

# Distribution of genetic diversity in forest tree species: role of demography and selection

Giovanni G. Vendramin

*IGV-CNR*

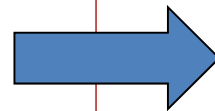
Delphine Grivet, Mario Zabal Aguirre, Santiago C. González-Martínez  
Federico Sebastiani, Sara Torre  
Hadrien Lalague, Bruno Fady

***CNR-FAO “International year of forests: reasearch needs and perspective” – Rome, 17 October 2011***



The importance of trees for sustaining life in general and biodiversity in particular can hardly be overstated

- 27% of the terrestrial surface is covered by forest
- trees make up around 90% of the biomass on our planet
- forest harbour the vast majority of the world's terrestrial biodiversity



The impact of modern global change is likely to depend to a great extent on the reaction of trees and the ecosystem they sustain

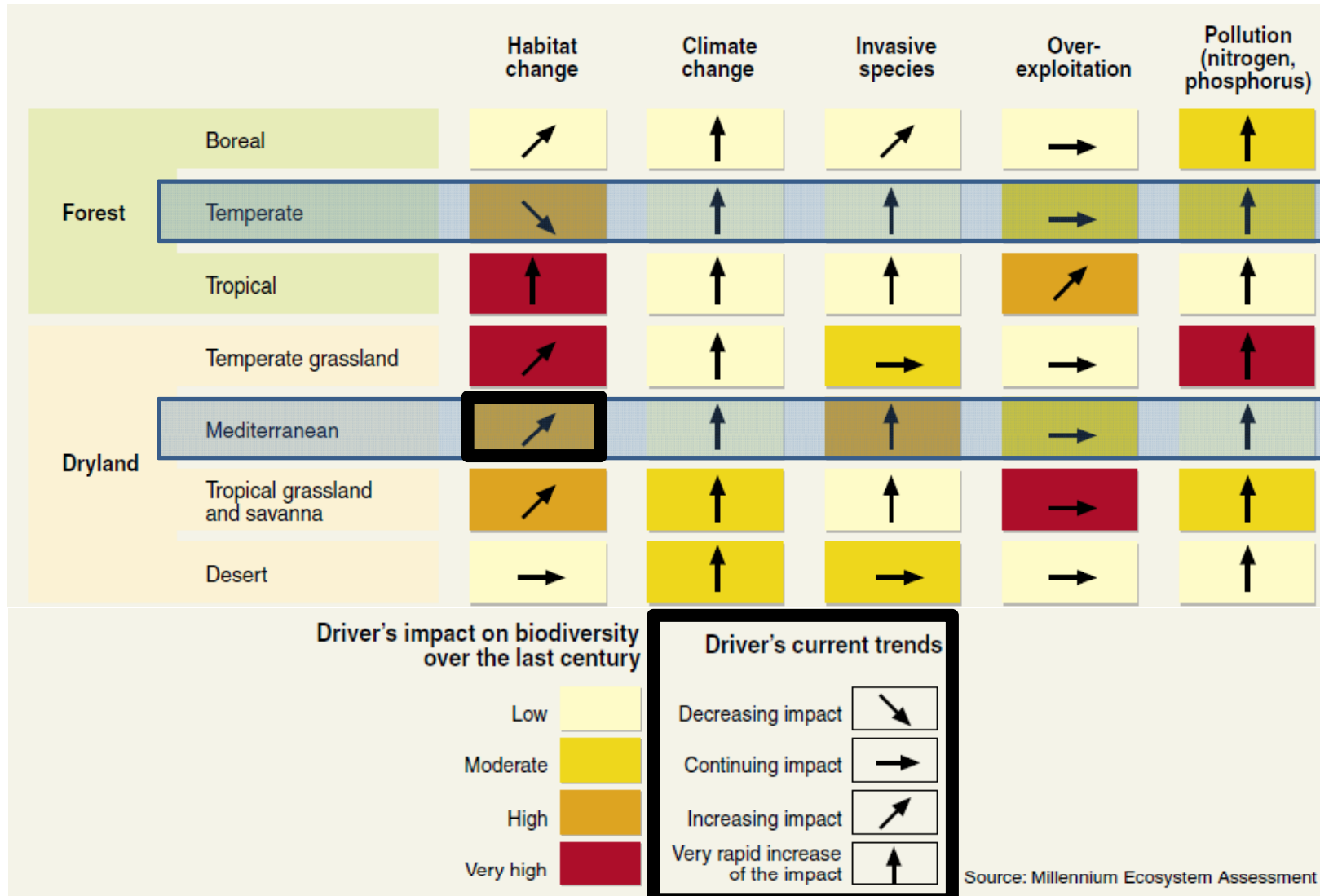
Estimates of global tree species richness range from a low 60.000 to 100.000 taxa, that is, as much as 15 to 25% of the 350-450.000 vascular plants

Currently, around 10% of the world's tree species are threatened to extinction, especially in the tropics, and many more tree populations will likely succumb



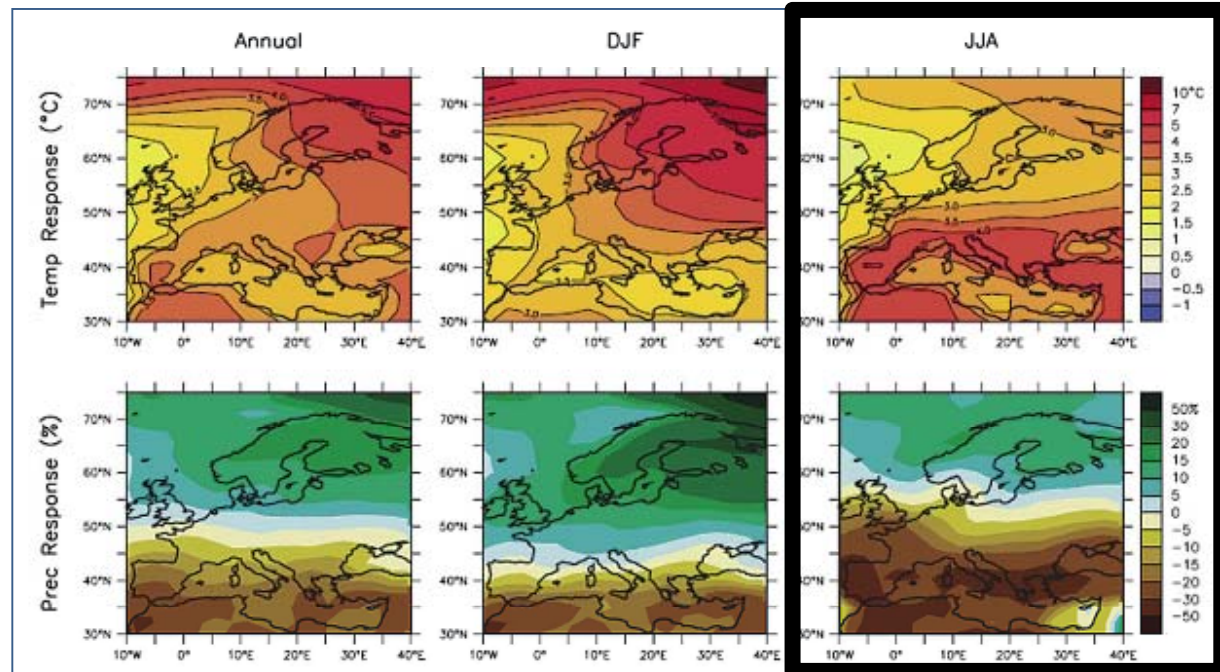
This figures underpin the need to study tree biodiversity and evolution

# Predictions for the Mediterranean region are dire





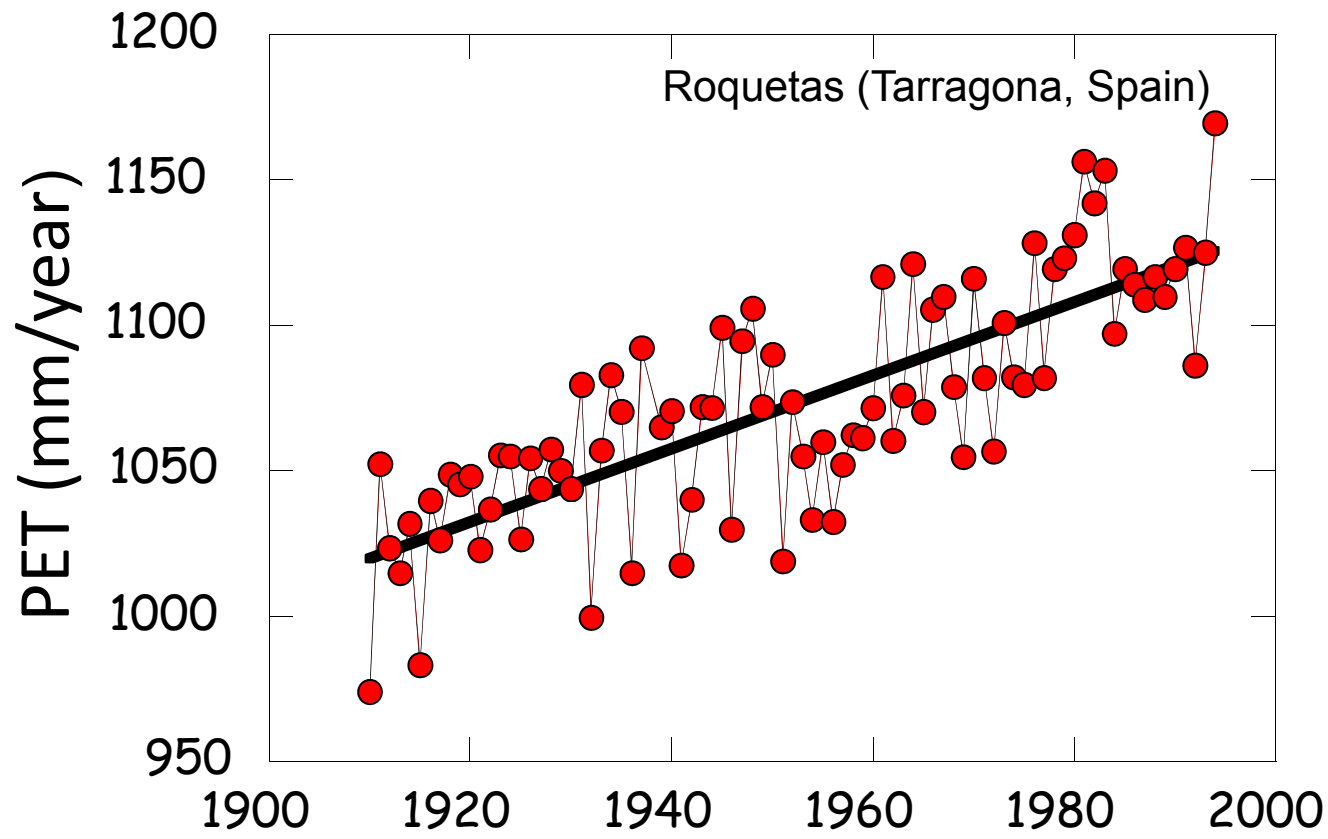
Temperature increment from 1850-1899 to 2001-2005 has been  $0.76^{\circ}\text{C}$  on a global scale; in some Mediterranean regions, however, the increment from 1971 to 2000 was  $1.53^{\circ}\text{C}$  –a higher value than the  $1.2^{\circ}\text{C}$  predicted by the climate models.



Mediterranean Forest  
Research Agenda  
2010-2020 - EFIMED

In particular, it is noticeable the predicted increase in temperature and decrease in precipitation in summer

Aridity has increased in the Mediterranean area during the last century



(Piñol et al. 1998)

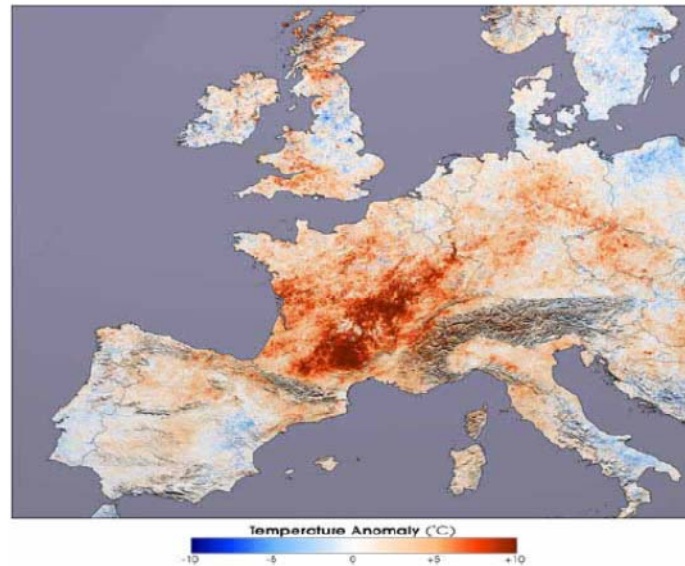
**Observed evolution:**  
Aridity has increased.  
PET is higher

**Temperature increment**

Global scale: 0.76°C  
(1859-99 to 2001-05)  
Spain Mediterranean:  
1.53°C (1971-2000)  
(>1.2 °C predicted by models)

# Increase in the risk of extreme events. It has already happened in Europe

## Heatwave 2003



## Increase in the risk of extreme events. What is predicted ? An example with France

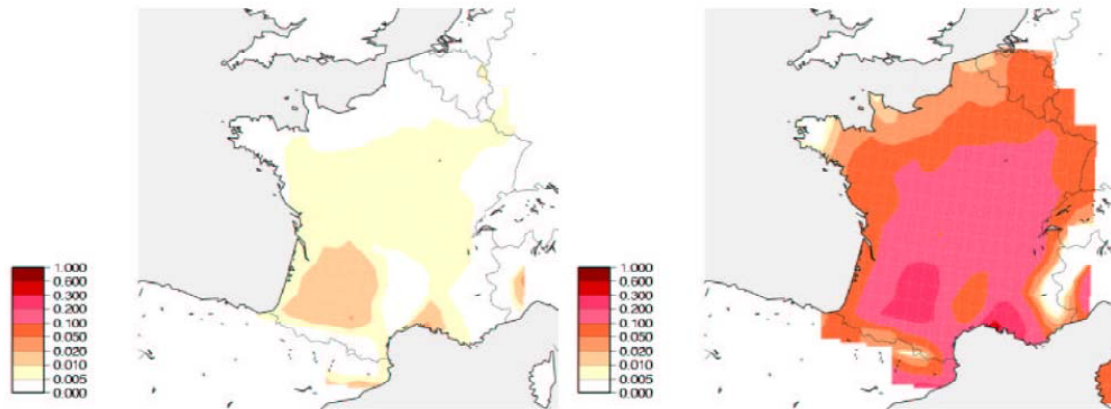
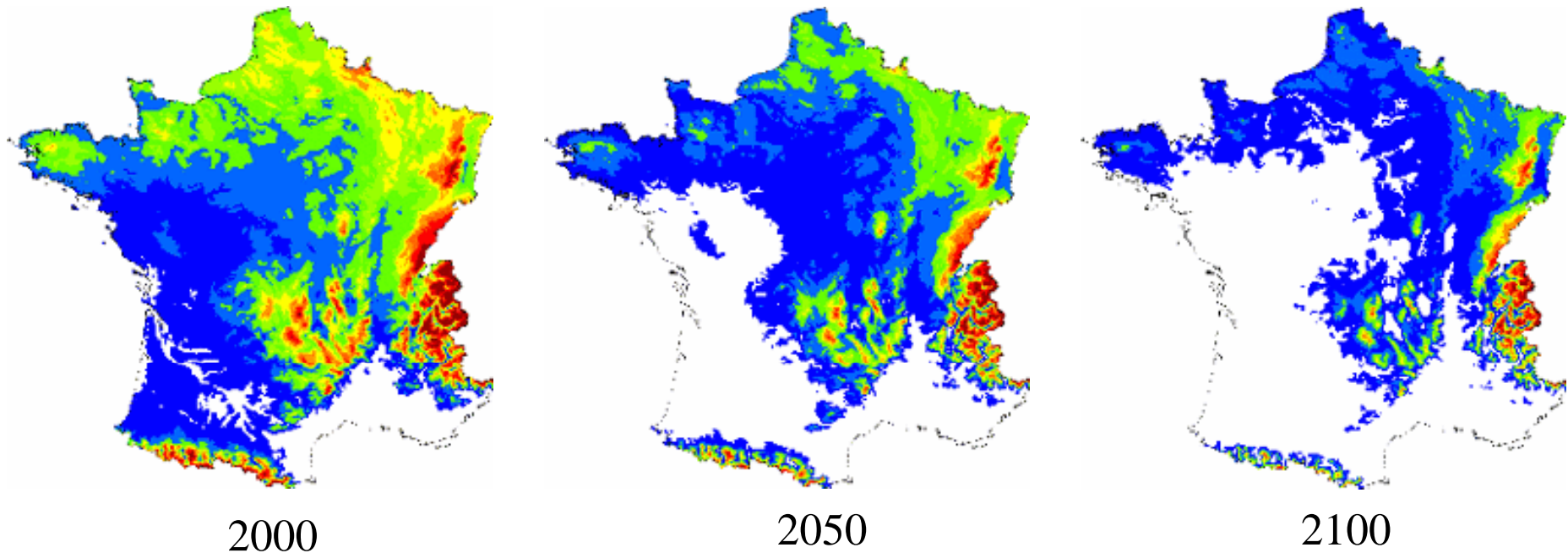


Figure 4: Probabilité d'avoir une température maximale supérieure à 35°C en été pour la période 1961-1990 (à gauche) et la période 2071-2100 (à droite).

**The limits of plasticity:  
predicted effect of climate change on tree distributions**

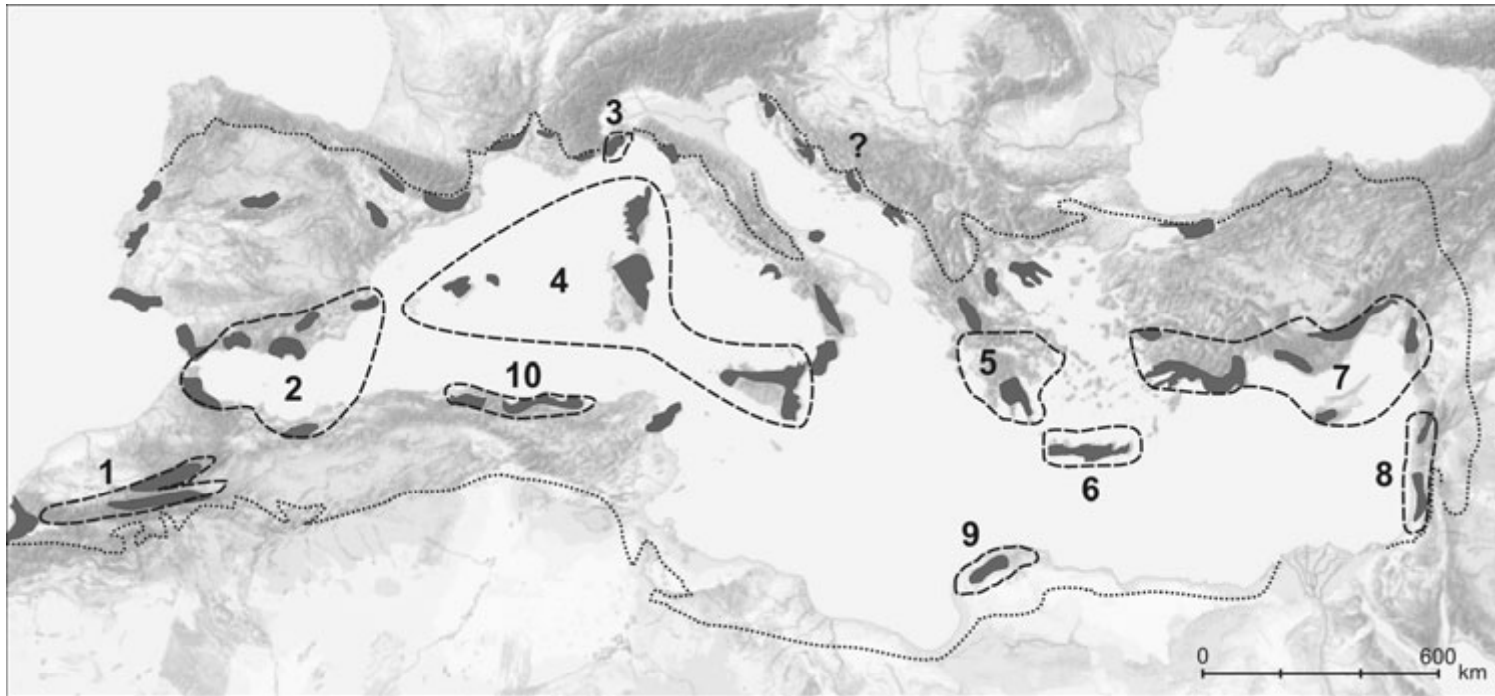
**Predicted evolution of the geographic distribution of *Fagus sylvatica* in France under climate warming (AURELHY climate model)**





# The Mediterranean forest: a threatened hotspot of diversity

A biodiversity hotspot with over 11,000 endemic plants



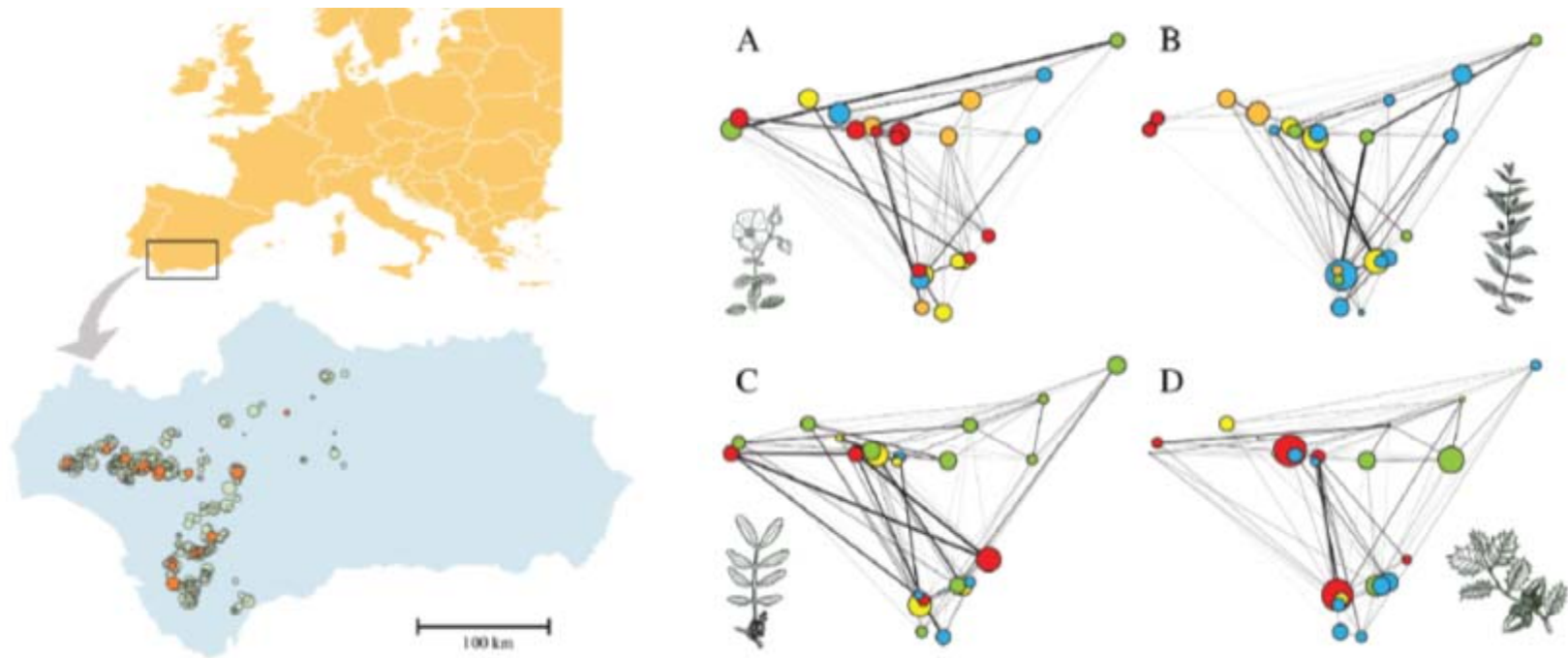
- 52 putative refugia within the Mediterranean region (resulting from the analysis of the phylogeographical patterns of 82 plant species, including 41 trees)
- 10 regional hotspots of plant biodiversity

(Médail & Diadema 2009 J. Biogeogr.)

# Some genetic particularities of Mediterranean forests

## Highly fragmented landscapes

(Fortuna et al. 2009 PNAS)

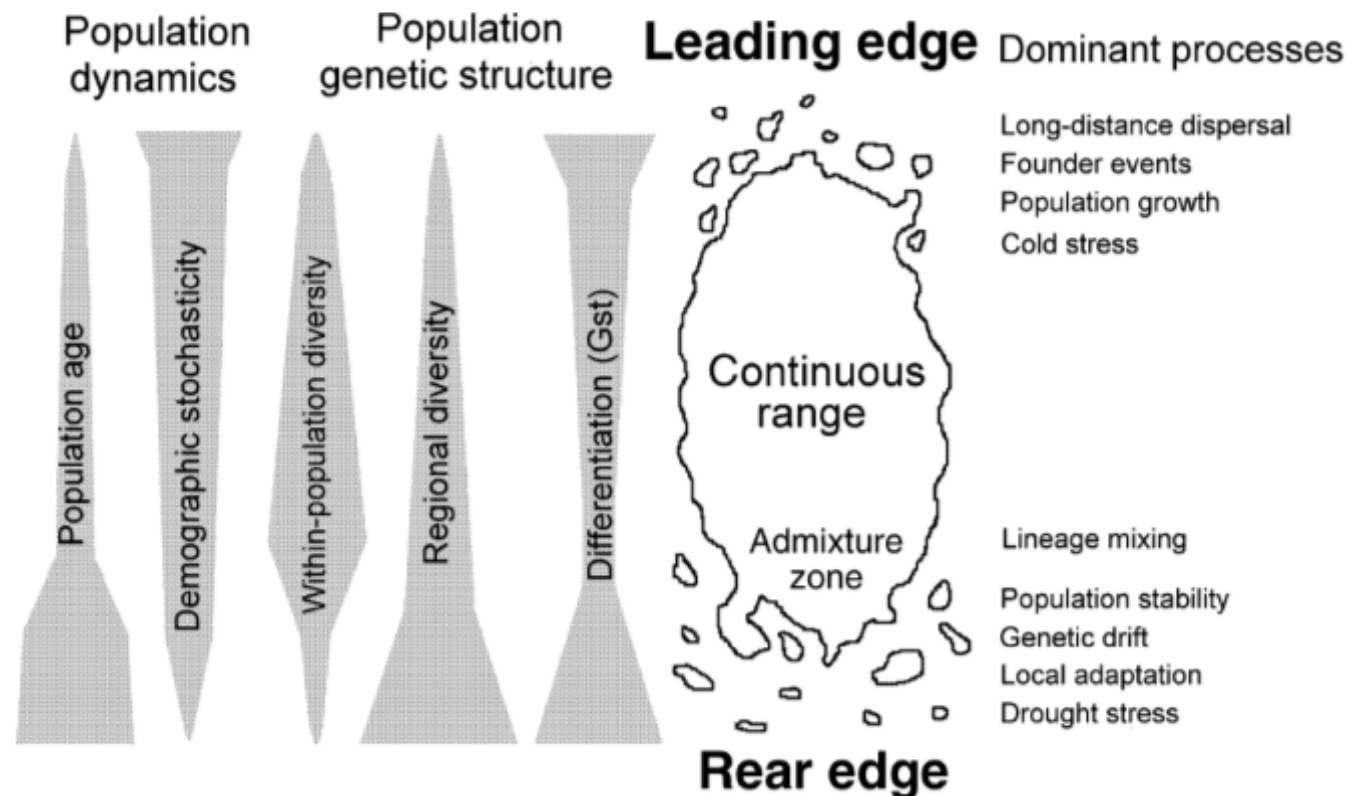


Fragmentation can reduce the evolutionary potential of Mediterranean forests; some fragments are key to maintain connectivity via gene flow.

# Some genetic particularities of Mediterranean forests

‘rear edge’ vs ‘leading edge’ populations

(Hampe & Petit 2006 Ecology Letters)

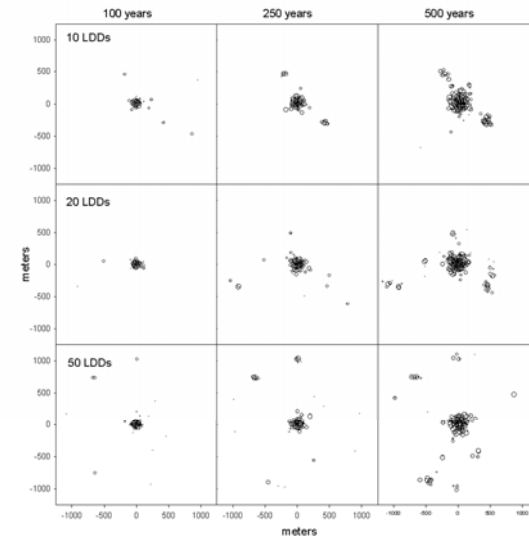


Genetic variation at European southern limits (i.e. Mediterranean environments) may contain relevant adaptations to warmer and drier conditions

# Confronted with global change **forest trees** will have either **to migrate** or **to adapt in situ**

Rates of spread of forest trees (<100 m/year) are far below what would be necessary (3,000 to 5,000 m/year) for species migration to track future climatic warming (Petit et al. 2008 Science).

**Example:** Spread rates in Mediterranean *P. pinaster* based on best-fitting dispersal kernels from natural populations: just a few meters of forest advance per year!

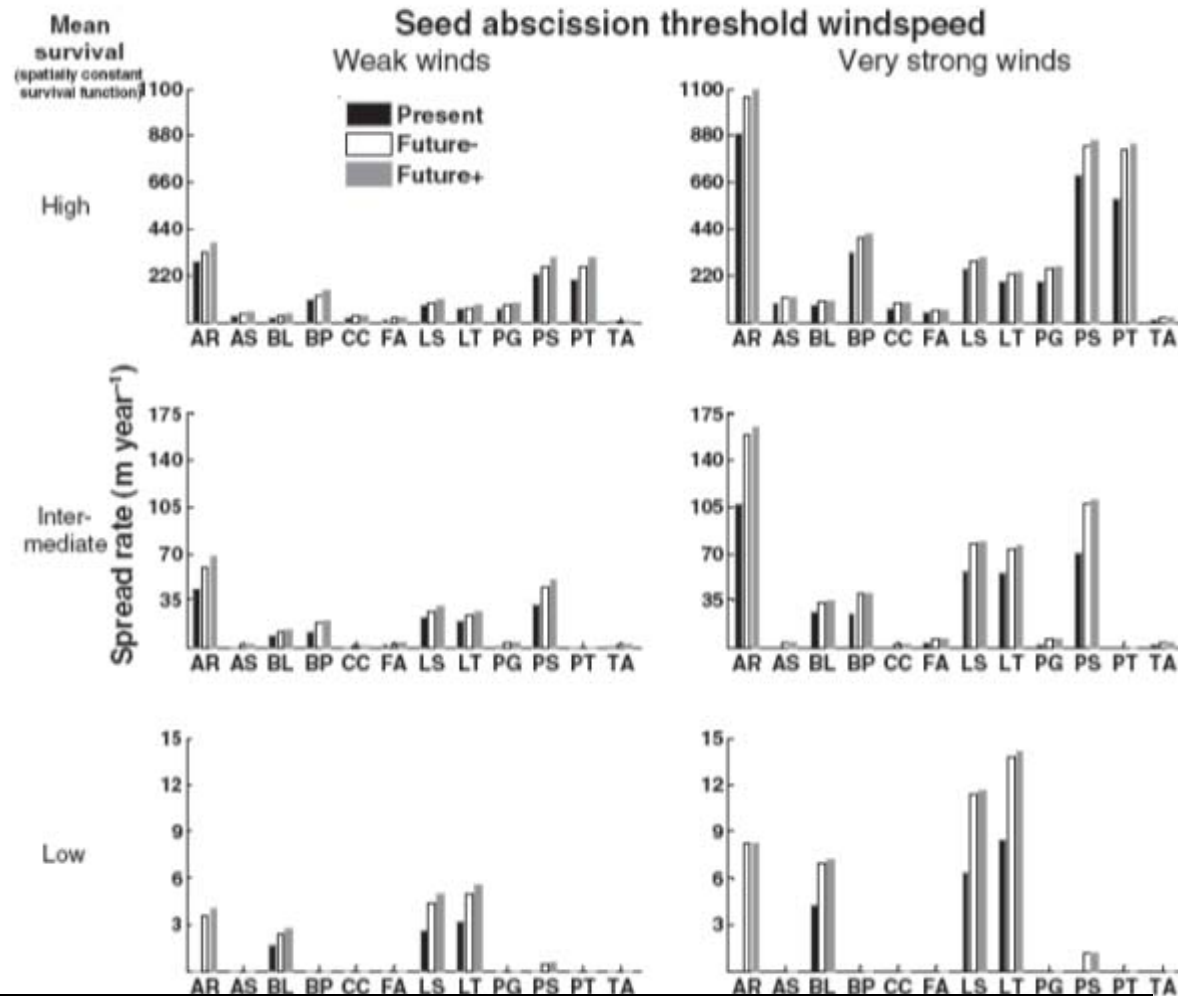


# Confronted with global change **forest trees** will have either **to migrate** or **to adapt in situ**

Rates of spread of forest trees are far below what would be expected (e.g., 5,000 m/year) for specific future climatic warming (IPCC Science).

Example: *Pinaster* dispersal kernels from meters of

Only in a scenario where trees have very high mean survival they would be able to track environmental optima  
(Nathan et al. 2011 Ecology Letters, for North American trees)

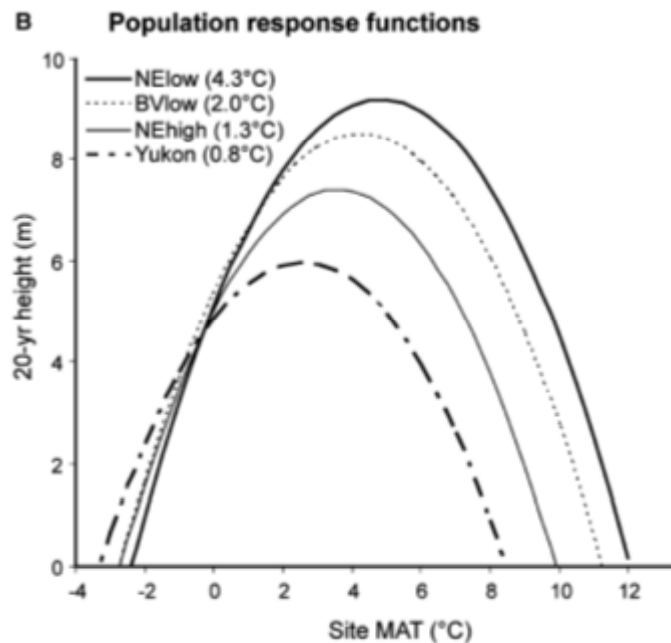


**Still, long-distance dispersal may save the day!**



# Thus, forest trees will probably have **to adapt in situ** to survive

To adapt in situ, plant species rely mostly on **standing genetic variation** (and not in de novo mutations), so that the level of genetic variation present in natural populations is important.



## Some advantages of trees are:

- They are undomesticated and still contain large amounts of genetic variation.
- Reaction norms from forest trees show high level of plasticity.
- Normally, tree populations have big effective sizes (i.e. better response to selection).

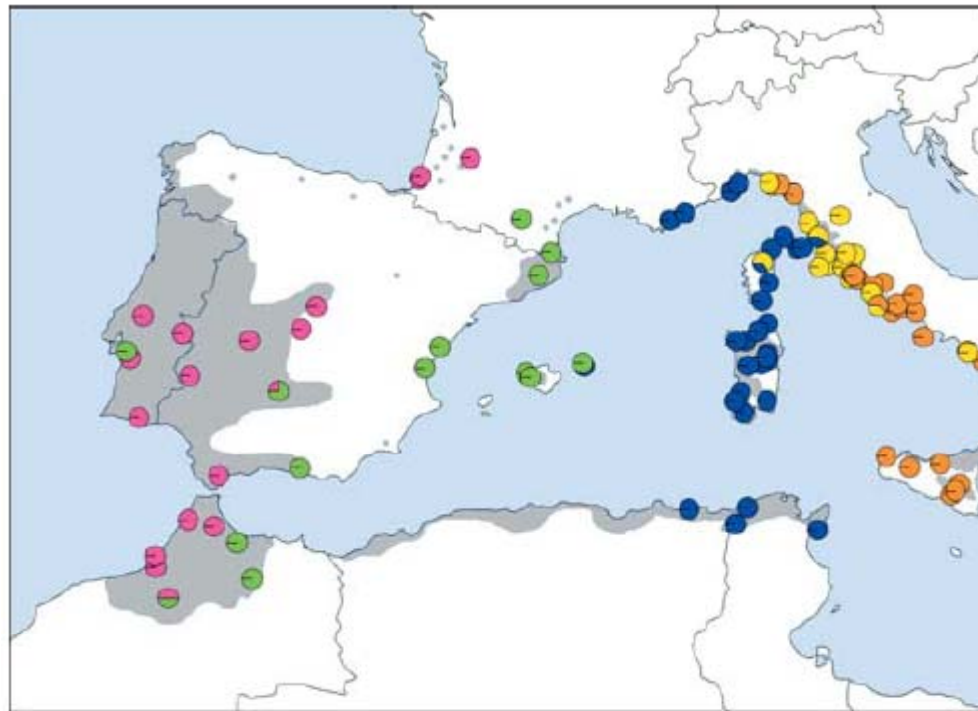
(Wang et al. 2006 Global Change Biology, for *Pinus contorta*)

# Complex system that results is significant population structure and opportunities for local adaptation

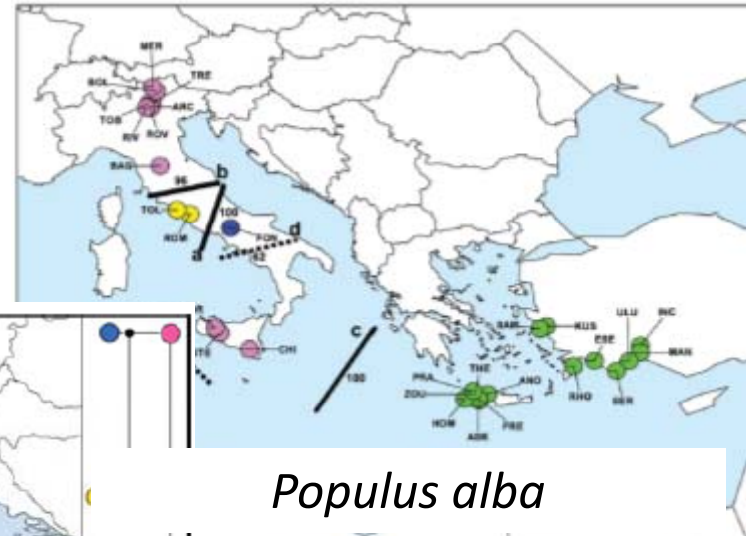
*Pinus pinaster*



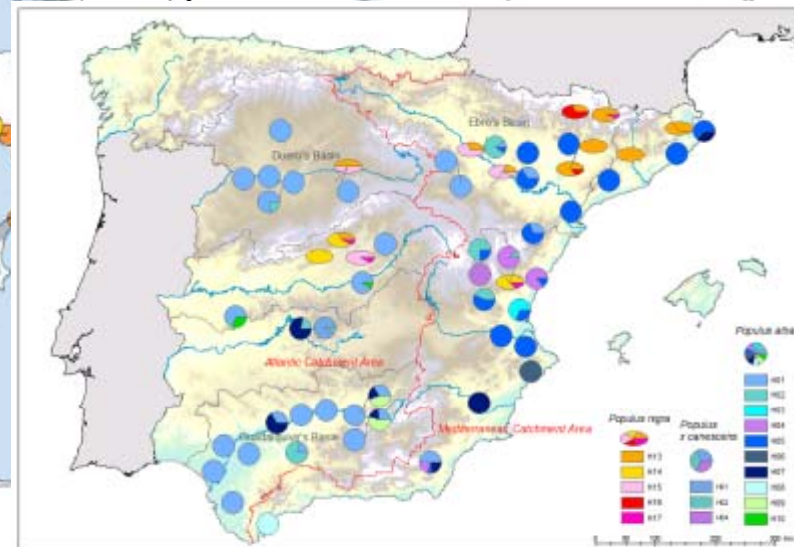
*Quercus suber*



*Cupressus sempervirens*

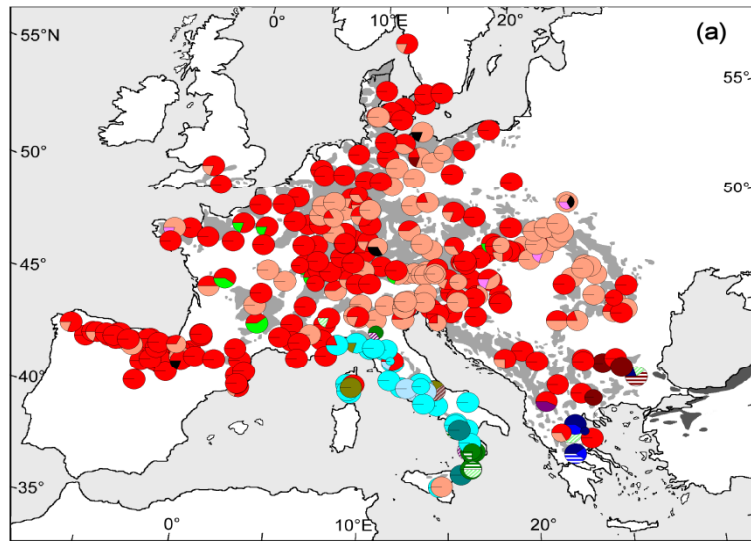


*Populus alba*

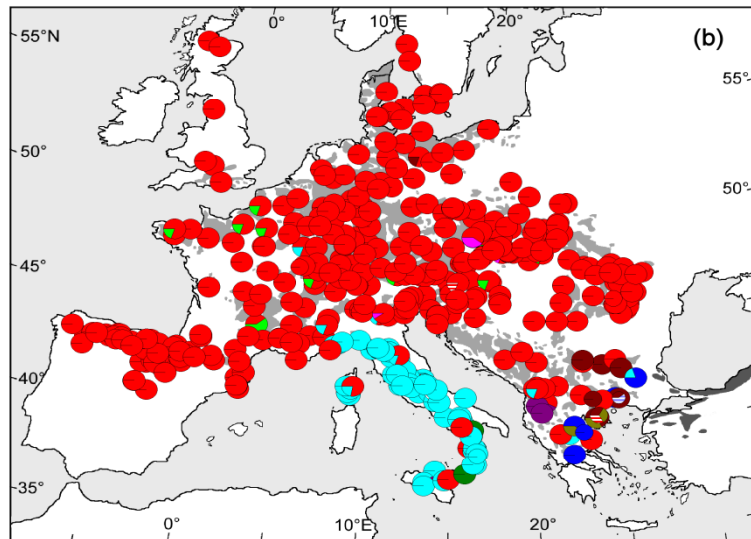


(Magri et al. 2007 Mol. Ecol.)

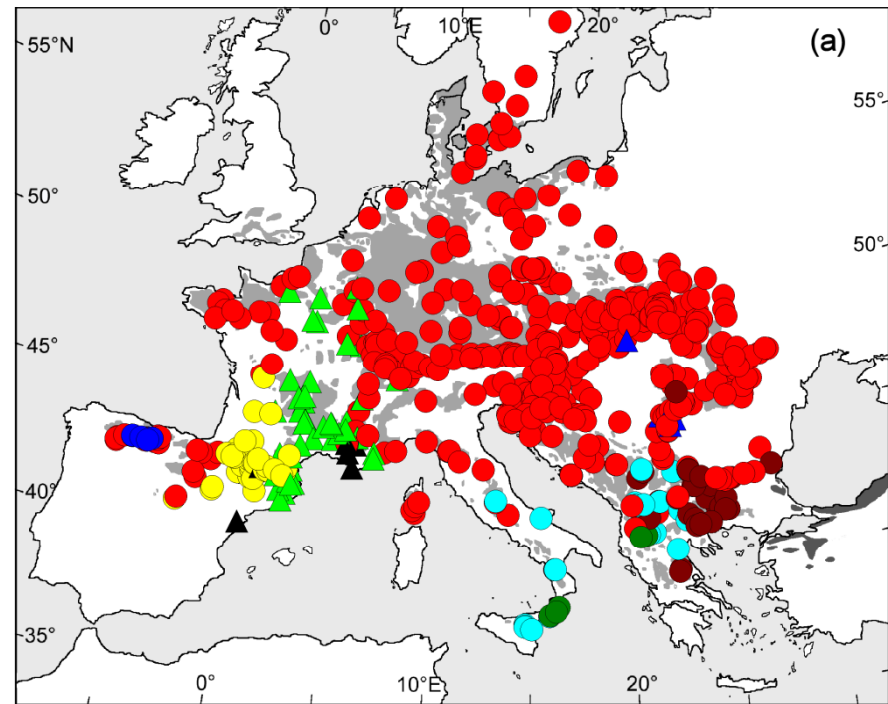
# *Fagus sylvatica* (Magri *et al.* 2006)



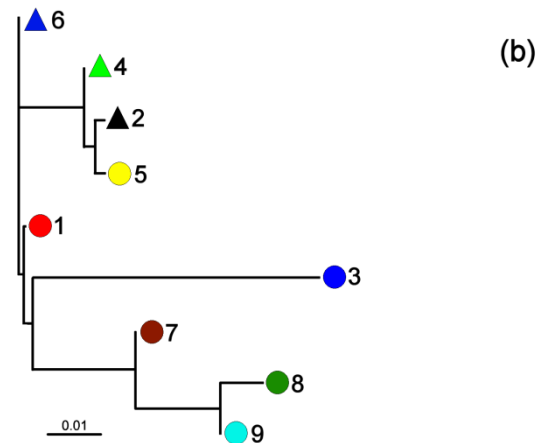
haplotype 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20



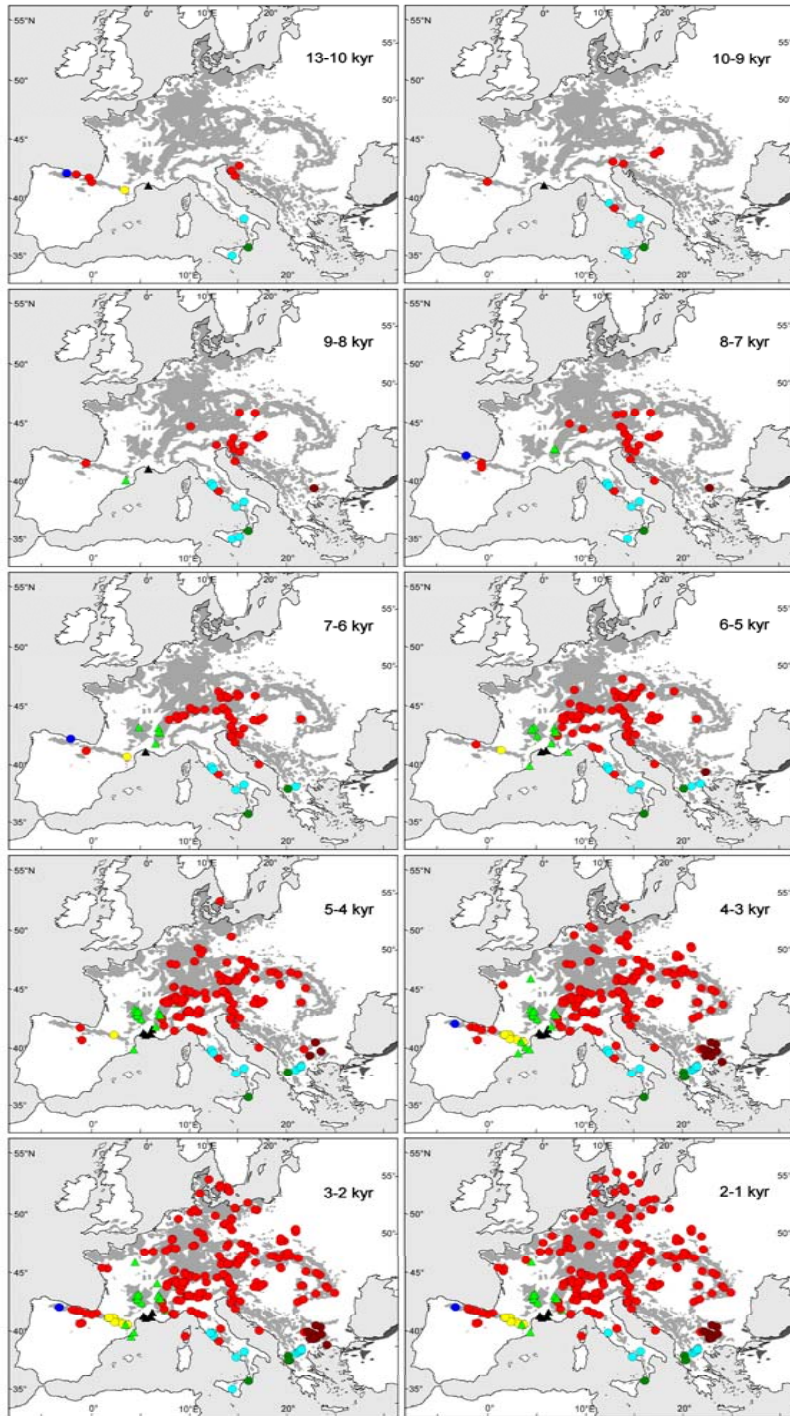
haplotype 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20



isozyme group 1 2 3 4 5 6 7 8 9

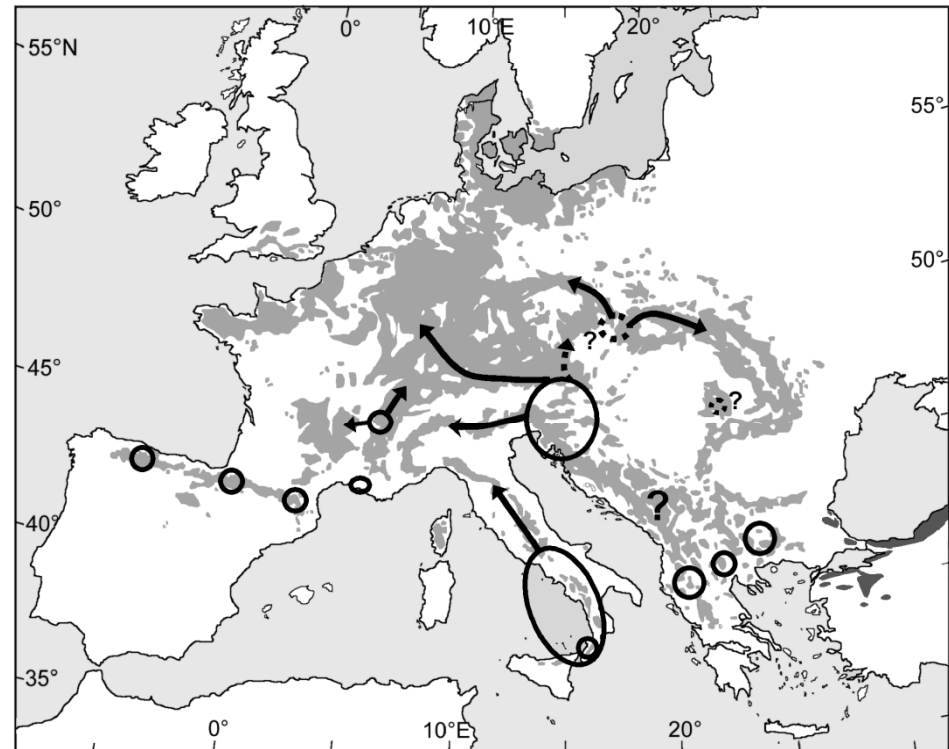


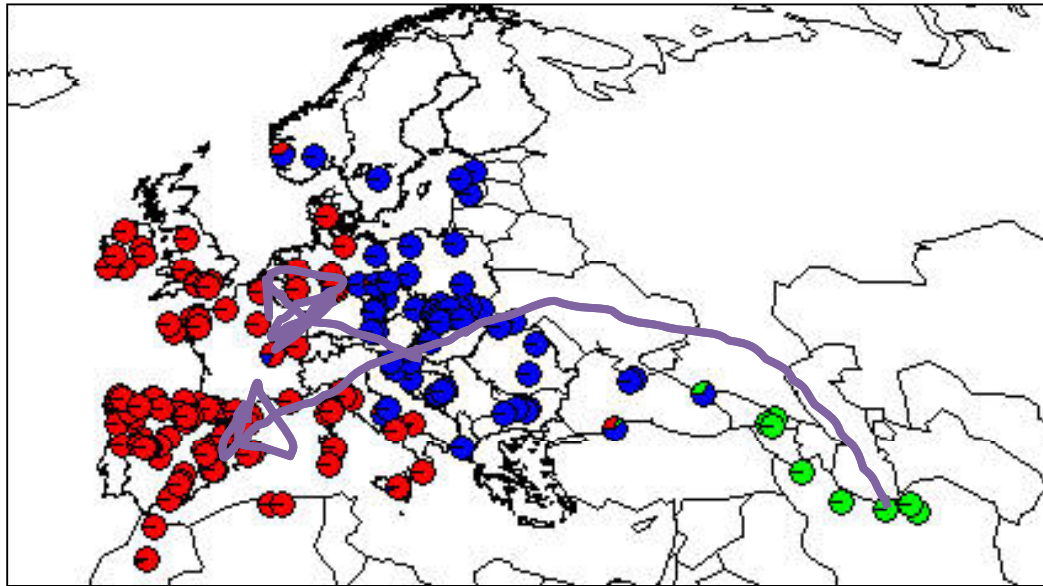




# *Fagus sylvatica*

(Magri *et al.*, 2006)

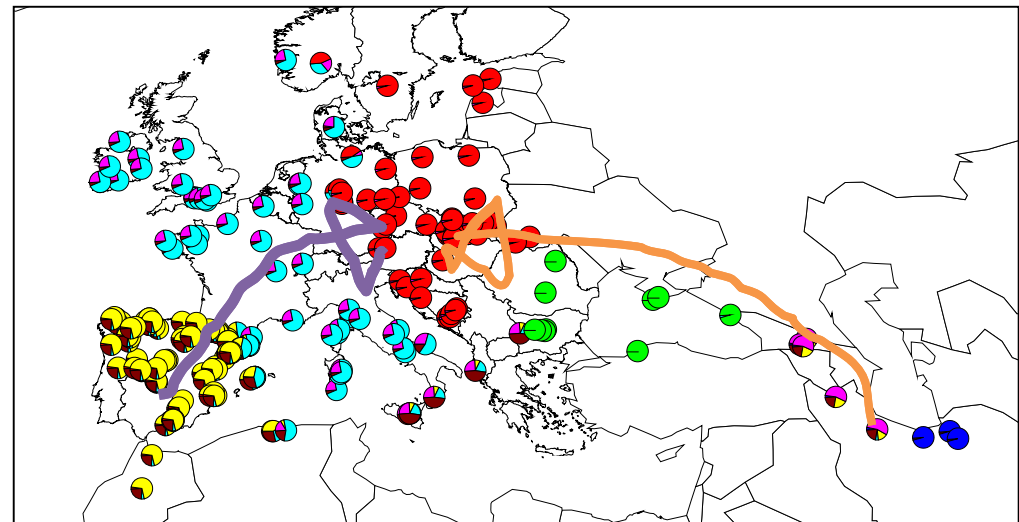




## *Taxus baccata*

**BAPS** (Bayesian  
Analysis of Population  
Structure)

**Geneland**

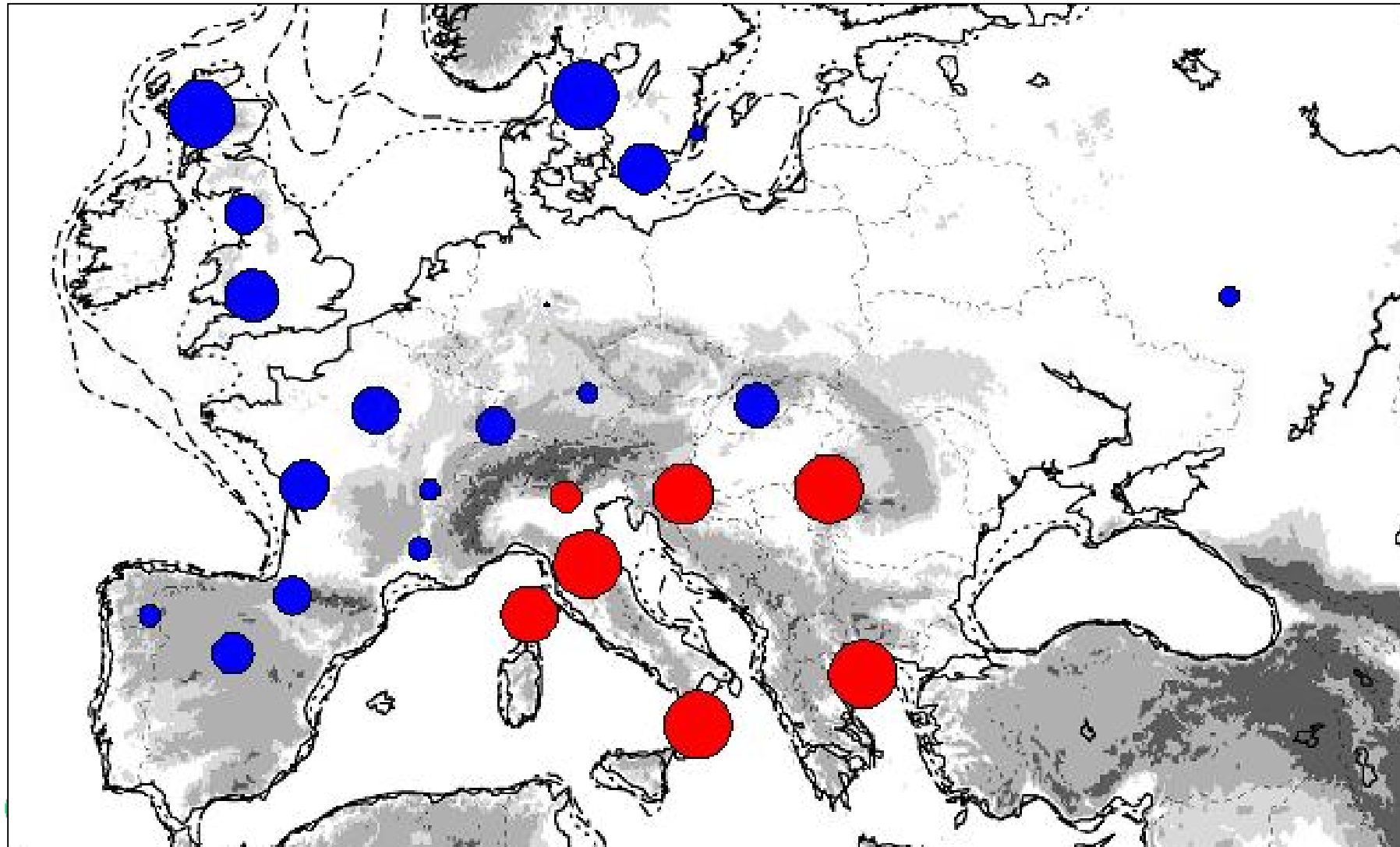




Refugia are characterised by **high divergence** but low levels of within population diversity



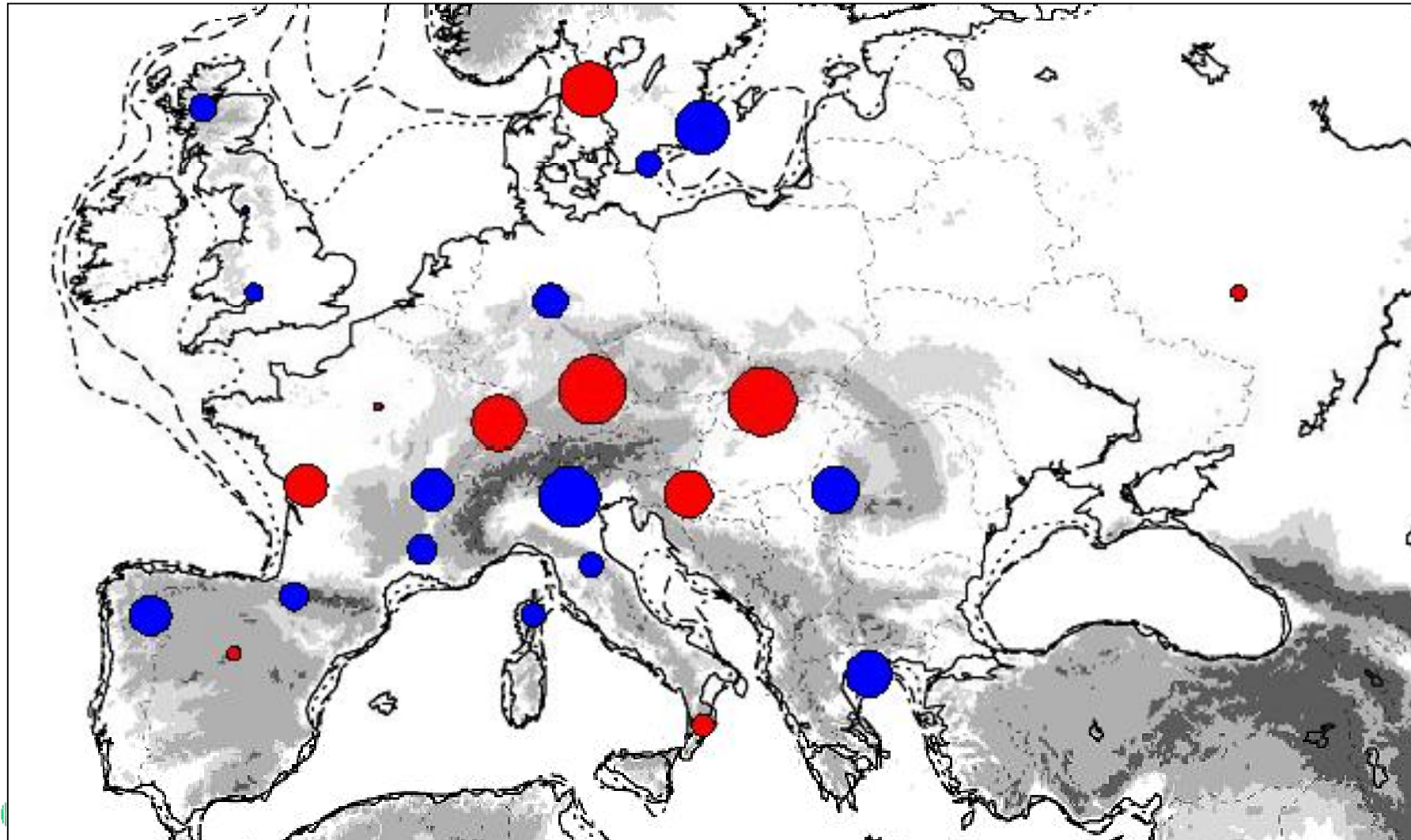
Petit et al., 2003



Intrapopulation **diversity** increases immediately north of the southern refugia (admixture, hybridisation...)

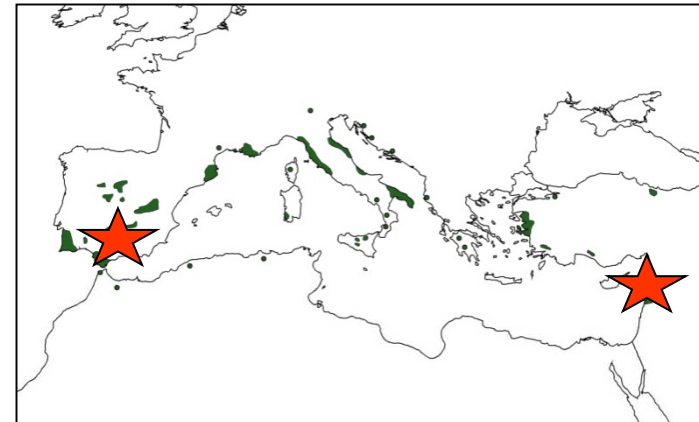


Petit et al., 2003

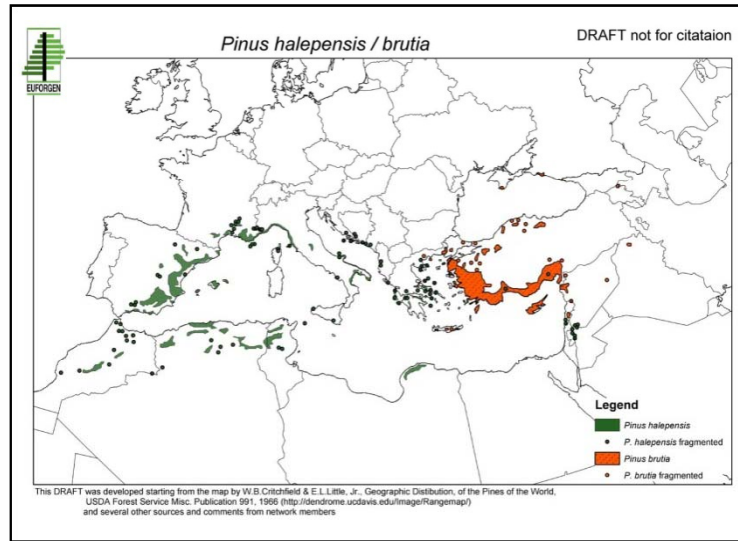


## *Pinus pinea*

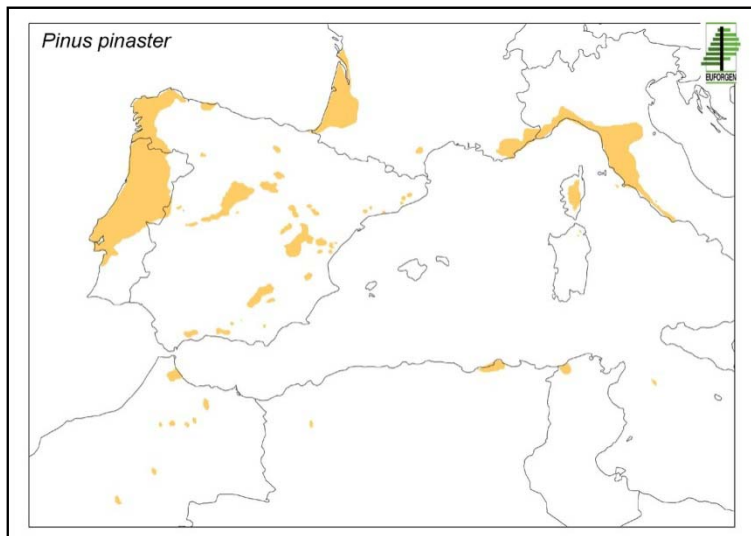
- This tree has been used since ancient times for timber, landscape architecture and for its large edible seeds. Consequently, it was extensively planted around the Mediterranean by Etruscans, Greeks, Romans and Arabs
- Due to its intense cultivation since well before the pre-Roman age, its natural range is now difficult to circumscribe
- Palaeobotanical records pre-dating its cultivation are very scarce, a possible indication that its large present-day distribution is of recent and artificial nature; however, macrofossils and charcoals indicate that *P. pinea* was present in **Spain** before the last glacial maximum and in **Lebanon** long before the coming of man



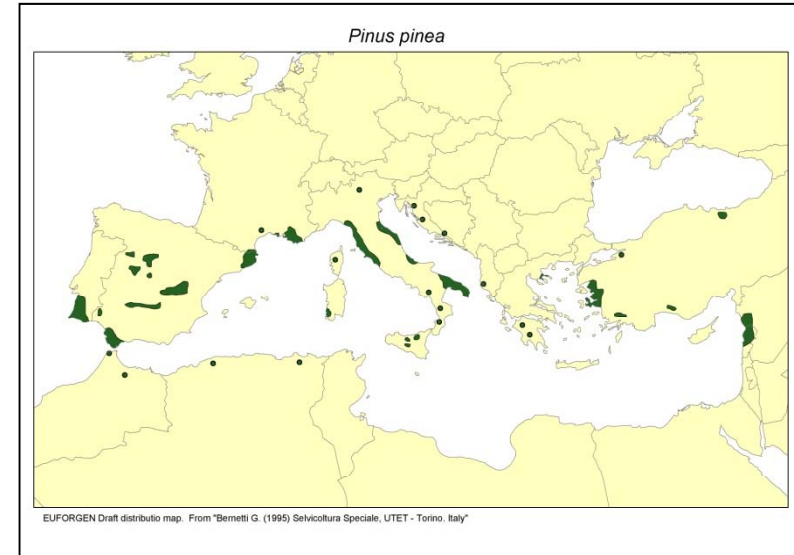
# *Pinus halepensis*



# *Pinus pinaster*



# *Pinus pinea*

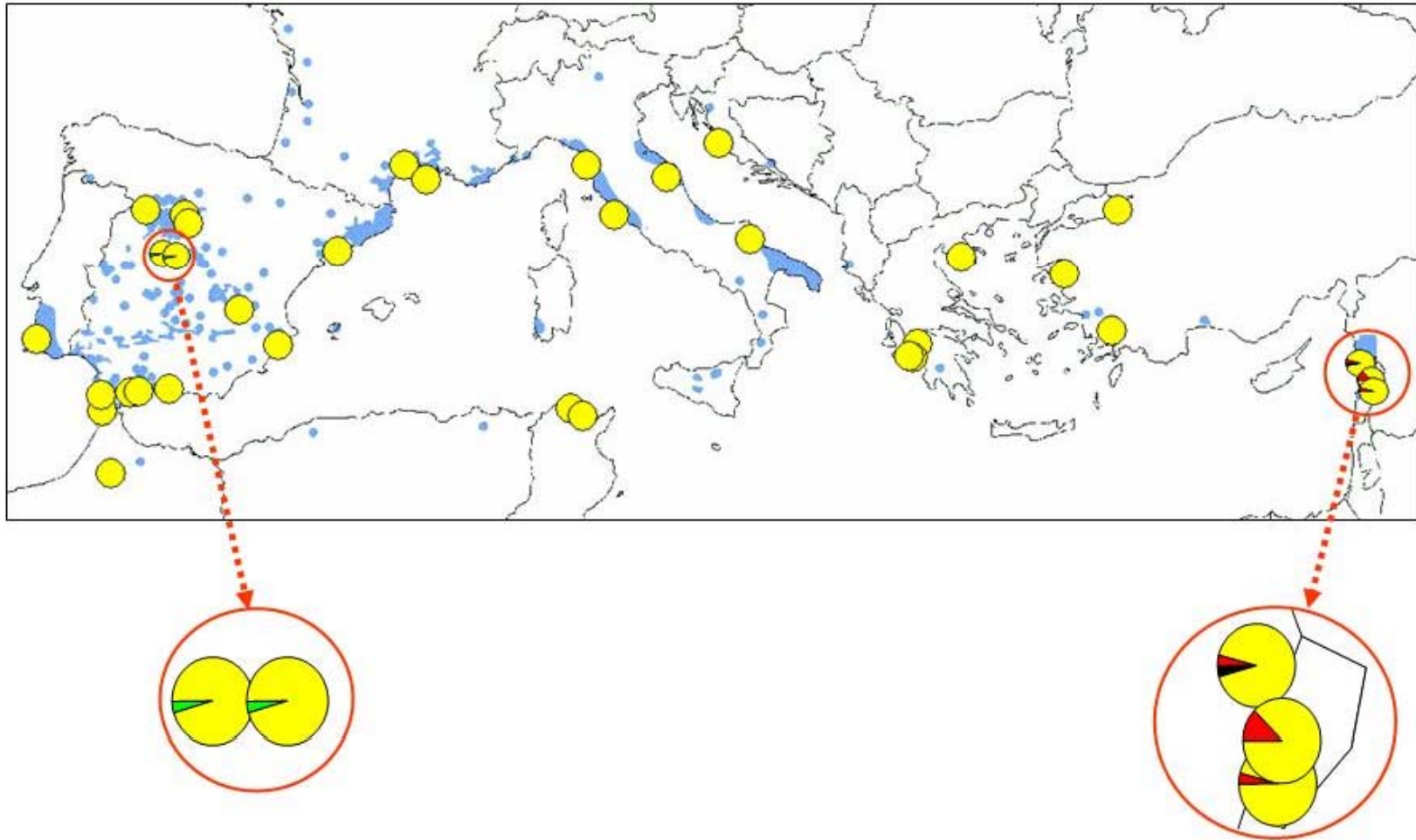


- 55 populations analysed
- 1400 individuals analysed
- 13 cpSSRs used (displaying high variation in other pine species)



# Chloroplast SSR variation in *Pinus pinea*

*Vendramin et al. (2008), Evolution*





# Variation at chloroplast microsatellites in *Pinus pinea* compared to that in other conifers



	<i>Pinus pinea</i>	Other conifers <sup>1</sup>
Number of species	1	8
Number of loci per species (>6 repeats)	13	4.9
Proportion of polymorphic loci <sup>2</sup>	23%	71%
<i>H</i> (haplotypic diversity)	0.019	0.37
$\hat{A}$ (mean number of alleles/locus)	1.23	4.62
$\hat{A}_{[100]}$ <sup>3</sup>	1.06	3.90
$\hat{A}_{[100;10]}$ <sup>4</sup>	1.08	2.58
Mean number of repeats per locus	10.0	13.0
Maximum number of repeats	14.0	15.1

<sup>1</sup> Data for eight species combined: *Abies alba*, *Picea abies*, *Pinus brutia*, *P. cembra*, *P. halepensis*, *P. lambertiana*, *P. mugo* and *P. pinaster*; see ref. 33.

<sup>2</sup> Proportion of polymorphic loci computed without any threshold.

<sup>3</sup> Mean number of alleles per locus in equal-sized samples of 100 individuals.

<sup>4</sup> Mean number of alleles per locus in equal-sized samples of 100 individuals after standardization to a mean number of repeats per locus of 10.



Vendramin et al. (2008), *Evolution*

# List of genetically depauperated but widespread plant species, excluding clonal and self-fertilising plants, ranked by order of increasing diversity

Species <sup>1</sup>	Family	Stature <sup>2</sup>	Habit <sup>3</sup>	Pop <sup>4</sup>	Loci <sup>5</sup>	$H_{es}$ <sup>6</sup>	Reference
<b><i>Pinus resinosa</i></b>	Pinaceae	32	W	2	27	0.001	(27)
<b><i>Berchemia berchemiaefolia</i></b>	Rhamnaceae	17	W	4	14	0.001	(64)
<i>Schwalbea americana</i>	Scrophulariaceae	0.6	H	13	15	0.006	(65)
<b><i>Pinus pinea</i></b>	Pinaceae	30	W	17	32	0.015	(16)
<i>Lespedeza capitata</i>	Fabaceae	1.6	H	12	34	0.020	(66)
<b><i>Juglans cinerea</i></b>	Juglandaceae	30	W	9	12	0.029	(67)
<i>Heuchera americana</i>	Saxifragaceae	0.6	H	12	14	0.039	(68)
<i>Desmodium nudiflorum</i>	Fabaceae	0.29	H	5	13	0.043	(69)
<b><i>Tsuga canadensis</i></b>	Pinaceae	35	W	17	10	0.043	(70)

<sup>1</sup> Trees in heavy font

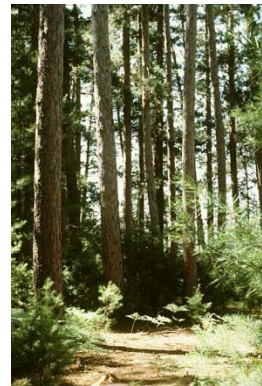
<sup>2</sup> Stature in meters

<sup>3</sup> Habit (W: woody, H: herbaceous)

<sup>4</sup> Pop: number of populations sampled

<sup>5</sup> Loci: number of loci scored

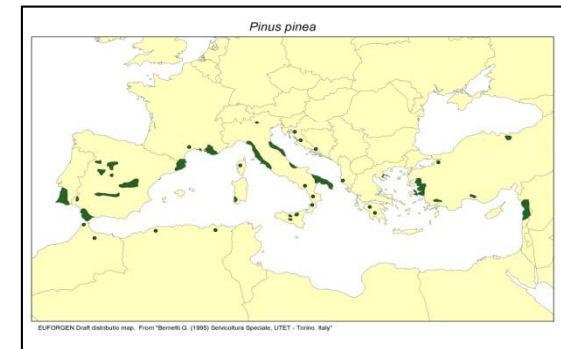
<sup>6</sup>  $H_{es}$ : expected heterozygosity at the species level



*Vendramin et al. (2008), Evolution*

*This extreme situation is probably the result of a combination of factors that have acted in the same direction*

➤ fragmented range



➤ mating system of the species

➤ ability to disperse seeds

➤ human impact

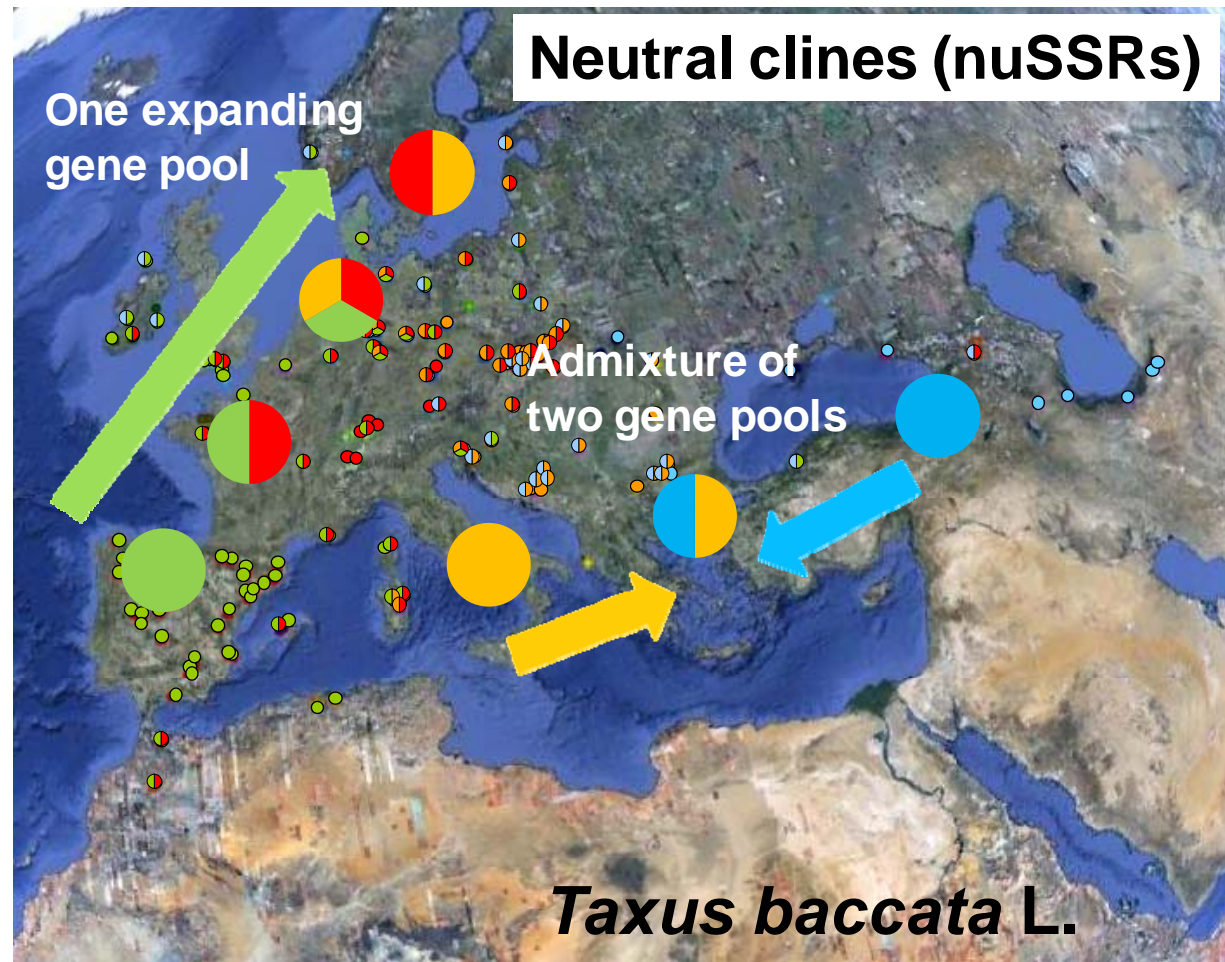




# Some genetic particularities of Mediterranean forests

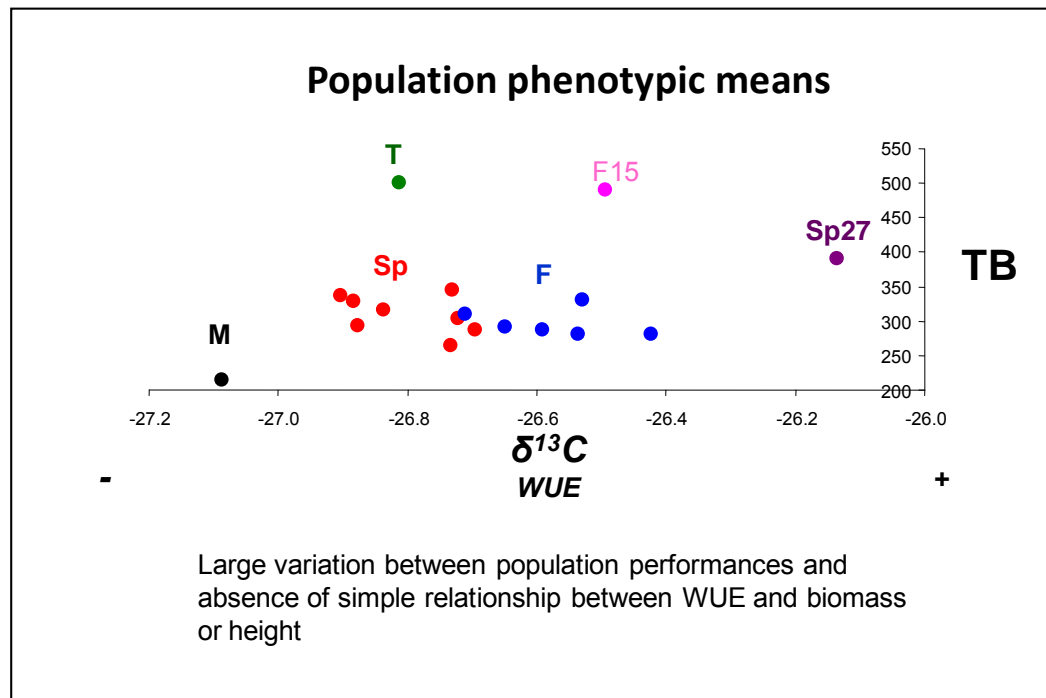
Establishment of 'neutral' genetic clines  
(i.e. due to historical factors)

Historical factors are often confounded with adaptive clines tracking variation of environmental variables (e.g. variation of photoperiod with latitude)



# Traditional common garden experiments have provided valuable insights for Mediterranean trees

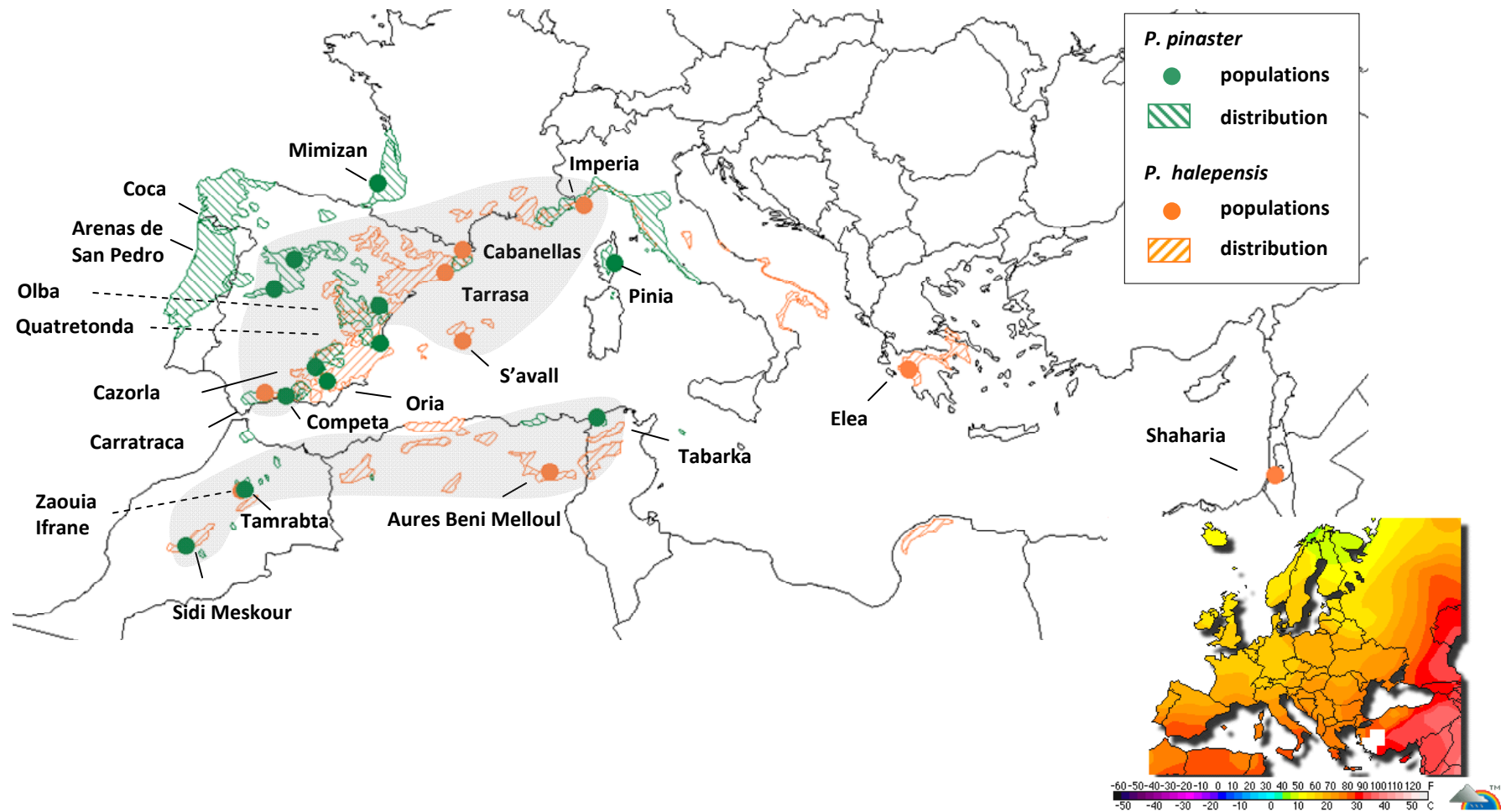
- Large amount of genetic variation is contained within populations.
- In addition, for most species there is also relevant geographical variation.
- Good and relatively fast response to selection.
- The local population is not always the one with the best performance.



**Example: *P. pinaster*  
multisite combined  
provenance-progeny tests**



# Molecular approaches based on putatively functional markers (= candidate genes)







# OBJECTIVES



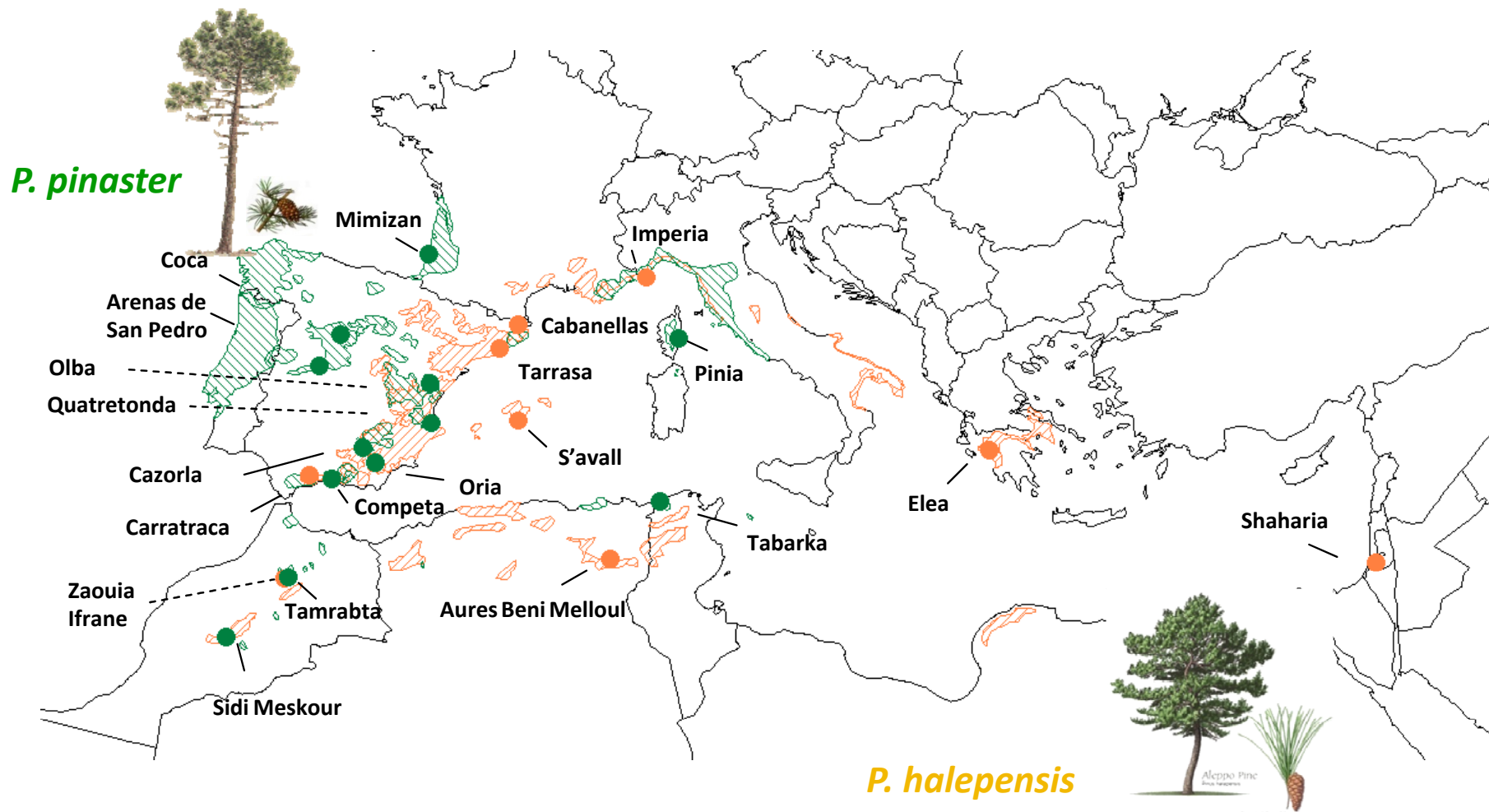
Using a combined strategy we investigated patterns of polymorphism for a set of drought-related candidate genes in maritime pine (*Pinus pinaster* Ait.) and in Aleppo pine (*Pinus halepensis* Mill.)

In details:

- i) examine level and distribution of diversity
  - ii) confirm that some of these genes are under positive selection using new approaches
  - ii) examine which environmental factors are relevant to explain molecular signatures of selection.

Our pluralistic approach provides insights on the adaptive strategy of two conifers that live under the same Mediterranean climate, but present distinct demographic, (re)colonization and life histories (Barbéro et al. 1998)

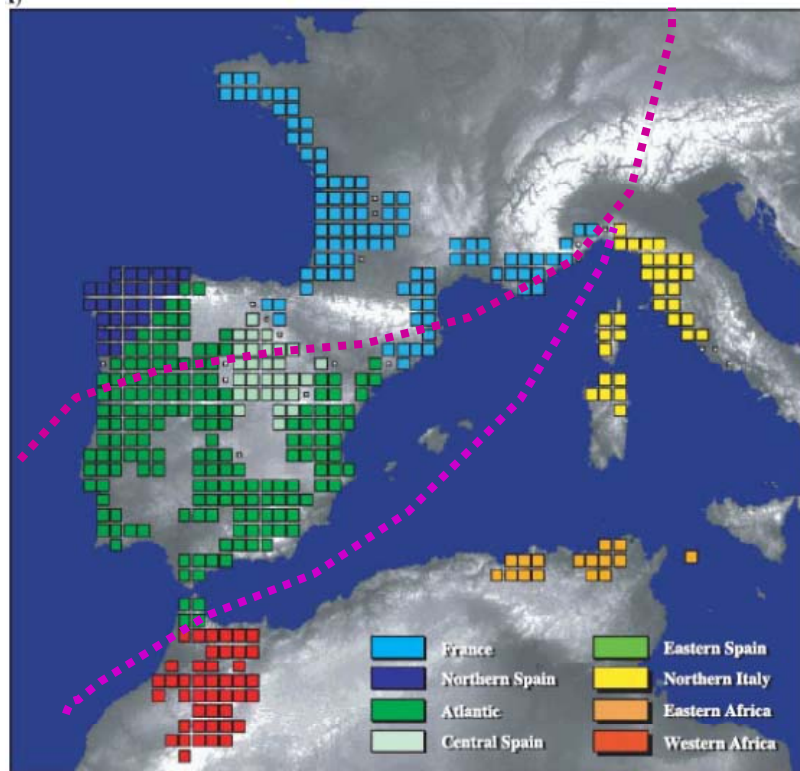
Our **model system**: two widespread Mediterranean pines with contrasting demography: the Aleppo pine, *Pinus halepensis*, and the maritime or cluster pine, *P. pinaster*



**Maritime pine** is a paleo-endemic from the western Mediterranean with marked population structure

**Aleppo pine** is a relatively recent colonizer in this range with less complex population structure

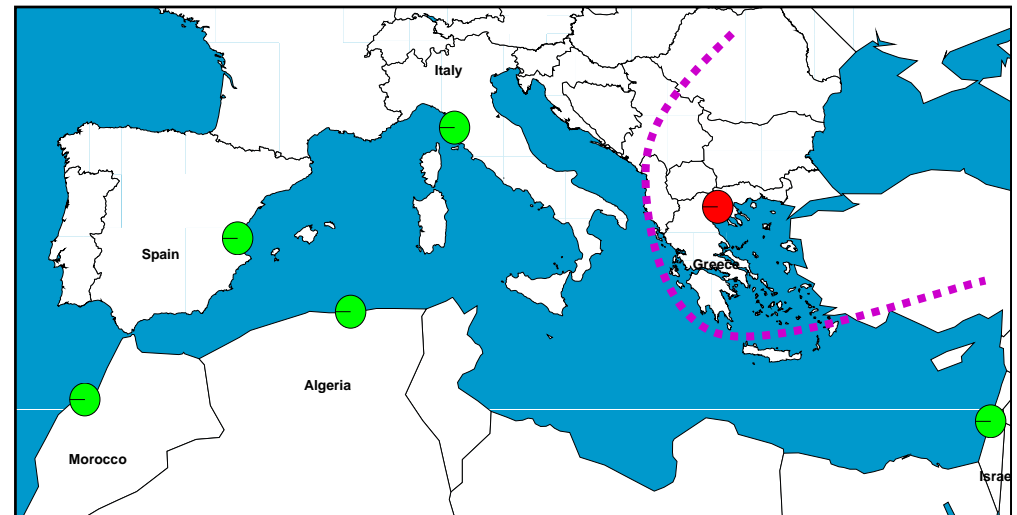
*P. pinaster*



(Bucci et al. 2007 Mol. Ecol., based on cpSSRs)



*P. halepensis*



(Grivet et al. 2009 New Phytol., based on cpSSR)

## Selecting candidate genes to drought tolerance

Information transferred from: *Pinus taeda*  
*Pinus sylvestris*

- Dehydrin 2 (dh-2) PT
  - Dehydrin 2 (dhn-2) PS
  - Dehydrin 5 (dhn-5) PS
- } dehydrin
- ABA and WDS induced gene (lp3.1) PT
  - ABA and WDS induced gene (lp3.3) PT
- } ABA- and WDS-induced (ASR family)
- 4 coumarate coenzyme A ligase (4CL) PT
- } 4-coumarate: CoA ligase

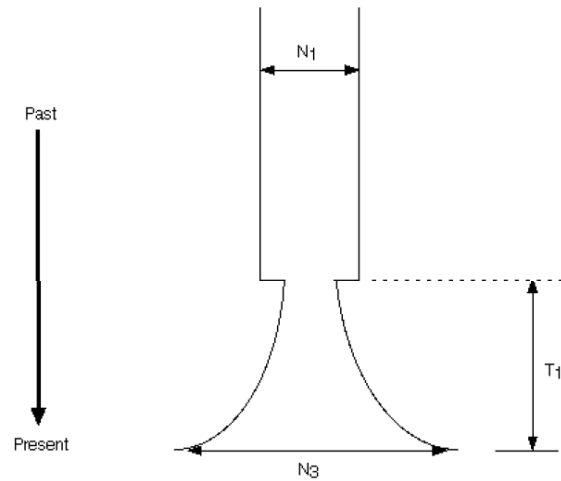
12 amplicons - 533 kb *P. pinaster* - 131 SNP  
- 380 kb *P. halepensis* - 65 SNP

# Some approaches to detect signatures of natural selection on candidate gene loci

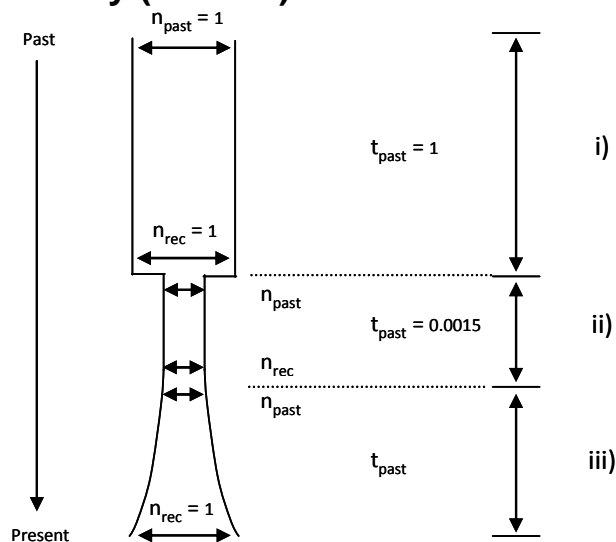
- Neutrality tests with or without outgroups (normally based on the Site Frequency Spectrum, SFS, of mutations)
- Compound tests combining different summary statistics (see Zeng et al. 2007 MBE)
- Ratios of within-species polymorphism to among-species divergence
- Environmental correlations



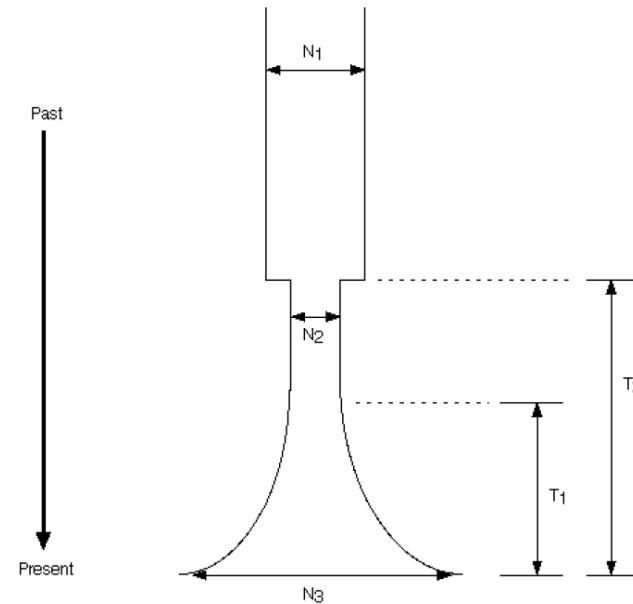
## Bottleneck followed by population growth (BM) model



Recent ( $t_1=0.0005$ ;  $\sim 15,000$ BP) and high-intensity ( $f=0.001$ ) bottleneck



## Bottleneck followed by population growth (BM) model

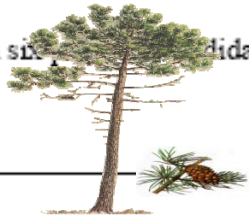



Recent ( $t_1=0.0005$ ;  $\sim 15,000$ BP) to old ( $t_2=0.003$ ;  $\sim 90,000$ BP) and high-intensity ( $f=0.001$ ) bottleneck

Estimates:

- Tajima's D
- Fay & Wu's H

Table 2. Gene diversity (all sites) for 12 amplicons from six candidate genes across the twelve *P. pinaster* and the nine *P. halepensis* populations.

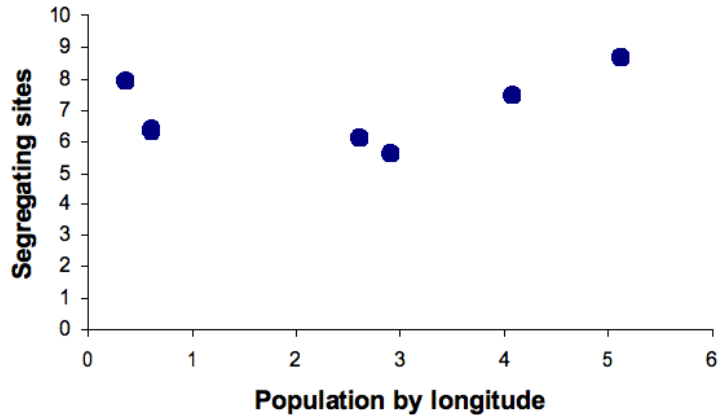
Amplicon	<i>P. pinaster</i>							<i>P. halepensis</i>						
	<i>N</i>	<i>L</i>	<i>S</i>	$\theta_{\pi}$	$\theta_w$	<i>K</i>	<i>He</i>	<i>N</i>	<i>L</i>	<i>S</i>	$\theta_{\pi}$	$\theta_w$	<i>K</i>	<i>He</i>
<i>lp31-Pt</i>														
<i>a</i>	122	456	12	6.26	4.89	6	0.717	79	353	3	2.48	1.72	4	0.460
<i>b</i>	111	560	7	3.91	2.37	3	0.446	74	488	7	3.93	2.94	4	0.464
<i>lp33-Pp</i>	97	449	9	4.29	3.89	9	0.798	90	375	1	0.17	0.53	2	0.065
<i>dhn2-Pp</i>														
<i>a</i>	120	472	12	6.06	4.74	11	0.807	89	448	2	0.20	0.88	3	0.088
<i>b</i>	85	596	21	8.42	7.03	11	0.811	93	346	4	0.55	2.26	5	0.105
<i>dhn2-Ps</i>														
<i>a</i>	92	743	27	8.10	7.13	14	0.899	na	na	na	na	na	na	na
<i>b</i>	80	513	15	5.92	5.90	14	0.855	92	457	11	1.93	4.73	6	0.204
<i>dhn5-Ps</i>	77	449	4	4.00	1.81	5	0.735	na	na	na	na	na	na	na
<i>4cl-Pt</i>														
<i>a</i>	91	551	13	5.67	4.64	4	0.678	88	461	10	5.88	4.30	13	0.805
<i>b</i>	92	256	7	3.56	5.37	5	0.291	79	688	21	10.67	6.18	13	0.793
<i>c</i>	87	543	4	0.69	1.46	4	0.341	78	528	3	2.47	1.15	5	0.697
<i>d</i>	na	na	na	na	na	na	na	72	452	3	2.80	1.37	4	0.484
<b>Total</b>	<b>1054</b>	<b>5588</b>	<b>131</b>	<b>-</b>	<b>-</b>	<b>86</b>	<b>-</b>	<b>834</b>	<b>4596</b>	<b>65</b>	<b>-</b>	<b>-</b>	<b>59</b>	<b>-</b>
<b>Mean</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>5.17</b>	<b>4.48</b>	<b>-</b>	<b>0.671</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>3.11</b>	<b>2.61</b>	<b>-</b>	<b>0.416</b>

Higher level of diversity in candidate genes for drought response were present in *P. pinaster* than in *P. halepensis*, despite its narrower range in the Mediterranean

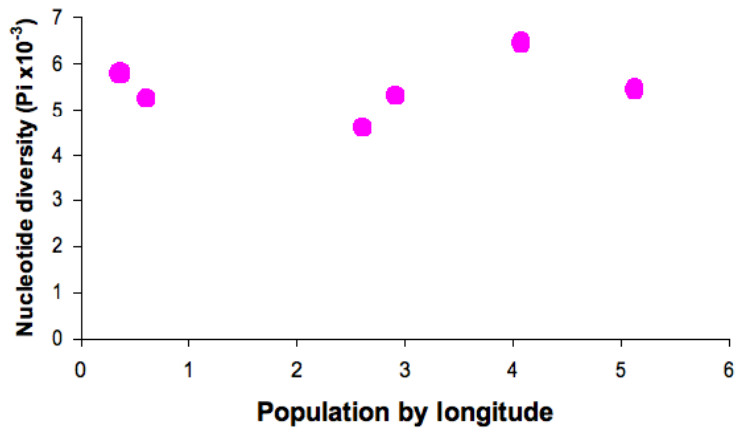
# *Pinus pinaster*



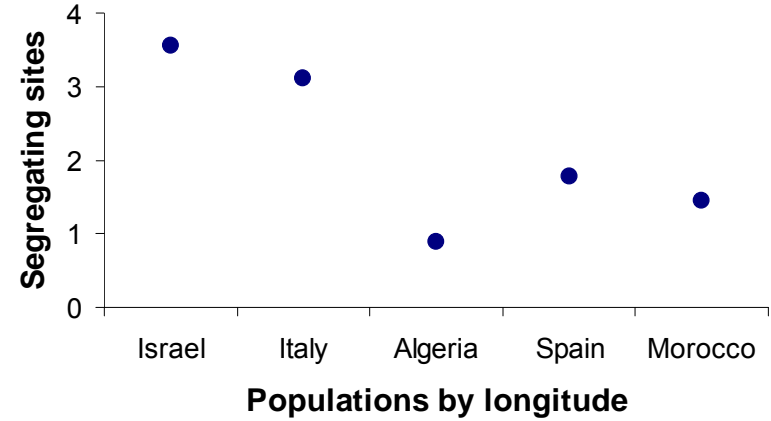
Quatretonda  
Olba  
Oria  
Cazorla  
Coca  
Arenas de San Pedro



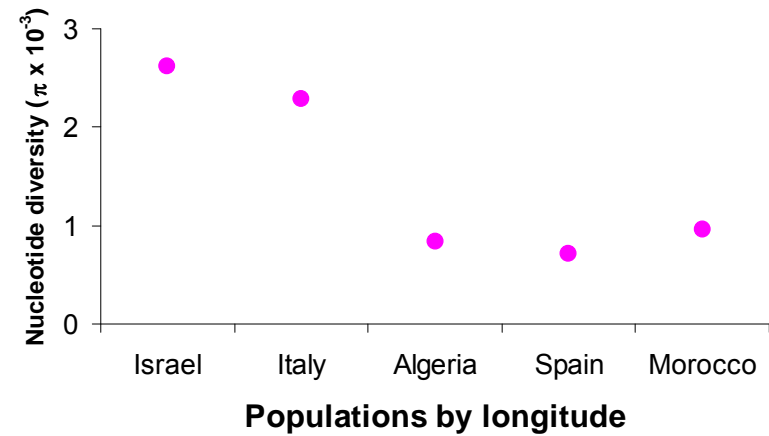
Ancient  Recent



# *Pinus halepensis*



Ancient  Recent

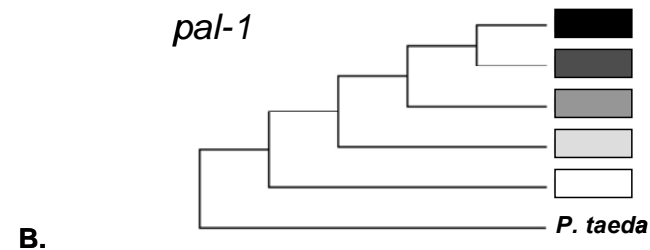
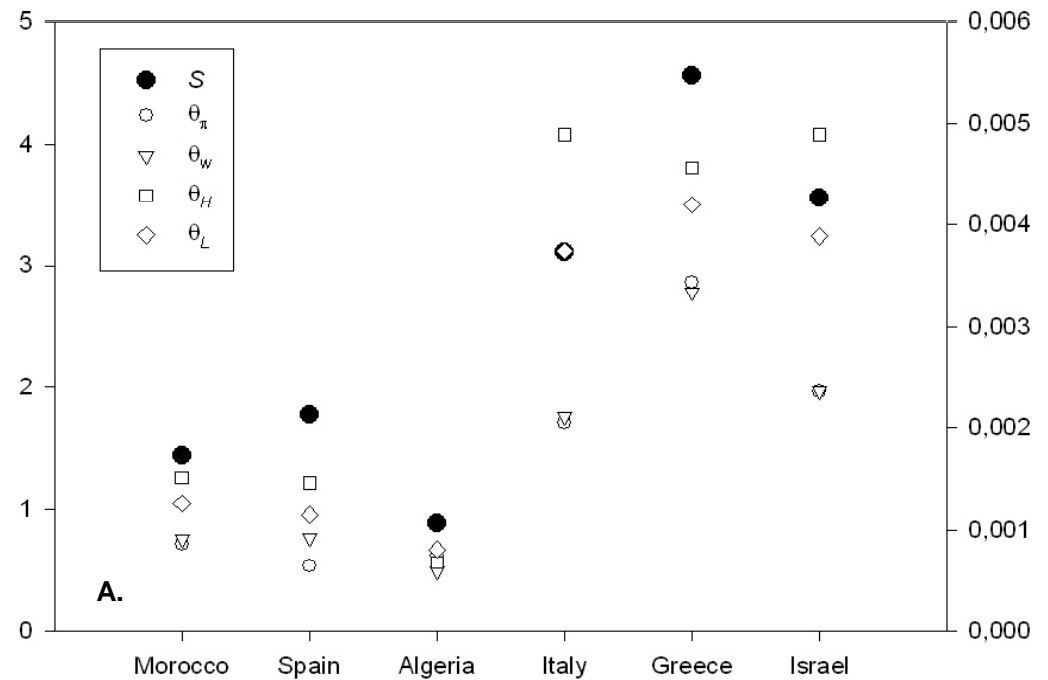


# In addition in *P. halepensis*

(Grivet et al. 2009 New Phytol.)

The low level of diversity in *P. halepensis* is more noticeable in its western distribution where most genes were fixed or almost fixed for particular haplotypes

This is probably a consequence of long-range colonization of the WM from ancient populations in the easternmost edge

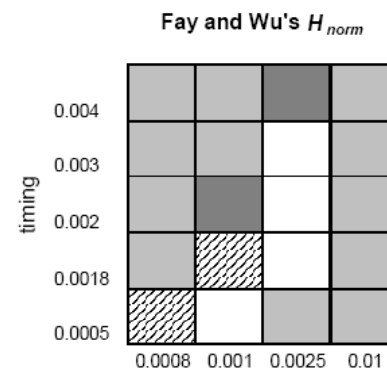
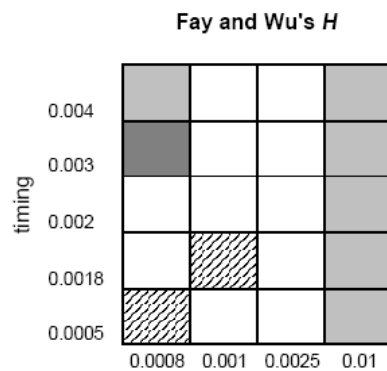
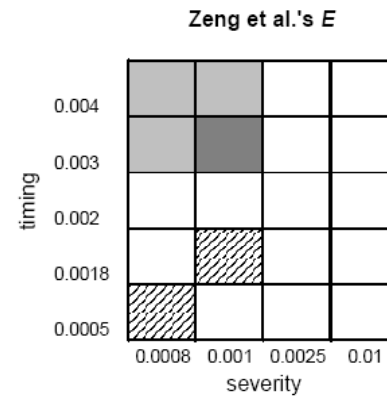
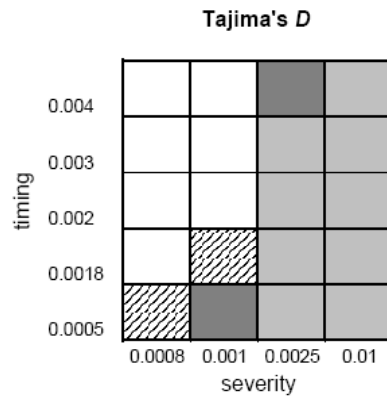




# Demography and selection in Aleppo and maritime pines



Signatures of intense and very recent bottlenecks in Aleppo pine, not so old as the one estimated for maritime pine



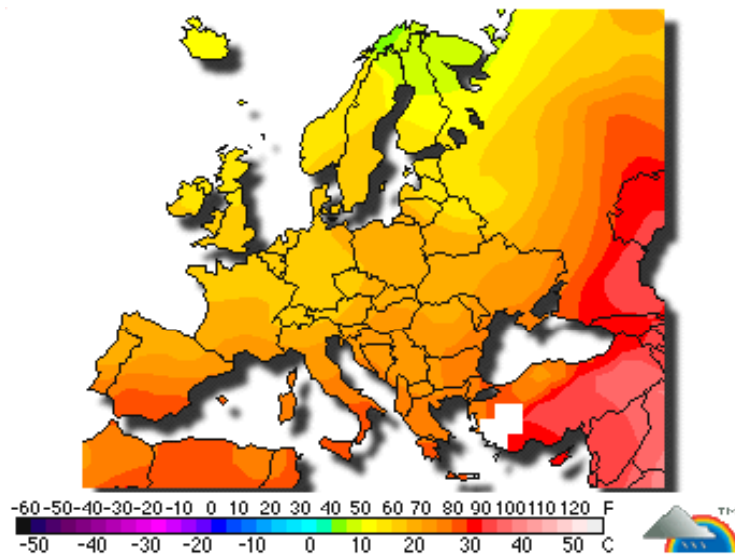
Maritime pine seems to harbour large amount of diversity due to a more stable demography

In addition, because of its mesic distribution, higher environmental heterogeneity would have resulted in contrasted selective pressure that may have increased general levels of diversity



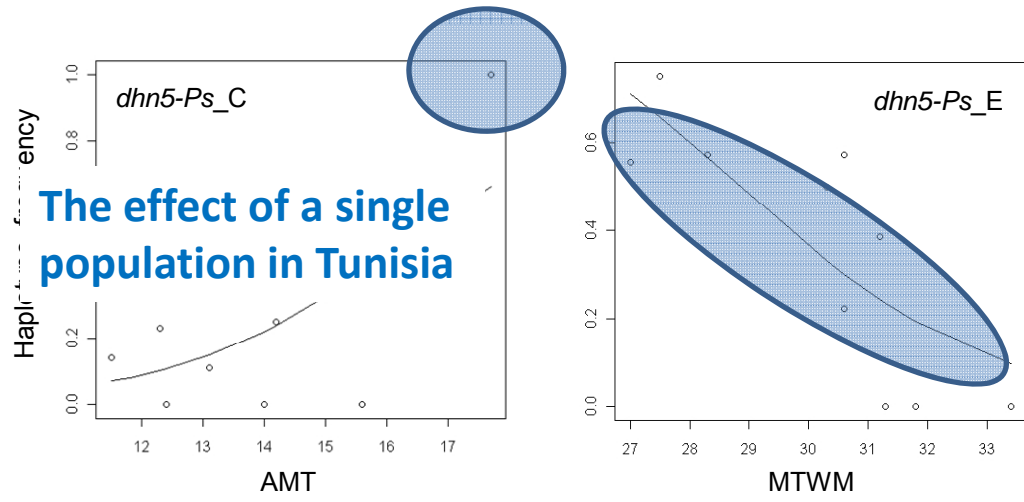
# Correlation with environmental variables at the wide-range scale in pine

Logistic regression for candidate genes related to drought tolerance in maritime and Aleppo pines (Grivet et al. 2010 MBE)



- Analyses both at the SNP and the haplotype level
- Confounding effects → correction for latitudinal gradients caused by historical factors (e.g. postglacial migrations) using **neutral markers as covariates** in the models

# *P. pinaster*

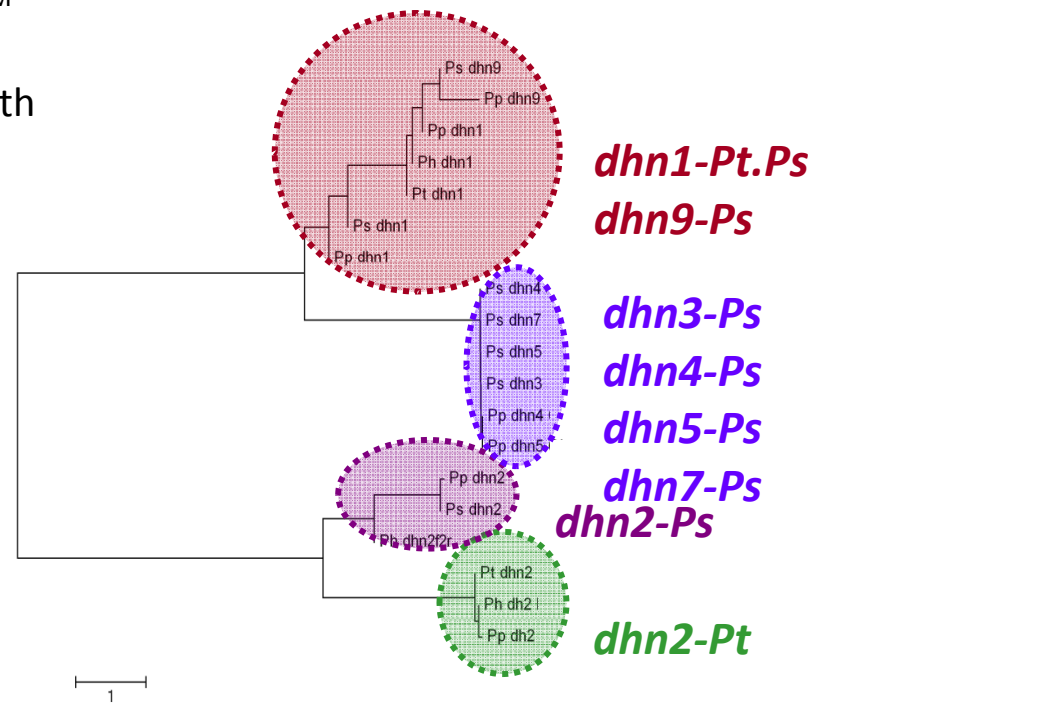


The effect of a single population in Tunisia

But good correlation in other cases

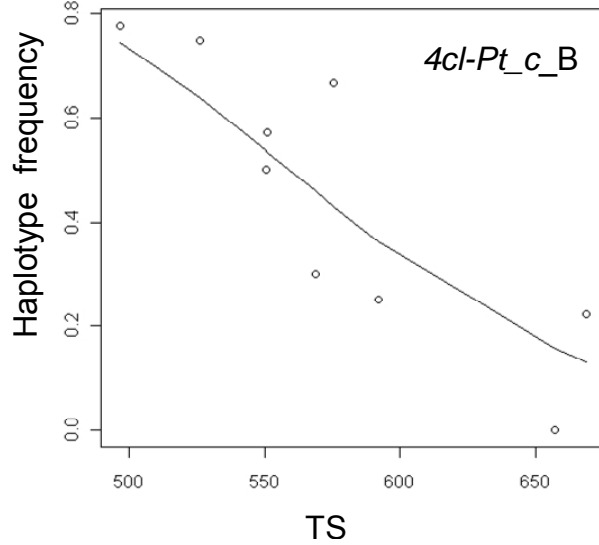
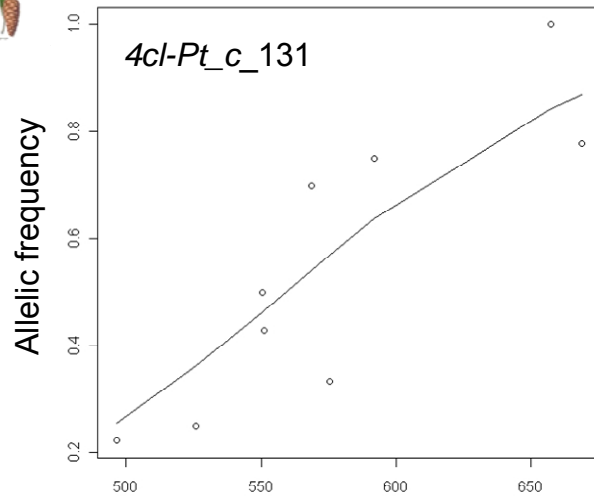
**AMT:** annual mean temperature  
**MTWM:** maximum temperature of the warmest month

**Dhn**, a gene family related to response to abiotic stress, including drought and cold temperature

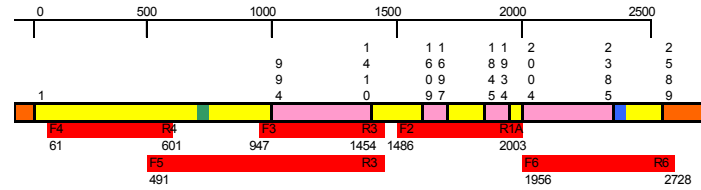




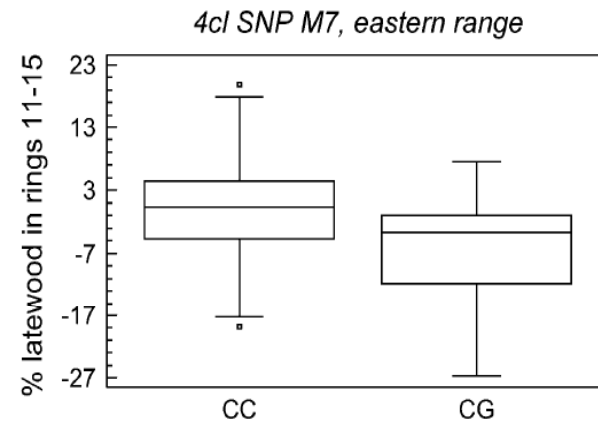
# *P. halepensis*



TS: temperature seasonality



**4cl**, a key gene in lignin formation (and other important metabolic processes)...



... found to be associated with wood properties in pine (González-Martínez et al. 2007 Genetics)

	<i>D</i>	<i>H<sub>norm</sub></i>	<i>DHEW</i>	<i>ML-HKA</i>	Association
<i>lp31-Pt_a</i>		✓			
<i>lp31-Pt_b</i>	✓	✓			
<i>lp33-Pp</i>					
<i>dhn2-Pp_a</i>					
<b><i>dhn2-Pp_b</i></b>				✓	
<i>dhn2-Ps_a</i>					
<i>dhn2-Ps_b</i>					
<b><i>dhn5-Ps</i></b>					✓
<i>4cl-Pt_a</i>					
<b><i>4cl-Pt_b</i></b>		✓	✓		
<i>4cl-Pt_c</i>					✓
<i>4cl-Pt_d</i>					

*P. pinaster*

Western Mediterranean

North African



**5 amplicons potentially under selection**

	<i>D</i>	$H_{norm}$	<i>DHEW</i>	<i>ML-HKA</i>	Association
<i>lp31-Pt_a</i>					
<i>lp31-Pt_b</i>					
<b><i>lp33-Pp</i></b>		✓		✓	
<i>dhn2-Pp_a</i>					
<i>dhn2-Pp_b</i>					
<i>dhn2-Ps_a</i>					
<i>dhn2-Ps_b</i>					
<i>dhn5-Ps</i>					
<i>4cl-Pt_a</i>					
<b><i>4cl-Pt_b</i></b>				✓	
<b><i>4cl-Pt_c</i></b>					✓
<b><i>4cl-Pt_d</i></b>	✓	✓			

*P. halepensis*

Western Mediterranean

North African



**4 amplicons potentially under selection**



## Conclusions

- Multiple refugia located in southeastern Spain, northern Africa and the Atlantic coast of Portugal
- Restricted spatial range
- No loss of variation during recolonization
- Signatures of bottleneck



- Few refugia mainly located in Greece
- Wide spatial range
- Loss of variation during recolonization
- Signatures of intense bottleneck



## Conclusions

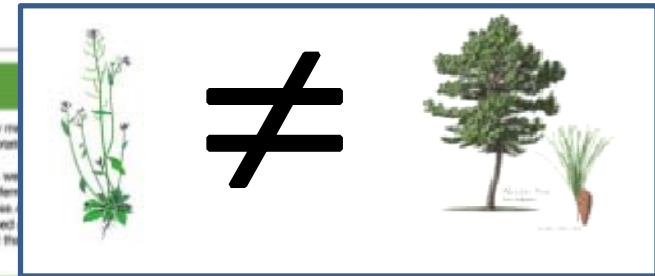
- Some of these events differ between the two species reflecting **individual histories** (recolonization, demography, adaptation)
- Other events are shared, which translates partly as a **common history** of these closely-related and partially sympatric Mediterranean pines



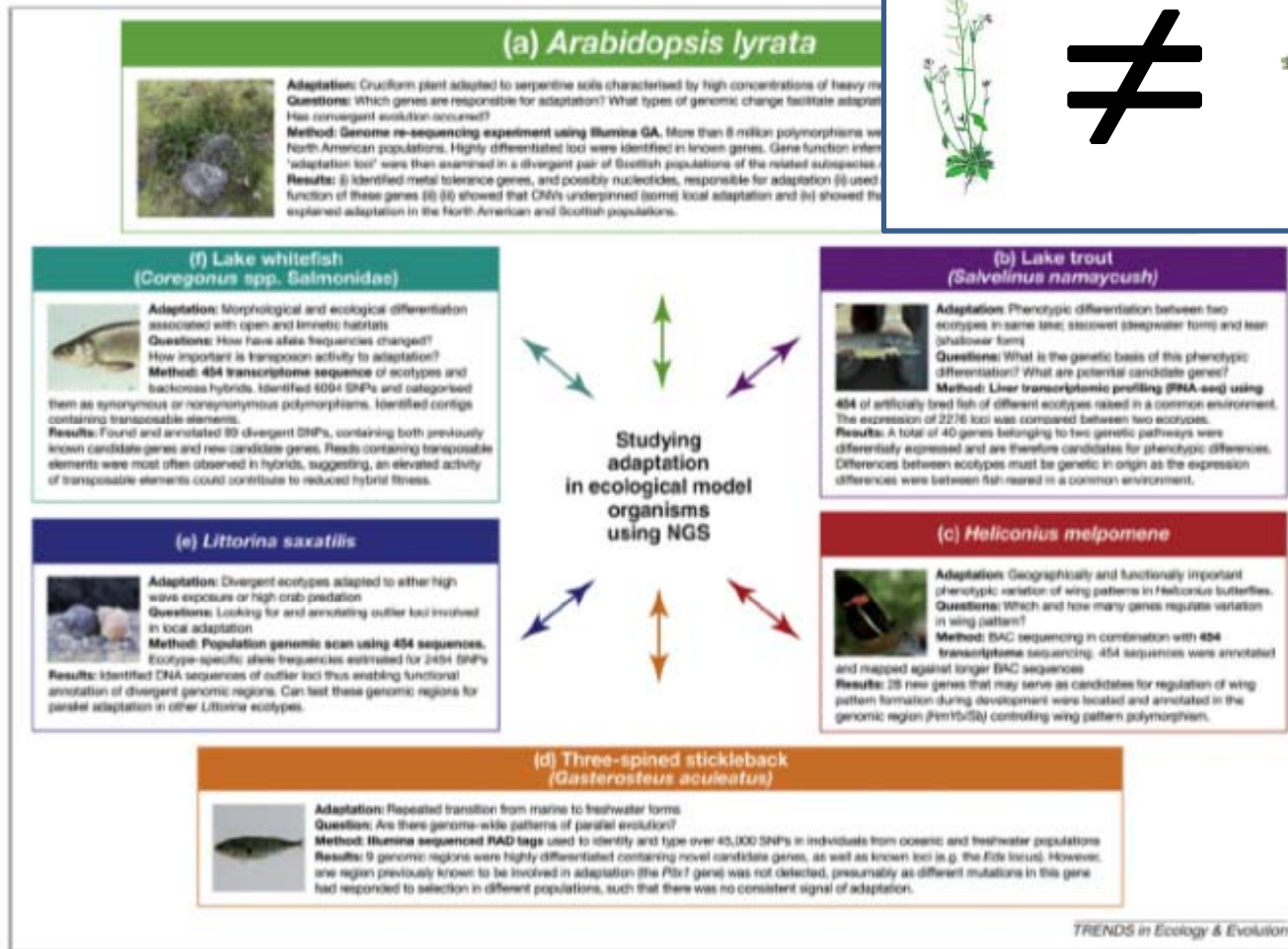
## Conclusions & perspectives

- **Several approaches available** → need basic knowledge on life-history traits, demography and population structure of the target species.
- **Environmental correlation** approaches are promising to identify candidate genes and environmental variables relevant for local adaptation of tree populations.
- Combination with common garden experiments → **association genetics**. Development of a clonal population in maritime pine and a 1536-SNP array (EU NovelTree project, [www.noveltree.eu](http://www.noveltree.eu)).
- Use of **HT-sequencing** and new technologies (454, GAll, Illumina) to upscale research based on population genetics of candidate genes (see, for example, Turner et al. 2010 Nature Genetics). **CRIEC** (S.González-Martinez & G.G. Vendramin): 280 candidate genes in 18 conifers.

# Need to increase availability of genomic resources for non-model species



(Stapley et al. 2010 TREE)



High throughput (HT) Sequencing (454, GAllx, Illumina) → An excellent opportunity for forest trees!



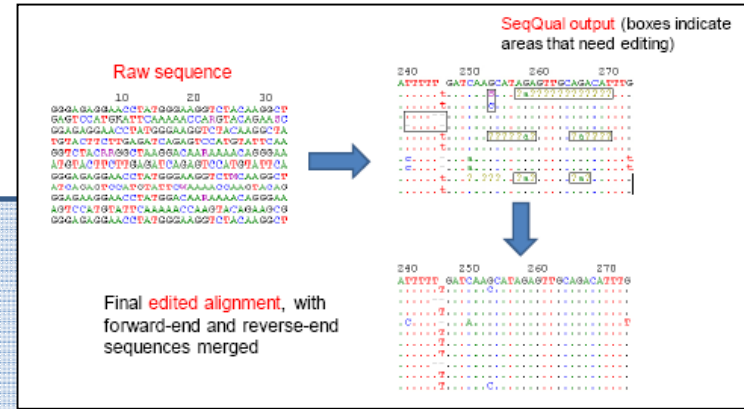
# Resequencing in Conifers (CRIEC initiative)



280 amplicons

(in collaboration with UC Davis (David Neale) ACE-SAP project)

	<i>N</i>	# of polymorphic amplicons (%)	# of 'well-placed' SNPs
<i>Abies pinsapo</i>	3	6 (13)	
<i>Juniperus communis</i>	3	2 (4)	
<i>Taxus baccata</i>	3	2 (4)	
<i>Pinus leucodermis</i>	7	19 (40)	
<i>Pinus brutia</i>	12	28 (60)	
<i>Pinus halepensis</i>	48	17 (36)	
<i>Pinus nigra</i>	12	36 (77)	46



## Generating candidate gene resources in non-model conifers: The CRIEC initiative



Santiago C. González-Martínez<sup>1</sup>, Giovanni G. Vendramin<sup>2</sup> and CRIEC consortium



<sup>1</sup>Department of Forest Ecology and Genetics, Center of Forest Research, CIFOR-INIA, Madrid, Spain (santiago@inia.es)

<sup>2</sup>Plant Genetics Institute (IGV), CNR, Sesto Fiorentino, Firenze, Italy (giovanni.vendramin@igv.cnr.it)



### Goals and description of CRIEC

In the framework of the EvoTree EU Network of Excellence, the CRIEC consortium has promoted the transference of conserved candidate genes that are expressed under different environmental stresses to several non-model conifers. With this initiative we intend to provide the forest genetics community with tools to conduct demography and selection studies in conifers with ecological importance. We have focused on



# Full genome and transcriptome sequencing



*Taxus baccata*



*Pinus sylvestris*



*Pinus halepensis*



*Pinus pinaster*



Federico



Hadrien



Bruno



David



Santiago

...and for your attention