Use of DNA barcoding for host plant identification

Barbara Gravendeel
Marcel Eurlings
Theodoor Heijerman

KEY WORDS:
DNA barcoding, host plant, *rbc*L, *Scolytus multistriatus*, *Scolytus laevis*, *Scolytus scolytus*, *trn*L, *Ulmus glabra*

Entomologische Berichten 69 (2): 30-35

We present a molecular approach for identification of Dutch elm (*Ulmus glabra*) as host plant of some species of European elm beetles (*Scolytus*) using DNA sequences of the chloroplast *rbc*L gene and *trn*L intron. The samples analysed were obtained from a piece of infested bark collected from a bridge in France. Molecular identification was consistent with microscopical analysis and with reference material collected in The Netherlands. When sequences were included of all trees surrounding the bridge, the same results were obtained. DNA barcoding therefore seems a promising new tool for increasing knowledge of plant-insect interactions.

Introduction

Plants are used by insects as food, as refuge, as substrate for mating and egg-laying and other purposes. Many insect-plant interactions are still unknown and field observations are often scarce and incomplete. To increase knowledge on the ecology of insects and plants, more data on plant-insect interactions are urgently needed.

Identification of plant parts in which insects are collected can be done with multiple techniques. The simplest method consists of identification using field guides. This often requires flowering or fruiting of the host plant. If only vegetative plant material is available, identification requires more complex techniques such as microscopical analyses.

The development of the polymerase chain reaction (PCR) and automated DNA sequencing has been revolutionary for identification of biodiversity. Researchers around the world are now committed to find the appropriate DNA segments that allow accurate species identifications in the so-called Consortium of the Barcode of Life (CBOL) initiative (http://barcoding.si.edu). DNA barcodes are relatively short and omnipresent within a particular taxonomic group. They mutate relatively fast, which results in significant sequence variation between species, and, in principle, relatively low sequence variation within species.

In June 2007 we collected specimens of *Scolytus multistriatus* (Marsham) (smaller European elm bark beetle) and S. *laevis* Chapuis (middle elm tree split bark beetle) from a freshly cut log

1. Wooden bridge near La Cabanasse (42.504° N; 2.130° E; 1400 m altitude) infested with elm beetles in June 2008. Photo: Theodoor Heijerman
1. Houten bruggetje bij La Cabanasse (42.504° N; 2.130° O; 1400 m) geïnfecteerd met iepenspintkevers in juni 2008.

2. Gallery system with dead specimen of *Scolytus laevis* (upper) and *S. scolytus* (below). The pictures shows a maternal gallery with entrance orifice and mating chamber. Perpendicular to the maternal galleries the larval galleries can be seen, which are now filled with excrements. Photo: Theodoor Heijerman
2. Gangsysteem van *Scolytus laevis* (boven) en *S. scolytus* (beneden) met daarin een dood exemplaar. Op de foto's is de moedergang te zien met het boorgat en de paringskamer. Loodrecht op de moedergang staan de larvengangen die nu gevuld zijn met uitwerpselen.

3. *Scolytus multistriatus*, smaller European elm beetle. Photo: Theodoor Heijerman
3. *Scolytus multistriatus*, kleine iepenspintkever.

4. *Scolytus laevis*, middle European elm beetle. Photo: Theodoor Heijerman
4. *Scolytus laevis*, middelste iepenspintkever.

that was used to construct steps for a wooden bridge in a steep valley near La Cabanasse, France (42.504° N; 2.130° E; 1400 m altitude; figure 1). We collected a piece of bark from the log for identification. In June 2008, when we had the opportunity to visit the location again, we collected some additional bark samples. Underneath the bark were many gallery systems (figure 2). In some of these we found dead specimens of three species of Scolytidae, viz. *S. multistriatus* (figure 3), *S. laevis* (figure 4) and *S. scolytus* (Fabricius) (large elm bark beetle).

Host plants for *S. laevis* include multiple tree genera: elm (*Ulmus*), maple (*Acer*), alder (*Alnus*), hazel (*Corylus*), beech (*Fagus*), crabapple (*Malus*), oak (*Quercus*) and lime (*Tilia*) whereas *Scolytus multistriatus* and *S. scolytus* are mainly reported from elm and a small range of other hosts (Wood & Bright 1992, Bright & Skidmore 1997). Trees surrounding the bridge included ash (*Fraxinus*), elm, oak, maple, hazel, lime, alder and aspen (*Populus*). Superficially, the bark of these tree species all looked very similar (figure 5). The piece of infested bark (figure 6) could not be identified by professional botanists based on morphological characters only.

Methods for the molecular identification of wood have been developed in the past ten years (Dumolin-Lapegue *et al.* 1999, Ohyama *et al.* 2001, Deguilloux *et al.* 2002, Eurlings & Gravendeel 2005). In this study, we investigated whether (1) DNA sequences of the chlorplast *rbc*L gene and *trn*L intron could be obtained from the infested piece of bark; (2) whether these could be used for taxonomic identification of the tree which originally produced the bark and (3) whether the DNA barcodes obtained were consistent with microscopical identifications.

Material & Methods

DNA extraction

Bark and leaf material were ground with a mortar and pestle in liquid nitrogen to dust in a special ancient DNA facility at Leiden commited to extract tiny amounts of DNA from very old or highly degraded material. To avoid contamination, all tools were thoroughly rinsed with a 0.04% bleach solution and 70% ethanol after each usage. Total genomic DNA was extracted from 40–100 mg of dust using the DNeasy Plant mini kit of Qiagen and following the manufacturer's protocols (Qiagen, Inc.).

PCR amplification

Part of the *rbc*L gene was amplified using the primers 636F and 1460R (Savolainen 2000). The *trn*L intron was amplified using the primers c and d (Taberlet *et al.* 1991). Polymerase chain reactions were carried out on a T3 Thermocycler (Westburg) in a 50 microliter volume containing 0.1–50 ng of genomic DNA, 0.1 M of each primer, 10 M of each dNTP, Qiagen PCR buffer (50 mM KCl, 10 mM TRIS-HCl pH 8.7, 1.5 mM $MgCl_2$) and 1.5 units of Taq

5. Bark details of nine tree species growing in the vicinity of the wooden bridge at La Cabanasse in June 2008. First row, from left to right: Dutch elm (*Ulmus glabra*), Norway maple (*Acer platanoides*), trembling aspen (*Populus tremula*) and common alder (*Alnus glutinosa*). Second row, from left to right: sycamore (*Acer pseudoplatanus*), lime (*Tilia platyphyllos*), European mountain ash (*Sorbus aucuparia*) and chess apple (*Sorbus aria*). Third row: common hazel (*Corylus avellana*). Photo's: Theodoor Heijerman
5. Details van de schors van negen boomsoorten rond het met Iepenspintkevers geïnfecteerde bruggetje. Eerste rij van links naar rechts: gladde iep, noorse esdoorn, ratelpopulier en zwarte els; tweede rij van links naar rechts: gewone esdoorn, zomerlinde, lijsterbes, meelbes; derde rij: hazelaar.

6. Detail of bark of steps of wooden brigde at La Cabanasse, june 2008. Photo: Theodoor Heijerman
6. Details van de schors van de tredes van het bruggetje bij La Cabanasse, juni 2008.

DNA polymerase (Qiagen, Inc.). BSA (Amersham Biosciences) was generally necessary for amplification and added to a final concentration of 1–2 ng/microliter. The thermal cycling profile started with a 5 min denaturation step of 94 °C, then comprised 35 cycles each with 20 seconds denaturation at 94 °C, 20 sec annealing at 52–55 °C and 45 seconds elongation at 72 °C, and the program ended with 5 min extension at 72 °C.

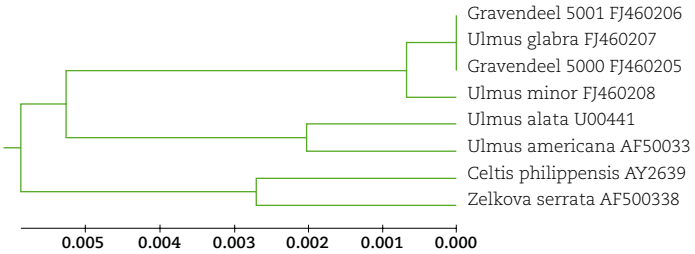
DNA sequencing

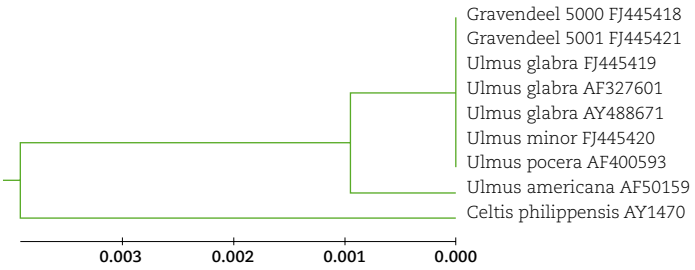
Amplification products were separated on a 1.5% agarose/TBE gel, purified using the QIAquick PCR Purification Kit (Qiagen, Inc.) and sequenced on an ABI 377 automated sequencer using standard dye-terminator chemistry following the manufacturers protocols (Applied Biosystems, Inc.). Complementary strands were assembled and edited with Sequencer version 4.01 (Gene Codes Corporation).

Voucher [1] Accessie [2]	Species Soort	Origin Herkomst	GenBank accession GenBank referentie *rbc*L	*trn*L
Gravendeel 5000 (B)	*U. glabra*	La Cabanasse, France	FJ460205	FJ445418
Gravendeel 5001 (L)	*U. glabra*	La Cabanasse, France	FJ460206	FJ445421
Gravendeel 5011 (L)	*U. glabra*	Wageningen, NL	FJ460207	FJ445419
Gravendeel 5010 (L)	*U. minor*	Wageningen, NL	FJ460208	FJ445420

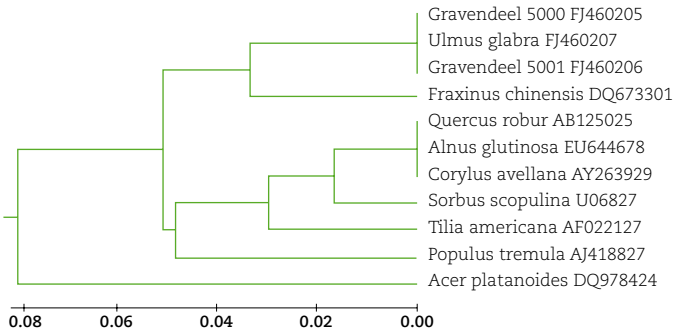
[1] L=DNA extracted from leaves; B=DNA extracted from bark
[2] L=DNA geïsoleerd uit blad; B=DNA geïsoleerd uit schors

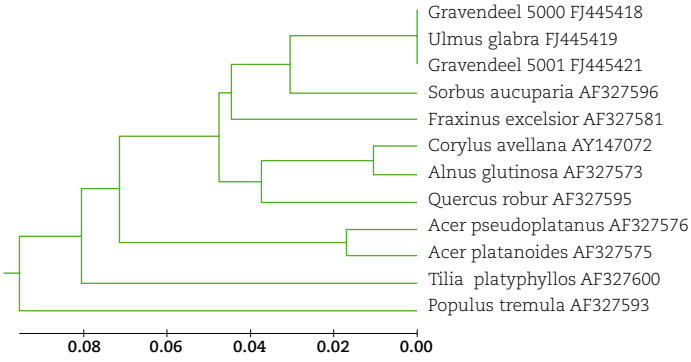
Table 1. NCBI GenBank accessions of DNA sequences obtained for this study. Voucher specimens are deposited at the Leiden branch of the National Herbarium of The Netherlands.
Tabel 1. NCBI GenBank accessies van de in dit onderzoek gegenereerde DNA sequenties. Vouchers zijn opgeslagen in de Leidse vestiging van het Nationaal Herbarium Nederland.





7. Results of matches between *rbc*L (above) and *trn*L (below) sequences obtained from bark piece analyzed with NCBI GenBank database sequences found by BLAST searches. Numbers on the x-axis represent percentage of DNA sequence difference.
7. Resultaten van vergelijkingen tussen *rbc*L (boven) en *trn*L (beneden) sequenties gegenereerd uit het geanalyseerde stukje schors met DNA sequenties in de NCBI GenBank database gevonden na BLAST analyses. De nummers op de x-as corresponderen met het percentage afwijkende nucleotiden in de DNA sequenties.





8. Results of matches between *rbc*L (above) and *trn*L (below) sequences obtained from bark with NCBI GenBank database sequences of trees surrounding the infected bridge. Numbers on the x-axis represent percentage of DNA sequence difference.
8. Resultaten van vergelijkingen tussen *rbc*L (boven) en *trn*L (beneden) sequenties gegenereerd uit de schors met DNA sequenties in de NCBI GenBank database van bomen uit de omgeving van het bruggetje. De nummers op de x-as corresponderen met het percentage afwijkende nucleotiden in de DNA sequenties.

Sequence analysis

Firstly, NCBI GenBank BLAST searches were used to determine the differences among DNA sequences generated in this study with those already deposited in this public database. Identifications with significant matches were checked against Sosef *et al.* (2007) and Saule (2002) to determine their presence or absence in the area from which the bark was collected. To find further support for the molecular identifications found using BLAST searches, we collected additional *rbc*L and *trn*L sequences of English elm and Dutch elm from accessions in the Botanical Gardens of Wageningen University and a local tree growing in the surroundings of the infested bridge (table 1). To check whether DNA barcoding would also be capable of finding the closest matching sequence amongst a large set of sequences from more distantly related species, a second analysis was performed. This analysis included DNA sequences of all trees surrounding the bridge. For local tree species of which no *rbc*L and *trn*L sequences were available in NCBI GenBank, sequences of another species of the same genus were used.

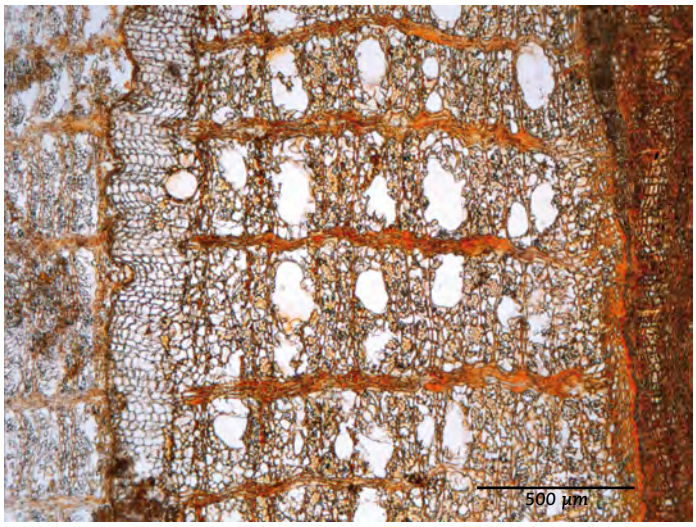
Microscopy

Bark samples were sectioned and macerated according to the standard techniques described by Baas & Zhang (1986) for light microscopy.

Results

DNA sequences of both the *rbc*L gene and *trn*L intron were successfully generated from DNA extracted from the piece of bark (table 1). BLAST searches of these DNA sequences against the data present in the NCBI GenBank matched most closely with elm (figure 7). The DNA sequences generated from additional material collected from Dutch elm (Gravendeel 5011) matched those generated from the piece of bark and local elm tree most closely (figure 7). When the sequences generated from the piece of bark and local elm tree were compared with sequences from all surrounding trees, they also matched most closely with Dutch elm (figure 8).

Microscopical investigations revealed that the anatomy of the bark resembled that of elm mostly since the tangential



9. Wood anatomical section of bark piece analyzed. Photo: Bertie-Joan van Heuven
9. Houtanatomisch preparaat van het geanalyseerde stukje schors.

fibre-bands alternated with layers of sieve tubes, and the companion cells and phloem parenchyma were interspersed with large mucilage cells and cavities. The phloem rays showed hardly any dilatation, but in the cortex more or less triangular dilatation zones were present (figure 9). These features are characteristic for the wood of elm (Trockenbrodt 1989).

Discussion

According to Sosef *et al.* (2007) and Saule (2002), three species of Ulmaceae occur in the area where the bark was collected: English elm (*Ulmus minor*), Dutch elm (*Ulmus glabra*) and European hackberry (*Celtis australis*). Unfortunately, reference DNA sequences of the *rbc*L gene and *trn*L intron of these three species were not yet present in the NCBI GenBank database when this study was carried out. Multiple DNA sequences of other species of these two genera had been deposited in this database, though. The best matches between the DNA barcodes obtained from the piece of bark with those in GenBank were all from elm. It seems therefore very likely that the bark was produced by a species of elm instead of Hackberry. Comparison with DNA sequences obtained from additionally collected material of Dutch elm (Gravendeel 5011) and English elm (Gravendeel 5010) in The Netherlands confirmed this. When DNA sequences were included of all trees surrounding the bridge, the DNA barcoding method was also capable of finding Dutch elm as the closest matching sequence amongst a large set of sequences from other tree species.

The *rbc*L gene evolves relatively slow and cannot always provide an unambiguous resolution of plant identity to the species level. DNA barcodes of the *trn*L intron, however, are often well suited for this. When adequate databases of local plant reference sequences are available, DNA barcodes can therefore be a valuable tool for taxonomic identification of host plants when traditional methods such as field guides fail.

Wood and bark – just like many museum specimens – are challenging templates to generate DNA barcodes from as DNA quality and quantity are usually low. Even when DNA is degraded, though, sufficiently long barcodes can often be recovered for effectively identifying both plant and insect specimens (Eurlings & Gravendeel 2005, Hajibabaei *et al.* 2006). Targets of DNA barcoding identification therefore nowadays not only comprise host plants (Singer *et al.* 2008, this study), but also include incomplete museum specimens lacking morphological structures critical for identification keys such as genitalia, immature stages of insects such as eggs or caterpillars (Van Velzen *et al.* 2007) and diet components (Miller *et al.* 2006, Matheson *et al.* 2006, Pons 2006).

Not all museum curators are keen to sacrifice rare specimens in their collections for DNA extraction. Non-destructive extraction protocols which are recently developed make it possible to retain voucher specimens (Singer *et al.* 2008, Hunter *et al.* 2008). With this method, specimen material is not ground but only soaked in an extraction buffer with low amounts of demineralising reagents. Afterwards, the buffer is subsequently processed for DNA extraction whereas the original specimen can be dried and restored again for future morphological and or molecular studies.

Conclusions

We successfully generated DNA barcodes from a piece of bark infested with three species of *Scolytus*. With these barcodes, the tree which originally produced the bark could be identified as Dutch elm. This molecular identification was congruent with microscopical data and with reference DNA sequences collected from additional plants that could be identified based on flowers or fruits. DNA barcoding therefore seems a promising new tool for host plant identification when host material is vegetative and more traditional identification methods cannot be used.

Acknowledgements

We thank Pieter Baas and Bertie Joan van Heuven (Nationaal Herbarium Nederland) for preparation of the wood anatomical sections , Michelle Spierings and Jelle Zandveld (biology students at Leiden University) for generating part of the DNA sequences, Erik van Nieukerken (Naturalis) for literature references and Mirjam Lemmens (Botanical Gardens of Wageningen University) for assistance in obtaining additional leaf material.

References

Baas P & Zhang X 1986. Wood anatomy of trees and shrubs from China. I. Oleaceae. International Association of Wood Anatomists Bulletin 7: 195-220.

Bright DE & Skidmore RE 1997. A catalog of Scolytidae and Platypodidae (Coleoptera). Supplement 1 (1990-1994) NRC Research Press, Ottawa, Ontario.

Deguilloux MF, Pemonge MH & Petit RJ 2002. Novel perspectives in wood certification and forensics: dry wood as a source of DNA. Prococeedings of the Royal Society of London B 269: 1039-1046.

Dumolin-Lapegue S, Pemonge MH, Gielly L, Taberlet P & Petit RJ 1999. Amplification of oak DNA from ancient and modern wood. Molecular Ecology 8: 2137-2140.

Eurlings MCM & Gravendeel B 2005. TrnL-trnF sequence data imply paraphyly of *Aquilaria* and *Gyrinops* (Thymelaeaceae) and provide new perspectives for agarwood identification. Plant Systematics and Evolution 254: 1-12.

Hajibabaei M, Smith MA, Janzen FH, Rodriguez JJ, Whitfield JB & Hebert PDN 2006. A minimalist barcode can idenfity a specimen whose DNA is degraded. Molecular Ecology Notes 6: 959-964.

Hunter SJ, Goodall TI, Walsh KA, Owen R & Day JC. 2008. Nondestructive DNA extraction from blackflies (Diptera: Simuliidae): retaining voucher specimens for DNA barcoding projects. Molecular Ecology Resources 8: 56-61.

Matheson CD, Muller GC, Junnila A, Vernon K, Hausmann A, Miller MA, Greenblatt C & Schlein Y 2006. A PCR method for detection of plant meals from the guts of insects. Organisms Diversity & Evolution 7: 294-303.

Miller MA, Mueller GC, Kravchenko VD, Junnila A, Vernon KK, Matheson CD & Hausmann A 2006. DNA-based idenfitication of Lepidoptera larvae and plant meals from

their gut contents. Russian Entomological Journal 15: 427-432.
Ohyama M, Baba K & Itoh T 2001. Wood identification of Japanese *Cyclobalanopsis* species (Fagaceae) based on DNA polymorphism of the intergenetic spacer between trnT and trnL 5'exon. Journal of Wood Science 47: 81–86.
Pons J 2006. DNA-based identifications of preys from non-destructive, total DNA extractions of predators using arthropod universal primers. Molecular Ecology Notes 6: 623-626.
Saule M 2002. La grande Flore illustreé des Pyrénées. Editions Milan, France.
Savolainen V 2000. Phylogeny of the eudicots: a nearly complete familial analysis based on rbcL gene sequences. Kew Bulletin 55: 257-309.
Singer RB, Gravendeel B, Cross H & Ramirez SR 2008. The use of orchid pollinia for taxonomic identification. Selbyana 29(1): 6-19.
Sosef M, Van den Berg RG & Van der Maesen LJG 2007. Handleiding botanisch veldpracticum Les Pyrenees Orientales. Reader BIS-20803, Wageningen Universiteit.
Taberlet P, Gielly L, Patou G & Bouvet J 1991. Universal primers for amplification of three noncoding regions of chloroplast DNA. Plant Molecular Biology 17: 1105-1109.
Trockenbrodt M 1989. Entwicklung, Variabilität und taxonomische Bedeutung der Rindenstrukturen von *Quercus robur* L., *Ulmus glabra* Mill., *Populus tremula* L. und *Betula pendula* Roth. PhD thesis, Hamburg.
Van Velzen R, Bakker FT & Van Loon JJA 2007. DNA barcoding reveals hidden species diversity in *Cymothoe* (Nymphalidae). Proceedings of the Netherlands Entomological Society Meeting 18: 95-103.
Wood SL & Bright DE 1992. A catalog of Scolytidae and Platypodidae (Coleoptera). Part 2: Taxonomic index. Great Basin Naturalist Memoirs 13: [i-iv], 1-833, [i-iv], 835-1553.

Ontvangen: 21 november 2008
Geaccepteerd: 14 januari 2009

Samenvatting

Het gebruik van DNA barcodering voor de determinatie van waardplanten
Verschillende soorten iepenspintkevers (*Scolytus*) werden in de schors van een met stammetjes gerepareerd houten bruggetje aangetroffen. Omdat we wilden weten welke boomsoort hier als waardplant gebruikt werd, vergeleken we de schors van de geïnfecteerde stammetjes met de schors van bomen uit de omgeving. Het bleek echter voor divere professionele plantentaxonomen onmogelijk om op deze manier de identiteit van het hout vast te stellen. Met behulp van DNA sequenties van het chloroplast *rbc*L gen en *trn*L intron kon de schors echter geïdentificeerd worden als ruwe iep (*Ulmus glabra*). Determineren met DNA lijkt een veelbelovende nieuwe techniek om waardplanten mee op naam te brengen als bloemen en/of vruchten ontbreken en traditionele determinatiesleutels te kort schieten.

Barbara Gravendeel
Universiteit Leiden
Nationaal Herbarium Nederland & Instituut Biologie Leiden
Einsteinweg 2
2333 CC Leiden
gravendeel@nhn.leidenuniv.nl

Marcel Eurlings
Universiteit Leiden
Moleculair lab Instituut Biologie Leiden, Naturalis & Nationaal Herbarium Nederland
Kaiserstraat 63
2311 GP Leiden

Theodoor Heijerman
Universiteit Wageningen
Leerstoelgroep Biosystematiek, sectie Diertaxonomie
Generaal Foulkesweg 37
6703 BL Wageningen