



A new resource for a better understanding of the complex genetic dynamics within the oak syngameon

Erwin Dreyer  based on peer reviews by **Oliver Brendel**, **Hilke Schröder**, **Ricardo Alia** and **Komlan Avia**

Tiange Lang, Pierre Abadie, Valérie Léger, Thibaut Decourcelle, Jean-Marc Frigerio, Christian Burban, Catherine Bodénès, Erwan Guichoux, Grégoire Le Provost, Cécile Robin, Naoki Tani, Patrick Léger, Camille Lepoittevin, Veronica A. El Mujtar, François Hubert, Josquin Tibbits, Jorge Paiva, Alain Franc, Frédéric Raspail, Stéphanie Mariette, Marie-Pierre Reviron, Christophe Plomion, Antoine Kremer, Marie-Laure Desprez-Loustau, Pauline Garnier-Géré (2021) High-quality SNPs from genic regions highlight introgression patterns among European white oaks (*Quercus petraea* and *Q. robur*). Missing preprint_server, ver. Missing article_version, peer-reviewed and recommended by Peer Community in Forest and Wood Sciences. <https://doi.org/10.1101/388447>

Submitted: 03 March 2020, Recommended: 22 February 2021

Cite this recommendation as:

Dreyer, E. (2021) A new resource for a better understanding of the complex genetic dynamics within the oak syngameon. *Peer Community in Forest and Wood Sciences*, 100003. [10.24072/pci.forestwoodsci.100003](https://doi.org/10.24072/pci.forestwoodsci.100003)

Published: 22 February 2021

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

When we launched the multidisciplinary [PCI Forest and Wood Sciences](<https://forestwoodsci.peercommunityin.org/PCIForestWoodSci>), we were well aware that this PCI devoted to a broad object (forests represent 1/3 of the emerged land area) and gathering multiple disciplines (from biology to social sciences), would be confronted with some challenges with respect to the range of pre-prints we would have to consider. Among the first ones submitted to the new PCI was this very interesting one (Lang et al, 2021) in the area of tree genomics and tree population genetics with a strong background in bioinformatics. This could well have been submitted to the PCI in Genomics, or the PCI in Evolutionary Biology, but finally fits very well in PCI Forest and Wood Sciences due to its finality: provide reliable genomic resources to tackle questions that are central to our understanding of the dynamics of forest ecosystems and the adaptation of trees to local conditions. The phylogenetic status of the two sympatric oak species **Quercus robur** L. and **Q. petraea** Matt. Liebl. is an important question that was the center of in depth investigations by the BioGeCo research group at Bordeaux in cooperation with many European groups. The two species are major forest tree species

in western Europe, structuring important ecosystems and providing high quality wood and other important ecosystem services. These species are considered to build up what is now called a “syngameon” together with a few other western European white oaks (Cannon and Petit, 2019). A syngameon is a complex of species exchanging genes, while still maintaining some degree of interspecific diversity enabling to differentiate, on the basis of leaf morphology for instance, the originally described species (Kremer et al, 2002 among many others dealing with the identification of oak species on the basis of their phenotype). The full sequencing of the genome of *Q. robur* (Plomion et al, 2018) opened an avenue for the development of new genomic tools to investigate the concept of species in this syngameon and to identify gene flows between populations and species. To date, whole genome sequencing and annotation is completed for *Q. robur*, *Q. lobata* Née, *Q. suber* L., and *Fagus sylvatica* L. (Plomion and Martin, 2020). The preprint of Lang et al (2021) builds up on this background and provides a very interesting and important contribution to this area. First, it provides access to a very large data base of genomic sequences, based on 800 gene fragments and identified 12500 Single Nucleotide Polymorphisms (SNPs) as well as 1500 insertions deletions. 11 individuals assigned to the species *Q. robur*, 13 assigned to *Q. petraea* and 1 to *Q. ilex* L. were used to establish this data base. This is an important asset for developing future research on species identification based on their genomes. In full agreement to the policy of all PCIs about data availability, full access is provided to the data under a free use license CC-BY. The manuscript has already a long history as a preprint and it was not that easy to find reviewers willing to do the job of reviewing such a large amount of information with a lot of technical details. At the end, we succeeded and obtained three very enthusiastic and positive reviews (attached to this recommendation). All reviewers agreed that this was an important contribution in the field, insisted on the quality of the writing and of the provided data, and finally agreed that the manuscript was to be recommended after a careful revision. The revision was readily done by the authors. One of the reviewers found that the preprint is close to a “data paper”, i.e., a paper providing the required information to enable (and advertise) the reuse of an important data set, which is only partly true. Indeed, the authors nicely used their data to address one of the most debated questions about the population genetics: how comes that in this syngameon, with inter-fertile individuals, clearly distinguishable species (from the phenotype at least) have been maintained? The main result of the preprint is that indeed there are some (only few) genomic differences among the two species in the patterns of distribution of the SNPs, but also that the patterns reveal a clear introgression of the genome of one of the species into the other one. This was based on a Bayesian clustering analysis reported figure 3, evidencing that 2 individuals displaying an unambiguous *Q. petraea* phenotype could not be assigned to any of the species. The usual hypothesis postulates an introgression of the *Q. robur* genome, the pioneer species able to colonize open areas during the post-glacial recolonization into that of *Q. petraea*, due to hybridization and successive events of back cross with inflowing *Q. petraea* pollen. Such an introgression was hypothesized decades ago, and receives here some degree of confirmation. Recent findings showed the importance of such an introgression for the adaptation of local oak populations to altitudinal gradients (Leroy et al, 2020). Given the importance of the contribution (both from the point of view of the development of new genomic tools and from that of our understanding of the population genetics in the oak syngameon, I do recommend this preprint in the frame of the PCI Forest and Wood Sciences after an insightful peer reviewing and a very careful revision leading to the current version 4 of this preprint.

References:

- Cannon CH, Petit RJ (2020) The oak syngameon: more than the sum of its parts. *New Phytologist*, 226, 978–983. <https://doi.org/10.1111/nph.16091>
- Kremer A, Dupouey JL, Deans JD, Cottrell J, Csaikl U, Finkeldey R, Espinel S, Jensen J, Kleinschmit J, Dam BV, Ducousso A, Forrest I, Heredia UL de, Lowe AJ, Tutkova M, Munro RC, Steinhoff S, Badeau V (2002) Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across western European mixed oak stands. *Annals of Forest Science*, 59, 777–787. <https://doi.org/10.1051/forest:2002065>

Lang T, Abadie P, Léger V, Decourcelle T, Frigerio J-M, Burban C, Bodénès C, Guichoux E, Provost GL, Robin C, Tani N, Léger P, Lepoittevin C, Mujtar VAE, Hubert F, Tibbits J, Paiva J, Franc A, Raspail F, Mariette S, Reviron M-P, Plomion C, Kremer A, Desprez-Loustau M-L, Garnier-Géré P (2021) High-quality SNPs from genic regions highlight introgression patterns among European white oaks (**Quercus petraea** and **Q. robur**). bioRxiv, 388447, version 4 peer-reviewed and recommended by PCI Forest and Wood Sciences. <https://doi.org/10.1101/388447>

Leroy T, Louvet J-M, Lalanne C, Provost GL, Labadie K, Aury J-M, Delzon S, Plomion C, Kremer A (2020) Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist*, 226, 1171–1182. <https://doi.org/10.1111/nph.16095>

Plomion C, Aury J-M, Amselem J, Leroy T, Murat F, Duplessis S, Faye S, Francillon N, Labadie K, Le Provost G, Lesur I, Bartholomé J, Faivre-Rampant P, Kohler A, Leplé J-C, Chantret N, Chen J, Diévert A, Alaeitabar T, Barbe V, Belser C, Bergès H, Bodénès C, Bogeat-Triboulot M-B, Bouffaud M-L, Brachi B, Chancerel E, Cohen D, Couloux A, Da Silva C, Dossat C, Ehrenmann F, Gaspin C, Grima-Pettenati J, Guichoux E, Hecker A, Herrmann S, Huguene P, Hummel I, Klopp C, Lalanne C, Lascoux M, Lasserre E, Lemainque A, Desprez-Loustau M-L, Luyten I, Madoui M-A, Mangenot S, Marchal C, Maumus F, Mercier J, Michotey C, Panaud O, Picault N, Rouhier N, Rué O, Rustenholz C, Salin F, Soler M, Tarkka M, Velt A, Zanne AE, Martin F, Wincker P, Quesneville H, Kremer A, Salse J (2018) Oak genome reveals facets of long lifespan. *Nature Plants*, 4, 440–452. <https://doi.org/10.1038/s41477-018-0172-3>

Plomion C, Martin F (2020) Oak genomics is proving its worth. *New Phytologist*, 226, 943–946. <https://doi.org/10.1111/nph.16560>

Reviews

Evaluation round #1

DOI or URL of the preprint: [10.1101/388447v2](https://doi.org/10.1101/388447v2)

Version of the preprint: 2

Authors' reply, 11 January 2021

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Erwin Dreyer](#) , posted 29 September 2020

Minor revision of the preprint before final and positive recommendation.

Dear Pauline,

I do apologize for the long time it took to reach this decision. We had trouble finding suitable referees and had to solicit a large number of colleagues to do the job. This is linked to the fact that the concept of PCI is still not fully integrated by the potential referees and our research community; we still have a large communication effort to do.

The reviewers produced very positive assessments of the manuscript and praised the precision and details of the description of the data and the usefulness of the attached data based. One of the reviewers stated that the preprint sounded like a data paper, but added that the text provided many useful and precise discussions of the data that go beyond what is expected from a data paper. The different reviewers suggested some minor

changes in the text that I am not going to detail again here, and that you will find in their comments. I guess from what I read that the required changes are very minor and should not be too difficult to introduce. Once these changes are made and the new version of the preprint is deposited in the archive, I will be able to provide a very positive recommendation of the preprint, which would be the first (or possibly the second one) for the PCI Forest&Wood Science.

Thank you anyway for your patience.

I am looking forward to consider the revised version and write the final recommendation.

With best regards

Erwin Dreyer

INRAE, UMR Silva (Université de Lorraine, AgroParisTech & INRAE), Nancy.

Recommender for this preprint

Reviewed by **Oliver Brendel**, 28 September 2020

Review of “High-quality SNPs from genic regions highlight introgression patterns among 2 European white oaks (*Quercus petraea* and *Q. robur*). ” from Lang et al.

This review is based on a version downloaded from the BioRxiv server : <https://www.biorxiv.org/content/10.1101/388447v2>

My expertise is in forest ecophysiology, especially oaks, with some knowledge on the genetics of these two oak species. So I will not be able to comment the sequencing techniques that were used in this manuscript as well as the finer points of the genetic implications of the results.

Overall, this manuscript describes a large data base of nucleotide polymorphisms of overall 24 trees (13 *Q. robur*, 11 *Q. petraea*) distributed over the geographical range and over 4 of the 5 cpDNA lineages along white oaks recolonization routes, with 2 to 4 individuals per location. The large number of detected nucleotide polymorphisms are then described in terms of their quality, coverage, introgression patterns, exclusive polymorphisms, polymorphism typology and relation to gene ontology. The presented catalogue of nucleotide diversity will not only be very useful for future research in landscape ecology of these two and other genetically close oak species, but has been analysed in terms of differences between the two species. These confirm earlier results, but using a large covering of the geographic range and a high number of nucleotide polymorphisms.

The abstract is very detailed and descriptive on the experimental set-up and the results in terms of detected polymorphisms, numbers of SNPs, the quality of the data provided for future research. There is only one short allusion on the results of differences in diversity between the two species where “these patterns are discussed in the context”. This is unfortunate, as the discussion broaches interesting subjects such as the introgressions detected and their relation to leaf morphology. Also two hypotheses are discussed for the detected higher diversity in *Q. petraea* (as compared to the one mentioned in the abstract). This gives the reader rather the impression of a data paper, which is less the case once the whole manuscript is read.

The introduction gives a quite complete context of establishing such a catalogue in terms of sequencing techniques, available genomic resources in oaks, knowledge on species differentiation and species introgression pattern. Their main objective was to provide a detailed characterization of sequence variation in *Q. petraea* and *Q. robur*. Using these data, they investigated the species differences and suggest future research possibilities.

Given my expertise, I have only few comments on the techniques described in the material and methods section as well as the very detailed genetic results presented in that section. The representativity of the 24 trees in terms of covering the geographic range as well as the genetic diversity is nicely presented in this section. It would however be helpful if the affiliation to the different postglacial migration routes could be included into table 1.

A large part of the discussion describes again the results on the detected nucleotide polymorphisms, in terms of availability in other species, the representativity of the experimental setup, the sequencing method used, the quality of the data, polymorphism typology, and the content and formats of the catalogue, by comparing

with already available data in the literature, for oaks or other forest tree species. They mention that the chosen locations cover 4 of the 5 cpDNA lineages along white oaks recolonization routes. As it is argued that 8 gametes were sufficient to represent most polymorphisms. Therefore, given the 40 gametes across both species (line 386), I wonder if it would not be possible to do a comparative analysis of the diversity in the four cpDNA lineages. This could add an interesting discussion of differences between these lineages. The authors also discuss in more detail the observed introgression in relation to the morphological species attribution, and they confirm on their small data set, that morphological *Q. petraea* were more likely to be introgressed individuals than morphological *Q. robur*. Further, they observed a higher diversity of *Q. petraea* for genes classed as "abiotic stress" in gene ontology terms. The authors then discuss two non-exclusive hypotheses proposed in the literature, one pertaining to the difference in life-history strategies for colonizing new stands, the other pertaining to higher selective constraints in *Q. robur*. The discussion finishes with an outlook for a future usage of the presented nucleotide diversity catalogue.

This manuscript not only provides a solid foundation for future landscape ecology studies but also presents interesting insights into the genetic difference between two closely related, hybridizing oak species. I therefore recommend strongly the publication of this manuscript.

Detailed comments : Abstract : Probably not all readers will know what Sanger sequencing is, similarly for the "phi-pi" (line 51).

Introduction : line 88 "species pairs" : This refers probably to the four species cited above and their capability to hybridize ? Also it is not clear what "scenarios" refers to, scenarios for the species distribution ? This section might need some editing to clarify these points.

Material and Methods line 144 : It would help the reader at this point if it could be mentioned that the à priori morphological species assignments will be supplemented with genetic assignments and an analysis of introgression further down in the manuscript. Lines 162-163 : It is not clear to me what these 146 individuals from 3 French regions relate to. Were these the basis for the EST ? This seems to be a much larger choice than the 25 individuals described at the beginning of the M&M section.

Results : In Figure 3, the limits of the attribution of the individual to one or the other species should be visible. It is not clear which 4 individuals were excluded from the following analyses.

Reviewed by [Ricardo Alia](#), 17 July 2020

The paper is a complete characterization of SNPs in oak species, and the application of these resources to two important topics of research: characterization of sequence variation in *Quercus petraea* and *Quercus robur*, and the analysis of introgression asymmetry. The paper is very well written, with a detailed description of the methods applied and a complete discussion of the results. The resources developed are important for the study of the species, and the only concern is that the different objectives of the paper (description of the new resources, characterization, and application to the analysis of the introgression) are quite different. Then, the paper is quite long and in some cases difficult to follow, but as the information is relevant, I do not have any significant recommendation to the authors.

Reviewed by [Komlan Avia](#), 17 July 2020

[Download the review](#)

Reviewed by [Hilke Schröder](#), 01 July 2020

I am impressed by the amount of data, analyses and interpretation presented in this paper. The authors not only made an exhaustive effort with Sanger resequencing of a lot of interesting genes, they also gave a kind of review for already existing data, interpreted them and made recommendations for further studies. The data sets will provide scientist working with oaks with unprecedented possibilities for future projects in a

broad range. This is a very valuable paper as well how it is organized as because of the extensive range it is covering. So, I can only suggest to recommend it.

Only one small request: Already in the introduction („across a large part of both species geographic range“) and also in the discussion, the authors stated that they used *Q. robur* and *Q. petraea* populations „across a large part of their geographic range“. I would like to have this statement qualified because the most Eastern population used by the authors is in Hungary. The further South-Eastern distribution is missing. Maybe it can be stated as „Western and Central Europe“ as already mentioned in the chapter „sample collection“.