

FORGENIUS

Improving access to FOREst GENetic resources Information and services for end-Users

Deliverable D3.1

***Genetic, phenotypic and environmental
characterization of two GCUs for at least one
species.***

Planned delivery date (as in DoA): M23 30/11/2022

Actual submission date: 09/12/2022, month M24

Workpackage: WP3

Workpackage leader: INIA

Deliverable leader: INRAE

Version: 1.0

Project funded by the European Commission within the Horizon 2020 Programme	
Dissemination Level	
PU Public	PU
CI Classified, as referred to Commission Decision 2001/844/EC	
CO Confidential, only for members of the consortium (including the Commission Services)	

Research and Innovation action: GA no. 862221

Start date of the project: January 1st, 2021



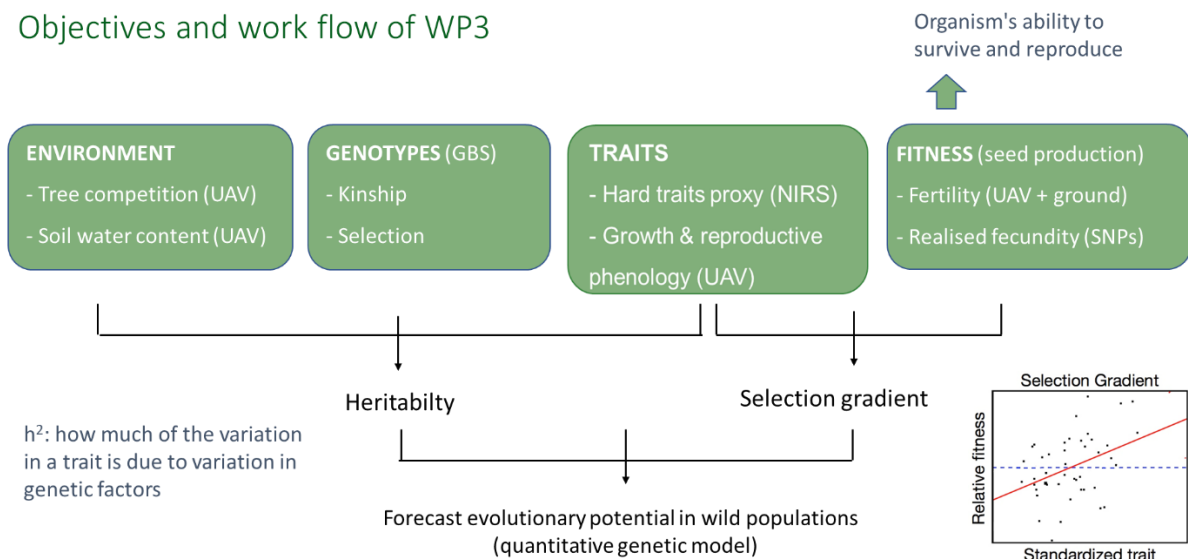
TABLE OF CONTENTS

1	Summary.....	3
2	Introduction.....	3
3	Results	4
3.1	Genetic data	4
3.2	Phenotypic data.....	5
3.3	Environmental data	11
3.4	Integration of the data	12
4	Conclusions.....	12
5	Partners involved in the work	12
6	Annexes	13

1 Summary

In the first two-year period of the project, WP3 has focused on producing three datasets (genetic, phenotypic¹ and environmental data) for two forest tree species (*Populus nigra* and *Pinus pinaster*). The initial idea was to focus on only one species, but we changed our approach to take advantage of the fact that data are cumulatively produced. This strategy has paid off as we could get nearly complete datasets for the two focal species using a number of innovative methods. The still partial datasets will be completed beginning of 2023. These datasets will contribute to the assessment of several population-level estimators (e.g. fitness, kinship, heritability) that will be incorporated into quantitative models aiming at predicting population adaptation in the wild (see the schematic representation of the objectives and work flow below), and thereby produce dynamic descriptors of Genetic Conservation Units (GCUs).

Objectives and work flow of WP3



2 Introduction

For each species (*Populus nigra* and *Pinus pinaster*) two GCUs with contrasted environmental settings have been characterized, each composed of 500 adult trees (georeferenced at the individual level, see Figure 1) and 250 juvenile trees (georeferenced at the patch or individual level). The GCUs characteristics are described in the Table 1 and Figure 1.

Table 1. Characteristics of the eight GCUs sampled in 2021-2022.

Species	EUGIS code	Country	Latitude	Longitude	# adults	# juveniles
<i>P. nigra</i>	Nationalpark Donau-Auen	Austria	48,18068-48,15283	16,49609-16,92818	500	250
<i>P. nigra</i>	ESP0395	Spain	41.584117	-0.762574	500	250
<i>P. pinaster</i>	FRA00051	France	44°57'48.6"N	1°09'53.9"W	500	250
<i>P. pinaster</i>	ITA00019	Italy	43.131365	11.248482	500	250

¹ Phenotypic refers to the set of observed characteristics or traits of the trees

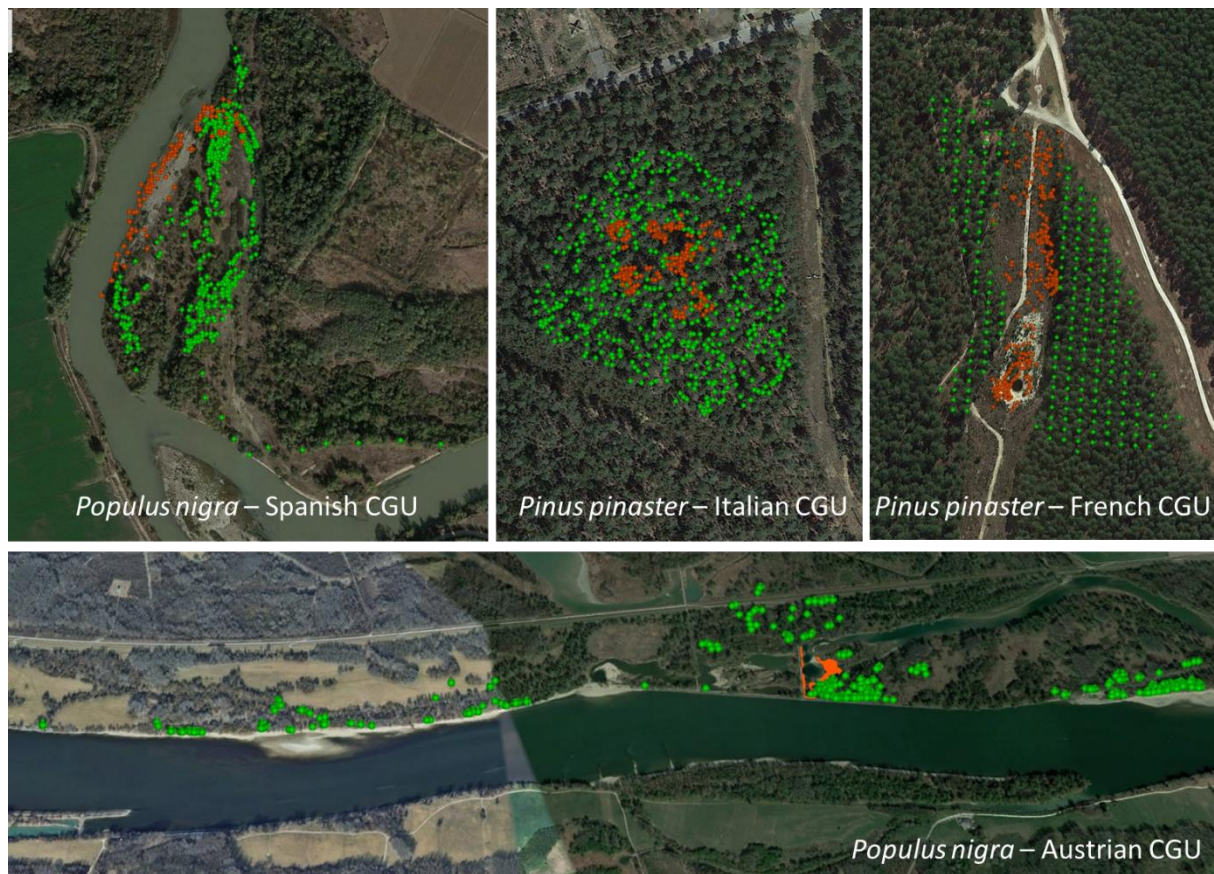


Figure 1: Distribution map for the four GCUs sampled in 2021 (green: adults; red: juveniles). Note that the regular spatial pattern observed of individual trees is linked to the low accuracy of GPS coordinates measured by local partners. More accurate positions were obtained by UMR using dGPS and tachymeters (see the “Phenotypic data” section below).

3 Results

3.1 Genetic data

Leaf/needle tissue were collected by local partners for 2×500 adults and 2×250 juvenile trees per species. DNA was extracted by local partners then sent to CNR for handling and shipping to ThermoFisher Scientific (USA) for genotyping. For both species, the genotyping technology was the Axiom 50K microarray, and the SNP set 4TREE (developed in a previous European project B4EST). The delivery dates of the genotypes by the ThermoFisher Scientific company are indicated in Table 2. Production of genotyping data was slightly delayed because of technical demands by the provider of the sequencing service (for the Austrian *P. nigra* and Italian *P. pinaster* genotypes) and of DNA quality issues (for the French *P. pinaster* genotypes)). The method used here has been validated multiple times and we foresee high-quality genotypes for all datasets.

Table 2. Genotyping advances for adult and juvenile trees for the eight GCUs.

Species	Country	500 adult genotypes	250 juvenile genotypes
<i>P. nigra</i>	Austria	December 2022- January 2023	December 2022- January 2023
<i>P. nigra</i>	Spain	June 2022	June 2022
<i>P. pinaster</i>	France	January-February 2023	January-February 2023
<i>P. pinaster</i>	Italy	December 2022- January 2023	December 2022- January 2023

⇒ Datasets are held by Partner 2 (CNR).

3.2 Phenotypic data

Measurement were taken and samples were collected on adult trees by local teams to get access to different phenotypes:

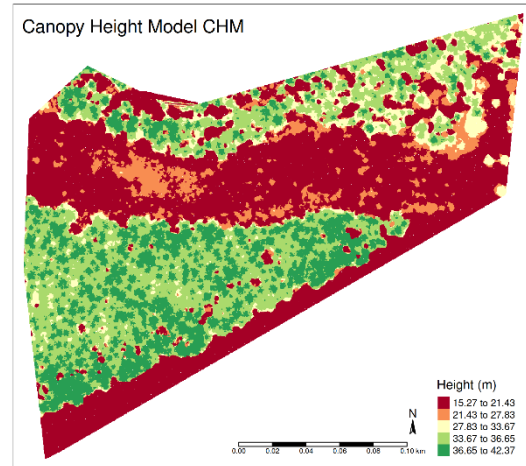
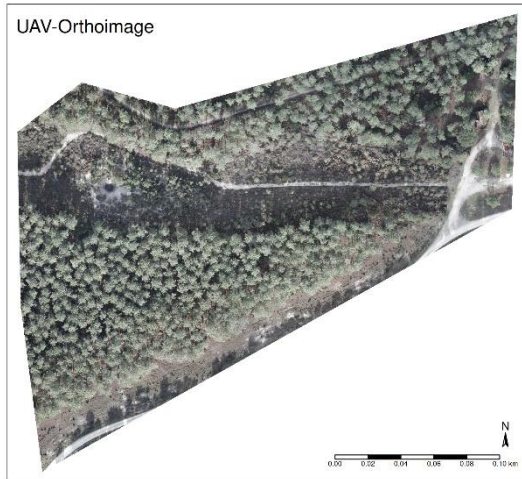
- **Diameters of 2x500 trees**
- **Tree height of 2x500 trees**
- **Crown size of 2x500 trees**

Tree height and crown size were assessed in an innovative way based on unmanned aerial vehicle (UAV), i. e., drone flights. Specifically, we created optimized flight plans using the R software-based flight planning application *uavRmp*. In each test area, seven ground control points (GCP) were surveyed using total stations, in such a way as to georeference sufficiently well in 3D the photogrammetric processing (Figure 2). On this basis, a 3D point cloud² and an orthoimage mosaic (cone classification) were reproducibly generated via Agisoft Metashape and the Ortho+ toolbox³. The 3D point cloud was classified into soil / non-soil points using the *lidR* software package. Based on spline interpolation and iterative filtering, a Digital Surface Model (DTM) was calculated from the soil points. After normalization of the values, a CHM was generated by using a *pitfree* algorithm. The delineation of the trees was based on another algorithm (*dalponte* algorithm) and on subsequent planar projection to create the 2D crown segments. All height values from the point cloud were assigned to the segments. Tree height and crown size were derived based on the DTM. This complete and original reproducible workflow is available at the GitHub repository *forgenius-pp*.

² 3D point cloud are characterized by segmentation, Digital Surface Model (DTM), Canopy Height Model (CHM)

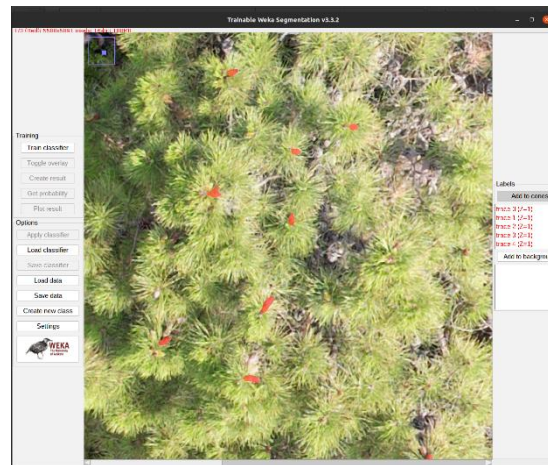
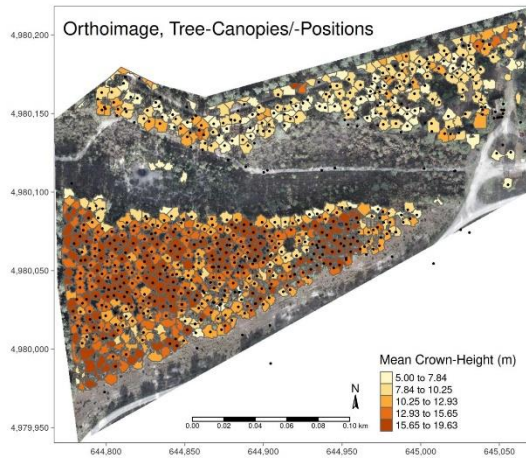
³ Ortho+ toolbox is an advanced supporting tool for reproducible creation of point clouds and orthoimages

- **Seed cones of 2x500 trees.** Here also counting was innovately based on UAV flights. The trainable segmentation plugin WEKA fastRandomForest algorithm in FIJI (ImageJ, Figure 2) was used to segment and classify pine cones based on the orthomosaic⁴. Also, control counting by visual inspection was performed for both data sets. Classified and visual inspection results produced highly comparable results.



High-resolution orthoimage

High-resolution Canopy Height Model (10 cm)



Segmentations of tree canopies (10 cm) and mid-canopy height. Dots indicate stem position measured with a total station.

Processing of cones via fastRandomForest algorithm in FIJI (ImageJ)

Figure 2: data retrieved from UAV flights in the French GCU of *P. pinaster*.

⁴ Orthomosaic is a large, map-quality image with high detail and resolution made by combining many smaller images called orthophotos

- **Short increment cores from 2x500 trees** for *P. nigra*, and **drilldust from 2x500 adult trees** for *P. pinaster* (Figure 3). From this material, Near InfraRed Spectroscopy (NIRS) measurements were performed (Figure 4). Using a correlative approach linking the wood traits measured in WP2 across the full species range (150 trees) and the NIRS measured in WP3 (1,000 trees), we will explore whether we could infer wood traits for the 1,000 trees/species in WP3.

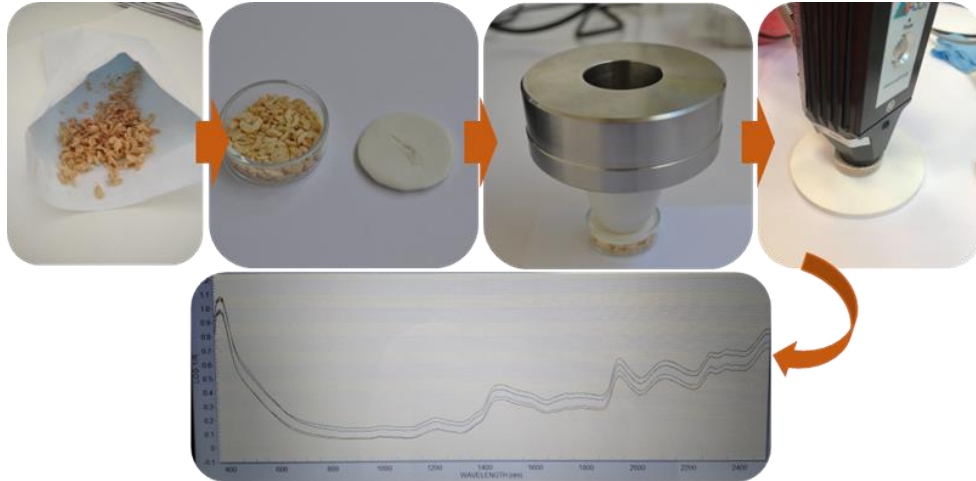


Figure 3: Steps to analyse a drilldust sample of *P. pinaster* with NIRS (CETEMAS).

- **Short cores of 10x15 trees** for *P. nigra* and *P. pinaster* were collected in WP2 and sent to INRAE in WP3 to collect NIRS spectra on solid and ground samples.

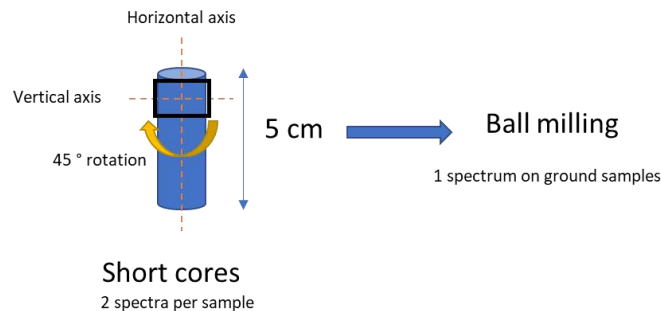


Figure 4: NIRS spectra acquisition on short cores (INRAE)

- **Branches of 10x15 trees** for *P. nigra* and *P. pinaster* were collected and provided by WP2 team for NIRS measurements (Figure 5). Near InfraRed Spectra were collected on the cross-section of the branches or on the xylem after debarking the branches. Ball milling was then used to finely grind the samples. Calibration models are then developed based on hard traits measured in WP2.

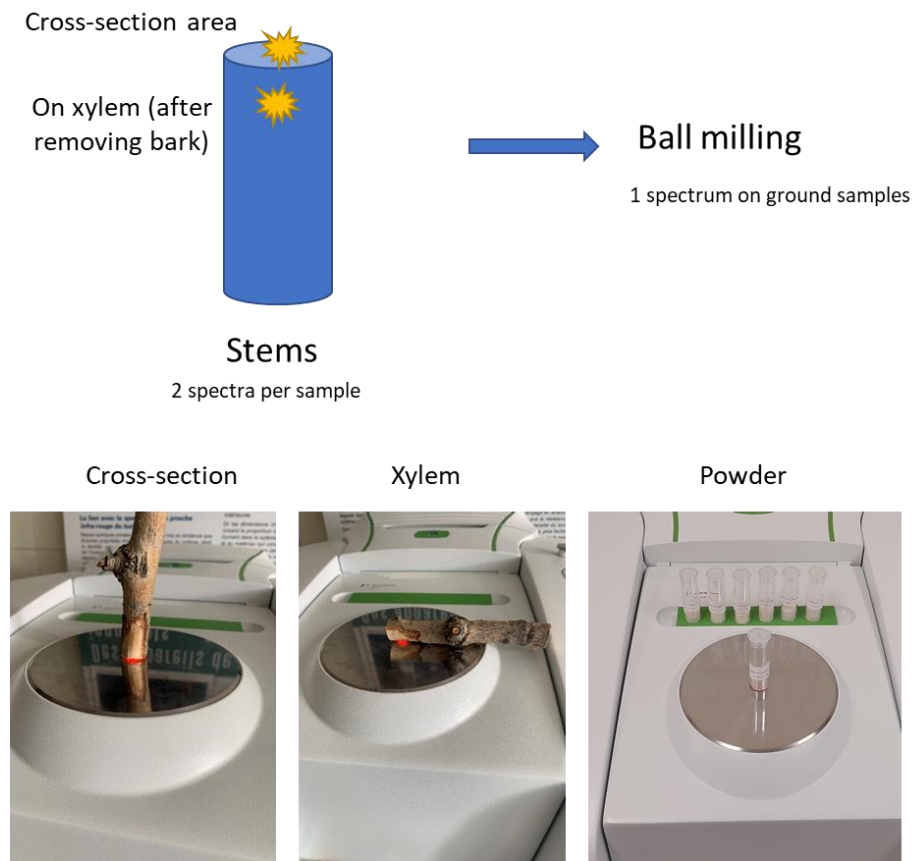


Figure 5: NIRS spectra acquisition on branche (INRAE)

- **Long increment cores of 2x50 trees.** From this material we measured X-ray microdensity profiles for the full core as well as for the outermost 5 cm. Using a correlative approach linking microdensity measured in WP2 across the full species range (150 trees) and NIRS measured in WP3 (1,000 trees), we will explore whether we can indirectly infer wood density for the 1,000 trees/species in WP3 (Figure 6).

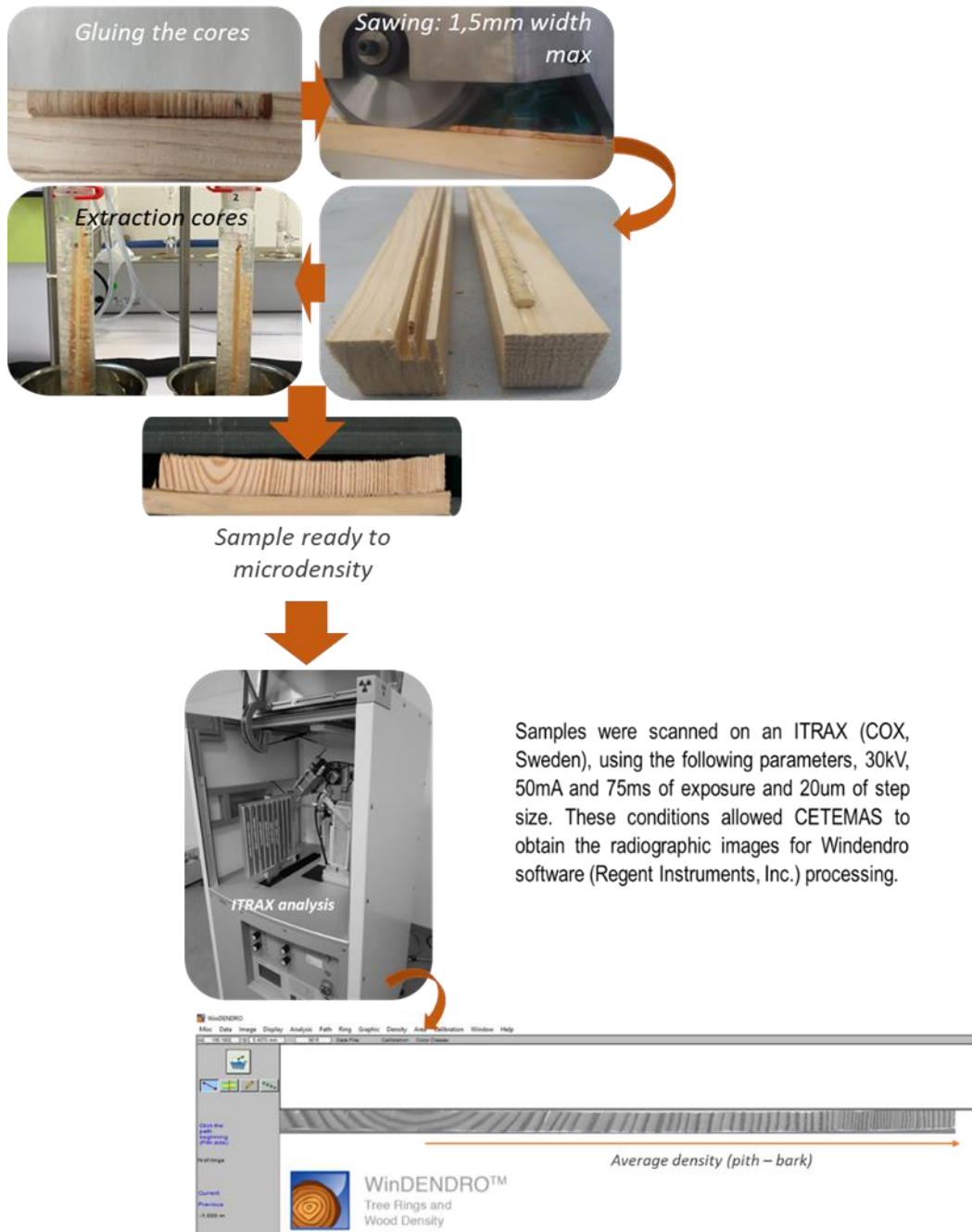


Figure 6: Steps for microdensity determination of cores with X-ray technique (CETEMAS).

A summary of the phenotypes available from the different partners is presented in Table 3.

Table 3. Summary of the traits measured in the eight GCUs.

Species	Country	500 diameters	500 wood NIRS spectra
<i>P. nigra</i>	Austria	BFW	INRAE
<i>P. nigra</i>	Spain	INIA-CSIC	INRAE
<i>P. pinaster</i>	France	INRAE	CETEMAS
<i>P. pinaster</i>	Italy	CNR	CETMAS

Species	Country	150 wood NIRS spectra for calibration	200 Wood microdensity for calibration	10 wood traits WP2 for calibration	150 branch NIRS spectra for calibration
<i>P. nigra</i>	Austria	INRAE	CETEMAS	CREAF	INRAE
<i>P. nigra</i>	Spain	INRAE	CETEMAS	CREAF	INRAE
<i>P. pinaster</i>	France	CETEMAS	CETEMAS	CREAF	INRAE
<i>P. pinaster</i>	Italy	CETEMAS	CETEMAS	CREAF	INRAE

Species	Country	Seed counts per tree	Tree height	Crown size	Exact position total station
<i>P. nigra</i>	Austria	Spring 2023	Spring 2023	Spring 2023	Spring 2023
<i>P. nigra</i>	Spain	Spring 2023	Spring 2023	Spring 2023	Spring 2023
<i>P. pinaster</i>	France	UMR	UMR	UMR	UMR
<i>P. pinaster</i>	Italy	UMR	UMR	UMR	UMR

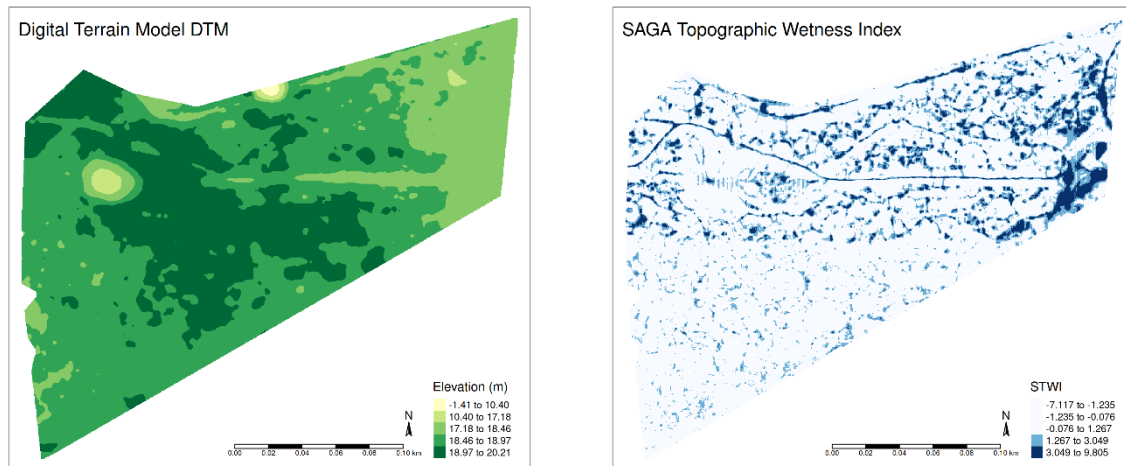
Delay for the UAV flights over the *P. nigra* GCUs was linked to a staff problem at UMR at the beginning of the project. Flights will be performed in the next flowering-fruiting season, in spring 2023.

As mentioned above, in addition to this sampling strictly planned within WP3, additional phenotypes collected by WP2 are also available for the two species, for 10 trees out of the 500 within a GCU. These wood traits include wood density, wood capacitance, cavitation (P50, P12, P88). They were sampled across the full distribution of the species, and will be used to infer the correlation between hard-to-measure phenotypes in WP2 (150 trees per species) and the easier-to-measure NIRS phenotypes in WP3 (1,000 trees per species). Initial analyses of phenotypic and NIRS data are encouraging as preliminary results show significant relationships between hard-to-measure traits and NIRS (these outputs are due by M32 and will be reported in D3.3).

⇒ Datasets are held by Partners 1 (INRAE), Partner 3 (CETEMAS), Partner 13 (UMR), and Partner 4 (CREAF).

3.3 Environmental data

The integration of environmental data into the models is necessary to consider microenvironmental variation at the individual level. Given that the sandy soils at the two investigation sites for now are very uniform (i.e. *Pinus pinaster*), microtopography is the single most important microenvironmental determinant with regards with both water availability and insolation.



High-resolution Terrain Model (10 cm)

Modified SAGA Topographic Wetness Index (10 cm)

Figure 7: data retrieved from UAV flights in the French GCU of *P. pinaster*

The novel calculation of the derived environmental parameters is based on the high-resolution terrain model explained in the previous section in combination with established tools from SAGA/GRASS/QGIS and various R scripts (Figure 7). In particular, these are:

- Total irradiance surface
- Direct irradiance surface
- Diffuse irradiance surface
- The modified topographic wetness index SAGATWI (NDVI Sentinel 5 years)

In addition, the soil moisture and exemplarily the soil respiration were repeatedly recorded at 50 points systematically sampled each with a soil moisture kit (SM150-Kit from Delta-T), and with a new, soon to be published, technology from the University of Freiburg, respectively. The soil moisture measurements are being used to ground truth the topographic wetness index.

Also, based on the tree-segmentation and the height and crown sizes we have calculated tree-competition indices.

For cross-comparison, we have acquired satellite imagery with a 10-m resolution tiles from Sentinel-2 (<https://eos.com/find-satellite/sentinel-2/>) by synergy with WP1. From these images, tree competitions indices were extracted via a vegetation density index (NDVI: Normalized Difference Vegetative Index) while soil water content could also be obtained.

A summary of the available environmental data is presented in Table 4.

Table 4. Environmental data available for the eight GCUs (UAV and satellite).

Species	Country	Tree competition UAV	Water soil content UAV	Sentinel-2 tiles
<i>P. nigra</i>	Austria	Spring 2023	Spring 2023	not available
<i>P. nigra</i>	Spain	Spring 2023	Spring 2023	available
<i>P. pinaster</i>	France	November 2022	November 2022	available
<i>P. pinaster</i>	Italy	November 2022	November 2022	available

Delay for the UAV flights over the *P. nigra* GCUs is linked to a staff problem at UMR at the beginning of the project. Flights will be performed in the next flowering-fruiting season, in spring 2023. Sentinel-2 data could not be retrieved for the Austrian GCU because satellite data are not available for this country.

⇒ Datasets are held by partners by Partner 13 (UMR) for UAV data and by Partner 1 (INRAE) for satellite data.

3.4 Integration of the data

A general meeting was organized on November 29th with most of the partners involved in WP3 to discuss the processing of the different datasets (kinship and fitness inferred from genetic data, fitness inferred from seed imagery, NIRS calibration, environmental data from UAV/satellites) and their integration into quantitative genetics models (see minutes of the meeting in the Annexes).

4 Conclusions

During the two first years of the project, most of the sampling objectives have been achieved, or will be by the beginning of 2023. These measures were an opportunity to implement original and efficient methods. Instead of having a full dataset for a single species, we have produced a near complete dataset for two species. These significant progresses should guarantee the completion of the planned activities in the next phases of the project.

5 Partners involved in the work

	<i>Populus nigra</i>		<i>Pinus pinaster</i>	
	Austria	Spain	France	Italy
Leaf/needle, wood sampling	BFW	INIA-CSIC, CITA	INRAE	CNR
DNA extraction	BFW	INIA-CSIC	INRAE	CNR
DNA genotyping	CNR	CNR	CNR	CNR
Basic traits	BFW, UMR	INIA-CSIC, CITA, UMR	INRAE, UMR	CNR, UMR
Trunk/drilldust NIRS	INRAE	INRAE	CETEMAS	CETEMAS
Branch NIRS	INRAE	INRAE	INRAE	INRAE
Trunk Microdensity	CETEMAS	CETEMAS	CETEMAS	CETEMAS
WP2 traits	CREAF, INRAE	CREAF, INRAE	CREAF, INRAE	CREAF, INRAE
Environmental data	UMR, INRAE	UMR, INRAE	UMR, INRAE	UMR, INRAE

6 Annexes

Minutes Task 3.4 Meeting. 29/11/2002. Videoconference

Version revised by most partners. Updated version available on the FORGENIUS platform.

Objective: Establish a road map for completion of the task.

Deliverables to be produced:

D3.7: Estimation of GCUs adaptability (quantitative genetics linking trait proxy and genomics) for at least one species. [Initially planned at M37]

Decisions:

- ⇒ This deliverable will be produced in month 36.
- ⇒ This deliverable will be based on *Pinus pinaster* results (INRAE Bordeaux, CNR Italy).

D3.8: Estimation of GCUs adaptability (quantitative genetics linking trait proxy and genomics) for the remaining species. [56]

- ⇒ No decision at this time. We have revised the different steps for completing this deliverable. See next steps.

Data availability:

Fitness estimation. (UMR not present in the meeting because it was unaware of the meeting due to a problem in the mailing list).

Different fitness component estimations will be available:

- Cone count, tree height and crown size (with exact position of each tree) (UMR).
- Fitness estimation based on parentage analysis. The methods (MEMM) are robust and correct for the spatial distribution of adults and seedlings.
 - ⇒ *Pinus pinaster*: drones and counting from the ground done during 2022 is completed and will be made available in January, as currently data curation is ongoing. The progress can be monitored here: <https://gitlab.uni-marburg.de/reudenba/forgenius-pp>
 - ⇒ *Populus nigra*: delayed to spring 2023
 - ⇒ *Fagus sylvatica*: late summer 2023
 - ⇒ *Pinus sylvestris*: Fall/Winter 2023

Genotyping (CNR).

- ⇒ *Populus nigra*: SNPs available for the Spanish GCU, samples for the Austrian GCU pending genotyping by ThermoFisher (TF).

- ⇒ *Pinus pinaster*: samples for the Italian GCU pending genotyping by TF, DNA extraction ongoing for the French GCU.
- ⇒ *Fagus sylvatica*: ongoing DNA extraction. Implications for a possible delay in a PhD thesis (INRAE-GIS). Necessary to complete the genotyping before next summer, so that the PhD student can still work on it.
- ⇒ *Pinus sylvestris*: ongoing DNA extraction.

NIRS (CETEMAS, INRAE).

Pinus pinaster: microdensity done (CETEMAS). NIRS on trunk and branches done (CETEMAS, INRAE). WP2 traits done (CREAF, INRAE). Ongoing: NIRS calibration. Preliminary results show no correlation between drilldust and traits but promising correlation between plain wood and traits (CETEMAS).

Populus nigra: microdensity done (CETEMAS). NIRS on trunk and branches done (CETEMAS, INRAE). WP2 traits done (CREAF, INRAE). Ongoing: NIRS calibration. Preliminary results show good correlation with P12, P50, P88 (INRAE).

- ⇒ Meeting in early 2023 to check the progress on these two species, and the phenotypes of the 500 adult trees in order to compute genetic parameters for deliverable D3.7.

Fagus sylvatica and *Pinus sylvestris*: ongoing analyses of microdensity and NIRS. For the former species, same concern to get the phenotypes available by summer for the PhD student to work on this species.

Environment (INRAE, UMR).

- ⇒ WP1 (INRAE) provides information at the site and tree level for normalized differential vegetation index (NDVI). This information will complement the UAV data and be integrated mainly in a multispecies analysis.
- ⇒ Unmanned aerial vehicle (UAV) was used to get data at tree level by UMR in the two GCUs of *Pinus pinaster*. The data will be made available in January 2023. The process can be accessed here: <https://gitlab.uni-marburg.de/reudenba/forgenius-pp>
- ⇒ *Populus nigra* flights are planned for Spring 2023.
- ⇒ *Fagus sylvatica* in late summer 2023 and *Pinus sylvestris* in fall/winter 2023, see answers above.

Theoretical considerations:

- Different possible methods to estimate phenotypic and selection gradients (articles were shared among partners and are available on the Forgenius platform in the WP3 folder).
- Effect of existing genetic structure in the sites (kinship) to efficiently estimate genetic covariances.

- Necessary to consider the multiple dimensions of the phenotypes. *There are almost 20 traits from WP2, but we only use proxies from NIRS based on wood and branch. Can we also use NIRS itself as a source of traits? Combining genomic with NIRS data to get a better characterization (article on that topic was shared among partner). Three levels of analyses could be considered:*
 - a) Use NIRS to derive relationships with wood traits only in those cases where it is logical to expect that NIRS bands reflect functional variations in a trait. In this case, we obtain an estimate of the phenotype for that trait specifically related to drought resistance. Likely with a high accuracy.
 - b) Use NIRS to derive correlations with all the other traits available from WP2 on the assumptions that spectra of phenotypic traits (structural physiological, biochemical properties of leaf, wood or a mixture of the two) may also be related to NIRS spectral bands. In this case, we obtain an estimate of several components of the phenotype or some multidimensional phenotypic index. Probably with low to medium accuracy.
 - c) One can also expand this correlative avenue by doing, say, a PCA on all phenotypic traits and then attempt to relate NIRS bands to the phenotypic PCA axes. In this case, we obtain an estimate of NIRS gradients directly, probably with very high accuracy but without a link, functional or otherwise, to the ‘classical’ concept of the phenotype.
 - ⇒ Sharing of relevant papers. A folder is available on the Forgenius platform within the WP3.
 - ⇒ Sharing and discussion of papers can also be done on the Forgenius Slack
 - ⇒ Discuss a review paper based on these topics (INIA-CSIC) for the next meeting.

Publications:

- ⇒ Review paper (to decide in the next meeting).
- ⇒ Specific papers for the different species (PhD students and Postdocs in the group).
- ⇒ Multispecies analysis (to decide in a next step of the task).
- ⇒ NIRS evaluation: not considered in the meeting.

Next Meeting:

- ⇒ Before the next Forgenius annual meeting in 2023.

Please answer this survey: <https://doodle.com/meeting/participate/id/e3QgNxOb>

Partners involved: Partners involved and related tasks in WP3 and other WPs.

Task 3.4. INIA-CSIC, INRAE, GIS, Luke and UHEL.

Related to other WP and tasks in WP:

- Phenotyping- WP2-CREAF, Task 3.3.
- Genotyping- WP4-CNR, Task 3.2

Participants in the meeting

Delphine Grivet (INIA-CSIC), Ricardo Alia (INIA-CSIC), Sanna Olsson (INIA-CSIC), Aida Sole (INIA-CSIC), Ivan Scotti (INRAE), Santiago C. Gonzalez (INRAE), Nassim Belmokhtar (INRAE), Mara Arrojo (CETEMAS), Juan Majada (CETEMAS), Berthold Heinze (BFW), Maurizio Mencuccini (CREAF), Tanja Pyhajarvi (UHEL), Marjana Westergren (GIS), Giuseppe Vendramin (CNR), Camilla Avanzi (CNR), Andrea Piotti (CNR).

UMR was not included in the WP3 mailing-list and therefore did not participate since they were unaware of the meeting.

Maurizio Marchi (CNR) excused his presence.

List of PhD and postdoc students involved in the tasks

- Adelaïde Théraroz (INRAE-GIS): selection gradients/responses to selection in *Pinus pinaster* and beech.
- Aida Sole (INIA-CSIC): review paper, selection gradients/responses to selection in *Populus nigra*.
- Christian Mestre—Runge (UMR): UAV/GIS-based counting of cone production and environmental Characterization.
- Elliot Shayle (UMR): Individual tree identification and morphological characterization with 3-D data.

Background (as in DOA)

Phenotypic data on FS sites, as provided by WP2, will be combined with corresponding heritability estimates (T3.2) and (multivariate) selection gradients (T3.3) to estimate the adaptive potential of the GCUs of the FS group. In particular:

(a) through the relationship between trait values and individual fitness (selection gradient) we will estimate average GCU-level fitness and variance thereof, as direct indices of, respectively, adaptedness and adaptability; this task will be led by INIA.

(b) through the combination of heritability and selection gradients with population-level, quantitative genetic modelling algorithms, we will predict the adaptive potential (viability, persistence of variance in fitness) of each GCU under climate change. Notice that: (i) because we will have obtained two estimates of both heritability and selection gradient per species (T3.2, T3.3), expected heritability and selection gradient for individual GCUs will be obtained by interpolation between these two estimates, based on their (climatic and geographical) position relative to the GCUs used in T3.1- T3.3; (ii) projection of future selection gradients for each GCU, as determined by future climate, will be extrapolated from currently observable selection gradients. This task will be coordinated by INIA.