Use of DNA-markers for tracing illegal logging

Bernd Degen[1], Matthias Fladung[1]

Abstract

Illegal logging is one of the main reasons of deforestation in natural forests and is the cause of high ecological and economic damage. Genetic methods are useful to infer species identity and are promising tools to control the geographic origin of logged timber. DNA barcoding and multilocus approaches using nuclear and chloroplast microsatellites as well as Single Nucleotide Polymorphism (SNPs) are the main methods in use to determine species identity. Due to recolonisation after the last glacial periods and limited pollen and seed dispersal we observe a spatial genetic structure for most species in natural forests. Genetic inventories with extensive and systematic samples over the whole species distribution area are the basis to identify the country of timber origin. In order to control for the origin of timber on the level of a logging concessions genetic inventories with a higher spatial resolution are needed. DNA extracted from wood generally is degraded in comparison to that from fresh leaf material. Thus for the practical application only gene markers that can be easily amplified in DNA extracted from unprocessed and processed timber are usable for geographic timber tracking. This has been sucessfully tested for microsatellites (cpSSRs, nSSRs) with amplified DNA fragments usually not larger than 500 bp. The power and possible spatial resolution of gene markers to identify the geographic origin of timber depends on the spatial genetic structure in the species distribution area and the quality of the genetic reference data base (the number and distribution of sampled populations, sample size within each sampled population) and the geographic distance between the geographic origin of the unknown timber probe and the next sampled population in the reference data base. Here computer simulations are very useful to design suitable sample strategies for timber tracking.

[1] Johann Heinrich von Thünen Institute (vTI), Institute of Forest Genetics,
Sieker Landstrasse 2, 22927 Grosshansdorf, Germany.
bernd.degen@vti.bund.de

Keywords: Chloroplast, DNA barcoding, gene marker, identification, illegal logging, micro-satellite, recolonisation, refugia, spatial genetic structure, simulation studies, tropics

1 Introduction

Illegal logging and trade with illegal timber and wood products are the cause for many economic and ecological problems both in the producer and in the consumer countries. Illegal logging is believed to be one of the chief causes of worldwide deforestation and trade with illegal timber and wood products creates market disadvantages for products from sustainable forestry. Moreover, fallow land produced by illegal logging contributes to climate change by releasing greenhouse-relevant gases, The OECD assesses global damages through illegal timber at approx. € 150 billion per year. According to estimates, approx. 50 % of timber exports from the Amazon Basin, Central Africa, South-East Asia and the Russian Federation originate from illegal logging. Since illegal logging poses also a major threat to forest biodiversity, the 9th Conference of the Parties of the Convention on Biological Diversity (CBD) most recently in May 2008 has urged countries to strengthen forest law enforcement and governance and to prevent illegal logging and related trade.

So far, wood anatomical methods have been made available to identify species identity for many of the traded tree species (see Koch et al. this issue). We lack, however, practicable control mechanisms to identify the species identity of a remaining set of important timber tree species and we need methods to trace the geographic origin of timber and wood products. Genetic methods have the potential to address the topic of tracing illegal logging on three scales:

1.1. Species identification

It is relatively common for tropical producer countries to prohibit the logging or export of certain species in certain forms. The Convention on International Trade in Endangered Species (CITES) has also the objective to protect a set of endangered tree species. CITES controls seldom

ban trade entirely, but do require that special legal documents be presented at export and import. A common way in which criminal traders seek to circumvent CITES controls is by providing a false declaration of the species involved. To overcome this problem genetic methods to control species identity would be very helpful.

1.2. Control of the geographic origin of timber

The control of the geographic origin in timber is an important issue as well. The falsification of the country of origin is another well-documented area of illegality in the trade in tropical timber. This occurs at the point of import for timber that is in international trade, and usually involves the production of false paperwork such as phyto-sanitary certificates, invoices and certificates of origin. An actual example is the ban of the EU and the USA on teak from Birma.

Another common problem of illegal logging on smaller spatial scales is the false declaration of timber that has been logged outside a registered concession, or within a protected area. On this scale certified forest companies might have an economic interest to apply genetic fingerprints to proof their efforts of sustainable forest management.

2 Genetic approaches

2.1. DNA extraction

A number of different DNA extraction protocols from dried and processed wood from different tree species have been developed in the last 10 years (De Filippis and Magel 1998, Deguilloux et al. 2002, Rachmayanti et al. 2006, Asif and Cannon 2007). These protocols have in common the challenge to avoid contamination with external DNA and to minimise DNA degradation (Deguilloux et al. 2002). In an early study, De Filippis and Magel (1998) demonstrated for *Robinia* that the procedures and protocols developed for leaves are applicable to wood and that RAPD-PCR technology is a versatile and sensitive method of detecting genomic changes in trees. However, all proceeding paper claim that markers of choice to circumvent both problems are those that (a) show species specificity and (b) are of small size following PCR amplification and are present in high copy number (chloroplast, mitochondrial, or repeated nuclear sequences; Deguilloux et al. 2002, Rachmayanti et al. 2006).

DNA markers of choice fulfilling the mentioned requirements and that were already successfully applied in paternity analyses and species identification are microsatellite (SSR) markers (Ziegenhagen et al. 2003, Ziegenhagen and Fladung 2004). The fact that dry wood is accessible for molecular genetic investigations opens a wide horizon of potential downstream applications. In a study to control the geographic origin of oak wood destined for the French barrel industry Deguilloux et al. (2004) detected the existence of unlabeled oak woods originating from Eastern Europe and the incorrect use of the names of famous French forests. We extracted DNA from wood in another study to detect foreign genes in wood sampled from genetically modified trees (Fladung et al. 2004). In comparison to stem wood taken from herbarium specimens, Asif and Cannon (2007) investigated the possibility of identification of an endangered tropical timber species using sequencing technology for wood DNA.

2.2. Species identification

DNA-barcoding is a genetic approach to distinguish between different species. Here differences of the nucleotide sequence at specific target DNA-regions are used for identification. These target DNA regions show genetic differences among different species but not or only to a small degree within different individuals of a species (Taberlet et al., 2007, Kress et al., 2005, Hebert et al., 2003, 2005). Other characteristics required from these target DNA-regions are that they are present in most taxa and that they are easy to sequence. For species that perform photosynthesis the so called ITS-region (Internal Trancribed Spacer) of the nucleus has been successfully used for taxonomic purposes during the last ten years (Syring et al., 2007, Barker et al., 2007, Mort et al., 2007, Kenicer et al., 2005, Erikson et al., 2003). The genome of the chloroplast of plants is highly conserved in terms of size (120-170kb), structure and linear order of genes. With a few exceptions the chloroplast genome includes two so called “Inverted repeat Regionen (IR a/b)” These two regions are interrupted by a SSC- (small single copy) and a LSC (large singly copy)-region (Shaw et al., 2007). Taxonomists have used sequence differences in standardised target regions within the chloroplast genome to distinguish among species. The LSC-region has been in the focus of these studies (Taberlet et al. 1991, Timme et al., 2007; Butcher et al., 2007; Feldberg et al., 2007; Tsai et al., 2006).

In many cases the highly variable nuclear micosatellites can be applied for different species within a genus or even within a family (see for example White and Powell 1997). Microsatellites, or Simple Sequence Repeats (SSRs), are polymorphic loci present in nuclear and organelle DNA that consist of repeating units of 1-6 base pairs in length. In most cases there are strong differences of the allele frequencies between species of the same taxonomic genus. Thus multilocus approaches have the potential to assign individuals to a given species (Duminil el al. 2006, Hertel u. Degen 2000). The advantage of this approach is that the classification is not restricted to the pure species approach. Thus even different levels of hybridisation between species can be detected.

In a recent study SNP-markers have been made available to differentiate between different *Populus* species and to detect hybrids between different *Populus* species (Fladung 2006). The "problem" with *Populus* species is the high introgressive hybridisation capacity. The genus *Populus* is divided into five sections. From about 35 poplar species of the northern hemisphere only three species are native to Central Europe: *Populus nigra* (black poplar), *Populus alba* (white poplar) and *Populus tremula* (aspen). Four of the five sections of *Populus* are represented in North America which are important for gene introgression studies when species are transferred to Central Europe because members of these sections are freely interbreeding with European poplar species. Introgressive hybridisation and gene flow from domesticated poplar species into their wild relatives can have a profound effect on the persistence and evolution of wild populations. Today, particularly the European black poplar (*P. nigra)* is a threatened species mainly caused by gene introgression from the North-American cottonwoods *P. deltoides* and *P. trichocarpa* (Vanden Broeck et al. 2004; EUFORGEN, www.ipgri.cgiar.org/networks/euforgen/euf_home.asp).

However, the risk of loosing species purity through hybridisation concerns not only black poplar, but also the two other Central Europe poplar species with wide distribution, aspen and white poplar.

Single nucleotide polymorphisms (SNPs) are DNA sequence variations at the DNA level occurring when a single nucleotide differs at a homologous position between members of a species, genus or family. SNPs are more easily detected since fast and high-throughput sequencing and genotyping techniques are available. In the last decade SNPs have been used for widespread applications, for example, in the construction of high-density genomic maps (Cho et al. 1999), in gene detection and identification, and the analysis of the genetic structure of populations (Garcia-Gil et al. 2003). Further, genetic diversity within a species which is believed to be fundamental for the ability of individuals to adapt to different/changing environments is most commonly realized on the basis of nucleotide differences. For the human genome, for instance, it has been estimated that SNPs occur in at least 10 million nucleotide positions (Wang et al. 1998), in mean every 200 to 300 bp a SNP.

In a pilot study the gene for polyphenoloxidase (PPO) was used to distinguish between *P. tremula* and *P. tremuloides*. In the EMBL-data bank two sequences from *Populus* species are available. Both sequences are 1689 bp long but differ in 66 SNPs. A fragment of 834 bp of the PPO gene was sequenced of six to 13 different genotypes per *Populus* species and "consensus" sequences specific for each species were designed. Based on these results the two aspen species could clearly be differentiated: at position 381 and 411, *P. tremula* carries a "T" and a "G" instead of a "C" and an "A" in *P. tremuloides*, respectively (red letters in Table 1, Figure 1). The interspecific hybrid of the two species is heterozygous in the two loci (Figure 1).

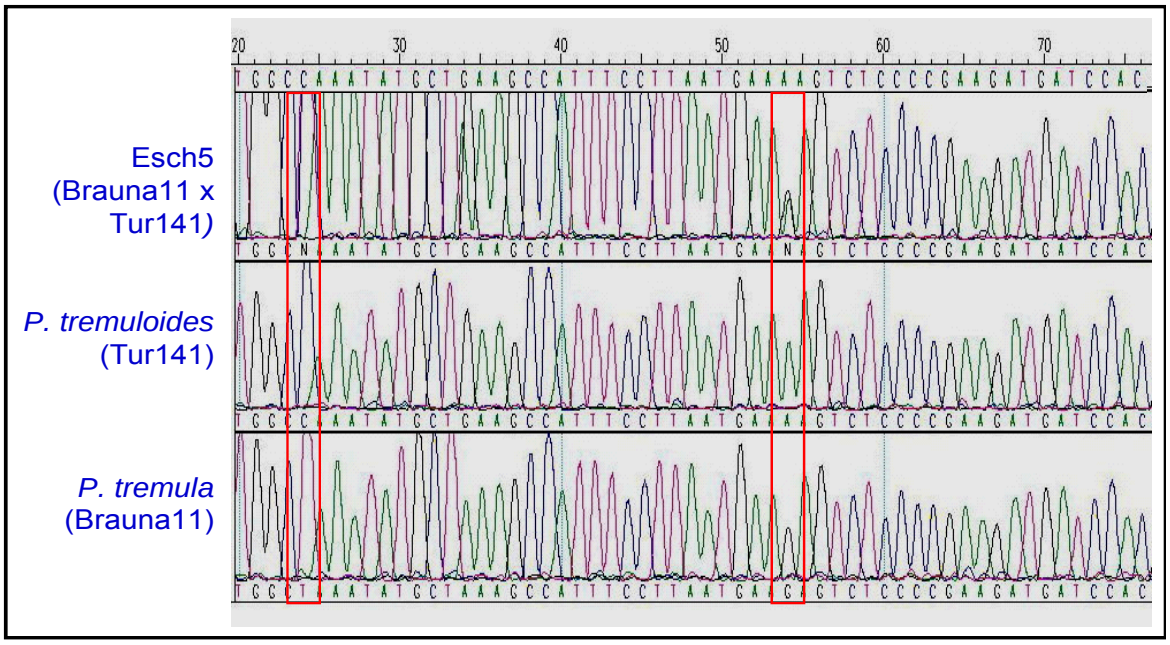

Figure 1:
Differentiation of the two aspen species by SNPs in partial sequences of the polyphenoloxidase (PPO) gene in *P. tremula* (line Brauna11) and *P. tremuloides* (line Tur141) and an interspecific hybrid (line Esch5)

Analyzing the same PPO fragment *P. alba* (green letters in Table 1), *P. nigra* (blue letters in Table 1), *P. deltoides* (brown letters in Table 1) can be distinguished from *P. trichocarpa* as well as from the other *Populus* species. A similar approach has already been performed with five more genes (Fladung and Buschbom, unpublished).

Table 1:
Partial "consensus" nucleotides of the PPO gene of six different *Populus* species. For each species six to 13 different trees have been sequenced. The sequence of *P. trichocarpa* serves as a control

	381	411	405	673	444	781	342	511	378	424
P. trichocarpa	C	A	A	G	C	C	T	G	C	A
P. tremula	**T**	**G**	A	G	C	C	T	G	G	G
P. tremuloides	**C**	**A**	A	G	C	C	T	G	G	G
P. alba	C	A	**G**	**A**	C	C	T	G	G	G
P. nigra	C	A	A	G	**T**	**A**	T	G	G	G
P. deltoides	C	A	A	G	C	C	**A**	**A**	G	G
	381	411	405	673	444	781	342	511	378	424

2.3. Controlling the geographic origin within a species

In natural forests usually a genetic structure at local and regional spatial scales can be observed. In temperate forests as well as in tropical forests, the glacial periods changed the vegetation drastically. In the temperate zone large areas were covered by ice and were free of any vegetation and in the tropics former rain forests were transformed to dry savannas. After each glacial period trees recolonised their distribution area starting from different refugia. As a result of this recolonisation in many cases a clear genetic differentiation can be identified between tree populations from different regions. The extent of genetic differences depends on the recolonisation routes and the genetic differences in these refugia. Chloroplast gene markers and nuclear microsatellites have been successfully used to elaborate reference data about this genetic differentiation (Petit et al. 1997, Caron et al. 2000, Dutech et al. 2000, 2003, Koenig et al. 2003). Figure 2 shows one example for the tropical tree species *Carapa guianensis* in the Amazon (Cloutier et al. 2005). Limited pollen and seed dispersal are the main factors casing spatial genetic structure on smaller scales in natural tree populations (Degen et al. 2001, 2004; Hardy et al. 2006). The spatial resolution of a possible control of timber origin depends on local and regional genetic differences and on the number of sampled populations used to generate a genetic reference data base (Degen et al. 2001, Cavers et al. 2005). There are already a few good data sets for the spatial genetic structure of tree populations on a large scale with a spatial resolution of 50 to 200 km (Caron et al. 2000, Dutech et al. 2000, 2003). And even more studies have found spatial genetic structure on very small scales with a resolution of a few meters to a few hundred meters (Degen et al. 2001, figure 3, 4).

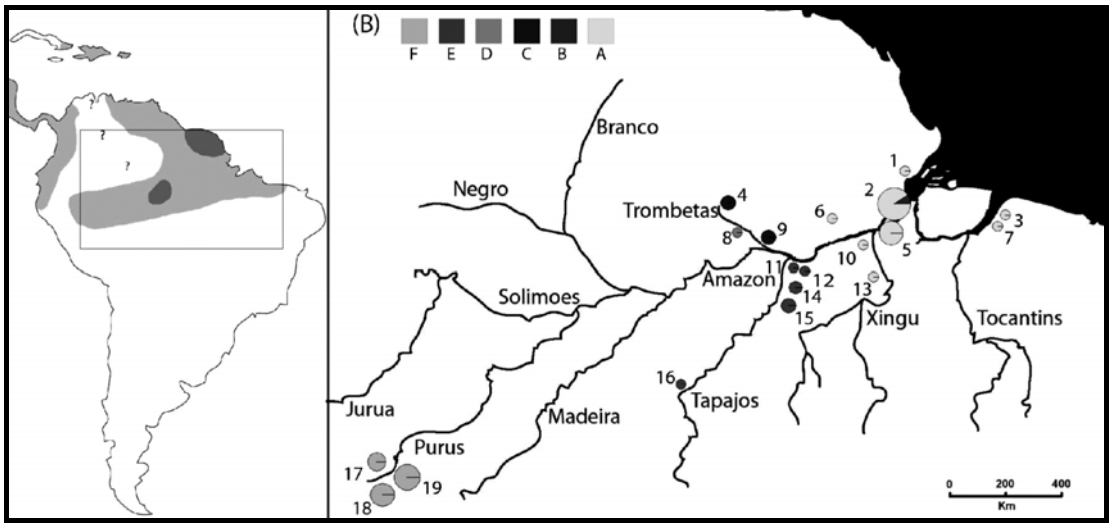

Figure 2:
(a) Geographic range of *Carapa guianensis* (light grey) and areas where both *C. guianensis* and *C. procera* are found (dark grey). (b) Map of the 19 locations sampled for chloroplast DNA variation in *Carapa guianensis* in the Amazon basin. Each location is represented by a numbered circle of size proportional to the number of individuals sampled, and each haplotype is represented by a different colour (Cloutier et al. 2005)

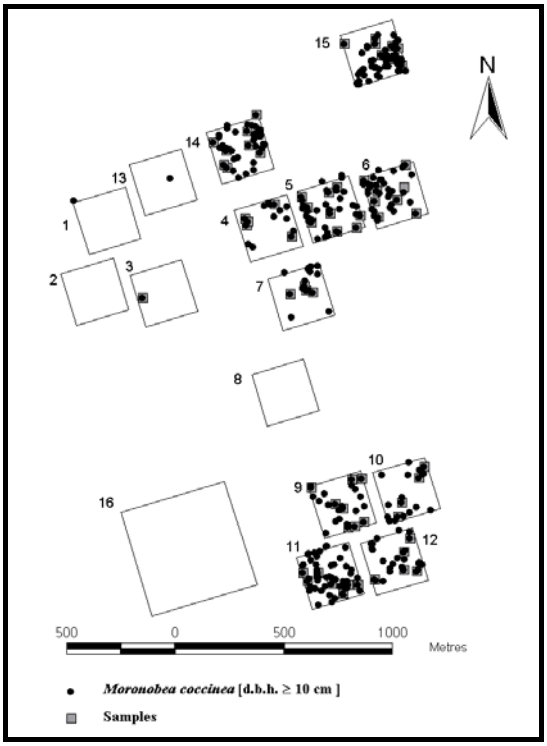

Figure 3:
Position of all trees of the tropical tree species *Moronobea coccinea* [d.b.h. > 10 cm] and position of trees that have been sampled for genetic studies in the experimental trial Paracou in French Guiana (Degen et al. 2001)

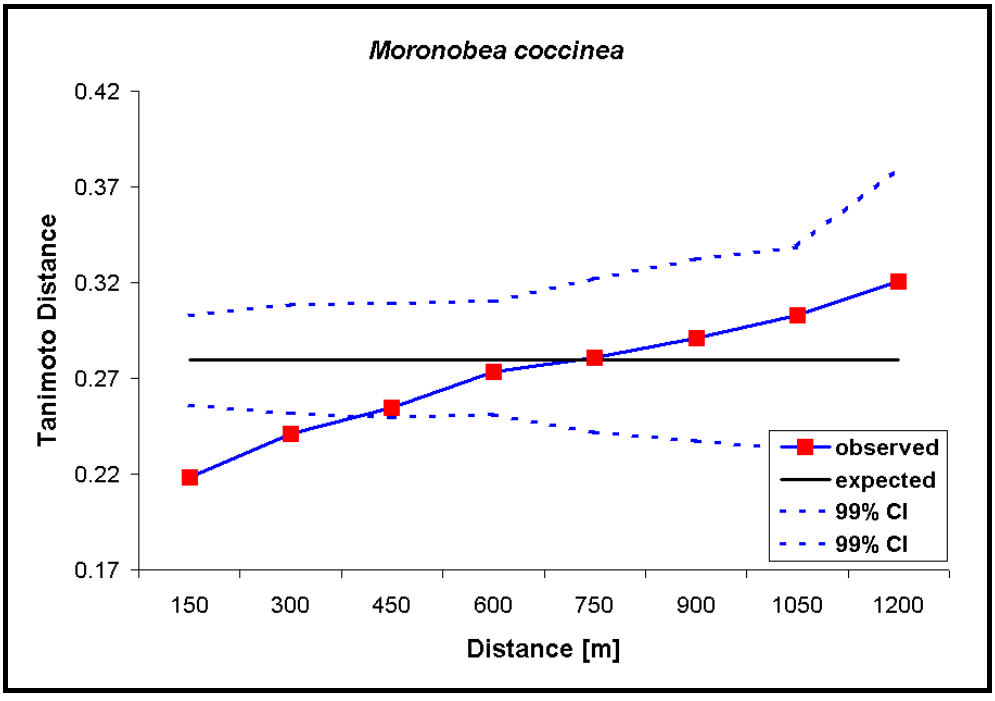

Figure 4:
On a scale of up to 1200m increasing genetic distance (Tanimoto) is detected with increasing spatial distance among Monorobea trees in Paracou in French Guiana (Degen et al. 2001)

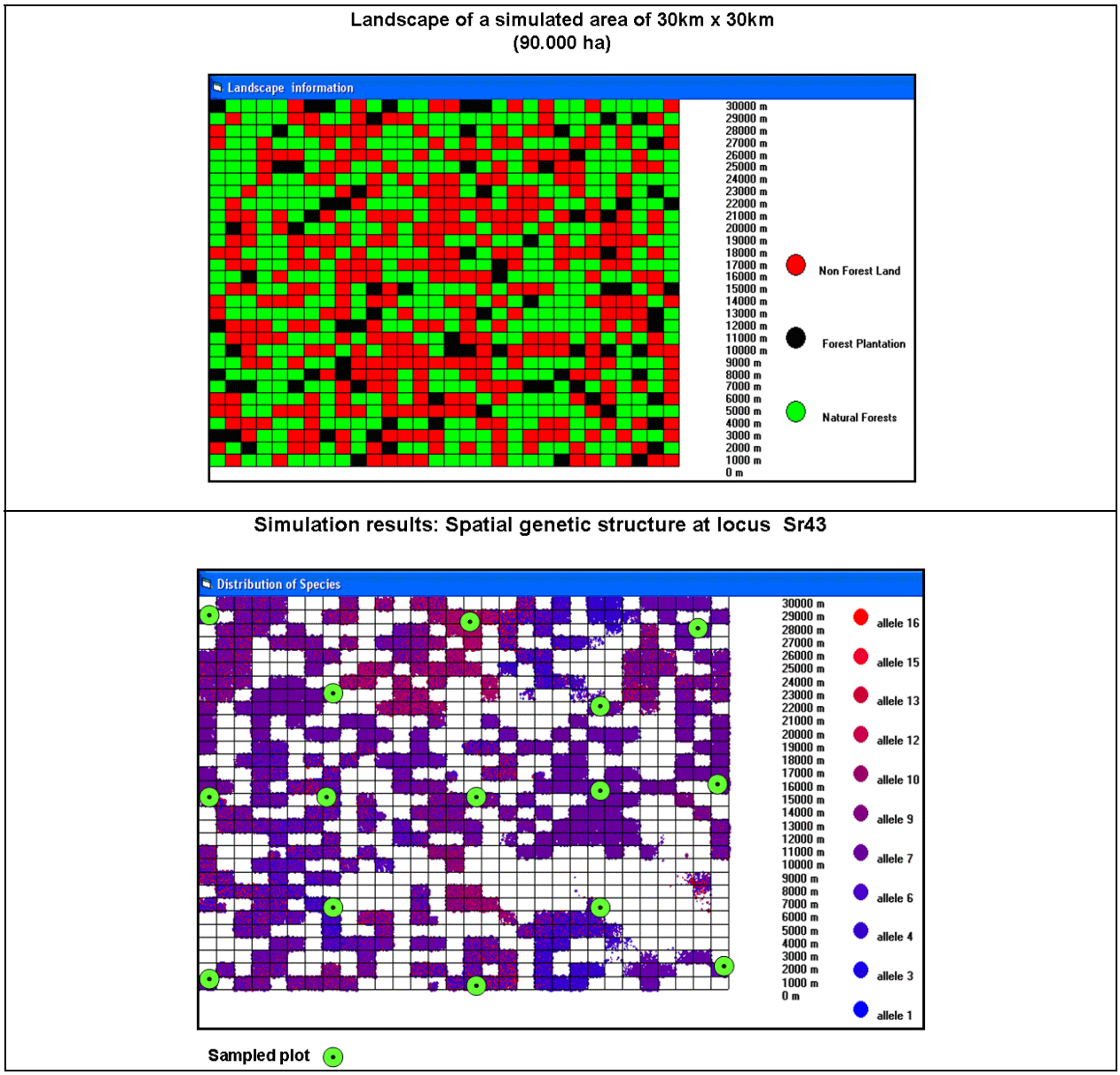

Figure 5:
a): Distribution of natural forest, forest plantation and non forest land in a large scale simulation with the computer model Eco-Gene. The simulated area covers 90.000 ha. Each grid cell represents 100 ha. (b) After 3000 simulated years clear genetic differences (illustrated for one locus as different colours) were observed. Based on this spatial genetic structure at several gene loci a distribution of 15 sample plots along transects (green circles) is expected to be a good starting point for successful assignment of randomly selected individuals to the next sample plot.

So far missing are results on genetic patterns at spatial scales that can be used to control the origin of timber from a logging concession with a usual size of several thousand hectares. We used the simulation model Eco-Gene (Degen et al. 2006) to run first simulations on the expected spatial genetic pattern for a tropical tree species at a scale of 30 km by 30 km (figure 5). The simulations included typical allele frequencies for nuclear microsatellites, densities of a commercial tropical tree species, diameter distribution and important population genetic processes like re-colonisation, pollen and seed dispersal. The simulation results suggested that there is a correlation between genetic and geographic distance up to a few kilometres. In these cases genetic assignment of unknown individuals has a chance of success if there is a sampled population in the reference data base every 5 km (figure 5b).

Here we need more simulation studies to consider the special situation of each timber species and timber concession.

The power of genetic fingerprinting to trace the geographic origin of timber depends on the spatial genetic pattern in the natural tree populations, the sample design for the genetic reference data base and the possibility to amplify the gene markers from DNA extracted from wood.

3 References

Asif MJ, Cannon CH (2007) DNA extraction from processed wood: a case study for the Identification of an endangered timber species (Gonystylus bancanus). Plant Mol Biol Report 23:185-192

Barker NP, Galley C, Verboom GA, Mafa P, Gilbert M, Linder HP (2007) The phylogeny of the austral grass subfamily Danthonioideae: Evidence from multiple data sets . Plant Sys Evol 264:135-156

Caro H, Dumas S, Marque G, Messier C, Bandou E, Petit RJ, Kremer A (2000) Spatial and temporal distribution of chloroplast DNA polymorphism in a tropical tree species. Mol Ecol 9:1089-1098

Cavers S, Degen B, Caron H, Lemes MR, Margis R, Salgueiro F, Lowe AJ (2005) Optimal sampling strategy for estimation of spatial genetic structure in tree populations. Heredity 95:281-289

Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, et al (1999) Genome-wide mapping with biallelic markers in Arabidopsis thaliana. Nature Genetics 23:203-207

Cloutier D, Povoa JSR, Procopio LC, Leao NVM, Wadt LHO, Ciampi AY, Schoen DJ (2005) Chloroplast DNA variation of Carapa guianensis in the Amazon basin. Silvae Genet 54:270-274

De Filippis L, Magel E (1998) Differences in genomic DNA extracted from bark and from wood of different zones in Robinia trees using RAPD-PCR. Trees 12:377-384

Degen B, Bandou E, Caron H (2004) Limited pollen dispersal and biparental inbreeding in Symphonia globulifera in French Guiana. Heredity 93:585-591

Degen B, Caron H, Bandou E, Maggia L, Chevallier MH, Leveau A, Kremer A (2001) Fine-scale spatial genetic structure of eight tropical tree species as analysed by RAPDs. Heredity 87:497-507

Degen B, Petit R, Kremer A (2001) SGS - Spatial genetic software: A computer program for analysis of spatial genetic structures of individuals and populations. J Hered 92:447-448

Deguilloux MF, Pemonge MH, Petit RJ (2002) Novel perspectives in wood certification and forensics: dry wood as a source of DNA. Proc R Soc Lond, Ser B-Biol Sci 269:1039-1046

Deguilloux MF, Pemonge MH, Petit RJ (2004) DNA-based control of oak wood geographic origin in the context of the cooperage industry. Ann For Sci 61:97-104

Duminil J, Caron H, Scotti I, Cazal SO, Petit RJ (2006): Blind population genetics survey of tropical rainforest trees. Mol Ecol 15:3505-3513

Dutech C, Maggia L, Joly HI (2000) Chloroplast diversity in Vouacapoua americana (Caesalpinniaceae), a neotropical forest tree. Mol Ecol 9:1427-1432

Dutech C, Maggia L, Tardy C, Joly HI, Jarne P (2003) Tracking a genetic signal of extinction-recolonization events in a neotropical tree species: Vouacapoua americana aublet in french guiana. Evolution 57:2753-2764

Erikson T, Hibbs MS, Yodrer AD, Delwiche CF, Donoghues MJ (2003) The phylogeny of rosideae (Rosaceae) based on sequences of the internal transcribed spcers (ITS) of nuclear ribosomal DNA and the TRNL/F region of chloroplast DNA. Int. J. Plant Sci. 164:197-211

Fladung M (2006) Entwicklung von SNP-Markern in putativ anpassungsrelevanten Genen. Vortr Pflanzenzüchtg Göttingen 70:139-146

Fladung M, Buschbom J (submitted) Identification of Single Nucleotide Polymorphisms (SNPs) in different Populus species.

Fladung M, Nowitzki O, Ziegenhagen B, Markussen T (2004) Identification of transgenes from wood of genetically transformed poplar trees. Wood Sci Technol 38:207-215

Garcia-Gil MR, Mikkonen M, Savolainen O (2003) Nucleotide diversity at two phytochrome loci along a latitudinal cline in Pinus sylvestris. Mol Ecol 12:1195–1206

Hardy OJ, Maggia L, Bandou E, Breyne P, Caron H, Chevallier MH, Doligez A, Dutech C, Kremer A, Latouche-Halle C, Troispoux V, Veron V, Degen B (2006) Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. Mol Ecol 15:559-571

Hebert PDN, Cywinska A, Ball SL, de Waard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond, Ser B-Biol Sci 270:313-321

Hebert PDN, Gregory, TR (2005) The promise of DNA barcoding for taxonomy. Syst Biol 54:852-859

Hertel, H, Degen B (2000) Unterscheidung von Stiel- und Traubeneichen (Quercus robur L. und Q. petraea [Mattuschka] Liebl.) mit Hilfe von genetischen und morphologischen Merkmalen. For Snow Landsc Res 75:169-183

Indrioko S, Gailing O, Finkeldey R (2006) Molecular phylogeny of Dipterocarpaceae in Indonesia based on chloroplast DNA. Plant Syst Evol 261:99-115

Kencier GJ, Kajita T, Pennington RT, Murata J (2005) Systematics and biogeography of Lathyrus (Leguminosae) based on internal transcribed spacer and cpDNA sequence data. Am. J. of Bot. 92:1199-1209

Konig AO, Ziegenhagen B, Dam BCV, Csaikl UM, Coart E, Degen B, Burg K, Vries SMGD, Petit RJ (2002) Chloroplast DNA variation of oaks in western Central Europe and genetic consequences of human influences. For Ecol Manage 156:147-166

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of barcodes to identify flowering plants. Proc. Natl. Acad. Sci. USA 102:8369-8374

Lemes MR, Gribel R, Proctor J, Grattapaglia D (2003) Population genetic structure of mahogany (Swietenia macrophylla King, Meliaceae) across the Brazilian Amazon, based on variation at microsatellite loci: implications for conservation. Mol Ecol 12:2875-2883

Mort ME, Archibald JK, Randle CP, Levsen ND, O'Leary TR, Topalov K, Wiegand CM, Crawford DJ (2007) Inferring phylogeny at low taxonomic levels: Utility of rapidly evolving cpDNA and nuclear ITS loci. Am. J. of Bot. 94:173-183

Novick RR, Dick CW, Lemes MR, Navarro C, Caccone A, Bermingham E (2003) Genetic structure of Mesoamerican populations of Big-leaf mahogany (Swietenia macrophylla) inferred from microsatellite analysis. Mol Ecol 12:2885-2893

Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducousso A, Kremer A 1997: Chloroplast DNA footprints of postglacial recolonization by oaks. Proc. Natl. Acad. Sci. USA 94:9996-10001

Rachmayanti Y, Leinemann L, Gailing O, Finkeldey R (2006) Extraction, amplification and characterization of wood DNA from Dipterocarpaceae. Plant Mol Biol Report 24:45-55

Shaw J, Lickey E, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for Phylogenetic studies in angiosperms: The Tortoise and the Hare III. Am J Bot 94:275-288

Syring J, Farrell K, Businsky R, Cronn R, Liston A (2007) Widespread genealogical nonmonophyly in species of Pinus subgenus Strobus. Syst Bio 56:163-181

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105-1109

Vanden Broeck A, Storme V, Cottrell JE, Boerjan W, Van Bockstaele E, Quataert P, Van Slycken J (2004) Gene flow between cultivated poplars and native black poplar (Populus nigra L.): a case study along the river Meuse on the Dutch–Belgian border. For Ecol Manage 197:307–310

Wang DG, Fan JB, Siao CJ, Berno A, et al (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077-1082

White G, Powell W (1997) Cross-species amplification of SSR loci in the Meliaceae family. Mol Ecol 6:1195-1197

Ziegenhagen B, Fladung M (2004) DNA markers for identification and evaluation of genetic resources in forest trees – Case studies in Abies, Picea and Populus. In: Lörz H, Wenzel G (eds) Molecular Marker Systems in Plant Breeding and Crop Improvement. Springer Verlag. pp 413-429

Ziegenhagen B, Liepelt S, Kuhlenkamp V, Fladung M (2003) Molecular identification of individual oak and fir trees from maternal tissues of their fruits or seeds. Trees 17:345-350