Université Paul Cézanne

Faculté des Sciences et Techniques

Master 2 Professionnel : Expertise Ecologique et Gestion de la Biodiversité

Avenue Escadrille Normandie Niémen

F-13.397 Marseille Cedex 20

Investigation of the recent recolonisation of Beech on Mont Ventoux using historical records, vegetation analyses from satellite image and landscape genetics.



Prouillet-Leplat Hélène

hprouillet@gmail.com

Institut National de Recherche Agronomique

Site Agroparc Domaine Saint Paul

84914 AVIGNON Cedex 9

Directrice de stage : Dr Sylvie Oddou-Muratorio

Acknowledgements

Fist of all, thank to my great supervisor, Sylvie!, for her help during my work at INRA Avignon and the all her time spent on my project (even during the weekend!). Tack så mycket Sylvie!

Thank to all the team of the Archives Départementales du Vaucluse of the Popes' Palace in particular Rozet Jean-Michel, Carlettti Laurent and Auvray Nicolas for their help and their good advices during my historical researches.

Thank as well to forestry workers of the URFM unit (Olive, Medhi, Anthony...), for their team spirit, their solar beam humour and especially their kindness.

Thank to Hendrik for all the hours spent on the Spot satellite images analysis and your fight with GRASS and to Gwendal for his essential second reading especially on the genetic part of my report: your advices were more than welcome!

Many thanks to Sandrine annd Aurore for their good moon and their presence in the hard times of my training period, Hej girls! don't forget to come and visit me in Umeå!

Thank to all my family and friends, you were indispensable!

Finally, I would like to thank Nicole for her very warm welcome to Sweden and her support during the last days of the writing of my master thesis report.

Table of contents

Acknowledgements....2
Table of contents....3
Index of Appendix....4
Index of Tables and figures....5
Introduction....6
Materials and methods....9
A. Study species and study site....9
A.1) The European Beech (Fagus sylvatica)....9
A.2) Mont Ventoux....9
B. Historical research on Beech recolonisation history....10
C. Digital satellite image processing....10
C.1) Spot satellite images....10
C.2) Radiometric preprocessing....11
C.3) Image classification and analysis....11
D. Genetic structure of Beech populations....12
D.1) Sampling design....12
D.2) Genotyping....13
D.2.a. Extraction of total DNA....13
D.2.b. Microsatellite markers....13
D.2.c. DNA Amplification....13
D.2.d. Allele scoring....14
D.3) Genetic analysis....14
D.3.a. Estimation of null allele frequencies....14
D.3.b. Standard population genetics analyses on the 38 plots....14
D.3.c. Delimitation of genetic units using Geneland....15
Results....17
A. Past Beech stands dynamic inferred from historical records....17
B. Present distribution of Beech stands inferred by SPOT images analyses....18
C. Analysis of population genetic diversity and structure....19
C.1) Quality of the microsatellite data set....19
C.2) Population genetic analysis on the 38 plots....19
Genetic diversity within plots and over the studied area....19
Genetic differentiation among plots....20
C.3) Detection of genetic units using Geneland....20
Discussion....21
A. Main results and drawbacks of the different approaches used in this study....21
B. Patterns of genetic differentiation among Beech stands....22
C. Future response of Beech population to climate change (CC)....23
Glossary....25
References....27

Index of Appendixes

Appendix 1: Detailed information of the thirteen microsatellites primers used in DNA amplication by multiplex PCR.

Appendix 2 : Details on GeneLand Clustering method

Appendix 3: Pictures taken in 1888 and 1903 and showing the deforestation impact

Appendix 4 : Estimation of allele null frequencies for each locus in each experimental plots.

Appendix 5. Map of posterior probability of an individual t belong to one cluster, based on the highest probability run at Kmax=8.

Index of Tables and Figures

Figure 1A: European Beech stand and Beech shoot bearing a young cupule

Figure 1B: Eastern part of the study site.

Figure 2: Mont Ventoux and its geological location

Figure 3: SPOT satellite images used for vegetation classification analyses, with summer (top) and winter (bottom) views.

Figure 4: Map of the "AU" (units of analyses) of vegetation established by ONF: top in the Western sector (RBI), and bottom in the Eastern sector (VTX) of the North face of Mont Ventoux.

Figure 5: Sampling design for the 38 plots of the landscape genetic study.

Box 1: Geneland hidden model of spatial organisation through Voronoi tesselation

Figure 6: Map on the forested areas of the Mont Ventoux, 29 July 1876.

Figure 7: Zoom on the SPOT images showing the area under study in A: summer and B: winter.

Figure 8: Variation of the Normalized Difference Vegetation Index (NDVI) computed over the studied area from SPOT summer image (see material and method for details). Dark lines delimit the UA.

Figure 9: Results of unsupervised vegetation classification using green channel and NDVI as variables for the classification process.

Figure 10: Correlogram of pairwise population differentiation, as measured by Fst/1-Fst, against logarithm of distance.

Figure 11: Posterior density distribution of the number of clusters estimated from GeneLand analysis at Kmin= 1, Kmax= 8 after 1000 000 MCMC interations.

Figure 12**:** Spatial representation of the five significant genetic clusters inferred by Geneland

Figure 13: Identification of population discontinuities in the Western zone of the study area

Table 1: Significant index of the heterozygous disequilibrium (Fis*) at the nominal level of 5% with the Bonferroni correction.

Table 2: Allele diversity per locus over all the population.

Table 3: Estimates of population genetics summary statistics for each population.

Table 4: Number and frequency of private alleles

Table 5 Nested analysis of molecular variation for genetic variation among young and old trees within the 38 Beech plots

Introduction

According to the last report of the IPCC (Intergovernmental Panel on Climate Change), eleven of the twelve last years (1995-2006) rank among the twelve warmest years in the instrumental record of global surface temperature since 1850 (IPCC 2007). The temperature increase is widespread over the globe. The global warming is now unequivocal, and its connection with the increase in global greenhouse gas emissions since pre industrial time is now widely accepted by the scientific community. Due to human activities (huge consumption of fossil fuel such as charcoal and natural gas, changes in the land cover such as deforestation...), the global greenhouse gas emissions have increased by 70% between 1970 and 2004 (IPCC 2007). Together with the increase in the atmospheric and ocean mean temperatures, the ongoing global change (GC) is mainly characterized by a massive ice and snow melting, a sea level rise and changes in the wind and precipitation patterns. Observational evidence from all the continents and most oceans show that many natural ecosystems are impacted by regional climate change, particularly temperatures increase (Parmesan 2006). Mountain, arctic and Mediterranean ecosystems are the most affected by the recent global warming.

Among the most documented and worrisome ecological consequences of GC is the the precocity of spring events (earlier bud burst, flowering, breaking hibernation, migrating, breeding), (Roman-Amat 2007, Parmesan 2006). These phenological changes have been shown to increase mortality (risk of early or late plant frost) and to affect species' biological cycles such as the timing between the life cycles of pollinator insects and flowering plant. In addition, GC has already triggered species distribution shifts in many parts of the world (Thuiller *et al.*, 2005), with simultaneous extension of warm-adapted communities and contraction of range-restricted species (particular polar and mountaintop species). Such major shifts in species' locations alter species composition within communities and thus species interactions.

European forests are important reservoirs of biological diversity (at gene, individual, and community level) as a consequence of their complex history and environmental variation at local and regional scales. Trees are keystone organisms in European ecosystems; they directly support rich plant and animal communities that rely on them and mediate nutrient and water ecological cycles. Impacts of GC on European forests are expected to be acute, resulting in notable changes in species range, ecosystem functioning and in the interactions among species (Thuiller *et al.*, 2005). Because they are long-lived and sessile, trees can either disappear, disperse to other places or adapt in situ to the ongoing GC over a reduced number of generations.

The study of past recolonization events following glacial periods have shown the important migration abilities of many tree species (e.g. McLachlan *et al.*, 2005). Moreover, many modern trees display adaptive differentiation in relation to latitude or elevation throughout their range (Davis & Shaw 2001). It is thus likely that in the close past, the interplay between selection and gene flow contributed to rapid adaptation of environmental sensitivities of populations throughout species range in conjunction with post-glacial migrations (Davis & Shaw 2001). However, the unprecedented rates of climate changes anticipated to occur in the future challenges this process by imposing stronger selection and by distancing populations from environments to which they are adapted According to several bioclimatics models, tree species would have to shift their ranges over several hundred of kilometres Northward by 2100, without adaptation, and without altitudinal compensation (Badeau *et al.*, 2004). However, a recent study based on molecular marker on the postglacial recolonization of two North-American trees, red maple (*Acer rubrum*) and American Beech (*Fagus grandofolia*) across North America, suggests that many temperate tree species expanded at rates < 100 m/yr during the early Holocene (McLachlan *et al.*, 2005). The huge

discrepancy between the tree migration rates estimated from post-glacial recolonisation studies and the migration rate that would be required to track 21st-century warming has lead some authors to propose the use of non-native provenances or species for restoration plantations (O'Brien *et al.*, 2007). However, such recommendations could be unnecessary if other evolutionary processes (adaptation, long distance gene flow) could mitigate GC consequences.

Different approaches are available to gauge the respective roles of migration and adaptation in the response of tree populations to GC (Reuch and Wood 2007): experimental selection under future climatic conditions, temporal *in situ* studies tracking evolutionary change through time, or spatial studies of populations differentiation across spatial environmental gradients to infer retrospectively past evolution. This last approach is clearly the most feasible in many species, especially in long-lived species like trees. In this context, URFM laboratory has initiated a long-term multidisciplinary research program (ecophysiology, ecology and genetic) aiming at understanding the European Beech (*Fagus sylvatica*) and silver Fir (*Abies alba*) functional and evolutionary response to an attitudinal gradient on Mont Ventoux, South Eastern France. By 1850, the site was almost entirely deforested due to over-grazing by sheep and goat. Mixed *A. alba–F. sylvatica* forests were reduced to small forest islands in the most inaccessible parts of the mountain. The decrease of grazing combined with the reforestation program launched in the late 19th century (using mostly pine and spruce) made it possible for *A. alba* and *F. sylvatica* to gradually recolonize the planted stands. In few generations, Beech stands for instance spread over an elevation belt ranging from 900 m to 1500 m, and some evidence of adaptive differentiation for budburst phenology could already be observed (Davi comm. pers.).

The objective of our study is to document the recent colonization history of European Beech on Mont Ventoux, and its impact on the genetic structure on present populations by combining historical researches and a landscape genetic approach. This is a first step in the retrospective study of the respective roles of migration and adaptation in the response of Beech populations to a climatic gradient. The originality of this study is to focus on a local, landscape geographical scale, and on a recolonization process that date back to the 1960's, that is to maximum 3 generations ago. Indeed, these spatial and temporal scales are particularly relevant to address ecological processes related to population dynamics. Moreover predictions and management guidelines are required at the scale of a few generations in a regional context.

In this study, we first conducted a series of historical researches to identify the distribution and the importance of relict Beech populations during the over-logging period, and to detect possible restoration plantation events of Beech. Then we used SPOT satellite photographs combined with vegetation classification analysis to map the present distribution of Beech. Finally, we used a landscape genetic approach to document the recolonization history of Beech based on patterns of genetic differentiation among 38 populations. All these investigations were conducted on the North face of Mont Ventoux.

Landscape genetics (Manel *et al.*, 2003) has emerged as a new research are that integrates landscape ecology, spatial statistics and population genetics. This discipline provides information about the interaction between landscape features, genetic discontinuity, gene flow and genetic population structure. The two keys steps are the detection of genetic discontinuities (geographic zone of sharp genetic change), and the correlation of these discontinuity with landscape and environmental features (biotic, climatic, edaphic and other conditions that compromise the immediate habitat of an organism).

The genetic structure of the Beech stands was analysed using software, the Bayesian clustering approach of Guillot *et al* (2005b). This approach allows inference of the location of genetic discontinuities from individual georeferenced multilocus genotypes, without a priori

knowledge on population units and limits (Guillot *et al.*, 2005a). The objectives of these clustering analysis were: 1) to retrieve the positions of the different refugees of the Beech stands during the period of over-logging and over-grazing and 2) to estimate the intensity of drift due to the reduction in population size during this period and 3) to infer the re-colonisation routes followed from these refugees. In particular, we want to assess whether diversities from different refugees mixed in some particular crossways (highly structured case) or whether long distance dispersal (LDD) led to a much more even mixing at the scale of the forest (structured refugees and uniformly distributed diversity elsewhere).

Materials and methods

A. Study species and study site

A.1) The European Beech (*Fagus sylvatica*)

The European Beech is a deciduous tree belonging to the family Fagaceae. It has a typical lifespan of 150 to 200 years, exceptionally 300 years. Trees typically grow up to a height of 20-40 m but they can reach 40 m height and 1.5 m diameter (Tessier du Cros *et al.,* 1981) (Figure 1). As all the other members of the Fagaceae family, Beech tree is monoecious: each individual bears male and female inflorescence. The catkins appear shortly after the leaves in spring (from the middle of April to the beginning of may). The European Beech starts to flower when it is between 60-80 years old and sometimes until 200 years. The fecundation give birth to two small, sharply nuts inside a husk known as cupule (Figure 1). Beechnuts are mature in September-October, then the cupules release the seeds. The flowers are pollinated by wind (amenophily) and the most common mode of reproduction is allogamic. Root sprouting occur quite often. Beech nuts are relatively heavy and the median seed dispersal distance is in the order of 10 m (Sagnard *et al.,* 2007; Bontemps 2008).

European Beech trees have great ecological amplitude in particular a large edaphic plasticity. However, the species is seldom found on strongly acid hydromorphic soils, and where soil is too dry (Tessier du Cros *et al.,* 1981). At early stages, Beech seedlings need to grow in a shady place: they dread any excessive sunlight (shade tolerant species). At adult stage, Beech trees are usually dominant or codominant , and it is a powerful competitor against other tree species. The range of Beech forests cover wide areas in Europe: from the zonal vegetation from Southern Sweden to the Alps and from central France to Western Poland and the Carpathians (Rob 1992). Beech tree is today the most important natural tree species in central Europe. In the Northern part of France it is mainly located in plains and rather found in mountain region in the South. European Beech has a great economic importance. It is an excellent firewood and it is one of the best hardwood in the manufacture of the paper pulp and of numerous objects and implements.

A.2) Mont Ventoux

The study was carried out in Mont Ventoux, a French mountain located in the Provence region. It is the highest peak of the Vaucluse department. Culminating at 1912 meters, it rises from the South-western end of the Southern Prealps, extending some 24 km from East to West and 15 km from North to South (Merle & Guende 1978) (Figure 2). Most of the mountain is made up of Cretaceous limestone. Mont Ventoux is subject to a dominant Mediterranean regime but the elevation range results in a great diversity of climate: from Mediterranean it evolves toward a temperate climate at mid-slope and to a mountain continental climate at the top. The wind is particularly strong and the precipitations are especially abundant during spring and winter time.

The mountain shows very rich flora and fauna, an arctic species is even found on the scree slopes at the summit: the purple saxifrage, *Saxifrage oppositifolia*. Many species are rare and protected like the Orsini's viper, *vipera ursini*. Some of them are endemic: *Leucojum fabrei* is a very narrow endemic species, discover by Jean-Henri Fabre in 1880, only found in three locations on the Southern slope. In 1990 the UNESCO recognized the biological distinctiveness of the massif, creating as a part of the program Man And Biosphere the biosphere reserve of the Mont Ventoux, a protected area of 89 408 ha. The core area has been defined through six biotope protection ordered

zones.

The study sites are located on the Northern slope of Mont Ventoux, at an elevation range from 1000 to 1500 m. There, the climate is cold and wet, it is at this level of the massif that the precipitations are the most abundant (1293 mm/yr in average at Mont Serein meteorological station). The North face is very steep, especially on its oriental sector, which is broken by large steep rocks like at the Serre Gros location where the elevation drops by 1,500 m on only three kilometers. Avalanches occur quite often in this eastern sector, resulting in a mixing of forest stands and rock slides. Numerous Beech tree relics can be found, saved from deforestation of the last centuries. Most often, Beech occurs in mixed stands, with silver Firs and mountain Pine. A part of the site is inside a biotope protection ordered zone of 98 ha and two ZNIEFF (natural zone of ecological faunistical and floristical interest) preserve the area. On the western sectors, the study area mostly correspond to Bedouin communal forest, and it is submitted to classical forest management operations.

B. Historical research on Beech recolonisation history

Forest of Mont Ventoux are highly heterogeneous and the official forest maps are not accurate enough to give all the potential sources of beeches. In order to try to identify all the Beech stands over the North face of Mont Ventoux, different archives were investigated. Old printed materials (books, maps, newspapers, periodicals, photographies from the RTM...), forest registers and local administration records were consulted at the Vaucluse and the Drome local historical archives and at the ONF branch in Avignon and Valence. Part of the historical researches has been done at the Requien Museum (natural history museum in Avignon) with the curator Marie Hélène Grabié to have access to the scientific library. I also exchanged some informations with some researchers, historians and writers working on the botanical aspects of the Mont Ventoux.

C. Digital satellite image processing

In order to identify all the present Beech stands over the North face of Mont Ventoux, digital image analyses were run on two multi spectral Spot 5 satellite images using the Geographic Information System (GIS) software GRASS (Figure 3). The first satellite image was taken in winter and the second one in summer (inside and outside vegetation period) in order to clearly identify the Beech stands from the Fir and Pine (conifer) whose reflectance signals might be very similar to those of Beech (deciduous).

The results of our image processing were confronted to a vegetation map established by ONF (the French National Forest Office) and called Analysis Units Maps (AU map). To establish this AU map, Mont Ventoux has been divided in 2201 homogeneous plots (units) established from aerial photographs and of their crossing with the land register. Then, random systematic field inventories have been done in order to identify the tree species in each plots. Thus, the map of stands where Beech is the main species (or the second main) are available under GRASS (see Figure 4).

C.1) Spot satellite images

The spatial resolution under the multi-spectral mode is 10 m for both of the satellite images (60 km x 60 km). The two images were acquired on the 22 December 2006 at 10:43 and on the 23 Mai 2007 at 10:18 with an angle of incidence of 3.6° and 28.1° respectively (Figure 3). The Spot satellites images were delivered under the Lambert 93 coordinate reference system. The 4 spectral bands are the infra-red IR (Red channel: 780-890 nm), red (Green channel: 610-680), green (Blue

channel: 500-590 nm) and the middle infra-red MIR (Alpha channel: 1.580-1.750 nm).

C.2) Radiometric preprocessing

The slope and its aspect, the atmospheric conditions, the sun location, could cause variations in the brightness reflectance, which can lead to wrong classification results (Neteler & Mitosova 2008). The slope effects can be extremely high on the very steep North face of Mont Ventoux. To overcome this problem, an absolute image calibration (illumination correction) was performed under GRASS to produce a normalized spectral reflectance* ρ, that we call thereafter the corrected Spot image (Meygret 2006).

$$\rho_k = \frac{X_k}{A_k G_{mk}} \frac{\pi}{E_k co(\theta) \times u}$$

X_k is the raw spot satellite image and k the spectral band. A_k is the absolute calibration coefficient and G_{mk} is the analog gain depending on the gain number m. The value of the product $A_K.G_{mk}$ was given with the ancillary data of Spot image as the physical gain. E_k is the normalized solar irradiance* calculated from an elevation model. It has to be corrected by u, the earth-sun distance (depending on the orbit eccentricity and the mean rotation angle). Θ, the solar zenith angle* was calculated from sun elevation EL (Θ =90- EL). All the calculations were done using *r.mapcalculator* and *r.sun* modules under GRASS.

C.3) Image classification and analysis

Common multispectral classification algorithms use the multichannel images as explicative variables and the class composition as output variables. In our case, the resulting classes correspond to the main tree species over the forested areas of the North face. The first step in analysis of image satellite data is to identify which spectral bands are the most informative based on the pixel values; this is done by looking at each channel histogram showing the frequencies of grey levels in a given spectral bands. When classifying multispectral satellite data, the image data set is analysed at pixel level, with the values of all informative bands being taken into account for each pixel. Within the feature space, the classification algorithm tries to separate similar spectral signatures which vary depending on the observed object types as soil, vegetation, roads....Similar spectral signatures will be assigned to the same class. All classes are finally stored in a thematic map where each class describes the dominating land use type.

The objective was first to use an unsupervised classification on our satellite images. This is a fully automated method based on image statistics (such as spectral band, reflectance index...). This method consists of two main steps. First, image channels of interest are collected into an image group using *i.group* module. A clustering algorithm groups pixel values with similar statistical properties depending on minimum cluster size and the number of clusters (*i.cluster* module). The pixel clusters consist in image categories that are expected to be related to our land cover types. Second, the classification is achieved using the maximum likelihood algorithm (*i.malix* module): all the pixels in the satellite image are assigned to the spectral signatures (classes) derived by the previous clustering process. Then, correspondences between classes of spectral signature and vegetation type (e.g. species) can be searched. For these unsupervised classification, we first used each band separately then all the bands together. Finally we tested classification combining a Normalized Difference Vegetation Index (NDVI) and the unsupervised classification previously done. NVID was computed as:

$$NDVI = \frac{NIR - R}{NIR + R}$$

with R the reflectance and NIR the Near Infra-Red absorbance

Another common method for the classification of remote sensing data is the supervised classification. In this case, training areas covering known land use have to be digitized. Image statistics are automatically derived from these training areas and used for the final classification. As training areas, we intended to use the UA. However, a short test showed us that we could not use supervised classification, because in one UA the Beech stands are often mixed with Fir and Pine and therefore the land cover is not locally homogeneous. We could have used training area smaller than the UA, based for instance on aerial photographs available on some site. But in this case, the low resolution of SPOT satellite image would have been limiting.

D. Genetic structure of Beech populations

The objective of this part of our study was to characterize the level and the organization of the neutral genetic diversity at the level the North face of Mont Ventoux on the basis of 38 experimental plots and a total number of individuals of 1599 (Figure 5).

D.1) Sampling design

The sampling design of the present Beech stands was defined to cover the relic areas of the oriental sectors (refuges areas during the over-logging of the forest) and the main recolonization axes over the North face (Figure 5). It consists in two groups of plots, respectively identified as RBI-x and VTX-x (x being the number of the plot).First, 24 experimental plots (noted RBI-x) were chosen inside the biotope protection ordered zone, from a network of plots previously installed by ONF . Close to the Mont Serein station, three additional new plots were set up in order to make a link between the East and the West sector of the North face (plots referred to as RBI-ext-X). On each of these 27 plots, 30 adult Beeches were selected within the limit of the number of observed individuals and according to their diameter: 15 “old individuals” (the biggest of the stand) and 15 “young individuals” (small trees with a diameter higher than 7 cm) were collected. In case of resprouts , only the larger stems was sampled. The average number of individuals per experimental plots ranged between 7 and 30, with a mean value of 28. Each tree was georeferenced. Diameters were measured at breast height. Some plant material (leaves) was collected on each tree and stored at INRA Avignon laboratory at – 20°C. The ages of one tree of each cohort in each plot were estimated by extracting incremental cores and counting tree rings. All the collected individuals were Sampling was done between the 11th May and the 14th June 2009.

Second, to cover the West sector of the North face of Mont Ventoux, 916 individuals were sampled in14 experimental plots (called VTX-X) from a network of plots previously set up by Ph. Dreyfus in the frame of ECOFOR project. In 11 of these 14 plots, 30 individuals were selected following the same rules as for the RBI-plots. In 3 of the 14 plots corresponding to Intensive Study Plots (VTX_N2, VTX_384 and VTX_257_2), individuals were exhaustively sampled, with 98, 293 and 176 individuals respectively. All these trees were previously mapped, their diameter was measured and fresh material already collected and stored at – 20°C at INRA Avignon laboratory (sampled in 2008, projects Ecoger and Evoltree). The average number of individuals per

experimental plots ranged between 7 and 293, with 65,5 trees on average.

D.2) Genotyping

D.2.a. Extraction of total DNA

Approximately 50 mg wet weight (optimal amount of starting material) of frozen leaves was placed into collection microtubes racks* of 96 wells each. The purification of total DNA from frozen the tissues was done following to the DNeasy® 96 plant procedure from the firm QIAGEN. DNeasy® Plants kits provide a fast and easy way to purify DNA from plant tissue and provide pure total DNA for reliable Polymerase Reaction Chain, PCR*. The principle of the DNA extraction relies on the binding of the DNA on a silica column which has the faculty to link selectively the DNA molecule in the presence of a high salt content and at optimal pH (7.5). After elution, extracted DNA solutions were stored in microtubes at – 20°C. The concentration and the purity of DNA were estimated by measuring the absorbance at 260 nm and 280 nm in a spectrophotometer (sample dilution for PCR should be adjusted accordingly), and double-checked using pulse-field gel electrophoresis (PFGE*) on agarose gel.

D.2.b. Microsatellite markers

Microsatellites or SSRs (Single Sequence Repeats) are defined as loci (regions within DNA sequences) where shorts sequences of DNA of 1-6 base pairs are repeated 10-100 times in tandem arrays. The length of sequences used most often di-, tri-, tetra- nucleotides. Microsatellites are considered as neutral markers* and co-dominant*. They have a high mutation rates compared to other regions of DNA, 10^{-6} to 10^{-2} mutation per genome and per generation (De Vienne 1998). The number of time the sequence is repeated often varies between individuals, even within population, and this polymorphism in the number of repeats can be easily revealed using Polymerase Chain Reaction (PCR) and high definition electrophoresis (see below). Microsatellite thus usually display a high degree of polymorphism, allowing to "fingerprint" easily individuals. They have become in the last decade very powerful molecular markers in population genetic. The set of thirteen microsatellite markers used in this study was adapted from markers developed for *Fagus sylvatica* : Fs1_15, Fs3_04 (Pastorelli et al., 2003) or for other close species such as *Fagus crenata, Fagus japonica* : sfc0007_2, sfc0018, sfc0036, sfc1063, sfc1143 and mfc7 (Asuka et al., 2004; Samper 2007) and *Quercus robur: cfs_05, cfs_06, cfs_25, cfs_29, cfs_31(*unpublished microsatellite sequences from the Italian laboratory CNR, Consiglio Nazionale della Ricerche of Florence*)*. The information on marker sequences and amplification condition is given in Appendix 3.

D.2.c. DNA Amplification

The sequences flanking the microsatellite region are used to design specific, around 20 bp primers* for microsatellite loci amplification by PCR. PCR amplifications were performed with the QIAGEN® multiplex PCR kit and using thermal cyclers Eppendorf® (Applied Biosystems). Multiplex PCR is a powerful technique that enables the amplification of two or more primers in parallel in a single sample. This technique requires that the different primer pairs do not make hybrids and that the annealing temperature is the same for all the primers of a multiplex. Moreover, it is necessary to verify that the primer sequences are not complementary. Two multiplexes were realized, 7 and 5 markers respectively. The binding of fluorochromes* on the primers (Fam- green, Tamra- yellow and Hex- bleu) has been chosen to avoid that the overlapping allele sequences appear with the same colour on the allele profile (A. Roig, INRA Avignon).

According to the QIAGEN® protocol a PCR mix was realized for each sample with the

extracted DNA, a QIAGEN® reaction mix and the primer mix. PCR programs for all loci consisted of an initial activation step (activation of the HotstarTaq® DNA polymerase*) at 94°C for 15 min, 30 cylces of 94°C DNA denaturation for 30s, 90s of primer annealing at 60°C and 60s of primer extension at 72°C. A final step at 72°C for 30 min was used to a complete primer extension.

D.2.d. Allele scoring

DNA fragment separation was done by electrophoresis on a MegaBACE™ 1000 sequencer at the molecular biology laboratory of INRA Avignon. The genotypes were scored with the MegaBACE Genetic Profiler ©Amersham Biosciences 2003 software version 2.2 against an internal size standard* (ET400 DNA size markers). Automatic allele assignation was checked and revised manually twice to ensure consistency of genotyping.

D.3) Genetic analysis

D.3.a. Estimation of null allele frequencies

One of the main limitations of the use of microsatellite as molecular markers in population genetic is the possible occurrence of "null alleles" , where microsatellites fail to amplify during PCR reactions. Null alleles are usually considered to result from some mutations in the DNA sequence where the hybridization of the primer occurs. Consequently, the primer fail to hybridize : there is no microsatellite amplification. Heterozygous individuals AN (N being the null allele and therefore not detected) are scored as a homozygous AA and homozygous individuals NN appear as missing data in the population genotype profile. Null alleles presence lead to an overestimation of the homozygous at the locus level and therefore the estimation of the population genetic diversity can be biased (Kalinowski & Taper 2006).

Thus, the frequency of null alleles was estimated for each locus with the method of Kalinowski *et al.* (2006) implemented in ML-NULL software . The advantage of this method is that it accounts both for heterozygous deficiency and for missing data to estimate null allele frequencies following a maximum-likelihood procedure. The percentages of individuals with missing genotypes as well as the number of population affected were computed. To remove missing data due to amplification failure or human error, we performed a third multiplex PCR on the individuals having no amplification at some locus from the multiplex 1 and 2.

D.3.b. *Standard population genetics analyses on the 38 plots*

- *Genetic diversity within population*

The expected heterozygosity assuming panmixy hypothesis (*He*) and the observed heterozygosity (*Ho*) were estimated for each population using GeneALEx 6 software (Peakall and Smouse 2006). The fixation index *Fis* of Wright (1969), describing the heterozygous deficiency in a population, was also estimated and its significance level was assessed using the Bonferroni procedure* implemented in FSTAT software (Goudet 1995). For each population, the number of alleles (*Na)* and the number of private alleles (*Np*) were computed. Finally, linkage disequilibrium, the non random associations of alleles at different loci, was measured for each pair of loci using Genetix software (Belkhir et al.)

$$He = 1 - \Sigma (pi)^2$$ Genetic diversity index of Nei (1973)

pi being the the frequency of the allele *I*

$$Fis = 1 - (Ho/He)$$

- *Genetic structure among populations*

The average level of differentiation among populations was measured by Wright Fst, and estimated using analysis of molecular variance (AMOVA) as implemented in GeneALEx 6 software (Peakall and Smouse 2006). Pairwise Fst values were also computed among every population pairs using FSTAT.

Under isolation by distance (IBD) in a set of populations, genetic differentiation among populations increases with their geographical distance. According to the theoretical works of Rousset (1997), in an IBD model in two dimensions, the relation between *Fst* /1-*Fst* and the logarithm of the distance between populations is approximately linear: $y = ax + b$. The slope coefficient *a* provide an estimator of the product $4\pi Ds^2$ where *D* is the effective population density and s^2 the axial variance of gene dispersal distance.

The hypothesis of IBD was tested by plotting pairwise population genetic differentiation (measured as *Fst* / 1-*Fst*) against the logarithm of the distance using the SPAGeDI 1.2 software (Hardy & Vekemans 2002). The mean differentiation values were calculated over 8 distances classes (100 m). The 95% confidence interval of the autocorrelogram was calculated via a Mantel test* through 5000 permutations of the individual genotypes on their geographical coordinates.

Finally, temporal differentiation among young and old trees within each plot was investigated using hierarchical AMOVA design.

D.3.c. Delimitation of genetic units using Geneland

The Bayesian approach of Guillot et al (2005b) uses individual georeferenced multilocus genotypes to partition individuals into K panmictic populations without a priori knowledge on populations limits and units. Particularly interesting features of GeneLand model are its ability to: (i) deal with an unknown number of population simultaneously with the estimation of the other parameters (location of individuals, number and location of tiles…) (ii) estimate the number of populations in the studies areas, (iii) assign individuals to their population of origin and (iv) detect migrants in populations (Guillot *et al.*, 2005a).

The model

The individual is the operational unit of study. The whole set of geo-referenced individuals is viewed as belonging to one of several populations at Hardy-Weinberg equilibrium (HWE), and departure from HWE leads to the splitting of the population into K sub-populations. Individuals within populations are assumed to be randomly located and linkage equilibrium is assumed between loci. Allele frequencies f_{kl} at each locus *l* and in each population *k* are drawn from Dirichlet distributions which are assumed to be either independent or non-independent among populations. In the last case, the model specifies the relation between the frequencies in the present populations to those of the ancestral population through a drift factor *d*. The populations are assumed to be spatially organized through the so-called colored Poisson-Voronoi tessellation. This model consider that the K populations occupy some sub-domains $\Delta_1, \ldots, \Delta_K$, of the total spatial domain Δ. The spatial domain Δ can thus be approximated by an union of convex polygons, denoted tiles. As one color *c* is attributed to all the tiles of the same population, this approach is denoted “coloured Voronoi tessellation”. The amount of spatial dependence depends on how the spatial sub-domains Δ_k are fragmented in smaller polygons. Briefly, low values of *m* (the total number of tiles, an random

parameter to estimate) correspond to weakly fragmented partitions of Δ and thus to strong dependence of the hidden spatial organization of populations, whereas large values of *m* correspond to high fragmentation and weak spatial dependence (see Box 1).

To summary, the different unknown parameters to estimate from the data (that is the set of georeferenced genotypes) are the following:

K : the number of populations

m: the number of sub-domains per population,

c: the colors of the domains

Δ: the spatial organisation of the domain

f_{kl}: the allelic frequencies at each locus within each populations

s: the true coordinates of the individuals, because the observed coordinates are assumed to be affected by measurement errors.

Inference is performed via simulation of the posterior distribution of parameters by Markov Chain Monte-Carlo techniques (MCMC). The MCMC algorithm is detailed in the Appendix 2**.**

Parameter values used for Geneland procedure

The MCMC was ran ten times (to check the consistency of the results), with the following parameters: 1,000 000 MCMC iterations of which one every 100 was saved, no uncertainty attached to the spatial coordinates, maximum rate of Poisson point process fixed to 100, maximum number of nuclei in the Poisson-Voronoi tessellation fixed to 300, a minimum of K fixed to 1 and a maximum K fixed to 8. We used the Dirichlet model as a model for allelic frequencies. The mean logarithm of posterior probability for each of the 10 runs was calculated and the run with the highest value was selected. This procedure was run first on all the 1599 individuals from the 38 plots, and then on 382 relic-trees from the 38 plots. Only results for the run with 1599 individuals were reported because the second series of run on the 382 relic-trees provided no results.

In order to map the geographical limits of each cluster, a post process was done on the selected run. The posterior probability membership for each pixel of the spatial domain was computed for each of these five runs with 350 pixels along the X axis and 150 along the Y axis (the Mont Ventoux being along an East-West axis). The number of clusters was determined as well as the potential migrants in the inferred populations. Then, using GRASS software, the inferred clusters were put in relation with the digital elevation model to identify if topography can explain physical barriers or corridors leading to the colonization routes and the pattern of the different populations.

Figure 1A: European beech stand and beech shoot bearing a young cupule





Figure 1B: Eastern part of the study site.



Figure 2 : Geographical location and elevation range of Mont Ventoux.

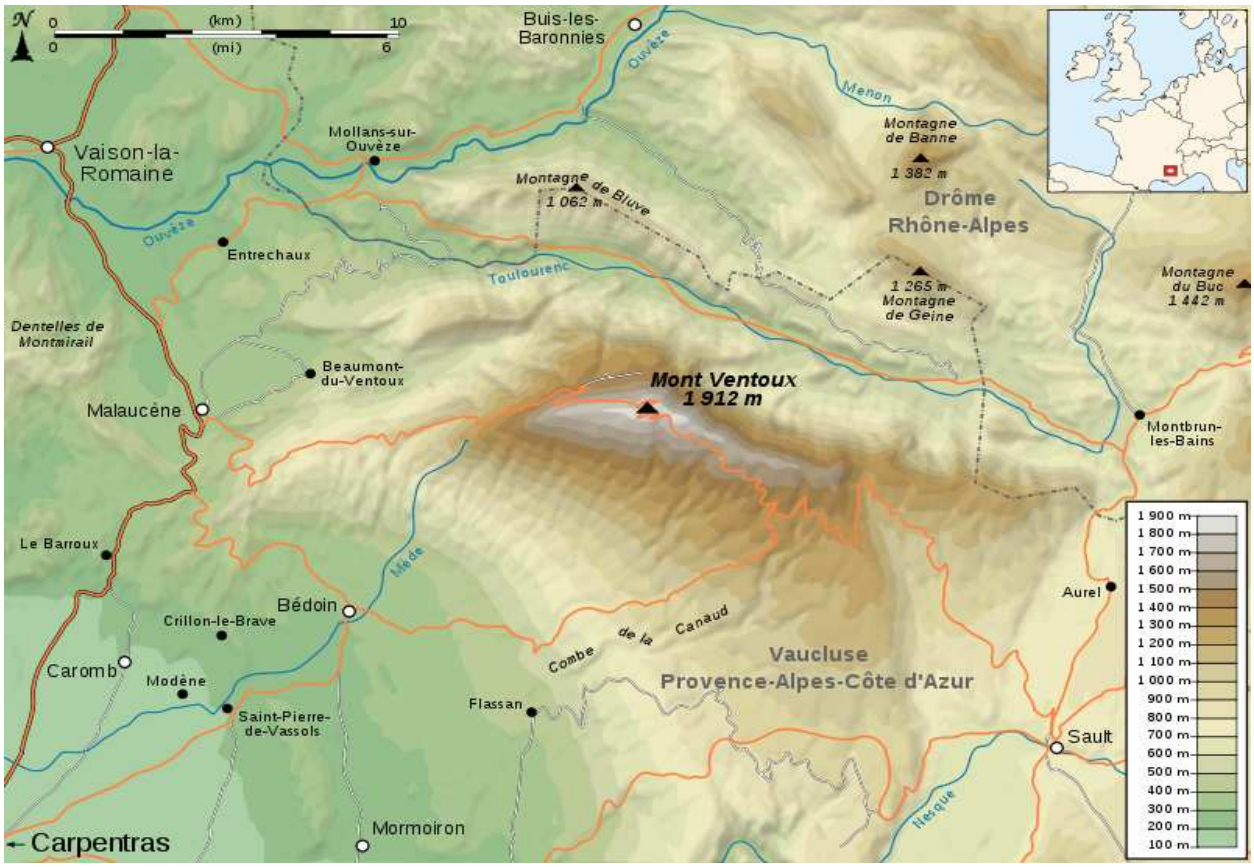
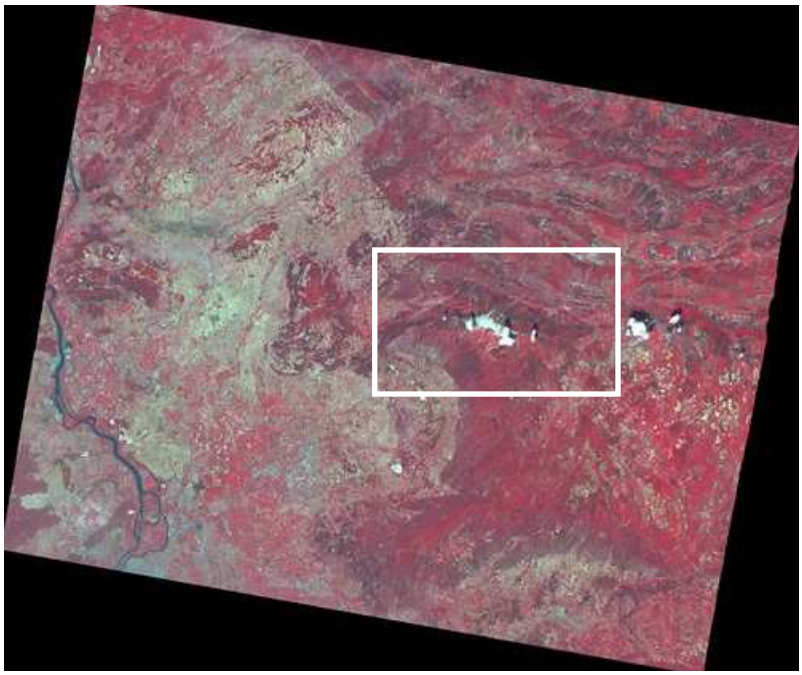


Figure 3: SPOT satellite images used for vegetation classification analyses, with summer (A) and winter (B) views.

« Summer » Spot satellite image acquired on the 23 Mai 2007 with an angle of incidence of 28.1°. Mont Ventoux is located by the white rectangle. Some thick clouds covered the peak and a small part of the North face. « Winter » Spot satellite image acquired on the 22 December 2006 with an angle of incidence of 3.6°. Mont Ventoux is located by the black circle. In addition to some thick clouds, the whole North face was in the shade.

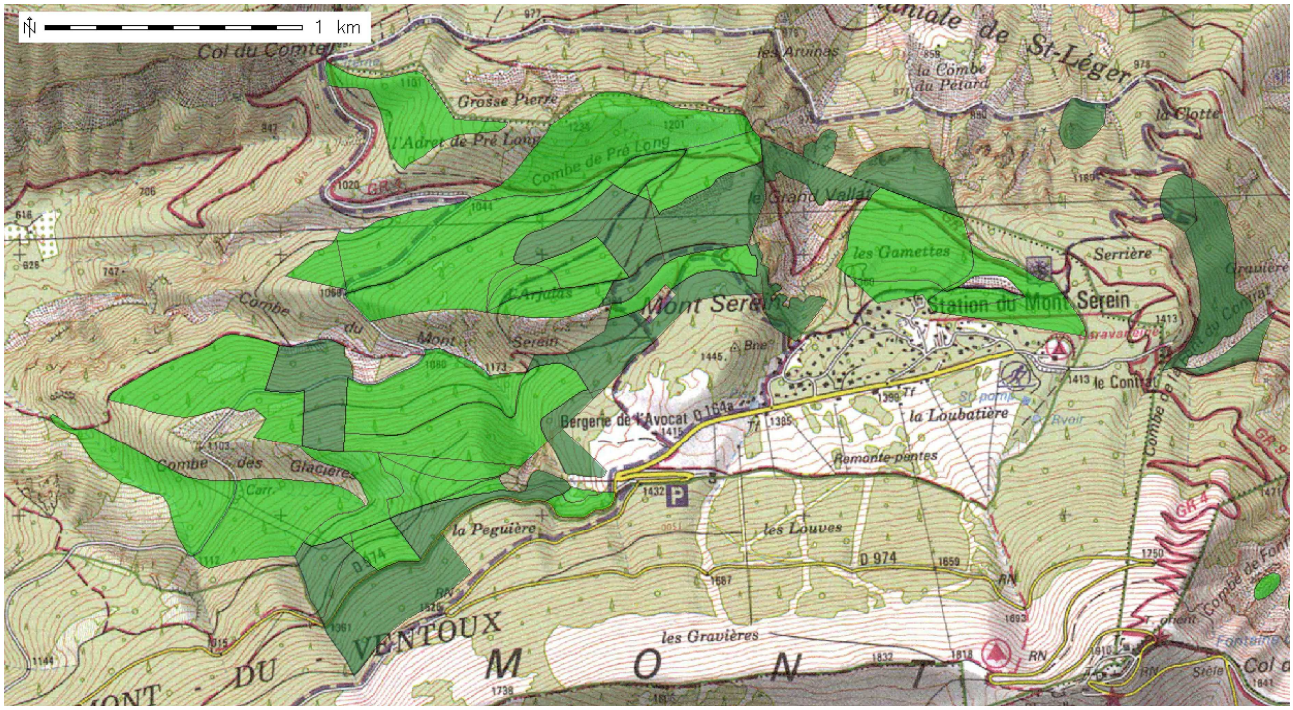


A



B

Figure 4: Map of the "AU" (units of analyses) of vegetation established by ONF: Top: in the Western sector (RBI), and Bottom: in the Eastern sector (VTX) of the North face of Mont Ventoux. Dark green correspond to areas where Beech is the main species and light green to areas where Beech is the second main species



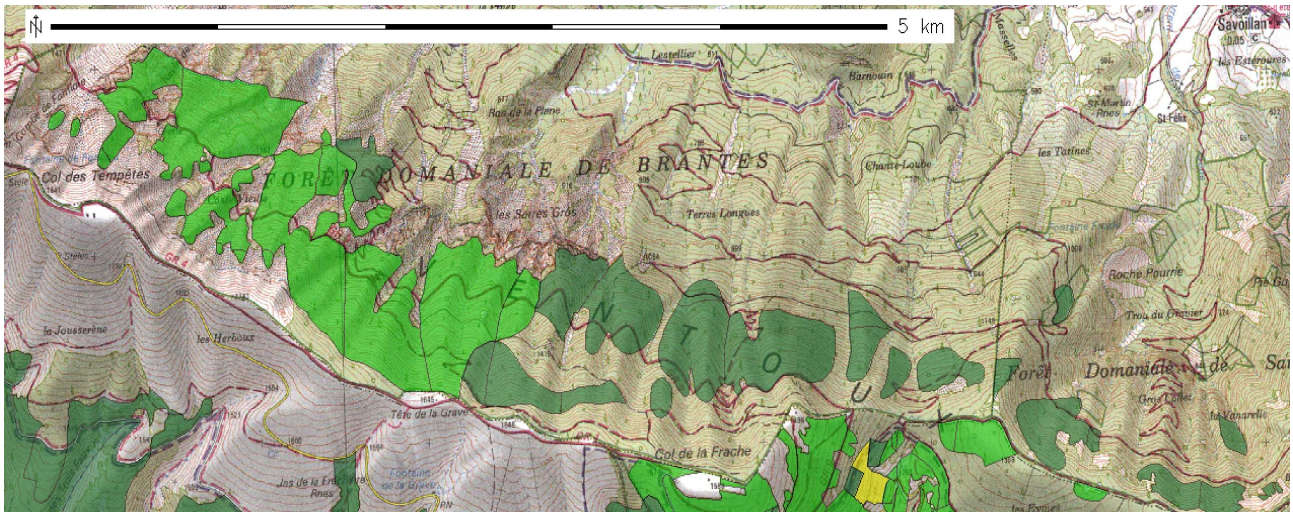
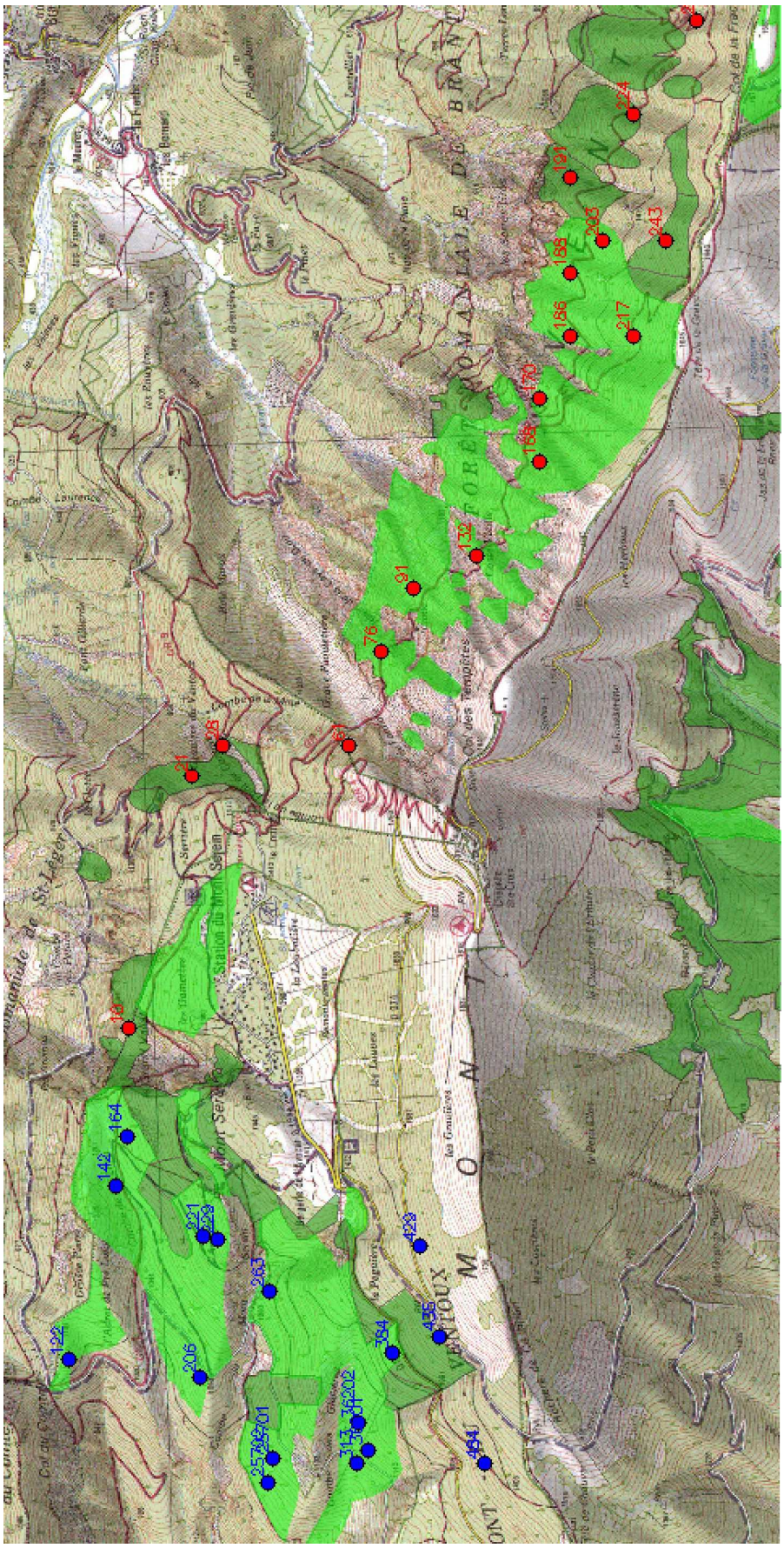


Figure 5: Sampling design for the 38 plots of the landscape genetic study.

Red dots correspond to plots within the biotope protection ordered zone (RBI)

Blue dots correspond the western zone of study in Bedouin forest (VTX)

Note: this map needs to be re-edited, as some plots were finally not included, and other were added.

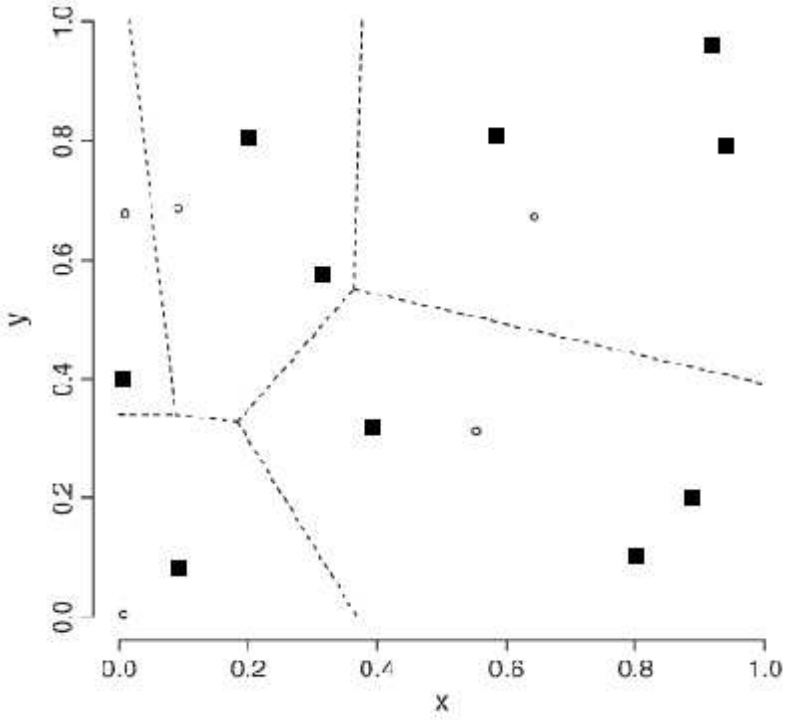

MONT VENTOUX
Mont Serein
Station du Mont Serein
Chalet Reynard
Col des Tempêtes
Col de la Madeleine
Col de la Fromagère
Col du Comte
Col de Macuègne
Col de Veyrun
Col de la Croix
Col de la Chaudière
Col de la Fontaine
Col de la Fromagère
Col de la Madeleine
Col de la Chaudière
Col de la Croix

Box 1: Geneland hidden model of spatial organisation through Voronoi tesselation

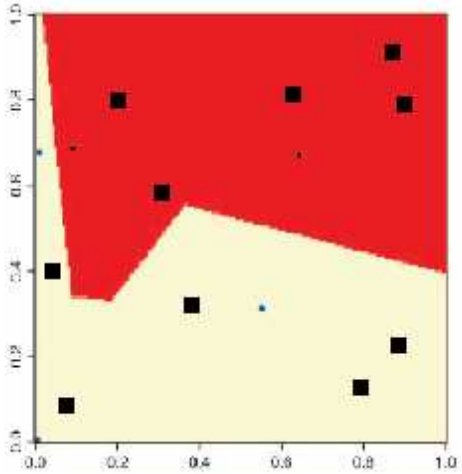
Each population *k* of the set of K populations is supposed to occupy a subdomain Δ_k for which the borders are unknown. In order to approximate the true pattern of populations spread across space, the spatial domain Δ is divided in polygons (also denoted tiles) induced by a set of random points. The vector *u* of the nucleus of all the *m* tiles (m being itself random) is assumed to follow an homogenous Poisson point process of parameter λ. Each random point belongs to a polygon *p* consisting of all the points closer to *p* than to any other polygon (top, Voronoi tesselation). The geographical coordinates of each individual will be represented in this space. Each polygon belongs to one of the *k* populations depending on the individuals inside it. Each population is coded by a colour for the graphical representation and its territory correspond to the union of the polygon of the same colour (bottom). The probability of two individuals to belong to the same population depend on their localization and therefore on the geographical distance between them. The colour and the localization of the Voronoi polygons has to be estimated from the data set.

Random tesselation of a unit square into two spatial domain through a colored Voronoi tiling (Figures modified from Guillot et al., 2005a). The black square pattern is the spatial spread of individuals and the tiny black dots are the location of the centre of the polygons. In this example, the number of polygons m=5, the number of individuals is n=10 and the number of population is K=2

Top: Realization of a Poisson point process with Vornoi tesselation with the random points spread across space.



Bottom: Partition obtained after the assignation of individuals into populations and the union of tiles belonging to the same population (coded as two colours).



Results

A. Past Beech stands dynamic inferred from historical records

Mont Ventoux history is intimately related to human presence. Beech stands showed a complex dynamics of regression-recolonisation. Indeed, through time and human activities, these forested landscapes have underwent various extension and regression phases.

After the last glacial time, 12 000 years ago, the top of Mont Ventoux was colonized by a mixed population of European silver fir, *Abies alba*, Italian maple, *Acer opalus*, Scots pine, *Pinus sylvestr*is and Mountain pine, *Pinus uncinata* (Thinon 2007). A forest of deciduous oaks was present on the slopes of the mountain depending on the topographic and exposure conditions. Thus, at this time, European Beech was absent on the mountain. Indeed during the last glaciation, the populations of the common Beech were had taken refuge in the Southern Europe. And it is just 11 000 years ago that he common Beech begun to colonize the continent from these meridional refuges (Magri *et al.*, 2006).

Around 5000 years ago, the rich vegetation made of deciduous oaks was confined to lower altitude by the arrival of European Beech and the spread of one other mountain tree species, the European silver fir (Barbero & Quezel 1987). At this time, in the early neolithic age, human pressure started to be significant on the forested areas of the massif with first clearing mainly located down the valley. Then, the clearing reached all the part of the mountain with a widespread use of the stubble-burning culture at high altitude. Beech stands were progressively overlogged to serve the huge demand for of wood from the shipbuilders of the naval port of Toulon from the XIIth century, but also because of the rise of metallurgy that required the charcoal burner's clearing and of the use of Beech as fire wood. No information were found on the exact extent of Beech stands on the massif during the firsts centuries and no paleoecological studies were performed on the North face of the Mont Ventoux.

A French botanist, Michel Darluc, with its climbing in 1778 brought a valuable evidence of the significant regression of the Beech stands on the both faces: they would be just present in the ravines on the Southern part of the massif whereas the stand on the North face looked more preserved (Darluc 1782). Comparing to the amount of botanical records found on the Southern part of the Mont Ventoux, few archives were present on the North face to cover the XVIIIth century. Indeed, there were relatively few botanical expeditions and climbings done on the North face of the massif due to its accessibility (very steep slopes). In early XIX century, it seems that the whole Beech stands was reduced to the more inaccessible on the North part of the massif. Nevertheless, according to Martins (1838), there were still some stunted Beeches from 900 m to 1550m on the North face (from 1100 to 1660m on the South side of the massif) and the last old Beech stands were found at the level of the Mont Serein. Up to 1500m, Beech stands seemed to be very scarce and dominated by the mountain pine. Beech spatial dynamic was confirmed in part by the ONF records (Série B of the local forest archives of Brantes, St Léger, Beaumont du Ventoux, 1821-1831).

Several documents found at the local historical archives (reports on the forest reclamation of the Mont Ventoux) as well as a map established in 1876 (Figure 6) provide some informations on the extent and the composition of the forests over the North face at the end of the XIX century. In the South-East part of the massif, the local forest of Brantes was made up of 281 ha of a slight continuous populations of Beech and Fir aged from 1 to 200 years old and located under the some scarce mountain pines. The Beech stands of the local forest of St Léger, situated underneath of the Mont Serein, appeared as a fragmented population of 35 ha of Beech aged from 1 to 100 years old

mixed up some mountain pine. On the West part of the massif, 250 ha of Beech, mountain Pine and Fir were found among 793 ha of forested land on the local forest of Beaumont and Malaucène.

A statement under oath drawn up by the forest rangers on the 17 June 1889 brought a confirmation on the localization of the main Beech refuge areas: " The Beech stands were found on the East part of the massif in a very thin transversal band between the very steep escarpments and the stony summit cap. On the West part they were present at the level of the Mont Serein and in the upper part of the local forest of Beaumont and Malaucène". Documents from the land registers give evidence that in this relic zones, in no case the Beech populations were exclusive of any other tree species.

The reforestation of the massif was officially launched with the ministerial decree of the 22 march 1861 as part of the vast project called RTM, Restauration des Terrains de Montagne (mountain land reclamation). Some pictures taken during the first phases (1888 and 1903) of the RTM show the dramatically deforestation done on the North face of the massif and the very scarce Beech stands left (Appendix 3). In 1890, Beech stands were only represented as solitary trees scattered in the higher part of the massif and from 1550 down to 1400m the Beech stands was a stunted forest (Rouis 1895). The reforestation over the North face were launched after the decree of the 26 July 1892, around 30 years after the first works done on the South face. On North face, the European black pine (*Pinus nigra*) was the most introduced species as well as Scots pine and mountain pine. According to some documents and publications found on the RTM works, it seems that none significant plantations of Beech were performed on the North face (Ningre 1991, Carmantrand 1995, Jean 2008 and Vaucluse local historical archives records). Nevertheless, three limited attempts of Beech reintroductions have to be reported: at the Font du contrat close to the Mont Serein (1892-1894) and above 1200 m at the Tempête and Frache pass (1895).

B. Present distribution of Beech stands inferred by SPOT images analyses

In order to clearly identify the Beech stands from other tree species over the North face of Mont Ventoux, we intended to use both winter and summer Spot satellite images. However, the winter image was too shady and consequently very little informative (Figure 7 bottom). In summer thick clouds were also present (Figure 7 top), but their area did not hamper the analyses. This is why all the spectral bands of the satellite images taken during the winter were excluded from the study and all taken during summer were used. On the summer Spot satellite image, the spectral bands giving most of the information was the green channel.

First of all, a simple automatic classification was done with images both uncorrected and corrected for slope effects, to determine if the slope correction improves the results. The conclusion was satisfactory: the unsupervised classification done on the corrected Spot images were much more homogeneous than those with the initial raw images (uncorrected for slope effects). But the first unsupervised classification over all the spectral bands of the summer Spot satellite images gave not reliable results. The maximum likelihood classification algorithm fail to identify correctly the beech stands and quite often tree species were confused between each other and between other landscape features such as bare soil. Moreover no correlation was found with the UA map.

To improve the results, we analysed the link between species composition and the NDVI vegetation index. To that aim, NDVI was computed for each tree species occurring in the study area (Figure 8), and the resulting map confronted to the UA map**.** Beech appeared as a species significantly different from coniferous and other broadleaves trees ($p < 0{,}0001$) using Turkey's test. In other words, beech stands were characterized by high canopy density and therefore high NDVI. Using the NDVI as additional variable to determine the species composition thus appeared to be a good option.

Therefore, we have incorporated the NDVI in a new unsupervised classification shown in Figure 9 using in addition of the green band the NDVI map. The beech stands were identified with a quite good accuracy at stand levels (pink and red pattern on the classification map). However, it was not possible to establish an accurate map identifying the other tree species over the North face. These results could still be improved because the GRASS classification process confused some beech stands with other landscape features for instance some pink and red pattern were present at the bottom of the valley of the North face, in a location where no beech stands occur (Figure 8**)**. Some identification error could be due to the presence of the clouds hidden a part of the beech stands as is was shown Figure 7**.**

C. Analysis of population genetic diversity and structure

C.1) Quality of the microsatellite data set

The null allele frequencies estimated for all the loci and within each population ranged from 0.002 to 0.364 (Appendix 6). Loci sfc_1063 and sfc_18 seemed to be particularly affected by null alleles, with respective mean frequencies of 0.055 and 0.034. Twenty-six populations over 38 had null allele frequency different from 0 at the locus sfc_1063.

In order to test null allele frequencies significance, *Fis*-values were estimated for each plot and their significance tested with FSTAT. Overall, only four loci (FS1_15, Sfc_18, Sfc_1063, and cfs_5) had a significant positive *Fis-value* (indicating a heterozygous deficit) in 3 plots (VTX_N2, VTX_257_2, VTX_384) (Table 1). These 3 plots corresponded to very large Beech populations with a high level of genetic relatedness expected as the result of biparental inbreeding. In addition, none of the loci had a significant *Fis* over all the plots (Table 2). Thus, none the loci were considered to be significantly affected by null alleles, and all microsatellite markers were kept for the following genetic analysis.

The allelic diversity was high for most loci (Table 2). The number of allele per locus ranged from 9 to 22 (with 12 on average), the loci FS1_15 being the most polymorphic. The effective number of alleles ranged from 2 to 7.5, with an average value of 3.7. This discrepancy among Na and Nae value is typical of the L-shaped distribution of microsatellite alleles, with a combination of few frequent alleles and numerous rare alleles. The occurrence of missing data was overall small. There was no significant linkage disequilibrium over all the loci, there is no microsatellite marker associations (data not shown).

C.2) Population genetic analysis on the 38 plots

Genetic diversity within plots and over the studied area

Genetic diversity was high within the different plots (Table 3). The observed heterozygosity frequency Ho ranged from 0,605 to 0,763, while Nei's expected heterozygosity He ranged from 0,626 to 0,718 (Table 3). The average Fis value over all population was 0,015, and was not significantly different from 0. Only VTX_384 plot has a significantly positive Fis-value (average Fis = 0,046). Twelve of the 38 plots had at least one private allele with a total number of private allele Np=24 (Table 4). VTX_164 plot had the highest Np (5, among which 4 at locus sfc_1063). VTX_384 plot also had a high Np-value (4), but this can be due again to its large population size.

Genetic differentiation among plots

Genetic differentiation among plots was small with a mean Fst-value of 2,57%. The pairwise differentiation between plots ranged from 0,6 % (between the plots VTX_N2 and VTX_257_2 separated by 420 m) to 5,8 % (between RBI_256 and VTX_164, separated by 8 km).

Isolation by Distance (IBD) hypothesis was tested by plotting mean pairwise differentiation values calculated within 9 distances classes against distance (Figure 10). Though the differentiation value in the first distance class (for population separated by less than 150 m) was significantly lower than expected by chance (see the 95% confidence interval), there was no overall significant IBD pattern. Thus, linear distance did not contribute significantly to the genetic differentiation among the studied Beech plots. Consequently, the dispersal distance could not be estimated from the slope of the autocorrelogram.

The analyses of differentiation among cohorts (young versus old trees within each plot) showed that the "among cohort" component of variation within plot was low (1.1% of the total variation), though significant (Table 5). The variation among plots was still important (4.5%). Using these covariance components to estimate genetic differentiation, we obtained respectively: Rrt = 4.5% (among plots), Rsr = 1.1% (among cohort/plot) and Rst = 5.5%. This means that 1.1% of the variation within a subplot is because of temporal differenciation.

C.3) Detection of genetic units using Geneland.

Running Geneland model on all 1599 young or relic trees , the posterior distribution of the number of cluster K on the 10 runs displayed a clear mode at K=5 (Figure 11). By contrast, Running GeneLand model only on the relic-trees of the plots (trees with a diameter equal or higher than 1000 cm) gave no significant results. In the folowing, all the result are based on the full set of 1599 individuals. Individuals assignment into the five clusters are mapped on Figure 12 and 13. Three main clusters (1-3) were identified with 382, 840 and 372 individuals respectively and two marginal clusters with only 3 and 2 individuals respectively. No individuals were assigned to clusters 6, 7 and 8. According to Guillot et al (2008) we disregarded these populations which are no modal populations for none of the individuals (ghost populations) and considered the number of non empty populations as the correct K. The maps of the posterior probability of beech trees to belong to each cluster is shown on the Appendix 7. All individuals were assigned with a high posterior probability, especially those of the cluster 2. Note that maps presented in appendix 5 show posterior probabilities to belong to a given cluster cumulated over all individuals; therefore, in case where just one individual has a significant posterior probability to belong to cluster X whereas all other individual belong to another one, cluster X may appear as poorly supported on these cumulated maps. No potential migrants were clearly identified.

Almost all the RBI plots (Eastern sector) belonged to cluster 2, while the main genetic discontinuities appeared at the level of the VTX plots. Looking in details in VTX sector (Figure 13), most of the individuals were assigned to the cluster 2, predominantly in the North part of the zone. The two large populations, VTX_257_2 and VTX_N2 were assigned to cluster 3, which was poorly represented in the eastern part of the massif. Very few individual belonged to cluster 4 and none to cluster 5. The genetics discontinuities of the western zone were confronted to the landscape features. No physical barriers such as roads can explain the pattern of the different population (Figure 13). The inferred populations were also confronted to the digital elevation model. No clear elevation barriers (very steep slopes) among group of individuals assigned to different cluster could be identified. The same analyses were done on the East part of the massif at the level of the RBI plots and analogue results were found: the genetic discontinuities could not be explain neither by the topography or other landscapes features.

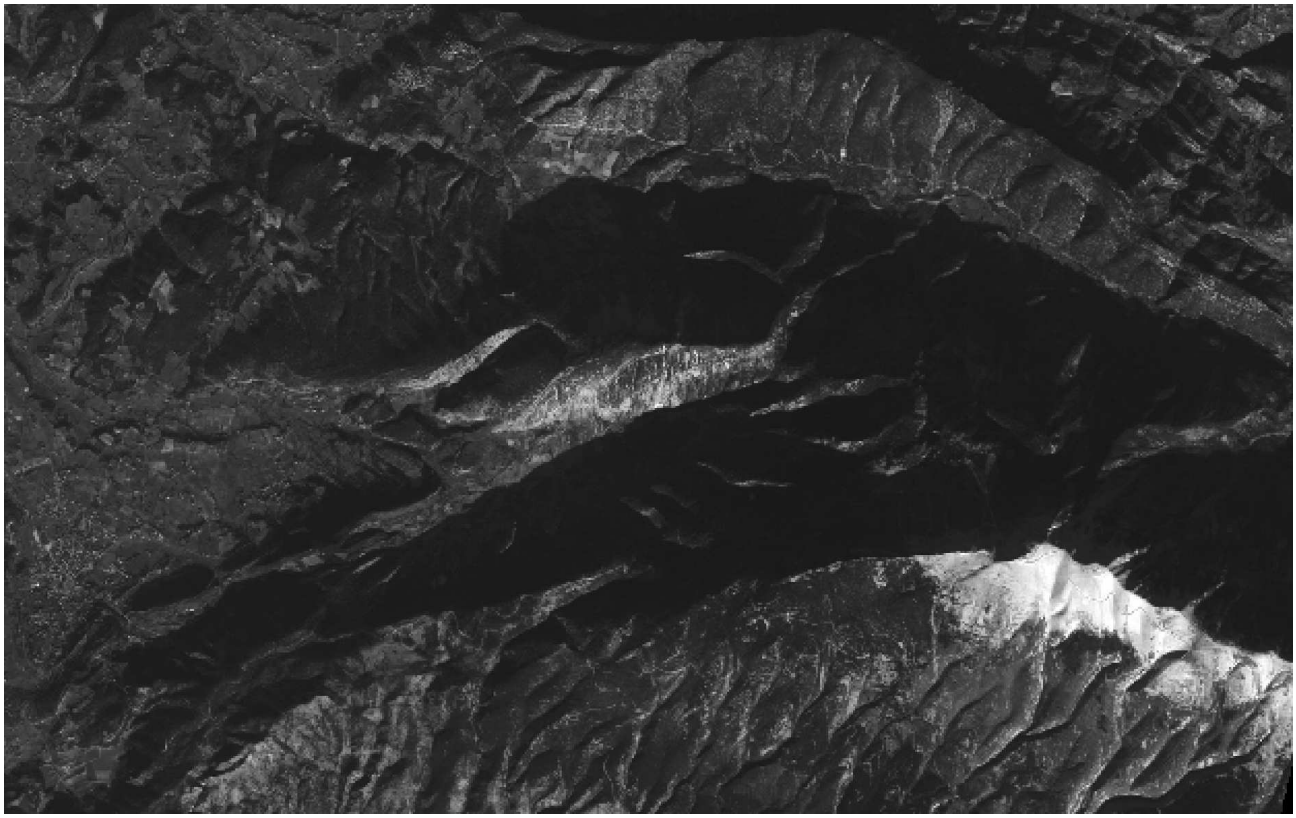
Figure 6: Map on the forested areas of the Mont Ventoux, 29 July 1876. The relic zones of the beech groves are located with the white squares (Vaucluse local historical archives).



Figure 7: Zoom on the SPOT images showing the area under study in A: summer and B: winter.



A



B

Figure 8: Variation of the Normalized Difference Vegetation Index (NDVI) computed over the studied area from SPOT summer image (see material and method for details). Dark lines delimit the UA.

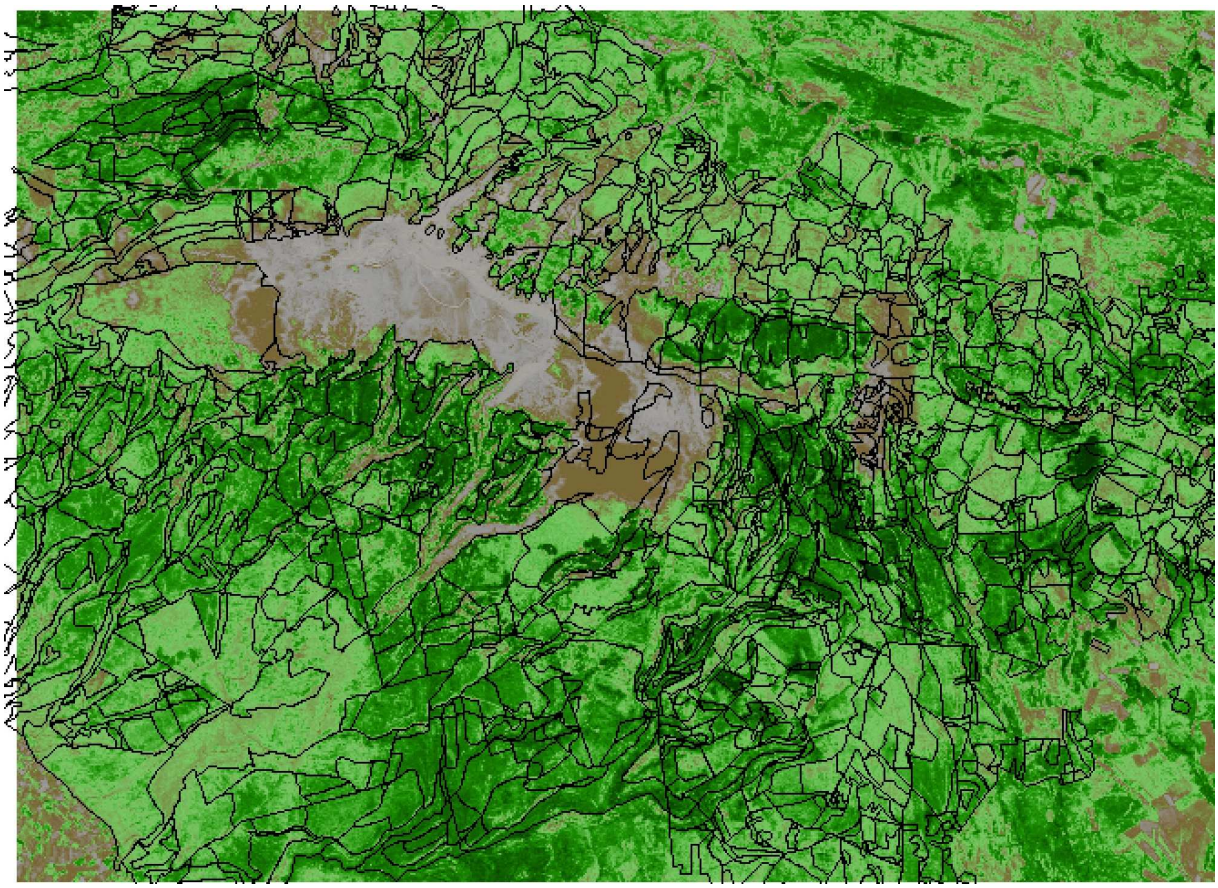
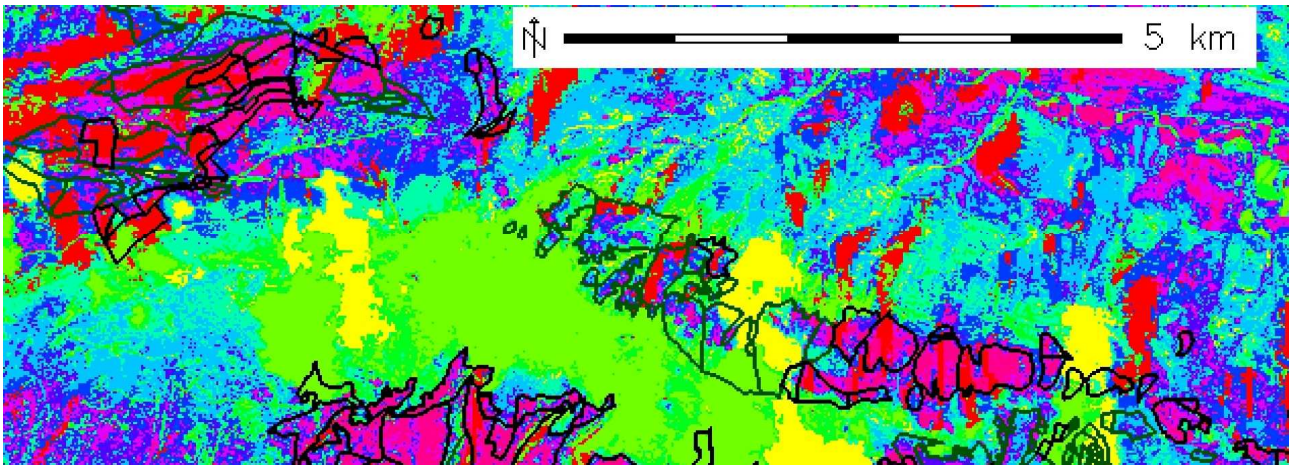


Figure 9: Results of unsupervised vegetation classification using green channel and NDVI as variables for the classification process.



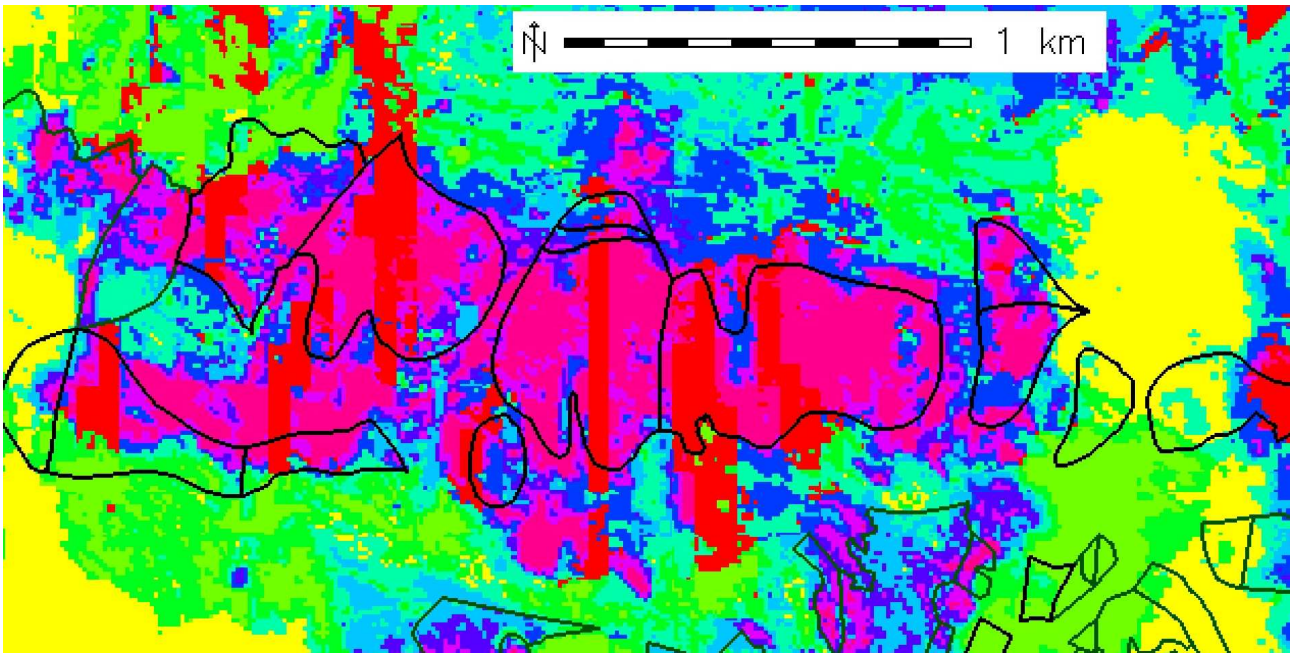


Table 1: Fixation index (*Fis*) and null allele frequencies estimates in locus per plot combinations where null allele presence were suspected. Null allelle frequencies (NAF) were estimated with ML-NULL Significance of *Fis* values were tested using Bonferroni correction implemented in FSTAT.

Locus	Plot	NAF	Fis*
FS1_15	VTX_N2	0,096	0,207
Sfc_18	VTX_257_2	0,070	0,170
Sfc_1063	VTX_384	0,139	0,242
cfs_5	VTX_384	0,138	0,264

Table 2: Allele diversity per locus over all the populations. Mas: Mean allele size (Pb), Na: mean number of alleles per locus, Ne: mean effective number of alleles, Fis: index of the heterozygous disequilibrium Abs %: percentage of individuals with missing genotypes, Nb Pop: number of population affected by these missing genotypes..

Locus	MAS (Pb)	Na	Nae	Fis	Abs %	Nb Pop
FS1_15	215.3	22	7,5	0.006	0,6	11
FS3_04	203.2	5	2	-0.066	0,3	5
Sf07_2	152.4	9	4,8	-0.011	0,6	10
Sfc_18	168.7	14	4	0.059	0,2	4
Sfc_36	205.0	12	3,1	0.089	0,2	4
Sfc_1063	200.3	11	5,1	-0.048	1,4	23
Sfc_1143	123.8	16	5,3	-0.033	0,2	4
cfs_25	173.3	11	2,2	-0.050	1,1	19
cfs_29	145.0	11	2,2	-0.059	1,4	23
cfs_31	118.5	12	3,8	-0.014	1,1	19
cfs_5	166.1	10	4	-0.027	1,4	24
cfs_6	217.7	10	6,4	-0.043	1,2	21
mfc7	116.6	13	3,1	-0.016	0,9	15
All loci	**169,7**	**12**	**3,7**		**0,8**	**12,2**

Table 3: Estimates of population genetics summary statistics for each population. Ho: observed heterozygosities, He: expected under HWE and Fis: fixation index (* indicates ignificant Fis at the nominal 5%level with the Bonferroni correction).

Plot	Na	Ho	He	Fis
RBI_10	6	0,682	0,662	-0,012
RBI_132	6.8	0,651	0,656	0,023
RBI_168	6.8	0,763	0,717	-0,047
RBI_170	6	0,703	0,682	-0,011
RBI_184	6.2	0,724	0,704	-0,011
RBI_186	6.5	0,718	0,688	-0,026
RBI_188	6.4	0,682	0,682	0,017
RBI_191	6.2	0,613	0,631	0,045
RBI_203	5.6	0,733	0,658	-0,090
RBI_21	6.2	0,640	0,663	0,051
RBI_217	6.2	0,666	0,646	-0,013
RBI_224	6.3	0,605	0,660	0,101
RBI_243	6.7	0,687	0,665	-0,016
RBI_256	4.4	0,714	0,640	-0,040
RBI_259	6.3	0,710	0,687	-0,017
RBI_267	6.4	0,682	0,670	-0,001
RBI_274	6.4	0,703	0,688	-0,004
RBI_28	6.2	0,723	0,697	-0,021
RBI_61	6.1	0,697	0,687	0,003

Plot	Na	Ho	He	Fis
RBI_76	6	0,679	0,664	-0,005
RBI_91	7	0,730	0,684	-0,050
RBI_ext_a	5.9	0,732	0,712	-0,010
RBI_ext_b	6.3	0,662	0,664	0,020
RBI_ext_c	5.9	0,728	0,685	-0,046
VTX_122	4.6	0,615	0,626	0,093
VTX_142	6.8	0,671	0,678	0,027
VTX_164	6.5	0,655	0,660	0,027
VTX_206	6.2	0,686	0,705	0,049
VTX_221	7.2	0,719	0,705	-0,007
VTX_229	7.1	0,742	0,718	-0,021
VTX_257_1	6.3	0,678	0,692	0,035
VTX_257_2	8.2	0,688	0,705	0,027
VTX_263	6.8	0,684	0,701	0,038
VTX_313	6.8	0,663	0,689	0,054
VTX_362_1	6.4	0,697	0,677	-0,017
VTX_362_2	6.8	0,691	0,687	0,008
VTX_384	8.5	0,669	0,700	0,046*
N2	7.8	0,697	0,710	0,024
Mean	**6.7**	**0,69**	**0,68**	**15**

Table 4: Number and frequency of private alleles

Plot	Locus	Np	Freq
N2	Sfc_1143	2	0,010
RBI_10	l025	1	0,017
RBI_132	Sfc0007_2	1	0,017
	Sfc_1143	1	0,017
RBI_168	Sfc0007_2	2	0,033
RBI_191	FS1_15	1	0,017
	l029	1	0,017
RBI_21	Sfc_0018	1	0,033
RBI_224	FS1_15	1	0,017
RBI_28	Sfc_0036	1	0,017
VTX_164	Sfc_1063	4	0,111
	mfc7	1	0,019
VTX_229	mfc7	1	0,012
VTX_362_2	l029	1	0,014
	l031	1	0,013
VTX_384	Sfc_1063	1	0,002
	l05	3	0,005

Figure 10: Correlogram of pairwise population differentiation, as measured by Fst/1-Fst, against logarithm of distance. Abscise values correspond to the upper limit of the distance intervals. Red dashed envelopes correspond to 95% confidence intervals obtained through 10,000 random permutation of genotypes on spatial locations. Confidence intervals around observed Fst/1-Fst values were obtained through a jacknife procedure consisting in deleting each locus at a time.

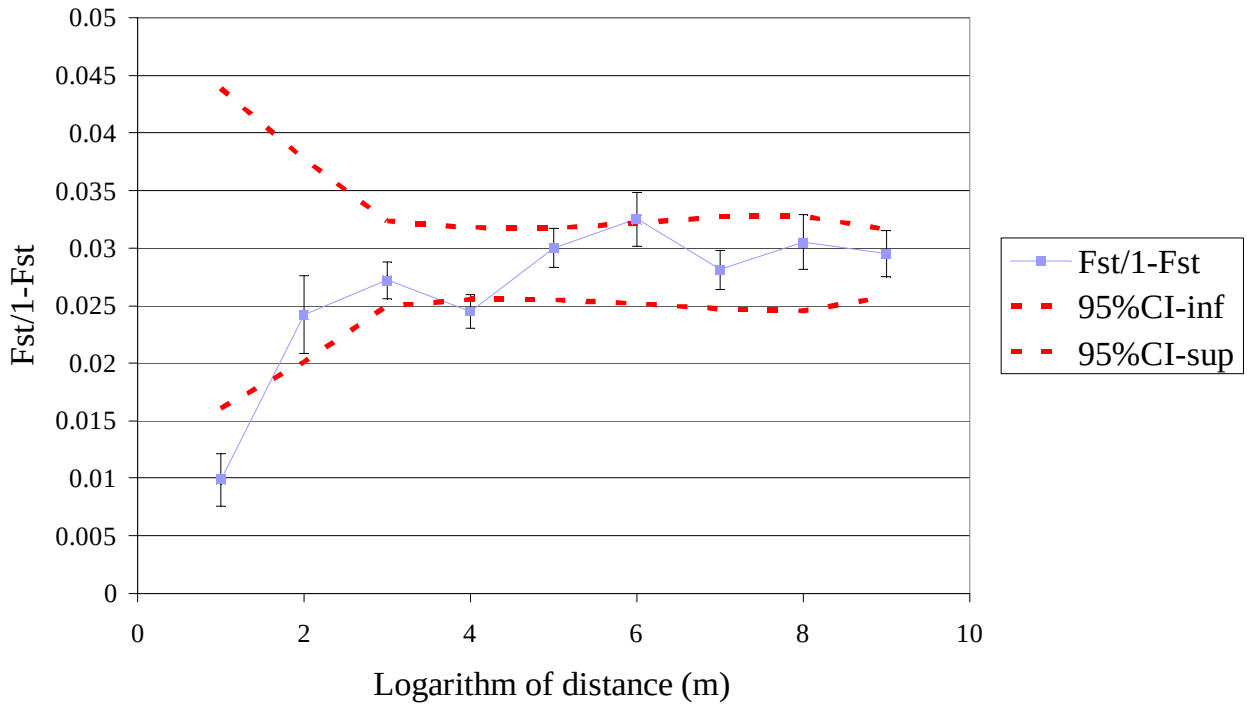


Table 5 Nested analysis of molecular variation for genetic variation among young and old trees within the 38 Beech plots

Source of variation	**df**	**SS**	**MS**	**Est. Var.**	**% total variance**
Among plots	32	828.462	25.889	0.429	4.5%
Among cohort within plots	35	376.204	10.749	0.102	1.1%
Within cohort	1097	9946.639	9.067	9.067	94.4%
Total	1164	11151.305	45.705	9.597	

Figure 11: Posterior density distribution of the number of clusters estimated from GeneLand analysis at Kmin= 1, Kmax= 8 after 1000 000 MCMC iterations.

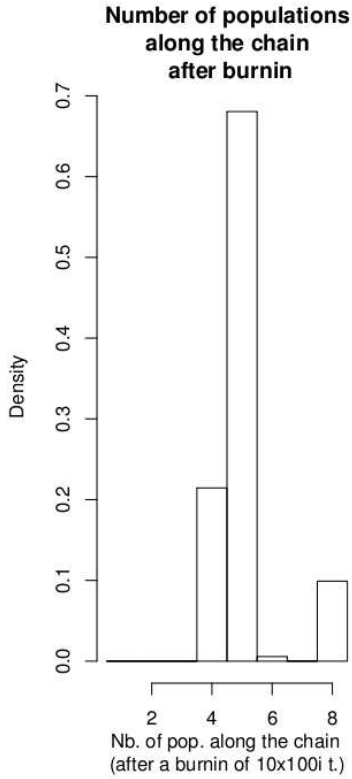


Figure 12: Spatial representation of the five significant genetic clusters inferred by Geneland.

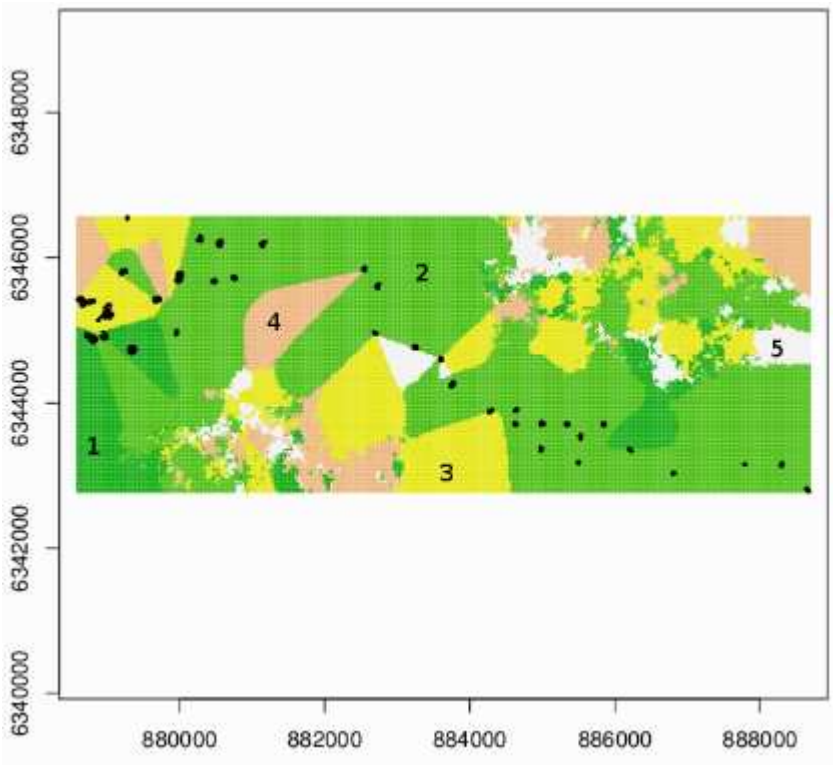
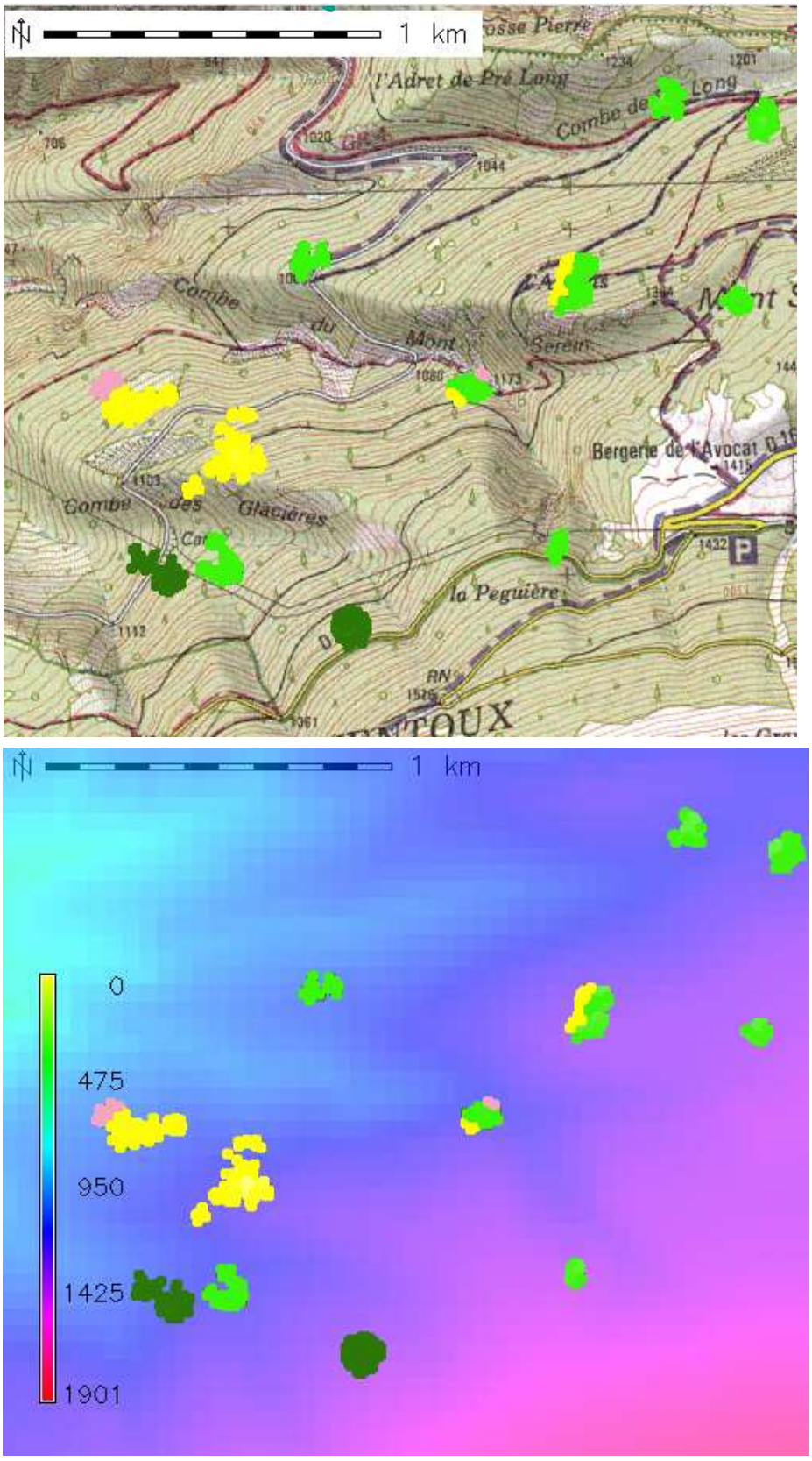


Figure 12: Identification of population discontinuities in the western zone of the study area.



Discussion

The combination of historical and geographical approaches with genetics data analyses highlighted the recolonisation history of the beech stands over the North face of the Mont Ventoux. This study illustrated the power of landscape genetics approach to provide valuable knowledge on the delimitation of population units and give reliable informations on how various landscape features influence genetic variation within and between population.

A. Main results and drawbacks of the different approaches used in this study.

Past records investigation : The investigation of past history of Beech recolonisation on North face of Mont Ventoux allowed us to get a better qualitative image of Beech stands distribution from the late XIX century to now. The main refuges areas, spared during the deforestation period, were clearly identified. However, it was not possible to estimate population sizes of Beech stands before and during the over-logging period. Our results highlighted the difficulties in getting quantitative figures from past records of forest management and history, even in a site like Mont Ventoux where a lot of information has been stored for a long time. A proper historicist approach in collaboration with specialists may help us to exploit with better efficiency the historical records. However, such detailed information as tree population size for a given species are particularly difficult to retrieve, considering that 1) the different species were often not distinguished from each other in past forest records (unless for plantation record, but Beech was not planted) 2) the area covered by Mont Ventoux is quite large. Records were particularly lacunar for the North face of Mont Ventoux, as compared to the South face, probably because 1) records documents relative to North face were stored in three different places (Avignon Pope Palace, Brantes and Valence) whereas those related to South face were always stored in Avignon; 2) much more attention was paid to the South slope where deforestation and soil erosion were more important than on North slope.

SPOT image analyses and vegetation classification assisted by GRASS : The unsupervised classification combining the vegetation index and the green spectral band allowed us to identify the current beech stands on the North face with a correct accuracy, as revealed by the comparison with the previous vegetation map of the ONF (AU map). Our result show the usefulness of satellite image for vegetation cartography at stand scale, even in mixed forest. Note however that if the objective is to distinguish among species at the level of the individual tree, higher resolution images would have to be used, such as aerial photographies. In our case, using only aerial photographies to map Beech stand at the scale of Mont Ventoux would have been clearly impossible, because of the high number of photographs required to cover the area and especially of the time required to analyse them. However, a good strategy may be supervised classification, with a first step based on few aerial photographies analyses in order to establish Beech spectral signature with accuracy in some standq of the North face, and then vegetation supervised classification of satellite image.

Landscape genetic approach: the spatial Bayesian clustering method implemented in Geneland was applied on our data set of 1599 individuals scored at 13 microsatellite loci. Despite the low overall differentiation among plots (2.57%), five different clusters were clearly identified, with a main population of 840 individuals (cluster 2). By contrast running Geneland on the sub-sample of 382 relic trees revealed no significant population structure (one single cluster was inferred). Additionnal analyses need to be done to understand whether these results indicate 1/ low accuracy of the set of 13 markers and 382 individuals to detect differentiation, or alternatively 2/ if all the different relic trees an be considered to belong to the same population. To tell among these different hypothesis, we intend to use simulations based on an individual model already developed in URFM and dedicated to the dynamics of Beech and Fir populations on Ventoux.

The main genetic discontinuities identified by Geneland were located in VTX sector (western sector), in particular at the level of the two large populations, VTX_257_2 and VTX_N2 assigned to the cluster 3 (rare in the other part of the massif). However, it was difficult to identify whether some landscape feature contributed to the observed patterns of genetic discontinuities. We could show that slope did not result in significant differentiation. Additional analyses have to be done to integrate more complex landscape features, such as forest management units.

B. Patterns of genetic differentiation among Beech stands

Overall, standard population genetic analyses indicated a low, though significant genetic differentiation among studied plots (Fst = 2.57% among the 38 plots). This low level of differentiation could be expected considering the local geographical scale of the the study, and very recent recolonisation of Mont Ventoux by Beech. However, the combination of historical investigation with landscape genetic analyses brought several new insights on the observed patterns of differentiation. First, historical records showed us that Beech stands were confined to a thin belt at around 1400 m in elevation below the Mont Serein, and to scattered fragmented stands in the communal forests of Beaumont and Malaucène (see Figure 6). Though reduced in size, Beech populations were thus never totally isolated from each other even at the period where forest cover was minimal. In particular, Beech population located in the inaccessible parts of the massif (on the top of a stony cliff, down a very steep slope) seemed to have been completely spared during the period of over-logging and over-grazing. Note that this was not the case for other tree species at lower elevation, as well as for Beech of the South face. The distribution of relic Beech population in a more or less continuous belt probably favoured genetic connection among fragments It is likely that pollen mediated gene flow contributed to a large extend to genetic exchanges between fragments, as pollen is expected to be efficiently dispersed at medium and long distance in this wind-pollinated species (Smouse and Sork 2004).

Consistently with the patterns of low genetic differentiation, no clear re-colonisation routes could be identified at the scale of the North face. In the eastern part of the studied sector (RBI) in particular, most of the trees belonged to a single genetic cluster (cluster 2), and there were no clear physical or geographical barriers between trees assigned to this main cluster and the other rare trees assigned to other clusters. The main genetic discontinuities identified by Geneland were located in western sector (VTX) at the level of administrative boundaries between the local forest of Beaumont and the local forest of Malaucène. The differentiation observed could be explained by past different managements, but this hypotheses need to be confirmed by detailed analyses of forest management records in this area.

Finally, historical record and patterns of genetic differentiation both agree on the absence of successful Beech plantations on the North face of Mont Ventoux. All the present beech-trees thus originate from the same weakly differentiated genetic pool. Considering the large extend of present Beech stands (see the UA map, Figure 4, and the results of vegetation classification on Figure 9), this mean that the recolonisation process was quite fast and efficient. The absence of significant isolation by distance also indicate that migration abilities did not limit the recolonisation process, unless the limiting component of distance was due to complex landscape features not accounted for here. Recolonisation required both the large production of seeds, and their successful dispersal and establishment. Previous studies of Beech seed dispersal abilities revealed a combination of preferential dispersal at short distance (median dispersal distance = 6.5 in Sagnard et al. 2007) with non negligible event of long-distance dispersal (Bontemps 2008): for instance, up to 40% of immigrant seeds were detected in a 4 ha plot (VTX_N2). However, the discrepancy between the scale of the dispersal process inferred through parentage analyses (Bontemps 2008) and the low differentiation overs ten of kilometres as observed in this study suggest that long-distance seed

dispersal alone is not sufficient to explain the weak spatial genetic structure.

The global picture of the recolonisation process of Beech on Mont Ventoux based on present and previous studies indicate that Beech progressively recolonised the North face of Mont Ventoux from a set of small but well connected relic-populations. Besides low spatial and genetic differentiation among Beech relic sources of seedlings, the continuous distribution of Pine plantations also provided overall important availability of sites for successful Beech seedling establishment.

C. Future response of Beech population to climate change (CC)

The past response of Beech population to a massive, human induced population decline illustrated the important migration and recolonization capacities of this species. While at the end of the XIX century, beech stands were represented as solitary groups of trees scattered in the higher part of the massif, after a maximum of 3 generations, Beech is present almost everywhere from an elevation from 950 to 1500 m on the North face. This is reassuring, considering that in the next future, Beech populations will have to face increased habitat fragmentation, and reduction in population size. However our results showed that the fast recolonisation pattern observed on the North face of Mont Ventoux resulted from the combination of good seed production and dispersal abilities with favourable conditions for seedlings establishment, and continuous distribution of relic populations.

The future response of Beech population to the observed and predicted climate change (CC) is difficult to predict. The rate and the absolute magnitude of CC is expected to be greater than that inferred at least for the last four million years. Tree responses to this CC is expected to result from the interplay between natural selection, which can lead to local adaptation, and gene flow (migration), that can bring adapted or maladapted genes within locally adapted populations (Savolainen et al. 2008). Moreover, identical genotypes may differ in their phenotypic responses due to variations of gene expression among different environment conditions (genetic plasticity). These different evolutionary processes will take place at different levels, individual, population and species.

Patterns of neutral genetic variation observed in this study over the North face of Mont Ventoux showed high level of diversity and low levels of differentiation. However, neutral molecular markers can not be readily used as a proxy for adaptive variation, and the molecular variation in genomic sequences involved in adaptation needs to be investigated. Such studies will be done in the next future, but they require high genomic resources currently not available in Beech. However, the present study show that the recent demographic history of population contraction and expansion on Mont Ventoux did not significantly reduced the genetic diversity or increased genetic differentiation among populations. This low impact of population history is likely to hold also for adaptive genes.

Conclusion

This study brought significant new insights on the recolonisation history of Beech on Mont Ventoux in the late XIX-early XXth century. Using of combination of historical record investigation and a landscape genetic approach we showed that 1) the relic Beech populations during the over-logging and over-grazing period were distributed as small but not isolated forest fragments on a belt at around 1400 m in elevation 2) the size of beech relic populations was reduced but not enough or not long enough to let genetic drift results in an important lost of genetic variability and the establishment of well structure refuges areas and 3) the recolonisation process was quick and resulted in a low pattern of genetic differentiation among the present Beech populations at the scale on Mont Ventoux. These results concerning the past response of Beech to a major population decline suggest that beech population of the North face of the Mont Ventoux are able to resist to quite level high level of disturbance without major reduction in genetic diversity of increase in population differentiation.

Glossary

Bonferroni correction: A correction used when several statistical tests are being performed simultaneously (since while a given á-value may be appropriate for each individual comparison, it is not for the set of all comparisons). In order to avoid a lot of spurious positives, the á-value needs to be adjusted to account for the number of comparisons being performed. Suppose we are testing for Hardy–Weinberg proportions at 20 loci. Instead of using the traditional 0.05 á-level, we would test at á of 0.05/20 = 0.0025 level. This insures that the overall chance of making a Type I error is still less than 0.05.

Co-dominant: In the case where both alleles contribute to the phenotype.

Fluorochromes: Molecular label conjugated to a primer that emits fluorescent light within a measurable color spectrum in response to a specific wavelength of laser light or chemical interaction(Fam- green, Tamra- yellow and Hex- bleu)

HotstarTaq® DNA polymerase: Taq polymerase is an enzyme isolated from bacteria that catalyse the polymerase chain reaction.Hotstart polymerases are modified versions of the original taq polylerase, modified in such a way that they need a 5-10 minute activation prior to PCR, which is convenient for reducing aspecific amplification and room temperature set up.

HWE: The Hardy–Weinberg principle states that both allele and genotype frequencies in a population remain constant, that is, they are in equilibrium from generation to generation unless specific disturbing influences are introduced. Those disturbing influences include non-random mating, mutations, selection, limited population size, random genetic drift and gene flow.

Mantel test: Statisitical test of the correlation between two matrices, commonly used in ecology, where the data are usually estimaes of the "distance" between organisms. For example, one matrix might contain estimates of the genetic distancs between all possible pairs of species in the study while the other might contain estimates of the geographical distance between the ranges of each species and every other species.

Neutral markers: Molecular marker not affected by natural selection

NDVI: Normalized Difference Vegetation Index, simple numerical indicator that can be used to analyze remote sensing measurements, and to assess whether the target being observed contains live green vegetation or not. Computed as

$$NDVI = \frac{NIR - R}{NIR + R}$$ with R the reflectance and NIR the Near Infra-Red absorbance.

PCR: Technique to amplify a single or few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. The method relies on thermal cycling, consiting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA.

PFGE: Pulsed-field gel electrophoresis (PFGE) is a method used to separate high molecular weight linear DNA fragments that are greater than 50 kilobases (kb) in size. Fragments are separated by pulsing an electrical field across an agarose gel rather than applying the constant field used in conventional electrophoresis.

Preprocessing level of 1A: Radiometric correction of distortions due to differences in sensitivity of the elementary detectors of the viewing instrument.

Primer: Strand of nucleic acid that serves as a starting point for DNA replication. They contain

sequences complementery to the DNA target region along with a DNA polymerase to enable selective and repeated amplification.

Spectral reflectance: (of a surface). At each wavelength in the visible spectrum, the proportion of incident light the surface reflects at that wavelength.

Size standard:

Solar irradiance: The amount of solar energy that arrives at a specific area at a specific time

Zenith angle: angle between the local zenith and the line of sight to the sun

References

Asuka Y., Tani N., Tsumusa Y., Tomaru N., 2004. Development and characterization of micorsatellite markers for *Fagus crenata.* Blume, *Molecular Ecology Notes*, 4: 101-103

Barbero M., Quezel P., 1987. La végétation du Ventoux, diversité, stabilité actuelles des ecosystèmes. *Etudes Vauclusiennes*, 3: 79-84

Badeau V, Dupouey J-L, Cluzeau C, Drapier J, Le Bas C (2004) Modélisation et cartographie de l'aire climatique potentielle des grandes essences forestières françaises. In: V*olet D1 du Rapport Final du projet CARBOFOR* « Séquestration de Carbone dans les grands écosystèmes forestiers en France. Quantification, spatialisation, vulnérabilité et impacts de différents scénarios climatiques et sylvicoles. » (ed. D. L), p. 49.

Belkhir k. et al. GENETIX, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Boisseau P., 2000. Gestion forestières du Ventoux. *Les carnets du Ventoux*, 29: 101-107

Bontemps A (2008) Caractérisation de la dispersion efficace du hêtre commun, Fagus sylvatica, sur le Mont Ventoux :Comparaison d'approches génétiques directes et indirectes, Université P. Cézanne - Aix-Marseille III.

Carmatrand R., 1955. Le reboisement du Ventoux. *Bulletin de la société forestière de Franche-Comté et des provinces de l'est*, 19 pp

Coulon A., Cosson J.F., Angibault J.M.A., Aulagnier S., Cargnelutti B., Galan M., Hewison A.J.M., 2006. Genetic structure is influenced by landscape features: empirical evidence from roe deer population. *Molecular ecology,* 15:1669-1679

Darluc M., 1782. Histoire naturelle de la Provence. Avignon, *Niel impr* T II: 523p

Davis MB, Shaw RG (2001) Range Shifts and Adaptive Responses to Quaternary Climate Change. Science 292, 673-679.

De Vienne D., 1998. Les marqueurs moléculaires en génétique et biotechnologies végétales. *INRA editions*, Paris, 201 pp

Du Merle P., Guende G., 1978. Presentation du Mont Ventoux. *La terre et la vie*, 32: 11-20

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. J*ournal of Heredity,* 86: 485-486.

Guillot G., 2008. Inference of structure in subdivided populations at low levels of genetic differentiation. The correlated allele frequencies model revisited. *Bioinformatics*, 24: 2222_2228

Guillot G., Estoup A., Mortier F., Cosson J.F., 2005a. A spatial statisitical model for landscape genetics. *Genetics*, 170: 1261-1280

Guillot G., Mortier F., Estoup A., 2005b. GeneLand: a computer package for landscape genetics. *Molecular Ecology Notes*, 5: 712-715

Hardy O. J., Vekemans X., 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes,* 2:618-620.

IPCC (2007) Climate Change 2007. Fourth Assessment Report. (ed. Nations U).

Jean H., 2008. Reboisement du Ventoux et restauration des terrains en montagne dans la vallée du Toulourenc. *Les carnets du ventoux*, 61: 72-75

Kalinowski S. T., Wagner A. P., Taper M.L., 2006. *ML-Relate*: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, 6:576-579

Kalinowski S., Taper M., 2006. Maximum likelohood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics*, 7: 991-995

Magri D., Vendramin G., Comps B., Dupanloup I., Geburek T., Gomory D., Latalowa M., Litt T., Paule L., Roure J., Tantau I., van der Knaap W., Petit R. & de Beaulieu, J., 2006. A new scenario for the Quaternary history of European Beech populations: palaeobotanical evidence and genetic consequences. *New Physiologis*, 171: 199-22

Manel S., Schwartz M. K., Luikart G., Taberlet P., 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, 18: 189-197

Martins C., 1838. Essai sur la topographie botanique du Mont Ventoux en Provence. *Annale des Science Naturelles*, 10: 129-150

McLachlan JS, Clark JS, Manosa PS (2005) Molecular indicators of tree migration capacity under rapid climate change. Ecology 86, 2088–2098.

Meygret A., 2006. Spot absolute calibration: Synthesis. *CNES edtions*, 29 pp

Neteler M., Mitasova H., 2008. Open source GIS: A GRASS GIS approach. Third edition. *Springer*, 406 pp

Ningre J.M., 1991. Le reboisement du Ventoux au siècle dernier. *Les carnets du Ventoux*, 12: 49-58

O'Brien E.K., Mazanec R.A., Krauss S.L., 2007. Provenance variation of ecologically important traits of forest trees: implications for restoration. *Journal of Applied Ecology* 44, 583-593.

Parmesan C (2006) Ecological and Evolutionary Responses to Recent Climate Change. Annual Review of Ecology, Evolution, and Systematics 37, 637-669.

Pastorelli P., Smulders M. J. M., Van't Westende W. P. C., Vosman B., Giannini R., Vettori C., Vendramin G., 2003. Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis*. Lipsky. *Molecular Ecology Notes*, 3: 76-78

Peakall R., Smouse P. E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288-295

Roman-Amat B (2007) Préparer les forêts françaises au changement climatique – Rapport, 125 p. Ministère de l'agriculture et de la Pêche, Paris.

Reusch T.B.H., Wood T.E., 2007. Molecular ecology of global change. *Molecular Ecology* 16, 3973-3992.

Rob P., 1992. Ecology of Beech forests in the Northern hemisphere. *Wageningen*, 122 pp.

Rouis E., 1895. Note sur la flore phanérogamique des environs de Carpentras, du Ventoux et des monts du Vaucluse, 14, vol 1

Sagnard F., Pichot C., Dreyfus P., Jordano P., Fady B., 2007. Modelling seed dispersal to predict seedling recruitment; recolonization dynamics in a plantation forest. *Ecological modelling*, 203: 464-474

Samper P., 2007. Développement de marqueurs microsatellites pour caractériser des populations de hêtre (*Fagus sylvatica*) sur le Mont Ventoux. Rapport de Licence. 28 pp

Savolainen O, Pyhajarvi T, Knurr T (2007) Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics* 38, 595-619.

Smouse PE, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management* 197, 21-38.

Tamaka K., Tsumara Y., Nakamura T., 1999. Development and polymorphism of micorsatellite markers for Fagus crenata and the closely related species F. japonica. *Theoretical and applied genetics*, 99:11-15

Tessier du Cros E., Le lacon F., Nepveu G., Pardé J., Timbal J.E., 1981. Le hêtre. Département des recherches forestières, *INRA editions*, Paris, 613 pp.

Thinon M., 2007. La vegetation du Ventoux au cours des derniers millénaires. *Forêt méditérranéenne*, TXXVIII, **4**: 289-294

Thuiller W, Lavorel S, Araujo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *PNAS* 102, 8245-8250.

Wright, S., 1969. Evolution and the Genetics of Populations, Vol. II. The Theory of Gene Frequencies. *University of Chicago Press*, Chicago.

Appendixes

Appendix 1: Detailed information of the thirteen microsatellites primers used in DNA amplication by multiplex PCR.

Locus	Reference	Repeat sequence	Primer sequence (5'-3')	Size range (bp)
FS1_15	Pastorelli et al., 2003	(GA)26	TCAAACCCAGTAAATTTCTCA GCCTCAATGAACTCAAAAAC	95-135
FS3_04		(GCT)5(GTT)3 (GCT)6	AGATGCACCACTTCAAATTC TCTCCTCAGCAACATACCTC	201-207
sfc0007_2	Asuka et al ., 2004	(AG)24	TGTCGCAAACATTGACAAGG GTGGATGTGAGGTCGTTGG	149–157
sfc0018		(AG)17	GAAGCAGAGCATTGTATTGG CATCTGTTTCAGTTCTGTAAAGG	161–191
sfc0036		(TC)23	CATGCTTGACTGACTGTAAGTTC TCCAGGCCTAAAAACATTTATAG	96–142
sfc1063		(CT)13	TTTCCAACTACAACTTCATTG AGTGCTCGCATCGTATG	188–222
sfc1143		(AG)21	TGGCATCCTACTGTAATTTGAC ATTCCACCCACCATCTGTC	96–136
mfc7	Tanaka et al ., 1999	(GA)9	AAAATACACTGCCCCCAAAA CAGGTTTTGGTTTCTTACAC	111-129
cfs_ 05	unpublished primer sequences	(TA)10	- -	184
cfs_ 06		(AG)13	- -	128
cfs_25		(AT)11	-	177
cfs_ 29		(CT)11	- -	142
cfs_ 31		(AG)12	- -	116

Appendix 2 : Details on GeneLand Clustering method

Maximum-Likelihood estimate of the parameter of interest

The vector of parameters to be estimated by the GeneLand MCMC model is θ =(*K*, *m*, *u*, *c*, *d*, *f*, *fa*, *s*), with *K*, the number of population, *m*, the number of domain per population, *u*, the location of the different domain, *c* the colours of tiles, *d* the drift parameters, *f*... and *fa* the allele frequencies in the present and ancestor population respectively. Denoting *z* the true location of individual (geographical coordinates of the individuals) and *t* the genotype of the sampled individuals, the likelihood of the data π (*t*,*z*|θ *)* can be expressed as:

$$\pi(t, z|\theta) = \pi(t|\theta)\pi(z|t, \theta) = \pi(t|\theta)\prod_{i=1}^{n}\prod_{l=1}^{L}\pi(z_{i,l}|\theta).$$

The term of the product is given by the allelic frequencies,

$$\pi(z_{i,l} = (\alpha, \beta)|\theta) = \begin{cases} 2f_{kl\alpha}f_{kl\beta} & \text{if } \alpha \neq \beta \\ f^2_{kl\alpha} & \text{if } \alpha = \beta. \end{cases}$$

Bayesian principle of inference by MCMC under GeneLand

All parameters are randomly initialized from uniform prior. Then, the moves of the MCMC chains are proposed as follows :

1/ update d

2/ upadta f_{Al}

3/ update f_{kl}

4/ update c_j for j=1,...,m

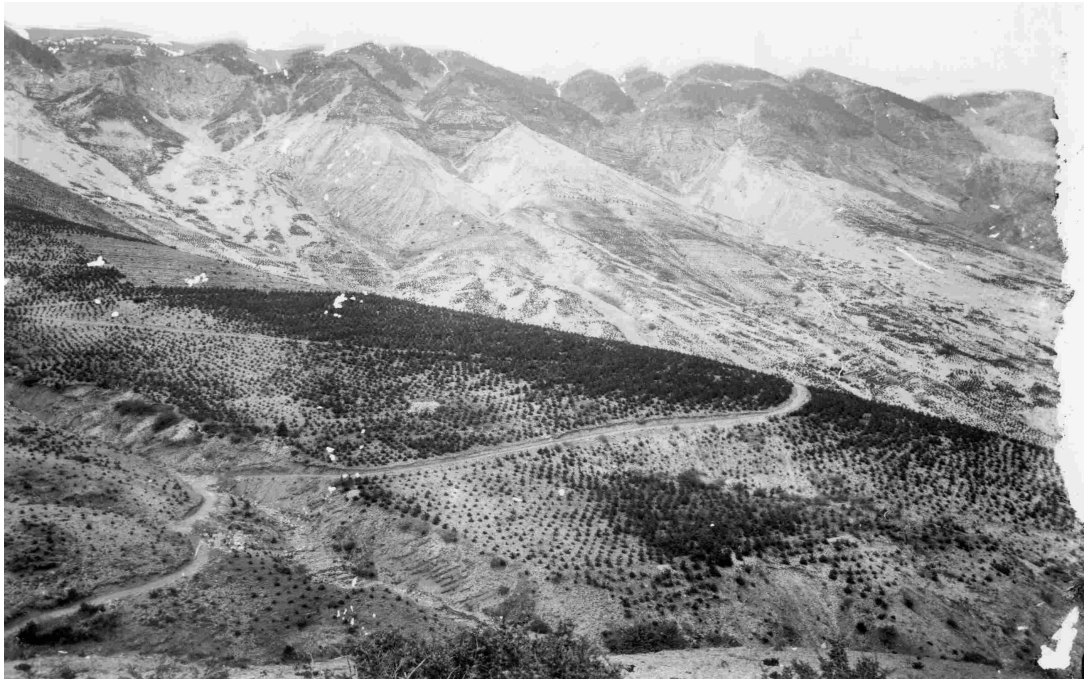
5/ update u_j for j=1,...,m

6/ update s_i for j=1,...,n

7/discard or add a tile (increase or decrease m by 1)

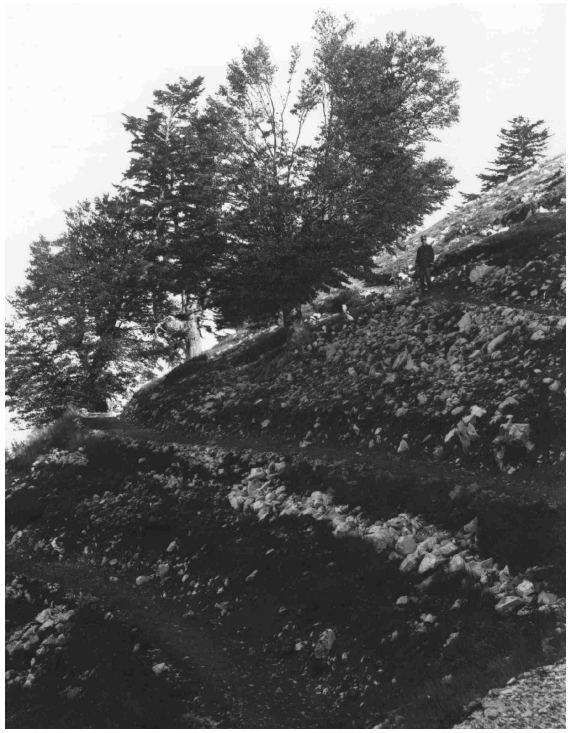
8/ Split one existing population into two or merge two to one (for instance increase or decrease K by 1)

Appendix 3: Pictures taken in 1888 and 1903 and showing the deforestation impact



The north face of the massif in 1903. On the foreground the first artificial plantations done during the RTM efforts are clearly visible. On the background, the deforestation impact is obvious: there is no trees left below an elevation of 1850m. The Beech stands is reduced to the higher part of the massif mixed up with fir just below some scarce pines close to the summit cap of the massif. However at least at this part of the mountain, the Beech stands was not totally fragmented into isolated populations, there was still a slight continuous vegetation band.





Two isolated Beech saved from the overexploitation period on the north face. Around them the soil is definitly stony presenty no suitable condition for a rapid recolonization. A man is standing up

close to one tree, showing that the relic-trees were well developed.

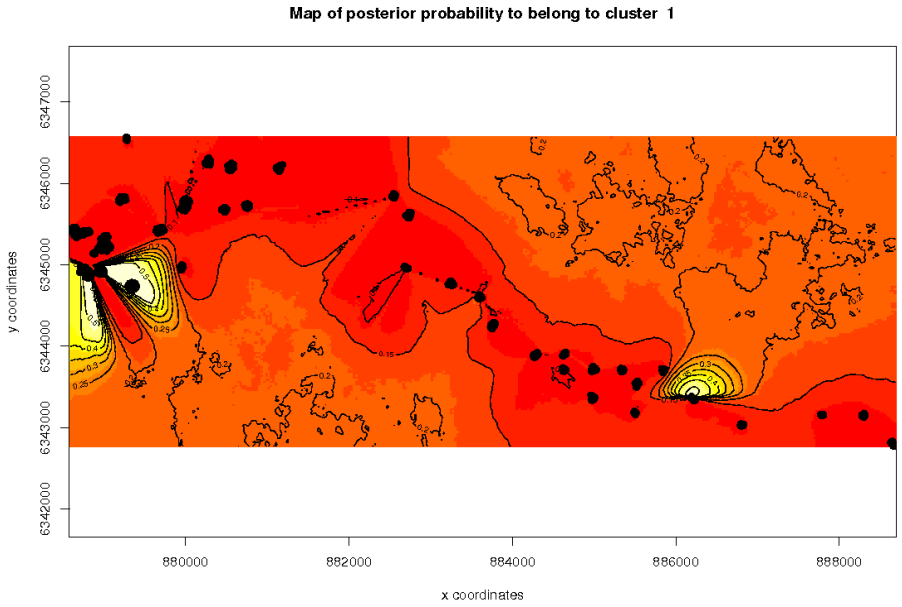
Appendix 4 : Estimation of allele null frequencies for each locus in each experimental plots.

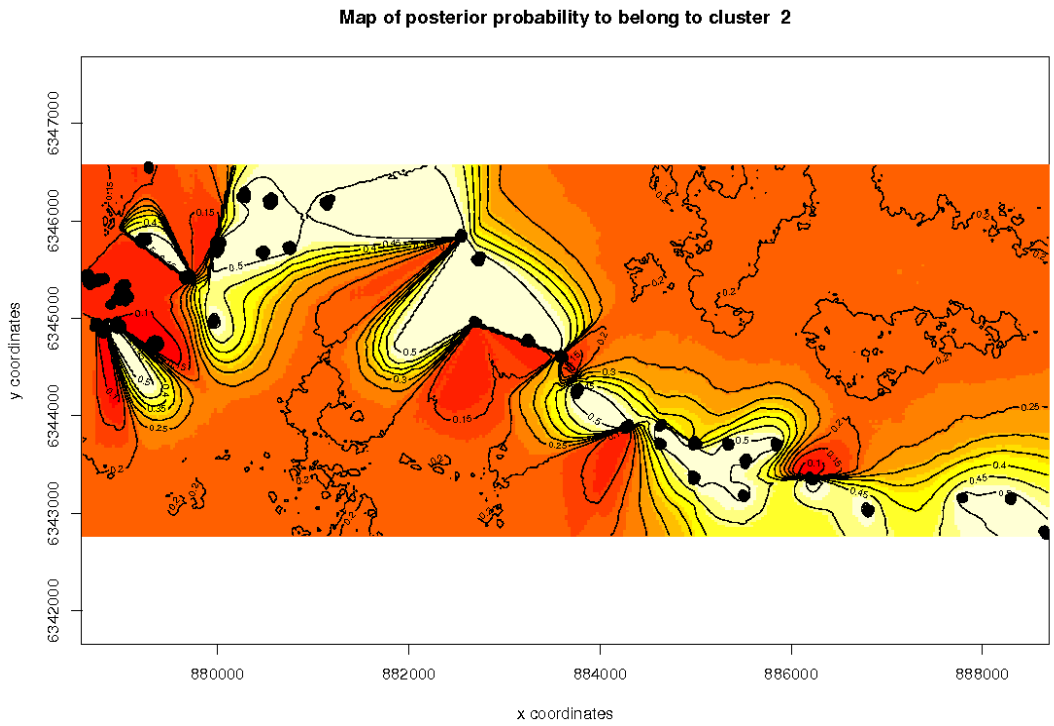
Values are given in percentage. - indicate a null allele null frequencies.

	FS1_15	FS3_04	Sf07_2	Sfc_18	Sfc_36	Sfc_1063	Sfc_1143	cfs_29	cfs_25	cfs_31	cfs_5	cfs_6	mfc7
VTX_N2	9,6	-	-	5	-	1,1	-	-	1,6	1,9	-	1,4	0,6
RBI_10	4,8	5,4	-	-	4,8	2,5	1,4	-	-	-	7,3	-	2,2
RBI_132	-	4,4	-	3,9	-	2,2	1,5	-	-	-	0,7	2,7	-
RBI_168	-	2,1	-	-	-	-	-	-	-	-	-	1,2	-
RBI_170	-	-	3,4	-	-	-	2,5	3,1	-	-	-	-	-
RBI_184	-	2,6	-	-	-	-	-	4,3	-	-	1,9	-	-
RBI_186	-	-	2,5	-	-	-	-	-	-	-	-	8	-
RBI_188	-	3,6	-	2,5	-	2,5	-	-	-	-	4,7	1,9	-
RBI_191	4,4	11,1	3,6	-	6,1	6,7	-	-	-	-	3,4	-	-
RBI_203	-	-	-	-	-	-	-	-	19,9	-	-	-	-
RBI_21	-	-	1,5	2,9	4,9	7,2	2,7	-	8,6	-	-	0,9	3,3
RBI_217	-	-	3,6	-	-	-	-	-	-	-	-	-	-
RBI_224	-	-	-	6	0,5	36,4	0,8	-	13,9	-	1,4	-	-
RBI_243	-	-	-	7,7	1	-	-	2,4	1,1	-	-	-	-
RBI_256	-	-	5,9	1,8	-	-	-	-	-	-	-	-	-
RBI_259	5	-	-	1,2	3,3	-	-	-	1,1	1,8	-	-	-
RBI_267	-	6,7	-	7,7	-	4,7	-	-	-	-	-	-	6,1
RBI_274	3	-	-	2,7	1,3	2,4	-	-	-	-	0,7	2,7	-
RBI_28	2,6	-	-	-	2,3	-	-	-	7	-	-	1,1	-
RBI_61	3,9	-	-	14,4	-	7,3	-	-	-	3,5	-	2,4	-
RBI_76	-	11,5	-	3	-	6,1	-	2,1	-	-	2,4	-	-
RBI_91	7,1	-	-	-	4,4	-	-	-	-	-	-	-	-
RBI_ext_a	1,4	-	2,1	6,8	-	0,7	-	-	-	4,1	-	2,8	-
RBI_ext_b	-	3,2	-	-	-	5,3	-	3,1	-	-	1,9	-	9
RBI_ext_c	2,3	-	6,5	0,3	3	-	-	-	-	-	-	-	-
VTX_122	-	9,6	-	12,2	-	1,8	-	4,7	-	-	-	-	7,2
VTX_142	1,6	-	4,2	2,9	-	6,6	1,7	-	-	-	3,7	3,6	-
VTX_164	8,3	-	9,2	-	-	11,6	-	6,9	-	-	-	-	-
VTX_206	-	-	5,9	5,7	-	5,3	9,4	2,4	-	5	17,7	-	-
VTX_221	1,4	-	4,5	-	-	27,2	-	-	-	-	-	-	-
VTX_229	-	-	-	-	-	4,3	-	-	3,4	0,6	-	-	-
VTX_257_1	8,5	-	2,9	13,7	-	8,5	-	-	-	-	-	-	-
VTX_257_2	6,1	-	2,2	7,1	1,7	4,1	-	-	-	-	-	0,9	-
VTX_263	-	-	3,8	4,4	-	18,1	-	5,6	-	5,8	-	-	-
VTX_313	-	-	-	5,4	5	6,2	0,2	8,2	18,5	-	12,1	-	-
VTX_362_1	-	-	0,6	1,3	-	9,1	0,2	-	-	-	-	-	-
VTX_362_2	-	-	-	7,3	6	8,5	2	-	-	4,4	-	-	-
VTX_384	-	-	-	2,7	-	13,9	2	-	-	-	13,8	-	-
Mean	**1,84**	**1,58**	**1,64**	**3,38**	**1,17**	**5,53**	**0,64**	**1,13**	**1,98**	**0,71**	**1,89**	**0,78**	**0,75**

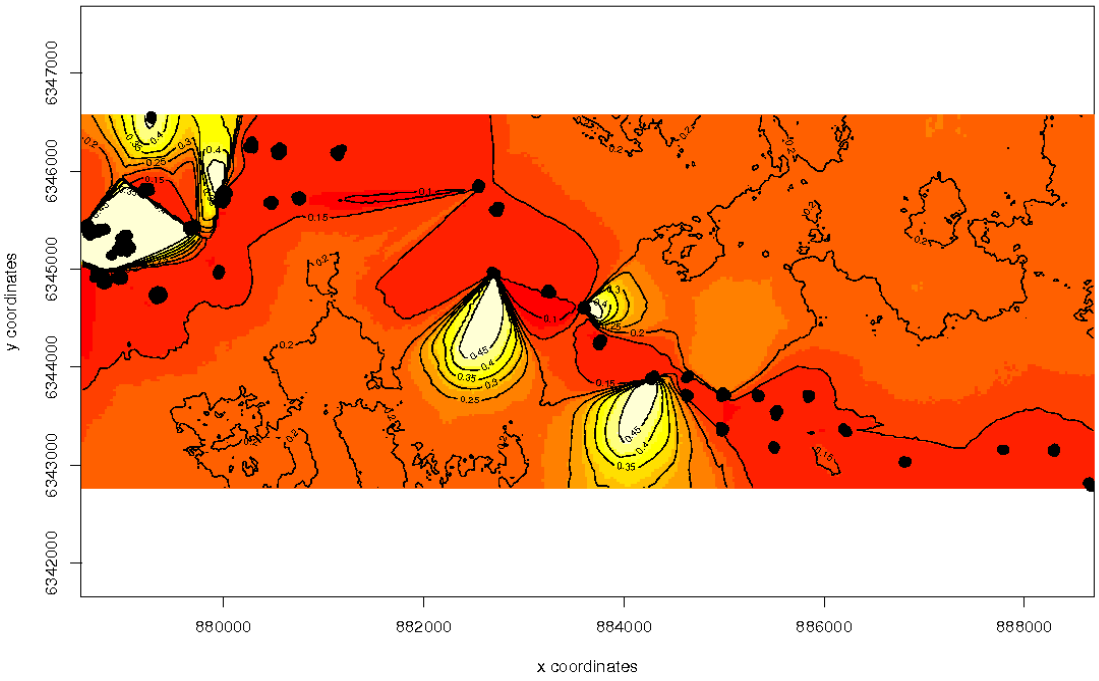
Appendix 5. Map of posterior probability of an individual t belong to one cluster, based on the highest probability run at Kmax=8.

The assignment of pixels to the cluster 6 to 8 are not shown, as no individuals were assigned to them (ghost populations). The highest membership values are in light yellow and the curves illustrate the spatial changes in assignment values.

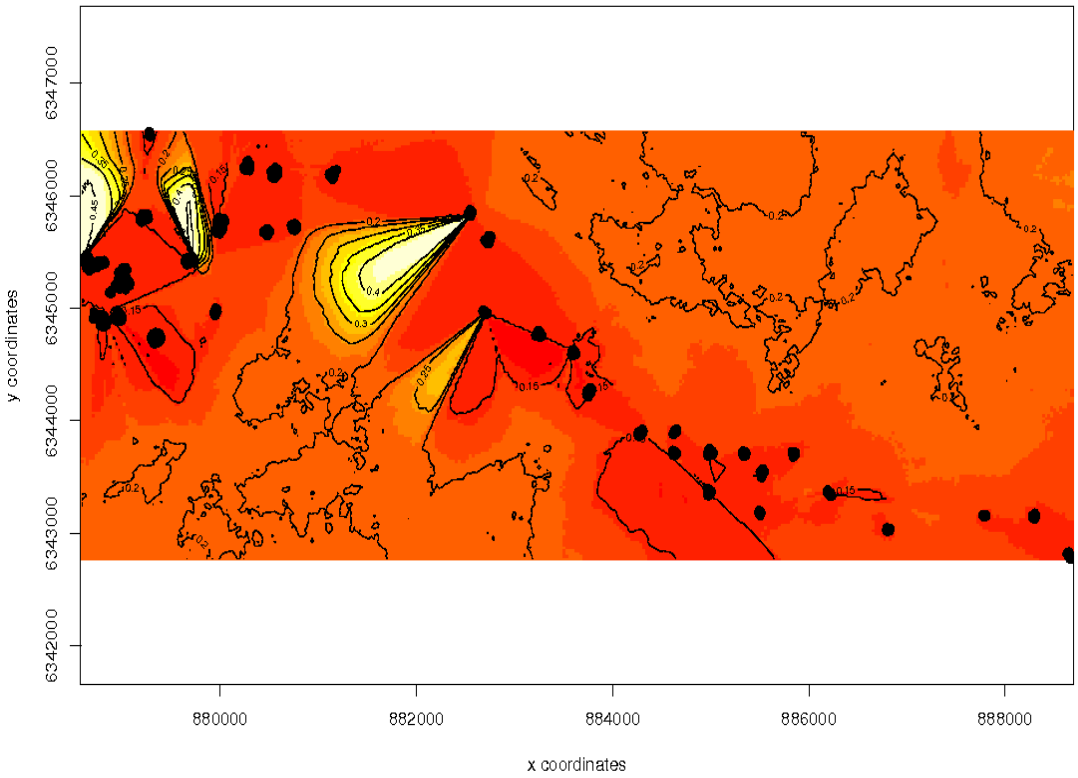




Map of posterior probability to belong to cluster 3



Map of posterior probability to belong to cluster 4



Abstract

In this study, we investigated the genetic structure and the recolonisation process of the European beech (*Fagus sylvatica*) over the north face of the Mont Ventoux Mountain, using of combination of historical record investigation, vegetation mapping from satellite image and unsupervised classification process, and a landscape genetic approach. Mont Ventoux has undergone large deforestation phases until the XIXth century due to over-grazing and over-logging for woofd supply. Historical records showed that the relic Beech populations were distributed as small but not isolated forest fragments on a belt at around 1300 m in elevation. Vegentation classification based on satellite images allowed to map with a correct accuracy the present distrribution of Beech on North face of Mont Ventoux, and confirmed its presence almost everywhere between 950m and 1500 m in elevation. Finally, we used the landscape genetic approach implemented in Geneland software in order to identify the delimitation of population units and to correlate the genetic discontinuities to landscape features. Over 38 experimental plots, we sampled 1599 individuals divided in two cohorts (young beech stands and relic-trees), and genotyped them at 13 microsatellite loci. Standard population genetic analyses showed small but significant genetic differentiation among experimental plots (Fst = 2.57%). Moreover no significant pattern of isolation by distance among plots was found. However the GeneLand model success to identified 3 main clusters and 2 marginal populations. No obvious correlation were found between the genetic discontinuities and the topography or the landscapes features. These results concerning the past response of Beech to a major population decline suggest that beech population of the north face of the Mont Ventoux are able to resist to quite level high level of disturbance without major impact on their genetic structure.