Differentiation of six *Eucalyptus* trees grown in Mexico by ITS and six chloroplast barcoding markers

By M. Fladung[1),*)], H. Schroeder[1)], C. Wehenkel[2)] and B. Kersten[1)]

(Received 6th July 2015)

Abstract

Different species of the genus *Eucalyptus*, originally native to Australia, are being cultivated in different parts of the world due to their fast growth and beneficial wood properties. In Mexico, probably up to 25 different *Eucalyptus* species (many of them with unknown species declaration) were introduced early in the 20th century. Many *Eucalyptus* species are cross compatible and information about provenances of the single eucalypt species is rare. In this study, an experimental plantation established in 1984 and located in Northeast of Mexico was chosen as example to re-assign the species name of six randomly selected *Eucalyptus* trees growing in this plantation. First, a phylogenetic tree was constructed from complete chloroplast sequences of 31 *Eucalyptus* species available in the NCBI database. The phylogenetic tree includes three of the nine *Eucalyptus* species known to be introduced to Mexico, namely *E. camaldulensis*, *E. saligna* and *E. grandis*, which belong to a clade named "Symphyomyrts". By employing combined BLASTN and UPGMA analyses of six chloroplast (cp) regions, three of the six unknown eucalypt samples (Euc4, 5, 6) cluster together with *E. microtheca* and *E. cladocalyx*, whereas the other three (Euc1, 2, 3) were more similar to a group containing *E. camaldulensis*, *E. grandis* and *E. saligna*. UPGMA analysis of the ITS region overall shows the same rough clustering, but provide more detailed information for two samples being most likely assigned to *E. camaldulensis*.

Keywords: *matK*, *rbcL*, intergenic linker, *psbA_matK*, *psbK_psbI*, *trnG_psbK*, *matK_trnK*, chloroplast, Internal transcribed spacer.

[1)] Thünen-Institute of Forest Genetics, Sieker Landstr. 2, D-22927 Grosshansdorf.

[2)] Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Km 5.5 Carretera Mazatlán, 34120 Durango, Mexico.

[*)] Author for correspondence: Matthias Fladung. E-Mail: matthias.fladung@thuenen.de

Introduction

The genus *Eucalyptus* is originally native to Australia, comprising more than 700 different species (Brooker, 2000). Most eucalypt species are evergreen and fast-growing, and prefer warm and dry climates. The wood of *Eucalyptus* is extremely hard and durable, and is used, e.g., in railway sleepers, fencing, and flooring. Due to their fast growth and beneficial wood properties, few eucalypt species were brought to different parts of the world, including Mexico (Ceccon and Martinez, 1999; Bernal-Lugo et al., 2009). For Mexico, a potential of about 11 Million hectares of eucalypt plantations is estimated (Ruiz et al., 2006).

Species of the genus *Eucalyptus* were introduced early in the 20th century to Mexico but conduction of controlled trials only began at Chapingo arboretum in 1948. Plantations of eucalypts, especially of *E. camaldulensis*, started in 1956 at El Rancho Casas Blancas, and covered about 2,000 ha in 1967. In 1967, seeds of several provenances of *E. camaldulensis*, *E. citriodora*, *E. tereticornis*, *E. botryoides*, *E. saligna*, *E. grandis*, *E. microtheca* and *E. deanei* were imported from Australia, together with seeds of 17 other eucalypt species. *E. camaldulensis* and *E. tereticornis* provenances from northern Queensland and Western Australia have been shown to grow fast. *E. citriodora* and *E. botryoides* also look promising for special uses (FAO, 1981).

Tree breeding programs for *Eucalyptus* in Mexico are still in its infancy. Today in large-scale plantations in Mexico, however, three species of eucalypts are basically used (*E. urophylla*, *E. grandis* and *E. saligna*). These plantations were established following selection and cloning of a small number of genotypes originated from natural populations (Couto and Betters, 1995; in Ceccon and Martínez, 1999). But unfortunately, information about provenances of the single eucalypt species is rare.

Further, many *Eucalyptus* species are cross compatible, thus hybridisation is possible and hybrids putatively exist. Morphologically, only in few cases, eucalypt trees can be determined as hybrids beyond a reasonable doubt. However, due to environmental conditions and genomic plasticity the phenotype is often extremely variable, e.g. in *E. globulus* (JORDAN et al., 1993) and *E. saligna* (PASSIOURA and ASH, 1993).

Molecular marker systems offer an appropriate tool for identification and taxonomic classification of organisms. Taxonomists or conservationists but also public authorities are increasingly applying molecular markers for species classification purposes (also called "DNA barcoding"; HEBERT et al., 2003; HOLLINGSWORTH, 2009). In the following years, "The International Barcode of Life project" (iBOL) and other genetic barcoding programs were founded, with the iBOL project nowadays representing one of the largest bio-diversity genomics initiatives (http://www.ibol.org/; http://www.barcodeoflife.org/). In principle, all sufficiently variable DNA regions that occur in the three subcellular compartments of the plant cell, nucleus (n), chloroplast (cp) and the mitochondria (mt), can be used for species differentiation and breeding purposes. As molecular marker systems, microsatellites or simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) have successfully been adopted in many plant and tree species (for review: ZIEGENHAGEN and FLADUNG, 2004; AGARWAL et al., 2008; but see also SCHROEDER and FLADUNG, 2010; HOELTKEN et al., 2012; SCHROEDER et al., 2012; SCHROEDER and FLADUNG, 2014).

Nuclear ribosomal internal transcribed spacers (ITS) are part of the ribosomal DNA (rDNA) region in the genome that codes for ribosomal RNA. In eukaryotes, the rDNA consists of 18S, 5.8S and 28S genes, interrupted by ITS1 (18S–5.8S) and ITS2 (5.8S–28S). The ITS1 and ITS2 marker have already successfully been used as a universal DNA barcode marker for, e.g., fungi (SCHOCH et al., 2012) and red macroalgae (HU et al., 2009). For plants, many of the phylogenetic hypotheses formulated include or are based exclusively on ITS sequences (ALVAREZ and WENDEL, 2003). This confirms the near-universal usage of ITS sequence data in plant systematics and phylogenetic studies. Alternatively for the differentiation of land plant species, molecular markers based on the DNA variation in two cp regions (*matK* and *rbcL*) have been proposed as a two-locus barcode (HOLLINGSWORTH et al., 2009). For *Eucalyptus*, whole chloroplast (cp) genome sequences are already available for 31 different species (http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=2759&opt=plastid).

In this study, we aim to assign the species of six randomly selected *Eucalyptus* trees planted in 1984 on an experimental plantation located in Northeast of Mexico (BRISEÑO-URIBE et al., 2015) by using the nuclear ITS region as well as six chloroplast regions, including the two highly-polymorphic "official" barcoding marker matK and rbcL, and four intergenic linker regions (*psbA_matK*, *psbK_psbI*, *trnG_psbK*, *matK_trnK*). All sequences are blasted against available information in public databases. The plantation was established to test survival, wood volume, and pulp production of the trees, and no records on species name and origin of plant material are available anymore.

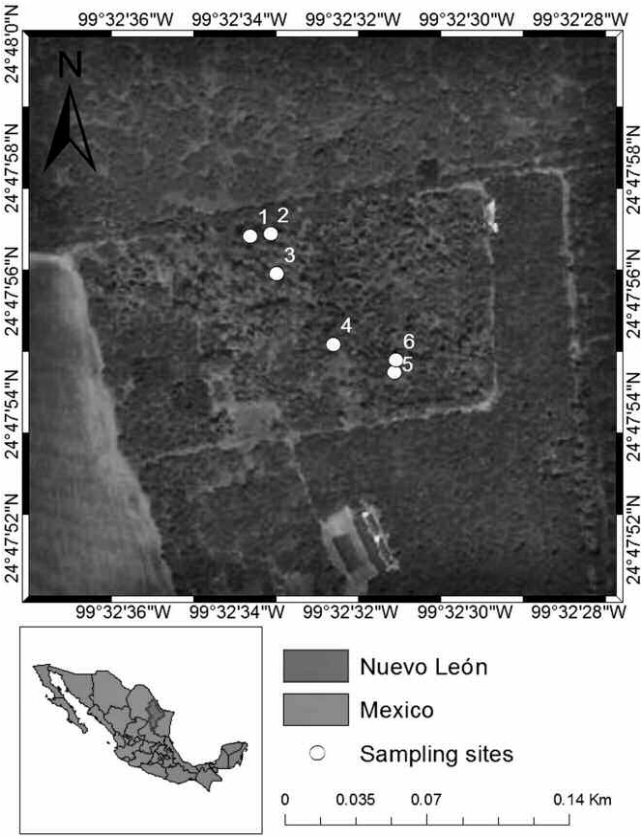

Figure 1. – Map of study area located in Northeast of Mexico (24° 47'N; 99° 32'W). Numbered dots indicate sampling positions of the three putative *Eucalyptus camaldulensis* (Euc1 to Euc3) and three putative *E. microtheca* (Euc4 to Euc6) trees. Source: ESRI Inc. (1999-2012). ArcGIS for Desktop 10. USDA Natural Resources Conservation Service.

Material and Methods

Study area, plantation system and plant material

The study area is located in Northeast Mexico (24°47'N; 99°32'W; *Figure 1*) at 350 m above sea level in a subtropical and semi-arid climate, with warm summer rain between April and November, and the presence of a period of summer drought, with temperatures up to 45 °C during the summer, appearing monthly averages of 14.7 °C in January to 22.3 °C in August; mean annual precipitation is 805 mm, with a bimodal distribution (YÁÑEZ-DÍAZ et al., 2014). The potential evapotranspiration is rated as 1,150 mm. The rainy season occurs in summer interrupted by a dry period (NAVAR and BRYAN, 1994). The native vegetation is known as Tamaulipan thornscrub. The origin of rocky-type soils is upper cretaceous being rich in calcite and dolomite. The dominant soils are deep, dark grey, lime-clay vertisol resulting from alluvial and colluvial processes with a pH of 7.5, electric conductivity of 101.6 µS cm^{-1}, organic matter content 6.0%, organic carbon 3.53%, and phosphorus 0.1 mg L^{-1}. The experimental plantation was established in 1984 with the objective to determine the species survival, wood volume, and expected production of pulp for paper industry. Putative *E. camaldulensis* and *E. microtheca* were planted in a grid of 3 × 3 m (1,111 plants ha^{-1}). Diameter at breast height varied from 0.15 (*E. camaldulensis* (EC)) to 0.19 m (*E. microtheca* (EM)), tree height from 9.46 (EC) to 10.54 m (EM), and tree volume varied from 0.02 (EC) to 0.06 m^3 (EM). Mean volume per hectare ranged from 19 (EC) to 71 m^3 ha^{-1} (EM). Growing rate values ranged from 0.76 (EC) to 2.73 m^3 ha^{-1} $year^{-1}$ (EM) (BRISEÑO-URIBE et al., 2015).

Samples were taken from three putative *E. camaldulensis* (Euc1-Euc3) and three putative *E. microtheca* trees (Euc4-Euc6) randomly distributed over the study area (red dots in *Figure 1*). Dried leaf samples of *E. microtheca* were kindly provided by L. Hasz, Arboretum UC Davis, Davis, USA.

Molecular methods

Genomic DNA was extracted from leaves of the six eucalypt trees under study and from dried *E. microtheca* leaves using the standard ATMAB protocol of DUMOLIN et al. (1995) and stored at –20 °C. Aliquots were diluted to a concentration of 10 ng/µL and stored at 4 °C until use. For the PCR reaction, forward and reverse primers were mixed with a concentration of 5 µM each. Sequences of the primer pairs amplifying parts of the six cp regions as well as the ITS nuclear region are given in *Table 1*. PCR reactions contained 2.5 µL 10 × reaction buffer BD (provided together with Taq-polymerase by DNA Cloning Service, Hamburg, Germany), 2.0 µL $MgCl_2$ (25 mM), 1.0 µL

Table 1. – List of primers used in the study.

Organelle / nucleus	Region*	Primer forward (5'-3')	Primer reverse (5'-3')
Cp	*psbA_matK*[1]	5'-CAG TGC TGG TTA TCC AAT TAC AG-3'	5'-CGT AAA CAC AAA AGT ACT GTA CG-3'
	psbK_psbI[1]	5'-CCA ATC GTA GAT TTT ATG CCA G-3'	5'-GGA TTA CGT CCT GGA TCA TTA G-3'
	trnG_psbK[1]	5'-GAA GGA TTC GAA CCT CCG AAT A-3'	5'-CTG GCA TAA AAT CTA CGA TTG G-3'
	matK_trnK[1]	5'-CGA AAG ATA AGT GGG TAT AGG-3'	5'-GCT AAC TCA ATG GTA GAG TAC-3'
	matK[2]	5'-CGT ACA GTA CTT TTG TGT TTA CGA G-3'	5'-ACC CAG TCC ATC TGG AAA TCT TGG TTC-3'
	rbcL[2]	5'-ATG TCA CCA CAA ACA GAG ACT AAA GC-3'	5'-GTA AAA TCA AGT CCA CCA CG-3'
rDNA	ITS	5'-CAT TGT CGA ATC CTG CCC AG-3'	5'-ATT CGC GGA GCG TGC GAT GA-3'

* Origin of primer sequences is indicated by reference: [1] SCHROEDER et al., 2012; [2] HOLLINGSWORTH et al., 2009.

Table 2. – ITS sequences from NCBI (http://www.ncbi.nlm.nih.gov/) included in the UPGMA analyses.

rDNA	*Eucalyptus* species	NCBI accession number(s)
ITS	*E. camaldulensis* Dehnh.	HQ995672, AF058473, AF100363, HM596038
	E. cladocalyx F. Muell.	EF488229
	E. deglupta Blume	DQ499111, AF190362, AF390518, AF390519, AF390521, HM596043, AF390520
	E. globulus Labill.	AY615680, AY615679, AY615678, AY615677, AY615676, AY615675, AY615674, AY615673, AY615672, AY615671, HM596049, EF694711, AF058468, AF058463
	E. grandis W.Hill ex Maiden	HM596050, AF390472, AF390471, AF058475
	E. melliodora A.Cunn. ex Schauer	AF390514
	E. radiata Sieber ex DC.	KM064958
	E. saligna Sm.	KM064960
	E. torquata Leuhm.	AF390499
	E. urophylla ST Blake	AF390489

dNTP-Mix (5 mM), 1.5 µL primer mix (5 µM), 0.5 µL dimethyl sulfoxide, 0.2 µL DNA polymerase (5 U/µL) and 0.8 to 1.0 µL DNA (100 ng/µL). The PCR reaction mixture was filled up to 25 µL with water. The PCR program was started with an initial denaturation for 3 min at 94°C. Forty PCR cycles followed, with 30 s at 94°C, 45 s at the respective annealing temperature, and 60 s at 72°C. The reaction was completed by a final elongation for 10 min at 72°C. Annealing temperatures were calculated on the basis of the primer sequences. Five µL of each PCR product were visualized on 1.2% agarose gel (120 V, 1:20 h) stained with the DNA fluorescence additive Roti-Safe Gel Stain (Carl Roth, Karlsruhe, Germany).

Unimorphic PCR products (one distinct band) were prepared as recommended and sequenced by StarSeq, Mainz, Germany. The PCR products were purified with lithium chloride (5 µL LiCl) before sequencing, and 130 µL of absolute ethanol were added to precipitate DNA overnight at –70°C. The obtained sequences were aligned with available eucalypt NCBI database sequences (*Table 2*, *3*) using the software SeqMan 12.2.0 (DNAStar, Madison, USA).

For *E. microtheca*, only one ITS sequence is available in NCBI public database, and neither a complete cp genome sequence nor any sequence of the cp regions studied could be found. Therefore, all six cp as well as the

Table 3. – cp regions *matK*, *rbcL*, *psbA_matK*, *psbK_psbI*, *trnG_psbK*, and *matK_trnK*, taken from complete genome sequences for different Eucalyptus species (NCBI; http://www.ncbi.nlm.nih.gov/), were included in UPGMA analyses.

Cp region	*Eucalyptus* species	NCBI accession number
Complete genome	*E. camaldulensis*	KC180791
	E. cladocalyx	KC180786
	E. deglupta	KC180792
	E. globulus	KC180787
	E. grandis	HM347959
	E. melliodora	KC180784
	E. radiata	KC180770
	E. saligna	KC180790
	E. torquata	KC180794

Table 4. – cp regions mat K, *rbcL*, *psbA_matK*, *psbK_psbI*, *trnG_psbK*, and *matK_trnK* for *E. microtheca* submitted to NCBI database were included in UPGMA analyses.

Eucalyptus species	Cp region	NCBI accession number
E. microtheca	*matK*	KT224661
	rbcL	KT224662
	psbA_matK	KT224664
	psbK_psbI	KT224665
	trnG_psbK	KT224666
	matK_trnK	KT224663
E. microtheca	ITS region	KT186365

nuclear ITS regions were PCR amplified and sequenced as described before. Sequences obtained were submitted to NCBI (accession numbers are given in *Table 4*) and included in the eucalypt sequence alignments (*Figs. 3* and *4*).

Bioinformatic analyses

BLASTN analyses were performed by using the obtained sequences as nucleotide query and submitting them to the NCBI webtool (http://blast.ncbi.nlm.nih.gov) (NCBI BLAST 2008).

To create phylogenetic trees, (1) the complete chloroplast genome sequences of all 31 *Eucalyptus* species currently available at the NCBI Organelle Genome Resources (http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=2759&opt=plastid; last page view June, 19, 2015) were aligned, (2) the six cp region sequences of the six *Eucalyptus* samples together with 10 selected *Eucalyptus* species available at NCBI (for selection criteria, see results and Discussion) were aligned, and (3) the ITS sequences of the six *Eucalyptus* samples together with eleven selected *Eucalyptus* species available from NCBI were aligned using the "Create Alignment"-tool of CLC Genomics Workbench (CLCbio, Aarhus, Denmark; v8.0.1). Based on the result of these multiple alignments, phylogenetic trees were created by computing distance estimates using UPGMA (Unweighted Pair Group Method with Arithmetic Mean) in combination with Jukes-Cantor distance correction ("Create tree"-tool of CLC Genomics Workbench. Bootstrapping with 100 replicates for the complete chloroplast sequences and 1,000 replicates for the other two alignments were performed.

Results and Discussion

Six different cp regions including the "official" barcoding regions *matK* and *rbcL*, successfully applied in *Populus* (SCHROEDER et al., 2012; SCHROEDER and FLADUNG, 2014) and *Swietenia* (HOELTKEN et al., 2012), have been tested for their suitability to assign the *Eucalyptus* species to six sample trees grown on the study area located in Northeast of Mexico (*Fig. 1*), putatively belonging to *E. camaldulensis* and *E. microtheca*. Following BLASTN analyses in the NCBI public database (http://www.ncbi.nlm.nih.gov/), inconsistent results were obtained for the potential assignment of the samples to an *Eucalyptus* species when comparing the different regions among each other (6 cp regions and ITS region; see subsections on BLASTN analyses). This was the reason why UPGMA analyses were also performed (see subsection on Phylogenetic tree analyses), however not including all eucalypt species as indicated from BLASTN analyses but a few selected species. The selected species included all those species mentioned in the introduction for which complete cp sequences or, at least, sequences of the regions of interest were available. For *E. microtheca*, the regions of interest were newly sequenced in this study. Additionally, species that represent major clades of *Eucalyptus* were selected based on a phylogenetic tree constructed from all completely sequenced *Eucalyptus* cp genomes at NCBI (see subsection on alignment of 31 eucalypt cp genomes).

BLASTN analyses of Eucalyptus sequences of six cp regions

The two official barcode marker *matK* and *rbcL* revealed to be rarely polymorphic in *Eucalyptus*. Overall, the different *Eucalyptus* individuals gave heterogeneous BLASTN analysis

Table 5. – Summary of the results of the BLASTN analyses for six cp regions in the six *Eucalyptus* samples. Additional details are described in the text.

Cp region	Euc1	Euc2	Euc3	Euc4	Euc5	Euc6
matK	No 100%-similarity to any *Eucalyptus* species	100% similar to 10 different species (e.g., *E. camaldulensis*, *E. globulus* and *E. saligna*)	100% similar to *E. urophylla*	100% similar to *E. bosistoana*	No 100%-similarity to any *Eucalyptus* species	100% similar to *E. bosistoana*
rbcL	100%-similar to *E. grandis*, *E. camaldulensis*, *E. deglupta*, *E. diversicolor*	100%-similar to *E. grandis*, *E. camaldulensis*, *E. deglupta*, *E. diversicolor*	100%-similar to *E. grandis*, *E. camaldulensis*, *E. deglupta*, *E. diversicolor*	100%-similar to *E. cladocalyx*, *E. polybractea*, *E. melliodora*	100%-similar to *E. cladocalyx*, *E. polybractea*, *E. melliodora*	100%-similar to *E. cladocalyx*, *E. polybractea*, *E. melliodora*
psbK-psbI	100%-similar to *E. melliodora*	No 100%-similarity to any *Eucalyptus* species	100%-similar to *E. camaldulensis*, *E. grandis*, *E. saligna*	100%-similar to *E. cladocalyx* and *E. melliodora*	No 100%-similarity to any *Eucalyptus* species	No 100%-similarity to any *Eucalyptus* species
psbA_matK	100%-similar to *E. grandis*, *E. cladocalyx*, *E. polybractea*	No 100%-similarity to any *Eucalyptus* species	100%-similar to *E. grandis*, *E. deglupta*, *E. polybractea*	100%-similar to *E. grandis*, *E. polybractea*, *E. deglupta*	100%-similar to *E. grandis*, *E. polybractea*, *E. deglupta*,	100%-similar to *E. grandis*, *E. polybractea*, *E. deglupta*,
matK_trnK	100%-similar to *E. cladocalyx*	No 100%-similarity to any *Eucalyptus* species	No 100%-similarity to any *Eucalyptus* species	No 100%-similarity to any *Eucalyptus* species	100%-similar to *E. cladycalyx*	No 100%-similarity to any *Eucalyptus* species
trnG_psbK	100%-similar to *E. cladocalyx*, *E. melliodora*, *E. polybractea*	100%-similar to *E. saligna*	100%-similar to *E. camaldulensis*, *E. grandis*, *E. saligna*	No 100%-similarity to any *Eucalyptus* species	100%-similar to *E. diversicolor*, *E. cladocalyx*, *E. melliodora*, *E. polybractea*	No 100%-similarity to any *Eucalyptus* species

results which are summarized in *Table 5*. The *matK* sequence of Euc1 revealed highest similarity to up to 15 different *Eucalyptus* species, but all with three mismatches at the same positions. For Euc5, highest similarity was found to *E. bosistoana* with one mismatch. The results of the second barcode marker *rbcL* are different to the *matK* data. The *rbcL* sequences of Euc1 to Euc3 and Euc4 to Euc6 are identical each, with 100%-similarity to three *Eucalyptus* species each (*Table 5*).

Also for the four other cp regions, the situation seems confusing (*Table 5*). In Euc1, the *psbK_psbI* sequence was identical only to *E. melliodora*, whereas *psbA_matK* was identical to three other *Eucalyptus* species, or *matK_trnK* only to *E. cladocalyx*. In Euc2, only *trnG_psbK* was 100% similar to *E. saligna*. The other three cp regions showed highest similarity (one to three mismatches) to *E. camaldulensis* and *E. grandis*. In Euc3, three cp regions were 100% similar to at least one *Eucalyptus* species (*Table 5*), while *matK_trnK* showed one mismatch to *E. camaldulensis*. In Euc4, two cp regions were 100% similar to at least one *Eucalyptus* species (*Table 5*), while for *matK_trnK* nine mismatches and eight gaps were found to the best hits (*E. cladycalyx*, *E. deglupta*, *E. melliodora*, and *E. polybractea*), and for *trnG_psbK* three mismatches to *E. cladocalyx*, *E. melliodora*, and *E. polybractea*. Three cp regions were 100% similar to at least one *Eucalyptus* species in Euc5 (*Table 5*), but the *psbK_psbI* sequence is highly similar to *E. cladocalyx*, *E. melliodora*, and *E. polybractea* (with one mismatch each). Finally, Euc6 revealed 100%-similarity only for *psbA_matK* (*Table 5*), but for *psbK_psbI* one mismatch to the best hits (*E. cladocalyx* and *E. melliodora*), for *matK_trnK* nine mismatches and eight gaps to

the best hits (*E. cladocalyx*, *E. deglupta*, *E. melliodora*, and *E. polybractea*), and for *trnG_psbK* one mismatch to *E. cladocalyx*, *E. melliodora*, and *E. polybractea*.

BLASTN analyses of Eucalyptus Internal transcribed spacer (ITS) sequences

Blast analyses of ITS sequences of the six eucalypt samples against NCBI database also revealed no uniform results. The ITS sequences of Euc1 and Euc3 are differing in only 1 SNP, and are highly similar to *E. camaldulensis* (HM596038) and *E. glaucina* (HM596048) (with one mismatch each), but surprisingly also with one mismatch to *Terminalia bellirica* (FM887018), a member of the *Combretaceae* family. For Euc2, highest similarities were found to *E. lockyeri* (AF390488; with one mismatch), *E. brassiana* (HM596037) and *E. tereticornis* (AF390482), both with two mismatches, followed by *E. camaldulensis* (HM596038), *E. glaucina* (HM596048) and *E. vicina* (HM116971) with three mismatches. Euc4 seems to be a hybrid between two eucalypt species, because of sequence overlapping. Not overlapping sequences showed highest similarities 97 to 98 % to *E. porosa* (HM116969) and *E. melliodora* (KP142207). Highest similarities of Euc5 and Euc6 were found to *E. porosa* (HM116969) with four mismatches and to *E. woodwardii* (AF058479) with eight mismatches.

Alignment of 31 eucalypt cp genomes available from NCBI database

A phylogenetic tree constructed from complete chloroplast sequences of 31 *Eucalyptus* species resolved 5 major clades (*Fig. 2*). Three main groups previously described (Johnson and Briggs, 1984; Ladiges et al., 1995) were recovered: the "Monocalypt" group (Monocalypts; clades 1 and 4), the "Symphyomyrt" group (Symphyomyrts; clades 2 and 3), and the "Eudesmid" group (Eudesmids; clade 5). These groups were also resolved in previous phylogenetic studies based on chloroplast sequences (Bayly et al., 2013) or ITS sequences (e.g., Steane et al., 1999). The study of Bayly et al. (2013) basing on complete chloroplast genome sequences of the same 31 *Eucalyptus* species; provided similar results as our study with two main differences: All species included in the small Symphyomyrts$_B$ clade (clade 3; *Fig. 2*) were assigned to one large Symphyomyrts clade by Bayly et al. (2013) where they form a subclade of earliest divergence within the clade. *E. curtisii* that was assigned to the Monocalypts-representing clade as a sister of all other members of this clade (Bayly et al., 2013) resolved as a separate clade (clade 4) in our analysis. It is not unexpected as *E. curtisii* is morphologically considered as an unusual relictual species (Bayly et al., 2013). The described differences of the results between both studies may be due to the different tree

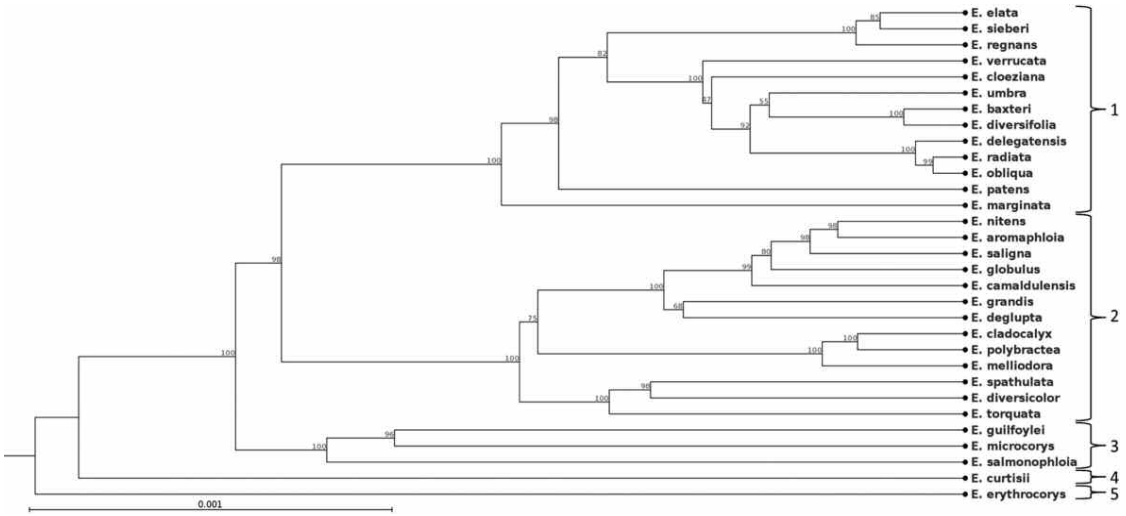

Figure 2. – UPGMA (Unweighted Pair Group Method with Arithmetic Mean) analysis including the complete chloroplast sequences of 31 *Eucalyptus* species. The five clades were named based on the assignment of the related species to the main *Eucalyptus* groups as defined previously (Johnson and Briggs, 1984; Ladiges et al., 1995): "Monocalypts$_A$" (clade 1), "Symphyomyrts$_A$" (clade 2), "Symphyomyrts$_B$" (clade 3), "Monocalypts$_B$" (clade 4), "Eudesmids" (clade 5). Bootstrap support values (%) are shown above branches.

construction methods applied and the inclusion of additional species of other genera of the eucalyptus group (*Corymbia* and *Angophora*) and of outgroups in the study of BAYLY et al. (2013).

The phylogenetic tree (*Fig. 2*) includes 3 of the 9 species known to be introduced to Mexico, namely *E. camaldulensis*, *E. saligna* and *E. grandis*, which belong to the main Symphyomyrts clade (clade 2, *Fig. 2*). These three species together with *E. radiata* that represents the Monocalypts (clade 1, *Fig. 2*) were included beside others in the further analyses.

Phylogenetic tree analyses including all six cp marker

To obtain a visual picture, UPGMA were performed for all six cp marker (*matK*, *rbcL*, *psbA-matK*, *psbK-psbI*, *matK-trnK*, *trnG-psbK*) together (*Fig. 3*). *E. radiata* appeared to cluster as a kind of outgroup (*Fig. 3*) which is in consensus with the clades of 31 *Eucalyptus* species (*Fig. 2*), because *E. radiata* belongs to the "Monocalypt" group, whereas all other species used in this study belong to the "Symphyomyrt" group. Within the "Symphyomyrt" group in

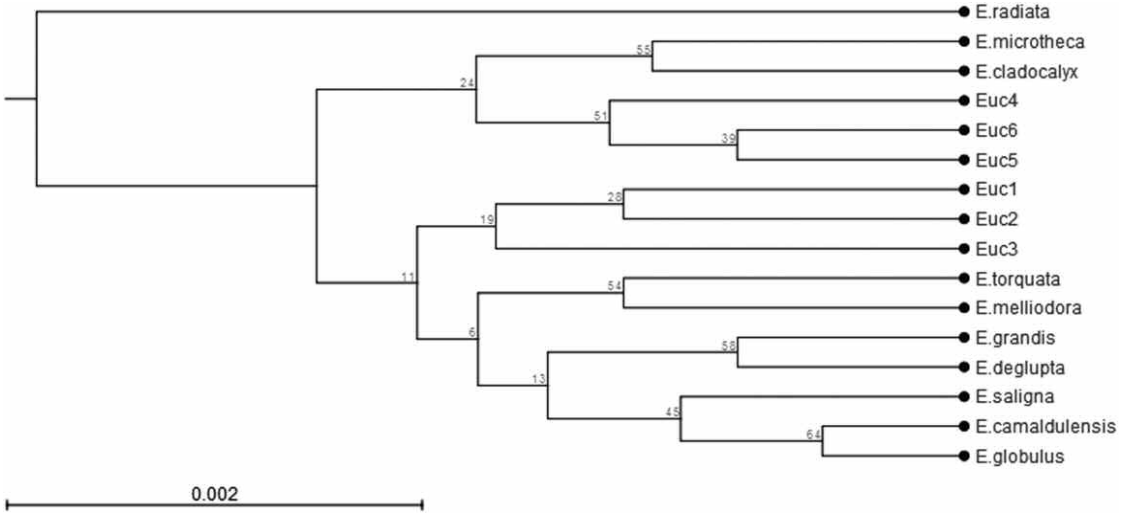

Figure 3. – UPGMA (Unweighted Pair Group Method with Arithmetic Mean) analysis including all six cp marker of the six eucalypt samples. Bootstrap support values (%) are shown above branches.

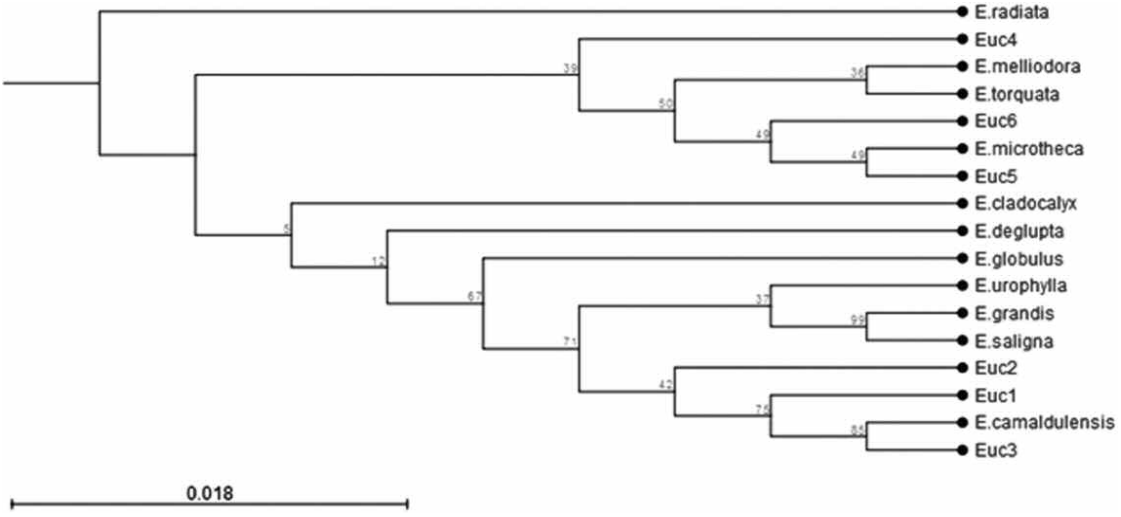

Figure 4. – UPGMA (Unweighted Pair Group Method with Arithmetic Mean) analysis of the ITS sequences of the six eucalypt samples. Bootstrap support values (%) are shown above branches.

Fig. 3, two further clusters are to be seen assigning the three unknown samples Euc4, 5 and 6 to *E. cladocalyx* and *E. microtheca* and Euc1, 2 and 3 to a group containing seven eucalypt species including *E. camaldulensis*. This result is in accordance with the BLASTN analyses of all single cp regions given above. Overall, none of the analysed samples revealed to be consistently assigned to any of the ten *Eucalyptus* species used for this study in all six cp regions.

Taken together, combined BLASTN (versus the public NCBI database) results of sequences of the two official cp barcode marker *matK* and *rbcL* obtained from all six eucalypt samples investigated in this study did not reveal any clear indication to one eucalypt species. Adding additional four cp markers to the analyses increased the confusion. Also the UPGMA analyses did not provide a clear picture about the assignment of the six eucalypt samples to any eucalypt species. Either the cp markers are not sufficient for species determination within the genus *Eucalyptus* or the public databases did not contain cp sequences of all eucalypt species existing in Mexico. The latter option is in particular evident for *E. urophylla*, i.e., no complete cp genome has been sequenced so far from this species, and only sequences for *rbcL* and *matK* are available in NCBI, but not for the other four cp regions. Thus, this species could not be included in the analyses.

Phylogenetic tree analyses using ITS marker

Altogether, phylogenetic tree analysis based on the nuclear ITS region gave similar clusters as for the cp regions. Again, *E. radiata* is an outgroup, and the three samples Euc1, 2 and 3 are in one clade and the other three samples Euc4, 5 and 6 are in another (*Fig. 4*). The three samples Euc1, 2 and 3 cluster together and build a subclade with *E. camaldulensis*. The difference between *E. camaldulensis* and Euc1 is based on two heterozygote sites and on one further SNP, and Euc2 is distinguished from *E. camaldulensis* by the same two heterozygote sites plus three further SNPs. From the sample Euc3, *E. camaldulensis* differed in two heterozygote sites only (*Fig. 4*). The cluster with Euc4, 5 and 6 is more heterogeneous because of the higher number of SNPs between the three samples Euc4, 5 and 6, and other *Eucalyptus* species. The two species *E. melliodora* and *E. torquata* cluster together with *E. microtheca* and *E. cladocalyx*.

Thus, considering that all reference species declared as different species used in this study have shown three and more SNPs between them, it is most likely that Euc3 is assigned to *E. camaldulensis* and maybe also Euc1 can be considered belonging to this species. For the samples Euc4, 5 and 6 it is not possible to assign them to one of the *Eucalyptus* species considered in this study.

Conclusions

Unfortunately, the tested cp marker regions (including the two official barcode marker) could not sufficiently unravel species affiliation of the six eucalypt samples (Euc1 to Euc6) from the study area located in Northeast of Mexico. By UPGMA analysis of the 6 cp regions, however, Euc1, 2, 3 could be assigned to a cluster with *E. camaldulensis*, *E. grandis* and *E. saligna*, and Euc4, 5, 6 to a cluster with *E. microtheca* and *E. cladocalyx*. UPGMA analysis of the ITS region gave more detailed information for two samples, assigning Euc1 and Euc3 most likely to *E. camaldulensis*, however, Euc2, and Euc4 to 6 cannot per-se be assigned to one of the *Eucalyptus* species available in public databases. DNA barcoding is a useful tool for molecular systematics, however, cannot be applied to unravel phylogenetic relationships in *Eucalyptus*.

Acknowledgements

We deeply thank KATRIN GROPPE for helpful technical assistance in the lab, ARTEMIO CARRILLO-PARRA, Mexico, for sampling of Euc1 to Euc6, and L. HASZ, Arboretum UC Davis, Davis, USA, for providing leaf samples for *E. microtheca*.

References

AGARWAL, M., N. SHRIVASTAVA and H. PADH (2008): Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reports **27**: 617–631.

ÁLVAREZ, I. and J. F. WENDEL (2003): Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution **29**: 417–434.

Bayly, M. J., P. Rigoult, A. Spokevicius, P.Y. Ladiges, P. K. Ades, C. Anderson, G. Bossinger, A. Merchant, F. Udovicic, I. E. Woodrow and J. Tibbits (2013): Chloroplast genome analysis of Australian eucalypts – *Eucalyptus*, *Corymbia*, *Angophora*, *Allosyncarpia* and *Stockwellia* (*Myrtaceae*). Molecular Phylogenetics and Evolution **69**: 704–716.

Bernal-Lugo, I., C. Juárez Palacios, R. Santos-Gally, L. Vázquez-Silva, H. A. Zavaleta-Mancera, H. S. Azpiroz-Rivero, C. Parraguirre-Lezama and M. D. P. D. Garza-López de Lara (2009): Identificación del agente patógeno del cancro del eucalipto en plantaciones del Sureste de México. Ciencia forestal en México **34**: 17–35.

Briseño-Uribe, K. C., A. Carrillo-Parra, V. Bustamante-García, H. González-Rodríguez and R. Foroughbachk (2015): Firewood production, yield and quality of charcoal from *Eucalyptus camaldulensis* and *E. microtheca* planted in the semiarid land of northeast Mexico. International Journal of Green Energy **12**: 961–969.

Brooker, M. I. H. (2000): A new classification of the genus *Eucalyptus* L'Hér. (*Myrtaceae*). Australian Systematic Botany **13**: 79–148.

Ceccon, E. and M. R. Martínez (1999): Aspectos ambientales referentes al establecimiento de plantaciones de eucalipto de gran escala en áreas tropicales: aplicación al caso de México. Interciencia **24**: 352–359.

FAO (1981): El eucalipto en la repoblación forestal. FAO: estudios de silvicultura y productos forestales No.11. FAO, Roma. http://www.fao.org/DOCREP/004/AC459S/AC459S00.html: 03/18/2015.

Fritz C. (2015): *Eucalyptus* Trees in Mexico. http://www.gardenguides.com/112509-eucalyptus-trees-mexico.html (last page view March 3, 2015).

Hebert, P. D. N, A. Cywinska, S. L. Ball and J. R. de Waard (2003): Biological identifications through DNA barcodes. Proceedings of the Royal Society, Series B, Biological Science **270**: 313–321.

Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei and et al. (2009): A DNA barcode for land plants. Proceedings of the National Academy of Sciences USA **106**: 12794–12797.

Hoeltken, A. M., H. Schroeder, N. Wischnewski, B. Degen, E. Magel and M. Fladung (2012): Development of DNA-based methods to identify CITES-protected timber species: A case study in the *Meliaceae* family. Holzforschung **66**: 97–104.

Hu, Z., M. D. Guiry and D. Duan (2009): Using the ribosomal internal transcribed spacer (ITS) as a complement marker for species identification of red macroalgae. Hydrobiologia **635**: 279–287.

Johnson, L. A. S. and B. G. Briggs (1984): *Myrtales* and *Myrtaceae* – a phylogenetic analysis. Annals of the Missouri Botanical Gardens **71**: 700–756.

Johnson, M., I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis and T. Madden (2008): NCBI BLAST: a better web interface. Nucleic Acids Res **36** (Web Server issue): W5-9.

Jordan, G. J., B. M. Potts, J. B. Kirkpatrick and C. Gardiner (1993): Variation in the *Eucalyptus globulus* complex revisited. Australian Journal of Botany **41**: 763–785.

Ladiges, P.Y., F. Udovicic and A. N. Drinnan (1995): Eucalypt phylogeny – Molecules and morphology. Australian Systematic Botany **8**: 483–497.

Návar, J. and R. B. Bryan (1994): Fitting the analytical model of rainfall interception of Gash to individual shrubs of semi-arid vegetation in northeastern Mexico. Agricultural and Forest Meteorology **68**: 133–143.

Passioura, J. A. and J. E. Ash (1993): Phenotypic, Genetic and Ecological Variation in the *Eucalyptus saligna*, *E. botryoides* Complex. Australian Journal of Botany **41**: 393–412.

Ruiz, R. M., H. S. Azpíroz-Rivero, V. M. Cetina-Alcalá and M. A. Gutiérrez Espinoza (2006): Importancia de las plantaciones forestales de *Eucalyptus*. Ra Ximhai: Revista Científica de Sociedad, Cultura y Desarrollo Sostenible **2**: 815–846.

Schoch, C. L., K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen and Fungal Barcoding Consortium (2012): Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proceedings of the National Academy of Sciences **109**: 6241–6246.

Schroeder, H. and M. Fladung (2010): SSR and SNP markers for the identification of clones, hybrids and species within the genus *Populus*. Silvae Genetica **59**: 257–262.

Schroeder, H. and M. Fladung (2014): Differentiation of *Populus* species by chloroplast SNP markers for barcoding and breeding approaches. iForest Doi: 10.3832/ifor1326-007

Schroeder, H., A. M. Hoeltken and M. Fladung (2012): Differentiation of *Populus* species using chloroplast single nucleotide polymorphism (SNP) markers – essential for comprehensible and reliable poplar breeding. Plant Biology **14**: 374–381.

Steane, D. A., G. E. McKinnon, R. E. Vaillancourt and M. M. Potts (1999): ITS sequence data resolve higher level relationships among the eucalypts. Molecular Phylogenetics and Evolution **12**: 215–223.

Yáñez-Díaz, M. I., I. Cantú-Silva, H. González-Rodríguez and J. I. Uvalle-Sauceda (2014): Redistribución de la precipitación en tres especies arbustivas nativas y una plantación de eucalipto del noreste de México. Tecnología y Ciencias del Agua **2**: 71–84.

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