodges, S.A., Kramer, E.M., Ren, Y., Kong, H., 2013. Disruption of the petal identity gene *APETALA3-3* is highly correlated with loss of petals within the buttercup family (Ranunculaceae). *Proc. Natl. Acad. Sci.* 110, 5074–5079.

Zohary, M., 1966. Flora Palaestina. Part I. Israel Academy of Sciences and Humanities, Jerusalem.

Zohary, M., 1983. The genus *Nigella* (Ranunculaceae) – a taxonomic revision. Plant Syst. Evol. 142, 71–105.

Table 1

Specimens from P Herbarium used for genotyping. The genotype is recorded as presence/absence of the MITE insertion in the second intron of the *NdAP3-3* gene. nr: not reproducible. Images of the specimens are accessible online using the following link: <https://science.mnhn.fr/institution/mnhn/collection/p/item/search>

Specimen barcode	Floral morph	Year of collection	Country of collection	Allele with a MITE insertion
P02391601	[T]	1847	Italy	nr
P02391571	[T]	1876	Czech Republic	-/-
	[P]			-/-
P03174950	[T]	1903	France	+/-
	[P]			-/-
P02376183	[T]	1904	Canada	+/-
P02570110	[T]	1906	France	+/+
P02559213	[T]	1914	France	+/-
P03175016	[T]	1917	France	+/-
P02391524	[T]	1918	Spain	+/+
	[P]			-/-
P02557922	[P]	1930	Algeria	-/-
P02559215	[P]	1930	France	-/-
P02819054	[P]	1993	Tunisia	-/-
P02557822	[P]	1975	Italy	-/-

Fig. 1. Herbarium flowers from (A) the [P] morph (herbarium barcode: P02305504; *Gouget s.n.*) and (B) the [T] morph (P02575144; *B. de Retz 284*). High-resolution images of the specimens can be viewed and downloaded using <http://coldb.mnhn.fr/catalognumber/mnhn/p/p02305504> and - p02575144.

Fig. 2. Likely distribution area of *Nigella damascena* in the broad Mediterranean Basin reconstructed based on 2009 georeferenced herbarium specimens collected in the wild from 1793 onwards, for (A) the morph [P], and (B) the morph [T]. Increasing collection densities are represented using 2500 km² squares of different colours. Yellow: 1–3 collections, green: 4–10, light blue: 11–30, dark blue: >31.

Fig. 3. Frequency of [P] and [T] morphs of *Nigella damascena* in a Museum time-series spanning more than two centuries. The total number of specimens collected in the wild and for which the floral morph could be determined is given for each time-period.

Fig. 4. Total distribution area of *Nigella damascena* in the broad Mediterranean Basin from 1800 onwards, broken up in 50-year periods. (A–E) Distribution area of the [P] morph. (F–J) Distribution area of the [T] morph. The number of specimens used to build each distribution map is specified.

Fig. 5. PCR amplification using the U9-R9 (A) and U7-R9 and U9-R9 primer pairs (B) in herbarium specimens. The size marker is Low molecular weight DNA Ladder (Biolabs, Ref: N3233S). (A) the amplification patterns indicate the presence of the MITE insertion in P03175016, P02376183 and P03174950, all [T] morphs, and its absence in P02559215 and P03174950 (both [P] morphs), and both morphs in P02391571.

(B) PCR were done with two DNA amounts (5 μ L and 10 μ L, respectively the two first and two last wells for each specimen morph) and using the two pairs of primers (first and third well: U7-R9 and second and fourth well: U9-R9) for each sample. The two [P] morph samples give a band at \sim 95 pb with U7-R9 only (both DNA concentrations for P03174950, only the highest one for P02391524) indicating absence of the MITE insertion. The 344 bp only was amplified in the P02391524 [T] morph with the U7-R9 pair, indicating the presence of the MITE inserted allele; this result is confirmed by the amplification of the \sim 95 bp band with U9-R9 pair that is specific of the MITE insertion.

Nc corresponds to negative extraction control and shows no amplification except for primer dimer (about 50 bp), validating the absence of DNA contamination in the specimen extracts.

Supplementary Figure 1

PCR amplification using the two pairs of primers U7-R9 (first bracket) and U9-R9 (second bracket) in two modern accessions (a, b). In each accession, the genotypes used for PCR amplification are successively: PP, Pp, pp. The U7-R9 primer pair amplifies both alleles in the heterozygous Pp genotypes, while the U9-R9 pair amplifies a fraction of the MITE insertion in both heterozygous Pp and pp genotypes.

Supplementary Figure 2

PCR amplification results for eight herbarium specimens and a modern accession (first comb: four [P] morphs (successively P02557922, P02819054, P02557822, P0229215), empty well then pp and PP modern accession genotypes; second comb: four [T] morphs (successively P03175016, P02559213, P02376183, P02391601), empty well then negative extraction control showing only primer dimers at \sim 50 bp. For each sample, the two first wells used 5 μ L DNA amount in the PCR mix with respectively U7-R9 then U9-R9 primer pairs, and the last wells

used 10 µL DNA amount, with the same primer pairs. P02391601 gave no amplification results in this experiment. Comparison of this figure with Fig. 5A shows that P03175016 and P02376183 (both [T] morphs) amplify the ~95 bp with both U7-R9 (no MITE insertion) and U9-R9 (with MITE insertion), indicating that these specimens are heterozygous for the presence / absence of the MITE.

Supplementary Figure 3

PCR amplification results for five herbarium specimens. First comb: P02557922, P02819054 (both [P] morphs) and P02376183 ([T] morph). The PCR amplifications are obtained from the second DNA extraction for these specimens: the patterns were the same as for the first extraction (used in Supplementary Fig. 2) for the two first specimens, while there was no amplification in P02376183. Second comb: P02570110 ([T] morph) and P02391571 ([T] then [P] morphs). The design is the same as in Supplementary Fig. 2 (four wells per specimen, two DNA concentrations and two primer pairs). P02570110 appears homozygous for the MITE insertion, while both morphs of P02391571 do not have the insertion and the [T] morph of P03174950 is heterozygous for the insertion. The lowest band corresponds to primer dimers.

Supplementary data Table S1

Contribution of each of the 42 herbaria to our database. The codes of the 42 herbaria listed in the table follow Thiers 2020.

Supplementary data Table S2

Database gathering the information found on the 3294 herbarium specimens of *Nigella damascena* from 42 herbaria. The data gathered from the biological specimen and from the label are organized into 16 fields: (A) herbarium code; (B) barcode or reference in the

herbarium; (C) collector's name; (D) collection number; (E–G) day, month, and year of collection; (H) country of origin, transcribed from the label or interpreted; current country names are provided; (I) collection locality; (J) other notes on the label (in the original language); (K) label mentioning that the specimen was in cultivation at the time of collection (yes/no); (L) assumption that the specimen was in cultivation at the time of collection (yes/no); (M) presence of flower(s) (yes/no); (N) flower morph ([P]/[T]); (O) presence of capsule(s) (yes/no); (P) GPS coordinates in decimal degrees, in the WGS 84 system. In some cases, GPS coordinates were available directly from the label. In most cases, localities were georeferenced by us using Google Maps (Google Inc., Mountain View, CA, USA). The collection locality could be precisely specified on the label allowing precise GPS coordinates to be determined (in red). Less precisely described collection sites, for example just mentioning the closest city, were also attributed GPS coordinates (in blue). In other cases, e.g. when only the country of origin is specified or when the plant was in cultivation when collected with no mention of the seed origin, localization data were considered unavailable and (P) cells were marked with a double-slash (/). Except for (H) and (L), all data are transcribed from the label without interpretation or come from the observation of the specimen.

Supplementary data Table S3

Primers used to amplify *NdAP3-3* second intron from DNA extracted from herbarium specimens and modern accessions.

Supplementary data Alignment

Alignment of the sequences obtained from the modern accessions and some of the herbarium specimens.

Supplementary data Floras

List of the floristic treatments including information about the distribution and ecology of

Nigella damascena.

Figure 1.

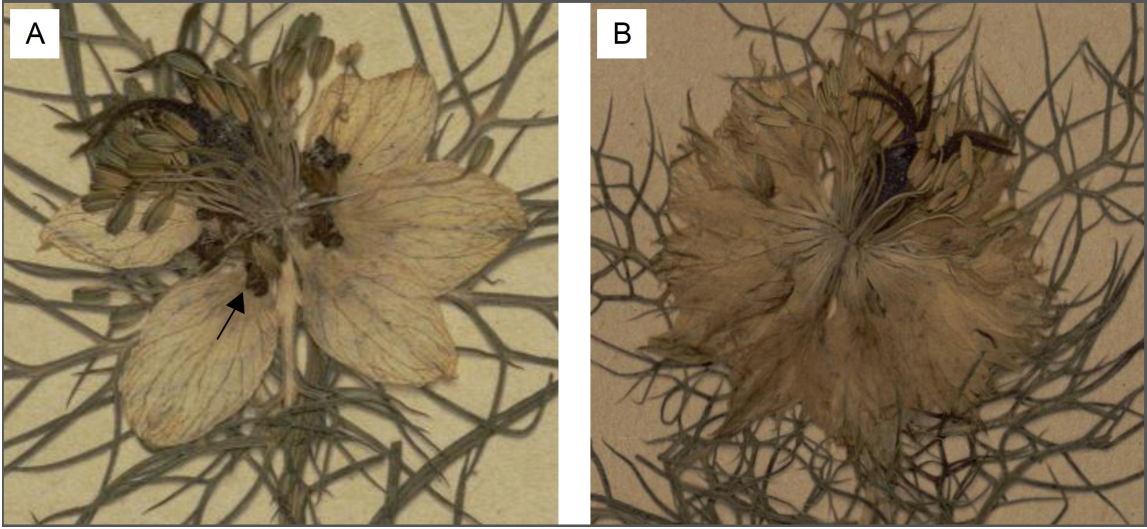


Figure 2.

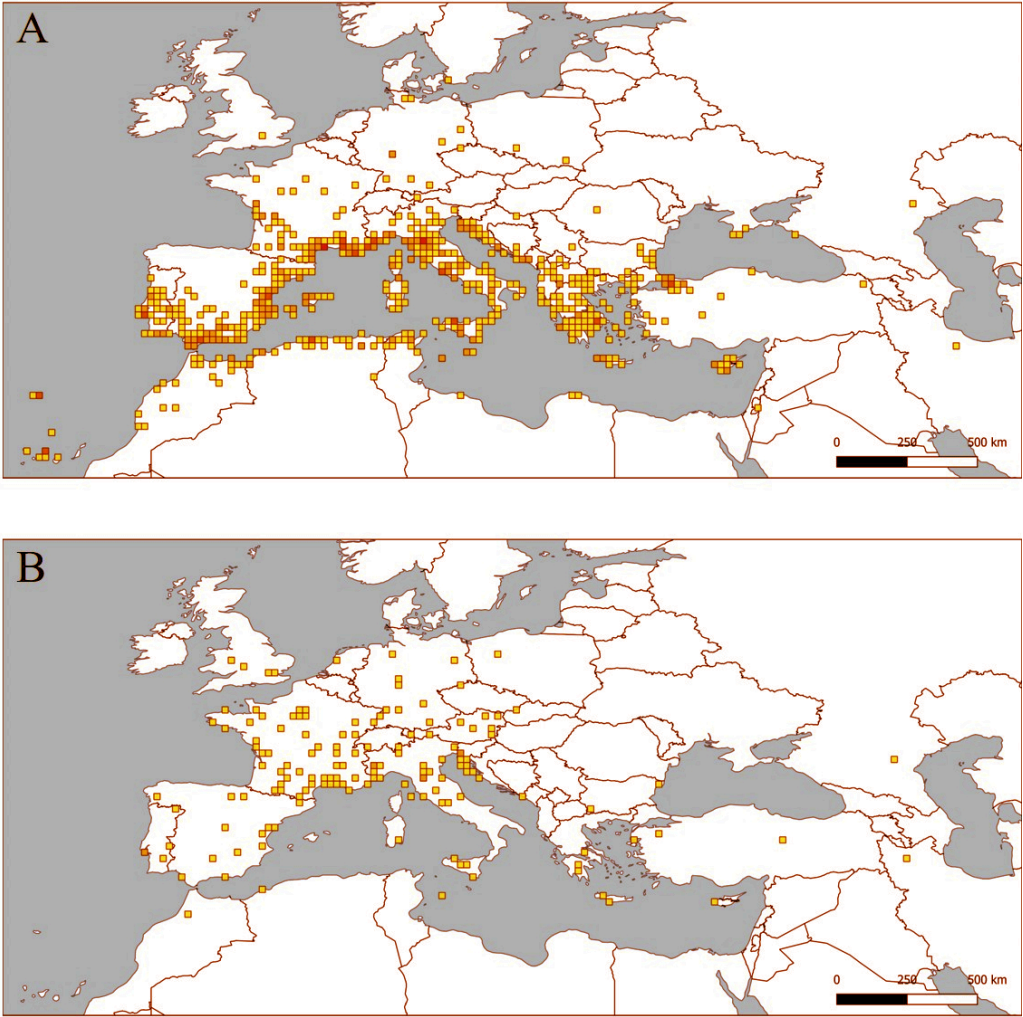


Figure 3.

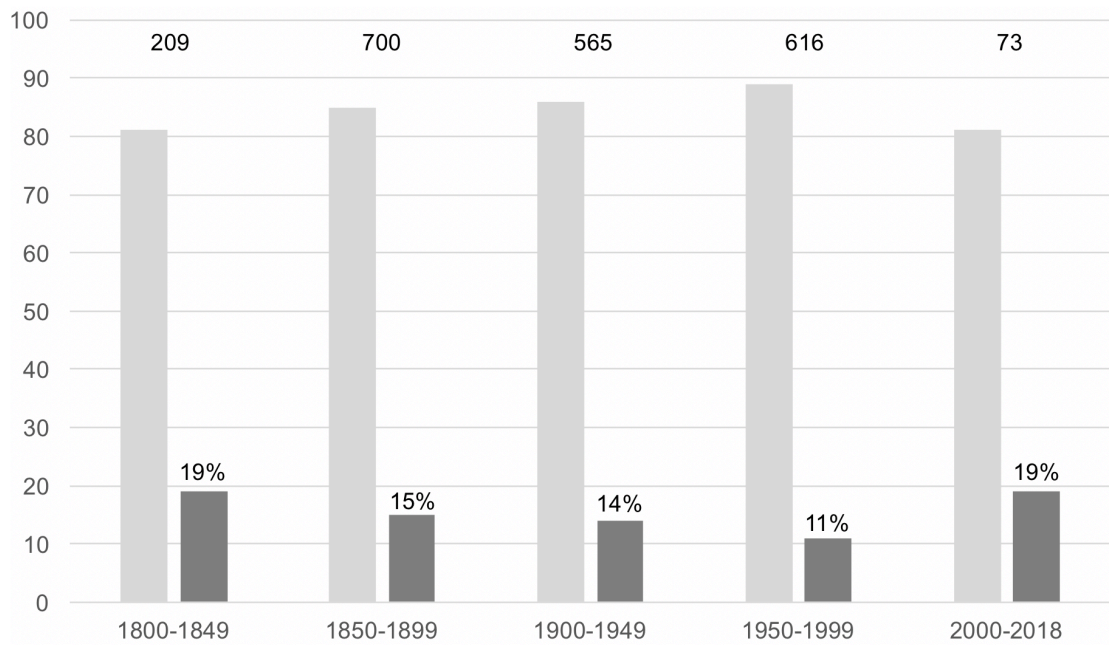


Figure 4.

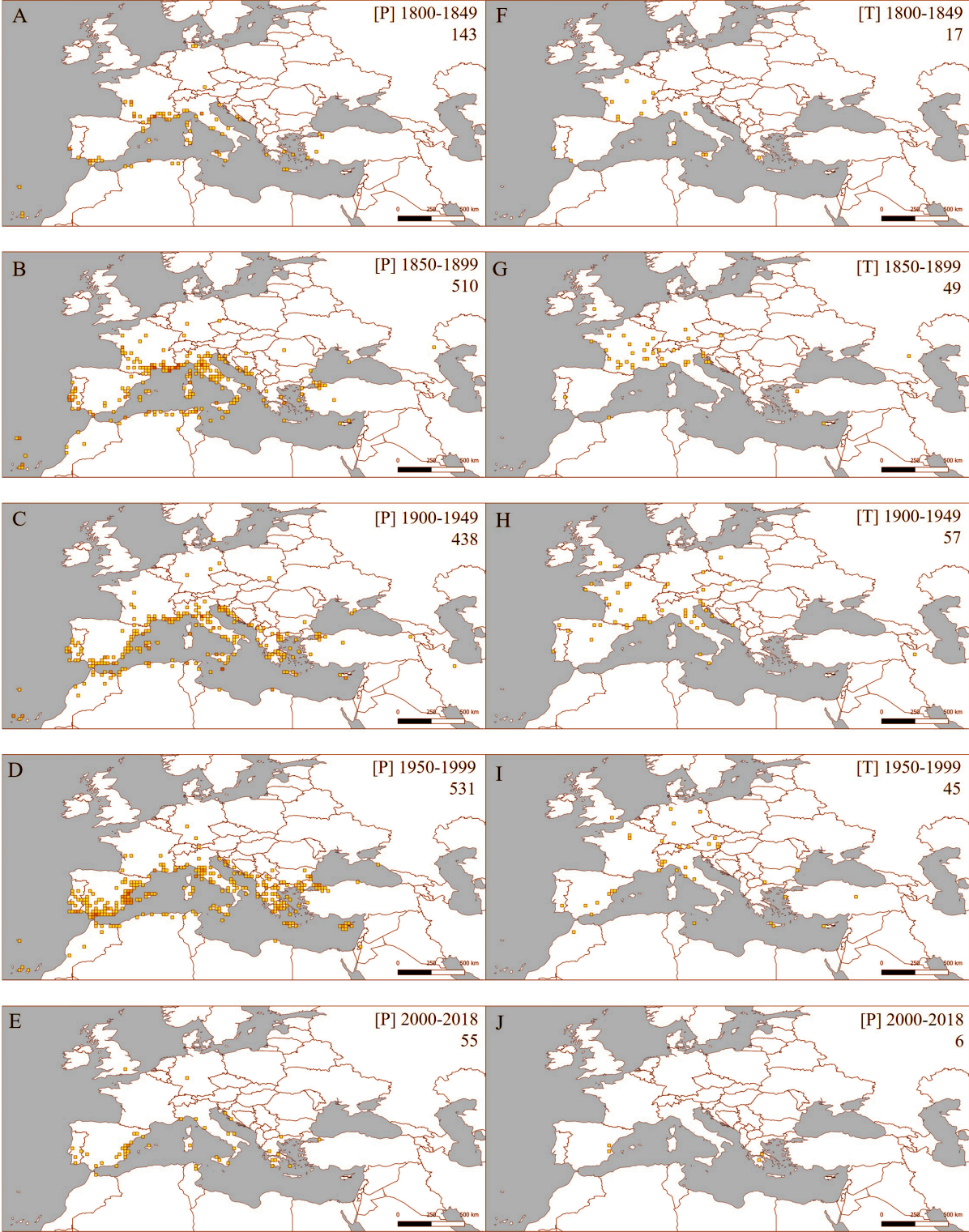
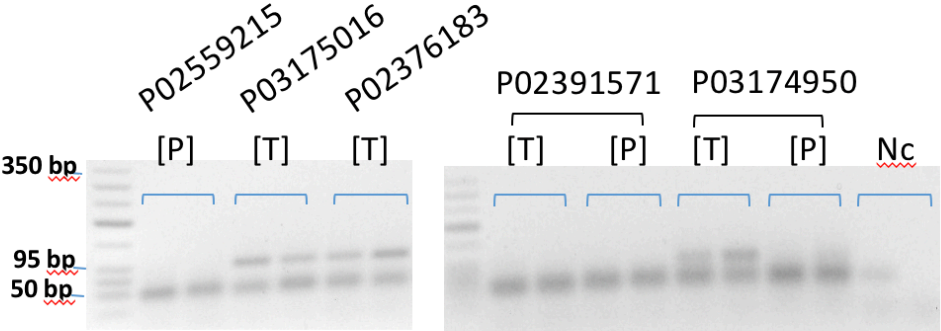
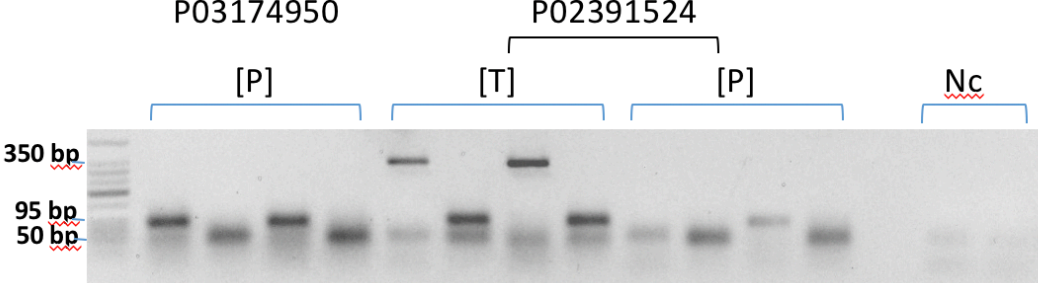


Figure 5.

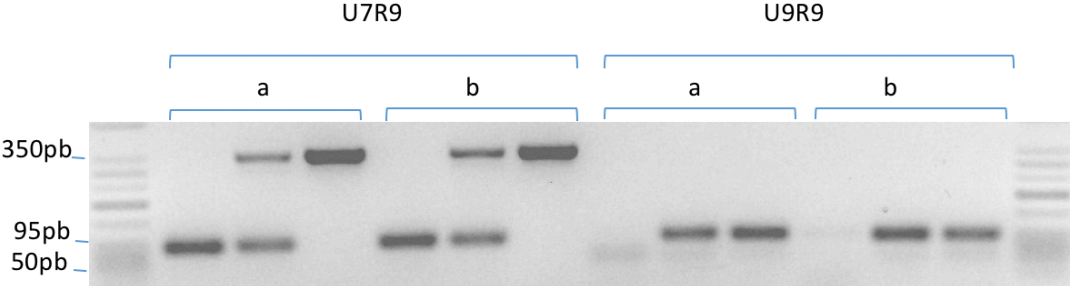
A



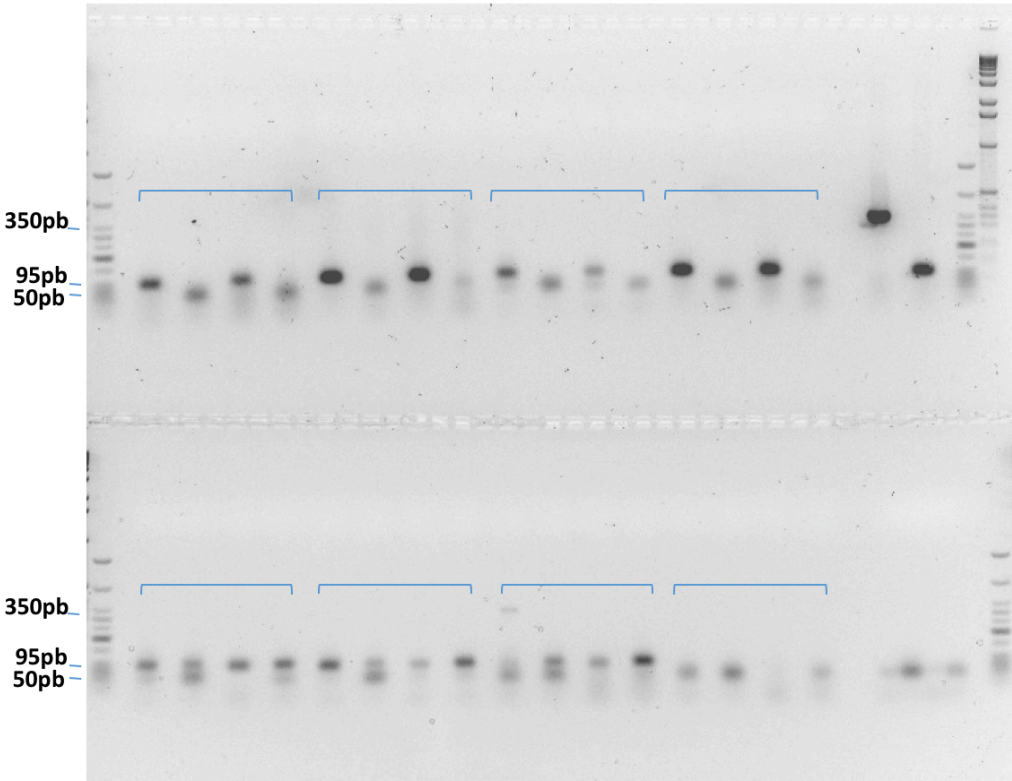
B



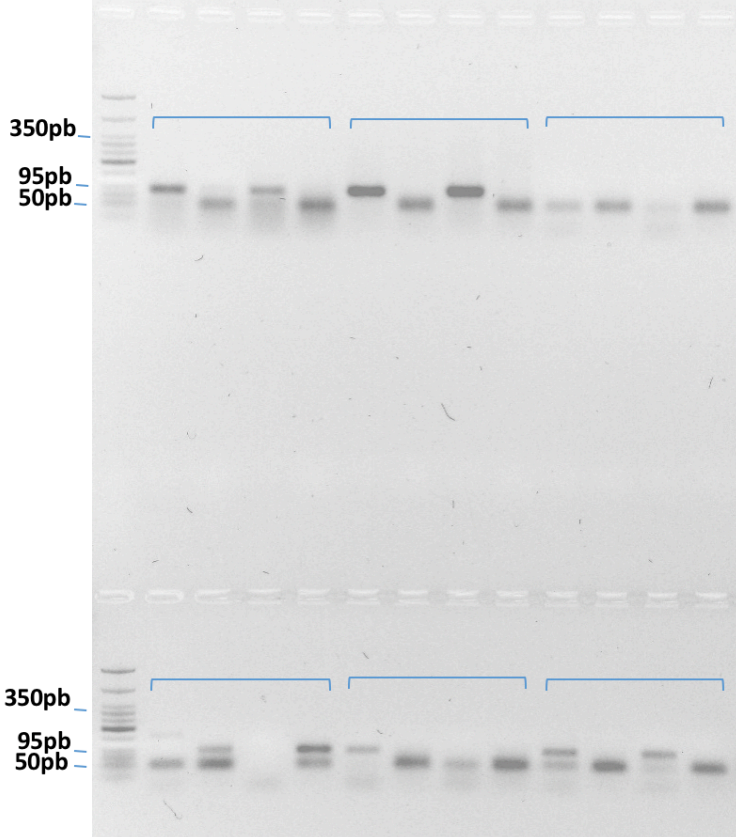
Supplementary Figure 1.



Supplementary Figure 2.



Supplementary Figure 3.



Supplementary data Table S1.

Code of Herbarium	Location of Herbarium	Number of specimens included in database
AD	Adelaide, Australia	2
ATHU	Athens, Greece	6
B	Berlin, Germany	169
BC	Barcelona, Spain	133
BCN	Barcelona, Spain	10
BEI	Beirut, Lebanon	6
BM	London, UK	230
BUC	Bucharest, Romania	4
CAI	Cairo, Egypt	5
CANB	Canberra, Australia	4
CNHM	Zagreb, Croatia	3
COI	Coimbra, Portugal	31
E	Edinburgh, UK	99
FI	Florence, Italy	390
FR	Frankfurt, Germany	80
G	Geneva, Switzerland	187
GAZI	Ankara, Turkey	1
HO	Hobart, Australia	3
HUH	Cambridge, USA	4
HUJ	Jerusalem, Israel	23
ISTF	Istanbul, Turkey	38
K	Kew, UK	209
LINN	London, UK	2
LISU	Lisbon, Portugal	46
LP	La Plata, Argentina	2
MA	Madrid, Spain	209
MEL	Melbourne, Australia	27
MELU	Melbourne, Australia	3
MPU	Montpellier, France	146
NE	Armidale, Australia	2
NHMS	Split, Croatia	3
NSW	Sydney, Australia	4
NY	New York, USA	3
P	Paris, France	623
PAL	Palermo, Italy	11
S	Stockholm, Sweden	88
SIENA	Siena, Italy	39
TO	Turin, Italy	86
UPS	Uppsala, Sweden	65
VAL	Valencia, Spain	120
W	Vienna, Austria	146
ZA	Zagreb, Croatia	44

Supplementary data Table S2. (See supplementary file)

Supplementary data Table S3.

Primer name	Primer sequence 5' → 3'
U7	CATCCCATTTTTCCATCATTTT
U9	CAATTTTTGCTGGTGGAAATCT
R9	CAAGTTTTCTTGCATTTTCTGC

Supplementary data Alignment.

```

>genomic_P
CATCCCATTTTTCCATCATTTTATTTATGAGGTTTGAAGATATTAATATACATAT
A-----
-----TAGTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG
>genomic_T
CATCCCATTTTTCCATCATTTTATTTATGAGGTTTGAAGATATTAATATACATAT
AGGCCCTGCTTGTTTTACCCGGAAAAGTGCCGGAAAGTAGAGCCACGTCACCTTT
CCCTCTTTTTCTGATGGAAGGATGCTTGGAAAGGCAGTTATTGGCGTGGGGTCCA
CGCAAAACTTTTACAAAATCTGAACAAAATCAGCTCCTTTTTTTGCTGTTTCAAAA
CGACTGGAAACTTTTGCTTTTGGCTGCAATTTTGGCTGGTGGAAATCTTTGTTTTT-
CTT-GTTTTT-
CCCTGCATCCAAGCAGGGCCATAGTTCTTGTAATTTGCAGAAAATGCAAGAAAAC
TTG
>Primers: _U7--R9
CATCCCATTTTTCCATCATTTT-----
-----
-----GCAGAAAATGCAAGAAAACCTTG
>Primers: _U9--R9
-----
-----
-----CAATTTTTGCTGGTGGAAATCT-----
-----GCAGAAAATGCAAGAAAACCTTG
>Modern_accession__T_U7R9_PCR-consensus
CATCCCATTTTTCCATCATTTTATTTATGAGGTTTGAAGATATTAATATACATAT
AGGCCCTGCTTGTTTTACCCGGAAAAGTGCCGGAAAGTAGAGCCACGTCACCTTT
CCCTCTTTTTCTGATGGAAGGATGCTTGGAAAGGCAGTTATTGGCGTGGGGTCCA
CGCAAAACTTTTACAAAATCTGAACAAAATCAGCTCCTTTTTTTGCTGTTTCAAAA
CGACTGGAAACTTTTGCTTTTGGCTGCAATTTTGGCTGGTGGAAATCTTTGTTTTT-
CTT-GTTTTT-
CCCTGCATCCAAGCAGGGCCATAGTTCTTGTAATTTGCAGAAAATGCAAGAAAAC
TTG
>Modern_accession_T_U9R9

```


-----CAATTTTGGCTGGTGGAAATCTTTGTTTTT-CTT-GTTTTT-
CCCTGCATCCAAGCAGGGCCATAGTTCTTGTAATTTGCAGAAAATGCAAGAAAAC
TTG

>Modern_accession_P_U7R9

CATYCCATTTTCCATCATTTTTATTTATGAGGTTTGAAGATATTAATATACATW---

-----ATAGTTCTTGTAATTTGGCGAAAATGCAAGAAAACCTTG

>P02557922_1P_U7R9

CATCCCATTTTCCATCATTTTTATTTATGAGGTTTGAAGATATTAATATACATAT--

-----ACTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG

>P02819054_2P_U7R9

CATCCCATTTTCCATCATTTTTATTTATGAGGTTTGAAGATATTAATATACATAT--

-----AGTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG

>P02557822_3P_U7R9

CATCCCATTTTCCATCATTTTTATTTATGAGGTTTGAAGATATTAATATACATAT--

-----AGTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG

>P02559215_4P_U7R9

CATCCCATTTTccaTCaTTTTATTTATGAGGTTTGAAGATATTAATATACATAT----

-----AGTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG

>P03175016_5T_U7R9

CATCCCATTTTCCATCATTTTTATTTATGAGGTTTGAAGATATTAATATACATAT--

-----GTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG

>P03175016_5T_U9R9

-----CAATTTTGGCTGGTGGAAATCTTTGTTTTT-CTTTGTTTTT-
CCCTGCATCCAAGCAGGGCCATAGTTCTTGTAATTTGCAGAAAATGCAAGAAAAC
TTG

>P02376183_7T_U7R9

CATCCCATTTTCCATCATTTTTATTTATGAGGTTTGAAGATATTAATATACATAT--

-----GTTCTTGTAATTTGCAGAAAATGCAAGAAAACCTTG

>P02376183_7T_U9R9

-----CAATTTTGGCTGGTGGAAATCTTTGTTTTT-CTT-

GTTTTTCCCTGCATCCAAGCAGGGCCATAGTTCTTGTAATTTGCAGAAAATGCA
AGAAAAGTTG
>P03174950_14T_U9R9

-----CAATTTTGGCTGGTGGAAATCTTTGTTTTT-CTT-GTTTTT-
CCCTGCATCCAAGCAGGGCCATAGTTCTTGTAATTTGCAGAAAATGCAAGAAAAC
TTG

Supplementary data Floras.

- Aeschiman, D., Burdet, H.M., 1989. Flore de la suisse et des territoires limitrophes. Le nouveau Binz. Editions du Griffon, Neuchâtel.
- Ball, J., 1878. Spicilegium florum maroccanarum. J. Linn. Soc. 16, 281–772.
- Barceló y Combis, F., 1879. Flora de las islas baleares. Segunda de un diccionario de las plantas espontáneas y de las cultivadas. PJ Gelabert, Palma de Mallorca.
- Battandier, J.-A., Trabut, L.-C., 1902. Flore analytique et synoptique de l'Algérie et de la Tunisie. Edition Vve Girault, Alger.
- Chambouleyron, M., 2012. Contribution à la connaissance de la flore de la Péninsule Tingitane (Maroc). Lagasalia 32, 35–227.
- Chater, A.O., Walters, S.M., Akeroyd, J.R., 1993. Flora Europaea. Cambridge University Press, Cambridge.
- Clapham, A.R., Tutin, T.G., Warburg, E.F., 1952. Flora of the British Isles. CUP Archive, Cambridge.
- Coste, H., 1900. Flore descriptive et illustrée de la France, de la Corse et des contrées limitrophes. Klincksieck, Paris.
- Coutinho, A.X.P., 1939. A Flora de Portugal (Plantas vasculares). Disposta em chaves dicotómicas. 2e Edition. Bertrand (Irmaos), Ltd, Lisboa.
- Davis, P.H., 1965. Flora of Turkey and the East Aegean Islands. Edinburgh University Press, Edinburgh.
- Debeaux, M.O., 1894. Flore de la Kabylie du Djurdjura (Algérie). Klincksieck, Paris.
- Degen, A., 1937. Flora Velebitica. Academy of Sciences, Budapest.
- Dostál, J., Novák, F.A., 1950. Květena ČSR. Přírodovědecké nakladatelství, Praha.
- Castroviejo, S., 1986. Flora Ibérica. Real Jardín Botánico. CSIC, Madrid.
- Hayek, A., 1927. Prodromus Florae Peninsulae Balcanicae. Dahlem-Berlin.
- Jahandiez, E., Maire, R., 1932. Catalogue des plantes du Maroc. Minerva, Paris: Lechevalier, Alger.
- Jalas, J., Suominen, J., 1988. Atlas Florae Europaeae: Distribution of Vascular Plants in Europe (Vol. 3). Cambridge University Press, Cambridge.
- Jeanmonod, D., Gamisans, J., 2007. Flora corsica. Edisud, Aix-en-Provence.
- Jordanov, D., 1970. Flora Reipublicae Popularis Bulgaricae. Aedibus Academiae Scientiarum Bulgaricae, Serdicae.
- Knoche, H., 1921. Etude phytogéographique sur les Iles Baléares. Roumégous et Déhan, Montpellier.
- Lawalrée, A., 1955. Flore générale de Belgique. Spermatophytes. Ministère de l'Agriculture, Jardin Botanique de l'Etat, Bruxelles.
- Maire, R., 1952. Flore d'Afrique du Nord. Lechevalier, Paris.
- Meikle, R.D., 1977. Flora of Cyprus. The Bentham-Moxon Trust, Royal Botanic Gardens Kew.

- Mouterde, P., 1966. Nouvelle flore du Liban et de la Syrie. Tome I. Texte. Dar-el-Machreq, Beyrouth.
- Mouterde, P., 1970. Nouvelle flore du Liban et de la Syrie. Tome II. Texte. Dar-el-Machreq, Beyrouth.
- Nevski, S.A., 1937. Ranunculaceae. In: Komarov, V.L. (ed.) Flora of the U.S.S.R. Akademii Nauk SSSR, Leningrad, pp. 79–143.
- Oberdorfer, E., 1990. Pflanzensoziologische Excursion Flora. Eugen Ulmer GmbH & Co, Hohenheim.
- Pignatti, S., 1982. Flora d'Italia. Edagricole, Bologna.
- Post, G.E., 1896. Flora of Syria, Palestine, and Sinai, from the Taurus to Ras Muhammad, and from the Mediterranean sea to the Syrian Desert. Syrian Protestant College, Beirut.
- Pottier-Alapetite, G., 1979. Flore de la Tunisie. Publications scientifiques Tunisiennes, Imprimerie Officielle de la République Tunisienne, Tunis.
- Qaiser, M., 1984. Flora of Libya: Vol. 108 Ranunculaceae. Al Faateh University, Faculty of Science, Department of Botany, Tripoli.
- Quézel, P., Santa, S., 1962. Nouvelle flore de l'Algérie et des régions désertiques méridionales. CNRS, Paris.
- Riedl, H., 1992. *Nigella* L. In: Iranshahr, M., Rechinger, K.H., Riedl, H. (Eds.) Flora Iranica, Vol. 171. Akademische Druck- und Verlagsanstalt, Vienna, pp. 24–34.
- Riedl, H., Nasir, Y.J. 1991. Flora of Pakistan. No. 193. Ranunculaceae. University of Karachi, National Herbarium, Pakistan Agricultural Research Council, Islamabad.
- Săvulescu, T., Buia, A., 1953. Flora Republicii Populare Române (Vol. 2). Editura Academiei Republicii Populare Române, București.
- Schlosser, J.C., Farkas-Vukotinović, L., 1869. Flora Croatica. Alberc & Fiedler, Zagreb.
- Stojanov, N., Stefanov, B., 1948. Flora na Bălgarija, ed. 3. Nauka I Izkustvo, Sofia.
- Strid, A., Tan, K., 2002. Flora Hellenica Volume 2. Koeltz Scientific Books, Koenigstein.
- Täckholm, V., 1956. Students' flora of Egypt. Anglo-Egyptian Bookshop, Cairo.
- Tison, J.-M., de Foucault, B., 2014. Flora Gallica. Flore de France. Société Botanique de France. Biotope Editions, Mèze.
- Zohary, M., 1966. Flora Palaestina. Part I. Israel Academy of Sciences and Humanities, Jerusalem.