Molecular approaches reveal speciation between red and blue flowered
plants in the Mediterranean *Lysimachia arvensis* and *Lysimachia monelli*
(Primulaceae)

Jiménez-López F.J., Viruel J., Arista M. , Ortiz P.L. and Talavera M.

Abstract

Flower colour constitutes a pivotal evolutionary force in speciation. The Mediterranean *Lysimachia arvensis* and *L. monelli* are morphologically variable species having both blue or red flowered plants. Previous studies suggested that *L. arvensis* plants differing in colour are diverging lineages, but this variation has not been considered in a phylogenetic context. We reconstruct the phylogenetic signal and the ancestral states of flower colour of Mediterranean *Lysimachia* species by using nuclear (ITS) and three plastid markers. Blue and red specimens are nested in two independent clades in the ITS tree, thus supporting that *L. arvensis* and *L. monelli* are polyphyletic, whereas low phylogenetic resolution was found in plastid markers. Blue-flowered *L. arvensis* is reconstructed sister to *L. talaverae* in a monophyletic clade sister to the remaining *Lysimachia*. Red-flowered *L. arvensis* is reconstructed sister to red-flowered *L. monelli* in a monophyletic clade sister to blue-flowered *L. monelli* and *L. foemina*. Our results suggest that colour lineages in *L. arvensis* and *L. monelli* constitute different species, but flower colour did not promote the separation of these lineages. We propose a new name for blue-flowered *L. arvensis* (*L. loeflingii*) and a new combination for red-flowered *L. monelli* (*L. collina*).

Key words: *Anagallis*, nuclear and plastid DNA regions, phylogeny, molecular taxonomy

Introduction

The systematic circumscription of the taxa belonging to the *Lysimachia* complex has been thoroughly studied since Linnaeus descriptions (Harborne, 1968; Källersjö, Bergqvist & Anderberg, 2000; Martins, Oberprieler & Hellwig, 2003; Hao et al., 2004; Manns & Anderberg, 2007a, 2007b; Zhang et al., 2012). Traditionally, *Anagallis* L. had always been considered a genus closely related to *Lysimachia* L. However, phylogenetic analyses based on the nuclear ITS and plastid data resolved species belonging to either *Anagallis* and *Lysimachia* intermingled across different clades (Manns & Anderberg, 2005). For example, the Mediterranean *A. arvensis* L. and *A. monelli* L. were more closely related to species of *Lysimachia* and other genera such as *Pelletiera* A.St.-Hil., *Asterolinon* Hoffmanns. & Link, and *Trientalis* L., than to other *Anagallis* species (Anderberg, Manns, & Källersjö, 2007). These phylogenetic results led to the inclusion of some taxa previously included in *Anagallis* within *Lysimachia* (Manns & Anderberg, 2007b, 2009); and new combinations were subsequently proposed for *L. arvensis* (L.) U. Manns & Anderb. and *L. monelli* (L.) U. Manns & Anderb. (Manns & Anderberg, 2009).

All these systematic readjustments are congruent with the high phenotypic variability found within the former genus *Anagallis*, in which morphological traits, such as leaf shape or colour, size and presence of glands in flowers, hinder the definition of taxonomic boundaries. The Mediterranean *L. arvensis* and *L. monelli* show a high morphological and ecological diversity, and consequently a wide number of infraspecific taxa have been described within each species. For example, diploid plants ($2n=20$; Šveřepová, 1972; Kress, 1969; García Pérez et al., 1997) with small size and small flowers, and specific ecology (Bolòs & Vigo, 1996; Gibbs & Talavera,

2001) originally recognized as *A. parviflora* Hoffmanns. & Link (Hoffmannsegg & Link, 1813-1820), were classified as infraspecific taxa of *L. arvensis*, or currently recognized as *L. talaverae* L. Sáez & Aymerich (Aymerich & Sáez, 2015). Likewise, plants morphologically very similar to *L. arvensis* but with four cells in the marginal glands of the petals (instead of three) were first recognized as *A. arvensis* subsp. *foemina* (Mill.) Schinz & Thell., and later described as a different species: *L. foemina* (Mill.) U. Manns & Anderb. (Manns & Anderberg, 2009). This taxonomic reorganization was based on phylogenetic results that reconstructed *L. monelli* as sister to *L. foemina* (Manns & Anderberg, 2007a). Several taxa have also been recognized from the infraspecific variability within *L. monelli*, such as plants with narrow and linear leaves which were considered as *L. monelli* subsp. *linifolia* (L.) Peruzzi (Peruzzi, 2010). Likewise, plants with small and wide leaves and short internodes were recognized as *L. monelli* subsp. *maritima* (Mariz) Carlón et al. (Carlón et al., 2014).

Both *L. arvensis* and *L. monelli* show flower colour polymorphism with blue and red-flowered plants (Ferguson, 1972; Pujadas, 1997). This trait further increased the taxonomic complexity because numerous taxa have been described in relation to flower colour. The original description of *A. arvensis* was based on a red-flowered specimen (Linnaeus 1753), whereas the blue morph was first described as *A. latifolia* L. (Linnaeus 1753), then as *A. caerulea* L. (Linnaeus 1759) but it is now accepted as *L. arvensis* var. *caerulea* (L.) Turland & Bergmeier (Bergmeier et al., 2011). Blue and red plants of *L. arvensis* show different geographical distributional pattern, pure blue or mixed populations appear chiefly in dryer Mediterranean localities while pure red populations predominate in more temperate areas (Arista et al., 2013). In addition, studies using microsatellite markers showed the separation of both colour morphs in

two groups (Jiménez-López et al., 2019a). Both colour lineages of *L. arvensis* differ in other traits as flower phenology or type of herkogamy (Arista et al., 2013; Jiménez-López et al., 2020), and salmon-flowered plants resulting from the cross between red and blue individuals are rare in mixed populations (Jiménez-López et al., 2020).

In the case of *L. monelli*, the species description was based on a blue-flowered specimen (Linnaeus 1753), and the red-flowered individuals were first described as *A. collina* Schousb. (Schousboe, 1800), then recognized by various authors as infraspecific taxa of *Anagallis monelli* or *A. linifolia* L. (Linnaeus, 1762) (see taxonomic implications). Currently, these names are accepted within *Lysimachia monelli*. Populations of *L. monelli* with blue-flowered plants are restricted to the Iberian Peninsula, Morocco, Algeria, northwest of Libia (Tripolitania) and Tunisia (Jahandiez & Maire, 1934) while the red-flowered plants grow mainly in Morocco, Sardinia and northeast of Spain (Willkomm, 1870; Jahandiez & Maire, 1934).

The huge morphological variability of the Mediterranean *Lysimachia* complex is currently reduced to four species: *L. arvensis*, *L. monelli*, *L. foemina* and *L. talaverae* (Manns & Anderberg, 2009; Aymerich & Sáez, 2015). *Lysimachia arvensis* is a Mediterranean species that at present is distributed across the world. It is annual, self-compatible (Gibbs & Talavera, 2001) and tetraploid ($2n=40$; revised in Pastor, 1992; García Pérez et al., 1997; Monein, Atta & Shehata, 2003). *Lysimachia monelli* is distributed exclusively in the west of the Mediterranean Region. It is perennial, self-incompatible (Gibbs & Talavera, 2001; Talavera et al., 2001; Freyre & Griesbach, 2004) and diploid ($2n=20$; Kress, 1969; Valdés, 1970; Šveřepová, 1972; García Pérez et al., 1997). *Lysimachia foemina* is also a Mediterranean species but, as *L. arvensis*, at present is distributed across the world. It is annual, self-compatible (Marsden-Jones, 1935; Marsden-Jones & Weiss, 1938) and tetraploid ($2n=40$; revised in Pastor,

1992). *Lysimachia talaverae* is endemic to the western Mediterranean, it is annual, self-compatible (Gibbs & Talavera, 2001) and diploid ($2n=20$; Šveřepová, 1972; García Pérez et al., 1997).

Previous phylogenetic studies have scarcely explored the potential implications of colour polymorphism in taxa delimitation because it was considered as part of the intraspecific variation, and usually only one colour per species was sampled in molecular analyses (Martins et al., 2003; Manns & Anderberg, 2005; Anderberg et al., 2007; Manns & Anderberg, 2007b; Yan et al., 2018). However, Manns & Anderberg (2007a) reconstructed a phylogenetic tree including samples of both colours of *L. arvensis* from a single population, which were resolved as monophyletic using plastid markers, but appeared in independent clades in the ITS-based phylogenetic tree. All these evidences suggest that larger sampling of the two colour lineages for both *L. arvensis* and *L. monelli* is required in phylogenetic studies to clarify the implications of flower colour in taxa boundaries. In this context, our study aims to (1) define the phylogenetic identity of the blue and red lineages of *L. arvensis* and *L. monelli*, (2) establish the role of flower colour in the divergence of these taxa and (3) propose new names or combinations in accordance with the results of phylogenetic analyses where necessary.

Material & methods

Plant material

Freshly leaf material from 28 populations of *L. arvensis*, 12 of *L. monelli*, three of *L. foemina* (LF) and four of *L. talaverae* (LT) (Fig. 1, Table S1) were collected and dried in silica-gel (Chase & Hills, 1991). This sampling represents nine pure red and 12 pure blue populations and seven mixed (red and blue) populations of *L. arvensis*

(LA_R and LA_B, respectively). Sampling of *L. monelli* included four pure red (LM_R) and eight pure blue populations (LM_B). Also, one sample of *L. azorica* Hornem. ex Hook. (LZ), and one sample of two populations of *L. linum-stellatum* (L) Hoffmanns. & Link (LLS). Two additional related species were included in the phylogenetic analyses and used as outgroups (Table S1).

DNA isolation, amplification and sequencing

Total genomic DNA was isolated with Invisorb Vegetal DNA Kit HTS 96 (Invitex, Spain), with modifications following Jiménez-López et al. (2016). The quality of the extracted DNAs was checked in 1% TAE-agarose gel, and DNA concentration was estimated using a NanoDrop DS-11 Spectrophotometer (DeNovix).

Both ITS and plastid markers were used for phylogenetic reconstructions. The ribosomal ITS region partial18S–5.8S–partial28S was amplified using ITS5 and ITS4 primers published by White et al. (1990). We tested 19 plastid markers in eight samples of *L. arvensis* (four blue and four red specimens), but only three of them were polymorphic. The polymorphic plastid markers *rps16-trnK*, *rpl32-trnL* were amplified following Shaw et al. (2007), and *trnH-psbA* following Hamilton (1999). PCR amplifications were performed following Jiménez-López et al. (2016) using 4ng of DNA as input and the amplified fragments were checked in 2% TAE-agarose gels. An Applied Biosystems Veriti™ Thermal Cycler was used with the following programmes: for *rps16-trnK* and *rpl32-trnL*: first step, one cycle of 80 °C for 5 min; a second step, 35 cycles with three consecutive conditions: 95 °C for 1 min, increase from 50 to 65 °C during 1 min and 65°C for 3 min; third step, one cycle of 65°C for 5 min and then 4 °C. For *trnH-psbA* the conditions were first step, one cycle of 96 °C for 5 min; second step, 35 cycles with three consecutive conditions: 95 °C for 45 sec, 53 °C for 1 min and 72 °C for 30 sec; third step, 1 cycle of 72 °C for 5 min and then 4

°C. After purification with ExoSAP-IT (Roche, Spain), fragments were sequenced in an ABI 3730 machine, BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA), at STAB Vida Lda. (Oeiras, Portugal). Sequences were edited using Geneious R10 (Biomatters Ltd, Auckland, New Zealand). Alignments were conducted in MEGA ver. 6 (Tamura et al., 2013) using the algorithm Clustal IW (Thompson, Gibson & Higgins, 2002), and manually edited.

Phylogenetic inference

The Akaike information criterion (AIC) was used to estimate the evolution model that better fit each DNA matrix using jModeltest 2.1.4 (Darriba et al., 2012). Phylogenetic relationships were inferred based on maximum parsimony inference using PAUP ver. 4.0b 10 (Swofford, 2003) and maximum likelihood methods using RAxML Ver. 8 (Stamatakis, 2014). Bootstrap support (BS) was estimated with 500,000 bootstrap replicates following full heuristic searches. Bayesian inference was conducted using MrBayes on X_{SEDE} (3.2.7a) (CIPRES Science Gateway; Miller, Pfeiffer & Schwartz, 2010) and BEAST v.2.5 (Bouckaert et al., 2019). We use coalescent models (Heled & Drummond, 2010) and we ran 30 independent analyses with 50 million generations, sampling every 5000. The phylogenetic reconstructions were calculated for each plastid and nuclear markers independently, and for a combined dataset including the three plastid datasets. Following Viruel et al. (2016) clades with bootstrap support (BS) of 75-100 or posterior probabilities (PPS) of 0.95-1.0 were considered moderately to strongly supported. Phylogenetic reconstruction with ITS was performed with and without the sample LA19_B according to the results of the recombination analysis (see below and results section).

In order to test for incongruent topologies between the independent DNA data matrices, a partition homogeneity test was performed (Farris et al., 1994,

Symonds & Lloyd, 2003). The Incongruence Length Difference (ILD) test of Farris et al. (1994) implemented in PAUP ver.4.0b 10 (Swofford, 2003) was performed through 1,000 random-order-entry replicates to estimate if the nuclear and the three plastid data sets were significantly different from random partitions of the same size.

Recombination Analysis

To test possible recombination events in the ITS sequences of all the samples, the seven methods (RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan & 3SEQ) implemented in RDP4 v.484 were applied to determinate parental and recombinant sequences with their probability scores for hypothetical recombination events (Martin et al., 2010; Viruel et al., 2018).

Ancestral state reconstruction

To estimate the basal state of flower colour, we performed ancestral state reconstruction of colour classes using stochastic character mapping (SIMMAP; Bollback, 2006) with the package ‘phytools’ (Revell, 2012) implemented in R (R Core Development Team, 2015). We performed the analyses on the RAxML phylogenetic tree obtained from ITS dataset. Because we included more than one sample per species, we also pruned the additional samples for the same species in the case of monophyly. For SIMMAP analyses, we ran 1000 simulations per tree. Ancestral state reconstructions for all additional nodes were estimated with parsimony and SIMMAP methods only. Ancestral states reconstruction was shown only on major clades.

Divergence-time estimation

In addition, we used molecular clock approaches to estimate divergence times among the main clades. We selected one representative sequence per species, together with sequences of other *Lysimachia* and related species from GenBank (Table S1).

In this case, the phylogenetic reconstruction was inferred with maximum likelihood methods with RAxML (Stamatakis, 2014). We used a Bayesian uncorrelated-lognormal relaxed-clock approach (Drummond et al., 2006) as implemented in BEAST v.2.5 (Bouckaert et al., 2019). Node ages and absolute branch substitution rates were estimated using two relaxed clock methods that represent two contrasting approaches to estimate divergence times (Magallón, Hilu & Quandt, 2013). We used both Yules speciation and birth–death tree priors (Heled & Drummond, 2012). We ran 30 independent analyses (50 million generations each, sampled every 5000). Tracer 1.7.1 software (Rambaut et al., 2018) was used for convergence and satisfactory Effective sample size (ESS) values (>200) and verify that a burn limit of 10% was appropriate. Subsequent parameter distributions were obtained by combining the independent MCMCs with LogCombiner 1.10.4, and calculating the maximum credibility tree using TreeAnnotator 1.10.4 (Rambaut et al., 2018).

Fossil records are scarce because these plants lack woody parts, and both seed and pollen production are low. The genus *Maesa* Forssk is at the root of the *Primulaceae* phylogeny (Anderberg et al., 2007; Yesson, Toomey & Culham, 2009) and its pollen is known from the Early Eocene (55.8 Ma; Brook, Burney & Coward, 1990; Kershaw, Bretherton & Kaars, 2007) according to The Paleobiology Database (<http://paleodb.org/>). Furthermore, fossil seeds similar to those of *Lysimachia vulgaris* L. and its close relatives (Oh et al., 2008) are found from the Middle Miocene (12–16 Ma; Friis, 1985). We therefore used the pollen record to set the age prior for rooting and seed fossil to set the age prior for *L. vulgaris*. A secondary calibration was

applied to the crown age of core *Lysimachia* using an estimated mean age of 30 Ma from previous phylogenetics reconstructions (Yesson et al., 2009, Zanne et al., 2014). The most generally appropriate prior for fossil calibrations is the lognormal distribution (Ho & Phillips, 2009). Then, for the crown age of core *Lysimachia* node we used a standard lognormal prior distribution with mean of 30 Ma and standard deviation (SD) of 4 Ma following Yang et al. 2018. In order to test the effect of the different calibration points and tree prior approaches, we also performed analyses using only one fossil calibration, only secondary calibration and both calibration points. We then performed a regression between the node ages of same clades from all analyses.

Results

Phylogenetic reconstructions and chronogram inference

Fifty-two high quality sequences were obtained for ITS, and 37 for the plastid regions. The pDNA regions were 675-750 bp long for *rps16-trnK*, 580-630 bp for *rpl32-trnL* and 365-445 bp for *trnH-psbA*. The ITS1 and ITS2 fragments ranged from 600 to 737 bp. The G+C content in all samples was 22.22% for *rps16-trnK*, 27.25% for *rpl32-trnL*, 29.04% for *trnH-psbA* and 57.88% for ITS on average. The G+C contents were significantly different between *L. arvensis* and *L. monelli* for ITS ($p=0.001$; 58.86% and 56.46%, respectively) and for *rps16-trnK* ($p=0.015$; 22.17% and 22.30%, respectively). No significant differences were found on the G+C contents among blue and red individuals of *L. arvensis* ($p=0.482$) or *L. monelli* ($p=0.338$).

The parsimony, RAxML and Bayesian trees reconstructed the same topology for all analysis for the ITS region (Fig. 2). Likewise, for each pDNA region separately and

for the combined matrix, the tree reconstructed showed a very similar topology in all analyses and markers (Figs. 3). However, the ILD test showed significant differences between nuclear and plastid regions ($p=0.010$), indicating that there was an inconsistency between these markers. Thus, phylogenetic analyses were performed for the ITS region separately from the concatenated dataset representing the three plastid markers.

The phylogenetic tree reconstructed with the concatenated pDNA matrix showed two clades (Figs. 3). The first clade (94 BS, 1.000 PPS) comprised all the red and some blue individuals of *L. arvensis* and *L. talaverae*. The second clade (86 BS, 0.890 PPS) included blue individuals of *L. arvensis*, *L. monelli* (blue and red) and *L. foemina*. This topology was also observed for each plastid marker independently, although the support bootstrap was lower in all cases.

The RAxML tree constructed with ITS showed two main clades (Fig. 2). In the Clade I (C1, 98 BS, 1.000 PPS), *L. arvensis* blue samples (95 BS, 1.000 PPS) were sister to *L. talaverae* (89 BS, 0.999 PPS), whereas the samples of *L. arvensis* from SW of Spain formed a clade (82 BS, 0.980 PPS). The Clade II (C2, 87 BS, 1.000 PPS) comprised two subclades, named A and B. In the subclade A (SCA, 85 BS, 0.935), red *L. arvensis* samples (84 BS, 0.976 PPS) were sister to the red *L. monelli* samples (85 BS, 1.000 PPS). In the subclade B (89 BS, 0.990), blue *L. monelli* samples (83 BS, 0.984 PPS) were sister to *L. foemina* (95 BS, 1.000 PPS). Therefore, blue and red individuals of both *L. arvensis* and *L. monelli* appeared in different groups. It is noteworthy that blue and red individuals from mixed populations of *L. arvensis* appeared in independent clades (Figs. 2). In addition, most of the ancestral reconstruction states showed high probability of blue-flowered plants until the SCA, where red flowers appeared (Fig. 2).

Similar time estimates and topologies were obtained using the three different calibration ages and different tree priors for the stem age of *Lysimachieae* tribe (Fig. 4). Divergence time results based on these different strategies were all very similar to each other. Our primary analyses (Fig. 4) used the full tree with the Yule prior and all calibration points. The correlation between the node ages of clades from the primary tree and alternative trees based on different calibration schemes and tree priors were significant for all cases ($r^2 = 0.978-0.891$, $p=0.000$) The maximum clade credibility (MCC) tree of *Lysimachieae* tribe showed three clades in *Lysimachia* (Fig. 4): one with *L. arvensis*, *L. monelli* and related species (BS 100%; 1.000 PPS, Clade A), other with *L. vulgaris*, *L. punctata* and related species (0.940 PPS, Clade B), and a third clade of *Lysimachia*, with *L. tenella* or *L. minima* and other related species (BS 100%; 0.990 PPS, Clade C). The latter clade appeared closest to *Cyclamen* sp. (BS 100%; 1.000 PPS). According to the relaxed molecular clock and Yules tree priors, the common ancestor of blue plants of *Lysimachia arvensis* and *L. talaverae* diverged ca. 1.5 Mya (95% HPDs 0.2-2.5), and the divergence of blue plants of *L. monelli* and *L. foemina* likely occurred ca. 0.6 Mya (95% HPDs 0.1-1.6) and the split between red plants of *L. arvensis* and red plants of *L. monelli* was estimated ca. 1.2 Mya. (95% HPDs 0.2-2.4). Furthermore, the common ancestor between *L. arvensis* red-*L. monelli* red and *L. monelli* blue/*L. foemina* subclades likely diverged ca. 1.4 Mya (95% HPDs 0.6-2.9) and the divergence between these clades and *L. arvensis*/*L. talaverae* clade occurred ca. 3.6 Mya (95% HPDs 1.3-7.5) (Figs. 4 and S7). Two events of divergence occurred during Miocene period between the different clades of *Lysimachia*. The first one happened about 28.9 Mya (95% HPDs 16.9-47.2), when diversifying the clade C from the other *Lysimachia* clades, and the second one happened about 24.2 Mya (95% HPDs 13.1-42.6) dividing clades A and B (Fig. 4).

Recombination Analyses

Four of the seven recombination methods detected two recombinant regions in the sample LA19_B (*L. arvensis* blue). An insertion of 18 nucleotides at position 16 from *L. arvensis* and an insertion of 25 nucleotides at position 133 from *L. foemina* were observed. Therefore, this sample was discarded from the tree in Figure 2.

Discussion

Our phylogenetic results based on the nuclear ITS region demonstrates that blue and red individuals of *Lysimachia arvensis* and *L. monelli* are independent taxa (Fig. 2). The presence of different ITS sequences in the red and blue flowered individuals of *L. arvensis* collected from the same populations had already been reported by Manns & Anderberg (2007a). The high support of our ITS results (89 BS, 0.999 PPS) suggests that blue and red lineages of *L. arvensis* belong to different taxa, and in addition, both colour taxa are much more close to other species than between them. On one hand, the blue flowered-plants of *L. arvensis* are sister to *Lysimachia talaverae* (Clade I). This last species has already been recognized with the rank of species based only on morphological, ecological and karyological traits (Aymerich & Sáez, 2015) but not on molecular tools. Our study further validates the species rank status of *L. talaverae* using phylogenetic inference. On the other hand, red individuals of *L. arvensis* are sister to red individuals of *L. monelli* (Subclade A). Likewise, the two colour lineages of *L. monelli* should be also considered as different taxa, being the blue lineage of *L. monelli* sister to *L. foemina* (Subclade B), supporting the independence of *L. foemina* from *L. arvensis*, in accordance with Manns & Anderberg (2007a).

Colour lineages of both species appeared together with pDNA markers. Hybridization has been proposed as causes of incongruences between pDNA and nuclear

phylogenies (Wendell & Doyle, 1998; Semerikova & Semerikov 2016). Hybridization involves past or present contact or hybrid zones (Petit, Bretagnolle & Felber, 1999), where lineages with divergent genomes have the potential to exchange genes (Souissi et al., 2017). However, as a result of this exchange it would be difficult to find a clear separation between lineages, contrasting with the strong separation between *Lysimachia* species using the nuclear ITS (Fig. 2). Only one sample (LA19B, blue *L. arvensis*) showed hybridization evidence according to the recombination analyses. The position of this sample in the ITS tree would suggest a possible hybridization between blue *L. arvensis* and *L. foemina*. These two taxa co-occur in some populations in Europe and their flowers are very similar, thus they possibly share pollinators. However, the recombination analysis pointed that this hybrid resulted from the crossing between red *L. arvensis* and *L. foemina*. Thus, that hybridization was originated from an individual with blue flowers very similar to the blue *L. arvensis* (as it was initially considered in the sampling). This hybrid (*L. arvensis* x *L. foemina*) has already been described as *A. x doerfleri* Ronniger (Dörfler, 1903). (= *Lysimachia* x *doerfleri* (Ronniger) Stace). However, hybrid natural populations stabilisation seems difficult, since experimental crosses carried out by other authors frequently resulted in sterile F₁ progeny (Marsden-Jones, 1935; Kollmann & Feinbrun, 1968; Šveřepová, 1972).

Our results corroborate the maternally haplotype inheritance of the plastomes (Wolfe, Li & Sharp, 1987), and their sequences are frequently identical among closely related species in many plant taxa (Wolf, Murray & Sipes, 1997; Matsumura et al., 2009). The incongruent pattern observed between the two phylogenetic trees (Degnan & Rosenberg, 2009) can be also explained by the coalescent theory (Nordborg, 2001; Wakeley, 2009) because two or more lineages can coexist in the same ancestral

population (Degnan & Rosenberg, 2009). Thus, the common ancestor of this group would have had the two pDNA haplotypes, which are currently present in one of them in red individuals of *L. arvensis* and *L. talaverae*, the other in *L. monelli* and *L. foemina*, and both haplotypes in blue individuals of *L. arvensis*. Later, a segregation of colour lineages of *L. arvensis* and *L. monelli* likely happened. That segregation could have been promoted by geographic separation, assortative mating mediated by pollinators or by differential tolerance to abiotic factors as has been described in other species (Kirkpatrick, 2000; Strauss & Whittall, 2006; Hopkins & Rausher, 2012; Wang et al., 2013). In *L. monelli* geographic separation could be contributing to lineage separation. However, in *L. arvensis* genetic flow is potentially possible between lineages in mixed populations although the reproductive isolation between them is high (Jiménez-López et al., unpublished data); in fact, the results of nuclear phylogenetic reconstruction points towards a clear isolation of both colour lineages.

Flower colour constitutes a pivotal evolutionary force to speciation in several groups of plants (Carlson & Holsinger, 2015; Ellis & Field, 2016; Takahashi, Takakura & Kawata, 2016; Narbona et al., 2018) and it has been proposed as a “magic trait”, that is a trait “encoded by genes subjected to divergent selection that affect pleiotropically reproductive isolation” (Servedio et al., 2011). However, according to ancestral state reconstruction, flower colour does not seem the trait promoting divergence between lineages of *L. arvensis* or in those of *L. monelli*. Ancestral states reconstruction of flower colour points to a common blue-flowered ancestor of this group and the transition to red-flowered plants occurred only once, at the divergence node of red *L. arvensis* and *L. monelli*. This kind of transition from blue to red in flowers is quite frequent due to the inactivation of a branch of the anthocyanin pathway (Rausher, 2008). In other plant groups, red-flowered species usually derived from blue-flowered

species (Kay et al., 2005; Wilson et al., 2007; Rausher, 2008; Wessinger & Rausher, 2012). In our study, the blue ancestor likely originated two lineages, one entirely blue (which includes the blue lineage of *L. arvensis*) and another that subsequently separated into a blue and another red subclade (which includes the red lineage of *L. arvensis* and *L. monelli*). This result indicates that flower colour cannot have been the trigger for the current speciation of *L. arvensis* or *L. monelli* colour lineages.

The location of each colour lineage of *L. arvensis* and *L. monelli* in the phylogenetic reconstruction suggests independent origins for colour lineages of both species. According to the phylogenetic tree constructed with ITS, both a diploid and a tetraploid taxa appear together in each clade: blue lineage of *L. monelli* (2n) with *L. foemina* (4n), *L. talaverae* (2n) with blue lineage of *L. arvensis* (4n), and red lineage of *L. monelli* (2n) with red lineage of *L. arvensis* (4n). That configuration strongly suggests the role of polyploidy in the speciation process of this complex. Frequently speciation events in flowering plants involve hybridization processes and subsequent modification of divergent genomes (Stebbins, 1971). Polyploid speciation implies the complete duplication of the genome after hybridization (allopolyploidy). In addition, new species can be produced not by hybridization, but by duplication of complete genomes which also leads to polyploidy (autopolyploidy). The configuration observed in the phylogenetic reconstruction suggests that both colour lineages of *L. arvensis*, which are tetraploid, have been originated by autopolyploidy. Red *Lysimachia arvensis* could derive from red *L. monelli* (diploid) and blue *L. arvensis* could derive from *L. talaverae* (diploid with blue flowers). Moreover, *L. foemina* (tetraploid with blue flowers) could derive from blue *L. monelli* (diploid). Although this is a hypothesis, at least for *L. arvensis*, both colour lineages present a karyotype with four equal copies from a set of chromosomes (Monein et al., 2003), supporting the

autopolyploid origin of these taxa. This would indicate the presence of multiple polyploidization events in the clade studied, something that has previously been found in other genera (eg. Rieseberg et al., 2003; Wang et al., 2006; Wood et al., 2009; Soltis et al., 2015; Padilla-García et al., 2018). Divergence-time estimation between the diploid and tetraploid taxa, ca 1Mya (Fig. 4) and the support of the phylogenetic reconstruction (Fig. 2) point to non-recent tetraploidization events, and the stability in the genomic sequences studied also suggest an autopolyploid origin, in accordance with current evolutionary theories on polyploids (Otto & Whitton, 2000; Soltis, Visger & Soltis, 2014; Soltis et al., 2015). However, this role of polyploidy in the evolution of the studied *Lysimachia* would need a more directed study.

In conclusion, the two colour morphs of *L. arvensis* constitute independent lineages and their current isolation is evident. Blue *L. arvensis* is a more adapted taxon to the arid conditions of the Mediterranean while red lineage is a more mesic taxon with a greater distribution area, co-occurring with the blue lineage in some Mediterranean areas (Arista et al., 2013). These two colour lineages also separate when using nuclear microsatellite markers (Jiménez-López et al., 2019a). Similarly, the two colour lineages of *L. monelli* are also independent entities with allopatric distribution. A phylogeographic approach would be necessary to clarify the geographical origin of both colour lineages. Although our results indicated that flower colour has not triggered recent speciation in this group, it can have an important role as a reinforcement mechanism that prevents the formation of maladapted hybrids (Hopkins, 2012). This could be the situation at least in *L. arvensis*, where pollinators prefer the blue lineage and barely made transitions between flowers of different colours (Ortiz et al., 2015; Jiménez-López et al., 2019b). The low frequency of hybrids in mixed-colour populations (Jiménez-López et al., 2020) and the lack of

hybrid samples found in the present study support the importance of colour as reinforcement mechanism.

Taxonomic implications

Our results have taxonomic implications for the colour lineages of *L. arvensis* and *L. monelli* as each lineage should be defined as different taxa with morphological, genetic and geographic identity. Thus, we propose the following names or combinations in accordance with molecular analyses:

-Red plants of *Lysimachia arvensis* should maintain this name because Linnaeus in 1753 described the species from red-flowered plants.

Lysimachia arvensis (L.) U. Manns & Anderb. in Willdenowia, 39(1):51.
(2009) ≡ *Anagallis arvensis* L., Sp. Pl.: 148. 1753, [basion.] ≡ *Anagallis arvensis* var. *phoenicea* Gouan, Fl. Mons.: 24 (1764), nom. Illeg. ≡ *Anagallis phoenicea* (Gouan) Scop., Fl. Carn. ed. 2, 1: 139 (1777), nom. Illeg. ≡ *Anagallis arvensis* subsp. *phoenicea* (Gouan) Wallmann in Ber. Bayer. Bot Ges. 9: 44 (1904), nom. inv.

Ind. Loc.: “Habitat in Europae arvis”.

Lectotype designated by Dyer & al. 1963: 14: Herb. Linn. 208.1 (LINN)

-Blue plants of *Lysimachia arvensis* should be called with the specific epithet *latifolia* as it was the first name employed by Linnaeus in 1753 for plants with blue flowers. However, the epithet *latifolia* already exists within *Lysimachia* referred to a different taxon [*Lysimachia latifolia* (Hook.) Cholewa in Phytoneuron 28: 1-2 (2014) = *Trientalis latifolia* Hook., Fl. Bor. Amer. 2(9): 121 (1839), a plant described from Washington]. Therefore, we have selected the name *Lysimachia loeflingii* because

Linnaeus in 1753 described blue-flowered plants from materials collected by Loeffling from Spain.

***Lysimachia loefferlingii* F.J. Jiménez-López & M. Talavera, nom. nov. ≡**

Anagallis latifolia L., Sp. Pl. 1: 149 (1753) [syn. subst.] ≡ *A. arvensis* subsp. *latifolia* (L.) Arcang., Comp. Fl. Ital., ed.2: 456 (1894) - *Anagallis arvensis* L. subsp. *arvensis* var. *caerulea* sensu Kollmann & Feinbrun in Notes Royal Bot. Gard. Edingurg 27: 176 (1968), non *A. arvensis* L. var. *caerulea* (L.) Gouan, Fl. Monsp.: 30 (1765)

Ind. loc.: "Habitat in Hispania Loeff."

Lectotype designated here: Herb. Linn. n° 208.3, "H.U.3. latifolia. ex Hispanica folij amplexicauly" [m. Linnaeus]

Anagallis latifolia L. was described by Linnaeus in the 1st edition of his Species Plantarum (1753: 149) indicating: "3. Anagallis foliis cordatis amplexicaulibus, caulibus compressis latifolia/ Anagallis hispanica, latifolia, maximum flore. Tournef. Inst. 142/ Cruciata Montana minor, flore caeruleo. Barr. Ic. 584/ Habitat in Hispania Loeffl. "

Linnaeus described that plant indicating the colour of the flowers, "*Corolla caerulea, fondo purpurascens*" and that of the stamens, "*Filamenta purpurea. Antheris oblongis, flavis*".

The reference made by Linnaeus to Tournefort (1719: 142) corresponds, according to Sampaio (1900: 57), to a plant collected by Tournefort in "Ultra San Joan de Foz ad ostium durii" on his travel through Portugal in 1689, as it appears in a manuscript that Tournefort left in the Museu Botanico da Universidade (Coimbra). This name phrase of Tournefort was used by Sampaio (1900: 58) as a synonym of *Anagallis hispanica*

Sampaio, which we consider here synonymous of *Anagallis monelli* L. (see below this species).

The icon 584 of the Lam. 157 by Barrelier (1714: 17), indicated by Linnaeus very faithfully, represents the upper part of a branch with large flowers and with all the lanceolate leaves arranged in whorls of 4 in each node, each with its flower. This plant can also be identified as *Anagallis monelli* L. Therefore, none of the synonyms that Linnaeus gives in *Anagallis latifolia* could be selected as type since both the Tournefort plant and the Barrelier icon have characters very different from those described by Linnaeus in *Anagallis latifolia*. Therefore, the type should be chosen among the material of herbarium that Loeffling sent to his Master from Spain, as Linnaeus indicated in the locotypical indication.

In the different Linnaean herbaria, we did not find any material of *Anagallis latifolia* collected by Loeffling in Spain. In the main herbarium of Linnaeus (LINN) there is a sheet No. 208.3 with the indication, handwritten by Linnaeus "HU" at the base of the plant and at the base of the sheet "3 latifolia". 3 is the order that Linnaeus gave to his description of *Anagallis latifolia*, that is, we are dealing with a material that comes from the crops of the Botanical Garden of Uppsala and that was marked by Linnaeus as material of the first edition of his *Species Plantarum* of 1753, and therefore the material could be the type of the species. In addition, in an attached cut-out to the sheet and handwritten by Linnaeus "Ex Hispanica folio amplexicauly" is indicated. With this statement Linnaeus related the cultivated plant in Hortus Upsaliensis with one from Spain that Loeffling sent him. The material that is in the sheet consists of the upper half of a flowered plant of 20 cm, with quadrangular stems and internodes 2-5 cm; opposite leaves, being the bigger 3 x 2 cm, broadly ovate-lanceolate, sharp, cordate at the base; solitary, axillary flowers, with capillary pedicels of 20-30 mm,

almost the length of the leaves; blue corolla, with petals of c. 7 mm in length, and immature capsules smaller than the calyx. It is evident that this material comes from cultivation as Linnaeus indicated. Linnaeus also made a detailed description of the species on the living plant, since some characters, especially those related to the colours of flowers and filament and anther of the stamens, cannot be described with material found in 'Linnaeus' main herbal sheet. This material from the Uppsala Garden came from seeds that Loeffling periodically sent to Linnaeus (López González, 1990: 43). For all the above, we propose as lectotype of *Anagallis latifolia* L. the only plant found in Linnaeus' main herbarium (LINN) n° 308.3.

-Plants with very small blue flowers that live in wetland environment have been included sometimes in *Anagallis arvensis* thus, a typification proposal is made.

Lysimachia talaverae L. Sáez & Aymerich in Orsis 29: 48 (2015) $\equiv$ *Anagallis parviflora* Hoffmanns. & Link, Fl. Portug. 1 (11) : 325, 326. Tab. 64 (1813-1820) [syn. subst.] $\equiv$ *Anagallis arvensis* subsp. *parviflora* (Hoffmanns. & Link) Arcangeli, Comp. Fl. Ital. ed.2: 456 (1894) = *Anagallis latifolia* “raza” *parviflora* (Hoffmanns. & Link) Merino in Broteria, sér. Bot. 14: 162 (1916), nom. illeg.

Ind. loc.: “Dans les lieux sablonneux aux environs de Comporta”.

Lectotype designated here (iconotype): *Anagallis parviflora* Tab. 64 in Hoffmannsegg & Link, Fl. Portug. 1 (13): 326 (1813-1820).

Epitype: Huelva. Hinojos, Las Porqueras. Lagoon edge. 37° 17' 37" N-6° 25' 15"W. 108 m. 5/5/2014. Leg. F.J. Jiménez & S. Talavera. SEV286467. A sheet of this sample has been used in this study by establish the phylogeny with a nuclear marker (ITS) and three plastic markers.

Lysimachia monelli should be referred exclusively to the blue lineage because it was used by Linnaeus in 1753 to describe this species for the first time.

Lysimachia monelli (L.) U. Manns & Anderb. in Willdenowia, 39(1): 52 (2009) $\equiv$ *Anagallis monelli* L., Sp. Pl. 1: 148 (1753), [basion.].

Ind. Loc.: not indicated by Linnaeus in 1753; "Habitat in Verona" in Linnaeus, Sp. Pl. ed. 2, 1: 212. 1762; As indicated Carlos Pau (1915) the locotypical indication was clearly exposed by Linnaeus when he described the species when indicating his previous synonym, both in his Hortus cliffortianus, 52. 1738: "Crecit forte Gadibus und seminis accepit Joh. Monellus Tornacensis, atque eaden cum Clussius comunicavit anno 1602", as in Hortus upsaliensis, 38. 1748: "Habitatio incerta. Johannes Monellus 1662 Gadibus semina Clussius misit. Hospitatur in Terpidario, perennes".

Lectotype designated by Manns & Anderberg in Willdenowia 39 (1): 52 (2009): Hort. Cliff. 52.2, *Anagallis* 2 (BM000557969).

The lectotype (photo seen) is composed by two flowering stems of c. 30 cm, one undivided and the other branched almost from the base, with numerous flowers but only those of the apex in anthesis, the others in postanthesis; internodes 2-5 cm, longer than the leaves; opposite inferior leaves, of 20-25 x 5 mm, elliptical, attenuate at the base, acute in the apex, superior in whorls of 3 leaves of 15-20 x 3-5 mm, elliptical, each one with a long pedicelled flower; pedicels of 25-35 mm, generally larger than the leaves. These branches come from the garden of George Clifford III (Hartekamp Garden, Holland).

= *Anagallis linifolia* L. Sp. Pl. ed. 2, 1: 212 (1762)

≡ *Anagallis monelli* subsp. *linifolia* Maire in Jahandiez & Maire, Cat. Pl. Maroc.: 562 (1934)

Ind. Loc.: “Habitat in Lusitania, Hispania. Claud Alstroemer”

Lectotype designated here: “linifolia [m. Linnaeus] A [Alstroemer]”: Herb. Linn. N° 208.4 (LINN-ML)

The lectotype is formed by a single, apparently, annual plant of 10 cm high, erect, with 3 short branches with fruits in the upper half, and other branches emerging near to the axonomorph root; stem with very short internodes, surpassed by the leaves; leaves of 6-10 x 0.1-0.2 mm, narrowly linear, truncated at the base, obtuse; pedicels all with fruits, recurved, rigids; fruits shorter than the calyx. The characters of this plant coincide with those indicated by Linnaeus when he described the species, so there is no doubt that this plant is the same as Linnaeus saw. The indication “Lusitania” could have been made by Linnaeus from a plant from Portugal described by Tournefort and cited as a synonym.

Annual plants such as the type are very frequent on the coast of the province of Cádiz near the capital, so it is not surprising that Linnaeus thought that *A. linifolia* was annual. According to Fernández Pérez (1990), Class Alstroemer toured these territories in the spring-summer of 1760. Probably he collected this plant in late spring when it did not flowers. That is why Linnaeus only described the fruits, as seen in the type material. So the type probably comes from this region of Cádiz.

-Red plants of *Lysimachia monelli* should be called *Lysimachia collina* because plants with this colour were described by Schousboe in 1800 as *Anagallis collina*. Due to *Anagallis* now are included in *Lysimachia*, we do a new combination.

***Lysimachia collina* (Schousb.) FJ. Jiménez-López, com.nov.** $\equiv$ *Anagallis collina* Schousb., Jagttag. Vextrig. Marokko.: 78. (1800), basion. $\equiv$ *Anagallis monelli* subsp. *collina* (Schousb.) H. Lindb. in Acta Soc. Sci. Fenn., ser. B, Opera Biol. 1(2): 115. (1932). $\equiv$ *Anagallis monelli* var. *collina* (Schousb.) Pau in Mem. Real. Soc. Esp. Hist. Nat. 12 (5): 360 (1924) $\equiv$ *Anagallis linifolia* var. *collina* (Schousb.) Ball in J. Linn. Soc. 16: 562 (1878). $\equiv$ *Anagallis linifolia* f. *rubriflora* Batt. in Battandier & Trabut, Fl. Algérie, Dicot.: 723 (1890), [syn. subst.].

Ind. loc.: “Frequens occurrit in collibus aridis provinciae Haha”

Lectotype designated here: Mogador [Haha] Schousboe [m. Schousboe]: C 10001180.

The type material is composed by a single woody stem of 15 cm length, without leaves in the lower half, branched in the upper half; each branch short, with some oval-lanceolate leaves, acute, shorter than the pedicels and the majority of the flowers in anthesis; isoelectotype C 10001181.

Acknowledgements

We thank the following people for their contribution in the sample collection and data analysis: F. Balao, R. Berjano, M. Buide, I. Casimiro-Soriguer, J.A. Mejías, M.A. Ortiz and S. Talavera. We thank to S. Talavera for comments on the manuscript. This work was funded by Spanish with FEDER funds from the Ministerio de Ciencia e Innovación and MINECO (CGL2008-02531-E, CGL2012-33270 & CGL2015-63827 to M. Arista), the fellowship BES-2013-062859 from MINECO to F.J. Jiménez-López and the VPPI-US to M. Talavera. We thank the Herbarium of the University of Sevilla (CITIUS) for their collaboration and the use of its new DNA Bank, and to Royal Botanic Gardens, Kew.

References

- Anderberg AA, Manns U, Källersjö M. 2007. Phylogeny and floral evolution of the *Lysimachieae* (Ericales, *Myrsinaceae*). Evidence from *ndhF* sequence data. *Willdenowia*, **37**: 407-421.
- Arcangeli G. 1894. *Compendio della Flora Italiana* 2: 456. Torino and Roma.
- Arista M, Talavera M, Berjano R, Ortiz PL. 2013. Abiotic factors may explain the geographical distribution of flower colour morphs and the maintenance of colour polymorphism in the scarlet pimpernel. *Journal of Ecology*, **101**: 1613-1622.
- Aymerich P, Sáez LI. 2015. Comentaris i precisions previs a la Checklist de la flora de Catalunya (Nord-Est de la península Ibèrica). *Orsis*, **29**: 23-90.
- Barrelier J. 1714. *Plantae per Galliam, Hispaniam et Italiam*. Ed. Antoine de Jousieu. Parisii.
- Bergmeier E, Blockeel T, Böhling N, Fournaraki C, Gotsiou P, Jahn R, Lansdown R & Turland N. 2011. An inventory of the vascular plants and bryophytes of Gavdopoula island (S Aegean, Greece) and its phytogeographical significance. *Willdenowia*, **41**: 179-190.
- Bollback JP. 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC bioinformatics*, **7**: 88.
- Bolòs O, Vigo J. 1996. *Flora dels Països Catalans*. 3. Ed. Barcino. Barcelona.
- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, MendesFK, Müller NF, Ogilvie HA, Plessis L, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology*, **15**: e1006650.
- Brook GA, Burney DA, Coward J. 1990. Desert paleoenvironmental data from cave speleothems with examples from the Chihuahuan. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **76**:311-330.
- Carlón L, Laínz M, Moreno Moral G, Rodríguez Berdasco JM, Sánchez Pedraja O. 2014. Contribuciones al conocimiento de la flora cantábrica, IX. *Jardín Botánico Atlántico*.

- Carlson JE, Holsinger KE. 2015. Extrapolating from local ecological processes to genus-wide patterns in colour polymorphism in South African *Protea*. *Proceedings of the Royal Society of London B: Biological Sciences*, **282**: 20150583.
- Chase MW, Hills HH. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon*, **40**: 215-220.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**: 772.
- Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, **24**: 332-340.
- Demesure B, Sodji N, Petit RJ. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**: 129-134.
- Dörfler I. 1903. Herbarium Normale XLV: 143.
- Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS biology*, **4**: e88.
- Dyer RA. 1963. Primulaceae. In: Dyer RA, Codd IE, Rycroft HB. eds. *Flora of Southern Africa*. Pretoria. 9-14.
- Ellis TJ, Field DL. 2016. Repeated gains in yellow and anthocyanin pigmentation in flower colour transitions in the Antirrhineae. *Annals of Botany*, **117**: 1133-1140.
- Farris JS, Kellersjo M, Kluge AG, Bult C. 1994. Testing significance of congruence. *Cladistics*, **10**: 315-320.
- Ferguson LF. 1972. *Anagallis* L. In: Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA. eds. *Flora Europea 3, Diapensiaceae-Myoporaceae*. Cambridge University Press. Great Britain. 28-29.
- Fernández Pérez J. 1990. Tres apóstoles de Linné en Cádiz: Pehr Osbeck, Pehr Löfving y Clas Alströmer. In: Pelayo López F. ed. *Pehr Löfving y la expedición al Orinoco*. Sociedad Estatal V Centenario. Madrid. 51-102.

- Freyre R, Griesbach RJ. 2004. Inheritance of flower color in *Anagallis monelli* L. *HortScience*, **39**: 1220-1223.
- Friis EM. 1985. Angiosperm fruits and seeds from the Middle Miocene of Jutland (Denmark). *Det Kongelige Danske Videnskaberne Selskab*, **24**: 1-165.
- García Pérez L, Talavera S, Ortiz PL. & Arista M. 1997. Números cromosómicos de plantas occidentales 734-737. *Anales del Jardín Botánico de Madrid*, **55**: 344-345.
- Gibbs PE, Talavera S. 2001. Breeding system studies in three species of *Anagallis* (Primulaceae): self-incompatibility and reduced female fertility in *A. monelli* L. *Annals of Botany*, **88**: 139-144.
- Hamilton MB. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8**: 521-523.
- Jahandiez E, Maire R. 1934. Catalogue des Plantes du Maroc. Vol 3. Imprimerie Minerva. Alger. 563
- Hao G, Yuan YM, Hu CH, Ge CH, Zhao NX. 2004. Molecular phylogeny of *Lysimachia* (*Myrsinaceae*) based on chloroplast *trnL-F* and nuclear ribosomal ITS sequences. *Molecular Phylogenetics and Evolution*, **31**: 323-339.
- Harborne JB. 1968. Comparative biochemistry of the flavonoids – VII. Correlations between flavonoid pigmentation and systematics in the family Primulaceae. *Phytochemistry*, **7**: 1215-1230.
- Heled J, Drummond A. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**: 570-580.
- Heled J, Drummond AJ. 2012. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Systematic biology*, **61**: 138-149.
- Ho SY, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**: 367–380.
- Hoffmannsegg JCG, Link JHF. 1813-1820. Flore Portugaise, 1 (11): 325, Tab. 64. Impr Charles Frédéric Amelang. Berlin.
- Hopkins R, Rausher MD. 2012. Pollinator-mediated selection on flower color allele drives reinforcement. *Science*, **335**: 1090-1092.
- Hopkins, R. (2013). Reinforcement in plants. *New Phytologist*, 197(4), 1095-1103.

- Jiménez-López FJ, Ortiz MA, Berjano R, Talavera S, Terrab A. 2016. High population genetic substructure in *Hypochaeris leontodontoides* (Asteraceae), an endemic rupicolous species of the Atlas Mountains in NW Africa. *Alpine Botany*, **126**: 73-85.
- Jiménez-López FJ, Ortiz PL, Talavera M, Arista M. 2019a Selfing maintains flower colour polymorphism in *L. arvensis* despite high inbreeding depression. *bioRxiv* doi: 10.1101/761122
- Jiménez-López FJ, Matas L, Arista M, Ortiz PL. 2019b. Flower colour segregation and flower discrimination under the bee vision model in the polymorphic *Lysimachia arvensis*. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, 1-9.
- Jiménez-López FJ, Ortiz PL, Talavera M, Pannell JR, Arista M. 2020. The role of lateral and vertical herkogamy in the divergence of the blue-and red-flowered lineages of *Lysimachia arvensis*. *Annals of Botany*, **125**: 1127-1135
- Källersjö M, Bergqvist G, Anderberg AA. 2000. Generic realignment in primuloid families of the Ericales s.l.: a phylogenetic analysis based on DNA sequences from three chloroplast genes and morphology. *American Journal of Botany*, **87**: 1324-1341.
- Kay KM, Reeves PA, Olmstead RG, Schemske DW. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. *American Journal of Botany*, **92**: 1899-1910.
- Kershaw AP, Bretherton SC, Kaars S. 2007. A complete pollen record of the last 230 ka from Lynch's Crater, north-eastern Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **251**:23-45.
- Kirkpatrick M. 2000. Reinforcement and divergence under assortative mating. *Proceedings of the Royal Society of London B: Biological Sciences*, **267**: 1649-1655.
- Kollmann F, Feinbrun N. 1968. A cyto-taxonomic study in Palestinian *Anagallis arvensis* L. *Notes of the Royal Botanic Garden Edinburgh*, **27**: 173-186.
- Kress A. 1969. Cytotaxonomic studies on Primulaceae. *Phyton (Austria)*, **13**: 211-225.
- Lindberg H. 1932. Beetle, collected in Latvia 1931. *Folia Zoologica et Hydrobiologica*, **4**: 163-166.
- Linnaeus C. 1753. Species Plantarum 1. Stockholm, Sweden. 148-149.

- López González G. 1990. La obra botánica de Löfling en España. In: Pelayo López F. ed. *Pehr Löfling y la expedición al Orinoco*. Sociedad Estatal V Centenario. Madrid. 33-49.
- Magallón S, Hilu KW, Quandt D. 2013. Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *American Journal of Botany*, **100**: 556-573.
- Manns U, Anderberg AA. 2005. Molecular phylogeny of *Anagallis* (Myrsinaceae) based on ITS, *trnL*-F, and *ndhF* sequence data. *International Journal of Plant Sciences*, **166**: 1019-1028.
- Manns U, Anderberg AA. 2007a. Relationships of *Anagallis foemina* and *A. arvensis* (Myrsinaceae): new insights inferred from DNA sequence data. *Molecular Phylogenetics and Evolution*, **45**: 971-980.
- Manns U, Anderberg AA. 2007b. Character evolution in *Anagallis* (Myrsinaceae) inferred from morphological and molecular data. *Systematic Botany*, **32**: 166-179.
- Manns U, Anderberg AA. 2009. New combinations and names in *Lysimachia* (Myrsinaceae) for species of *Anagallis*, *Pelletiera* and *Trientalis*. *Willdenowia*, **39**: 49-54.
- Mariz J. 1899-1900. Subsidy for the Portuguese flora studio. Primulaceae and Gentianaceae. *Boletín de la Sociedad Broteriana*, **16**: 156-195.
- Marsden-Jones EM. 1935. The genetic of *Anagallis arvensis* Linn. and *Anagallis foemina* Mill. *Proceedings of the Linnean Society of London*, **147**: 105-106.
- Marsden-Jones EM, Weiss FE. 1938. The essential differences between *Anagallis arvensis* Linn. and *Anagallis foemina* Mill. *Proceedings of the Linnean Society of London*, **150**: 146-155.
- Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuve P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*, **26**: 2462-2463.
- Martins L, Oberprieler C, Hellwig FH. 2003. A phylogenetic analysis of *Primulaceae* s.l. based on internal transcribed spacer (ITS) DNA sequence data. *Plant Systematics and Evolution*, **237**: 75-85.
- Matsumura S, Yokoyama J, Fukuda T, Maki M. 2009. Origin of the disjunct distribution of flower colour polymorphism within *Limonium wrightii* (Plumbaginaceae) in the Ryukyu Archipelago. *Biological Journal of the Linnean Society*, **97**: 709-717.

- Miller MA, Pfeiffer W, Schwartz T. 2010. "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA. 1 - 8.
- Monein A, Atta II, Shehata AA. 2003. On the delimitation of *Anagallis arvensis* L. (*Primulaceae*) 1. Evidence based on micromorphological characters, Palynological features and Karyological studies. *Pakistan Journal of Biological Sciences*, **6**: 29-35.
- Narbona E, Wang H, Ortiz PL, Arista M, Imbert E. 2018. Flower colour polymorphism in the Mediterranean Basin: Occurrence, maintenance and implications for speciation. *Plant Biology*, **20**: 8-20.
- Nordborg M. 2001. Coalescent theory. In: Balding DJ, Bishop M, Cannings C. eds. *Handbook of Statistical Genetics*. Wiley. 179-212.
- Oh IC, Anderberg AL, Schönenberger J, Anderberg AA. 2008. Comparative seed morphology and character evolution in the genus *Lysimachia* (*Myrsinaceae*) and related taxa. *Plant systematics and Evolution*, **271**: 177-197.
- Ohsako T, Ohnishi O. 2000. Intra-and interspecific phylogeny of wild *Fagopyrum* (*Polygonaceae*) species based on nucleotide sequences of noncoding regions in chloroplast DNA. *American Journal of Botany*, **87**: 573-582.
- Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual review of genetics*, **3**: 401-437.
- Padilla-García N, Rojas-Andrés BM, López-González N, Castro M, Castro S, Loureiro J, Albache DC, Machon N, Martínez-Ortega MM. 2018. The challenge of species delimitation in the diploid-polyploid complex *Veronica* subsection Pentasepalae. *Molecular Phylogenetics and Evolution*, **119**: 196-209.
- Pastor J. 1992. *Atlas Cromosómico de la Flora Vascular de Andalucía Occidental*. Universidad de Sevilla. 283-284.
- Pau C. 1915. About *Anagallis monelli* L. *Boletín de la Sociedad Aragonesa de Ciencias Naturales*, **14**: 100-107.
- Peruzzi L. 2010. Notulae alla checklist della Flora Vascolare Italiana 9 (1623-1681). *Informatore Botanico Italiano*, **42**: 383-384.

- Petit C, Bretagnolle F, Felber F. 1999. Evolutionary consequences of diploid–polyploid hybrid zones in wild species. *Trends in Ecology & Evolution*, **14**: 306-311.
- Pujadas A. 1997. *Anagallis* L. In: Castroviejo S, Aedo C, Laínz M, Morales R, Muñoz Garmendia F, Nieto Feliner G, Paiva J. eds. *Flora iberica*, V. Ebenaceae-Saxifragaceae. Real Jardín Botánico, CSIC. Madrid. 57-62.
- Rambaut A. 2012. FigTree v. 1.4.4. <http://tree.bio.ed.ac.uk/software/figtree/>
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*. syy032.
- Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences*, **169**: 7-21.
- Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in ecology and evolution*, **3**: 217-223.
- Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**: 1211-1216.
- Sampaio G. 1900. Studies on the Flora of Porto's Surroundings. I. Primulaceae (Vent.) *Annaes de Sciences Naturaes (Porto)* **6**: 51-62.
- Sang TD, Crawford J, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (*Paeoniaceae*). *American Journal of Botany*, **84**: 1120-1136.
- Schousboe P. 1800. *Observations on the Growth War in Morocco*. Copenhagen: Printed for Fofatteren has KH. Seidelin. Copenhagen.
- Semerikova SA. & Semerikov VL. 2016. Phylogeny of firs (genus *Abies*, Pinaceae) based on multilocus nuclear markers (AFLP). *Russian Journal of Genetics*, **52**: 1164-1175.
- Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**: 275-288.

- Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Judd WS. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species?. *Taxon*, **5**: 13-30.
- Soltis DE, Visger CJ, Soltis PS. 2014. The polyploidy revolution then... and now: Stebbins revisited. *American journal of botany*, **10**: 1057-1078.
- Soltis PS, Marchant DB, Van de Peer Y, Soltis DE. 2015. Polyploidy and genome evolution in plants. *Current Opinion in Genetics & Development*, **35**: 119-125.
- Souissi A, Gagnaire PA, Bonhomme F, Bahri-Sfar L. 2017. Introgressive hybridization and morphological transgression in the contact zone between two Mediterranean Solea species. *Ecology and Evolution*, **7**: 1394-1402.
- Stamatakis A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics*, **30**: 1312-1313.
- Stebbins GL. 1971. Adaptive radiation of reproductive characteristics in angiosperms, II: seeds and seedlings. *Annual Review of Ecology and Systematics*, **2**: 237-260.
- Strauss SY. & Whittall JB. 2006. *Non-pollinator agents of selection on floral traits*. In: Harder LD, Barrett SCH eds. *Ecology and Evolution of Flowers*. Oxford University Press, Oxford, UK. 120-138.
- Šveřepová G. 1972. On the cytotaxonomy of the genus Anagallis L. (K cytotaxonomi rodu Anagallis L.). *Preslia*, Praha, **44**: 219-226.
- Swofford DL. 2003. *PAUP**. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Symonds VV, Lloyd AM. 2003. An analysis of microsatellite loci in Arabidopsis thaliana: mutational dynamics and application. *Genetics*, **165**: 1475-1488.
- Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**: 1105-1109.
- Takahashi Y, Takakura KI, Kawata M. 2016. Spatial Distribution of Flower Color Induced by Interspecific Sexual Interaction. *PloSone*, **11**: e0164381.

- Talavera S, Gibbs PE, Fernández-Piedra MP, Ortiz-Herrera MA. 2001. Genetic control of self-incompatibility in *Anagallis monelli* (Primulaceae: Myrsinaceae). *Heredity*, **87**: 589-597.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA 6: Molecular Evolutionary Genetics Analysis v. 6.0. *Molecular Biology and Evolution*, **30**: 2725-2729.
- Thompson JD, Gibson T, Higgins DG. 2002. Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, **2** Unit 2.3.
- Tournefort JP. 1719. Institutions of botany, ed.3, 1. Paris.
- Valdés B. 1970. Números cromosómicos de algunas plantas españolas. *Boletín de la Real Sociedad Española de Historia Natural (Biología)*, **68**:193-197.
- Viruel J, Forest F, Paun O, Chase MW, Devey D, Couto RS, Segarra-Moragues JG, Catalán P, Wilkin P. 2018. A nuclear Xdh phylogenetic analysis of yams (Dioscorea: Dioscoreaceae) congruent with plastid trees reveals a new Neotropical lineage. *Botanical Journal of the Linnean Society*, **187**: 232-246.
- Viruel J, Segarra-Moragues JG, Raz L, Forest F, Wilkin P, Sanmartín I, Catalán P. 2016. Late Cretaceous–Early Eocene origin of yams (*Dioscorea*, *Dioscoreaceae*) in the Laurasian Palaearctic and their subsequent Oligocene–Miocene diversification. *Journal of Biogeography*, **43**: 750-762. .
- Wakeley J. 2009. *Coalescent theory: an introduction*. Roberts and Company Publishers, Greenwood Village, Colorado.
- Wang H, Conchou L, Bessière JM, Cazals G, Schatz B, Imbert E. 2013. Flower color polymorphism in *Iris lutescens* (Iridaceae): biochemical analyses in light of plant–insect interactions. *Phytochemistry*, **94**: 123-134.
- Wang J, Tian L, Lee HS, Chen ZJ. 2006. Non-additive regulation of FRI and FLC loci mediates flowering-time variation in Arabidopsis allopolyploids. *Genetics*, **173**: 965-974.
- Wendel JF, Doyle JJ. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis DE, Soltis PS, Doyle JJ. eds. *Molecular Systematics of Plants II*. Springer, Boston, MA. 265-296.

- Wessinger CA, Rausher MD. 2012. Lessons from flower colour evolution on targets of selection. *Journal of Experimental Botany*, **63**: 5741-5749.
- White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a Guide to Methods and Applications*, **18**: 315-322.
- Willkomm HM. 1870. Primulaceae. In: Willkomm HM, Lange JM, eds. *Prodromus Florae Hispanicae*, 2. Stuttgart. 635-650.
- Wilson P, Wolfe AD, Armbruster WS, Thomson JD. 2007. Constrained lability in floral evolution: counting convergent origins of hummingbird pollination in *Penstemon* and *Keckiella*. *New Phytologist*, **176**: 883-890.
- Wolf PG, Murray RA, Sipes SD. 1997. Species-independent, geographical structure of chloroplast DNA haplotypes in a montane herb *Ipomopsis* (Polemoniaceae). *Molecular Ecology*, **6**: 283-291.
- Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences*, **84**: 9054-9058.
- Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, **106**: 13875-13879.
- Yan HF, Zhang CY, Anderberg AA, Hao G, Ge XJ, Wiens JJ. 2018. What explains high plant richness in East Asia? Time and diversification in the tribe *Lysimachieae* (Primulaceae). *New Phytologist*, **219**: 436-448.
- Yesson C, Toomey NH, Culham A. 2009. *Cyclamen*: time, sea and speciation biogeography using a temporally calibrated phylogeny. *Journal of Biogeography*, **36**: 1234-1252.
- Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlinn DJ, O'Meara BC, Moles AT, Reich PB, Royer DL, Soltis DE, Stevens PF, Westoby M, Wright IJ, Aarssen L, Bertin RI, Calaminus A, Govaerts R, Hemmings F, Leishman MR, Oleksyn J, Soltis PS, Swenson NG, Warman L, Beaulieu JM. 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature*, **506**: 89-92.

Zhang CY, Wang FY, Yan HF, Hao G, Hu CM, Ge XJ. 2012. Testing DNA barcoding in closely related groups of *Lysimachia* L. (*Myrsinaceae*). *Molecular Ecology Resources*, **12**: 98-108.

Figure legends

Figure 1. – Geographical distribution of the studied populations of *Lysimachia arvensis* (circles) and *Lysimachia monelli* (triangles). Colours correspond to flower colour of the studied populations: blue and red for pure blue and red populations, respectively, and pink for mixed populations.

Figure 2. – Ancestral state reconstruction of flower colour and ploidy level of *L. arvensis* and *L. monelli* and related taxa (see Table S1). RAxML tree based on ITS sequences. The samples of *L. arvensis* from mixed populations are highlighted in bold. Bootstrap values > 50% are given above the branches and Posterior probabilities of Bayesian approach are given below branches: a (0.995-1.000 PPS), b (0.950-0.994 PPS) and c (0.850-0.949). C1: Clade1; C2: Clade 2; SCA: Subclade A and SCB: Subclade B. * (83 BS, 0.984 PPS).

Figure 3. – (1) Phylogenetic network and (2) Maximum likelihood (RAxML) tree based on concatenated plastid sequence data (*trnH-psbA*, *rps16-trnK* and *rpl32-trnL*) of *L. arvensis*, *L. monelli* and related taxa (see Table S1). Main clades in trees from ML (RAxML) and Bayesian analysis (MrBayes) are identical. Bootstrap values > 50% and posterior probabilities of Bayesian tree are given above and below the

branches respectively. Samples of *L. arvensis* from mixed populations are highlighted in bold.

Figure 4. – Phylogenetic reconstruction for Lysimachiae tribe. Maximum Clade Credibility (MCC) tree reconstructed with BEAST based on Internal transcribed spacer (ITS) data of *Lysimachia* species and related genera of tribe Lysimachieae. The root was calibrated following Yan et al. 2018 and a *Maesa* sp. fossil (Kershaw et al. 2007), and a Yules speciation tree prior. Bootstrap support > 65% of ML estimate (RAxML) and Posterior probabilities of Bayesian approach (BEAST; a [0.995-1.000 PPS], b [0.950-0.994 PPS] and c [0.850-0.950]) are given above branches.

SCB

C1

0.04

LA3_R
LA20_R
LA22_R
LA24_R
LA27_R
LA17_R
LA10_R
LA26_R
LA16_R
LA28_R
LA1_R
LA23_R
LM10_R
LM11_R
LM9_R
LM12_R

LF3
LF1
LF2
LM6_B
LM2_B
LM1_B
LM4_B
LM5_B
LM3_B
LM7_B
LM8_B
LA7_B
LA13_B
LA4_B
LA6_B
LA11_B
LA12_B
LA5_B
LA8_B
LA15_B
LA9_B
LA20_B
LA26_B
LA21_B
LA25_B
LA27_B
LA2_B
LA18_B
LA28_B

LT3
LT2
LT4
LT1
LZ1
LLS1
LLS1a
L_tyrrenhia
L_tenella

L. arvensis
(Red) [4n]

L. monelli
(Red) [2n]

L. foemina [4n]

L. monelli
(Blue) [2n]

L. arvensis
(Blue) [4n]

L. talaverae [2n]

Outgroup