SCIENTIA PLENA
www.scientiaplena.org.br
VOL. 16, NUM. 6 2020
doi: 10.14808/sci.plena.2020.060202

Comparative genomics plastomes of the Amaryllidaceae family species

Genômica comparativa de plastomas de espécies da família Amaryllidaceae

H. J. Jimenez[1]; A. D. F. da Silva[1*]; L. S. S. Martins[2]; R. de Carvalho[2]; R. M. de Moraes Filho[1]

[1]*Programa de Pós Graduação em Melhoramento Genético de Plantas, Departamento de Agronomia, Universidade Federal Rural de Pernambuco, 52171-900, Recife – PE, Brasil.*
[2]*Programa de Pós Graduação em Melhoramento Genético de Plantas, Departamento de Biologia, Universidade Federal Rural de Pernambuco, 52171-900, Recife – PE, Brasil.*

** a.deyws09@gmail.com*

(Recebido em 07 de fevereiro de 2020; aceito em 19 de junho de 2020)

The genus *Allium* covers more than 800 species, signaling among the largest among monocotyledons. The genus contains many economically important species, including garlic, leeks, onions, chives and Chinese chives. Due to the high conservation of chloroplast genomes compared to nuclear genomes and mitochondrial genome, sequence of chloroplasts in Amaryllidaceae have been consistently used for species identification and various *in silico* programs and strategies have been used to identify, characterize and compare plastid genome regions. Plastome from 15 species of the Amaryllidaceae family revealed similarity in both sequences and in the organization of their gene regions. The base pairs (bp) number ranged from 145,819 (*A. paradoxum*) to 159,125 (*A. ursinum*). In respect the GC content, the species presented a variation between 36.7% (*A. schoenoprasum* and *A. sativum*) and 37.5% (*A. coddii*) and the gene space ranged from 84.760 (*A. paradoxum*) to 94.766 (*A. sativum*). The genes that encode proteins showed values between 78 (*A. paradoxum*) to 89 (*A. cepa*). Phylogenetic trees acquired through alignment of complete plastomas and the plastidial *matK* gene revealed similarity to the proposed classification for the family. For the genus *Allium*, there was the formation of three clades with perfect correspondence of the clusters to the three evolutionary lines of the genus.
Keywords: Phylogeny, *Allium*, Plastome.

O gênero *Allium* abrange mais de 800 espécies, uma das maiores entre as monocotiledôneas. O gênero contém muitas espécies economicamente importantes, incluindo alho, alho-poró, cebola, cebolinha e cebolinha chinesa. Devido á alta conservação dos genomas dos cloroplastos em comparação com os genomas nucleares e mitocondriais, sequencia de genomas de cloroplastos em Amaryllidaceae tem sido constantemente utilizada para identificação de espécies. Vários programas e estratégias *in silico* tem sido utilizados no sentido de identificar, caracterizar e comparar regiões do genoma plastidial. Plastomas de 15 espécies da família Amaryllidaceae revelaram similaridade tanto em sequencias quanto na organização de suas regiões gênicas. O número de pares de base (pb) variou de 145.819 (*A. paradoxum*) a 159.125 (A. *ursinum*). Com relação ao conteúdo GC, as espécies apresentaram uma variação entre 36,7% (A. *schoenoprasum* e A. *sativum*) e 37,5% (A. *coddii*) e o espaço gênico variou de 84.760 (A. *paradoxum*) a 94.766 (A. *sativum).* Os genes que codificam proteínas apresentaram valores entre 78 (*A. paradoxum) a 89 (Allium cepa).* As árvores filogenéticas adquiridas através do alinhamento dos plastomas completos, e do gene plastidial matK, revelaram semelhança com a proposta de classificação para a família. Para o gênero *Allium*, Houve a formação de três clados com perfeita correspondência dos agrupamentos às três linhas evolutivas do gênero.
Palavras Chave: Filogenia, *Allium*, Plastoma.

1. INTRODUCTION

The Amaryllidaceae family is represented by about 80 genus with approximately 1600 species widely distributed in tropical and subtropical regions [1, 2]. In this family the *Allium* genus (subfamily Allioideae) is one of the largest genera of monocotyledons comprising more than 750 species [3, 4], which are distributed almost exclusively in the Northern Hemisphere. It is widely distributed in nature, and adapted to different habitats in all regions [4]. Its main diversity center is located between the southwest and Central Asia and the Mediterranean region, which must also be

the main center of diversification of *Allium*, in addition to a second one existing in North America [5]. Most species produce remarkable amounts of cysteine sulfoxides, causing the specific smell and taste of onion and garlic [6].

The *Allium* genus is currently divided into 15 subgenres and 72 sections [6]. The classification of the genus has proved to be very difficult, where many ambiguities remain in the phylogeny of *Allium* [7, 8]. In addition, there is significant morphological diversity at the intraspecific level in species such as *Allium cepa*, *A. sativum* and *A. porrum*. These variations must be measured at the molecular level for their proper characterization, which will be beneficial for future breeding programs. The internal spacer of transcribed nuclear DNA (ITS) and several regions of the plastidial genome (*trnL - trnF*, *matK*, *rbcl* and *rpl16*) have been frequently used for phylogenetic analysis of the species. A first study of the *Allium* genus by molecular markers was carried out by Linne et al. (1996) [9], where it was possible to confirm the subgeneric classification based on the association of morphological and other methods, but it was found that the subgenus *Amerallium* were not distinguished. Currently, molecular studies have concentrated on the classification and phylogeny of the entire genus *Allium* [7, 6] or exclusive subgenus such as *Amerallium* [4], *Melanocrommyum* [10] and *Rhizirideum* [4]. Other authors have focused on the origins and evolution of the main species of *Allium* [4, 11, 12], the phylogeny of the *Cepa* (Mill.) section [13] and the phylogenetic configuration of Western North American species [14].

Researchers have sought to use complete sequences of the chloroplast genome to obtain information about plants, including examining phylogenetic relationships. Phylogenetic analysis, however, requires substantial taxa sampling, and the use of whole genomes to infer phylogeny has been limited by the lack of sequenced complete genomes [15, 16]. But with the emergence of relatively fast and cheaper cloning and sequencing techniques [17, 18], we have seen a recent wave of sequenced plastid genomes. This rapid growth in the availability of complete chloroplast genome sequences has provided a wealth of new data for phylogenetic analyzes between species.

Chloroplasts are essential organelles in plant cells and play a crucial role in maintaining life [19]. Chloroplast genomes are mainly inherited from the maternal parent. The cp genome has a circular double-stranded molecular structure; a length of 120-220kb; and 120-140 protein coding genes [20]. The quadripartite structure of the chloroplast genome contains a single large copy region (LSC), a single small copy region (SSC) and two copies of an inverted repeat region (IRA and IRB) [20]. Due to the high conservation of chloroplast genomes compared to nuclear and mitochondrial genomes, sequences of chloroplast genomes have been constantly used for phylogenetic studies and species identification [21]. Several *in silico* strategies and programs have been used in order to identify, characterize and compare regions of the plastidial genome for phylogenetic analyzes of the studied species [22, 23]. Together, this information can help the genetic improvement programs of the culture, either through conventional methods or bioengineering.

In this context, plastidial DNA sequences of 14 species belonging to the genus *Allium*, available in public databases, were analyzed based on *in silico* tools with the following objectives: 1) to perform a comparative analysis among the plastid genomes of *Allium* species, based on the survey of complete plastidial genome sequences in the Gen Bank database; 2) perform a phylogenetic reconstruction of *Allium* species, based on plastidial genome sequences to understand the relationships among species in the group; 3) perform phylogenetic analysis based on a plastidial gene (*matK*) for species of the Amaryllidaceae family and understand their relationships at the subfamily level.

2. MATERIAL AND METHODS

2.1. Recovery and characterization of chloroplast genome sequences from the Amaryllidaceae family species

The chloroplast genome sequences of the species were retrieved from the NCBI (National Center for Biotechnology Information) database, in FASTA format, containing 15 species of the Amaryllidaceae and *Yucca filamentosa* (GenBank accession number (NC_035971) that belongs to the family Asparagaceae included as an outgroup.

2.2. Phylogenetic alignment and reconstruction

The plastomes were aligned using the mLAGAN algorithm based on the mVISTA server [24]. Standard parameters were applied, and the annotation structure of the A. cepa chloroplast genome was used as a reference. The percentage of identity between each plastome, all related to *A. cepa*, were later visualized through a VISTA graph [25]. The plastoma-based phylogeny was reconstructed for the fifteen species of Amaryllidaceae, using the total plastoma alignment generated by mLAGAN. Plastome of the species *Yucca filamentosa* (Asparagaceae) was also included as an outgroup. Using MEGA 7.0 software, the number of variable sites was calculated, according to the species *A. cepa* used as a reference. The linear representation of A. *cepa* plastoma was obtained by the server OGDRAW (Organelar Genome DRAW) [26].

The alignment was imported into MEGA7 software, version 7.0 (Molecular Evolutionary Genetics Analysis) [27], for the phylogenetic analysis using the Maxima Likelihood method (MPsearch level 3). The replacement model was the GTR + G type obtained by JmodelTEST [28]. Statistical support was obtained through bootstrap, using 1000 replicates. Bootstraps of 90-100 were considered strongly supported, 80-89 moderately supported and 50-79 poorly supported. For the phylogeny of the Amaryllidaceae family, based on the sequence of the plastidial gene *matK*, in addition o the fifteen sequences obtained from the complete genomes, seventy-three sequences were used, all belonging to the Amaryllidaceae families and three as an outgroup of the Xanthorrohoeaceae family (Asphodelus aestivus, Hemerocallis fulva and Hemerocallis dumortieri.

The sequences were aligned using the ClustalW algorithm and a phylogenetic tree was generated using the Beast software, using the Bayesian method. The replacement model was the GTR + G type obtained by JmodelTEST [28] and the statistical support was calculated using 1000 replicates.

3. RESULTS

Table 1 shows 15 species belonging to the Amaryllidaceae family with their information retrieved from the NCBI database. The number of base pairs (bp) ranged from 145,819 (*A. paradoxum*) to 159,125 (*A. ursinum*). Regarding the GC content, the species varied between 36.7% (*A. schoenoprasum* and *A. sativum*) and 37.5% (*A. coddii*) and the gene space varied from 84760 (*A. paradoxum*) to 94766 (*A. sativum*). The genes encoding proteins showed values between 78 (*A. paradoxum*) to 89 (*Allium cepa*).

Table 1. Amaryllidaceae family species and their respective characteristics obtained through the NCBI database.

Specie	Ref.seq	PB	GC %	Protein	rRNA	tRNA	Gene	Genic Space
Agapanthus coddii	NC_035971	157055	37,5	87	8	38	133	94443
Allium obliquum	NC_037199	152387	36,8	89	8	38	135	94575
Allium sativum	NC_031829	153131	36,7	89	8	38	135	94766
Allium prattii	NC_037432	154482	37	85	8	38	131	94446
Allium victorialis	NC_037240	154074	37	86	8	38	132	94407
Allium cepa	KM088013	153529	36,8	89	8	38	135	94536
Allium macleanii	LT699703	152633	36,9	84	8	37	129	94212
Allium fistulosum	LT674586	152862	36,9	83	8	39	130	94755
Allium schoenoprasum	LT699700	152806	36,7	81	8	39	128	94536
Allium nutans	LT799837	153456	36,9	83	8	39	130	94518
Allium cepa (N)	NC024813	153538	36,8	83	8	39	130	94542
Allium cepa (S)	KF728079	153355	36,8	83	8	39	130	94611
Allium ursinum	MH157875	159125	37,3	85	8	37	130	93410
Allium paradoxum	LT622239	145819	37,2	78	8	38	124	84760
Allium platyspatum	LT673892	152458	36,8	81	8	39	128	94524

3.1. Genomic comparison of the plastoma sequences of species of the Amaryllidaceae family

The number of variable sites in relation to *A. cepa* ranged from 451 to 8690. As shown in figure 1, the blue regions are the coding areas and the pink regions show the non-coding areas preserved according to the *Allium* species record strain, obtained through the mVISTA server. The rpoC2 and ycf1 genes showed the greatest points of divergence. The *Y. filamentosa* species was used as an outgroup and showed divergence in comparison with the other species presented. The plastomas organization similarity from 5 species of the Amaryllidaceae family is visible in the mVISTA graph.

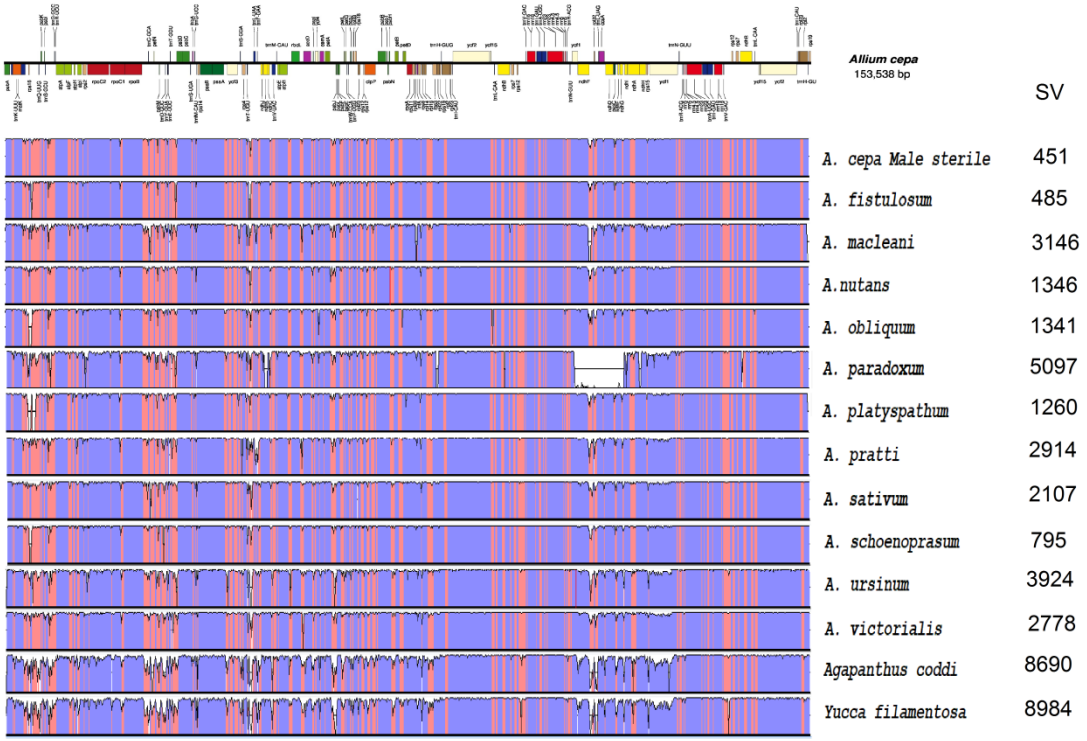


Figure 1. Alignment of complete plastoma sequences from 14 species of the Amaryllidaceae family. Blue regions identify coding regions according to the annotation for A. cepa, obtained by the mVISTA server. The linear representation of A. cepa plastoma was obtained by the server OGDRAW [26]). Pink regions represent conserved non-coding regions. The Y. filamentosa species was used as an outgroup. SV = Variable sites.

3.2. Phylogenetic analyzes

The phylogeny for Amaryllidaceae family species through the complete genome showed a well-supported monophyletic tree with a 100% robust *bootstrap* support (Figure 2). Three clades were formed, where the first clade covered the subgenus *Amerallium* (*A. ursinum* and *A. paradoxum*). The second clade (BS = 100%) included the subgenera *Anguinum* (*A. pratti* and *A. victorialis*) and *Melanocrommyum* (*A. macleani*). The third clade was the largest (BS = 100%) and covered the subgenus *Cepa* (*A. cepa*, *A. fistulosum* and *A. schoenoprasum*), *Polyprason* (*A. obliquum* and *A. platyspathum*), *Rhizidireum* (*A. nutans*) and *Allium* (*A. sativum*). The species of the outgroup *Y. filamentosa* formed a separate clade with 100% support (Figure 2).

The Amaryllidaceae family phylogeny through sequence of the plastidial gene *matK*, revealed a phylogenetic tree with the separation of the three subfamilies Agapanthoideae, Allioideae and Amaryllidoideae (Figure 3). The subfamily Agapanthoideae with 3 species represented with only 1 genus Agapanthus showed a support PP = 100%. The subfamily Allioideae comprised 22 species distributed between the Gilliesiae and Allieae tribes (PP = 100%). The Gilliesiae tribe represented by the genus *Leucocoryne*, *Nothoscordum* and *Tristagma* with 7 species (PP = 100%) and the Allieae tribe with only 1 genus *Allium* (PP = 100%). The subfamily Amaryllidoideae with 56 species distributed among 13 tribes (PP = 100). The Amaryllideae tribe with the genus *Nerine*,

Cybistetes, *Crinum*, *Boophone* and *Amaryllis* (PP = 100%). The Haemantheae tribe with the genus *Haemanthus*, *Scadoxus*, *Gethyllis*, *Apodolirion* and *Clivia* (PP = 100%). The Calostemmateae tribe with only 1 *Calostemma* genus (PP = 100%). The Hippeastreae tribe with the genus *Hippeastrum*, *Habranthus* and *Zephyranthes* (PP = 100%). The Eustephieae tribe with the genus *Chlidanthus* and *Eustephia* (PP = 100%). The Narcisseae tribe with the genus *Narcissus* (PP = 60%). The Galantheae tribe with the genus *Acis*, *Galanthus* and *Hannonia* (PP = 100%).

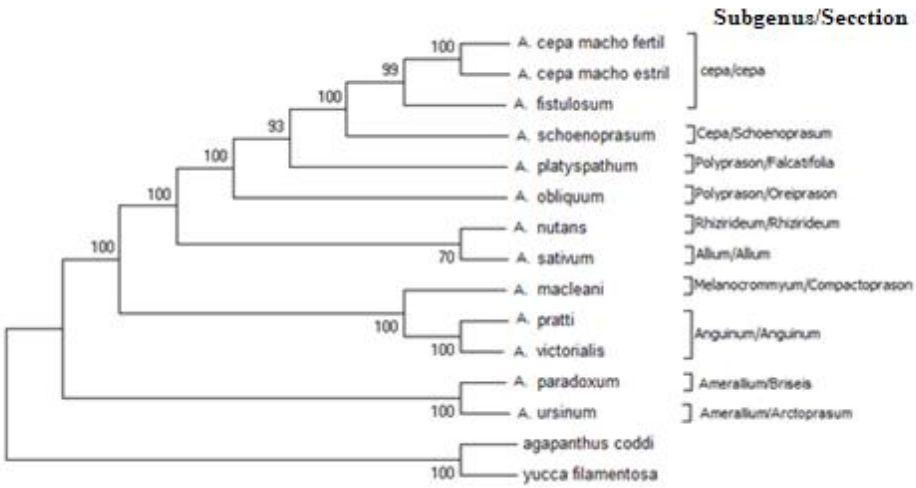
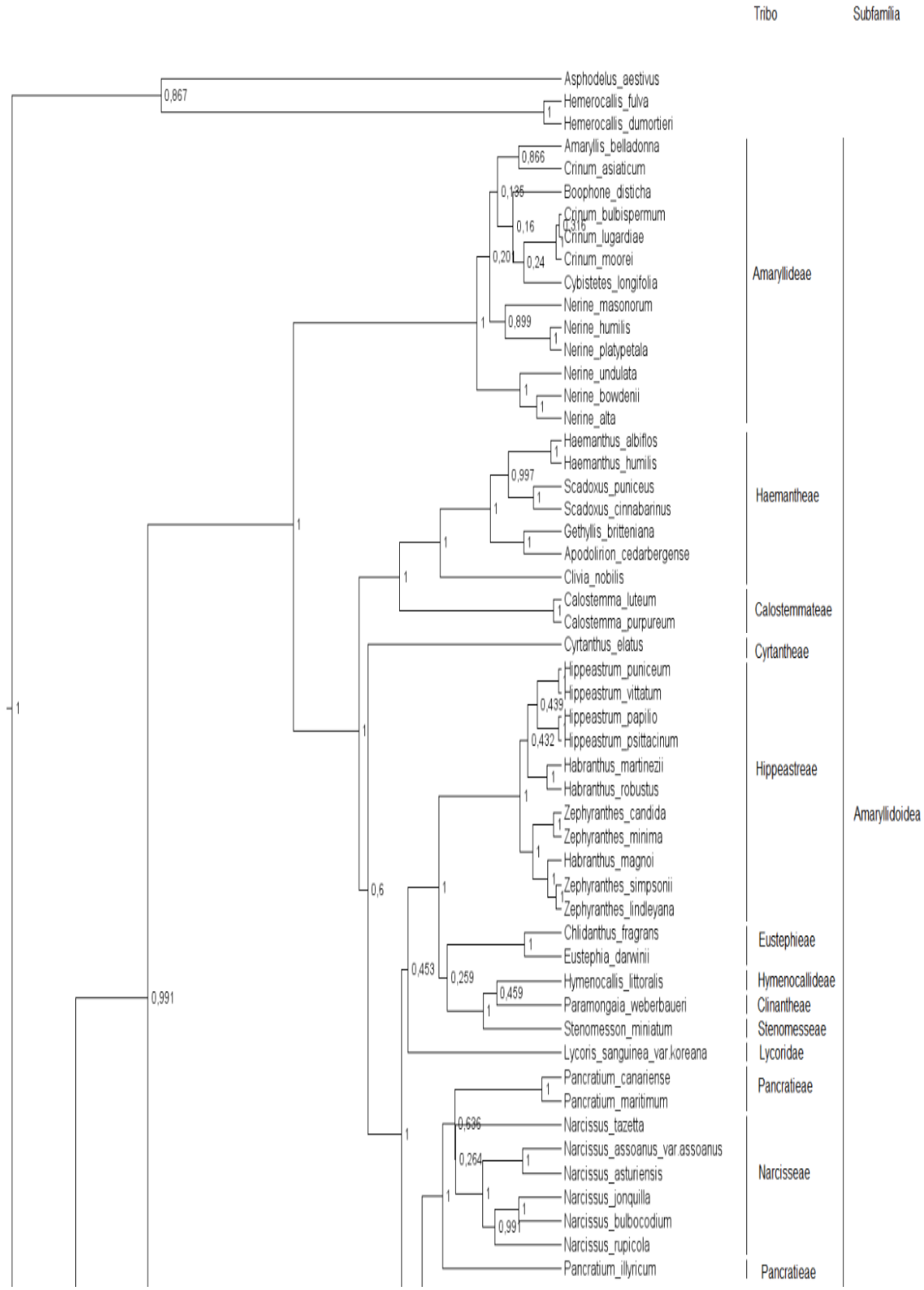


Figure 2. Allium genus phylogeny based on complete plastid genomes. Plastomas from A. obliquum (NC_037199.1), A. sativum (NC_031829.1), A. prattii (NC_037432.1), A. victorialis (NC_037240.1), A. cepa (KM088013.1), A. macleanii (LT699703.1), A. Fistulosum (LT674586.1), A. schoenoprasum (LT699700.1), A. nutans (LT799837.1), A. fertile male strain (NC_024813.1), A. sterile male strain (KF728079.1), A. ursinum (MH157875.1), A. paradoxum (LT622239.1) and A. platyspathum (LT673892.1). Yucca filamentosa specie (NC_032712.1) was included as an outgroup. The phylogenetic tree was generated by the MEGA7 software, using the Maximum Likelihood method, and the bootstrap support was calculated using 1000 replicates. Bootstrap support values (%) are shown at the intersection of each branch.


Tribo
Subfamilia
Asphodelus_aestivus
Hemerocallis_fulva
Hemerocallis_dumortieri
Amarylis_belladonna
Crinum_asiaticum
Boophone_disticha
Crinum_bulbispermum
Crinum_lugardiae
Crinum_moorei
Cybistetes_longifolia
Nerine_masonorum
Nerine_humilis
Nerine_platypetala
Nerine_undulata
Nerine_bowdenii
Nerine_alta
Amaryllideae
Haemanthus_albiflos
Haemanthus_humilis
Scadoxus_puniceus
Scadoxus_cinnabarinus
Gethyllis_britteniana
Apodolirion_cedarbergense
Clivia_nobilis
Haemantheae
Calostemma_luteum
Calostemma_purpureum
Calostemmateae
Cyrtanthus_elatus
Cyrtantheae
Hippeastrum_puniceum
Hippeastrum_vittatum
Hippeastrum_papilio
Hippeastrum_psittacinum
Habranthus_martinezii
Habranthus_robustus
Zephyranthes_candida
Zephyranthes_minima
Habranthus_magnoi
Zephyranthes_simpsonii
Zephyranthes_lindleyana
Hippeastreae
Chlidanthus_fragrans
Eustephia_darwinii
Eustephieae
Hymenocallis_littoralis
Hymenocallideae
Paramongaia_weberbaueri
Clinantheae
Stenomesson_miniatum
Stenomesseae
Lycoris_sanguinea_var.koreana
Lycoridae
Pancratium_canariense
Pancratium_maritimum
Pancratieae
Narcissus_tazetta
Narcissus_assoanus_var.assoanus
Narcissus_asturiensis
Narcissus_jonquilla
Narcissus_bulbocodium
Narcissus_rupicola
Narcisseae
Pancratium_illyricum
Pancratieae
Amaryllidoidea
0,867
0,866
0,135
0,16
0,315
0,201
0,24
0,899
0,997
0,439
0,432
0,6
0,453
0,259
0,459
0,991
0,636
0,264
0,99

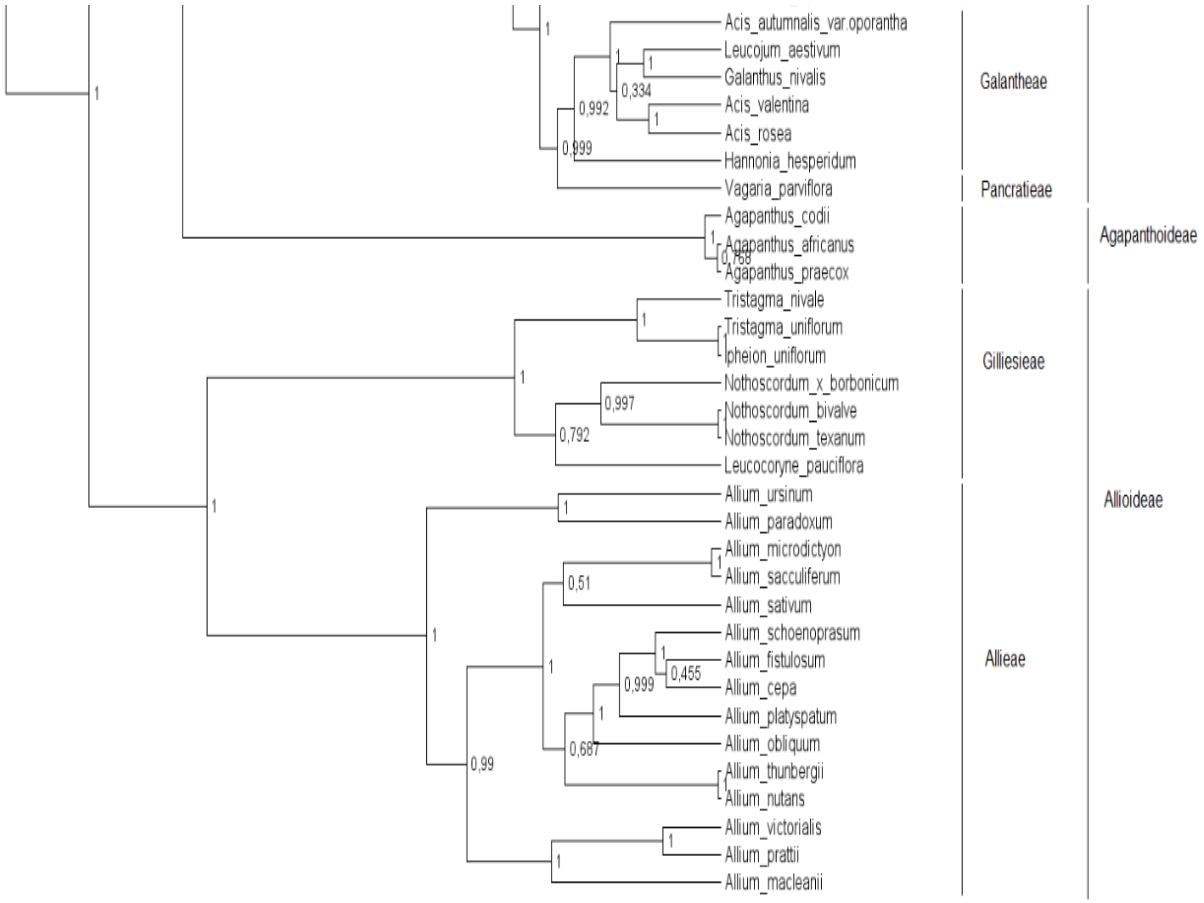


Figure 3. Amaryllidaceae family phylogeny based on sequences of the plastidial gene matK. The phylogenetic tree was generated by the Beast software, using the Bayesian method and the statistical support was calculated using 1000 replicates. PP support values (%) are shown at the intersection of each branch.

4. DISCUSSION

The genome size (bp) from the analyzed species ranged from 145,819 to 161,172 bp and was similar to those found in monocotyledons phylogenetically close to Amaryllidaceae as species of the genus *Polygonatum* (153,821 - 155,580 bp) [29] and *Asparagus officinalis* (156,699 bp) [30], both from the family Asparagaceae and *Iris gatesii* (153,441) [31] from the family Iridaceae. Genomes size variation in Angiosperm have been suggested as a common evolutionary phenomenon caused by contractions or expansions in the quadripartite structure of the chloroplast genome [32].

Plastid genomes showed GC% content values (Table 1) very close to other monocotyledons representatives, as in the study by Sheng et al. (2017) [30] with *Asparagus officinalis* (GC = 37.76%) and Wilson et al. (2014) [31] with *Iris gatesii* (CG = 39.7%). The low GC content is a significant feature of the plastid genomes, which are formed after endosymbiosis by DNA replication and repair [33]. The number of genes encoding proteins was close to that found by Floden and Schilling (2018) [29] in Polygonateae, where 128 genes were observed with 75 genes encoding proteins, 4 rRNA genes and 31 tRNA genes and by Sheng et al. (2017) [30] in *Asparagus officinalis*, where the author observed a total of 136 predicted genes in the genome including 78 protein coding genes, 30 tRNA genes and 4 rRNA genes.

Multiple complete genomes from accessible Amaryllidaceae chloroplast offer an opportunity to compare sequence variation within the family at the genome level. The *A. cepa* genome was similar to the other species of the Amaryllidaceae family, indicating that the Amaryllidaceae chloroplast genomes are quite conserved, with identical contents and orders, although some divergent regions

are found between these genomes. The points of divergence occurred mostly in non-coding regions, with the *rpoC2* and *ycf1* genes showing the greatest points of divergence. These results were very similar to those found by Nie et al. (2012) [34] and Eguiluz et al. (2017) [23], where it was confirmed that the location of points of divergence occurs mostly in intergenic regions. It is also known that non-chloroplast regions are competent molecular markers for phylogenetic studies in angiosperms [35] and that these regions are associated with repetitive sequences [36]. It is possible that repeated sequences also correlate with the genomic rearrangement in Amaryllidaceae genomes.

The Amaryllidaceae taxonomic limits have varied a lot over the last decades, causing intense debates among taxonomists about the genus that comprise it. Phylogenetic trees based on the chloroplast genome for species of the *Allium* genus revealed monophyletism with well-resolved relationships between their species and completely separated from the outgroup (BS = 100%). There was a formation of three main clades in phylogeny where each clade has been indicated as an important evolutionary line for the formation of the *Allium* genus. Li et al. (2010) [4] analyzed phylogeny and biogeography of *Allium* (Allieae) based on *ITS* and *rps16* markers and found very similar results.

The first clade was represented by the species of the subgenus *Amerallium* (*A. ursinum* and *A. paradoxum*) that indicate being part of the first evolutionary line of the genus. *Amerallium* is monophyletic, being extremely diverse morphologically and ecologically [4]. The species in this clade are located in three geographic groups: one containing species of *Allium* from North America (New World) and the rest comprising two smaller groups from the Mediterranean region and East Asia (Old World) [4]. The subgenus is characterized by having narrow, elongated bulbs, smooth and flat leaves with a single row of vascular bundles and subglobous seeds [37].

The second clade included representatives of the subgenus *Anguinum* (*A. pratti* and *A. victorialis*) and *Melanocrommyum* (*A. macleani*) who suggest they are part of the second evolutionary line of the genus. Species of the subgenus *Anguinum* have an area of occurrence in southwestern Europe, eastern Asia and northeastern North America [7], and have particular root anatomical characters [38], leaf and bulb organization [39]. The species of the subgenus Melanocrommyum occurs close to the Mediterranean and Middle East [40, 41], being characterized by presenting very advanced leaf sheaths, with very short development time and by having several anatomical properties [38].

In the third clade Li et al. (2010) [4] obtained the same species representing the subgenus Cepa (*A. cepa*, *A. fistulosum* and *A. schoenoprasum*), *Polyprason* (*A. obliquum* and *A. platyspathum*), *Rhizidireum* (*A. nutans*) and *Allium* (*A. sativum*), being part of the third evolutionary line of the genus. *Rhizidireum* representatives are diverse in southern Siberia and Mongolia [38]. Species of the subgenus *Allium*, *Cepa* and *Polyprasum* comprise the largest clade in the third evolutionary line, where some studies indicate that these subgenus are not monophyletic, with the systematic position of some species having to be reviewed [4].

Phylogeny based on the *matK* marker for species in the Amaryllidaceae family revealed well-resolved relationships among their species and completely separated from the outgroup (PP = 100%) with the separation of the three subfamilies: Agapanthoideae, Allioideae and Amaryllidoideae. By unsing the *matK* gene in the subfamily Amaryllidoideae, it was possible to resolve the Calostemmateae and Haemantheae tribe as sisters (PP = 100%) and it was not possible to establish the relationship of these two tribes with the Cyrtantheae tribe. The Amaryllidoideae subfamily taxonomy has been discussed by several authors, where Meerow et al. (2006) [42] and Ronsted et al. (2012) [43] working with phylogeny in Amaryllidaceae with ITS and *ndhF* markers suggested placing the Calostemmateae and Haemantheae tribe as sisters of the Cyrtantheae tribe. However, in terms of morphology, there may be some questioning about this proximity, since the indehiscent capsule of Calostemmateae has more similarity with the indehiscent fruit of Haemantheae than with the dehiscent capsule of Cyrtanthus [43].

The Amaryllideae tribe (PP = 100%) in this work is recommended as sister group of the others Amaryllidoideae (PP = 100%), corroborating the results presented by Ronsted et al. (2012) [43], where the author obtained a high Bayesian (PP = 100%) and Boostrap (BS = 100%) support to resolve the African tribe Amaryllideae as sister of the Amaryllidoideae. The clade encompassing the Hippeastreae tribe with the genus *Hippeastrum*, *Habranthus* and *Zephyranthes*, Eustephieae with the *Chlidanthus* and *Eustephia* genus, Hymenocallideae with the *Hymenocallis* genus,

Lycoridae with the *Lycoris* genus, Clinanthomen with the s *Clinanthus* genuwith a low support (PP = 49%), therefore it is still necessary to carry out further analyzes to elucidate the phylogenetic relationships in the Hippeastreae tribe.

5. CONCLUSION

The phylogenetic trees for Amaryllidaceae species showed similarity with the proposed classification for the family. This recommendation is based on both phylogeny based on complete plastoma alignment and phylogeny based on the *matK* gene.

The Amaryllidaceae chloroplast genomes are very conserved, as the *A. cepa* genome showed similarity in comparison with the other family species.

6. REFERENCES

1. Souza VC, Lorenzi H. Botânica Sistemática: guia ilustrado para identificação das famílias de Fanerógamas nativas e exóticas no Brasil, baseado em APG III. 3ª ed. Nova Odessa: Instituto Plantarum; 2012. 768 p.
2. Stevens PF. Angiosperm Phylogeny Website, Version 14. Acesso em: Jan 2019, Disponível em: http://www.mobot.org/MOBOT/research/APweb/>.
3. Seregin AP. *Allium marmoratum* (Amaryllidaceae), a new species of section *Falcatifolia* from Chimgan Massif, Eastern Uzbekistan. Phytotaxa. 2015 Abr;25(3):211-214, doi:10.11646/phytotaxa.205.3.9
4. Li QQ, Zhou SD, He XJ, Yu Y, Zhang YC, Wei XQ. Phylogeny and biogeography of *Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast rps16 sequences, focusing on the inclusion of species endemic to China. Ann Bot. 2010 Nov;106(5):709-733, doi:10.1093/aob/mcq177
5. Choi HJ, OH BU. A partial revision of *Allium* (Amaryllidaceae) in Korea and north-eastern China. Bot J Linn Soc. 2011 Out;167(2):153-211, doi: org/10.1111/j.1095-8339.2011.01166.x
6. Friesen N. Fritsch RM. Blattner FR. Phylogeny and new intrageneric classification of *Allium* (Alliaceae) based on nuclear ribosomal DNA ITS sequences. Aliso. 2006 Abr;22(1):372-395, doi:10.5642/aliso.20062201.31
7. Fritsch RM. *Allium* crop science: recent advances. 1st ed. Wallingford (United Kingdom): CABI Publishing; 2002. Chapter 2, Evolution, domestication and taxonomy; p. 5-30.
8. Gurushidze M, Fritsch RM, Blattner FR. Phylogenetic analysis of *Allium* subgen. *Melanocrommyum* infers cryptic species and demands a new sectional classification. Mol Phylogenet Evol. 2008 Dez; 49(3):997-1007, doi: 10.1016/j.ympev.2008.09.003
9. Linne von Berg G, Samoylov A, Klaas M, Hanelt P. Chloroplast DNA restriction analysis and the infrageneric grouping of *Allium* (Alliaceae). Plant Syst Evol. 1996;200:253-261.
10. Fritsch RM, Blattner FR, Gurushidze M. New classification of *Allium* L. subg. *Melanocrommyum* (Webb & Berthel) Rouy (Alliaceae) based on molecular and morphological characters. Phyton. 2010;Jan;49(2):145-220.
11. Friesen N, Pollner S, Bachmann K, Blattner FR. RAPDs and noncoding chloroplast DNA reveal a single origin of the cultivated *Allium fistulosum* from *A. altaicum* (Alliaceae). Am J Bot. 1999 Abr;86(4):554-562.
12. Li MJ, Tan JB, Xie DF, Huang DQ, Gao YD, He XJ. Revisiting the evolutionary events in *Allium* subgenus *Cyathophora* (Amaryllidaceae): insights into the effect of the Hengduan Mountains Region (HMR) uplift and Quaternary climatic fluctuations to the environmental changes in the Qinghai-Tibet Plateau. Mol. Phylogenet. Evol. 2016 Jan;94:802-813, doi: 10.1016/j.ympev.2015.10.002.
13. Gurushidze M, Mashayekhi S, Blattner FR, Friesen N, Fritsch RM. Phylogenetic relationships of wild and cultivated species of *Allium* section *Cepa* inferred by nuclear rDNA ITS sequence analysis. Plant Syst Evol. 2007 Dez; 269(3-4):259-269.
14. Nguyen NH, Driscoll HE, Specht CD. A molecular phylogeny of the wild onions (*Allium*; Alliaceae) with a focus on the western North American center of diversity. Mol Phylogenet Evol. 2008 Jun; 47(3):1157-1172, doi: 10.1016/j.ympev.2007.12.006
15. Soltis DE, Albert VA, Savolainen V, Hilu K, Qiu Y-Q, Chase MW, Farris JS, Stefanović S, Rice DW, Palmer JD, Soltis PS. Genome-scale data, angiosperm relationships, and 'ending incongruence': a cautionary tale in phylogenetics. Trends Plant Sci. 2004 Out;9(10):477-483, doi: 10.1016/j.tplants.2004.08.008.

16. Jansen, RK Cai Z. Raubeson LA. Daniell H. Depamphilis CW. Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S-B, Peery R, McNeal JR, Kuehl JV, Boore JL. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA. 2007 Dez;104(49):19369-19374, doi: 10.1073 / pnas.0709121104
17. Jansen RK, Raubeson LA, Boore JL, De Pamphilis CW, Chumley TW, Haberle RC, Wyman SK, Alverson AJ, Peery R, Herman SJ, Fourcade HM, Kuehl JV, McNeal JR, Leebens-Mack J, Cui L. Methods for obtaining and analyzing whole chloroplast genome sequences. Methods Enzymol. 2005 Abr;395:348-384, doi:10.1016/S0076-6879(05)95020-9
18. Barrett CF, CD Bacon, Antonelli A, Cano A, Hofmann T. An introduction to plant phylogenomics with a focus on palms. Bot J Linn Soc. 2016 Out;182(2): 234-255, doi:org/10.1111/boj.12399
19. Xiong AS. Peng RH. Zhuang J. Gao F. Zhu B, Fu XY, Xue XY, Jin XF, Tian YS, Zhao W, Yao QH. Gene duplication, transfer, and evolution in the chloroplast genome. Biotechnol Adv. 2009 Jul-Ago;27(4):340-347, doi:10.1016/j.biotechadv.2009.01.012
20. Rogalski M, Vieira LN, Fraga HP, Guerra MP. Plastid genomics in horticultural species: importance and applications for plant population genetics, evolution, and biotechnology. Front Plant Sci. 2015 Jul;6:586, doi: 10.3389/fpls.2015.00586
21. Song Y, Dong W, Liu B, Xu C, Yao X, Gao J, Corlett RT. Comparative analysis of complete chloroplast genome sequences of two tropical trees *Machilus yunnanensis* and *Machilus balansae* in the family Lauraceae. Front Plant Sci. 2015 Ago;6:662, doi: 10.3389 / fpls.2015.00662
22. Morton CM, Telmer C. New subfamily classification for Rutaceae. Ann Missouri Bot. 2014;99:620-641.
23. Eguiluz M, Rodrigues NF, Guzman F, Yuyama P, Margis R. The chloroplast genome sequence from *Eugenia uniflora*, a Myrtaceae from Neotropics. Plant Syst. Evol. 2017 Nov;303(9):1199-1212, doi: 10.1007/s00606-017-1431-x
24. Brudno M, Do CB, Cooper GM, KIm MF, Davydov E. LAGAN and Multi LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. Genome Res. 2003;13:721-731, doi: 10.1101/gr.926603
25. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 2004;32:273-279.
26. Lohse MD, Oliver KS. Organellar Genome DRAW - a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic acids research. 2013;41. doi: 10.1093/nar/gkt289
27. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33:1870-1874.
28. Posada D. jModelTest: Phylogenetic Model Averaging. Mol Biol Evol. 2008 Jul;25(7):1253-1256, doi:org/10.1093/molbev/msn083
29. Floden A, EE Schilling. Using phylogenomics to reconstruct phylogenetic relationships within tribe Polygonateae (Asparagaceae), with a special focus on Polygonatum. Mol Phylogenet Evol. 2018 Dez; 129:202-213, doi: org/10.1016/j.ympev.2018.08.017
30. Sheng W, Chai X, Rao Y, Tu X, Du S. 2017. The Complete Chloroplast Genome Sequence of Asparagus (*Asparagus officinalis* L.) and its Phylogenetic Positon within Asparagales. Int J Plant Biol Res. 2017;5(3):1075.
31. Wilson CA. The complete plastid genome sequence of Iris gatesii (Section Oncocyclus), a bearded species from southeastern Turkey. Aliso. 2014 Jun;32(1):47-54, doi: 10.5642/aliso.20143201.03
32. Davis JI, Soreng RJ. Migration of endpoints of two genes relative to boundaries between regions of the plastid genome in the grass family (Poaceae). Am J Bot. 2010;97:874-892, doi:10.3732/ajb.0900228
33. Howe CJ, Barbrook AC, Koumandou VL, Nisbet RE, Symington HA, Wightman TF. Evolution of the chloroplast genome. Philos Trans R Soc Lond B Biol Sci. 2003 Jan;358(1429):99-106.
34. Nie X, Lv S, Zhang Y, Du X, Wang L, Biradar SS, Tan X, Wan F, Weining S. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). PloS One. 2012 May;7(5):e36869, doi: 10.1371/journal.pone.0036869
35. Wu F-H, Chan M-T, Liao D-C, Hsu C-T, Lee Y-W, Daniell H, Duvall MR, Lin C-S 2010. Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. BMC Plant Biol. 2010 Abr;10(68):1-12, doi:10.1186/1471-2229-10-68
36. Yang JB, Tang M, Li HT, Zhang ZR, Li DZ. Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evol Biol. 2013 Abr;13:84, doi:10.1186/1471-2148-13

37. Blazewicz-Woźniak M, Michowska A. The growth, flowering and chemical composition of Leaves of three ecotypes of *Allium ursinum* L. Acta Agrobot. 2011;64(4):171-180.
38. Fritsch RM. The genus *Allium* – Taxonomic problems and genetic resource. 1st ed. Gatersleben (Germany): IPK - Proceedigns of International Symposium; 1992. Chapter, Septal nectaries in the genus *Allium* L; p. 77-85.
39. Pastor J, Valdes B. Bulb structure in some species of *Allium* (*Liliaceae*) of the Iberian Peninsula. Ann Mus Goulandris. 1985;7:249-262.
40. Khassanov FO, Fritsch RM. New taxa in *Allium* L. subg. *Melanocrommyum* (Webb & Berth.) Rouy from Central Asia. Linzer Biol Beitr. 1994 Dez;26(2):965-990.
41. Mes THM, Fritsch RM, Pollner S, Bachmann K. Evolution of the chloroplast genome and polymorphic ITS regions in *Allium* subg. *Melanocrommyum*. Genome. 1999 Abr;42(2):237-247, doi:10.1139/g98-123
42. Meerow A, Francisco-Ortega J, Kuhn D, Schnell R. Phylogenetic Relationships and Biogeography within the Eurasian Clade of Amaryllidaceae Based on Plastid *ndh*F and nrDNA ITS Sequences: Lineage Sorting in a Reticulate Area? Syst Bot. 2006 Jan;31(1):42-60, doi.org/10.1600/036364406775971787
43. Rønsted N, Symonds, MRE, Birkholm T, Christensen SB, Meerow AW, Molander M, Mølgaard P, Petersen G, Rasmussen N, Van Staden J, Stafford GI, Jäger AK. Can phylogeny predict chemical diversity and potential medicinal activity of plants? A case study of Amaryllidaceae. BMC Evol Biol. 2012 Set;12:182, doi: org/10.1186/1471-2148-12-182