FORGENIUS

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

Deliverable D3.3

Trait proxy identification (relationship of phenotypes from WP2 and NIRS from WP3) & heritability estimates of trait proxy for at least one species

Planned delivery date (as in DoA):M32 30/08/2023

Actual submission date: 02/10/2023, month M34

Workpackage: WP3

Workpackage leader: INIA

Deliverable leader: INIA

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

One of the objectives of WP3 is to forecast the adaptability of forest genetic resources by estimating selection gradient (trait-fitness relationships) and heritability of selected traits of four species (*Populus nigra*, *Pinus pinaster*, *Fagus sylvatica*, *Pinus sylvestris*) collected in two natural populations growing in contrasted environments. To reach that objective, quantitative-genetic theory approaches are being used (e.g. Alexander et al. 2020), which require three sets of data: phenotypic data, genetic data and environmental data. The originality of the project lies in producing proxies of hard-to-measure traits (measured in WP2 in a subset of samples) in thousands of individuals from natural populations, via Near Infrared Spectra (NIRS). At this stage of the project data production is heterogeneous across sites and species and we present preliminary results for the most advanced species, *Populus nigra*.

2 Introduction

Traditional methods to estimate heritabilities employed progenies with known pedigree and sharing common environmental conditions (*ex situ* common garden and/or controlled environmental conditions). Although this estimates usually present more precise estimates of heritability compared to *in situ* studies, they could not accurately represent the additive genetic variance that is expressed under natural environments. Under these conditions is where the evolution acts, and thus *in situ* estimates are expected to better represent the ability of populations to adapt to their natural environment and to future conditions.

However, in situ estimates underline different problems to reliable estimates of heritability and selection under those environments. These problems are related to the estimation of the pedigree (or genetic relationship matrix) among the different trees. Results in forest trees have shown that parentage analysis in natural conditions can detect only some relationships among the trees, and the error estimates of heritability are depending in the number of offspring from the same family. A second problem is the estimation of the phenotype, that in in situ natural conditions are depending on the age of the tree, and the environmental conditions among other factors. For instance, age has a major effect on height, diameter, basic density, ring width and other phenotypic traits related to growth. But also age is related to the competition among trees of different ages, and therefore in the acquisition of resources. Environment also is variable within the area, and these variability could be associated to the distribution of the related genotypes, confounding the genotypic and environmental effects. Therefore, in situ estimation should be taken with caution in absence of a detailed analysis of all these factors.

Trait proxy identification and estimation of their heritability in natural populations is challenging as it requires the production and processing of multiple datasets in order to control for various factors that may bias the output: phenotypic data together with genetic and environmental data. The production of these datasets requires the coordination of many partners and the use of various approaches:

- Phenotypic traits: hard traits (together with NIRS measurements) are produced by WP2 in a subset of 10 individuals across 15 populations distributed in the species distribution range. The relationship between hard traits and NIRS is then applied to the 500 individuals of each GCU in order to produce hard traits for the 500 individuals.
- Genetic data: genomic data are produced by WP4 for the 500 adults and 250 juveniles of each GCU. This data is used to estimate the genetic relationship across adults

required to estimate the heritability. It is also used to produce individual fecundity estimates (i.e. fitness) comparing adults and juveniles genetic data.

- Environmental data: measured by UAV (e.g. soil humidity, competition index, etc…), they are needed to control the effect of environmental heterogeneity in the field.

The collaboration of different partners in this project allows to rule out some of the factors affecting the precision of the heritability estimates. In natural populations, the spatial distributions of individuals could be not random and thus, environmental, spatial and genetic parameters could be correlated. Accurate information on the genetic relationship among individuals is essential to elucidate the genetic basis of complex traits and the capacity of the populations to genetically respond to environmental changes. Furthermore, considering the spatial distribution and the environmental resemblance among individuals improves the precision and accuracy of the estimated additive genetic variance.

Following the methodological framework employed by Alexandre et al., 2020 we considered the autocorrelation of the spatial distribution of the individuals and include the environmental variables collected by UAV flights in the populations as covariates in the models to account for the environmental effects over the resemblance of phenotypic values among individuals. This procedure allow to better characterize the contribution of genes to the observed phenotypes.

3 Results

3.1 Trait proxy identification (relationship of phenotypes from WP2 and NIRS from WP3)

In agreement with WP2, a list of 10 hard traits is contemplated within WP3 (level 1 traits are more relevant and are measured on wood):

Function	Traits	Ranking	Definition
Survival during drought	*P50*	1	Water potential causing 50% of embolism.
Survival during drought	*Slope*	1	Slope parameters of the vulnerability curve to cavitation.
Growth and productivity	*wood density*	1	wood density.
Growth and productivity	*Ksmax*	1	Specificl hydraulic conductivity (per unit sapwood area).
Survival during drought	*HV*	1	Huber value: sap wood area to leaf area ratio.
Survival during drought	*Capacitance*	1	Branch capacitance.
Survival during drought	*Capacity_branch*	1	Amount of water stored in the branch at saturation.
Survival during drought	*TLP*	2	Turgor loss point.
Survival during drought	*Gmin*	2	Minimum or cuticular conductance.
Growth and productivity	*SLA*	2	specific leaf area.
Growth and productivity	*N*	2	Nitrogen content.

Production of NIRS trait proxy is explained in D3.1. Besides these 10 hard traits, others soft traits are available: DBH, height, slenderness, crown size, seed count and wood density. A subset of hard and soft traits has been used to test the methodological approach of heritability

estimation (see below section 3.2). As soon as others traits will be available, their heritability will be estimated following the same procedure (end of 2023).

3.2 Heritability estimates

Following the methodological approach employed by Alexandre et al., (2020), we implemented animal models to estimate the additive genetic variance (V_a) of the phenotypic traits and evaluate their heritability (h^2) and evolvability (I_a) with the following model:

$$Y = \mu + \alpha + X\beta + W\gamma + Za + \varepsilon$$

Where **Y** is the vector of phenotypic traits, **μ** is the population mean, **X, W, Z** and **S** are incidence matrix related to each effect, **β** is the fixed environmental effect, **γ** is the random spatial effect that accounts for the autocorrelation in the phenotypic trait values among individuals as a result of their spatial proximity, **a** is the random additive genetic effect associated to the genetic relationship matrix and **ε** is the residual effect.
To study the heritability of phenotypic traits, the main parameter we need to estimate is the variance of the additive genetic effect (V_a), which is extracted from the animal model. The way to estimate the heritability is by dividing V_a by the phenotypic variance (V_p). There are two ways to calculate heritability:

a) Calculated heritability (h^2_{calc}): dividing V_a by the variance of all the random effects included in the model.
b) Observed heritability (h^2_{obs}): dividing V_a by the variance computed directly form the phenotypes.

Besides, we evaluated the influence of the spatial and environmental effects on heritability estimates by comparing models including or not these effects. Furthermore, in a dioecious species such as *Populus nigra*, we decided to investigate differences in heritability estimates among sexes. Lastly, we estimated the evolvability of phenotypic traits as V_a/x^2, where x is the population mean phenotypic value. This estimate is less dependent on the environmental variance than the heritability estimates (Table 1, Figure 1).

These approaches were used to estimate the heritability of a subset of traits in the Spanish GCU. Because of some delay in genetic and environmental data production in the Austrian GCU, estimation of trait heritability has been delayed (see below). It is expected that the same procedure is applicable for the rest of populations and species once all three datasets will be available.

Table 1. Calculated and observed heritability estimates and evolvability of *Populus nigra* population in La Alfranca (Spain). The null models only include the genetic effect, the spatial models additionally include the spatial effect and the full model additionally include the spatial effect and environmental covariates.

	P50			Height		
h^2_{calc}	Total	Female	Male	Total	Female	Male
Null model	0.384	0.617	0.528	0.488	0.574	0.761
Spatial model	0.205	0.550	0.393	0.051	0.194	0.042
Full model	0.170	0.534	0.444	0.058	0.201	0.048
h^2_{obs}	Total	Female	Male	Total	Female	Male
Null model	0.453	0.815	0.667	0.585	0.740	1.015
Spatial model	0.221	0.706	0.471	0.066	0.274	0.058
Full model	0.174	0.673	0.541	0.063	0.232	0.050
Evolvability (%)	Total	Female	Male	Total	Female	Male
Null model	0.576	1.098	0.793	5.931	6.886	11.043
Spatial model	0.281	0.951	0.560	0.666	2.550	0.626
Full model	0.221	0.906	0.643	0.639	2.163	0.549

	DBH			Slenderness		
h^2_{calc}	Total	Female	Male	Total	Female	Male
Null model	0.397	0.557	0.656	0.462	0.455	0.457
Spatial model	0.044	0.264	0.062	0.305	0.289	0.267
Full model	0.046	0.284	0.022	0.300	0.290	0.272
h^2_{obs}	Total	Female	Male	Total	Female	Male
Null model	0.472	0.718	0.897	0.569	0.558	0.575
Spatial model	0.044	0.358	0.064	0.348	0.356	0.301
Full model	0.041	0.418	0.019	0.325	0.330	0.297
Evolvability (%)	Total	Female	Male	Total	Female	Male
Null model	10.489	15.383	20.468	13.069	7.434	18.785
Spatial model	0.979	7.669	1.455	7.986	4.734	9.825
Full model	0.913	8.967	0.443	7.461	4.390	9.713

	PC1 NIRs		
h^2_{obs}	Total	Female	Male
Null model	0.315	0.428	0.381
Spatial model	0.027	0.008	0.051
Full model	0.028	0.005	0.059

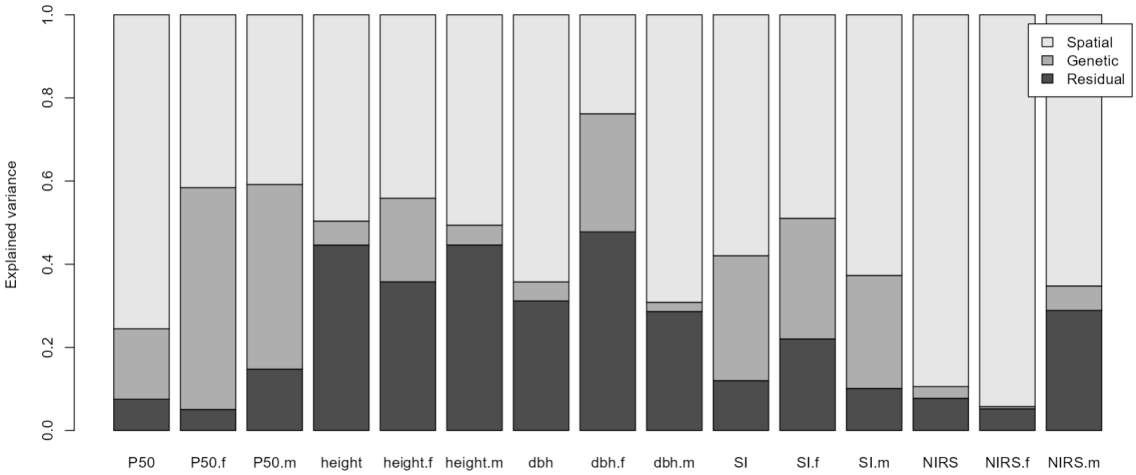

Figure 1. Proportion of explained variance for the studied traits between sexes and within females and males (denoted by "f" and "m", respectively) by spatial effect, genetic effect and residuals.

Beyond assessing the heritability of the studied traits, we are investigating the genetic correlations among them (Table 2). Positive and negative correlations among traits are expected to accelerate or delay the rate of evolution of the population. Given the striking differences found among sexes in the heritability estimates, assessing the genetic correlation separately by sex could provide insights into the genetic architecture of the traits and the possible different evolutionary paths that males and females will undergo.

Table 2. Trait-trait Pearson genetic (lower diagonal) and phenotypic correlation (upper diagonal). R values and p-values (in brackets) are presented for all individuals (total) and separated by sex.

Total	P50	height	dbh	Slenderness	NIRS
P50	1	-0.23 (<0.001)	-0.27 (<0.001)	0.15 (0.002)	-0.3 (<0.001)
height	-0.07 (0.168)	1	0.52 (<0.001)	0.22 (<0.001)	0.01 (0.810)
dbh	-0.22 (<0.001)	0.3 (<0.001)	1	-0.55 (<0.001)	-0.03 (0.569)
Slender.	0.08 (0.118)	0.44 (<0.001)	-0.5 (<0.001)	1	0.03 (0.517)
NIRS	-0.13 (0.006)	0.15 (0.002)	-0.16 (<0.001)	0.17 (<0.001)	1

Females	P50	height	dbh	Slenderness	NIRS
P50	1	-0.23 (<0.001)	-0.25 (<0.001)	0.12 (0.068)	-0.26 (<0.001)
height	-0.26 (<0.001)	1	0.54 (<0.001)	0.2 (0.003)	-0.09 (0.178)
dbh	-0.22 (0.001)	0.49 (<0.001)	1	-0.61 (<0.001)	-0.12 (0.073)
Slender.	0.02 (0.717)	0.36 (<0.001)	-0.53 (<0.001)	1	0.02 (0.719)
NIRS	-0.32 (<0.001)	0.11 (0.107)	-0.06 (0.386)	0.14 (0.045)	1

Males	P50	height	dbh	Slenderness	NIRS
P50	1	-0.21 (0.002)	-0.29 (<0.001)	0.18 (0.010)	-0.34 (<0.001)
height	0.15 (0.037)	1	0.5 (<0.001)	0.24 (<0.001)	0.11 (0.109)
dbh	-0.11 (0.116)	0.14 (0.047)	1	-0.54 (<0.001)	0.07 (0.349)
Slender.	0.16 (0.019)	0.5 (<0.001)	-0.54 (<0.001)	1	0.04 (0.582)
NIRS	-0.3 (<0.001)	0.01 (0.891)	-0.01 (0.916)	0.06 (0.410)	1

3.3 Delay and corrective measures

At this stage of the project all raw measurements are available for the two GCUs of *Populus nigra*, except for seed count and environmental data for which partial data are available (Table 3). This delay is mainly due to the challenge in capturing seed production in poplar using unmanned aerial vehicle, as seeds are released sequentially in time, and due to bad flight conditions during the drone campaigns. Other flights are planned for this species as soon as the UMR agenda will allow it (summer and autumn 2023).

Table 3. Available datasets at the time of writing the Deliverable. Projected availability of processed data (i.e. data ready to use for the heritability estimates) are indicated in parenthesis.

	Genomic data		Phenotypic data			Environmental data (UAV)
	Genotypes (adults + juveniles)	Parentage relationship & fecundity	Soft traits	NIRS proxy	Seed count (UAV)	
Spain	2022	2023	2022	Ongoing (end 2023)	Ongoing (end 2023)	Ongoing (end 2023)
Austria	Summer 2023	Ongoing (autumn 2023)	2022	Ongoing (end 2023)	Ongoing (end 2023)	Ongoing (end 2023)

4 Conclusions

- Comparison across models.

Results show the implications of including different kind of data in the models to estimate heritability. The importance of including the spatial distribution of the studied individuals drastically reduce the proportion of additive genetic variance, implying that a great part of the phenotypic and genetic resemblance between individuals is due to their spatial location. In the studied population, the environmental covariates have a lower impact on the heritability estimates than the spatial distribution.

- Comparisons between males and females.

Different genetic architecture between sexes is common in animal species, however little is known in plant species. The very few investigations on the variation of heritability estimates among sexes in plants are mostly focused on reproductive traits. The inclusion of a dioecious species in the project has shown that, for most of the studied traits, the heritability is higher in females than in males, suggesting that females may respond faster to natural selection. The

processes behind the observed differences among males and females are still unknown, but we are performing further analysis to understand the nature of this differences. For now, we have checked for differences in phenotypic values and the quantity of genetic relations between sexes and have found no evidence of this being the cause of the differences in heritabilities (see Annex 1). Further analysis of differences in genetic correlations between sexes will give insights in the factors behind this process. Besides, the comparison with the Austrian population when data is available will allow to investigate if the different heritabilities between sexes are consistent across the species range or if they are characteristic of the studied population.

- Genetic correlations across traits.

The presented results show that most correlations are similar among genetic and phenotypic trait values. However, there are some contrasts such as the lack of correlation of the phenotypic values of the NIRS PCA with any measured trait except P50, while genetic correlation highlight its relation with most of the studied traits when the whole population is considered. Besides, there is a change in sign in males corresponding to the correlation of height and P50, negative for the phenotypic values but positive for the genetic ones.

As evolution act on whole phenotypes, both, the phenotypic and genetic correlation may affect the rate and path of evolutionary processes and deserve further investigation.

- Methodological considerations and perspectives for the other site, and for the others species.

Statistical methods employed in the Spanish GCU will be easily transferrable to the new data that will be available in other GCUs and for different species. It is important to remark that, given the heterogeneity of the species under study, not all traits, fitness estimates and environmental variables will match among species and populations. In any case, these results will give information on the adaptive capacity and innovative *in situ* methods to assess it within and across species.

- Use of NIRS directly as a compound trait.

The use on NIRS as proxy of hard traits is an innovative procedure and could facilitate phenotyping individuals in natural populations. Furthermore, the reduction of the whole NIRs spectra into PCA components and studying it as a compound trait itself has, to our knowledge, never been done. The interpretation of this parameter as a functional trait is not straight-forward, however this approximation will open up the door for further research which would help in understanding the implication of this trait on the adaptation of tree species.

5 Partners involved in the work

INRAE, INIA-CSIC, CNR, UMR, BFW, CREAF

6 Annexes

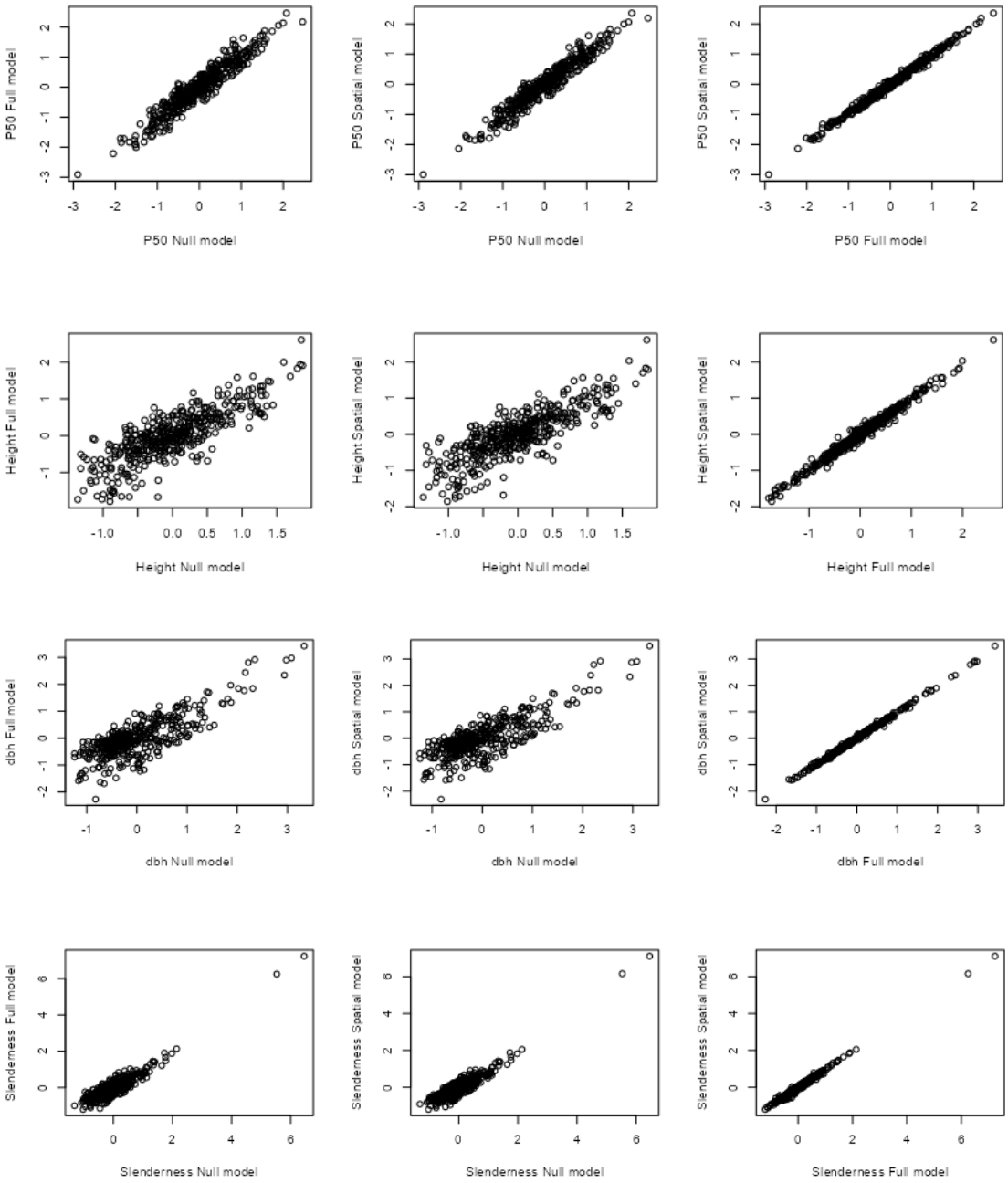

Figure A1. Scatter plot of the residuals of the null, full and spatial models by trait.

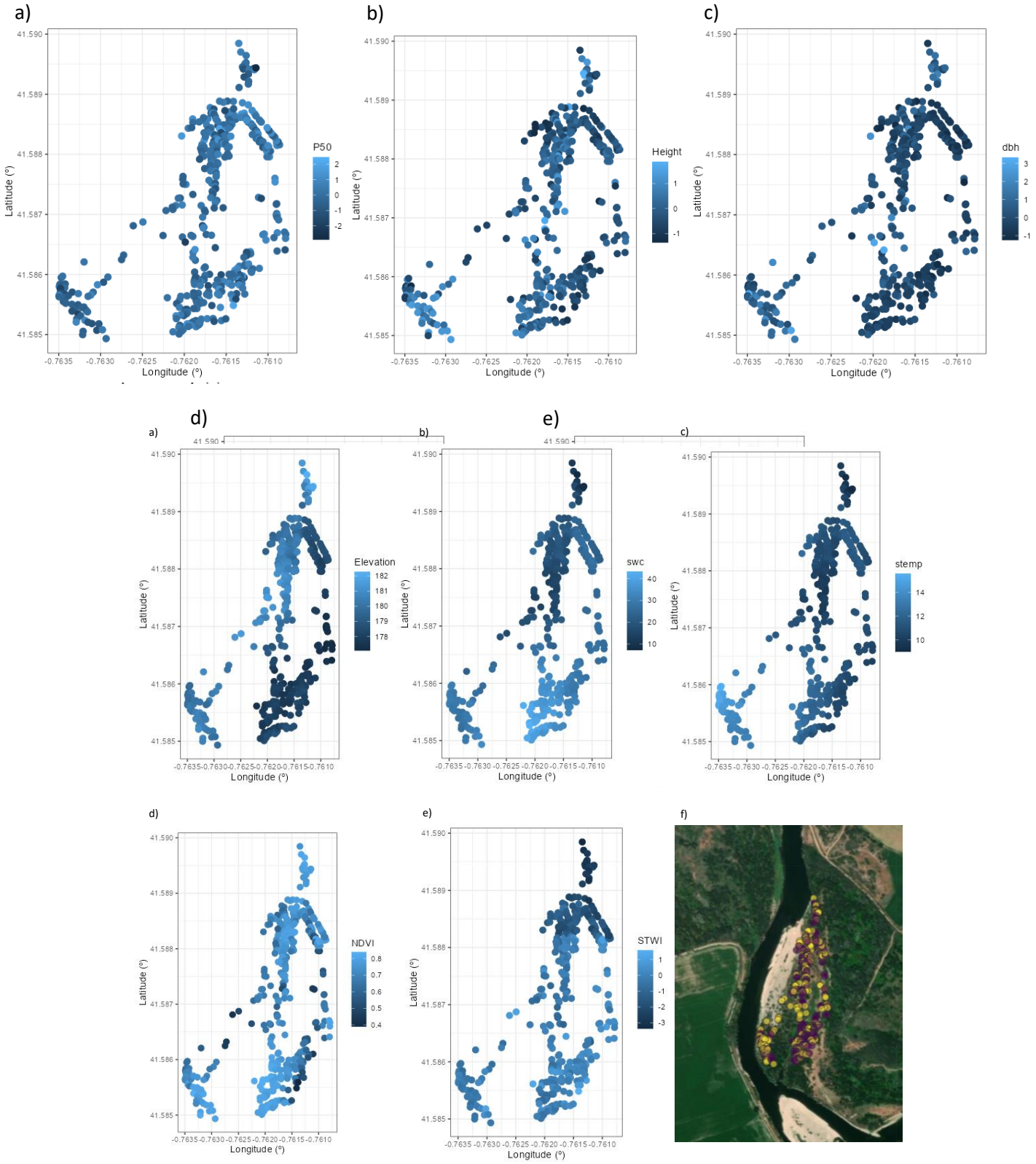

Figure A2. Spatial distribution of residuals from the null model for the studied traits

Table A1. Summary statistics, mean and standard deviation, of studied phenotypic traits by sex. No significant differences were found for any trait between sexes.

	Females		**Males**	
	Mean	Std. dev.	Mean	Std. dev.
P50	-2.121	0.246	-2.158	0.235
Height	12.839	3.917	13.444	4.435
dbh	26.814	12.414	28.667	13.694
Slenderness	53.402	19.486	54.925	31.404
NIRS	-11.030	18.814	-11.575	18.531
Fecundity	0.123	1.606	0.041	0.664
Genetic relatedness	0.006	0.063	0.006	0.069