



# Genomics of Mediterranean forest trees

A. Marchesini<sup>1,2</sup>, P. Pollegioni<sup>1,2</sup>,  
C. Mattioni<sup>1,2</sup>

*<sup>1</sup>Research Institute on Terrestrial Ecosystems,  
National Research Council, Porano (TR), Italy*

*<sup>2</sup>National Biodiversity Future Center, Palermo, Italy*



**NATIONAL  
BIODIVERSITY  
FUTURE CENTER**

## Genetic diversity and structure of *Alnus cordata* and *Quercus trojana* populations in the native range of southern Italy: an essential step towards the definition of management and conservation strategies

C. Mattioni, A. Marchesini, M. Gaudet, F. Chiocchini, L. Leonardi, M. Cherubini, P. Pollegioni P.

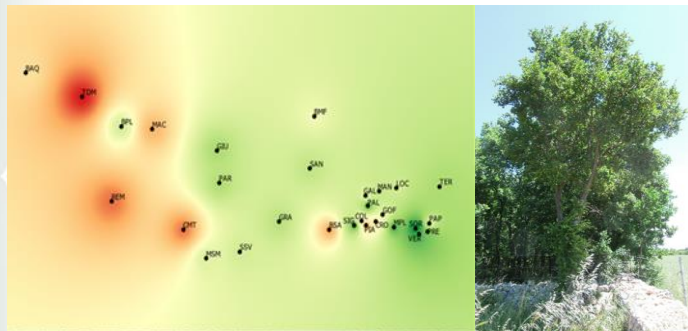
### Microsatellite genotyping for:

- Understanding demography and genetic structure of natural populations
- Identify hot spots of neutral diversity and conservation units
- Reconstruct evolutionary history
- integrate the genetic/genomic data with climatic and topographic data

Priority conservation areas

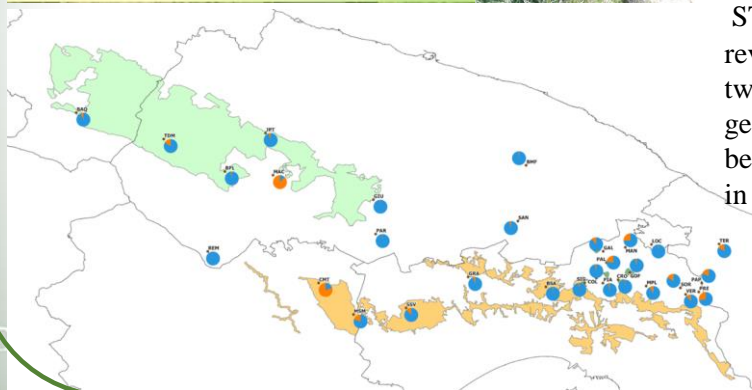
Management strategies

### *Quercus trojana*



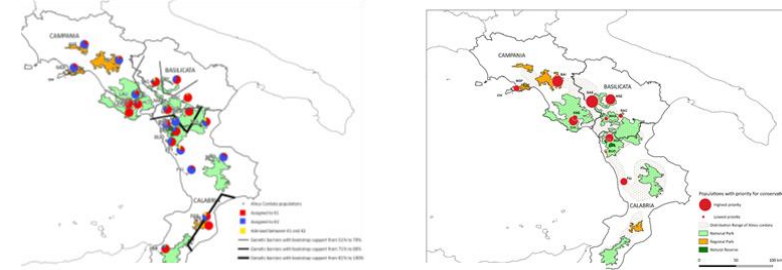
30 populations,  
552 samples,  
12 SSRs markers

STRUCTURE analysis revealed the presence of two main gene pools, genetic divergence between two populations in the core of distribution



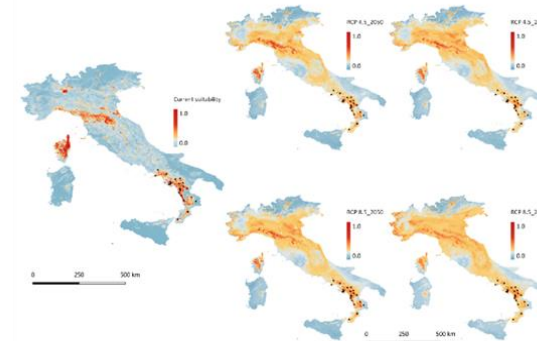
STRUCTURE identified two genetic clusters weakly correlated with the elevation

### *Alnus cordata*



23 populations,  
517 samples,  
10 SSRs markers

Indication of priority conservation areas



Predicted current (A) and future potential geographic distribution of *Alnus cordata* by 2050 (range 2041-2060)



# Genomics of Mediterranean forest trees at CNR-IRET:

## Activities:

**(1) GENERATION OF REFERENCE GENOMES**

**(2) FOREST GENOMIC AND PHYLOGENOMIC STUDIES**

## Target organisms:

- (a) Mediterranean oaks (*Quercus* spp.)
- (b) Italian alder (*Alnus cordata*)
- (c) European chestnut (*Castanea sativa*)

# GENERATION OF REFERENCE GENOMES

## Reference genomes are key scientific resources

- allow to perform genomic studies based on methods which rely on the mapping of whole-genome sequence reads to a known reference sequence
- provide the functional and genomic contexts for regions under selection, enabling the association of such loci with phenotypes important for adaptation and resilience
- preserve genomics information for species at risk of extinction
- contribute to new discoveries

Brandies et al. (2019). The value of reference genomes in the conservation of threatened species. *Genes*, 10(11), 846

Formenti et al. (2022). The era of reference genomes in conservation genomics. *Trends in ecology & evolution*, 37(3), 197-202.

Mc Cartney et al. (2024). The European Reference Genome Atlas: piloting a decentralised approach to equitable biodiversity genomics. *npj Biodiversity*, 3(1), 28.

## Generation of reference genomes: *Castanea sativa*

**Objective:** generating the first reference genome for the sweet chestnut *Castanea sativa* Mill., the only native *Castanea* species in Europe, a tree of high economic value

### Activity in collaboration with:

- Fondazione Edmund Mach, Research and Innovation Center (San Michele all'Adige, TN)
- CNR-IPSP (Sesto Fiorentino)
- University of Bologna, Dept. of Agricultural and Food Science)
- University of Turin, Dept. of Agricultural, Forest and Food Sci)
- Marche Polytechnic University, Dep. of Agricultural, Food and Env.Sci

**Methods:** DNA of the ancient Italian variety 'Marrone di Chiusa Pesio'; genome assembled using a combination of:

- long reads sequencing (Oxford Nanopore)
- short reads sequencing (Illumina PE150; 89 Gb)
- Hi-C sequencing to capture chromatin conformation (Omni-C; Dovetail Genomics)

### Results:

- high-quality genome; 12 chromosomes, for a total genome size of about 715 Mb;
- BUSCO assessment: 98.6% of the genome matched the embryophyte dataset, highlighting good completeness
- Gene prediction and annotation: a total of 57,653 and 58,146 genes were predicted in the two haplotypes, and approximately 73% of the genes were functionally annotated

## Generation of reference genomes: *Castanea sativa*

- Our annotated de novo assembly was published and deposited in GenBank as reference genome for the species: GCA\_040712315.1

Bianco et al. *BMC Genomic Data* (2024) 25:64  
<https://doi.org/10.1186/s12863-024-01245-7> BMC Genomic Data

**DATA NOTE** **Open Access**



**The de novo, chromosome-level genome assembly of the sweet chestnut (*Castanea sativa* Mill.) Cv. Marrone Di Chiusa Pesio**

Luca Bianco<sup>1†</sup>, Paolo Fontana<sup>1†</sup>, Alexis Marchesini<sup>2,3†</sup>, Sara Torre<sup>4</sup>, Mirko Moser<sup>1</sup>, Stefano Piazza<sup>1</sup>, Sara Alessandri<sup>5</sup>, Vera Pavese<sup>6</sup>, Paola Pollegioni<sup>2,3</sup>, Cristiano Vernesi<sup>1</sup>, Mickael Malnoy<sup>1</sup>, Daniela Torello Marinoni<sup>6</sup>, Sergio Murolo<sup>7</sup>, Luca Dondini<sup>5</sup>, Claudia Mattioni<sup>2,3</sup>, Roberto Botta<sup>6</sup>, Federico Sebastiani<sup>4</sup>, Diego Micheletti<sup>1</sup> and Luisa Palmieri<sup>1\*</sup>

**Genome assembly ASM4071231v1** reference

[Download](#) [datasets](#) [URL](#) [FTP](#)

Submitted GenBank assembly GCA\_040712315.1

Taxon [Castanea sativa](#) (European chestnut)

### Assembly statistics

	GenBank
Genome size	715.9 Mb
Total ungapped length	715.9 Mb
Number of chromosomes	12
Number of scaffolds	12
Scaffold N50	59.9 Mb
Scaffold L50	5
Number of contigs	77
Contig N50	25 Mb
Contig L50	9
GC percent	35.5
Genome coverage	200.0x
Assembly level	Chromosome





## Generation of reference genomes: *Quercus trojana*

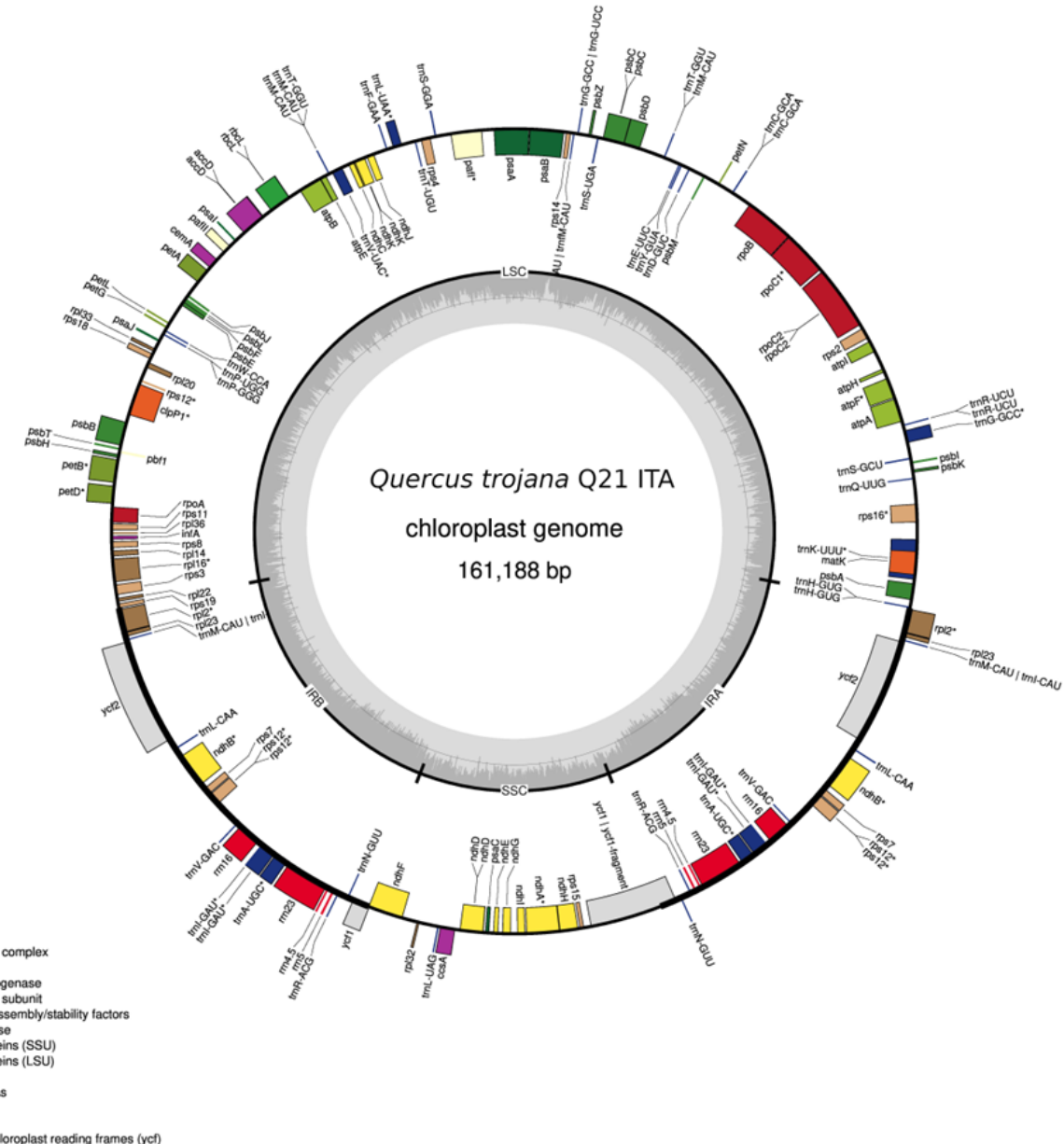
**Objective:** generating the first chloroplast and nuclear reference genomes for the Macedonian oak (*Quercus trojana*), an eastern Mediterranean endemism

### Chloroplast genome

- Whole-genome sequencing (Illumina NovaSeq; PE150; coverage 5x)
- genome assembly using GetOrganelle (Jin et al. 2020); genome annotation using Geseq (Tillich et al. 2017)
- the whole chloroplast genome was assembled and annotated for the first time (sample: Q2; Italy, Murge)

### Nuclear genome

- planned for 2025; in collaboration with: University of Udine, Istituto di Genomica Applicata (IGA), Udine



## Generation of reference genomes: *Alnus cordata*

**Objective:** generating the first reference genome for the Italian alder (*Alnus cordata* Loisel.), endemic to south-western Apennines and Corsica. Note: due to its nitrogen-fixing ability, the species is ideal for reforestation, improvement of degraded lands and reducing the risk of landslides (NBS)

### Activity in collaboration with:

- UNIUD, Dept. of Agricultural, Food, Environmental and Animal Sciences;
- Istituto di Genomica Applicata, IGA, Udine

### Methods:

- sequences generated with Oxford Nanopore technology (ONT)
- haplotype-resolved assembly with *hifiasm*

### Coming up next:

- reordering contigs; refining the assembly
- gene prediction and annotation (for high accuracy, we need RNA-seq data)

### Draft genome

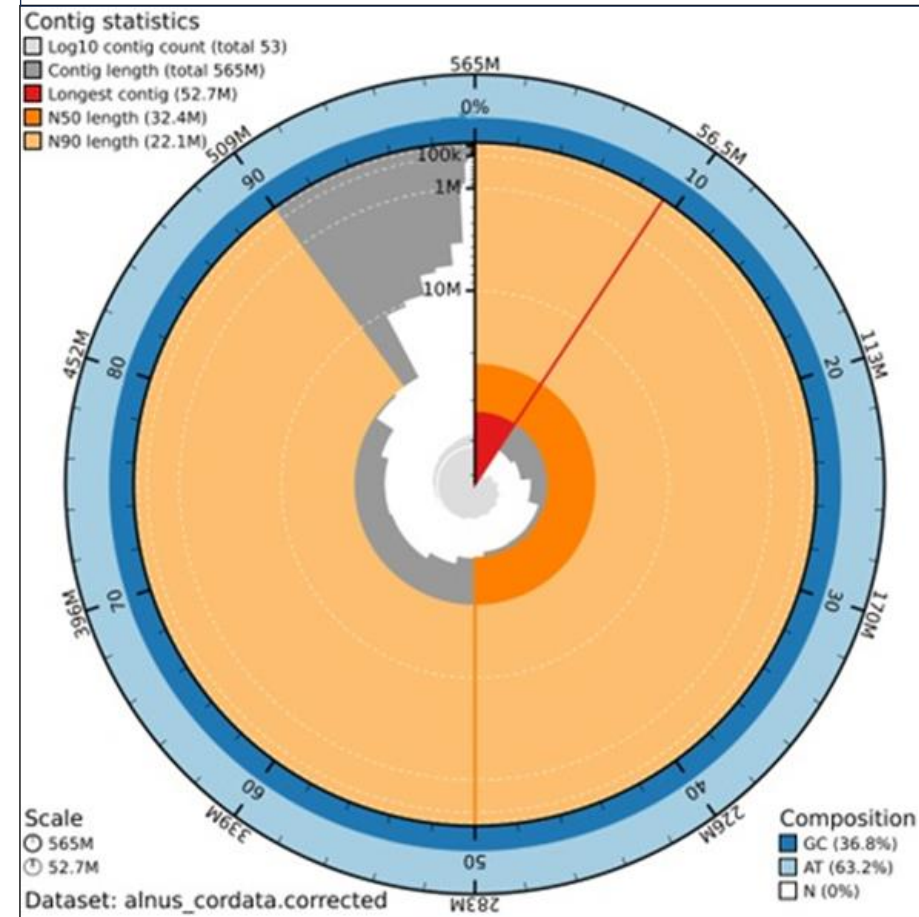
No. of contigs: 53  
No. of T2T contigs: 15  
(52,7 e 22,1 Mbp)  
N50 (Mbp): 32  
Total assembly length (Mbp): 565  
No. haplotype-resolved chromosomes: 14

### ONT long reads

10,16 M reads  
73.68 Gb

### ONT ultra-long reads

2,86 M reads  
75.51 Gb





# Genomic and phylogenomic studies: *Quercus* sect. *Cerris*

**Study species:** western Eurasian oaks of sect. *Cerris*; special focus on *Q. trojana*, an eastern Mediterranean endemism

**Methods:** phylogenomics and phylogeographic analysis based on **whole chloroplast genomes**, assembled starting from **genome skimming** data (i.e., WGS, about 5x coverage)

**Activity in collaboration with:** CNR-IPSP; Università degli Studi della Tuscia, Viterbo (DAFNE)

## Oak species included in this study (bold)

Table from:  
Simeone et al. (2018).  
PeerJ 6: e5793



**Table 1** Species and taxa included in *Quercus* Section *Cerris*. Nomenclature followed [Govaerts & Frodin \(1998\)](#); species investigated in the present study are bolded. Taxonomic remarks and species distributions according to [\\*Govaerts & Frodin \(1998\)](#) and [\\*\\*Menitsky \(2005\)](#).

Taxon	Taxonomic remarks	Distribution
<i>Q. afares</i> Pomel		Endemic, Algerian and Tunisian Tell Atlas
<i>Q. brantii</i> Lindl.		S and SE Anatolia to Iran, Lebanon
<i>Q. castaneifolia</i> C.A. Mey		Endemic; SE Caspian Sea, Azerbaijan to Iran
<i>Q. cerris</i> L.		E and C Mediterranean, Balkans
<i>Q. crenata</i> Lam.	*Poorly known	Endemic, Italian peninsula
<i>Q. libani</i> Oliv.		SE Anatolia to Iran
<i>Q. look</i> Kotschy	*Synonym of <i>Q. ithaburensis</i> ssp. <i>ithaburensis</i> ; **hybrid <i>Q. ithaburensis</i> × <i>Q. libani</i>	Endemic, Lebanon to Anti-Lebanon mountain range
<i>Q. ithaburensis</i> Decne.	**Including subsp. <i>macrolepis</i> (Kotschy), distributed in the European part of the range, and subsp. <i>ithaburensis</i> (Decaisne), in the Middle East	E Mediterranean, SE Italy to Palestine
<b><i>Q. trojana</i> Webb</b>	*Including subsp. <i>trojana</i> and subsp. <i>euboica</i> (Papaioann.) K.I.Chr., endemic of Euboea (Greece)	Anatolia, Aegean to SE Italy
<i>Q. suber</i> L.		C and W Mediterranean
<i>Q. acutissima</i> Carruth.		E and SE Asia
<i>Q. chenii</i> Nakai		E Asia
<i>Q. variabilis</i> Blume		E and SE Asia, Japan

## Genomic and phylogenomic studies: *Quercus* sect. *Cerris*

- **Dataset:**

south-eastern Mediterranean species (most of them are endemics):

*Q. trojana* (20 provenances; 11 Italian samples, covering the whole Italian range - Murge region)

*Q. trojana* subsp. *euboica* (2 provenances)

*Q. cerris* (4 provenances)

*Q. suber* (2 provenances)

*Q. macrolepis* (3 provenances)

*Q. ithaburensis* (3 provenances)

*Q. afares* (1 provenance)

*Q. brantii* (1 provenance)

*Q. castaneifolia* (1 provenance)

*Q. crenata* (1 provenance)

*Q. libani* (1 provenance)

*Q. look* (1 provenance)

- **Results:** 40/41 samples successfully assembled (whole chloroplast genomes)

### Geographic distribution of oak samples included in this study

(red ellipse)

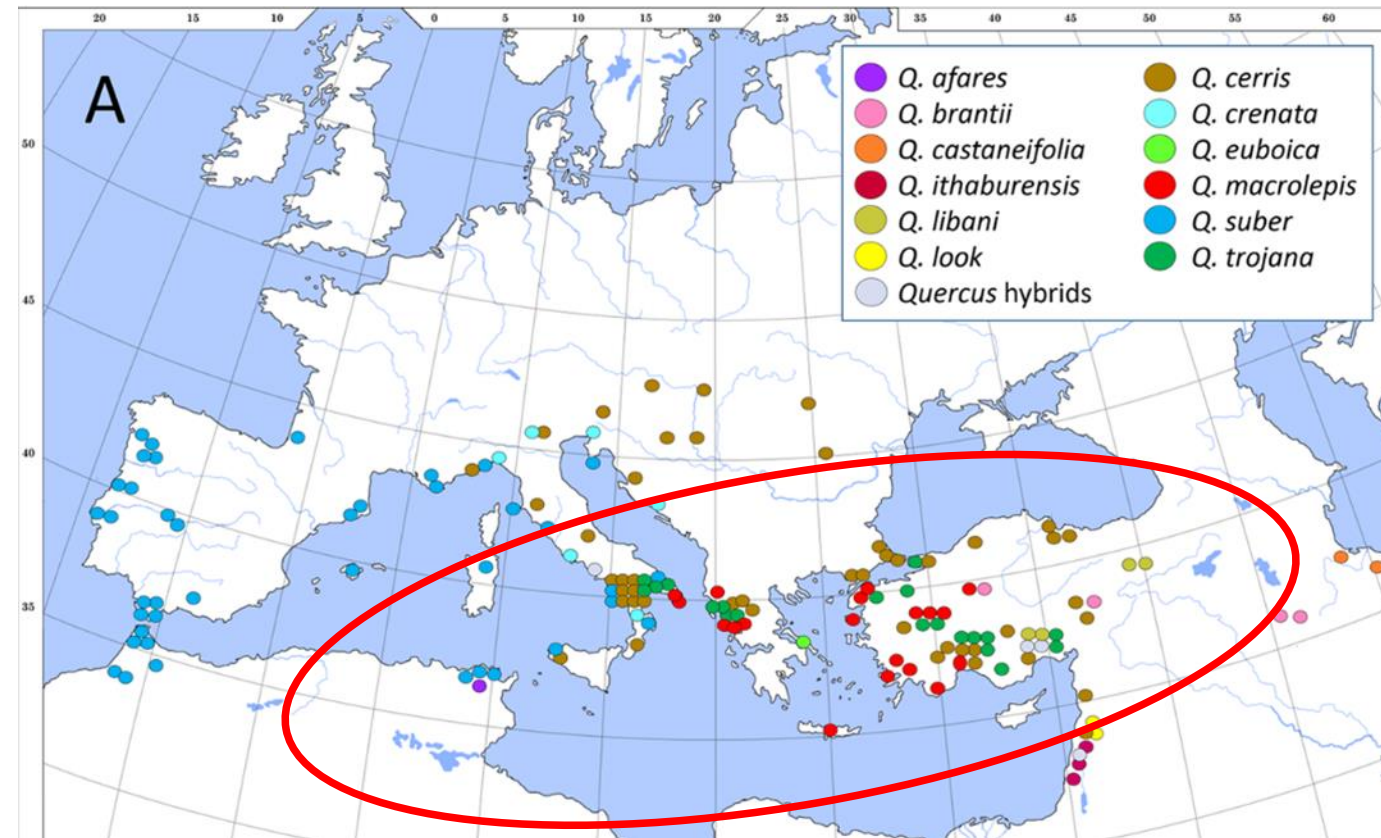
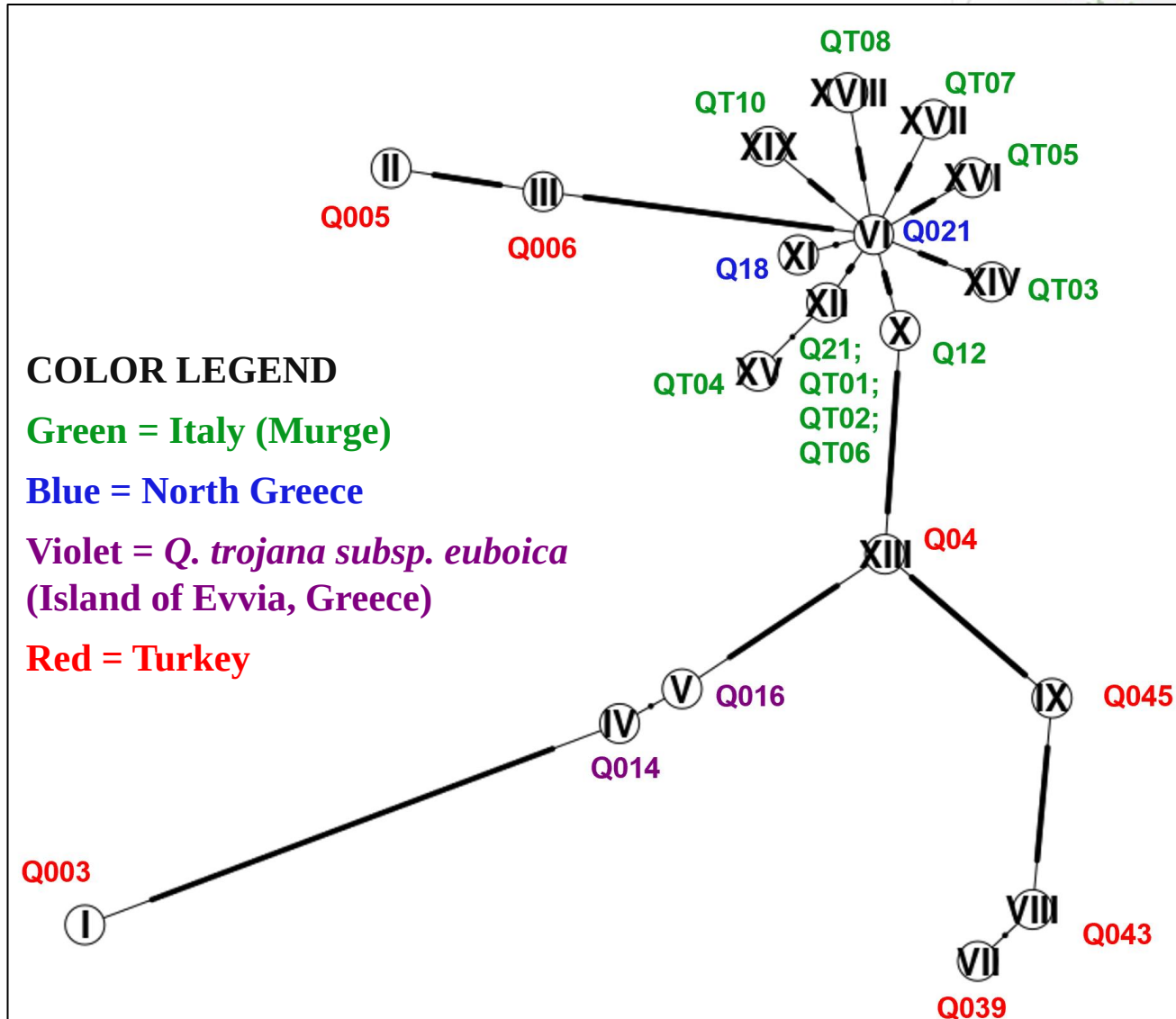


Figure from: Simeone et al. (2018). PeerJ 6: e5793

## Genomic and phylogenomic studies: *Quercus* sect. *Cerris*

- All *Q. trojana* Italian samples belong to the same haplogroup;
- However, there is relatively high diversity (8 different haplotypes)
- Italian samples seem to be closely related to samples from Northern Greece
- Anatolian samples belong to different clades spread throughout the network; this is not surprising: they correspond to geographically distant and isolated areas
- The phylogenetic relationships between the different *Q. trojana* haplogroups will be better investigated through Bayesian phylogenetic tree, including all the other *Q. sect. cerris* species (work in progress)

## Parsimony network based on whole cp genomes (focus on *Q. trojana*)





## Genomic and phylogenomic studies: *Quercus* sect. *Ilex*

- **Study species:** Mediterranean evergreen oaks of sect. *Ilex*
- **Methods:** phylogenomic and phylogeographic analysis based on whole chloroplast genomes, assembled starting from genome skimming data (i.e., WGS, about 5x coverage)
- **Activity in collaboration with:** Università degli Studi della Tuscia, Viterbo (DAFNE)
- **Preliminary results:** all the 12 sequenced samples were successfully assembled (whole chloroplast genomes);
- **Coming up next:** comparative genomics (with publicly available cp genomes of non-Mediterranean *Quercus* sect. *Ilex*) and phylogenomics

sample	species/provenance	circular cp genome
L090	<i>Q. look</i>	X
QS01	<i>Q. alnifolia</i> (Aln 3)	X
QS02	<i>Q. aucheri</i> (Au 3)	X
QS03	<i>Q. coccifera</i> (Manavgat; Turkey)	X
QS04	<i>Q. coccifera</i> (Ikaria; Greece)	X
QS05	<i>Q. ilex</i> (Anadolu; Turkey)	X
QS06	<i>Q. coccifera</i> (Portugal)	X
QS07	<i>Q. coccifera</i> (Caltagirone; Sicily)	X
QS08	<i>Q. coccifera</i> (Tunisia)	X
QS09	<i>Q. ilex</i> (Collesalveti; Italy)	X
QS10	<i>Q. ilex</i> (Assemini; Sardinia)	X
QS11	<i>Q. ilex</i> Leon (Spain)	X
QS12	<i>Q. ilex</i> (Algeria)	X

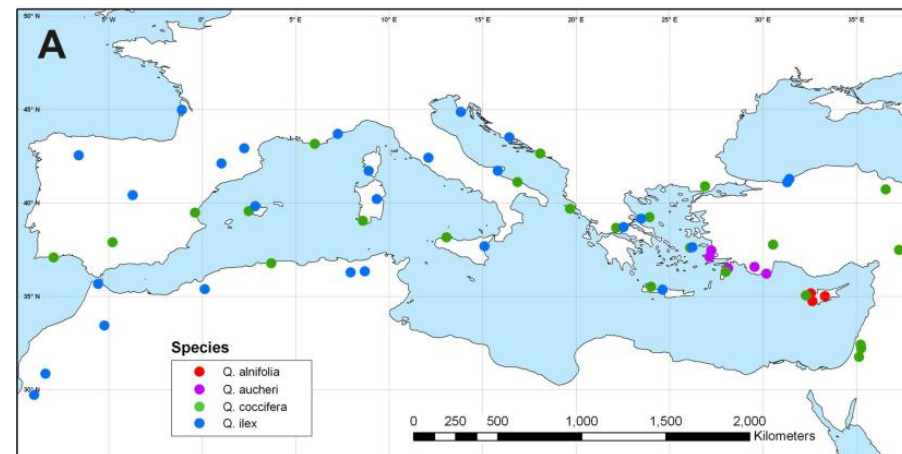


Figure from: Simeone et al. (2016). PeerJ, 4, e1897

# Genomic and phylogenomic studies: *Alnus cordata* and *Alnus glutinosa*

## Study species and sampling design:

- 89 *Alnus cordata* samples, collected from 19 locations covering the whole Italian distributional range;
- 26 *Alnus glutinosa*, from forest stands in close proximity to *A. cordata*, were included in the analysis to control for potential hybridization/introgression events

## Methods:

- whole-genome sequencing (WGS, about 20x coverage);
- de novo plastome assembly + analysis of nuclear variants (SNPs), after mapping reads on ref. genome

## Objectives:

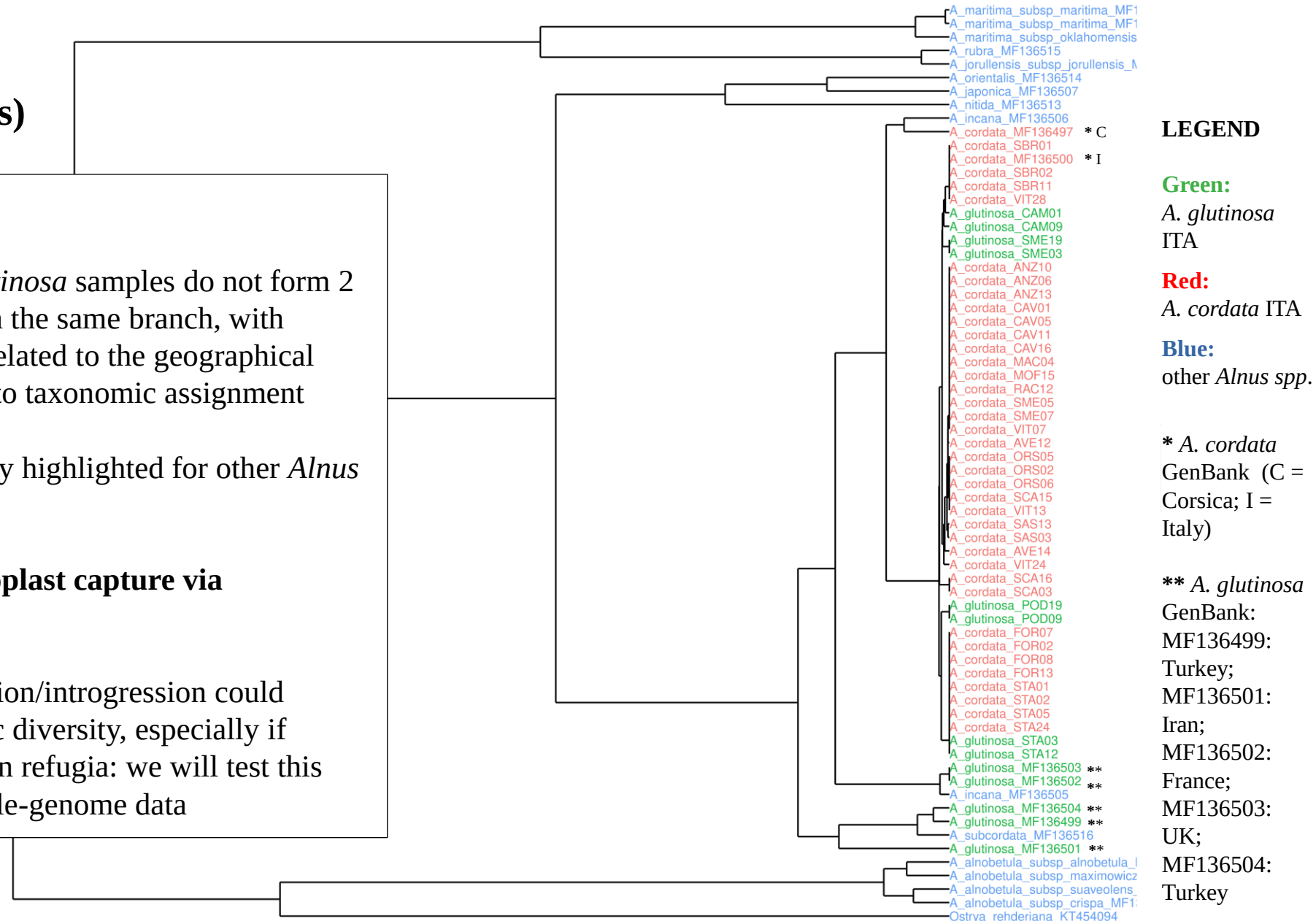
- investigating genome-wide genetic diversity and population structure in *Alnus cordata* populations
- identifying Genetic Conservation Unit (GCU), diversity hotspots and conservation priorities
- checking for potential hybridization/introgression events with the sympatric *Alnus glutinosa*, and evaluating their effects on *A. cordata* genome-wide diversity

# Genomic and phylogenomic studies: *Alnus cordata* and *Alnus glutinosa*

## NJ tree (whole cp genomes)

### Preliminary results

- Italian *A. cordata* and *A. glutinosa* samples do not form 2 distinct clades: they are all in the same branch, with substructure which is more related to the geographical provenance of samples than to taxonomic assignment
- a similar scenario was already highlighted for other *Alnus* spp. by Gryta et al. (2017)
- potential explanation: **chloroplast capture via introgressive hybridization**
- Coming up next:** hybridization/introgression could contribute to increase genetic diversity, especially if several species co-occurred in refugia: we will test this hypothesis with nuclear whole-genome data







## Genomic and phylogenomic studies: *Castanea sativa*

### Sampling design:

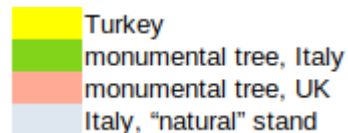
- 4 Italian, natural chestnut stands (Calabria, Lazio, Sicilia, Toscana); 5 individuals per stand;
- 6 monumental trees in the surrounding areas
- 1 sample from West Turkey, 1 sample from East Turkey; 2 samples from England (same tree)

### Methods:

whole-genome sequencing (WGS, about 10x coverage);  
plastome assembly + analysis of nuclear variants (SNPs), after mapping reads on our assembled ref. genome

### Activity in collaboration with:

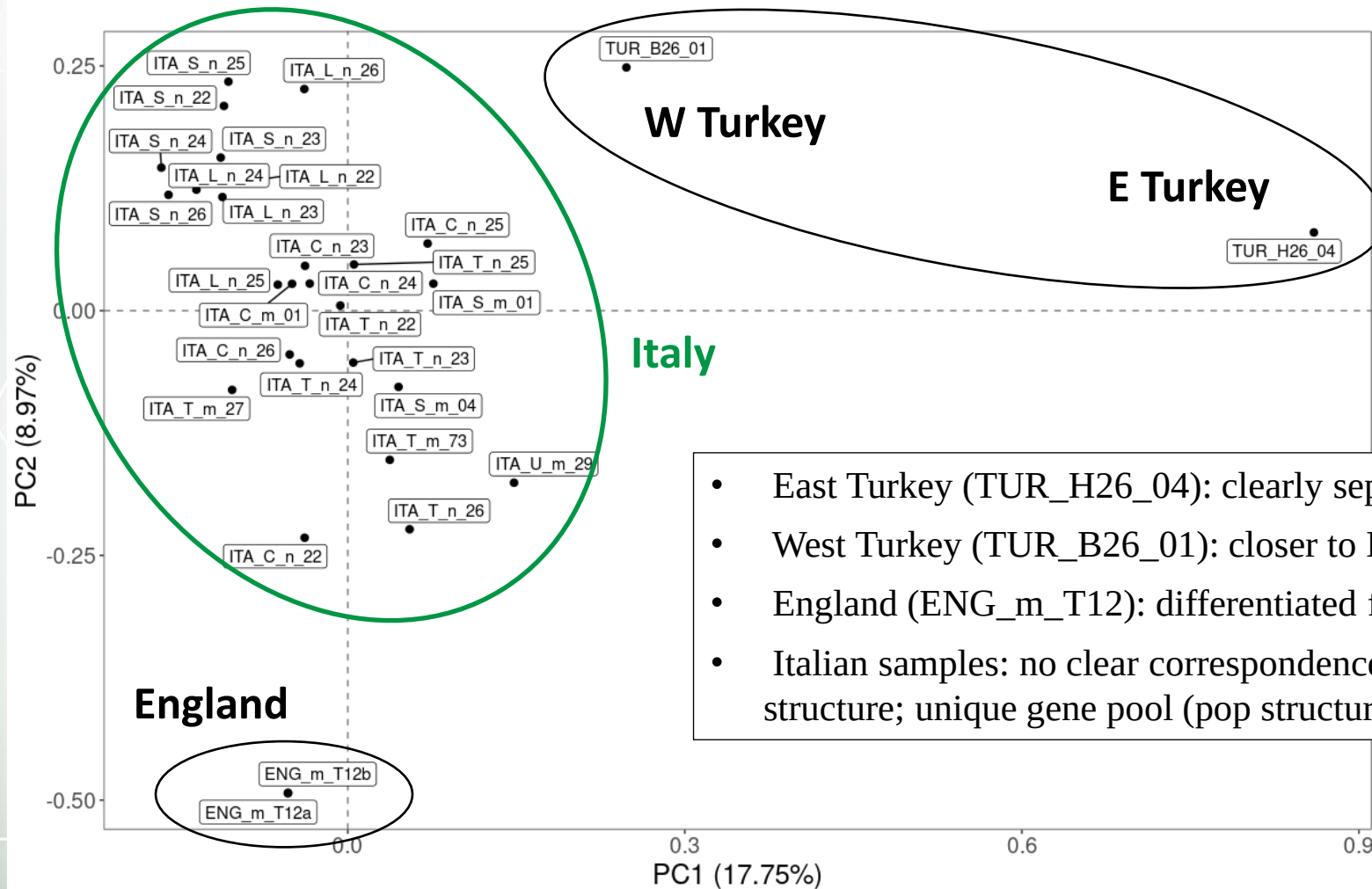
CNR-IPSP; Fondazione Edmund Mach



Raw reads name	Recoded as	Provenance
B26_1	TUR_B26_01	Turkey West, Bursa
H26_4	TUR_H26_04	Turkey East, Jopa
M1_cal	ITA_C_m_01	Calabria
M1_nave	ITA_S_m_01	Sicily, Nave
M27	ITA_T_m_27	Tuscany, Castagno Miraglia
M29	ITA_U_m_29	Umbria, Stroncone 1
M4	ITA_S_m_04	Sicily, Cento Cavalli
S73M_B	ITA_T_m_73	Tuscany, Voltiglione
T12a	ENG_m_T12a	UK, Torthwort chestnut
T12b	ENG_m_T12b	UK, Torthwort chestnut
C22	ITA_C_n_22	Italy, Calabria
C23	ITA_C_n_23	Italy, Calabria
C24	ITA_C_n_24	Italy, Calabria
C25	ITA_C_n_25	Italy, Calabria
C26	ITA_C_n_26	Italy, Calabria
L22	ITA_L_n_22	Italy, Latium
L23	ITA_L_n_23	Italy, Latium
L24	ITA_L_n_24	Italy, Latium
L25	ITA_L_n_25	Italy, Latium
L26	ITA_L_n_26	Italy, Latium
S22	ITA_S_n_22	Italy, Sicily
S23	ITA_S_n_23	Italy, Sicily
S24	ITA_S_n_24	Italy, Sicily
S25	ITA_S_n_25	Italy, Sicily
S26	ITA_S_n_26	Italy, Sicily
T22	ITA_T_n_22	Italy, Tuscany
T23	ITA_T_n_23	Italy, Tuscany
T24	ITA_T_n_24	Italy, Tuscany
T25	ITA_T_n_25	Italy, Tuscany
T26	ITA_T_n_26	Italy, Tuscany

## Genomic and phylogenomic studies: *Castanea sativa*

**Nuclear genome:** PCA based on whole-genome SNPs

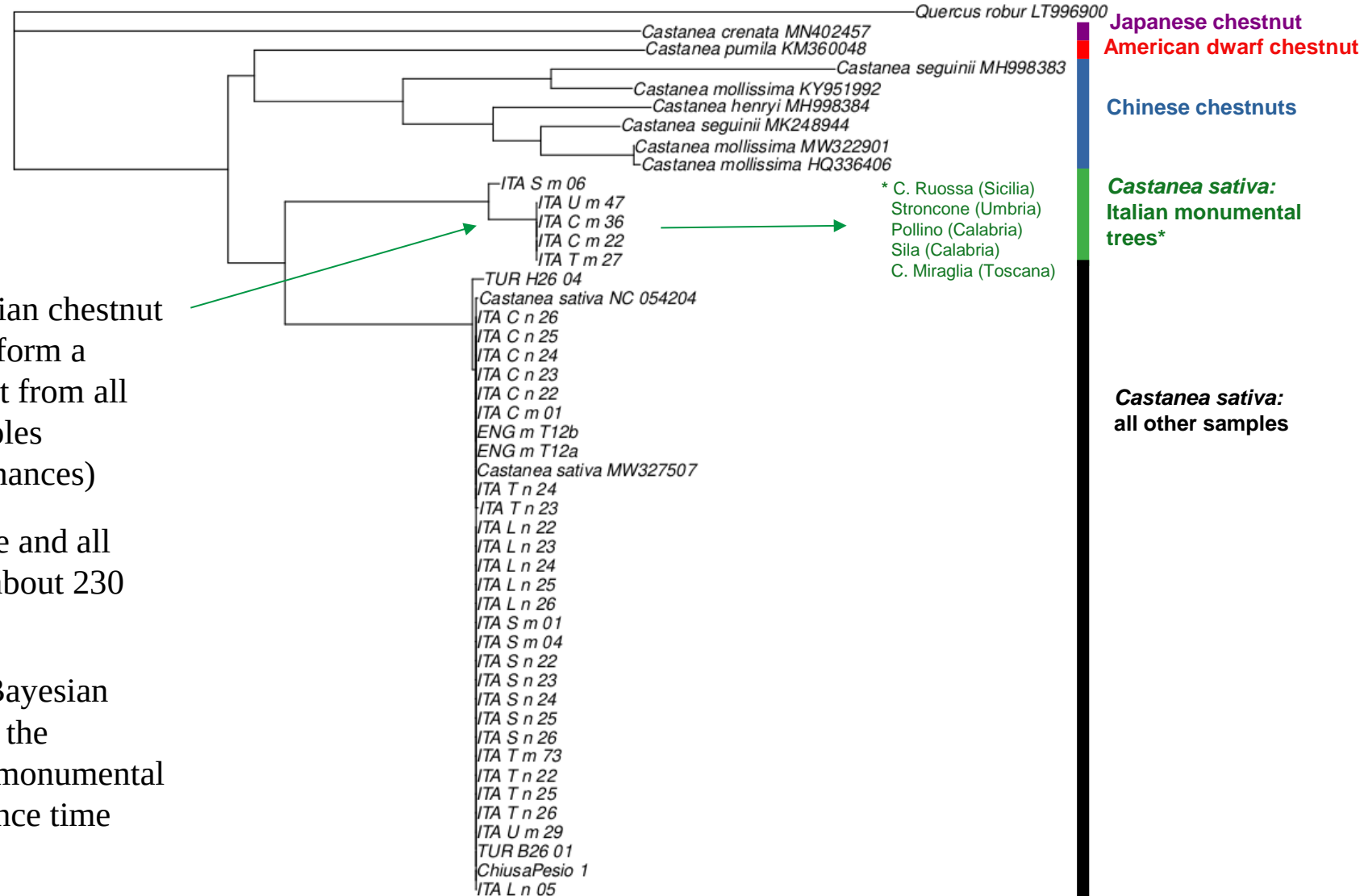


- East Turkey (TUR\_H26\_04): clearly separated from all the others on PC1 axis
- West Turkey (TUR\_B26\_01): closer to Italian samples
- England (ENG\_m\_T12): differentiated from Italian samples on PC2 axis
- Italian samples: no clear correspondence between genetic and geographical structure; unique gene pool (pop structure need to be better investigated)

# Genomic and phylogenomic: *Castanea sativa*

## Chloroplast genome: Neighbor joining (NJ) tree

- The cp haplotypes of some Italian chestnut samples (5 monumental trees) form a separate clade, highly divergent from all the other *Castanea sativa* samples (including the 2 Turkish provenances)
- The distance between this clade and all other *C. sativa* is remarkable: about 230 nucleotide differences
- Coming up next:** ML and/or Bayesian phylogeny to better investigate the evolutionary history of Italian monumental chestnuts and estimate divergence time





# Why study and monitoring genetic diversity?

nature

[Explore content](#) ▾

[About the journal](#) ▾

[Publish with us](#) ▾

[nature](#) > [articles](#) > article

Article | [Open access](#) | Published: 29 January 2025

## Global meta-analysis shows action is needed to halt genetic diversity loss

[Robyn E. Shaw](#), [Katherine A. Farquharson](#), [Michael W. Bruford](#), [David J. Coates](#), [Carole P. Elliott](#), [Joachim Mergeay](#), [Kym M. Ottewell](#), [Gernot Segelbacher](#), [Sean Hoban](#), [Christina Hvilsom](#), [Sílvia Pérez-Espona](#), [Dainis Runģis](#), [Filippos Aravanopoulos](#), [Laura D. Bertola](#), [Helena Cotrim](#), [Karen Cox](#), [Vlatka Cubric-Curik](#), [Robert Ekblom](#), [José A. Godoy](#), [Maciej K. Konopiński](#), [Linda Laikre](#), [Isa-Rita M. Russo](#), [Nevena Veličković](#), [Philippine Vergeer](#), [Carles Vilà](#), [Vladimir Brajkovic](#), [David L. Field](#), [William P. Goodall-Copestake](#), [Frank Hailer](#), [Tara Hopley](#), [Frank E. Zachos](#), [Paulo C. Alves](#), [Aleksandra Biedrzycka](#), [Rachel M. Binks](#), [Joukje Buiteveld](#), [Elena Buzan](#), [Margaret Byrne](#), [Barton Huntley](#), [Laura Iacolina](#), [Naomi L. P. Keehnen](#), [Peter Klinga](#), [Alexander Kopatz](#), [Sara Kurland](#), [Jennifer A. Leonard](#), [Chiara Manfrin](#), [Alexis Marchesini](#), [Melissa A. Millar](#), [Pablo Orozco-terWengel](#), [Jente Ottenburghs](#), [Diana Posledovich](#), [Peter B. Spencer](#), [Nikolaos Tourvas](#), [Tina Unuk Nahberger](#), [Pim van Hooft](#), [Rita Verbylaite](#), [Cristiano Vernesi](#) & [Catherine E. Grueber](#) 

# Thank you!



NATIONAL  
BIODIVERSITY  
FUTURE CENTER

## ACKNOWLEDGEMENTS

- NBFC (National Biodiversity Future Center) - Project funded by the European Union - NextGenerationEU
- Marcello Cherubini, Luca Leonardi, Michele Mattioni (CNR-IRET Porano): sampling, lab work and IT support
- Muriel Gaudet, Francesca Chiocchini (CNR-IRET Porano); Paola Mairota (UNIBA): *Alnus* spp. and *Quercus* spp. sampling
- Marco Cosimo Simeone (UNITUS, DAFNE); Mariangela Santorsola (UNIPV): *Quercus* spp. sampling and phylogenomics
- Federico Sebastiani, Sara Villa (CNR-IPSP): phylogenomic analysis
- Janik Fernando (MSc student), Tiziana Castrignanò (UNITUS): *Alnus* genomics analysis
- Luca Bianco, Paolo Fontana, Stefano Piazza, Mirko Moser, Cristiano Vernesi, Mickael Malnoy, Diego Micheletti, Luisa Palmieri (Fondazione Edmund Mach, San Michele A./A., TN); Federico Sebastiani, Sara Torre (CNR-IPSP Sesto Fiorentino); Luca Dondini, Sara Alessandri (UNIBO, Dept. of Agricultural and Food Sciences); Roberto Botta, Vera Pavese, Daniela Torello Marinoni (UNITO, Dept. of Agricultural, Forest and Food Sci.); Sergio Murolo (UNIVPM, Dept. of Agricultural, Food and Env.Sci.): *Castanea sativa* reference genome assembly and genomic analysis
- Alberto Biscontin<sup>1</sup>, Michele Morgante<sup>1,2</sup> (<sup>1</sup>UNIUD, Dept. of Agricultural, Food, Environmental and Animal Sciences; <sup>2</sup>Istituto di Genomica Applicata, IGA, Udine): *Alnus cordata* reference genome assembly
- G-BikE (Genomic Biodiversity knowledge for Resilient Ecosystems) COST Action