RESEARCH ARTICLE

The Polyploid Series of the *Achillea millefolium* Aggregate in the Iberian Peninsula Investigated Using Microsatellites

Sara López-Vinyallonga[1], Ignasi Soriano[2], Alfonso Susanna[1]*, Josep Maria Montserra[3], Cristina Roquet[4], Núria Garcia-Jacas[1]

1 Botanic Institute of Barcelona (IBB-CSIC-ICUB), Pg. del Migdia, Barcelona, Spain, 2 Department of Plant Biology, University of Barcelona, Barcelona, Spain, 3 Barcelona Botanical Garden (Consortium of the Museum of Natural History of Barcelona), Barcelona, Spain, 4 Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Grenoble Alpes, Grenoble, France

* asusanna@ibb.csic.es

Abstract

The *Achillea millefolium* aggregate is one of the most diverse polyploid complexes of the Northern hemisphere and has its western Eurasian boundary in the Iberian Peninsula. Four ploidy levels have been detected in *A. millefolium*, three of which have already been found in Iberia (diploid, hexaploid and octoploid), and a fourth (tetraploid) reported during the preparation of this paper. We collected a sample from 26 Iberian populations comprising all ploidy levels, and we used microsatellite markers analyzed as dominant in view of the high ploidy levels. Our goals were to quantify the genetic diversity of *A. millefolium* in the Iberian Peninsula, to elucidate its genetic structure, to investigate the differences in ploidy levels, and to analyse the dispersal of the species. The lack of spatial genetic structure recovered is linked to both high levels of gene flow between populations and to the fact that most genetic variability occurs within populations. This in turn suggests the existence of a huge panmictic yarrow population in the Iberian Peninsula. This is consistent with the assumption that recent colonization and rapid expansion occurred throughout this area. Likewise, the low levels of genetic variability recovered suggest that bottlenecks and/or founder events may have been involved in this process, and clonal reproduction may have played an important role in maintaining this genetic impoverishment. Indeed, the ecological and phenologic uniformity present in the *A. millefolium* agg. in Iberia compared to Eurasia and North America may be responsible for the low number of representatives of this complex of species present in the Iberian Peninsula. The low levels of genetic differentiation between ploidy levels recovered in our work suggest the absence of barriers between them.

Introduction

Achillea millefolium agg. is one of the most diverse polyploid complexes of the Northern hemisphere in terms of morphological, genetic and ecological features [1–4]. This group includes

Achillea millefolium, together with a set of Eurasian and North American related lineages (ca. 24 species; [2]), some of them naturalized in temperate and cold areas on other continents. In Eurasia, the group has its southwestern boundary in the Iberian Peninsula, where it is widely distributed throughout the extra-Mediterranean areas of the northern half, reaching some southern mountain ranges. Yarrows are plants of economic importance, widely used in agriculture and horticulture, as well as in pharmacology and folk medicine due to their numerous therapeutic properties. For example, they are considered to contain antihypertensive, anti-inflammatory, antimalarial and antimicrobial compounds, and have a long list of other uses beside [5–11].

According to [4], the complexity of the aforementioned group is the result of multiple processes of hybridization, polyploidization and evolution linked with different types of habitats. Furthermore, the naturalization of alien strains, introduced either intentionally or accidentally, contributed to this. The existence of different auto- and allopolyploids representing four ploidy levels ($2x$, $4x$, $6x$ and $8x$) is widely documented [4], and analysis of genetic diversity carried out using AFLP markers revealed substantial polymorphism, significantly higher in polyploid than in diploid strains [3]. All our knowledge of the polyploidy and genetics of this species is based on taxa and populations originating from central Eurasia, where the main diversity of the group is concentrated, and, to date, genetic studies on the Iberian populations have not been undertaken. Based on their own data from the northern slopes of the Pyrenees and on few previously published chromosome counts from Portugal [13] and France [14], [3,12] suggest the existence of three entities on the Iberian Peninsula corresponding to at least three ploidy levels whose distribution is poorly known: $2x$ (often considered a different species, *A. ceretanica*, endemic to the eastern Pyrenees), $6x$ (*A. millefolium s. str.*) and $8x$ (sometimes considered another different species, *A. monticola*). The hexaploid and the octoploid cytotypes are difficult to distinguish based on morphological criteria (I. Soriano in *Flora Iberica*, unpublished report), and genetic data show no clear differences [1], [12]. The presence of the three karyologic entities in Iberia has been confirmed by indirect methods (flow cytometry and stomatal measurements of herbarium material) together with chromosome counts (I. Soriano, University of Barcelona, unpublished data; see a breakthrough in [15]). The distribution of the three cytotypes is still unclear, although the $8x$ seem to be much more common than the $6x$ populations. The tetraploid cytotype had not been reported from the Iberian Peninsula, but we have recorded its presence there (albeit very scarce) during the research carried out for this paper (A. Susanna et al., Botanic Institute of Barcelona, unpublished data). Hexaploid strains are considered weeds and display a sub-cosmopolitan distribution [3,16]. In the Iberian Peninsula, *A. millefolium* agg. does not display the great ecological diversity present in Eurasia and North America. Instead, their populations tend to occupy ruderal and nitrified environments and disturbed pastures.

The lack of genetic studies on Iberian yarrows motivated us to investigate this species aggregate at population level using SSR markers (microsatellites). Microsatellites are widely used to investigate genetic variation, structure and dynamics both at population and species level (e. g., [17,18]). The main advantages of microsatellites are: i) high polymorphism derived from a high mutation rate; ii) co-dominant Mendelian inheritance; and iii) multi-allelic nature [19]. In polyploid organisms, the number of alleles for each locus is not known in partial heterozygotes. In these cases, the number of microsatellite DNA alleles displayed is less than the possible maximum number for the ploidy level of a given individual at a given locus [20,21]. Therefore, it is not possible to calculate some of the usual parameters of genetic diversity [22–25]. For this reason, despite polyploidy being common in plants [26–30], few studies on population genetics using microsatellites focus on these organisms [20]. However, different

approaches have been proposed to overcome these problems [20–23,31–34] and microsatellites are an adequate tool for this purpose.

The aims of the present work are: (i) to quantify the genetic diversity of *A. millefolium* in the Iberian Peninsula; (ii) to elucidate its genetic structure and gene flow; (iii) to investigate the putative genetic differentiation between ploidy levels; and (iv) to evaluate alternative dispersal methods for this species within the reported area.

Materials and Methods

Plant material

Our sampling covered the whole area of *A. millefolium* in the Iberian Peninsula (Fig 1), with only one exception, namely, that despite intense searching, we were not able to include the populations reported from southern Iberia. Some of these citations are perhaps best considered cases of naturalization, involving individuals that have escaped from cultivation [Gabriel Blanca, University of Granada (Spain), pers. comm.]. A total of 26 populations, each containing 12–16 individuals, were sampled for the study (Table 1). Species of the *Achillea millefolium* aggregate are not endangered or protected in Spain and the EU, and our sampling was not carried out in any protected area. Thereafter, no specific permissions were required for our collections.

Chromosome counts

Chromosome preparations were obtained using root meristems gathered from germinating seeds or from wild-collected plants cultivated at the Botanic Garden of Barcelona. Root tips were first pretreated with 0.002 M 8-hydroxyquinoline at 4°C for 8 h or in 0.02% colchicine at room temperature for 3 h, then the material was fixed with Carnoy at low temperatures for 24 h. Root tips were then hydrolysed with 5N HCl at room temperature for 45 m, after which they were stained in 1% acetic orcein or in Schiff's reagent at room temperature for more than 2 hours. The root tips were then mounted in 45% acetic acid, macerated and then squashed by hand under a coverslip. Somatic metaphases were then examined from a minimum of five metaphase preparations from 3–6 different individuals using an Olympus microscope U-TV1-X and a C3030 camera. Suitable preparations were then made permanent by freezing in solid CO_2 for removing the cover slip, dehydrating in ethanol and mounting in Canada balsam.

Stomata measurements

Ploidy level was also indirectly estimated by measurement of stomata size. For the five populations for which chromosome counts were unavailable by lack of adequate material (living plants or achenes), measurements were carried out on five individuals from desiccated leaves of the same individuals used for microsatellite analysis. For the rest of populations, one to three individuals were used. Foliar segments were excised from the middle part of the leaf, softened in NaOH 0.5 M for 24 h, and bleached in a saturated solution of chloral hydrate for 6–12 h. Epidermis fragments were excised from the adaxial side of the distal half of the segments, semipermanently mounted in glycerol sealed with nail varnish, and measured with a micrometric ocular mounted in an Olympus CH-2 optic microscope at 400x. Two measures, long axis (stoma length or SL) and short axis (stoma width or SW), were made on an ideal ellipse adapted to the stoma. Measurement of SL equals the length of the guard cell. For each mounted slide, measurements were made on different zones of the epidermis of the segment, discarding stomata on the margins of the leaf. A total of 30–63 stomata were measured per population.
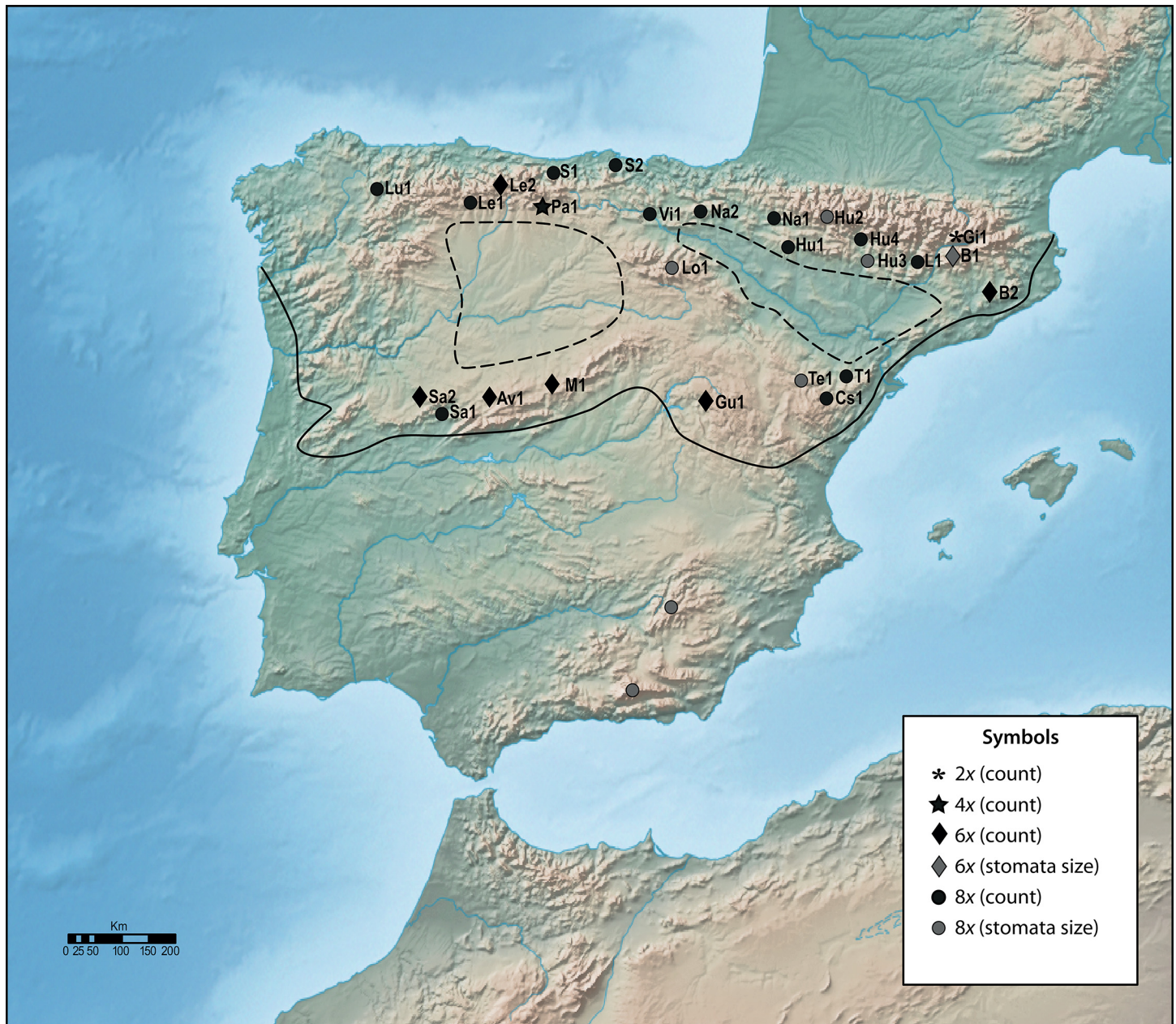
Fig 1. Locations of sampling with indication of ploidy level. A solid line indicates the southern limit of the area of *A. millefolium*. Within this area, dotted lines delimit the areas where the species is absent. The two points in South Iberia represent the most probable localities where the species currently grows (see "Plant material" for details). Base map from Natural Earth (http://www.naturalearthdata.com/).

doi:10.1371/journal.pone.0129861.g001

For the five populations for which we lacked chromosome counts, 100–120 stomata were measured.

DNA isolation and microsatellite loci

Total genomic DNA was extracted from desiccated leaves using the CTAB method of [35], as modified by [36] and [37], and stored at -21°C prior to genotyping. Sixteen SSR markers previously developed for *A. millefolium* from Iran [38] were tested for amplification. Seven of these

Table 1. Populations of *Achillea millefolium* investigated in the present study. K: chromosome count; S: stomatal size. Altitudes are expressed in meters above sea level.

Code	Location	Coordinates and Altitude	Habitat	Ploidy	Method
Av1	ÁVILA: between Poveda and Pradosegar	40° 33.6480'N 5° 5.5512'W; 1,171 m	Meadows and dry siliceous grasslands	6x	K + S
B1	BARCELONA: Pobla de Lillet, heliport	42° 14.6460'N 1° 57.8028'E, 887 m	Meso-xerophilous calcareous pastures and edges of paths.	6x	S
B2	BARCELONA: Montseny massif	41° 48.3552'N 2° 21.2484'E; 1,326 m	Mesophilous siliceous pastures and thickets.	6x	K + S
Cs1	CASTELLÓ: road CV15 Vilafranca-coll d'Ares	40° 27.2088'N 0° 8.6106W; 1,145 m	Damp patches on calcareous fringes	8x	K + S
Gi1	GIRONA: Alp, La Molina	42° 20.0496'N 1° 55.9392'E; 1,647 m	Mesophilous siliceous pastures and ski slopes.	2x	K + S
Gu1	GUADALAJARA: Orea	40° 33.,8544N 1° 44.5692W; 1,466 m	Field margins and slopes, near a stream.	6x	K + S
Hu1	HUESCA: Loarre Castle	42° 19.3258'N 0° 37.0019'W; 1,040 m	Waste places and roadsides on calcareous soil	8x	K + S
Hu2	HUESCA: Valle de Tena, Barranco del Petruso	42° 47.7942'N 0° 24.0918'W, 1,700 m	Waste places and roadsides on calcareous soil	8x	S
Hu3	HUESCA: Escales reservoir	42° 20.3377'N 0° 43.6887E; 816 m	Meso-xerophilous calcareous pastures.	6x	S
Hu4	HUESCA: Plan	42° 34.8500'N 0° 20.5110 E; 1,085 m	Oak fringes and edges of paths.	8x	K
L1	LLEIDA: Vall de la Vansa, Clot de les Fonts	42° 14.3322'N 1° 26.0190E; 870 m	Waste places on calcareous soil	8x	K
Le1	LEÓN: Pola de Gordón	42° 51.0504'N 5° 39.8982'W; 1,107 m	Clearings, slopes and meadows, near a river.	8x	K
Le2	LEÓN: Mampodre massif	43° 2.2848' 5° 12,0774'W; 1,810 m	Mesophilous calcareous pastures.	6x	K
Lo1	LA RIOJA: Viniegra de Abajo, Collado Ocejo	42° 12.5798'W 2° 54.0270'W; 1,788 m	Mesophilous siliceous pastures.	8x	S
Lu1	LUGO: Bullán, Quinta de Cancelada, Becerreá	42° 58.0902'N 7° 5.0967W; 725 m	Mesophilous siliceous pastures	8x	K
M1	MADRID: Sierra de Guadarrama, Fuente de los Geólogos	40° 46.5521'N 4° 0.3374'W; 1,743 m	Scots pine fringes and slopes, near a pic-nic area.	6x	K + S
Na1	NAVARRA: Bigüezal	42° 41.0916N 1° 8.3124'W; 880 m	Meadows on calcareous soil.	8x	K + S
Na2	NAVARRA: Sierra de Urbasa, south from Alsasua	42° 52.3212N 2° 10.9782W; 683 m	Oak and beechwood fringes and edges of paths.	8x	K + S
Pa1	PALENCIA: Cervera de Pisuerga	42° 52.0194N 4° 30.1638W; 1029 m	Clearings with dry meadows, near the road	4x	K
S1	CANTABRIA: road CA-182 to Cabuérniga	43° 15.1908'N 4° 23.0844W; 155 m	Meadows, slopes and roadsides.	8x	K + S
S2	CANTABRIA: Laredo	43° 24.8262'N 3° 24.2904'W; 26 m	Meadows and edges of paths, near the coastline.	8x	K + S
Sa1	SALAMANCA: between Candelario and La Garganta	40° 20.3934'N 5° 46.0698'W; 1,220 m	Oakwood fringes and edges of paths.	8x	K
Sa2	SALAMANCA: La Alberca	40° 29.3124'N 6° 6.9036'W; 1,085 m	Scots pine fringes and edges of paths.	6x	K + S
T1	TARRAGONA: Port de Tortosa, Cova de les Avellanes	40° 47.3028'N 0° 18.3984'E; 990 m	Oakwood fringes and edges of paths.	8x	K
Te1	TERUEL: road N-420 between Valdeconejos and Utrillas	40° 46.616'N 0° 49.863W; 1,258 m	Waste places on calcareous soil.	8x	S
Vi1	ÁLAVA: between Paul and Salinas de Añana	42° 47.8578N 2° 58.1724W; 676 m	Slopes and roadsides.	8x	K + S

were polymorphic and successfully amplified for yarrow populations from Iberia. All SSR loci were amplified using FAM, NED, PET and VIC fluorescently labeled forward primers as explained by [39]. Different profiles were used for the amplification following the conditions established for each locus in the original publication. Given the importance of clonal propagation in *A. millefolium*, plants were collected at least 1 meter apart to avoid duplication. For this reason, we obtained data for 12–16 random individuals per site, accounting for a total of 375 individuals belonging to 26 populations (Fig 1) and representing the entire distribution area of the species and the four cytotypes known to occur within the Iberian Peninsula. Vouchers were deposited at the herbarium of the Botanic Institute of Barcelona (BC). Genotyping was performed by means of an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA) and LIZ600 size standard at the Interdisciplinary Center for Biotechnology Research (ICBR) facility at the University of Florida. Fragment analysis was performed using GENE-MARKER 1.5 software (SoftGenetics, LLC, State College, PA) and data was scored manually.

Statistical analysis

As mentioned in the introduction, polyploidy frustrates the statistical treatment of microsatellite data and prevents the performance of many analyses and calculations, especially when individuals with different ploidy levels co-exist in the same matrix, as in the present work [22–24,31–34]. We intended to use peak heights to determine the multilocus genotypes of each individual [40], but we were not able to do so for individuals other than for homozygous and fully heterozygous individuals. For this reason, and since there were four different ploidy levels ($2x$, $4x$, $6x$ and $8x$), we followed the approach of [32] and recorded the banding patterns observed at each locus as a binary presence/absence matrix: alleles were coded as a separate "locus" into 0 (allele absent) vs. 1 (allele present); the so called 'allele phenotypes'. Therefore, despite using co-dominant markers for amplification and genotyping, data were scored as dominant loci in order to render them suitable for conventional population genetics analysis software. Although this method may imply some bias, we considered it the best way to treat individuals having different ploidy levels in the same matrix when dosage effect was not evident. However, measures of genetic diversity, such as total number of alleles and number of private alleles, would not be biased [32].

All the genetic diversity parameters were computed using GenAlEx 6.5 [41,42]. At population level, the total number of alleles (k), the mean number of alleles (Na), the number of effective alleles (Ne), the number of private alleles (PA) and the unbiased diversity (uh) were calculated. These parameters were also computed at ploidy level. The genetic relationship between pairwise populations and ploidy levels was estimated by means of Nei's unbiased genetic distance (D) and Φ_{PT}, a measure of population genetic differentiation analogous to F_{ST} (GenAlEx documentation). The significance level of Φ_{PT} was computed through 99 permutations.

Four different approaches were used in order to gain insights into the patterns of spatial genetic structure of *A. millefolium* in Iberia. Firstly, the Bayesian method implemented in the software STRUCTURE 2.2 [43] was used to infer the number of genetic clusters present in the data set. The iterations were conducted using the admixture model with correlated allele frequencies, since we assumed that a certain amount of gene flow occurred between populations. The analyses were executed with both prior information on population structure and with no a-priori structure to evaluate the consistency of the groups obtained. We carried out an exploratory run using the number of clusters (K) ranging from 1 to 28 (the total number of populations of the study plus two), using only 10^4 Markov MonteCarlo (MCMC) replications. This analysis did not yield any structure in the data. We carried out a second analysis to see whether we could improve the results, restricting K to a range varying from K = 1 to K = 15. Each K was

estimated 10 times, the length of burn-in period comprising 10^5 iterations and 10^6 MCMC replications. The optimal number of clusters was determined using the ΔK statistical approach of [44]. STRUCTURE assumes Hardy-Weinberg equilibrium within populations and linkage equilibrium between loci within populations [43] and, given the polyploid nature of our data, these tests could not be computed. In order to confirm the results yielded by STRUCTURE, we also applied a non-model-based approach: nonhierarchical K-means clustering [45], as implemented in [46]. This approach, free from a priori population genetic assumptions, assigns individuals to a defined number of genetic groups (K) in order to maximize the intergroup variance (measured here as the inertia [47]). The analysis was performed using R (package "stats" [48]), based on the script of [46]. We performed 100000 independent runs for each assumed value of K (ranging from 1 to 15). We computed I(K) and sd(K), the mean and standard deviation of inter-group inertia for each K value and then I1(K) and I2(K), the first and second order derivatives of I(K). We computed ΔK as L2(K) divided by sd(K) and used it as a criterion for selecting the most likely number of groups in our clustering analyses.

The remaining three methods were performed using GenAlEx 6.5. software. An analysis of molecular variance (AMOVA) was computed with respect to three different factors: 1) between and within populations; 2) between and within ploidy levels; and 3) between and within soil type (calcareous vs. siliceous). In addition, a Principal Co-ordinate Analysis (PCoA) was performed based on the pair-wise genetic distances of individuals and populations. Finally, isolation-by-distance between populations was investigated by computing the correlation between the matrix of pair-wise population genetic distance (Φ_{PT}) and the matrix of geographical distances, by applying the Mantel test (1000 permutations).

Results

Chromosome numbers are summarized in Table 1, and partly illustrated in Fig 2. Stomata sizes and statistics of the measurements are shown in supplementary information S1, S2 and S3 Tables, and S1 Fig. The seven SSR primer pairs used yielded 107 alleles. The overall unbiased diversity (*uh*) for *A. millefolium* was 0.077. At population level (Table 2), the total number of alleles (*k*) ranged from 16 (Hu1) to 35 (B1, Le1, Lu1 and Na2), the mean number of alleles (*Na*) ranged from 0.224 (Hu1) to 0.645 (B1) and the number of effective alleles (*Ne*) ranged from 1.032 (Hu1) to 1.169 (B1). Forty-five exclusive alleles were detected ranging from one (B2, Cs1, Hu2, Le1, Pa1, S2 and Te1) to six (HU3), all of them occurring at low frequencies (mean frequency 0.097). Only five of the populations surveyed lacked private alleles (Hu1, Le2, Lo1, Na1 and Sa1). The lowest value for unbiased diversity was found in population Hu1 (*uh* = 0.022), whereas the highest value was found in population B1 (*uh* = 0.112). Nei's genetic distance (*D*) between pairs of populations ranged from 0.002 between L1 and Sa1 to 0.049 between Gi1 and Lo1 (see Table 3 for details). Most values of pair-wise population genetic differentiation Φ_{PT} were significant (P < 0.05) and ranged from 0.044 between B1 and S2 to 0.334 between Lu1 and S1 (see Table 3 for details).

Regarding ploidy level (Table 2), the total number of alleles (*k*) ranged from 22 (2*x*) to 77 (8*x*), the mean number of alleles (*Na*) ranged from 0.383 (2*x*) to 0.567 (6*x*) and the number of effective alleles (*Ne*) ranged from 1.106 (8*x*) to 1.141 (4*x*). Fifty-five exclusive alleles were detected ranging from one (4*x*) to 28 (8*x*). The lowest value for unbiased diversity belonged to 2*x* and 8*x* (*uh* = 0.071) cytotypes, whereas the highest value belonged to 4*x* and 6*x* (*uh* = 0.088) cytotypes. Nei's genetic distance (*D*) between ploidy levels ranged from 0.002 between 6*x* and 8*x* to 0.030 between 2*x* and 4*x* (see Table 4 for details). All values of genetic differentiation $\Phi_{PT \, between}$ ploidy levels were significant (P < 0.05) and ranged from 0.005 between 6*x* and 8*x* to 0.105 between 2*x* and 4*x* cytotypes (see Table 4 for details).
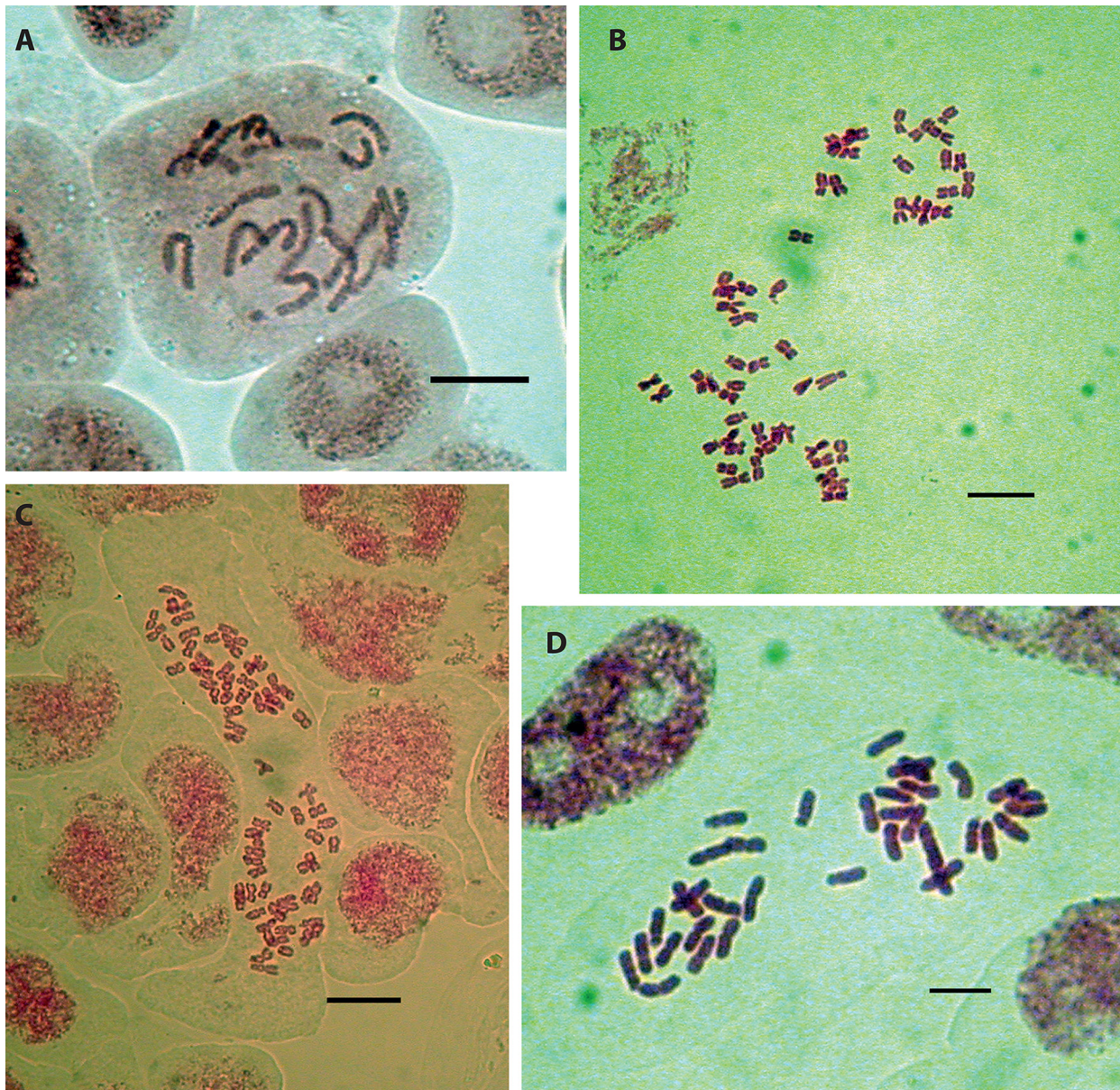
Fig 2. Selected metaphases of the four ploidy levels. A) Diploid plate with $2n = 18$ (pop. Gi1); B) Hexaploid plate with $2n = 54$ (pop. Sa2); C) Octoploid plate with $2n = 72$ (pop. Cs1); D) Tetraploid plate with $2n = 36$ (pop. Pa1). Scale bar = 10 µm

The results of the second STRUCTURE simulation were consistent between runs and, according to [44], K = 3 was the most likely number of clusters for our data (Fig 3). Despite this, no population clusters were generated in any of the runs performed; instead each population was subdivided into three groups (Fig 4). The same pattern was recovered for all the Ks tested. In contrast, the nonhierarchical K-means clustering analysis yielded K = 7 as the most likely according to [44] (Fig 5), but again the resulting groups yielded no population clusters (Fig 6). This absence of population structuring was also found in the other Ks assessed. When AMOVA was performed at population level, the majority of variations (87%) was significantly partitioned within populations, while 13% of the variation was detected between populations ($\Phi_{PT} = 0.126$; $P < 0.005$). Similarly, when the partition was based on the ploidy level, most of

Table 2. Main parameters of genetic diversity for the 26 populations surveyed and the four ploidy levels of Spanish *A. millefolium*. N, number of individuals; *k*, total number of alleles; *Na*, mean number of alleles; *Ne*, number of effective alleles; *PA*, number of private alleles; *uh*, unbiased diversity. For abbreviations of populations, see Table 1.

Population	N	k	Na	Ne	PA	uh
Av1	15	34	0.636	1.158	2	0.105
B1	12	35	0.645	1.169	2	0.112
B2	15	24	0.402	1.108	1	0.066
Cs1	15	27	0.458	1.124	1	0.077
Gi1	11	22	0.383	1.109	2	0.071
Gu1	15	32	0.589	1.119	3	0.084
Hu1	14	16	0.224	1.032	0	0.022
Hu2	15	30	0.533	1.108	1	0.074
Hu3	15	34	0.607	1.095	6	0.070
Hu4	15	32	0.589	1.144	2	0.094
L1	12	24	0.421	1.100	2	0.068
Le1	14	35	0.598	1.095	1	0.070
Le2	15	29	0.505	1.125	0	0.082
Lo1	15	20	0.327	1.081	0	0.051
Lu1	15	35	0.617	1.123	2	0.083
M1	15	30	0.542	1.127	3	0.084
Na1	16	22	0.383	1.082	0	0.054
Na2	15	35	0.626	1.141	2	0.094
Pa1	14	29	0.505	1.141	1	0.088
S1	15	31	0.533	1.098	2	0.071
S2	15	30	0.523	1.125	1	0.080
Sa1	13	28	0.505	1.138	0	0.091
Sa2	15	33	0.607	1.162	4	0.104
T1	15	26	0.458	1.117	3	0.075
Te1	14	24	0.411	1.102	1	0.064
Vi1	15	33	0.570	1.094	3	0.067
Total	**375**	**107**	**0.508**	**1.116**	**45**	**0.077**
2x	11	22	0.383	1.109	2	0.071
4x	14	29	0.505	1.141	1	0.088
6x	102	76	0.567	1.133	25	0.088
8x	248	77	0.485	1.106	26	0.071

doi:10.1371/journal.pone.0129861.t002

the variation was again detected within groups (97%), while only 3% of the variation was detected between ploidy levels (Φ_{PT} = 0.026; P < 0.005). Finally, the same pattern was detected when considering soil type, since 99% of the variation was detected within groups, while only 1% of the variation was detected between the two soil types (Φ_{PT} = 0.012; P < 0.05). With regard to the PCoA analysis, the first three components explain 43.27% and 64.55% of total variance of data when considering individuals and populations, respectively. At individual level, in the scatter-plot for the two first components (35.70% of the variance), individuals are distributed across the graph without any grouping pattern, regardless of the population to which they belong (S2 Fig). Similarly, a lack of clustering pattern is again recovered when considering populations (Fig 7) in the scatter-plot for the two first components (53.49% of the variance). No clusters of individuals or populations were detected based either on their geographic distribution, or their chromosome number (Fig 7). According to the Mantel test (Fig 8), no correlation

Table 3. Pair-wise population matrix. Nei's unbiased genetic distance (D) is shown below diagonal and values for population genetic differentiation (Φ_{PT}) are shown above diagonal. For abbreviations of populations, see Table 1. *P<0.05.

	Av1	B1	B2	Cs1	Gi1	Gu1	Hu1	Hu2	Hu3	Hu4	L1	Le1	Le2	Lo1	Lu1	M1	Na1	Na2	Pa1	S1	S2	Sa1	Sa2	T1	Te1	Vi1
Av1	0.000	0.131*	0.062*	0.169*	0.105*	0.187*	0.168*	0.089*	0.189*	0.106*	0.025	0.140*	0.091*	0.191*	0.216*	0.199*	0.125*	0.113*	0.032	0.101*	0.164*	0.000	0.106*	0.056	0.158*	0.106*
B1	0.020	0.000	0.088*	0.120*	0.101*	0.062*	0.136*	0.071*	0.069*	0.075*	0.123*	0.079*	0.147*	0.142*	0.103*	0.099*	0.122*	0.046*	0.165*	0.250*	0.044*	0.151*	0.110*	0.059*	0.068*	0.088*
B2	0.021	0.017	0.000	0.111*	0.104*	0.095*	0.118*	0.012	0.121*	0.015	0.050	0.074*	0.076*	0.116*	0.133*	0.147*	0.125*	0.037	0.086*	0.160*	0.069*	0.039	0.119*	0.084*	0.054	0.022
Cs1	0.027	0.018	0.024	0.000	0.180*	0.104*	0.209*	0.095*	0.130*	0.149*	0.194*	0.148*	0.205*	0.187*	0.120*	0.156*	0.174*	0.076*	0.214*	0.295*	0.101*	0.183*	0.158*	0.128*	0.111*	0.094*
Gi1	0.038	0.021	0.040	0.045	0.000	0.141*	0.175*	0.082*	0.175*	0.140*	0.069	0.180*	0.124*	0.237*	0.221*	0.155*	0.148*	0.099*	0.105*	0.165*	0.146*	0.107*	0.112*	0.087*	0.148*	0.133*
Gu1	0.031	0.009	0.014	0.019	0.032	0.000	0.118*	0.061*	0.030	0.132*	0.181*	0.067*	0.161*	0.137*	0.020	0.073*	0.189*	0.050*	0.200*	0.267*	0.046*	0.180*	0.123*	0.082*	0.073*	0.079*
Hu1	0.026	0.023	0.023	0.036	0.031	0.020	0.000	0.067*	0.101*	0.190*	0.173*	0.182*	0.254*	0.194*	0.135*	0.226*	0.151*	0.116*	0.220*	0.316*	0.112*	0.164*	0.145*	0.100*	0.154*	0.047*
Hu2	0.024	0.015	0.015	0.020	0.034	0.014	0.013	0.000	0.036	0.044	0.067	0.094*	0.083*	0.097*	0.097*	0.110*	0.075*	0.033	0.119*	0.177*	0.058*	0.074*	0.109*	0.065*	0.067*	0.000*
Hu3	0.027	0.017	0.018	0.020	0.035	0.007	0.015	0.004	0.000	0.146*	0.168*	0.094*	0.175*	0.097*	0.053*	0.150*	0.112*	0.057*	0.214*	0.274*	0.066*	0.196*	0.141*	0.074*	0.067*	0.065*
Hu4	0.028	0.012	0.015	0.016	0.035	0.011	0.035	0.020	0.016	0.000	0.081	0.083*	0.075*	0.103*	0.165*	0.150*	0.140*	0.043	0.116*	0.189*	0.097*	0.082*	0.179*	0.085*	0.042	0.089*
L1	0.015	0.006	0.012	0.022	0.018	0.008	0.012	0.020	0.007	0.018	0.000	0.090*	0.011	0.193*	0.215*	0.179*	0.100*	0.080*	0.026	0.032	0.164*	0.004	0.170*	0.054	0.154*	0.105*
Le1	0.030	0.015	0.013	0.015	0.046	0.005	0.027	0.020	0.011	0.013	0.011	0.000	0.075*	0.174*	0.077*	0.103*	0.167*	0.030	0.109*	0.174*	0.076*	0.136*	0.155*	0.079*	0.089*	0.109*
Le2	0.036	0.011	0.020	0.020	0.042	0.009	0.035	0.020	0.014	0.011	0.016	0.005	0.000	0.226*	0.215*	0.142*	0.159*	0.093*	0.046	0.028	0.179*	0.051	0.235*	0.095*	0.149*	0.159*
Lo1	0.032	0.022	0.017	0.023	0.049	0.019	0.026	0.015	0.013	0.030	0.018	0.026	0.031	0.000	0.163*	0.225*	0.216*	0.140*	0.257*	0.323*	0.169*	0.208*	0.184*	0.156*	0.200*	0.121+
Lu1	0.028	0.015	0.017	0.025	0.042	0.009	0.035	0.015	0.007	0.015	0.007	0.005	0.011	0.000	0.000	0.147*	0.211*	0.066*	0.256*	0.334*	0.056*	0.217*	0.159*	0.118*	0.112*	0.094*
M1	0.047	0.020	0.037	0.019	0.040	0.019	0.044	0.030	0.029	0.028	0.024	0.026	0.020	0.019	0.023	0.000	0.220*	0.091*	0.178*	0.235*	0.132*	0.182*	0.178*	0.123*	0.145*	0.168*
Na1	0.024	0.019	0.023	0.026	0.034	0.027	0.019	0.011	0.013	0.022	0.013	0.028	0.026	0.031	0.024	0.041	0.000	0.015	0.143*	0.237*	0.159*	0.163*	0.184*	0.069*	0.085*	0.097*
Na2	0.021	0.012	0.017	0.022	0.031	0.011	0.025	0.018	0.012	0.014	0.006	0.007	0.010	0.026	0.009	0.024	0.015	0.000	0.123*	0.207*	0.011	0.105*	0.097*	0.058*	0.017	0.021*
Pa1	0.019	0.019	0.016	0.010	0.030	0.020	0.026	0.020	0.021	0.020	0.013	0.011	0.016	0.037	0.022	0.027	0.020	0.012	0.000	0.019	0.206*	0.020	0.190*	0.076	0.176*	0.168*
S1	0.033	0.016	0.022	0.023	0.028	0.008	0.025	0.017	0.008	0.026	0.009	0.007	0.011	0.029	0.014	0.016	0.025	0.012	0.010	0.000	0.298*	0.034	0.298*	0.132*	0.277*	0.258*
S2	0.019	0.007	0.013	0.021	0.019	0.009	0.021	0.017	0.011	0.015	0.006	0.007	0.013	0.025	0.007	0.027	0.019	0.008	0.015	0.012	0.000	0.175*	0.092*	0.111*	0.004	0.032*
Sa1	0.009	0.009	0.017	0.016	0.025	0.009	0.010	0.013	0.010	0.024	0.002	0.012	0.016	0.019	0.008	0.024	0.020	0.007	0.011	0.009	0.010	0.000	0.175*	0.070	0.175*	0.114*
Sa2	0.006	0.022	0.023	0.033	0.022	0.025	0.026	0.024	0.025	0.031	0.013	0.027	0.034	0.031	0.028	0.037	0.029	0.022	0.019	0.024	0.018	0.011	0.000	0.132*	0.138*	0.089*
T1	0.018	0.009	0.022	0.014	0.023	0.008	0.018	0.014	0.008	0.014	0.008	0.018	0.018	0.026	0.011	0.025	0.015	0.011	0.022	0.015	0.012	0.010	0.023	0.000	0.095*	0.099*
Te1	0.026	0.014	0.014	0.018	0.035	0.015	0.030	0.012	0.012	0.007	0.014	0.014	0.014	0.033	0.017	0.032	0.009	0.012	0.016	0.019	0.007	0.025	0.028	0.000	0.000	0.038*
Vi1	0.013	0.019	0.012	0.021	0.037	0.019	0.010	0.009	0.010	0.026	0.005	0.018	0.020	0.018	0.015	0.042	0.010	0.012	0.015	0.018	0.014	0.007	0.020	0.016	0.016	0.000

Table 4. Pair-wise ploidy matrix. Nei's unbiased genetic distance (D) is shown below diagonal and values for genetic differentiation (Φ_{PT}) are shown above diagonal. *P<0.05.

	2*x*	4*x*	6*x*	8*x*
2*x*	0.000	0.105*	0.065*	0.086*
4*x*	0.030	0.000	0.080*	0.093*
6*x*	0.024	0.010	0.000	0.005*
8*x*	0.026	0.011	0.002	0.000

doi:10.1371/journal.pone.0129861.t004

was detected between genetic and geographic distance, and therefore, the hypothesis of isolation by distance (IBD) was rejected (Rxy = 0.059; P = 0.198).

Discussion

Our results for Bayesian clustering (Figs 4 and 6), nonhierarchical clustering, AMOVA (Table 4) and PcoA (Fig 7 and S2 Fig), as well as the low values of pair-wise population differentiation, pair-wise genetic distance and the absence of IBD (Fig 8) all suggest the existence of a substantial panmictic yarrow population in the Iberian Peninsula: although this species presents displays clonal reproduction, outbreeding predominates [14], and this is consistent with panmixia. The weak geographical genetic structure recovered is linked to the high levels of gene flow between populations [46] and to the fact that most of the genetic variability is found within populations. These results are also consistent with the assumption that a recent colonization event and rapid expansion of the species occurred throughout the Iberian Peninsula involving especially octoploid and hexaploid cytotypes. Population structure usually develops by means of population genetic differentiation, either by adapting to local conditions or via random genetic drift. However, this process is slow and there has been too little time since colonization for this to be reflected in neutral markers such as microsatellites [49]. With reference to local adaptation, *A. millefolium* in Iberia grows only in ruderal and nitrified environments, and this ecological uniformity may, in turn, have introduced similar adaptive pressures to the entire colonized area and may, in part, have been responsible for the genetic uniformity of this species. Moreover, a study investigating the colonization of North America [50], where yarrow displays maximum ecological diversity within its extensive distribution range, recovered low levels of genetic diversity and negligible genetic structure (associated with varietal identity, geographical distribution or ploidy level), as well as maximum variation in terms of intra-population differences. To summarize, virtually the same genetic pattern that we recovered in the Iberian Peninsula was found for North America [50]. Given that the great taxonomic diversity present in North America is interpreted to be mainly due to ecological adaptation to habitat differences (such as in the case of ecological specialists like *A. millefolium* var. *gigantea* or *A. millefolium* var. *puberula*) and phenological differences (for example, between *A. millefolium* var. *arenicola* and *A. millefolium* var. *littoralis*), the few taxa comprising Iberian *A. millefolium* agg. may be due to their uniform ecological preferences (i.e., ruderal and disturbed habitats).

Similarly, the low levels of genetic variability recovered suggest that bottlenecks and/or founder events may have been involved in this process [51,52]. When a population passes through a bottleneck, its genetic diversity is expected to diminish to a greater or lesser degree depending on both the size of the founding population and the rate of population growth. However, when population size increases, genetic diversity may increase by the occurrence of new mutations [51]. In our opinion, the presumed genetic influence of population size and growth following introduction of *A. millefolium* to the Iberian Peninsula has not had sufficient time to produce new variations and to develop a clear genetic structure [49]. Furthermore, this
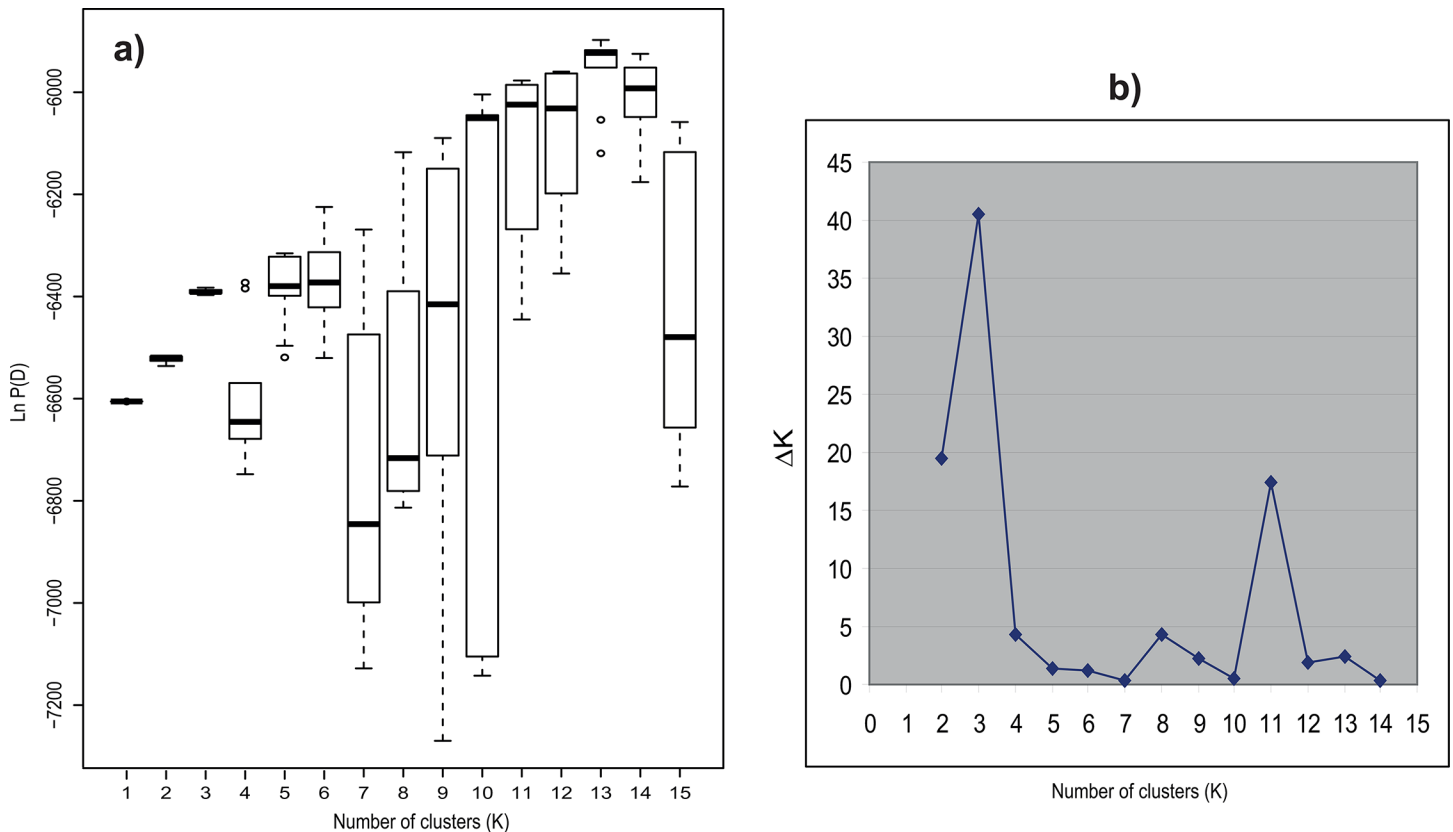
Fig 3. a) Ln P(D), mean (±standard deviation) of log-likelihood values for each value of K = 1–15 (10 independent runs per K). b) Mean absolute difference of the second order rate of change with respect to K (ΔK, following [44]). Most supported K value: K = 3.

doi:10.1371/journal.pone.0129861.g003

process may have been counter-balanced by the influence of clonal reproduction, which may also have played an important role in maintaining genetic impoverishment [53]. Previous studies of *A. millefolium* agg. phylogeny based on both nuclear and plastid genomes found low levels of genetic variation and limited resolution [12]. According to [54], the genetic consequences of the colonization of a new area are influenced not only by the number of introductions, but also by the number of propagules introduced at each introduction event. Low levels of genetic variation are commonly explained by few introductions and small founding population sizes [54,55]. Conversely, a pattern of great genetic diversity within populations is explained by the occurrence of multiple introductions from different genetic lineages present in
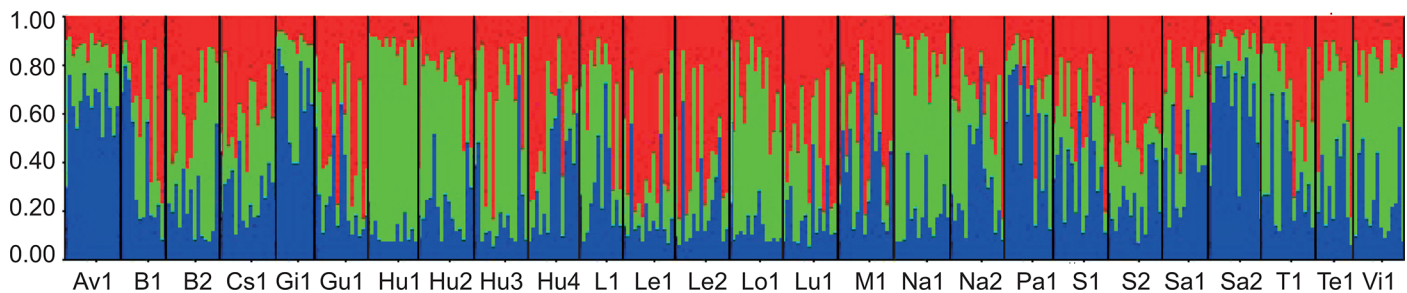
Fig 4. Percentage assignment of each individual (represented by vertical bars) to each of the three genetic clusters (represented by different colours) inferred by the program Structure [43]. See Table 1 for population codes.

doi:10.1371/journal.pone.0129861.g004
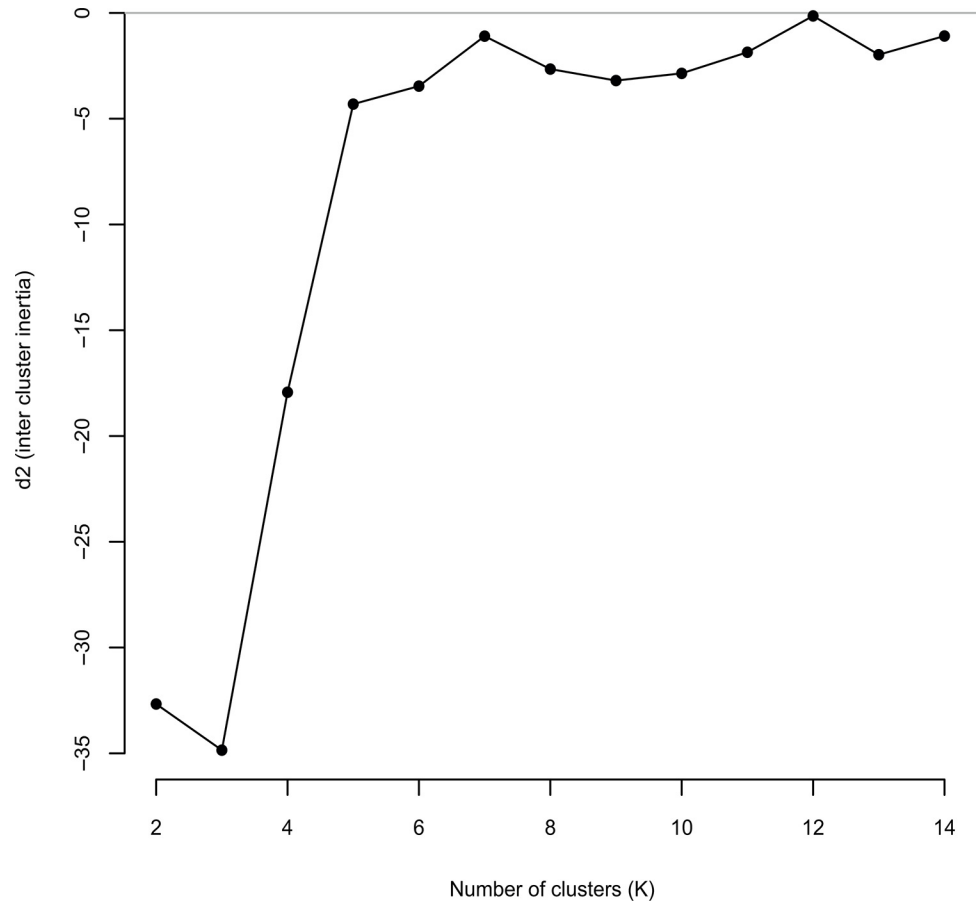
Fig 5. Screening for the most likely number of groups (K) with non-hierarchical K-means clustering [45], performed with 100000 independent runs for each value of K. The runs maximising ΔK values were initially considered as optimal (K = 7 and K = 12), and because K = 12 added little gain to ΔK, we kept K = 7 as the optimal final value.

doi:10.1371/journal.pone.0129861.g005

the indigenous area [54,56–60]. Our results suggest that the number of both colonization events and propagules may have been low during the Iberian colonization by *A. millefolium*,
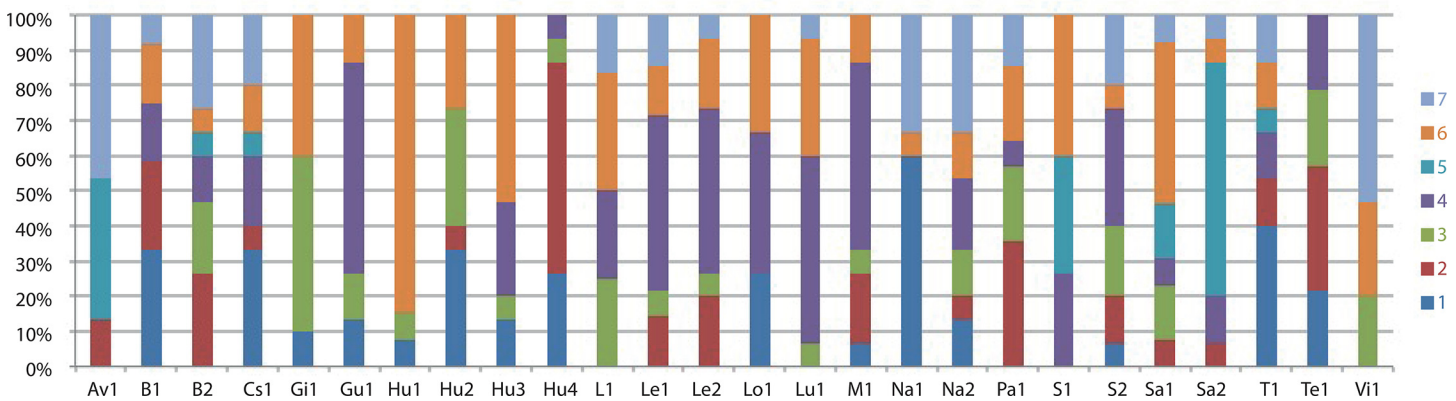
Fig 6. Percentage of individuals that belong to each of the seven genetic clusters (represented by different colours) inferred by the nonhierarchical K-means clustering analysis [46]. Each vertical column corresponds to a population. See Table 1 for population codes.

doi:10.1371/journal.pone.0129861.g006
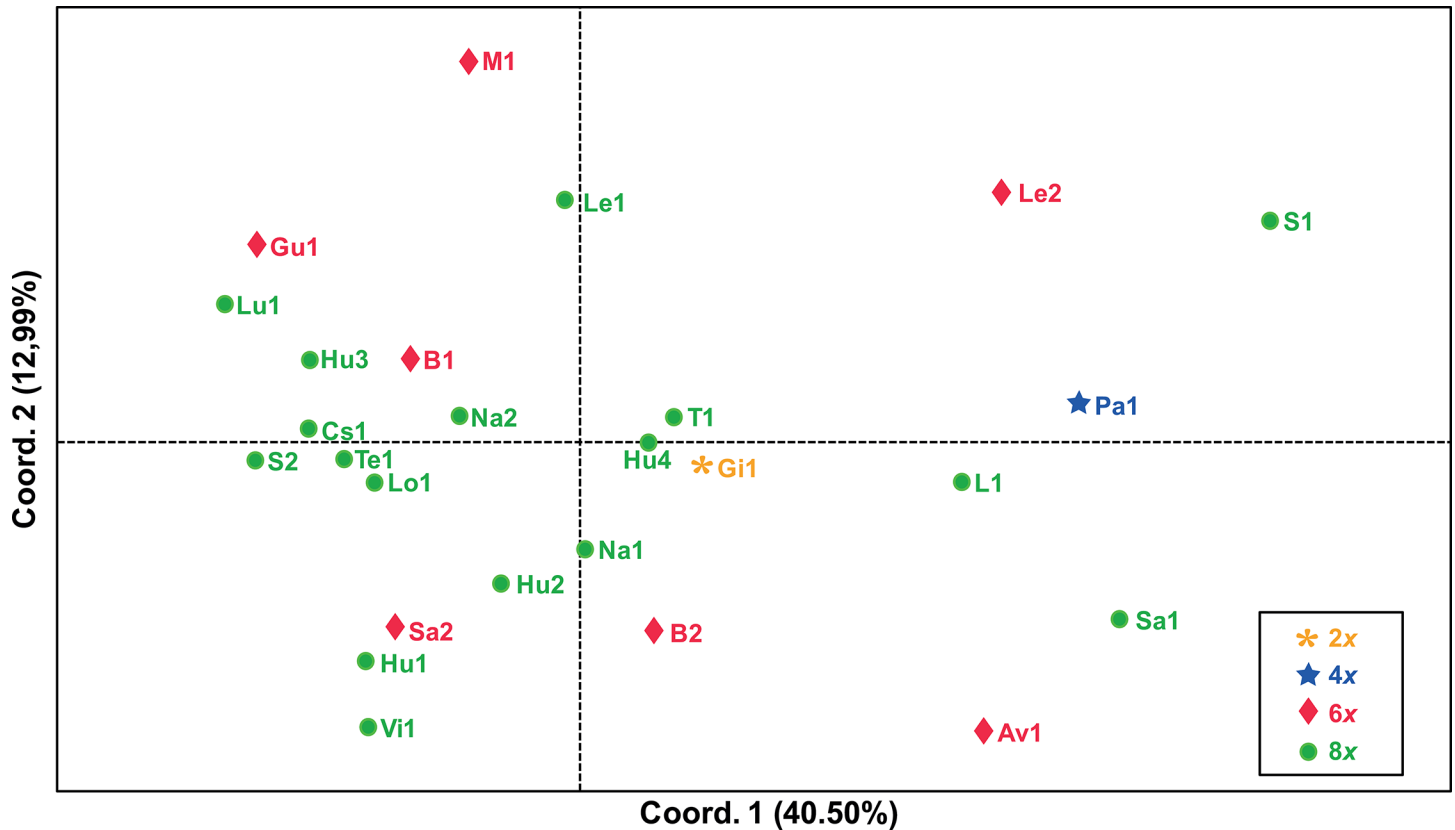
Fig 7. Principal component analysis performed from pairwise Nei's genetic distances between populations. Orange asterisk: 2*x*; blue star: 4*x*; pink diamond: 6*x*; green dot: 8*x*. See Table 1 for population codes.

doi:10.1371/journal.pone.0129861.g007
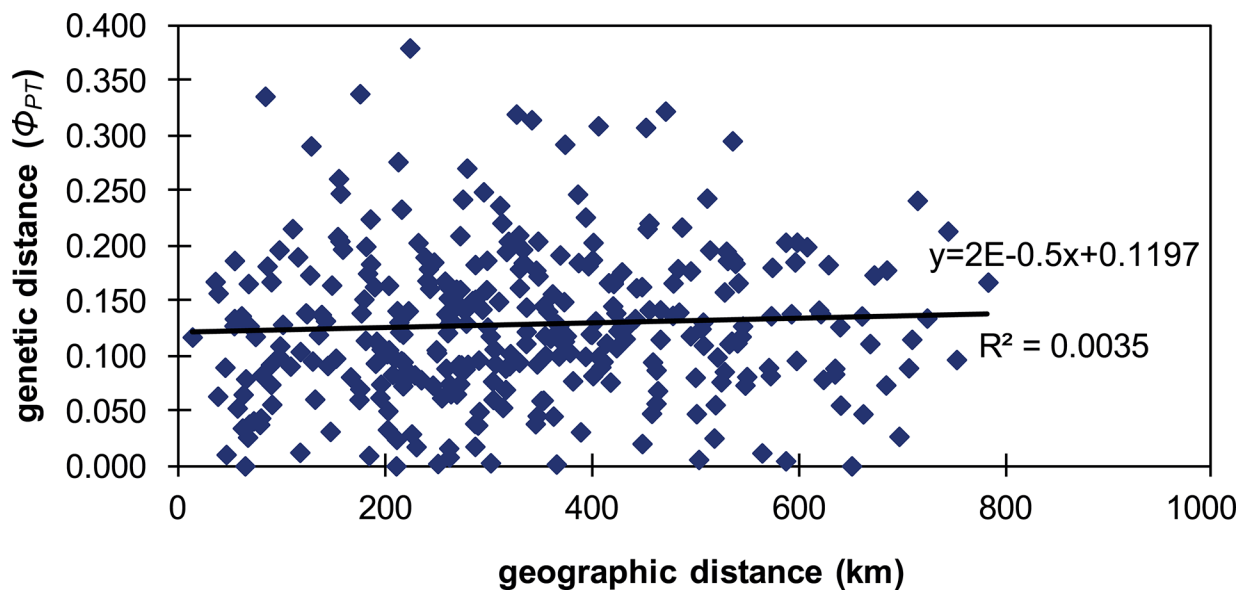
Fig 8. Mantel test correlating geographic distance and genetic distance. Results of the IBD test searching for a correlation between geographic distance and genetic distance.

doi:10.1371/journal.pone.0129861.g008

and thus, the Iberian Peninsula should be considered both in evolutionary and geographical terms, a dead-end for this species. However, given its therapeutic properties and ornamental value, several independent, human-mediated introductions (intentionally or accidentally by livestock) cannot be ruled out, and this may explain both the low levels of genetic diversity and the presence of exclusive alleles found in most of the populations studied.

Another factor that may affect genetic variability is invasiveness. For example, *Rubus alceifolius* displays least genetic diversity in areas where it has become a serious weed [55], and it has been suggested that a genotype well-adapted to a specific biotope may spread rapidly through an area by means of asexual reproduction. *Achillea millefolium* is considered such a weed and meets the parameters attributed to invasive plants [61], namely, great asexual reproductive potential and high ploidy levels. According to our data, the behavior of *A. millefolium* during the colonization of the Iberian Peninsula may be considered as invasive.

With reference to the correlation between ploidy level and genetic diversity in *A. millefolium*, it is expected to be greater for polyploids. Polyploidy allows more recombination events to occur during meiosis and therefore results in increased diversity of the genome, thus allowing adaptive plasticity [24,62] and reducing the cost of inbreeding [55]. According to our results, yarrow from the Iberian Peninsula does not display this pattern. Consistently, [3] found that throughout Europe, the number of AFLP bands per individual were not significantly greater in polyploid genomes than in those of diploids and suggested that this may indicate genome down-sizing or sequence suppression during polyploidization. These authors also discovered that the number of polymorphic bands was significantly greater for polyploid populations compared to diploid populations, suggesting a greater degree of genetic polymorphism in polyploids. Indeed, the low levels of genetic differentiation between ploidy levels recovered in our work also suggest an absence of barriers between them, even between diploids and the remaining cytotypes. Negligible genetic differentiation was also observed to exist between $4x$ and $6x$ ecotypes of *A. millefolium* agg. in America [50] and this [3] highlighted the existence of hybrid contacts and several independent lines of auto- and allopolyploidy both within the same and between different ploidy levels of *A. millefolium* agg. This may have masked the separation between sympatric populations and highlights the lack of barriers between polyploids, even those of different species. Consequently, previously published karyological data [63–65], prove that karyotypes of *A. millefolium* agg. are quite similar, show limited structural differentiation and lack obvious obstacles to hybridization.

Conclusions

All of our results indicate a very recent colonization of the Iberian Peninsula by the *Achillea millefolium* aggregate, and this involved a limited number of individuals. The species behaves in Iberia as an invasive plant colonizing ruderal and disturbed habitats. The low genetic variability of populations and lack of differentiation between the different ploidy levels all indicate that the Iberian Peninsula constitutes a geographical cul-de-sac for the expansion of the aggregate and a genetic and evolutionary dead-end for the Iberian populations.

Supporting Information

S1 Dataset. Binary-coded matrix of the microsatellite results for the studied populations.
0 = absence; 1 = presence. See Table 1 for population codes.
(TXT)

S1 Fig. Box-plots of the measurements of the stomata. SL = Stomata length;
SW = Stomata width.
(EPS)

S2 Fig. Principal component analysis performed from pairwise Nei's genetic distances between individuals. See Table 1 for population codes.
(EPS)

S1 Table. Dimensions of stomata for individual populations of known ploidy (by count). See Table 1 for population codes.
(DOCX)

S2 Table. Dimensions of stomata for ploidy level. See Table 1 for population codes.
(DOCX)

S3 Table. Dimensions of stomata and estimated ploidy level for populations withouth counts. See Table 1 for population codes.
(DOCX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: IS JMM NGJ AS. Performed the experiments: SLV. Analyzed the data: SLV CR IS. Contributed reagents/materials/analysis tools: IS JMM AS NGJ. Wrote the paper: AS IS NGJ SLV.

References

1. Ehrendorfer F, Guo Y-P (2006) Multidisciplinary studies on *Achillea* sensu lato (Compositae-Anthemideae): New data on systematics and phylogeography. Willdenowia 36: 69–87.

2. Guo Y-P, Saukel J, Mittermayr R, Ehrendorfer F (2005) AFLP analyses demonstrate genetic divergence, hybridization and multiple polyploidization in the evolution of *Achillea* (Asteraceae-Anthemideae). New Phytol 166: 273–290. PMID: 15760370

3. Guo Y-P, Saukel J, Ehrendorfer F (2008) AFLP trees versus scatterplots: evolution and phylogeography of the polyploid complex *Achillea millefolium* agg. (Asteraceae). Taxon 57: 153–169.

4. Guo Y-P, Wang S- Z, Vogl C, Ehrendorfer F (2012) Nuclear and plastid haplotypes suggest rapid diploid and polyploid speciation in the N Hemisphere *Achillea millefolium* complex (Asteraceae). BMC Evol Biol 12: 2. doi: 10.1186/1471-2148-12-2 PMID: 22214230

5. Candan F, Unlu M, Tepe B, Daferera D, Polissiou M, Sökmen A, et al. (2003) Antioxidant and antimicrobial activity of the essential oil and methanol extracts of *A. millefolium*. J Ethnopharmacol 87: 215–220. PMID: 12860311

6. Stojanović G, Radulović N, Hashimoto T, Palić R (2005) In vitro antimicrobial activity of extracts of four species: the composition of *A. millefolium* L. (Asteraceae) extract. J Ethnopharmacol 101: 185–190. PMID: 15978758

7. Yaeesh S, Jamal Q, Khan AU, Gilani AH (2006) Studies on hepatoprotective, antispasmodic and calcium antagonist activities of the aqueous-methanol extract of *A. millefolium*. Phytother Res 20: 546–551. PMID: 16619341

8. Lehane AM, Saliba KJ (2008) Common dietary flavonoids inhibit the growth of the intraerythrocytic malaria parasite. BMC Res Notes 1: 26–30. doi: 10.1186/1756-0500-1-26 PMID: 18710482

9. Dall'Acqua S, Bolego C, Cignarella A, Gaion RM, Innocenti G (2011) Vasoprotective activity of standardized *A. millefolium* extract. Phytomedicine 18: 1031–1046. doi: 10.1016/j.phymed.2011.05.005 PMID: 21684130

10. De Souza P, Gasparotto A, Crestani S, Stefanello M, Marques M, Kassuya CA (2011) Hypotensive mechanism of the extracts and artemetin isolated from *A. millefolium* L. (Asteraceae) in rat. Phytomedicine 18: 819–825. doi: 10.1016/j.phymed.2011.02.005 PMID: 21420289

11. Akram M (2013) Minireview on *Achillea millefolium* Linn. J Membr Biol 246: 661–663. PMID: 23959026

12. Guo Y-P, Ehrendorfer F, Samuel R (2004) Phylogeny and systematics of *Achillea* (Asteraceae-Anthemideae) inferred from the nrITS and plastid trnL-F DNA sequences. Taxon 53: 657–672.

13. Queirós M (1973) Contribuição para o conhecimento citotaxonómico das Spermatophyta de Portugal. II. Compositae, Supl. 1. Bol Soc Brot, sér. 2 47: 299–314.

14. Vetter S, Lambrou M, Franz CH, Ehrendorfer F, Saukel J (1996) Chromosome numbers of experimental tetraploid hybrids and selfpollinated progenies within the *Achillea millefolium* complex (Compositae). Caryologia 49(3/4): 227–231.

15. Soriano I, Bernal M, Flordelis G, Vallès J (2013) Una aproximación multidisciplinar al complejo de *Achillea millefolium* (Asteraceae) en el norte de la península ibérica y áreas próximas. 10è Colloque de Botanique Pyrénéo-cantabrique, Bagnères-de-Luchon 2013. Résumés des communications. pp. 18.

16. Meusel H, Jäger EJ (1991) Vergleichende Chorologie der zentraleuropäischen Flora, vol. 3. Jena: G. Fischer.

17. Harbaugh DT, Wagner WL, Percy DM, James HF, Fleischer RC (2009) Genetic structure of the polymorphic *Metrosideros* (Myrtaceae) complex in the Hawaiian Islands using nuclear microsatellite data. PLoS ONE 4(3): e4698. doi: 10.1371/journal.pone.0004698 PMID: 19259272

18. Lepais O, Muller SD, Ben Saad-Limam S, Benslama M, Rhazi L, Belouahem-Abed D, et al. (2013) High Genetic Diversity and Distinctiveness of Rear-Edge Climate Relicts Maintained by Ancient Tetraploidisation for *Alnus glutinosa*. PLoS ONE 8(9): e75029. doi: 10.1371/journal.pone.0075029 PMID: 24098677

19. Zhang D, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol Ecol 12: 563–584. PMID: 12675814

20. Esselink GD, Nybom H, Vosman B (2004) Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. Theor Appl Genet 109: 402–408. PMID: 15085263

21. Markwith SH, Stewart DJ, Dyer JL (2006) TETRASAT: a program for the population analysis of allotetraploid microsatellite data. Mol Ecol Notes 6: 586–589.

22. Espinoza NR, Noor MAF (2002) Population genetics of a polyploid: Is there hybridization between lineages of *Hyla versicolor*? J Hered 93: 81–85. PMID: 12140266

23. Bruvo R, Michiels NK, D'Souza T, Schulenburg H (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol Ecol 13: 2101–2106. doi: 10.1111/j.1365-294X.2004.02209.x PMID: 15189230

24. Andreakis N, Kooistra WHCF, Procaccini G (2009) High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. Mol Ecol 18: 212–226. doi: 10.1111/j.1365-294X.2008.04022.x PMID: 19192176

25. Huang Y, Li Y, Hu X, Ge XJ, Zhang CQ, Long CL (2009) Development of twelve polymorphic microsatellite loci in polyploid endangered *Omphalogramma vincaeflora* Franch. (Primulaceae). Conserv Genet 10: 515–517.

26. Masterson J (1994) Stomatal size in fossils: evidence for polyploidy in majority of angiosperms. Science 264: 421–424. PMID: 17836906

27. Soltis DE, Soltis PS (1993) Molecular data and the dynamic nature of polyploidy. Crit Rev Plant Sci 12: 243–273.

28. Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. Trends Ecol Evol 14: 348–352. PMID: 10441308

29. Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng CF, et al. (2009) Polyploidy and angiosperm diversification. Am J Bot 96: 336–348. doi: 10.3732/ajb.0800079 PMID: 21628192

30. Jiao YN, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, et al. (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473: 97–100. doi: 10.1038/nature09916 PMID: 21478875

31. Samadi S, Mavárez J, Pointier JP, Delay B, Jarne P (1999) Microsatellite and morphological analysis of population structure in the parthenogenetic freshwater snail *Melanoides tuberculata*: insights into the creation of clonal variability. Mol Ecol 8: 1141–1153. PMID: 10447855

32. Becher SA, Steinmetz K, Weising K, Boury S, Peltier D, Renou JP (2000) Microsatellites for cultivar identification in *Pelargonium*. Theor Appl Genet 101: 643–651.

33. Saltonstall K (2003) Microsatellite variation within and among North American lineages of *Phragmites australis*. Mol Ecol 12: 1689–1702. PMID: 12803624

34. Babaei A, Tabaei-Aghdaei SR, Khosh-Khui M, Omidbaigi R, Naghavi MR, Esselink GD, et al. (2007) Microsatellites analysis of Damask rose (*Rosa damascena* Mill.) accessions from various regions in Iran reveals multiple genotypes. BMC Plant Biol 7: 12. doi: 10.1186/1471-2229-7-12 PMID: 17346330

35. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11–15.

36. Cullings KW (1992) Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. Mol Ecol 1: 233–240.

37. Tel-Zur N, Abbo S, Myslabodski D, Mizrahi Y (1999) Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera *Hylocereus* and *Selenicereus* (Cactaceae). Plant Mol Biol Rep 17: 249–254.

38. Rahimmalek M, Sayed Tabatabaei ME, Arzani A, Khorrami M (2011) Development and characterization of microsatellite markers for genomic analysis of yarrow (*Achillea millefolium* L.). Genes Genomics 33: 475–482.

39. López-Vinyallonga S, Arakaki M, Garcia-Jacas N, Susanna A, Gitzendanner MA, Soltis DE, et al. (2010) Isolation and characterization of novel microsatellite markers for *Arctium minus* L (Compositae). Am J Bot Notes Protoc Plant Sci 97: e4–e6. doi: 10.3732/ajb.0900376 PMID: 21622378

40. Landergott U, Naciri Y, Schneller JJ, Holderegger R (2006) Allelic configuration and polysomic inheritance of highly variable microsatellites in tetraploid gynodioecious *Thymus praecox* agg. Theor Appl Genet 109 113: 453–465. PMID: 16786342

41. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288–295.

42. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537–2539. PMID: 22820204

43. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959. PMID: 10835412

44. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation Study. Mol Ecol 14: 2611–2620. PMID: 15969739

45. Hartigan JA, Wong MA (1979) Algorithm AS-136: A K-means clustering algorithm. Appl Stat-J Roy St C 28: 100–108.

46. Arrigo N, Felber F, Parisod C, Buerki S, Alvarez N, David J, et al. (2010) Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. New Phytol 187: 1170–1180. doi: 10.1111/j.1469-8137.2010.03328.x PMID: 20561204

47. Legendre P, Legendre L (1998) Numerical ecology. Amsterdam: Elsevier.

48. R Development Core Team (2013) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

49. Marrs RA, Sforza R, Hufbauer RA (2008a) When invasion increases population genetic structure: a study with *Centaurea diffusa*. Biol Invasions 10: 561–572.

50. Ramsey J, Robertson A, Husband B (2008) Rapid adaptive divergence in New World *Achillea*, an auto-polyploid complex of ecological races. Evolution 62: 639–653. PMID: 18039326

51. Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution 29: 1–10.

52. Husband BC, Barrett SCH (1991) Colonisation history and population genetic structure of *Eichornia paniculata* in Jamaica. Heredity 66: 287–296.

53. Poulin J, Weller SG, Sakai AK (2005) Genetic diversity does not affect the invasiveness of fountain grass (*Pennisetum setaceum*) in Arizona, California and Hawaii. Divers Distrib 11: 241–247.

54. Marrs RA, Sforza R, Hufbauer RA (2008b) Evidence for multiple introductions of *Centaurea stoebe micranthos* (spotted knapweed, Asteraceae) to North America. Mol Ecol 17: 4197–4208. PMID: 19378400

55. Amsellem L, Noyer JL, Le Bourgeois T, Hossaert-McKey M (2000) Comparison of genetic diversity in the invasive weed Rubus alceifolius Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Mol Ecol 9: 443–455. doi: 10.1046/j.1365-294x.2000.00876.x PMID: 10736047

56. Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? Proc Natl Acad Sci, USA 97: 7043–7050. PMID: 10860969

57. Durka W, Bossdorf O, Prati D, Auge H (2005) Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. Mol Ecol 14: 1697–1706. PMID: 15836643

58. Gaskin JF, Zhang D-Y, Bon M-C (2005) Invasion of *Lepidium draba* (Brassicaceae) in the Western United States: distributions and origins of chloroplast haplotypes. Mol Ecol 14: 2331–2341. PMID: 15969718

59. Genton BJ, Shykoff JA, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. Mol Ecol 14: 4275–4285. PMID: 16313592

60. Facon B, Pointier J-P, Jarne P, Sarda V, David P (2008) High genetic variance in life-history strategies within invasive populations by way of multiple introductions. Curr Biol 18: 363–367. doi: 10.1016/j.cub. 2008.01.063 PMID: 18334202

61. Barrett SCH, Richardson BJ (1986) Genetic attributes of invading species. In: Groves RH, Burdon JJ, editors. Ecology of Biological Invasions: an Australian Perspective. Canberra: Australian Academy of Science. pp. 21–33.

62. Roose ML, Gottlieb LD (1976) Genetic and biochemical consequences of polyploidy in *Tragopogon*. Evolution 30: 818–830.

63. Schweizer D, Ehrendorfer F (1983) Evolution of C-band pattern in Asteraceae-Anthemideae. Biol Zent Bl 102: 637–655.

64. Ehrendorfer F (1986) Chromosome differentiation and evolution in angiosperm groups. In: Iwatsuki K, Raven PH, Bock WJ, editors. Modern aspects of Species. Tokyo: University Press. pp. 59–86.

65. Lambrou M, Ehrendorfer F (2000) DAPI-banded karyotypes and hybrid compatibility in diploid and tetraploid taxa of *Achillea millefolium* agg. Botanikertagung Jena, 17–22 September 2001, Tagungsband. p. 207.