

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**BİTKİ MOLEKÜLER FİLOGENİSİNDE KULLANILAN BAZI GEN VE
YÖNTEMLERİN KARŞILAŞTIRILMASI**

BEHCET İNAL

YÜKSEK LİSANS TEZİ

TARLA BİTKİLERİ ANABİLİM DALI

2011

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**BİTKİ MOLEKÜLER FİLOGENİSİNDE KULLANILAN BAZI GEN VE
YÖNTEMLERİN KARŞILAŞTIRILMASI**

BEHCET İNAL

YÜKSEK LİSANS TEZİ

TARLA BİTKİLERİ ANABİLİM DALI

**Bu tez 2011.02.0121.002 proje numarası ile Akdeniz Üniversitesi Bilimsel
Araştırma Projeleri Koordinasyon Birimi tarafından desteklenmiştir.**

2011

T.C.
AKDENİZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİTKİ MOLEKÜLER FİLOGENİSİNDE KULLANILAN BAZI GEN VE
YÖNTEMLERİN KARŞILAŞTIRILMASI

BEHCET İNAL

YÜKSEK LİSANS TEZİ

TARLA BİTKİLERİ ANABİLİM DALI

Bu tez 04/07/2011 tarihinde aşağıdaki jüri tarafından (98) not takdir edilerek oybirliği ile kabul edilmiştir.

Doç. Dr. Mehmet KARACA (Danışman)



Prof. Dr. Kenan TURGUT



Prof. Dr. A. Naci ONUS



ÖZET

BİTKİ MOLEKÜLER FİLOGENİSİNDE KULLANILAN BAZI GEN VE METODLARIN KARŞILAŞTIRILMASI

Behcet İNAL

Yüksek Lisans Tezi, Tarla Bitkileri Anabilim Dalı

Danışman: Doç. Dr. Mehmet KARACA

Temmuz 2011, 213 Sayfa

Canlıların sınıflandırılması ve canlı birliklerinin sınırlarının çizilmesi gözleme ve deneye dayalı sistemli bilgi üretmeye başlanmasıyla birlikte karşılaşılan en karmaşık problemlerden biri olmuştur. Bu amaçla araştırmacılar birçok kuram ve metot geliştirerek var olan canlı çeşitliliğini saptamaya çalışmışlardır. Bu çeşitliliği ortaya koyan moleküler sistematik çalışmaları özellikle son 20 yıldır hızla gelişmektedir. Bu gelişme süreci içinde dizi analizlerinin kullanımı ve yeni filogenetik analiz metotlarının bulunması moleküler sistematığe de katkı sağlamaya başlamıştır. Filogenetik bilgi açısından morfolojik karakterlerin yetersiz olduğu fosil kaynaklarının bulunmadığı durumlarda dizi analizleri filogenetik analizler için çok faydalı olmaktadır.

Bu çalışmada, filogenetik çalışmalarda etkin olarak kullanılabilen gen, gen bölgesi, gen sayısı, genin hücresel lokasyonu (kloroplast, nükleus), gen veya genler için hizalama yöntemleri (alignment) ve hizalama yöntemleri ile hizalanmış veriler yaygın filogenetik analiz yöntemleri ve programları kullanılarak değerlendirilmiştir. Çalışmada dört farklı alignment programı (CLUSTAL W, T-COFFEE, MAFFT, MUSCLE) ve bu programlarda bulunan iki farklı algoritma kullanılmıştır.

Filogenetik yöntemlerin karşılaştırılması için uzunluk ("Distance"), Maksimum Olası (Maksimum Likelihood, ML), Maksimum Parsimoni (Maximum Parsimony, MP) ve Bayes-MCMC (BI) yöntemleri iki farklı veri seti kullanılarak karşılaştırılmıştır. Veri setlerinden birincisi 44 pamuk (*Gossypium*) türünden oluşturulmuş olup burada ITS1,

ITS2, ITS1-ITS2 n klear ribozomal RNA'ya ait DNA dizileri, ikinci veri seti ise  alı, a a sı, otsu, monokotiledon ve dikotiledondan olu an toplam 60 taksondan ve 5 farklı lokustan (*matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB*) olu turulmu tur.

 alı mamız sonucunda yaptığımız analizlere dayanarak kullanılan alignment (hizalama) programlarından olan MAFFT'in mevcut sekans verilerimizi en uygun bir  ekilde hizalayarak do ru *Gossypium* a a  topolojisini vermi tir. Aynı zamanda filogenetik metodları analiz etmek i in mevcut veriye en uygun substit syon modeli olarak GTR ve TN93 modeli  nerilmi tir.  alı mada kullanılan atmış bitki taksonu i in ise bitki gruplarını y ksek posterior probability de erleri ayırt etti inden *matK*, ve *trnH-psbA* ilgili bitki grubu i in markır geni olarak belirlenmi tir.

ANAHTAR KELİMELELER: Filogeni, molek ler sistematik, biyoinformatik, filogenetik analiz

J R : Do . Dr. Mehmet KARACA (Danı man)
Prof. Dr. Kenan TURGUT
Prof. Dr. A. Naci ONUS

ABSTRACT

COMPARISON OF SOME GENES AND METHODS USED IN PLANT MOLECULAR PHYLOGENY

Behcet İNAL

M.Sc. Thesis, Department of Field Crops

Supervisor: Assoc. Prof. Dr. Mehmet KARACA

July 2011, 213 pages

Drawing the boundaries of the classification of living organisms and living units to produce systematic knowledge based on observations and experiments with the introduction has been one of the most complex problems. For this purpose, researchers have developed many theories and methods to determine the diversity of the livings. The molecular systematic studies that reveal biodiversity have been rapidly evolving in the last 20 years. In this developing process using sequence analysis and methods of phylogenetic analysis of molecular systematic have also started to contribute the phylogenetic studies. In terms of phlogenetic information when morphological characters and fasils are insufficient for phylogenetic studies. DNA sequence analysis is very useful for phylogenetic studies.

In this study, genes, gene regions, the cellular location of the gene (chloroplast, nucleus), the number of genes, which use in phylogenetic studies efficiently, four alignment programs (CLUSTAL W, T-COFFEE, MAFFT, MUSCLE), aligned data sets with these alignment methods and algorithms using common phylogenetic analysis programs were evaluated.

In order to compare phylogenetic methods Distance (D), Maximum Likelihood (ML), Maximum Parsimony (MP), and Bayesian-MCMC (BI) were compared using two different data sets. Hereby, the first data set was composed of 44 cotton (*Gossypium*)

species. This set was assessed using ITS1, ITS2, and ITS1-ITS2 DNA sequences of nuclear ribosomal RNA. The second set composed of 60 taxa containing shrubs, woody, herbaceous, monocotyledon and dicotyledon were evaluated using 5 different chloroplast genes (*matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB*).

The consequence of this study, based on our analysis, MAFFT alignment software aligned our data the best fit and had gave the right tree topology for *Gossypium* species. At the same time the most appropriate substitution models were GTR and TN93 for the current data. Based on the bootstrap and posterior probability values *matK* and *trnH-psbA* genes were determined as the marker genes for the sixty plant taxa used in the study.

KEYWORDS: Phylogeny, molecular systematic, bioinformatic, phylogeny analysis.

COMMITTEE: Assoc. Prof. Dr. Mehmet KARACA (Supervisor)

Prof. Dr. Kenan TURGUT

Prof. Dr. A. Naci ONUS

ÖNSÖZ

Bitki moleküler sistematigi özellikle son 20 yıldır hızla gelişmektedir. Bu gelişme dizi analizlerinin kullanılmasıyla ve yeni filogenetik analiz metotlarının bulunmasıyla moleküler sistematige de katkı sağlamaya başlamıştır. Filogenetik bilgi açısından morfolojik karakterlerin ve fosil kaynakları yetersiz olduğu zamanlarda dizi analizleri filogenetik analizler için çok faydalı olmaktadır. Çünkü filogenetikçiler genellikle dizilerin filogenisinin organizmaların filogenisine çok yakın olduğunu farz etmektedirler. Dizi analizi yöntemleri canlıların coğrafik orijinlerinin bulunmasından canlıların filogenilerini moleküler olarak ispatlamaya kadar birçok alanda kullanılmaktadır.

Bu Yüksek Lisans Tez çalışmasında giderek önemi artan moleküler filogeni çalışmalarında kullanılabilecek gen, gen bölgeleri, gen lokasyonları farklı alignment programları ile analiz edilmiş ve elde edilen veriler etkin nükleotit süstitüsyon modelleri ve filogenetik yöntemler kullanılarak analiz edilmiştir.

Çalışmalarımı yönlendiren, araştırmalarımın her aşamasında bilgi, öneri ve yardımlarını esirgemeyerek akademik ortamda olduğu kadar beşeri ilişkilerde de engin fikirleriyle yetişme ve gelişmeye katkıda bulunan danışman hocam Sayın Doç. Dr. Mehmet KARACA'ya en içten teşekkürlerimi sunarım. Her türlü bilgi ve deneyimi hiç esirgmeden benimle paylaşan değerli hocam Sayın Dr. Ayşe Gül İNCE'ye, yüksek lisans çalışmamın her aşamasında desteklerini esirgemeyen, yorulduğum anlarda her zaman yanımda olan sevgili arkadaşım Sayın Araş. Gör. Adnan AYDIN'a teşekkür ederim.

Değerli katkıları için jüri üyelerim Sayın Prof. Dr. Kenan TURGUT (Akdeniz Üniversitesi) ve Sayın Prof. Dr. A. Naci ONUS'a (Akdeniz Üniversitesi) teşekkür ederim. Tanıdığım ilk günden beri maddi ve manevi desteğini hiç esirgemeyen ve pozitif düşünceleri ile beni hep bilim dünyasına bağlayan değerli Sayın Yrd. Doç. Dr. M. Emre EREZ'e (Siirt Üniversitesi), bana tezim sırasında tanıdığı her türlü imkânlardan dolayı Sayın Yrd. Doç. Dr. Osman KARABACAK'a (Siirt Üniversitesi) ve

tezim sırasında desteklerini hep yanımda hissetiğim Siirt Üniversitesi Biyoloji bölümünün geri kalan tüm öğretim üyeleri ve öğretim elemanlarına teşekkür ederim.

Bu çalışmaya 2011.02.0121.002 no'lu araştırma projesi ile maddi destek sağlayan Akdeniz Üniversitesi Rektörlüğü Bilimsel Araştırma Projeleri Koordinasyon Birimi'ne, maddi ve manevi desteklerini hiçbir zaman esirgemeyip, bu günlere gelmemi sağlayan AİLEME sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	iii
ÖNSÖZ	v
İÇİNDEKİLER	vii
SİMGELER ve KISALTMALAR DİZİNİ	ix
ŞEKİLLER DİZİNİ.....	xii
ÇİZELGELER DİZİNİ	xix
1. GİRİŞ	1
2. KURAMSAL BİLGİLER ve KAYNAK TARAMALARI	5
3. MATERYAL ve METOT	36
3.1. Pamuk Taksonları ve ITS	36
3.2. Dış Grup Olarak Kullanılan Taksonlar	36
3.3. Diğer Taksonlar ve Genler	36
3.4. Dış Grup Olarak Ele Alınan Taksonların DNA İzolasyonu.....	36
3.5. DNA Kalitesi ve Miktarının Belirlenmesi.....	38
3.6. Polimeraz Zincir Reaksiyonları.....	38
3.7. Amplikonların Jelden İzolasyonu.....	39
3.8. Transformasyon.....	39
3.9. Amplikonların Vektöre Aktarılması ve Sekans Analizleri.....	40
3.10. Ligasyon	40
3.11. Rekombinant Klonların Analizi	40
3.12. Plazmit DNA Ekstraksiyon Protokolü	41
3.13. Amplikonların Sekans Analizleri	42
3.14. Verilerin Analizi.....	42
3.14.1. Genlere ait sekans dizilerinin gen bankasından indirilmesi.....	42
3.14.2. Elde edilen sekansların alignment (Hizalama) yapılması.....	42

3.14.3. Filogenetik analiz için kullanılan metodlar	43
4. BULGULAR ve TARTIŞMA.....	44
4.1. Çalışmada Kullanılan Bitki Taksonları, DNA Dizileri ve Elde Ediliş Şekilleri ..	44
4.2. Çalışmada Kullanılan Hizalama (Alignment) Yazılım Programları ve Seçilen Parametrelerin Belirlenmesi	51
4.3. Filogenetik Analiz İçin Substitüsyon Modellerinin Belirlenmesi	53
4.4. Kullanılan Filogeni Yöntemleri ve Oluşturdukları Filogenetik Ağaçlar.....	55
4.4.1. Maksimum Olası (ML) metodu ve elde edilen ağaçlar	56
4.4.2. Bayes analizi sonucu elde edilen ağaçlar.....	61
4.4.3. Maksimum parsimony (MP) analizi sonucu elde edilen ağaçlar.....	66
4.4.4. Distance analizi sonucu elde edilen ağaçlar	71
4.5. Uygun Markır (barkod) Geninin Belirlenmesi.....	76
4.5.1. Markır gen için yapılan maksimum likelihood (ML) analizi ve elde edilen ağaçlar.....	77
4.5.2. Markır gen için yapılan Bayes analizi ve oluşturulan ağaçlar	85
5. SONUÇ	93
6. KAYNAKLAR	96
7. EKLER.....	101
8. ÖZGEÇMİŞ	

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler:

°C	Santigrat derece
%	Yüzde
g	Gram
kb	Kilobaz
LiCl	Lithium klorür
M	Molarite
MgCl ₂	Magnezyum klorür
mL	Mililitre
mM	Milimolar
mmol	Milimol
NaCl	Sodyum klorür
nm	Nanometre
pH	“Potentia hydrogeny”
v	Hacim
v/v	Hacim/hacim
w/v	Ağırlık/hacim
µg	Mikrogram
µL	Mikrolitre

Kısaltmalar

<i>accD</i>	Asetil CoA karboksilaz geni
AFR	Afrika
AIC	“Akaike Information Criterion”
AMR	Amerika
AVS	Avustralya
BIC	“Bayesian Information Criterion”
Bknz	Bakınız
BLAST	“Basic Local Alignment Search Tool”
BST	“Bootstrap”
D	“Distance”
DNA	“Deoxyribonucleic Acid” (Deoksiribonükleik Asit)
DNTP	“Deoxyribonucleotide triphosphate”
EB	Ekstraksiyon çözeltisi
EDTA	“EthyleneDiaminteTraaceticAcid”
GEP	“Gap Extensiyon Penalty”
GI	“Gene Identifier”
GOP	“Gap Opening Penalty”
GTR	“General-Time Reversible”
ITER	“Iteration”
ITS	“Internal Transcribed Spacer”
K2P	Kimura’s 2 Parameter
LB	Lezyon çözeltisi
LRT	“Likelihood Ratio Test”
MA	Moleküler Ağırlık
MAFFT	“Multiple Alignment with Fast Fourier Transform”
<i>matK</i>	Maturaz geni
MCMC	“Markov Chain Monte Carlo”
ML	“Maximum Likelihood”
MP	“Maximum Parsimony”
MUSCLE	“Multiple Sequence Comparison by Log-Expectation”
NCBI	“National Center Biotechnology Informative”
PAUP	“Phylogenetic Analysis Using Parsimony”
PP	“Posterior Probability”
PRMT	Parametre
PVPP	“Polyvinylpolypyrrolidone, Insoluble”
PZR	Polimeraz Zincir Reaksiyonu
<i>rpoB</i>	RNA polimeraz B alt ünitesi geni
SBT	Subtitüsyon
SLS	“Sodium N-Lauroyl Sarcosinate”
TN93	Tamura-Nei 1993

TRIS	“Tris (hydroxymethyl) aminomethane”
Triton X-100	“Octyl phenol ethoxylate”
TTP	Tetraploid

ŞEKİLLER DİZİNİ

Şekil 2.1.	UPGMA ağacında dal uzunlukları hesaplanması.....	15
Şekil 2.2.	ITS1 ve ITS2 bölgelerinin şematik görünüşü.....	27
Şekil 2.3.	<i>matK</i> lokusunun şematik görünüşü	28
Şekil 4.1.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi	58
Şekil 4.2.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	59
Şekil 4.3.	MAFFT programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki ML analizi	60
Şekil 4.4.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	63
Şekil 4.5.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	64
Şekil 4.6.	MAFFT programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki Bayes analizi	65
Şekil 4.7.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Parsimony analizi	68
Şekil 4.8.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Parsimony analizi	69
Şekil 4.9.	MAFFT programının 1. parametresi ile hizalanmış ITS2-ITS2 sekansının GTR SBT modelindeki Parsimony analizi	70
Şekil 4.10.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Distance analizi	72
Şekil 4.11.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Distance analizi	74
Şekil 4.12.	MAFFT programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki Distance analizi	75
Şekil 4.13.	MAFFT programının 1. parametresi ile hizalanmış <i>matK</i> gen sekansının GTR SBT modelindeki ML analizi	80
Şekil 4.14.	MAFFT programının 1. parametresi ile hizalanmış <i>accD</i> gen sekansının GTR SBT modelindeki ML analizi	81
Şekil 4.15.	MAFFT programının 1. parametresi ile hizalanmış <i>rpoC1</i> gen sekansının GTR SBT modelindeki ML analiz	82
Şekil 4.16.	MAFFT programının 1. parametresi ile hizalanmış <i>trnH-psbA</i> gen sekansının GTR SBT modelindeki ML analizi	83
Şekil 4.17.	MAFFT programının 1. parametresi ile hizalanmış <i>rpoB</i> gen sekansının GTR SBT modelindeki ML analizi	84
Şekil 4.18.	MAFFT programının 1. parametresi ile hizalanmış <i>matK</i> gen sekansının GTR SBT modelindeki Bayes analizi	88

Şekil 4.19.	MAFFT programının 1. parametresi ile hizalanmış <i>accD</i> gen sekansının GTR SBT modelindeki Bayes analizi	89
Şekil 4.20.	MAFFT programının 1. parametresi ile hizalanmış <i>rpoC1</i> gen sekansının GTR SBT modelindeki Bayes analizi	90
Şekil 4.21.	MAFFT programının 1. parametresi ile hizalanmış <i>trnH-psbA</i> gen sekansının GTR SBT modelindeki Bayes analizi	91
Şekil 4.22.	MAFFT programının 1. parametresi ile hizalanmış <i>rpoB</i> gen sekansının GTR SBT modelindeki Bayes analizi	92
Şekil Ek-1.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	102
Şekil Ek-2.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	103
Şekil Ek-3.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	104
Şekil Ek-4.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	105
Şekil Ek-5.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	106
Şekil Ek-6.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	107
Şekil Ek-7.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	108
Şekil Ek-8.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	109
Şekil Ek-9.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	110
Şekil Ek-10.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	111
Şekil Ek-11.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	112
Şekil Ek-12.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	113
Şekil Ek-13.	MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	114
Şekil Ek-14.	MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	115
Şekil Ek-15.	MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	116
Şekil Ek-16.	MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	117
Şekil Ek-17.	MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	118

Şekil Ek-18.	MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	119
Şekil Ek-19.	MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	120
Şekil Ek-20.	MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	121
Şekil Ek-21.	MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	122
Şekil Ek-22.	MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	123
Şekil Ek-23.	MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	124
Şekil Ek-24.	MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	125
Şekil Ek-25.	T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	126
Şekil Ek-26.	T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	127
Şekil Ek-27.	T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	128
Şekil Ek-28.	T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	129
Şekil Ek-29.	T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi	130
Şekil Ek-30.	T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi.....	131
Şekil Ek-31.	T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi	132
Şekil Ek-32.	T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi	133
Şekil Ek-33.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	134
Şekil Ek-34.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	135
Şekil Ek-35.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	136
Şekil Ek-36.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	137
Şekil Ek-37.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	138
Şekil Ek-38.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	139

Şekil Ek-39.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	140
Şekil Ek-40.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	141
Şekil Ek-41.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	142
Şekil Ek-42.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	143
Şekil Ek-43.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	144
Şekil Ek-44.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analiz	145
Şekil Ek-45.	MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	146
Şekil Ek-46.	MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	147
Şekil Ek-47.	MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	148
Şekil Ek-48.	MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	149
Şekil Ek-49.	MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	150
Şekil Ek-50.	MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	151
Şekil Ek-51.	MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	152
Şekil Ek-52.	MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	153
Şekil Ek-53.	MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	154
Şekil Ek-54.	MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	155
Şekil Ek-55.	MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	156
Şekil Ek-56.	MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	157
Şekil Ek-57.	T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	158
Şekil Ek-58.	T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	159
Şekil Ek-59.	T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	160

Şekil Ek-60.	T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	161
Şekil Ek-61.	T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi	162
Şekil Ek-62.	T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi	163
Şekil Ek-63.	T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi	164
Şekil Ek-64.	T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi	165
Şekil Ek-65.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi	166
Şekil Ek-66.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi	167
Şekil Ek-67.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi	168
Şekil Ek-68.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi	169
Şekil Ek-69.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi	170
Şekil Ek-70.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi	171
Şekil Ek-71.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi	172
Şekil Ek-72.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi	173
Şekil Ek-73.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi	174
Şekil Ek-74.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi	175
Şekil Ek-75.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi	176
Şekil Ek-76.	MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi	177
Şekil Ek-77.	MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi	178
Şekil Ek-78.	MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi	179
Şekil Ek-79.	MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi	180
Şekil Ek-80.	MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi	181

Şekil Ek-81.	MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi.....	182
Şekil Ek-82.	MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi.....	183
Şekil Ek-83.	MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi.....	184
Şekil Ek-84.	MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi.....	185
Şekil Ek-85.	MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi.....	186
Şekil Ek-86.	MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi.....	187
Şekil Ek-87.	MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi.....	188
Şekil Ek-88.	MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi.....	189
Şekil Ek-89.	T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansını GTR SBT modelindeki distance analizi.....	190
Şekil Ek-90.	T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansını K2P SBT modelindeki distance analizi.....	191
Şekil Ek-91.	T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansını GTR SBT modelindeki distance analizi.....	192
Şekil Ek-92.	T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansını K2P SBT modelindeki distance analizi.....	193
Şekil Ek-93.	T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansını GTR SBT modelindeki distance analizi.....	194
Şekil Ek-94.	T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansını K2P SBT modelindeki distance analizi.....	195
Şekil Ek-95.	T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansını GTR SBT modelindeki distance analizi.....	196
Şekil Ek-96.	T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansını K2P SBT modelindeki distance analizi.....	197
Şekil Ek-97.	ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	198
Şekil Ek-98.	ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	199
Şekil Ek-99.	ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	200
Şekil Ek-100.	ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	201
Şekil Ek-101.	MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	202

Şekil Ek-102. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	203
Şekil Ek-103. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	204
Şekil Ek-104. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	205
Şekil Ek-105. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	206
Şekil Ek-106. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	207
Şekil Ek-107. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	208
Şekil Ek-108. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	209
Şekil Ek-109. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	210
Şekil Ek-110. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	211
Şekil Ek-111. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi.....	212
Şekil Ek-112. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi.....	213

ÇİZELGELER DİZİNİ

Çizelge 2.1. Yaygın olarak kullanılmakta olan alignment yazılım programları.....	8
Çizelge 2.2. Alignment ve oluşturduğu rehber ağaç.....	11
Çizelge 2.3. Yaygın bazı filogenetik programları ve kullandıkları metodlar	12
Çizelge 2.4. Yaygın olarak kullanılan analiz yöntemleri.....	13
Çizelge 2.5. Filogenetik çalışmalarda yaygın olarak kullanılan moleküler markırlar....	26
Çizelge 4.1. Çalışmada kullanılan pamuk türleri ve GenBank GI numaralar	46
Çizelge 4.2. Çalışmada kullanılan atmış taksona ait <i>matK</i> , <i>rpoB</i> ve <i>rpoC1</i> genlerinin GI numaraları.....	47
Çizelge 4.3. Çalışmada kullanılan atmış taksona ait <i>accD</i> ve <i>trnH-psbA</i> genlerinin GI numaraları.....	49
Çizelge 4.4. Çalışmada kullanılan alignment programları ve parametreleri	51
Çizelge 4.5. Gen bölgesine ve hizalama parametrelerine bağlı olarak test modellerin seçtiği, mevcut verilere uygun en iyi substitüsyon (SBT) modelleri.....	53
Çizelge 4.6. Farklı programlarla hizalanmış gen bölgelerine ait parsimonik olarak bilgi verici baz sayıları	66
Çizelge 4.7. Farklı programlarla hizalanmış gen bölgelerine ait parsimonik olarak bilgi verici baz sayıları	66

1. GİRİŞ

Organizmaların sınıflandırılması, birbirleri ve ilgili diğer gruplarla ilişkilerinin saptanması, geçmişteki ve gelecekteki konumlarının belirlenmesi bilim insanlarının çok uzun süredir ilgilendiği konuların başında gelmektedir. Bu amaçla araştırmacılar birçok kuram ve metod geliştirerek var olan canlı çeşitliliğini saptamaya çalışmışlardır. Bu çeşitliliği ortaya koyan moleküler sistematik çalışmaları özellikle son 20 yıldır hızla gelişmektedir. Bu gelişme süreci içinde moleküler dizi analizinin kullanımı ve yeni filogenetik analiz metodlarının bulunması moleküler sistematığe de katkı sağlamaya başlamıştır. Filogenetik bilgi açısından fosil kaynaklarının bulunmaması ve morfolojik karakterlerin de çoğu durumda yetersiz olması nedeni ile moleküler dizi analizleri filogenetik analizler için çok faydalı olmaktadır.

Evrin insan oğlunun yaşamında doğrudan gözlemleyeceği süre içinde meydana gelen bir biyolojik süreç olmayıp çok uzun süreler ve şartlar gerektiren bir biyolojik süreç ve gerçektir. Evrim süreçler dizini olup canlıların nesilden nesile değişmesi, canlıların yeni özellikler kazanması ve bazı özelliklerin kaybetmesi şeklinde tanımlanabilir. Organizmaların veya taksonomik gruplarının evrimsel geçmişini, hiyerarşik sistemlerini, genetik ilişkilerini fosil, morfolojik veriler veya genetik analizlerle ortaya konulması ise “filogeni” olarak tanımlanmaktadır.

Biyolojik bir türün türlerin veya bir organizma grubunun zaman süreci içerisinde bir dizi morfolojik, fizyolojik, genetiksel değişimden geçerek şekillenmesi filogeneze olarak tanımlanır. Filogeneze süreci filogenetik yöntemler kullanılarak filogeni şeklinde ortaya konmaktadır. Filogenetik yöntemlerde evrimsel ilişkiler veya diğer bir söylem ile filogeneze bir dallanma süreci olarak düşünülür. Bu bağlamda filogenetik bir ağaç, organizmanın veya bir grup organizmanın tarihsel bir geçmişini gösterir. Örneğin zaman içinde ne kadar değişime uğradığını ve bunun sonucu olarak farklı dallar halinde nasıl türleştiği, birbiri ile melezlendiğini ve/veya tükenerek son bulduğunu gösterebilir.

Genel olarak filogeneze iki yoldan gerçekleşmektedir. Bu yollar dikey (vertikal) gen transferi ve yatay (horizontal veya lateral) gen transferidir. Filogeneze sürecine dikey ve yatay gen transferi farklı düzeylerde etkili olmaktadır. Dikey gen transferinde genler ebeveynlerden döllerle aktarılmakta veya diğer bir söylem ile transfer edilmektedir.

Dikey gen transferi genellikle birbirileri ile çok yakın ilişkili türler arasında ve türler içerisinde yaygın olan gen transfer şeklidir.

Horizontal gen transferi ise, birbirleri ile ilişkisiz organizmalar arasında genlerin akışı ile meydana gelir. Horizontal gen transferi prokaryotlarda sık rastlanan bir olgu olmakla birlikte rekombinant gen transferi ile elde edilen transgenik organizmalarda (GDO) horizontal gen transferi ile oluşturulmaktadır. Horizontal gen transferi organizmaların filogenisinin belirlenmesini zorlaştırmaktadır. Bu neden ile etkin bir filogenetik analizde horizontal gen transfer özelliği olmayan genlerin kullanılması önemlidir. Genlerin arasında transfer yönü bakımından farklılıklar bulunmaktadır. Örneğin ribozomal RNA (rRNA) genleri horizontal gen transfer özelliğine sahip değildir. Bu özelliği nedeni ile en yaygın kullanılan filogenetik dizilerdir. Horizontal gen transfer özelliğinin yanı sıra ribozomal gen dizilerinin moleküler saat hipotezine daha yakın bir özellik gösterdiği bilinmektedir. Filogeni çalışmalarında kullanılan genlerin özellikleri yanı sıra taksonlara ait diziler arasında oluşturulan hizalama da (alignment) önemlilik arz etmektedir. Filogeni çalışmalarında kullanılan genlerin ve alignment işlemlerinin dışında önemli olan bir diğer nokta ise dizilerin birbirileri ile karşılaştırılmasında ve filogenetik ağaçların oluşturulmasında kullanılacak nükleotid substitüsyon modellerinin belirlenmesidir. Filogeni çalışmalarında önemli olan diğer bir unsur ise en uygun filogeniyi tahmin edecek yöntemin ve programın kullanılmasıdır.

Bu araştırmanın amacı filogeni çalışmalarında etkin olarak kullanılabilecek gen, gen bölgesi, gen sayısı, genin hücresel lokasyonu (Kloroplast, Nükleus) hizalama yöntemleri (alignment), filogenetik analiz yöntemlerini ve programlarını belirlemektir.

Çalışmada toplam 6 gen sekansı kullanılmıştır. Gen bölgelerinin filogeniye etkilerini belirlemek için ITS1, ITS2, ITS1-ITS2 nüklear ribozomal RNA'ya ait dizilerini ve gen sayısını ve çeşidinin etkisini ortaya koymak için ise *matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB*, kloroplast genlerine ait dizilerden yararlanılmıştır. Ayrıca bu genlerdeki nükleotid substitüsyon modelleri araştırılmıştır.

Çalışmada kullanılan toplam bitki takson sayısı 105 olup bunlardan 44 tanesi *Gossypium* türlerinden (birinci veri seti), geri kalanı ise 15 familyanın temsil edildiği 60 taksondan (ikinci veri seti) ve dış grup olarak kullanılan *Origanum vulgare* subsp.

hirtum taksonlarından oluşturulmuştur. Çalışmada *Gossypium* taksonların kullanım amacı bu cinse ait mevcut filogenetik bilgilerinin fosil, morfolojik ve moleküler diziler kullanılarak ortaya konulmuş olması nedeni ile seçilmiştir. Çalışmada kullanılan 60 taksonun seçimi ise bu taksonların bütününde *matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB* gen dizilerinin belirlenmiş olması ve filogenik açıdan belirgin farklılıklar içermelidir.

Çalışmada bilgisayar ve süre göz önüne alınarak 4 farklı alignment yöntemi (ClustalW, MAFFT, MUSCLE, T-COFFEE) ve her bir yöntem için 2 değişik algoritma kullanılmıştır.

Filogenetik yöntemlerin karşılaştırılması için 3 farklı filogeni programı (PAUP, MetaPIGA ve MrBayes) kullanılmıştır ve bu programlarda var olan D, MP, ML, BI filogeni yöntemleri kullanılmıştır.

Çalışmada elde edilen verilerin bir arada kullanıldığında MAFFT hizalama programının diğer 3 programa göre daha doğru filogeni tahminini ortaya koyan alignment programı olduğu tesbit edilmiştir.

Araştırmada farklı nükleotid substitüsyon modelleri karşılaştırılmış ve bu modeller arasında GTR (“General Time Reversible”), K2P (“Kimura's 2 parameter”) ve TN93 (Tamura-Nei 1993) diğer substitüsyon modellerine göre daha uygun olduğu tesbit edilmiştir.

Nüklear ITS1, ITS2, ITS1-ITS2 dizi verilerinin filogeni çalışmaları uygun olduğu ancak 5.8 RNA dizi kısmının uzaklaştırılmasından sonra ITS1-ITS2 bölgesinin daha etkin sonuç oluşturduğu belirlenmiştir. Bu durum birden fazla gen bölgesinin daha doğru filogeni tahmini yaptığı belirlenmiştir. ITS1 ve ITS2 dizileri arasında substitüsyon model farklılığı tesbit edilmiştir. ITS1 için TN93, K2P. ITS2 için GTR ve K2P modeli önerilmiştir. ITS1-ITS2 dizilerinin birlikte analiz edildiğinde de GTR modeli kullanılmıştır.

Kloroplast genlerinden *matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB* karşılaştırıldığı çalışmalarımızda en etkin filogeni tahmininin *matK* ve *trnH-psbA* gen dizilerinden elde

edildiđi ve bu genlerin filogeni alıřmalarında markır gen olarak kullanılabileceđi teyit edilmiřtir.

Bu tez alıřmasında elde edilen veriler z olarak verilmek istenirse bitki filogenisinde kullanılan yntemlerin, lokusların, karřılařtırılması yapılarak etkin ve yntemler belirlenmiřtir. alıřma ıktıları arasında diđer nemli bir nokta ise pamuk filogenisi ile ilgili yeni veriler elde edilmiř ancak alıřma konusundan uzaklařmamak adına detaylandırılmamıřtır. Sonu olarak molekler filogenide kullanılan yntemlerden birden fazlasının kullanımının yararlı olacađı grlmřtr. Ayrıca tek gen dizisi yerine 3 farklı gen dizisinin birlikte kullanımının daha dođru topoloji tahminini ortaya koyacađı tesbit edilmiřtir.

2. KURAMSAL BİLGİLER ve KAYNAK TARAMALARI

Evrım doğrudan gözlemlenemeyecek kadar uzun süreler içinde meydana gelen bir biyolojik süreçtir ve evrimsel değişim filogeni ile ortaya konulur. Genel bir tanımlama ile organizmaların taksonomik gruplarının evrimsel geçmişini, hiyerarşik sistemlerini, genetik ilişkilerini fosil, morfolojik veriler veya genetik analizlerle ortaya konulması “filogeni” olarak tanımlanmaktadır.

Evrimsel değişiklikler genel olarak fosil kaynaklar kullanılarak anlaşılabilmektedir. Ancak fosil kaynaklarının sayılarının yetersiz olmasının yanında analizlerinde de bazı sorunlar bulunmaktadır. Biyolojik bir türün, türlerin veya bir organizma grubunun zaman süreci içerisinde bir dizi morfolojik, fizyolojik, genetiksel değişimlerden geçerek şekillenmesi filogeneze olarak tanımlanır. Filogeneze süreci filogenetik yöntemler kullanılarak filogeni şeklinde ortaya konmaktadır. Filogenetik yöntemlerde evrimsel ilişkiler veya diğer bir söylem ile filogeneze bir dallanma süreci olarak düşünülür. Bu bağlamda filogenetik bir ağaç, organizmaların tarihsel bir geçmişini gösterir. Örneğin zaman içinde ne kadar değişime uğradığını ve bunun sonucu olarak farklı dallar halinde nasıl türleştiği, birbiriyle melezlendiğini ve tükenerek nasıl son bulduğunu gösterebilir.

Filogeninin bir ağaç şeklinde (yaprak, dal, sap ve kök) gösterilmesi ile filogenetik ağaç oluşturulmaktadır (Suarez-Diaz vd 1998). Filogenetik ağaçlar, monofiletik, parafiletik ve polifiletik olabilmektedir. Monofiletik ilişkileri gösteren ağaçlarda, organizmaların türediği ortak atasal kaynak tektir. Bu tip taksonomik gruplama veya filogenetik ağaç monofiletiktir. Diğer taraftan bir taksonomik gruplamada, birden fazla ortak atanın bulunması durumunda polifiletik kavram ortaya çıkar bu tip filogenetik ilişkilerde grup üyeleri için en yakın ortak ata belirlenememiştir.

Parafiletik filogenetik ilişkilerde veya gruplamalarda ise ortak ataya sahip bir grup organizma belirlenmiştir ancak bu grup içerisindeki bazı organizmaların bütün ataları bulunmamaktadır. Örneğin modern sürüngenler bir atadan gelmişlerdir ancak kuşlar bu grupta yer almaz. Sıcakkanlı hayvanlarda hem memeliler vardır hem de kuşlar vardır. Kuşlar ve memelilerin en yakın ortak atası bu grupta yer almaz ve dolayısıyla polifiletik bir filogenetik ilişki vardır.

Lamarck (1809) tarafından ilk kez evrimsel ağaç (filogeni) ortaya konmuş ve bu analizde balık, kuş, memeliler ve bazı omurgasız hayvan grupları ağaç ilişkileri olarak ilk kez gösterilmiştir. Bu tür filogeni uygulamaları yabani veya kültür bitkilerinin taksonomisinde günümüzde yaygın olarak kullanılmaktadır. Günümüzde DNA ve protein dizileri bakteriden insana kadar farklı organizmaların evrimsel tarihlerini veya geçmişlerinin araştırılmasında önemli bir araçtır. Moleküler filogenetik çalışmalarda homolog genlere ait aminoasit ve DNA dizileri taksonomik sınıflandırmaların yapılmasına, taksonlar arasındaki genetik ve biyocografik ilişkilerin saptanmasında, protein kodlamayan diziler de dahil olmak üzere genlerin, proteinlerin ve diğer moleküler dizilerin gruplandırılmasında, epidemik araştırmalarda ve parazit konukçu arasındaki paralel evrimleşme çalışmalarında kullanılmaktadır (Smith vd 1981).

Moleküler filogenetiğin uygulama alanlarına örnek olarak herbisit (“roundup: glyphosate isopropylamine”) ve pestisit (*Bacillus thuringiensis* toksin) dayanıklılık genlerin doğal bitki ya da organizmalarda dayanıklılığın kazanılmasının saptanması ile kullanılan herbisit veya pestisit dozlarının ve ekim alanlarının artırılması ile horizontal gen akışının belirlenmesi gösterilebilir. Filogeniden yararlanılarak ekonomik açıdan önemli olan kültür bitkilerinde var olan genetik kaynakların daha etkin kullanılabilmesi için gerekli genetik bilginin elde edilmesinde, ıslah metotlarının uzun zaman içerisinde genom yapısında meydana getirdiği değişikliklerin durumunun saptanmasında ve bu değişikliklerin eriştiği nokta hakkında bilgi edinmede de oldukça yaygın olarak kullanılmaktadırlar.

Moleküler filogeninin başka ve yaygın bir kullanım alanı da bitki moleküler sistematığıdır. Bitki moleküler sistematığı özellikle son 20 yıldır hızla gelişmektedir (Wen vd 1997). Bu gelişme dizi analizlerinin kullanılmasıyla ve yeni filogenetik analiz metotlarının bulunmasıyla (Ro vd 2007) moleküler sistematığe de kapsamlı katkı sağlamaya başlamıştır. Filogenetik bilgi açısından morfolojik karakterlerin yetersiz olduğu zamanlarda dizi analizleri filogenetik analizler için çok faydalı olmaktadır (Yokoyama vd 2000). Çünkü filogenetikçiler genellikle dizilerin filogenisinin organizmaların filogenisine çok yakın olduğunu farz etmektedirler. Dizi analizi yöntemleri canlıların coğrafik orijinlerinin bulunmasından (Allan vd 2004) canlıların filogenilerini moleküler olarak ispatlamaya kadar birçok alanda kullanılmaktadır

(Cohen ve Weydmann 2005). Angiospermilerin moleküler filogenileri hakkında yapılan çalışmalarda genellikle kloroplast DNA'sı (Graham ve Olmstead 2000) mitokondri DNA'sı (Qui vd 1999) veya çok tekrarlı nüklear ribozomal DNA genleri kullanılmıştır. Son zamanlarda yapılan çalışmalara göre bu üç genomdaki DNA dizileri farklı oranlarda değişime uğramışlardır. Nüklear genomda değişim daha hızlıyken mitokondri ve plastid (kloroplast) genomlarının daha yavaş değiştiği tahmin edilmektedir.

Nüklear ribozomal DNA (nrDNA) bölgeleri arasındaki farklı mutasyon oranları birçok taksonomik seviyede filogenetik sonuç alınmasına olanak sağlar. ITS, (ITS1+ 5.8 ITS2) nrDNA'nın en çok çeşitlilik gösteren bölümüdür ve bitki filogenetiğinin yeniden yapılanmasında önemli bir lokus olduğu kanıtlanmıştır (Liston vd 1999). Bitki moleküler filogenetiğini ortaya koymak için önemli ve vazgeçilmez bir araç olarak biyoinformatik, moleküler filogenetik yöntemlerinin kullanılmasında, proteinlerin primitif formlardaki durumunu belirlemede veya gelecekte dönüşebilecekleri formlarının araştırılmasına kadar biyolojinin hemen her alanında kullanılmaya başlanmıştır. Örneğin aminoasit dizisi bilinen ancak fonksiyonu bilinmeyen bir protein diğer proteinlerle karşılaştırılabilmekte ve benzerlik ağaçları oluşturulabilmektedir. Bu analiz sonucunda proteinin olası fonksiyonları tahmin edilebilmektedir.

Moleküler filogenetik çalışmaların temelinde homoloji gösteren moleküler dizilerin birbirleri ile karşılaştırılarak hizalanması (alignment) ilk adımdır. Alignment işleminde, karakterler arasında oluşturulacak işlemlerde homolog diziler arasında var olabilecek dizi (karakter) değişiklikleri [substitüsyon; transisyon ve transversiyon], eklemler (insersiyon), çıkarmalar (delesyon), aynı dizi üzerinde farklı kısımlarda bulunan düzenleyici dizilerin varlığı veya yokluğu, değişim oranlarının farklılıkları, dizilerin birbirlerine paralog, ortolog olabilmeleri, homoplasi durumlarının varlığı nedeni ile alignment işlemi önemlilik arz etmektedir. Günümüze değin çok sayıda biyoinformatiksek yazılım programı geliştirilerek alignment işlemlerinin doğruluğu artırılmaya çalışılmaktadır (Çizelge 2.1)

Çizelge 2.1. Yaygın olarak kullanılmakta olan alignment yazılım programları

Program Adı	Dizi Tipi	Alignment Tipi	Referans
BALI-PHY	DNA/Protein	Global	Redelings ve Suchard 2010
CLUSTALW	DNA/Protein	Lokal/ Global	Thompson vd 1994
KALIGN	DNA/Protein	Global	Lassmann 2005
MAFFT	DNA/Protein	Lokal/Global	Katoh vd 2005
MAVID	DNA/Protein	Global	Bray ve Pachter 2004
MSA	DNA/Protein	Lokal/Global	Lipman vd 1995
MSAPROBS	Protein	Global	Liu vd 2010
MULTALIN	DNA/Protein	Lokal/Global	Corpet 1998
MULTI-LAGAN	DNA/Protein	Global	Brudno vd 2003
MUSCLE	DNA/Protein	Lokal/Global	Edgar 2004
OPAL	DNA/Protein	Lokal/Global	Wheeler ve Kececioğlu 2007
PECAN	DNA	Global	Paten vd 2008
POA	Protein	Lokal/Global	Lee 2002
PROBALIGN	Protein	Global	Roshan ve Livesay 2006
PROBCONS	Protein	Lokal/Global	Do vd 2005
PRRN/PRRP	Protein	Lokal/Global	Totoki 1991
PSALIGN	DNA/Protein	Lokal/Global	Sze vd 2006
SAGA	Protein	Lokal/Global	Notredame vd 1998
STATALIGN	DNA/Protein	Global	Novak vd 2008
STEMLOC	RNA	Lokal/Global	Holmes 2005
T-COFFEE	DNA/Protein	Lokal/Global	Notredame vd 2008

Homoloji (Homology) analizinde bir atadan gelmiş bireylerde ortak olan özellikler (değişim göstermeyen ve korunmuş karakterler), farklı olan özellikler ve yeni oluşmuş özellikler (mutasyon, rekombinasyon) belirlenebilmektedir. Bu homolog özellikler paralog veya ortolog özelliklerdir. Genler bu özelliklerine göre paralog ve ortolog olarak sınıflandırmaktadır.

Morfolojik özellikler arasında (anatomik ve morfolojik karakterler) paralog ve ortolog yapıların veya unsurların belirlenmesi göreceli olarak daha kolay olmasına rağmen, bu özelliklerin çevreden daha fazla etkilenmeleri, gelişim dönemlerine bağlı olmaları ve sayıca az olmaları nedenleri ile kullanımlarında sorunlar bulunmaktadır. Diğer taraftan morfolojik karakterlere göre moleküler dizilerin filogenide kullanımı daha etkindir.

Ortolog diziler (genler) farklı türlerde bulunan ve benzer dizilere veya fonksiyonlara sahip homolog diziler iken paralog diziler veya genler duplikasyon sonucu oluşmuş veya aynı tür içerisinde atadan gelen homolog dizilerdir. Ancak farklı genler üzerinde bulunan nükleik asitlerin [toplamda 4: A, G, T, C] veya aminoasitlerin [toplamda 20 adet standart aminoasit] salt dizi sıralanışına bakarak hangisinin diğerine paralog veya ortolog olduğu basit hizalama testleri ile belirlenmesi kolay bir işlem değildir. Bu nedenle günümüze değin birçok alignment yazılım programları ve bu programlarda kullanılan farklı algoritmik ve istatistik yöntemleri geliştirilmiştir (Çizelge 2.1).

Basit bir tanımlama ile en az iki molekül diziyi (pairwise alignment) veya sayı olarak daha fazla dizileri (multi-alignment) birbirlerine göre niteliksel konumlandırılmaları alignment olarak nitelendirilebilir. Alignment sonucu oluşan diziler birbirlerine homolog olabilirler ancak her zaman birbirlerine göre hizalanan diziler homolog olmayabilir. Bu durum alignment işlemlerinin basit konumlama işlemi olmadığını açıkça göstermektedir.

Günümüze değin geliştirilmiş olan bazı alignment programları aynı anda mutasyonları (insersyon, delesyon ve sübtitüsyonları) ortaya koyabilecek niteliktedirler. Bu tür programlarda farklı diziler birebir olarak alignment edebilmektedir. Aynı zamanda bu programlarda mutasyonlar göz önüne alınabilmektedir. Bu tür yazılım programlarına ClustalW (BioEdit ve MEGA), Kalign, T-COFFEE, Sequencher, MAFFT ve MUSCLE programları örnek olarak gösterilebilir. Diğer bazı algoritmalarda olasılıklı alignment analizleri, dizilerin bütününde (global) veya bir kısmında (lokal) ayrı ayrı alignment analizleri gerçekleştirilebilmektedir. Global algoritmalar bir hizalamada dizinin bütününe göz önüne alırken bölgesel algoritmalar hizalamada dizilerin belirli bir kısmını dikkate almaktadır.

Çoklu sekans hizalaması (ÇSH) biyolojik sekansı konu alan birçok karşılaştırmalı çalışmalarda çok önem taşımaktadır. ÇSH, aynı zamanda filogenetik çalışmalarda, korunmuş ve korunmamış bölgelerin tanınmasında, kodlanmayan RNA (ncRNA) bölgelerin belirlenmesinde ve daha birçok alanda yaygın bir biçimde kullanılmaktadır (Kato 2008). ÇSH kullanılarak, diziler arasında homolojiden yararlanarak evrimsel ilişkiler ortaya konabilmektedir. ÇSH yöntemleri ile nokta mutasyonları ve delesyonlar

(indel) tespit edilebilmektedir. ÇSH işlevsel olarak önemli konumların tespit edilmesinde kullanılabilir. Korunmuş bölgelerin belirlenmesinde örneğin, bağlanma bölgeleri, aktif bölgeler veya diğer anahtar fonksiyonlara karşılık gelen konumlar bulunabilmektedir. ÇSH yöntemlerinin kullanımında farklı özelliklerin ya da parametrelerin göz önüne almak önemlilik arz etmektedir (Durbin vd 1998). Çoklu hizlama işlemi yukarıda bahsedilen işlemlerin yanı sıra filogeni çıkarımlar için oldukça güçlü bir araç olup filogeni çalışmalarında oldukça yaygın olarak kullanılmaktadır (Gotoh 2003). Günümüze değin geliştirilmiş olan çok sayıda alignment yazılım programları mevcut olup bunların bir kısmı Windows işletim sistemine uyumlu iken bir kısmı Mac, Linux ve ağ tabanlı programlardır.

Karşılıklı Dizilerlerin Hizalaması (Pairwise Alignment) yönteminde çoklu diziler içerisinde seçilmiş bir dizi ile diğer bir dizi karşılaştırılmakta ve daha sonra sırasıyla diğer diziler karşılaştırılmaktadır. İki dizi arasındaki uzaklık (Distance) boşluklar dikkate alınmadan iki dizi arasında birbirleri ile aynı olan unsurların (baz veya aminoasit) toplam sayılarının bütün dizi toplamına bölünmesi ile belirlenmektedir. Örneğin A, B ve C dizileri arasında uzaklık matrisi (Bkz. Çizelge 2.2) de gösterilmektedir.

Alignment işlemlerinden sonra ikinci işlem olarak matriks bilgileri kullanılarak bir ağaç oluşturulmaktadır (Bkz. Çizelge 2.2). Bu işlemin amacı çalışmada kullanılan dizilerin birbirleri ile olan ilişkilerini ortaya koymaktır. Bu ağaç rehber ağaç olarak isimlendirilmektedir. Ancak bu rehber ağacın filogenetik açıdan bir değeri yoktur. DNA kullanılarak yapılan birçok analizde araştırmacılar genellikle grupta analizleri ile filogeni analizlerini karıştırmakta ve grupta sonuçlarını filogeni olarak sunmaktadır.

Alignment işlemlerinde üçüncü adımda rehber ağacına bakılarak veya kullanılarak çalışmada kullanılmakta olan diziler yeniden sıralanmaktadır. Örneğin Çizelge 2.2'deki örnekte önce A dizisi sonra B dizisi ve en son olarak C dizisi sıralanır. Bu sıralama gerekirse manuel olarak değiştirilebilmektedir. Alignment çalışmalarında çoğu durumda manuel eklemeler çıkarmalara özellikle heliks, plaka veya diğer düzenleyici bölgelerde yapılarak oluşturulan hizalamanın fonksiyonel özelliği artırılmaktadır.

Manuel düzeltme oldukça önemli bir olgudur. Yanlış manuel düzeltmeler bazı durumlarda yanlış sonuçların alınmasına neden olduğu gibi manuel düzeltmenin yapılmadığı durumlarda da yanlış sonuçlar alınabilmektedir. Manuel düzeltmelerin zorunluluğu yazılım programlarına sıralı olarak verilen bir dizi grubunda hangisinin atasal dizi ve hangilerinin ise türemiş dizi olduğunun bilinmemesinden kaynaklanmaktadır. Örneğin ClustalW programında genellikle alignment bir dizi kullanılarak diğer dizilerin karşılaştırılması ile gerçekleştirilmekte ve rehber dizi üzerinde sadece boşluklar eklenerek diğer diziler eşleştirebilmektedir.

Çizelge 2.2. Alignment ve oluşturduğu rehber ağaç

Örnekler	A	B	C	
A	1.00			
B	0.87	1.00		
C	0.59	0.60	1.00	

Alignment işlemlerinin tamamlanmasından sonra filogenetik çalışmalara geçilmektedir. Son yıllarda DNA dizi veya gen fragmentlerin filogenetik analizleri için çok sayıda yazılım programları ve analiz yöntemleri geliştirilmiştir. Günümüze değin çok sayıda filogenetik programlar geliştirilmiş olmasına rağmen bu programlardan hangisinin veya hangilerinin kullanımının seçimi oldukça önemlidir (Çizelge 2.3).

Genel bir kural olarak filogenetik çalışmaların en az iki farklı program veya yöntemle gerçekleştirilmesinde yarar bulunmaktadır. Ancak çalışmada elde edilen verilere en uygun olan metodun seçilmesi gerekmektedir Çünkü bir metot ya da yöntemde var olmayan bir algoritma diğer bir yöntemde bulunabilmektedir (Çizelge 2.4). Birçok analizi bir arada gerçekleştiren etkin bir filogenetik programının geliştirilmesinde büyük yararlar bulunmaktadır.

Günümüze değin filogenetik çalışmalarda PAUP yazılım programının kullanılması nedeni bu programda birçok algoritma veya kriterin bulunmasıdır. Ancak PAUP

programını alignment yapamamaktadır. Diğer taraftan MetaPIGA gibi yeni programlar dizi model testlerini gerçekleştirirken uzaklık matrislerini yapamamaktadır. Bu durum moleküler filogenetik çalışmalarında birden fazla programın kullanılmasını zorunlu kılmaktadır.

Çizelge 2.3. Yaygın olarak kullanılan bazı filogenetik programları ve kullandıkları metodlar

Filogenetik Program	Kullandığı Metot
BATWING	Bayes Yöntemi
BAYESPHYLOGENIES	Bayes Yöntemi
BOSQUE	Distance, ML
FASTDNAML	ML
GENEIOUS	NJ, UPGMA
HYPHY	ML, NJ, Distance
IQPNNI	ML, NJ
MEGA	Distance, MP, ML
MESQUITE	MP, Distance, ML
MOLPHY	ML
MRBAYES	Bayes Yöntemi
NETWORK	MedianJoining, Reduced Median, Steiner Network
NONA	MP, Ratchet
PAML	ML
PAUP*	MP, Distance, ML
PHYLIP	MP, Distance, ML
PHYLOQUART	Quartet Method
QUICKTREE	NJ
RAXML-HPC	ML, MP
SEMPHY	ML, NJ
TNT	MP
TREEGEN	Distance
TREEALIGN	Distance, PM
TREEFINDER	ML, Distance
TREE-PUZZLE	ML
T-REX	Distance, MP, ML
WINCLADA	MP
XRATE	Dal Uzunluğu Tahmini
MVSP	Distance

Çizelge 2.4. Yaygın olarak kullanılan bazı analiz yöntemleri

UPGMA	NJ	MP	ML	BI
Dal Uzunluğunu Verir	Dal Uzunluğunu Verir	Dal Uzunluğunu Vermez	Dal Uzunluğunu Verir	Dal Uzunluğunu Verir
Atasal Kök	Atasal Köksüz	Atasal Köksüz	Atasal Köksüz	Atasal Köksüz
Moleküler Saat	-	-	Evrimsel Modeller	
Hızlı Analiz	Hızlı Analiz	Yavaş	Yavaş	Hızlı Analiz
Divergen Diziler için İyi Sonuç Vermez	Divergen Diziler için İyi Sonuç Vermez	Yakın Diziler İçin Uygun Sonuç Verir	Yakın Diziler İçin Uygun Sonuç Verir	Yakın Diziler İçin Uygun Sonuç Vermez
Zayıf Güvenilirlik	Oldukça Hızlı ve Güvenilir	Orta Düzeyde Güvenilir	Oldukça Yüksek Düzeyde Güvenilir	Oldukça Yüksek Düzeyde Güvenilir

Moleküler filogenetik analizleri karakter tabanlı ve numerik tabanlı analizlerden oluşabilmektedir. Numerik analizlerde DNA markır analizlerine göre (varlık veya yokluk markırları) yapılan analizler olup bu tip analizlerde uzaklık matrisleri kullanılmaktadır. Numerik metotlarda yaygın olarak kullanılan yöntemler ise Neighbor Joining (NJ), Unweighted Pair Group Method with Arithmetic Mean (UPGMA) metotlarıdır.

Karakter tabanlı analizlerde morfolojik karakterler ve moleküler karakterler (nükleik asitlerde ve aminoasitlerde insersyon, delesyon, substitüsyon) kullanılarak yapılan analizler olup bu tip analizler yaygın olarak kullanılmaktadır. Karakter tabanlı analizlerde yaygın olarak kullanılan başlıca analiz yöntemleri ise maksimum ihtiyat analizi (Maximum Parsimony, MP), maksimum olası analizleri (Maximum Likelihood, ML) ve Bayes (Bayesian) analizleridir. Bu temel analiz metotlarına değişik istatistiksel ve algoritmalar eklenerek çeşitlendirilmişlerdir (Swofford 2002, Helaers 2010).

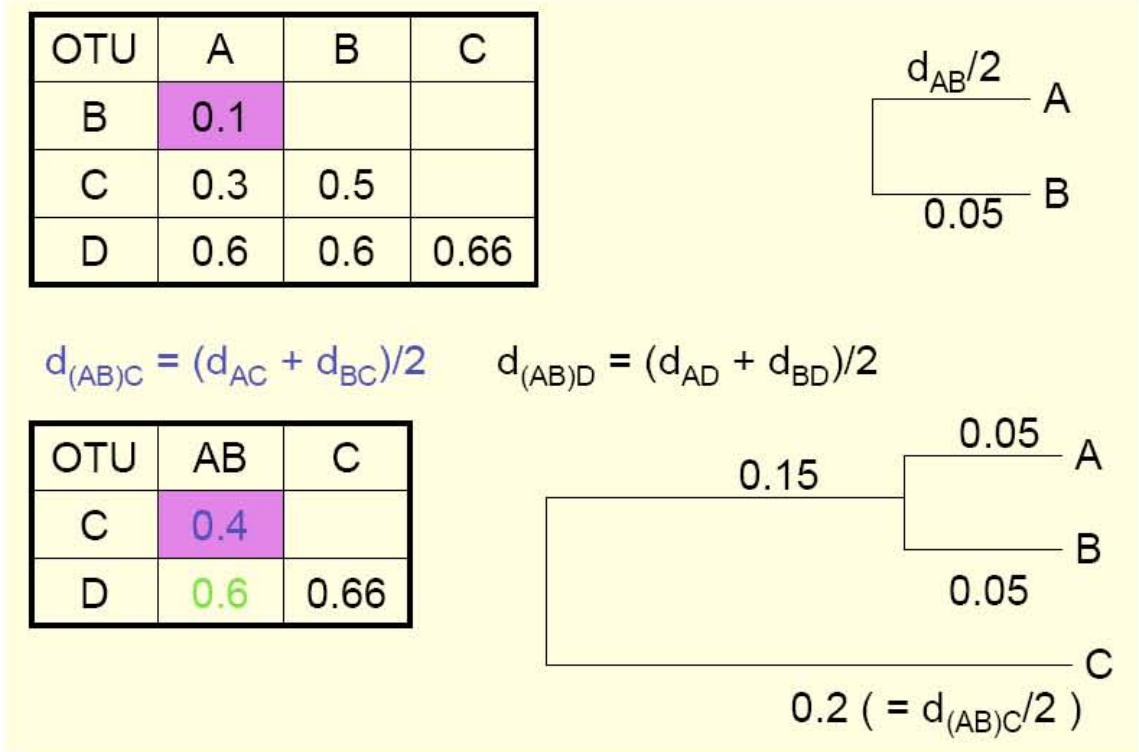
Analiz metodu ne olursa olsun filogenetik ağaçların ve topolojilerin elde edilmesinde ve yorumlanmasında başlıca iki temel işlem bulunmaktadır. Bu işlemler ağaç topolojisinin tahmini olarak nitelenen bir ağacın dallanma patterni ve diğeri ise dal uzunluklarının tahmin edilmesidir. Topoloji bilindiği zaman dal uzunluklarının tahmini veya belirlenmesi değişik istatistiksel yöntemler kullanılarak kolay bir şekilde belirlenebilmektedir. Filogenetik çalışmalarda zor olan topolojinin belirlenmesidir. Örnek sayısının artması olası topoloji sayısını oldukça yüksek rakamlara çıkarmakta ve bu durumda doğru olan topolojiyi belirlemek oldukça zorlaşmaktadır. Genel bir kural

olarak örnek sayısı n ile gösterilir ve olası toplam köklendirilmemiş topoloji sayısı = $[(2n-5)!$ olduğundan çok sayıda topoloji oluşturulmaktadır.

Yaygın olarak kullanılan UPGMA, NJ yöntemleri numerik metotlar olarak adlandırılırken dizi tabanlı yöntemler ise ML, MP, Bayes ve diğer olasılık yöntemleridir. Moleküler filogenetik analizlerinde kullanılan yöntemlere bağlı olarak değişik varsayımların kabulü zorunludur. Bu varsayımlara örnek olarak; i) karşılaştırılan dizilerin homolog diziler olması şarttır. İki dizinin birbirlerine homolog olup olmadıkları aralarında BLAST yazılım programının da yer aldığı birçok biyoinformatiksel programlarla protein ve DNA dizilerinde belirlenebilmektedir, ii) birden fazla dizi birbirlerine hizalanabilmektedir (“multiple alignment”), iii) karşılaştırılacak dizilerin aynı tipte dizilerden (örneğin nükleik asit dizileri, aminoasit dizileri) oluşurlar, iv) karşılaştırmaya alınan diziler birbirlerinden bağımsız olarak örneklenmiştir, v) karşılaştırılan diziler örneğin genler veya proteinler farklı olarak evrimleşmişlerdir, vi) nükleik asit ve aminoasit dizileri içerisinde bulunan karakterler birbirlerinden bağımsız evrimleşmiştir.

Moleküler evrimsel genetikte moleküler saat kavramı (MS) oldukça önemlidir ancak moleküler saat kavramının doğruluğu onlarca yıldır tartışılmaktadır. Moleküler saat hipotezine göre belirli bir gen ve özellikte değişim zaman sürecinde belirli bir değişim süreci gösterir ve organizmanın geçmişi ile günümüzdeki durumu zaman içerisinde belirli bir düzende değiştiği kabul edilmektedir.

MS kavramı değişik programlar tarafından kullanılmaktadır, örneğin UPGMA yönteminde. UPGMA yönteminde MS hipotezi kabul edilir. MS hipotezinin doğru olduğu çalışmalarda bu yöntem kullanılmalıdır. UPGMA yöntemi MS hipotezine bağlı olması nedeni ile yaygın olarak kullanılmamaktadır. UPGMA yönteminde NJ yönteminden farklı olarak köklendirilmiş ağaçlar üretilir. UPGMA yönteminde bir matriks kullanılarak dendogramlar üretilir. Bu yöntem morfolojik, biyokimyasal, numerik ve dizi analizlerinde kullanılabilir. Bu yöntemde veriler birebir incelenir ve birbirine en yakın iki grup belirlenir ve daha sonra bu gruba diğer bir grup eklenerek bir üst gruba taşır (Şekil 2.1).



Şekil 2.1. UPGMA ağacında dal uzunlukları hesaplanması

İki grup arasındaki bütün uzaklıklar veya yakınlıkların ortalaması alınarak gruplamalar yapılmaktadır. Örneğin A örneği ile B örneği arasındaki uzaklık 0.1 ve A ile B ağaçta en sondaki yaprak ise, UPGMA ağacında A'dan köke 0.05 ve B'den köke 0.05 ve A'dan B'ye dal uzunluğu 0.1 dir. UPGMA gruplama yönteminde uygulanan yaklaşım aşağıdaki tabloda örneklendirilmiştir. Çok sayıda taksonun kullanıldığı örneklerde önce en yakın olan gruplar belirlenir ve daha sonra bu gruplar birbirleri ile karşılaştırılarak ağaç formunda dönüştürülür.

Moleküler filogenetik çalışmalarda karakterlerin evrimsel değişiminin eşit olduğunu varsayan ve eşit olmadığını varsayan yaklaşımlar kullanılmaktadır. Örneğin Jukes ve Cantor yöntemlerinde değişimlerin eşit oranda olduğu varsayılmaktadır. Bazı karakterler diğerlerine oranla daha fazla değişim göstermektedir. Bu durumda karakterler arasında dengenin sağlanması amacı ile ağırlık yapılabilir. Genel olarak transisyon (adeninden guanine veya guaninden adenine) transversiyona (guaninden sitozine, timinden adenine) göre ağırlıklandırılır. Örneğin F84, Kimura ve diğerleri nükleik asit analizlerinde ağırlık kullanılmaktadır, diğer taraftan PAM (Point

Accepted Mutations) yönteminde ise aminoasit dizi analizlerinde ağırlık yöntemlerini kullanılmaktadır.

Algoritmik (numerik) metotlardan NJ ve UPGMA yaygın olarak kullanılmakta olup bu tip yöntemlerde tek bir ağaç özel bir algoritmaya bağlı olarak belirlenir. Algoritmik metotlardan farklı olarak Eniyilik kriterlerinin (“optimality criterion”) kullanıldığı yöntemler ise MP, ML ve Bayes metotlarıdır. Eniyilik kriterlerinin kullanıldığı filogeni yaklaşımlarıyla elde edilen topoloji ve ağaçlar aynı olmamaktadır. Bu durum özellikle karmaşık filogenetik çalışmalarda daha belirgindir. Birden fazla ağacın elde edilmesini önleyen veya birden fazla ağaç içerisinde doğruya daha yakın olan ağaçların belirlenmesi için değişik yöntemler geliştirilmiştir. Genel olarak başvuru yöntemlerden simülasyon, yöntemlerinden yararlanılarak substitüsyon matriksleri ile gerçek verilerin karşılaştırılmaktadır. MP ve ML yöntemlerinde substitüsyon matriksleri oluşturulmakta ve daha sonra bu matriksler yardımı ile gerçek veriler analiz edilmekte ve filogenetik ağaçlar oluşturulmaktadır. Distance yöntemlerinde ise Distance matriksi oluşturulmakta ve daha sonra filogenetik ağaçlar oluşturulmaktadır.

Uzaklık Metotları (“Distance Method”) yardımıyla oluşturulan topolojilerde üzerinde çalışılan dizilerin kendileri ve bu dizilerin elde edildiği genler ve bu genleri taşıyan organizmaların özellikleri oldukça etkili olmaktadır. Uzaklık metotlarının kullanıldığı analizlerde eğer diziler arasında ve diziler üzerinde farklı düzeylerde transisyon, transversiyon farklı düzeylerde gerçekleşiyor olması veya delesyon ve insersiyonun farklı düzeylerde gerçekleşiyor olması veya substitüsyonların farklı düzeylerde gerçekleşmesi nedenleri ile elde edilen topolojiler gerçeği yansıtmayabilmektedir.

Bunlara ek olarak alignment işlemlerindeki uyumsuzluklar veya hizalama işlemlerinin gerçeği yansıtmaması durumlarında da uzaklık tabanlı filogenetik çalışmalar güvenilir sonuçlar ortaya koyamamaktadır. Bu gibi durumlardan dolayı genel olarak uzaklık tabanlı filogenetik çalışmalarda cinsler veya türler arasındaki ilişkiler orta düzeyde saptanılabilirken yakın tür veya çeşitler arasında sorunlar yaşanabilmektedir.

Uzaklık metotlarında evrimsel uzaklık belirlenebilmektedir. Uzaklık metotlarındaki evrimsel uzaklık kavramı ise belirli bir gen veya DNA fragmanında veya aminoasit dizilerinde var olan değişikliklerdir. Diziler arasındaki değişiklikler karakter farklılığı (substitüsyon), karakter ekleme (insersyon) veya karakter çıkarma (delesyon) olayları sonucunda oluşmaktadır. Uzaklık yöntemleri ile oluşturulmuş bir filogenetik topoloji, dizilerin birebir karşılıklı karşılaştırmalardan elde edilen veriler bir algoritma veya optimizeleştirme işlemi kullanılarak elde edilmektedir. En yaygın kullanılan algoritmalara örnek olarak unweighted pair-group method with arithmetic averages (UPGMA) Fitch-Margoliash (FM) ve neighbor-joining method (NJ) verilebilir. Optimizasyon yöntemlerinden ise yaygın olarak kullanılmamakla birlikte minimum evrim metodudur. Uzaklık tabanlı filogenetik ağaçlar ise hesaplanmış olan evrimsel uzaklığın kriter olarak kullanılmasıyla elde edilmektedir. Çok sayıda kıstas veya kriter olmasına rağmen çoğu kriter genel olarak minimum evrim (ME) veya en düşük kareler (LS) kriterlerinden türetilmiştir.

En Düşük Kareler Yöntemine [Least-Squares (LS)] göre seçilmiş ağaç topografyasında seçim birebir (pairwise) uzaklık hesaplarından elde edilen verilerden en düşük kareler toplamını vermesi prensibine göre gerçekleştirilmektedir. Bu tür topoloji bulma prensipleri DNA ve protein dizilerinde uygulanabilmektedir. LS yönteminde değişik modifikasyonlar bulunmakta olup bunlar ağırlıklı LS metodu veya genelleştirilmiş LS metodu olarak adlandırılmaktadır. LS yöntemiyle elde edilen topolojilerdeki dal uzunluklarında bazı durumlarda negatif değerlerin elde edilmesi bu yöntemin en büyük dezavantajıdır.

Minimum Evrim Metodunda (Minimum Evolution (ME)) yine birebir karşılaştırmalı uzaklık değerleri elde edilir ve bir ağacın dal uzunluğu belirli algoritmalar kullanılarak tahmin edilmektedir. Bu yöntemde toplam dal uzunluğu en az olan ağaç topografyası seçilir. Bu yöntemde dal uzunluklarının hesaplanmasında aralarında LS metodunun bulunduğu farklı metotlarda kullanılmaktadır. LS yönteminde olduğu gibi ME yönteminde de elde edilen ağacın gerçek ağaç olduğu kesin değildir. ME yönteminin LS'de olduğu gibi negatif değerler vermesi değişik algoritmalarla önlenmiş olmasına

karşın bu modifiye yöntemlerin büyük zaman alması ve güçlü analiz işlemlerinin gerektirmesi gibi dezavantajlar bulunmaktadır.

Neighbor Joining (NJ) yöntemi 1973 yılında Sneath ve Sokal tarafından oluşturulan rakamsal gruptama analizi olup daha sonra Saito ve Nei (1987) tarafından geliştirilerek filogenetik çalışmalarda kullanılmaya başlanmıştır. Bu yöntemde ilk olarak bütün dizi veriler birebir olarak karşılaştırılarak bir uzaklık matrisi oluşturulur ve NJ algoritması kullanılarak ağaç oluşturulur. Bu yöntemde moleküler saat hipotezi kullanılmaz ancak ME yönteminin bir türevi olan FM veya diğer yaklaşımlar kullanılır. Bu yöntemle alignment sonucu oluşturulmuş diziler doğrudan kullanılarak gruptama gerçekleştirilir. NJ yönteminde analiz başında içsel dalların uzunluğunun sıfır olarak kabul edilir ve analiz sırasında dallar arasında en küçük uzunluklar birleştirilerek analiz devam edilir. Örneğin eğer 10 takson varsa ($N=10$) maksimum dal sayısı toplam örnek sayısının bir eksigidir ($N-1$). Uzaklık matrisleri genellikle FM yöntemi kullanılarak hesaplanır. Örnek sayısı ve örneklere ait verilerin büyüklüğü ile NJ topolojisi arasında bir ilişkinin varlığı ve örnek sayısının artması ile NJ topolojisinin doğruluğu arasında bir ilişki tespit edilmiştir. NJ metodu bazı araştırmacılara göre birkaç potansiyel topoloji yerine tek bir ağaç üretmekte ve ağaçtaki dalların istatistiksel olarak desteklenmesi durumunda güvenilir sonuçlar elde edilebilmektedir. Ancak NJ topolojisinde istatistiksel olarak desteklenmeyen dalların fazlalığı durumunda diğer metotlar kullanılmalıdır. Yaygın olarak NJ metodundan elde edilen ağaçların doğruluğu bootstrap yöntemiyle belirlenmektedir.

Evrimsel uzaklık hesaplamaları ve oluşturulan topolojilerin örnekler arasındaki değişimlerin bilinmesi durumunda doğru olarak hesaplanmaktadır. Ancak gerçekte bu değişimler tam olarak bilinmez ve bunların hesaplanması için değişik tahmin yöntemleri geliştirilmiştir. Değişik araştırmacılar tarafından değişik yöntemler geliştirilmiş olup bütün bu yöntemlerin ortak amacı gerçekte var olan değişimleri (ekleme, çıkarma veya değiştirme) ortaya koymaya yöneliktir. Örneğin Kimuar veya Hasegawa tarafından geliştirilen yöntemlerde değişimin hesaplanması (d mesafesi) basit orantısal hesaplama (düzeltilmemiş p mesafesi) yöntemine göre daha doğru sonuç vermektedir. Mesafe d hesaplamalarında ortaya çıkan yüksek varyasyon Jukes-Cantor mesafesi ile daha da

azaltılmıştır. Diğer bir söylemle doğru topoloji daha doğru olarak ortaya konabilmektedir.

Maksimum parsimoni metodu [Maximum Parsimony (MP)] yöntemi ilk olarak Eck ve Dayhoff 1966 tarafından kullanılmıştır MP yönteminde filogenetik farklılık en küçük karakter farklılığına dayanmakta ve böylece elde edilen veya tahmin edilen filogeni nükleik asit veya aminoasit değişimlerini yansıtmaktadır. Algoritmik yöntemlerin aksine MP yönteminde bir istatistiksel özet ile değil de, filogenetik ilişki karakter farklılığına dayanmaktadır. MP yönteminde elde edilen ağaç en kısa ağaç ve taksonların ise birbirlerinden en az farklı olanları birleştirilmektedir. Ancak elde edilebilecek ağaç veya topoloji sayısı dizi uzunluğu ve özellikle örnek sayısına bağlı olduğundan, en iyi ağacın bulunması hem zaman açısından uzun bir süre gerektirmekte hem de elde edilen ağacın doğruluğunun kesin olmaması durumları bu yöntemin dezavantajıdır. Bu yöntemde yaygın olarak dal değişim yöntemi kullanılmakta olup bu yaklaşımda başlangıç aşamasında bir ağaç tahmin edilmekte ve sonra bu ağaç iki dala ayrılmakta ve sırasal eklemeli olarak yenileri eklenerek en uygun ağaç belirlenebilmektedir.

Ağaç arama algoritmalarında genel bir kural olarak en az varsayımı kullanan yöntemin daha doğru olduğu düşünülür. Bu nedenle varsayımı en az olan algoritmalar tercih edilmektedir. Örneğin ağaç ayırma ve tekrar tanıma (tree bisection and reconnection) bu tip algoritmalara örnektir. Genel olarak ağaç arama yöntemlerinde ilk elde edilen ağaç sonraki ağaçların oluşumunda büyük öneme sahip olması nedeni ile genellikle tekrarlamalı aramalar (addition sequence replicates) kullanılmaktadır. Bu yöntemde nükleik asit dizisi veya aminoasit dizisi kullanılarak belirli varsayımlar altında filogenetik ilişkiler incelenmektedir. Bu analizde kullanılan varsayımlarda nükleik asitlerde ve aminoasitlerde her yöne mutasyonun varlığı dikkate alınmakta ve en düşük değişim göz önüne alınarak olası topolojiler hesaplanmaktadır. Mümkün olan olası topolojiler arasında en az değişimi gösteren topoloji gerçek ağaç olarak kabul edilmektedir.

MP yöntemi kullanılan diziler arasında, dizilerde çok sayıda varyasyonun bulunmaması ancak yeterli bir düzeyde değişimin bulunması durumunda doğru topolojiyi tespit edilebilmektedir. Pratikte nükleik asitlerde oluşan değişimler geriye

veya paralel deęişimler (transisyon veya transversiyon) bulunmaktadır. Bu durumun varlığı MP analizinde sorunlar oluşturmaktadır. Ancak filogenetik analizler arasında MP ekleme ve çıkarmaları (indels: insersiyon ve delesyon) en iyi deęerlendiren bir topoloji yöntemidir. MP topolojilerinin belirlenmesinde yaygın olarak iki algoritma kullanılmaktadır. Bu algoritmalarından biri dal-ve-sınır olarak adlandırılan (branch-and-bound) yöntemdir. Bu yöntemde çok sayıda topoloji ayrı ayrı analiz etmek yerine dal uzunlukları kısa olan topolojiler kullanılmaktadır. Dięer bir yöntem ise deneyimsel (heuristic) metot olup çok sayıda alt parametrelerden oluşmaktadır. Deneyimsel metotla birlikte dal-deęişimi (branch swapping) genetik algoritmalar, (genetic algorithms), simülasyon bağlamaları (simulated annealing) kullanılmaktadır.

MP yönteminde dal deęişim algoritması kullanıldığında başlangıçta bir ağaç tahmin edilmekte ve bu ağaca basamaksal olarak diziler eklenmekte ve burada oluşturulan ağaçlar alt ağaçlar olarak birbirleri üzerine yerleştirilir. Doğal olarak bir analizde ne kadar az varsayım kullanılırsa o analiz dahi iyi bir analizdir. Algoritmalarından ağaç-ayırımı ve tekrar bağlama olarak adlandırılan (“tree bisection and reconnection”) yönteminde az sayıda varsayım kullanılmaktadır. Bu tür yöntemlerde en iyi topolojiyi belirlemede ağaç arama başlangıçta varsayılan veya kurulan ağaca bağımlı olacağından ağaç aramaları çok sayıda tekrarlanmaktadır. Başlangıç noktasına eklemeli tekrarlamalar (“addition sequence replicates”) olarak isimlendirilmektedir. Filogenetik çalışmalarda MP analizleri ile oluşturulan topolojilerde bazı durumlarda birden fazla topoloji elde edilebilmektedir. Bu ağaçların birden fazlasında en kısa dal uzunluğu bulunabilmektedir (“multiple equally-parsimonious trees”). Bu gibi durumlarda araştırmacılar bir kesin konsensus ağaç veya analizlerde ortak olarak bulunan (“a majority rule consensus”) ağaçlarını kullanabilmektedir.

Genel olarak MP analizleri transisyon ve transversiyon durumunda ve diziler arasında yüksek deęişim oranları bulunan örneklerin analizinde zorluklar ortaya koymaktadır. Gerçekte bu farklı deęişim oranının varlığı söz konusudur ancak MP analizi bu deęişimlerin eşit olduğunu varsayması nedeniyle sonuçlar farklı topolojiler ortaya koyabilmektedir. Bu durumu ortadan kaldırmak amacıyla transversiyon veya genomdaki daha az deęişimler içeren bölgelere daha ağırlık vermek veya daha hızlı

değişim gösteren genomik bölgelere daha az ağırlık vermek şeklinde ortadan kaldırılabilir. Ancak çoğu durumda üzerinde çalışılan dizilerin hızlı veya yavaş değişim bölgelerini teşhis etmek mümkün değildir ancak dinamik ağırlıklı parsimoni yöntemleri geliştirilmiştir (“Dynamically Weighted Parsimony”). MP analizleri ile elde edilen ağaçlar minimum sayıda aminoasit veya nükleik asit dizilerindeki değişimi gösterir. Parsimoni analizinde iki temel karşılaştırma bulunmaktadır. Birinci yaklaşım Henning (1966) tarafından geliştirilmiş olup bu yöntemde verilerin paylaşılmış karakterler ve elde edilmiş karakterler olmak üzere iki gruba bölünür ve sadece elde edilmiş ortak karakterler (“synapomorphs”) parsimoni analizi için kullanılır. Ancak bu yaklaşım moleküler verilerde kullanılamaz çünkü normalde atasal ve elde edilmiş (türetilmiş) karakterler bilinmez. Parsimoni yöntemlerinde homoplasi, evrimsel sürecin dönüşümlü (“unrooted”) veya dönüşümsüz (“rooted”) ve çoklu karakter durumları göz önüne alınarak alternatif yöntemler geliştirilmiştir.

Parsimoni yöntemlerine örnek olarak Camin ve Sokal parsimoni metodu, Dollo Parsimonisi ve Fitch-Wagner Parsimoni verilebilir. Camin ve Sokal parsimoni metodu tipik bir parsimoni metodu olup Camin ve Sokal (1965) tarafından geliştirilmiştir. Bu yöntemde atasal durum mutlaka bilinmesi gerekmekte ve tek bir karakter değişimi vardır ve bu değişim sadece tek bir yönedir. Bu yöntem geri dönüşümü olmayan küçük DNA delesyonlarının varlığı durumunda filogenetik çalışmalarda kullanılmaktadır. Ancak dizi analizlerine uygun bir yöntem değildir. Dollo parsimonisi yöntemi Dollo yasasına dayanmaktadır. Bu metoda göre dönüşmüş bir karakter tekrar dönüşmez diğer bir söylemle evrim geriye dönüşmez. Bu yöntem restriksiyon enzim fragman analizlerine uygun bir yöntemdir. Fitch-Wagner Parsimoni yöntemde geriye dönüşümlü evrimsel süreçte kabul edilir ve analizler bu yönde gerçekleştirilir. Bu yöntem dizi analizlerinde kullanılabilmektedir. Bu metot PHYLIP paketinde bulunan DNAPARS ve DNAPROT programlarıyla veya diğer bazı programlarla gerçekleştirilebilmektedir.

Maksimum Olası [Maximum Likelihood (ML)] metodu parametrelerin tahmininde kullanılan gelişmiş bir istatistiksel yöntemdir. Filogenide ML’nin kullanımı ilk olarak Cavalli-Sforza ve Edwards (1967) tarafından öngörülmüş ve daha sonra J. Felsenstein (2002) tarafından PHYLIP ve DNAML program paketiyle birlikte moleküler dizilerde kullanılmıştır. ML ile ilgili değişik yöntemler FastDNAmI programında bulunmaktadır. ML yönteminde en yüksek olasılıklı ağaç substitüsyon modellerinin kullanılmasıyla elde

edilmektedir. DNA ve protein dizilerinde ML yöntemi yaygın olarak kullanılmaktadır. Örneğin MOLPHY protein filogenetik analizlerinde kullanılan yaygın bir ML metodudur. Bu yöntem belirli bir olasılıkta en düşük varyasyonu fazla sayıda örneğin kullanılması durumunda ortaya koymaktadır.

Filogenetik çalışmalarında ML yöntemi kullanıldığında en iyi ağaç topolojisini elde etmek için en yüksek (maksimum) olasılık kullanılır. Ancak farklı topolojiler farklı olasılık düzeylerinde elde edilebilir ve bu nedenle elde edilen en yüksek olasılıklı topoloji bile bazı durumlarda doğru topolojiyi ortaya koyamamaktadır. ML metodunun amacı olası mümkün bütün topolojiler içerisinde en yüksek olasılığa (maksimum olasılıklı) sahip topolojinin belirlenmesine yöneliktir. ML analizlerinde analizlere başlamadan önce aminoasit veya nükleotidler için uygun bir modelin belirlenmesi gerekmektedir. Günümüze değin çok sayıda model geliştirilmiştir ve bu modellerin büyük bir kısmı *ijModelTest*, *PAUP* veya *MetaPIGA* gibi programlarda mevcuttur. Bu modellerin seçimi büyük önemlilik arz etmektedir. Modellerin belirlenmesinde değişik yöntemler kullanılmaktadır. Günümüzde değişik programlar kullanılarak uygun modeller belirlenebilmektedir. Örnek programlar olarak nükleik asitler için *Modeltest* veya *MrModeltest* verilebilir. ML yönteminin çok uzun süre alması dezavantajıdır ancak *GARLI* ve *RAxML* algoritmaları kullanılarak analiz süresinde kısaltmalar gerçekleştirilmektedir. ML yönteminde elde edilen topolojiler *Bootstrap* ve *Jackknife* yöntemleri ile incelenebilmekte, elde edilen dalların istatistiksel yönden doğruluğu belirlenebilmektedir.

Bayes Filogeni Tahmini (BI) yöntemi ML yöntemi ile tam olarak ortaya konamayan filogenetik ilişkilerin ortaya konmasında kullanılmaktadır. Bu yöntemde analizler olasılığı en yüksek olan bir *a priori* topoloji ile başlar ve ağaçlar Markov Zinciri Monte Carlo (MCMC) yöntemi kullanılarak simüle edilir ve en iyi topolojiler yakalanır ve posterior olasılıkları yüksek olan ağaçlar ağaç dosyalarına yazılır. Bayes yönteminde dikkat edilmesi gereken önemli noktalar;

- i) Çok sayıda yakın ilişkili diziler kullanılmamalı aksi takdirde çok kısa dallar elde edilmekte,

- ii) Birbirlerine çok uzak olan diziler kullanılmamalı aksi takdirde çok uzun dallar oluşarak “Felsenstein-zone” oluşmakta ve elde edilen ağaçların güvenilirliği azalmaktadır. Bayes yönteminin güvenilirliğini artırmak adına çok yakın diziler kendi aralarında bir dışsal dizi veya diziler (“outgroup”) analiz edilmelidir,
- iii) Birbirleri ile aynı olan diziler uzaklaştırılmalıdır. Mümkün olan bütün taksonlar analize dahil edilmelidir aksi durumda bir taksonun yokluğu bütün bir analizi yanlış kılabilir, bütünüdür,
- iv) Etkin takson sayısı ile yapılan analizlerde topoloji simetrik değildir. Türleşme ve yok olma bir denge halindedir. Kompleks topolojinin varlığı yüksek oranda türleşme veya yok olmayı göstermektedir.

BI metodu hizalanmış dizi verileri ile evrimsel bir model ve bir filogenetik ağacın MCMC ile birlikte filogenetik çalışmalarda kullanılmaktadır. Moleküler filogenetikte Bayes yaklaşımı ile türlerin filogenisi ve evrimsel değişimleri belirlenebilmektedir. Bayes-MCMC yöntemi ML yöntemi ile karşılaştırıldığında hız bakımından oldukça üstündür. ML yönteminde olduğu gibi Bayes-MCMC yönteminde de birden fazla topoloji elde edilmektedir. Bu durum ML ve Bayes-MCMC çok sayıda ağacın konsensus analizleri ile özetlenmesi ile sonuçlar yorumlanabilmektedir. Bayes-MCMC yönteminde özellikle “star” ağaç olarahtan isimlendirilen çok dallı ve gerçekte olmayan ağaç topolojilerinin oluşumu gibi dezavantaja sahiptir. Ancak bootstrap yönteminin kullanılması ile bu durumun saptanması sağlanabilmektedir.

Günümüze değin yapılan çalışmalarda NJ, UPGMA, ML, MP, BI en yaygın kullanılan yöntemlerdir ancak bu yöntemlerin dışında oldukça fazla diğer yaklaşımlarda bulunmaktadır. Filogenetik yöntemlerin birbirleri ile karşılaştırıldığında tutarlı, etkin ve güçlü sonuç vermesi yönlerinden karşılaştırılabilir. Filogenetik yönteminin bir evrimsel modelde tutarlı sonuç vermesi demek bu yöntemin verilerinden daha az etkilenmesi anlamına gelir. Filogenetik yöntem etkinliği çok veri kullanımında dahi analiz süresinin kısalığı ve doğru ağacı vermesidir. Filogenetik yöntem güçlüdür demek evrimsel modellerin öngördüğü varsayımların ihlali durumunda dahi doğru ağacı verebilmesidir (Hillis 1995).

Filogenetik yöntemlerin kullanırken yöntem için uygun bilgisayar platformunda (uygun işletim sisteminde), uygun veri formatında, uygun modeller altında gerçekleştirilmelidir. Filogenetik yöntemleri i) hesaplama hızı, ii) oluşturulan topolojinin tutarlılığı, iii) oluşturulan topolojinin doğruluk derecesi ve iv) dal uzunluklarının doğruluğu yönünden karşılaştırılabilir. Bu yöntemler hesaplama hızı yönünden karşılaştırıldığında NJ metodu bootstrap analizi dahil diğer yöntemler oranla hızlıdır. Diğer taraftan MP, LS, ME, BI ve ML metotlarında ağaçların elde edilmesi için bütün olası topolojiler incelenmektedir bu da analizin hızına oldukça yavaşlatmaktadır. Ancak bazı değişik ağaç arama algoritmaları geliştirilmiş olup bu yöntemlerin hızları artırılmıştır. Oluşturulan topolojinin tutarlılığı genel olarak kullanılan DNA dizisinin sayısının yeterli olması ve değişimlerin doğru bir şekilde tanımlanmasıyla NJ, ME ve LS metotları tutarlı topolojiler oluşturabilmektedir. Pratikte bu genelde var olmayan bir durum olduğundan dolayı iyi bir ağaç programında tutarlılık tam anlamıyla elde edilememektedir. Günümüzde en yaygın istatistiksel model testi NJ ve ME topolojileri için kullanılabilmektedir. Ayrıca kesin istatistiksel testler geliştirilmiş LS metodunda da uygulanmaktadır. Diğer taraftan ML topolojilerinde yaygın olarak kullanılan testler bulunmamaktadır.

Yapılan çalışmalarda yeterli sayıda genin kullanılmasıyla oluşturulmuş topolojilerin karşılaştırılmasında NJ, ML ve MP metotları arasında farklı sonuçların alınmadığı tespit edilmiştir. Bu çalışmalarda farklı algoritmalara gerek duyulmadan NJ, ML ve MP benzer sonuçlar üretmişlerdir. Ancak bazı genlerin filogenetik çalışmalarda kullanımı ile NJ, ML ve MP farklı sonuçlar üretmişlerdir. Bu durum bazı genlerin filogenetik çalışmalarda kullanımının daha avantajlı olduğu ve sonuçların daha güvenilir olduğunu ortaya koymaktadır. Yapılan çalışmalarda eğer uygun genlerin kullanılmaması durumunda yüksek düzeyde algoritmaların kullanılmasında dahi doğru topolojilerin elde edilemediği belirlenmiştir. Diğer taraftan yapılan çalışmalarda filogeni çalışmalarında kullanılan gen veya DNA fragmanların birbirlerine yakın olmaları durumunda daha güvenilir topolojilerin alınabildiği gözlenmiştir.

Çalışmalarda elde edilen ağaçların doğru ağaç olup olmadıkları analiz sonucuna göre belirlenemez ancak en iyi sonucu üretebilecek modellerin kullanılması ile en doğru ağacın elde edilebileceği düşünülmelidir. Bilgisayar simülasyon çalışmalarında iyi bir

topolojinin elde edilmesinde en önemli faktörün nükleik asit veya aminoasit dizilerinin büyüklüğünün oldukça önemli olduğu ortaya konmuştur. Bu nedenle filogenetik çalışmalarda kullanılacak gen dizi sayısının uzunluğunun önemi bir kez daha ortaya konulmuştur. Elde edilen topolojilerin istatistiksel olarak değerlendirilmesi iki kategori altında incelenebilmektedir;

1) Ağacın veya topolojinin güvenilirliğinin test edilmesi,

2) Aynı veriden elde edilen iki veya daha fazla ağacın topolojilerinin test edilmesidir.

Bir ağacın doğruluğunun test edilmesinde iç dalların test edilmesiyle ortaya konabilir. Bu yöntem özellikle uzunluk metotlarıyla oluşturulan ağaçların test edilmesinde kullanılan bir yöntemdir. İç dalların testi ile ilgili birkaç algoritma ve analizler geliştirilmiştir. Bu testlerde dal uzunlukları istatistiksel olarak analiz edilerek normal dağılım gösterip göstermemesine göre belirlenmektedir. İç dalların testinde kullanılan diğer bir test yöntemi ise bootstrap testidir. Bootstrap testinin birkaç farklı uygulaması bulunmaktadır. Bootstrap testinde karakterler veya diziler rastlantısal olarak belirli bir sayıda tekrar tekrar analiz edilir. Her bir dal uzunluğu belirli bir topoloji için analiz edilir. Bu analizde genellikle 1000 veya daha fazla oranda topoloji oluşturulur. Bu topolojilerden en yaygın olan topolojiler gerçek ilişkiyi ortaya koyan topoloji olarak düşünülür. Genel bir kural olarak bootstrap değeri 0.90 veya %90 değerinin üstündeki değerler güçlü değerler, 0.75 ile 0.89 arasındaki değerler ise orta düzeyde destek değerleri olarak kabul edilmektedir.

ML ağaçlarında mümkün olan topoloji sayısı oldukça büyüktür. Bu durum ML ağaçlarının istatistiksel olarak değerlendirilmesinde sorunlar ortaya çıkarmaktadır. ML ağaçlarındaki yüksek düzeydeki topolojilerden en doğrusunu seçmek genellikle bootstrap değerlerinden yararlanılarak gerçekleştirilmektedir. MP topolojilerinde en düşük sayıdaki değişimlerin belirlenmesinin zorluğu nedeni ile parametrik bir testin oluşturulması zordur. Bunun yanında Felsenstein bootstrap testi en yaygın olarak MP ile birlikte kullanılmaktadır.

Filogenetik çalışmalarda yaygın olarak kullanılmakta olan NJ, ME, MP ve ML metotları bir arada incelendiğinde simülasyon verilerine dayanılarak bu metotlardan

güvenilir sonuçların alınabileceği belirlenmiş olup burada en önemli unsurun çalışılan gen ve aminoasit dizilerinin sayısı ve lokus tipi olduğu tespit edilmiştir (Çizelge 2.5).

Çizelge 2.5. Filogenetik çalışmalarda yaygın olarak kullanılan moleküler markırlar

Organizma	Gen Lokasyonu (Sitozol/Mitokondri/Kloroplast)	Nüklear
Bakteri	16S, 23S rRNA, kodlayan diziler	
Fungus	12S rRNA	18S, 28S, ITS rRNA, <i>RPB</i> , <i>ATP6</i> , <i>EF1-α</i>
Bitkiler	<i>rbcL</i> , <i>atpB</i> , <i>matK</i> , <i>NDHF</i> , <i>psbA-trnH</i> , <i>rpl32-trnL</i> , <i>trnK-matK</i>	ITS, 26S rRNA, CT
Hayvanlar (Omurgasızlar)	mtCOI, mtCytB, mtCR, 16S rRNA	18S, 28S, ITS rRNA, <i>EF1-α</i>
Hayvanlar (Omurgalılar)	mtCytB, ND1, ND2, mtCOI, 12S, 16S rRNA	18S, 28S rRNA, RAG1, IRBP

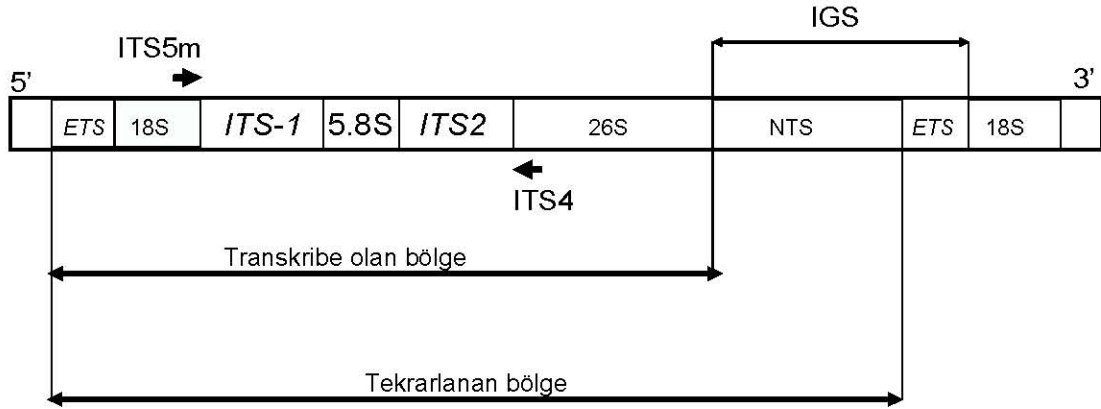
Günümüzde yaygın olarak kullanılan lokuslar (Çizelge 2.5) bitki moleküller filogenisinde kullanılan yaygın gen bölgeleri (kloroplast, mitokondri, nükleus) verilmiştir. Bu genlerin her birinin filogenide kullanılmasının farklı amaçlarının olmasının yanında filogenik olarak çözüm getirdikleri noktalarda farklı olabilmektedir. Bu genlerden bitki moleküller filogenisinde kullanılan en yaygın gen ITS olarak bildirilmiştir. Çünkü İnternal transkripsiyon sekansları (ITS) sayısız sistematik çalışmada ve çok geniş bitki çeşitinde cins ve tür seviyesinde kullanılmıştır.

Ökaryotlarda ITS iki iç boşluk ITS1 ve ITS2, 5.8S, 18S ve 26S nüklear ribozomal RNA (nrRNA) alt ünitelerini kodlayan genlerin arasında yerleşmiştir. ITS1 ve ITS2 boşlukları ve bunlara ek olarak 5.8S geni ITS bölgesi olarak adlandırılır (Şekil 2.2). ITS1 ve ITS2 bölgeleri ribozomal transkripsiyonel ünitenin birer parçası olsa da bu diziler olgun ribozomların yapısına katılmazlar. ITS bölgesinin her iki tarafında da korunmuş diziler vardır böylece, evrensel primerler kullanılarak çoğaltılıp dizi analizleri yapılabilir. ITS bölgeleri, nüklear ribozomal RNA'ların (nrRNAs) transkripsiyonlarında ve postranskripsiyonlarında etkili bölgelerdir.

rpoC1 geni ise, sistematikte ilgi duyulan bir diğer gendir. Çünkü şu ana kadar yapılan çalışmalarda; *Marchantia* (Koyun otu), black pine (siyah bananas) ve birçok angiospermde genelde bu genin 750 bp uzunluğunda tek bir intron tarafından ayrıldığı bulunmuştur. DNA sekansı bu intronun *Marchantia polymorpha* L. bitkisinde olduğu

açıklanmıştır (Ohyama vd 1986) Ama bu olay her bitkide görülmez. Örneğin çimlerde ve ağaç formunda bir alt familya olan cacti (*Cactoideae*) da bu intron kaybolmuştur.

matK geni önceden ORF K (“Open Reading Frame”) açık okunabilir çerçeve olarak biliniyordu. Bu gen pseudo gen olmaya doğru hızlı ve anlamlı nükleotid ve aminoasit substitüsyona sahiptir. Ancak moleküler çalışmalar, *matK* geninin ifade edilir ve fonksiyonel olduğunu göstermiştir. *matK* geni kloroplast genomunda büyük tek kopya şeklinde yer alır. *trnK* ekzonunda 5’-3’ arasında bir yerde yer alır. *matK* geninin yaklaşık uzunluğuve 1500 bp ve 500 adet aminoasit’e karşılık gelir. *matK* geninde her alanda bulunan nükleotid varyasyonu *rbcL* geninkinden üç kat daha fazladır (Soltis ve Soltis 1998). Aminoasit substitüsyonu ise *rbcL* de yaklaşık 6 kat daha fazladır (Olmstad ve Polmer 1999).



Şekil 2.2. ITS1 ve ITS2 bölgelerinin şematik görünüşü

matK geninde yüksek substitüsyon hızı, tüm üçlü kodonlarda neredeyse eşit hızda substitüsyonların olmasını sağlamaktadır (Hilu ve Liang 1997). Aynı zamanda bu gendeki yüksek substitüsyon hızı parsimonik olarak informatif (bilgi verici) bölgelerin sağlamasında ve güçlü filogenetik sinyallerin yanında birçok takson seviyesinde evrimsel geçmişin tahmin edilmesinde yardımcı olmaktadır (Johnson ve Soltis 1994). *matK* geninin sistematik çalışmalarda ki önemli görevinin yanına; kloroplast genomunda varsayılan intron II maturaz I kodlayan tek gendir.

Maturazlar prematüre RNA lardan intronları kodlayan nonkatalitik reaksiyonları kataliz eder. Yapılan son çalışmalar *matK* nın transkript ve translasyona uğradığını

göstermiştir. *matK* geninin oluşturduğu protein bitkiler için önemli olan ışık düzenlenmesinde görev alır (Barhet ve Hilu 2007).



Şekil 2.3. *matK* lokusunun şematik görünüşü

trnH-psbA aralığı bitkilerde barkot gen (markır gen) olarak önerilir. Bu barkotluk görevini hem yalnız başına hemde diğer sekanslar/genler ile beraber yerine getirir. Son çalışmalara göre *trnH-psbA* geni yakın türleri ayırt etmek için önemli bir ek kaynak oluşturmaktadır. *trnH-psbA* geni tersine çevrilmiş tekrar dizilerin yanında yer alır. Birçok bitki soyundaki *trnH-psbA* geninin özelliği barkot için istenen bir gen olduğu ortaya çıkmıştır. Bu genin uzunluğu genelde 500 bp den küçük olmakla birlikte bir bitki genomunda aynı anda birden fazla yerde görülebilir. *trnH-psbA* geninin tür içi yüksek sekans polimorfizmi ve çoğaltımında üniversal primerlerin kullanılabilmesi bu genin kolay bir şekilde amplifikasyonunun yapılmasını sağlar. Ancak bazı bitkilerde bitkilerde bu gen iyi bir şekilde amplifiye olamamakta veya çoklu bant verecek şekilde amplifiye olabilir. Zaman zaman normal uzunluğundan daha uzun sekanslarda verebilmektedir. Bunun nedeni muhtemelen mononükleotid tekrarlara sahip olduğundan doğru bir biçimde sekanslamasını zorlaştırıyor. Veya başka genlere ait insersiyonlar içermesi bazı gruplardaki yakın türleri ayırt etmek için *trnH-psbA* geni yeteri kadar değişken olmayabilir ama türler arasındaki varyasyonu çok yüksektir.

rpoB geni RNA polimerazın beta altbirimini kodlar. Hızlı evrimleşen ve yeterince korunmuş bir bölgedir. Bu gen grup I gene aittir *Mycobacterium* cinsinin filogenisini ortaya koymak için kullanılan *rpoB* geni (RNA polimerazın beta altbirimini kodlayan gen) bu cinsin filogenisini doğru ve güvenilir bir şekilde sonuçlandırmıştır (Kim vd 1999). Buna ek olarak *rpoB* geninin *Mycobacterium* cinsindeki bazı türleri örneğin *Mycobacterium kansasii* ve *Mycobacterium gastri* türlerini ayırt etmek için 16S rDNA'dan daha etkili ve kullanışlı olduğu bulunmuştur (Kim vd 1999). Filogenetik analizlerde bu genin (*rpoB*), rDNA'dan avantajlı kılan bazı nedenler vardır. Birincisi *rpoB* geninin tek kopyaya sahip olmasından ileri gelmektedir (Dahllof vd 2000).

İkincisi *Streptomyces* cinsi için doğrudan filogenetik analiz için bu genin (*rpoB*) sekansı üniversal olarak uygulanabilir niteliktedir. Ama buna karşılık 16S rDNA geninin birçok kopyaya sahip olduğu ve ender görüldüğü düşünülür. Bazı *streptomyces* türlerinde bu gene ait sekansa ait birçok farklı dizilim görülebilmektedir. Bu durumda belirsiz sonuçlar vereceğinden 16S rDNA bölgesinin direkt sekansının uygulanması mümkün değildir. İkincisi *rpoB* geni protein kodlamaktadır. Bu yüzden DNA sekansına ek olarak aminoasit dizilimine de sahiptir. Bu da *streptomyces* cinsinde bulunan bazı türlerin ayrıntılı analizlerinde kullanılabilir. Üçüncüsü *rpoB* geninin çoklu diziliminde (multiple alignment) boşluk ve fazlalıklar görülmemiştir. Bu da demektir ki *rpoB* sekansında herhangi bir boşluk silinmeden, *rpoB* genine ait bütün sekans verisi filogenetik analiz için düşünülebilir. Bu durum 16S rDNA için mümkün değildir.

Birçok çiçekli bitkinin kloroplast genomunda bulunan *accD* geni asetil-CoA karboksilaz geninin altünitesi olan beta karboksil transferaz enzimini kodlar (Kode vd 2005, Wakasugi vd 2001). Asetil-CoA karboksilaz asetil-CoA dan melonil-coA oluşumunu kataliz eder. Ayrıca bu genin yağasidi sentezini düzenleyen enzim olduğu düşünülür (Sasaki ve Nagano 2004). *accD* geni aynı zamanda fotosentetik olmayan bitkilerde de bulunur. *accD* geni hücre bölünmesi sırasında büyüme ve gelişmeye katkıda bulunur (Kode vd 2005). *Escherichia coli* de *accD* geni temel (essasiyel) bir gendir. *accD* geni nükleusa bağımlı RNA polimeraz tarafından transkripte edilir. Ters genetik çalışmaları sonucunda bütün bitkisindeki *accD* geni bitki gelişimi için temel bir gen olduğu belirlenmiştir (Kode vd 2005).

Filogenetik yöntemler kullanılırken üzerinde çalışılan organizmalardan elde edilen moleküler veriler arasında genden gene değişen bazı özelliklerin varlığı çalışmanın sonucu üzerinde oldukça etkilidir. Örneğin nükleik asitlerde bazı karakterlerin (nükleotidlerin) frekansları diğerlerine göre farklılık göstermektedir. Örneğin bir dizide yer alan belirli nükleik asit sayı veya aminoasit sayısı, substitüsyon sayısı, transversiyon ve transisyon farklılıkları bulunabilmektedir. Bu nedenlerden dolayı filogenetik çalışmalarda ilk işlem olarak uzaklık matrislerinin kullanılmasında yararlar bulunmaktadır. Günümüze değin filogenetik yöntemler kullanılarak yapılan çok sayıda çalışma bulunmaktadır.

Clegg vd (1994) yaptıkları bir çalışmada bitkilerdeki kloroplast genomu (cpDNA), bitki moleküller evrimi ve bitki moleküller sistematığını kapsayan çalışmalarda ilgi noktası haline geldiğini ve bu genomun birçok özelliğinden dolayı moleküler evrim analizlerini kolaylaştırdığını belirtmişler. Bu özelliklerden birincisi, küçük olması ve hücresel DNA'nın büyük bir miktarını oluşturmamasıdır. İkincisi, moleküler seviyede geniş bir biçimde karakterize edilmiş olması ve böylece karşılaştırmalı evrimsel çalışmaları destekleyecek temel bilgileri sağlıyor olması. Üçüncüsü, cp DNA'sındaki nükleotid substitüsyonların nisbeten yavaş olması, özellikle bitki filogenisinde derin bir evrimsel ilişkiyi ortaya koyan çalışmalarda bir çözüm penceresi oluşturduğunu vurgulamışlardır.

Schadt vd (1998) yaptıkları çalışma sonucu, maksimum likelihood modellerinin dört önemli avantajı olduğunu belirtmişlerdir. Birincisi; maksimum likelihood modelleri mevcut tüm veriler de kullanılabilir. İkincisi; gerçekçi evrimsel modeller tanımlanabilir. Üçüncüsü, parametrelerin ve ilgili standart hatanın tahminini yapar. Dördüncüsü ise; maksimum likelihood istatistiği boyunca, farklı evrimsel ağaçları birbiriyle karşılaştırabilir olmasıdır. Bu durumun çalışmalarında, tek bir nükleotid bölgesindeki substitüsyonları incelemek için, parametre zenginliği ve analitik kolaylaştırma arasında bir kolaylık sağladığını bildirmişlerdir. Bunun için Kimura'nın ihtimali tartışmalarını daha da genişleterek, Kimura'nın iki parametre olan modelini sekiz parametreye çıkarmışlardır. Ayrıca, Likelihood ve türevlerini aynı anda hesaplayan başka bir algoritma da ortaya koymuşlardır. Bu sonuçlara ulaşmak için uygulanabilir bir model olan HIV sekans verilerini ve RNA verilerini kullanmışlardır.

Bakker vd (2000) yaptıkları çalışmada *Pelargonium* (Geraniaceae) cinsinin tamamının filogenisini, mitokondrinin kodlanan *nad1* b/c ekzonunun ve kloroplastın kodlanan *trnL* (UAA) 5'exon-*trnF* (GAA) ekzonunun sekanslarından yararlanarak ortaya koymuşlardır. *Geranium* ve *Sarcocaulon vanderetiae* türlerini de dış grup olarak değerlendirmişlerdir. Çalışmalarında *nad1* bölgesinin b ve c bölgelerinin arasında bulunan intron II bölgesinin, *Pelargonium*, *Geranium*, ve *Sarcocaulon*, *Erodium* türlerinde bulunmadığını belirterek, mitokondrial bu bölgenin bu türlerde bulunmadığını kaydeden ilk çalışma niteliğinde olmuştur. Filogenetik analizlerini, mitokondri ve

kloroplast sekanslarını kullanarak hem ayrı ayrı hemde kombine bir şekilde ortaya koymuşlardır. Mitokondri ve kloroplast sekanslarına dayalı analizler kendi aralarında yüksek bir uyumlukta sonuç verdiğini görmüşlerdir ama kombine olarak yapılan analizde ise daha yüksek bir topolojik çözümleme ortaya koyduğunu kaydetmişlerdir. Ortaya çıkan topolojiler sonucu taksonlar, büyük ve küçük kromozom sayılarına göre iki temel gruba ayrılmışlardır. Filogentik analizlerini PAUP 3.1.1. (Swofford 1993), MacClade3.04 (Maddison ve Maddison 1992) ve Parsimony Jackknifer 4.22 (Farris vd 1996) biyoinformatik programlarını kullanarak sonuca ulaşmışlardır.

Kress vd (2002) yaptıkları bir çalışmada Nüklear ITS ve plastid *matK* gen bölgesini kullanarak *Zingiberaceae* familyasının kladogramını oluşturmuşlardır. Bu çalışma sonucunda yeni bir sistem önerdiler. Yeni sistem de *Zingiberaceae* familyası dört altfamilya ve altı soy (tribe)'den oluştuğunu bildirmişlerdir. *Siphonochiloideae* (*Siphonochileae*), *Tamijioideae* (*Tamijieae*), *Alpinioideae* (*Alpinieae*, *Riedelieae*), *Zingiberoideae* (*Zingibereae*, *Globbeae*). Aynı zamanda sonuç birçok genera olduğunu da gösterdi. *Alpina*, *Amomum*, *Boesenbergia*, *Curcuma*, *Etlingera* ve *Globba* cinslerinin monofiletik olmadığını savunmuşlardır. *Alpinieae* soyunu (tribe) oluşturan *Pomereschea* ve *Rhynchanthus*, *Zingibereae* soyunda (tribe) yer aldığını ortaya koymuşlardır. *Curcuma* cinsi üç gruptan oluştuğunu görmüşlerdir. Birinci grubun *Hitchenia* genusu ile kardeş olduğunu ikinci kardeş grubun *Smithatris* olduğunu ve son grup olarak ta *Stahlianthus* olduğunu öne sürmüşlerdir.

Williams vd (2004) yaptıkları çalışmada *Globba* cinsini ve *Globbeae* soyunu çalışmışlardır. Nüklear ITS bölgesini ve kloroplastın *trnK-matK* bölgelerinin sekansını kullanarak Parsimony ve Bayes algoritmaları ile değerlendirmiştir. Sonuçta, *Gagnepainia* ve *Hemiorchis* ve *Mantisia*'nın monofiletik olduğunu görmüşlerdir. *Gagnepainia* ve *Hemiorchis* in kardeş gruplar olduğunu ve geri kalanlarla da temel bir grup oluşturduğunu bulmuşlardır. *Mantis*'anın *Globba*'nın içerisinde yer aldığını, *Globba*'nın *Mantisia* bölümüne ("section") indirgediğini ve en sonunda *Globba*'nın *Ceratanthera*, *Globba* ve *Mantisia* adında üç ayrı altcins ayrıldığını bulmuşlardır. Aynı zamanda *Globba*'nın sekiz bölüme ayrıldığını da söylemişlerdir.

Lunter vd (2005) hesaplamalı biyolojide hizalama ve filogenetik ilişkinin ortaya çıkarması aşamalarının birbirinden bağımsız olmadığı, aksine birbirleriyle çok ilişkili olduğunu ifade etmişlerdir. Bunun için hem hizalama hemde filogeniyi birlikte tahmin eden Bayesiyen Markov ve Chain Monte metodunu geliştirmişlerdir. Bunu geliştirmelerinde iki algoritmadan faydalanmışlardır. Birincisi: “indeel peeling algorithm” adında insersiyon ve delesyon içeren filogenetik olarak bilgi verici (informative) olan evrimsel olgular teşkil eden bir algoritmadır. Aynı zamanda Felsenstein’nin peeling algoritmasına da benzer. İkinci algoritma: hizalama için etkili bir bağımsız, kısmi bir metropolize modelidir. Bu iki algoritmayı birleştirerek hizalama ve filogeni ikilisinin aynı anda düşünme problemi için tam bir Bayes birlikte tahmin prosedürünü çıkarmışlardır.

Duangjai vd (2006) yaptıkları bir çalışmada kloroplast genomunun *atpB*, *matK*, *ndhF*, *trnK* intron bölgesini, *trnL* intron bölgesini ve *trnL-trnF* bölgelerini kullanarak tropik olan *Ebenaceae* familyasını geniş bir şekilde araştırmışlardır. Yaptıkları örneklemeler *Ebenaceae* familyasında bulunan bütün cinsleri temsil edecek şekilde yapmışlardır. *Ebenaceae* familyası; *Diospyros*, *Euclea*, ve *Lissocarpa* cinslerden oluşur ama daha önce cins seviyesinde *Cargillia*, *Gunisanthus*, *Maba*, *Macreightia*, *Royena*, ve *Tetracelis* gibi cinslere sınıflandırılmıştır. Bu çalışmanın sonucunda güçlü bir şekilde *Ebenaceae*’nın monofiletik olduğu ortaya koymuşlardır ve çalışmanın sonucu şunu göstermiştir ki daha önce yapılan familyalar arasında sınıflandırmada monofiletik grup olmadığı belirlenmiştir. Ama yeni familyalar arası sınıflandırmada filogenetik yaklaşımlara dayanarak bu çalışma sonucunda *Ebenaceae* familyasının, *Lissocarpoideae* ve *Ebenoideae* adında iki tane alt familyaya ayrıldığı ve *Lissocarpa*, *Euclea*, *Royena*, ve *Diospyros* adında dört tane cinse ayrıldığını bulmuşlardır. Bu bulgularına maksimum parsimoni ve bayes algoritmalarını kullanarak aynı zamanda PAUP 4.0b10 (Swofford 2003) versiyonunu kullanmışlardır. Çalışmada kullanılan üç tane kodlanan gen bölgelerinin (*atpB*, *matK*, *ndhF*) filogenetik ilişkiyi, kodlanmayan *trnK* intronu, *trnL* intronu, ve *trnL-trnF* bölgelerinde daha net ortaya koyduğunu bulmuşlardır.

Hughes vd (2006) çalışmalarında yakın zamanda, *American Journal of Botany*, *Molecular Phylogenetic and Evolution*, *Systematic Botany and Taxon* gibi bilimsel

dergilerde son zamanlarda yayınlanmış olan 136 çalışma ve bu çalışmalar içinde ele alınan toplam 345 birey matrislerinden yararlanılarak çalışmalarını yapmışlardır. Özellikle 226 tane son çalışılmış tür düzeyindeki taksonlarda belirlenmiş matrisleri kullanmışlardır. Bu matrisler başlıca; çalışmaların amaçlarını, taksonların örneklemelerini, karakter örneklerini, kullandıkları sekansların DNA bölgeleri ve elde edilen çözüm derecesini, çözümleme noktasındaki başarının derecesini belirlemek olmuştur. Bu 226 bitkide kullanılan nrDNA, nDNA, kodlanmayan CpDNA karşılaştırılmasını da yapmışlardır aynı zamanda bu türler içerisinde aynı tür sayılabilenleri de tartışmışlardır. Sonuç olarak bitkilerde tür düzeyinde teşhis için en çok kullanılan kodlanmayan birkaç CpDNA gen bölgesi ve nrDNA bölgesidir. Bu lokuslar araştırılmış olan 345 matrisin %87'sini oluşturmuştur. Bu bölgeler düşük yani aşağı taksonomik seviyelerde değişiklik göstermeleri, aynı zamanda üniversal primerler ile çoğaltılması kolay olduğu için birçok çalışmada kullanılmıştır. Araştırılmış çalışmaların ancak %11'i tamamen çözümlenmiş filogenetik ağaçlardan oluşmuştur. Yapılan çalışmaların yarısından fazlası çok az çözümlenmiş olan filogenetik ağaç elde etmişlerdir. Sonuçta tür seviyesinde filogenetiği ortaya koyan bu çalışmaların toplamında kullanılan CpDNA lokuslarının parsimonik olarak bilgi verici karakter 0.2-8.5 arasında iken, ITS (nrDNA) için bilgi verici karakter yüzdesi 0.3-14 arasında olduğunu bulmuşlardır. Bu değerler, yüksek takson teşhisinin yapıldığı taksonlarda daha düşük olduğunu ve genelde hayvanlar için kullanılan mtDNA'sından da daha düşük olduğunu ifade etmişlerdir.

Ogden ve Rossenberg (2006) yaptıkları çalışmada filogenetik ağaç oluşturulmasında spesifik sekans hizalamasının ağacı oluşturan metotlardan çok daha önemli olduğunu vurgulamışlardır. Filogenetik sonuç için hizalamanın doğruluğunu göstermek için insersiyon ve delesyon içeren sekansların simülasyonunu yapmışlardır. Sonuç olarak veri setleri, ultrametrik eşit dal uzunluğu ve ultrametrik rastgele dal uzunluğu gibi farklı koşullar altında pektinat, dengeli ve rastgele ağaç şekilleri için simüle etmişlerdir. Hipotez hizalama ile doğru hizalama ("alignment")'nın karşılaştırılmasını yapmışlardır. Hizalama doğruluğunu ölçmek için iki kriterin olduğunu saptamışlardır. Birincisi, total veri seti ikincisi, filogenetik ağaçta bulunan birey dallarının olduğunu ifade etmişlerdir. Sonuçta hizalama yanlışları artarken topolojik doğruların azalmakta olduğu sonucuna varmışlardır. Bu eğilimi daha çok pektinat topolojilerden meydana gelen verilerde

bulmuşlardır. Buna karşılık dengeli ağaç şekilleri için olan veri setlerinde ultrametrik, eşit dal uzunluğu şekli ve hizalama yanlışları ağaç yapmak için çok az etkilerinin var olduğunu söylemişlerdir. Ayrıca dal uzunlukları ve komşu dalların uzunlukları arttığında hizalama doğruluğunun azaldığını ifade etmişlerdir.

Carroll vd (2007) yaptıkları bir çalışmada insersiyon ve delesyon için GOP (“Gap Open Penalty”) ve GEP (“Gap Extension Penalty”) cezalandırma boşluklarını kullanmışlardır. GOP ve GEP’in etkisine bakmak için altı tane nüklear ribozomal genlerden oluşan veri kullanmışlardır. Bu boşluk cezalandırmaların etkisini, garanti optimal maximum parsimony filogensini bulmak için geniş bir filogeni pametresi/ölçeği kullanmışlardır. Değişen bu parametreler sonucunda sadece değişik optimal filogeni meydana gelmemiş aynı zamanda muhtemel filogeni skorların dağılımında da etkili olmuştur.

Lahaye vd (2007) yaptıkları bir çalışmada Orta Amerika ve Güney Afrika da taksonomik olarak problemlili türler içeren iki büyük alandan tür koleksiyonu oluşturmuşlardır. 1600 den fazla tür içeren bu koleksiyon bireylerini 8 tane potansiyel barkod gen bölgesi kullanarak bir karşılaştırma gerçekleştirmişlerdir. Önceki çalışmalardan yararlanıp tür içi ve türler arası olması gereken barkod boşluğunun ne kadar olması gerektiğini değerlendirmişlerdir. Her türden birçok bireyi çalışmaya almışlardır. Sonuç olarak; polimorfizm bazında yeterince değişiklik gösterebilen, kolayca çoğalabilen ve hızlanabilen gen bölgesi kloroplast geni olan *matK* olduğunu bulmuşlardır. Bu yüzden *matK* gen bölgesini çiçekli bitkilerde barkod gen olarak tanımlamışlardır. Ayrıca biyoçeşitlilik araştırmacıları için barkod gen uygulanabilirliğini göstermek için, 1000 nin üzerinde Orta Amerika orkid türlerini analiz ederek ortaya koymuşlardır. Sonuç olarak *matK* geninin belirsiz türleri birbirinden ayırt edebilmiş ve uluslararası tehlike altında ki tür listesini çıkarmada kullanışlı bir barkod gen olduğunu ifade etmişlerdir.

Yue vd (2009) yaptıkları çalışmada çoklu hizalamayı ve en iyi hizalamayı gösteren ağacı aynı anda bulan yeni bir algoritma keşfetmişlerdir. Aynı zamanda C adında bir program paketi gibi bir algoritma daha ortaya çıkarmışlardır. Bu algoritma hem DNA hemde protein verilerini kullanabildiğini ifade etmişler. Ayrıca PAM250 ve

BLOSUM62 gibi kompleks substitüsyon matrislerini de alabildiğini vurgulamışlardır. Yeni metodun performansını daha önce yaygın bir biçimde kullanılan diğer çoklu hizalama yöntemleri (ClustalW ve T-Coffee) ile karşılaştırmışlardır. Sonuçta bu algoritmanın hem filogeni doğruluğu ve hizalama kalitesi bakımından çok daha iyi bir performans sergilediğini bulmuşlardır. Thomas (2010) yapmış olduğu bir çalışmada Doğu Asya'nın 540'tan fazla tür içeren *Begonia* L. (*Begoniaceae*) florasını araştırmıştır. Bu türlerde çok az apomorfik karakter olduğu için bu türlerin belirlenmesi ve karakterize edilmesi oldukça güç bir problem olarak görülmüştür. Şu an ki *Begonia* L. Biyocoğrafyasını, evrimini ve özel taksonomisini yapan çalışmalar hala eksik durumda olduğunu belirtmiştir. Bu çalışmada *Begonia* L. (*Begoniaceae*) florasının biyocoğrafyasını ve filogenetiğini ortaya koymak için 115 taksonda *ndhA* intron bölgesini, *ndhF-rpl32* arabölgesini ve 89 taksonda ITS ("internal transcribed spacer") nüklear ribozom bölgesinin sekans dizilimlerinden faydalanarak MP, ML, ve BI analiz yöntemlerini yardımı ile Doğu Asya *Begonia* L. florasının filogenetiğini düzenlemiş ve *Begonia* L.'nin monofiletik olduğunu belirlemiştir.

3. MATERİYAL ve METOT

3.1. Pamuk Taksonları ve ITS

Bu çalışmada, bitki moleküler filogenisinde kullanılan gen bölgeleri ve analiz metotlarını araştırmak için iki farklı bitki grubu ele alınmıştır. Birinci bitki grubu Afrika, Amerika, Avustralya, Asya orijinli ağaçsı, çalimsı ve otsu olmak üzere kırk dört pamuk türünün temsil edildiği 44 *Gossypium* taksonundan oluşturulmuştur. Aynı zamanda bu türler içerisinde ticari olarak dünyada %90-95 oranında üretimi yapılan *Gossypium hirsutum*, %5-8 oranında üretimi yapılan *G. barbadense* ve %2 oranında ticari üretimi yapılan *G. arboreum* türleri de yer almaktadır. Bu kırk dört bitki örneğine ait "Internal transcribed spacer" (ITS) gen bölgesine ait DNA dizileri, National Center for Biotechnology [NCBI (<http://www.ncbi.nlm.nih.gov/>)] Gen Bankasından indirilmiştir.

3.2. Dış Grup Olarak Kullanılan Taksonlar

Çalışmada dış grup olarak kullanılmak üzere *Origanum vulgare*, *Thymus vulgaris*, *Hibiscus surattensis* türleri kullanılmıştır. *Origanum vulgare*, ve *Thymus vulgaris*'e ait ITS dizileri laboratuvar çalışmaları ile belirlenmiştir.

3.3. Diğer Taksonlar ve Genler

Toplam on beş familyadan oluşan ve aralarında monokotil, dikotil, ağaçcı, otsu bitkiler bulunan altmış bitki taksonu bu grupta yer almaktadır. Bu bitki grubu, bitki moleküller filogenisinde yaygın olarak kullanılan kloroplast genom bölgelerinden maturaze geni (*matK*), RNA polimeraz geni (*rpoB*) ve (*rpoC1*), asetil CoA karboksilaz geni (*accD*), intergenik kloroplast geni (*trnH-psbA*) oluşturulmuştur. Bu gen bölgelerine ait sekanslar yine NCBI Gen Bank asından indirilmiştir.

3.4. Dış Grup Olarak Ele Alınan Taksonların DNA İzolasyonu

DNA izolasyonu Karaca vd 2005'e göre yapılmıştır. Buna göre aşağıda belirtilen aşamalarda DNA izolasyonu gerçekleştirilmiştir.

- 1) Her bir bitkiden 2-3 yaprak alınıp havan-havaneli yardımıyla bu örnekler sıvı azot ile toz haline gelen kadar iyi bir şekilde ezilmiştir. Her bir örnek için her bir havanda 100 mg PVPP önceden eklenmiştir.
- 2) Toz haline getirilen örnekler 2 mL'lik 2 ayrı steril tüpe paylaştırılmış ve üzerine 1000 µL EB (Ekstraksiyon Buffer) çözeltisi eklenmiş ve 3-5 dakika bir süreyle vorteks edildikten sonra 10000 xg de 10 dakikalık bir sürede santrifüj edilmiştir.
- 3) Tüplerdeki süpernatant (üst faz) dikkatlice dökülmüş ve alttaki pellet (kalıntı) 350 µL LB (Lezyon Buffer) çözeltisinde vorteks edilerek çözülmüş ve üzerine 350 µL 8M LiCl eklenerek tekrar vorteks edilerek iyi bir şekilde karışması sağlanmıştır.
- 4) Her bir tüplere daha sonra 25 µL. %20' lik SLS ("sodium lauryl sarcosinate") çözeltisi eklenmiş ve tüpler para film ile kapatılarak 65 °C'ye ayarlanmış bir su banyosunda 2 saat bir süreyle bekletilmiştir. Ancak her 20 dakika da bir 3-5 dakika vorteks edilmiş ve en son su banyosundan çıkartıldıktan sonra tekrar vorteks edilerek karıştırılmıştır.
- 5) Su banyosundan alınan her bir örnek üzerine 1:1 oranında CIS (Kloroform İzooamil Alkol 1:24) eklenmiş ve 10 dakika bir sürede oda sıcaklığında (25°C-28°C) bekletilmiştir.
- 6) Tüpler 10000 xg de 10 dakika bir sürede santrifüj edilerek 3 faz oluşturulmuş ve en üst fazdan (süpernatant) alınabildiği kadar hacimdeki (genellikle bu miktara 600 µL) solüsyon yeni bir tüpe alınmıştır.
- 7) Her bir tüpten alınan süpernatant miktarı kadar üzerine CIS eklenerek (600 µL) 10 dakika bir sürede oda sıcaklığında arada bir çalkalanarak bekletilmiştir.
- 8) Oda sıcaklığında (20 °C - 25 °C) bekletilmiş olan tüpler daha sonra 10000 xg de 10 dakikalık bir sürede santrifüj edilmiştir.
- 9) Üst faz temiz bir tüpe alınmış ve alınan miktar kadar üzerine izoproponal (2-proponal) eklenmiştir.
- 10) Her bir tüpe ilk alınan hacmin 1/20 si kadar 5 M NaCl eklenmiş ve 5 dakika bir süreyle oda sıcaklığında bekletildikten sonra 5000 xg de 5 dakika santrifüj edilmiş ve üst faz dikkatlice uzaklaştırılmıştır.

- 11) DNA pelleti üzerine 390 µL TE çözeltisi [(10 mM Tris ve 1 mM EDTA, pH 8.0)] ve 10 µL RNaz (10 µg) eklenmiş ve vorteks edilerek 30 dakika bir süreyle 37°C de bekletilmiştir.
- 12) Bu sürenin sonunda 40 µL 3 M sodyum asetat (NaAc) eklenmiş ve bu karışıma da 1.2 mL etanol eklendikten sonra karıştırılmış ve 5000 xg de 5 dakika santrifüj edilmiştir.
- 13) Süpernatant dikkatlice dökülmüş ve pellet 0.75 ml %80 etanol ile yıkanmış ve 5000 xg de 5 dakika santrifüj edilmiştir. Bu işlem en az 2 kez tekrarlanmıştır.
- 14) Süpernatant dökülmüş ve tüpler temiz bir havlu üzerinde ters çevrilerek etanol kalıntısı uzaklaştırıldıktan sonra DNA pelleti 0.2 mL pH: 8.0 TE'de çözülmüştür. Örnekler ileriki çalışmalara kadar +4 °C veya derin dondurucuda bekletilmiştir.

3.5. DNA Kalitesi ve Miktarının Belirlenmesi

Ekstraksiyonu tamamlanmış olan DNA'nın miktar ve kalitesi yaygın olarak kullanılan iki yöntem olan spektrofotometrik ve agaroz jel elektroforez yöntemleri kullanılarak belirlenmiştir.

3.6. Polimeraz Zincir Reaksiyonları

Polimeraz Zincir Reaksiyonları (PZR) hacmi 25 µL ye ayarlanmış ve içerisinde 120 nanogram toplam genomik DNA, 0.5 µM ITS primer çifti, 80 mM Tris-HCl (pH 8.8), 19 mM (NH₄)₂SO₄ (pH 8.8), %0.009 Tween-20 (g/h), 0.28 mM her bir dNTP, 3 mM MgCl₂ ve 2 ünite *Taq* DNA polimeraz (Bioron) içeren solüsyonda gerçekleştirilmiştir. PZR profilinde ise 94°C'de 3 dakika bekletilen tüpler önce 10 döngülük ön-PZR'de denatürasyon için 94°C'de 1 dakika, primer bağlanması için 50°C'de 45 saniye ve sentezleme aşaması için 72°C'de 2 dakikada süre uygulanmıştır.

Ön-PZR profilinin primer bağlama sıcaklığı döngü başına 0.5°C azaltılarak gerçekleştirilmiştir. Otuz döngülük ikinci PZR profilinde ise ön-PZR döngüsüne ait veriler kullanılmış ancak primer bağlama sıcaklığı 45°C olarak sabitlenmiştir. PZR çalışmaları tamamlanmış örnekler -20°C'de agaroz jel elektroforesis işlemlerine kadar bekletilmiştir.

3.7. Amplikonların Jelden İzolasyonu

Aşağıda belirtilen bütün işlemler Ince 2010'a göre gerçekleştirilmiş olup ana işlemler özetlenmiştir.

1. PZR ile çoğaltılmış ürünler steril bir şekilde metafor agaroz jel ile elektroforez edildikten sonra steril bisturi ucu ile kesilerek içerisinde 1 mL QX1 tamponu bulunan 1.5 mL'lik Ependof tüpleri içerisine alınmıştır. Nazikçe vorteks edilmiş ve 50°C'de 5 dakika bekletilerek jelin çözünmesi sağlanmıştır.
2. Jel çözündükten sonra içerisine hafifçe vorteks edilerek süspanse edilmiş olan QIAEX II çözeltisinden 10 µL eklenerek nazikçe karıştırılmış ve tüp 50 °C'de 10 dakika inkübe edilmiştir. Bu aşamada her 2 dakikada bir kez nazikçe vorteks edilmiştir.
3. Sürenin sonunda 12000 xg hızda oda sıcaklığında 1 dakika santrifüj edilerek süpernatant dikkatlice pipet kullanılarak uzaklaştırılmıştır.
4. Pellet üzerine 750 µL QX1 tamponu eklenerek nazikçe vorteks edilmiş ve 12000 xg hızda oda sıcaklığında 1 dakika santrifüj edildikten sonra süpernatant dikkatlice uzaklaştırılmıştır.
5. Pellet üzerine 750 µL PE tamponu eklenmiş ve nazikçe vorteks edildikten sonra 12000 xg hızda oda sıcaklığında 1 dakika santrifüj edildikten sonra süpernatant dikkatlice hiçbir kalıntı kalmayacak şekilde uzaklaştırılmış ve pellet beyaz renk alıp tamamen kuruyuncaya kadar oda sıcaklığında bekletilmiştir.
6. Pellet üzerine 50 µL TE pH 8.5 tamponu eklenmiş ve DNA'nın silikadan ayrılarak tampon içinde çözünmesi için oda sıcaklığında 15 dakika bekletilmiş ve tüp 12000 xg hızda oda sıcaklığında 1 dakika santrifüj edilerek süpernatant (saf DNA) dikkatlice yeni bir tüpe aktarılmıştır. Saflaştırmanın başarıyla yapıp yapılmadığını araştırmak için 5 µL DNA tekrar agaroz jelde yürütülerek varlığı teyit edilmiştir.

3.8. Transformasyon

E. coli JM107 bakteri ırkı ana stoğu hazırlanırken kurutularak dondurulmuş kompetan hücrelerin bulunduğu tüpe 0.5 mL sıvı LB-Sıvı (pH 7.0) kültür ortamı ilave edilmiş ve karıştırılmıştır. İki yöntemle bakteri ana stoğu çoğaltılmıştır. Birinci yöntemde LB-Katı (pH 7.0) plakanın (Petri kabı) üzerine hazırlanan kompetan hücre

stoğundan 20 µL ve 80 µL alınarak plakamın 5 bölgesine pipetle yerleştirilmiş ve L-Baget yayma çubuğu ile solüsyon plaka üzerini tamamen eşit oranda kaplayacak şekilde dağıtılmıştır.

Dağıtma işleminin ardından 100 µL LB-SIVI (pH 7.0) kültür ortamı plaka üzerine eşit oranda eklenmiş ve tüm yüzeye yayılmıştır. Bu şekilde hazırlanmış olan plakalar sıcaklığı 37°C'ye ayarlanmış olan etüv içerisinde düz bir şekilde 12 saat inkübe edilmiştir.

3.9. Amplikonların Vektöre Aktarılması ve Sekans Analizleri

Klonlama işlemi Fermentas firmasının bir ürünü olan InsTAclone™ Cloning Kit kullanılarak Ligasyon, Transformasyon ve Rekombinant klonların teyit edilmesi olarak 3 farklı analiz ile gerçekleştirilmiştir. Bu kit 0.055 µg/µL konsantrasyonunda 2886 bp uzunluğunda pTZ57R/T vektörü içermektedir.

3.10. Ligasyon

Ligasyon reaksiyonları 0.2 mL'lik ince kenarlı PZR tüplerinde gerçekleştirilmiş, vektör ve 5X ligasyon çözeltisi karışım halinde hazırlanarak PZR ürünü üzerinde ilave edildikten sonra *T4* DNA Ligaz enzim eklenerek tamamlanmış sonrasında hafifçe vorteks edilmiş ve santrifüj işleminden sonra 4 saat 22°C'de inkübe edilmiştir.

3.11. Rekombinant Klonların Analizi

Transformantları taşıyan rekombinant kolonilerin seçiminde LB Amfisilin Katı (pH 7.5) ortamına ilave edilen IPTG ve Xgal kimyasallarından yararlanılmıştır. Ortama ilave edilen IPTG'nin Beta-galaktosidaz proteini üretimini teşvik ettiği Xgal'ın ise ortamda Beta-galaktosidaz proteinin bulunması halinde Beta-galaktosidaz ile kimyasal tepkimeye girerek mavi renk oluşturduğu gözlenmiştir. Mavi ve beyaz renkli koloniler rekombinant bakteri hücrelerinin seçiminde kullanılmıştır. Ekme işleminin ardından plakalar üzerinde koloni gelişimi gözlenmiştir. Koloni büyüklüğü arttıkça mavi ve beyaz renk ayırt edilmiştir.

On beş veya 18 saatlik sürenin sonunda beyaz renkteki tek koloniden (rekombinant vektör taşıyan koloniler) lup ile bir kez dokunarak alınmış ve steril ortamda 100 ml'lik Erlenmayer içerisinde 10 mL LB Sıvı (pH 7.5) ortamı konulup içerisine 5 µL Amfisilin (100 µg/µL) ilave edilerek hazırlanmış olan ortam içerisine lup daldırılarak Erlenmayerin ağzı parafilm ile sıkı almayacak şekilde sarılmış ve bir bakteri hücresinin çoğaltımı yapılmıştır. Çoğaltım için 37 °C'ye ayarlanmış olan orbital çalkalayıcı üzerinde 12 saat 250 RPM hızda çalkalanmıştır. On iki saatlik çoğaltma süresinin sonunda bakterilerden plazmit DNA ekstraksiyonu geçilmiştir.

3.12. Plazmit DNA Ekstraksiyon Protokolü

1. 50 mL Erlenmayer içerisinde 10 mL'lik LB Amfisilin (pH 7.5) sıvı ortamında 37°C'de 250 RPM hızda 12 saatlik kültür sonrası çoğaltılmış olan bakteri kolonisi steril kabin içerisinde 50 mL'lik steril test tüpleri içerisine transfer edilmiştir.
2. Test tüpleri 6000 xg hızda 1 dakika santrifüjlenmiş ve süpernatant dökülerek uzaklaştırıldıktan sonra oluşan pellet 1 mL LB Sıvı (pH 7.0) ortamında çözülmüştür.
3. Çözülme işlemi tamamlandıktan sonra 2 mL'lik Ependorf tüplerine transfer edilmiş ve 12000 xg de 1 dakika santrifüj işleminden sonra üst faz uzaklaştırılmıştır.
4. Pellet üzerine RNaz içeren süspansiyon çözeltisinden 250 µL ilave edilerek karıştırılarak süspanse edilmiş ve sonra üzerine 250 µL lezyon çözeltisinden eklenerek 6 kez ters düz edilerek 5 dakika oda sıcaklığında bekletilmiştir.
5. Sürenin sonunda üzerine 350 µL bağlanma çözeltisi ilave edilmiş ve 6 kez ters düz edilerek 5 dakika oda sıcaklığında inkübe edilmiştir. Ardından 13000 xg de 10 dakika santrifüj edilmiş ve üst faz yeni bir tüpe (kolleksiyon) alınmıştır.
6. Kolleksiyon tüpü 13000 xg de 1 dakika santrifüj işlemi sonrası, tüp kartuşdan (silika) kaldırılarak alt hazne dökülerek atık uzaklaştırılmıştır. Üzerine 500 µL Yıkama çözeltisi I ilave edilmiş ve Kolleksiyon tüpü 13000 xg de 1 dakika santrifüj edildikten sonra tüp kartuşu kaldırılmış alt hazne dökülerek atık uzaklaştırılmıştır. Kolleksiyon tüp kartuşu üzerine 700 µL Yıkama çözeltisi II ilave edilmiş ve Kolleksiyon tüpü 13000 xg de 1 dakika santrifüj işleminden sonra tüp kartuşu kaldırılmış alt hazne dökülerek atık uzaklaştırılmıştır. Kartuş boş şekilde Kolleksiyon tüpü son kez 13000 xg de 1 dakika santrifüj işlemi sonrasında tüp kartuşu kaldırılmış alt hazne dökülerek atık kalıntılarda uzaklaştırılmıştır.

7. Kartuş yeni steril tüpün içerisine konularak 100 µL elüsyon çözeltisi tam silika tabanın üzerini kaplayacak şekilde ilave edilmiş ve 5 dakika oda sıcaklığında bekletildikten sonra 13000 xg de 1 dakika santrifüj işlemi ile kartuştaki solüsyon yeni bir tüpe alınmıştır. Pürifikasyonu (arıtma) işlemleri tamamlanan örnekler etiketlenerek -20°C’de saklanmıştır.

3.13. Amplikonların Sekans Analizleri

Çalışmada DNA Dizi Analizi Macrogene (Güney Kore) Firmasına yaptırılmıştır. Sekanslama işlemi Sanger metodu temel alınarak gerçekleştirilmiştir. Sekanslama işleminde universal M13F-pUC primer seti (İleri sekans: 5’-GTTTCCAGTCACGAC Geri sekans: 5’-CAGGAAACAGCTATGAC) kullanılmıştır. Her bir pozitif vektör en az iki kez DNA dizi analizlerine tabi tutulmuş ve sonuçlar Sequencher 4.8 biyoinformatik programı ile analiz edilmiştir.

3.14. Verilerin Analizi

3.14.1. Genlere ait sekans dizilerinin gen bankasından indirilmesi

Pamuk türlerine ait ITS gen dizileri NCBI veri tabanından BLAST programı yardımı ile belirlenerek alınmıştır. DNA sekans verileri NCBI veri tabanındaki "Batch Entrez" bağlantısı ile sekanslar FASTA formatında indirilerek hard disk ortamına yedekli olarak kayıt edilmiştir.

İkinci bitki grubunu oluşturan altmış türe ait sekansları NCBI’nın Taxonomy bağlantısı yardımı ile tek çenekliler (tek çim yapraklılar: monokotiledonlar), çift çenetliler (çift çim yapraklılar: dikotiledonlar), ağaçsı ve otsu türleri temsil eden taksonlar seçilmiş ve sekanslar FASTA formatında indirilerek hard disk ortamına kayıt edilmiştir.

3.14.2. Elde edilen sekansların alignment (Hizalama) yapılması

Çoklu Hizalama (Multiple Alignment) çalışmaları toplam dört farklı çoklu hizalama programı kullanılarak yapılmıştır. Çoklu hizalama programlarından en yaygın olarak

kullanılan ClustalW, MUSCLE (Multiple Sequence Comparison by Log- Expectation), T-COFFE, MAFFT (Multiple Alignment with Fast Fourier Transform) parogramları ile hizalama yapılmıştır ve her bir program için iki farklı parametre kullanılmıştır.

3.14.3. Filogenetik analiz için kullanılan metodlar

Çalışmada 4 farklı filogenetik yöntem ile analiz yapılmıştır ve bu yöntemlerde değişik sayıda nükleotit modelleri kullanılmıştır. MetaPIGA 2 (Helaers ve Milinkovitch 2010), MR. BAYES v.3.1.2 (Ronquist ve Huelsenbeck 2003), Phylogenetic Analysis Using Parsimony (PAUP v.4.0b10 (Swofford 2002) ve UGENE 1.9.1 (Fursov vd 2011) programları kullanılmış olup Maksimum Likelihood (ML), Bayes (BI), Maksimum Parsimoni (MP) ve Distance (D) metotları kullanılmıştır. Bu programlardan 43 adet ML, 43 adet BI, 17 adet MP ve 33 adet Distance filogenetik ağaç FigTreev1.3.1. (Rambaut, 2009) ve Tree View (Page 2000) programları ile görselleştirilmiştir.

4. BULGULAR ve TARTIŞMA

4.1. Çalışmada Kullanılan Bitki Taksonları, DNA Dizileri ve Elde Ediliş Şekilleri

Çalışmanın amacına ulaşmak için iki farklı veri grubu kullanılmıştır. Bu gruplarından ilki *Gossypium* L. cinsine giren 4 altcinsin temsil edildiği toplam 44 türden ve bir dış grup taksonundan oluşturulmuştur. Dış grup olarak kullanılan *Origanum vulgare* subsp. *hirtum* laboratuvar şartlarında genomik DNA'sı elde edilmiş ve ITS1-5.8-ITS2 DNA dizisi klonlandıktan sonra dizi analizi ticari olarak sentezletilmiştir.

Gossypium L. alt cinsleri, seksiyonları ve türleri Çizelge 4.1'de gösterilmiştir. Çalışmaya dahil edilen türler GenBank veri tabanında bulunan daha önceki çalışmalarda kullanılan türlerdir. Bilgimiz dahilinde pamukta daha önce 44 türün bir arada kullanıldığı ve ITS1-ITS2 dizlerinin kullanıldığı bir filogeni çalışması bulunmamaktadır.

Kırkdört pamuk türüne ait ITS1-5.8-ITS2 DNA dizileri (Çizelge 4.2) Sequencher 4.8 programı ve referans ITS1 ve ITS2 dizileri yardımı ile analiz edilerek ITS1 ve ITS2 dizilerine dönüştürülmüş, ayrı ayrı analiz edilmiş veya ITS1-ITS2 DNA dizileri birleştirilerek analiz edilmiştir. Çalışmada kullanılan diğer bitki grubu 60 taksondan oluşmuş ve her bir takson için bitki moleküler filogenisinde yaygın olarak kullanılan 5 farklı kloroplast gen bölgesi kullanılmıştır. Çizelge 4.3 ve Çizelge 4.4'de gösterildiği üzere çalışmada kullanılan lokuslar *matK*, *trnH-psbA*, *accD*, *rpoB*, *rpoC*'den oluşturulmuştur. Bu genlerden bitki moleküler filogenik çalışmalarda en etkin olanın belirlenmesi hedeflenmiştir.

Elde edilen DNA dizileri Sequencher 4.8 yazılım programı kullanılarak nükleotid dizi okumalarının doğrulukları kontrol edilmiş ve diziler arasında yapılan konsensus analizi sonucunda *Origanum vulgare* subsp. *hirtum* ait 640 bp uzunluğunda ITS1-5.8-ITS2 DNA dizisi elde edilmiştir. Çalışmada kullanılan bütün bitkilere ait DNA dizileri FASTA formatına dönüştürüldükten sonra ClustalX2 veya ClustalW yazılım programı ile CLUSTAL formatında (ALG) kayıt altına alınmıştır. Elde edilen veriler daha sonra çoklu hizalama programlarında kullanılmış ve filogeni çalışmalarında kullanılmak üzere NEXSUS dosyalarına dönüştürülmüştür.

Çizelge 4.1. Çalışmada kullanılan pamuk türleri ve ait oldukları biyocoğrafyalar

<i>Gossypium</i> L.		Biyocoğrafya
1-Altıncı Sturtia (R. Brown) Todaro	Seksiyon I: <i>Sturtia</i> (C genomu)	
	<i>G. sturtianum</i> J. H. Willis (C1 genomu)	Avustralya
	<i>G. robinsonii</i> F. Mueller (C2 genomu)	Avustralya
	Seksiyon II: <i>Grandicalyx</i> Fryxell (K genomu)	
	<i>G. costulatum</i> Todaro (K genomu)	Avustralya
	<i>G. cumminghamii</i> Todaro (K genomu)	Avustralya
	<i>G. exiguum</i> Fryxell, Craven & Stewart K	Avustralya
	<i>G. pilosum</i> Fryxell	Avustralya
	<i>G. rotundifolium</i> Fryxell, Craven & Stewart	Avustralya
	<i>G. enthyale</i> Fryxell, Craven & Stewart (K genomu)	Avustralya
	<i>G. nobile</i> Fryxell, Craven & Stewart	Avustralya
	<i>G. pulchellum</i> (C. A. Gardner) Fryxell	Avustralya
	<i>G. londonerriense</i> Fryxell, Craven & Stewart	Avustralya
	<i>G. marchantii</i> Fryxell, Craven & Stewart	Avustralya
	<i>G. populifolium</i> (Bentham) F. Mueller ex Todaro	Avustralya
	<i>G. anapoides</i>	Avustralya
	Seksiyon III <i>Hibiscoidea</i> Todaro	
	<i>G. australe</i> F. Mueller (G genomu)	Avustralya
	<i>G. nelsonii</i> Fryxell	Avustralya
	<i>G. bickii</i> Prokhanov (G1 genomu)	Avustralya
2-Altıncı Houzingenia (Fryxell)	Seksiyon I: <i>Houzingenia</i>	
	Altseksiyon I: <i>Houzingenia</i>	
	<i>G. thurberi</i> Todaro (D1 genomu)	Meksika, Arizona
	<i>G. trilobum</i> (DC.) Skovsted (D8 genomu)	Meksika
	Altseksiyon II: <i>Integrifolia</i> (Todaro) Todaro	
	<i>G. davidsonii</i> Kellogg (D3-d genomu)	Meksika
	<i>G. klotzschianum</i> Andersson (D3-k genomu)	Galapagos Adaları
	Altseksiyon III: <i>Caducibracteolata</i> Mauer	
	<i>G. armourianum</i> Kearney (D2-1 genomu)	Meksika
	<i>G. harknessii</i> Brandegees (D2-2 genomu)	Meksika
	<i>G. turneri</i> Fryxell (D10 genomu)	Meksika
	Seksiyon II: <i>Erioxylum</i> (Rose & Standley) Prokhanov	
	Altseksiyon I: <i>Erioxylum</i> (D genomu)	
	<i>G. aridum</i> (Rose & Standley ex Rose) Skovsted (D4 genomu)	Meksika
	<i>G. lobatum</i> H. Gentry (D7 genomu)	Meksika
	<i>G. laxum</i> Phillips (D9 genomu)	Meksika
	<i>G. schwendimannii</i> Fryxell & S. Koch (D11 genomu)	Meksika
	Altseksiyon II: <i>Selera</i> (Ulbrich) Fryxell (D genomu)	
	<i>G. gossypoides</i> (Ulbrich) Standley (D6 genomu)	Meksika
	Altseksiyon III: <i>Austroamericana</i> Fryxell	
3-Altıncı Gossypium	<i>G. raimondii</i> Ulbrich (D5 genomu)	Peru
	Seksiyon I: <i>Gossypium</i> (A, B, E, F genomları)	
	<i>G. herbaceum</i> L. (A1 genomu)	Asya-Afrika
	<i>G. arboreum</i> L. (A2 genomu)	Asya-Afrika
	Altseksiyon I: <i>Anomala</i> Todaro (B genomu)	
	<i>G. anomalum</i> Wawra & Peyritsch (B1 genomu)	Afrika
	<i>G. triphyllum</i> (Harvey & Sonder) Hochreutiner (B2 genomu)	Afrika
	<i>G. capitata-viridis</i> Mauer (B3 genomu)	Cape Verde
	Altseksiyon II: <i>Longiloba</i> Fryxell	
	<i>G. longicalyx</i> J. B. Hutchinson & Lee (F1 genomu)	Afrika
	Altseksiyon III: <i>Pseudopambak</i> (Prokhanov) Fryxell (E genomu)	
	<i>G. stocksii</i> Masters in Hooker (E1 genomu)	Arabistan
	<i>G. somalense</i> (Gurke) J. B. Hutchinson (E2 genomu)	Afrika
	<i>G. areysianum</i> Deflers (E3 genomu)	Arabistan
4-Altıncı Karpas	<i>G. incanum</i> (Schwartz) Hillcoat (E4 genomu)	Arabistan
	<i>G. hirsutum</i> L. (AD1 genomu)	Amerika
	<i>G. barbadense</i> L. (AD2 genomu)	Amerika
	<i>G. tomentosum</i> Nuttall ex Seemann (AD3 genomu)	Havai
	<i>G. mustelinum</i> Miers ex Watt (AD4 genomu)	Brezilya
	<i>G. darwinii</i> Watt (AD5 genomu)	Galapagos Adaları

Çizelge 4.2. Çalışmada kullanılan pamuk türleri ve GenBank GI numaralar

	Takson	ITS GI
1.	<i>G. anapoides</i>	3093922
2.	<i>Gossypium anomalum</i> Wavr & Pevr.	238477918
3.	<i>Gossypium arboreum</i> L.	295292537
4.	<i>Gossypium areysianum</i> (Defl.) Hutch.	1654278
5.	<i>Gossypium aridum</i> (Rose & Standl) Skov.	530851
6.	<i>Gossypium armourianum</i> Kearns	530844
7.	<i>Gossypium australe</i> F. Muell	3093919
8.	<i>Gossypium barbadense</i> L.	295292549
9.	<i>Gossypium bickii</i> Prokh.	3093913
10.	<i>Gossypium capitis-viridis</i> Maner.	238477917
11.	<i>Gossypium costulatum</i> Tod.	1654281
12.	<i>Gossypium cunninghamii</i> Tod.	3093927
13.	<i>Gossypium darwinii</i> Watt	530835
14.	<i>Gossypium davidsonii</i> Kell.	530848
15.	<i>Gossypium enthyle</i> Fryx.	3093929
16.	<i>Gossypium exiguum</i> Fryx.	1654285
17.	<i>Gossypium gossypoides</i> (Ulbr.) Standl	530843
18.	<i>Gossypium harknessii</i> (Brandg)	530846
19.	<i>Gossypium herbaceum</i> L.	295292565
20.	<i>Gossypium hirsutum</i> L.	238477922
21.	<i>Gossypium incanum</i> (SchWartz) Hille	16542830
22.	<i>Gossypium klotzschianum</i> Anderss	530847
23.	<i>Gossypium laxum</i> Phillips	530849
24.	<i>Gossypium lobatum</i> Gentry	530850
25.	<i>Gossypium londonderriense</i> Fryx.	3093936
26.	<i>Gossypium longicalyx</i> Hutch. & Lee	238477920
27.	<i>Gossypium marchantii</i> Fryx.	3093941
28.	<i>Gossypium mustelinum</i> Watt	530833
29.	<i>Gossypium nelsonii</i> Fryx.	3093916
30.	<i>Gossypium nobile</i> Fryx.	3093942
31.	<i>Gossypium pilosum</i> Fryx.	3093947
32.	<i>Gossypium populifolium</i> (Benth.) Tod.	3093952
33.	<i>Gossypium pulchellum</i> (C. A Gard.) Fryx	3093957
34.	<i>Gossypium robinsonii</i> F. Muell.	3093911
35.	<i>Gossypium rotundifolium</i> Fryx.	3093961
36.	<i>Gossypium schwendimanii</i> Fryx.	530853
37.	<i>Gossypium somalense</i> (Gurke) Hutch.	238477916
38.	<i>Gossypium stocksii</i> Mast. Hook	238477921
39.	<i>Gossypium sturtianum</i> J. H. Willis	3093909
40.	<i>Gossypium thurberi</i> Tod.	530830
41.	<i>Gossypium tomentosum</i> Seem.	530836
42.	<i>Gossypium trilobum</i> (D C.) Skov.	530842
43.	<i>Gossypium triphyllum</i> (Harv & Sond.) Hochr.	530840
44.	<i>Gossypium turneri</i> Fryx.	530845

Çizelge 4.3. Çalışmada kullanılan atmış taksona ait *matK*, *rpoB* ve *rpoCl* genlerinin GI numaraları

Familya	Tür adı	<i>matK</i>	<i>rpoB</i>	<i>rpoCl</i>
Anacardiaceae	<i>Rhus gueinzii</i> Sond.	167890190	167891222	167890715
Anacardiaceae	<i>Rhus gueinzii</i> Sond.	167890192	167891224	167890717
Anacardiaceae	<i>Rhus leptodictya</i> Diels	167890196	167891228	167890721
Anacardiaceae	<i>Rhus leptodictya</i> Diels	167890198	167891230	167890723
Anacardiaceae	<i>Rhus transvaalensis</i> Engl.	167890202	167891234	167890727
Anacardiaceae	<i>Rhus transvaalensis</i> Engl.	167890204	167891238	167890731
Apiaceae	<i>Heteromorpha arborescens</i> Cham. & Schltdl.	167890164	167891196	167890689
Apiaceae	<i>Heteromorpha arborescens</i> Cham. & Schltdl.	167890166	167891198	167890691
Apiaceae	<i>Steganotaenia araliacea</i> Hochst.	167890212	167891246	167890739
Apiaceae	<i>Steganotaenia araliacea</i> Hochst.	167890214	167891248	167890741
Arecaceae	<i>Hyphaene coriacea</i> Gaertn.	167890172	167891204	167890697
Arecaceae	<i>Hyphaene coriacea</i> Gaertn.	167890174	167891206	167890699
Arecaceae	<i>Hyphaene petersiana</i> Klotzsch ex Mart	167890180	167891212	167890705
Arecaceae	<i>Hyphaene petersiana</i> Klotzsch ex Mart	167890182	167891214	167890707
Asteraceae	<i>Brachylaena huillensis</i> O. Hoffm.	167890074	167891104	167890597
Asteraceae	<i>Brachylaena huillensis</i> O. Hoffm.	167890076	167891106	167890599
Combretaceae	<i>Combretum collinum</i> Fresen.	167890086	167891116	167890609
Combretaceae	<i>Combretum collinum</i> Fresen.	167890088	167891118	167890611
Combretaceae	<i>Combretum apiculatum</i> Sond.	167890080	167891110	167890603
Combretaceae	<i>Combretum apiculatum</i> Sond.	167890082	167891112	167890605
Euphorbiaceae	<i>Croton megalobotrys</i> Müll. Arg.	167890102	167891134	167890627
Euphorbiaceae	<i>Croton megalobotrys</i> Müll. Arg.	167890104	167891136	167890629
Euphorbiaceae	<i>Croton gratissimus</i> Burch	167890096	167891128	167890621
Euphorbiaceae	<i>Croton gratissimus</i> Burch	167890098	167891130	167890623
Fabaceae	<i>Acacia tortilis</i> Hayne	167890060	167891090	167890583
Fabaceae	<i>Acacia tortilis</i> Hayne	167890062	167891092	167890585
Loganiaceae	<i>Strychnos decussata</i> (Pappe) Gilg	167890220	167891254	167890747
Loganiaceae	<i>Strychnos decussata</i> (Pappe) Gilg	167890221	167891256	167890749
Loganiaceae	<i>Strychnos madagascariensis</i> Spreng. ex Baker	167890223	167891260	167890753
Loganiaceae	<i>Strychnos madagascariensis</i> Spreng. ex Baker	167890225	167891262	167890755
Malvaceae	<i>Grewia flavescens</i> Juss.	167890152	167891184	167890677
Malvaceae	<i>Grewia flavescens</i> Juss.	167890154	167891186	167890679
Malvaceae	<i>Grewia bicolor</i> Juss.	167890146	167891178	167890671
Malvaceae	<i>Grewia bicolor</i> Juss.	167890148	167891180	167890673
Moraceae	<i>Ficus abutilifolia</i> Miq	167890128	167891160	167890653
Moraceae	<i>Ficus abutilifolia</i> Miq	167890130	167891162	167890655
Moraceae	<i>Ficus glumosa</i> Delile	167890134	167891166	167890659

Çizelge 4.3. Devamı.

Familya	Tür adı	matK	rpoB	rpoCl
Moraceae	<i>Ficus glumosa</i> Delile	167890136	167891168	167890661
Myrothamnaceae	<i>Myrothamnus flabellifolia</i> Welw.	167890184	167891216	167890709
Myrothamnaceae	<i>Myrothamnus flabellifolia</i> Welw.	167890186	167891218	167890711
Orchidaceae	<i>Clowesia warczewiczii</i> (Lindl. & Paxton) Dodson	167889940	167890968	167890461
Orchidaceae	<i>Clowesia warczewiczii</i> (Lindl. & Paxton) Dodson	167889942	167890970	167890463
Orchidaceae	<i>Encyclia alata</i> (Bateman) Schltr.	167889960	167890988	167890481
Orchidaceae	<i>Encyclia alata</i> (Bateman) Schltr.	167889962	167890990	167890483
Orchidaceae	<i>Ponthieva racemosa</i> (Walter) C. Mohr	167890008	167891038	167890529
Orchidaceae	<i>Ponthieva racemosa</i> (Walter) C. Mohr	1678900010	167891040	167890531
Orchidaceae	<i>Prosthechea fragrans</i> (Sw.) W. E. Higgins	167890016	167891046	167890539
Orchidaceae	<i>Prosthechea fragrans</i> (Sw.) W. E. Higgins	167890018	167891048	167890541
Orchidaceae	<i>Stanhopea saccata</i> Bateman	167890032	167891062	167890555
Orchidaceae	<i>Stanhopea saccata</i> Bateman	167890034	167891064	167890557
Orchidaceae	<i>Trichopilia tortilis</i> Lindl.	167890036	167891066	167890559
Orchidaceae	<i>Trichopilia tortilis</i> Lindl.	167890038	167891068	167890561
Proteaceae	<i>Faurea rochetiana</i> Chiov. ex Pic. Serm.	167890116	167891148	167890641
Proteaceae	<i>Faurea rochetiana</i> Chiov. ex Pic. Serm.	167890118	167891150	167890643
Proteaceae	<i>Faurea saligna</i> Harv.	167890122	167891154	167890647
Proteaceae	<i>Faurea saligna</i> Harv.	167890124	167891156	167890649
Solanaceae	<i>Solanum panduriforme</i> Drège ex Dunal	167890206	167891240	167890733
Solanaceae	<i>Solanum panduriforme</i> Drège ex Dunal	167890208	167891242	167890735
Velloziaceae	<i>Xerophyta retinervis</i> Baker	167890236	167891274	167890767
Velloziaceae	<i>Xerophyta retinervis</i> Baker	167890238	167891276	167890769

Çizelge 4.4. Çalışmada kullanılan atmış taksona ait *accD* ve *trnH-psbA* genlerinin GI numaraları

Familya	Tür adı	<i>accD</i>	<i>trnH-psbA</i>
Anacardiaceae	<i>Rhus gueinzii</i> Sond.	167891761	167890911
Anacardiaceae	<i>Rhus gueinzii</i> Sond.	167891763	167890912
Anacardiaceae	<i>Rhus leptodictya</i> Diels	167891767	167890914
Anacardiaceae	<i>Rhus leptodictya</i> Diels	167891769	167890915
Anacardiaceae	<i>Rhus transvaalensis</i> Engl.	167891773	167890917
Anacardiaceae	<i>Rhus transvaalensis</i> Engl.	167891777	167890919
Apiaceae	<i>Heteromorpha arborescens</i> Cham. & Schltdl.	167891735	167890905
Apiaceae	<i>Heteromorpha arborescens</i> Cham. & Schltdl.	167891737	167890906
Apiaceae	<i>Steganotaenia araliacea</i> Hochst.	167891785	167890923
Apiaceae	<i>Steganotaenia araliacea</i> Hochst.	167891787	167890924
Arecaceae	<i>Hyphaene coriacea</i> Gaertn.	167891743	167890844
Arecaceae	<i>Hyphaene coriacea</i> Gaertn.	167891745	167890845
Arecaceae	<i>Hyphaene petersiana</i> Klotzsch ex Mart	167891751	167890848
Arecaceae	<i>Hyphaene petersiana</i> Klotzsch ex Mart	167891753	167890849
Asteraceae	<i>Brachylaena huillensis</i> O. Hoffm.	167891643	167890860
Asteraceae	<i>Brachylaena huillensis</i> O. Hoffm.	167891645	167890861
Combretaceae	<i>Combretum collinum</i> Fresen.	167891655	167890866
Combretaceae	<i>Combretum collinum</i> Fresen.	167891657	167890867
Combretaceae	<i>Combretum apiculatum</i> Sond.	167891649	167890863
Combretaceae	<i>Combretum apiculatum</i> Sond.	167891651	167890864
Euphorbiaceae	<i>Croton megalobotrys</i> Müll. Arg.	167891673	167890875
Euphorbiaceae	<i>Croton megalobotrys</i> Müll. Arg.	167891675	167890876
Euphorbiaceae	<i>Croton gratissimus</i> Burch	167891667	167890872
Euphorbiaceae	<i>Croton gratissimus</i> Burch	167891669	167890873
Fabaceae	<i>Acacia tortilis</i> Hayne	167891629	167890855
Fabaceae	<i>Acacia tortilis</i> Hayne	167891631	167890856
Loganiaceae	<i>Strychnos decussata</i> (Pappe) Gilg	167891793	167890926
Loganiaceae	<i>Strychnos decussata</i> (Pappe) Gilg	167891795	167890927
Loganiaceae	<i>Strychnos madagascariensis</i> Spreng. ex Baker	167891799	167890929
Loganiaceae	<i>Strychnos madagascariensis</i> Spreng. ex Baker	167891801	167890930
Malvaceae	<i>Grewia flavescens</i> Juss.	167891723	167890900
Malvaceae	<i>Grewia flavescens</i> Juss.	167891725	167890901
Malvaceae	<i>Grewia bicolor</i> Juss.	167891717	167890897
Malvaceae	<i>Grewia bicolor</i> Juss.	167891719	167890898
Moraceae	<i>Ficus abutilifolia</i> Miq	167891699	167890888
Moraceae	<i>Ficus abutilifolia</i> Miq	167891701	167890889
Moraceae	<i>Ficus glumosa</i> Delile	167891705	167890891

Çizelge 4.4. Devamı.

Familya	Tür adı	accD	trnH-psbA
Moraceae	<i>Ficus glumosa</i> Delile	167891707	167890892
Myrothamnaceae	<i>Myrothamnus flabellifolia</i> Welw.	167891755	167890908
Myrothamnaceae	<i>Myrothamnus flabellifolia</i> Welw.	167891757	167890909
Orchidaceae	<i>Clowesia warczewiczii</i> (Lindl. & Paxton) Dodson	167891507	167890789
Orchidaceae	<i>Clowesia warczewiczii</i> (Lindl. & Paxton) Dodson	167891509	167890790
Orchidaceae	<i>Encyclia alata</i> (Bateman) Schltr.	167891527	167890799
Orchidaceae	<i>Encyclia alata</i> (Bateman) Schltr.	167891529	167890800
Orchidaceae	<i>Ponthieva racemosa</i> (Walter) C. Mohr	167891577	167890823
Orchidaceae	<i>Ponthieva racemosa</i> (Walter) C. Mohr	167891579	167890824
Orchidaceae	<i>Prosthechea fragrans</i> (Sw.) W. E. Higgins	167891585	167890828
Orchidaceae	<i>Prosthechea fragrans</i> (Sw.) W. E. Higgins	167891587	167890829
Orchidaceae	<i>Stanhopea saccata</i> Bateman	167891601	167890836
Orchidaceae	<i>Stanhopea saccata</i> Bateman	167891603	167890837
Orchidaceae	<i>Trichopilia tortilis</i> Lindl.	167891605	167890838
Orchidaceae	<i>Trichopilia tortilis</i> Lindl.	167891607	167890839
Proteaceae	<i>Faurea rochetiana</i> Chiov. ex Pic. Serm.	167891687	167890882
Proteaceae	<i>Faurea rochetiana</i> Chiov. ex Pic. Serm.	167891689	167890883
Proteaceae	<i>Faurea saligna</i> Harv.	167891693	167890885
Proteaceae	<i>Faurea saligna</i> Harv.	167891695	167890886
Solanaceae	<i>Solanum panduriforme</i> Drège ex Dunal	167891779	167890920
Solanaceae	<i>Solanum panduriforme</i> Drège ex Dunal	167891781	167890921
Velloziaceae	<i>Xerophyta retinervis</i> Baker	167891813	167890936
Velloziaceae	<i>Xerophyta retinervis</i> Baker	167891815	167890937

4.2. Çalışmada Kullanılan Hizalama (Alignment) Yazılım Programları ve Seçilen Parametrelerin Belirlenmesi

Bu çalışmada alignment (hizalama) yazılım programlarının ve parametrelerinin gen, gen bölgeleri ve uzunlukları, filogenetik yöntemler, substitüsyon modelleri ve ağaç topolojisi üzerinde etkisinin olup olmadığını tespit etmek için moleküler filogenide kullanılan yaygın dört adet çoklu alignment yazılım programı kullanılmıştır. Çizelge 4.5’de gösterilmiş olan parametreler ClustalW, T-COFFEE, MUSCLE ve MAFFT çoklu hizalama programlarında uygulanmıştır (Çizelge 4.5). Hizalama çıktılarının gözle incelenmesi yapıldığında hizalamanın program ve parameterlerden etkilendiği tespit edilmiştir.

Çalışmada kullanılan hizalama programlarının farklı hizalama oluşturması kullandıkları algoritmaların farklı olması yanında uygulanan parametrelerin de etkin olmasından kaynaklanmıştır. Kullanılan paramaterlerden cezalandırma boşlukları (GOP), ekleme cezalandırma (GEP), hizalama tekrar sayıları (ITER), tekrarlama tipi ve ağırlık matrislerini etkili olduğu tespit edilmiştir. Çoklu sekans hizalamada değişik parametrelerin etkisinin olup olmadığını araştırmak için ilk kez Fitch ve Smith (1983) tarafından araştırılmış ve bu parametrelerin önemi vurgulanmıştır. Çoklu sekans hizalamanın temeli, indelleri yani sekansta bulunan insersiyon ve delesyonları başka bir deyişle boşlukları (gaps) modellemektir. En yaygın kullanılan model ise GOP ve GEP cezalandırma boşluklarıdır. GOP, indeler için oluşan boşlukları sınırlandırmak için kullanılır. Yüksek GOP değeri boşluk sayıları sınırlandırır. GEP ise, olan boşlukları uzatır. GEP değeri 0-10 arası değişir. Yüksek GEP değeri boşlukları kısaltır.

Çizelge 4.5. Çalışmada kullanılan alignment programları ve parametreleri

Program	PARAMETRE 1	PARAMETRE 2
CLUSTALW	GOP=15 GEP=6.66 WM=CLUSTALW IT=ALIGN ITER=100	GOP=15 GEP=6.66 WM=CLUSW IT=TREE ITER=100
MAFFT	GOP=1.5 GEP=0 ITER=100	GOP=3 GEP=1 ITER=100
MUSCLE	MUSCLE ITER=100, SEKANSLARI TEKRAR DÜZELTMESİN	MUSCLE ITER=100, SEKANSLARI TEKRAR DÜZELTSİN
T-COFFEE	GOP= -100 GEP=10 ITER=100	GOP= -200 GEP=20 ITER=100

GOP: Gap opening penalty, GEP: Gap extensiyon penalty, WM = Weight matrix, ITER= Iteration, IT= Iteration type

Çizelge 4.5’de verilen alignment programları ve parametreleri ITS1 ve ITS2 dizileri kullanılarak hizalanmış ve toplam 33 ML ağacı oluşturulmuştur (Bknz. Ekler). Kırkdört pamuk türü ve bir dış grup taksonun kullanılması ile elde edilen filogenilerin karşılaştırılması sonucu MAFFT hizalama programının ClustalW, T-COFFEE, MUSCLE programlarından daha etkin hizalama oluşturduğu tespit edilmiştir.

ClustalW parogramı verilen sekansı hızlı bir şekilde hizalarken, önce eldeki dizilerin ikili benzerlikleri kullanılarak tahmini bir filogenetik ağaç (rehber) oluşturmuş ve daha sonra bu bilgiden yararlanılarak, aşamalı olarak diziler hizalanmıştır. Bu yöntemin bu çalışmada seçilmemiş olması büyük bir ihtimalle yerel minimuma takılmış olması ve problemi doğru şekilde modelleyecek parametrelerin belirlenememiş olmasıdır (Thompson vd 2009).

T-COFFE programı verilen sekansı hizalarken birçok kritere göre ClustalW’den çok daha doğrulayıcı olmasına rağmen hızı ClustalW daha düşük bulunmuştur. T-COFFEE ile ClustalW arasında T-COFFEE’nin farklı birçok olanakları bulunmaktadır. Örneğin eğer çok ilişkili sekansları dizmek istediğinizde T-COFFEE ClustalW’den çok daha hızlı bir seçeneğe sahiptir. Bu seçenek iki sekans arasında en iyi çaprazlanmayı düşünerek çalışmaktadır (Notredame 2008). Ancak çalışmamızda T-COFFEE programının oluşturmuş olduğu hizlama beklenen filogenik ilişkileri tam anlamı ile ortaya koyamamıştır. T-COFFEE hizalama için çok zaman harcamış ve karmaşık boşluklar oluşturmuştur (Edgar 2004). Çalışmada kullanılmış olan diğer bir hizalama programı olan MUSCLE, hizalama için T-COFFEE den daha hızla olmasına rağmen ondan aynı zamanda daha az doğrulayıcı sonuçlar vermiştir.

MAFFT programı hızlı bir şekilde çoklu hizalama yapmak için geliştirilmiştir. ‘Fast Fourier Transform (FFT)’ algoritması ve uzaklık hesaplama algoritması hızlı bir şekilde hizalamayı kolaylaştırmaktadır. MAFFT aynı zamanda eklemeli (iterative) bir şekilde çoklu hizalama yaparak her yeni eklenen sekansla birlikte önceden hizalandıkları sekansları tekrar hizalayabilmektedir bu özelliği ile T-COFFEE’den daha doğrulayıcı ama yavaş sonuç vermiştir (Katoh vd 2008). Verilerin bir arada incelenmesi sonucu çalışmada kullanılan veriler için MAFFT programı seçilmiştir.

4.3. Filogenetik Analiz İçin Substitüsyon Modellerinin Belirlenmesi

Bu çalışmada verilerimize uygun model oluşturmak için MetaPIGA (Helaers ve Milinkovitch, 2010) programının bünyesinde bulunan üç farklı model test (LRT, AIC, BIC) kullanılarak önce ITS dizilerinde daha sonrada diğer 5 farklı genin analizlerinde kullanılacak substitüsyon modelleri belirlenmiştir. Çizelge 4.6’da gösterildiği üzere değişik alignment programları ve parametreleri doğrultusunda üç ayrı model (GTR, TN93, K2P) belirlenmiştir. Filogenetikte mümkün olan en iyi evrimsel uzaklık tahminini yapmak, uygun filogenetik ağaç ve doğru dal uzunluklarını oluşturmak aynı zamanda yüksek bootstrap değerlerini elde etmek için eldeki veriye en uygun substitüsyon modeli seçilmelidir. Model testlerin seçtiği substitüsyon modellerine bakıldığında seçilen model, alignment programı çeşidi ve oluşturulan alignment parametrelerinden çok etkilendiği bulunmuştur. Aynı zamanda gen bölgesi, gen uzunluğununda model seçiminde etkili olduğu görülmüştür.

Çizelge 4.6. Gen bölgesine ve hizalama parametrelerine bağlı olarak test modellerin seçtiği, mevcut verilere uygun en iyi substitüsyon (SBT) modelleri

Gen	Hizalama Parametresi	LRT	AIC	BIC
ITS1	CLUSTALW 1. PRMT	TN93	K2P	K2P
	CLUSTALW 2. PRMT	TN93	K2P	K2P
	MAFFT 1. PRMT	TN93	K2P	K2P
	MAFFT 2. PRMT	TN93	TN93	K2P
	MUSCLE 1. PRMT	TN93	TN93	K2P
	MUSCLE 2. PRMT	TN93	K2P	K2P
	T-COFFEE 1.PRMT	TN93	TN93	K2P
	T-COFFEE 2.PRMT	TN93	TN93	K2P
ITS2	CLUSTALW 1. PRMT	GTR	GTR	K2P
	CLUSTALW 2. PRMT	GTR	GTR	K2P
	MAFFT 1. PRMT	GTR	GTR	K2P
	MAFFT 2. PRMT	GTR	GTR	K2P
	MUSCLE 1. PRMT	GTR	GTR	K2P
	MUSCLE 2. PRMT	GTR	GTR	K2P
	T-COFFEE 1. PRMT	GTR	GTR	K2P
	T-COFFEE 2. PRMT	GTR	GTR	K2P

LRT: Likelihood ratio test, AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion, GTR: General-Time Reversible model, TN93: Tamura-Nei 1993 model.

Subtitüsyon modelleri sekans evriminin, genetik uzaklık tahminini doğru bir şekilde elde edilmesini sağlamaktadır (Yang 1996). Subtitüsyon modeline ait transversiyon ve transisyon oranı filogenetik ağaç topolojisi üzerinde çok büyük etkisi olduğu bilinmektedir (Yang 1994). Filogenide kullanılan ve içiçe (“nested”) modellerin oluşturulduğu GTR ailesi toplam 64 modeli kapsamaktadır. En yaygın olarak kullanılanlar modeller Jukes-Cantor (JC) , K2P, Felsenstein 1984 (F84), Hasegawa-Kishhono-Yano (HKY85), TN93 modelleri olup bunların büyük bir çoğunluğu GTR modelinin farklı parametrelerinin bir kombinasyonu sonucu oluşturulmuşlardır (Hillis 1996).

HKY85 ve GTR modelleri farklı olarak dört nükleotidin farklı frekanslarda oluşmasına (değişiminin) izin verir. Buna göre F84 ve HKY85 modelleri birbirine çok yakın iki modeller olarak tanımlanabilmektedir. Model için farklı hesaplar kullanılsa da esasında aynı şeylerdir. Bu iki model, transisyon ve transversiyonun farklı hızlarda gerçekleştiğini varsaymaktadırlar (Hillis 1996).

GTR modeli ise bir simetrik subtitüsyon matriksini varsayar. Başka bir deyiş ile A’nın T’ye değişimi ile T’nin A’ya değişimini aynı hızda olduğunu kabul eder. Her bir nükleotid çiftinin subtitüsyonun farklı hızlarda meydana geldiğini varsayar. Nükleotidlerin ise farklı hızlarda meydana gelebileceğini ön görmektedir.

TN93 modeli dört farklı frekans parametresini kullanır. Bu model, transisyon ve transversiyonlar arasında farklılık olduğunu varsayar. Aynı zamanda iki çeşit transisyon arasında (Pürin-Pürin, Pirimidin-Primidin) farklılık olduğunu varsayar. Sonuç olarak bu modeller çok yakın ilişkilidir. İç içe modeller, (nested) genel modellerin özel bir durumunu temsil etmektedirler (Hillis 1996, Li 1997).

Bu modellerin çoğu benzer olsada kendi içerinse varsaydıkları algoritmalar farklı olabilmektedir. JC ve K2P modelleri bütün nükleotidlerin eşit frekansta meydana geldiğini varsaymaktadır. JC modeli ayrıca bütün subtitüsyonların eşit ve bütün nükleotidlerin aynı frekansta meydana geldiğini varsayar. Bu model subtitüsyon hızı (alpha) adında sadece bir parametre kullanmaktadır. K2P modeli ise transversiyon ve

transisyon hızını farklı olduğunu, bütün nükleotidlerin aynı frekansta meydana geldiğini varsaymaktadır (Li 1997).

4.4. Kullanılan Filogeni Yöntemleri ve Oluşturdukları Filogenetik Ağaçlar

Çalışmada en iyi alignment yöntemi, parametersi, substitüsyon modeli, filogeni yöntemini ve filogeni programını belirlemek için *Gossypium* L. cinsinin oluşturduğu taksonlar kullanılmıştır. *Gossypium* L. cinsinin taksonomisini seksiyon ve genom büyüklüğü bazında veren kriterler, parametreler ve metodlar en uygun yöntem ve metot olarak varsayılmıştır.

Wendel vd (2010) göre: *Gossypium* L. cinsi genel olarak dört altcins altında sınıflandırılır. Birinci altcins *Sturtia* (R. Brown) Todaro dir. Avustralya kıtasında yayılış gösteren bu altcinsin ait üç bölüm (seksiyon) bulunmaktadır.

Birinci seksiyon genomu C ile temsil eden *Sturtia*'dır. İkinci seksiyon *Grandicalyx* Fryxell olup genomu K ile temsil edilmektedir. Üçüncü seksiyon ise *Hibiscoidea* Todaro olup yine genomu C ile temsil edilen türler içermektedir (Wendel vd 2010).

İkinci altcins *Houzingenia* (Fryxell) olup Amerika kıtasında yayılış gösteren taksonlardan oluşmuştur. Bu altcins temel olarak iki seksiyondan oluşmaktadır. Birinci seksiyon *Houzingenia* olup üç alt seksiyondan oluşur ve genom büyüklüğü bakımından D genomunu temsil eder. İkinci seksiyon ise, *Erioxylum* (Rose ve Standley) Prokhanov olup üç altseksiyondan oluşmaktadır. Bu seksiyon da genom büyüklüğü bakımından yine D genomunu temsil eden taksonlardan oluşmaktadır.

Üçüncü altcins *Gossypium* olup Afrika/Arabistan bölgesinde yayılış gösteren A, B, E, ve F genom büyüklerini kapsayan taksonları içerir. Bu altcins aynı zamanda iki seksiyondan oluşmaktadır.

Birinci seksiyon *Gossypium*, ikinci seksiyon ise, *Serrata* Fryxell dir. Dördüncü ve son altcins *Karpas* Rafinesque olup tetraploid türlerini içeren taksonlardan oluşmaktadır.

4.4.1. Maksimum Olası (ML) metodu ve elde edilen ağaçlar

ML analizinde pamuk türlerine ait farklı alignment parametreleri ile hizalanmış ITS1, ITS2, ITS1-ITS2 gen bölgelerinin sekansları MetaPIGA (Helaers ve Milinkovitch 2010) programı kullanılarak toplam üç farklı SBT modelinde gerçekleştirerek toplam 33 ağaç elde edilmiştir (Bkz. Ekler). Ancak pamuk türlerini kıtalara (Afrika, Avustralya, Amerika) uygun bir topoloji ile veren MAFFT programının birinci parametresi ve GTR SBT modeli sonucu oluşan ağaçlar aşağıda verilmiş olan şekillerde verilmiştir. Bu durumda oluşturulan MAFFT programının birinci parametresi ve GTR substitüsyon modeli, mevcut pamuk türlerinin ağaç topolojisini en uygun bir şekilde ve yüksek bootstrap (ML) veya posterior probability (BI) değerleri ile ortaya çıkardığı görülmüştür. Bu analiz sonucunda öngörülen bazı bilgiler ışığında diğer analizlerin de MAFFT'ın birinci parametresi ve GTR substitüsyon modeli ile değerlendirilmesi gerektiği sonucuna varılmıştır.

Şekil 4.1'de MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi sonucu elde edilen ağaç topolojisi gösterilmiştir. TN93 modeli kullanıldığında Avustralya pamukları iki ana grup içerisinde toplanmıştır. Birinci grubun en dış türü *G. nobile*, ikinci grupta bulunan Avustralya kökenli pamuk türleri ile Afrika kökenli pamuk türlerinin ilişkili olduğu görülmektedir. Afrika pamuklarında Avustralya pamuklarında görülen iki grup altında toplanmış ancak Amerika orijinli olan *G. mustelinum* ve Afrika orijinli *G. herbaceum* ve *G. arboreum*'a yakın bir tür olarak belirlenmiştir. *G. mustelinum* allotetraploid bir tür olup genomunda *G. herbaceum* ve *G. arboreum* kalıntıları bulunmaktadır. Amerika pamuk türleri de Avustralya, Afrika türlerine benzer şekilde iki alt grup altında toplanmıştır. ITS1 dizisinin yalnız kullanıldığı analizlerde Amerikan pamuk türlerinin özellikle yeni gelişmiş olan allotetraploid türlerden dolayı diğer pamuk türlerinden belirgin olarak farklı olduğu tespit edilmiştir.

Çalışmanın Ekler kısmında verilen ağaç topolojilerinin birlikte incelenmesi ile ITS1 ve ITS2 dizilerinin tek başlarına kullanıldığında (Şekil 4.1 ve Şekil 4.2) oluşan topolojilerde sorunlar olduğu tespit edilmiştir. Örneğin Şekil 4.2'de görüldüğü üzere *G. gossypoides* dış gruptan daha dış bir tür olarak belirlenmiştir. Bu nedenle pamuk

filogeni çalışmalarında salt ITS1 veya ITS2 nin kullanımı yerine ITS1-ITS2 dizilerinin birlikte kullanılması gerektiği ortaya konulmuştur.

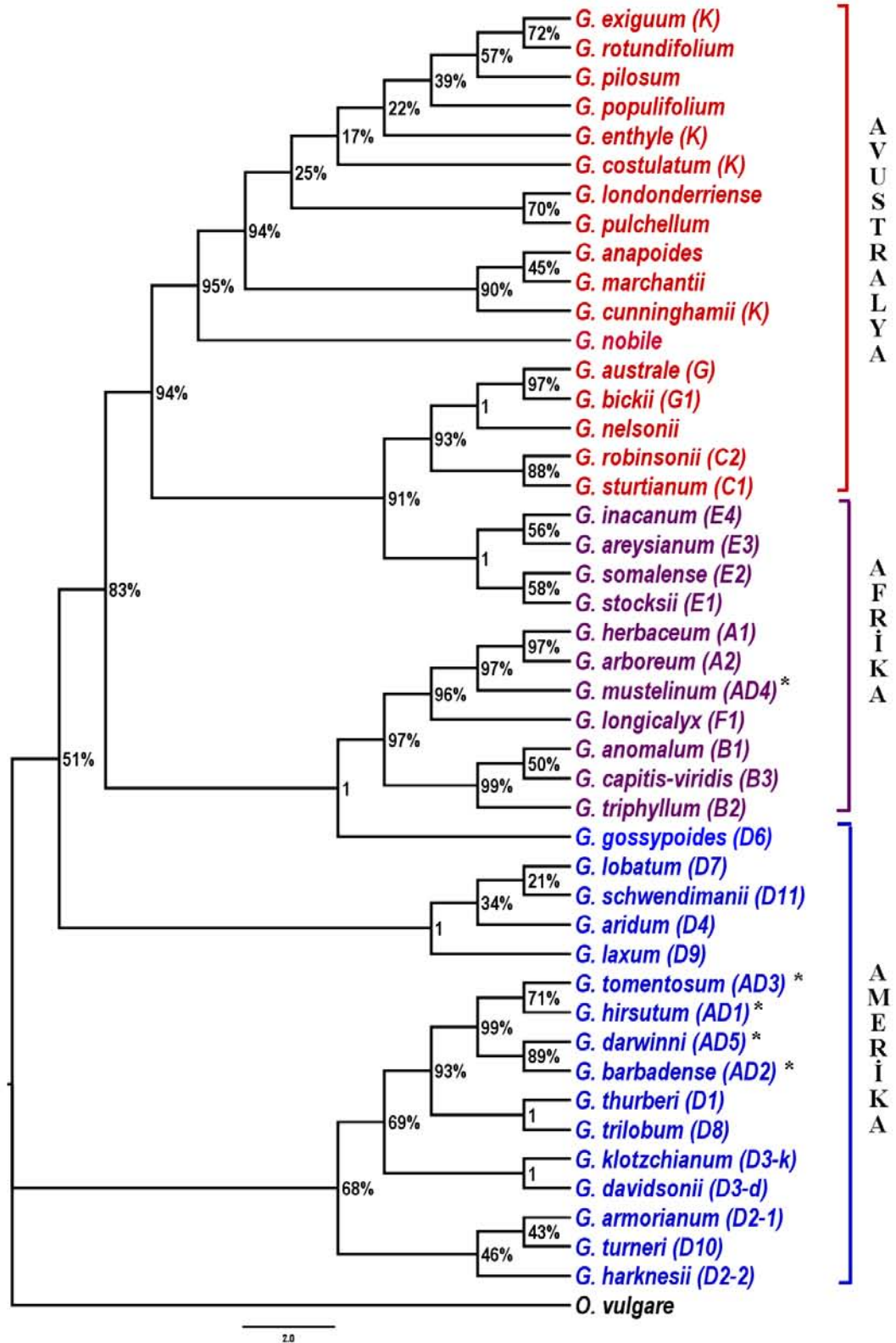
Substitüsyon modeli TN93 ve alignment yöntemi MAFFT ve parametre olarak birinci parametre ML analizi ile kullanıldığında ITS1 ve ITS2 gen filogenileri arasında belirgin farklılıklar ortaya çıkmıştır. ITS1 gen filogesinin ITS2'ye oranla daha doğru tahmin yaptığı belirlenmiştir.

ITS1-ITS2 genleri birlikte MAFFT programının birinci parametresi ile hizalanmış ve GTR substitüsyon modeli ile ML analizi gerçekleştirilmiş ve böylece bir ağaç topolojisi de burdan elde edilmiştir. Şekile 4.3'de MAFFT alignment programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki ML analizi topolojiyi gösterilmektedir. Bu topoloji günümüzde pamuk türleri arasında kabul edilmiş olan ilişkiyi tam olarak ortaya koyan ağaçtır.

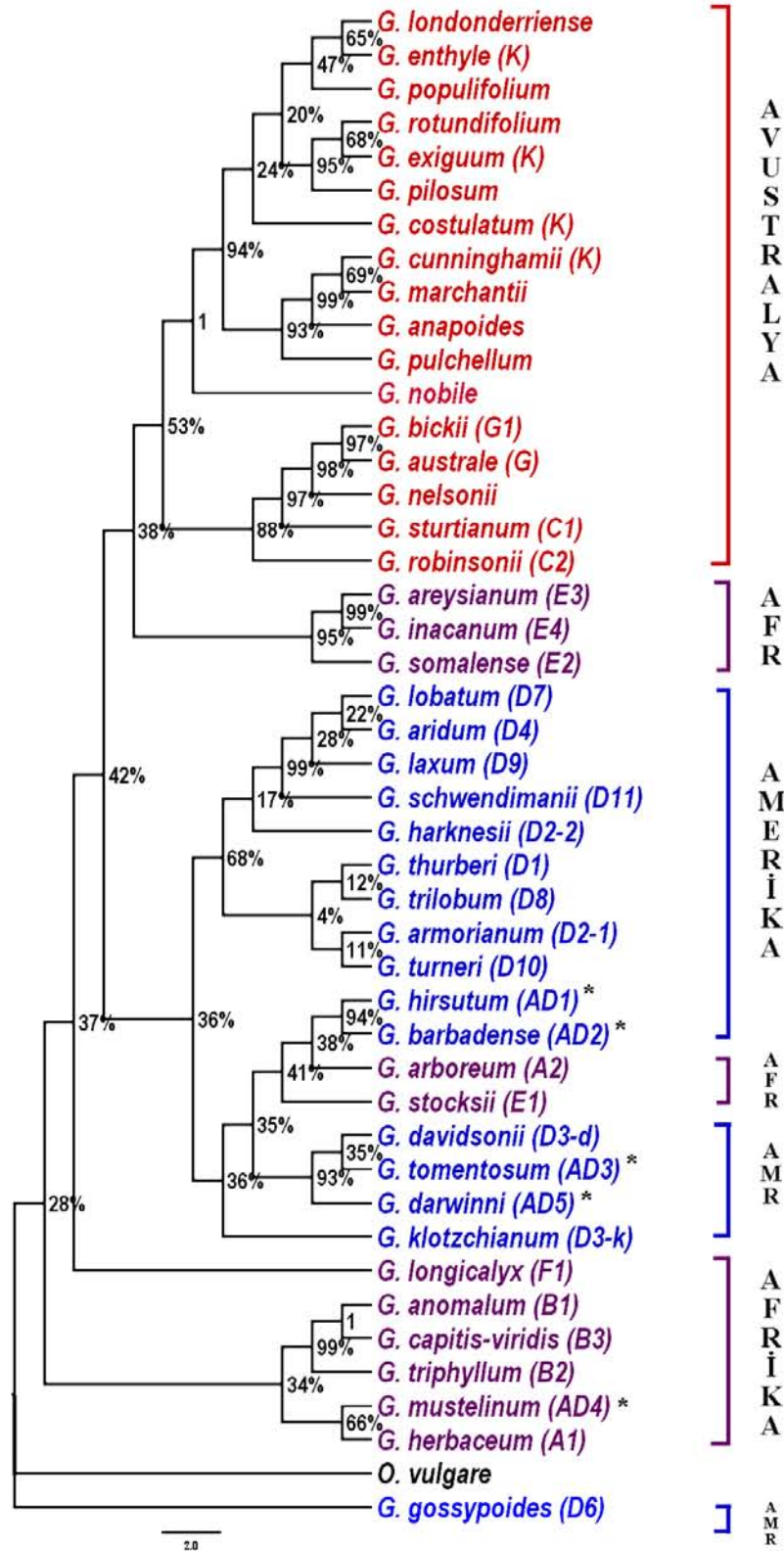
Bu topolojiye göre günümüzde 52 olarak belirlenen pamuk türlerinin Afrika kökenli pamuklardan türediği, Amerikan pamuk türlerinin *G. gossypoides* veya bu türe benzeyen yakın bir atadan türediği, kıtaların kayması ile birlikte anakarada bulunan pamuk türlerinin eşzamanlı olarak Avustralya, Amerika ve kıtalar arasındaki küçük adalar üzerinde türleştiği elde edilen topoloji verileri ile desteklenmiştir.

Bu çalışmada elde edilen filogenik ilişki kullanılarak Avustralya pamuklarının büyük çoğunluğunun K genom yapısında olduğu görülmüştür. Avustralyanın türlerinde K genomu dışında G ve C genomlarında bulunduğu ve bu genomların Afrikan türlerinden E genomlu türlere akraba olduğu belirlenmiştir.

Amerikan pamuk türlerinin MAFFT, parameter 1 ve ML analiz sonuçları yönünden değerlendirildiğinde modern allotetraploid türlerin Amerika kıtasında geliştiği ve bu modern türlerdeki D genom atalarının yine Amerikan diploid D genomlu türlerden aldığı bunun yanı sıra Afrikan A genomlu türlerin de modern pamuğun oluşumunda katkılar sağladığı teyit edilmiştir (Şekil 4.1, 4.2 ve 4.3).

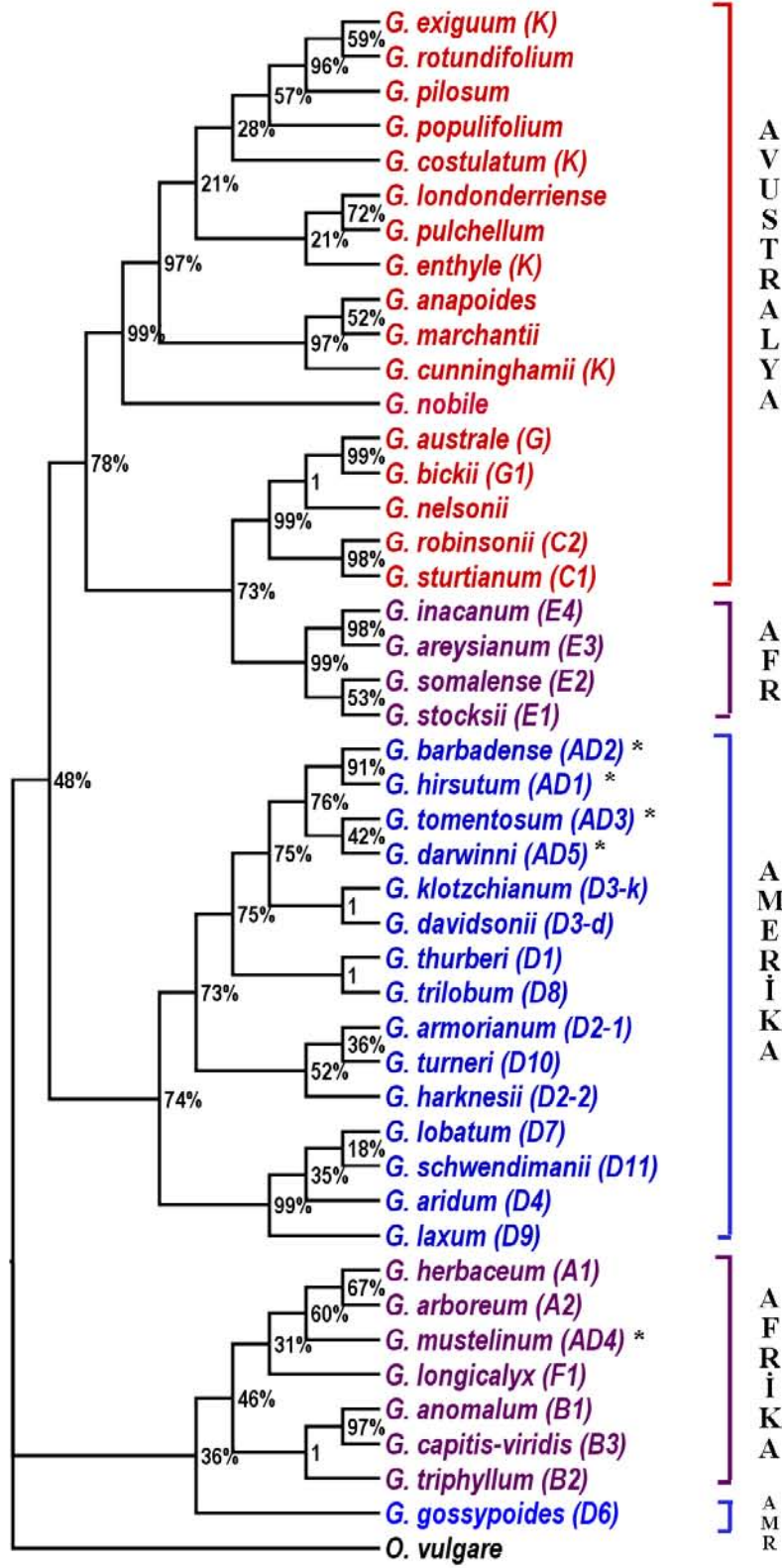


Şekil 4.1. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)



(*): Tetraploid türler

Şekil 4.2. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap probability değerleridir)



(*): Tetraploid türler

Şekil 4.3. MAFFT programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)

4.4.2. Bayes analizi sonucu elde edilen ağaçlar

Çalışmada Bayes (BI) analizi gerçekleştirilirken, ITS1 ve ITS2 bölgelerinin sekansları çalışmada kullanılmış olan her bir hizalama programına ait iki farklı parametre ile hizalanmıştır. Hizalama sonucu ITS1 ve ITS2 verilerine ait toplam onaltı dosya elde edilmiştir. Elde edilmiş bu hizalanmış dosyaların her biri GTR ve K2P substitüsyon modelleri kullanarak toplamda 32 Bayes analizi gerçekleştirilmiştir. Ayrıca ITS1-ITS2 genleri birlikte MAFFT programının birinci parametresi ile hizalanmış ve GTR substitüsyon modeli ile Bayes analizi gerçekleştirilmiş ve böylece bir ağaç topolojisi de burdan elde edilmiştir. Ancak mevcut verilerimize en uygun topolojiyi ortaya koyan yöntem ve modelere ait ağaç topolojilerinin bulgular ve tartışma kısmına yer alması uygun görülmüştür. Bahsedilen uygun ağaç topolojileri Şekil 4.4, Şekil 4.5, Şekil 4.6'da verilmiştir.

Şekil 4.4'de MAFFT programının birinci parametresi ile hizalanmış ITS1 sekansının GTR substitüsyon modeli sonucu elde edilen ağaç topolojisine bakıldığında, dış grup olarak kullanılan *O. vulgare* belirgin bir şekilde diğer taksonlardan ayrı bir şekilde gösterilmiştir. Genel olarak ağaç topolojisi üç ayrı gruba ayrılmıştır. Birinci grupta Afrika-Asya pamuklarından oluşmaktadır. İkinci grupta ise, Avustralya, Afrika-Arabistan pamuklarından oluşmuştur. Üçüncü grup ise, sadece Amerika kıtasında yayılış gösteren pamuklardan oluşmuştur.

Şekil 4.5 de yapılmış Bayes analizi MAFFT programının birinci parametresi ile hizalanmış ITS2 gen bölgesinin sekansının GTR substitüsyon modeli ile değerlendirmesi sonucu elde edilmiştir. Ağaç topolojisine bakıldığında, *O. vulgare* dış grup olarak diğer taksonlardan ayrı olarak en dışta yer almıştır. Geriye kalan ağaç topolojisi beş ayrı gruba ayrılmıştır.

Birinci grupta Avustralya pamukları yer almıştır. İkinci grupta ise, Afrika-Arabistan pamukları yer almıştır. Üçüncü grupta ise yine Avustralya pamuklarından oluşmuştur. Dördüncü grupta ise sadece D genomuna sahip Amerika pamukları yer almıştır. Son ve beşinci grupta ise, Afrika-Asya pamukları yer almıştır.

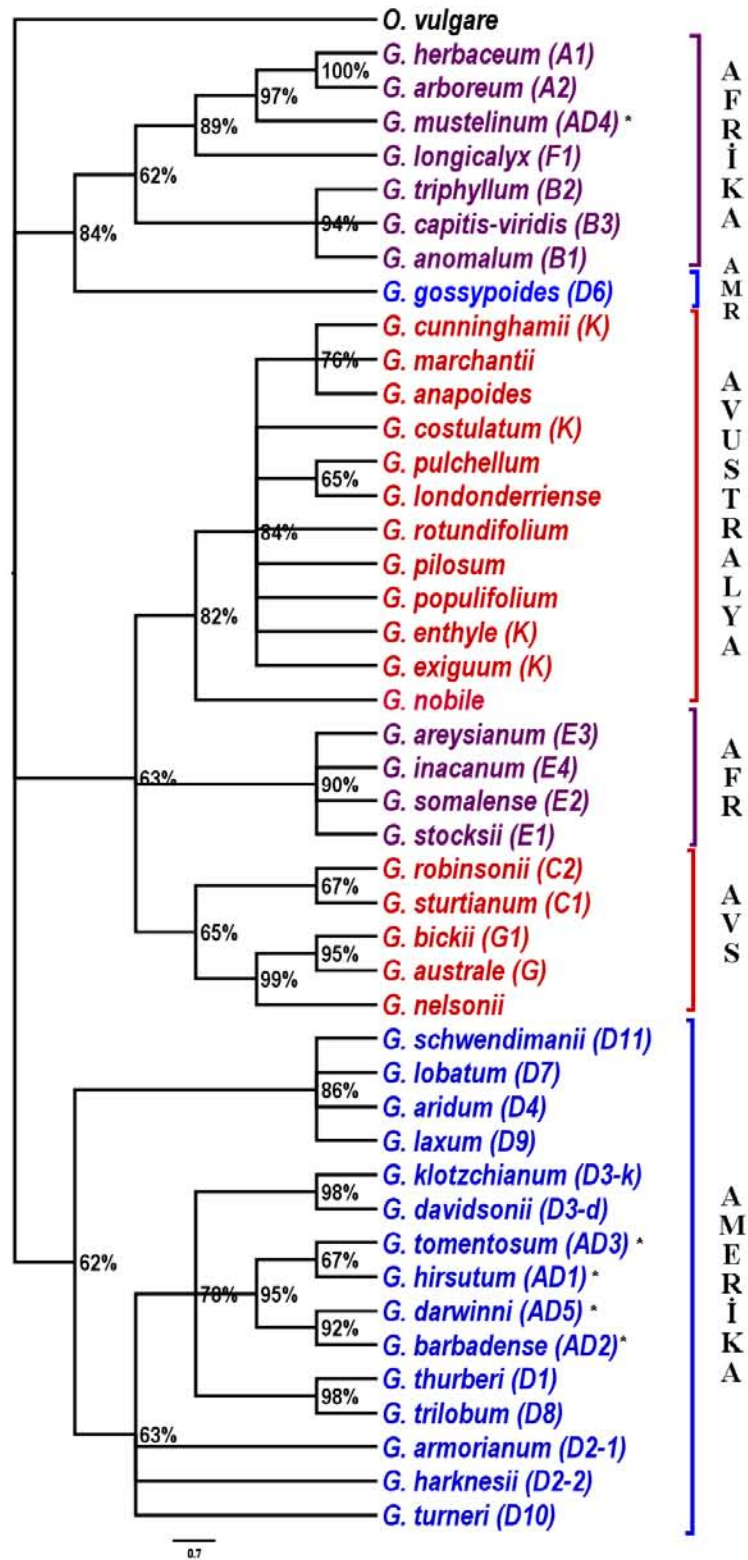
Şekil 4.6’da yine MAFFT parogramının birinci parametresi ile hizalanmış ITS1-ITS2 gen bölgelerinin GTR substitüsyon modelini kullanarak Bayes analizi sonucu elde edilmiştir.

Ağaç topolojisinde dış grup olan *O. vulgare* diğer taksonlardan ayrı olarak belirgin bir şekilde ayrılmıştır. Geriye kalan taksonlar üç ayrı grup şeklinde topolojiyi oluşturmuştur. Yukarıdan aşağı doğru birinci grup, Afrika-Asya pamuklarından oluşmuştur. İkinci grup, D genom büyüklüğüne sahip Amerika pamuklarından oluşmuştur. Üçüncü grup ise, Avusturya ve Afrika-Arabistan pamuklarından oluşmuştur.

Sonuçta yukarıda bahsedilen Bayes analizi sonucu elde edilmiş olan üç topoloji birlikte değerlendirildiğinde topoloji yapısı üzerinde gen, gen sayısı ve algoritma çeşidinin ağaç topolojisine etki ettiği belgindir.

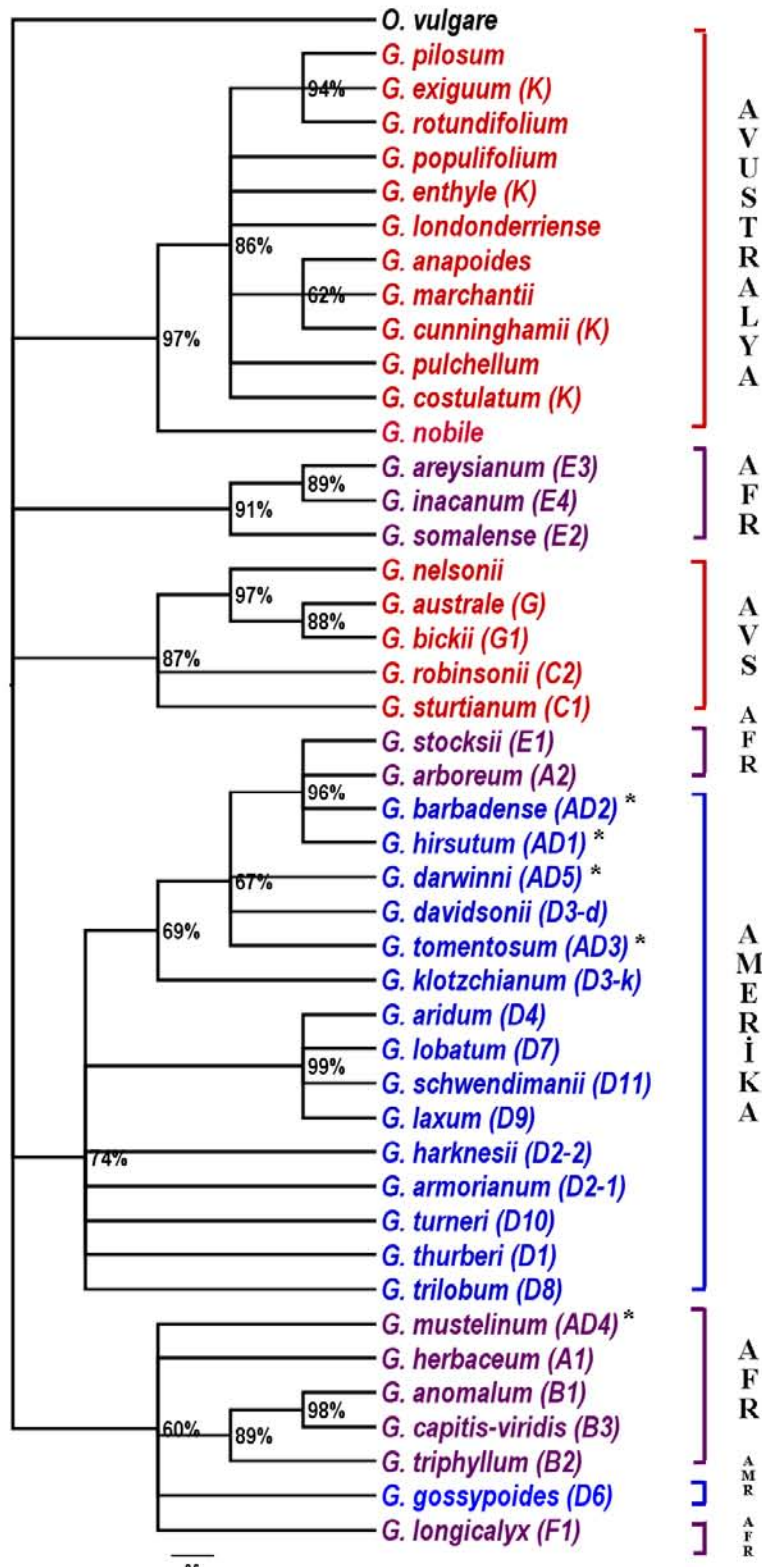
Şekil 4.4, Şekil 4.5, Şekil 4.6’da yeralan ağaç topolojisinin üçünde de bir tetraploid tür olan *G. mustelinum* ML analizinde olduğu gibi A genom büyüklüğünü temsil eden türlere yakın çıkmıştır. Geriye kalan dört tetraploid türlerde ise yine ML analizinde olduğu gibi Amerika pamukları arasında çıkmıştır. D genom büyüklüğüne sahip *G. gossypoides* de her üç topolojide Afrika pamukları arasında çıkmıştır. Bunun nedeni olarak D genomlu *G. gossypoides* türünün diğer D genomlu Amerika pamuklarına sadece bazı morfolojik, karyolojik verilere dayanarak benzediği görülmüştür. Ancak *G. gossypoides*’in rDNA sekansı beklenmedik bir şekilde Amerikanın geriye kalan D genomlarına benzemediği aksine bir Afrika genomu A genomu ile büyük oranda rekombinasyon gerçekleştirdiği görülmüştür (Zhao vd 1998).

Genel olarak bir değerlendirme yapıldığında ITS1-ITS2 dizilerinin birlikte kullanımı ile elde edil ML ağaç topolojisi BI ağaç topolojisine benzer şekilde oluşmuştur. Analizi hızı yönünden BI analizi ML analizinden daha hızlı oluşu ve benzer sonuçlar ortaya koymaları nedeni ile bu çalışmada BI analizinin ML analizine oranla daha etkin bir yöntem olduğu görülmüştür.

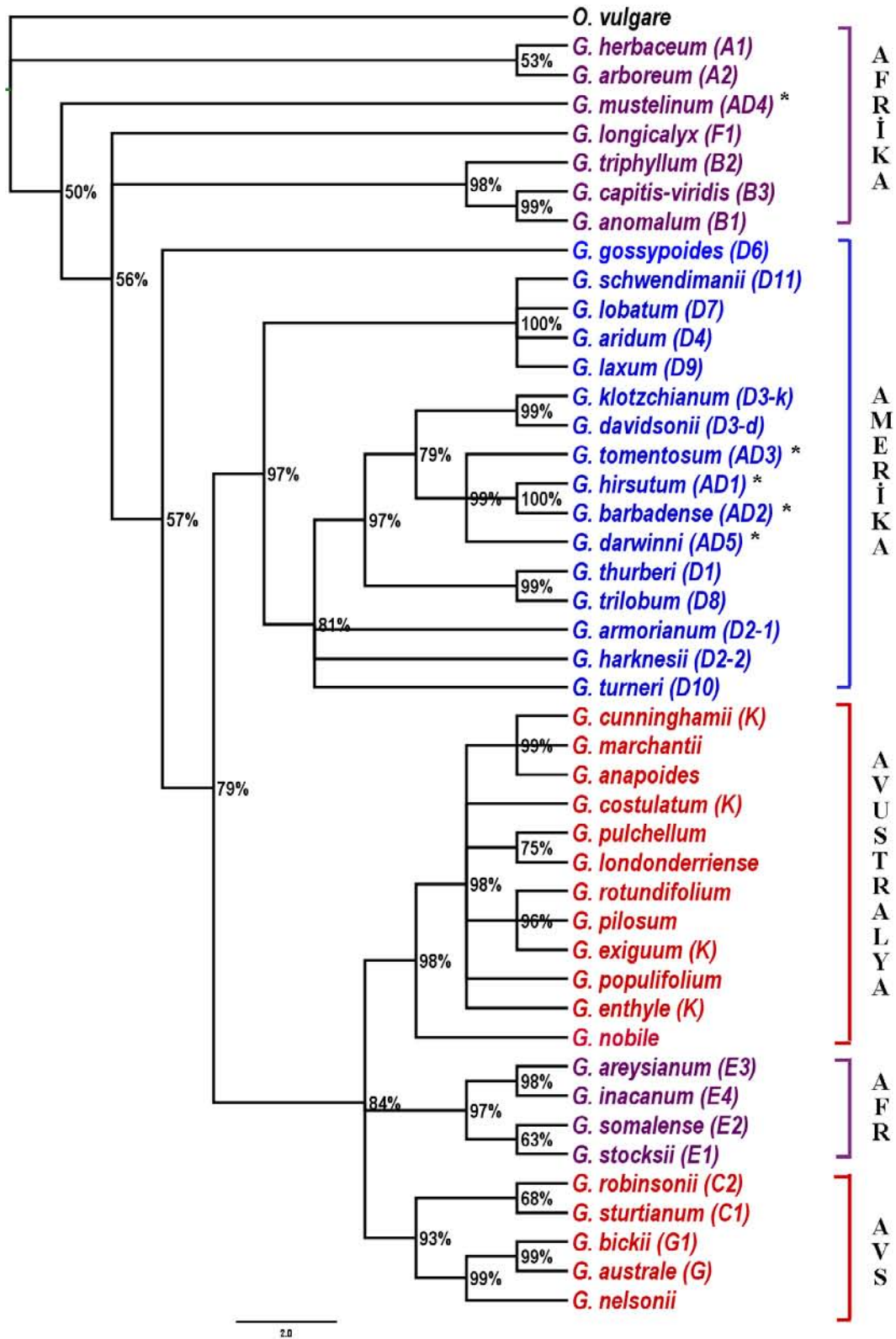


(*): Tetraploid türler

Şekil 4.4. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)



Şekil 4.5. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)



(*): Tetraploid türler

Şekil 4.6. MAFFT programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)

4.4.3. Maksimum parsimony (MP) analizi sonucu elde edilen ağaçlar

MP yöntemi verileri analiz ederken iki birim (örneğin farklı iki takson) arasında en az değişiklik gerektiren ağaç en makul olarak varsaymaktadır. Yani en az farklı olan birimler, birbirine en yakın ve benzeyen birimlerdir. Ağacın belirlenmesinde işlemin daha kısa zamanda tamamlanması için en zengin filogenetik bilgiye sahip olan az sayıdaki bölgeler kullanılır bu bölgelere parsimonik olarak bilgi verici (informative) bölgeler olarak adlandırılmaktadır. MP analizi sonucunda bu bölgelerin sayısı gen bölgesi ve kullanılan hizalama programına bağlı olarak değişiklik gösterdiği tesbit edilmiştir (Çizelge 4.7).

Çizelge 4.7. Farklı programlarla hizalanmış gen bölgelerine ait parsimonik olarak bilgi verici baz sayıları

PROGRAM VE PARAMETRELER	ITS1	ITS2
CLUSTALW (1. PRMT)	72	49
CLUSTALW (2. PRMT)	72	45
MAFFT (1. PRMT)	74	54
MAFFT (2. PRMT)	72	52
MUSCLE (1. PRMT)	72	49
MUSCLE (2. PRMT)	71	53
T-COFFEE (1. PRMT)	76	54
T-COFFEE (2. PRMT)	76	54

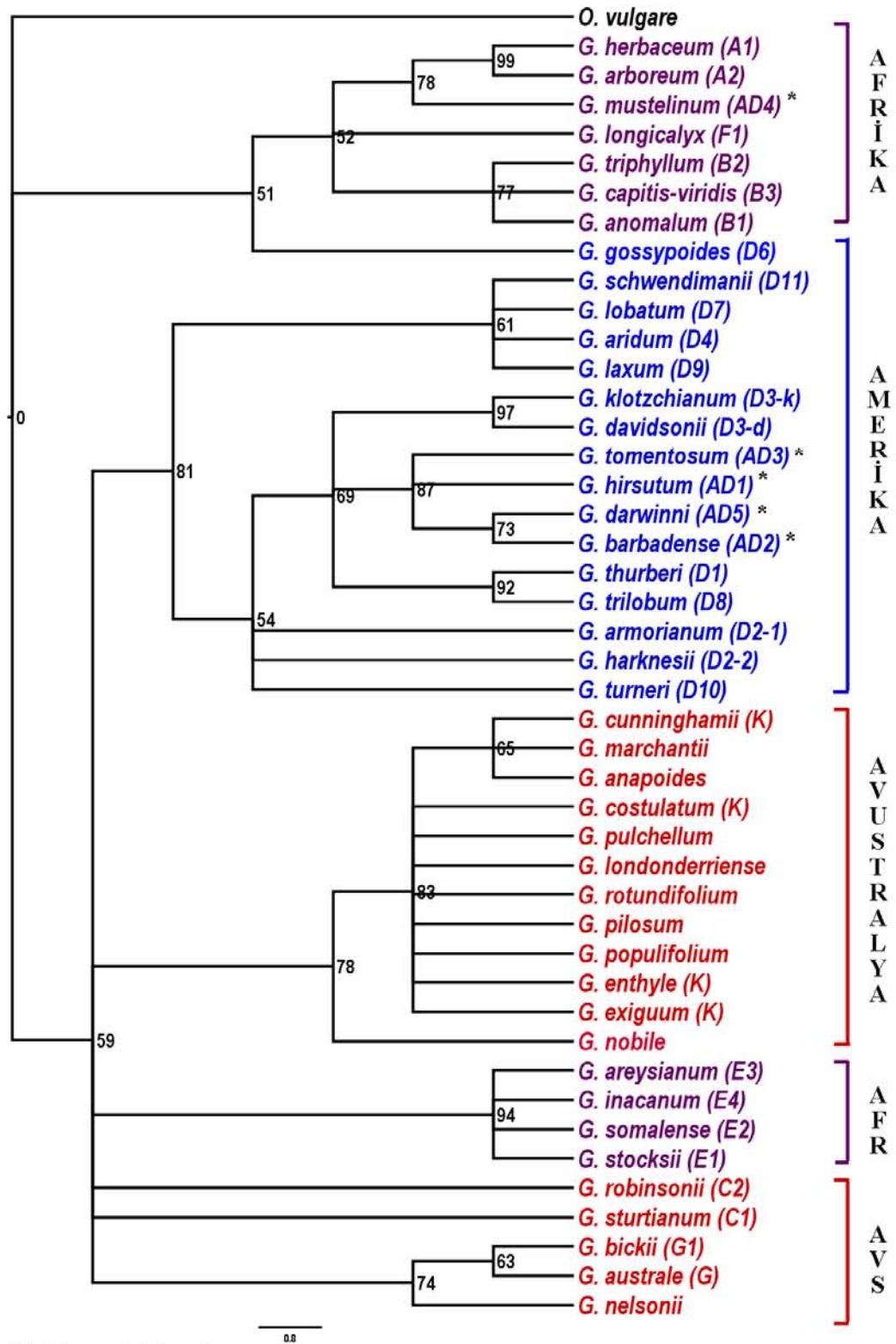
Çalışmada MP analizi gerçekleştirilirken, ITS1 ve ITS2 bölgelerinin sekansları çalışmada kullanılmış olan her bir hizalama programına ait iki farklı parametre ile hizalanmıştır. Hizalama sonucu ITS1 ve ITS2 verilerine ait toplam onaltı dosya elde edilmiştir. Elde edilmiş bu hizalanmış dosyaların her biri deneysel yöntemi ile MP analizi yapılmıştır. Analiz sonucunda toplam onaltı farklı ağaç topolojisi elde edilmiştir (Bkz. Ekler). Ayrıca ITS1-ITS2 genleri birlikte MAFFT programının birinci parametresi ile hizalanmış hōristik (deneyimsel) yöntemi ile 1000 tekrarlı bootstrap değeri ile MP analizi gerçekleştirilmiş ve böylece bir ağaç topolojisi de burdan elde edilmiştir (Bkz. Şekil 4.9). Bahsedilen uygun ağaç topolojileri Şekil 4.7, Şekil 4.8, Şekil 4.9 da verilmiştir. Şekil 4.7’deki MP analizi MAFFT programının birinci parametresi ile hizalanmış ITS1 verisinin MP analizi sonucu oluşturduğu ağaç topolojisine bakıldığında dış grup olarak *O. vulgare* belirgin bir şekilde diğer taksonların dışında çıkmıştır. Kalan

diğer pamuk taksonlarını beş ayrı filogenetik gruba ayırmıştır. Yukardan aşağı inildiğinde birinci grubun Afrika pamuklarından oluştuğu görülmektedir. Bu gruba daha önceki analizlerde olduğu gibi D6 genom büyüklüğüne sahip bir Amerika pamuk türü olan *G. gossypoides* de %51 bootstrap desteği ile dahil olmuştur.

İkinci grup *G. mustelinum* dışında diğer dört tetraploid tür ve Asya-Afrika pamukları ile beraber D genom büyüklüğünü temsil eden Amerika pamuklarından oluşmuştur. Üçüncü grup ise Avustralya pamuklarından oluşmuştur. Dördüncü grup ise Afrika/Arabistan pamuklarından oluşmuştur. Son grup olan beşinci grup da Avustralya pamuk türlerinden oluşmuştur. Yalnız bu son grup ta *G. robinsonii* ve *G. sturtianum* türleri kalan Avustralya türleri ile aynı kümeye girmemişler aksine çatallama yapmış olup direkt ana gövde ile bağlantı yapmışlardır. Bu durum MP analizinde sekans benzerliğinin düşüklüğünden kaynaklanan bir durum olabilmektedir (Kabaoglu 2007).

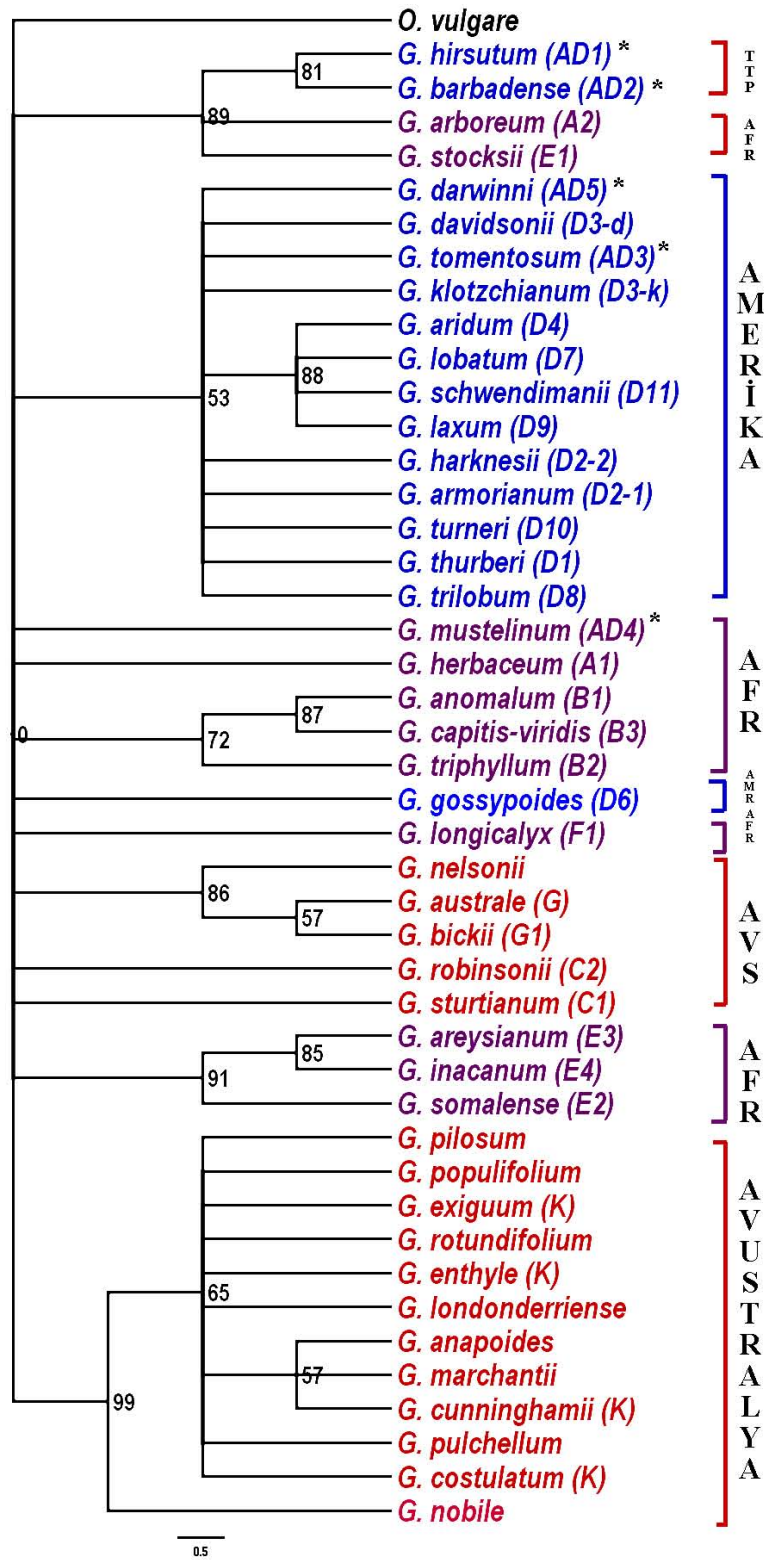
Şekil 4.8 deki ITS2 gen bölgesinin sağladığı veri ile oluşan ağaç topolojisine bakıldığında *O. vulgare* yine dış grup olarak diğer taksonlardan belirgin bir şekilde ayrılmıştır. Bu topolojideki türlerin oluşturduğu çatallama sayısı ITS1 sekansının oluşturduğu topolojideki çatallama sayısından belirgin bir şekilde fazladır. Bu topolojide D genomlu Amerika pamukları ayrı bir monofiletik grup oluşturmuştur. Benzer bir biçimde Avustralya pamuklarından bir kısım %99 bir bootstrap desteği ile bir monofiletik grup oluşturmuşlardır. Diğer iki monofiletik grup ise Afrika/Arabistan ve Afrika/Asya pamuklarının oluşturduğu gruplardır. Son monofiletik grup ise, G genomunu temsil eden Avustralya pamuklarından oluşmuştur.

Şekil 4.9'daki ITS1-ITS2 verisinin kullanıldığı ve MP analizi sonucu oluşan ağaca bakıldığında taksonla arası çatallamanın hiç olmadığı görülmüştür. Bu topolojide de *O. vulgare* yine dış grup olarak en dışta yer almıştır. Kalan pamuk taksonları üç ayrı gruba ayrılmıştır. Birinci grubu Afrika pamuklarından oluşmuştur. İkinci grup ise D genomunu temsil eden Amerika pamuklarından oluşmuştur. Üçüncü grup ise Avustralya, Afrika pamuklarından oluşmuştur. Sonuç olarak yukarda bahsedilen MP analizi sonucu elde edilen ağaçları birlikte değerlendirilğinde, en uygun topolojiyi veren ITS1-ITS2 verisinin kullanıldığı ağaç olduğu görülmüştür.



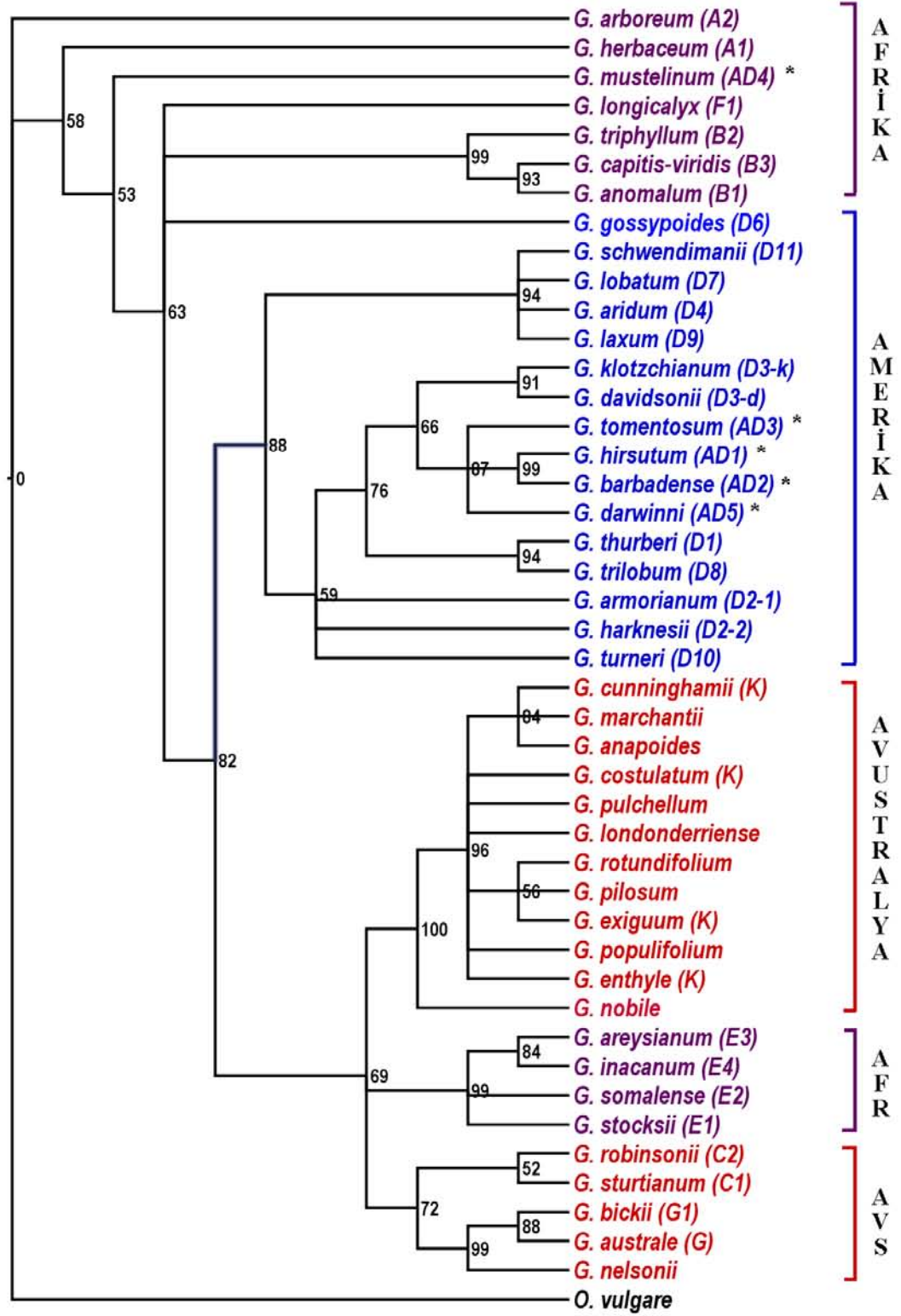
(*): Tetraploid türler

Şekil 4.7. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBTmodelindeki Parsimony analizi (Nodlarda bulunan değerler bootstrap değerleridir)



(*): Tetraploid türler

Şekil 4.8. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Parsimony analizi (Nodlarda bulunan değerler bootstrap değerleridir)



(*): Tetraploid türler

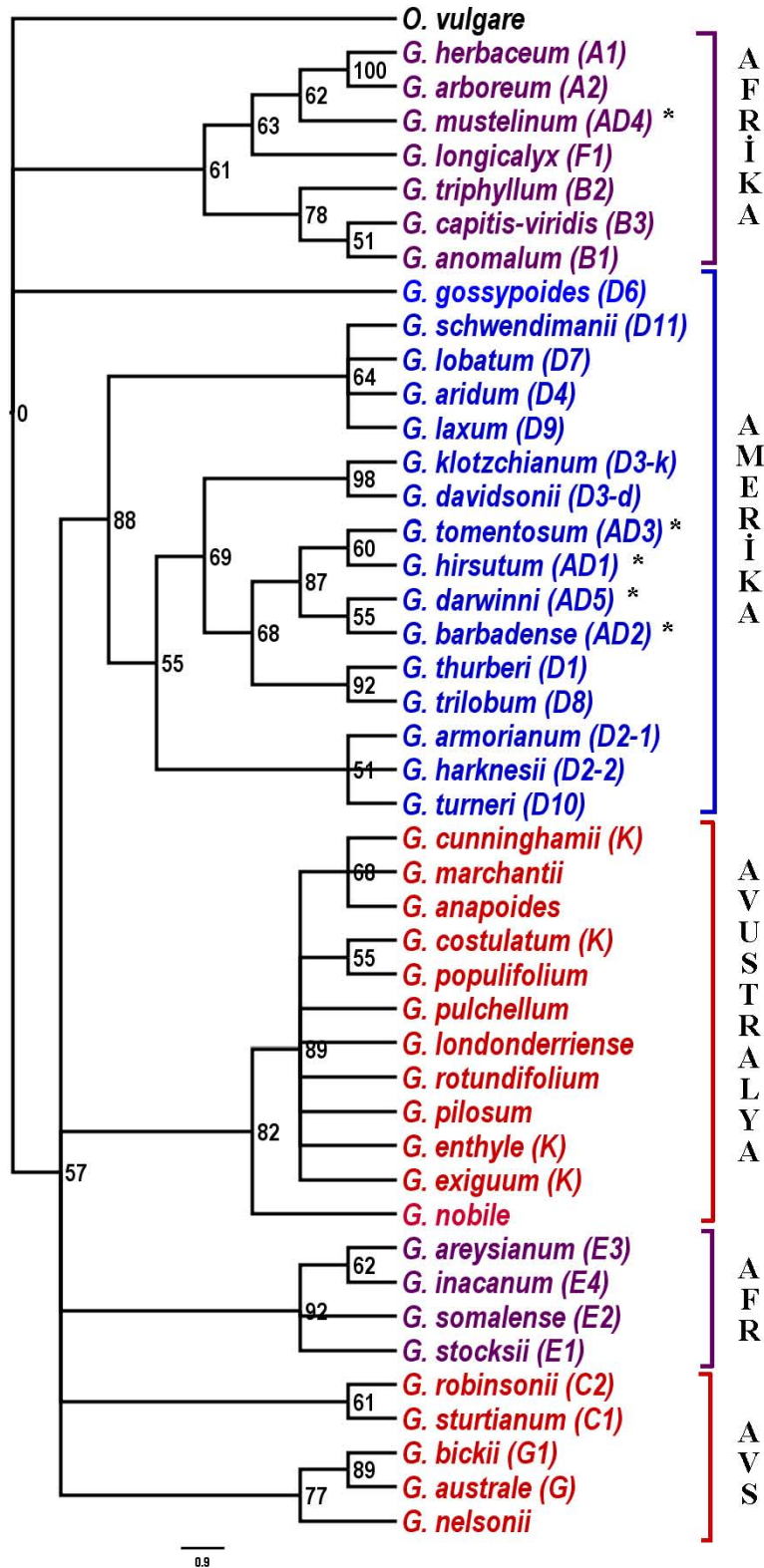
Şekil 4.9. MAFFT programının 1. parametresi ile hizalanmış ITS2-ITS2 sekansının GTR SBT modelindeki Parsimony analizi (Nodlarda bulunan değerler bootstrap değerleridir)

4.4.4. Distance analizi sonucu elde edilen ağaçlar

Çalışmada 1000 tekrarlı bootstrap Neighbour-joining (NJ) algoritması uygulanarak Distance analizi gerçekleştirilirken, ITS1 ve ITS2 bölgelerinin sekansları çalışmada kullanılmış olan her bir hizalama programına ait iki farklı parametre ile hizalanmıştır. Hizalama sonucu ITS1 ve ITS2 verilerine ait toplam onaltı dosya elde edilmiştir. Elde edilmiş bu hizalanmış dosyaların her biri GTR ve K2P substitüsyon modelleri kullanılarak Distance analizleri gerçekleştirilmiştir. Analiz sonucunda toplam otuziki farklı ağaç topolojisi elde edilmiştir (Bkz. Ekler). Ayrıca ITS1-ITS2 genleri birlikte MAFFT programının birinci parametresi ile hizalanmış ve GTR substitüsyon modeli ile Distance analizi gerçekleştirilmiş ve böylece bir ağaç topolojisi de burdan elde edilmiştir (Bkz. Şekil 4.12). Seçilmiş olan ağaç topolojileri Şekil 4.10, Şekil 4.11, Şekil 4.12’de verilmiştir.

Şekil 4.10'daki Distance analizi MAFFT programının birinci parametresi ile hizalanmış ITS1 verisinin Distance analizi sonucu oluşturduğu ağaç topolojisine bakıldığında dış grup olarak *O. vulgare* belirgin bir şekilde diğer taksonların dışında çıkmıştır. Kalan diğer pamuk taksonlarını beş ayrı filogenetik gruba ayırmıştır. Yukarıdan aşağı inildiğinde birinci grubunu Afrika pamukları oluşturmuştur. İkinci grup *G. mustelinum* dışında diğer dört tetraploid tür ve Asya-Afrika pamukları ile bearaber D genom büyüklüğünü temsil eden Amerika pamuklarından oluşmuştur. Üçüncü grup ise Avustralya pamuklarından oluşmuştur. Dördüncü grup ise Afrika/Arabistan pamuklarından oluşmuştur. Son grup olan beşinci grup da Avustralya pamuk türlerinden oluşmuştur. Yalnız bu son grup ta *G. robinsonii* ve *G. sturtianum* türleri kalan Avustralya türleri ile aynı kümeye girmemişlerdir aksine çatallama yapmış olup direkt ana gövde ile bağlantı yapmışlardır. Bu durum Distance analizinde sekans benzerliğinin düşüklüğünden kaynaklanan bir durum olabilmektedir.

Şekil 4.11'deki ITS2 gen bölgesinin sağladığı veri ile oluşan ağaç topolojisine bakıldığında *O. vulgare* yine dış grup olarak diğer taksonlardan belirgin bir şekilde ayrılmıştır. Bu topolojideki türlerin oluşturduğu çatallama sayısı ITS1 sekansının oluşturduğu topolojideki çatallama sayısından belirgin bir şekilde fazladır.



(*): Tetraploid türler

Şekil 4.10. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Distance analizi (Nodlarda bulunan değerler bootstrap değerleridir)

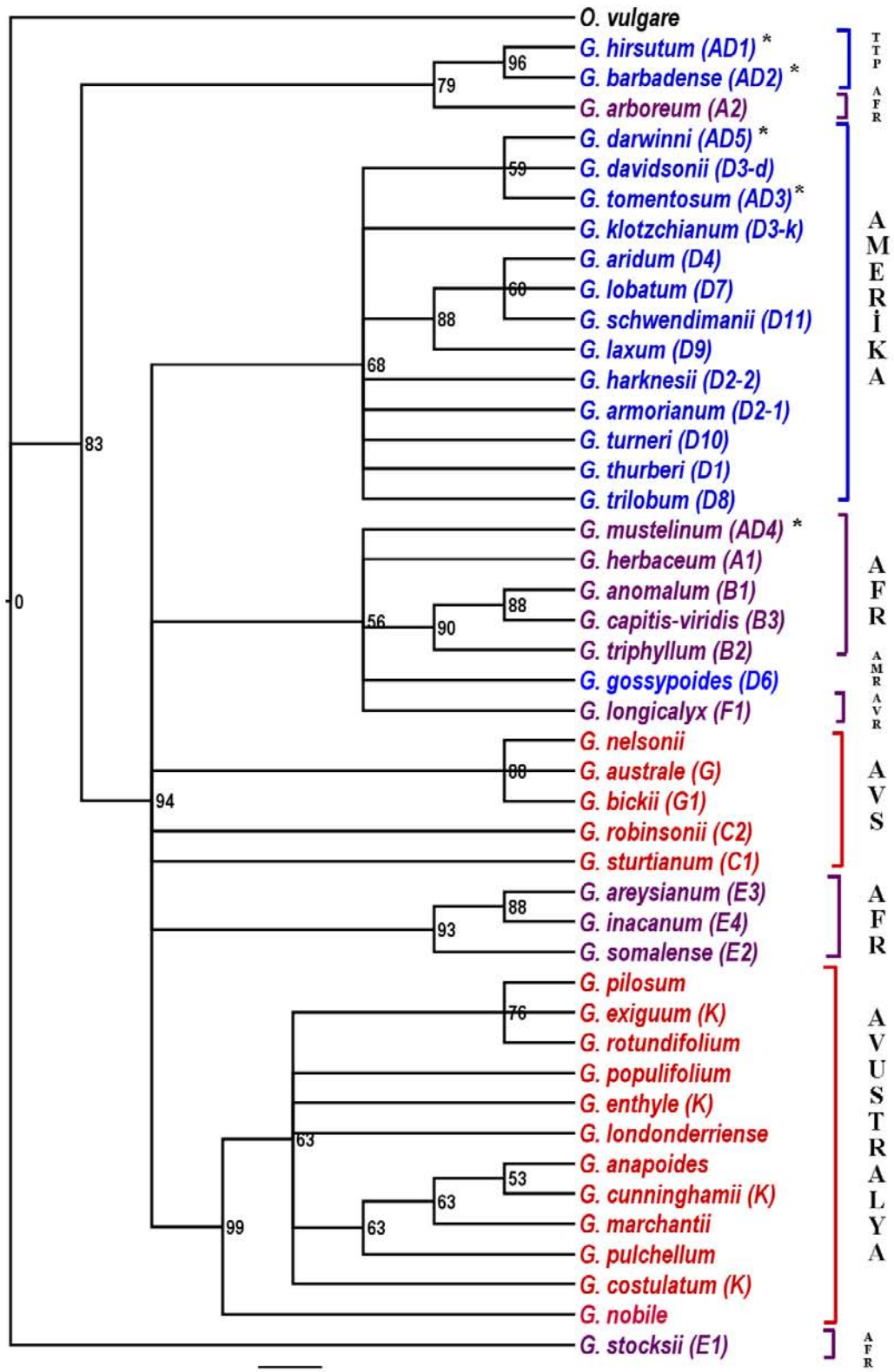
Bu topolojide D genomlu Amerika pamukları ayrı bir monofiletik grup oluşturmuştur. Benzer bir biçimde Avustralya pamuklarından bir kısım %99 bir bootstrap desteği ile grup oluşturmışlardır. Diğer iki grup ise Afrika/Arabistan ve Afrika/Asya pamuklarının oluşturduğu gruplardır. Son grup ise, G genomunu temsil eden Avustralya pamuklarından oluşmuştur.

Şekil 4.12'deki ITS1-ITS2 verisinin kullanıldığı ve Distance analizi sonucu oluşan ağaca bakıldığında taksonla arası çatallamanın hiç olmadığı görülmüştür. Bu topolojide de *O. vulgare* yine dış grup olarak en dışta yer almıştır. Kalan pamuk taksonları üç ayrı monofiletik gruba ayrılmıştır. Birinci grubu Afrika pamuklarından oluşmuştur. İkinci grup ise D genomunu temsil eden Amerika pamuklarından oluşmuştur. Üçüncü grup ise Avustralya, Afrika pamuklarından oluşmuştur.

Sonuç olarak yukarda bahsedilen Distance analizi sonucu elde edilen ağaçları birlikte değerlendirilğinde, en uygun topolojiyi veren MP analizinde de olduğu gibi ITS1-ITS2 verisinin kullanıldığı ağaç olduğu görülmüştür. ITS1 verisinin oluşturduğu topoloji de ITS2 den çok daha uygun ve düzenli bir topoloji ortaya koymuştur.

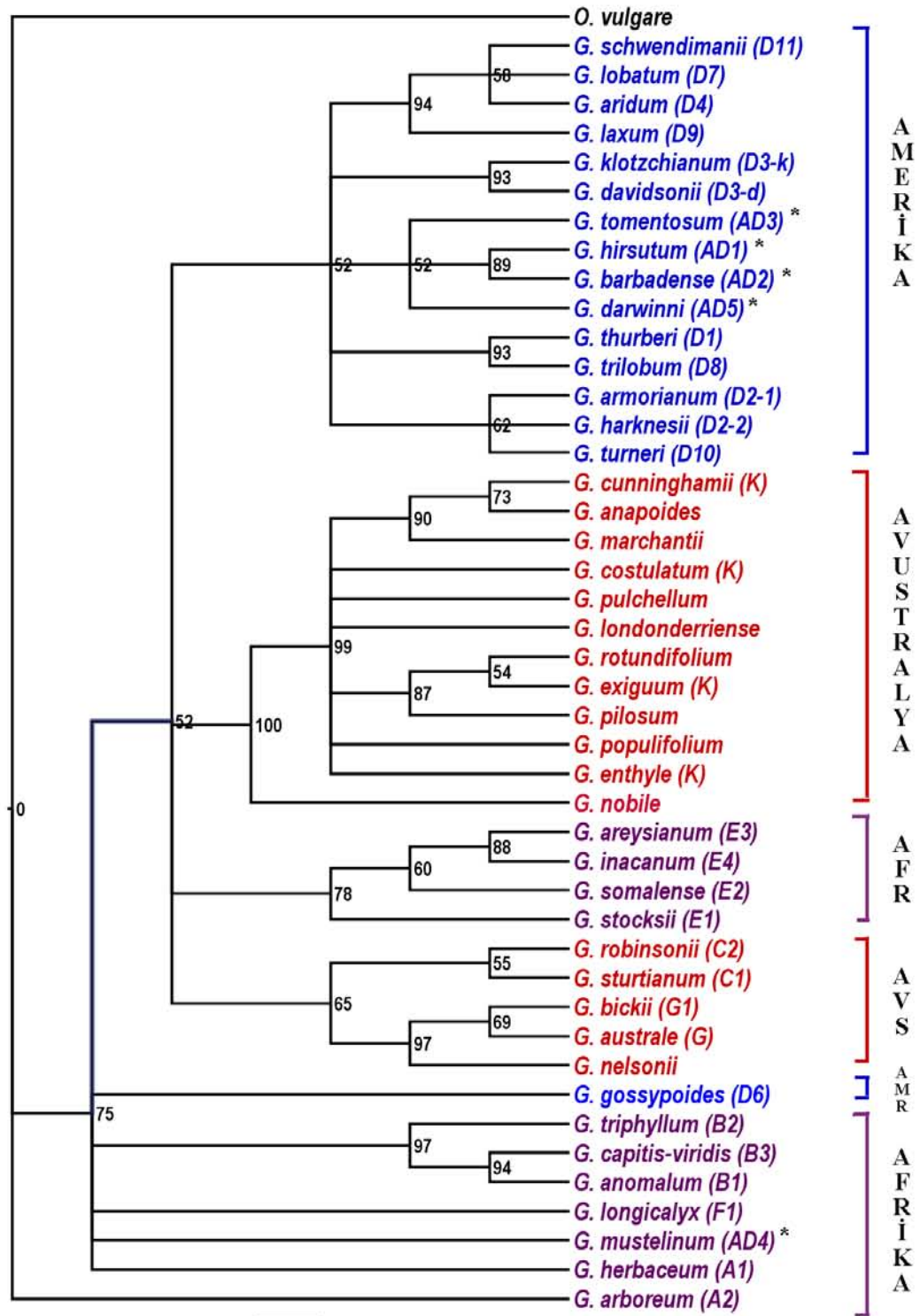
Distance analizi sonucu meydana gelen ağaç topolojileri, MP analizi sonucu oluşan ağaç topolojileri ile benzer olduğu görülmüştür. Sadece nodlarda seç-bağla (bootstrap) değerleri farklılık göstermiştir.

Genel bir değerlendirme yapıldığında D ve MP analizleri ML ve BI analizlerine göre dah az bilgi verici analizler olarak tespit edilmiştir. Ancak D ve MP analizleride ML ve BI analizlerinde olduğu gibi alignment, alignment parameterleri, gen sayısı ve gen bölgesinden etkilenmektedir. MP ve D analizleri karşılaştırıldığında ise MP analizlerinin D analizlerine göre daha az filogeni bilgisi içermesine rağmen bir ihtiyat analizi olması nedeni ile tespit etmiş olduğu filegenik ilişkinin D analizlerine göre daha kesin olduğu söylenebilir.



(*): Tetraploid türler

Şekil 4.11. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Distance analizi (Nodlarda bulunan değerler bootstrap değerleridir)



(*): Tetraploid türler

Şekil 4.12. MAFFT programının 1. parametresi ile hizalanmış ITS1-ITS2 sekansının GTR SBT modelindeki Distance analizi (Nodlarda bulunan değerler bootstrap değerleridir)

4.5. Uygun Markır (barkod) Geninin Belirlenmesi

Markır gen, bir bitki grubu/gruplarına ait takson seviyelerini (Alem, Şube, Sınıf, Takım, Familya, Cins, Tür, Altür, Varyete) tanımlamak için belirleyici bir tekniktir. Bu teknik, klasik yöntemlerle teşhis edilemeyen koruma biyoloji ve ekolojisi çalışmalara büyük katkı sağlamaktadır. Markır gen çalışmaları hayvanlar aleminde iyice gelişmişken bitkiler aleminde ise henüz universal olarak kabul edilen gen bulunmamaktadır (Lahaye vd 2007). Markır gen taksonları teşhis etme özelliği olan DNA'nın küçük bir bölgesidir. Fakat çok geniş bir taksonomik çalışma için uygun markır gen bulmak ya da ortaya koymak kolay değildir. Markır genin özelliği, taksonlar arası sekans değişiminin yeterince olması olması gerekmektedir yani taksonlar arasında polimorfik bir özellik taşınmalıdır ki taksonları birbirinden kolay bir şekilde edebilsin. Ancak polimorfizm seviyesinin tür içinde düşük olması istenirken türler arası polimorfizmin ise yüksek olması beklenir.

Markır genin taksonomik amaçlı kullanıldığı çalışmalarda, yakın ilişkili bir bitki grubundan, uzak ilişkili bir grubuna kadar çeşitli sayıda bitki grubunun olmasına ihtiyaç duyar. Bu çalışmada: Otsu, odunsu, monokot, dikot özelliklerini taşıyan ve çiçekli bitki grubunu büyük oranda temsil eden çeşitli sayılarda tür içeren onbeş familya için markır gen belirlenmeye çalışılmıştır. Çalışmada seçilen bitki gruplarının çeşitliliği ve bu grupların değişik habitatları bakımından ayrıcalıklı bir çalışma niteliğinde olmuştur. Çalışmada kullanılan orkid grupların teşhisinin çok zor olduğu bilinmektedir. Bu yüzden özellikle bitki steril olduğu durumlarda markır gen belirlemek için çok uygun bir bitki grubu olma niteliğini taşır (Lahaye vd 2007).

Bu çalışmada her bir markır gen (*matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB*) için ML ve Bayes filogenetik analiz yöntemleri ve nodlarda ki posterior probability değerleri göz önüne alınarak değerlendirilmiştir. Markır gen belirlenirken mevcut taksonların familya, cins ve tür seviyesinde kümeleşme yapıp yapmadıkları ve aynı zamanda türlerin monofiletik olma kriterleri göz önüne alınmıştır.

4.5.1. Markır gen için yapılan maksimum likelihood (ML) analizi ve elde edilen ağaçlar

Markır gen bulmak için yapılan ML analizi için toplam beş gen (*matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB*) kullanılmıştır. Genlere ait sekanslar MAFFT programı kullanılarak hizalama yapılmıştır. Ayrıca verileri analiz etmek için GTR substitüsyon modeli kullanılmıştır. ML analizi sonucu çıkan ağaç topolojileri incelendiğinde Şekil 4.13 de *matK* genine ait ağaç topolojisi verilmiştir. *matK* genin analiz sonucu ortaya koyduğu ağaç topolojisinde, çalışmada kullanılmış olan onbeş familyanın ve cinslerin tümü kendi aralarında uyumlu bir şekilde kümelenmiştir. Oluşan kümeler yüksek posterior probability değerleri ile desteklenmiştir. Ayrıca monokotil bitki grubunu oluşturan *Orchidaceae*, *Arecaceae*, *Velloziaceae* familyalarını %100 bir posterior probability desteği ile bir arada göstermiştir. Ancak, *Proteaceae* familyasına ait *F. rochetina* türünü ve *Arecaceae* familyasına ait *H. coriacea* türlerini polifiletik göstermiş olup kardeş türleri ile tek kümede olacak şekilde göstermemiştir.

Şekil 4.14'de *accD* gen bölgesinin ML analizi sonucu oluşan ağaç topolojisi verilmiştir. Bu topolojide, familyalar bazında uyumlu bir kümeleşme görülmüştür. Ancak monokotil bitki grubunu oluşturan *Orchidaceae*, *Arecaceae*, *Velloziaceae* familyaların ikisini *Arecaceae*, *Velloziaceae*'yi %84 posterior probability desteği ile bir arada gösterirken *Orchidaceae* familyasını bunlardan ayrı göstermiştir. Ayrıca *Moraceae*, *Proteaceae*, *Orchidaceae*, *Anacardiaceae* familyasındaki türleri ve *Anacardiaceae* familyasında ki cinsleri polifiletik göstermiştir.

Şekil 4.15'de *rpoC1* gen bölgesinin oluşturduğu ağaç topolojisi verilmiştir. Bu topolojide de familyalar tek kümede gösterilmişken, *Malvaceae* familyasına ait *G. flavescens* türü, *Euphorbiaceae* familyasına ait *C. gratissimus* ve *C. megalobotrys* türleri, *Orchidaceae* familyasına ait *E.alata*, *P. fragrans* türleri, *Arecaceae* familyasına ait *H. petersiana* türünü ve *Proteaceae* familyasına ait *F. rochetiana* türünü polifiletik göstermiş olup kardeş türü/türleri ile aynı küme altında göstermemiştir. Ayrıca *Orchidaceae* familyasına ait *Encylia* ve *Prosthechea* cinsleri de polifiletik göstermiştir.

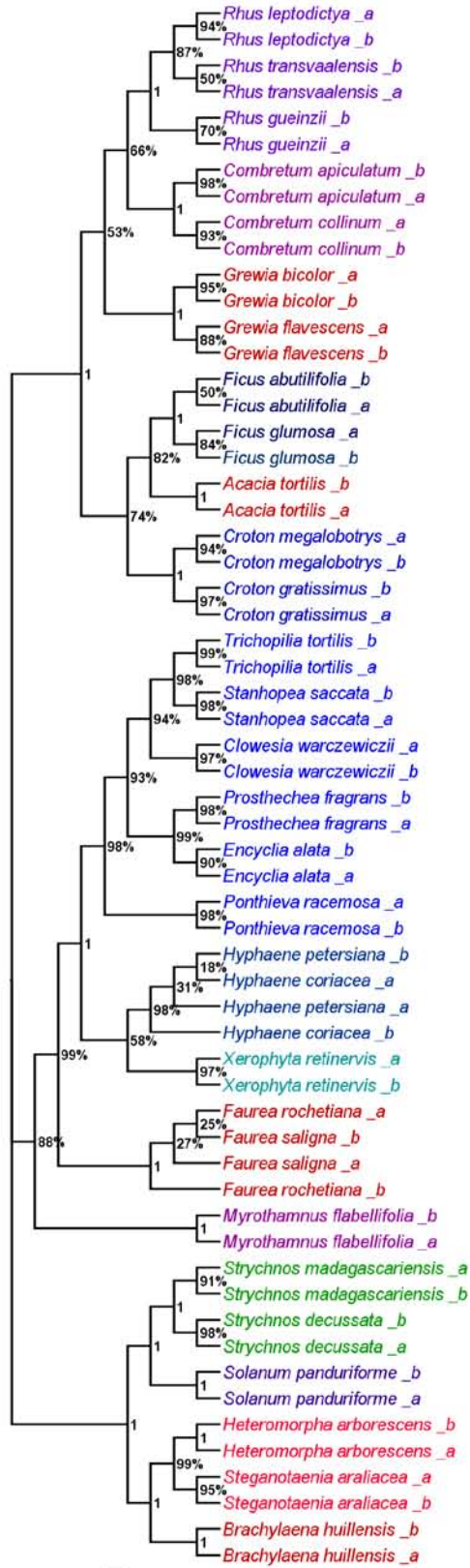
Şekil 4.16'da *trnH-psbA* gen bölgesinin analizi sonucu elde edilen ağaç topolojisi verilmiştir. Bu topoijide de familyalar uyumlu olacak şekilde bir arada kümeleşmişlerdir. Ayrıca monokotil bitki grubunu oluşturan *Arecaceae*, *Velloziaceae*, *Orchidaceae* familyaları %95 posterior probability desteği ile bir arada göstermiştir. Cins ve tür bazında büyük oranda bir monofiletiklik görülmüştür. Ancak *Proteaceae* familyasına ait *F. saligna* ve *F. rochetina* türleri ile *Moraceae* familyasına ait *F. glumosa*, *F. abutifolia* türlerinde ve *Anacardiaceae* familyasına ait *R. transvaalensis* türünde bir polifiletiklik görülmüştür.

Şekil 4.17'de ise *rpoB* gen bölgesinin analizi sonucu elde ağaç topolojisi verilmiştir. bu ağaç topolojisinde de familyalar uyumlu bir kümeleme oluşturmuşlardır. Ayrıca monofiletik bitki grubunu oluşturan *Arecaceae*, *Velloziaceae*, *Orchidaceae* familyalarını da %96 posterior probability desteği ile bir arada göstermiştir. Ancak *Anacardiaceae* familyasına ait *R. transvalensis*, *Malvaceae* familyasına ait *G. bicolor*, *Arecaceae* familyasına *H. petersiana*, *Proteaceae* familyasına ait *F. saligna* türlerini polifiletik göstermiş olup türleri kardeş türleri ile aynı kümede göstermemiştir.

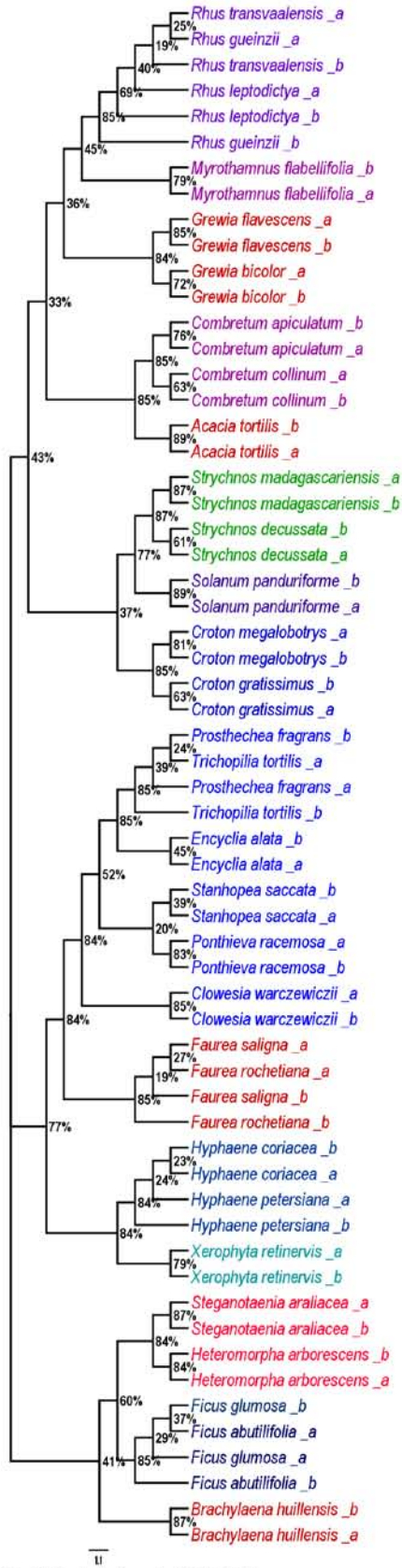
Sonuç olarak markır gen bulmak için yapılmış olan ML analizi değerlendirildiğinde *matK* ve *trnH-psbA* genlerinin markır gen olma başarılarının birbirine çok yakın olduğu görülmüştür. İkisi de biri ortak iki familya türlerinde bir polifiletiklik göstermiştir. Ancak *matK* geni monokotil gruplarını %100 posterior probability desteği ile bir arada göstermişken *trnH-psbA* geni ise bu grupları %95 posterior probability desteği ile bir arada göstermiştir.

Krees vd (2007) yaptıkları bir çalışmada *trnH-psbA* geninin karasal bitkiler için uygun bir markır geni olduğunu belirtmişlerdir. Aynı zamanda Lahey vd (2007) yaptıkları bir çalışmada *trnH-psbA* gen bölgesinin türleri birbirinden ayırmada başarılı olduğunu belirtmişlerdir. Ayrıca Kress vd (2007) *matK* gen bölgesinin için türleri birbirinden ayırmak için uygun bir markır gen olmadığı ifadesinin tersine Lahaye vd (2007) yaptıkları çalışmada *matK* geninin türleri ayırmada uygun bir markır gen olduğunu belirtmişlerdir. Diğer taraftan bu çalışma sonuçların dayanarak iki cins ve toplam dokuz türde polifiletiklik göstermiş olan *accD* gen bölgesi bu çalışma açısından markır gen olarak uygun görülmemiştir. Bunu takip eden *rpoC1* gen bölgesi ise iki cins

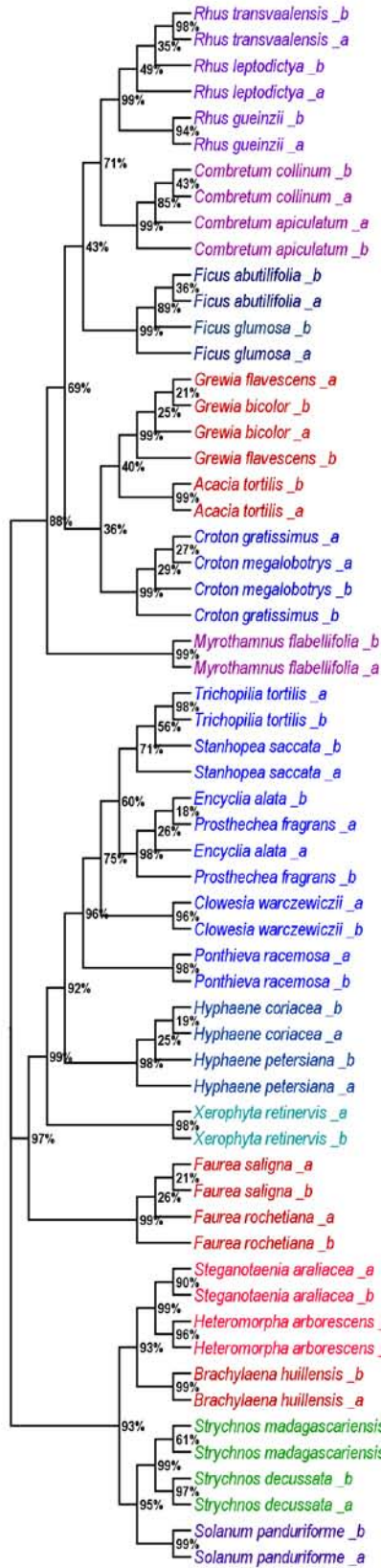
ve toplam yedi türde polifiletiklik göstermiştir. *rpoB* de toplam beş türde polifiletiklik göstermiştir. Ayrıca Chase vd (2007) de yaptıkları bir çalışmada *rpoC1*, *rpoB* gen bölgelerinin markır gen için kullanışlı olmadığı ifadeleri sonuçlarımızı desteklemiştir. Ama sonuç olarak kloroplast genomu yavaş evrimleştiği için marker gen sağlama konusunda birden fazla kloroplast geninin birleştirilmiş olarak kullanılmasına ihtiyaç duyulmuştur (Kress ve Chase 2007).



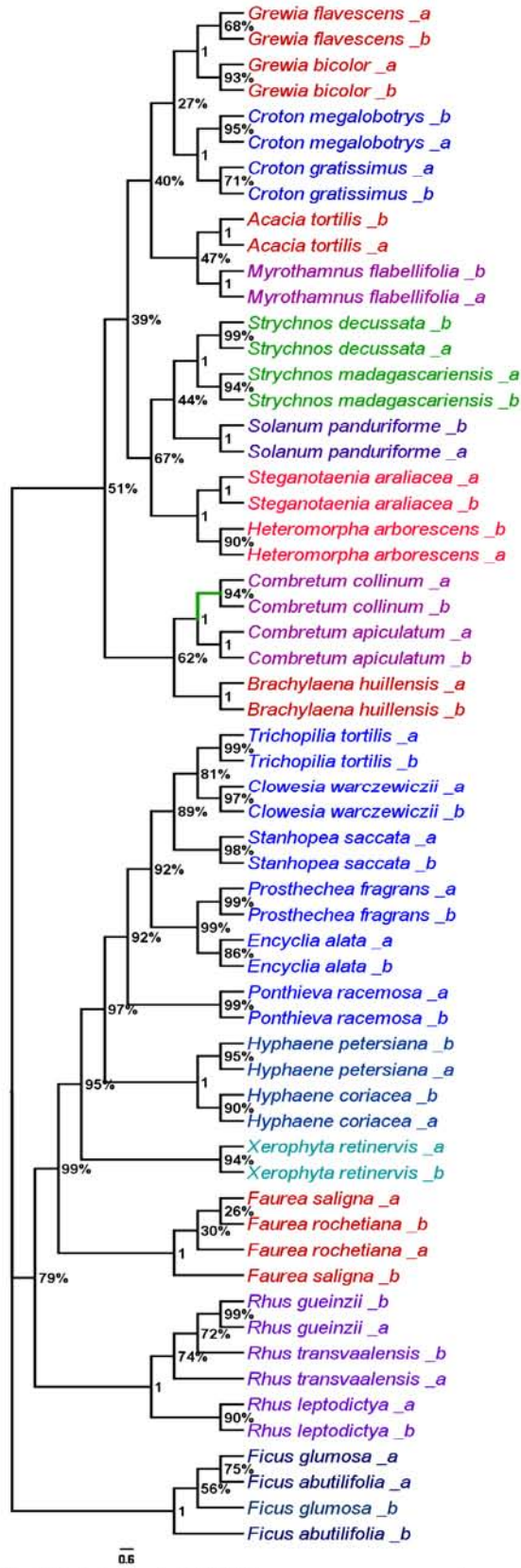
Şekil 4.13. MAFFT programının 1. parametresi ile hizalanmış *matK* gen sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)



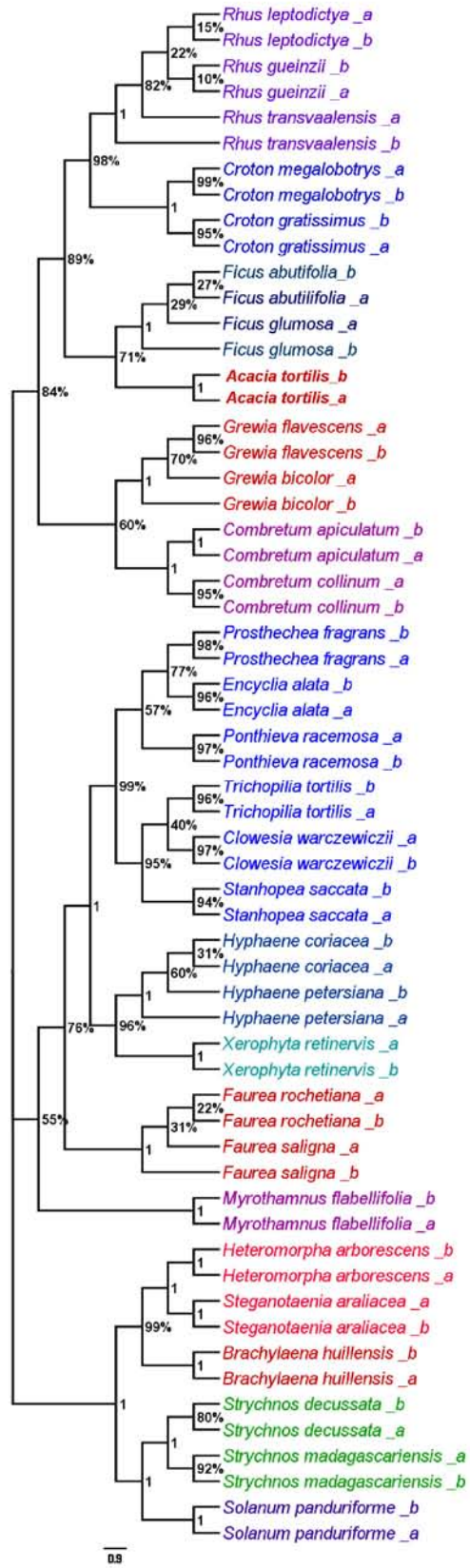
Şekil 4.14. MAFFT programının 1. parametresi ile hizalanmış *accD* gen sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)



Şekil 4.15. MAFFT programının 1. parametresi ile hizalanmış *rpoC1* gen sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)



Şekil 4.16. MAFFT programının 1. parametresi ile hizalanmış *trnH-psbA* gen sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)



Şekil 4.17. MAFFT programının 1. parametresi ile hizalanmış *rpoB* gen sekansının GTR SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler bootstrap değerleridir)

4.5.2. Markır gen için yapılan Bayes analizi ve oluşturulan ağaçlar

Markır gen bulmak için yapılan Bayes analizi için toplam beş gen (*matK*, *trnH-psbA*, *accD*, *rpoC1*, *rpoB*) kullanılmıştır. Genlere ait sekanslar MAFFT programı kullanılarak hizalama yapılmıştır. Ayrıca verileri analiz etmek için GTR substitüsyon modeli kullanılmıştır. Bayes analizi sonucu çıkan ağaç topolojileri incelendiğinde Şekil 4.18’de *matK* genine ait ağaç topolojisi verilmiştir. *matK* genin analiz sonucu ortaya koyduğu ağaç topolojisinde, çalışmada kullanılmış olan *Orchidaceae* dışında geride kalan ondört familya ve cinslerin tümü kendi aralarında uyumlu bir şekilde kümelenmiştir. Oluşan kümelerin büyük bir kısmı yüksek posterior probability değerleri ile desteklenmiştir. Ancak monokotil bitki grubunu oluşturan *Orchidaceae*, *Arecaceae*, *Velloziaceae* familyaları kendi aralarında bir arada bir kümeleşme göstermemiştir. Ayrıca *Arecaceae* ve *Proteaceae* familyalarına ait bütün türlerde, *Anacardiaceae* familyasına ait *R. gueinzii* ve *R. transvaalensis* türlerinde, *Moraceae* familyasına ait *F. abutilifolia* türünde, *Orchidaceae* familyasına ait *C. warczewiczii* türünde polifiletik bir durum göstermiştir. Ağaç topoloji genel, olarak tür seviyesinde yirmibir tane monofiletik küme göstermiştir.

Şekil 4.19’da ise *accD* geninin analizi sonucu elde edilmiş olan ağaç topolojisi gösterilmiştir. Topolojide, çalışmada kullanılmış olan *Orchidaceae* dışında geride kalan ondört familya ve cinslerin tümü kendi aralarında uyumlu bir şekilde kümelenmiştir. Monokotil bitki grubunu oluşturan *Arecaceae* ve *Velloziaceae* %99 posterior probability desteği ile bir arada gösterilmişken *Orchidaceae* familyası ise ayrı bir kümeleşme göstermiştir. Ayrıca bu genin analizinin oluşturduğu topolojide toplam onyediy tür bir polifiletiklik durum göstermiştir. Ayrıca tür seviyesinde toplam onüç monofiletik küme göstermiştir.

Şekil 4.20’de *rpoC1* gen bölgesinin analizi sonucu elde edilen ağaç topolojisi verilmiştir. Söz konusu ağaç topolojisinde *Orchidaceae* dışında diğer familyalar tek bir küme şeklinde gösterilmiştir. Monokotil bitki grubunu oluşturan familyalar bir arada kümeleşmemiş bir şekilde farklı pozisyonlarda yer almışlardır. Topolojide yer alan toplam onsekiz tür polifiletik bir durum göstermiş olup geride kalan türler oniki monofiletik küme oluşturmuşlardır.

Şekil 4.21'de ise *trnH-psbA* gen bölgesinin analizi sonucu elde ağaç topolojisi verilmiştir. Bu topolojide de *Orchidaceae* familyası dışında diğer familyalar uyumlu kümeleşme göstermişlerdir. Monokotil bitki grubunu temsil eden familyaları bir arada gösterilmemiştir. *trnH-psbA* gen bölgesinin oluşturduğu bu topolojide, toplam beş türde polifiletiklik durumu görülmüşken; geride kalan türler toplam yirmibeş monofiletik küme olarak gösterilmiştir.

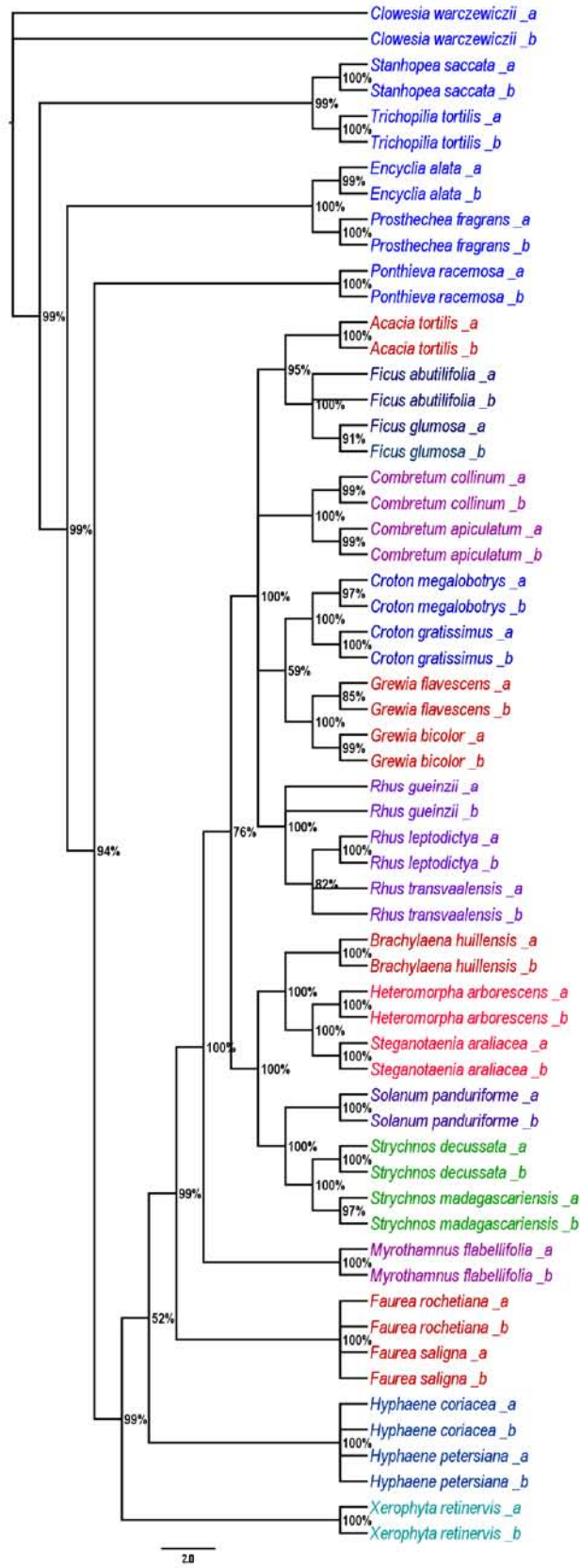
Şekil 4.22'de ise *rpoB* gen bölgesinin analizi sonucu elde edilmiş ağaç topolojisi verilmiştir. Bu topolojide de diğerlerinde olduğu gibi *Orchidaceae* familyası dışında diğer familyalar uygun bir şekilde kümeleşmişlerdir. Bu topolojide monokotil bitki grubunu temsil eden *Arecaceae* ve *Velloziaceae* %78 posterior probability desteği ile bir arada bulunurken; *Orchidaceae* familyası ise farklı bir pozisyonda yer almıştır. Ağaç topolojisi genel olarak on iki türde polifiletiklik bir durum sergilemişken; geride kalan türler ise toplam onsekiz ayrı monofiletik küme oluşturmuşlardır.

Sonuç olarak genel bir değerlendirme yapıldığında, Bayes metodu ile yapılan analizler sonucunda önerilecek uygun markır gen şüphesiz *trnH-psbA* gen bölgesi olmuştur. Çünkü bu gen bölgesinin analizi sonucu elde edilmiş ağaç topolojisi diğer gen bölgelerinin oluşturduğu topolojiler karşısında en az polifiletik tür grupları ve en fazla monofiletik tür kümeleşmesini göstermiştir. Ancak monokotil familyaları bir arada göstermemiştir. *trnH-psbA* takip ikinci uygun markır geni ise *matK* geni olarak önerilmiştir. Bu gen toplam sekiz türde bir polifiletiklik durum sergilemişken toplam olarak yirmibir tane monofiletik tür grubunu oluşturmuştur.

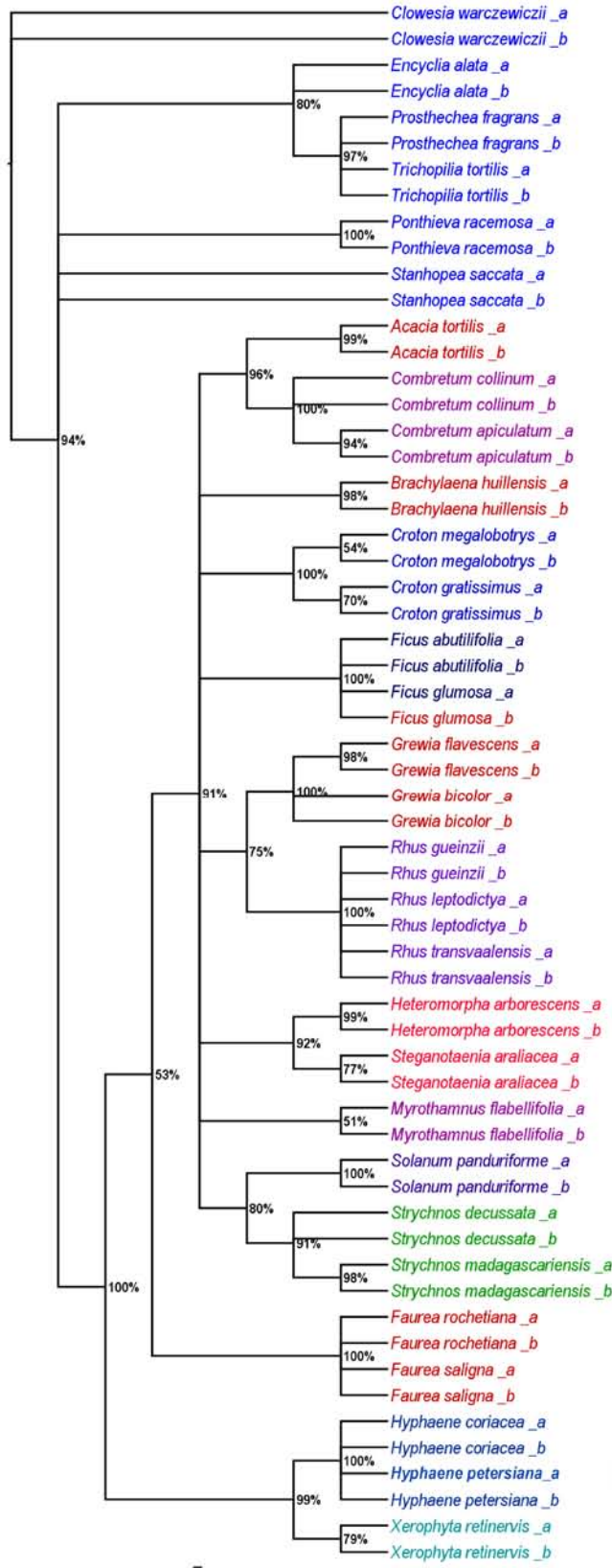
Çalışmada kullanılan verileri ve kullanılan metodlar temel alınarak en başarısız markır genler olarak onsekiz tane türde polifiletik durum göstren *rpoC1* geni ve onyediy türde polifiletik durum sergileyen *accD* geni olarak önerilmiştir. Bayes analizi sonucu elde edilen tüm topolojilerde *Orchidaceae* familyasına ait *C. warczewiczii* türü familyanın diğer türlerinden ayrı pozisyonda kalmıştır. Bu durumda Lahaye vd (2007) ifade edildiği üzere *Orchidaceae* familyasını oluşturan türleri taksonomik olarak belirlemek zordur ifadesini desteklemiştir.

Çalışmamızın temel amacı olan metot ve yöntem karşılaştırmaları çalışmanın bu bölümünde de ayrıntıları ortaya çıkarmıştır. Markır gen bulmak için yapılan ML ve Bayes filogenetik metodların ortaya koydukları sonuçlar yeterince çarpıcı olmuştur. ML analizlerin genelinde oluşan ağaç topolojileri daha fazla monofiletik küme oluşturmuştur. Aynı zamanda monokotil grubunu oluşturan familyaları aynı kümede vermiştir. Ayrıca önerdiği markır gen uygunluğunda da farklılık görülmüştür. Yani ML analizi sonucu uygun bir markır geni net olarak belirlenememişken Bayes analizler sonucunda ise büyük bir olasılıkla *trnH-psbA* geni markır gen olarak belirlenmiştir. Aynı zamanda Bayes analizi sonucu elde edilmiş olan ağaç topolojilerinde taksonlar daha fazla bir polifiletik durum görülmüştür ve monokotil bitki grubunu oluşturan familyaların hepsini birden bir kümede gösterememiştir. Ancak Bayes analizleri sonucunda elde edilen topolojilerin hiç birinde cins düzeyinde bir polifiletiklik durum gözlenmemiştir.

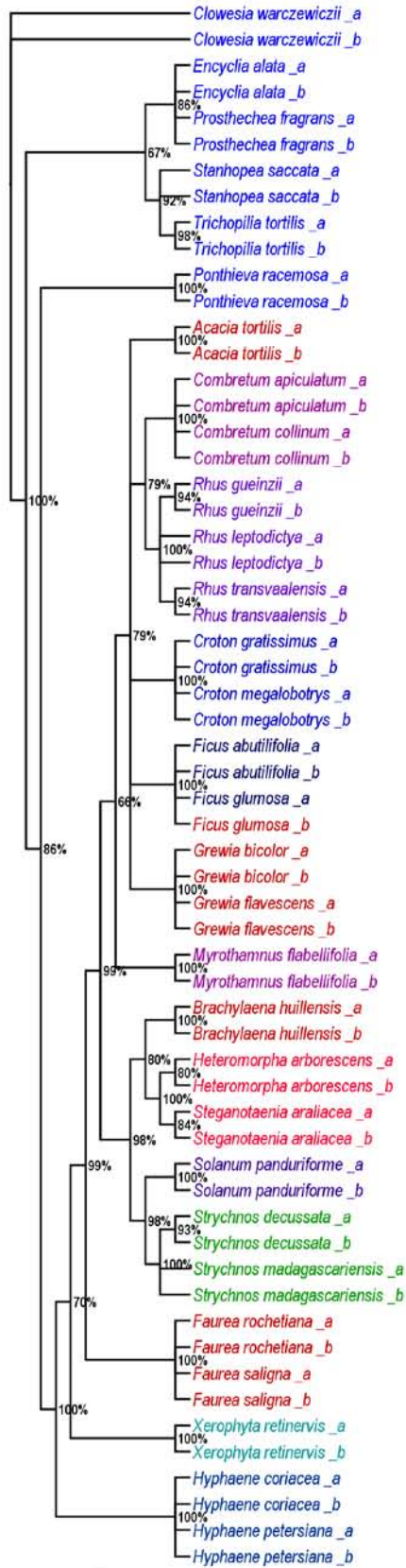
Görüldüğü gibi metot, yöntem ve genler arasındaki karşılaştırmalarda kayda değer farklılıklar ve veriler elde edilmiştir. Yalnız çalışmanın bu ayağı için istenilen amaca daha fazla yaklaşmak için şüphesiz, familya sayıları daha da çeşitlendirilmeli kardeş tür sayıları daha da artırılmalı, kullanılan genler çoğaltılmalı ve bu genler kendi aralarında birleştirilerek analiz edilmeli ve daha fazla filogenetik analiz metodları ve yöntemleri kullanılmalıdır.



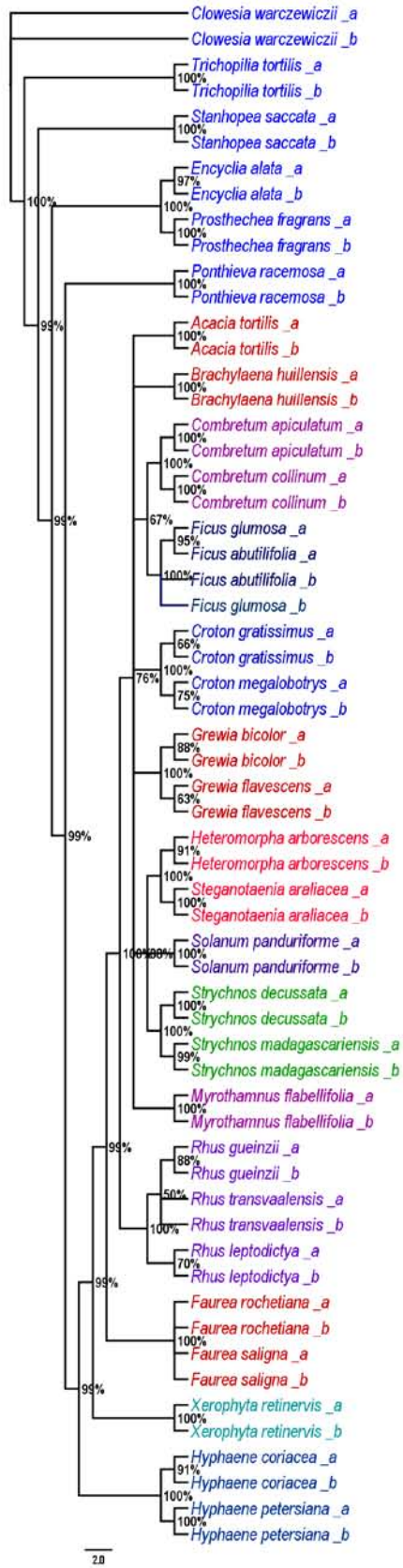
Şekil 4.18. MAFFT programının 1. parametresi ile hizalanmış *matK* gen sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)



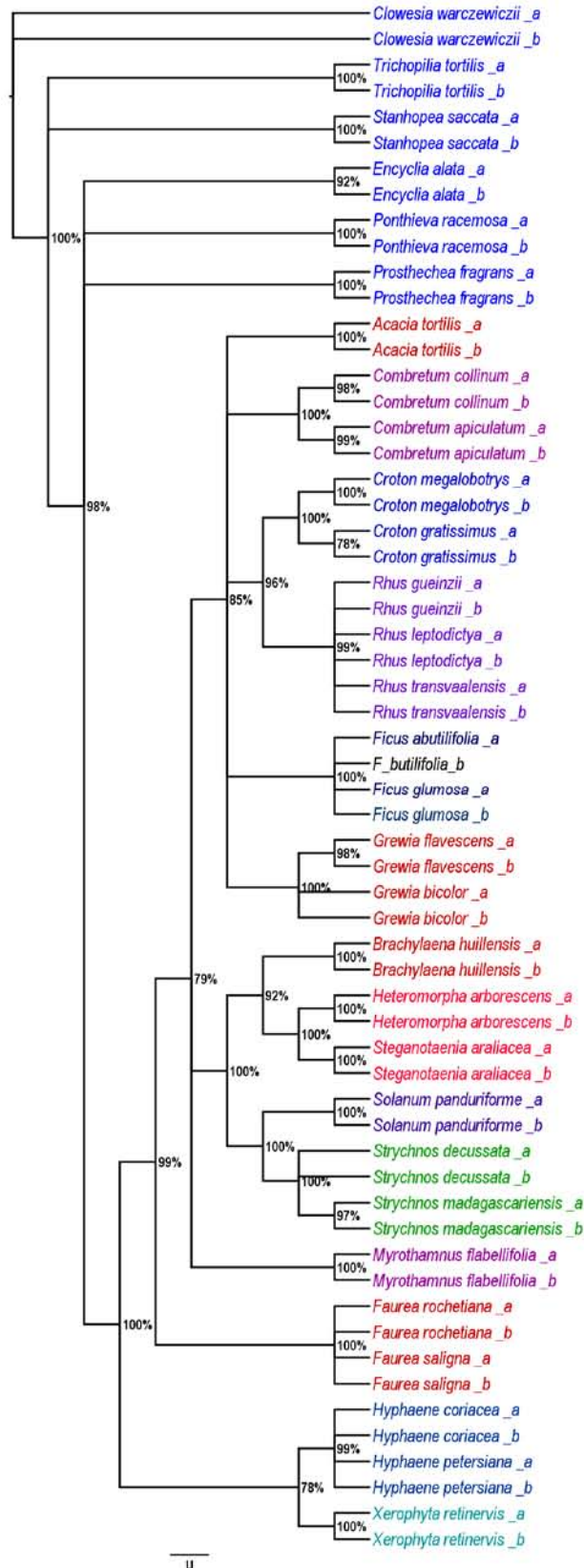
Şekil 4.19. MAFFT programının 1. parametresi ile hizalanmış *accD* gen sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)



Şekil 4.20. MAFFT programının 1. parametresi ile hizalanmış *rpoC1* gen sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)



Şekil 4.21. MAFFT programının 1. parametresi ile hizalanmış *trnH-psbA* gen sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)



Şekil 4.22. MAFFT programının 1. parametresi ile hizalanmış *rpoB* gen sekansının GTR SBT modelindeki Bayes analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)

5. SONUÇ

Taksonomi yeni taksonların tanımlanması, adlandırılması ve organizmaların uygun bir sınıflandırma sistemi içerisinde düzenlenmesi gibi bir dizi alanı kapsar (Simpson 1961). Sistematik, biyolojik çeşitliliği sınıflandırmakla yetinmez ve “neden” ve “nasıl” sorularını cevaplayabilmek için organizmaların evrimi, özellikle türleşme olgusunu teorik ve pratik yönleri ile ele alır. Bu yönüyle sistematik taksonomiye de içine alan bir alandır.

İlk defa Willi Hennig (1950) tarafından formüle edilen filogenetik sistematik taksonomiye filogenetik yönüyle inceleyen alandır. Bir takson ya da takson grubunun filogenilerinin belirlenmesi demek, zamansal olarak (önce-sonra) onların birbiri ile ve diğer taksonlarla ortak ata temelinde durumlarının ortaya konması demektir. Günümüzde filogeninin bitki taksonomi açısından zorunluluğu hemen hemen tüm sistematikçilerce kabul edilmektedir. Filogeni salt taksonomiye ilgilendirmemekte aralarında tarımsal üretiminde yer aldığı birçok alanda kullanım olanağı bulmaktadır. Bu nedenle özellikle fosil kaynaklarının bulunmadığı durumlarda moleküler filogeni filogenetik çalışmalarda önemlilik arz etmektedir.

Bu çalışmada bitki moleküler filogenetiğini konu alan çalışmalarda etkin olarak kullanılabilecek gen, gen bölgesi, gen sayısı, genin hücresel lokasyonu (kloroplast, nükleus), gen veya genler için hizalama yöntemleri, substitüsyon modelleri ve hizalama yöntemleri ile hizalanmış veriler yaygın bazı filogenetik analiz yöntemleri ile değerlendirilmesi gerçekleştirilmiştir.

Çalışmada gerek kullanılan bilgisayar donanım-yazılımları (Windows ve Linux) ve gerekse de süre göz önüne alınmış ve kullanılan alignment, filogeni, takson ve gen sayıları az tutulmak zorunda kalmıştır. Çalışmada sayısı günümüzde yüze yaklaşan alignment programlarından sadece ClustalW, MAFFT, MUSCLE, T-COFFEE programları ve programlarda bulunan algoritmalarından ikisi (parametreleri) kullanılmıştır.

Çalışma kapsamında kullanılan alignment programları arasında belirgin farklılıklar bulunmuş ve bu farklılıklar gerek analiz süresi ve gereksede oluşturulan alignmentler

arasındaki farklılıklar olarak ortaya çıkmıştır. T-COFFEE dışındaki programlar Windows tabanlı işletim sistemlerinde kullanılabilirken T-COFFEE programı Linux/Unix tabanlı işletim sistemlerinde kullanılabilir. Kullanılan alignment programlarının farklı algoritma içermeleri nedeni ile sadece ortak olan paramaterlerin seçimine özen gösterilmiştir.

Çalışmada elde edilen veriler üç farklı filogeni programı kullanılarak analiz edilmiştir. Bu programlar PAUP, MetaPIGA ve MrBayes programlarıdır. Filogenetik yöntemlerin karşılaştırılmasında PAUP yazılım programında uzunluk (Distance) ve MP analizleri gerçekleştirilmiş iken MetaPIGA ve MrBayes programlarında ML analizleri uygulanmıştır. Çalışmada kullanılan gen ve gen bölgeleri için uygun nükleotit modelleri MetaPIGA programına eklenmiş olan ModelTest programı yardımı ile belirlenmiştir. Çalışmada ModelTest yardımı ile GTR (General time reversible), K2P (Kimura's 2 parameter model), TN93 (Tamura-Nei 1993) yöntemlerinin kullanılan sekanslar için en uygun model olduğu saptanmıştır.

Çalışmada alignment programlarından MAFFT programının ITS dizisi kullanılarak diğer alignment yöntemlerinden daha iyi sonuç ürettiği tespit edilmiştir. ITS1 ve ITS2 dizilerinin ayrı ayrı, bir arada ve ITS1+5.8S+ITS2 dizilerinin hepsi bir arada kullanımının farklı sonuçlar ortaya koyduğu saptanmıştır. Özellikle ITS bölgelerinin arasında bulunan 5.8S rRNA kısmının gerek farklı nükleotit modeli kullanımını zorunlu kılması ve de analiz süresini çok uzatması bakımında kullanımının uygun olmadığı belirlenmiştir. ITS1 ve ITS2 dizilerinin birbirleri ile karşılaştırmaları yapıldığında uzunluk açısından ITS1 dizisinin daha uzun olduğu ve varyasyonun nükleotit değişimi yönünden ortaya koyduğu saptanırken ITS2 sekansının ITS1'e göre daha kısa olduğu ancak varyasyonun daha çok uzunluk açısından var olduğu belirlenmiştir.

ITS1 ve ITS2 bölgelerinin ayrı ayrı kullanıldığı analizlerde elde edilen sonuçlarla bir arada kullanıldığında elde edilen sonuçlar birbirlerinden farklı olduğu tespit edilmiştir. Bu farklılığın özellikle ITS1 bölgesinin TN93 modeline uygun iken ITS2'nin GTR modeline uygun olmasından kaynaklanmış olabileceği öngörülmüştür.

Çalışmada kullanılmış ikinci veri setinde filogeni için en uygun barkod genin (markır geninin) tespitine yönelik çalışmalar gerçekleştirilmiştir. Bu amaca ulaşmak için sekiz farklı hizalama algoritmasından en uygununu seçilmiş, ML ve BI metodları uygulanarak sonuca varılmıştır. Çalışmada kullanılan atmış bitki taksonu içinse bitki gruplarını yüksek bootstrap ve posterior probability değerleri nedeni ile *matK*, ve *trnH-psbA* ilgili bitki grubu için markır genler olarak belirlenmiştir.

Bundan sonraki çalışmalara ışık tutması amacıyla bu tezde elde edilen veriler doğrultusunda aşağıda listelenmiş öneriler elde edilmiştir. Bunlar araştırmacıların kullandığı gen sekanslarının güvenilirliği konusunda kesin bilgi sahibi olması (salt GenBank verilerine bağlı kalınılması), filogeni çalışmaları için kullanılacak takson sayısının mümkün olan en yüksek sayıya genişletilmesi (eksik takson son filogeni tahminini doğrudan etkilemesi nedeni ile), alingment (hizalama) yazılım programlarının daha da geliştirilmişleri (alignment programlarının sayısı her gün artmaktadır), bu programlarda kullanılacak parametre spektrumu oldukça geniş tutulması, model testlerin seçtiği substitüsyon modellerinin de sayısının artırılması (günümüzde toplam 64 farklı model bulunmaktadır), çalışmalarda kullanılan gen çeşitlerinin (gen bölgeleri, gen lokasyonları, gen sayıları) ve gen uzunluğunun yeterli düzeyde tutulması, model belirlemelerinde mümkün olan en yüksek doğruluğun tespitinin gerekli olduğu ortaya konulmuştur.

6. KAYNAKLAR

- ALLAN, G.J., FRANCISCO-ORTEGA, J., SANTOS-GUERRA, A., BOERNER, E., and ZIMMER, E.A., 2004. Molecular Phylogenetic Evidence for the Geographic Origin and Classification of Canary Island Lotus (Fabaceae: Loteae). *Molecular Phylogenetics and Evolution*, 32: 123-138.
- BAKKER, F.T., CULHAM, A., PANKHURST, C.E. and GIBBY, M. 2000. Mitochondrial and Chloroplast DNA-Based Phylogeny of *Pelargonium* (Geraniaceae) *American Journal of Botany* 87: 727-734.
- BALDWIN, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Molecular Phylogeny Evolution*, 1: 3-16.
- CAMIN, J. H. and SOKAL, R. R. 1965. A method for deducing branching sequences in phylogeny. *Evolution*, 19: 311-326.
- CAVALLI-SFORZA, L. L. and EDWARDS, A. W. F. 1967. Phylogenetic analysis: Models and estimation procedures. *American Journal Human Genetic*, 19: 233-257.
- CHASE, M.W., COWAN, R.S., HOLLINGSWORTH, P.M., VAN DEN BERG, C., MADRINAN, S., PETERSEN, G., SEBERG, O., JORGENSEN, T., CAMERON, K.M. and CARINE, M., 2007. A proposal for a standardised protocol to barcode all land plants. *Taxon* 56: 295-299.
- CLEGG, M.T., GAUT, B S. LEARN, G.H., J.R., and MORTON, B.R. 1994. Rates and patterns of chloroplast DNA evolution. *Proc. Nati. Acad. Sci. USA*, Vol. 91, pp. 6795-6801.
- COHEN, B.L. and WEYDMANN, A., 2005. Molecular Evidence That Phoronids Are a Subtaxon of Brachiopods (Brachiopoda: Phoronata) and That Genetic Divergence of Metazoan Phyla Began Long before the Early Cambrian. *Organisms Diversity & Evolution*, 5: 253:273.
- DRUMMOND, A.J. and RAMBAUT, A. 2007. BEAST: Bayesian evolutionary analyses by sampling trees. *BMC Evolutionary Biology* 7: 214-221.
- DUANGJAI, S., WALLNO, B., SAMUEL, R., MUNZINGER, J. and CHASE, M.W. 2006. Generic Delimitation and Relationships in Ebenaceae Sensu Lato: Evidence from six plastid DNA regions. *American Journal of Botany*, 93: 1808-1827
- DURBIN, R. EDDY, S. KROGH, A. and MITCHISON, G. 1998. Biological sequence analysis: probabilistic models of proteins and nucleic acids, Cambridge University Press,

- EDGAR, R.C. 2004 MUSCLE Low-complexity multiple sequence alignment with T-Coffee accuracy [Short Paper]. *ISMB/ECCB*, pp: 1210-1230.
- FELSENSTEIN, J. 1989. PHYLIP – Phylogeny Inference Package (Version3.2). *Cladistics*, 5:164-166.
- FITCH, W. M. and SMITH, T. F. 1983. Optimal sequence alignments. *National Academy of Sciences of the United States of America*, 80: 1382–1386.
- GOTOH, O 2003. Multiple sequence alignment: Algorithms and applications Saitama Cancer Center Research Institute, Ina-machi, Saitama 362-0806.
- GRAHAM, S.W. and OLMSTEAD, R.G., Systematics 2000. Utility of 17 Chloroplast Genes for Inferring the Phylogeny of the Basal Angiosperms. *American journal of Botany*. 87: 1712-1730.
- HELAERS, R. , MILINKOVITCH, M.C. 2010. MetaPIGA v2.0: maximum likelihood large phylogeny estimation using the metapopulationgenetic algorithm and other stochastic heuristics *BMC Bioinformatics* 11: 379-390.
- HENNIG, W. 1966. Phylogenetic Systematics. Univ. Illinois Press, Urbana. Hillis, D. M. 1995. Approaches for assessing phylogenetic accuracy. *Systematic Biology*, 44: 3-16.
- HILU, K.W. 1994. Evidence from RAPD markers in the evolution of *Echinochloa* Millets (Poaceae). *Plant Systematics and Evolution* 189: 247–257.
- HILU, K.W. and LIANG, H. 1997. The matK gene sequence variation and application in plant systematics. *American Journal of Botany* 84: 830–839.
- HILLIS, D.M. MORITZ, C. and MABLE, B.K. 1996. *Molecular Systematics*, 2nd Ed. Sinauer Associates, Inc. Sunderland, MA.
- HUELSENBECK, J.P.F., RONQUIST, R., NIELSEN, and BOLLBACK J.P. 2001. Bayesian inference of phlogeny and its impact on evolutionary biology. *Science* 294:2310-2314.
- HUGHES, C.E., EASTWOOD, R.J. and BAILEY, C.D. 2006. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction *Philosophical Transactions. R. Soc. B*, 361: 211-225.
- HUSON, D.H. 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* 14: 68-73.
- INCE, A.G., KARACA, M., ONUS, A.N. 2010. A reliable gender diagnostic PCR assay for jojoba(*Simmondsia chinensis* (Link) Schneider) *Genet Resour Crop Evol* 57:773–779.

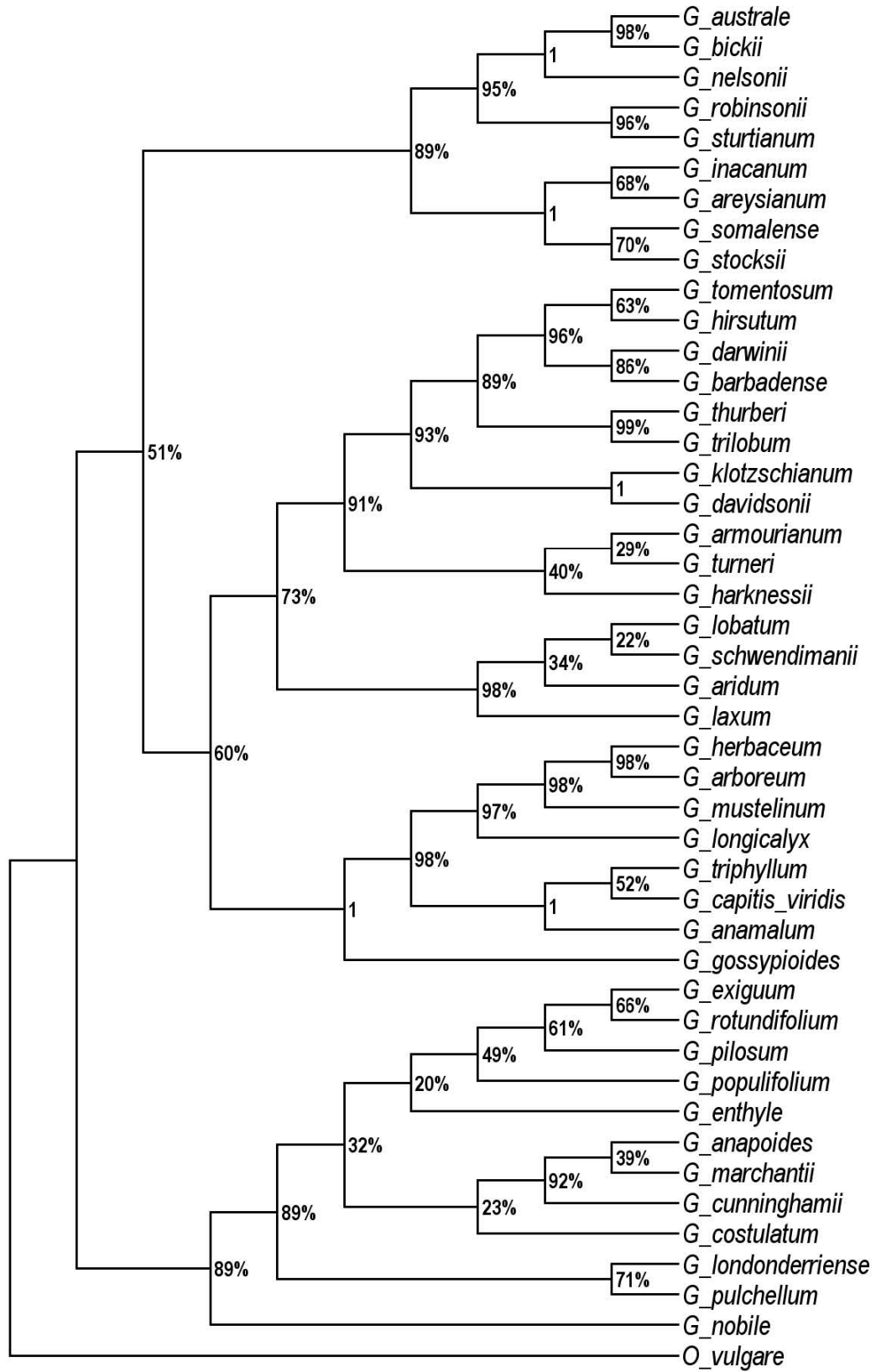
- İNCE, A.G. 2010. Doku/Organ Spesifik Mikrosatellit DNA Gen İçeriklerinin *Capsicum* cDNA Kütüphanelerinde *In Silico* ve *In Vitro* Yaklaşımlarla Belirlenmesi. Doktora tezi, Akdeniz Üniversitesi, Fen Bilimleri Enstitüsü, 300ss.
- JOHNSON, L.A., and SOLTIS, D.E. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany* 19: 143–156.
- KABAOĞLU, A. 2007. Türkiyede bulunan *Hypogymnia* (Nyl.) Nyl. türlerinde bulunan rDNA ITS bölgesi DNA dizi analizi ile çeşitliliğin tanımlanması. Yüksek lisans tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, ss: 86.
- KARACA, M., INCE, A. G., ELMASULU, S. Y., ONUS, A. N. & TURGUT, K. 2005. Coisolation of genomic and organelle DNAs from 15 genera and 31 species of plants. *Analytical Biochemistry*, 343: 353-355.
- KISHINO, H. and HASEGAWA, M. 1989. Evaluation of the maximum likelihood model estimates of the evolutionary tree topologies from sequence data, and the branching order in Homonoidea. *Journal of Molecular Evolution*. 29: 170-179.
- KRESS, W.J. and ERICKSON D.L. 2007. A Two-Locus Global DNA Barcode for Land Plants: The Coding *rbcL* Gene Complements the Non-Coding *trnH-psbA* Spacer Region. *PLoS One*, 44: 18621–18626
- LAHAYE, R., BANK, M.V.D., BOGARIN, D., WARNER, J., PUPULIN, F., GIGOT, G., MAURIN, O., DUTHOIT, S., BARRACLOUGH, T.G. and SAVOLAINEN, V., 2007. DNA barcoding the floras of biodiversity hotspots. *Pnas*, 8:2923-2928
- LAMARCK, J. B. 1809. Zoological Philosophy: An exposition with regard to the natural history of animals. Translated by H. Elliot (Macmillan, London 1914. Reprinted by University of Chicago Press, 1984).
- LISTON, A., ROBINSON, W.A., PINERO, D. and ALVAREZ-BUYLLA, E.R. 1999. Phylogenetics of Pinus (Pinaceae) Based on Nuclear Ribosomal DNA Internal Transcribed Spacer Region Sequences. *Molecular Phylogenetics and Evolution*, 11: 95-109.
- LUNTER, G., MIKLOS I., DRUMMOND, A. JENSEN J.L. and HEIN, J. 2005. Bayesian coestimation of phylogeny and sequence alignment *BMC Bioinformatics*, 6: 83-93.
- NEI, M. and KUMAR, S. 2000. Molecular Evolution and Phylogenetics. Oxford.
- NEUHAUS, H. and LINK, G. 1987. The chloroplast tRNA^{LYS}(UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. *Current Genetics* 11: 251–257.

- NOTREDAME, C. and HIGGINS, D.G. and HERINGA, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal Molecular Biology*, 302: 205-217.
- OGDEN, T.H., and ROSENBERG, M.S. 2006. Multiple Sequence Alignment Accuracy and Phylogenetic Inference *Systematic Biology*. 55: 314-32.
- POSADA, D. and KRANDAL, K.A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatic*, 14: 817-818.
- QUI, Y.L., LEE, J.H., BERNASCONI-QUADRONI, F., SOLTIS, D.E., SOLTIS, P.S., ZANIS, M., ZIMMER, E.A., CHEN, Z.D., SAVOLAINEN, V. and CHASE, M.W. 1999. The Earliest Angiosperms, Evidence from Mitochondrial, Plastid and Nuclear Genoms. *Nature*, 25: 402-404.
- RAMBOUT, A. and DRUMMOND, A.J. 2003. Tracer v1.3. Available form: <http://evolve.zoo.ox.ac.uk/>.
- RO, K.E., KEENER, C.S. and MCPHERON, B.A. 1997 Molecular Phylogenetic Study of the Ranunculaceae: Utility of the Nuclear 26s Ribosomal DNA in Inferring Intrafamilial Relationships. *Molecular Phylogenetics and Evolution*, 8: 117-127.
- SIMPSON, G.G., 1961. Principles of Animal Taxonomy, New York., Columbia University Press
- SAITO, N. and NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular. Biology. Evolution*. 4: 406-25.
- SCHADT, E.E., SINSHEIME, J.S. and LANGE, K. 1998. Computational Advances in Maximum Likelihood Methods for Molecular? Phylogeny *Genome Res*. 88: 222-233.
- SCHMITT, I. and F. KEITH, B. 2009. Phylogenetic methods in natural product research *Nat. Prod. Rep.*, 26: 1585–1602.
- SHIMODAIRA, H. and HASEGAWA M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, 16: 1114–1116.
- SMITH, T.F., WATERMAN, M.S. 1981. Identification of common molecular subsequences. *Journal of Molecular Biology*, 147: 195-97
- SNEATH, P. H. A. and SOKAL, R. R. 1973. Numerical taxonomy. Freeman, San Francisco.

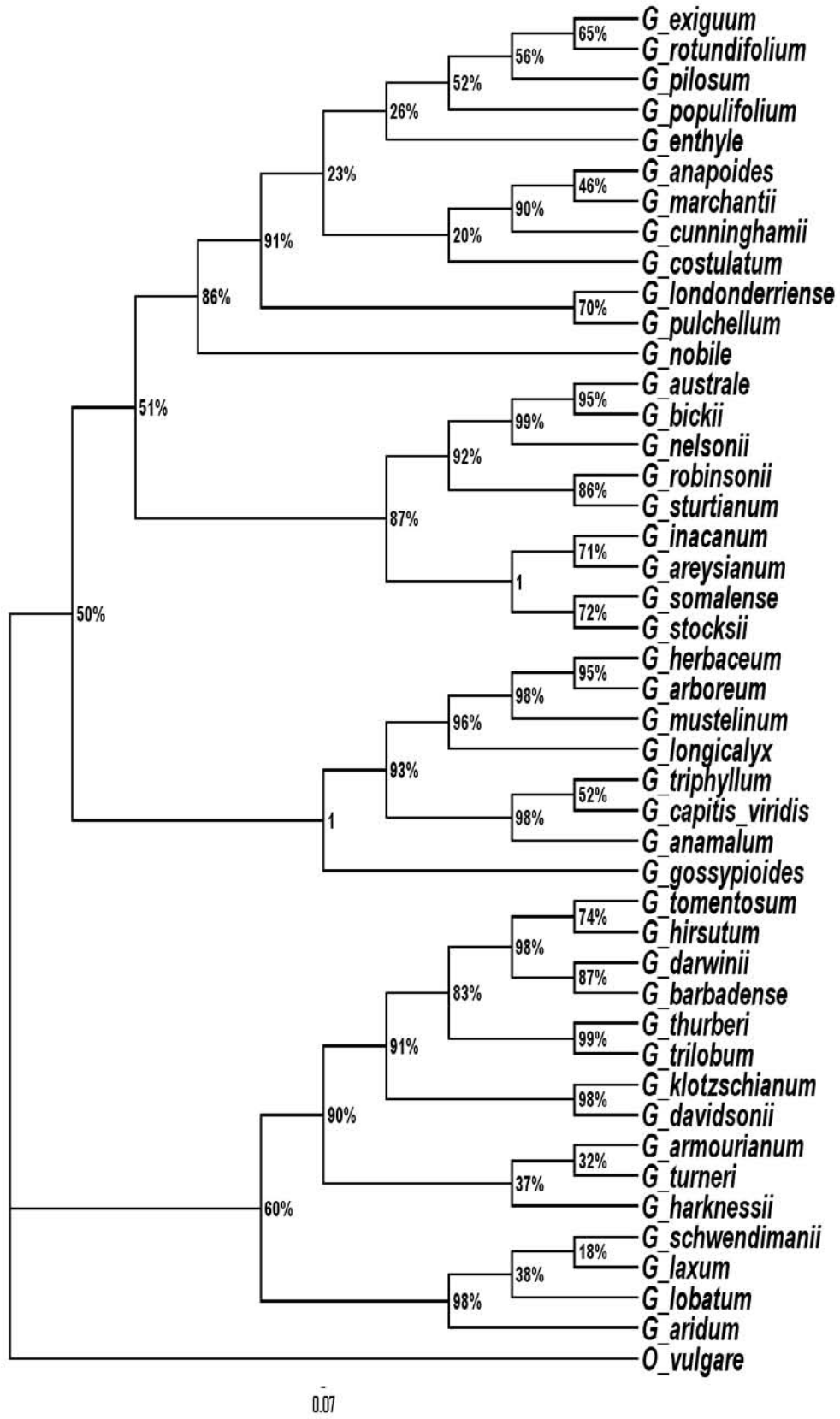
- STAMATAKIS, A., LUDWIG, T. , MEIER, H.: RAXML, A. 2005. Fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* 21: 456–463.
- SUAREZ-DIAZ, E. and ANAYA-MUÑOZ, V.H. 2008. History, objectivity, and the construction of molecular phylogenies. *Stud. Hist. Phil. Biol. & Biomed. Sci.* 39: 451–468.
- SWOFFORD, D.L. 2002. PAUP* Pylogenetic Analysis Using Parsimony (* and other methods). v. 4.0 beta. Sunderland, M. A: Sinauer Associates.
- TAMURA, K., DUDLEY, J. NEI, M., KUMAR, S. 2007. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24: 1596-1599
- TAYLOR, W.R., MUNRO, R.E. 1997. Multiple sequence threading: conditional gap placement. *Fold Des* 2: 33-39.
- THOMAS, D.C., 2010. Phylogenetics and historical biogeography of Southeast Asian Begonia L.(Begoniaceae) Ph.D. Thesis 192p University of Glasgow.
- THOMPSON, J.D. and HIGGINS, D.G. and GIBSON, T.J: CLUSTAL W 1994. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
- WEN, J., VANEK-KREBITZ, M., HOFFMANN-SOMMERGRUBER, K., SCHEINER, O. and BREITENEDER, H. 1997. The Potential Of betv1 homologues, a Nuclear Multigene Family, as Phylogenetic Markers in Flowering Plants. *Molecular Phylogenetics and Evolution*, 8: 317-333.
- WHEELIS, M. 2007. Darwing: not the first to sketch a tree. *Science* 315: 597. 4680.
- YOKOYAMA, J., SUZUKI, M., IWATSUKI, K., and HASEBE, M. 2000. Molecular Phylogeny of Coriaria, with Special Emphasis on the Disjunct Distribution. *Molecular Phylogenetics and Evolution*, 14: 11-19.
- YUE, F., SHI, J. and TANG, J. 2009. Simultaneous phylogeny reconstruction and multiple sequence alignment. *BMC Bioinformatics* 11: 1471-2105.
- WENDEL, J.F., BRUBAKE, C. L. and SEELANAN, T. 2010 The Origin and Evolution of *Gossypium*. *Physiology of Cotton*, 10: 978-990.
- ZHAO, X. -P., SI, Y., HANSON, R. E., CRANE, C. F., PRICE, H. J., STELLY, D. M., WENDEL, J. F., AND PATERSON, A. H. 1998. Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Res.* 8: 479–492.

7. EKLER

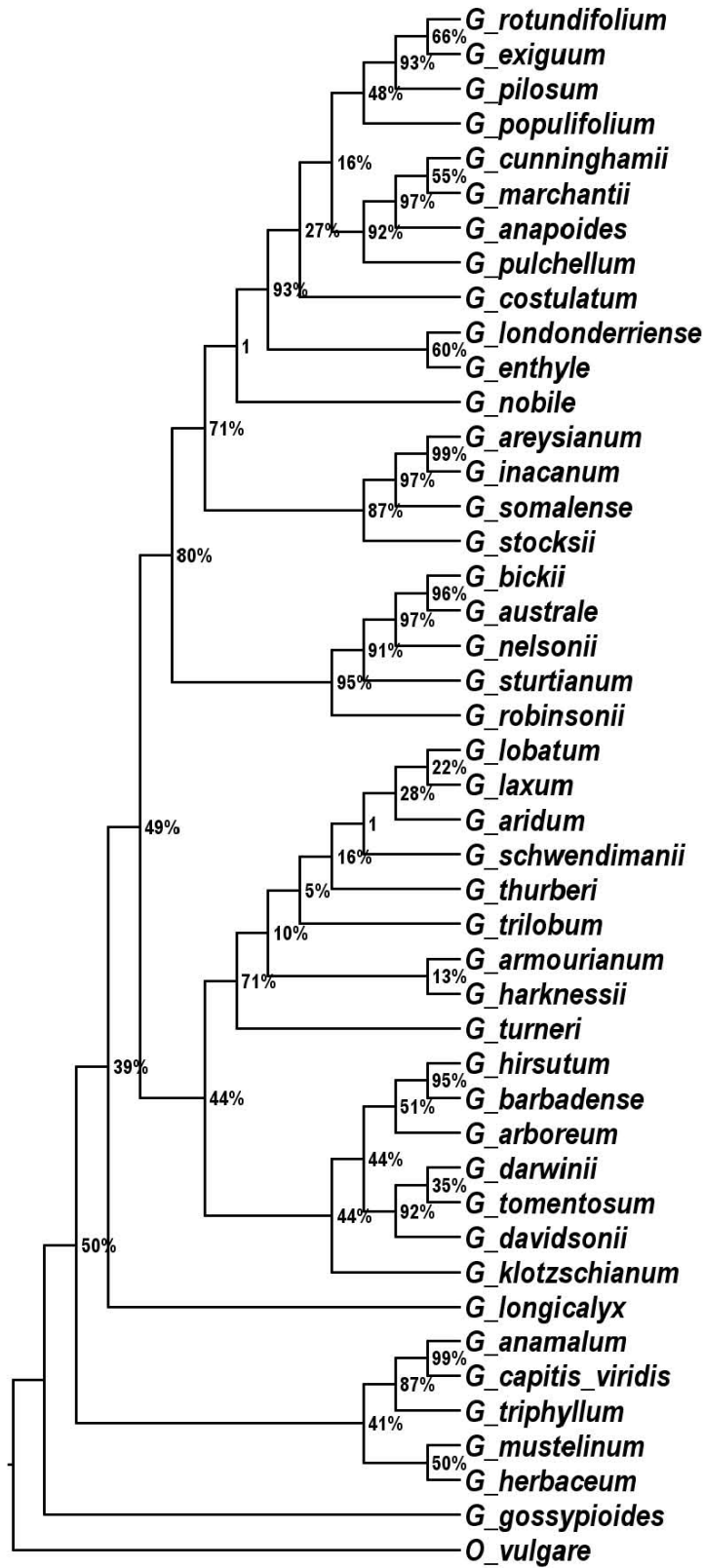
Çalışmada kullanılan farklı alignment yöntemleri ve alignment analiz parametreleri, substitüsyon modellerinden, gen bölgelerinden ve farklı filogeni yöntemleri ile programlarından elde edilen filogeniler bu bölümde sunulmuştur. Bu bölümde verilen filogeni ağaçlarının nodlarda bulunan yüzde değerler ML, MP ve D analizleri için bootstrap değerleri iken BI filogenilerindeki değerler ise posterior probability değerleridir.



Şekil Ek-1. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi

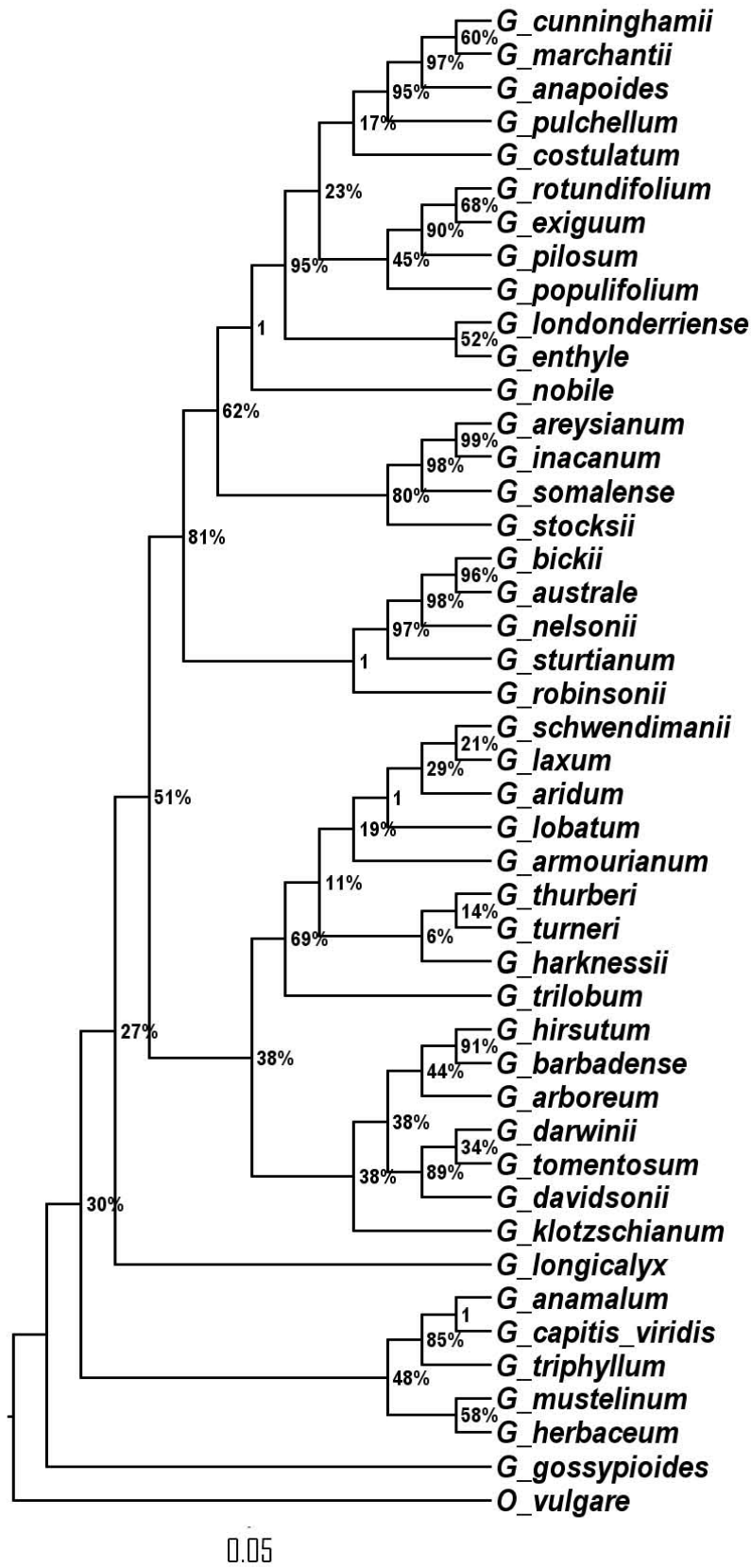


Şekil Ek-2. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi

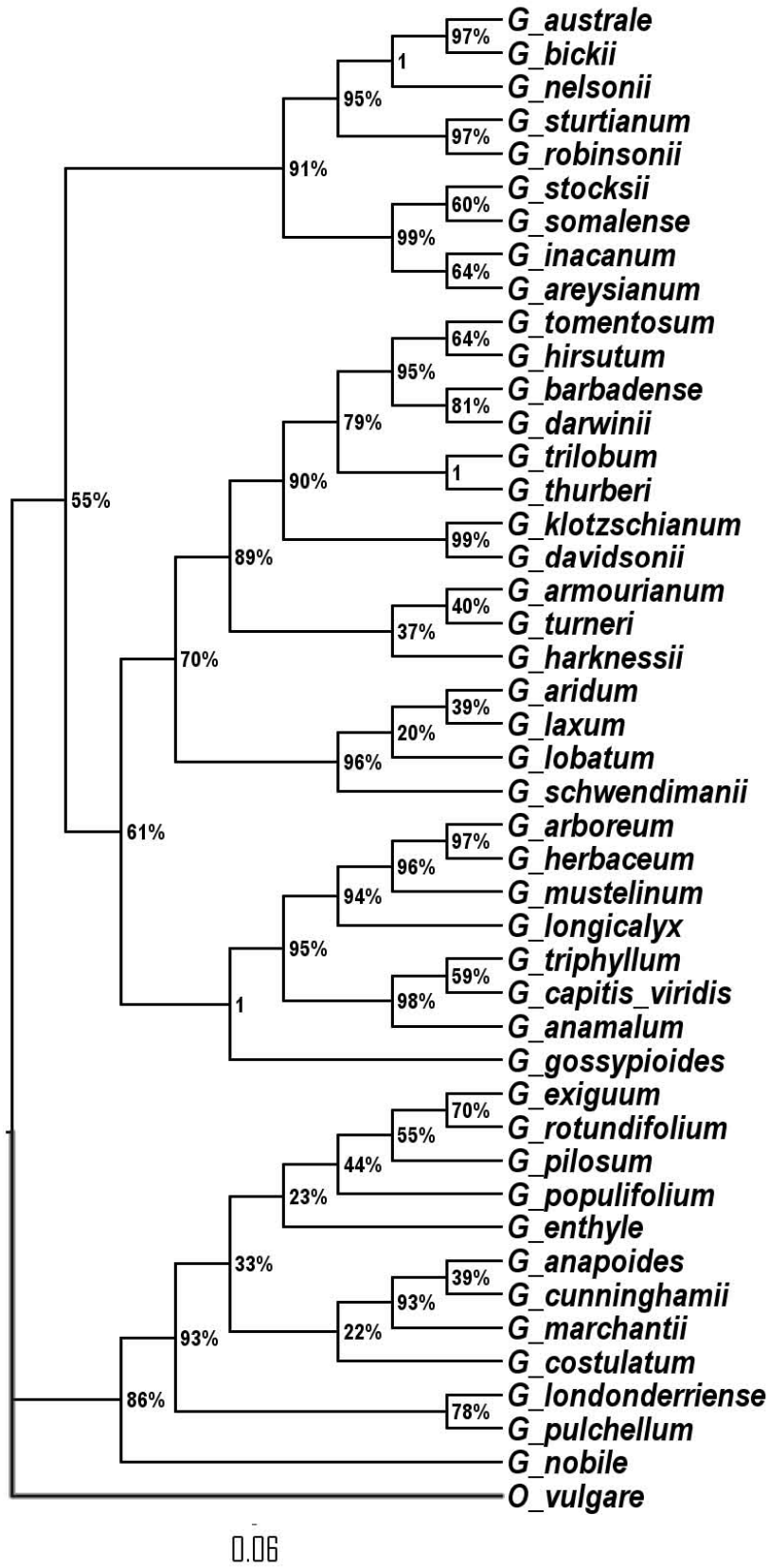


0.05

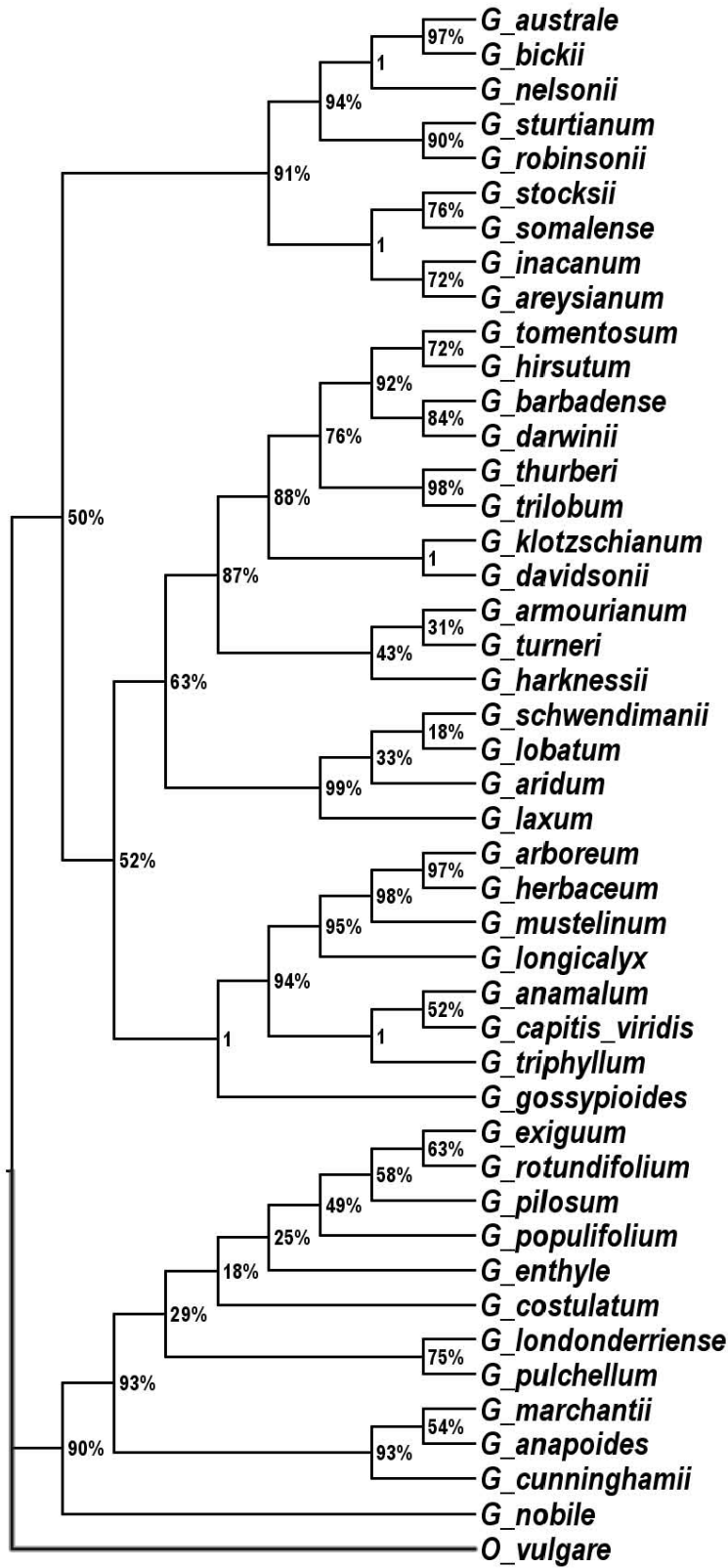
Şekil Ek-3. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi



Şekil Ek-4. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi

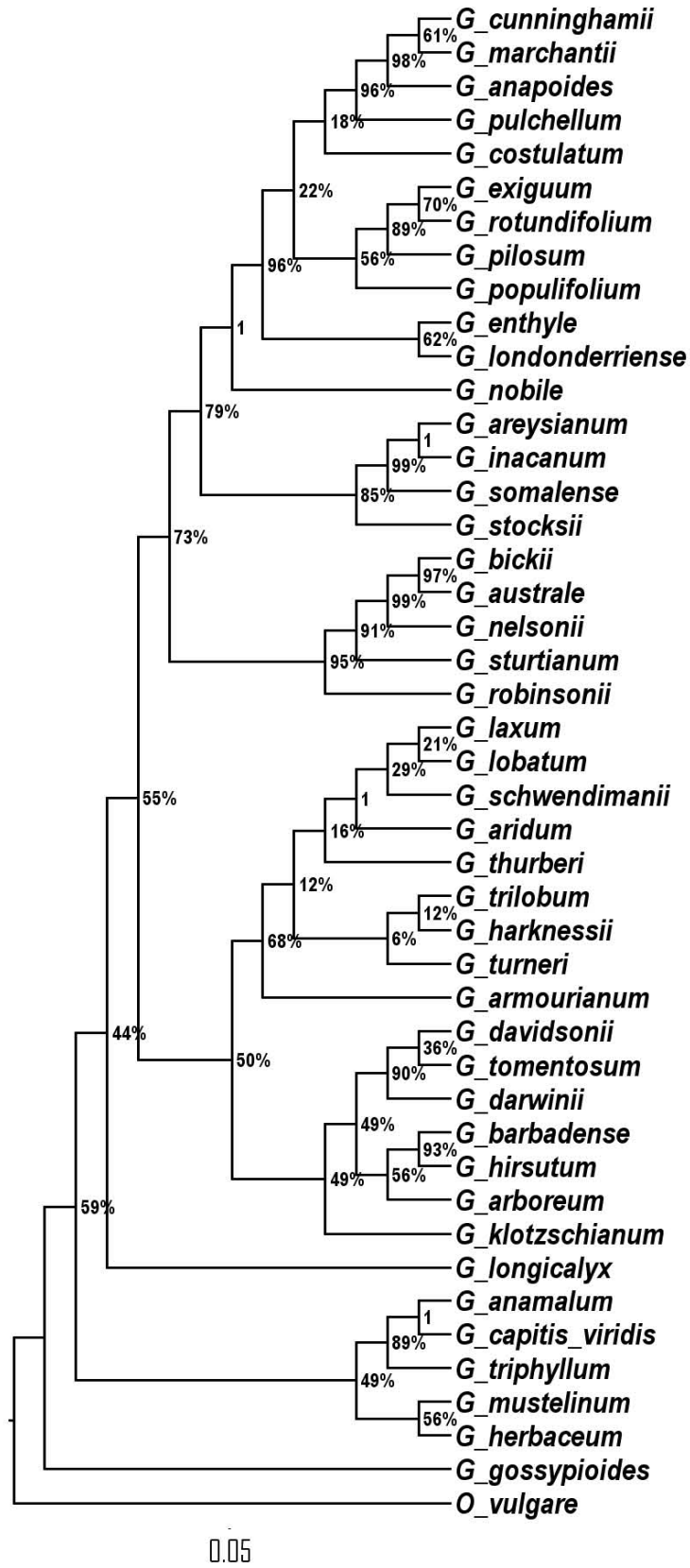


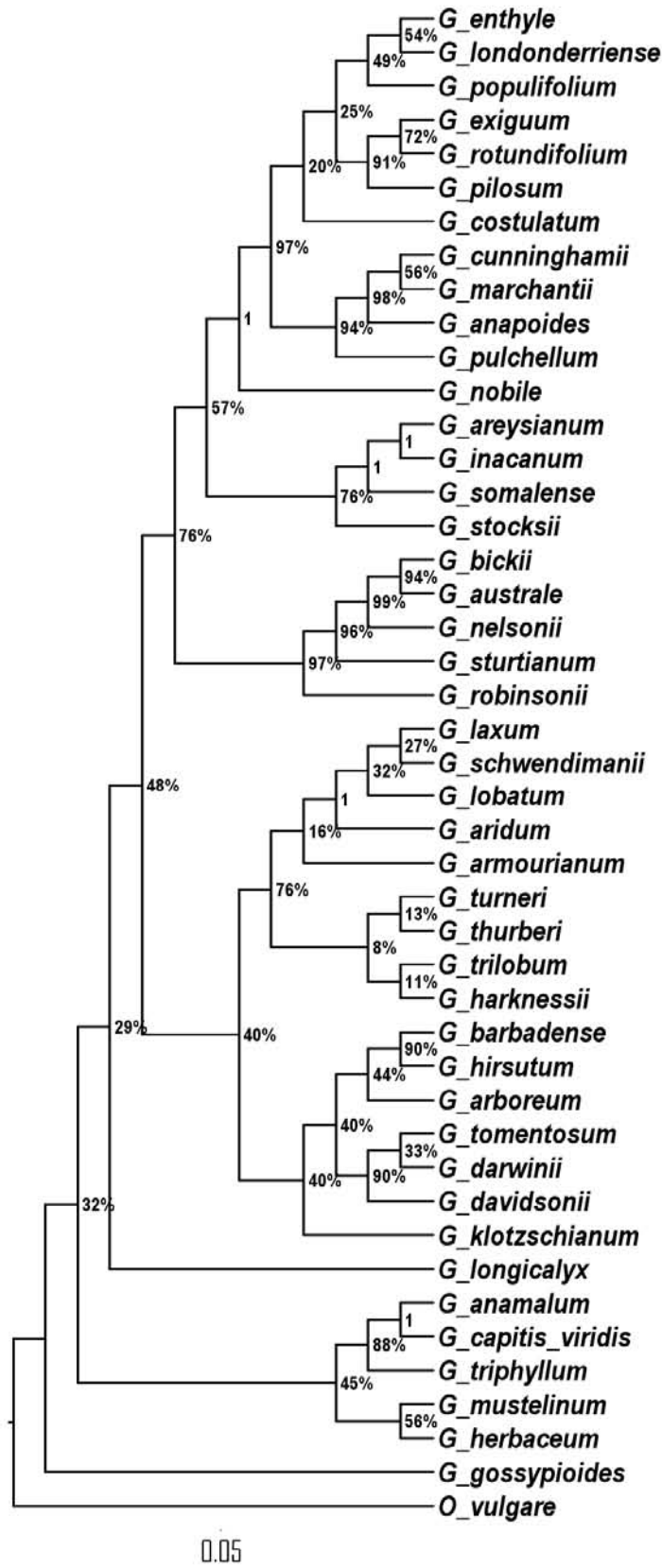
Şekil Ek-5. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi



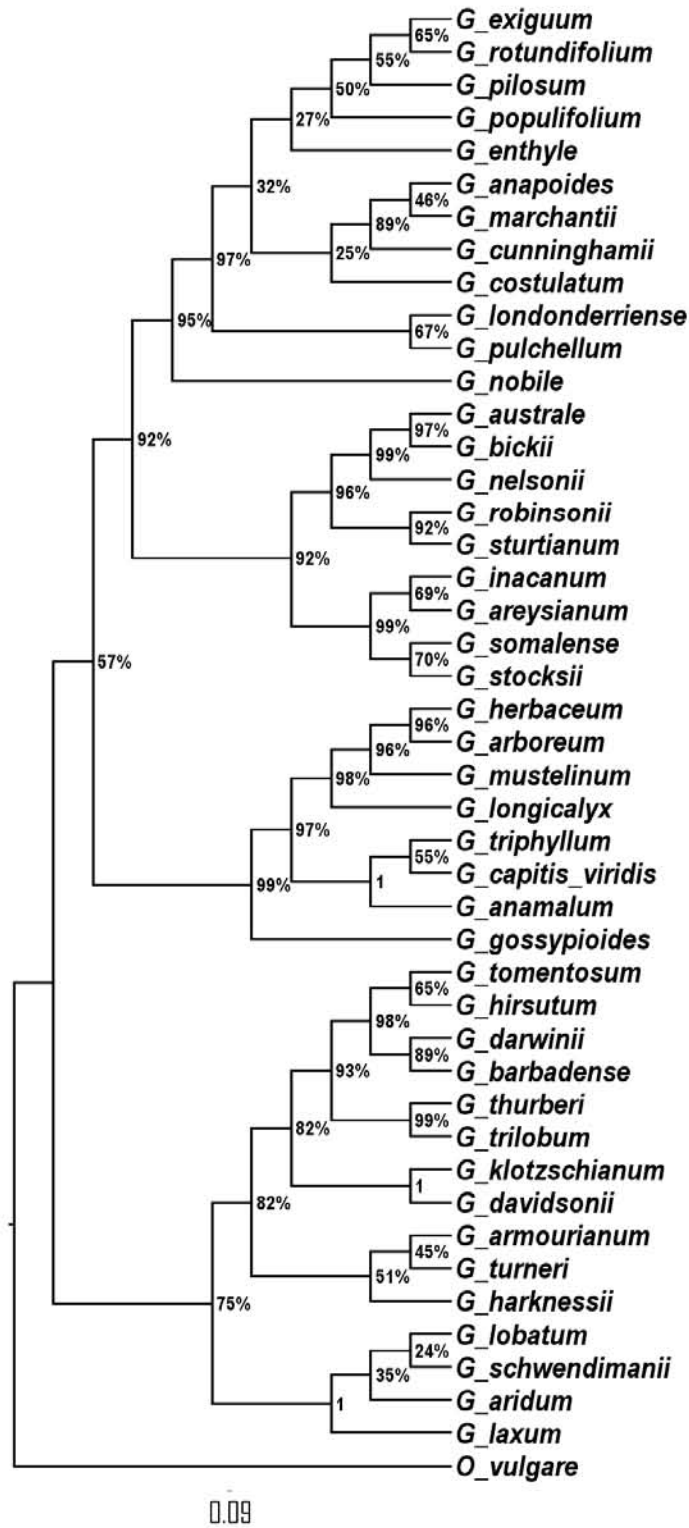
0.07

Şekil Ek-6. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi

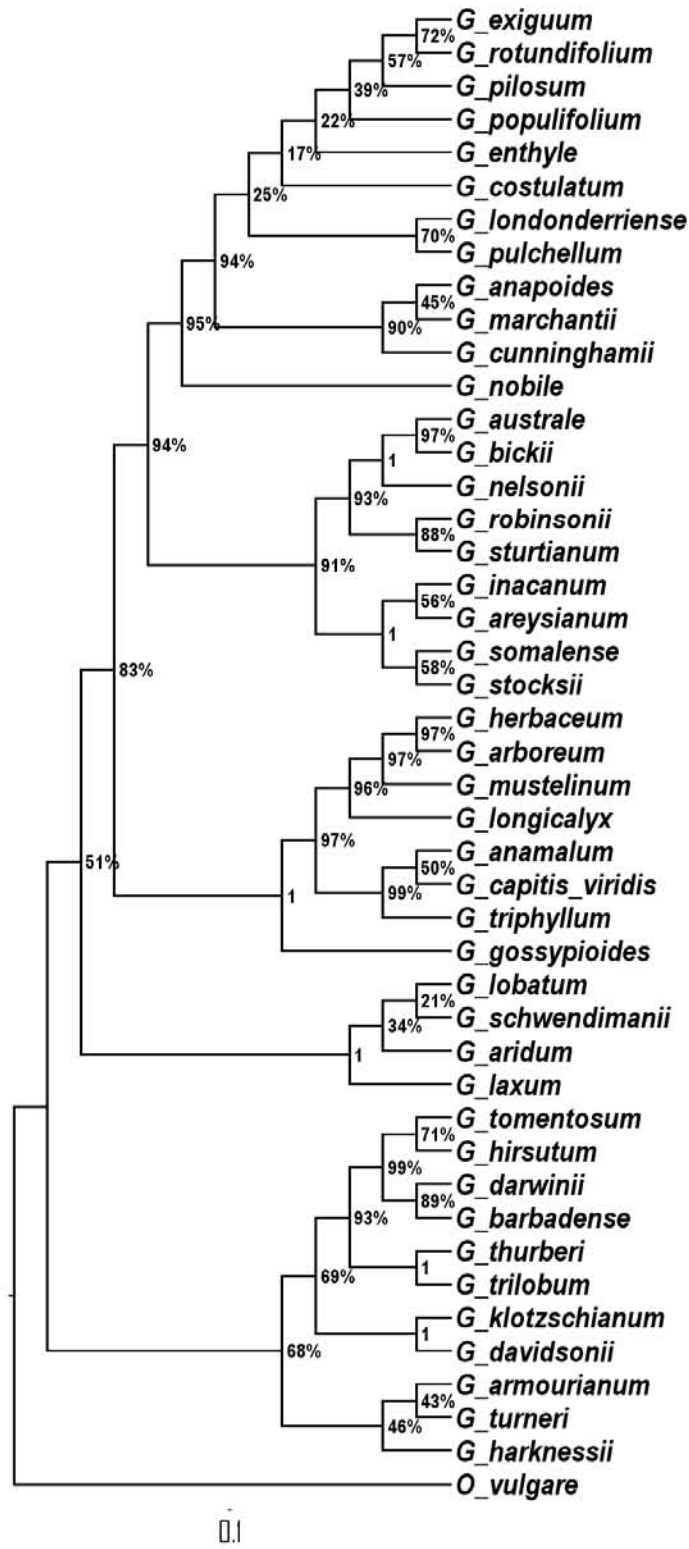




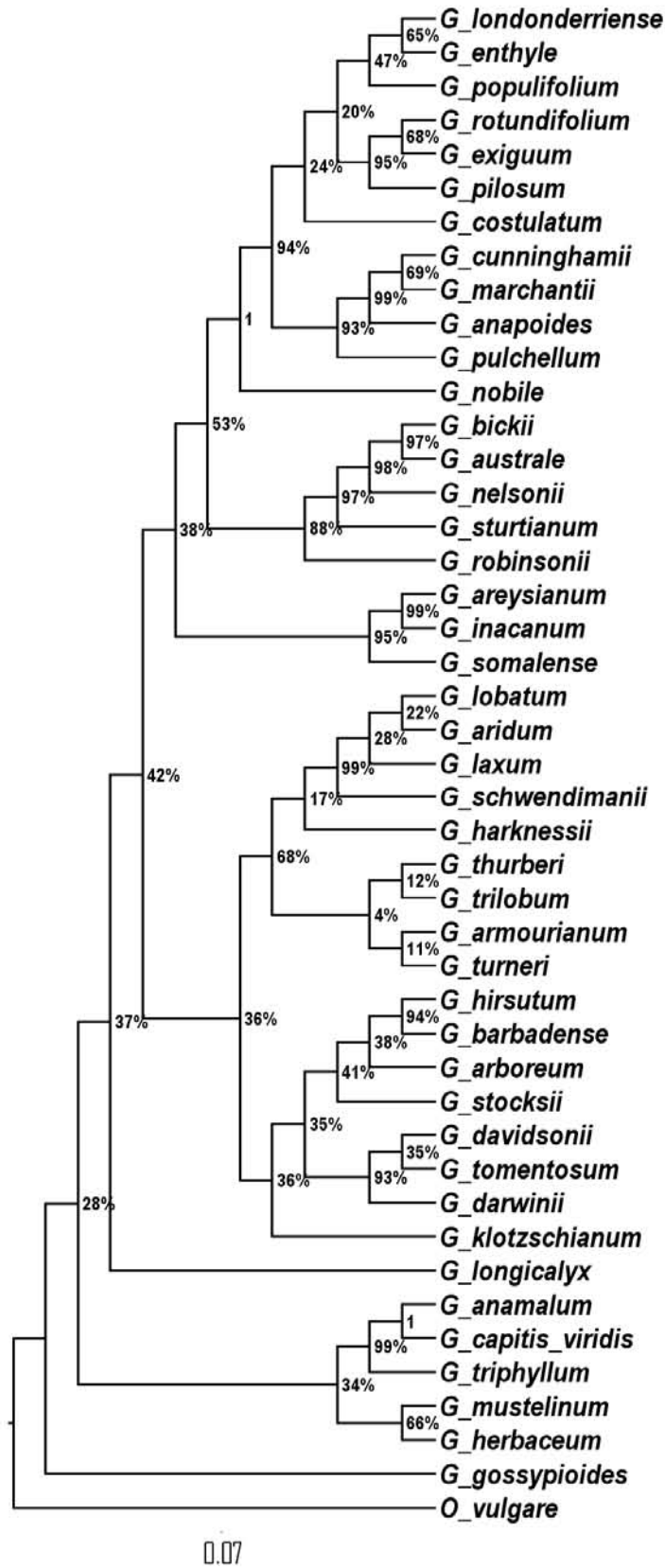
Şekil Ek-8. ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



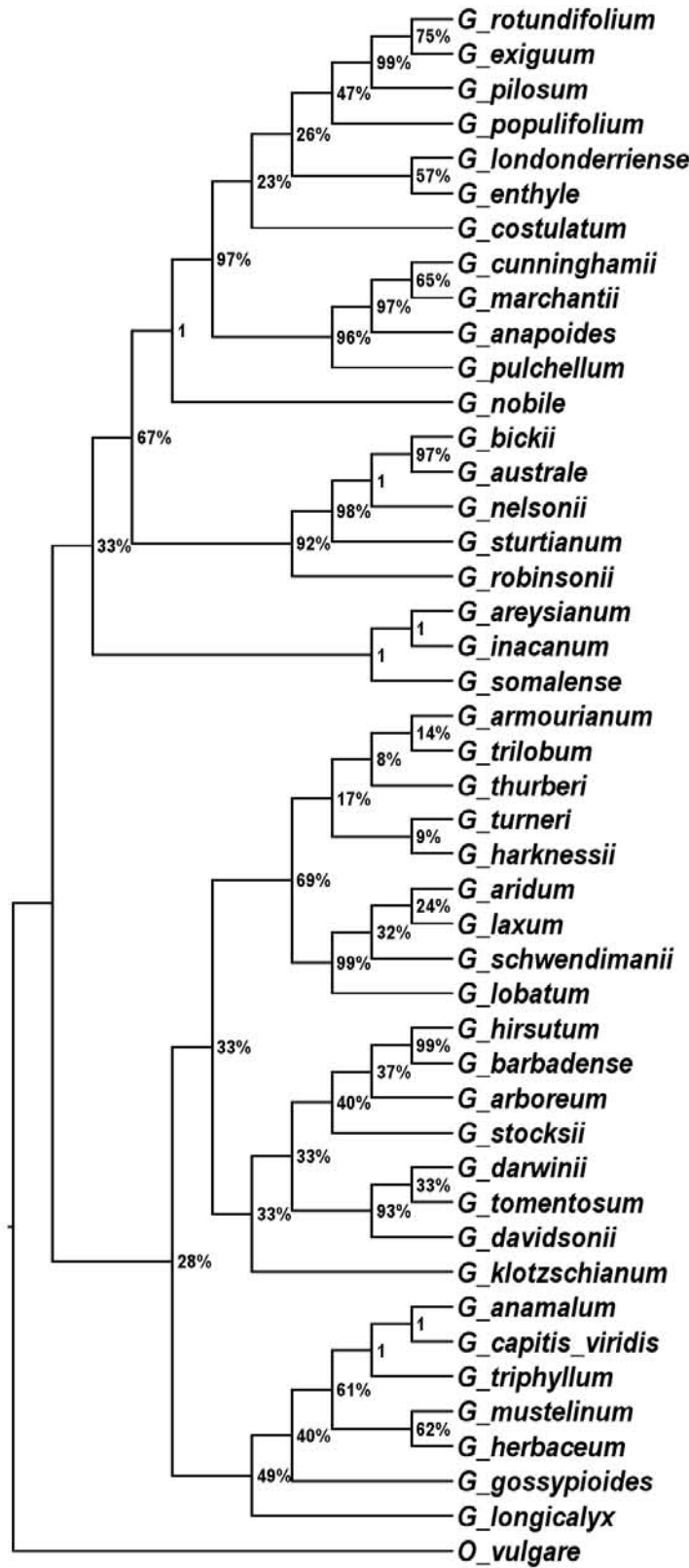
Şekil Ek-9. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi



Şekil Ek-10. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi (Nodlarda bulunan yüzde değerler posterior probability değerleridir)

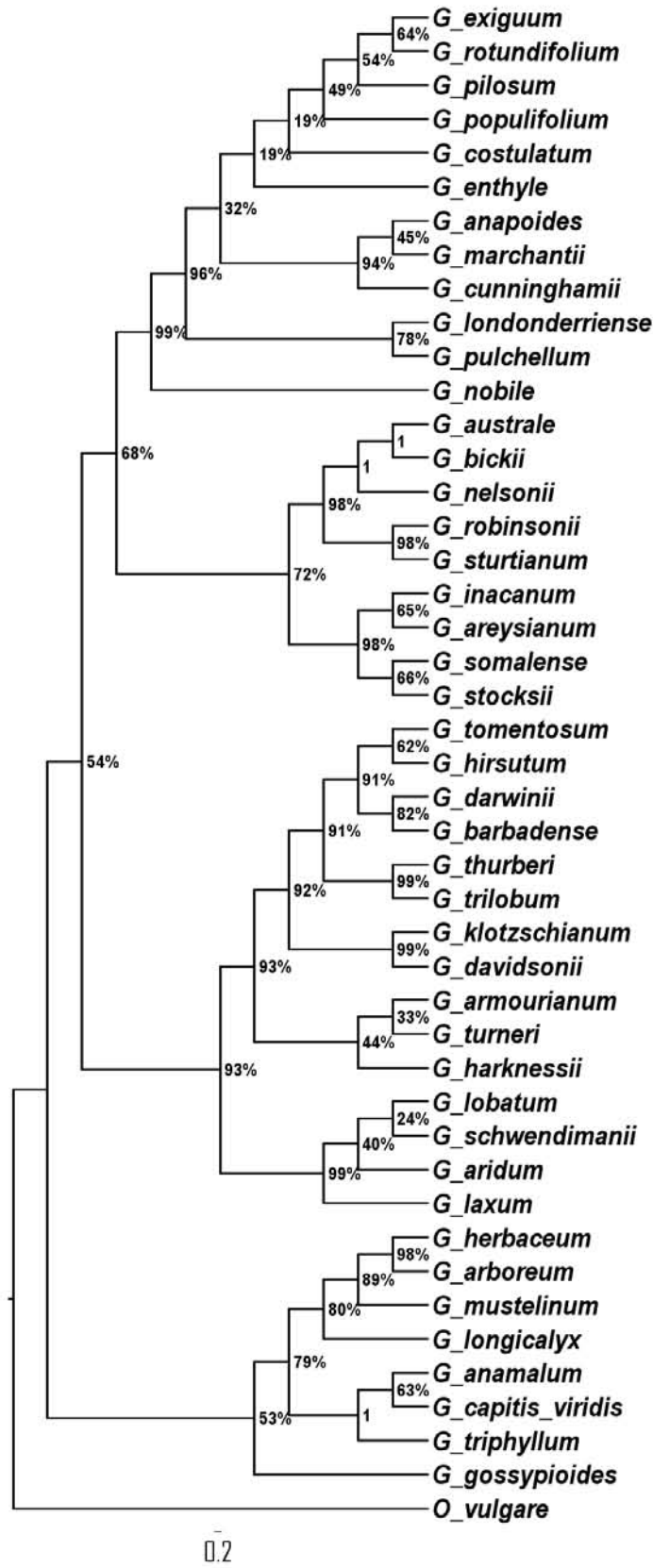


Şekil Ek-11. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi

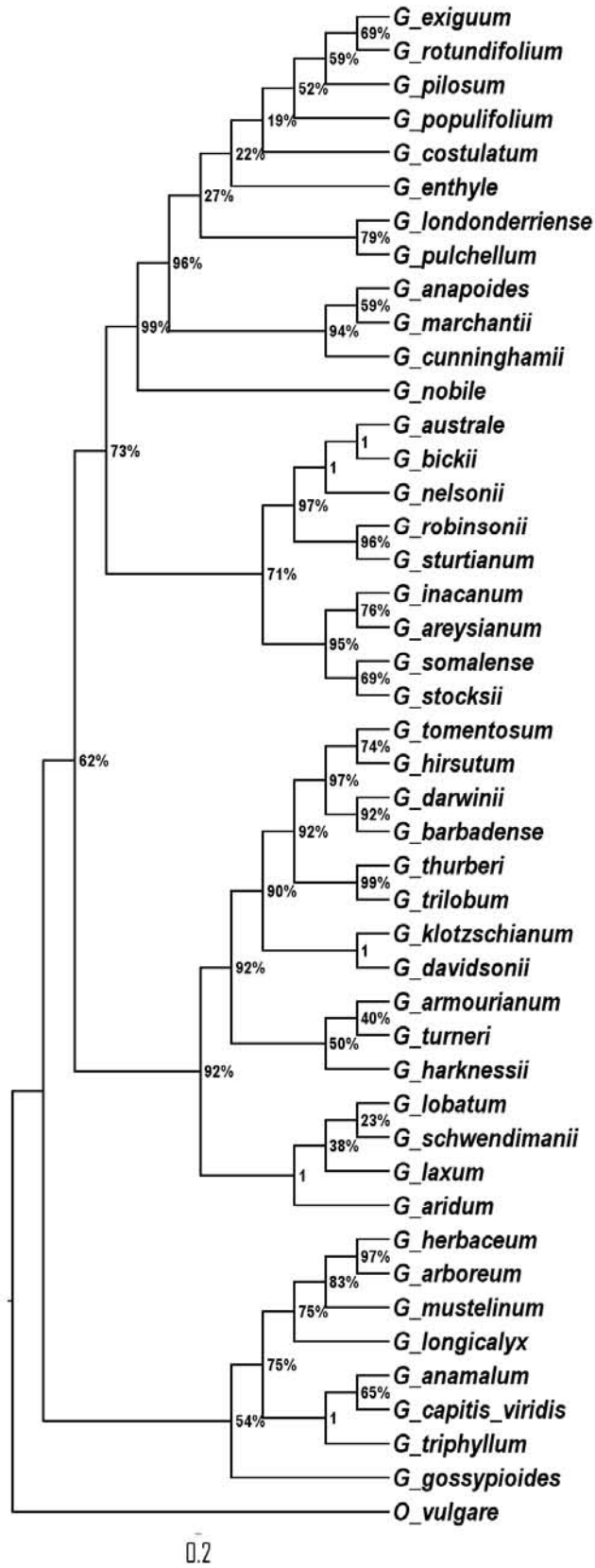


0.06

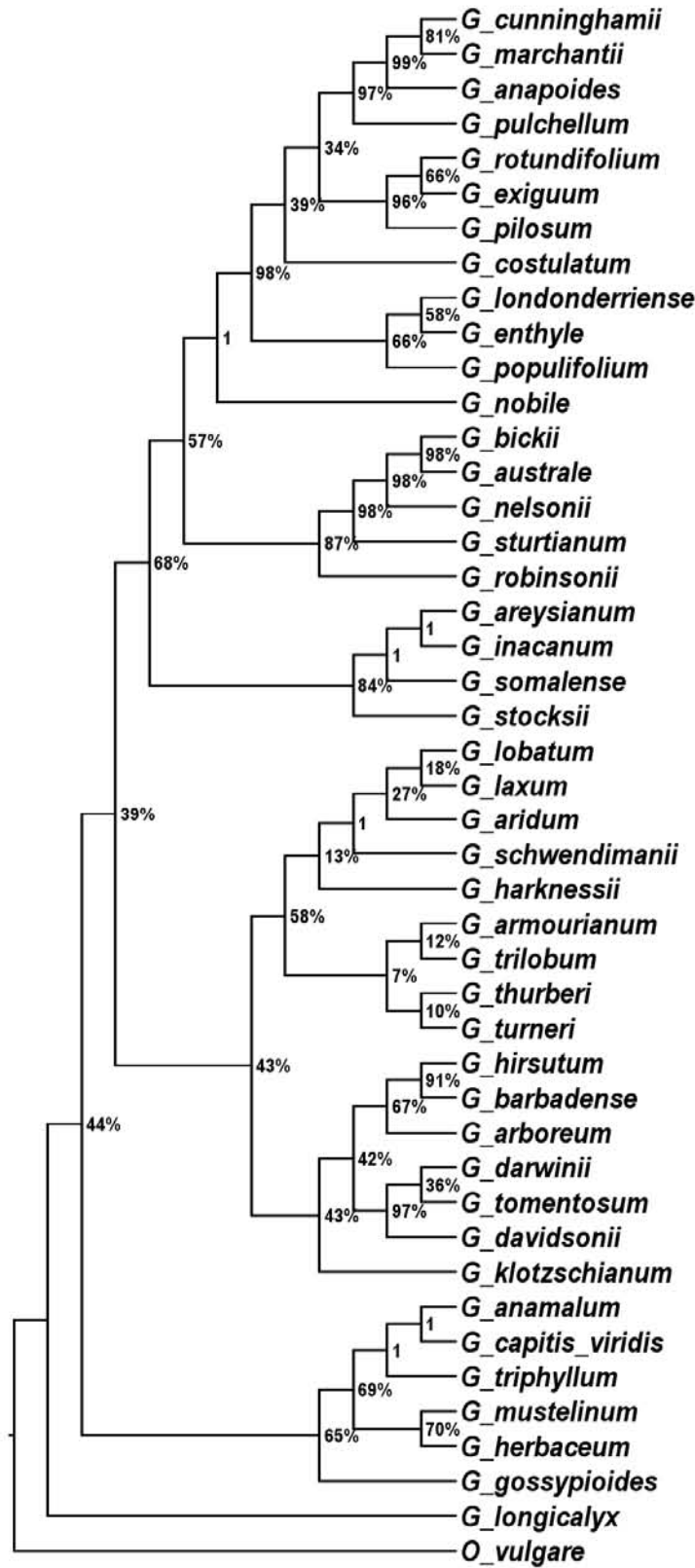
Şekil Ek-12. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



Şekil Ek-13. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi

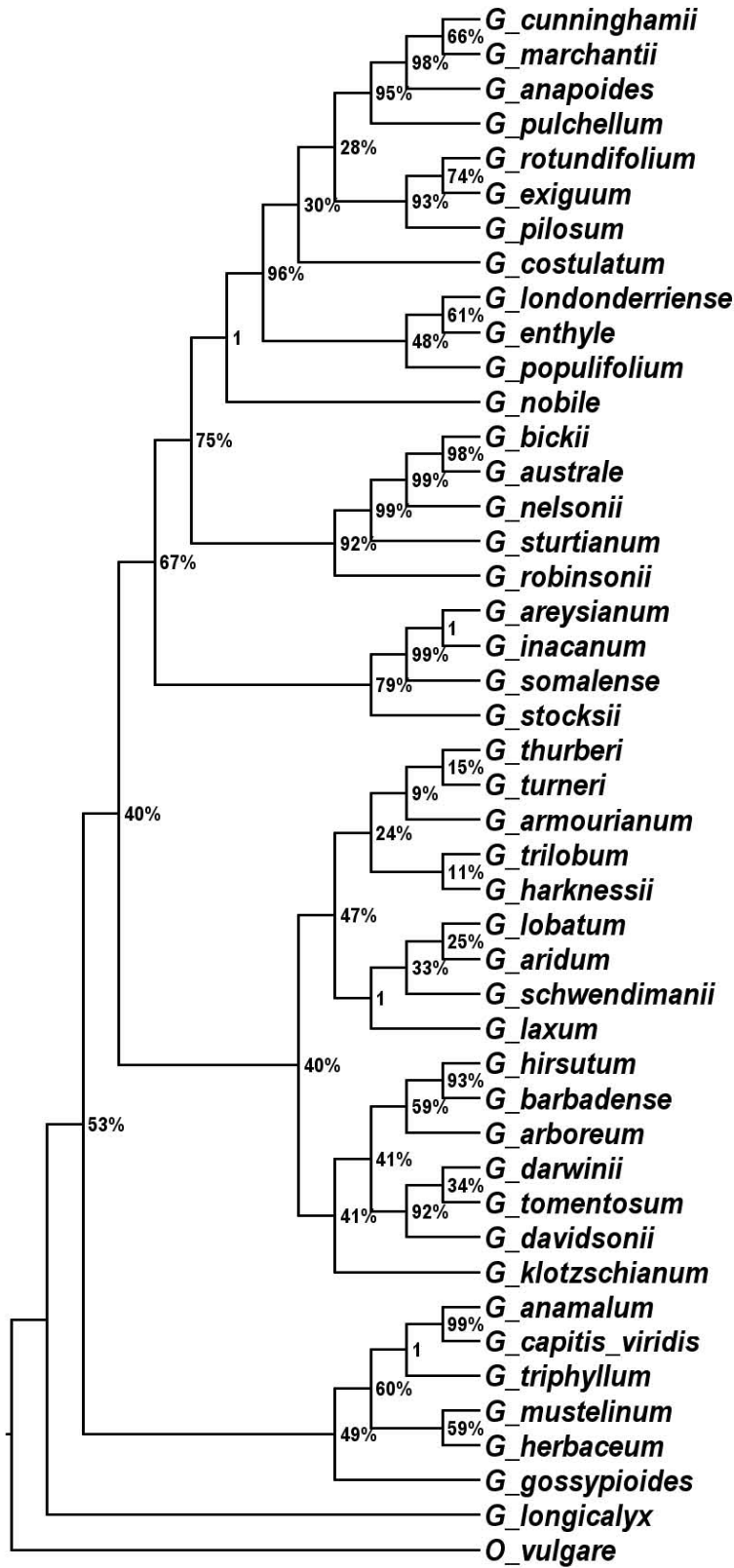


Şekil Ek-14. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi



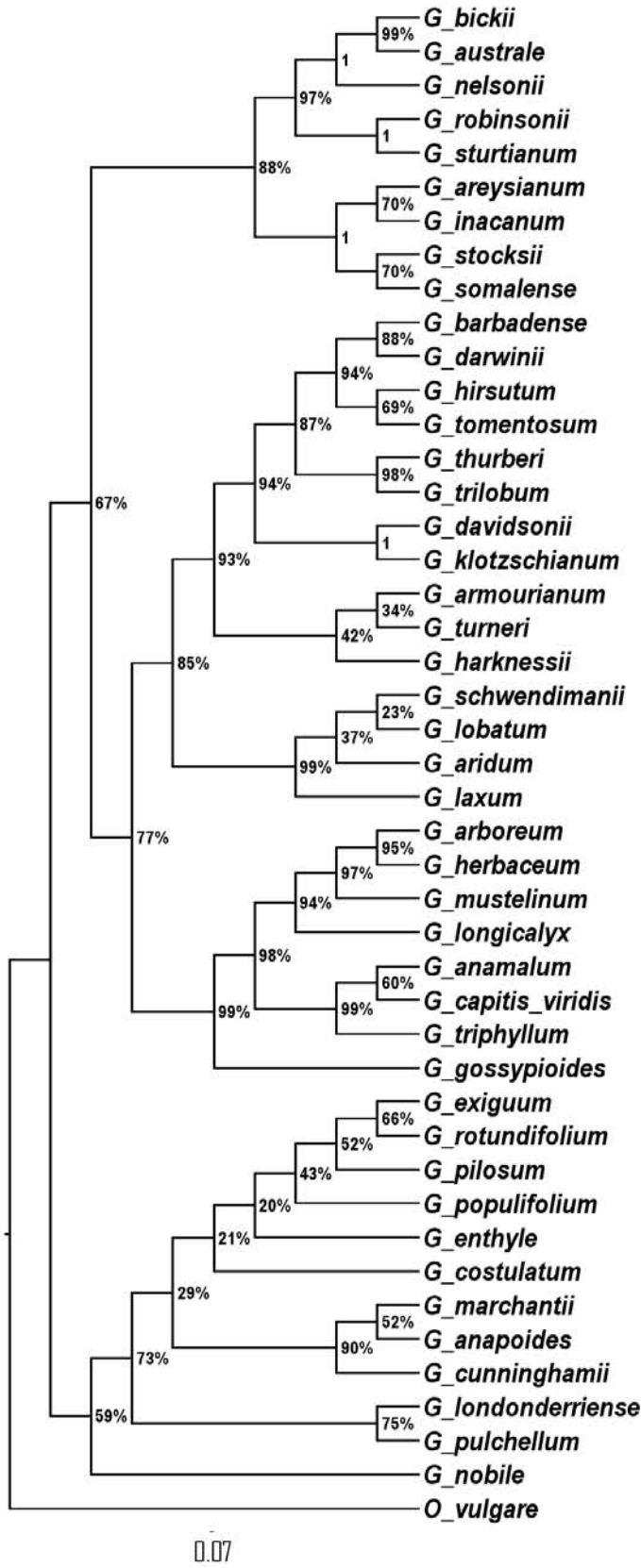
0.1

Şekil Ek-15. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi

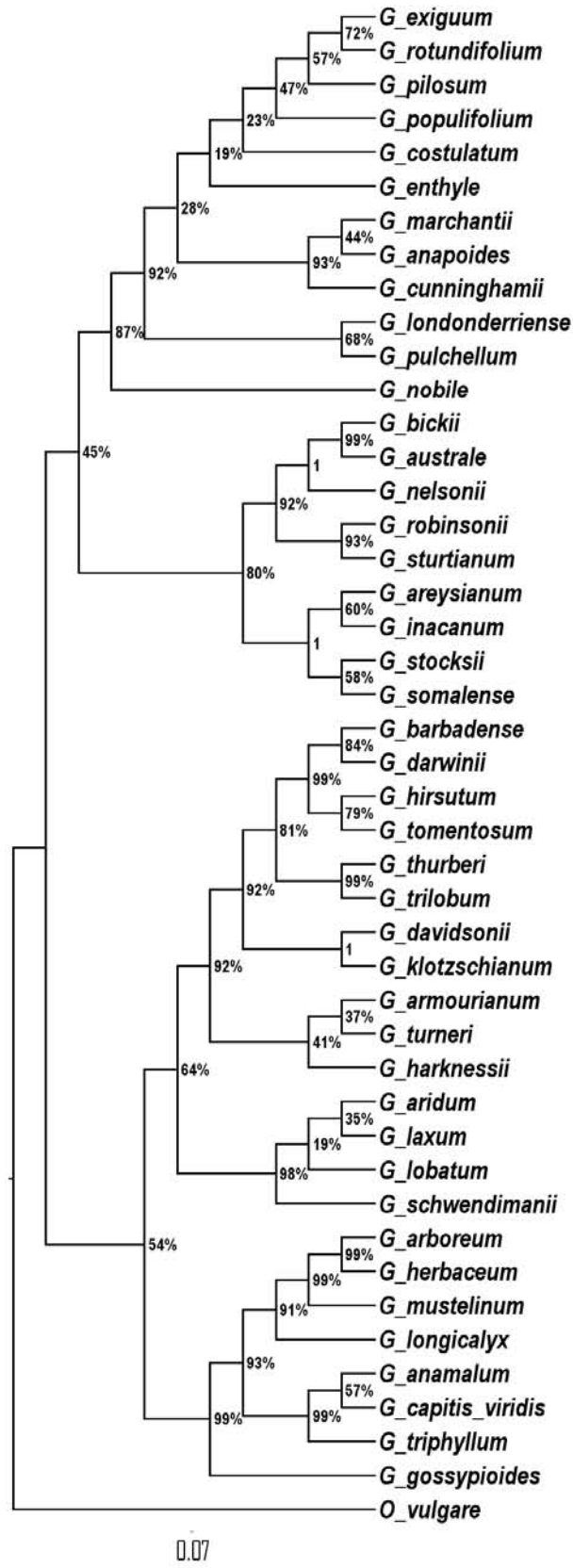


0.1

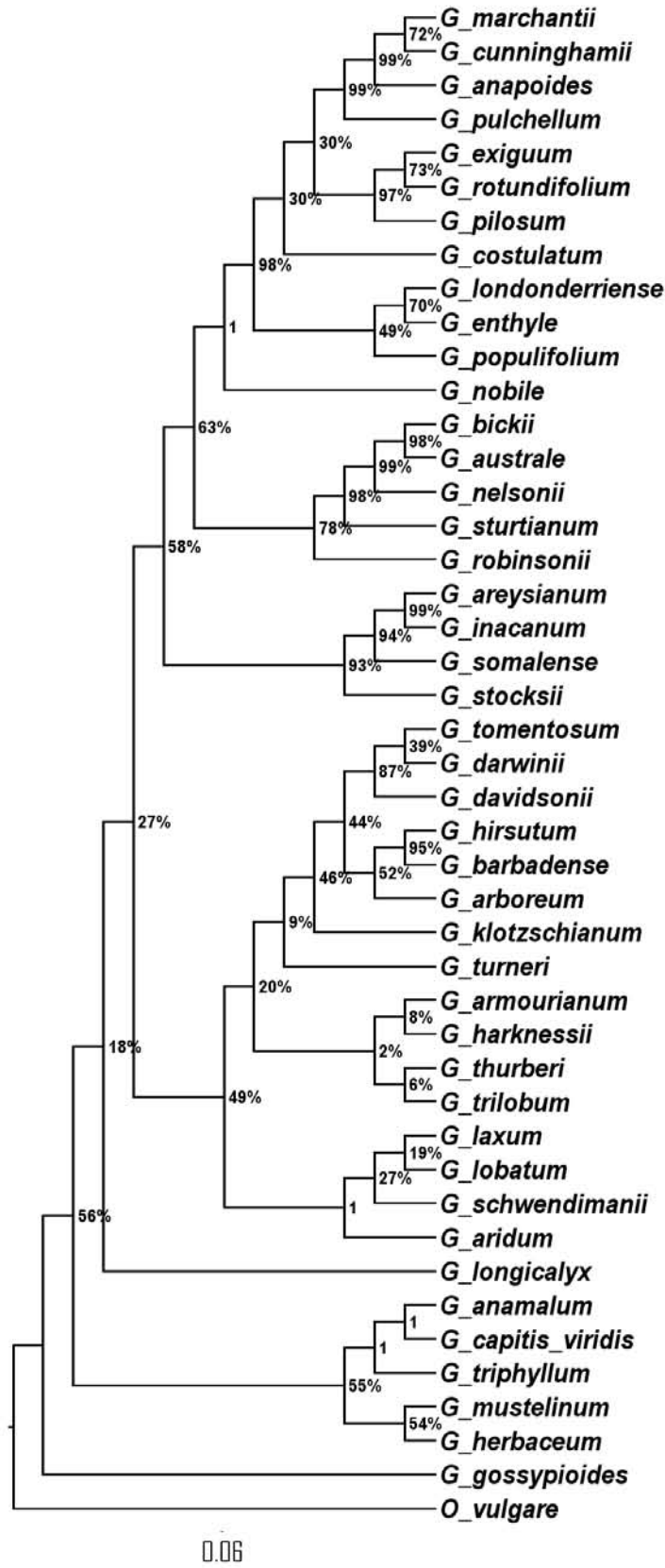
Şekil Ek-16. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



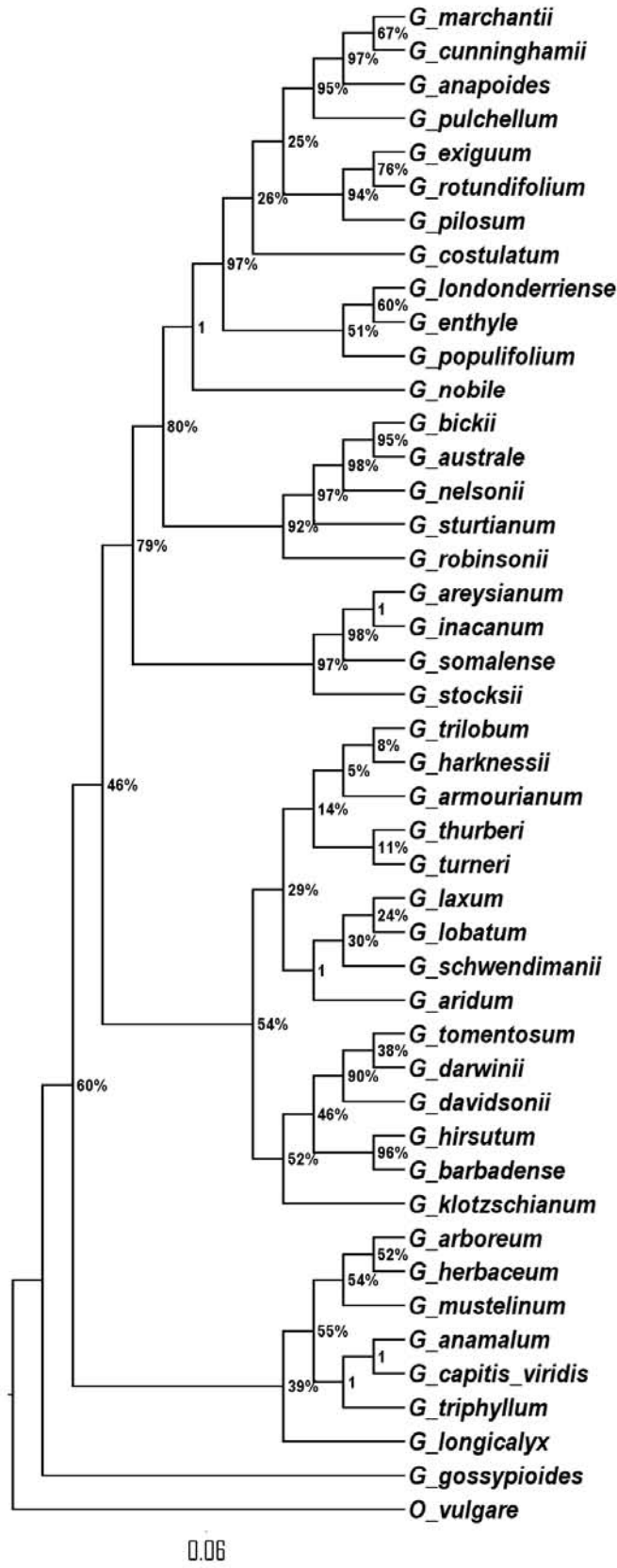
Şekil Ek-17. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi



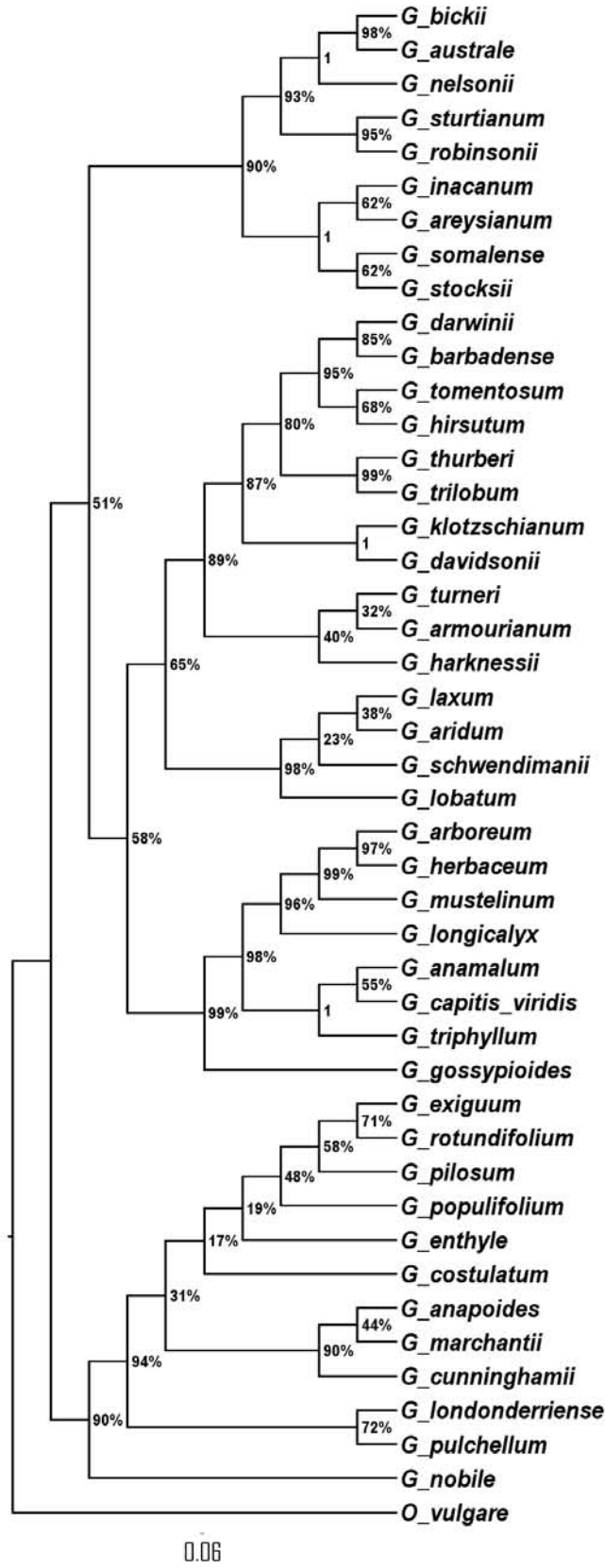
Şekil Ek-18. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi



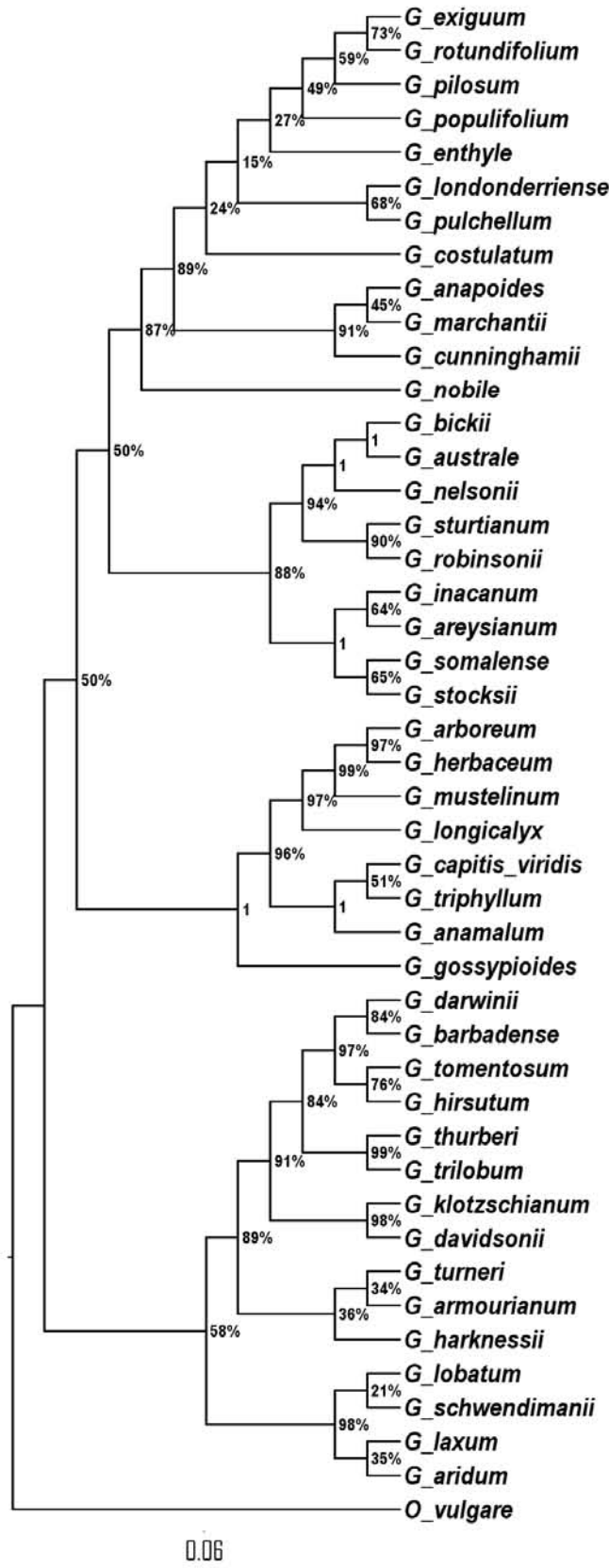
Şekil Ek-19. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi



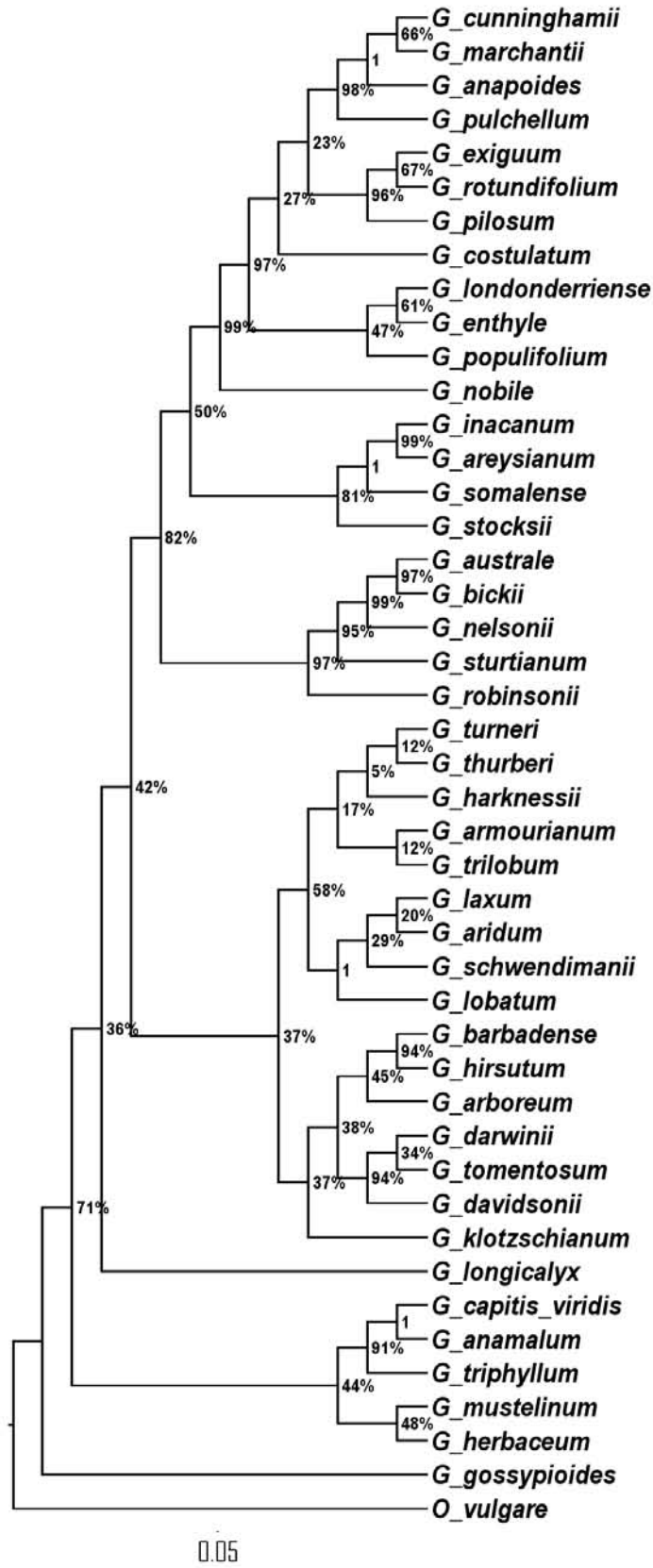
Şekil Ek-20. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



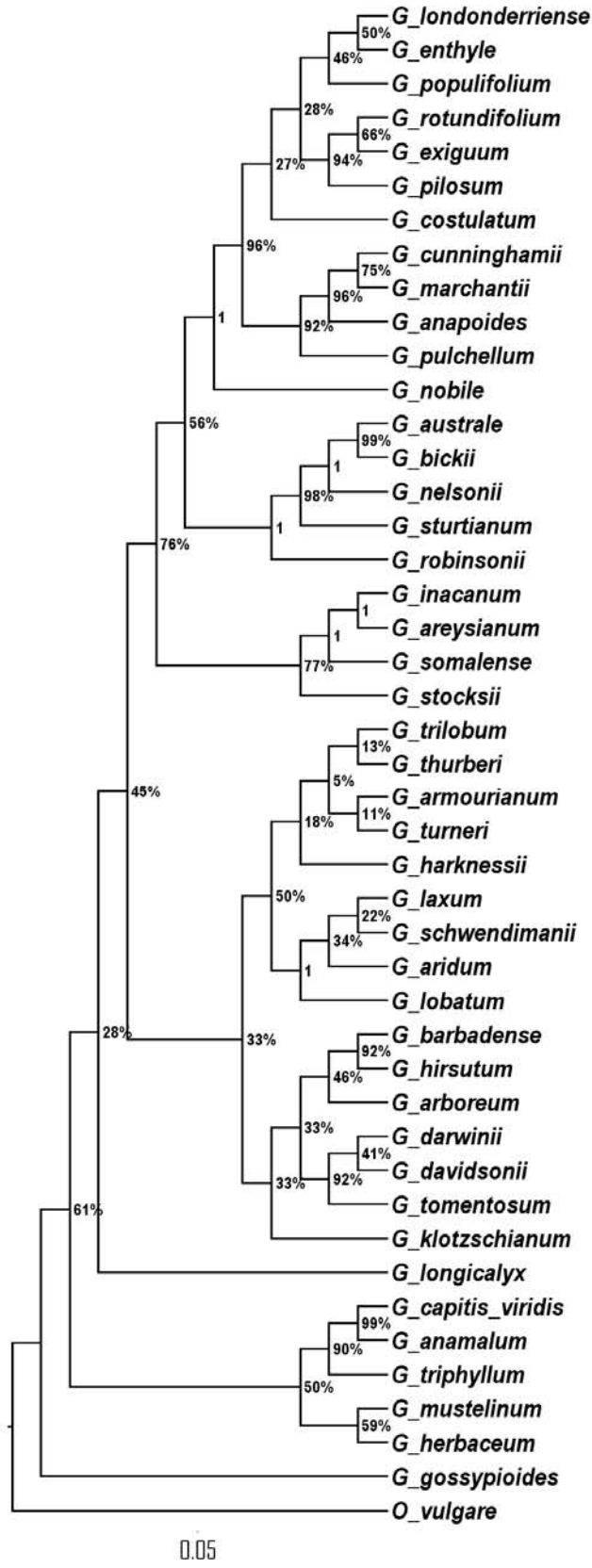
Şekil Ek-21. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi



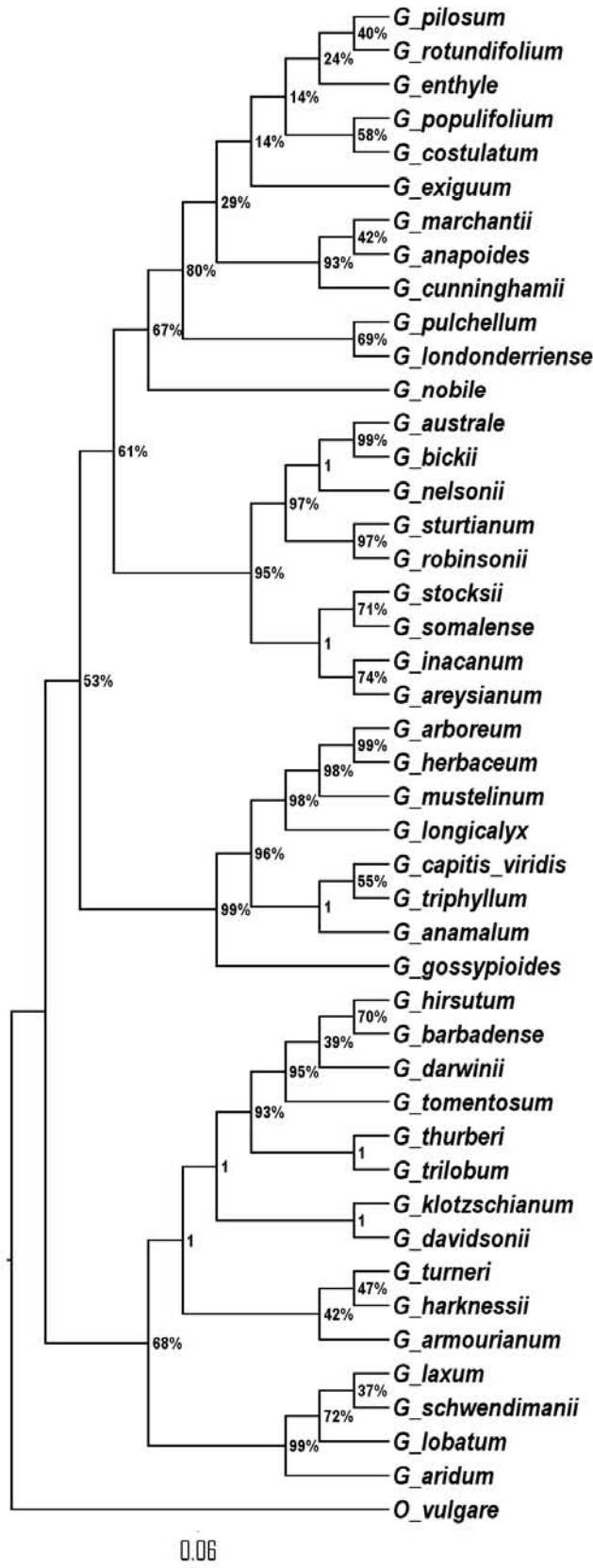
Şekil Ek-22. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi



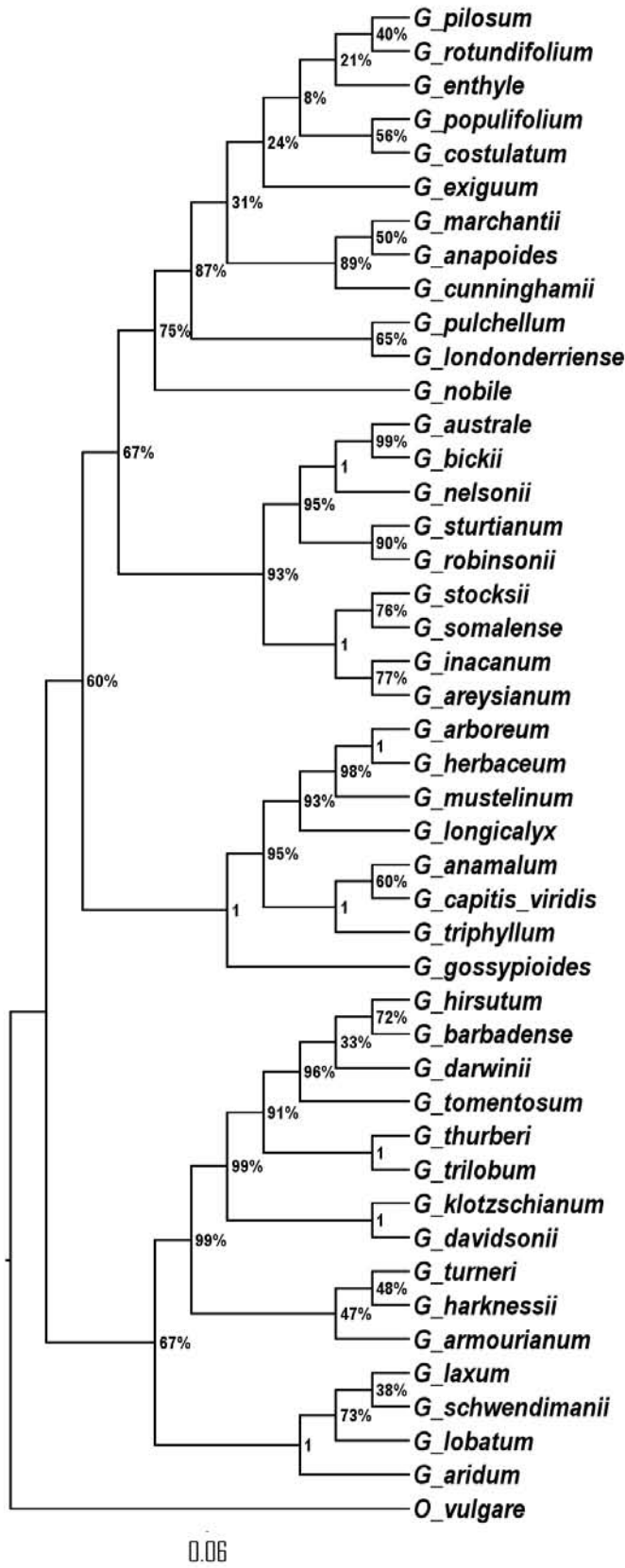
Şekil Ek-23. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi



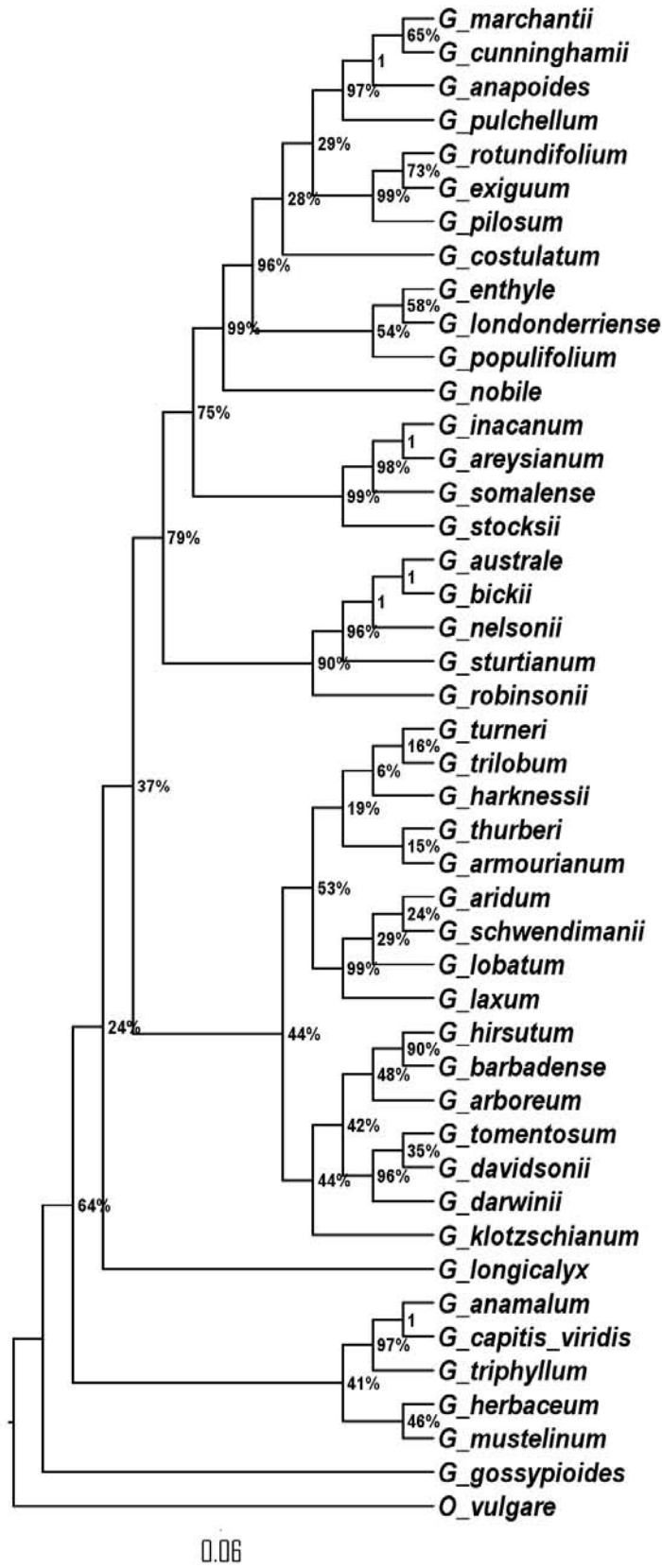
Şekil Ek-24. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



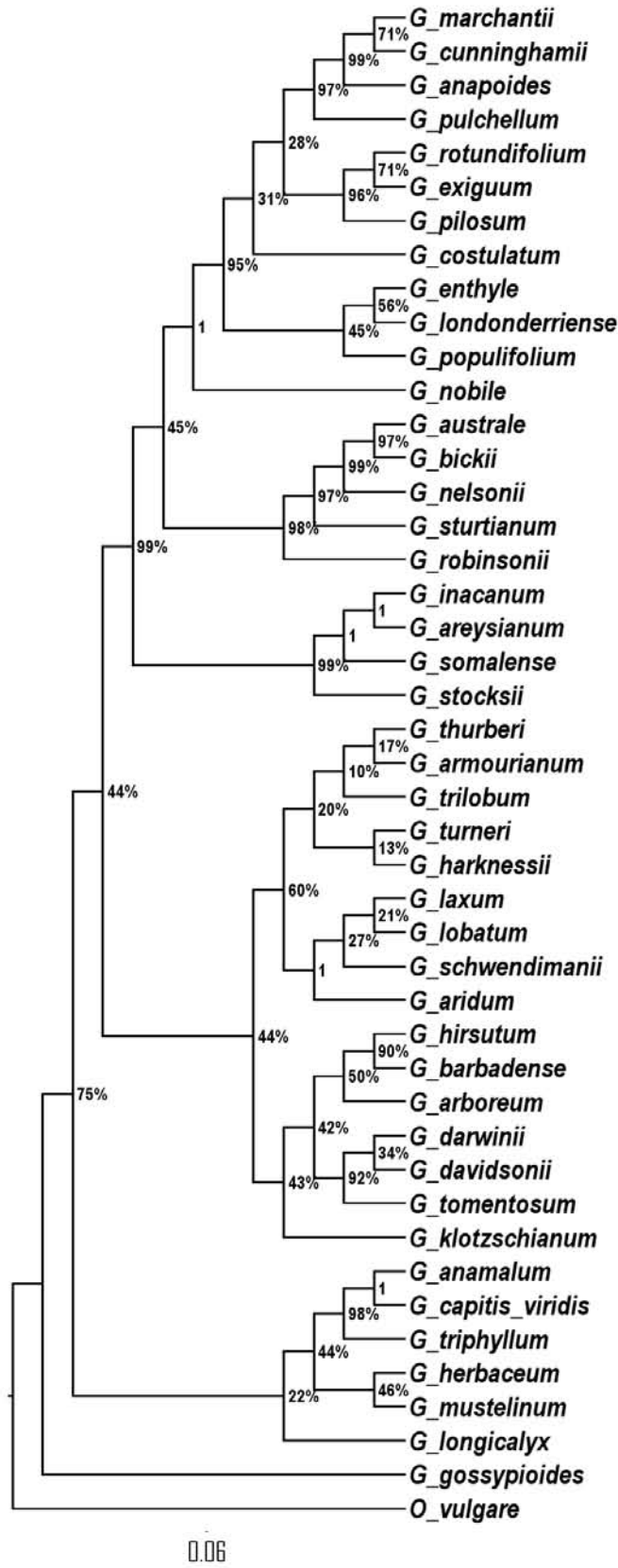
Şekil Ek-25. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi



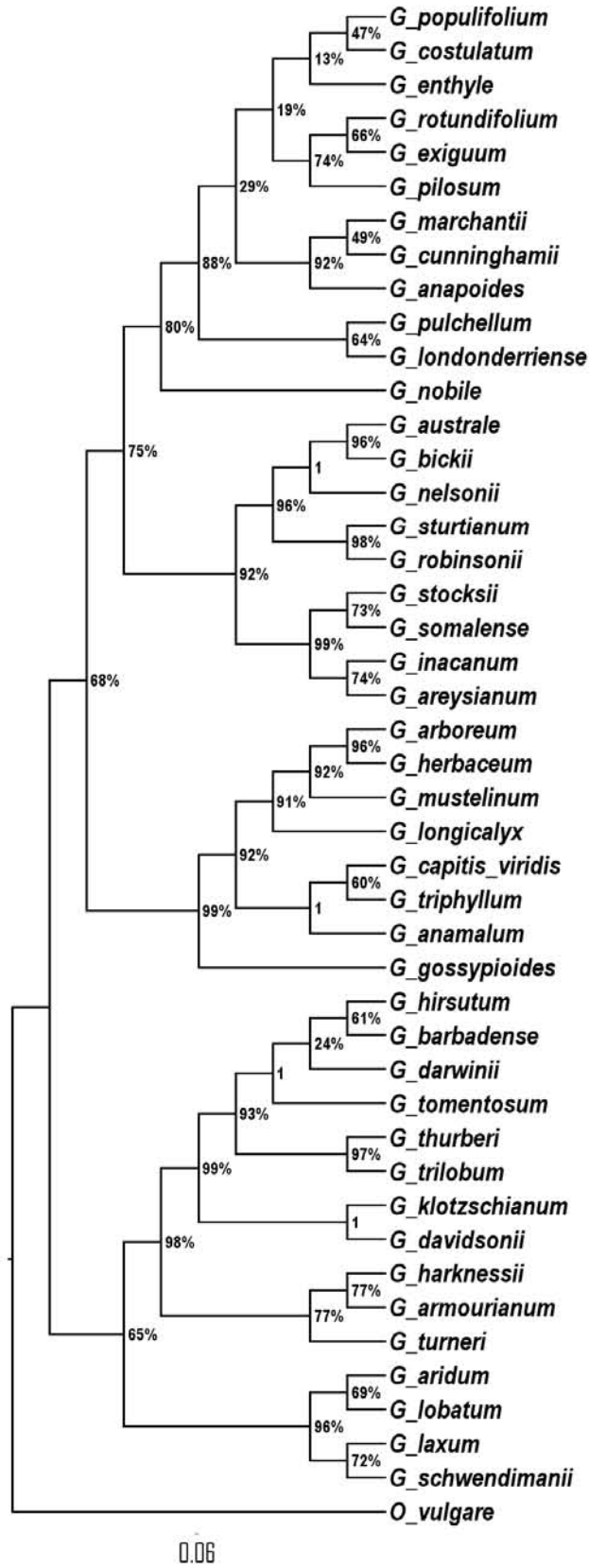
Şekil Ek-26. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi



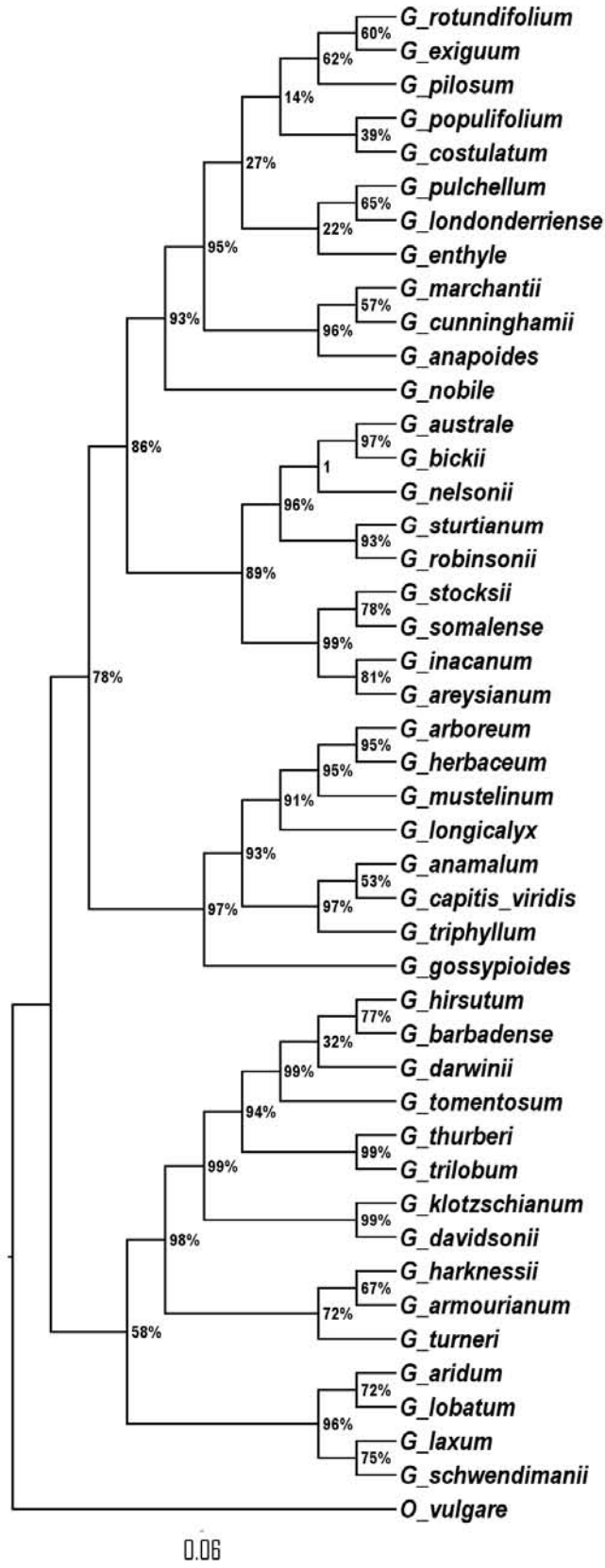
Şekil Ek-27. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi



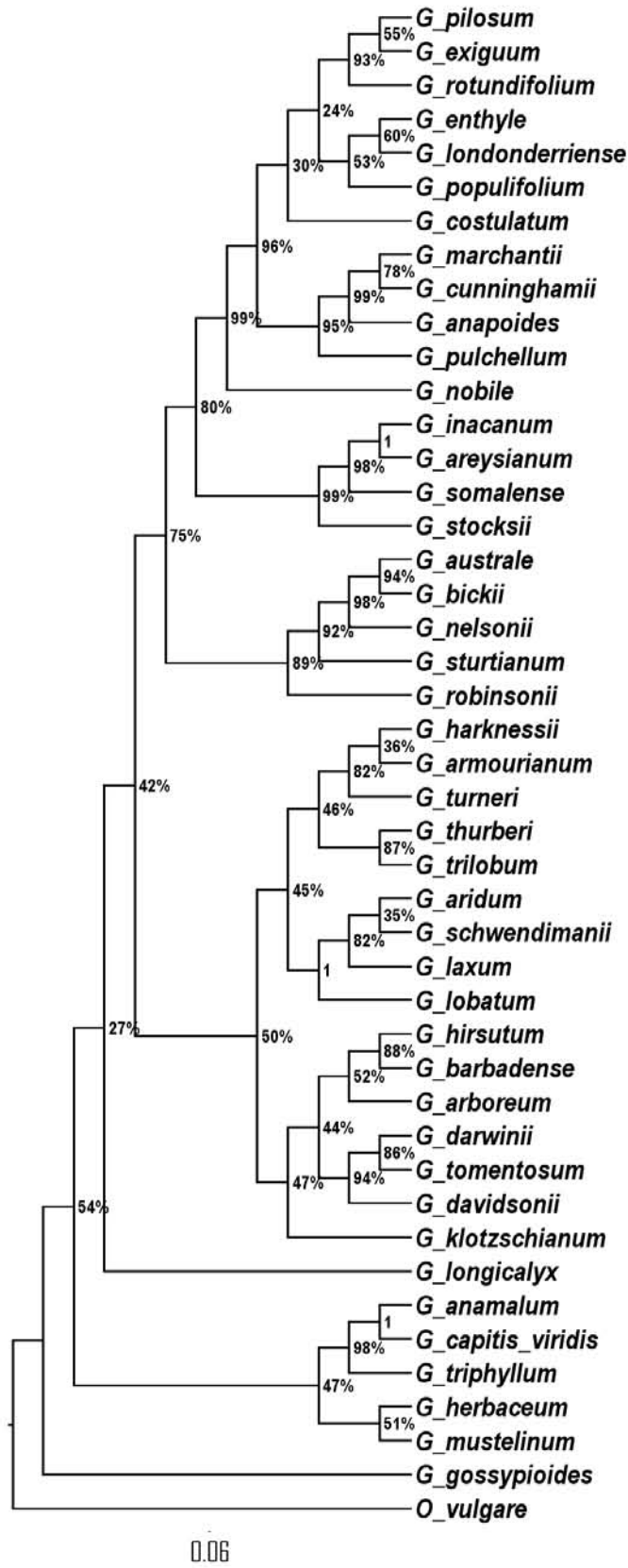
Şekil Ek-28. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



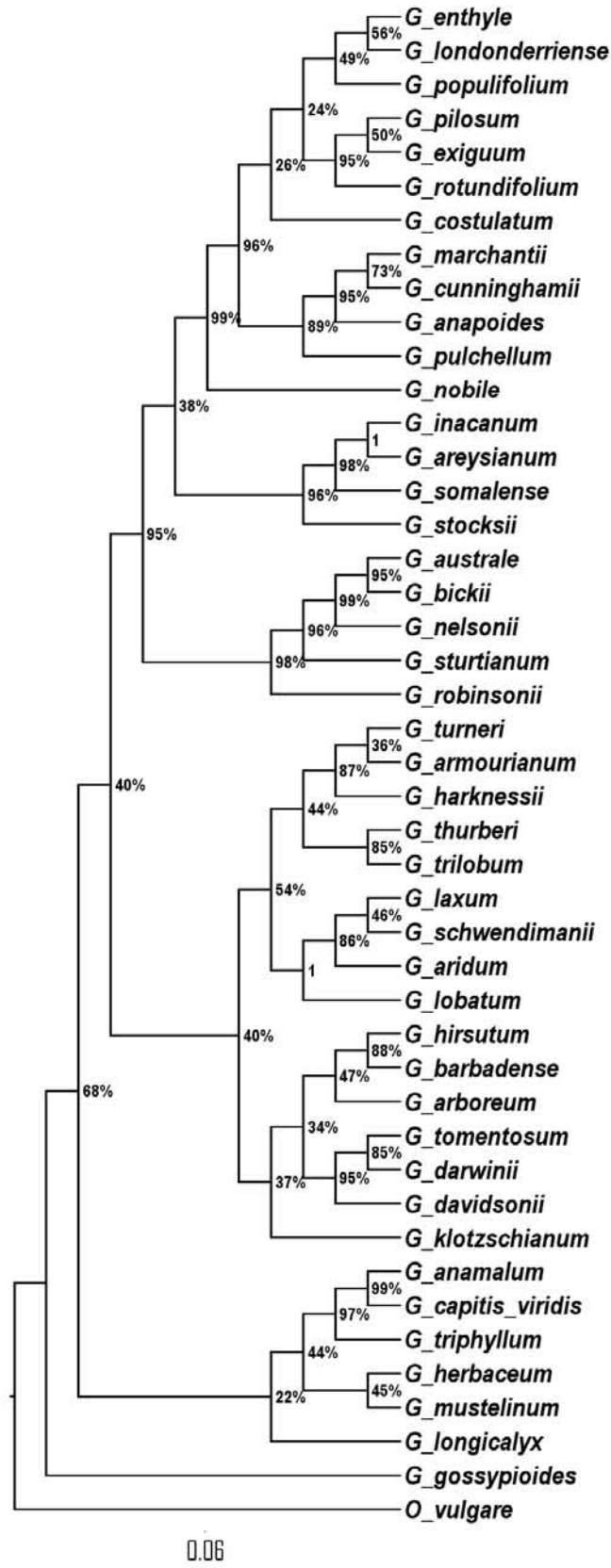
Şekil Ek-29. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki ML analizi



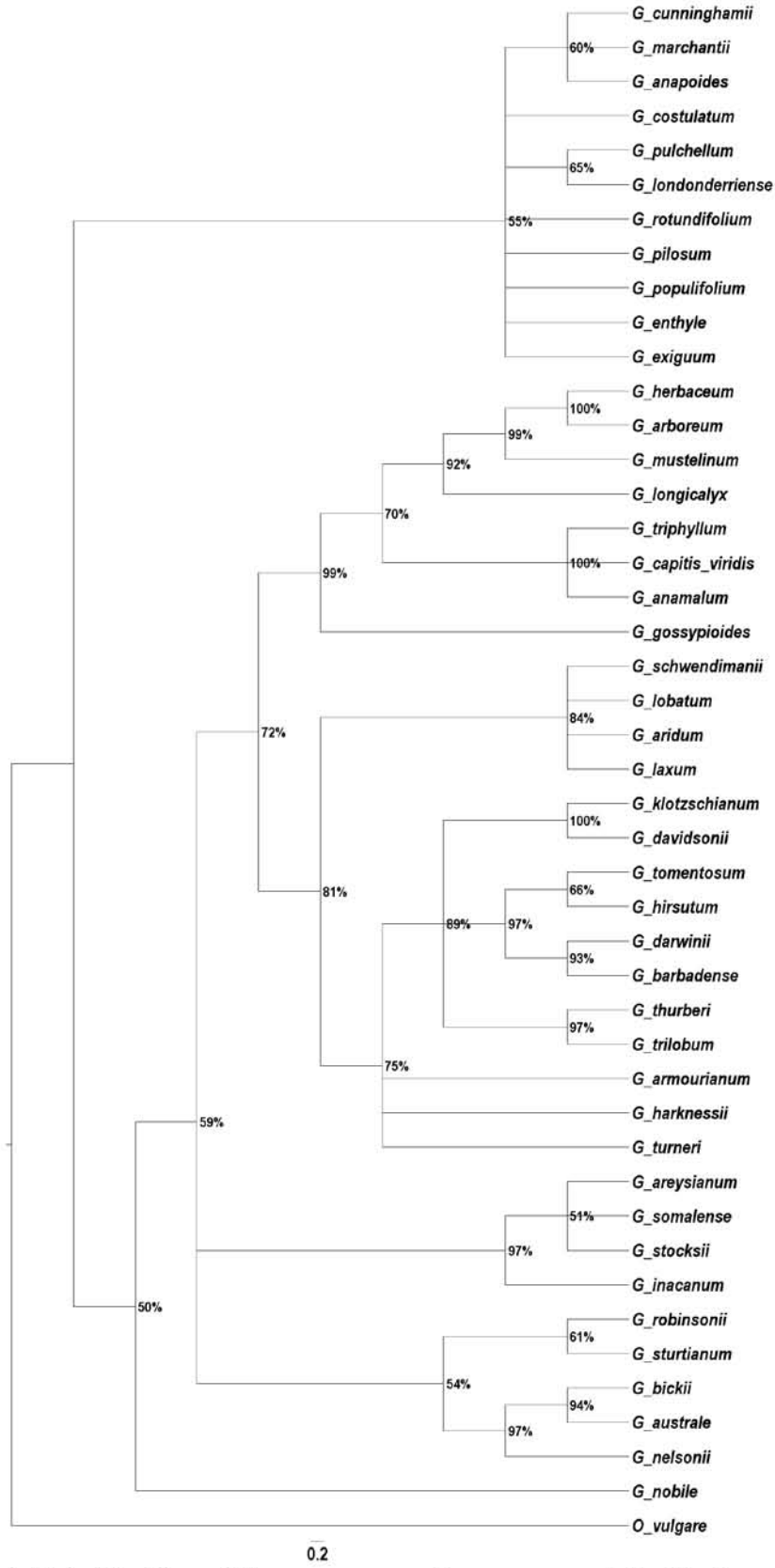
Şekil Ek-30. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının TN93 SBT modelindeki ML analizi



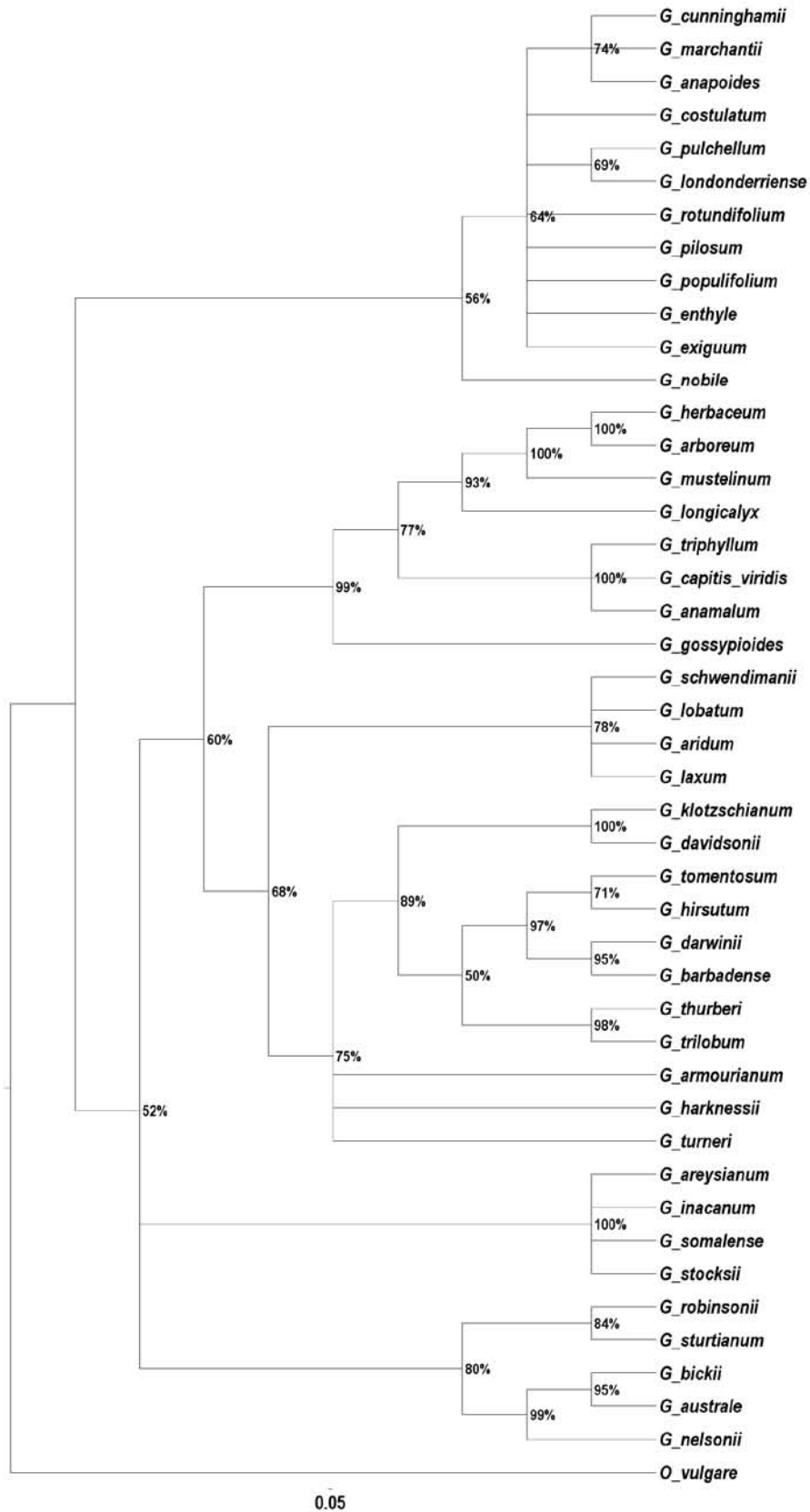
Şekil Ek-31. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki ML analizi



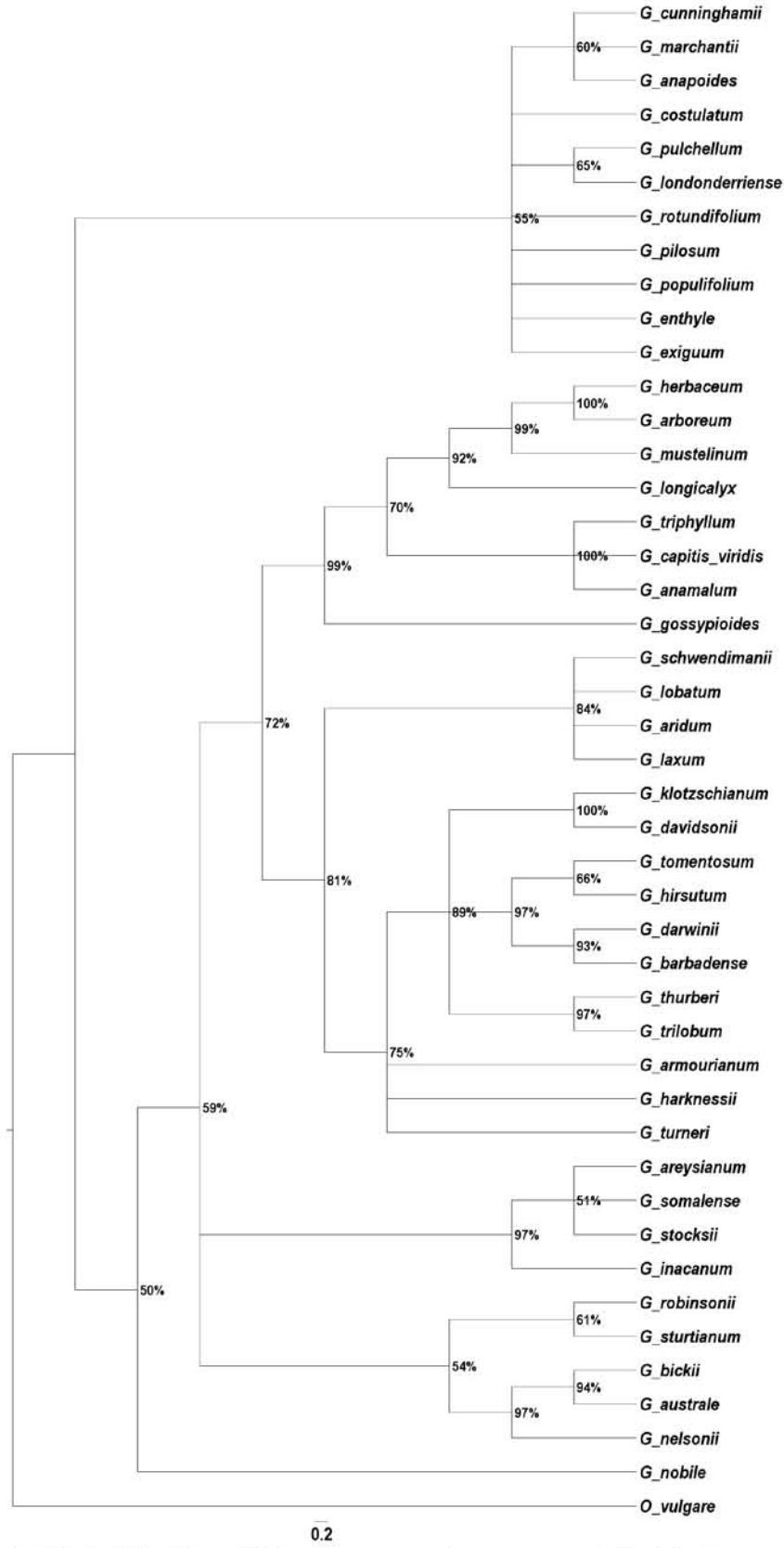
Şekil Ek-32. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki ML analizi



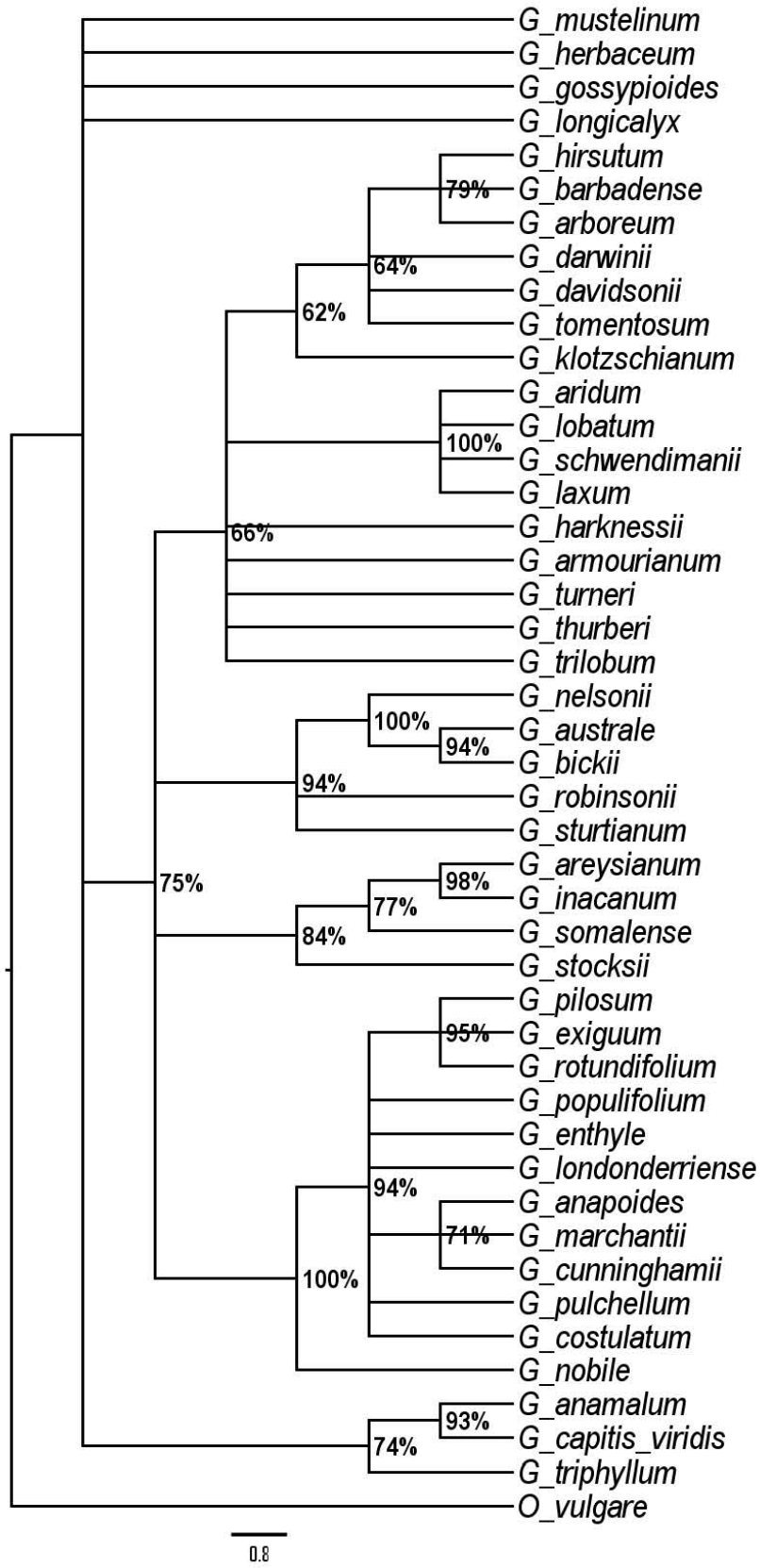
Şekil Ek-33. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



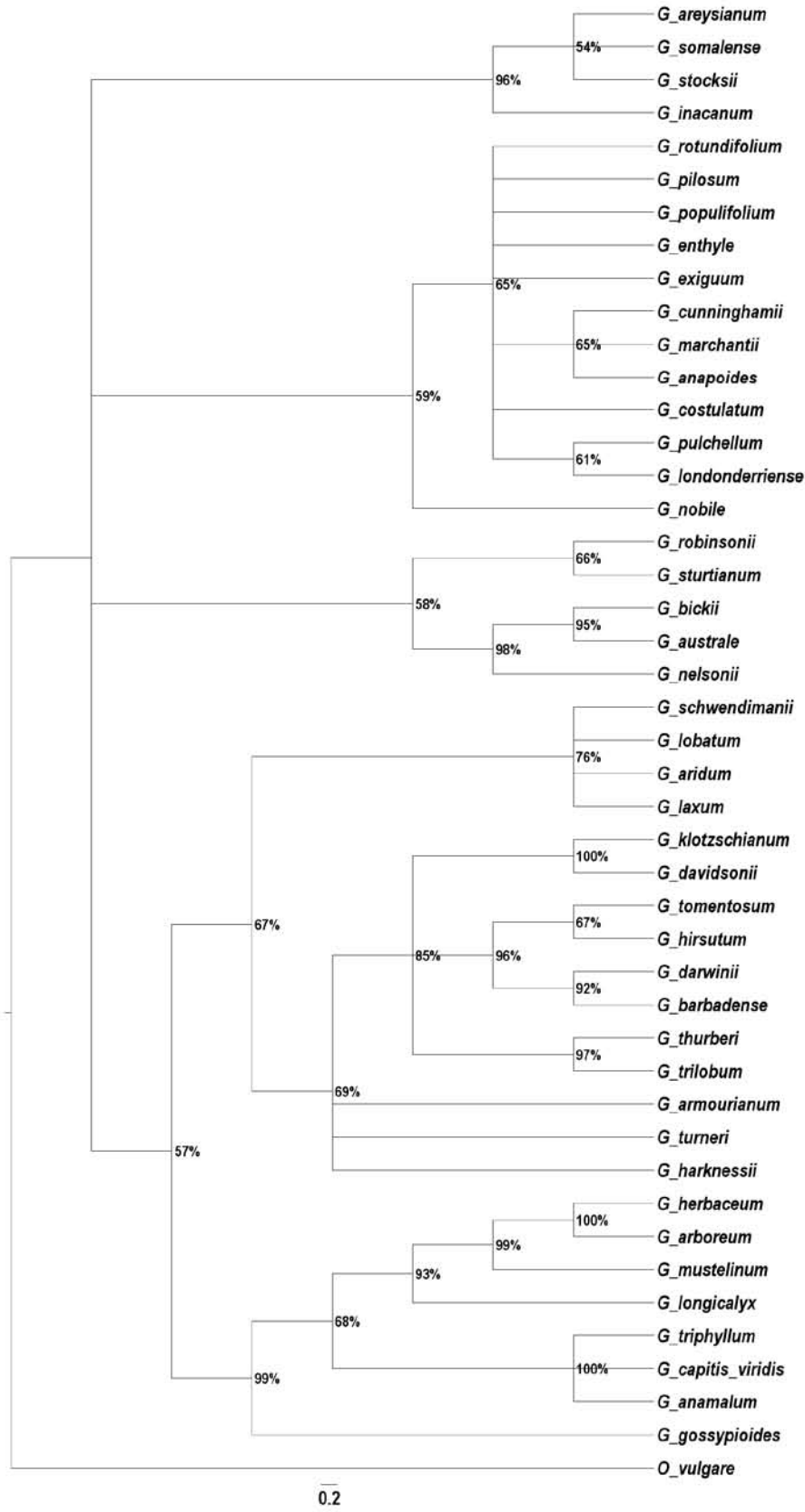
Şekil Ek-34. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



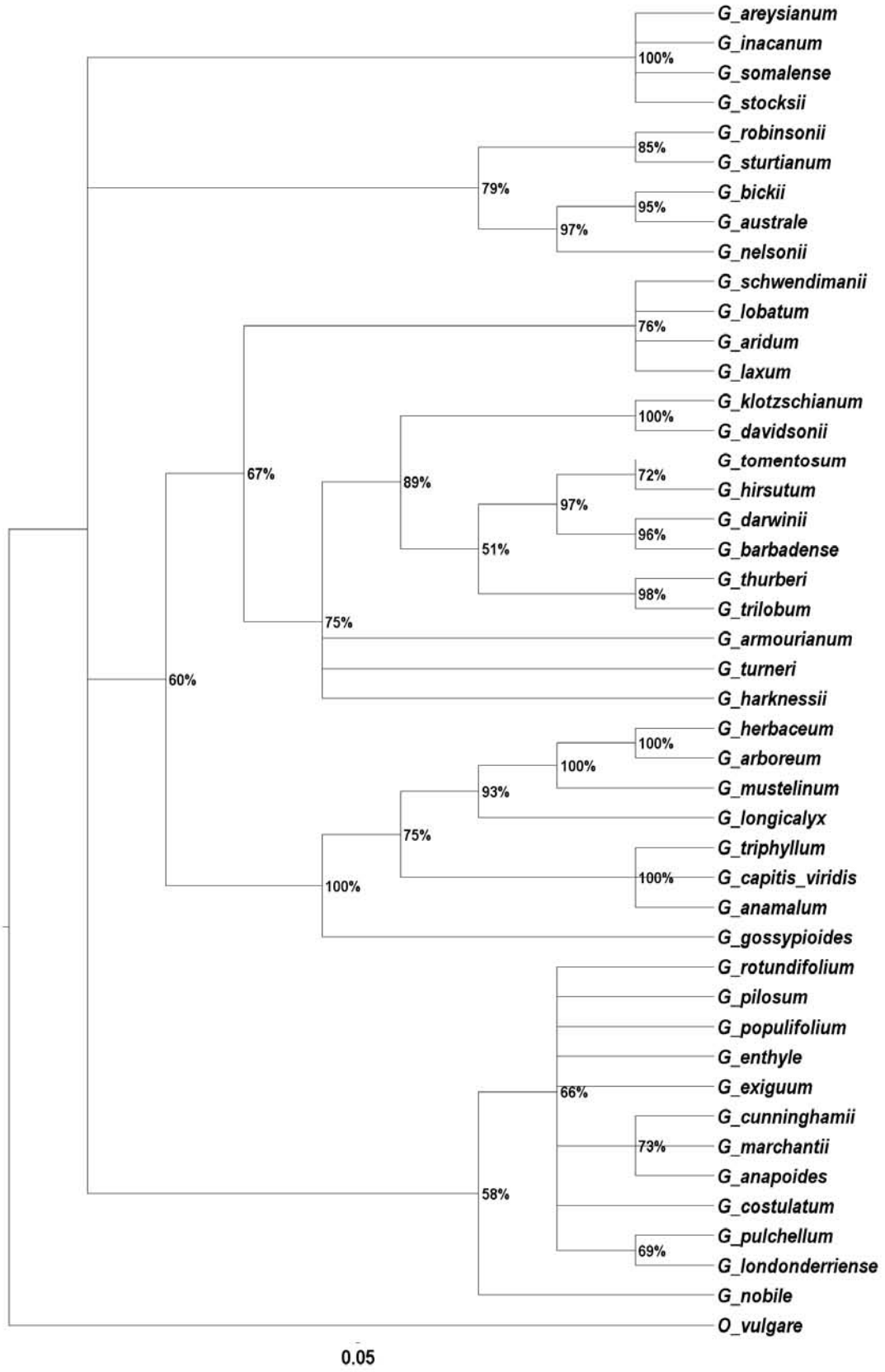
Şekil Ek-35. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



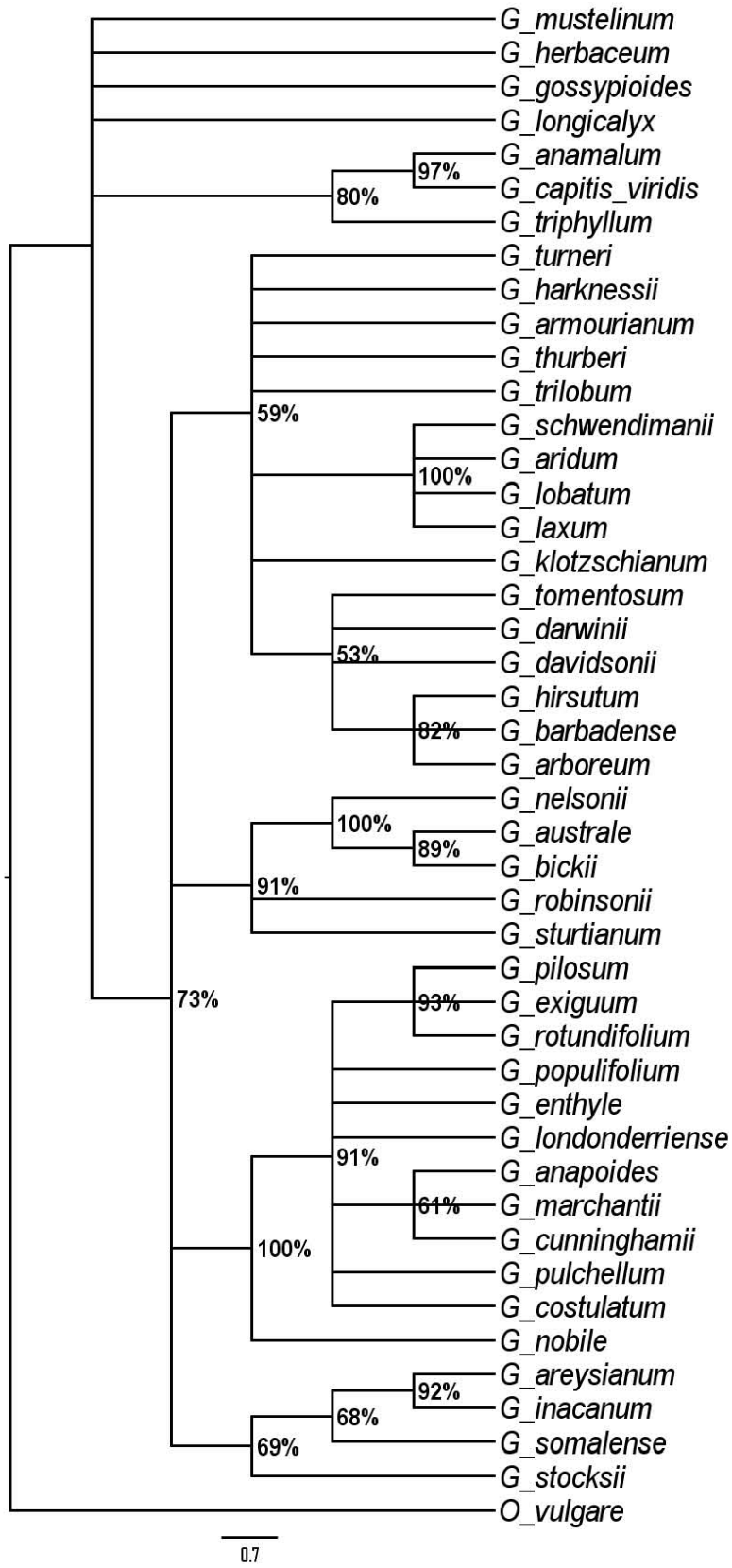
Şekil Ek-36. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



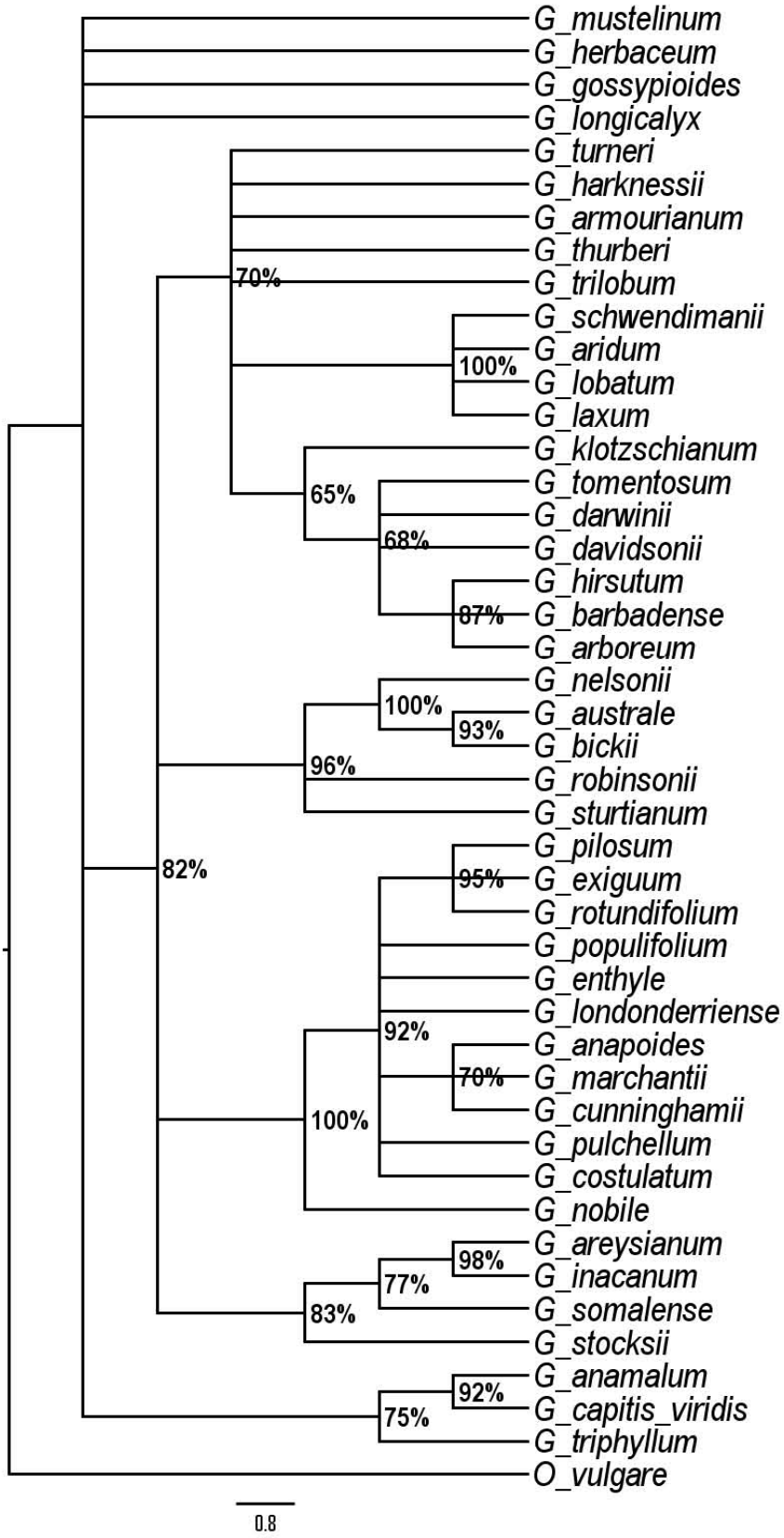
Şekil Ek-37. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



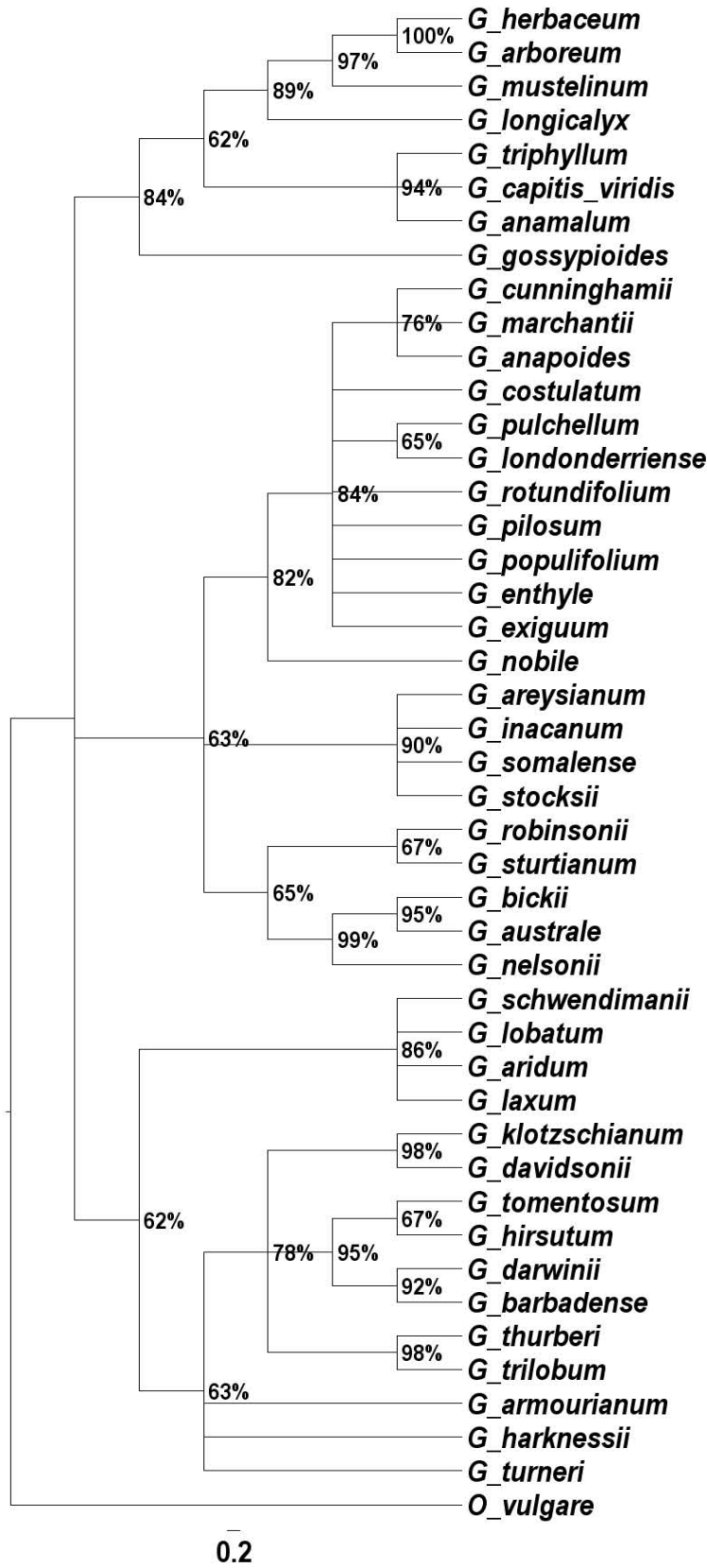
Şekil Ek-38. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



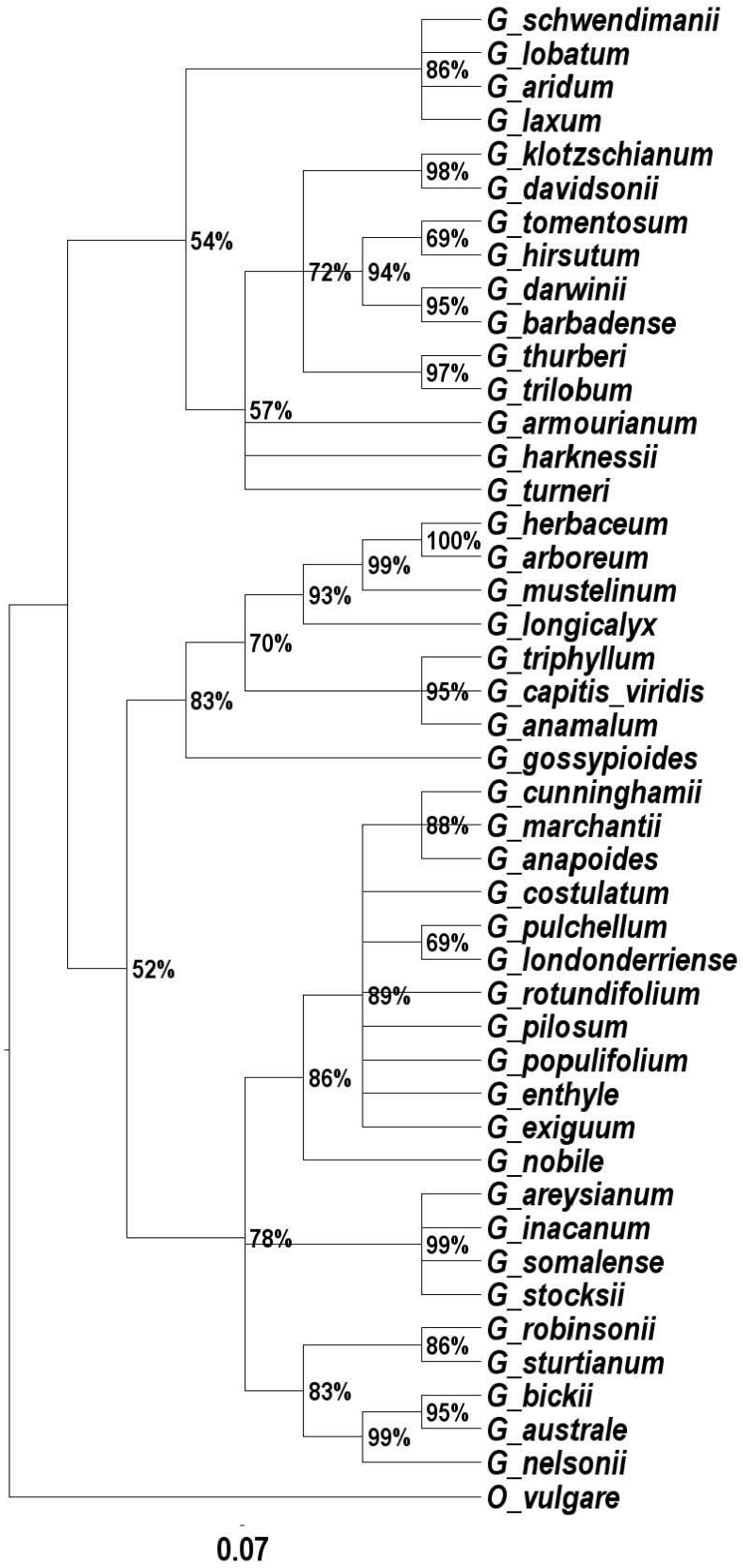
Şekil Ek-39. ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



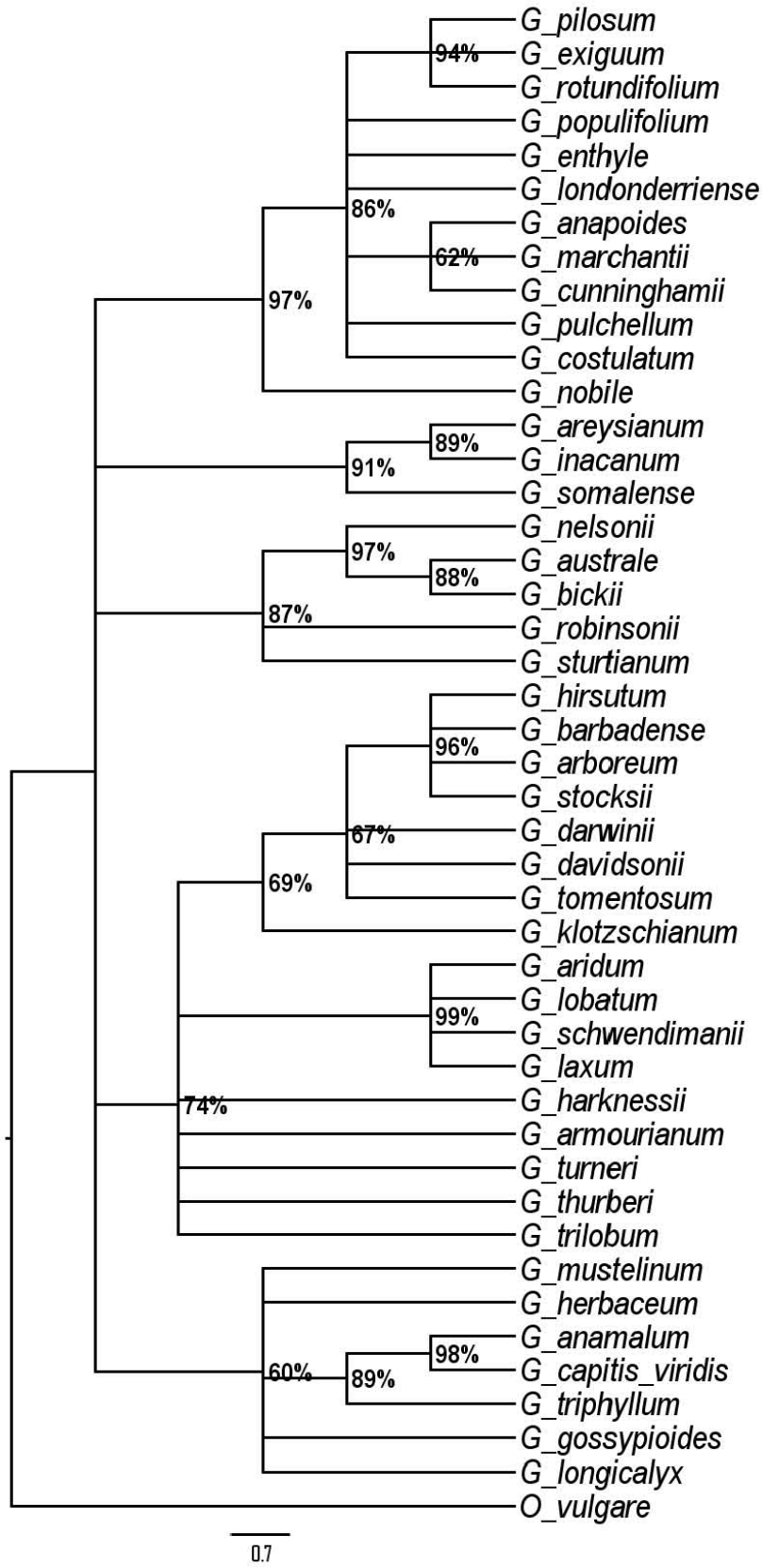
Şekil Ek-40. ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



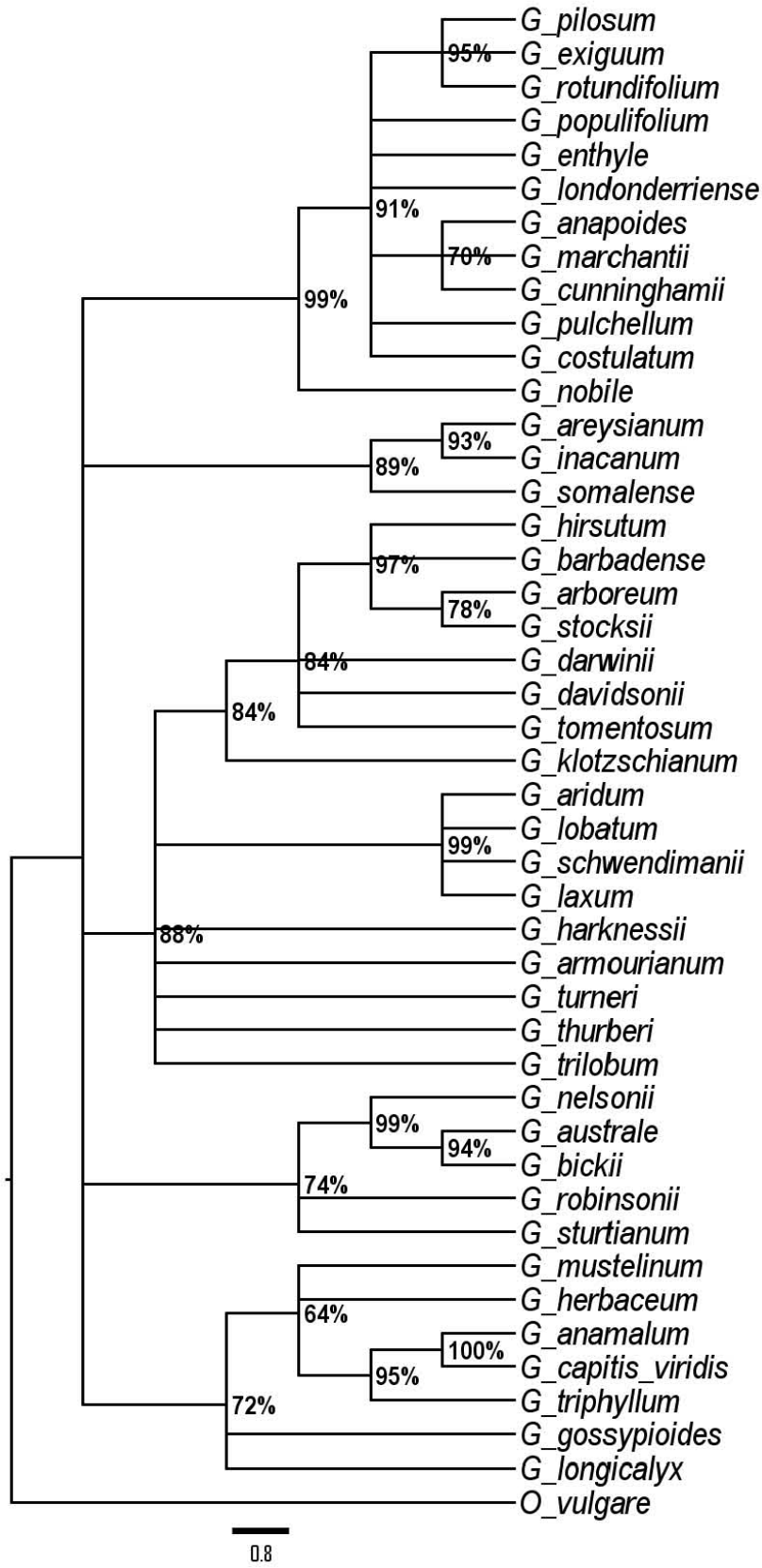
Şekil Ek-41. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



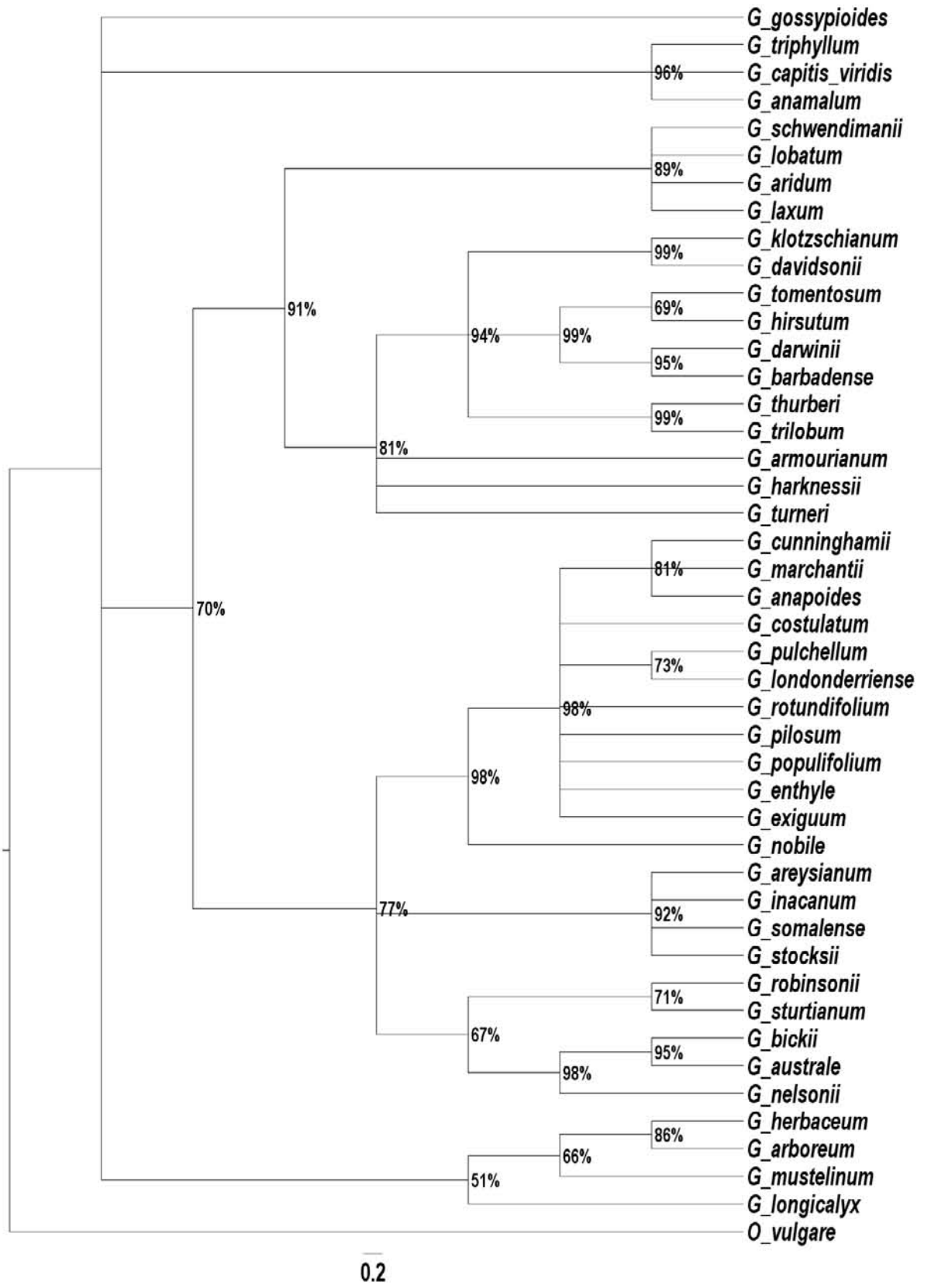
Şekil Ek-42. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



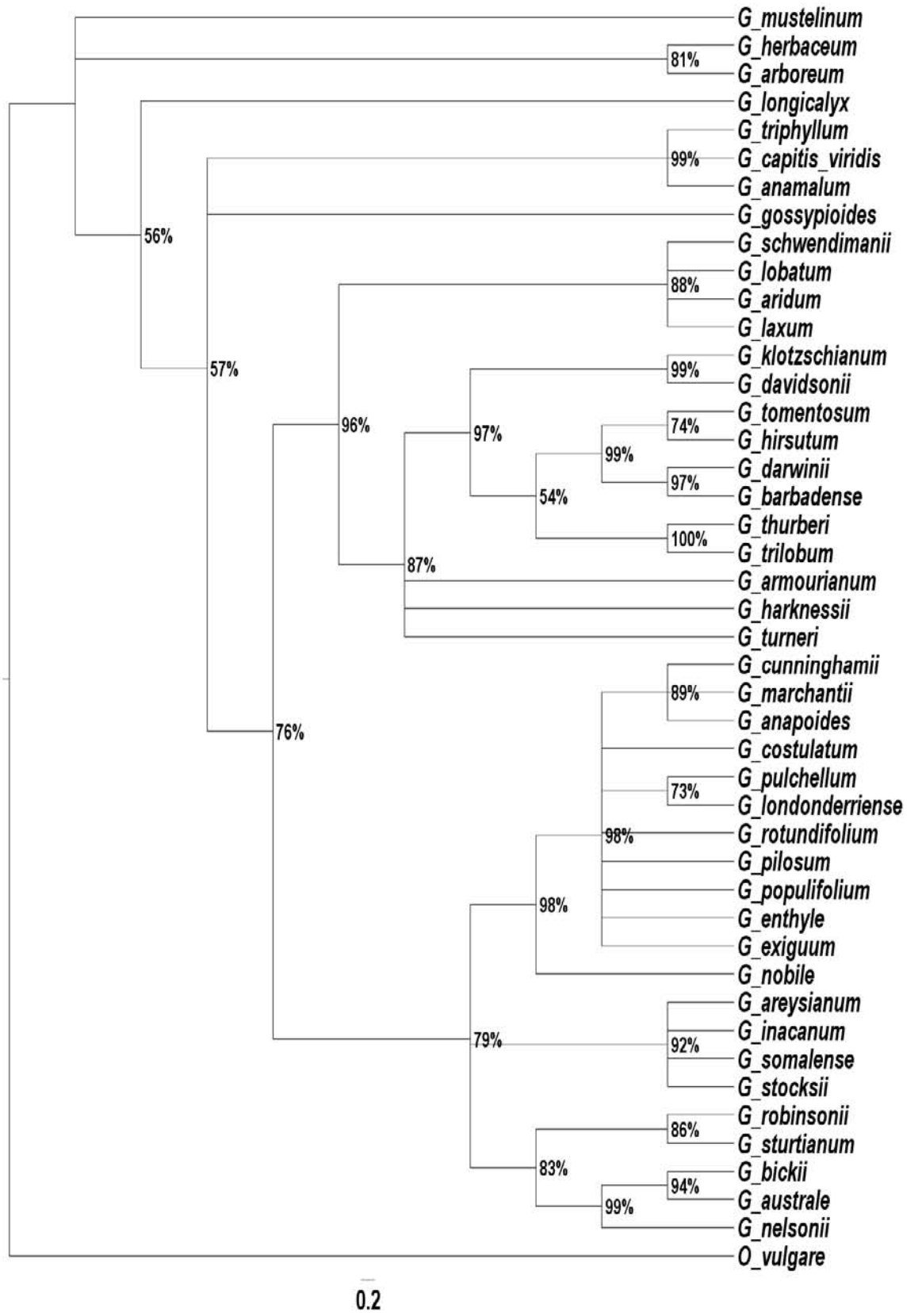
Şekil Ek-43. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



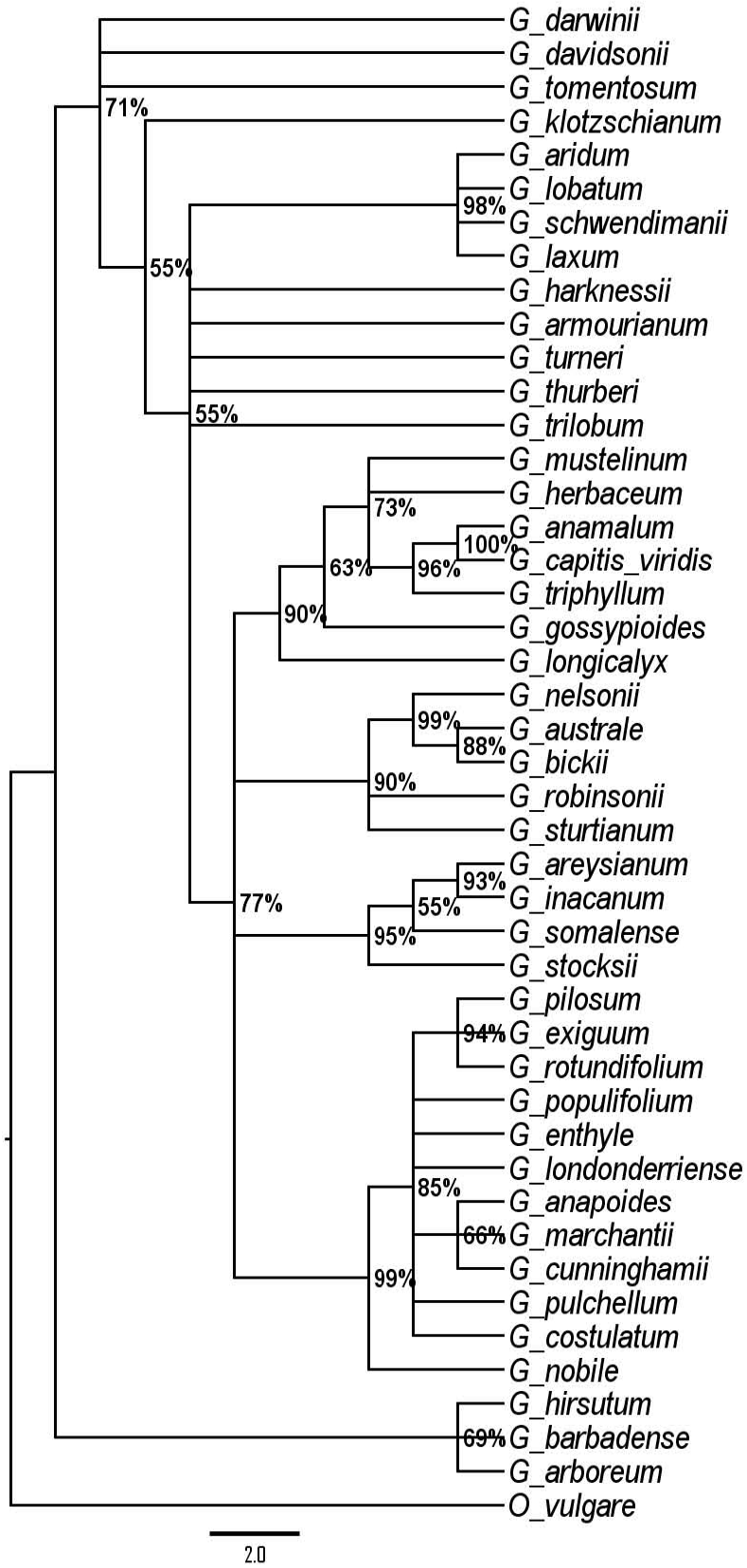
Şekil Ek-44. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analiz



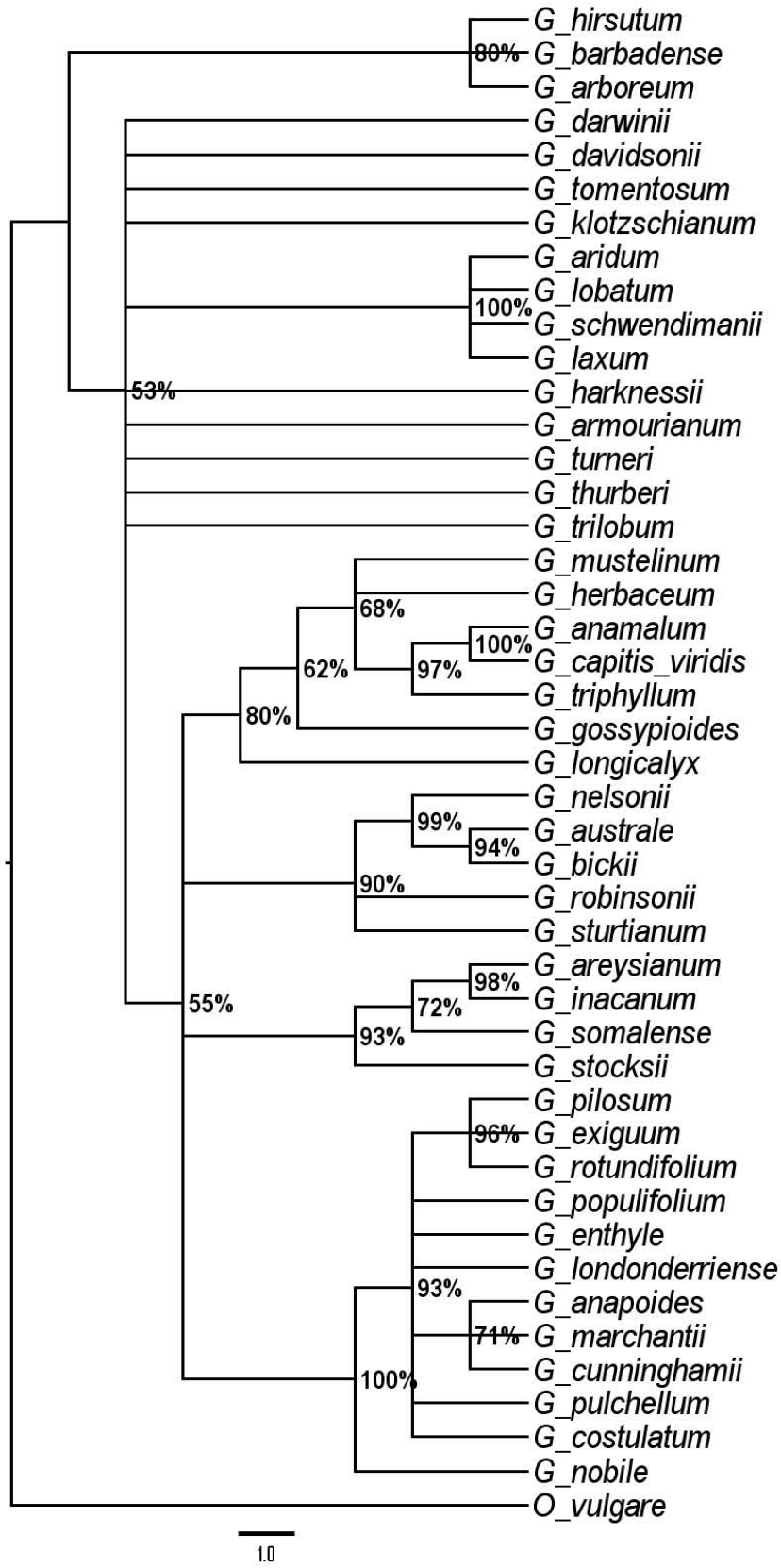
Şekil Ek-45. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



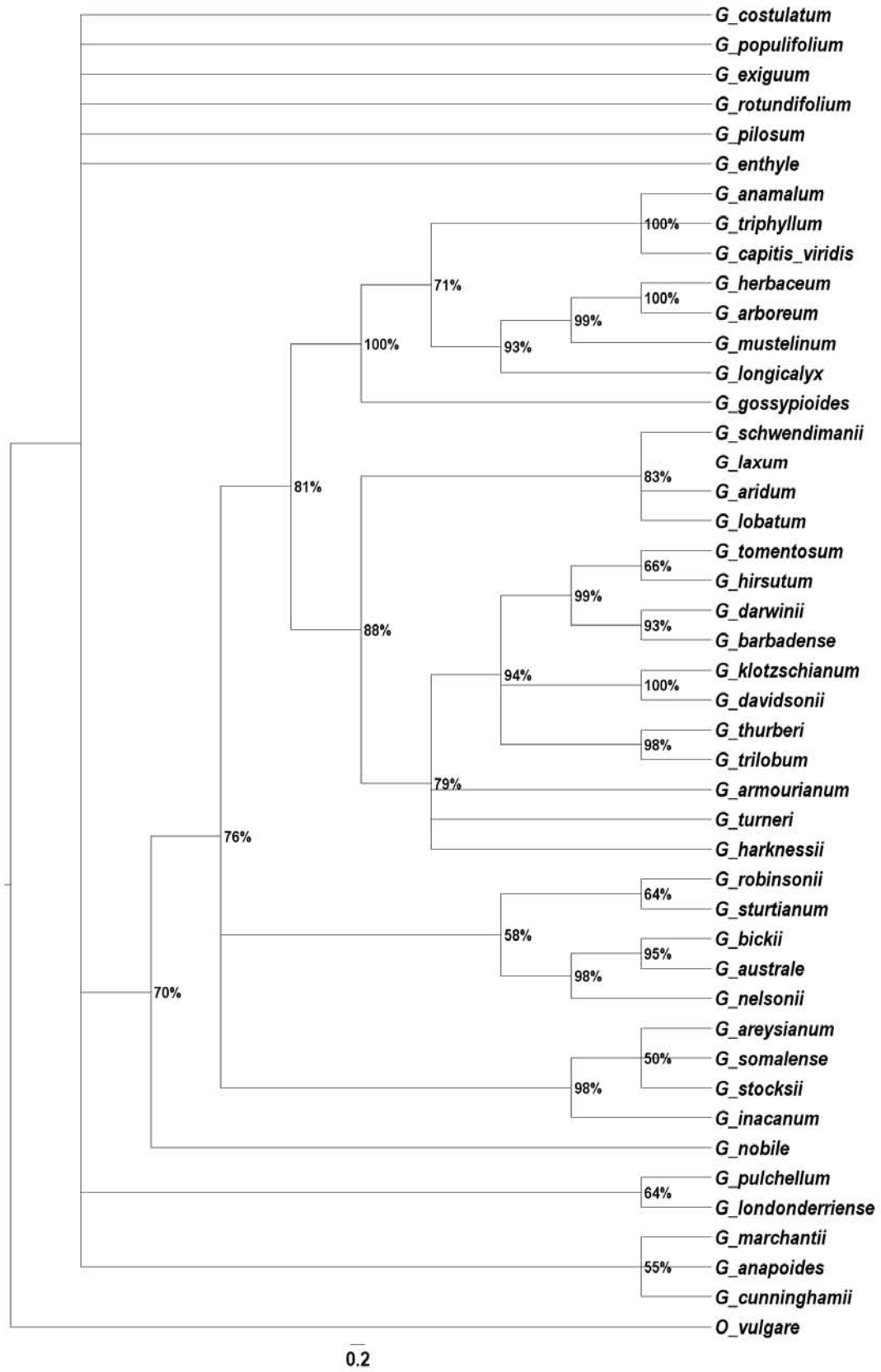
Şekil Ek-46. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



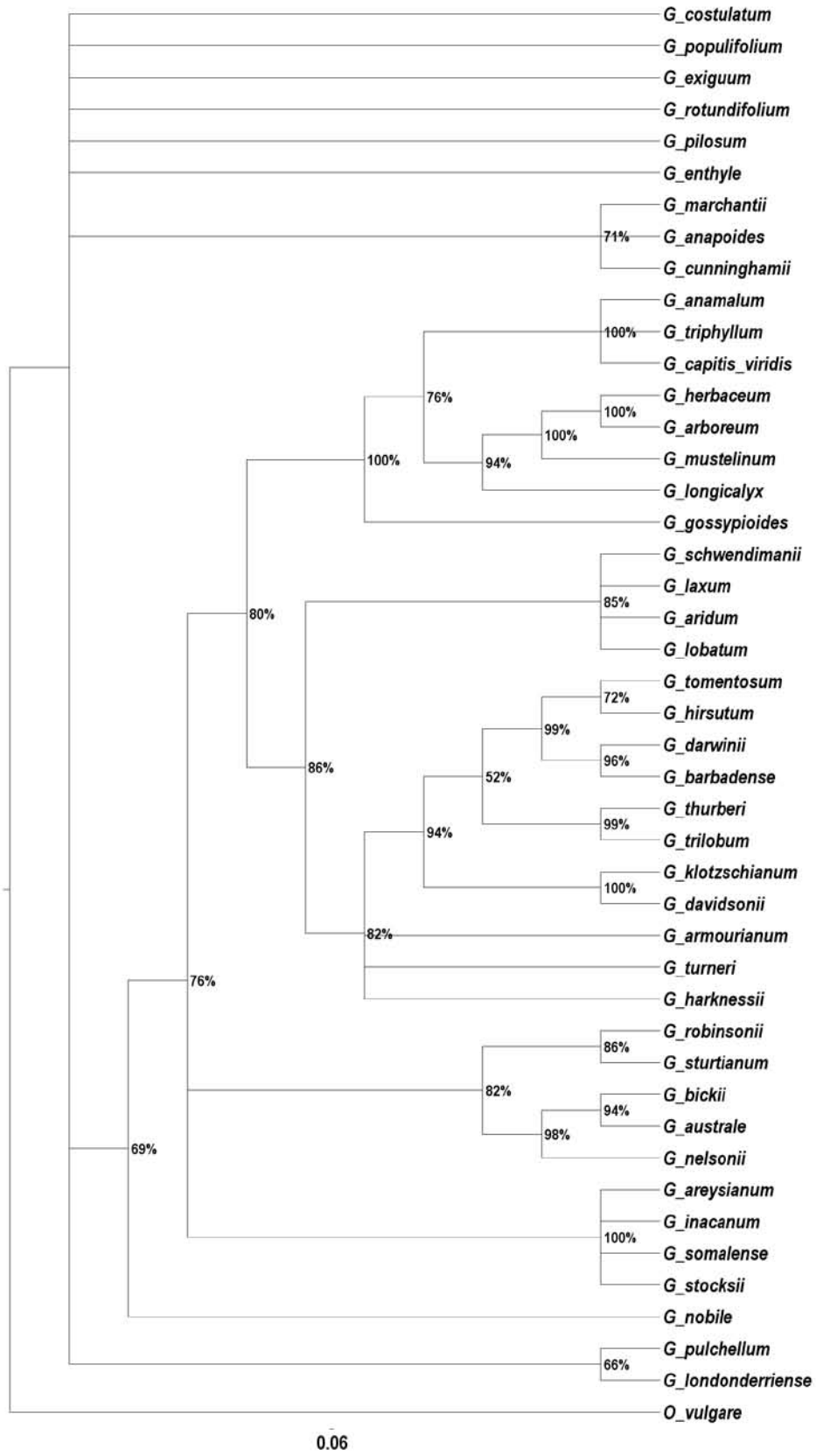
Şekil Ek-47. MAFFT programının 2.parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



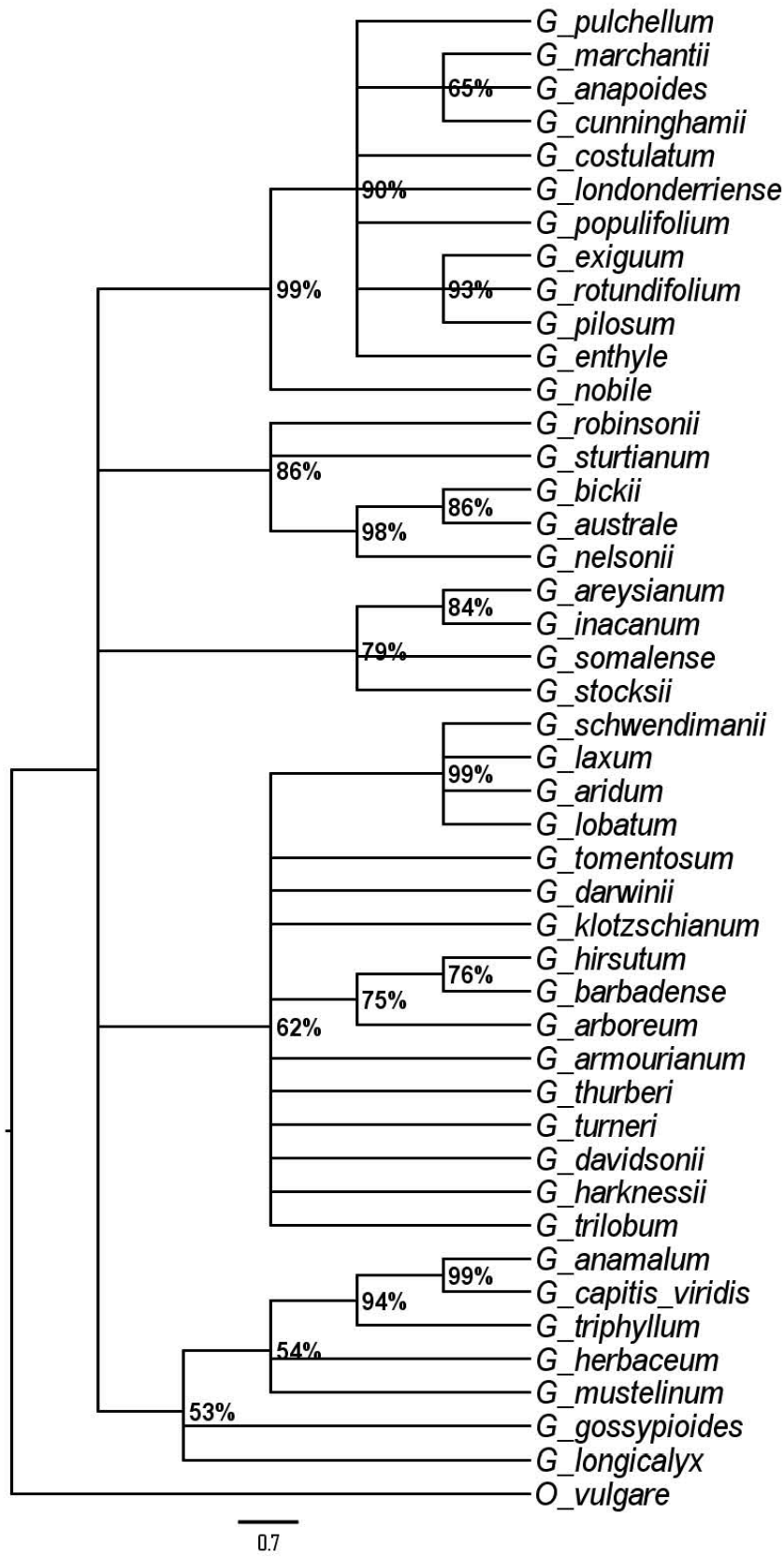
Şekil Ek-48. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



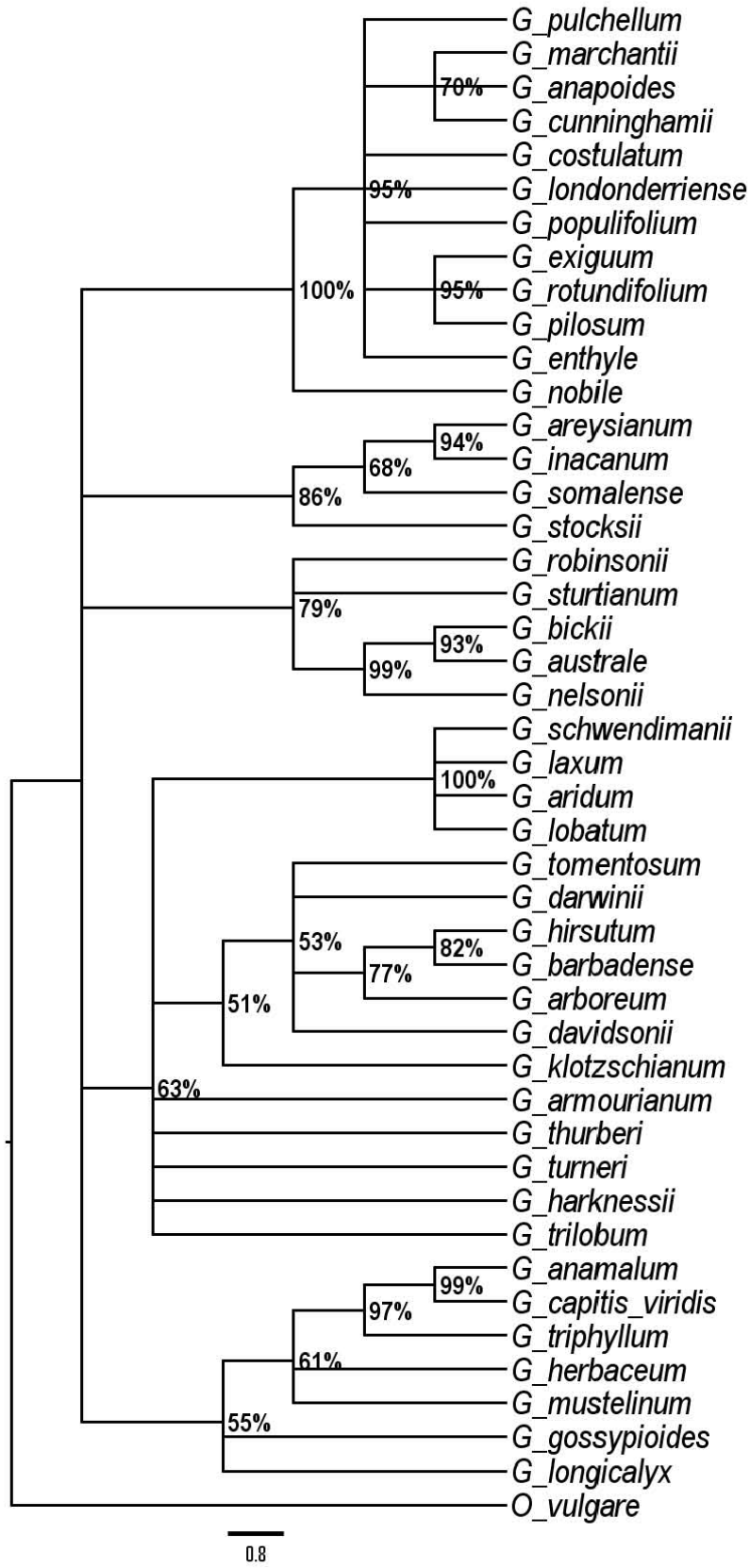
Şekil Ek-49. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



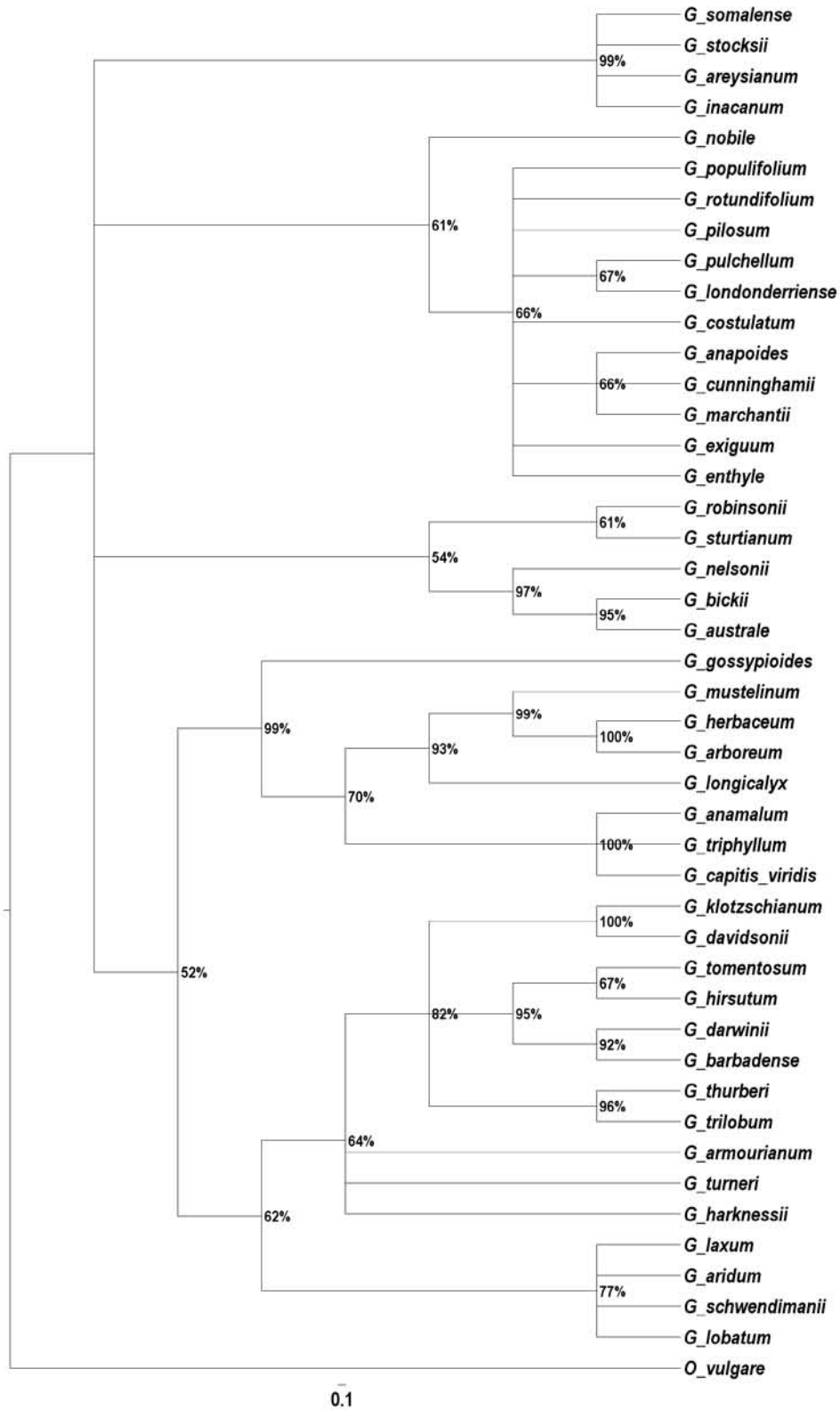
Şekil Ek-50. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



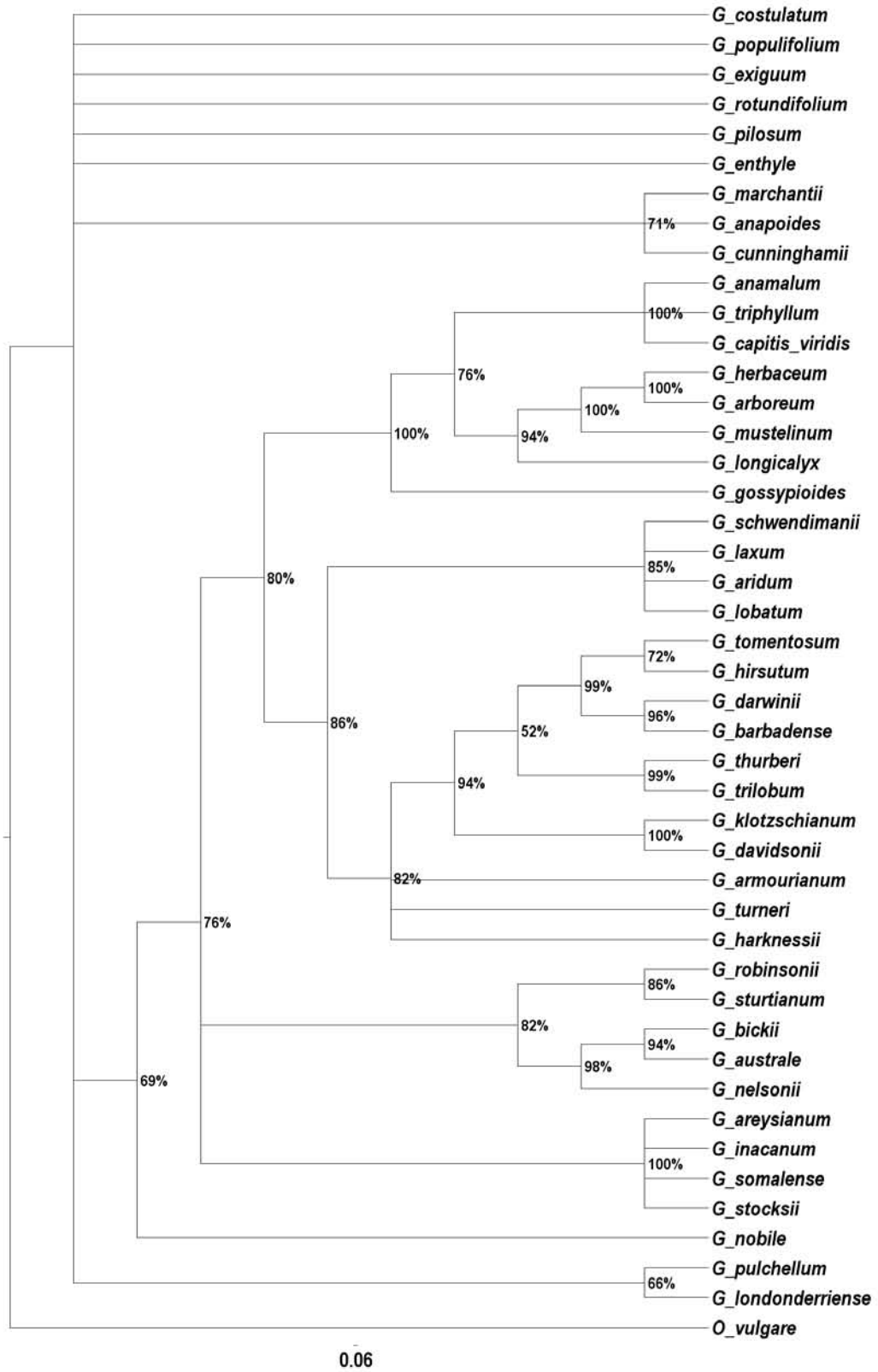
Şekil Ek-51. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



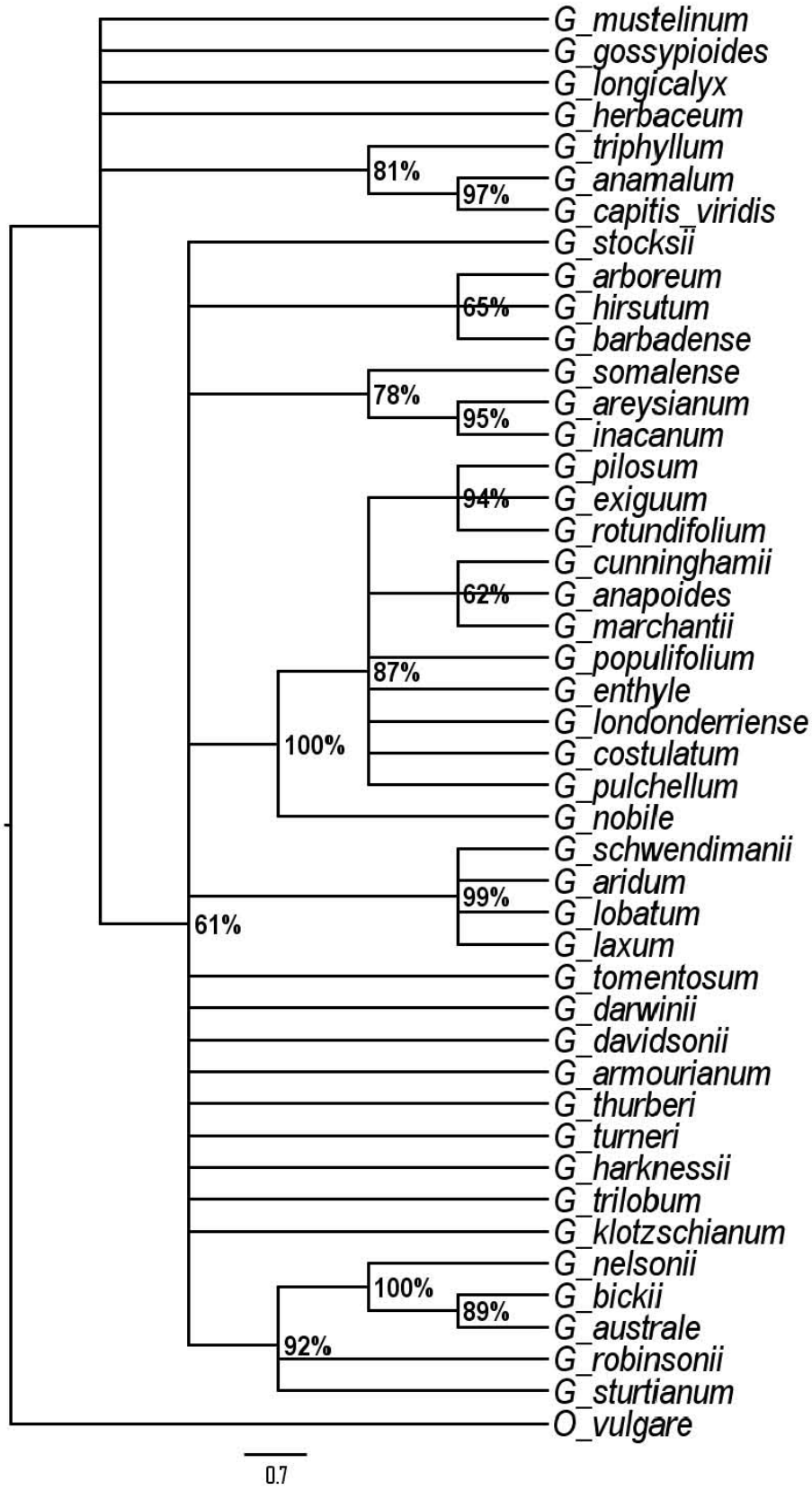
Şekil Ek-52. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



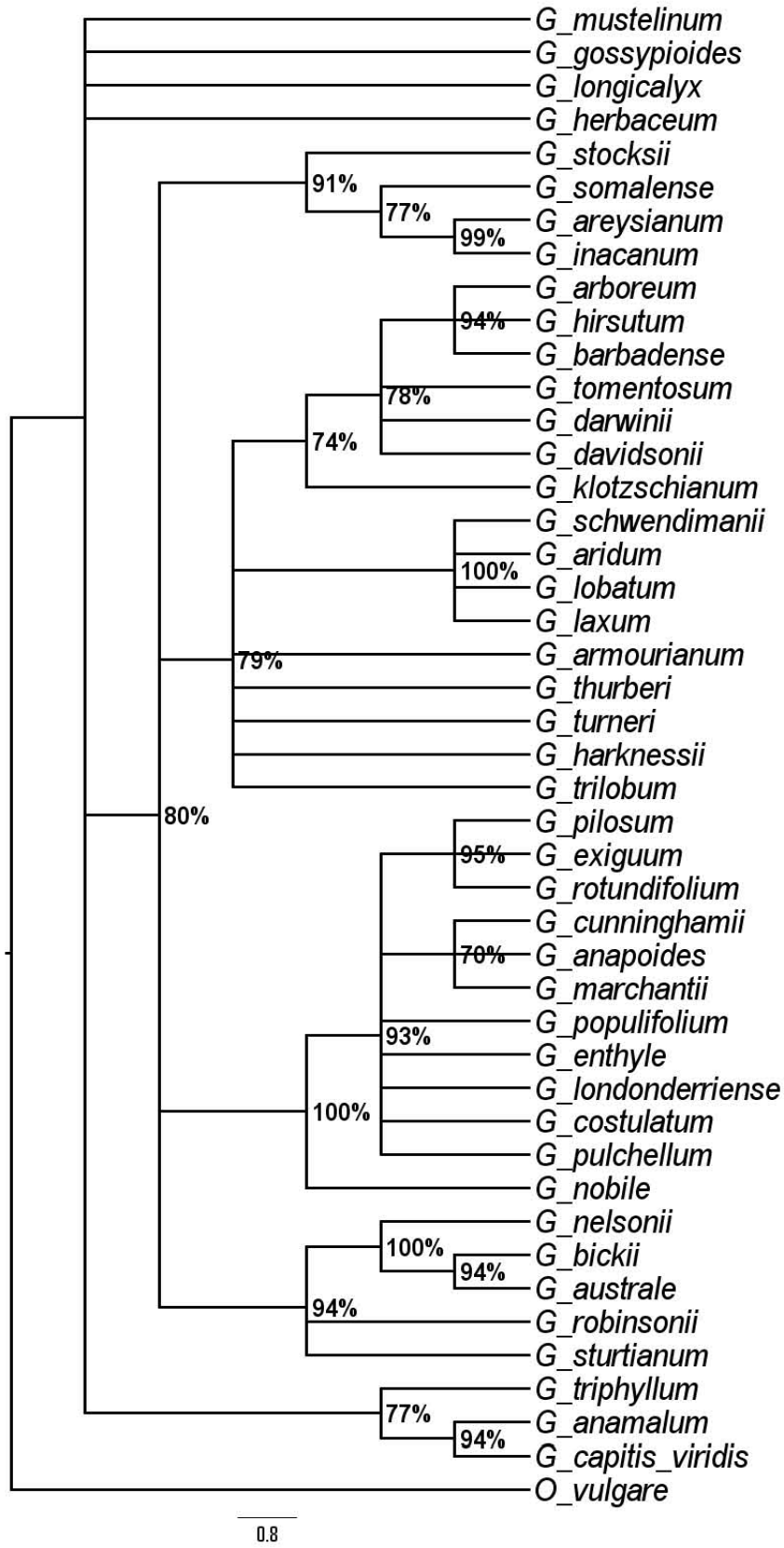
Şekil Ek-53. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



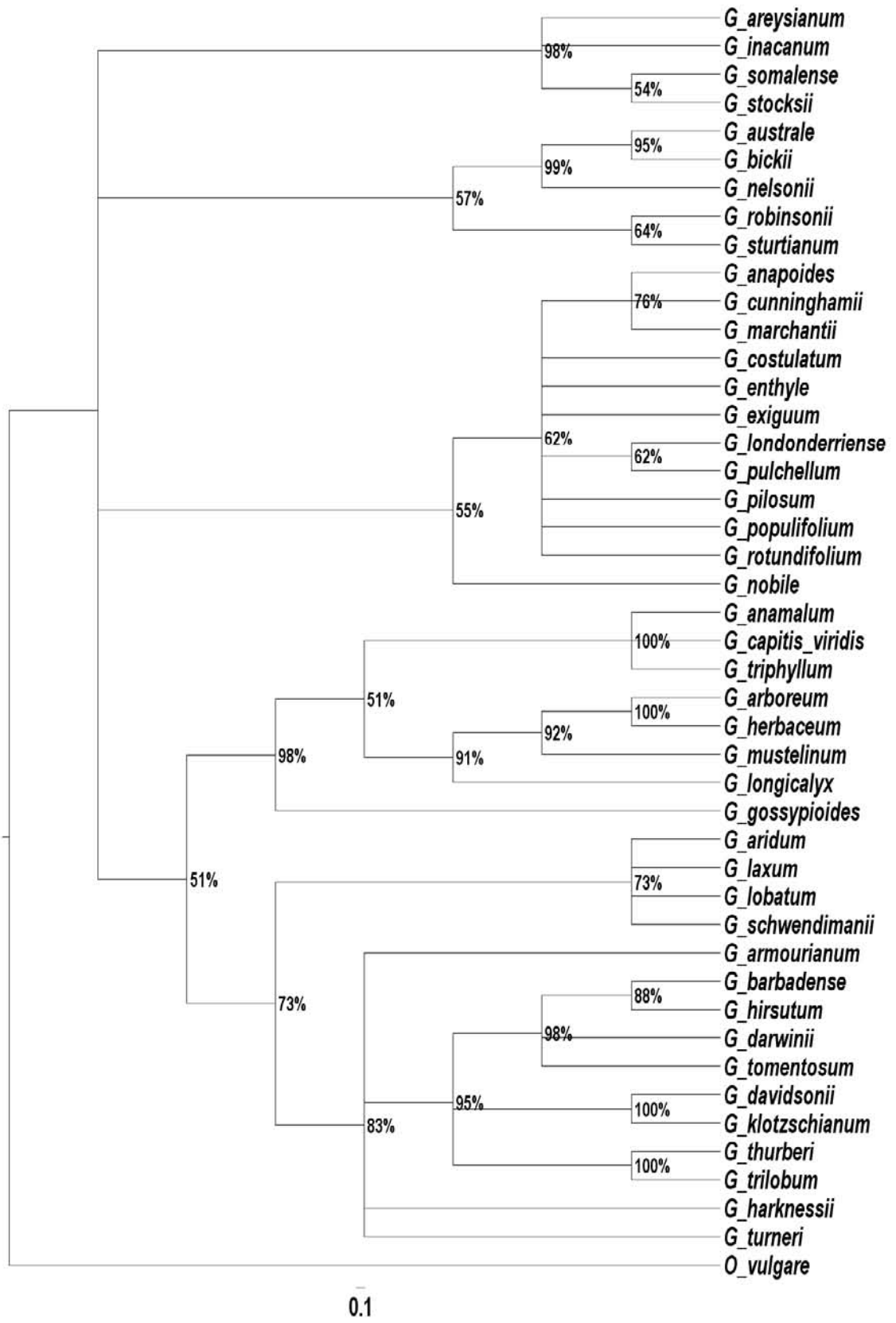
Şekil Ek-54. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



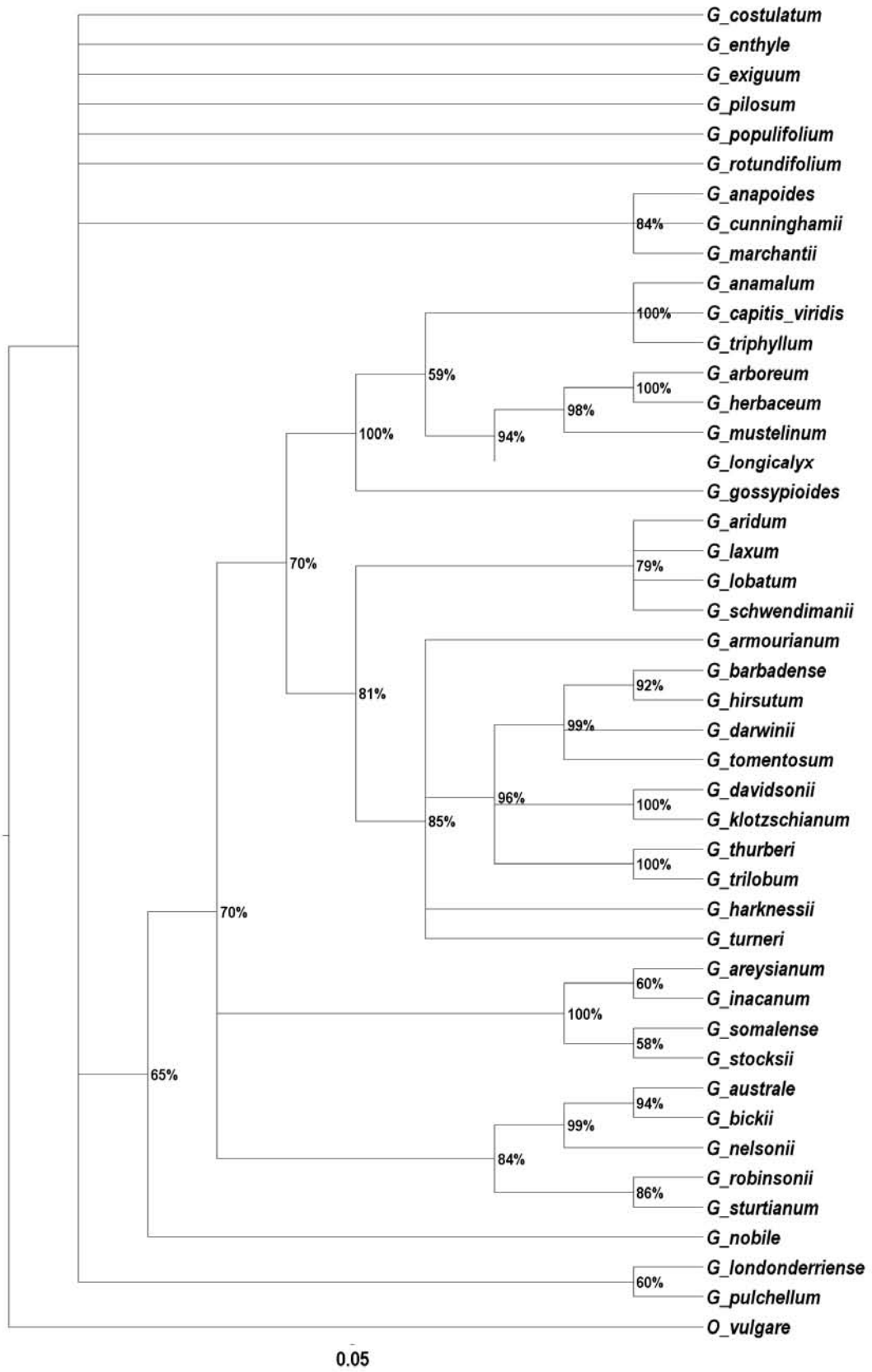
Şekil Ek-55. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



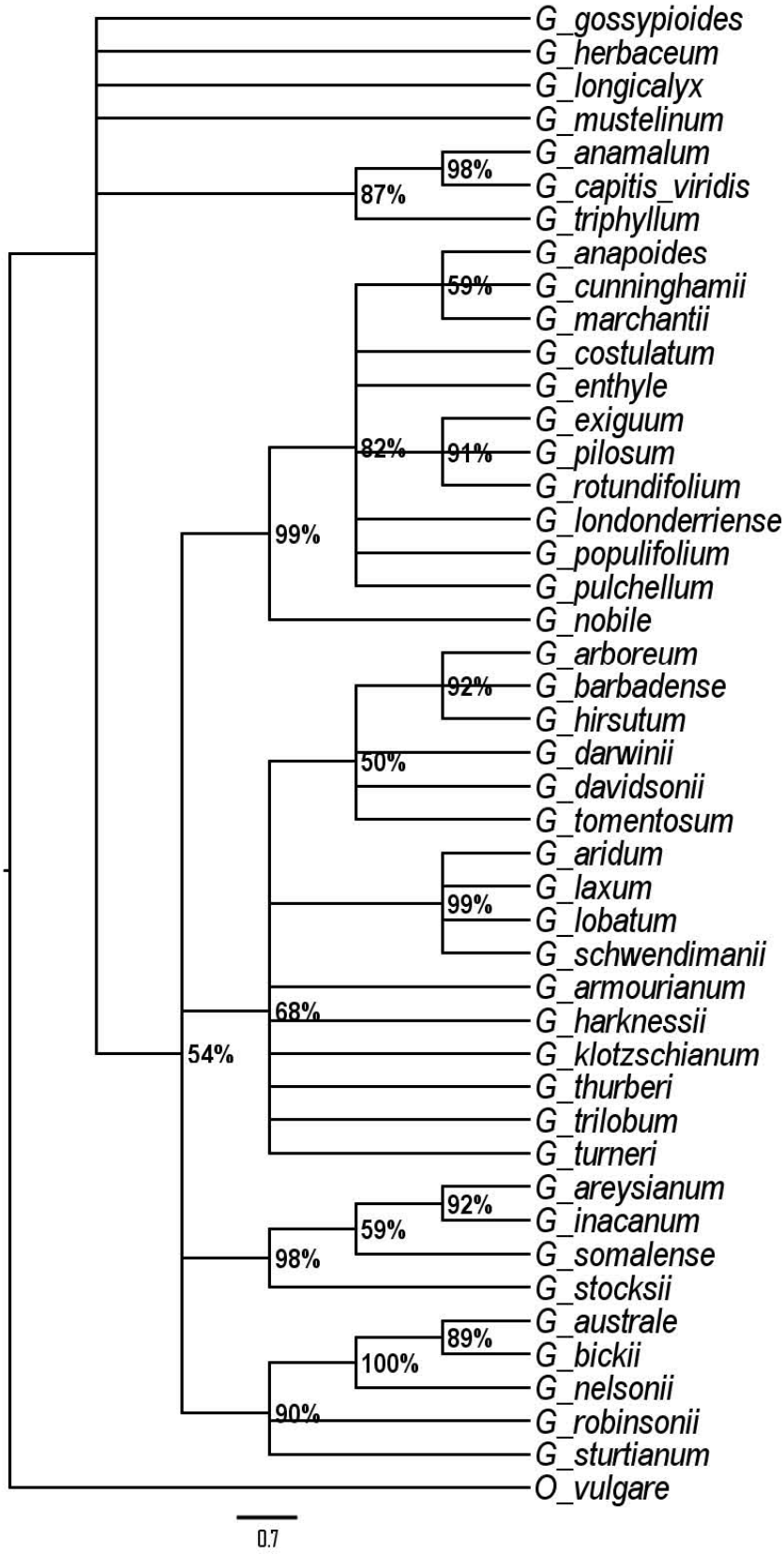
Şekil Ek-56. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



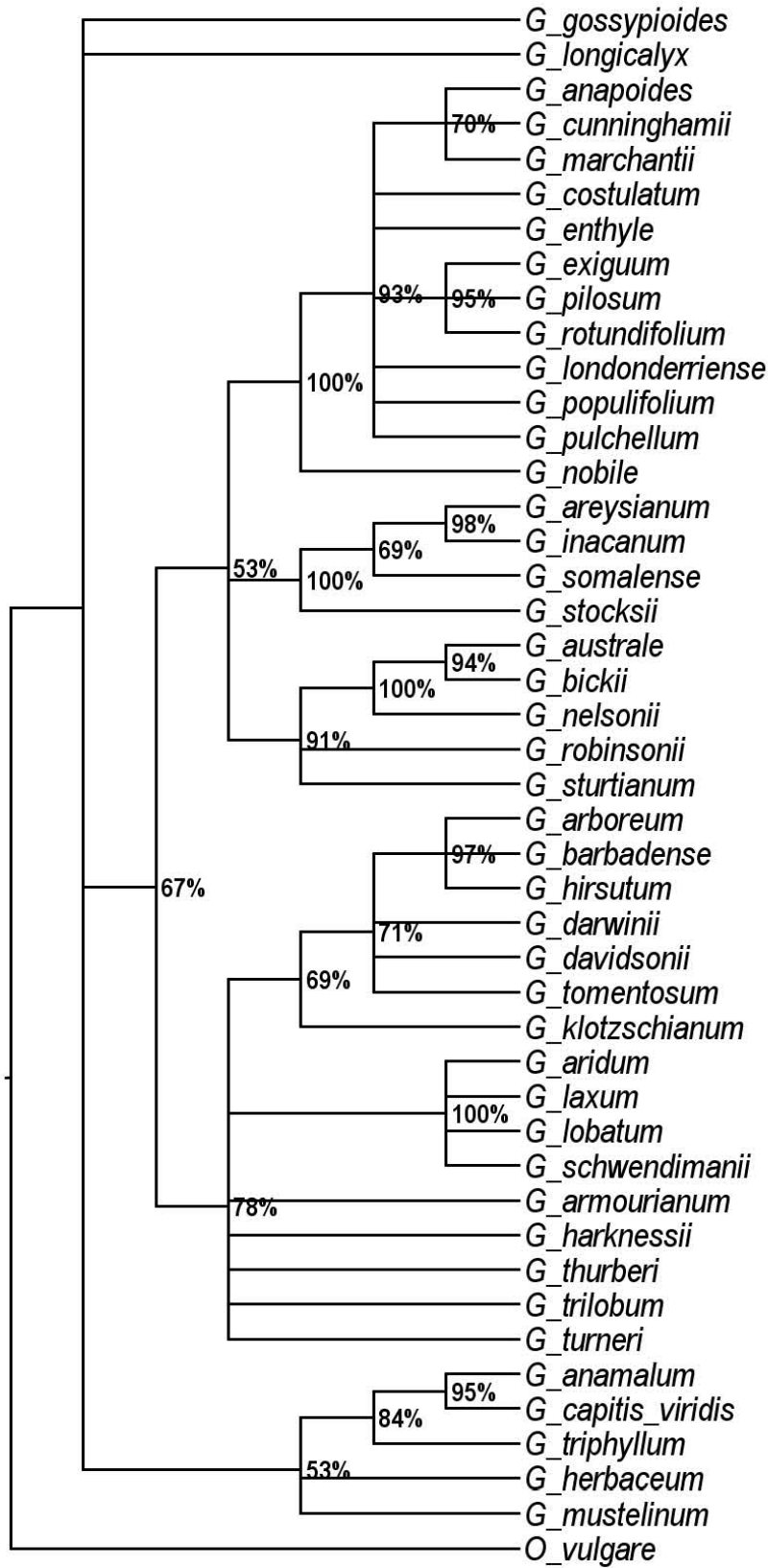
Şekil Ek-57. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



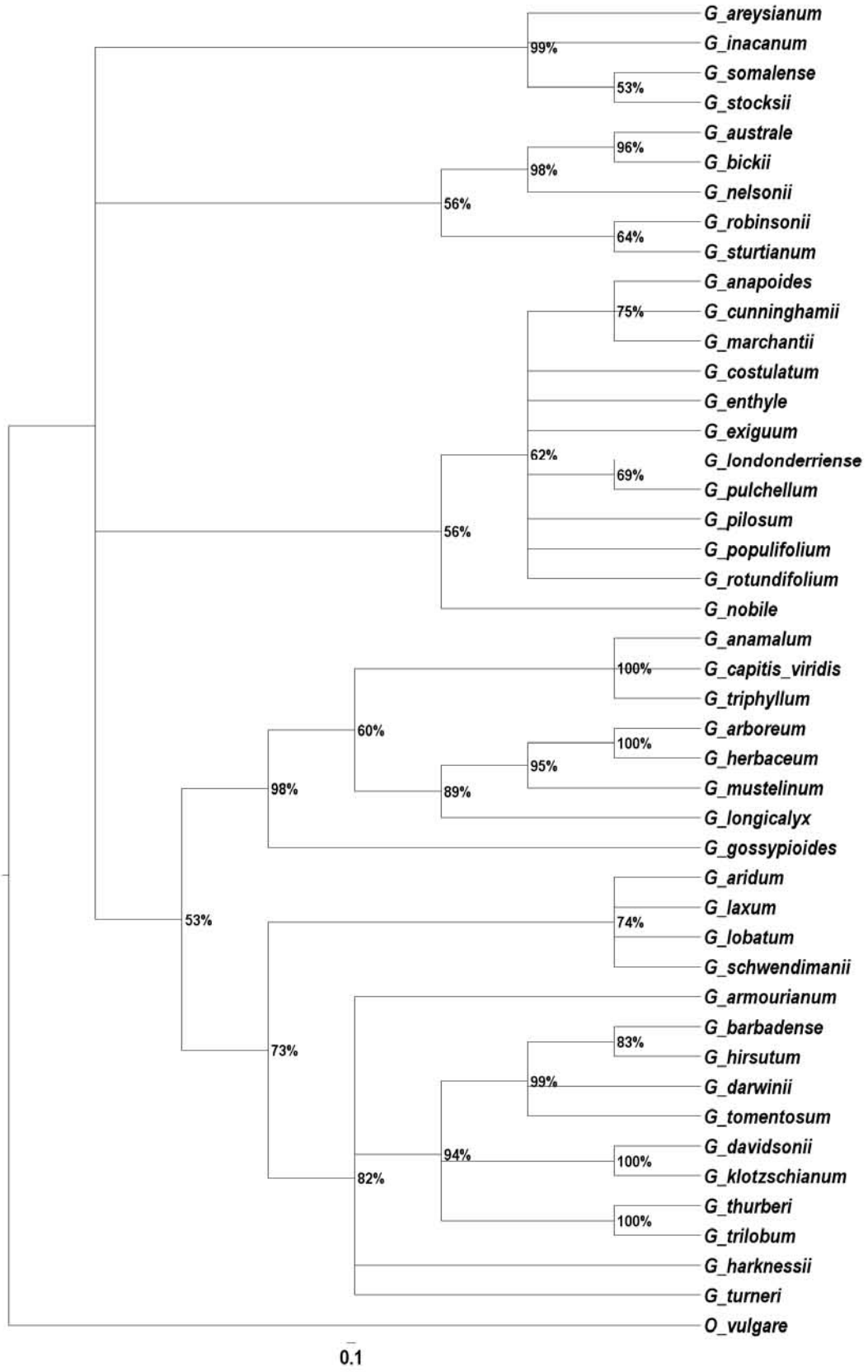
Şekil Ek-58. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



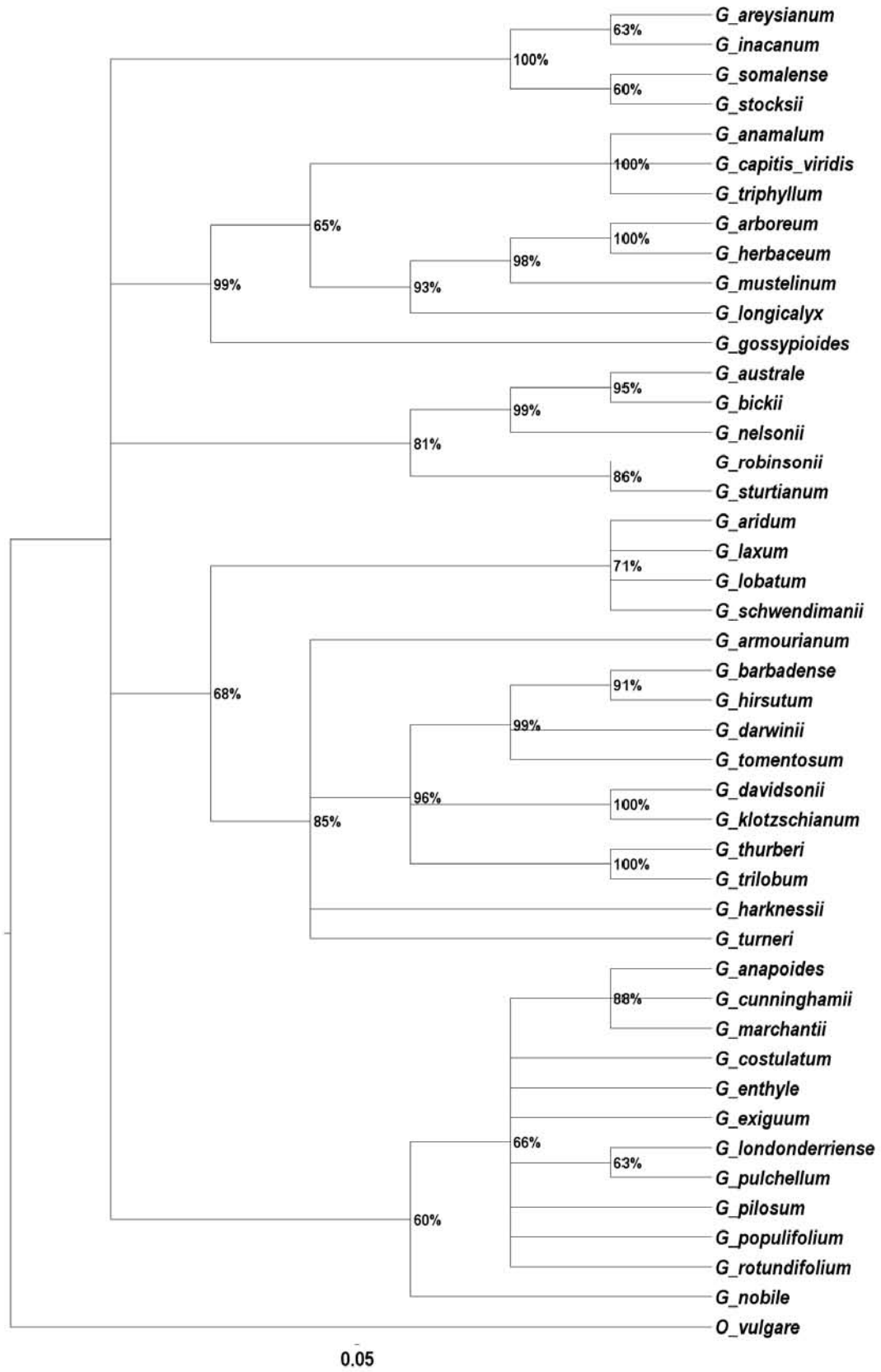
Şekil Ek-59. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



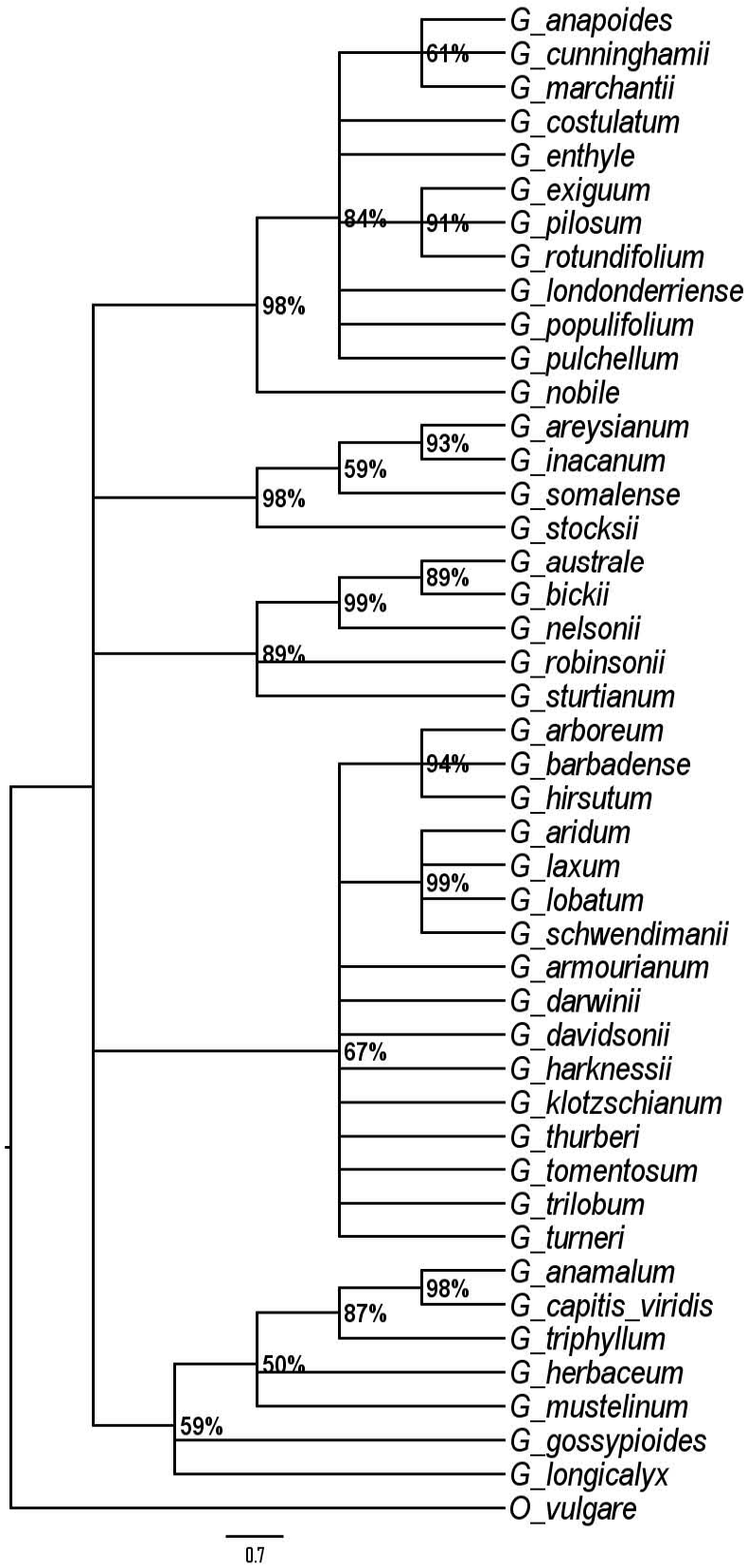
Şekil Ek-60. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



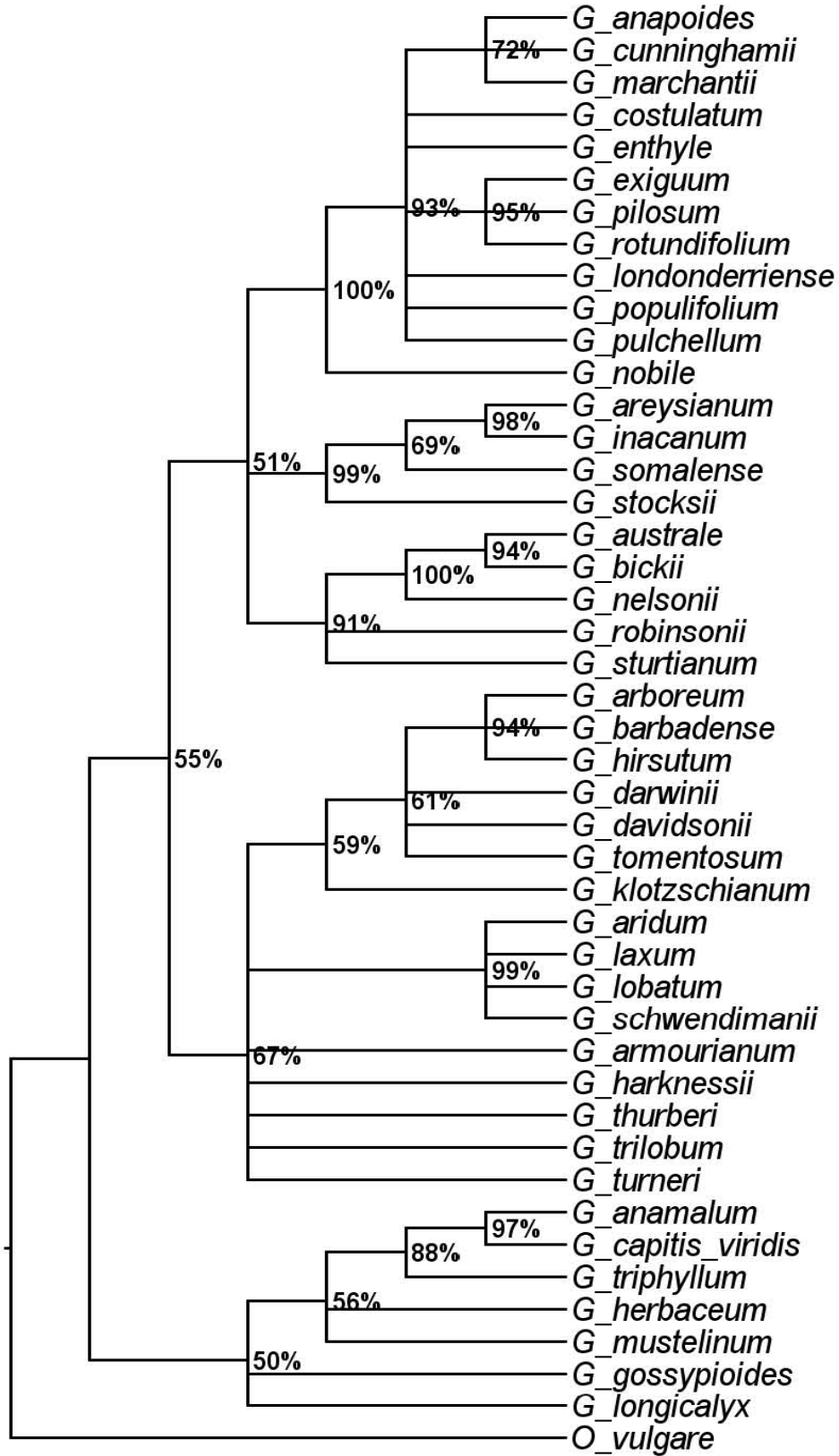
Şekil Ek-61. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki Bayes analizi



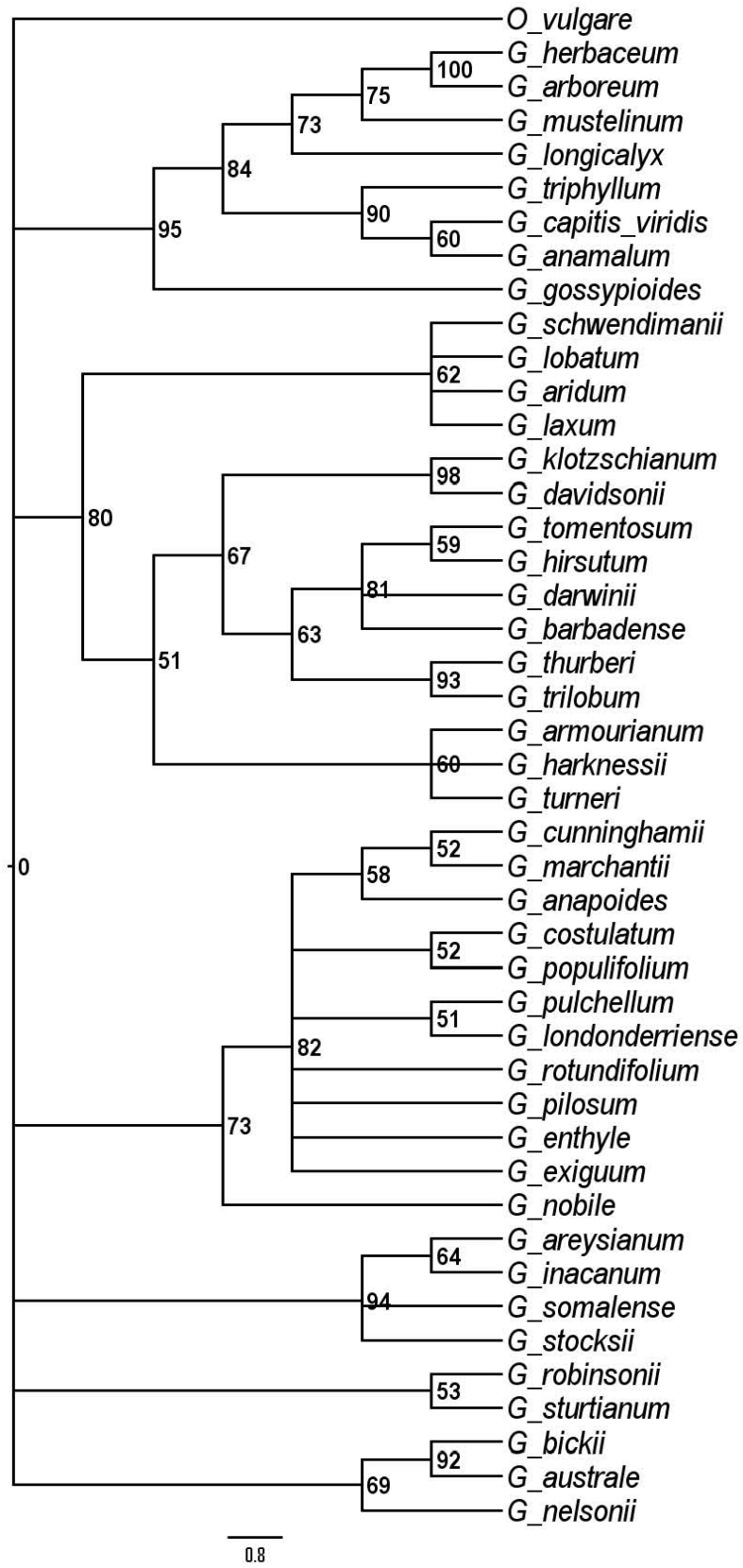
Şekil Ek-62. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki Bayes analizi



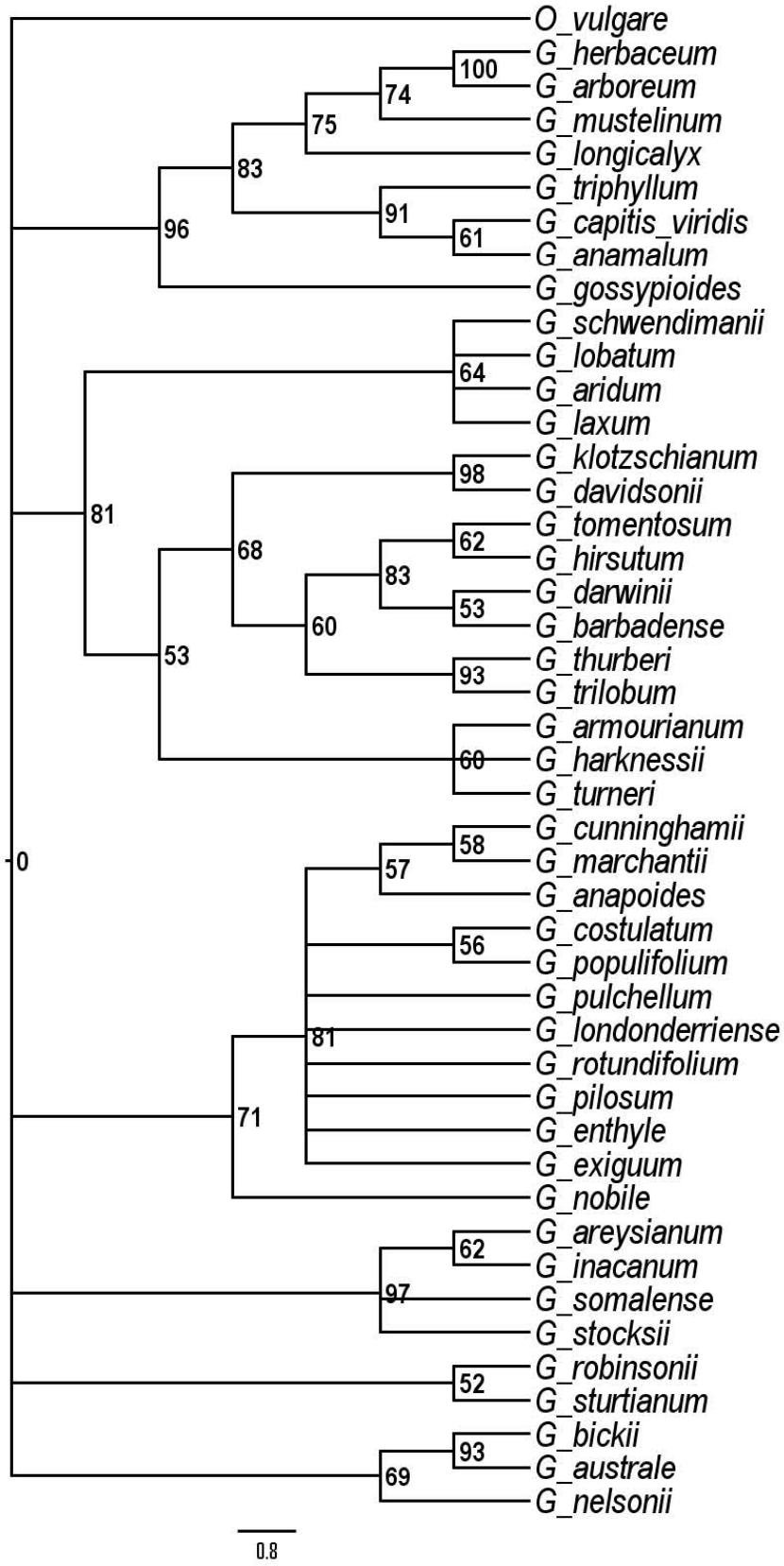
Şekil Ek-63. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki Bayes analizi



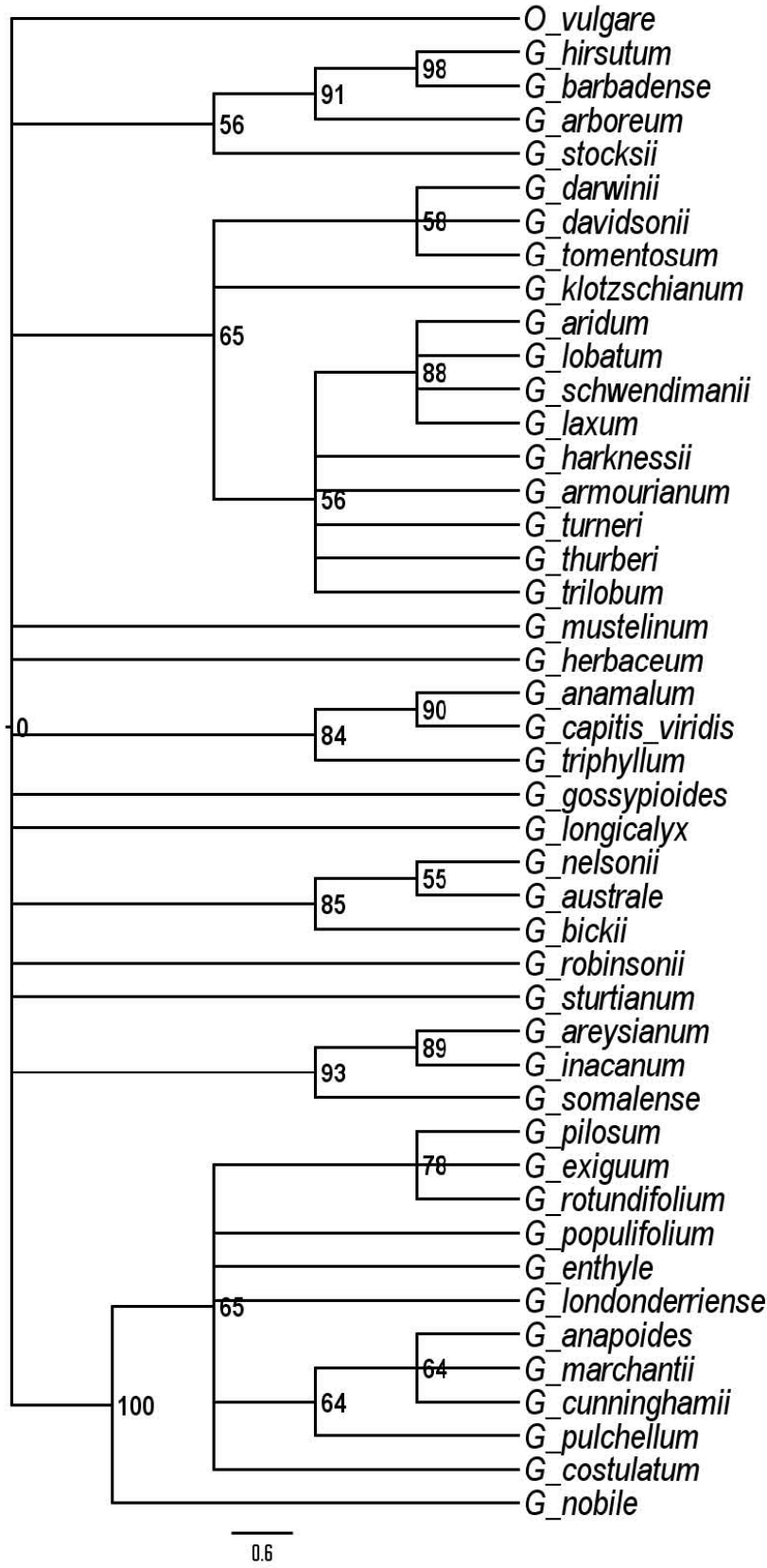
Şekil Ek-64. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki Bayes analizi



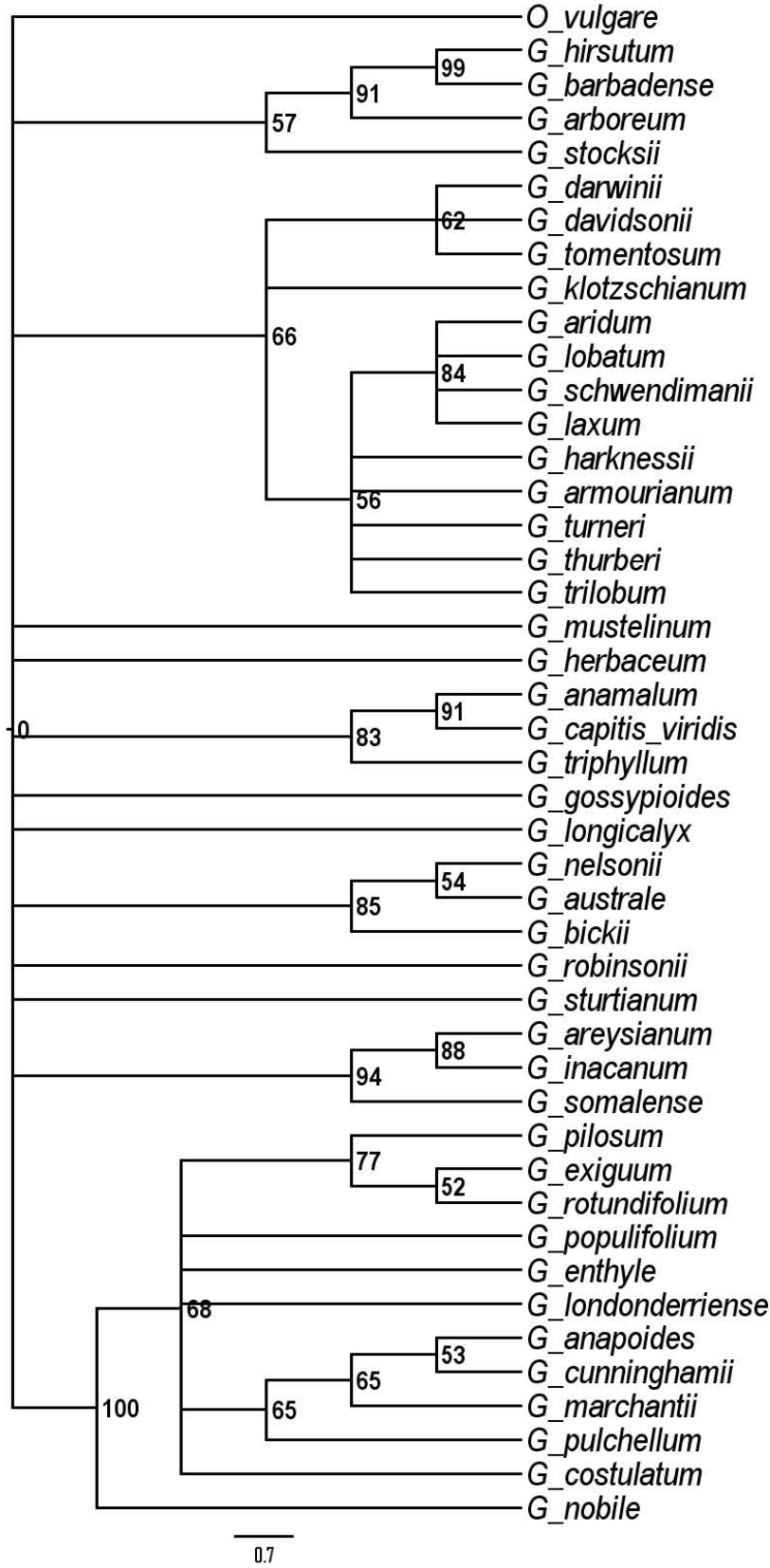
Şekil Ek-65. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi



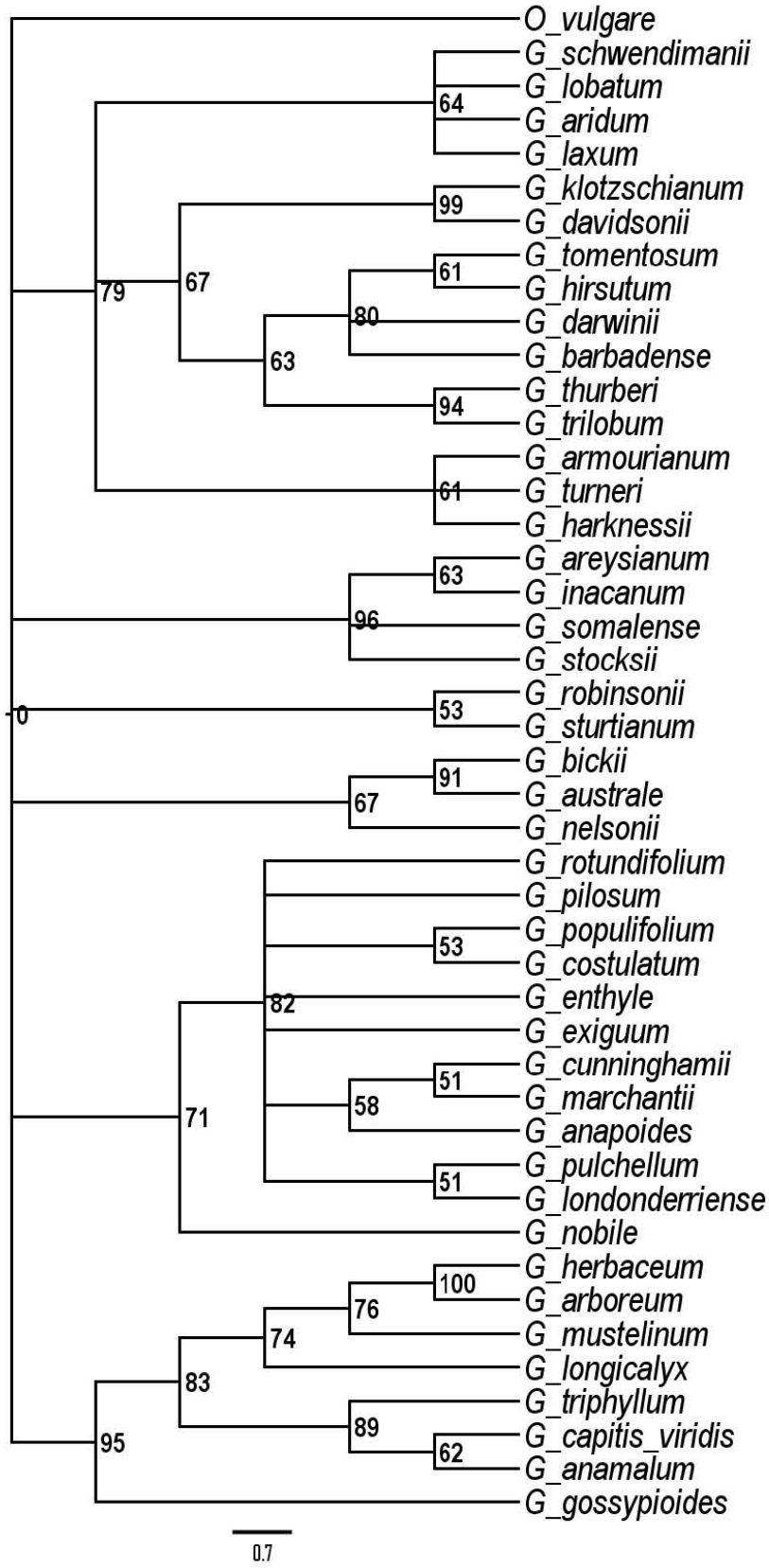
Şekil Ek-66. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi



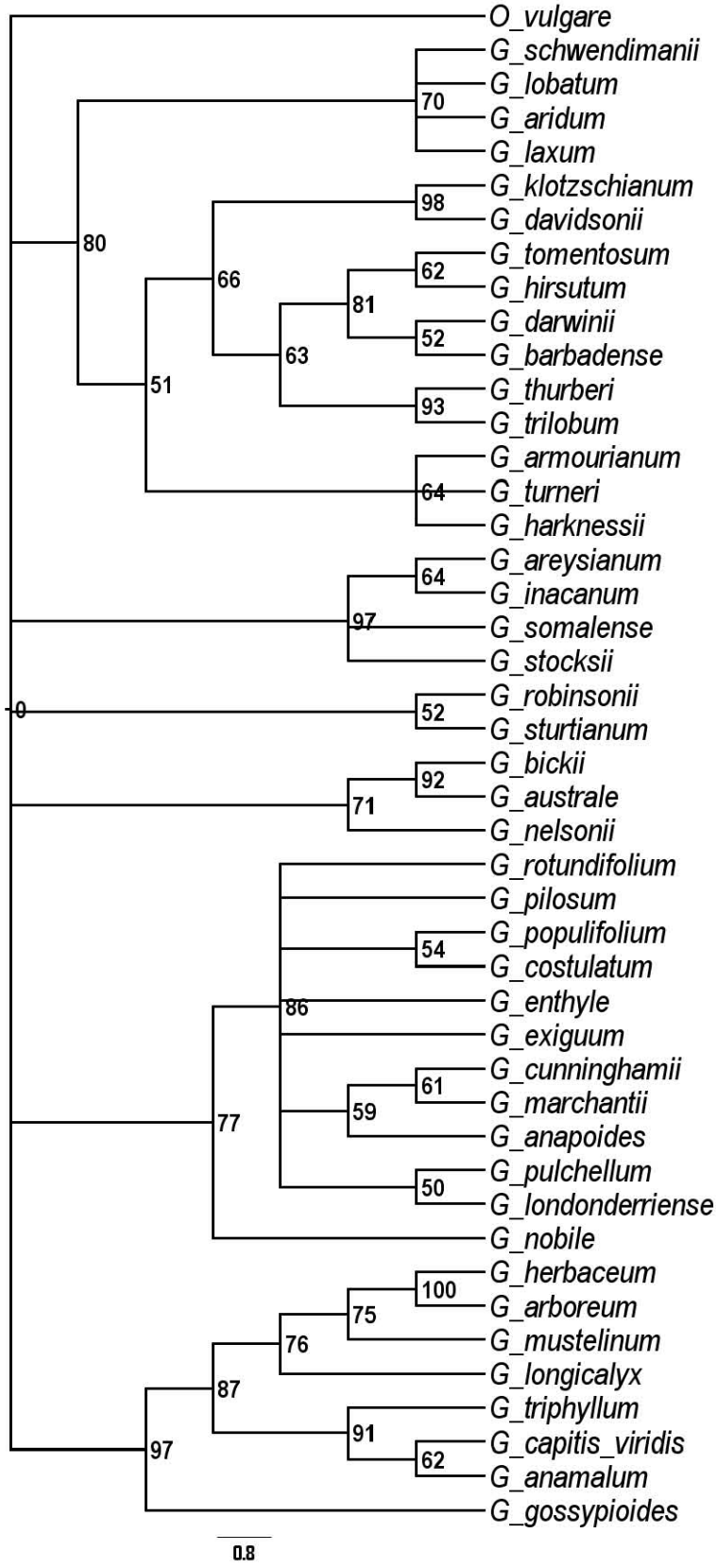
Şekil Ek-67. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi



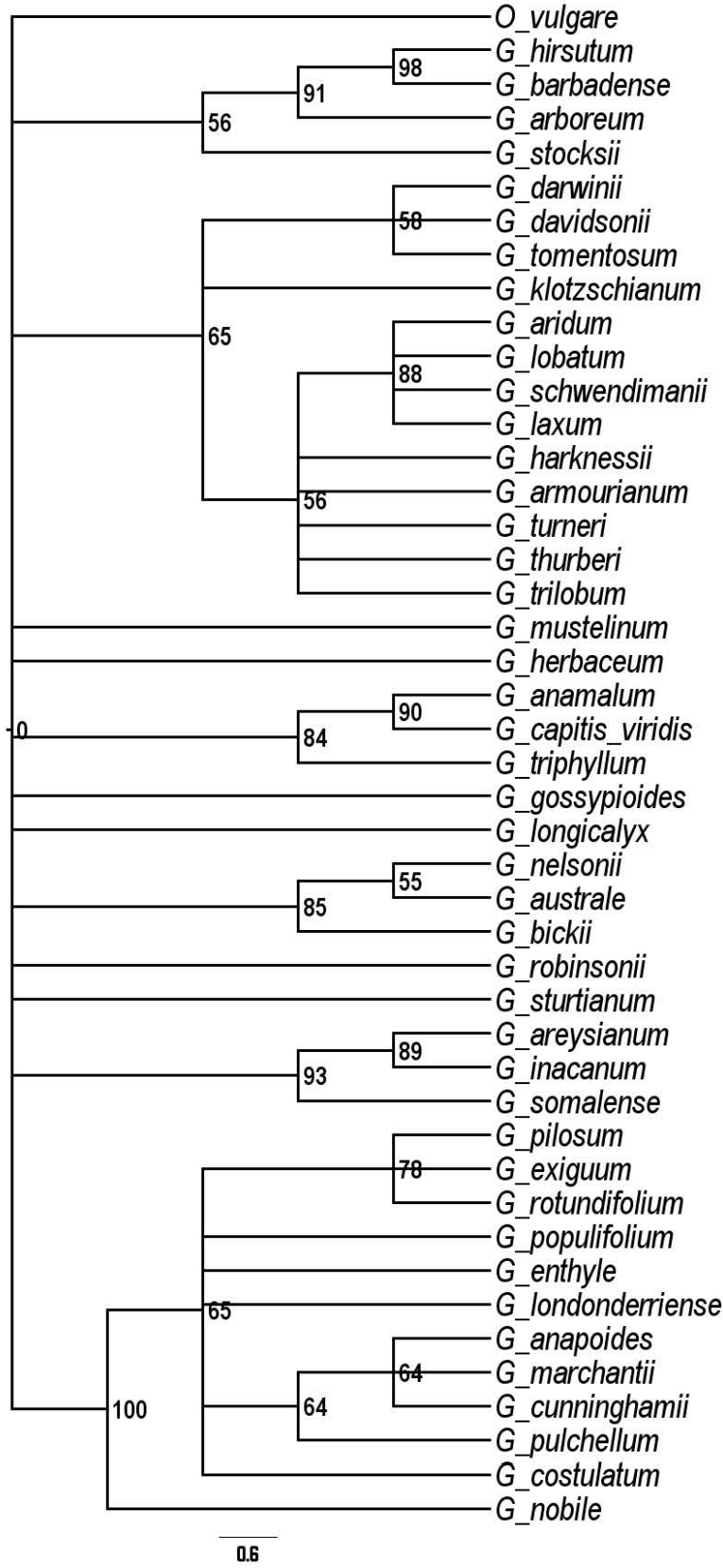
Şekil Ek-68. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi



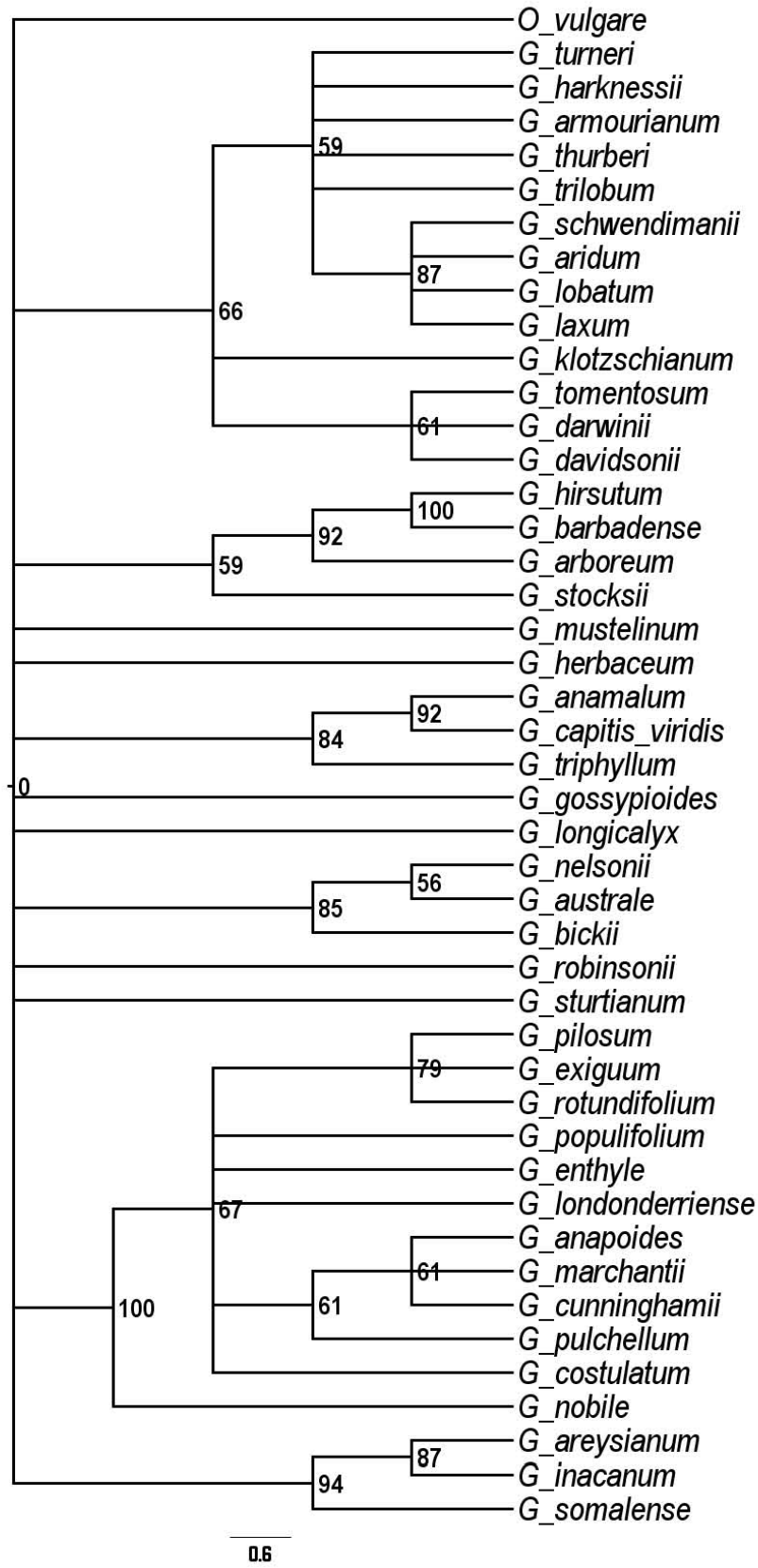
Şekil Ek-69. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi



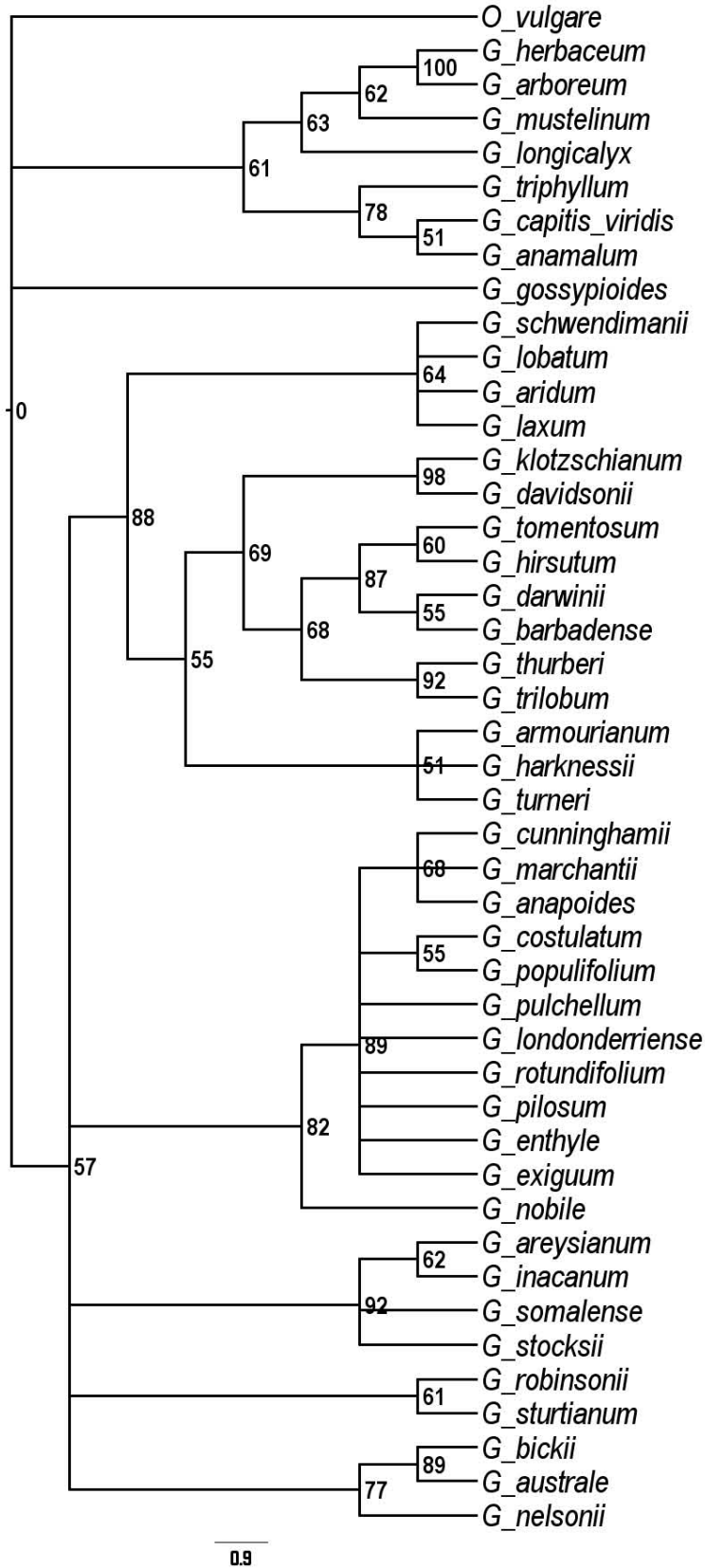
Şekil Ek-70. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi



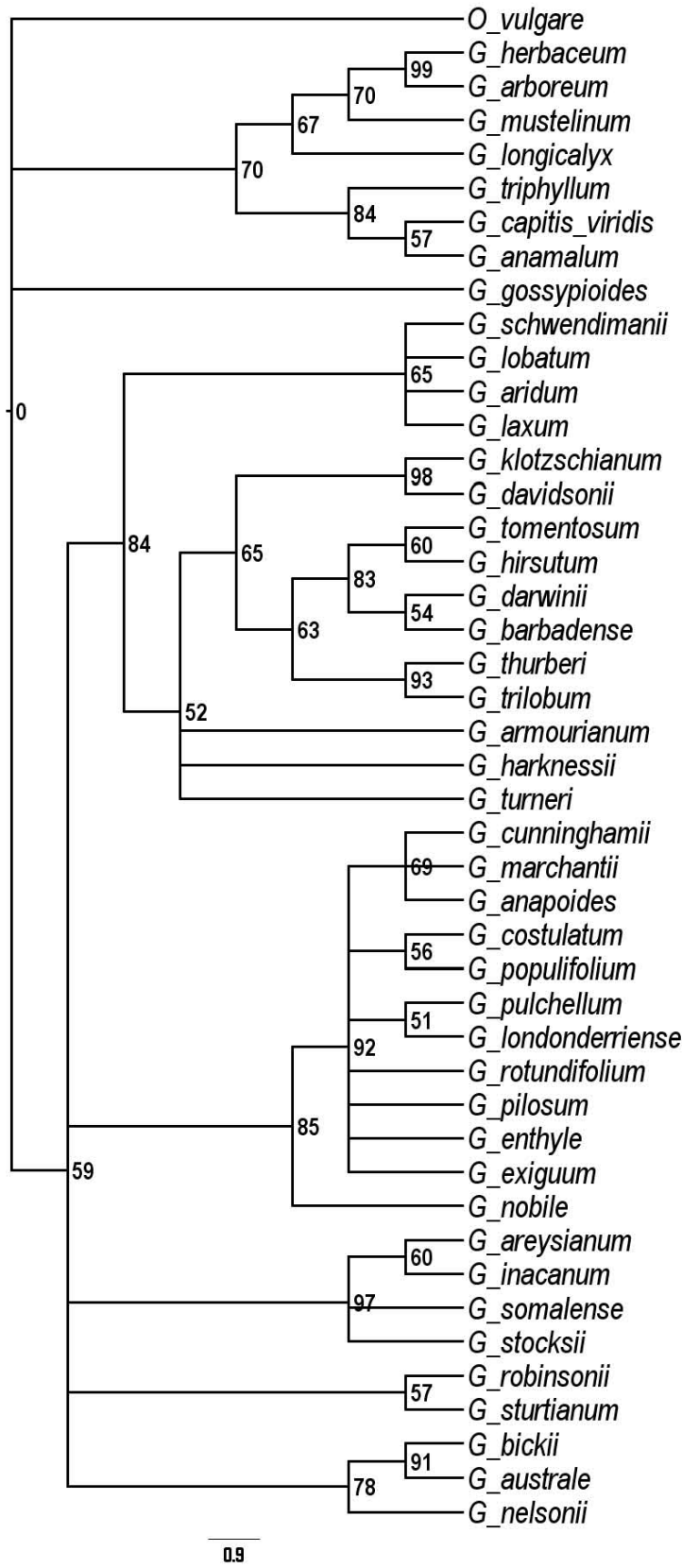
Şekil Ek-71. ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi



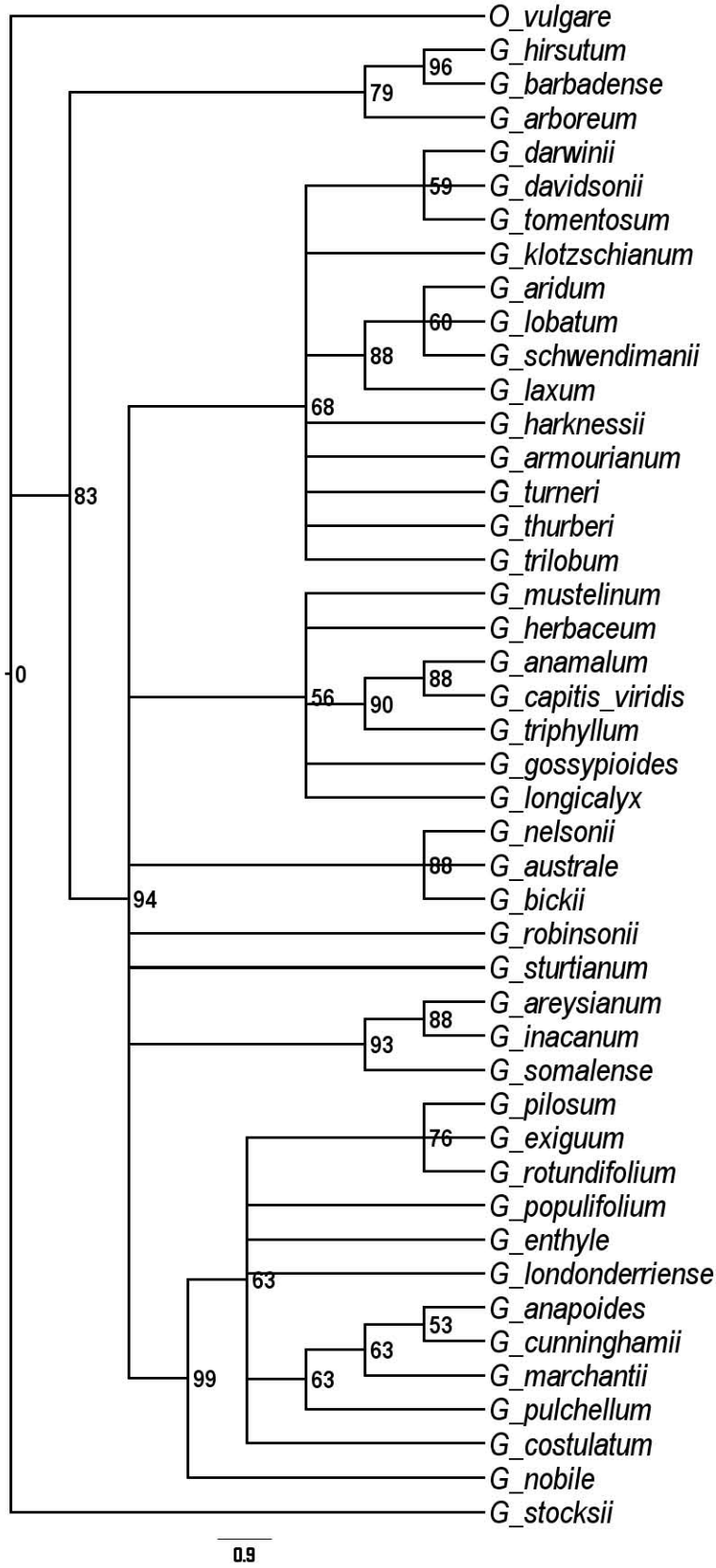
Şekil Ek-72. ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi



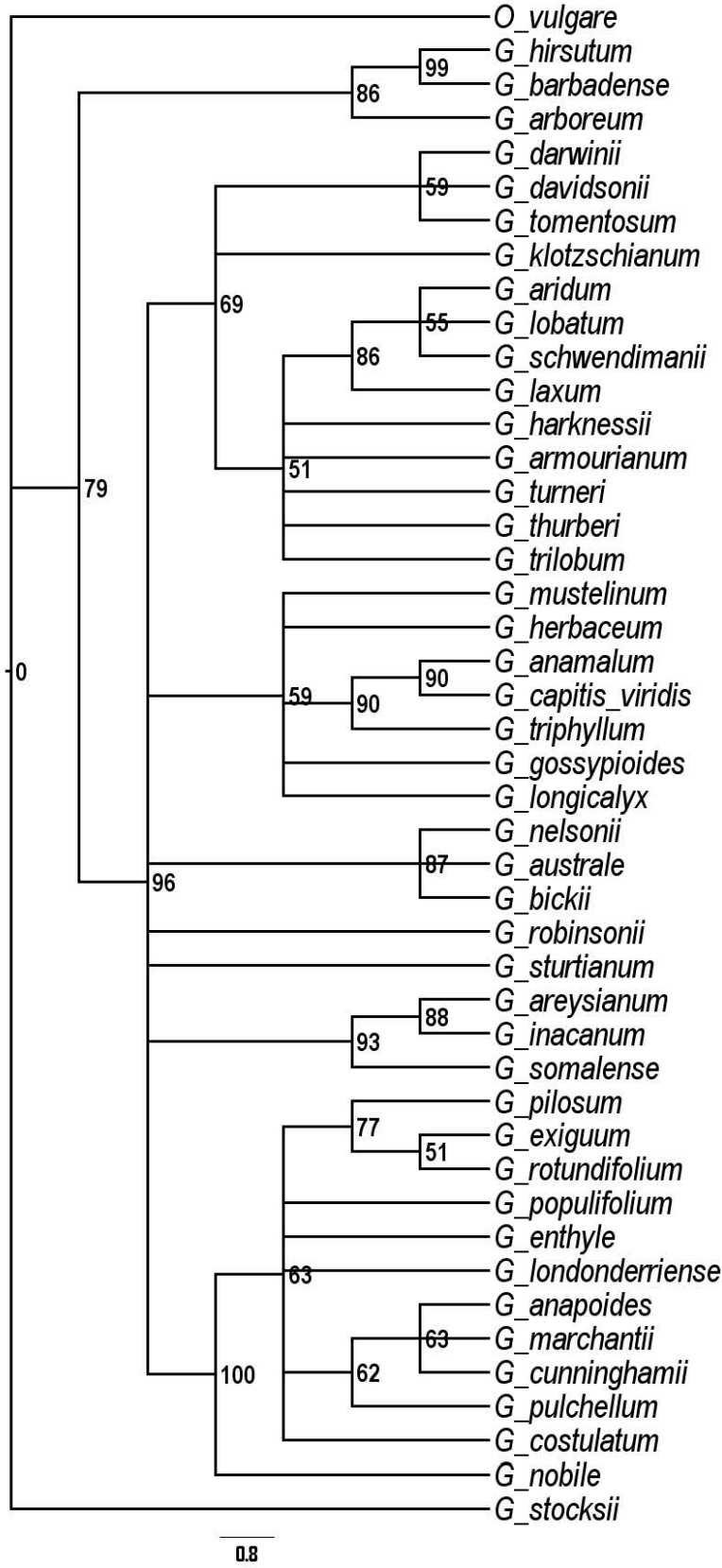
Şekil Ek-73. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi



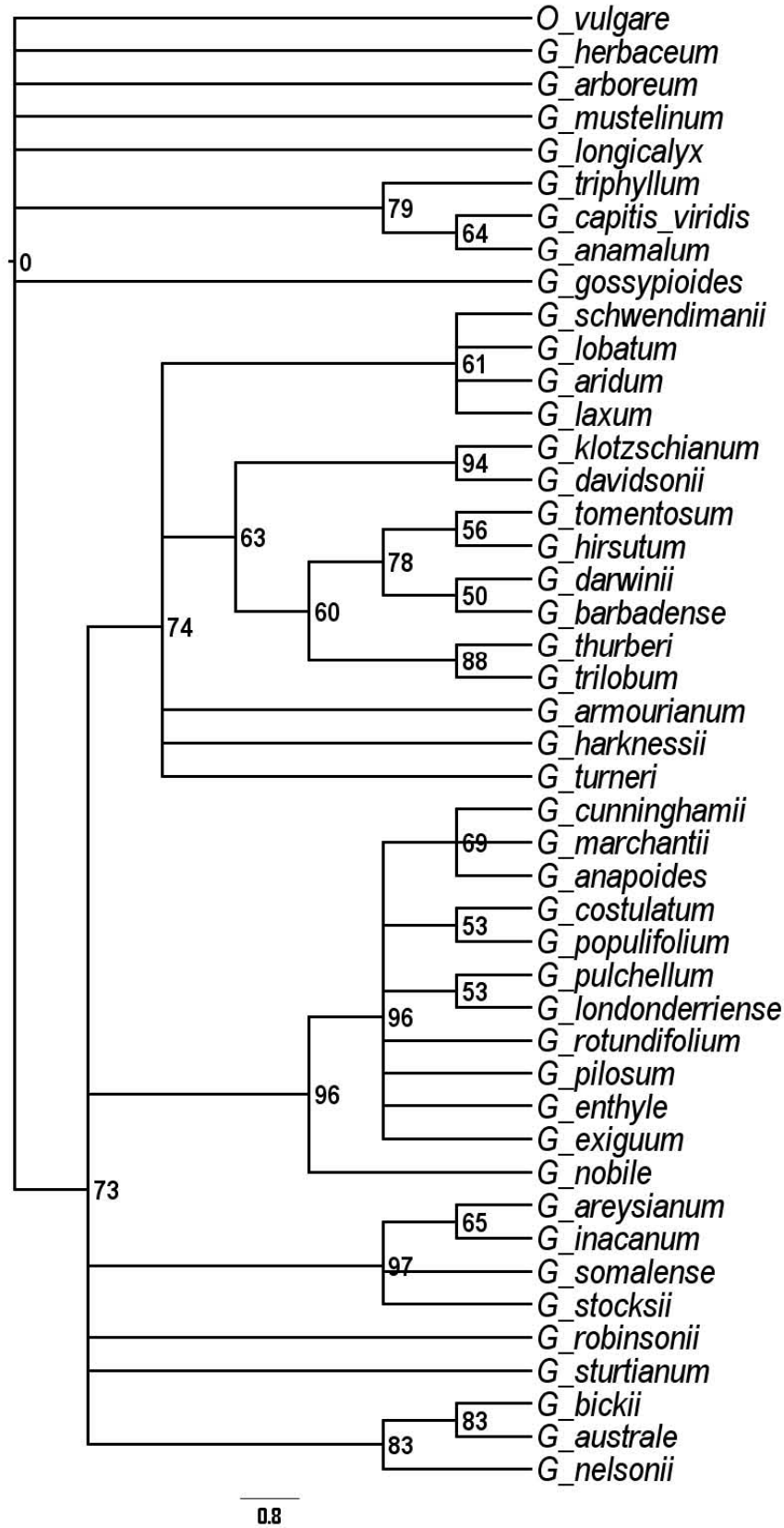
Şekil Ek-74. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi



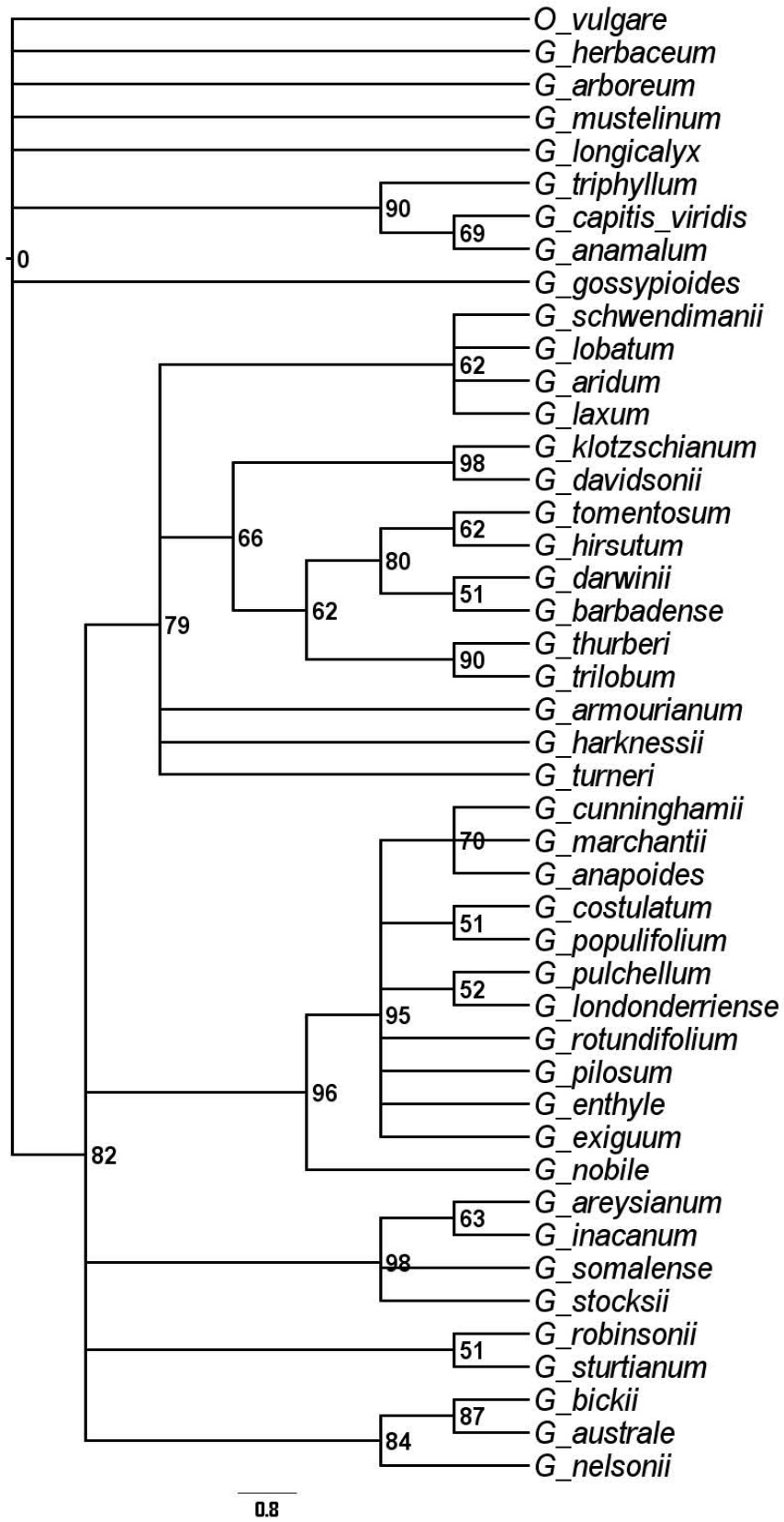
Şekil Ek-75. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi



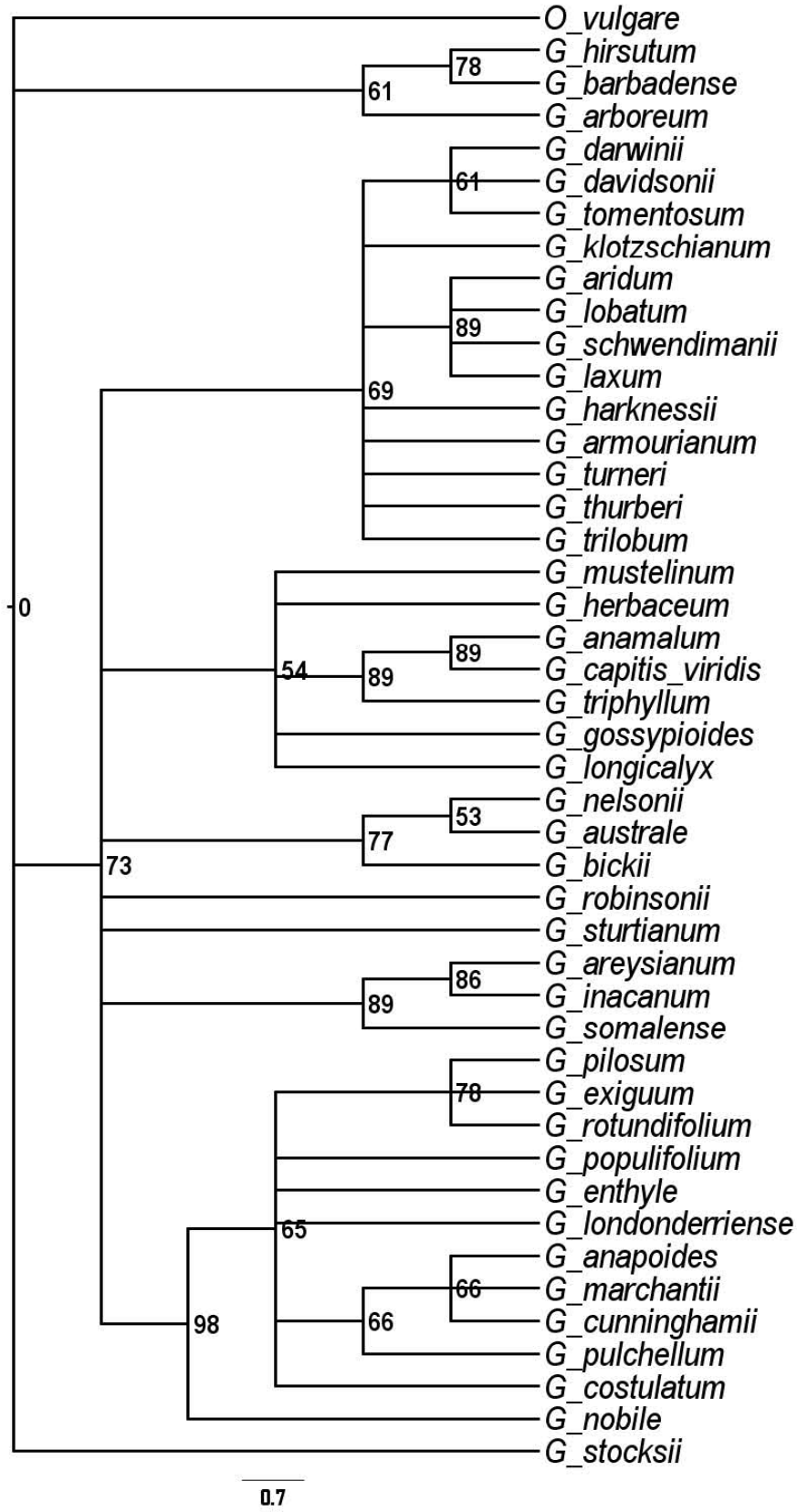
Şekil Ek-76. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi



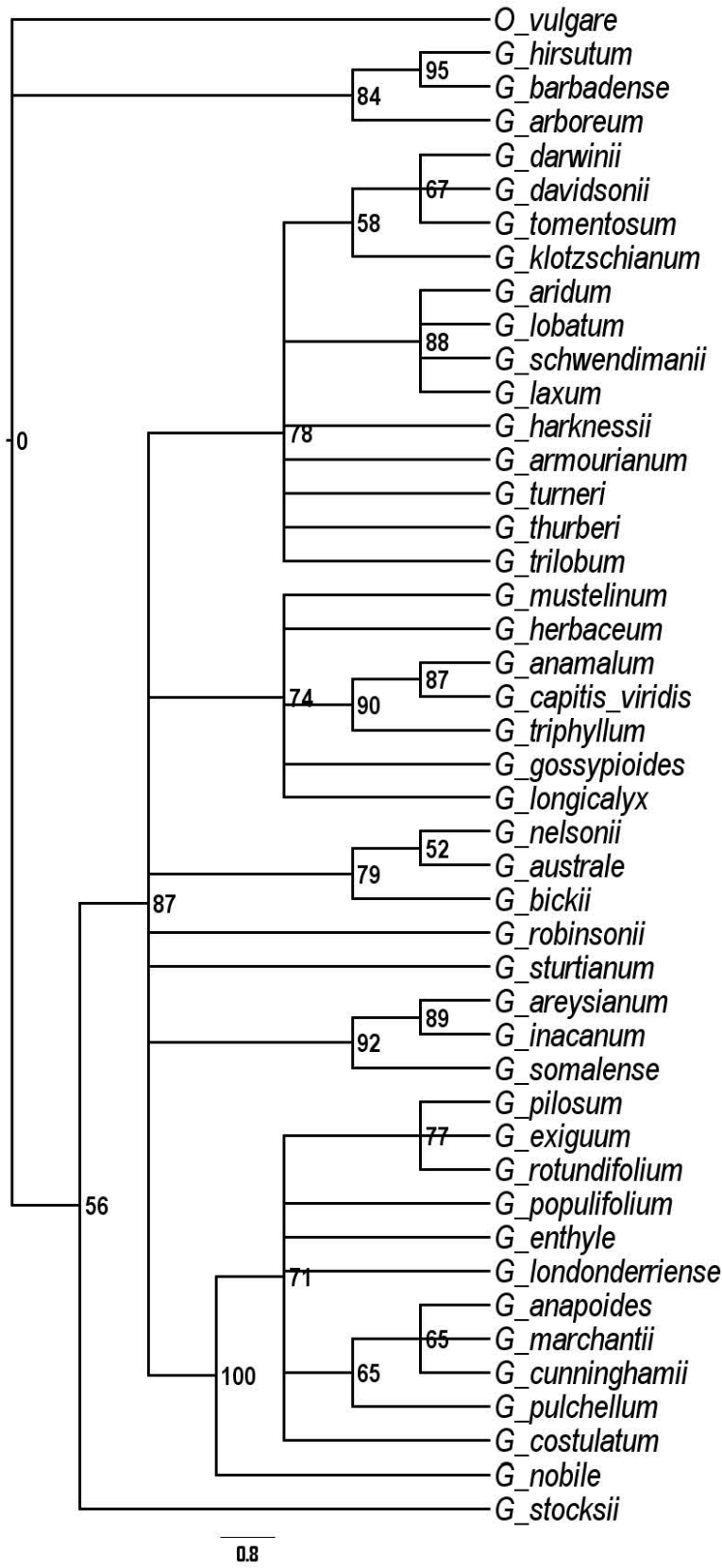
Şekil Ek-77. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi



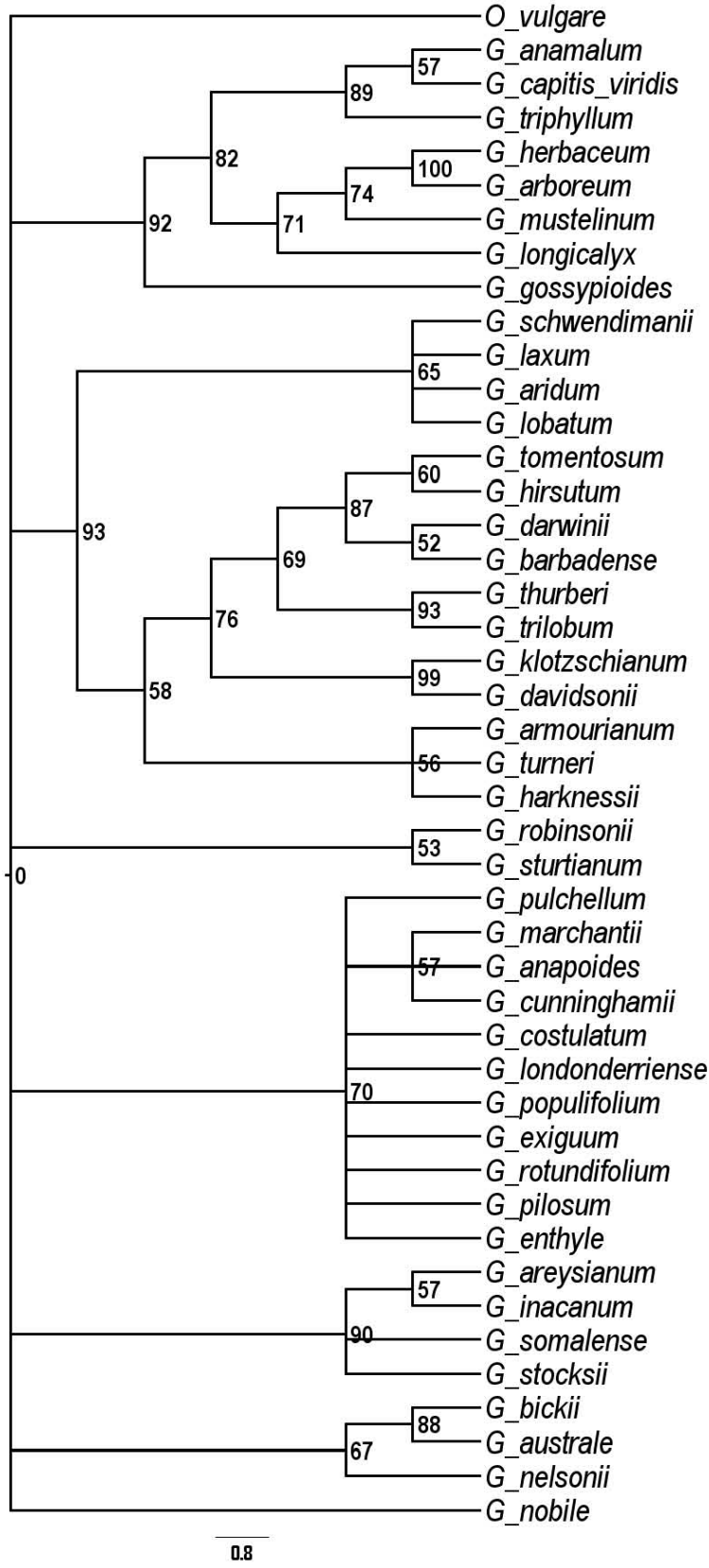
Şekil Ek-78. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi



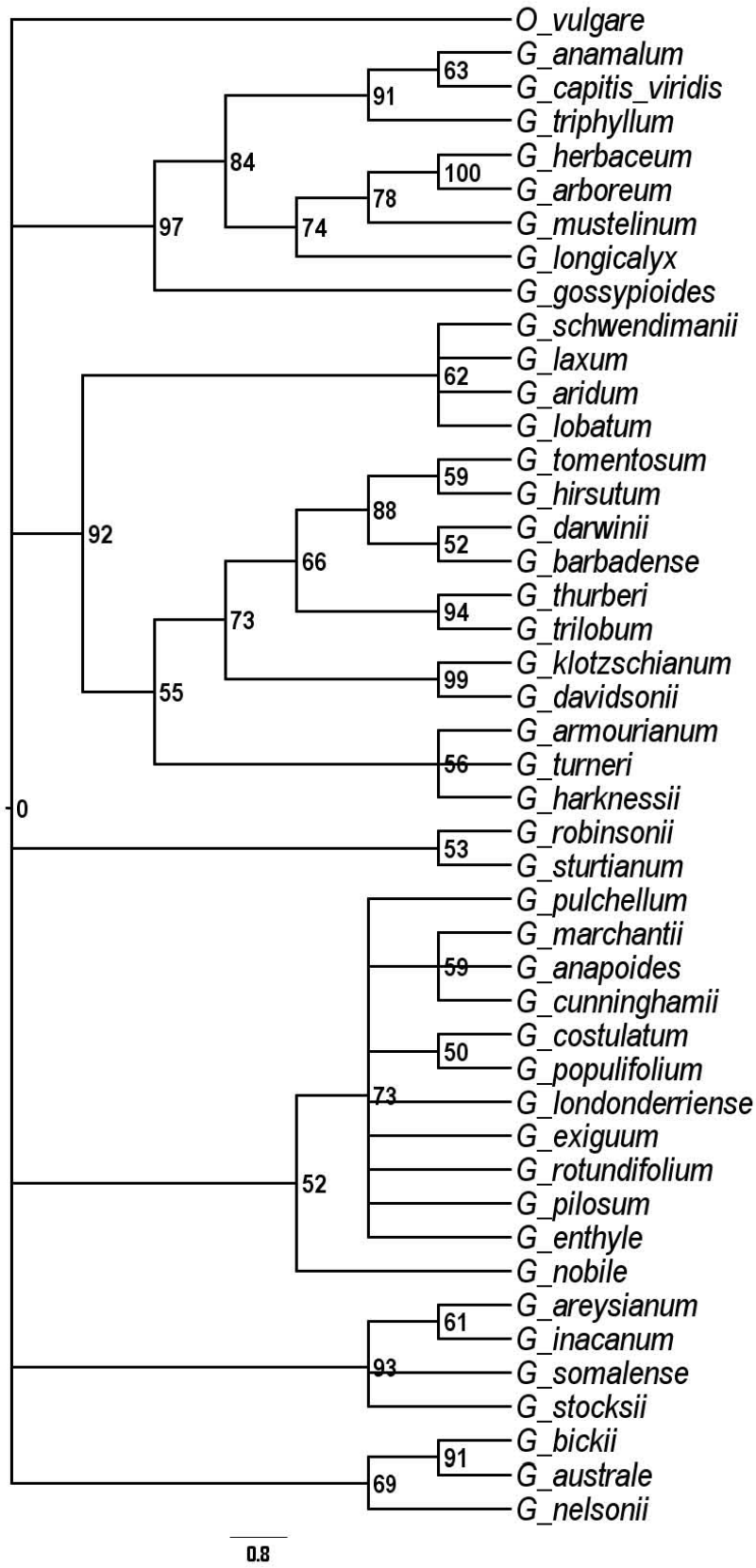
Şekil Ek-79. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi



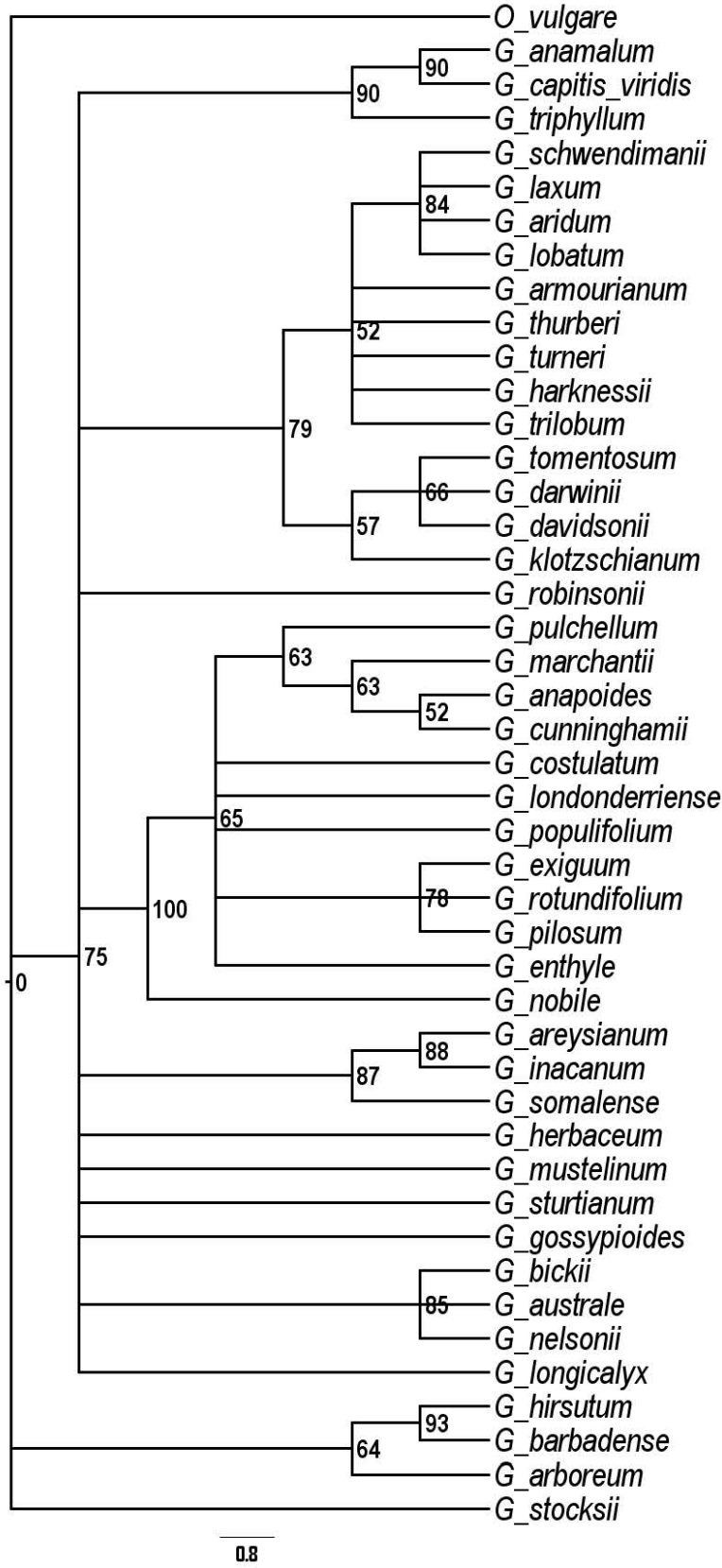
Şekil Ek-80. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi



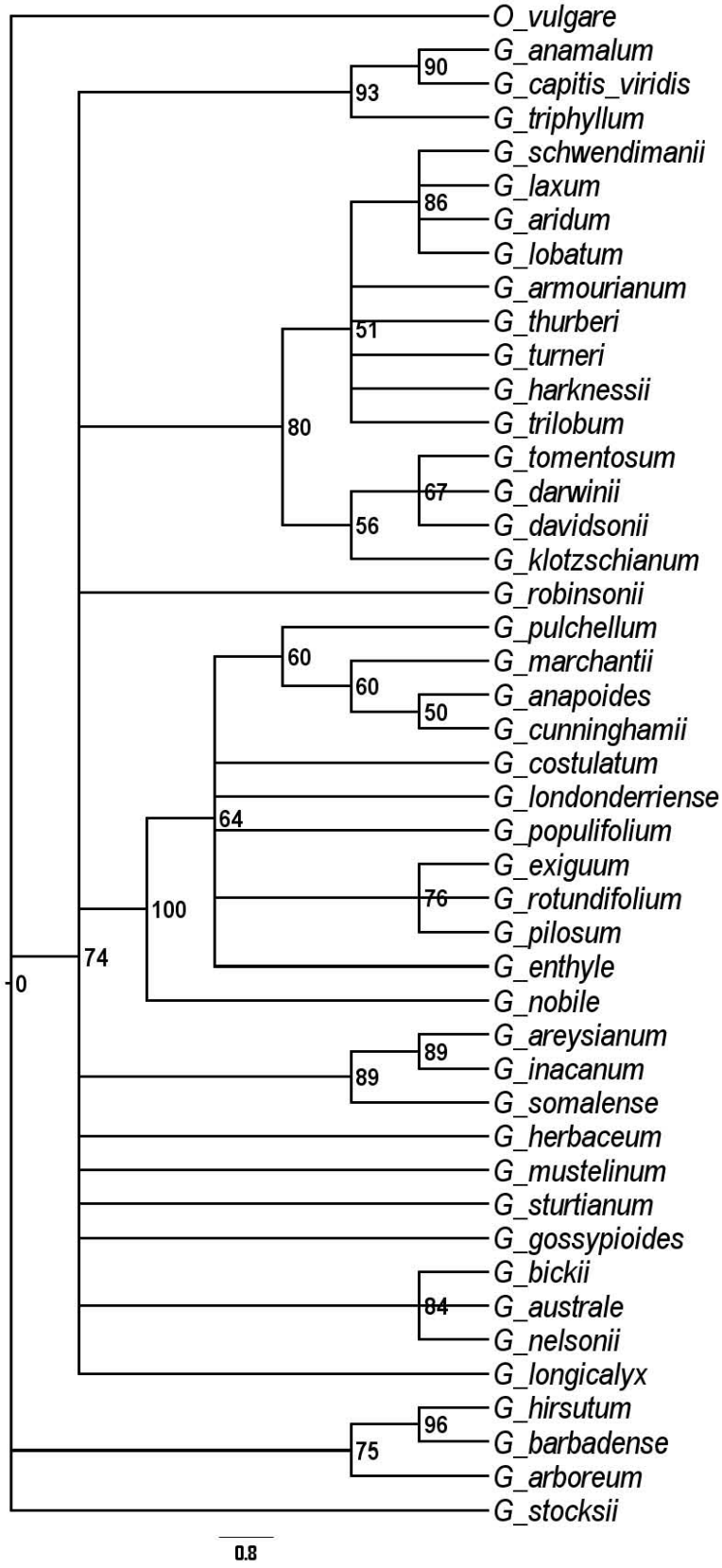
Şekil Ek-81. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi



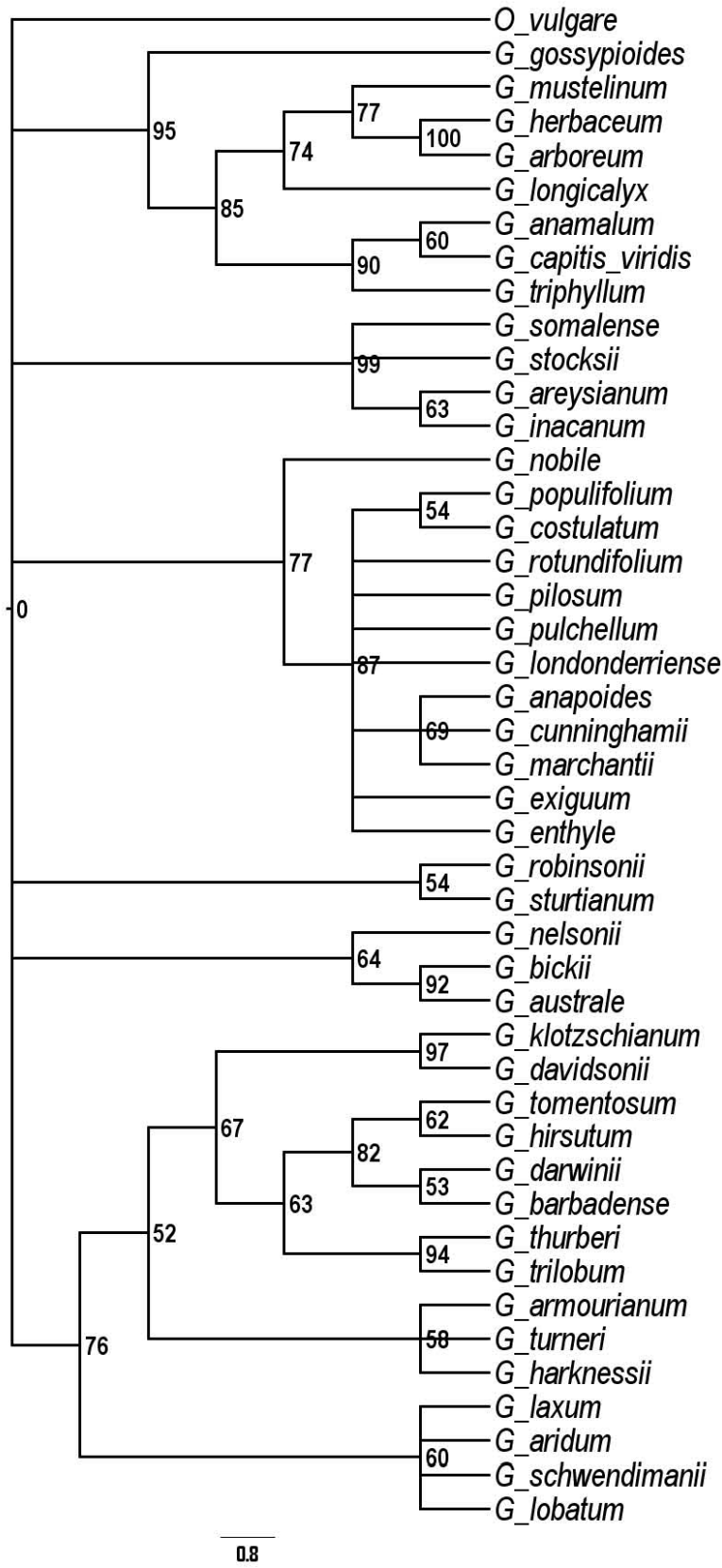
Şekil Ek-82. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi



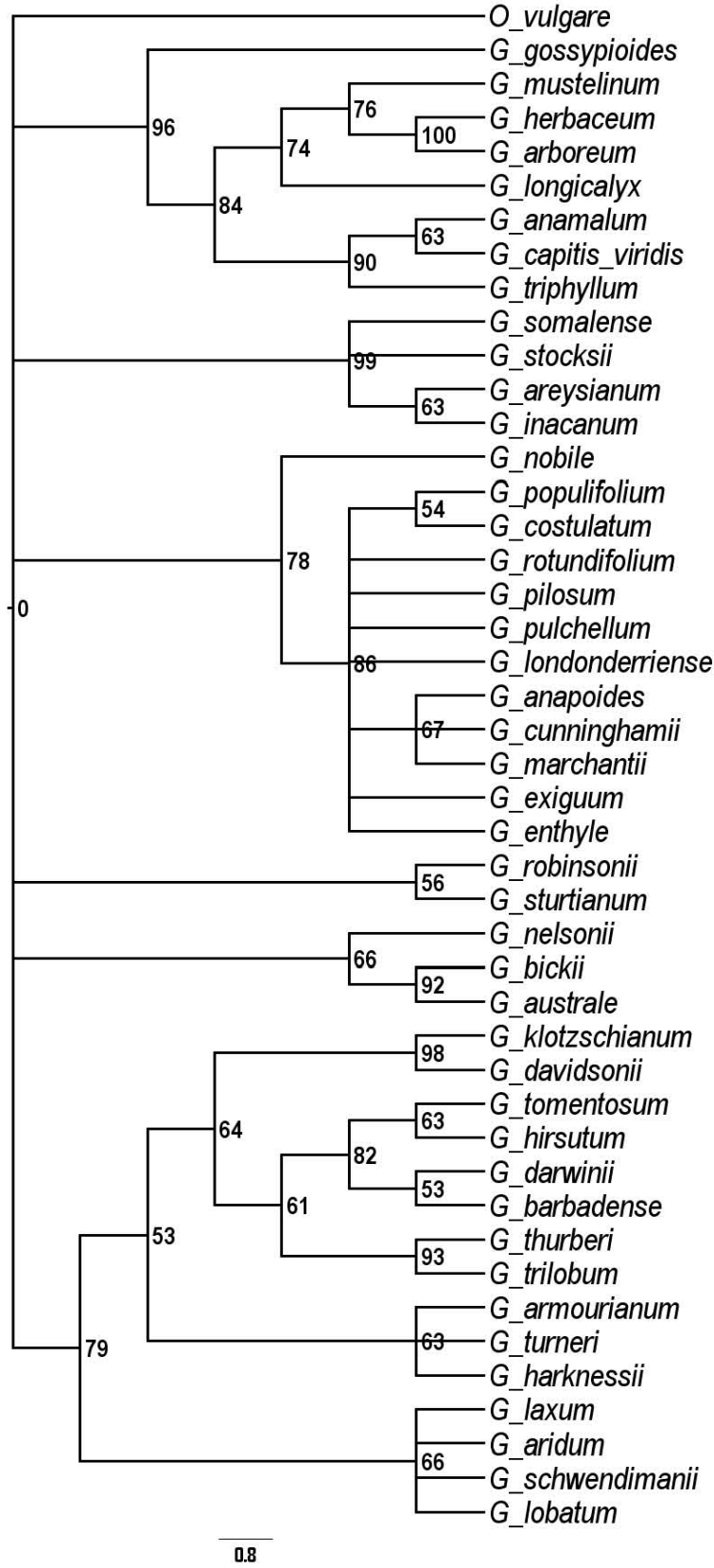
Şekil Ek-83. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi



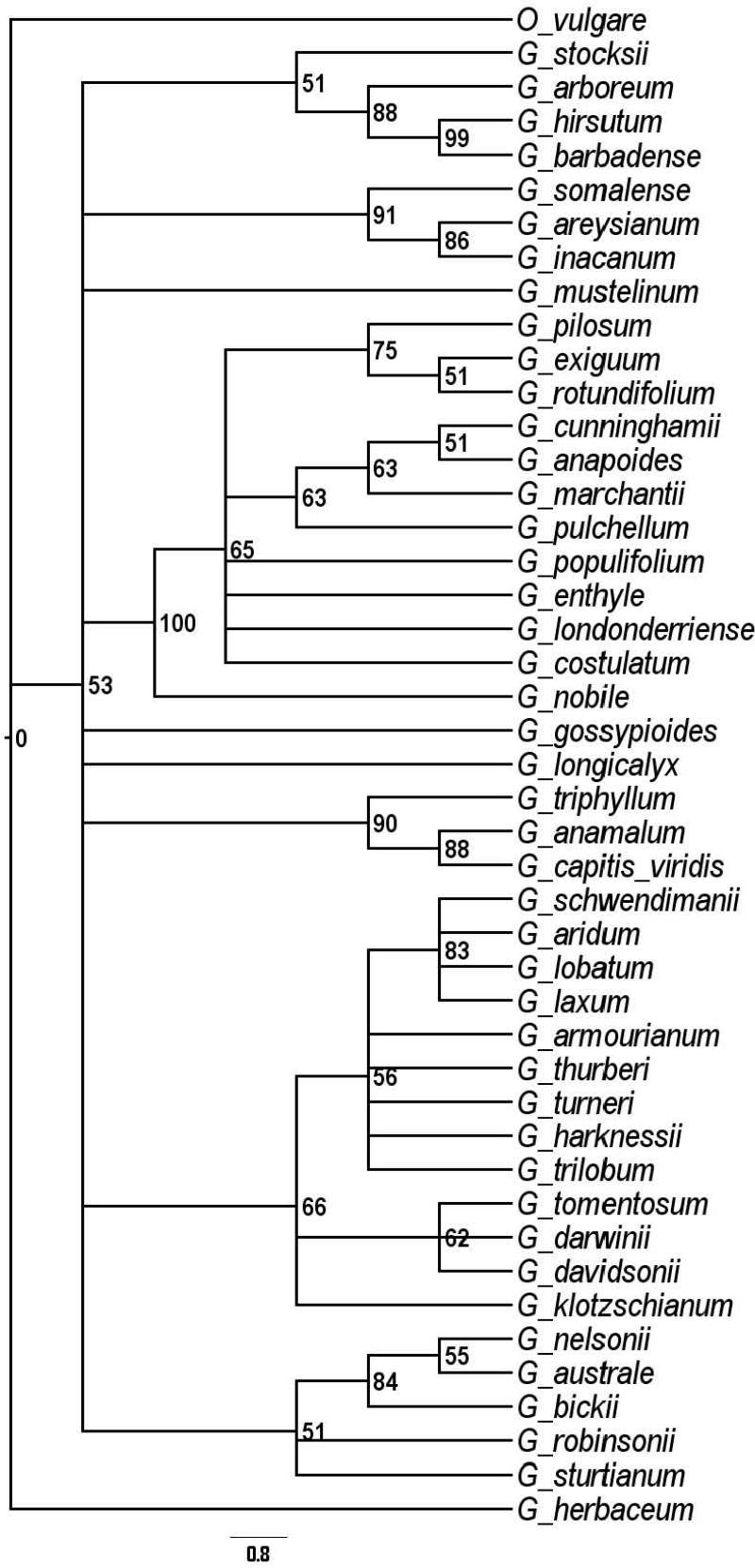
Şekil Ek-84. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansınınK2P SBT modelindeki distance analizi



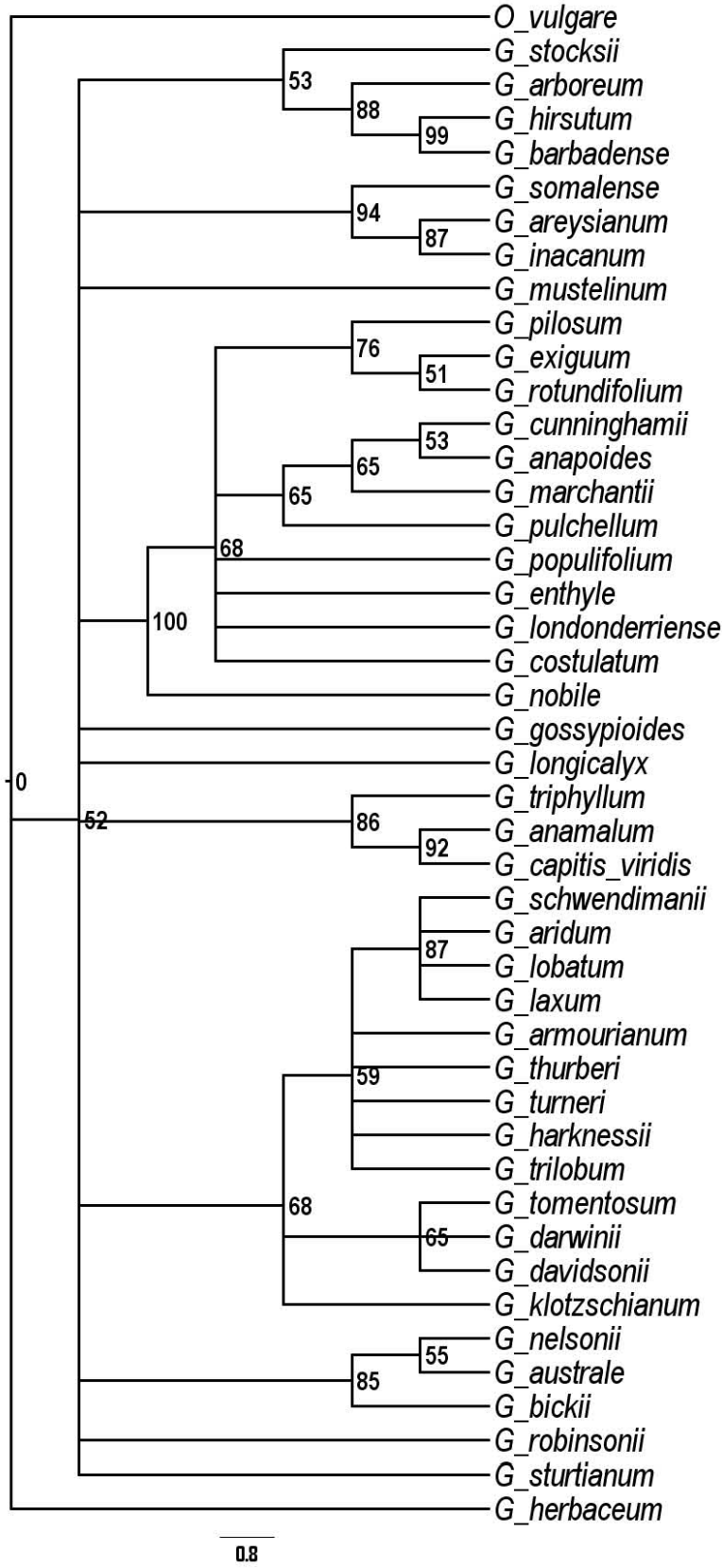
Şekil Ek-85. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının GTR SBT modelindeki distance analizi



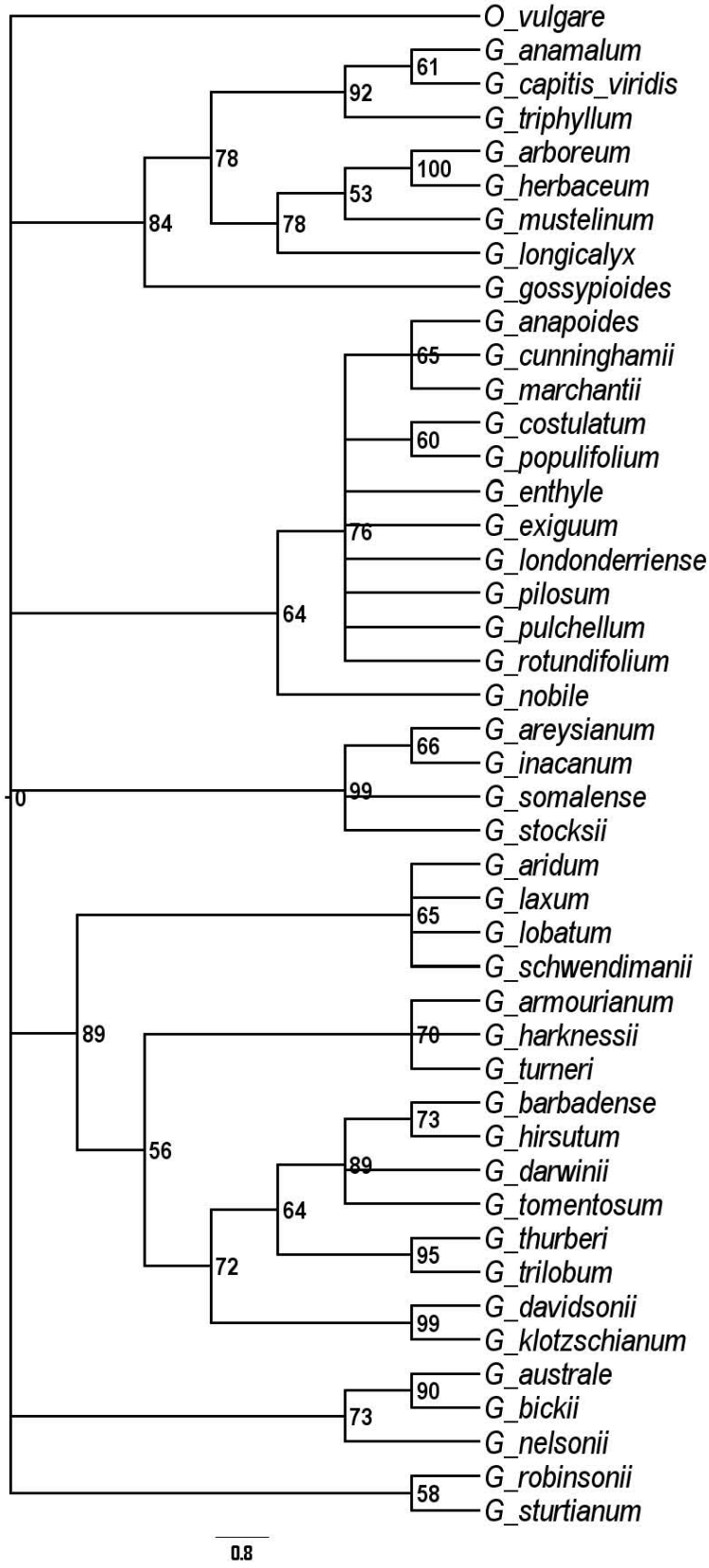
Şekil Ek-86. MUSCLE programının 2. parametresi ile hizalanmış ITS1 sekansının K2P SBT modelindeki distance analizi



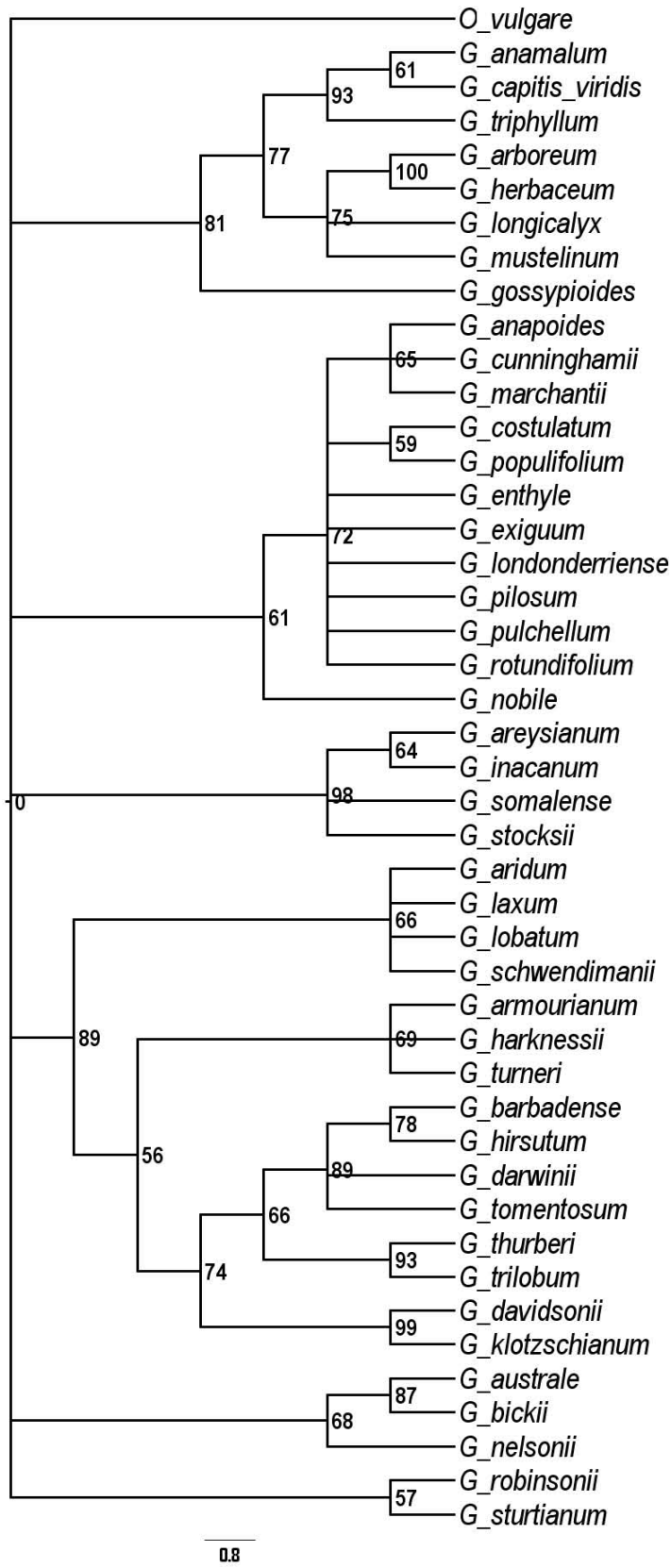
Şekil Ek-87. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının GTR SBT modelindeki distance analizi



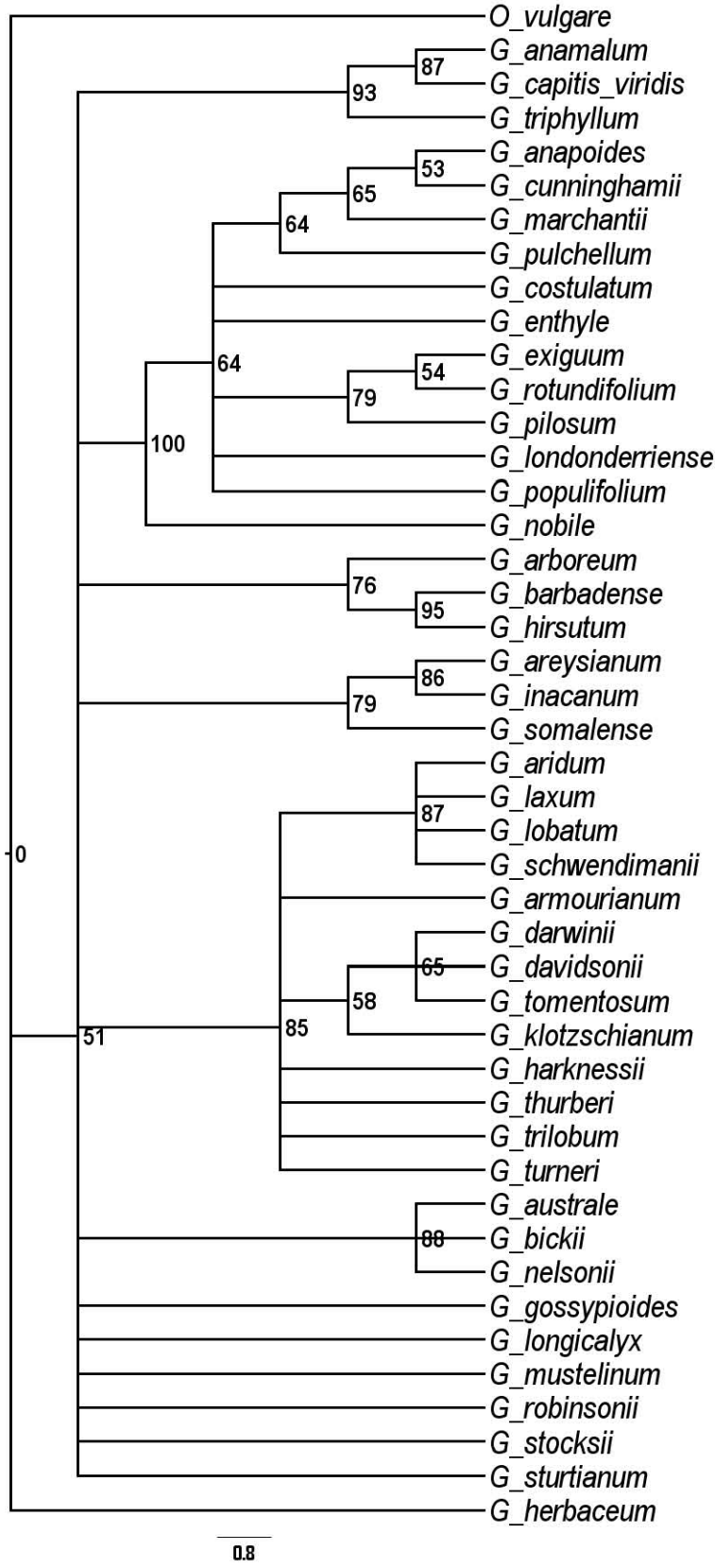
Şekil Ek-88. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının K2P SBT modelindeki distance analizi



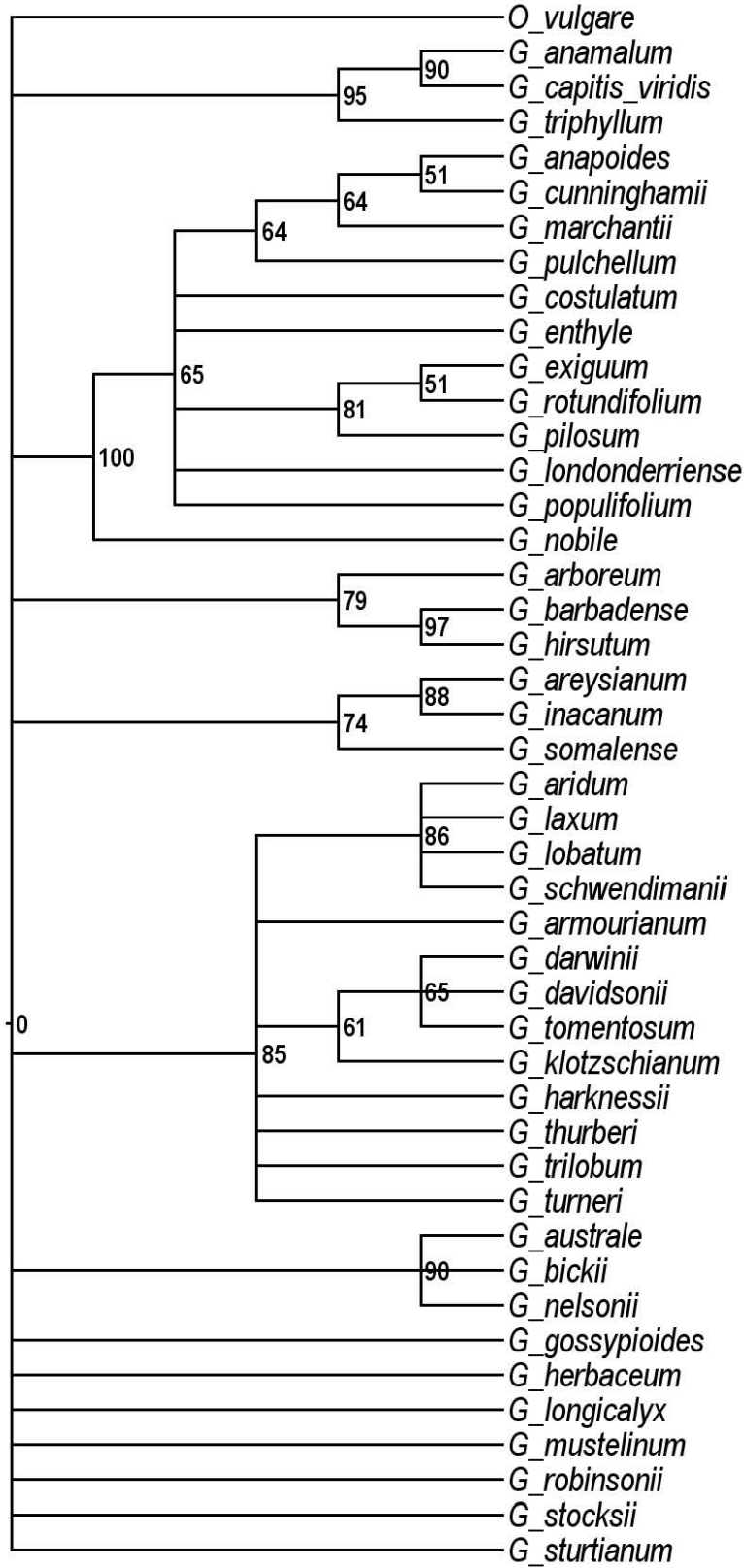
Şekil Ek-89. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansını GTR SBT modelindeki distance analizi



Şekil Ek-90. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansını K2P SBT modelindeki distance analizi

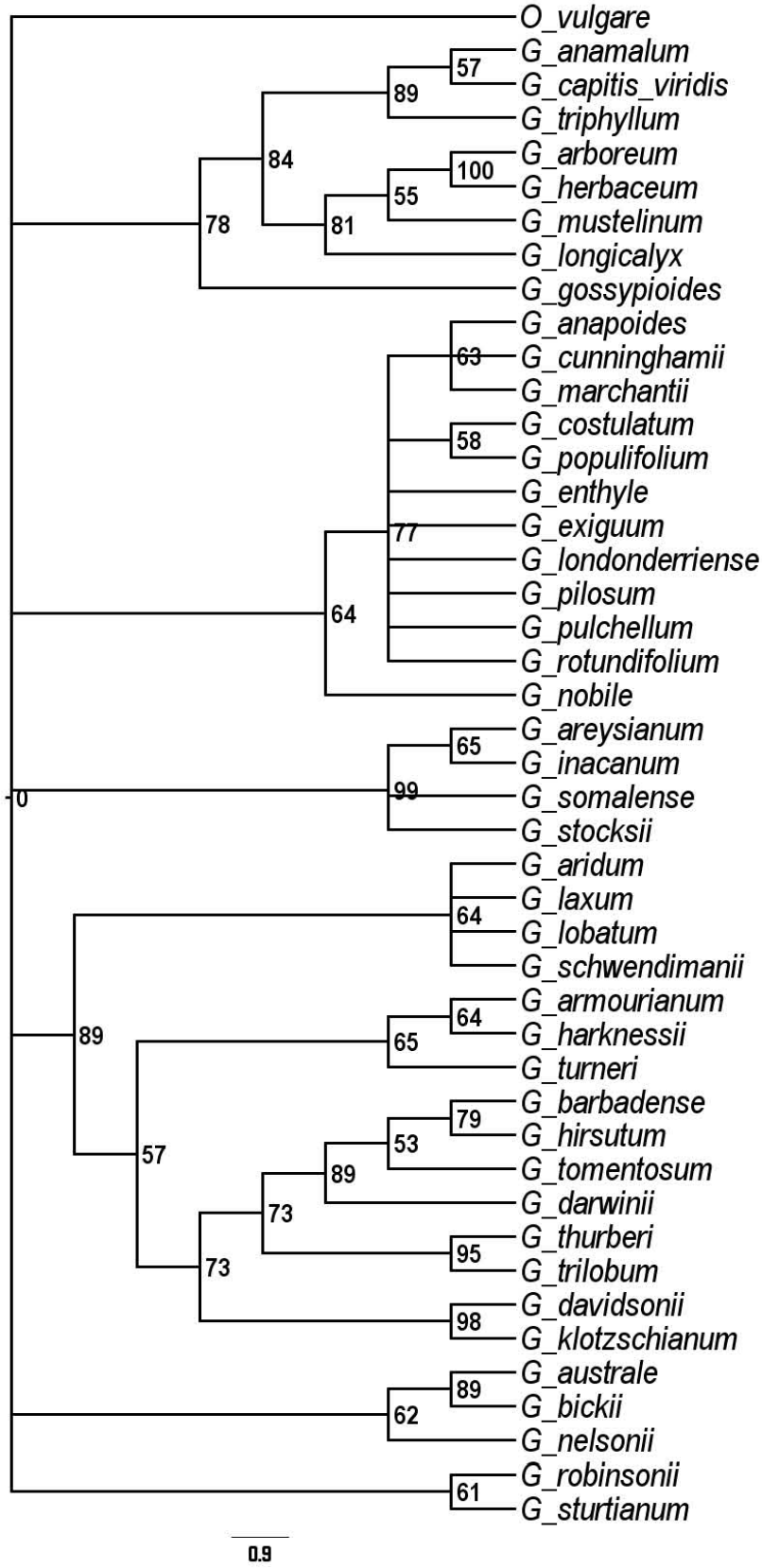


Şekil Ek-91. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansını GTR SBT modelindeki distance analizi

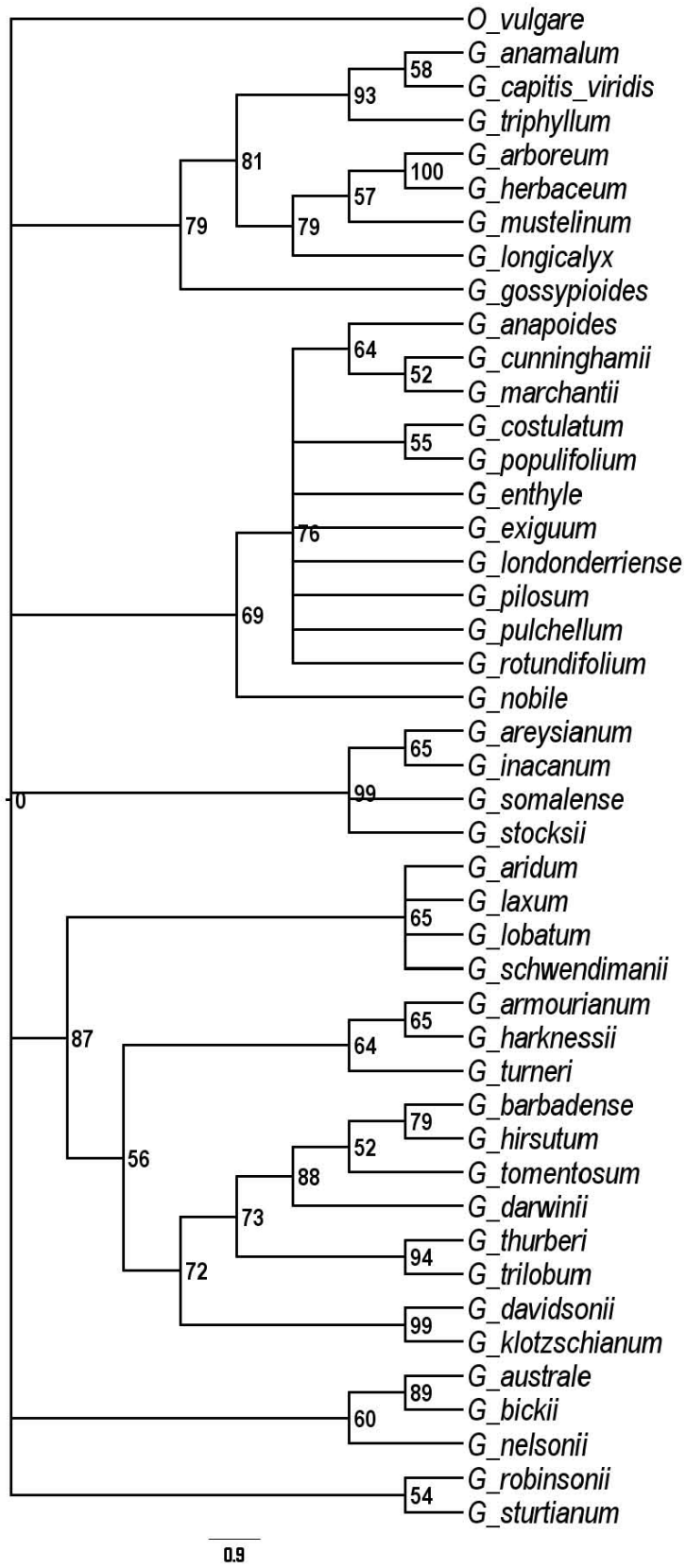


0.7

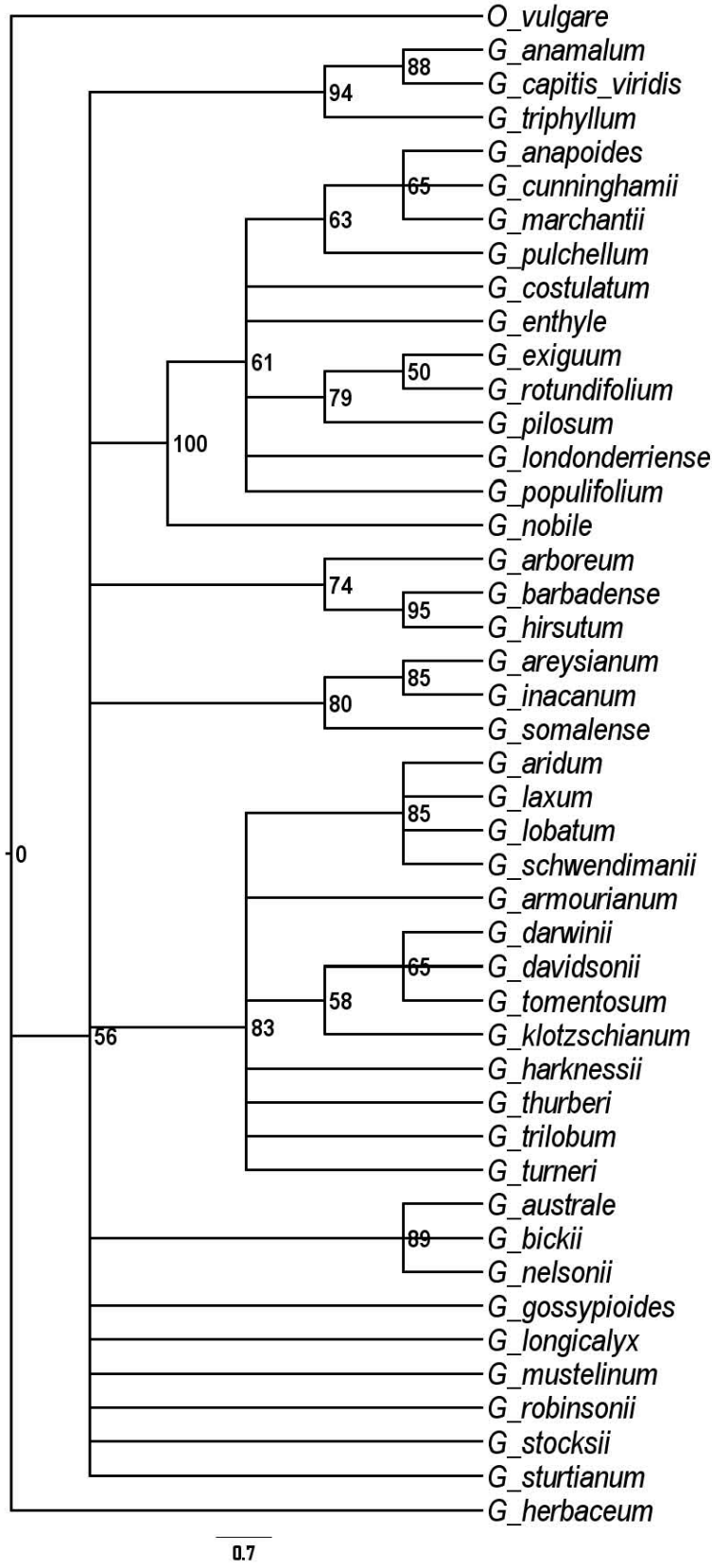
Şekil Ek-92. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansını K2P SBT modelindeki distance analizi



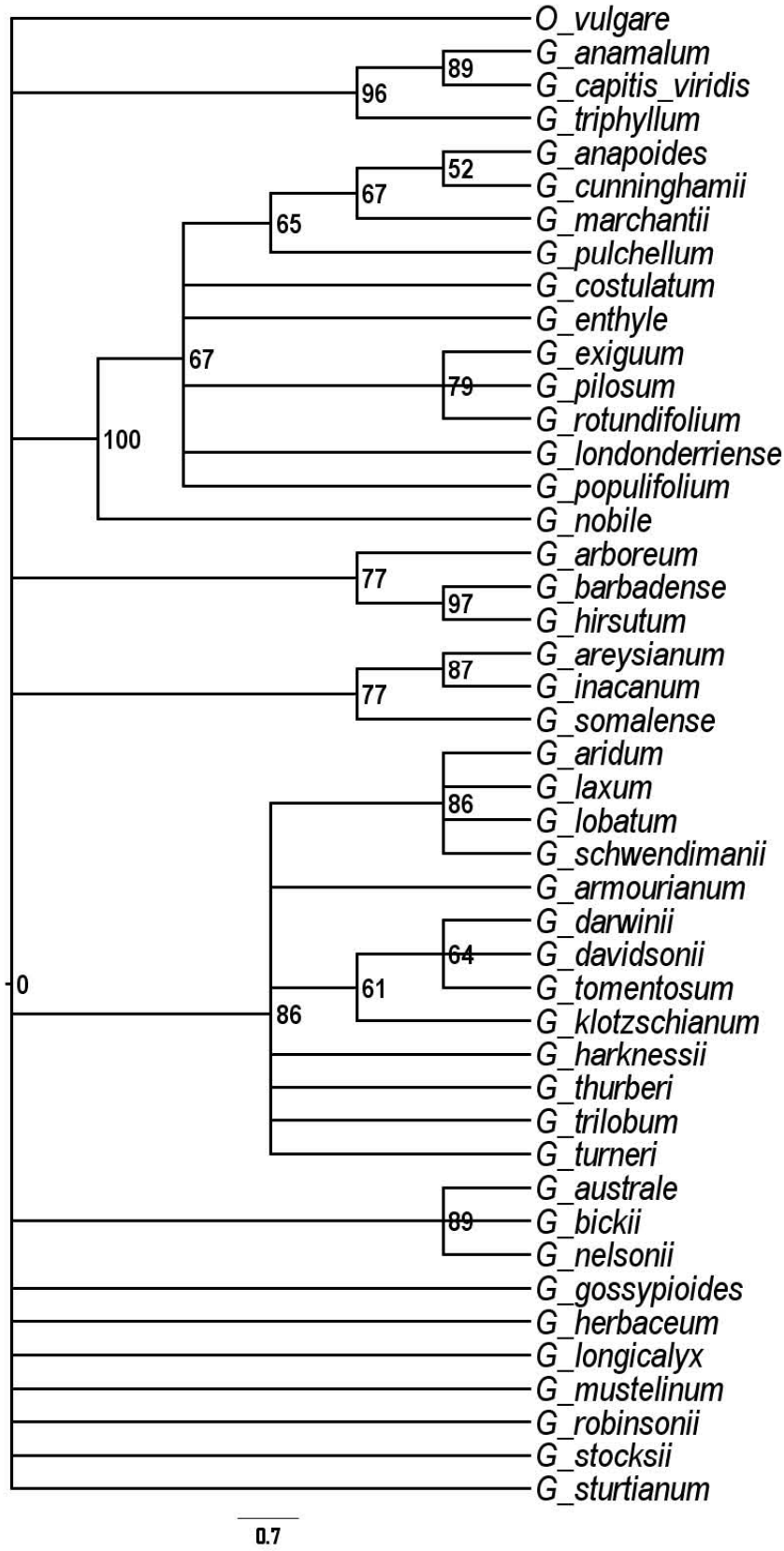
Şekil Ek-93. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansını GTR SBT modelindeki distance analizi



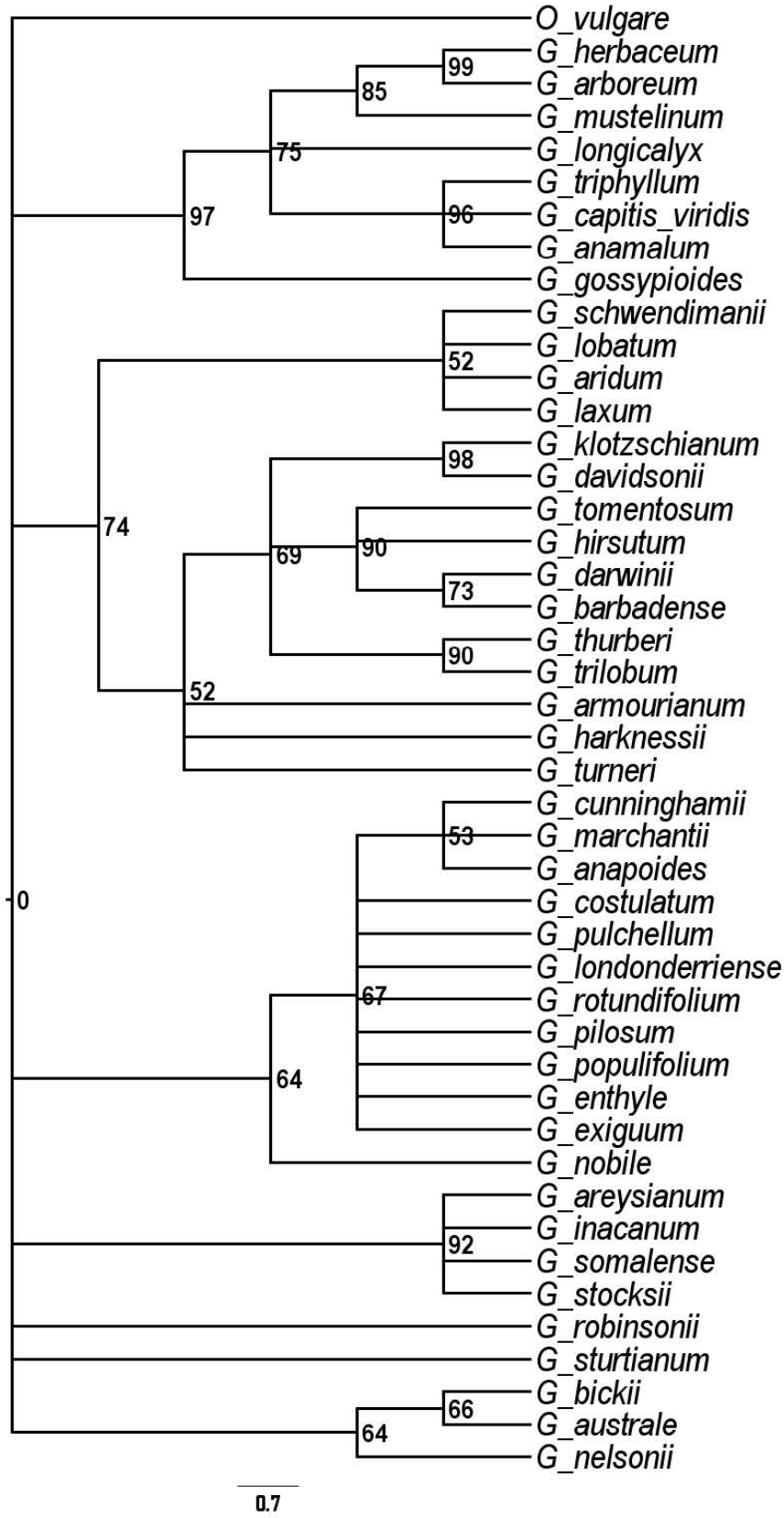
Şekil Ek-94. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansını K2P SBT modelindeki distance analizi



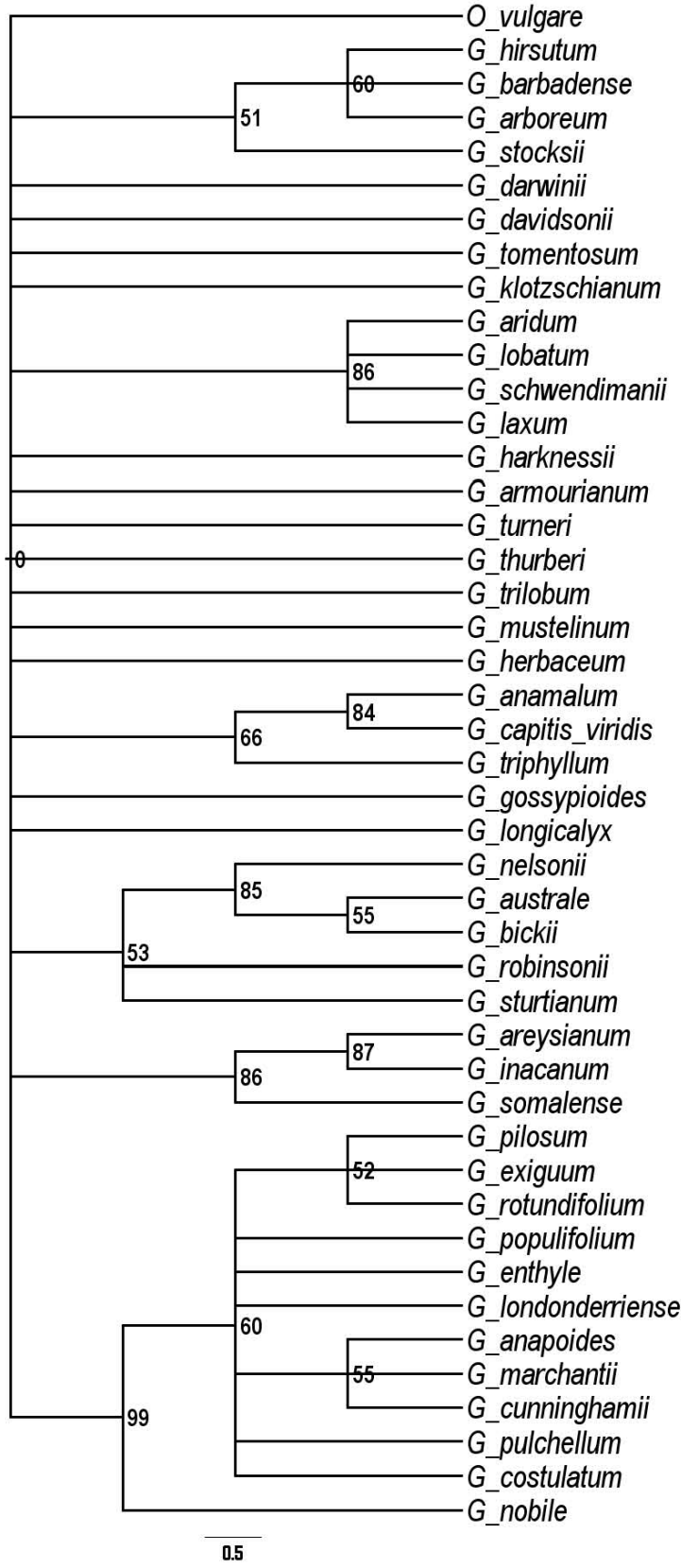
Şekil Ek-95. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansını GTR SBT modelindeki distance analizi



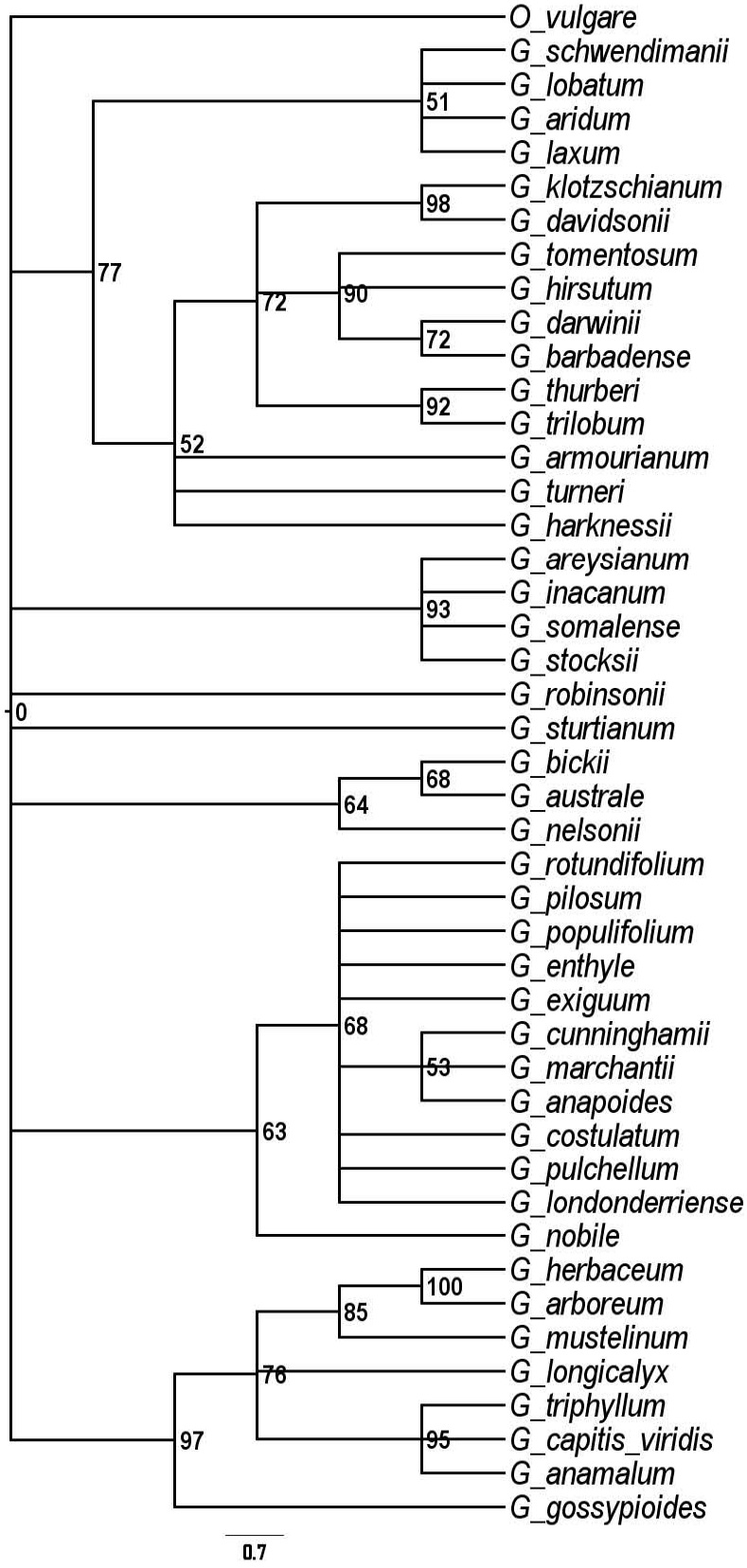
Şekil Ek-96. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansını K2P SBT modelindeki distance analizi



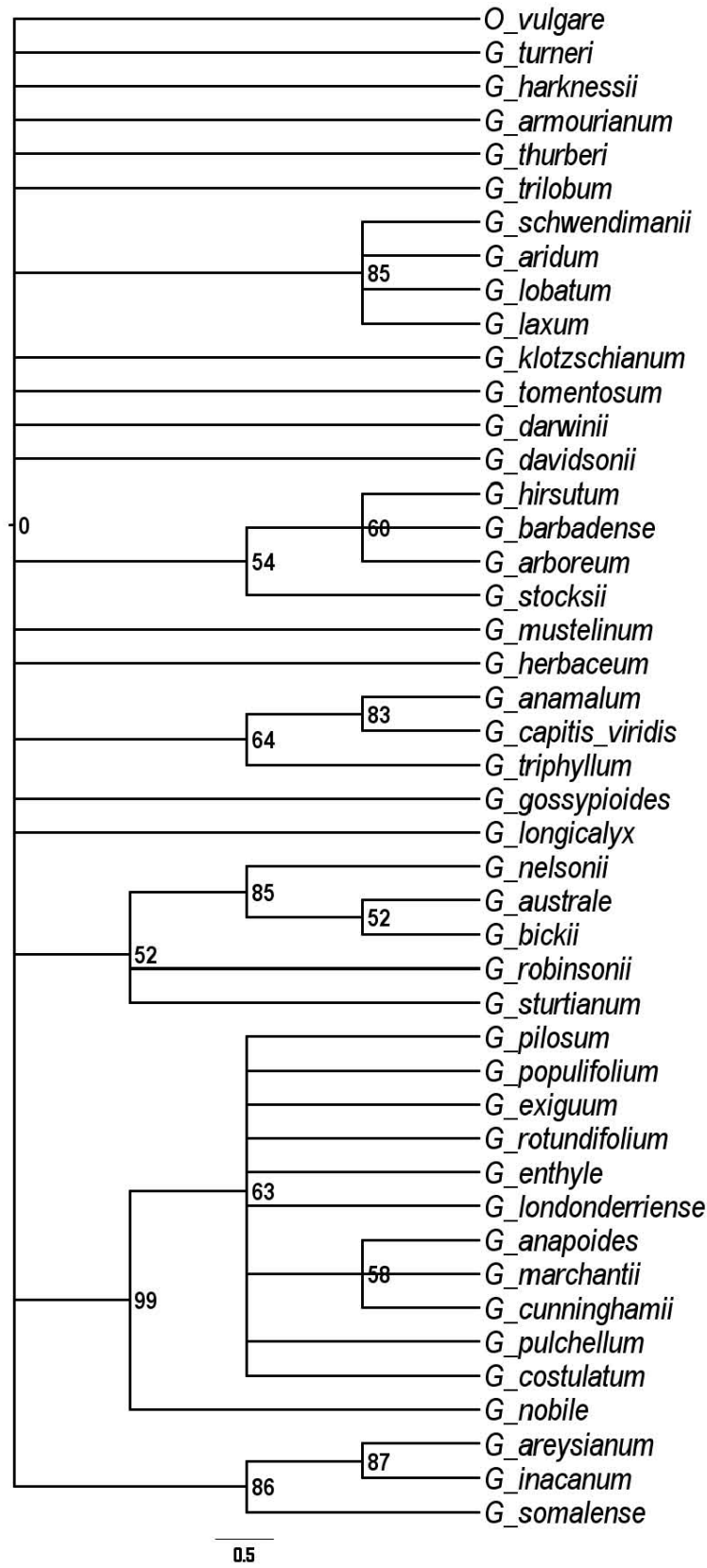
Şekil Ek-97. ClustalW programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



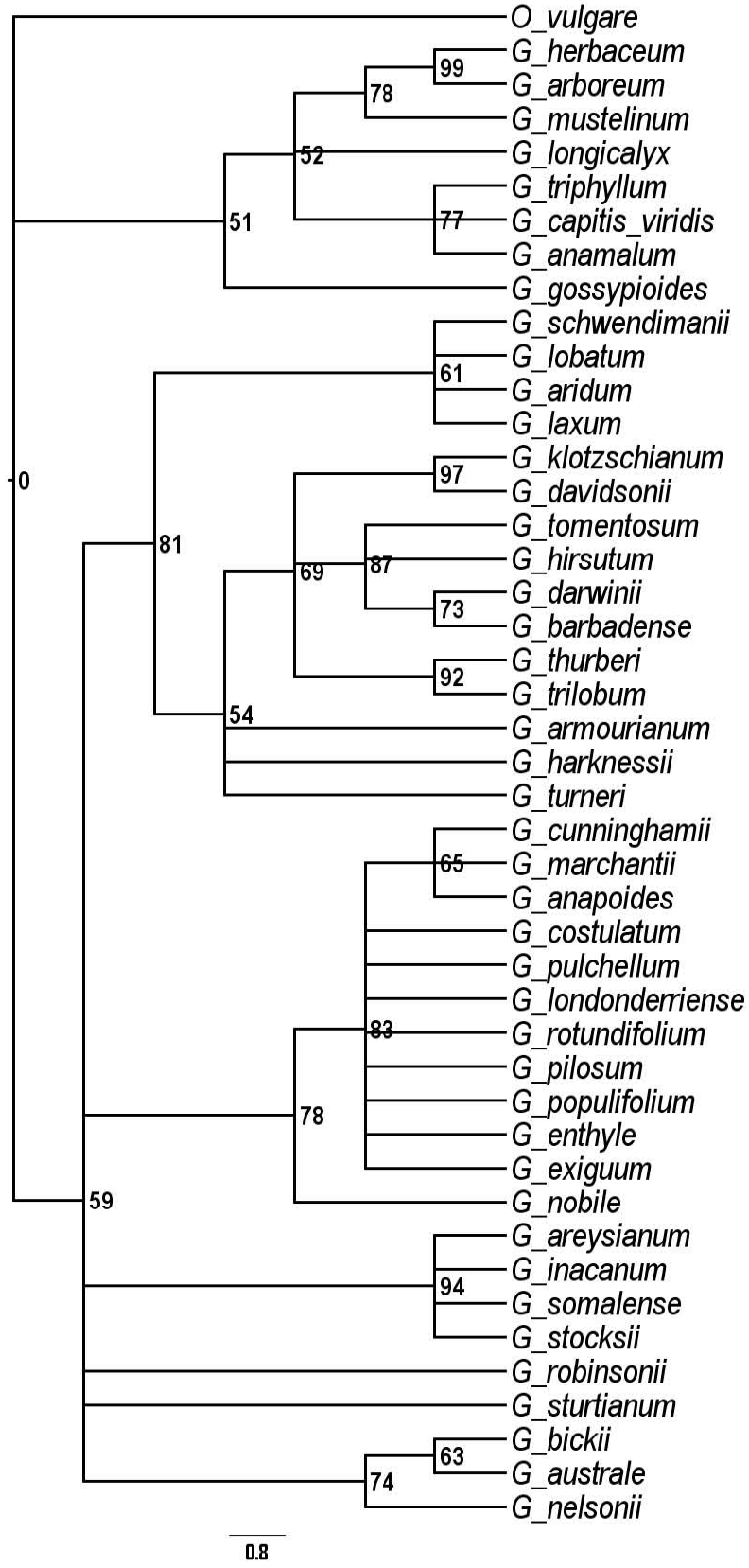
Şekil Ek-98. ClustalW programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi



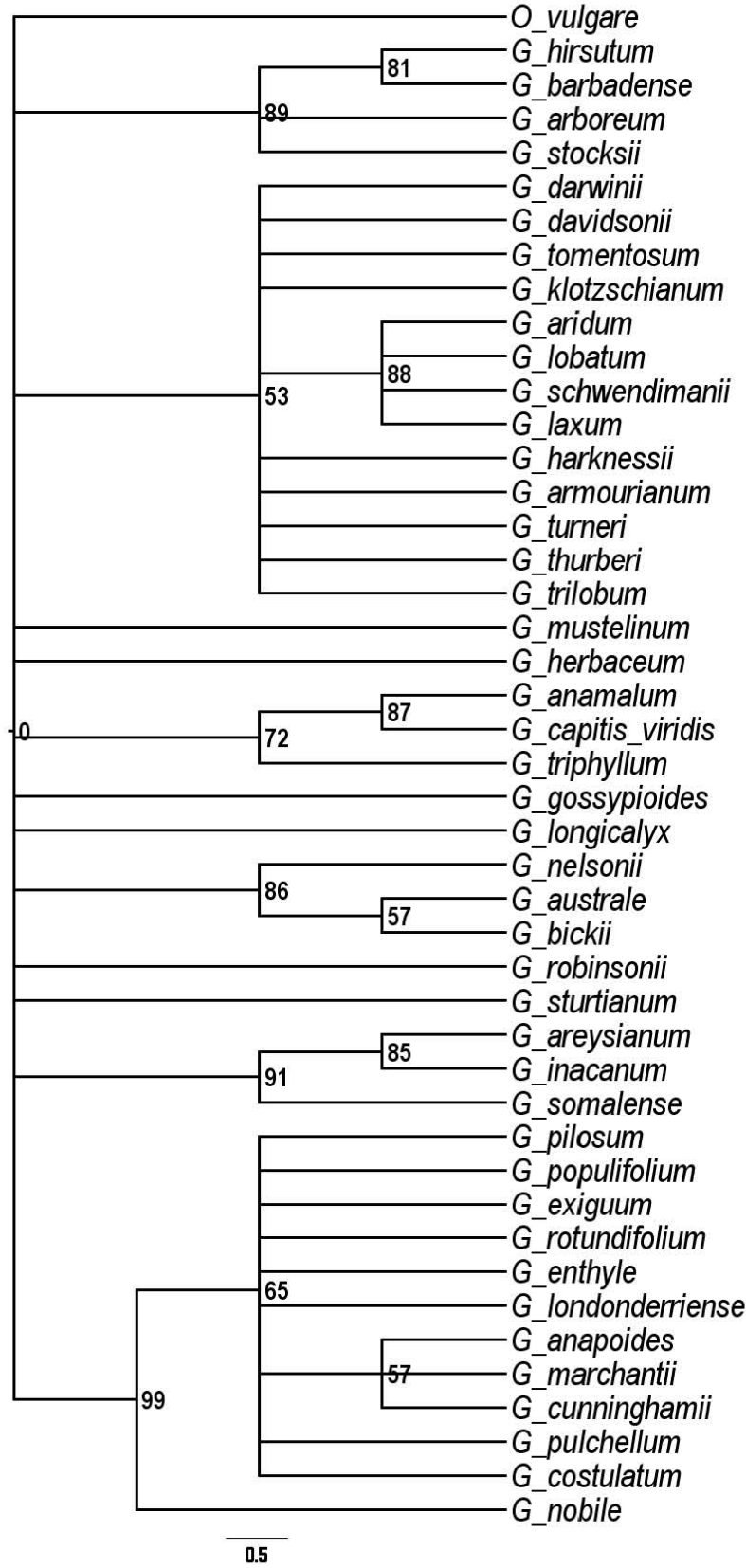
Şekil Ek-99. ClustalW programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



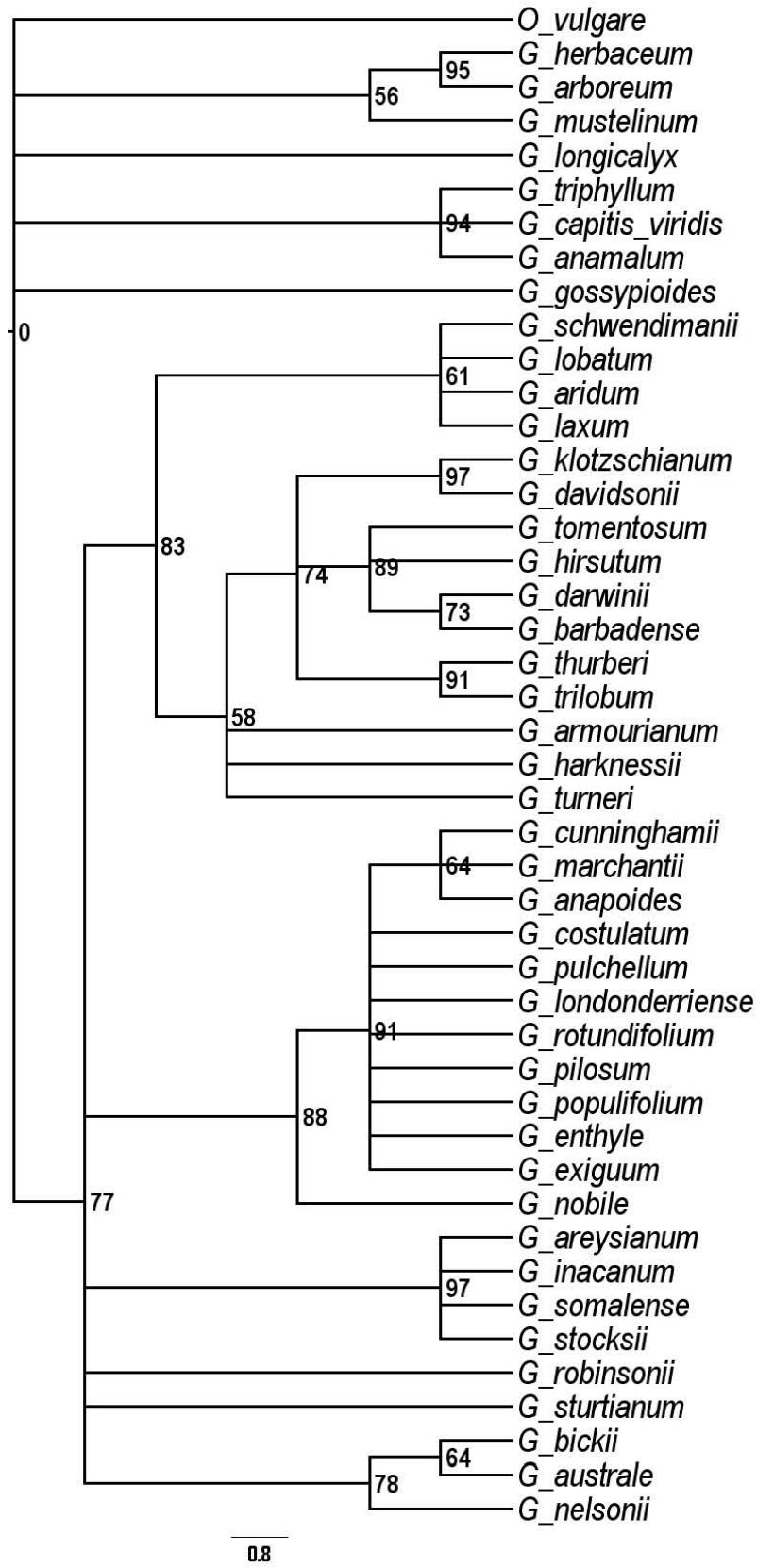
Şekil Ek-100. ClustalW programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi



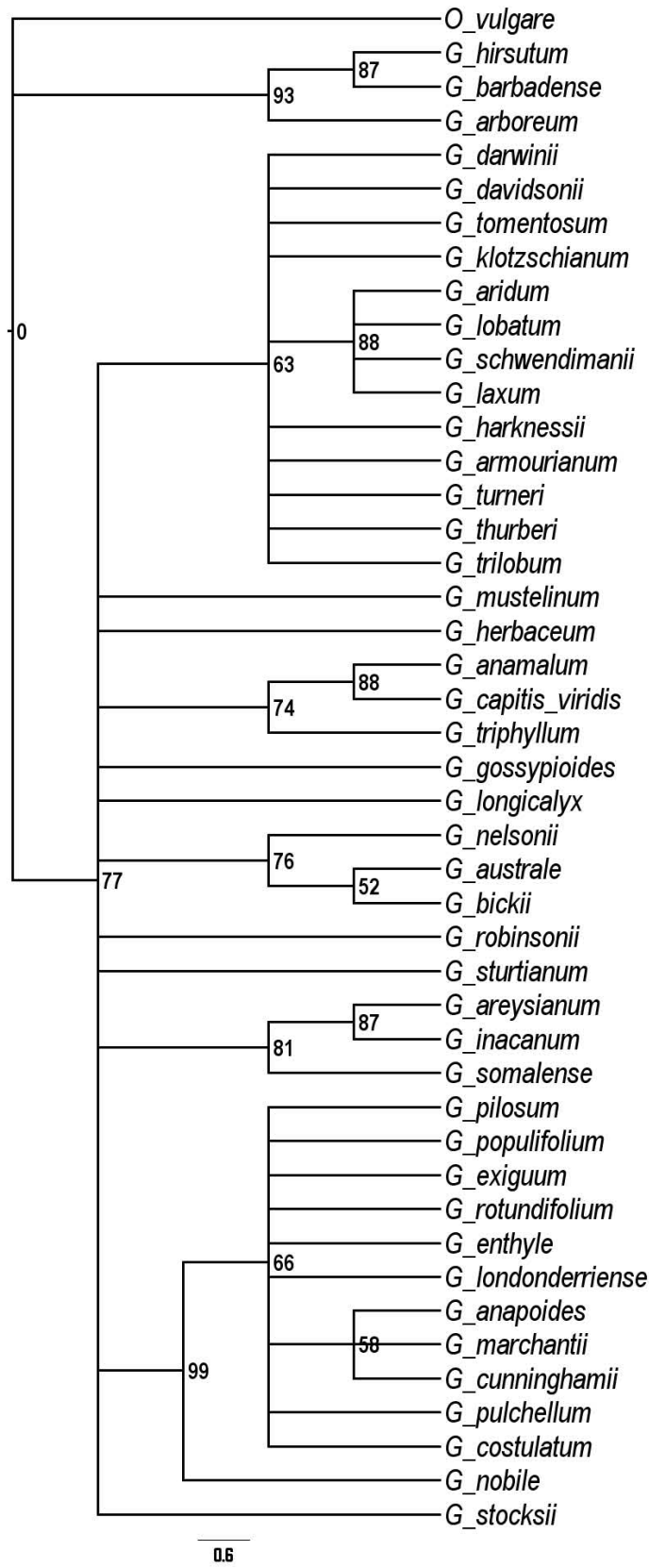
Şekil Ek-101. MAFFT programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



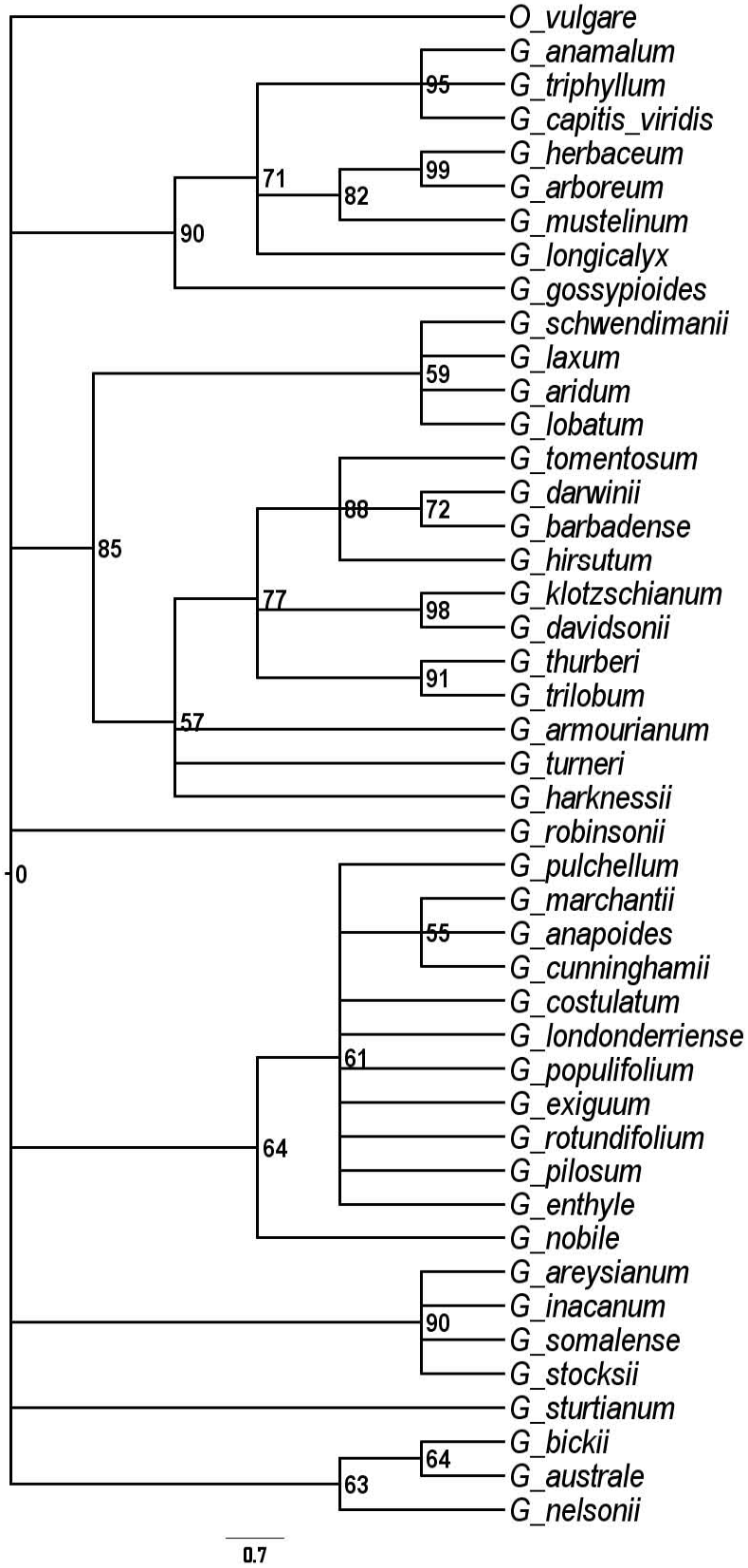
Şekil Ek-102. MAFFT programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi



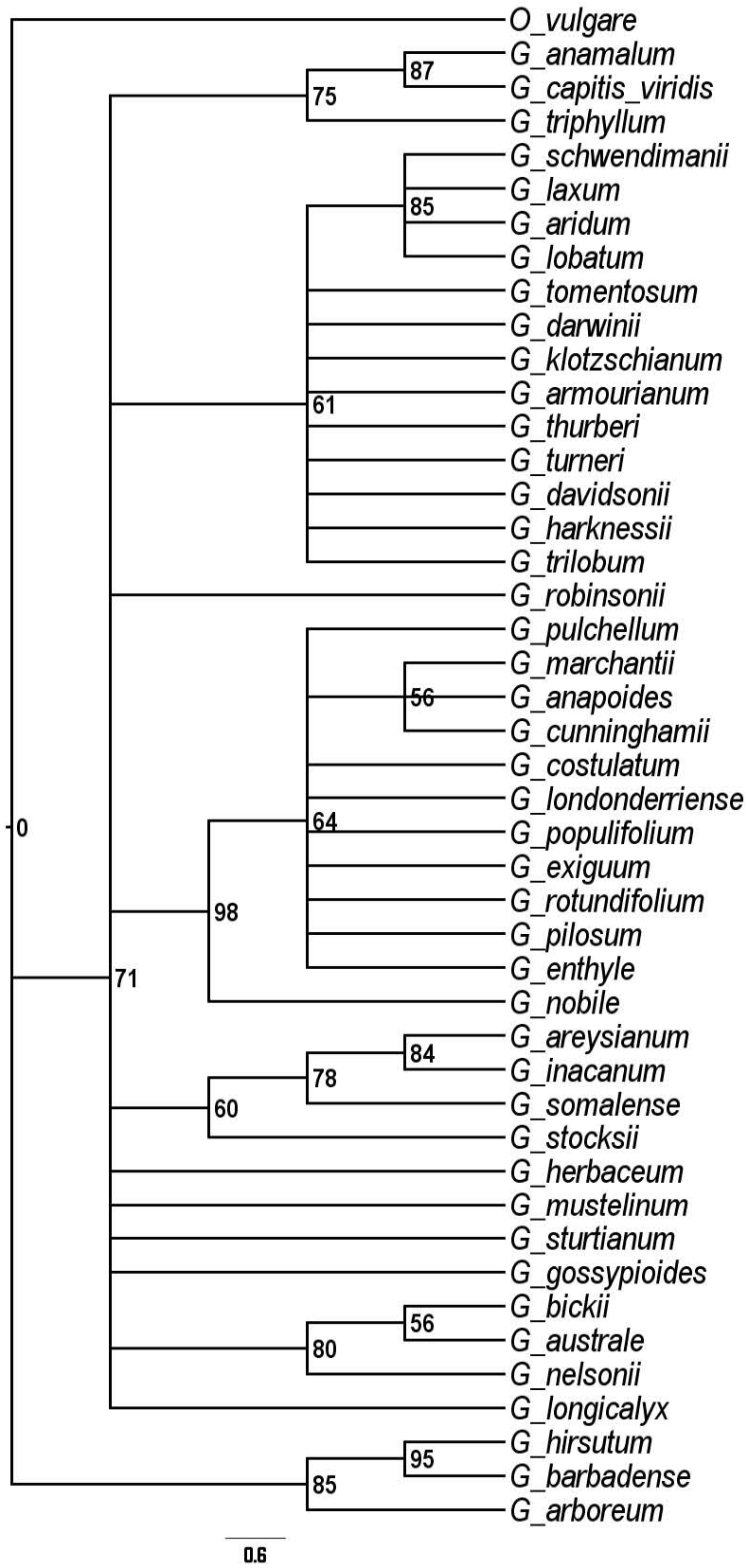
Şekil Ek-103. MAFFT programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



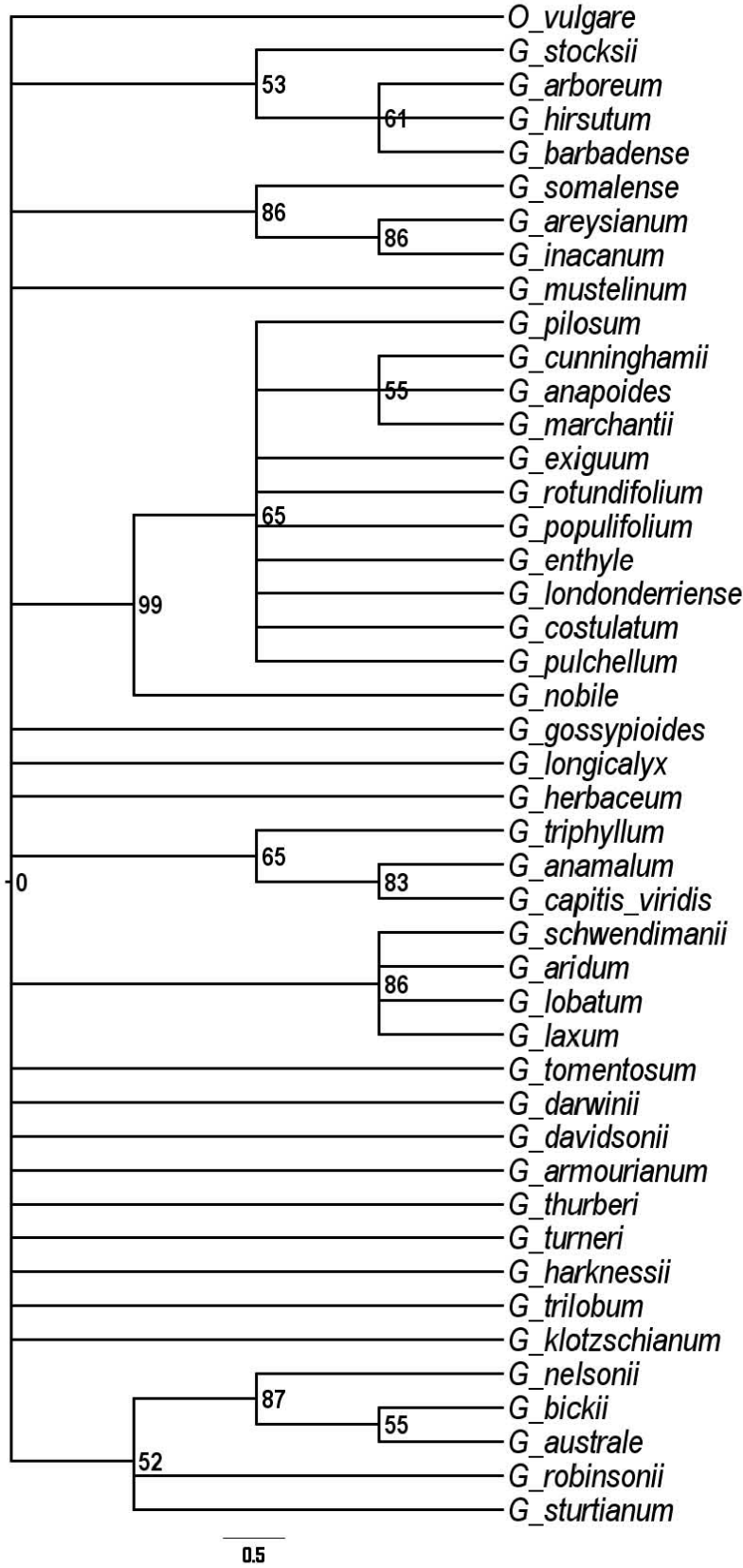
Şekil Ek-104. MAFFT programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi



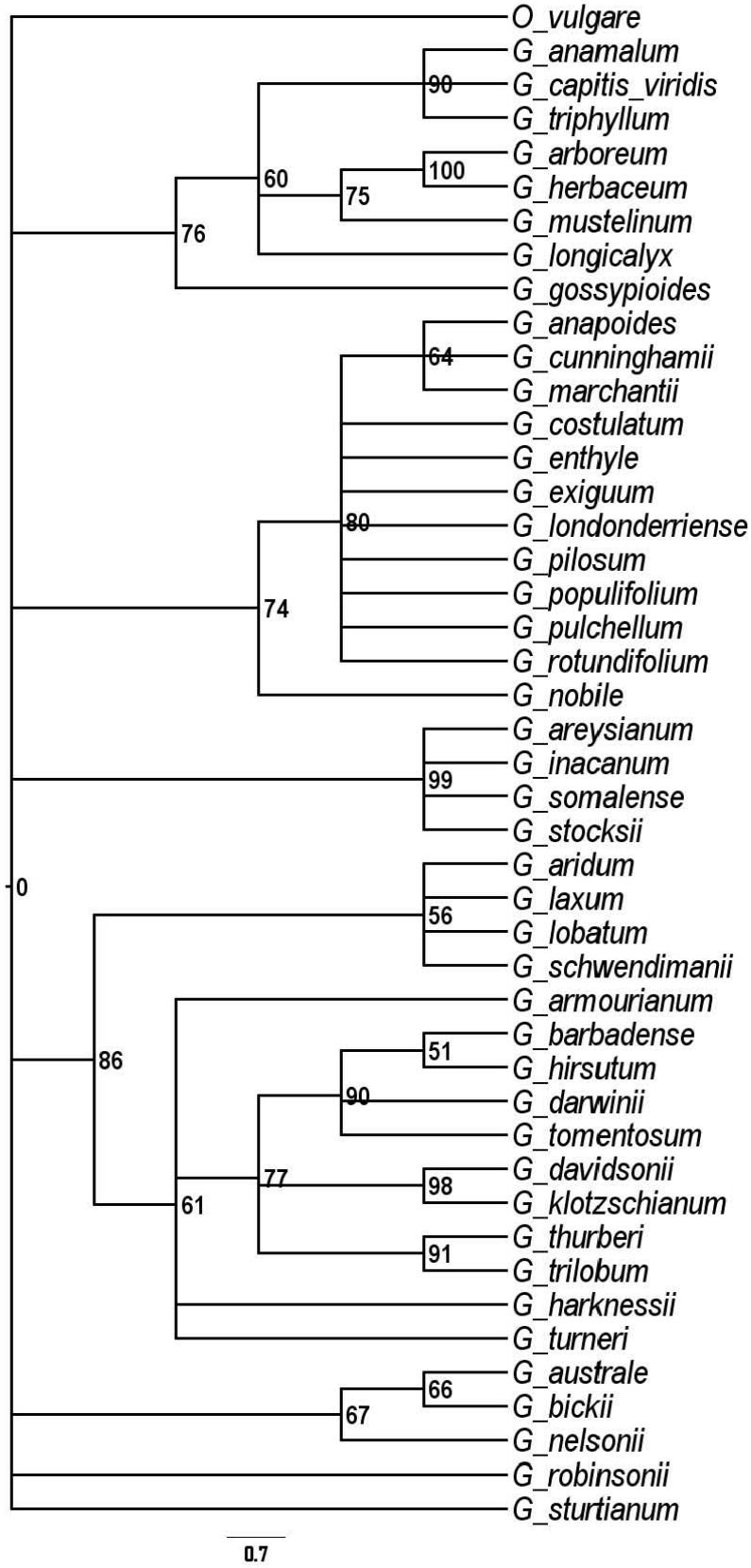
Şekil Ek-105. MUSCLE programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



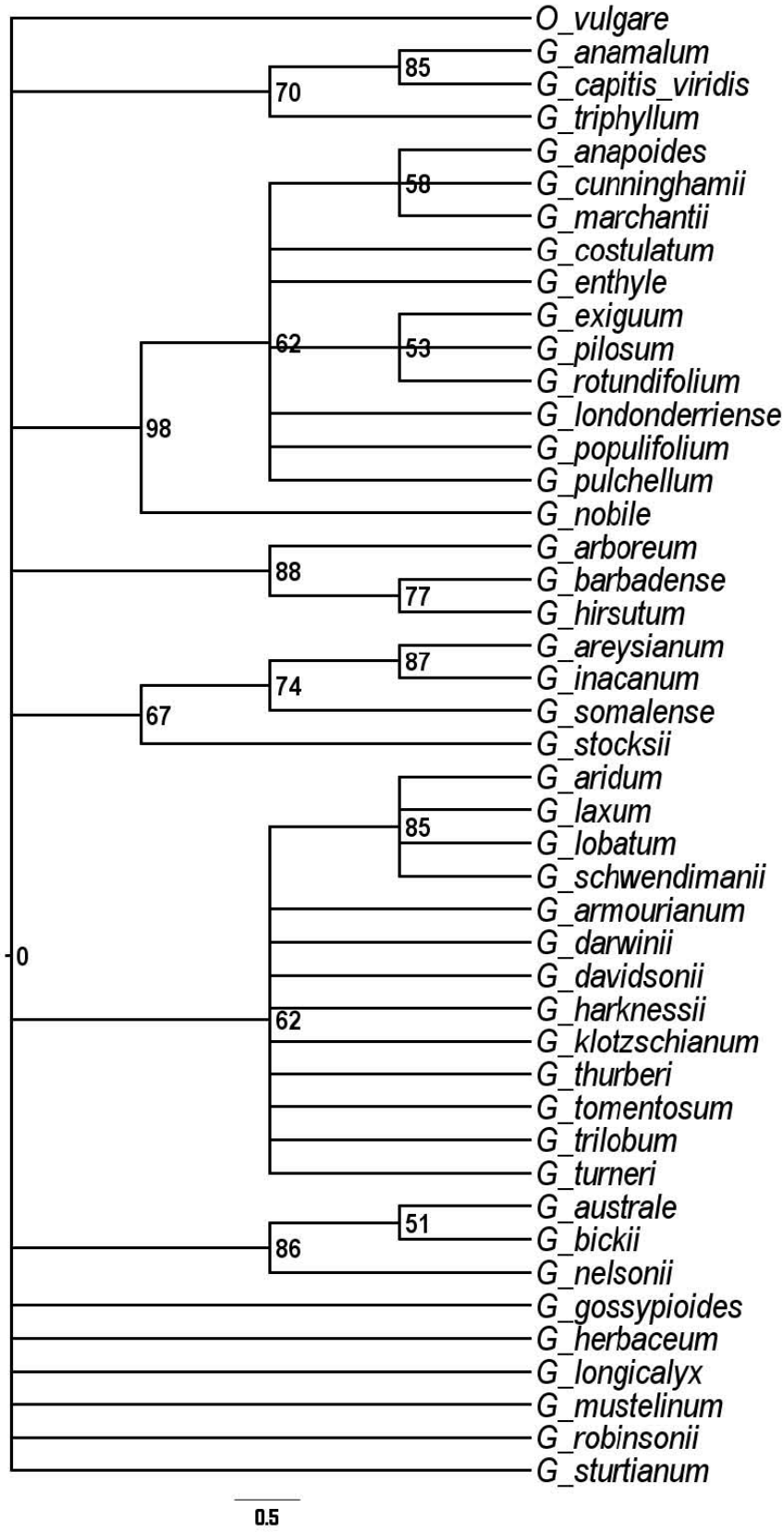
Şekil Ek-106. MUSCLE programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi



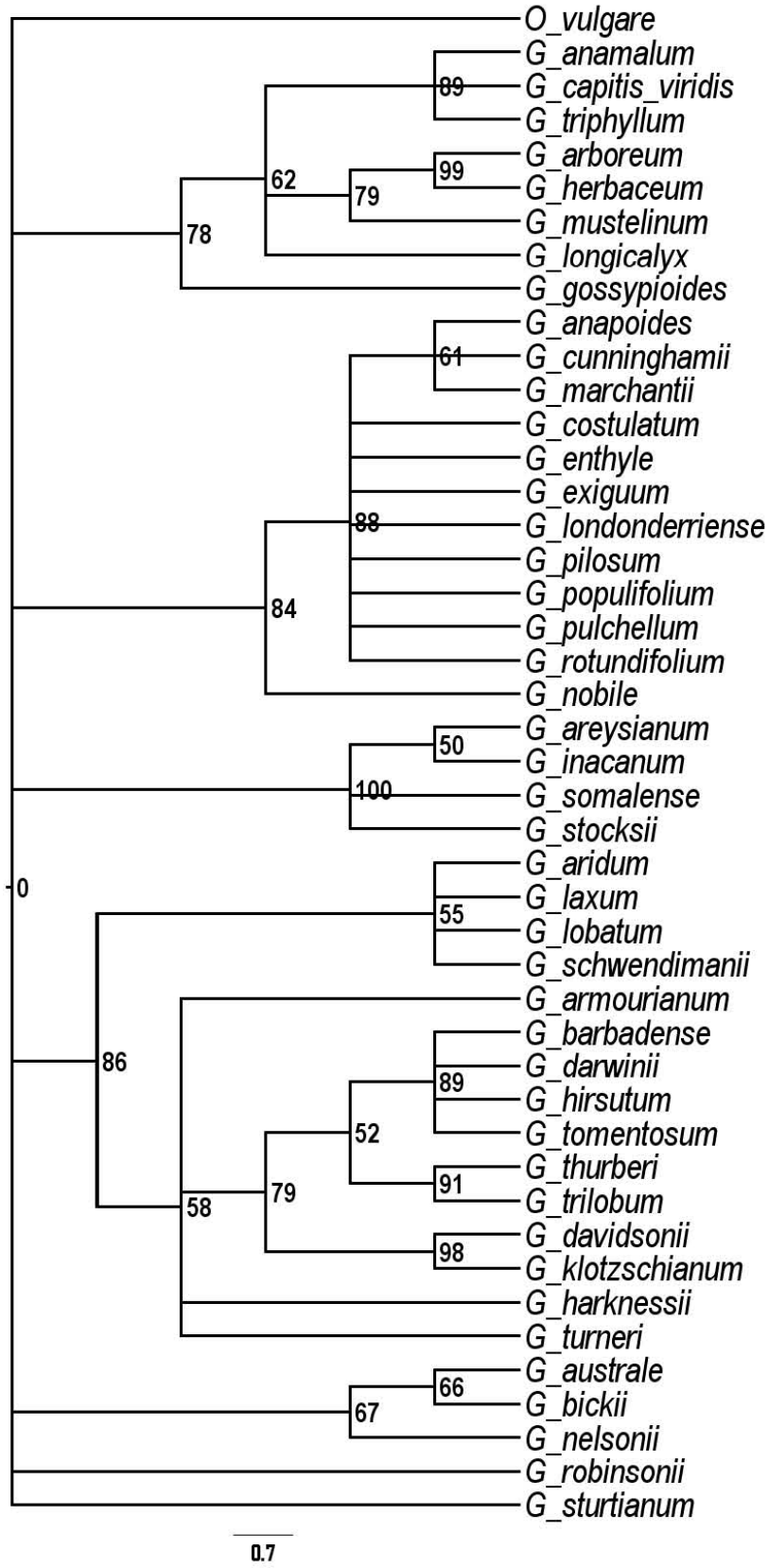
Şekil Ek-108. MUSCLE programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi (Nodlarda bulunan değerler bootstrap değerleridir)



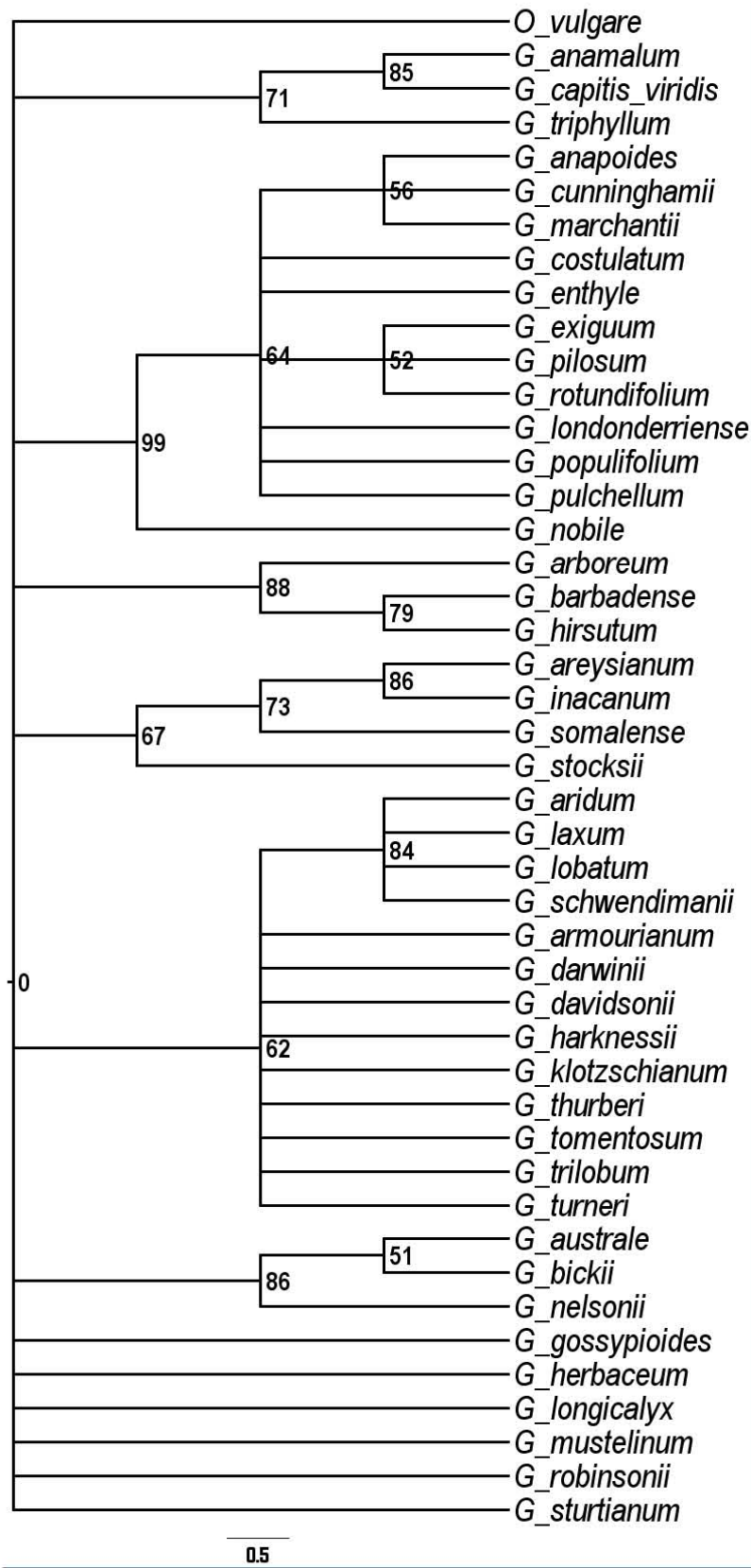
Şekil Ek-109. T-COFFEE programının 1. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



Şekil Ek-110. T-COFFEE programının 1. parametresi ile hizalanmış ITS2 sekansının parsimony analizi



Şekil Ek-111. T-COFFEE programının 2. parametresi ile hizalanmış ITS1 sekansının parsimony analizi



Şekil Ek-112. T-COFFEE programının 2. parametresi ile hizalanmış ITS2 sekansının parsimony analizi

ÖZGEÇMİŞ

1985 yılında Şırnak'ta doğdu. İlk ve lise öğrenimini Şırnak'ta tamamladı. 2005 yılında girdiği ÖSS sınavında Akdeniz Üniversitesi Fen Edebiyat Fakültesi Biyoloji Bölümüne yerleşti. 2009'da Akdeniz Üniversitesi Fen Edebiyat Fakültesi Biyoloji bölüm birincisi olarak mezun oldu. 2009 yılında Akdeniz Üniversitesi Ziraat Fakültesi Tarla Bitkileri Anabilim Dalında yüksek lisansa başladı. 2011 Mart ayında Siirt Üniversitesi Fen Edebiyat Fakültesi Biyoloji Bölümüne Araştırma Görevlisi olarak atandı. Halen Siirt Üniversitesi Fen Edebiyat Fakültesi Biyoloji Bölümünde Araştırma Görevlisi olarak çalışmaktadır.