Can nucleotide diversity patterns in Pinus pinaster lignification transcription factors be explained by demography or selection?

Camille Lepoittevin, François Hubert, Christophe Plomion, Pauline Garnier-géré

To cite this version:

Camille Lepoittevin, François Hubert, Christophe Plomion, Pauline Garnier-géré. Can nucleotide diversity patterns in Pinus pinaster lignification transcription factors be explained by demography or selection?. 12. Congress of the European Society for Evolutionary Biology, Aug 2009, Turin, Italy. 2 p., 2009. hal-02813943

HAL Id: hal-02813943
https://hal.inrae.fr/hal-02813943v1

Submitted on 6 Jun 2020

12th Congress of the European Society for Evolutionary Biology, August 24-29, 2009 –Turin, Italy

Can nucleotide diversity patterns in *Pinus pinaster* lignification transcription factors be explained by demography or selection?

Camille Lepoittevin[1,2], François Hubert[1], Christophe Plomion[1], Pauline Garnier-Géré[1]

[1] INRA, UMR BIOGECO, 69 route d'Arcachon, 33612 Cestas, France
(lepoittevin@pierroton.inra.fr)
[2] FCBA, Laboratoire de Biotechnologie, Domaine de l'Etancon, 77370 Nangis, France.

Lignification is an essential process in trees, playing a key role in sap conducting, mechanical support or biotic and abiotic stress resistance. Although structural genes coding for enzymes of the lignin biosynthetic pathway are well known, we are only beginning to understand the mechanisms that regulate this process. This study aimed at assessing the nucleotide diversity of 9 transcription factors which are thought to play key regulatory roles in conifer lignification, namely *MYB1*, *2*, *4*, *5*, *8* and *14*, *HDZip31*, *LIM2* and *SCL1*. Individuals were sampled in *Pinus pinaster* natural populations from the French Atlantic coast. The "*cis*-regulatory evolution" theory states that transcription factors possess little variation due to the large number of downstream target genes that would be affected, and are submitted to strong purifying selection. In line with this theory, nucleotide diversity observed in the 9 TFs was twice lower than for other structural genes from the lignin pathway. We then modelled a large range of bottleneck *scenari* and tested whether patterns of diversity at each gene departed from expectations. We will present the contrasting results across the different genes and discuss the consistency of the best fitting models with maritime pine species demographic history.

Disentangling diversity patterns due to demography or natural selection in *Pinus pinaster* transcription factors involved in wood formation

Camille Lepoittevin[1,2], François Hubert[1], Luc Harvengt[2], Christophe Plomion[1], Pauline Garnier-Géré[1]

[1] INRA, UMR BIOGECO, 69 route d'Arcachon, 33612 Cestas, France (lepoittevin@pierroton.inra.fr); [2] FCBA, Laboratoire de Biotechnologie, Domaine de l'Etancon, 77370 Nangis, France;

Aims

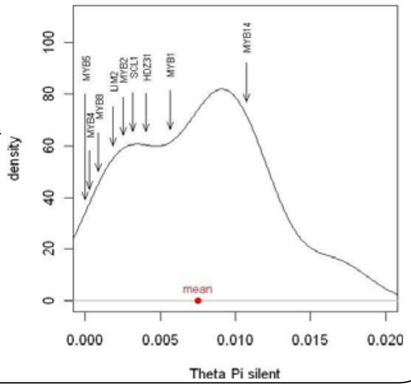
1/ Estimate the nucleotide diversity of 9 transcription factors (TFs) which are thought to play key regulatory roles in conifer lignification

2/ Test whether diversity patterns at each gene depart from neutral expectations.

Sampling and sequencing

- ✓ 42 seeds from the Aquitaine *P. pinaster* population (unstructured)
- ✓ DNA extracted from megagametophytes (haploid tissue)
- ✓ Sequencing of 9 TFs with known or putative implication in conifer wood formation: *MYB1, 2, 5, 4, 8, 14, HDZ31, LIM2* and *SCL1.*

Distribution of $\theta_{\Pi s}$ estimates for the Aquitaine population (based on 25 candidate genes)



Low level of nucleotide diversity in TFs

$\theta_{\Pi s}$ generally lower than that observed for structural genes in the same species…

Cis-regulatory evolution theory : Strong purifying selection on TFs because of the large number of downstream target genes with which they can interact, and thus possible deleterious pleiotropic effects in case of mutation.

Significant departures from the SNM

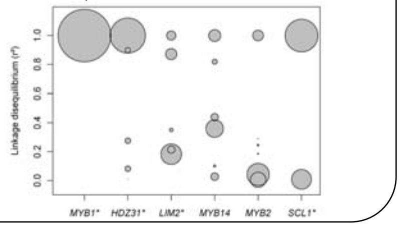
Neutrality tests statistics for 6 TFs

Genes	Tajima's D	Fay & Wu's H	Fu's Fs
Myb1	3.31***	-0.17^{ns}	11.86**
HDZ31	2.27**	0.40^{ns}	8.8**
LIM2	1.71*	-0.49^{ns}	5.49**
Myb14	-0.19^{ns}	-8.36*	7.08**
Myb2	-0.45^{ns}	-5.02*	0.11^{ns}
SCL1	-0.54^{ns}	-1.47^{ns}	1.99^{ns}

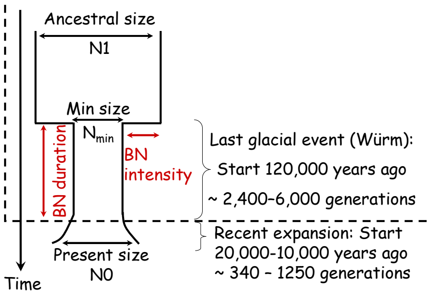
Strong departures from neutral expectations :
✓ Excess of intermediate frequency variants
✓ Lack of haplotypes
✓ High levels of LD

→ Could be due to bottleneck or balancing selection…

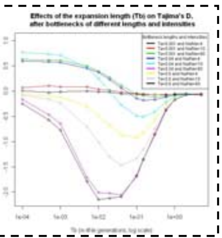
Pairwise linkage disequilibrium values (r^2) between informative sites within six TFs.



Demographic models



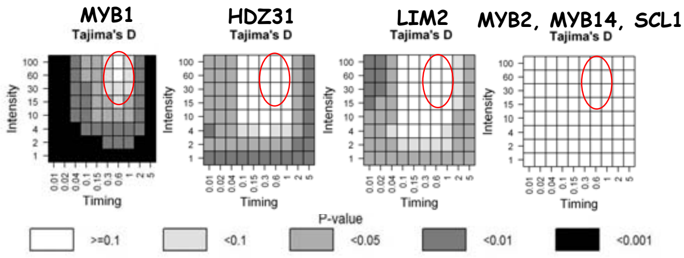
✓A model without expansion is more conservative than a model with an expansion period…



✓ The expansion period is probably too short to leave traces in diversity patterns

✓ Our model : simple contraction, N0 = Nmin

✓ 2 parameters: duration (in 4N0 generations) & intensity (N1/N0)



Best fitting models :

0.15 < Duration < 2 & Intensity > 10

Hypothesis 1:

$$\underbrace{0.0075}_{\text{Mean } \theta\pi s} = 4.\ N0 \times \underbrace{1.7 \times 10^{-9}}_{\mu} \times \underbrace{35}_{\text{Generation time}} \Rightarrow Ne = 31{,}500$$

⇒ 0.7 MYA < Bottleneck timing < 9 MYA

Implies that the population size did not change during the last glacial period!!!

N0 is probably over-estimated: large CI for *Ks*, *θπs* and generation time, fossil calibrations, assumptions of demographic equilibrium after divergence clearly unrealistic…

Hypothesis 2: Start of the BN ~ 2,400-6,000 generations ago > 0.15-2 x 4.N0 generations ⇒ **750 < N0 < 10,000**

Bottleneck during the last glacial period + small N0 more plausible…

Acknowledgements :

This research is currently supported by grants from ANR (GNP05013C project).

We are grateful to the Aquitaine Region for providing fundings to the wood quality (20030306002A) and genotyping (20030304002FA) facilities.

ESEB 2009 – Turin, ITALY