2011

Molecular Genetic Studies on Complex Evolutionary Processes in *Archaesolanum* (Solanum, Solanaceae)

Péter K. Poczai

University of Pannonia

UNIVERSITY OF PANNONIA

GEORGIKON FACULTY

DOCTOR OF PHILOSOPHY

(PHD)

THESIS

PÉTER K. POCZAI

KESZTHELY, HUNGARY

2011

UNIVERSITY OF PANNONIA

GEORGIKON FACULTY

DOCTORAL SCHOOL IN CROP PRODUCTION AND HORTICULTURAL SCIENCES

PLANT GENETICS AND BIOTECHNOLOGY PROGRAMME

HEAD OF THE DOCTORAL SCHOOL

PROF. RICHARD GÁBORJÁNYI, DSC

MOLECULAR GENETIC STUDIES ON COMPLEX EVOLUTIONARY PROCESSES IN *ARCHAESOLANUM* (*SOLANUM*, SOLANACEAE)

DOCTOR OF PHILOSOPHY (PHD) THESIS

WRITTEN BY

PÉTER K POCZAI
DOCTOR OF PLANT HEALTH (MSC)

SUPERVISIOR

DR. JÁNOS TALLER, PHD

KESZTHELY, HUNGARY

2011

MOLECULAR GENETIC STUDIES ON COMPLEX EVOLUTIONARY PROCESSES IN *ARCHAESOLANUM* (*SOLANUM*, SOLANACEAE)

Írta:
Poczai K. Péter

Készült a Pannon Egyetem Növénytermesztési és Kertészeti Tudományok Doktori Iskolája, Növénynemesítés, Genetika és Agrárbiotechnológia Alprogramja keretében

Témavezető: Dr. Taller János

Elfogadásra javaslom (igen / nem)

(aláírás)

A jelölt a doktori szigorlaton …. %-ot ért el,

Az értekezést bírálóként elfogadásra javaslom:

Bíráló neve: …..... …..... igen /nem

…….
(aláírás)

Bíráló neve: …..... …..... igen /nem

…….
(aláírás)

A jelölt az értekezés nyilvános vitáján ….....%-ot ért el.

Keszthely, …….
a Bíráló Bizottság elnöke

A doktori (PhD) oklevél minősítése….....

……
Az EDHT elnöke

TABLE OF CONTENT

ABSTRACT

Molecular genetic studies on complex evolutionary proceses in *Archaesolanum* (*Solanum*, Solanaceae)

Kangaroo apples, subgenus *Archaesolanum*, are a unique and still poorly known group within the genus *Solanum*. The subgenus is composed of eight species with a characteristic chromosome number based on n = x = 23 and distribution restricted to the South Pacific. This subgenus is an isolated group of *Solanum*, and its phylogenetic relationships are still poorly known. This study represents an approach to analyze genetic relationships within this group. In this context, seven species were examined in a pilot study using random amplified polymorphic DNA (RAPD) as well as start codon targeted (SCoT) and intron targeting (IT) markers. In subsequent analysis, the amplification products of two chloroplast regions (*trn*S-*trn*G and *rbc*L) were studied with polymerase chain reaction (PCR) restriction fragment length polymorphism (RFLP) method. Screening for the presence of unique mitochondrial rearrangements was also carried out using universal mitochondrial primers for the detection of fragment length polymorphisms. The pilot study revealed two major groups within the subgenus; one was composed of the members of ser. *Avicularia* and *Laciniata*, while the other was formed by species belonging to ser. *Similia*. It is suggested that the taxonomic status of series within the *Archaesolanum* clade should be revised. The hybrid origin of *S. laciniatum* was also studied, and two hypotheses regarding its phylogeny are presented.

In further studies, we aimed to reveal phylogeny, historical biogeography and age of diversification of *Archaesolanum*. We sampled all recognized species of the group and sequenced three chloroplast regions, the *trn*T-*trn*L spacer, *trn*L intron and *trn*L-*trn*F spacer to calibrate a molecular clock to estimate the age of the group. Distributional data were combined with the results of phylogenetic analysis to track the historical processes responsible for the current range of the group. Our analysis supported the monophyly of the kangaroo apples and the biogeographical disjunction between the two subclades within the group. Based on the divergence time estimates the most recent common ancestor of kangaroo apples is from the late Miocene age (~ 9 MY). Based on the age estimate the common ancestors of the kangaroo apples are presumed to have arrived in Australia by long-distance dispersal. The two distinct lineages within the group have most likely separated during the aridification of the continent and further speciated in the brief resurgence of rainforests during the Pliocene.

KIVONAT

Archaesolanum (*Solanum*, Solanaceae) fajok komplex evolúciós folyamatainak molekuláris genetikai vizsgálata

A kenguru-almák, *Archaesolanum* subgenus, a *Solanum* genus egyedi és kevéssé ismert csoportja. A subgenusba nyolc, kizárólag a csendes-óceáni térségben előforduló faj tartozik, melyek közös jellemzője az n = x = 23 alap kromoszóma szám. Ezen izolált *Solanum* fajokkal kapcsolatban molekuláris genetikai módszerekre alapozott filogenetikai ismereteink teljes mértékben hiányoznak. Jelen tanulmány ezen ismeretek bővítést hivatott szolgálni. Elő kísérleteink során a csoport hét faját vizsgáltuk RAPD (Random Amplified Polymorphic DNA), SCoT (Start Codon Targeted Polymorphism) és IT (Intron targeting) markerek segítségével. További vizsgálataink során két amplifikált kloroplasztisz régiót vizsgáltunk PCR-RFLP (Polymerase Chain Reaction-Restriction Fragment Length Polymorphism) módszerrel. Univerzális primerek segítségével vizsgáltuk továbbá egyedi mitokondriális átrendeződések jelenlétét a csoportban. Ezen elő kísérletek során két szubkládot különítettünk el a subgenuson belük. Az egyik szubkládot az *Avicularia* és *Laciniata* sorozat, míg a másik szubkládot a *Similia* sorozat fajai alkották. Javaslataink alapján az *Archaesolanum* kládba tartozó sorozatok taxonómiáját felül kell vizsgálni. Eredményeinkre alapozva két lehetséges hipotézist fogalmaztunk meg a *S. laciniatum* hibrid eredetét illetően.

További vizsgálataink során tanulmányoztuk az *Archaesolanum* csoport történeti biogeográfiáját és megállapítottuk diverzifikációjának becsült korát. Ehhez vizsgálatainkba a csoport összes ismert faját valamint további külcsoportbeli tagokat is bevontunk és megállapítottuk három kloroplasztisz régió (*trn*T-*trn*L spacer, *trn*L intron és a *trn*L-*trn*F spacer) szekvenciáját, melyekre alapozva molekuláris órát kalibráltunk. A filogenetikai eredményeinket elterjedési adatokkal együtt elemezve vizsgáltuk a csoport jelenlegi elterjedését valószínűsítő biogeográfiai folyamatokat. Eredményeink alátámasztották a kenguru almák monofiletikus eredetét és a két szubklád biogeográfiai diszjunkcióját. Divergencia becsléseink alapján a kenguru almák közös őse a Miocén korból (~ 9Mya) való. Ezen megközelítésre alapozva a közös ős egy nagy távolságú elterjedési esemény következtében került Ausztráliába. A kenguru almákat a kontinens belső területeinek kiszáradási folyamata két különálló kisebb csoportra osztotta, melyeken belül a Pliocén kori esőerdők rövid felvirágzása további fajképzősi folyamatokat indított be.

ZUSAMMENFASSUNG

Molekular Genetische Untersuchungen an komplexe evolutionäre Prozesse in *Archaeasolanum* (*Solanum*, Solanaceae)

Die Kängurusäpfel Arten (Untergattung *Archaesolanum*) sind eine wenig bekannte spezielle Gruppe der Gattung *Solanum*, mit acht Arten, die ausschließlich in der Region des Pazifischen Ozeans vorkommen. Das gemeinsame Merkmal dieser Arten ist eine Basischromosomenzahl von x = 23. Die phylogenetischen Beziehungen dieser isolierten *Solanum* Arten sind nicht bekannt. Diese Studie wird der fehlenden Kenntnisse erweitern dienen. Sieben Arten der Gruppe wurden in unserem Voraufsatz mit RAPD (Random Amplified Polymorphic DNA), SCoT (Start Codon Targeted Polymorphism) und IT (Intron Targeting) Markern vervolgt. Das Vorhandensein der einmaligen mitochondrialen Neuordnung wurde mit universalen Primern untersucht. In dieser Arbeit wurden zwei Gruppen innerhalb der Untergattung getrennt. Eine von diesen wird durch die Arten der *Avicularia* und *Laciniata* Serie gebildet, die Andere durch die Arten der *Similia* Serie. Es wird angeraten die Taxonomie der zur *Archaesolanum* Untergattung gehörenden Serien zu überprüfen. Auf Grund unserer Ergebnisse wurden zwei mögliche Hypothesen über den Ursprung der *S. laciniatun* Hibryde abgefasst.

Die geschichtliche Biogeographie der *Archaesolanum* Untergattung wurde während unserer weiteren Untersuchungen studiert sowie das Zeitalter der Diversifikation festgestellt. In diese Untersuchung wurden alle bekannten Arten der Untergattung sowie weitere Pflanzarten außerhalb der Gruppe einbezogen und die Sequenzen drei informativer Chloroplastenegionen (*trn*T-*trn*L spacer, *trn*L intron und *trn*L-*trn*F spacer) bestimmt. Eine Molekularuhr, die auf dieser Untersuchung basiert, wurde kalibriert. Die phylogenetischen Kenntnisse wurden zusammen mit den geographischen Daten analysiert um die biogeographischen Vorgänge zu klären, die zu der Verbreitung der Gruppe führten. Unsere Ergebnisse stützten den monophyletischen Ursprung der Arten und die biogeographische Disjunktion der zwei Subkladen. Nach unserer Schätzung stammt der jüngste gemeinsame Vorfahr der Kängurusäpfel aus dem Miozän (~ 9 Millionen Jahre). Aufgrund dieser Einschätzung wird davon ausgegangen dass der gemeinsame Vorfahr der Kängurusäpfel durch eine weitläufige biogeographische Ausbreitung nach Australien kam. Der Trocknungsprozess des inneren australischen Gebietes teilte die Kängurusäpfel in zwei kleinere Gruppen auf. Auf diesen kleineren Flächen führte das Erscheinen der pliozänischen Regenwälder zu die weiteren Artbildungprozessen.

LIST OF ABBREVIATIONS

AAD – Arbitrarily Amplified DNA
AFLP – Amplified Fragment Length Polymorphism
BIC – Bayesian Information Criterion
bp – base pair
cpDNA – Chloroplast DNA
DEM – digital elevation map
DIVA – dispersal-vicariance analysis
EPT – Equally parsimonious tree
ESS – Effective sample size
EST – Expressed sequence tag
GISH – Genomic *in situ* hybridization
GSW – weight gain steps
GTR – General Time Reversible
ISSR – Inter-simple sequence repeat
IT – Intron targeting
kb – kilo base
KT – Cretaceous-Tertiary
LDD – Long-distance dispersal
LTT – Lineages-through time
LWS – weight loss steps
M – Million
MCMC – Markov chain Monte Carlo
MRCA – most recent common ancestor
mtDNA – Mitochondrial DNA
MY – Million years
Mya – Million years ago
NJ – Neighbor Joining
NW/SE – North-West/South-East
P – Pleistocene
PCR-RFLP – Polymerase chain reaction-Restriction Fragment Length Polymorphism
PI – probability indexes
PIC – Polymorphic Information Content
PL – Pliocene
PP – posterior probability
RAPD – Random Amplified Polymorphic DNA
rDNA-ITS – ribosomal DNA Internal transcribed spacer
SCoT – Start Codon Targeted Polymorphism
SD – standard deviation
sect. – section
SEM – Scanning Electron Microscope
ser. – series
subg. – subgenus
SW– South West
tMRCA – the most recent common ancestor
UCLD – uncorrelated lognormal distributed relaxed molecular clock
WAAA – weight-ancestral area analysis
WE – West-East
WGD – whole genome duplication

INTRODUCTION

Naturally occurring variation of the wild relatives of crop plants is an underexploited resource in plant breeding. All cultivated plants were once wild. Plant evolution under domestication has led to increased productivity, but, at the same time, domestication has dramatically narrowed the genetic basis of the species in cultivation. Considering that flowering plants evolved over 135 million years ago, crop plants, as we know them, have existed for the mere blink of an evolutionary eye. Agricultural productivity met the demand to increase crop yields which can maintain the current level of outputs for an increasing human population. According to some forecasts the human population will reach ca. 9 billion by the year 2030 (Brown 1994). It is unlikely that new farmland will become available in the near future. This makes genetic improvement of crops the best approach to keep pace with the anticipated growth of demand. To this end the promising new pool of genetic variation provided by wild species must be harnessed and successfully utilized. This genetic variation is the engine that propels breeding to meet future challenges.

Solanum is a genus of nearly 1,400 species and only a small portion of the species is cultivated. It is quite obvious to pursue to use the traits provided by the large and unexplored diversity. This genus is important from an agricultural perspective as well as from an evolutionary standpoint. The genus consists of valuable crop plants like eggplant, tomato and especially potato. Potato ranks as the world's third most important crop, with increasing production worldwide (www.potato2008.org). *Solanum* species represent nearly 1% of the world's angiosperm flora, which might be attributed to its great antiquity and an extraordinary rate of speciation. This huge diversity in one genus is quite exceptional in angiosperms, making *Solanum* interesting from an evolutionary standpoint as well as for its great economic importance.

Exotic germplasm resources, which include wild species and landraces, often carry many agriculturally desirable alleles and unknown evolutionary history. Wild *Solanum* species have been known for long time to be resistant to many pathogens. Therefore wild resistant species have been used in studies aimed at the identification of disease resistance genes (Vleeshouwers et. al. 2001).

Due to the large diversity of solanaceous plants, wild species holding important traits could also be important minor food crops or ornamentals in some parts of the world, while the same species can become invasive weeds of agricultural and rural habitats. In addition, these plants could also serve as alternative hosts for major diseases for crop plants.

Recent progress in understanding the phylogeny of the economically important plant family Solanaceae makes this an ideal time to develop models for linking the new data on plant genomics with the huge diversity of naturally occurring species of the family. Phylogenetics provides the framework to investigate these linkages. However, critical, good species-level descriptive resources for the Solanaceae are not available because many groups of the genus are still poorly known. In many cases the comprehensive molecular phylogenetic treatment to reconstruct evolutionary history of taxa has not yet been made. Even phylogenies of species having great economic importance are still debated e.g. the origin of eggplant has only recently been revealed (Weese and Bohs 2010). Much less is known about small groups of the genus, leaving questions open and providing exceptional opportunities for important evolutionary and applied research.

The species that belong to subg. *Archaesolanum* (kangaroo apples), are a distinctive group with no obvious close relatives. The subgenus includes eight species, which occur only in the South Pacific (New Guinea, Australia, Tasmania, New Zealand). Besides their restricted occurrence, they possess many unique characters like unusual chromosome number based on n = x = 23, instead of n = x = 12 typical in other members of the genus. The genetic relationships and how this interesting chromosome number has developed are also still unknown. In addition, the phylogeny of this group has not yet been studied utilizing molecular tools despite the fact that many molecular studies on phylogenetic relationships within the genus *Solanum* have included species representing this subgenus. Compared to other *Solanum* clades still very little is known about the evolutionary dynamics, biogeography, dispersal, radiation and genetic diversity of the group. Several studies have hinted the complexity of solanaceaus plants, because of polyploidy, hybridization etc. The reported complexity – which is unambiguously based on the biological processes given above – remarkably affects also practical research (e.g. plant breeding programs).

On the other hand, it makes the *Archaesolanum* interesting from an evolutionary standpoint to plant scientists. This small group of Australian species can serve as an ideal

group to study these processes to understand evolutionary complexity. Presumably interesting evolutionary processes have acted in this plant group and thus they would serve as ‘model organisms’ to study fundamental and important processes of plant biology. Despite of this no serious research program has yet been started to explore their biology and phylogeny.

OBJECTIVES OF THE STUDY

The aim of the present work is to clarify the taxonomy of kangaroo apples and to investigate speciation processes by using multi-locus markers and chloroplast DNA sequences.

The aims can be summarized as:

- Define taxonomic boundaries, and provide basis for a new systematic scheme based on molecular data
- Analyze phylogenetic relationships in the subgenus to reveal relationships of kangaroo apples
- Test putative hybridization patterns between *Solanum laciniatum, S. vescum, S. multivenosum,* and *S. aviculare*.
- Estimate the age of the most recent common ancestor of the group using a molecular clock
- Evaluate the biogeographical history of the group in Australia

CHAPTER 1

Background

1.1. Phylogeny of the genus *Solanum*

The great abundance of *Solanum* L. species represents nearly 1% of the world's angiosperm flora (Whalen and Caruso 1983). The extreme diversity *Solanum* may be attributed to its great antiquity, but in addition to an extraordinary rate of speciation (Whalen 1979a). The current infrageneric subdivisions within *Solanum* have been challenged in several studies. *Solanum* is the largest genus in the Solanaceae with approximately 1,400 species, and is one of the largest genera of flowering plants (Olmstead and Palmer 1997). *Solanum* is a taxonomical paradox, exhibiting both uniformity and extreme diversity in its morphology (Roe 1972). This hyperdiversity in one genus is quite unusual in angiosperms, making *Solanum* interesting from an evolutionary standpoint as well as for its usefulness to humans (Knapp et al. 2004). The genus is widely distributed throughout the world, with major species diversity in America, Australia and Africa (Bukenya and Carasco 1995).

Early workers such as Dillenius (1732) and Linnaeus (1753) were the first to study the taxonomy of the genus. Linnaeus (1753) divided *Solanum* into two groups, *Spinosa* and *Inermia*, based on the presence or absence of spines (Bohs 2005), while Dunal (1813, 1816) described two categories, *Aculeata* and *Inermia* in his monographs. Bitter (1912, 1913, 1917, 1919, 1921, 1922, 1923) who has been criticized for splitting the genus excessively, described more than 60 new *Solanum* species from the Americas (Edmonds 1977).

The works of Seithe (1962), Danert (1970) and Gilli (1970) provided elements for D'Arcy's (1972, 1991) scheme which is widely used today. According to D'Arcy (1972, 1991) *Solanum* is divided into seven subgenera [*Archaesolanum* Marzell, *Bassovia* (Aubl.) Bitter, *Leptostemonum* (Dunal) Bitter, *Lyciosolanum* Bitter, *Minon* Raf. (*Brevantherum* (Seithe) D'Arcy), *Potatoe* (G. Don) D'Arcy and *Solanum* Seithe] and 60 to 70 sections. Well defined and probably monophyletic subgenera and sections exist along with a plethora of poorly circumscribed groups. Significant number of *Solanum* species have no conclusive

subgeneric or sectional affiliation. Even for well-characterized infrageneric groups phylogenetic relationships with other groups are unknown (Bohs and Olmstead 1997).

Solanum is ridden with taxonomic confusion (Lester 1997). The difficulty of associating the species names of *Solanum* used by earlier taxonomists like Linnaeus is due to the fact that many of them are very difficult to typify (Hepper 1979). In addition to this the early descriptions are brief, often vague and frequently lacking in characters now considered to be diagnostic (Bukenya and Carasco 1995).

Several authors have provided schemes for infrageneric groups (e.g. Child and Lester 2001, Nee 1999, Hunzinker 2001). Analyses of morphological characters provided information for *Solanum* sect. *Androceras* (Nutt.) Marzell (Whalen 1979a,b), Solanum sect. *Lasiocarpa* (Dunal) D'Arcy (Whalen et al. 1981, Whalen and Caruso 1983, Bruneau et al. 1995), the *S. nitidum* Ruiz & Pav. group [*Solanum* sect. *Holophylla* (G. Don) Walp. pro parte Knapp 1989], the *S. sessile* Ruiz & Pav. group [Solanum sect. *Geminata* (G. Don) Walp. pro parte, Knapp 1991], *Solanum* subg. *Leptostemonum* (Dunal) Bitter (Whalen 1984), *Solanum* subg. *Potatoe* (G. Don) D'Arcy (Spooner et al. 1993), *Solanum* sect. *Brevantherum* Seithe (Roe 1972) and Solanum sect. *Solanum* Seithe (Edmonds 1972, 1977, 1979).

The advent of molecular data has revolutionized the field of plant systematics and has led to new insights into phylogenetic relationships at all taxonomic levels (Bohs 2005). Molecular techniques were used in the studies of Olmstead and Palmer (1997) and Olmstead et al. (1999) to investigate relationships of Solanaceae. Other studies using molecular techniques provide information at the subgeneric and sectional levels (e.g. Bohs and Olmstead 1997, 1999, 2001; Levin et al. 2005, 2006; Jacoby et al. 2003, Stedje and Bukenya-Ziraba 2003, Furini and Wunder 2004).

The study of Bohs (2005) based on the use of molecular data from chloroplast *ndh*F sequences and a broad spectrum of samples from different subgroups identified about 13 major clades within *Solanum.* Sampling included all the seven subgenera listed in D'Arcy's (1972) conspectus, and 40 of the 62 sections in D'Arcy (1991). The study recognized several new clades such as the *Dulcamaroid* and *Morelloid* clades which include species from different taxonomical groups. Even though the clades identified by Bohs (2005) are well supported, they need to be corroborated by data from other genes and both morphological and biochemical characters should be examined together. Moreover, new formal taxonomic

designations for infrageneric categories in *Solanum* are still poorly defined without more extensive data and sampling (Bohs 2005). Later Weese and Bohs (2007) provided the major scheme for classification of *Solanum* based on sequence data of three separate genes. This scheme is widely applied today (Fig.1).

Most papers provide information about phylogenetic relationships using single-locus methods (Bohs and Olmstead 1997, Bohs 2005). Other studies use DNA sequence data from nuclear regions such as ITS and granule-bound starch synthase gene (GBSSI or waxy) or chloroplast regions (*trn*T-*trn*F and *trn*S-*trn*G) or combinations of these data (Levin et al. 2005, 2006). Several studies have used multi-locus techniques to investigate phylogenetic relationships in *Solanum*, including the AFLP analysis of *S. melongena* L. and its wild relatives (Mace et al. 1999), and *S. retroflexum* Dun. and related species (Jacoby et al. 2003). Furini and Wunder (2004) used AFLP´s to analyze infrageneric relationships. RAPD data was used in several studies [e.g. Stedje and Bukenya-Ziraba (2003), Berg et al. (2002), Spooner et al. (1996, 1997), Miller and Spooner (1999), Karihaloo et al. (1995)] to clarify phylogenetic relationships. In most cases multi-locus methods were applied to explore relationships at sectional levels. This approach has been rarely used for the analysis of infrageneric groups.

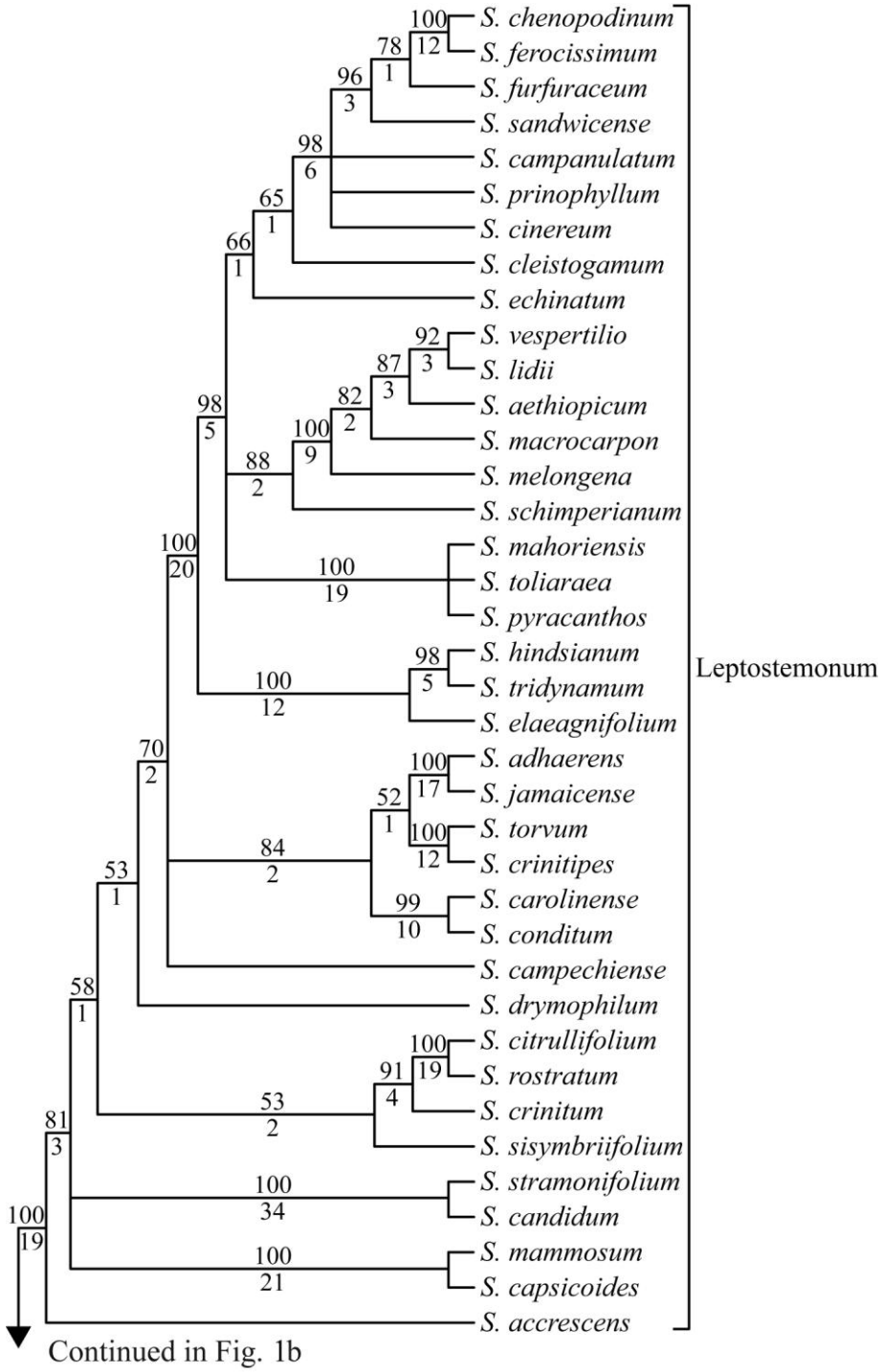

Fig.1a. Major clades in the genus *Solanum* by Weese and Bohs (2007). Figures are kindly provided by L. Bohs (University of Utah).

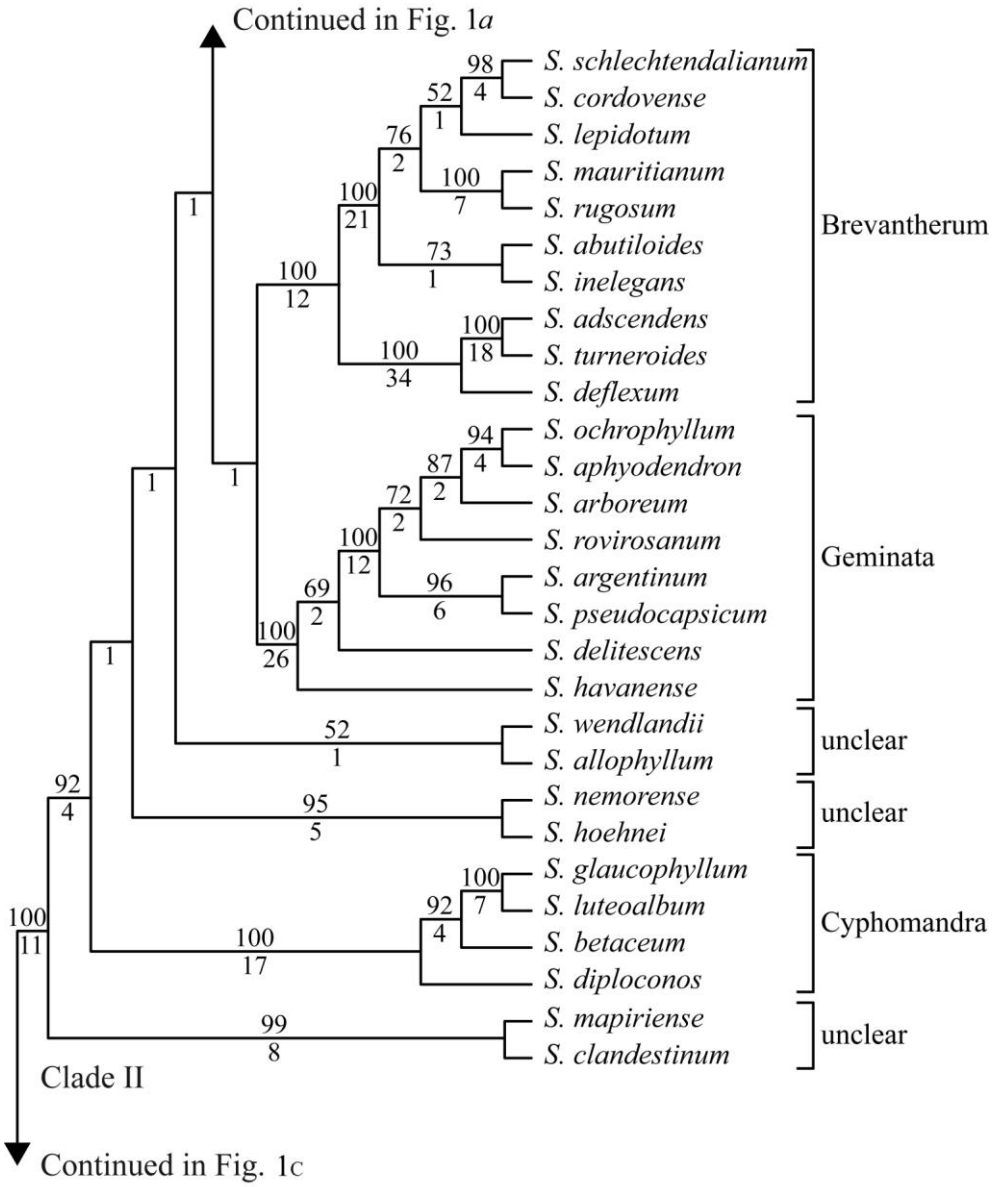

Fig.1b. Continued figure showing major clades in the genus *Solanum*. Strict consensus of 21,017 most parsimonious trees obtained from the combined analysis of the *trn*T-F, *ndh*F, and *waxy* data. Numbers above branches are bootstrap values over 50% based on 1,000 random addition replicates; numbers bellow branches are decay values.

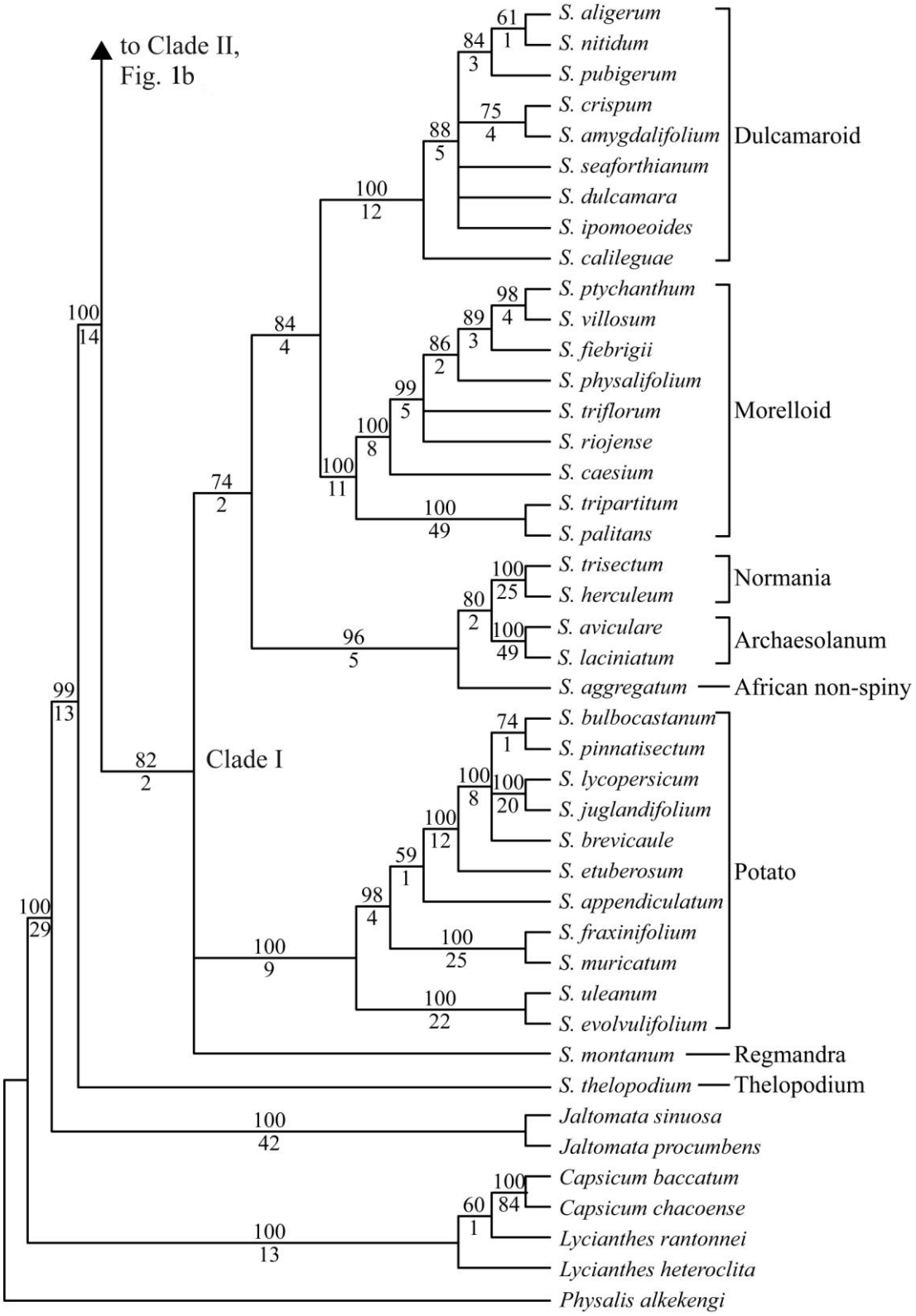

Fig.1c. Continued figure showing major clades in the genus *Solanum*.

1.2. Kangaroo apples (subg. *Archaesolanum*)

Solanum L. subgenus *Archaesolanum* Bitter ex Marzell[1], often called kangaroo apples, is composed of eight species occurring only in the SW Pacific region (Australia, Tasmania, New Zealand, Papua New Guinea). The group is characterized by its unique chromosome number (x = 23), possibly resulting from an aneuploid loss from a polyploid (x = 24) ancestor (Randell and Symon 1976). This unique feature makes *Archaesolanum* particularly interesting from an evolutionary standpoint. However, genetic relationships and how this interesting chromosome number has developed are unknown. In addition, phylogeny of this group has not yet been studied utilizing molecular tools despite the fact that many molecular studies on phylogenetic relationships within the genus *Solanum* have included one or two representative species from the subgenus (Bohs and Olmstead 2001; Bohs 2005; Weese and Bohs 2007; Poczai et al. 2008). Compared to other clades, evolutionary dynamics, biogeography, dispersal, radiation and genetic diversity of the group are poorly known.

The earliest record of the kangaroo apples is by Forster (1786a) describing *Solanum aviculare* based on the specimen collected during Captain James Cook´s second voyage to New Zealand. After Forster's first report, during the 18th and 19th century, at least three species were collected from coastal Australia or New Zealand and brought to Europe where they were cultivated in botanical gardens serving as material for confusing descriptions (for example Lamarck 1792; L'Héritier 1805) whilst others continuously published new names (e.g. Mueller 1855; Hooker 1857) with varying success. These early studies have resulted in extensive synonymization and unsettled taxonomic concepts. As summarized by Spooner (2009), species concepts in *Solanum* have been so controversial that even experienced taxonomists have provided different identifications for identical collections. However, many problematic cases have been settled by Baylis (1963) and Symon (1994).

Species concepts have been controversial within *Archaesolanum* and also its status as a separate taxonomic unit has been ambiguous. Dunal (1852) did not recognize the distinctiveness of the group and placed it under subsection *Dulcamara.* Later Bitter (1927), in his last contribution to Gustav Hegi's book, published the name subgenus *Archaesolanum*, typified by *S. aviculare*. However, Bitter died shortly before his work was published and it

[1] Basionym of *Solanum* section *Archaesolanum* (Bitter ex Marzell) Danert

was edited by Marzell. The name given to the group suggests an ancient origin possibly attributed to the free oriented stamens and the presence of abundant stone cell mass in the fruit flesh. Since Bitter, the group has been recognized at sectional level (Danert 1970). It was elevated to subgeneric level by D'Arcy (1972, 1991), and again treated as a section (Nee 1999) then once more as a subgenus (Hunziker 2001). In studies utilizing sequence level characters (Bohs 2005, Weese and Bohs 2007) the group has been distinguished as the *Archaesolanum* clade. All treatments agree that it is distincti from other members of the genus, and that the species form a unique and coherent group within *Solanum*. So far very little work has been done to characterize further subdivisions within the group. Gerasimenko (1970) described three series, and this division, as modified, is also applied today (Symon 1994).

1.2.1. Taxonomy and typification

The species of subg. *Archaesolanum* are short-lived soft-wooded shrubs, 1–3 m tall, becoming straggly with age, glabrescent, with large (up to 30 cm) deeply lobed leaves in the juvenile phase, becoming smaller (up to 10 cm) and entire in the adult stage, with violet–purple flowers in cymes growing at the axils of branches (Symon 1985). The fruits are greenish, yellowish, or scarlet; the succulent berries produce numerous seeds (approx. 100–600). White or yellowish stone cell aggregates are present in the dried contents of the fruits, mixed with seeds. The fruits are eaten by birds, which are probably responsible for their dispersal throughout Australia, New Zealand, Tasmania, and New Guinea (Symon 1981).

As mentioned above plants of this group were first collected by Forster in Australia during the second voyage of Captain James Cook. Forster (1786a) was the first to publish the name *Solanum aviculare* in the ''*Dissertatio inauguralis botanico-medica de plantis esculentis insularum oceani australis*'', a record of his collections from New Zealand and two other records from Australia. However, the correct citation for the name *S. aviculare* is often confused since there are three publications which appeared slightly after each other in the same year. The *Plantis Esculentis* (Forster 1786b) and the *Flora insularum* (1786c) was published after the *Dissertatio inauguralis* (1786a) which predates all other publication and fulfill the criteria of the Botanical Code.

However, the distinction of the group has been debated and also its taxonomic rank - whether it is should be treated as a section or a subgenus the group was divided to further series by Gerasimenko (1970). This system was further modified by Baylis (1963) and Symon (1994), than further synonyms were clarified by the PBI: *Solanum* project (Knapp et al. 2004).

Table 1. Taxonomic schemes of kangaroo apples

Taxonomic scheme by Gerasimenko (1970)	**Modified scheme by Baylis (1963) and Symon (1994)**
Series *Avicularia* Geras.	**Series *Avicularia* Geras.**
S. aviculare Forst. 1786	*S. aviculare* Forst. 1786
S. cheesemanii Geras. 1971	*S. multivenosum* Symon 1985
S. baylisii Geras. 1971	
S. brisbanense (Geras.) Geras. 1971	
Series *Laciniata* Geras.	**Series *Laciniata* Geras.**
S. laciniatum Ait. 1789	*S. laciniatum* Ait. 1789
S. linearifolium Geras. 1965	*S. linearifolium* Geras. ex Symon 1981
S. vescum F. Muell 1855	*S. vescum* F. Muell 1855
Series *Similia* Geras.	**Series *Similia* Geras.**
S. capsiciforme (Domin) Baylis 1963	*S. capsiciforme* (Domin) Baylis 1963
S. simile F. Muell. 1855	*S. simile* F. Muell. 1855
S. symonii Eichler 1963	*S. symonii* Eichler 1963

The further subdivision made by Gerasimenko (1970) were sparse and had several weaknesses, but it pointed out that further and relevant genetic variability may exist within this group to make further subdivisions. Later studies reduced the number of species and currently only eight species are included in the group (Fig. 2A-F)

Fig. 2. 2A-Flowers of *Solanum laciniatum*; 2B-*Solanum linearifolium*; 2C-D- *Solanum simile*; 2E- *Solanum symonii*; 2F- *Solanum vescum*

1.2.2. Chromosome numbers and polyploidy

The genus *Solanum* is an interesting target group to study complex polyploid genome evolution. It seems obvious that a genus which represents nearly 1% of the angiosperms (Whalen and Caruso 1983) has had at least one polyploid ancestor at some point in its evolution. The age of Solanaceae is estimated to be ca. 40 million years (Myr; Wikström et al. 2001) and they possibly diverged from an ancestral diploid with x = 12 (Wu et al. 2006). Molecular clock estimates suggests that an ancient duplication in potato is most likely shared with tomato, and represents a whole genome duplication (WGD) event early in the evolution of the Solanaceae (Schlueter et al. 2004). The genetic maps (Bonierbale et al. 1988; Tanksley et al. 1992, Doganlar et al. 2002) also support this view. The members of the related family Rubiaceae (coffee family) are also diploid with x = 11 or x = 12, implying that this WGD event possibly occurred before their divergence (Wang et al. 2008). Moreover, another duplication has occurred in the evolution of potato about ~12-13 My ago (Gebhardt et al. 2003; Schlueter et al. 2004). While another estimate shows that the first ancient WGD coincides with the Cretaceous-Tertiary (KT) boundary extinction events circa 65 Mya (Fawcett et al. 2009). The authors also proposed that this polyploidization may have contributed to the survival and propagation of this plant lineage during the KT extinction event, due to advantages such as altered gene expression manifested in hybrid vigor and an increased set of genes and alleles available for selection leading to a better adaptation in the drastically changed environment (Fawcett et al. 2009).

Polyploidy is not restricted only to the potato-tomato lineage. It seems to have happened in almost all of the 13 clades (Bohs 2005; Weese and Bohs 2007) of the *Solanum*, i.e. Morelloid (sect. *Solanum*; Edmonds 1977); Potato (sect. *Petota*; Hawkes 1990), and Leptostemonum clades (sect. *Melongena*; Moscone 1992). Based on the x = 12 chromosome number polyploid series are frequent, while some species are anomalous aneuploids, i.e. x = 11 (*S. mamosum* L.), 15 (*S. bullatum* Vell.) and 23 (subg. *Archaesolanum* Bitter ex Marzell) (Acosta et al. 2005). There are many other examples which could be mentioned, but kangaroo apples (subg. *Archaesolanum*) are a distinct case among these. Interestingly, species have generated a further secondary polyploid series. Consequently, these diploids (x = 46, e.g. *S. aviculare*) could be better regarded as "tetraploids" – in terms of the x = 12, typical basic chromosome number of the genus – while tetraploids of the group (x = 92, e.g. *S. multivenosum*) are therefore better understood as "octoploids" (see Table 2). In other words, it

is presumed that the early ancestor of this group has gone through a simple ploidy increase accompanied by a chromosome loss and then the duplication has been repeated. Despite the name, suggesting an archetypal *Solanum*, the chromosome number indicates a derived condition which has itself become polyploid (Symon 1979), probably reached by aneuploid loss from n = x = 24 (Randell and Symon 1976). It is clear that all species based on secondary gametic numbers are polyploid; in the case of secondary polyploidy (Hair 1966) but how this interesting structure developed presently can only be speculated.

Table 2. Chromosome numbers of kangaroo apples (subg. *Archaesolanum*).

Diploids (n = x = 23)	**Tetraploids (n = x = 46)**
S. aviculare	*S. laciniatum*
S. capsiciforme	*S. multivenosum*
S. linearifolium	*S. symonii*
S. simile	
S. vescum	

1.2.3. Distribution

The *Archaesolanum* clade represents an isolated group and its closest relatives have not yet been identified (Bohs 2005). Putative ancestors have certainly not been recognized in Australia, nor are any extra-Australian relatives apparent (Symon 1970, 1979, 1984). Although red-fruited species, such as *Solanum dunalianum* Gaudich., *S. viride* Spreng., and *S. incanoalabastrum* Symon, occur in New Guinea, these all have stellate hairs and no stone cells; they are not related to the *Archaesolanum* group. According to the most widely accepted hypothesis (Olmstead and Palmer 1997) on the biogeography of *Solanum* the *Archaesolanum* clade presents an ambiguous case, either representing an early dispersal event in the genus, or a plausible case of vicariance dating to a time preceding the separation of South America and Australia. Hawkes and Smith (1965) suggested a Gondwanan origin for the family.

However, in these studies species of *Archaesolanum* are specifically not mentioned. Later studies by Symon (1986, 1991, 1994) supported the southern origin of kangaroo apples considering that the establishment predated the Gondwanan break up. However, there is no information on how the species reached their current distribution. We do not know what the driving forces of speciation and diversification were during the evolution of the group.

However, the current range of each species is well documented thanks to the monograph of Symon (1994) and the information accessible through the Global Biodiversity Information Facility (GBIF) portal (http://data.gbif.org/species/). Based on these data and other records the current range of kangaroo apple species are presented in Fig. 3.

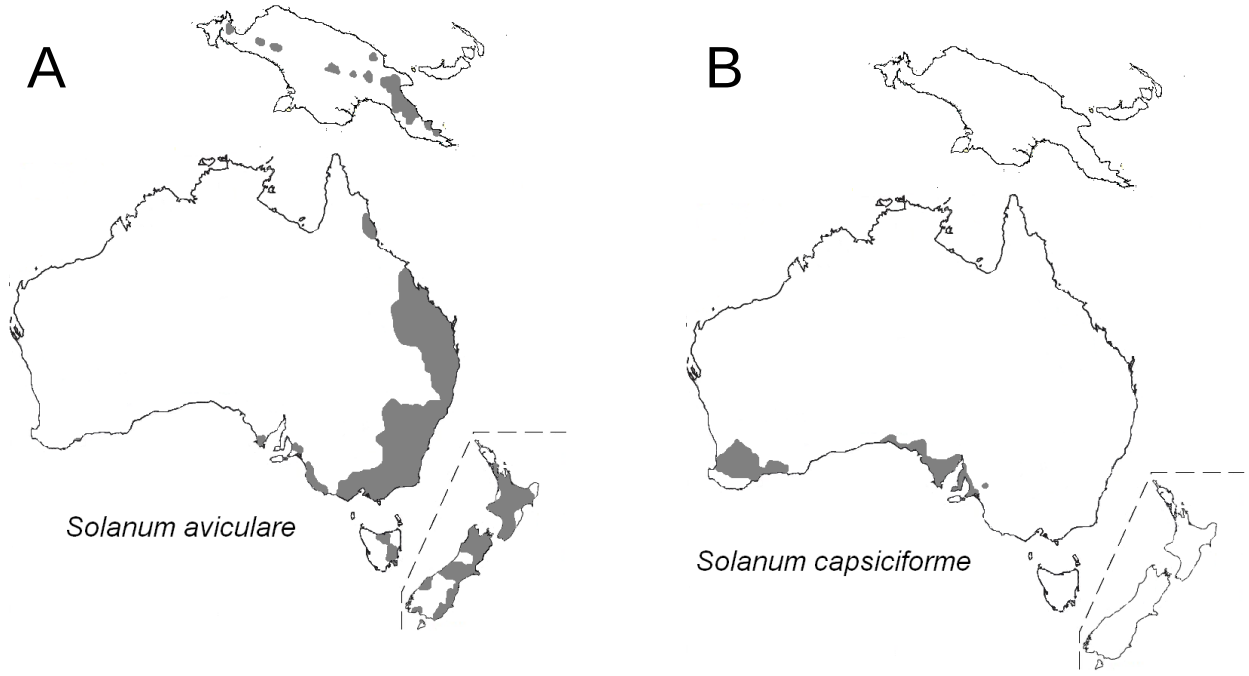

Fig.3. Distribution of kangaroo apples (*Solanum* subg. *Archaesolanum*) in the South Pacific region (Poczai et al. 2011a). Maps are based on the surveys of Symon (1994) and on the records registered by GBIF portal. **A**- *Solanum aviculare*; **B**- *Solanum capsiciforme* (continued on next page).

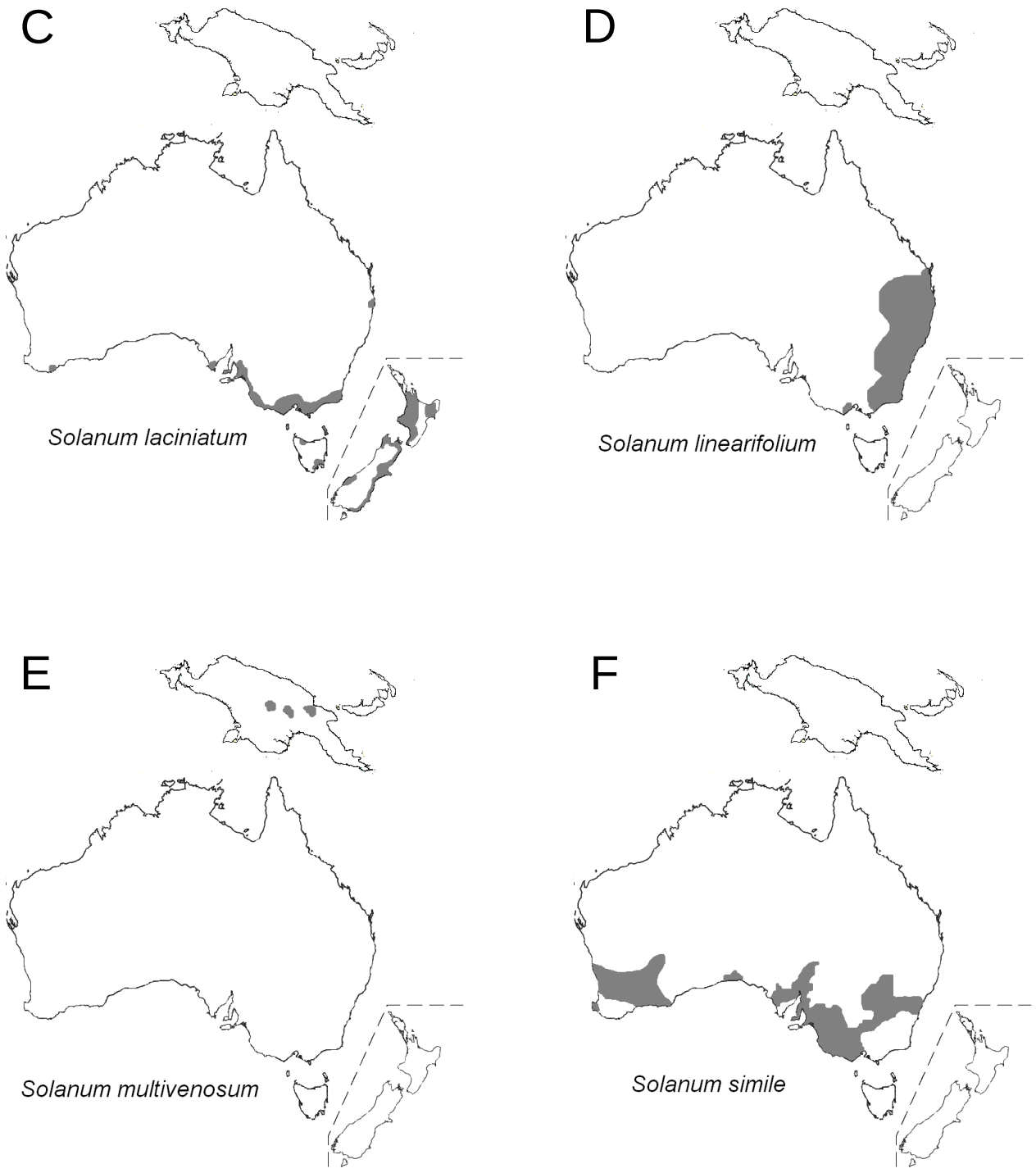

Fig.3. Distribution of kangaroo apples (continued). **C**- *Solanum laciniatum*; **D**- *Solanum linearifolium*; **E**-*Solanum multivenosum*; **F**- *Solanum simile*.

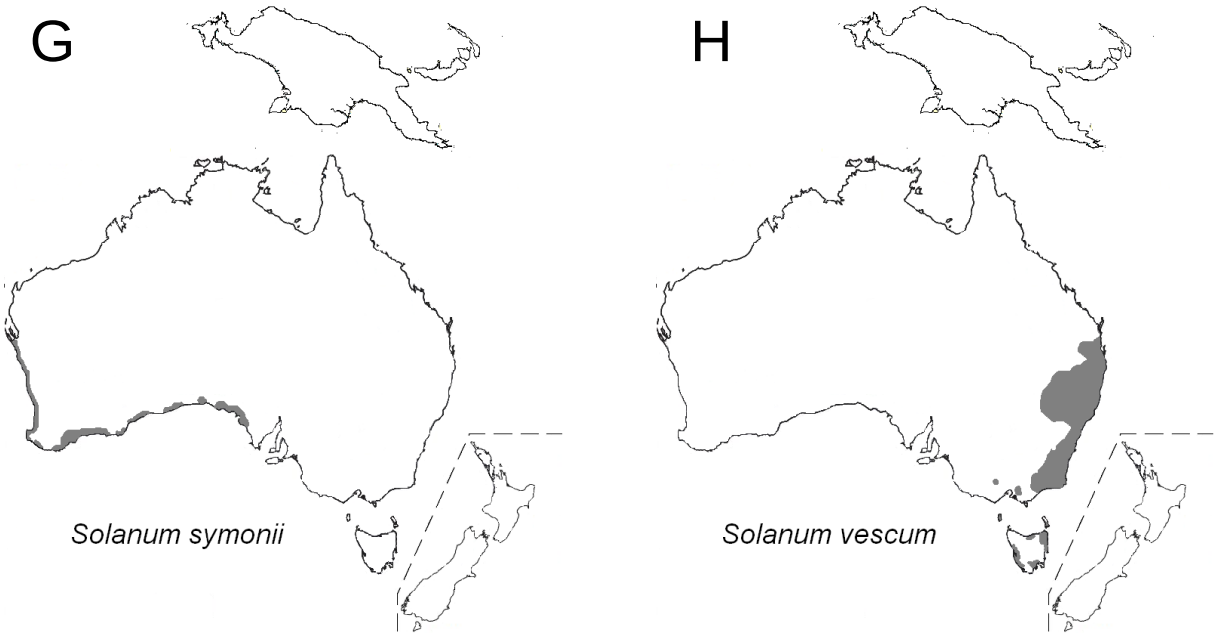

Fig.3. Distribution of kangaroo apples (continued). **G**- *Solanum symonii*; **H**- *Solanum vescum*.

1.2.4. Utilization

Alkaloids are one of the most important groups of secondary metabolites due to the great number of isolated products and their pharmacological activity. Solasodine, a steroidal alkaloid mainly found in solanaceous plants, has been considered as a potential alternative to diosgenin for commercial steroid drug synthesis (Galanes et al., 1984). It is a precursor for commercial production of steroidal hormones. Solasodine has been reported to accumulate in relatively high concentrations in a number of *Solanum* species, but it is found in highest quantity in kangaroo apples. Species of the group were considered as drug crops to exploit their alkaloid content for the pharmaceutical industry.

Species of kangaroo apples are also used as a remedy by locals in New Zealand and Australia (Mueller 1855). These early observations probably served as a starting point to utilize kangaroo apples in the drug industry. Several reports appeared during the ‘70s and ‘80s on cultivation experiments in Australia, New Zealand and also in Hungary (Collins et al. 1976; Mann 1979; Szalay-Marzsó and Nagy 1969).

However, the biggest effort to cultivate kangaroo apples was done by researchers in the former USSR (Korneva and Matveenko 1979). The history of these efforts is well

summarized by Symon (1994). Overall none of these activities have persisted, due to perhaps to variability in supplies, alkaloid content, or, to lack of demand after the development of synthetic production (Knapp 2006). Other reasons like variable germination of seeds (Pólya and Pólya Borsos 1966), great variation found in alkaloid content strongly depending on environmental conditions (Máthé et al. 1964; Máthé 1982) as well as unsuccessful cultivation (Verzár Petri 1964) all contributed for further experiments to cease with these plants. It was also reported that several pathogens – mostly viruses – affect cultivated kangaroo apples making production cumbersome (Gáborjányi 1969; Gáborjányi and Nagy 1970). In Hungary cultivation experiments were mostly unsuccessful due to unfavorable weather conditions (Máthé et al. 1986) which resulted in the loss of flowers (Bernáth 1970; Bernáth 1971) and low alkaloid content (Verzár Petri et al. 1967). However, other research groups continued to develop new varieties in Thailand (Vocel and Horn 1992) while others started *in vitro* experiments. In the past two decades, several studies have focused on the production of steroidal compounds from *in vitro* cultures, especially callus and cell suspension, of a number of *Solanum* species (Khanna et al. 1977); root cultures have been largely overlooked due to a low accumulation of solasodine in the roots of intact plants. This, however, changed dramatically when transformed 'hairy' root culture was introduced as a new route of secondary metabolite production (Flores et al., 1987). These experiments continued with variable success (Subroto and Doran 1994; Yu et al. 1996; Jasik et al. 1997; Kittipongpatana et al. 1998; Vaněk et al. 1999).

Besides, their utility in pharmaceutical sciences no other reports have been published on the utilization of kangaroo apples as genetic resources in plant breeding. However, Takács (2001) briefly studied pathogens affecting *S. laciniatum* and *S. aviculare* and reported that no resistance response was found.

The anomalous chromosome numbers found in the group would also prevent any other possibilities to utilize gene resources originating from kangaroo apples in conventional breeding programs.

Some kangaroo apple species were introduced very early to Europe where they were cultivated as ornamentals (Máthé és Földesi 1965). The most successful species in this respect is undoubtedly *S. laciniatum* which is also available nowadays in commercial trade in Europe.

1.3. Molecular markers in plant phylogenetics

Molecular data sets are undoubtedly the most important resources of phylogenetic reconstruction. In plants molecular markers are one of the most valuable resources for phylogenetic analyses. Their utility in determining genetic diversity and to reconstruct evolutionary processes is also well known. The detection and the analysis of these events enable us to understand the molecular basis of various biological phenomena in plants. Systematics (taxonomy) has been totally transformed during last decades because of i) adoption of cladistic methodology, ii) development of numerical methods and related powerful algorithms, iii) steadily increasing computing resources, and finally iv) recent development in molecular methods that have led to exponential growth of data available. Databases such as GenBank have become an essential resource also for systematics. In recent years, a new class of advanced techniques has emerged, primarily derived from combination of earlier basic techniques (Agarwal et al. 2008). There is also a wide range of different marker systems that can be applied in different ways in phylogenetics and genetic diversity analyses (Calonje et al. 2009).

1.3.1. Utility of chloroplast markers in plant phylogenetics

Chloroplast DNA (cpDNA) has been used extensively to infer plant phylogenies at different taxonomic levels (Gielly and Taberlet 1994). The advantages and disadvantages of using chloroplast characters – both structural and DNA sequence data – in phylogenetic reconstructions are well known (see Soltis and Soltis 1999). The first advantage of cpDNA might be its relative small size, since the chloroplast genome varies little in size, structure, and gene content among angiosperms (Olmstead and Palmer 1994). The typical chloroplast genome in angiosperms ranges in size from 135 to 160 kb and is characterized by a large, ca. 25-kb inverted repeat, which divides the reminder of the genome into a large and one small single copy region (Palmer 1985; Sugiura 1989,1992). However, smaller genomes have been documented in which one copy of the inverted repeat is missing (DePamphilis and Palmer 1990). Substantially larger chloroplast genomes (217 kb) have also been documented, but in most cases the size increase is due to inverted repeats and not to an increase in genome

complexity (Palmer 1987). The second advantage is that in the chloroplast genome most genes are essentially single-copy (Palmer 1985, 1987), in contrast many nuclear genes belong to multi-copy gene families e.g. rDNA-ITS (Poczai and Hyvönen 2010).

The conservative evolution of the chloroplast genome can be an advantage or even a disadvantage for phylogenetic analysis, but in these reconstructions it should also be considered that different regions of the cpDNA evolve at different rate (Palmer 1985). This feature of chloroplasts can be very useful for alignment of sequences at higher level, but this might be a disadvantage at lower level phylogenetic analyses because there is not enough variation. The second disadvantage might be that chloroplast phylogenies only represent maternal lineages since in land plants the chloroplast genome is mostly maternally inherited (Gillham 1978) but there are well know exceptions of paternal (Wagner et al. 1987; Szmidt et al. 1987) or even biparental inheritance (Stubbe 1984; Metzlaff et al. 1981). Another, third disadvantage can be the potential occurrence of chloroplast transfer: the movement of a chloroplast genome from one species to another by introgression (Soltis and Soltis 1999). Although chloroplast capture, if undetected, will bias estimates of phylogeny, it can, when recognized, be very informative about evolutionary processes (Soltis and Soltis 1999).

1.3.2. The utility of the chloroplast *trn*T-*trn*F region

The *trn*T-*trn*F region is located in the large single-copy regions of the chloroplast genome, approximately eight kb downstream of *rbc*L (Jigden et al. 2010). This region consists of three highly conserved transfer RNA genes namely tRNA genes for threonine (UGU), leucine (UAA) and phenylalanine (GAA) (Borsch et al. 2003,2007). These exons are separated by two intergenic spacers (*trn*T-L and *trn*L-F) while the *trn*L gene is split by a group I intron (Borsch et al. 2003).

As the advent of molecular methods has revolutionized the field of plant systematics (Panwar et al. 2010; Liu et al. 2010; Ciarmiello et al. 2010; Wang et al. 2010; Pamidimarri et al. 2010; Grativol et al. 2010) this region became widely used due to its high variability. We have chosen this region for our sequence level based investigations, because the evolution of the *trn*T-F has been thoroughly analyzed and is well understood (Borsch et al. 2003), and it can be used to calibrate a molecular clock. More recently, it was also shown that this region

comprises more phylogenetic structure per informative character than *mat*K (Müller et al. 2006), another widely used chloroplast region in phylogenetics. Based on its high variability it was used in studies to address relationships at the species and genus levels (e.g. Taberlet et al. 1991; Sang et al. 1997; Bakker et al. 2000). Moreover, this region has been quite informative in phylogenetic studies of the families like Asteraceae (Bayer and Starr 1998), Arecaceae (Asmussen and Chase 2001) and orders like Laurales (Renner 1999) and Magnoliales (Sauquet et al. 2003) or even across angiosperms (Borsch et al. 2003). The region has been frequently used in systematic studies of Solanaceae (Olmstead and Sweere 1994; Fukuda et al. 2001; Garcia and Olmstead 2003; Santiago-Valentin and Olmstead 2003; Clarkson et al. 2004; Montero-Castro et al. 2006; Lorenz-Lemke et al. 2010) and to infer relationships in the genus *Solanum* (Bohs 2004; Levin et al. 2005; Miller and Diggle 2007; Weese and Bohs 2007; Stern et al. 2010; Weese and Bohs 2010; Poczai and Hyvönen 2011).

1.3.3. Arbitrarily amplified DNA markers (AAD) for phylogenetic inference

Collectively, techniques, such as AFLP, ISSR and RAPD, have been termed as arbitrarily amplified dominant (AAD) markers (Karp et al. 1996; Wolfe and Liston. 1998). AAD markers have also been a source for phylogenetic inference and systematic studies at various levels, in both distance- and parsimony-based analyses (Winter and Kahl. 1995; Gupta et al. 1999; John et al. 2005; Simmons et al. 2007). The major advantage of the above mentioned dominant markers is based on the fact, that there is no need to have any preliminary sequence information from the analyzed organism. Moreover, dominant markers are generated randomly all over the whole genome sampling multiple loci at one time, providing large amount of data for analyses.

These methods generate a relatively large number of markers per sample in a technically easy and cost effective way. However, AAD markers have been criticized by their negative features. Gorji et al. (in press) summarizes these as: i) homoplasy, the comigration of same size fragments originating from independent loci among different analyzed samples; ii) non-homology, comigrating bands are paralogous (map to different positions in different individuals) instead of being orthologous (map to the same genomic location); iii) nested priming – amplicons result from overlapping fragments; iv) heteroduplex formation – products are also generated from alternate allelic sequences and/or from similar duplicated

loci; v) collision – the occurrence of two or more equally sized, but different fragments within a single lane; vi) non-independence – a band is counted more than once, due to co-dominant nature or nested priming; vii) artefactual segregation distortions, caused by loci mis-scoring, undetected codominance or poor gel resolution (Gort et al. 2009; Bussel et al. 2005; Simmons et al. 2007).

We chose arbitrary amplified DNA (AAD) markers for our pilot studies to produce fragments that are generated by random amplified polymorphic DNA (RAPD) or start codon targeted (SCoT) over the whole genome. Recent studies (Jacobs et al. 2008; Kingston et al. 2009; Rubio-Moraga et al. 2009; Croll and Sanders 2009) have shown that AAD markers can solve phylogenetic relationships of closely related, recently radiated taxa at low taxonomic levels (Davierwala et al. 2001; Awasthi et al. 2004; Sica et al. 2005). However, one of the arguments against the use of AADs is that they are homoplastic – co-migration of non-identical bands – causing noise instead of phylogenetic signal in the datasets as discussed above (Jones et al. 1997; Meudt and Clarke 2007).

The species of the subg. *Archaesolanum* are assumed to be very closely related and homoplasy becomes a greater problem where distantly related species are involved; it is less likely to cause problems for studies of very closely related species (Jacobs et al. 2008; Koopman 2005). This assumption certainly applies to other *Solanum* taxa where the utility of multi-locus methods in phylogenetic reconstruction have repeatedly been used at species level (Kardolus et al. 1998; Berg et al. 2002; McGregor et al. 2002; Lara-Cabrera and Spooner 2004; Spooner et al. 2005; Poczai et al. 2008; Poczai et al. 2010; Poczai and Hyvönen 2011).

1.3.4. Start Codon Targeted (SCoT) Polymorphism

Based on the rapid increase of genomic research many new advanced techniques have emerged. In the recent years there has been a trend away from random DNA markers towards gene-targeted markers (Andersen and Lubberstedt 2003). Molecular markers from the transcribed region of the genome can facilitate various applications in plant genotyping as they reveal polymorphisms that might be directly related to gene functions (functional markers; De Keyser et al. 2009). The novel marker system called Start Codon Targeted (SCoT) Polymorphism was described by Collard and Mackill (2009), based on the

observation that the short conserved regions of plant genes are surrounded by the ATG translation start codon (Sawant et al. 1999; Joshi et al. 1997). The technique uses single primers designed to anneal to the flanking regions of the ATG initiation codon on both DNA strands. The generated amplicons are possibly distributed within gene regions that contain genes on both plus and minus DNA strands (Collar and Mackill 2009).

1.3.5. Intron targeting (IT) Polymorphism

In the solanaceous plants, the relatively conserved nature of the gene structures makes it possible to use intron sequences as molecular markers. This high degree of conservation may be due to Solanaceae genomes having undergone relatively few genomic rearrangements and duplications and therefore having similar gene content and order (Mueller et al. 2005). The close proximity of introns to exons makes them especially well suited for linkage disequilibrium studies that have potential to add a powerful new dimension to understanding and improvement of crop gene pools. One effective strategy for exploiting this information and generating gene-specific codominant markers is a method called Intron Targeting (IT). This method was first applied by Choi et al. (2004) to construct a linkage map of the legume *Medicago truncatula* Gaertn. The basic principle of IT relies on the fact that intron sequences are generally less conserved than exons, and they display polymorphism due to length and/or nucleotide variation in their alleles. Primers designed to anneal in conserved exons to amplify across introns can reveal length polymorphism in the targeted intron.

Such primers can be designed for potato using the available sequences of known genes or by exploiting expressed sequence tag (EST) records from the NCBI database. These marker systems may provide new valuable tools for genetic diversity assessment of germplasm collections as well as in other fields of plant science and breeding. However, little effort has been invested to address the utility of these markers for the above mentioned goals.

CHAPTER 2

Materials and methods

2.1. Laboratory techniques and sampling used in the pilot study

2.1.1. Plant material and DNA extraction

Taxon sampling for the pilot studies included seven species belonging to subg. *Archaesolanum*, with two accessions from each of the species *S. aviculare*, *S. laciniatum* and *S. simile* and one from *S. linearifolium*, *S. capsiciforme*, *S. symonii* and *S. vescum*. Seven *Solanum* species, representing different subgenera, were included in the analysis as outgroups. An accession outside *Solanum, Capsicum annuum* L. was also added in the experiments, according to the results reported by Olmstead et al. (1999) and Bohs and Olmsted (2001). Although we used only one accession for each species in this study, our ongoing continuous studies within different lineages of *Solanum* show that intraspecific variation does not adversely affect phylogenetic analyses between sections of *Solanum*, as this has been also shown previously in other lineages of the genus by Spooner and Systma (1992). Voucher specimens are deposited in the herbarium of the University of Pannonia, Keszthely, Hungary. Information about the accessions can be found in Table 3.

Genomic DNA was extracted from approximately 50 mg of young fresh leaves using the modified procedure of Walbot and Warren (1988). Further details are found in Appendix 1. RAPD fingerprints were obtained from DNA bulks, according to Spooner et al. (1997), where five plants from each accession were bulked for DNA extraction. Although, fragments present in <15% of individuals comprising the DNA bulk are often observed to be lost from the binding patterns of bulked samples (e.g. Divaret et al. 1999). However, we designed our study to sample as many alleles as possible within the accessions. Additionally, the aim was to examine more populations, rather than more individuals within a population. Thus the bulking strategy described by Michelmore et al. (1991) was considered to be useful to generate a group (e.g. population or accession) fingerprint by combining DNA from a number of individuals. This strategy may reduce the noise in the dataset due to markers segregating

within the groups (Bussel et al. 2005) and has been successfully used for species of *Solanum* (Miller and Spooner 1999; Rodríguez and Spooner 1997; Spooner et al. 1991,1993, 1995, 1997; Clausen and Spooner 1998).

2.1.2. RAPD amplification

In the RAPD analysis 20 primer pairs were used. Each reaction was performed twice to verify reproducibility. The primers were paired arbitrarily, but palindromes and complementarity within and between primers were avoided. The sequence of each primer was generated randomly, comprising 12 base oligonucleotides and 50-70% GC content. The sequences of the primers are found in Appendix 2. PCR was carried out on a 96-well RoboCycler (Stratagene, USA) using a 20 µl reaction mix which contained the following: 10 µl sterile ion exchanged water, 5 ng template DNA, 1 µM of each primer, 0.2 mM dNTP (Fermentas, Lithuania), 2 µl 10×PCR buffer (1 mM Tris-HCl, pH 8.8 at 25˚C, 1.5 mM $MgCl_2$, 50 mM KCl and 0.1% Triton X-100) and 0.5 U of DyNazyme II (Finnzymes, Finland) polymerase. Reaction conditions were 1 min at 94˚C, followed by 35 cycles of 30 s at 94˚C, 1 min at 37˚C and 2 min at 72˚C. A final amplification for 5 min at 72˚C was applied. Amplification products were separated on 1.5% agarose gels (Promega, USA) in 0.5×TBE buffer (300V, 1.5 h) and post-stained with ethidium-bromide. The gels were documented using the GeneGenius Bio Imaging System (Syngene, UK). The binding patterns were evaluated and annotated with the program GeneTools (Syngene, UK).

Table 3. Accessions used in the pilot analysis (Poczai et al. 2011b).

Taxonomic position[b]	Taxon	Collector	Accession number	Collection locality	Origin
Ingroup					
Subg. *Archaesolanum*					
Ser. *Avicularia* Geras.	*S. aviculare* (1) Forst.	Unknown	874 750 027	A	I.
	S. aviculare (2) Forst. var. *latifolium* Bayl.	GTS Baylis	844 750 003	NZ	I.
Ser. *Laciniata* Geras.	*S. laciniatum* Ait. (1)[a]	Unknown	A24 750 011	Unknown	I.
	S. laciniatum Ait. (2)[a]	DE Symon	A24 750 098	A	I.
	S. linearifolium Geras.	Unknown	814 750 056	Unknown	I.
	S. vescum F. Muell.	D Martin	904 750 174	A	I.
Ser. *Similia* Geras.	*S. capsiciforme* (Domin) Bayl.	Unknown	884 750 213	Unknown	I.
	S. simile F. Muell. (1)[a]	CR Alcock	A24 750 094	A	I.
	S. simile F. Muell. (2)[a]	Unknown	894 750 053	A	I.
	S. symonii Eichler	N Lovett	844 750 004	A	I.
Outgroup					
Subg. *Leptostemonum*					
Sect. *Acanthophora*	*S. atropurpureum* Schrank.	Unknown	UHBG211-1471	Unknown	II.
Androceras	*S. rostratum* Dunal	Unknown	HU1GEO20060029	Unknown	III.
Cryptocarpum	*S. sisymbriifolium* Lam.	Unknown	HU1GEO20060053	Unknown	III.
Subg. *Minon*					
Sect. *Bravantherum*	*S. abutiloides* Bitter & Lillo	Unknown	UHBG211-1455	Unknown	II.
Subg. *Potatoe*					
Sect. *Dulcamara*	*S. dulcamara* L.	P Poczai	HU1GEO20060017	H	III.
Subg. *Solanum*					
Sect. *Solanum*	*S. americanum* Miller.	BG Redwood	904 750 023	USA	I.
	S. physalifolium Rusby var. *nitidibaccatum* (Bitter) Edm.	Unknown	894 750 076	G	I.
Genus *Capsicum*	*Capsicum annuum* L.	Unknown	884 750 092	Unknown	I.

Note: Collection locality abbreviations: **A**-Australia, **G**- Germany, **H**- Hungary, **NZ**-New Zealand, **USA**- United States of America. Origin abbreviations: **I.-** Botanical and Experimental Garden of the Radboud University, The Netherlands **II**.- Botanical Garden of the University of Hohenheim, Stuttgart, Germany **III.**- Georgikon Botanical Garden of the University of Pannonia, Keszthely, Hungary

[a] Numbers are abbreviations used in the further text, tables and figures to distinguish accessions.

[b] According to D'Arcy (1972, 1992).

2.1.3. SCoT amplification

Twelve SCoT primers were selected for further study, after a screening and optimization process with 25 primers for each method. The selected primers yielded stabile and reproducible banding patterns. Amplification reactions were performed in 10 µl volumes in 384-well plates containing: 5 µl NFW (Nuclease Free Water, Promega), approx. 20 ng template DNA, 0.5 µM of each primer, 0.2 mM dNTP (Fermentas, Lithuania), 1 µl 10xPCR buffer (1 mM Tris-HCl, pH 8.8 at 25˚C, 1.5 mM $MgCl_2$, 50 mM KCl and 0.1% Triton X-100) and 0.5 U of DyNazyme II (Finnzymes, Finland) polymerase. All reactions were performed with a MasterCycler ep384 (Eppendorf, Germany) with the following conditions: 2 min at 94˚C for initial denaturation, 35 cycles of 30 s denaturation at 94˚C, 1 min annealing at 50˚C, and 2 min extension at 72˚C, followed by a final extension for 5 min at 72˚C. Amplification products were separated on 1.5% agarose gels (Promega, USA) in 0.5X TBE buffer (300V, 1.5 h) and post-stained with ethidium-bromide. The gels were documented using the GeneGenius Bio Imaging System (Syngene, UK). Although, SCoTs were found to be reproducible in our experiments - due to higher annealing temperature and longer anchored primers - replicate experiments were performed, containing one negative and positive control, to check the reliability of the primers and the patterns produced. Details about the primers are found in the Appendix 2.

2.1.4. Intron targeting (IT) primer design, amplification and analysis

The intron targeting primer design procedure utilized the blastn search mode of the blastall program from NCBI (http://www.ncbi.nlm.nih.gov/) to find marker candidates with the E-value set to 10^{-20} (up to August 2010). In the first step, sequences of 340 potato genes known to play a role in resistance mechanisms (e.g catalase) and metabolic pathways (e.g. sucrose-synthesis) were screened to find primer candidates. We selected single and low-copy genes as primary targets. In the cases, where the exon-intron structures of the genes were known, flanking primers were designed for the exons to amplify across the intercalated introns. In the second step, ~ 270 *S. tuberosum* EST sequences with unidentified exon-intron structures were analyzed as follows: potato EST sequences were compared with known genes of tomato or *Arabidopsis thaliana* (L.) Heynh in TAIR, The Arabidopsis Information Resource using a set of algorithms implemented in the SPLIGN software tool (Kapustin et al., 2008) to find the putative exon and intron sequences. This was achieved by aligning the

spliced transcript sequence with its parent genomic sequence to identify correct exon-intron junctions. We preferred a product size within the range of 200-1200 bp, and used filtering parameters accordingly to select suitable exons. After locating the precise positions of the introns, the joined sequences of the flanking exons were passed to the primer designer program PRIMER3 (http://frodo.wi.mit.edu/primer3/). Finally a total of 56 oligonucleotide primer pairs were designed from which 29 proved to be useful.

Thirty individuals of a potato F_1 segregating population originating from a cross between the cultivar White Lady and breeding line S440 and twenty-two individuals from representative *Solanum nigrum* L. populations were used to demonstrate the utility of the designed primers and to filter possible disequilibrium. Black nightshade (*Solanum nigrum*) individuals represented three natural populations located in Hungary (Keszthely, Szolnok) and Croatia (Pula) and served as preliminary models for amplification, optimizations and transferability tests before final amplifications in the kangaroo apple taxon set. Voucher specimens were deposited in the University of Pannonia, Keszthely, Hungary (Appendix 3).

After test amplification all 29 designed IT primers were applied in the taxon set used in the RAPD and SCoT analysis. PCR was carried out in 20-µL volume containing 20 ng of genomic DNA, 0.2 mM of dNTPs, 2 µL 1×PCR buffer, 20 pmol of each primer and 0.4 Unit DynaZyme DNA polymerase (Finnzymes, Espoo, Finland). PCR was performed on a RoboCycler 96 (Stratagene, La Jolla, CA) with an initial 3 min of denaturation at 94°C, followed by 35 cycles of 94°C for 1 min, appropriate annealing temperature (Appendix 3) for 1 min, 72°C for 1 min, and a final extension at 72°C for 7 min. The PCR products were separated on 1.5% agarose gel (Promega, Madison, WI) in 0.5×TBE (Tris-HCl-Boric acid-EDTA) buffer (220 V; 1.5 h) and stained with ethidium bromide. In many cases where the polymorphisms manifested as small size difference, the PCR products were separated on 2.5 % Metaphor agarose gels (Cambrex, East Rutherford, NJ). The genetic statistics based on the potato F_1 population of 30 individuals and on the three *Solanum nigrum* populations of 22 individuals were calculated using the program POPGENE (version 1.31; Yeh et al., 1999) including the number of alleles (A), observed heterozygosity (H_O) and expected heterozygosity (H_E). Sequences of the designed intron targeting (IT) primers, population genetic parameters obtained are also found in Appendix 3.

2.1.5. Chloroplast region amplification and restriction digestion

*trn*S-*trn*G region: The chloroplast intergenic spacer between *trn*S and *trn*G was amplified using the primers described by Hamilton (1999). All PCR reactions were performed in a MasterCycler ep384 (Eppendorf, Germany) with the same composition as in the RAPD analysis, except that the $MgCl_2$ concentration was adjusted to 2 mM. The thermal cycler program included an initial denaturation at 94˚C for 4 min; 40 cycles of 94˚C for 45 s, 52˚C for 1 min, 72˚C for 1 min; with a final extension at 72˚C for 7 min, as described by Levin et al. (2006). The *trn*S-*trn*G amplification products were digested with the restriction endonuclease enzymes: *Hinf*I, *Dde*I, *Mbo*I, *Msp*I, *Rsa*I, *Taq*$^{\alpha}$I and *Alu*I (New England Biolabs Inc., USA). The reaction conditions recommended by the supplier were used for all enzymes. The restriction fragments were separated on 2.3% high resolving MetaPhor agarose gel (Cambrex Bio Science Rockland, Inc., USA), after which the electrophoresis products were visualized by ethidium-bromide staining.

*rbc*L1-*rbc*L2 region: The sequence of the large subunit of the ribulose-1,5-bisphosphate carboxylase gene (*rbc*L) was amplified using the primers described by Demesure et al. (1995). The 20 µl reaction solution was the same as that described above. The thermal cycler program was the following: 94˚C for 1 min; 30 s at 94˚C, 1 min at 60˚C and 2 min at 72˚C for 35 cycles; and a final cycle of 4 min at 72˚C. Further information about the primers used in the study is given in Table 4. The *rbc*L1-*rbc*L2 region was digested with the same enzymes except that instead of *Taq*$^{\alpha}$I the *Alu*I and *Hha*I restriction endonucleases were used in the analysis. The separation and visualization procedure was the same as that described above. The enzymes for the analysis were selected based on the virtual digestion of the sequence data of the fragment *trn*S-*trn*G from *S. aviculare*, submitted to the NCBI database by Levin et al. (2005) with the accession number AY555458.

Table 4. Details of the primers used in the study of chloroplast and mitochondrial regions (Poczai et al. 2011b). [a] *trn*S-*trn*G, intergenic spacer between Ser-tRNA and Gly-tRNA; *rbc*L1-*rbc*L2, subunit of the ribulose-1,5-bisphosphate carboxylase gene; *atp*6 (or *atp*F), F0-ATPase subunit 6 gene; *cob*, apocytochrome b gene; *cox*1 (or *cox*I), cytochrome c oxidase subunit 1 gene; *nad*3 (or *nad*C, *nadh*C, *nadh*3, or *nd*3), NADH-ubiquinone oxidoreductase subunit 3 gene; *nad*5a (or *nad*F, *ndh*F, *ndh*5, *nd*5), NADH-ubiquinone oxidoreductase subunit 5 gene (intron 1); *nad5*dF (or *nad*F, *ndh*F, *ndh*5,*nd*5), NADH-ubiquinone oxidoreductase subunit 5 gene (intron 2); *rps*14, ribosomal protein subunit 14 gene; *nad*4exon1 (or *nad*D, *ndh*D, *ndh*4, *nd*4) NADH-oxidoreductase subunit 4 gene (intron 1). [b] Inosine was used in the synthesis instead of the corresponding nucleotide because of the nucleotide variation between plant sequences.

Gene[a]		Primer pairs	Size (bp)	G + C (%)	Annealing tempe-rature (°C)	PCR product (bp)
Chloroplast primers						
*trn*S-*trn*G	*trn*S	5'-GCCGCTTTAGTCCACTCAGC-3'	20	60	52	~700-735
	*trn*G	5'-GAACGAATCACACTTTTACCAC-3'	22	41		
*rbc*L1-*rbc*L2	*rbc*L1	5'-ATGTCACCACCACAAACAGAGACT-3'	24	46	60	~1371
	*rbc*L2	5'-CTTCACAAGCAGCAGCTAGTTCAGGACTCC -3'	31	52		
Mitochondrial primers						
*atp*6F-*atp*6R	*atp*6F[b]	5'-GGAGG(A=I)GGAAA(C=I)TCAGT(A=I)CCAA-3'	22	48	58	~589-610
	*atp*6R	3'-TAGCATCATTCAAGTAAATACA-5'	22	27		
*cob*F-*cob*R	*cob*F	5'-AGTTATTGGTGGGGGTTCGG-3'	20	55	58	~290-313
	*cob*R	3'-CCCCAAAAGCTCATCTGACCCC-5'	22	59		
*cox*1F-*cox*1R	*cox*1F[b]	5'-GGTGCCATTGC(T=I)GGAGTGATGG-3'b	22	59	58	~1466
	*cox*1R	3'-TGGAAGTTCTTCAAAAGTATG-5'	21	33		
*nad*3F-*nad*3R	*nad*3F	5'-AATTGTCGGCCTACGAATGTG-3'	21	48	58	~237
	*nad*3R	3'-TTCATAGAGAAATCCAATCGT-5'	21	33		
*nad*5aF-*nad*5aR	*nad*5aF	5'-GAAATGTTTGATGCTTCTTGGG-3'	22	41	58	~1000
	*nad*5aR	3'-ACCAACATTGGCATAAAAAAAGT-5'	23	30		
*nad*5dF-*nad*5dR	*nad*5dF	5'-ATAAGTCAACTTCAAAGTGGA-3'	21	33	58	~1095-1136
	*nad*5dR	3'-CATTGCAAAGGCATAATGAT-5'	20	35		
*rps*14F-*rps*14R	*rps*14F	5'-ATACGAGATCACAAACGTAGA-3'	21	38	58	~114
	*rps*14R[b]	3'-CCAAGACGATTT(C=I)TTTATGCC-5'	21	38		
*nad*4exon1-*nad*4exon2a	*nad*4exon1	5'-CAGTGGGTTGGTCTGGTATG-3'	20	55	58	~2058
	*nad*4exon2a	3'-TCATATGGGCTACTGAGGAG-5'	20	50		

2.1.6. Mitochondrial region amplification

The universal primers described by Demesure et al. (1995) for the amplification of different mitochondrial regions were tested to detect fragment length polymorphism between the accessions. The contents and concentrations used in the reaction mixture were the same as described in for the RAPD analysis. The PCR program was the following: 94˚C for 1 min; 30s at 94˚C, 1 min at 58˚C and 2 min at 72˚C for 35 cycles; and a final cycle of 4 min at 72˚C. The amplified regions and further information about the primers is summarized in Table 4.

2.2. Data analysis

2.2.1. Band scoring

The evaluation of the RAPD and SCoT binding patterns was carried out with the program GeneTools (Syngene, UK). Kingston and Rosel (2004) described a conservative scoring protocol that was used also here to prevent problems associated with multi-locus methods, e.g. uneven amplification among samples and poor amplification of larger fragments for degraded DNA samples. Only well-resolved, distinct bands were scored. Amplicons found in replicate reactions were considered reliable. The amplified fragments were coded as absent/present (0/1). It was presumed that fragments with equal length had been amplified from homologous loci and represent a single, dominant locus with two possible alleles. To measure the information content detected with each primer the Polymorphic Information Content (*PIC*) value was calculated, according to Botstein et al. (1980). The heterozigosity (*H*) value was also calculated according to Liu (1998). For all calculation the test version of the online program PICcalc was used (Nagy et al. 2008). As *PIC* and *H* are both influenced by the number and frequency of alleles, the maximum number for a dominant marker is 0.5, since two alleles per locus are assumed in the analysis (Henry 1997; De Riek et al. 2001; Bolaric et al. 2005). For all data about the primers and calculated values see Appendix 2.

Intron targeting banding patterns together with the cpDNA PCR-RFLP patterns and the structural mitochondrial PCR amplification fragments were also coded as absent/present (0/1). The only difference was that the Kingston and Rosel (2004) conservative scoring protocol was not applied.

2.2.2. Parsimony analysis of binary data sets

Data obtained from the multi-locus AAD analysis generated by RAPD and SCoT primers were united with the data matrix produced by intron targeting (IT) markers; while the cpDNA restriction patterns were united with the matrix produced by the structural amplified mitochondrial primers. The two data sets (chloroplast-mitochondrial and nuclear multi-locus) were analyzed separately and in combination. Most studies in which AADs have been subject to cladistic parsimony analysis have used Wagner parsimony criterion (Kluge and Farris 1969). The Fitch parsimony criterion (Fitch and Margoliash 1967) is equivalent to Wagner parsimony for binary data (Kitching 1992). This is appropriate where the probability of character state change is unknown or symmetrical (Swofford and Olson 1990; Kitching 1992). Symmetrical characters are freely reversible and changes from 1→0 and 0→1 are defined as equally probable. AAD characters are not freely reversible and may not be suited to Wagner parsimony analysis, since there are many more ways of losing than gaining a fragment (Backeljau et al. 1995). The Dollo parsimony method (DeBry and Slade 1985) was used to overcome these difficulties due to inequity of loss and gain probabilities in AAD data as it was previously suggested (Stewart and Porter 1995; Furman et al. 1997; Harvey et al. 1997). It is constrained by the conditions that each apomorphic character state must be uniquely derived, and that all homoplasy must be accounted for by reversals to more plesiomorphic (ancestral) states (Swofford and Olson 1990). Dollo parsimony has been applied in phylogenetic analysis of RFLP data where there are also skewed probabilities for gain or loss: independent gain of a restriction site in different lineages is so unlikely that taxa sharing a site are presumed to have inherited it from a common ancestor (Holsinger and Jansen 1993). The Dollo criterion is restrictive, effectively precluding the (albeit remote) possibility that a product is gained twice independently (Backeljau et al. 1995). All calculations were carried out using the program package PHYLIP (Phylogeny Inference Package; Felsenstein 1989). The further analysis was carried out using the branch and bound algorithm of the DOLPENNY program to find all of the most parsimonious trees implied by the data.

The data analysis was performed with the use of the Dollo parsimony method using 10,000 bootstrap replications. The program was run with default setting modified to report every 100 trees and 1000 groups. From the resulting output trees a consensus tree was built with the CONSENSE program using the majority rule criterion.

2.2.3. Distance-based analysis of the binary data set

Distance-based methods were included because the parsimony criterion, in particular, may be inappropriate for use with dominant, anonymous (AAD) markers due to the inherent faulty assumption of homology among shared absent markers and the possible parsimonious, but incorrect, reconstruction in which no markers are assigned to an ancestor at a given internal node (Blackeljau et al. 1995; Swofford and Olsen 1990). However, as discussed above Dollo parsimony may overcome these problems. We used the two methods in parallel to check whether the two different assumptions produce same topologies for the obtained dataset. The joined (RAPD, SCoT and IT) presence/absence matrix of homologous bands was used to calculate a distance matrix according to Nei and Li (1979) based on Dice's similarity coefficient (Dice 1945). A dendrogram was constructed using the Neighbor Joining method described by Saitou and Nei (1987); the original matrix was bootstrapped 1,000 times in order to check the reliability of the branching patterns, and the quality of the resulting phylogenetic groups. These bootstrap values are shown at the nodes of the dendrogram as percentages. The FAMD program (Schlüter and Harris 2006) was used for all calculations. The tree obtained using FAMD was visualized and edited using the TreeView program (Page 1996).

2.3. Phylogenetic and laboratory treatment used in sequence analysis studies

2.3.1. Taxon Sampling

A different taxon set was used in sequence based analysis. The outgroup terminals were extended by including species from the genus *Solanum* and other taxa outside the genus, but belonging to the family Solanaceae. The ingroup terminal set was also extended by including herbarium material from kangaroo apples. Overall, three accessions per species were analyzed for *Solanum aviculare*, *S. laciniatum*, *S. linearifolium*, *S. simile*, *S. symonii*, *S. vescum* and two accessions for *S. capsiciforme*. Only one accession was sampled from the rare *S. multivenosum* with only few herbarium records. This is due to its restricted occurrence in the high altitude (> 2,500 m) mountain ranges of Papua New Guinea where kangaroo apples have still been poorly collected. The outgroup exemplars from other *Solanum* subgenera and outside the genus were selected following the results of Weese and Bohs (2007) and Olmstead et al. (2008). For the molecular clock analyses a further outgroup (*Ipomoea purpurea*) was included to represent the split between Solanaceae and Convolvulaceae following Paape et al. (2008). Further information about the terminals is summarized in Table 5.

2.3.2. DNA extraction, PCR amplification

Total genomic DNA was extracted from 50 mg of fresh leaves following the modified protocol of Walbot and Warren (1988). From the herbarium specimens extractions were made with the NucleoSpin 96 Plant Kit (Machery-Nagel) or with a CTAB protocol used by the Biotechnology Group, University of Pannonia (see Appendix 1). Absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) was measured for each DNA sample using the NanoDrop 2000 (Thermo Fisher Scientific, USA) spectrophotometer.

Table 5. Plant material used in the study (Poczai et al. 2011a). [a]Subgeneric names are according to D'Arcy (1972,1991); [b]Major clades after Weese and Bohs (2007). [c]These genera are now nested within the *Solanum* genus. Accession numbers in bold are provided by this study.

Taxa	Subgeneric Rank[a]	Major clade within *Solanum*[b]	Collection Country	Voucher	GenBank accession number
Ingroup					
S. aviculare Forst.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-12	**HM006836**
S. aviculare Forst.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-29	**HM006853**
S. aviculare Forst.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-30	**HM006854**
S. capsiciforme (Domin) Bayl.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-15	**HM006839**
S. capsiciforme (Domin) Bayl.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-35	**HM006859**
S. laciniatum Ait.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-11	**HM006835**
S. laciniatum Ait.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-27	**HM006851**
S. laciniatum Ait.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-28	**HM006852**
S. linearifolium Geras.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-10	**HM006833**
S. linearifolium Geras.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-25	**HM006849**
S. linearifolium Geras.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-26	**HM006850**
S. multivenosum Symon	*Archaesolanum*	Archaesolanum	Papua New Guinea	Symon 13889	**HM006834**
S. simile F.Muell.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-13	**HM006837**
S. simile F.Muell.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-31	**HM006855**
S. simile F.Muell.	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-32	**HM006856**
S. symonii Eichler	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-14	**HM006838**
S. symonii Eichler	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-33	**HM006857**
S. symonii Eichler	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-34	**HM006858**
S. vescum F.Muell	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-09	**HM006832**
S. vescum F.Muell	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-23	**HM006847**
S. vescum F.Muell	*Archaesolanum*	Archaesolanum	Australia	ISZ 10-24	**HM006848**
Outgroup					
S. abutiloides (Griseb.) Bitt.&Lillo	*Minon*	Brevantherum		ISZ 10-06	**HM006829**
S. aggregatum Jacq.	*Lyciosolanum*	African non-spiny	South Africa	RGO 99-25 (WTU)	DQ180460
S. betaceum Cav.	Genus *Cyphomandra*[c]	Cyphomandra	Bolivia	ISZ 10-07	**HM006830**
S. brevicaule Bitt.	*Potatoe*	Potato	Bolivia	Hawkes et al. 6701	DQ180443
S. caesium Griseb.	*Solanum*	Morelloid	Bolivia	ISZ 10-19	**HM006843**
S. citrullifolium A.Braun	*Leptostemonum*	Leptostemonum	Mexico	ISZ 10-03	**HM006826**
S. dulcamara L.	*Potatoe*	Dulcamaroid	Hungary	ISZ 10-16	**HM006840**
S. etuberosum Lindl.	*Potatoe*	Potato	Chile	UAC 1322	DQ180463
S. glaucophyllum Desf.	*Solanum*	Cyphomandra	Argentina	ISZ 10-08	**HM006831**

S. herculeum Bohs	Genus *Triguera*[c]	Normania	Morocco	Jury 13742 (RNG)	DQ180466
S. lycopersium L.	Genus *Lycopersicon*[c]	Potato	Hungary (cult.)	ISZ 10-17	**HM006841**
S. mauritianum Scop.	*Minon*	Brevantherum	Australia	ISZ 10-05	**HM006828**
S. melongena L.	*Leptostemonum*	Leptostemonum	Hungary (cult.)	ISZ 10-04	**HM006827**
S. nitidum Ruiz&Pav.	*Minon*	Dulcamaroid	Bolivia	Nee 31944 (NY)	DQ180451
S. trisectum Dun.	*Potatoe*	Normania	France	Bohs 2718 (UT)	DQ180471
S. tuberosum L.	*Potatoe*	Potato	Hungary (cult.)	ISZ 10-18	**HM006842**
S. villosum L.	*Solanum*	Morelloid	Hungary	ISZ 10-20	**HM006844**
Other genera					
Capsicum baccatum L.	-	-	Bolivia	ISZ 10-21	**HM006845**
Capsicum chacoense Hunz.	-	-	Bolivia	ISZ 10-22	**HM006846**
Ipomea purpurea (L.)Roth	-	-	Unknown	Unknown	EU118126
Jaltomata procumbens (Cav.) J. L. Gentry	-	-	Mexico	Davis 1189A	DQ180419
Jaltomata sinuosa (Miers) Mione	-	-	Mexico	Nee et al. 51830 (NY)	DQ180418
Physalis alkekengi L.	-	-	Hungary (cult.)	ISZ 10-01	**HM006824**
Physalis alkekengi L.	-	-	Hungary (cult.)	ISZ 10-02	**HM006825**

Sample concentration was calculated by the NanoDrop nucleic acid application module using Beer's law, and assuming 50 ng cm/ml absorbance for dsDNA, A_{260}/A_{280} ratios averaged 1.79±0.12 SD. Each sample was diluted to 20ng/µl final concentration. The complete *trn*T-*trn*F chloroplast region (Fig.4) was amplified in three overlapping fragments with the universal primers designed by Taberlet et al. (1991). The *trn*T^{UGU}-5'*trn*L^{UAA} spacer region was amplified with primers *A* (5'-CAT TAC AAA TGC GAT GCT CT-3') and *B* (5'-TCT ACC GAT TTC GCC ATA TC-3'); 5'*trn*L^{UAA}-3'*trn*L^{UAA} intron with primers *C* (5'-CGA AAT CGG TAG ACG CTA CG-3') and *D* (5'-GGG GAT AGA GGG ACT TGA AC-3'), and finally the 3'*trn*L^{UAA}-*trn*F^{GAA} spacer region with primers *E* (5'-GGT TCA AGT CCC TCT ATC CC-3') and *F* (5'-ATT TGA ACT GGT GAC ACG AG-3'), respectively.

Amplification reactions were performed in 50 µl volumes containing: 25 µl NFW (Nuclease Free Water), approx. 20 ng template DNA, 0.5 µM of each primer, 0.2 mM dNTP

(Fermentas, Lithuania), 5 µl 10xPCR buffer (1 mM Tris-HCl, pH 8.8 at 25˚C, 1.5 mM $MgCl_2$, 50 mM KCl and 0.1% Triton X-100) and 0.5 U of DyNazyme II (Finnzymes, Finland) polymerase. All reactions were done in a MasterCycler ep96 (Eppendorf, Germany) with the following settings: 2 min at 94˚C for initial denaturation, 35 cycles of 30 s denaturation at 94˚C, 1 min annealing at 50˚C, and 2 min extension at 72˚C, followed by a final extension for 5 min at 72˚C. Amplification products were separated on 1.5% agarose gels (GE Healthcare, UK) in 0.5X TBE buffer (220V, 0.5 h) and stained with ethidium-bromide.

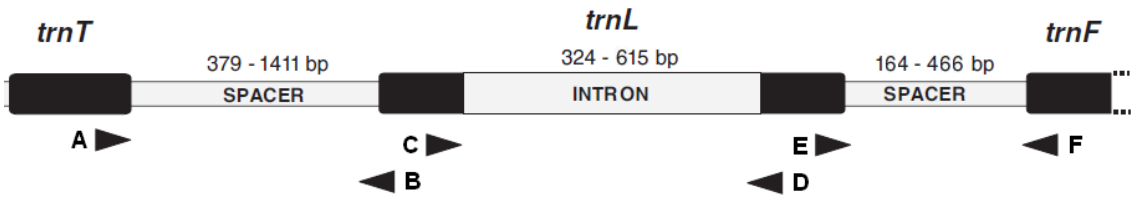

Fig.4. Structure of the *trn*T-*trn*F region in basal angiosperms and other seed plants based on the study of Borsch et al. (2003). tRNA genes (*trn*T and *trn*F are each 73 bp long) and exons (*trn*L-5' is 35 bp and 3' is 50 bp) are represented by black boxes. The spacers and the intron are illustrated as grey bars after Löhne et al. (2008). Minimum and maximum sizes of the spacers and intron among taxa are indicated above the bars. Positions of primers by Taberlet et al. (1991) are marked by arrows. (Original figure kindly provided by C. Löhne).

2.3.3. Cloning and sequencing of PCR products

Fragments excised from agarose gels or direct PCR products were cleaned with NucleoSpin Extract II Kit (Machery-Nagel, Germany) and cloned to JM107 competent *Escherichia coli* strains using ColneJET PCR Cloning Kit (Fermentas, Lithuania) and the Transform Aid Bacterial Transformation Kit (Fermentas, Lithuania). The procedure was carried out according to the manufacturer's instructions.

Plasmids were extracted from the selected colonies holding the desired insert with NucleoSpin Plasmid DNA Kit (Machery-Nagel, Germany). All DNA sequencing was performed on an ABI 3730XL automated sequencer using ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit v. 3.0 (Perkin-Elmer/Applied Biosystems, California, USA) with the pJET1.2 forward and reverse primers. The sequencing of three separate

plasmids per each analyzed terminal was done in both directions to detect possible ambiguities. Detailed protocols for PCR product purification and cloning are found in Appendix 4-8.

2.3.4. Sequence alignment and phylogenetic analysis

Sequences were aligned with ClustalW implemented in BIOEDIT (Hall 1999) using the default parameters of a match score of 5, a mismatch penalty of 4, a gap open penalty of 15 and a gap extension penalty of 6.66. The matrix is available from TreeBASE (Submission ID 10328).

Phylogenetic analyses with parsimony as the optimality criterion were performed using Nona (Goloboff 1999) within the Winclada (Nixon 2002) shell. Because sequences of species with multiple entries did not show any variation within species, only one for each species was included in the analyses, thus resulting in 31 terminals. Four different analyses were made with the following settings: hold* (number of trees held in memory, * denoting as many as the memory allows), mult 5,000 (number of search replicates), hold/2 (number of randomly chosen starting trees per replicate), and using mult*max* (multiple tree-bisection-reconnection algorithm). The second analysis was performed with the same settings, only with an increased number of starting trees (hold/20). We employed also the parsimony ratchet (Nixon 1999) in two additional analyses with the following settings: 1,000 replicates, two trees hold per iteration, 400 characters (ca. 20%) reweighted, and with amb-poly = (default setting of Nona: if any of the reconstructed states are shared between ancestor and descendant node, the branch is collapsed). Another ratchet analysis was performed with the same settings but with an increased number of trees held per iteration (hold/10), and with 600 characters (ca. 30%) reweighted.

Jackknife (Farris et al. 1996) support values were calculated with 10,000 replicates using the Mac OSX version of the program TNT (Goloboff et al. 2008).

2.3.5. Molecular clock and divergence time estimation

For the selection of the appropriate nucleotide substitution model we used jModelTest (Posada 2008) to calculate the probabilities of changes between nucleotides along branches of phylogenetic trees using the Bayesian Information Criterion (BIC). For the parsimony analyses the original combined chloroplast sequence matrix was pruned for molecular clock

dating, keeping only one representative sequence from each species of subg. *Archaesolanum*, since there was no intraspecific variation. This was done to shorten the computational time needed for each run. For this data matrix the GTR+Γ (General Time Reversible + Gamma) was determined as the best-fitting statistical model. Tree topology, node ages and substitution rates were simultaneously estimated with Bayesian Markov chain Monte Carlo (MCMC) simulations using BEAST v. 1.5.2 (Drummond and Rambaut 2007). We employed an uncorrelated and lognormal distributed relaxed clock (UCLD), implemented in BEAST, which allowed different branches of the trees to have independent clock rates, making no assumption about the correlation between substitution rates in the tree. The value of mean substitutions per million years (MY) was fixed at 0.0007, according to the estimates of previous studies concerning coding and non-coding regions of cpDNA (Palmer 1991; Schnabel and Wendell 1998). We assumed a constant rate of speciation per linage and selected the Yule Speciation Process as the tree prior, which has been recommended for species-level phylogenetics (Drummond et al. 2006). Operators were auto-tuned and the starting tree was chosen to be randomly generated. Fossil dates were used as calibration points to reduce the bias and generate more accurate age estimations (Crepet et al. 2004). However, fossil records of the Solanaceae are very limited; the earliest ones of *Solanum*- and *Physalis*-like seeds are from the mid-Miocene and for Convolvulaceae from the lower Eocene (Benton 1993). These records have been previously regarded as sufficiently reliable (Paape et al. 2008) but conclusive analyses about their relationships have not been made. The clades of interest were defined via most recent common ancestors (MRCA), and they could thus have varying taxon composition in the posterior.

Normal distribution priors were applied to the calibration points at nodes: the MRCA of *Solanum* and *Physalis*, with a mean age of 10 MY and a standard deviation (SD) of 4 MY; and the MRCA of *Ipomoea purpurea*, with a mean age of 52 MY and an SD value of 5.2 MY to represent the split between Solanaceae and Convolvulaceae as a minimum age constraint (Magallón et al. 1999). The SD values represent the upper and lower bounds of geological epochs from which the fossils were obtained as previously proposed by Paape et al. (2008). Two separate runs were performed with 10 M generations sampling every 1,000th tree, with a burn-in of 1 M generations each. The results of the individual runs were checked for convergence and analyzed with Tracer v. 1.5 (Rambaut and Drummond 2007), then combined into one consensus log file with LogCombiner v. 1.5.2 (Drummond and Rambaut 2007), as

recommended for phylogenetic MCMC analysis, instead of running single but considerably longer chains (Beiko et al. 2006). The effective sample sizes (ESS) for all estimated parameters were well above 100. The resulting trees were finally combined into one consensus tree with TreeAnnotator 1.5.2 (Drummond and Rembaut 2007) and edited with FigTree v. 1.3.1 (Rambaut 2008). In order to evaluate temporal variation in rates of diversification a lineages-through time plot was calculated based on BEAST estimates.

2.3.6. Geospatial analysis

GenGIS version 1.08 (Parks et al. 2009) was used to combine distributional data with the phylogenetic information obtained from the parsimony analysis. The location set (simple sites) file is available as Appendix 9. This was done to track the historical processes that might be responsible for the current distribution of lineages as well as to analyze and visualize past events that might be inferred. To this end GenGIS provided a 3D graphical interface to combine georeferenced genetic data into a cartographic display to yield a clear view of the relationships between phylogenetic relationships and distribution. In the first step digital map data were obtained with GenGIS MapMaker (Parks and Beiko 2010) using the world maps provided by Natural Earth (http://www.naturalearthdata.com/). The region of interest within the world map was selected and cropped in order to create georeferenced maps specific to the kangaroo apple data set. The location file was generated manually in a comma-separated format containing the taxon set with a unique location identifier and their decimal degrees of latitude and longitude.

We used the consensus tree of the final parsimony analysis as an input phylogenetic tree. It should be noted that for *Archaesolanum* all of the EPTs were completely congruent and thus the consensus tree was fully resolved. The digital elevation map (DEM) of the South Pacific region together with the location and phylogenetic tree file was loaded to GenGIS. To determine the optimal tree layout, GenGIS uses a branch-and-bound algorithm to find the optimal ordering of leaves and internal nodes of a tree that minimizes the number of generated crossings (Parks et al. 2009). Crossings are the mismatches between tree ordering and sample sites measuring the fit between geography and phylogeny. Linear geographic axes were defined arbitrarily and used for the 2D transformation of cladograms. To examine the possible relationships between terminals and latitude we performed a linear regression analysis based on these transformations. This was done by employing a permutation test in 1,000 iterations

implemented in GenGIS to determine whether the fit of tree leaves to geography is significantly better than random, based on randomization of leaf nodes (Parks et al. 2009).

2.3.7. Historical biogeography

In order to explain the current distribution (Fig. 3) of kangaroo apples we inferred the probable location of an ancestral geographic range by performing event-based dispersal-vicariance analysis (DIVA; Ronquist 1997) and weight-ancestral area analysis (WAAA; Huasdorf 1998). Geographical areas were defined based on the distribution of endemic taxa and previously recognized regions in Australia for kangaroo apples by Symon (1994). The areas considered were the Coastal Southern West and South Australia (A), Inland Southern West and South Australia (B), South New South Wales (C), East New South Wales (D), East Victoria (E), Tasmania (F), South Victoria (G), Queensland (H), New Zealand (I), and Papua New Guinea (J). Distributional data were derived from the monograph of Symon (1981) and from the Global Biodiversity Information Facility (GBIF) portal (http://data.gbif.org/species/).

For the WAAA analysis an area cladogram was obtained from the results of the BEAST and parsimony analyses (both resulted in identical topologies). The taxon-area cladogram was constructed by replacing the names of the terminals with the names of the areas in which they occur. Areas were treated as single transformation series.

WAAA estimates ancestral areas using reversible parsimony, and weights areas in plesiomorphic branches more than apomorphic branches (Hausdorf 1998). Therefore, gains and losses of an area are counted separately. The number of weight gain steps (GSW) and weight loss steps (LSW) on the area cladogram were calculated for each area at each node. The final probability indices (PI) were calculated as the ratio of GSW and LSW for every area.

Input Nexus files for DIVA 1.2 (Ronquist 1997) were evaluated manually and contained different area characters as absent (0)/present (1) for each species. DIVA assumes vicariance as the default option of speciation and counts steps for dispersal without signing any cost for vicariance. We constrained the distributional areas to include a maximum of four areas because increasing the number beyond this would provide results with innumerable combinations.

The distributions of climatic conditions affecting the group were analyzed and visualized using SEEVA v.0.33 (Heiberg 2008). Data files from SEEVA were prepared in an Excel file, containing the distribution presence (1) /absence (0) matrix of species in four predefined climatic zones [semi-arid (A), temperate (B), subtropical (C) and tropical (D)]. Areas were delimited based on the updated Köppen-Geiger climate map of Australia (Peel et al. 2007) and the distributional data of the species (Symon 1994). The Excel file together with the NEXUS file containing the pruned phylogenetic tree used previously for DIVA was loaded into SEEVA. Impact indices were calculated for each node on the tree according to Heiberg and Struwe (2008). The percentages of climatic zone categories in each sister group were mapped onto the phylogenetic tree.

CHAPTER 3

Results

3.1. Results of the pilot study

3.1.1. Multi-locus (RAPD, SCoT and IT) analysis

The 20 RAPD primer pairs generated 295 reliable fragments from all the accessions analyzed. The 12 SCoT and the 29 IT primers yielded 95 and 109 scored characters, respectively. AAD analysis using primer combinations clearly separated the accessions used in the study. The joined matrix contained 499 characters from which 249 were parsimony informative (145 RAPD; 55 SCoT; IT 49). The Dollo parsimony analysis conducted with the AAD dataset resulted in 35 equally parsimonious trees. The resulting trees of the parsimony and NJ analysis were identical. The resulting NJ tree is presented in Figure 5. Two main groups can be observed on Fig. 5: the first group includes the two accessions of *Solanum aviculare* the two accessions of *S. laciniatum*; *S. vescum* and *S. linearifolium* which exhibit close affinity based on the bootstrap values. *S. vescum* is sister to this group composed of *S. aviculare* and *S. laciniatum*; *S. linearifolium* occupies a basal position in this clade.

In the other main clade *S. capsiciforme* is separated from the other members of the group. *S. symonii* and the two accessions of *S. simile* are grouped together. This clade has a 100% bootstrap value.

3.1.2. Chloroplast and mitochondrial region analysis

One amplification product per sample was obtained for all the accessions examined. The size of the fragments amplified with *trn*S-*trn*G primers was approximately 710 bp among the species. The size of the *rbc*L fragment was approximately 1,370 bp and showed no variation.

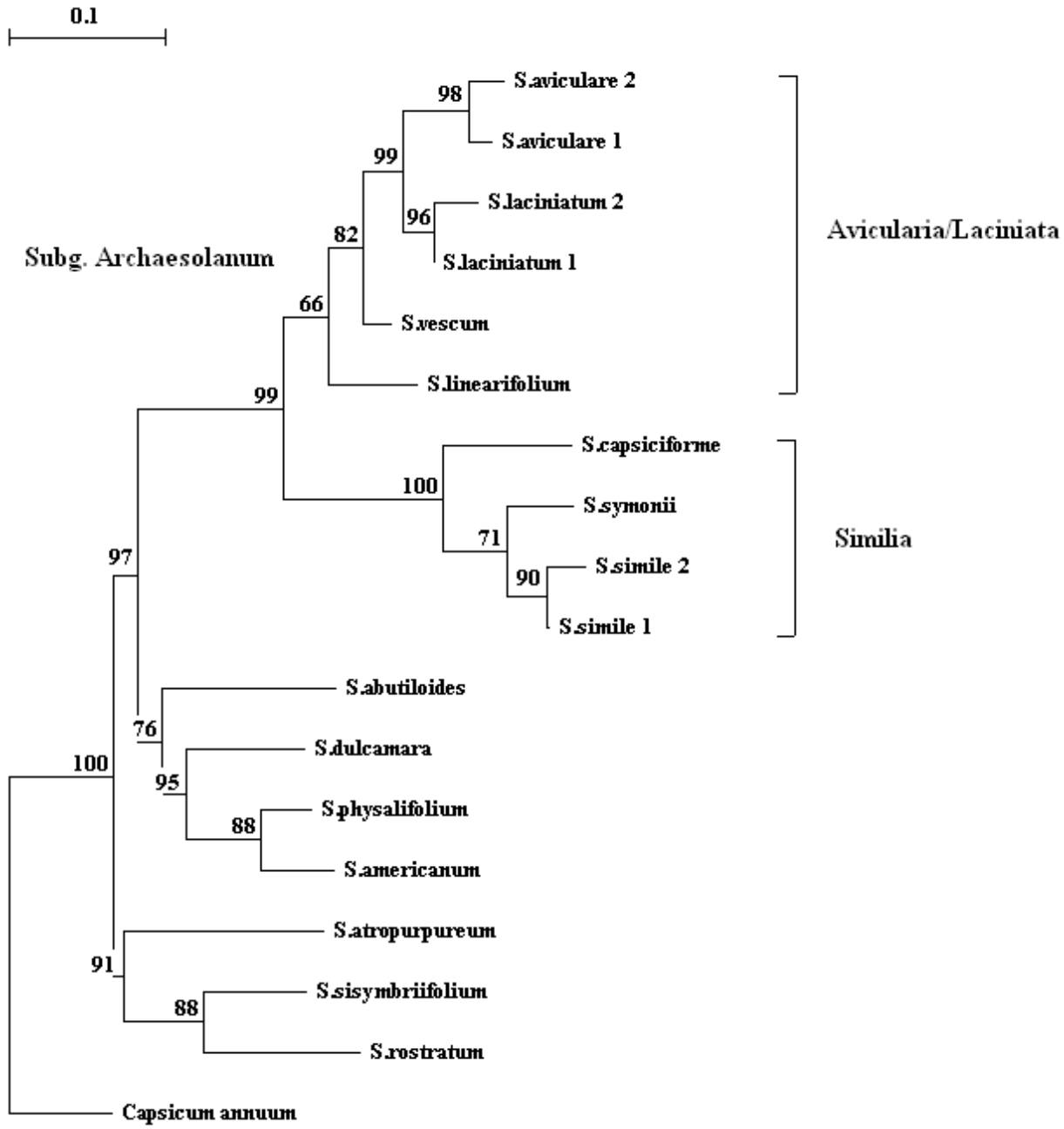

Fig. 5. Dendrogram constructed with the Neighbor Joining (NJ) method from the AAD (RAPD, SCoT and IT) data matrix calculated using the formula given by Nei and Li (1979).

Restriction endonuclease enzyme assay with seven different enzymes resulted in variable fragments in the case of the *trn*S-*trn*G region. Approximately ten bp differences were observed between the restriction fragments. The PCR-RFLP fragments of the region clearly

separate two groups in the subgenus. *S. laciniatum*, *S. aviculare* and *S. vescum* compose a group with larger fragments, while *S. simile*, *S. symonii* and *S. capsiciforme* gave smaller restriction products. *S. linearifolium* occupied an intermediate position between the two groups according to the fragment lengths of the PCR-RFLP analysis. An example is shown at Figure 6. The observed variability in the restriction products indicates that there are differences between the Archaesolanum species in the *trn*S-*trn*G region. The restriction patterns are summarized in Table 6.

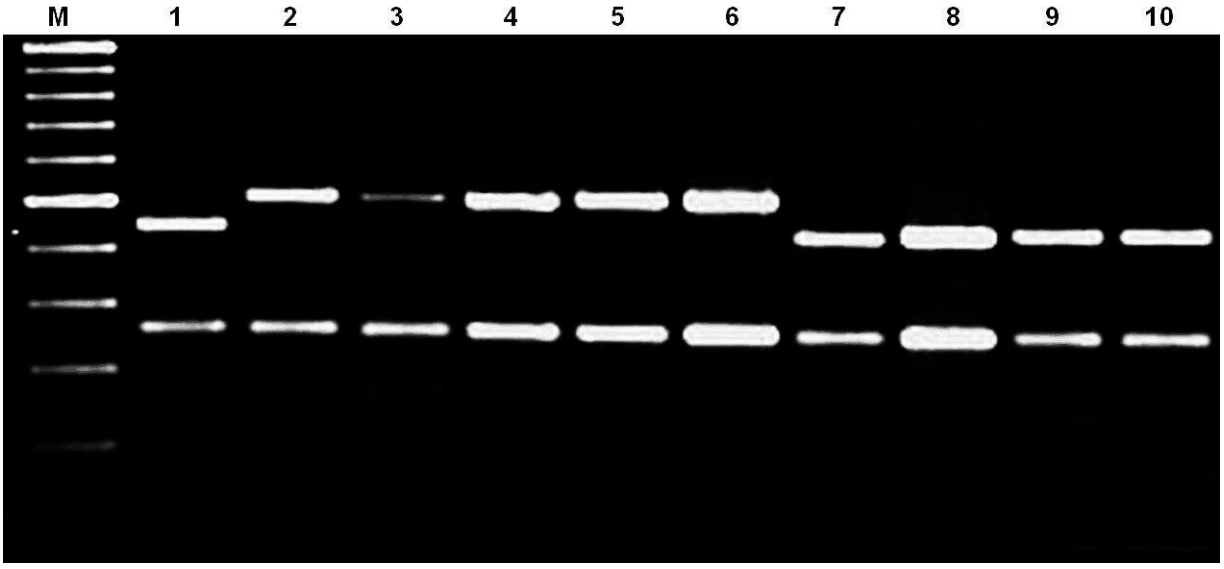

Fig.6. Restriction patterns of the *trn*S-*trn*G chloroplast region digestion with *Mbo*I obtained for kangaroo apple species (subg. *Archaesolanum*). The restriction patterns show an approx. 10 bp difference between the members of the *Avicularia/Laciniata* clade and the taxa of the *Similia* clade. **M**- Molecular weight size marker (100bp to 1,000bp). **1.** *Solanum linearifolium*; **2**. *S. aviculare* (1), **3.** *S. aviculare* (2); **4.** *S. vescum*; **5.** *S. laciniatum* (1); **6.** *S. laciniatum* (2); **7.** *S. capsiciforme*; **8.** *S. symonii*; **9.** *S. simile* (1); **10.** *S. simile* (2). Figure by Poczai et al. (2011b).

The *rbc*L1-*rbc*L2 region was uninformative in the case of subg. *Archaesolanum*. After digestion with six restriction endonuclease enzymes, no polymorphism was detectable. However, the selected enzymes had four to nine restriction sites in the 1,370 bp sequence.

This homology between *Archaesolanum* species and their difference from the outgroup *Capsicum annuum* indicates, that, although this region is uninformative at the subgeneric level, it could be useful in higher taxonomic analyses. Digestion with the *Mbo*I, *Hha*I and

*Rsa*I enzymes resulted in polymorphic restriction patterns in *C. annuum*. As the *rbc*L region was uninformative in the case of kangaroo apples it is not discussed below.

Table 6. Restriction patterns of *trn*S-*trn*G; number of fragments obtained in the analysis (Poczai et 2011b).

Accession	Restriction enzyme						
	Hinf I	*Dde* I	*Mbo* I	*Msp* I	*Rsa* I	*Taq*$^{\alpha}$I	*Alu* I
Ingroup							
S. aviculare (1)	2	2	2	2	3	5	4
S. aviculare (2)	2	2	2	2	3	5	4
S. capsiciforme	2	2	2	2	3	4	3
S. laciniatum (1)	2	2	2	2	3	5	3
S. laciniatum (2)	2	2	2	2	3	5	3
S. linearifolium	2	2	2	2	3	5	4
S. simile (1)	2	2	2	2	3	4	4
S. simile (2)	2	2	2	2	3	4	4
S. symonii	2	2	2	2	2	4	3
S. vescum	2	2	2	2	3	5	4
Outgroup							
S. abutiloides	2	3	0	2	3	4	2
S. americanum	2	3	2	2	2	3	2
S. atropurpureum	2	2	0	0	2	5	0
S. dulcamara	2	2	2	2	3	4	2
S. physalifolium	2	3	2	2	3	4	2
S. rostratum	2	2	0	2	3	6	3
S.sisymbriifolium	2	2	0	2	3	4	3
C. annuum	2	2	2	2	3	3	2

Eight mitochondrial regions were amplified to detect different fragment length patterns which could be attributed to the unique replication of the mitochondrial genome. For this reason the universal primers described by Demesure et al. (1995) were used, to reveal this type of variation. The size of the PCR products and the sequence of the primers are given in Table 4. Polymorphic fragments were detected with the use of *nad*5aF-*nad*5aR while the others resulted in monomorphic patterns.

The *nad*5a primers are designed for the amplification of an intron between exon 1 (*nad*5aF) and exon 2 (*nad*5aR) of the NADH-ubiquinone oxidoreductase subunit 5 gene. While all the species had a 1,000 bp fragment, for *S. vescum*, *S. laciniatum* and *S. linearifolium* an approx. 690 bp size band was also detected. Beside these, *S. linearifolium* also had a unique fragment of approximately 880 bp (Fig.7).

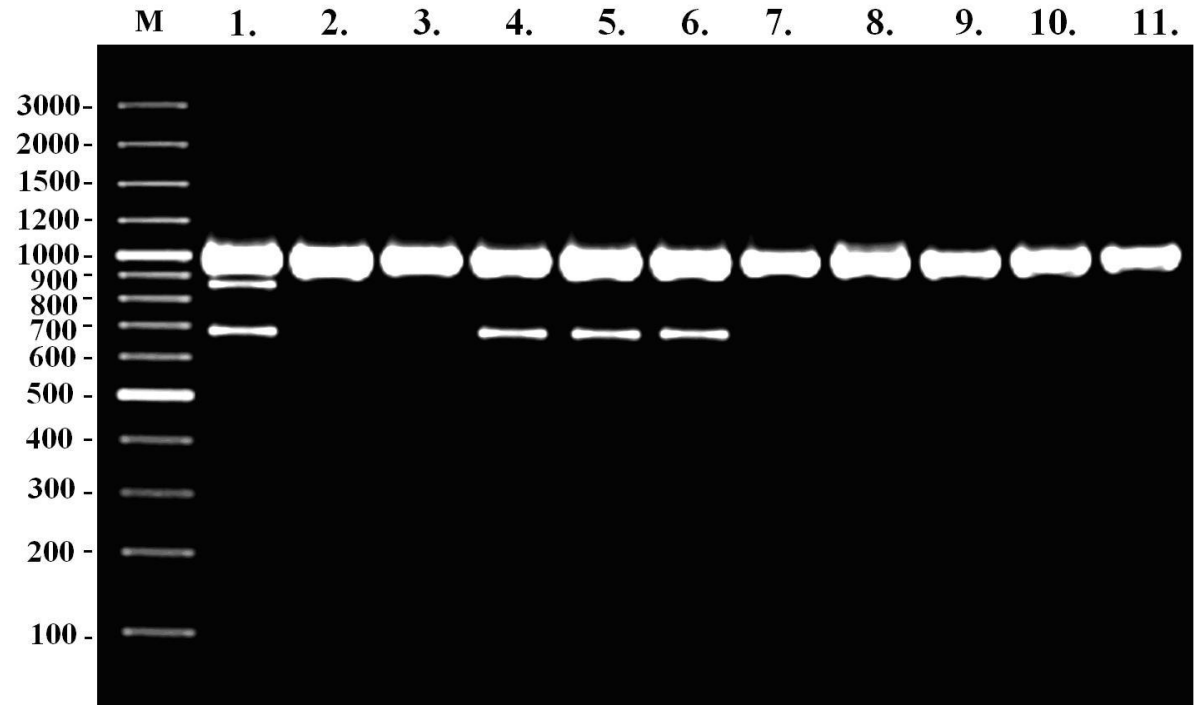

Fig.7. Amplification products of *nad*5aF-*nad*5aR. **M**- Molecular weight size marker (100bp), the numbers indicate ladder size in bp. **1.** *Solanum linearifolium*; **2**. *S. aviculare* (1), **3.** *S. aviculare* (2) **4.** *S. vescum*; **5.** *S. laciniatum* (1); **6.** *S. laciniatum* (2); **7.** *S. capsiciforme*; **8.** *S. symonii*; **9.** *S. simile* (1); **10.** *S. simile* (2*);* **11.** *Capsicum annuum.* Figure by Poczai et al. (2011b).

3.1.3. Results of the cpDNA and mtDNA based parsimony analysis

The phylogenetic tree obtained from the analysis of the joined chloroplast and mitochondrial data set is given in Figure 8. All analysis performed with the joined dataset composed of 60 characters from the cpDNA PCR-RFLP (22 *trn*S-trnG; 22 *rbc*L) and mtDNA amplifications (16) resulted in the same set of four equally parsimonious trees (EPTs).

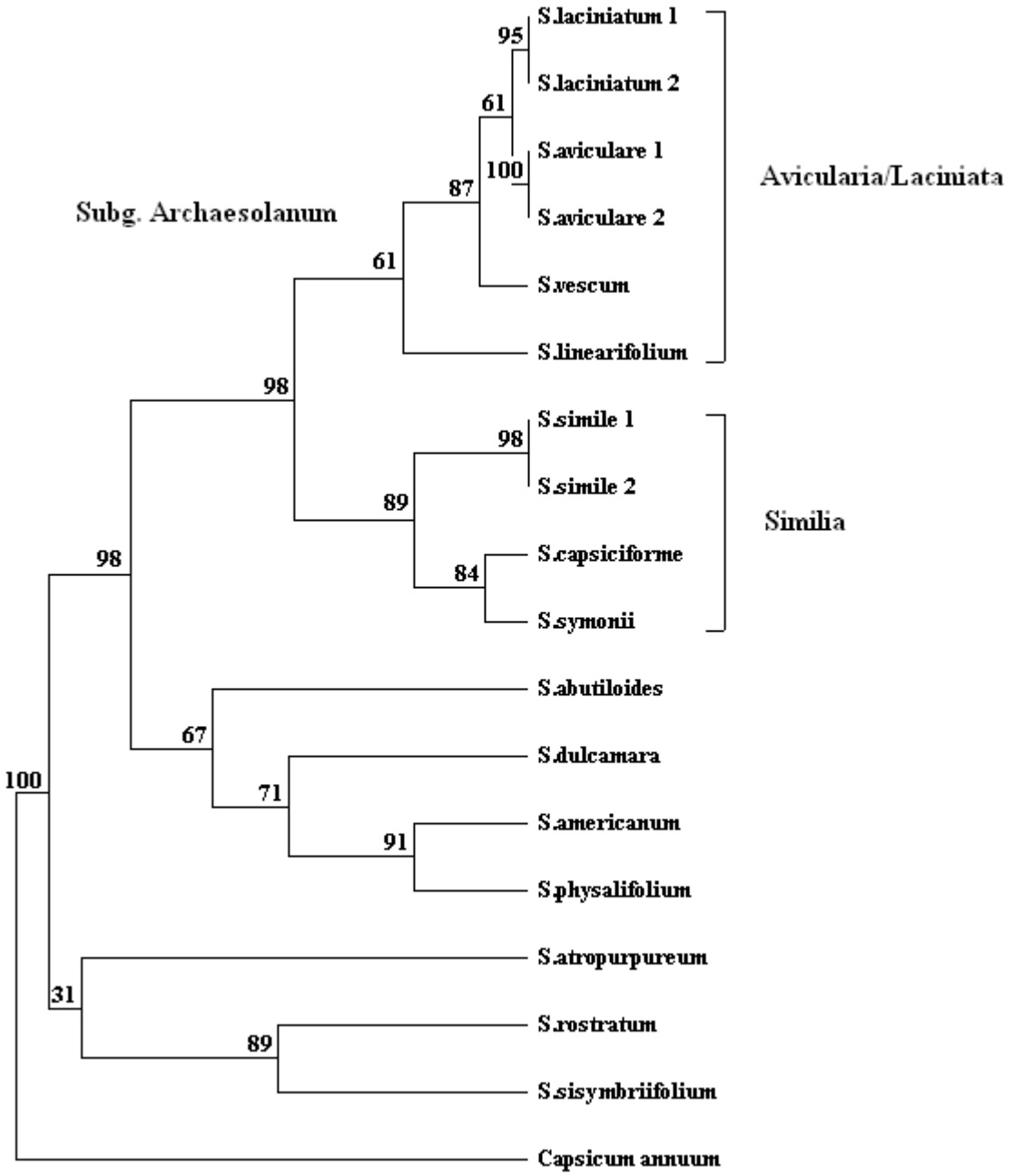

Fig.8. The majority rule consensus tree constructed from the four equally parsimonious trees (EPTs) showing phylogenetic relationships and major clades in subg. *Archaesolanum*. The dataset contained 60 characters and obtained from the cpDNA PCR-RFLP analysis (22 *trn*S-trnG; 22 *rbc*L) as well as from the mtDNA amplifications (16) using universal primer sets designed to anneal in plant mtDNA regions. The numbers above tree nodes are bootstrap replicates in percentage (Poczai et al. 2011b).

The topology of the tree constructed from the chloroplast region restriction data was almost identical to that of the AAD data. The major difference is that *S. symonii* and *S.*

capsiciforme formed a group together in the case of the chloroplast-mitochondrial tree. This arrangement was due to the absence of a restriction site with *Rsa*I in the *trn*S-*trn*G sequence unique for these species (Fig. 9).

The simultaneous analysis of the cpDNA and mtDNA dataset with the AAD data did not modify the resulting topology from those shown in Figure 5. This may be due to that major portion of this dataset were generated by the AAD markers and compared to these the chloroplast-mitochondrial dataset contained less characters with less informative sites.

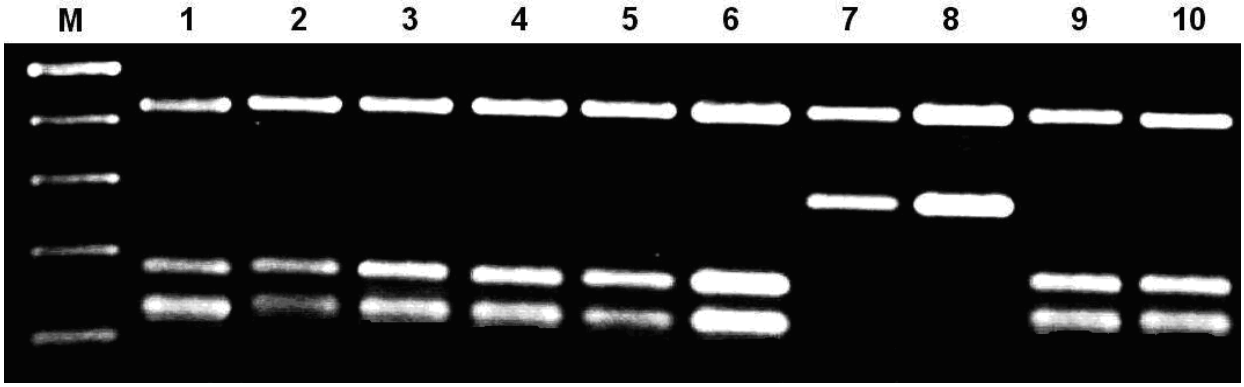

Fig.9. Restriction patterns of the *trn*S-*trn*G chloroplast region digestion with *Rsa*I showing a unique cleavage shared by *Solanum symonii* and *S. capsiciforme*. **M**- Molecular weight size marker (100bp to 500bp). **1.** *S. linearifolium*; **2**. *S. aviculare* (1) **3.** *S. aviculare* (2) **4.** *S. vescum*; **5.** *S. laciniatum* (1); **6.** *S. laciniatum* (2); **7.** *S. capsiciforme*; **8.** *S. symonii*; **9.** *S. simile* (1); **10.** *S. simile* (2). After Poczai et al. (2011b) modified.

3.2. Results of chloroplast sequence analysis

3.2.1. Phylogeny and age estimations

The parsimony analysis resulted in 138 distinct equally parsimonious trees of length 853 steps (Fig. 10). The values of consistency (Kluge and Farris 1969) and retention (Farris 1989) indices were 0.85 and 0.78, respectively. The consensus tree (Nixon and Carpenter 1996) based on the 138 trees did not differ from the one based on 16 trees obtained in the first analysis. Both the parsimony and BEAST analyses resulted in congruent and completely resolved phylogenies of the kangaroo apples. The topology of the outgroup nodes was largely consistent with previously published phylogenetic trees (Bohs 2005; Weese and Bohs 2007;

Poczai et al. 2008). The results are illustrated in the chronogram (Fig. 11). The estimated value of the standard deviation of the uncorrelated lognormal relaxed clock (ucld.stdev) was 1.537 (95% HPD: 1.104-1.959) while the coefficient of variation was 1.909 (95% HPD: 0.967-2.864). These values indicate the presence of substitution rate heterogeneity across the tree and that the chosen relaxed clock model is the most appropriate. The Yule prior indicating the birth rate (i.e., speciation rate) was 0.071 (95%HPD: 0.039-0.106). The tree separates two well resolved clades: i) a *Similia* clade (*S. simile*, *S. symonii*, *S. capsiciforme*); ii) an *Avicularia-Laciniata* clade with two subclades composed of *S. aviculare*, *S. laciniatum*, *S. multivenosum*, and *S. linearifolium* plus *S. vescum*. All of these clades had high Bayesian posterior probabilities, with values ranging from 0.9008 to 1.0. The MCMC runs gave an estimated tMRCA of 17.95 MY (95% HPD: 12.41-23.86) for the genus *Solanum* and 8.85 MY (95% HPD: 4.45-14.03) tMRCA for the subgenus *Archaesolanum*. The separation of Convolvulaceae and Solanaceae was estimated to be 61.95 MY (95% HPD: 38.91-91.59).

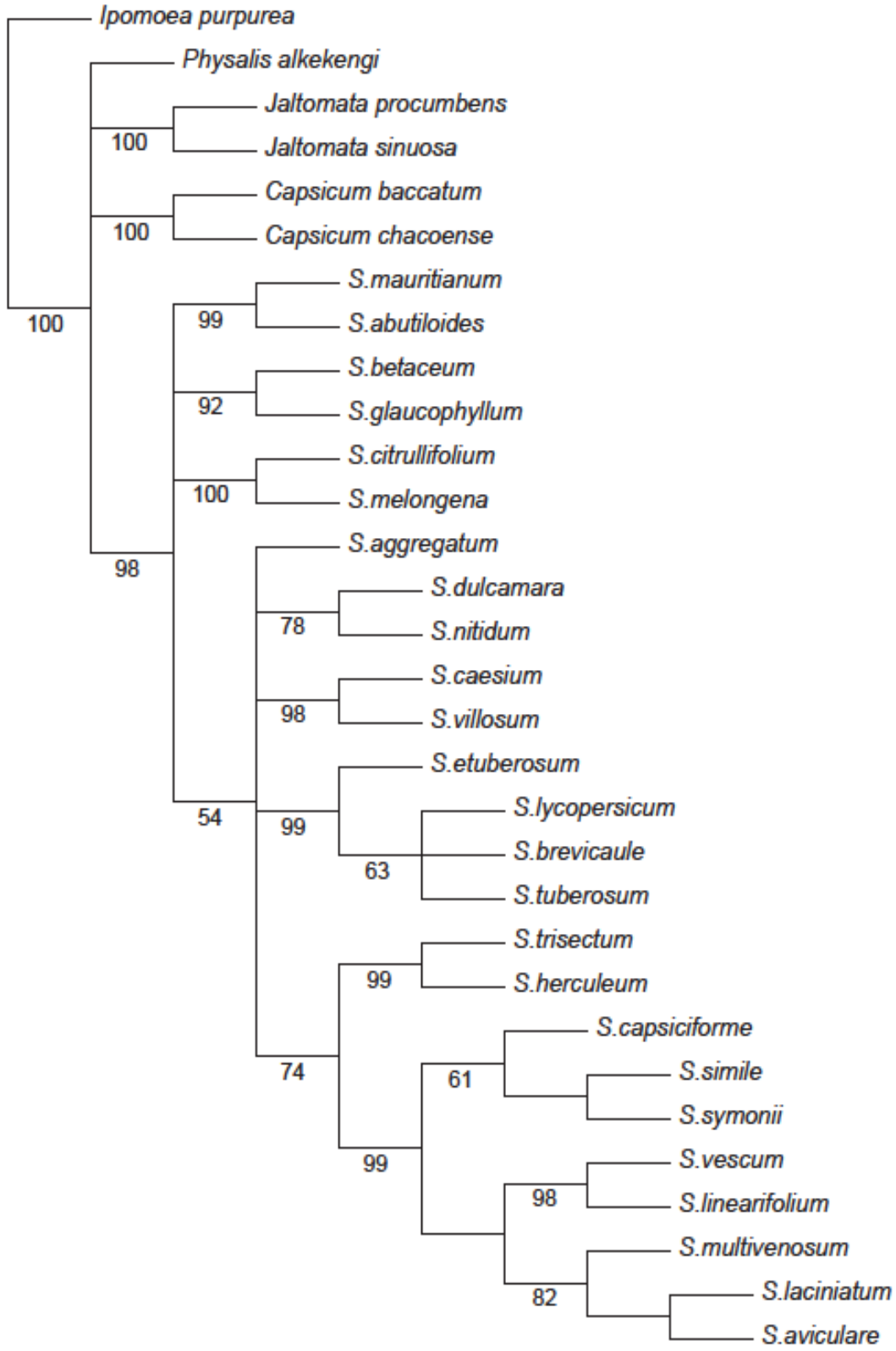

Fig.10. Consensus based on 68 equally parsimonious trees (EPTs) obtained in the four parsimony analyses (see the text for further details) (Poczai et al. 2011a).

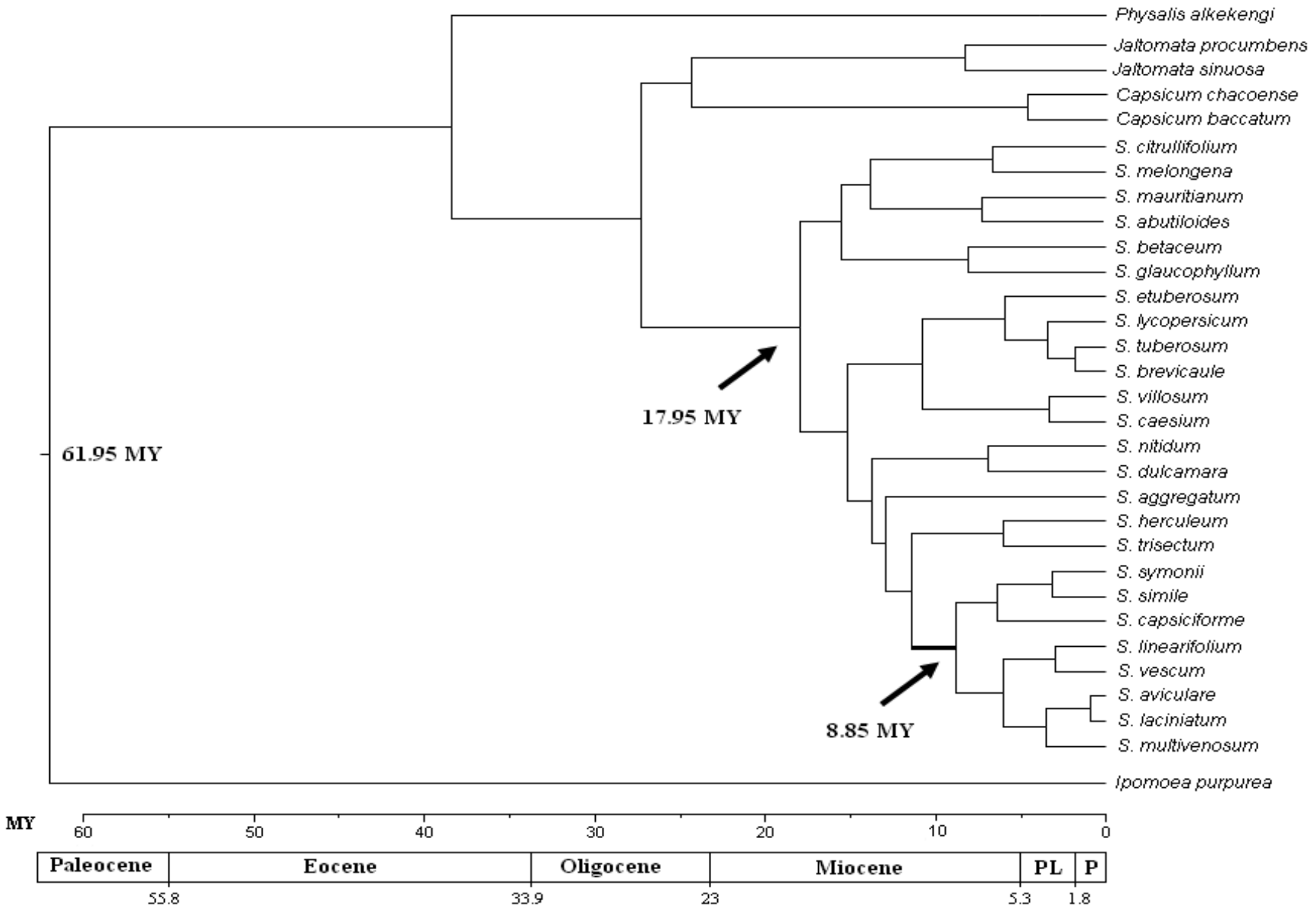

Fig.11. Bayesian chronogram from the BEAST analysis. Geological time-scale abbreviations: PL, Pliocene; P, Pleistocene (Poczai et al. 2011a).

3.2.2. Geospatial analysis

Georeferenced trees were drawn in two (Fig. 12) and in three dimensions to visualize the ecological and phylogenetic relationships among samples collected from different sites. In order to find the optimal matches between the obtained phylogenetic tree and geography we defined geographical polylines in different orders and calculated *p*-values expressing the significance of each choice.

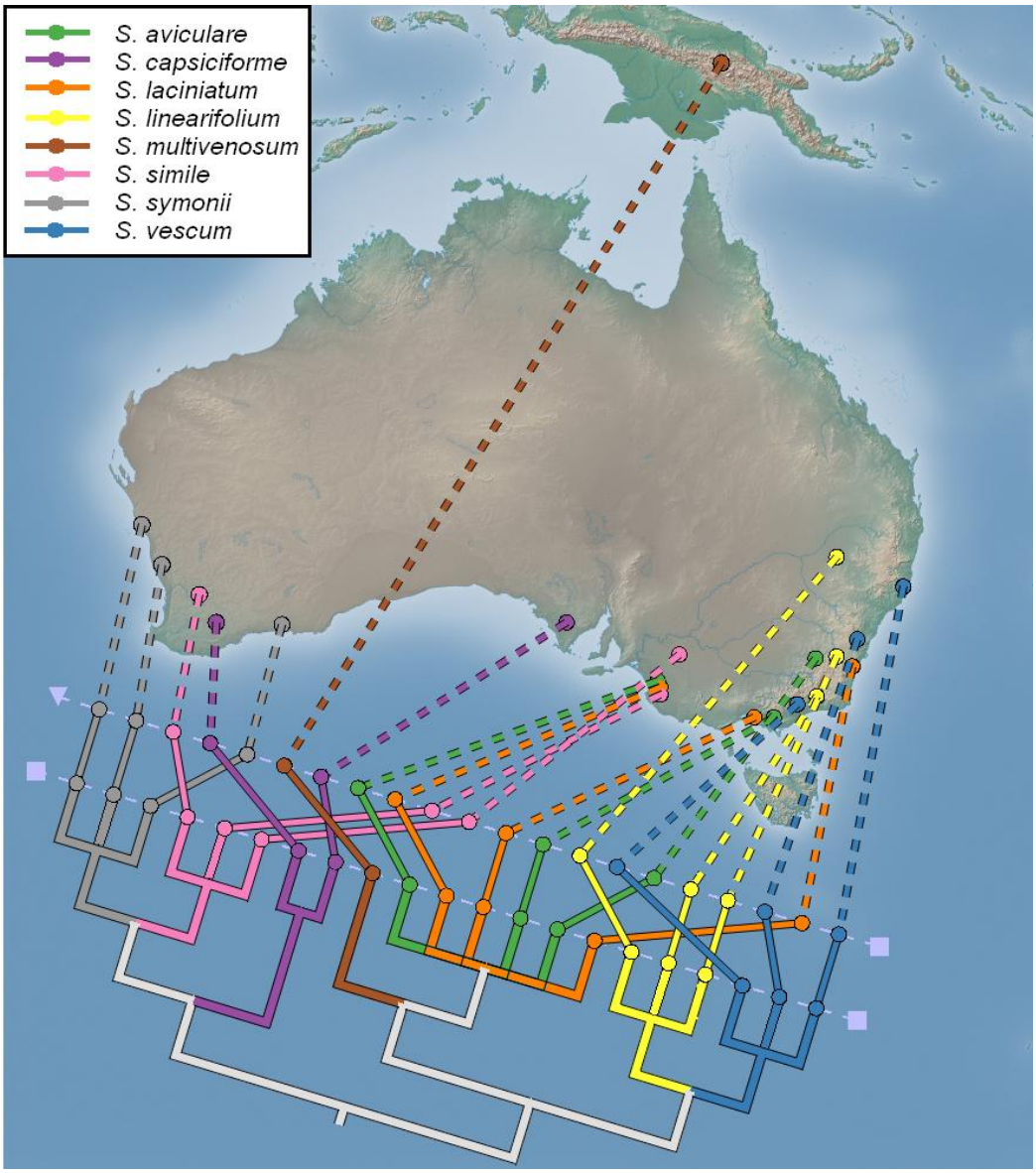

Fig.12. Geospatial representation of the phylogeny of the subgenus *Archaesolanum*. The optimal tree layout with 20 crossings is shown, with coloured branches and leaves, according to terminals taxonomic identity (Poczai et al. 2011a).

When geographical locations were mapped on the parsimony tree, a global optimum of 20 crossings was observed for the *Archaesolanum* clade. Defining a strict west-east (WE) axis, a permutation test on the cladogram yielded a *p*-value of 0.00095. Comparing this to the typical $\alpha = 0.05$ threshold level the null hypothesis can be rejected, indicating that geography is indeed an explaining factor for phylogeny of kangaroo apples. However, as yet, the program GenGIS does not allow broad testing of encoded hypotheses with automatic polyline enumeration. This is why we conducted permutation tests manually on a larger proportion of different linear axes. All of these defined in several directions, yielded *p*-values far greater than the 0.05 threshold level with observed crossings between 30 and 53, higher than the strict WE axis. The only significantly fitting orientation was a north-west/south-east (NW/SE) axis (24 crossings, $p = 0.001$). As the key principle in the tree construction through the defined linear geographical axis results in a dimensional gradient of similarity, the extent of data fit is expressed by the goodness of fit where mismatches are leading to crossings between the linking lines (Parks et al. 2009). The tree with the smallest number of crossings represents the best fitting hypothesis between geography and phylogeny. Accordingly, the strict WE axis (20 crossings) applies better to our dataset than the NW/SE axis (24 crossings). A closer inspection of subclades within the selected WE polyline was conducted to understand the relative contribution of geography and habitat. The permutation test inferred from 1,000 iterations was not significant in the case of the *Similia* subclade (6 crossings, $p = 0.535$) but yielded significant results with the *Avicularia/Laciniata* subclade (10 crossings, $p = 0.047$). Within this major clade permutation tests were also made for the two smaller subclades, one composed of *S. multivenosum*, *S. laciniatum* and *S. aviculare* (2 crossings, $p = 0.845$) and the other formed by *S. vescum* and *S. linearifolium* (3 crossings, $p = 0.668$), but neither of them had statistical significance. The SEEVA analysis was (Fig. 13) mostly congruent with the results GenGIS. However, impact indices were not so robust. In general the correlation between the analyzed environmental variables can be regarded as representing trends and tendencies within phylogenetic lineages, and not as the definite cause of lineage divergence.

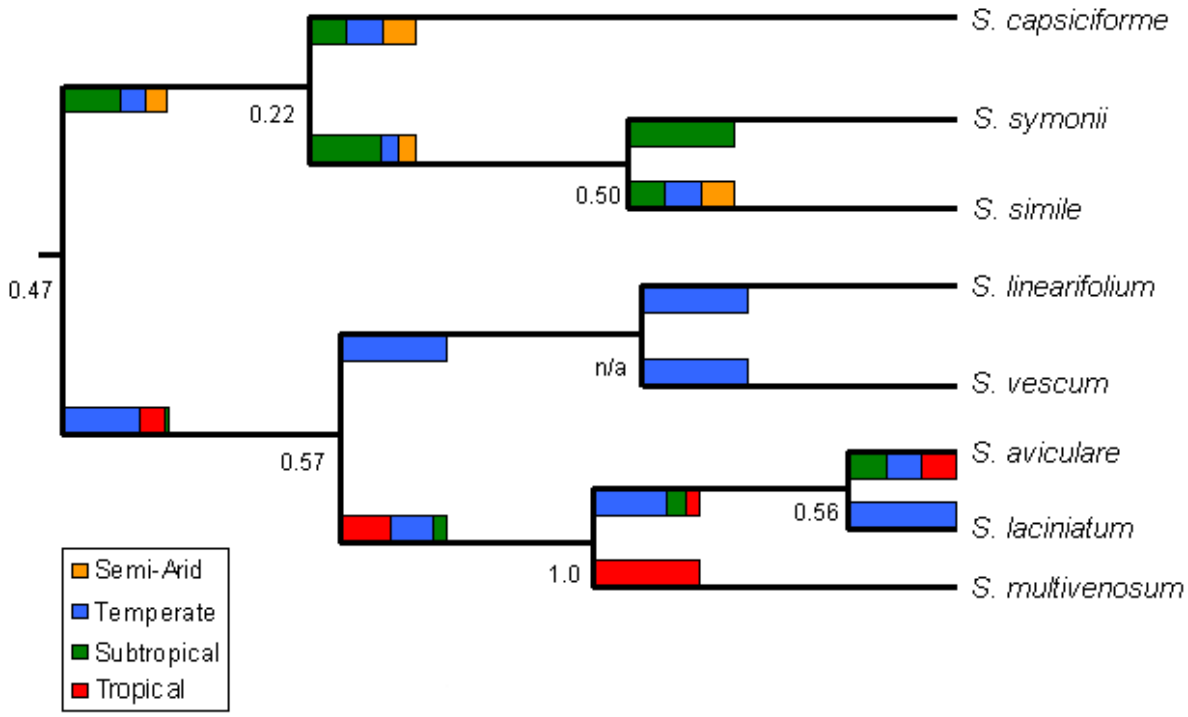

Fig.13. Differences in distribution of climatic variables through the phylogeny of the subgenus *Archaesolanum*. The numbers are impact indices for each node (see the text for further details); the bars indicate percentage of categories present in each group (Poczai et al. 2011a).

3.2.3. Historical biogeography of kangaroo apples

The DIVA and WAAA results were largely congruent. The inferred ancestral areas are shown in Fig. 14 (summarized in Table 8). The DIVA reconstruction required at least 14 dispersal events to explain the present-day distribution of the group when the maximum number of areas was restricted to four at each node. There were four equally optimal reconstructions. Both DIVA and WAAA indicate that the most likely ancestral area was in Australia and New Guinea, and indicates a dispersal event to New Zealand. Three vicariance events are indicated in *Archaesolanum*, one separating the *Similia* and *Avicularia-Laciniata* clades, another connected to *S. multivenosum* (resulting in further speciation), and one separating *S. vescum* + *S. linearifolium*.

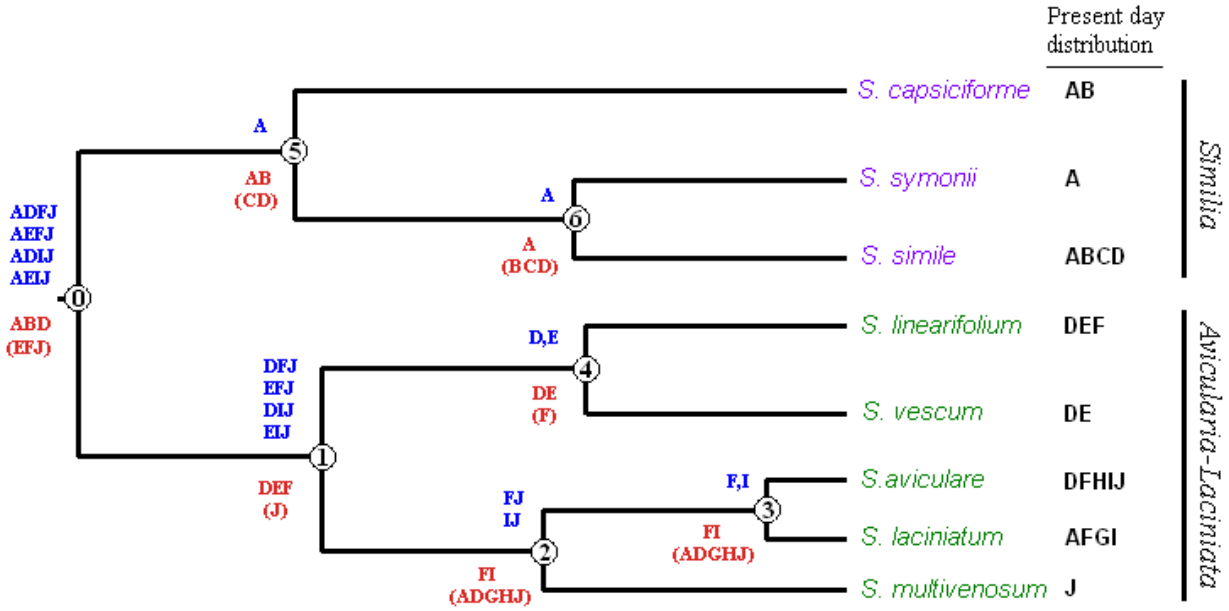

Fig.14. Mapping ancestral distribution of the subgenus *Archaesolanum*. Blue letters (above lines) represent all alternative distributions at each node according to DIVA, while red letters (bellow lines) show WAAA results. Brackets indicate areas with lower probability indices. Letters right to each terminal indicate present day distribution of taxa. Areas abbreviated are: Coastal Southern West and South Australia (A), Inland Southern West and South Australia (B), South New South Wales (C), East New South Wales (D), East Victoria (E), Tasmania (F), South Victoria (G), Queensland (H), New Zealand (I), and Papua New Guinea (J) (Poczai et al. 2011a)

Table 8. Mean posterior estimates of divergence times (MY) in the Archaesolanum clade using BEAST, with the results of the ancestral area reconstructions using event based dispersal-vicariance analysis (DIVA) and weighted ancestral area analysis (WAAA).

Node	Age estimates	DIVA	WAAA										Bayesian
	Mean (MY)		A	B	C	D	E	F	G	H	I	J	PP
0	8.85	ADFJ AEFJ ADIJ AEIJ	1.15	0.66	0.18	0.58	0.33	0.39	0.12	0.12	0.20	0.33	1.0
1	6.1	DFJ EFJ DIJ EIJ	0.17	-	-	1.51	1	1	0.17	0.17	0.33	066	0.99
2	3.5	FJ IJ	0.33	-	-	0.33	-	1	0.33	0.33	1	0.33	0.90
3	0.9	F, I	1	-	-	1	-	∞	1	1	∞	1	1.0
4	3.0	D, E	-	-	-	∞	∞	1	-	-	-	-	0.99
5	6.4	A	∞	3	0.33	0.33	-	-	-	-	-	-	0.90
6	3.2	A	∞	1	1	1	-	-	-	-	-	-	0.99

Node numbers refer to Fig.14. Posterior mean ages are shown in millions of years before present. DIVA reconstructions are shown for optimality criteria maxareas = 4. WAAA reconstructions are shown with probability indices (PI). Biogeographical areas are indicated by capital letters of Fig.14. Area codes are discussed in the Materials and Methods section (Poczai et al. 2011a).

CHAPTER 4

Discussion

4.1. Preliminary phylogeny of kangaroo apples: pilot studies

4.1.1. Relationships within subg. *Archaesolanum*

This study represents an approach using genomic DNA fingerprint markers to study genetic relationships in subg. *Archaesolanum*. The AAD data obtained in this study are a random sample, representing all the polymorphic RAPDs, SCoTs and ITs in the germplasm examined. The high bootstrap values indicate that a different sample of AAD loci would be unlikely to give a different result. Thus, the high bootstrap probability observed in the study gives strong statistical support for the conclusions. AADs have been shown to be useful as taxonomic markers for closely related species through concordant results using other molecular marker systems for closely related taxa elsewhere, for example in sect. *Petota* (Cisneros and Quiros 1995; Spooner et al. 1996). In addition, the results indicate that *Solanum aviculare* and *S. laciniatum* are closely related, which is strongly supported by bootstrap values. *S. aviculare* is easily confused with *S. laciniatum*. Previously, *S. laciniatum* was treated as a variety of *S. aviculare* under the name *S. aviculare* var. *laciniatum* (Aiton) Domin. Both species were commonly cultivated in the former Soviet Union, Australia and New Zealand and they have been used in the alkaloid industry. Although they are difficult to distinguish, there are some morphological parameters in which they differ. *S. aviculare* is diploid (2n = 2x = 46) with bright orange or red mature fruits containing approx. 600 seeds per fruit, while *S. laciniatum* is tetraploid (2n = 4x = 92), with fruits that are first green, later turning to yellow or orange-yellow and containing approx. 200 seeds per fruit. Baylis (1954) describes some quantitative characters to distinguish these two: *S. laciniatum* has larger pollen grains, flowers, seeds and stone-cell masses (in the fruit pulp) than *S. aviculare*; its corolla is deeper in colour with relatively shallow lobes, the margins of which flatten more completely, producing an emarginated apex.

The analysis carried out provide results in re-examining the putative hybrid origin of *S. laciniatum* indicating that it is not of recent hybrid origin, because there is a lack of additivity in the AAD patterns. In the trees constructed from AAD data, *S. laciniatum* formed a group together with *S. aviculare* and *S. vescum,* and the same topology was present in the tree from PCR-RFLP patterns. This topology is in agreement with the crosses summarized by Symon (1994), who suggested that *S. laciniatum* may be a hybrid of *S. aviculare* and *S. vescum*. *S. aviculare* and *S. vescum* hybridize spontaneously when the species are grown next to each other. The hybrids are only fertile when *S. vescum* is involved as female parent (Baylis 1963).

The presence of a specific fragment obtained from the amplification of the mtDNA region *nad*5a also confirms these results. This fragment is present in *S. vescum,* the hypothetical female parent, but it is absent from *S. aviculare* the putative male parent of *S. laciniatum*. According to the maternal inheritance of organelle genomes, this fragment in *S. vescum* and in *S. laciniatum* represents a unique mitochondrial structure, which supports the hybridization theory. Although this type of mitochondrial structure is present also in *S. linearifolium,* it separates from them forming another fragment of approximately 880 bp. This fragment could be a promising item for the further analysis of the subgenus.

Preliminary phylogeny constructed from the crosses summarized by Symon (1994) and based on morphology suggests *S. multivenosum* as a possible parent of *S. laciniatum*. However, the role of *S. multivenosum* in this phylogenetic concept still remains uncertain, because the pilot study did not include samples from this species. It was included in further investigations on the phylogenetic relationships in subg. *Archaesolanum* based on chloroplast DNA sequences. Only herbarium material was available from *S. multivenosum,* which hampered the amplifications of unambiguous AAD fragments. However, some RAPD and IT fragments amplified in test reactions, but it was ambiguous that their presence or even their absence is due to DNA lesions and degradation caused by long term storage or perhaps the amplified fragments are clearly missing from the taxon in question. Herbarium collections are valuable sources of genetic information available for phylogenetic studies, but DNA obtained from these specimens is often highly fragmented and present in small quantities (Lambertini at el. 2008).

High purity genomic DNA is required for AAD markers to avoid altered patterns in herbarium samples, which may be due to degradation rather than genetic polymorphism (Vos

et al. 1995; Blears et al. 1998). Due to these difficulties herbarium samples were excluded from our pilot studies with AAD markers and used only in DNA sequence analysis.

Another problem with the samples from *S. multivenosum* is that this species is endemic in high altitude (>2,000 m) sites in Papua New Guinea so access to good material is very difficult. Possibly for this reason, no plant material or herbarium specimen was recorded in the Solanaceae Source, a global project for taxonomy, nor by the Botanical and Experimental Garden of Radboud University Nijmegen.

Solanum linearifolium formed a basal branch of the clade composed of *S. aviculare*, *S. laciniatum* and *S. vescum* in all trees from obtained in the pilot studies receiving less bootstrap support (Figs. 5 and 8). According to morphological parameters its closest relative is *S. vescum*. The well-developed sinus tissues are diagnostic in the case of *S. linearifolium* and the strongly winged stems (from the sessile, decurrent leaves) in the case of *S. vescum* (Baylis 1954).

The other separated clade is composed of *S. capsiciforme*, *S. symonii* and *S. simile* (two accessions). These species belong to ser. *Similia*. The separation of this group based on morphology was confirmed by data derived from the present investigation using AAD markers. *S. symonii* is often confused with *S. simile* as the two species have same habit, green fruits and small flowers, and they can be very difficult to distinguish from each other. This morphological similarity can be detected at DNA level according to the AAD data. In our pilot studies the two species form a monophyletic group with high bootstrap values of 100%. Despite the morphological similarity the chromosome numbers of the two species are not the same; *S. symonii* is tetraploid (2n = 4x = 92) and *S. simile* is diploid (2n = 2x = 46). Crosses between *S. simile* and *S. symonii* and the other members of the subgenus have been made, but no fertile hybrids could be obtained (Baylis 1963).

S. capsiciforme is sister to all the other members of this group in Symon's (1994) preliminary phylogeny based on morphology. In the present study it occupied a distinctive place in the cluster composed of members of ser. *Similia*. In the analysis of the restriction patterns of the two chloroplast regions *trn*S-*trn*G it formed a group with *S. symonii*.

However, the AAD data separated *S. capsiciforme* from the members of the ser. *Similia*, and *S. symonii*. It is grouped together with the two accessions of *S. simile*. This difference

between the results can be explained by the different nature of the two methods. AADs amplify fragments from the whole genome, but mostly from the nuclear genome, while the restriction analysis in this study focused on specific regions of the chloroplast genome, detecting site variations. Based on this *S. symonii* is more closely related to *S. simile* at the nuclear genomic level than to *S. capsiciforme*. However, there is evidence from the *trn*S-*trn*G restriction site variation that both *S. symonii* and *S. capsiciforme* lack a cleavage site of the *Rsa*I endonuclease enzyme. This indicates that *S. capsiciforme* and *S. symonii* could share a common maternal ancestor, but additional data will be required to prove this hypothesis.

Series *Similia* can be easily distinguished morphologically from the other series of subg. *Archaesolanum.* Such a clear distinction cannot be made between the members of ser. *Avicularia* and *Laciniata*, which form a single clade with 100% bootstrap values at the base of the tree in the RAPD data, while the same topology could be observed in the tree obtained from the chloroplast data. *S. aviculare*, the type species for ser. *Avicularia* and *S. laciniatum*, the type species for ser. *Laciniata*, formed a separate subclade together. From this topology it is concluded that the existence of these two taxonomic groups must be reconsidered. New formal taxonomic designations for the series in subg. *Archaesolanum* will be needed. Raising the series to sectional level should be considered, since the *Archaesolanum* group is recognized as a subgenus. Both the AAD and cpDNA region analysis separated the two groups, and it is suggested that they form two sections in subg. *Archaesolanum*. These might be sect. *Similia* consisting of former members of ser. *Similia*. The *Avicularia/Laciniata* clade consists of both members from ser. *Avicularia* and ser. *Laciniata*, a new section should be formed uniting these two groups, which could be sect. *Avicularia*, since, *S. aviculare* was the first name published by Forster (1786a).

4.1.2. The origin of *Solanum laciniatum*: recent autopolypoidy or ancient hybridization?

Hybrid speciation could occur in at least two ways in this case: it is possible, that *S. laciniatum* was developed through diploid hybrid speciation involving *S. aviculare* and *S. vescum* as female parent and the entire genome is duplicated through autopolyploidy. This alternative hybridization can only be achieved if *S. laciniatum* is an ancient hybrid. The unique mitochondrial rearrangement supports this hypothesis. The diploid hybrid speciation could result from a normal sexual event where each gamete has a haploid complement of the nuclear chromosomes from its parent, but gametes that form the zygote come from different species (in this case from *S. vescum* and *S. aviculare*). From the crosses reported by Baylis (1963) it is known that partial fertility exists between *S. vescum* as female parent and of *S. aviculare*, and backcross is often possible like in the typical case of diploid hybrid speciation. However, this hypothesis could not explain the extensive genomic homology detected between *S. aviculare* and *S. laciniatum*. In the data set generated by the selected polymorphic AAD primers the number of polymorphic bands was very low which was also reported in our previous analysis with a different set of RAPD primers (Poczai et al. 2008).

Another alternative hypothesis might be that *S. laciniatum* was developed through autopolyploid, where the normal genome of *S. aviculare* is duplicated entirely and produced tetraploid off-spring which were postzygotically isolated from their parent. Crosses made between *S. aviculare* and *S. laciniatum* produced very few seeds and no successful germination could be detected (Gerasimenko 1969). This hypothesis could be an explanation to the lack of additive bands in *S. laciniatum* from *S. vescum* and *S. aviculare* in the AAD and profiles, and it also could be a reason to the presence of the high genetic similarity between *S. aviculare* and *S. laciniatum*. However, the latter hypothesis suggesting an autopolyploid speciation event for *S. lacianiatum* seems considerably more reasonable, than a hybrid introgression through *S. vescum*.

To clarify the relationship of *S. laciniatum* with the other members of the group will require a much intensive phylogenetic analysis. If *S. laciniatum* resulted from an ancient hybridization, additional evidence on the hybrid origin could be provided with particular analysis of nuclear and chloroplast DNA sequences.

To resolve these events reticulate evolution should be taken in consideration for the further investigations and a cytogenetic analysis should also be carried out to resolve the

origin of *S. laciniatum*. In addition, an improved taxonomic sampling and/or the use of more sensitive markers would be required for a more comprehensive understanding of the evolutionary history of *S. laciniatum*/*S. aviculare.* Such study based on alternative rDNA ITS sequence typing is currently performed and will be summarized in a companion paper.

4.1.3. The utility of intron targeting (IT) markers in plant systematics

The results indicated that 29 loci showed clear polymorphic patterns (Appendix 3) while others could not amplified. The number of alleles ranged from two to seven in the analyzed populations. The observed heterozygosity ranged from 0 to 0.833 and the expected heterozygosity ranged from 0 to 0.750 (Appendix 3). The results from cross-species amplification in *Solanum nigrum* populations revealed that all 29 primers generated specific and polymorphic products. These results demonstrate the potential of these newly developed IT markers for further genetic analysis in potato and related *Solanum* species. The sequence-based IT markers developed here may have wider application within different populations of potato and related species in the genus *Solanum.* These primers may potentially useful in applied research fields such as solanaceous crop breeding programs, as further primers could easily be designed based on the methods described here. Due to their polymorphism and close association with functional genes, these markers will also be useful in studies of genetic diversity in natural populations and germplasm collections, as well as for genetic mapping of potato and related species. In conclusion, the IT method is a simple and efficient marker system that can be adapted for different purposes in plant molecular genetics. It is simple to use due to being agarose-based and produces co-dominant markers for potato research and breeding as well as for genetic diversity analysis.

4.2. Phylogeny of kangaroo apples based on chloroplast sequences

4.2.1. Monophyly and relationships within subgenus Archaesolanum

The results of the BEAST analysis were congruent with the parsimony analysis and supported a robust topology for a monophyletic *Archaesolanum* clade. At this point we propose to maintain *Archaesolanum* at the subgeneric level following Bitter's (1927) original conspectus in concordance with D'Arcy (1972,1991) and Hunziker (2001), until higher level relationships of the whole genus are resolved and the closest relatives of the group are determined.

Analysis of the chloroplast sequence data generated here separates two well-supported subclades within kangaroo apples. The *Similia* clade with *S. capsiciforme* + *S. symonii* + *S. simile*, and the *Avicularia-Laciniata* clade, further divided into separate *S. vescum* + *S. linearifolium* and *S. laciniatum* + *S. aviculare* + *S. multivenosum* subclades. Based on morphological characters Symon (1994) also recognized a separate *Avicularia-Laciniata* clade, but *S. aviculare* was a sister species of *S. multivenosum*. In Symon (1994) a separate *Similia* clade was not accepted because *S. capsiciforme* did not have a consistent position in the most parsimonious trees, either occupying a basal position relative to all other members, or else forming a clade with *S. simile* and *S. symonii*. Poczai et al. (2008) separated *S. simile* from *S. aviculare* and *S. laciniatum* and found evidence to suggest the existence of two separate subclades within *Archaesolanum*. The topology obtained in the present analysis is congruent with those obtained in a pilot study based on multi-locus genetic markers (Poczai et al. 2011a). However, the position of *S. multivenosum* was not determined because only herbarium material of this taxon was available. The position of *S. multivenosum* within *Archaesolanum* has been ambiguous since its description (Symon 1985). According to our results an early split between *S. aviculare* and *S. laciniatum* dates back to 3.5 Mya, while a more ancient (6.1 Mya) split separated *S. vescum* + *S. linearifolium*, the other members of the *Avicularia-Laciniata* clade (Fig. 15). The relationships discussed above were proposed by Symon (1994) who examined the seed surface with scanning electron microscopy (SEM). This analysis led him to postulate a group of *S. capsiciforme* + *S. symonii* + *S. simile* and a separate *S. aviculare* + *S. multivenosum* + *S. laciniatum* group with *S. vescum* + *S. linearifolium* in an intermediate position.

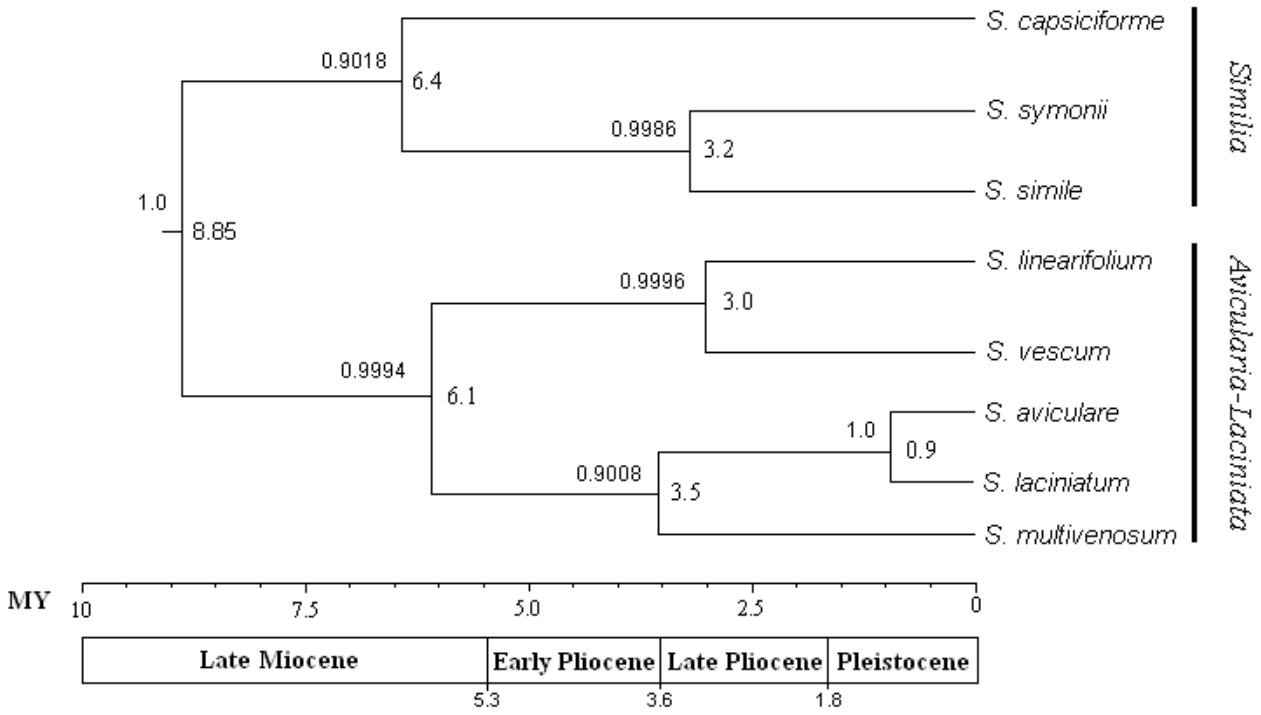

Fig.15. Bayesian chronogram of the subgenus *Archaesolanum.* (Poczai et al. 2011a).

The morphological characters together with the robust topology presented here (Figs. 10 and 15) do not agree with the scheme of series proposed by Gerasimenko (1970). However, the recognition of a separate group of ser. *Similia* (*S. simile*, *S. symonii*, *S. capsiciforme*) seems to be valid while the separation of ser. *Avicularia* (*S. aviculare*, *S. cheesmanii*[2], *S. baylisii*[2], *S. brisbanense*[2]) from ser. *Laciniata* (*S. laciniatum*, *S. vescum*, *S. linearifolium*) seems unwarranted. Based on the phylogenetic hypothesis we present, ser. *Laciniata* would be paraphyletic. According to our results a new taxonomic scheme should be devised for further subdivision of *Archaesolanum*.

Questions concerning the natural variation within this group and the status of previously published species and intraspecific taxa have not been intensively studied and remain open even after our analyses.

The controversy is due to the studies based on the collections of Kondratenko and Kibaltchich in 1968. Many intraspecific names have been published based on material

[2] *S. brisbanense* Geras. is now recognized as synonym of *S. aviculare*; while *S. cheesmanii* Geras. refers to the white flowered (var. *albiflorum*) variety and *S. baylisii* Geras. to the broad leaf (var. *latifolium*) mutant of *S. aviculare*.

possibly originating from these collections (Gerasimenko 1969, 1971; Korneva et al. 1969; Korneva and Balakhova 1973). However, they were never validly published as the descriptions lack Latin diagnoses and type specimen citations (Knapp 2006), thus not fulfilling the criteria of the International Code of Botanical Nomenclature (McNeill et al. 2006). In addition, these varieties would be better treated as horticultural forms and most of them (e.g. *S. baylisii, S. brisbanense*) have been synonymised by the global Planetary Biodiversity Inventory: *Solanum* Project (Knapp et al. 2004; http://www.nhm.ac.uk/research-curation/research/projects/solanaceaesource/). On the other hand, these publications undoubtedly indicate that there must be considerable, and as yet unexplored, variation within some species of the group, which should be investigated further and documented properly.

4.2.2. Implications on higher level relationships: an unresolved case

The unknown origin and lack of plausible common ancestors of *Archaesolanum* and other species of *Solanum* have been previously recognized (Hawkes 1990; Bohs 2005). However, some attempts have already been made to find even distant relatives of the group utilizing molecular markers. The first studies based on chloroplast restriction site variation and *ndh*F sequences, sampling only *S. aviculare* and *S. laciniatum*, indicated monophyly and a basal position for subg. *Archaesolanum* (Bohs and Olmstead 1997), and relationships to other clades composed of members of sect. *Solanum*, *Dulcamara* and *Jasminosolanum* (Olmstead and Palmer 1997). Bohs (2005), using a broader taxonomic sampling for the *ndh*F gene, also supported the basal position of this group, with the closest relatives remaining unsettled. The latest analysis based on three genes (*ndh*F, *trn*T-F, and *waxy*) representing the broadest sampling of *Solanum* to date, indicated a surprisingly close relationship of *Archaesolanum*, *Normania* and the so called African non-spiny clade (Weese and Bohs 2007). However, this analysis included only two species of kangaroo apples. The present analysis including all recognized species of the group confirms the close relationship of the clades *Archaesolanum* and *Normania* despite the great morphological differences between the two groups.

However, the topology could still be misleading due to very limited sampling of both *Normania* and the African Non-Spiny clades. The *Normania* clade comprises only three

species and thus the error, if due to small sampling, is most likely due to the inadequate sampling of the African non-spiny group.

Finding the higher-level relatives of kangaroo apples could be a problematic task. If the present conclusions and those by Weese and Bohs (2007) are correct the closest relatives are to be found in *Normania*, or in the African non-spiny group. Obtaining good material for the Macaronesian *Normania* group is relatively difficult, because of the extreme rarity of these plants. For example only two living plants have been found from *Solanum nava* Webb & Berth. since its original description (Francisco-Ortega 1993; Bohs and Olmstead 2001). On the other hand, the African non-spiny clade is also poorly characterized both morphologically and at the molecular level. It is obvious that in order to find the closest relatives of kangaroo apples a more thorough sampling is needed of both of the groups given above. Symon (1994) suggested that the closest relatives of *Archaesolanum* are to be found within subg. *Solanum*. This is consistent with the fact that the African non-spiny clade is composed of members from sect. *Afrosolanum*, *Quadrangulare* and *Benderianum*, all belonging to subg. *Solanum*. In summary, our current sampling is not sufficient to settle the controversy concerning the closest relatives of *Archaesolanum*.

4.2.3. Reaching Australian Shores: Vicariance or Long-Distance Dispersal?

Vicariance hypotheses have been challenged by previous botanical and molecular studies which aimed to study the historical biogeography of Solanaceae (D'Arcy 1991; Olmstead et al. 2008). The most widely accepted hypothesis by Olmstead and Palmer (1997) on the biogeography of *Solanum* species assumes that *Archaesolanum* presents an ambiguous case, either representing an early dispersal event in the genus, or a plausible case of vicariance dating to a time preceding the separation of South America and Australia. Our divergence time estimates suggest that the most recent common ancestor of kangaroo apples is from the late Miocene (~ 9 Mya).

The observed split of the subgenus is unlikely to be the result of the break-up of Gondwanaland, since the division began already in the Jurassic, 180-150 Mya (Scotese et al. 1988). The separation of Australia from Antarctica started in the late Cretaceous (90 Mya) and was completed in the late Eocene (35 Mya) with the opening of the South Tasman Sea (Pigram and Davies 1987). However, some authors suggest a much earlier (~50 Mya)

separation (Woodburne and Case 1996). This geological break-up of the southern continent does not fit with our estimate dating back only to the late Miocene. Because, South America and Antarctica separated 30-28 Mya, while the distinct Australian continent collided with the Asian Plate (10 Mya; see Sanmartín and Ronquist 2004). These geological events together with our divergence estimates indicate that kangaroo apples probably do not represent a plausible case of vicariance dating to a time preceding the separation of South America and Australia as previously proposed (Olmstead and Palmer 1997). These are in accordance with the hypothesis of D'Arcy (1991) suggesting that the continental break-up was too early to have carried *Solanum* precursors. Vicariance requires that speciation and the corresponding fragmentation occur simultaneously, and thus information concerning absolute timing of speciation events is crucial for evaluating such scenarios (Queiroz 2005). However, molecular clock dating has been criticized as being inaccurate for a variety of reasons (see Pulqério and Nichols 2007). One of the major faults could be the inclusion of fossil records with ambiguous taxonomic affinities as clearly presented by Crepet et al. (2004).

However, assuming that the presence of *Archaesolanum* in Australia is due to vicariance would challenge not only our molecular clock analyses but also those presented for larger sampling of angiosperms (Wikström et al. 2001) or those concentrated especially on Solanaceae (Paape et al. 2008). Therefore, long-distance dispersal (LDD) with subsequent speciation is the most likely explanation for the occurrence of kangaroo apples in the SW Pacific. This might have taken place through trans-oceanic transmission by migrating birds, because only a few intercontinental plant disjunctions can be explained by water- or wind-mediated transports (Carlquist 1967; Mummenhoff and Franzke 2007), while other studies suggest that long-distance plant dispersal by birds is by far the most important vector (Winkworth et al. 2002). However, not many birds are capable of retaining seeds for such a long time, although some studies report that viable seeds have been recovered from the gizzards of migratory birds after 200 to 360 hours; for example, viable seeds of *Rhus glabra* L. have been recovered after they were in the digestive tract of a killdeer (*Charadrius vociferous* L.) for more than 14 days (Proctor 1968).

The soft and sticky, fallen ripe fruits of kangaroo apples are known to be eaten by birds (Keighery 1984; Symon 1994), which might support a long-distance dispersal scenario. Although, there is very little information available on exactly which bird species are responsible for their distribution. The first record of fruit predation by birds was made by

Forster (1786a), which inspired him to give the telling name "*Aviculare*" (bird dispersed) for the first described species of the group. Later it was confirmed by other authors that birds play an important role in the distribution of seeds among the group (Keighery 1984; Symon 1994) especially in the case of species having red or orange colored fruits (e.g. *S. aviculare*, *S. laciniatum*) which are attractive to frugivorous birds. Previous aviary experiments with different plant species have demonstrated that migrant birds often exhibit color preferences, when factors such as taste, nutrition, and accessibility to food sources are equal (Willson et al. 1990). Fruit colors as red, orange and yellow are commonly considered to increase the conspicuousness of a ripe fleshy crop and attract birds that eat fruits and disperse the enclosed seeds (Janson 1983; Willson et al. 1989). However there are some records which indicate that mammals (e.g. rabbits, dingos) and macropods (e.g. wallabies, quokkas) may also be important for local dispersal (Bell et al. 1987).

Regarding the long-distance dispersal hypothesis, specific fruit traits might have played a role in the arrival of the group to Australia. The fruits of all kangaroo apples contain a relatively conspicuous and abundant proportion of stone cell aggregates. However, stone cell granules are commonly also present in some other groups of *Solanum* (e.g. sect. *Solanum* or sect. *Pachyphylla*; see Bitter 1911). The quantity and quality of these may vary from species to species as affected by environmental conditions. The anatomy and distribution of stone cells have been repeatedly studied (e.g. Bitter 1911; Danert 1970), but no final interpretation regarding their function has been presented. According to our divergence time estimates the character of abundant stone cell formation seems to be an ancient trait in *Archaesolanum*. The abundant stone cell mass produced by kangaroo apples might have had an important role in bird dispersal by protecting the seeds in the gizzard, or perhaps even by helping seeds to adhere to legs or plumage of birds.

Intercontinental dispersal from the African continent to Australia seems unlikely at first sight, considering that these areas are separated by more than 8,000 km of open ocean. Knowing that the shores of different continents have always been linked since the evolution of migratory birds (Carlquist 1983); makes this assumption more likely. These birds regularly travel across landmasses and they might carry internally or externally attached seeds from Africa/South America to Australia or even to New Zealand (Winkworth et al. 2002; Mummenhoff and Franzke 2007). A similar case of intercontinental dispersal of mucilaginous *Lepidium* L. seeds was demonstrated by Carlquist (1983). Mummenhoff et al. (2004) using

chloroplast *trn*T-F and rDNA ITS sequences concluded that these Australian *Lepidium* s.s. (Brassicaceae) species were developed through allopolyploidization following not only one but two separate trans-oceanic dispersals from California and Africa to Australia and gave birth to the development of 26 modern species during the Quaternary. In another study based on also *trn*L-F chloroplast sequences and AFLP analysis the origin of a disjunct Australian *Microseris* D.Don (Asteraceae) lineage was revealed (Vijverberg et al. 1999; 2000). It was concluded that Australian *M. lanceolata* (Walp.) Sch.Bip. is the outcome of a single introduction of a new hybrid formed in North America.

Other potential candidates as trans-oceanic long-distant dispersal events from Africa and South America to Australia are also known from the Solanaceae. The ancestors of *Lycium australe* F.Muell. are assumed to have arrived from Africa (Fukuda et al. 2001; Miller 2002). In this genus bird dispersal probably plays an important role, because *Lycium* L. species have relatively small succulent, orange, red berries which are very attractive to birds. Because of this feature at least three long-distant dispersal events are assumed to have occurred in the genus, resulting in the current range from North America to the Pacific Islands (e.g. Hawaii) as well as to South Africa, and more interestingly from South Africa to Australia, and possibly to the Eurasian region (Fukuda et al. 2001). The single African taxa *Nicotiana africana* Merxm. is member of sect. *Suaveolentes* an Australian group of genus *Nicotiana* (Marks 2010). A recent study by Chase et al. (2003) based on rDNA ITS sequences and genomic *in situ* hybridizations (GISH) concludes that the monophyletic species of *Nicotiana* were derived from an allotetraploid progenitor from South America which reached the Australian continent by a single long-distance dispersal event. Unfortunately, this hybrid species left no trace in the area of origin.

The mentioned examples outline that transoceanic long-distance dispersal events are often accompanied with hybridization (e.g. *Lepidium*) and/or allopolyploid formation (e.g. *Nicotiana*). This might open the possibility of multiple origins for kangaroo apples. It seems reasonable to assume that kangaroo apples were developed in a plausible hybridization/allopolyploidization event. The question is when and where did this happen: were they formed by multiple introductions from separate continents? Did the characteristic chromosome formation occur in the new Australian environment or in the place of origin? These questions can be answered by revealing the origin of the most recent common ancestor of the group, but until that these questions will remain open.

4.2.4. Australian east-west disjunction

Our results indicate that kangaroo apples most likely diversified in the Miocene, thus representing an Australian floral element that arrived after the isolation of the continent and then radiated. During the Late Eocene, as the continent moved north, the Australian climate dried out, with more pronounced seasonality. This resulted from the action of the circumpolar currents and the formation of ice caps in the Antarctic (Kemp 1978). The climate became cooler and drier, and kangaroo apples radiated to more humid habitable areas (SE and SW Australia; NW Queensland). The lineage through time plot (Fig. 16) suggests that the diversification of kangaroo apples proceeded steadily in the early stages of their evolution. Two new lineages then emerged in the early Pliocene, and they diversified again in the early/late Pliocene. This coincides with the continental aridification of Australia and later with the warming phase coupled with a brief resurgence of the rainforests which continued until ca. 3 Mya. The earlier drying phase has possibly given rise to the *Similia* and *Avicularia-Laciniata* clades ca. 6.5 Mya, while the later short rainforest resurgence led to speciation within these two lineages. Biogeographic analyses (DIVA, WAAA) together with molecular clock results indicate that the two major subclades arose and radiated during the Miocene in the period when the Australian environment underwent notable aridification that significantly affected other plant groups (e.g. tribes of Fabaceae) and the environmental changes caused expansion or contraction of their ranges (Crisp et al. 2004).

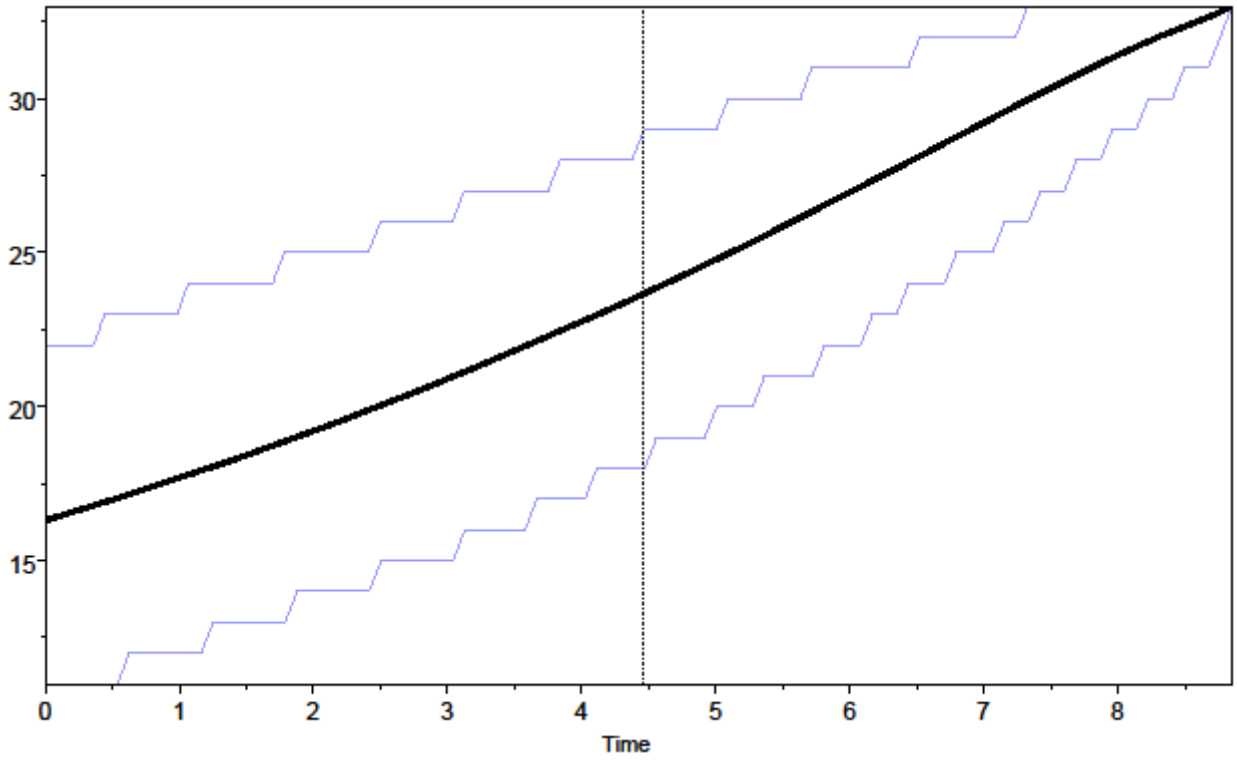

Fig.16. Lineages-through time plot (LTT) for the subgenus *Archaesolanum* calculated across age estimates obtained for the tMRCA of the group with BEAST. Blue graphs represent upper and lower 95% HPDs. Time scale ages are shown in millions of years (MY) before present. The plot shows a steady rate of diversification, without significant changes or shifts (Poczai et al. 2011a).

The two lineages within *Archaesolanum* were obviously separated due to the fragmentation of their once wider habitat by the growing interior of arid deserts in the Australian landscape. This assumption is clearly supported by all biogeographical analysis carried out here (Fig. 17). The *Similia* clade restricted to the SE regions, while the range of *Avicularia-Laciniata* clade to the SW region, as the climate became progressively drier during the Miocene. Taxa that required more humid climates were restricted to refugia in the highlands, or to small favourable habitats (*Aviculara-Laciniata* clade), while others adapted to more arid conditions (*Similia* clade). In the case of *Solanum symonii*, the distribution area became highly restricted to the SE-SW Australian coastlines suggesting a strong maritime shoreline influence. The same pattern is also encountered in the disjunctive present day distribution of the *Similia* clade in Western and Southern Australia, where a substantial gap is preserved within the populations of *S. simile* and *S. capsiciforme*, while this gap is much smaller in the case of *S. symonii* (see Fig. 3). No specimens belonging to the *Avicularia-*

Laciniata clade have been collected in Western Australia, or just some sporadic records are known.

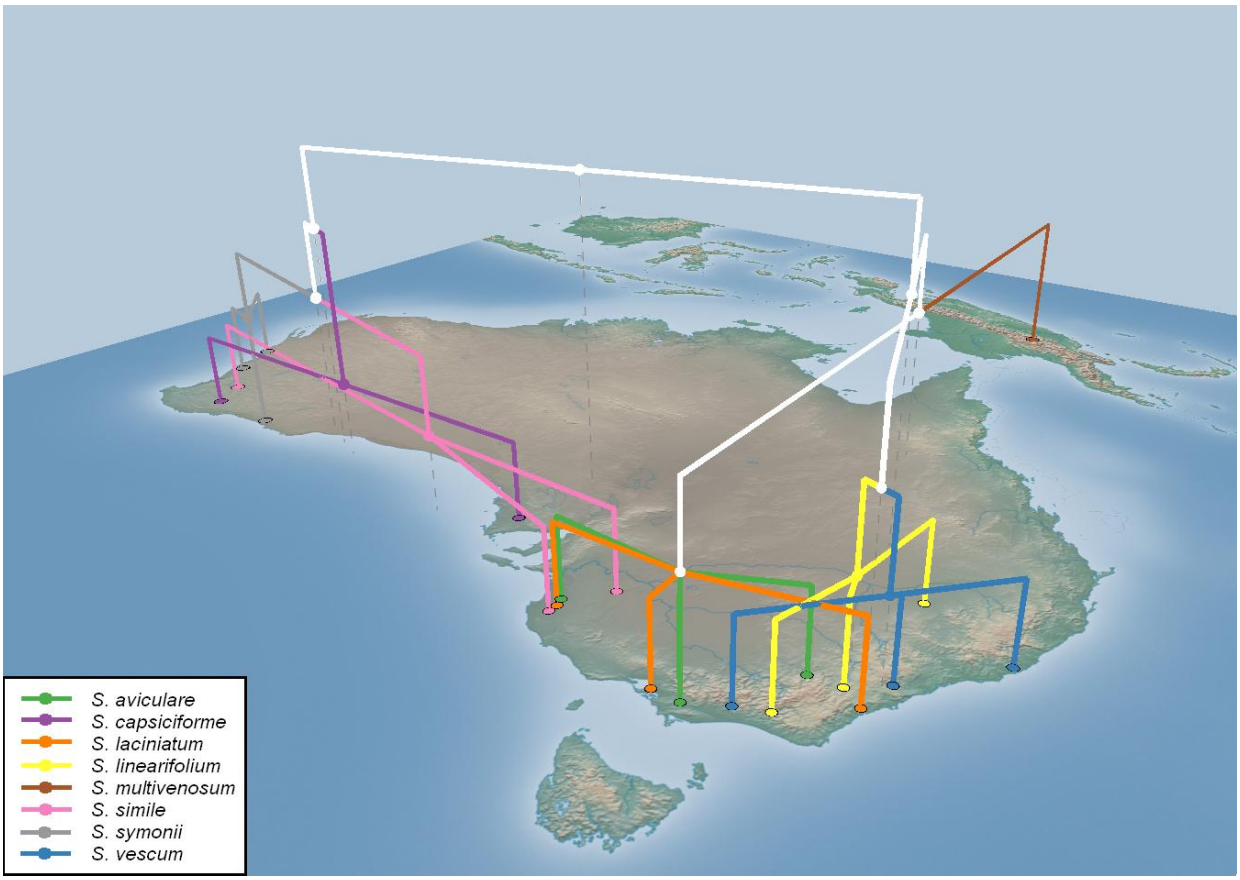

Fig.17. Geospatial representation of the phylogeny of the subgenus *Archaesolanum* in 3D. Phylogenetic tree of kangaroo apples is based on the BEAST and Maximum Parsimony phylogeny obtained from the combined *trn*T-*trn*F chloroplast region. Different species are indicated with unique colouring of branches (Poczai et al. 2011a).

4.2.5. Diversification in Papua New Guinea

The *trn*T-F data together with the DIVA and WAAA reconstructions supports the hypothesis of an early Australian and New Guinean origin of kangaroo apples followed by a habitat fragmentation and diversification in the eastern region reaching Tasmania and New Zealand. The uplands of New Guinea are very young, originating from a recent vertical uplift along the axis of the island (Pain and Ollier 1983). New Guinea was clearly part of the Australian region as exemplified by its floristic affinities. The extensive ever-wet uplands of the present day did not exist in its current form through most of the period of angiosperm history of the region (Barlow 1994). *Solanum multivenosum* presumably diversified through vicariance during the continuous uplift of New Guinea and evolved in isolation from other species of the group, adapting to tropical climatic conditions. This could be associated with the recent vertical uplift of the Central Range orogeny of New Guinea, which developed from several geological arcs during the Late Miocene to the Holocene (Polhemus and Polhemus 1998). This orogenesis formed a long mountainous backbone (ca. 1,300 km) with some peaks over 5,000 m (Beebe and Cooper 2002). These mountain ranges presumably promoted speciation by acting as physical barriers and creating a hot, wet climate associated with annual rainfall of over 2,500 mm. This process is analogous to the uplift of the Andes in South America – the centre of genetic diversity for many core groups of *Solanum* lineages - which also resulted in well known important speciations within the genus (see Spooner 2009). Therefore other undiscovered species and/or intra- and infraspecific taxa may exist in the high altitude ranges of New Guinea where the subgenus *Archaesolanum* is poorly known. *S. aviculare*, the species nowadays also known to be present in New Guinea, probably originates from a fairly recent colonization, possibly from the Northern Queensland refugia as indicated by DIVA and WAAA.

SUMMARY AND FUTURE DIRECTIONS

Kangaroo apples represent an early radiation in the evolutionary history of *Solanum*, possibly dating back to the late Miocene. This dispersal - one of the most ancient in the genus - led to the formation of a unique isolated group which split into two separate lineages during the continental aridification of Australia. Despite of the robust phylogenetic hypothesis presented here, many questions regarding the evolutionary history of kangaroo apples remain open. In future studies higher level relatives of this group should be determined by increasing sampling of the poorly characterized African non-spiny group to define its taxonomic limits, and to indentify the precise origin of the *Archaesolanum* clade. In general, the natural variation shown by the group should be explored in order to settle the taxonomy. It is also possible that further undescribed species exist, especially at the high altitudes in the New Guinean mountains where kangaroo apples are still poorly collected. Further field studies are also needed to reveal population genetic processes (e.g. population dynamics, gene flow), especially in New Zealand where the white flowered and ovoid leaf varieties of *S. aviculare* occur. This would give us a more precise view of speciation in this isolated group as well as genetic diversity within and among populations. Such a research program has already been initiated, and the results are in preparation (G.Weavers, personal comm.). The unique aneuploid chromosome structure - the most interesting feature of this group, indeed in the whole genus - should be explored. Some unusual chromosome numbers are also known in other lineages of *Solanum*, but kangaroo apples represent a biologically poorly known case which deserves more attention. Further molecular studies based on biparentally inherited nuclear genes are tools that could be used to investigate possible hybridization and reticulate evolutionary process of polyploidy within members of the group, as well as to test the obtained hypothesis of phylogenetic relationships.

REFERENCES

Acosta MC, Bernardello G, Guerra M, Moscone EA (2005) Karyotype analysis in several South American species of *Solanum* and *Lycianthes rantonnei* (Solanaceae). Taxon 54:713-723

Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep 27:617-631

Andersen JR, Lübberstedt T (2003) Functional markers in plants. Trend Plant Sci 8:554-560

Asmussen CB, Chase MW (2001) Coding and noncoding plastid DNA in palm systematics. Am J Bot 88:1103-1117

Awasthi AK, Nagaraja GM, Naik GV, Kanginakdudru S, Thangavelu K, Nagaraju J (2004) Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays**.** BMC Genetics 5:16

Backeljau T, De Bruyn L, De Wolf H, Jordaens K, Van Dongen S, Verhagen R, Winnepenninckx B (1995) Random amplified polymorphic DNA (RAPD) and parsimony methods. Cladistics 11:119-130

Bakker FT, Culham A, Gomez-Martinez R, Carvalho J, Campton J, Dawtrey R, Gibby M (2000) Patterns of nucleotide substitution in angiosperm cpDNA *trn*L (UAA)-*trn*F (GAA) regions. Mol Biol Evol 17:1146-1155

Barlow BA (1994) Phytogeography of the Australian region. In: Groves RH (Ed) Australian vegetation. Cambridge, UK, University Press pp. 3-36

Bayer RJ, Starr JR (1998) Tribial phylogeny of the Asteraceae based on two non-coding chloroplast sequences the *trn*L intron and *trn*L/*trn*F intergenic spacer. Ann Mo Bot Gard 85:242-256

Baylis GTS (1954) Chromosome number and distribution of *Solanum aviculare* Forst. and *S. laciniatum* Ait. Trans Proc Roy Soc New Zealand 82: 639-643

Baylis GTS (1963) A cytogenetical study of the *Solanum aviculare* species complex. Aust J Bot 11: 168-177

Beebe NW, Cooper RD (2002) Distribution and evolution of the *Anopheles punctulatus* group (Diptera: Culicidae) in Australia and Papua New Guinea**.** Int J Parasitol 32**:**563-574

Beiko RG, Keith JM, Harlow TJ, Ragan MA (2006) Searching for convergence in phylogenetic Markov chain Monte Carlo. Syst Biol 55:553–565

Bell DT, Moredoundt JC, Longeragan WA (1987) Grazing pressure by tamer (*Macropus eugenii* Desm.) on the vegetation of Garden Island, W Aust J Roy Soc 69:89-94

Benton MJ (1993) The fossil record 2. London, UK, Chapman and Hall.

Berg RGVD, Groenduk-Wilders N, Kardolus JP (1996) The wild ancestors of cultivated potato: the brevicaule complex. Act. Bot. Neerl. 45:157-171

Berg RGVD, Bryan G, del Rio A, Spooner DM (2002) Reduction of species in the wild potato *Solanum* section *Petota* series *Longipedicellata*: AFLP, RAPD and chloroplast SSR data. Theor Appl Genet 105:1109-1114

Bernáth J (1970) Az orvosi csucsor (*Solanum laciniatum* Ait.) bogyókötődését és a generatív szervek elrúgását kiváltó meteorológiai tényezők I. Herb Hung 9:35

Bernáth J (1971) Az orvosi csucsor (*Solanum laciniatum* Ait.) bogyókötődését és a generatív szervek elrúgását kiváltó meteorológiai tényezők II. Herb Hung 10:21

Bitter G (1911) Steinzellkonkretionen im Fruschtfleisch beerentragender Solanaceaen und deren systematische Bedeutung. Bot Jahr 45:483-507

Bitter G (1912) *Solana nova vel minus cognita*. Feddes Repert10: 529-565

Bitter G (1913) *Solana Africana* 1. Bot Jahr 49: 560-569

Bitter G (1917) *Solana Africana* 3. Bot Jahr 54: 416-506

Bitter G (1919) Die papuasischen Arten von Solanum. Bot Jahr 55: 69-89

Bitter G (1921) Solana Africana 3. Bot Jahr 57: 248-286

Bitter G (1922) *Solana nova vel minus cognita*. Feddes Repert 18: 301-307

Bitter G (1923) *Solana Africana* 4. Feddes Repert 16: 1-320

Bitter G (1927) *Archaesolanum* In: Hegi G (Ed) Illustrierte Flora von Mitter-Europa. Band V. Tiel. 4. München, JF Lehmann

Blears MJ, De Grandis SA, Lee H, Trevors JT (1998) Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. J Ind Microbiol Biotechnol 21:99-114

Bohs L (2004) A chloroplast DNA phylogeny of *Solanum* section *Lasiocarpa*. Syst Bot 29:177-187

Bohs L (2005) Major clades in *Solanum* based on *ndh*F sequences. In: Keating RC, Hollowell VC, Croat TB (Eds) A festschrift for William G. D'Arcy: the legacy of a taxonomist. Monographs in Systematic Botany from the Missouri Botanical Garden, Vol. 104. St. Louis, USA, Missouri Botanical Garden Press, pp. 27-49

Bohs L, Olmstead RG (1997) Phylogenetic realationships in *Solanum* (Solanaceae) based on *ndh*F sequences. Syst Bot 22:5-17

Bohs L, Olmsted RG (1999) *Solanum* phylogeny inferred from chloroplast DNA sequence data. In: Nee M, Symon DE, Lester RN, and Jessop JP (Eds) Solanaceae IV: Advances in Biology and Utilization. Royal Botanic Gardens, Kew pp. 97-110

Bohs L, Olmstead RG (2001) A reassessment of *Normania* and *Triguera* (Solanaceae). Plant Syst Evol 228:33-48

Bolaric S, Barth S, Melchinger AE, Posselt UK (2005) Genetic diversity in European perennial rygrass cultivars investigated with RAPD markers. Plant Breeding 124:161-166

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095-1103

Borsch T, Hilu KW, Quandt D, Wilde V, Neinhuis C, Barthlott (2003) Noncoding plastid *trn*T-*trn*F sequences reveal a well resolved phylogeny of basal angiosperms. J Evol Biol 16:558-576

Borsch T, Hilu KW, Wiersema JH, Löhne C, Barthlott W, Wilde V (2007) Phylogeny of *Nymphaea* (Nymphaeaceae): evidence from substitutions and microstructural changes in the chloroplast *trn*T-*trn*L region. Int J Plant Sci 168:639-671

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Brown LR (1994) State of the World, Norton, New York

Bruneau A, Dickson EE, Knapp S (1995) Configurance of chloroplast DNA restriction site characters with morphological and isozyme data in *Solanum* sect. *Lasiocarpa*. Can J Bot73: 1151-1167

Bukenya ZR, Carasco JF (1995) *Solanum* (Solanaceae) in Uganda. Bothalia 25:43-59

Bussell JD, Waycott M, Chappill JA (2005) Arbitrarily amplified DNA markers as characters for phylogenetic inference. Persp Plant Ecol Evol Syst 7:3-26

Calonje M, Martín-Bravo S, Dobeš C, Gong W, Jordon-Thaden I, Kiefer C, Kiefer M, Paule J, Schmickl R, Koch MA (2009) Non-coding nuclear DNA markers in phylogenetic reconstruction. Plant Syst Evol 282:257-280

Carlquist S (1967) The biota of long-distance dispersal. V. Plant dispersal to the Pacific Islands. B Torrey Bot Club 94:129-162

Carlquist S (1983) Intercontinental dispersal. In: K Kubitzki (Ed) Dispersal and distribution. Hamburg, Germany, Paul Parey pp. 37-47

Chase MW, Knapp S, Cox AV, Clarkson JJ, Butsko Y, Joseph J, Savolainen V, Parokonny AS (2003) Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). Ann Bot 90:1446-1454

Child A, Lester RN (2001) Synopsis of the genus *Solanum* L. and its infrageneric taxa. In: Berg, RGVD, Barendse WM, Weerden WMVD, Mariani C (Eds) Solanaceae V: Advances in Taxonomy and Utilization. Nijmegen Univ. Press, The Netherlands pp. 39-52

Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun J-H, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB, Cook DR. (2004) A sequence-based genetic map of *Medicago trunculata* and comparison of marker colinearity with *M. sativa*. Genetics 166:1463-1502

Ciarmiello LF, Piccirillo P, Pontecorvo G, De Luca A, Kafantaris I et al. (2010) A PCR based SNPs marker for specific characterization of English walnut (*Juglans regia* L.) cultivars. Mol Biol Rep doi: 10.1007/s11033-010-0223-y

Cisneros PL, Quiros CF (1995) Variation and phylogeny of the triploid cultivated potato *Solanum shaucha* Juz. et. Buk. based on RAPD and isozyme markers. Genet Res Crop Evol 42:373-386

Clarkson JJ, Knapp S, Garcia VF, Olmstead RG, Leitch AR, Chase MW (2004) Phylogenetic relationships in *Nicotiana* (Solanaceae) inferred from multiple plastid DNA regions. Mol Phylogenet Evol 33:75-90

Clausen AM, Spooner DM (1998) Molecular support for the hybrid origin of the wild potato species *Solanum* × *rechei* (*Solanum* sect. *Petota*). Crop Sci 38:858-865

Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) Polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol Rep 27:86-93

Collins DJ, Eastwood FW, Swan JM, Fryer C (1976) A steroid industry in Australia. Search 7:378-383

Crepet WL, Nixon KC, Gandolfo MA (2004) Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from cretaceous deposits. Am J Bot 91:1666-1682

Crisp M, Cook L, Steane D (2004) Radiation of the Australian flora: what can comparisions of molecular phylogenies accross multiple taxa tell us about the evolution of diversity in present-day communities? Philos T R Soc B 359:1551-1571

Croll D, Sanders IR (2009) Recombination in *Glomus intraradices*, a supposed ancient asexual arbuscular mycorrhizal fungus. BMC Evol Biol 9:13

D'Arcy WG (1972) Solanaceae studies II: typification of subdivisions of *Solanum*. Ann Mo Bot Gard 59:262-278

D'Arcy WG (1991) The Solanaceae since 1976, with a review of its biogeography. In: Hawkes JG, Lester RN, Nee M, Estrada-Ramos N (Eds) Solanaceae III: taxonomy, chemistry, evolution. Kew, UK, Royal Botanic Gardens pp. 75-137.

Danert S. (1970) Infragenerische Taxa der Gattung *Solanum* L. Kulturpflanze 18:253-297

Davierwala AP, Ramakrishna W, Chowdari V, Ranjekar PK, Gupta VS (2001) Potential of $(GATA)_n$ microsatellites from rice for inter- and intra-specific variability studies. BMC Evol Biol 1:7

De Keyser ED, De Riek J, Van Bockstaele E (2009) Discovery of species-wide EST-derived markers in *Rhododendron* by intron-flanking primer design. Mol Breeding 23:171-178

De Riek J, Calsyn E, Everaert I, Van Bockstaele E, De Loose M (2001) AFLP based alternatives for the assessment of Distinctness, Uniformity and Stability of sugar beet varieties. Theor Appl Genet 103: 1254-1265

DeBry RW, Slade NA (1985) Cladistic analysis of restriction endonuclease cleavage maps within a maximum-likelihood framework. Syst Zool 34:21-34

DePamphilis CW, Palmer JD (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348:337-339

Demesure BN, Sodzi R, Petit J (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129-131

Dice LR (1945) Measuring of amount of ecological association between species. Ecology 26:297-302

Dillenius JJ (1732) *Hortus Elthamensis* II pp. 344-366

Divaret I, Margale E, Thomas G (1999) RAPD markers on seed bulks efficiently assess the genetic diversity of a *Brassica oleracea* L. collection. Theor Appl Genet 98:1029-1035

Doganlar S, Frary A, Daunay M-C. Lester RN, Tanksley SD (2002) A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. Genetics 161:1697-1711

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biology 4:e88

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees**.** BMC Evol Biol 7**:**214

Dunal MF (1813**)** Historie des *Solanum*, et des genres qui ont été confundus avec eux. Koenig, Paris

Dunal MF (1816) *Solanorum generumque affinium synopsis*, Ed. 2. Renaud, Montpellier

Dunal MF (1852) Solanaceae In: De Candolle AP (Ed) Prodromus systematis naturalis regni vegetabilis, Vol. 13. Paris: Treuttel and Würz

Edmonds JM (1972) A synopsis of the taxonomy of *Solanum* sect. *Solanum* (*Maurella*) in South America. Kew Bull 27: 95-114

Edmonds JM (1977) Taxonomic studies of *Solanum* sect. *Solanum* (*Maurella*). Bot J Linn Soc 75: 141-178

Edmonds JM (1979) Biosystematics of *Solanum* L. section *Solanum* (*Maurella*). In: Hawkes JG, Lester RN, Skedling AD (Eds) The Biology and Taxonomy of the Solanaceae, Academic Press London pp. 529-548

Felsenstein J (1989) PHYLIP—Phylogeny Inference Package (Version 3.2). Cladistics 5:164 -166

Farris JS (1989) The retention index and homoplasy excess. Syst Zool 38: 406-407

Farris JS, Albert VA, Källersjö M, Lipscomb D, Kluge AG (1996) Parsimony jackknifing outperforms neighbor-joining. Cladistics 12: 99-124

Fawcett JA, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. PNAS USA 106:5737-5742

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155:279-284

Flores HE, Hoy HW, Pickard JJ (1987) Secondary metabolites from root cultures. Trend Biotech 5:64-69

Forster JGA (1786a) *Dissertatio inauguralis botanico-medica de plantis esculentis insularum oceani australis*. Halle, Typis Frankianus

Forster JGA (1786b) *De plantis esculentis insularum oceani australis commentatio botanica*. Hande & Spener, Berlin

Forster JGA (1786c) *Florae insularum australium prodromus*. J.C. Dietrich, Göttingen

Francisco-Ortega J, Hawkes JG, Lester RN, Acebes-Ginovés JR (1993) *Normania*, an endemic Macarinesian genus distinct from *Solanum* (Solanaceae). Plant Syst Evol 185:189-205

Fukuda T, Yokoyama J, Ohashi H (2001) Phylogeny and biogeography of the genus *Lycium* (Solanaceae): inferences from chloroplast DNA sequences. Mol Phylogenet Evol 19:246-258

Furini A, Wunder J (2004) Analysis of eggplant (*Solanum melongena*)-related germplasm: morphologial and AFLP data contribute to phylogenetic interrelations and germplasm utilization, Theor Appl Genet 108:197-208

Furman BJ, Grattapaglia D, Dvorak WS, O'Malley DM (1997) Analysis of genetic relationships of Central American and Mexican pines using RAPD markers that distinguish species. Mol Ecol 6:321-331

Garcia VF, Olmstead RG (2003) Phylogenetics of tribe Anthocercideae (Solanaceae) based on *ndh*F and *trn*L/F sequence data. Syst Bot 18:609-615

Gáborjányi R (1969) Dohány mozaik virus fertőzés *Solanum laciniatum*on. Herb Hung 8:73-79

Gáborjányi R, Nagy F (1970) A *Solanum laciniatum*ról izolált vírusbetegségek. Herb Hung 9:67-72

Galanes IT, Webb DT, Rosario O (1984) Steroid production by callus and cell suspension cultures of *Solanum aviculare* Forst. J Nat Prod 47:373-376

Gebhardt C, Walkemeier B, Heselewski H, Barakat A, Delseny M, Stüber K (2003) Comparative mapping between potato (*Solanum tubereosum*) and *Arabidopsis thaliana* reveal structurally conserved domain and ancient duplications in the potato genome. Plant J 34:529-541

Gerasimenko II (1969) Inter and intraspecific hybridization in the genus *Solanum* subgenus *Archaesolanum* Bitter ex Marzell. Genetika 5:51-60

Gerasimenko II (1970) *Conspectus subgeneris Archaesolanum* Bitt. ex Marz. generis *Solanum* L. Novosti Sist. Vyssh. Rast. 7:270-275

Gerasimenko II (1971) Interpscific variation in *Solanum laciniatum* Ait. Rast. Resur. 7:363-371

Gielly L, Taberlet P (1994) The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences Mol Biol Evol 11:769-777

Gillham NW (1978) Organelle heredity. Raven Press, New York, USA

Gilli A (1970) Bestimmungsschlüssel der Subgenera und Sektionen der Gattung *Solanum*. Feddes Repert 81:429-435

Goloboff PA (1999) NONA, version 2.0. www.cladistics.com.

Goloboff PA, Farris JS, Nixon KC (2008) TNT: Tree analysis using new technology. Program and documentation available from the authors and at http://www.zmuc.dk/public/phylogeny/TNT/

Gort G, van Hintum T, van Eeuwijk F (2009) Homoplasy corrected estimation of genetic similarity from AFLP bands, and the effect of the number of bands on the precision of estimation. Theor Appl Genet 119:397-416

Grativol C, Lira-Medeiros CDF, Hemerly AS, Ferreira PCG (2010) High efficiency and reliability of inter-simple sequence repeats (ISSR) markers for evaluation of genetic diversity in Brazilian cultivated *Jatropha curcas* L. accessions. Mol Biol Rep doi: 10.1007/s11033-010-0547-7

Gupta PK, Varshney RK, Sharma PC, Ramesh B (1999) Molecular markers and their application in wheat breeding. Plant breed 118:369-90

Hair JB (1966) Biosystematics of the New Zealand Flora, 1945-64. New Zeal J Bot 4:559-595

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acid Symp Ser 41:95-98

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol Ecol 8:513-525

Harvey CF, Gill GP, Crossman C, Fraser LG (1997) Assessing relationships of kumara cultivars by RAPD analysis. New Zeal J Bot 35:479-485

Hausdorf B (1998) Weighted ancestral area analysis and a solution of the redundant distribution problem. Syst Biol 47:445-456

Hawkes JG, Smith P (1965) Continental drift and the age of angiosperm genera. Nature 207:48-50

Hawkes JG (1990) The potato: evolution, biodiversity and genetic resources. Whasington DC, USA, Smithsonian Institution Press

Heiberg E (2008) SEEVA ver. 0.33. Software for Spatial Evolutionary and Ecological Vicariance Analysis. Available from the author at http://seeva.heiberg.se

Heiberg E, Struwe L (2008) SEEVA manual. On-line publication, Rutgers University. Available at http://seeva.heiberg.se and http://www.rci.rutgers.edu/~struwe/seeva

Henry RJ (1997) Practical applications of plant molecular biology. Chapman & Hall, London, pp 59–98

Hepper FN (1979) On typifying Linnean names of Solanaceae. Linnean Society Symposium 7: 131-134

Holsinger KE, Jansen RK (1993) Phylogenetic analysis if restriction site data. Methods Enzymol 224:439-455

Hooker JD (1857) The botany of the Antarctic voyage. III. Flora Tasmaniae. Part 1 (4). London, UK, Reeve.

Hunzinker AT (2001) *Genera Solanacearum*. A. R. G. Ganter, Ruggell. pp 49-85

Jacobs MMJ, van den Berg RG, Vleeshouwers VGAA, Visser M, Mank R, Sengers M, Hoeskatra R, Vosman B (2008) AFLP analysis reveals a lack of phylogenetic structure within *Solanum* section *Petota***.** BMC Evol Biol 8:145

Jacoby A, Labuschagne MT, Viljoen C. D. (2003) Genetic relationship between Southern African *Solanum retroflexum* Dun. and other related species measured by morphological and DNA markers, Euphytica 132: 109-113

Janson CH (1983) Adaptation of fruit morphology to dispersal agents in a neotropical forest. Science 219:187-189

Jasik J, Boggetti B, Caricato G, Mantell S (1997) Characterization of morphology and root formation in the model woody perennial shrub *Solanum aviculare* Forst. expressing *rol*ABC genes of *Agrobacterium rhizogenes*. Plant Sci 124:57-68

Jigden B, Wang H, Samadan N, Yang D-C (2010) Molecular identification of oriental medical plant *Anemarrhena asphoides* Bunge (‘Jimo’) by multiplex PCR. Mol Biol Rep 37:955-960

John DB, Michelle W, Jennifer AC (2005) Arbitrarily amplified DNA markers as characters for phylogenetic inference. Persp Plant Ecol, Evol Syst 7:3-26

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Wiel C van de, Bredemejer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert E, Volckaert G, Rueda J, Lincero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breed 3:381-190

Joshi C, Zhou H, Huang X, Chiang VL (1997) Context sequences of translation initiation codon in plants. Plant Mol Biol 35:93–1001

Kapustin Y, Souvorov A, Tatusova T, Lipman D (2008) SPLIGN: algorithms for computing spliced alignments with identification of paralogs. *Biology Direct* 3:20

Kardolus JP, van Eck HJ, Berg RGVD (1998) The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy. Plant Syst Evol 210:87-103

Karihaloo JL, Brauner S, Gootlieb LD (1995) Random amplified polymorphic DNA variation in the eggplant *Solanum melongena* L (Solanaceae) Theor Appl Genet 90:767-770

Karp A, Seberg O, Buiatti M (1996) Molecular techniques in the assessment of botanical diversity. Ann Bot 78:143-149

Keighery G (1984) *Solanum* berries and wattlebirds. West Aust Nat 16:22

Kemp EM (1978) Tertiary climatic evolution and vegetation history in the Southeast Indian Ocean region. Paleogeogr Paleoecol 24:169-208

Khanna P, Sharma GL, Rathore AK, Manot SK (1977) Effect of cholesterol on in vitro suspension tissue cultures of *Costus speciousus* (Kown) Sm., *Dioscorea floribunda* Mart. & Gal., *Solanum aviculare* Forst. and *Solanum xanthocarpum* Schard & Wendl. Indian J Exp Biol 15:1025-1027

Kingston SE, Adams LD, Rosel PE (2009) Testing mitochondrial sequences and anonymous nuclear markers for phylogeny reconstruction in a rapidly radiating group: molecular systematics of Delphininae (Cetacea:Odontoceti:Delphinidae). BMC Evol Biol 9:245

Kingston SE, Rosel PE (2004) Genetic differentiation among recently diverged Delphinid taxa determined using AFLP markers. J Hered 95:1-10

Kitching IL (1992) Tree-building techniques. In: Forey PL, Humphries CJ, Kitching IL, Scotland RW, Siebert DJ, Williams DM (Eds), Cladistics: a practical course in systematics. Clarendon, Oxford, pp. 44-71

Kittipongpatana N, Hock RS, Porter JR (1998) Production of solasodine by hairy root, callus, and cell suspension cultures of Solanum aviculare Forst. Plant Cell Tissue and Organ Culture 52:133-143

Kluege AR, Farris JS (1969) Quantitative phyletics and the evolution of anurans. Syst Zool 18:1-32

Knapp S (2006) *Solanum aviculare.* In: Solanaceae Source, 2010. Available via http://www.nhm.ac.uk/research-curation/projects/solanaceaesource/taxonomy/description-detail.jsp?spnumber=1213

Knapp S, Bohs L, Nee M, Spooner DM (2004) Solanaceae: a model for linking genomics with biodiversity. Comp Funct Genom 5:285-291

Knapp S (1989) A revision of the *Solanum nitidum* group (section *Holophylla* pro parte): Solanaceae. Bull Brit Mus (Nat Hist) Bot 19:63-112

Knapp S (1991) A revision of the *Solanum sessile* species group (section *Geminata* pro parte: Solanaceae). Bot J Linn Soc 105:179-210

Koopman WJM (2005) Phylogenetic signal in AFLP data sets. Syst Biol 54:197-217

Korneva EI, Balakhova OF (1973) Polyploid hybrids of *Solanum aviculare* var. *brisbanense*, *S. aviculare* var. *albiflorum* in relation to the appearance of *S. laciniatum* type. Genetika 9:47-53

Korneva EI, Matveenko LF (1979) Geographic localization of steroids containing *Solanum* subg. *Archaesolanum* Bit. with respect to breeding. Genetika 15:1127-1129

Korneva EI, Ostretsova IN, Kondratenko PN (1969) Characteristic biological features of *S. laciniatum* and its intraspecific hybrids. Rast Resur 5:197-201

L'Héritier CL de B (1805) *Solanum.* In: Persoon CH, Synopsis plantarum, vol.1. Paris, CF Cramer

Lamarck JPAPM de B (1792) *Encyclopédie méthodique, Botanique*, vol. 3(2). Liége, Plomteux.

Lambertini C, Frydenberg J, Gustafson MHG, Brix H (2008) Herbarium specimens as a source of DNA for AFLP fingerprinting of *Phragmites* (Poaceae): possibilities and limitations. Plant Syst Evol 272:223-231

Lara-Cabrera SI, Spooner DM (2004) Taxonomy of North and Central American diploid wild potato (*Solanum* sect. *Petota*) species: AFLP data. Plant Syst Evol 248:129-142

Lester RN (1997) Typification of nineteen names of African *Solanum* species described by A. Richard and others, including *S. campylacanthum* and *S. panduriforme*. Bot J Linn Soc 125: 273-293

Levin RA, Myers N, Bohs L (2006) Phylogenetic relationship among the „Spiny Solanums" (*Solanum* subgenus *Leptostemonum*, Solanaceae) Am J Bot 93:157-169

Levin RA, Watson K, Bohs L (2005) A four-gene study of evolutionary relationships in *Solanum* section *Acanthophora*. Am J Bot 92:603–612

Linnaeus C (1753) *Species plantarum*, 2 vols. Stockholm

Liu BH (1998) Statistical genomics: Linkage, mapping and QTL analysis. CRC Press; pp 158-159

Liu QL, Zhang N, Li L, Liu J (2010) Identification of *Elymus* (Triticaceae, Poaceae) and its related genera genomes by RFLP analysis of PCR-amplified *Adh* genes. Mol Biol Rep 37:3249-3257

Lorenz-Lemke AP, Togni PD, Mäder G, Kriedt RA, Stehmann JR, Salzano FM, Bonatto SL, Freitas LB (2010) Diversification of plant species in a subtropical region of eastern South American highlands: a Phylogeographic perspective on native *Petunia* (Solanaceae). Mol Ecol 19:5240-5251

Löhne C, Borsch T, Jacobs SWL, Hellquist CB, Wiersema JH (2008) Nuclear and plastid DNA sequences reveal complex reticulate patterns in Australian water-lilies (*Nymphaea* subgenus *Anecphya*, Nymphaeaceae). Aust Syst Bot 21:229-250

Mace ES, Lester RN, Gebhardt CG (1999) AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). Theor Appl Genet 99:626-633

Magallón S, Crane PR, Herendeen PS (1999) Phylogenetic pattern, diversity, and diversification of eudicots. Ann Mo Bot Gard 86:297-372

Mann JD (1979) Production of solasodine for the pharmaceutical industry. Adv Agron 30:207-245

Marks CE (2010) Definition of South Pacific taxa of *Nicotiana* section *Suaveolentes*. Muelleria 28:74-94

Máthé I, Földesi D (1965) Az orvosi csucsor – *Solanum laciniatum* Ait. Magyarország Kultúrflórája 24., Akadémiai Kiadó, Budapest.

Máthé I Jr, Máthé I, Botz L, Koch L (1986) Possibilities of *Solanum* alkaloid production in European temperate zone. Acta Hort 188:193-202

Máthé I, Földesi D, Szabó Z, Sárkány I, Tétényi P (1964) Szteroid hormonalapanyagokat szolgáltató néhány *Solanum* faj botanikai és kémiai összehasonlító vizsgálata. Herb Hung 3:293-307

Máthé I, Jr (1982) Research on the alkaloid of *Solanum* in Hungary. Solanaceae Newsletter 2:36-39

McGregor CE, van Treuren E, Hoekstra R, van Hintum ThJL (2002) Analysis of the wild potato germplasm of the series *Acaulia* with AFLPs: implications for ex situ conservation. Theor Appl Genet 104:146-156

McNeill J, Barrie FR, Burdet HM, Demoulin V, Hawksworth DL et al. (2006) International Code of Botanical Nomenclature (Vienna Code). Ruggell, Liechtenstein, ARG Ganter Verlag

Metzlaff M, Borner T, Hagemann R (1981) Variations of chloroplast DNAs in the genus Pelargonium and their biparental inheritance. Theor Appl Genet 60:37-41

Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends Plant Sci 12:106-117

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. PNAS USA 88:9828-9832

Miller JT, Spooner DM (1999) Collapse of species boundaries in the wild potato *Solanum* brevicaule complex (Solanaceae sect. *Petota*): molecular data. Plant Syst Evol 214:103-130

Miller JS (2002) Phylogenetic relationships and the evolution of gender dimorphism in *Lycium* (Solanaceae). Syst Bot 27:416-428

Miller JS, Diggle PK (2007) Correlated evolution of fruit size and sexual expression in andromonoecious *Solanum* sctions *Acanthophora* and *Lasiocarpa* (Solanaceae). Am J Bot 94:1706-1715

Montero-Castro JC, Delgado-Salinas A, De Luna E, Eguiarte LE (2006) Phylogenetic analysis of *Cestrum* section *Habrothamnus* (Solanaceae) based on plastid and nuclear DNA sequances. Syst Bot 31:843-850

Moscone EA (1992) Estudios de cromosomas meiticos en Solanaceae de Argentina. Darwiniana 31:261-197

Mueller F (1855) Account of the gunyang: a new indigenous fruit of Victora. In: Transactions and Proceedings of the Victorian Institute for the Advancement of Science. Melbourne, Australia, G. Robertson pp. 67-70

Mueller LA, Solo TH, Taylor N, Skwarecki B, Buels R, Binns J, Lin C, Wright MH, Ahrens R, Wang Y, Herbst EV, Keyder ER, Menda N, Zamir D, Tanksley SD (2005) The SOL genomics network: A comparative resource for Solanaceae biology and beyond. Plant Physiol 138:1310-1317

Müller K, Borsch T, Hilu KW (2006) Phylogenetic utility of rapidly evolving DNA at high taxonomic levels: contrasting *matK*, *trn*T-F, and *rbc*L in basal angiosperms. Mol Phylogenet Evol 41:99-117

Mummenhoff K, Franzke A (2007) Gone with the bird: late Tertiary and Quaternary intercontinental long-distance dispersal and allopolyploidization in plants. Syst Biodivers 5:255-260

Mummenhoff K, Linder P, Friesen N, Bowman JL, Lee J.-Y, Fanzke A (2004) Molecular evidence for biocontinental hybridsogenous genomic constitution in *Lepidium* sensu strict (Brassicaceae) species from Australia and New Zealand. Am J Bot 91:254-261

Nagy S, Cernák I, Poczai P, Taller J (2008) PICcalc: Online Polymorphic Information Content Calculator. Available at: http://www.georgikon.hu/pic/english/default.aspx

Nee M (1999) Synopsis of *Solanum* in the New World. In: Nee M, Symon DE, Lester RN, Jessop JP (Eds) Solanaceae IV: advances in biology and utilization. Kew, UK, Royal Botanic Gardens pp. 285-333

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. PNAS USA 76:5269-5273

Nixon KC (1999) The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15:407-414

Nixon KC (2002) Winclada, version 1.00.08. www.cladistics.com

Nixon KC, Carpenter JM (1996) On consensus, collapsibility, and clade concordance. Cladistics 12:305-321

Olmstead RG, Sweere JA (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in Solanaceae. Syst Biol 43:467-481

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematics: a review of methods and data analysis. Am J Bot 81:1205-1224

Olmstead RG, Palmer JD (1997) Implications for the phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. Syst Bot 22:19-29

Olmstead RG, Sweere JA, Spangler RE, Bohs L, Palmer JD (1999) Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. In: Nee M, Symon D E, Lester RN, Jessop JP (Eds) Solanaceae IV: Advances in Biology and Utilization. Royal Botanic Gardens, Kew pp. 111-137

Olmstead RG, Bohs L, Migid HA, Santiago-Valentin E, Garcia VF, Collier SM (2008) A molecular phylogeny of the Solanaceae. Taxon 57:1159-1181

Paape T, Igric B, Smith SD, Olmstead E, Bohs L, Kohn JR (2008) A 15-Myr-old genetic bottleneck. Mol Biol Evol 25:655-663

Pain CF, Ollier CD (1983) Drainage patterns and tectonics around Milne Bay, Eastern Papua New Guinea. Rev Geomorph Dyn 32:113-120

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357-358

Palmer JD (1985) Evolution of chloroplast and mitochondrial DNA in plants and algae. In: MacIntyre RJ (Ed) Molecular evolutionary genetics. Plenum, New York, USA, pp.131-240

Palmer JD (1987) Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. Am Natur 130:S6-S29

Palmer JD (1991) Plastid chromosome: structure and evolution. In: Bogorad L, Vasil IK, (Eds) The molecular biology of plastids. San Diego, USA, Academic Press pp. 5–53

Pamidimarri DVNS, Mastan SG, Rahman H, Reddy MP (2010) Molecular characterization and genetic diversity analysis of *Jatropha curcas* L. in India using RAPD and AFLP analysis. Mol Biol Rep 37:2249-2257

Panwar P, Saini RK, Sharma N, Yadav D, Kumar A (2010) Efficiency of RAPD, SSR and cytochrome P450 gene based markers in accessing genetic variability amongst finger millet (*Eleusine coracana*) accessions. Mol Biol Rep 37: 4075-4082

Parks DH, MacDonald NJ, Beiko RG (2009) Tracking the evolution and geographic spread of influenza A. PLoS Currents: Influenza RRN1014

Parks D, Beiko R (2010) GenGIS MapMaker v1.0, Available at http://kiwi.cs.dal.ca/GenGIS/MapMaker

Parks DH, Porter M, Churcher S, Wang S, Blouin C, Whalley J, Brooks S, Beiko RG. (2009) GenGIS: A geospatial information system for genomic data. Genome Res 19:1896-1904

Peel MC, Finlayson BL, McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. Hydrol Earth Syst Sc 11:1633-1644

Pigram, CJ, Davies HL (1987) Terranes and the accretion history of the New Guinean orogen. BMR J Austr Geol Geophy 10:193-211

Poczai P, Taller J, Szabó I (2008) Analysis of phylogenetic relationships in the genus *Solanum* (Solanaceae) as revealed by RAPD markers. Plant Syst Evol 275:59-67

Poczai, Mátyás K, Taller J, Szabó I (2010) Study of the origin of the rearly cultivated edible Solanum species: morphological and molecular data. Biol Plant 54:543-546

Poczai P, Hyvönen J (2010) Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. 37:1897-1912

Poczai P, Hyvönen J (2011) On the origin of *Solanum nigrum*: can networks help? Mol Biol Rep 38:1171-1185.

Poczai P, Hyvönen J, Symon DE (2011a) Phylogeny of kangaroo apples (*Solanum* subg. *Archaesolanum*). Mol Biol Rep doi:10.1007/s11033-011-0675-8

Poczai P, Cseh A, Taller J, Symon DE (2011b) Genetic diversity and relationships in *Solanum* subg. *Archaesolanum* (Solanaceae) based on RAPD and chloroplast PCR-RFLP analyses. Plant Syst Evol 291:35-47

Polhemus DA, Polhemus JT (1998) Assembling New Guinea: 40 million years of island arc accretion as indicated by the distributions of aquatic Heteroptera (Insecta). In: Hall R, Holloway JD (Eds) Biogeography and Geological Evolution of SE Asia. Leiden, Netherlands: Buckhuys pp. 327-340

Pólya L, Pólya Borsos I (1966) A *Solanum laciniatum* Ait. magnyugalmáról. Acta Biol Debr 4:107-116

Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253-1256

Proctor VW (1968) Long-distance dispersal of seeds by retention in digestive tract of birds. Science 160:321-322

Pulquério MJF, Nichols RA (2007) Dates from the molcecular clock: how wrong can we be? Trend Ecol Evol 22:180-184

Queiroz De A (2005) The reconstruction of oceanic dispersal in historical biogeography. Trend Ecol Evol 20:68-73

Rambaut A (2008) FigTree ver. 1.3.1 http://tree.bio.ed.ac.uk/software/figtree/

Rambaut A, Drummond AJ (2007) Tracer ver. 1.5 http://tree.bio.ed.ac.uk/software/tracer/

Randell BR, Symon DE (1976) Chromosome numbers in Australian *Solanum* species. Aust J Bot 24: 369-379

Renner SS (1999) Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data. Am J Bot 86:1301-1315

Rodríguez A, Spooner DM (1997) Chloroplast DNA analysis of *Solanum bulbocastanum* and *S. cardiophyllum*, and evidence for the distinctiveness of *S. cardiophyllum* subsp. *ehrenbergii* (sect. *Petota*). Syst Bot 22:31-43

Roe KE (1972) A revision of *Solanum* section *Brevantherum* (Solanaceae) Britonia 24:239-278

Ronquist F (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst Biol 46:195-203

Rubio-Moraga A, Castillo-López R, Gómez-Gómez L, Ahrazem O (2009) Saffron is a monomorphic species as revealed by RAPD, ISSR and microsatellite analyses. BMC Research Notes 2:189

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am J Bot 84:1120-1136

Sanmartín I, Ronquist F (2004) Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst Biol 53:216-243

Santiago-Valentin E, Olmstead RG (2003) Phylogenetics of the Antillean Goetzeoideae (Solanaceae) and their relationships within the Solanaceae based on shloroplast and ITS DNA sequence data. Syst Bot 28:452-460

Sauquet H, Doyle JA, Scharaschkin T, Borsch T, Hilu KW, Chatrou LW, Thomas AL (2003) Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. Bot J Linn Soc 142:125-186

Sawant SV, Singh PK, Gupta SK, Madnala R, Tuli R (1999) Conserved nucleotide sequences in highly expressed genes in plants. J Genet 78:123-31

Schlüter PM, Harris SA (2006) Analysis of multilocus fingerprinting data sets containing missing data. Mol Ecol Not 6:569-572

Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC (2004) Mining EST database to resolve evolutionary events in major crop species. Genome 47:868-876

Schnabel A, Wendell JF (1998) Cladistic biogeography of *Gleditsia* (Leguminosae) based on *ndh*F and *rpl*16 chloroplast gene sequences. Am J Bot 85:1753-1765

Scotese CR, Gahagan LM, Larson RL (1988) Plate tectonic reconstructions of the Cretaceous and Cenozoic ocean basins. Tectonophysics 155:27-48

Seithe A (1962) Die Haarten der Gattung *Solanum* L. und ihre taxonomische Verwertung. Bot Jahr 81:261-336

Sica M, Gamba G, Montieri S, Gaudio L, Aceto S (2005) ISSR markers show differentiation among Italian populations of *Asparagus acutifolius* L. BMC Genetics 6:17

Simmons MP, Zhang L-B, Webb CT, Müller K (2007) A penalty of using anonymous dominant markers (AFLPs, ISSRs and RAPDs) for phylogenetic inference. Mol Phylogenet Evol 42:528-542

Soltis DE, Soltis PS (1999) Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis DE, Soltis PS, Doyle JJ (Eds) Molecular Systematics of Plants II: DNA sequencing. Kluwer Academic Publ, London, UK pp. 1-43

Spooner DM, Sytsma KJ, Conti E (1991) Chloroplast DNA evidence for genome differentiation in wild potatoes (*Solanum* sect. *Petota*: Solanaceae). Am J Bot 78:1354-1366

Spooner DM, Systma KJ (1992) Reexamination of series relationships of Mexican and Central American wild potatoes (*Solanum* sect. *Petota*): evidence from chloroplast DNA restriction site variation. Syst Bot 17:432-448

Spooner DM, Anderson GJ, Jansen RK (1993) Chloroplast DNA evidence for the interrelationship of tomatoes, potatoes, and pepinos (Solanaceae). Am J Bot 80:676-688

Spooner DM, Tivang J, Nienhuis J, Miller JT, Douches DS, Contreras MA (1995) Comparison of four molecular markers in measuring relationships among the wild potato relatives *Solanum* section *Etuberosum* (subgenus *Potatoe*). Theor Appl Genet 92:532-540

Spooner DM, Tivang J, Nienhuis J, Miller JT, Douches DS, Contreras MA (1996) Comparison of four molecular markers in measuring relationships among the wild potato relatives *Solanum* section *Etuberosum* (subgenus *Potatoe*). Theor Appl Genet 92:532-540

Spooner DM, Ugarte ML, Skroch PW (1997) Species boundaries and interrelationships of two closely related sympatric diploid wild potato species, *Solanum astleyi* and *S. boliviense*, based on RAPDs. Theor Appl Genet 95:764-771

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. PNAS USA 102:14694-14699

Spooner DM (2009) DNA barcoding will frequently fail in complicated groups: an example in wild potatoes. Am J Bot 96:1177-1189

Stedje B, Bukenya-Ziraba R (2003) RAPD variation in *Solanum anguivi* Lam. and *S. aethiopicum* L. (Solanaceae) in Uganda. Euphytica 131:283-297

Stern SR, Weese T, Bohs LA (2010) Phylogenetic relationships in *Solanum* section *Androceras* (Solanaceae). Syst Bot 35:885-893

Stewart CNJ, Porter DM (1995) RAPD profiling in biological conservation: an application to estimating clonal variation in rare and endangered *Ilimna* in Virginia. Biol Conserv 74:132-142

Stubbe W (1984) Biparental plastid inheritance in *Oenothera organensis* Munz. Acta Bot Neerl 33:229

Subroto MA, Doran PM (1994) Production of steroidal alkaloids by hairy roots of *Solanum aviculare* and the effect of gibberellic acid. Plant Cell Tiss Organ Cult 38:93-102

Sugiura M (1989) The chloroplast chromosomes in land plants. Ann Rev Cell Biol 5:51-70

Sugiura M (1992) The chloroplast genome. Plant Mol Biol 19:149-168

Swofford DL, Olson GJ (1990) Phylogeny reconstruction. In: Hillis DM, Moritz C, Mable BK (Eds), Molecular systematics. Sinauer Associates, Sunderland, pp. 411-501

Symon DE (1970). Dioecious Solanums. Taxon 19:909-910

Symon DE (1979) The genus *Solanum* in Australia. In: Hawkes JG, Lester RN, Skelding AD (Eds) The biology and taxonomy of The Solanaceae. Academic Press, London, pp. 125-127

Symon DE (1981) A revision of *Solanum* in Australia. J Adel Bot Gard 4:1-367

Symon DE (1984) *Solanum* sect. *Archaesolanum* (Bitt. Ex Marzell) Danert. In: van Balgooy MMJ (ed) Pacific plant areas, vol 4. National Herbarium of the Netherlands, pp 215–245

Symon DE (1985) Solanaceae of New Guinea. J Adel Bot Gard 8:1-171

Symon DE (1986) The phytogeography of New Guinea *Solanum* (Solanaceae). Blumea 31:319-328

Symon DE (1991) Gondwanan elements of the Solanaceae, In: Hawkes JG, Lester RN, Nee M, Estrada-Ramos N (Eds) Solanaceae III: Taxonomy, chemistry, evolution. Kew, UK Royal Bot Gard Press

Symon DE (1994) Kangaroo apples: *Solanum* sect. *Archaesolanum*. Keswick, South Australia, Published by the author

Szalay-Marzsó L, Nagy F (1969) A burgonyabogár (*Leptinotarsa decemlineata* Say) elleni védekezés *Solanum laciniatum* kultúrákban. Herb Hung 8:93-99

Szmidt AE, Alden T, Hallgren JE (1987) Paternal inheritance of chloroplast DNA in *Larix*. Plant Mol Biol 9:59-64

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105-1109

Takács AP (2001) A burgonya Y-vírus (Potato Y Potyvirus, PVY) jellemzése, az NTN törzzsel szembeni rezisztencia a vad *Solanum* fajokban. PhD thesis, University of Pannonia, Keszthely

Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141-1160

Vaněk T, Valterová I, Vaňková R, Vaisar T (1999) Biotransformation of (-)-limonene using *Solanum aviculare* and *Dioscorea deltoidea* immobilized plant cells. Biotechnol Lett 21:625-628

Verzár Petri G (1964) Anatomy of *Solanum laciniatum* Ait. Ann Uni Sci Bud Sect Biol 7:241-254

Verzár Petri G, Sárlány S, Szentpéteri G, Gracza P, Stieber J (1967) A *Solanum laciniatum* Ait. hajtásának posztmortális vizsgálata. Acta Pharma Hung 1:112-117

Vijverberg K, Kuperus P, Breeuwer JAJ, Bachmann K (2000) Incipient adaptive radiation of New Zealand and Australian *Microseris* (Asteraceae): an amplified fragment length polymorphism (AFLP) study. J Evol Biol 13:997-1008

Vijverberg K, Mes THM, Bachmann K (1999) Chloroplast DNA evidence for the evolution of *Miscroseris* (Asteraceae) in Australia and New Zealand after long-distant dispersal from western North America. Am J Bot 86:1448-1463

Vleeshouwers VGAA, Martens A, Dooijeweert WV, Colon LT, Govers F, Kamoun S (2001) Ancient diversification of the *Pto* kinase family preceded speciation in *Solanum*. *MPMI* 14:996-1005

Vocel H, Horn AJW (1992) Variability of yield characters and selection in *Solanum laciniatum* Ait. Plant Breeding 109:192-197

Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Re 23:4407-4414

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, Allard RW (1987) Chloroplast DNA polymorphisms in lodge pole and jack pines and their hybrids. PNAS USA 84:2097-2100

Walbot V, Warren C (1988) Regulation of *Mu* element copy number in maize lines with an active or inactive Mutator transposable element system. Mol Gen Genet 211:27-34

Wang H, Sun H, Kwon W-S, Jin H, Yang D-C (2010) A PCR-based SNP marker for specific authentication of Korean ginseng (*Panax ginseng*) cultivar "Chunpoong". Mol Biol Rep 37:1053-1057

Wang Y, Diehl A, Wu F, Vrebalov J, Giovannoni J, Siepel A, Tanksley SD (2008) Sequencing and comparative analysis of a conserved syntenic segment in the Solanaceae. Genetics 180:391-408

Weese TL, Bohs L (2007) A three-gene phylogeny of the genus *Solanum* (Solanaceae). Syst Bot 32:445-463

Weese TL, Bohs L (2010) Eggplant origins: out of Africa, into the Orient. Taxon 59:49-56

Whalen MD (1979a) Taxonomy of *Solanum* section *Androceras*. Gentes Herb. 11:359-426

Whalen MD (1979b) Speciation in *Solanum*, section *Androceras* In: Hawkes JG, Lester RN, Skedling AD (Eds) The Biology and Taxonomy of the Solanaceae. Academic Press, London pp. 581-596

Whalen MD (1984) Conspectus of species groups in *Solanum* subgenus *Leptostemonum*. Gentes Herb 12:179-282

Whalen MD, Caruso E (1983) Phylogeny in *Solanum* sect. *Lasiocarpa* (Solanaceae). Congruence of morphological and molecular data. Sys Bot 8:369-380

Whalen MD, Coshch DE, Heiser CB (1981) Taxonomy of *Solanum* section *Lasiocarpa*. Gentes Herb 12:41-129

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc R Soc Lond B 268:2211-2220

Willson MF, Graff DA, Whelan CJ (1990) Color preferences of frugivorous birds in relation to the colors of fleshy fruits. Condor 92:545-555

Willson MF, Irvine AK, Walsh NG (1989) Vertebrate dispersal syndromes in some Australian and New Zealand plant communities, with geographic comparisons. Biotropica 21:133-147

Winkworth RC, Wagstaff SJ, Glenny D, Lockhart PJ (2002) Plant dispersal N.E.W.S. from New Zealand. Trend Ecol Evol 17:514-520

Winter P., Kahl G (1995) Molecular marker technologies for plant improvement. World J Micr Biotechnol 11:438-448

Wolfe AD, Liston A (1998) Contribution of PCR based methods to plant systematic and evolutionary biology. In: Soltis PS, Soltis DE, Doyle JJ (Eds), Molecular systematic of plants II. Chapman and hall, New York, pp. 43-86

Woodburne MO, Case JA (1996) Dispersal, vicariance, and the late Cretaceous to early Tertiary land mammal biogeography from South America to Australia. J Mammal Evol 3:121-161

Wu F, Mueller LA, Crouzillat D, Petiard V, Tanksley SD (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade. Genetics 174:1407-1420

Yeh FC, Yand RC, Boyle T (1999) POPGENE (Version 1.31): Microsoft Window-bases freeware for population genetic analysis, University of Alberta and the Centre for International Forestry Research. http://www.ualberta.ca/~fyeh/index.htm

Yu S, Kwok KH, Doran PM (1996) Effect of sucrose, exogenous product concentration, and other culture conditions on growth and steroidal alkaloid production by *Solanum aviculare* hairy roots. Enzym Micr Technol 18:238-243

THESIS POINTS

1) Two hypotheses are suggested regarding the origin of *Solanum laciniatum*. Hybrid speciation could have occurred in at least two ways: it is possible, that *S. laciniatum* was developed through diploid hybrid speciation involving *S. aviculare* and *S. vescum* as female parent, and then the entire genome is duplicated through autopolyploidy. This alternative hybridization can only be achieved if *S. laciniatum* is an ancient hybrid. Another alternative hypothesis might be that *S. laciniatum* was developed through autopolyploid, where the normal genome of *S. aviculare* is duplicated in its entirety and produced tetraploid offspring which were postzygotically isolated from their parent.

2) The monophyletic origin of the *Archaesolanum* clade suggests maintaining it at the subgeneric level, until higher level relationships of the whole genus are resolved and the closest relatives of the group are determined.

3) Divergence time estimates suggest that the most recent common ancestor of kangaroo apples is from the late Miocene (~ 9 Mya), thus representing an Australian floral element that arrived after the isolation of the continent and then radiated. It is suggested that kangaroo apples probably do not represent a case of vicariance dating to a time preceding the separation of South America and Australia during the brake of Gondwana.

4) Long-distance dispersal (LDD) with subsequent speciation is the most likely explanation for the occurrence of kangaroo apples in the SW Pacific. This might have taken place through trans-oceanic dispersal.

5) Two new well-supported subclades within kangaroo apples have been identified: the *Similia* clade with *S. capsiciforme* + *S. symonii* + *S. simile*, and the *Avicularia-Laciniata* clade, further divided into separate *S. vescum* + *S. linearifolium* and *S. laciniatum* + *S. aviculare* + *S. multivenosum* subclades.

6) The two subclades within subg. *Archaesolanum* emerged in the Early Pliocene during the continental aridification of Australia and they further diversified in the Early/Late Pliocene during the brief resurgence of the rainforest ca. 3 Mya. The earlier drying phase has possibly given rise to the *Similia* and *Avicularia-Laciniata* clades ca. 6.5 Mya, while the later short rainforest resurgence led to speciation within these two lineages.

7) *Solanum multivenosum* presumably diversified through vicariance during the continuous uplift of New Guinea and evolved in isolation from other species of the group, adapting to tropical climatic conditions.

ACKNOWLEDGEMENTS

I would like to thank all the people who helped me during my PhD years without whom this thesis would not have been possible.

First, I would like to thank Dr. János Taller for establishing a laboratory at Georgikon (Keszthely, Hungary) where I picked up the micropipette and spent my early years since my BSc and where I've conducted my pilot studies with *Solanum* species and many other topics.

I also thank the previous and present members of the Biotechnology Group at the University of Pannonia, Hungary for making my time in the Lab enjoyable. I also thank for their attitude tolerating my difficult personality. Special thanks goes to Dr. András Cseh and Dr. István Cernák for the early years in the Lab making 'midnight' cloning experiments together; Ahmad Mosapour Gorji for everyday discussions and Iranian music. Thanks for the moments of fun we had together. I will always remember the nice time I had at the Biotech Lab.

Many thanks go to my colleagues Dr. Kincső Decsi, Kinga Mátyás at the Department of Plant Sciences and Biotechnology, University of Pannonia for being always kind and supportive.

I also thank Dr. Zsolt Polgár and the staff of the Potato Research Centre, University of Pannonia for their excellent assistance and work maintaining kangaroo apples in their greenhouses.

I want to thank Dr. David E. Symon, the greatest expert of kangaroo apples of Australia for seeds, herbarium sheets provided for this work, and also for his traditional letters. Thank you for your ideas and help making the molecular phylogeny of *Archaesolanum* real.

Many thanks are due to several botanical gardens and collectors for their assistance providing *Solanum* seeds not just from *Archaesolanum* but from many other groups. This thesis would not have been completed without your huge effort, since I was waiting for years to get proper material for my experiments.

A very special thank go to Prof. Jaakko Hyvönen at the University of Helsinki for helping me to move to Finland where I have continued and finished large parts of my thesis. Thank you for treating me as a colleague and friend. Thanks for the discussions and all the help with analyses made on the kangaroo apple data set as well as for other efforts and help get funding for experiments and easing to get settled in Helsinki. I will never forget your help and kindness.

I also thank CIMO and the Finnish Governmental Scholarship provided for me to make my experiments possible.

Many thanks are due to colleagues working in the Molecular Ecology and Systematics Laboratory at the Viikki Biocampus for facilities. Special thanks are due to Dr. Kalevi Trontti for general full time assistance during my work in Helsinki related to my PhD or even those

which were not related to it. I also thank the staff of the MES Lab for providing everything what I needed in time and order. I'm also grateful for Dr. Éva Tas for making the cloning experiments together at the 'Cloning Corner' of MES. For everybody: thank you for your time, effort, patience and support.

I also want thank Dr. Gunilla Ståhls-Mäkelä and all others working in the Laboratory of the Zoological Museum, Helsinki for very great and memorable time spent there.

I also want to thank colleagues Dr. Neil Bell, Dr. Hidetsugu Miwa and Mohammad Sohrabi from the Botanical Museum, Helsinki for discussions and great time.

I thank my family for supporting me when it was needed.

Finally, I also thank my Ildikó for supporting me more than anyone before, providing me perfect balance needed for a researcher to complete his work. Thanks for being beside me during hard times and bringing light to my life that I can follow.

Many thanks for the help to all of you!

Kaikille suuri kiitos avusta!

Köszönöm a segítséget mindenkinek!

PUBLICATION LIST

Publications related to the PhD thesis

1. Refereed International Journals

POCZAI P, TALLER J, SZABÓ I (2008) Analysis of phylogenetic relationships in the genus *Solanum* (Solanaceae) as revealed by RAPD markers. **Plant Systematics and Evolution** 275:59-67. **IF:1.492**

POCZAI P, HYVÖNEN J (2010) Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. **Molecular Biology Reports** 37:1897-1912. **IF:2.038.**

POCZAI P, CERNAK I, GORJI AM, NAGY S, TALLER J, POLGAR Z (2010) Development of intron targeting (IT) markers for potato and cross-species amplification in *Solanum nigrum* (Solanaceae). **American Journal of Botany** 97: e142-e145. **IF:2.684**

POCZAI P, CSEH A, TALLER J, SYMON DE. (2011) Genetic diversity and relationships in *Solanum* Subg. *Archaesolanum* (Solanaceae) based on RAPD and chloroplast PCR-RFLP analyses. 291:35-47. **Plant Systematics and Evolution IF: 1.410**.

POCZAI P, HYVÖNEN J, SYMON DE (2011) Phylogeny of kangaroo apples (*Solanum* Subg. *Archaesolanum*) doi:10.1007/s11033-011-0675-8 **Molecular Biology Reports IF: 2.038.**

2. National Publications (In Hungarian)

POCZAI P (2011) A fekete csucsorfélék (*Solanum* L. sectio *Solanum*) biológiája, taxonómiája és felhasználásának lehetőségei. **Botanikai Közlemények**, *megjelenés alatt*.

3. Conferences

International conferences

POCZAI P, TALLER J, HYVÖNEN J (2010) Genetic diversity of black nightshades (*Solanum* sect. *Solanum*) In: Green Plant Breeding Technologies Vienna, Austria Febr. 2-5. 2010, Abstract book pp. 80.

National Conferences (*In Hungarian*)

POCZAI P (2008) Tévedés vagy történelem- Kitaibel Pál herbáriumának molekuláris genetikai vizsgálata. Pro Scientia Aranyérmesek IX. Konferenciája, Kaposvár, pp. 79-83.

Publications not related to the PhD thesis

1. Refereed International Journals

POCZAI P, MÁTYÁS K., TALLER J, SZABÓ I (2010) Study of the origin of the rarely cultivated edible *Solanum* species: morphological and molecular data. **Biologia Plantarum** 54(3):543-546. **IF:1.656**

POCZAI P, HYVÖNEN J. (2011) On the origin of *Solanum nigrum*: can networks help? **Molecular Biology Reports** 38:1171-1185. **IF:2.038**

GORJI AM, **POCZAI P**, POLGAR Z, TALLER J (2011) Efficiency of arbitrarily amplified dominant markers (SCOT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. **American Journal of Potato Research** doi:10.1007/s12230-011-9187-2 **IF:0.667**

POCZAI P, MÁTYÁS KK, SZABÓ I, VARGA I, HYVÖNEN J, CERNÁK I, GORJI AM, DECSI K, TALLER J (2011) Genetic variability of thermal *Nymphaea* (Nymphaeaceae) populations based on ISSR markers: implications on relationships, hybridization and conservation. **Plant Molecular Biology Reporter** doi:10.1007/s11105-011-0302-9. **IF:0.711**

2. Refereed National Publications (In English)

POCZAI P, TALLER J, SZABÓ I (2009) Molecular genetic study on a historical Solanum (Solanaceae) herbarium specimen collected by Paulus Kitaibel in the 18th century. **Acta Botanica Hungarica** 51(3-4):337-346.

3. Conferences,Posters

Posters

MÁTYÁS K**, POCZAI P**, CERNÁK I, SZABÓ I, TALLER J (2009) Genetic diversity in thermal *Nymphaea* (Nymphaeaceae) populations based on ISSR markers: implications on phylogeny, hybridization and conservation. IAVS European Vegetation Survey 18th Workshop Rome, Italy.

SZABÓ I, KERCSMÁR V, **POCZAI P**, TALLER J (2010) Morphological and molecular genetic traits of the Hungarian Bear's Ear (*Primula hungarica* Borbás 1895 s. ampl.) 19th International Workshop of European Vegetation Survey. 27 Apr- 02 May, 2010. University of Pécs, Hungary.

4. National Publications (In Hungarian)

National conferences

POCZAI P, SZABÓ I, TALLER J (2005) Különböző *Solanum* vonalak molekuláris genetikai vizsgálata. XI. Ifjúsági Tudományos Fórum, Keszthely, CD kiadvány.

POCZAI P, SZABÓ I, TALLER J (2006) A *Solanum melanocerasum* All. és alakkörének jellemzése molekuláris genetikai módszerekkel. XII. Ifjúsági Tudományos Fórum, Keszthely, CD kiadvány.

POCZAI P (2007) A *Solanum* genus taxonómiai változásai napjainkban – irodalmi feldolgozás. XIII. Ifjúsági Tudományos Fórum, Keszthely, CD kiadvány.

MÁTYÁS K, **POCZAI P**, CERNÁK I, SZABÓ I, TALLER J (2009) A hévízi tündérrózsa fajok monitorozása és populációgenetikai vizsgálata. XV. Ifjúsági Tudományos Fórum, Keszthely, CD kiadvány.

SZÜCS R, NÉMETH SZ, BUDAHÁZI A, **POCZAI P**, ZIJIE L., ORBAN L (2009) Két sziklahal faj (*Scorpaena porcus* és *S. scrofa*) elkülönítése molekuláris genetikai módszerekkel. XV. Ifjúsági Tudományos Fórum, Keszthely, CD kiadvány.

SZÜCS R, NÉMETH SZ, BUDAHÁZI A, **POCZAI P**, ZIJIE L, ORBÁN L (2009), A ponty pikkelymintázat öröklődése, és az abból következő néhány gyakorlati következmény újragondolása. *XXXIII. Halászati Tudományos Tanácskozás*, *Szarvas,* Abstract book.

POCZAI P (2010) Egy ígéretes új módszer, a filogenetikai hálózatok alkalmazásának lejtőségei a komplex evolúciós események rekonstrukciójában. In: MTA Veszprémi Területi Bizottsága Tudományos Előadások 2010, VSP Nyomda, Pétfürdő, pp. 39-46.

TALLER J, CERNÁK I, MÁTYÁS KK, **POCZAI P**, CSEH A (2010) Az ürömlevelű parlagfű (*Ambrosia artemisiidolia* L.) atrazin és ALS-gátló hatóanyagú herbicidekkel szembeni rezisztenciájának molekuláris genetikai vizsgálata. XX. Keszthelyi Növényvédelmi Fórum, 2010. jan. 27-29, Keszthely, pp. 151.

SZÜCS R, NÉMETH SZ, ZIJIE L, **POCZAI P**, ORBÁN L, BERCSÉNYI M (2010), Sziklahal fajok (Scorpaenidae, Teleostei) elkülönítése két molekuláris genetikai módszerrel. *XXXIV. Halászati Tudományos Tanácskozás, Szarvas,* Abstract book

5. Program, Software

NAGY S, CERNÁK I, **POCZAI P**, TALLER J (2008) PICcalc: An on-line polymorphic information content calculator. http://www.georgikon.hu/pic/english/default.aspx

APPENDIX 1

DNA extraction from fresh plant tissues

1. 50 mg of fresh young leaves were suspended in 1 ml of Solution I. The tissues were crushed and homogenized, for complete cell lyses and disruption, using a mixer mill (Retsch MM 301, Germany) and sterile steal beads for 10 min, 30 s^{-1}.
2. The lysated samples were centrifuged at room temperature for 5 min at 8,000 rpm.
3. The supernatant was discarded and the remaining pellet was dissolved in 300 µl of 20T-10E buffer and 20 µl 10% SDS solution was added. The samples were incubated for 15 min at 70˚C and gently shaken a few times.
4. After the incubation 150 µl of 7.5 M sodium acetate was added, and the samples were placed on ice for 1 h and centrifuged for 10 min, 15,000 rpm at 4˚C.
5. The flow through (approx. 400 µl) was transferred to new Eppendorf tubes and 500 µl of ice-cold isopropanol was added to precipitate DNA. The samples were kept at 4˚C for 30 min and then centrifuged for 5 min, 8,000 rpm at 4˚C.
6. The flow through was removed; DNA was recovered as a pellet and dissolved in 500 µl of TE buffer equal volume of CIA (chloroform:isoamyl-alcohol, 24:1) was added and the samples were shaken for 15 min at 300 rpm and then centrifuged for 15 min, 15,000 rpm at room temperature.
7. The upper phase was transferred to a new Eppendorf tube. 1 ml of 99.9% ethanol and 40 µl of ammonium acetate was added to the samples and then centrifuged for 10 min, 15,000 rpm at 4˚C.
8. The supernatant was removed and the pellet was washed again with 500 µl of 70% ethanol and centrifuged for 5 min, 15,000 rpm at 4˚C.
9. The tubes were placed upside down and air dried. Finally the DNA was dissolved in 200 µl TBE (Tris-Borate-EDTA pH 8.0) and stored at –20˚C.
10. a) The DNA concentration was measured with fluorodensitometric analysis (GeneTools, Syngene UK) after electrophoresis and ethidium-bromide post-staining using the GeneRuler 1 kb Ladder (Fermentas, Lithuania) separated on 1% agarose gel.
    b) Absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) was measured for each DNA sample using the NanoDrop 2000 (Thermo Fisher Scientific, USA) spectrophotometer. Sample concentration was calculated by the NanoDrop nucleic acid application

module using Beer's law, and assuming 50 ng cm/ml absorbance for dsDNA, A_{260}/A_{280} ratios averaged 1.79±0.12 SD.

11. Each sample was diluted to 20ng/µl final concentration.

Required Stock Solutions

Solution I: 50mM Tris-HCl, 50mM EDTA, 500mM NaCl, 15% sucrose, adjust pH 8.0.

20T-10E buffer: 20mM Tris-HCl, 10mM EDTA, adjust pH 8.0

TE buffer: 10mM Tris-HCl, 1mM EDTA, pH 8.0

DNA extraction from dried herbarium plant tissues with spin-columns

(Macherey-Nagel, NucleoSpin Plant II Kit)

1. 100 mg of dried plant tissue was suspended in 500 µl Buffer C1 for 48 hours and kept on room temperature.
2. The tissues were crushed and homogenized, for complete cell lyses and disruption, using a mixer mill (Retsch MM 301, Germany) and sterile steal beads for 10 min, 30 s^1.
3. Steal beads were removed and 12 µl of RNase A stock solution was added. Samples were incubated for 30 min at 60°C in a water bath.
4. The tubes were centrifuged at 11,000 × g for 5 min at room temperature. The clear supernatant was transferred to a fresh tube.
5. 300 µl of Buffer C4 and 200 µl of 95% ethanol were added. Tubes were inverted 2–4 times.
6. Spin-columns were placed into a 2-ml collection tube. Samples were applied to the spin columns and centrifuged at 11,000 × g for 1 min at room temperature.
7. Flow-through was discarded and centrifugation steps were repeated until all of the lysate had passed through the column.
8. 400 µl of Buffer CW was added to the spin-column and centrifuged at 11,000 × g for 1 min. Flow-through was discarded. This wash step removed additional contaminants.

9. 700 µl of Buffer C5was added to the spin-column and centrifuged at 11,000 × g for 1 min. The flow-through was discarded.
10. The spin-columns were reinserted into the collection tubes. 200 µl of Buffer C5 was added to the spin-columns and centrifuged at 11,000 × g for 2 min to remove any traces of Buffer C5 from the membrane filter. This was done since residual ethanol from Buffer C5 inhibits enzymatic reactions and must be removed completely by performing this centrifugation step.
11. The spin-columns were placed in a clean 1.5-ml microcentrifuge tube. Elute The DNA was eluted with 50–100 µl of Buffer CE (warmed to 70°C). Samples were also incubated at room temperature for 5 min and centrifuge at 11,000 × g for 1 min at room temperature.
12. Absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) was measured for each DNA sample using the NanoDrop 2000 (Thermo Fisher Scientific, USA) spectrophotometer. Sample concentration was calculated by the NanoDrop nucleic acid application module using Beer's law, and assuming 50 ng cm/ml absorbance for dsDNA, A_{260}/A_{280} ratios averaged 1.79±0.12 SD.

DNA extraction from dried herbarium plant tissues

The procedure is based Nan et al. (2003) and Doyle and Doyle (1987) with modifications:

1. Dried leaf material was kept in 800 µl of 2 × CTAB extraction solution (2% cetyl-trimethyl-ammonium bromide, 0,2% mercaptoethanol, 1.4 mM NaCl, pH 8.0) for 1 week.
2. Samples were ground in a Mixer Mill (Retsch MM 301, Germany) for 5 min at 30 1/s.
3. The samples were incubated for 1 hour at 64°C.
4. Subsequently, 600 µl of chloroform:isoamyl-alcohol (24:1 v/v) was added, and the mixture was shaken gently for 2 min and centrifuged at 8.000 rpm for 15 min at room temperature.
5. The supernatant was mixed with 2/3 volume ice-cold isopropanol. For accurate DNA precipitation the mixture was stored for 5 days at 4°C.
6. DNA was then recovered as a pellet by centrifugation at 15.000 rpm for 10 min and washed with 1 ml washing buffer (99.9 % ethanol, 10 mM ammonium acetate).
7. The supernatant was removed and the pellet was washed again with 400 µl 76 % ethanol, dried, then re-suspended in 50 µl TE buffer [10 mM Tris (hydroxymethyl)-aminomethane, 1 mM EDTA, pH 8.0].
8. DNA extraction was done for each sample in three repetitions, to allow the detection of sporadic contaminants, and to detect further changes caused by miscoding DNA lesions due to low numbers of template molecules.
9. Each extraction process included one extraction control, which did not contain any sample material, but was treated identically.

References

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of resh leaf tissue. Phytochemical Bulletin 19:11-15.

Nan P, Shi SH, Peng ST, Tian CJ, Zhong Y (2003) Genetic diversity in *Primula obconica* (Primulaceae) from central and south-west China as revealed by ISSR markers. Ann Bot 91: 329–333.

APPENDIX 2

Sequences of RAPD primer pairs used in the pilot studies.

Primer numer	Sequince 5'-3'	Primer number	Sequence 5'-3'
1	GGCATGACCTGT	36	5GGGACGATGGTG
3	CTGATGCGTGTG	39	GGCTAACCGATG
5	AGATGCAGCCAG	45	GTAAGGCGCATG
9	CTATGCCGACAC	47	CCATTCCGAGTC
13	GAATGCGACCAA	46	CCGTTAGCGTGC
16	GAATGGCACCTG	51	TCTTAGGCGGCT
17	ACCGATGCTGAG	55	TCCGCATACCAG
20	TGATGCCGCTGC	77	CCCACTAGACTC
21	GATGCGACGGTA	85	TGCCGTGAGACC
23	GTGAATGCGGAG	88	ACGGACGTCACC
26	TGCGATGCGAAC	90	TCACCAGCCAGG
29	CCATGCGGAGTC	96	CTACGGCTTCGG
33	GTCATGCGACGA	97	CGCGAATTCC
35	GGCTTATGCCGT		

Primer combinations: 9+85; 13+29; 16+26; 17+77; 20+1; 23+47; 33+39; 35+88; 36+46; 45+51; 55+5; 91+97; 96+3; 96+21; 21+1; 13+47; 51+50; 90+97; 13+35; 33+13.

Sequences of SCoT primers used in the pilot studies.

Primer name	Primer Sequence 5'-3'	GC%	Annealing Temp. (ºC)	*H*	*PIC*
SCoT Primers					
1. SCOT01	CAACAATGGCTACCACCA	50	50	0,31	0,25
2. SCOT02	CAACAATGGCTACCACCC	56	50	0,35	0,31
3. SCOT04	CAACAATGGCTACCACCT	50	50	0,29	0,26
4. SCOT05	CAACAATGGCTACCACGA	50	50	0,28	0,25
5. SCOT11	AAGCAATGGCTACCACCA	50	50	0,3	0,26
6. SCOT13	ACGACATGGCGACCATCG	61	50	0,4	0,38
7. SCOT14	ACGACATGGCGACCACGC	67	50	0,3	0,26
8. SCOT31	CCATGGCTACCACCGCCT	67	50	0,32	0,29
9. SCOT33	CCATGGCTACCACCGCAG	67	50	0,26	0,24
10. SCOT34	ACCATGGCTACCACCGCA	61	50	0,34	0,31
11. SCOT35	CATGGCTACCACCGGCCC	72	50	0,37	0,31
12. SCOT36	GCAACAATGGCTACCACC	56	50	0,42	0,35

APPENDIX 3

Representative voucher specimens deposited at the Georgikon Faculty, University of Pannonia (GF-UP).

Taxon — Voucher specimens and locality.

Solanum nigrum — Hungary, Keszthely — UPG0013.1, UPG0013.2, UPG0013.3, UPG0013.4, UPG0013.5, UPG0013.6, UPG0013.6, UPG0013.7, UPG0013.8, UPG0013.9, UPG0013.10;
Hungary, Szolnok — UPG0015.1, UPG0015.2, UPG0015.3, UPG0015.4, UPG0015.5, UPG0015.6, UPG0015.7;
Croatia, Pula — UPG0016.1, UPG0016.2, UPG0016.3, UPG0016.4, UPG0016.5

Detailed information of intron targeting primers used in this study

TABLE 1. Characterization of 29 intron targeting primers in *Solanum nigrum*. Shown for each primer are the forward and reverse sequences, size of the original fragment (bp), annealing temperature (T_a), reported putative function acquired by searching NCBI database with blastx with expected value < 10^{-20}, allele number and size range in *S. nigrum* L., and GenBank accession numbers. All values are based on 30 samples of the F_1 potato population and 22 samples representing natural populations from Hungary (Keszthely and Szolnok), and Croatia (Pula).

Primer	Sequences	Size	T_a	Putative function	No. of alleles and size ranges	GenBank
Adk-242	F: TGCTTTTAAAGTCGCACCA R: TTATATCCGGAGCATGTCCAC	242	55	Adenylate kinase gene	4 (112-242)	AJ276864.1
Adk-795	F: GCATGGTTCTTTCCTTCCTG R: TGGGCCAGGAATTTTGCTATC	795	54	Adenylate kinase gene	4 (805-385)	AJ276864.1
Cat-232	F: AGGAGGCGGATCTAGCCTTA R: TGTCAAGAAAGGGGTGTCGT	232	55	Potato catalase gene	2 (210-232)	Z37106.1
Cat-260	F: TGACAACAAATGCTGGTGGT R: AAGGTGGCAAGCTTCTCAAT	260	53	Potato catalase gene	3 (260-310)	Z37106.2
GPSS-275	F: CTTTTGATGGGGGCAGATTA R: CAGCTTCCTGTCAGCATCAG	275	53	ADP-glucose pyrophosphorylase small-subunit gene	3 (275-125)	L36648.1

GPSS-943	F: TCATTGGTGAAGGTTGTGTGA R. ACCACGGAATGGTGAATCTT	943	53	ADP-glucose pyrophosphorylase small-subunit gene	7 (1043-150)	L36648.2
INHWI-509	F: TGAAACTCTCTTGGCACGAA R: TTCTGGCCACCTTTGTTTTC	509	54	Wound-inducible proteinase inhibitor I gene	2 (509-205)	M17108.1
INHWI-545	F: TCAAGTTTGCTCACATCTTGT R: TCGTGCCAAGAGAGTTTCAA	545	54	Wound-inducible proteinase inhibitor I gene	4 (685-425)	M17108.1
InvG-220	F: ACAGGAATCACACCTGCACA R: TCTGCACCCTTAAGTCCACA	220	54	Invertase gene	3 (280-150)	AJ133765.1
InvG-262	F: TTCTCATGTGCTCAGATGCT R: GAGGGCTTGTACATTGACTTCA	262	53	Invertase gene	3 (182-262)	AJ133765.1
InvG-393	F: TGGTTACCATTTTCAGCCAGA R: CCATTGAAATACATTGGTGCTG	393	55	Invertase gene	3 (320-393)	AJ133765.2
LBr-G9	F: TGGATCTGAAGATGGCACTG R: TTGCTCTCAAATCCCACACA	652	55	Transducin family protein	3 (1100-652)	CO267873.1
LBr-4D6	F: GAGTATTCATTCGGGCTTGG R: CTCTACCGACCCGTAGCAAG	196	54	Plastidic ATP/ADP-transporter protein	4 (597-196)	CO267884.1
PatI-433	F: TCAAGCTCGTCATTCACAAAA R: TCAGACGCATCATCCATTTC	433	54	Potato patatin class I. gene	3 (510-285)	M18880
PatI-838	F: CGAACATGGCCCTCATATTT R: TGCACACGAGTTTCTCCAAG	838	53	Potato patatin class I. gene	5 (938-145)	M18880
Pat-In3	F: CAGAAAGTTGCCATCTCAAGC R: GCTGCTGCTGTGGAATAACA	581	53	Potato patatin gene Intron 3	4 (595-220)	X03932.1
Poni1a-718	F: GGTGGTGGTGGTAGCTCAGT R: CCCARRGGCATTAACACTCC	718	55	Potato membrane protein	3 (850-718)	AJ309301.1
Poni1a-442	F: TTTGCCTCGGAACTCTTCAG R: GCCTCAGAAGCAAAGCAAAT	442	54	Potato membrane protein	3 (510-295)	AJ309301.2
Ry1-In3	F: AATGCAGAAGGTGCAACGAT R: TGGGCGAAATTTCATTAACA	199	54	Ry1 resistance gene-like function protein intron 3	2 (450-199)	AJ300266
Ry1-In4	F: TCGAAAAATTCTCAAATGCAAA R: GATTGCTTCGATAGCCTTGG	477	55	Ry1 resistance gene-like function protein intron 4	6 (125-590)	AJ300267
Ry1-In5	F: CCAGCAGAGTTCACTGTTTCA R: GTTGCACAGCTGCTGAGAT	481	54	Ry1 resistance gene-like function protein intron 5	2 (481-501)	AJ300268
Ry1-In6	F: GCTCTCGTCTCCACTTCTGC R: AACTCCTCAGCAACTGCACA	741	56	Ry1 resistance gene-like function protein intron 5	7 (751-110)	AJ300269
S2-317	F: CGGCCAGTTACAATTCTGC R: AATCCAGTGGTGGTCCAGAG	202	55	Self-incompatibility locus linked stylar Rnase gene	6 (455-202)	X62727

STAC1-226	F: GTCTTCCCGCTTTCAAAGAT R: TCAGCAAGGCAAAACATGAG	226	53	1-Aminocyclo-propane-1-carboxylate synthase gene	4 (135-295)	Z27235.1
Suc16-321	F: TCACCGCAATGAGATACTGC R: TATCCCTTTTCCGTGGCTTT	321	54	Sucrose synthase gene	5 (270-385)	U24087.1
Suc16-349	F: TGACGTTGAGAATGACGAACA R: CCAACCTTGCCATTGTGAAT	349	55	Sucrose synthase gene	4 (605-275)	U24087.1
UBQ-627	F: TCTCAATTGCCTTCAATTTCTC R: TCCGGTGAGAGTTTTCACAA	627	53	Polyubiquitin gene	6 (715-520)	U26831.1
Ure-242	F: TGCTTTTAAAGGTCGCACCA R: TTATATCCGGAGCATGTCCAC	242	56	Ure gene for Urease	2 (242-257)	AJ276865.1
Ure-271	F: GAGCAGCCACGAGATTTGA R: CACAAATCAATGCCCAAGC	271	55	Ure gene for Urease	2 (271-451)	AJ276865.2

TABLE 2. Results of initial primer screening in potato (*S. tuberosum*) and black nightshade (*S. nigrum*). Number of alleles (*A*), observed heterozygosity (H_o), and expected heterozygosity (H_E) for each population are shown.

	F1 potato (*N* = 30)			**Keszthely (*N* = 10)**		
Locus	**A**	**H_o**	**H_E**	**A**	**H_o**	**H_E**
Adk-242	4	0.364	0.723	4	0.500	0.708
Adk-795	4	0.545	0.636	4	0.667	0.583
Cat-232	2	0.500	0.375	2	0.833	0.486
Cat-260	3	0.273	0.455	3	0.273	0.247
GPSS-275	3	0.636	0.498	3	0.636	0.749
GPSS-943	4	0.667	0.597	3	0.455	0.567
INHWI-509	2	0.333	0.444	2	0.167	0.153
INHWI-545	4	0.667	0.597	4	0.545	0.636
InvG-220	3	0.455	0.567	3	0.428	0.562
InvG-262	3	0.451	0.684	3	0.500	0.640
InvG-393	3	0.429	0.571	3	0.400	0.480
LBr-G9	3	0.636	0.749	3	0.273	0.455
LBr-4D6	4	0.667	0.583	4	0.455	0.567
PatI-433	3	0.818	0.775	3	0.714	0.670
PatI-838	3	0.833	0.569	2	0.833	0.486
Pat-In3	4	0.667	0.583	4	0.714	0.626
Poni1a-718	3	0.727	0.589	3	0.364	0.450
Poni1a-442	3	0.714	0.670	3	0.455	0.526
Ry1-In3	2	0.091	0.091	2	0.182	0.314
Ry1-In4	3	0.273	0.247	3	0.636	0.498
Ry1-In5	2	0.455	0.567	2	0.285	0.498

Ry1-In6	3	0.364	0.671	4	0.667	0.750
S2-317	4	0.667	0.625	2	0.285	0.244
STAC1-226	4	0.667	0.625	4	0.500	0.708
Suc16-321	2	0.143	0.337	3	0.333	0.653
Suc16-349	4	0.500	0.708	4	0.667	0.583
UBQ-627	3	0.285	0.255	3	0.500	0.568
Ure-242	2	0.200	0.500	2	0.200	0.180
Ure-271	2	0.300	0.420	2	0.285	0.244

	Szolnok (*N* = 7)			**Pula (*N* = 5)**		
Locus	**A**	**H_o**	**H_E**	**A**	**H_o**	**H_E**
Adk-242	4	0.545	0.636	4	0.545	0.636
Adk-795	4	0.667	0.583	4	0.545	0.636
Cat-232	2	0.500	0.375	2	0.833	0.486
Cat-260	3	0.200	0.500	3	0.200	0.500
GPSS-275	3	0.429	0.571	3	0.636	0.749
GPSS-943	3	0.451	0.684	3	0.429	0.571
INHWI-509	2	0.333	0.444	2	0.333	0.444
INHWI-545	4	0.667	0.625	4	0.500	0.708
InvG-220	3	0.400	0.340	3	0.100	0.255
InvG-262	3	0.428	0.540	3	0.200	0.180
InvG-393	3	0.285	0.255	3	0.425	0.532
LBr-G9	3	0.455	0.526	3	0.167	0.292
LBr-4D6	4	0.667	0.750	4	0.333	0.681
PatI-433	3	0.833	0.569	3	0.167	0.486
PatI-838	2	0.833	0.486	2	0.833	0.486
Pat-In3	4	0.545	0.636	4	0.500	0.542
Poni1a-718	3	0.636	0.749	3	0.116	0.278
Poni1a-442	3	0.451	0.684	3	0	0
Ry1-In3	2	0.100	0.255	2	0	0
Ry1-In4	3	0.451	0.684	3	0.333	0.500
Ry1-In5	2	0.273	0.247	2	0.500	0.611
Ry1-In6	4	0.667	0.583	4	0.284	0.791
S2-317	2	0.200	0.180	2	0.273	0.247
STAC1-226	4	0.636	0.749	4	0.364	0.723
Suc16-321	3	0.273	0.247	3	0.182	0.274
Suc16-349	4	0.667	0.583	4	0.667	0.583
UBQ-627	3	0.285	0.255	3	0.182	0.512
Ure-242	2	0.273	0.247	2	0.091	0.091
Ure-271	2	0.285	0.255	2	0	0

APPENDIX 4

DNA fragment purification from agarose gel with spin-columns

(Machery-Nagel NucleoSpin Extract II Kit)

1. Desired fragments from 1% agarose gels were excised with a scalpel. The gel slice was excised carefully to minimize the gel volume. UV exposure time was also minimized to avoid damaging the DNA.
2. The weight of the gel slice was determined and transferred into a clean tube.
3. For each 100 mg of agarose gel 200 µl Buffer NT was added.
4. Samples were incubated from 5-20 min at 50 °C. Samples were vortexed briefly every 2-3 min until the gel slice is completely resolved.
5. The spin-column was placed into a collection tube (2 ml) and samples were loaded.
6. Tubes were centrifuged for 1 min at 11,000 × g. Flow-through was discarded and the spin-column was placed back to the collection tube.
7. To wash the silica membrane 700 µl Buffer NT3 was added to the spin-column, than the tubes were centrifuged for 1 min at 11,000 × g. Flow-through was discarded and the column was placed back into the collection tube.
8. Tubes were centrifuged for 2 min at 11,000 × g to completely remove Buffer NT3.
9. Spin-columns were placed into new 1.5 ml tubes. 15-50 µl Buffer NE was added and incubated at room temperature for 3 min, than centrifuged for 1 min at 11,000 × g.
10. Absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) was measured for each DNA sample using the NanoDrop 2000 (Thermo Fisher Scientific, USA) spectrophotometer. Sample concentration was calculated by the NanoDrop nucleic acid application module using Beer's law, and assuming 50 ng cm/ml absorbance for dsDNA, A_{260}/A_{280} ratios averaged 1.79±0.12 SD.

DNA fragment purification from agarose

1. Desired fragments from 1% agarose gels were excised with a scalpel. The gel slice was excised carefully to minimize the gel volume. UV exposure time was also minimized to avoid damaging the DNA.
2. The weight of the gel slice was determined and transferred into a clean tube
3. For each 100 mg of agarose gel 200 µl of TE buffer was added.
4. Agarose was melted at 65°C for 30 min in a water bath. Samples were vortexed briefly every 2-3 min until the gel slice is completely resolved.
5. Equal volume of phenol was added to the melted agarose solution, buffered to pH 8.0 with 0.1M Tris-HCl and mixed by vortexing or shaking for 10 min. Note: the longer the mixing the cleaner it gets.
6. Tubes were centrifuged for 15 min at 10,000 rpm.
7. The upper aqueous phase was collected into a new clean 1.5 ml tube. Note: Do not recover the white internal precipitate.
8. The aqueous fraction was precipitated with 100-200 µl of ice-cold isopropanol and 15-30 µl 3M sodium acetate.
9. Tubes were centrifuged for 10 min at 10,000 rpm. The flow-through was removed and the 500 µl of TE buffer and 200 µl cold 70% ethanol was added.
10. Tubes were centrifuged for 15 min at 10,000 rpm. The flow-through was removed.
11. Tubes were place upside-down and air dried.
12. Re-suspend samples in 50 µl TE buffer.
13. Absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) was measured for each DNA sample using the NanoDrop 2000 (Thermo Fisher Scientific, USA) spectrophotometer. Sample concentration was calculated by the NanoDrop nucleic acid application

module using Beer's law, and assuming 50 ng cm/ml absorbance for dsDNA, A_{260}/A_{280} ratios averaged 1.79±0.12 SD.

11. Samples were stored at -70°C until further downstream applications.

APPENDIX 5

PCR product purification with spin-columns

(Machery-Nagel NucleoSpin Extract II Kit)

1. 1 volume of PCR product was mixed with 2 volumes of Buffer NT.
2. Spin-columns were placed onto collection tubes (2 ml) and samples were loaded.
3. Tubes were centrifuged for 1 min at 11,000 × g. Flow-through was discarded and columns were placed back into the collection tube.
4. 700 µl Buffer NT3 was added to the spin-columns and centrifuged for 1 min at 11,000 × g. Flow-through was discarded and spin-columns were placed back into the collection tubes.
5. The silica membrane was dried with an additional centrifugation step for 2 min at 11,000 × g to remove Buffer NT3 completely.
6. DNA was eluted by placing the spin-columns into a new 1.5 ml tube and adding 15-50 µl Buffer NE. Samples were incubated at room temperature for 1 min, than centrifuged for 1 min at 11,000 × g.
7. Absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) was measured for each DNA sample using the NanoDrop 2000 (Thermo Fisher Scientific, USA) spectrophotometer. Sample concentration was calculated by the NanoDrop nucleic acid application module using Beer's law, and assuming 50 ng cm/ml absorbance for dsDNA, A_{260}/A_{280} ratios averaged 1.79±0.12 SD.

APPENDIX 6

Preparation of fresh competent cells

This protocol is recommended for the production of fairly high efficiency competent cells for reliable cloning of single inserts from digested genomic DNA in library construction experiments.

1. Grow overnight culture of desired strain in 10 ml of LB broth (without antibiotic), 2 days before the intended use of the cells.
2. Dilute 1.5 ml of the overnight culture into 40 ml of LB broth preheated to 37°C.
3. Shake at 37°C until the OD600 reaches 0.4-0.6 (about 2.5-3.0 h).
4. Transfer the cells to a 50 ml centrifuge tube (e.g., Corning) and chill on ice for 20 min.
5. Centrifuge the cell suspension for 15 min at 3000 rpm at 4°C.
6. Carefully discard the supernatant and re-suspend the pellet by gently pipetting 20 ml of sterile, ice-cold 50 μM $CaCl_2$. Use the tip of the pipette to gently re-suspend the cells.
7. Chill on ice for 20 min.
8. Centrifuge the cell suspension for 15 min at 3000 rpm at 4°C.
9. Carefully discard the supernatant and re-suspend the pellet by gently pipetting 4 ml of sterile, ice-cold 100 μM $CaCl_2$. Use the tip of the pipette to very gently re-suspend the cells.
10. Place on ice and keep in the refrigerator for use next morning.

Stock Solutions

50 mM $CaCl_2$

STOCK	**100ml**	**200ml**	**300ml**	**400ml**	**500 ml**
1.0 M $CaCl_2$	5 ml	10 ml	15 ml	20 ml	25 ml
ddH_2O	95 ml	190 ml	285 ml	380 ml	475 ml

100 mM $CaCl_2$

STOCK	**100 ml**	**200 ml**	**300 ml**	**400 ml**	**500 ml**
1.0 M $CaCl_2$	10 ml	20 ml	30 ml	40 ml	50 ml
ddH_2O	90 ml	180 ml	270 ml	360 ml	450 ml

Preparation of fresh competent cells with Transform-Aid Kit

(TransformAid Bacterial Transformation Kit, Fermentas)

1. Inoculate a LB plate with bacteria using streak plate method. Incubate the plate overnight at 37°C.
2. Pre-warm culture tubes containing the TransformAid C-Medium (1.5 ml for every 2 transformations) at 37°C.
3. Add small portion of bacterial culture (4x4 mm) from overnight LB plate using an inoculating loop into the pre-warmed C-Medium (1.5 ml).
4. Suspend the culture by gentle mixing and incubate the tubes in a shaker at 37°C for 2 hours. Note: *The colonies on LB plates can be stored at 4°C and used for inoculating fresh cultures within 10 days*.

Preparation of frozen competent cells

This protocol is recommended for the production of large amounts of competent cells of medium efficiency for rapid sub-cloning of single inserts.

1. Grow overnight culture of desired strain in 5 ml of LB broth (without antibiotic).
2. Dilute the overnight culture 1:100 with LB broth (without antibiotic) and shake at 37°C until the OD600 reaches 0.3-0.4.
3. Transfer the cells to 250 ml centrifuge bottles and chill on ice for 10 minutes.
4. Centrifuge the cells for 7 min at 3500 rpm at 4°C.
5. Carefully discard the supernatant and re-suspend the pellet by gently pipetting 5 ml of sterile, ice-cold 10 μM $MgCl_2$. After cells are re-suspended, add an additional 120 ml of 10 μM $MgCl_2$.
6. Centrifuge the cells for 7 min at 3500 rpm at 4°C.

7. Carefully discard the supernatant and re-suspend the pellet by gently pipetting 5 ml of sterile, ice-cold 50 µM $CaCl_2$, 20% glycerol. After the cells are re-suspended, add an additional 5 ml of 50 µM $CaCl_2$, 20% glycerol.
8. Place on ice for at least 1 h.
9. Transfer 400 µl aliquots of cells to individual, sterile 500 µl microfuge tubes.
10. Quick freeze cells in a dry ice/ethanol bath (or in ethanol at -80°C) and store at -80°C until use.

Stock Solutions

10 mM $MgCl_2$

STOCK	**100 ml**	**200 ml**	**300 ml**	**400 ml**	**500 ml**
1.0 M $MgCl_2$	1 ml	2 ml	3 ml	4 ml	5 ml
ddH_2O	99 ml	198 ml	297 ml	396 ml	495 ml

50 mM $CaCl_2$, 20 % glycerol

STOCK	**100 ml**	**200 ml**	**300 ml**	**400 ml**	**500 ml**
1.0 M $CaCl_2$	5 ml	10 ml	15 ml	20 ml	25 ml
Glycerol	20 ml	40 ml	60 ml	80 ml	100 ml
ddH_2O	75 ml	150 ml	225 ml	300 ml	375 ml

APPENDIX 7

Sticky-End Cloning Protocol

(ClonJET PCR Cloning Kit, Fermentas)

- For cloning PCR products with 3'-dA overhangs generated by *Taq* DNA polymerase or enzyme mixtures containing *Taq* DNA polymerase.

1. Set up the blunting reaction **on ice**:

Component	**Volume**
2X Reaction Buffer	10 µl
Purified PCR product/other sticky-end DNA fragment	1 µl
Water, nuclease-free	to 17 µl
DNA Blunting Enzyme	1 µl
Total volume	18 µl

Vortex briefly and centrifuge for 3-5 s.

2. Incubate the mixture at 70°C for 5 min. Chill on ice.
3. Set up the ligation reaction **on ice**. Add the following to the blunting reaction mixture:

Component	**Volume**
pJET1.2/blunt Cloning Vector (50 ng/µl)	1 µl
T4 DNA Ligase	1 µl
Total volume	20 µl

Vortex briefly and centrifuge for 3-5 s to collect drops.

4. Incubate the ligation mixture at room temperature (22°C) for 5 min.

Note: *For PCR products >3 kb, ligation can be prolonged to 30 min. Ligation times longer than 30 min are not recommended and may decrease cloning efficiency*.

4. Use the ligation mixture directly for transformation. **Note:** *Keep the ligation mixture at -20^{o}C if transformation is postponed. Thaw on ice and mix carefully before transformation.*

Transformation

1. Pre-warm LB-Ampicillin agar plates (37°C for 20 min).
2. Prepare TransformAid™ T-Solution by mixing equal volumes of T-Solution (A) and T-Solution (B) (500 µl for every 2 transformations). Keep on ice.
3. Dispense 1.5 ml of fresh culture into a tube and spin down at maximum speed for 1 min.
4. Discard the supernatant and resuspend the pelleted cells in 300 µl TransformAid T-Solution. incubate on ice for 5 min.
5. Spin down the cells again for 1 min and then remove the supernatant.
6. Resuspend the cells in 120 µl of TransformAid T-Solution and incubate on ice for 5 min.
7. Dispense 1 µl supercoiled DNA (10-100 pg) or 2.5 µl ligation mixture (10-20 ng of vector DNA) into new microcentrifuge tubes and place on ice for 2 min.
8. Add 50 µl re-suspended cells to each tube containing DNA and incubate on ice for 5 min.
9. Plate the cells on pre-warmed LB-Ampicillin agar plates. Incubate the plates overnight at 37°c.

Additional stock solutions

Ampicillin stock solution (50 mg/ml)

1. Dissolve 2.5 g ampicillin sodium salt in 50 ml of deionized water.
2. Filter sterilize and store in aliquots at 4°C.

LB-ampicillin plates

1. Prepare LB-agar Medium (1 liter), weigh out:
   Bacto Tryptone 10 g,
   Bacto Yeast extract 5 g,
   NaCl 5 g. 2.

2. Dissolve in 800 ml of water, adjust pH to 7.0 with NaOH and add water to 1000 ml.
3. Add 15 g of agar and autoclave.
4. Allow the medium to cool to 55°C.
5. Add 2 ml of ampicillin stock solution (50 mg/ml) to a final concentration of 100 μg/ml.
6. Mix gently and pour plates.

APPENDIX 8

Plasmid isolation with spin-columns

(Macherey-Nagel, NuceloSpin Plasmid Kit)

1. Grow 1-2 ml overnight culture in LB broth at 37°C.
2. Use 1 ml of saturated *E. coli* LB culture, pellet cells in a standard benchtop microcentrifuge for 30 s at 11,000 x *g*. Discard the supernatant and remove as much of the liquid as possible.
3. Add 250 µL Buffer A1. Re-suspend the cell pellet completely by vortexing or pipetting up and down. Make sure no cell clumps remain before addition of Buffer A2!
4. Add 250 µl Buffer A2. Mix gently by inverting the tube 6 – 8 times. Do not vortex to avoid shearing of genomic DNA. Incubate at room temperature for up to 5 min or until lysate appears clear.
5. Add 300 µl Buffer A3. Mix thoroughly by inverting the tube 6 – 8 times. Do not vortex to avoid shearing of genomic DNA!
6. Centrifuge for 5 min at 11,000 x *g* at room temperature. Repeat this step in case the supernatant is not clear!
7. Place a spin-column in a collection Tube (2 ml) and decant the supernatant from step 3 or pipette a maximum of 750 µl of the supernatant onto the column. Centrifuge for 1 min at 11,000 x *g*. Discard flow-through and place the spin-column back into the collection tube. Repeat this step to load the remaining lysate.
8. Add 600 µL Buffer A4. Centrifuge for 1 min at 11,000 x *g*. Discard flow-through and place the spin-column back into the empty collection tube. Note: *If plasmid DNA is prepared from host strains containing high levels of nucleases (e.g., HB101 or strains of the JM series), it is strongly recommended performing an additional washing step with 500 µL Buffer AW preheated to 50 °C and centrifuge for 1 min at 11,000 x* g *before proceeding with Buffer A4. Additional washing with Buffer AW will also*

increase the reading length of DNA sequencing reactions and improve the performance of critical enzymatic reactions.

9. Centrifuge for 2 min at 11,000 x *g* and discard the collection tube.
10. Place the spin-column in a 1.5 ml microcentrifuge tube (not provided) and add 50 µl Buffer AE. Incubate for 1 min at room temperature. Centrifuge for 1 min at 11,000 x *g*.

Plasmid isolation protocol

1. Grow 1 ml overnight culture in LB broth at 37°C.
2. Harvest cells by centrifuging entire culture in a 15 ml centrifuge tube for 5 min at full speed in a table-top centrifuge (1300-1500 x g). Discard supernatant.
3. Re-suspend cell pellet thoroughly by vortexing before adding 200 µl of solution I containing 5mg/ml lysozyme (add lysozyme within 1 h of use). Vortex and leave at room temperature for 5 min. It is easier to re-suspend cells if they are vortexed before adding the lysozyme mix.
4. Add 400 µl of solution II, mix gently (no vortex), and incubate 10 min on ice (solution should be clear).
5. Add 300 µl of solution III, mix gently (no vortex), and incubate 15 min on ice.
6. Centrifuge 15 min at full speed in table-top centrifuge; pour off supernatant into 1.5 ml microfuge tube.
7. Add 600 µl ice-cold isopropanol; mix and leave at -20°C for 1 h or at -70°C for 30 min. Centrifuge 5 min at full speed in microfuge (~12,000 rpm); drain and dry tube. Re-dissolve pellet in 190 µl dH2O. It may be placed on a vortex for 45 min, but use gentle vortexing.
8. Add 5 µl of 1 mg/ml RNAse A and 5 µl of 500 U/ml RNAse T1. Incubate at 37°C (or RT) for 15 min.
9. Add 10 µl of 5 mg/ml Proteinase K. Incubate at 37°C (or RT) for 20 min.
10. Extract with 200 µl phenol [or 200 µl chloroform-isoamil-alcohol (24:1)].
11. Centrifuge for 4 min at full speed in microfuge (~12,000 rpm). Transfer aqueous (upper) phase to new microfuge tube.
12. Add 100 µl 7.5 M NH4OAc to precipitate the DNA.

13. Add 800 µl ice-cold absolute EtOH; mix gently and incubate at -80°C for 30 min. Centrifuge 5 min at full speed in microfuge and pour off the supernatant.
14. Wash pellet with 1 ml 75% EtOH; centrifuge 4 min in microfuge. Pour off supernatant and dry tube in vacuum desiccator (for 20-30 min).
15. Dissolve pellet in 50 µl TE-8.0.

Stock Solutions

Solution I: 25 mM Tris-8.0, 10 mM EDTA, 50 mM glucose

STOCK	**10 ml**	**20 ml**	**30 ml**	**40 ml**	**50 ml**
1.0 M Tris-8.0	250 µl	500 µl	750 µl	1000 µl	1250 µl
0.5 M EDTA-8.0	200 µl	400 µl	600 µl	800 µl	1000 µl
Glucose	90 mg	180 mg	270 mg	360 mg	450 mg

NOTE: Solution I may be prepared as a 10X stock solution and stored -20°C in small aliquots for later use. Before using: thaw, dilute, and add lysozyme.

Solution II: 0.2 M NaOH, 1.0% SDS

STOCK	**100 ml**	**200 ml**	**300 ml**	**400 ml**	**500 ml**
1.0 M NaOH	20 ml	40 ml	60 ml	80 ml	100 ml
20% SDS	5 ml	10 ml	15 ml	20 ml	25 ml

Solution III: 3 M KOAc, pH 5.5

Dissolve 29.5 g potassium acetate in 60 ml dH_2O. Add enough glacial acetic acid to bring pH to 5.5 (approx. 11 ml). Bring final volume to 100 ml.

PCR amplification of inserts from bacterial cultures

1. Scrape a fresh single colony from a culture plate with a toothpick, or use 2 µl of an overnight culture or 2 µl of a glycerol stab.
2. Suspend in 50 µl of TTE buffer in a 0.5 ml microfuge tube.
3. Incubate at 95°C for 10 min to produce bacterial lysate.
4. Spin down bacterial debris for 5 min and use 2.5 µl of the supernatant for PCR amplification reaction. This lysate may be kept at 4°C for further uses.

TTE buffer

STOCK [FINAL]	**25 ml**	**100 ml**
ddH_2O	24.15 ml	96.6 ml
Triton X - 100 1 %	0.25 ml	1.0 ml
1 M Tris HCl - 8.5 20 mM	0.50 ml	2.0 ml
0.5 M EDTA - 8.0 2 mM	0.10 ml	0.4 ml

Sterilize and aliquot into 1.5 ml tubes or 2 ml Sarsted tubes. Store at 4°C.

APPENDIX 9

Location set (simple sites) file for GenGIS

Site Id,Site Name,Latitude,Longitude,Geographic Region,Species
S.aviculare_A,S.aviculare_A,-38.16,146.12,Melbourne SA,Aviculare
S.aviculare_B,S.aviculare_B,-36.32,140.32,Fairview CP,Aviculare
S.aviculare_C,S.aviculare_C,-35.20,148.21,Meadow Ck,Aviculare
S.laciniatum_A,S.laciniatum_A,-36.65.,140.45,Robe SA,Laciniatum
S.laciniatum_B,S.laciniatum_B,-38.15.,145.15,Melbourne,Laciniatum
S.laciniatum_C,S.laciniatum_C,-35.60.,150.05,Nelligen,Laciniatum
S.vescum_A,S.vescum_A,-37.49,147.3,Howe Ran,Vescum
S.vescum_B,S.vescum_B,-34.25,150.3,Near Sydney,Vescum
S.vescum_C,S.vescum_C,-31.61,152.6,Kendall,Vescum
S.linearifolium_A,S.linearifolium_A,-35.10,149.27,Lake George,Linearifolium
S.linearifolium_B,S.linearifolium_B,-37.10,148.27,Stockdale,Linearifolium
S.linearifolium_C,S.linearifolium_C,-30.12,149.27,Yarrie Lake,Linearifolium
S.multivenosum,S.multivenosum,-5.41,143.50,Enga,Multivenosum
S.capsiciforme_A,S.capsiciforme_A,-33.40,135.76,Eyre,Capsiciforme
S.capsiciforme_B,S.capsiciforme_B,-33.40,118.26,South of Kukerin,Capsiciforme
S.simile_A,S.simile_A,-32.01,117.41,Location,Simile
S.simile_B,S.simile_B,-37.01,140.45,Near Glen Groy CP,Simile
S.simile_C,S.simile_C,-35.01,141.4,Big Desert,Simile
S.symonii_A,S.symonii_A,-33.51,121.54,Pink Lake,Symonii
S.symonii_B,S.symonii_B,-28.51,114.54,Northampton,Symonii
S.symonii_C,S.symonii_C,-30.51,115.54,Location,Symonii