

# ***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***

---

José Carlos del Valle

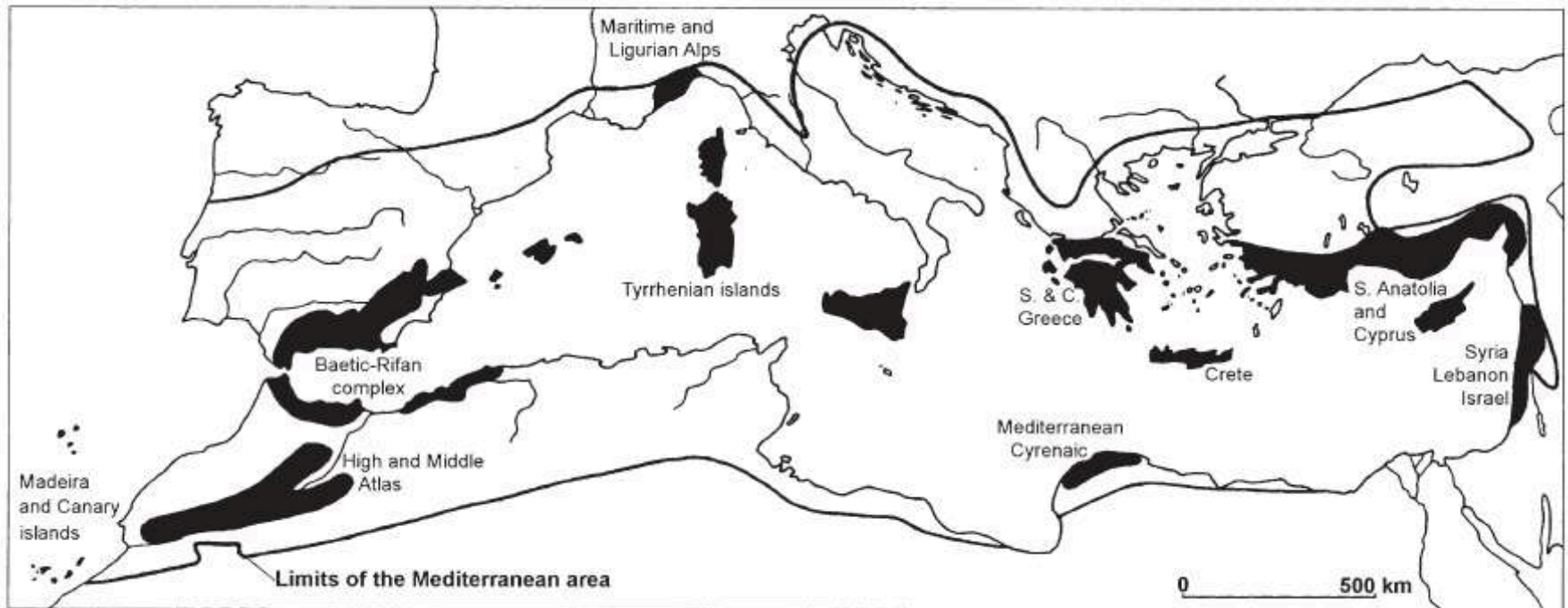
[jvgarcia@us.es](mailto:jvgarcia@us.es)

**I Simposio Anual de  
Botánica Española**

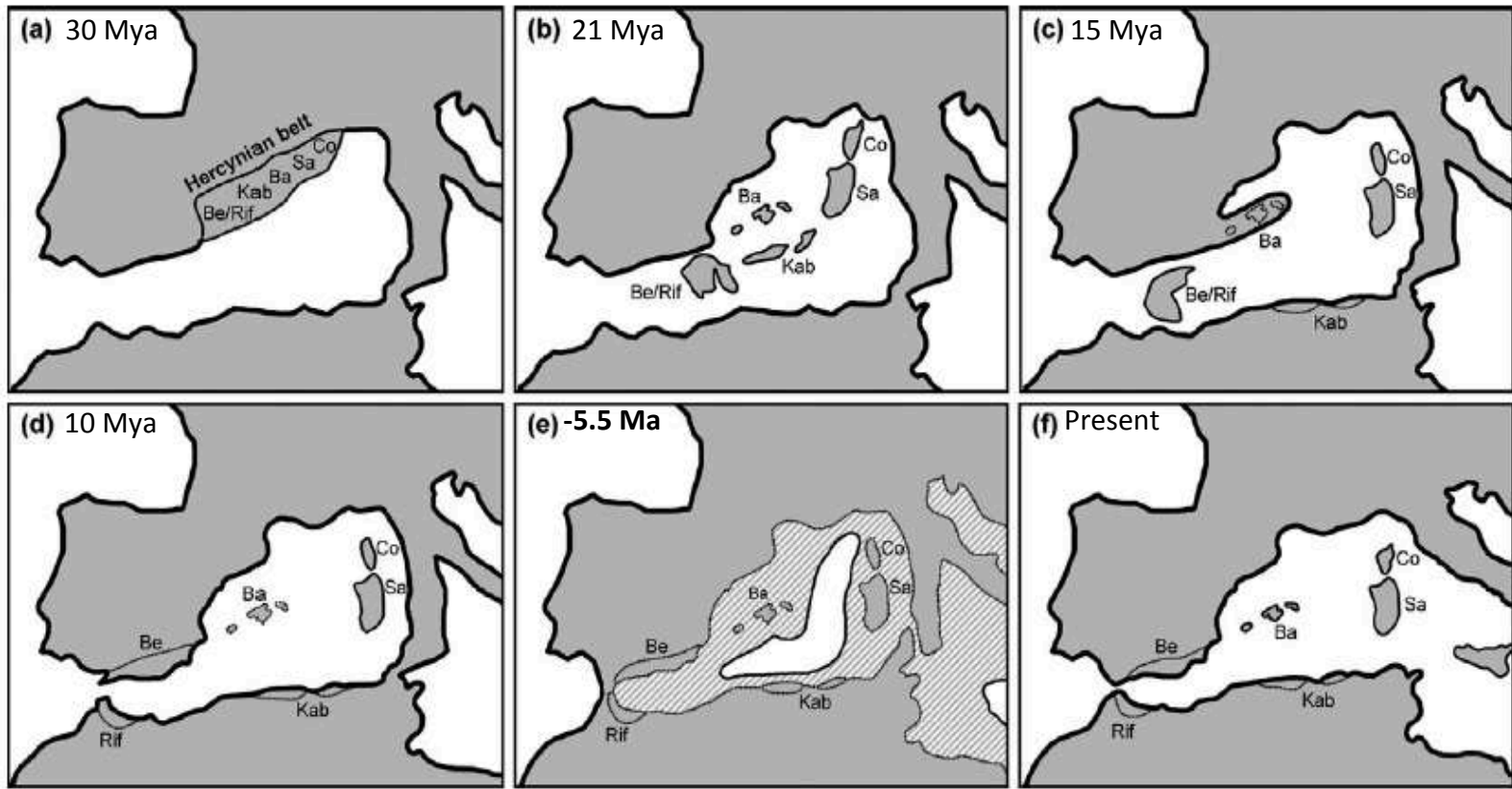


## Mediterranean Basin: hotspot of biodiversity

More than 20,000 plant species → 60% (13,000 spp.) are endemic of this region

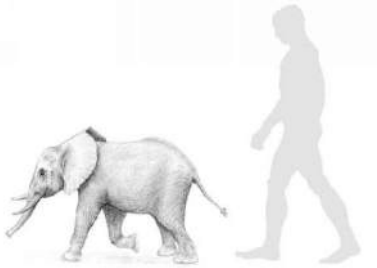


The Balearic Islands are especially rich in endemics, making them excellent models for understanding speciation

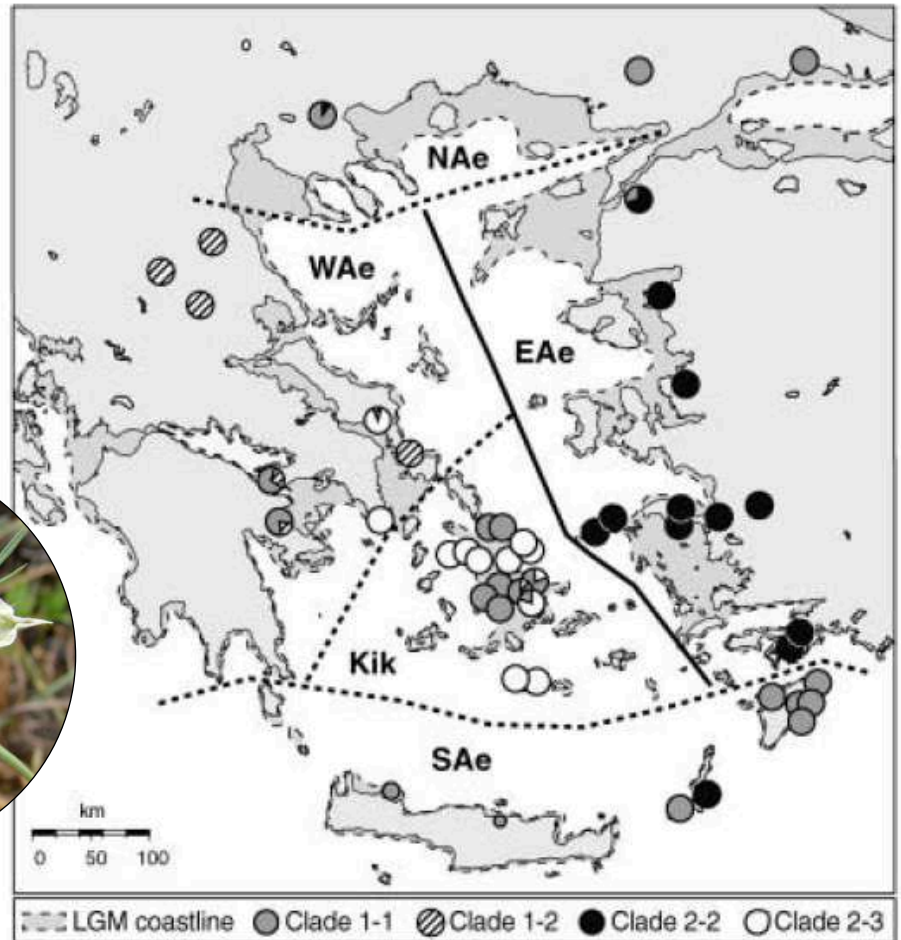
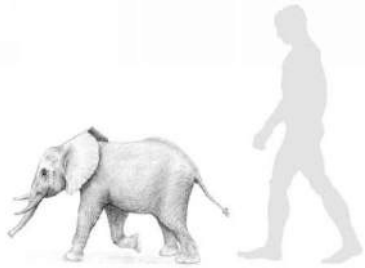
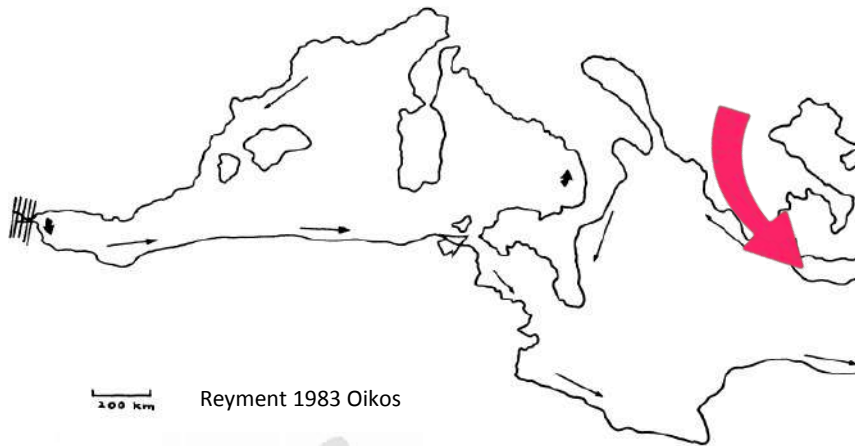


The isolation of Balearic Islands contrasts with the colonization events in other Mediterranean islands

The isolation of Balearic Islands contrasts with the colonization events in other Mediterranean islands

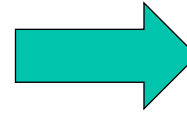


The isolation of Balearic Islands contrasts with the colonization events in other Mediterranean islands

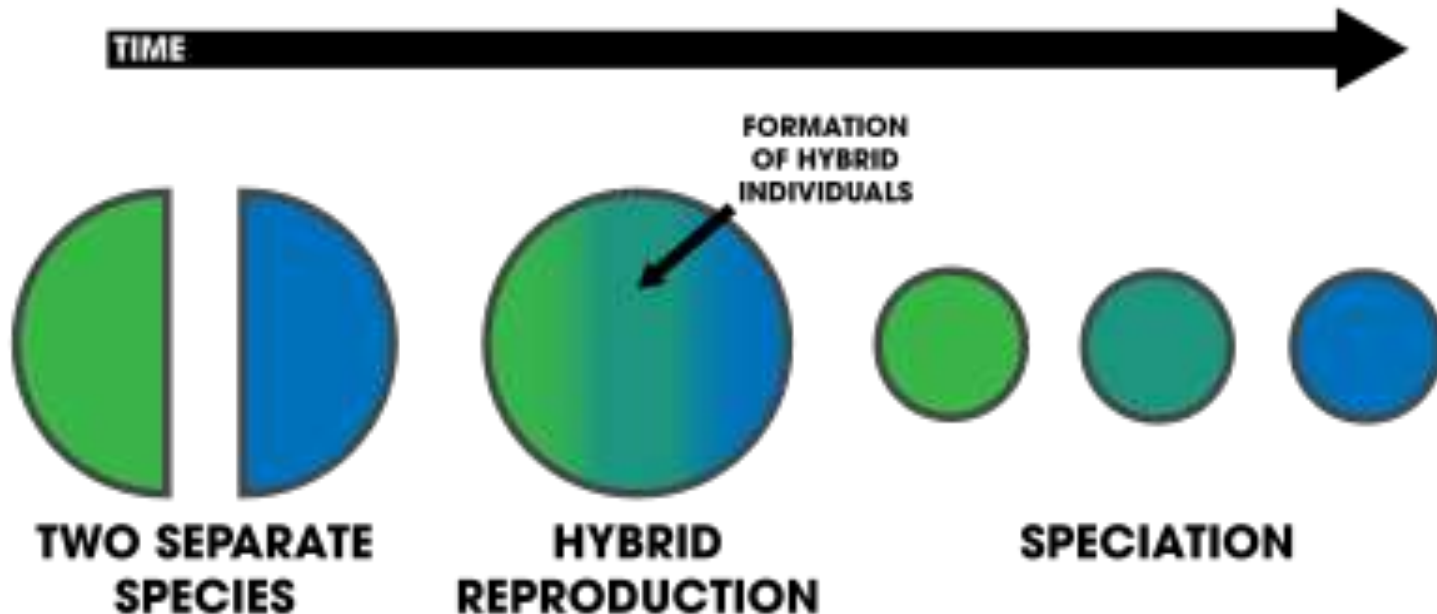


In the Iberian Peninsula, dramatic geological and climatic changes have repeatedly caused fragmentation, contraction, and expansion of species ranges

In the Iberian Peninsula, dramatic geological and climatic changes have repeatedly caused fragmentation, contraction, and expansion of species ranges



**Increased chances for hybridization and introgression**



Hybridization events are important events in the tribe Sileneae  
(Caryophyllaceae)

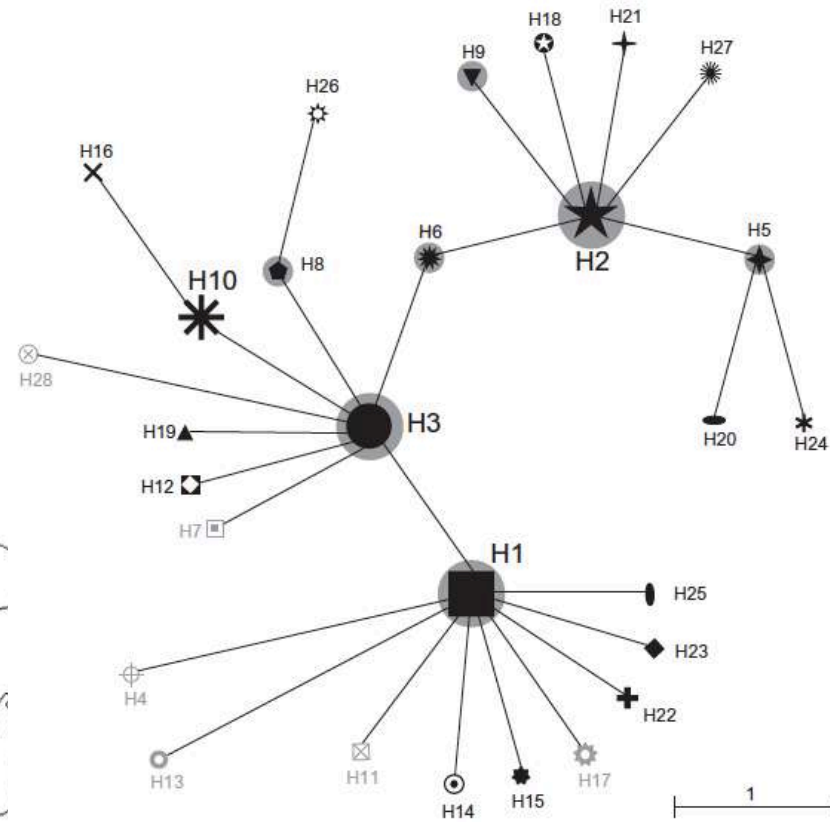
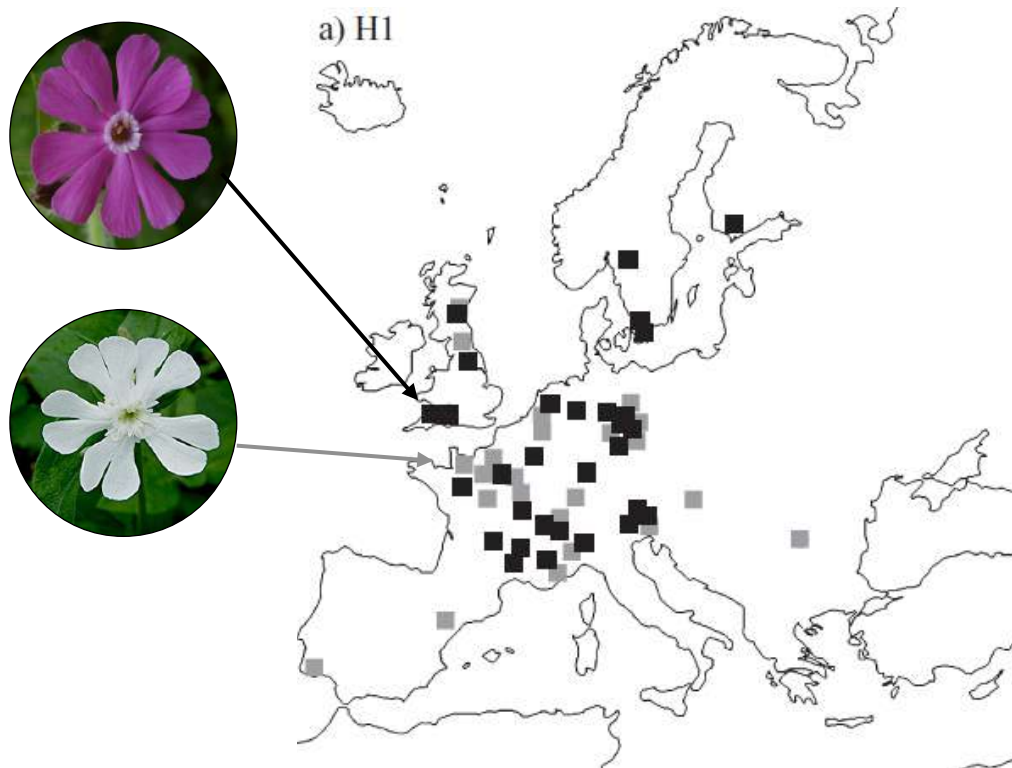
## Hybridization events are important events in the tribe Sileneae (Caryophyllaceae)

The ability of *Silene latifolia* and *S. dioica* to hybridize is one of the best examples of incomplete reproductive isolation in this group



# Hybridization events are important events in the tribe Sileneae (Caryophyllaceae)

The ability of *Silene latifolia* and *S. dioica* to hybridize is one of the best examples of incomplete reproductive isolation in this group

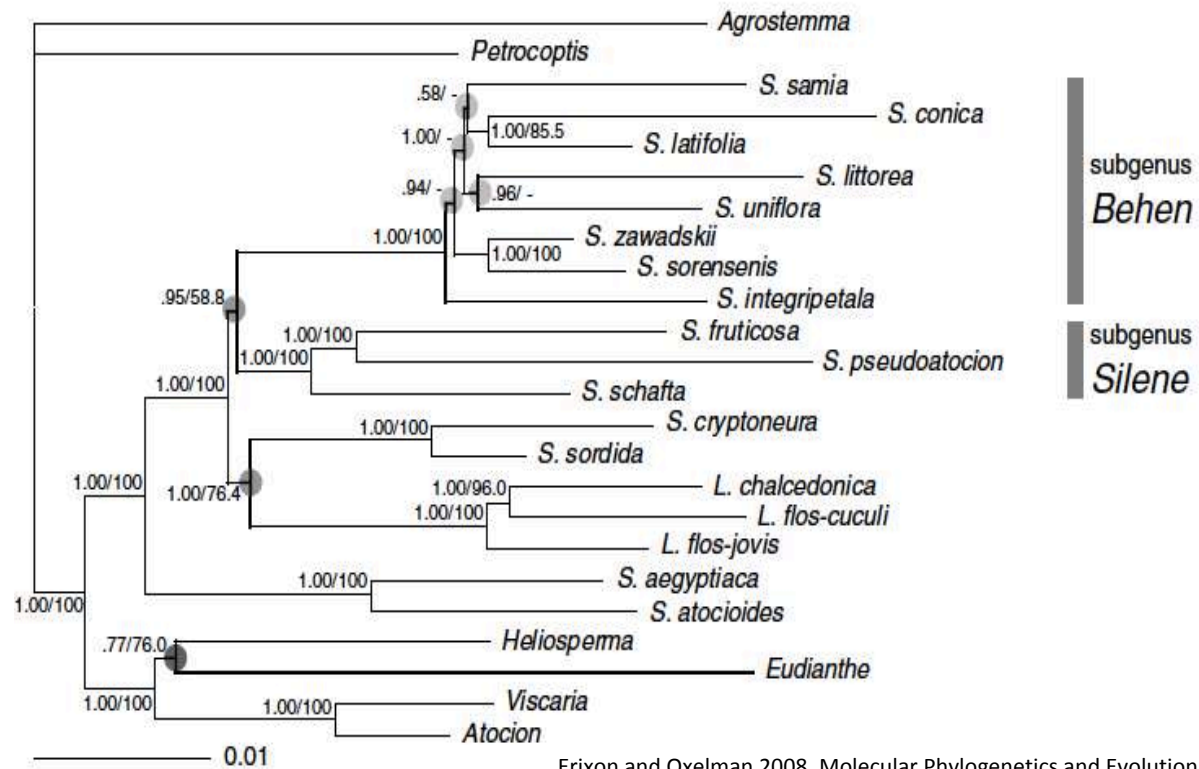


Phylogenies in Sileneae are frequently incongruent with morphology-based classifications

## Phylogenies in Sileneae are frequently incongruent with morphology-based classifications

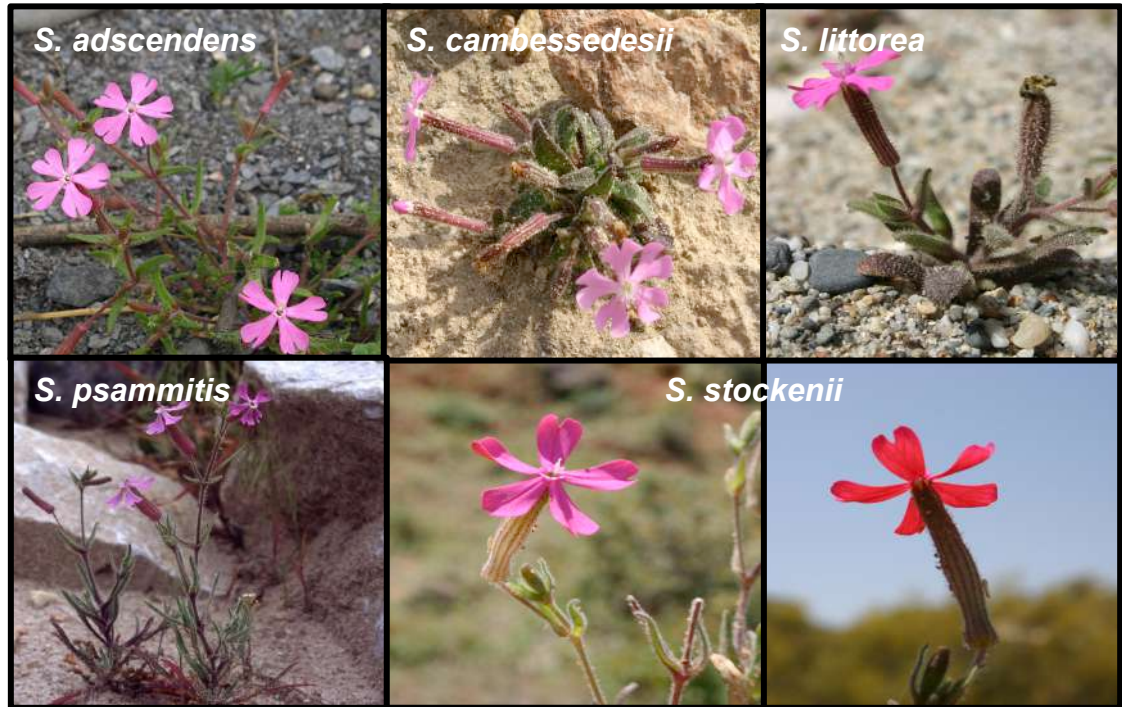
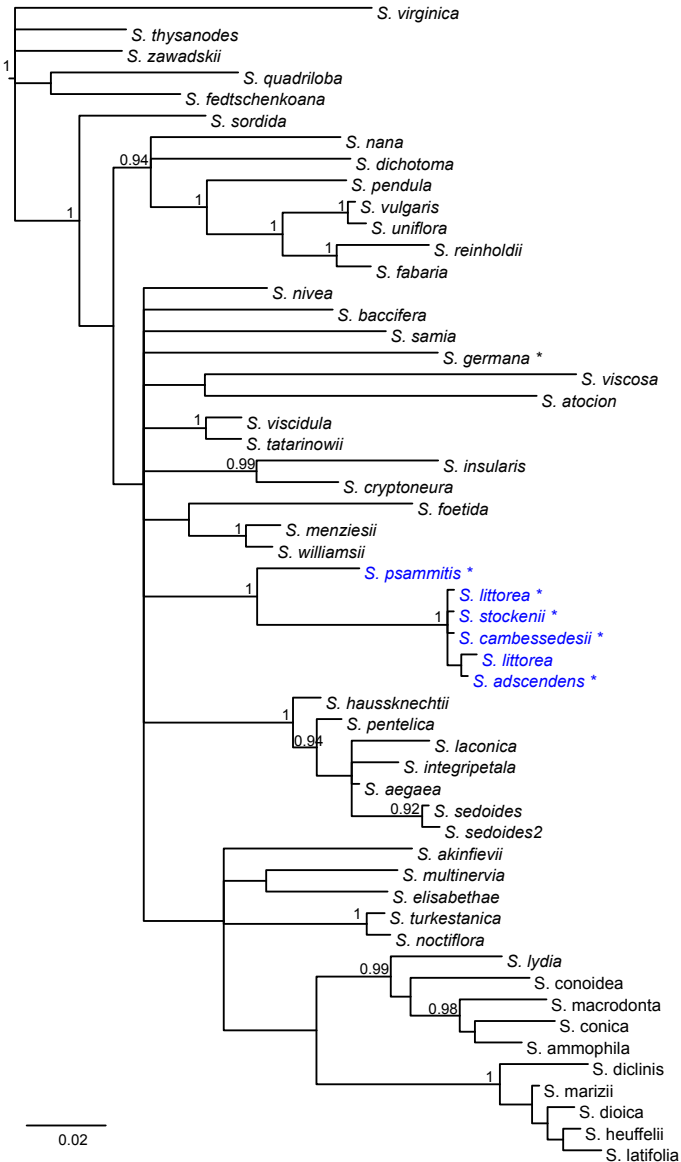
The tribe Sileneae is subdivided into eight genera → *Silene* is the most diverse, with approximately 470 spp.

- Subgenus *Silene*
- Subgenus *Behenantha*

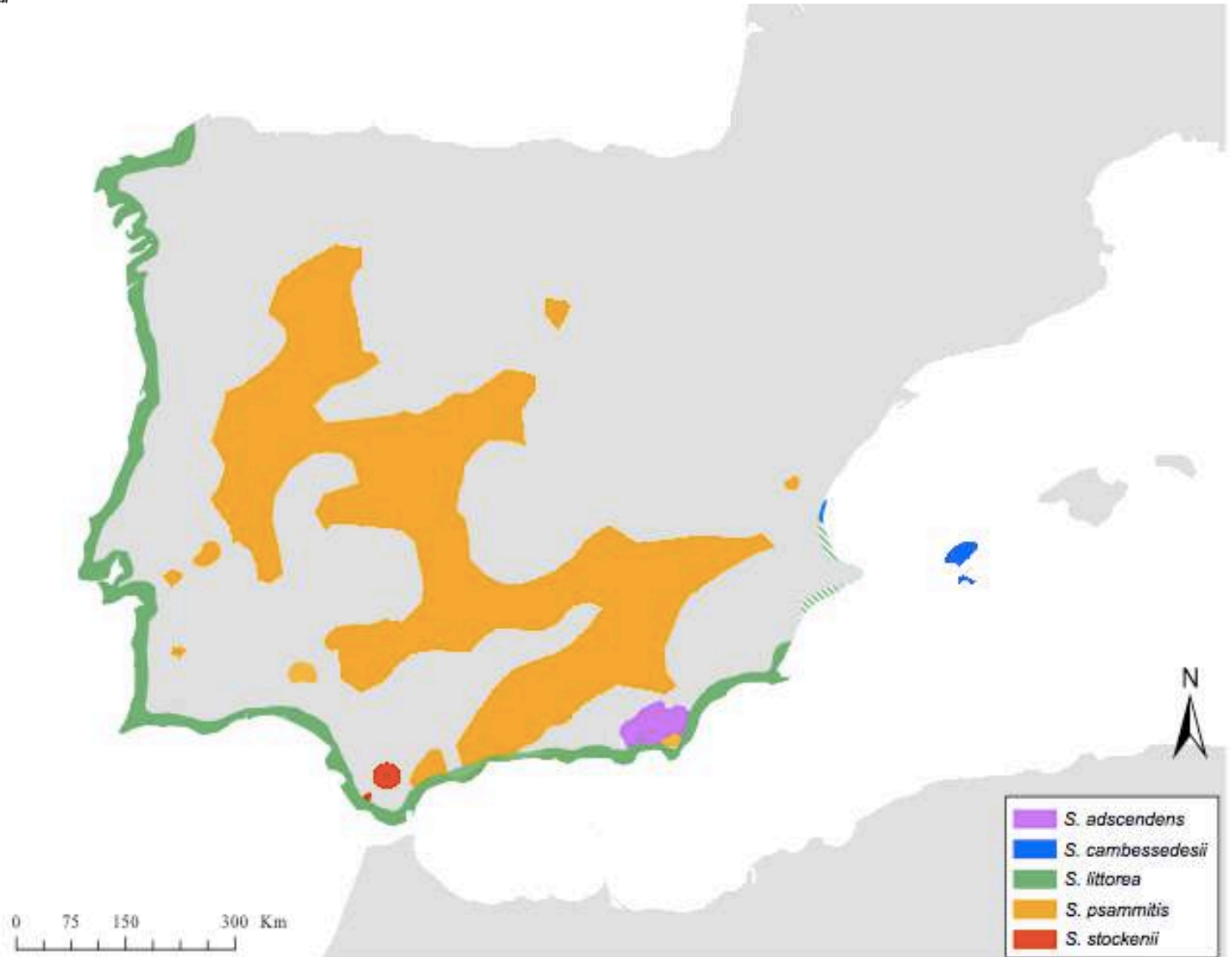


## *Silene* Sect. *Psammophilae*

- Five species endemic to the Iberian Peninsula and Balearic Islands
- Self-compatible, mainly pollinated by insects
- Distinct edaphic affinities



***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***



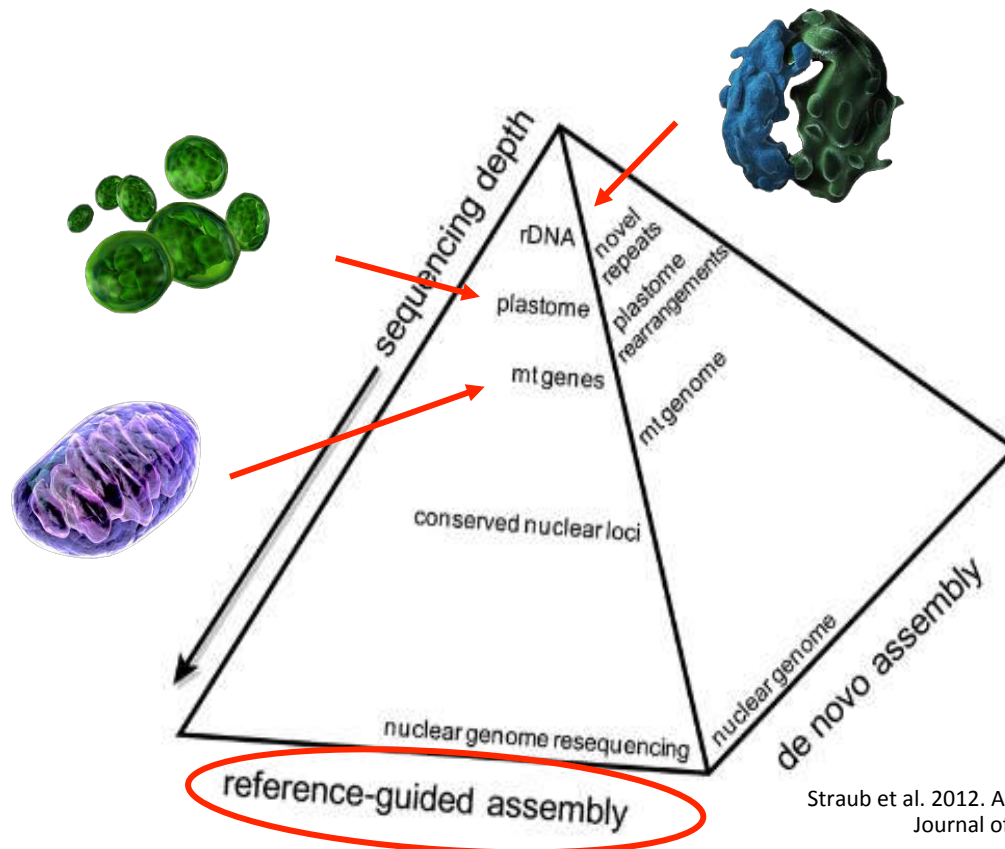
## **What are the aims of this study?**

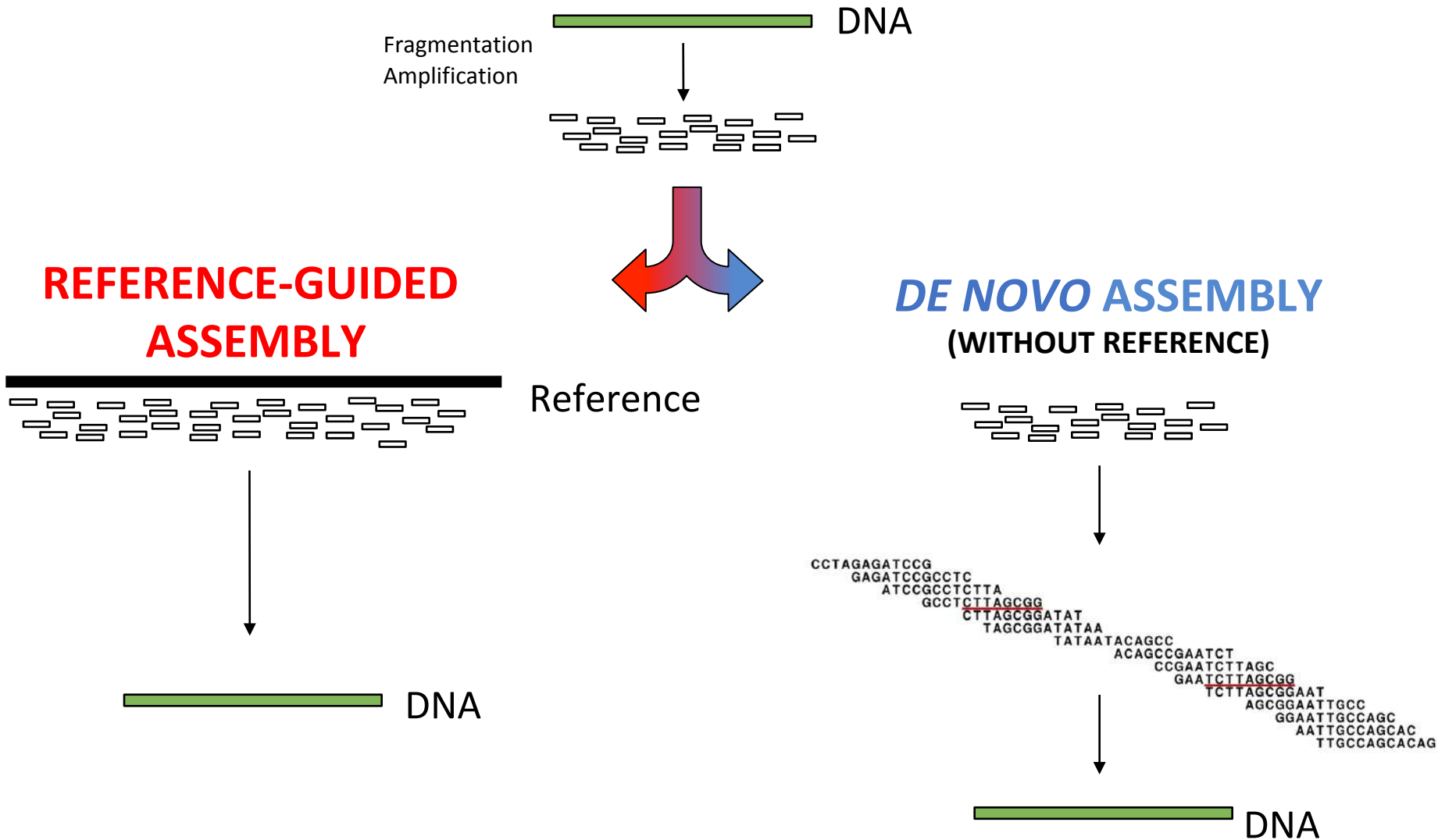
To explain the phylogeographic patterns of *Silene* Sect. *Psammophilae*

## What are the aims of this study?

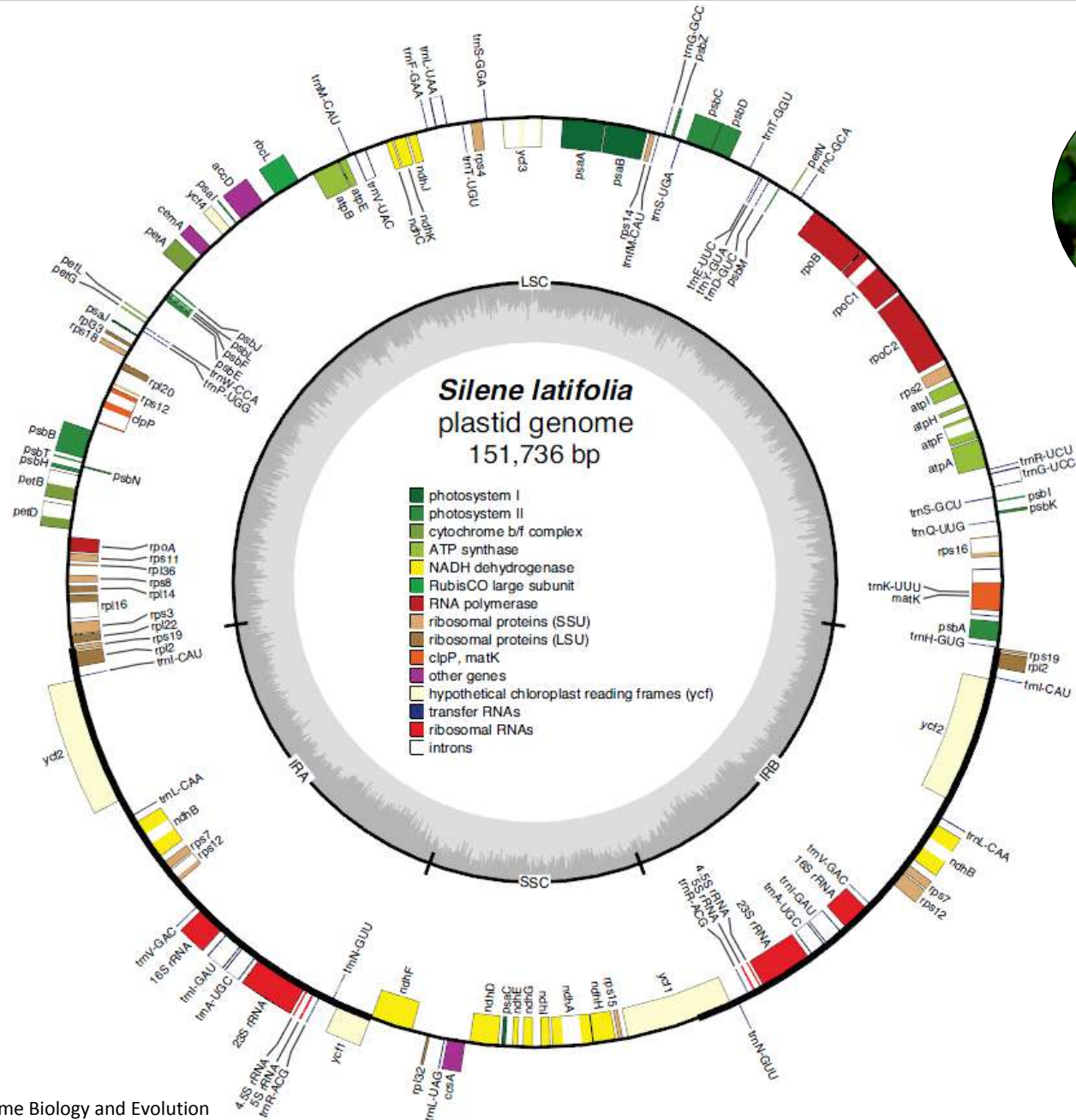
To explain the phylogeographic patterns of *Silene* Sect. *Psammophilae*

How? By using a **genome skimming** approach



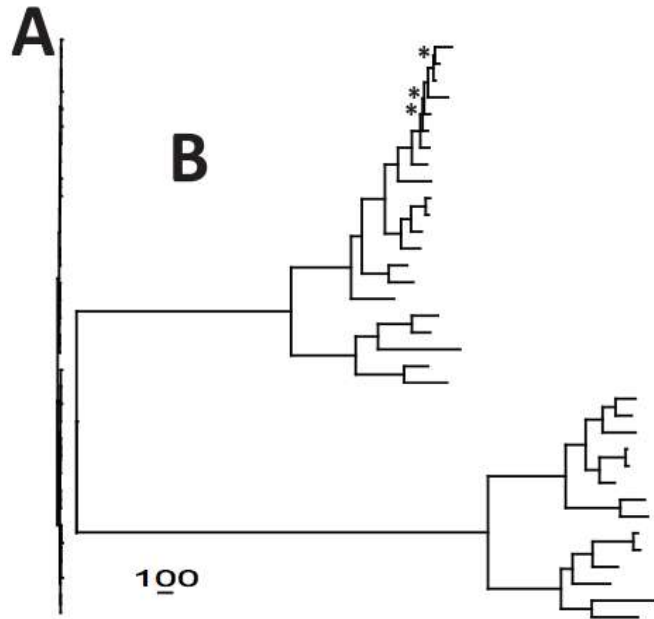


# La secuenciación del plastoma completo en la sección *Psammophilae* (*Silene*) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares



## Why next generation sequencing (NGS)?

## Why next generation sequencing (NGS)?



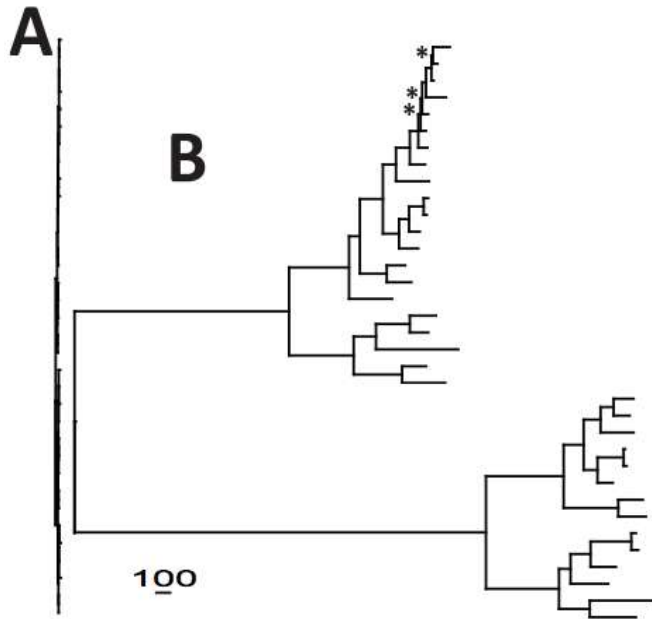
Phylogenetic resolution:

A) 2 loci

B) Whole chloroplast

(increase of ~ 60x phylogenetic informative characters)

## Why next generation sequencing (NGS)?



Parks et al. 2009. BMC Biology



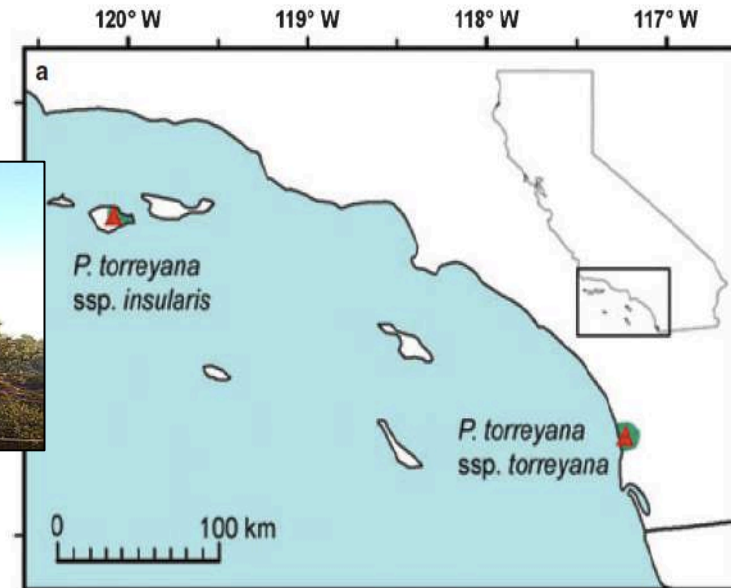
Phylogenetic resolution:

A) 2 loci

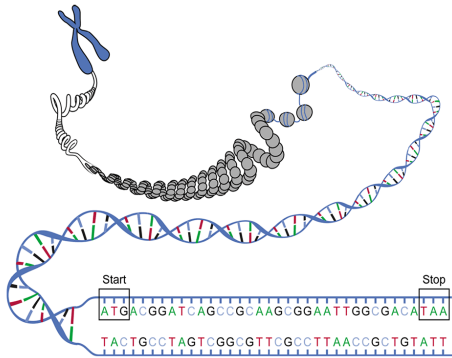
B) Whole chloroplast

(increase of ~ 60x phylogenetic informative characters)

Low level of divergence in *Pinus torreyana*  
(1 polymorphism / ~21 kb)



Whittall et al. 2010 Molecular Ecology



Library  
preparation

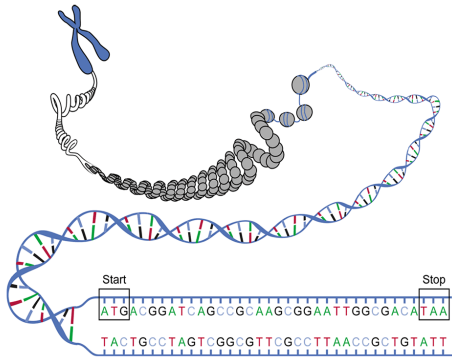


Sequencing

Data

gctacctaag acttcgtcaaa  
acttcgtcaaa acgtaccgtaa  
gctacctaag  
acctaggcctt gctacctaag  
acgtaccgtaa acctaggcctt

Data analysis



Data

```
gctacctaag  actctgtcaaa
actctgtcaaa  acgtaccgtaa
gctacctaag
acctaggcctt  gctacctaag
acgtaccgtaa  acctaggcctt
```

Library  
preparation

Sequencing

Data analysis

## Library preparation and sequencing details:

- DNA extractions: 5 individuals per population
- Nextera Kit w/barcodes
- Sequenced on HiSeq @USC's Epigenome Center (CA)
  - Single-end: 4 samples
  - Paired-end: 22 samples

illumina®



# 1. Phylogenetic reconstruction



# 1. Phylogenetic reconstruction

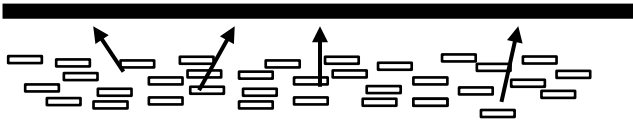


- Trim of raw data

## 1. Phylogenetic reconstruction



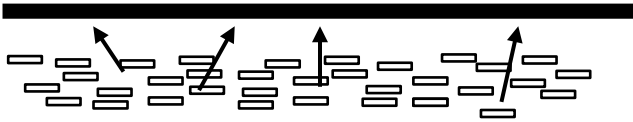
- Trim of raw data
- Reference-guided assembly



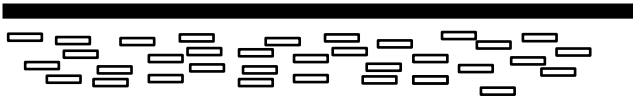


## 1. Phylogenetic reconstruction

- Trim of raw data
- Reference-guided assembly



- Extraction of consensus sequence



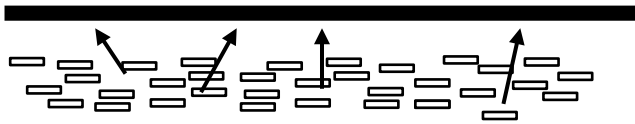






# 1. Phylogenetic reconstruction

- Trim of raw data
- Reference-guided assembly



- Extraction of consensus sequence

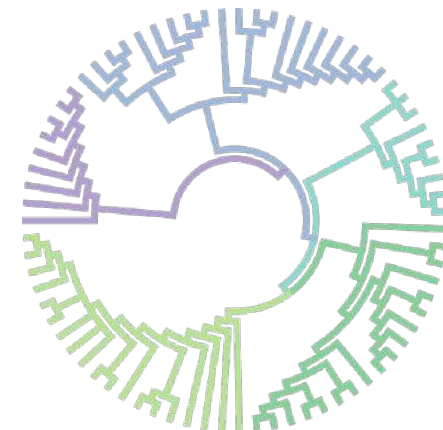
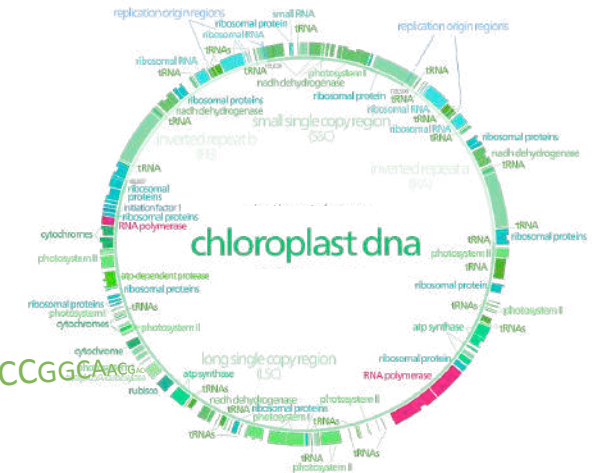


- Alignment (MAFFT)
- Phylogenetic reconstruction

**RAXML**



**Mr Bayes**  
Bayesian Inference  
of Phylogeny

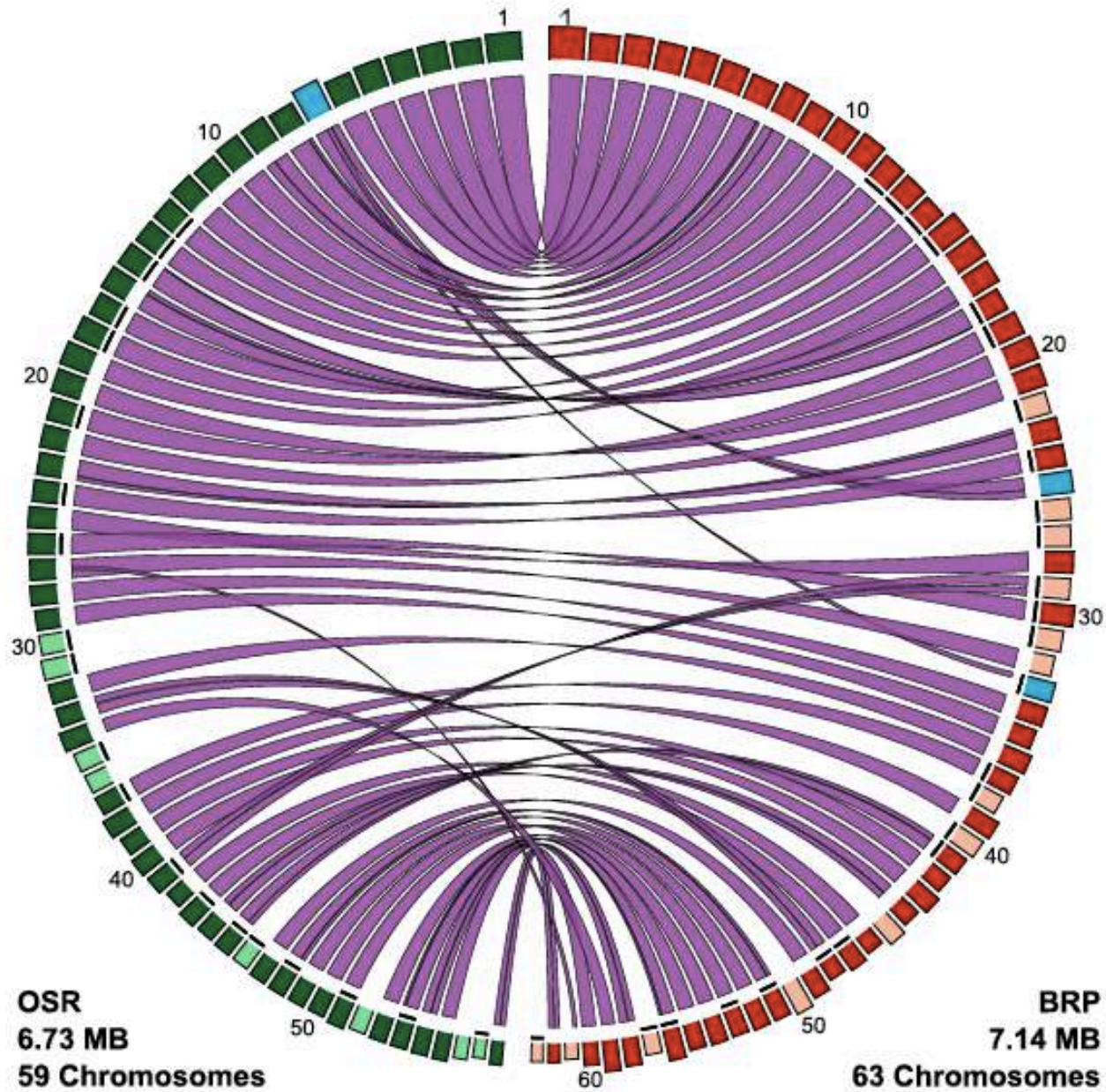


## Sequencing results and cpDNA, mtDNA and nrDNA assemblies' information:

	Average sequencing depth (median)	Assembly length	Recovered sequence	Variable sites
<b>cpDNA</b>	359.91X (153.95X )	154,199 bp	~99%	6,322
<b>mtDNA*</b>	33.4X (10.5X)	254,270 bp	~38%	-
<b>nrDNA</b>	2,765.2X (2,046.7X)	6,415 bp	100%	257

---

***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***



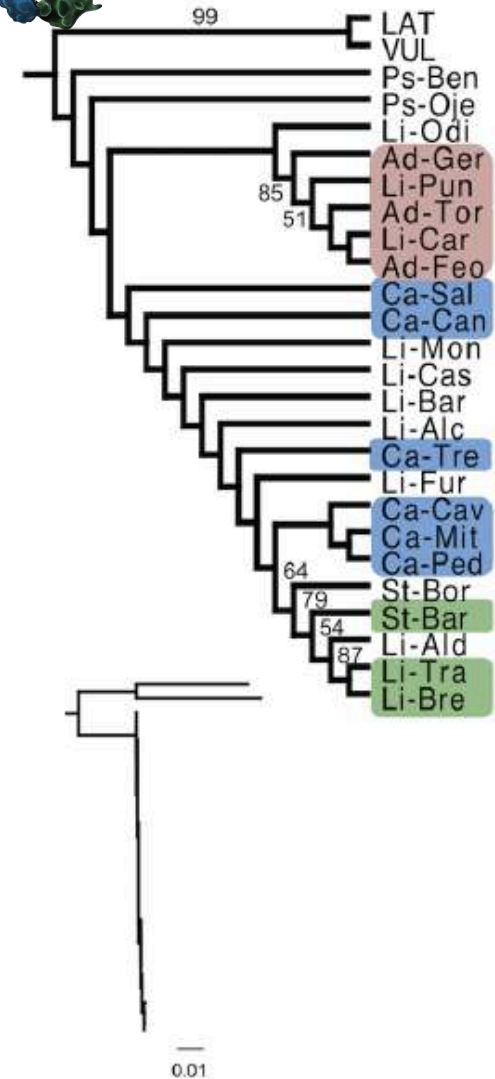
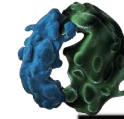
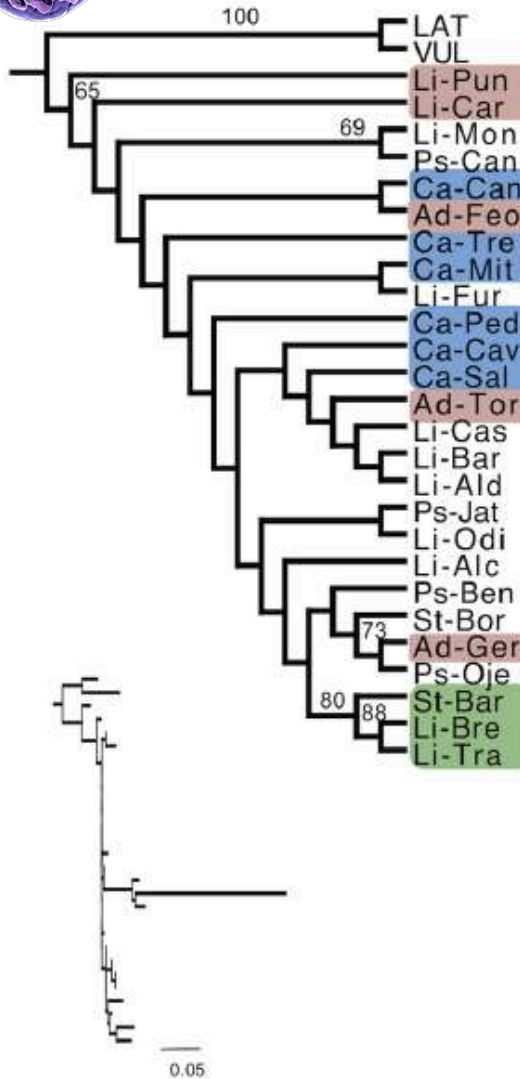
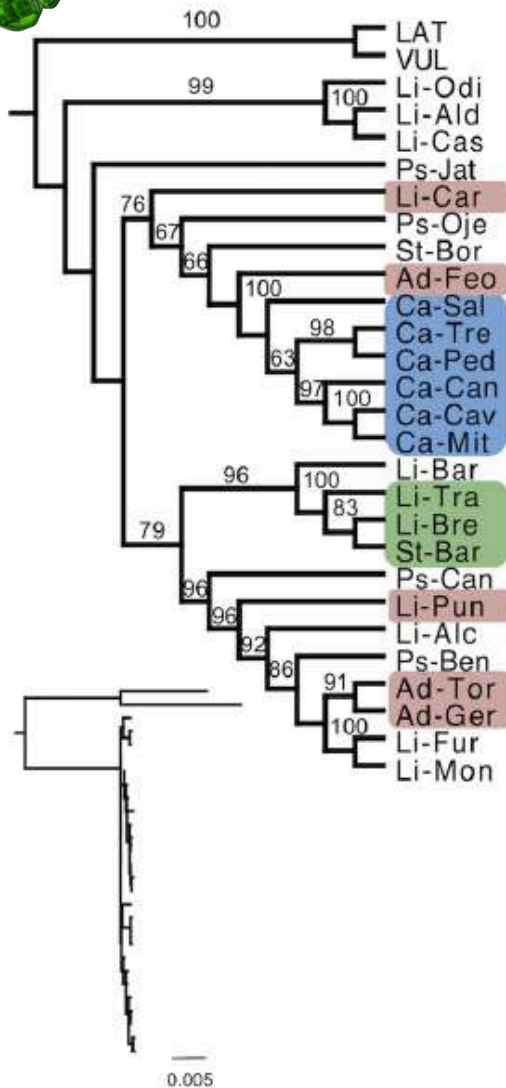
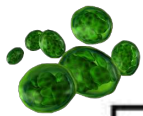
## Sequencing results and cpDNA, mtDNA and nrDNA assemblies' information:

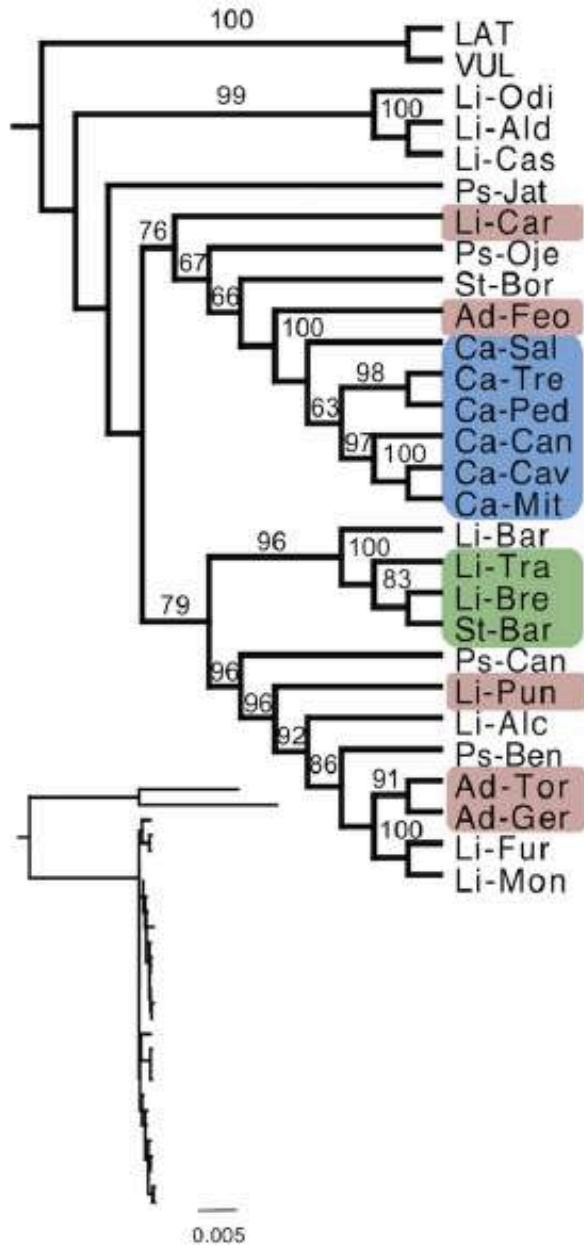
	Average sequencing depth (median)	Assembly length	Recovered sequence	Variable sites
<b>cpDNA</b>	359.91X (153.95X )	154,199 bp	~99%	6,322
<b>mtDNA*</b>	33.4X (10.5X)	254,270 bp	~38%	-
<b>nrDNA</b>	2,765.2X (2,046.7X)	6,415 bp	100%	257

---

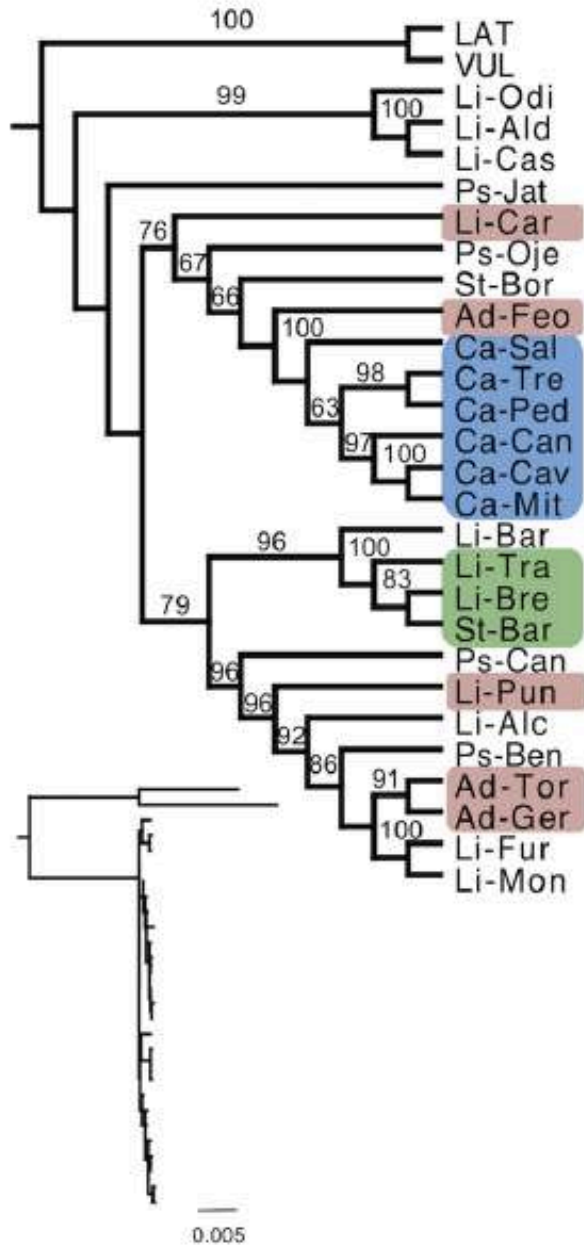
*\* We selected a concatenation of 5,648 bp from six mitochondrial genes (130 variable sites)*

## Phylogenetic results:



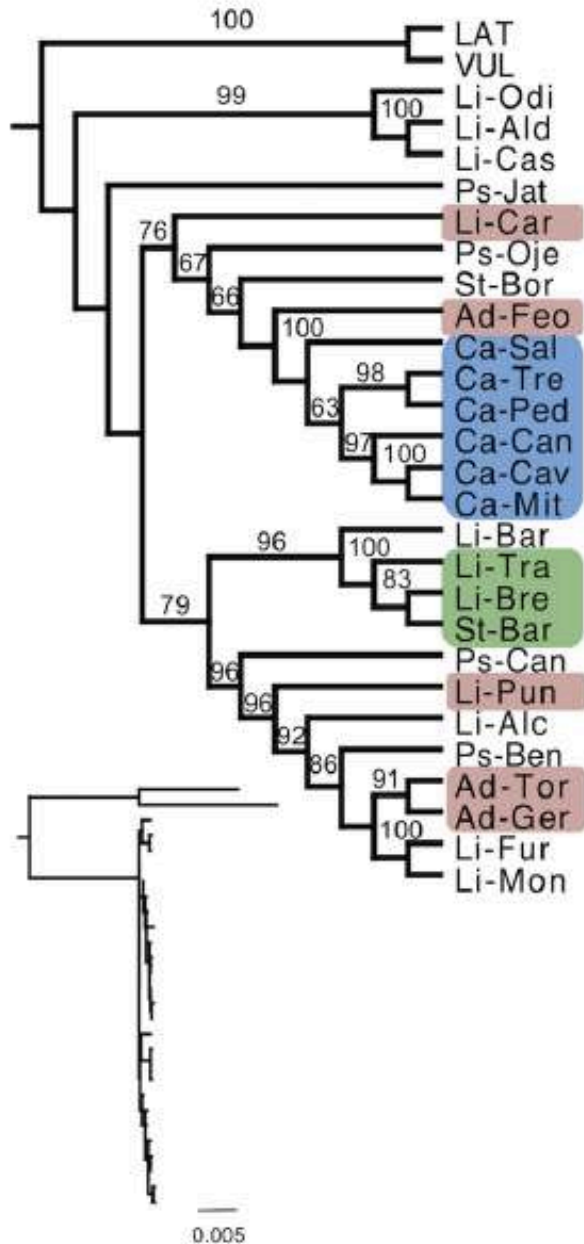


cpDNA:



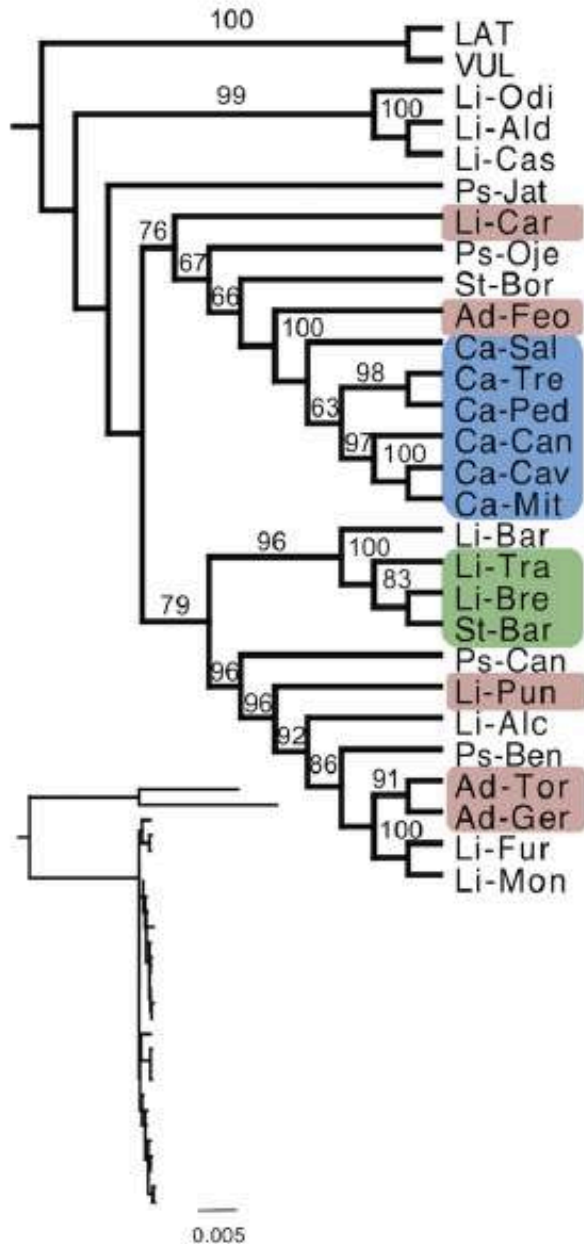
### cpDNA:

- ML and Bayesian analysis showed mostly congruent topologies



cpDNA:

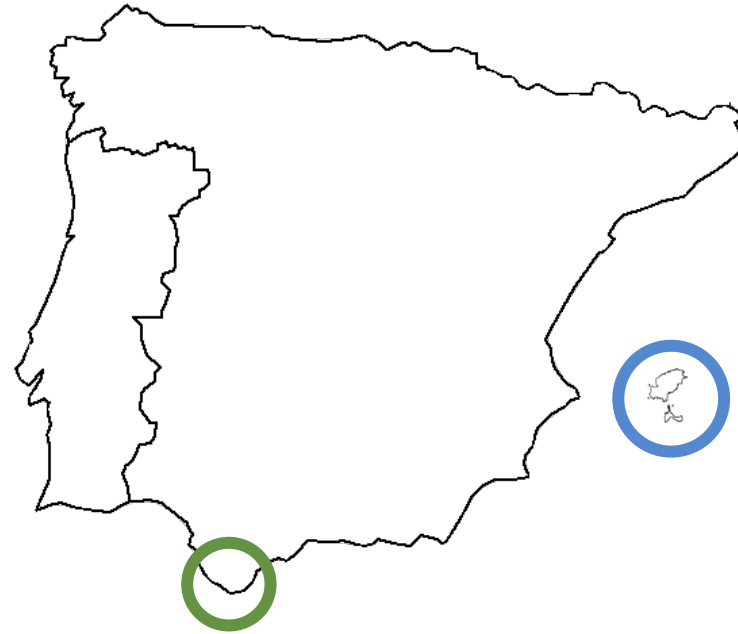
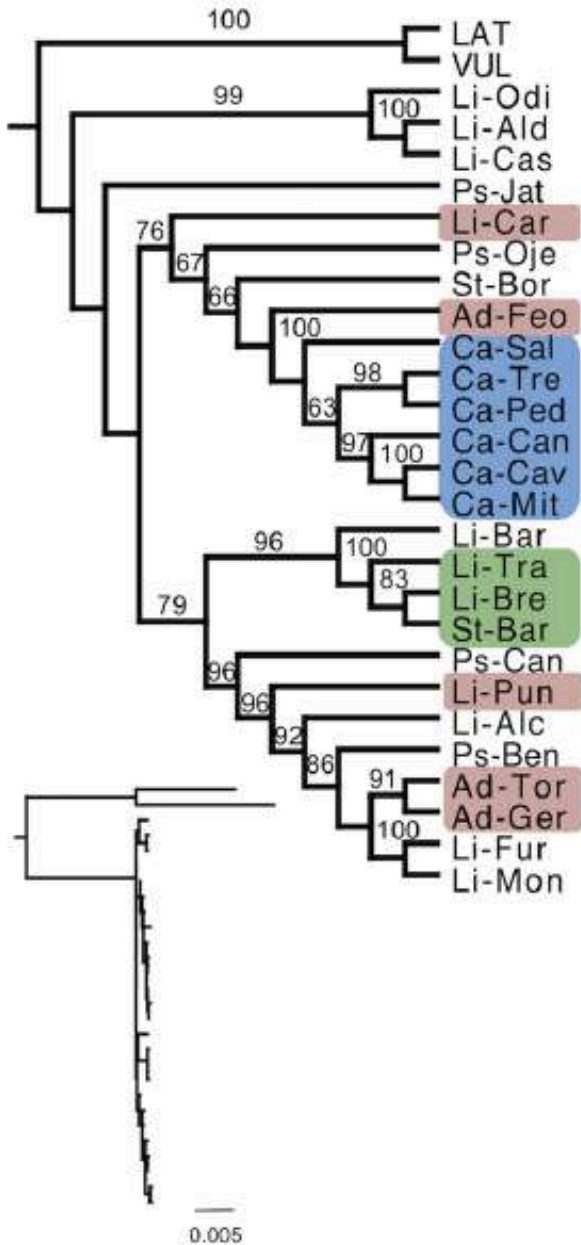
- ML and Bayesian analysis showed mostly congruent topologies
- 17 of 27 (63%) nodes with BS > 70



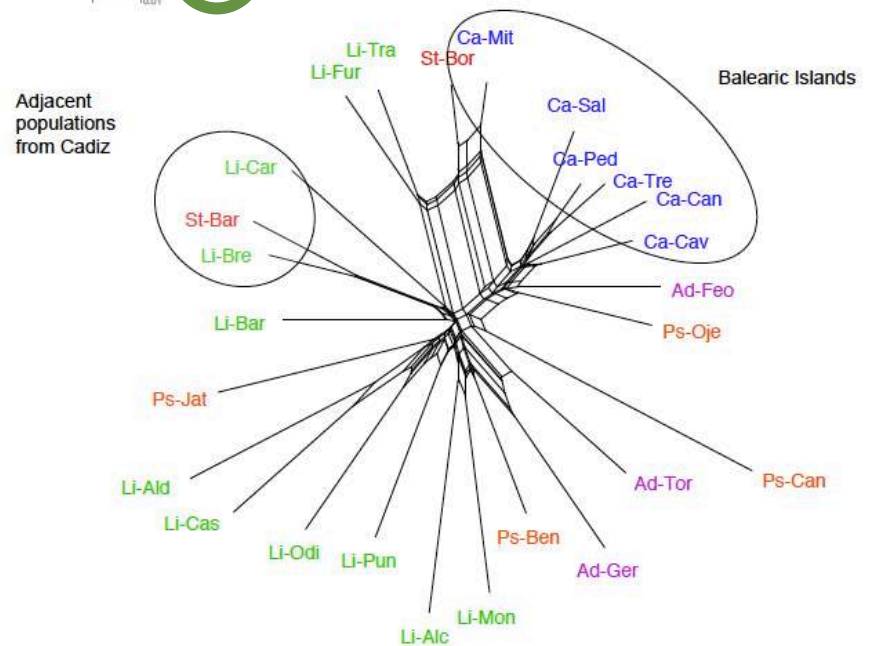
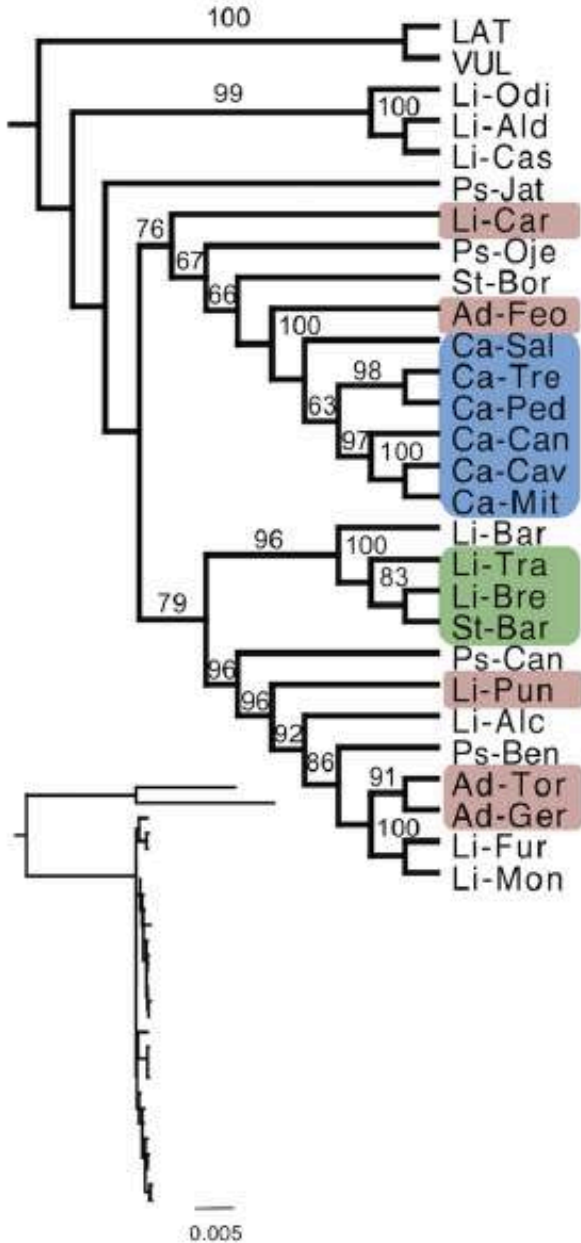
### cpDNA:

- ML and Bayesian analysis showed mostly congruent topologies
- 17 of 27 (63%) nodes with BS > 70
- Very few geographic patterns:
  - Balearic Islands
  - Three populations from Cadiz

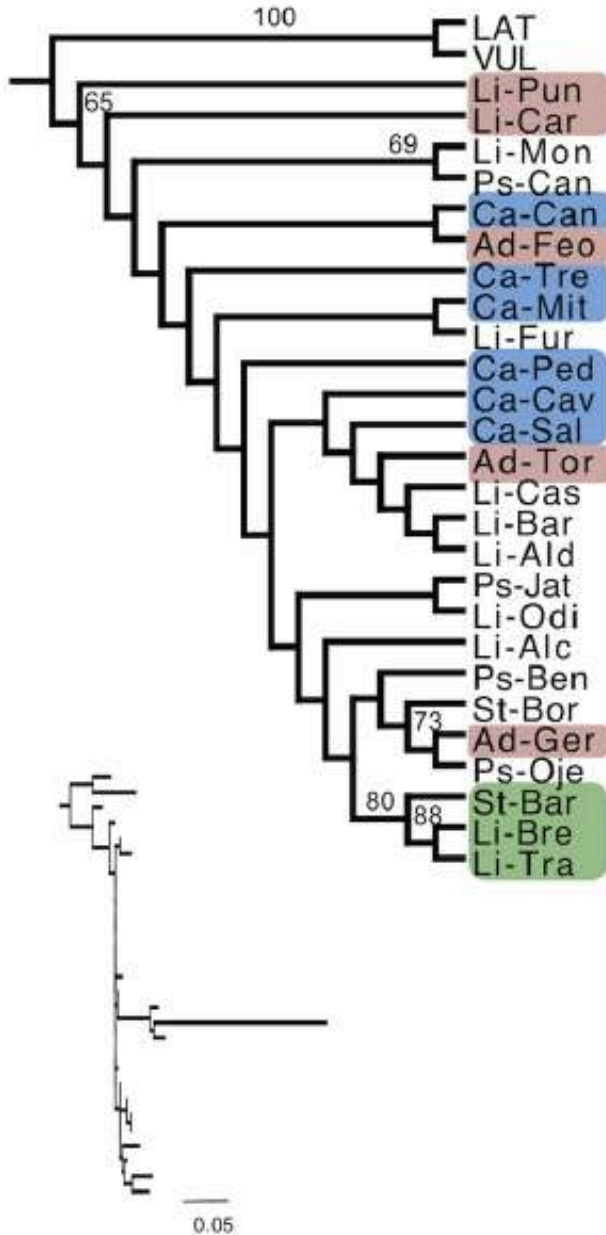
***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***



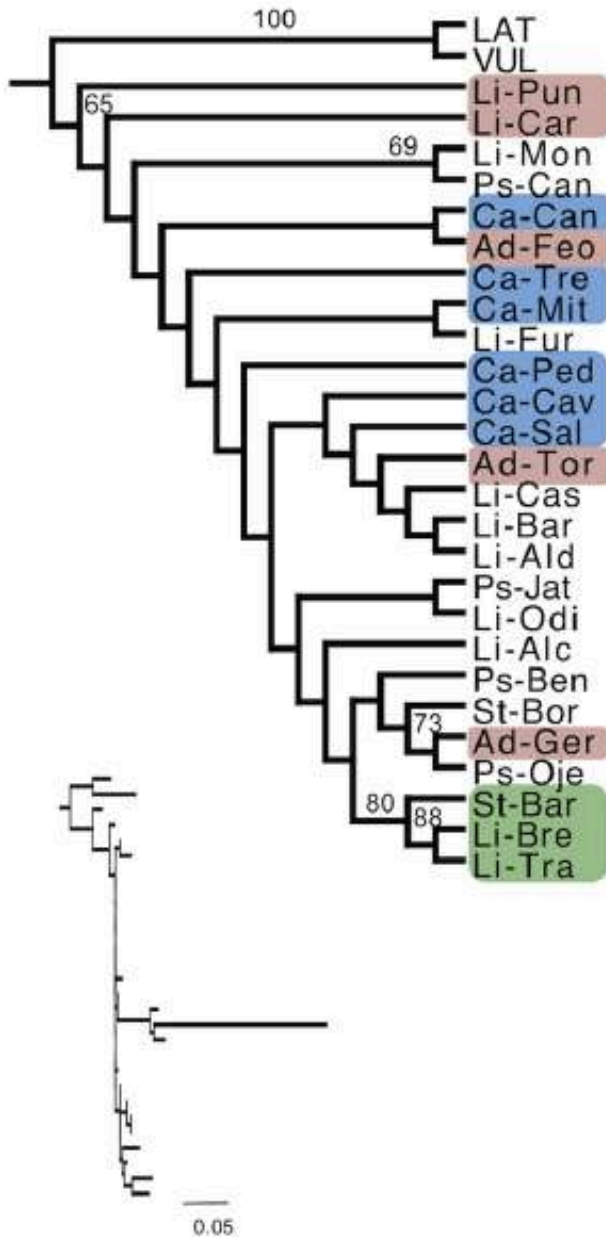
# La secuenciación del plastoma completo en la sección *Psammophilae* (*Silene*) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares



mtDNA:

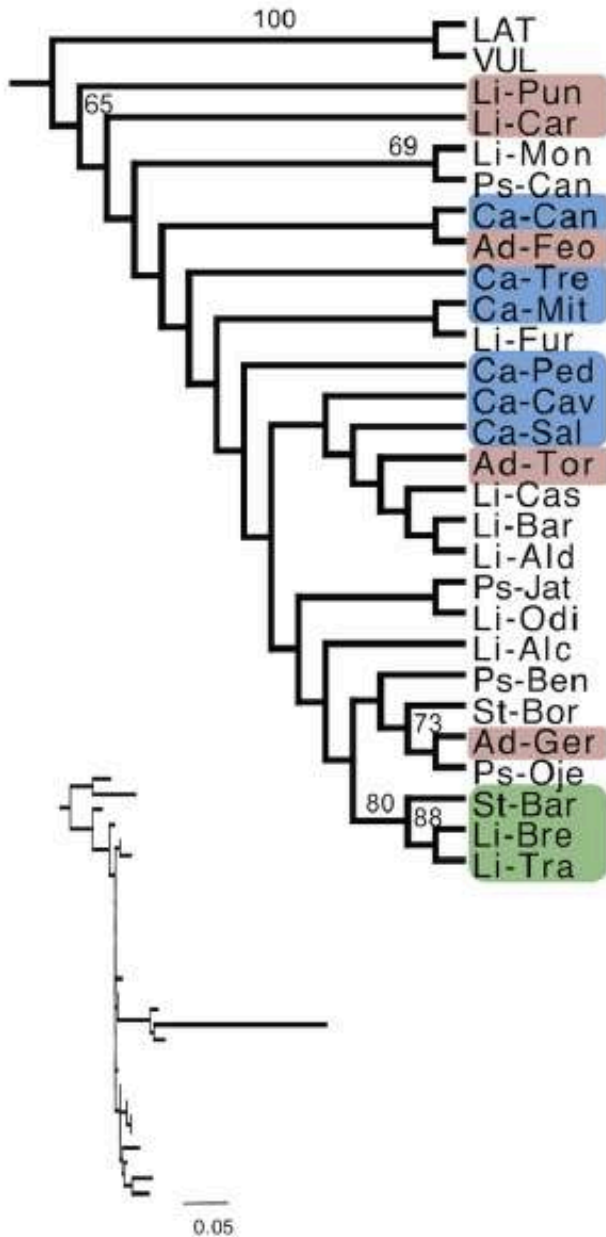






### mtDNA:

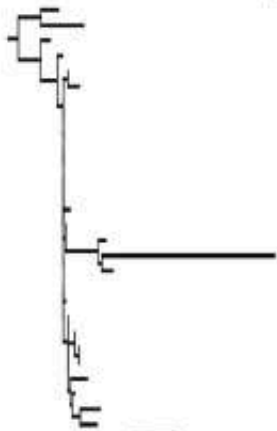
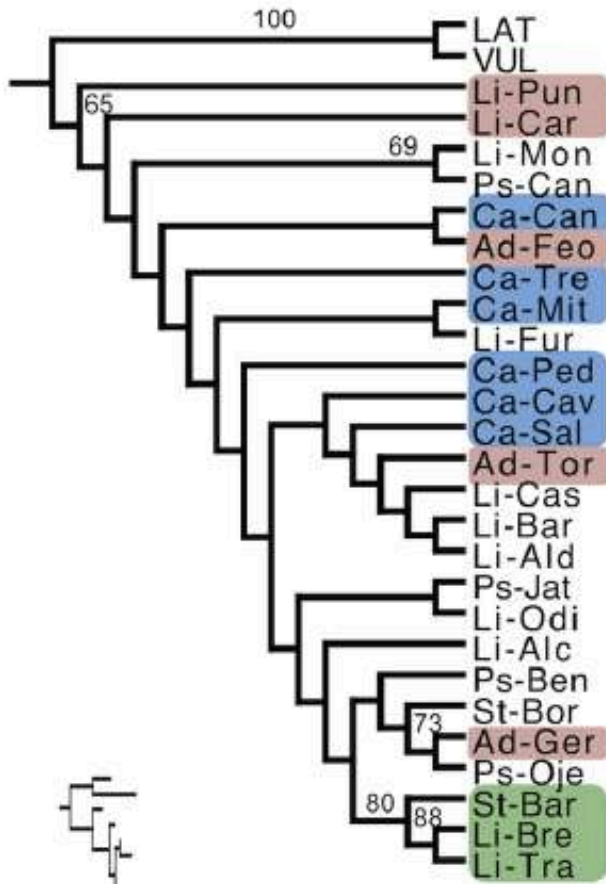
- ML and Bayesian analysis showed mostly congruent topologies
- 4 of 27 (15%) nodes with BS > 70



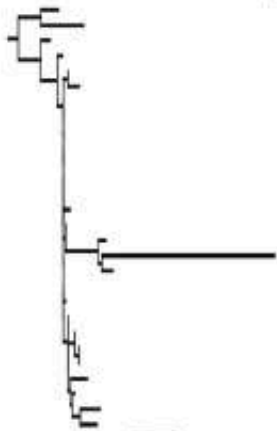
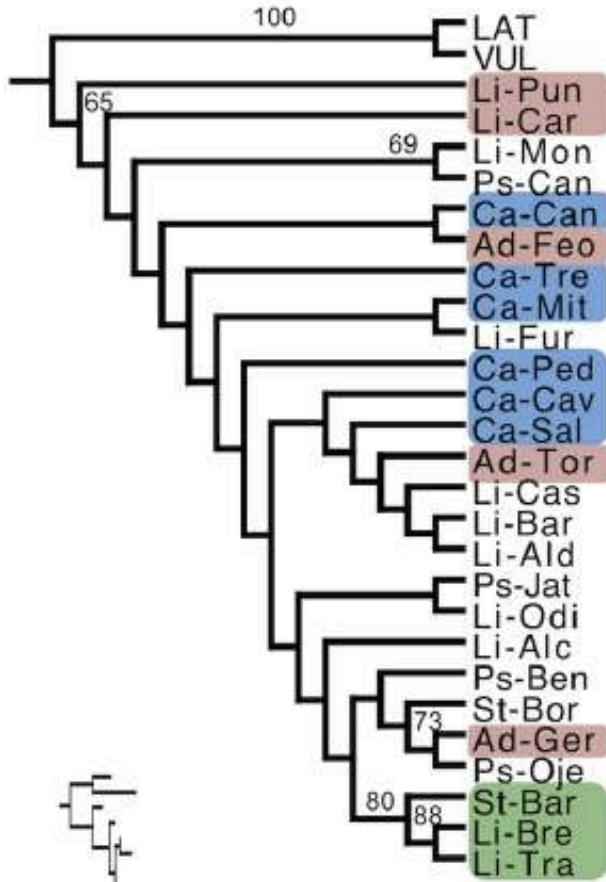
### mtDNA:

- ML and Bayesian analysis showed mostly congruent topologies
- 4 of 27 (15%) nodes with BS > 70
- The mtDNA-based trees did not show any clear phylogeographic pattern, except for populations from Cadiz

***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***

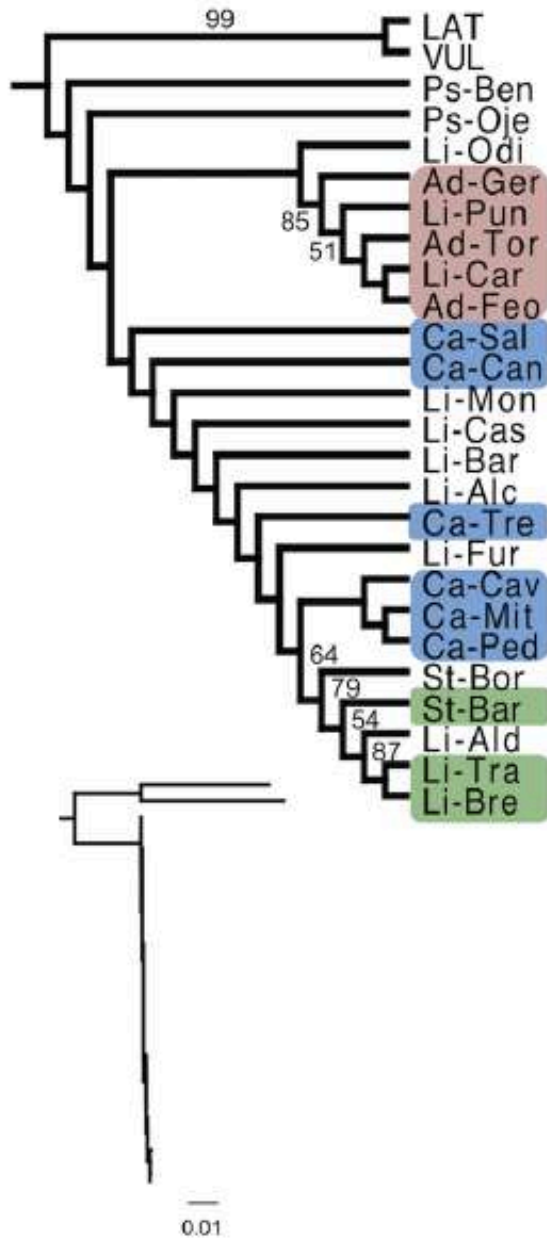


**La secuenciación del plastoma completo en la sección *Psammophilae* (*Silene*) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares**

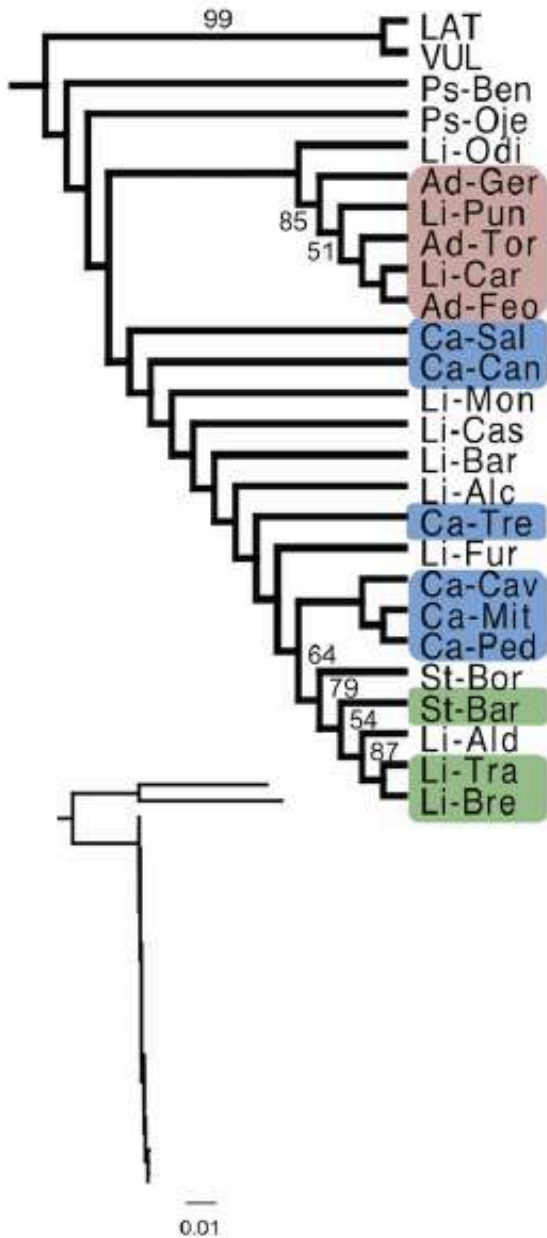


0.05



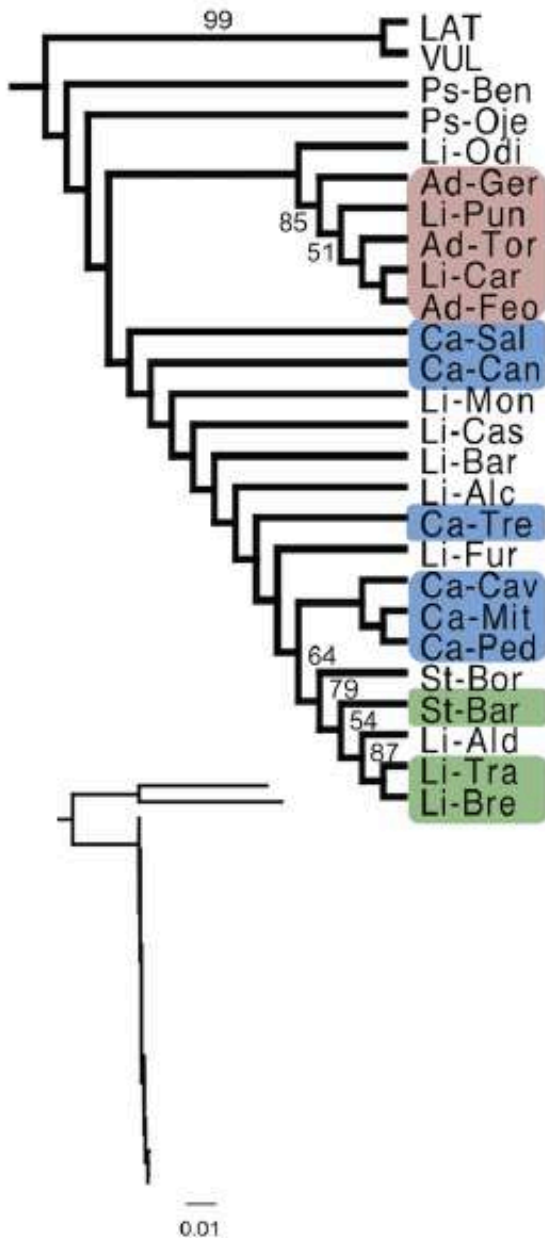


nrDNA:



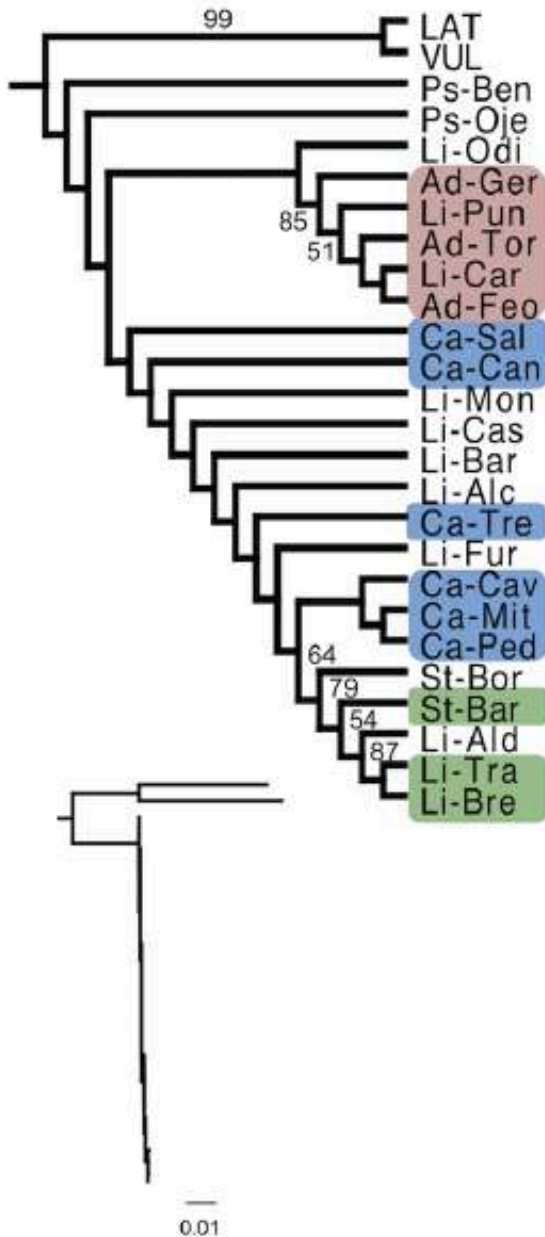
nrDNA:

- ML and Bayesian analysis showed mostly congruent topologies



nrDNA:

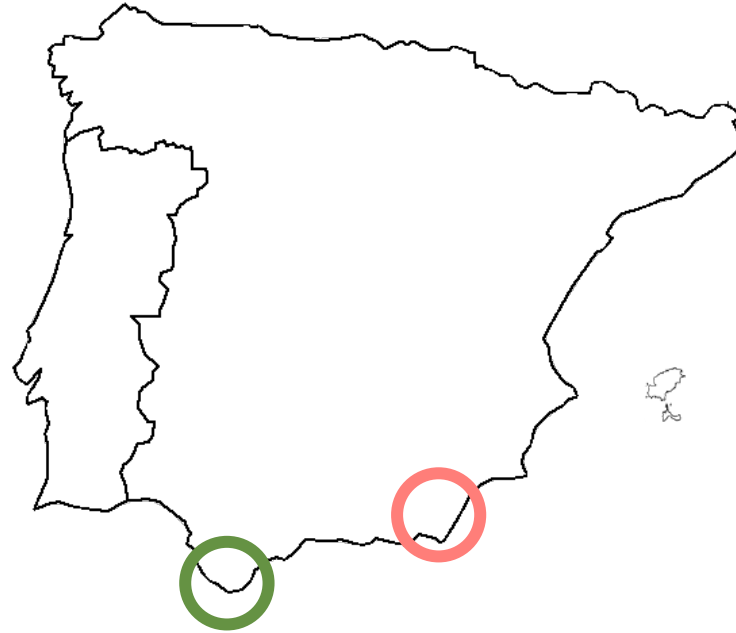
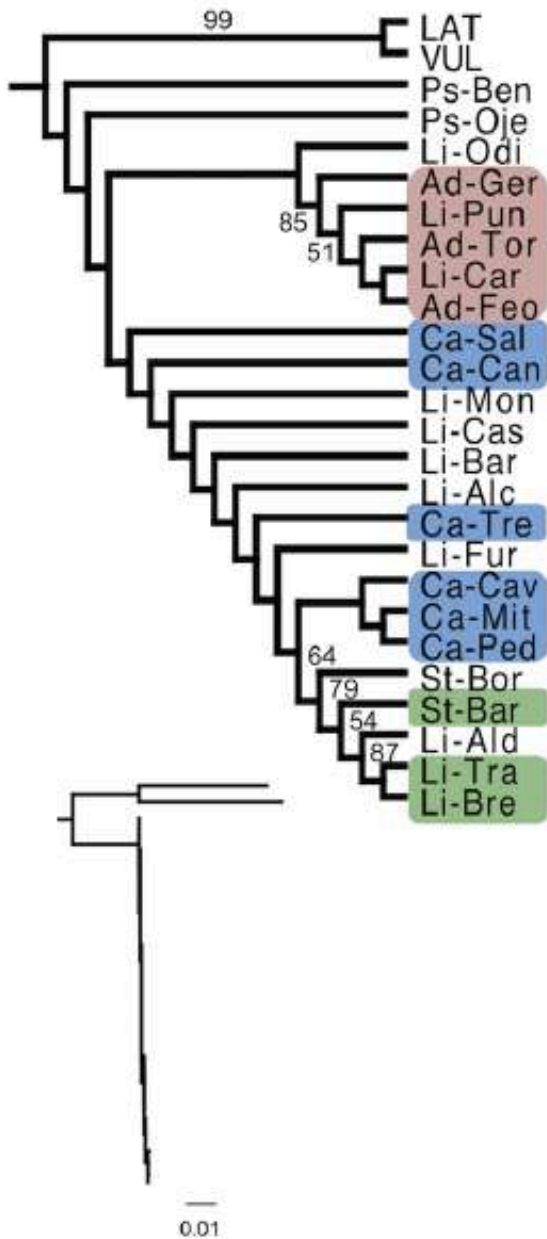
- ML and Bayesian analysis showed mostly congruent topologies
- 4 of 27 (15%) nodes with BS > 70



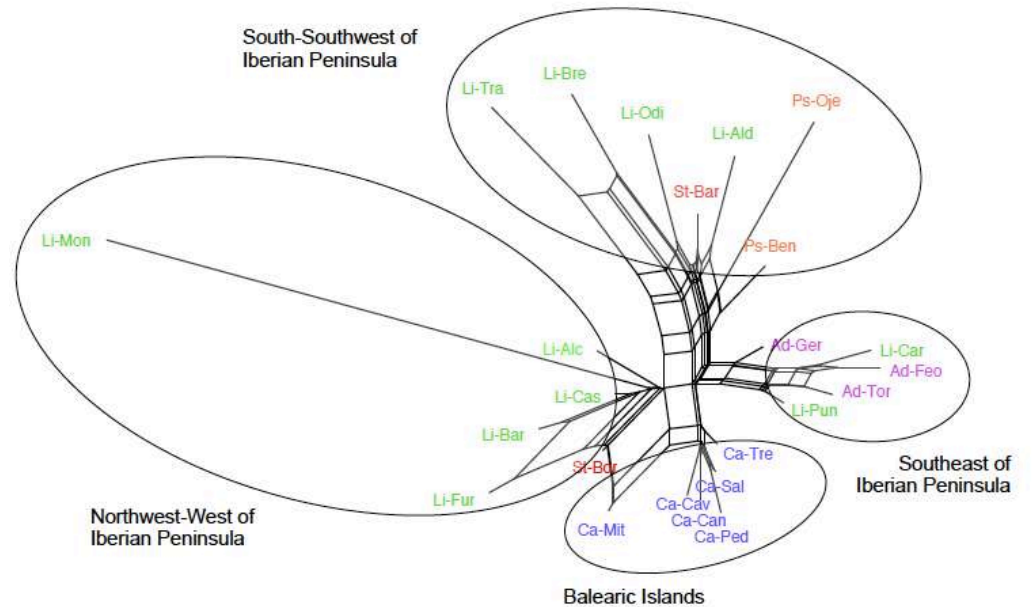
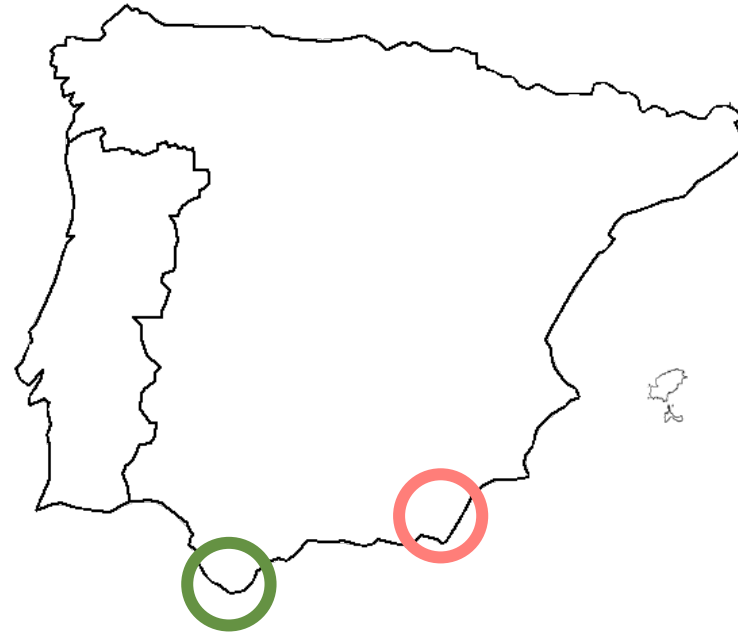
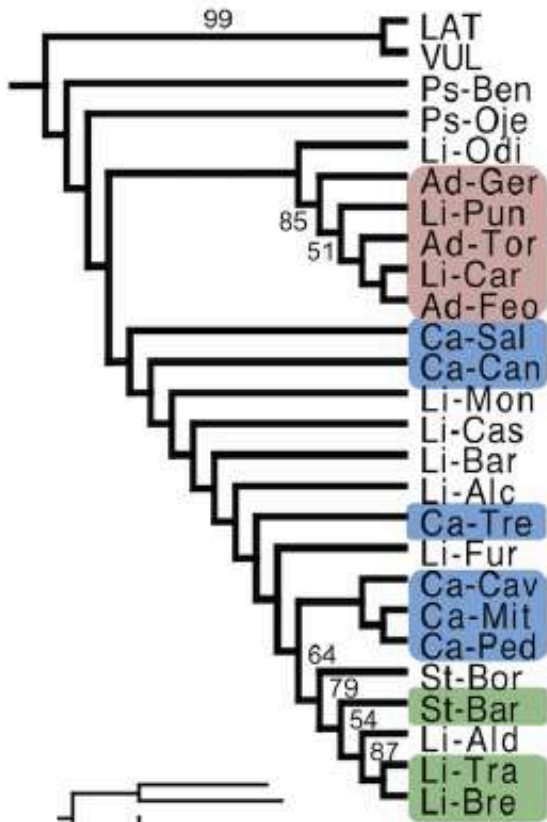
### nrDNA:

- ML and Bayesian analysis showed mostly congruent topologies
- 4 of 27 (15%) nodes with BS > 70
- Very few geographic patterns:
  - Five populations from Almeria
  - Three populations from Cadiz + one population from Malaga

# *La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares*



***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***



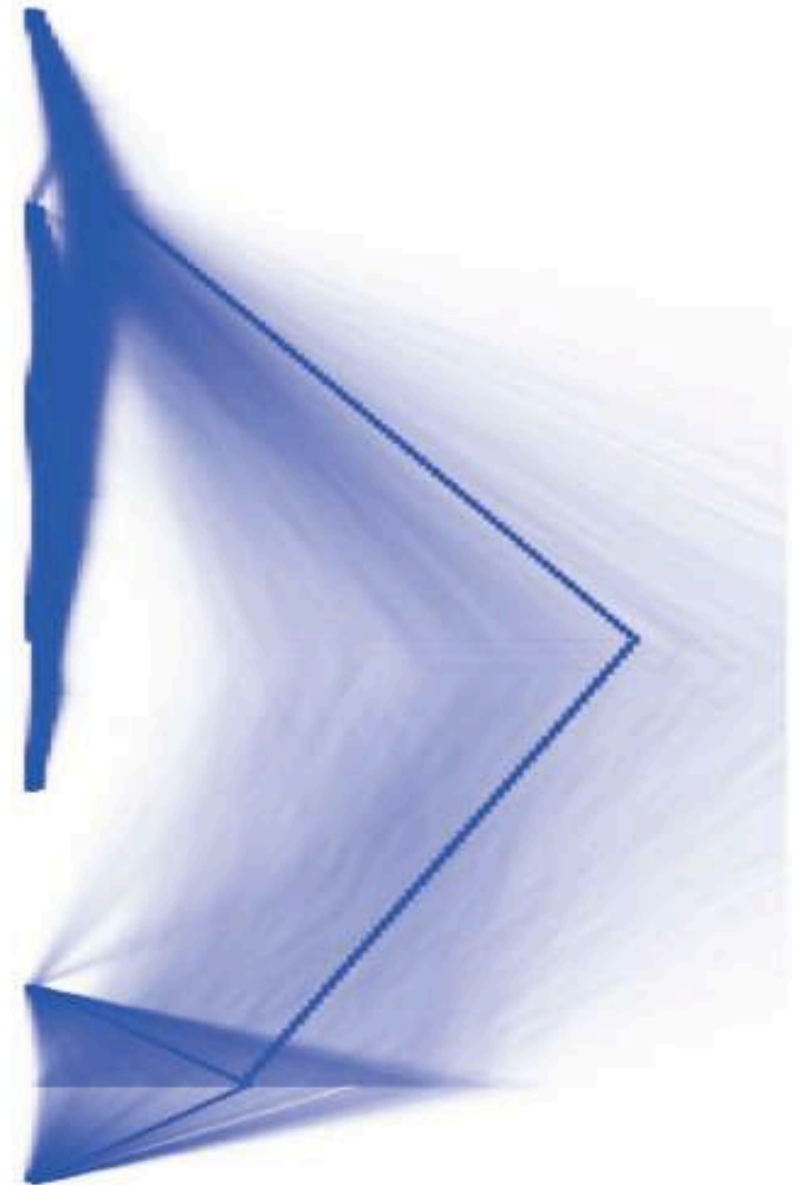
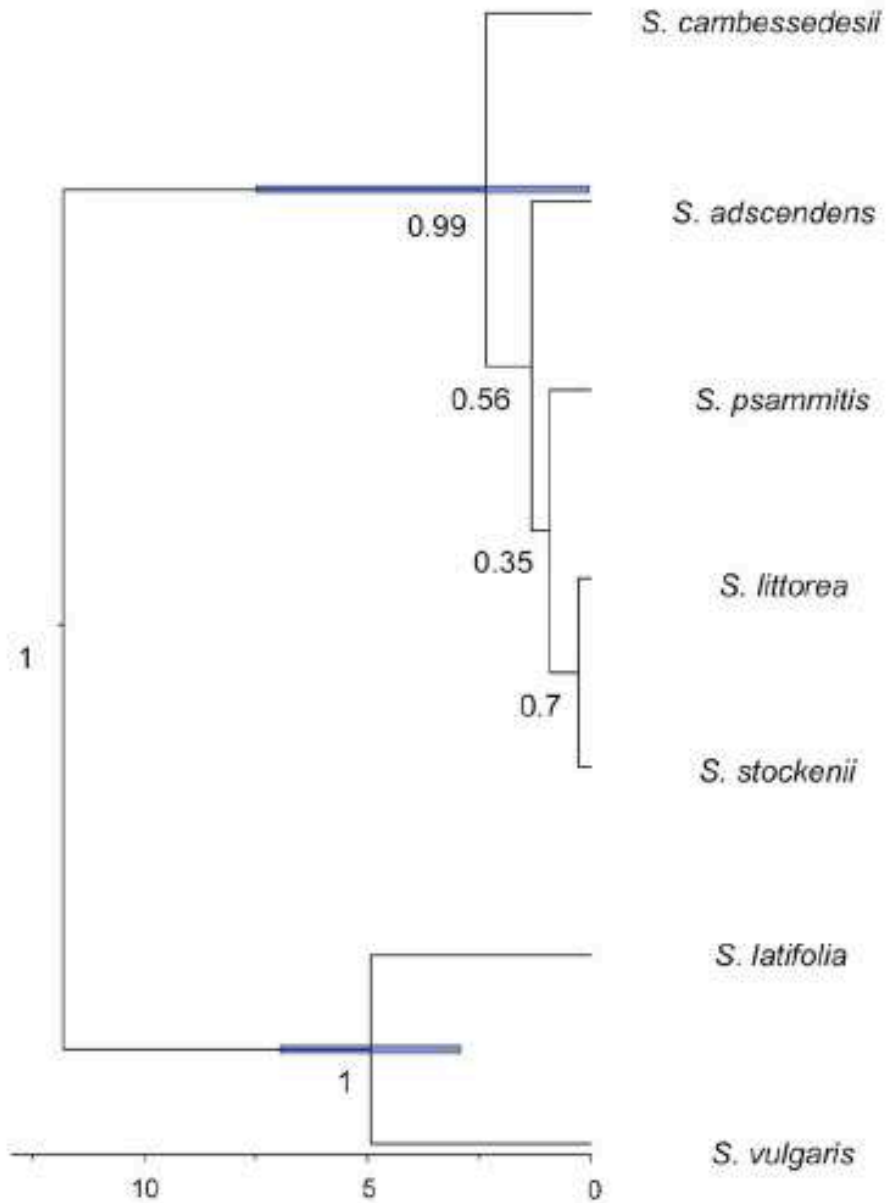
0.01

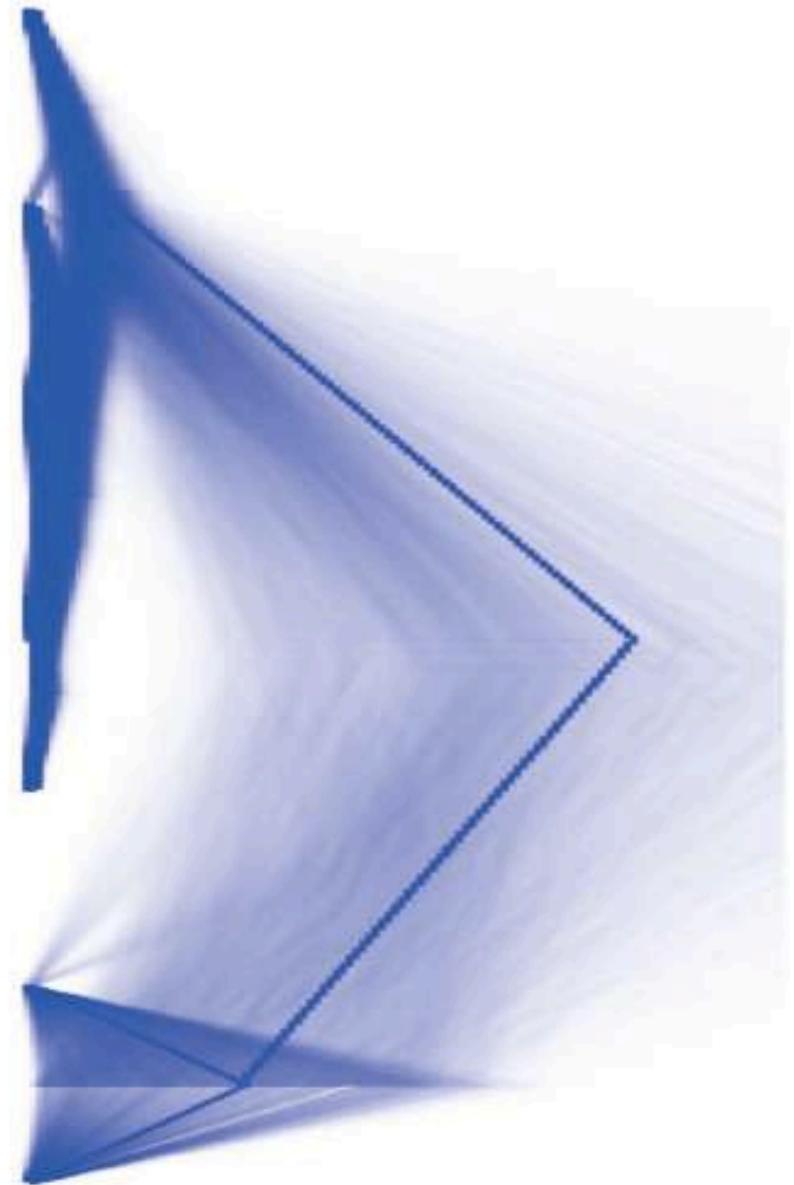
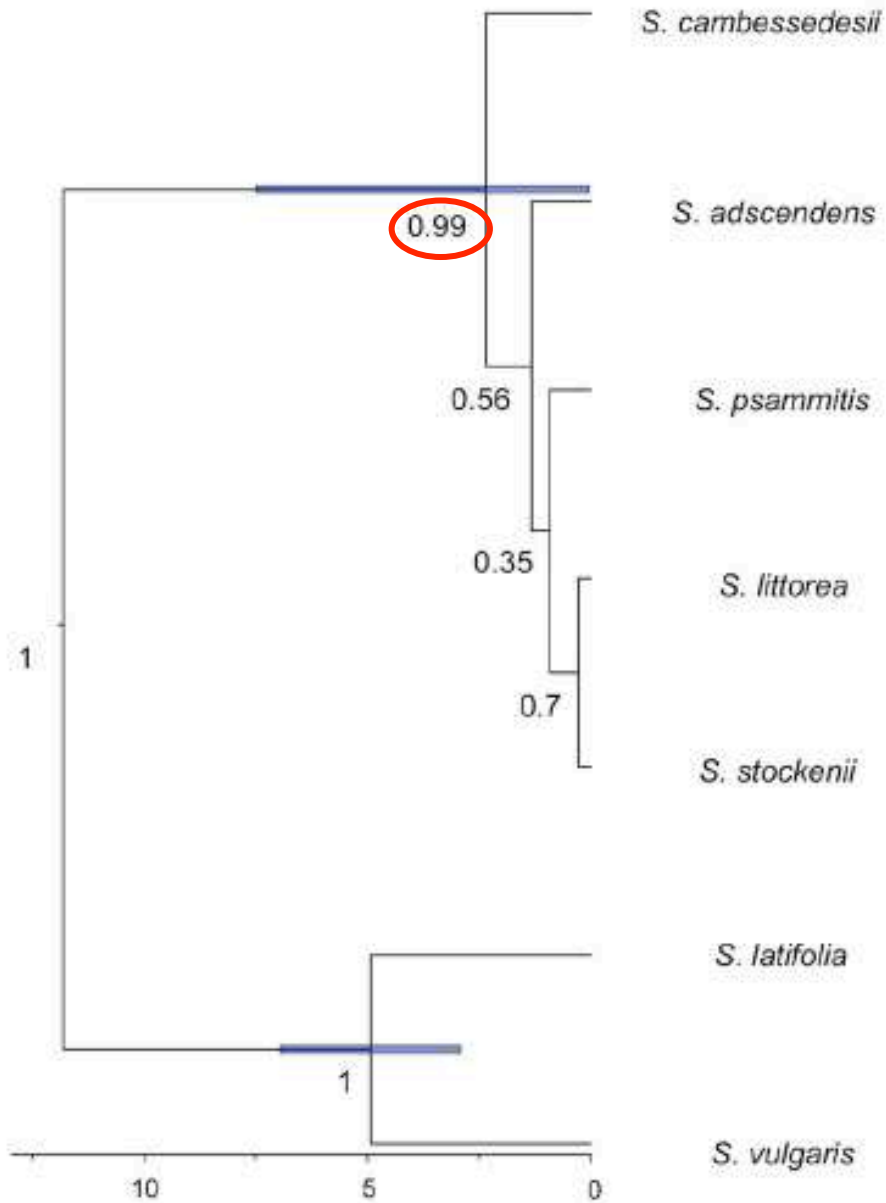
1. Phylogenetic reconstruction

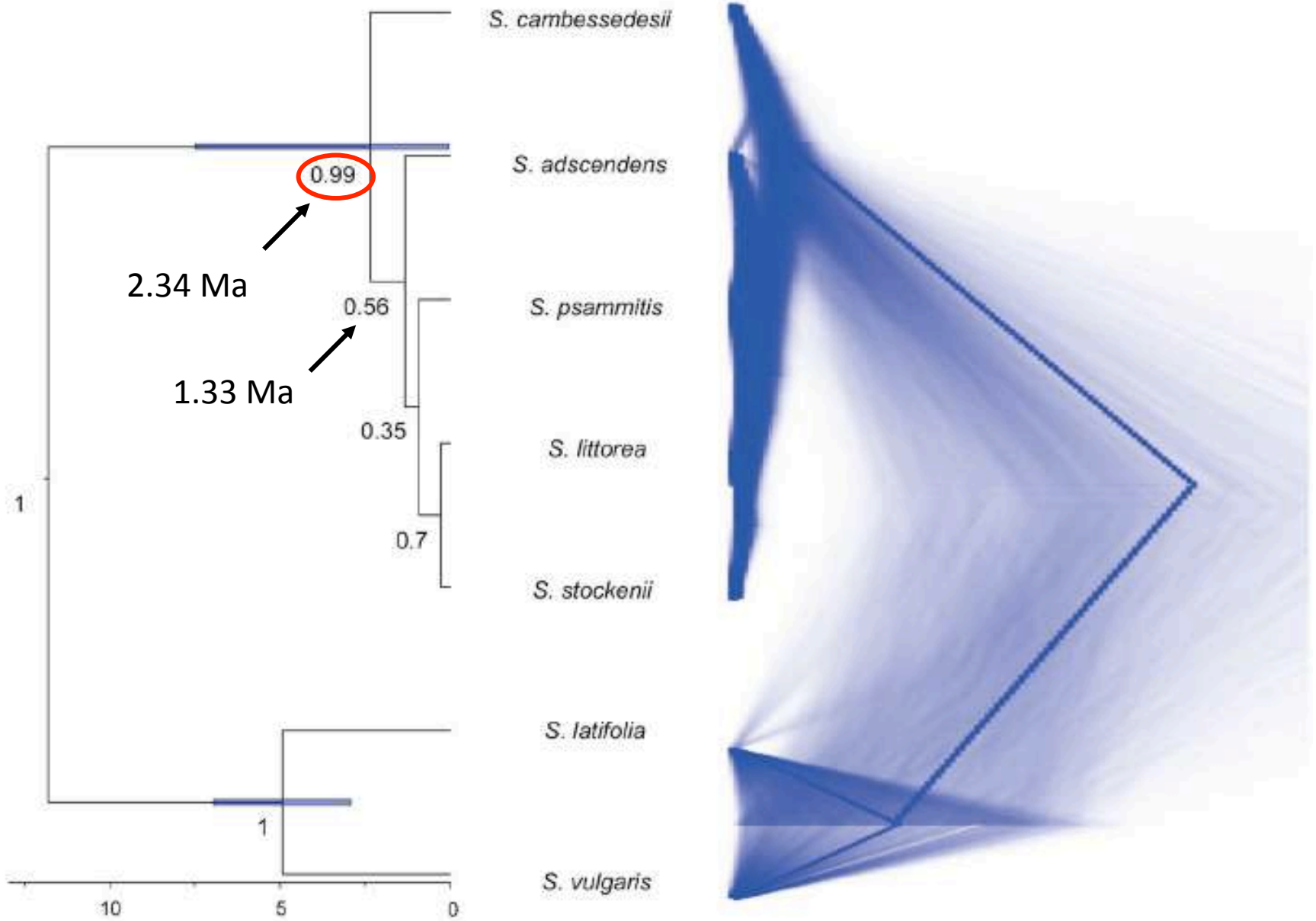
geneious

2. Gene tree-Species tree reconciliation



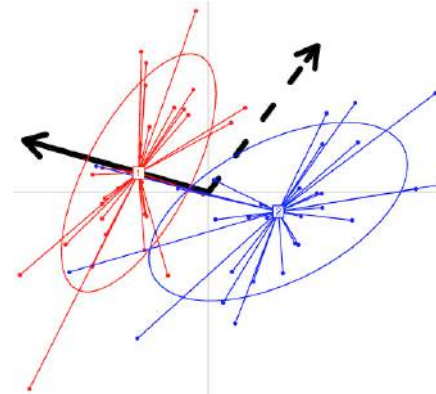






1. Phylogenetic reconstruction
2. Gene tree-Species tree reconciliation
3. Population genetic analyses

geneious



## BAPS: Population structuring

## BAPS: Population structuring

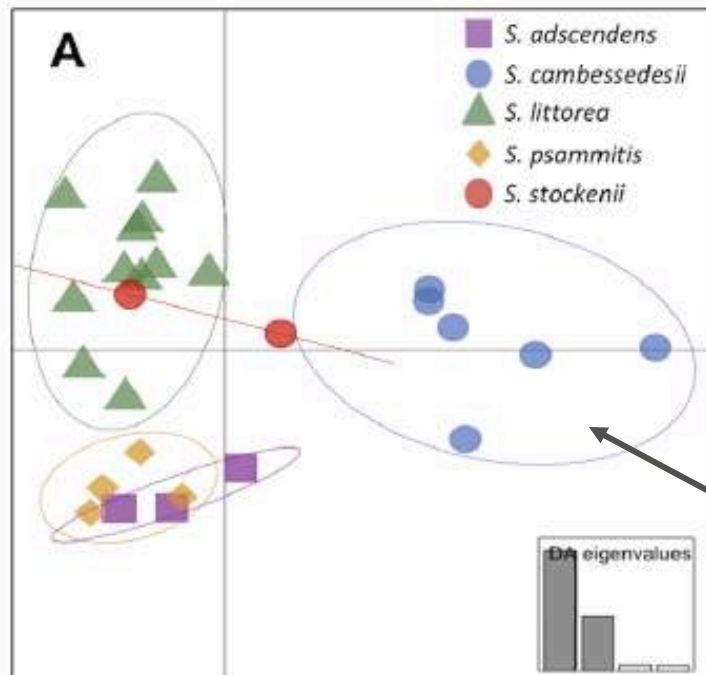
There is just one genetic cluster for the cpDNA data ( $K_1 = -37,598.87$ ,  $K_2 = -39,489.59$ ,  $K_3 = -39,576.70$ ,  $K_4 = -40,069.07$ )

## BAPS: Population structuring

There is just one genetic cluster for the cpDNA data ( $K_1 = -37,598.87$ ,  $K_2 = -39,489.59$ ,  $K_3 = -39,576.70$ ,  $K_4 = -40,069.07$ )

## Discriminant Analysis of Principal Components (DAPC)

Species as priors:



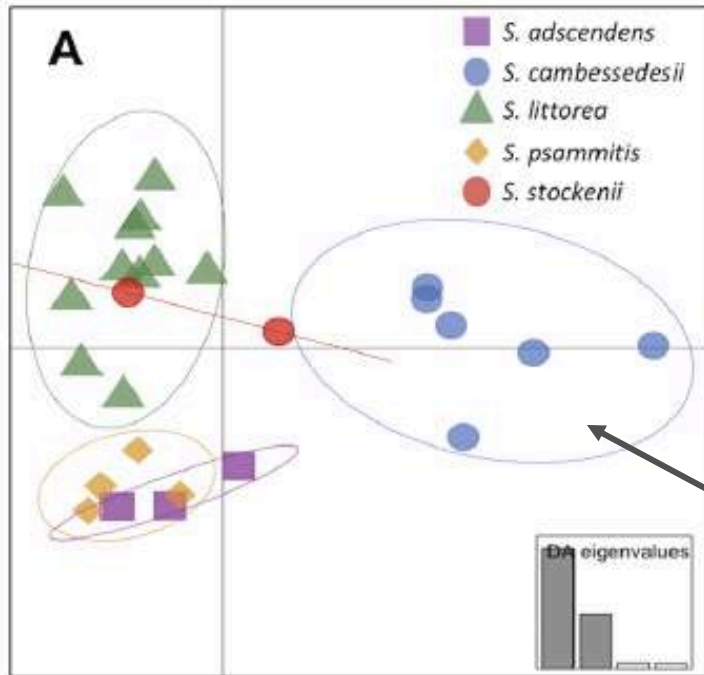
Balearic populations

## BAPS: Population structuring

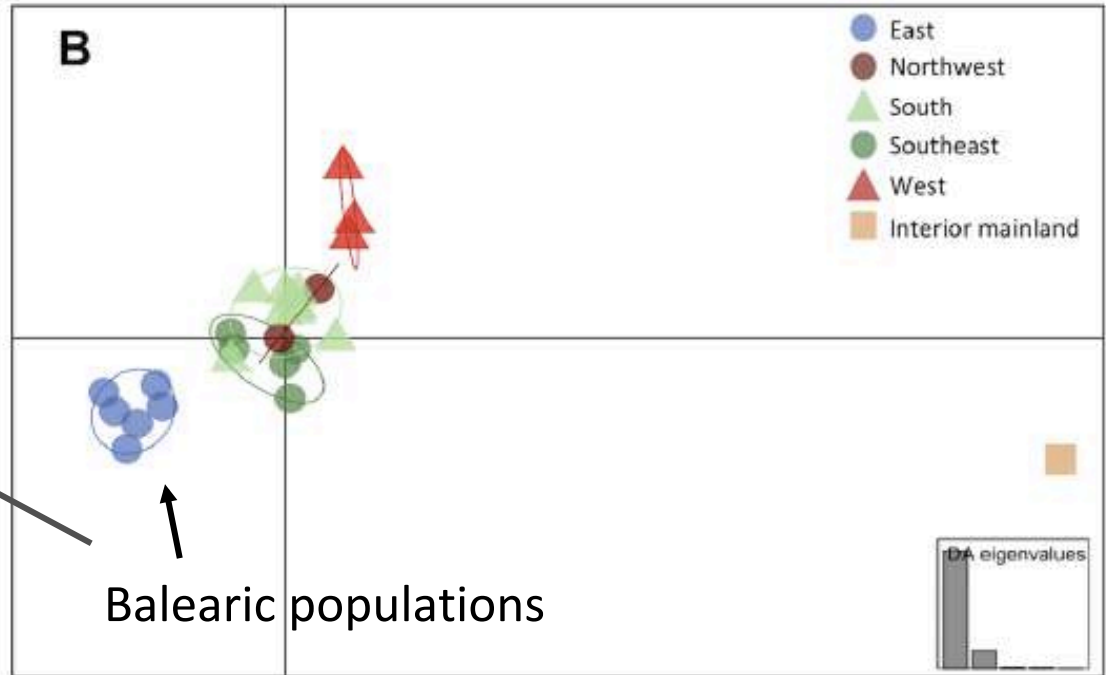
There is just one genetic cluster for the cpDNA data ( $K_1 = -37,598.87$ ,  $K_2 = -39,489.59$ ,  $K_3 = -39,576.70$ ,  $K_4 = -40,069.07$ )

## Discriminant Analysis of Principal Components (DAPC)

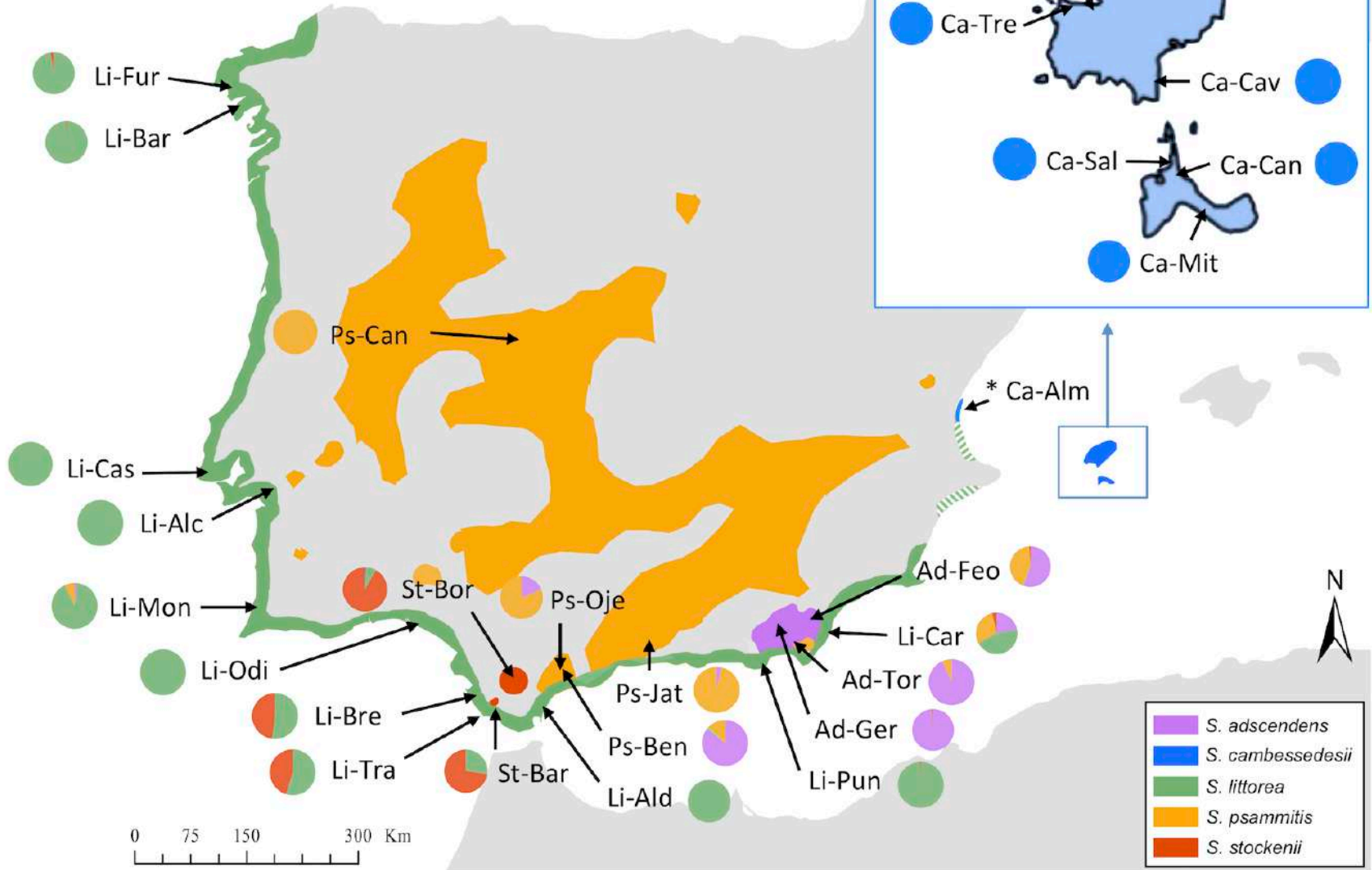
Species as priors:



Location as priors:



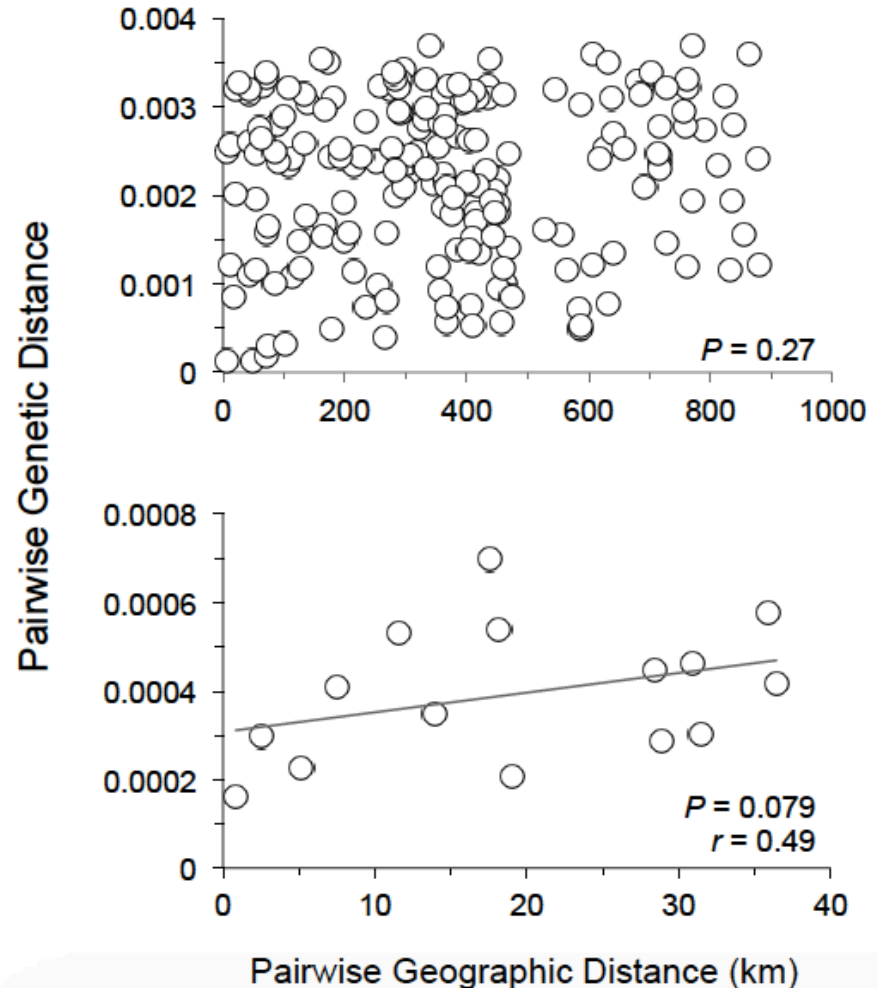
DAPC results when using species as priors:



Partial Mantel test: correlation between genetic variation and geographic distance

## Partial Mantel test: correlation between genetic variation and geographic distance

- All populations:  $P < 0.001$ ,  $r = 0.35$
- The correlation was not significant when analyses were conducted separately

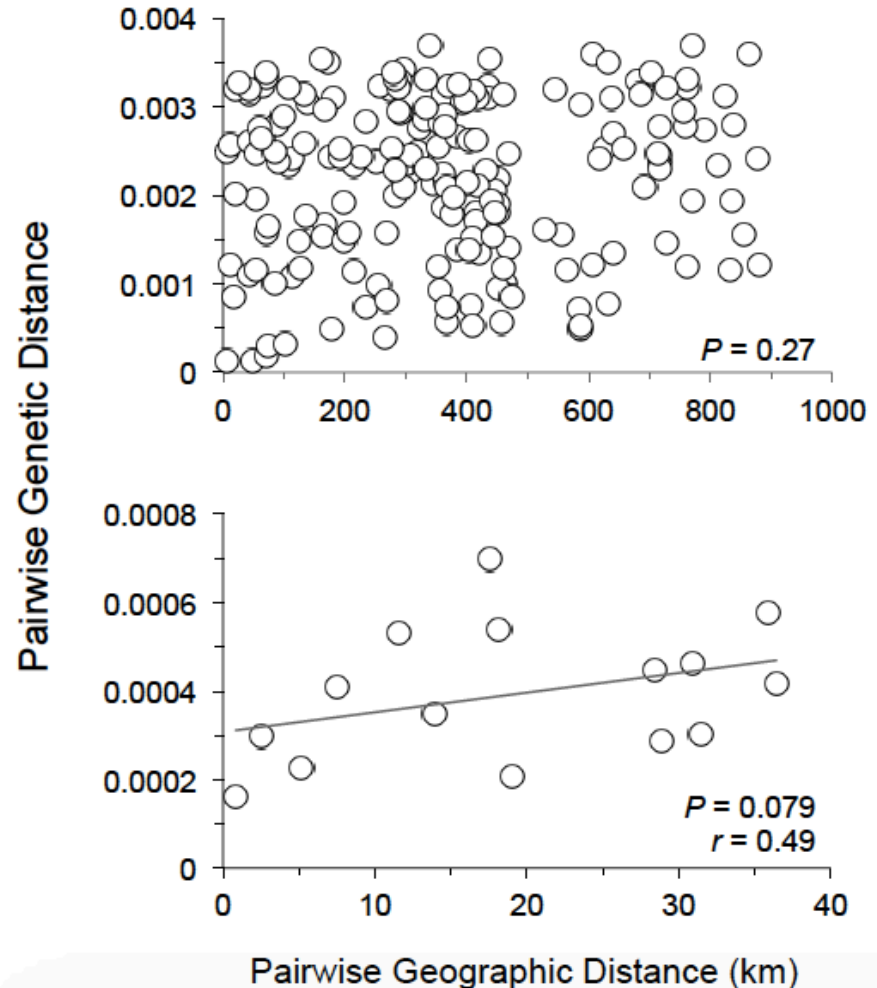


## Partial Mantel test: correlation between genetic variation and geographic distance

- All populations:  $P < 0.001$ ,  $r = 0.35$
- The correlation was not significant when analyses were conducted separately

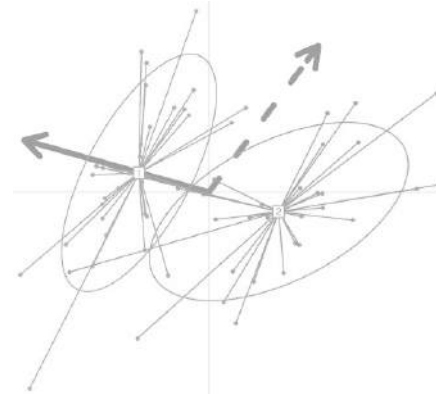
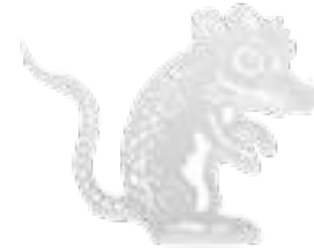
## AMOVA: genetic differentiation among species

- All populations:  $F_{ST} = 0.23$ ;  $P < 0.001$ . Most of the genetic variation was concentrated within species (77.3%)
- No genetic differentiation neither among mainland species nor among Balearic islands

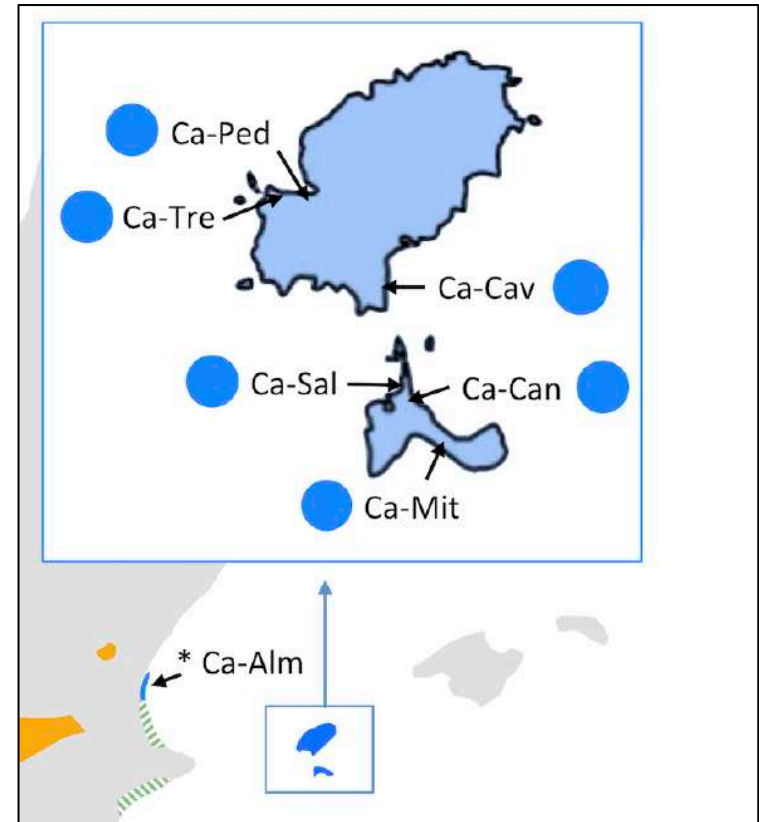


1. Phylogenetic reconstruction
2. Gene tree-Species tree reconciliation
3. Population genetic analyses
4. Similarity of *S. cambessedesii* plants from mainland and Balearic populations

geneious

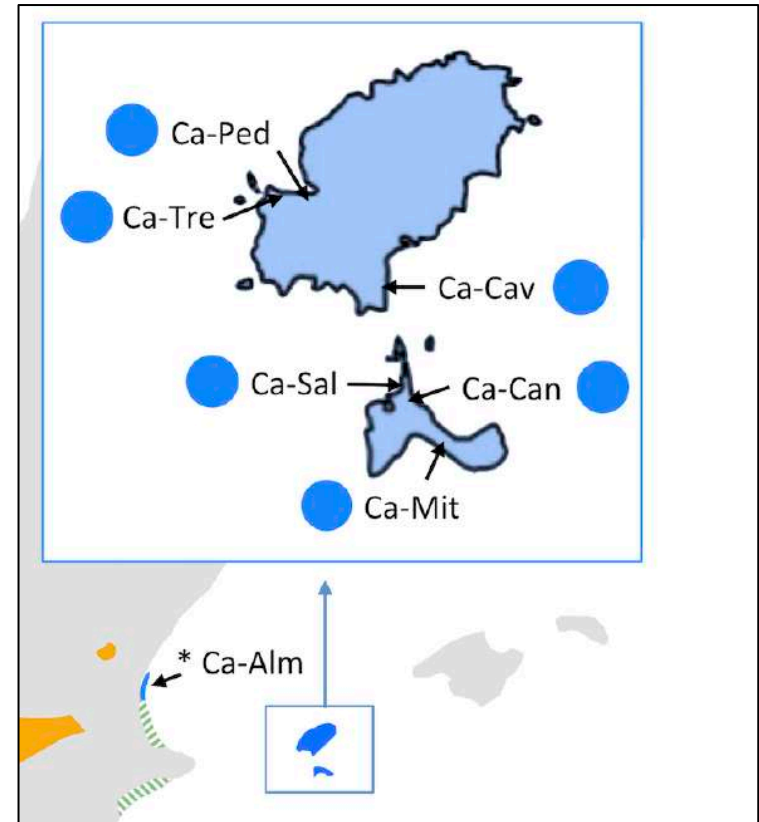
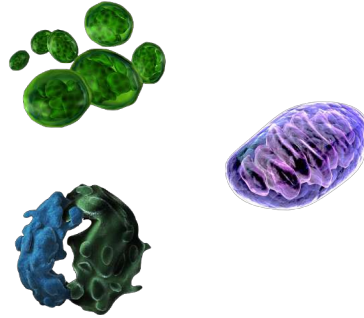


Sanger sequencing of four regions:



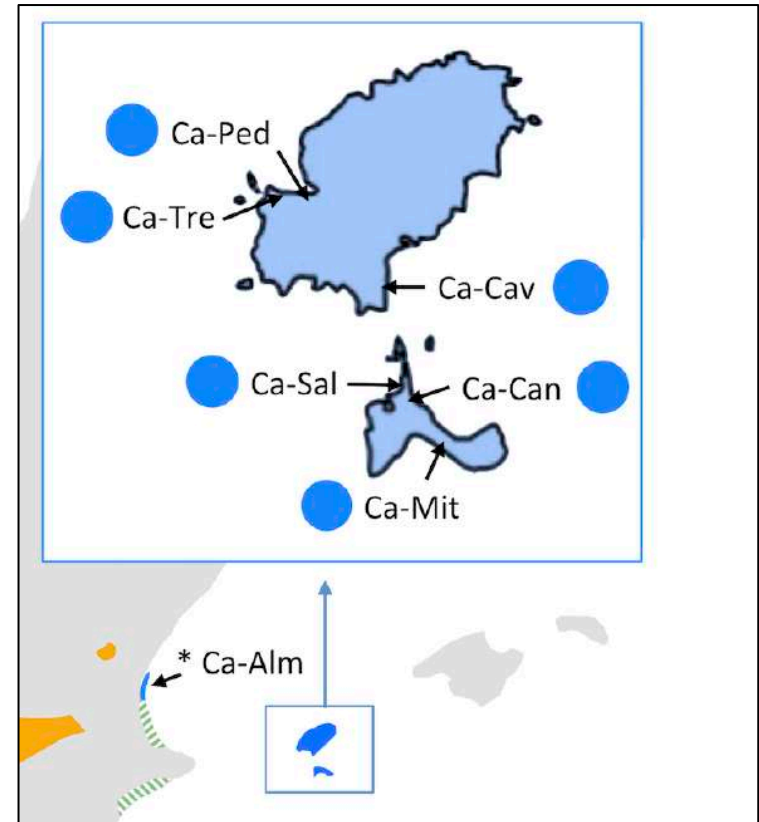
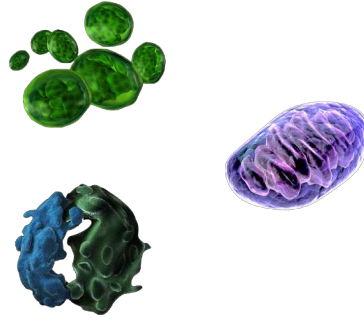
Sanger sequencing of four regions:

- **trnK**: 814 bp
- **ycf1**: 871 bp
- **atp1**: 753 bp
- **ITS**: 714 bp



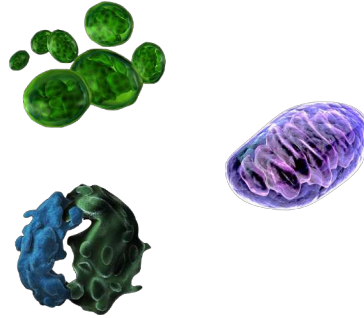
Sanger sequencing of four regions:

- **trnK**: 814 bp
  - **ycf1**: 871 bp
  - **atp1**: 753 bp
  - **ITS**: 714 bp
- 99.2% of sequences were identical (the remaining 0.8% correspond to ambiguities that were partially validated)

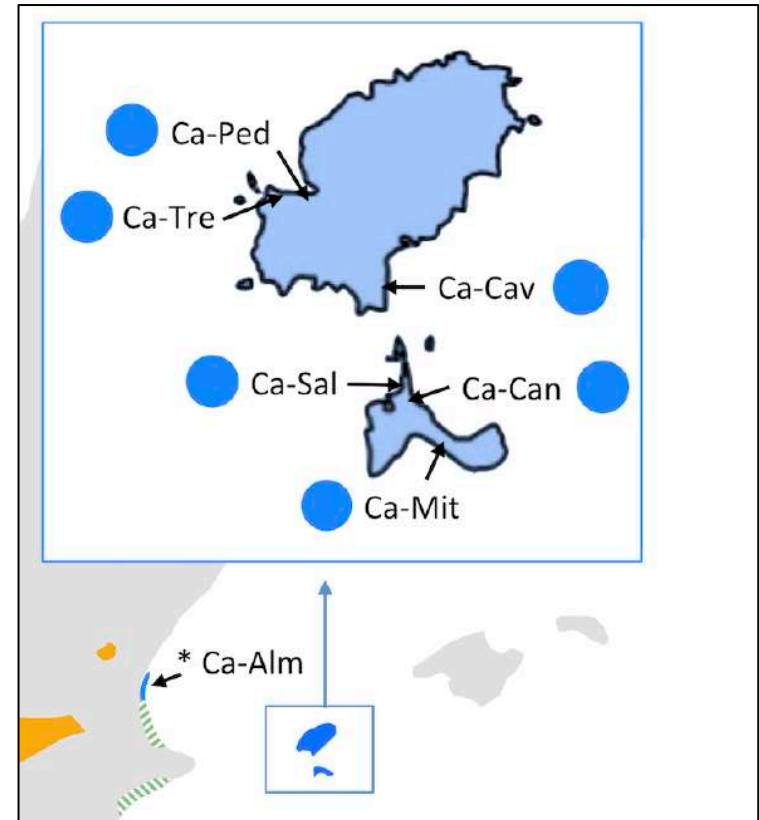


Sanger sequencing of four regions:

- **trnK**: 814 bp
- **ycf1**: 871 bp
- **atp1**: 753 bp
- **ITS**: 714 bp

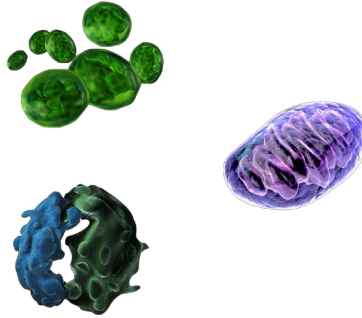


- 99.2% of sequences were identical (the remaining 0.8% correspond to ambiguities that were partially validated)
- Most SNPs were also present in some populations from the S and SE of the Iberian Peninsula. However:

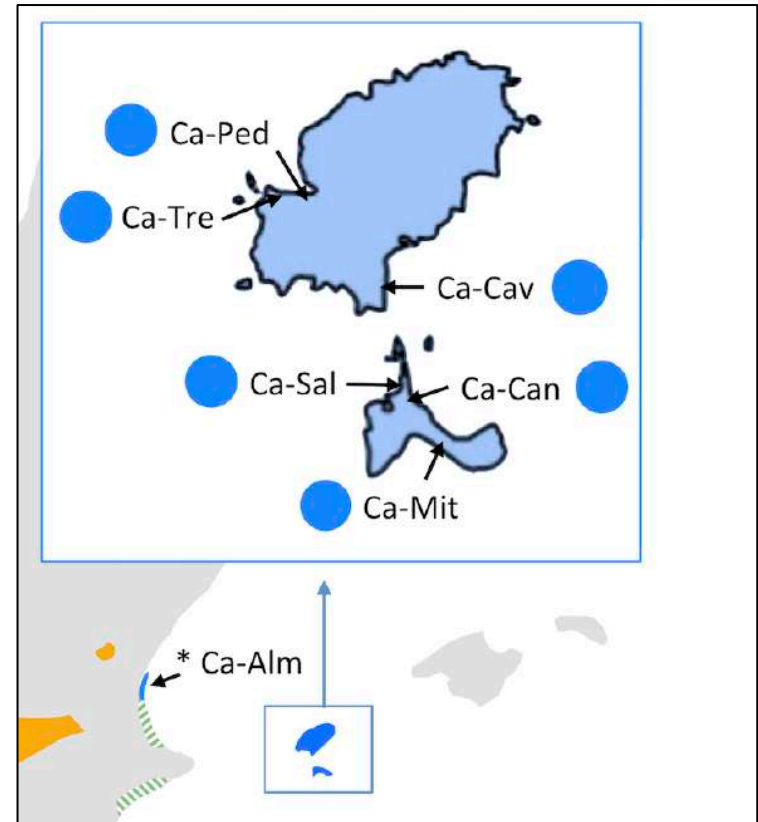


Sanger sequencing of four regions:

- **trnK**: 814 bp
- **ycf1**: 871 bp
- **atp1**: 753 bp
- **ITS**: 714 bp

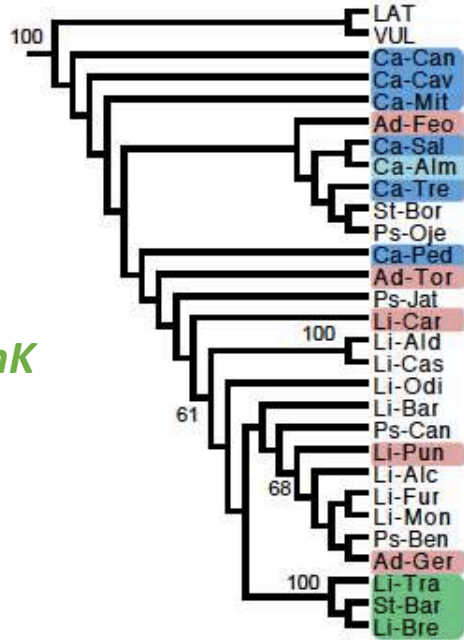


- 99.2% of sequences were identical (the remaining 0.8% correspond to ambiguities that were partially validated)
- Most SNPs were also present in some populations from the S and SE of the Iberian Peninsula. However:
  - **ITS1/ITS2**: 2 SNPs present only in Balearic populations
  - **ycf1**: 1 SNP was exclusively found in all *S. cambessedesii* samples

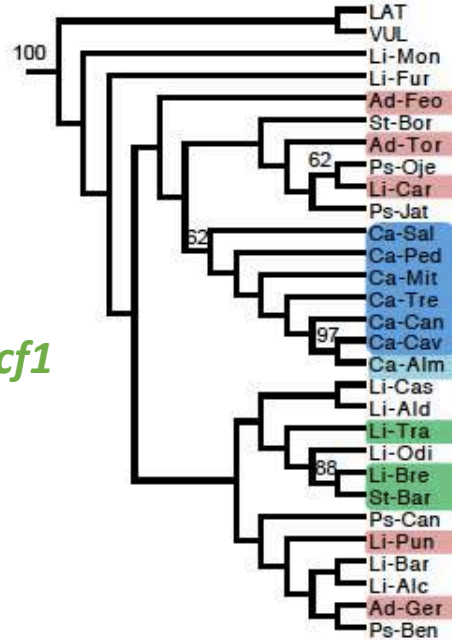


# La secuenciación del plastoma completo en la sección *Psammophilae* (*Silene*) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares

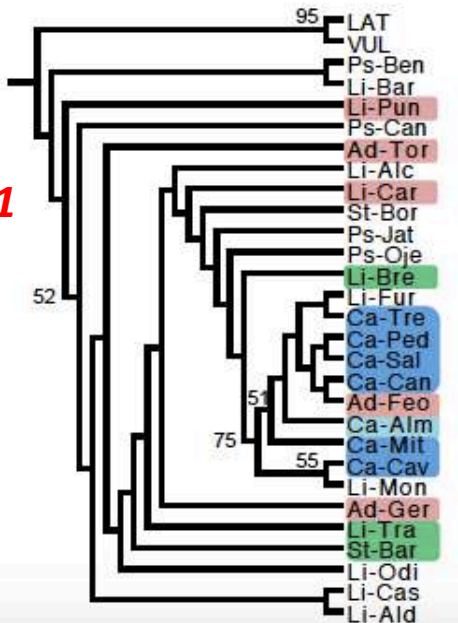
*trnK*



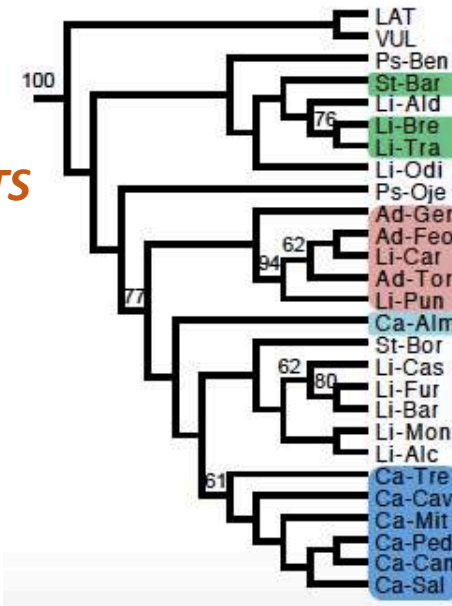
*ycf1*



*atp1*

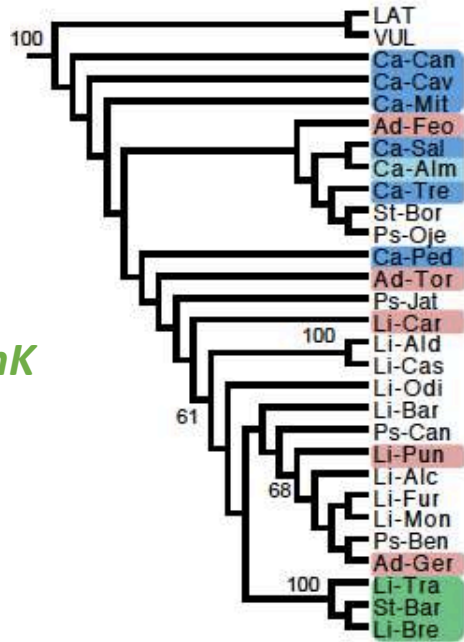


*ITS*

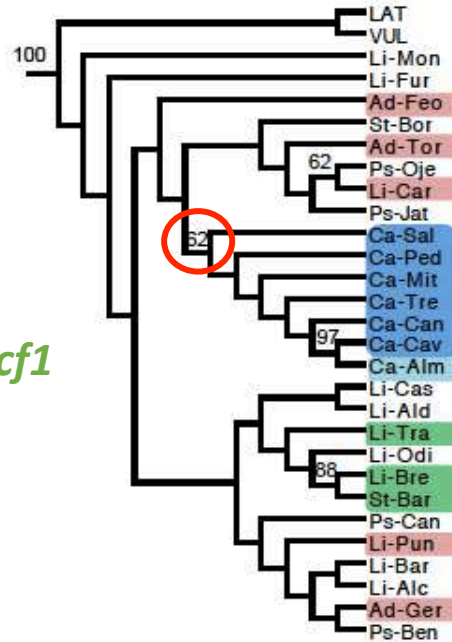


- Very low number of internal branches with strong support

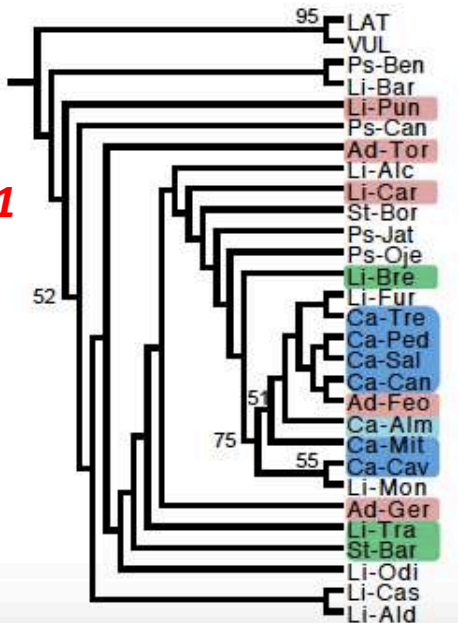
*trnK*



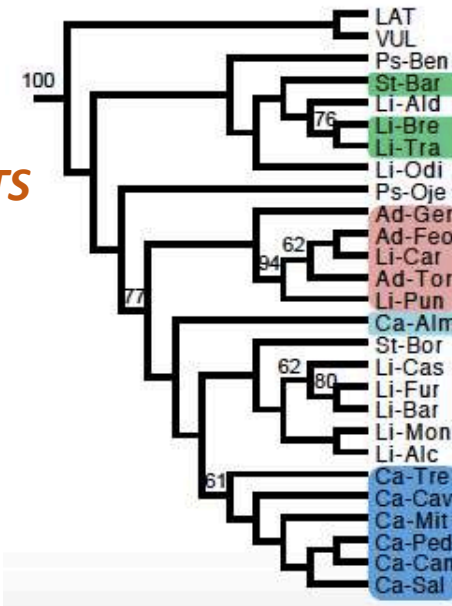
*ycf1*



*atp1*



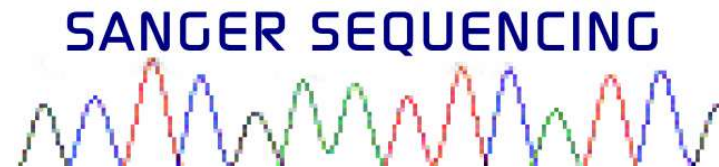
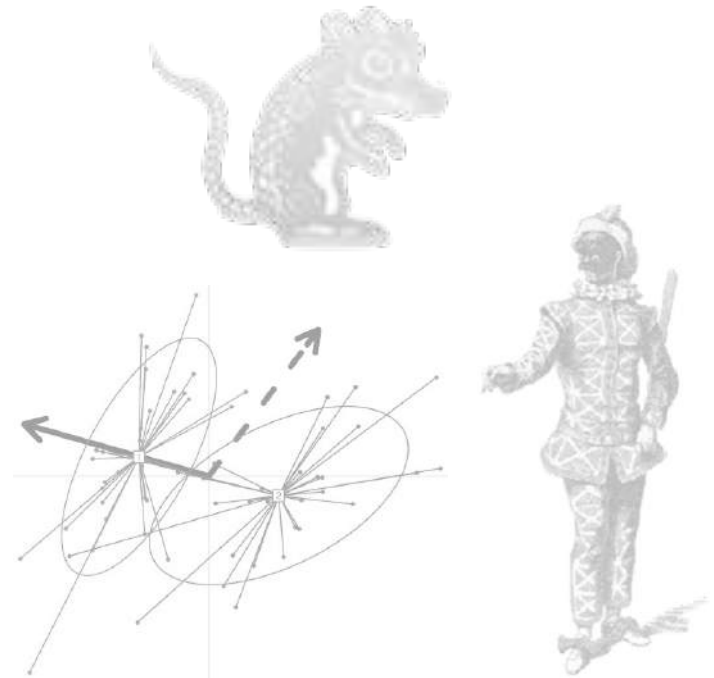
*ITS*



- Very low number of internal branches with strong support
- Only in the *ycf1*-based tree, the Almenara population formed a cluster with Balearic populations, but with but node was only moderately supported

1. Phylogenetic reconstruction
2. Gene tree-Species tree reconciliation
3. Population genetic analyses
4. Similarity of *S. cambessedesii* plants from mainland and Balearic populations
5. SNPs validation

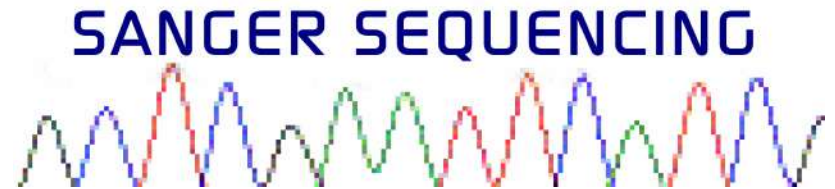
geneious



Sanger sequencing was performed in order to validate putative SNPs:



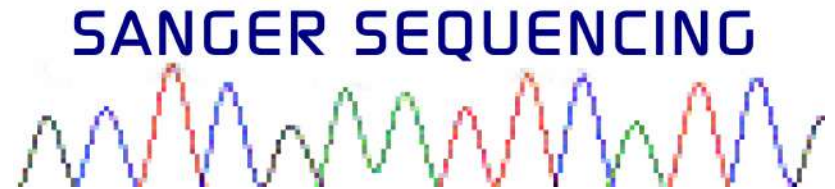
VS



Sanger sequencing was performed in order to validate putative SNPs:



VS

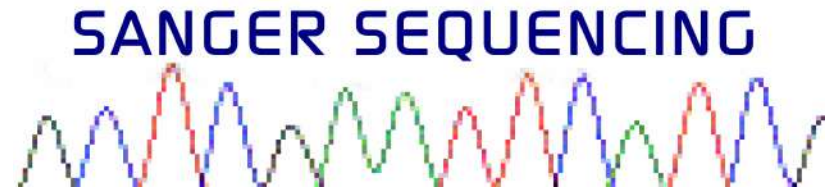


- PCR of many samples as possible from all populations

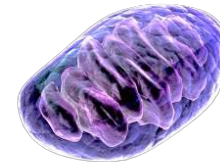
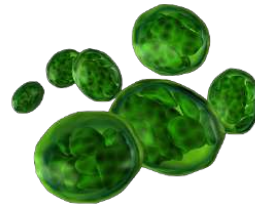
Sanger sequencing was performed in order to validate putative SNPs:



VS



- PCR of many samples as possible from all populations
- Amplification and sequencing of specific regions with high number of phylogenetically informative sites
  - **trnK**: 814 bp – 60 SNPs
  - **atp1**: 753 bp – 125 SNPs



We sequenced 16,648 bp (*trnK*) and 15,813 bp (*atp1*) by Sanger Seq:

*trnK* = 99.1% confirmed

- 3 SNPs confirmed at 100%
- 47 SNPs confirmed at 50%
- 10 SNPs not confirmed

*atp1* = 98.1% confirmed

- 11 SNPs confirmed at 100%
- 98 SNPs confirmed at 50%
- 16 SNPs not confirmed

We sequenced 16,648 bp (*trnK*) and 15,813 bp (*atp1*) by Sanger Seq:

*trnK* = 99.1% confirmed

- 3 SNPs confirmed at 100%
- 47 SNPs confirmed at 50%
- 10 SNPs not confirmed

*atp1* = 98.1% confirmed

- 11 SNPs confirmed at 100%
- 98 SNPs confirmed at 50%
- 16 SNPs not confirmed

NGS data →

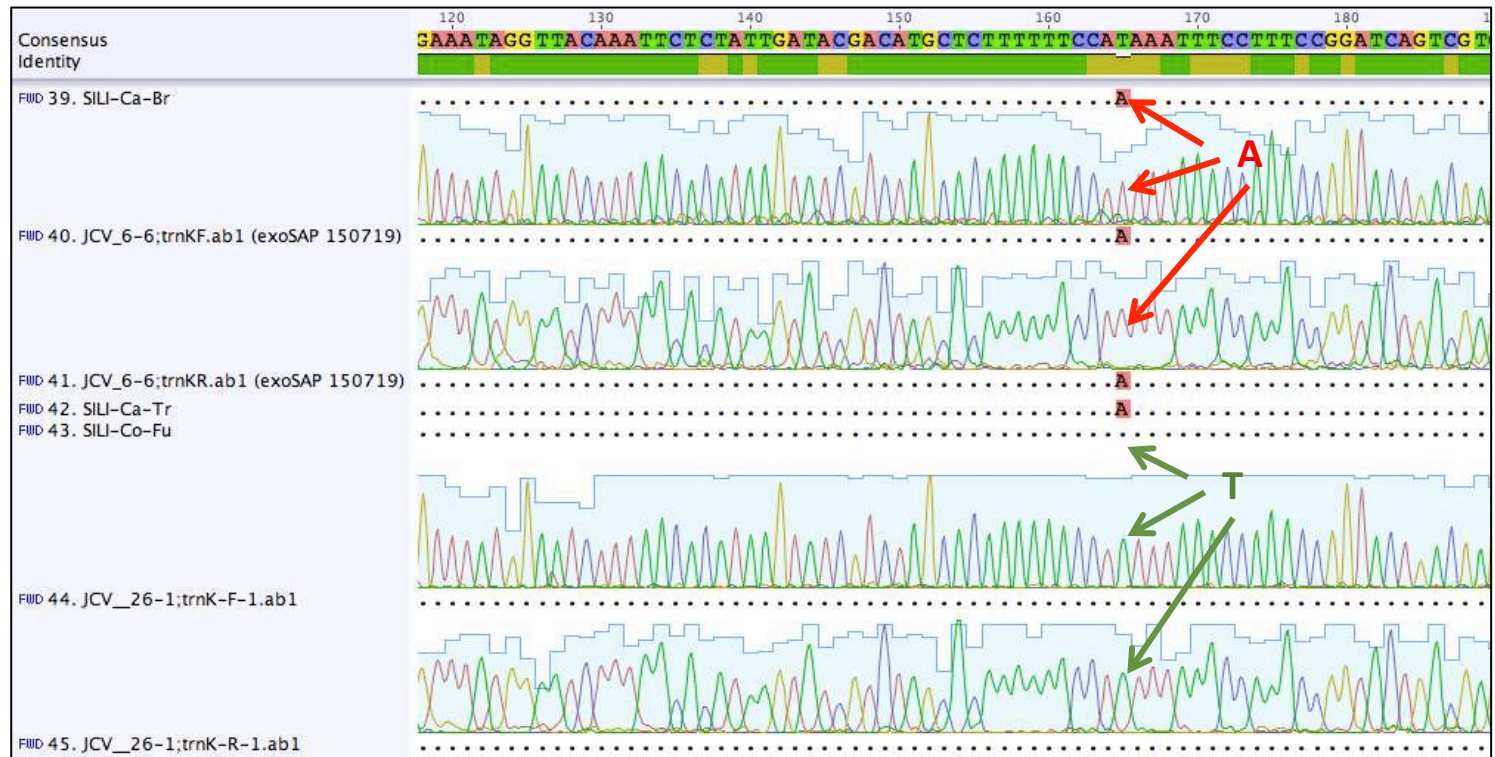
Sanger-F →

Sanger-R →

NGS data →

Sanger-F →

Sanger-R →



# Main conclusions



# ① Phylogenetic relationships and hybridization in Iberian *Silene*

## ① Phylogenetic relationships and hybridization in Iberian *Silene*

- There is an incongruence between taxonomy and phylogeographic results. In fact, geography was a better predictor of relatedness than either morphology-based species boundaries or edaphic preferences

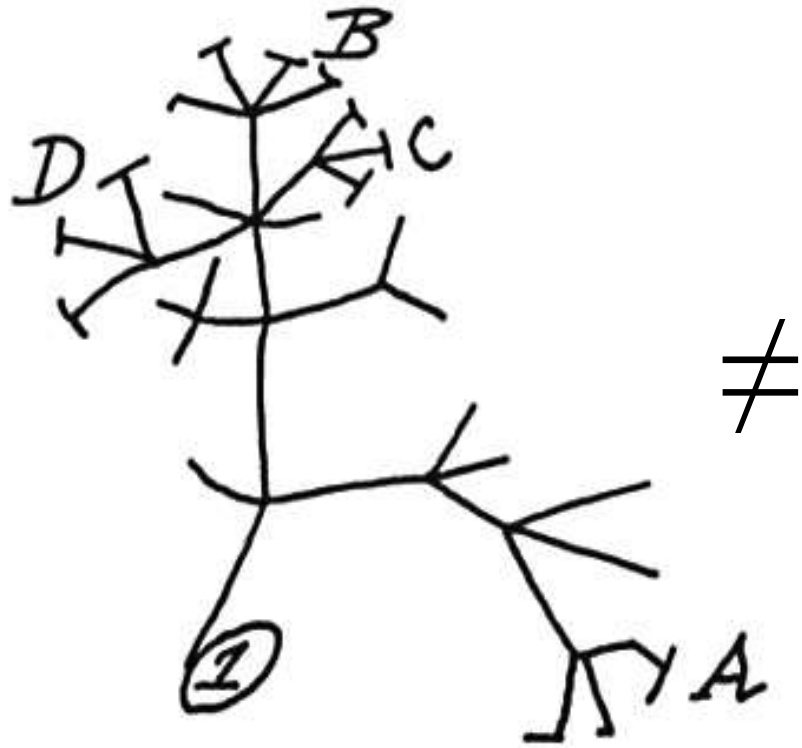
## ① Phylogenetic relationships and hybridization in Iberian *Silene*

- There is an incongruence between taxonomy and phylogeographic results. In fact, geography was a better predictor of relatedness than either morphology-based species boundaries or edaphic preferences
- Incongruence among genomes and the intricate relationships among species suggest hybridization is acting in *Silene* Sect. *Psammophilae*

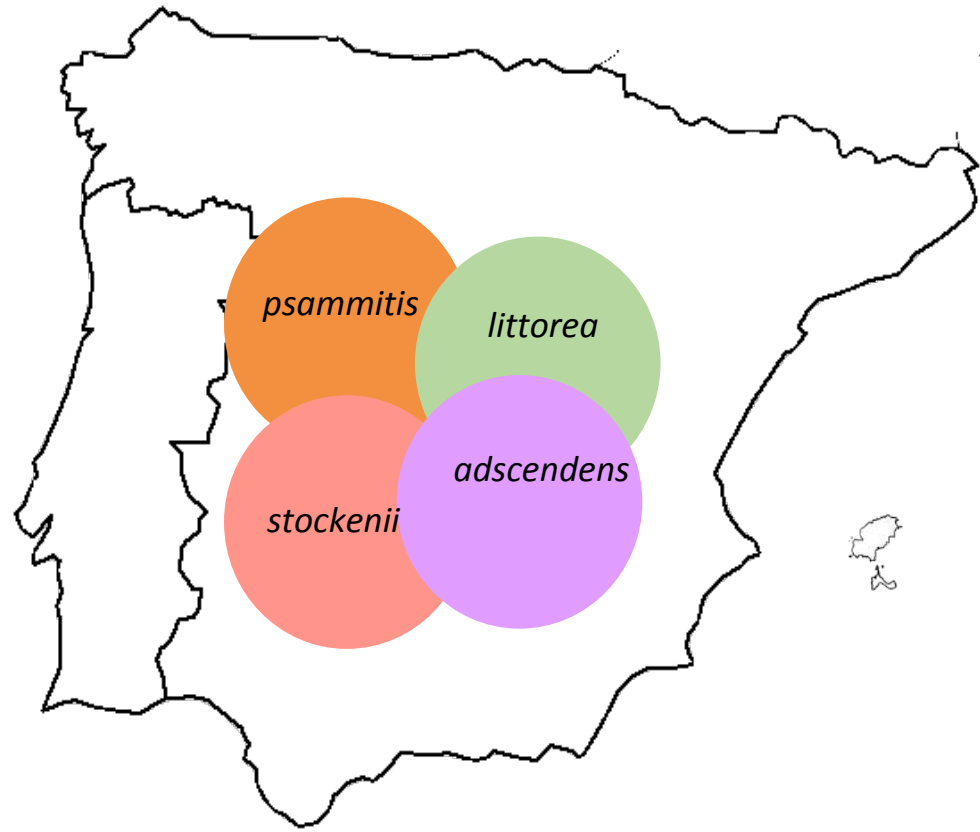
## ① Phylogenetic relationships and hybridization in Iberian *Silene*

- There is an incongruence between taxonomy and phylogeographic results. In fact, geography was a better predictor of relatedness than either morphology-based species boundaries or edaphic preferences
- Incongruence among genomes and the intricate relationships among species suggest hybridization is acting in *Silene* Sect. *Psammophilae*
- Geographic proximity of morphologically distinct species increases the probability of interspecific hybridization or introgression

**La secuenciación del plastoma completo en la sección *Psammophilae* (*Silene*) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares**

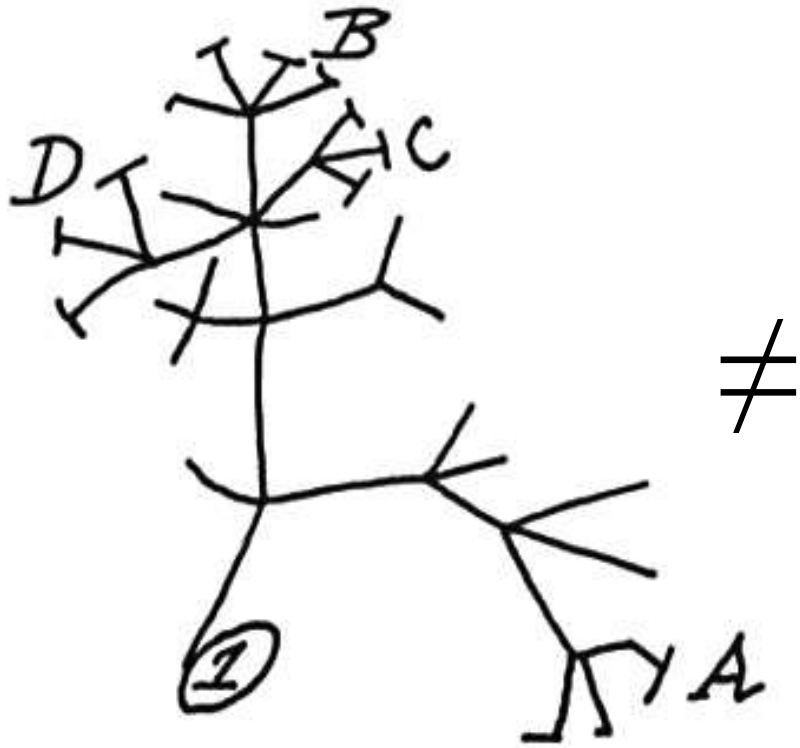


≠

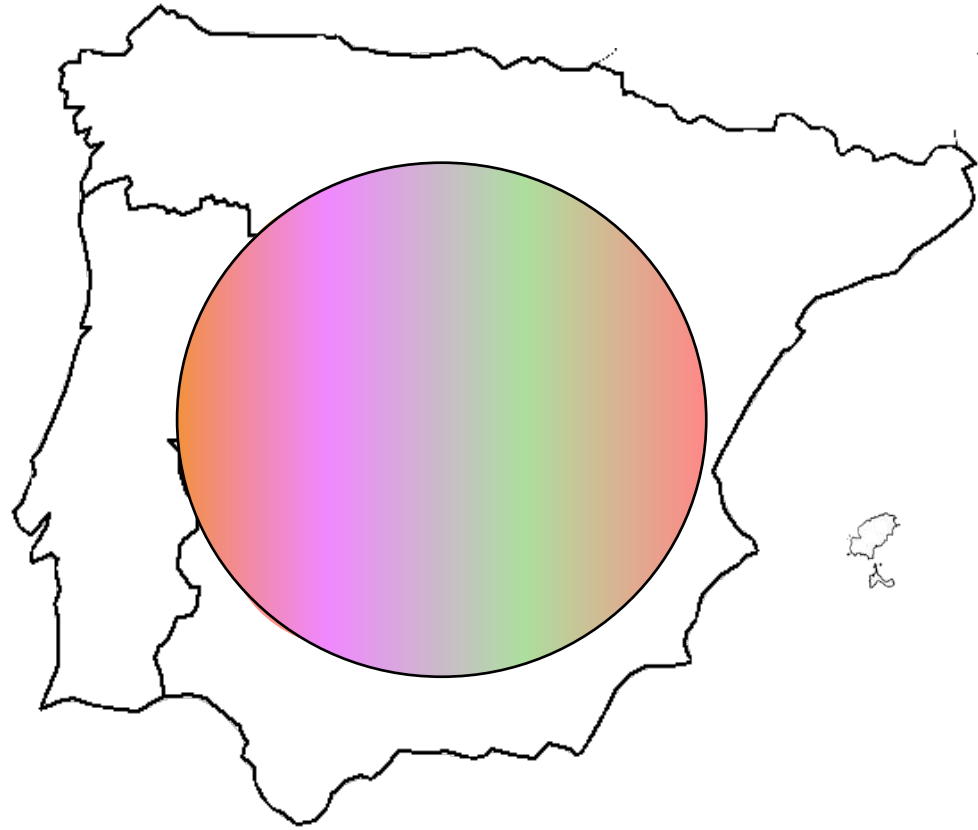


***La secuenciación del plastoma completo en la sección Psammophilae (Silene) revela un proceso de hibridación en poblaciones interiores y divergencia en las poblaciones de las Islas Baleares***

---

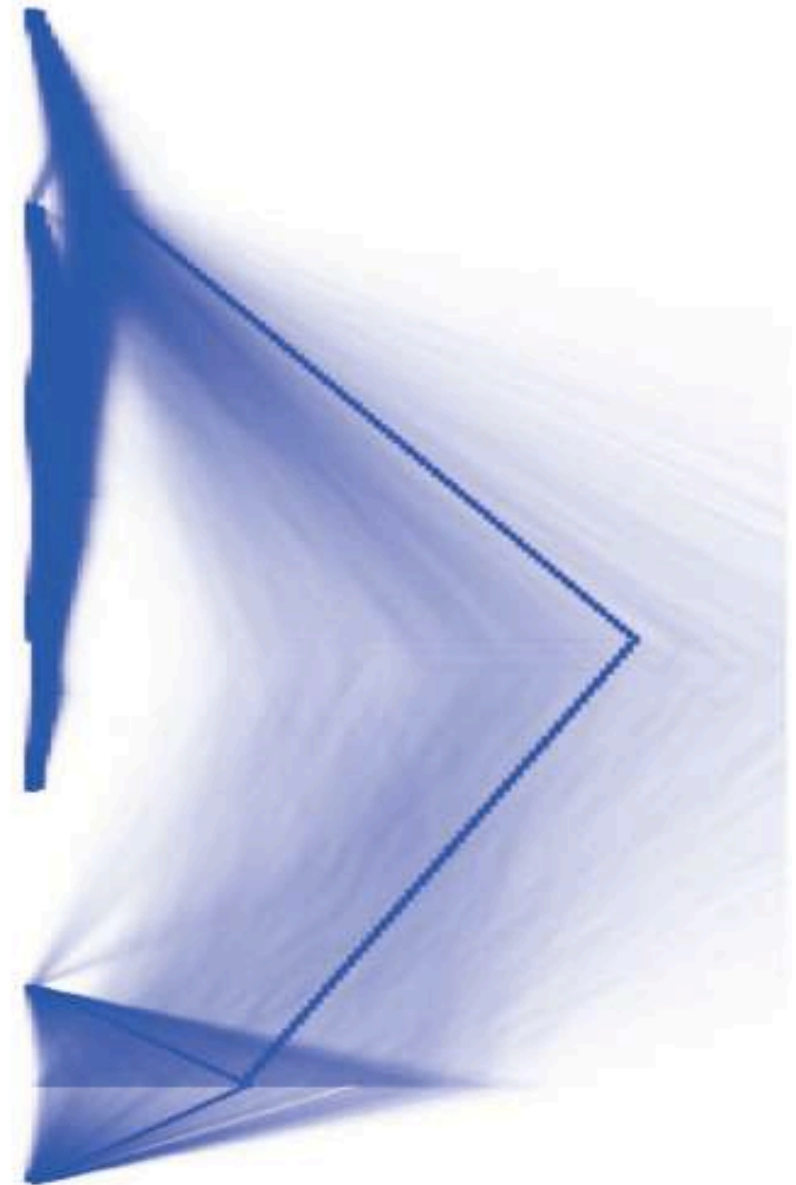
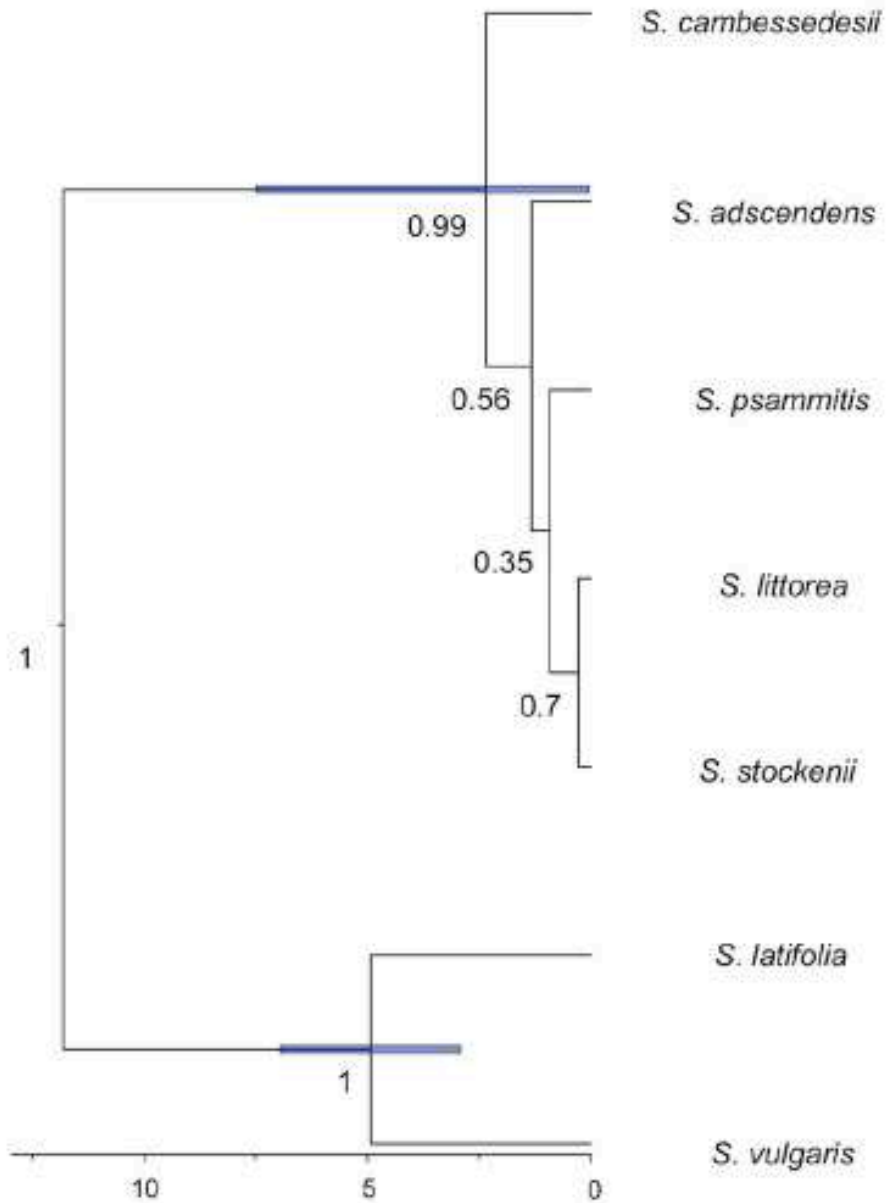


≠



## ① **Phylogenetic relationships and hybridization in Iberian *Silene***

- There is an incongruence between taxonomy and phylogeographic results. In fact, geography was a better predictor of relatedness than either morphology-based species boundaries or edaphic preferences
- Incongruence among genomes and the intricate relationships among species suggest hybridization is acting in *Silene* Sect. *Psammophilae*
- Geographic proximity of morphologically distinct species increases the probability of interspecific hybridization or introgression
- **Are species of *Silene* Sect. *Psammophilae* a single species?**



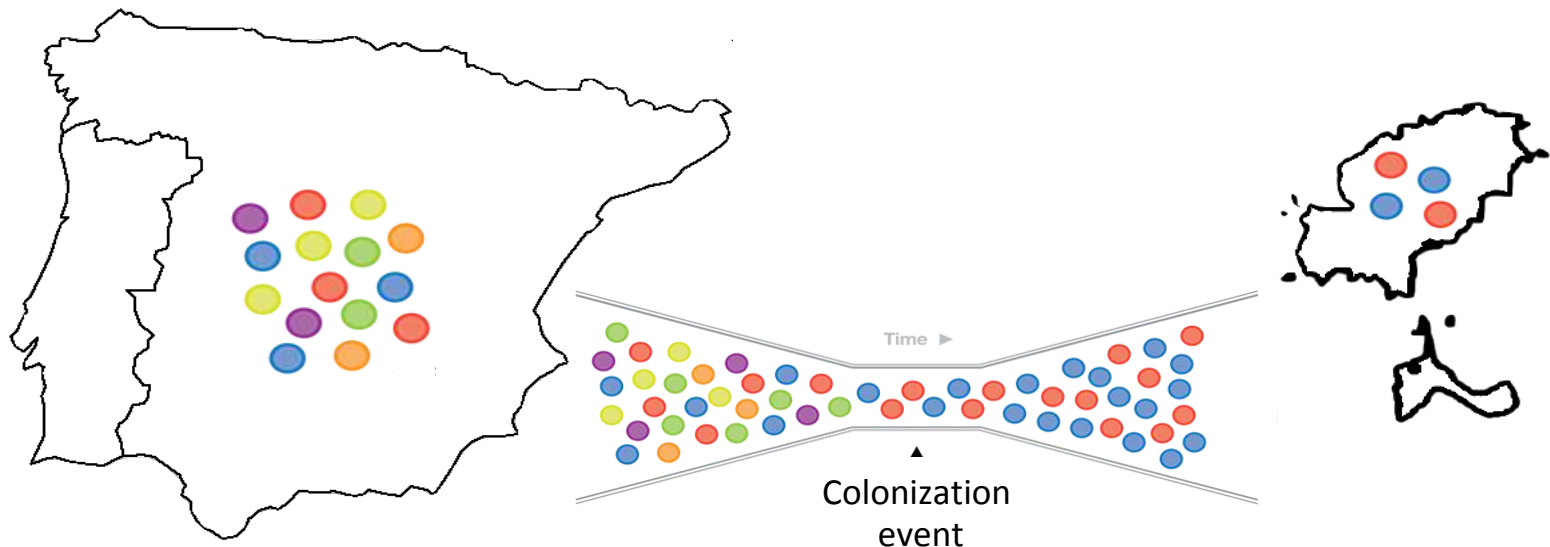
## ② Geographic isolation promotes genetic differentiation of Balearic populations

## ② Geographic isolation promotes genetic differentiation of Balearic populations

- Moderate genetic differentiation of *S. cambessedesii* with respect to mainland species

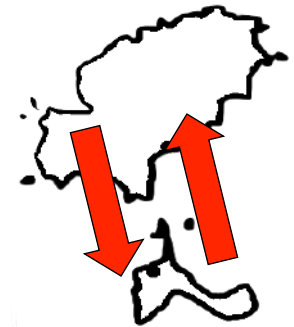
## ② Geographic isolation promotes genetic differentiation of Balearic populations

- Moderate genetic differentiation of *S. cambessedesii* with respect to mainland species
- This genetic differentiation is probably due to founder effects following the colonization of the Balearic Islands by a small number of individuals and/or subsequent bottlenecks



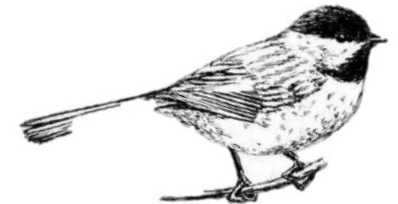
## ② Geographic isolation promotes genetic differentiation of Balearic populations

- Absence of genetic differentiation between populations from Ibiza and Formentera



## ② Geographic isolation promotes genetic differentiation of Balearic populations

- Absence of genetic differentiation between populations from Ibiza and Formentera
- The origin of *S. cambessedesii* could be explained by a long distant dispersal event from mainland to the Balearic islands



## ② Geographic isolation promotes genetic differentiation of Balearic populations

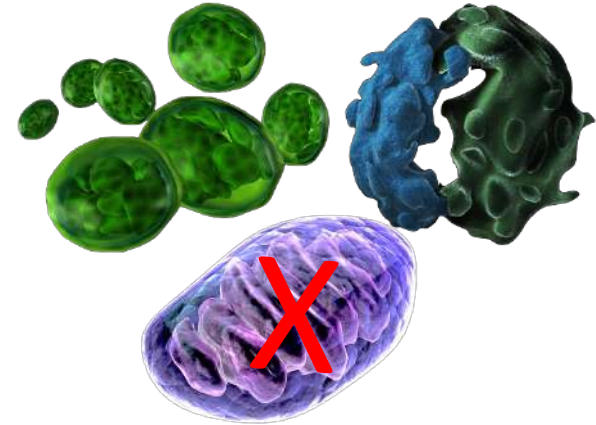
- Absence of genetic differentiation between populations from Ibiza and Formentera
- The origin of *S. cambessedesii* could be explained by a long distant dispersal event from mainland to the Balearic islands
- Genetic traces observed in DNA sequences of Almenara individuals are the result of at least one dispersal event from the islands back to the mainland, followed by introgression with mainland populations



### ③ On the relative accuracy of reference- guided assembly of each genome

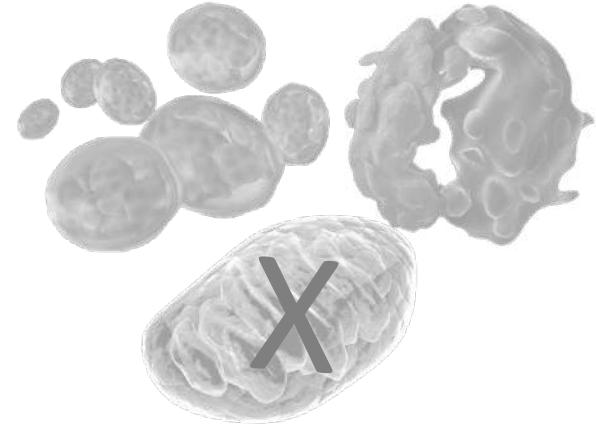
### ③ On the relative accuracy of reference- guided assembly of each genome

- Genome skimming is an efficient approach to generate the majority of the chloroplast genome, nrDNA cistron, and some mitochondrial coding sequences



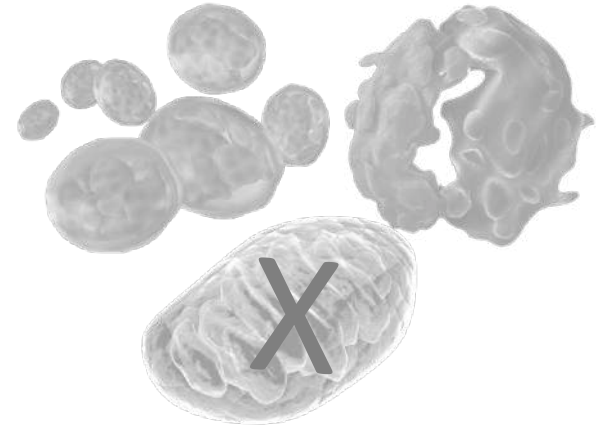
### ③ On the relative accuracy of reference- guided assembly of each genome

- Genome skimming is an efficient approach to generate the majority of the chloroplast genome, nrDNA cistron, and some mitochondrial coding sequences
- Larger fragments of the mitochondrial genome, including introns and intergenic regions, will be necessary for further biogeographic analyses



### ③ On the relative accuracy of reference- guided assembly of each genome

- Genome skimming is an efficient approach to generate the majority of the chloroplast genome, nrDNA cistron, and some mitochondrial coding sequences
- Larger fragments of the mitochondrial genome, including introns and intergenic regions, will be necessary for further biogeographic analyses
- In *Silene*, lineages from hybridization have phylogenetic histories that maybe cannot be easily solved, even with high amount of data from NGS



# Acknowledgements

Thank you for your help:

- Justen B. Whittall – Santa Clara University, Santa Clara, CA, USA)
- Inés Casimiro-Soriguer – Universidad Pablo de Olavide, Seville, Spain)
- Eduardo Narbona – Universidad Pablo de Olavide, Seville, Spain
- M<sup>a</sup> Luisa Buide – Universidad Pablo de Olavide, Seville, Spain
- Daryn Baker – Santa Clara University, Santa Clara, CA, USA)
- Marcial Escudero – Universidad de Sevilla, Seville, Spain)
- Thunder – Santa Clara University, Santa Clara, CA, USA)
- Ágata Cardoso

... and thank you for your attention!



CGL2009-08257, CGL2012-37646 and CGL2015-63827-P



