Molecular Phylogeny of *Artemisia thuscula*

TRABAJO DE FIN DE MÁSTER

Máster en Biomedicina

Curso 2019-2020

DIANA MESA MORILLO

Aceptación de los tutores

Dr. **Luis Fabián Lorenzo Díaz,**

Profesor Ayudante Doctor del Área de Genética, Departamento de Bioquímica, Microbiología, Biología Celular y Genética de la Universidad de La Laguna,

Dr. **Mario Andrés González Carracedo,**

Profesor Ayudante Doctor del Área de Genética, Departamento de Bioquímica, Microbiología, Biología Celular y Genética de la Universidad de La Laguna,

autorizan la presentación, defensa y evaluación de la presente Memoria de Fin de Máster.

En La Laguna a 02 de julio de 2020.

Fdo. Dr. Luis Fabián Lorenzo Díaz   Fdo. Dr. Mario Andrés González Carracedo

Abstract

The *Artemisia* genus includes several plant species with high economic and ecological value, being two examples *A. annua* and *A. thuscula.* While the first one is well characterized since is currently exploited for Artemisinin production, the second represents an endemism from the Canary Islands that has been poorly studied at molecular level, but with high potential for its future industrial exploitation. In the present work, DNA sequences corresponding to two different chloroplast *loci* (*mat*K and *rbc*L) have been obtained for both plant species, which represents the first report of *A. thuscula mat*K and *rbc*L DNA Barcodes. Sequence comparisons revealed some differences between both species. Moreover, a phylogenetic analysis including the *A. thuscula* obtained sequences together with those available for other *Artemisia* species, showed a high degree of conservation throughout the *Artemisia* genera. However, unique genetic features have been revealed for *A. thuscula,* which are not present in other *Artemisia* species included in the present study. Finally, in order to increase micromorphological knowledge about *A. thuscula*, the plant surface of both species have been studied by Scanning Electron Microscopy, discovering significant morphological differences.

Resumen

El género *Artemisia* incluye numerosas especies de plantas con elevado valor económico y ecológico, siendo dos ejemplos *A. annua* y *A. thuscula.* Mientras que la primera especie ha sido muy bien caracterizada debido a su actual uso para la producción de Artemisinina, la segunda representa un endemismo de las Islas Canarias que ha sido poco estudiado a nivel molecular, pero con un elevado potencial para su futura explotación industrial. Es este

trabajo, se han obtenido secuencias de ADN correspondientes a dos *loci* del cloroplasto (*mat*K y *rbc*L), lo que representa la primera vez que se obtiene un código de barras de ADN para estos genes en el caso de *A. thuscula*. La comparación de estas secuencias ha revelado algunas diferencias entre ambas especies. Además, un análisis filogenético donde se ha incluido las secuencias obtenidas para *A. thuscula* junto con aquellas disponibles para otras especies del género *Artemisia*, ha mostrado una alta conservación de estos dos genes entre las diferentes especies del género. Sin embargo, también ha revelado características genéticas únicas de *A. thuscula*, que no están presentes en ninguna otra de las especies estudiadas. Finalmente, con el objetivo de incrementar el conocimiento sobre las características micromorfológicas de *A. thuscula*, la superficie de ambas especies ha sido estudiada mediante Microscopía Electrónica de Barrido, lo que ha permitido descubrir importantes diferencias morfológicas.

Introduction

The genus *Artemisia*: biological significance and taxonomy.

The *Artemisia* genus (tribe *Anthemideae,* family *Asteraceae*) contains more than 500 plant species, several with high economic value and ecological significance (Hayat *et al.,* 2009). A great number of *Artemisia* species have shown beneficial properties and, therefore, have been exploited in different fields, such as medicine, chemistry or food industry (Hussain *et al.,* 2017).The importance of this genus can be highlighted by a search in the PubMed database, including "*Artemisia*" as query, which yielded 2668 publications in the last decade (**Fig. 1**).

Figure 1. Scientific relevance of the *Artemisia* genus and its taxonomical classification. The number of scientific publications contained in PubMed database, corresponding to last years, are shown. Searches included "*Artemisia*" (black) or *"Artemisia AND phylogeny"* (grey) as query.
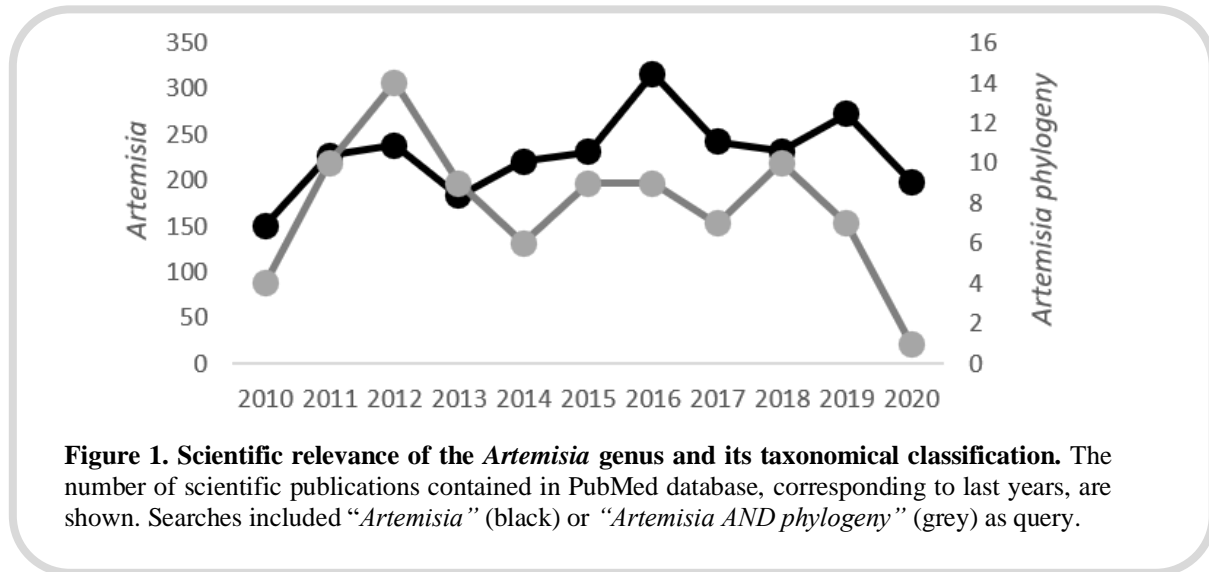
During this time, several efforts have been carried out to obtain the so called "DNA Barcodes" for all groups of living organisms (Fazekas *et al.,* 2012; Rich, Trinder and Long, 2015), since this information not only supports their taxonomical characterization, but also improves the biodiversity conservation and supports the correct management of species (Rich, Trinder and Long, 2015). In this sense, numerous studies have attempted to taxonomically classify the *Artemisia* genus, starting from botanical characters and, during the last three decades, including molecular phylogenetic and phylogenomic approaches (Vallès *et al.,* 2011). In fact, 86 publications were found in PubMed, since 2010, when "*Artemisia AND* phylogeny" was searched (**Fig. 1**).

Botanical approaches have been mainly based on morphology, which allowed the *Artemisia* subclassification into five subgenera (*Artemisia*, *Absinthium*, *Seriphidium*, *Dracunculus* and *Tridentatae*), but with some controversy related to the consideration of *Seriphidium* as an independent genus (Hussain, Potter, *et al.,* 2019). Moreover, alternative approaches have been carried out, in base of the distribution and morphology of foliar trichomes (Hayat *et al.,* 2009) and other epidermal anatomical characteristics (Hussain,

Hayat, *et al.,* 2019). These studies have suggested that morphology of foliar characters, studied by Scanning Electron Microscopy (SEM), can be used as good taxonomic markers to resolve the subclassification of the *Artemisia* genus (Hussain, Hayat, *et al.,* 2019).

However, the most important advances in this field have been achieved by molecular phylogenies based on Internal Transcribed Spacer (ITS) sequences (Kornkven, Watson, and Estes. 1998; Torrell *et al.,* 1999; Watson *et al.,* 2002; Vallès *et al.,* 2003), chloroplast DNA restriction variation (Kornkven, Watson and Estes, 1999) or ITS combined with External Transcribed Spacer (ETS) (Sanz *et al.,* 2008; Pellicer *et al.,* 2010), and/or with other nuclear and chloroplast DNA markers (Garcia *et al.,* 2011; Riggins and Seigler, 2012; Hobbs and Baldwin, 2013; Haghighi *et al.,* 2014; Malik *et al.,* 2017; Hussain, Potter, *et al.,* 2019). In addition, Maturase K (*mat*K) based phylogenies and haplotype network analysis have been recently performed including several *Artemisia* species (Turuspekov *et al.,* 2018). Therefore, due to the advances in Next Generation Sequencing technology, complete chloroplast genomes have been assembled for different *Artemisia* species during last years (Meng *et al.,* 2019; X. Shen *et al.,* 2017; Lu *et al.,* 2020; Shahzadi *et al.,* 2020; Iram *et al.,* 2019; Lim *et al.,* 2018; Min *et al.,* 2019; P. Li and Jia, 2019; Nangong, He and Huang, 2020; Peng *et al.,* 2018; Y. S. Lee *et al.,* 2016; Kang *et al.,* 2016). Furthermore, the complete genome of *A. annua* has been recently sequenced (Q. Shen *et al.,* 2018). However, genomic-level information is still limited to a reduced number of *Artemisia* species, and a higher number of taxa need to be sequenced to obtain genome-based phylogenies that overall represent the different species contained in this genus.

Nowadays, after different taxonomic rearrangements, the five subgenera of the *Artemisia* genus are mainly accepted, in spite of subgenera *Tridentatae* have been recently re-included as a subgenera, with molecular evidence (Torrell *et al.,* 1999; Watson *et al.,* 2002). However, the subclassification of *Artemisia* genus remains not fully understood, because assignments of some taxa by molecular phylogenies are not consistent with morphology-based classification, and molecular data for some species are not available or are incomplete. Moreover, phylogenetic inference based on nuclear and chloroplast *loci*, especially when chloroplast-coding regions are compared with non-coding nuclear regions, usually reports several incongruences (Pellicer *et al.,* 2018; Hussain *et al.,* 2017; Hussain, Potter, *et al.,* 2019).

Artemisia annua highlights the importance of the *Artemisia* genus.

Some taxa of the *Artemisia* genus, especially *A. annua*, have been extensively studied from both ethnobotanical and molecular phylogenetic approaches, mainly due to its medicinal properties (Efferth, 2017). *A. annua* is native from Asia, being part of the steppe populations of plants from the provinces of Chahar and Suiyuan, in the northeast China (Acosta de la Luz and Castro Armas, 2009), and has been used back from almost 2,000 years by the Chinese medicine as an antimalarial treatment (Acosta de la Luz and Castro Armas, 2009; Willcox *et al.,* 2004; Tang and Eisenbrand, 2013). The importance of *A. annua* is related to its ability to synthetize different secondary metabolites with high biological activity, such as sesquiterpenoids, flavonoids, phenolic acids and coumarins, among others (Acosta de la Luz and Castro Armas 2009; Bryant *et al.,* 2015; Croteau, 1986). In fact, *A. annua* has been extensively exploited to produce the chemically-active sesquiterpene Artemisinin (Tang and Eisenbrand, 2013), which nowadays is being used as the

basis of the artemisinin-combination therapies (ACTs). These therapies are currently considered as the standard treatment worldwide for *Plamodium falciparum*, the causal agent of malaria disease in humans (Duffy and Mutabingwa, 2006).

On the other hand, numerous publications support the therapeutic action of *A. annua* against parasitic diseases, such as Schistosomiasis (Gold *et al.,* 2017), Parasitemia (Raffetin *et al.,* 2018) ,Toxoplasmosis (Chorlton, 2017) and Coccidiosis (Fatemi, Asasi, and Razavi, 2017), against viral diseases, such as Human Papillomavirus (HPV) (Satish Kumar *et al.,* 2015) and cytomegalovirus (Drouot, Piret, and Boivin, 2016), or against bacterial infections, such as Lyme disease (Puri, Hakkarainen-Smith, and Monro, 2017). Moreover, *A. annua* biochemical derivatives have been recently used as treatment for autoimmune diseases, such as lupus and multiple sclerosis (Liang *et al.,* 2018), and several studies have tested their antimicrobial activity against different human pathogens. In particular, the so called "essential oil" extracted from this plant has been shown to inhibit the growth of broad spectrum bacteria and fungi, and also to reduce the cytotoxic effects caused by their infection (Ćavar *et al.,* 2012). More recently, these compounds have also been tested as antitumoral agents (Im *et al.,* 2018), and results suggest that are able to reduce tumoral growth of lung, colorectal and intestinal cancer cells (X. Li *et al.,* 2018), being currently involved in phase-II of different clinical trials.

A. annua has been botanically characterized long time ago in the Canary Islands, as can be confirmed by the presence of well conserved specimens in Herbario-TFC (SEGAI-ULL) (**Fig. 2a**). Moreover, *A. annua* has been started to be industrially exploited in this region during the last decade. Currently, greenhouse crops of *A. annua* are grown for commercial uses, which involves

sesquiterpenoids production, especially Artemisinin (**Fig. 2b**). However, the potential use of other *Artemisia* species for similar purposes is currently under study. In this sense, the use of *Artemisia* species which represents endemisms of the Canary Islands is of highly interest, since these endemic species could have unique genetic characteristics that make them more profitable, or even allow the purification of novel bioactive compounds of high economic interest.

Figure 2. Botanical characterization and industrial crops of *A. annua* a) Herbarium voucher specimen (Courtesy of *Herbario TFC, SEGAI-ULL*) and classical botanical description of *A. annua*. b) Cultivars of *A. annua* for industrial Artemisinin production (Courtesy of *Biotech Tricopharming Research S.L., Valle de Guerra, Tenerife, Spain*).

***Artemisia thuscula* represents an endemism with high ecological value and potential for its industrial exploitation.**

Several efforts have been carried out to test the potential of other *Artemisia* species for production of Artemisinin and its derivatives (Pellicer *et al.,* 2018; Mannan *et al.,* 2010; Hussain *et al.,* 2017). Interestingly, recent results

showed that Artemisinin is produced by several *Artemisia* species, including the macaronesian endemism *A. thuscula* (Pellicer *et al.,* 2018), which opens the doors to its potential exploitation for Artemisinin production in the future. *Artemisia thuscula* Cav. *(A. canariensis* Less.*)* is considered as an endemism of the macaronesian region (BDBC, 2020), which grows in semi-arid zones of the Canary Islands (Sainz *et al.,* 2017). Natural populations can be found in the islands of Tenerife, La Gomera, El Hierro, La Palma and Gran Canaria (**Fig. 3a**) (BDBC, 2020). Therefore, the biological value of *A. thuscula* is not only related with its potential use for Artemisinin production, but also with its important role in the conservation of regional biodiversity.

Figure 3. Geographical distribution, morphological features and industrial crops of *A. thuscula.* a) Distribution of wild populations of *A. thuscula* in the Canary Islands. Light blue areas represent coverage with 100% certainty, dark blue dots indicate coverage certainty of 66% and green squares represent coverage certainty of 33%. b) Morphological features of wild *A. thuscula* populations, the arbustive growth (left), leaf morphology (center) and typical inflorescences (right) are shown. Images were taken from (Juanillo, 2020). c) Herbarium voucher specimen (Courtesy of *Herbario TFC, SEGAI-ULL*) and classical botanical description of *A. thuscula.*

A. thuscula grows as a woody branched shrub (up to one meter tall) with persistent foliage. Leaves produce a very strong and characteristic odor, and leaf morphology is highly variable, with flat lobes and gray-green color. It shows

grouped brownish-yellow terminal or subterminal inflorescences, of about four millimeters in diameter (Juanillo, 2020) (**Fig. 3b**). In fact, *A. thuscula* has been taxonomically characterized in base to botanical approaches, and perfectly preserved specimens can be found in various collections, such as the Herbario-TFC (SEGAI-ULL) (**Fig. 3c**).

Interestingly, *A. thuscula* has been extensively used in traditional medicine as diuretic tonic for treatment of digestive colic, flatulence, and for the expulsion of intestinal worms (Benjumea *et al.,* 2005). Moreover, it has also been applied topically as a pomade to reduce inflammation and articular pain (Pérez de Paz and Hernández Padrón, 1999). Other traditional uses are related with agriculture, in which the plant itself acts as a repellent, to protect stored potatoes from insects (Cruz, 2007).

Therefore, as in the case of *A. annua*, traditional medicine suggests potential therapeutic uses of *A. thuscula* and, in fact, the diuretic activity of *A. thuscula* has been evaluated and confirmed during last years (Hernández-Luis *et al.,* 2014). Currently, several industries traditionally focused in *A. annua* exploitation for Artemisinin production (**Fig. 2b**), are starting to study *A. thuscula* in the Canary Islands to this purpose, a fact that increases the economic value of this endemism, and also requires a more extensive study of this species at taxonomical and molecular levels.

Other recent studies related with *A. thuscula* involves the isolation of endophytic fungi, and the study of the relationship between the endophytic content with biodiversity (Cosoveanu, Rodriguez Sabina and Cabrera, 2018). However, to the best of our knowledge, no molecular phylogenetic studies have been focused on *A. thuscula*. Only one *A. thuscula* ITS sequence has been found

at the NCBI database, from a work where 133 different taxa were studied to compare different *Artemisia* arctic lineages and related species (Tkach *et al.,* 2008). Therefore, the complete DNA barcode of *A. thuscula* remains unsolved, since other DNA markers need to be sequenced to obtain a more robust phylogenetic characterization of this endemism.

Finally, as explained before, there are several micromorphological studies of *A. annua,* which include the use of the Scanning Electron Microscopy (SEM) to study the plant surface micromorphology in order to describe the architecture of the leaves, stem, or inflorescences (Wetzstein *et al.,* 2014), suggesting their use as taxonomic markers (Hayat *et al.,* 2009; Hussain, Hayat, *et al.,* 2019). However, as far as we know, these kinds of studies have not been carried out with *A. thuscula*. Therefore, SEM-based micromorphological comparative analysis of epidermal anatomical characteristics could be used to contrast the phylogenetic conclusions about *A. thuscula*.

Hypothesis and Objectives

Hypothesis

Artemisia thuscula represents an endemic plant species from Canary Islands with high industrial, biotechnological and biomedical potential, as has been demonstrated for *Artemisia annua*. Due to the particular biogeographical characteristics of *A. thuscula*, the increase of knowledge at molecular and micromorphological levels could reveal specific genetic and ethnobotanical features that will be exploited in the future for industrial purposes.

Objectives

Therefore, the main objectives of the present work are as follows:

1. Collect *A. thuscula* samples and sequence two different chloroplast DNA barcodes (*mat*K and *rbc*L).

2. Analyze of the differences between *A. annua* and A. *thuscula* at molecular level and study the phylogenetic relationships.

3. Explore the trichome micromorphology of *A. annua* and *A. thuscula* by Scanning Electron Microscopy.

Methods

Plant collection and species identification.

  A. thuscula and *A. annua* specimens were provided by *Biotech Tricopharming Research S.L.,* from a population maintained in a greenhouse for industrial Artemisinin production (**Fig. 2b**). Exploitation is located at the municipality of San Cristobal de La Laguna, in the northeast of Tenerife, Canary Islands, Spain (28°31'31.1"N 16°22'09.6"W). Samples were collected at the greenhouse (*Artennua*) and stored in sterile polyethylene sealing bags. Plants were in one-month adult phase and both leaves and stem were collected for analysis. Species identification was confirmed by trained personnel in base to morphological comparison with voucher specimens, deposited at Herbario-TFC (SEGAI-ULL), Tenerife, Spain (**Fig. 2a, 3c**). Once received at the laboratory, samples were stored at 4ºC until DNA extraction and SEM examination, for a maximum of 24 hours.

Genomic DNA extraction and quantification.

  For DNA extraction, 50 mg of fresh plant leaves were frozen with liquid nitrogen and homogenized inside a 1.5 ml tube, with the aim of

micropipette tip previously sealed over a flame. Genomic DNA was purified with the E.Z.N.A. Plant DNA Kit (Omega Biotek), following the manufacturer recommendations for fresh vegetal samples. Each sample was eluted in 100 µl of Elution Buffer, and DNA was preserved at −20°C for further analysis.

The presence of high molecular weight genomic DNA fragments was confirmed by agarose gel electrophoresis, using a 1% low-melting point agarose (VWR) gels prepared in 1X TAE buffer, as explained elsewhere (P. Y. Lee *et al.,* 2012). 1X GelRed (Biotium) was added to the agarose gel to visualize DNA, and 0.1 µg of 1 Kb DNA Ladder (Panreac) was also loaded as molecular weight reference. Electrophoresis conditions were 30 min at 80V (120 mA). A UV-transiluminator TFM20 (UVP) was used and image analysis was performed with the software Fiji (Schindelin *et al.,* 2012).

A fluorescence-based quantification was performed using Qubit 4 fluorimeter (ThermoFisher) and the dsDNA BR Assay Kit (ThermoFisher), following the manufacturer instructions. Each sample was measured by triplicate.

Amplification of *mat*K and *rbc*L chloroplast-coding regions by PCR.

Following the Consortium for the Barcode of Life (CBOL) recommendations (Hollingsworth *et al.,* 2009), two chloroplast coding genes were selected for PCR amplification and sequencing, as DNA Barcodes for *A. thuscula*. The Maturase-K (*mat*K) gene was amplified with primers matK-KIM3F (5′-CGTACAGTACTTTTGTGTTTACGAG-3′) and matK-KIM1R (5′-ACCCAGTCCATCTGGAAATCTTGGTTC-3′) (Fazekas *et al.,* 2012), while the Ribulose bifosfate carboxylase large subunit (*rbc*L) was amplified with primers rbcLa-F (5′-ATGTCACCACAAACAGAGACTAAAGC-3′) and

rbcLa-R (5′-GTAAAATCAAGTCCACCRCG-3′) (Fazekas *et al.,* 2012). Primers were purchased as desalted, from Metabion international AG.

PCR reactions were carried out with VWR Taq DNA polymerase Kit (VWR). Each PCR reaction contains 1.0 µl of each primer (10 pmol/µl), 0.15 µl of Taq DNA polymerase (5.0 U/µl), 2.5 µl of Taq Key Buffer (10X; 15 mM MgCl$_2$), 2.5 µl of deoxynucleotide Mix (2.0 mM each) and 10 µl of genomic DNA template (1.0 or 10.0 ng/µl, as indicated). Reaction volume was adjusted to 25 µl by adding 7.85 µl of H$_2$O, and negative controls with 10 µl of H$_2$O instead of DNA were included for each primer combination. Amplifications were performed in an iCycler Thermal Cycler (BioRad), with an initial denaturation step of 95°C (2 min), followed by 35 amplification cycles of 95°C (0.5 min), 55°C (0.5 min) and 72°C (1 min). A final extension step at 72°C (10 min) was also included. The presence of a single band from each amplification reaction, and the absence of contamination, were confirmed by agarose gel electrophoresis, as explained before.

Purification of PCR products and Sanger Sequencing.

Amplicons were purified using the EXOSAP-IT PCR Product Cleanup Kit (Affimetrix-USB), and sequenced with the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems), following the manufacturer instructions. Briefly, 5 µl of each PCR product was mixed with 1 µl of EXOSAP cleaning reagent, incubated at 37ºC (15 min) and then at 80ºC (5 min). Two sequencing reactions were prepared for each amplicon, by mixing 1.0 µl of the purified PCR product with 2.0 µl of BigDye Terminator v3.1 sequencing reagent, and 1.0 µl of one of the primers previously used for the amplification (1.6 pmol/µl). Sequencing reactions were then subjected to thermal cycles, precipitated by

ethanol-EDTA and resuspended in 10 µl of HiDi Formamide (ThermoFisher), exactly as described in the BigDye Terminator v3.1 kit manual. Capillary electrophoresis was performed with a 3500 Genetic Analyzer (Applied Biosystems), using POP7 polymer and a 50 cm capillary array, at the Servicio de Genómica (SEGAI-ULL), Tenerife, Canary Islands, Spain.

Electropherogram analysis and consensus sequence generation.

Electropherograms were manually curated using the software Chromas V2.6.5 (Technelysium Pty Ltd), by trimming both ends to maintain only those base calls with high quality. Curated sequences are listed in the **Supplementary data A and B**. When possible, forward and reverse sequences were aligned with the Clustal Omega server (Sievers *et al.,* 2011) to obtain a unique contiguous sequence (contig) either for *rbc*L or *mat*K. Primer sequences were then removed from each contig if present.

Sequence comparison between *A. thuscula* and *A. annua.*

Contig sequences obtained for *A. thuscula* and *A. annua* (*mat*K or *rbc*L, shown in **Supp. data A and B**, respectively) were aligned using the Clustal Omega server (Sievers *et al.,* 2011) to obtain the identity and similarity percentages, as shown in **Table 1**. Differences at nucleotide level, between the two species, were manually detected from the alignments. Sequences were then transferred to MEGA-X software (Sudhir Kumar *et al.,* 2018) and translated, to obtain the corresponding protein sequences.

Sequence retrieving from GenBank.

To obtain the *mat*K and *rbc*L-based phylogenies, available sequences with species-level definition corresponding to *Artemisia* genus (subtribe

Artemisiinae, tribe *Anthemideae*) were downloaded from the NCBI Taxonomy browser (Federhen, 2012). Searches were carried out with terms "txid4219[Organism:exp]", followed by the text "*mat*K" or "*rbc*L". In both cases, results were filtered to sequences from plant species, with length in the range of 10 – 3,000 bp, and those sequences without species-level definition were excluded. In addition, *mat*K and *rbc*L sequences from *Chrysanthemum indicum* and *Ajania fruticulosa* were retrieved from the NCBI Taxonomy browser. *Chrysanthemum* and *Ajania,* as well as *Artemisia*, represents different genera of the subtribe *Artemisiinae*, and were included in the alignments to obtain an overall vision of phylogenetic differences at the subtribe level.

Finally, two outgroup species were selected in base to sequence availability for *mat*K and *rbc*L and considering their taxonomy relationship with the *Artemisia* genus. Outgroup sequences correspond to *Anthemis arvensis* (subtribe *Anthemidinae*, tribe *Anthemideae*) and *Achillea millefolium* (subtribe *Matricariinae*, tribe *Anthemideae*). *Anthemis* and *Achillea* represents two *Artemisia*-related genera that belong to different subtribes but are included into the same tribe as the *Artemisia* genus. The complete list of retrieved sequences, indicating their species definition and corresponding NCBI accessions are included in **Supp. data C.**

Multiple sequence alignments.

As the first step, two independent alignments were obtained, either for *mat*K or *rbc*L. Initial sequence sets were constructed in MS-Excel, to include both the *A. thuscula* sequence obtained in the present work and the outgroups. Each sequence set was FASTA formatted and then transferred to MEGA-X software (Sudhir Kumar *et al.,* 2018). The *mat*K and *rbc*L preliminary

alignments were generated with the CLUSTALW algorithm (Thompson, Higgins, and Gibson 1994; Larkin *et al.,* 2007). Alignments were manually trimmed from both ends, excluding those regions that aligned outside the *A. thuscula* sequence. During this step, sequences which were too short were excluded, but the alignment length and the species number were maintained as higher as possible. Moreover, when more than two sequences were retrieved for the same species, only one representative was included at random, prioritizing sequences without indeterminations. The final sequence sets are listed in **Supp. data C**. At the end of this step, independent *mat*K and *rbc*L curated alignments were obtained with the MUSCLE algorithm (Edgar, 2004), followed by manual revision of gap positions when necessary.

In a second step, a concatenated dataset was generated for the two chloroplast coding sequences (*rbc*L-*mat*K). During this step, the independent *mat*K and *rbc*L alignments were exported in FASTA format from MEGA-X, and transferred to MS-Excel to exclude those species only present in either *mat*K or *rbc*L alignment. Therefore, only those species present in both *rbc*L and *mat*K datasets were maintained in the concatenated sequence matrix, which was then re-transferred to MEGA-X and also to MrBayes v3.1.2 (Ronquist *et al.,* 2012), to continue with the phylogenetic analysis. Sequences included in the concatenated alignment are listed in **Supp. data C.** Details of each alignment, as the alignment length, number of sequences or number of variable and informative sties can be found in **Table 2**.

Phylogenetic analysis

The independent alignments (*mat*K and *rbc*L) were analyzed by Maximum Parsimony (MP) and Maximum Likelihood (ML) methods, while the

concatenated sequence matrix (*mat*K-rbcL) was analyzed by ML, MP and Bayesian Inference (BI).

The MP trees were obtained with MEGA-X software (Sudhir Kumar *et al.,* 2018), using the Subtree-Pruning-Regrafting (SPR) algorithm with search level 1, in which the initial trees were obtained by random addition of sequences (10 replicates) (Nei and Kumar, 2000). The consistency index, the retention index, and the composite index were used to evaluate overall support for optimal trees, and the MP consensus tree was inferred from the most parsimonious trees. In the case of the ML-based phylogenies (Felsenstein, 1981), the best evolutionary substitution model was determined (independently for each alignment) on the basis of Bayesian Information Criterion (BIC) with MEGA-X (Sudhir Kumar *et al.,* 2018) (Table 2). The initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and the topology with superior log likelihood value was selected. In both cases (MP and ML phylogenies) statistical measures of clade support included the calculation of bootstrap values from 1000 replicates, and branches supported in less than 50% of the replicated trees were collapsed (Felsenstein, 1985).

Bayesian Inference (BI) method was also used to corroborate the topology of the MP and ML phylogenetic trees. In this case, JModelTest v2.1.5 software (Darriba *et al.,* 2012) was used to determine the nucleotide substitution models for each marker independently, under the Bayesian Information Criterion (BIC). Obtained models were JC69 (Jukes and Cantor, 1969) or F81+I (Felsenstein, 1981) for *rbc*L and *mat*K, respectively. Phylogenetic tree were constructed for each marker separately and for the concatenated sequences

using BI method with MrBayes v.3.2.1 software (Ronquist *et al.,* 2012), showing the same tree topology (not shown). Two parallel runs were applied with four Markov Monte Carlo Metropolis Coupled (MCMCMC) chains each, and with 107 of generations and a sampling frequency every 100 generations. Of the resulting 100,000 trees, the first 25,000 were discarded as "burned" and the next 75,000 were used to estimate the topology and parameters of the consensus tree. The percentage of times that nodes appeared in those 75,000 trees was interpreted as the posterior probability (PP) of each node. In the construction of the tree based on the concatenated sequences, the parameters obtained for each marker were estimated independently.

Phylogenetic *mat*K-*rbc*L tree was formatted and visualized with the iTOL server (Letunic and Bork 2019). Bootstrap (from ML and MP methods) and PP (from BI phylogeny) values were manually included for each branch, with the Adobe Acrobat-DC.

Analysis by Scanning Electron Microscopy.

To obtain high-resolution images of the leaf surface of both species (*A. annua* and *A. thuscula*), plant tissue samples were first dehydrated with ethanol solutions at increasing concentrations, for a period of 24 hours. Samples were completely dried by immersion in a saturated hexamethyldisilazane solution and then shaded with a 15 nm silver-coat, by the use of a QUORUM Q150R ES-PLUS instrument. Images were obtained in a ZEISS-EVO15 Scanning Electron Microscope, with a resolution of 2 nm and a microanalyzer of X-ray dispersive energies (EDX) Oxford X-MAX. Sample preparation and image obtention were carried out at Servicio de Microscopia Electrónica (SEGAI-ULL).

Results and discussion

PCR amplification of *rbc*L and *mat*K from *A. thuscula* and *A. annua*.

Genomic DNA was successfully extracted either from *A. annua* or *A. thuscula* leaves at the first attempt. The integrity of the genomic DNA was slightly higher in case of *A. thuscula*, since some degradation was detected in the case of *A. annua* (**Fig. 4a**). In case of *A. thuscula*, the total amount of genomic DNA recovered reached 5,78 µg, while this value was 3,88 µg in case of *A. annua,* which represents 32.9% less DNA recovery in the last sample. However, both samples yielded enough DNA concentration for PCR amplification (**Fig. 4b**). For each plant species, two amounts of template DNA were tested in each PCR, since the presence of PCR inhibitors could affect the amplification yield, especially when plant genomic DNA is used as template (Hollingsworth *et al.,* 2009).

Two chloroplast coding genes, Maturase-K (*matK*) and Ribulose bifosfate carboxylase large subunit (*rbc*L), were selected for PCR amplification and sequencing, since represent the classical chloroplast markers used for DNA barcoding of land plants recommended by the CBOL (Fazekas *et al.,* 2012; Hollingsworth *et al.,* 2009). In addition, *matK* based phylogeny has been recently reported about *Artemisia* species from Kazakhstan (Turuspekov *et al.,* 2018). Results showed that *matK* and *rbc*L regions were successfully PCR-amplified using both 1 or 10 ng of template DNA, without significant differences in the PCR product yield. In the case of *mat*K amplification, the expected DNA fragment of about 900 bp was successfully amplified with primers matK-KIM1R and matK-KIM3F, from both plant species (**Fig. 4c**). The same result was obtained for the amplification of *rbc*L, since only the expected

600 bp DNA fragment was amplified with primers rbcLa-F and rbcLa-R (**Fig. 4d**). Moreover, no amplification was detected when H_2O was included instead of template DNA, therefore confirming the absence of contamination during the PCR preparation step (**Fig. 4c and d**).
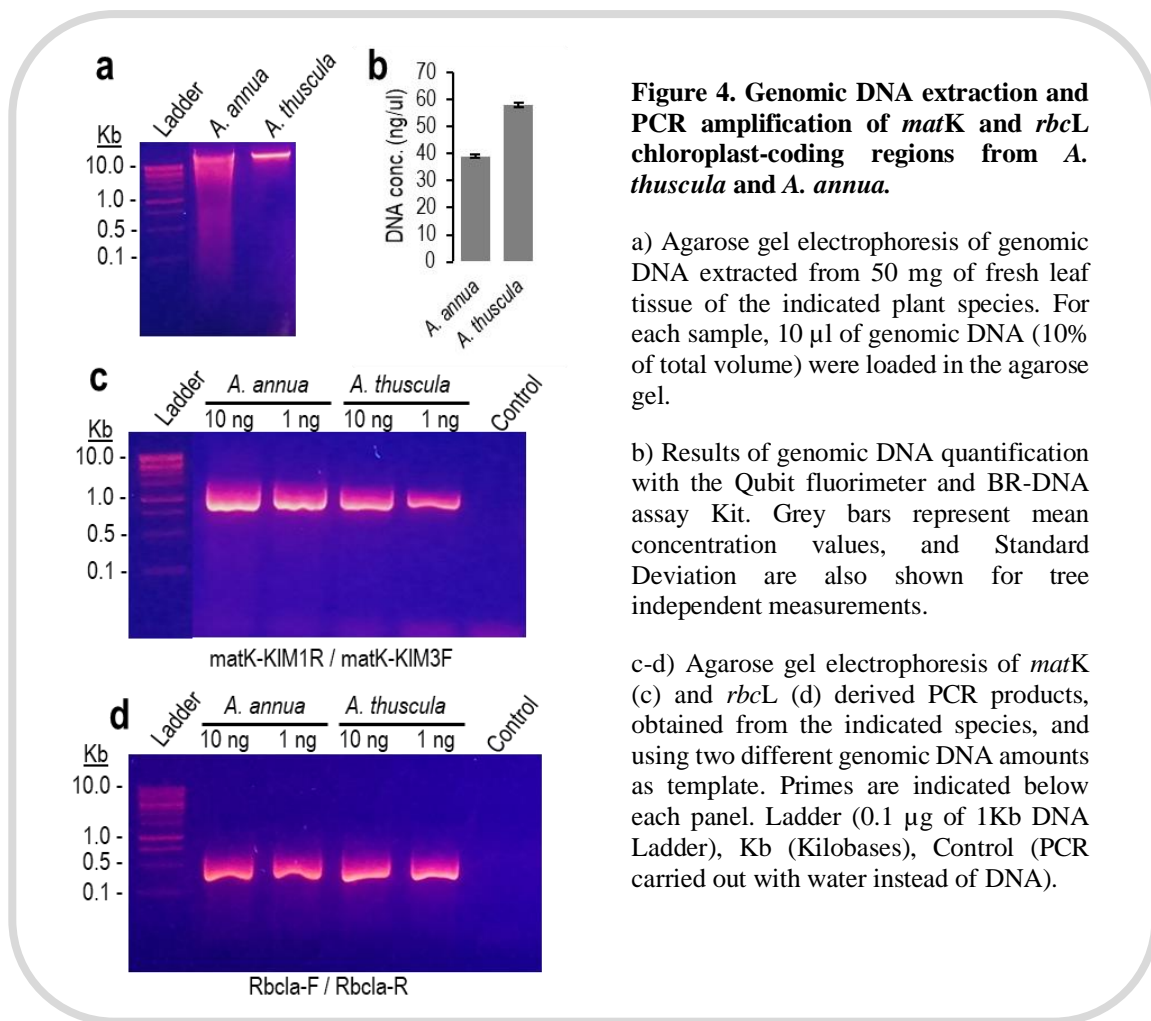
Figure 4. Genomic DNA extraction and PCR amplification of *mat*K and *rbc*L chloroplast-coding regions from *A. thuscula* and *A. annua.*

a) Agarose gel electrophoresis of genomic DNA extracted from 50 mg of fresh leaf tissue of the indicated plant species. For each sample, 10 µl of genomic DNA (10% of total volume) were loaded in the agarose gel.

b) Results of genomic DNA quantification with the Qubit fluorimeter and BR-DNA assay Kit. Grey bars represent mean concentration values, and Standard Deviation are also shown for tree independent measurements.

c-d) Agarose gel electrophoresis of *mat*K (c) and *rbc*L (d) derived PCR products, obtained from the indicated species, and using two different genomic DNA amounts as template. Primes are indicated below each panel. Ladder (0.1 µg of 1Kb DNA Ladder), Kb (Kilobases), Control (PCR carried out with water instead of DNA).

Sanger sequencing and sequence analysis.

The PCR products were purified sequenced from both ends, with the primers previously used for the PCR amplification (see methods). In the case of *rbc*L *locus*, forward and reverse sequences were obtained for each PCR product. Therefore, sequences were aligned to obtain a unique high-quality contig

sequence of 553 bp, either for *A. thuscula* or *A. annua* (**Supp. data A**). In the case of *mat*K, also two sequences were obtained for *A. thuscula* and, after their alignment, a unique high-quality contig sequence of 842 bp was generated (**Supp. data B**). However, only low-quality sequences were obtained after several sequencing attempts of *mat*K-derived PCR product of *A. annua*. The fail of *mat*K sequencing in the case of *A. annua* seems not to be related with PCR conditions, since amplicons showed only a specific band of the expected length, and the absence of nonspecific amplification products was confirmed in the agarose gels (**Fig. 4c**). Therefore, it could be related with low amounts of contaminants (probably polysaccharides and phenolic compounds) present in the genomic DNA preparation, which not affect the PCR amplification, but remains present in the PCR product and inhibits the sequencing reaction. In order to solve this problem, a nucleotide BLAST search was carried out at the NCBI, using the *A. thuscula mat*K sequence as query, and limiting the search to *A. annua*. Eight *mat*K sequences were retrieved and aligned, to obtain a unique consensus *mat*K sequence for *A. annua* of 842 bp (**Supp. data B**).

As far as we know, the present work represents the first report in which DNA barcodes from chloroplast-coding regions have been obtained for *A. thuscula*, and therefore supposes a clear increase in the molecular knowledge for this species. On the other hand, the generated *mat*K and *rbc*L sequences allows us to study its taxonomical characterization, which nowadays is only based in non-coding nuclear ITS sequences (Tkach *et al.,* 2008), and to compare both phylogenies in order to determinate if differences exist when both kind of genetic markers are used to reconstruct the evolutionary history of this species. Moreover, *mat*K and *rbc*L sequences of *A. thuscula* will be available for

scientific community, which allows the molecular identification of this species in base to these DNA barcodes in the future.

Comparison of *A. thuscula* and *A. annua* obtained DNA sequences.

As a first attempt to study if *A. thuscula* could show specific genetic features prpoduced by the evolutive adaptation to its restricted biogeographical distribution, the obtained *mat*K and *rbc*L sequences were compared with those obtained for *A. annua*. Results showed an identity percentage of 99.5% and 99.8% for the *rbc*L and *mat*K sequences, respectively (**Table 1**).

Table 1. Comparison of rbcL and matK DNA barcodes obtained for *A. annua* and *A. thuscula*. Sequences were aligned with Clustal Omega server. Identity and gap percentages were obtained from the alignments

DNA Barcode	Lenght (bp)	Identity	Gaps
*rbc*L	553	99.5%	0.0%
*mat*K	842	99.8%	0.0%

As expected, the conservation degree found between the two DNA Barcodes was extremely high, since these sequences represent chloroplast-coding regions that are usually conserved due to the restriction in nucleotide substitution necessary to keep intact the protein function. However, even the high degree of conservation found for *rbc*L and *mat*K, some Single Nucleotide Polymorphisms (SNPs) were detected.

In the case of *rbc*L, three variants were found in *A. thuscula* (**Supp. data A**). Two of them, c.139T>G and c.543T>C (named taking as reference the *A. annua* complete *rbc*L coding sequence), were found to be synonymous, since affect the third base of their respective codons but not change the protein sequence. Nevertheless, the third SNP, named c.271C>T, was found to be non-synonymous, since causes an amino acid change from proline in *A. annua* to

leucine in *A. thuscula* (p.Pro91Leu). On the other hand, when *mat*K sequences of both species were compared, also two synonymous SNPs were detected, c.975T>C and c.1134A>G (named with respect to *A. annua* complete *mat*K coding sequence) (**Supp. data B**).

The presence of these substitutions in *mat*K and *rbc*L sequences of *A. thuscula*, specially the non-synonymous mutation c.271C>T, detected in the highly conserved protein *rbc*L, allows us to speculate that *A. thuscula* could have developed unique genetic features as a cause of its evolutionary adaptation to its particular environment.

Alignments, estimation of evolutionary distances and analysis of *A. thuscula* variants.

Therefore, to better investigate this hypothesis and to afford the phylogenetic analysis of *A. thuscula*, sequence alignments were obtained to compare *A. thuscula* DNA barcodes with other *mat*K and *rbc*L sequences from as much as possible *Artemisia* species.

Two independent alignments were obtained, each based in a different chloroplast-coding region (*mat*K or *rbc*L). Therefore, 218 *mat*K and 275 *rbc*L nucleotide sequences were retrieved from the NCBI, including all sequences with species-level definition for the *Artemisia* genus (58 different *Artemisia* species, at the time of writing), and also *Chrysanthemum indicum* and *Ajania fruticulosa*, which represents two *Artemisia*-allied genera of subtribe *Artemisiinae* (**Table 2, Supp. data C**). As outgroups, *mat*K or *rbc*L sequences from *Anthemis arvensis* (subtribe *Anthemidinae*) and *Achillea millefolium* (subtribe *Matricariinae*) were included in their respective sets, as well as sequences from *A. thuscula* obtained in the present work. Thus, the initial *mat*K

and *rbc*L sequence sets included 221 and 278 sequences, respectively, comprising 61 ingroup and two outgroup species (**Table 2**).

After its manual refinement (see methods), the final *mat*K data matrix was reduced to 58 ingroup sequences (including *A. thuscula*), each from a different species. Therefore, 95% of species-level records retrieved from the NCBI were represented in the curated *mat*K alignment, which contains 654 sites and covers 77.7% of the *A. thuschula mat*K sequenced region (842 bp). If the two outgroups were excluded, 626 (95.7%) of the aligned positions remain conserved, 28 sites (4.3%) were variable and only 10 positions (1.5%) were parsimony informative sites (**Table 2**). In the case of *rbc*L, the final alignment comprises 54 sequences from their respective species (including *A. thuscula*), thus including 88.5% of all species retrieved from NCBI. The *rbc*L data matrix contains 510 sites (92.2% of the *A. thuscula rbc*L sequenced region), where only 6 positions (1.2%) were parsimony informative, 10 positions (2.0%) were variable, and the rest (500 sites, 98.0%) remains conserved (**Table 2**). Therefore, for both chloroplast-coding regions, low level of phylogenetic information was detected, and high degree of sequence conservation was found for the *Artemisia* genus. Indeed, when mean evolutionary distances were calculated with the best-fit model for each alignment, excluding the outgroups (**Table 2**), this value only reaches $3,2 \times 10^{-3}$ substitutions $\cdot$ site^{-1} in case of *mat*K, and $1,8 \times 10^{-3}$ substitutions $\cdot$ site^{-1} for *rbc*L. As expected, when outgroups were included in the estimations, mean distances slightly increase, reaching 4.1×10^{-3} and 2.6×10^{-3} substitutions $\cdot$ site^{-1} for *mat*K and *rbc*L, respectively. Taken together, these results show that the two chloroplast-coding regions analyzed in the present work (*mat*K and *rbc*L) remains highly conserved in the *Artemisia* genus, thus exhibiting a low level of evolutionary divergence.

The concatenated alignment (*rbc*L-*mat*K) was also obtained, in an attempt to increase the amount of phylogenetic information. Therefore, only those species present in both *mat*K and *rbc*L alignments were maintained in the concatenated sequence matrix, which comprises 43 ingroup species (including *A. thuscula*), and the two outgroups. Therefore, the *mat*K-*rbc*L alignment contains 70.5% of all species retrieved from NCBI (**Table 2, Supp. data C**), and comprises 1164 sites (83.4% of *mat*K and *rbc*L sequenced regions from *A. thuscula*). If the two outgroups were excluded, 1134 (97.4%) of the aligned positions remain conserved, 30 sites (2.5%) were variable, and 11 (0.9%) were parsimony informative sites (**Table 2**).

Table 2. Sequences and alignment features. For each alignment (matK, rbcL, and matK-rbcL), different characteristics are shown. Details, as NCBI sequence accesion numbers, can be found in Supp. data C. a) Number of sequences retrieved from NCBI (No. of outgroups between brackets). b) Number of different species represented in the NCBI, including *A. thuscula*, but excluding outgroups. c) Number of species contained in the final alignments, and percentage respect to (b). d) Length of sequence obtained for *A. thuscula* in the present work (bp; basepairs). e) Alignment length and percentage of coverage respect to the *A. thuscula* sequence. f, g and h) Number of conserved, variable and parsimony informative sites, respectively, and percentage respect to the alignment length. i) Best-fit evolutionary model calculated with MEGA X software, where G indicates Gamma discrete distribution. j) Bayesian Information Criterion (BIC) value obtained for the best-fit model. k and l) Average of evolutionary distances calculated with the best-fit model, excluding (k) or including (l) outgroups in calculations

	*mat*K	*rbc*L	*mat*K-*rbc*L
a) No. sequences retrieved from NCBI	218 (2)	275 (2)	N/A
b) No. different species in NCBI database	61	61	61
c) No. species aligned (%)	58 (95.0)	54 (88.5)	43 (70.5)
d) *A. thuscula* sequence length	842 bp	553 bp	1395 bp
e) Alignment length (%)	654 (77.7)	510 (92.2)	1164 (83.4)
f) Conserved sites (%)	626 (95.7)	500 (98.0)	1134 (97.4)
g) Variable sites (%)	28 (4.3)	10 (2.0)	30 (2.5)
h) Parsimony informative sites (%)	10 (1.5)	6 (1.2)	11 (0.9)
i) Best evolutionary model	Tamura 3 param. (G = 0.08)	Yukes and Cantor	Tamura 3 param. (G = 0.05)
j) BIC	$3,72 \times 10^{6}$	$2,84 \times 10^{6}$	$5,06 \times 10^{6}$
k) Mean evolutionary distance (excl. outgroups)	$3,20 \times 10^{-3}$	$1,77 \times 10^{-3}$	$2,58 \times 10^{-3}$
l) Mean evolutionary distance (incl. outgroups)	$4,11 \times 10^{-3}$	$2,60 \times 10^{-3}$	$3,72 \times 10^{-3}$

The mean evolutionary distance was also estimated with the best-fit model for the concatenated alignment (**Table 2**), reaching 2,5 x 10^{-3} substitutions · site^{-1} when outgroups were excluded, and increasing to 3,7 x 10^{-3} substitutions · site^{-1} including the outgroups. As expected, evolutionary distances were in the same range of those previously obtained independently for *mat*K and *rbc*L. However, BIC values showed 1.4 and 1.8-fold increase when compared with those values obtained independently for the *mat*K and *rbc*L, respectively, which indicate a better adjustment of the evolutionary models to the data.

Taken together, these results confirm the high degree of conservation for *rbc*L and *mat*K sequences in the *Artemisia* genus, suggesting that specific genetic variants could be only the product of evolutive adaptation to very specific conditions and, therefore, should be scarce. Interestingly, when the five SNPs found by the comparison of *A. annua* and *A. thuscula* sequences were localized in the alignments, two of them were found to be unique genetic variants found in *A. thuscula* (**Table 3**).

Table 3. Analysis of *A. thuscula* SNP variants prevalence between different Artemisia species. The five SNPs that were found by comparison of *A. thuscula* and *A. annua* DNA barcodes were studied, to obtain the percentage of Artemisia species that contain the *A. thuscula* variants. The alignment includes 42 Artemisia species and two outgroups, since *A. thuscula* was excluded from calculations.

SNP	*Locus*	Substitution type	No. of species with each variant				No. of species with *A. thuscula* variant (%)
			T	**G**	**C**	**A**	
c.139T>G	*rbc*L	Synonymous	44	0	0	0	0 (0.0%)
c.272C>T	*rbc*L	Non-synonymous	6	0	38	0	6 (13.3%)
c.543T>C	*rbc*L	Synonymous	1	0	43	0	43 (95,6%)
c.975T>C	*mat*K	Synonymous	43	1	0	0	0 (0.0%)
c.1134A>G	*mat*K	Synonymous	1	35	2	6	35 (77,8%)

One *A. thuscula* unique variant (c.139T>G) is localized in the *rbc*L coding sequence, and the other (c.975T>C) in the *mat*K gene. Interestingly, both unique variants correspond to synonymous mutations, which not alter the protein sequence, while the non-synonymous mutation (c.272C>T) found in the *rbc*L gene was present in about 13% of the analyzed *Artemisia* species. In spite of these results requires further studies to confirm the unique genetic features of *A. thuscula*, they increase the attractiveness of *A. thuscula* as a potential species to obtain novel bioactive compounds for industrial exploitation.

Phylogenetic analysis and taxonomical classification of *A. thuscula.*

As the first step to obtain the *A. thuscula* phylogeny, *mat*K-based cladograms were generated with ML and MP methods (**Supp. data D**). Some differences in tree topologies were found when both methods were compared, which could be related with the higher error rate assumed by the MP method (Lin and Nei, 1991). In addition, *mat*K cladograms were incompletely resolved, since only 48.3% of species were grouped with at least another, by both methods (bootstrap cut-off of 50%). Unfortunately, *A. thuscula* remains together with the group of unsolved species, thus limiting its taxonomical classification in base to this marker (**Supp. data D**). Different conclusions were obtained from the *rbc*L-based cladograms, since ML and MP methods generate identical tree topologies. In this case, in spite of cladograms were incompletely resolved, (44.6% of species were grouped with at least another, with 50% bootstrap), *A. thuscula* was included in a separate clade from the rest of *Artemisia* species, together with *A. argy*, *A. sibirica, A. igniaria* and *A. lactiflora* (**Supp. data D**). Therefore, *rbc*L-based results suggest *A. thuscula* could be included in the subgenera *Artemisia*, as well as the other four *Artemisia* species contained in this clade. However, the bootstrap value obtained for the ML phylogeny

remains slightly low (51%) and, therefore, these results should be considered with caution.

Since phylogenetic inference from *mat*K and *rbc*L markers, at least when treated independently, was not able to separate the majority of *A. species*, the phylogenetic tree was obtained from the concatenated alignment. In this case, ML and MP methods generate identical tree topologies, and phylograms showed a better degree of resolution, since, 34 out of 43 ingroup species (79%) were grouped resolved (**Fig. 5**). In addition, the Bayesian Inference (BI) method was also carried out to confirm results, and exactly the same tree topology was obtained (**Fig. 5**). The obtained phylogeny showed that all *Artemisia* species formed one large clade separately from outgroups, in which nine species were not resolved, while the others were grouped into eight subclades. The first subclade included *A. anethoides* and *A. sieversiana*, two species from subgenera *Absinthium*, as well as the second subclade, which included *A. frigida* and *A. rupestrisfrom*. Likewise, the third subclade includes *A. alaska*, *A. biennis* and *A. norvegica*, being the first within the subgenera *Absinthium* and the last two in the *Artemisia* subgenera. The fourth subclade contains two species from the subgenera *Artemisia* (*A. arctica* and *A. hyperborea*) and two species from subgenera *Tridentate* (*A. globularia* and *A. tridentata*) (**Fig. 5**).

Moreover, the rest of species that conforms the *Artemisia* subgenera are grouped in subclades five, six and seven. *A. roxburghiana*, *A. abrotanum*, *A. gmelinii* and *A. pontica* are grouped together (subclade five), as well as *A. annua, A. lucdovicina, A. michauxiana, A. tilsesii, A. sacrorum* and *A. suksdorfi* (subclade seven). Interestingly, *A. thuscula* was found to be grouped with *A. argyi*, *A. igniaria* and *A. lactiflora* (sixth subclade), being all of them species included in the *Artemisia* subgenera, as explained before. Finally, for the last

subclade, *A. dracunculus*, *A. pubescens*, *A. borealis*, *A. campestrus*, *A. japonica* and *A. scoporia* are included within the subgenera *Dracunculus* (subclade eight). Therefore, of the eight subclades obtained, two of them (first and second) contains only species which belong to the subgenera *Absinthium*, three subclades (fifth, sixth, and seventh) contains only species from the subgenera *Artemisia*, and only one subclade (eight) included all species from subgenera *Dracunculus*. Subclades three and four contains species from more than one subgenera, and *Anthemis arvensis* (subtribe *Anthemidinae*) and *Achillea millefolium* (subtribe *Matricariinae*) were grouped together in a separate clade, which indicates a higher evolutive divergence from the rest of *Artemisia* species from subtribe *Artemisiinae* (**Fig. 5**).

In view of the results obtained by the different phylogenetic methods, *A. thuscula* seems to be included in the *Artemisia* subgenera, and this classification is well supported by ML and MP bootstrap values (88% and 59%, respectively). Nevertheless, BI-based posterior probabilities remain slightly low (0.753), which is probably caused by the high conservation of *rbc*L and *mat*K detected for the *Artemisia* genus. Therefore, additional chloroplast non-coding markers should be analyzed to confirm these results. To the best of our knowledge, the present work represents the first time in which a phylogeny of *A. thuscula* has been obtained in base to *mat*K and *rbc*L. However, as far as we know, *A. thuscula* ITS1-2 regions has been only sequenced once, in a work focused in evolutionary pattern of arctic *Artemisia* species (Tkach *et al.,* 2008), in which *A. thuscula* was included in subgenera *Absinthium*. Therefore, nuclear and chloroplast phylogenetic histories seems to be different, at least in the case of *A. thuscula*. This incongruence could be related with the biogeographical

isolation of this endemism, since chloroplast phylogeny is restricted to maternal inheritance, while it is not the case of nuclear ITS sequences.
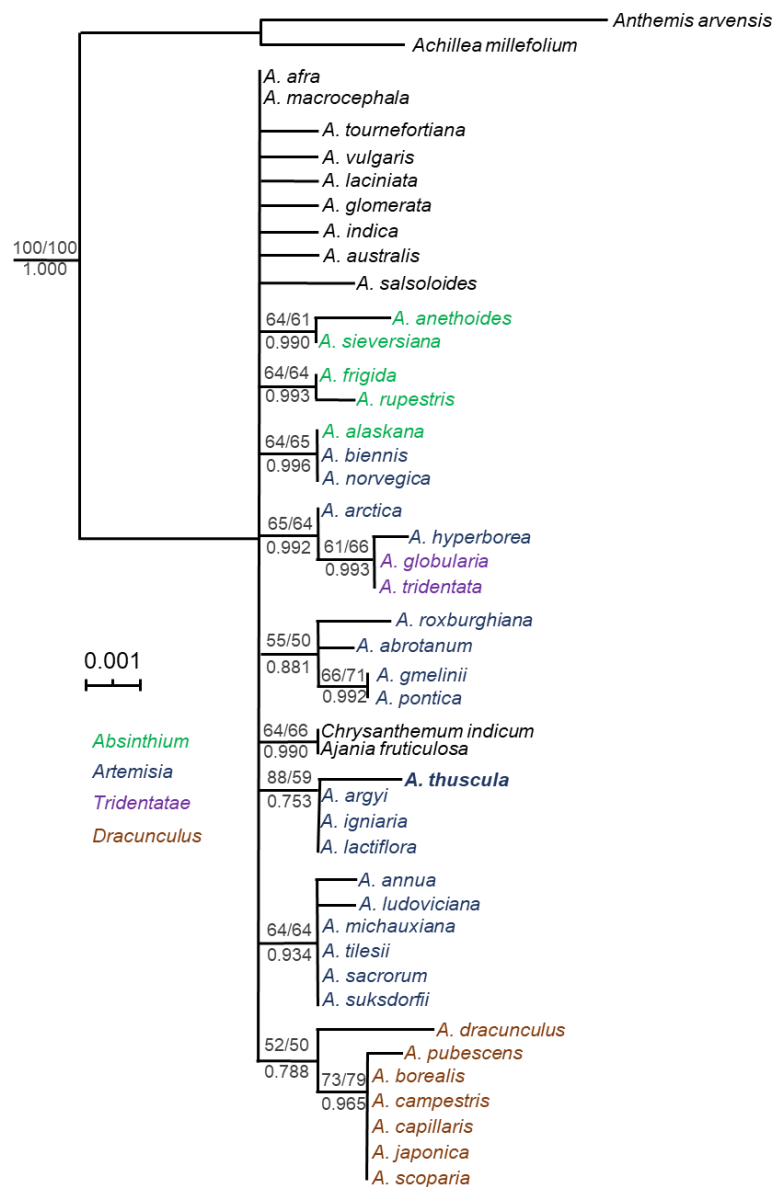
Figure 5. Evolutionary analysis in base to *mat*K-*rbc*L concatenated alignment. The evolutionary history was inferred by the ML, MP and BI methods. The three methods reproduce the same phylogeny. Therefore, only the ML-based tree with the highest log likelihood (-2038.30) is shown. ML method was performed with Tamura 3-parameter model ($G = 0.05$). Pairwise distances were estimated using the Maximum Composite Likelihood (MCL) approach. The MP consensus tree was inferred from 9 most parsimonious trees. The consistency index was 0.88, the retention index was 0.94, and the composite index was 0.89. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown above branches (ML/MP). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. Posterior probabilities from BI method are shown below branches.

Finally, recent studies suggest the use of *mat*K as an useful tool for taxonomical classification of the *Artemisia* genus (Turuspekov *et al.,* 2018). However, in the present work, we have obtained limited resolution using *mat*K, wich have been notably incleased with its combination with *rbc*L. In this sense, other studies have shown that the combination of both markers are suitable for the classification of several plant species (Jaén-Molina *et al.,* 2015), and this seems to be the case for the *Artemisia* genus. However, due to the high degree of consevation found in these two genes for the *Artemisia* genus, it is important to emphasize that it seems to be neccesary to include more chloroplast markers to confirm these results. Moreover, it could be nesessary to explore other taxonomical features, as SEM-based micromorphology (Hayat *et al.,* 2009; Hussain, Hayat, *et al.,* 2019) or bothanical taxonomic characteres (Ferri *et al.,* 2015).

Comparative micromorphology study by Scanning Electron Microscopy.

In order to identify taxonomic features that facilitate the identification and taxonomical characterization of *A. annua,* and specially of *A. thuscula*, micromorphology of the leaves and the stem of both plants was studied by Scanning Electron Microscopy (**Fig. 6 and 7**). Microscopic analysis of *A. annua* leaves allowed to distinguish the pluricellular trichomes, a typical character of the *Artemisia* subgenera (Hayat *et al.,* 2009; Hussain, Hayat, *et al.,* 2019), which were present mainly at leaf beam and undersides (**Fig. 6a-e**). Trichomes with scattered distribution were also observed in the stem of the analyzed samples of *A. annua* (**Fig. 6f-h**). With a higher resolution, it was even possible to study how the base of the pluricellular trichomes are attached to the leaf surface (**Fig. 6i and j**). In the case of *A. thuscula*, SEM images were also obtained from tail and leave surfaces (**Fig. 7**). As a clear difference with *A. annua,* it can be seen

that both leaves and the stem are completely covered by trichomes (**Fig. 7a-f**), and trichome mophology was found to be clearly different. Therefore, the number of thricomes represents a clear difference between both species, since *A. thuscula* leaves and stem are completely covered, while *A. annua* showed a more discreet distribution, depending on the surface studied. Finally, the trichome morphology seemed to be different, since *A. annua* showed pluricellular trichomes, these kind of structures were not present, or at least were not visible, in case of *A. thuscula*, which showed mainly long trichomes with a filiform morphology (**Fig. 7**).

Figure 6. SEM microphotographs obtained for *A. annua* leaves and stem surfaces. a-b) Images of the plant leaves beam. c-e) Images of underside plant leaves. f-h) Stem image. i-j) Pluricellular trichomes images.
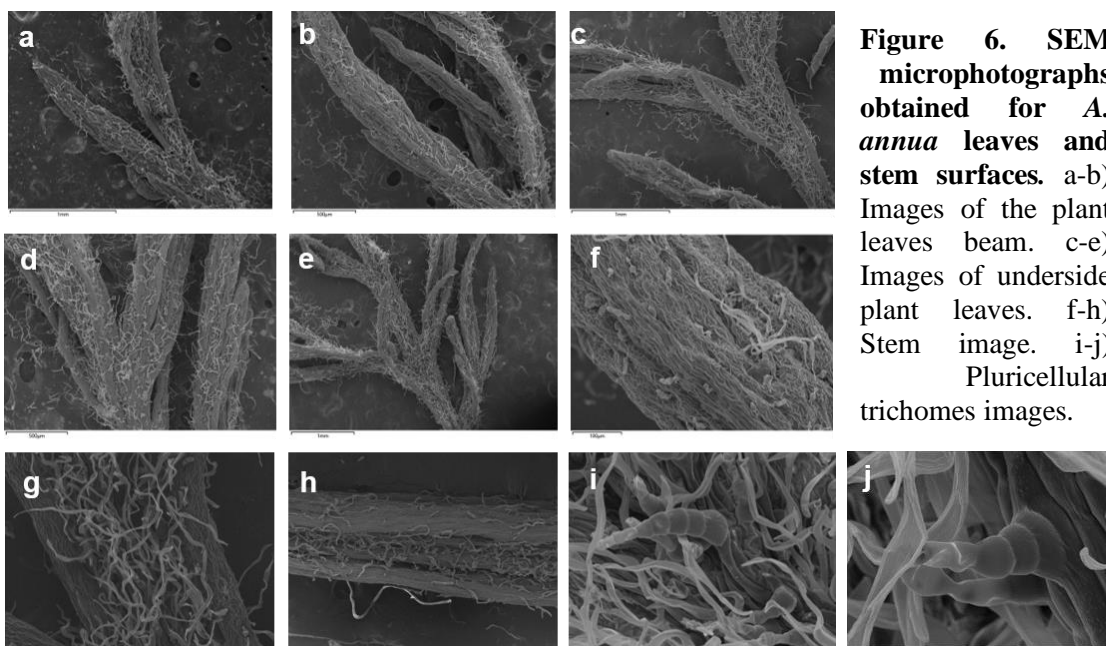
Unfortunately, the high density of trichomes present in *A. thuscula* caused that the morfological features of pluricellular trichomes, if present, were not visible and, therefore, no taxonomic information could be obtained from these experiments.
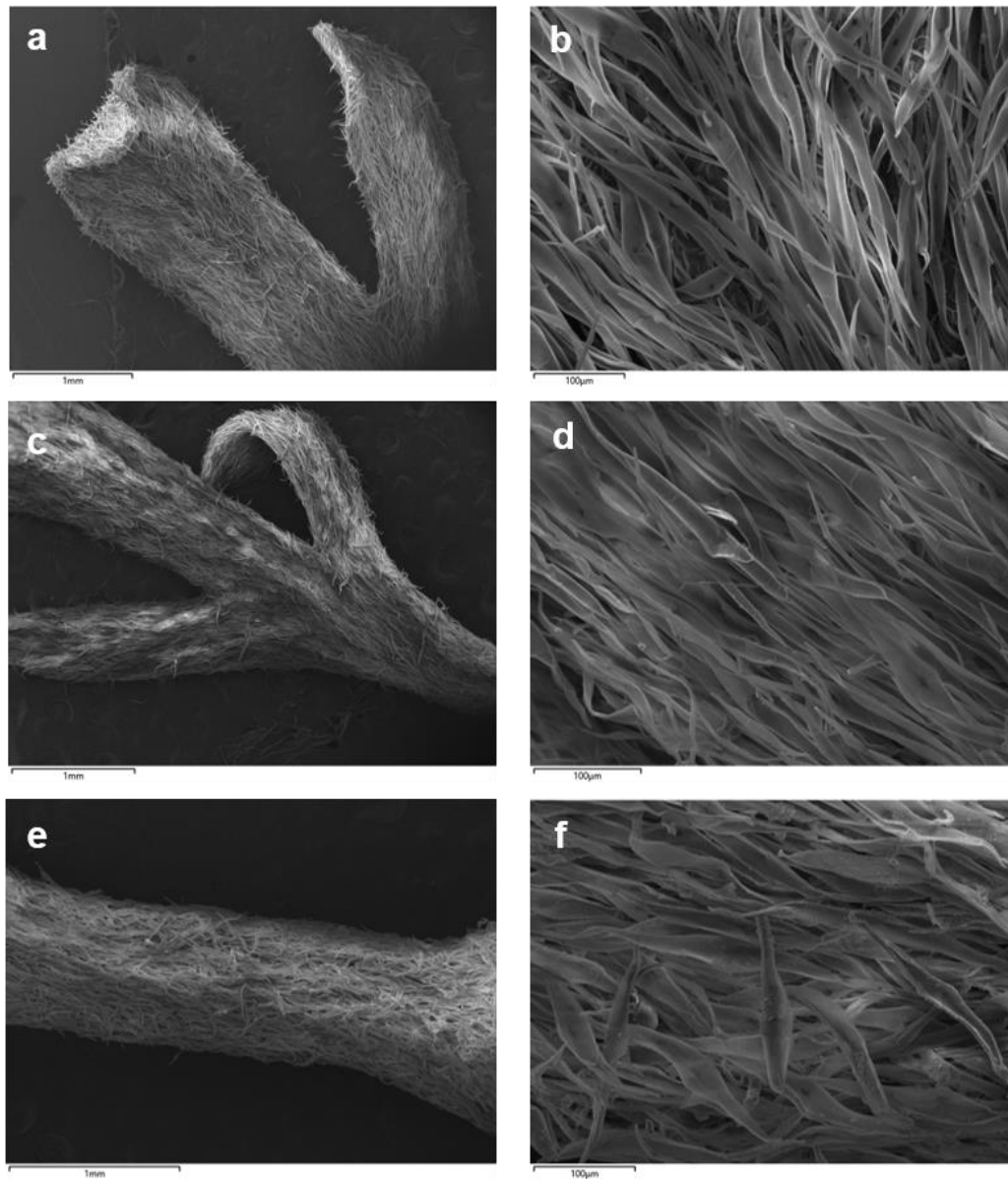
Figure 7. SEM microphotographs obtained for *A. thuscula* leaves and stem surfaces. a-b) Image of the plant leaves beam. c-d) images of leaf underside. e-f) Images of plant stem.

Conclusions

1. The genetic knowledge of *A. thuscula* has been increased by PCR amplification and sequencing of two chloroplast DNA barcodes (*mat*K and *rbc*L), for the first time.

2. The comparison of the two DNA barcodes obtained for *A. annua* and *A. thuscula* have shown five different SNPs, three of them (c.139T>G; c.272C>T and c.543T>C) present in the *rbc*L gene, and the other two (c.975T>C and c.1134A>G) in *mat*K gene coding region.

3. High conservation was detected for the *mat*K and *rbc*L genes in the *Artemisia* genus. However, the c.272C>T and c.975T>C variants detected in the *rbc*L and *mat*K genes, respectively, seems to be unique genetic features of *A. thuscula*, since were not detected in any other *Artemisia* species included in the present work.

4. Phylogenetic inference based in *mat*K-*rbc*L concatenated alignment suggest a taxonomical classification of *A. thuscula* and *A. annua* in the subgenera *Artemisia*, but the high degree of conservation detected requires the study of other chloroplast markers to confirm these results.

5. Morphological differences in the number of trichomes and morphology between *A. thuscula* and *A. annua* surfaces were detected by SEM, for the first time in the present study, but the obtained information was not enough to confirm the taxonomical classification of *A. thuscula*.

References

Acosta de la Luz, Lérida, and Ricardo Castro Armas. 2009. "Botánica, Biología, Composición Química y Propiedades Farmacológicas de *Artemisia Annua* L." *Revista Cubana de Plantas Medicinales* 14 (4): 0.

Bandelt, Hans-Jurgen, Peter Forster, and Arne Röhl. 1999. "Median-Joining Networks for Inferring Intraspecific Phylogenies." *Molecular Biology and Evolution* 16 (1): 37–48.

BDBC. 2020. Banco de Datos de Biodiversidad de Canarias (BDBC). Gobierno de Canarias. 2020. https://www.biodiversidadcanarias.es.

Benjumea, D, S Abdala, F Hernandez-Luis, P Pérez-Paz, and D Martin-Herrera. 2005. "Diuretic Activity of *Artemisia Thuscula*, an Endemic Canary Species." *Journal of Ethnopharmacology* 100 (1–2): 205–9.

Bryant, Laura, Brian Flatley, Chhaya Patole, Geoffrey D Brown, and Rainer Cramer. 2015. "Proteomic Analysis of *Artemisia Annua*–towards Elucidating the Biosynthetic Pathways of the Antimalarial pro-Drug Artemisinin." *BMC Plant Biology* 15 (1): 175.

Ćavar, Sanja, Milka Maksimović, Danijela Vidic, and Adisa Parić. 2012. "Chemical Composition and Antioxidant and Antimicrobial Activity of Essential Oil of *Artemisia Annua* L. from Bosnia." *Industrial Crops and Products* 37 (1): 479–85.

Chorlton, Sam D. 2017. "Toxoplasma Gondii and Schizophrenia: A Review of Published RCTs." *Parasitology Research* 116 (7): 1793–99.

Cosoveanu, Andreea, Samuel Rodriguez Sabina, and Raimundo Cabrera. 2018. "Fungi as Endophytes in *Artemisia Thuscula*: Juxtaposed Elements of Diversity and Phylogeny." *Journal of Fungi (Basel, Switzerland)* 4 (1).

Croteau, R. 1986. "Biochemistry of Monoterpenes and Sesquiterpenes of Essential Oils." *Herbs, Spices, and Medicinal Plants: Recent Advances in Botany, Horticulture, and Pharmacology*.

Cruz, J. 2007. "Más de 100 Plantas Medicinales." *Medicina Popular Canaria, Monografía. Las Palmas de Gran Canaria*, 19–35.

Darriba, Diego, Guillermo L Taboada, Ramón Doallo, and David Posada. 2012. "JModelTest 2: More Models, New Heuristics and Parallel Computing." *Nature Methods* 9 (8): 772.

Drouot, Emilien, Jocelyne Piret, and Guy Boivin. 2016. "Short Communication Artesunate Demonstrates in Vitro Synergism with Several Antiviral Agents against Human Cytomegalovirus." *Antiviral Therapy* 21: 535–39.

Duffy, Patrick E., and Theonest K. Mutabingwa. 2006. "Artemisinin Combination Therapies." *Lancet*. Elsevier Limited. https://doi.org/10.1016/S0140-6736(06)68900-9.

Edgar, Robert C. 2004. "MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput." *Nucleic Acids Research* 32 (5): 1792–97.

Efferth, Thomas. 2017. "From Ancient Herb to Modern Drug: *Artemisia Annua* and Artemisinin for Cancer Therapy." In *Seminars in Cancer Biology*, 46:65–83. Elsevier.

Fatemi, Ahmadreza, Keramat Asasi, and Seyed Mostafa Razavi. 2017. "Anticoccidial Effects of *Artemisia Annua* Ethanolic Extract: Prevention, Simultaneous Challenge-Medication, and Treatment." *Parasitology Research* 116 (9): 2581–89.

Fazekas, Aron J, Maria L Kuzmina, Steven G Newmaster, and Peter M Hollingsworth. 2012. "DNA Barcoding Methods for Land Plants." In *DNA Barcodes*, 223–52. Springer.

Federhen, Scott. 2012. "The NCBI Taxonomy Database." *Nucleic Acids Research* 40 (D1): D136–43.

Felsenstein, Joseph. 1981. "Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach." *Journal of Molecular Evolution* 17 (6): 368–76.

Felsenstein, Joseph. 1985. "Confidence Limits on Phylogenies: An Approach Using the Bootstrap." *Evolution* 39 (4): 783–91.

Ferri, Gianmarco, Beatrice Corradini, Francesca Ferrari, Anna Laura Santunione, Federica Palazzoli, and M Alu. 2015. "Forensic Botany II, DNA Barcode for Land Plants: Which Markers after the International Agreement?" *Forensic Science International: Genetics* 15: 131–36.

Garcia, Sònia, E Durant McArthur, Jaume Pellicer, Stewart C Sanderson, Joan Vallès, and Teresa Garnatje. 2011. "A Molecular Phylogenetic Approach to Western North America Endemic *Artemisia* and Allies (Asteraceae): Untangling the Sagebrushes." *American Journal of Botany* 98 (4): 638–53.

Gold, Daniel, Mohammed Alian, Avraham Domb, Yara Karawani, Maysa Jbarien, Jacques Chollet, Richard K Haynes, Ho Ning Wong, Viola Buchholz, and Andreas Greiner. 2017. "Elimination of Schistosoma Mansoni in Infected Mice by Slow Release of Artemisone." *International Journal for Parasitology: Drugs and Drug Resistance* 7 (2): 241–47.

Haghighi, Ahmad Razban, Ali Osman Belduz, Mohammad Moghaddam Vahed, Kamil Coskuncelebi, and Salih Terzioglu. 2014. "Phylogenetic Relationships among Artemisia Species Based on Nuclear ITS and Chloroplast PsbA-TrnH DNA Markers." *Biologia* 69 (7): 834–39.

Hayat, Muhammad Qasim, Muhammad Ashraf, Mir Ajab Khan, Tariq Mahmood, Mushtaq Ahmad, and Shazia Jabeen. 2009. "Phylogeny of Artemisia L.: Recent Developments." *African Journal of Biotechnology* 8 (11): 2423–28.

Hernández-Luis, Felipe, Susana Abdala, Dora Benjumea, Sandra Dévora, and Domingo Martín-Herrera. 2014. "Application of Diuretic Power Index to Evaluate the Diuretic Activity of Medicinal Plants or Drugs." *Phytochemistry* 12.

Hobbs, Christopher R, and Bruce G Baldwin. 2013. "Asian Origin and Upslope Migration of Hawaiian Artemisia (Compositae–Anthemideae)." *Journal of Biogeography* 40 (3): 442–54.

Hollingsworth, Peter M, Laura L Forrest, John L Spouge, Mehrdad Hajibabaei, Sujeevan Ratnasingham, Michelle van der Bank, Mark W Chase, Robyn S Cowan, and David L Erickson. 2009. "A DNA Barcode for Land Plants." *Proceedings of the National Academy of Sciences* 106 (31): 12794–97.

Hussain, Adil, Muhammad Q Hayat, Sumaira Sahreen, Qurrat Ul Ain, and Syed A I Bokhari. 2017. "Pharmacological Promises of Genus Artemisia (Asteraceae): A Review." *Proceedings of the Pakistan Academy of Sciences: B. Life and Environmental Sciences* 54 (4): 265–87.

Hussain, Adil, Muhammad Qasim Hayat, Sumaira Sahreen, and Syed Ali Imran Bokhari. 2019. "C(Asteraceae) from Northeast (Gilgit-Baltistan), Pakistan." *International Journal of Agriculture and Biology* 21 (3): 630–38.

Hussain, Adil, Daniel Potter, Sangtae Kim, Muhammad Q Hayat, and Syed A I Bokhari. 2019. "Molecular Phylogeny of Artemisia (Asteraceae-Anthemideae) with Emphasis on Undescribed Taxa from Gilgit-Baltistan (Pakistan) Based on NrDNA (ITS and ETS) and CpDNA (PsbA-TrnH) Sequences." *Plant Ecology and Evolution* 152 (3): 507–20.

Im, Eunji, Changhwan Yeo, Hyo-Jeong Lee, and Eun-Ok Lee. 2018. "Dihydroartemisinin Induced Caspase-Dependent Apoptosis through Inhibiting the Specificity Protein 1 Pathway in Hepatocellular Carcinoma SK-Hep-1 Cells." *Life Sciences* 192: 286–92.

Iram, Shabina, Muhammad Qasim Hayat, Muhammad Tahir, Alvina Gul, and Ibrar Ahmed. 2019. "Chloroplast Genome Sequence of Artemisia Scoparia: Comparative Analyses and Screening of Mutational Hotspots." *Plants* 8 (11): 476.

Jaén-Molina, Ruth, Águedo Marrero-Rodríguez, J Alfredo Reyes-Betancort, Arnoldo Santos-Guerra, José Naranjo-Suárez, and Juli Caujapé-Castells. 2015. "Molecular Taxonomic Identification in the Absence of a 'Barcoding Gap': A Test with the Endemic Flora of the C Anarian Oceanic Hotspot." *Molecular Ecology Resources* 15 (1): 42–56.

Juanillo. 2020. "Atras Rural de Gran Canaria." Recursos Turísticos, Culturales, Naturales y Socio-Económicos Del Medio Rural de Gran Canaria. 2020.

Jukes, Thomas H, and Charles R Cantor. 1969. "Evolution of Protein Molecules." *Mammalian Protein Metabolism* 3 (21): 132.

Kang, Sang-Ho, Kyunghee Kim, Jeong-Hoon Lee, Byoung Ohg Ahn, So Youn Won, Seong-Han Sohn, and Jung Sun Kim. 2016. "The Complete Chloroplast Genome Sequence of Medicinal Plant, Artemisia Argyi." *Mitochondrial DNA Part B* 1 (1): 257–58.

Kornkven, Amy B, Linda E Watson, and James R Estes. 1998. "Phylogenetic Analysis of Artemisia Section Tridentatae (Asteraceae) Based on Sequences from the Internal Transcribed Spacers (ITS) of Nuclear Ribosomal DNA." *American Journal of Botany* 85 (12): 1787–95.

Kornkven, Amy B, Linda E Watson, and James R Estes. 1999. "Molecular Phylogeny of Artemisia Section Tridentatae (Asteraceae) Based on Chloroplast DNA Restriction Site Variation." *Systematic Botany*, 69–84.

Kumar, Satish, Lingaraja Jena, Maheswata Sahoo, Mrunmayi Kakde, Sangeeta Daf, and Ashok K Varma. 2015. "In Silico Docking to Explicate Interface between Plant-Originated Inhibitors and E6 Oncogenic Protein of Highly Threatening Human Papillomavirus 18." *Genomics & Informatics* 13 (2): 60.

Kumar, Sudhir, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura. 2018. "MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms." *Molecular Biology and Evolution* 35 (6): 1547–49.

Larkin, Mark A, Gordon Blackshields, Nigel P Brown, R Chenna, Paul A McGettigan, Hamish McWilliam, Franck Valentin, Iain M Wallace, Andreas Wilm, and Rodrigo Lopez. 2007. "Clustal W and Clustal X Version 2.0." *Bioinformatics* 23 (21): 2947–48.

Lee, Pei Yun, John Costumbrado, Chih-Yuan Hsu, and Yong Hoon Kim. 2012. "Agarose Gel Electrophoresis for the Separation of DNA Fragments." *JoVE (Journal of Visualized Experiments)*, no. 62: e3923.

Lee, Yun Sun, Jee Young Park, Jin-Kyung Kim, Hyun Oh Lee, Hyun-Seung Park, Sang-Choon Lee, Jung Hwa Kang, Taek Joo Lee, Sang Hyun Sung, and Tae-Jin Yang. 2016. "Complete Chloroplast Genome Sequence of Artemisia Fukudo Makino (Asteraceae)." *Mitochondrial DNA Part B* 1 (1): 376–77.

Letunic, Ivica, and Peer Bork. 2019. "Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments." *Nucleic Acids Research* 47 (W1): W256–59.

Li, Peng, and Guolun Jia. 2019. "Characterization of the Complete Chloroplast Genome of Artemisia Capillaris (Campanulales: Asteraceae Bercht), a Unique Economic Plant to China." *Mitochondrial DNA Part B* 4 (1): 1808–9.

Li, Xinyang, Shiyan Gu, Donglei Sun, Huangmei Dai, Hongyu Chen, and Zunzhen Zhang. 2018. "The Selectivity of Artemisinin-Based Drugs on Human Lung Normal and Cancer Cells." *Environmental Toxicology and Pharmacology* 57: 86–94.

Liang, Ning, Yanchun Zhong, Jie Zhou, Bihao Liu, Ruirui Lu, Yezhi Guan, Qi Wang, Chunlin Liang, Yu He, and Yuan Zhou. 2018. "Immunosuppressive Effects of Hydroxychloroquine and Artemisinin Combination Therapy via the Nuclear Factor-κB Signaling Pathway in Lupus Nephritis Mice." *Experimental and Therapeutic Medicine* 15 (3): 2436–42.

Lim, Chae Eun, Goon-Bo Kim, Se-A Ryu, Hee-Ju Yu, and Jeong-Hwan Mun. 2018. "The Complete Chloroplast Genome of Artemisia Hallaisanensis Nakai (Asteraceae), an Endemic Medicinal Herb in Korea." *Mitochondrial DNA Part B* 3 (1): 359–60.

Lin, J, and M Nei. 1991. "Relative Efficiencies of the Maximum-Parsimony and Distance-Matrix Methods of Phylogeny Construction for Restriction Data." *Molecular Biology and Evolution* 8 (3): 356–65.

Lu, Ke, Wei Mao, Zhongyu Du, Yiming He, Cheng Fan, Kai Zhang, Lijun Wang, Guoming Liu, and Yizhong Duan. 2020. "The Complete Chloroplast Genome Sequence of Artemisia Ordosica." *Mitochondrial DNA Part B* 5 (2): 1663–64.

Malik, Sadia, Daniel Vitales, Muhammad Qasim Hayat, Aleksandr A Korobkov, Teresa Garnatje, and Joan Vallès. 2017. "Phylogeny and Biogeography of Artemisia Subgenus Seriphidium (Asteraceae: Anthemideae)." *Taxon* 66 (4): 934–52.

Mannan, Abdul, Ibrar Ahmed, Waheed Arshad, Muhammad F Asim, Rizwana A Qureshi, Izhar Hussain, Bushra Mirza, *et al.,* 2010. "Assessment of Artemisinin Contents in Selected Artemisia Species from Tajikistan (Central Asia)." *Malaria Journal* 9 (1): 310. https://doi.org/10.3390/medicines6010023.

Meng, Dong, Zhou Xiaomei, Ku Wenzhen, and Zhenggang Xu. 2019. "Detecting Useful Genetic Markers and Reconstructing the Phylogeny of an Important Medicinal Resource Plant, Artemisia Selengensis, Based on Chloroplast Genomics." *PloS One* 14 (2).

Min, Juhyeon, Jongsun Park, Yongsung Kim, and Woochan Kwon. 2019. "The Complete Chloroplast Genome of Artemisia Fukudo Makino (Asteraceae): Providing Insight of Intraspecies Variations." *Mitochondrial DNA Part B* 4 (1): 1510–12.

Nangong, Zhenying, Xiangxiang He, and Fangliang Huang. 2020. "The Complete Chloroplast Genome Sequence of Medicinal Plant, Artemisia Lavandulaefolia YC." *Mitochondrial DNA Part B* 5 (2): 1194–95.

Nei, Masatoshi, and Sudhir Kumar. 2000. *Molecular Evolution and Phylogenetics*. Oxford university press.

Pellicer, Jaume, Sònia Garcia, M A Canela, Teresa Garnatje, Aleksandr A Korobkov, J D Twibell, and Joan Vallès. 2010. "Genome Size Dynamics in Artemisia L.(Asteraceae): Following the Track of Polyploidy." *Plant Biology* 12 (5): 820–30.

Pellicer, Jaume, C Haris Saslis-Lagoudakis, Esperança Carrió, Madeleine Ernst, Teresa Garnatje, Olwen M Grace, Airy Gras, Màrius Mumbrú, Joan Vallès, and Daniel Vitales. 2018. "A Phylogenetic Road Map to Antimalarial Artemisia Species." *Journal of Ethnopharmacology* 225: 1–9.

Peng, Jiao, Yunlin Zhao, Chaoyang Li, and Zhenggang Xu. 2018. "The Complete Chloroplast Genome and Phylogeny of Artemisia Selengensis in Dongting Lake." *Mitochondrial DNA Part B* 3 (2): 907–8.

Pérez de Paz, Pedro L, and Consuelo E Hernández Padrón. 1999. "Plantas Medicinales o Útiles En La Flora Canaria."

Puri, B K, J S Hakkarainen-Smith, and J A Monro. 2017. "The Effect of Artesunate on Short-Term Memory in Lyme Borreliosis." *Medical Hypotheses* 105: 4–5.

Raffetin, A, Fabrice Bruneel, Camille Roussel, M Thellier, P Buffet, E Caumes, and S Jauréguiberry. 2018. "Use of Artesunate in Non-Malarial Indications." *Medecine et Maladies Infectieuses* 48 (4): 238–49.

Rich, T C, S A Trinder, and C Long. 2015. "DNA Barcoding for Plants." *Methods in Molecular Biology (Clifton, NJ)* 1245: 101–18.

Riggins, Chance W, and David S Seigler. 2012. "The Genus Artemisia (Asteraceae: Anthemideae) at a Continental Crossroads: Molecular Insights into Migrations, Disjunctions, and Reticulations among Old and New World Species from a Beringian Perspective." *Molecular Phylogenetics and Evolution* 64 (3): 471–90.

Ronquist, Fredrik, Maxim Teslenko, Paul Van Der Mark, Daniel L Ayres, Aaron Darling, Sebastian Höhna, Bret Larget, Liang Liu, Marc A Suchard, and John P Huelsenbeck. 2012. "MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space." *Systematic Biology* 61 (3): 539–42.

Sainz, Paula, Ángel Cruz-Estrada, Carmen Elisa Díaz, and Azucena González-Coloma. 2017. "The Genus Artemisia: Distribution and Phytochemistry in the Iberian Peninsula and the Canary and Balearic Islands." *Phytochemistry Reviews* 16 (5): 1023–43.

Sanz, María, Roser Vilatersana, Oriane Hidalgo, Núria Garcia-Jacas, Alfonso Susanna, Gerald M Schneeweiss, and Joan Vallès. 2008. "Molecular Phylogeny and Evolution of Floral Characters of Artemisia and Allies (Anthemideae, Asteraceae): Evidence from NrDNA ETS and ITS Sequences." *Taxon* 57 (1): 66–78.

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, and Benjamin Schmid. 2012. "Fiji: An Open-Source Platform for Biological-Image Analysis." *Nature Methods* 9 (7): 676–82.

Schwarz, Gideon. 1978. "Estimating the Dimension of a Model." *The Annals of Statistics* 6 (2): 461–64.

Shahzadi, Iram, Furrukh Mehmood, Zain Ali, Ibrar Ahmed, and Bushra Mirza. 2020. "Chloroplast Genome Sequences of Artemisia Maritima and Artemisia Absinthium: Comparative Analyses, Mutational Hotspots in Genus Artemisia and Phylogeny in Family Asteraceae." *Genomics* 112 (2): 1454–63.

Shen, Qian, Lida Zhang, Zhihua Liao, Shengyue Wang, Tingxiang Yan, Pu Shi, Meng Liu, Xueqing Fu, Qifang Pan, and Yuliang Wang. 2018. "The Genome of Artemisia Annua Provides Insight into the Evolution of Asteraceae Family and Artemisinin Biosynthesis." *Molecular Plant* 11 (6): 776–88.

Shen, Xiaofeng, Mingli Wu, Baosheng Liao, Zhixiang Liu, Rui Bai, Shuiming Xiao, Xiwen Li, Boli Zhang, Jiang Xu, and Shilin Chen. 2017. "Complete Chloroplast Genome Sequence and Phylogenetic Analysis of the Medicinal Plant Artemisia Annua." *Molecules* 22 (8): 1330.

Sievers, Fabian, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, and Johannes Söding. 2011. "Fast, Scalable Generation of High-quality Protein Multiple Sequence Alignments Using Clustal Omega." *Molecular Systems Biology* 7 (1).

Stamatakis, Alexandros. 2006. "RAxML-VI-HPC: Maximum Likelihood-Based Phylogenetic Analyses with Thousands of Taxa and Mixed Models." *Bioinformatics* 22 (21): 2688–90.

Tang, Weici, and Gerhard Eisenbrand. 2013. *Chinese Drugs of Plant Origin: Chemistry, Pharmacology, and Use in Traditional and Modern Medicine.* Springer Science & Business Media.

Thompson, Julie D, Desmond G Higgins, and Toby J Gibson. 1994. "CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice." *Nucleic Acids Research* 22 (22): 4673–80.

Tkach, Natalia V, Matthias Heinrich Hoffmann, Martin Roser, Alexander Alexandrovich Korobkov, and Klaus Bernhard von Hagen. 2008. "A." *Evolution; International Journal of Organic Evolution* 62 (1): 184–98. https://doi.org/10.1111/j.1558-5646.2007.00270.x.

Torrell, Montserrat, Núria Garcia-Jacas, Alfonso Susanna, and Joan Vallès. 1999. "Phylogeny in Artemisia (Asteraceae, Anthemideae) Inferred from Nuclear Ribosomal DNA (ITS) Sequences." *Taxon* 48 (4): 721–36.

Turuspekov, Yerlan, Yuliya Genievskaya, Aida Baibulatova, Alibek Zatybekov, Yuri Kotuhov, Margarita Ishmuratova, Akzhunis Imanbayeva, and Saule Abugalieva. 2018. "Phylogenetic Taxonomy of Artemisia L. Species from Kazakhstan Based on MatK Analyses." In *Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences.*, 72:29–37. Sciendo.

Vallès, Joan, Sònia Garcia, Oriane Hidalgo, Joan Martín, Jaume Pellicer, María Sanz, and Teresa Garnatje. 2011. "Biology, Genome Evolution, Biotechnological Issues and Research Including Applied Perspectives in Artemisia (Asteraceae)." In *Advances in Botanical Research*, 60:349–419. Elsevier.

Vallès, Joan, Montserrat Torrell, Teresa Garnatje, Núria Garcia-Jacas, Royal Vilatersana, and Alfonso Susanna. 2003. "The Genus Artemisia and Its Allies: Phylogeny of the Subtribe Artemisiinae (Asteraceae, Anthemideae) Based on Nucleotide Sequences of Nuclear Ribosomal DNA Internal Transcribed Spacers (ITS)." *Plant Biology* 5 (03): 274–84.

Watson, Linda E, Paul L Bates, Timothy M Evans, Matthew M Unwin, and James R Estes. 2002. "Molecular Phylogeny of Subtribe Artemisiinae (Asteraceae), Including Artemisia and Its Allied and Segregate Genera." *BMC Evolutionary Biology* 2 (1): 17.

Wetzstein, Hazel Y, Justin A Porter, Jules Janick, and Jorge F S Ferreira. 2014. "Flower Morphology and Floral Sequence in Artemisia Annua (Asteraceae)." *American Journal of Botany* 101 (5): 875–85.

Willcox, Merlin, Gerard Bodeker, Geneviève Bourdy, Vikas Dhingra, Jacques Falquet, Jorge F S Ferreira, Bertrand Graz, Hans-Martin Hirt, Elisabeth Hsu, and Pedro Melillo de Magalhães. 2004. "Artemisia Annua as a Traditional Herbal Antimalarial." *Traditional Medicinal Plants and Malaria* 4: 43–59.

Supplementary data. Bioinformatics analysis of *rbc*L and *mat*K sequences from *A. thuscula* and *A. annua*.

A) *rbc*L **bioinformatics analysis**

1. rbcL sequences obtained from *A. thuscula* (At) and *A. annua* (Aa).

>At_Rbcla-F
TATTATACTCCTGAGTATGAAACCAAGGATACTGATATCTTGGCAGCATTTCGAGTAACTCCTCAACCGGGA
GTTCCGCCTGAAGAAGCAGGGGCCGCAGTAGCTGCCGAATCTTCTACTGGTACATGGACAACTGTGTGGAC
CGATGGACTTACGAGCCTTGATCGTTACAAAGGGCGATGCTATGGAATTGAGCCTGTTCTTGGAGAAGAGA
ATCAATATATTTGCTATGTAGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTC
CATTGTAGGTAACGTATTTGGTTTCAAAGCCCTGCGTGCTCTACGTCTGGAAGATTTGCGAATTCCTACTGC
GTATGTTAAAACTTTCCAAGGTCCGCCTCACGGTATCCAAGTTGAAAGAGATAAATTGAACAAGTATGGTC
GTCCTCTGTTGGGATGTACTATTAAACCTAAATTGGGGTTATCCGCTAAAAACTACGGTAGAGCTGTTTATG
AATGTCTTCGTGGTGGACTTGATTTTACA

>At_Rbcla-R
TATGTCACCACAAACAGAGACTAAAGCAAGTGTTGGATTCAAAGCTGGGGTTAAAGATTATAAATTGACTT
ATTATACTCCTGAGTATGAAACCAAGGATACTGATATCTTGGCAGCATTTCGAGTAACTCCTCAACCGGGAG
TTCCGCCTGAAGAAGCAGGGGCCGCAGTAGCTGCCGAATCTTCTACTGGTACATGGACAACTGTGTGGACC
GATGGACTTACGAGCCTTGATCGTTACAAAGGGCGATGCTATGGAATTGAGCCTGTTCTTGGAGAAGAGAA
TCAATATATTTGCTATGTAGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCC
ATTGTAGGTAACGTATTTGGTTTC

>**Aa_Rbcla-F**

CCTGAGTATGAAACCAAGGATACTGATATCTTGGCAGCATTTCGAGTAACTCCTCAACCTGGAGTTCCGCCT
GAAGAAGCAGGGGCCGCAGTAGCTGCCGAATCTTCTACTGGTACATGGACAACTGTGTGGACCGATGGACT
TACGAGCCTTGATCGTTACAAAGGGCGATGCTATGGAATTGAGCCTGTTCCTGGAGAAGAGAATCAATATA
TTTGCTATGTAGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGTAGG
TAACGTATTTGGTTTCAAAGCCCTGCGTGCTCTACGTCTGGAAGATTTGCGAATTCCTACTGCGTATGTTAA
AACTTTCCAAGGTCCGCCTCACGGTATCCAAGTTGAAAGAGATAAATTGAACAAGTATGGTCGTCCTCTGTT
GGGATGTACTATTAAACCTAAATTGGGGTTATCTGCTAAAAACTACGGTAGAGCTGTTTATGAATGTCTTCG
TGGTGGACTTGATTTTACA

>**Aa_Rbcla-R**

ATGTCACCACAAACAGAGACTAAAGCAAGTGTTGGATTCAAAGCTGGGGTTAAAGATTATAAATTGACTTA
TTATACTCCTGAGTATGAAACCAAGGATACTGATATCTTGGCAGCATTTCGAGTAACTCCTCAACCTGGAGT
TCCGCCTGAAGAAGCAGGGGCCGCAGTAGCTGCCGAATCTTCTACTGGTACATGGACAACTGTGTGGACCG
ATGGACTTACGAGCCTTGATCGTTACAAAGGGCGATGCTATGGAATTGAGCCTGTTCCTGGAGAAGAGAAT
CAATATATTTGCTATGTAGCTTACCCATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCA
TTGTAGGTAACGTATTTGGTTTCAAAGCCCTGCGTGCTCTACGTCTGGAAGATTTGCGAATTCCTACTGCGT
ATGTTAAAACTTTCCAAGGTCCGCCTCACGGTATCCAAGTTGAAAGAGATAAATTGAACAAGTATGGTCGT
CCTCTGTTGGGATGTACTATTAAACCTAAA

2. Consensus sequences for *rbcL*.

SNPs that not alters the protein sequence (synonymous mutations) are shown in blue (c.139T>G and c.543T>C, with respect to *A. annua* complete *rbc*L coding sequence). Non-synonymous substitution (c.272C>T; p.Pro91Leu) is shown in red.

>**At_*rbc*L(553bp)**

AAGTGTTGGATTCAAAGCTGGGGTTAAAGATTATAAATTGACTTATTATACTCCTGAGTATGAAACCAAGG
ATACTGATATCTTGGCAGCATTTCGAGTAACTCCTCAACC**G**GGAGTTCCGCCTGAAGAAGCAGGGGCCGCA
GTAGCTGCCGAATCTTCTACTGGTACATGGACAACTGTGTGGACCGATGGACTTACGAGCCTTGATCGTTAC
AAAGGGCGATGCTATGGAATTGAGCCTGTTC**T**TGGAGAAGAGAATCAATATATTTGCTATGTAGCTTACCC
ATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGTAGGTAACGTATTTGGTTTCAAA
GCCCTGCGTGCTCTACGTCTGGAAGATTTGCGAATTCCTACTGCGTATGTTAAAACTTTCCAAGGTCCGCCT
CACGGTATCCAAGTTGAAAGAGATAAATTGAACAAGTATGGTCGTCCTCTGTTGGGATGTACTATTAAACCT
AAATTGGGGTTATC**C**GCTAAAAACTACGGTAGAGCTGTTTATGAATGTCTT

>**Aa_*rbc*L(553bp)**

AAGTGTTGGATTCAAAGCTGGGGTTAAAGATTATAAATTGACTTATTATACTCCTGAGTATGAAACCAAGG
ATACTGATATCTTGGCAGCATTTCGAGTAACTCCTCAACC**T**GGAGTTCCGCCTGAAGAAGCAGGGGCCGCA
GTAGCTGCCGAATCTTCTACTGGTACATGGACAACTGTGTGGACCGATGGACTTACGAGCCTTGATCGTTAC
AAAGGGCGATGCTATGGAATTGAGCCTGTTC**C**TGGAGAAGAGAATCAATATATTTGCTATGTAGCTTACCC
ATTAGACCTTTTTGAAGAAGGTTCTGTTACTAACATGTTTACTTCCATTGTAGGTAACGTATTTGGTTTCAAA
GCCCTGCGTGCTCTACGTCTGGAAGATTTGCGAATTCCTACTGCGTATGTTAAAACTTTCCAAGGTCCGCCT
CACGGTATCCAAGTTGAAAGAGATAAATTGAACAAGTATGGTCGTCCTCTGTTGGGATGTACTATTAAACCT
AAATTGGGGTTATC**T**GCTAAAAACTACGGTAGAGCTGTTTATGAATGTCTT

B) *mat*K bioinformatics analysis

1. *mat*K sequences obtained from *A. thuscula* (At) and *A. annua* (Aa).

*Only low-quality sequences were obtained for *A. annua*, after several attempts. Therefore, these sequences were excluded from the analysis.
**Sequences retrieved from NCBI-GenBank to obtain the *mat*K consensus sequence for *A. annua*,

>**At_MatK-KIM3F**

AGTCGAAGTATATACTTTACTCGATACAAACTCTTTTTTTTTGAAGATCCACTATGATAATGAGAAAGATTT
CTGTATATACGCCCAAAGCGCTCAATAATATCAGAATCTGATAAATCGGTCCAAATCGCCTTGCCAATAGG
ATGCCCCAATGCGTTACAAAATTTCGATTTAGCCAGTGATCCAATCAGAGGCATAATTGGAACAAGAGTAT
CAAACTTCTTAATAGCATTATCGATTAGAAATGCATTTTCTAGCATTTGACTGCGTACCGTTGAAGGATTTA
GCCGCACACTTGAACGATAACCCAGAAAGTCAAGGGAATGGTTGGATAATTGGTTTATATAAATCCTTCCT
GGTTGAGGCCACAGGTAAAAATAATATTTCCAGAAATTGACAAAGTAATATTTCCATTTATTCATCAAAAG

AAACGTCCCTTTTGAAGCAAGAATGGATTTTCCTTGATACCTAACATAATGCATGAAAGGATCTTTGAACAA
CCATAAATTTGCTTGAAAAGACCTGACAAAGACTTCTGCAAGATGCTCTATTTTTCCATAGAAATTTATTCG
TTCAATAAGGGCTCCAGAAGATGTTGATCGTAAGTGAGAAGACTGGTTACGGAGAAAGAGGAAGCCAGAT
TCATATTCACATACATGAAAAGTATATAGGAAGAAGAATAATCTGTTTTTTCTTTTTGAAAAAGAAGAACTA
ACTTTCTTTGAATTTGAAGTAATAAGACTATCCCAATTATGACACTCATGGAGAAAGAATCTTAATAAATGC
AAAGAGGAAGCATCTTTTATCCAATAGCGAAGAGCCTGAACCAAGATTTCCAGAATGGACTGGGT

>At_MatK-KIM1R
TGTGTTTACGAGCTAAAGTTCTAGCACAAGAAAGTCGAAGTATATACTTTACTCGATACAAACTCTTTTTTTT
TGAAGATCCACTATGATAATGAGAAAGATTTCTGTATATACGCCCAAAGCGCTCAATAATATCAGAATCTG
ATAAATCGGTCCAAATCGCCTTGCCAATAGGATGCCCCAATGCGTTACAAAATTTCGATTTAGCCAGTGATC
CAATCAGAGGCATAATTGGAACAAGAGTATCAAACTTCTTAATAGCATTATCGATTAGAAATGCATTTTCTA
GCATTTGACTGCGTACCGTTGAAGGATTTAGCCGCACACTTGAACGATAACCCAGAAAGTCAAGGGAATGG
TTGGATAATTGGTTTATATAAATCCTTCCTGGTTGAGGCCACAGGTAAAAATAATATTTCCAGAAATTGACA
AAGTAATATTTCCATTTATTCATCAAAAGAAACGTCCCTTTTGAAGCAAGAATGGATTTTCCTTGATACCTA
ACATAATGCATGAAAGGATCTTTGAACAACCATAAATTTGCTTGAAAAGACCTGACAAAGACTTCTGCAAG
ATGCTCTATTTTTCCATAGAAATTTATTCGTTCAATAAGGGCTCCAGAAGATGTTGATCGTAAGTGAGAAGA
CTGGTTACGGAGAAAGAGGAAGCCAGATTCATATTCACATACATGAAAAGTATATAGGAAGAAGAATAAT
CTGTTTTTTCTTTTTGAAAAAGAAGAACTAACTTTCTTTGAATTTGAAGTAATAAGACTATCCCAATTATGAC
ACTC

>AaMatK-KIM3F*
NNNNNN

>AaMatK-KIM1R*
NNNNNN

** KY085890.1; MF623173.1; HM989753.1; KX581897.1; KX581896.1; KX581895.1; MK509452.1; KJ499926.1.

2. Consensus sequences for *mat*K.

SNPs that not alters the protein sequence (synonymous mutations) are shown in blue (c.975T>C and c.1134A>G, with respect to *A. annua* complete *mat*K coding sequence).

>At_*mat*K(842bp)
AGGCTCTTCGCTATTGGATAAAAGATGCTTCCTCTTTGCATTTATTAAGATTCTTTCTCCATGAGTGTCATAA
TTGGGATAGTCTTATTACTTCAAATTCAAAGAAAGTTAGTTCTTCTTTTTCAAAAAGAAAAAACAGATTATT
CTTCTTCCTATATACTTTTCATGTATGTGAATATGAATCTGGCTTCCTCTTTCTCCGTAACCAGTCTTCTCACT
TACGATCAACATCTTCTGGAGCCCTTATTGAACGAATAAATTTCTATGGAAAAATAGAGCATCTTGCAGAAG
TCTTTGTCAGGTCTTTTCAAGCAAATTTATGGTTGTTCAAAGATCCTTTCATGCATTATGTTAGGTATCAAGG
AAAATCCATTCTTGCTTCAAAAGGGACGTTTCTTTTGATGAATAAATGGAAATATTACTTTGTCAATTTCTGG
AAATATTATTTTTACCTGTGGCCTCAACCAGGAAGGATTTATATAAACCAATTATCCAA**C**CATTCCCTTGAC
TTTCTGGGTTATCGTTCAAGTGTGCGGCTAAATCCTTCAACGGTACGCAGTCAAATGCTAGAAAATGCATTT
CTAATCGATAATGCTATTAAGAAGTTTGATACTCTTGTTCCAATTATGCCTCTGATTGGATCACTGGCTAAAT
C**G**AAATTTTGTAACGCATTGGGGCATCCTATTGGCAAGGCGATTTGGACCGATTTATCAGATTCTGATATTA
TTGAGCGCTTTGGGCGTATATACAGAAATCTTTCTCATTATCATAGTGGATCTTCAAAAAAAAAGAGTTTGT
ATCGAGTAAAGTATATACTTCGACTTTCTTGTGCTAGAACTTTAG

>Aa_*mat*K(842bp)
AGGCTCTTCGCTATTGGATAAAAGATGCTTCCTCTTTGCATTTATTAAGATTCTTTCTCCATGAGTGTCATAA
TTGGGATAGTCTTATTACTTCAAATTCAAAGAAAGTTAGTTCTTCTTTTTCAAAAAGAAAAAACAGATTATT
CTTCTTCCTATATACTTTTCATGTATGTGAATATGAATCTGGCTTCCTCTTTCTCCGTAACCAGTCTTCTCACT
TACGATCAACATCTTCTGGAGCCCTTATTGAACGAATAAATTTCTATGGAAAAATAGAGCATCTTGCAGAAG
TCTTTGTCAGGTCTTTTCAAGCAAATTTATGGTTGTTCAAAGATCCTTTCATGCATTATGTTAGGTATCAAGG
AAAATCCATTCTTGCTTCAAAAGGGACGTTTCTTTTGATGAATAAATGGAAATATTACTTTGTCAATTTCTGG
AAATATTATTTTTACCTGTGGCCTCAACCAGGAAGGATTTATATAAACCAATTATCCAA**T**CATTCCCTTGAC
TTTCTGGGTTATCGTTCAAGTGTGCGGCTAAATCCTTCAACGGTACGCAGTCAAATGCTAGAAAATGCATTT
CTAATCGATAATGCTATTAAGAAGTTTGATACTCTTGTTCCAATTATGCCTCTGATTGGATCACTGGCTAAAT
C**A**AAATTTTGTAACGCATTGGGGCATCCTATTGGCAAGGCGATTTGGACCGATTTATCAGATTCTGATATTA
TTGAGCGCTTTGGGCGTATATACAGAAATCTTTCTCATTATCATAGTGGATCTTCAAAAAAAAAGAGTTTGT
ATCGAGTAAAGTATATACTTCGACTTTCTTGTGCTAGAACTTTAG

C) List of sequences retrieved from NCBI taxonomy browser

1. *rbc*L sequences.

1.1. *rbc*L sequences from the *Artemisia* genus retrieved from NCBI taxonomy browser to obtain a preliminary alignment. Only secuences with species-level definition were selected.

Species	Ac. No.	Species	Ac. No.	Species	Ac. No.	Species	Ac. No.
A. abrotanum	KX783830.1	*A. borealis*	KC474118.1	*A. hyperborea*	MG224211.1	*A. scoparia*	KX582029.1
A. abrotanum	MG222415.1	*A. borealis*	KC474119.1	*A. hyperborea*	KC474125.1	*A. scoparia*	KX282550.1
A. abrotanum	MG223606.1	*A. borealis*	KC474120.1	*A. hyperborea*	KC474126.1	*A. scoparia*	KX282551.1
A. abrotanum	MN167228.1	*A. borealis*	KC474121.1	*A. hyperborea*	KC474127.1	*A. scoparia*	KX282552.1
A. absinthium	KX581993.1	*A. borealis*	KC474122.1	*A. hyperborea*	KC474128.1	*A. scoparia*	GU724242.1
A. absinthium	KX581994.1	*A. borealis*	KC474123.1	*A. hyperborea*	KC474129.1	*A. serrata*	MK525242.1
A. absinthium	KX581995.1	*A. borealis*	KC474124.1	*A. hyperborea*	KC482050.1	*A. sibirica*	KX527325.1
A. absinthium	KX581996.1	*A. borealis*	KC482037.1	*A. hyperborea*	KC482051.1	*A. sieversiana*	KX582030.1
A. absinthium	KX679031.1	*A. borealis*	KC482038.1	*A. hyperborea*	KC482052.1	*A. sieversiana*	KX582031.1
A. absinthium	MG222186.1	*A. borealis*	KC482039.1	*A. hyperborea*	KC482053.1	*A. sieversiana*	KX582032.1
A. absinthium	MG223720.1	*A. borealis*	KC482040.1	*A. igniaria*	JQ173396.1	*A. sieversiana*	MF158791.1
A. absinthium	MG946820.1	*A. borealis*	KC482041.1	*A. indica*	LC413432.1	*A. sieversiana*	JQ173398.1
A. absinthium	MG946821.1	*A. borealis*	KC482042.1	*A. indica*	MH116070.1	*A. sinanensis*	LC377038.1
A. absinthium	MK525237.1	*A. borealis*	KC482043.1	*A. indica*	MH116071.1	*A. stelleriana*	MG223384.1
A. absinthium	MK348958.1	*A. borealis*	KC482044.1	*A. japonica*	LC364390.1	*A. suksdorfii*	KX677904.1
A. absinthium	JN890797.1	*A. borealis*	KC482045.1	*A. japonica*	KF476063.1	*A. suksdorfii*	MG224174.1
A. absinthium	JN891748.1	*A. borealis*	KC482046.1	*A. judaica*	KX709619.1	*A. tilesii*	MG223886.1
A. absinthium	JN892095.1	*A. borealis*	KC482047.1	*A. laciniata*	MG221405.1	*A. tilesii*	MG224715.1
A. absinthium	HE963336.1	*A. borealis*	KC482048.1	*A. laciniata*	MG223389.1	*A. tilesii*	JN862215.1
A. afra	AM234849.1	*A. borealis*	KC482049.1	*A. laciniata*	MG224635.1	*A. tilesii*	KC474130.1
A. afra	JQ412318.1	*A. campestris*	MG221482.1	*A. lactiflora*	GU724217.1	*A. tilesii*	KC474131.1
A. alaskana	MG222455.1	*A. campestris*	MG221985.1	*A. lactiflora*	GU724218.1	*A. tilesii*	KC474132.1
A. alaskana	MG222786.1	*A. campestris*	MG222107.1	*A. lactiflora*	GU724219.1	*A. tilesii*	KC474133.1
A. alaskana	MG223570.1	*A. campestris*	MG222160.1	*A. lavandulifolia*	GQ436484.1	*A. tilesii*	KC482054.1
A. alaskana	JN862213.1	*A. campestris*	MG222725.1	*A. longifolia*	MG222197.1	*A. tilesii*	KC482055.1
A. anethoides	KX581997.1	*A. campestris*	MG223436.1	*A. ludoviciana*	MG221701.1	*A. tilesii*	KC482056.1
A. anethoides	KX581998.1	*A. campestris*	MG224287.1	*A. ludoviciana*	MG222341.1	*A. tournefortiana*	KX582033.1
A. anethoides	KX581999.1	*A. campestris*	MG224659.1	*A. ludoviciana*	MG223621.1	*A. tournefortiana*	KX582034.1
A. annua	KX582000.1	*A. campestris*	MK525240.1	*A. ludoviciana*	MG224300.1	*A. tournefortiana*	KX582035.1
A. annua	KX582001.1	*A. campestris*	MK925165.1	*A. ludoviciana*	JX848405.1	*A. tournefortiana*	KX582036.1
A. annua	KX582002.1	*A. campestris*	MK925212.1	*A. macrocephala*	KX582014.1	*A. tridentata*	KU905016.1
A. annua	MG221862.1	*A. campestris*	JN890800.1	*A. maritima*	JN892340.1	*A. tridentata*	KU905017.1
A. annua	MG222743.1	*A. campestris*	JX848403.1	*A. maritima*	KF997353.1	*A. tridentata*	KX677988.1
A. annua	MG224658.1	*A. campestris*	KJ746262.1	*A. michauxiana*	KX678802.1	*A. tridentata*	MG221902.1
A. annua	MH087481.1	*A. cana*	MG221884.1	*A. michauxiana*	MG221182.1	*A. tridentata*	MG223132.1
A. annua	MH051919.1	*A. cana*	MG222338.1	*A. monosperma*	KX709618.1	*A. tridentata*	MG668936.1
A. annua	MK525238.1	*A. capillaris*	JF949967.2	*A. myriantha*	LT576796.1	*A. tridentata*	MH025371.1
A. annua	MK903549.1	*A. capillaris*	JQ173395.1	*A. norvegica*	MF963097.1	*A. tridentata*	MH048919.1
A. annua	JF949966.2	*A. cina*	MK895573.1	*A. norvegica*	MG221961.1	*A. tridentata*	MH048920.1
A. annua	JQ173392.1	*A. douglasiana*	KF613101.1	*A. norvegica*	MG223362.1	*A. tridentata*	JN862216.1
A. annua	JQ173393.1	*A. douglasiana*	KF613102.1	*A. norvegica*	MG224148.1	*A. tripartita*	MG222196.1
A. annua	JQ173394.1	*A. dracunculus*	KX582003.1	*A. norvegica*	MK925043.1	*A. vulgaris*	LT576797.1
A. annua	DQ006057.1	*A. dracunculus*	KX582004.1	*A. norvegica*	MK925241.1	*A. vulgaris*	KX582037.1
A. annua	KJ667633.1	*A. dracunculus*	KX582005.1	*A. norvegica*	MK925401.1	*A. vulgaris*	KX582038.1
A. annua	KJ667647.1	*A. dracunculus*	KX582006.1	*A. nova*	KY584343.1	*A. vulgaris*	KX582039.1
A. annua	KJ667651.1	*A. dracunculus*	KX582007.1	*A. pontica*	KX582015.1	*A. vulgaris*	KX582040.1
A. annua	KJ667662.1	*A. dracunculus*	KX582008.1	*A. pontica*	KX582016.1	*A. vulgaris*	KX582041.1
A. arctica	MG223305.1	*A. dracunculus*	KX582009.1	*A. pontica*	KX582017.1	*A. vulgaris*	KX582042.1
A. arctica	MG224274.1	*A. dracunculus*	MG222966.1	*A. pontica*	KX582018.1	*A. vulgaris*	KX582043.1
A. arctica	MG224306.1	*A. dracunculus*	MG224134.1	*A. pontica*	MG221625.1	*A. vulgaris*	KX582044.1
A. arctica	MG224589.1	*A. dracunculus*	MG224512.1	*A. pontica*	MG224035.1	*A. vulgaris*	MG222521.1
A. arctica	JN862217.1	*A. dracunculus*	MF158804.1	*A. princeps*	KM218339.1	*A. vulgaris*	MG224447.1
A. argyi	GQ436428.1	*A. dracunculus*	MK525241.1	*A. pubescens*	KX582019.1	*A. vulgaris*	MK525243.1
A. argyi	GQ436429.1	*A. dracunculus*	MN167229.1	*A. pubescens*	KX582020.1	*A. vulgaris*	HQ593908.1
A. australis	MH755603.1	*A. dracunculus*	HQ593182.1	*A. pubescens*	KX582021.1	*A. vulgaris*	HQ593183.1
A. australis	MH755604.1	*A. frigida*	MG222519.1	*A. roxburghiana*	KT280075.1	*A. vulgaris*	HQ594578.1

A. biennis	MG222339.1	A. frigida	MG223945.1	A. roxburghiana	KJ372409.1	A. vulgaris	HQ596606.1
A. biennis	MG224137.1	A. frigida	MG224121.1	A. rupestris	KX582022.1	A. vulgaris	JN890801.1
A. biennis	MK525239.1	A. frigida	JN862214.1	A. rupestris	KX582023.1	A. vulgaris	JN891751.1
A. borealis	MG221553.1	A. frigida	JX848404.1	A. rupestris	MG221497.1	A. vulgaris	JN892237.1
A. borealis	MG221903.1	A. globularia	MG223490.1	A. rupestris	MG222280.1	A. vulgaris	HE963337.1
A. borealis	JN862218.1	A. glomerata	MG223735.1	A. sacrorum	JQ173397.1	A. vulgaris	KC870884.1
A. borealis	KC474111.1	A. glomerata	MG224016.1	A. salsoloides	MF694666.1	A. vulgaris	KF589298.1
A. borealis	KC474112.1	A. gmelinii	KX582010.1	A. salsoloides	MF694951.1	A. vulgaris	KF639960.1
A. borealis	KC474113.1	A. gmelinii	KX582011.1	A. scoparia	KX582024.1	A. vulgaris	KF664584.1
A. borealis	KC474114.1	A. gmelinii	KX582012.1	A. scoparia	KX582025.1	A. vulgaris	KM360653.1
A. borealis	KC474115.1	A. gmelinii	KX582013.1	A. scoparia	KX582026.1		
A. borealis	KC474116.1	A. gmelinii	GQ436432.1	A. scoparia	KX582027.1		
A. borealis	KC474117.1	A. herba-alba	KX282549.1	A. scoparia	KX582028.1		

1.2. *rbc*L **sequences selected as representatives for each species, to obtain the curated alignment.** *A. thuscula* **sequence obtained in the present work, as well as outgroups, are shown:**

Species	Ac. No.	Species	Ac. No.	Species	Acc. No.	Species	Acc. No.
A. abrotanum	MN167228.1	A. dracunculus	KX582003.1	A. macrocephala	KX582014.1	A. sibirica	KX527325.1
A. afra	JQ412318.1	A. frigida	MG222519.1	A. maritima	KF997353.1	A. sieversiana	KX582030.1
A. alaskana	MG222455.1	A. globularia	MG223490.1	A. michauxiana	KX678802.1	A. stelleriana	MG223384.1
A. anethoides	KX581997.1	A. glomerata	MG223735.1	A. monosperma	KX709618.1	A. suksdorfii	KX677904.1
A. annua	KX582000.1	A. gmelinii	KX582010.1	A. myriantha	LT576796.1	**A. thuscula**	**Present work**
A. arctica	MG224306.1	A. hyperborea	MG224211.1	A. norvegica	MF963097.1	A. tilesii	MG223886.1
A. argyi	GQ436428.1	A. igniaria	JQ173396.1	A. nova	KY584343.1	A. tournefortiana	KX582033.1
A. australis	MH755603.1	A. indica	MH116070.1	A. pontica	KX582015.1	A. tridentata	KU905016.1
A. biennis	MG222339.1	A. japonica	LC364390.1	A. pubescens	KX582019.1	A. tripartita	MG222196.1
A. borealis	KC482038.1	A. laciniata	MG221405.1	A. roxburghiana	KT280075.1	A. vulgaris	LT576797.1
A. campestris	MG222107.1	A. lactiflora	GU724219.1	A. rupestris	MG221497.1	*Ajania fruticulosa	KX527160.1
A. cana	MG222338.1	A. lavandulifolia	GQ436484.1	A. sacrorum	JQ173397.1	*Chrysanthemum indicum	JN867592.1
A. capillaris	JQ173395.1	A. longifolia	MG222197.1	A. salsoloides	MF694951.1	**Anthemis arvensis	MG222653.1
A. douglasiana	KF613102.1	A. ludoviciana	MG223621.1	A. scoparia	GU724242.1	**Achillea millefolium	EU384938.1

Chrysanthemum and *Ajania* represents two different genera of the subtribe *Artemisiinae*. These two sequences were included to obtain an overall vision of phylogenetic differences at the subtribe level.

** Outgroup sequences. *Anthemis* (subtribe *Anthemidinae*) and *Achillea* (subtribe *Matricariinae*) represents two *Artemisia*-related genera that belong to different subtribes, but to the same tribe as *Artemisia* (subtribe *Artemisiinae*, tribe *Anthemideae*).

2. *mat*K **sequences.**

2.1. matK sequences from the *Artemisia* **genus, retrieved from NCBI taxonomy browser to obtain a preliminary alignment. Only secuences with species-level definition were selected.**

Species	Ac. No.	Species	Ac. No.	Species	Ac. No.	Species	Ac. No.
A. abrotanum	KX783637.1	A. borealis	KC474117.1	A. gmelinii	KX581908.1	A. scoparia	HM989797.1
A. abrotanum	MN167188.1	A. borealis	KC474118.1	A. gmelinii	MG282059.1	A. scoparia	KX581919.1
A. absinthium	JN894044.1	A. borealis	KC474119.1	A. gurganica	MG282058.1	A. scoparia	KX581920.1
A. absinthium	JN894750.1	A. borealis	KC474120.1	A. herba-alba	KX758475.1	A. scoparia	KX581921.1
A. absinthium	HE970675.1	A. borealis	KC474121.1	A. hyperborea	KC474125.1	A. scoparia	KX581922.1
A. absinthium	KX581888.1	A. borealis	KC474122.1	A. hyperborea	KC474126.1	A. scoparia	KX581923.1
A. absinthium	KX581889.1	A. borealis	KC474123.1	A. hyperborea	KC474127.1	A. scoparia	KX581924.1
A. absinthium	KX581890.1	A. borealis	KC474124.1	A. hyperborea	KC474128.1	A. scopiformis	MG282054.1
A. absinthium	KX581891.1	A. campestris	JN894047.1	A. hyperborea	KC474129.1	A. serrata	MK509458.1
A. absinthium	KX677578.1	A. campestris	MG224837.1	A. hyperborea	MG225028.1	A. sieversiana	JQ173391.1
A. absinthium	MG225207.1	A. campestris	MG224910.1	A. hyperborea	MG225303.1	A. sieversiana	KX581925.1
A. absinthium	MG946952.1	A. campestris	MG224929.1	A. igniaria	JQ173389.1	A. sieversiana	KX581926.1
A. absinthium	MG946953.1	A. campestris	MG224936.1	A. indica	MH116552.1	A. sieversiana	KX581927.1
A. absinthium	MK509451.1	A. campestris	MG225037.1	A. indica	MH116553.1	A. sieversiana	MF158701.1

A. afra	JQ412200.1	*A. campestris*	MG225161.1	*A. japonica*	HM989768.1	*A. sieversiana*	MK435681.1
A. alaskana	MG224989.1	*A. campestris*	MG225228.1	*A. japonica*	KF530805.1	*A. sublessingiana*	MG282053.1
A. alaskana	MG225045.1	*A. campestris*	MG225314.1	*A. japonica*	MK435679.1	*A. suksdorfii*	KX676605.1
A. alaskana	MG225187.1	*A. campestris*	MK509454.1	*A. judaica*	KX758474.1	*A. terrae-albae*	MG282052.1
A. anethoides	KX581892.1	*A. campestris*	MK925653.1	*A. kotuchovii*	MG282057.1	*A. tilesii*	KC474130.1
A. anethoides	KX581893.1	*A. campestris*	MK926386.1	*A. kruhsiana*	FN668454.1	*A. tilesii*	KC474131.1
A. anethoides	KX581894.1	*A. capillaris*	JQ173388.1	*A. laciniata*	MG224824.1	*A. tilesii*	KC474132.1
A. annua	MK509452.1	*A. cina*	MK898774.1	*A. laciniata*	MG225153.1	*A. tilesii*	KC474133.1
A. annua	HM989753.1	*A. dracunculus*	HQ593182.1	*A. laciniata*	MG225371.1	*A. tilesii*	MG225236.1
A. annua	HM989754.1	*A. dracunculus*	KX581898.1	*A. lactiflora*	HM989727.1	*A. tilesii*	MG225381.1
A. annua	JQ173387.1	*A. dracunculus*	KX581899.1	*A. lactiflora*	HM989728.1	*A. tournefortiana*	KX581928.1
A. annua	KJ499926.1	*A. dracunculus*	KX581900.1	*A. lactiflora*	HM989729.1	*A. tournefortiana*	KX581929.1
A. annua	KJ499961.1	*A. dracunculus*	KX581901.1	*A. ludoviciana*	MK509457.1	*A. tournefortiana*	KX581930.1
A. annua	KJ499964.1	*A. dracunculus*	KX581902.1	*A. macrocephala*	KX581909.1	*A. tournefortiana*	KX581931.1
A. annua	KX581895.1	*A. dracunculus*	KX581903.1	*A. michauxiana*	KX677380.1	*A. transiliensis*	MG282051.1
A. annua	KX581896.1	*A. dracunculus*	KX581904.1	*A. michauxiana*	MG224780.1	*A. tridentata*	AF456776.1
A. annua	KX581897.1	*A. dracunculus*	MG225073.1	*A. michauxiana*	MG225022.1	*A. tridentata*	KX676677.1
A. arctica	MG225310.1	*A. dracunculus*	MG225256.1	*A. michauxiana*	MG225087.1	*A. vulgaris*	HQ593183.1
A. arctica	MG225363.1	*A. dracunculus*	MG225292.1	*A. norvegica*	MG224906.1	*A. vulgaris*	JN894048.1
A. arctica	FN668453.1	*A. dracunculus*	MG225350.1	*A. norvegica*	MG225147.1	*A. vulgaris*	JN894753.1
A. arctisibirica	FN668458.1	*A. dracunculus*	MF158713.1	*A. norvegica*	MG225295.1	*A. vulgaris*	HE967349.1
A. argyi	HM989725.1	*A. dracunculus*	MK509455.1	*A. norvegica*	MK926361.1	*A. vulgaris*	KC870883.1
A. argyi	HM989726.1	*A. dracunculus*	MK800537.1	*A. norvegica*	MF963479.1	*A. vulgaris*	KF604887.1
A. atrovirens	MK435678.1	*A. dracunculus*	MN161189.1	*A. parviflora*	MK435680.1	*A. vulgaris*	KF648716.1
A. australis	MH755557.1	*A. frigida*	MG225000.1	*A. pontica*	KX581910.1	*A. vulgaris*	KF664585.1
A. australis	MH755558.1	*A. frigida*	MG225250.1	*A. pontica*	KX581911.1	*A. vulgaris*	KR231888.1
A. biennis	MG224892.1	*A. frigida*	MG225289.1	*A. pontica*	KX581912.1	*A. vulgaris*	KX581932.1
A. biennis	MG224968.1	*A. frigida*	MG225356.1	*A. pontica*	KX581913.1	*A. vulgaris*	KX581933.1
A. biennis	MG225294.1	*A. frigida*	MK509456.1	*A. pubescens*	KX581914.1	*A. vulgaris*	KX581934.1
A. biennis	MG225328.1	*A. furcata*	FN668456.1	*A. pubescens*	KX581915.1	*A. vulgaris*	KX581935.1
A. biennis	MK509453.1	*A. furcata*	MG224783.1	*A. pubescens*	KX581916.1	*A. vulgaris*	KX581936.1
A. borealis	FN668457.1	*A. globularia*	MG224997.1	*A. radicans*	MG282056.1	*A. vulgaris*	KX581937.1
A. borealis	MG224848.1	*A. globularia*	MG225024.1	*A. roxburghiana*	KJ372399.1	*A. vulgaris*	KX581938.1
A. borealis	MG224883.1	*A. glomerata*	FN668455.1	*A. roxburghiana*	KT280182.1	*A. vulgaris*	KX581939.1
A. borealis	MG224896.1	*A. glomerata*	MG224975.1	*A. rupestris*	KX581917.1	*A. vulgaris*	MG225341.1
A. borealis	KC474111.1	*A. glomerata*	MG225211.1	*A. rupestris*	KX581918.1	*A. vulgaris*	MF770237.1
A. borealis	KC474112.1	*A. glomerata*	MG225264.1	*A. rupestris*	MG224838.1	*A. vulgaris*	MK509459.1
A. borealis	KC474113.1	*A. gmelinii*	GQ434109.1	*A. rupestris*	MG224957.1		
A. borealis	KC474114.1	*A. gmelinii*	KX581905.1	*A. sacrorum*	JQ173390.1		
A. borealis	KC474115.1	*A. gmelinii*	KX581906.1	*A. salsoloides*	MF694828.1		
A. borealis	KC474116.1	*A. gmelinii*	KX581907.1	*A. santolinifolia*	MG282055.1		

2.2. *mat*K sequences selected as representatives for each species, to obtain the curated alignment. *A. thuscula* sequence obtained in the present work, as well as outgroups, are shown:

Species	Ac. No.	*Species*	Ac. No.	*Species*	Ac. No.	*Species*	Ac. No.
A. abrotanum	MN167188.1	*A. dracunculus*	KX581898.1	*A. ludoviciana*	MK509457.1	*A. serrata*	MK509458.1
A. absinthium	JN894044.1	*A. frigida*	MG225000.1	*A. macrocephala*	KX581909.1	*A. sieversiana*	JQ173391.1
A. afra	JQ412200.1	*A. furcata*	FN668456.1	*A. michauxiana*	MG225087.1	*A. sublessingiana*	MG282053.1
A. alaskana	MG224989.1	*A. globularia*	MG224997.1	*A. norvegica*	MG225147.1	*A. suksdorfii*	KX676605.1
A. anethoides	KX581892.1	*A. glomerata*	MG224975.1	*A. parviflora*	MK435680.1	*A. terraealbae*	MG282052.1
A. annua	HM989753.1	*A. gmelinii*	KX581907.1	*A. pontica*	KX581910.1	***A. thuscula***	**Present work**
A. arctica	FN668453.1	*A. gurganica*	MG282058.1	*A. pubescens*	KX581914.1	*A. tilesii*	KC474130.1
A. arctisibirica	FN668458.1	*A. hyperborea*	KC474125.1	*A. radicans*	MG282056.1	*A. tournefortiana*	KX581931.1
A. argyi	HM989726.1	*A. igniaria*	JQ173389.1	*A. roxburghiana*	KJ372399.1	*A. transiliensis*	MG282051.1
A. atrovirens	MK435678.1	*A. indica*	MH116552.1	*A. rupestris*	MG224957.1	*A. tridentata*	KX676677.1
A. australis	MH755557.1	*A. japonica*	HM989768.1	*A. sacrorum*	JQ173390.1	*A. vulgaris*	JN894753.1
A. biennis	MG224892.1	*A. kotuchovii*	MG282057.1	*A. salsoloides*	MF694828.1	*Ajania fruticulosa*	KX526529.1
A. borealis	KC474111.1	*A. kruhsiana*	FN668454.1	*A. santolinifolia*	MG282055.1	*Chrysanthemum indicum*	JN867592.1
A. campestris	JN894047.1	*A. laciniata*	MG225371.1	*A. scoparia*	KX581919.1	**Anthemis arvensis*	JN895748.1
A. capillaris	JQ173388.1	*A. lactiflora*	HM989728.1	*A. scopiformis*	MG282054.1	**Achillea millefolium*	KX677060.1

Chrysanthemum and *Ajania* represents two different genera of the subtribe *Artemisiinae*. These two sequences were included to obtain an overall vision of phylogenetic differences at the subtribe level.

** Outgroup sequences. *Anthemis* (subtribe *Anthemidinae*) and *Achillea* (subtribe *Matricariinae*) represents two *Artemisia*-related genera that belong to different subtribes, but to the same tribe as *Artemisia* (subtribe *Artemisiinae*, tribe *Anthemideae*).

3. *mat*K-*rbc*L sequence set. Only those species that were present in both *mat*K and *rbc*L alignments were included.

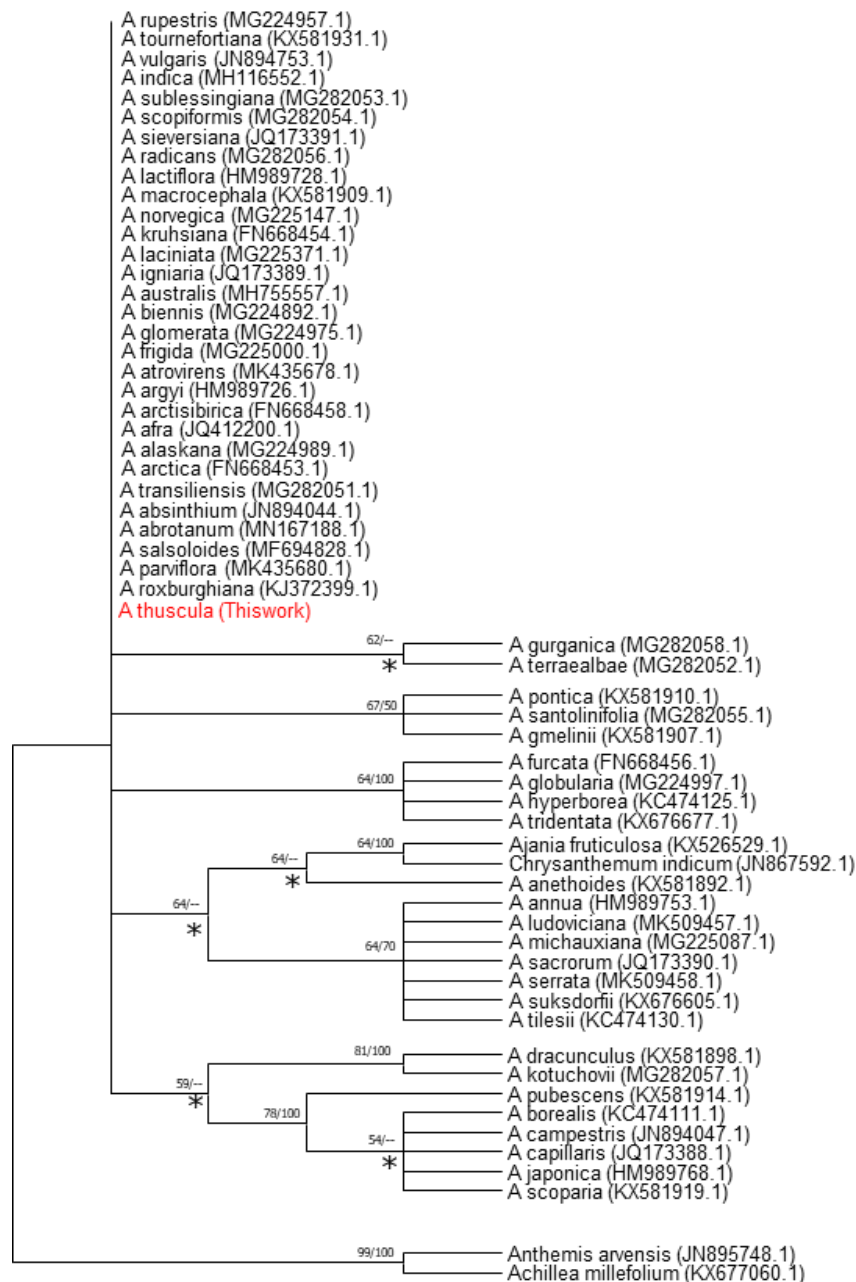| *Species*	*rbc*L Ac. No.	*mat*K Ac. No.	*Species*	*rbc*L Ac. No.	*mat*K Ac. No.
A. abrotanum	MN167188.1	MN167228.1	*A. ludoviciana*	MK509457.1	MG223621.1
A. afra	JQ412200.1	JQ412318.1	*A. macrocephala*	KX581909.1	KX582014.1
A. alaskana	MG224989.1	MG222455.1	*A. michauxiana*	MG225087.1	KX678802.1
A. anethoides	KX581892.1	KX581997.1	*A. norvegica*	MG225147.1	MF963097.1
A. annua	HM989753.1	KX582000.1	*A. pontica*	KX581910.1	KX582015.1
A. arctica	FN668453.1	MG224306.1	*A. pubescens*	KX581914.1	KX582019.1
A. argyi	HM989726.1	GQ436428.1	*A. roxburghiana*	KJ372399.1	KT280075.1
A. australis	MH755557.1	MH755603.1	*A. rupestris*	MG224957.1	MG221497.1
A. biennis	MG224892.1	MG222339.1	*A. sacrorum*	JQ173390.1	JQ173397.1
A. borealis	KC474111.1	KC482038.1	*A. salsoloides*	MF694828.1	MF694951.1
A. campestris	JN894047.1	MG222107.1	*A. scoparia*	KX581919.1	GU724242.1
A. capillaris	JQ173388.1	JQ173395.1	*A. sieversiana*	JQ173391.1	KX582030.1
A. dracunculus	KX581898.1	KX582003.1	*A. suksdorfii*	KX676605.1	KX677904.1
A. frigida	MG225000.1	MG222519.1	**A. thuscula**	**Present work**	**Present work**
A. globularia	MG224997.1	MG223490.1	*A. tilesii*	KC474130.1	MG223886.1
A. glomerata	MG224975.1	MG223735.1	*A. tournefortiana*	KX581931.1	KX582033.1
A. gmelinii	KX581907.1	KX582010.1	*A. tridentata*	KX676677.1	KU905016.1
A. hyperborea	KC474125.1	MG224211.1	*A. vulgaris*	JN894753.1	LT576797.1
A. igniaria	JQ173389.1	JQ173396.1	**Chrysanthemum indicum*	JN867592.1	JN867592.1
A. indica	MH116552.1	MH116070.1	**Ajania fruticulosa*	KX526529.1	KX527160.1
A. japonica	HM989768.1	LC364390.1	***Anthemis arvensis*	JN895748.1	MG222653.1
A. laciniata	MG225371.1	MG221405.1	***Achillea millefolium*	KX677060.1	EU384938.1
A. lactiflora	HM989728.1	GU724219.1			

*Chrysanthemum and *Ajania* represents two different genera of the subtribe *Artemisiinae*. These two sequences were included to obtain an overall vision of phylogenetic differences at the subtribe level.

** Outgroup sequences. *Anthemis* (subtribe *Anthemidinae*) and *Achillea* (subtribe *Matricariinae*) represents two *Artemisia*-related genera that belong to different subtribes, but to the same tribe as *Artemisia* (subtribe *Artemisiinae*, tribe *Anthemideae*).

D) Cladograms obtained from *mat*K and *rbc*L alignments

1. *mat*K-based cladogram.
Cladogram was inferred by Maximum Likelihood (ML) and Maximum Parsimony (MP) methods. Both methods reproduce phylogenyes with significant differences. The ML-based tree with the highest log likelihood (-1223.71) is shown, and branches which were not supported by the MP method are marked with an asterisk. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown over branches (ML/MP). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. ML method was performed with Tamura 3-parameter model with discrete Gamma distribution, including 5 categories (G = 0.0826; BIC = 3.72 x 10^6; AICc = 2,69 x 10^6). Pairwise distances were estimated using the Maximum Composite Likelihood (MCL) approach. The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (10 replicates), and consensus tree was inferred from 9 most parsimonious trees.

A rupestris (MG224957.1)
A tournefortiana (KX581931.1)
A vulgaris (JN894753.1)
A indica (MH116552.1)
A sublessingiana (MG282053.1)
A scopiformis (MG282054.1)
A sieversiana (JQ173391.1)
A radicans (MG282056.1)
A lactiflora (HM989728.1)
A macrocephala (KX581909.1)
A norvegica (MG225147.1)
A kruhsiana (FN668454.1)
A laciniata (MG225371.1)
A igniaria (JQ173389.1)
A australis (MH755557.1)
A biennis (MG224892.1)
A glomerata (MG224975.1)
A frigida (MG225000.1)
A atrovirens (MK435678.1)
A argyi (HM989726.1)
A arctisibirica (FN668458.1)
A afra (JQ412200.1)
A alaskana (MG224989.1)
A arctica (FN668453.1)
A transiliensis (MG282051.1)
A absinthium (JN894044.1)
A abrotanum (MN167188.1)
A salsoloides (MF694828.1)
A parviflora (MK435680.1)
A roxburghiana (KJ372399.1)
A thuscula (Thiswork)

62/-- A gurganica (MG282058.1)
* A terraealbae (MG282052.1)

67/50 A pontica (KX581910.1)
A santolinifolia (MG282055.1)
A gmelinii (KX581907.1)

64/100 A furcata (FN668456.1)
A globularia (MG224997.1)
A hyperborea (KC474125.1)
A tridentata (KX676677.1)

64/100 Ajania fruticulosa (KX526529.1)
64/-- Chrysanthemum indicum (JN867592.1)
* A anethoides (KX581892.1)
64/-- A annua (HM989753.1)
* A ludoviciana (MK509457.1)
A michauxiana (MG225087.1)
64/70 A sacrorum (JQ173390.1)
A serrata (MK509458.1)
A suksdorfii (KX676605.1)
A tilesii (KC474130.1)

81/100 A dracunculus (KX581898.1)
A kotuchovii (MG282057.1)
59/-- A pubescens (KX581914.1)
* A borealis (KC474111.1)
78/100 A campestris (JN894047.1)
54/-- A capillaris (JQ173388.1)
* A japonica (HM989768.1)
A scoparia (KX581919.1)

99/100 Anthemis arvensis (JN895748.1)
Achillea millefolium (KX677060.1)

2. *rbcL*-based cladogram.

Cladogram was inferred by Maximum Likelihood (ML) and Maximum Parsimony (MP) methods. Both methods reproduce the same tree topology. Therefore, only the ML-based tree with the highest log likelihood (-859.58) is shown. ML method was performed with Jukes and Cantor model, with uniform distribution (BIC = 2,84 x 10^6; ALCc = 1,94 x 10^6). Pairwise distances were estimated using the Maximum Composite Likelihood (MCL) approach. The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (10 replicates), and consensus tree was inferred from 10 most parsimonious trees. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown over branches (ML/MP) . Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed.