



DNA-fingerprints to determine the origin of forest reproductive material: A case study in pedunculate oak (*Quercus robur*)

Bernd Degen¹ & Aki Höltnen²

1

**J.H. von Thünen-Institute (vTI)
Institute for Forest Genetics
Grosshansdorf
Germany**

2

**Plant Genetic Diagnostics GmbH
Grosshandorf
Germany**



1) Introduction

2) Case study with *Quercus robur*

3) Conclusions for the practice



Economic damage due to the use of unsuitable forest reproductive material:

- Low growth rates
- Bad timber quality
- Low resistance against environmental factors, insects and fungi
- Pollen contamination in neighbouring stands



IUFRO – Scots pine provenance trial 1982/84 at age 20 at plots in Waldsieversdorf



Legal requests – Reproductive material

German Act on Forest Reproductive Material (FoVG) 22 May 2002

Art. 1 Purpose and scope of the law

*(1) It is the purpose of the law to maintain and improve the forest with its manifold positive effects and its genetic diversity by providing high-quality forest reproductive material **with secured identity** as well as to promote the forestry and its performance....*

Programme for the Endorsement of Forest Certification (PEFC)

*Sixty four percent (64%) of all private forests in Germany are now certified as sustainable managed under PEFC and the PEFC authorities in Germany have recently imposed a condition that all plants used in PEFC certified forests in Germany must be of **controlled origin** (PEFC-Standards 4.4).*



Need for control of forest reproductive material:

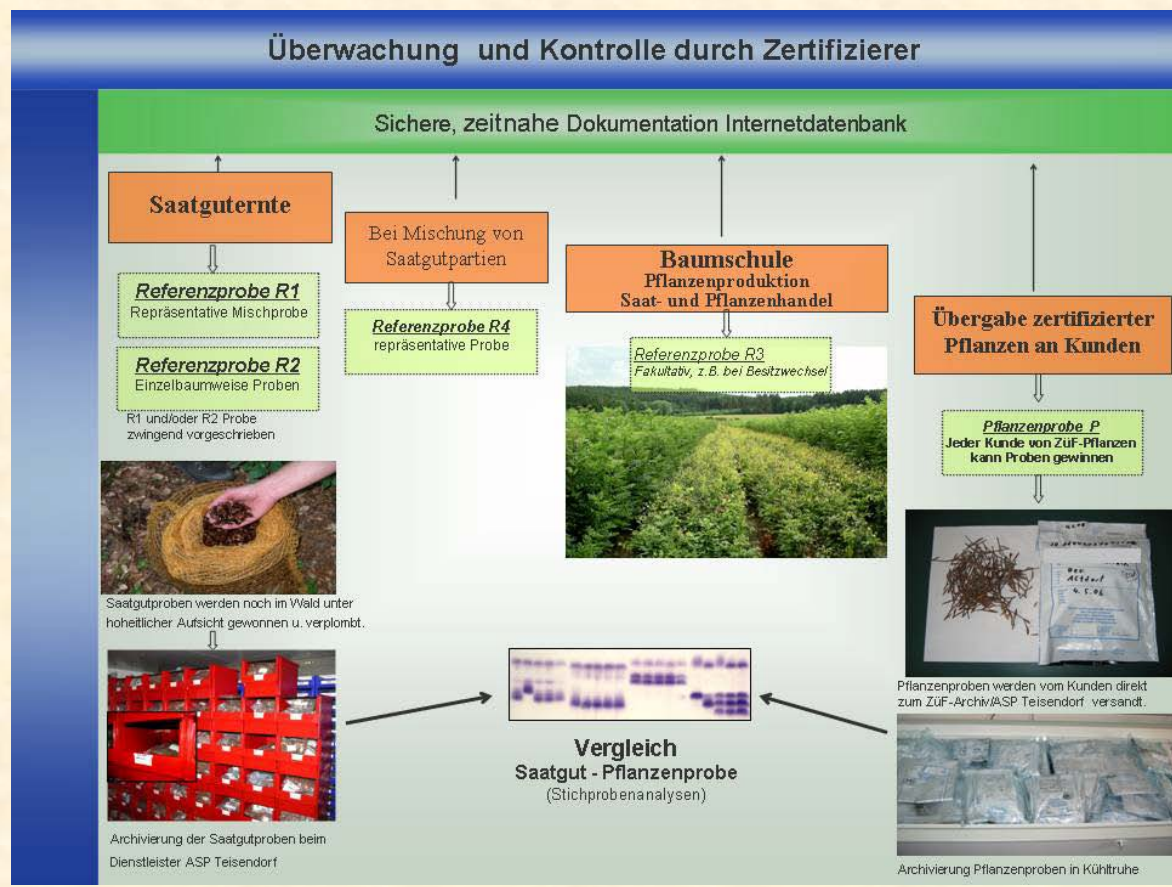
- Important differences in species performance at a given stand e.g. *Quercus robur* versus *Quercus petraea* => **Identification of tree species and level of species mixture, level of hybridisation**
- High quality clones as result of intensive tree improvement e.g. *Populus sp.* => **verification of clone identity**
- Selected material from seed orchards and seed stands => **test origin of the material**





ZÜF (Zertifizierungsring für überprüfbare Forstliche Herkunft Süddeutschland e. v)

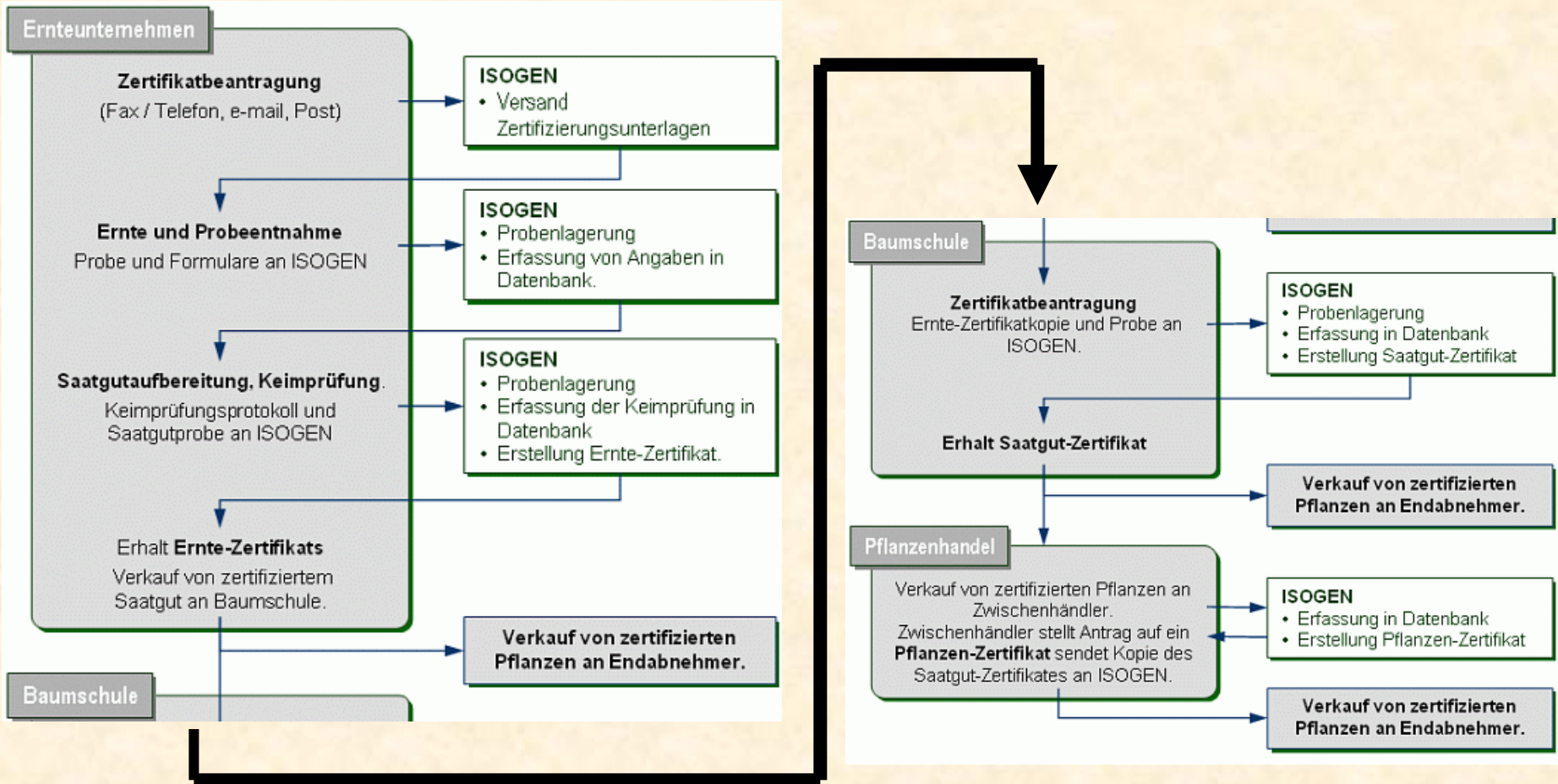
<http://www.zuef-forstpflanzen.de>





Certification by ISOGEN

<http://www.isogen.de/>





Contribution of Evoltree for a better control of forest reproductive material:

- Optimisation of DNA-extraction protocols
- Microsatellite markers + SNPs for assignment tests
- Simulation models to design optimal sampling strategies
- Repository Centre to store reference samples
- Information on genetic differentiation among populations for different kind of markers / sequences

2. Case study with *Quercus robur*



Sampling

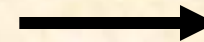
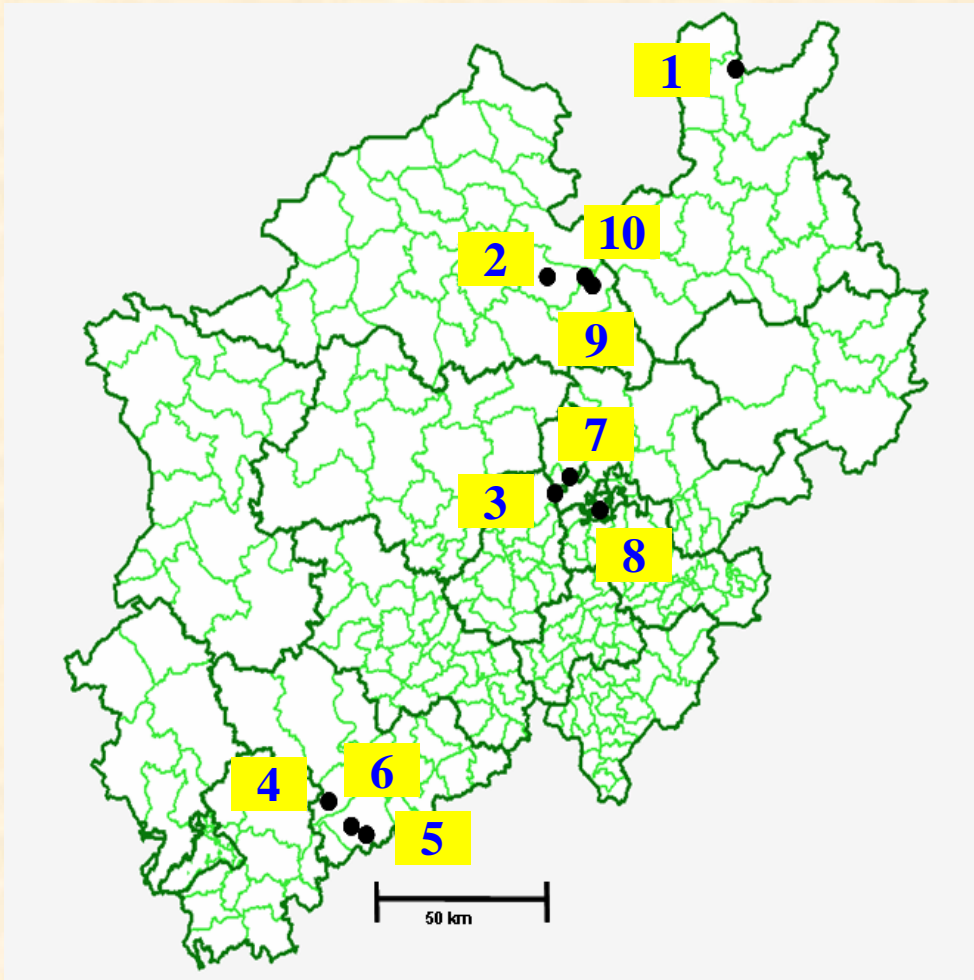
- 10 registered seed stands with 2 ha to 27 ha
- all stands have been artificially regenerated
- 822 adult trees sampled (56 to 91 per stand)
- 320 seeds / seedlings (18-35 per stand)
- genotyping at 9 microsatellites (nSSRs)



2. Case study with *Quercus robur*



Genetic study in 10 registered seed stands in Germany

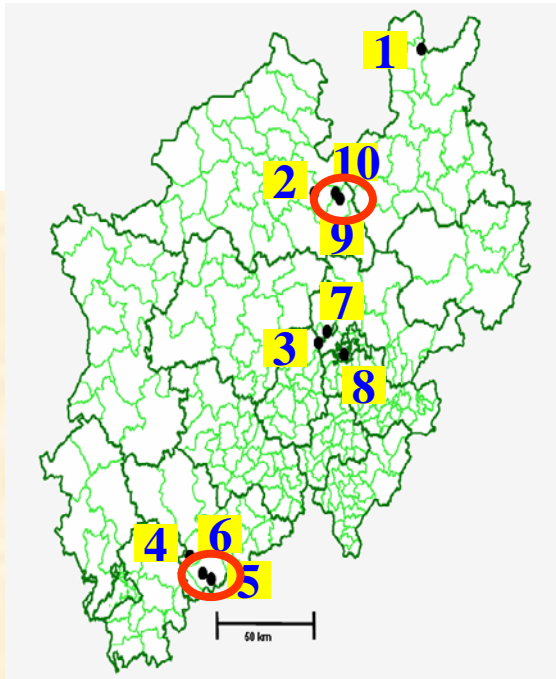


2. Case study with *Quercus robur*



Probabilities for genetic differences among stands

	1	2	3	4	5	6	7	8	9	10
1		1.000	0.964	0.878	1.000	0.919	1.000	1.000	1.000	1.000
2			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3				0.910	1.000	0.959	1.000	1.000	1.000	1.000
4					0.999	0.890	1.000	1.000	1.000	1.000
5						0.565	1.000	1.000	1.000	1.000
6							0.991	1.000	1.000	1.000
7								1.000	1.000	1.000
8									1.000	1.000
9										0.491
10										



Stands 5 and 6 as well as stands 9 and 10 are from equal or similar seed sources!



Assignment of seeds/ seedlings to adult stands

Using a Bayesian multilocus-approach (Rannala & Mountain 1997)

- Calculate observed frequencies for each allele in a all stands
- Estimate likelihood of a diploid genotype at each locus
- Multiply likelihoods for each locus together to yield overall likelihood for genotype in reference population
- Multiply likelihoods of each analysed seed/ seedling of the group
- Rank the seed stands according to their likelihoods



Assignment of seeds/ seedlings to adult stands

```
#####
#### SCORES ####
#####
```

	P (1)	P (2)	P (3)	P (4)	P (5)	P (6)	P (7)	P (8)	P (9)	P (10)
Seeds_1	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Seeds_2	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Seeds_3	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Seeds_4	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
Seeds_5	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000
Seeds_6	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
Seeds_7	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
Seeds_8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
Seeds_9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
Seeds_10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.627	0.373

3. Conclusions for the practice



?



?

Need for a reliable method to control origin of reproductive material

DNA fingerprints using nuclear microsatellites are available for most important tree species

Superior statistical power of SNPs for assignment will be tested soon

Field test with *Quercus robur* and simulation studies demonstrated that the assignment of multilocus-genotypes for groups is a very robust method

Need for a centralised and standardised reference data base

Method could be used in existing certification systems => reduction of sampling and storage efforts



State forest service of NRW for financial and logistic support => Martin Rogge



EvoTree for financial support

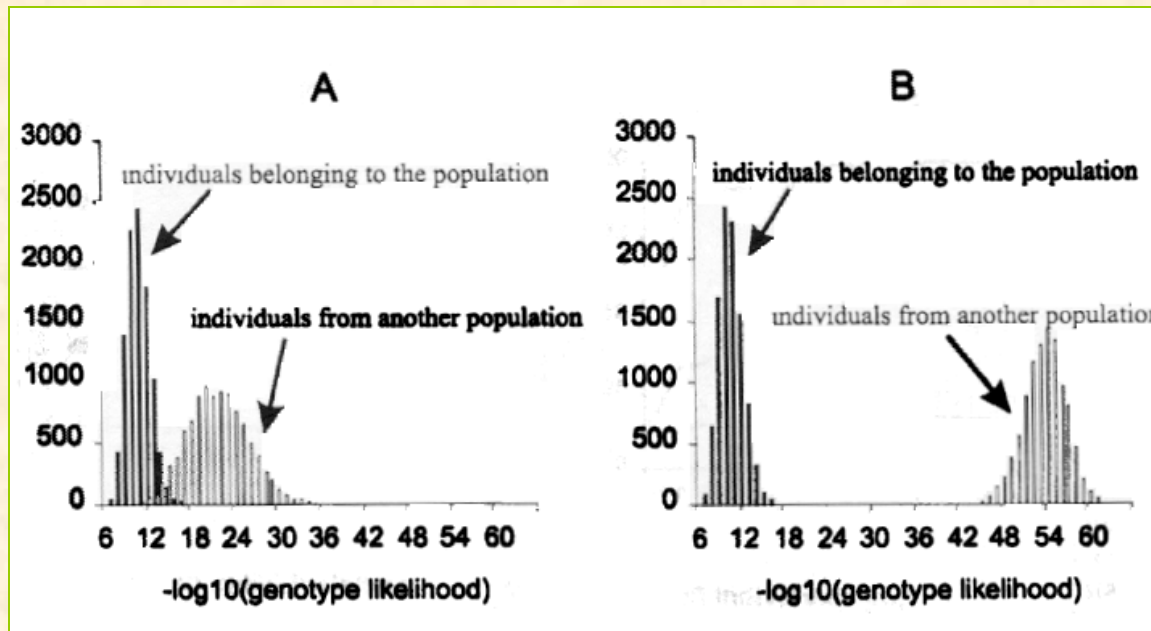
Alexandre Sebbenn (Sao Paulo Forest Institute) for his contributions to the simulation studies

Inga Michalczyk, Vivian Kuhlenkamp, Alexandra Meyer and Inge Schulze (vTI) for genotyping / lab work



Calculation of the likelihoods

- a) Frequency approach
- b) Bayesian approach
- c) Genetic distance approach





Estimate likelihood of a diploid genotype:

$$k = k'$$

$$A_k A_{k'} = \frac{(n_{ijk}+1/K_j+1)(n_{ijk}+1/K_j)}{(n_{ij}+2)(n_{ij}+1)}$$

$$k \neq k'$$

$$A_k A_{k'} = \frac{2(n_{ijk}+1/K_j)(n_{ijk'}+1/K_j)}{(n_{ij}+2)(n_{ij}+1)}$$

i = population ; j = locus; k = allele

n_{ijk} = absolute frequency of allele k

n_{ij} = sample size

K_j = number of different alleles sampled in all pops



Simulation study on sample design for control of origin oak seeds in NRW

Error I	Right pop not assigned
Error II	Prob of right pop not in 99% confidence interval
Error III	Prob of wrong pop in 99% confidence interval

Standard	
Pops	20
N Pop	150
Seeds	30
Loci	9 Quercus loci
MD	0%
GE	0%

N per Pop	Error I	Error II	Error III
50	0.0	94.0	0.0
100	0.0	60.0	0.0
150	0.0	28.0	0.0
200	0.0	20.0	0.0
250	0.0	15.0	0.0

N seeds	Error I	Error II	Error III
10	0.0	15.0	8.0
20	0.0	22.0	0.0
30	0.0	27.0	0.0
40	0.0	35.0	0.0
50	0.0	32.0	0.0

N_Loci	Error I	Error II	Error III
3	0.0	14.0	4.0
6	0.0	27.0	1.0
9	0.0	26.0	0.0
12	0.0	24.0	0.0
12	0.0	22.1	0.0



Simulation study on sample design for control of origin oak seeds in NRW

Error I	Right pop not assigned
Error II	Prob of right pop not in 99% confidence interval
Error III	Prob of wrong pop in 99% confidence interval

Standard	
Pops	20
N Pop	150
Seeds	30
Loci	9 Quercus loci
MD	0%
GE	0%

MD %	Error I	Error II	Error III
0	0.0	36.0	0.0
5	0.0	11.0	16.0
10	0.0	7.0	83.0
15	0.0	12.0	94.0
20	0.0	43.0	94.0

GE %	Error I	Error II	Error III
0	0.0	31.0	0.0
1	0.0	58.0	0.0
2	0.0	64.0	0.0
3	0.0	69.0	0.0
4	0.0	86.0	0.0