UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

AVALIAÇÃO DE GENES NUCLEARES PARA FILOGENIA E PLASTOMA DE LENTIBULARIACEAE

Afonso Pinto Fançony

Engenheiro Agrônomo

2024

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

AVALIAÇÃO DE GENES NUCLEARES PARA FILOGENIA E PLASTOMA DE LENTIBULARIACEAE

Afonso Pinto Fançony

Orientador: Prof. Dr. Vitor Fernandes Oliveira de Miranda

Coorientadora: Dra. Saura Rodrigues da Silva

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Genética e Melhoramento de Plantas).

2024

F199a

Fançony, Afonso Pinto

Avaliação de genes nucleares para filogenia e plastoma de Lentibulariaceae / Afonso Pinto Fançony. -- Jaboticabal, 2024

81 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Vitor Fernandes Oliveira de Miranda

Coorientadora: Saura Rodrigues da Silva

1. Lentibulariaceae. 2. Filogenia. 3. Genlise-Utricularia. 4. Plastoma de Pinguicula. 5. Genes nucleaes. I. Título.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: AVALIAÇÃO DE GENES NUCLEARES PARA FILOGENIA E PLASTOMA DE LENTIBULARIACEAE

AUTOR: AFONSO PINTO FANÇONY
ORIENTADOR: VITOR FERNANDES OLIVEIRA DE MIRANDA
COORIENTADORA: SAURA RODRIGUES DA SILVA

Aprovado como parte das exigências para obtenção do Título de Mestre em Agronomia (Genética e Melhoramento de Plantas), pela Comissão Examinadora:

Prof. Dr. VITOR FERNANDES OLIVEIRA DE MIRANDA (Participaçao Presencial)
Departamento de Biologia / FCAV UNESP Jaboticabal

Dr. GUILHERME CAMARA SEBER (Participaçao Presencial)
Seber Ltda - Comércio de Café Brasil / Torrinha/SP

Prof. Dr. DANIEL GUARIZ PINHEIRO (Participaçao Presencial)
Departamento de Biotecnologia Agropecuaria e Ambiental / FCAV / UNESP Jaboticabal

Jaboticabal, 29 de fevereiro de 2024

DADOS CURRICULARES DO AUTOR

AFONSO PINTO FANÇONY, nascido em 30 de janeiro de 1991 na cidade no Município do Amboim, Província do Cuanza Sul (Angola), filho de Augusto Fançony e Mecilia Pinto. É Engenheiro Agrônomo graduado pelo Instituto Superior Politécnico do Cuanza Sul (Angola), com ingresso em março de 2012 e término em agosto de 2017. Em agosto de 2021 ingressou no Programa de Pós-graduação em Agronomia (Genética e Melhoramento de Plantas), nível mestrado, área de concentração em Genética e Biologia Evolutivas, pela Faculdade de Ciências Agrárias e Veterinárias (FCAV) da Universidade Estadual Paulista "Júlio de Mesquita Filho" – Unesp, Câmpus de Jaboticabal.

Okwijiva kwafumbu okwaha

Tradução: O saber suplanta a beleza

Interpretação: Não basta ter beleza é preciso lutar pela aprendizagem para aumentar a sua elegância

Ka mwinle ongongo ka kolele

Tradução: Quem não sofreu não amadureceu.

Interpretação: A aprendizagem, exige sacrifício

MUEZU UA MUADIAKIMI, A-U-SUNGA NI NDUNGE

Tradução: Barbas de homem respeitável, com jeito se puxam

Interpretação: Com brandura tudo se consegue, basta só o desejo e o sacrifício para ir ao encontro das coisas.

(*In* provérbios angolanos)

À minha recente falecida avó, **Isabel António**, mãe de minha mãe,
pelos ensinamentos e conselhos transmitidos e que não pude estar
com ela nos últimos dias de sua vida, para ela, eu
OFEREÇO

À **Aziela Alberto Fançon**y, mãe do meu pai,
que muito precisou e precisa de mim, a ti eu
DEDICO

AGRADECIMENTOS

A dissertação de mestrado além de muita responsabilidade exige muita dedicação e persistência. A sua realização, só foi possível através da entrega e colaboração de diferentes pessoas e entidades, pelas quais dirijo o meu sincero reconhecimento:

A Deus, pelo seu amor infinito, por me amparar nos momentos difíceis, pelas oportunidades que colocou em minha vida, por ser o autor do meu destino, meu guia, protetor e por nunca ter me abandonado. Eu creio na sua existência. Além disso, agradeço a Deus por ter arquitetado de uma forma exuberante o meio ambiente, principalmente os vegetais pelas quais me apaixono todos os dias e que tem sido o foco da minha pesquisa. Pois “*Deus cria, e o homem classifica*”.

A toda minha família, principalmente ao casal Augusto Fançony e Mecilia Pinto, por serem meus maiores exemplos de superação. Apesar da distância continental que nos separa, sempre estiveram comigo, velando pela minha saúde psicológica. Agradeço por sempre me apoiarem nas minhas decisões e sonhos, e por estarem compartilhando dessa minha jornada, ajudando a me manter segura e firme sempre nos momentos que sempre precisei. Vocês são os maiores orgulhos da minha vida.

Aos meus orientadores, na pessoa do Prof. Dr. Vitor Fernandes Oliveira de Miranda e da Dra. Saura Rodrigues Silva, por me terem aceitado como orientando e me mostrarem o caminho para trilhar a vida acadêmica, científica e profissional. Os conhecimentos que recebi persistentemente sobre Biologia molecular, Bioinformática, Genômica e Filogenia, me fazem enxergar que já sou um novo diamante lapidado para o mercado. Grandes amigos que levarei para vida toda.

A toda equipe do Laboratório de Sistemática Vegetal (LSV) da Faculdade de Ciências Agrárias e Veterinária da UNESP-Jaboticabal, pelos momentos vividos quer seja científico-profissional, como os de níveis culturais e de lazer, pois cada um segundo a sua particularidade marcou o meu coração.

Aos responsáveis, pesquisadores e técnicos do Laboratório de Genética de Bactérias e Biotecnologia Aplicada, pela disponibilização da infraestrutura para a realização das análises moleculares (PCR e Eletroforese).

Aos professores Prof. Dr. Daniel Guariz Pinheiro e Prof. Dr. Alessandro Mello Varani pelas discussões e sugestões levantadas no Exame de Qualificação.

Ao Instituto Nacional de Bolsa de Estudo de Angola (INAGBE), pela concessão da bolsa de estudo e ao Instituto Superior Politécnico do Cuanza Sul (Angola), pela dispensa de formação.

A Faculdade de Ciências Agrárias e Veterinárias - (FCAV) da Universidade Estadual Paulista (UNESP), principalmente ao Programa de Pós-Graduação em Agronomia (Genética e Melhoramento de Plantas), pelas disciplinas cursadas e outras questões inerente a minha formação.

Aos meus amigos de Angola e aos africanos residentes em Jaboticabal, com os quais dividi segredos, momentos de descontração e boas gargalhadas, pois foram também foram essenciais para que aqui chegasse.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradeço fortemente a qualquer outra pessoa que tenha colaborado de forma direta ou indireta para o desenvolvimento deste trabalho.

O meu muito obrigado!

SUMÁRIO

Página

AVALIAÇÃO DE GENES NUCLEARES PARA FILOGENIA E PLASTOMA DE LENTIBULARIACEAE

RESUMO

As relações filogenéticas dos gêneros *Pinguicula* L., *Genlisea* A.St.-Hil. e *Utricularia* L. dentro da família Lentibulariaceae foram bem estudadas. Pesquisas indicam um grupo monofilético, no qual *Pinguicula* é irmão do clado *Genlisea-Utricularia*. Porém, algumas lacunas entre os gêneros ainda existem. Essas incluem as delimitações e relações entre seções ou subgêneros e circunscrição infra-genérica, principalmente as espécies descobertas recentes. Além disso, as espécies desta família, apresentam uma grande variedade de habitats, formas de vida, mecanismos e especializações de captura das presas e características fenotípicas, o que aumenta a dificuldade de identificação, sobretudo no uso exclusivo de características morfológicas. As abordagens filogenéticas embasadas em caracteres moleculares com os genes nucleares e plastoma podem contribuir nas discrepâncias taxonômicas da família. Sendo assim, o presente estudo teve como objetivos principais (i) testar os genes álcool desidrogenase e nitroreductase para filogenia de *Genlisea* e *Utricularia (*Lentibulariaceae) e (ii) montar e caracterizar o plastoma de *Pinguicula moranensis*. Tanto os genes nucleares como o genoma de cloroplasto de *Pinguicula moranensis* foram caracterizados nesse estudo. Nossos resultados abrem pistas que contribuem para discussões evolutivas dos genes álcool desidrogenase e nitroreductase nas espécies de Lentibulariaceae. Além disso, o nosso estudo apresenta pela primeira vez o plastoma montado de *Pinguicula moranensis* e fornece novos *insights* para a compreensão sobre a evolução, filogenia e genômica de Lentibulariaceae.

Palavras chaves: Angiospermas, evolução, filogenia molecular, genoma, ptDNA

EVALUATION OF NUCLEAR GENES FOR LENTIBULARIACEAE PHYLOGENY AND CHLOROPLAST GENOME

ABSTRACT

The phylogenetic relationships of the genera *Pinguicula* L., *Genlisea* A.St.-Hil., and *Utricularia* L. within the Lentibulariaceae family have been well studied. Research indicates a monophyletic group, in which *Pinguicula* is sister to the *Genlisea-Utricularia* clade. However, some gaps between the genera still exist. These include the delimitations and relationships between sections or subgenera and infra-generic circumscriptions, especially for recently discovered species. Additionally, species in this family exhibit a wide variety of habitats, life forms, prey capture mechanisms, and phenotypic characteristics, which increases the difficulty of identification, especially when relying solely on morphological characteristics. Phylogenetic approaches based on molecular characters with nuclear and plastome genes can contribute to taxonomic discrepancies within the family. Therefore, the present study aimed to (i) test the alcohol dehydrogenase and nitroreductase genes for the phylogeny of *Genlisea* and *Utricularia* (Lentibulariaceae) and (ii) assemble and characterize the plastome of *Pinguicula moranensis*. Both the nuclear genes and the chloroplast genome of *Pinguicula moranensis* were characterized in this study. Our results provide clues that contribute to evolutionary discussions of the alcohol dehydrogenase and nitroreductase genes in Lentibulariaceae species. Additionally, our study presents for the first time the assembled plastome of *Pinguicula moranensis* and provides new insights into understanding the evolution, phylogeny, and genomics of Lentibulariaceae.

Key-words: Angiosperms, Evolution, Molecular phylogeny, genome; ptDNA

CAPÍTULO 1 - Considerações gerais

1- Introdução geral

A monofilia da família Lentibulariaceae Rich (Lamiales) encontra-se totalmente resolvida quanto à posição filogenética dos seus três gêneros constituintes (*Pinguicula* L., *Genlisea* A.St.-Hil., and *Utricularia* L.). Pois, diversos estudos baseados em caracteres morfológicos, moleculares atestam e comprovam esta abordagem (Taylor, 1989; Müller et al., 2004 e 2006; Silva et al., 2018a; Fonseca, 2021). No entanto, do ponto de vista taxonômico, especialmente nos táxons pertencentes ao clado *Genlisea - Utricularia*, ainda existem algumas lacunas que necessitam ser resolvidas. Isso inclui a definição taxonômica de seções ou subgêneros, identificação imprecisa e as relações entre espécies em nível infragenérico (Jobson et al., 2003; Müller e Borsch, 2005; Bobrov et al., 2017; Westermeier et al., 2017; Silva et al., 2018a e 2023; Shimai et al., 2021). Além disso, a crescente descoberta de novas espécies pertencente à família Lentibulariaceae pode provocar alterações ao perfil filogenético ou atualizações na taxonomia, quer seja intra ou extra seções, visto que as espécies desta família são caracterizadas pela diversidade fenotípica, *habitats* e formas de vida.

Essas lacunas destacam a importância de realizar pesquisas taxonômicas mais abrangentes nos três gêneros que constituem a família Lentibulariaceae. Entretanto, com base nas iniciativas tecnológicas, em particular no sequenciamento genômico, diversas técnicas têm sido empregadas para promover a identificação de táxons mais confiáveis e a reconstrução de hipóteses filogenéticas sólidas, contribuindo assim para uma taxonomia natural e mais defensável (Silva et al., 2018a). Os caracteres moleculares, combinados com ferramentas atuais baseadas em sistemática filogenética, podem ser muito úteis para sugestões nas delimitações taxonômicas, evolução, circunscrição dos gêneros de Lentibulariaceae e outras famílias de angiospermas.

Nos últimos anos, espécies de Lentibulariaceae foram sequenciadas (DNA cloroplastidial, nuclear e mitocondrial), sendo que algumas dessas possuem seus genomas montados e caraterizados (Wicke et al., 2013; Ibarra-Laclette et al., 2013; Silva et al., 2016, 2018a, 2018 e 2019; Li e Liu, 2022; Matos et al., 2022). O que tem possibilitado a identificação e seleção dos genes úteis para estudos filogenéticos e evolutivos das espécies da família. Além disso, as funções e as estruturas desses genes têm ajudado a esclarecer os padrões taxonômicos e evolutivos, bem como as características específicas, como características fenotípicas e adaptação em ambientes adversos.

Os representantes da família Lentibulariaceae, principalmente as espécies do clado *Genlisea-Utricularia*, possuem os menores genomas nucleares dentre todas as angiospermas sequenciadas até o presente momento, *Genlisea tuberosa* (61 Mpb, *Genlisea aurea* (63,6 Mb) e *Utricularia gibba* (100 Mb) respetivamente (Ibarra-Laclette et al., 2013; Lan et al., 2017), além de possuírem altas taxas de mutação nos seus compartimentos intracelulares (Greilhuber et al., 2006; Silva et al., 2023), tornando-as fortes candidatas para modelos em estudos sobre filogenia e evolução.

Em Spermatophyta, incluindo Lentibulariaceae, como em todas as angiospermas, os genes cloroplastidiais e nucleares têm sido os mais utilizados em estudos sobre filogenia de espécies e história evolutiva genética, em detrimento dos genes mitocondriais (Palmé et al., 2009; Cai e Ma et al., 2016; Silva et al., 2016, 2018a 2018 e 2019; Fonseca, 2021; Li e Liu, 2022).

Os genes nucleares apresentam vantagens para filogenia por possuírem herança biparental, possibilitando o resgate da ancestralidade completa das gerações (Small et al., 2004; Palmé et al., 2009; Cai e Ma et al., 2016; Huang et al., 2023; Zhang et al., 2024). Adicionalmente, alguns genes do compartimento nuclear demonstram taxas evolutivas consideravelmente elevadas, proporcionando uma maior precisão na distinção entre espécies intimamente relacionadas (Wolfe et al., 1987). É importante destacar que os genes nucleares de angiospermas sofrem poliploidia ou duplicações, o que permite o surgimento de genes homólogos quer seja artólogo ou parálogos, com dificuldade de identificação (Soltis et al., 2011; Guo et al., 2023). Portanto, torna-se importante avaliar que genes nucleares se tornam viáveis e congruentes para estudos filogenéticos nos organismos fotossintéticos (Cai e Ma et al., 2016).

Os plastomas são fundamentais para a filogenia e outros estudos genômicos, pois apresentam um elevado grau de polimorfismo e conservação genética, facilitando para o desenvolvimento de códigos de barras genéticos, estudos de genética populacional, reconstrução de árvores filogenéticas de alta resolução e conservação de espécies ameaçadas e resolução de discrepâncias taxonômicas (Downiwe e Palmer 1992; Daniell et al., 2016; Robbins e Kelly, 2023). Esses genomas têm sido bem explorados por diversos pesquisadores e estudos publicados recentemente revelam uma quantidade de >13.000 sequências depositadas e/ou publicadas no banco de dados público do NCBI (National Center for Biotechnology Information) (Wang et al., 2024). Porém para a família Lentibulariaceae há poucas sequências genômicas de plastoma, sendo apenas 7 para *Genlisea* e 7 para *Utricularia* (Silva et al., 2023). *Pinguicula* é o que menos registrou investigações de plastoma com quatro sequências publicadas, *Pinguicula vulgaris* L. (NC_084256.1), *Pinguicula jackii* Barnhart (NC_068629.1)*, Pinguicula ehlersiae* Speta & F.Fuchs (Wicke et al., 2013) e *Pinguicula alpina* L. (Li e Liu, 2022), sendo as duas últimas com seus genomas cloroplastidiais montados e caracterizados.

Nesse contexto, diante da relevância da filogenia na classificação de plantas, juntamente com a utilidade de Lentibulariaceae como modelo para estudos filogenômicos e evolutivos e de outra natureza, esta dissertação teve como objetivos, divididos em dois capítulos:

Capítulo 1: Testar os genes álcool desidrogenase e nitrorredutase para filogenia de *Genlisea* e *Utricularia (*Lentibulariaceae)

Capítulo 2: Montar e caracterizar o plastoma de *Pinguicula moranensis* e inferir sua posição filogenética dentro da família Lentibulariaceae;

2- Revisão da literatura

2.1- Considerações sobre família Lentibulariaceae Rich.

A família Lentibulariaceae Rich. é um grupo de plantas pertencente à ordem Lamiales das angiospermas (Bremer et al., 2002; APG IV; 2016; Fonseca, 2021). É composta por três gêneros monofiléticos (fig 1): *Pinguicula* L., *Genlisea* A.St.-Hil. e *Utricularia* L. (Taylor, 1989; Müller et al., 2004; Silva et al., 2018a).

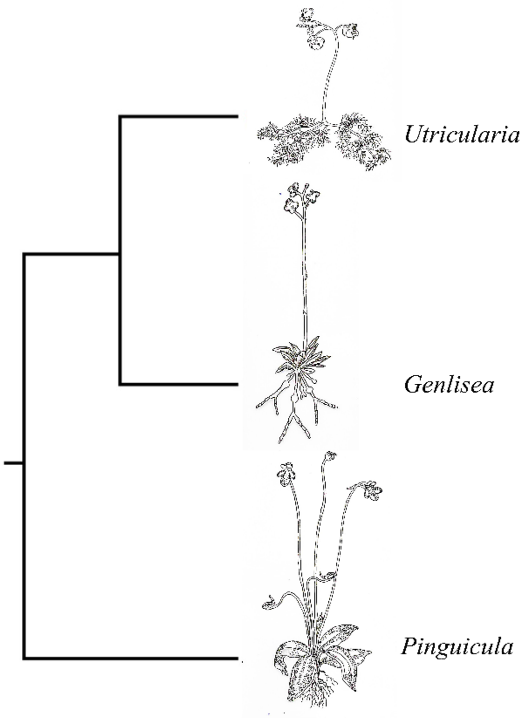


Fig. 1. Relações filogenéticas de Lentibulariaceae. Extraído e adaptado de Müller et al. (2006).

Das famílias botânicas com plantas carnívoras em angiospermas, a Lentibulariaceae é a mais diversificada quanto a quantidade de espécies, *habitats*, forma de vida e tipos de armadilhas (fig. 2). Os três gêneros abrigam mais de 350 espécies, distribuídos nos três gêneros em regiões tropicais, subtropicais e regiões áridas (Taylor, 1989; Casper, 1996; Baleeiro et al., 2017), em ambientes aquáticos e terrestres e as formas de vidas podem ser aquáticas livres, aquáticas flutuantes, litofítica, epífitas e reófitas (Taylor, 1989; Poppinga et al., 2016; Miranda et al., 2021). Cada gênero apresenta um padrão de armadilha (fig. 2), sendo que em *Pinguicula* as armadilhas, são as próprias folhas, que possuem uma mucilagem adesiva especializada em captura de presas, e *Genlisea* as armadilhas são folhas submersas metamorfoseadas em forma de "Y" invertido, finalmente o gênero *Utricularia* apresenta armadilhas em forma de vesículas, chamadas de utrículos (Taylor, 1989; Casper, 1996; Adamec, 2007; Albert et al., 2009; Fleischmann e Rocia 2018; Silva et al., 2018a; Reut et al., 2021). O Brasil consta dos maiores centros de diversidade de Lentibulariaceae, com os registros das espécies de *Genlisea* e *Utricularia* (Miranda et al., 2015; Poppinga et al., 2016; Guedes et al., 2020).

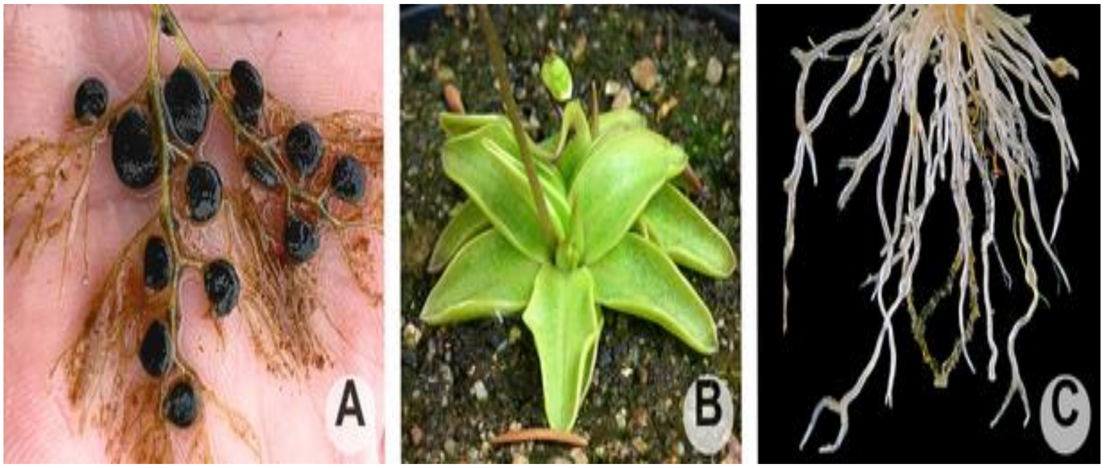


Fig 2: Diversidade de armadilhas em Lentibulariaceae: **A-** armadilhas em *Utricularia*; **B-** armadilhas em *Pinguicula*; **C-** armadilhas em *Genlisea*. Crédito-Unesp/LSV-Jaboticabal).

2. 1.1- Gênero *Pinguicula* L.

Originada do latim "pinguis" (gorduroso ou untuoso ao toque) (Heslop-Harrilson, 2004), o gênero Pinguicula (fig. 3) é composto por plantas herbáceas com distribuição cosmopolita, encontradas principalmente no Hemisfério Norte, África, América do Sul e nas regiões do Caribe e América Central. É nessas últimas que se verifica a maior diversidade específica, sendo consideradas o centro de diversidade do gênero (Casper, 1996; Legendre, 2000; Casper e Stimper, 2009; Domínguez et al., 2012; 2013; 2023). O gênero apresenta duas formas de crescimento distintas: uma temperada, caracterizada por um botão de inverno a partir do qual a nova planta pode brotar na primavera subsequente, e uma tropical, na qual a roseta fica mais estreita (Casper, 1996; Alcalá e

Domínguez, 2005; Grob et al., 2007). Em *Pinguicula* armadilha é constituída por folhas rosetadas e adesivas ou muciladas ocasionado pelas quantidades de tricomas glandulares, pedunculados ou sésseis com capacidade de fixar e digerir as presas com maior destaque os artrópodes (Heslop-Harrilson, 2004; Müller et al., 2004; Fleischmann e Rocia, 2018).

Fig. 4*:* Representante do gênero *Pinguicula L.:* **A-** *P. alpina* L.*;* **B**- *P. moranensis* Kunth*;* **C-** *P. vulgaris* L*.;* **D**- *P. jackii* Barnhart*;* **E-** roseta de *P. ehlersiae* Speta & F.Fuchs*;* ***F-*** *P.* flor de *P. ehlersiae* Speta & F.Fuchs*;* ***G***- *r*oseta e armadilhas de *P. vulgais* L. *Cr*éditos: https://cpphotofinder.com/. com exceção *P. vulgais* L. (Crédito-Unesp/LSV-Jaboticabal).

Em número de espécies, é o segundo maior gênero da família Lentibulariaceae, depois de *Utricularia,* albergando mais ou menos 100 espécies (Legendre, 2013; Rondondi et al., 2010; Fleischmann, 2021; Shimai et al., 2021; Li e Liu, 2022). Filogeneticamente *Pinguicula*, é grupo irmão de *Genlisea-Utricularia*, com características plesiomórficos para as características raízes e armadilhas, já que *Genlisea* e *Utricularia* revelam caráter apomórficos com ausência de raízes e com armadilhas complexas (Rutishauser e Isler 2001; Jobson *et al.*, 2003; Müller et al., 2004; Shimai et al., 2021). Com base no estudo realizado por Casper (1996), *Pinguicula* encontra-se dividida em três subgêneros: *Isoloba* Barnhart, *Pinguicula L.* e *Temnoceras* Barnhart.

2. 1.2- Gênero *Genlisea* A.St.-Hil*.*

Em termos de diversidade, *Genlisea* é o terceiro gênero dentro da família Lentibulariaceae, composto por 31 espécies, distribuídas desde o continente africano até a América do Sul e Central, com maior diversidade concentrada no sudeste do Brasil (Fleischmann et al., 2010; Silva et al., 2020; Adamec et al., 2021). No Brasil, foram registradas até o momento 19 espécies, das quais 12 são endêmicas, incluindo as do subgênero *Tayloria*, que é endêmico do leste do país (Guedes et al., 2024). No subgênero *Tayloria* destaca-se a espécie *Genlisea hawking*ii S.R.Silva, B.J.Plachno & V.Miranda, descoberta e descrita recentemente pelos pesquisadores da Faculdade de Ciências Agrárias e Veterinária afeto a Universidade Estadual Paulista - Brasil e seus colaboradores (Silva et al., 2020).

Os representantes deste gênero, são pequenas ervas (fig. 4) terrestres sem raízes, cujas folhas subterrâneas parecem um “Y” invertido, torcidas helicoidalmente especializadas em capturar, digerir e absorver presas, principalmente, bactérias, protozoários, algas e integrantes da microfauna, como nematóides, rotíferos, anelídeos e pequenos artrópodes (Płachno et al., 2007; Caravieri et al., 2014; Silva et al., 2020 e 2023). Taxonomicamente, este gênero agrupa dois subgêneros, *Tayloria* e *Genlisea,* diferenciados pelas características fenotípicas dos frutos, do pólen e das glândulas (Fromm-Trinta, 1977; 981; Płachno et al., 2007; Silva et al., 2023). Além disso, as espécies de *Genlisea* têm despertado interesse entre diversos pesquisadores por possuírem os menores tamanhos de genomas nucleares já sequenciados entre as plantas com flores, como é o caso da *Genlisea tuberosa* Rivadavia, Gonella & A.Fleischm (Ibarra-Laclette et al., 2013; Fleischmann et al., 2014; Lan et al., 2017).

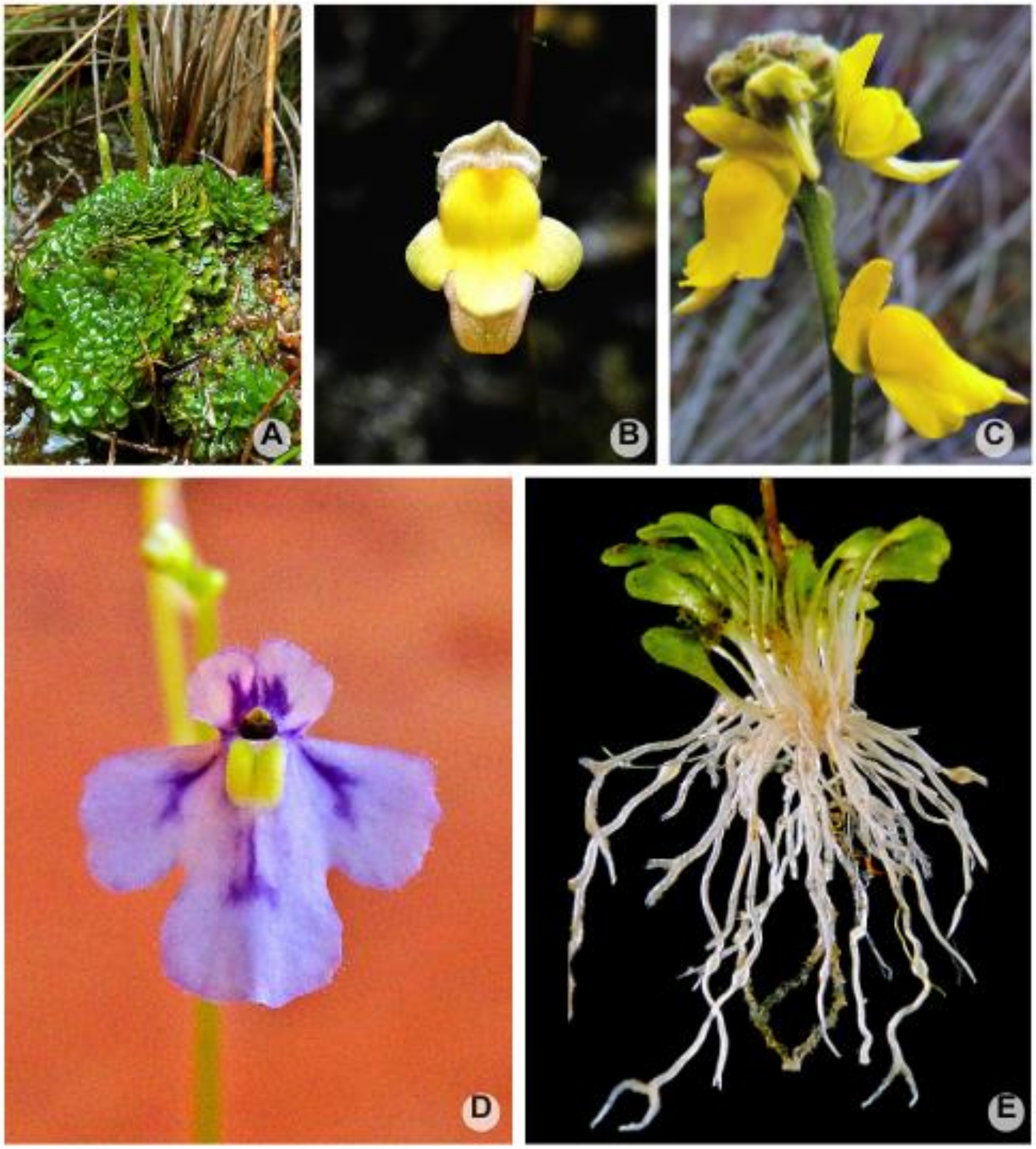


Fig. 4. Representantes do Gênero *Genlisea* A.St.-Hil. A- Hábito de *G. aurea* A.St.-Hil; B- flor de *G. filiformis* A.St.-Hil; C- flor de *G. aurea*. D- flor de *G. violacea* A.St.-Hil; E- roseta e armadilhas de *G. filiformis* (Crédito-LSV-Jaboticabal).

2. 1.3- Gênero *Utricularia* L.

Utricularia é o gênero que apresenta a maior diversidade de espécies na família Lentibulariaceae, além de abrigar a maior quantidade de plantas carnívoras dentro das angiospermas, totalizando mais de 250 espécies, o que equivale a cerca de 35% de todas as espécies de plantas carnívoras existentes (Taylor, 1989; Müller et al., 2004; Silva et al., 2018; Jobson et al., 2018; Miranda et al., 2021). Além disso, destaca-se pela significativa variação morfológica, tanto intra quanto interespecífica (fig. 5), uma característica que permite que suas espécies exploram uma ampla gama de habitats e formas de vida, desde ambientes aquáticos e terrestres até epífitas, crescendo sobre rochas e em outros ambientes com baixa disponibilidade de nutrientes (Taylor, 1989).

As armadilhas do gênero *Utricularia,* são pequenas vesículas denominadas utrículos, que apresentam estratégias de sugar a presa e digeri-las com ajuda de enzimas hidrolíticas, como proteases, fosfatases ácidas e esterases (Adamec et al., 2010; Rutishauser, 2016; Poppinga et al., 2016; Miranda et al., 2021). A sua distribuição é cosmopolita, ocorrendo em zonas tropicais e subtropicais, com a exceção de ilhas oceânicas e regiões áridas (Taylor, 1989). No Brasil, o gênero *Utricularia,* contempla 71 espécies (Guedes et al., 2024), deste número, 21 são endêmicas, presentes no Norte, Nordeste, Centro-Oeste, Sul e Sudeste (Taylor, 1989 e Guedes et al., 2024).

Trata-se de plantas com hábito herbáceo, com órgãos bem definidos, com variações florais, o que aumenta a dificuldade de identificação, até mesmo pelos especialistas (Rutishauser e Isler, 2001; Jobson et al., 2003). Suas estruturas florais, são muitas vezes confundidas com as de *Genlisea,* pois são quase parecidas, principalmente na coloração da corola, diferenciando-se em seus cálices, em que em *Genlisea* há cinco sépalas (pentalobado), enquanto *Utricularia* é bilobado, com duas ou quatro sépalas (Taylor, 1989).

Estudos, baseados em filogenética molecular, revelam o agrupamento do gênero *Utricularia* em três subgêneros, *Polypompholyx* (Lehm.) P.Taylor, *Bivalvaria* S.Kurz e *Utricularia,* sendo *Utricularia e Bivalvaria* compõe um grupo monofiléticos e

Polypompholyx é grupo irmãos do clado composto por *Utricularia e Bivalvaria* (Taylor, 1989; Müller et al., 2004; Müller e Borsch, 2005; Silva et al., 2018a; Silva et al., 2023), com 35 seções (Taylor, 1989). No entanto, ainda persistem várias discussões sobre esta classificação.

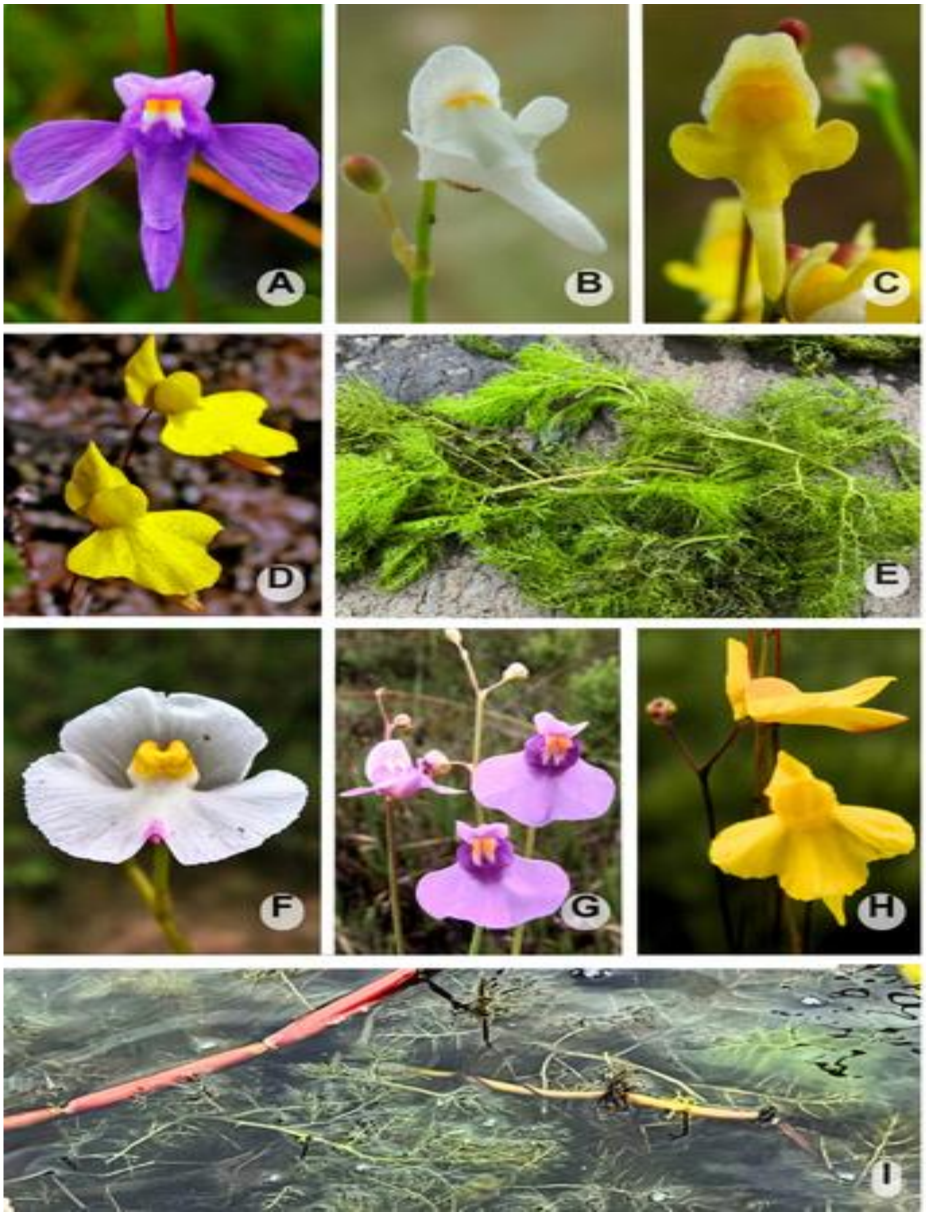


Fig. 5. Representante do gênero *Utricularia* L.: **A-C-** Diversidade floral de *Utricularia amethystina* Salzm. ex A.St.-Hil. & Girard *de* acordo com a cor da corola (A) *U. amethystina* roxa, (B) *U. amethystina branca e* (C) *U. amethystina* amarelo (Crédito: Silva, 2019); **D-** U. *subulata* L.; **E-** *U.foliosa* L.; **F-** U. *poconsensis* Fromm; **G-** U. *tricolor A.St.-Hil*; **H-** *U. triloba Benj;* **I-** *U. foliosa* L. no seu habitat natural (Crédito-Unesp/LSV-Jaboticabal).

2.2- Considerações sobre genoma na taxonomia de Lentibulariaceae

Tanto para as plantas quanto para outros organismos, o estudo do genoma é essencial, pois contém informações fundamentais para o controle, funcionamento, crescimento e reprodução dos organismos. Nele estão contidos todos os genes responsáveis pelas informações genéticas e proteicas. As plantas e outros seres eucarióticos são caracterizados pela presença de três compartimentos genômicos: nuclear, mitocondrial e cloroplastidial. *Arabidopsis thaliana* (L.) Heynh foi a primeira planta das angiospermas que recebeu o sequenciamento no ano de 2000 e até hoje é um modelo no estudo de genomas de plantas cultivadas e não cultivadas (The Arabidopsis Genome Initiative, 2000). Desde lá centenas de plantas têm sido sequenciadas, para aplicações em áreas como melhoramento genético, proteômica, medicina, taxonomia e outras.

A família Lentibulariaceae, apesar de possuir poucos estudos genômicos, registrou dados de sequenciamento dos genomas de algumas espécies, tanto de genoma nuclear, mitocondrial e cloroplastidial, sendo este último com maior registro (Silva et al., 2023). O fato de que alguns desta família apresentarem menores genomas nDNA como por exemplo *Genlisea tuberosa (*61 Mpb), menor que *Arabidopsis thaliana* com registro de 135 - 157 Mpb *(*The Arabidopsis Genome Initiative, 2000*;* Bennett & Leitch, 2005*),* além desses táxons apresentam altas taxas de mutação nos seus compartimentos intracelulares (Greilhuber et al., 2006; Ibarra-Laclette et al., 2013; Lan et al., 2017; Silva et al., 2023), este fato torna a família Lentibulariaceae um modelo para estudos taxonômicos para outras angiospermas. Nos últimos anos diversos estudos revelaram as resoluções de problemas taxonômicos dos três gêneros de Lentibulariaceae (Silva et al., 2018a, 2019; 2020; Shimai et al., 2021).

Considerando a filogenia, em Lentibulariaceae os genes dos genomas organelas têm sido mais utilizados em detrimento dos genes do compartimento nuclear. O genoma mitocondrial (mitogenoma), caracteriza-se por sua complexidade estrutural, possuindo um sistema genético semi-autônomo, com o tamanho variável, que hospedam genes que codificam proteínas responsáveis para o normal funcionamento da planta, relacionado como o metabolismo energético e homeostase celular (Gualberto et al., 2014). Estudos sobre mitogenomas realizados com espécies de Lentibulariaceae, revelam uma mobilidade genética entre os compartimentos e contração com perda (Silva et al., 2017; Matos et al., 2022). O genoma plastidial (plastoma) é uma estrutura quadripartida, composta duas regiões de cópias única, a longa (LSC) e pequena (SSC) e duas regiões de repetição invertida (IRa e IRb) (Downiwe & Palmer 1992; Daniell et al., 2016). Diversos estudos filogenéticos foram feitos na família Lentibulariaceae envolvendo os plastomas (e.g. Silva et al., Silva et al., 2018a, 2019; 2020; Shimai et al., 2021), revelando um alto grau de polimorfismo, identidade e similaridade entre sequências, que os tornam candidatos para estudos filogenéticos da família, servindo de modelos para outras angiospermas.

Referências bibliográficas

ADAMEC, L. Dark respiration of leaves and traps of terrestrial carnivorous plants: are there greater energetic costs in traps? **Open Life Sciences**, v. 5, n. 1, p. 121–124, 2010.

ADAMEC, L. Oxygen Concentrations Inside the Traps of the Carnivorous Plants *Utricularia* and *Genlisea* (Lentibulariaceae). **Annals of Botany**, v. 100, n. 4, p. 849–856, 2007.

ADAMEC, L.; MATUŠÍKOVÁ, I.; PAVLOVIČ, A. Recent ecophysiological, biochemical and evolutional insights into plant carnivory. **Annals of Botany**, v. 128, n. 3, p. 241–259, 2021.

ALBERT, V. A.; JOBSON, R. W.; MICHAEL, T. P.; *et al*. The carnivorous bladderwort (*Utricularia*, Lentibulariaceae): a system inflates. **Journal of Experimental Botany**, v. 61, n. 1, p. 5–9, 2009.

ALCALÁ, R. E., DOMINGUEZ, C. A. Diferential selection for carnivory traits along an environmental gradient *Pinguicula moranensis*. Ecology,86(10), p. 2652–2660, 2005.

ANGIOSPERM PHYLOGENY GROUP APG IV (APG IV). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. **Botanical Journal of the Linnean Society**, v. 181, n. 1, p. 1–20, 2016.

BALEEIRO, P. C; MOREIRA, A. D. R.; SILVA, N. G.; *et al*. Flora do Rio de Janeiro: Lentibulariaceae. **Rodriguésia**, v. 68, n. 1, p. 59–71, 2017.

BENNETT, M. D. Nuclear DNA Amounts in Angiosperms: Progress, Problems and Prospects. **Annals of Botany**, v. 95, n. 1, p. 45–90, 2005.

BOBROV, A.A.; VOLKOVA, P. A.; KOPYLOV-GUSKOV, Y. O.; *et al*. Unknown sides of *Utricularia* (Lentibulariaceae) diversity in East Europe and North Asia or how hybridization explained old taxonomical puzzles. **Perspectives in Plant Ecology, Evolution and Systematics**, v. 54, p. 125649, 2022.

BREMER, B.; B. E.R, Kåre; H., N.; *et al*. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. **Molecular Phylogenetics and Evolution**, v. 24, n. 2, p. 274–301, 2002.

CARAVIERI, F. A.; FERREIRA, A. J.; FERREIRA, A.; *et al*. Bacterial community associated with traps of the carnivorous plants *Utricularia hydrocarpa* and *Genlisea filiformis*. **Aquatic Botany**, v. 116, p. 8–12, 2014.

CASPER, S. J.; STIMPER, R. Chromosome numbers in *Pinguicula* (Lentibulariaceae): survey, atlas, and taxonomic conclusions. **Plant Systematics and Evolution**, v. 277, n. 1–2, p. 21–60, 2009.

CASPER, S. J. Monographie der Gattung *Pinguicula* L. Bibliotheca Botanica v. 127-128, n. 1, p. 123–145, 1966.

DANIELL, H.; LIN, C.; YU, M.; *et al*. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. **Genome Biology**, v. 17, n. 1, 2016.

DOMÍNGUEZ, Y.; DA SILVA, S. R.; PANFET, V. C. M.; *et al*. Inter- and intra-specific diversity of Cuban *Pinguicula* (Lentibulariaceae) based on morphometric analyses and its relation with geographical distribution. **Plant Ecology & Diversity**, v. 7, n. 4, p. 519–531, 2013.

DOMÍNGUEZ, Y.; PANFET, V. C.M.; MIRANDA, V. F. O. New features of Cuban endemic *Pinguicula filifolia* (Lentibulariaceae) and considerations on its habitat and ecology. **Flora - Morphology, Distribution, Functional Ecology of Plants**, v. 207, n. 11, p. 838–841, 2012.

DOMINGUEZ, Y.; TEMPLE, P.; PANČO, I.; *et al*. **Biogeographical Patterns of *Pinguicula* L. (Lentibulariaceae) in the Americas Revealed by Endemicity and Potential Distribution Analyses**. [s.l.]: Elsevier BV, 2023. Disponível em: <http://dx.doi.org/10.2139/ssrn.4500281>. Acesso em: 16 Feb. 2024.

DOWNIE, S. R.; PALMER, J. D. Use of Chloroplast DNA Rearrangements in Reconstructing Plant Phylogeny. *In*: **Molecular Systematics of Plants**. Boston, MA: Springer US, 1992, p. 14–35. Disponível em: <http://dx.doi.org/10.1007/978-1-4615-3276-7_2>. Acesso em: 16 Feb. 2024.

FLEISCHMANN, A. On the infrageneric classification of *Pinguicula*. **Carnivorous Plant Newsletter**, v. 50, n. 4, p. 174–188, 2021.

FLEISCHMANN, A.; MICHAEL, T. P.; RIVADAVIA, F.; *et al*. Evolution of genome size and chromosome number in the carnivorous plant genus *Genlisea* (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. **Annals of Botany**, v. 114, n. 8, p. 1651–1663, 2014.

FLEISCHMANN, A.; ROCCIA, A. **Systematics and evolution of Lentibulariaceae: I. Pinguicula**. [s.l.]: Oxford University Press, 2018. Disponível em: <http://dx.doi.org/10.1093/oso/9780198779841.003.0006>. Acesso em: 16 Feb. 2024.

FLEISCHMANN, A.; SCHÄFERHOFF, B.; HEUBL, G.; *et al*. Phylogenetics and character evolution in the carnivorous plant genus *Genlisea* A. St.-Hil. (Lentibulariaceae). **Molecular Phylogenetics and Evolution**, v. 56, n. 2, p. 768–783, 2010.

FONSECA, L. H. M. Combining molecular and geographical data to infer the phylogeny of Lamiales and its dispersal patterns in and out of the tropics. **Molecular Phylogenetics and Evolution**, v. 164, p. 107287, 2021.

GUEDES, F. B. M.; GONRLLA, P. M.; DOMINGUEZ, Y.; MOREIRA, A. D. R.; SILVA, S. R.; DÍAZ, Y. C. A.; FLEISCHMANN, A.; MENEZESs, C. G.; RIVADACIA, F.; MIRANDA, V. F. O. *Utricularia in* **Flora e Funga do Brasil.** Jardim Botânico do Rio de Janeiro.Disponível em: <https://floradobrasil.jbrj.gov.br/FB8570>. Acesso em: 16 Feb. 2024.

GREILHUBER, J.; BORSCH, T.; MÜLLER, K.; *et al*. Smallest Angiosperm Genomes Found in Lentibulariaceae, with Chromosomes of Bacterial Size. **Plant Biology**, v. 8, n. 6, p. 770–777, 2006.

GROB, V.; PFEIFER, E.; RUTISHAUSER, R. Sympodial Construction of Fibonacci-type Leaf Rosettes in *Pinguicula moranensis* (Lentibulariaceae). **Annals of Botany**, v. 100, n. 4, p. 857–863, 2007.

GUALBERTO, J. M.; MILESHINA, D.; WALLET, C.; *et al*. The plant mitochondrial genome: Dynamics and maintenance. **Biochimie**, v. 100, p. 107–120, 2014.

HESLOP-HARRISON, Y. *Pinguicula* L. **Journal of Ecology**, v. 92, n. 6, p. 1071–1118, 2004.

HUANG, J.; XU, W.; ZHAI, Junwen; *et al*. Nuclear phylogeny and insights into whole-genome duplications and reproductive development of Solanaceae plants. **Plant Communications**, v. 4, n. 4, p. 100595, 2023.

IBARRA-LACLETTE, E.; LYONS, E; HERNÁNDEZ-GUZMÁN, G.; *et al*. Architecture and evolution of a minute plant genome. **Nature**, v. 498, n. 7452, p. 94–98, 2013.

JOBSON, R. W; BALEEIRO, P. Molecular phylogenetic study of *Utricularia* section Oligocista in Australia and a new Cape York endemic species. **Telopea**, v. 26, 2023.

JOBSON, R. W.; BALEEIRO, P. C.; GUISANDE, C. **Systematics and evolution of**

Lentibulariaceae: III. *Utricularia*. [s.l.]: Oxford University Press, 2018. Disponível em: <http://dx.doi.org/10.1093/oso/9780198779841.003.0008>. Acesso em: 16 Feb. 2024.

LAN, T.; RENNER, T.; IBARRA-LACLETTE, E.; *et al*. Long-read sequencing uncovers the adaptive topography of a carnivorous plant genome. **Proceedings of the National Academy of Sciences**, v. 114, n. 22, 2017.

LEGENDRE, L. The genus *Pinguicula* L. (Lentibulariaceae): an overview. **Acta Botanica Gallica**, v. 147, n. 1, p. 77–95, 2000.

LI, J.; LIU, Z. L. Complete chloroplast genome sequence of the carnivorous herb *Pinguicula alpina* (Lentibulariaceae). **Mitochondrial DNA Part B**, v. 7, n. 6, p. 1061–1062, 2022.

MATOS, R. G.; SILVA, S. R.; PŁACHNO, B. J.; *et al*. The complete mitochondrial genome of carnivorous *Genlisea tuberosa* (Lentibulariaceae): Structure and evolutionary aspects. **Gene**, v. 824, p. 146391, 2022.

MIRANDA, V. F. O.; SILVA, S. R.; REUT, M. S.; *et al*. A Historical Perspective of Bladderworts (U*tricularia*): Traps, Carnivory and Body Architecture. **Plants**, v. 10, n. 12, p. 2656, 2021.

MIRANDA, V. F. O.; MENEZES, C. G.; SILVA, S. R.; DÍAZ, Y. C. A.; RIVADAVIA, F. Lentibulariaceae in Lista de Espécies da Flora do Brasil. Jard. Botânico do Rio Janeiro 12 Jan 2015 Available from: http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB146

MÜLLER, K.; BORSCH, T.; LEGENDRE, L.; *et al*. Evolution of Carnivory in Lentibulariaceae and the Lamiales. **Plant Biology**, v. 6, n. 4, p. 477–490, 2004.

MÜLLER, K. F.; BORSCH, T.; LEGENDRE, L.; *et al*. Recent Progress in Understanding the Evolution of Carnivorous Lentibulariaceae (Lamiales). **Plant Biology**, v. 8, n. 6, p. 748–757, 2006.

MÜLLER, K.; BORSCH, T. Phylogenetics of *Utricularia* (Lentibulariaceae) and molecular evolution of the trnK intron in a lineage with high substitutional rates. **Plant Systematics**

and Evolution, v. 250, n. 1–2, p. 39–67, 2004.

PALMÉ, A. E.; PYHÄJÄRVI, T.; WACHOWIAK, W.; *et al*. Selection on Nuclear Genes in a Pinus Phylogeny. **Molecular Biology and Evolution**, v. 26, n. 4, p. 893–905, 2009.

PŁACHNO, B. J.; KOZIERADZKA-KISZKURNO, M.; ŚWIĄTEK, P. Functional Utrastructure of *Genlisea* (Lentibulariaceae) Digestive Hairs. **Annals of Botany**, v. 100, n. 2, p. 195–203, 2007.

POPPINGA, S.; WEISSKOPF, C.; WESTERMEIER, A. S.; *et al*. Fastest predators in the plant kingdom: functional morphology and biomechanics of suction traps found in the largest genus of carnivorous plants. **AoB PLANTS**, v. 8, 2015.

REUT, M. S.; ŚWIĄTEK, P.; MIRANDA, V. F. O.; *et al*. Living between land and water – structural and functional adaptations in vegetative organs of bladderworts. **Plant and Soil**, v. 464, n. 1–2, p. 237–255, 2021.

ROBBINS, E. H. J; KELLY, S. The Evolutionary Constraints on Angiosperm Chloroplast Adaptation. **Genome Biology and Evolution**, v. 15, n. 6, 2023.

RODONDI, G.; BERETTA, M.; ANDREIS, C. Pollen morphology of alpine butterworts (*Pinguicula* L., Lentibulariaceae). **Review of Palaeobotany and Palynology**, v. 162, n. 1, p. 1–10, 2010.

RUTISHAUSER, R. Developmental Genetics and Morphological Evolution of Flowering Plants, Especially Bladderworts (*Utricularia*): Fuzzy Arberian Morphology Complements Classical Morphology. **Annals of Botany**, v. 88, n. 6, p. 1173–1202, 2001.

RUTISHAUSER, R. Evolution of unusual morphologies in Lentibulariaceae (bladderworts and allies) and Podostemaceae (river-weeds): a pictorial report at the interface of developmental biology and morphological diversification. **Annals of Botany**, v. 117, n. 5, p. 811–832, 2015.

SHIMAI, H.; SETOGUCHI, H.; ROBERTS, D. L.; *et al*. Biogeographical patterns and speciation of the genus *Pinguicula* (Lentibulariaceae) inferred by phylogenetic analyses.

PLOS ONE, v. 16, n. 6, p. 0252581, 2021.

SILVA, S. R.; DIAZ, Y. C. A.; PENHA, H. A.; *et al*. The Chloroplast Genome of *Utricularia reniformis* Sheds Light on the Evolution of the ndh Gene Complex of Terrestrial Carnivorous Plants from the Lentibulariaceae Family. **PLOS ONE**, v. 11, n. 10, p. e0165176, 2016.

SILVA, S. R.; ALVARENGA, D. O.; ARANGUREN, Y.; *et al*. The mitochondrial genome of the terrestrial carnivorous plant *Utricularia reniformis* (Lentibulariaceae): Structure, comparative analysis and evolutionary landmarks. **PLOS ONE**, v. 12, n. 7, p. e0180484, 2017.

SILVA, S. R.; GIBSON, R.; ADAMEC, L.; *et al*. Molecular phylogeny of bladderworts: A wide approach of *Utricularia* (Lentibulariaceae) species relationships based on six plastidial and nuclear DNA sequences. **Molecular Phylogenetics and Evolution**, v. 118, p. 244–264, 2018a.

SILVA, S. R.; MICHAEL, T. P.; MEER, E. J.; *et al*. Comparative genomic analysis of *Genlisea* (corkscrew plants-Lentibulariaceae) chloroplast genomes reveals an increasing loss of the ndh genes. **PLOS ONE**, v. 13, n. 1, p. e0190321, 2018.

SILVA, S. R.; PINHEIRO, D. G.; PENHA, H. A.; *et al*. Intraspecific Variation within the *Utricularia amethystina* Species Morphotypes Based on Chloroplast Genomes. **International Journal of Molecular Sciences**, v. 20, n. 24, p. 6130, 2019.

SILVA, S. R.; PŁACHNO, B. J.; CARVALHO, S. G. M.; *et al*. *Genlisea hawkingii* (Lentibulariaceae), a new species from Serra da Canastra, Minas Gerais, Brazil. **PLOS ONE**, v. 15, n. 1, p. e0226337, 2020.

SILVA, S. R.; MIRANDA, V. F. O.; MICHAEL, T. P.; *et al*. The phylogenomics and evolutionary dynamics of the organellar genomes in carnivorous *Utricularia* and *Genlisea* species (Lentibulariaceae). **Molecular Phylogenetics and Evolution**, v. 181, p. 107711, 2023.

SMALL, R. L.; CRONN, R. C.; WENDEL, J. F. Use of nuclear genes for phylogeny reconstruction in plants. **Australian Systematic Botany**, v. 17, n. 2, p. 145, 2004.

SOLTIS, D. E.; SMITH, S. A.; CELLINESE, N.; *et al*. Angiosperm phylogeny: 17 genes, 640 taxa. **American Journal of Botany**, v. 98, n. 4, p. 704–730, 2011.

TAYLOR, P. The Genus *Utricularia* – A Taxonomic Monograph. Kew Bulletin Additional Series XIV. Royal Botanic Gardens, Kew. London, 1989

WANG, J.; KAN, S.; LIAO, X.; *et al*. Plant organellar genomes: much done, much more to do. **Trends in Plant Science**, 2024.

WESTERMEIER, A. S; FLEISCHMANN, A.; MÜLLER, K.; *et al*. Trap diversity and character evolution in carnivorous bladderworts (*Utricularia*, Lentibulariaceae). **Scientific Reports**, v. 7, n. 1, 2017.

WICKE, S.; SCHÄFERHOFF, B..; DEPAMPHILIS, C. W.; *et al*. Disproportional Plastome-Wide Increase of Substitution Rates and Relaxed Purifying Selection in Genes of Carnivorous Lentibulariaceae. **Molecular Biology and Evolution**, v. 31, n. 3, p. 529–545, 2013.

WOLFE, K. H; LI, W. H; SHARP, P. M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. **Proceedings of the National Academy of Sciences**, v. 84, n. 24, p. 9054–9058, 1987.

WOLFE, K. H; LI, W. H; SHARP, P. M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. **Proceedings of the National Academy of Sciences**, v. 84, n. 24, p. 9054–9058, 1987.

ZHANG, G., & MA, H. Nuclear phylogenomics of angiosperms and insights heir relationships and evolution. **Journal of Integrative Plant Biology**. J. Integr. Planta Biol.00:1–33, 2024.

CAPÍTULO 2- Phylogeny of *Genlisea-Utricularia* (Lentibulariaceae) based on alcohol dehydrogenase and nitroreductase genes.

Abstract

The clade *Genlisea-Utricularia* (Lentibulariaceae) is the most diverse group of carnivorous angiosperm plants. They are characterized by a diversity of habitats (nutrient-poor soils of polar, temperate, tropical and subtropical regions, both in terrestrial and aquatic oligotrophic environments), life forms (lithophytic, epiphytic and rheophytic) and phenotypic characteristics, thus showing strategies for survival in unfavorable environments. On the other hand, alcohol dehydrogenase and nitroreductase genes participate in the regulation and reduction of substances that cause instability in plant growth and development. The main aim of this study is to test the alcohol dehydrogenase and nitroreductase genes for the phylogeny of *Genlisea* and *Utricularia* (Lentibulariaceae), using Maximum Likelihood and Bayesian phylogenetic analyses. In addition, we evaluated the protein domains, residual structure and motif composition of both genes. Our results revealed a phylogeny of genes reflecting the phylogeny of Lentibulariaceae, congruently and with monophyletic topology in both genes. The characteristics of the genes identified by bioinformatic techniques indicate a similarity and conservation of these nuclear genes, which influenced the phylogenetic resolution. This is the first study to evaluate the behavior of these nuclear orthologous genes in Lentibulariaceae, thus opening up avenues for discussions on perspectives for future studies involving more representation of this plant family as well as other angiosperms.

Keywords: Angiosperms, Gene conservation and evolution, Nuclear genes, Phylogenetic congruence.

1- Introduction

The Lentibulariaceae (Lamiales) family is the most diverse group of carnivorous plants, with a total of 350 species discovered and published, distributed in just three monophyletic genera: *Pinguicula* L., *Genlisea* A.St.-Hil. and *Utricularia* L., the latter two being considered as a sister group (Taylor, 1989; Müller et al., 2004; Fleischmann et al., 2010; Fleischmann and Rocia 2018; Silva et al., 2023), characterized by the diversity of traps (one type for each genus), specialized in capturing prey and converting them into nutrients necessary for their survival during the life cycle. Their species are cosmopolitan, and can be found in polar, temperate, tropical and subtropical regions inhabiting oligotrophic terrestrial and aquatic environments, with litholytic, epiphytic and rheophytic forms of life, thus reflecting a high degree of polymorphism (Taylor, 1989; Poppinga et al., 2016; Miranda et al., 2021). While species of *Pinguicula* have roots, those of the *Genlisea-Utricularia* clade do not, constituting a synapomorphy for the latter two genera. Phylogenetic studies point to *Pinguicula* as a sister group of the genera *Genlisea* and *Utricularia* (Müller et al., 2006; Fleischmann et al., 2010; Silva et al., 20118a e 2023), which means that in evolutionary history, *Genlisea-Utricularia* species lost roots over time. This synapomorphy means that species explore different habitats, growth habits and adaptation strategies, present greater species diversity as well as phenotypic diversity, serving as models for research on physiological, genetic, morphological aspects and evolutionary history in other species of angiosperms (Müller et al., 2004; Albert et al., 2009; Clivati et al., 2014; Silva et al., 2023).

In recent years, several genes from the mitochondrial, chloroplast and nuclear genome of eukaryotes have been used in order to phylogenetically describe the evolutionary history of taxa in angiosperms, through the characteristics of homologies that can be orthologous and paralogous. It is important to highlight that nuclear genome genes present advantages over chloroplast and mitochondria genes in evolutionary studies, as they present higher mutational rates, as well as because they present biparental inheritance, allowing the recovery of a complete ancestry of lineages (Wolfe et al., 1987; Small et al., 2004; Huang et al., 2023; Zhang et al., 2024). Nuclear genes with one or a few copies are of great importance for the phylogenetic escape of plants. The fact that these genes have biparental inheritance, unlike maternally inherited plastid and mitochondrial markers, allows us to glimpse hybridization events that may occur throughout plant evolution (Small et al., 2004; Linder and Rieseberg, 2004; Shen et al., 2021).

Among the various genes localized in the nucleus genome are the alcohol dehydrogenase and nitroreductase genes, which can be found in prokaryotes as well as in eukaryotes including angiosperms (Van et al., 1991; Speirs et al., 1998; Hannink et al., 2010; Dabravolski et al., 2020; Zeng et al., 2020; Shen et al., 2021).

These genes perform specialized functions that are essential for cellular maintenance and the adaptation of organisms in their characteristic environments or in adverse environments. However, studies centered on descriptions of the evolutionary history of alcohol dehydrogenase and nitroreductase genes in higher plants are rare (Sang et al., 1997; Charlesworth et al., 1998**;** Dabravolski et al., 2020; Li e al., 2022), and for the Lentibulariaceae family, no research has been developed or published involving the evolutionary history of these genes or the phylogeny of this gene class.

Alcohol dehydrogenases (ADHs) are a gene family belonging to the glycolytic pathway, which play a number of important roles in the various physiological processes of plants, enabling survival in anaerobic environments and functional maintenance of organism (Harberd, 1982; Thompson et al., 2007; Shen et al., 2021). With zinc-binding characteristics, alcohol dehydrogenase is a huge family, divided into three subfamilies: (i) short-chain dehydrogenase/reductase ((SDR)-ADH)), which contains about 250 amino acid residues, (ii) medium-chain dehydrogenase/reductase ((MDR)-ADH) with about 350 residues and (iii) long-chain dehydrogenase/reductase ((LDR)-AD))H which generally has 385~900 amino acid residues or about amino acid residues), the latter being absent in plant organisms (Speirs et al., 1998; Kim et al., 2009; Alka et al., 2013). Structurally, it is made up of two conserved domains: (i) the zinc-binding domain (ADH_zinc_N) with 206~340 amino acids and (ii) a GroES-like domain (ADH_N) with 35-164 amino acids which forms the enzyme catalytic domain (Murzin, 1996; Svensson et al., 1998; Zeng et al., 2020; Shen et al., 2021).

In plants, the activation of these domains not only allows plants to survive in anaerobic conditions, but also participates in other catabolic processes, catalysing reciprocal transformations between alcohol and aldehyde and flavour synthesis during fruit development, as well as the formation of alcohols during fruit ripening and resistance to aldehydes under hypoxic conditions or biotic stress (Chase, 1999; Dixon and Hewett, 2000; Zeng et al., 2020; Shen et al., 2021; Li e al., 2022).

However, research into the evolutionary studies of alcohol dehydrogenase has been carried out on a considerable number of plants, including angiosperms such as Phrymaceae, Poaceae, Brassicaceae, Fabaceae, Ranunculaceae, Rosaceae and others, reporting the occurrence of duplications and divergence (Kouse et al., 1995; Sang et al., 1997; Thompson et al., 2010; Zeng et al., 2020; Li e al., 2022). For example, in a study carried out by Zeng et al (2020), with genes from this family in Rosaceae species, the authors noticed frequent events of duplication of a single gene and expansion of other genes from this gene family. Therefore, this is a candidate gene family for in-depth studies, given the importance of duplications in plant evolutionary processes.

Nitroreductases (NRL) are a group of genes subdivided into two groups: (i) type I which catalyzes the reduction of organic nitro compounds; and (ii) type II which catalyzes a one-electron reduction of the nitro group to produce nitrate anion radicals which can react with oxygen to form superoxide and cause oxidative stress (Miškinien et al 1998; Akiva et al., 2017; Kim et al., 2021). Its structural domain is made up of flavin adenine mononucleotide (FMN), capable of metabolizing nitro compounds, using nicotinamide dinucleotide (NAD+) as an electron source and NAD (Nicotinamide adenine dinucleotide) and NADP+ (Nicotinamide adenine dinucleotide phosphate) acting as reducing agents in the genome (Miškinien et al 1998; Akiva et al., 2017; Dabravolski et al., 2020; Kim et al., 2021).

Interestingly, the nitroreductase gene, as well as being hydrophobic, forms homodimeric structures, which can be enzymes that are sensitive or insensitive to oxygen (Carvalho et al., 2020), fundamental aspects of plant survival strategies. There are reports on the evolution of nitroreductases, which occur divergently and radially from a minimal flamine-binding framework (Akiva et al., 2017). However, for plants, despite being scarce, evolutionary studies with this gene family in some Viridiplantae species (including angiosperms such as *Arabidopsis lyrata* L. and *Gossypium barbadense* L., suggest conservation of the functions of the coding genes and a definite high level of similarity in their structures (Dabrovolski et al., 2020), thus providing clues for more detailed research into the evolutionary history of genes in flowering plants.

Given their structural, functional, adaptive and also evolutionary importance, the genes encoding alcohol dehydrogenase and nitroreductase are potential candidates for the phylogeny of Lentibulariaceae species, and may provide insights for the taxonomy, biogeography and evolution not of only Lentibulariaceae but also for other angiosperms. Thus, our work aimed to test the alcohol dehydrogenase and nitroreductase genes for the phylogeny of *Genlisea* and *Utricularia* (Lentibulariaceae).

2- Materials and methods

2.1- Obtaining gene sequences, domains and residues

The sequences of *Genlisea* and *Utricularia* of the alcohol dehydrogenase and nitroreductase genes used in this work were extracted from previously assembled genomes (not yet published) belonging to the Laboratory of Plant Systematics (LSV) database. In order to investigate the domains of each gene, the sequences extracted were translated and aligned using the Mega-X tool (Kumar et al., 2018). These sequences were then submitted to the Interpro database (https://www.ebi.ac.uk/interpro/) to confirm the presence or absence of the characteristic domains of the alcohol dehydrogenase and nitroreductase genes. This programme also made it possible to identify the location of the protein domains of the respective genes.

2.2- Structural modeling and prediction of gene motifs

Structural modeling was carried out using the SWISS-MODEL web platform (Waterhouse et al., 2018), available at (https://swissmodel.expasy.org/interactive), a 3D structure visualization programme showing the angles, distance and location of hydrogen bonds. In this study, we modeled one species from each genus (*Genlisea nigrocaulis* and *Utricularia geminiloba*) and *Arabidopsis thaliana*, as it is a previously studied species and its structure is known. These two species were selected because of some differences in their sequences at the time of alignment. The main objective of this modeling was to observe the target sites of amino acid residues. In addition, this programme made it possible to observe the functional domains of each group of genes. Therefore, two structural models were built and compared, the monomer and the homotetramer model.

The conserved motifs were determined using the MEME Suite v.5.3.0 tool (Bailey et al., 2009), which can be accessed at https://meme-suite.org/meme/tools/meme. To this end, the protein sequences of the six Lentibulariaceae species used in the modeling analyses were accomplished and compared with *Arabidopsis thaliana*. The main objective of this analysis was to verify the recurrence of motifs, that is, short sequences of functional importance of the alcohol dehydrogenase and nitroreductase genes (D’haesleer, 2006; Boeva, 2016), in species of the *Utricularia-Genlisea* clade, and compare them with the motifs commonly found in *Arabidopsis thaliana*.The parameters used for this search were those proposed by the platform, changing only the maximum number of motifs to 15.

2.3- Phylogenetic analyses of the alcohol dehydrogenase and nitroreductase genes

All the sequences extracted and the presence of the domains that characterize the alcohol dehydrogenase and nitroreductase genes of the *Utricularia* and *Genlisea* species were aligned using Muscle Codon (Edgar, 2004) within the Mega-X software (Kumar et al., 2018). Phylogenetic and evolutionary inferences for the two genes were made using Maximum Likelihood (ML) and Bayesian (BI) methods.

The construction of the Maximum Likelihood tree was calculated using the online software IQ-Tree 1.6.12 (Nguyen et al., 2014) with the best-of-fit model TIM2+F+I+G4, according to the BIC criteria (Akaike, 1974) calculated with the program ModelFinder v.2 (Kalyaanamoorthy et al., 2017). Clade support was estimated using the ultrafast bootstrap (UFBoot) and SH-aLRT algorithms (Hoang et al., 2018) with 1,000 replicates.

Bayesian analyses were carried out using the MrBayes v. 3.2.2 program (Ronquist et al., 2012) on the CIPRES platform. Here, 10^{10} generations were produced and the posterior probabilities (clade support) calculated using the MCMC (Markov Chain Monte Carlo) algorithm. Trees were sampled every 100 generations, discarding 25% of the initial trees (burnin). The evolutionary model used was the same as that used for the ML tree.

Both Maximum Likelihood (ML) and Bayesian (BI) trees were rooted with species *Arabidopsis thaliana L.* (NM_180926.3 for alcohol dehydrogenase; for NM_100082.4 nitroreductase), Vitis vinifera L. (XM_034835507.1; CP146242.1), Coffea canephora Pierre ex A.Froehner (GW397363.1; EU164537.1)*, Mimulus guttatus DC (XM_013002161.1; CV522154.1)* and *Solanum lycopersicum* L. *(XM_004239582.4*; XM_004229754.4) as outgroup. The phylogenetic trees were visualized and drawn using iTOL - Interactive Tree of Life v5 (Letunic e Bork, 2021) and edited using the Inkscape v1.3.2 application (https://inkscape.org). All sequences of the alcohol dehydrogenase and nitroreductase genes used as external groups were obtained from previously assembled genomes and deposited in the GenBank database incorporated into NCBI (National Center for Biotechnology Information) (<http://www.ncbi. nlm.nih.gov>), through keyword searches and BLAST (Basic Local Alignment Search Tool) searches.

3- RESULTS

3.1- Gene identification and characterisation

3.1.1- Alcohol dehydrogenase

The searches carried out by the Interpro (https://www.ebi.ac.uk/interpro/) identified the presence of the alcohol dehydrogenase gene domains in all the sequences of

Utricularia and *Genlisea* species, comparing them with those of *Arabidopsis thaliana*, which served as a model. These domains fulfill biological and structural functions, containing the residues and motifs of alcohol dehydrogenase. The two main domains that characterize these genes are zinc-binding (ADH_zinc_N) and GroES-like (ADH_N) (Murzin, 1996; Svensson et al., 1998; Zeng et al., 2020; Shen et al., 2021). In addition, other orthologous domains belonging to alcohol dehydrogenase were found, namely: NAD (P)-biding domain superfamily (NAD(P)-bd_dom_sf), Alcohol dehydrogenase-like, N-terminal (ADH-like _N); GroES-like superfamily (GroES-like_sf); Alcohol dehydrogenase, zinc-type (ADH_Zn_CS), Alcohol dehydrogenase-like, C-terminal (ADH-like_C), Polyketide synthase, enoylreductase domain (PKS_ER). They recorded variations in the locations of occurrence in almost all the species, showing a slight similarity between species of the same genus, although many of these are overlapping. The main domains were found in regions between 66 and 487, and in some species they overlapped.

The structural diversity of alcohol dehydrogenase genes constructed using SWISS-MODEL (Waterhouse et al., 2018). The model was constructed by homology of the modular monomer and homotetramer structures in *Genlisea nigrocaulis* and *Utricularia geminiloba*, using *Arabidopsis thaliana* (AF-Q93ZM6-F1) as template. A comparative analysis of the modular structures of the *Genlisea-Utricularia* species used in these studies revealed a structural identity. Notably, in Lentibulariaceae the structural variations occur mainly in the top regions and in the zinc-binding sites which preferentially contain the ALA, VAL, HIS and SER residues. In addition, variations in the position of amino acid residues were also identified. Curiously, the Lentibulariaceae species showed the same stability performance, mainly in the modular monomer, while the homotetramer model revealed few differences, thus revealing a possible protein conservation of the gene. Overall, the results show an identity of 60 to 90% of the protein domains between the sequences of the species, with only a slight divergence between the genera. The main variations were found in the zinc-binding regions, mainly in monomers, corroborating previous studies (Thompson et al., 2007; Zeng et al., 2020).

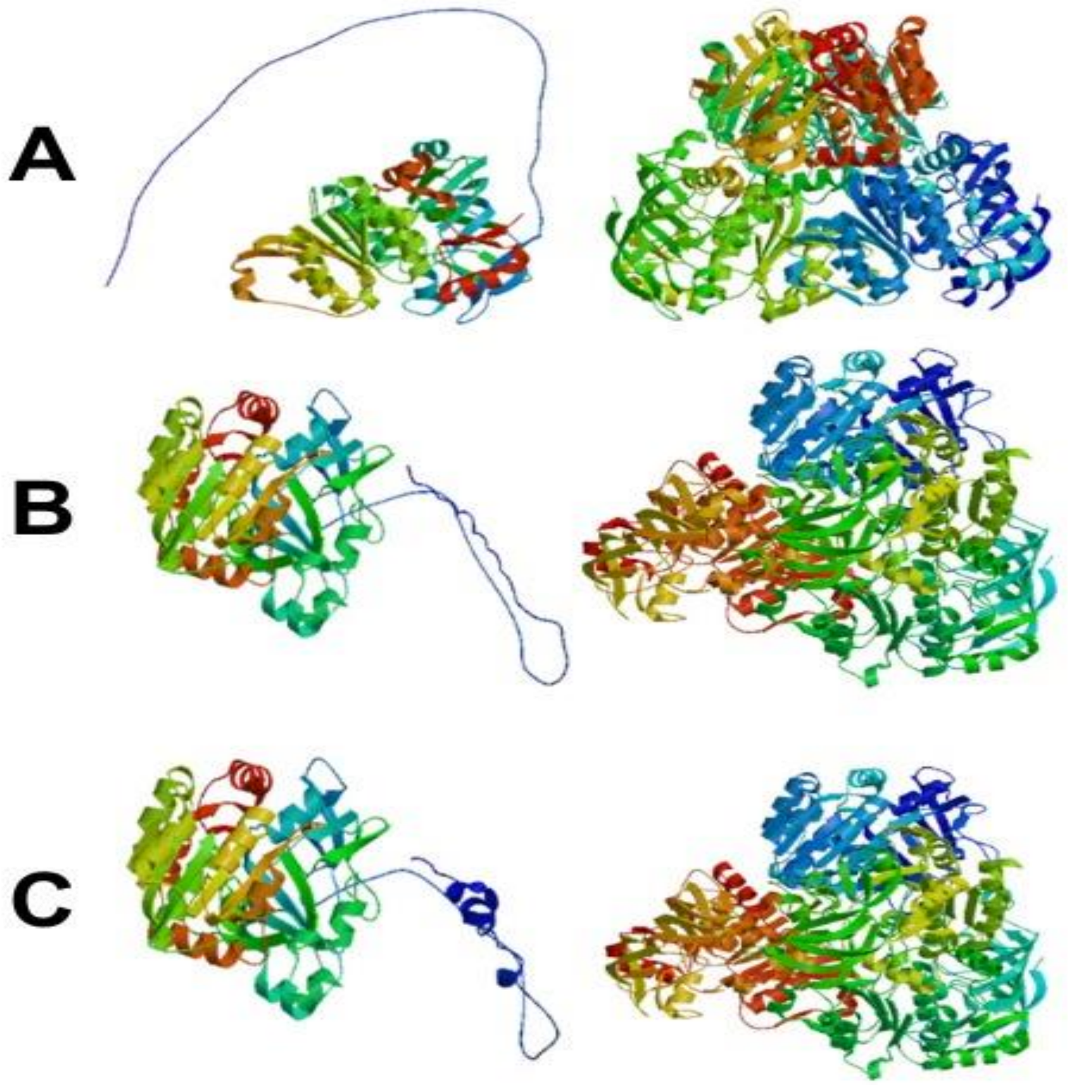


Figure 1: Modeled structures of alcohol dehydrogenase from the species: *Arabidopsis thaliana* (A), *Genlisea nigrocaulis* (B) and *Utricularia geminiloba* (C). The Monomer model is on the left and the homotetramer model is on the right. The colors indicate the position of the amino acid residues of the alcohol dehydrogenase representative domains.

When comparing the motifs of the sequences using the MEME Suite v.5.3.0 software (Bailey et al., 2009), certain divergences in the numbers of motifs were noted between the genera and within the genera (fig. 2).

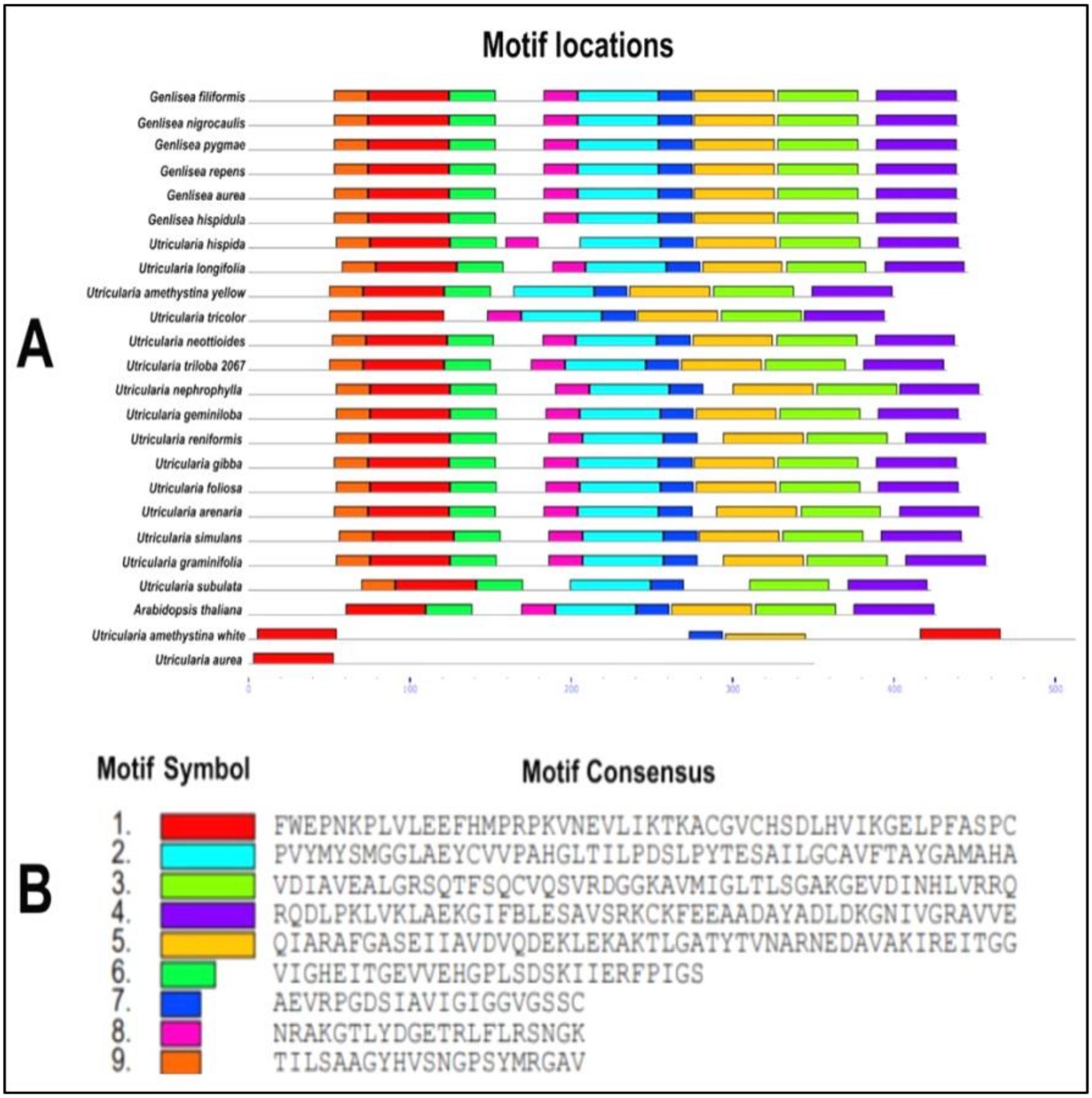


Fig. 2: Representation of the alcohol dehydrogenase motifs: (A) Locations of the motifs of the 23 sequences from Lentibulariaceae and Arabidopsis thaliana (B) consensus of the motifs of the species represented.

The genus *Genlisea* seems to have the highest similarity of motifs both in number and position, and of the nine motifs recovered, and without positional variation on the sequences of the genus. The opposite was true for the genus *Utricularia*. Although certain species had nine motifs, some varied even between phylogenetically close species, such as *Utricularia amethystina* “yellow”, which had 8 motifs, and *Utricularia amethystina* “white”, which only had motif one (1). Surprisingly, it was found that the species with the highest and lowest number of amino acids of all the sequences used in these studies (*Utricularia amethystina* with 690 aa and *Utricularia aurea* with 350 aa), were the ones with the lowest number of motifs. In terms of localisation, the motifs that varied the most were motifs one and four, while the most stable were motifs two and three. The widths of the motifs vary from a maximum of 50 to a minimum of 21. Notably all the motifs are located in the preserved regions of the domains.

3.1.2- Nitroreductase

The Interpro (https://www.ebi.ac.uk/interpro/) identified in all the sequences the presence of the nitroreductase domain with FMN (Flavin mononucleotide), FAD (Flavin adenine dinucleotide), NADH (Nicotinamide adenine dinucleotide) and NADPH (Nicotinamide adenine dinucleotide phosphate) molecules corroborating other authors (Miškinien et al 1998; Akiva et al., 2017; Dabravolski et al., 2020; Kim et al., 2021). The domain is found in regions 95 to 620. An orthologue of nitroreductase (nitroreductase-like) was also identified in all sequences, and in some species (e.g. *Utricularia geminiloba* and *Utricularia arenaria*) it overlaps with the main domain. Curiously, some species showed double nitroreductase, positioned in the first and last regions.

The model was constructed by homology of the modular monomer and homotetramer structures in *Genlisea nigrocaulis* and *Utricularia geminiloba*, using *Arabidopsis thaliana* (A0A5D1ZY20) as template. As with the alcohol dehydrogenase gene in this study, the results of the modelling (fig 3) of the nitroreductase gene referring to the monomer and homotetramer models of the species *Arabidopsis thaliana*, *Genlisea nigrocaulis* and *Utricularia geminiloba*, revealed structural identity in the order of 46 to 87%. The main differences were observed in *Genlisea nigrocaulis*, mainly in the structural monomer. The locations of the amino acid residues also varied in all the species, but in Lentibulariaceae they tended to be closer.

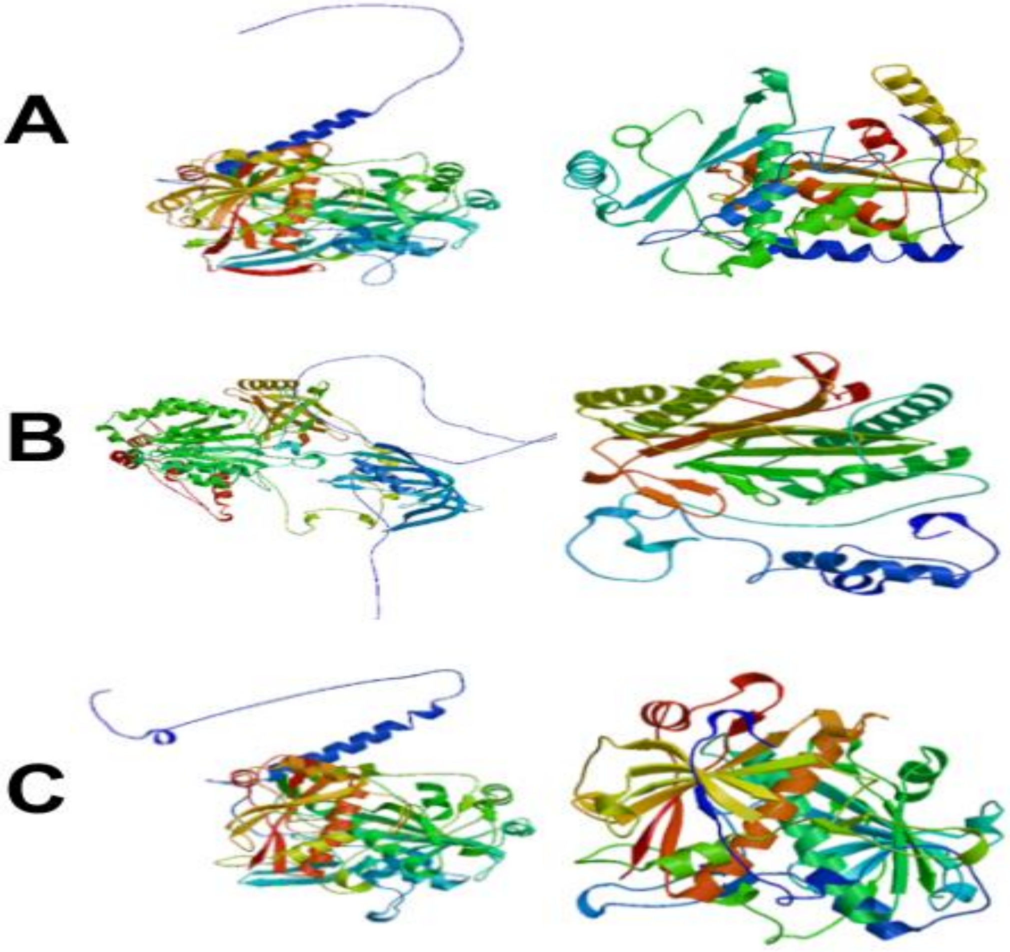
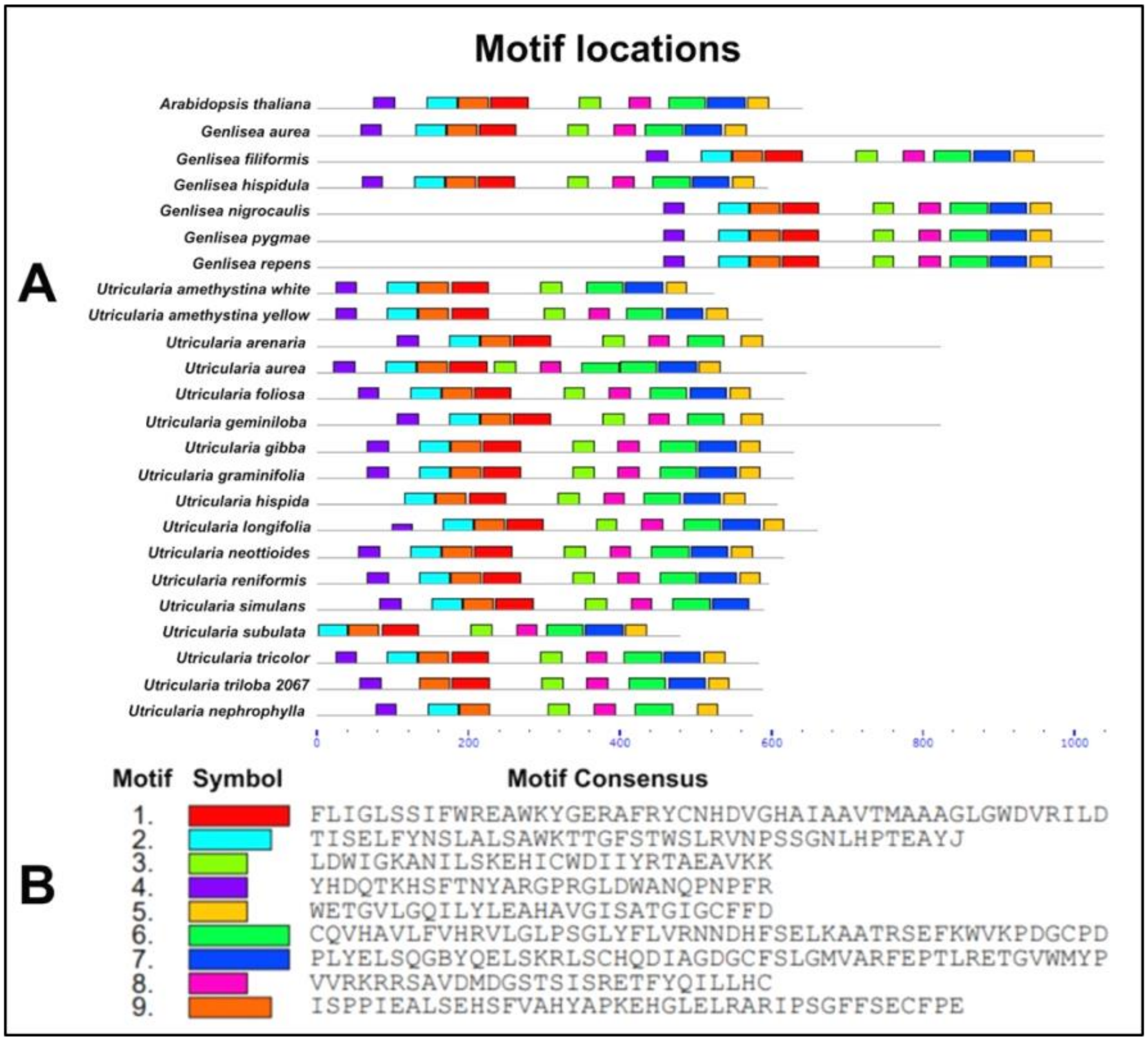


Figure 3: Modeled structures of the nitroreductase gene from the species: *Arabidopsis thaliana* **(A)**, *Genlisea nigrocaulis* **(B**) and *Utricularia geminiloba* **(C)**. The Monomer model is on the left and the homotetramer model is on the right. The colors indicate the position of the amino acid residues of the nitroreductase representative domains.

For the nitroreductase gene, the motifs identified by the MEME tool (https://meme-suite.org/meme/tools/meme) are all located in the conserved regions of the domains. The maximum width was 50 aa for motifs one, six and seven, while the minimum width was 29 aa for motifs three, four, five and eight. Motifs eight and two had a width of 41 aa. These results reveal significant conservation between the sequences, as well as their similarity. Interestingly, the regions of occurrence of the motifs are grouped together, except for *Genlisea nigrocaulis*, *Genlisea pygmaea* and *Genlisea repen*s, which are slightly further apart, with the nine motifs being positioned in the regions of 450 to 950 apparently. These residue changes contribute to amino acid rearrangements (Thompson et al., 2007).



Motif	Symbol	Motif Consensus
1.	🟥	FLIGLSSIFWREAWKYGERAFRYCNHDVGHAIAAVTMAAAGLGWDVRILD
2.	🟦	TISELFYNSLALSAWKTTGFSTWSLRVNPSSGNLHPTEAYJ
3.	🟩	LDWIGKANILSKEHICWDIIYRTAEAVKK
4.	🟪	YHDQTKHSFTNYARGPRGLDWANQPNPFR
5.	🟨	WETGVLGQILYLEAHAVGISATGIGCFFD
6.	🟩	CQVHAVLFVHRVLGLPSGLYFLVRNNDHFSELKAATRSEFKWVKPDGCPD
7.	🟦	PLYELSQGBYQELSKRLSCHQDIAGDGCFSLGMVARFEPTLRETGVWMYP
8.	🟪	VVRKRRSAVDMDGSTSISRETFYQILLHC
9.	🟧	ISPPIEALSEHSFVAHYAPKEHGLELRARIPSGFFSECFPE

Fig 4: Representation of the motifs of the nitroreductase gene: (A) Locations of the motifs of the 23 sequences from Lentibulariaceae and Arabidopsis thaliana (B) consensus of the motifs of the species represented.

3.2- Phylogenetic analyses of alcohol dehydrogenase and nitroreductase genes in Lentibulariaceae

Phylogenetic analysis based on 23 sequences of alcohol dehydrogenase and nitroreductase genes from Lentibulariaceae species were shown in Figures 5 and 6, respectively. The phylogenetic trees performed by maximum likelihood and Bayesian methods for both genes showed very similar and completely congruent trees. The only differences were only detected in the support values in both Bayesian (posterior probabilities) and Maximum likelihood (bootstrap and SH-aLRT). Therefore, in this study, we adopted a more conservative approach to propose evolutionary hypotheses, selecting trees derived from the Bayesian method.

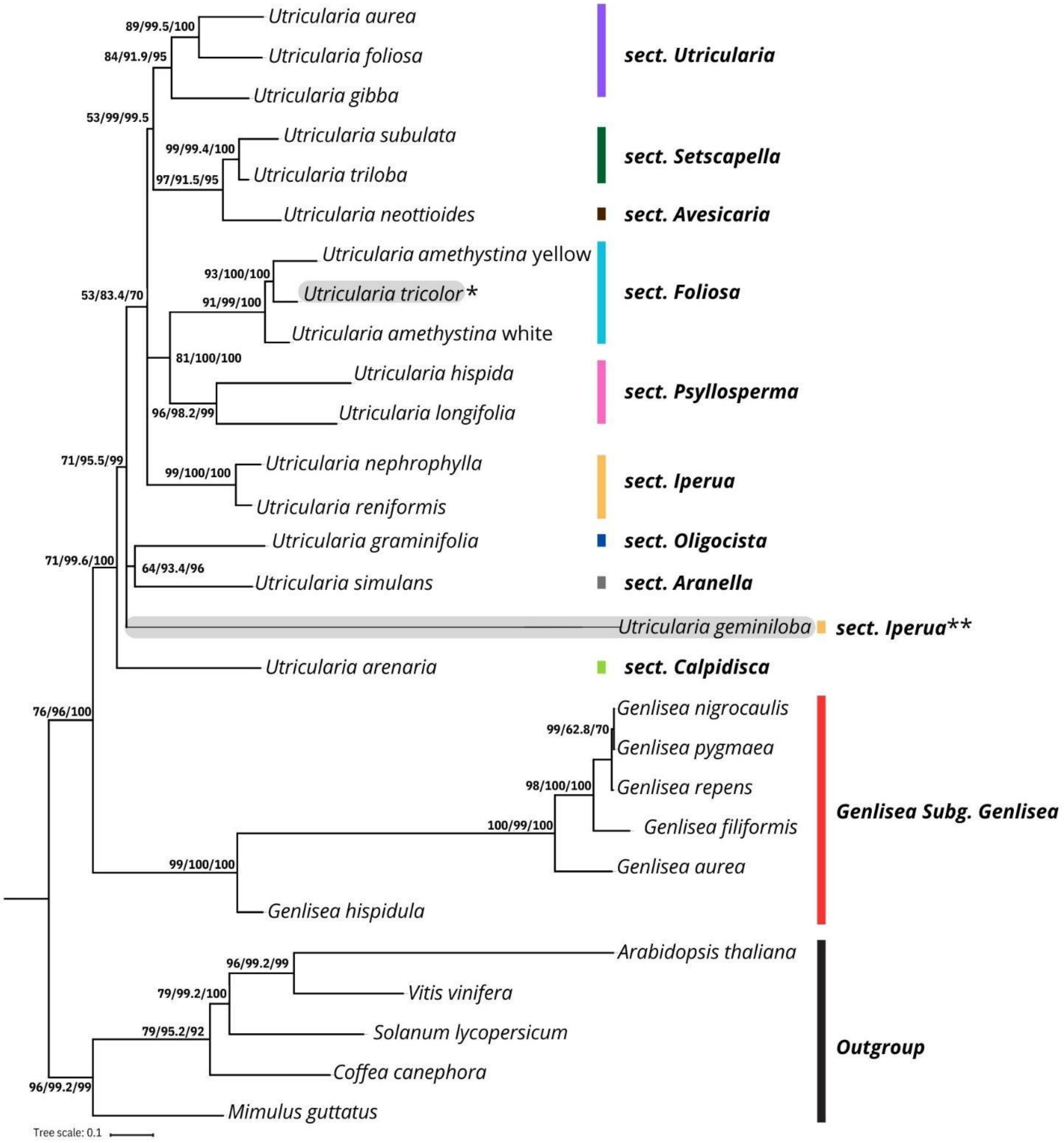


Fig. 5: Cladogram of Bayesian analysis based on the alcohol dehydrogenase gene. The support values (posterior probabilities, bootstrap of the ML analysis and SH-aLRT) of each clade are indicated next to the

respective nodes. The infrageneric classification (Taylor, 1989; Fleischmann, 2012) is shown on the right. The gray background color in the terminal branch denotes incongruence of the taxa in the clades in which they are positioned.

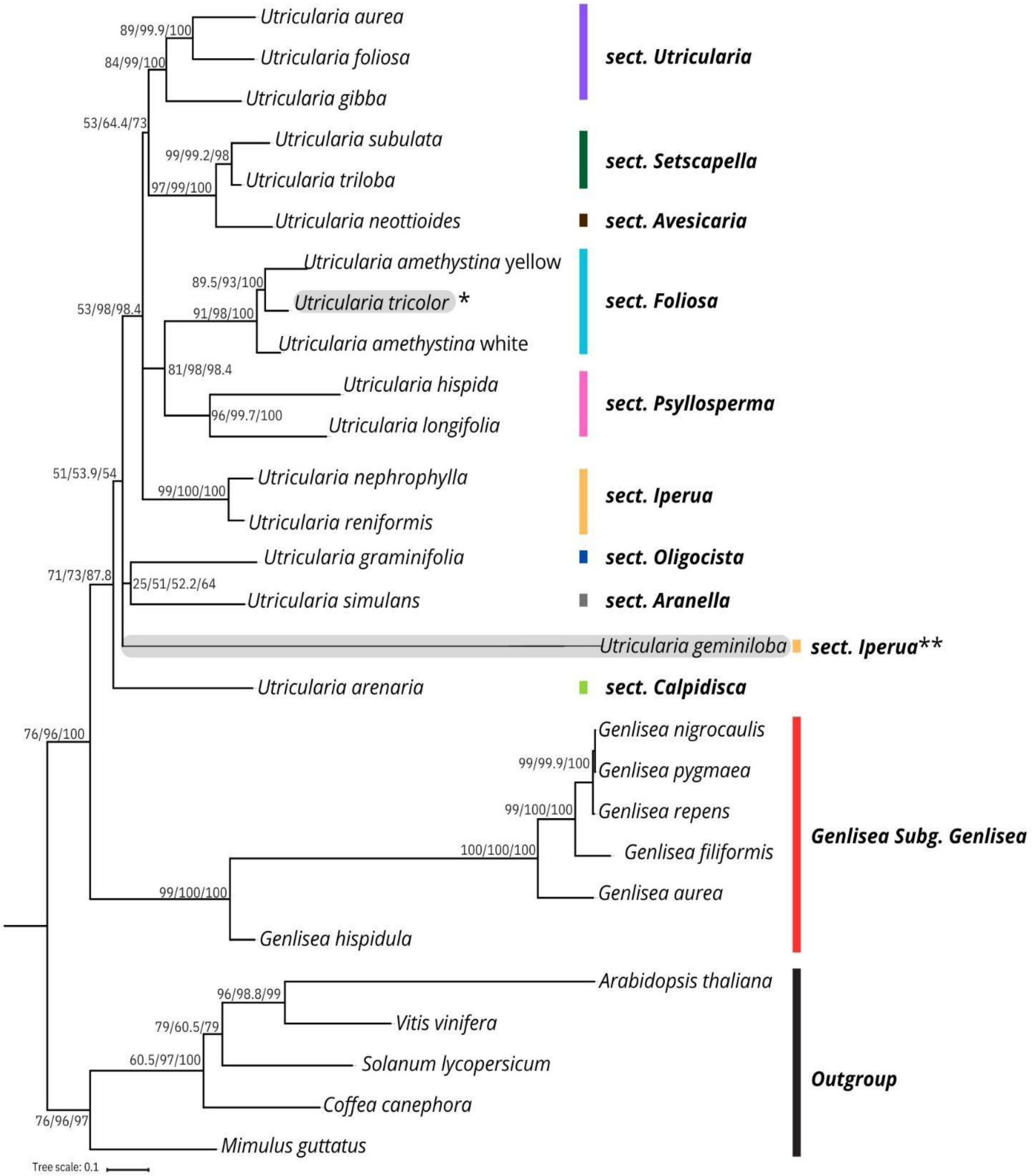


Fig. 6: Cladogram of Bayesian analysis based on the nitroreductase gene. The support values (posterior probabilities, bootstrap of the ML analysis and SH-aLRT) of each clade are indicated next to the respective nodes. The infrageneric classification (Taylor, 1989; Fleischmann, 2012) is shown on the right. The gray background color in the terminal branch denotes incongruence of the taxa in the clades in which they are positioned.

4- Discussion

The motivation behind this study was to test the feasibility of reconstructing the phylogenetic trees using alcohol dehydrogenase and nitroreductase genes located in the nuclear genome using Lentibulariaceae as a model. Therefore, this is the first study involving these genes in this family. Before the phylogenetic analyses, the genes were first structurally characterized in terms of their domains, model structure and motif composition.

Protein motifs are short and conserved sequences, that is, recurring patterns that can be identified in several different proteins, but with some type of relationship and are often associated with a specific structure or function, with the purpose of interacting with target genes of protein factors. transcription, recognition and molecular resistance, catalytic site, cellular targeting, interactions of protein pathways and with other proteins essential for the proper functioning of organisms, while the domains are broader and may even encompass one or more motifs that determine the specificity in the formation of protein dimers and transcriptional activation, the modular structure allows knowing the variations or not of the amino acid residues responsible for the formation of proteins or a probable gene or rearrangement, being able to form new homologous genes (Egea-Cortines et al., 1999; Pařenicová et al., 2003; D'haesleer, 2006; Thompson et al., 2010; Boeva, 2016). Knowledge of these and other aspects is essential in phylogenetic studies, especially when there is a lack of information about the respective genes of the organisms under study, as it allows evaluating the quality of the sequences used for phylogeny.

The alcohol dehydrogenase gene exhibits various functional domains, including zinc-binding domains (ADH_zinc_N) and GroES-like domains (ADH_N), along with orthologues within its gene family. Zinc-binding domains, which are typically dimeric or tetrameric proteins, facilitate the connection of two zinc atoms by coordinating cysteine and histidine residues, crucial for catalyzing the reversible oxidation of alcohols into acetaldehyde or ketone, while reducing NAD. The stability of these proteins is maintained by additional residues, contributing to molecular stability (Thorn et al., 1995; Pauly et al., 2003; Klein et al., 2023).

On the other hand, GroES-like domains, characterized by their irregular β-chaperonin-10-containing protein structure, function as quinone oxidoreductases and glucose dehydrogenases. These domains are prone to Ile to Val mutations and are hydrophobic, aiding in the conservation of residues necessary for maintaining the three-dimensional structure, regulating gene transcription, and ensuring overall integrity (Murzin, 1996; Taneja and Mande, 1999; Rutten et al., 2009).

Given these biological functions, the alcohol dehydrogenase gene holds significant importance for plant survival in anaerobic conditions and other adverse habitats, including those encountered by Lentibulariaceae species. Lentibulariaceae species, particularly within the *Genlisea-Utricularia* clade, exhibit diverse habitats, such as poor and swampy soils, polar, temperate, tropical and subtropical regions, and oligotrophic terrestrial and aquatic environments. These environments are often associated with water and present various stresses (Taylor, 1989; Poppinga et al., 2016; Miranda et al., 2021). Therefore, it is suggested that further studies on alcohol dehydrogenase genes be conducted with a broad representation of Lentibulariaceae species, including *Pinguicula*, to better understand their role in adapting to and surviving in such challenging environments.

In this study, through structural modeling and the search for motifs, we were able to understand the diversity of residues and motif structure. Despite a slight difference, our results show mostly a high identity, similarity and genetic conservation along the main regions in almost all the sequences of the Lentibulariaceae species and the model species *Arabidopsis thaliana*, as has been seen in other studies with alcohol dehydrogenase genes involving other eudicotyledons (Kouse et al., 1995; Sang et al., 1997; Svensson et al., 1998; Thompson et al., 2007 and 2010; Zeng et al., 2020; Shen et al., 2021; Li et al., 2022). These aspects are fundamental for solving phylogenetic problems. Therefore, these results provide clues for considering the alcohol dehydrogenase gene as a candidate for phylogenetic studies of Lentibulariaceae and other angiosperms.

This study also characterized the nitroreductase gene, with a view to phylogenetic reconstruction of some species belonging to the *Genlisea-Utriculari*a clade (Lentibulariaceae). The conserved domain and an orthologue of (nitroreductase-like) were identified in all species, although some showed a double domain, especially *Utricularia geminiloba* and *Utricularia arenaria*, which probably influenced the topological position of the clade (see chapter 3.2). It is important to note that nitroreductase genes can contain partial copies of both domains, which in many cases cannot be identified (Oliveira et al., 2007; Lee et al., 2008), which probably and can be present in other sequences in this study.

Therefore, in all the sequences the molecules FAD (Flavin adenine dinucleotide), NADH (Nicotinamide adenine dinucleotide) and NADPH (Nicotinamide adenine dinucleotide phosphate). The presence of these molecules in nitroreductases promote the nitro-substituted compounds metabolization (Miškinien et al 1998; Akiva et al., 2017; Dabravolski et al., 2020; Kim et al., 2021), thus enabling the survival and stability of organisms in unfavorable environments. In addition, nitroreductase is responsible for the conversion of nitro compounds into amines, which may be important for the detoxification of xenobiotics and the production of secondary metabolites, participating in the response of plants to abiotic and biotic stress, as well as in the regulation of plant growth and development (Zhu et al.,2018; Dabravolski et al., 2020; Carvalho et al., 2020).

Furthermore, the structural identity of *Genlisea* and *Utricularia*, as well as the similarity and conservation of nitroreductase motifs were revealed in this study. Given the habitats and the wide lifestyles of the *Genlisea* and *Utricularia* species, as well as the specialization of their traps in prey capture and nutrient digestion, evolutionary studies with nitroreductase genes become challenging. *Genlisea* traps are underground, forming forks with slits where soil microorganisms are inserted and end up being digested, while in *Utricularia*, its traps are small vesicles commonly called utricles, with the capacity to suck in prey (Barthlott et al., 1998; Albert et al., 2009; Fleischmann et al., 2014; Poppinga et al., 2016; Silva et al., 2023) and, therefore, the traps may contribute to the probable presence of nitro compounds as several subproducts as the result of the prey digestion in

Lentibulariaceae species. Therefore, the nitroreductase gene can be also a strong candidate for evolutionary studies in Lentibulariaceae and maybe to other angiosperms.

Notably, phylogenetic analysis of alcohol dehydrogenase and nitroreductase genes based on *Genlisea* and *Utricularia* (Lentibulariaceae) species, performed using Maximum Likelihood and Bayesian inference, showed the same topology and the same evolutionary history of the genes. Our results (Figs. 5 and 6) reveal a phylogeny of genes that reflects the phylogeny of Lentibulariaceae, monophyletic and congruent within the family, as reported by other authors (Müller et al., 2006; Silva et al., 2018a; Silva et al., 2023). Despite the slight congruences when these trees are compared previous phylogenies of this genus (Müller et al., 2006; Silva et al., 2018a), both genes have potential for recovering the phylogeny of the species. The incongruences found are possibly due these genes do not totally match the evolutionary history of the species.

Although in some clades the support values (posterior probabilities, bootstrap and SH-aLRT) were slightly low, as well as slight inconsistencies within *Utricularia*, the phylogenetic reconstructions based on these two genes have phylogenetic signal and in the majority the phylogenetic history of Lentibulariaceae. In the genus *Genlisea*, there was 100% congruence regarding the positioning of the species in the respective clades. The main incongruences found include the positioning of *Utricularia geminiloba*, which was separated from other species, *Utricularia nephrophylla* and *Utricularia reniformis*, all belonging to sect. *Iperua*, as well as the proximity of *Utricularia amethystin*a yellow and *Utricularia tricolor* as opposed to *Utricularia amethystina* white, both belonging to sect. *Foliosa* (Taylor, 1989; Silva et al., 2018a).

These incongruences, especially in *Utricularia geminiloba*, are probably attributed to the different phylogenetic history of the genes. These can be caused by different factors, as gene duplication, incomplete lineage sorting and the coalescence processes, which can be critical when nuclear genes are used in phylogeny of species (Yijie et al., 2019; Zeng et al., 2020). As the case of gene duplications, changes and rearrangements of amino acid residues lead to the emergence of new functions and characteristics, while gene conservation is important in ensuring the maintenance of essential characteristics for the survival of species (Dabravolski et al., 2020; Zeng et al., 2020; Shen et al., 2021). However, the phylogenetic relationship of the species of the sect. *Foliosa* represented in this study (*Utricularia amethystina* yellow, *Utricularia tricolor* and *Utricularia amethystina* white) merits an in-depth evaluation with further studies of the genus, with a wider range of phylogenetically related species. In addition, the phylogeny of genes does not always represent the real phylogeny of species, as plant evolution sometimes involves several reticulation events (Linder and Rieseberg, 2004).

Therefore, based on the results obtained in this study, it is possible to state that phylogenetic studies using the nuclear, alcohol dehydrogenase and nitroreductase genes in Lentibulariaceae, especially *Utricularia* and *Genlisea* species, are feasible and mostly defensible and consistent.

5- Conclusion

For the first time, the phylogeny of the alcohol dehydrogenase and nitroreductase genes has been used in the phylogeny of Lentibulariaceae, specifically in species of the genera *Genlisea* and *Utricularia*. Our results revealed a phylogenetic evolution of the genes, reflecting the actual phylogeny of the Lentibulariaceae species, with the trees of both genes presenting a congruent and monophyletic topology for both genes, thus attesting to the congruence and consistency of these genes for phylogenetic studies. In addition, we identified and characterized the genes in terms of their protein domains, residual structure and motif composition. The results indicate similarity, identity and genetic conservation. This study has provided important information for comparative and functional analyses of these two genes for these plant groups as well as in all angiosperms. Therefore, with this study we conclude that the alcohol dehydrogenase and nitroreductase genes are potential candidates for the phylogeny of Lentibulariaceae and other plant groups.

References

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N. & Csaki, F. (Eds). **Second International Symposium on Information Theory**. Budapest: Akademiai Kiado. pp. 267–281, 1974.

AKIVA, El; COPP, Janine N.; TOKURIKI, Nobuhiko; *et al*. Evolutionary and molecular foundations of multiple contemporary functions of the nitroreductase superfamily. **Proceedings of the National Academy of Sciences**, v. 114, n. 45, 2017.

ALBERT, V. A.; JOBSON, R. W.; MICHAEL, T. P.; *et al*. The carnivorous bladderwort (*Utricularia*, Lentibulariaceae): a system inflates. **Journal of Experimental Botany**, v. 61, n. 1, p. 5–9, 2009.

ALKA, K.; WINDLE, H. J.; CORNALLY, D.; *et al*. A short chain NAD(H)-dependent alcohol dehydrogenase (HpSCADH) from Helicobacter pylori: A role in growth under neutral and acidic conditions. **The International Journal of Biochemistry & Cell Biology**, v. 45, n. 7, p. 1347–1355, 2013.

BARTHLOTT, W.; POREMBSKI, S.; FISCHER, E.; *et al*. First protozoa-trapping plant found. **Nature**, v. 392, n. 6675, p. 447–447, 1998.

BOEVA, V. Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells. **Frontiers in Genetics**, v. 7, 2016.

CARVALHO, S.M.; BEAS, J. Z.; VIDEIRA, M. A. M.; *et al*. Defenses of multidrug resistant pathogens against reactive nitrogen species produced in infected hosts. *In*: **Advances in Microbial Physiology**. [s.l.]: Elsevier, 2022, p. 85–155.

CASPER, S. J. Monographie der Gattung *Pinguicula* L. Bibliotheca Botanica v. 127-128, n. 1, p. 123–145, 1966.

CHARLESWORTH, D.; LIU, F. L.; ZHANG, L. The evolution of the alcohol dehydrogenase gene family by loss of introns in plants of the genus Leavenworthia (Brassicaceae).

Molecular Biology and Evolution, v. 15, n. 5, p. 552–559, 1998.

CHASE, T. Alcohol Dehydrogenases: Identification and Names for Gene Families. **Plant Molecular Biology Reporter**, v. 17, n. 4, p. 333–350, 1999.

CLIVATI, D.; CORDEIRO, G. D.; PŁACHNO, B. J.; *et al*. Reproductive biology and pollination of *Utricularia reniformis* A.St.-Hil. (Lentibulariaceae). **Plant Biology**, v. 16, n. 3, p. 677–682, 2013.

DABRAVOLSKI, S. A. Evolutionary aspects of the Viridiplantae nitroreductases. **Journal of Genetic Engineering and Biotechnology**, v. 18, n. 1, p. 60, 2020.

OLIVEIRA, I. M.; HENRIQUES, J. A. P.; BONATTO, D. In silico identification of a new group of specific bacterial and fungal nitroreductases-like proteins. **Biochemical and Biophysical Research Communications**, v. 355, n. 4, p. 919–925, 2007.

DEANE, C. M.; BLUNDELL, T. L. Protein comparative modelling and drug discovergy. *In*: **The Practice of Medicinal Chemistry**. [s.l.]: Elsevier, 2003, p. 445–458.

D'HAESELEER, P. What are DNA sequence motifs? **Nature Biotechnology**, v. 24, n. 4, p. 423–425, 2006.

DIXON, J.; HEWETT, E. W. Factors affecting apple aroma/flavour volatile concentration: A Review. **New Zealand Journal of Crop and Horticultural Science**, v. 28, n. 3, p. 155–173, 2000.

EDGAR, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**, v. 5, n. 1, 2004.

EGEA-CORTINES, M. Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. **The EMBO Journal**, v. 18, n. 19, p. 5370–5379, 1999.

FLEISCHMANN, A.; MICHAEL, T. P.; RIVADAVIA, F.; *et al*. Evolution of genome size and chromosome number in the carnivorous plant genus *Genlisea* (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. **Annals of Botany**, v. 114, n. 8, p. 1651–1663, 2014.

Fleischmann, A. & Roccia, A. 2018. Systematics and evolution of Lentibulariaceae: I. *Pinguicula*. In: Ellison AM, Adamec L (eds) Carnivorous plants: physiology, ecology and evolution. Oxford University Press, 2018.

FLEISCHMANN, A.; SCHÄFERHOFF, B.; HEUBL, G.; *et al*. Phylogenetics and character evolution in the carnivorous plant genus *Genlisea* A. St.-Hil. (Lentibulariaceae). **Molecular Phylogenetics and Evolution**, v. 56, n. 2, p. 768–783, 2010.

FLEISCHMANN, A. Monograph of the genus *Genlisea*. Redfern Natural History Productions, 2012

HANNINK, N.; ROSSER, S. J.; FRENCH, C. E.; *et al*. Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. **Nature Biotechnology**, v. 19, n. 12, p. 1168–1172, 2001.

HARBERD, N. P.; EDWARDS, K. J. R. The efect of a mutation causing alcohol dehydrogenase deficiency on flooding tolerance in barley. **New Phytologist**, v. 90, n. 4, p. 631–644, 1982.

HOANG, D. T.; CHERNOMOR, O.; VON HAESELER, A. t; *et al*. UFBoot2: Improving the Ultrafast Bootstrap Approximation. **Molecular Biology and Evolution**, v. 35, n. 2, p. 518–522, 2017.

HUANG, J.; XU, W.; ZHAI, J.; *et al*. Nuclear phylogeny and insights into whole-genome duplications and reproductive development of Solanaceae plants. **Plant Communications**, v. 4, n. 4, p. 100595, 2023.

Bailey, T. L.; Boden, M.; Buske, F. A.; Frith, M.; Grant, C. E.; Clementi. L; Ren, J.;, Li, W. W.; Noble, W. S. MEME Suite: Tools for motif discovery and searching. **Nucleic Acids Research** 37: W202–W208, 2009.

KALYAANAMOORTHY, S.; MINH, B. Q..; WONG, T. K. F; *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. **Nature Methods**, v. 14, n. 6, p. 587–589, 2017.

KIM, S. H.; PARK, S.; PARK, E.; *et al*. Structure and substrate specificity determinants of NfnB, a dinitroaniline herbicide–catabolizing nitroreductase from Sphingopyxis sp. strain HMH. **Journal of Biological Chemistry**, v. 297, n. 4, p. 101143, 2021.

KIM, Y. Jin; SHIM, J. S.; LEE, J. H.; *et al*. Isolation and characterization of a novel short-chain alcohol dehydrogenase gene from Panax ginseng. **BMB Reports**, v. 42, n. 10, p. 673–678, 2009.

KLEIN, T.; FUNKE, F.; ROSSBACH, O.; *et al*. Investigating the Prevalence of RNA-Binding Metabolic Enzymes in E. coli. **International Journal of Molecular Sciences**, v. 24, n. 14, p. 11536, 2023.

Kosuge, K.; SAWADA, T. D.; ADACHI, J.; WATANABE, K. Phylogenetic relationships of some genera in the Ranunculaceae based on alcohol dehydrogenase genes. **Retrieved from Springer Vienna**, 1995.

KUMAR, S.; STECHER, G.; LI, M.; *et al*. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547–1549, 2018.

LEE, S. W.; MITCHELL, D. A.; MARKLEY, A. L.; *et al*. Discovery of a widely distributed toxin biosynthetic gene cluster. **Proceedings of the National Academy of Sciences**, v. 105, n. 15, p. 5879–5884, 2008.

LETUNIC, I.; BORK, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. **Nucleic Acids Research**, v. 49, n. W1, p. W293–W296, 2021.

LI, Y.; WANG, R.; PEI, Y.; *et al*. Phylogeny and functional characterization of the cinnamyl alcohol dehydrogenase gene family in *Phryma leptostachya*. **International Journal of Biological Macromolecules**, v. 217, p. 407–416, 2022.

LINDER, C. R.; RIESEBERG, L. H. Reconstructing patterns of reticulate evolution in plants. **American Journal of Botany**, v. 91, n. 10, p. 1700–1708, 2004.

MIRANDA, V. F. O.; SILVA, S. R.; REUT, M. S.; *et al*. A Historical Perspective of Bladderworts (*Utricularia*): Traps, Carnivory and Body Architecture. **Plants**, v. 10, n. 12, p. 2656, 2021.

MIŠKINIENĖ, V.; ŠARLAUSKAS, J.; JACQUOT, J. P.; *et al*. Nitroreductase reactions of *Arabidopsis thaliana* thioredoxin reductase. **Biochimica et Biophysica Acta (BBA) - Bioenergetics**, v. 1366, n. 3, p. 275–283, 1998.

MÜLLER, K.; BORSCH, T.; LEGENDRE, L.; *et al*. Evolution of Carnivory in Lentibulariaceae and the Lamiales. **Plant Biology**, v. 6, n. 4, p. 477–490, 2004.

MÜLLER, K. F.; BORSCH, T.; LEGENDRE, L.; *et al*. Recent Progress in Understanding the Evolution of Carnivorous Lentibulariaceae (Lamiales). **Plant Biology**, v. 8, n. 6, p. 748–757, 2006.

MURZIN, A. G. Structural classification of proteins: new superfamilies. **Current Opinion in Structural Biology**, v. 6, n. 3, p. 386–394, 1996.

NGUYEN, L. T.; SCHMIDT, H. A.; VON, H. A.; *et al*. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. **Molecular Biology and Evolution**, v. 32, n. 1, p. 268–274, 2014.

PAŘENICOVÁ, L.; DE FOLTER, S.; KIEFFER, M.; *et al*. Molecular and Phylogenetic Analyses of the Complete MADS-Box Transcription Factor Family in *Arabidopsis*. **The Plant Cell**, v. 15, n. 7, p. 1538–1551, 2003.

PAULY, T. A.; EKSTROM, J. L.; BEEBE, D. A.; *et al*. X-Ray Crystallographic and Kinetic Studies of Human Sorbitol Dehydrogenase. **Structure**, v. 11, n. 9, p. 1071–1085, 2003.

POPPINGA, S.; WEISSKOPF, C.; WESTERMEIER, A. S.; *et al*. Fastest predators in the plant kingdom: functional morphology and biomechanics of suction traps found in the largest genus of carnivorous plants. **AoB PLANTS**, v. 8, 2015.

RONQUIST, F.; TESLENKO, M.; VAN DER MARK, P.; *et al*. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. **Systematic Biology**, v. 61, n. 3, p. 539–542, 2012.

RUTTEN, L.; RIBOT, C.; TREJO-AGUILAR, B.; *et al*. A single amino acid change (Y318F) in the L-arabitol dehydrogenase (LadA) from Aspergillus niger results in a significant increase in affinity for D-sorbitol. **BMC Microbiology**, v. 9, n. 1, 2009.

SANG, T.; DONOGHUE, M. J.; ZHANG, D. Evolution of alcohol dehydrogenase genes in peonies (Paeonia): phylogenetic relationships of putative nonhybrid species. **Molecular Biology and Evolution**, v. 14, n. 10, p. 994–1007, 1997.

SHEN, C.; YUAN, J.; OU, X; *et al*. Genome-wide identification of alcohol dehydrogenase (ADH) gene family under waterlogging stress in wheat (Triticum aestivum). **PeerJ**, v. 9, p. e11861, 2021.

SILVA, S. R.; GIBSON, R.; ADAMEC, L.; *et al*. Molecular phylogeny of bladderworts: A wide approach of *Utricularia* (Lentibulariaceae) species relationships based on six plastidial and nuclear DNA sequences. **Molecular Phylogenetics and Evolution**, v. 118, p. 244–264, 2018

SILVA, S. R.; MIRANDA, V. F. O.; MICHAEL, T. P.; *et al*. The phylogenomics and evolutionary dynamics of the organellar genomes in carnivorous *Utricularia* and *Genlisea* species (Lentibulariaceae). **Molecular Phylogenetics and Evolution**, v. 181, p. 107711, 2023.

SMALL, R. L.; CRONN, R. C.; WENDEL, J. F. Use of nuclear genes for phylogeny reconstruction in plants. **Australian Systematic Botany**, v. 17, n. 2, p. 145, 2004.

SPEIRS, J.; LEE, E.; HOLT, K.; *et al*. Genetic Manipulation of Alcohol Dehydrogenase Levels in Ripening Tomato Fruit Affects the Balance of Some Flavor Aldehydes and Alcohols1. **Plant Physiology**, v. 117, n. 3, p. 1047–1058, 1998.

SVENSSON, S.; HEDBERG, J. J.; HÖÖG, J. O. Structural and functional divergence of class II alcohol dehydrogenase. **European Journal of Biochemistry**, v. 251, n. 1–2, p. 236–243, 1998.

TAYLOR, P. The *Genus Utricularia* – A Taxonomic Monograph. Kew Bulletin Additional Series XIV. Royal Botanic Gardens, Kew: London, 1989

TANEJA, B.; MANDE, S. C. Conserved structural features and sequence patterns in the GroES fold family. **Protein Engineering, Design and Selection**, v. 12, n. 10, p. 815–818, 1999.

THOMPSON, C. E.; SALZANO, F. M.; DE SOUZA, O. N.; *et al*. Sequence and structural aspects of the functional diversification of plant alcohol dehydrogenases. **Gene**, v. 396, n. 1, p. 108–115, 2007.

THOMPSON, C. E.; SALZANO, F. M.; DE SOUZA, O. N.; *et al*. Sequence and structural aspects of the functional diversification of plant alcohol dehydrogenases. **Gene**, v. 396, n. 1, p. 108–115, 2007.

THORN, J. M.; BARTON, J. D.; DIXON, N. E.; *et al*. Crystal Structure of *Escherichia coli* QOR Quinone Oxidoreductase Complexed with NADPH. **Journal of Molecular Biology**, v. 249, n. 4, p. 785–799, 1995.

VAN DER STRAETEN, D.; POUSADA. R, R. A.; GIELEN, J.; *et al*. Tomato alcohol dehydrogenase. **FEBS Letters**, v. 295, n. 1–3, p. 39–42, 1991.

WATERHOUSE, A.; BERTONI, M.; BIENERT, S.; *et al*. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. W1, p. W296–W303, 2018.

WOLFE, K. H; LI, W. H; SHARP, P. M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. **Proceedings of the National Academy of Sciences**, v. 84, n. 24, p. 9054–9058, 1987.

YIJIE. F.; HUIHUI, G.; JUNMEI, W.; Yao, Z.; JIANFEI, W.; FANCHANG, Z., Meiosis, the master driver of gene duplication in higher plants? **Biochemical and Biophysical Research Communications**, v. 514, n. 3, p. 756–758, 2019.

ZHANG, G.; MA, H. Nuclear phylogenomics of angiosperms and insights into their relationships and evolution. **Journal of Integrative Plant Biology**, 2024.

ZENG, W.; QIAO, X.; LI, Q.; *et al*. Genome-wide identification and comparative analysis of the ADH gene family in Chinese white pear (*Pyrus bretschneideri*) and other Rosaceae species. **Genomics**, v. 112, n. 5, p. 3484–3496, 2020.

ZHU, B.; HAN, H.; FU, X.; *et al*. Degradation of trinitrotoluene by transgenic nitroreductase in *Arabidopsis* plants. **Plant, Soil and Environment**, v. 64, n. 8, p. 379–385, 2018.

CAPÍTULO 3- Comparative analysis of the chloroplast genome of *Pinguicula moranensis* Kunth (Lentibulariaceae)

Abstract

In recent years, researchers have tended to apply molecular approaches to fill the gaps when morphology has not satisfactorily filled in systematic and taxonomic questions or for studies on genetics and diversity. For that, genomes can be an essential source of information, and thus plant species have been sequenced, assembled, and annotated. Within the Lentibulariaceae family, the genus *Pinguicul*a L. is the one with the fewest published chloroplast genomes. In this study, we assembled, annotated, and described the complete chloroplast (pt) genome of *Pinguicula moranensis* Kunth, a carnivorous plant species that occurs in Mexico, Guatemala, and El Salvador and has tropical growth characteristics. With a total size of 147,072 bp, is currently the smallest chloropastid genome of the genus Pinguicula, compared to those published to date. It has a total guanine-cytosine (GC) content of 38.3% and presents a quadripartite structure composed of three regions: single copy region (LSC), short single copy region (SSC) and a short single copy region (SSC). It contains a pair of inverted repeat (IR) regions, typical of pt genomes. These regions occupied 81,393 bp, 11,537 bp, and 27,071 bp of the genome for LSC, SSC, and IR, respectively. 110 genes were found, of which 77 are protein-coding, 27 tRNA, four rRNA, and two pseudo genes. When comparing the identities of *Pinguicula vulgari*s and *P. jackii*, the results show that they are conserved and mostly similar. The maximum likelihood phylogenetic tree was generated with 69 protein-coding genes and recovered a monophyletic clade of *Pinguicula,* sister to a clade formed by *Genlisea - Utricularia.* In our analysis, *P. moranensis was* grouped with *P. jackii,* thus showing a close relationship. This study adds to the scarce genomes data on *Pinguicula* and provides clues to the evolution of the pt genomes in the Lentibulariaceae family.

Keywords: Genomic, Lentibulariaceae evolution, Phylogenetic analysis, ptDNA.

1- Introduction

Pinguicula L., together with *Genlisea* A.St.-Hil. and *Utricularia* L., represent the three genera of the carnivorous plant family Lentibulariaceae of the Lamiales (Taylor, 1989). The species in the genus *Pinguicula* have rosette-shaped leaves covered by sticky glands that secrete mucilage (Degtjareva and Sokoloff, 2012). These glands are adapted to capture small insects such as flies, mosquitoes, and other arthropods (Casper, 1996; Rondondi et al., 2010) which are attracted to the sticky substance and become trapped on the leaves. The digestive enzymes produced by the glands act to break down the prey macromolecules (main proteins) and thus the leaves absorb nutrients into the plant for its survival and the continuity of other vital functions. With around 100 species, the genus *Pinguicula* L. is the second most diverse in the family, after *Utricularia* L. (Legendre, 2013; Rondondi et al., 2010; Fleischmann, 2021; Shimai et al., 2021; Li and Liu, 2022). It occurs in the Northern Hemisphere and some taxa in South America and Africa, with the greatest specific diversity found in the Caribbean-Central American region, which is considered the center of diversity (Casper, 1996; Casper and Stimper, 2009; Domínguez et al., 2012; 2013; 2023). While the *Genlisea-Utricularia* clade lacks a root system, *Pinguicula* has adventitious roots (Corrêa et al., 2002), thus corroborating phylogenetic studies based on DNA which point to *Pinguicula* as a sister group of the *Genlisea-Utricularia* clade (Legendre, 2013; Muller et al., 2004; Silva et al., 2016; 2018; 2023).

Belonging to the section *Mesoamerican*, *Pinguicula moranensis* Kunth is a perennial plant found the high regions of Mexico, Guatemala and El Salvador (Zamudio, 2001; Grob et al., 2007; Shimai et al., 2021). Their habitats are preferably rocky, humid, and shaded (Villegas and Alcalá, 2017). Given its occurrence in tropical regions, it flowers almost year-round (Zamudio and Rzedowsky, 1991), presenting two rosettes with dimorphic leaves (Rueda-Almazán, 2021). Therefore, in the months of June to October, which correspond to summer, the species produces a rosette called a capture rosette, with 6-8 tiny leaves. In winter, specifically from November to May, the rosettes have around 60 to 100 leaves of limited length (Zamudio, 1999). It has hermaphrodite flowers, with a style measuring 15–38 mm; corollas vary in length from 60 to 64 mm (Zamudio, 2001; Grob et al., 2007; Villegas and Alcalá, 2017), with zygomorphic flowers of purple color; the flowers are trilobal, despite the upper lips being bilobed. Five to six weeks after pollination by dipterans, fruiting appears, which takes place from July to September (Alcalá and Domínguez, 2005; Grob et al., 2007). The fruits are in the form of capsules that house around 400 seeds (Villegas and Alcalá, 2017).

Several studies based on molecular and morphological characteristics have been important for the taxonomic classification of species in this family (Muller et al., 2004; Rutishauser, 2015; Silva et al., 2016a). In recent years, some species of Lentibulariaceae had their genomes sequenced and assembled, where was found that some taxa in this family have smaller genomes and present high rates of nucleotide substitutions in their genomes, mainly in species of the *Genlisea-Utricularia* clade (Albert et al., 2009, Silva et al., 2018 e 2019, Johnson and Albert, 2005). Genetic information is essential for managing information. and mapping evolutionary trajectories (Li and Liu, 2022). These molecular tools have strongly contributed to taxonomic resolution or adjustment, the discovery of new species, and providing clues about the evolution, similarity, and genetic behavior of species (Wang et al., 2024). Therefore, research on the nuclear, mitochondrial, and chloroplast genomes of plants is extremely important, as it allows knowledge about the genetic information and evolution of the genes that make up the species' genomes, as well as the phylogenetic relationships of intra- and interspecific.

The chloroplast genome (ptDNA) has a quadripartite structure, composed of two inverted repeat regions (IRa and IRb) and two cyclically unique regions, one of large size (LSC) and the other of small size (SSC) (Palmer, 1985). Its polymorphism plays a crucial role in resolving taxonomic and phylogenetic discrepancies, contributing to the development of genetic barcodes, population genetics studies, the construction of high-resolution phylogenetic trees, and the conservation of threatened species (Downie and Palmer 1992; Daniell et al., 2016; Robbins and Kelly, 2023; Wang et al., 2024). ptDNAs contain a double-stranded circular genome with a size ranging from 107 kb to 218 kb, which encodes about 120 genes, including ribosomal RNA (rRNA) and transfer RNA (tRNA), and genes encoding different proteins (Daniell et al., 2016). Studies on the chloroplast genome have registered a progression, because until 2023 approximately 13,000 studies had been published and of that number, flowering plants cover the largest percentage with around 93.4% (Wang et al., 2024). Chloroplast genomes provide crucial information for phylogenetic studies. However, for the genus *Pinguicul*a to date, only *Pinguicula ehlersiae* Speta & F.Fuchs (Wicke et al., 2013), *Pinguicula alpina* L. (Li and Liu, 2022), *Pinguicula vulgaris* L. (NC_084256.1)*,* and *Pinguicula jackii* Barnhart (NC_068629.1) had their genomes assembled or available on the NCBI platform. In this research, we analyzed the genome of *Pinguicula moranensis* and inferred its phylogenetic position within the family Lentibulariaceae.

2- Materials and methods

2.1- Assembly and annotation of the chloroplast genome *Pinguicula moranensis*

The raw DNA sequencing reads of *Pinguicula moranensis* were extracted from the database SRA of NCBI (National Center for Biotechnology Information) with the identification SRP390263, sequenced on an Illumina Hi-Seq platform in paired-end, 2 × 150 bp format (Procko et al., 2023). The annotations of the chloroplast genes of *P. moranensis* were made using the GESeq tool of the CHLOROBOX Web Service software (Tillich et al., 2017) using previously published Lentibulariaceae genomes. Subsequently, the TBL created by CHLOROBOX, Sequin 10.3 was used to validate and correct the start and stop codons as well as the intron and exon boundaries for the genes with software annotation errors. The genome maps were drawn using the Organellar GenomeDraw v. 1.6.4 (OGdraw) software (Lohse et al., 2013).

The tRNAs were predicted using the tRNAscan-SE v2.0.7 software (Wood, Lu and Langmead, 2019), while the rRNAs were predicted using RNAmmer v.1.2 (Lagesen et al., 2007), both implemented by the CHLOROBOX GeSeq platform. Pseudogenes were characterized according to the absence of a start and/or stop codon, frameshift, and genes with >50% coverage and >50% coding region identity. Final adjustments were made using

Blastn and tBlastx on ptDNAs from *Utricularia foliosa* L. (KY025562.1) and *Pinguicula alpina* L. (NC_056190.1).

2.2- Identity and synteny analyses

To find out the degree of identity or variation of the ptDNA of *Pinguicula moranensis* within the genome, we used the online software mVISTA v2.0 (http://genome.lbl.gov/vista/index.shtml) (Frazer et al., 2004) where sequences similarity can be assessed by comparing coding regions with non-coding regions, introns, and exons. Therefore, we compared the ptDNA of two species belonging to the same genus: *P. jackii* and *P. vulgaris*). The genome of *P. vulgaris* was used as a reference. To detect the arrangements between the genes, the analysis in mVISTA was carried out using the Shuffle-LAGAN.

2.3- Repetitive regions

The repetitions that occur in the ptDNA of *P. morenensis* were compared with the ones of *P. jackii* (NC_068629.1) and *P. vulgaris* (NC_084256.1). The MISA perl script (Beier et al., 2017) was used to calculate simple sequence repeats (SSRs), following parameters with limits of seven repetitions for mononucleotide SSRs, four repetitions for di- and trinucleotide SSRs, and three repetitions for tetra SSRs, penta SSRs, and hexanucleotides. The REPuter software v2 (Kurtz et al., 2001) was applied for detecting longer repeats (forward, reverse, complementary, and palindromic), using the parameters of Maximum Computed Repeats 90, Minimal Repeat Size > 30, and a Hamming distance of 0.01.

2.3- Phylogenetic reconstruction of P*inguicula moranensis*

Phylogenetic analyzes were estimated using the maximum likelihood (ML) approach with a concatenated matrix generated by FASCONCAT-G v1.05.1(Kück and Longo, 2014) composed of 69 coding genes (*acc*D, *atp*A, *atp*B, *atp*E, *atp*F, *atp*H, *atp*I, *ccs*A, *cem*A, *clp*P, *inf*A, *mat*K, *pet*A, *pet*B, *pet*D, *pet*G, *pet*L, *pet*N, *psa*A, *psa*B, *psa*C, *psa*I, *psa*J, *psb*A, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*I, *psb*J, *psb*K, *psb*L, *psb*M, *psb*N, *psb*T, *psb*Z, *rbc*L, *rpl*14, *rpl*16, *rpl*2, *rpl*20, *rpl*22, *rpl*23, *rpl*32, *rpl*33, *rpl*36, *rpo*A, *rpo*B, *rpo*C1, *rpo*C2, *rps*11, *rps*12, *rps*14, *rps*15, *rps*16, *rps*18, *rps*19, *rps*2, *rps*3, *rps*4, *rps*7, *rps*8, *ycf*1, *ycf*15, *ycf*2, *ycf*3 and *ycf*4.). Sequences were extracted from 12 species of the Lentibulariaceae family using the Geneious v9.0 tool (Kearse et al., 2012), followed by individually alignment in MAFFT online v.7 (Katoh, 2017). The species used were: *Pinguicula moranensis* (this study), *Pinguicula jackii* (NC_068629.1), *Pinguicula alpina* (NC_056190.1), *Pinguicula ehlersiae* (HG803178.1), *Pinguicula vulgaris* (NC_084256.1), *Utricularia reniformis* (NC_029719.2), *Utricularia foliosa* (KY025562.1), *Utricularia amethystin*a (MN223722.1), *Genlisea aurea* (NC_037078.1), *Genlisea tuberosa* (NC_037082.1), *Genlisea repen*s (NC_037081.1) and *Genlisea violacea* (NC_037083.1).

The ML tree was calculated using IQ-Tree 1.6.12 (Nguyen et al., 2014) with the best-of-fit model TIM2+F+I+G4, according to BIC criteria (Akaike, 1974) calculated with the ModelFinder v.1.4.2 software (Kalyaanamoorthy et al., 2017). Clade support was estimated with the ultrafast bootstrap (UFBoot) and SH-aLRT algorithms (Hoang et al., 2018) with 1,000 replicates. To root the phylogenetic tree, the following species were used as an outgroup: *Incarvillea arguta* Royle (NC_045915.1), *Tanaecium tetragonolobum* (Jacq.) L.G.Lohmann (NC_027955.1) and *Sesamum indicum* L. (NC_016433.2). The visualization and editing of the generated file “treefile” of the phylogenetic tree were carried out with the aid of the iTOL -Interactive Tree of Life v5 (Letunic and Bork, 2021) and Inkscape v1.3.2 (https://inkscape.org).

To assess the phylogenetic signal and thus the potential level of contribution of each gene for the phylogenetic tree, we determined the number of parsimony informative characters (PICs) using the ptDNA sequences extracted from the respective genomes, using PAUP* v. 4.0a169 (Swofford, 1998).

3- Results and discussion

3.1- Genome structure and content

The ptDNA of *Pinguicula moranensis*, assembled in the present study, has a size of 147,072 bp (Fig. 1). As expected, this genome has a circular, quadripartite structure resembling other angiosperms (Palmer, 1985; Daniell et al., 2016; Silva et al., 2019; Li and Liu, 2022). Our results demonstrate a large single-copy region (LSC) with 81,393 bp, a small single-copy region (SSC) with 11,537 bp, and two inverted regions (IR) with 27,071 bp each. The overall GC content of the genome is 38.3%.

Figure 1. Genomic map of the *Pinguicula moranensis* chloroplast genome. Genes shown on the outside of the map are transcribed clockwise, whereas genes on the inside are transcribed counter-clockwise. Genes are color coded by their function in the legend.

Recently, species from the Lentibulariaceae family had their chloroplast genomes assembled, mainly taxa from the *Genlisea-Utricularia* clade (Park et al., 2022; Silva et al., 2016; 2018, 2019; Wicke et al., 2013). Within the genus *Pinguicula*, the species *P. ehlersiae* (Wicke et al., 2013), *P. alpina* (Li and Liu, 2022), *P. jacki*i (NC_068629.1) and *P. vulgaris* (NC_084256.1) are the example.

Interestingly, the results of this research place the genome of *Pinguicula moranensi*s as the smallest within the genus of *Pinguicula* when compared to the first assembled genomes and/or sequences published to date. This is followed by *P. ehlersiae* (Wicke et al., 2013) with 147,147 bp, *P. alpina* (Li and Liu, 2022) with 147,479 bp, *P. vulgaris* (NC_084256.1) with 149,996 and *P. jackii* (NC_068629.1) is the largest with 152,435 bp. Regarding the GC content of the LSC, SSC and IRs regions, it is 36.3%, 31.6% and 42.7%, respectively. Despite a slight superiority, the GC values obtained in *P. moranensis* are very close to those presented by the species *P. alpina* which presented content of 36.1%, 31.4% and 43.0% for the the regions LSC, SSC and IRs, respectively (Li and Liu, 2022).

3.1.1- List of genes encoded by the *Pinguicula moranensis* ptDNA

The genomes of angiosperm chloroplasts contain 30 to 120 genes that mainly participate in photosynthesis, transcription and translation (Daniell et al., 2016). We found a total of 113 genes for *Pinguicula moranensis* (Fig. 1; Table 1) which 77 are protein coding, 30 tRNA, 4 rRNA and 2 pseudogenes (*ndh*A and *ndh*F). These tRNAs represent the possible amino acid anticodons, including two methionines (Ile), predicted by tRNAscan-SE software (Wood, Lu and Langmead, 2019). In the repeat regions (IRs), a total of 19 duplicated genes were observed, 9 of which can encode proteins (*rps*7, rps12, *ndh*B, *rpl*2, *rpl*23, *ycf*1, *ycf*2 and *ycf*15), 7 tRNA genes (trnN-GUU, trnA-UGC, trnI-CAU, trnI-GAU, trnL-CAA, trnR-ACG and trnV-GAC), in addition to the four genes linked to the ribosomal complex, that is, the rRNA (rrn4.5, rrn5, rrn16 and rr23). It is important to highlight that the genes rps16, atpF, *rpo*C1, *clp*P, petB, *pet*D, *rpl*2, *ndh*B, which encode proteins, as well as the RNA transporters trnI-GAU, trnA-UGC presented an intron, while the genes clpP and *ycf*3 (*paf*I) had two introns in their structure, as found in other Lentibulariaceae (Silva et al., 2016). All of these genes participate, whether directly or indirectly, in the process of photosynthesis, protein biosynthesis, transport of substances necessary for photosynthesis, as well as in the structural maintenance of the genome itself.

Table 1: Annotated *Pinguicula moranensis* genes.

Category	**Genes annotated**
Photosystem I	*psaA, psaB, psaC, psaI, psaJ*
Photosystem II	*psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN* (*pbf*1), *psbT, psbZ*
Cytochrome b/f complex	*petA, petB•, petD•, petG, petL, petN*
ATP synthase	*atpA, atpB, atpE, atpF•, atpH, atpI*
NADH dehydrogenase	*ndhAΨ, ndhB (2×), ndhD, ndhE, ndhF Ψ, ndhG, ndhH, ndhI, ndhJ, ndhK*
RubisCO large subunit	*rbcL*
RNA polymerase	*rpoA, rpoB, rpoC1•, rpoC2*
Ribosomal proteins (SSU)	*rps2, rps3, rps4, rps7 (2x), rps8, rps11, rps12* ^•(2×), *rps14, rps15, rps16 •, rps18, rps19*
Ribosomal proteins (LSU)	*rpl2* (2×)•, *rpl14, rpl16, rpl20, rpl22, rpl23 (2×), rpl32, rpl33, rpl36*
Other genes	*ccsA, clpP •, matK, accD, cemA, infA*
hypothetical chloroplast reading frames	*ycf1* (2×*), ycf2* (2×)*, ycf3 (pfa*1*), ycf4, ycf15* (2×)
Transfer RNAs	*trnA-UGC* (2×*)•, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA , trnG-UCC•, trnH-GUG, trnI-CAU* (2×), *trnI-GAU* (2×)•, *trnK-UUU•, trnL-CAA* (2×), *trnL-UAG, trnM-CAU, trnN-GUU (2×), trnP-UGG, trnQ-UUG, trnR-ACG (2×), trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC* (2×), *trnW-CCA, trnY-GUA, trnfM-CAU*
Ribosomal RNAs	*rrn4.5* (2×), *rrn5* (2×), *rrn16* (2×), *rrn23* (2×)

Ψ pseudogene
^ trans-splicing
• intron-containing gene

It is also important to highlight that in our results of the 11 genes of the *ndh* complex (NADH dehydrogenase), eight are present as intact copies and two are pseudogenes (*ndh*A and *ndh*F), only missing *ndh*C. Taking into account the importance of the gene complex (*ndh*) that consists of stabilizing stress and carbon dioxide concentrations (Silva et al., 2018), and considering that Lentibulariaceae taxa generally develop aquatic ecosystems with high stresses of different orders (Ellison and Adamec, 2018), the presence of NAD(P)H-dehydrogenase complex genes are of extreme importance for the development of highly specialized leaves, as well as the survival of plants through nutrients that are digested from prey captured in leaf traps (Adamec et al., 2021). The two pseudogenes found in this work (ndhA and ndhF) were also detected in other species of Lentibulariaceae, as fragmentation of ndh genes generally occurs in this family (Silva et al., 2016), factors that probably occurred with this genome.

3.2- Analysis of identity or variation of genes in the genome

As expected, the results on identity analysis of the *Pinguicula moranensis* ptDNA assembled in this study, presented a highly conserved and similar structure compared to the genomes of *P. vulgaris* (used as reference) and *P. jackii* (Fig 2). These are phylogenetically close species within the genus and have the same habitat characteristics, which are poor and rocky soils. (Cieslak, 2005; Degtjareva et al.; 2006; Lustofin et al., 2020; Shimai et al., 2021) The results show the identity of the nucleotide sequences of these species is greater than 50% in almost all protein-coding genes. High degrees of variations were detected in RNA transporter genes such as trnH-GUG, trnQ-UUG, trnF-GAA, trnE-UUC, trnC-GCA, trnD-GUC, trnG-UCC, trnS-GGA and in intergenic spacers trnK -rps16, rps16-trnQ, psbK-psbI, trnL-trnF and trnH-psb, atpH-atpT, *atp*B-rbcl, psaA-ycf3, petA-psbJ, rbcL-aacD, rpoC1-rpoB and petN-psbM.

These variations in non-coding genes possibly occurred due to the fact that in Lentibulariaceae, the non-coding regions of the plastid genome have high substitution rates, and some genes have experiences relaxed purifying selection in comparison to other closely related lamiids (Silva et al., 2023). Furthermore, some coding genes such as *ycf*1, *rpl*2, *atp*F, *rpo*C1, *pet*B, *pet*D, *rps*12, *rps*19, *ccs*A, *rpo*A *rpl*16 and almost all of the NDH complex showed certain similar variations or divergences. Generally, the protein-coding genes of taxa of the Lentibulariaceae family present high degrees of non-synonymous genetic substitutions, mainly those of the *Utricularia-Genlisea* clade (Palmer, 1985 and Wicke et al., 2013).

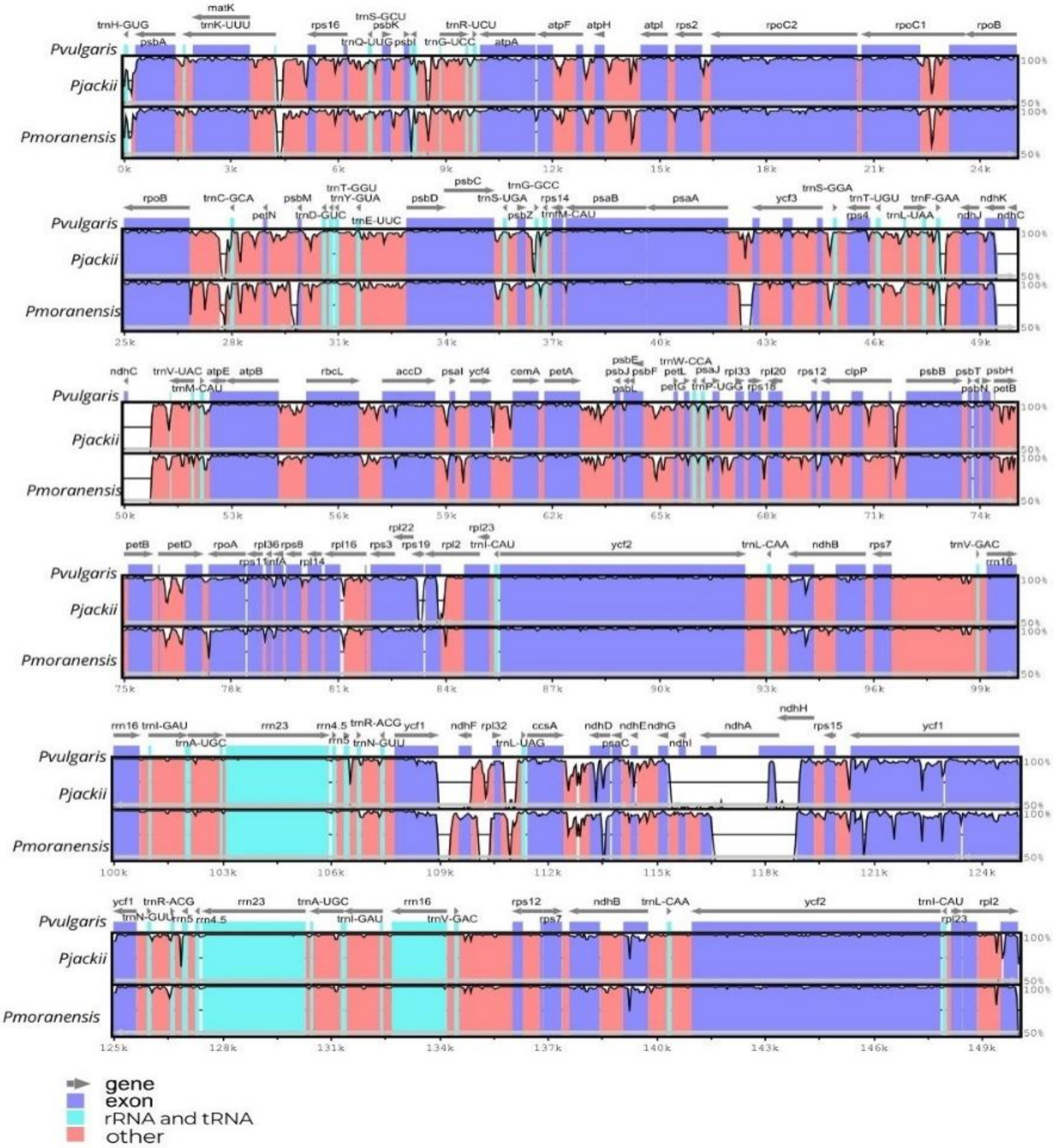


Fig 2. Comparative analysis of identity or variations of the chloroplast genes of *Pinguicula moranensis* with *P. jackii* and *P. vulgaris*, obtained by the mVISTA software, based on the Shuffle-LAGAN alignment. The *Pinguicula vulgaris* chloroplast genome sequence was used as a reference. Regions colored in purple represent exons. Regions colored red represent non-coding sequences. The light blue regions represent introns. Arrows below gene names indicate their orientation. The axis represents the identity measurements between each of the sequences and the reference (*P. vulgaris*).

As can be seen in figure 2, the degree of similarity in the compared species was also influenced by the absence or presence of certain genes of the ndh complex (NADH dehydrogenase) or pseudogenization. In *Pinguicula vulgari*s all genes of the NADH complex are present, in *Pinguicula moranensis* only the ndhC gene is reported as absent, while in *Pinguicula jackii* only the ndhA gene was recorded in ptDNA. The progressive degradation or loss of plastid ndh genes in Lentibulariaceae has been discussed in previous studies (Silva et al., 2018; 2023; Wicke et al., 2013).

3.3- Analysis of repeats in the genome

Chloroplast genome is characterized by the presence of repetitive regions, among them are palindromic, complementary, forward, reverse repeats and SSRs (simple repeated sequences) or microsatellites and they play essential roles in DNA replication and genome stability, in species identification, species conservation, and as genetic markers (Downie and Palmer 1992; Saltonstall and Lambertini, 2012; Clivati et al., 2012; Xu et al., 2023). In this study, the number of microsatellites identified in the chloroplast genome of three *Pinguicula* species by the MISA perl script (Beier et al., 2017) are represented in table 2. The program identified a total of 339, 371 and 363 microsatellites in *P. moranensis*, *P.jackii* and *P.vulgaris* respectively.

In the three species, the mononucleotide repeats stood out the most with more than 86%, that is, 295, 323 and 314 of the repeats in the A/T and C/G motifs. Dinucleotides are the second most represented, with more than 10%, having presented 36 repetitions for *Pinguicula moranensis* and 40 for *P. jackii* and *P. vulgaris*. Tri and tetranucleotides represent approximately 4% of the repeats of the total microsatellites detected. These present an almost similar amount between species. Penta- and hexanucleotide repeats were not found in any species. These results once again indicate the close similarity of the studied species Silva et al. (2019), studying the microsatellites in three morphotypes of *U. amethystina*, concluded that the mono- and direct nucleotidic sequences were the most represented. It is important to highlight that these repeats were found both in coding regions and in intergenic regions and in RNA transport genes. This is probably caused by the high substitution rates that the genomes of species from the Lentibulariaceae family present, helping in their genomic rearrangement. It is known in advance that in areas of genetics, microsatellites are fundamental, as they serve for the development of markers for research into evolutionary biology, domestication and genetic composition (Lopes et al., 2021). Previous studies confirm that the pt genomes of species belonging to the Lentibulariaceae family have a high number of microsatellites. Using microsatellites from species in this family, primer pairs were developed for taxons in the same genus, with the possibility of being used in studies of other genera within Lentiburaliaceae (Clivati et al., 2012; Aranguren-Díaz et al., 2017).

Table 2. Distribution of microsatellites found in the chloroplast genome of *Pinguicula moranensis, P. jackii* and *P. vulgaris* by the MISA software, in terms of their composition in terms of motif type.

SSR	***P.moranensis***	***P.jackii***	***P.vulgaris***
A/T	268	299	291
C/G	27	24	23
AC/GT	2	2	2
AG/CT	16	18	18
AT/AT	17	19	19
CG/CG	1	1	1
AAT/ATT	3	5	5
AAAT/ATTT	2	0	1
AATC/ATTG	1	0	0
AAAC/GTTT	0	1	1
AATC/ATTG	0	2	2
AATT/AATT	1	0	0
AGGAT/ATCC	1	0	0

The palindromic repeats, complementary, forward and reverse of the species *Pinguicula moranensis*, *P.jackii* and *P.vulgaris* detected by the REPuter software, are shown in tables 3, 4 and 5. Interestingly, repeats were detected in all quadripartite regions of the chloroplast genome, as happened in other studies involving Lentibulariaceae species (Silva et al., 2016a). The *ycf*2 and *ycf*3 genes have the longest repeats in all three species. In addition, the ndhK-trnK(UAC) spacer in *P. jackii* and *ndh*B in *P. vulgaris* have different repeats, showing high mutations and altering the entire genome.

Table 3. Sequence repeats in *P. moranensis*. Type, length, region, location and plastome quadripartite region for each repeat are indicated (Type: F = Forward repeat; P= Palindromic repeat; C = Complementary; R= Reverse)

Repeat 1	Type	Length (bp)	Repeat 2	Gene	Location	Region
81393	P	27071	120001	*ycf2*	CDS	IR
71682	P	42	71682	*psbT-Pbf1*	IGS	LSC
111266	P	35	111308	*ccsA-ndhD*	IGS	SSC
2945	F	33	77979	*petN-trnK(UUU)*	IGS	LSC
7851	P	30	44357	*psbI-trnS(GCU)*	IGS	LSC
79627	P	30	79627	*rpl16*	intron	LSC
112927	P	36	112927	*ndhD-psaC*	CDS	SSC
88554	F	37	88578	*cf2*	CDS	IR
88554	P	37	139850	*Ycf2*	CDS	IR
88578	P	37	139874	*Ycf2*	CDS	IR
139850	F	37	139874	*cf2*	IGS	IR
15838	F	36	15839	*rps2-rpaC2*	IR	LSC
43199	F	36	95500	*pafI*	intron	LSC
43199	P	36	132929	*pafI*	intron	LSC
43202	F	35	92466	*pafI*	intron	LSC
43202	P	35	135964	*pafI*	intron	LSC
38509	F	33	40733	*psaB*	CDS	LSC
7849	F	32	35332	*psbI-trnS(GCU)*	IGS	LSC
56959	P	31	56987	*accD-psaI*	IGS	LSC
8677	C	30	63697	*trnS(GCU)-trnR(UCU)*	CDS	LSC
9290	F	30	36297	*trnS(GCU)-trnR(UCU)*	IGS	LSC
35334	P	30	44357	*psbC-trns(UGA)*	IGS; intron	LSC

60943	P	30	60978	*petA-psbJ*	IGS	LSC
86175	F	30	86217	*ycf2*	CDS	IR
86175	P	30	142218	*ycf2*	CDS	IR
86217	P	30	142260	*ycf2*	CDS	IR
142218	F	30	142260	*ycf2*	CDS	IR

Table 4. Sequence repeats in *P. jackii*. Type, length, region, location and plastome quadripartite region for each repeat are indicated (Type: F = Forward repeat; P= Palindromic repeat; C = Complementary; R= Reverse)

Repeat 1	**Type**	**Length (bp)**	**Repeat 2**	**Gene**	**Location**	**Region**
82995	P	33116	119319	*ycf2*	CDS	IR
0	P	409	82586	*ycf2*	CDS	IR
50077	R	55	50077	*ndhK-trnK(UAC)*	CDS	IR
50083	R	50	50083	*ndhK-trnK(UAC)*	IGS	LSC
50077	F	49	50084	*ndhK-trnK(UAC)*	IGS	LSC
50077	R	48	50077	*ndhK-trnK(UAC)*	IGS	LSC
57903	P	47	57903	*accD- psaI*	IGS	LSC
50090	R	43	50090	*ndhK-trnK(UAC)*	IGS	LSC
50077	F	42	50091	*ndhK-trnK(UAC)*	IGS	LSC
50077	R	41	50077	*ndhK-trnK(UAC)*	IGS	LSC
50097	R	36	50097	*ndhK-trnK(UAC)*	IGS	LSC
72842	P	45	72842	*ndhK-trnK(UAC)*	IGS	LSC
50077	F	35	50098	*ndhK-trnK(UAC)*	IGS	LSC
50077	R	34	50077	*ndhK-trnK(UAC)*	IGS	LSC
114144	P	34	114144	*psbT-psbN*	IGS	LSC
114144	F	34	121252	*psbT-psbN*	IGS	LSC
121252	P	34	121252	*ndhA*	CDS	IR

76395	P	43	76395	*petD-rpoA*	IGS	LSC
3352	F	33	79171	*matk-trnK(UU)*	IGS	LSC
8322	P	30	45219	*psbI-trnS(GCU)*	IGS	LSC
80821	P	30	80821	*Rpl16*	intron	LSC
118942	P	36	118942	*ndhD-psaC*	IGS	IR
89750	F	37	89774	*Ycf2*	CDS	IR
89750	P	37	145619	*Ycf2*	CDS	IR
89774	P	37	145643	*Ycf2*	CDS	IR
145619	F	37	145643	*Ycf2*	CDS	IR
50071	F	34	50099	*ndhK-trnK(UAC)*	IGS	LSC
44050	F	36	96691	*Ycf3*	intron	LSC
44050	P	36	138703	*Ycf3*	intron	LSC
50103	R	30	50104	*ndhK-trnK(UAC)*	IGS	LSC
44053	F	35	93662	*Ycf3*	intron	LSC
44053	P	35	141733	*Ycf3*	intron	LSC
39110	F	33	41334	*psaB,psaA*	CDS; IGS	LSC
115759	P	30	115759	*ndhE-trnL(UAG)*	CDS; IGS	IR
115759	F	30	119641	*ndhE-trnL(UAG)*	CDS; IGS	IR
119641	P	30	119641	*psaC-ndhE*	CDS; IGS	SSC
8320	F	32	35919	*psbT-trnS(GCU)*	IGS	LSC
57903	R	31	57911	*accD-psaI*	IGS	LSC
9767	F	30	36886	*trnG(UCC)-*	tRNA	LSC
35921	P	30	45219	*psbC-trns(UGA)*	IGS; intron	LS
87371	F	30	87413	*ycf2*	CDS	IR
87371	P	30	147987	*ycf2*	CDS	IR
87413	P	30	148029	*ycf2*	CDS	IR
147987	F	30	148029	*ycf2*	CDS	IR

Table 5. Sequence repeats in *P. vulgaris.* Type, length, region, location and plastome quadripartite region for each repeat are indicated (Type: F = Forward repeat; P= Palindromic repeat; C = Complementary; R= Reverse).

Repeat 1	**Type**	**Length (bp)**	**Repeat 2**	**Gene**	**Location**	**Region**
95457	P	13492	124436	*ndhB; rps7; ycf1*	CDS; IGS	LSC
95748	P	13201	124436	*ndhB; rps7; ycf1*	CDS; IGS	LSC
83389	P	10294	139702	*psaA-ycf3*	CDS;intron	LSC
94181	P	869	138335	*psaA-ycf3*	CDS	IR
94265	P	785	138335	*ndhB*	CDS	IR
95304	P	441	137640	*ndhB*	CDS	IR
95223	P	231	137931	*ndhB*	CDS	IR
95124	P	178	138083	*ndhB*	CDS	IR
93947	P	147	139291	*ndhB*	CDS	IR
93986	P	135	139264	*ndhB*	CDS	IR
93829	P	116	139440	*ndhB*	CDS	IR
93875	P	109	139401	*ndhB*	CDS	IR
94166	P	97	139122	*ndhB*	CDS	IR
93734	P	78	139573	*ndhB*	CDS	IR
95052	P	66	138267	*ndhB*	CDS	IR
94097	P	66	139222	*ndhB*	CDS	IR
93757	P	55	139573	*ndhB*	CDS	IR
93818	P	55	139512	*ndhB*	CDS	IR
94124	P	55	139206	*ndhB*	CDS	IR

110903	F	50	110944	*Rpl32-trnL(UAG)*	IGS	LSC
113683	P	36	113683	*ndhD-psaC*	IGS	SSC
77344	P	45	77344	*petD-rpoA*	IGS	LSC
28229	F	35	28245	*trnC-petN*	IGS	LSC
112462	P	35	112504	*ccsA-petN*	IGS	LSC
97383	F	35	116811	*Rps12-trnV(GAC)*	IGS	LSC
116811	P	35	135967	*ccsA*	CDS	SSC
93716	P	40	139629	*ndhB*	CDS	IR
2899	F	30	80111	*mat-trnK(UUU)*	IGS	LSC
8069	P	30	44956	*Ycf2*	CDS	IR
90533	F	37	90557	*Ycf2*	CDS	IR
90533	P	37	142791	*Ycf2*	CDS	IR
90557	P	37	142815	*Ycf2*	CDS	IR
142791	F	37	142815	*Ycf2*	CDS	IR
43767	F	36	97385	*Ycf3*	intron	LSC
43767	P	36	135964	*Ycf3*	intron	LSC
43770	F	35	94352	*Ycf3*	intron	LSC
43770	P	35	138998	*Ycf3*	intron	LSC
120280	P	32	120280	*rps15-ycf1*	IGS	LSC
68971	R	31	68971	*rpsl20-rps12*	IGS	LSC
38775	F	33	40999	*psaB*	CDS	LSC
76562	P	30	76562	*petD*	intron	LSC
8067	F	32	35613	*psbI-trnS(GCU)*	IGS; intron	LSC
120294	P	31	120294	*rps15-ycf1*	IGS	LSC
7806	P	30	8021	*psbK-psbI*	IGS	LSC
9605	F	30	36568	*trnG(UCC)*	tRNA	LSC

35615	P	30	44956	*psbC-trnS(UGA)*	IGS	LSC
43770	F	30	116816	*Ycf3*	CDS	LSC
51601	P	30	68965	*trnV(UAC)-trnM(CAU)*	IGS	LSC
88154	F	30	88196	*Ycf2*	CDS	IR
88154	P	30	145159	*Ycf2*	CDS	IR
88196	P	30	145201	*Ycf2*	CDS	IR
120280	F	30	120299	*rsp15-ycf1*	IGS	LSC
145159	F	30	145201	*Ycf2*	CDS	IR

The results indicate a total of 27, 43 and 53 repeats in *Pinguicula moranensis, P.jackii* and *P.vulgaris*, respectively (Fig3). Of these, palindromic repeats (with 15, 19 and 37) and Forward repeats (11, 15 and 15) were the most detected. These occurred mainly in intergenus regions and in tRNAs. Observing the results, it can be deduced that the forward repeats seem to be more conserved among the chloroplast genomes of these species, as there was little variation between them, these repeats were between 11 (*P. moranensis*) and 15 (*P. jackii* and *P. vulgaris*) repeats in the genome.

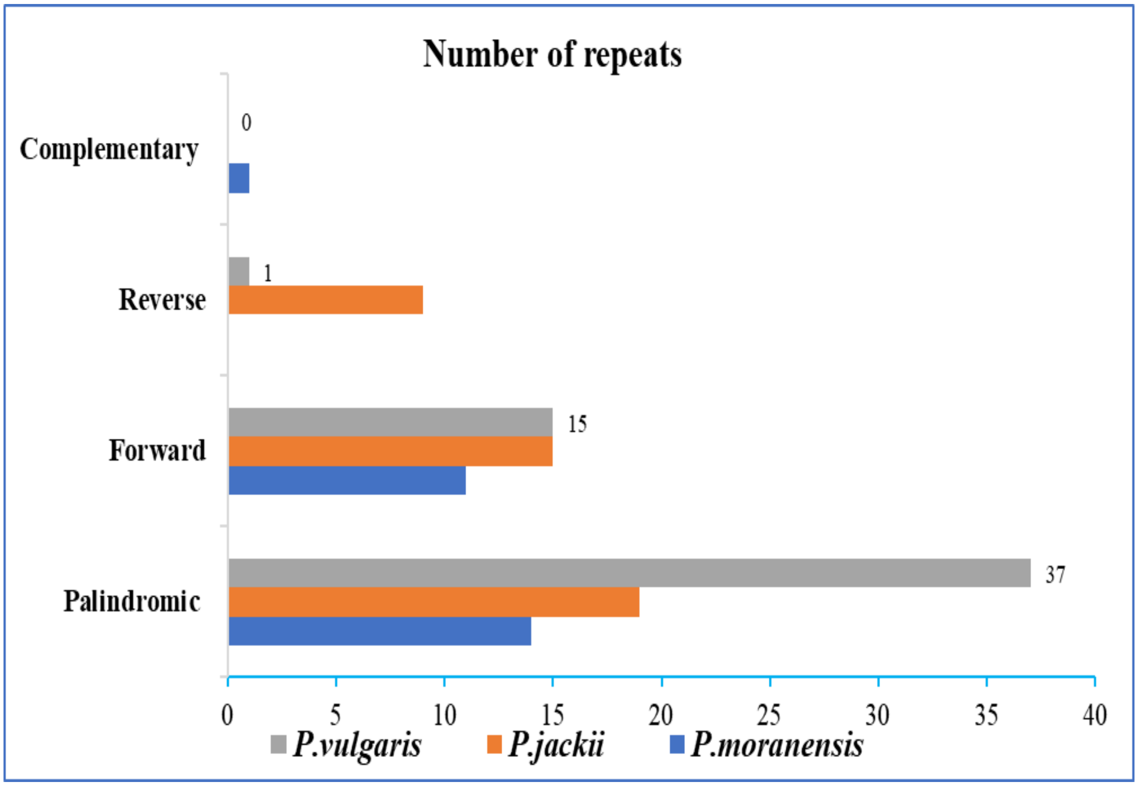


Fig 3: Number of palindromic repeats, complementary, forward and reverse in the pt genome of *Pinguicula moranensis, P.jackii* and *P.vulgaris*.

3.4- Phylogenetic hypothesis of *Pinguicula moranensis* within Lentibulariaceae

Phylogenetic inferences based on sequences of 69 chloroplast genes of the species used, using Maximum Likelihood approaches, resulted in totally congruent topologies within the Lentibulariaceae family (Fig. 4), as the chloroplast sequences of Lentibulariaceae species presented a monophyletic group, converging with studies by other authors (Müller et al., 2006; Silva et al., 2018). All clades exhibited identical phylogenetic topologies, and bootstrap support values based on 1000 replications were 100% across all branches, making the tree highly congruent. Despite the low representation of the family's species in the present study, the results show that the groups formed monophyletic clades, supporting the idea that the genus *Pinguicula* is a sister clade to *Utricularia-Genlisea* (Taylor, 1989; Casper, 1996, Silva et al, 2017; Fleischmann, 2021), and the monophyly of each genus with multiple species was strongly supported.

Our results place *Pinguicula moranensis* as a monophyletic group with *P. jackii*, indicating a close relationship between them within the genus and family.

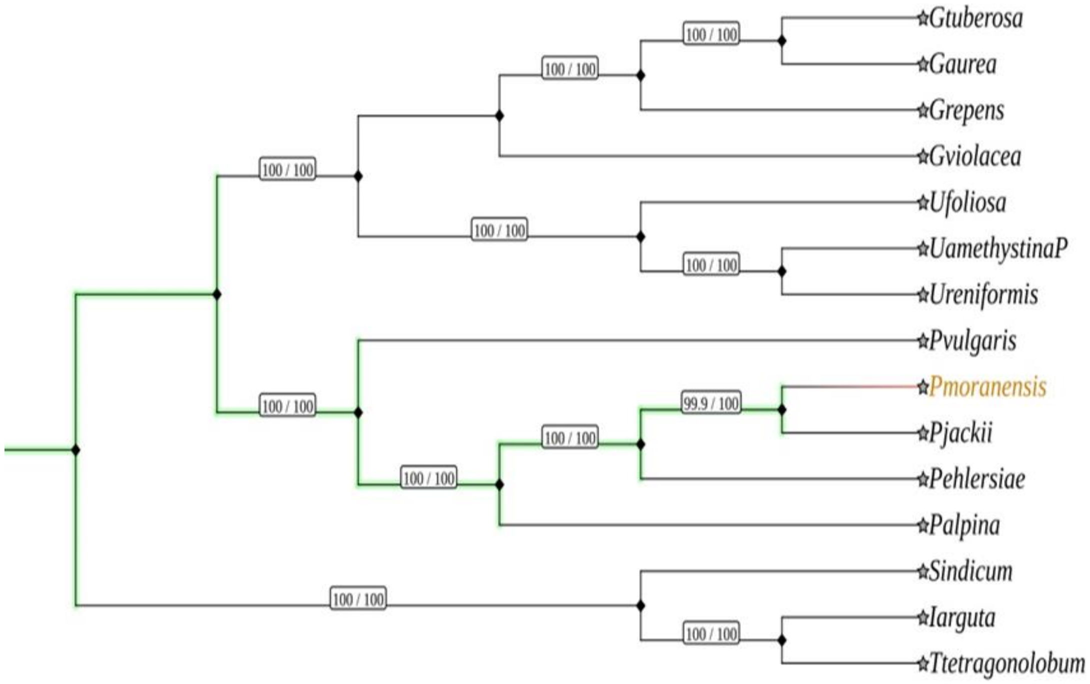


Fig. 4. Maximum likelihood phylogenetic tree constructed from the 14 chloroplast genome sequences in Lentibulariaceae. Bootstrap support values and SH-aLRT based on 1000 replications were 100% across all branches (indicated next to the respective nodes). The support values (posterior probabilities, bootstrap of the ML analysis and SH-aLRT) of each clade are indicated next to the respective nodes.

The taxonomy of the genus *Pinguicula* has undergone updates since the first description made by De Candolle (1844), which presented 3 sections (sect. *Orcheosanthus, Pionophyllum* and *Brandonia*) (Shimai et al., 2021), placing *P. moranensis* in sect. *Orcheosanthus* (Zamudio, 2001; Shimai et al., 2021). Generally, these descriptions have been based on morphological characters, mainly on flower structures and color (Casper, 1996). However, Shimaiem (2021), based on molecular characters such as matK, trnK, rpl32-trnL, ITS1, 5.8S, and ITS2, proposed a taxonomic system with 11 sections and placed *P. moranensi*s, together with *P. ehlersiae* (which in this study it forms a sister group with the clade *P. moranensis-P. jacki*i.) in sect. *Mesoamerican* Shimai and *P. jacki*i in sect. *Caribensis*. Morphologically, the two species show strong similarities, mainly in the lobed calyx, zygomorphism and floral coloration (Zamudio, 2001; Shimai et al., 2021; Rueda-Almazán, 2021; Shimai et al., 2021).

One of the hypotheses to explain the result obtained by phylogenetic analysis is related to the habitats in which these species evolved. *P. moranens*is is distributed from northeastern Mexico to Central America. On the other hand, *P. jackii* is endemic to Cuba (Zamudio, 2001; Grob et al., 2007; Villagas and Alcalá, 2017; Shimai et al., 2021; Temple et al., 2023). They are eco-geographically close regions with little climate variation, which influences the temporal awareness of vegetative and reproductive phenology. It is interesting to note a grouping of American species in a clade sister to *P. alpina*, which is a Eurasian taxon, corroborating a study carried out by Cieslak (2005), and this clade is a sister group to *P. vulgaris*, which despite a high degree of geographic representation has been considered a European species.

The topology of the cladogram obtained in this work leads us to believe that the classification proposed by Shimai et al. (2021), is the most defensible. These results provide clues for studies on phylogeny and biogeography of *Pinguicula* species. Furthermore, despite the consistency of our results, given the alternative taxonomic classification proposed to *Pinguicula*, as well as the new discoveries of species in this genus, combined with the results of this research, it still raises certain questions about the real phylogeny of *Pinguicula*. Therefore, further investigation with a wider range of taxa could help address this gap.

3.3.1- Phylogenetically informative characters (PICs)

It is known that the evolutionary history of plants can be studied through phylogenetic analyses, which seek to understand the kinship relationships between different species of the genus or family. Parsimony informative characters (PICs) are genetic SNPs that can provide information about the evolutionary relationship between different species, giving useful information about phylogenetic inferences and in reconstructing the evolutionary history of organisms (Silva et al., 2018a).

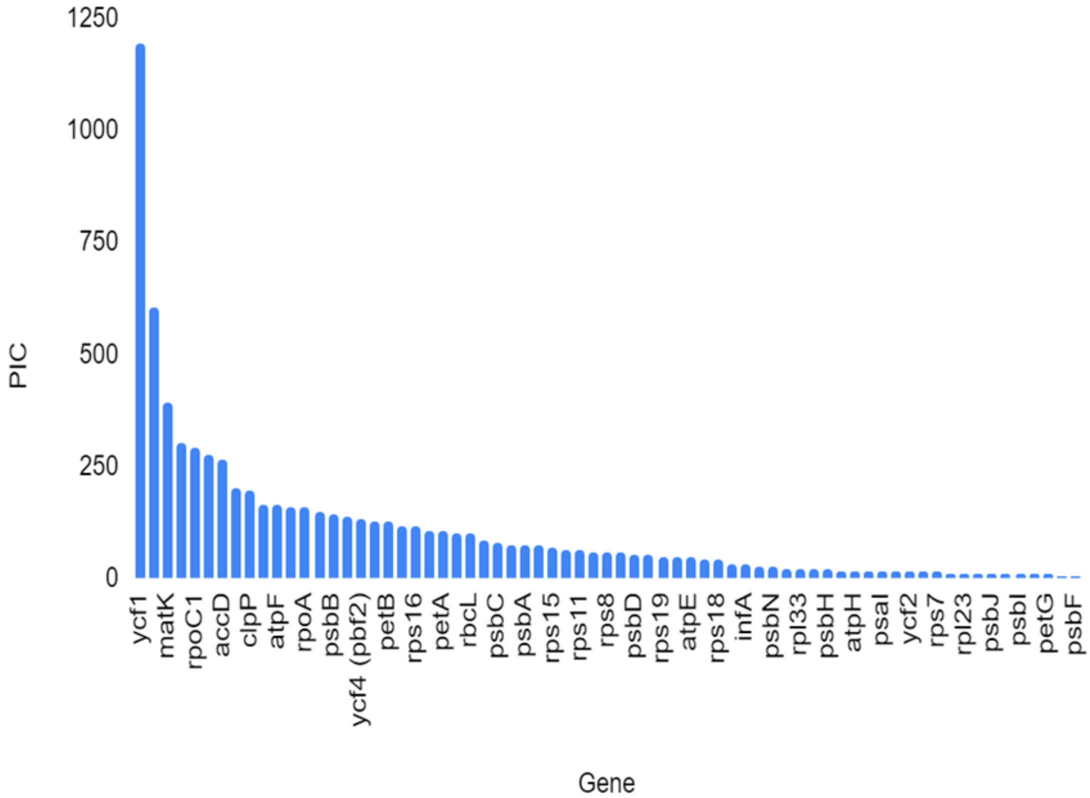


Fig 5. Phylogenetically informative characters (PIC) of chloroplast genes from 15 Lamiales species used in the phylogenetic tree of this research.

Therefore, in this research, we determined the number of PICs, through PAUP (Swofford, 1998), based on 69 chloroplast gene sequences of the species used in the phylogenetic construction, previously concatenated by FASCONCAT-G (Kück and Longo, 2014). The results denote that the phylogenetically informative characters suggest that the main regions with the highest number of PICs are the *ycf*1, *rpo*C2, *mat*K, *rpo*B and *rpo*C1 genes (Fig 3), corroborating Hollingsworth (2011) and Silva et al., (2018a).

4- Conclusions

For the first time, the ptDNA genome of *Pinguicula moranensis* was assembled and evaluated through this study, which contributed to updating the genomic data on Lentibulariaceae. Our results place the ptDNA genome of *P. moranensis* (147,072 bp) as the smallest within the genus, when compared to previously published taxa. The quadripartite structure, genetic material, and genomic composition do not differ much from the other chloroplast genomes in the family published so far. In addition, we compared the newly assembled genome with the sequences of *P. vulgaris* and *P. jackii* in terms of degree of identity and repeats. Regarding identity, the results indicate highly conserved and similar structures between the species, with non-significant divergences only in the spacer regions, RNA transporter genes, inverted region genes, and genes of the NDH complex, which is typical in the genomes of Lentibulariaceae species. Regarding palindromic repetitions in *P. vulgaris*, there is some divergence. The phylogenetic analyses reveal a close relationship between *P. moranensis* and *P. jackii*, corroborating other authors based on geographical hypotheses.

These results awaken new research visions to understand the evolution and biogeography of different taxa of this family and contribute to the taxonomic resolution of *Pinguicula*, which has been undergoing strong updates and debate in scientific literature.

References

ADAMEC, L.; MATUŠÍKOVÁ, I.; PAVLOVIČ, A. Recent ecophysiological, biochemical and evolutional insights into plant carnivory. **Annals of Botany**, v. 128, n. 3, p. 241–259, 2021

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N. & Csaki, F. (Eds). **Second International Symposium on Information Theory**. Budapest: Akademiai Kiado. pp. 267–281, 1974.

ALBERT, V. A.; JOBSON, R. W.; MICHAEL, T. P.; *et al*. The carnivorous bladderwort (*Utricularia*, Lentibulariaceae): a system inflates. **Journal of Experimental Botany**, v. 61, n. 1, p. 5–9, 2009.

ALCALÁ, R. E.; DOMÍNGUEZ, C. A. Diferential selection for carnivory traits along an environmental gradient *Pinguicula moranensis*. Ecology,86(10), 2005, pp. 2652–2660. **Ecology**, v. 86, n. 10, p. 2652–2660, 2005.

ARANGUREN-DÍAZ, Y. C.; VARANI, A. M.; MICHAEL, T. P.; *et al*. Development of microsatellite markers for the carnivorous plant *Genlisea aurea* (Lentibulariaceae) using genomics data of NGS. **Molecular Biology Reports**, v. 45, n. 1, p. 57–61, 2017.

BEIER, S.; THIEL, T.; MÜNCH, T.; *et al*. MISA-web: a web server for microsatellite prediction. **Bioinformatics**, v. 33, n. 16, p. 2583–2585, 2017.

CASPER, S. J.; STIMPER, R. Chromosome numbers in *Pinguicula* (Lentibulariaceae): survey, atlas, and taxonomic conclusions. **Plant Systematics and Evolution**, v. 277, n. 1–2, p. 21–60, 2009.

CASPER, S. J. Monographie der Gattung *Pinguicula* L. Bibliotheca Botanica v. 127-128, n. 1, p. 123–145, 1966.

CIESLAK, T.; POLEPALLI, J. S.; WHITE, A.; *et al*. Phylogenetic analysis of *Pinguicula* (Lentibulariaceae): chloroplast DNA sequences and morphology support several geographically distinct radiations. **American Journal of Botany**, v. 92, n. 10, p. 1723–1736, 2005.

CLIVATI, D.; GITZENDANNER, M. A.; HILSDORF, A. W. S.; *et al*. Microsatellite markers developed for *Utricularia reniformis* (Lentibulariaceae). **American Journal of Botany**, v. 99, n. 9, 2012.

CORRÊA, M. A., MAMEDE, M. C. H. Lentibulariaceae In: Wanderley, M.G.L., Shepherd, G.J., Giulietti, A.M., Melhem, T.S., Bittrich, V., Kameyama, C. (eds.) Flora Fanerogâmica do Estado de São Paulo. Instituto de Botânica, São Paulo, vol. 2. ISBN 85-7523-053-0, 2002.

DANIELL, H.; LIN, C.; YU, M.; *et al*. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. **Genome Biology**, v. 17, n. 1, 2016.

DEGTJAREVA, G. V.; SOKOLOFF, D. D. Inflorescence morphology and flower development in *Pinguicula alpina* and *P. vulgaris* (Lentibulariaceae: Lamiales): monosymmetric flowers are always lateral and occurrence of early sympetaly. **Organisms**

Diversity & Evolution, v. 12, n. 2, p. 99–111, 2012.

DEGTJAREVA, G. V.; CASPER, S. J.; HELLWIG, F. H.; *et al*. Morphology and nrITS Phylogeny of the *Genus Pinguicula* L. (Lentibulariaceae), with Special Attention to Embryo Evolution. **Plant Biology**, v. 8, n. 6, p. 778–790, 2006.

DOMÍNGUEZ, Y.; PANFET, V. C. M.; DE MIRANDA, V. F. O. New features of Cuban endemic *Pinguicula filifolia* (Lentibulariaceae) and considerations on its habitat and ecology. **Flora - Morphology, Distribution, Functional Ecology of Plants**, v. 207, n. 11, p. 838–841, 2012.

DOMÍNGUEZ, Y.; SILVA, S. R.; PANFET VALDÉS, C. M.; *et al*. Inter- and intra-specific diversity of Cuban *Pinguicula* (Lentibulariaceae) based on morphometric analyses and its relation with geographical distribution. **Plant Ecology & amp; Diversity**, v. 7, n. 4, p. 519–531, 2013.

DOMINGUEZ, Y.; TEMPLE, P.; PANČO, I.; *et al*. Biogeographical Patterns of *Pinguicula* L. (Lentibulariaceae) in the Americas Revealed by Endemicity and Potential Distribution Analyses. [s.l.]: **Elsevier BV,** 2023. http://dx.doi.org/10.2139/ssrn.4500281.

DOWNIE, S. R.; PALMER, J. D. Use of Chloroplast DNA Rearrangements in Reconstructing Plant Phylogeny. *In*: **Molecular Systematics of Plants**. Boston, MA: Springer US, 1992, p. 14–35

ELLISON, A. M, ADAMEC, L. 2018. Introduction: what is a carnivorous plant? In: Adamec L, Ellison AM eds. Carnivorous Plants: Phys-iology, Ecology, and Evolution. London: **Oxford University Press**.p.3-6, 2018

FLEISCHMANN, A. On the infrageneric classification of *Pinguicula*. **Carnivorous Plant Newsletter**, v. 50, n. 4, p. 174–188, 2021.

FRAZER, K. A.; PACHTER, L.; POLIAKOV, A.; *et al*. VISTA: computational tools for comparative genomics. **Nucleic Acids Research**, v. 32, n. Web Server, p. W273–W279, 2004.

GROB, V.; PFEIFER, E.; RUTISHAUSER, R. Sympodial Construction of Fibonacci-type Leaf Rosettes in *Pinguicula moranensis* (Lentibulariaceae). **Annals of Botany**, v. 100, n. 4, p. 857–863, 2007.

HOANG, D. T.; CHERNOMOR, O.; VON HAESELER, A.; *et al*. UFBoot2: Improving the Ultrafast Bootstrap Approximation. **Molecular Biology and Evolution**, v. 35, n. 2, p. 518–522, 2017.

HOLLINGSWORTH, P. M. Refining the DNA barcode for land plants. **Proceedings of the National Academy of Sciences**, v. 108, n. 49, p. 19451–19452, 2011.

JIN, J. J.; YU, W. B.; YANG, J. B.; *et al*. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. **Genome Biology**, v. 21, n. 1, 2020.

JOBSON, R. W.; ALBERT, V. A. Molecular Rates Parallel Diversification Contrasts between Carnivorous Plant Sister Lineages1. **Cladistics**, v. 18, n. 2, p. 127–136, 2002.

KALYAANAMOORTHY, S.; MINH, B. Q..; WONG, T. K. F; *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. **Nature Methods**, v. 14, n. 6, p. 587–589, 2017.

KATOH, K.; ROZEWICKI, J.; YAMADA, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. **Briefings in Bioinformatics**, v. 20, n. 4, p. 1160–1166, 2017.

KEARSE, M.; MOIR, R.; WILSON, A.; *et al*. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, v. 28, n. 12, p. 1647–1649, 2012.

KÜCK, P.; LONGO, G. C. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. **Frontiers in Zoology**, v. 11, n. 1, 2014.

KURTZ, S. REPuter: the manifold applications of repeat analysis on a genomic scale. **Nucleic Acids Research**, v. 29, n. 22, p. 4633–4642, 2001.

LAGESEN, K.; HALLIN, P.; RØDLAND, E. A.; *et al*. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. **Nucleic Acids Research**, v. 35, n. 9, p. 3100–3108, 2007.

LEGENDRE, L. The genus *Pinguicula* L. (Lentibulariaceae): an overview. **Acta Botanica Gallica**, v. 147, n. 1, p. 77–95, 2013.

LETUNIC, I.; BORK, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. **Nucleic Acids Research**, v. 49, n. W1, p. W293–W296, 2021.

LI, J.; LIU, Z. L. Complete chloroplast genome sequence of the carnivorous herb *Pinguicula alpina* (Lentibulariaceae). **Mitochondrial DNA Part B**, v. 7, n. 6, p. 1061–1062, 2022.

LOHSE, M.; DRECHSEL, O.; KAHLAU, S.; *et al*. OrganellarGenomeDRAW-a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. **Nucleic Acids Research**, v. 41, n. W1, p. W575–W581, 2013.

LOPES, A. S.; GOMES, P. T.; NASCIMENTO, S. O.; *et al*. Plastid genome evolution in Amazonian açaí palm (*Euterpe oleracea* Mart.) and Atlantic forest açaí palm (Euterpe edulis Mart.). **Plant Molecular Biology**, v. 105, n. 4–5, p. 559–574, 2021.

LUSTOFIN, K..; ŚWIĄTEK, P.; STOLARCZYK, P.; *et al*. Do food trichomes occur in *Pinguicula* (Lentibulariaceae) flowers? **Annals of Botany**, v. 126, n. 6, p. 1039–1048, 2020.

MÜLLER, K.; BORSCH, T.; LEGENDRE, L.; *et al*. Evolution of Carnivory in Lentibulariaceae and the Lamiales. **Plant Biology**, v. 6, n. 4, p. 477–490, 2004.

NGUYEN, L. T.; SCHMIDT, H. A.; VON, H. A.; *et al*. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. **Molecular**

Biology and Evolution, v. 32, n. 1, p. 268–274, 2014.

PALMER, J. D. Comparative organization of chloroplast genomes. **Annual Review of Genetics**, v. 19, n. 1, p. 325–354, 1985.

PARK, J.; XI, H.; KIM, Y.; *et al*. The complete chloroplast genome of *Utricularia tenuicaulis* Miki (Lentibulariaceae) isolated in Korea. **Mitochondrial DNA Part B**, v. 7, n. 6, p. 1143–1145, 2022.

Procko, C; Chory, J; and Pirro, S. The Genome Sequences of 17 Species of Carnivorous Plants. Biodiversity Genomes, November, 2023.

ROBBINS, E. H. J; KELLY, S. The Evolutionary Constraints on Angiosperm Chloroplast Adaptation. **Genome Biology and Evolution**, v. 15, n. 6, 2023.

RODONDI, G.; BERETTA, M.; ANDREIS, C. Pollen morphology of alpine butterworts (*Pinguicula* L., Lentibulariaceae). **Review of Palaeobotany and Palynology**, v. 162, n. 1, p. 1–10, 2010.

RUEDA-ALMAZÁN, J. E.; HERNÁNDEZ, V. M.; ALCALÁ. M. J. R.; *et al*. Spatial and temporal differences in the community structure of endophytic fungi in the carnivorous plant *Pinguicula moranensis* (Lentibulariaceae). **Fungal Ecology**, v. 53, p. 101087, 2021.

RUTISHAUSER, R. Evolution of unusual morphologies in Lentibulariaceae (bladderworts and allies) and Podostemaceae (river-weeds): a pictorial report at the interface of developmental biology and morphological diversification. **Annals of Botany**, v. 117, n. 5, p. 811–832, 2015.

SALTONSTALL, K.; LAMBERTINI, C. The value of repetitive sequences in chloroplast DNA for phylogeographic inference: A comment on Vachon & Freeland 2011. **Molecular Ecology Resources**, v. 12, n. 4, p. 581–585, 2012.

SHIMAI, H.; SETOGUCHI, H.; ROBERTS, D. L.; *et al*. Correction: Biogeographical patterns and speciation of the genus *Pinguicula* (Lentibulariaceae) inferred by phylogenetic analyses. **PLOS ONE**, v. 16, n. 12, p. e0261600, 2021.

SILVA, S.R.; PINHEIRO, D. G.; MEER, E. J.; *et al*. The complete chloroplast genome sequence of the leafy bladderwort, *Utricularia foliosa* L. (Lentibulariaceae). **Conservation Genetics Resources**, v. 9, n. 2, p. 213–216, 2016a.

SILVA, S. R.; DIAZ, Y. C. A.; PENHA, H. A.; *et al*. The Chloroplast Genome of *Utricularia reniformis* Sheds Light on the Evolution of the ndh Gene Complex of Terrestrial Carnivorous Plants from the Lentibulariaceae Family. **PLOS ONE**, v. 11, n. 10, p. e0165176, 2016.

SILVA, S. R.; ALVARENGA, D. O.; ARANGUREN, Y.; *et al*. The mitochondrial genome of the terrestrial carnivorous plant *Utricularia reniformis* (Lentibulariaceae): Structure, comparative analysis and evolutionary landmarks. **PLOS ONE**, v. 12, n. 7, p. e0180484, 2017.

SILVA, S. R.; MICHAEL, T. P.; MEER, E. J.; *et al*. Comparative genomic analysis of Genlisea (corkscrew plants-entibulariaceae) chloroplast genomes reveals an increasing loss of the ndh genes. **PLOS ONE**, v. 13, n. 1, p. e0190321, 2018a.

SILVA, S. R.; GIBSON, R.; ADAMEC, L.; *et al*. Molecular phylogeny of bladderworts: A wide approach of *Utricularia* (Lentibulariaceae) species relationships based on six plastidial and nuclear DNA sequences. **Molecular Phylogenetics and Evolution**, v. 118, p. 244–264, 2018.

SILVA, S. R.; PINHEIRO, D. G.; PENHA, H. A.; *et al*. Intraspecific Variation within the *Utricularia amethystina* Species Morphotypes Based on Chloroplast Genomes. **International Journal of Molecular Sciences**, v. 20, n. 24, p. 6130, 2019.

SILVA, S. R.; PŁACHNO, B. J.; CARVALHO, S. G. M.; *et al*. *Genlisea hawkingii* (Lentibulariaceae), a new species from Serra da Canastra, Minas Gerais, Brazil. **PLOS ONE**, v. 15, n. 1, p. e0226337, 2020.

SILVA, S. R.; MIRANDA, V. F. O.; MICHAEL, T. P.; *et al*. The phylogenomics and evolutionary dynamics of the organellar genomes in carnivorous *Utricularia* and *Genlisea* species (Lentibulariaceae). **Molecular Phylogenetics and Evolution**, v. 181, p. 107711, 2023.

SWOFFORD, D. L. PAUP. **Phylogenetic Analysis Using Parsimon**y (*and other metods). Sinauer Associates, Sunderland, 1999.

TAYLOR, P. The Genus *Utricularia* - A Taxonomic Monograph. Kew Bulletin Additional Series XIV. Royal Botanic Gardens, Kew: London, 1989.

TEMPLE, P.; PANČO, I.; VALDÉS, C. M. P.t; *et al*. *Pinguicula lithophytica* is not conspecific with *Pinguicula jaraguana* or a variety of *Pinguicula jackii.* **Carnivorous Plant Newsletter**, v. 52, n. 2, p. 68–83, 2023.

TILLICH, M.; LEHWARK, P.; PELLIZZER, T.; *et al*. GeSeq – versatile and accurate annotation of organelle genomes. **Nucleic Acids Research**, v. 45, n. W1, p. W6–W11, 2017.

VILLEGAS, S. G.; ALCALÁ, R. E. Reproductive ecology of the carnivorous plant *Pinguicula moranensis* (Lentibulariaceae). **Plant Biology**, v. 20, n. 2, p. 205–212, 2017.

WANG, J.; KAN, S.; LIAO, X.; *et al*. Plant organellar genomes: much done, much more to do. **Trends in Plant Science**, 2024.

WICKE, S.; SCHÄFERHOFF, B.; DEPAMPHILIS, C. W.; *et al*. Disproportional Plastome-Wide Increase of Substitution Rates and Relaxed Purifying Selection in Genes of Carnivorous Lentibulariaceae. **Molecular Biology and Evolution**, v. 31, n. 3, p. 529–545, 2013.

WOOD, D. E.; LU, J.; LANGMEAD, B. Improved metagenomic analysis with Kraken 2. **Genome Biology**, v. 20, n. 1, 2019.

XU, S; TENG, K.; ZHANG, H.i; *et al*. Chloroplast genomes of four Carex species: Long repetitive sequences trigger dramatic changes in chloroplast genome structure. **Frontiers in Plant Science**, v. 14, 2023.

ZAMUDIO, R.S., RZEDOWSKIi, J. Dos especies nuevas de *Pinguicula* (Lentibulariaceae) del estado de Oaxaca, Mexico. **Acta Botanica Mexicana**.;14:23–32, 1991.

ZAMUDIO, R. S. **Revision de la Orcheosanthus del género *Pingüicula* (Lentibulariaceae**). (Tesis de Maestría). Universidad Nacional Autónoma de México, México, 2001.

ZAMUDIO, R. S. Notas sobre la identidad de *Pinguicula moranensis* H. B. K., con la descripción de una variedad nueva. ***Acta Botanica Mexicana*** 49: 23–34, 1999.

CONSIDERAÇÕES FINAIS

Os objetivos inicialmente propostos para dissertação foram alcançados. Nossos resultados revelam: (i) os genes nucleares álcool desidrogenase e nitroreductase são viáveis para estudos filogenéticos de Lentibulariaceae; pois apresentaram uma congruência evolutiva do clado *Genlisea-Utricularia;* (ii) O genoma ptDNA de *Pinguicula moranensis* possui o tamanho de 147.072 pb com um conteúdo total de guanina-citosina (GC) de 38,3%, possuindo 109 genes, dos quais 76 são codificadores de proteínas, 27 tRNA, quatro rRNA e dois pseudogenes. Os plasmas de *Pinguicula* (Lentibulariaceae) utilizados nesse estudo permitiram posicionar *Pinguicula moranensis* como grupo irmão de *Pinguicula jackii*.

Esse estudo permitiu caracterizar os genes álcool desidrogenase e nitroreductase em *Genlisea-Utricularia* bem como o plastoma de *Pinguicula moranensis,* disponibilizando assim, novas informações para a compreensão e comparação dos processos evolutivos e genômicos dos táxons de Lentibulariaceae. Portanto, novos estudos são sugeridos envolvendo um maior número de táxons desta família de plantas, para dar maior sustentabilidade a estes resultados.