
# UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Mechanisms of speciation in bryophytes: evolutionary studies in the moss *Funaria hygrometrica* along environmental gradients

Mecanismos de especiación en briófitos: estudios evolutivos en el musgo *Funaria hygrometrica* a lo largo de gradientes ambientales

**Dña. Elena de la Cruz Martínez**

**2021**

# UNIVERSIDAD DE MURCIA

Mechanisms of speciation in bryophytes:
evolutionary studies in the moss *Funaria hygrometrica*
along environmental gradients

Mecanismos de especiación en briófitos:
estudios evolutivos en el musgo *Funaria hygrometrica*
a lo largo de gradientes ambientales

**Elena de la Cruz Martínez**

Esta tesis doctoral fue realizada en el Departamento de Biología Vegetal de la Universidad de Murcia, dirigida por:

**Dr. Olaf Franziskus Werner** (Universidad de Murcia)

**Dr. Rosa María Ros Espín** (Universidad de Murcia)

**Dr. Jairo Patiño Llorente** (Universidad de La Laguna)

**Murcia, 2021**

## FUENTES DE FINANCIACIÓN

La investigación llevada a cabo en el marco de esta Tesis Doctoral fue financiada por los proyectos CGL2014-52579-R del Ministerio de Economía y Competitividad (cofinanciado con fondos FEDER de la Unión Europea) y 19308/PI/14 de la Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia.

Asimismo, Elena de la Cruz Martínez disfrutó de una beca-contrato del Programa de Formación de Personal Investigador (referencia BES-2015-073780) del Ministerio de Economía y Competitividad, desde diciembre de 2015 hasta noviembre de 2019, que incluyó una ayuda a la movilidad predoctoral para la realización de una estancia breve en la Universidad de La Laguna, Tenerife, España, en 2019 (3 meses).

# AGRADECIMIENTOS

Durante estos casi seis años, muchas han sido las personas que han contribuido de una u otra forma a que alcanzase mi meta de terminar con éxito mi tesis doctoral. A todas ellas, mi más sincero agradecimiento.

En primer lugar, quiero hacer especial mención a la inestimable ayuda y labor realizada por mis directores de tesis en la Universidad de Murcia, Olaf Franziskus Werner y Rosa María Ros Espín, esta última también mi tutora de tesis, así como a mi director en la Universidad de La Laguna, Jairo Patiño Llorente, de no ser por ellos no tendría ahora entre mis manos esta tesis doctoral. Todas aquellas vivencias y largas horas compartidas con ellos, tanto dentro como fuera del laboratorio, quedan plasmadas de alguna forma en las páginas de este trabajo y grabadas en mi memoria y corazón. Gracias por todo.

Asimismo, mostrar mi agradecimiento a todas aquellas personas e instituciones que han colaborado en este trabajo científico: A Marta Nieto Lugilde, Jan Kučera y Petr Koutecký por la realización de los análisis de ploidía en la Universidad del Sur de Bohemia, República Checa. A Susana Rams Sánchez por las recolecciones de muestras en las zonas altas de Sierra Nevada. A Misha Ignatov por la donación de duplicados de *Funaria* procedentes de Rusia de los herbarios MHA y MW. A Bernard Goffinet por el préstamo de muestras americanas del herbario CONN y por compartir con nosotros datos inéditos sobre el genoma de *Funaria hygrometrica* antes de su publicación. A los gestores del Espacio Natural de Sierra Nevada por permitirnos la recolección de muestras en los Parques Nacional y Natural, y a la Consejería de Agua, Agricultura, Ganadería, Pesca y Medio Ambiente de la Comunidad Autónoma de la Región de Murcia por su apoyo en este trabajo.

Echando la mirada atrás, no puedo olvidar a los profesores que me han acompañado desde que comencé la etapa de educación secundaria hasta que obtuve la titulación de grado y máster en la Universidad de Murcia, haciendo especial mención

(aunque no fueron los únicos) a Juanjo Vila, Juan Antonio Delgado y Cecilio Vidal, que siempre creyeron en mí. Gracias.

Me gustaría destacar el amor y el apoyo incondicional de mi familia, que no han permitido que en ningún momento me desmoronase, alentándome a seguir adelante y alcanzar mi objetivo. En especial a mis padres y a mi pareja, que han sufrido y celebrado conmigo cada uno de los momentos de este largo camino. Os quiero mucho.

Quisiera agradecer también a mis amigos su comprensión, apoyo y compañía, tanto los que están cerca como los que se encuentran lejos, sin olvidar a mis compañeros de laboratorio y departamento, tanto en la Universidad de Murcia como en la Universidad de La Laguna. Sin todos vosotros esto habría sido bastante más difícil.

Y en último lugar y no menos importante, agradecer a Pepe Galián (coordinador del Programa de Doctorado en Biodiversidad y Gestión Ambiental) su gran ayuda en los procesos burocráticos de la titulación. Así como al antiguo Ministerio de Economía y Competitividad, a la Universidad de Murcia y al Departamento de Biología Vegetal (Botánica) por darme la oportunidad de realizar este doctorado y formar parte de esta institución universitaria como investigadora en formación.

Muchas gracias a todos.

*Elena*

# INDEX

# I.

# RESUMEN

Un número representativo de especies de briófitos presentan amplios rangos de distribución y pueden ser tratadas como cosmopolitas. Tales rangos se explican a menudo por eventos de dispersión a larga distancia que pueden alcanzar distancias intercontinentales. Evidencias empíricas basadas en estudios filogeográficos han puesto de manifiesto que muchas de estas especies *a priori* muestran señales de especiación alopátrica. Algunos estudios también han mostrado que hay variabilidad genética en poblaciones de la misma especie de briófitos a lo largo de gradientes ambientales, como por ejemplo, altitudinales, cuyas condiciones climáticas exhiben patrones de variación tanto globales como locales, donde la capacidad de adaptación de las especies es fundamental. Los gradientes altitudinales representan, por tanto, escenarios de estudio ideales, porque proporcionan sistemas donde los hábitats y las condiciones ambientales suelen experimentar una rápida renovación. Sin embargo, solo en unos pocos casos de plantas se ha observado variabilidad genética a lo largo de un gradiente altitudinal en poblaciones de la misma especie, como ha ocurrido en el musgo cosmopolita *Funaria hygrometrica* Hedw. En concreto, Magdy Abdallah Awad (2013) y Magdy et al. (2016) detectaron que en esta especie existe variabilidad genética a lo largo del gradiente altitudinal de las montañas de Sierra Nevada (España). Además, datos de secuenciación de ADN, extraídos a partir de marcadores nucleares, cloroplásticos y mitocondriales sugirieron un cierto grado de hibridación entre los linajes encontrados en Sierra Nevada.

Comprender los procesos de especiación no es tarea fácil debido a la multitud de factores y mecanismos evolutivos que entran en juego, tales como el aislamiento geográfico, la adaptación a un nuevo nicho ecológico o la aparición de variaciones genómicas y morfológicas, entre otras. La poliploidía, fenómeno frecuente e importante en la naturaleza, especialmente en plantas, también es un agente importante en la especiación ya que genera una fuerte barrera reproductiva que da origen a nuevas especies.

Tradicionalmente, las especies son identificadas en base a variaciones morfológicas, sin embargo, en algunos casos las alteraciones genéticas y ecológicas no

se correlacionan con las diferencias morfológicas y dan lugar a especies crípticas, dentro de linajes aparentemente uniformes. Mediante la combinación de técnicas moleculares (p. ej., marcadores genéticos, estudio de la ploidía, genética del paisaje, análisis del transcriptoma) con estudios morfológicos, de distribución e incluso ecológicos, se puede arrojar luz sobre los mecanismos de especiación y permitir el reconocimiento de taxones crípticos.

El objetivo general de esta tesis doctoral fue contribuir al conocimiento de la importancia de la geografía y el clima en la configuración de la variación genética y su papel como mecanismos de especiación en briófitos con amplias distribuciones. Se utilizó como modelo de estudio el musgo *F. hygrometrica* y se analizó en cuatro áreas geográficas de España con diferentes regímenes climáticos: las provincias de Asturias y León en el norte, influenciadas por el océano Atlántico, y la provincia de Murcia y el sistema montañoso de Sierra Nevada en el sur, con clima Mediterráneo.

Los objetivos específicos fueron los siguientes: i) conocer si existen diferentes niveles de ploidía en *F. hygrometrica* en las áreas estudiadas; ii) describir los niveles de diferenciación genética entre las poblaciones de *F. hygrometrica* de las cuatro áreas geográficas estudiadas utilizando la técnica *Genotyping by sequencing* (GBS); iii) confirmar la presencia de genotipos diferentes e híbridos que coexisten en el rango de distribución de la especie en Sierra Nevada, como se propuso en un estudio previo (Magdy Abdallah Awad, 2013); la ausencia de híbridos sugeriría que los diferentes linajes son ya demasiado distintos genéticamente para formar un acervo genético común, asumiendo que se puede descartar un evento de colonización reciente resultante de la expansión demográfica; iv) buscar indicios de expansión o contracción demográfica reciente bajo el supuesto de neutralidad selectiva, a fin de comprender la historia demográfica reciente de los linajes inferidos; v) conocer cómo la estructura genética de las poblaciones de *F. hygrometrica* se alinea con las distancias geográficas (IBD) frente a la diferenciación ambiental (IBE), en un esfuerzo por desentrañar el papel de la adaptación en la evolución de los briófitos; vi) comparar la expresión génica entre muestras estudiadas en condiciones control y de estrés térmico a fin de proporcionar información de referencia sobre los mecanismos moleculares que regulan la amplia distribución geográfica de *F. hygrometrica*; vii) evaluar la

correlación entre las presuntas diferencias genéticas, de ploidía y morfológicas bajo una perspectiva sistemática y en el caso de una falta de variación morfológica constatar la existencia de taxones crípticos.

Para abordar estos objetivos se recolectaron 183 muestras de *F. hygrometrica* en 82 localidades repartidas en las cuatro áreas estudiadas. Asimismo se utilizaron varios especímenes de otras especies del género *Funaria* prestados de otros herbarios (*F. aequidens, F. arctica*, *F. flavicans* y *F. polaris*) y otro de *F. hygrometrica* de Chile. Se generaron cultivos *in vitro* de todas las muestras recolectadas y de algunas de las de las otras especies. A partir de esporas aisladas de cada muestra y germinadas en el laboratorio se obtuvieron individuos que fueron usados para los análisis posteriores.

Se estimó la ploidía de una selección de muestras de Asturias, Murcia y Sierra Nevada mediante citometría de flujo, la cual reveló la existencia de cuatro citotipos. El más frecuente (67,4% de las muestras) se correspondió con individuos haploides, otros dos citotipos fueron diploides (16,3% y 11,6%), pero con distinto tamaño del genoma y un cuarto citotipo (4,7%), exclusivo de la provincia de Asturias, tenía más del doble del tamaño que el primero. Las ligeras variaciones entre los grupos de individuos con el mismo nivel de ploidía podrían ser consecuencia de la presencia de cromosomas accesorios, adicionales al cariotipo regular de una especie, como se ha descrito en el caso de varios briófitos.

La secuenciación Sanger de ADN se empleó para conocer las secuencias de tres marcadores moleculares en una selección de muestras de Sierra Nevada: la región ITS del genoma nuclear, el espaciador intergénico *rpl5-rpl16* del genoma mitocondrial y la región *rps3-rpl16* del genoma cloroplástico. La técnica genómica GBS fue empleada para obtener una representación de fragmentos cortos de ADN repartidos aleatoriamente por todo el genoma de todas las muestras recolectadas. Y la técnica RNAseq se utilizó para llevar a cabo un estudio del transcriptoma de determinadas muestras de Sierra Nevada bajo condiciones de estrés térmico en laboratorio y de control.

Los resultados obtenidos con todos los análisis genéticos confirmaron la existencia en *F. hygrometrica* de dos genotipos claramente diferenciados dentro del territorio

estudiado y de manera más específica, a lo largo de un gradiente altitudinal en Sierra Nevada. El genotipo A quedó restringido a individuos de las zonas altas de Sierra Nevada (por encima de los 1600 m s.n.m.) y a unos pocos individuos de Asturias, León y Murcia. Y el genotipo B, más generalista, estuvo presente a lo largo de todo el gradiente altitudinal de Sierra Nevada y fue el predominante en el resto de áreas estudiadas. Las secuencias ITS de *F. hygrometrica* de Sierra Nevada mostraron una mayor proximidad filogenética entre el genotipo A y la muestra de *F. hygrometrica* de Chile, mientras que el genotipo B estuvo más relacionado con las otras especies estudiadas.

No se observaron evidencias de hibridación entre los dos genotipos en las montañas de Sierra Nevada. La ausencia de híbridos podría explicarse por el hecho de que los dos linajes genéticos son demasiado distintos para formar un acervo genético común, siempre que se pueda descartar un evento de colonización reciente. No obstante los análisis filogenéticos mostraron que algunas de las muestras agrupadas en ambos clados obtenidos procedían del mismo punto de muestreo. Esto parece indicar que los dos linajes inferidos pueden ocurrir en la vecindad cercana, potencialmente en contacto, pero que no se cruzan (o casi nunca). Una posible explicación de que no se esté produciendo la recombinación entre individuos con diferentes genotipos, aunque estén muy cerca unos de otros (en la misma localidad), podría ser que la reproducción sexual está ocurriendo por autofecundación.

A partir de los datos GBS se llevaron a cabo análisis de genética de poblaciones para estimar el número de grupos genómicos de individuos. Todos ellos consistentemente apuntaron a la existencia de dos grupos. También se realizaron estudios de genética del paisaje para conocer la importancia de la geografía (aislamiento por distancia, IBD) *versus* la ecología (aislamiento por ambiente, IBE) para explicar esta diferenciación genética dentro de las cuatro áreas estudiadas. Según los análisis demográficos realizados, la colonización y expansión recientes en Sierra Nevada no están respaldadas por el conjunto de datos GBS. Los análisis de resistencia del paisaje revelaron la no existencia de aislamiento ambiental en poblaciones de Asturias, León y Murcia, mientras que a lo largo del gradiente altitudinal de Sierra

Nevada la variable climática isotermalidad constituyó el modelo que mejor explicó la estructura genética observada en este sistema montañoso.

Los resultados del análisis del transcriptoma sugirieron que podría haber diferencias en la tolerancia al estrés por calor de los genotipos A y B. De manera sorprendente las muestras estudiadas de *F. aequidens*, *F. arctica* y *F. polaris* evidenciaron pocas diferencias en la expresión génica entre ellas, y por tanto hubo una mayor diferencia de expresión génica intraespecífica que interespecífica. Además, estas especies parecen estar un poco más cerca de los individuos de *F. hygrometrica* con genotipo B, como también apuntaron los resultados de secuenciación Sanger. El análisis de ontología génica (GO) indicó que una gran parte de las diferencias observadas entre los genotipos A y B en condiciones de control correspondían al sistema de traducción. Especialmente interesante es el hecho de que la categoría GO denominada "Unión a proteína desplegada" se expresa diferencialmente bajo condiciones de estrés térmico pero que esta expresión diferencial es incluso significativa en el genotipo B bajo estrés térmico *versus* genotipo A sometido a estrés térmico. Muchas de las proteínas descritas en esta categoría pertenecen a las *heat shock proteins* y en condiciones normales tienen el papel de conducir el plegamiento de las proteínas para obtener la conformación correcta. En condiciones de estrés térmico las proteínas pueden perder su forma funcional y las *heat shock proteins* pueden ayudar a que vuelvan a funcionar correctamente, o, si esto ya no es posible, marcarlas para su degradación.

Asimismo se llevó a cabo un estudio biométrico de plantas de Sierra Nevada, en el que se incluyeron también las otras especies de *Funaria* y el espécimen de *F. hygrometrica* de Chile. Los resultados no mostraron una diferenciación morfológica significativa entre los individuos de Sierra Nevada con diferente genotipo ni de distinta altitud. Tampoco se encontró una correlación clara de los datos morfológicos con los datos de ploidía ni con los genéticos. Esta falta de variación morfológica apunta a que los dos genotipos podrían ser especies crípticas, que parecen tener una evolución o divergencia más rápida a nivel molecular que a nivel morfológico. Todo ello sugiere que las poblaciones de *F. hygrometrica* de Sierra Nevada representan un caso de especiación incipiente.

Dado que las dos especies crípticas de *F. hygrometrica* aparentemente coexisten y compiten al menos parcialmente por los mismos recursos y muestran hábitats superpuestos, se realizó una simulación en R de sus patrones de distribución. El principio de exclusión competitiva declara que dos especies no pueden ocupar el mismo nicho ecológico al mismo tiempo de forma estable, aunque no se sabe si se puede aplicar de forma estricta en este caso, ya que el conocimiento de las relaciones competitivas entre especies de briófitos es poco conocido. No obstante, la simulación realizada mostró que dos especies independientes pueden compartir el mismo hábitat, incluso si compiten fuertemente por los mismos recursos y una de ellas tiene cierta ventaja selectiva en ese hábitat. La condición es que llegue a este hábitat un número suficiente de migrantes desde otras localidades donde se invierte la ventaja selectiva. Este requisito no sería ningún obstáculo para los briófitos en general dada su elevada capacidad de dispersión a larga distancia, ni para *F. hygrometrica* en particular, lo que es apoyado por la mayor proximidad filogenética de la muestra de Chile con los individuos de Sierra Nevada con genotipo A y el pequeño tamaño de sus esporas.

# II.

# ABSTRACT

A representative number of bryophyte species exhibit wide ranges of distribution and can be treated as cosmopolitan. Such ranges are often explained by long-distance dispersal events that can reach intercontinental distances. Empirical evidence based on phylogeographic studies has shown that many of these species *a priori* show signs of allopatric speciation. Genetic variability in populations of the same bryophyte species along environmental gradients is also shown in some studies, for example, altitudinal gradients, whose climatic conditions exhibit both global and local patterns of variation, where the adaptive capacity of the species is fundamental. Therefore, altitudinal gradients represent ideal study scenarios, because they provide systems where habitats and environmental conditions often undergo rapid renewal. However, only in a few cases, genetic variability along an altitudinal gradient is observed in populations of the same plant species, as in the cosmopolitan moss *Funaria hygrometrica* Hedw. Specifically, Magdy Abdallah Awad (2013) and Magdy et al. (2016) detected that in this species there is genetic variability along the gradient of altitude in the Sierra Nevada Mountains (Spain). Besides, DNA sequencing data, extracted from nuclear, chloroplastic and mitochondrial markers, suggested a certain degree of hybridization between the lineages found in the Sierra Nevada Mountains.

Understanding the processes of speciation is not an easy task, due to the multitude of factors and evolutionary mechanisms that come into play, such as geographic isolation, adaptation to a new ecological niche or the appearance of genomic and morphological variations, among others. Polyploidy, a frequent and important phenomenon in nature, especially in plants, is also an important agent in speciation since it generates a strong reproductive barrier that gives rise to new species.

Traditionally, species are identified based on morphological variations, nevertheless, in some cases genetic and ecological alterations are not correlated with morphological differences and give rise to cryptic species, within apparently uniform lineages. By combining molecular techniques (e.g., genetic markers, ploidy study, landscape genetics, transcriptome analysis) with morphological, distributional and

even ecological studies, it is possible to shed light on the mechanisms of speciation and allow the recognition of cryptic taxa.

The general objective of this Ph.D. thesis was to contribute to the knowledge on the importance of geography and climate in shaping genetic variation and their role as speciation mechanisms in bryophytes that exhibit wide distributions. The cosmopolitan moss *F. hygrometrica* was used as a study model in four geographical areas of Spain with different climatic regimes: Asturias and Leon provinces in the north, influenced by the Atlantic Ocean, and Murcia province and the Sierra Nevada Mountains in the south, with Mediterranean climate.

The specific objectives were the followings: i) to know whether there exist different ploidy levels in *F. hygrometrica* in the studied areas; ii) to describe the levels of genetic differentiation among *F. hygrometrica* populations from the four geographical areas studied using the Genotyping by sequencing (GBS) technique; iii) to confirm the presence of different and hybrid genotypes coexisting in the distribution range of the species across the Sierra Nevada Mountains, as proposed in a previous study (Magdy Abdallah Awad, 2013); the absence of hybrids would suggest that the different lineages are genetically already too distinct to form a common gene pool, assuming that a recent colonization event resulting from demographic expansion can be ruled out; iv) to look for signatures of recent demographic expansion or contraction under the assumption of selective neutrality, in order to understand the recent demographic history of the lineages inferred; v) to know how the genetic structuring of the populations of *F. hygrometrica* is aligned with geographical distances (IBD) compared with environmental differentiation (IBE), in an effort to disentangle the role of adaptation in bryophyte evolution; vi) to compare gene expression among samples studied under natural conditions and thermal stress in order to provide baseline information on the molecular mechanisms that regulate the broad geographic distribution of *F. hygrometrica*; vii) to evaluate the presumed genetic, ploidy and morphological differences from a systematic perspective and, in the case of a lack of morphological variation, verify the existence of cryptic taxa.

To address these objectives, 183 samples of *F. hygrometrica* were collected in 82 localities distributed in the four areas studied. Likewise, several specimens of other

species of the genus *Funaria* loaned from other herbaria (*F. aequidens*, *F. arctica*, *F. flavicans* and *F. polaris*) and another of *F. hygrometrica* from Chile were used. *In vitro* cultures were generated from all the collected samples and some of the other species. Individuals were obtained from spores isolated from each sample and germinated in the laboratory, which were used for subsequent analyses.

The ploidy of a selection of samples from Asturias, Murcia and the Sierra Nevada Mountains was estimated by flow cytometry, which revealed the existence of four cytotypes. The most frequent (67.4% of the samples) corresponded to haploid individuals, two other cytotypes were diploid (16.3% and 11.6%), but with different genome sizes and a fourth cytotype (4.7%), exclusive to the province of Asturias, it was more than twice the size of the first. Slight variations between groups of individuals with the same level of ploidy could be a consequence of the presence of accessory chromosomes, in addition to the regular karyotype of a species, as has been described in the case of several bryophytes.

Sanger DNA sequencing was used to determine the sequences of three molecular markers in a selection of samples from the Sierra Nevada Mountains: the ITS region of the nuclear genome, the *rpl5-rpl16* intergenic spacer of the mitochondrial genome and the *rps3-rpl16* region of the chloroplast genome. The GBS genomic technique was used to obtain a representation of short DNA fragments randomly distributed throughout the genome of all the samples collected. And the RNAseq technique was used to carry out a study of the transcriptome of certain samples from the Sierra Nevada Mountains under thermal stress and control conditions in laboratory.

The results obtained with all the genetic analyses confirmed the existence of two clearly differentiated genotypes in *F. hygrometrica* within the territory studied and, more specifically, along an gradient of altitude in the Sierra Nevada Mountains. Genotype A was restricted to individuals from the highlands of the Sierra Nevada Mountains (above 1600 m a.s.l.) and to a few individuals from Asturias, Leon and Murcia. And genotype B, more generalist, was present throughout the entire altitude gradient of the Sierra Nevada Mountains and was the predominant one in the rest of the areas studied. The ITS sequences of *F. hygrometrica* from the Sierra Nevada Mountains showed greater phylogenetic proximity between genotype A and the

sample of *F. hygrometrica* from Chile, while genotype B was more related to the other species studied.

Hybridization between both genotypes could not be confirmed in the Sierra Nevada Mountains. The absence of hybrids could be explained by the fact that the two genetic lineages are too different to form a common gene pool, provided that a recent colonization event can be ruled out. However, the phylogenetic analyses showed that some of the samples grouped in both clades obtained came from the same sampling point. This seems to indicate that the two inferred lineages may occur in close vicinity, potentially in contact, but that they do not (or hardly so) cross. One possibility to explain that recombination is not taking place between individuals with different genotypes, although they are very close to each other (in the same locality), is that the predominant way of sexual reproduction might be self-fertilization.

Population genetic analyses were carried out from GBS data to estimate the number of genomic groups of individuals. All of them consistently pointed to the existence of two groups. Landscape genetics studies were also conducted to determine the importance of geography (isolation by distance, IBD) *versus* ecology (isolation by environment, IBE) to explain this genetic differentiation within the four areas studied. Based on demographic analyses performed, a recent colonization and expansion in the Sierra Nevada Mountains is not supported by the GBS data set. The landscape resistance analyses revealed the non-existence of environmental isolation in populations of Asturias, Leon and Murcia, while along the gradient of altitude in the Sierra Nevada Mountains the climatic variable isothermality constituted the model that best explained the genetic structure observed in this mountain range.

The results of the transcriptome analysis suggested that there could be differences in thermal stress tolerance of genotypes A and B. Surprisingly, the studied samples of *F. aequidens*, *F. arctica* and *F. polaris* showed few differences in gene expression between them, and therefore there was a greater difference in *F. hygrometrica* intra-specific than inter-specific gene expression. Furthermore, these species seem to be a little closer to the individuals of *F. hygrometrica* with genotype B, as also pointed out by the Sanger sequencing results. The analysis of gene ontology (GO) indicated that a large part of the differences observed between genotypes A and B under control

conditions corresponded to the translation system. Especially interesting is the fact that the GO category called "Unfolded protein binding" is differentially expressed under thermal stress conditions but that this differential expression is even significant in genotype B under heat stress *versus* genotype A under heat stress. Many of the proteins described in this category belong to the heat shock proteins and under normal conditions have the role of driving the folding of the proteins to obtain the correct conformation. Under conditions of thermal stress, proteins can lose their functional form and heat shock proteins can help them to function properly again, or, if this is no longer possible, mark them for degradation.

Likewise, a biometric study of plants from the Sierra Nevada Mountains was carried out, which also included the other species of *Funaria* and the specimen of *F. hygrometrica* from Chile. The results did not show a significant morphological differentiation between the Sierra Nevada individuals with different genotypes or of different altitude. Neither a clear correlation of morphological data with ploidy data or genetic data was found. This lack of morphological variation point to that the two genotypes could be cryptic species, which seem to have a faster evolution or divergence at the molecular level than at the morphological level. All this suggests that the populations of *F. hygrometrica* from the Sierra Nevada Mountains represent a case of incipient speciation.

Since the two cryptic species of *F. hygrometrica* apparently coexist and compete at least partially for the same resources and show overlapping habitats, an R simulation of their distribution patterns was performed. The principle of competitive exclusion states that two species cannot occupy the same ecological niche at the same time in a stable way, although it is not known if it can be applied strictly in this case, since the knowledge of the competitive relationships between bryophyte species is little known. However, the simulation carried out showed that two independent species can share the same habitat, even if they compete strongly for the same resources and one of them has a certain selective advantage in that habitat. The condition is that a sufficient number of migrants arrive in this habitat from other localities where the selective advantage is inverted. This requirement would not be an obstacle for bryophytes in general due to their high long-distance dispersal

capacity, nor for *F. hygrometrica* in particular, which is supported by the greater phylogenetic proximity of the sample from Chile with individuals from the Sierra Nevada Mountains with genotype A and the small size of its spores.

# III.

# INTRODUCTION

## MAIN BRYOPHYTE CHARACTERISTICS

Bryophytes are small land plants, whose life cycle has a haploid dominant generation (gametophyte), and a diploid spore producing sporophyte, which remains attached to the gametophyte during the whole life span (Vanderpoorten & Goffinet, 2009). Liverworts, hornworts and mosses are traditionally included in the so-called bryophyte group (but see discussions in Hedges et al., 2018 and Morris et al., 2018) and there are around 20,000 species worldwide (Medina et al., 2011; The Plant List, 2013; Patiño & Vanderpoorten, 2018).

Their gametophyte is green and photoautotrophic, and although they are considered as non vascular plants (i.e. lacking lignin, tracheids and xylem vessels), the internal transport of nutrients and water is carried out by specialized tissues (Ligrone et al., 2000; Goffinet et al., 2009; Glime, 2017). Their poikilohydric nature makes them dependent on ambient humidity, because their internal water content is equilibrated rapidly to the water potential of the environment, due to the lack of mechanisms and structures to prevent desiccation (Medina et al., 2011; Patiño & Vanderpoorten, 2018). However, in general, their tolerance to desiccation (and freezing) allows them to maintain latent life (quiescence) after drying at the cellular level, and to recover a normal metabolism upon rehydration (Proctor, 2000; Mishler, 2001; Oliver et al., 2005).

Unlike most of the vascular plants, bryophytes are rarely found as single isolated individuals (i.e. shoots) in nature, normally growing together to form colonies or assemblage of individuals with a specific growth type, which gives them the great advantage of better retaining water inter-individually and being able to cope with unfavourable conditions (Richardson, 1981; Glime, 2017). Many species thus form cushions, turfs or tapestries (mats) on rocks, wood, or directly on the ground (Smith, 1982; Crandall-Stotler & Bartholomew-Began, 2007).

These and other characteristics enhance their ability to adapt to differents climates and to changes in environmental conditions (Vanderpoorten & Goffinet,

2009; Medina et al., 2011). Some metabolic adaptions allow them to survive under adverse conditions of temperature and dryness, such as through the production of substances that offer them a protective mechanism, at the expense of restricting growth rate (Vanderpoorten & Goffinet, 2009). The production of these substances can be triggered by abscisic acid (ABA), whose endogenous level is increased six-fold during the slow drying process, preparing cells to survive to fast drying processes in the future (Werner et al., 1991). Sometimes, molecular and genetic differences between bryophyte populations can allow adaptation to certain environmental conditions (Szövényi et al., 2009), both natural and human-influenced conditions (Vanderpoorten & Goffinet, 2009). For example, Shaw (1988, 1994, 1999) found genetic variations associated with heavy metal tolerance within moss populations, where a small number of loci could be involved in the tolerance mechanism (McDaniel, 2009). There are even populations with genome duplication, which seems to give them some advantage to survive under a specific environment, as occurs in some species of the genus *Sphagnum* L. (Glime & Bisang, 2017).

## BRYOPHYTE BIOGEOGRAPHY

Despite their dependence on water, bryophytes are present in all environments in which vascular plants may appear, with the exception of marine aquatic media (Richardson, 1981). They can be found growing (depending on the genus and species) from hot and dry deserts to cold and wet mountains, likewise from sea level to alpine highlands, being widespread and abundant across rainforests (Frahm, 2008; Vanderpoorten & Goffinet, 2009; Medina et al., 2011).

Dispersal is achieved by means of diaspores, which can be sexual spores or asexually produced structures, originated from gametophyte fragments and/or specialized vegetative structures (Frahm, 2008; Vanderpoorten & Goffinet, 2009; Medina et al., 2011). In the different parts of the thallus (in thallose liverworts and hornworts), caulidia, phyllidia or rhizoids (in foliose liverworts and mosses), the formation of propagules can give rise to a new complete plant and identical to its parent individual (Frey & Kürschner, 2011). Diaspores are dispersed at even

intercontinental distances mainly by air currents, although there are other mechanisms such as water or animal transportation, being able to play an important role in the distribution pattern of certain species (Muñoz et al., 2004; Frahm, 2008; Vanderpoorten & Goffinet, 2009; Medina et al., 2011; Patiño & Vanderpoorten, 2018). Within the bryophytes, there are monoicous species that usually present sporophytes; and dioicous species, that rarely reproduce sexually, perhaps due to the difficulty of finding a male and a female plant relatively close (Vanderpoorten & Goffinet, 2009; Patiño & Vanderpoorten, 2018). Despite this, in those cases where spore dispersal is not possible, asexual propagule dispersion seems equally viable and successful (van Zanten & Pócs, 1981; Frahm, 2008; Vanderpoorten & Goffinet, 2009; Laenen et al. 2016). In this respect, whereas spores might significantly contribute to long-distance dispersal (LDD), specialized vegetative propagules are, due to the larger size and lack of release mechanisms, thought to mainly contribute to short-distance dispersal and population persistence (Patiño & Vanderpoorten, 2018). Laenen et al. (2016), however, proposed the key role of specialized vegetative propagulae in shaping distribution ranges in liverworts, based on phylogenetic evidence. These authors found a significant correlation between the production of asexual diaspores and the size of geographical ranges.

Many bryophyte species, indeed, exhibit wide distribution ranges, and can be found in more than one continent or even in all of them (cosmopolitan) (Schuster, 1983; Frahm, 2008; Vanderpoorten & Goffinet, 2009). In fact, more than 75% of the bryophyte families can be found in both hemispheres (Tan & Pócs, 2000). However, the number of really ubiquitous species is not very high, perhaps below 1% in bryophytes (Frahm, 2008; Medina et al., 2011); that is the case, for example, of the moss species *Bryum argenteum* Hedw., *Ceratodon purpureus* (Hedw.) Brid. and *Funaria hygrometrica* Hedw. (Schuster, 1983; Ochyra et al., 2008; Vanderpoorten & Goffinet, 2009).

In turn, there are also species with a restricted distribution or even considered endemic to discrete areas (Frahm, 2008; Vanderpoorten & Goffinet, 2009). Unlike flowering plants, bryophytes display low levels of endemism, as well as differing slightly in their worldwide patterns of endemic hot-spots (Myers et al., 2000;

Vanderpoorten & Hallingbäck, 2009; Patiño et al., 2014). For example, the Mediterranean areas are considered very important hot-spots for vascular plants, however, this is not observed in bryophytes; in contrast, several temperate areas, such as New Zeland, Patagonia and Japan, and some tropical areas such as part of the Malesio-Indonesian area, among others, stand out as areas with high levels of bryophyte species endemism (higher than 15%) (Vanderpoorten & Hallingbäck, 2009; Patiño & Vanderpoorten, 2018). Despite this, in many cases, levels of endemism in angiosperms are more than double that in bryoflora, a tendency observed above all on many oceanic and continental islands (Vanderpoorten & Hallingbäck, 2009; Patiño & Vanderpoorten, 2018). For example, the percentage of endemicity in the flora of the Canary Islands (Spain) reaches only the 1.5% in bryophytes in comparison with the 40% in angiosperms (Patiño et al., 2014), while in peninsular and balearic Spain this percentage is only a scant 0.5% in bryoflora (Infante et al., 2017) and it reaches 15.5% in vascular plants (Aedo et al., 2013). Galapagos Islands are another example of high endemicity level, that reach the 50% in angiosperms but only the 13% in liverworts and 6% in mosses (Vanderpoorten et al., 2010; Medina et al., 2011). Some cases of Canarian endemic bryophyte species are *Grimmia curviseta* Bouman from Tenerife and La Palma islands, or *Orthotrichum handiense* F. Lara, Garilleti & Mazimpaka from Fuerteventura island (Garilleti & Albertos, 2012; Patiño et al., 2013).

## BRYOPHYTE ECOLOGY AND ELEVATIONAL GRADIENTS

Environmental gradients impose a strong variation on the ecological and biological forces acting locally as filters of different nature (Bello et al., 2013; Abbott, 2017; Tylianakis & Morris, 2017; Riesch et al., 2018). One of the most studied and complex gradients is the elevation, whose climatic conditions exhibit both global and local patterns of variation (Ohsawa & Ide, 2008). Elevational gradients represent thus ideal study arenas, because they support systems where habitats and enviornmental conditions typically undergo a rapid turnover (Nogués-Bravo et al., 2008). A mounting number of studies have found that changes in habitats along elevational gradients shape diversity, community composition, species ranges and evolutionary

rates (Ohsawa & Ide, 2008; Pisa et al., 2013; Morueta-Holme et al., 2015), but very often neglect patterns of genetic diversity within species that are tempting of being treated as signatures of adaptation and ongoing speciation (Shaw, 2001; Jiggins, 2019).

In various bryological studies, variability in species diversity has been observed along an altitudinal gradient, being in several cases greater at high altitudes compared to the species diversity found in lowlands, in other words, species diversity increases with altitude (e.g. Ah-Peng et al., 2007; Sim-Sim et al., 2015). In turn, it is just a few cases where some genetic variability has also been observed along that gradient in populations of the same species (e.g. McDaniel, 2009). The data of Magdy et al. (2016) suggest a correlation of genetic data of the moss *F. hygrometrica* with environmental conditions along the gradient of altitude in the Sierra Nevada Mountains, in the south of Spain that may be due to ecological specialization. Moreover, the sequence data of Madgy Abdallah Awad (2013) covering nuclear, chloroplast and mitochondrial markers suggest a certain degree of hybridization between the present genetic lineages of *F. hygrometrica* found in Sierra Nevada Mountains. Anderson (1980) points out that $F_1$ gametophyte hybrids do not exist and that the gametophytes produced by an $F_1$ sporophyte show $F_2$ segregational variation (and are not uniform like $F_1$ sporophytes). In his paper he suggests to use the term recombinant for the gametophytes produced after an hybridization event. In this thesis the term hybrid is used in a broader sense referring to all the offspring generations after an initial hybridization event. Also *B. argenteum* shows genetic variability based on elevation or environmental conditions, where one genotype was found above 2000 m a.s.l. in Sierra Nevada Mountains, different from lowland populations and similar to known sequences of populations occuring in the Antartic region (Pisa et al., 2013, 2014). In addition, the genus *Ceratodon* Brid. shows a clear speciation process in the same mountain range, where a new species was recently described (*Ceratodon amazonum* Nieto-Lugilde, O. Werner, S.F. McDaniel & Ros), which is found above 1300 m a.s.l. approximately. Specifically, *C. amazonum* differs at a genetic and morphological level from the cosmopolitan *C. purpureus* (Nieto-Lugilde et al., 2018a, 2018b). A greater genetic richness of several species of bryophytes has also been observed at high altitudes in other regions, such as on Madeira Island (Sim-Sim et al., 2015).

## SPECIATION AND CRYPTIC SPECIES IN BRYOPHYTES

Understanding the processes of speciation is not an easy task, due to the multitude of factors and evolutionary mechanisms that come into play in the origin of a new species (Turelli et al., 2001; Seehausen et al., 2014; Jiggins, 2019). The geographic isolation of a population (allopatric speciation) reduces or prevents gene flow with others (reproductive isolation), which allows, over time, genetic variations with respect to the original population that can finally generate additional reproductive barriers (Turelli et al., 2001; Coyne & Orr, 2004; Seehausen et al., 2014). Furthermore, adaptation to a new ecological niche (peripatric or parapatric speciation) or even the appearance of genomic, ethological or other variations (sympatric speciation) can also give rise to a new species (Turelli et al., 2001; Lenormand, 2002; Seehausen et al., 2014). The study of speciation in bryophytes is interesting, because their diversification seems to depend more on genetic, ecological and demographic factors, and less on geographic factors (McDaniel et al., 2010).

Polyploidy is a phenomenom frequent and important in nature, especially in plants, generating a strong reproductive barrier at the beginning, which give origin to new species (Grant, 1981; Newton, 1984; Otto & Whitton, 2000; Turelli et al., 2001; Soltis et al., 2004, 2015; Marques et al., 2018). In bryophytes the dominant generation is usually haploid (n), but it is not uncommon to find cases of polyploidy (two or more complete sets of chromosomes in each cell) (Grant, 1981; Newton, 1984; Otto & Whitton, 2000; Soltis & Soltis, 2009). When polyploidy comes from the same species, for example by duplicating its genome, it is named "autoploidy" or “autopolyploidy”, which is a common phenomenon in bryophytes (Soltis et al., 2004; Perley & Jesson, 2015; Glime & Bisang, 2017). But there is also polyploidy that results from the combination of two genomes of different taxa through hybridization, called “allopolyploidy” (Soltis et al., 2004; Glime & Bisang, 2017). Within mosses, the genus *Sphagnum* is an example, where it is common to find polyploid species, either by duplication of their own genome or by hybridization of males and females of two

different species of the same genus (secondary contact) (Ricca et al., 2011; Glime & Bisang, 2017).

Traditionally, species are identified based on morphological variations, nevertheless, in some cases the genetic and ecological variations in bryophytes are not correlated with the accumulation of morphological differences, and as a consequence there are probably hidden taxa within morphologically uniform species (Shaw, 2001; Hedenäs et al., 2014; Vanderpoorten & Shaw, 2014; Struck et al., 2018; Renner, 2020). This is the case of cryptic or sibling species, characterized by a complete or nearly complete lack of differences at morphological level (Shaw, 2001; Bączkiewicz et al., 2017; Renner, 2020). The recognition of cryptic species is possible by means of molecular techniques, such as the use of genetic markers, study of ploidy, etc. (e. g. McDaniel & Shaw, 2003; Bickford et al., 2007; Hedenäs & Eldenäs, 2007; Bączkiewicz et al., 2017; Manukjanová et al., 2018; Hedenäs, 2020; Renner, 2020). Although bryophytes are less speciose than vascular plants, perhaps the species number is, to some extent, underestimated by the existence of these cryptic species (Medina et al., 2012; Patiño & Vanderpoorten, 2018).

## THE MOSS GENUS *FUNARIA*

The genus *Funaria* Hedw. belongs to the family Funariaceae. Its name is derived from the Latin word *"funis"* which means "rope", alluding to the cord-like seta which twist and curl in response to changes in humidity (McIntosh, 2007; Martin, 2016). According to Fife (1985), the morphological diagnostic characteres of *Funaria* genus are the capules sulcate when dry due to vertical bands and strongly cuneate cells alternating with firm-walled, non cuneate cells; the double peristome; teeth strongly sigmoid and fused at apices by a lattice and annulus compound and revoluble.

Phylogenetic inferences from 10 loci from all three genomic compartments carried out by Y. Liu et al. (2012a) confirmed the monophyly of the genus and the systematic concept proposed by Fife (1985) based on the sporophytic characters. On the contrary, this was not the case with other genera such as *Entosthodon* Schwägr.,

*Physcomitrella* Bruch & Schimp., and *Physcomitrium* (Brid.) Brid., for which these authors suggested a polyphyletic nature.

According to Fife & Seppelt (2012) this cosmopolitan genus includes c. 200 species worldwide. But there is no updated comprehensive taxonomical revision of *Funaria* at global scale. In Europe, according to Hodgetts et al. (2020), four species are recognized (*F. aequidens* Lindb. ex Broth., *F. arctica* (Berggr.) Kindb., *F. hygrometrica* and *F. microstoma* Bruch ex Schimp.), and up to 17 species are reported from the African flora (O'Shea, 2006; Ros et al., 2013a), having in common with the European continent only *F. hygrometrica* and *F. microstoma*). The Russian flora (Ignatova & Fedosov, 2017) shares the four European species plus *F. polaris* Bryhn, and in the rest of the Asian continent, only considering the Chinese flora (http://legacy.tropicos.org/Project/MFC), three more species are present (*F. attenuata* (Dicks.) Lindb., *F. discelioides* Müll. Hal. and *F. japonica* Broth.). In North America (McIntosh, 2007) eight species are accepted, including all the European except *F. aequidens*, the non European *F. flavicans* Michx. and three exclusive to the USA (*F. apiculatopilosa* Cardot*, F. serrata* Brid. and *F. americana* Lindb.). At least six more species are accepted in the Chilean Moosflora (http://www.mobot.org/MOBOT/moss/Chile/speclist.htm), some of them not known from other continents, such as *F. chilensis* (Thér.) Thér., *F. commixta* Thér., *F. costesii* Thér., *F. macrocarpa* (Schimp.) Zand., and *F. porteri* Thér. With certainty the American list of *Funaria* species is much wider. In contrast, in Australia only *F. hygrometrica* and *F. microstoma* are reported (Fife & Seppelt, 2012).

Therefore it is clearly deduced that the most widespread and abundant species in the genus is *F. hygrometrica* (Figs. 1 and 2), which grows on cracks of rock and pavement, shady and moist soils, damp walls, recently burnt sites and disturbed soils (Casas et al., 2006; McIntosh, 2007; Frahm, 2008; Brugués & Ruiz, 2010; Martin, 2016). In summary, this species seems to exhibit a certain affinity for natural and anthropogenic disturbed areas, but perhaps more important for the overreaching goal of the present thesis, *F. hygrometrica* can cover broad environmental gradients, for instance in elevation (R.M. Ros & O. Werner, personal observation).

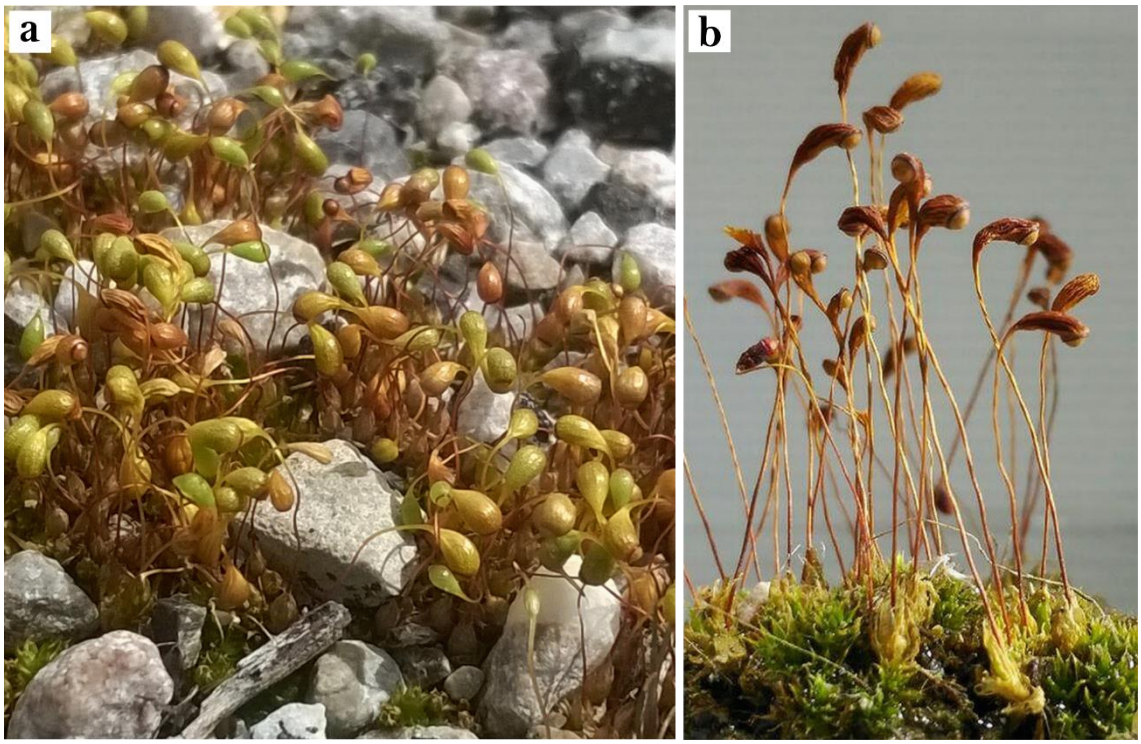



**Figure 1.** *Funaria hygrometrica*. **a:** Plants forming a grass between gravel in which capsules begin to mature. **b:** Plants with mature and dry closed capsules. Source: own elaboration.

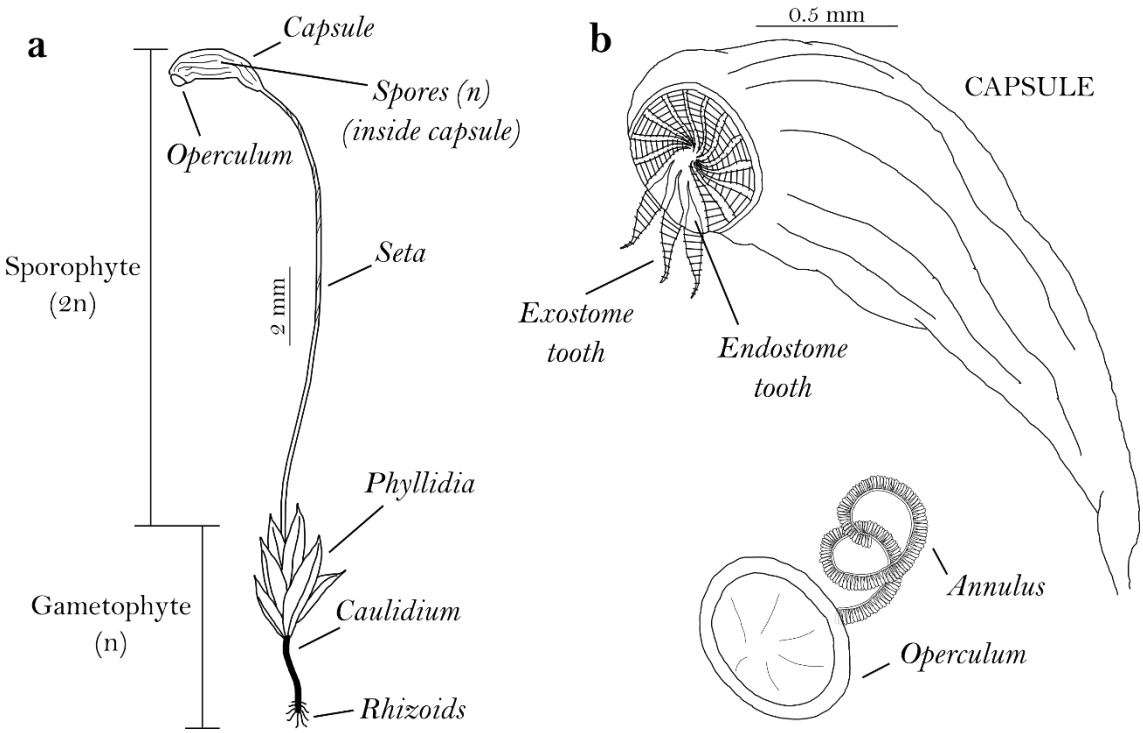



**Figure 2.** *Funaria hygrometrica*. **a**: Complete plant showing gametophyte and sporophyte generations with indication of the different parts of each one. **b**: Detail of the capsule. Source: own elaboration.

*Funaria hygrometrica* can be characterized as follows (Casas et al., 2006; McIntosh, 2007; Brugués & Ruiz, 2010; Ignatova & Fedosov, 2017): plants ephemeral, gregarious or scattered. Rhizoids light brown or reddish. Caulidium red brownish, 4-10 mm high. Phyllidia light green to yellowish green, grouped in the upper part of the caulidium, 2-4 mm long, smaller proximally, oblong-ovate to broadly obovate, deeply concave; apex acute to short-acuminate; margins entire or weakly serrulate or dentate distally, plain to incurvate; nerve ending in or below apex to short-excurrent; distal laminal cells lax, thin-walled and inflated, hexagonal or oblong-hexagonal, 30-50 μm wide. Autoicous. Seta usually (12-)20-45(-80) mm long, yellowish orange, slender and flexuose, curved or cygneous, usually hygroscopic and twisted. Capsule stegocarpous, exerted, 2.0-3.5 mm long, asymmetric, pyriform, curved to straight, reddish-golden when mature, sulcate when mature and dry, from horizontal to pendent or inclined or nearly erect with respect to the seta, with a distint neck, cryptopore stomata in the neck, mouth very oblique. Annulus revoluble. Peristome double, exostome of 16 sigmoid teeth brown, lanceolate, acuminate, papillose, striate, strongly trabeculate, appendiculate and fused at apices; endostome of 16 lanceolate teeth, hyaline, opposite to the exostome teeth, cohesive at the base, about 2/3 as long as the exostome teeth papillose or slightly papillose-striated. Operculum slightly convex. Calyptra 4-7 mm long, deciduous, cucullate and inflated, smooth. Spores 12-23 μm in diameter, from yellowish to light brownish, spherical to ovoid, finely papillose.

*Funaria hygrometrica* is widely employed as model of study in plant physiology, biochemical synthesis and molecular ecology, among others. Its popular use in experimental studies can be explained by its ubiquity and easy recognition in the field, as well as higher growth rates than other species in laboratory conditions (Nakosteen & Hughes, 1978; Frahm, 2008; Proctor, 2009). Among the available examples, it is worth mentioning studies dealing with a broad variety of topics such as: germination and growth of *in vitro* cultivated mosses (Kofler, 1959); hormonal regulation of protonema development (Lehnert & Bopp, 1983); metal tolerance (Shaw, 1988, 1990); desiccation tolerance in mosses (Werner et al., 1991; Pressel & Duckett, 2010); generation-biased gene expression in land plants (Szövényi et al., 2010); or genetic

variability in populations of cosmopolitan mosses (Magdy et al., 2016). The complete sequence of its genome was recently published (Kirbis et al., 2020).

## MOLECULAR ECOLOGY OF BRYOPHYTES

Thanks to advances in molecular genetic technologies in recent years, the field of molecular ecology makes use of molecular tools and approaches to solve ecological and evolutionary problems (Eguiarte et al., 2007). Among the molecular compartiments in molecular ecology, nuclear DNA stands out as the most widely used genetic data source, but also RNA, mitochondrial DNA and chloroplast DNA) are studied in many occasions (Freeland, 2020). For example, the DNA sequences and genome size of several moss populations of the genus *Ceratodon*, together with morphometric data, were employed to discover a new species (*C. amazonum*) in the Sierra Nevada Mountains (Nieto-Lugilde et al., 2018a, 2018b). On the other side, the hydration-dehydration-rehydration transcriptomic data in *B. argenteum* were analysed to establish a transcriptomic atlas of desiccation tolerance, revealing protein domains that appear to be directly involved with this key process (Gao et al., 2017).

### Landscape genetics

Landscape genetics was formally defined by Manel et al. (2003), who stated that this discipline "aims to provide information about the interaction between landscape features and microevolutionary processes, such as gene flow, genetic drift and selection". Since then, it has become evident that studying how landscape characteristics (geographical and environmental) structure the genetic variation of a population (Balkenhol et al., 2015) is important to understand evolutionary processes, mainly gene flow and adaptation. One branch of landscape genetics is landscape genomics, which uses representations of the genome of each individual in a population (employing a large number of genotyped molecular markers per individual) to study the influence of different landscape factors on adaptive genetic diversity in a population or even in a species (Manel & Holderegger, 2013; Jia et al., 2020). Various methods can help us to determine which factor(s) could best explain

geographic patterns of genetic structure. In addition, genetic distance can be employed to estimate gene flow between demes (biological populations), and can be aligned with geographic distance (isolation by distance, IBD; Wright, 1943), generally assuming landscape homogeneity. The approach isolation by resistance (IBR) takes into account landscape heterogeneity, using least-cost paths analyses (McRae, 2006; McRae & Beier, 2007). Specifically, isolation by environment (IBE) is a more recent approach, which predicts the relationship between spatial variation in gene flow and environmental heterogeneity (Sexton et al., 2013; Wang & Bradburd, 2014).

### DNA sequencing techniques

During the mid-1970s, Frederick Sanger and collaborators developed a DNA sequencing method based on the selective incorporation of chain-terminating dideoxynucleotides using a purified DNA polymerase during *in vitro* DNA replication (Sanger & Coulson, 1975; Sanger et al., 1977). These modified di-deoxynucleotide triphosphates (ddNTPs) lack the 3' hydroxyl (OH) group needed to form the phosphodiester bond between one nucleotide and the next during DNA strand synthesis. By performing the polymerization separately for each of the ddNTPs (ddATP, ddTTP, ddGTP, ddCTP), DNA fragments of different lengths are obtained, and can be separated by size using gel or capillary tube electrophoresis. This allows to identify the order of the nucleotides in a DNA sequence. Because Sanger sequencing was improved and automated years later, this method has been widely used for more than 30 years (Schuster, 2008; Heather & Chain, 2016).

Between the 1990s and the 2000s, several new methods for DNA sequencing were developed, especially for large-scale and automated genome analyses, replacing the Sanger method. The "Next-Generation" sequencing (NGS) techniques perform reading of millions of small DNA fragments in parallel (massive parallel sequencing), greatly increasing the volume of information to obtain and reducing the time employed (L. Liu et al., 2012; Behjati & Tarpey, 2013; Heather & Chain, 2016; Shendure et al., 2017). Today, specific regions are generally amplified by PCR and then sequenced on an automatic sequencer. A large proportion (90%) of NGS information worldwide is generated on Illumina platforms using Sequencing by

Synthesis (SBS) technology (according to data from Illumina Inc. in 2015, https://www.illumina.com/). This method is similar to Sanger sequencing, but it uses modified dNTPs containing a terminator with a fluorescent label, which can be detected by a camera. After adding a new dNTP, the images are recorded and the terminators are removed ("reversible terminators"), continuing with more dNTPs addition cycles. In this way, the single bases are added uniformly, producing a set of DNA sequence reads of uniform length (Turcatti et al., 2008; Heather & Chain, 2016). Within the applications of high-performance genomic technologies such as NGS, genotyping allows the exploration and detection of genetic variations such as single nucleotide polymorphisms (SNPs) and large structural changes in DNA (Rapley & Harbron, 2012).

The Genotyping by sequencing (GBS) technique (Elshire et al., 2011) allows to discover sequences of short fragments distributed throughout the genome of a large number of samples without reference genome. This is achieved through sequencing a reduced representation library of genomic DNA of a high number of DNA fragments previously digested by restriction enzymes. Unlike other methologies, like restriction-site associated DNA (RAD) tagging (Miller et al., 2007; Baird et al., 2008), GBS has the advantages of requiring less DNA per sample and simpler laboratory processing. A variant of GBS (Poland et al., 2012) was also developed, using the combination of two restriction enzymes (instead of one) and libraries formed with barcoded forward adapters and a single common reverse adapter. This approach aimed to reduce its complexity properly and uniformly, and it was developed and tuned in plants (Poland & Rife, 2012), so it could be a good candidate to be used for population genetic research in mosses.

### Genome studies in bryophytes

Within bryophytes, the first sequenced genome belonged to the moss species *Physcomitrium patens* (Hedw.) Mitt. (=*Physcomitrella patens* (Hedw.) Bruch & Schimp.) (Rensing et al., 2008), and it is currently in its third version (Lang et at., 2018). New species and genera of bryophytes have been more recently incorporated into complete DNA and genome databases. At present, the genomes of *C. purpureus*

and several species of the genus *Sphagnum* are also available on the Joint Genome Institute (JGI) portal website (https://genome.jgi.doe.gov/portal/). Szövényi (2016) published the complete genome sequence of *Anthoceros agrestis* Paton. And Bowman et al. (2017) presented the first liverwort genome (*Marchantia polymorpha* L.). As a consequence, now the genomes of representatives of all three major bryophyte lineages are available.

In addition to genomic DNA, it is also interesting to study the DNA found in cytoplasmic organelles such as mitochondria and chloroplasts. These non-nuclear DNA molecules are small and circular, very similar to the DNA of a typical bacterium, there are usually numerous copies and they have an uniparental inheritance (normally maternal) (Birky, 1995; Freeland, 2020). Indeed, the chloroplast genome of *M. polymorpha* was the first of all plant species that was completely sequenced (Ohyama et al., 1986) and, six years later, the complete mitochondrial genome of this species became available (Oda et al., 1992).

### Transcriptome studies in bryophytes

Transcriptome is the set of all RNA molecules (transcripts) present in a cell or group of cells, in a tissue or even in an organism, including messenger RNA (mRNA), non-coding RNAs and small RNAs, at a given moment (Alberts et al., 2002; Zvelebil & Baum, 2008; Wang et al., 2009). These mRNA molecules reflect the gene expression in the biological sample studied at a specific time, providing important information about functional genes (Freeland, 2020). The mRNA expression varies according to the tissue studied, but it also differs according to the conditions under which it has been extracted (e.g., light or dark, temperature stress, water stress) (Wang et al., 2009, Gao et al., 2017). The reference transcriptomes of *C. purpureus*, *P. patens*, and some *Sphagnum* species are also available on the JGI website (https://genome.jgi.doe.gov/portal/).

## OBJECTIVES OF THE THESIS

The general objective of this Ph.D. thesis was to contribute to the knowledge on the importance of geography and climate in shaping genetic variation an their role as speciation mechanisms in bryophyte that exhibit wide distributions.

To achieve this aim, the cosmopolitan moss *F. hygrometrica* was used as a study model in four geographical areas of the Iberian Peninsula with different climatic regimes, including Asturias, Leon, and Murcia provinces and Sierra Nevada Mountains.

The specific objectives were the followings:

(i) To know whether there exist different ploidy levels in *F. hygrometrica* in the studied areas.

(ii) To describe the levels of genetic differentiation among *F. hygrometrica* populations from the four geographical areas studied using a GBS technique.

(iii) To confirm the presence of different and hybrid genotypes coexisting in their distribution range across the Sierra Nevada Mountains, as proposed in a previous study (Magdy Abdallah Awad, 2013). The absence of hybrids would suggest that the different lineages are genetically already too distinct to form a common gene pool, assuming that a recent colonization event resulting from demographic expansion can be ruled out.

(iv) To look for signatures of recent demographic expansion or contraction under the assumption of selective neutrality, in order to understand the recent demographic history of the lineages inferred.

(v) To know how the genetic structuring of the populations of *F. hygrometrica* is aligned with geographical distances (IBD) compared with environmental differentiation (IBE), in an effort to disentangle the role of adaptation in bryophyte evolution.

(vi) To compare gene expression among samples studied under natural conditions and thermal stress in order to provide baseline information on the molecular mechanisms that regulate the broad geographic distribution of *F. hygrometrica*.

(vii) To evaluate the presumed genetic, ploidy and morphological differences from a systematic perspective and, in the case of a lack of morphological variation, verify the existence of cryptic taxa.

# IV.

# MATERIAL AND METHODS

## STUDY AREAS

Four geographical areas in the Iberian Peninsula with different climatic regimes were chosen as study areas for this research: (i) Asturias and (ii) Leon provinces in the north of Spain, influenced by the Atlantic Ocean; (iii) Murcia province and (iv) the Sierra Nevada Mountains, the latter belonging to Granada and Almeria provinces in the south of Spain, which are mostly influenced by the Mediterranean Sea (Fig. 3). The Asturias province has a temperate, oceanic to hiperoceanic bioclimate; the Leon province is situated in a transition zone between temperate oceanic and mediterranean pluvioseasonal-oceanic bioclimates; the Murcia province exhibits in most of its territory a mediterranean xeric-oceanic bioclimate, having some areas in the NE and NW with a mediterranean pluvistacional oceanic bioclimate, another small area in the center with a mediterranean xeric continental bioclimate and even a mediterranean desertic oceanic bioclimate in the most southern Mediterranean coast. The Sierra Nevada Mountains present a mediterranean pluvioseasonal-continental bioclimate in the highest altitudes and a mediterranean pluvioseasonal-oceanic to a mediterranean xeric-oceanic bioclimate in the rest of the mountain system (Rivas-Martínez et al., 2004; López Fernández et al., 2008).

More specifically, the climate of the four study areas is characterized by significant differences in temperature, rainfall and snow precipitation, as follows: the Asturias province has mild winters and summers, with an average annual temperature around 13°C and a total average precipitation of 800-1400 mm/year in the northern half of the province; snow is present around 0-3 days/year in the sampled areas (AEMET, 2011; Chazarra et al., 2018). The Leon province has cold-long winters and hot-short summers, with an annual temperature of 5-12°C (depending on altitude) and a precipitation in the north and west fringes of 600-1200 mm/year; the average number of days with snow in the sampled locations is in the range of 5-20 days/year, depending on the altitude (AEMET, 2011; Chazarra et al., 2018). The Murcia province has mild winters and hot summers, with an average temperature of 18°C in coastal areas and around 15°C in the central area; rainfall is scarce and concentrated in a few

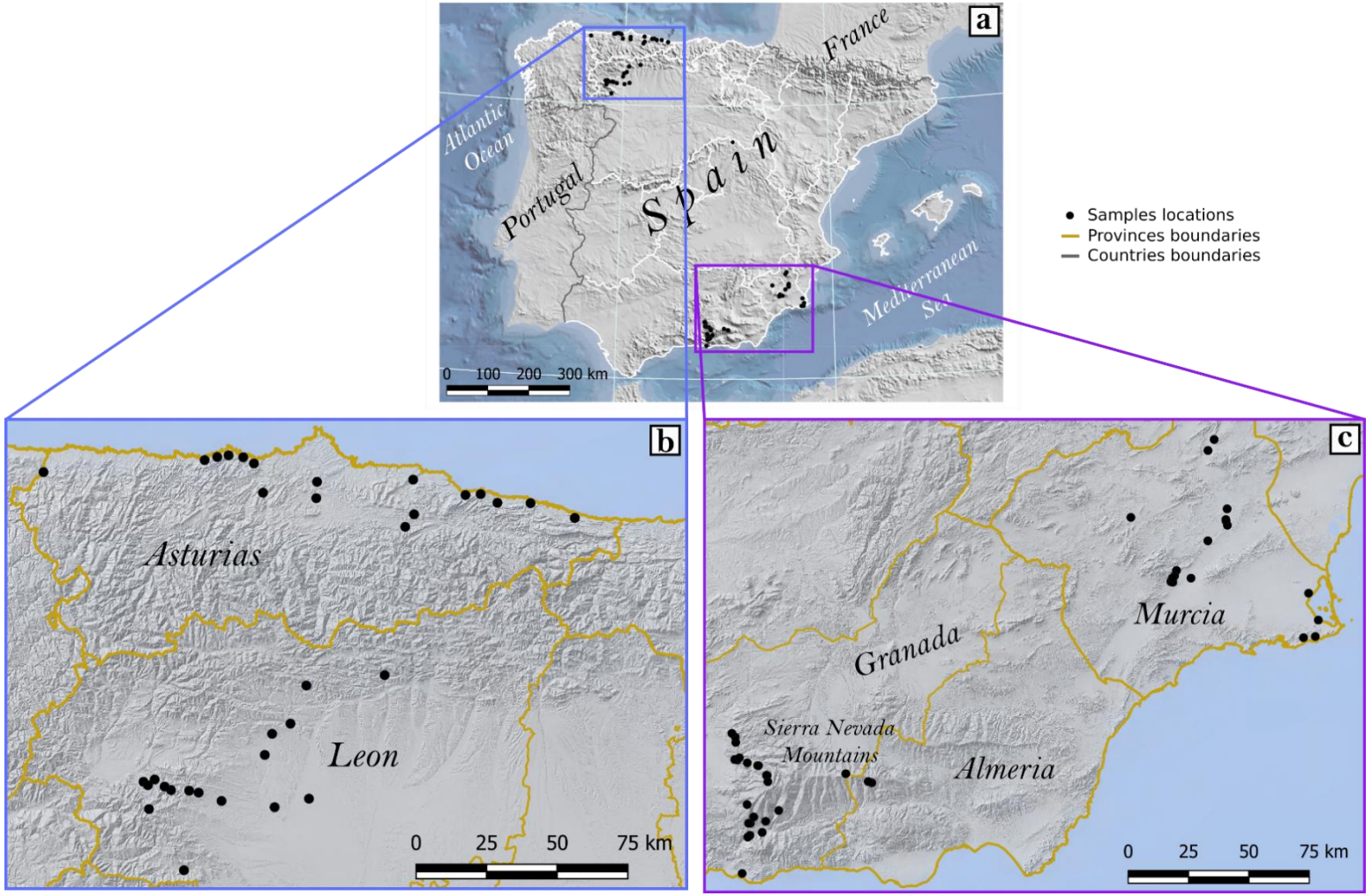



**Figure 3.** Study areas and geographic location of samples studied. **a**: Situation in Spain of the geographical areas studied. **b**: Detail of Asturias and Leon provinces locations (in the north of Spain). **c**: Detail of Sierra Nevada Mountains and Murcia province locations (in the south of Spain). Maps edited from IDEE (https://www.idee.es/es_ES/web/guest/directorio-de-servicios, http://servicios.idee.es/wms-inspire/mdt ).

days a year, with a precipitation of 250-350 mm/year in the sampled locations; snow is only present 0-1 days/year (Rodríguez Llopis et al., 2006; AEMET, 2011). In the Sierra Nevada Mountains, the annual temperature is in the range of 12-16°C (below 1500 m), 8-12°C (1500-2000 m), 4-8°C (2000-2500 m) and below 2-0°C (above 2500 m); while the precipitation ranges from 400 to 800 mm/year (depending on zone); snow is present 5-40 days/year above 1000 m (AEMET, 2011; Polo et al., 2019).

## PLANT MATERIAL AND SAMPLING METHOD

A total of 183 samples of *F. hygrometrica*, collected from 82 locations in the four areas studied (Fig. 3), were employed to carry out the molecular analysis. The locations were prospected during several field trips from 2010 to 2017, since the collection campaigns in the Sierra Nevada Mountains and the Murcia province had to be repeated several years. This was due to the difficulty of finding field plants with mature sporophytes that were necessary to obtain *in vitro* cultured plants, and also to the null or low germination rates of the spores obtained from some field collections (see *In vitro* cultures subsection). The geographic distribution of the samples collected was as follows: 36 samples were collected from Asturias, between 0 and 400 m a.s.l.; 37 from Leon, between 600 and 1500 m a.s.l.; 37 from Murcia, between 0 and 600 m a.s.l.; and 73 from Sierra Nevada Mountains, between 0 and 2700 m a.s.l. In terms of habitats, most sampling locations were situated in the proximities of roads and paths. When possible up to ten samples were collected from each location. The minimum distance between samples from the same location was one meter. During the sampling, environmental variables such as altitude, geographical coordinates, cormophytic and bryophytic vegetation were recorded from each location (see Appendix 1), as well as a brief description of the microhabitat for each sample. All field-collected samples are kept at MUB herbarium (University of Murcia, Spain).

In addition, six herbarium specimens were studied with the aim to complement the morphological and molecular analyses (see Appendix 1). Four were donated by Misha Ignatov from the MHA (Main Botanical Garden of Russian Academy of Sciences, Russia) and MW (Moscow University, Russia) herbaria, and two were provided on loan conditions by Bernard Goffinet from CONN herbarium (University of Connecticut, USA). Only one of the herbarium samples belonged to *F. hygrometrica* (from Chile) and the rest to four additional *Funaria* species as follows: one of *F. aequidens*, one of *F. arctica*, one of *F. flavicans*, and two of *F. polaris*.

## *IN VITRO* CULTURES

In order to obtain axenic plant material for DNA and RNA extractions, *in vitro* cultures were generated from all the field collections of *F. hygrometrica*, as well as from the four Russian specimens of *F. aequidens*, *F. arctica* and *F. polaris*. One mature capsule of each sample was surface-sterilized for 5 min using diluted store bought bleach (37 g/l active chlorine, diluted 1:4) and then washed with sterile distilled water. With the aim that capsules were fully hydrated, isolated capsules were kept in an Eppendorf tube for 4-8 h submerged in 500 µl sterile distilled water. Spores were then released in water by breaking the capsule with tweezers. Two dilutions of the spore suspension (1:10 and 1:100) were made and 30 µl of each one were sown on respective polystyrene ø 90 x 14 mm Petri dishes, in order to decrease the concentration of spores and facilitate their isolation. Gelzan (5 g/L, G1910 SIGMA) was used as gelling agent to which Murashige and Skoog Basal Salt Micronutrient Solution (0.23x final concentration, M0529 SIGMA) was added. Alternatively the macronutrients $KH_2PO_4$ (0.25 g/L, 145020 PROBUS), KCl (0.25 g/L, 131494 Panreac), $MgSO_4$ (0.5 g/L, 121320 PROBUS), $Ca(NO_3)_2$ (1 g/L, 131231 Panreac), and Fe EDTA (0.64 mg/L, 11275 SERVA) were used in a modification of the culture protocol of Ros et al. (2013b). A cellophan layer was put on medium in order to prevent that rhizoids penetrate the culture medium and facilitate transferring protonemata and making subcultures (Bopp & Brandes, 1964).

When spores began to germinate after keeping in culture chamber during 1-2 weeks, from each Petri dish incipient protonemata originating from clearly isolated individual spores (IIS) in any of the two dilutions made, were transferred to a new ø 90 x 14 mm Petri dish. Although only one protonema per capsule was needed for DNA isolation, five protonemata were taken in order to have spare ones in case any became contaminated or died. After two weeks, only two of the IIS from each Petri dish (one again as spare) were taken separately and put in respective polystyrene ø 55 x 14 mm Petri dishes, with the same medium and nutrients solution as before and covered also by a cellophan layer. The non-used IIS were discarded after having obtained one well developed individual. About 2-3 fragments of each IIS were taken in separated Petri dishes and left in the culture chamber for 4-8 weeks in order to obtain enough plant material to perform DNA extractions and not lose cultured samples in case of contamination. As minimum, there were three Petri dishes for each IIS.

After performing DNA extraction (see DNA extraction subsection), and with the aim of getting enough plant *in vitro* material to perform RNA extraction, all samples were transferred to new ø 55 x 14 mm Petri dishes. They contained culture medium with agar-agar fibers (20 g/L, SERVA) as gelling agent, to which the previously named macronutrients were added. Subcultures were performed again, taking about 2-3 fragments of each sample and putting them in separated Petri dishes with agar-agar medium. They were left in the growth chamber for 4-12 weeks.

All Petri dishes were kept in a growth chamber (Binder KB 720, Tuttlingen, Germany) at 22 ± 3°C and with 16/8 h light/dark, supplied by cool-white fluorescent tubes (Lumilux, Osram Germany) at a photon fluency rate of 33.5–55.0 μmol $m^{-2}$ $s^{-1}$. The Petri dishes were randomly put in the culture chamber to avoid the effects of small differences in light, temperature or other factors.

## PLOIDY ANALYSIS

Flow cytometry (FCM) technology was employed to estimate nuclear DNA content of 43 *in vitro* cultured individuals obtained by IIS. They belonged to Asturias (22

individuals), Murcia (5 individuals), and Sierra Nevada Mountains (16 individuals) (see Appendix 1). Leon specimens were still not collected when this analysis was carried out and, therefore, this area was not included. The protocol used by Nieto-Lugilde et al. (2018a) was followed. From each individual one stem was taken and introduced in 1 mL of LB01 buffer (Doležel et al., 1989) together with the internal standard (*Carex acutiformis* Ehrh., 1C = 0.41 pg; Lipnerová et al., 2013 or *Bellis perennis* L., 1C = 1.56 pg, the latter used according to our own calibration against *C. acutiformis*). The samples were then chopped with a razor blade. Propidium iodide as fluorochrome and RNase IIa (both at final concentration 50 µg/mL) were added and the samples were then stained during ≥ 10 min. The FCM analyses were performed at University of South Bohemia (Czech Republic), using a Partec CyFlow SL Flow cytometer. It was equipped with a 532 nm (green) diode-pumped solid-state laser (100 mW output); the fluorescence intensity of 12,000 particles was recorded. The fluorescence histograms were processed using FlowJo version 10.2 (https://www.flowjo.com/).

## DNA EXTRACTION

Protonemata and young gametophores grown *in vitro* were employed to perform DNA extractions using a protocol with CTAB and chloroform (Doyle & Doyle, 1990) with the addition of 1% PVP to the extraction buffer. DNA was quantified using a Qubit 2.0 fluorometer, with Qubit dsDNA BR Assay kit (Q32853, Invitrogen). The DNA concentration of each sample was adjusted to 10 µg/µl.

## DNA SANGER SEQUENCING

### Laboratory protocol

DNA Sanger sequencing was employed to know certain sequences of the nuclear (nr), mitochondrial (mt) and chloroplastic (cp) genome of 47 *in vitro* cultured individuals obtained by IIS from Sierra Nevada Mountains (see Appendix 1). Due to resource

limitations and wanting to focus the study on samples from these mountains, samples from the other three study areas were not used. As representative of the nr genome the ITS region was amplified with the primers AB 101F and AB 102R (Douzery et al., 1999), the mt *rpl5-rpl16* intergenic spacer was amplified with the primers rpl16-F1 and rpl5-R1 (Y. Liu et al., 2012b) and the cp *rps3-rpl16* region with the primers rps3F1-Fun (Magdy Abdallah Awad, 2013) and rpl16R4-Fun (AAGCAATAGAATTACCTCGAGTAGC; newly designed for this study) were used. 1 µl of stock DNA was used as template in 25 µl reaction volume. 200 µM of each dNTP (Canvax, Cordoba, Spain), 2.5 mM $MgCl_2$, 1 unit Taq polymerase (Horse-PowerTaq, Canvax, Cordoba, Spain), and the buffer (provided by the enzyme supplier) were added. The amplification conditions were as follows: 3 min at 94°C, 35 cycles with 30 sec at 94°C, 30 sec at 55°C and 1 min at 72°C, and a final 5 min extension step at 72°C. Amplification products were controlled on 1% agarose gels and successful reactions were cleaned by adding 1 u FastAP (Thermo Fischer) and 10 u ExoI (ThermoFischer) for one hour at 37°C followed by a thermal inactivation of 15 min at 85°C. Cycle sequencing was performed by Macrogen, using the amplification primers.

### Data analysis

The sequences were edited using Bioedit 5.0.9 (Hall, 1999) and aligned manually. We used gappy (Donath & Stadler, 2018) in order to identify reliable indels and recodify the indels in the data matrix as 0/1 character states. The data were analysed by Bayesian inference implemented with MrBayes 3.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003; Ronquist et al., 2012). Gaps were treated as standard data. Trees were sampled across the substitution model space in Bayesian MCMC analysis itself (Huelsenbeck et al., 2004) using the option nst=mixed. Therefore *a priori* model testing was not necessary. Indels were treated as separate unlinked partition, using the restriction site model (F81). When combining data from different genetic regions (nr, cp and mt) rates were allowed to be different across partitions with the command: prset applyto=(all) ratepr = variable. Two runs with four chains were conducted with 4,000,000 generations. Trees were sampled every 1,000$^{th}$ generation and the first 1,000,000 generations were discarded (burn-in) in order to exclude the trees before the chain reached the stationary phase. We checked for

stationarity of the log likelihood values, that the Potential Scale Reduction Factor (PSRF) was close to "1" (0.99 < PSRF < 1.01), and that the estimated sample size was above 200. The final trees were edited with TreeGraph2 (Stöver & Müller, 2010).

## GENOTYPING BY SEQUENCING

### Laboratory protocol

GBS libraries were prepared using 192 individuals obtained by IIS (see Appendix 1) following the protocol described in Poland et al. (2012), using two restriction enzymes: PstI (CTGCA^G) and Eco47I (=AvaII, G^GWCC). This technique was employed to obtain two multiplexed DNA libraries, each one generated with DNA fragments of 96 *in vitro* cultured individuals, with barcoded forward adapters (96 different barcodes, 4-9 bp in length) and a common reverse adapter. Variable length barcodes avoid problems with the Illumina software to detect or to reduce phasing errors during sequencing (Elshire et al., 2011). Nine samples were used as technical replicates.

In the restriction/ligation step 200 ng (20 µl) of genomic DNA of each sample were used. The final restriction/ligation mix contained 25 µM of the barcoded PstI adapter, 250 µM of the common Eco47I adapter, 10 u of PstI (FD ThermoFisher Scientific), 10 u of Eco47I (ThermoFisher Scientific), 5 Weiss units of T4 DNA Ligase, and 0.5 mM ATP in 40 µl final volume. The samples were held at 37°C for 2 h. The temperature was then risen to 65°C for 20 min in order to inactivate the enzymes. After this step the 10 µl of each sample were united in an Eppendorf tube and cleaned with the GeneJET Gel extraction kit (ThermoFisher Scientific) with the elution volume adjusted to 50 µl. For each library two independent cleanings were carried out and a total of eight amplification reactions of 25 µl were prepared (four for each clean-up) We used the Q5 High-Fidelity DNA Polymerase (NEB) for PCR amplification of the fragments. Each 25 µl reaction was set with 1 µl of the cleaned DNA fragments, 0.5 µl of 10 mM dNTPs (ThermoFisher Scientific), 1,25 µl of forward and reverse primers (10 mM) and 0.5 µl of Q5 polymerase. The cycling conditions

were as follows: initial denaturation of 30 s at 98°C, 20 cycles of 98°C for 10 s, 62°C for 20 s and 72°C for 20 s and a final extension at 72°C for 2 min. DNA fragments with size ≥ 700 bp and ≤ 150 bq were removed following the Double Size Selection Protocol of Select-a-Size DNA Clean & Concentrator Kit (D4080, Zymo Research). Finally, the size selected DNA of the two libraries was quantified again using Qubit 2.0 fluorometer, with Qubit dsDNA HS Assay kit (Q32854, Invitrogen) in order to guarantee the minimal DNA concentration required by the sequencing service. The libraries were sequenced on Illumina HiSeq 2500 instrument by Novogene.

### Single Nucleotide Polymorphism calling, clustering *de novo* and filtering

The sequencing process generated four fastq files (two files per library) with forward (R1) and reverse (R2) reads. Before starting to process raw sequence data, they were passed through a visor of quality control check using FastQC (Andrews, 2010). Quality scores should not be below 20 (Q20 is equivalent to the probability of a correct base call is 99%). Although with differences in quality, fastq files with both forward and reverse sequences could be employed.

Next, GBS data were demultiplexed, quality filtered and *de novo* clustered using ipyrad 0.9.28 (Eaton & Overcast, 2016; http://ipyrad.readthedocs.io/). Only reads with unambiguous barcodes and fewer than five low quality bases (Phred quality score < 20), were retained, and a strict filter was applied to remove Illumina adapter contaminants. A first analysis was undertaken with all individuals to assess patterns of relatedness among individuals. However, the data quality was limited, and during this first round of analyses, only R1 reads could be analysed. Then, a reference genome for *F. hygrometrica* was provided by Bernard Goffinet´s lab (Kirbis et al., 2020) and the performance significantly improved. We then chose the option of *de novo* + reference clustering method, which provided a much better resolution. For this analysis, parameters were set to the following values following García-Olivares et al. (2019) approach: clust_threshold (the sequence similarity threshold) = 0.90, max_SNPs_locus (the maximum number of SNPs allowed in a locus) = 20, and min_samples_locus (the minimum number of samples that must have data at a given

locus for that locus to be retained in the final data set) = 60%, with all remaining parameter values set to their default values.

In summary, the ipyrad parameters were set to default, except for the following: "gbs" was indicated as data type; the restriction overhang was "TGCAG, GWCC", because two restriction enzymes (PstI and Eco47I) were employed during digestion; "5" as minimum depth for statistical base calling; "3" as minimum depth for majority-rule base calling; clustering threshold set to 0.90 (90%) for *de novo* assembly; one allele as maximum number of alleles per site in consensus sequences (haploid); in our case, at least 80% of our samples (145 of 183) was the minimum number of samples per locus for output; filter for adapters/primers in its most restrictive option ("2"); and "*" was indicated to make all output datatypes. Upon identifying the appropriate parameter combination, replicate samples were used to estimate genotyping error (Mastretta-Yanes et al., 2015).

Raw reads of each library were demultiplexed separately by step 1 of ipyrad, because the barcodes that identify each sample were the same to form both libraries. Afterwards, all demultiplexed reads were processed following the general workflow of this pipeline for paired-end sequencing. Adapters/primers and edges of poor-quality reads were removed (step 2), reads were clustered within samples and aligned without reference genome (step 3), consensus allele sequences were estimated (steps 4-5), consensus sequences were clustered and aligned across all of the samples (step 6), and the final alignments were filtered (as indicated our parameter settings) and saved in different out-put formats (step 7).

Before carrying out phylogenetic analyses of output data, VCFtools v0.1.15 (Danecek et al., 2011; http://vcftools.sourceforge.net) was used to perform additional quality checking steps. Our data in vcf format were analysed to evaluate the potential genotyping errors and individual levels of missing data. Based on this, the *loci* and the individuals that did not sequence well were removed.

## RNA EXTRACTION

Twenty-six *in vitro* cultured individuals obtained by IIS coming from the four studied areas were employed to study their mRNA (see Appendix 1). They were selected based on the DNA analysis data (see Results section). Two Petri dishes of each individual were employed. One Petri dish of each indivivual was exposed to thermal stress by heating at 40 ± 0.4°C for 16 h in a laboratory oven (Heraeus B6 incubator, Kendro Laboratory Products, Germany), specifically near the limit temperature at which *F. hygrometrica* can survive in humid state (45°C), as observed in our previous assays (unpublished results). The other Petri dish was used as a control sample, by maintaining it under culture chamber conditions (at 22 ± 3°C) for 16 h. Since the laboratory oven had no available light inside, both conditions (control and thermal stress) were in darkness to minimize other differences apart from the temperature. In the case of two individuals technical replicates were prepared for both conditions. Another four individuals of three *Funaria* species were included in the RNA study: one of *F. aequidens*, one of *F. arctica* and two of *F. polaris* (only subjected to control conditions) (see Appendices 1 and 2). Therefore, in a total of 60 samples were mRNA sequenced. The plant samples were then blotted dry on an absorbent paper sheet and immediantely shock frozen in liquid nitrogen at the same time in the morning and stored at -75°C until RNA extraction.

RNA was extracted with two protocols. The first version used an extraction buffer consisting of 6 M guanidine thiocyanate in 25 mM Tris pH 6.4 and 10 mM EDTA pH 8.0 complemented with 1.3% (v/v) Triton X-100 (Boom et al., 1990). 1% mercaptoethanol was added inmediately prior to use. The solution was then extracted with chloroform/isoamyl alcohol (24/1) and centrifuged for 5 min at 12 500 x g. The supernatant was complemented with the same volume of 96% ethanol. The solution was applied to the NZY total RNA isolation kit according to the manual (NZYtech, MB13402).

The second RNA extraction protocol used a variant of the LogSpin method of Yaffe et al. (2012). For 30 mg – 100 mg of plant material 0.5 ml of phenol/chlroform/isoamyl alcohol (25/24/1), pH 4.3 was added. The material was

ground with mortar and pestle until no major plant fragments were observed. Then, 0.1 ml of 3M sodium acetate (pH 5.2) and 0.4 ml of $ddH_2O$ were added. The samples were mixed and then centrifuged for 5 min at 12,500 x g. The supernatant was transferred to a new tube and 1.5 volumes of 96% ethanol were added. The solution was then applid to a spin column of the GeneJET Gel Extraction Kit (ThermoFisher Scientific) and centrifuged for 1 min at 12,500 x g. The colums were then washed with 600 µl 3M sodium acetate, pH 8.0, 600 µl 70% ethanol and finally 250 µl 70% ethanol, each wash step followed by a centrifugation of 1 min at 12,500 x g. The RNA was finally eluted twice with a total of 70 µl $ddH_2O$.

## TRANSCRIPTOME STUDY

The RNA samples were send to Novogene Europe for massive sequencing (RNASeq). On an Illumina NovaSeq 6000 machine (2 x 150 bp read length). Low quality reads due to adapter contamination, high percentage of uncertain nucleotide calls (N > 10%) or a high number of low quality nucleotides (Phred score below 20 of more than 50% of the bases of the read) were excluded. Further quality control steps included the error rate distribution along reads and the A/T/G/C content distribution along reads. Theoretically, A should equal T and C should equal G.

After the initial filtering of low quality reads the gene transcripts were assembled *de novo* with the Trinity software package vs r20140413p1 (Grabherr et al., 2011; Haas et al., 2013) using the default settings with the exception of min_kmer_cov=2 and min_glue=2. This is necessary because the Illumina reads are much shorter than the length of most transcripts. The assembled short reads are called “contigs”. One problem associated with downstream applications like differential expression analysis is that programs like Trinity often report more than one contig for one gene because there are several transcript isoforms that are slightly different. Another source for different contigs reported for one gene are sequencing errors and artifacts that occur during the *de novo* assembly (Davidson & Oshlak, 2014). Corset (Davidson & Oshlak, 2014) was used to remove these redundancies from the Trinty results by clustering

contigs into genes. The longest transcript of each gene is called unigene and serves as basis for the following analyses.

Once the unigenes were identified, the gene function was deduced with searches in several databases: (1) NR, the NCBI non-redundant protein sequence database (https://www.ncbi.nlm.nih.gov/refseq/; O'Leary et al., 2016); (2) NT, the nucleotide sequence database of the NCBI (https://www.ncbi.nlm.nih.gov/nucleotide/); (3) Pfam (Protein family), a database that collects protein domains and families (http://pfam.xfam.org/; El-Gebali et al., 2019); (4) KOG/COG, databases that are based on evolutionary relationships of homologous genes (http://www.ncbi.nlm.nih.gov/COG); (5) Swiss-Prot, a high quality database, which is manually edited (Boeckmann et al., 2003); (6) KEGG (Kyoto Encyclopedia of Genes and Genome), a database with a complete annotation system (https://www.genome.jp/kegg/; Kanehisa, 2019; Kanehisa & Goto, 2000; Kanehisa et al., 2021); and (7) GO (Gene Onthology) for the functional annotation of gene products (http://geneontology.org/; Ashburner et al., 2000; The Gene Ontology Consortium, 2018). To search NR and Swiss-Prot we used DIAMOND v0.8.22 (Buchfink et al., 2015) with the e-value set to $1 \times 10^{-5}$. DIAMOND was also used in the case of KOG/COG but with the e-value set to $1 \times 10^{-3}$. KAAS r140224 (Moriya et al., 2007) was used for KEGG annotation with the e-value set to $1 \times 10^{-8}$. In the case of NT NCBI blast v2.2.28+ (Altschul et al., 1990) was used with an e-value of $1 \times 10^{-5}$. Pfam annotations were conducted with hmmscan (https://www.mankier.com/package/hmmer) and the e-value set to 0.01. Finally, blast2go b2g4pipe_v2.5 (Götz et al., 2008) with an e-value of $1 \times 10^{-6}$ was used for the GO annotation.

The transcript abundances were quantified with the help of RSEM (Li & Dewey, 2011). This program maps the reads of the samples back to the *de novo* transcriptome with the help of Bowtie2 (Langmead & Salzberg, 2012) and then gets the read count for individual samples and the normalized read counts. Normalization is necessary because a higher gene length rises the gene count and because different samples have slightly different data output in gigabases. Differentially expressed genes were identified with an adjusted p-value of 0.05 using DESeq2 (Love et al., 2014). DESeq2

was also used to calculate the heatmap od Euclidean sample distances, principal components analysis and cluster analysis.

The enrichment analysis is aimed to find genes related to certain biological functions or pathways whose expression is significantly changed between groups of treatment or other biological conditions like species. We searched for Gene Onthology (GO) enrichment of differentially expressed genes. This analysis was done using Goseq 1.32.0 (Young et al., 2010) and topGo 2.32.0 (Alexa & Rahnenfuhrer, 2020) with a corrected p-value 0f 0.05. Additionally we conducted a KEGG enrichment analysis with KOBAS v3.0 (Mao et al., 2005) with an adjusted p-value set to 0.05. This analysis is aimed to identify biochemical pathways associated with differentially expressed genes.

mRNA is longer than the coding sequences of genes because it contains 3' and 5' untranslated sequences and introns. If the unigenes match information contained in the NR and Swissprot databases the coding regions were extracted directly and translated into peptide sequences using the standard codon table. Unigenes with no hits were analysed with ESTScan 3.0.3 (Iseli et al., 1999) to predict their coding region and sequence direction.

SNPs (single nucleotide polymorphisms) are DNA sequence variations that affect a single nucleotide. The term INDEL refers to insertions or deletions of one or more nucleotides when comparing two or more sequences. We used samtools 1.0 /bcftols 1.10 (Li et al., 2009) to carry out SNP calling and INDEL calling. The settings were: QUAL > 20 (quality score); DP > 4 (sequencing depth) and MQ > 30 (mapping quality).

## POPULATION GENOMIC ANALYSES

### Principal Component Analysis and sparse non-negative matrix factorization cross-entropy analyses

To evaluate congruence between our results and former population genetic analyses (Madgy Abdallah Awad, 2013; Magdy et al., 2016), principal component analysis

(PCA) was used to describe the existence of genomic groups of individuals. We also used sparse non-negative matrix factorization (sNMF) cross-entropy to compare the number of inferred ancestral populations and provide ancestry coefficients for each analysed individual. PCA analyses were performed using the *adegenet* 2.1.0 package (Jombart & Ahmed, 2011) and sNMF using the *LEA* package (Frichot & François, 2015) in R 3.6.1 (R Core Team, 2019). The sNMF algorithm provides model-based least-squares estimates of ancestry proportions to infer individual ancestry and population clustering (Frichot et al., 2014). To obtain the best-fit number of ancestral populations (*K*) within the four studied areas, *K* values from 1–10 were evaluated with 1,000 replicates per *K*, using a cross-entropy criterion to identify the best fit value of *K*. In addition, the minimum spanning network (MSN) was employed to identify phylogenetic relationships in samples studied using the *poppr* package in R (https://github.com/zkamvar/poppr_msn_shiny). A Reynold's distance matrix was used as the distance calculation.

### Landscape variables

To test the influence of landscape climatic factors on genetic differentiation in *F. hygrometrica* within the four areas studied, six climatic variables were assessed on genomic relatedness among individuals. These variables were: annual mean temperature (°C x10), mean diurnal range (°C), isothermality, maximum temperature of warmest month (°C x10), minimum temperature of coldest month (°C x10), and annual precipitation (mm/year). The corresponding layers were downloaded in GEOtiff format from CHELSA climate dataset, v1.2 (Karger et al., 2017; https://chelsa-climate.org/), with a spatial resolution of 1 km × 1 km (30 arcseconds). As these layers are provided at worldwide scale, the six layers were cropped to the extent of each area, resampled to a 100 m resolution to make optimisation of resistance surfaces tractable and checked for collinearity. The layer extents of each study area (minimum and maximum longitude and latitude) at the time of making the cropping polygon, with extent format (xmin, xmax, ymin, ymax) were: Asturias (-7.200, -4.500, 43.020, 43.620), Leon (-7.050, -4.800, 42.000, 43.020), Murcia (-2.350, -0.683, 37.383, 38.750), and Sierra Nevada Mountains (-3.800, -2.560, 36.650, 37.450).

### Landscape patterns of genomic similarity: IBD *versus* IBE

Here we used landscape genomic analyses within each of the four areas studied to test among alternative hypotheses, including IBD and IBR, which here is interpreted as IBE and involves climate-driven isolation. In addition to the resistance surfaces, we assessed Euclidean distance alone (IBD) as well as an intercept only null model. To do this, we used the R package ResistanceGA 4.0 (Peterman, 2018) which, in general, explores parameter space to find the resistance surface values and transformations or the combination of them that maximize the statistical relationship between matrices of pairwise genetic distances (response) and cost-distances (predictor/s). Such a statistical relationship was assessed through a linear mixed effects model approach with maximum likelihood population effects (MPLE; Clarke et al., 2002), as implemented in the R package lme4 (Bates et al., 2014) and proposed in Peterman et al. (2014) and Peterman (2018). This approach circumvents typical issues such as spatial autocorrelation and high dimensionality that resistance surfaces can have, and subjectivity in assigning resistance values (Peterman et al., 2014; Peterman, 2018). Pairwise genomic distances between individuals, estimated following the approach proposed by Petkova et al. (2016), were used as the dependent variable, while scaled and centred circuit resistance distance matrices between individuals were used as independent variables. Analyses were applied to each area separately. Due to the results obtained in Sierra Nevada Mountains, two data treatments were applied to: the first treatment included all individuals, while the second treatment excluded two subsets of specimens that showed significant departures in their genetic distance with respect to the core of specimens with Sierra Nevada Mountains (see Results section).

To calculate pairwise resistance distances between individuals, we used the random-walk commute time algorithm (function “commuteDistance” in “gdistance”), a genetic algorithm to maximize fit of resistance surfaces similar to circuitscape (McRae, 2006; McRae et al., 2008). We used the wrapper function “all_comb” in ResistanceGA to implement single-surface and multiple-surface optimization, followed by a bootstrapping step. For the multiple-surface optimization, we simultaneously combined two resistance surfaces to optimize and create a novel composite resistance surface.

### Demographic analysis (Tajima's D neutrality test)

To test for signatures of recent demographic expansion under the assumption of selective neutrality, we used Tajima's D (Tajima, 1989) neutrality test conducted in DnaSP 6 (Rozas et al., 2017), using all SNPs. While some *loci* may be under selection, it is not unreasonable to assume that the great majority are behaving neutrally. Significantly negative values for Tajima's D point to an excess of rare polymorphisms in a given population, which can be an indication of a recent increase in population size. In contrast, significantly positive values indicate a recent population contraction.

## BIOMETRIC STUDY

In several former studies performed across Sierra Nevada Mountains (Madgy Abdallah Awad, 2013; Magdy et al., 2016), a clear segregation into two groups of *F. hygrometrica* samples according to their altitudinal range was inferred: lowland populations (below 1500 m a.s.l.) *versus* highland populations (between 1500 m and 2700 m a.s.l). The present thesis revisits the question of which mechanisms determine the genetic structuring by means of genomic and transcriptomic approaches (see Results section). In order to shed light on the possible existence of cryptic species, and whether morphological variation mirrors spatial genetic and ploidy variation, we performed a quantitative morphometric study of *F. hygrometrica* across the Sierra Nevada Mountain range.

A total of 30 collections of field plants were selected (see Appendix 1) across the two DNA groups. In each collection, whenever possible, three fructified plants were taken for analysis of both gametophytic and sporophytic characters. Only in two Sierra Nevada Mountains collections, two plants were used due to their scarcity in the field. The final number of plants studied was 88. Moreover, the six *Funaria* specimens received from CONN, MHA and MW herbaria from outside Sierra Nevada Mountains were studied morphologically to provide a taxonomical context. From these, 18 plants were analysed, three from each specimen, belonging to *F. aequidens*, *F. arctica*, *F. flavicans, F. hygrometrica* and *F. polaris* (see Appendix 1).

In total, 29 morphological characters were selected (16 quantitative and 13 qualitative), according to the diagnostic characters used in the identification keys of the genus *Funaria* in Europe, North America and Russia (Casas et al., 2006; McIntosh, 2007; Brugués & Ruiz, 2010; Ignatova & Fedosov, 2017), and also to our own observations (Table 1, Figs. 4, 5 and 6). First, capsule position with respect to the seta and presence/absence of striae in mature closed capsule were observed in dry conditions. The rest of characters were observed and measured after hydrating the plants for 8-12 h. Phyllidial characters were observed in nine phyllidia (three from each caulidium). The diameter of 15 spores was measured, five of each three capsules studied. Terms used are based on "Glossarium polyglottum bryologiae, a multilingual glossary for bryology" (Magill, 1990).

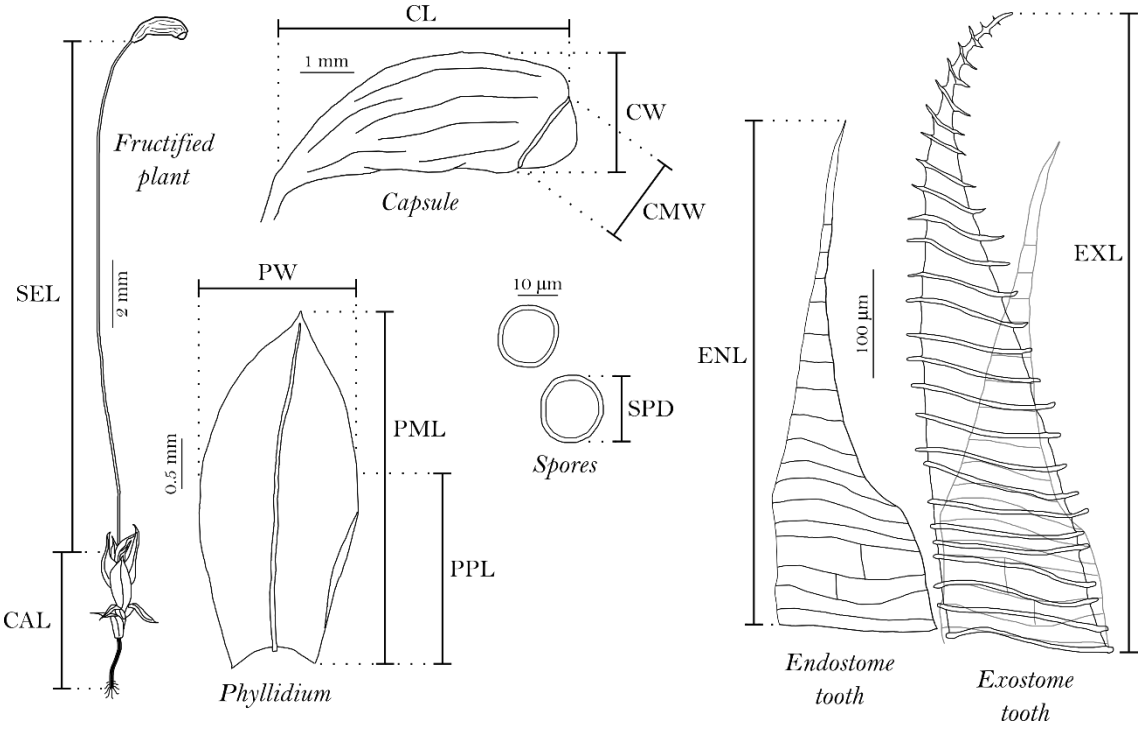



**Figure 4.** Quantitative morphological characters included in the biometric study of *F. hygrometrica*. Abbreviations are those used in Table 1. Source: own elaboration.

**Table 1.** Descriptive morphological characters included in the biometric study of *F. hygrometrica* with indication of the number of items studied, abbreviations used, type (QL = qualitatitive, QT = quantitative) and status characters.

| | Abbreviation | Character | Type and status character |
|---|---|---|---|
| **Caulidium** (from 3 fructified plants) | CAL | Caulidium length | QT (µm) |
| **Phyllidium** (3 phyllidia from each caulidium) | PP | Phyllidium posture when moist | QL: plain (1); keeled (2); concave (3) |
| | PML | Phylllidum maximal length (including apiculus) | QT (µm) |
| | PPL | Phyllidium partial length, from base to widest part | QT (µm) |
| | PPL/PML | Phyllidium partial length / Phylllidum maximal length ratio | QT (µm) |
| | PW | Phyllidium width at widest part | QT (µm) |
| | PML/PW | Phylllidum maximal length / Phyllidium width ratio | QT |
| | PSH | Phyllidium shape | QL: elliptical (1); oblong-ovate (2); obovate (3); oblong-lanceolate (4); oblanceolate (5) |
| | NL | Nerve length | QL: ending below the apex (1); excurrent (2) |
| | PA | Phyllidium apex | QL: acute (1); apiculate (2); obtuse (3) |
| | PM | Phyllidium margins at upper part of leaf | QL: entire (0); dentate (1); serrate (2) |
| | PMC | Phyllidium margins curvature | QL: plane (0); incurvate (1) |
| | PMR | Number of rows of marginal cells narrower than other phyllidium laminal cells | QT (number of rows) |
| **Seta** (from 3 fructified plants) | SEC | Seta contortion (when moist) | QL: straight (1); curved (2); strongly twisted (3) |
| | SEL | Seta length (when moist) | QT (mm) |
| **Capsule** (from 3 fructified plants) | CPS | Capsule position with respect to the seta (whendry) | QL: nearly straight (1); oblique (30-60°) (2); horizontal (90°) (3); Pendent (4) |
| | CSY | Capsule symmetry | QL: symmetric (1); nearly straight = weakly curved (2); asymmetric = curved (3) |
| | CSH | Capsule shape (seen from below) | QL: cylindrical (1); pyriform (2); ovoid (3); ellipsoidal (4) |
| | CL | Capsule length | QT (µm) |
| | CW | Capsule width | QT (µm) |
| | CMW | Capsule mouth width | QT (µm) |
| | CST | Striae in mature dryclosed capsule | QL: absent (0); present (1) |
| | ANN | Annulus revoluble | QL: absent (0); present (1) |
| | EXL | Exostome teeth length | QT (µm) |
| | ENL | Endostome teeth length | QT (µm) |
| | ENL/EXL | Endostome teeth length/exostome teeth length ratio | QT |
| | ENR | Endostome teeth roughness | QL: smooth (0); papilose (1) |
| | ENA | Endostome teeth apex | QL: pointed (1); bluntly 2-lobed (2); emarginated (3) |
| **Spores** (5 spores from each capsule) | SPD | Spore diameter | QT (µm) |

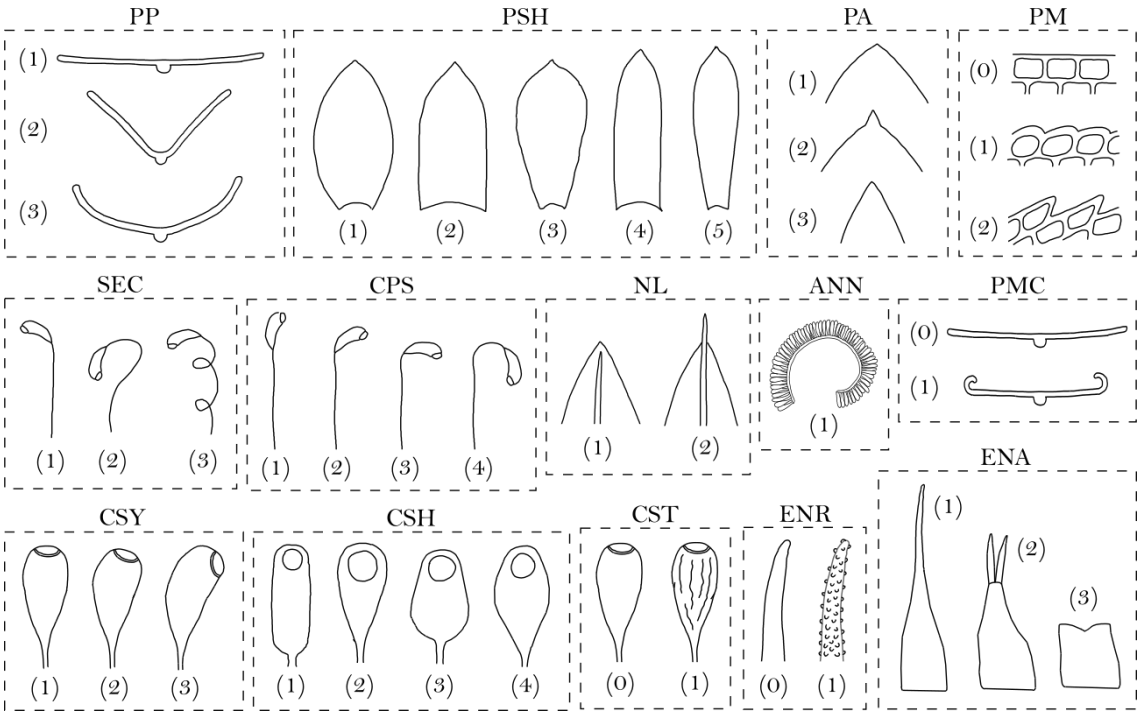
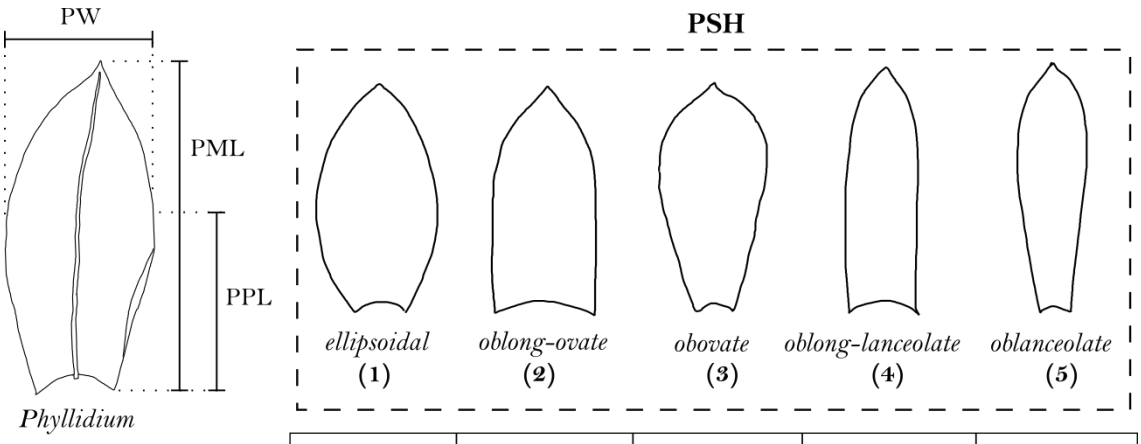



**Figure 5.** Qualitative morphological characters included in the biometric study of *F. hygrometrica*. Abbreviations and status character are those used in Table 1. Source: own elaboration.

To determine the shape of the leaves in the most objective and exact way possible, the values of characteristics PPL, PPL/PML and PML/PW were used, as shown in Fig. 6.




| | | | | | |
|---|---|---|---|---|---|
| $\frac{PPL}{PML}$ = | 0.5 | < 0.5 | > 0.5 | < 0.5 | > 0.5 |
| $\frac{PML}{PW}$ = | < 2<br>> 2 | < 2 | < 2 | > 2 | > 2 |

**Figure 6.** Phyllidia shape considered in the biometric study of *F. hygrometrica*. Abbreviations are those used in Table 1. Source: own elaboration.

The morphological parameters of caulidium, seta and capsule (except peristome) were studied and measured using a stereomicroscope (Leica A8APO) with a video camera connected (Leica camera: DFC295) in order to transfer the images to a computer. The Leica Application Suite v.4.1.0 was used for image analysis. The morphological parameters of phylllidia, peristomes and spores were measured and observed using a microscope (Olympus BH2) with a micrometer inserted in an ocular.

For statistical analyses, morphometric data were entered in a spreadsheet (Microsoft Excel) and then imported in R version 3.6.2 (R Core Team, 2019), in a tabular file format. The R functions employed for the morphometric analysis were based on MorphoTools (Koutecký, 2015). The qualitative characters were split in as many characters as number of status of that character. Furthermore, the values of the qualitative characters were transformed into percentages of presentiality in the plants studied. The values of characters that presented a non-normal distribution of the data were transformed into logarithms. The multivariate morphometric analyses carried out were PCA and Linear Discriminant Analysis (LDA). In the case of PCA, MorphoTools automatically standardizes the observed values in order to avoid the distorting effect of differing ranges or different scales. Within MorphoTools we also conducted a classificatory discriminant analysis with cross-validation. MorphoTools uses the lda function from the package MASS (Venables & Ripley, 2002).

## SIMULATION OF DISTRIBUTION PATTERNS

Because the molecular results (see Results section) suggested the possibility that two *Funaria* species with overlapping habitat requirements exist, a R script simulating this situation was used developed by Olaf Werner (https://github.com/olafumes/Vicariance). The basic assumptions of the simulation were that a linear geographic gradient exists along 100 equidistant localities. The fitness value of the two species was given by values between “0” and “1”. This fitness value was multiplied by the abundance of the species at each locality to calculate the offspring number. A fitness value of “0” means that no offspring is produced at this locality and “1” is the maximum number. The fitness values for the two species at

locality one and locality 100 were specified at the beginning and the values for the intermediate localities were calculated as equidistant steps along the gradient. Each generation a certain number of new individuals reach the neighboring localities where their parents lived. The function that describes this behavior is: $1/(10^*(d+1)^2)$ (d=distance units, neighboring localities have a distance of "1"). This means that at a distance of "1" the number of new individuals that arrives is 1/40 of the spores that arrive at the original locality, at a distance of "2", 1/90 of the spores of the original locality arrive, and so on. The growth in each locality is given by the logistic formula $N_1 = rN_0^*[(K-N_0)/K]$ ($N_1$ = number of individuals at time 1, r = growth parameter, $N_0$ = number of individuals one generation before time 1, K = capacity of locality), which results in the typical sigmoid graph. The two species are competing, because the sum of the number of individuals of both species in one locality is given by the capacity at this locality. In other words, if the locality is completely occupied, the number of individuals of one species can only increase if the other species decreases in number. At the beginning of the simulation, locality one and locality 100 are each occupied completely by the species with the highest fitness value there, all intermediate localities are empty. The results of the first and last generation are saved as "svg" files. Additionally, a video in "mp4" format with the sequence of all generations is saved to show the results.

# V.

# RESULTS

## GENOME SIZE

Four groups of cytotypes, three of them poorly differentiated, were obtained by FCM for the analysed 43 individuals of *F. hygrometrica*, sampled across the provinces of Asturias, Murcia, and the Sierra Nevada Mountains (Table 2 and Fig. 7).

When all individuals were considered together, the first cytotype (a) had a mean value of DNA content of 1C = 0.463 pg (standard deviation, SD = 0.020). The second (b) and third (c) groups showed a mean value of 1C = 0.895 pg (SD = 0.015) and 1C = 0.964 pg (SD = 0.014), respectively. When the second and third cytotypes were considered together (b+c), the mean value of DNA content was double with respect to first group (99.65% more), but with a larger standard deviation (SD = 0.038) than when they were considered separate. The fourth cytotype (d), with only two individuals had 1C = 1.100 pg mean value of DNA content (SD = 0.033). The first cytotype (a) is overrepresented in the data, with 67.44% of the individuals, with respect to the rest of cytotypes (16.28%, 11.63% and 4.65%, respectively). The third cytotype seems to be slightly more frequent in the Sierra Nevada Mountains than in the rest of the study areas. The only two individuals with the fourth cytotype occur in the same locality, in Asturias.

**Table 2.** *Funaria hygrometrica* nuclear DNA content as measured by FCM. Cytotypes considered, number of individuals used in the analysis (n), mean value of DNA, standard deviation, and range of DNA values obtained for each cytotype are given.

| | Cytotype | n | Mean DNA value (pg) | Standard deviation | Minimum DNA value (pg) | Maximum DNA value (pg) |
|---|---|---|---|---|---|---|
| Asturias province | a | 18 | 0.459 | 0.016 | 0.436 | 0.514 |
| | b | 2 | 0.897 | 0.016 | 0.886 | 0.909 |
| | c | - | | | | |
| | d | 2 | 1.100 | 0.033 | 1.077 | 1.124 |
| Murcia province | a | 2 | 0.486 | 0.057 | 0.445 | 0.526 |
| | b | 2 | 0.878 | 0.010 | 0.870 | 0.885 |
| | c | 1 | 0.984 | | | |
| Sierra Nevada Mountains | a | 9 | 0.466 | 0.015 | 0.449 | 0.487 |
| | b | 3 | 0.906 | 0.002 | 0.904 | 0.908 |
| | c | 4 | 0.959 | 0.009 | 0.947 | 0.969 |
| All the samples | a | 29 | 0.463 | 0.020 | 0.436 | 0.526 |
| | b | 7 | 0.895 | 0.015 | 0.870 | 0.909 |
| | c | 5 | 0.964 | 0.014 | 0.947 | 0.984 |
| | d | 2 | 1.100 | 0.033 | 1.077 | 1.124 |
| | b + c | 12 | 0.924 | 0.038 | 0.870 | 0.984 |
| **Total individuals:** | | **43** | | | | |

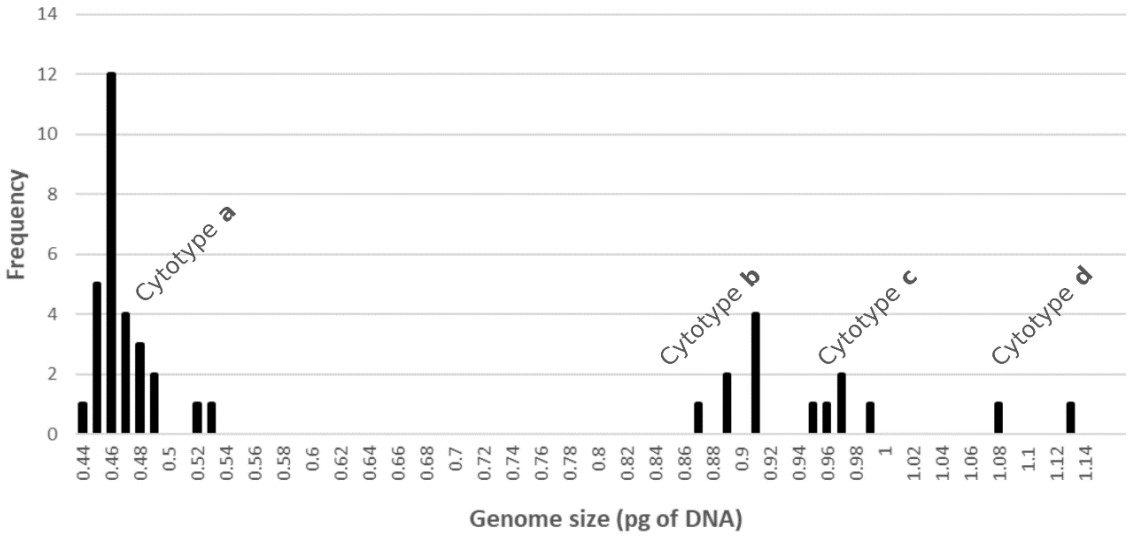



**Figure 7.** Histogram of genome sizes of the *F. hygrometrica* individuals studied generated by FCM. The four cytotypes obtained are indicated.

## DNA SANGER SEQUENCING

The complete alignment of all sequences had a length of 2,296 bp. The nr ITS sequences alignment had a length of 809 bp, the mt alignment 854 bp and the cp alignment 633 bp. Gappy identified 18 reliable indels in the ITS region, 2 in the cp *rps3-rpl16* region and 1 in the mt *rpl5-rpl16 region*. This mt indel is especially remarkable for its length, which stretches over 45 bp. In the case of the ITS region, several sequences of other *Funaria* species where available and added to the alignment, e.g. *F. aequidens*, *F. flavicans*, *F. microstoma*, *F. polaris* together with one sample of *F. hygrometrica* from Chile. The resulting tree (Fig. 8) is unrooted because one of the two Sierra Nevada Mountains (SN) clades seems to be more related to several of the remaining *Funaria* species than to the second SN clade and, as a consequence, no outgroup specimens could be established without doubt. Other available Funariaceae sequences like those of several *Entosthodon* species or *Funariella curviseta* (Schwägr.) Sérgio were too distant for unambiguously aligning a major part of the sequences. While the *F. hygrometrica* plant from Chile is close to a clade formed by samples from the higher altitudes of SN, the second clade of SN seems to be close to a clade composed of *F. aequidens*, *F. flavicans* and *F. polaris.* The samples of the second clade can be found all over SN from the lowlands to the highest altitudes.

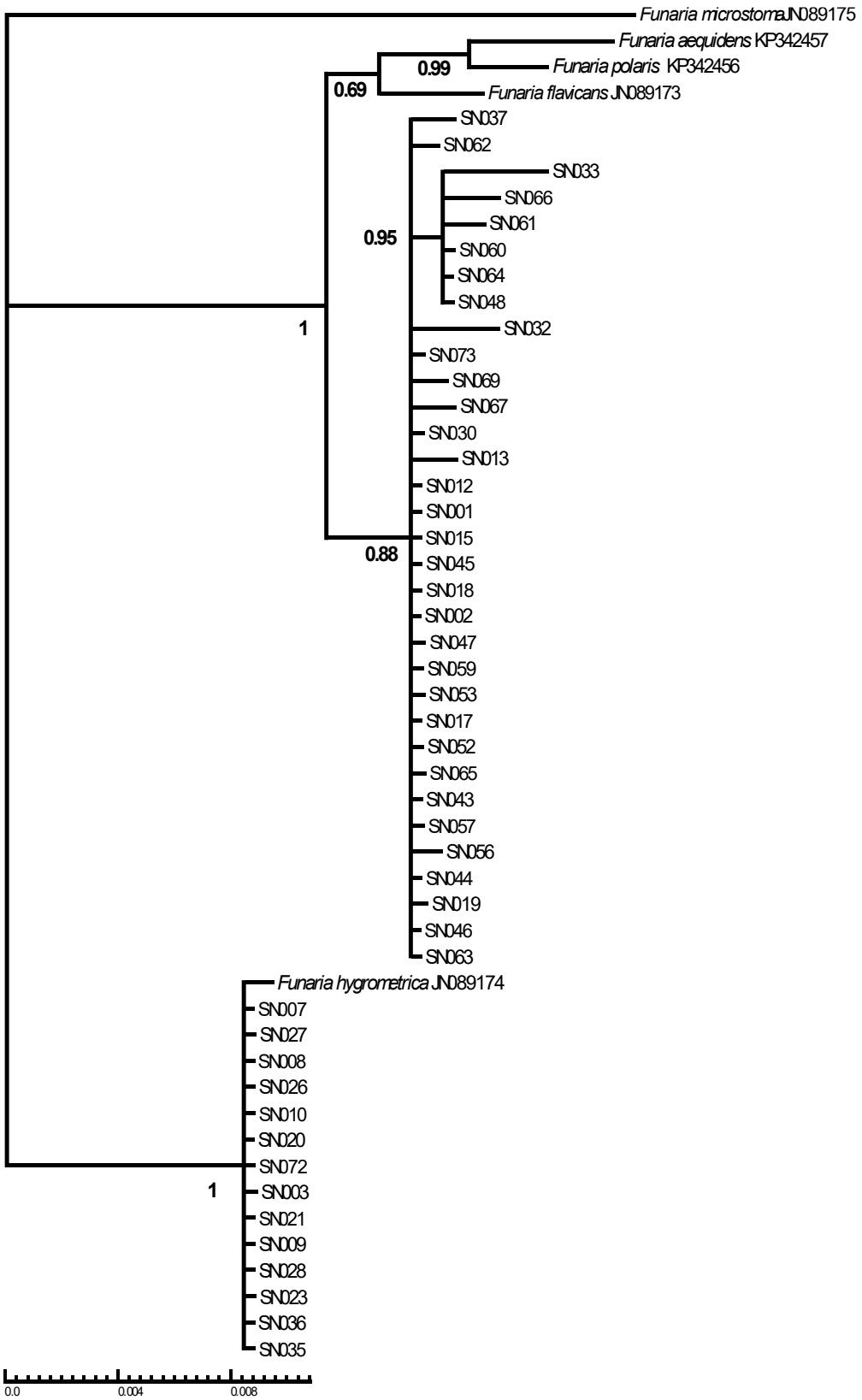



**Figure 8.** Unrooted tree based on nrITS sequences. The Sierra Nevada Mountains (SN) individuals are clearly separated in two clades, one of them closely related to a group of species formed by *F. aequidens*, *F. flavicans* and *F. polaris* while the second clade is closely related to a specimen of *F. hygrometrica* from Chile (JN089174). The bar at the bottom of the figure indicates the genetic distance. Numbers at the nodes indicate posterior probabilities estimated by MrBayes.

The sequence variation in the mt DNA was quite low. In addition to the already mentioned 45 bp indel there were 4 more indels of 1 bp length, but these were excluded because the default values of gappy deem 1 bp indels as unreliable. As a consequence, no phylogenetic tree based on the mt sequences was constructed. But there was a clear correlation of the nr ITS data with the mt data (results not shown). All individuals from the more numerous ITS clade shared the short version of the large mt indel and all specimens from the less numerous ITS clade, mostly restricted to higher elevations, shared the long version of the mt indel.

The phylogenetic analysis of the cp sequences (Fig. 9) produced exactly the same separation of individuals in two clades, with a posterior clade probability of 0.99. The mutations that support this separation are one indel of 17 bp length, one transition, and one transversion. Taken together these results show no indications of phylogenetic incongruence between nr, mt and cp datasets.

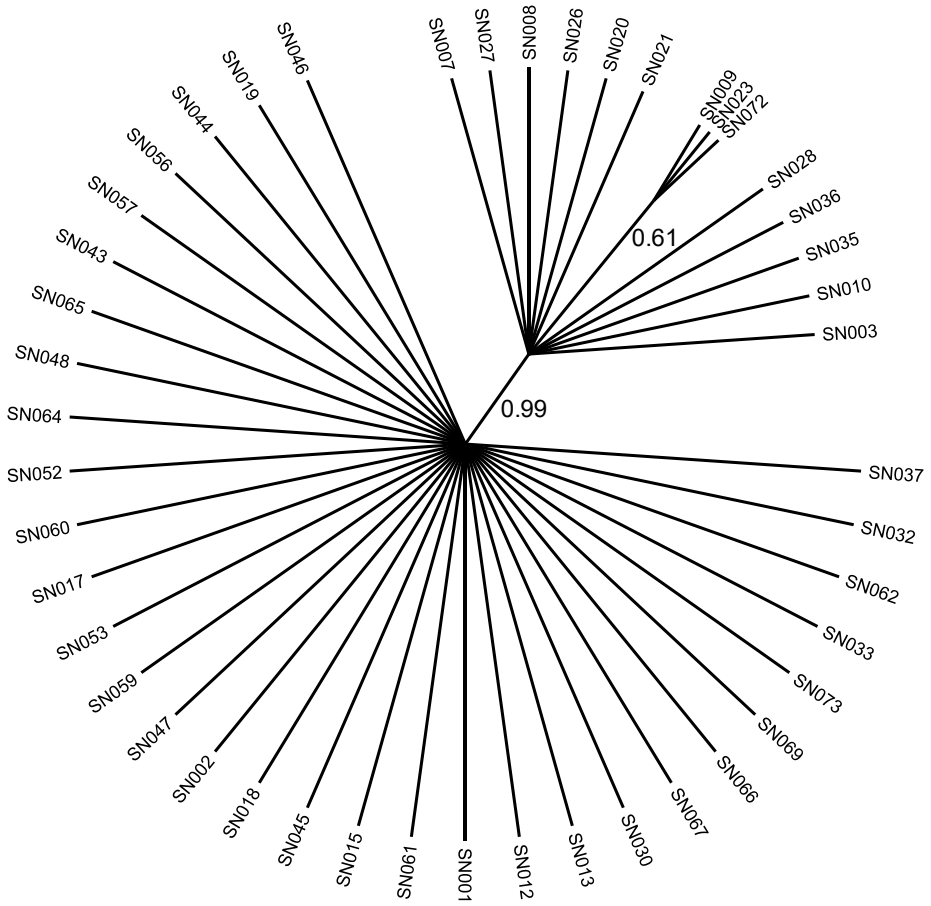



**Figure 9.** Unrooted tree based on the cp *rps3-rpl16* region. The two clades observed in the ITS tree separating the SN samples are reproduced with the cp data. Numbers at the clades indicate posterior probabilities estimated by MrBayes.

Based on this outcome we prepared finally a tree with the combined datasets of the nr ITS, cp *rps3-rpl16* , and mt *rpl5-rpl16* regions. The resulting tree confirms again the clear separation of the two clades, but with an almost absent structure within each clade (Fig. 10).

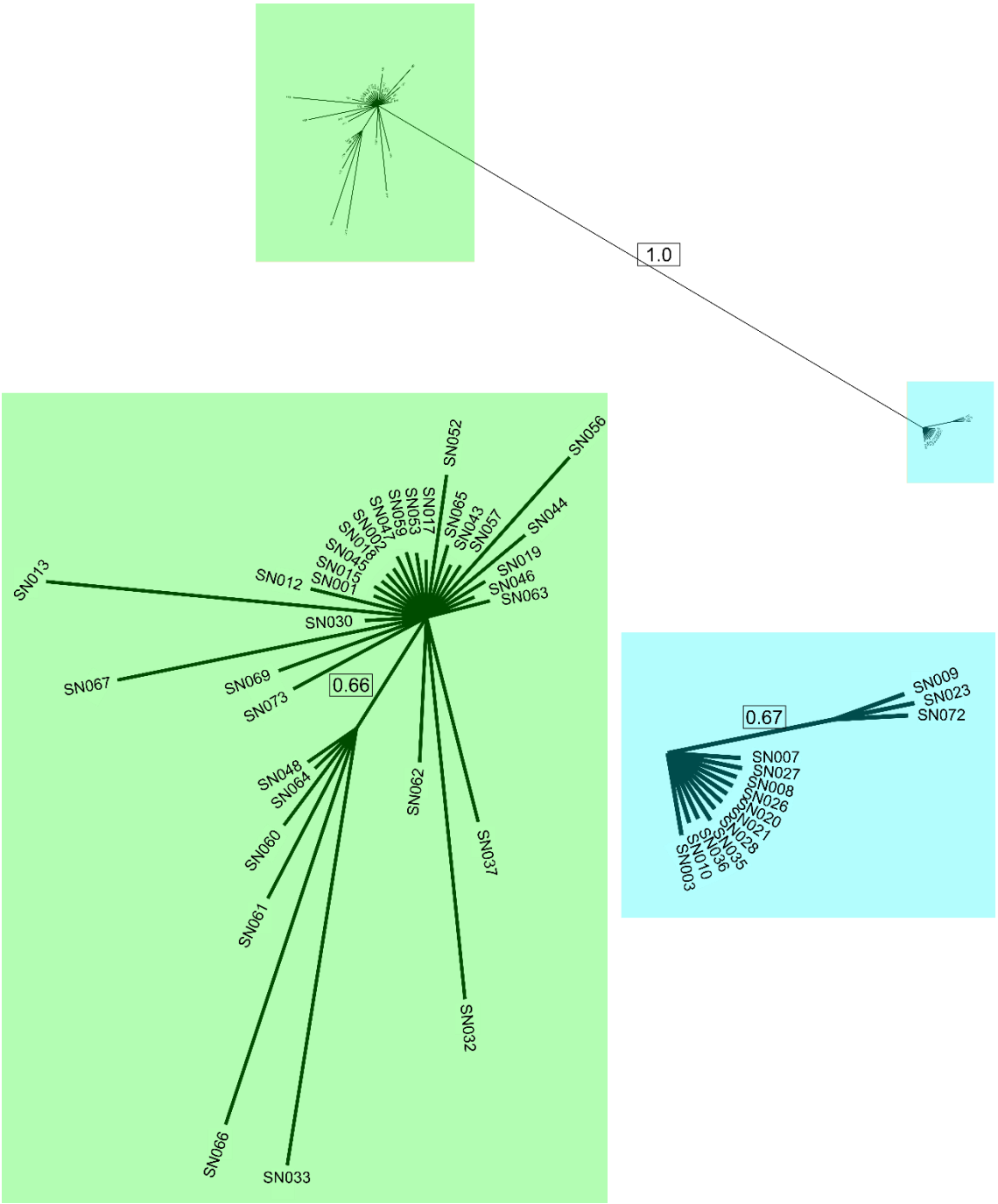



**Figure 10.** Unrooted Bayesian tree of the combined DNA sequence datasets. At the upper right part the clear distance between the two clades is evident and supported with a posterior probability of 1.0. At the lower part both clades are shown with more detail.

## GENOTYPING BY SEQUENCING AND POPULATION GENOMIC ANALYSES

A total of 192 individuals (nine of them technical replicates) and 784 million reads were sequenced for the GBS analyses. The dataset yielded an average of 3.3 million raw reads per individual, of which 2.7 million reads passed the quality filtering steps of ipyrad, with 550,000 reads per sample after filtering. After evaluation of the potential genotyping errors and individual levels of missing data through VCFtools, 17 individuals of *F. hygrometrica* that did not sequence well were removed: four from Asturias, two from Leon, four from Murcia, and seven from Sierra Nevada Mountains. Only variants successfully genotyped in 80% of individuals were kept, at a minor allele count of three. The final matrix comprised 171 individuals (five of them technical replicates, being a total of 166 different individuals), with 263 *loci* and 2.26% of missing data, which was used for the subsequent population genetic analysis.

PCA results for all individuals studied showed two clearly differentiated genomic groups (cluster A and cluster B) separated along the first component axis, with some outlier individuals (Fig. 11). When a sensitivity analysis procedure was carried out (for PCA and ancestry matrix), removing eight outlier individuals (AS006, AS008, AS012, MU002, SN013, SN033, SN041 and SN042) in terms of their genetic distance, the two genomic clusters remained consistent (Fig. 12).

When the SN samples were analysed separately from the rest of the study areas, without the four outlier individuals of this area (SN013, SN033, SN041 and SN042), therefore a total of 63 individuals analysed, again a clear separation of both genomic clusters along the first component was observed (Fig. 13). It is worth mentioning that, while the most numerous cluster seemed to exhibit low levels of variation (i.e. most of the specimens appear to be closely grouped in the PCA), the other cluster showed a larger variation, particularly along the second component axis (Fig. 13).

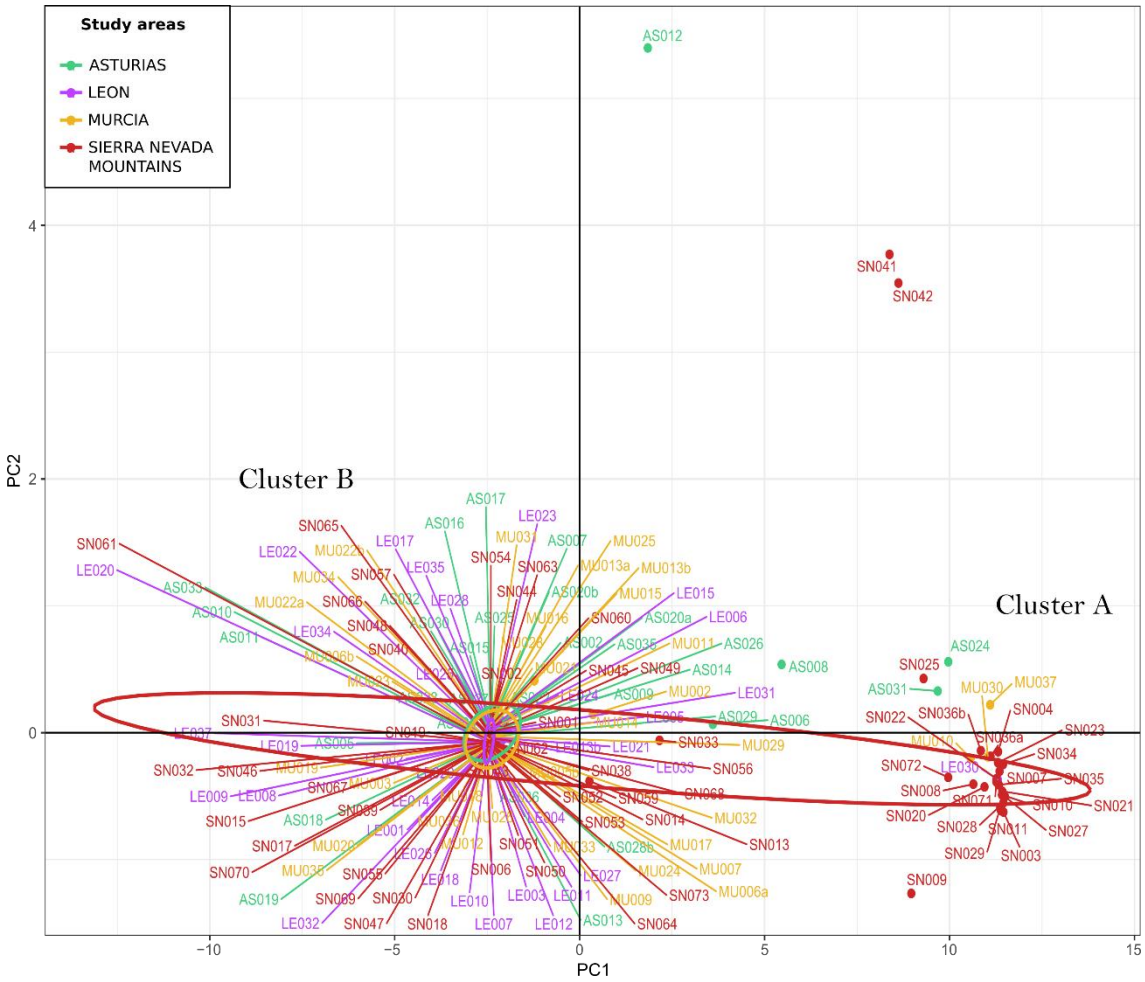



**Figure 11.** Principal component analysis (PCA) of multilocus genotypes for all individuals studied.

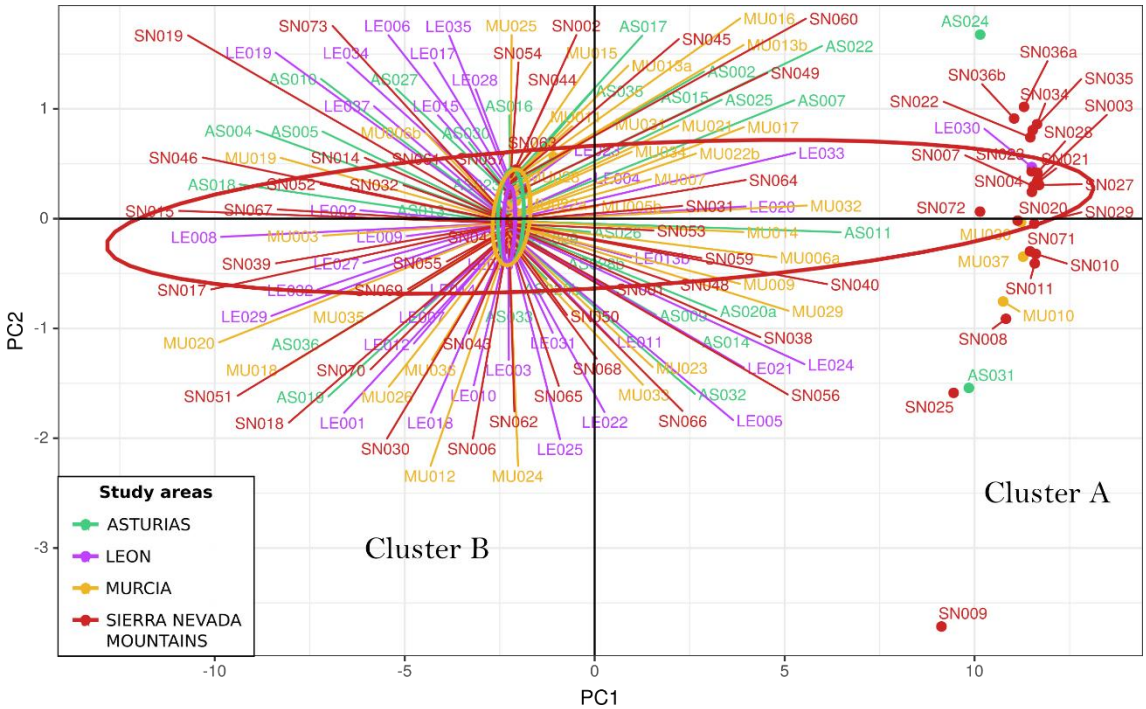



**Figure 12.** Principal component analysis (PCA), without outlier individuals (AS006, AS008, AS012, MU002, SN013, SN033, SN041 and SN042).

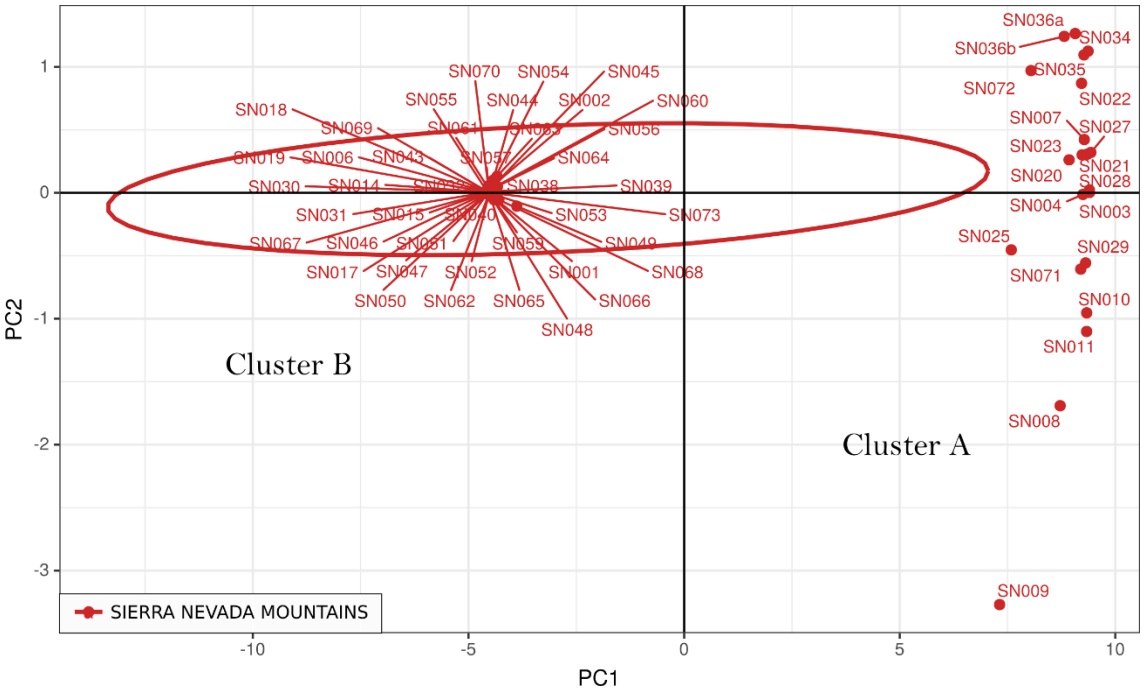



**Figure 13.** Principal component analysis (PCA) of SN individuals, without outliers of this area (SN013, SN033, SN041 and SN042).

Model-based ancestry estimation (sNMF cross-entropy) showed the lowest value for $K = 3$ when all samples were used and for $K = 2$ when outlier samples identified based on the PCA analysis were removed (Fig. 14). In both cases the most important change in cross-entropy was observed between $K = 1$ and $K = 2$ and the values did not change very much for $K > 2$. Therefore, $K = 2$ was chosen as the best number of clusters.

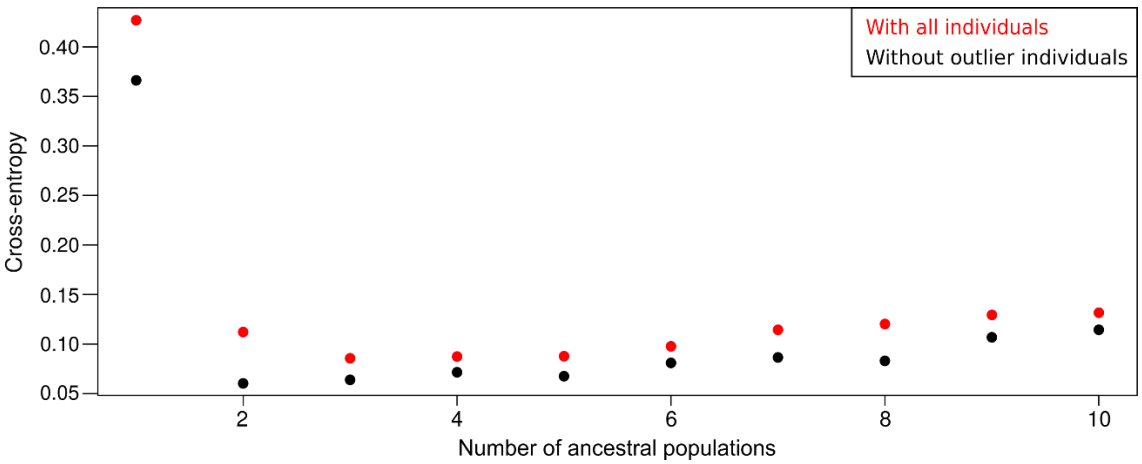



**Figure 14.** Values of the cross-entropy criterion to identify the best number of clusters for the individuals studied. In red, the values were obtained with all the individuals of the four study areas. In black, the values were obtained after removing eight outlier samples (AS006, AS008, AS012, MU002, SN013, SN033, SN041 and SN042).

Ancestry matrix results for all individuals studied (Fig. 15.a and Appendix 3.1), without outlier individuals (Fig. 15.b and Appendix 3.2) and focusing only on the SN populations (Fig. 15.c and Appendix 3.3), also showed two consistent genomic clusters. After removing the outlier individuals identified by PCA, all individuals were clearly assigned to one cluster (the major ancestry component ≥ 0.8).

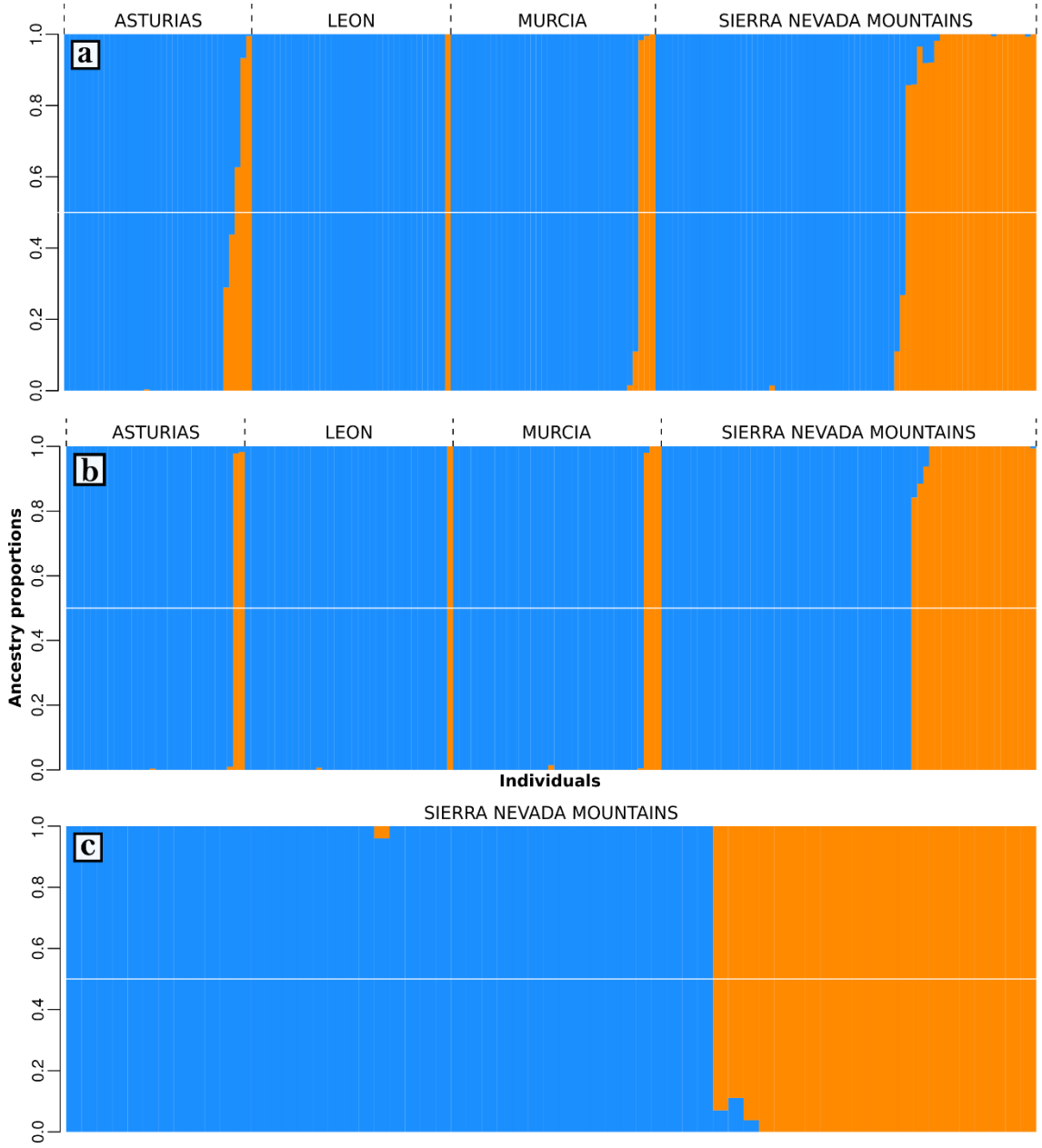



**Figure 15.** Ancestry assignment and individual clustering to ancestral populations A (orange) and B (blue). **a**: with all individuals studied. **b**: without outlier individuals (AS006, AS008, AS012, MU002, SN013, SN033, SN041 and SN042). **c**: only SN individuals, without outliers of this area (SN013, SN033, SN041 and SN042). They are ordered by study areas and assignment to each ancestral population. Abscissa axis indicate each individual according information given in Appendix 3.1 for graphic a, in Appendix 3.2 for graphic b, and in Appendix 3.3 for graphic c. Ordinate axis gives ancestry proportions for each graphic. The white line marks an ancestry proportion of 0.5.

MSN results indicated that individuals were not grouped according the geographic study area, which agrees with the results shown (Fig. 16). The group of individuals at the end of the right branch coincides with individual clustering to ancestral populations of cluster A (Fig. 15.a and Appendix 3.1).

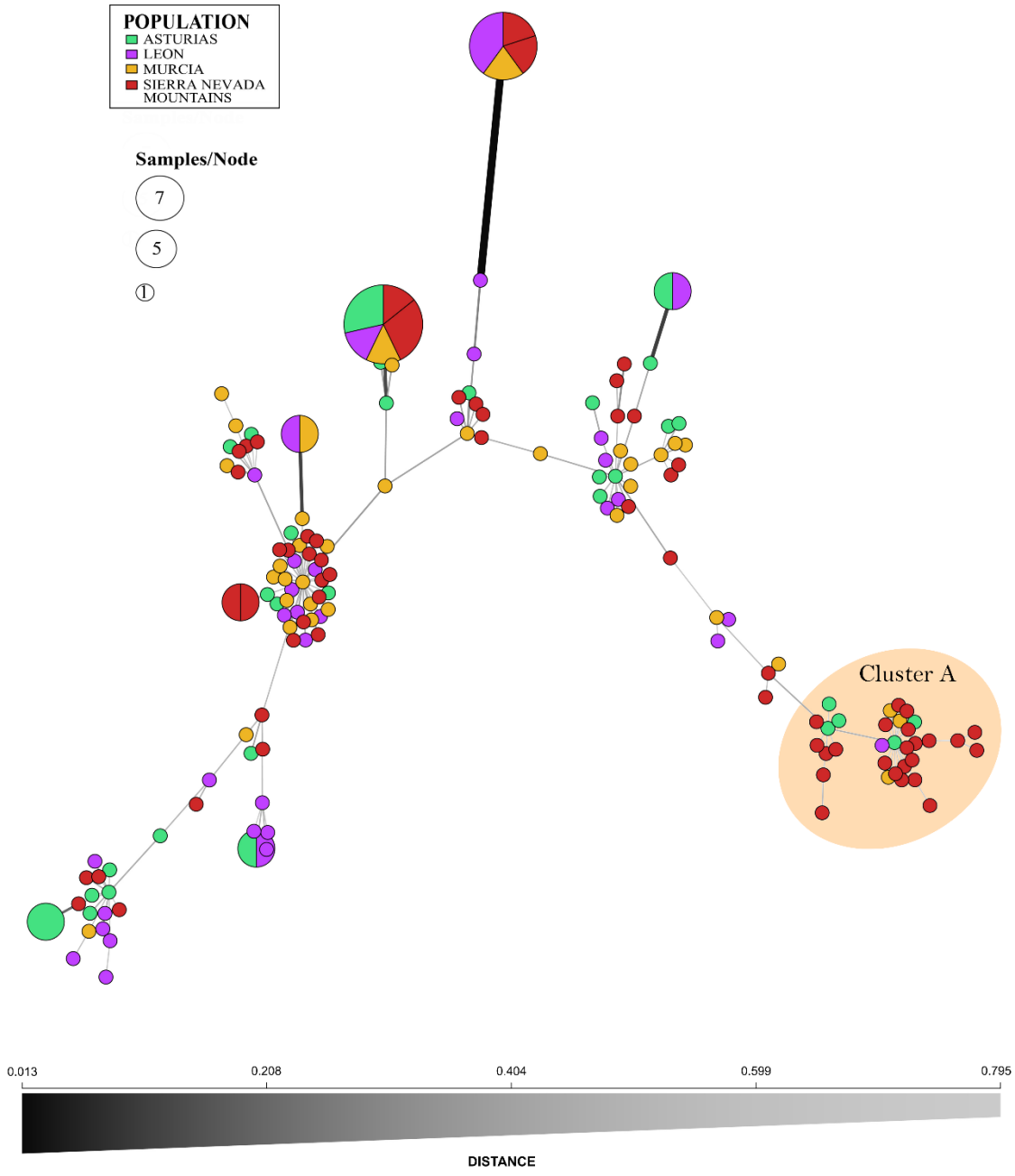



**Figure 16.** Minimum spanning network (MSN) analysis for all individuals studied. Each circle or pie represents an individual or group of individuals having the same SNP pattern. Colours indicate the belonging of each individual to any of the four areas studied. The genetic distance is represented by the lines between groups of individual samples, which become darker and thicker for more related nodes (according the scale bar at the bottom, which represents the Reynold genetic distance between individuals). The shaded ellipse on the right bottom includes the individuals that belong to cluster A, according to the ancestry assignment and individual clustering (Fig. 15.a and Appendix 13.1).

Therefore, each cluster was considered a different genotype within *F. hygrometrica*, in the four areas studied: genotype A and genotype B. In the Fig. 17 their geographical distributions are shown.

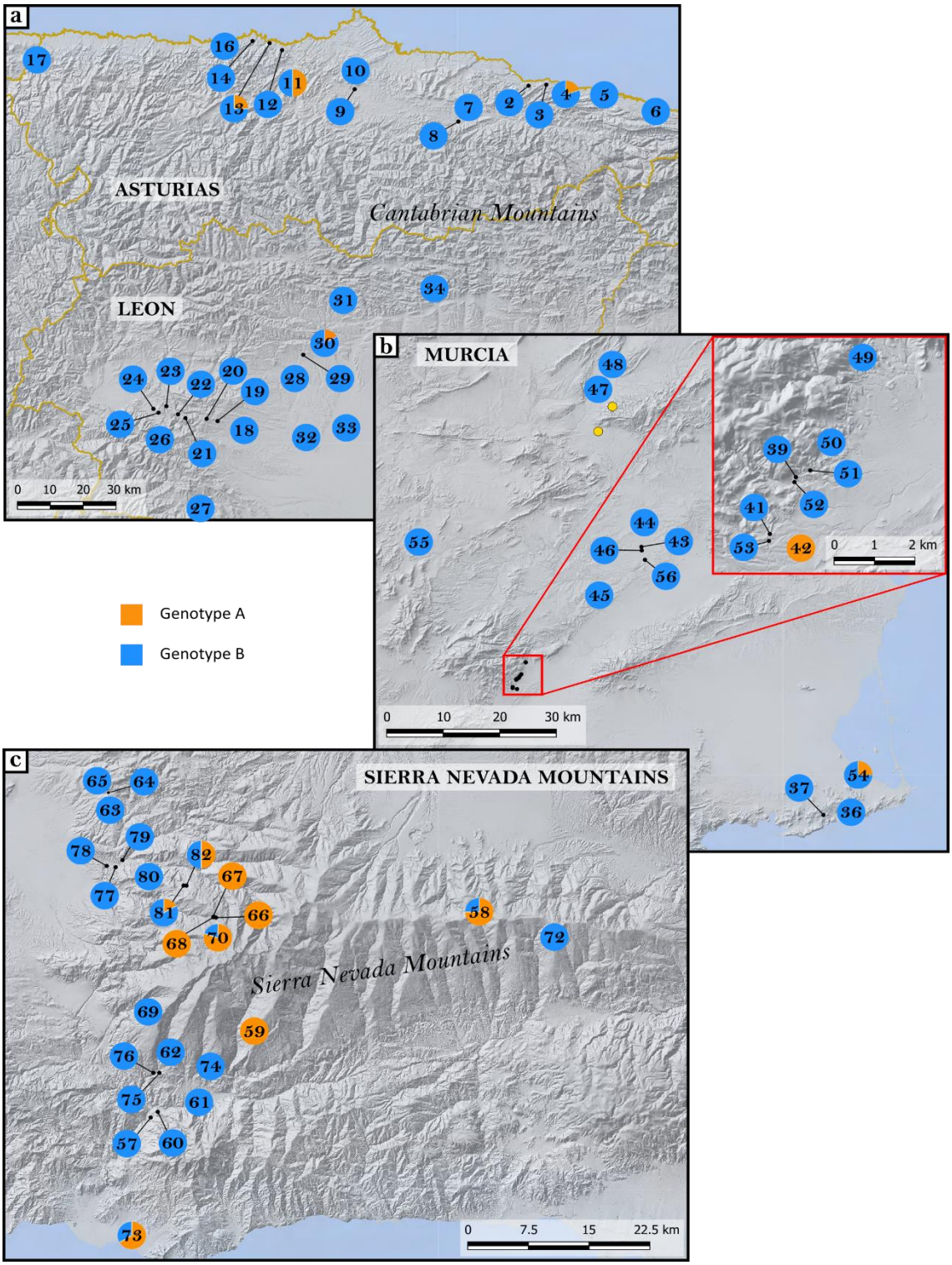



**Figure 17.** Locations sampled in the four study areas showing the percentage of the two genotypes detected within *F. hygrometrica*. **a**: Asturias and Leon locations. **b**: Murcia locations. **c**: Sierra Nevada Mountains locations. Genotype A = orange; genotype B = blue. For identification of locations number see Appendix 1.

While genotype A was poorly represented in Asturias, Leon and Murcia, genotype B resulted to be overrepresented across these areas. The two clearly differentiated genotypes were also found along the altitudinal gradient in the Sierra Nevada Mountains: one with a more restricted distribution (genotype A), which was found at high altitudes in Sierra Nevada Mountains (above 1600 m a.s.l.), and another more generalist (genotype B), which was found along the entire elevation gradient (Fig. 17). Most of the individuals studied above 2000 m (approximately 64%) showed genotype A. Below 1600 m, all individuals studied presented genotype B (except two individuals found at 58 m, collected in a plant nursery under palm trees, which were finally considered as outliers). In the rest of the studied areas (provinces of Asturias, Leon and Murcia), genotype B was found to be predominant, while genotype A appears very sporadically, with 1-4 individuals for each area.

The genotype of each individual was compared with their cytotype obtained through FCM (Fig. 18 and Table 3), showing that only the cytotypes b and c seem to belong to a unique genotype, B and A, respectively.

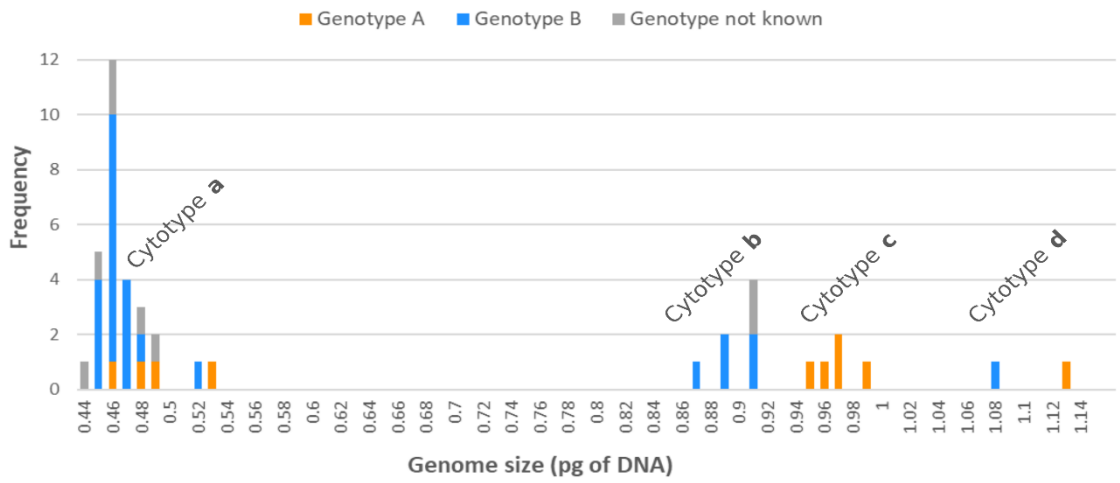



**Figure 18.** Histogram of genome sizes of the *F. hygrometrica* individuals studied generated by FCM, taking into account the genotype of each individual.

The Tajima's D tests were non-significantly positive in general (with all the specimens together across the four areas; D = 0.451, p-value > 0.243) and in the case of the SN populations (D = 0.158, p-value > 0.209). In turn, Asturias was not significantly negative (D = -0.257, p-value > 0.211), with the exception of Leon (D = -0.823, p-value > 0.032) and Murcia (D = -0.278, p-value > 0.044), which were significantly negative, respectively.

**Table 3.** Genome size, cytotype and genotype for all the individuals analysed by FCM. Indivuals are coded by acronyms and sequencing code (AS: Asturias province; LE: Leon province; MU: Murcia province; SN: Sierra Nevada Mountains; see Appendix 1). Also location number and altitude are indicated. Cytotype and genotype columns follows a colour key in order to highlight the belonging of each individual to each one. Individuals of the same location are separated by a thick line (*: individual MU008 showed to present genotype A in a preliminary analysis).

| Individuals code | Mean DNA value (pg) | Cytotype | Genotype | Location | Altitude (m a.s.l.) |
|---|---|---|---|---|---|
| **AS001** | 0.436 | a | - | L01 | 83 |
| **AS002** | 0.444 | a | B | L02 | 107 |
| **AS003** | 0.445 | a | - | L02 | 107 |
| **AS004** | 0.514 | a | B | L03 | 30 |
| **AS005** | 1.077 | d | B | L04 | 5 |
| **AS008** | 1.124 | d | A | L04 | 5 |
| **AS010** | 0.473 | a | B | L05 | 36 |
| **AS011** | 0.456 | a | B | L06 | 32 |
| **AS012** | 0.886 | b | B | L06 | 32 |
| **AS013** | 0.454 | a | B | L06 | 32 |
| **AS014** | 0.458 | a | B | L07 | 329 |
| **AS018** | 0.452 | a | B | L10 | 97 |
| **AS019** | 0.458 | a | B | L10 | 97 |
| **AS021** | 0.909 | b | - | L10 | 97 |
| **AS022** | 0.461 | a | B | L10 | 97 |
| **AS025** | 0.459 | a | B | L12 | 155 |
| **AS026** | 0.456 | a | B | L12 | 155 |
| **AS030** | 0.462 | a | B | L13 | 140 |
| **AS031** | 0.455 | a | A | L13 | 140 |
| **AS032** | 0.454 | a | B | L14 | 76 |
| **AS034** | 0.459 | a | - | L15 | 127 |
| **AS036** | 0.461 | a | B | L17 | 35 |
| **MU002** | 0.885 | b | B | L36 | 150 |
| **MU003** | 0.870 | b | B | L37 | 114 |
| **MU005** | 0.445 | a | B | L39 | 346 |
| **MU008** | 0.526 | a | A* | L40 | 350 |
| **MU010** | 0.984 | c | A | L42 | 350 |
| **SN002** | 0.461 | a | B | L57 | 287 |
| **SN001** | 0.904 | b | B | L57 | 287 |
| **SN009** | 0.969 | c | A | L59 | 1965 |
| **SN012** | 0.453 | a | - | L60 | 398 |
| **SN013** | 0.449 | a | B | L61 | 320 |
| **SN015** | 0.449 | a | B | L62 | 687 |
| **SN016** | 0.908 | b | - | L63 | 1086 |
| **SN018** | 0.456 | a | B | L64 | 1102 |
| **SN019** | 0.905 | b | B | L65 | 1170 |
| **SN021** | 0.962 | c | A | L66 | 2325 |
| **SN022** | 0.477 | a | A | L66 | 2325 |
| **SN023** | 0.947 | c | A | L66 | 2325 |
| **SN024** | 0.475 | a | - | L66 | 2325 |
| **SN025** | 0.487 | a | A | L67 | 2275 |
| **SN026** | 0.484 | a | - | L67 | 2275 |
| **SN029** | 0.958 | c | A | L68 | 2240 |

Landscape resistance analyses revealed that patterns of genomic relatedness among individuals within each area were best explained by different variables or scenarios (Table 4). An IBD model, where geographic isolation is favoured, and IBE model, where environmental isolation is favoured, were rejected in the set of best supported models for populations across Asturias, Leon and Murcia. In these three areas, a null model was favoured, likely pointing to a scenario where all the sampled populations form a large, panmictic population. In contrast, our model selection procedure results showed that the resistance surface defined by the climatic variable isothermality (i.e. isothermality quantifies how large the day-to-night temperatures oscillate relative to the summer to winter [annual] oscillations) was the best-supported model in Sierra Nevada Mountains (Table 4). Results showed that the combined surfaces were not the best supported model in any of the areas analysed. We performed sensitivity analyses in order to estimate the resistance surfaces with the subsets of data (SN1: without outliers SN041 and SN042; and SN2: without outliers SN013, SN033, SN041 and SN042). We obtained consistent results (Table 4), revealing that isothermality was the best fitting resistance surface explaining the spatial genetic structure within Sierra Nevada Mountains.

**Table 4.** Model selection results for both individual and composite surfaces for *F. hygrometrica* across the four study areas: Asturias, Leon and Murcia provinces, and Sierra Nevada Mountains. SN1 and SN2 represent two subsets of data from Sierra Nevada Mountains: SN1, without outliers SN041 and SN042; and SN2, without outliers, SN013 and SN033, SN041 and SN042). Only the two highest ranking models are presented in each case, with models identified as being optimal (ΔAICc ranges between 0–2) highlighted in bold. K = number of parameters; AICc = Akaike information criteria for small sample size; ΔAICc = delta AICc; Average weight = AICc weight across the 1,000 bootstrap iterations; Percentage top = percentage of the 1,000 bootstrap iterations in which each model was selected as the top model.

| Study areas | Surface | K | AIC | AICc | ΔAICc | Average weight | Percentage top |
|---|---|---|---|---|---|---|---|
| Asturias | **Null** | **1** | **-335.030** | **-338.901** | **0** | **0.673** | **59.4** |
| | Distance | 2 | -333.280 | -336.880 | 2.021 | 0.245 | 35.4 |
| Leon | **Null** | **1** | **-4183.930** | **-4187.808** | **0** | **0.695** | **39.3** |
| | Distance | 2 | -4181.930 | -4185.555 | 2.253 | 0.225 | 20.9 |
| Murcia | **Null** | **1** | **-3754.689** | **-3750.773** | **0** | **0.700** | **52.3** |
| | Distance | 2 | -3752.459 | -3748.648 | 2.125 | 0.120 | 20.4 |
| Sierra Nevada Mountains | **Isothermality** | **4** | **1639.423** | **1640.068** | **0** | **1** | **99.9** |
| | Temperature* | 4 | 1695.992 | 1696.637 | 56.569 | 5.20E-13 | 0.1 |
| SN1 | **Isothermality** | **4** | **1683.069** | **1683.736** | **0** | **0.996** | **99.9** |
| | Temperature* | 4 | 1749.127 | 1749.794 | 66.057 | 0.003 | 0.1 |
| SN2 | **Isothermality** | **4** | **1694.289** | **1694.761** | **0** | **0.996** | **99.9** |
| | Temperature* | 4 | 1755.340 | 1755.945 | 61.184 | 0.003 | 0.1 |

(*) Annual mean temperature

## TRANSCRIPTOME STUDY

For the 60 samples used for RNAseq, 29.8 – 41.3 million of raw reads were obtained, of which around 98% remained after quality filtering, adaptor trimming and short-read removal. In the assembly process without a reference genome, a total of 163,237 transcripts and 163,088 unigenes were obtained, of which about 27.8% had a length of 1,000-2,000 bp (Table 5). After *de novo* transcriptome reconstruction, a mean of 86% of reads were mapped per sample.

**Table 5.** Overview of the number of transcripts and unigenes in different length intervals.

| Transcript length interval | 200-500 bp | 500-1 kbp | 1k-2 kbp | >2 kbp | Total |
|---|---|---|---|---|---|
| Number of transcripts | 37,258 | 41,083 | 45,414 | 39,482 | 163,237 |
| Number of Unigenes | 37,113 | 41,079 | 45,414 | 39,482 | 163,088 |

According to the previous results of GBS and Sanger sequencing, these samples were divided into five groups, namely AC, AT, BC, BT and RU (see Appendix 2). The AC and BC groups were formed by samples with genotype A and genotype B, respectively, subjected to control condition. The AT and BT groups were formed by the same previous individuals but subjected to thermal stress. The RU group, consisted exclusively of four samples from species other than *F. hygrometrica* (*F. aequidens*, *F. arctica* and *F. polaris*), which were only subjected to control condition. A PCA plot transcription levels (Fig. 19) showed that the first component (36.18 %) separated the A and B genotypes very clearly while the second component (14.78%) separated the control and the heat stressed samples. Only the specimen SN018 was located at an intermediate position in between A and B genotypes, although according to GBS and Sanger sequencing data it belongs to the B genotype group. Interestingly, the three species from Russia (RU) were closely together and located near the control B genotype specimens.

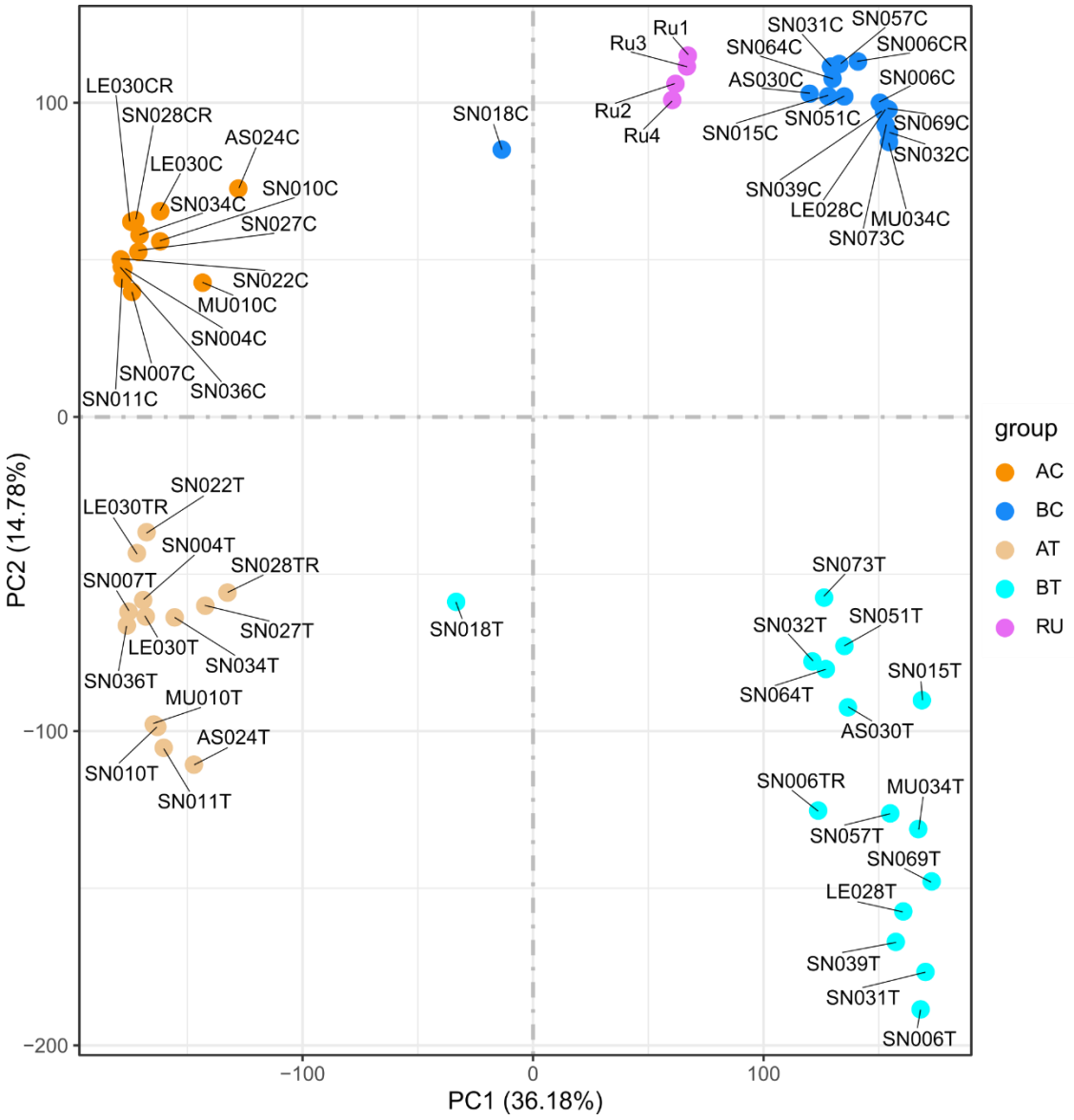



**Figure 19.** Principal Component Analysis (PCA) in 2D of different gene expression levels between samples. Abbreviations of groups within *F. hygrometrica* indicate: AC, genotype A in control condition; BC, genotype B in control condition; AT, genotype A in thermal stress; BT, genotype B in thermal stress. RU group includes samples from other species of *Funaria*, in control condition.

In Venn diagrams (Fig. 20), the *loci* expressed by each group and all the possible overlapping of *loci* sets among the different groups are shown. The groups of samples subjected to thermal stress condition show a slightly higher number of genes expressed with respect to the control condition, being this difference greater in samples with genotype B (AC = 44,652, BC = 42,982, AT = 45,067, BT = 47,758). Furthermore, the RU group showed the lowest number of expressed genes (33,850), but this may be due to the lower number of samples. Groups AC, BC, AT and BT had around 20% of the genes analysed in common, being expressed in both environmental

conditions and in both genotypes. About 29% of the genes were expressed in all the samples of control groups (including RU). As in Fig. 19, RU samples appear to be more closely related to genotype B than genotype A, since 79% of the expressed genes are in common with BC, while 64% are in common with AC.

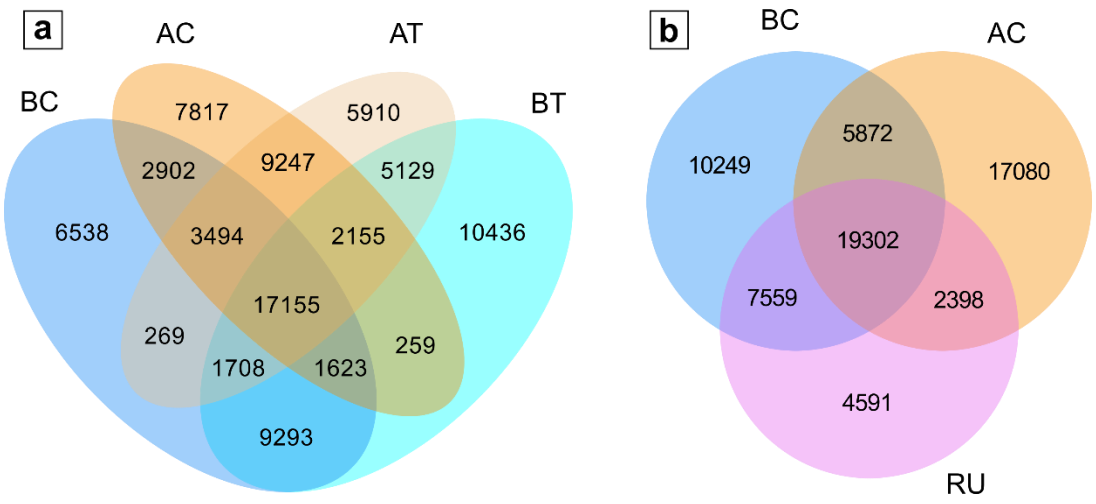



**Figure 20.** Venn diagrams of gene coExpression. **a**: shows the number of *loci* expressed by each group and all the possible overlapings of *loci* sets among the different groups within *F. hygrometrica* with A and B genotypes respectively subjected to thermal stress (T) and control (C) conditions. **b**: shows the three groups of samples that were subjected to a control condition (A genotype, B genotype and Russian samples. These diagrams show the relationships between sets of elements by means of closed lines or overlaped areas. In diagram a, 17,155 genes are found in the four transcriptomic groups, while 5,129 genes are shared by the AT and BT groups (subjected to thermal stress), but are not found in any of the other two control groups (AC and BC). In diagram b, it can be highligted that 19,302 genes are shared by AC, BC and RU groups, while 7,559 genes are shared by BC and RU groups. Abbreviations of groups as in Fig. 19.

Volcano plots were used to infer the overall distribution of differentially expressed genes (Fig. 21). A large number of genes showed significant differential expressions between samples of treatment group and control group (BT *vs* BC, AT *vs* AC), both over and under expressed genes. However, this difference was even greater when differentially expressed genes were compared between groups of samples that were subjected to the same environmental condition, but with different genotypes (BC *vs* AC, BT *vs* AT), which is accordance with the PCA data. This is further confirmed by the heatmap of the correlation coefficients of the gene expression levels (Fig. 22). According to Fig. 23, if differentially expressed genes are compared between RU and the other two control groups (AC and BC), the number of upregulated genes is slightly higher than the downregulated ones.

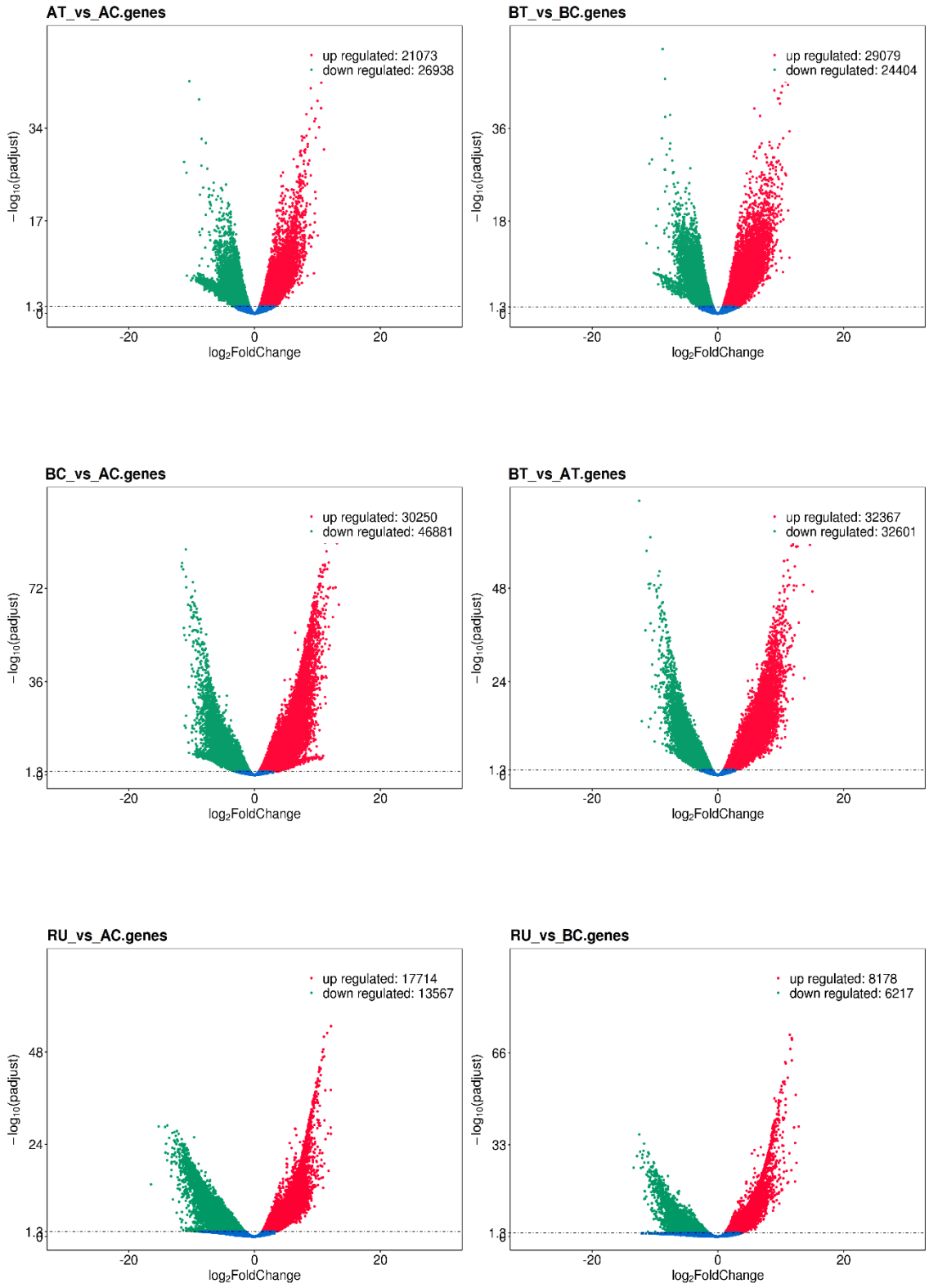



**Figure 21.** Volcano plots for differentially expressed genes. The abscissa axis shows the fold change in gene expression between different samples, and the ordinate axis shows the statistical significance of the differences. Significantly up- and downregulated genes are highlighted in red and green, respectively. Genes that did not express differently between treatment group and control group are in blue. Abbreviations of groups within *F. hygrometrica* as in Fig. 19.

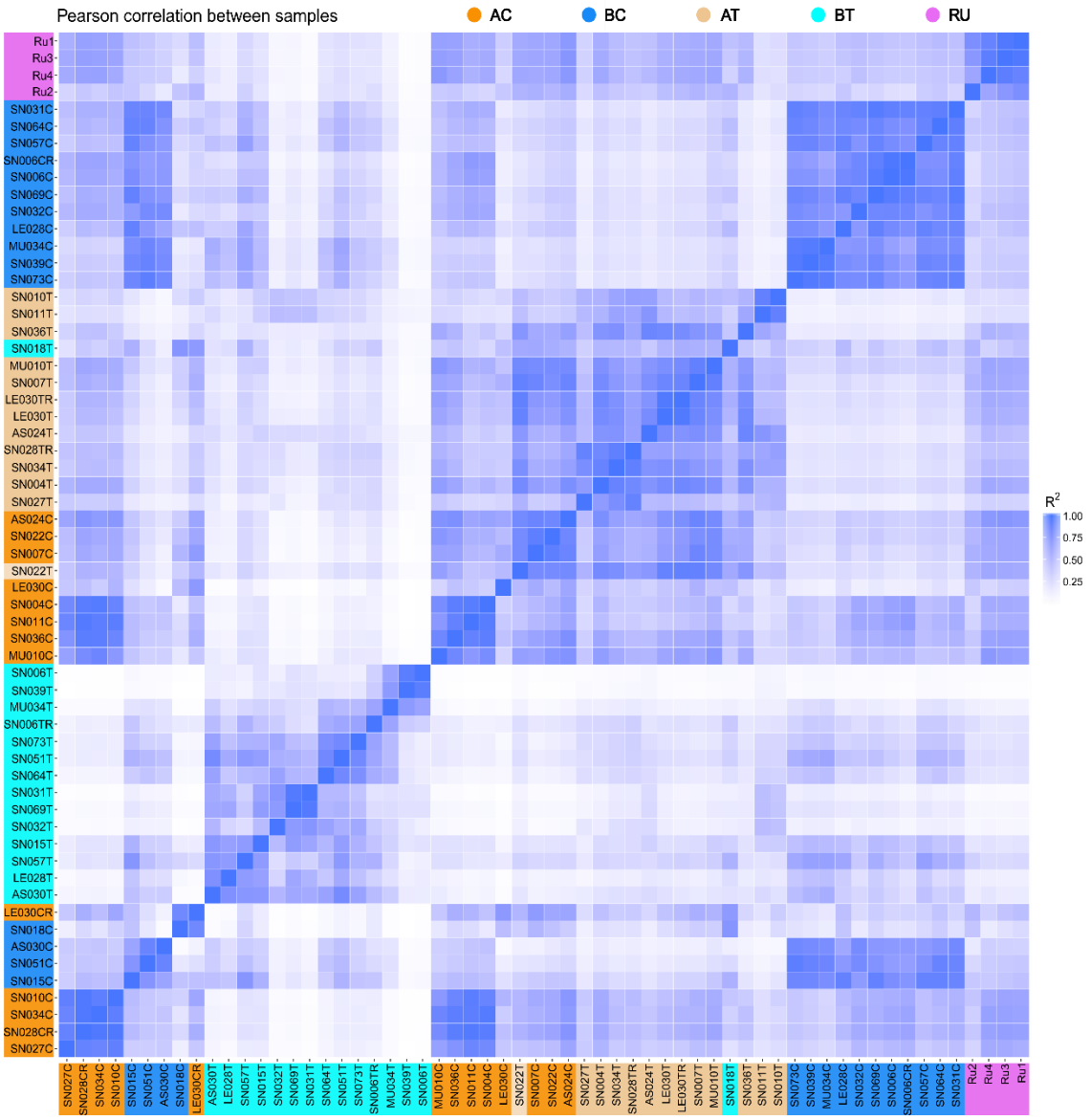



**Figure 22.** Different gene expression levels under different experiment conditions between samples. Abbreviations of groups within *F. hygrometrica* as in Fig. 19.

The count of genes differentially expressed between different transcriptomic groups is shown in Fig. 23, comparing treatment group with control group and with the same genotype (AT *vs* AC, BT *vs* BC), subjected to the same environmental condition (BC *vs* AC, BT *vs* AT), and finally the RU group with the other two control groups (RU *vs* AC, RU *vs* BC). A greater difference in genetic expression is also observed here between groups of samples subjected to the same environmental condition, but different genotypes. The analyses were carried out with all differentially expressed genes together (all-DE genes) and with up- and downregulated genes separately.

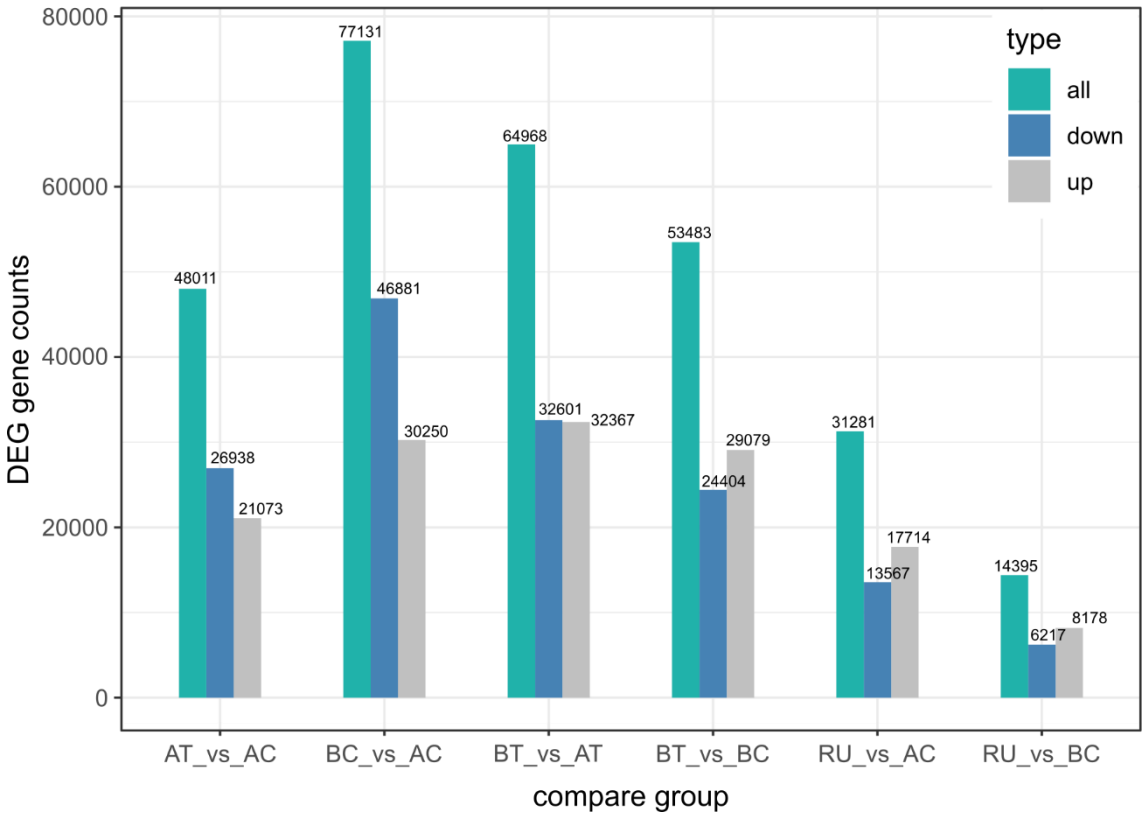



**Figure 23.** Differentially expressed gene count, comparing transcriptomic groups with each other. Abbreviations of groups within *F. hygrometrica* as in Fig. 19.

After performing the statistical analysis for enrichment, using as gene annotation Gene Ontology (GO), many biological functions were found associated with these differential expressed genes. Interestingly, with all-DE genes, there was no significant enrichment in comparison BC *vs* AC (genotype B *vs* genotype A under control conditions). This hardly changed when up- and downregulated genes were analysed separately, although here there was significant enrichment in four GO categories with upregulated genes and in two with downregulated genes. This changed dramatically when the samples were under heat stress (BT *vs* AT), when 17 GO categories showed significant enrichment analysing all-DE genes (14 GO categories with upregulated genes and 11 GO categories analysing downregulated genes, when separately). These categories include cellular component (ribosome, protein-containing complex, cytoplasm, organelle, intracellular, cell, thylakoid, mitochondrion, cell wall), molecular function (structural component of ribosome, structural molecular activity, unfolded protein binding) and biological process (translation, protein folding, photosynthesis, cellular protein modification process, homeostatic process). The heat stress did not lead to significant GO enrichment within the genotype A samples (AT

*vs* AC) when all-DE genes were analysed together. However, when the analysis was performed with the genes (up- and downregulated) separately, there were many categories with significant enrichment (19 categories with upregulated genes and eight categories with downregulated genes). Within the genotype B samples (BT *vs* BC) and with all-DE genes, the heat stress affected to a small degree with significant enrichment in three aspects, one biological process (cellular protein modification process) and two molecular functions (kinase activity, unfolded protein binding). In contrast with the above, 20 GO categories with upregulated genes and two categories with downregulated genes had significant enrichment when up- and downregulated genes were analysed separately.

The comparison of the Russian samples, belonging to the closely related species *F. aequidens*, *F. arctica* and *F. polaris*, showed some enrichment in comparison with the genotype B samples (five categories: thylakoid, ribosome, photosynthesis, structural constituent of ribosome, structural molecule activity). The enrichment was much more conspicuous in comparison to the genotype A samples with 15 categories. This was also observed when separately, with significant enrichment in nine categories and 12 categories in up- and downregulated genes (RU *vs* AC), respectively; and in five categories and three categories in up- and downregulated genes (RU *vs* BC), respectively.

The cellular components with significant enrichment were: ribosome, thylakoid, cytoplasm, protein-containing complex, cell wall, external encapsulating structure, intracellular, organelle and cell. The molecular functions are: structural constituent of ribosome and structural molecule activity. And finally the biological processes with enrichment are: translation, photosynthesis, reproduction and cell wall organization or biogenesis. More detail the GO categories with significant enrichment are given in the following table, when up- and downregulated genes were analysed separately (Table 6).

**Table 6.** Gene Ontology (GO) enrichment analysis results with biological processes, cellular components and molecular functions significantly over- and underexpressed, comparing the transcriptomic groups with each other. Abbreviations of groups within *F. hygrometrica* as in Fig. 19.

| | **AT *vs* AC** | | **BT *vs* BC** | | **BC *vs* AC** | | **BT *vs* AT** | | **RU *vs* AC** | | **RU *vs* BC** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **up** | **down** | **up** | **down** | **up** | **down** | **up** | **down** | **up** | **down** | **up** | **down** |
| Biosynthetic process | | | | | | | | | | ✓ | | |
| Cell | ✓ | | ✓ | | | | ✓ | ✓ | | ✓ | | |
| Cell morphogenesis | ✓ | | | | | | | | | | | |
| Cell wall | | | | | | | | | ✓ | | | |
| Cell wall organization or biogenesis | | | | | | | | | ✓ | | | |
| Cellular nitrogen compound metabolic process | | | ✓ | | | | | | | | | |
| Cellular protein modification process | | | | ✓ | | | | ✓ | | | | |
| Cytoplasm | ✓ | | ✓ | | | | ✓ | ✓ | | ✓ | | |
| DNA−binding transcription factor activity | | ✓ | | | | | | | | | | |
| Enzyme binding | ✓ | | | | | | | | | | | |
| Extracellular matrix organization | ✓ | | ✓ | | | | | | | | | |
| Generation of precursor metabolites and energy | | | | | | | ✓ | | | | | |
| GTPase activity | ✓ | | | | | | | | | | | |
| Helicase activity | ✓ | | | | | | | | | | | |
| Histone binding | ✓ | | ✓ | | | | | | | | | |
| Homeostatic process | | | ✓ | | | | | | | | | |
| Intracellular | ✓ | | ✓ | | | | ✓ | ✓ | | ✓ | | |
| Kinase activity | | | | ✓ | | | | | | | | |
| Lipid metabolic process | | ✓ | | | | | | | | | | |
| Microtubule organizing center | | ✓ | | | | | | | | | | |
| Mitochondrion | | | | | | | | ✓ | | | | |
| mRNA processing | ✓ | | ✓ | | | | | | | | | |
| Nucleocytoplasmic transport | ✓ | | | | | | | | | | | |
| Nucleoplasm | ✓ | | | | | | | | | | | |
| Nucleotidyltransferase activity | | | ✓ | | | | | | | | | |
| Nucleus | | ✓ | | | | ✓ | | | | | | |
| Organelle | | | ✓ | | | | ✓ | ✓ | | ✓ | | |
| Oxidoreductase activity | | ✓ | | | | ✓ | | | | | | |
| Photosynthesis | | | | | | | ✓ | | ✓ | ✓ | ✓ | |
| Protein folding | ✓ | | ✓ | | | | ✓ | | | | | |
| Protein−containing complex | ✓ | | ✓ | | | | ✓ | ✓ | ✓ | ✓ | | |
| Ribonucleoprotein complex assembly | ✓ | | ✓ | | | | | | | | | |
| Ribosome | ✓ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| RNA binding | | | ✓ | | | | | | | | | |
| Structural constituent of ribosome | ✓ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Structural molecule activity | ✓ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Thylakoid | | | | | | | ✓ | | ✓ | ✓ | ✓ | |
| Translation | | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | | |
| Transmembrane transport | | ✓ | | | | | | | | | | |
| Transmembrane transporter activity | | ✓ | | | | | | | | | | |
| Transport | | ✓ | | | | | | | | | | |
| Unfolded protein binding | ✓ | | ✓ | | | | ✓ | | | | | |
| Vesicle−mediated transport | | | ✓ | | | | | | | | | |
| **Total number** | **19** | **8** | **20** | **2** | **4** | **2** | **14** | **11** | **9** | **12** | **5** | **3** |

## BIOMETRIC STUDY

The biometric data of the 88 SN plants used for the biometric study were grouped for the statistical analyses according to the GBS results, which, as mentioned before, showed two clear genotype groups: genotype A, exclusive from SN highlands (43 plants), and the generalist genotype B, present along the whole mountain system (45 plants). Additionally the altitude where the samples were collected, considered as an ecological factor, was also taken into account for the genotype B in order to separate the highland populations (between 1500 m and 2700 m a.s.l.) and the lowland populations (below 1500 m a.s.l.). Therefore, in some analyses three groups of plants were considered: i) with genotype A, exclusive from SN highlands (43 plants); ii) with genotype B from SN highlands (24 plants); and iii) with genotype B from SN lowlands (21 plants) (Table 7).

The 18 plants outside SN, belonging to *F. aequidens*, *F. arctica*, *F. flavicans, F. hygrometrica* and *F. polaris* were separated in five different groups, one for each *Funaria* species (Table 8). It is worth mentioning that three plants of *F. hygrometrica* from Chile were also included in the biometric analysis.

From the 29 morphological characters studied, five had the same values among all plants analysed (PM, PMR, CSH, CST and ANN; Tables 7 and 8). All the plants observed had entire phyllidium margins at upper part of leaf, zero rows of narrower cells than the other phyllidium laminal cells, all capsules had a pyriform shape, striae present in mature dry closed capsule, and presence of annulus revoluble. These characters were removed for data analyses. Measurements of characteristics PPL and PPL/PML were also removed from the analyses, since they were only useful, together with PML/PW, to determine the shape of the leaf (as explained in Material and Methods section).

**Table 7.** Statistical results of quantitative (QT) and qualitative (QL) characters of *F. hygrometrica* specimens from Sierra Nevada Mountains included in the biometric study. Character abbreviations and status character are given in Table 1. For QT mean values ± SD [range] are given. For QL percentage of each status character (identified in brackets) is presented. The second column gives the values for all samples with genotype B and the third column the values for genotype A. The forth and fifth colums give the values for genotype B separated according to altitude. For the PCA and LDA analyses, the genotype B was alternatively grouped together (indicated by the green line) or samples were separated according to altitude (purple line). The number (n) of fructified plants examined for each group of samples is given. All measurements are given in µm, except for SEL in mm.

| Abbreviations of characters | SN highlands and lowlands Genotype B (n = 45) | SN highlands | | SN lowlands |
|---|---|---|---|---|
| | | Genotype A (n = 43) | Genotype B (n = 24) | Genotype B (n = 21) |
| CAL | 3139.08 ± 1134.66 [1850.00-7900.00] | 3414.09 ± 886.34 [2083.00-5693.00] | 2691.29 ± 602.78 [1850.00-4223.00] | 3421.92 ± 1236.05 [2318.00-7900.00] |
| PP | 42.96% (2)<br>57.04% (3) | 81.40% (2)<br>18.60% (3) | 37.50% (2)<br>62.50% (3) | 49.21% (2)<br>50.79% (3) |
| PML | 2837.91 ± 474.77 [1676.60-3999.60] | 2785.25 ± 471.41 [1777.60-4141.00] | 2760.03 ± 487.62 [1676.60-3999.60] | 2886.91 ± 417.15 [1737.20-3757.20] |
| PW | 1598.44 ± 268.21 [929.80-2222.00] | 1396.70 ± 322.27 [768.10-2424.00] | 1564.56 ± 264.36 [929.80-2222.00] | 1638.93 ± 288.63 [1091.50-2323.00] |
| PML/PW | 1.80 ± 0.32 [1.33-3.58] | 2.06 ± 0.40 [1.45-3.75] | 1.79 ± 0.36 [1.33-3.58] | 1.79 ± 0.26 [1.27-2.44] |
| PSH | 56.30% (1)<br>18.52% (2)<br>17.78% (3)<br>3.70% (4)<br>3.70% (5) | 37.98% (1)<br>22.48% (2)<br>1.55% (3)<br>31.48% (4)<br>5.43% (5) | 69.44% (1)<br>9.72% (2)<br>12.50% (3)<br>5.56% (4)<br>2.78% (5) | 41.27% (1)<br>28.57% (2)<br>23.81% (3)<br>1.59% (4)<br>4.76% (5) |
| NL | 94.81% (1)<br>5.19% (2) | 91.47% (1)<br>8.53% (2) | 97.22% (1)<br>2.78% (2) | 92.06% (1)<br>7.94% (2) |
| PA | 7.41% (1)<br>88.89% (2)<br>3.70% (3) | 18.60% (1)<br>79.07% (2)<br>2.33% (3) | 8.33% (1)<br>91.67% (2) | 6.35% (1)<br>85.71% (2)<br>7.94% (3) |
| PM | 100% (0) | 100% (0) | 100% (0) | 100% (0) |
| PMC | 45.19% (0)<br>54.81% (1) | 69.77% (0)<br>31.01% (1) | 34.72% (0)<br>65.28% (1) | 57.14% (0)<br>42.86% (1) |
| PMR | 0 ± 0 | 0 ± 0 | 0 ± 0 | 0 ± 0 |
| SEC | 2.22% (1)<br>88.89% (2)<br>8.89% (3) | 69.77% (2)<br>30.33% (3) | 95.83% (2)<br>4.17% (3) | 4.76% (1)<br>80.95% (2)<br>14.29% (3) |
| SEL | 15.79 ± 5.06 [8.50-31.00] | 20.09 ± 4.48 [11.50-32.00] | 16.48 ± 6.04 [8.50-31.00] | 14.08 ± 2.95 [8.50-20.00] |
| CPS | 26.67% (2)<br>46.67% (3)<br>26.67% (4) | 37.21% (2)<br>37.21% (3)<br>23.26% (4) | 20.83% (2)<br>58.33% (3)<br>20.83% (4) | 33.33% (2)<br>33.33% (3)<br>33.33% (4) |
| CSY | 100% (3) | 100% (3) | 100% (3) | 100% (3) |
| CSH | 100% (2) | 100% (2) | 100% (2) | 100% (2) |
| CL | 2613.42 ± 315.42 [2100.00-3259.00] | 2739.63 ± 373.50 [2050.00-3818.00] | 2630.29 ± 336.52 [2100.00-3259.00] | 2601.00 ± 319.61 [2102.00-3360.00] |
| CW | 1260.78 ± 206.60 [932.00-1766.00] | 1320.14 ± 220.72 [855.00-1883.00] | 1290.52 ± 250.44 [932.00-1766.00] | 1247.29 ± 141.56 [1048.00-1552.00] |
| CMW | 789.39 ± 119.71 [562.00-991.00] | 785.26 ± 96.99 [505.00-991.00] | 821.52 ± 127.54 [562.00-991.00] | 769.96 ± 95.41 [616.00-972.00] |
| CST | 100% (1) | 100% (1) | 100% (1) | 100% (1) |
| ANN | 100% (1) | 100% (1) | 100% (1) | 100% (1) |
| EXL | 603.63 ± 69.02 [451.40-781.82] | 594.46 ± 62.67 [395.00-741.52] | 627.67 ± 66.71 [451.40-781.82] | 561.66 ± 64.72 [411.10-717.34] |
| ENL | 493.13 ± 65.00 [362.70-612.56] | 469.36 ± 56.57 [282.10-588.38] | 512.69 ± 63.60 [362.70-612.56] | 461.16 ± 58.72 [346.60-588.38] |
| ENL/EXL | 0.82 ± 0.04 [0.73-0.92] | 0.79 ± 0.05 [0.67-0.92] | 0.82 ± 0.05 [0.73-0.92] | 0.82 ± 0.04 [0.74-0.90] |
| ENR | 100% (1) | 100% (1) | 100% (1) | 100% (1) |
| ENA | 100% (1) | 100% (1) | 100% (1) | 100% (1) |
| SPD | 13.70 ± 1.45 [11.00-18.00] | 15.35 ± 1.83 [12.00-20.00] | 14.07 ± 1.45 [11.00-18.00] | 13.63 ± 1.43 [11.00-17.00] |

**Table 8.** Statistical results of quantitative (QT) and qualitative (QL) characters of *Funaria* specimens from outside Sierra Nevada Mountains, included in the biometric study. Character abbreviations and status character are given in Table 1. For QT mean values ± SD [range] are given. For QL percentage of each status character (identified in brackets) is presented. Number (n) of fructified plants examined for each species is given. All measurements are given in μm, except for SEL in mm. The *F. flavicans* sporophytic measurements come from three incomplete plants, while gametophytic measurements come from a single plant. The measurements of *F. aequidens* phyllidia come from two plants and the other characters from three plants.

| Abbreviation of characters | *F. aequidens* (n = 3) | *F. arctica* (n = 3) | *F. flavicans* (n = 3) | *F. hygrometrica* (n = 3) | *F. polaris* (n = 6) |
|---|---|---|---|---|---|
| CAL | 2646.33 ± 682.76 [1861.00-3099.00] | 1231.67 ± 435.09 [730.00-1506.00] | 5514.00 | 3177.67 ± 286.41 [2879.00-3450.00] | 2406.67 ± 403.40 [1697.00-2778.00] |
| PP | 100% (2) | 100% (3) | 100% (2) | 100% (2) | 11.76% (2)<br>88.24% (3) |
| PML | 2094.07 ± 332.08 [1717.00-2626.00] | 987.80 ± 214.78 [644.80-1289.60] | 2437.47 ± 23.32 [2424.00-2464.40] | 3309.91 ± 443.55 [2666.40-3737.00] | 2107.93 ± 227.52 [1797.80-2646.20] |
| PW | 747.85 ± 155.54 [606.40-1051.10] | 798.83 ± 136.04 [564.20-983.30] | 1098.20 ± 99.68 [1030.80-1212.70] | 1221.40 ± 154.77 [1010.60-1374.50] | 1293.38 ± 184.94 [1010.60-1696.80] |
| PML/PW | 2.82 ± 0.25 [2.50-3.14] | 1.23 ± 0.15 [1.05-1.51] | 2.23 ± 0.17 [2.03-2.35] | 2.71 ± 0.17 [2.37-2.90] | 1.64 ± 0.16 [1.31-1.88] |
| PSH | 100% (4) | 88.89% (1)<br>11.11% (3) | 100% (1) | 42.86% (1)<br>57.14% (4) | 76.47% (1)<br>11.76% (2)<br>11.76% (3) |
| NL | 100% (1) | 100% (1) | 100% (2) | 100% (1) | 100% (1) |
| PA | 66.67% (1)<br>33.33% (2) | 100% (3) | 33.33% (1)<br>66.67% (3) | 100% (2) | 5.88% (1)<br>94.12% (2) |
| PM | 100% (0) | 100% (0) | 100% (0) | 100% (0) | 100% (0) |
| PMC | 66.67% (0)<br>33.33% (1) | 100% (0) | 100% (0) | 42.86% (0)<br>57.14% (1) | 52.94% (0)<br>47.06% (1) |
| PMR | 0 ± 0 | 0 ± 0 | 0 ± 0 | 0 ± 0 | 0 ± 0 |
| SEC | 100% (1) | 66.67% (1)<br>33.33% (2) | 100% (1) | 66.67% (1)<br>33.33% (2) | 16.67% (1)<br>83.33% (2) |
| SEL | 14.17 ± 4.75 [9.50-19.00] | 5.83 ± 1.04 [5.00-7.00] | 12.00 | 26.67 ± 4.01 [22.50-30.50] | 12.17 ± 1.83 [10.00-14.50] |
| CPS | 66.67% (1)<br>33.33% (4) | 100% (3) | 100% (4) | 100% (4) | 50.00% (1)<br>16.67% (2)<br>16.67% (3)<br>16.67% (4) |
| CSY | 100% (1) | 100% (3) | 100% (3) | 100% (3) | 100% (1) |
| CSH | 100% (2) | 100% (2) | 100% (2) | 100% (2) | 100% (2) |
| CL | 2276.67 ± 431.32 [1805.00-2651.00] | 1403.33 ± 131.25 [1277.00-1539.00] | 1850.00 ± 229.10 [1688.00-2012.00] | 3237.67 ± 216.09 [3065.00-3480.00] | 2279.33 ± 168.28 [2129.00-2562.00] |
| CW | 1220.00 ± 254.28 [927.00-1383.00] | 657.33 ± 25.77 [634.00-685.00] | 916.00 ± 141.42 [816.00-1016.00] | 1384.67 ± 123.51 [1262.00-1509.00] | 1068.67 ± 63.27 [991.00-1150.00] |
| CMW | 549.00 ± 80.91 [462.00-622.00] | 412.67 ± 21.73 [393.00-436.00] | 489.00 ± 48.08 [455.00-523.00] | 659.67 ± 24.03 [635.00-683.00] | 461.50 ± 36.90 [413.00-493.00] |
| CST | 100% (1) | 100% (1) | 100% (1) | 100% (1) | 100% (1) |
| ANN | 100% (1) | 100% (1) | 100% (1) | 100% (1) | 100% (1) |
| EXL | 388.00 ± 32.64 [340.00-434.00] | 312.78 ± 12.53 [300.00-340.00] | 430.00 ± 56.86 [360.00-500.00] | 442.46 ± 13.02 [427.20-467.50] | 326.00 ± 25.46 [302.00-360.00] |
| ENL | 288.50 ± 42.11 [244.00-346.00] | 239.22 ± 15.46 [216.00-266.00] | 138.29 ± 25.07 [110.00-176.00] | 325.13 ± 18.02 [298.30-354.70] | 179.75 ± 10.97 [170.00-195.00] |
| ENL/EXL | 0.74 ± 0.06 [0.67-0.81] | 0.76 ± 0.03 [0.72-0.81] | 0.32 ± 0.03 [0.29-0.36] | 0.73 ± 0.04 [0.69-0.79] | 0.55 ± 0.03 [0.52-0.58] |
| ENR | 100% (1) | 100% (1) | 100% (1) | 100% (1) | 100% (0) |
| ENA | 100% (2) | 100% (1) | 100% (3) | 100% (1) | 100% (2) |
| SPD | 19.53 ± 1.36 [18.00-22.00] | 21.93 ± 2.09 [18.00-26.00] | 21.47 ± 1.92 [18.00-25.00] | 20.60 ± 1.59 [17.00-24.00] | 19.67 ± 1.18 [18.00-22.00] |

The results did not show a great differentiation in the measurements or qualitative characteristics between the SN samples with different genotypes. Small differences were observed in the values of the characters PP, PMC, SEL and SPD, but not statistically significant. Keeled and flat margins phyllidia appeared in a smaller percentage in specimens with genotype B. The average seta length was greater in high mountain areas of SN than in the rest, although this average was even greater in specimens with genotype A. The average spore diameter was lower in SN lowlands. The quantitative characters CAL, PML, PW, CL, CW, CMW, EXL and ENL were homogeneus among the SN samples.

On the other hand, the results did show differences in the measurements or qualitative characteristics between the specimens studied of the different species of the genus *Funaria*. The plants belonging to the species *F. aequidens* and *F. polaris* were the only ones with asymmetric capsules. *Funaria arctica* had a much shorter seta length (and as a whole, the rest of the plant was much smaller, too) than the rest of the studied species. However, the average diameter of their spores was slightly higher than the rest. The ENL/EXL ratio was greater than 2/3 in all cases, except for studied samples of *F. flavicans* and *F. polaris*.

The values of the characters CAL, PML/PW and SPD were transformed into logarithms for the following analyses, since the original values presented a non-normal distribution.

The differences between the studied plants from the different groups are shown as boxplots in Figs. 24 and 25. The boxplots confirm high similarity within SN plants, but large differences with other species. Somewhat surprisingly, the *F. hygrometrica* plants from SN are different with respect to those from Chile, above all in the seta length and spore diameter quantitative characteristics.

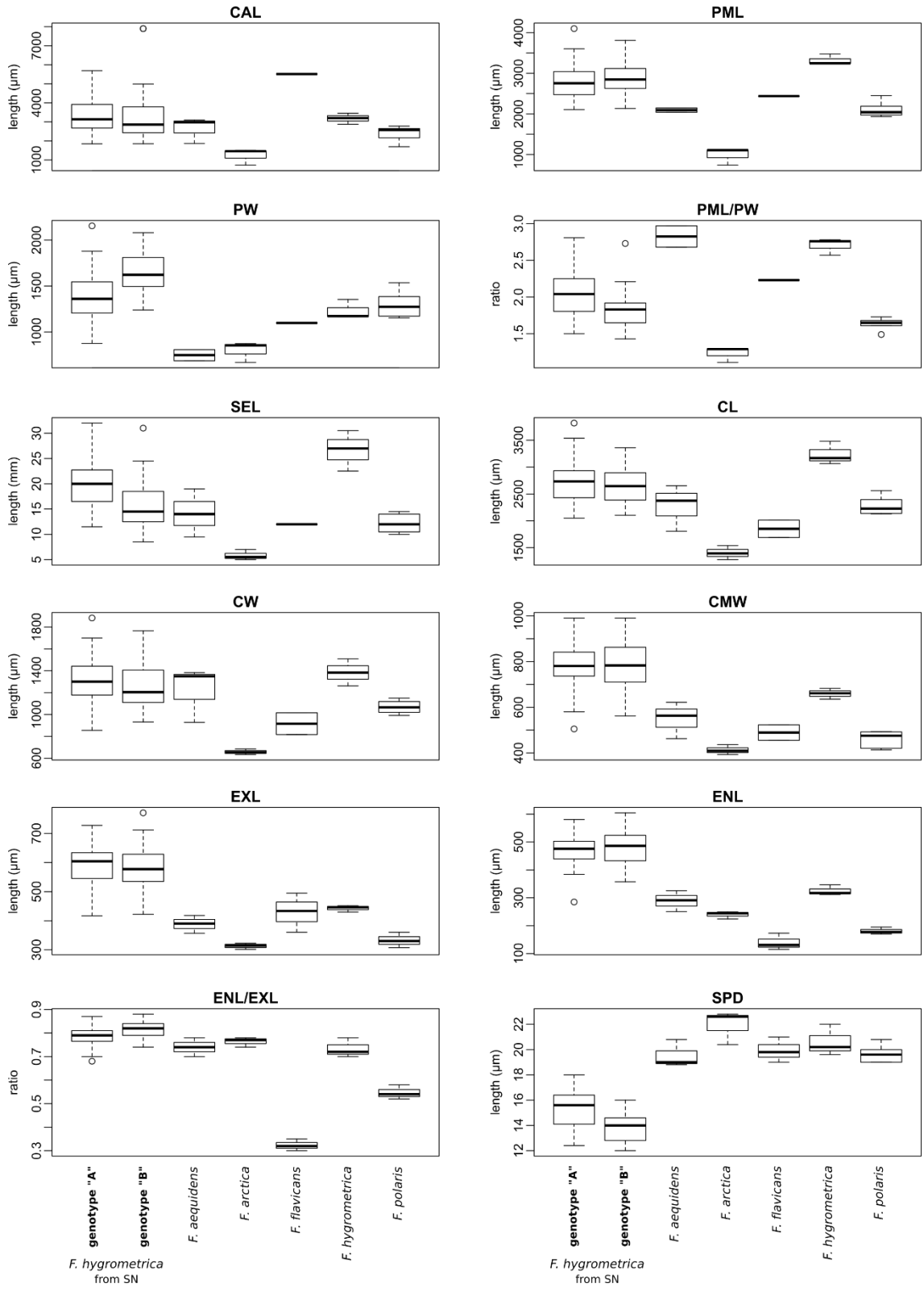



**Figure 24.** Boxplots of 12 morphological quantitative characters studied: caulidium length (CAL), phyllidium maximal length (PML), phyllidium width at widest part (PW), phyllidium maximal length/ phyllidium width ratio (PML/PW), seta length (SEL), capsule length (CL), capsule width (CW), capsule mouth width (CMW), exostome teeth length (EXL), endostome teeth length (ENL), endostome teeth length / exostome teeth length ratio (ENL/EXL) and spore diameter (SPD). The *F. hygrometrica* plants from SN are grouped by genotype.

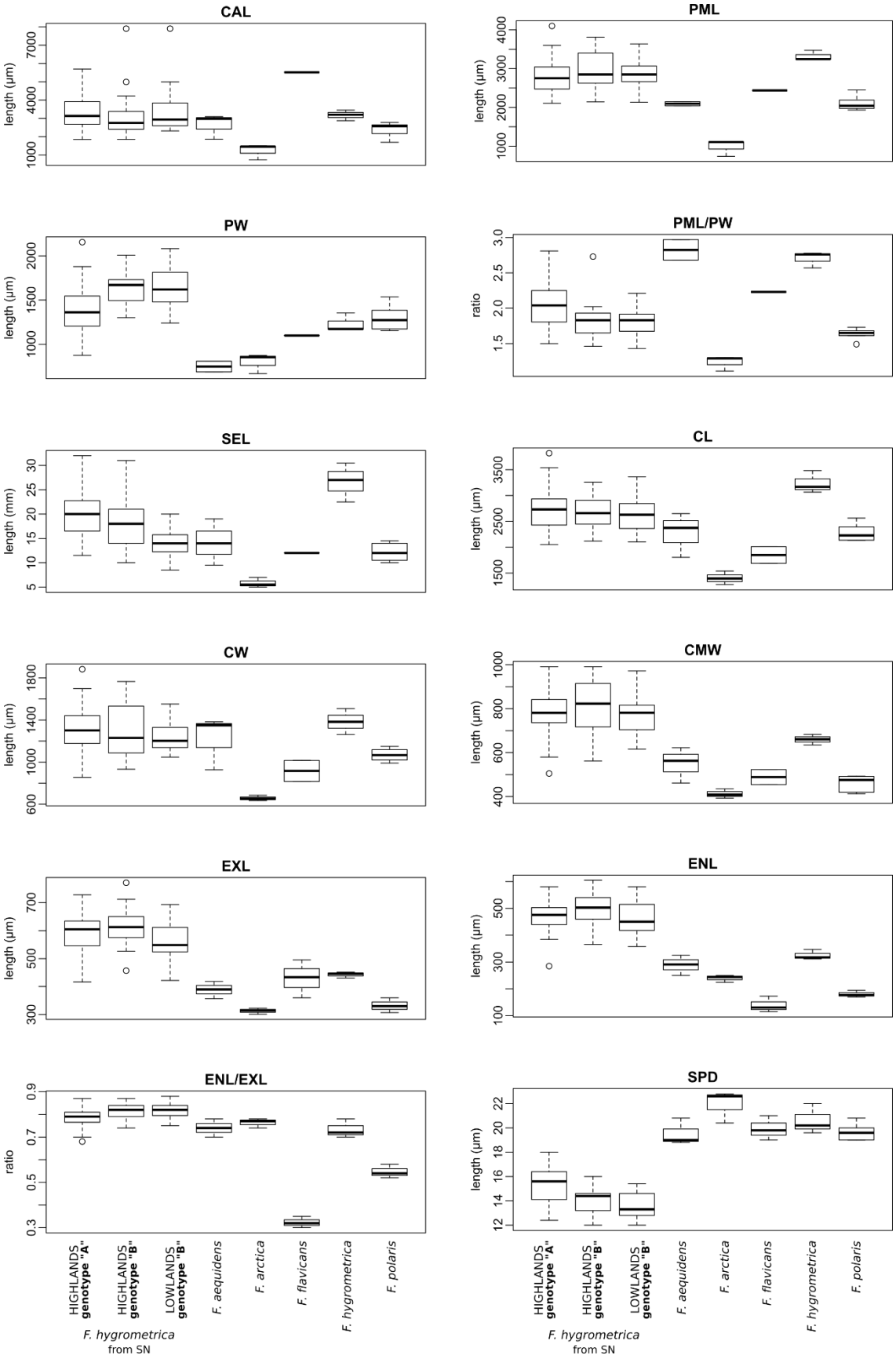



**Figure 25.** Boxplots of 12 morphological quantitative characters studied (abbreviations as in Fig. 24). The *F. hygrometrica* plants from SN are grouped by genotype and altitude.

Next, for the PCA, LDA and erroneous assignments analyses, only the plants that had data all the studied characters were used. The rest of the plants were eliminated by MorphoTools for these analyses. For this reason, only one of the three *F. flavicans* plants, two of the three *F. aequidens* plants and three of the six *F. polaris* plants were employed.

The PCA plots (Figs. 26 and 27) further support the high level of differentiation between the *Funaria* plants from SN in comparison with the plants of other species. The data points corresponding to the SN samples with different genotypes are slightly separated along the second principal component, but with a broad zone of overlapping. The plants of *F. hygrometrica* from Chile are close to the lower extreme of SN genotype B plants.

Similarly occurs on the LDA plots (Figs. 28 and 29). The studied SN plants are broadly overlapping, not being able to differentiate groups of plants. In these graphs, the *F. hygrometrica* plant from Chile is further away from the SN plants than in the PCA plots. The SN samples are clearly separated from the rest of the species of the genus *Funaria*.

In many samples from SN the prior classification based on ecological and genetic data and the posterior classification based on morphological data are contradictory (Figs. 30 and 31). There is no clear morphological differentiation between the two genotypes of the SN samples, based on morphological data. With the exception of a plant from the SN highlands, which is assigned to the *F. hygrometrica* group from Chile, the rest of the plants of each species are assigned to *a priori* defined group (species). In the case of *F. flavicans*, as there was only one complete plant, it was not assigned to any group. The erroneous assignments within *F. hygrometrica* from SN were the followings: 14 plants when grouped by genotype (Fig. 30) and 27 plants when grouped by genotype and altitude (Fig. 31).

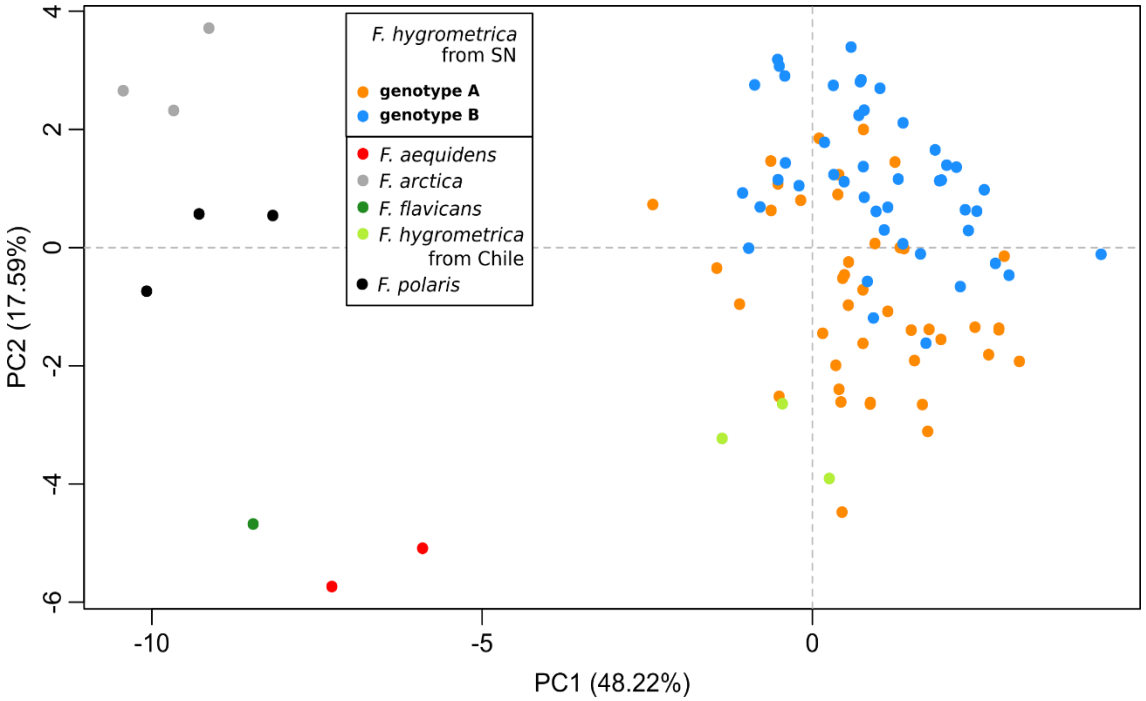



**Figure 26.** Principal Component Analysis (PCA) plots of *Funaria* based on morphometric data. The *F. hygrometrica* plants from SN are grouped by genotype.

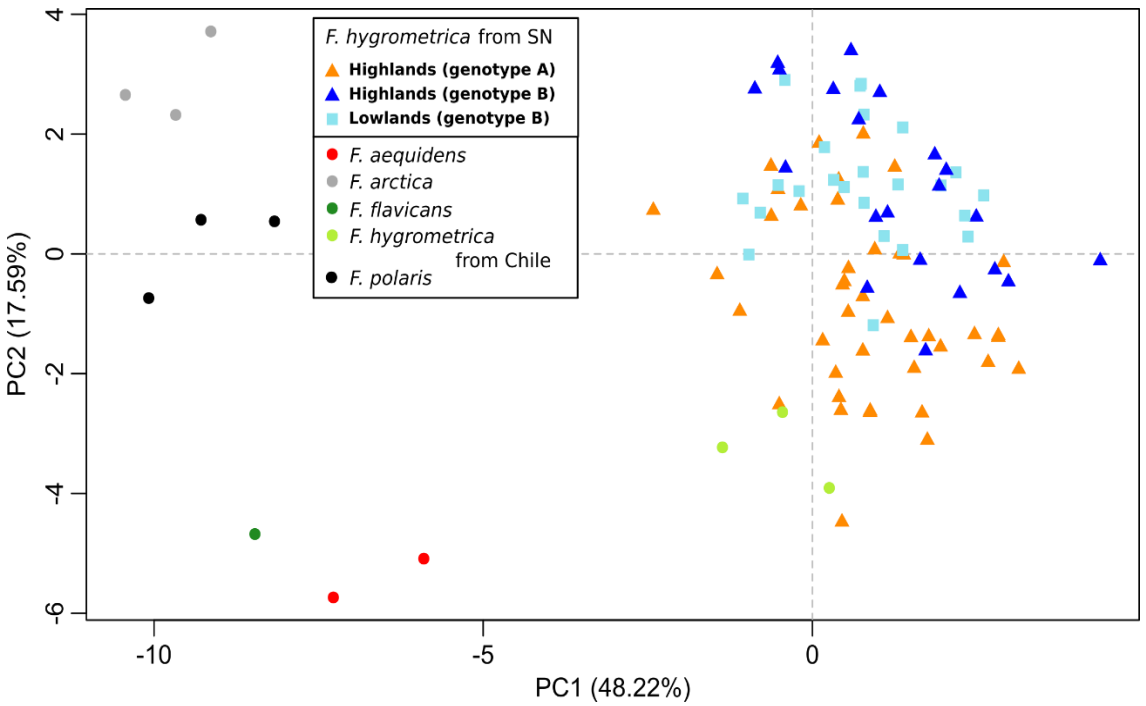



**Figure 27.** Principal Component Analysis (PCA) plots of *Funaria* based on morphometric data. The *F. hygrometrica* plants from SN are grouped by genotype and altitude.

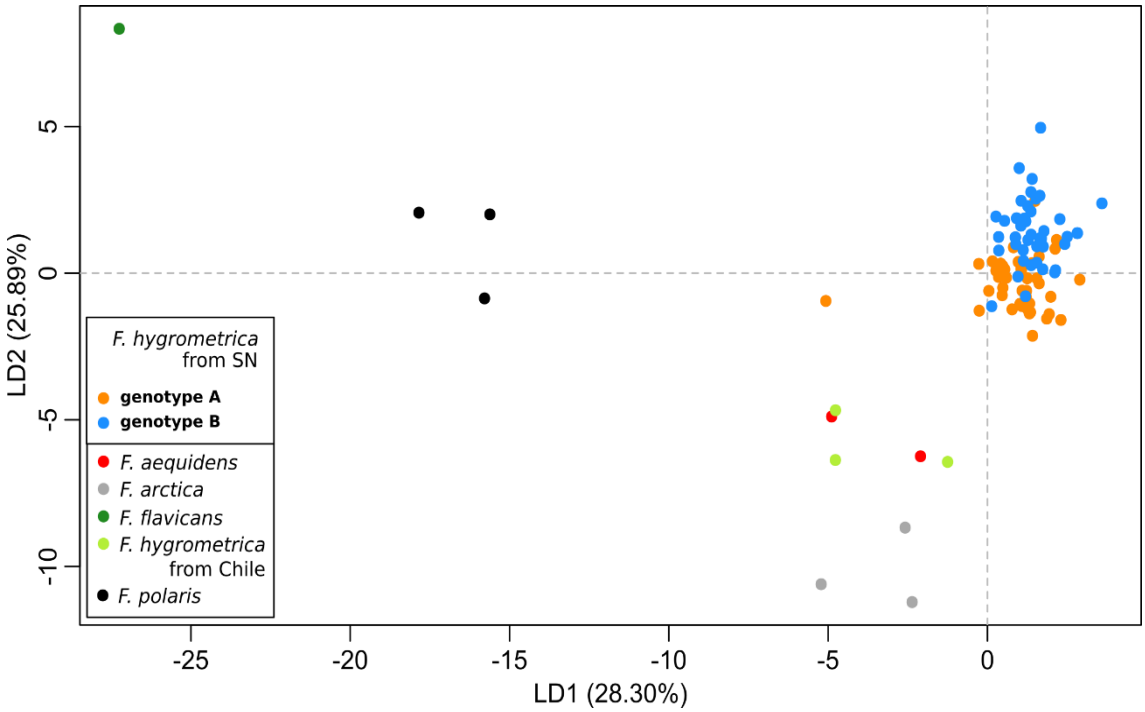



**Figure 28.** Linear Discriminant Analysis (LDA) plots of *Funaria* based on morphometric data. The *F. hygrometrica* plants from SN are grouped by genotype.

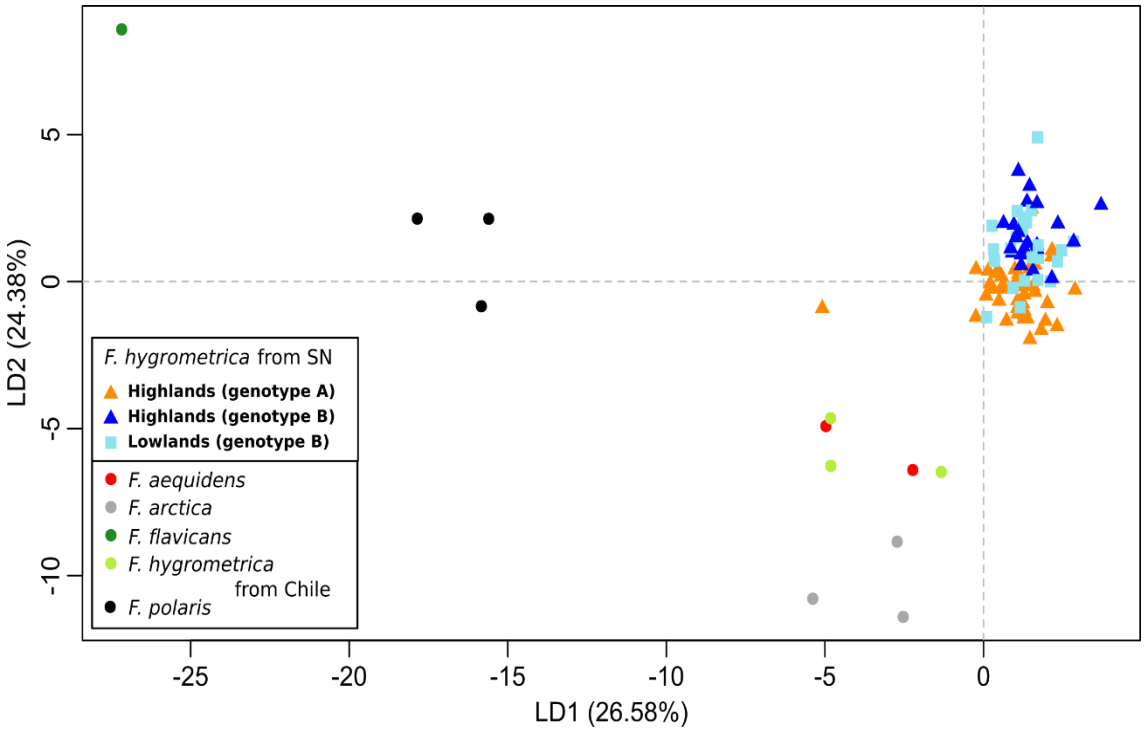



**Figure 29.** Linear Discriminant Analysis (LDA) plots of *Funaria* based on morphometric data. The *F. hygrometrica* plants from SN are grouped by genotype and altitude.

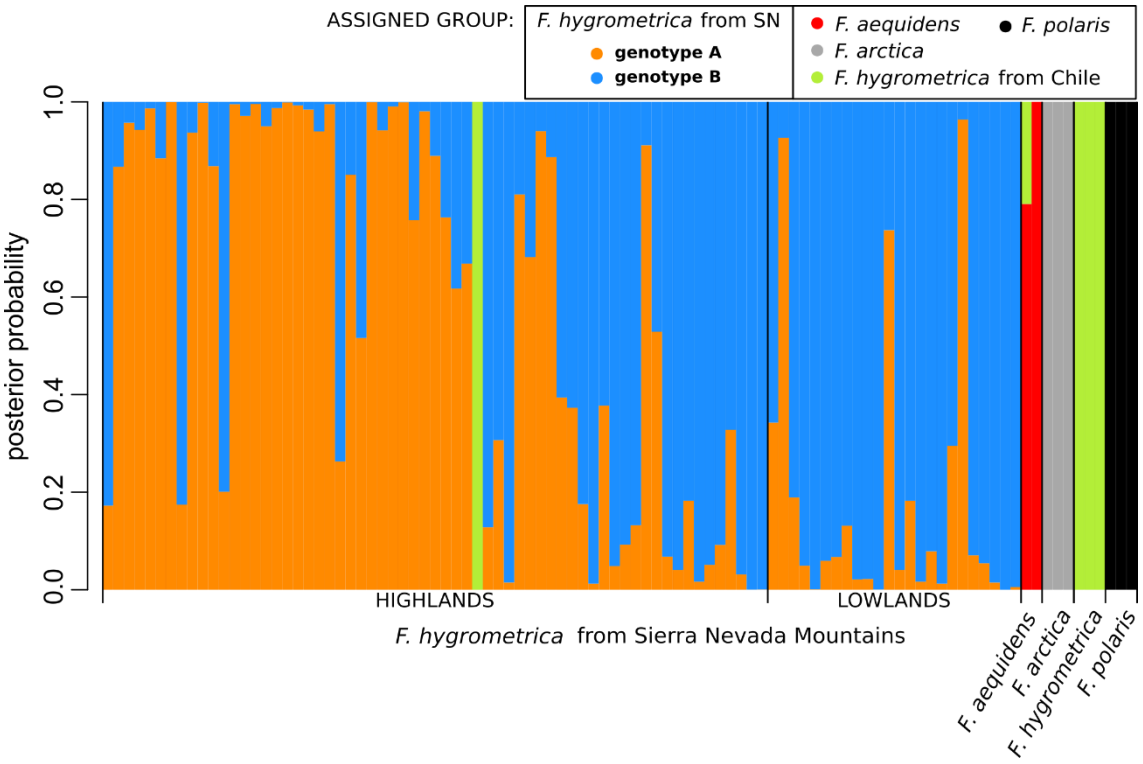



**Figure 30.** Barplot of posterior probabilities of assignment of the studied plants to predefined groups, according to the classificatory discriminant analysis. The *F. hygrometrica* plants from SN are grouped by genotype. *Funaria flavicans* was not assigned to any group as only one complete plant was studied.

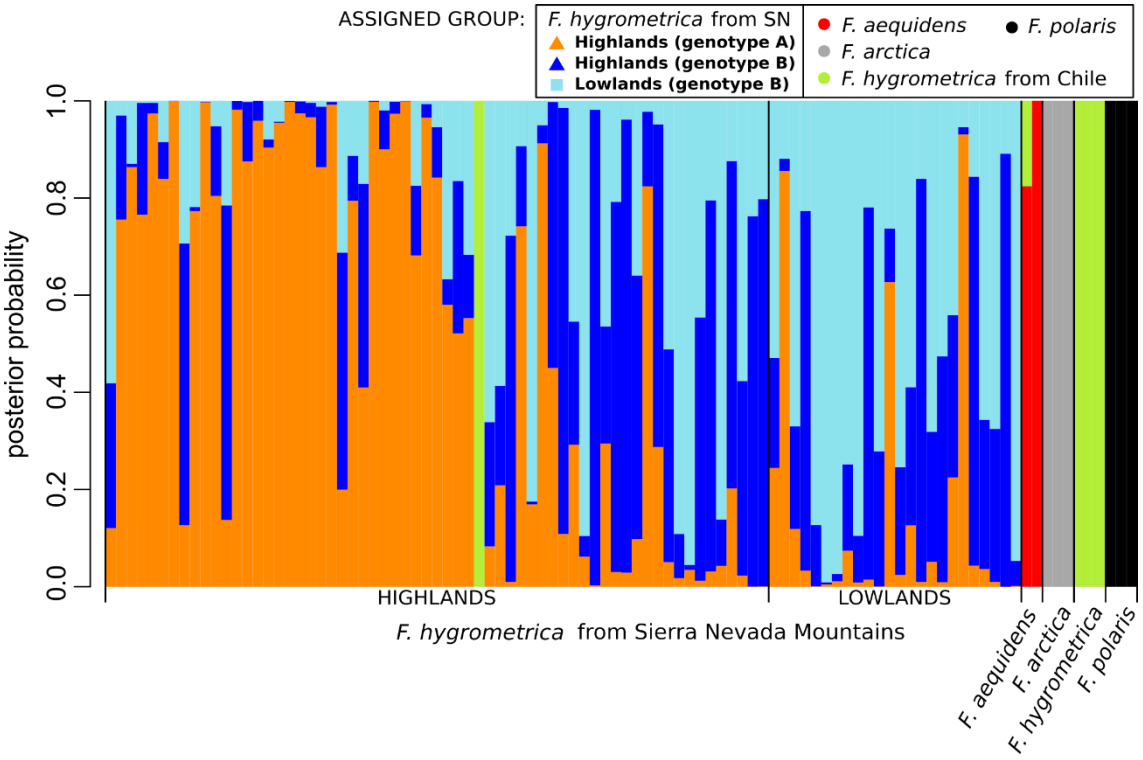



**Figure 31.** Barplot of posterior probabilities of assignment of the studied plants to predefined groups, according to the classificatory discriminant analysis. The *F. hygrometrica* plants from SN are grouped by genotype and altitude. *Funaria flavicans* was not assigned to any group as only one complete plant was studied.

## SIMULATION OF DISTRIBUTION PATTERNS

As explained in the Material and Methods section, the simulations were carried out to test if theoretically two species competing for the same resources can coexist under the assumptions that (i) the two species grow along an ecological gradient, each species having a selective advantage at one extreme of the gradient, and (ii) that there is a considerable number of migrants that reach neighboring locations. Several parameters for the fitness values were tried. If the fitness values were not too different the two species coexisted along the total of the gradient, but each species dominating the area where it had the higher fitness value (Fig. 32).

Additionally, to better visualize the results, a video in "mp4" format with the sequence of all generations can be downloaded at the following link: (https://github.com/olafumes/Vicariance/raw/main/simulacion_1_0_5_A.mp4).

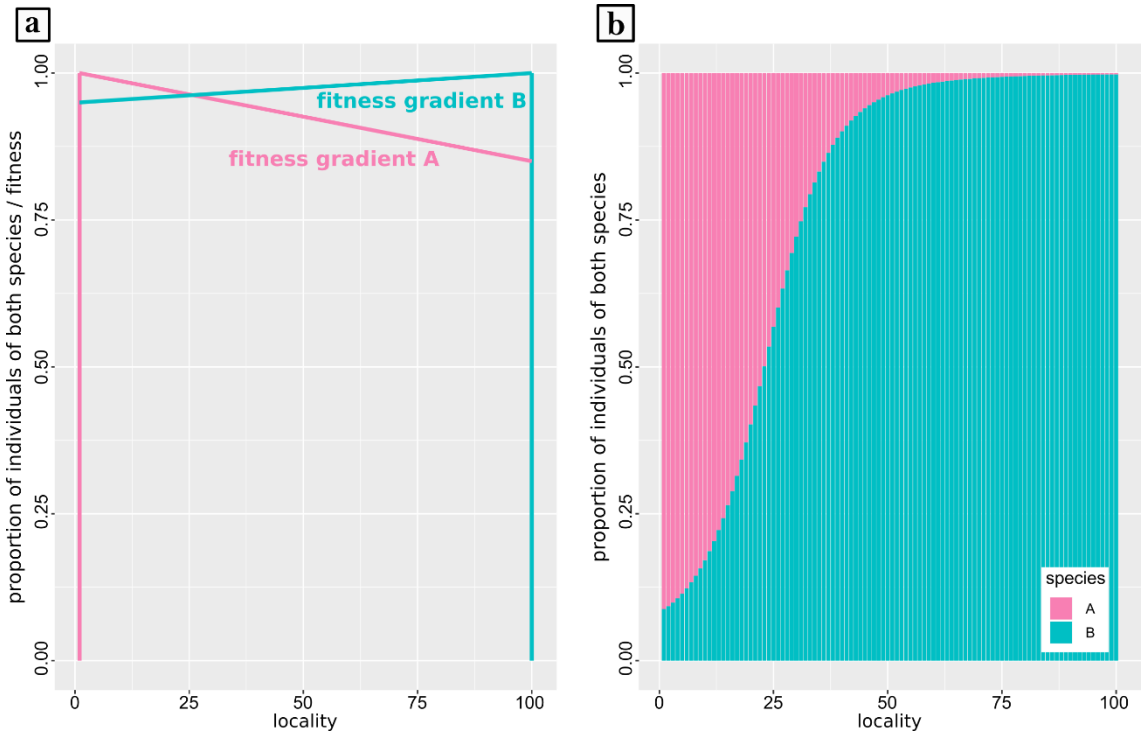



**Figure 32.** Results of the simulation study. **a**: shows the beginning of the simulation. **b**: shows the end of the simulation. The ordinate axis gives the relative proportion of individuals of both species, and in diagram a it also indicates the fitness value assigned to each species. Two species A and B can coexist along an environmental gradient even if they compete for the same resources if they have different selective advantages along this gradient.

# VI.

# DISCUSSION

## GENETIC VARIABILITY IN THE STUDIED AREAS

Cosmopolitan species are defined by being widely distributed throughout the world and it has been commonly thought that their populations are homogeneous not only morphologicaly but also lacking genetic spatial structure due to their presumed large population sizes, capability to easily disperse across intercontinental distances and high investment in reproduction (Tan & Pócs, 2000; Shaw, 2001; Vanderpoorten & Goffinet, 2009; Patiño et al., 2014). However, when their genetic diversity is studied, in some cases a diversity among populations greater than expected is observed and examples of single panmictic worldwide distributed lineages or species appear to be rather rare (Zaccara et al., 2020). That is the case of *F. hygrometrica,* for which the GBS results of this study confirmed the existence of two clearly differentiated genotypes (A and B) within a limited territory that accounts for a low portion of its total range across the studied Spanish areas, and, in a more interesting way, along a local altitudinal gradient in the Sierra Nevada Mountains. Intriguingly, there are similar cases observed in this mountain range, such as in the moss *Bryum argenteum*, which also showed a unique genotype restricted to areas above 2000 m a.s.l. (Pisa et al., 2013). Another similar case was identified in the moss *Ceratodon purpureus*, where an even new species was recently described (*C. amazonum*) above 1300 m a.s.l., which differs at both genetic and morphological levels from the widespread species *C. purpureus* (Nieto-Lugilde et al., 2018a, 2018b).

In a previous study focused on *F. hygrometrica* populations from Sierra Nevada Mountains (Magdy Abdallah Awad, 2013), a certain degree of hybridization between two genetic lineages was found in sequence data covering nuclear, chloroplast and mitochondrial markers. In our case study, this evidence of hybridization between the two mentioned genotypes A and B was not detected by any of the methods employed including genomic and transcriptomic sequencing technologies. Our results do not exclude the possibility of ocasional hybridization but suggest that the frequency is probably lower than initially thought. However, as highlighted in specific objective iii, the absence of hybrids can be explained by the fact that the two lineages are

genetically already too distinct to form a common gene pool, provided that a recent colonization event can be ruled out. Based on the demographic analyses performed in order to detect recent changes in population size, it seems that a scenario of recent colonization and expansion in Sierra Nevada Mountains is not supported by the GBS dataset (see Tajima's D tests in Results section). The latter, along with the lack of clear diagnostic morphological characters that allow to distinguish the two genotypes (see Morphology and possible cryptic species subsection below), raises the question of to what extent the two genotypes observed in *F. hygrometrica* populations represent a case of incipient speciation.

In a sexually reproducing plant species, nuclear genome *loci* are supposed to recombine freely with cytoplasmatic (mitochondria and plastids) *loci* while cytoplasmatic *loci* are often linked by uniparental inheritance (Olson & McCauly, 2000). Available evidence suggests that cytoplasmic inheritance in mosses is maternal and recombination is therefore practically absent (Natcheva & Cronberg, 2007). Here no indications of recombination between the nuclear ITS marker and the cytoplasmatic markers used in this study were found, although some of the specimens grouped in the two clades obtained in the phylogenetic analyses came from the same sampling location. This seems to indicate that the two inferred lineages can occur in the near vicinity, potentially in contact, but do not (or hardly so) cross. Perhaps this recombination between individuals of different genotypes, even though they occur very close (same location), is not taking place in favor of self-fertilization. It should be stressed that the consequences of self-fertilization depend strongly on the reproductive system of the organisms. In vascular plants, the reduction in heterozygosity of the offspring reaches 50% (Klekowski, 1979). According to Fisher (1949), a decrease in heterozygosity can have negative effects because recessive alleles with harmful effects are manifested. But if these harmful alleles do not exist (e.g. because they were eliminated by several cycles of self-fertilization), individuals capable of selfing would exhibit evolutionary advantages, since they can fertilise their own female gametes and other individuals (Fisher, 1949). But in bryophytes, due to their different biological cycle, the situation is different. The haploid gametophyte is the dominant generation, exposing deleterious alleles and eliminating them (Wyatt,

1982; Hedrick, 1987; Charlesworth, 1991). Furthermore, in seed plants, there are always two different gametophytes, the male pollen grain and the female gametophyte consisting of the ovule and a low number of acompanying cells. Therefore seed plants are always dioicous at the gametophyte level. As a consequence, even in monoicous species (at sporophyte level) as long as the sporophyte maintains some level of heterocygosity, the gametophytes produced after meiosis will be genetically different. In the case of bryophytes, the equivalent of self-fertilization in seed plants is the intergametophytic fertilization between two different sibling gametophytes originating from two spores produced by the same sporophyte (Eppley et al., 2007). In monoicous bryophytes however, such as *F. hygrometrica*, the same gametophyte produces both antheridia and archegonia and self-fertilization can occur intragametophytically (Hedrick, 1987), a situation not existing in seed plants. In these cases, a 100% homozygous sporophyte is produced, since both egg and sperm contain exactly the same genetic information (Hedrick, 1987; Eppley et al., 2007). As a consequence, all spores produced after meiosis are genetically identical, so that through an apparently sexual process the effective result is the production of a large number of clones (Haig, 2016).

However, there are few studies that have demonstrated a reduction of heterozygosity by selfing in *Funaria*. For instance, Shaw (1991) studied two populations of *F. hygrometrica* and found no genetic variability within each using isoenzymes as markers. Taylor et al. (2007) did not observe any negative effect due to selfing in the same species. If this species self-fertilises in a generalised way in Sierra Nevada Mountains, a working hypothesis would be that these two evolutionary distant lineages A and B are composed of individuals with clonal multiplication, possibly undergoing incipient speciation favoured by reproductive isolation and, a likely option, mostly adapted to contrasting local environmental regimes. But at this point it not clear if the low rate of crossing between the two lineages is a consequence of prezygotic (i.e. sperm cells do not reach neighboring individuals) o postzygotic (i.e. eggs fertilized by sperm of other lineage do not develop) barriers. The results of von Wettstein (1923), who produced artificial crosses between different species of the Funariaceae belonging to different genera, among them *F. hygrometrica*, suggest that

hybrids may occur in nature, although possibly with a reduced viability and a high percentage of sterile spores.

The appreciation that geographic features can impose an important barrier to the dispersal and posterior establishment of plants and animals dates back to historical biogeographers such as Augustin-Pyramus de Candolle and Alfred Russell Wallace, who already pointed to the thesis that distribution disjunctions can often be explained by abiotc phenomena (Juárez-Barrera et al., 2018). Furthermore, the new sequencing technologies to obtain nuclear DNA information, as the GBS dataset achieved in the present study, is increasingly used to detect patterns of effective dispersal in plants across the landscape, due to the massive numbers of independently segregating genomic regions and much greater statistical power. Our plant landscape genetics approach supports the hypothesis that there were effects of landscape factors on patterns of spatial genetic structure, and geographic distance was a poor predictor of genetic relatedness patterns among populations, at least across the Sierra Nevada Mountains.

The only studied area for which we found significant effects of IBR on gene flow is thus the Sierra Nevada Mountains. In this mountainous system, isothermality and temperature seem to determine patterns of genetic isolation. It is thus tempting to propose that the higher exposure to contrasted temperatures mirrored by the importance of isothermality might render lowland habitats less suitable for dispersal and establishment of the genotype A. Effects of isolation by ecological factors on gene flow has been related to local adaptation in plants (Freeland et al., 2010; Gould et al., 2015) and other organisms (Peterman et al., 2014; García-Olivares et al., 2019; Salces-Castellano et al., 2020). However, because we were not able to identify GBS *loci* under selection (results not shown), the potential role of a natural selection on gene flow patterns is completely dependent on the transcriptomic dataset (see Transcriptome subsection below). Alternatively, it is also possible that the failure to detect any *loci* under selection in the GBS dataset is because we did not recover any SNPs within or closely linked to potential genes under selection. A more efficient procedure being able to recover a larger number of SNPs than in the present study could shed light on

the molecular mechanisms that are behind the spatial genetic structuring observed in *F. hygrometrica*.

Based on the ITS sequences of the two clades of the *F. hygrom*etrica specimens from Sierra Nevada Mountains, the Chilean sample of *F. hygrometrica* is very closely related to the highland phylogenetic clade of Sierra Nevada (corresponding to genotype A), being a presumed indication of the existence of transcontinental dispersal in this species. This also seems to have occurred in the other study areas in the Iberian Peninsula, where 1-4 scattered samples in each area belong to genotype A. Perhaps, as a hypothesis, spores of genotype A plants could travel those long distances and germinate in these places, and did not interbreed with other genotypes due to an extremely high rate of self-fertilization or genetic barriers. As a consequence, a certain number of less adapted individuals might maintain populations under suboptimal conditions where the population size should be conditioned by dispersal capacity allowing establishment of individuals arriving from neighbouring regions and local reproduction (although lower than that of better adapted individuals).

According to transcriptomic data and ITS sequences, there is little genetic and gene expression distance between genotype B and the other studied species of *Funaria* (*F. aequidens*, *F. arctica, F. flavicans*, *F. polaris*), that can easily be distinguished by morphological characters. Genotype A is further away from them, supporting its consideration as a clearly different genetic lineage.

## GENOME SIZE AND PLOIDY

The flow cytometry data showed an unclear correlation with the obtained genotypes, perhaps due to the small number of samples studied. However, some interesting information was obtained. Of the four cytotypes observed, cytotype a seems to represent haploid individuals, while cytotypes b and c appear to represent two types of diploidy, with approximately twice the genome size of individuals with cytotype a.

Cytotype d, formed by only two individuals, with slightly different genome sizes, is between 133% and 142% larger than the mean genome size of cytotype a.

According to our results, haploidy was found in both genotypes, being mostly exclusive to genotype B. Diploid individuals were also found in both genotypes, but cytotype b appeared exclusively in genotype B, while cytotype c was found exclusively in genotype A. It is known that in several bryophyte species one to several small accessory chromosomes are occasionally found (Wigh, 1973). By definition these chromosomes are supernumerary and their number can vary within different individuals of one species (Wigh, 1973). These chromosomes could possibly account for the minor differences at the diploid level. It is also known that in the moss *Ceratodon purpureus* rearangements occur at the chromosome level (McDaniel et al. 2013). Likewise, Nieto-Lugilde et al. (2018a) found clear differences in DNA content in two closely related species of the genus *Ceratodon*, underlining the fact that although bryophytes show very little variation in genome size across species in comparison with other groups of land plants (Temsch et al. 1998; Voglmayr, 2000; Bainard & Villarreal, 2013; Bainard et al., 2013), there is a condiderable dynamics in genome size within species or between closely related species.

In other species of bryophytes, such as the moss *Tortula muralis* Hedw. (Košnar et al., 2012), no clear relationship between genetic sequences and the level of ploidy was found. In angiosperms, an evolutionarily association between polyploidy and clonality was observed (Castro & Loureiro, 2014; Vallejo-Marin, 2014), although the fact that clonality promotes polyploidy is not clear (van Drunen & Husband, 2019). Little can be said about cytotype d, because there are only two samples studied, and each one belongs to a different genotype.

Due to the lack of clear signatures of hybrids, we would favour that our findings in terms of cytotypes might be the result of autopolyploidy mechanisms. Currently, it is broadly recognized that large-scale or whole genome duplication promoted by autopolyploidy or allopolyploidy can drive speciation in plants (Soltis et al., 2016), with implications for morphological or physiological differentiation (Rensing, 2014; Nieto-Lugilde et al., 2018a, 2018b).

## TRANSCRIPTOME

The studied samples of the species *F. aequidens*, *F. arctica* and *F. polaris* appear grouped, showing few differences in gene expression between them and clustered in the corresponding PCA. This is surprising, since a greater difference in the inter-specific gene expression with respect to the intra-specific one was expected. Furthermore, these seemed to be a little closer to *F. hygrometrica* samples with genotype B, as occurs with the Sanger sequencing methodology.

The results presented in the volcano plots and the number counts of differentially expressed genes show that a high percentage of all genes were expressed differentially between the A and B genotypes and that the differences are even higher than the expression level differences caused by extreme thermal stress.

The gene onthology analysis indicates that a major part of the differences observed between A and B genotypes under control conditions corresponds to the translational system (GO categories: ribosome, structural component of ribosome, translation). Especially interesting is the fact that the GO category Unfolded Protein Binding is differentially expressed under thermal stress conditions (AT *vs* AC, BT *vs* BC) but that this differential expression is even significant in comparison BT *vs* AT, that is B genotype under thermal stress *vs* A genotype under thermal stress. The proteins described under this term are mainly heat shock proteins and chaperones. Protein folding is a complex process that in the case of most proteins cannot proceed spontaneously to result in the correct final protein conformation (Vabulas et al., 2010). Molecular chaperones aid the newly formed protein to reach their final conformation (Hartl, 1996). Under stress and especially heat stress chaperones act by preventing protein aggregation (Vabulas et al., 2010). Other functions under stress conditions are their involvement in the lysis of terminally damaged proteins (Chiang et al., 1989; Dice, 1990; Arndt et al., 2007; Kon & Cuervo, 2010).

Together the transcriptome data points to the direction that there might be differences in the heat stress tolerance of A and B genotypes. These differences might

be very subtle and difficult to detect in physiological experiments or even less so in transplantation experiments. But with the help of some statisticals (e.g. the pwr package of R) it is possible to calculate the number of samples necessary to detect statistically significant differences between populations. For example, if we have a small effect size of $d = 0.2$ and wish to use a significance level alpha = 0.05 with a statistical power of 80%, 393 individuals of each group are needed. The effect size cannot easily be transformed in selective pressure but a selective disadvantage of as low as 1% less survival per generation will lead to near extinction in 1,000 generations, given that $0.99^{1000} = 4.3 \times 10^{-5}$. That means that starting from an equal proportion of individuals of both species or lineages, after 1,000 generations (a very short time in evolutionary terms) the proportion will be less than one individual of the species with disadvantage for 2,000 individuals of the species with a slight selective advantage.

The data obtained for this study will be used in the near future for additional analysis. The transcriptome sequences will be used to reconstruct phylogenetic trees with a high number of markers. Furthermore the sequences will be studied to detect signals of selection.

## MORPHOLOGY AND POSSIBLE CRYPTIC SPECIES

In the biometric study carried out with the Sierra Nevada Mountains samples and a few specimens of *F. aequidens*, *F. arctica*, *F. flavicans* and *F. polaris*, and even a Chilean sample of *F. hygrometrica*, it was verified that the diagnostic characters of *Funaria* genus (Fife, 1985; Casas et al., 2006; McIntosh, 2007; Brugués & Ruiz, 2010; Ignatova & Fedosov, 2017) did not show significant variation. All the studied plants showed entire phyllidium margins at upper part, with no rows of narrower cells than the other phyllidium laminal cells, pyriform, striate when mature and dry capsules, and revoluble annulus. Also many of the morphological characteristics described for *F. hygrometrica* were homogeneous among the Sierra Nevada Mountains samples, such as caulidium length, phyllidum size and shape, capsule size, capsule mouth width, exostome teeth length, and the most important, endostome teeth length, as it is considered the differential character from *F. microstoma* (McIntosh, 2007; Ignatova

& Fedosov, 2017) and also from *F. flavicans* (McIntosh, 2007). Only a few characters showed some differences within Sierra Nevada samples but with no significant support. Phyllidia of plants with genotype B were less frequently keeled and with plane margins than in genotype A. The seta length was very variable, but the specimens with genotype A were observed to present the longest, followed by those of highlands. And the spores were smaller in plants of lowlands. On the contrary, the Chilean plants of *F. hygrometrica* showed significant differences with the Sierra Nevada Mountains samples in sporophytic characters, such as seta length and spore size.

Neither PCA nor LDA analyses support a morphological differentiation between the two genotypes. As expected, the other species of the genus included in the biometric study could be easily distinguished morphologically, although according to the Sanger sequence data they rather resemble the genotype B samples from the Sierra Nevada lowlands and highlands. The lack of morphological variation between genotypes A and B from Sierra Nevada samples point to that genotypes A and B could be cryptic species within the morphospecies *F. hygrometrica*, that seem to have a faster evolution or divergence at the molecular level than at the morphological one. This view is supported by the findings of Rahmatpour et al. (2021), who came to the conclusion that the family Funariaceae is generally characterized by a high genetic variability, which contrasts with a low morphological variation.

The geographical distribution of genotype A is more reduced than in the generalist genotype B, being restricted to high altitudes in Sierra Nevada Mountains and some isolated samples in Asturias, Leon and Murcia provinces. Perhaps other high mountain areas in the Iberian Peninsula should be studied, to verify if this new detected genotype has a broader geographical distribution.

Cryptic species have been very frequently described in bryophytes (see Renner 2020 for a review). There are already numerous molecular studies that indicate the existence of cryptic biological species that show genetic divergences but that are morphologically similar to other phylogenetically close ones and consequently cannot be identified using the traditional concept of morphological species (e. g. Geiser et al., 1998; Shaw, 2001; Hebert et al., 2004; Bickford et al., 2007; Heinrichs et al., 2011).

A remarkable example of cryptic bryophyte species is related to the circumpolar wide-boreal, almost cosmopolitan liverwort *Aneura pinguis* (L.) Dumort. for which Bączkiewicz et al. (2017) identified at least ten sibling species by means of DNA barcoding, but they also differed in geographic distribution and habitat preferences. Inside the moss *Racomitrium lanuginosum* (Hedw.) Brid., three possible cryptic species were found, which presented genetic and geographical divergences, but not morphological ones (Hedenäs, 2020). In the case of *Conocephalum salebrosum* Szweyk., Buczk. & Odrzyk. it was firstly considered a cryptic species complex based on molecular differences, but after some decades new studies allowed its distinction at morphological level from *C. conicum* (L.) Dumort (Szweykowski & Krzakowa, 1979; Szweykowski et al., 2005).

According to Struck et al. (2018) and Renner (2020), when defining a genetic lineage as a cryptic species it is necessary to carry out a very careful interdisciplinary study, combining molecular data (genetic variability, ploidy, isozymes, transcriptome, etc.) with morphological data. Vigalondo et al. (2019) advocate for using operational species delimitation criteria (e.g. morphological differentiation, monophyly, intrinsic reproductive isolation, geographic and ecological information). In this study we also used an integrative taxonomy approach to explore biodiversity in *F. hygrometrica*, obtaining genome size data, genetic data from DNA Sanger sequencing, genomic and transcriptomic sequencing, and also morphological data. We even analyze the distribution data. Nevertheless, results of all these analyses did not point in the same direction, indicating data of genetic technologies and distribution towards the existence of two clearly differentiated genetic lineages, but not the ploidy analysis neither the morphometric data.

Species concept and delimitation are recurrent topics in biology. There are different ways to define a species depending on the features that are taken into account. Some of the most employed concepts, although there are many more, are the followings (De Queiroz, 2007; Aldhebiani, 2018). A "biological" species is a group of individuals that can interbreed with each other, but are reproductively isolated from other groups. A “morphospecies” is distinguished exclusively by morphological characters, ignoring ethological, ecological, geographical and even genetic

characteristics. An "ecological" species is a set of individuals that live in a specific ecological niche. However, currently, an attempt is made to bring together all these characteristics when defining a new species, mainly due to recent technological advances in molecular analyses (Turelli et al., 2001; Seehausen et al., 2014; Jiggins, 2019). The vast majority of species have been identified by morphological criteria, probably underestimating the number of species due to ignoring other criteria (Wyatt, 1985; Shaw, 2001; Duminil & Di Michele, 2009; Harper et al., 2009).

According to Kadereit et al. (2012), almost none of the cryptic species mentioned in the literature are formally described and named by their discoverers, nevertheless these authors argue in favour of naming cryptic taxa despite their similar morphology, in order to avoid merging genetical divergent lineages. Because of the traditional morphology-based taxonomy, the identification of these taxa would not be possible without specialized laboratory facilities and skills to carry out molecular analyses. For instance the use of a molecular barcode would allow to reliably identify this cryptic species. Already previous authors (e.g. Hedenäs et al., 2014; Lang et al., 2014; Patiño et al., 2017) proposed to use DNA barcodes to identify closely related species, even in those that although not cryptic, distinction based only in morphological features is very difficult.

## SIMULATION OF DISTRIBUTION PATTERNS

The competitive exclusion principle states that "complete competitors cannot coexist" (Hardin, 1960). This principle was based on observations of unicellular organisms, specifically yeasts and mathematically substantiated by Gause (1934). In our simulation, two species were complete competitors in the sense that the maximum number of individuals at one locality did not change, independently if it was occupied by one species, the other or both. The results of the simulation study, although admittedly very simplified, show that two independent species can share the same habitat even if they compete strongly for the same resources and one of the species has a certain selective advantage at this habitat. The condition is that a sufficient number of migrants reaches this habitat from other localities where the selective

advantage is reversed. In the case of bryophytes, intercontinental dispersal has been observed (e. g. van Zanten, 1978; van Zanten & Pocs, 1981; Muñoz et al., 2004; Shaw et al., 2003, Biersma et al., 2017). In our work the ITS sequence of a sample from Chile was very close to samples from Sierra Nevada Mountains, indicating a high dispersal capacity in this species. As a consequence, the conditions for the coexistence of two different species are possibly given in the case of the two *Funaria* lineages. This is especially the case in the Sierra Nevada Mountains, where within few kilometers the environmental conditions change drastically. Our model also explains the possibility that two lineages in a given locality occur with the same abundance although one of them is better adapted to the local conditions. According to the simulation, this can be in the case that more migrants of the less adapted lineage reach this locality than from the better adapted lineage.

## FUTURE PERSPECTIVES

The results of this study suggest that within the morphospecies *F. hygrometrica* at least two genetically distinct lineages exist, which are to some degree reproductively isolated. Although no clear morphological separation was found between the two lineages using the characters applied in *Funaria* systematics, it cannot be excluded that subtle differences exist. In future studies computer based image processing programs like MASS (Chuanromanee et al., 2019), which are designed to analyse size in shape of objects in digital images might reveal those differences. Another line of possible investigations is exploiting additional characters like those of the protonema, although in this case they are of less value to actually determine field collected samples.

The inclusion of more genetic markers in future studies is also desirable. Our GBS did reveal a lower number of SNPs than initially expected. This may in part be due to the selection of the restriction enzyme combination used. After our experiments were carried out, a draft version of the genome of *F. hygromerica* became available (Kirbis et al., 2020) and with the help of programs like SimRAD (Lepais &

Weir, 2014) it is now possible to calculate the number of expected fragments of restriction cuts in silico and aid the choice of appropriate enzyme combinations.

A possible alternative to GBS is target enrichment, where genomic regions of interest are captured before NGS sequencing. Breinholt et al. (2021) published a target enrichment probe set that allows to analyse 451 exons corresponding to 248 genes found in land plants simultaneously.

To understand the evolutionary history of *Funaria*, it is important to know the grade up to which self-fertilization occurs in natural populations. This can be done by studying the sequence variability among gametophytes with origin from the same sporophyte. As explaind above, in the case of self-fertilization all gametohytes will be genetically identical. Techniques like GBS, with an appropriate combination of restriction enzymes, should be adequate to solve this problem.

The availability of the complete genome of *F. hygrometrica* allows to study selection directly at the sequence level. Basically it is necessary to sequence a number of samples of both lineages and identify mutants that are fixed in one group and absent in the other. Mutations that follow this pattern are more frequent in groups of genes that are under selection. To test the significance of fixed mutations the relation of synomymous to non-synomymous mutations is taken into account by the test of McDonald & Kreitman (1991).

Epigenetic changes may contribute to ecological speciation (Smith & Ritchie, 2013). Recent advances in technology allow now to sequence a high number of genetic markers in a variation of the GBS protocol (Trucchi et al., 2016; van Gurp et al., 2016; Werner et al., 2020). *Funaria hygrometrica* might be an especially interesting subject of research in this field as the dominat haploid phase simplifies genetic analyses.

*Funaria hygrometrica* is a cosmopolitan found in all continents including the Antarctic (Schuster, 1983). To understand the evolutionary history of this morphospecies and to gain a deeper insight in the biogeography of different genetic lineages a worldwide sampling scheme is necessary. Such a project might also reveal the existence of more genetic lineages with different habitat preferences.

# VII.

# CONCLUSIONS

The main conclusions arising from the results of the present Doctoral Thesis are as follows:

1. The presence of two clearly differentiated genetic lineages in the cosmopolitan moss species *F. hygrometrica* is confirmed across the four studied Spanish areas and, even, along a local altitudinal gradient in the Sierra Nevada Mountains.

2. Hybridization between both lineages could not be confirmed in the Sierra Nevada Mountains, even in co-occuring localities, suggesting that they are to some degree reproductively isolated. A scenario of recent colonization and expansion in Sierra Nevada Mountains is not supported by a demographic analysis. Alternatively, the absence of recombination between individuals of different genotypes might be explained by a phenomenon of self-fertilization as dominant mode of sexual reproduction in both genomic clusters.

3. Landscape resistance does not contribute to genetic differentiation in the two genetic clusters in three out of the four areas studied. Only in Sierra Nevada Mountains, findings suggest that an isolation-by-environment (namely isothermality) cannot be ruled out, which points to a potential barrier that might restrict dispersal or effective establishment, ultimately promoting ecological isolation.

4. Transcriptomic data points to the direction that there might be differences in the heat stress tolerance between the two genetic linages of *F. hygrometrica*, as differentially expressed gene under thermal stress conditions belong to the gene ontology category Unfolded protein binding, which consists mainly of heat shock proteins.

5. Haploid and diploid individuals are present in both genetic lineages but genome size of diploids varied slightly. A possible mechanism to explain these differences are variable numbers of accessory chromosomes, known to occur within other moss species.

6. The lack of significant morphological differentiation among the Sierra Nevada individual groupings classified according to their genetic identity, elevation and ploidy level, points to the hypothesis that the two genotypes correspond to cryptic species within the morphospecies *F. hygrometrica*, that seem to have a faster evolution or divergence at the molecular level than at the morphological one.

7. The simulation carried out showed that the two cryptic species could share the same habitat, even where one of them has a slighltly lower fitness, as long as sufficient number of migrants arrive in this habitat from other localities, where the selective advantage is inverted.

8. The present work underscores the importance of considering multiple sources of empirical evidence when inferring evolutionary trajectories and developing taxonomical criteria to better account for species concepts in broadly distributed bryophyte species.

# VIII.

# REFERENCES

Abbott, R. J. (2017). Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *Journal of Systematics and Evolution*, *55*(4), 238-258. https://doi.org/10.1111/jse.12267

Aedo, C., Medina, L., & Fernández-Albert, M. (2013). Species richness and endemicity in the Spanish vascular flora. *Nordic Journal of Botany*, *31*(4), 478-488. https://doi.org/10.1111/j.1756-1051.2012.00009.x

AEMET (Eds.) (2011). *Atlas climático ibérico/Iberian climate atlas*. Retrieved from https://www.aemet.es/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/detalles/Atlas-climatologico

Ah-Peng, C., Chuah-Petiot, M., Descamps-Julien, B., Bardat, J., Stamenoff, P., & Strasberg, D. (2007). Bryophyte diversity and distribution along an altitudinal gradient on a lava flow in La Réunion. *Diversity and Distributions*, *13*(5), 654-662. https://doi.org/10.1111/j.1472-4642.2007.00393.x

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Molecular Biology of the Cell* (4th ed.). New York, NY: Garland Science.

Aldhebiani, A. Y. (2018). Species concept and speciation. *Saudi Journal of Biological Sciences*, *25*(3), 437-440. https://doi.org/10.1016/j.sjbs.2017.04.013

Alexa, A., & Rahnenfuhrer, J. (2020). topGO: Enrichment analysis for Gene Ontology. R package version 2.40. 0 [Computer software]. Retrieved from https://bioconductor.org/packages/release/bioc/html/topGO.html

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anderson, L. E. (1980). Cytology and reproductive biology of mosses. In R. J. Taylor & L. E. Levotin (Eds.), *The mosses of North America* (pp. 37-76). San Francisco: California Academy of Sciences.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data [Computer software]. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Arndt, V., Rogon, C., & Höhfeld, J. (2007). To be, or not to be — molecular chaperones in protein degradation. *Cellular and Molecular Life Sciences*, *64*(19), 2525-2541. https://doi.org/10.1007/s00018-007-7188-6

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25-29. https://doi.org/10.1038/75556

Bączkiewicz, A., Szczecińska, M., Sawicki, J., Stebel, A., & Buczkowska, K. (2017). DNA barcoding, ecology and geography of the cryptic species of *Aneura pinguis* and their relationships with *Aneura maxima* and *Aneura mirabilis* (Metzgeriales, Marchantiophyta). *PloS ONE*, *12*(12), e0188837. https://doi.org/10.1371/journal.pone.0188837

Bainard, J. D., Forrest, L. L., Goffinet, B., & Newmaster, S. G. (2013). Nuclear DNA content variation and evolution in liverworts. *Molecular Phylogenetics and Evolution*, *68*(3), 619-627. https://doi.org/10.1016/j.ympev.2013.04.008

Bainard, J. D., & Villarreal, J. C. (2013). Genome size increases in recently diverged hornwort clades. *Genome*, *56*(8), 431-435. https://doi.org/10.1139/gen-2013-0041

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., ... Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS ONE*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Balkenhol, N., Cushman, S., Storfer, A., & Waits, L. (2015). *Landscape genetics: concepts, methods, applications*. Chichester, United Kingdom: John Wiley & Sons.

Bates, D. M., Maechler, M., Bolker, B. M., & Walker, S. (2014). lme4: Linear mixed - effects models using Eigen and S4 (Version R package version 1.1 - 6). Retrieved from https://cran.r-project.org/web/packages/lme4/

Behjati, S., & Tarpey, P. S. (2013). What is next generation sequencing?. *Archives of Disease in Childhood. Education and Practice Edition*, *98*(6), 236–238. https://doi.org/10.1136/archdischild-2013-304340

Bello, F. D., Lavorel, S., Lavergne, S., Albert, C. H., Boulangeat, I., Mazel, F., & Thuiller, W. (2013). Hierarchical effects of environmental filters on the functional structure of plant communities: a case study in the French Alps. *Ecography*, *36*(3), 393-402. https://doi.org/10.1111/j.1600-0587.2012.07438.x

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., ... Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*(3), 148-155. https://doi.org/10.1016/j.tree.2006.11.004

Biersma, E. M., Jackson, J. A., Hyvönen, J., Koskinen, S., Linse, K., Griffiths, H., & Convey, P. (2017). Global biogeographic patterns in bipolar moss species. *Royal Society Open Science*, *4*(7), 170147. https://doi.org/10.1098/rsos.170147

Birky, C. W. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences*, *92*(25), 11331-11338. https://doi.org/10.1073/pnas.92.25.11331

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., ... Schneider, M. (2003). The SWISS-PROT protein knowledgebase and its supplement

TrEMBL in 2003. *Nucleic Acids Research*, *31*(1), 365-370. https://doi.org/10.1093/nar/gkg095

Boom, R. C. J. A., Sol, C. J., Salimans, M. M., Jansen, C. L., Wertheim-van Dillen, P. M., & Van der Noordaa, J. P. M. E. (1990). Rapid and simple method for purification of nucleic acids. *Journal of Clinical Microbiology*, *28*(3), 495-503.

Bopp, M., & Brandes, H. (1964). Versuche zur Analyse der Protonemaentwicklung der Laubmoose: II. Über den Zusammenhang zwischen Protonemadifferenzierung und Kinetinwirkung bei der Bildung von Moosknospen. *Planta*, *62*, 116-136.

Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., ... Schmutz, J. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, *171*(2), 287-304. https://doi.org/10.1016/j.cell.2017.09.030

Breinholt, J. W., Carey, S. B., Tiley, G. P., Davis, E. C., Endara, L., McDaniel, S. F., ... Burleigh, J. G. (2021). A target enrichment probe set for resolving the flagellate land plant tree of life. *Applications in Plant Sciences*, e11406. https://doi.org/10.1002/aps3.11406

Brugués, M., & Ruiz, E. (2010). *Funaria* Hedw. In J. Guerra, M. Brugués, M. J. Cano, & R. M. Cros (Eds.), *Flora Briofítica Ibérica*, Vol. 4, *Funariales, Splachnales, Schistostegales, Bryales, Timmiales* (pp. 59-63). Murcia, Spain: Universidad de Murcia, Sociedad Española de Briología.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59-60. https://doi.org/10.1038/nmeth.3176

Casas, C., Brugués, M., Cros, R. M., & Sérgio, C. (2006). *Handbook of mosses of the Iberian Peninsula and the Balearic Islands: illustrated keys to genera and species*. Barcelona, Spain: Institut d'Estudis Catalans. Retrieved from https://floramontiberica.files.wordpress.com/2014/01/musgos-peninsula-iberica-casas-al-2006-ocr.pdf

Castro, S., & Loureiro, J. (2014). El papel de la reproducción en el origen y la evolución de las plantas poliploides. *Ecosistemas*, *23*(3), 67-77. https://doi.org/10.7818/ECOS.2014.23-3.09

Charlesworth, B. (1991). When to be diploid. *Nature*, 351, 273-274.

Chazarra, A., Flórez, E., Peraza, B., Tohá, T., Lorenzo, B., Criado, E., ... Botey, R. (2018). *Mapas climáticos de España (1981-2010) y ETo (1996-2016)*. Madrid: Spain. Agencia Estatal de Meteorología. NIPO: 014-18-004-2. Retrived from https://www.aemet.es/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/detalles/MapasclimaticosdeEspana19812010

Chiang, H. L., Terlecky, S. R., Plant, C. P., & Dice, J. F. (1989). A role for a 70-kilodalton heat shock protein in lysosomal degradation of intracellular proteins. *Science*, *246*(4928), 382-385. https://doi.org/10.1126/science.2799391

Chuanromanee, T. S., Cohen, J. I., & Ryan, G. L. (2019). Morphological Analysis of Size and Shape (MASS): An integrative software program for morphometric analyses of leaves. *Applications in Plant Sciences*, *7*(9), e11288. https://doi.org/10.1002/aps3.11288

Clarke, R. T., Rothery, P., & Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*(3), 361. https://doi.org/10.1198/108571102320

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

Crandall-Stotler, B. J., & Bartholomew-Began, S. E. (2007). Morphology of mosses (Phylum Bryophyta). In Flora of North America Editorial Committee (Eds.), *Flora of North America North of Mexico*, Vol. 27, *Bryophytes: Mosses,* part 1 (pp. 3-13). Retrieved from http://flora.huh.harvard.edu/FloraData/001/WebFiles/fna27/FNA27-Chapter1.pdf

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158. https://doi.org/10.1093/bioinformatics/btr330

Davidson, N. M., & Oshlack, A. (2014). Corset: enabling differential gene expression analysis for *de novo* assembled transcriptomes. *Genome Biology*, *15*(7), 1-14. https://doi.org/10.1186/s13059-014-0410-6

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879-886. https://doi.org/10.1080/10635150701701083

Dice, J. F. (1990). Peptide sequences that target cytosolic proteins for lysosomal proteolysis. *Trends in Biochemical Sciences*, *15*(8), 305-309. https://doi.org/10.1016/0968-0004(90)90019-8

Doležel, J., Binárová, P., & Lucretti, S. (1989). Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum,* *31*(2), 113–120. https://doi.org/10.1007/BF02907241

Donath, A., & Stadler, P. F. (2018). Split-inducing indels in phylogenomic analysis. *Algorithms for Molecular Biology*, *13*(1), 12. https://doi.org/10.1186/s13015-018-0130-7

Douzery, E. J., Pridgeon, A. M., Kores, P., Linder, H. P., Kurzweil, H., & Chase, M. W. (1999). Molecular phylogenetics of Diseae (Orchidaceae): a contribution from nuclear ribosomal ITS sequences. *American Journal of Botany*, *86*(6), 887-899. https://doi.org/10.2307/2656709

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*, 13-15.

Duminil, J., & Di Michele, M. (2009). Plant species delimitation: a comparison of morphological and molecular markers. *Plant Biosystems*, *143*(3), 528-542. https://doi.org/10.1080/11263500902722964

Eaton, D. A. R. & Overcast, I. (2016). ipyrad: interactive assembly and analysis of RADseq data sets [Computer software]. Retrieved from http://ipyrad.readthedocs.io/

Eguiarte, L. E., Souza, V., & Aguirre, X. (compilers) (2007). *Ecología Molecular*. Coyoacán, México: Instituto Nacional de Ecología, SEMARNAT.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., ... Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*, *47*(D1), D427-D432. https://doi.org/10.1093/nar/gky995

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS ONE*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Eppley, S. M., Taylor, P. J., & Jesson, L. K. (2007). Self-fertilization in mosses: a comparison of heterozygote deficiency between speciesrahm with combined *versus* separate sexes. *Heredity*, *98*(1), 38-44. https://doi.org/10.1038/sj.hdy.6800900

Fife, A. J. (1985). A generic revision of the *Funariaceae* (Bryophyta: Musci). Part I. *Journal of the Hattori Botanical Laboratory*, *58*, 149-196.

Fife, A. J. & Seppelt, R. D. (2012), *Australian Mosses Online 67. Funariaceae.* Canberra, Australia: Australian Biological Resources Study. Retrieved from http://www.anbg.gov.au/abrs/Mosses_online/67_Funariaceae.html

Fisher, R. A. (1949). *Theory of Inbreeding*. London, United Kingdom: Oliver and Boyd.

Frahm, J. P. (2008). Diversity, dispersal and biogeography of bryophytes (mosses). *Biodiversity and Conservation, 17*(2), 277-284. https://doi.org/10.1007/s10531-007-9251-x

Freeland, J. R. (2020). *Molecular Ecology* (3rd ed.). Hoboken, NJ: Wiley-Blackwell.

Freeland, J. R., Biss, P., Conrad, K. F., & Silvertown, J. (2010). Selection pressures have caused genome-wide population differentiation of *Anthoxanthum odoratum* despite the potential for high gene flow. *Journal of Evolutionary Biology*, *23*(4), 776-782. https://doi.org/10.1111/j.1420-9101.2010.01947.x

Frey, W., & Kürschner, H. (2011). Asexual reproduction, habitat colonization and habitat maintenance in bryophytes. *Flora*, *206*(3), 173-184. https://doi.org/10.1016/j.flora.2010.04.020

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925-929. https://doi.org/10.1111/2041-210X.12382

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*(4), 973-983. https://doi.org/10.1534/genetics.113.160572

Gao, B., Li, X., Zhang, D., Liang, Y., Yang, H., Chen, M., … Wood, A. J. (2017). Desiccation tolerance in bryophytes: The dehydration and rehydration transcriptomes in the desiccation-tolerant bryophyte *Bryum argenteum*. *Scientific Reports*, *7*(1), 1-14. https://doi.org/10.1038/s41598-017-07297-3

García-Olivares, V., Patiño, J., Overcast, I., Salces-Castellano, A., López de Heredia, U., Mora-Márquez, F., … Emerson, B. C. (2019). A topoclimate model for Quaternary insular speciation. *Journal of Biogeography*, *46*, 2769–2786. https://doi.org/10.1111/jbi.13689

Garilleti, R., & Albertos, B. (Coords.) (2012). *Atlas y Libro Rojo de los Briófitos Amenazados de España*. Madrid, Spain: Organismo Autónomo Parques Nacionales. Retrieved from https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/briofitos_tcm30-198033.pdf

Gause, G. F. (1934). *The Struggle for Existence* (1st ed.). Baltimore, MD: Williams & Wilkins. Retrieved from https://asantos.webs.ull.es/The%20Struggle%20for%20Existence.pdf

Geiser, D. M., Pitt, J. I., & Taylor, J. W. (1998). Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. *Proceedings of the National Academy of Sciences*, *95*(1), 388-393. https://doi.org/10.1073/pnas.95.1.388

Glime, J. M. (2017). Water Relations. Chapt. 7. In J. M. Glime (Ed.), *Bryophyte Ecology*, Vol. 1, *Physiological Ecology*. [Ebook sponsored by Michigan Technological University and the International Association of Bryologists]. Retrieved from http://digitalcommons.mtu.edu/bryophyte-ecology/

Glime, J. M., & Bisang, I. (2017). Sexual strategies. Chapt. 3. In J. M. Glime(Ed.), *Bryophyte Ecology*, Vol. 1, *Physiological Ecology*. [Ebook sponsored by Michigan Technological University and the International Association of Bryologists]. Retrieved from http://digitalcommons.mtu.edu/bryophyte-ecology/

Goffinet, B., Buck, W. R., & Shaw, A. J. (2009). Morphology, anatomy, and classification of the Bryophyta. In B. Goffinet, & A. J. Shaw (Eds.), *Bryophyte Biology* (2nd ed.) (pp. 53–138). Cambridge, United Kingdom: Cambridge University Press.

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., ... Conesa, A. (2008). High-throughput functional annotation and data mining with the

Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420-3435. https://doi.org/10.1093/nar/gkn176

Gould, B., McCouch, S., & Geber, M. (2015). *De novo* transcriptome assembly and identification of gene candidates for rapid evolution of soil Al tolerance in *Anthoxanthum odoratum* at the long-term park grass experiment. *PLoS ONE*, *10*(7), e0124424. https://doi.org/10.1371/journal.pone.0124424

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644-652. https://doi.org/10.1038/nbt.1883

Grant, V. (1981). *Plant speciation*. New York, NY: Columbia University Press. https://doi.org/10.7312/gran92318

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494-1512. https://doi.org/10.1038/nprot.2013.084

Haig, D. (2016). Living together and living apart: the sexual lives of bryophytes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1706), 20150535. https://doi.org/10.1098/rstb.2015.0535

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95-98.

Hardin, G. (1960). The competitive exclusion principle. *Science*, *131*(3409), 1292-1297.

Harper, J. T., Gile, G. H., James, E. R., Carpenter, K. J., & Keeling, P. J. (2009). The inadequacy of morphology for species and genus delineation in microbial eukaryotes: an example from the parabasalian termite symbiont *Coronympha*. *PLoS ONE*, *4*(8), e6577. https://doi.org/10.1371/journal.pone.0006577

Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. *Nature*, *381*(6583), 571-580. https://doi.org/10.1038/381571a0

Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, *107*(1), 1-8. https://doi.org/10.1016/j.ygeno.2015.11.003

Hebert, P. D., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, *101*(41), 14812-14817. https://doi.org/10.1073/pnas.0406166101

Hedenäs, L. (2020). Cryptic speciation revealed in Scandinavian *Racomitrium lanuginosum* (Hedw.) Brid. (Grimmiaceae). *Journal of Bryology*, 42(2), 117-127. https://doi.org/10.1080/03736687.2020.1722923

Hedenäs, L., Désamoré, A., Laenen, B., Papp, B., Quandt, D., González-Mancebo, J. M., … Stech, M. (2014). Three species for the price of one within the moss *Homalothecium sericeum* sl. *Taxon*, *63*(2), 249-257. https://doi.org/10.12705/632.16

Hedenäs, L., & Eldenäs, P. (2007). Cryptic speciation, habitat differentiation, and geography in *Hamatocaulis vernicosus* (Calliergonaceae, Bryophyta). *Plant Systematics and Evolution*, *268*(1-4), 131-145. https://doi.org/10.1007/s00606-007-0529-y

Hedges, S. B., Tao, Q., Walker, M., & Kumar, S. (2018). Accurate timetrees require accurate calibrations. *Proceedings of the National Academy of Sciences*, *115*(41), E9510-E9511. https://doi.org/10.1073/pnas.1812558115

Hedrick, P. W. (1987). Population genetics of intragametophytic selfing. *Evolution*, *41*(1), 137-144. https://doi.org/10.1111/j.1558-5646.1987.tb05776.x

Heinrichs, J., Kreier, H. P., Feldberg, K., Schmidt, A. R., Zhu, R. L., Shaw, B., ... Wissemann, V. (2011). Formalizing morphologically cryptic biological entities: new insights from DNA taxonomy, hybridization, and biogeography in the leafy liverwort *Porella platyphylla* (Jungermanniopsida, Porellales). *American Journal of Botany*, *98*(8), 1252-1262. https://doi.org/10.3732/ajb.1100115

Hodgetts, N. G., Söderström, L., Blockeel, T. L., Caspari, S., Ignatov, M. S., Konstantinova, N. A., … Porley, R. D. (2020). An annotated checklist of bryophytes of Europe, Macaronesia and Cyprus. *Journal of Bryology*, *42*(1), 1-116. https://doi.org/10.1080/03736687.2019.1694329

Huelsenbeck, J. P., Larget, B., & Alfaro, M. E. (2004). Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution*, *21*(6), 1123-1133. https://doi.org/10.1093/molbev/msh123

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*(8), 754-755. https://doi.org/10.1093/bioinformatics/17.8.754

Ignatova, E. A. & Fedosov, V. E. (2017). *Funariaceae*. In M. S. Ignatov (Ed.), *Moss flora of Russia,* Vol. 2, *Oedipodiales–Grimmiales* (pp. 134-174). Retrieved from http://arctoa.ru/Flora/taxonomy-ru/taxonomy-ru.php

Infante, M., Puelles, L. M., Albertos, B., Garilleti, R., & Heras, P. (2017). View on bryophyte conservation in Peninsular and Balearic Spain: analysis of Red Lists and legal protection. *Cryptogamie, Bryologie*, *38*(1), 19-51. https://doi.org/10.7872/cryb/v38.iss1.2017.19

Iseli, C., Jongeneel, C. V., & Bucher, P. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proceedings. International Conference on Intelligent Systems for Molecular Biology*, *99*, 138-148.

Jia, K. H., Zhao, W., Maier, P. A., Hu, X. G., Jin, Y., Zhou, S. S., ... Mao, J. F. (2020). Landscape genomics predicts climate change-related genetic offset for the widespread

*Platycladus orientalis* (Cupressaceae). *Evolutionary Applications*, *13*(4), 665-676. https://doi.org/10.1111/eva.12891

Jiggins, C. D. (2019). Can genomics shed light on the origin of species?. *PLoS Biology*, *17*(8), e3000394. https://doi.org/10.1371/journal.pbio.3000394

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Juárez-Barrera, F., Bueno-Hernández, A., Morrone, J. J., Barahona-Echeverría, A., & Espinosa, D. (2018). Recognizing spatial patterns of biodiversity during the nineteenth century: the roots of contemporary biogeography. *Journal of Biogeography*, *45*(5), 995-1002. https://doi.org/10.1111/jbi.13218

Kadereit, G., Piirainen, M., Lambinon, J., & Vanderpoorten, A. (2012). Cryptic taxa should have names: Reflections in the glasswort genus *Salicornia* (Amaranthaceae). *Taxon*, *61*(6), 1227-1239. https://doi.org/10.1002/tax.616005

Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Science*, *28*(11), 1947-1951. https://doi.org/10.1002/pro.3715

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Research*, *49*(D1), D545-D551. https://doi.org/10.1093/nar/gkaa970

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27-30. https://doi.org/10.1093/nar/28.1.27

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*, 170122. https://doi.org/10.1038/sdata.2017.122

Kirbis, A., Waller, M., Ricca, M., Bont, Z., Neubauer, A., Goffinet, B., & Szövényi, P. (2020). Transcriptional landscapes of divergent sporophyte development in two mosses, *Physcomitrium (Physcomitrella) patens* and *Funaria hygrometrica*. *Frontiers in Plant Science*, *11*, 747. https://doi.org/10.3389/fpls.2020.00747

Klekowski, E. J. (1979). The genetics and reproductive biology of ferns. In A. F. Dyer (Ed.). *The Experimental Biology of Ferns* (pp 133–170). London, United Kingdom: Academic Press.

Kofler, L. (1959). Contribution à l'étude biologique des mousses cultivées *in vitro*. Germination des spores, croissance et dévelopment du protonema chez *Funaria hygrometrica*. *Revue Bryologique et Lichénologique*, *24*, 1-199.

Kon, M., & Cuervo, A. M. (2010). Chaperone-mediated autophagy in health and disease. *FEBS Letters*, *584*(7), 1399-1404. https://doi.org/10.1016/j.febslet.2009.12.025

Košnar, J., Herbstová, M., Kolář, F., Koutecký, P., & Kučera, J. (2012). A case study of intragenomic ITS variation in bryophytes: assessment of gene flow and role of polyploidy in the origin of European taxa of the *Tortula muralis* (Musci: Pottiaceae) complex. *Taxon*, *61*(4), 709-720. https://doi.org/10.1002/tax.614001

Koutecký, P. (2015). MorphoTools: a set of R functions for morphometric analysis. *Plant Systematics and Evolution*, *301*(4), 1115-1121. https://doi.org/10.1007/s00606-014-1153-2

Laenen, B., Machac, A., Gradstein, S. R., Shaw, B., Patiño, J., Désamoré, A., … Vanderpoorten, A. (2016). Geographical range in liverworts: does sex really matter?. *Journal of Biogeography*, *43*(3), 627-635. https://doi.org/10.1111/jbi.12661

Lang, A. S., Kruijer, J. D., & Stech, M. (2014). DNA barcoding of Arctic bryophytes: an example from the moss genus *Dicranum* (Dicranaceae, Bryophyta). *Polar Biology*, *37*(8), 1157-1169. https://doi.org/10.1007/s00300-014-1509-7

Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., … Rensing, S. A. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal*, *93*(3), 515-533. https://doi.org/10.1111/tpj.13801

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357. https://doi.org/10.1038/nmeth.1923

Lehnert, B., & Bopp, M. (1983). The hormonal regulation of protonema development in mosses I. Auxin-cytokinin interaction. *Zeitschrift für Pflanzenphysiologie*, *110*(5), 379-391. https://doi.org/10.1016/S0044-328X(83)80189-5

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183-189. https://doi.org/10.1016/S0169-5347(02)02497-7

Lepais, O., & Weir, J. T. (2014). Sim RAD: an R package for simulation-based prediction of the number of loci expected in RAD seq and similar genotyping by sequencing approaches. *Molecular Ecology Resources*, *14*(6), 1314-1321. https://doi.org/10.1111/1755-0998.12273

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, *12*(1), 323. https://doi.org/10.1186/1471-2105-12-323

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Ligrone, R., Duckett, J. G., & Renzaglia, K. S. (2000). Conducting tissues and phyletic relationships of bryophytes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1398), 795-813. https://doi.org/10.1098/rstb.2000.0616

Lipnerová, I., Bures, P., Horová, L., & Smarda, P. (2013). Evolution of genome size in *Carex* (Cyperaceae) in relation to chromosome number and genomic base composition. *Annals of Botany, 111*(1), 79-94. https://doi.org/10.1093/aob/mcs239

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., … Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International*, *2012*, 251364. https://doi.org/10.1155/2012/251364

Liu, Y., Budke, J. M., & Goffinet, B. (2012a). Phylogenetic inference rejects sporophyte based classification of the Funariaceae (Bryophyta): rapid radiation suggests rampant homoplasy in sporophyte evolution. *Molecular Phylogenetics and Evolution*, *62*(1), 130-145. https://doi.org/10.1016/j.ympev.2011.09.010

Liu, Y., Moskwa, N. L., & Goffinet, B. (2012b). Development of eight mitochondrial markers for Funariaceae (Musci) and their amplification success in other mosses. *American Journal of Botany*, *99*(2), e62-e65. https://doi.org/10.3732/ajb.1100402

López Fernández, M. L., Piñas, S., & López-Fernández, M. S. (2008). Macrobioclimas, Bioclimas y Variantes Bioclimáticas de la España Peninsular y Balear, y su Cartografía. Retrieved from http://hdl.handle.net/10171/8097

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Magdy, M., Werner, O., McDaniel, S. F., Goffinet, B., & Ros, R. M. (2016). Genomic scanning using AFLP to detect loci under selection in the moss *Funaria hygrometrica* along a climate gradient in the Sierra Nevada Mountains, Spain. *Plant Biology, 18*(2), 280-288. https://doi.org/10.1111/plb.12381

Magdy Abdallah Awad, M. (2013). *Genetic variability in mosses and its relation to climate change adaptation processes in Mediterranean environments*. (Doctoral dissertation). Retrieved from http://hdl.handle.net/10201/35602

Magill, R. E. (1990). Glossarium polyglottum bryologiae. A multilingual glossary for bryology. *Monographs in Systematic Botany*, *33, 1-297*.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614-621. https://doi.org/10.1016/j.tree.2013.05.012

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189-197. https://doi.org/10.1016/S0169-5347(03)00008-9

Manukjanová, A., Košnar, J., & Kučera, J. (2018). Microsatellite primers for the cryptic species of the moss *Hamatocaulis vernicosus* and methods for their quick barcoding. *Journal of Bryology*, *40*(3), 302-305. https://doi.org/10.1080/03736687.2018.1450199

Mao, X., Cai, T., Olyarchuk, J. G., & Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*, *21*(19), 3787-3793. https://doi.org/10.1093/bioinformatics/bti430

Marques, I., Loureiro, J., Draper, D., Castro, M., & Castro, S. (2018). How much do we know about the frequency of hybridisation and polyploidy in the Mediterranean region?. *Plant Biology*, *20*, 21-37. https://doi.org/10.1111/plb.12639

Martin, A. (2016). *The magical world of moss gardening*. Portland, OR: Timber Press.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*(1), 28-41. https://doi.org/10.1111/1755-0998.12291

McDaniel, S. F. (2009). The genetic basis of natural variation in bryophyte model systems. *Annual Plant Reviews online, 36*, 16-41. https://doi.org/10.1002/9781119312994.apr0385

McDaniel, S. F., Neubig, K. M., Payton, A. C., Quatrano, R. S., & Cove, D. J. (2013). Recent gene-capture on the UV sex chromosomes of the moss *Ceratodon purpureus*. *Evolution*, *67*(10), 2811-2822. https://doi.org/10.1111/evo.12165

McDaniel, S. F., & Shaw, A. J. (2003). Phylogeographic structure and cryptic speciation in the trans-antarctic moss *Pyrrhobryum mnioides*. *Evolution*, *57*(2), 205-215. https://doi.org/10.1111/j.0014-3820.2003.tb00256.x

McDaniel, S. F., Von Stackelberg, M., Richardt, S., Quatrano, R. S., Reski, R., & Rensing, S. A. (2010). The speciation history of the *Physcomitrium—Physcomitrella* species complex. *Evolution: International Journal of Organic Evolution*, *64*(1), 217-231. https://doi.org/10.1111/j.1558-5646.2009.00797.x

McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*, *351*(6328), 652-654. https://doi.org/10.1038%2F351652a0

McIntosh, T. T. (2007). *Funariaceae*. In Flora of North America Editorial Committee (Eds.), *Flora of North America North of Mexico*, Vol. 27, *Bryophytes: Mosses,* Part 1 (pp. 180-199). Retrieved from http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=10352

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*(8), 1551-1561. https://doi.org/10.1111/j.0014-3820.2006.tb00500.x

McRae, B. H., & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences*, *104*(50), 19885-19890. https://doi.org/10.1073/pnas.0706568104

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*(10), 2712-2724. https://doi.org/10.1890/07-1861.1

Medina, N. G., Draper, I., & Lara, F. (2011). Biogeography of mosses and allies: Does size matter? In D. Fontaneto (Ed.), *Biogeography of Microscopic Organisms: Is Everything Small Everywhere?* (pp. 209-233). Cambridge, United Kingdom: Cambridge University Press. https://doi.org/10.1017/CBO9780511974878.012

Medina, R., Lara, F., Goffinet, B., Garilleti, R., & Mazimpaka, V. (2012). Integrative taxonomy successfully resolves the pseudo-cryptic complex of the disjunct epiphytic moss *Orthotrichum consimile* s.l. (Orthotrichaceae). *Taxon*, *61*(6), 1180-1198. https://doi.org/10.1002/tax.616002

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, *17*, 240–248. https://doi.org/10.1101/gr.5681207

Mishler, B. D. (2001). The biology of bryophytes—bryophytes aren't just small tracheophytes. *American Journal of Botany, 88*, 2129-2131. https://doi.org/10.2307/3558438

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, *35*(suppl_2), W182-W185. https://doi.org/10.1093/nar/gkm321

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., … Donoghue, P. C. J. (2018). The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences*, *115*(10), E2274-E2283. https://doi.org/10.1073/pnas.1719588115

Morueta-Holme, N., Engemann, K., Sandoval-Acuña, P., Jonas, J. D., Segnitz, R. M., & Svenning, J. C. (2015). Strong upslope shifts in Chimborazo's vegetation over two centuries since Humboldt. *Proceedings of the National Academy of Sciences*, *112*(41), 12741-12745. https://doi.org/10.1073/pnas.1509938112

Muñoz, J., Felicísimo, Á. M., Cabezas, F., Burgaz, A. R., & Martínez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, *304*(5674), 1144-1147. https://doi.org/10.1126/science.1095210

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853-858. https://doi.org/10.1038/35002501

Nakosteen P. C., & Hughes K. W. (1978). Sexual life cycle of three species of Funariaceae in culture. *The Bryologist, 81*, 307-314. https://doi.org/10.2307/3242191

Natcheva, R., & Cronberg, N. (2007). Maternal transmission of cytoplasmic DNA in interspecific hybrids of peat mosses, *Sphagnum* (Bryophyta). *Journal of Evolutionary Biology*, *20*(4), 1613-1616. https://doi.org/10.1111/j.1420-9101.2007.01341.x

Newton, M. E. (1984). The cytogenetics of bryophytes. In A. F. Dyer, & J. G. Duckett (Eds.), *The experimental biology of bryophytes* (pp. 65–96). London, United Kingdom: Academic Press.

Nieto-Lugilde, M., Werner, O., McDaniel, S. F., Koutecký, P., Kučera, J., Rizk, S. M., & Ros, R. M. (2018a). Peripatric speciation associated with genome expansion and female-biased sex ratios in the moss genus *Ceratodon*. *American Journal of Botany 105*(6), 1–12. https://doi.org/10.1002/ajb2.1107

Nieto-Lugilde, M., Werner, O., McDaniel, S. F., & Ros, R. M. (2018b). Environmental variation obscures species diversity in southern European populations of the moss genus *Ceratodon*. *Taxon*, *67*(4), 673-692. https://doi.org/10.12705/674.1

Nogués-Bravo, D., Araújo, M. B., Romdal, T., & Rahbek, C. (2008). Scale effects and human impact on the elevational species richness gradients. *Nature*, *453*(7192), 216-219. https://doi.org/10.1038/nature06812

Ochyra, R., Lewis, S., & Bednarek-Ochyra, H. (2008). *The illustrated moss flora of Antarctica*. Cambridge, United Kingdom: Cambridge University Press.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., … Ohyama, K. (1992). Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. *Journal of Molecular Biology*, *223*(1), 1-7. https://doi.org/10.1016/0022-2836(92)90708-R

Ohsawa, T., & Ide, Y. (2008). Global patterns of genetic variation in plant species along vertical and horizontal gradients on mountains. *Global Ecology and Biogeography*, *17*(2), 152-163. https://doi.org/10.1111/j.1466-8238.2007.00357.x

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., … Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature*, *322*(6079), 572-574. https://doi.org/10.1038/322572a0

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., ... Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733-D745. https://doi.org/10.1093/nar/gkv1189

Oliver, M. J., Velten, J., & Mishler, B. D. (2005). Desiccation tolerance in bryophytes: a reflection of the primitive strategy for plant survival in dehydrating habitats?. *Integrative and Comparative Biology*, *45*(5), 788-799. https://doi.org/10.1093/icb/45.5.788

Olson, M. S., & McCauley, D. E. (2000). Linkage disequilibrium and phylogenetic congruence between chloroplast and mitochondrial haplotypes in *Silene vulgaris*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1454), 1801-1808. https://doi.org/10.1098/rspb.2000.1213

O'Shea, B. J. (2006). Checklist of the mosses of sub-Saharan Africa (version 5, 12/06). *Tropical Bryology Research Reports*, *6*, 1-252. Retrieved from https://websites.rbge.org.uk/bbs/Activities/tbg/Files/O'Shea%202006%20Checklist%20of%20sub-Saharan%20African%20mosses.pdf

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*(1), 401-437. https://doi.org/10.1146/annurev.genet.34.1.401

Patiño, J., Carine, M., Fernández-Palacios, J. M., Otto, R., Schaefer, H., & Vanderpoorten, A. (2014). The anagenetic world of spore-producing land plants. *New Phytologist*, *201*(1), 305-311. https://doi.org/10.1111/nph.12480

Patiño, J., Hedenäs, L., Dirkse, G. M., Ignatov, M. S., Papp, B., Müller, F., ... Vanderpoorten, A. (2017). Species delimitation in the recalcitrant moss genus *Rhynchostegiella* (Brachytheciaceae). *Taxon*, *66*(2), 293-308. https://doi.org/10.12705/662.1

Patiño, J., Medina, R., Vanderpoorten, A., González-Mancebo, J. M., Werner, O., Devos, N., … Ros, R. M. (2013). Origin and fate of the single-island endemic moss *Orthotrichum handiense*. *Journal of Biogeography*, *40*(5), 857-868. https://doi.org/10.1111/jbi.12051

Patiño, J., & Vanderpoorten, A. (2018). Bryophyte biogeography. *Critical Reviews in Plant Sciences*, *37*(2-3), 175-209. https://doi.org/10.1080/07352689.2018.1482444

Perley, D. S., & Jesson, L. K. (2015). Hybridization is associated with changes in sexual system in the bryophyte genus *Atrichum*. *American Journal of Botany*, *102*(4), 555-565. https://doi.org/10.3732/ajb.1400494

Peterman, W. E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, *9*(6), 1638-1647. https://doi.org/10.1111/2041-210X.12984

Peterman, W. E., Connette, G. M., Semlitsch, R. D., & Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Molecular Ecology*, *23*(10), 2402-2413. https://doi.org/10.1111/mec.12747

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94-100. https://doi.org/10.1038/ng.3464

Pisa, S., Biersma, E. M., Convey, P., Patiño, J., Vanderpoorten, A., Werner, O., & Ros, R. M. (2014). The cosmopolitan moss *Bryum argenteum* in Antarctica: recent colonisation or in situ survival?. *Polar Biology*, *37*(10), 1469-1477. https://doi.org/10.1007/s00300-014-1537-3

Pisa, S., Werner, O., Vanderpoorten, A., Magdy, M., & Ros, R. M. (2013). Elevational patterns of genetic variation in the cosmopolitan moss *Bryum argenteum* (Bryaceae). *American Journal of Botany*, *100*(10), 2000-2008. https://doi.org/10.3732/ajb.1300100

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. *PLoS ONE* *7*(2), e32253. https://doi.org/10.1371/journal.pone.0032253

Poland, J. A., & Rife, T. W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. *The Plant Genome Journal*, 5(3), 92. https://doi.org/10.3835/plantgenome2012.05.0005

Polo, M. J., Herrero, J., Pimentel, R., & Pérez-Palazón, M. J. (2019). The Guadalfeo Monitoring Network (Sierra Nevada, Spain): 14 years of measurements to understand the complexity of snow dynamics in semiarid regions. *Earth System Science Data*, *11*(1). https://doi.org/10.5194/essd-11-393-2019

Pressel, S., & Duckett, J. G. (2010). Cytological insights into the desiccation biology of a model system: moss protonemata. *New Phytologist*, *185*(4), 944-963. https://doi.org/10.1111/j.1469-8137.2009.03148.x

Proctor, M. C. (2000). The bryophyte paradox: tolerance of desiccation, evasion of drought. *Plant Ecology*, *151*(1), 41-49. https://doi.org/10.1023/A:1026517920852

Proctor, M. C. (2009). Physiological ecology. In B. Goffinet, & A. J. Shaw (Eds.), *Bryophyte Biology* (2nd ed.) (pp. 237–268). Cambridge, United Kingdom: Cambridge University Press.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [Computer software]. Retrieved from https://www.R-project.org/

Rahmatpour, N., Perera, N. V., Singh, V., Wegrzyn, J. L., & Goffinet, B. (2021). High gene space divergence contrasts with frozen vegetative architecture in the moss family Funariaceae. *Molecular Phylogenetics and Evolution*, *154*, 106965. https://doi.org/10.1016/j.ympev.2020.106965

Rapley, R., & Harbron, S. (Eds.). (2012). *Molecular analysis and genome discovery* (2nd ed.). Hoboken, NJ: Wiley-Blackwell.

Renner, M. A. (2020). Opportunities and challenges presented by cryptic bryophyte species. *Telopea*, *23*, 41-60. http://dx.doi.org/10.7751/telopea14083

Rensing, S. A. (2014). Gene duplication as a driver of plant morphogenetic evolution. *Current Opinion in Plant Biology*, *17*, 43-48. https://doi.org/10.1016/j.pbi.2013.11.002

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., … Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*, *319*(5859), 64-69. https://doi.org/10.1126/science.1150646

Ricca, M., Szövényi, P., Temsch, E. M., Johnson, M. G., & Shaw, A. J. (2011). Interploidal hybridization and mating patterns in the *Sphagnum subsecundum* complex. *Molecular Ecology*, *20*(15), 3202-3218. https://doi.org/10.1111/j.1365-294X.2011.05170.x

Richardson, D. H. S. (1981) *The Biology of Mosses*. Oxford, United Kingdom: Blackwell Scientific Publications.

Riesch, R., Plath, M., & Bierbach, D. (2018). Ecology and evolution along environmental gradients. *Current Zoology*, *64*(2), 193-196. https://doi.org/10.1093/cz/zoy015

Rivas-Martínez, S., Penas, A., & Díaz, T. E. (2004). *Bioclimatic Map of Europe, Bioclimates*. Leon, Spain: Cartographic Service, University of Leon. Retrieved from http://www.ucm.es/info/cif/form/maps.htm

Rodríguez Llopis, M. (Dir.), & Martínez Carrión, J. M. (Coord.) (2006). *Atlas histórico ilustrado de la Región de Murcia y su antiguo Reino*. Murcia, Spain: Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*(12), 1572-1574. https://doi.org/10.1093/bioinformatics/btg180

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*(3), 539-542. https://doi.org/10.1093/sysbio/sys029

Ros, R. M., Mazimpaka, V., Abou-Salama, U., Aleffi, M., Blockeel, T. L., Brugués, M., … Werner, O. (2013a). Mosses of the Mediterranean, an annotated checklist. *Cryptogamie, Bryologie*, *34*(2), 99-283. https://doi.org/10.7872/cryb.v34.iss2.2013.99

Ros, R. M., Werner, O., & Pérez-Álvarez, J. R. (2013b). *Ex situ* conservation of rare and threatened Mediterranean bryophytes. *Flora Mediterranea 23*, 223-235. Retrieved from https://www.herbmedit.org/flora/23-223.pdf

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, *34*(12), 3299-3302. https://doi.org/10.1093/molbev/msx248

Salces-Castellano, A., Patiño, J., Alvarez, N., Andújar, C., Arribas, P., Braojos-Ruiz, J. J., … Emerson, B. C. (2020). Climate drives community-wide divergence within species over a limited spatial scale: evidence from an oceanic island. *Ecology Letters*, *23*(2), 305-315. https://doi.org/10.1111/ele.13433

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, *94*(3), 441-448. https://doi.org/10.1016/0022-2836(75)90213-2

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463-5467. https://doi.org/10.1073/pnas.74.12.5463

Schuster, R. M. (1983). Phytogeography of the bryophyta. In R. M. Schuster (Ed.), *New Manual of Bryology* (pp. 463-626). Nichinan, Japan: The Hattori Botanical Laboratory.

Schuster, S. (2008). Next-generation sequencing transforms today's biology. *Nature Methods*, *5*, 16–18. https://doi.org/10.1038/nmeth1156

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., ... Brelsford, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176-192. https://doi.org/10.1038/nrg3644

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2013). Genetic isolation by environment or distance: which pattern of gene flow is most common?. *Evolution*, *68*(1), 1-15. https://doi.org/10.1111/evo.12258

Shaw, A. J. (1988). Genetic variation for tolerance to copper and zinc within and among populations of the moss, *Funaria hygrometrica* Hedw. *The New Phytologist, 109*(2), 211-222. Retrieved from http://www.jstor.org/stable/2434840

Shaw, A. J. (1990). Metal tolerances and cotolerances in the moss *Funaria hygrometrica*. *Canadian Journal of Botany*, *68*(10), 2275-2282. https://doi.org/10.1139/b90-290

Shaw, A. J. (1991). The genetic structure of sporophytic and gametophytic populations of the moss, *Funaria hygrometrica* Hedw. *Evolution*, *45*(5), 1260-1274. https://doi.org/10.1111/j.1558-5646.1991.tb04391.x

Shaw, A. J. (1994). Adaptation to metals in widespread and endemic plants. *Environmental Health Perspectives*, *102*(suppl 12), 105-108. https://doi.org/10.1289/ehp.94102s12105

Shaw, A. J. (1999). The evolution of heavy metal tolerance in plants: Adaptations, limits, and costs. In V. E. Forbes (Ed.), *Genetics and ecotoxicology* (pp. 9–30). Philadelphia, PA: Taylor & Francis. Retrieved from https://books.google.es/books?id=vU4rEqu8_7UC&printsec=frontcover&hl=es

Shaw, A. J. (2001). Biogeographic patterns and cryptic speciation in bryophytes. *Journal of Biogeography*, *28*(2), 253-261. https://doi.org/10.1046/j.1365-2699.2001.00530.x

Shaw, A. J., Werner, O., & Ros, R. M. (2003). Intercontinental Mediterranean disjunct mosses: morphological and molecular patterns. *American Journal of Botany*, *90*(4), 540-550. https://doi.org/10.3732/ajb.90.4.540

Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., & Waterston, R. H. (2017). DNA sequencing at 40: past, present and future. *Nature*, *550*(7676), 345-353. https://doi.org/10.1038/nature24286

Sim-Sim, M., Lopes, T., Ruas, S., & Stech, M. (2015). Does altitude shape molecular diversity and richness of bryophytes in Madeira's natural forest? A case study with four bryophyte species at two altitudinal levels. *Plant Ecology and Evolution*, *148*(2), 171-180. https://doi.org/10.5091/plecevo.2015.1041

Smith, A. (Ed.). (1982). *Bryophyte Ecology*. London, United Kingdom: Chapman and Hall. https://doi.org/10.1007/978-94-009-5891-3

Smith, G., & Ritchie, M. G. (2013). How might epigenetics contribute to ecological speciation?. *Current Zoology*, *59*(5), 686-696. https://doi.org/10.1093/czoolo/59.5.686

Soltis, P. S., Marchant, D. B., Van de Peer, Y., & Soltis, D. E. (2015). Polyploidy and genome evolution in plants. *Current Opinion in Genetics & Development*, *35*, 119-125. https://doi.org/10.1016/j.gde.2015.11.003

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, *60*, 561-588. https://doi.org/10.1146/annurev.arplant.043008.092039

Soltis, D. E., Soltis, P. S., & Tate, J. A. (2004). Advances in the study of polyploidy since plant speciation. *New Phytologist*, *161*(1), 173-191. https://doi.org/10.1046/j.1469-8137.2003.00948.x

Soltis, D. E., Visger, C. J., Marchant, D. B., & Soltis, P. S. (2016). Polyploidy: pitfalls and paths to a paradigm. *American Journal of Botany*, *103*(7), 1146-1166. https://doi.org/10.3732/ajb.1500501

Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, *11*(1), 7. https://doi.org/10.1186/1471-2105-11-7

Struck, T. H., Feder, J. L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V. I., ... Stedje, B. (2018). Finding evolutionary processes hidden in cryptic species. *Trends in Ecology & Evolution*, *33*(3), 153-163. https://doi.org/10.1016/j.tree.2017.11.007

Szövényi, P. (2016). The genome of the model species *Anthoceros agrestis*. In S. A. Rensing (Ed.), *Advances in botanical research,* Vol. 78, *Genomes and Evolution of Charophytes, Bryophytes, Lycophytes and Ferns* (pp. 189-211). London, United Kingdom: Elsevier. https://doi.org/10.1016/bs.abr.2015.12.001

Szövényi, P., Hock, Z., Korpelainen, H., & Shaw, A. J. (2009). Spatial pattern of nucleotide polymorphism indicates molecular adaptation in the bryophyte *Sphagnum fimbriatum*.

*Molecular Phylogenetics and Evolution*, *53*(1), 277-286. https://doi.org/10.1016/j.ympev.2009.06.007

Szövényi, P., Rensing, S. A., Lang, D., Wray, G. A., & Shaw, A. J. (2010). Generation-biased gene expression in a bryophyte model system. *Molecular Biology and Evolution*, *28*(1), 803-812. https://doi.org/10.1093/molbev/msq254

Szweykowski, J., Buczkowska, K., & Odrzykoski, I. J. (2005). *Conocephalum salebrosum* (Marchantiopsida, Conocephalaceae)–a new Holarctic liverwort species. *Plant Systematics and Evolution*, *253*(1-4), 133-158. https://doi.org/10.1007/s00606-005-0301-0

Szweykowski, J., & Krzakowa, M. (1979). Variation of four enzyme systems in Polish populations of *Conocephalum conicum* (L.) Dumort. (Hepaticae, Marchantiales). *Bulletin de l'Academie Polonaise des Sciences, Série des Sciences Biologiques*, *27*, 21-35.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), 585-595.

Tan, B. C., & Pócs, T. (2000). Bryogeography and conservation of bryophytes. In A. J. Shaw, & B. Goffinet (Eds.), *Bryophyte Biology* (pp. 403–448). Cambridge, United Kingdom: Cambridge University Press.

Taylor, P. J., Eppley, S. M., & Jesson, L. K. (2007). Sporophytic inbreeding depression in mosses occurs in a species with separate sexes but not in a species with combined sexes. *American Journal of Botany*, *94*(11), 1853-1859. https://doi.org/10.3732/ajb.94.11.1853

Temsch, E. M., Greilhuber, J., & Krisai, R. (1998). Genome size in *Sphagnum* (peat moss). *Botanica Acta*, *111*(4), 325-330. https://doi.org/10.1111/j.1438-8677.1998.tb00716.x

The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, *47*(D1), D330-D338. https://doi.org/10.1093/nar/gky1055

The Plant List (2013). The Plant List. A working list of all plant species. Version 1.1. [Database]. Retrieved from http://www.theplantlist.org/

Trucchi, E., Mazzarella, A. B., Gilfillan, G. D., Lorenzo, M. T., Schönswetter, P., & Paun, O. (2016). Bs RAD seq: Screening DNA methylation in natural populations of non-model species. *Molecular Ecology*, *25*(8), 1697-1713. https://doi.org/10.1111/mec.13550

Turcatti, G., Romieu, A., Fedurco, M., & Tairi, A. P. (2008). A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Research*, *36*(4), e25-e25. https://doi.org/10.1093/nar/gkn021

Turelli, M., Barton, N. H., & Coyne, J. A. (2001). Theory and speciation. *Trends in Ecology & Evolution, 16*(7), 330-343. https://doi.org/10.1016/S0169-5347(01)02177-2

Tylianakis, J. M., & Morris, R. J. (2017). Ecological networks across environmental gradients. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 25-48. https://doi.org/10.1146/annurev-ecolsys-110316-022821

Vabulas, R. M., Raychaudhuri, S., Hayer-Hartl, M., & Hartl, F. U. (2010). Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harbor Perspectives in Biology*, *2*(12), a004390. https://doi.org/10.1101/cshperspect.a004390

Vallejo-Marin, M. (2014). La correlación entre la poliploidía y la reproducción asexual. *Ecosistemas*, *23*(3), 78-82. https://doi.org/10.7818/ECOS.2014.23-3.10

van Drunen, W. E., & Husband, B. C. (2019). Evolutionary associations between polyploidy, clonal reproduction, and perenniality in the angiosperms. *New Phytologist*, *224*(3), 1266-1277. https://doi.org/10.1111/nph.15999

van Gurp, T. P., Wagemaker, N. C., Wouters, B., Vergeer, P., Ouborg, J. N., & Verhoeven, K. J. (2016). epiGBS: reference-free reduced representation bisulfite sequencing. *Nature Methods*, *13*(4), 322-324. https://doi.org/10.1038/nmeth.3763

van Zanten, B. O. (1978). Experimental studies on trans-oceanic long-range dispersal of moss spores in the Southern Hemisphere. *Journal of the Hattori Botanical Laboratory*.

van Zanten, B. O. & Pócs, T. (1981) Distribution and disperal of bryophytes. In W. Schultze-Motel (Ed.), *Advances in Bryology,* Vol. 1 (pp 479-562). Vaduz, Liechtenstein: J. Cramer.

Vanderpoorten, A., & Goffinet, B. (2009). *Introduction to Bryophytes*. Cambridge, United Kingdom: Cambridge University Press. https://doi.org/10.1017/CBO9780511626838

Vanderpoorten, A., Gradstein, S. R., Carine, M. A., & Devos, N. (2010). The ghosts of Gondwana and Laurasia in modern liverwort distributions. *Biological Reviews*, *85*(3), 471-487. https://doi.org/10.1111/j.1469-185X.2009.00111.x

Vanderpoorten, A., & Hallingbäck, T. (2009). Conservarion biology of bryophytes. In B. Goffinet, & A. J. Shaw (Eds.), *Bryophyte Biology* (2nd ed.), pp. 487–533). Cambridge, United Kingdom: Cambridge University Press.

Vanderpoorten, A., & Shaw, A. J. (2014). The application of molecular data to the phylogenetic delimitation of species in bryophytes: A note of caution. *Phytotaxa*, *9*(1), 229-237. http://dx.doi.org/10.11646/phytotaxa.9.1.12

Venables W. N., & Ripley B. D. (2002). *Modern Applied Statistics with S* (4th ed.). New York, NY: Springer.

Vigalondo, B., Garilleti, R., Vanderpoorten, A., Patiño, J., Draper, I., Calleja, J. A., ... Lara, F. (2019). Do mosses really exhibit so large distribution ranges? Insights from the integrative taxonomic study of the *Lewinskya affinis* complex (Orthotrichaceae, Bryopsida). *Molecular Phylogenetics and Evolution*, *140*, 106598. https://doi.org/10.1016/j.ympev.2019.106598

Voglmayr, H. (2000). Nuclear DNA amounts in mosses (Musci). *Annals of Botany*, *85*(4), 531-546. https://doi.org/10.1006/anbo.1999.1103

von Wettstein, F. (1923). Kreuzungsversuche mit multiploiden Moosrassen. *Biologisches Zentralblatt*, *43*, 71-83.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649-5662. https://doi.org/10.1111/mec.12938

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57-63. https://doi.org/10.1038/nrg2484

Werner, O., Prudencio, Á. S., de la Cruz-Martínez, E., Nieto-Lugilde, M., Martínez-Gómez, P., & Ros, R. M. (2020). A cost reduced variant of epi-genotyping by sequencing for studying DNA methylation in non-model organisms. *Frontiers in Plant Science*, *11*, 694. https://doi.org/10.3389/fpls.2020.00694

Werner, O., Ros Espín, R. M., Bopp, M., & Atzorn, R. (1991). Abscisic-acid-induced drought tolerance in *Funaria hygrometrica* Hedw. *Planta*, *186*(1), 99-103. https://doi.org/10.1007/BF00201503

Wigh, K. (1973). Accessory chromosomes in some mosses. *Hereditas*, *74*(2), 211-224. https://doi.org/10.1111/j.1601-5223.1973.tb01122.x

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Wyatt, R. (1982). Population ecology of bryophytes. *Journal of the Hattori Botanical Laboratory*, *52*(1), 179-198.

Wyatt, R. (1985). Species concepts in bryophytes: input from population biology. *Bryologist*, 182-189. https://doi.org/10.2307/3243027

Yaffe, H., Buxdorf, K., Shapira, I. Ein-Gedi, S., Zvi, M. M. B., Fridman, E., … Levy, M. (2012). LogSpin: a simple, economical and fast method for RNA isolation from infected or healthy plants and other eukaryotic tissues. *BMC Research Notes 5*(1), 45. https://doi.org/10.1186/1756-0500-5-45

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*, *11*(2), R14. https://doi.org/10.1186/gb-2010-11-2-r14

Zaccara, S., Patiño, J., Convey, P., Vanetti, I., & Cannone, N. (2020). Multiple colonization and dispersal events hide the early origin and induce a lack of genetic structure of the moss *Bryum argenteum* in Antarctica. *Ecology and Evolution*, *10*(16), 8959-8975. https://doi.org/10.1002/ece3.6601

Zvelebil, M. J., & Baum, J. O. (2008). *Understanding Bioinformatics*. Adingdon, United Kingdom: Garland Science.

# IX.

# APPENDICES

**Appendix 1.** Voucher information for the specimens studied. Specimens collected from the Iberian Peninsula are ordered according to their belonging to the four geographical areas studied, identified by acronyms and sequencing code (AS: Asturias province; LE: Leon province; MU: Murcia province; SN: Sierra Nevada Mountains). Specimens studied from outside the Iberian Peninsula are ordered according to their belonging to species within *Funaria* genus. Also the next information is given: location number, location name, geographical coordinates, altitude, collection date, herbarium code, and its use for: *in vitro* culture, ploidy analysis, DNA Sanger sequencing, GBS and RNA analyses, and biometric study.

*FUNARIA HYGROMETRICA* SPECIMENS FROM THE IBERIAN PENINSULA

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AS001 | L01 | Carretera AS-256 de Gijón a Santander, km 9, desvío a Bedriñana (Villaviciosa) | 43° 30′ 24.6″ N; 05° 25′ 39.2″ W | 83 | 19/05/2015 | MUB 57133 | ✓ | ✓ | | ✓ | | |
| AS002 | L02 | Polígono industrial de Carrales, en carretera N-632 de Gijón a Santander (Caravia) | 43° 27′ 44.4″ N; 05° 11′ 47.2″ W | 107 | 19/05/2015 | MUB 57139 | ✓ | ✓ | | ✓ | | |
| AS003 | L02 | Polígono industrial de Carrales, en carretera N-632 de Gijón a Santander (Caravia) | 43° 27′ 44.4″ N; 05° 11′ 47.2″ W | 107 | 19/05/2015 | MUB 57140 | ✓ | ✓ | | ✓ | | |
| AS004 | L03 | Mina Santa Ana, cruce a Playa de Vega, en la carretera N-632 (Ribadesella) | 43° 27′ 59.7″ N; 05° 07′ 47.9″ W | 30 | 19/05/2015 | MUB 57143 | ✓ | ✓ | | ✓ | | |
| AS005 | L04 | Llovio, estación de FEVE (Ribadesella) | 43° 26′ 24.7″ N; 05° 03′ 20.7″ W | 5 | 19/05/2015 | MUB 57145 | ✓ | ✓ | | ✓ | | |
| AS006 | L04 | Llovio, estación de FEVE (Ribadesella) | 43° 26′ 24.7″ N; 05° 03′ 20.7″ W | 5 | 19/05/2015 | MUB 57146 | ✓ | | | ✓ | | |
| AS007 | L04 | Llovio, estación de FEVE (Ribadesella) | 43° 26′ 24.7″ N; 05° 03′ 20.7″ W | 5 | 19/05/2015 | MUB 57147 | ✓ | | | ✓ | | |
| AS008 | L04 | Llovio, estación de FEVE (Ribadesella) | 43° 26′ 24.7″ N; 05° 03′ 20.7″ W | 5 | 19/05/2015 | MUB 57149 | ✓ | ✓ | | ✓ | | |
| AS009 | L04 | Llovio, estación de FEVE (Ribadesella) | 43° 26′ 24.7″ N; 05° 03′ 20.7″ W | 5 | 19/05/2015 | MUB 57150 | ✓ | | | ✓ | | |
| AS010 | L05 | Cementerio entre Villahormes y Cardoso, en la carretera AS-263, km 14 (Llanes) | 43° 26′ 30.1″ N; 04° 54′ 39.9″ W | 36 | 19/05/2015 | MUB 57155 | ✓ | ✓ | | ✓ | | |
| AS011 | L06 | Carretera N-634, km 297, a 3 km de Riego desde Llanes (Llanes) | 43° 23′ 51.4″ N; 04° 42′ 58.3″ W | 32 | 19/05/2015 | MUB 57157 | ✓ | ✓ | | ✓ | | |
| AS012 | L06 | Carretera N-634, km 297, a 3 km de Riego desde Llanes (Llanes) | 43° 23′ 51.4″ N; 04° 42′ 58.3″ W | 32 | 19/05/2015 | MUB 57158 | ✓ | ✓ | | ✓ | | |
| AS013 | L06 | Carretera N-634, km 297, a 3 km de Riego desde Llanes (Llanes) | 43° 23′ 51.4″ N; 04° 42′ 58.3″ W | 33 | 19/05/2015 | MUB 57164 | ✓ | ✓ | | ✓ | | |
| AS014 | L07 | La Curciada, carretera AS-255 de Villaviciosa a Infiesto, cruce a Piñera, km 11 (Cabranes) | 43° 23′ 46.6″ N; 05° 25′ 08.1″ W | 329 | 20/05/2015 | MUB 57166 | ✓ | ✓ | | ✓ | | |
| AS015 | L08 | Carretera N-634, km 372, próximo a Nava (Nava) | 43° 21′ 20.6″ N; 05° 27′ 21.6″ W | 223 | 20/05/2015 | MUB 57168 | ✓ | | | ✓ | | |
| AS016 | L09 | Posada. Carretera AS-17 de Oviedo a Avilés, km 16 (Llanera) | 43° 26′ 17.0″ N; 05° 50′ 56.0″ W | 176 | 20/05/2015 | MUB 57171 | ✓ | | | ✓ | | |
| AS017 | L09 | Posada. Carretera AS-17 de Oviedo a Avilés, km 16 (Llanera) | 43° 26′ 17.0″ N; 05° 50′ 56.0″ W | 176 | 20/05/2015 | MUB 57172 | ✓ | | | ✓ | | |
| AS018 | L10 | Carretera AS-17 de Oviedo a Avilés, km 10 (Corvera de Asturias) | 43° 29′ 24.5″ N; 05° 50′ 51.7″ W | 97 | 20/05/2015 | MUB 57174 | ✓ | ✓ | | ✓ | | |
| AS019 | L10 | Carretera AS-17 de Oviedo a Avilés, km 10 (Corvera de Asturias) | 43° 29′ 24.5″ N; 05° 50′ 51.7″ W | 97 | 20/05/2015 | MUB 57175 | ✓ | ✓ | | ✓ | | |
| AS020 | L10 | Carretera AS-17 de Oviedo a Avilés, km 10 (Corvera de Asturias) | 43° 29′ 24.5″ N; 05° 50′ 51.7″ W | 97 | 20/05/2015 | MUB 57176 | ✓ | | | ✓ | | |
| AS021 | L10 | Carretera AS-17 de Oviedo a Avilés, km 10 (Corvera de Asturias) | 43° 29′ 24.5″ N; 05° 50′ 51.7″ W | 97 | 20/05/2015 | MUB 57177 | ✓ | ✓ | | ✓ | | |
| AS022 | L10 | Carretera AS-17 de Oviedo a Avilés, km 10 (Corvera de Asturias) | 43° 29′ 24.5″ N; 05° 50′ 51.7″ W | 97 | 20/05/2015 | MUB 57179 | ✓ | ✓ | | ✓ | | |
| AS023 | L11 | Carretera AS-236 de Grullos a Pravia, km 5, cruce San Tirso-Aces y Cornellana (Candamo) | 43° 26′ 58.1″ N; 06° 04′ 57.3″ W | 30 | 20/05/2015 | MUB 57185 | ✓ | | | ✓ | | |
| AS024 | L11 | Carretera AS-236 de Grullos a Pravia, km 5, cruce San Tirso-Aces y Cornellana (Candamo) | 43° 26′ 58.1″ N; 06° 04′ 57.3″ W | 30 | 20/05/2015 | MUB 57186 | ✓ | | | ✓ | ✓ | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AS025 | L12 | Carretera N-632, km 120 entre Piñera y Muros de Nalón (Pravia) | 43° 32′ 28.9″ N; 06° 07′ 37.3″ W | 155 | 21/05/2015 | MUB 57188 | ✓ | ✓ | | ✓ | | |
| AS026 | L12 | Carretera N-632, km 120 entre Piñera y Muros de Nalón (Pravia) | 43° 32′ 28.9″ N; 06° 07′ 37.3″ W | 155 | 21/05/2015 | MUB 57189 | ✓ | ✓ | | ✓ | | |
| AS027 | L12 | Carretera N-632, km 120 entre Piñera y Muros de Nalón (Pravia) | 43° 32′ 28.9″ N; 06° 07′ 37.3″ W | 155 | 21/05/2015 | MUB 57192 | ✓ | | | ✓ | | |
| AS028 | L13 | Carretera N-632 km 125 entre Las Dueñas y Artedo (Cudillero) | 43° 33′ 35.4″ N; 06° 10′ 29.6″ W | 140 | 21/05/2015 | MUB 57195 | ✓ | | | ✓ | | |
| AS029 | L13 | Carretera N-632 km 125 entre Las Dueñas y Artedo (Cudillero) | 43° 33′ 35.4″ N; 06° 10′ 29.6″ W | 140 | 21/05/2015 | MUB 57196 | ✓ | | | ✓ | | |
| AS030 | L13 | Carretera N-632 km 125 entre Las Dueñas y Artedo (Cudillero) | 43° 33′ 35.4″ N; 06° 10′ 29.6″ W | 140 | 21/05/2015 | MUB 57197 | ✓ | ✓ | | ✓ | ✓ | |
| AS031 | L13 | Carretera N-632 km 125 entre Las Dueñas y Artedo (Cudillero) | 43° 33′ 35.4″ N; 06° 10′ 29.6″ W | 140 | 21/05/2015 | MUB 57202 | ✓ | ✓ | | ✓ | | |
| AS032 | L14 | Carretera N-632a km 133 entre Soto de Luiña y Albuerne (Cudillero) | 43° 33′ 50.7″ N; 06° 14′ 22.0″ W | 76 | 21/05/2015 | MUB 57203 | ✓ | ✓ | | ✓ | | |
| AS033 | L14 | Carretera N-632a km 133 entre Soto de Luiña y Albuerne (Cudillero) | 43° 33′ 50.7″ N; 06° 14′ 22.0″ W | 76 | 21/05/2015 | MUB 57205 | ✓ | | | ✓ | | |
| AS034 | L15 | Carretera N-632a. Estación de FEVE de Novellana, km 140 (Cudillero) | 43° 33′ 27.0″ N; 06° 17′ 20.0″ W | 127 | 21/05/2015 | MUB 57206 | ✓ | ✓ | | ✓ | | |
| AS035 | L16 | Carretera N-632a km 149, Tablizo (Valdés) | 43° 32′ 45.7″ N; 06° 20′ 36.2″ W | 129 | 21/05/2015 | MUB 57207 | ✓ | | | ✓ | | |
| AS036 | L17 | Carretera N-640 km 10 de Vilavedelle a Vegadeo, Mirador sobre la ría de Ribadeo (Castropol) | 43° 29′ 05.9″ N; 07° 02′ 45.8″ W | 35 | 21/05/2015 | MUB 57215 | ✓ | ✓ | | ✓ | | |
| LE001 | L18 | El Ganso, Astorga. Camino de Santiago en dirección a Ponferrada, cerca de carretera LE-142 (Brazuelo) | 42° 27′ 47.0'' N; 06° 12′ 44.9'' W | 1015 | 05/06/2017 | MUB 57228 | ✓ | | | ✓ | | |
| LE002 | L18 | El Ganso, Astorga. Camino de Santiago en dirección a Ponferrada, cerca de carretera LE-142 (Brazuelo) | 42° 27′ 47.0'' N; 06° 12′ 44.9'' W | 1015 | 05/06/2017 | MUB 57230 | ✓ | | | ✓ | | |
| LE003 | L19 | Ponferrada, carretera LE-142, km 25 (Santa Colomba de Somoza) | 42° 29′ 11.9'' N; 06° 18′ 45.0'' W | 1277 | 06/06/2017 | MUB 57242 | ✓ | | | ✓ | | |
| LE004 | L20 | Ponferrada, carretera LE-142, km 29 (Santa Colomba de Somoza) | 42° 29' 30.8'' N; 06° 21' 12.1'' W | 1494 | 06/06/2017 | MUB 57254 | ✓ | | | ✓ | | |
| LE005 | L20 | Ponferrada, carretera LE-142, km 29 (Santa Colomba de Somoza) | 42° 29' 30.8'' N; 06° 21' 12.1'' W | 1494 | 06/06/2017 | MUB 57258 | ✓ | | | ✓ | | |
| LE006 | L21 | Ponferrada, carretera LE-142, km 37 (Molinaseca) | 42° 29' 30.4'' N: 06° 25' 54.6'' W | 1412 | 06/06/2017 | MUB 57261 | ✓ | | | ✓ | | |
| LE007 | L21 | Ponferrada, carretera LE-142, km 37 (Molinaseca) | 42° 29' 30.4'' N: 06° 25' 54.6'' W | 1412 | 06/06/2017 | MUB 57262 | ✓ | | | ✓ | | |
| LE008 | L21 | Ponferrada, carretera LE-142, km 37 (Molinaseca) | 42° 29' 30.4'' N: 06° 25' 54.6'' W | 1412 | 06/06/2017 | MUB 57268 | ✓ | | | ✓ | | |
| LE009 | L22 | El Acebo, carretera LE-142, km 40 (Molinaseca) | 42° 30' 05.8'' N; 06° 27' 34.8'' W | 1122 | 06/06/2017 | MUB 57271 | ✓ | | | ✓ | | |
| LE010 | L22 | El Acebo, carretera LE-142, km 40 (Molinaseca) | 42° 30' 05.8'' N; 06° 27' 34.8'' W | 1122 | 06/06/2017 | MUB 57273 | ✓ | | | ✓ | | |
| LE011 | L23 | Carretera LE-142, km 47 (Molinaseca) | 42° 31' 22.7'' N; 06° 30' 14.1'' W | 783 | 06/06/2017 | MUB 57275 | ✓ | | | ✓ | | |
| LE012 | L23 | Carretera LE-142, km 47 (Molinaseca) | 42° 31' 22.7'' N; 06° 30' 14.1'' W | 783 | 06/06/2017 | MUB 57276 | ✓ | | | ✓ | | |
| LE013 | L24 | Sales de los Barrios, Carretera LE-192/21, km 5 (Ponferrada) | 42° 30' 50.7'' N; 06° 33' 05.2'' W | 642 | 06/06/2017 | MUB 57277 | ✓ | | | ✓ | | |
| LE014 | L25 | Carretera LE-192/21, km 8 (Ponferrada) | 42° 30' 14.0'' N; 06° 31' 55.1'' W | 909 | 06/06/2017 | MUB 57278 | ✓ | | | ✓ | | |
| LE015 | L26 | Carretera LE-192/21, km ¿*? (Ponferrada) | 42° 25' 41.4'' N; 06° 31' 22.6'' W | 1507 | 06/06/2017 | MUB 57284 | ✓ | | | ✓ | | |
| LE016 | L26 | Carretera LE-192/21, km ¿*? (Ponferrada) | 42° 25' 41.4'' N; 06° 31' 22.6'' W | 1507 | 06/06/2017 | MUB 57286 | ✓ | | | ✓ | | |
| LE017 | L27 | Carretera LE-192/21, km 20 (Truchas) | 42° 14' 16.0'' N; 06° 21' 44.1'' W | 1072 | 06/06/2017 | MUB 57287 | ✓ | | | ✓ | | |
| LE018 | L27 | Carretera LE-192/21, km 20 (Truchas) | 42° 14' 16.0'' N; 06° 21' 44.1'' W | 1072 | 06/06/2017 | MUB 57289 | ✓ | | | ✓ | | |
| LE019 | L28 | Sueros de Cepeda, carretera LE-451, km 19 (Villamejil) | 42° 36' 52.9'' N; 06° 01' 59.7'' W | 948 | 07/06/2017 | MUB 57303 | ✓ | | | ✓ | | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LE020 | L28 | Sueros de Cepeda, carretera LE-451, km 19 (Villamejil) | 42° 36' 52.9'' N; 06° 01' 59.7'' W | 948 | 07/06/2017 | MUB 57306 | ✓ | | | ✓ | | |
| LE021 | L28 | Sueros de Cepeda, carretera LE-451, km 19 (Villamejil) | 42° 36' 52.9'' N; 06° 01' 59.7'' W | 948 | 07/06/2017 | MUB 57307 | ✓ | | | ✓ | | |
| LE022 | L28 | Sueros de Cepeda, carretera LE-451, km 19 (Villamejil) | 42° 36' 52.9'' N; 06° 01' 59.7'' W | 948 | 07/06/2017 | MUB 57310 | ✓ | | | ✓ | | |
| LE023 | L28 | Sueros de Cepeda, carretera LE-451, km 19 (Villamejil) | 42° 36' 52.9'' N; 06° 01' 59.7'' W | 948 | 07/06/2017 | MUB 57312 | ✓ | | | ✓ | | |
| LE024 | L29 | Carretera LE-451, km 29 (Quintana del Castillo) | 42° 40' 57.7'' N; 06° 00' 17.6'' W | 1162 | 07/06/2017 | MUB 57319 | ✓ | | | ✓ | | |
| LE025 | L30 | Carretera LE-451, km 41, cruce a La Utrera y Las Omañas (Valdesamario) | 42° 43' 01.8'' N; 05° 55' 38.7'' W | 984 | 07/06/2017 | MUB 57323 | ✓ | | | ✓ | | |
| LE026 | L30 | Carretera LE-451, km 41, cruce a La Utrera y Las Omañas (Valdesamario) | 42° 43' 01.8'' N; 05° 55' 38.7'' W | 984 | 07/06/2017 | MUB 57325 | ✓ | | | ✓ | | |
| LE027 | L30 | Carretera LE-451, km 41, cruce a La Utrera y Las Omañas (Valdesamario) | 42° 43' 01.8'' N; 05° 55' 38.7'' W | 984 | 07/06/2017 | MUB 57326 | ✓ | | | ✓ | | |
| LE028 | L30 | Carretera LE-451, km 41, cruce a La Utrera y Las Omañas (Valdesamario) | 42° 43' 01.8'' N; 05° 55' 38.7'' W | 984 | 07/06/2017 | MUB 57328 | ✓ | | | ✓ | ✓ | |
| LE029 | L30 | Carretera LE-451, km 41, cruce a La Utrera y Las Omañas (Valdesamario) | 42° 43' 01.8'' N; 05° 55' 38.7'' W | 984 | 07/06/2017 | MUB 57330 | ✓ | | | ✓ | | |
| LE030 | L30 | Carretera LE-451, km 41, cruce a La Utrera y Las Omañas (Valdesamario) | 42° 43' 01.8'' N; 05° 55' 38.7'' W | 984 | 07/06/2017 | MUB 57331 | ✓ | | | ✓ | ✓ | |
| LE031 | L31 | Carretera CL-626, km 69 (Los Barrios de Luna) | 42° 50' 25.6'' N; 05° 51' 50.4'' W | 1074 | 07/06/2017 | MUB 57337 | ✓ | | | ✓ | | |
| LE032 | L32 | Carretera N-120 de Astorga a León, km 343.7 (San Justo de la Vega) | 42° 26' 57.6'' N; 05° 58' 58.5'' W | 917 | 08/06/2017 | MUB 57340 | ✓ | | | ✓ | | |
| LE033 | L32 | Carretera N-120 de Astorga a León, km 343.7 (San Justo de la Vega) | 42° 26' 57.6'' N; 05° 58' 58.5'' W | 917 | 08/06/2017 | MUB 57341 | ✓ | | | ✓ | | |
| LE034 | L33 | Carretera N-120 de Astorga a León, km 330.6 (Santa Marina del Rey) | 42° 28' 48.6'' N; 05° 50' 10.6'' W | 845 | 08/06/2017 | MUB 57347 | ✓ | | | ✓ | | |
| LE035 | L33 | Carretera N-120 de Astorga a León, km 330.6 (Santa Marina del Rey) | 42° 28' 48.6'' N; 05° 50' 10.6'' W | 845 | 08/06/2017 | MUB 57348 | ✓ | | | ✓ | | |
| LE036 | L33 | Carretera N-120 de Astorga a León, km 330.6 (Santa Marina del Rey) | 42° 28' 48.6'' N; 05° 50' 10.6'' W | 845 | 08/06/2017 | MUB 57352 | ✓ | | | ✓ | | |
| LE037 | L34 | Carretera LE-315, km 5 (Vegacervera) | 42° 52' 53.3'' N; 05° 31' 34.7'' W | 1037 | 08/06/2017 | MUB 57373 | ✓ | | | ✓ | | |
| MU001 | L35 | Los Narejos, Avda. de la Isla del Barón (San Javier) | 37° 45' 26.8'' N; 00° 51' 01.4'' W | 135 | 29/03/2010 | MUB 43894 | ✓ | | | ✓ | | |
| MU002 | L36 | Atamaría, Casa de las Cenizas, Rambla de la Carrasquilla (La Unión) | 37° 35' 42.3'' N; 00° 49' 25.8'' W | 150 | 07/04/2010 | MUB 43895 | ✓ | ✓ | | ✓ | | |
| MU003 | L37 | Sierra Minera, El Gorguel, Pantano Galo, ladera E (La Unión) | 37° 35' 30.8'' N; 00° 52' 45.9'' W | 114 | 07/04/2010 | MUB 43903 | ✓ | ✓ | | ✓ | | |
| MU004 | L38 | Alhama de Murcia, polígono industrial (Alhama de Murcia) | 37° 49' 20.9'' N; 01° 24' 08.8'' W | 166 | 09/05/2013 | MUB 43923 | ✓ | | | ✓ | | |
| MU005 | L39 | Carretera del canal del margen derecho del Trasvase Tajo-Segura entre Totana y Alhama (Alhama de Murcia) | 37° 49' 29.0'' N; 01° 29' 21.8'' W | 346 | 09/05/2013 | MUB 43932 | ✓ | ✓ | | ✓ | | |
| MU006 | L39 | Carretera del canal del margen derecho del Trasvase Tajo-Segura entre Totana y Alhama (Alhama de Murcia) | 37° 49' 29.0'' N; 01° 29' 21.8'' W | 346 | 09/05/2013 | MUB 43934 | ✓ | | | ✓ | | |
| MU007 | L39 | Carretera del canal del margen derecho del Trasvase Tajo-Segura entre Totana y Alhama (Alhama de Murcia) | 37° 49' 29.0'' N; 01° 29' 21.8'' W | 346 | 09/05/2013 | MUB 43936 | ✓ | | | ✓ | | |
| MU008 | L40 | Carretera de servicio del trasvase Tajo-Segura entre Totana y Alhama (Alhama de Murcia) | 37° 49′ 45.1″ N; 01° 28′ 58.0″ W | 350 | 06/04/2015 | MUB 55904 | ✓ | ✓ | | ✓ | | |
| MU009 | L41 | Carretera de servicio del trasvase Tajo-Segura entre Totana y Alhama (Totana) | 37° 48′ 42.3″ N; 01° 29′ 49.2″ W | 350 | 06/04/2015 | MUB 55896 | ✓ | | | ✓ | | |
| MU010 | L42 | Carretera de servicio del trasvase Tajo-Segura entre Totana y Alhama (Totana) | 37° 48′ 30.1″ N; 01° 29′ 18.1″ W | 350 | 06/04/2015 | MUB 55899 | ✓ | ✓ | | ✓ | ✓ | |
| MU011 | L43 | Camino de San Juan, en paralelo a la N-344 (Las Torres de Cotillas) | 38° 02′ 19.6″ N; 01° 13′ 59.6″ W | 72 | 28/05/2015 | MUB 55906 | ✓ | | | ✓ | | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MU012 | L44 | Polígono industrial La Serreta, calle Uruguay (Molina del Segura) | 38° 04' 42.8" N; 01° 13' 34.7" W | 98 | 28/05/2015 | MUB 55908 | ✓ | | | ✓ | | |
| MU013 | L45 | Carretera RM-C1, de Alcantarilla a Barqueros, (Murcia) | 37° 57' 38.0'' N; 01° 19' 11.6'' W | 270 | 11/04/2017 | MUB 55836 | ✓ | | | ✓ | | |
| MU014 | L45 | Carretera RM-C1, de Alcantarilla a Barqueros, (Murcia) | 37° 57' 38.0'' N; 01° 19' 11.6'' W | 270 | 11/04/2017 | MUB 55837 | ✓ | | | ✓ | | |
| MU015 | L45 | Carretera RM-C1, de Alcantarilla a Barqueros, (Murcia) | 37° 57' 38.0'' N; 01° 19' 11.6'' W | 270 | 11/04/2017 | MUB 55838 | ✓ | | | ✓ | | |
| MU016 | L45 | Carretera RM-C1, de Alcantarilla a Barqueros, (Murcia) | 37° 57' 38.0'' N; 01° 19' 11.6'' W | 270 | 11/04/2017 | MUB 55839 | ✓ | | | ✓ | | |
| MU017 | L45 | Carretera RM-C1, de Alcantarilla a Barqueros, (Murcia) | 37° 57' 38.0'' N; 01° 19' 11.6'' W | 270 | 11/04/2017 | MUB 55842 | ✓ | | | ✓ | | |
| MU018 | L45 | Carretera RM-C1, de Alcantarilla a Barqueros, (Murcia) | 37° 57' 38.0'' N; 01° 19' 11.6'' W | 270 | 11/04/2017 | MUB 55843 | ✓ | | | ✓ | | |
| MU019 | L46 | Camino de San Juan de Loma a San Pedro (Las Torres de Cotillas) | 38° 01' 58.4'' N; 01° 13' 57.4'' W | 68 | 18/04/2017 | MUB 55874 | ✓ | | | ✓ | | |
| MU020 | L47 | Carretera N-344, km 44, cruce con carretera RM-A10, entre Estación de Blanca y Jumilla (Abarán) | 38° 17' 51.9'' N; 01° 18' 41.6'' W | 311 | 18/04/2017 | MUB 55876 | ✓ | | | ✓ | | |
| MU021 | L48 | Carretera N-344, km 49, entre Estación de Blanca y Jumilla (Jumilla) | 38° 20' 18.3'' N; 01° 16' 53.6'' W | 333 | 18/04/2017 | MUB 55881 | ✓ | | | ✓ | | |
| MU022 | L49 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 51' 07.4'' N; 01° 28' 12.2'' W | 350 | 12/04/2017 | MUB 55844 | ✓ | | | ✓ | | |
| MU023 | L50 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 49' 57.5'' N; 01° 28' 45.4'' W | 350 | 12/04/2017 | MUB 55851 | ✓ | | | ✓ | | |
| MU024 | L50 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 49' 57.5'' N; 01° 28' 45.4'' W | 350 | 12/04/2017 | MUB 55852 | ✓ | | | ✓ | | |
| MU025 | L51 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 49' 34.5'' N; 01° 29' 07.3'' W | 350 | 12/04/2017 | MUB 55853 | ✓ | | | ✓ | | |
| MU026 | L52 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 49' 24.9'' N; 01° 29' 23.5'' W | 350 | 12/04/2017 | MUB 55861 | ✓ | | | ✓ | | |
| MU027 | L52 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 49' 24.9'' N; 01° 29' 23.5'' W | 350 | 12/04/2017 | MUB 55862 | ✓ | | | ✓ | | |
| MU028 | L53 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 48' 36.5'' N; 01° 29' 50.2'' W | 350 | 12/04/2017 | MUB 55864 | ✓ | | | ✓ | | |
| MU029 | L53 | Camino de servicio canal de Trasvase entre Alhama-Totana (Alhama de Murcia) | 37° 48' 36.5'' N; 01° 29' 50.2'' W | 350 | 12/04/2017 | MUB 55865 | ✓ | | | ✓ | | |
| MU030 | L54 | Finca Lo Poyo en la carretera RM-F54, entre Los Nietos y Los Urrutias (Cartagena) | 37° 39' 22.9'' N; 00° 48' 28.3'' W | 2 | 19/04/2017 | MUB 55884 | ✓ | | | ✓ | | |
| MU031 | L54 | Finca Lo Poyo en la carretera RM-F54, entre Los Nietos y Los Urrutias (Cartagena) | 37° 39' 22.9'' N; 00° 48' 28.3'' W | 2 | 19/04/2017 | MUB 55887 | ✓ | | | ✓ | | |
| MU032 | L54 | Finca Lo Poyo en la carretera RM-F54, entre Los Nietos y Los Urrutias (Cartagena) | 37° 39' 22.9'' N; 00° 48' 28.3'' W | 2 | 19/04/2017 | MUB 55888 | ✓ | | | ✓ | | |
| MU033 | L54 | Finca Lo Poyo en la carretera RM-F54, entre Los Nietos y Los Urrutias (Cartagena) | 37° 39' 22.9'' N; 00° 48' 28.3'' W | 2 | 19/04/2017 | MUB 55889 | ✓ | | | ✓ | | |
| MU034 | L55 | Carretera del Chaparral, entre Bullas y Cehegín, cerca del cruce con RM-15 (Bullas) | 38° 03' 09.7'' N; 01° 40' 53.1'' W | 600 | 16/04/2017 | MUB 55869 | ✓ | | | ✓ | ✓ | |
| MU035 | L55 | Carretera del Chaparral, entre Bullas y Cehegín, cerca del cruce con RM-15 (Bullas) | 38° 03' 09.7'' N; 01° 40' 53.1'' W | 600 | 16/04/2017 | MUB 55870 | ✓ | | | ✓ | | |
| MU036 | **L44** | Polígono industrial La Serreta, calle Uruguay (Molina del Segura) | 38° 04' 42.8" N; 01° 13' 34.7" W | 98 | 02/05/2017 | MUB 55892 | ✓ | | | ✓ | | |
| MU037 | L56 | Avda. del Mediterráneo, en puente sobre carretera N-344, salida Medialegua en dirección La Loma (Las Torres de Cotillas) | 38° 01' 02.8'' N; 01° 13' 35.5'' W | 81 | 11/05/2017 | MUB 55893 | ✓ | | | ✓ | | |
| SN001 | L57 | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51' 56.5" N; 03° 28' 59.6" W | 287 | 05/02/2010 | MUB 44008 | ✓ | ✓ | ✓ | ✓ | | ✓ |
| SN002 | L57 | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51' 56.5" N; 03° 28' 59.6" W | 287 | 05/02/2010 | MUB 44022 | ✓ | ✓ | ✓ | ✓ | | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SN003 | L58 | ALMERÍA, Sierra Nevada, cercanías al Puerto de La Ragua (cara Sur), Pilar de La Yegua (Bayárcal) | 37° 06' 12.0'' N; 03° 01' 40.0'' W | 2048 | 03/08/2015 | MUB 55937 | ✓ | | ✓ | ✓ | | ✓ |
| SN004 | L58 | ALMERÍA, Sierra Nevada, cercanías al Puerto de La Ragua (cara Sur), Pilar de La Yegua (Bayárcal) | 37° 06' 12.0'' N; 03° 01' 40.0'' W | 2048 | 03/08/2015 | MUB 55940 | ✓ | | | ✓ | ✓ | ✓ |
| SN005 | L58 | ALMERÍA, Sierra Nevada, cercanías al Puerto de La Ragua (cara Sur), Pilar de La Yegua (Bayárcal) | 37° 06' 12.0'' N; 03° 01' 40.0'' W | 2048 | 03/08/2015 | MUB 55941 | ✓ | | | ✓ | | |
| SN006 | L58 | ALMERÍA, Sierra Nevada, cercanías al Puerto de La Ragua (cara Sur), Pilar de La Yegua (Bayárcal) | 37° 06' 12.0'' N; 03° 01' 40.0'' W | 2048 | 03/08/2015 | MUB 55942 | ✓ | | | ✓ | ✓ | |
| SN007 | L58 | ALMERÍA, Sierra Nevada, cercanías al Puerto de La Ragua (cara Sur), Pilar de La Yegua (Bayárcal) | 37° 06' 12.0'' N; 03° 01' 40.0'' W | 2048 | 03/08/2015 | MUB 55943 | ✓ | | ✓ | ✓ | ✓ | ✓ |
| SN008 | L59 | GRANADA, Sierra Nevada, alrededores de la Casa de Prado Grande (Capileira) | 36° 57' 56.0'' N; 03° 20' 24.0'' W | 1965 | 19/08/2015 | MUB 55944 | ✓ | | ✓ | ✓ | | |
| SN009 | L59 | GRANADA, Sierra Nevada, alrededores de la Casa de Prado Grande (Capileira) | 36° 57' 56.0'' N; 03° 20' 24.0'' W | 1965 | 19/08/2015 | MUB 55945 | ✓ | ✓ | ✓ | ✓ | | |
| SN010 | L59 | GRANADA, Sierra Nevada, alrededores de la Casa de Prado Grande (Capileira) | 36° 57' 56.0'' N; 03° 20' 24.0'' W | 1965 | 19/08/2015 | MUB 55946 | ✓ | | ✓ | ✓ | ✓ | ✓ |
| SN011 | L59 | GRANADA, Sierra Nevada, alrededores de la Casa de Prado Grande (Capileira) | 36° 57' 56.0'' N; 03° 20' 24.0'' W | 1965 | 19/08/2015 | MUB 55947 | ✓ | | | ✓ | ✓ | ✓ |
| SN012 | L60 | GRANADA, carretera A-346, km 7, hacia Órgiva, apartadero de coches (Órgiva) | 36° 52′ 19.8″ N; 03° 28′ 24.3″ W | 390 | 22/04/2016 | MUB 55949 | ✓ | ✓ | ✓ | ✓ | | |
| SN013 | L61 | GRANADA, carretera A-346, km 0, cruce a Órgiva, apartadero de coches borde carretera (Órgiva) | 36° 53′ 00.4″ N; 03° 24′ 58.5″ W | 320 | 22/04/2016 | MUB 55950 | ✓ | ✓ | ✓ | ✓ | | |
| SN014 | L62 | GRANADA, carretera GR-3204 de Melepís a Chite, km 1, Valle de Lecún (Lecrín) | 36° 56′ 27.4″ N; 03° 33′ 24.2″ W | 687 | 22/04/2016 | MUB 55951 | ✓ | | | ✓ | | |
| SN015 | L62 | GRANADA, carretera GR-3204 de Melepís a Chite, km 1, Valle de Lecún (Lecrín) | 36° 56′ 27.4″ N; 03° 33′ 24.2″ W | 687 | 22/04/2016 | MUB 55952 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| SN016 | L63 | GRANADA, carretera A-4002, km 7, Puerto Lobo desvío a Cortijo Buenos Aires (antigua carretera Granada-Murcia) (Víznar) | 37° 13′ 08.6″ N; 03° 32′ 30.7″ W | 1086 | 23/04/2016 | MUB 55956 | ✓ | ✓ | | ✓ | | |
| SN017 | L63 | GRANADA, carretera A-4002, km 7, Puerto Lobo desvío a Cortijo Buenos Aires (antigua carretera Granada-Murcia) (Víznar) | 37° 13′ 08.6″ N; 03° 32′ 30.7″ W | 1086 | 23/04/2016 | MUB 55960 | ✓ | | ✓ | ✓ | | |
| SN018 | L64 | GRANADA, Sierra de Huétor, barranco de Víznar, carretera GR-3101, km 5 (Víznar) | 37° 14′ 21.1″ N; 03° 32′ 40.2″ W | 1102 | 23/04/2016 | MUB 55963 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| SN019 | L65 | GRANADA, Sierra de Huétor, carretera GR-3103, km 7 (Alcafar) | 37° 15′ 10.1″ N; 03° 33′ 37.3″ W | 1170 | 23/04/2016 | MUB 55967 | ✓ | ✓ | ✓ | ✓ | | |
| SN020 | L66 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 46.0″ N; 03° 23′ 38.0″ W | 2325 | 25/06/2016 | MUB 55969 | ✓ | | ✓ | ✓ | | ✓ |
| SN021 | L66 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 46.0″ N; 03° 23′ 38.0″ W | 2325 | 25/06/2016 | MUB 55970 | ✓ | ✓ | ✓ | ✓ | | ✓ |
| SN022 | L66 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 46.0″ N; 03° 23′ 38.0″ W | 2325 | 25/06/2016 | MUB 55972 | ✓ | ✓ | | ✓ | ✓ | ✓ |
| SN023 | L66 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 46.0″ N; 03° 23′ 38.0″ W | 2325 | 25/06/2016 | MUB 55973 | ✓ | ✓ | ✓ | ✓ | | ✓ |
| SN024 | L66 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 46.0″ N; 03° 23′ 38.0″ W | 2325 | 25/06/2016 | MUB 55974 | ✓ | ✓ | | ✓ | | |
| SN025 | L67 | GRANADA, Sierra Nevada, Estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 50.0″ N; 03° 23′ 50.0″ W | 2275 | 25/06/2016 | MUB 55976 | ✓ | ✓ | | ✓ | | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SN026 | L67 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 50.0″ N; 03° 23′ 50.0″ W | 2275 | 25/06/2016 | MUB 55977 | ✓ | ✓ | ✓ | ✓ | | |
| SN027 | L68 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 47.0″ N; 03° 23′ 51.0″ W | 2240 | 25/06/2016 | MUB 55979 | ✓ | | ✓ | ✓ | ✓ | |
| SN028 | L68 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 47.0″ N; 03° 23′ 51.0″ W | 2240 | 25/06/2016 | MUB 55980 | ✓ | | ✓ | ✓ | ✓ | ✓ |
| SN029 | L68 | GRANADA, Sierra Nevada, estación de Esquí alrededores de Prado Llano (Monachil) | 37° 05′ 47.0″ N; 03° 23′ 51.0″ W | 2240 | 25/06/2016 | MUB 55982 | ✓ | ✓ | | ✓ | | ✓ |
| SN030 | L69 | GRANADA, Sierra Nevada, barranco Juan de La Villa (Nigüelas) | 36° 59' 14.0'' N; 03° 29' 15.0'' W | 1972 | 16/07/2016 | MUB 55984 | ✓ | | ✓ | ✓ | | |
| SN031 | L69 | GRANADA, Sierra Nevada, barranco Juan de La Villa (Nigüelas) | 36° 59' 14.0'' N; 03° 29' 15.0'' W | 1972 | 16/07/2016 | MUB 55985 | ✓ | | | ✓ | ✓ | ✓ |
| SN032 | L69 | GRANADA, Sierra Nevada, barranco Juan de La Villa (Nigüelas) | 36° 59' 14.0'' N; 03° 29' 15.0'' W | 1972 | 16/07/2016 | MUB 55986 | ✓ | | ✓ | ✓ | ✓ | ✓ |
| SN033 | L70 | GRANADA, Sierra Nevada, Laguna de la estación de esquí de Borreguiles (Monachil) | 37° 04' 22.0'' N; 03° 23' 28.0'' W | 2622 | 16/07/2016 | MUB 55992 | ✓ | | ✓ | ✓ | | |
| SN034 | L70 | GRANADA, Sierra Nevada Laguna de la estación de esquí de Borreguiles (Monachil) | 37° 04' 22.0'' N; 03° 23' 28.0'' W | 2622 | 16/07/2016 | MUB 55994 | ✓ | | | ✓ | ✓ | ✓ |
| SN035 | L70 | GRANADA, Sierra Nevada, Laguna de la estación de esquí de Borreguiles (Monachil) | 37° 04' 22.0'' N; 03° 23' 28.0'' W | 2622 | 16/07/2016 | MUB 55995 | ✓ | | ✓ | ✓ | | ✓ |
| SN036 | L70 | GRANADA, Sierra Nevada Laguna de la estación de esquí de Borreguiles (Monachil) | 37° 04' 22.0'' N; 03° 23' 28.0'' W | 2622 | 16/07/2016 | MUB 55996 | ✓ | | ✓ | ✓ | ✓ | ✓ |
| SN037 | L71 | ALMERÍA, Sierra Nevada, Casa de Las Minillas (Láujar de Andarax) | 37° 04' 10.0'' N; 02° 54' 23.0'' W | 2152 | 28/07/2016 | MUB 55997 | ✓ | | ✓ | ✓ | | |
| SN038 | L72 | ALMERÍA, Sierra Nevada, Minas del Collado de Gabiarra (Láujar de Andarax) | 37° 04' 27.0'' N; 02° 55' 24.0'' W | 2124 | 28/07/2016 | MUB 56000 | ✓ | | | ✓ | | ✓ |
| SN039 | L72 | ALMERÍA, Sierra Nevada, Minas del Collado de Gabiarra (Láujar de Andarax) | 37° 04' 27.0'' N; 02° 55' 24.0'' W | 2124 | 28/07/2016 | MUB 56002 | ✓ | | | ✓ | ✓ | ✓ |
| SN040 | L73 | GRANADA, carretera N-340, km 334, de Almería a Motril (Motril) | 36° 43' 47.4'' N; 03° 30' 30.8'' W | 58 | 29/04/2017 | MUB 56003 | ✓ | | | ✓ | | |
| SN041 | L73 | GRANADA, carretera N-340, km 334, de Almería a Motril (Motril) | 36° 43' 47.4'' N; 03° 30' 30.8'' W | 58 | 29/04/2017 | MUB 56005 | ✓ | | | ✓ | | |
| SN042 | L73 | GRANADA, carretera N-340, km 334, de Almería a Motril (Motril) | 36° 43' 47.4'' N; 03° 30' 30.8'' W | 58 | 29/04/2017 | MUB 56006 | ✓ | | | ✓ | | |
| SN043 | **L57** | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51′ 56.5″ N; 03° 28′ 59.6″ W | 287 | 29/04/2017 | MUB 56010 | ✓ | | ✓ | ✓ | | |
| SN044 | **L57** | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51′ 56.5″ N; 03° 28′ 59.6″ W | 287 | 29/04/2017 | MUB 56011 | ✓ | | ✓ | ✓ | | |
| SN045 | **L57** | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51′ 56.5″ N; 03° 28′ 59.6″ W | 287 | 29/04/2017 | MUB 56012 | ✓ | | ✓ | ✓ | | ✓ |
| SN046 | **L57** | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51′ 56.5″ N; 03° 28′ 59.6″ W | 287 | 29/04/2017 | MUB 56015 | ✓ | | ✓ | ✓ | | |
| SN047 | **L57** | GRANADA, carretera A-346, km 9, dirección a Órgiva (Órgiva) | 36° 51′ 56.5″ N; 03° 28′ 59.6″ W | 287 | 29/04/2017 | MUB 56017 | ✓ | | ✓ | ✓ | | |
| SN048 | **L60** | GRANADA, carretera A-346, km 7, hacia Órgiva, apartadero de coches (Órgiva) | 36° 52′ 19.8″ N; 03° 28′ 24.3″ W | 390 | 29/04/2017 | MUB 56020 | ✓ | | ✓ | ✓ | | |
| SN049 | L74 | GRANADA, Sierra Nevada, carretera A-4132, km 7, próximo a Soportújar (Carataunas) | 36° 55' 29.4'' N; 03° 24' 01.3'' W | 851 | 29/04/2017 | MUB 56024 | ✓ | | | ✓ | | |
| SN050 | L75 | GRANADA, Sierra Nevada, afueras de Lanjarón, carretera A-348 dirección a Órgiva (Lanjarón) | 36° 55' 01.4'' N; 03° 28' 17.0'' W | 657 | 29/04/2017 | MUB 56027 | ✓ | | | ✓ | | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SN051 | L75 | GRANADA, Sierra Nevada, afueras de Lanjarón, carretera A-348 dirección a Órgiva (Lanjarón) | 36° 55' 01.4'' N; 03° 28' 17.0'' W | 657 | 29/04/2017 | MUB 56028 | ✓ | | | ✓ | ✓ | ✓ |
| SN052 | L76 | GRANADA, Sierra Nevada, Lanjarón, centro ciudad (Lanjarón) | 36° 55' 01.6'' N; 03° 28' 47.9'' W | 630 | 29/04/2017 | MUB 56030 | ✓ | | ✓ | ✓ | | |
| SN053 | L77 | GRANADA, Sierra Nevada, alrededores de Cenes de la Vega, camino paralelo a carretera (Cenes de la Vega) | 37° 09' 12.6'' N; 03° 32' 01.7'' W | 762 | 30/04/2017 | MUB 56031 | ✓ | | ✓ | ✓ | | |
| SN054 | L77 | GRANADA, Sierra Nevada, alrededores de Cenes de la Vega, camino paralelo a carretera (Cenes de la Vega) | 37° 09' 12.6'' N; 03° 32' 01.7'' W | 762 | 30/04/2017 | MUB 56033 | ✓ | | | ✓ | | ✓ |
| SN055 | L78 | GRANADA, Sierra Nevada, alrededores de Cenes de la Vega, frente a gasolinera Galp a orillas del río (Cenes de la Vega) | 37° 09' 17.6'' N; 03° 32' 46.2'' W | 755 | 30/04/2017 | MUB 56034 | ✓ | | | ✓ | | |
| SN056 | L78 | GRANADA, Sierra Nevada, alrededores de Cenes de la Vega, frente a gasolinera Galp a orillas del río (Cenes de la Vega) | 37° 09' 17.6'' N; 03° 32' 46.2'' W | 755 | 30/04/2017 | MUB 56035 | ✓ | | ✓ | ✓ | | |
| SN057 | L79 | GRANADA, Sierra Nevada, carretera GR-420, km 1, de Cenes de la Vega a Pinos Genil (Pinos Genil) | 37° 09' 41.9'' N; 03° 31' 27.2'' W | 757 | 30/04/2017 | MUB 56037 | ✓ | | ✓ | ✓ | ✓ | |
| SN058 | L79 | GRANADA, Sierra Nevada, carretera GR-420, km 1, de Cenes de la Vega a Pinos Genil (Pinos Genil) | 37° 09' 41.9'' N; 03° 31' 27.2'' W | 757 | 30/04/2017 | MUB 56038 | ✓ | | | ✓ | | |
| SN059 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56040 | ✓ | | ✓ | ✓ | | |
| SN060 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56041 | ✓ | | ✓ | ✓ | | ✓ |
| SN061 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56042 | ✓ | | ✓ | ✓ | | |
| SN062 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56043 | ✓ | | ✓ | ✓ | | |
| SN063 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56045 | ✓ | | ✓ | ✓ | | |
| SN064 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56046 | ✓ | | ✓ | ✓ | ✓ | |
| SN065 | L80 | GRANADA, Sierra Nevada, carretera A-395, km 15, proximidades del restaurante "La Higuera" (Güejar-Sierra) | 37° 08' 34.1'' N; 03° 29' 14.5'' W | 1295 | 30/04/2017 | MUB 56048 | ✓ | | ✓ | ✓ | | ✓ |
| SN066 | L81 | GRANADA, Sierra Nevada, carretera A-395, km 16, proximidades de los restaurantes "Los Jamones", "El Mirador de Güejar-Sierra" y "El Desvío" (Güejar-Sierra) | 37° 07' 58.0'' N; 03° 26' 20.6'' W | 1650 | 30/04/2017 | MUB 56051 | ✓ | | ✓ | ✓ | | |
| SN067 | L81 | GRANADA, Sierra Nevada, carretera A-395, km 16, proximidades de los restaurantes "Los Jamones", "El Mirador de Güejar-Sierra" y "El Desvío" (Güejar-Sierra) | 37° 07' 58.0'' N; 03° 26' 20.6'' W | 1650 | 30/04/2017 | MUB 56052 | ✓ | | ✓ | ✓ | | ✓ |
| SN068 | L81 | GRANADA, Sierra Nevada, carretera A-395, km 16, proximidades de los restaurantes "Los Jamones", "El Mirador de Güejar-Sierra" y "El Desvío" (Güejar-Sierra) | 37° 07' 58.0'' N; 03° 26' 20.6'' W | 1650 | 30/04/2017 | MUB 56056 | ✓ | | | ✓ | | |

**Appendix 1.** Continued.

| Code | Location number | Location name | Geographical coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium code | *In vitro* culture | Ploidy analysis | DNA Sanger sequencing | GBS | RNA analysis | Biometric study |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SN069 | L81 | GRANADA, Sierra Nevada, carretera A-395, km 16, proximidades de los restaurantes "Los Jamones", "El Mirador de Güejar-Sierra" y "El Desvío" (Güejar-Sierra) | 37° 07' 58.0'' N; 03° 26' 20.6'' W | 1650 | 30/04/2017 | MUB 56059 | ✓ | | ✓ | ✓ | ✓ | ✓ |
| SN070 | L81 | GRANADA, Sierra Nevada, carretera A-395, km 16, proximidades de los restaurantes "Los Jamones", "El Mirador de Güejar-Sierra" y "El Desvío" (Güejar-Sierra) | 37° 07' 58.0'' N; 03° 26' 20.6'' W | 1650 | 30/04/2017 | MUB 56060 | ✓ | | | ✓ | | |
| SN071 | L81 | GRANADA, Sierra Nevada, carretera A-395, km 16, proximidades de los restaurantes "Los Jamones", "El Mirador de Güejar-Sierra" y "El Desvío" (Güejar-Sierra) | 37° 07' 58.0'' N; 03° 26' 20.6'' W | 1650 | 30/04/2017 | MUB 56061 | ✓ | | | ✓ | | ✓ |
| SN072 | L82 | GRANADA, Sierra Nevada, carretera A-395, Centro de Interpretación de la Naturaleza "El Dornajo" (Güejar-Sierra) | 37° 07' 57.9'' N; 03° 26' 06.2'' W | 1667 | 15/06/2017 | MUB 56062 | ✓ | | ✓ | ✓ | | |
| SN073 | L82 | GRANADA, Sierra Nevada, carretera A-395, Centro de Interpretación de la Naturaleza "El Dornajo" (Güejar-Sierra) | 37° 07' 57.9'' N; 03° 26' 06.2'' W | 1667 | 15/06/2017 | MUB 56065 | ✓ | | ✓ | ✓ | ✓ | ✓ |

SPECIMENS STUDIED FROM OUTSIDE THE IBERIAN PENINSULA

| Code | Location name | Geographic coordinates | Altitude (m a.s.l.) | Collection date (dd/mm/aaaa) | Herbarium data | Collector data | *In vitro* culture | DNA Sanger sequencing | RNA analysis | Biometric study | GenBank accession number (ITS) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Funaria aequidens* | | | | | | | | | | | |
| Ru2 | RUSSIA: Caucasus, Kabardino-Balkaria | | | 25/08/2005 | MW 9040124 | *Ignatov & Ignatova #05-1766* | ✓ | ✓ | ✓ | ✓ | KP342457 |
| *Funaria arctica* | | | | | | | | | | | |
| Ru4 | RUSSIA: Taimyr Autonomous Distr., south part, Krasnoyarsk Territory, Taimyr Distr. | 71° 15' 00.0'' N; 105° 37' 00.0'' E | 201 | 02/08/2013 | MHA 9033205 | *Fedosov #13-3-0726* | ✓ | | ✓ | ✓ | |
| *Funaria flavicans* | | | | | | | | | | | |
| Am1 | USA: Orange county. Durham. Duke University Campus | | | 12/03/2006 | CONN | *Bernard Goffinet 9345* | | ✓ | | ✓ | JN089173 |
| *Funaria hygrometrica* | | | | | | | | | | | |
| Am2 | CHILE: BIO-BIO: Prov. of Bio-Bio. E of Los Angeles, Parque Nacional Laguna de la Laja | 37° 23' 00.0'' S; 71° 23' 50.0'' W | 1200 | 24/10/1998 | CONN | *Bernard Goffinet 5576* | | ✓ | | ✓ | JN089174 |
| *Funaria microstoma* | | | | | | | | | | | |
| | AUSTRALIA: Western Australia. Coastal Highway, just N of Leeman township. | | | 18/09/2002 | CONN | *Beever 95-09* | | ✓ | | | JN089175 |
| *Funaria polaris* | | | | | | | | | | | |
| Ru1 | RUSSIA: Krasnoyarsk Territory, Taimyr Distr. | 71° 13' 00.0'' N; 105° 37' 00.0'' E | 267 | 30/07/2013 | MW 9040195 | *Fedosov #13-3-0670* | ✓ | ✓ | ✓ | ✓ | KP342456 |
| Ru3 | RUSSIA: Krasnoyarsk Territory, Taimyr Distr. | 71° 13' 00.0'' N; 105° 37' 00.0'' E | 267 | 30/07/2013 | MHA 9033366 | *Fedosov #13-3-0668* | ✓ | | ✓ | ✓ | |

**Appendix 2.** Voucher information for the specimens studied in the transcriptome analysis. Specimens collected from the Iberian Peninsula are ordered according to their belonging to the four geographical areas studied, identified by acronyms and sequencing code (AS: Asturias province; LE: Leon province; MU: Murcia province; SN: Sierra Nevada Mountains). Specimens studied from outside the Iberian Peninsula are ordered according to their belonging to species within *Funaria* genus. Also the next information is given: *Funaria* species, code, code in RNA analysis, environmental conditions, raw reads, clean reads, mapped reads (to *de novo* transcriptome) and transcriptomic group (each one highlighted in a different colour). Specimens marked with (*) are technical replicates.

| *Funaria* species | Code | Code in RNA analysis | Environmental conditions | Raw reads | Clean reads | Mapped reads (to *de novo* transcriptome) | Transcriptomic group |
|---|---|---|---|---|---|---|---|
| *F. hygrometrica* | AS024 | AS024C | CONTROL | 34322748 | 33975220 | 27128771 (79.85%) | AC |
| | | AS024T | THERMAL STRESS | 37145313 | 36813917 | 31518956 (85.62%) | AT |
| | AS030 | AS030C | CONTROL | 39919119 | 39566984 | 30380757 (76.78%) | BC |
| | | AS030T | THERMAL STRESS | 35754619 | 35398940 | 29449421 (83.19%) | BT |
| | LE028 | LE028C | CONTROL | 33605661 | 33325398 | 30020690 (90.08%) | BC |
| | | LE028T | THERMAL STRESS | 36723071 | 36298418 | 30222255 (83.26%) | BT |
| | LE030 | LE030C | CONTROL | 35441669 | 34969770 | 31277755 (89.44%) | AC |
| | | LE030T | THERMAL STRESS | 41290913 | 40936832 | 36157908 (88.33%) | AT |
| | LE030 (*) | LE030CR | CONTROL | 35733703 | 35398450 | 31676221 (89.48%) | AC |
| | | LE030TR | THERMAL STRESS | 31148454 | 30811786 | 27035485 (87.74%) | AT |
| | MU010 | MU010C | CONTROL | 35072128 | 34727513 | 30192797 (86.94%) | AC |
| | | MU010T | THERMAL STRESS | 33409044 | 33029245 | 28238553 (85.50%) | AT |
| | MU034 | MU034C | CONTROL | 32802903 | 32312574 | 28658674 (88.69%) | BC |
| | | MU034T | THERMAL STRESS | 33934623 | 33285319 | 26913724 (80.86%) | BT |
| | SN004 | SN004C | CONTROL | 31719170 | 31281924 | 27790974 (88.84%) | AC |
| | | SN004T | THERMAL STRESS | 31983838 | 31528554 | 27249520 (86.43%) | AT |
| | SN006 | SN006C | CONTROL | 30244914 | 29973476 | 27106309 (90.43%) | BC |
| | | SN006T | THERMAL STRESS | 34218773 | 33467982 | 27262960 (81.46%) | BT |
| | SN006 (*) | SN006CR | CONTROL | 32233208 | 31967981 | 28691277 (89.75%) | BC |
| | | SN006TR | THERMAL STRESS | 32628998 | 31864658 | 26053877 (81.76%) | BT |
| | SN007 | SN007C | CONTROL | 30687088 | 30249839 | 20198205 (66.77%) | AC |
| | | SN007T | THERMAL STRESS | 29844019 | 29513377 | 25942933 (87.90%) | AT |
| | SN010 | SN010C | CONTROL | 37486884 | 37055680 | 31648057 (85.41%) | AC |
| | | SN010T | THERMAL STRESS | 35287375 | 34879129 | 30500604 (87.45%) | AT |
| | SN011 | SN011C | CONTROL | 32040262 | 31386444 | 28063278 (89.41%) | AC |
| | | SN011T | THERMAL STRESS | 33696656 | 33186959 | 28789248 (86.75%) | AT |
| | SN015 | SN015C | CONTROL | 36855926 | 36591717 | 31217898 (85.31%) | BC |
| | | SN015T | THERMAL STRESS | 34775132 | 34512645 | 29335467 (85.00%) | BT |
| | SN018 | SN018C | CONTROL | 32882051 | 32652498 | 29589369 (90.62%) | BC |
| | | SN018T | THERMAL STRESS | 31299479 | 31072836 | 27799014 (89.46%) | BT |
| | SN022 | SN022C | CONTROL | 30842803 | 30287023 | 26738767 (88.28%) | AC |
| | | SN022T | THERMAL STRESS | 31861929 | 31333029 | 27342152 (87.26%) | AT |
| | SN027 | SN027C | CONTROL | 31289486 | 30674899 | 26366641 (85.96%) | AC |
| | | SN027T | THERMAL STRESS | 31058573 | 30585757 | 25747883 (84.18%) | AT |
| | SN028 | SN028C | CONTROL | 32582810 | 31917710 | 28143974 (88.18%) | AC |
| | | SN028T | THERMAL STRESS | 31671629 | 31134984 | 27040689 (86.85%) | AT |
| | SN031 | SN031C | CONTROL | 33132288 | 32857214 | 29012345 (88.30%) | BC |
| | | SN031T | THERMAL STRESS | 30923747 | 30671410 | 25695190 (83.78%) | BT |
| | SN032 | SN032C | CONTROL | 30345759 | 30061759 | 26909506 (89.51%) | BC |
| | | SN032T | THERMAL STRESS | 33771420 | 32194526 | 27228799 (84.58%) | BT |
| | SN034 | SN034C | CONTROL | 35246628 | 34806970 | 29843577 (85.74%) | AC |
| | | SN034T | THERMAL STRESS | 35551582 | 34854469 | 30095927 (86.35%) | AT |

**Appendix 2.** Continued.

| ***Funaria* species** | **Code** | **Code in RNA analysis** | **Environmental conditions** | **Raw reads** | **Clean reads** | **Mapped reads (to *de novo* transcriptome)** | **Transcriptomic group** |
|---|---|---|---|---|---|---|---|
| *F. hygrometrica* | SN036 | SN036C | CONTROL | 33787370 | 33248579 | 29801370 (89.63%) | **AC** |
| | | SN036T | THERMAL STRESS | 32735122 | 32436430 | 28936693 (89.21%) | **AT** |
| | SN039 | SN039C | CONTROL | 30720395 | 30441817 | 27322225 (89.75%) | **BC** |
| | | SN039T | THERMAL STRESS | 35680491 | 34978838 | 27989430 (80.02%) | **BT** |
| | SN051 | SN051C | CONTROL | 35664415 | 35321223 | 31261194 (88.51%) | **BC** |
| | | SN051T | THERMAL STRESS | 30362240 | 30005485 | 25600924 (85.32%) | **BT** |
| | SN057 | SN057C | CONTROL | 30850402 | 30580728 | 26481726 (86.60%) | **BC** |
| | | SN057T | THERMAL STRESS | 32360283 | 32046565 | 27729566 (86.53%) | **BT** |
| | SN064 | SN064C | CONTROL | 31803360 | 31520802 | 27375828 (86.85%) | **BC** |
| | | SN064T | THERMAL STRESS | 30159042 | 29881592 | 25071365 (83.90%) | **BT** |
| | SN069 | SN069C | CONTROL | 37591818 | 37307014 | 33023731 (88.52%) | **BC** |
| | | SN069T | THERMAL STRESS | 32016331 | 31697423 | 26658469 (84.10%) | **BT** |
| | SN073 | SN073C | CONTROL | 33808494 | 33488618 | 29787364 (88.95%) | **BC** |
| | | SN073T | THERMAL STRESS | 32784482 | 32414860 | 25919103 (79.96%) | **BT** |
| *F. aequidens* | Ru2 | Ru2 | CONTROL | 30616778 | 30359865 | 26148489 (86.13%) | **RU** |
| *F. arctica* | Ru4 | Ru4 | CONTROL | 32030187 | 31713265 | 27834031 (87.77%) | **RU** |
| *F. polaris* | Ru1 | Ru1 | CONTROL | 31048987 | 30775458 | 27236603 (88.50%) | **RU** |
| | Ru3 | Ru3 | CONTROL | 33719754 | 33389629 | 29251424 (87.61%) | **RU** |

(*) = Technical replicate

| | |
|---|---|
| CONTROL | 16 hours darkness in growth chamber (20-22ºC) |
| THERMAL STREES | 16 hours darkness in laboratory oven (40ºC) |

**Appendix 3.1.** Admixture coefficients calculated by sNMF for all individuals studied of *F. hygrometrica* for the number of ancestral populations $K = 2$. The individuals are assigned to cluster A or cluster B respectively when the ancestry coefficient of a given individual is greater than 0.5 for corresponding component, what is indicated by orange colour in individuals with a major component of cluster A and blue colour in individuals with a major component of cluster B. For identification of individuals see code in Appendix 1.

| | Cluster A | Cluster B | | Cluster A | Cluster B | | Cluster A | Cluster B |
|---|---|---|---|---|---|---|---|---|
| AS002 | 0.0001 | 0.9999 | LE026 | 0.0001 | 0.9999 | SN013 | 0.1023 | 0.8978 |
| AS004 | 0.0001 | 0.9999 | LE027 | 0.0001 | 0.9999 | SN014 | 0.0001 | 0.9999 |
| AS005 | 0.0001 | 0.9999 | LE028 | 0.0001 | 0.9999 | SN015 | 0.0001 | 0.9999 |
| AS006 | 0.4394 | 0.5606 | LE029 | 0.0001 | 0.9999 | SN017 | 0.0001 | 0.9999 |
| AS007 | 0.0001 | 0.9999 | LE030 | 0.9999 | 0.0001 | SN018 | 0.0001 | 0.9999 |
| AS008 | 0.6394 | 0.3606 | LE031 | 0.0001 | 0.9999 | SN019 | 0.0001 | 0.9999 |
| AS009 | 0.0001 | 0.9999 | LE032 | 0.0001 | 0.9999 | SN020 | 0.9969 | 0.0031 |
| AS010 | 0.0001 | 0.9999 | LE033 | 0.0001 | 0.9999 | SN021 | 0.9999 | 0.0001 |
| AS011 | 0.0001 | 0.9999 | LE034 | 0.0001 | 0.9999 | SN022 | 0.9999 | 0.0001 |
| AS012 | 0.2637 | 0.7363 | LE035 | 0.0001 | 0.9999 | SN023 | 0.9999 | 0.0001 |
| AS013 | 0.0001 | 0.9999 | LE037 | 0.0001 | 0.9999 | SN025 | 0.8838 | 0.1162 |
| AS014 | 0.0001 | 0.9999 | MU002 | 0.0856 | 0.9144 | SN027 | 0.9999 | 0.0001 |
| AS015 | 0.0001 | 0.9999 | MU003 | 0.0001 | 0.9999 | SN028 | 0.9999 | 0.0001 |
| AS016 | 0.0001 | 0.9999 | MU005b | 0.0001 | 0.9999 | SN029 | 0.9999 | 0.0001 |
| AS017 | 0.0001 | 0.9999 | MU006a | 0.0001 | 0.9999 | SN030 | 0.0001 | 0.9999 |
| AS018 | 0.0001 | 0.9999 | MU006b | 0.0001 | 0.9999 | SN031 | 0.0001 | 0.9999 |
| AS019 | 0.0001 | 0.9999 | MU007 | 0.0001 | 0.9999 | SN032 | 0.0001 | 0.9999 |
| AS020a | 0.0001 | 0.9999 | MU009 | 0.0001 | 0.9999 | SN033 | 0.2820 | 0.7180 |
| AS020b | 0.0001 | 0.9999 | MU010 | 0.9571 | 0.0429 | SN034 | 0.9999 | 0.0001 |
| AS022 | 0.0001 | 0.9999 | MU011 | 0.0001 | 0.9999 | SN035 | 0.9999 | 0.0001 |
| AS023 | 0.0001 | 0.9999 | MU012 | 0.0001 | 0.9999 | SN036a | 0.9999 | 0.0001 |
| AS024 | 0.9670 | 0.0330 | MU013a | 0.0001 | 0.9999 | SN036b | 0.9999 | 0.0001 |
| AS025 | 0.0001 | 0.9999 | MU013b | 0.0354 | 0.9646 | SN038 | 0.0001 | 0.9999 |
| AS026 | 0.0001 | 0.9999 | MU014 | 0.0001 | 0.9999 | SN039 | 0.0001 | 0.9999 |
| AS027 | 0.0001 | 0.9999 | MU015 | 0.0001 | 0.9999 | SN040 | 0.0001 | 0.9999 |
| AS028b | 0.0001 | 0.9999 | MU016 | 0.0001 | 0.9999 | SN041 | 0.9098 | 0.0902 |
| AS029 | 0.0001 | 0.9999 | MU017 | 0.0001 | 0.9999 | SN042 | 0.9060 | 0.0940 |
| AS030 | 0.0001 | 0.9999 | MU018 | 0.0001 | 0.9999 | SN043 | 0.0001 | 0.9999 |
| AS031 | 0.9422 | 0.0578 | MU019 | 0.0001 | 0.9999 | SN044 | 0.0001 | 0.9999 |
| AS032 | 0.0001 | 0.9999 | MU020 | 0.0001 | 0.9999 | SN045 | 0.0001 | 0.9999 |
| AS033 | 0.0001 | 0.9999 | MU021 | 0.0001 | 0.9999 | SN046 | 0.0001 | 0.9999 |
| AS035 | 0.0033 | 0.9967 | MU022a | 0.0001 | 0.9999 | SN047 | 0.0001 | 0.9999 |
| AS036 | 0.0001 | 0.9999 | MU022b | 0.0001 | 0.9999 | SN048 | 0.0001 | 0.9999 |
| LE001 | 0.0001 | 0.9999 | MU023 | 0.0001 | 0.9999 | SN049 | 0.0001 | 0.9999 |
| LE002 | 0.0001 | 0.9999 | MU024 | 0.0001 | 0.9999 | SN050 | 0.0001 | 0.9999 |
| LE003 | 0.0001 | 0.9999 | MU025 | 0.0001 | 0.9999 | SN051 | 0.0001 | 0.9999 |
| LE004 | 0.0001 | 0.9999 | MU026 | 0.0001 | 0.9999 | SN052 | 0.0001 | 0.9999 |
| LE005 | 0.0001 | 0.9999 | MU028 | 0.0001 | 0.9999 | SN053 | 0.0001 | 0.9999 |
| LE006 | 0.0001 | 0.9999 | MU029 | 0.0001 | 0.9999 | SN054 | 0.0001 | 0.9999 |
| LE007 | 0.0001 | 0.9999 | MU030 | 0.9999 | 0.0001 | SN055 | 0.0001 | 0.9999 |
| LE008 | 0.0001 | 0.9999 | MU031 | 0.0001 | 0.9999 | SN056 | 0.0001 | 0.9999 |
| LE009 | 0.0001 | 0.9999 | MU032 | 0.0001 | 0.9999 | SN057 | 0.0001 | 0.9999 |
| LE010 | 0.0001 | 0.9999 | MU033 | 0.0001 | 0.9999 | SN059 | 0.0001 | 0.9999 |
| LE011 | 0.0001 | 0.9999 | MU034 | 0.0001 | 0.9999 | SN060 | 0.0001 | 0.9999 |
| LE012 | 0.0001 | 0.9999 | MU035 | 0.0001 | 0.9999 | SN061 | 0.0001 | 0.9999 |
| LE013b | 0.0001 | 0.9999 | MU036 | 0.0001 | 0.9999 | SN062 | 0.0001 | 0.9999 |
| LE014 | 0.0001 | 0.9999 | MU037 | 0.9999 | 0.0001 | SN063 | 0.0001 | 0.9999 |
| LE015 | 0.0001 | 0.9999 | SN001 | 0.0001 | 0.9999 | SN064 | 0.0001 | 0.9999 |
| LE017 | 0.0001 | 0.9999 | SN002 | 0.0001 | 0.9999 | SN065 | 0.0001 | 0.9999 |
| LE018 | 0.0001 | 0.9999 | SN003 | 0.9999 | 0.0001 | SN066 | 0.0001 | 0.9999 |
| LE019 | 0.0001 | 0.9999 | SN004 | 0.9999 | 0.0001 | SN067 | 0.0001 | 0.9999 |
| LE020 | 0.0001 | 0.9999 | SN006 | 0.0001 | 0.9999 | SN068 | 0.0001 | 0.9999 |
| LE021 | 0.0001 | 0.9999 | SN007 | 0.9999 | 0.0001 | SN069 | 0.0001 | 0.9999 |
| LE022 | 0.0001 | 0.9999 | SN008 | 0.9840 | 0.0160 | SN070 | 0.0001 | 0.9999 |
| LE023 | 0.0001 | 0.9999 | SN009 | 0.8506 | 0.1494 | SN071 | 0.9990 | 0.0010 |
| LE024 | 0.0001 | 0.9999 | SN010 | 0.9999 | 0.0001 | SN072 | 0.9058 | 0.0942 |
| LE025 | 0.0001 | 0.9999 | SN011 | 0.9999 | 0.0001 | SN073 | 0.0001 | 0.9999 |

**Appendix 3.2.** Admixture coefficients calculated by sNMF for individuals studied of *F. hygrometrica* except outliers (AS006, AS008, AS012, MU002, SN013, SN033, SN041 and SN042) for the number of ancestral populations $K = 2$. Cluster assignement criterion as in Appendix 3.1. For identification of individuals see code in Appendix 1.

| | Cluster A | Cluster B | | Cluster A | Cluster B | | Cluster A | Cluster B |
|---|---|---|---|---|---|---|---|---|
| AS002 | 0.0001 | 0.9999 | LE026 | 0.0001 | 0.9999 | SN010 | 0.9999 | 0.0001 |
| AS004 | 0.0001 | 0.9999 | LE027 | 0.0001 | 0.9999 | SN011 | 0.9999 | 0.0001 |
| AS005 | 0.0001 | 0.9999 | LE028 | 0.0001 | 0.9999 | SN014 | 0.0001 | 0.9999 |
| AS007 | 0.0001 | 0.9999 | LE029 | 0.0001 | 0.9999 | SN015 | 0.0001 | 0.9999 |
| AS009 | 0.0001 | 0.9999 | LE030 | 0.9999 | 0.0001 | SN017 | 0.0001 | 0.9999 |
| AS010 | 0.0001 | 0.9999 | LE031 | 0.0001 | 0.9999 | SN018 | 0.0001 | 0.9999 |
| AS011 | 0.0001 | 0.9999 | LE032 | 0.0001 | 0.9999 | SN019 | 0.0001 | 0.9999 |
| AS013 | 0.0001 | 0.9999 | LE033 | 0.0001 | 0.9999 | SN020 | 0.9999 | 0.0001 |
| AS014 | 0.0001 | 0.9999 | LE034 | 0.0001 | 0.9999 | SN021 | 0.9999 | 0.0001 |
| AS015 | 0.0001 | 0.9999 | LE035 | 0.0001 | 0.9999 | SN022 | 0.9999 | 0.0001 |
| AS016 | 0.0001 | 0.9999 | LE037 | 0.0001 | 0.9999 | SN023 | 0.9999 | 0.0001 |
| AS017 | 0.0001 | 0.9999 | MU003 | 0.0001 | 0.9999 | SN025 | 0.8846 | 0.1154 |
| AS018 | 0.0001 | 0.9999 | MU005b | 0.0001 | 0.9999 | SN027 | 0.9999 | 0.0001 |
| AS019 | 0.0001 | 0.9999 | MU006a | 0.0001 | 0.9999 | SN028 | 0.9999 | 0.0001 |
| AS020a | 0.0040 | 0.9960 | MU006b | 0.0001 | 0.9999 | SN029 | 0.9999 | 0.0001 |
| AS020b | 0.0001 | 0.9999 | MU007 | 0.0001 | 0.9999 | SN030 | 0.0001 | 0.9999 |
| AS022 | 0.0001 | 0.9999 | MU009 | 0.0001 | 0.9999 | SN031 | 0.0001 | 0.9999 |
| AS023 | 0.0001 | 0.9999 | MU010 | 0.9797 | 0.0203 | SN032 | 0.0001 | 0.9999 |
| AS024 | 0.9831 | 0.0169 | MU011 | 0.0001 | 0.9999 | SN034 | 0.9999 | 0.0001 |
| AS025 | 0.0001 | 0.9999 | MU012 | 0.0001 | 0.9999 | SN035 | 0.9999 | 0.0001 |
| AS026 | 0.0001 | 0.9999 | MU013a | 0.0001 | 0.9999 | SN036a | 0.9999 | 0.0001 |
| AS027 | 0.0001 | 0.9999 | MU013b | 0.0047 | 0.9953 | SN036b | 0.9941 | 0.0059 |
| AS028b | 0.0001 | 0.9999 | MU014 | 0.0001 | 0.9999 | SN038 | 0.0001 | 0.9999 |
| AS029 | 0.0001 | 0.9999 | MU015 | 0.0001 | 0.9999 | SN039 | 0.0001 | 0.9999 |
| AS030 | 0.0001 | 0.9999 | MU016 | 0.0001 | 0.9999 | SN040 | 0.0001 | 0.9999 |
| AS031 | 0.9785 | 0.0215 | MU017 | 0.0001 | 0.9999 | SN043 | 0.0001 | 0.9999 |
| AS032 | 0.0001 | 0.9999 | MU018 | 0.0001 | 0.9999 | SN044 | 0.0001 | 0.9999 |
| AS033 | 0.0001 | 0.9999 | MU019 | 0.0001 | 0.9999 | SN045 | 0.0001 | 0.9999 |
| AS035 | 0.0096 | 0.9904 | MU020 | 0.0001 | 0.9999 | SN046 | 0.0001 | 0.9999 |
| AS036 | 0.0001 | 0.9999 | MU021 | 0.0141 | 0.9859 | SN047 | 0.0001 | 0.9999 |
| LE001 | 0.0001 | 0.9999 | MU022a | 0.0001 | 0.9999 | SN048 | 0.0001 | 0.9999 |
| LE002 | 0.0001 | 0.9999 | MU022b | 0.0001 | 0.9999 | SN049 | 0.0001 | 0.9999 |
| LE003 | 0.0001 | 0.9999 | MU023 | 0.0001 | 0.9999 | SN050 | 0.0001 | 0.9999 |
| LE004 | 0.0001 | 0.9999 | MU024 | 0.0001 | 0.9999 | SN051 | 0.0001 | 0.9999 |
| LE005 | 0.0001 | 0.9999 | MU025 | 0.0001 | 0.9999 | SN052 | 0.0001 | 0.9999 |
| LE006 | 0.0001 | 0.9999 | MU026 | 0.0001 | 0.9999 | SN053 | 0.0001 | 0.9999 |
| LE007 | 0.0001 | 0.9999 | MU028 | 0.0001 | 0.9999 | SN054 | 0.0001 | 0.9999 |
| LE008 | 0.0001 | 0.9999 | MU029 | 0.0001 | 0.9999 | SN055 | 0.0001 | 0.9999 |
| LE009 | 0.0001 | 0.9999 | MU030 | 0.9999 | 0.0001 | SN056 | 0.0001 | 0.9999 |
| LE010 | 0.0001 | 0.9999 | MU031 | 0.0001 | 0.9999 | SN057 | 0.0001 | 0.9999 |
| LE011 | 0.0001 | 0.9999 | MU032 | 0.0001 | 0.9999 | SN059 | 0.0001 | 0.9999 |
| LE012 | 0.0001 | 0.9999 | MU033 | 0.0001 | 0.9999 | SN060 | 0.0001 | 0.9999 |
| LE013b | 0.0064 | 0.9936 | MU034 | 0.0001 | 0.9999 | SN061 | 0.0001 | 0.9999 |
| LE014 | 0.0001 | 0.9999 | MU035 | 0.0001 | 0.9999 | SN062 | 0.0001 | 0.9999 |
| LE015 | 0.0001 | 0.9999 | MU036 | 0.0001 | 0.9999 | SN063 | 0.0001 | 0.9999 |
| LE017 | 0.0001 | 0.9999 | MU037 | 0.9999 | 0.0001 | SN064 | 0.0001 | 0.9999 |
| LE018 | 0.0001 | 0.9999 | SN001 | 0.0001 | 0.9999 | SN065 | 0.0001 | 0.9999 |
| LE019 | 0.0001 | 0.9999 | SN002 | 0.0001 | 0.9999 | SN066 | 0.0001 | 0.9999 |
| LE020 | 0.0001 | 0.9999 | SN003 | 0.9999 | 0.0001 | SN067 | 0.0001 | 0.9999 |
| LE021 | 0.0001 | 0.9999 | SN004 | 0.9999 | 0.0001 | SN068 | 0.0001 | 0.9999 |
| LE022 | 0.0001 | 0.9999 | SN006 | 0.0001 | 0.9999 | SN069 | 0.0001 | 0.9999 |
| LE023 | 0.0001 | 0.9999 | SN007 | 0.9999 | 0.0001 | SN070 | 0.0001 | 0.9999 |
| LE024 | 0.0001 | 0.9999 | SN008 | 0.9999 | 0.0001 | SN071 | 0.9999 | 0.0001 |
| LE025 | 0.0001 | 0.9999 | SN009 | 0.8418 | 0.1582 | SN072 | 0.9377 | 0.0623 |

**Appendix 3.3.** Admixture coefficients calculated by sNMF for SN individuals without outliers of this area (SN013, SN033, SN041 and SN042) for the number of ancestral populations $K = 2$. Cluster assignement criterion as in Appendix 3.1. For identification of individuals see code in Appendix 1.

| | Cluster A | Cluster B | | Cluster A | Cluster B |
|---|---|---|---|---|---|
| SN001 | 0.0001 | 0.9999 | SN039 | 0.0001 | 0.9999 |
| SN002 | 0.0001 | 0.9999 | SN040 | 0.0001 | 0.9999 |
| SN003 | 0.9999 | 0.0001 | SN043 | 0.0001 | 0.9999 |
| SN004 | 0.9999 | 0.0001 | SN044 | 0.0001 | 0.9999 |
| SN006 | 0.0001 | 0.9999 | SN045 | 0.0001 | 0.9999 |
| SN007 | 0.9999 | 0.0001 | SN046 | 0.0001 | 0.9999 |
| SN008 | 0.9566 | 0.0434 | SN047 | 0.0001 | 0.9999 |
| SN009 | 0.9636 | 0.0364 | SN048 | 0.0001 | 0.9999 |
| SN010 | 0.9999 | 0.0001 | SN049 | 0.0001 | 0.9999 |
| SN011 | 0.9999 | 0.0001 | SN050 | 0.0001 | 0.9999 |
| SN013 | 0.1041 | 0.8959 | SN051 | 0.0001 | 0.9999 |
| SN014 | 0.0001 | 0.9999 | SN052 | 0.0001 | 0.9999 |
| SN015 | 0.0001 | 0.9999 | SN053 | 0.0001 | 0.9999 |
| SN017 | 0.0001 | 0.9999 | SN054 | 0.0001 | 0.9999 |
| SN018 | 0.0001 | 0.9999 | SN055 | 0.0001 | 0.9999 |
| SN019 | 0.0001 | 0.9999 | SN056 | 0.0001 | 0.9999 |
| SN020 | 0.9999 | 0.0001 | SN057 | 0.0001 | 0.9999 |
| SN021 | 0.9999 | 0.0001 | SN059 | 0.0001 | 0.9999 |
| SN022 | 0.9999 | 0.0001 | SN060 | 0.0001 | 0.9999 |
| SN023 | 0.9999 | 0.0001 | SN061 | 0.0001 | 0.9999 |
| SN025 | 0.9357 | 0.0643 | SN062 | 0.0001 | 0.9999 |
| SN027 | 0.9999 | 0.0001 | SN063 | 0.0001 | 0.9999 |
| SN028 | 0.9999 | 0.0001 | SN064 | 0.0001 | 0.9999 |
| SN029 | 0.9999 | 0.0001 | SN065 | 0.0001 | 0.9999 |
| SN030 | 0.0001 | 0.9999 | SN066 | 0.0001 | 0.9999 |
| SN031 | 0.0001 | 0.9999 | SN067 | 0.0001 | 0.9999 |
| SN032 | 0.0001 | 0.9999 | SN068 | 0.0043 | 0.9957 |
| SN033 | 0.2273 | 0.7727 | SN069 | 0.0001 | 0.9999 |
| SN034 | 0.9999 | 0.0001 | SN070 | 0.0001 | 0.9999 |
| SN035 | 0.9999 | 0.0001 | SN071 | 0.9999 | 0.0001 |
| SN036a | 0.9999 | 0.0001 | SN072 | 0.9455 | 0.0545 |
| SN036b | 0.9811 | 0.0189 | SN073 | 0.0001 | 0.9999 |
| SN038 | 0.0001 | 0.9999 | | | |