J Plant Biotechnol (2018) 45:236-241
DOI:https://doi.org/10.5010/JPB.2018.45.3.236

ISSN 1229-2818 (Print)
ISSN 2384-1397 (Online)

Research Article

Evaluating phylogenetic relationships in the Lilium family using the ITS marker

Sina Ghanbari · Barat Ali Fakheri · Mohammad Reza Naghavi · Nafiseh Mahdinezhad

Received: 19 June 2018 / Revised: 23 July 2018 / Accepted: 23 July 2018
ⓒ Korean Society for Plant Biotechnology

Abstract Lilium is a perennial bulbous plant belonging to the liriotypes genus. Our aim was to study the phylogenetic relationships of the Lilium family. Two varieties of *Lilium ledebourii*, 44 varieties of the gene bank, and one variety from the Tulipa family served as the out group. In order to study the diversity between lilium masses, ITS regions were used to design the marker. The results showed that the guanine base is the most abundant nucleotide. Relatively high conservation was observed in the ITS regions of the populations (0.653). Phylogenetic analysis showed that sargentiae and hybrid varieties are older than other varieties of the Lilium family. Also, the location of *L. ledebourii* varieties (Damash and Namin) was identified in a phylogenetic tree by using the ITS marker. Overall, our research showed that ITS molecular markers are very suitable for phylogenetic studies in the Lilium family.

Keywords Molecular phylogeny, Lilium, ITS marker, Haplotype

Introduction

Lilium is a bulbous, perennial plant consisting of almost 100 species that are found mainly in the northern hemisphere (McRae 1998). The center of this species is in Southwest Asia and the Himalayas (Comber 1949). The lilium genus is divided into 5 to 11 groups (Branova 1988; Comber 1949). Branova (1998) divided the lilium genus into 11 groups. Based on this division, European lilies were categorized into four groups. The identification and classification of species is one of the most important parts of taxonomy and biological studies. The correct identification of species provides useful information about any living organism, such as its ecological, physiological, biochemical roles, social benefits and losses (Crous et al. 2009). The taxonomy consists of the classification and naming of parts, and aims at grouping living organisms into natural units that not only allows for the description of biodiversity and species identification, but also reflect kin relationships among different groups of living organisms (Crous et al. 2009; Aezanlou 2008). Taxonomy becomes controversial when molecular techniques are introduced and used. Particularly, combining the concept of phylogenetic species with DNA data results to the description of a new wave of new species that include a number of sibling species or twin species. It is impossible to identify twin species based on morphological characteristics (Hebert et al. 2003; Taylor et al. 2000; Davari et al. 2012; Termorshuizen and Arnolds 1997). Problems with morphology-based identification has caused taxonomy specialists to use molecular methods in plants. However, these methods have advantages such as identification in very small amounts, decrease in the required time of fraud detection, identification of species, subspecies and even lower, and accuracy in sample identification. DNA barcoding has become an important part of ecological research over the past 15-20 years, and has created a new attitude towards biodiversity and differences between plant groups (Hebert et al. 2003; Frezal and Leblois 2008). DNA barcoding is a very fast, accurate and automatable way for species (through) using a short and standard genetic area as inter-species indication that is a practical and reliable method requiring no high expertise. DNA barcoding is used as a tool for precise identification of species in ecological studies and inter-species variation (Hebert et al. 2003; Frezal

S. Ghanbari (✉) · B. A. Fakheri · N. Mahdinezhad
Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Zabol, Zabol, Iran
e-mail: sina_qanbari@yahoo.com or sina.qanbari@gmail.com

M. R. Naghavi
Department of Agronomy and Plant Breeding, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

S. Ghanbari · M. R. Naghavi
Plant Bank, Iranian Biological Resource Center (IBRC), ACECR, Tehran, Iran

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

and Leblois 2008). The ITS region has a large number of versions inside the genome with suitable variety among different groups of living organisms. It also has conserved areas for designing primers and various genes with different evolutionary velocity and is considered as reference and base for barcoding. ITS markers are based on internal transcribed spacer regions and part of the chromosomal regions that are located next to the regions of ribosomal genes, can be used as universal primers to evaluate the intra-and inter-species phylogenetic relationships. The ITS regions are among the regions conserved during evolution (Brasileiro et al. 2004; Gultepe et al. 2010). ITS regions are used to investigate new strains of pathogens, as well as kin relationships among organisms (Adams et al. 1998; Brasileiro et al. 2004; Joseph et al. 1999). The suitability of the ITS marker for phylogenetic studies is due to two-parent inheritance, ease of learning, intra-species sameness, interspecies variability and high copy numbers (Fang et al. 2002). A study investigated the interspecies diversity of *Erysiphe aquilegia* (Cunnington et al. 2004) and the diversity of the genus Fasciolidae (Adlard et al. 1993) through the sequencing of ITS regions and the results showed that the ITS marker was able to differentiate subspecies from each other. There are several molecular phylogenic studies on the Lilium genus, which include species belonging to the liriotypes genus. The ITS region is the most commonly-used nuclear marker (Dobohzat and Shinoda 1999). This region is used for monocots to explain the various aspects of taxonomic problems (Dizkiviei et al. 2010). Using the ITS marker, Ronsted et al. (2005) examined the 14 Lilium species. In addition, molecular phylogenetic studies, using the ITS marker, indicated that Fritillaria L., Namocharis Franch and Cardiocrinum are closely related to the Lilium genus (Dobohzat and Shinoda 1999; Nishikawa et al. 1999, 2001; Ronsted et al. 2005; Ikinci et al. 2006; Resetnik et al. 2007). Lilium genus can be considered as a valuable plant in corrective researches, thus plant breeding is necessary in order to domesticate and conserve the genetic reserves of this valuable plant. The aim of this research is to achieve genetic diversity status and structure among Lilium populations, as well as the possibility of differentiating, categorizing and determining species relationships and determining the location of *L. ledebourii*, as a rare and native plant of Iran, in the phylogenetic tree of the Lilium genus using the ITS molecular marker.

Materials and Methods

DNA was extracted from the bulbs of *L. ledebourii* varieties, namely Damash and Namin (GenBank accession numbers, respectively, KU679314, KU739031, Ghanbari et al. 2016). Genomic DNA extraction was performed according to the CTAB method (Doyle and Doyle 1987). The quality and quantity of DNA samples were determined using a spectrophotometric method at 260 and 280 nm wavelengths and 0.8% agarose gel electrophoresis. 20 µl PCR reaction, contained 2 µl of PCR buffer, 1 µl of $MgCl_2$, 1 µl of dNTP 1x mixture, 0.25 µl of each forward and reverse primer, 1 Taq unit, 1µl of template DNA and the reaction volume was made up to 20 microliter, using sterile deionized water. The marker used in this study was designed based on the ITS sequences found in the gene bank, using the Oligo Software (Forward: TCCTCCGCTTATTGATATGC; Reverse: GGAAGTAAAAGTCGTAACAAGG) and prepared by the Takapouzist Co. When the PCR reaction was performed, 5 µl of PCR solution with 5 µL loading buffer was poured into 1.5% agarose gel well containing TBE buffer. The samples were electrophoresed at 85V for 85 h. Then, the gel was stained for 30 min in 0.5 mg/l of ethidium bromide (EtBr), and replicated fragments were observed under UV light and gel document imaging was later carried out. The replicated products were sent to the MWG Co. (Germany) for sequencing, which was carried out based on both replicated fragments (forward and reverse). Sequences were extracted from the chromatograms of the company using Chromas ver. 2. The sequences were later received and Blast software confirmed that the fragment was lilium- specific. Sequences obtained from the ITS region of the Namin and Damash varieties of *L. ledebourii* samples and samples from the NCBI website (Table 1) were aligned using the Clustal W program. Thereafter, genetic indices were examined using MEGA, DNA SP, NETWORK 4.163 and HIV databases.

Table 1 GenBank accession numbers for DNA sequences used in this paper

HQ686065	HQ724820	HQ724819	HQ724817	HQ724816	HQ686068	KX865054
JF778866	HQ686064	HQ724818	AY616750	HQ724825	HQ724821	KX670792
KX865053	KX865056	KX865055	KR632775	KX865058	HQ686062	HQ724823
KX865057	HQ686073	HQ724827	HQ724822	AY616746	AY616749	AY616747
AF092516	EF042778	KU739031	AF092522	KX865068	AF092518	EF042779
AB035281	KX865067	AM292432	AM292423	AF092519	AB035280	AF092517
KU679314	KX495217	KX865062	KX865063	KU232889		

Results and Discussion

The results of analysis of the nucleotide sequences of the ITS gene in the Lilium family showed that the highest and lowest abundance rates belonged to the guanine base (32.9%) and adenine base (18.5%), respectively (Table 2).

When estimating nucleotide substitution in the Lilium family, high levels of pyrimidine substitution were obtained in such a way that this level obtained 29.65% and 22.87% for thymine-cytosine and cytosine-thymine conversion, respectively (Table 3). These levels were lower for purine bases and obtained as 15.89% and 9.05% for the adenine-guanine and guanine -adenine conversions, respectively. The results of this study are consistent with the results of other researchers who reported frequent cases of pyrimidine substitutions. These changes are probably as a result of cytosine methylation (Picoult et al. 1999).

After investigating the genetic indices of the ITS gene of the Lilium family, a total of 232 mutations were identified that were distributed differently throughout the genome. The number of polymorphic positions was observed in 206 positions, which indicates the process of positive selection of this gene (Table 4). The low level of nucleotide diversity is probably as a result of a sharp decline in the effective population of varieties belonging to this family and the confirmation of the risk and extinction of these varieties. The results of conserved DNA regions of the ITS gene in the Lilium family showed a 0.75 conservation threshold region, a minimum conservation length of 68 bases and sequence conservation of 0.653 (Table 5). These conserved regions include a small part of the sequence of the above gene, which indicates a different differentiation of this position as well as its susceptibility to nucleotide changes and mutations among different varieties, causing variability among varieties.

The numerical value of the dN / dS ratio of the Lilium family was 1:11 (Table 6), which indicates the positive selection of the ITS gene among the Lilium varieties during evolution. This type of selection has resulted to the emergence of new varieties on one hand and stabilization of better purification of their performance during the evolution, on the other hand, which is due to the conversion

Table 2 Abundance of nucleotides derived from a nucleotide sequence of ITS gene in the Lilium family

Nucleotide	A	T	C	G
Abundance	18.5	21	27.6	32.9

Table 3 Estimation matrix of the nucleotide substitution pattern of the ITS gene in the Lilium family

From/To	A	T	C	G
A	-	2.39	3.1	**15.89**
T	2.1	-	**29.65**	3.68
C	2.1	**22.87**	-	3.68
G	**9.05**	2.39	3.1	-

Each input may be substituted from one base (row) to another base (column). In this table, the percentage of transition mutations (purine-purine substitution, pyrimidine-pyrimidine substitution) and transversion mutations (purine-pyrimidine substitution and vice versa) are shown in bold and italics, respectively.

Table 4 Location of ITS gene polymorphism in the Lilium family

Population	S	H	Hd	Pi	Eta	K
Lilium	206	46	1.000	0.05675	232	34.62

S: The number of polymorphic positions, H: number of haplotypes, Pi: nucleotide diversity, Eta: total number of mutations, K: number of nucleotide differences between population or species (nucleotide divergence)

Table 5 Conserved DNA regions of the ITS gene in the Lilium family

Population	C	MWL	CT
Lilium	0.653	68	0.75

C: Sequence conservation, MWL: Minimum conservation Length, CT: Conservation threshold

Table 6 Identifying the natural selection process for the ITS gene in the Lilium family

Parameter	Numerical value
ds	1.9224
dn	2.1418
dn/ds*	1.11

* The numerical value of dn / ds represents the natural selection process

Table 7 Results obtained from the evaluation of the natural evolution of the ITS gene

Gene	Tajima's D	Fu's Fs
ITS	-1.258	-21.726

of non-coding gene regions to gene coding regions.

Neutrality tests, including Tajima's D and Fu's Fs, were calculated to identify any deviation from null hypothesis on the neutral evolution and identify the effects of natural selection on these genes in Lilium populations. Populations that have been affected by recent expansion or significant increase in the effective population size or the directional selection, generate negative and significant values of D and Fs. However, positive and significant values of D and Fs reflect the effects of genetic drift, genetic dilemma, or a balancing effect throughout the evolutionary history of the population. The results of this study showed that the values of D and Fs were negative (Table 7). The negative and positive results indicate a significant and a small difference between polymorphisms in terms of their frequency, respectively. The results of both neutrality tests indicate the effect of the recent expansion of the Lilium population in the world, or the effect of directional selection on this gene during evolution. In general, studies have shown that the Fu's Fs and Tajima's D tests are more effective for small-size and larger sized populations, respectively.

In this research, parsimony was used to plot the phylogenetic tree of the nucleotide sequences in 46 varieties and one sequence of the tulipa family, as an outside group, to determine the evolutionary process of the lily family. As shown in Figure 1, the ITS gene of tulipa is used as the

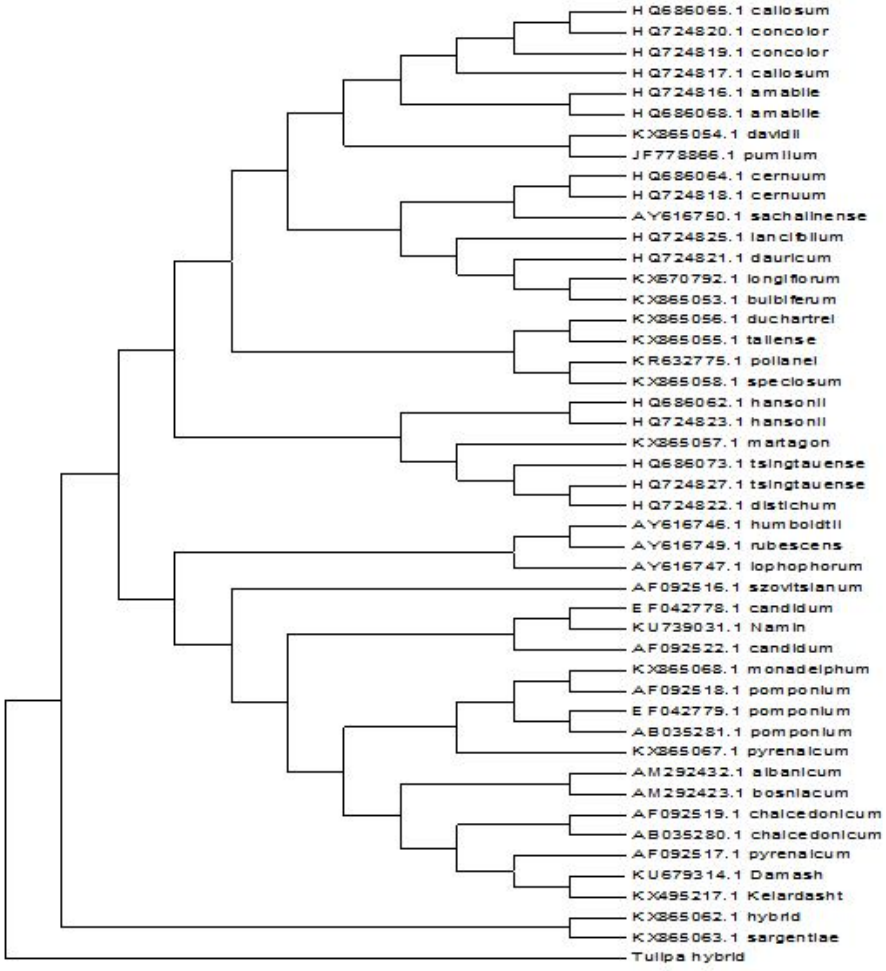

Fig. 1 The phylogenetic tree of the ITS gene using the parsimony method

Fig. 2 The interconnected grid of the studied varieties of the Lilium genus using the ITS gene

out-group, which is located in a separate branch, and has rooted the phylogenic tree. The phylogenic tree also has two distinct clusters. The results of phylogenetic analysis showed that sargentiae and hybrid varieties are older than other varieties of the Lilium family. Also, the location of *L. ledebourii* varieties (Damash and Namin) was determined in the phylogenetic tree using the ITS marker.

The determination of haplotype groups is one of the important steps in determining the geographical location of the studied breeds compared with other breeds. As shown in Figure 2, high-abundance and low-abundance haplotypes are shown with bigger and smaller circles, respectively. All 46 samples are placed in Haplotype group A. This haplogroup is the largest haplogroup in a variety of different species worldwide. Research results showed that Haplogroup A is commonly found in all continents, and Haplogroup B may also have originated from Asia, which has been proven by the findings of Luikart et al.'s studies (2003). Conventionally, the study of the genetic evolution of various species of lilies was carried out based on the study of a small fragment of the control area called the ITS region. The study showed that the maternal line of modern Lilium varieties is related to three basic populations, which are known today as Haplotypes A, B, and C. This shows that the lilies of the different regions of the world do not have a common maternal line and are related to a small number of lilies (Luikart et al. 2003). The results of the present research also showed that Haplogroup A is commonly found in all continents and Haplogroup B may also have originated from Asia.

The size of the circles is proportional to its abundance. There are techniques for detecting species and even varieties as raw and processed that are used in plant products (Ma et al. 2006). Molecular techniques have been widely used for fraud detection. These techniques are the most accurate methods for identifying plant material and samples because they utilize DNA for analysis. Several studies have shown sequence variability in the ITS region of the rDNA that is suitable for phylogenetic studies in many families of fungi and eukaryotes as a result of its two-parent inheritance, ease of learning, intraspecies sameness, interspecies variation and high copy numbers (Feng et al. 2009). The present study showed that the ITS marker has the ability to distinguish between subspecies and genotypes within a population. The spacer regions of rRNA transcripts are among sequences that remain conserved during evolution. This issue was revealed using the ITS sequence as a marker in *Lilium* genotypes (Ikinct, 2010). Due to incomplete ITS sequencing in *L. ledebourii*, polymorphism was observed throughout the replicated fragments. Variations in ITS regions have resulted to the formation of polymorphism. These variations could be diverse depending on the climatic conditions of each area (Ainouche and Bayer 1999), causing the masses to be separated from each other in different regions (Fig. 1). According to Vavilov's theory, the origin of the plants belongs to the centers having the highest diversity (Adams et al. 1998). Therefore, it is necessary to pay more attention to these geographic areas while collecting the Lilium germplasm. Since most of the genes of resistance to diseases, pests, environmental stresses, and genes affecting product quality are commonly found in diversity centers, therefore, if plant breeders have access to correct and accurate information on the genetic diversity of each plant, they can make more efficient use of genetic reserves and directly collect the required genetic reserves.

Acknowledgment

I warmly thank the staffs and managers of Iranian Biological Resource Center for providing necessary facilities to do this study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Author contribution statement SGH Conceived and designed research, wrote manuscript and acted as corresponding author. BAF and MRN Supervised development of work, analyzed the data, helped in data interpretation and manuscript evaluation. NMN Conducted experiments, contributed new reagents and drafted the manuscript. All authors read and approved the final manuscript.

References

Adams BJ, Burnell AM, Powers TO. (1998) A phylogenetic analysis of Heterorhabditis (Nemata: Rhabditidae) based on internal transcribed spacer 1 DNA sequence data. Nematol 30, 22-29

Adlard RD, Barker SC, Blair D, Cribb TH. (1993) Comparison ofthe second internal transcribed spacer (ribosomal DNA) from populations and species of Fasciolidae (Digenea). International J for Parasitol 23:423-425

Ainouche AK, Bayer RJ. (1999) Phylogenetic relationships in Lupinus (Fabacea: Papilionoideae) Based on internal transcribed spacer (ITS) of nuclear ribosomal DNA. American J Botany 86(4):590-607

Arzanlou M. (2008) Phylogeny, detection and mating behavior of Mycosphaerella spp occurring on banana. PhD thesis. Wageningen University: the Netherlands P 171

Baranova MV. (1988) A synopsis of the system of the genus Lilium (Liliaceae). Bot Žurn 73:1319-1329

Brasileiro B, Coimbra M, Morais JM, deOliveira NT. (2004) Genetic variability within Fusariumsolani species as revealed by PCR-Fingerprint based on PCR-markers. Brazilian J Microbial 35:205-210

Comber HF. (1949) A new classifi cation of the genus Lilium. Lily Year-Book 13:86-105

Crous PW, Verkley GJM, Groenewald JZ, Samson RA. (2004) Fungal biodiversity. CBS Laboratory Manual Series 1. Centraalbureau voor Schimmelcultures: Utrech P 269

Cunnington HJ, Lawrie CA, Pascoe GI. (2004) Unexpected ribosomal DNA internal transcribed spacer sequence variation within Erysipheaquilegiaesensulato. Fungal Diversity 16:1-10

Davari M, van Diepeningen AD, Babai-Ahari A, Arzanlou M, Najafzadeh MJ, van der Lee TA. (2012) Rapid identification of Fusarium graminearum species complex using Rolling Circle Amplification (RCA). J Microbiol Methods 89:63-70

Doyle JJ, Doyle JL. (1987) A rapid DNA isolation procedure for small quantitiesof fresh leaf tissue. Phytochem Bull 19:11-15

Dubouzet JG, Shinoda K. (1999) Phylogenetic analysis of the Internal Transcribed Spacer region of Japanese Lilium species. Theor Appl Genet 98:954-960

Feng T, Liu S, He X. (2009) Molecular authentication of the traditional Chinese medicinal plant Angelica sinensis based on internal transcribed spacer of nrDNA. Electronic J Biotech 13:1-13

Frézal L, Leblois R. (2008) Four years of DNA barcoding: Current advances and prospects. Infect Genet Evol 8:727-736

Gultepe M, Uzuner U, Coskuncelebi K, Belduz AS, Terzioglu S. (2010) Internal transcribed spacer (ITS) polymorphism in the wild*Primula*(Primulaceae) taxa of Turkey. Turkish J Botany. 34:147-157

Hebert PDN, Cywinska A, Ball SL, deWaard JR. (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313-321

Ikinci N. (2011) Molecular phylogeny and divergence times estimates of Lilium section Liriotypus (Liliaceae) based on plastid and nuclear ribosomal ITS DNA sequence data. Turk J Bot 35:319-330

İkinci N, Oberprieler C. (2010) Genetic relationships among NE Turkish Lilium L. (Liliaceae) species based on a Random Amplifi ed Polymorphic DNA (RAPD) analysis. Pl Syst Evol 284:41-48

Joseph N, Krauskopf E, Vera MI, Michot B. (1999) Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. Nucleic Acids Research 27:4533-4540

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. (2003) The power and promise of population genomics: from genotyping to genome typing. Nature reviews. Genetics 4(12):981-994

Ma X, Zhu D, Li S, Dong T, Tsim K. (2006) Authentic identification of stigma Croci (stigma of *Crocus sativus* L.) from its adulterants by molecular genetic analysis. Planta Med 2:183-186

McRae EA. (1998) Lilies: A Guide for Growers and Collectors. Oregon: Timber Press P 392

Nishikawa T, Okazaki K, Uchino T, Arakawa K, Nagamine T (1999) A molecular phylogeny of *Lilium* in the Internal Transcribed Spacer region of nuclear ribosomal DNA. J Mol Evol 49:238-249

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M. (1999) Mining SNPs from EST databases. Genome Research 9(2):167-174

Resetnik I, Liber Z, Satovic Z, Cigic P, Nikolic T. (2007) Molecular phylogeny and systematics of the Lilium carniolicum group (Liliaceae) based on nuclear ITS sequences. Pl Syst Evol 265: 45-58

Ronsted N, Law S, Th ornton H, Fay MF, Chase MW. (2005) Molecular phylogenetic evidence for the monophyly of Fritillaria and Lilium (Liliaceae; Liliales) and the infrageneric classification of Fritillaria. Mol Phyl Evol 35:509-527

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS. (2000) Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol. 31:21-32

Termorshuizen AJ, Arnolds EJM. (1997) On the nomenclature of the European species of the Armillaria mellea group. Mycotaxon 30:101-106