Solomon Benor Belay

Phylogeny of the Genus *Corchorus* (Malvaceae s.l.) and Diversity Analyses in Selected Species: Evidence from Morphology, Flow Cytometry, and Molecular Data

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

Phylogeny of the Genus *Corchorus* (Malvaceae s.l.) and Diversity Analyses in Selected Species: Evidence from Morphology, Flow Cytometry, and Molecular Data

Dissertation
Zur Erlangung des Grades
Doktor rerum naturalium (Dr.rer.nat)
Ökologische Agrarwissenschaften

an der
Universität Kassel

Vorgelegt von
Solomon Benor Belay
Diploma, B.Sc, M.Phil.

Gutachter
Prof Dr Karl Hammer
Dr Frank R. Blattner
Prof Dr Kurt Weising
Prof Dr Rainer Georg Joergensen

Kassel, 21 October 2011

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2011
Zugl.: Kassel, Univ., Diss., 2011

978-3-86955-924-7

Citation:

Benor S. 2011. *Phylogeny of the genus Corchorus (Malvaceae s.l.) and diversity analyses in selected species: evidence from morphology, flow cytometry, and molecular data*. Ph.D. dissertation, Kassel University, Germany.

Printed with support from the DAAD.

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2011
Gedruckt auf säurefreiem Papier

978-3-86955-924-7

Table of contents

List of Tables

List of Figures

List of publications from the PhD project

This dissertation is prepared in chapters referring to the following series of papers which are either published or in preparation to internationally referred journals.

1. Benor S, Blattner FR, Demissew S, Hammer K, 2010. Collection and ethnobotanical investigation of *Corchorus* species in Ethiopia: potential leafy vegetables for dry region. Genetic Resources and Crop Evolution. 57: 293-306. doi: 10.1007/s10722-009-9470-y.
2. Benor S, Fuchs J, Blattner FR, 2011. Genome size variation in *Corchorus olitorius* (Malvaceae s.l.) and correlation with elevation and phenotypic traits. Genome 54: 575-585. doi:10.1139/g11-021.
3. Benor S, Demissew S, Hammer, K, Blattner FR 2011. Genetic diversity and relationships in *Corchorus olitorius* (Malvaceae s.l.) inferred from molecular and morphological data. Genetic Resources and Crop Evolution. doi: 10.1007/s10722-011-9748-8
4. Benor S, Hammer K, Blattner FR. Molecular phylogeny and biogeography of the genus *Corchorus* (Grewioideae, Malvaceae s.l.) based on nuclear rDNA ITS sequences. Manuscript in preparation.

Dedication

This dissertation is dedicated to my newborn son, Nathanael Solomon, who was born two months before my PhD disputation.

“ደህን ልጅ እንዲሰጠኝ ጸለይሁ፤ እግዚአብሔርም የለመንሁትን ሰጠኝ። ስለዚህ እኔም ለእግዚአብሔር አሰጠዋለሁ፤ በሕይወት ዘመኑም ሁሉ ለእግዚአብሔር የተሰጠ ይሆናል።” 1ኛ ሳሙ: 27-28.

Chapter 1
General Background

1.1 Introduction to molecular systematics

Living organisms are all connected by ancestor-to-descendent relationships. Phylogeny, defined as the historical relationships among lineages of organisms, attempts to reveal the ancestral and evolutionary relationships of species or higher taxonomic groups of organisms (Williams and Ebach 2008). In the past, taxonomists have grouped organisms based mainly on morphological characters. Nowadays, DNA sequences are widely used to infer phylogenies of organisms. Morphological (or anatomical) and DNA sequence data are used to classify biological living organisms through phenetics or cladistics approaches. Phenetics advocates grouping by overall or raw phenotypic, anatomical or molecular similarities. Thus, phenetics recognizes taxa on the basis of percentage similarity, for the most part ignoring other phylogenetic considerations. This approach is used in my study and shown in Chapter 5. Whereas cladistics, which is among others employed in Chapter 3, advocates grouping by shared derived characters (Williams and Ebach 2008).

Amino acid or nucleotide sequences are the most widely and frequently used markers of making phylogenetic inferences, called molecular phylogenetics (also called molecular systematics). The internal transcribed space (ITS) region of nuclear ribosomal DNA is one of the most important molecular markers that have been used in numerous systematic studies, including in Chapter 3 of this dissertation. It consists of two internal spacers, ITS1 and ITS2 (Figure 1.1), located between genes encoding the 18S, 5.8S and 26S nuclear ribosomal DNA (nrDNA) subunits (Baldwin 1992). DNA sequences of the ITS region are preferred in phylogenetic analysis due to several reasons. The ITS region is highly repeated in plant nuclear genomes and its high copy number and high degree of homogeneity facilitates easy amplification and sequencing of the nrDNA from small quantities of DNA. It is also relatively small and is flanked by highly conserved sequences, the 18s and 26s nrDNA genes which aid the use of universal primers to amplify and sequence the ITS regions in most plant families (Baldwin et al. 1995). The ITS region within individuals of a single species mostly reveals little or no sequence divergence. On the other hand, alignments from different species, and although not always, between closely related species show good level of sequence divergence (Baldwin et al. 1992, Baldwin et al. 1995). These characteristics, especially the low-intra species sequence variability and inter-species sequence divergence make ITS one of the candidate genes for DNA barcoding, an

approach to determine species by specific DNA sequences (Stoeckle 2003, Kress et al. 2005, Chen et al. 2010).

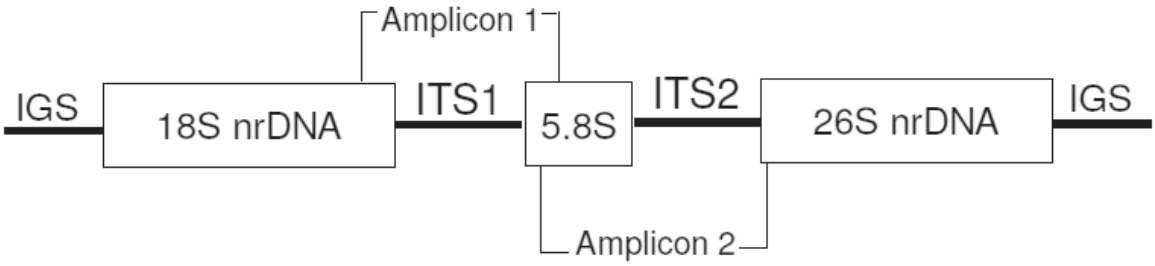

Figure 1.1. The internal transcribed spacer region of the nuclear ribosomal DNA.

1.2 Population genetics and diversity studies in plants

Population genetics is a branch of genetics which deals with the frequency of alleles responsible for genetic relationships or variations within populations or at species level. It considers the factors that determine the population size, the evolution of a population, such as natural selection, genetic drift, mutation, recombination, and gene flow (Halliburton 2004). Through population genetics tools, it is possible to study the naturally occurring genetic differences among organisms. Studying the level and patterns of genetic diversity is useful in several ways. At the most basic level, genetic diversity in a species is the raw stuff on which evolutionary processes are acting (May 2000). From a biodiversity point of view, genetic diversity studies help to generate information on origin and maintenance of the global set of species. Knowledge of the amount and distribution of genetic diversity is among the major criteria in locating sites for genetic reserves (Iriondo et al. 2008). Genetic variability in these reserves can also contribute in identification of genetic traits that foster tolerance to biotic or abiotic stresses.

In the field of population genetics, the level of genetic diversity of a given organism is usually investigated using molecular or morphological approaches. One of the methods frequently used in the molecular approaches, used in my study in Chapter 5, are amplified fragment length polymorphisms (AFLP) (Vos et al. 1995), which are used in several research areas, including the inference of phylogenetic relationships, investigation of genome wide genetic variation, individual ancestry, population structure, and in identifying hybrid individuals (Bensch and Akesson 2005, Falush et al. 2007, Pleines and Blattner 2008).

Morphological characterization is also used to investigate diversity of germplasm collections and contributes valuable information for agronomic and breeding programs (Chapter 5). In the field of plant systematics, morphometric analysis is also used to investigate morphological evolution, which is useful for

addressing species-level taxonomic questions. An understanding of evolutionary processes at the infraspecific level, and ultimately of speciation, requires the accurate characterization of patterns of variation within species (Gould and Johnston 1972, Thorpe 1976). One way of doing this is through multivariate morphometric techniques, which have the power to reveal complex patterns of population differentiation related to geographic divergence and evolutionary processes such as gene flow, selection and genetic drift (Gould and Johnston 1972).

1.3 Genome size variation in plants

The amount of DNA, including the non-coding DNA present within one copy of a genome of an organism is known as genome size (Gregory 2002). It is usually expressed in base pairs (bp) after converting the nuclear DNA content measured in picograms (pg). Nowadays, flow cytometry is the most widely used method of estimating nuclear DNA content because it is convenient, fast and reliable (Doležel et al. 2007). Genome size differences among species has been a puzzle in genetics for many decades, as relatively simple organisms could have 1000 times more nuclear DNA than more complex multicellular organisms. This phenomenon is known as the C-value paradox or C-value enigma (Gregory 2002), and could be explained by large fractions of non-coding DNA, which contribute to the total DNA content without increasing gene numbers. However, probable functional properties of non-coding DNA in living organisms are still an open question.

Estimation of nuclear DNA content can help to investigate ploidy level, which is useful in population biology, breeding, and quality control in seed production (Suda et al. 2007). Plants exhibit remarkable variation in genome size (Bennett et al. 2000, Bennett and Leitch 2005, Greilhuber et al. 2007). As a result, genome size has also been found to be a useful taxonomic marker and is employed in the field of systematics of various plant groups (Dimitrova et al. 1999, Mishiba et al. 2000, Zonneveld 2001, Doležel et al. 2007). This dissertation includes results of flow cytometric investigation of genome size in jute species, and its association with phenotypic traits and elevation is discussed in Chapter 4.

1.4 Systematics, distribution, and economic importance of jute

The genus *Corchorus* was first described by Linnaeus (1753) who recognized four species, namely *C. olitorius, C. capsularis, C. siliquosus*, and *C. hirsutus*. Later, several species belonging to this genus were identified and described by various authors. The genus was formerly classified under the traditional family Tiliaceae. However, this family is no longer recognized by the Angiosperm Phylogeny Group since 1998 (APG I 1998, APG II 20003, APG III 2009). Based on molecular evidence of chloroplast genes, *Corchorus* was included in the subfamily Grewioideae, which belongs to the family Malvaceae s.l. of the order Malvales (Alverson et al. 1999, Bayer et al. 1999, Whitlock et al. 2003). A reclassification of the broader Malvaceae was proposed by Heywood et al. (2007) and *Corchorus* was described within the family Sparrmanniaceae. However, this classification is not recognized by the Angiosperm Phylogeny Group (APG III 2009), thus the genus *Corchorus* at present maintains its phylogenetic position within the subfamily Grewioideae of the family Malvaceae.

The genus *Corchorus*, commonly known as jute, consists of shrubs, subshrubs, and annual and perennial herb species. Average plant height varies from ca. 0.09 m to over 4 m, the smallest recorded in *C. depressus* and *C. pumilio*, whereas the tallest are in the two cultivated species *C. olitorius* and *C. capsularis*. The leaves, mostly with palmate veins and serrated margins are simple, alternate and the shape varies from lanceolate, elliptical, ovate to palmate. Most species are morphologically similar and identification at vegetative phase is extremely difficult. Although shared by several species, a careful examination of the presence or absence, and the color of stipules are few of the morphological characters that help identify species at vegetative stage. The presence or absence of setae, a tail-like structure at the base of the blade, can also aid in species identification during the vegetative or reproductive stage (Figure 2.4). Stems are well developed with abundant fibres in the phloem tissue, which make several species a source of natural fibre, jute. Flowers (3-5-merus) are yellow usually with 4-5 free, acuminate sepals. Identification of *Corchorus* species is mainly possible during the reproductive stage, especially using capsule and seed morphology. Fruit (except in *C. capsularis*) is a straight or slightly curved, cylindric or spherical capsule that terminates in a beak with 2-7 valves (Figure 1.2). Seeds are mainly angular or ellipsoid in shape and the color varies from black, brown to gray (Figure 1.3). Most *Corchorus* species are diploids with 2n = 14, while few species, namely *C. cunninghamii, C. hirtus, C. junodii, C. orinocensis*, and *C. siliquosus* are tetraploids (Coleman 1982, Halford 1995, Islam et al. 1975, Rao and Datta 1953). Hexaploid (2n = 42) and aneuploid (2n = 13) species were also