# Innovative timber tracking for Swietenia macrophylla using genetic fingerprints

**CONFERENCE PAPER** · AUGUST 2010



**5 AUTHORS**, INCLUDING:

Bernd Degen
Thünen Institute
**60** PUBLICATIONS **1,167** CITATIONS
SEE PROFILE

Sheila Emily Ward
Mahogany for the Future, Inc., San Juan, P…
**12** PUBLICATIONS **87** CITATIONS
SEE PROFILE

Carlos Manuel Navarro
International Analog Forestry Network
**38** PUBLICATIONS **898** CITATIONS
SEE PROFILE

Stephen Cavers
Centre for Ecology & Hydrology
**57** PUBLICATIONS **499** CITATIONS
SEE PROFILE

Forensic Science International: Genetics 7 (2013) 55–62

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Forensic Population Genetics—Original Research

# Verifying the geographic origin of mahogany (*Swietenia macrophylla* King) with DNA-fingerprints

B. Degen [a,*], S.E. Ward [b], M.R. Lemes [c,1], C. Navarro [d], S. Cavers [e], A.M. Sebbenn [f]

[a] Johann Heinrich von Thünen-Institut (vTI), Institut für Forstgenetik, Sieker Landstrasse 2, D-22927 Grosshansdorf, Germany
[b] Mahogany for the Future Inc., San Juan, PR 00928, USA
[c] Laboratório de Genética e Biologia Reprodutiva de Plantas (LabGen), Instituto Nacional de Pesquisa da Amazônia (INPA), Av. André Araújo 2936, CEP 69083-000 Manaus, AM, Brazil
[d] Universidad Nacional, Calle 9, Avenidas 0 y 9, Costa Rica
[e] Centre for Ecology & Hydrology, Penicuik, Midlothian, Scotland EH26 0QB, United Kingdom
[f] Instituto Florestal de São Paulo, CP 1322, São Paulo, SP 01059-970, Brazil

ARTICLE INFO

*Article history:*
Received 13 March 2012
Received in revised form 18 May 2012
Accepted 3 June 2012

*Keywords:*
Illegal logging
Forensic
Mahogany
Microsatellite loci
Timber tracking
Tropical trees

ABSTRACT

Illegal logging is one of the main causes of ongoing worldwide deforestation and needs to be eradicated. The trade in illegal timber and wood products creates market disadvantages for products from sustainable forestry. Although various measures have been established to counter illegal logging and the subsequent trade, there is a lack of practical mechanisms for identifying the origin of timber and wood products. In this study, six nuclear microsatellites were used to generate DNA fingerprints for a genetic reference database characterising the populations of origin of a large set of mahogany (*Swietenia macrophylla* King, Meliaceae) samples. For the database, leaves and/or cambium from 1971 mahogany trees sampled in 31 stands from Mexico to Bolivia were genotyped. A total of 145 different alleles were found, showing strong genetic differentiation ($\delta_{Gregorious} = 0.52$, $F_{ST} = 0.18$, $G_{ST(Hedrick)} = 0.65$) and clear correlation between genetic and spatial distances among stands ($r = 0.82$, $P < 0.05$). We used the genetic reference database and Bayesian assignment testing to determine the geographic origins of two sets of mahogany wood samples, based on their multilocus genotypes. In both cases the wood samples were assigned to the correct country of origin. We discuss the overall applicability of this methodology to tropical timber trading.



## 1. Introduction

This study examines a method for identifying the source populations of mahogany (*Swietenia macrophylla*), one of the world's most valuable and widely traded tropical timbers, to the region of origin. Illegal logging and trade in illegal timber and wood products are the cause of many economic and ecological problems in both producer and consumer countries, resulting in the need to track imported timber back to its source. 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 [38,39]. Illegal logging is believed to be one of the chief causes of worldwide deforestation, and trade in illegal timber and wood products creates market disadvantages for products from legal and sustainable forestry. The WWF estimates the global costs of illegal timber at approx. €15 billion per year [http://wwf.panda.org/about_our_earth/about_forests/deforestation/forest_illegal_logging/].

In December 2010, new European Union (EU) timber regulations (No. 995/2010) came into force that will make it illegal, from March 2013, to place illegally harvested timber and timber products on the European market. The new rules prohibit the trade of illegally sourced timber and oblige operators to perform due diligence by seeking guarantees that the timber products they sell have been harvested in a sustainable way and according to the laws of the country of origin. To facilitate this, the EU is negotiating Voluntary Partnership Agreements (VPA) with individual timber-producing countries. VPA countries agree to export to the EU only verified legal timber with a FLEGT (Forest Law Enforcement, Governance and Trade) license. The new EU rules are modelled on similar legislation adopted in the USA in 2008, as an amendment of the Lacey Act. This amended act prohibits in the USA all trade in plants and plant products, including timber and timber products, that are illegally sourced

* Corresponding author at: Thünen-Institute of Forest Genetics (vTI), Sieker Landstrasse 2, D-22927 Grosshansdorf, Germany. Tel.: +49 0 4102 696 0; fax: +49 0 4102 696 200.
*E-mail addresses:* bernd.degen@vti.bund.de (B. Degen), seward@caribe.net (S.E. Ward), mlemes@inpa.gov.br (M.R. Lemes), carlosmanuelnavarro@gmail.com (C. Navarro), scav@ceh.ac.uk (S. Cavers), alexandresebbenn@yahoo.com.br (A.M. Sebbenn).
[1] Present address: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, Rua Pacheco Leão 915, Jardim Botânico, 22460-030 Rio de Janeiro, RJ, Brazil.

1872-4973/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigen.2012.06.003

from any US state or foreign country, and requires importers to declare the country of harvest and the species name of all plants contained in their products. In addition to legislation that targets the illegal timber trade, the Convention on International Trade in Endangered Species (CITES) aims to regulate international trade in selected animal and plant species to help ensure their survival. Selected species are listed in one of three appendices, depending on the degree of endangerment. With appropriate controls in place, commercial trade is allowed from natural and plantation sources of species listed in Appendix II, such as mahogany species (*Swietenia* spp.). Since August 2010 the EU has prohibited all imports of *S. macrophylla* from Bolivia and requested special permissions for imports from all other countries except Mexico and Guatemala. Because of problems with identification, CITES requires trade in look-alike species to be regulated as well. In combination, these regulations are intended to increase pressure on forest concession holders and timber traders to use credible methods to verify the source of their timber.

To meet these objectives, practical mechanisms are needed for identifying the origin of timber and wood products. Falsification of the country of origin is a well-documented problem in the international tropical timber trade. Typically, falsification occurs at the point of import and usually involves the production of false phytosanitary certificates (which verifies inspection and states that a specified disease was not found or does not occur in the area of production), invoices, or certificates of origin. Existing timber tracking systems use external electronic tags or paper-based documentation to trace back the timber origin, both vulnerable to falsification. Another common problem at smaller spatial scales is the false declaration of timber that has been logged outside a registered concession or within a protected area. Thus tools that can validate the point of harvest are also of interest to legitimate enterprises.

In this study, we test a tracking method based on DNA profiling for mahogany (*S. macrophylla*). DNA fingerprints are inherent natural characteristics of the wood itself and are therefore not falsifiable. We describe the development of a geo-referenced genetic reference database for *S. macrophylla* that allows wood samples to be traced back to their geographical regions of origin, following genotyping of the sample. The database uses a DNA profile based on six nuclear microsatellite loci (nSSRs) and draws on samples collected from locations across Latin America. The resolution and performance of the database were assessed by blind testing of two sets of mahogany wood samples using a multilocus assignment procedure. The genetic diversity, differentiation and the spatial genetic pattern of the genetic reference dataset were characterised and the statistical power of the assignment test was estimated.

## 2. Materials and methods

### 2.1. Species description

Big leaf mahogany (*S. macrophylla* King, Meliaceae) is the most valuable commercial hardwood species from the Neotropics. The species has a wide geographic range from Mexico through Central America and across the southern Amazon basin of Bolivia and Brazil [1]. It is a monoecious species with insect-pollinated flowers (bees, moths, and thrips [2]), wind-dispersed seeds [3] and a predominantly outcrossed mating system [4]. It can reach 40 m in height and is considered fast-growing but plantations are often inhibited by the larvae of *Hypsipyla grandella*, which destroy apical meristems.

Since the late 1700s, *S. macrophylla* has been heavily exploited [3,5] and today is threatened throughout its natural range. In Central America populations have declined by 80% in the last 50 years and some are already extinct [6]. In 2003, because of the high risks to population viability associated with over-exploitation and habitat destruction, *S. macrophylla* was listed on CITES Appendix II [7].

From previous studies using nuclear microsatellites, *S. macrophylla* has been shown to have high levels of genetic diversity and significant genetic differentiation among populations in both the Brazilian Amazon [8] and Central America [9,10]. More recently, phylogeographic analysis using chloroplast microsatellite markers found contrasting patterns of genetic structure between Central and South American populations [11].

## 3. Sampling

### 3.1. Samples for genetic reference database

Genotype data for this study was gathered from a total of 1971 individuals of *S. macrophylla* from 31 populations in 9 countries across Central and South America (Table 1, Fig. 1). These data were used to construct a genetic reference database (Table 1, Fig. 1). On average 63 (range of 24–400) individuals per population were sampled.

In all collections, leaves were sampled, apart from those from Belize, from which cambium tissue was sampled following [12].

### 3.2. Wood samples for testing

Two wood collections declared as *S. macrophylla* were used to validate the DNA profiling method on sawnwood. One was obtained from a timber trader in Hamburg and was comprised of end pieces from each of 20 wooden planks. The second was sent from Bolivia and contained 12 wood pieces from five small joineries in Ascención de Guarayos, Santa Cruz State, Bolivia, about 30 km from La Chonta concession.

## 4. Genetic analysis

### 4.1. DNA extraction

The genetic analyses were carried out at two laboratories: the Brazilian samples were analysed at the Plant Genetics Laboratory at INPA, Manaus, AM, Brazil and all other samples at vTI, Grosshansdorf, Germany. At INPA, total genomic DNA was extracted from leaves using a standard CTAB procedure [13]. At vTI, total genomic DNA was extracted from leaves and cambium using the ATMAB procedure [14]. DNA extraction of wood samples was carried out at vTI using the DNeasy Qiagen Plant Mini Kit as modified by Rachmayanti et al. [15].

### 4.2. Nuclear microsatellite analysis

All samples were genotyped at six previously developed, highly polymorphic nuclear microsatellite loci (sm01, sm22, sm31, sm32, sm40, and sm51) [16].

At INPA, microsatellite amplification was carried out in a final volume of 10 μL for single or 25 μL for multiplexed reactions, each containing 1.25–2.0 μM of each forward and reverse primer, 1 unit *Taq* DNA polymerase, 200 μM of each dNTP, 1× reaction buffer (10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$), BSA (2.5 mg/mL), 5.0 ng of template DNA, and ultrapure water. PCRs were performed in a MJ Research Incorporated PTC 200 thermal cycler under the following conditions for all loci: 96 °C for 2 min, followed by 30 cycles of 94 °C for 1 min, 56 °C for 1 min, and 72 °C for 1 min, and a final elongation step at 72 °C for 30 min. PCR products were added to GeneScan 500 ROX internal size standard (ABI, Inc.), electrophoresed on a 5% denaturing polyacrylamide gel in an ABI

**Table 1**
Location and sample size of *Swietenia macrophylla* populations used in the study; sample campaign (see description below).

| Sort | Country | Population | Sample campaign | Latitude | Longitude | Sample size |
|---|---|---|---|---|---|---|
| 1 | Bolivia | Bolpebra | 1 | −11.11 | −69.06 | 92 |
| 2 | Bolivia | Jorge Cruz | 1 | −17.98 | −63.47 | 103 |
| 3 | Bolivia | La Chonta_0 | 1 | −15.73 | −62.85 | 59 |
| 4 | Bolivia | La Chonta_1 | 5 | −15.78 | −62.92 | 80 |
| 5 | Bolivia | La Chonta_2 | 5 | −15.78 | −62.92 | 66 |
| 6 | Bolivia | San Borja | 1 | −15.01 | −67.10 | 69 |
| 7 | Bolivia | Sapecho | 1 | −15.63 | −67.17 | 52 |
| 8 | Bolivia | Yotau | 1 | −16.10 | −62.98 | 51 |
| 9 | Costa Rica | La Cruz | 1 | 11.07 | −85.49 | 50 |
| 10 | Costa Rica | Los Chiles | 1 | 10.96 | −84.74 | 51 |
| 11 | Costa Rica | Pocosol | 1 | 10.53 | −85.36 | 49 |
| 12 | Costa Rica | Sardinal | 1 | 10.09 | −84.83 | 52 |
| 13 | Guatemala | Peten | 2 | 17.00 | −89.66 | 56 |
| 14 | Honduras | Colon | 2 | 15.83 | −86.00 | 57 |
| 15 | Mexico | Veracruz | 2 | 17.66 | −94.91 | 98 |
| 16 | Mexico | Chiapas | 2 | 16.91 | −91.58 | 52 |
| 17 | Mexico | Campeche | 2 | 18.41 | −89.41 | 59 |
| 18 | Mexico | Quintana Roo | 2 | 19.50 | −88.08 | 50 |
| 19 | Nicaragua | Central | 2 | 13.66 | −84.16 | 53 |
| 20 | Nicaragua | Terra | 2 | 11.83 | −83.91 | 53 |
| 21 | Panama | Herrera | 2 | 7.58 | −80.66 | 26 |
| 22 | Panama | Darien | 2 | 8.83 | −78.00 | 37 |
| 23 | Belize | East Botes | 3 | 17.64 | −88.72 | 42 |
| 24 | Belize | Punta Gorda | 3 | 17.62 | −88.68 | 49 |
| 25 | Brazil | Agua Azul, PA | 4 | −6.9 | −50.27 | 31 |
| 26 | Brazil | Cach A, RO | 4 | −12.5 | −61.5 | 32 |
| 27 | Brazil | Pontes e Lacerda, MT | 4 | −15.08 | −59.15 | 24 |
| 28 | Brazil | Chico Mendes, AC | 4 | −10.42 | −69.3 | 34 |
| 29 | Brazil | Cach E, RO | 4 | −12.57 | −61.5 | 24 |
| 30 | Brazil | Pimenta Bueno, RO | 4 | −12.37 | −61.43 | 24 |
| 31 | Brazil | Marajoara, PA | 4 | −7.83 | −50.27 | 400 |

*Sample campaigns*:
(1) Samples provided from germplasm collection completed in 2000 and 2009 in collaboration with CIAT Bolivia and from a germplasm bank established in the forestry concession of Agroindustrial Forestal La Chonta.
(2) A total of 541 samples from 10 populations were collected from genetic trials in Puerto Rico. These trials had been established in 1965 and 1966 from seeds collected from populations ranging from Mexico to Panama.
(3) Samples from Belize collected during EU-INCO project contract ICA4-CT-2001-10101.
(4) Data was previously obtained from population genetic studies carried out in Brazil [8,4].
(5) Two collections of mahogany from Bolivia collected in a sampling mission in 2010.

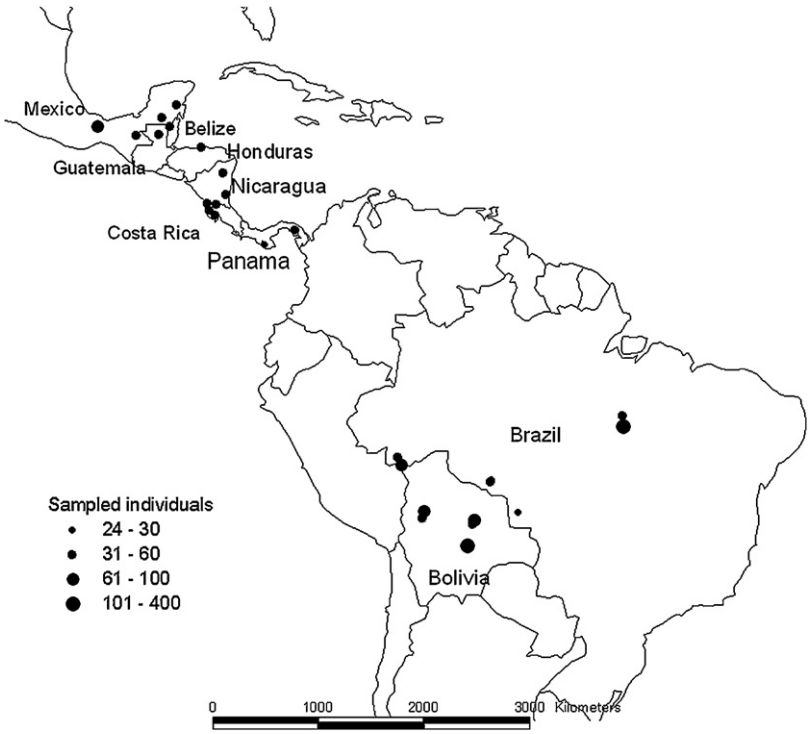


**Fig. 1.** Map showing the spatial position of the populations of *Swietenia macrophylla*. Size of the population marker reflects the sample size for each population.

Prism 377 sequencer (ABI Inc.) and analysed with Genescan and Genotyper software (ABI, Inc.).

At vTI, PCR was carried out in a final volume of 15 μL for both single and multiplexed reactions, containing 0.1–0.5 μM of each forward and reverse primer, 0.6 unit of Taq DNA polymerase, 200 mM of each dNTP, 1× reaction buffer (80 mM Tris–HCl, 20 mM $(NH_4)_2SO_4$, 0.02% Tween-20), 1.75 mM $MgCl_2$, 20 ng of template DNA, and ultrapure water. Amplifications were performed using a thermal cycler from Sensoquest using the following program: 94 °C for 3 min, followed by 30 cycles of 94 °C for 30 s, 54–60 °C (depending on annealing temperature of the primers) for 30 s, and 68 °C for 2 min, and a final elongation step at 68 °C for 15 min. PCR products were combined with 8 μL diluted size standard ET-400 ROX (GE Healthcare). Separation and detection of the fragments were carried out on a MegaBACE 1000 96 capillary automated sequencer (GE Healthcare) using Fragment Profiler software version 1.2 (GE Healthcare).

Genotyping of the wood samples was carried out in single PCR reactions with a final volume of 15 μL containing 0.1–0.5 μM of each forward and reverse primer, 0.375 unit Amplitaq Gold 360 DNA polymerase, 200 mM of each dNTP, 1× reaction buffer, 1.75 mM $MgCl_2$), 30 ng of template DNA, and ultrapure water. Amplifications were also performed using a thermal cycler from Sensoquest using the following program for all loci: initial denaturation at 95 °C for 10 min, followed by 45 cycles of 95 °C for 30 s, 54–60 °C for 30 s, 68 °C for 2 min and a final elongation step at 68 °C for 15 min.

### 4.3. Calibration

A common set of four individuals of *S. macrophylla* were genotyped for the six nuclear microsatellite loci in both INPA and in vTI laboratories and the results were used to calibrate the genotypes obtained in the two labs. In most cases the fragment sizes were exactly the same within 1 bp. For two loci there was a difference of one base pair and the data were adjusted by a systematic shift of one base pair for all data obtained in the vTI lab.

## 5. Data analysis

### 5.1. Genetic diversity and genetic differentiation

Using the program GDA_NT for each locus, the number of alleles ($N_A$), allele frequencies ($p_{ij}$), the observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), and the fixation index ($F = 1 - (H_o/H_e)$) were calculated as described by Weir [17]. Wright's $F_{ST}$ [17,28], the standardised $G_{ST(Hedrick)}$ [18], and delta [19] were computed as measures of fixation and genetic differentiation among populations. Genetic distance ($D_0$) was computed to measure the genetic differentiation between pairs of populations [20]. Numerical tests based on Monte Carlo methods were used to estimate the significance of $F$, delta, $F_{ST}$, $F_{ST(Hedrick)}$ and $D_0$ [21]. For all pairwise genetic distance ($D_0$) estimations, the genotypes in the two populations were pooled 10,000 times and from this joint group, the two individual populations were reformed by random sampling without replacement. The $D_0$ values were recalculated for both re-sampled populations ($i$ and $j$). For delta, $F_{ST}$, and $G_{ST(Hedrick)}$, the genotypes of the all populations were pooled 10,000 times and from this joint group the 31 (population-level) or 9 (country-level) individual populations were reformed by random sampling without replacement. The differentiation values were recalculated for all re-sampled populations. For each permutation, the genotypic structure of each population represented a random mixture of genotypes from all original populations. The probability of significance was estimated from the proportion of permutations with $D_0$, delta, $F_{ST}$ and $F_{ST(Hedrick)}$ that were smaller than the observed value. We calculated Spearman's rank coefficient (a non-parametric measure of statistical dependence between two variables) between the geographic and spatial distances among the populations.

### 5.2. Genotype assignment

Individual assignment tests and group assignment tests were performed using the Bayesian multilocus-approach [22] in Geneclass2 [23] and GDA_NT (Degen, unpublished). In addition to the parameters calculated with Geneclass 2, GDA_NT computes the probability of exclusion for group assignments. All individuals from the reference data were self-classified to the sampled populations using the leave-one-out approach (self-assignment) [24]. The group was compared with each reference population and assigned to the most probable. There was a most likely population in any reference set, to which the group could always be assigned. However, the set of reference populations might not include the true population of origin for the control group, resulting in a false positive assignment. Therefore, a measure of confidence was needed that the tested individual or tested group truly belonged to a given population [25]. This was achieved by comparing the likelihood value of the test group with the likelihood distribution of groups based on allele frequencies of the population and that were generated from 10,000 sampling runs with replacement. If the observed likelihood of the test group was well outside the 99.9% confidence interval of the distribution then this indicated that the group did not belong to the population.

As proposed by Marshall et al. [26] for paternity analysis, we used the difference in LOD scores (delta LOD) between the most-likely reference population and the next most-likely reference population as a test criterion for the exclusion probability [27]. In cases where the true population of origin was among the reference samples this difference was clearly larger than when the true population was not part of the reference samples. Again we used the distribution of delta LOD values generated from 10,000 re-sampling runs to locate the observed values.

## 6. Results

### 6.1. Genetic diversity

The mean number of alleles per nSSR locus over all populations was 9.13. We observed a total of 145 different alleles. Over all loci, mean expected heterozygosity ($H_e$) ranged from 0.61 to 0.83 among the 31 populations (Table 2). There was no significant correlation between the $H_e$ values and the sample size (Spearman's rank correlation = 0.068, $t$-test value for hypothesis $r = 0$ is 0.421, probability $> t = 0.6762$), indicating that this is a robust measure of genetic diversity insensitive to sample size differences. The Brazilian populations had relatively high values of $H_e$ while the populations from Nicaragua and Honduras had smaller values. As indicated by significant positive fixation indices, the number of homozygotes in 11 out of 31 populations exceeded Hardy–Weinberg Equilibrium (HWE) expectations, two populations (San Borja in Bolivia, and Colon in Honduras) had a significant excess of heterozygotes, and 17 populations had no significant departure from HWE (Table 2). Aggregating populations at the country level resulted in significant homozygote excess in six out of 9 cases, with one significant negative $F$-value and only two non-significant values.

### 6.2. Genetic differentiation and spatial genetic structure

All three measures of genetic differentiation and fixation – delta, $F_{ST}$ and $G_{ST(Hedrick)}$ – had very high and significant values for

**Table 2**
Average expected heterozygosity ($H_e$) average fixation index ($F$), probability of departure from Hardy–Weinberg equilibrium (Prob) for each population used in the study and for populations pooled by country.

| Sort | Country | Population | Sample size | $H_e$ | $F$ | Prob |
|---|---|---|---|---|---|---|
| 1 | Bolivia | Bolpebra | 92 | 0.710 | 0.019 | 0.758 |
| 2 | Bolivia | Jorge Cruz | 103 | 0.670 | 0.014 | 0.632 |
| 3 | Bolivia | La Chonta | 59 | 0.736 | 0.032 | 0.806 |
| 4 | Bolivia | La Chonta_1 | 80 | 0.735 | 0.003 | 0.478 |
| 5 | Bolivia | La Chonta_2 | 66 | 0.720 | 0.086 | **0.999** |
| 6 | Bolivia | San Borja | 69 | 0.690 | −0.040 | **0.965** |
| 7 | Bolivia | Sapecho | 52 | 0.672 | 0.036 | 0.792 |
| 8 | Bolivia | Yotau | 51 | 0.655 | 0.095 | **0.990** |
| 9 | Costa Rica | La Cruz | 50 | 0.610 | −0.041 | 0.887 |
| 10 | Costa Rica | Los Chiles | 51 | 0.622 | 0.083 | **0.982** |
| 11 | Costa Rica | Pocosol | 49 | 0.823 | 0.168 | **1.000** |
| 12 | Costa Rica | Sardinal | 52 | 0.642 | 0.155 | **1.000** |
| 13 | Guatemala | Peten | 56 | 0.704 | 0.136 | **1.000** |
| 14 | Honduras | Colon | 57 | 0.668 | −0.098 | **1.000** |
| 15 | Mexiko | Veracruz | 98 | 0.700 | 0.186 | **1.000** |
| 16 | Mexiko | Chiapas | 52 | 0.618 | 0.192 | **0.999** |
| 17 | Mexiko | Campeche | 59 | 0.669 | 0.017 | 0.637 |
| 18 | Mexiko | Quintana Roo | 50 | 0.714 | −0.009 | 0.673 |
| 19 | Nicaragua | Central | 53 | 0.612 | −0.055 | 0.915 |
| 20 | Nicaragua | North | 53 | 0.627 | −0.041 | 0.895 |
| 21 | Panama | Herrera | 26 | 0.611 | 0.206 | **1.000** |
| 22 | Panama | Darien | 37 | 0.722 | −0.044 | 0.904 |
| 23 | Belize | East Botes | 42 | 0.725 | 0.019 | 0.572 |
| 24 | Belize | Punta Gorda | 49 | 0.706 | −0.044 | 0.912 |
| 25 | Brazil | Agua Azul, PA | 31 | 0.806 | 0.015 | 0.568 |
| 26 | Brazil | Cach A, RO | 32 | 0.753 | −0.020 | 0.808 |
| 27 | Brazil | Pontes e Lacerda, MT | 24 | 0.793 | −0.003 | 0.829 |
| 28 | Brazil | Chico Mendes, AC | 34 | 0.706 | 0.040 | 0.817 |
| 29 | Brazil | Cach E, RO | 24 | 0.795 | 0.031 | 0.714 |
| 30 | Brazil | Pimenta Bueno, RO | 24 | 0.731 | 0.135 | **0.995** |
| 31 | Brazil | Marajoara, PA | 400 | 0.837 | 0.073 | **1.000** |
| | Belize | All | 84 | 0.718 | −0.010 | 0.687 |
| | Bolivia | All | 566 | 0.744 | 0.088 | **1.000** |
| | Brazil | All | 565 | 0.852 | 0.101 | **1.000** |
| | Costa Rica | All | 202 | 0.753 | 0.182 | **1.000** |
| | Guatemala | All | 56 | 0.704 | 0.136 | **1.000** |
| | Honduras | All | 57 | 0.668 | −0.098 | **1.000** |
| | Mexico | All | 258 | 0.743 | 0.187 | **1.000** |
| | Nicaragua | All | 106 | 0.628 | −0.035 | 0.937 |
| | Panama | All | 63 | 0.746 | 0.136 | **1.000** |

Bold numbers indicate significant values with a probaility > 0.95

both differentiation among populations and among countries for the nSSR dataset (Table 3). The differentiation among populations within countries was much smaller and only the $G_{ST(Hedrick)}$ for Nicaragua was statistically significant (Table 3).

The plot of pairwise genetic distances against spatial distance (km) among populations showed a clear positive correlation (Fig. 2). The high Spearman's rank correlation of 0.824 ($t$-test value for hypothesis $r = 0$ is 31.248; probability $> t = 0.0000$) indicated that genetic distance among populations increased significantly with spatial distance.

### 6.3. Assignment test

In order to probe the performance of the Bayesian method of [22], we used all six nSSR loci (sm01, sm22, sm40, sm31, sm32, sm51) to perform self-assignment tests for individuals and groups of individuals randomly selected from the reference dataset. When all 31 populations were used as individual reference populations, 70.7% of all individuals were correctly assigned. The expectation for correctly assigning individuals by chance was only 3.2%. When nine reference populations were created by pooling all populations within each of the nine countries, 82.2% of all individuals were assigned correctly. In this case the expectation of correct assignment by chance was 11.1%. The higher proportion of correct assignments with the populations pooled by country indicated that the blind test of the wood samples would be improved slightly using this strategy. The proportion of correctly assigned samples increased greatly when groups rather than single individuals were assigned. Nine groups of two, three, and four randomly selected individuals were correctly self-assigned at 85%, 90% and 95.6%, respectively, to the reference populations pooled by country.

### 6.4. Blind test with wood samples

All subsequent analyses were therefore conducted by assigning groups of individual wood samples from the same provider to one of the nine country-level reference populations. Amplification success for the nuclear microsatellites of the wood samples was lower than for the fresh material, resulting in a fragmented dataset, with 25% of the alleles missing for the wood samples from Hamburg and 35% of the alleles missing for the wood samples from Bolivia.

Alleles for at least one nSSR locus were amplified in 15 out of the 20 samples from the Hamburg trader. In three cases identical genotypes were found. Thus we derived 12 distinct multilocus genotypes for the assignment test. These samples were assigned with a high level of confidence (score = 100%) to the Guatemala reference population (Table 4). Exclusion probabilities based on LOD-values were 100% for all nine countries, but the exclusion probability based on delta LOD-values was not significant for Guatemala. Guatemala was confirmed by the timber trader as the tagged country of origin for the wood samples.

Alleles for at least one nSSR locus were amplified for 9 out of the 11 wood pieces sent from Bolivia. All genotypes were different; so

**Table 3**
Locus-by-locus and mean total genetic differentiation among *Swietenia macrophylla* populations and among groups of samples pooled by country for nSSRs.

| Level of aggregation | Locus | $\delta_{Gregorious}$ | Prob | $F_{ST}$ | Prob | $G_{ST(Hedrick)}$ | Prob |
|---|---|---|---|---|---|---|---|
| Populations ($N$ = 31) | sm01 | 0.4955 | 1.000 | 0.2176 | 1.000 | 0.6531 | 1.000 |
| | sm22 | 0.5456 | 1.000 | 0.2299 | 1.000 | 0.6827 | 1.000 |
| | sm40 | 0.3723 | 1.000 | 0.1255 | 1.000 | 0.4583 | 1.000 |
| | sm31 | 0.5700 | 1.000 | 0.1734 | 1.000 | 0.7093 | 1.000 |
| | sm32 | 0.5854 | 1.000 | 0.2409 | 1.000 | 0.7212 | 1.000 |
| | sm51 | 0.5567 | 1.000 | 0.1514 | 1.000 | 0.6825 | 1.000 |
| | Mean | 0.5209 | 1.000 | 0.1898 | 1.000 | 0.6512 | 1.000 |
| Countries ($N$ = 9) | sm01 | 0.3915 | 1.000 | 0.1138 | 1.000 | 0.4515 | 1.000 |
| | sm22 | 0.4149 | 1.000 | 0.1652 | 1.000 | 0.4907 | 1.000 |
| | sm40 | 0.3037 | 1.000 | 0.0729 | 1.000 | 0.3436 | 1.000 |
| | sm31 | 0.5429 | 1.000 | 0.1303 | 1.000 | 0.6846 | 1.000 |
| | sm32 | 0.4791 | 1.000 | 0.1672 | 1.000 | 0.5318 | 1.000 |
| | sm51 | 0.4746 | 1.000 | 0.1059 | 1.000 | 0.5568 | 1.000 |
| | Mean | 0.4344 | 1.000 | 0.1259 | 1.000 | 0.5098 | 1.000 |
| Belize (2 pops) | Mean | 0.1542 | 0.000 | 0.0059 | 0.331 | 0.0473 | 0.000 |
| Bolivia (8 pops) | Mean | 0.1459 | 0.521 | 0.0079 | 0.405 | 0.0689 | 0.385 |
| Brazil (7 pops) | Mean | 0.2187 | 0.864 | 0.0183 | 0.796 | 0.1563 | 0.829 |
| Costa Rica (4 pops) | Mean | 0.1961 | 0.916 | 0.0113 | 0.913 | 0.1066 | 0.871 |
| Mexico (4 pops) | Mean | 0.1719 | 0.736 | 0.0088 | 0.801 | 0.0857 | 0.763 |
| Nicaragua (2 pops) | Mean | 0.2411 | 0.901 | 0.0082 | 0.913 | 0.1271 | 0.954 |
| Panama (2 pops) | Mean | 0.2588 | 0.300 | 0.0093 | 0.427 | 0.1364 | 0.518 |

nine different multilocus genotypes were used for the assignment test. These samples were assigned with a high level of confidence (score = 100%) to the Bolivia reference population (Table 4). Exclusion probabilities based on LOD-values were 100% for all countries except for Bolivia which had an exclusion probability of 3% (Prob 1, Table 4). The exclusion probability based on delta LOD-values was also not significant for Bolivia.

## 7. Discussion

In this study a large, geographically widespread genetic reference database for *S. macrophylla* was compiled and used to test assignment procedures as well as source blind samples provided by the timber industry. Interest in DNA-based timber tracking has been growing in recent years, as demand and technology make the approach more feasible. Different approaches are being tested. Lowe et al. [29] have reported successful identification of individual merbau logs (*Intsia* spp.) from both the concession and the saw mill, using a 'DNA passport' approach to verify that illegal timber is not being added to shipments between harvest and processing. The alternative used in our study – assignment testing based on reference datasets – requires much more initial investment but is being pursued for several species. To our knowledge, reference data are being compiled and tested for *Aquilaria crassna* [30], *Neobalanocarpus heimii* [31], *Intsia palembanica* [32] and *Entandrophragma cylindricum* [27]. This approach has also been used to trace back forest reproductive material of oaks to the seed stand of origin [33].

### 7.1. Genetic differentiation and spatial genetic pattern

Strong genetic differentiation was found among the 31 mahogany populations at the nSSRs ($\delta_{Gregorious}$ = 0.52, $F_{ST}$ = 0.18, $G_{ST(Hedrick)}$ = 0.65) and genetic and spatial distances among stands were correlated. This clear pattern provided a robust basis for successful assignment of genotypes. In general, the accuracy of genotype assignment approaches increases with increasing genetic differentiation among populations [25]. The level of genetic differentiation in this study was a bit higher than that found in previous studies for mahogany using the same markers [8,10] probably due to the greater geographical scale covered.

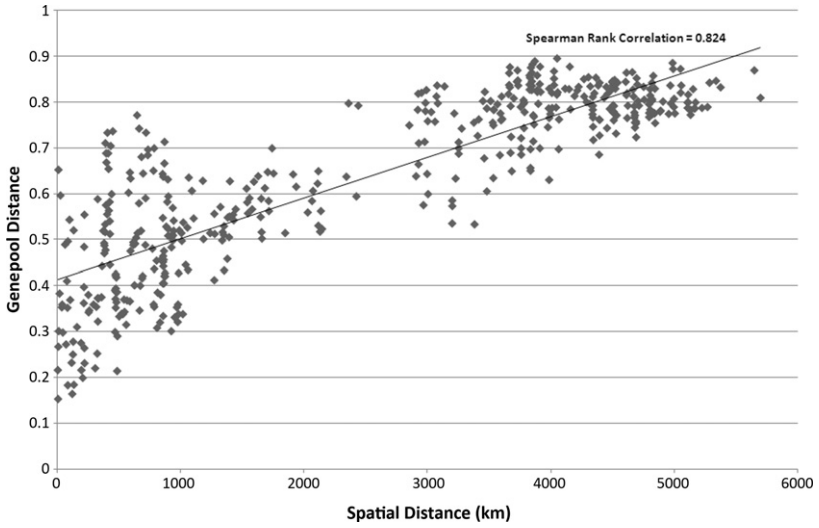


**Fig. 2.** Correlation between pairwise genetic distance [20] and among-population spatial distances for the 31 populations of *Swietenia macrophylla*.

**Table 4**
Results of the assignment test for two groups of wood samples of *Swietenia macrophylla*. Prob 1 = probability of exclusion based on LOD-values, Prob 2 = probability of exclusion based on DELTA LOD-values.

| Country | Samples Hamburg ($N = 12$) | | | | Samples South-America ($N = 9$) | | | |
|---|---|---|---|---|---|---|---|---|
| | −LOG($L$) | Scores | Prob 1 | Prob 2 | −LOG($L$) | Scores | Prob 1 | Prob 2 |
| Bolivia | 161.16 | 0.00 | 1.00 | | 41.65 | 1.00 | 0.03 | 0.18 |
| Costa Rica | 100.31 | 0.00 | 1.00 | | 101.48 | 0.00 | 1.00 | |
| Guatemala | 79.44 | 1.00 | 1.00 | 0.93 | 126.46 | 0.00 | 1.00 | |
| Honduras | 124.09 | 0.00 | 1.00 | | 153.51 | 0.00 | 1.00 | |
| Mexico | 86.6 | 0.00 | 1.00 | | 147.61 | 0.00 | 1.00 | |
| Nicaragua | 134.68 | 0.00 | 1.00 | | 174.28 | 0.00 | 1.00 | |
| Panama | 105.03 | 0.00 | 1.00 | | 88.37 | 0.00 | 1.00 | |
| Belize | 84.7 | 0.00 | 1.00 | | 126.69 | 0.00 | 1.00 | |
| Brazil | 204.26 | 0.00 | 1.00 | | 77.09 | 0.00 | 1.00 | |

### 7.2. Genotyping wood samples

Genotyping of timber samples required up to six repeats of PCR reactions because the amplification success of some nSSR loci was relatively low (<70%) and allelic drop-out occurred frequent. The amplification success and the detectable fragment size varied from locus to locus, which was probably due to degradation and fragmentation of DNA extracted from timber, as well as the presence of PCR inhibiting substances in the wood [34–37]. Lowe et al. [29] also encountered decreased success in amplifying long DNA fragments (100–350 bp) for merbau (*Intsia* sp.) wood, a timber from the Indonesian region.

From the results of the self-assignment tests, sufficient statistical power (>95% probabilities) was achieved when groups of at least 4 individuals were tested, rather than single individuals. For the blind tests, groups of 9 and 12 individuals were used, which made the assignment robust even when significant allele drop out was present. In most cases, sawn timber will be exported or imported in containers; each container will be filled with planks derived from multiple individuals. The associated trade documents will represent in most cases a claim for the true origin of the whole container load. Thus in practice the necessity for single individual assignment will be more an exception than the rule. It should also be noted that in practice a test will not be searching for the true origin of a timber shipment; instead it will be aiming to verify a match (or mismatch) between the stated origin and the source population. In other words, law enforcement agencies will be aiming to check whether a declared origin is correct or not. This approach is statistically less demanding. As long as the declared origin gets a non significant exclusion probability in the assignment test and has the highest assignment score compared to tested alternatives an authority could accept the statement of origin.

### 7.3. Blind test

The Bayesian methods of Rannala and Mountain [22] performed well in the blind tests for *S. macrophylla* timber origin. Calculation of exclusion probabilities is critical for avoidance of "false positive" assignments. Based on simulated genotypes, we calculated the distribution of LOD-values and compared it with observed values. This yielded a significant exclusion probability for the Guatemalan wood samples, and additional support for the origin of the Bolivian wood samples. It has been reported that this way to compute exclusion probabilities is too conservative [25]. Therefore we also used an approach based on differences between LOD values for the most likely and the second most likely reference population as a test statistic [26]. This less conservative approach gave non-significant exclusion probabilities in both cases. The difference in exclusion probabilities was also affected by the contrasting sample sizes for the reference databases from Bolivia and Guatemala. In Bolivia a total of 566 individuals from eight populations were used for the reference population but only 56 individuals from one population were available for Guatemala.

## Acknowledgements

The work was funded by Johann Heinrich von Thünen Institute (vTI) through budget given by the German Federal Ministry of Agriculture, Food and Consumer Protection (BMELV), the UK's Natural Environment Research Council and the EC (via contract ICA4-CT-2001-10101), and the Brazilian Council for Research (CNPq/CT-Amazonia Grant No. 554017/2006-7 to MRL). Access to RBCMA was granted by Programme for Belize. We are thankful for Vivian Kuhlenkamp and Caroline Hartleib for their excellent lab work, Nelson Repollet and Alberto Rodriguez for assistance in collecting the samples in Puerto Rico. We are grateful to Jimmy Grogan and José Ribeiro for field assistance in Brazil. We also acknowledge logistical support provided by the timber companies Serraria Marajoara, Madeireira Juary and Cacique Madeireira in Brazil. And we thank two anonymous reviewers for their helpful comments to improve the manuscript.

## References

[1] T.D. Pennington, A monograph of the Neotropical Meliaceae, in: T.D. Pennington, B.T. Styles, D.A.H. Taylor (Eds.), Flora Neotropica Monograph 28: Meliaceae, The New York Botanical Gardens, New York, USA, 1981.
[2] B.T. Styles, The flower biology of the Meliaceae and its bearing on tree breeding, Silvae Genet. 21 (1972) 175–182.
[3] F.B. Lamb, Mahogany of Tropical America: Its Ecology and Management, University of Michigan Press, Ann Arbor, MI, USA, 1966.
[4] M.R. Lemes, D. Grattapaglia, J. Grogan, J. Proctor, R. Gribel, Flexible mating system in a logged population of *Swietenia macrophylla* King (Meliaceae): implications for the management of a threatened neotropical tree species, Plant Ecol. 192 (2007) 169–179.
[5] B.D. Rodan, A.C. Newton, A. Veríssimo, Mahogany conservation: status and policy initiatives, Environ. Conserv. 19 (1992) 331–338.
[6] C. Navarro, J. Wilson, A. Gillies, M. Hernández, A new Mesoamerican collection of big-leaf mahogany, in: A.E. Lugo, J.C. Figueroa, M. Alayón (Eds.), Big Leaf Mahogany: Genetics, Ecology and Management, Ecological Studies, vol. 159, 2003, pp. 103–117.
[7] J. Grogan, P. Barreto, Big-leaf mahogany on CITES Appendix II: big challenge, big opportunity, Conserv. Biol. 19 (2005) 973–976.
[8] M.R. Lemes, R. Gribel, J. Proctor, D. Grattapaglia, 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 (2003) 2875–2883.
[9] A.C.M. Gillies, C. Navarro, A.J. Lowe, A.C. Newton, M. Hernández, J. Wilson, et al., Genetic diversity in Mesoamerican populations of mahogany (*Swietenia macrophylla*), assessed using RAPDs, Heredity 83 (1999) 722–732.
[10] R.R. Novick, C.W. Dick, M.R. Lemes, C. Navarro, A. Caccone, E. Bermingham, Genetic structure of Mesoamerican populations of big-leaf mahogany (*Swietenia macrophylla*) inferred from microsatellite analysis, Mol. Ecol. 12 (2003) 2885–2893.
[11] M.R. Lemes, C.W. Dick, C. Navarro, A.J. Lowe, S. Cavers, R. Gribel, Chloroplast DNA microsatellites reveal contrasting phylogeographic structure in mahogany (*Swietenia macrophylla* King, Meliaceae) from Amazonia and Central America, Trop. Plant Biol. 3 (2010) 40–49.

[12] N. Colpaert, S. Cavers, E. Bandou, H. Caron, G. Gheysen, A.J. Lowe, Sampling tissue for DNA analysis of trees: trunk cambium as an alternative to canopy leaves, Silvae Genet. 54 (2005) 265–269.
[13] J.J. Doyle, J.L. Doyle, Isolation of DNA from fresh plant tissue, Focus 12 (1987) 13–15.
[14] S. Dumolin, B. Demesure, R.J. Petit, Inheritance of chloroplast and mitochondrial genomes in pediculate oak investigated with an efficient PCR method, Theor. Appl. Genet. 91 (1995) 1253–1256.
[15] Y. Rachmayanti, L. Leinemann, O. Gailing, R. Finkeldey, Extraction, amplification and characterization of wood DNA from Dipterocarpaceae, Plant Mol. Biol. Rep. 24 (2006) 45–55.
[16] M.R. Lemes, R.P.V. Brondani, D. Grattapaglia, Multiplexed systems of microsatellite markers for genetic analysis in mahogany, *Swietenia macrophylla* King (Meliaceae), a threatened Netropical timber species, J. Hered. 93 (2002) 287–290.
[17] B.S. Weir, Genetic Data Analysis, Sinauer Associates, Sunderland, MA, 1990.
[18] P.W. Hedrick, A standardized genetic differentiation measure, Evolution 59 (2005) 1633–1638.
[19] H.R. Gregorius, B. Degen, A. Koenig, Problems in the analysis of genetic differentiation among populations – a case study in *Quercus robur*, Silvae Genet. 56 (2007) 190–199.
[20] H.R. Gregorius, The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance, Math. Biosci. 41 (1978) 253–271.
[21] B.F.J. Manly, Randomization and Monte Carlo Methods in Biology, Chapman and Hall, London, 1991.
[22] B. Rannala, J.L. Mountain, Detecting immigration by using multilocus genotypes, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 9197–9201.
[23] S. Piry, A. Alapetite, J.M. Cornuet, D. Paetkau, L. Baudouin, A. Estoup, GENECLASS2: a software for genetic assignment and first-generation migrant detection, J. Hered. 95 (2004) 536–539.
[24] B. Efron, Estimating the error rate of a prediction rule – improvement on cross-validation, J. Am. Stat. Assoc. 78 (1983) 316–331.
[25] J.M. Cornuet, S. Piry, G. Luikart, A. Estoup, M. Solignac, New methods employing multilocus genotypes to select or exclude populations as origins of individuals, Genetics 153 (1999) 1989–2000.
[26] T.C. Marshall, J. Slate, L.E.B. Kruuk, J.M. Pemberton, Statistical confidence for likelihood-based paternity inference in natural populations, Mol. Ecol. 7 (1998) 639–655.
[27] C. Jolivet, B. Degen, Use of DNA fingerprints to control the origin of sapelli timber (*Entandrophragma cylindricum*) at the forest concession level in Cameroon, Forensic Sci. Int. Genet. 6 (2012) 487–493.
[28] S. Wright, The genetical structure of populations, Ann. Eugen. 15 (1951) 323–353.
[29] A.J. Lowe, K.N. Wong, Y.S. Tiong, S. Iyerh, F.T. Chew, A DNA method to verify the integrity of timber supply chains; confirming the legal sourcing of merbau timber from logging concession to sawmill, Silvae Genet. 59 (2010) 263–272.
[30] M.C.M. Eurlings, H.H. van Beek, B. Gravendeel, Polymorphic microsatellites for forensic identification of agarwood (*Aquilaria crassna*), Forensic Sci. Int. 197 (2010) 30–34.
[31] L.H. Tnah, S.L. Lee, K.K.S. Ng, Q.Z. Faridah, I. Faridah-Hanum, Forensic DNA profiling of tropical timber species in Peninsular Malaysia, For. Ecol. Manage. 259 (2010) 1436–1446.
[32] K.N. Wong, W.L. Tan, F.T. Chew, Identification and characterization of microsatellite loci in *Intsia palembanica* (Leguminosae), a valuable tropical timber species, Mol. Ecol. Res. 9 (2009) 360–364.
[33] B. Degen, A. Höltken, M. Rogge, Use of DNA-fingerprints to control the origin of forest reproductive material, Silvae Genet. 59 (2010) 268–272.
[34] M.J. Asif, C.H. Cannon, DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*), Plant Mol. Biol. Rep. 23 (2005) 185–192.
[35] M.F. Deguilloux, M.H. Pemonge, R.J. Petit, Novel perspectives in wood certification and forensics: dry wood as a source of DNA, Proc. R. Soc. Lond. B: Biol. Sci. 269 (2002) 1039–1046.
[36] S. Dumolin-Lapègue, M.H. Pemonge, L. Gielly, P. Taberlet, R.J. Petit, Amplification of oak DNA from ancient and modern wood, Mol. Ecol. 8 (1999) 2137–2140.
[37] Y. Rachmayanti, L. Leinemann, O. Gailing, R. Finkeldey, DNA from processed and unprocessed wood: factors influencing the isolation success, Forensic Sci. Int. Genet. 3 (2009) 185–192.
[38] R. Li, J. Buongiorno, J.A. Turner, S. Zhu, J. Prestemon, Long-term effects of eliminating illegal logging on the world forest industries, trade and inventory, Forest Policy Econ. 10 (2008) 480–490.
[39] M.P. Goncalves, M. Panjer, T.S. Greenberg, W.B. Magrath, Justice for Forests – Improving Criminal Justice Efforts, World Bank Study, 2012. http://siteresources.worldbank.org/EXTFINANCIALSECTOR/Resources/Illegal_Logging.pdf.