UNIVERSIDADE FEDERAL DO PARANÁ

JOSUÉ OLIVEIRA CAMARGO

ANÁLISE FILOGENÉTICA DA REGIÃO CODIFICANTE DO GENOMA PLASTIDIAL: UMA ABORDAGEM LIVRE DE ALINHAMENTO UTILIZANDO O ALGORITMO SVECT

CURITIBA

2018

JOSUÉ OLIVEIRA CAMARGO

ANÁLISE FILOGENÉTICA DA REGIÃO CODIFICANTE DO GENOMA PLASTIDIAL: UMA ABORDAGEM LIVRE DE ALINHAMENTO UTILIZANDO O ALGORITMO SVECT

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Bioinformática, Setor de Educação Profissional e Tecnológica, Universidade Federal do Paraná, área de concentração Bioinformática.

Orientador: Prof° Dr° Fabio de Oliveira Pedrosa

Co-orientador: Prof° Dr° Roberto Tadeu Raittz

CURITIBA

2018

Catalogação na publicação
Sistema de Bibliotecas UFPR
Biblioteca de Educação Profissional e Tecnológica

Camargo, Josué Oliveira
C172 Análise filogenética da região codificante do genoma plastidial: uma abordagem livre de alinhamento utilizando o algoritmo SVect / Josué Oliveira Camargo. - Curitiba, 2018.
103 p.: il., tabs, grafs.

Orientador: Fabio de Oliveira Pedrosa
Co-orientador: Roberto Tadeu Raittz
Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Educação Profissional e Tecnológica, Curso de Pós-Graduação em Bioinformática.

1. Filogenia. 2. Plastídio. 3. Bioinformática. I. Pedrosa, Fabio de Oliveira. II. Raittz, Roberto Tadeu. III. Título. IV. Universidade Federal do Paraná.
CDD 576.88

Elaboração: Angela Pereira de Farias Mengatto - CRB 9/1002

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA

Pós-Graduação em Bioinformática WWW.BIOINFO.UFPR.BR
E-mail: bioinfo@ufpr.br **Tel: 41 33614906**

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em BIOINFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **JOSUÉ OLIVEIRA CAMARGO** intitulada: **"ANÁLISE FILOGENÉTICA DA REGIÃO CODIFICANTE DO GENOMA PLASTIDIAL: UMA ABORDAGEM LIVRE DE ALINHAMENTO UTILIZANDO O ALGORITMO SVECT"**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 27 de março de 2018.

Dr. Fábio de Oliveira Pedrosa
Presidente
Programa de Pós-graduação em Bioinformática – UFPR
Departamento de Bioquímica UFPR

Drª. Leila do Nascimento Vieira
Avaliadora Externa
Pós-doutoranda da Universidade Federal de Santa Catarina - UFSC

Dr. Dieval Guizelini
Avaliador Interno
Programa de Pós-graduação em Bioinformática - UFPR

À minha mãe Eloisa,
À minha esposa Joyce,
Ao meu irmão Jhonatan,
E a todos aqueles que acreditaram em mim.

AGRADECIMENTOS

Agradeço primeiramente ao eterno e bom Deus por mais essa etapa concluída em minha jornada... por ter me concedido vida, força de vontade, paciência, resiliência e persistência. Obrigado por guiar meus passos.

Agradeço à toda a minha família, e em especial minha mãe, Eloísa Helena de Oliveira Camargo, por tudo o que fez por mim. Por não ter deixado o meu sonho morrer, apesar das dificuldades e obstáculos pelo caminho. Obrigado por ter me fornecido meios para aqui chegar. E se aqui cheguei, tenha a certeza que foi principalmente por você.

Agradeço ainda ao meu irmão, Jhonatan Oliveira Camargo, por todo apoio em meus piores momentos. Creio que você me conhece de verdade, pois estava lá quando eu mais precisei. E sim, se aqui cheguei, foi porque você não me deixou desistir também.

Agradeço à minha esposa, Joyce Thaise Ramos de Souza Camargo, simplesmente por tudo. Da compreensão dispendida por todo tempo que deixei de passar contigo para dispender na execução e na escrita deste trabalho, até os puxões de orelha quando começava a me deixar levar pela preguiça. Se hoje me tornei um mestre, o mérito é seu também. Te amo muito!

Agradeço ao Clube de Desbravadores e a Igreja Adventista do Sétimo dia, por ter lançado os meus fundamentos morais, éticos, de respeito à natureza e de reverência a Deus. Se hoje estou em cima, é devido também a vocês.

Agradeço aos meus amigos Marcos, André, Carlos, Juliar e Italo, por me acompanharem na jornada da vida.

Agradeço aos meus orientadores, os professores Fábio de Oliveira Pedrosa e Roberto Tadeu Raittz, por me mostrarem o caminho por trilhar a vida acadêmica. Grandes amigos que levarei para sempre comigo.

Agradeço aos mestres, companheiros de turma Willian, Daniel, Bruno, Alan e Mattheus, e em especial, Dionata, Kelin, e Aryel por todo apoio, discussões, experiências vivenciadas e descontrações a caminho do RU. Vocês marcaram minha vida.

Agradeço a todos os integrantes do grupo de pesquisa de inteligência artificial aplicada à bioinformática, Bruno, Camila, Letícia, Mariane, Amanda e Diogo por me acolherem e me auxiliarem da forma que fizeram.

Agradeço à secretária do Programa de Pós-Graduação em Bioinformática da UFPR, Suzana de Azevedo Gobetti, por sua amizade, companheirismo, profissionalismo e eficiência.

Agradeço ao Programa de Pós-Graduação em Bioinformática da UFPR. Órgãos de fomento, colegiado do curso, colegas, funcionários, professores, enfim, a todos que de algum modo me auxiliaram, mesmo anonimamente, meu muito obrigado.

“É quando nos esquecemos de nós mesmos que
fazemos coisas que jamais serão esquecidas”

Ellen G. White

RESUMO

Com o advento do sequenciamento massivo paralelo de DNA, o campo da bioinformática tornou-se desafiador no que diz respeito a análises genômicas envolvendo *Big Data.* A filogenômica possui duas principais abordagens que utilizam os dados em larga escala para gerar reconstruções filogenéticas: super-árvores e super-matrizes. Para análises filogenéticas em geral, a limitação dos métodos baseados em alinhamento múltiplo para a comparação de sequências se torna evidente devido à complexidade algorítmica para se obter um alinhamento. Para tornar o método viável, poucas sequências são utilizadas para um grande número de organismos ou vice-versa. Podemos perder informação evolutiva ao usarmos um número mínimo de sequencias, ou o prejuízo pode ser ocasionado no tempo computacional. Há a necessidade de um método capaz de gerar filogenias consistentes a partir de grandes volumes de dados em tempo computacional não-proibitivo. A partir do SVect, criamos uma metodologia livre de alinhamento para análises filogenéticas em larga escala utilizando a redução de dimensionalidade. O método se mostrou rápido, eficiente, conciso e capaz de analisar um grande volume de dados, inviável para o alinhamento múltiplo de sequencias. Desenvolvemos a primeira árvore filogenética global de organismos portadores de genoma plastidial utilizando as regiões codificantes dos plastomas, e pretendemos disponibilizar tanto a árvore como a ferramenta de análise para uso científico.

Palavras-chave: Filogenia. Região codificante. SVect. Plastídio.

ABSTRACT

With the advent of large-scale genomic sequencing, the field of bioinformatics has become challenging with respect to genomic and proteomic analyzes involving Big Data. The phylogenomics have two main approaches that use the data in large scale to generate phylogenetic reconstructions: super-trees and super-matrices. For phylogenetic analyzes in general, the limitation of multiple alignment based methods for sequence comparison becomes apparent because of the algorithmic complexity to obtain an alignment. To make the method feasible, few sequences are used for a large number of organisms or vice versa. We can lose evolutionary information by using a minimal number of sequences, or the damage can be caused by computational time. Has the need for a method capable of generating consistent phylogenies from large data volumes in non-prohibitive computational time. From SVect, we created a free alignment methodology for large-scale phylogenetic analyzes using dimensionality reduction. The method was fast, efficient, concise and able to analyze a large volume of data, not feasible for multiple sequence alignment. We created the first global phylogenetic tree of organisms carrying plastids using the plastome's coding regions, and we intend to make available both the tree and the analysis tool for scientific use.

Key-words: Phylogeny. Coding Region. SVect. Plastid.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE QUADROS

LISTA DE SIGLAS

16S	*rRNA 16S ribosomal Ribonucleic Acid* (ácido ribonucleico ribosomal)
Aln	*Formato de arquivo de alinhamento*
DDBJ	*DNA database Japan (*Banco de dados de DNA do Japão*)*
DNA	*Desoxirribonucleic acid* (Ácido desoxirribonucleico)
ENA	*European Nucleotide Archive (Arquivo de nucleotídeos Europeu)*
faa	*Formato de arquivo multi-fasta de aminoácidos*
ftp	*File trasfer protocol (Protocolo de transferência de arquivos)*
GenBank	*Genetic sequence database* (Banco de dados de sequência genética)
kpb	Kilo pares de base
ML	Maximum likelihood (Máxima verossimilhança)
MP	Maximum Parsimony (Máxima parcimônia)
NCBI	National Center for Biotechnology Information (Centro nacional para informação biotecnológica)
NGS	*Next generation sequencing* (Sequênciamento de próxima geração)
NIH	*National Institutes of Health (*Institutos nacionais de saúde*)*
NJ	Neigbour-joinning (Junção de vizinhos)
OTU	*Operational Taxonomic Unit (Unidade taxonômica operacional)*
pb	Pares de base
RefSeq	*Reference Sequence Database*
SO	Sistema operacional
SSR	Simple Sequence Repeats (Sequência simples de repetição)
HTG	*Horizontal Transfer of Gene (*Transferência horizontal de genes)

LISTA DE ABREVIATURAS

Fig**.**	Figura
Tab.	Tabela
pág.	Página

LISTA DE SÍMBOLOS

®	Marca registrada
α	Nível de significância
<	Menor que
=	Igual a

SUMÁRIO

1. INTRODUÇÃO

Com o advento do sequênciamento massivo paralelo de DNA e o consequente aumento exponencial de sequências nos bancos de dados devido ao baixo custo e redução de complexidade técnica, o campo da bioinformática tornou-se desafiador no que diz respeito a análises genômicas envolvendo *Big Data* (GREENE et al., 2014; EMMS; KELLY, 2015; YANG et al., 2017). Para STEPHENS et al., (2015), o campo da genômica será um dos maiores desafios da próxima década.

A filogenômica, campo que surgiu com expansão do volume de dados genômicos e proteômicos disponíveis, no ano 2000 (PHILIPPE et al., 2017) possui duas principais abordagens que utilizam os dados em larga escala para gerar reconstruções filogenéticas: super-árvores e super-matrizes (FLEISCHAUER; BÖCKER, 2017). Enquanto na abordagem de super-árvores são utilizadas sobreposições de árvores filogenéticas, em formato de matrizes que codificam os nós da arvore, para inferir a topologia final da grande arvore filogenética, a abordagem de super-matrizes vale-se de um grande número de genes ou proteínas dos organismos estudados para montar uma matriz que pode ser analisada pelos métodos convencionais de filogenia (FLEISCHAUER; BÖCKER, 2017), tais com o Máxima Verossimilhança – ML (FELSENSTEIN, 1981), Máxima Parsimônia – MP (FITCH, 1971) e Inferência Bayesiana (YANG; RANNALA, 1997). Há ainda uma terceira abordagem, conhecida como Mega-filogenia. Muito similar à abordagem de super-matriz, combina alinhamento de perfis com a seleção manual de regiões gênicas de interesse (SMITH et al., 2009).

Com o passar do tempo, reconstruções filogenéticas utilizando um grande conjunto de dados vem se tornando cada vez mais comum. Como exemplos, a primeira inferência filogenética em larga escala baseada no gene plastidial *rbcL* abrangendo 499 espécies de Viridiplantae (CHASE et al., 1993); o dendrograma que engloba 916 espécies de morcegos, publicado por JONES et al., (2002); o trabalho de MCMAHON & SANDERSON, (2006), que reconstruiu as relações filogenéticas entre 2.228 espécies de leguminosas, a árvore filogenética de 13.533 espécies de plantas verdes baseadas no gene *rbcL* de SMITH et al., (2009); PYRON et al., (2013) ao gerar um dendrograma para Squamata com cerca de 4100 espécies de cobras e lagartos baseado em um conjunto de 12 genes; FITZJOHN et al., (2014), que reconstituiu as relações filogenética de aproximadamente 24 mil espécies de plantas lenhosas; a análise de 37 genomas completos para inferir a filogenia de Oomicetos publicado por MCCARTHY & FITZPATRICK, (2017); e a grande árvore de Viridiplantae de SMITH & BROWN, (2018), contendo 79.881 espécies.

Em seu trabalho, DE PIERRI (2017) propôs um algoritmo livre de alinhamento para comparação de sequências que foi aplicado em uma reconstrução filogenética para 6.811 proteomas mitocondriais, o SVect.

2. OBJETIVOS

2.1. Objetivo Geral

Propor e validar a aplicação da representação vetorial de sequências proteicas e a projeção em um hiperespaço como estratégia para identificar grupos de sequências que compartilham algum grau de semelhança em sua composição de aminoácidos, como uma eficiente alternativa para análises filogenéticas em grandes bases de dados, tais como banco de genomas completos, onde o alinhamento múltiplo de sequência é inviável.

2.2. Objetivos Específicos

Entender o desempenho computacional do algoritmo SVect e encontrar o melhor conjunto de parâmetros para análises filogenéticas baseadas em proteomas de plastídios;

Comparar alinhamento de sequências com o SVect;

Empregar o SVect no conjunto de sequências codificantes traduzidas dos portadores de genoma plastidial, a fim de obter uma filogenia global.

Validar o dendrograma global gerado utilizando o SVect na literatura;

3. JUSTIFICATIVA

Para análises filogenéticas em geral, a limitação dos métodos baseados em alinhamento múltiplo para a comparação de sequências se torna evidente devido à complexidade algorítmica para se obter um alinhamento ideal (VINGA; ALMEIDA, 2003). Para tornar o método viável, poucas sequências de aminoácidos ou nucleotídeos são utilizadas para um grande número de organismos (vide introdução) ou poucos organismos são necessários para se utilizar de genomas completos em análises. Podemos perder informação evolutiva ao usarmos um número mínimo de genes/proteínas/sítios ao inferir relações filogenéticas (DE PIERRI, 2017), ou o prejuízo pode ser sofrido no tempo computacional para obter os resultados da análise. Mediante a tais limitações da abordagem utilizando o alinhamento múltiplo de sequências, é evidenciado a

necessidade de um método capaz de gerar filogenias consistentes a partir de grandes volumes de dados em tempo computacional não-proibitivo.

4. REVISÃO DE LITERATURA

4.1. A Genômica

Por meio de avanços científicos, a genômica tem propiciado cada vez mais investigações a nível molecular (SHENDURE et al., 2004). A partir do momento em que o DNA foi descoberto e descrito na década de 1950, muitas tecnologias foram criadas e a sociedade em geral tem ganhado com esses avanços, que englobam entre outras áreas, a economia e a medicina (WATSON; CRICK, 1953; DINIZ, 2007).

A criação do projeto genoma humano em 1990 foi um símbolo de conquista de conhecimento no contexto da genômica, que passou a ser uma ciência relevante na área entre o meio das ciências biológicas (CATANHO et al., 2007; DINIZ, 2007). O emprego dos genomas completos ou parciais auxiliam no processo de comparação para entender a conformação de outro genoma (SHENDURE et al., 2004).

A Bioinformática aplicada à genômica é uma ciência que utiliza e maneja os dados oriundos do sequenciamento de amostras genéticas de diversas espécies de organismos. Isto posto, os dados resultantes do método de sequenciamento que estão fragmentados e embaralhados são arranjados afim de dar um entendimento mínimo em relação à configuração genômica de um determinado organismo (SIM et al., 2012).

Nesta sessão serão rapidamente revistos os passos para a obtenção de um genoma completo, um dos principais objetivos da genômica.

4.1.1. SEQUÊNCIAMENTO GENÔMICO

A etapa inicial para que possamos realizar a análise do genoma é o sequenciamento. A técnica de sequenciamento de genomas que consistia em fragmentar o DNA em pequenos pedaços conforme o tamanho do vetor usado para a clonagem era conhecida como técnica de Sanger. Essa técnica consiste em alongar primers de RNA sob uma fita de DNA utilizando didesoxinucleosídeos trifosfato além dos 4 nucleotídeos comuns. Assim, quando um didesoxinucleosídeo era adicionado na molécula em construção, interrompia a síntese (Fig 1A). Então, as bases que interrompiam a biossíntese eram identificadas por meio de eletroforese.

Essa técnica tornou-se possível após avanços tecnológicos da metodologia shotgun. (SANGER; COULSON, 1975)

Com as técnicas de sequênciamento de próxima geração (NGS), sequênciamento de segunda geração, ou sequenciamento massivo paralelo, houve uma modernização dos processos a ponto de dispor de uma clonagem *in vitro*, dispensando a clonagem laboratorial. Estes avanços, como a multiplexação e a marcação fluorescente de nucleotídeos (Fig 1B) ao invés dos métodos eletroforéticos resultaram em menores custos, maior velocidade de obtenção das sequências e um ganho considerável na qualidade do sequênciamento, além de chegar a sequenciar bilhões de pares de base em uma corrida (CARVALHO; SILVA, 2010; SHENDURE et al., 2017).

Atualmente, o sequênciamento em tempo real, de terceira geração ou de molécula única (diagramado simploriamente na Fig. 1C) reduziu ainda mais os custos e a facilidade de produção de sequências, visto que as amostras podem ser preparadas em campo. Além disso, um salto na qualidade das sequências foi conquistado, já que o tamanho dos fragmentos aumentou de cerca de 300-400pb nos sequenciadores de segunda geração, para aproximadamente 150kpb nos sequenciadores de terceira geração (BLEIDORN, 2016).

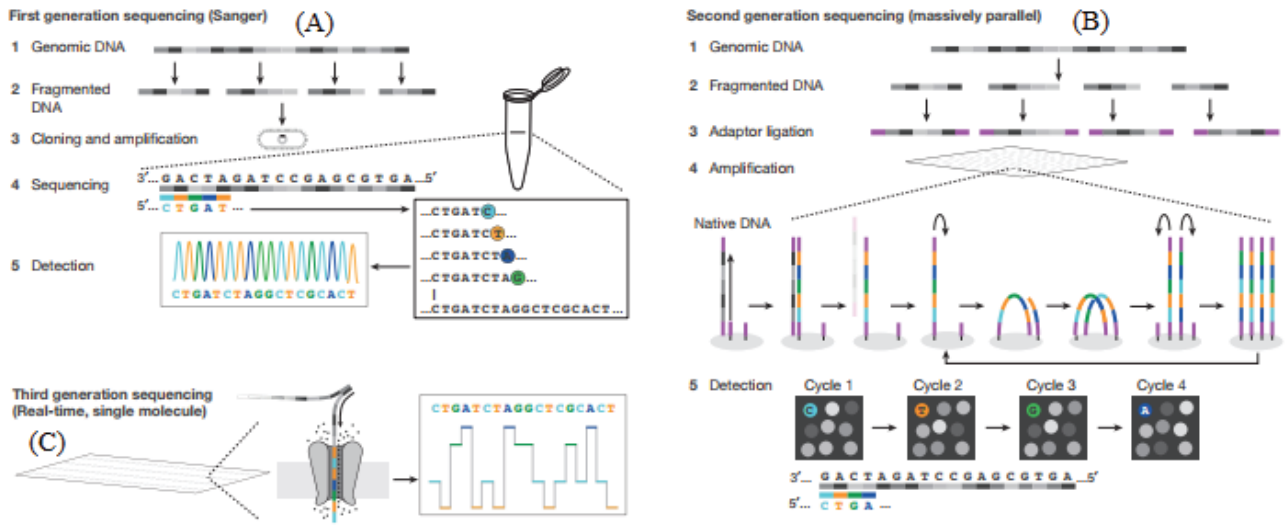


Figura 1: Tecnologias de sequenciamento. O esquema exemplifica o funcionamento dos sequenciadores de primeira, segunda e terceira gerações. Fonte: (BLEIDORN, 2016).

4.1.2. MONTAGEM GENÔMICA

A montagem de genomas é realizada por meio de agrupamento, em que sequências do DNA são compiladas utilizando como critério a sobreposição das *reads*. As *reads* são definidas

como fragmentos de sequência que são originados pela quebra do DNA no sequenciamento. São utilizados programas de Bioinformática para realizar o agrupamento das *reads* por superposição de sequencias idênticas, a fim de reconstruir o fragmento original (SIMPSON; POP, 2015).

Existem hoje muitos programas que realizam a montagem genômica, utilizando diferentes tipos de algoritmos para obter a ordem das sequências do DNA. Mesmo com uma gama de programas, ainda existem problemas que podem interferir negativamente na montagem, como agrupamentos repetidos, regiões com baixa qualidade de sequenciamento, o tamanho dos fragmentos e o próprio montador em si (SIMPSON; POP, 2015).

Os procedimentos que se baseiam no percentual de identidade para o reconhecimento de sequências de outros genomas para realizar a montagem são confiáveis, se o material norteador for de qualidade. Os elementos obtidos através da anotação gerados a fim de serem usados para mineração dos dados em banco de dados públicos, confrontando com genes já existentes, concluindo a montagem de forma adequada. É importante lembrar que a utilização de ferramentas de bioinformática deve ser realizada de forma integrada, para que a verificação de dados importantes seja favorecida (TROY et al., 2001).

4.1.3. ANOTAÇÃO GENÔMICA

Após a realização do sequenciamento, o próximo passo para a análise do genoma consiste na anotação e a busca por similaridade constitui o início dessa etapa. Neste ponto do trabalho, são usadas as sequências homólogas a fim de treinar os métodos estatísticos, para que esses sejam capazes de realizar o reconhecimento de genes. isso resulta em um arquétipo que será empregado para analisar o restante do genoma (TROY et al., 2001).

Para STEIN (2001), há três categorias na anotação de genoma que se destacam por serem mais relevantes: a anotação em nível de nucleotídeos, ou seja, quais os nucleotídeos que compõe a sequência; anotação em nível de proteína, que determina quais proteínas são codificadas pelos genes; e anotação em nível de processo, definindo que função a proteína desempenhará no organismo. Esses processos devem ser realizados de maneira a alcançar resultados de alta qualidade, para que se possa identificar corretamente os genes, os produtos eles codificam, ratificando as características em particular.

Montagens erradas podem levar à identificações erradas de genes e a anotações equivocadas, e se levarmos em consideração o fato de que futuras pesquisas também correm o risco de serem afetadas pelo erro cometido, a dimensão do problema pode fugir ao alcance. Mesmo

utilizando um bom preditor de genes e seguindo um controle de qualidade, ainda assim a anotação está sujeita erros (YANDELL; ENCE, 2012) .

4.2. Os Plastídios

Os plastídios são organelas membranosas que estão presentes em diversos grupos de organismos eucarióticos como as plantas, protozoários e algas. Estas organelas são classificadas conforme a função que exercem no organismo. São elas os proplastídios que são encontrados nos meristemas da planta e dão origem aos outros plastídios; como etioplastos, cloroplastos e leucoplastos. Os etioplastos são plastidios não portadores de clorofila, encontrados em cotilédones de plântulas de angiospermas e que podem ser convertidos em cloroplastos; Os cloroplastos são plastídios portadores de clorofila e realizam a fotossíntese; os cromoplastos são plastídios coloridos, derivados dos cloroplastos e funcionam como depósitos de carotenoides e outros pigmentos, estão presentes em pétalas de flores e frutas; Amiloplastos são derivados do leucoplastos e encontrados em tecidos de armazenamento, como tubérculos e endospermas de sementes; os elaioplastos são também derivados dos leucoplastos e cuja função é armazenar lipídeos; e os leucoplastos são plastidios sem pigmento encontrados nas células da raiz e dão origem aos amiloplastos, elaioplstos e proteinoplastos , (LOPEZ-JUEZ & PYKE, 2005 e EGEA et al., 2010).

Assim como as mitocôndrias, acredita-se que os plastídios foram resultado de um evento endossimbiótico. Um procarionte gram-negativo fotossintetizante foi fagocitado por um eucarioto anaeróbio, e a bactéria não foi digerida passando a trabalhar em conjunto com o hospedeiro. Acredita-se que parte do genoma do procarionte foi perdido ao longo da evolução. Outra parte desses genes foi incorporada ao genoma nuclear da célula eucariótica, tornando a cooperação irreversível e indispensável para a sobrevivência (WEEDEN, 1981; GOULD et al., 2008).

4.2.1. O CLOROPLASTO

Cloroplasto é uma organela de formato variado, revestido por dupla membrana. No estroma, plasma interno análogo à matriz mitocondrial, se localizam todos os orgânulos e enzimas metabólicas (BLANKENSHIP, 2013). Uma das estruturas presentes no estroma é chamado de tilacóide, que geralmente estão empilhados em forma de grana. O tilacóide é revestido por uma terceira membrana. Através desta membrana, um processo semelhante à cadeia transportadora de elétrons mitocondrial gera ATP e NADPH, que são usados como energia para realizar o metabolismo da planta (ANDERSON, 1975).

Além da produção de energia celular no tilacóide, uma enzima presente no estroma chamada *ribulose-1,5-bifosfato-carboxilase*, em plantas processa o gás carbônico atmosférico para produzir açucares na fase escura da fotossíntese, que serão modificados posteriormente no citoplasma, servindo como fonte de energia para outras partes da planta e como substância de reserva (CLELAND et al., 1998).

4.2.1.1 O GENOMA CLOROPLASTIDIAL (CPDNA)

O cloroplasto, de maneira similar à mitocôndria, também possui o seu próprio material genético (BURGESS, 1985). Em um cloroplasto jovem, podemos encontrar até 100 cópias do cpDNA no estroma, porém, de acordo com a idade, o número de cópias pode ser reduzida a 15 (HELDT; HELDT, 2005). É uma dupla fita comumente circular, tendo em média 150.000pb (CLEGG et al., 1994), podendo variar entre 120.000 e 170.000pb (SHAW et al., 2007) e contém *íntrons* (KREBS et al., 2014). Possui uma média de 120-130 genes (ROGALSKI et al., 2015; DANIELL et al., 2016) sendo o menor cpDNA composto de 4 genes (Monoraphidium neglectum) e o maior, de aproximadamente 273 - Pinus koraiensis (NCBI, 2018).

O cpDNA de plantas terrestres possui uma estrutura muito semelhante. Eles codificam 4 rRNAs, 30-31 tRNAs, essenciais para a síntese proteica dos cloroplastos, 21 proteínas que se associam ao ribossomo e 4 subunidades da RNA polimerase, além de 28 proteínas tilacóidais com função conhecida no sistema fotossintético (como a ATP sintase e a NADH desidrogenase) e a ribulose-1,5-bisfosfato-carboxilase, principal enzima do sistema (KREBS et al., 2014)

.

4.2.2 Bancos de Dados para cpDNA

O GenBank, vinculado ao NCBI (BENSON et al., 2017) é um banco de dados biológico de sequências genômicas mantido pelo NIH, e que faz parte de uma cooperação internacional com o DDBJ e o ENA. Essas três bases de dados trocam informações e são mutuamente atualizados todos os dias. O Genbank possui dados de todas as esferas de vida, sendo possível encontrar cpDNA e proteínas plastidiais, porém, os dados são redundantes e em sua maioria não passam por curadoria.

O ChloroplastDB (CUI et al., 2006) é uma base de dados especializada em cpDNA. Atualizado e mantido pelo GenBank, ele aplica técnicas de mineração genômica. A proposta do ChloroplastDB é disponibilizar ferramentas com uma interface amigável que facilite a análise dos dados de sequências cloroplastidiais, entretanto, esse banco está desatualizado, com sua última atualização em 2007, cujos números não ultrapassavam 100 genomas completos.

O GObase (O'BRIEN et al., 2009) é um banco de dados para genomas organelares que engloba plastídios e mitocôndrias e foi criado em 2008. Também se alimenta de dados disponibilizados pelo GenBank e possui diversas ferramentas para manipulação, análise e mineração de dados. Todavia, O GObase também se encontra desatualizado, com sua última atualização datando de 2010.

Outros dois bancos de dados importantes são o ChloroMitoSSRDB 2.0 (SABLOK et al., 2013, 2015) e o ChloroSSRdb (SHANKER, 2014). Com ênfase em microssatélites e SSRs, ambos contemplam genomas completos, no entanto, tais repositórios contêm anotações apenas de SSRs e microssatélites, não fornecendo anotações completas de genes (proteínas e demais produtos).

O *Plastid-encoded Protein Clusters* (ZVERKOV et al., 2012, 2015; LYUBETSKY et al., 2013) é um banco de dados de clusters proteicos de plastídios que aplica mineração em todos os organismos que contém plastídios sequenciados. O banco está sendo constantemente atualizado, entretanto, não disponibiliza os dados para download. Como o *Plastid-encoded Protein Clusters* é apenas para consulta, o pesquisador não consegue expandir suas pesquisas, ficando limitado às ferramentas do banco.

O Organelle Genome Resourses (WOLFSBERG et al., 2001), disponibiliza para descarga os dados do RefSeq (PRUITT et al., 2002, 2007; O'LEARY et al., 2016) – genomas, RNA's codificados e proteínas traduzidas, tanto para plastídios, quanto para as mitocôndrias. Esse banco é bimestralmente atualizado pelo NCBI e por esse motivo escolhemos trabalhar com essa base.

4.3 Filogenética

A filogenética, também conhecida como cladística, é a ciência que tem por objetivo criar classificações biológicas baseadas primordialmente em sequências genéticas, inferindo relações de parentesco e ancestralidade entre os organismos ao longo do tempo. O resultado de uma análise filogenética é um cladrograma, que é uma representação hierárquica das relações evolutivas em forma de uma árvore binária. Dessa forma, um ramo da árvore representa um grupo de organismos que tem um mesmo ancestral comum, sendo denominado clado ou grupo monofilético (PATWARDHAN et al., 2014; NICOLAU, 2017).

Há diversos métodos para se obter uma arvore filogenética, e os principais estão sendo tratados de maneira resumida no subtópico a seguir.

4.3.1 Métodos Filogenéticos

Os métodos filogenéticos são genericamente divididos em 2 grupos: os baseados em matrizes de distâncias e os baseados em características, ou dados discretos (PATWARDHAN et al., 2014). Os métodos baseados em distância tomam como premissa que a distância evolutiva real entre duas espécies pode ser calculada pela diferença entre suas sequências. Esses métodos empregam o número de alterações entre pares em um grupo de sequências para produzir uma árvore filogenética do grupo (FELSENSTEIN, 1996), e são subdivididos em baseados em algoritmos de agrupamento, e baseados em otimização (PATWARDHAN et al., 2014).

4.3.1.1 UPGMA

Um dos métodos baseados em matrizes de distâncias, o UPGMA, do inglês, *Unweighted Pair Group Method with Arithmetic mean* ou método não ponderado de agrupamento por média aritmética, consiste em um algoritmo de agrupamento (*clusterization*)hierárquica de baixo para cima (*bottom-up)*, ou seja, partindo das OTU's visa unir os grupos mais próximos criando um novo agrupamento até chegar ao todo, unindo o conjunto em um único grande grupo (SOKAL, R; MICHENER, 1958). É o algoritmo mais simples, e gera arvores enraizadas ultramétricas que refletem a estrutura da matriz de entrada.

4.3.1.2 Junção de vizinhos *(Neighbor-joining)*

Este método *de* junção de vizinhos, é um algoritmo de agrupamento de baixo para cima *que* considera como a melhor arvore, aquela que tem a menor soma total dos ramos (SAITOU; NEI, 1987). O NJ considera cada organismo como portador de uma taxa de evolução distinta, excluindo a hipótese do relógio molecular (ZUCKERKANDL; PAULING, 1965). Essa é uma vantagem desse método em relação ao UPGMA. O NJ consiste em 5 passos, segundo seus criadores:

1- Com base na matriz de distância atual, calcule a matriz Q.

2- Encontre os dois tipos distintos de i e j (ou seja, com $i \neq j$ para qual Q (i, j) tem o menor valor possível. Esses organismos estão unidos a um nó recém-criado, que é conectado ao nó central. Na Figura 2-B, as OTU's são unidos ao novo nó X.

3- Calcule a distância de cada um dos organismos do par para este novo nó.
4- Calcule a distância de cada um dos organismos fora desse par para o novo nó.
5- Comece o algoritmo novamente, substituindo o par de vizinhos unidos com o novo nó e usando as distâncias calculadas na etapa anterior.

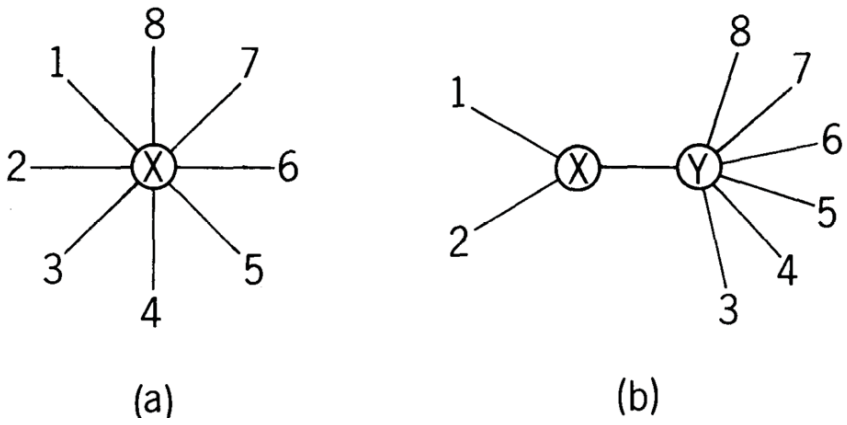


Figura 2: Inicio do processo de construção da arvore filogenética por NJ. (A) Arvore em formato de estrela sem hierarquia. (b) OTU's 1 e 2 unidas pelo nó X após o primeiro ciclo. Fonte: Adaptado de Saitou & Nei (1987).

Já os métodos baseados em características ou dados discretos, também conhecidos como métodos de busca, examinam individualmente cada coluna do alinhamento, procurando a árvore que melhor represente toda a informação. Destacamos aqui o Máxima parcimônia (DAY, 1987), Máxima verossimilhança (FELSENSTEIN, 1981) e a Inferência bayesiana (RANNALA; YANG, 1996). O funcionamento desses algoritmos são temas dos subtópicos a seguir.

4.3.1.3 Máxima Parcimônia

Máxima parcimônia é um método estatístico não-paramétrico que minimiza o número de alterações em uma árvore filogenética conferindo situações de caracteres a nós internos na árvore. O comprimento do caractere (ou do sitio) é o número mínimo de mudanças necessárias para esse sitio, enquanto a pontuação total da árvore é a soma dos comprimentos de caracteres em todos os sítios. A árvore de máxima parcimônia é a árvore que torna mínima a pontuação da árvore, ou seja, evolutivamente falando, a melhor árvore é aquela que exige o menor número de mudanças para explicar os dados observados (DAY, 1987; SWOFFORD, 2002). Uma ferramenta de destaque na utilização dessa abordagem é o PAUP (SWOFFORD, 2002).

4.3.1.4 Máxima Verossimilhança

Em uma análise filogenética por máxima verossimilhança, percebemos duas etapas de otimização: a otimização dos comprimentos dos ramos para calcular a pontuação da arvore e a otimização da busca pela arvore de máxima verossimilhança. Como é um método estatístico, a topologia da arvore torna-se um modelo, enquanto os comprimentos dos ramos e os parâmetros de substituição compõem os parâmetros desse modelo. Esse método consiste em comparar diversos modelos estatísticos com a mesma quantidade de parâmetros para encontrar a

topologia de máxima verossimilhança (FELSENSTEIN, 1981; YANG, 1996). Um software popular que implementa essa metodologia é o RAxML (STAMATAKIS, 2006).

4.3.1.5 Inferência Bayesiana

A inferência bayesiana é uma metodologia geral de inferência estatística. Quando comparada à máxima verossimilhança, os parâmetros no modelo são considerados variáveis aleatórias com distribuições estatísticas, diferindo do ML, que possui parâmetros constantes, fixos e desconhecidos. Antes de analisar os dados brutos, estes são combinados a uma distribuição anterior, para então estimar a distribuição posterior (RANNALA; YANG, 1996). Uma das aplicações mais populares da inferência bayesiana foi a implementação do algoritmo MCMC, *Markov chain Monte Carlo* ou Cadeia de Markov de Monte Carlo (YANG; RANNALA, 1997), implementado no software MrBayes (HUELSENBECK; RONQUIST, 2001).

O Mega 7.0 (KUMAR et al., 2016) é um pacote amplo, que disponibiliza tanto os métodos baseados em distância, quanto os métodos baseados em características.

4.4 INTELIGÊNCIA ARTIFICIAL

4.4.1 Teorema de Johnson-Linnestraus e Redução de Dimensionalidade

Esse teorema afirma que um pequeno conjunto de dados significativos em um espaço de alta dimensão pode ser projetado em um espaço consideravelmente menor, de forma que as distancias proporcionais entre os pontos é praticamente preservada. É possível manter a qualidade de um grande volume de dados, ao multiplicá-lo por um conjunto de vetores ortonormais aleatórios. Isso reduz o volume de dados a ser trabalhado computacionalmente, sem perder as propriedades dos dados originais (JOHNSON; LINDENSTRAUSS, 1984; FRANKL; MAEHARA, 1988)

5 MATERIAL E MÉTODOS

O trabalho foi desenvolvido de acordo com o fluxograma da Figura 3. Cada subtítulo descreve uma etapa de trabalho representado na imagem.

Figura 3: fluxograma do trabalho. O trabalho foi desenvolvido sob duas perspectivas: a construção da grande arvore a partir da aplicação do método proposto e a avaliação da consistência do resultado final, bem como uma simples análise de desempenho ao final do processo. Fonte: O autor.

5.1. Obtenção das Sequências Proteicas

Inicialmente, obtivemos as sequências proteicas de todos os organismos com proteoma cloroplastidial disponível nos arquivos do RefSeq (PRUITT et al., 2002, 2007; O'LEARY et al., 2016) por meio de ftp, acessados pelo Organelle Genome Resourses (WOLFSBERG et al., 2001) no formato *.FAA (Fasta de aminoácidos). A captura dos dados era realizada sempre que uma versão atualizada do banco era lançada, bimestralmente.

5.2 Concatenação das Proteínas

Em posse dos dados, concatenamos então as sequências de proteínas pertencentes ao mesmo organismo, inserindo um separador de cinco asteriscos (*****) entre uma proteína e outra, fato que será explicado no próximo parágrafo. Realizamos ainda a enumeração dos organismos por meio de um identificador sequencial para evitar problemas com dados duplicados, visto que os cabeçalhos das proteínas foram descartados. O resultado dessa etapa é exemplificado na Figura 4 :

```
>Chromera velia org 2222
MLFQRCTPLYLKYQSNTRLSGKKRQPLSVSRNLPKSKPARSYEVLYFIALVLLGCGLLYVLLSPLQKLGL
ARTSKRFFASLNLDAKTAKNVINCLVRIKSIARRGEVTLEDMLVIYKHIDHTLDVLSSPRHLVSLLPYDF
RALEVENIAGVLGVGPSAAKSLAEIFLTPVGNTTSSALVFQSEWSGLLPFGADYLMLCNAERQQVAFWWT
LSALKEIGPARQWDFTPRHTLGIPKKLAELKYKYILEEAAEALGQLNYIADQWQ*****MKTKRFTWMRE
ISRTVFNFVQNTDRRLYIGWFGLLMVPTLLLAVSCFIAAFVAAPPVDIDGIREPVSGSLLYGNNIISGAV
VPSSNAIGMHLYTVWDAMCIEEWLYNGGPYQFVVLHFLIGVASYLGREWELSYRLGMRPWIFVAFSAPVA
AASAVFLVYPIGQGSFSDGMPLGISGTFNFMLVFQAEHNILMHPFHMAGVAGVFGGSLFSAMHGSLVTST
LISMTSEDESPNYGYKFGQEYETYNIVAAHGYFGRLIFQFASFNNSRSLHFFLAIWPVVGIWLTAMGIST
MAFNLNGFNFNQSIIDAEGRVISSWADILNRANLGMEVMHERNAHNFPLDLA*****MEGIVAACSAFSA
GIAIGLGSIGPGIGQGILAGDAVSAISRQLEGEDKIRTTLLPSLAVLEAVTIYGLLIALVIGRVALLKVD
DSSKNPDQNSHEEKAIEPCYVVGAACSAFSYASVIDIIDVLCPDLNRTKNKEPKRDLVWYLQLIWYLIRY
LYFSAYWKYSNVNREW*****MLLSTFSRYTRRPLHKHAPPFRILNTTASILETGIKVIDLVTPYRIGGK
IGLFGGAGVGKTVVIMELITNIAKLHKGISVFAGVGERSREGFDLYKEMRESGVIDTVFPPDSKVSLVYG
QMNETPAARMRVALGAIAVAEFFRDEFFQNVLFFIDNIFRYVQAGAEISALLGRIPSAVGYQPTLATEMG
SLQERIASVSRLSLESQMPIISSITSIQAIYVPADDLTDPAPATNFAHLDATTVLSRNLAQKGIYPAIDP
LASSSLCLKPEVCGPFHYSITQKIIGVLQSARNIQDLLSIIGFEELSEEQKITIYRSRKIQNFLSQPFFV
AEVFNNIPGKLVNLSETVIDAQLILNGYADAIEDISLSYVGRLGEAIQKDYKNFYEKFFIVDKELFLTKW
LKEKAQIVDETFYRNDPQTKAPRNFLVEDLDQKVLNELDGSLPPLSRIWWLAERPGLVTNHLSVEDLLEE
LALPPESRTLPLSL*****MIITTIIIAPKSVPLSSVGKNWESWKSDNGIRFRLERKVVEMLLITFSRYT
RRRLNKHAPNILFSRRTPISLQQARQQSGMKNLAQQERVAVSLPQKPATRTGSSEILGLVALFMVVSVFI
AILYVRFSAGPKVGFAKNSKVVGYVDLPEWWHRKRKFLGEIIGRFVAAMANWARDCWKMYRRYIKYGWYP
PPSPRLDEILYQFGVELWEFCRQVYQDYLRSKTEGRPPV*****MDIEDPVQTFGEYEFDLDEVVYRLQH
LLFLSVQILCLSFIVSPIVSDVVFETWYFKEVSFMGQSVNTIFFFNFFVSFLLSVVILSSYIVFQIFLWL
DTAFLEDERLLYRFLNGFWLFFFLLALGAAIFLIFPLYWITVLQGIKFLNIQNFQSIVDLSQFTDLLAFL
VFSTWVWSLTPILQIVILYLNLITVENLISNWKYVIIFSMLGCGILTASADPLPQLLLALVINLFFIFSL
ILYKSLFHNMCVLFDEKSLTTEDDDDLFLMFLWE*****MFANFPIYAQQGNLIAREANGKLVCANCHLN
NYPIKIYGVRHLINNEVFNLKLSVDRPNSALTQSASLNGSKKFLGLGGIVILPEQFDVVYRPAAENPFIP
YSTEEDQTLVVGPLYTNKATLNIPVRSPAAKDGIPTTNMFFGGINRGRGQLTPLGVKTDTNKPELEFSNL
SVGQIWRSSKGHLYMTQANSENLYRLNGNLHLTWGDNSAVIVNSPNQGGFGQSEFSTSVLNLEYLVNYLF
FTLFCVSAQISLIIYKKDYLVRRSFDTWSYWCRFALPKRKIVVSEETVMVDPNFIVLTGKKYKR
```

Figura 4: Resultado da etapa de concatenação das proteínas. Um arquivo multifasta é gerado com a seguinte conFiguração: Cada cabeçalho contém o nome da espécie a qual pertencem as sequências proteicas (Chromera velia) e um identificador sequêncial (org 2222), enquanto as sequências se encontram concatenadas por cinco asteriscos (*****). Nenhuma ordenação foi usada na concatenação, sendo que a primeira proteína do concatenado é a primeira proteína do arquivo pertencente ao organismo em questão, e assim sucessivamente. Fonte: O autor.

5.3 Vetorização e Redução de Dimensionalidade

Em seguida, utilizamos o algoritmo SVect (DE PIERRI, 2017) para criar uma representação vetorial de cada proteoma. Esse algoritmo usa uma janela de tamanho cinco (k=5), com a terceira posição vazia, dividindo os aminoácidos selecionados em dois pares de dois resíduos. A janela percorre a sequência, conforme a Figura 5 (A) e uma matriz é gerada, de acordo com a ocorrência desses pares relacionados (Figura 5 B).

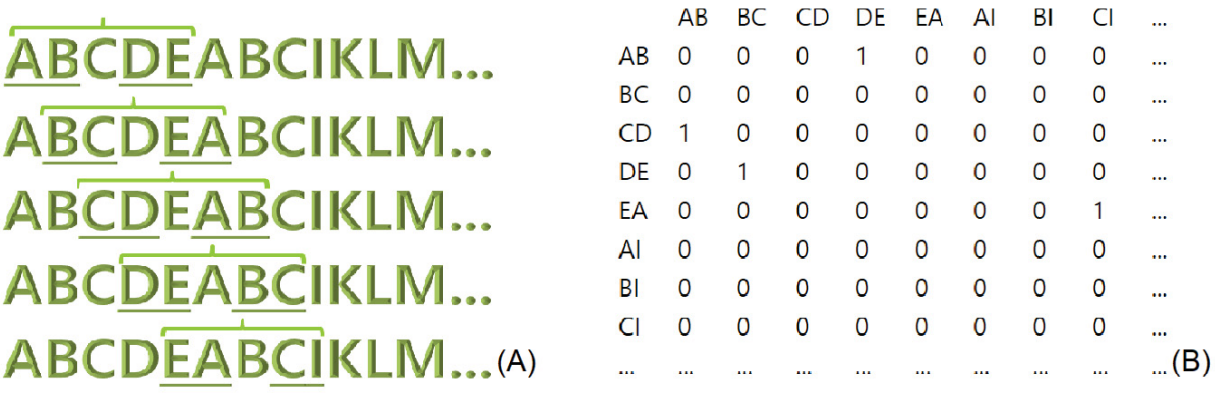


Figura 5: Esquema de funcionamento do algoritmo SVect (ir para a pág. 102). Uma janela móvel (k-mer) de tamanho cinco e espaçada na terceira posição percorre toda a sequência (A), marcando todas as combinações de tetra-aminoácidos em uma matriz binária quadrada de tamanho quatrocentos. Como resultados, temos uma matriz M_{mxn} = $M_{400x400}$ (B), tal que *m* e *n* são todas as possíveis combinações para pares de aminoácidos tradicionais. Os espaçadores entre as sequências de resíduos de aminoácidos inseridos na etapa anterior foram usados como recurso para evitar a criação de dados não-biológicos na matriz resultante. Fonte: O Autor.

Concatenamos então as colunas da matriz para gerar um vetor de 160 mil posições binárias. Essa quantidade de memória pode ser um problema quando se trata de um grande volume de dados. Implementamos então um algoritmo simples que aplica o teorema de Johnson-Linnestraus para reduzir o tamanho dos dados em memória. Após alguns testes, para fins de divulgação, usamos o vetor binário sem redução de dimensão e a projeção em um hiper-espaço de tamanho 600 para cada proteoma. Os testes que foram norteadores para a tomada de decisão encontram-se na sessão de Resultados e Discussão.

5.4 Construção

Com os proteomas vetorizados e projetados, extraímos a matriz de distâncias internas entre os pontos, usando o cálculo de distância euclidiana. Uma implementação do método de junção de vizinhos foi usada para a obtenção da arvore filogenética, dada a matriz de distancias como entrada para o algoritmo. Todos os passos realizados até então se desenvolveram em ambiente *MatLab*®, utilizando os parâmetros padrões.

5.5 Amostragem

Já em posse da filogenia, utilizamos scripts em Python3® para eleger 26 genes presentes em 21 organismos no conjunto de dados brutos não processados oriundos do RefSeq. Essas proteínas estão listadas no quad. 1. Ordenamos e concatenamos as proteínas selecionadas por espécie, utilizando os mesmos espaçadores da primeira etapa desse processo (*****). Todas as sequências foram salvas em um multi-fasta de aminoácidos.

5.6 Alinhamento Múltiplo do Grupo Controle

A etapa anterior nos forneceu dados bem controlados, afim de comparar a eficiência do método livre de alinhamento com o alinhamento de sequências tradicional. O alinhamento foi realizado em SO Linux, utilizando o algoritmo *Clustal Omega*® usando 20 *threads*, e o resultado foi direcionado para um arquivo *.aln.

5.7 Obtenção da Árvore Controle

Nosso arquivo de alinhamento foi então carregado no programa *SeaView4*® para que a matriz de distâncias fosse gerada a partir do cálculo de *Poisson*. A árvore filogenética também foi gerada no *SeaView4*®, usando o algoritmo junção de vizinhos (*neighbor-joining).* Foram rodados 1500 ciclos de reamostragem aleatória (*bootstrap)* para avaliar a consistência da árvore, e o consenso foi usado para as comparações.

O procedimento vetorial do item 3.2.3, também foi usado nesse conjunto amostral para obter um dendrograma. E de forma comparativa, pudemos avaliar a confiabilidade do método utilizado no trabalho.

5.8 Desempenho

Realizamos também testes de desempenho para demostrar a confiabilidade da estrutura das árvores filogenéticas obtidas pelo algoritmo SVect. Inicialmente, aferimos o tempo de processamento para a geração dos dendrogramas construídos a partir do Svect, baseados nas matrizes binária e projeções em um hiper-espaço de 1200, 1000, 800, 600, e 400 atributos. Pare esse teste, foram usados os parâmetros padrões do MATLAB® para a função NJ e demais.

Então, aferimos o tempo de processamento para, a partir da sequência, construir uma arvore filogenética, incluindo todos os pré-processamentos necessários. Dessa forma, comparamos as metodologias baseadas em alinhamento e livre de alinhamento no requisito velocidade de processamento. Dois computadores foram usados para este fim, e a descrição de hardware se encontra na Tabela 1.

Tabela 1: Configuração de hardware dos computadores utilizados nos testes

Servidor	**Notebook**
Processador Intel® Xeon® E5-2640 v4 40-Core @ 2,4GHz	Processador Intel® Core I7-4510U ® 2.6GHz
256 GB RAM	8 GB RAM

2 TB HDD	1 TB HDD
Linux 4.4.0-72 #93 Ubuntu SMP	Windows® 10 Pro

Fonte: O Autor.

Determinamos o tempo de processamento computacional a medida em que o número de de organismos com regiões codificantes inseridas na análise aumenta, visando entender a dinâmica de processamento do algoritmo. Geramos árvores filogenéticas com o mesmo conjunto de parâmetros para dados de proteomas liberados pelo RefSeq no período de julho de 2016 até dezembro de 2017. Por fim, buscamos entender a relação entre a matriz binária e as projeções em um hiper-espaço, para fixar um ponto de equilíbrio entre desempenho e qualidade da árvore. Geramos 30 projeções em um hiper-espaço entre 100 e 3000 atributos, incrementando de 100 em 100 para os quatro principais algoritmos de cálculos de distância que são aplicáveis a esse caso: distância euclidiana, distância de cossenos, distância de correlação e distância de Spearman. Utilizamos então a correlação de Pearson entre a projeção em um hiper-espaço peculiar em tamanho e tipo de distância usada (por exemplo, tamanho 1000 gerado a partir de distância dos cossenos) e a matriz binária original com as 160.000 coordenadas. Arbitramos um valor de $\alpha = 0,01$ para fins de corte.

6 RESULTADOS E DISCUSSÃO

Esta sessão será dividida em filogenia global, filogenia comparativa e desempenho de acordo com as etapas do trabalho. Discutiremos em cada subtópico os aspectos relevantes de cada fase do trabalho. No item 3.1, provamos que se pode gerar grandes filogenias sem exigir muito recurso computacional utilizando o algoritmo SVect. Em 3.2, mostramos e discutimos a comparação feita entre alinhamento e vetorização de sequências. Por fim, no item 3.3, apresentaremos a grande árvore filogenética dos plastídios.

6.1. Desempenho

Para demonstrar a vantagem da nossa ferramenta no que diz respeito a tempo de execução, fizemos testes em duas maquinas distintas, cujas configurações encontram-se listadas na tabela I.

Conforme observaremos na próxima secção, os resultados das metodologias contrastadas foram muito similares entre si. Fato que nos levou repetir o teste, aferindo o tempo total de processamento, tanto para o algoritmo SVect, quanto para a abordagem utilizando o

alinhamento múltiplo de sequências. A contagem de tempo computacional contemplou desde o tratamento da sequência bruta até o instante em que o dendrograma é montado, utilizando um computador pessoal.

Tipo de análise	Ordenação (S)*	Espaçadores (s)**	Processamento (s)	Montagem (s)	Total (s)
SVect	**-**	0,2	**0,2**	**0,4**	**0,8**
Alinhamento	643,59	**-**	18,2	9,2	670,99

Tabela 2: Tempo de processamento das abordagens de obtenção de árvores filogenéticas comparadas, rodando em um computador pessoal.

Fonte: O Autor. (*) Na fase de ordenação, as sequencias foram concatenadas ordenadamente em todos os organismos, de forma a simular a sintenia genética e favorecer o alinhamento. O SVect dispensa essa ordenação, pois a sitnenia não altera os seus resultados. (**) Tempo de pré-processamento para a inserção dos 5 asteriscos como espaçadores. No alinhamento, os espaçadores não são necessários.

Ao considerarmos o tempo de pré-processamento, os números da análise contidos na Tab. 2 nos mostraram que o SVect é cerca de 838 vezes mais rápido quando usamos os dados não projetados. Visto que a ordem de disposição das sequências é essencial para o alinhamento múltiplo de qualidade, tornou-se necessário um subprograma (*script*) em Phyton3 para ordenar e concatenar todas as sequências no conjunto dos organismos usados como controle. Esta etapa é custosa computacionalmente falando, e como os separadores Inter sequenciais utilizados no SVect permitem o algoritmo gerar os pontos da matriz independentemente em cada proteína, podemos obter a mesma matriz independente da ordem em que as sequencias estão dispostas ao longo do proteoma concatenado. Dessa forma, a vetorização de sequencias dispensa a etapa de ordenação, o que justifica a diferença gritante de tempo computacional de processamento.

E ainda que desconsiderássemos a etapa de ordenação dos dados pelos algoritmos em questão, o SVect se mantém cerca de 45 vezes mais ágil em relação ao alinhamento múltiplo quando executado nas configurações acima mencionadas.

Fizemos também testes para descobrir a curva do tempo de processamento em relação ao número de sequências analisadas sem projeção em um hiper-espaço. O resultado está representado na Figura 11, e nos mostra que o crescimento é relativamente linear, mas houve um aumento anormal no tempo de processamento com aproximadamente 2200 sequências, quando houve estouro de memória RAM e o computador passou a utilizar memória do disco rígido para concluir a análise.

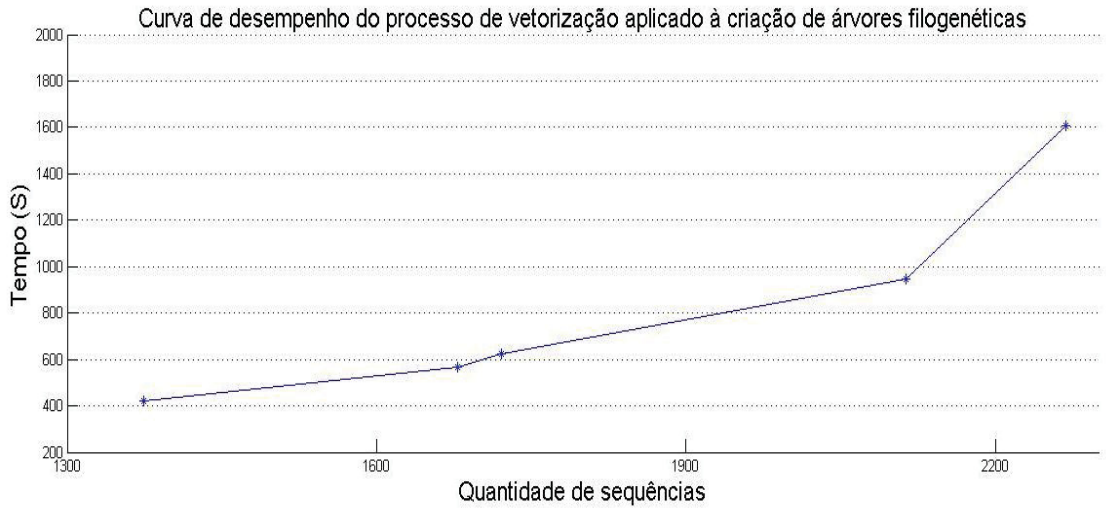


Figura 6: Gráfico de relação tempo/quantidade de sequências para obter a árvore filogenética em notebook usando como base o vetor binário não projetado. Fonte: O autor.

Apesar de nos fornecer os melhores resultados, a matriz binária contém 160 mil posições numéricas para cada proteoma/OTU processado. Esta matriz pode tomar muita memória RAM quando muitos organismos forem processados. Logo, a projeção em um hiper-espaço utilizando o teorema Johnsons-Lindenstrauss(JOHNSON; LINDENSTRAUSS, 1984) reduz o tamanho dos dados em memória, aumentando também a velocidade de processamento das informações.

Utilizando 4 algoritmos de cálculo de distância distintos: Spearman, Correlação (Pearson), Cossenos e Euclidiana, projetamos os dados processados pelo SVect em 30 vetores de tamanhos entre 100 e 3000, incrementando de 100 em 100, para inferir a melhor abordagem a ser trabalhada para as projeções em um hiper-espaço usando dados proteômicos de plastídios. Na Figura 12, o gráfico mostra a correlação de Pearson entre as projeções em um hiper-espaço e a matriz binária.

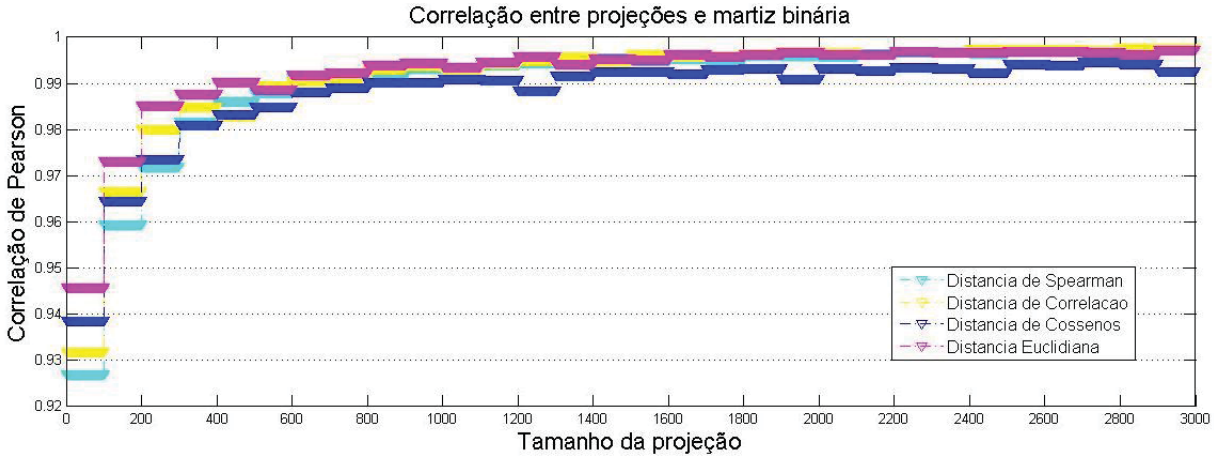


Figura 7: Gráfico comparativo dos algoritmos para cálculo de distância em função da correlação entre o tamanho da projeção em um hiper-espaço e a matriz binária. Fonte: O Autor.

A partir da Figura 12 podemos inferir que o algoritmo de distância euclidiana é superior na projeção em um hiper-espaço dos dados trabalhados, para a maior parte dos vetores projetados. Isso justifica a escolha desse algoritmo para todas as projeções em um hiper-espaço usadas nesse trabalho. Outro detalhe observado é o fato de que todos os vetores projetados a partir do tamanho 600 possuem $\alpha < 0,01$. Dessa forma, para sermos parcimoniosos entre a qualidade e o desempenho, usamos e recomendamos a projeção em um hiper-espaço de tamanho 600 para trabalhar com dados de proteomas plastidiais.

Tabela 3: Tempo de geração de dendrogramas utilizando diferentes projeções em um hiper-espaço em dois computadores distintos.

Projeção em um hiper-espaço	Dispositivo	Tempo(S)
Ausente	PC	706,3
	Servidor	393,8
1200	PC	236,6
	Servidor	92,5
1000	PC	217,7
	Servidor	81,8
800	PC	162,6
	Servidor	68,1
600	PC	141,5
	Servidor	59,8
400	PC	125,1
	Servidor	52,8

Fonte: O Autor.

Também contamos o tempo de processamento para construir uma árvore filogenética de 2113 proteomas de plastídios usando o algoritmo SVect a partir do momento em que o arquivo em formato *.faa é lido. Aferimos o tempo computacional nos dois dispositivos já mencionados para a matriz binária e para as projeções em um hiper-espaço com tamanho 1200, 1000, 800, 600 e 400. O resultado foi a obtenção dos cladogramas em tempos que variam entre 52 segundos a pouco mais de 11 minutos (Tab. 3), provando que são viáveis análises por vetorização em computadores pessoais, mesmo com um grande volume de dados.

As 3 tentativas de análise global dos proteomas por alinhamento não obtiveram sucesso pela ferramenta Mega7 (KUMAR et al., 2016), que permaneceu aproximadamente 37 horas

processando e exibiu uma mensagem de erro. O alinhamento de todos os proteomas foi realizado no Omega Clustal (SIEVERS; HIGGINS, 2014) após aproximadamente 45 horas de processamento, mas não foi possível gerar a filogenia, devido a diferença de tamanho dos proteomas alinhados e a má-qualidade do alinhamento por este programa.

6.2 Análise Filogenética Comparativa

A etapa amostral do trabalho desenvolvida em Python3®, nos forneceu um conjunto controlado e ordenado de proteínas, tornando possível o alinhamento. Consequentemente, uma árvore baseada na distância de sequências foi obtida. Nesta análise, 15 proteínas comuns e com função conhecida no metabolismo (Quadro 1) foram selecionadas para a comparação, totalizando 21 organismos usados como controle.

Proteínas utilizadas na análise	
photosystem II protein I	*ribosomal protein L20*
photosystem I subunit VIII	*ribosomal protein S12*
photosystem II protein J	*ribosomal protein S18*
photosystem II protein N	*cytochrome b6*
photosystem II protein D1	*cytochrome f*
photosystem II protein K	*maturase K*
photosystem II protein T	*ATP synthase CF1 alpha subunit*
photosystem I assembly protein	

Quadro I: Proteinas usadas na análise. Foram utilizadas sequências proteicas essenciais encontradas no maior número de organismos distintos. Residuos de aminoácidos funcionais na maquinaria de DNA e na fotossintese compõem a lista. Fonte: O autor.

Ao compararmos o dendrograma gerado por alinhamento e distância entre sequências (Figura 8-A) com o dendrograma irmão construído por nossa metodologia livre de alinhamento (Figura 8-B), verificamos a alteração de dois ramos entre as árvores (*Ludisia discolor e Megaleranthis saniculifolia*). Todavia, a topologia não foi alterada, visto que a espécies estão enraizadas na mesma posição em ambos os dendrogramas. Trata-se apenas de diferenças gráficas entre as representações.

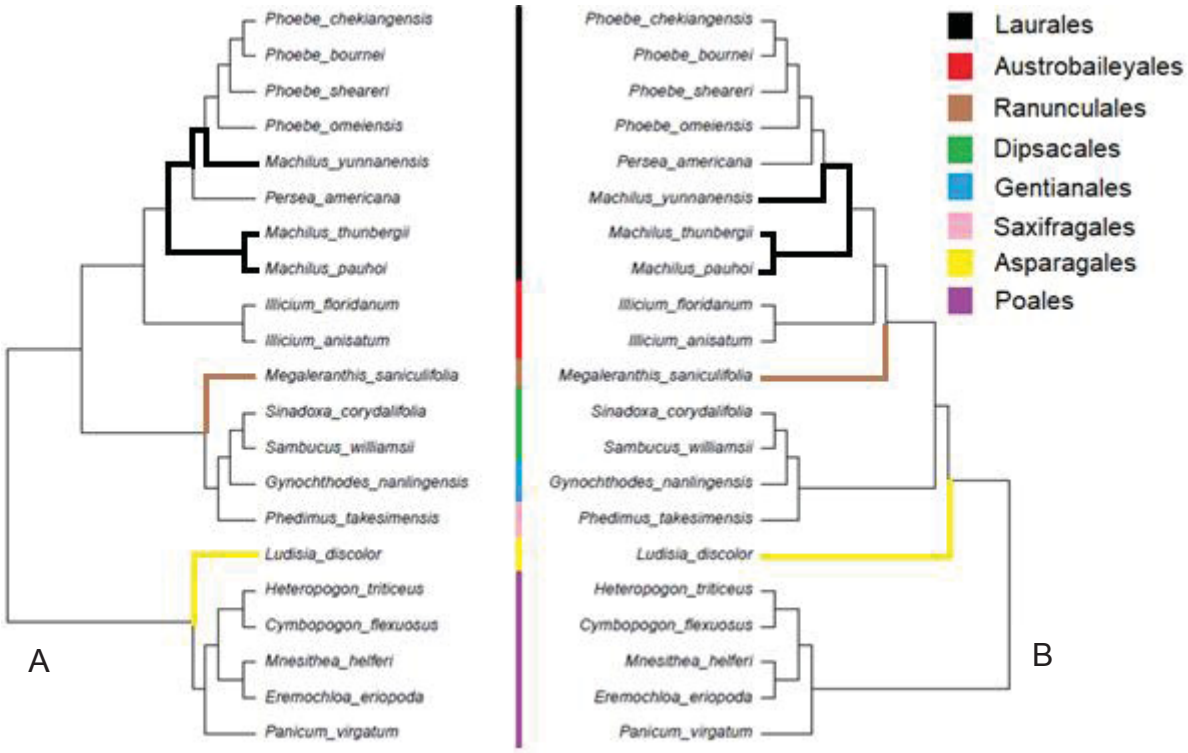


Figura 8: Comparação entre metodologia baseada em alinhamento e livre de alinhamento usando a matriz binária para obtenção de árvores filogenéticas. (A) Topologia construída com base em alinhamento. (B) Topologia gerada com base em vetorização de sequências. Fonte: O autor.

Notamos também que o dendrograma baseado em alinhamento (Fig 8A) dividiu o cluster pertencente ao gênero *Machilus*, inserindo o Abacate (*Persea americana*) entre os integrantes, tornando-o um grupo parafilético. Isso não ocorre no dendrograma baseado em vetorização de sequências (Fig 8B). O gênero *Machilus* coalesce em um único grupo parafilético.

Quando comparamos o dendrograma de baseado em alinhamento, com a árvore filogenética baseada na projeção em um hiper-espaço vetorial de tamanho 800 (Fig. 9), vemos o clado Laurales constituído de mesma topologia. A topologia da árvore muda em *Megaleranthis saniculifolia* (Ranunculales), que é colocado em um grupo monofilético com *Phedimus takesimensis* (Saxifragales). Como ambas as ordens são basais em relação a Dipsacales e Gentianales, não é possível inferir uma topologia como correta. Para tal, seria necessário a inclusão de mais espécies na análise.

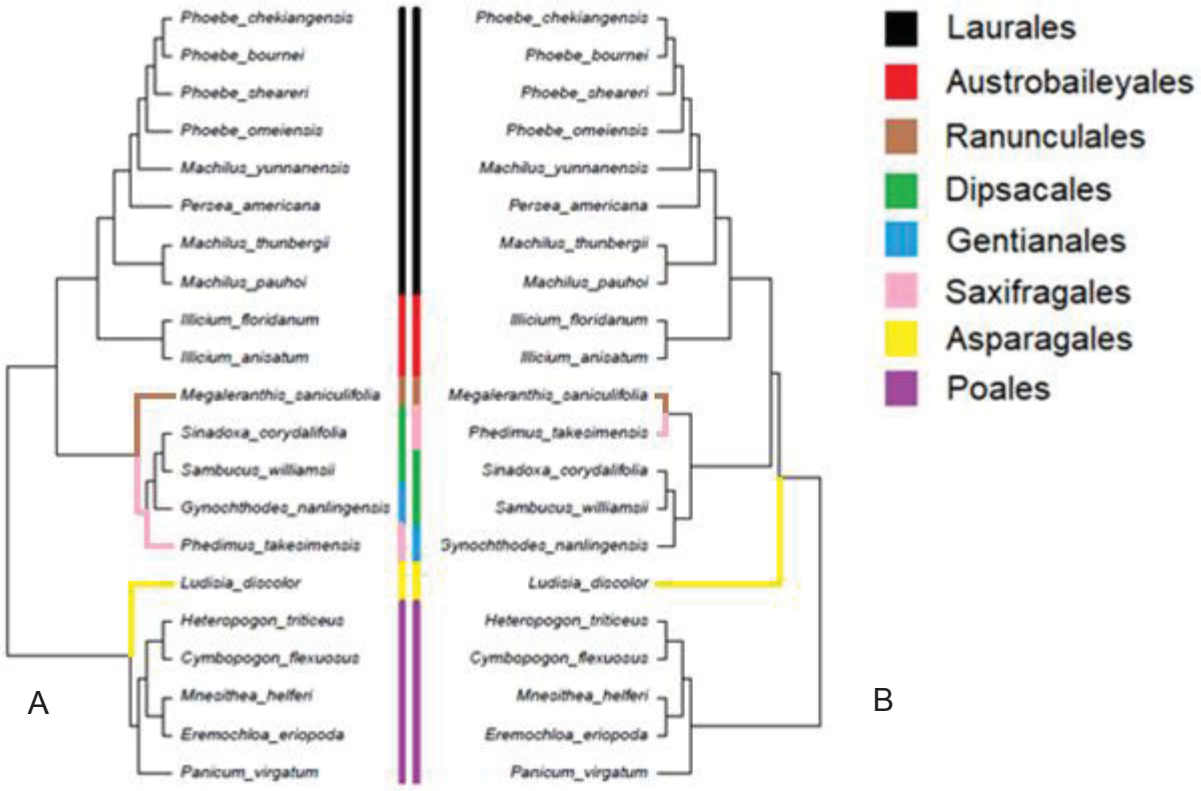


Figura 9: Comparação entre metodologia baseada em alinhamento e livre de alinhamento usando vetor de tamanho 800 para obtenção de árvores filogenéticas. (A) Topologia obtida com base no alinhamento. (B) Topologia construída por vetorização e projeção em um hiper-espaço. Fonte: O autor.

O resultado da comparação entre o dendrograma controle obtido por alinhamento e a árvore filogenética gerada pela projeção em um hiper-espaço do SVect de tamanho 600 (Fig. 10B) nos dá uma topologia idêntica à do dendrograma gerado pela matriz binária (Fig. 8B). A igualdade entre esses cladogramas é mais uma prova que fundamenta a nossa proposta de trabalho com projeções em um hiper-espaço de tamanho 600 no SVect.

Não podemos afirmar aqui que as topologias aqui apresentadas são corretas. Apenas mostramos que nossa abordagem nos fornece resultados muito similares aos do alinhamento múltiplo de sequências, e, em alguns casos, a qualidade da topologia proveniente da vetorização, quando validada na literatura, se mostrou mais consistente, visto que agrupou os plastídios pertencentes ao mesmo gênero (*Machilus*), o que não podemos observar quando utilizamos a metodologia baseada em alinhamento.

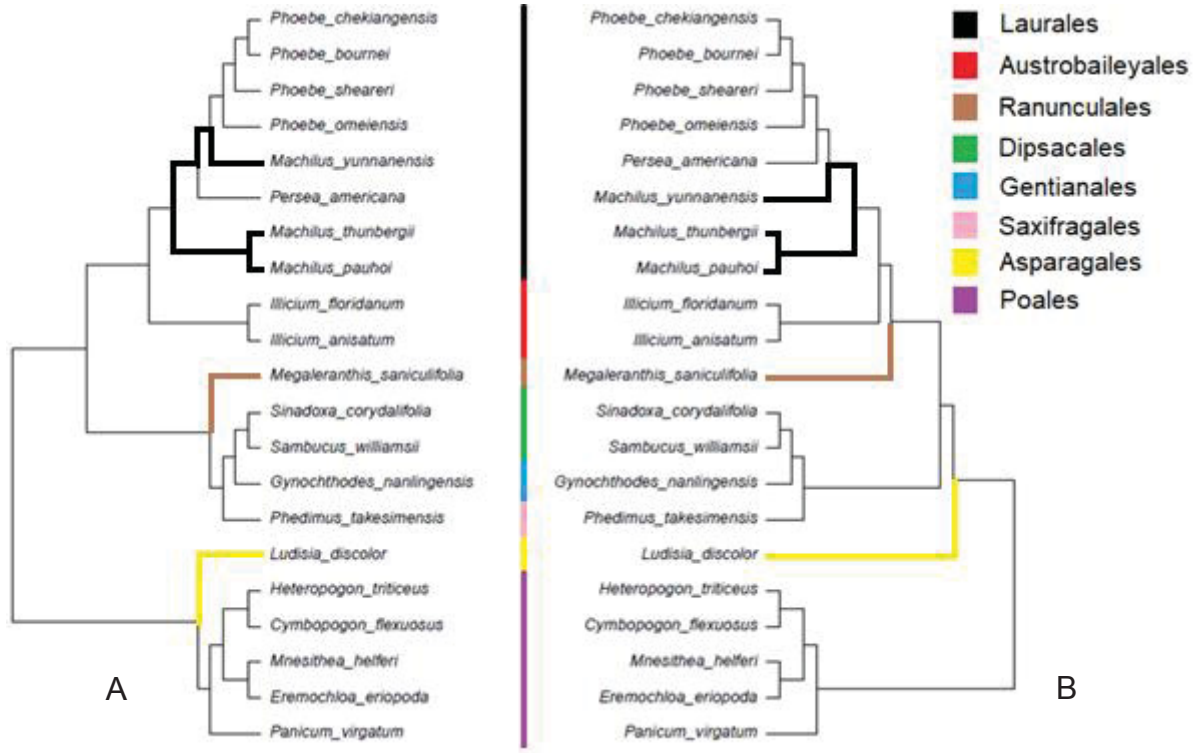


Figura 10: Comparação entre metodologia baseada em alinhamento e livre de alinhamento usando vetor de tamanho 600 para obtenção de árvores filogenéticas (A) Topologia obtida com base no alinhamento. (B) Topologia construída por vetorização e projeção em um hiper-espaço. Fonte: O autor.

6.3 Análise Filogenética Global

Baseada no conjunto de dados baixados no RefSeq/Organelle Genome Resourse versão 86, disponível no dia 12 de janeiro de 2018, nossa árvore filogenética possui 2493 OTU's, sendo 1962 de espécies portadoras de cloroplastos e 531 portadoras de plastídios indiferenciados e apicoplastos. A análise engloba todos grupos de organismos com plastídios sequênciados e depositados no banco de dados, tais como algas, protozoários, plantas terrestres vasculares e avasculares. Discutiremos partes da árvore em relação à sua coerência, validando os trechos conforme outros trabalhos já publicados. Por razões obvias, não discutiremos a validação de toda a árvore, todavia, podemos encontrar a filogenia em sua totalidade, de acordo com a topologia sua no Apêndice 1.

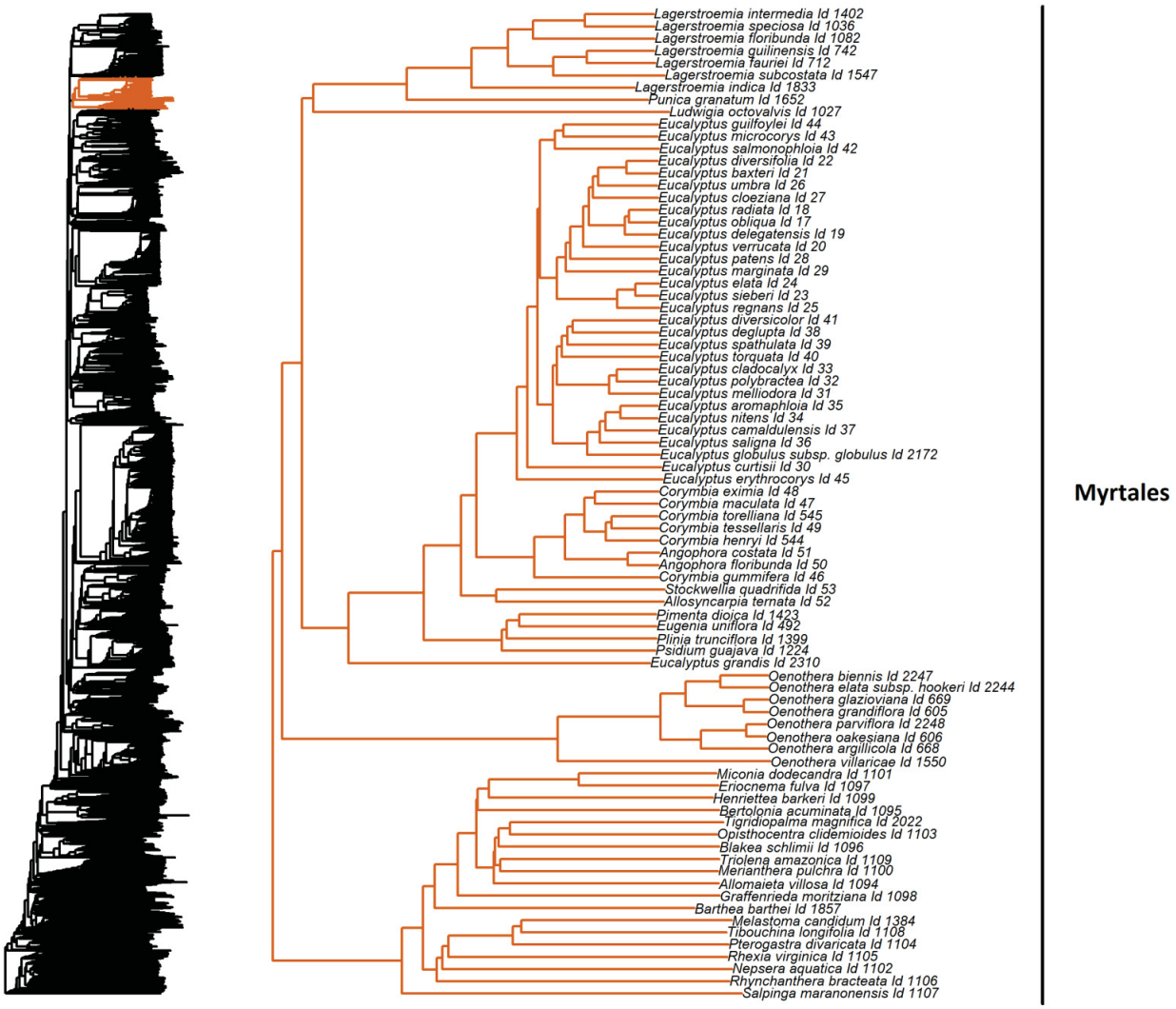


Figura 11: Posição da ordem Myrtales na árvore filogenética vetorial construída pelo método livre de alinhamento SVect. Fonte: O Autor.

A ordem Myrtales é composta por 9 famílias, 380 gêneros, e 13005 espécies, variando entre arbustivas e arbóreas. As folhas são na maioria dos casos opostas e inteiras (sem ramificações, simples), os plastídios encontrados em ductos de condução e armazenam apenas amido, possuem condutores de floema internos e suas flores são geralmente hermafroditas (STEVENS, 2012; BAYLY et al., 2013).

Na Figura 11, os gêneros *Corymbia* e *Angophora* são agrupados juntos. A mesma topologia é confirmada na Figura 12, extraída do estudo de BAYLY et al., (2013), que usa o genoma plastidial completo para inferir a filogenia. Também o gênero *Eucalyptus* em nossa análise encontra-se em similar posição em relação à BAYLY et al., (2013), sendo que os clusters de *E. regnans/E. elata/E. sieberi* e *E.radiata/E. obliqua/E. delegatensis,* estão similares em ambos os estudos (Fig. 12). Podemos então inferir que há coerência nos resultados da aplicação do SVect para análise filogenética quando tratamos desta ordem.

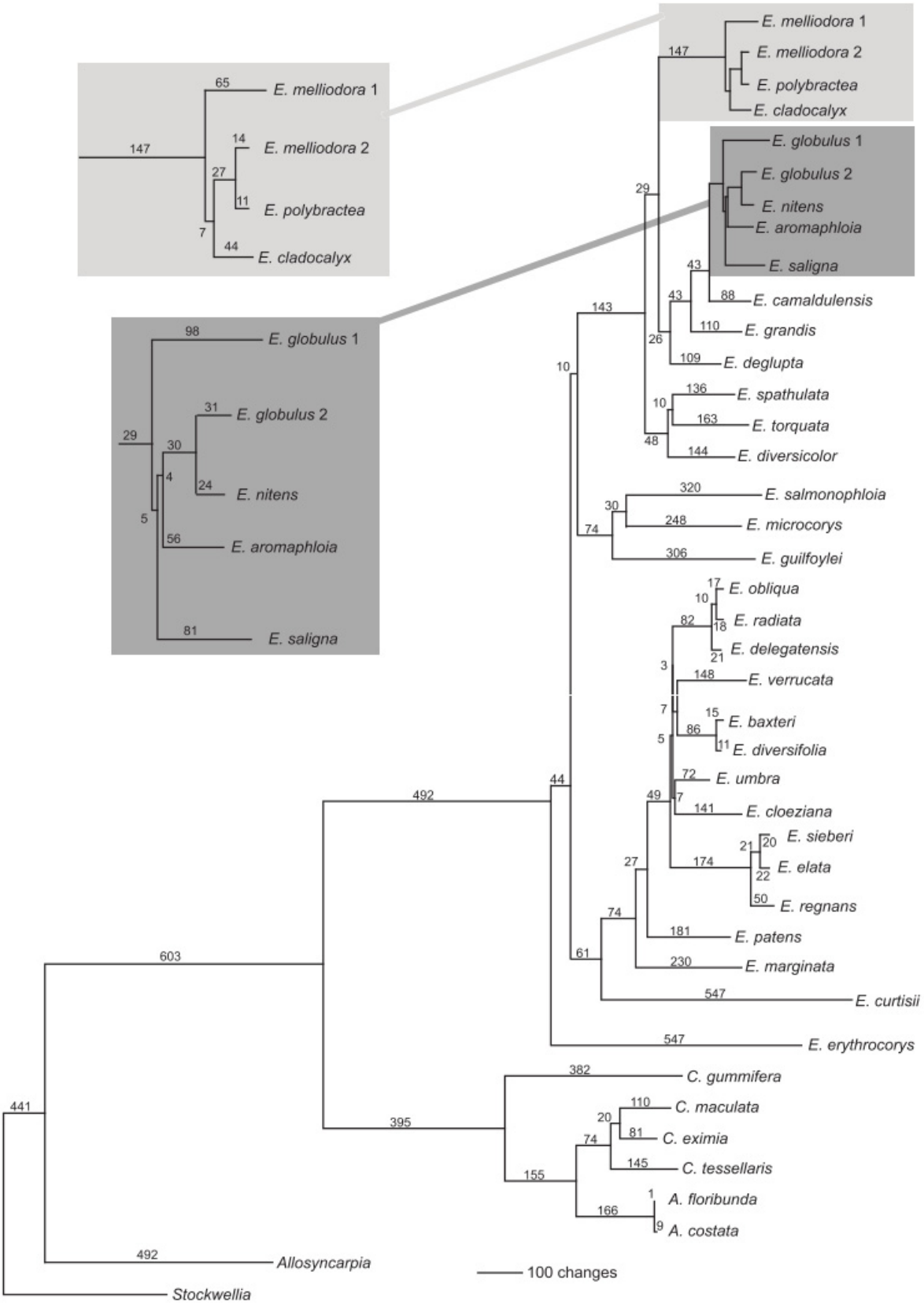


Figura 12: Árvore mais parcimoniosa com base na análise de genomas cloroplastidiais completos de 41 organismos da ordem Myrtales. Fonte: BAYLY et al., (2013)**.**

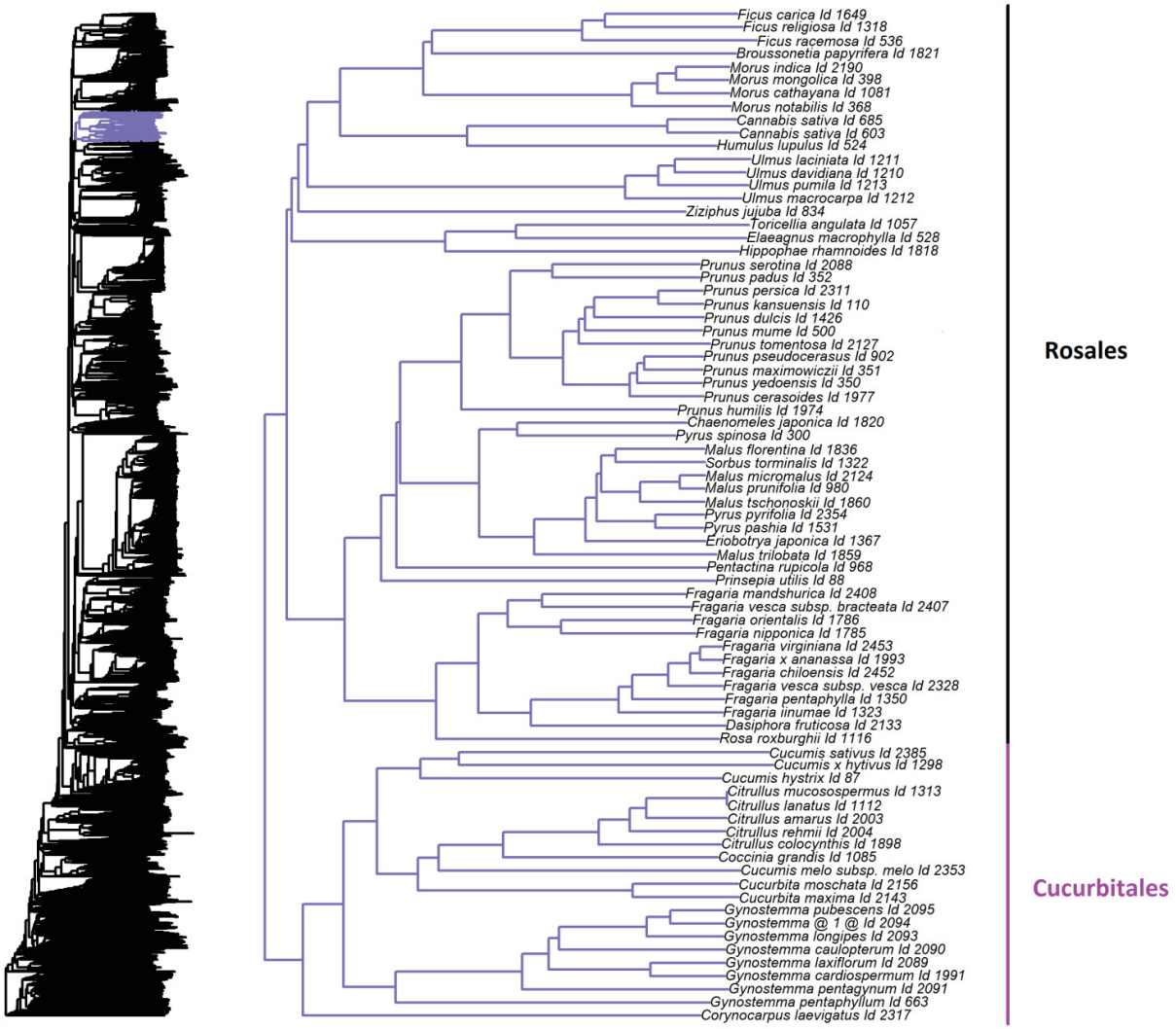


Figura 13: Posição das ordens Rosales e Cucurbitales na árvore filogenética vetorial construída pelo método livre de alinhamento Svect. Trecho da filogenia destacando as ordens em questão. Fonte: O autor.

A ordem das Rosales é composta por 9 famílias, 261 gêneros e 7725 espécies até o presente momento. Suas características sinapomórficas são a ausência de endosperma na semente e a presença de hipanto circundando o ovário (STEVENS, 2012). É o grupo irmão de Cucurbitales e Fabales, que junto com Fagales, formam o clado de fixação de nitrogênio.

A Figura 13 nos mostra o trecho da filogenia global correspondente à ordem Rosales. Ao compararmos a topologia da Figura 13 com o estudo realizado por LI et al., (2016) vemos as espécies arranjadas de forma similar, seguindo os clados de *Fragaria sp., Prumus sp* e *Morus sp.*, conforme Figura 14. A topologia também permanece consistente ao comparar com o dendrograma de ZHAO et al., (2017), conforme a Figura 15.

É notório que no trabalho de LI et. al., (2016) o grupo irmão de Rosales é colocado como Cucurbitales (*Corynocarpus* e *Cucumis*), corroborando com nosso estudo.

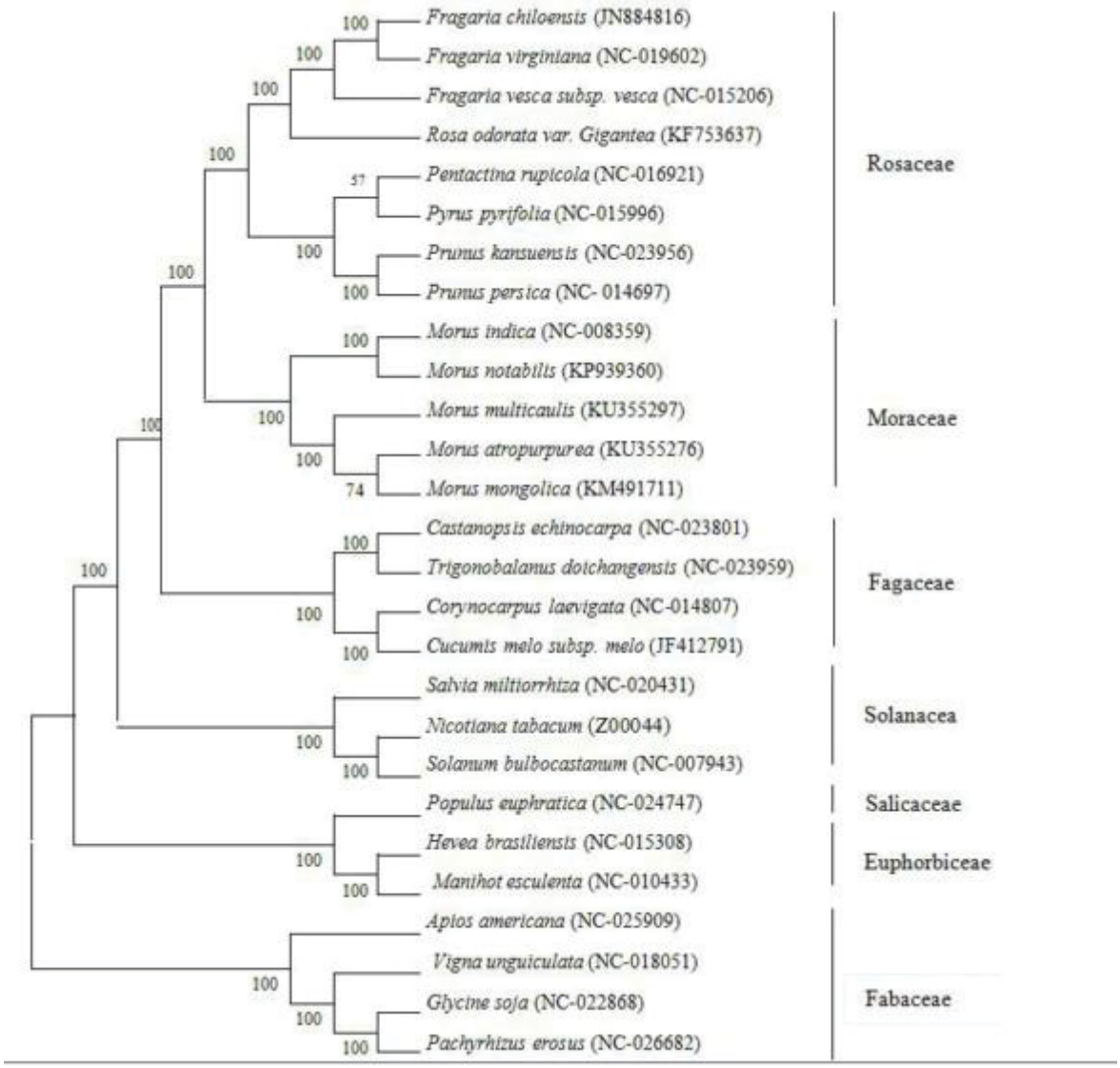


Figura 14: Análise filogenética de espécies de Amora baseado em genoma plastidial completo pelo método NJ. Fagaceae, Solanacea, Salicaceae, Euphorbiceae e fabaceae são incluídos como grupos externos. Fonte:(LI et al., 2016)**.**

Já a ordem Cucurbitales é formada por 7 famílias, 109 gêneros e 2935 espécies e suas principais características são flores unissexuadas, na maioria das vezes pentâmeras e polinizadas por insetos (MATTHEWS; ENDRESS, 2004; STEVENS, 2012).

Em nossa árvore filogenética, a colocação do clado Cucurbitales como grupo irmão das Rosales já foi fundamentado pelo estudo de ZHAO et al., (2017), e a pesquisa de RENNER & SCHAEFER, (2016) fundamenta o clado em nossa árvore colocando *Cucumis, Cucurbita, Citrullus e Coccinia* distante de *Gynostemma*, considerando-o gênero basal, conforme o dendrograma do estudo mostrado na Figura 16.

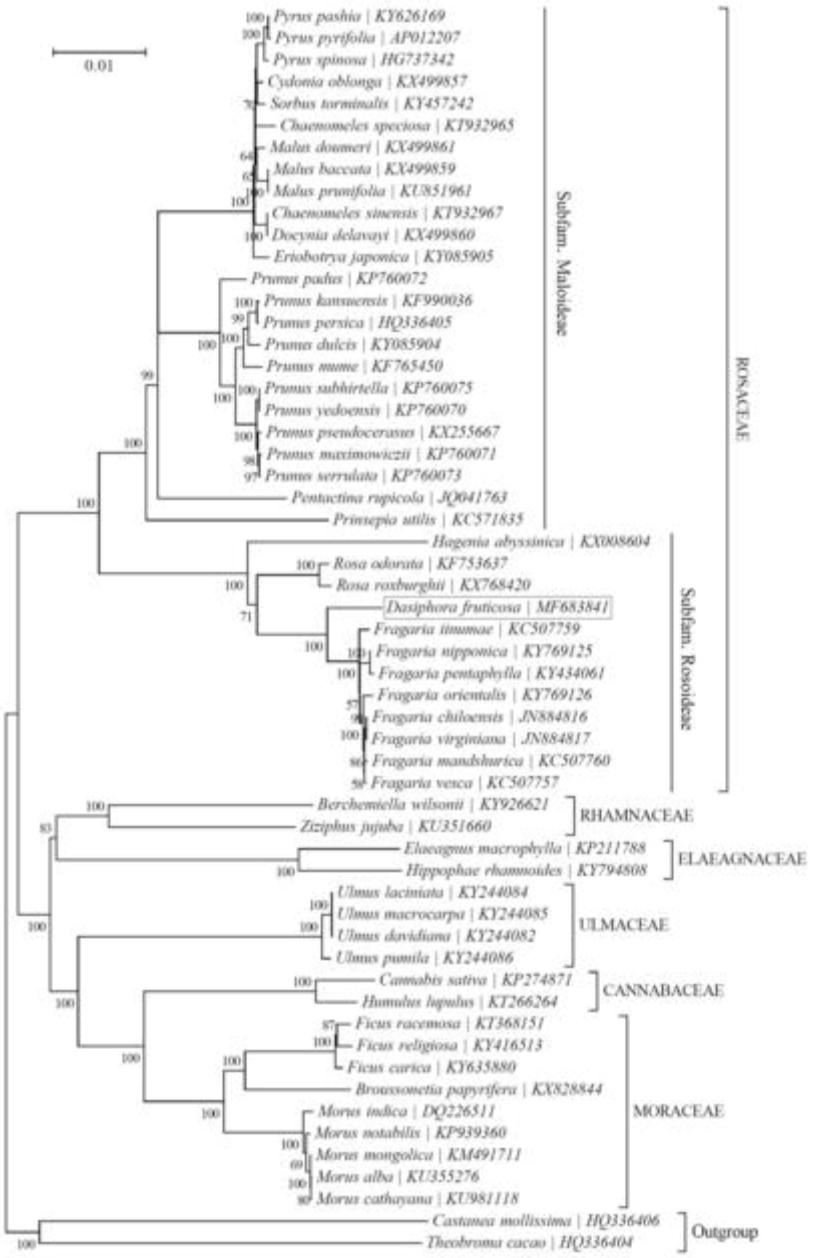


Figura 15: Filogenia de 55 espécies de Rosales gerado a partir de sequências codificadoras de proteína concatenadas, baseado no algoritmo NJ. os valores de bootstrap são baseados em 1000 replicações. Castanea mollissima e Theobroma cacao foram incluídos como grupo externo. Fonte: (ZHAO et al., 2017)**.**

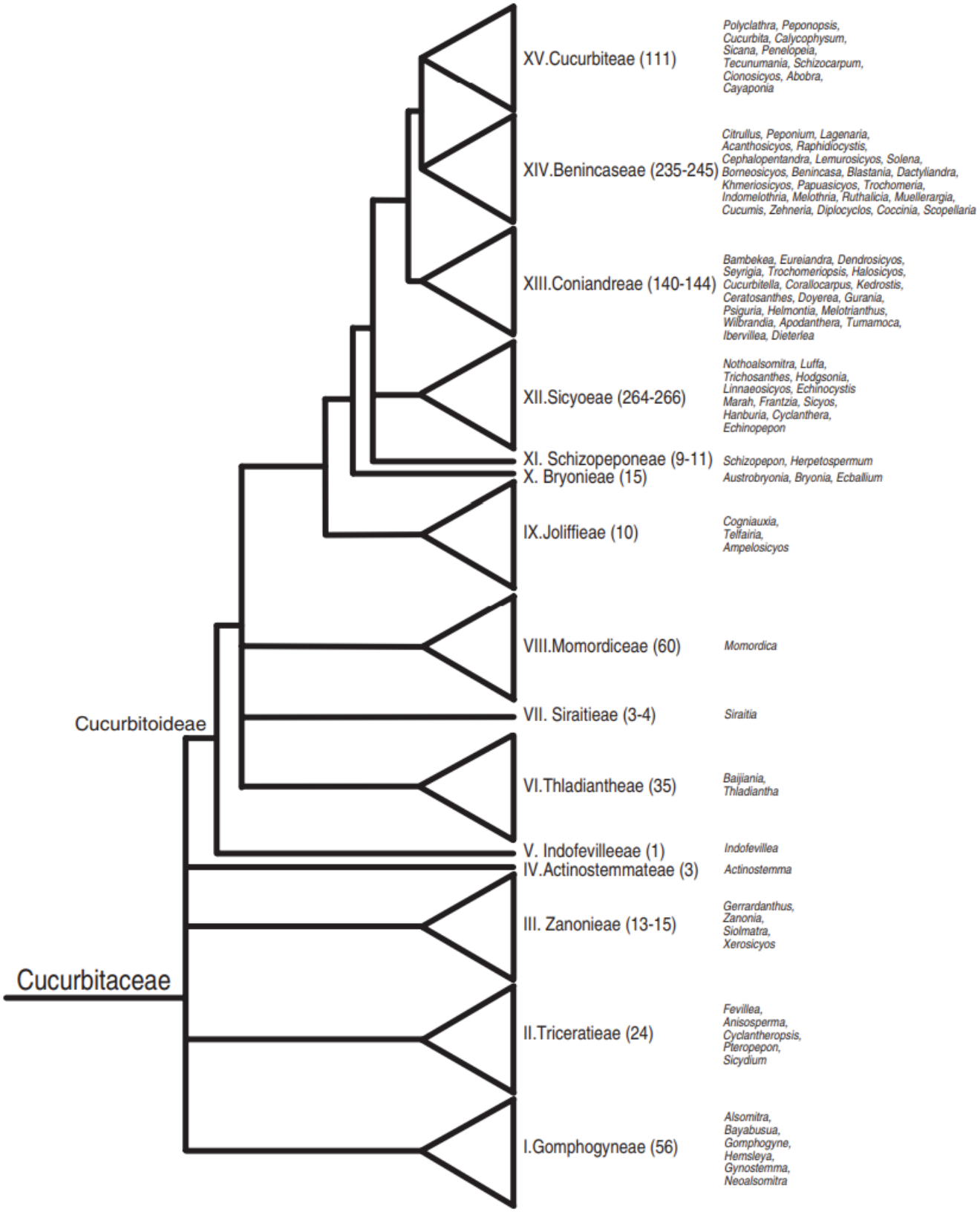


Figura 16: Cladograma baseado em 14 regiões de DNA mitocondrial, nuclear e de cloroplasto, mostrando as relações filogenéticas entre as tribos de Cucrbitaceae. Fonte: (RENNER; SCHAEFER, 2016).

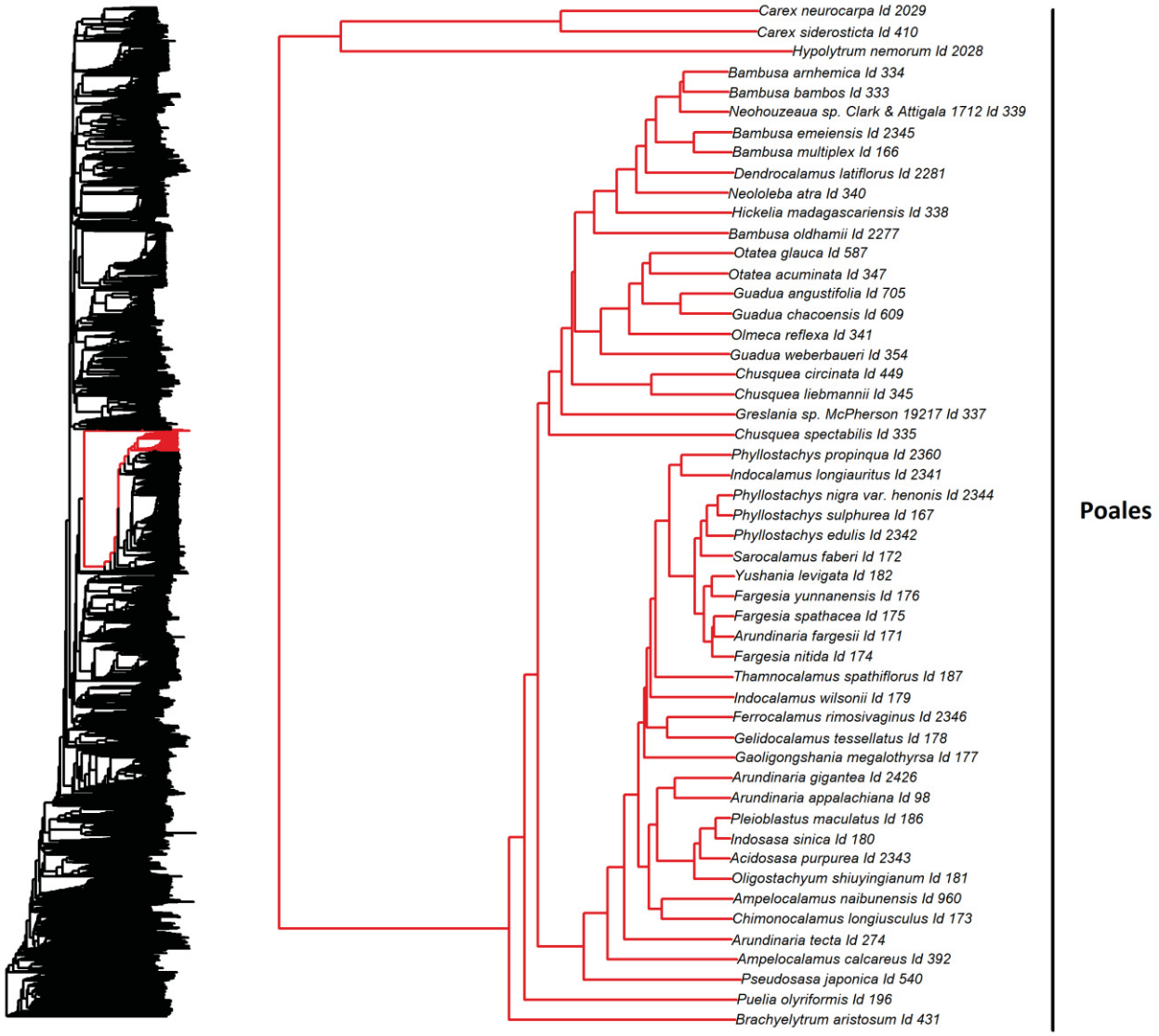


Figura 17: Posição da ordem Poales na árvore filogenética vetorial construída pelo método livre de alinhamento Svect. Trecho da filogenia destacando parte das espécies pertencentes à Poales. Fonte: O autor.

A gigante ordem Poales, contendo 15 famílias, 997 gêneros e 18.875 espécies possui grande importância econômica, pois engloba espécies como o trigo (*Triticum*), o Milho (*Zea*) e o Arroz (*Oryza*). Os espécimes deste táxon possuem pequenas flores revestidas por brácteas dispostas em inflorescências que geralmente são polinizadas por anemocoria, e suas sementes comumente acumulam amido (STEVENS, 2012).

A Fig. 17 representa uma parte das espécies de Poales utilizadas em nossa análise, e quando sua topologia é comparada com parte da árvore resultante do estudo de SAARELA et al. (2018) representado na Fig. 18, percebemos que os gêneros *Phyllostachis, Fargesia, Salocalamus,* e *Yushania* estão alocados no grupo irmão de *Arundinaria, Indososa, pleioblastus,* e *acidosasa,* que por sua vez, estão arranjados como grupo irmão de *Bambusa, Greslania,*

Neoleda Dendrocalamus, Otatea, Olmeca e *Guadua.* Assim, validamos a reconstrução filogenética livre de alinhamento também para a ordem Poales.

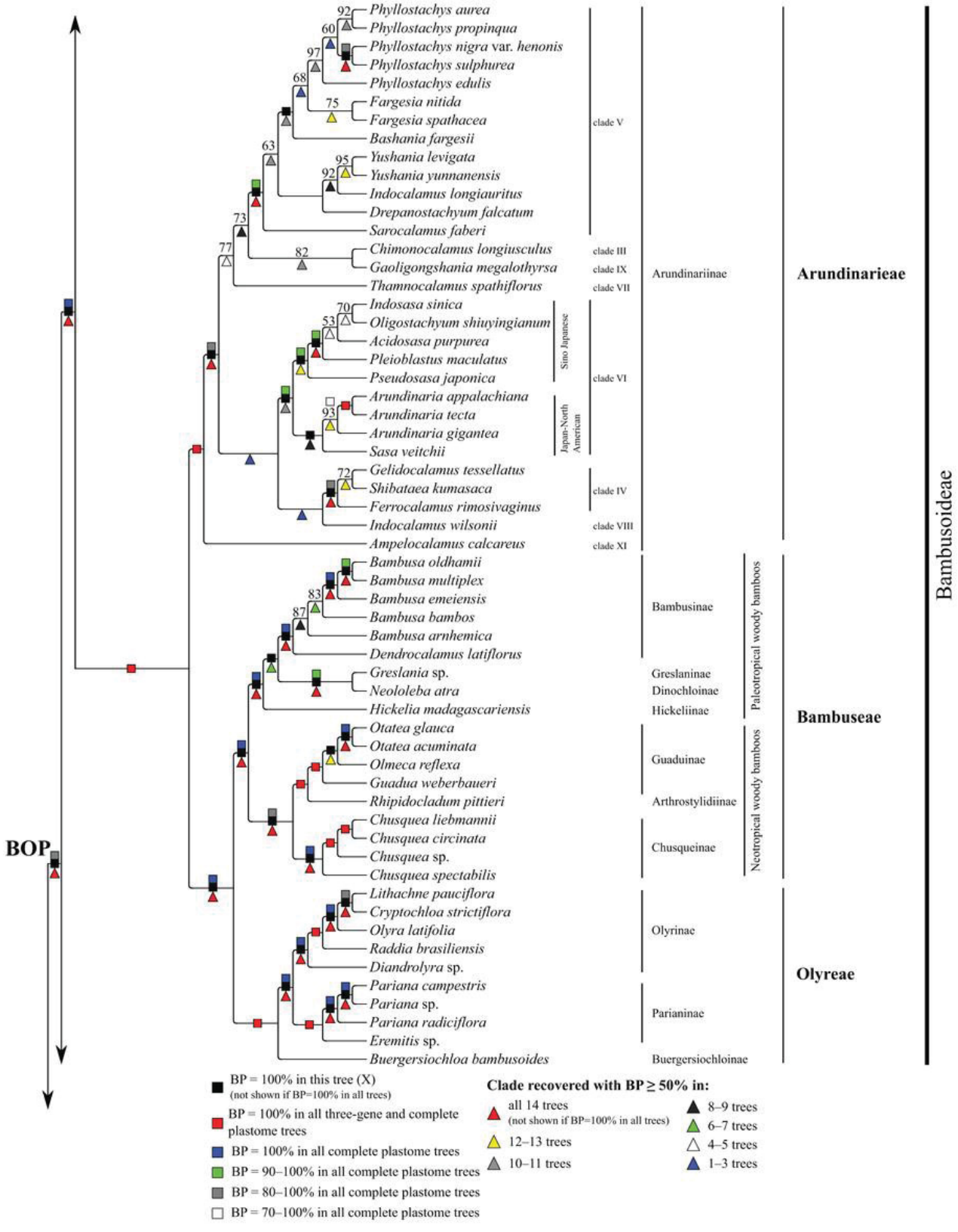


Figura 18: Clado Bambusoideae da árvore baseada em ML e construída por Saarela et. al. (2018) com 250 plastomas completos de espécies da ordem Poales. Fonte: (SAARELA et al., 2018)**.**

A Fig. 19 nos mostra parte do ramo das gimnospermas na grande árvore. As Gimnospermas são um grupo parafilético que inclui as plantas de “sementes nuas”. Pertencem a essa divisão as ordens: Pinales (7 famílias, 68 gêneros, 545 espécies), Gnetales (3 famílias, 3 gêneros, 93 espécies), Gingkoales (1 família, 1 gênero, 1 espécie) e Cycadales (2 famílias, 10 gêneros, 305 espécies), conforme STEVENS (2012).

Ao compararmos a topologia deste grupo na análise global com outros trabalhos, vemos relações muito coerentes. LU et al., (2014) usou os genes nucleares de cópia única LFY e NLY para resolver relações interfamiliares em gimnospermas. A Figura 20 nos permite afirmar que as ramificações e relações entre as ordens diferem sutilmente, mas as topologias a nível de espécie se encontram em consonância. Em Gnetales, por exemplo o clado *Welwítschia/Gnetum/Ephedra*, apesar de diferirem no número de OTU’s, possuem um posicionamento filogenético semelhante em relação a arvore global.

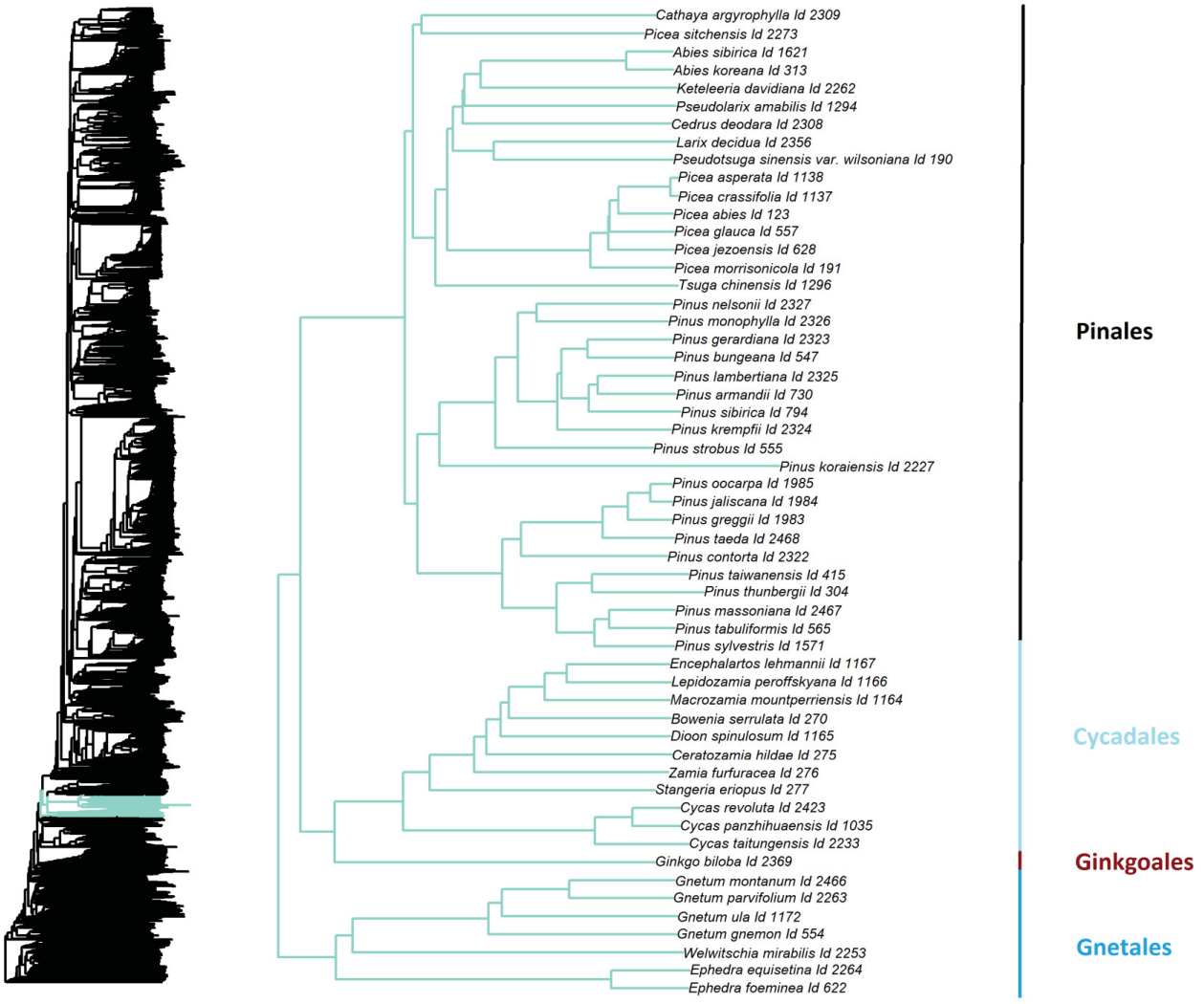


Figura 19: Posição das ordens de Gimnospermas na árvore filogenética vetorial construída pelo método livre de alinhamento SVect. Fonte: O Autor.

Ao compararmos o trecho do dendrograma de LU et al., (2014) equivalente à ordem Cycadales (Fig. 20) com o fragmento correspondente à mesma ordem em nossa filogenia, observamos que ambas as análises colocam o genero *Cycas* como o ramo mais basal do clado, E apesar das diferenças entre o número espécies utilizadas em cada estudo, o nó continua consistente, arranjando organismos estudados pelas duas análises de maneira similar. A mesma situação ocorre também com a ordem Pinales, conferindo confiança ao método e à reconstrução filogenética em questão.

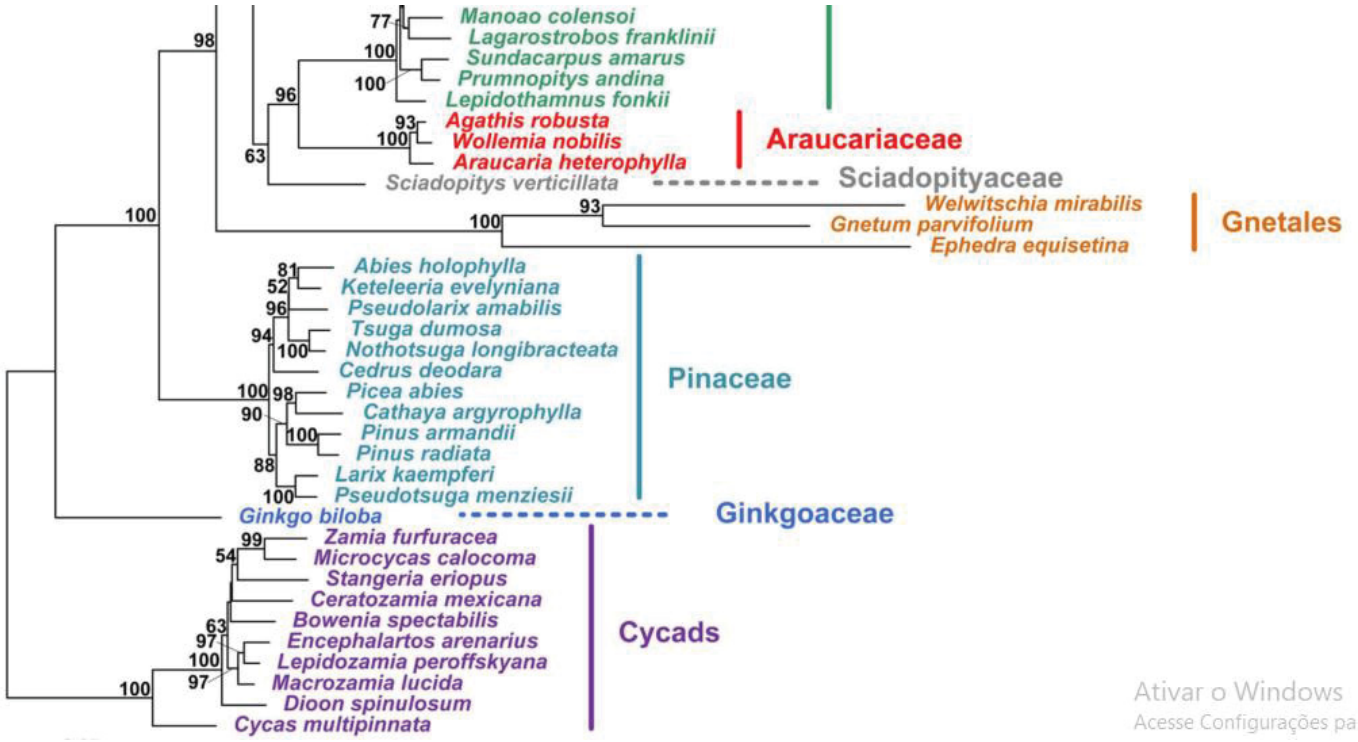


Figura 20: Parte da árvore filogenética das gimnospermas Fonte: Adaptado de LU et al., (2014)**.**

Para mostrar o posicionamento taxonômico de uma nova espécie de Pinus, YU et al., (2017) obteve uma análise filogenética Fig (21) que acorda com as posições dos organismos da árvore criada nesse estudo. As espécies *P. taiwanensis, P. thunbergii, P. tabulaeformis, P. massoniana, P. contorda* e *P. taeda* se encontram no táxon irmão de *P. koraiensis* e *P. lambertiana*. Assim, podemos afirmar que o arranjo da árvore global de cloroplastos e plastídios para gimnospermas também está em harmonia com outros estudos atuais.

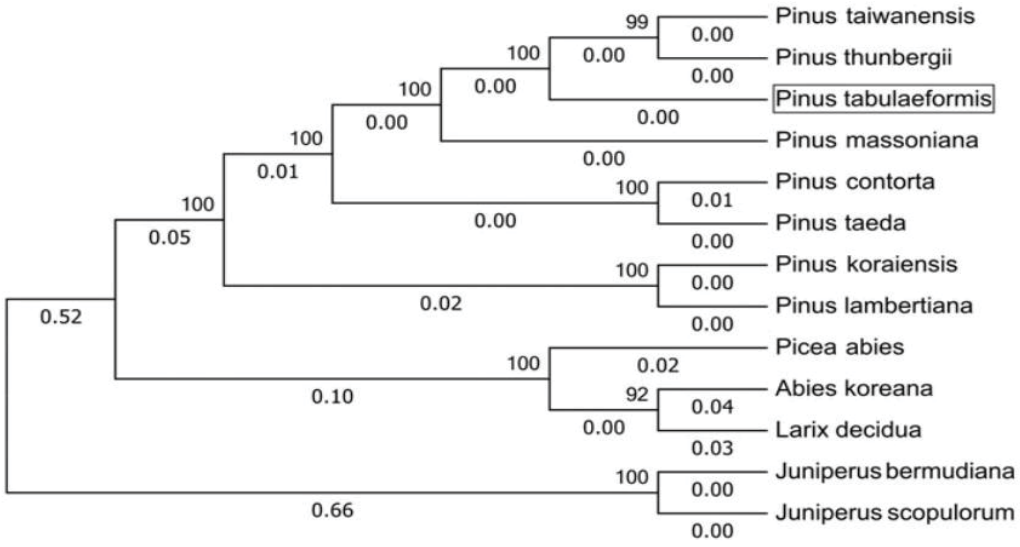


Figura 21: Topologia do cladograma de gimnospermas, com destaque para o gênero Pinus. Fonte: (YU et al., 2017).

O fragmento da árvore filogenética global para parte das algas vermelhas é representada pela Figura 22, e encontra respaldo nos cientifico nos recentes estudos de YANG et al., (2015) e PAIANO et al., (2018). Na Figura 23, YANG et al., (2015) coloca as espécies *Grateloupia taiwanensis* e *Sebdenia flabellata* no mesmo ramo, da mesma forma que o fixa como grupo irmão das Gracilariales, fenômenos também observados na arvore global. As Figuras 22 e 23 ainda acordam em relação à posição do clado Florideophyceae como grupo irmão de Bangiophyceae. Outro fato que confere validade a arvore global é a colocação da classe Cyanidiophyceae como táxon basal da análise, visto que se repete na reconstrução por vetorização. Da mesma forma a Figura 24 retrata uma topologia similar, apesar da diferença no número de espécies.

Diversos estudos também validam a posição das ordens na reconstrução filogenética proposta aqui (MYBURG et al., 2014; RUHFEL et al., 2014; QIAN; ZHANG, 2016; THE ANGIOSPERM PHYLOGENY GROUP VI, 2016; LIU et al., 2017), o que nos permite inferir que o modelo utilizado para as reconstruções filogenéticas são, além de computacionalmente rápidos, eficientes para análises filogenéticas em larga escala baseados em dados proteômicos de plastídios.

Alguns fragmentos da árvore global ficaram indefiníveis. A provável causa dos erros na configuração dos ramos é a má qualidade das sequências. Se os dados de entrada forem mal anotados ou mal sequenciados, toda a análise posterior será prejudicada(PHILIPPE et al., 2017). Além disso, o número de indivíduos de cada táxon presentes na análise tem influência direta sobre a análise filogenética. É necessário uma boa representatividade para uma análise de qualidade.

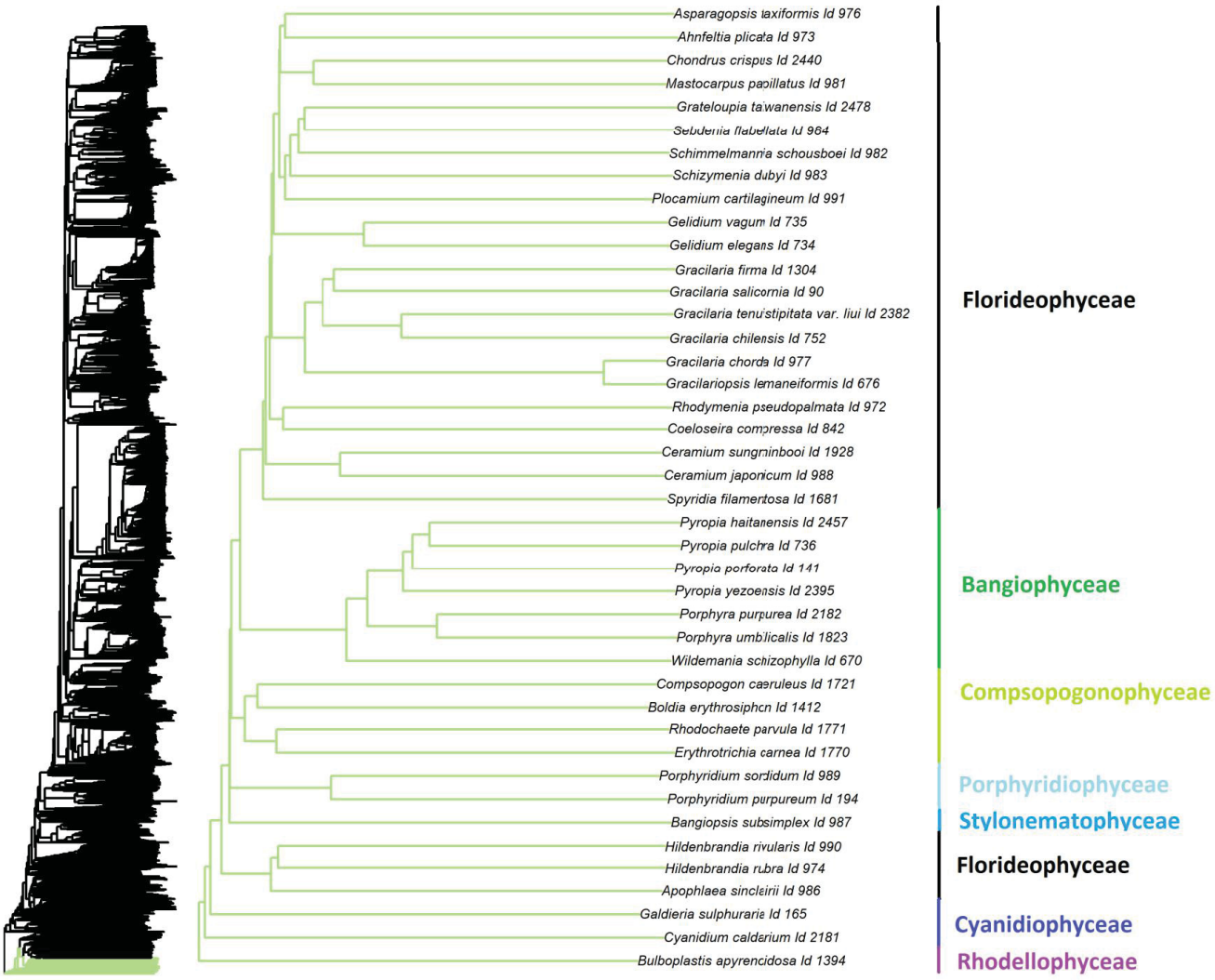


Figura 22: Posição das classes de algas vermelhas na árvore filogenética vetorial construída pelo método livre de alinhamento SVect. Fonte: O Autor.

A validação dos fragmentos da árvore filogenética global foi realizada em sua totalidade por meio da legenda interna de cada figura. As imagens contendo os trechos da árvore foram legendadas por classificação taxonômica de ordem, no caso de plantas vasculares e briófitas, e por classe, no caso das algas. As classificações foram obtidas por meio do *Angiosperm Phylogeny Website* (STEVENS, 2012) e do *Algaebase* (GUIRY; GUIRY, 2018).

Dessa forma, é possível identificar se determinado táxon avaliado tem um certo nível de coerência, pela ordem ou classe do grupo, em comparação às classificações dos organismos subjacentes.

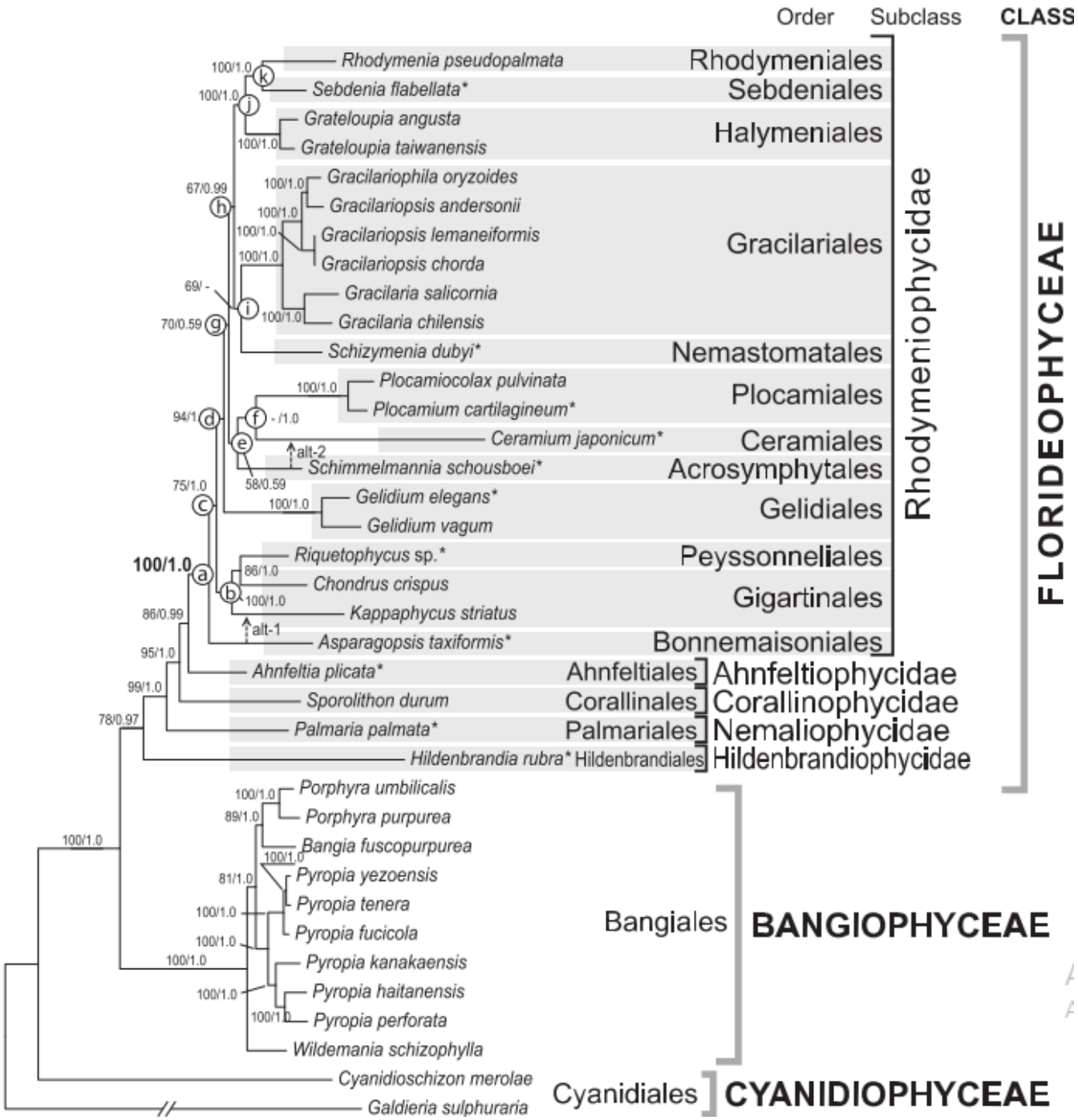


Figura 23: Reconstrução filogenética de algas vermelhas baseada em 24 genes mitocondriais concatenados. Fonte: (YANG et al., 2015)

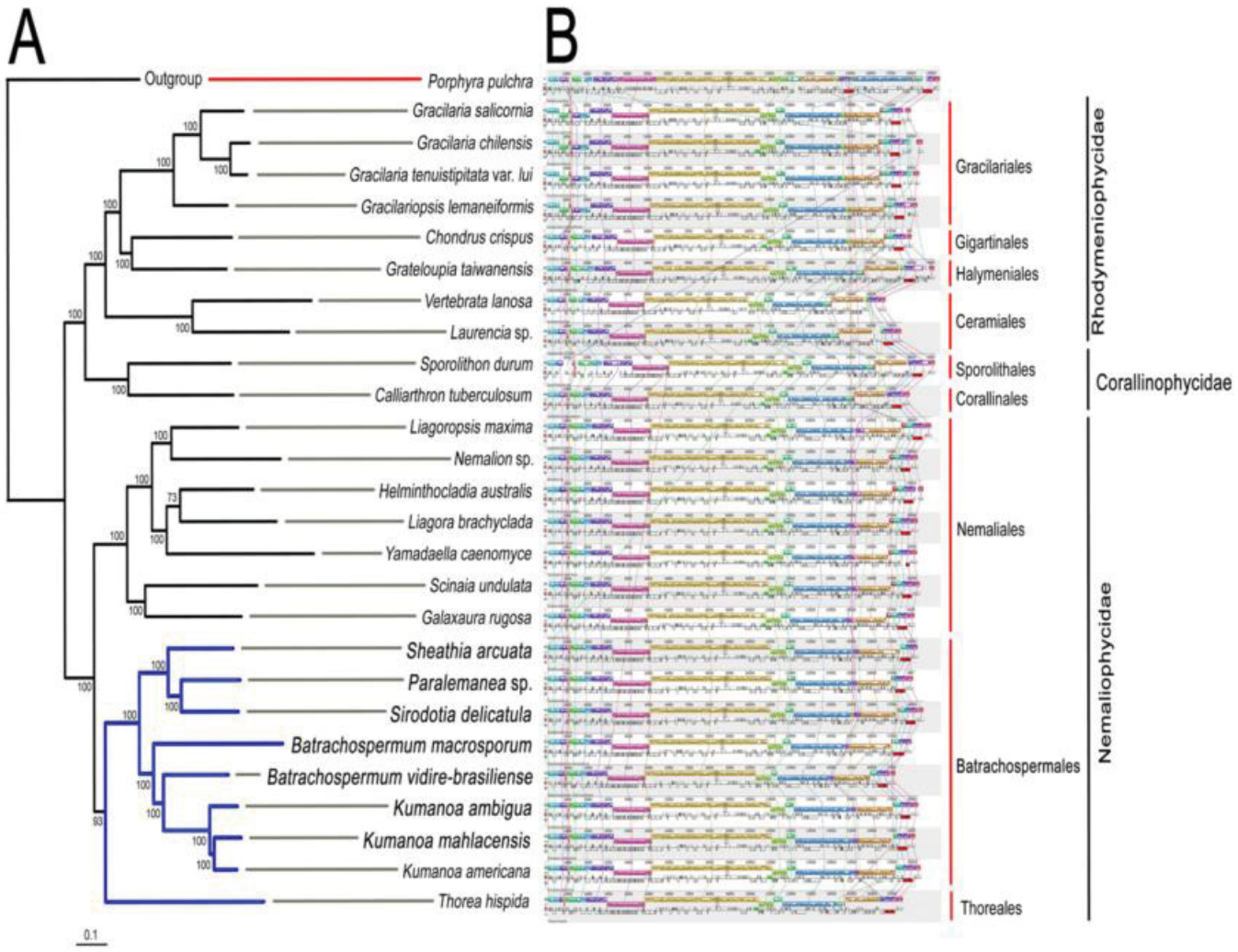


Figura 24: Relações filogenéticas e rearranjo plastidial de Florideophyceae, baseados em 177 genes concatenados.
Fonte: (PAIANO et al., 2018).

7. CONCLUSÃO

Por meio desse trabalho, podemos concluir que:

- A filogenômica enfrenta grandes desafios no que diz respeito à utilização do volume de dados disponíveis;

- Tanto a abordagem de super matiz quanto a abordagem de super árvore possuem limitações quando manipulam um grande conjunto de dados, pois utilizam alinhamentos múltiplos;

- O alinhamento múltiplo não atende as demandas atuais de desempenho computacional, e torna-se inviável quando trabalhamos com um conjunto grande de sequencias.

- A vetorização de sequências, somado à redução de dimensionalidade mostrou-se viável, rápida e eficiente quando aplicada a análises filogenéticas em larga escala;

- Utilizar a vetorização somada à redução de dimensão em alguns casos pode superar o alinhamento de sequencias quando se trata de qualidade de resultados;

- A metodologia proposta se mostrou 45 vezes mais rápida em relação ao alinhamento múltiplo de sequências;

- O SVect torna prescindível a concatenação ordenada para simular uma sintenia genética, o que exprime a vantagem do método em relação ao alinhamento múltiplo de sequências, visto na diferença de tempo de processamento das sequências..

- A qualidade da análise está condicionada a qualidade dos dados de sequência utilizados;

- A distância euclidiana e o tamanho do vetor em 600 pontos numéricos provaram ser o equilíbrio entre qualidade e desempenho para análises envolvendo redução de dimensionalidade em dados de regiões codificantes em genomas plastidiais;

- A árvore filogenética global de plastídios neste trabalho proposta é coerente, e foi validada com os dados já publicados.

8. PERSPECTIVAS FUTURAS

O próximo passo é disponibilizar uma ferramenta web para visualização da filogenia global de plastídios proposta nesse trabalho, com a funcionalidade de inserção de novos organismos na árvore. A ferramenta funcionará como classificador, devolvendo ao pesquisador a posição taxonômica dos espécimes cujas CDS's plastidiais sejam submetidas na ferramenta.

Outra necessidade é realizar um estudo detalhado para descobrir o efeito das HTG's (transferências horizontais de genes) nessas árvores resultantes da vetorização e eventual projeção em um hiper-espaço.

REFERÊNCIAS

ANDERSON, J. M. The molecular organization of chloroplast thylakoids. **BBA Reviews On Bioenergetics**, v. 416, n. 2, p. 191–235, 1975.

BAYLY, M. J.; RIGAULT, P.; SPOKEVICIUS, A.; et al. Chloroplast genome analysis of Australian eucalypts - Eucalyptus, Corymbia, Angophora, Allosyncarpia and Stockwellia (Myrtaceae). **Molecular Phylogenetics and Evolution**, v. 69, n. 3, p. 704–716, 2013. Elsevier Inc. Disponível em: <http://dx.doi.org/10.1016/j.ympev.2013.07.006>. .

BENSON, D. A.; CAVANAUGH, M.; CLARK, K.; et al. GenBank. **Nucleic Acids Research**, v. 45, n. D1, p. D37–D42, 2017.

BLANKENSHIP, R. E. **Molecular Mechanisms of Photosynthesis**. Oxford, UK: Blackwell Scientific, 2013.

BLEIDORN, C. Third generation sequencing: Technology and its potential impact on evolutionary biodiversity research. **Systematics and Biodiversity**, v. 14, n. 1, p. 1–8, 2016.

CARVALHO, M. C. D. C. G. DE; SILVA, D. C. G. DA. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural**, v. 40, n. 3, p. 735–744, 2010.

CATANHO, M.; MIRANDA, A. B. DE; DEGRAVE, W. Comparando genomas: bancos de dados e ferramentas computacionais para a análise comparativa de genomas procarióticos. **Reciis**, v. 1, n. 2, p. 335–358, 2007. Disponível em: <http://www.reciis.cict.fiocruz.br/index.php/reciis/article/view/105/121>. .

CHASE, M. W.; SOLTIS, D. E.; OLMSTEAD, R. G.; et al. Phylogenetics of Seed Plants: An Analysis of Nucleotide Sequences from the Plastid Gene rbcL. **Annals of the Missouri Botanical Garden**, v. 80, n. 3, p. 528, 1993. Disponível em: <http://www.jstor.org/stable/2399846?origin=crossref>. .

CLEGG, M. T.; GAUT, B. S.; LEARN, G. H.; MORTON, B. R. Rates and patterns of chloroplast DNA evolution. **Proceedings of the National Academy of Sciences of the United States of America**, v. 91, n. 15, p. 6795–67801, 1994.

CLELAND, W. W.; ANDREWS, T. J.; GUTTERIDGE, S.; HARTMAN, F. C.; LORIMER, G. H. Mechanism of Rubisco: The Carbamate as General Base $^{\chi}$. **Chemical Reviews**, v. 98, n. 2, p. 549–562, 1998. Disponível em: <http://pubs.acs.org/doi/abs/10.1021/cr970010r>. .

CUI, L.; VEERARAGHAVAN, N.; RICHTER, A.; et al. ChloroplastDB: the Chloroplast Genome Database. **Nucleic acids research**, v. 34, n. Database issue, p. D692-6, 2006. Disponível em: <http://nar.oxfordjournals.org/content/34/suppl_1/D692.abstract>. Acesso em: 14/6/2016.

DANIELL, H.; LIN, C.-S.; YU, M.; CHANG, W.-J. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. **Genome Biology**, v. 17, n. 1, p. 134, 2016. Genome Biology. Disponível em: <http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1004-2>. .

DAY, W. H. E. Computational complexity of inferring phylogenies from dissimilarity matrices. **Bulletin of Mathematical Biology**, v. 49, n. 4, p. 461–467, 1987.

DINIZ, F. Biotecnologia : modernas técnicas de engenharia genética unem setores em prol do desenvolvimento nacional. **Biotecnologia Ciência & Desenvolvimento**, , n. 37, p. 2–5, 2007.

EGEA, I.; BARSAN, C.; BIAN, W.; et al. Chromoplast differentiation: Current status and perspectives. **Plant and Cell Physiology**, v. 51, n. 10, p. 1601–1611, 2010.

EMMS, D. M.; KELLY, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. **Genome Biology**, v. 16, n. 1, p. 157, 2015. Genome Biology. Disponível em: <http://genomebiology.com/2015/16/1/157%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/26243257%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4531804>. .

FELSENSTEIN, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. **Journal of Molecular Evolution**, v. 17, n. 6, p. 368–376, 1981.

FELSENSTEIN, J. Inferring Phylogenies from Protein Sequences by Parsimony, Distance, and Likelihood Methods. . v. 266, p.418–427, 1996. Methods Enzymol.

FITCH, W. M. Toward Defining the Course for a Specific Tree. **Topology.Syst.Zool.**, v. 20, p. 406–416, 1971.

FITZJOHN, R. G.; PENNELL, M. W.; ZANNE, A. E.; et al. How much of the world is woody? **Journal of Ecology**, v. 102, n. 5, p. 1266–1272, 2014.

FLEISCHAUER, M.; BÖCKER, S. Bad Clade Deletion Supertrees: A Fast and Accurate Supertree Algorithm. **Molecular biology and evolution**, v. 34, n. 9, p. 2408–2421, 2017.

FRANKL, P.; MAEHARA, H. The Johnson-Lindenstrauss lemma and the sphericity of some graphs. **Journal of Combinatorial Theory, Series B**, v. 44, n. 3, p. 355–362, 1988.

GOULD, S. B.; WALLER, R. F.; MCFADDEN, G. I. Plastid evolution. **Annual Review of Plant Biology**, v. 59, p. 491–517, 2008. Disponível em: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=18315522%5Cnpapers2://publication/doi/10.1146/annurev.arplant.59.032607.092915%5Cnhttp://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&amp>. Acesso em: 14/6/2016.

GREENE, C. S.; TAN, J.; UNG, M.; MOORE, J. H.; CHENG, C. Big data bioinformatics. **Journal of Cellular Physiology**, v. 229, n. 12, p. 1896–1900, 2014.

GUIRY, M. D.; GUIRY, G. M. AlgaeBase. Disponível em: <http://www.algaebase.org>. Acesso em: 1/3/2018.

HELDT, W. H.; HELDT, F. **Plant Biochemistry**. 3rd ed. Academic Press, 2005.

HUELSENBECK, J. P.; RONQUIST, F. MRBAYES: Bayesian inference of phylogenetic trees. **Bioinformatics**, v. 17, n. 8, p. 754–755, 2001. Disponível em: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/17.8.754>. .

JOHNSON, W. B.; LINDENSTRAUSS, J. Extensions of Lipschitz mappings into a Hilbert space. **Contemporary Mathematics**, v. 26, n. September, p. 189–206, 1984. Disponível em: <http://www.ams.org/conm/026/>. .

JONES, K. E.; PURVIS, A.; MACLARNON, A.; BININDA-EMONDS, O. R. P.; SIMMONS, N. B. A phylogenetic supertree of the bats (Mammalia: Chiroptera). **Biological Reviews of the Cambridge Philosophical Society**, v. 77, n. 2, p. 223–259, 2002.

KREBS, J. E.; GOLDSTEIN, E. S.; KILPATRICK, S. T. **Lewin's Genes XI**. 2014.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular biology and evolution**, v. 33, n. 7, p. 1870–1874, 2016.

LI, Q. L.; GUO, J. Z.; YAN, N.; LI, C. C. Complete chloroplast genome sequence of cultivated Morus L . species. **Genetics and Molecular Research**, v. 15, n. 60, 2016.

LIU, M.; ZHAO, J.; WANG, J.; LIU, Z.; LIU, G. Phylogenetic analysis of 25 plant species

representing 19 angiosperm families and one gymnosperm family based on 390 orthologous genes. **Plant Systematics and Evolution**, v. 303, n. 3, p. 413–417, 2017. Springer Vienna.

LOPEZ-JUEZ, E.; PYKE, K. A. Plastids unleashed: Their development and their integration in plant development. **International Journal of Developmental Biology**, v. 49, n. 5–6, p. 557–577, 2005.

LU, Y.; RAN, J. H.; GUO, D. M.; YANG, Z. Y.; WANG, X. Q. Phylogeny and divergence times of gymnosperms inferred from single-copy nuclear genes. **PLoS ONE**, v. 9, n. 9, 2014.

LYUBETSKY, .V A.; SELIVERSTOV, A. V.; ZVERKOV, O. A. Построение разделяющих паралоги семейств гомологичных белков, кодируемых в пластидах цветковых растений. **Mathematical Biology and Bioinformatcs**, v. 8, n. 1, p. 225–233, 2013.

MATTHEWS, M. L.; ENDRESS, P. K. Comparative floral structure and systematics in Cucurbitales (Corynocarpaceae, Coriariaceae, Tetramelaceae, Datiscaceae, Begoniaceae, Cucurbitaceae, Anisophylleaceae). **Botanical Journal of the Linnean Society**, v. 145, n. 2, p. 129–185, 2004. Disponível em: <http://springerlink.metapress.com/openurl.asp?genre=article&id=doi:10.1007/s00606-003-0090-2>. .

MCCARTHY, C. G. P.; FITZPATRICK, D. A. Phylogenomic Reconstruction of the Oomycete Phylogeny Derived from 37 Genomes. **mSphere**, v. 2, n. 2, p. 1–17, 2017. Disponível em: <http://msphere.asm.org/content/2/2/e00095-17?utm_source=TrendMDmSphere&utm_medium=TrendMDmSphere&utm_campaign=trendmdalljournals_0>. .

MCMAHON, M. M.; SANDERSON, M. J. Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. **Systematic Biology**, v. 55, n. 5, p. 818–836, 2006.

MYBURG, A. A.; GRATTAPAGLIA, D.; TUSKAN, G. A.; et al. The genome of Eucalyptus grandis. **Nature**, v. 510, n. 7505, p. 356–362, 2014.

NICOLAU, P. B. História Da Classificação Biológica. , 2017. Disponível em: <https://repositorioaberto.uab.pt/bitstream/10400.2/6133/1/texto_apoio_1_Hist_classif_biologica.pdf>. Acesso em: 7/3/2017.

O'BRIEN, E. A.; ZHANG, Y.; WANG, E.; et al. GOBASE: An organelle genome database. **Nucleic Acids Research**, v. 37, n. SUPPL. 1, p. 946–950, 2009.

O'LEARY, N. A.; WRIGHT, M. W.; BRISTER, J. R.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. **Nucleic Acids Research**, v. 44, n. D1, p. D733–D745, 2016.

PAIANO, M. O.; CORTONA, A. DEL; COSTA, J. F.; et al. Organization of plastid genomes in the freshwater red algal order Batrachospermales (Rhodophyta). **Journal of Phycology**, v. 54, n. 1, p. 25–33, 2018.

PATWARDHAN, A.; RAY, S.; ROY, A. Biology Molecular Markers in Phylogenetic Studies-A Review. **J Phylogen Evolution Biol**, v. 2, n. 22, 2014. Disponível em: <http://dx.doi.org/10.4172/2329-9002.1000131>. Acesso em: 9/3/2017.

PHILIPPE, H.; VIENNE, D. M. DE; RANWEZ, V.; et al. Pitfalls in supermatrix phylogenomics. **European Journal of Taxonomy**, , n. 283, p. 1–25, 2017. Disponível em: <http://www.europeanjournaloftaxonomy.eu/index.php/ejt/article/view/407>. .

PIERRI, C. R. DE. **REPRESENTAÇÕES VETORIAIS DE PROTEOMAS : UM ESTUDO DE CASO REPRESENTAÇÕES VETORIAIS DE PROTEOMAS : UM ESTUDO DE CASO**, 2017. Universidade Federal do Paraná.

PRUITT, K.; BROWN, G.; TATUSOVA, T.; MAGLOTT, D. Chapter 18 : The Reference Sequence (RefSeq) Database Database. **The NCBI Handbook**, , n. Bethesda(MD): National Center for BIotechnology Information (US), p. 1–22, 2002.

PRUITT, K. D.; TATUSOVA, T.; MAGLOTT, D. R. NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. **Nucleic Acids Research**, v. 35, n. SUPPL. 1, p. 501–504, 2007.

PYRON, R.; BURBRINK, F. T.; WIENS, J. J.; et al. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. **BMC Evolutionary Biology**, v. 13, n. 1, p. 93, 2013. Disponível em: <http://bmcevolbiol.biomedcentral.com/articles/10.1186/1471-2148-13-93>. .

QIAN, H.; ZHANG, J. Are phylogenies derived from family-level supertrees robust for studies on macroecological patterns along environmental gradients? **Journal of Systematics and Evolution**, v. 54, n. 1, p. 29–36, 2016.

RANNALA, B.; YANG, Z. H. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. **Journal of Molecular Evolution**, v. 43, n. 3, p. 304–311, 1996.

REITH, M.; MUNHOITAND, J. Complete Nucleotide Sequence of the Porphyra purpurea Chloroplast Genome. **Plant Molecular Biology Reporter**, v. 13, n. 4, p. 333–335, 1995.

RENNER, S. S.; SCHAEFER, H. Phylogeny and Evolution of the Cucurbitaceae. **Plant Genetics and Genomics: Crops and Models**, p. 1–26, 2016. Disponível em: <springer.com/chapter/10.1007/7397_2016_14>. .

ROGALSKI, M.; NASCIMENTO VIEIRA, L. DO; FRAGA, H. P.; GUERRA, M. P. Plastid genomics in horticultural species: importance and applications for plant population genetics, evolution, and biotechnology. **Frontiers in Plant Science**, v. 6, n. July, p. 1–17, 2015. Disponível em: <http://www.frontiersin.org/Crop_Science_and_Horticulture/10.3389/fpls.2015.00586/abstract>. .

RUHFEL, B. R.; GITZENDANNER, M. A.; SOLTIS, P. S.; et al. From algae to angiosperms–inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. **BMC Evolutionary Biology**, v. 14, n. 1, p. 23, 2014. Disponível em: <http://bmcevolbiol.biomedcentral.com/articles/10.1186/1471-2148-14-23>. .

SAARELA, J. M.; BURKE, S. V.; WYSOCKI, W. P.; et al. A 250 plastome phylogeny of the grass family (Poaceae): topological support under different data partitions. **PeerJ**, v. 6, p. e4299, 2018. Disponível em: <https://peerj.com/articles/4299>. .

SABLOK, G.; MUDUNURI, S. B.; PATNANA, S.; et al. Chloromitossrdb: Open source repository of perfect and imperfect repeats in organelle genomes for evolutionary genomics. **DNA Research**, v. 20, n. 2, p. 127–133, 2013.

SABLOK, G.; PADMA RAJU, G. V.; MUDUNURI, S. B.; et al. ChloroMitoSSRDB 2.00: More genomes, more repeats, unifying SSRs search patterns and on-the-fly repeat detection. **Database**, v. 2015, p. 1–10, 2015.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, , n. January, 1987. Disponível em: <https://academic.oup.com/mbe/article/4/4/406/1029664/The-neighborjoining-method-a-new-method-for>. .

SANGER, F.; COULSON, A. R. A rapid method for determining sequences in DNA by primed

synthesis with DNA polymerase. **Journal of Molecular Biology**, v. 94, n. 3, p. 441–448, 1975. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/0022283675902132>. .

SHANKER, A. ChloroSSRdb: a repository of perfect and imperfect chloroplastic simple sequence repeats (cpSSRs) of green plants. **Database : the journal of biological databases and curation**, v. 2014, n. 2, p. 1–5, 2014.

SHAW, J.; LICKEY, E. B.; SCHILLING, E. E.; SMALL, R. L. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The Tortoise and the hare III. **American Journal of Botany**, v. 94, n. 3, p. 275–288, 2007.

SHENDURE, J.; BALASUBRAMANIAN, S.; CHURCH, G. M.; et al. DNA sequencing at 40: Past, present and future. **Nature**, v. 550, n. 7676, 2017. Nature Publishing Group. Disponível em: <http://dx.doi.org/10.1038/nature24286>. .

SHENDURE, J.; MITRA, R. D.; VARMA, C.; CHURCH, G. M. Advanced sequencing technologies: methods and goals. **Nature reviews. Genetics**, v. 5, n. 5, p. 335–344, 2004. Disponível em: <http://www.nature.com/nrg/journal/v5/n5/full/nrg1325.html>. .

SIEVERS, F.; HIGGINS, D. G. Clustal Omega. **Current Protocols in Bioinformatics**, v. 2014, n. December, p. 3.13.1-3.13.16, 2014.

SIM, A. Y. L.; MINARY, P.; LEVITT, M. Modeling nucleic acids. **Current Opinion in Structural Biology**, v. 22, n. 3, p. 273–278, 2012. Elsevier Ltd. Disponível em: <http://dx.doi.org/10.1016/j.sbi.2012.03.012>. .

SIMPSON, J. T.; POP, M. The Theory and Practice of Genome Sequence Assembly. **Annu. Rev. Genomics Hum. Genet**, v. 16, p. 153–72, 2015.

SMITH, S. A.; BEAULIEU, J. M.; DONOGHUE, M. J. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. **BMC Evolutionary Biology**, v. 9, n. 1, p. 1–12, 2009.

SMITH, S. A.; BROWN, J. W. Constructing a broadly inclusive seed plant phylogeny. **American Journal of Botany**, v. 105, n. 3, p. 1–13, 2018.

SOKAL, R; MICHENER, C. A statistical method for evaluating systematic relationships. **The University of Kansas science bulletin**, v. 38, n. 2, 1958.

STAMATAKIS, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics**, v. 22, n. 21, p. 2688–2690, 2006.

STEIN, L. GENOME ANNOTATION: FROM SEQUENCE TO BIOLOGY. **NATURE REVIEWS | GENETICS**, v. 2, p. 493–505, 2001. Disponível em: <http://www.nature.com/nrg/journal/v2/n7/pdf/nrg0701_493a.pdf>. Acesso em: 20/6/2016.

STEPHENS, Z. D.; LEE, S. Y.; FAGHRI, F.; et al. Big data: Astronomical or genomical? **PLoS Biology**, v. 13, n. 7, p. 1–11, 2015.

STEVENS, P. F. Angiosperm Phylogeny Website. Disponível em: <http://www.mobot.org/MOBOT/research/APweb/>. Acesso em: 1/3/2018.

SWOFFORD, DA. L. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). **Journal of Molecular Evolution**, 2002. Sinauer Associates, Sunderland, Massachusetts.

THE ANGIOSPERM PHYLOGENY GROUP VI. 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. 161, n. 2, p. 105–121, 2016.

TROY, C. S.; MACHUGH, D. E.; BAILEY, J. F.; et al. Genetic evidence for Near-Eastern origins of European cattle. **Nature**, v. 410, n. 6832, p. 1088–1091, 2001.

VINGA, S.; ALMEIDA, J. Alignment-free sequence comparison - A review. **Bioinformatics**, 2003.

WATSON, J. D.; CRICK, F. H. C. Molecular structure of nucleic acids. **Nature**, 1953. Disponível em: <http://www.nature.com/physics/looking-back/crick/%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/13054692>. .

WEEDEN, N. F. Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. **Journal of Molecular Evolution**, v. 17, n. 3, p. 133–139, 1981.

WOLFE, K. H.; MORDENT, C. W.; PALMERS, J. D. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant (chloroplast DNA/Epifagus virginiana/transtdon/rascriptlon/photosyntbesis). **Evolution**, v. 89, p. 10648–10652, 1992.

WOLFSBERG, T. G.; SCHAFER, S.; TATUSOV, R. L.; TATUSOVA, T. A. Organelle genome resources at NCBI. **Trends in Biochemical Sciences**, v. 26, n. 3, p. 199–203, 2001.

YANDELL, M.; ENCE, D. A beginner's guide to eukaryotic genome annotation. **Nature Publishing Group**, v. 13, 2012.

YANG, A.; TROUP, M.; HO, J. W. K. Scalability and Validation of Big Data Bioinformatics Software. **Computational and Structural Biotechnology Journal**, v. 15, p. 379–386, 2017. The Authors. Disponível em: <https://doi.org/10.1016/j.csbj.2017.07.002>. .

YANG, E. C.; KIM, K. M.; KIM, S. Y.; et al. Highly conserved mitochondrial genomes among multicellular red algae of the florideophyceae. **Genome Biology and Evolution**, v. 7, n. 8, p. 2394–2406, 2015.

YANG, Z. Phylogenetic analysis using parsimony and likelihood methods. **Journal of Molecular Evolution**, v. 42, n. 2, p. 294–307, 1996. Disponível em: <http://www.springerlink.com/index/10.1007/BF02198856>. .

YANG, Z.; RANNALA, B. Bayesian phylogentic inference using DNA sequences: A Markov chain monte carlo method. **Molecular Biology and Evolution**, v. 14, n. 7, p. 717–724, 1997. Disponível em: <http://mbe.oxfordjournals.org/cgi/content/abstract/14/7/717>. .

YU, Z.; PENG, S.; YANG, P. The complete chloroplast genome of the southern Chinese pine Pinus tabuliformis (Pinales: Pinaceae). **Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis**, v. 28, n. 1, p. 13–14, 2017.

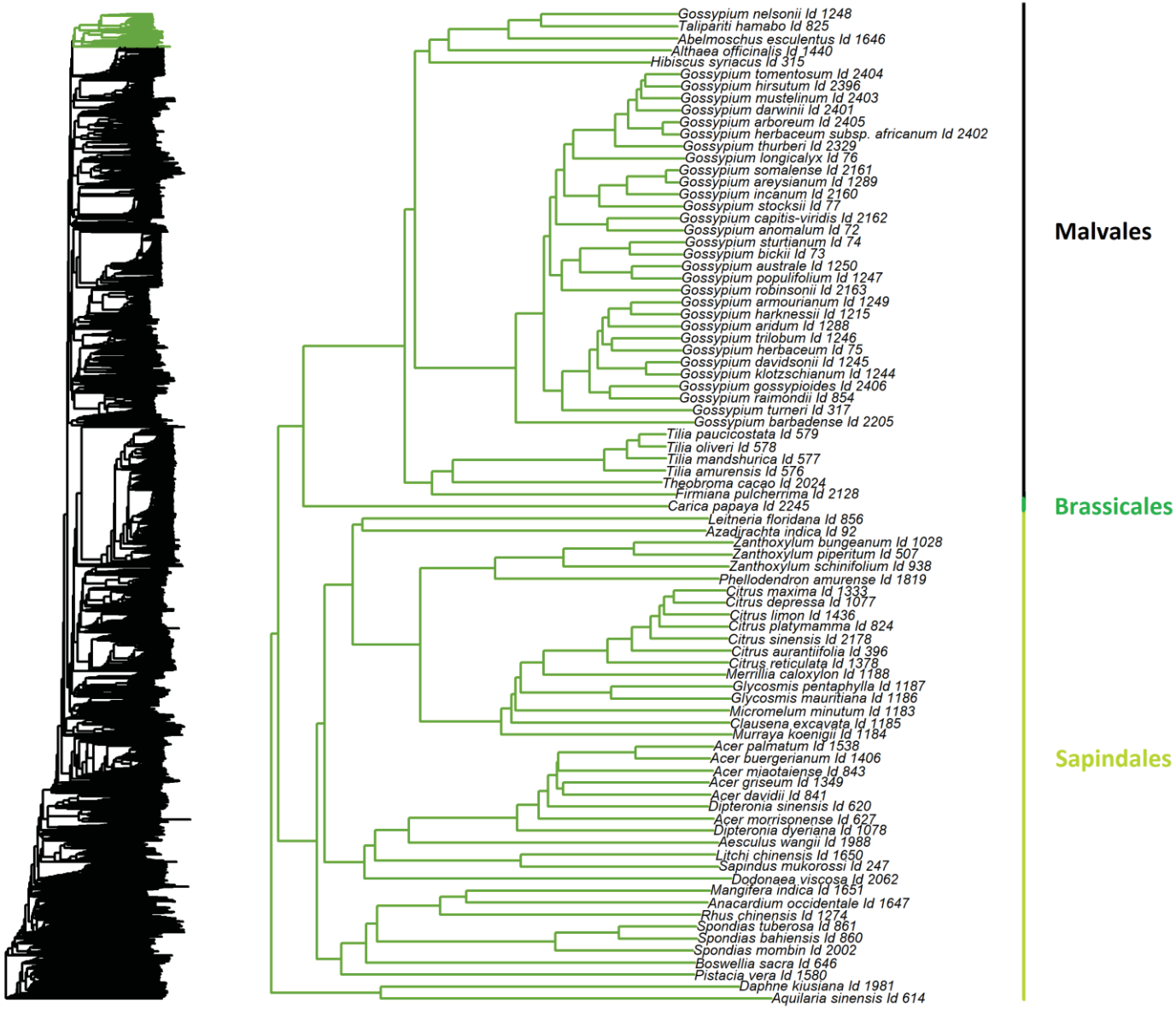
ZHAO, Y.; LU, D.; HAN, R.; WANG, L.; QIN, P. The complete chloroplast genome sequence of the shrubby cinquefoil Dasiphora fruticosa (Rosales: Rosaceae). **Conservation Genetics Resources**, v. 0, n. 0, p. 0, 2017. Springer Netherlands. Disponível em: <http://link.springer.com/10.1007/s12686-017-0899-6>. .

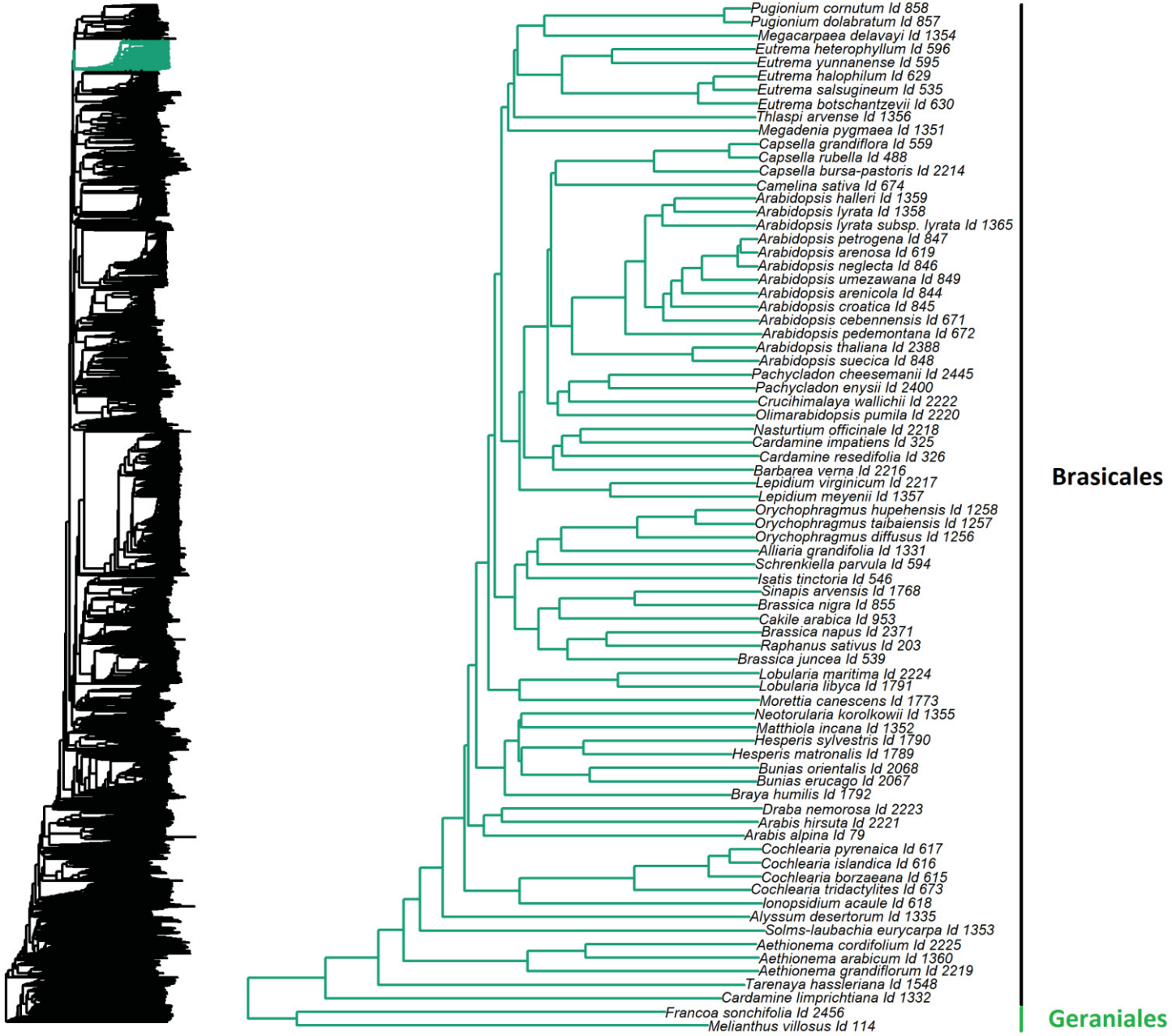
ZUCKERKANDL, E.; PAULING, L. Evolutionary divergence and convergence in proteins. **Evolving Genes and Proteins**, p. 97–166, 1965. Disponível em: <http://www.yanaiweb.com/genome/Clocks/Zuckerkandl_1965.pdf>. .

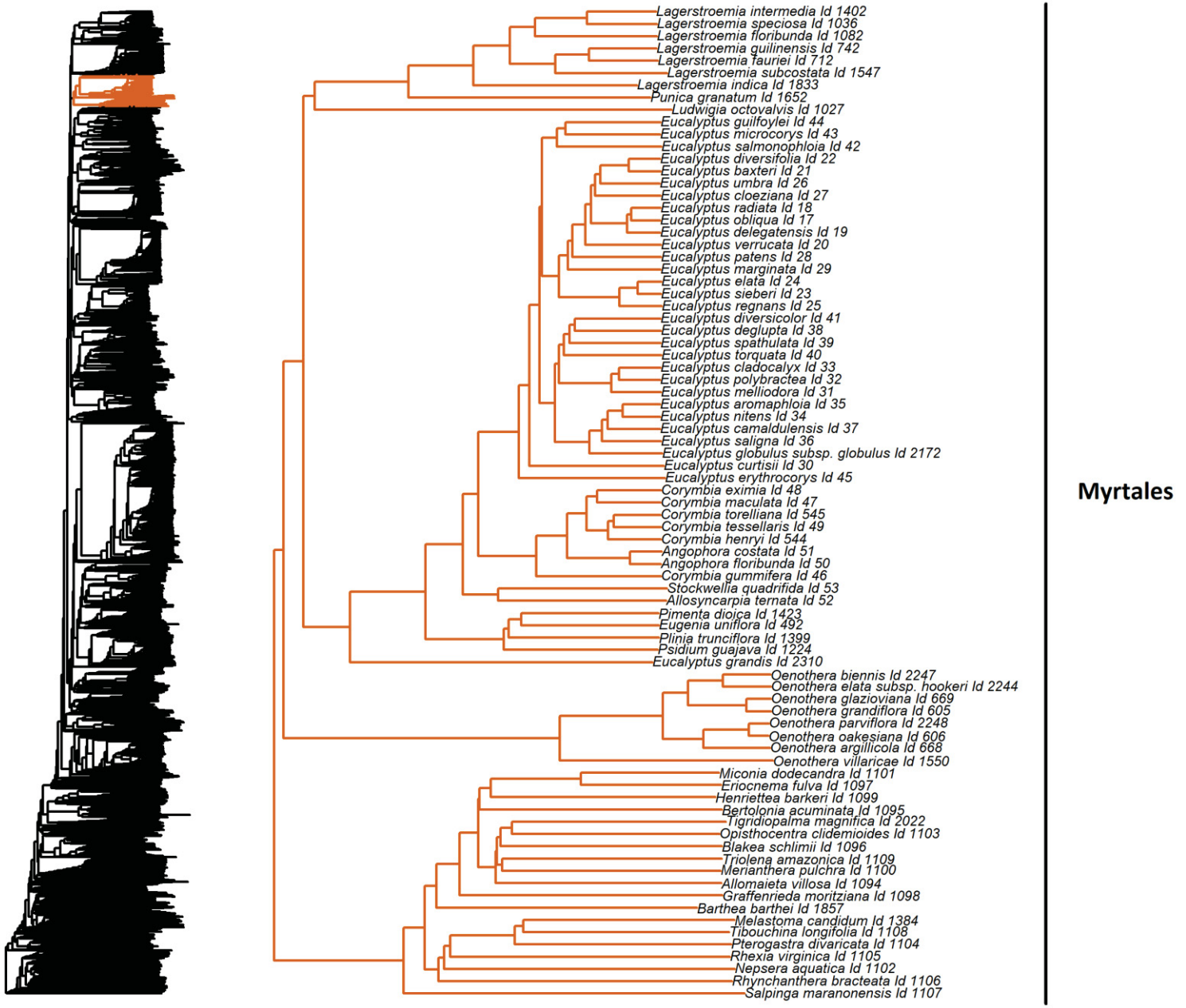
ZVERKOV, O. A.; SELIVERSTOV, A. V.; LYUBETSKY, .V A. PlastidEncoded Protein Families Specific for Narrow Taxonomic Groups of Algae and Protozoa. **Original Russian Text ©**, v. 46, n. 5, p. 717–726, 2012. Pleiades Publishing, Inc.

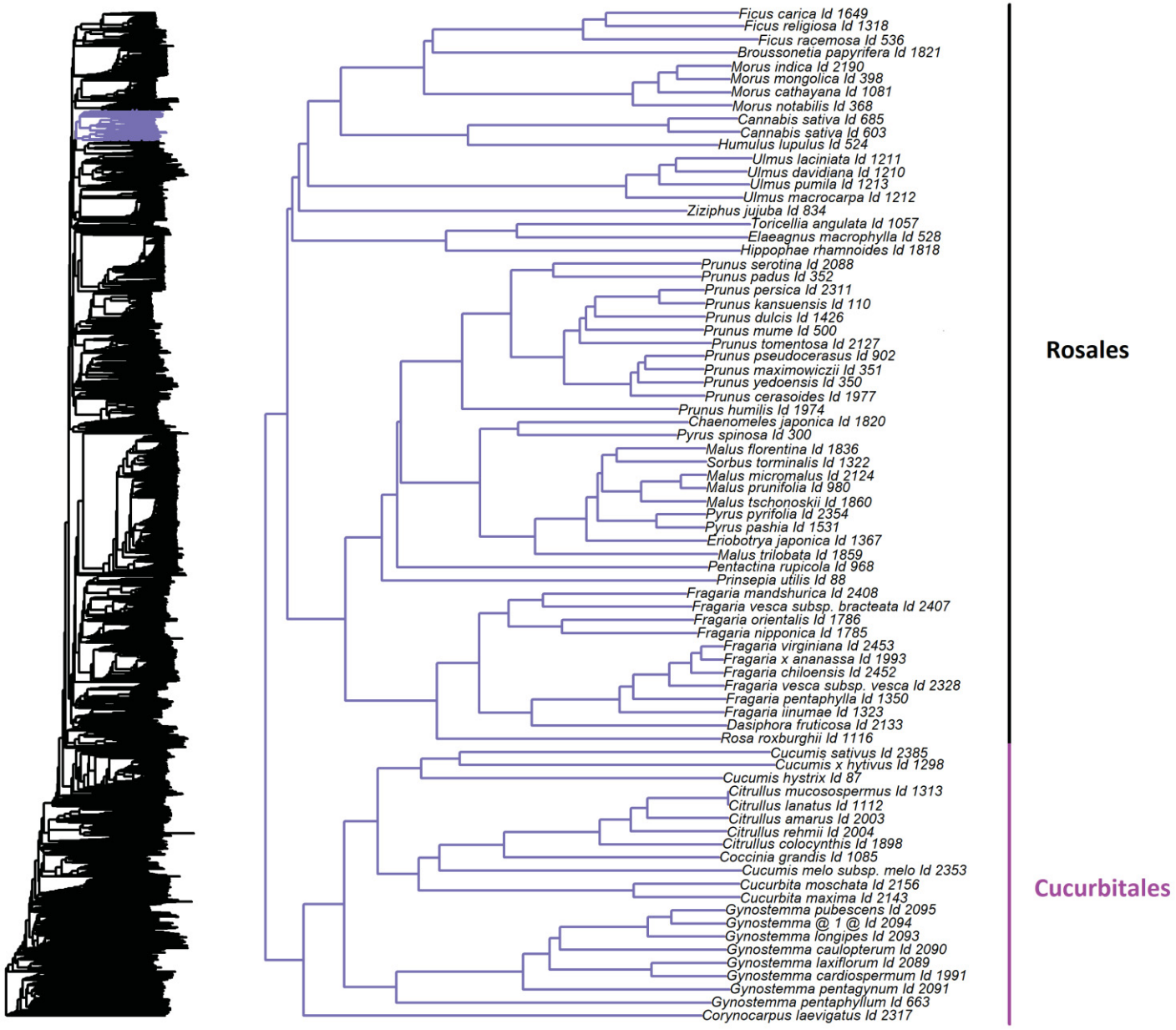
ZVERKOV, O. A.; SELIVERSTOV, A. V.; LYUBETSKY, V. A. A database of plastid protein families from red algae and apicomplexa and expression regulation of the moeB Gene. **BioMed Research International**, v. 2015, 2015.

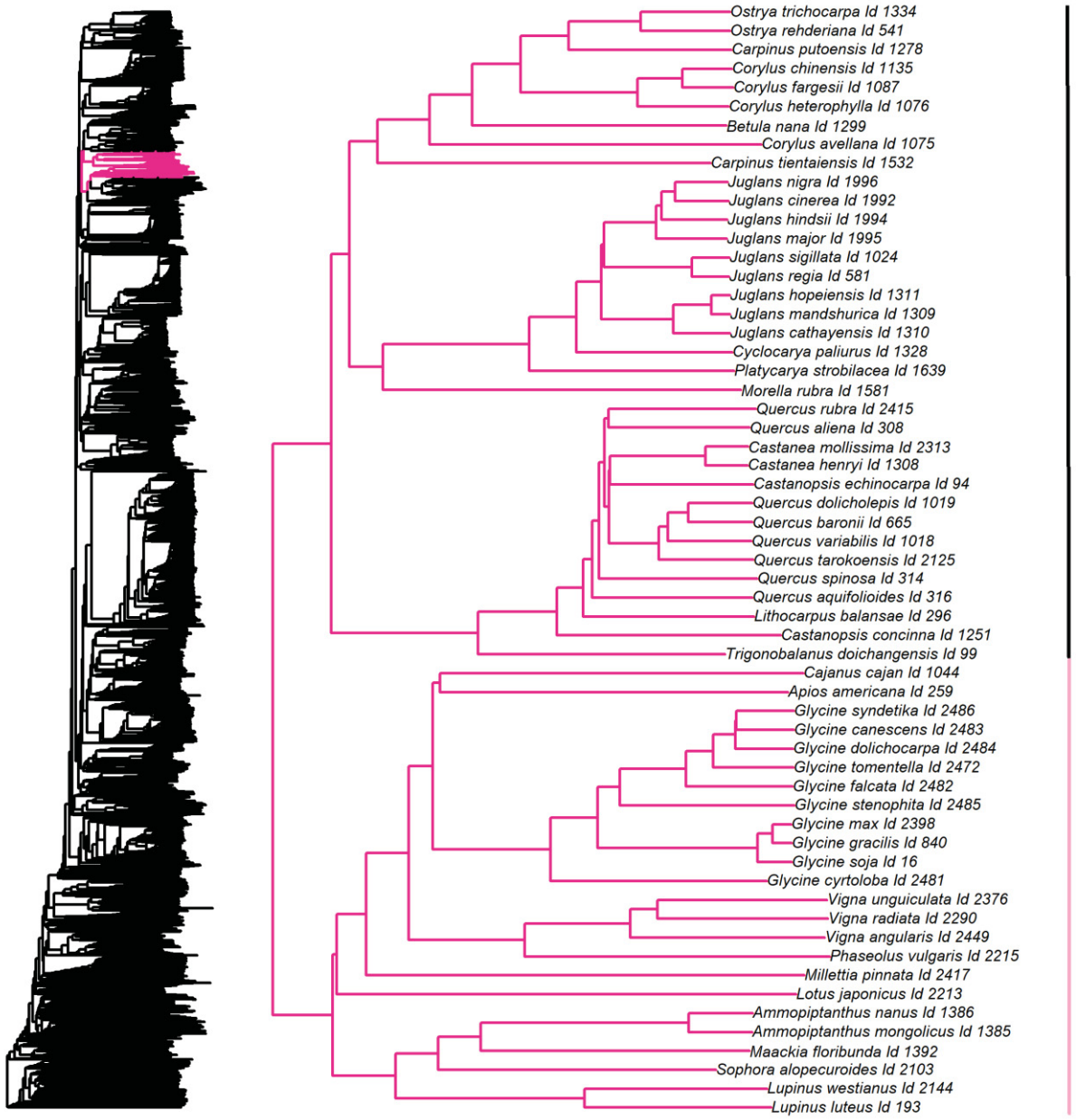
APÊNDICE I : FILOGENIA GLOBAL DE PLASTÍDIOS COMPLETA

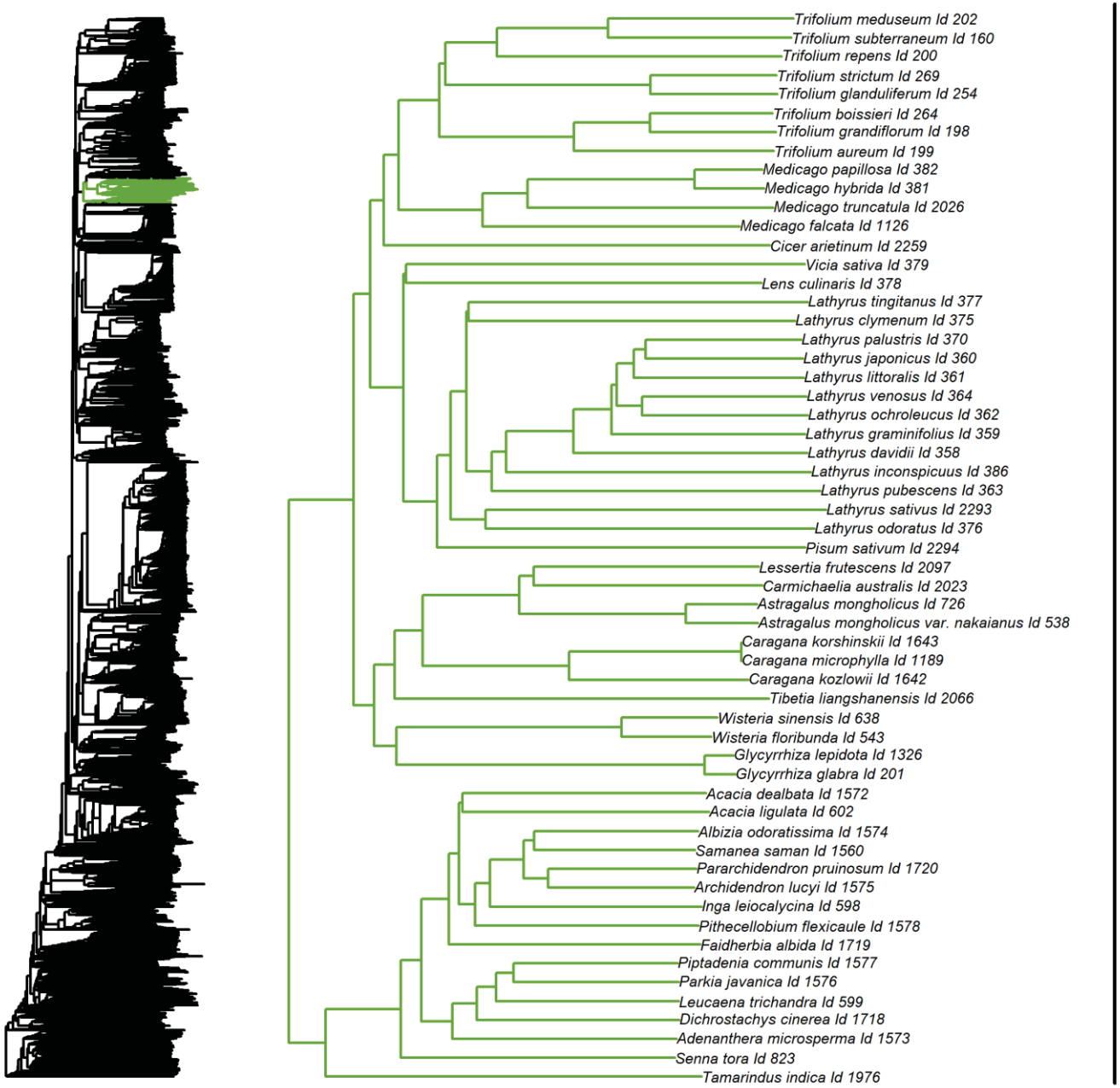


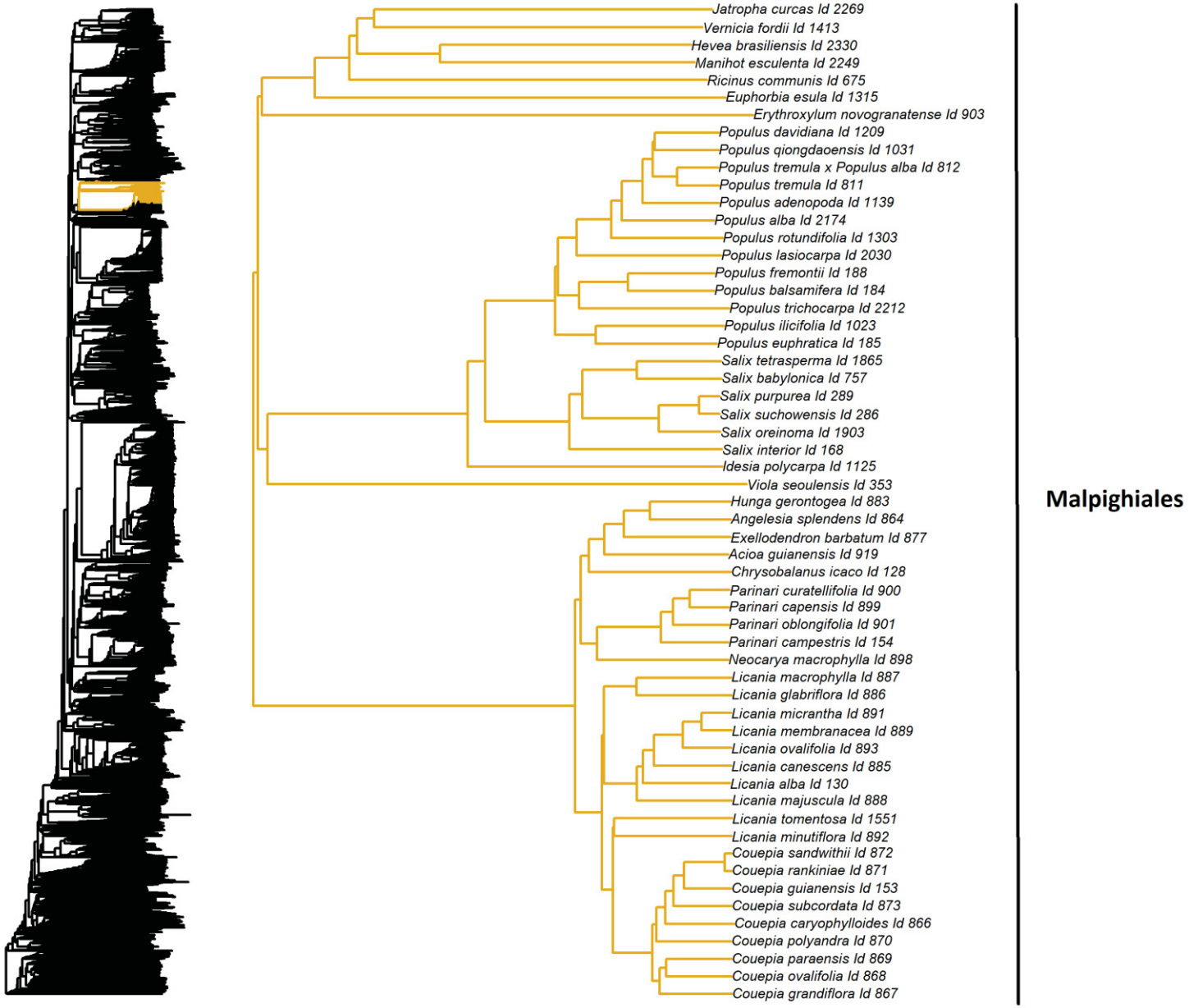

Pugionium cornutum Id 858
Pugionium dolabratum Id 857
Megacarpaea delavayi Id 1354
Eutrema heterophyllum Id 596
Eutrema yunnanense Id 595
Eutrema halophilum Id 629
Eutrema salsugineum Id 535
Eutrema botschantzevii Id 630
Thlaspi arvense Id 1356
Megadenia pygmaea Id 1351
Capsella grandiflora Id 559
Capsella rubella Id 488
Capsella bursa-pastoris Id 2214
Camelina sativa Id 674
Arabidopsis halleri Id 1359
Arabidopsis lyrata Id 1358
Arabidopsis lyrata subsp. lyrata Id 1365
Arabidopsis petrogena Id 847
Arabidopsis arenosa Id 619
Arabidopsis neglecta Id 846
Arabidopsis umezawana Id 849
Arabidopsis arenicola Id 844
Arabidopsis croatica Id 845
Arabidopsis cebennensis Id 671
Arabidopsis pedemontana Id 672
Arabidopsis thaliana Id 2388
Arabidopsis suecica Id 848
Pachycladon cheesemanii Id 2445
Pachycladon enysii Id 2400
Crucihimalaya wallichii Id 2222
Olimarabidopsis pumila Id 2220
Nasturtium officinale Id 2218
Cardamine impatiens Id 325
Cardamine resedifolia Id 326
Barbarea verna Id 2216
Lepidium virginicum Id 2217
Lepidium meyenii Id 1357
Orychophragmus hupehensis Id 1258
Orychophragmus taibaiensis Id 1257
Orychophragmus diffusus Id 1256
Alliaria grandifolia Id 1331
Schrenkiella parvula Id 594
Isatis tinctoria Id 546
Sinapis arvensis Id 1768
Brassica nigra Id 855
Cakile arabica Id 953
Brassica napus Id 2371
Raphanus sativus Id 203
Brassica juncea Id 539
Lobularia maritima Id 2224
Lobularia libyca Id 1791
Morettia canescens Id 1773
Neotorularia korolkowii Id 1355
Matthiola incana Id 1352
Hesperis sylvestris Id 1790
Hesperis matronalis Id 1789
Bunias orientalis Id 2068
Bunias erucago Id 2067
Braya humilis Id 1792
Draba nemorosa Id 2223
Arabis hirsuta Id 2221
Arabis alpina Id 79
Cochlearia pyrenaica Id 617
Cochlearia islandica Id 616
Cochlearia borzaeana Id 615
Cochlearia tridactylites Id 673
Ionopsidium acaule Id 618
Alyssum desertorum Id 1335
Solms-laubachia eurycarpa Id 1353
Aethionema cordifolium Id 2225
Aethionema arabicum Id 1360
Aethionema grandiflorum Id 2219
Tarenaya hassleriana Id 1548
Cardamine limprichtiana Id 1332
Francoa sonchifolia Id 2456
Melianthus villosus Id 114
Brasicales
Geraniales

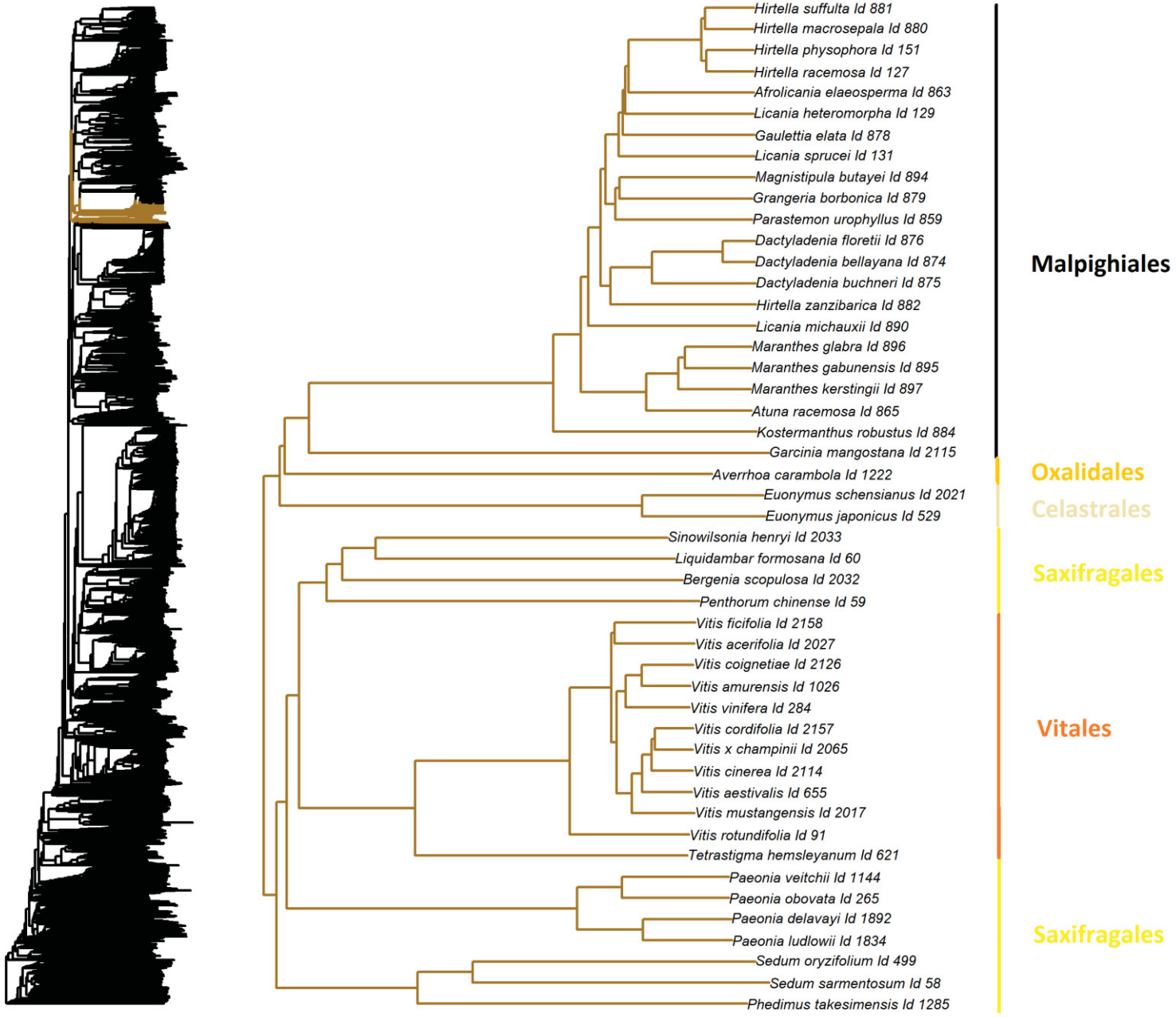

Lagerstroemia intermedia Id 1402
Lagerstroemia speciosa Id 1036
Lagerstroemia floribunda Id 1082
Lagerstroemia guilinensis Id 742
Lagerstroemia fauriei Id 712
Lagerstroemia subcostata Id 1547
Lagerstroemia indica Id 1833
Punica granatum Id 1652
Ludwigia octovalvis Id 1027
Eucalyptus guilfoylei Id 44
Eucalyptus microcorys Id 43
Eucalyptus salmonophloia Id 42
Eucalyptus diversifolia Id 22
Eucalyptus baxteri Id 21
Eucalyptus umbra Id 26
Eucalyptus cloeziana Id 27
Eucalyptus radiata Id 18
Eucalyptus obliqua Id 17
Eucalyptus delegatensis Id 19
Eucalyptus verrucata Id 20
Eucalyptus patens Id 28
Eucalyptus marginata Id 29
Eucalyptus elata Id 24
Eucalyptus sieberi Id 23
Eucalyptus regnans Id 25
Eucalyptus diversicolor Id 41
Eucalyptus deglupta Id 38
Eucalyptus spathulata Id 39
Eucalyptus torquata Id 40
Eucalyptus cladocalyx Id 33
Eucalyptus polybractea Id 32
Eucalyptus melliodora Id 31
Eucalyptus aromaphloia Id 35
Eucalyptus nitens Id 34
Eucalyptus camaldulensis Id 37
Eucalyptus saligna Id 36
Eucalyptus globulus subsp. globulus Id 2172
Eucalyptus curtisii Id 30
Eucalyptus erythrocorys Id 45
Corymbia eximia Id 48
Corymbia maculata Id 47
Corymbia torelliana Id 545
Corymbia tessellaris Id 49
Corymbia henryi Id 544
Angophora costata Id 51
Angophora floribunda Id 50
Corymbia gummifera Id 46
Stockwellia quadrifida Id 53
Allosyncarpia ternata Id 52
Pimenta dioica Id 1423
Eugenia uniflora Id 492
Plinia trunciflora Id 1399
Psidium guajava Id 1224
Eucalyptus grandis Id 2310
Oenothera biennis Id 2247
Oenothera elata subsp. hookeri Id 2244
Oenothera glazioviana Id 669
Oenothera grandiflora Id 605
Oenothera parviflora Id 2248
Oenothera oakesiana Id 606
Oenothera argillicola Id 668
Oenothera villaricae Id 1550
Miconia dodecandra Id 1101
Eriocnema fulva Id 1097
Henriettea barkeri Id 1099
Bertolonia acuminata Id 1095
Tigridiopalma magnifica Id 2022
Opisthocentra clidemioides Id 1103
Blakea schlimii Id 1096
Triolena amazonica Id 1109
Merianthera pulchra Id 1100
Allomaieta villosa Id 1094
Graffenrieda moritziana Id 1098
Barthea barthei Id 1857
Melastoma candidum Id 1384
Tibouchina longifolia Id 1108
Pterogastra divaricata Id 1104
Rhexia virginica Id 1105
Nepsera aquatica Id 1102
Rhynchanthera bracteata Id 1106
Salpinga maranonensis Id 1107
Myrtales

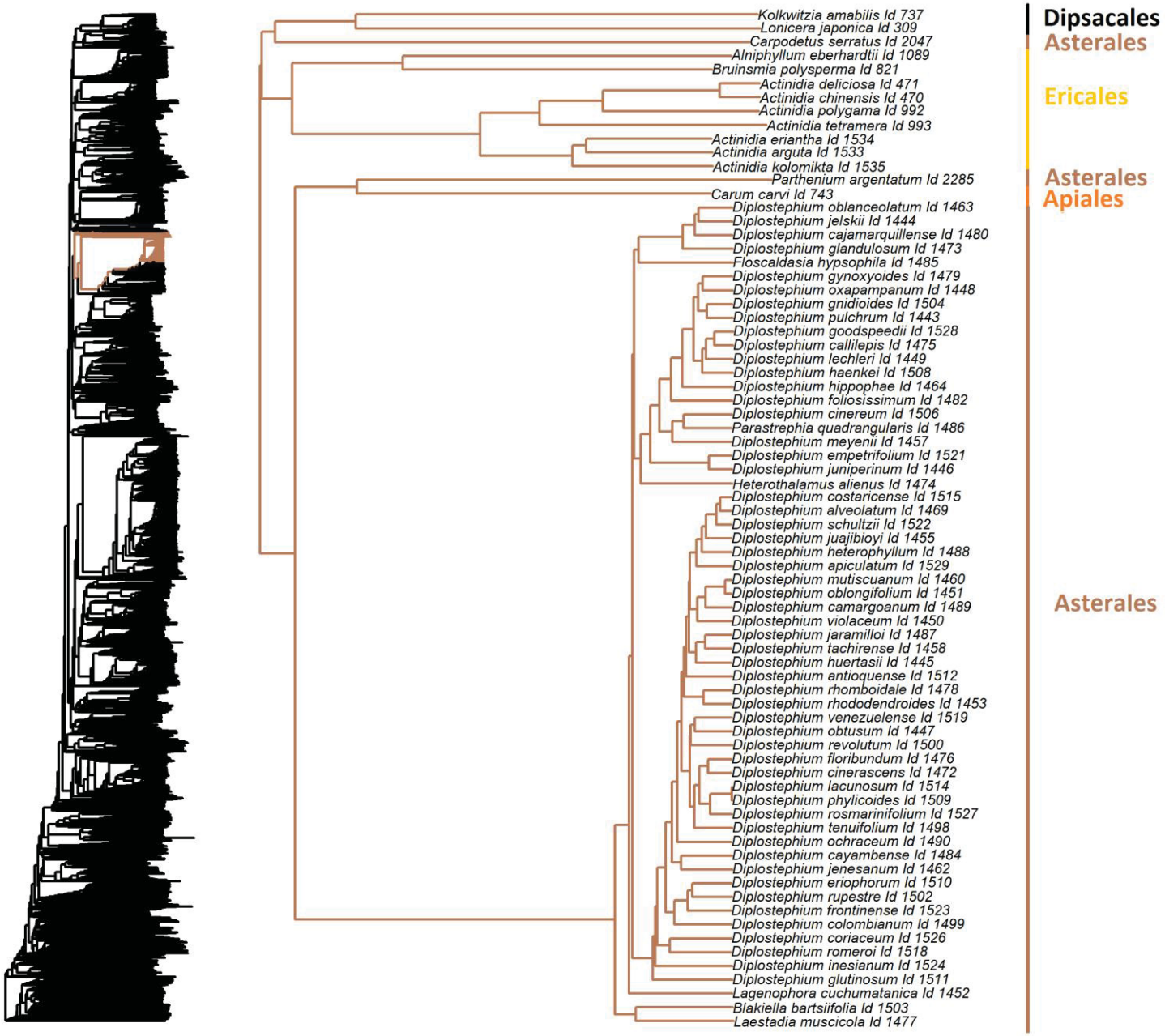

Ficus carica Id 1649
Ficus religiosa Id 1318
Ficus racemosa Id 536
Broussonetia papyrifera Id 1821
Morus indica Id 2190
Morus mongolica Id 398
Morus cathayana Id 1081
Morus notabilis Id 368
Cannabis sativa Id 685
Cannabis sativa Id 603
Humulus lupulus Id 524
Ulmus laciniata Id 1211
Ulmus davidiana Id 1210
Ulmus pumila Id 1213
Ulmus macrocarpa Id 1212
Ziziphus jujuba Id 834
Toricellia angulata Id 1057
Elaeagnus macrophylla Id 528
Hippophae rhamnoides Id 1818
Prunus serotina Id 2088
Prunus padus Id 352
Prunus persica Id 2311
Prunus kansuensis Id 110
Prunus dulcis Id 1426
Prunus mume Id 500
Prunus tomentosa Id 2127
Prunus pseudocerasus Id 902
Prunus maximowiczii Id 351
Prunus yedoensis Id 350
Prunus cerasoides Id 1977
Prunus humilis Id 1974
Chaenomeles japonica Id 1820
Pyrus spinosa Id 300
Malus florentina Id 1836
Sorbus torminalis Id 1322
Malus micromalus Id 2124
Malus prunifolia Id 980
Malus tschonoskii Id 1860
Pyrus pyrifolia Id 2354
Pyrus pashia Id 1531
Eriobotrya japonica Id 1367
Malus trilobata Id 1859
Pentactina rupicola Id 968
Prinsepia utilis Id 88
Fragaria mandshurica Id 2408
Fragaria vesca subsp. bracteata Id 2407
Fragaria orientalis Id 1786
Fragaria nipponica Id 1785
Fragaria virginiana Id 2453
Fragaria x ananassa Id 1993
Fragaria chiloensis Id 2452
Fragaria vesca subsp. vesca Id 2328
Fragaria pentaphylla Id 1350
Fragaria iinumae Id 1323
Dasiphora fruticosa Id 2133
Rosa roxburghii Id 1116
Cucumis sativus Id 2385
Cucumis x hytivus Id 1298
Cucumis hystrix Id 87
Citrullus mucosospermus Id 1313
Citrullus lanatus Id 1112
Citrullus amarus Id 2003
Citrullus rehmii Id 2004
Citrullus colocynthis Id 1898
Coccinia grandis Id 1085
Cucumis melo subsp. melo Id 2353
Cucurbita moschata Id 2156
Cucurbita maxima Id 2143
Gynostemma pubescens Id 2095
Gynostemma @ 1 @ Id 2094
Gynostemma longipes Id 2093
Gynostemma caulopterum Id 2090
Gynostemma laxiflorum Id 2089
Gynostemma cardiospermum Id 1991
Gynostemma pentagynum Id 2091
Gynostemma pentaphyllum Id 663
Corynocarpus laevigatus Id 2317
Rosales
Cucurbitales

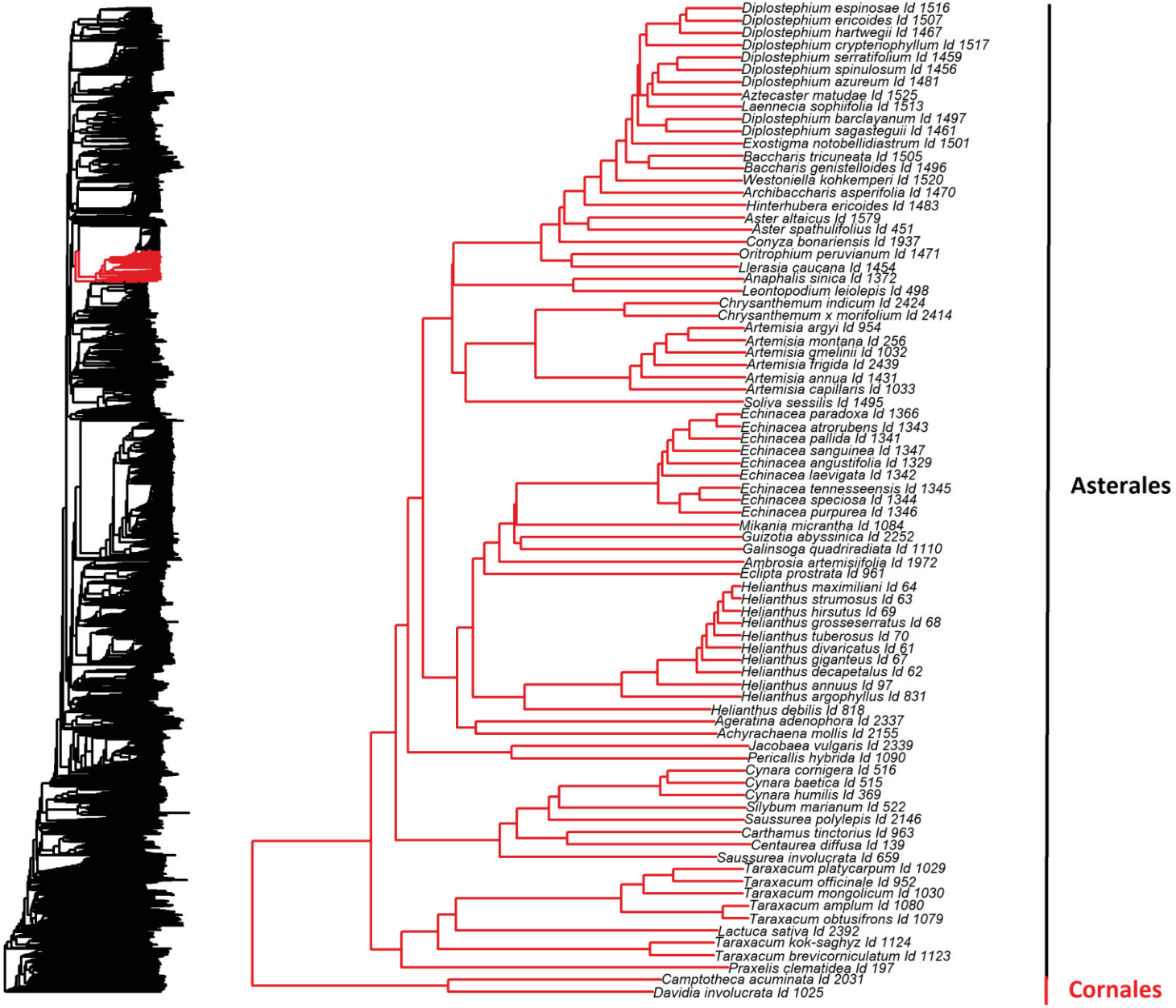

Ostrya trichocarpa Id 1334
Ostrya rehderiana Id 541
Carpinus putoensis Id 1278
Corylus chinensis Id 1135
Corylus fargesii Id 1087
Corylus heterophylla Id 1076
Betula nana Id 1299
Corylus avellana Id 1075
Carpinus tientaiensis Id 1532
Juglans nigra Id 1996
Juglans cinerea Id 1992
Juglans hindsii Id 1994
Juglans major Id 1995
Juglans sigillata Id 1024
Juglans regia Id 581
Juglans hopeiensis Id 1311
Juglans mandshurica Id 1309
Juglans cathayensis Id 1310
Cyclocarya paliurus Id 1328
Platycarya strobilacea Id 1639
Morella rubra Id 1581
Quercus rubra Id 2415
Quercus aliena Id 308
Castanea mollissima Id 2313
Castanea henryi Id 1308
Castanopsis echinocarpa Id 94
Quercus dolicholepis Id 1019
Quercus baronii Id 665
Quercus variabilis Id 1018
Quercus tarokoensis Id 2125
Quercus spinosa Id 314
Quercus aquifolioides Id 316
Lithocarpus balansae Id 296
Castanopsis concinna Id 1251
Trigonobalanus doichangensis Id 99
Cajanus cajan Id 1044
Apios americana Id 259
Glycine syndetika Id 2486
Glycine canescens Id 2483
Glycine dolichocarpa Id 2484
Glycine tomentella Id 2472
Glycine falcata Id 2482
Glycine stenophita Id 2485
Glycine max Id 2398
Glycine gracilis Id 840
Glycine soja Id 16
Glycine cyrtoloba Id 2481
Vigna unguiculata Id 2376
Vigna radiata Id 2290
Vigna angularis Id 2449
Phaseolus vulgaris Id 2215
Millettia pinnata Id 2417
Lotus japonicus Id 2213
Ammopiptanthus nanus Id 1386
Ammopiptanthus mongolicus Id 1385
Maackia floribunda Id 1392
Sophora alopecuroides Id 2103
Lupinus westianus Id 2144
Lupinus luteus Id 193
Fagales
Fabales

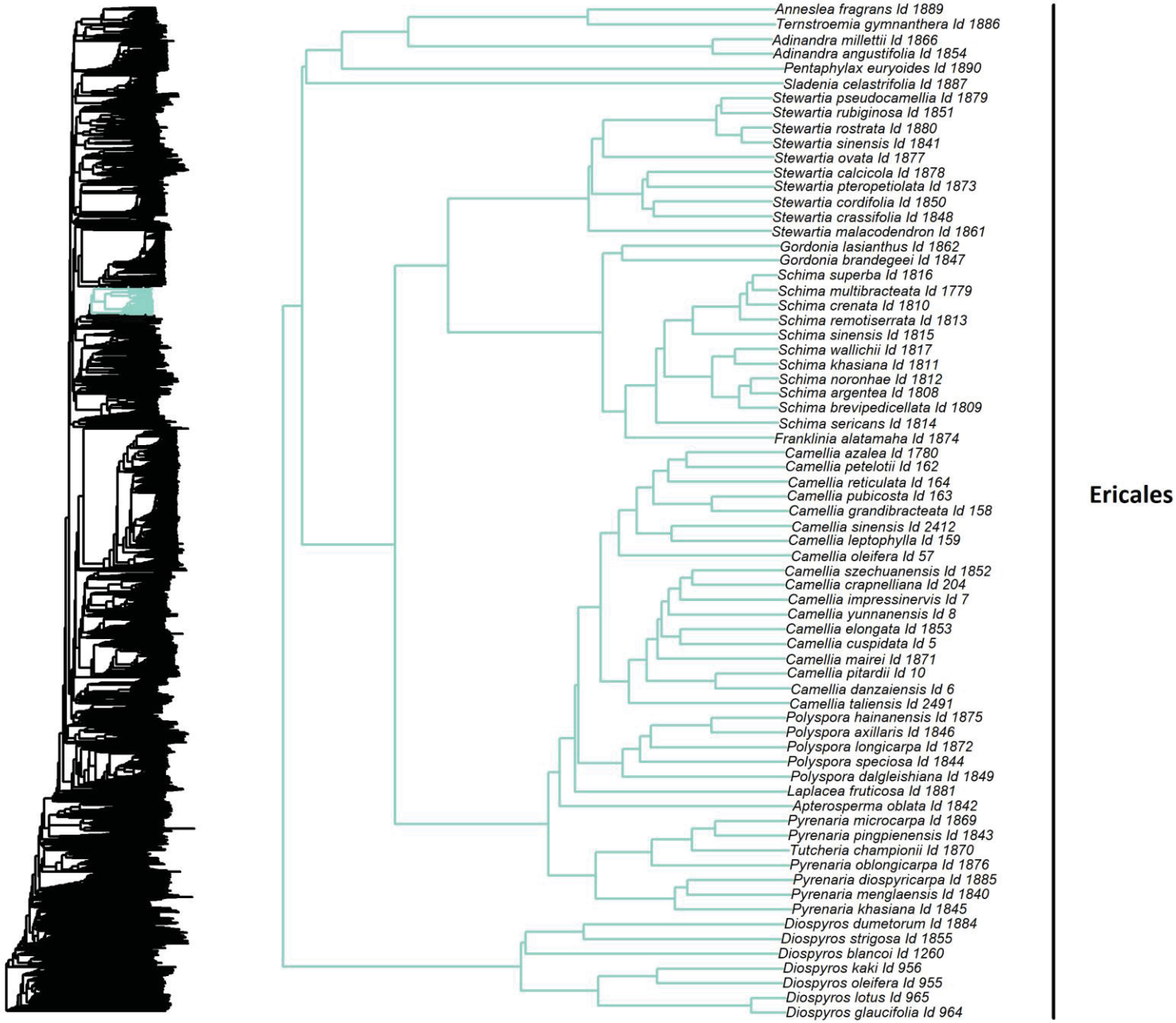

Trifolium meduseum Id 202
Trifolium subterraneum Id 160
Trifolium repens Id 200
Trifolium strictum Id 269
Trifolium glanduliferum Id 254
Trifolium boissieri Id 264
Trifolium grandiflorum Id 198
Trifolium aureum Id 199
Medicago papillosa Id 382
Medicago hybrida Id 381
Medicago truncatula Id 2026
Medicago falcata Id 1126
Cicer arietinum Id 2259
Vicia sativa Id 379
Lens culinaris Id 378
Lathyrus tingitanus Id 377
Lathyrus clymenum Id 375
Lathyrus palustris Id 370
Lathyrus japonicus Id 360
Lathyrus littoralis Id 361
Lathyrus venosus Id 364
Lathyrus ochroleucus Id 362
Lathyrus graminifolius Id 359
Lathyrus davidii Id 358
Lathyrus inconspicuus Id 386
Lathyrus pubescens Id 363
Lathyrus sativus Id 2293
Lathyrus odoratus Id 376
Pisum sativum Id 2294
Lessertia frutescens Id 2097
Carmichaelia australis Id 2023
Astragalus mongholicus Id 726
Astragalus mongholicus var. nakaianus Id 538
Caragana korshinskii Id 1643
Caragana microphylla Id 1189
Caragana kozlowii Id 1642
Tibetia liangshanensis Id 2066
Wisteria sinensis Id 638
Wisteria floribunda Id 543
Glycyrrhiza lepidota Id 1326
Glycyrrhiza glabra Id 201
Acacia dealbata Id 1572
Acacia ligulata Id 602
Albizia odoratissima Id 1574
Samanea saman Id 1560
Pararchidendron pruinosum Id 1720
Archidendron lucyi Id 1575
Inga leiocalycina Id 598
Pithecellobium flexicaule Id 1578
Faidherbia albida Id 1719
Piptadenia communis Id 1577
Parkia javanica Id 1576
Leucaena trichandra Id 599
Dichrostachys cinerea Id 1718
Adenanthera microsperma Id 1573
Senna tora Id 823
Tamarindus indica Id 1976
Fabales

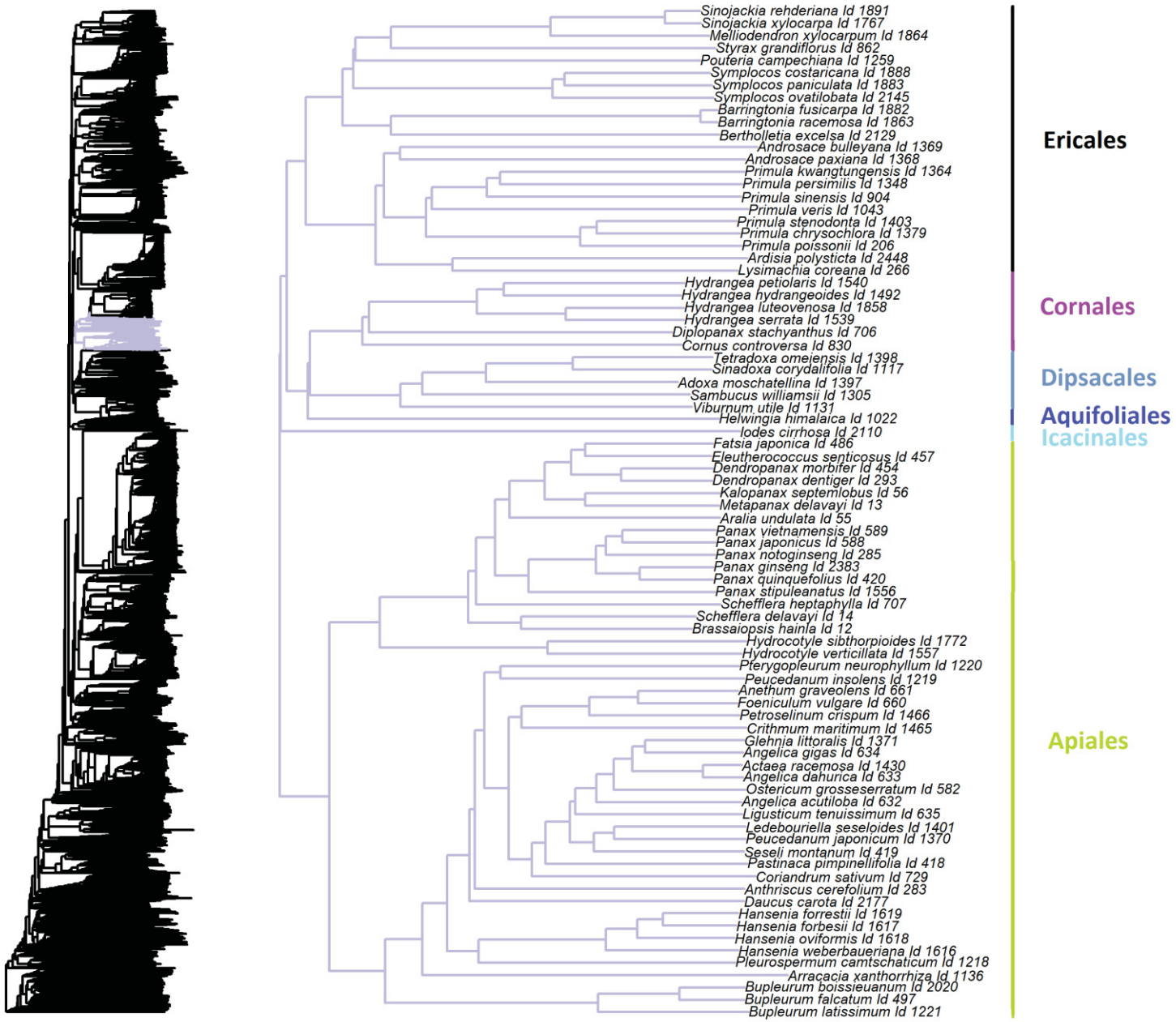

Jatropha curcas Id 2269
Vernicia fordii Id 1413
Hevea brasiliensis Id 2330
Manihot esculenta Id 2249
Ricinus communis Id 675
Euphorbia esula Id 1315
Erythroxylum novogranatense Id 903
Populus davidiana Id 1209
Populus qiongdaoensis Id 1031
Populus tremula x Populus alba Id 812
Populus tremula Id 811
Populus adenopoda Id 1139
Populus alba Id 2174
Populus rotundifolia Id 1303
Populus lasiocarpa Id 2030
Populus fremontii Id 188
Populus balsamifera Id 184
Populus trichocarpa Id 2212
Populus ilicifolia Id 1023
Populus euphratica Id 185
Salix tetrasperma Id 1865
Salix babylonica Id 757
Salix purpurea Id 289
Salix suchowensis Id 286
Salix oreinoma Id 1903
Salix interior Id 168
Idesia polycarpa Id 1125
Viola seoulensis Id 353
Hunga gerontogea Id 883
Angelesia splendens Id 864
Exellodendron barbatum Id 877
Acioa guianensis Id 919
Chrysobalanus icaco Id 128
Parinari curatellifolia Id 900
Parinari capensis Id 899
Parinari oblongifolia Id 901
Parinari campestris Id 154
Neocarya macrophylla Id 898
Licania macrophylla Id 887
Licania glabriflora Id 886
Licania micrantha Id 891
Licania membranacea Id 889
Licania ovalifolia Id 893
Licania canescens Id 885
Licania alba Id 130
Licania majuscula Id 888
Licania tomentosa Id 1551
Licania minutiflora Id 892
Couepia sandwithii Id 872
Couepia rankiniae Id 871
Couepia guianensis Id 153
Couepia subcordata Id 873
Couepia caryophylloides Id 866
Couepia polyandra Id 870
Couepia paraensis Id 869
Couepia ovalifolia Id 868
Couepia grandiflora Id 867
Malpighiales

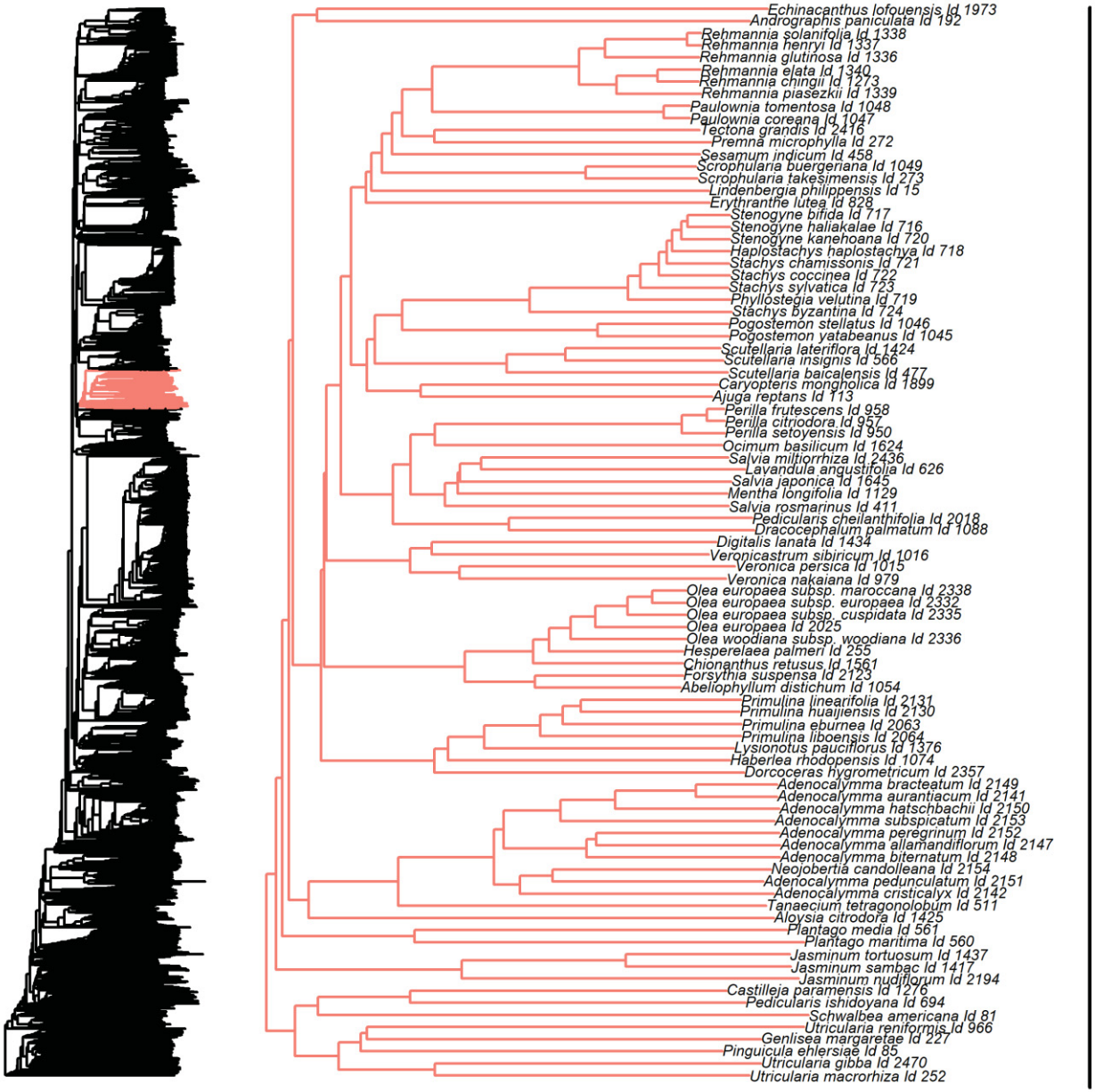

Hirtella suffulta Id 881
Hirtella macrosepala Id 880
Hirtella physophora Id 151
Hirtella racemosa Id 127
Afrolicania elaeosperma Id 863
Licania heteromorpha Id 129
Gaulettia elata Id 878
Licania sprucei Id 131
Magnistipula butayei Id 894
Grangeria borbonica Id 879
Parastemon urophyllus Id 859
Dactyladenia floretii Id 876
Dactyladenia bellayana Id 874
Dactyladenia buchneri Id 875
Hirtella zanzibarica Id 882
Licania michauxii Id 890
Maranthes glabra Id 896
Maranthes gabunensis Id 895
Maranthes kerstingii Id 897
Atuna racemosa Id 865
Kostermanthus robustus Id 884
Garcinia mangostana Id 2115
Averrhoa carambola Id 1222
Euonymus schensianus Id 2021
Euonymus japonicus Id 529
Sinowilsonia henryi Id 2033
Liquidambar formosana Id 60
Bergenia scopulosa Id 2032
Penthorum chinense Id 59
Vitis ficifolia Id 2158
Vitis acerifolia Id 2027
Vitis coignetiae Id 2126
Vitis amurensis Id 1026
Vitis vinifera Id 284
Vitis cordifolia Id 2157
Vitis x champinii Id 2065
Vitis cinerea Id 2114
Vitis aestivalis Id 655
Vitis mustangensis Id 2017
Vitis rotundifolia Id 91
Tetrastigma hemsleyanum Id 621
Paeonia veitchii Id 1144
Paeonia obovata Id 265
Paeonia delavayi Id 1892
Paeonia ludlowii Id 1834
Sedum oryzifolium Id 499
Sedum sarmentosum Id 58
Phedimus takesimensis Id 1285
Malpighiales
Oxalidales
Celastrales
Saxifragales
Vitales
Saxifragales

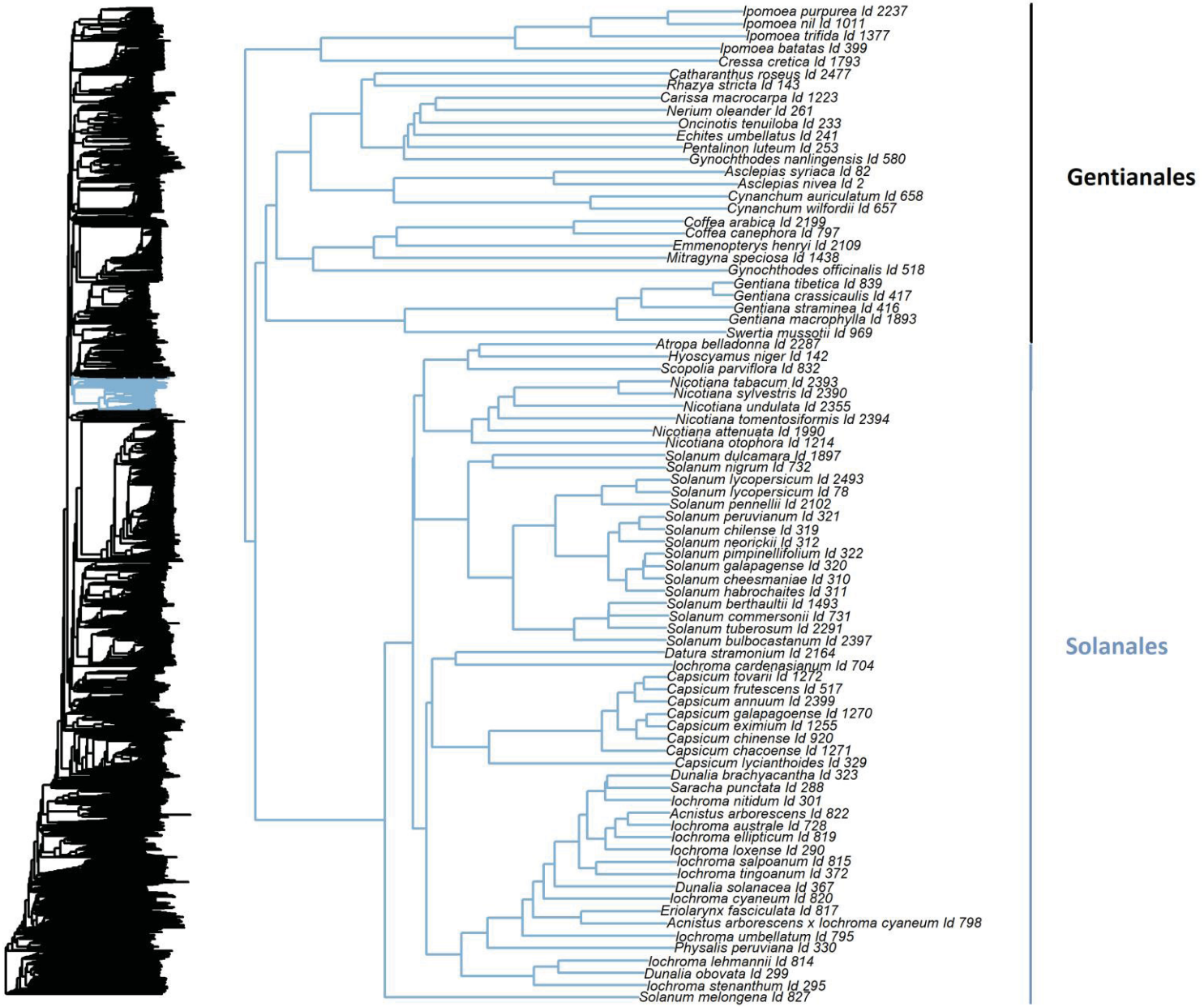

Kolkwitzia amabilis Id 737
Lonicera japonica Id 309
Carpodetus serratus Id 2047
Alniphyllum eberhardtii Id 1089
Bruinsmia polysperma Id 821
Actinidia deliciosa Id 471
Actinidia chinensis Id 470
Actinidia polygama Id 992
Actinidia tetramera Id 993
Actinidia eriantha Id 1534
Actinidia arguta Id 1533
Actinidia kolomikta Id 1535
Parthenium argentatum Id 2285
Carum carvi Id 743
Diplostephium oblanceolatum Id 1463
Diplostephium jelskii Id 1444
Diplostephium cajamarquillense Id 1480
Diplostephium glandulosum Id 1473
Floscaldasia hypsophila Id 1485
Diplostephium gynoxyoides Id 1479
Diplostephium oxapampanum Id 1448
Diplostephium gnidioides Id 1504
Diplostephium pulchrum Id 1443
Diplostephium goodspeedii Id 1528
Diplostephium callilepis Id 1475
Diplostephium lechleri Id 1449
Diplostephium haenkei Id 1508
Diplostephium hippophae Id 1464
Diplostephium foliosissimum Id 1482
Diplostephium cinereum Id 1506
Parastrephia quadrangularis Id 1486
Diplostephium meyenii Id 1457
Diplostephium empetrifolium Id 1521
Diplostephium juniperinum Id 1446
Heterothalamus alienus Id 1474
Diplostephium costaricense Id 1515
Diplostephium alveolatum Id 1469
Diplostephium schultzii Id 1522
Diplostephium juajibioyi Id 1455
Diplostephium heterophyllum Id 1488
Diplostephium apiculatum Id 1529
Diplostephium mutiscuanum Id 1460
Diplostephium oblongifolium Id 1451
Diplostephium camargoanum Id 1489
Diplostephium violaceum Id 1450
Diplostephium jaramilloi Id 1487
Diplostephium tachirense Id 1458
Diplostephium huertasii Id 1445
Diplostephium antioquense Id 1512
Diplostephium rhomboidale Id 1478
Diplostephium rhododendroides Id 1453
Diplostephium venezuelense Id 1519
Diplostephium obtusum Id 1447
Diplostephium revolutum Id 1500
Diplostephium floribundum Id 1476
Diplostephium cinerascens Id 1472
Diplostephium lacunosum Id 1514
Diplostephium phylicoides Id 1509
Diplostephium rosmarinifolium Id 1527
Diplostephium tenuifolium Id 1498
Diplostephium ochraceum Id 1490
Diplostephium cayambense Id 1484
Diplostephium jenesanum Id 1462
Diplostephium eriophorum Id 1510
Diplostephium rupestre Id 1502
Diplostephium frontinense Id 1523
Diplostephium colombianum Id 1499
Diplostephium coriaceum Id 1526
Diplostephium romeroi Id 1518
Diplostephium inesianum Id 1524
Diplostephium glutinosum Id 1511
Lagenophora cuchumatanica Id 1452
Blakiella bartsiifolia Id 1503
Laestadia muscicola Id 1477
Dipsacales
Asterales
Ericales
Asterales
Apiales
Asterales

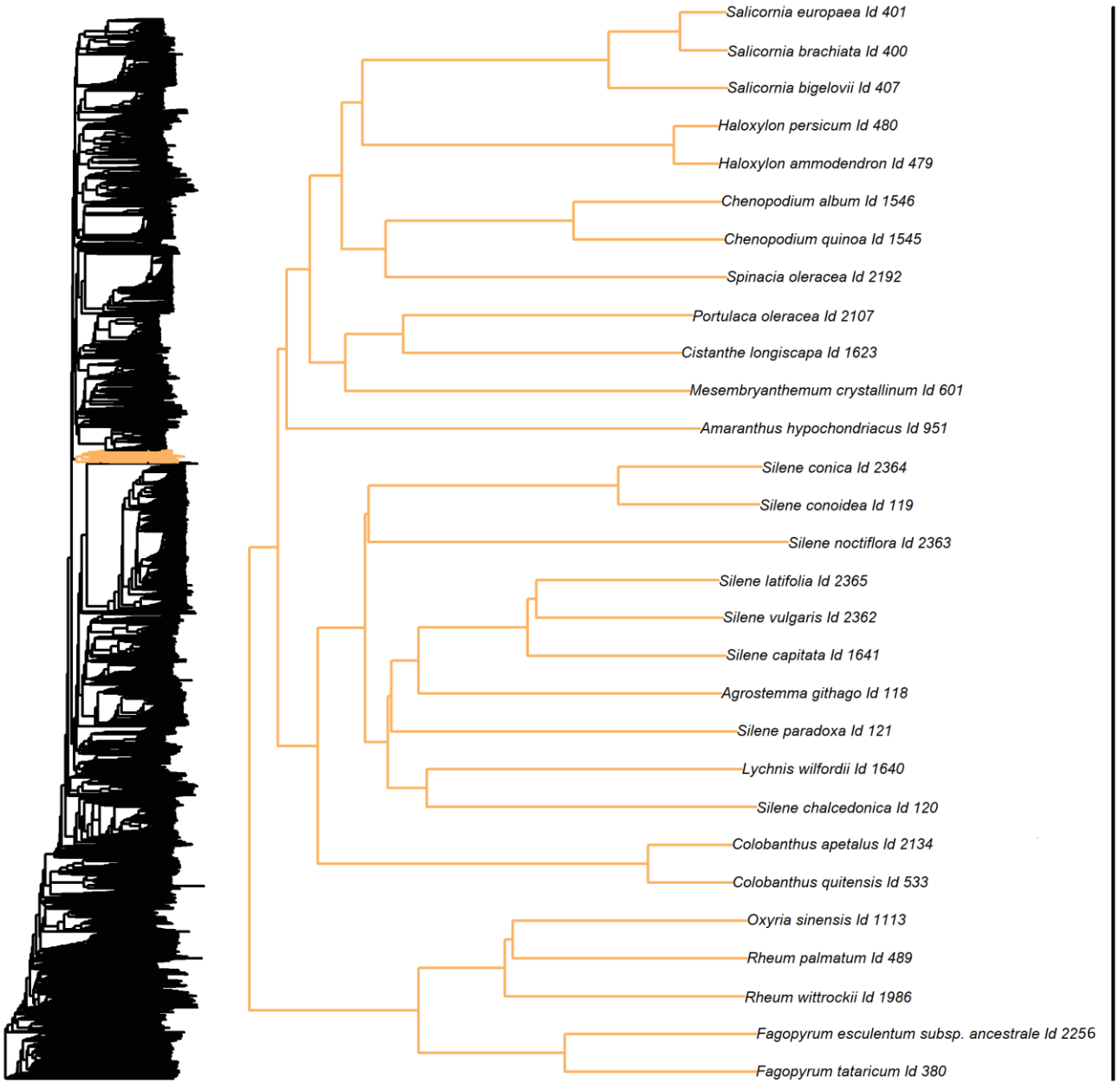

Diplostephium espinosae Id 1516
Diplostephium ericoides Id 1507
Diplostephium hartwegii Id 1467
Diplostephium crypteriophyllum Id 1517
Diplostephium serratifolium Id 1459
Diplostephium spinulosum Id 1456
Diplostephium azureum Id 1481
Aztecaster matudae Id 1525
Laennecia sophiifolia Id 1513
Diplostephium barclayanum Id 1497
Diplostephium sagasteguii Id 1461
Exostigma notobellidiastrum Id 1501
Baccharis tricuneata Id 1505
Baccharis genistelloides Id 1496
Westoniella kohkemperi Id 1520
Archibaccharis asperifolia Id 1470
Hinterhubera ericoides Id 1483
Aster altaicus Id 1579
Aster spathulifolius Id 451
Conyza bonariensis Id 1937
Oritrophium peruvianum Id 1471
Llerasia caucana Id 1454
Anaphalis sinica Id 1372
Leontopodium leiolepis Id 498
Chrysanthemum indicum Id 2424
Chrysanthemum x morifolium Id 2414
Artemisia argyi Id 954
Artemisia montana Id 256
Artemisia gmelinii Id 1032
Artemisia frigida Id 2439
Artemisia annua Id 1431
Artemisia capillaris Id 1033
Soliva sessilis Id 1495
Echinacea paradoxa Id 1366
Echinacea atrorubens Id 1343
Echinacea pallida Id 1341
Echinacea sanguinea Id 1347
Echinacea angustifolia Id 1329
Echinacea laevigata Id 1342
Echinacea tennesseensis Id 1345
Echinacea speciosa Id 1344
Echinacea purpurea Id 1346
Mikania micrantha Id 1084
Guizotia abyssinica Id 2252
Galinsoga quadriradiata Id 1110
Ambrosia artemisiifolia Id 1972
Eclipta prostrata Id 961
Helianthus maximiliani Id 64
Helianthus strumosus Id 63
Helianthus hirsutus Id 69
Helianthus grosseserratus Id 68
Helianthus tuberosus Id 70
Helianthus divaricatus Id 61
Helianthus giganteus Id 67
Helianthus decapetalus Id 62
Helianthus annuus Id 97
Helianthus argophyllus Id 831
Helianthus debilis Id 818
Ageratina adenophora Id 2337
Achyrachaena mollis Id 2155
Jacobaea vulgaris Id 2339
Pericallis hybrida Id 1090
Cynara cornigera Id 516
Cynara baetica Id 515
Cynara humilis Id 369
Silybum marianum Id 522
Saussurea polylepis Id 2146
Carthamus tinctorius Id 963
Centaurea diffusa Id 139
Saussurea involucrata Id 659
Taraxacum platycarpum Id 1029
Taraxacum officinale Id 952
Taraxacum mongolicum Id 1030
Taraxacum amplum Id 1080
Taraxacum obtusifrons Id 1079
Lactuca sativa Id 2392
Taraxacum kok-saghyz Id 1124
Taraxacum brevicorniculatum Id 1123
Praxelis clematidea Id 197
Camptotheca acuminata Id 2031
Davidia involucrata Id 1025
Asterales
Cornales

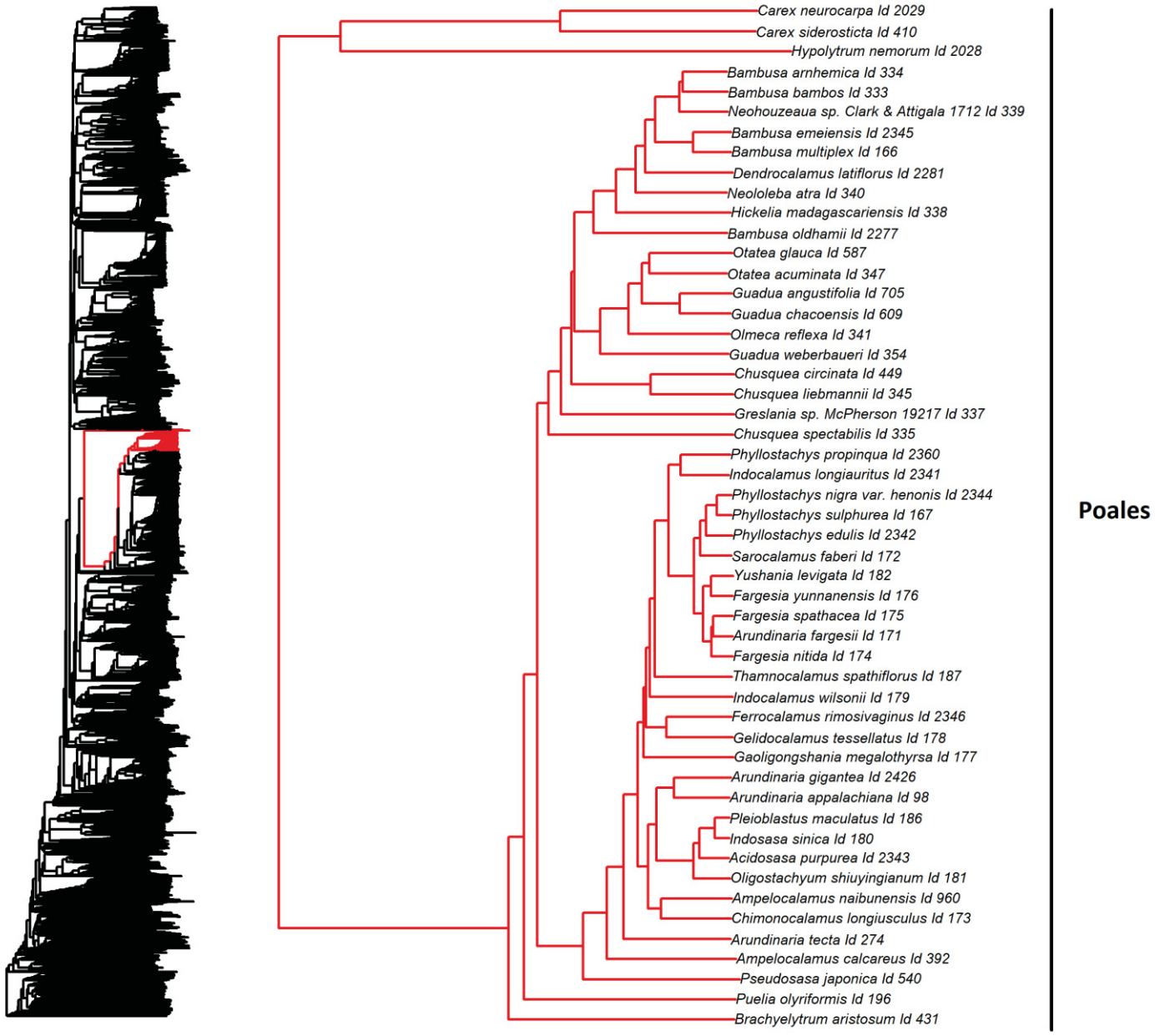

Anneslea fragrans Id 1889
Ternstroemia gymnanthera Id 1886
Adinandra millettii Id 1866
Adinandra angustifolia Id 1854
Pentaphylax euryoides Id 1890
Sladenia celastrifolia Id 1887
Stewartia pseudocamellia Id 1879
Stewartia rubiginosa Id 1851
Stewartia rostrata Id 1880
Stewartia sinensis Id 1841
Stewartia ovata Id 1877
Stewartia calcicola Id 1878
Stewartia pteropetiolata Id 1873
Stewartia cordifolia Id 1850
Stewartia crassifolia Id 1848
Stewartia malacodendron Id 1861
Gordonia lasianthus Id 1862
Gordonia brandegeei Id 1847
Schima superba Id 1816
Schima multibracteata Id 1779
Schima crenata Id 1810
Schima remotiserrata Id 1813
Schima sinensis Id 1815
Schima wallichii Id 1817
Schima khasiana Id 1811
Schima noronhae Id 1812
Schima argentea Id 1808
Schima brevipedicellata Id 1809
Schima sericans Id 1814
Franklinia alatamaha Id 1874
Camellia azalea Id 1780
Camellia petelotii Id 162
Camellia reticulata Id 164
Camellia pubicosta Id 163
Camellia grandibracteata Id 158
Camellia sinensis Id 2412
Camellia leptophylla Id 159
Camellia oleifera Id 57
Camellia szechuanensis Id 1852
Camellia crapnelliana Id 204
Camellia impressinervis Id 7
Camellia yunnanensis Id 8
Camellia elongata Id 1853
Camellia cuspidata Id 5
Camellia mairei Id 1871
Camellia pitardii Id 10
Camellia danzaiensis Id 6
Camellia taliensis Id 2491
Polyspora hainanensis Id 1875
Polyspora axillaris Id 1846
Polyspora longicarpa Id 1872
Polyspora speciosa Id 1844
Polyspora dalgleishiana Id 1849
Laplacea fruticosa Id 1881
Apterosperma oblata Id 1842
Pyrenaria microcarpa Id 1869
Pyrenaria pingpienensis Id 1843
Tutcheria championii Id 1870
Pyrenaria oblongicarpa Id 1876
Pyrenaria diospyricarpa Id 1885
Pyrenaria menglaensis Id 1840
Pyrenaria khasiana Id 1845
Diospyros dumetorum Id 1884
Diospyros strigosa Id 1855
Diospyros blancoi Id 1260
Diospyros kaki Id 956
Diospyros oleifera Id 955
Diospyros lotus Id 965
Diospyros glaucifolia Id 964
Ericales

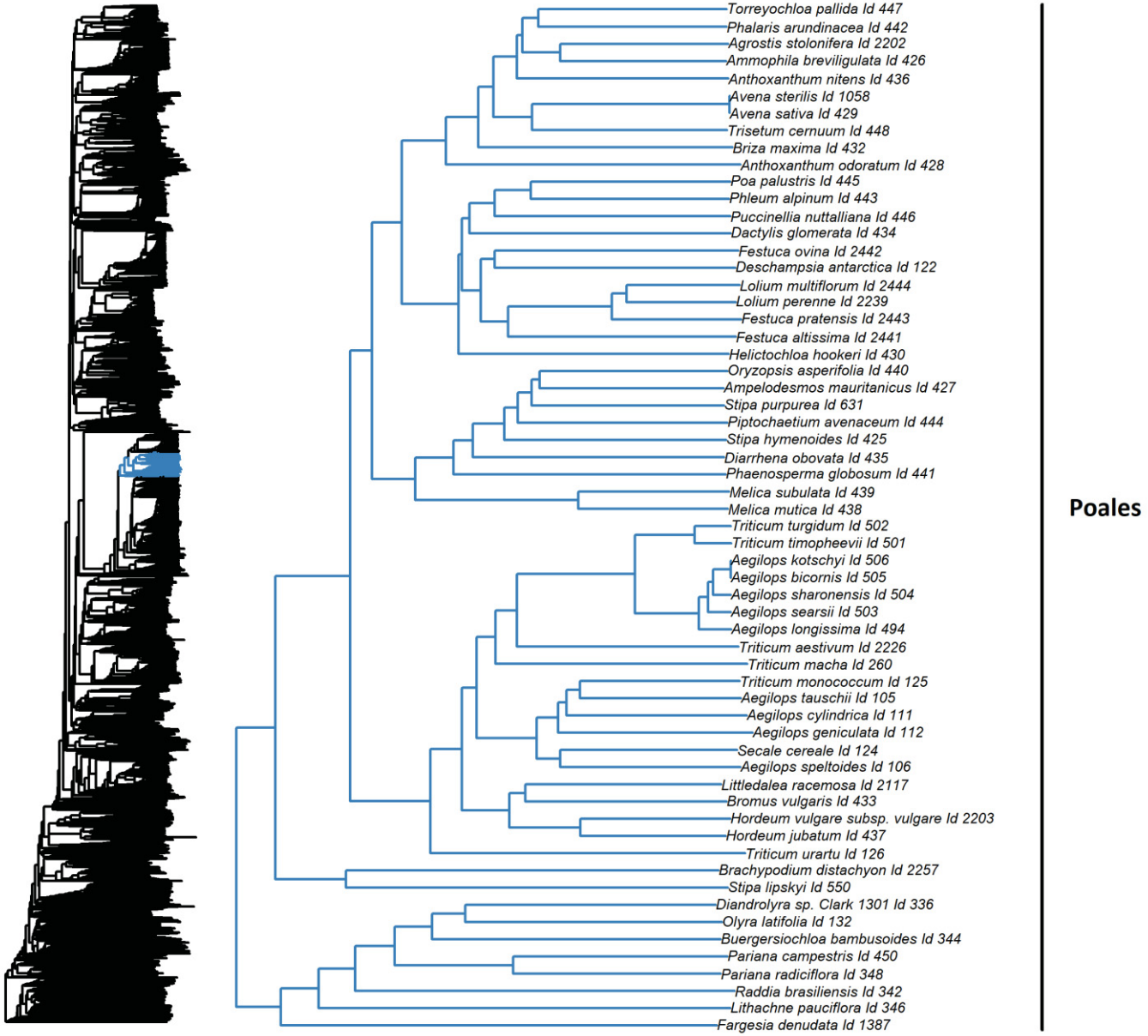

Sinojackia rehderiana Id 1891
Sinojackia xylocarpa Id 1767
Melliodendron xylocarpum Id 1864
Styrax grandiflorus Id 862
Pouteria campechiana Id 1259
Symplocos costaricana Id 1888
Symplocos paniculata Id 1883
Symplocos ovatilobata Id 2145
Barringtonia fusicarpa Id 1882
Barringtonia racemosa Id 1863
Bertholletia excelsa Id 2129
Androsace bulleyana Id 1369
Androsace paxiana Id 1368
Primula kwangtungensis Id 1364
Primula persimilis Id 1348
Primula sinensis Id 904
Primula veris Id 1043
Primula stenodonta Id 1403
Primula chrysochlora Id 1379
Primula poissonii Id 206
Ardisia polysticta Id 2448
Lysimachia coreana Id 266
Hydrangea petiolaris Id 1540
Hydrangea hydrangeoides Id 1492
Hydrangea luteovenosa Id 1858
Hydrangea serrata Id 1539
Diplopanax stachyanthus Id 706
Cornus controversa Id 830
Tetradoxa omeiensis Id 1398
Sinadoxa corydalifolia Id 1117
Adoxa moschatellina Id 1397
Sambucus williamsii Id 1305
Viburnum utile Id 1131
Helwingia himalaica Id 1022
Iodes cirrhosa Id 2110
Fatsia japonica Id 486
Eleutherococcus senticosus Id 457
Dendropanax morbifer Id 454
Dendropanax dentiger Id 293
Kalopanax septemlobus Id 56
Metapanax delavayi Id 13
Aralia undulata Id 55
Panax vietnamensis Id 589
Panax japonicus Id 588
Panax notoginseng Id 285
Panax ginseng Id 2383
Panax quinquefolius Id 420
Panax stipuleanatus Id 1556
Schefflera heptaphylla Id 707
Schefflera delavayi Id 14
Brassaiopsis hainla Id 12
Hydrocotyle sibthorpioides Id 1772
Hydrocotyle verticillata Id 1557
Pterygopleurum neurophyllum Id 1220
Peucedanum insolens Id 1219
Anethum graveolens Id 661
Foeniculum vulgare Id 660
Petroselinum crispum Id 1466
Crithmum maritimum Id 1465
Glehnia littoralis Id 1371
Angelica gigas Id 634
Actaea racemosa Id 1430
Angelica dahurica Id 633
Ostericum grosseserratum Id 582
Angelica acutiloba Id 632
Ligusticum tenuissimum Id 635
Ledebouriella seseloides Id 1401
Peucedanum japonicum Id 1370
Seseli montanum Id 419
Pastinaca pimpinellifolia Id 418
Coriandrum sativum Id 729
Anthriscus cerefolium Id 283
Daucus carota Id 2177
Hansenia forrestii Id 1619
Hansenia forbesii Id 1617
Hansenia oviformis Id 1618
Hansenia weberbaueriana Id 1616
Pleurospermum camtschaticum Id 1218
Arracacia xanthorrhiza Id 1136
Bupleurum boissieuanum Id 2020
Bupleurum falcatum Id 497
Bupleurum latissimum Id 1221
Ericales
Cornales
Dipsacales
Aquifoliales
Icacinales
Apiales

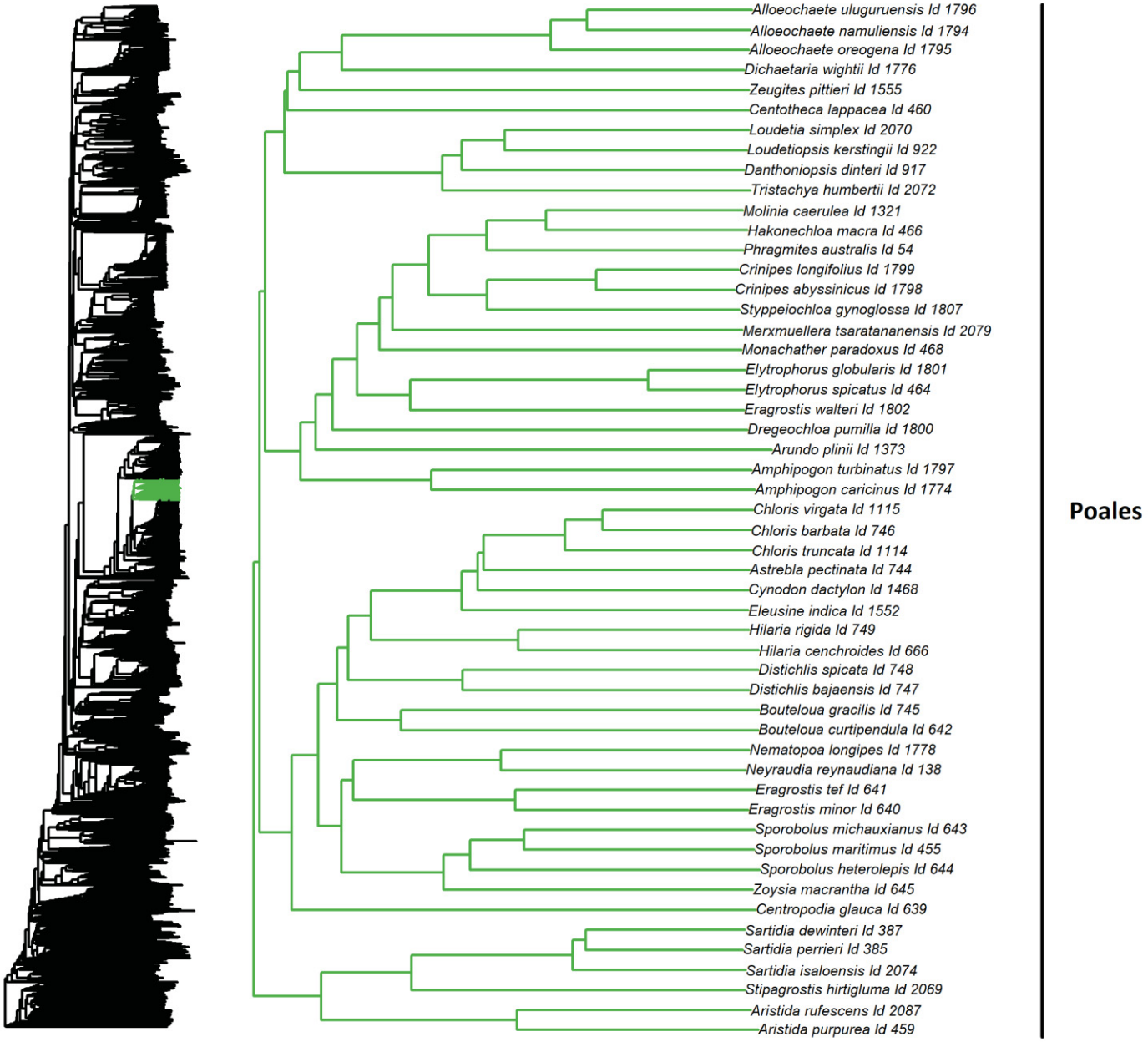

Echinacanthus lofouensis Id 1973
Andrographis paniculata Id 192
Rehmannia solanifolia Id 1338
Rehmannia henryi Id 1337
Rehmannia glutinosa Id 1336
Rehmannia elata Id 1340
Rehmannia chingii Id 1273
Rehmannia piasezkii Id 1339
Paulownia tomentosa Id 1048
Paulownia coreana Id 1047
Tectona grandis Id 2416
Premna microphylla Id 272
Sesamum indicum Id 458
Scrophularia buergeriana Id 1049
Scrophularia takesimensis Id 273
Lindenbergia philippensis Id 15
Erythranthe lutea Id 828
Stenogyne bifida Id 717
Stenogyne haliakalae Id 716
Stenogyne kanehoana Id 720
Haplostachys haplostachya Id 718
Stachys chamissonis Id 721
Stachys coccinea Id 722
Stachys sylvatica Id 723
Phyllostegia velutina Id 719
Stachys byzantina Id 724
Pogostemon stellatus Id 1046
Pogostemon yatabeanus Id 1045
Scutellaria lateriflora Id 1424
Scutellaria insignis Id 566
Scutellaria baicalensis Id 477
Caryopteris mongholica Id 1899
Ajuga reptans Id 113
Perilla frutescens Id 958
Perilla citriodora Id 957
Perilla setoyensis Id 950
Ocimum basilicum Id 1624
Salvia miltiorrhiza Id 2436
Lavandula angustifolia Id 626
Salvia japonica Id 1645
Mentha longifolia Id 1129
Salvia rosmarinus Id 411
Pedicularis cheilanthifolia Id 2018
Dracocephalum palmatum Id 1088
Digitalis lanata Id 1434
Veronicastrum sibiricum Id 1016
Veronica persica Id 1015
Veronica nakaiana Id 979
Olea europaea subsp. maroccana Id 2338
Olea europaea subsp. europaea Id 2332
Olea europaea subsp. cuspidata Id 2335
Olea europaea Id 2025
Olea woodiana subsp. woodiana Id 2336
Hesperelaea palmeri Id 255
Chionanthus retusus Id 1561
Forsythia suspensa Id 2123
Abeliophyllum distichum Id 1054
Primulina linearifolia Id 2131
Primulina huaijiensis Id 2130
Primulina eburnea Id 2063
Primulina liboensis Id 2064
Lysionotus pauciflorus Id 1376
Haberlea rhodopensis Id 1074
Dorcoceras hygrometricum Id 2357
Adenocalymma bracteatum Id 2149
Adenocalymma aurantiacum Id 2141
Adenocalymma hatschbachii Id 2150
Adenocalymma subspicatum Id 2153
Adenocalymma peregrinum Id 2152
Adenocalymma allamandiflorum Id 2147
Adenocalymma biternatum Id 2148
Neojobertia candolleana Id 2154
Adenocalymma pedunculatum Id 2151
Adenocalymma cristicalyx Id 2142
Tanaecium tetragonolobum Id 511
Aloysia citrodora Id 1425
Plantago media Id 561
Plantago maritima Id 560
Jasminum tortuosum Id 1437
Jasminum sambac Id 1417
Jasminum nudiflorum Id 2194
Castilleja paramensis Id 1276
Pedicularis ishidoyana Id 694
Schwalbea americana Id 81
Utricularia reniformis Id 966
Genlisea margaretae Id 227
Pinguicula ehlersiae Id 85
Utricularia gibba Id 2470
Utricularia macrorhiza Id 252
Lamiales


Ipomoea purpurea Id 2237
Ipomoea nil Id 1011
Ipomoea trifida Id 1377
Ipomoea batatas Id 399
Cressa cretica Id 1793
Catharanthus roseus Id 2477
Rhazya stricta Id 143
Carissa macrocarpa Id 1223
Nerium oleander Id 261
Oncinotis tenuiloba Id 233
Echites umbellatus Id 241
Pentalinon luteum Id 253
Gynochthodes nanlingensis Id 580
Asclepias syriaca Id 82
Asclepias nivea Id 2
Cynanchum auriculatum Id 658
Cynanchum wilfordii Id 657
Coffea arabica Id 2199
Coffea canephora Id 797
Emmenopterys henryi Id 2109
Mitragyna speciosa Id 1438
Gynochthodes officinalis Id 518
Gentiana tibetica Id 839
Gentiana crassicaulis Id 417
Gentiana straminea Id 416
Gentiana macrophylla Id 1893
Swertia mussotii Id 969
Atropa belladonna Id 2287
Hyoscyamus niger Id 142
Scopolia parviflora Id 832
Nicotiana tabacum Id 2393
Nicotiana sylvestris Id 2390
Nicotiana undulata Id 2355
Nicotiana tomentosiformis Id 2394
Nicotiana attenuata Id 1990
Nicotiana otophora Id 1214
Solanum dulcamara Id 1897
Solanum nigrum Id 732
Solanum lycopersicum Id 2493
Solanum lycopersicum Id 78
Solanum pennellii Id 2102
Solanum peruvianum Id 321
Solanum chilense Id 319
Solanum neorickii Id 312
Solanum pimpinellifolium Id 322
Solanum galapagense Id 320
Solanum cheesmaniae Id 310
Solanum habrochaites Id 311
Solanum berthaultii Id 1493
Solanum commersonii Id 731
Solanum tuberosum Id 2291
Solanum bulbocastanum Id 2397
Datura stramonium Id 2164
Iochroma cardenasianum Id 704
Capsicum tovarii Id 1272
Capsicum frutescens Id 517
Capsicum annuum Id 2399
Capsicum galapagoense Id 1270
Capsicum eximium Id 1255
Capsicum chinense Id 920
Capsicum chacoense Id 1271
Capsicum lycianthoides Id 329
Dunalia brachyacantha Id 323
Saracha punctata Id 288
Iochroma nitidum Id 301
Acnistus arborescens Id 822
Iochroma australe Id 728
Iochroma ellipticum Id 819
Iochroma loxense Id 290
Iochroma salpoanum Id 815
Iochroma tingoanum Id 372
Dunalia solanacea Id 367
Iochroma cyaneum Id 820
Eriolarynx fasciculata Id 817
Acnistus arborescens x Iochroma cyaneum Id 798
Iochroma umbellatum Id 795
Physalis peruviana Id 330
Iochroma lehmannii Id 814
Dunalia obovata Id 299
Iochroma stenanthum Id 295
Solanum melongena Id 827
Gentianales
Solanales

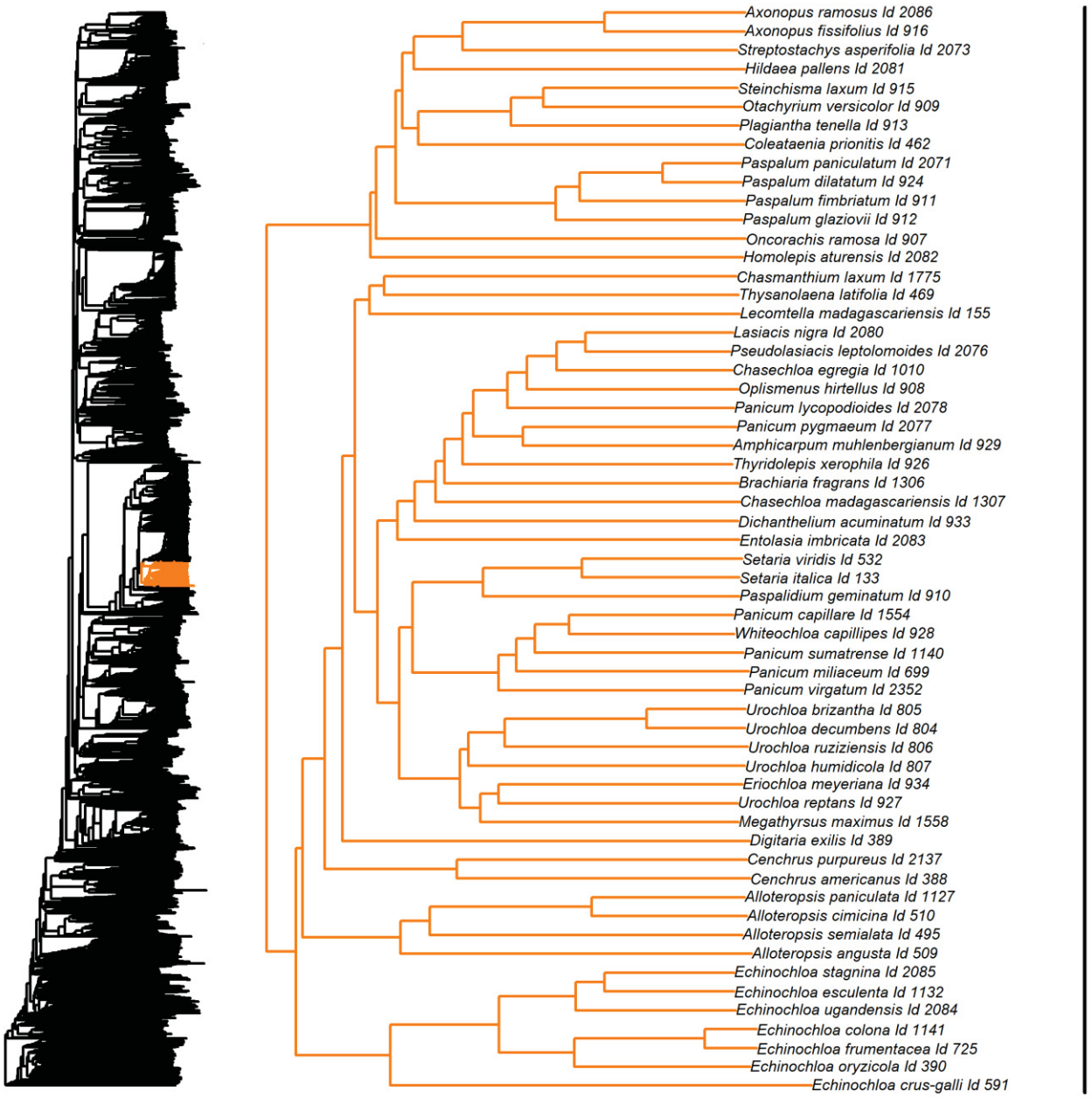

Salicornia europaea ld 401
Salicornia brachiata ld 400
Salicornia bigelovii ld 407
Haloxylon persicum ld 480
Haloxylon ammodendron ld 479
Chenopodium album ld 1546
Chenopodium quinoa ld 1545
Spinacia oleracea ld 2192
Portulaca oleracea ld 2107
Cistanthe longiscapa ld 1623
Mesembryanthemum crystallinum ld 601
Amaranthus hypochondriacus ld 951
Silene conica ld 2364
Silene conoidea ld 119
Silene noctiflora ld 2363
Silene latifolia ld 2365
Silene vulgaris ld 2362
Silene capitata ld 1641
Agrostemma githago ld 118
Silene paradoxa ld 121
Lychnis wilfordii ld 1640
Silene chalcedonica ld 120
Colobanthus apetalus ld 2134
Colobanthus quitensis ld 533
Oxyria sinensis ld 1113
Rheum palmatum ld 489
Rheum wittrockii ld 1986
Fagopyrum esculentum subsp. ancestrale ld 2256
Fagopyrum tataricum ld 380
Caryophyllales

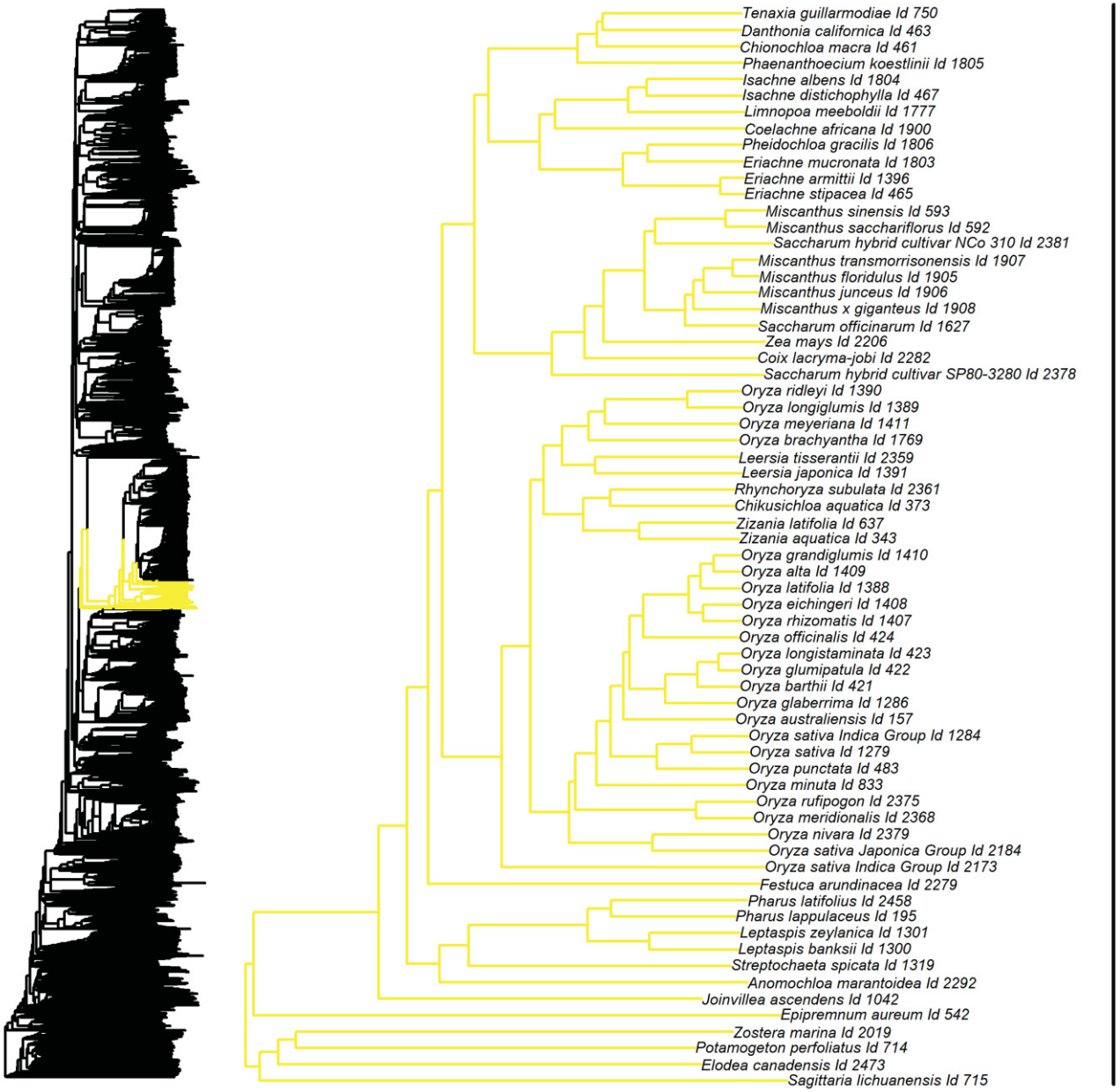

Carex neurocarpa Id 2029
Carex siderosticta Id 410
Hypolytrum nemorum Id 2028
Bambusa arnhemica Id 334
Bambusa bambos Id 333
Neohouzeaua sp. Clark & Attigala 1712 Id 339
Bambusa emeiensis Id 2345
Bambusa multiplex Id 166
Dendrocalamus latiflorus Id 2281
Neololeba atra Id 340
Hickelia madagascariensis Id 338
Bambusa oldhamii Id 2277
Otatea glauca Id 587
Otatea acuminata Id 347
Guadua angustifolia Id 705
Guadua chacoensis Id 609
Olmeca reflexa Id 341
Guadua weberbaueri Id 354
Chusquea circinata Id 449
Chusquea liebmannii Id 345
Greslania sp. McPherson 19217 Id 337
Chusquea spectabilis Id 335
Phyllostachys propinqua Id 2360
Indocalamus longiauritus Id 2341
Phyllostachys nigra var. henonis Id 2344
Phyllostachys sulphurea Id 167
Phyllostachys edulis Id 2342
Sarocalamus faberi Id 172
Yushania levigata Id 182
Fargesia yunnanensis Id 176
Fargesia spathacea Id 175
Arundinaria fargesii Id 171
Fargesia nitida Id 174
Thamnocalamus spathiflorus Id 187
Indocalamus wilsonii Id 179
Ferrocalamus rimosivaginus Id 2346
Gelidocalamus tessellatus Id 178
Gaoligongshania megalothyrsa Id 177
Arundinaria gigantea Id 2426
Arundinaria appalachiana Id 98
Pleioblastus maculatus Id 186
Indosasa sinica Id 180
Acidosasa purpurea Id 2343
Oligostachyum shiuyingianum Id 181
Ampelocalamus naibunensis Id 960
Chimonocalamus longiusculus Id 173
Arundinaria tecta Id 274
Ampelocalamus calcareus Id 392
Pseudosasa japonica Id 540
Puelia olyriformis Id 196
Brachyelytrum aristosum Id 431
Poales

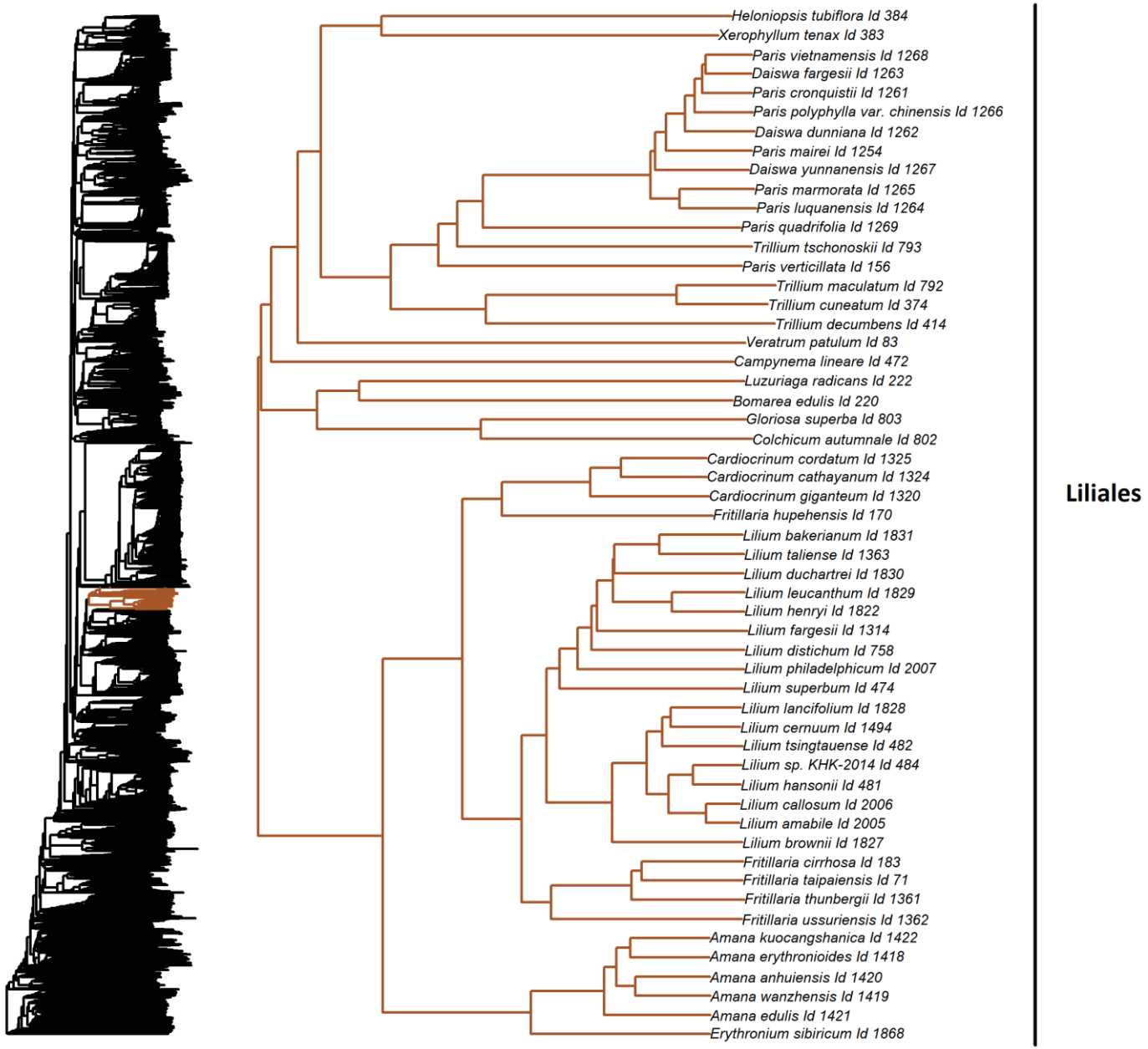

Torreyochloa pallida ld 447
Phalaris arundinacea ld 442
Agrostis stolonifera ld 2202
Ammophila breviligulata ld 426
Anthoxanthum nitens ld 436
Avena sterilis ld 1058
Avena sativa ld 429
Trisetum cernuum ld 448
Briza maxima ld 432
Anthoxanthum odoratum ld 428
Poa palustris ld 445
Phleum alpinum ld 443
Puccinellia nuttalliana ld 446
Dactylis glomerata ld 434
Festuca ovina ld 2442
Deschampsia antarctica ld 122
Lolium multiflorum ld 2444
Lolium perenne ld 2239
Festuca pratensis ld 2443
Festuca altissima ld 2441
Helictochloa hookeri ld 430
Oryzopsis asperifolia ld 440
Ampelodesmos mauritanicus ld 427
Stipa purpurea ld 631
Piptochaetium avenaceum ld 444
Stipa hymenoides ld 425
Diarrhena obovata ld 435
Phaenosperma globosum ld 441
Melica subulata ld 439
Melica mutica ld 438
Triticum turgidum ld 502
Triticum timopheevii ld 501
Aegilops kotschyi ld 506
Aegilops bicornis ld 505
Aegilops sharonensis ld 504
Aegilops searsii ld 503
Aegilops longissima ld 494
Triticum aestivum ld 2226
Triticum macha ld 260
Triticum monococcum ld 125
Aegilops tauschii ld 105
Aegilops cylindrica ld 111
Aegilops geniculata ld 112
Secale cereale ld 124
Aegilops speltoides ld 106
Littledalea racemosa ld 2117
Bromus vulgaris ld 433
Hordeum vulgare subsp. vulgare ld 2203
Hordeum jubatum ld 437
Triticum urartu ld 126
Brachypodium distachyon ld 2257
Stipa lipskyi ld 550
Diandrolyra sp. Clark 1301 ld 336
Olyra latifolia ld 132
Buergersiochloa bambusoides ld 344
Pariana campestris ld 450
Pariana radiciflora ld 348
Raddia brasiliensis ld 342
Lithachne pauciflora ld 346
Fargesia denudata ld 1387
Poales

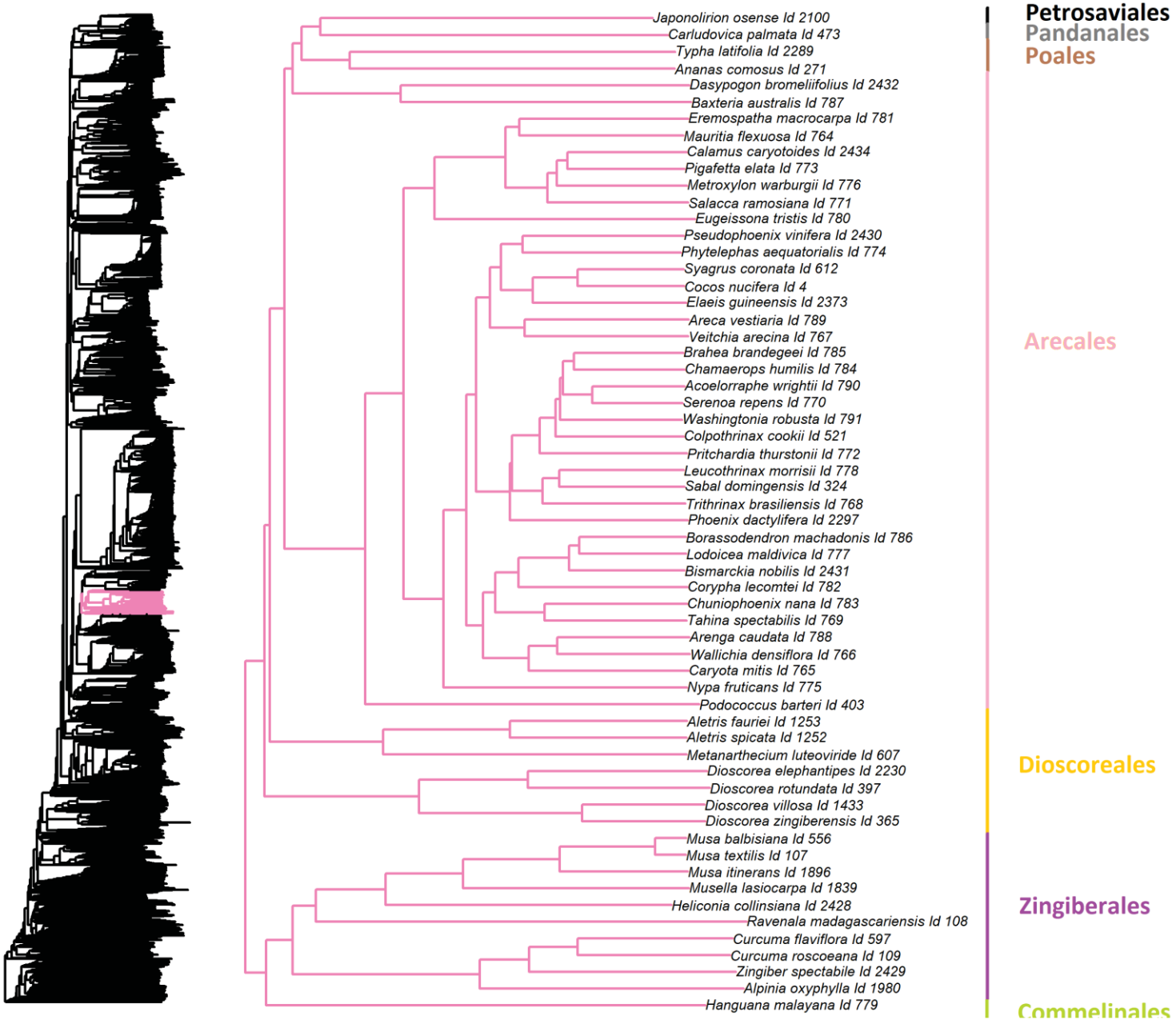

Alloeochaete uluguruensis Id 1796
Alloeochaete namuliensis Id 1794
Alloeochaete oreogena Id 1795
Dichaetaria wightii Id 1776
Zeugites pittieri Id 1555
Centotheca lappacea Id 460
Loudetia simplex Id 2070
Loudetiopsis kerstingii Id 922
Danthoniopsis dinteri Id 917
Tristachya humbertii Id 2072
Molinia caerulea Id 1321
Hakonechloa macra Id 466
Phragmites australis Id 54
Crinipes longifolius Id 1799
Crinipes abyssinicus Id 1798
Styppeiochloa gynoglossa Id 1807
Merxmuellera tsaratananensis Id 2079
Monachather paradoxus Id 468
Elytrophorus globularis Id 1801
Elytrophorus spicatus Id 464
Eragrostis walteri Id 1802
Dregeochloa pumilla Id 1800
Arundo plinii Id 1373
Amphipogon turbinatus Id 1797
Amphipogon caricinus Id 1774
Chloris virgata Id 1115
Chloris barbata Id 746
Chloris truncata Id 1114
Astrebla pectinata Id 744
Cynodon dactylon Id 1468
Eleusine indica Id 1552
Hilaria rigida Id 749
Hilaria cenchroides Id 666
Distichlis spicata Id 748
Distichlis bajaensis Id 747
Bouteloua gracilis Id 745
Bouteloua curtipendula Id 642
Nematopoa longipes Id 1778
Neyraudia reynaudiana Id 138
Eragrostis tef Id 641
Eragrostis minor Id 640
Sporobolus michauxianus Id 643
Sporobolus maritimus Id 455
Sporobolus heterolepis Id 644
Zoysia macrantha Id 645
Centropodia glauca Id 639
Sartidia dewinteri Id 387
Sartidia perrieri Id 385
Sartidia isaloensis Id 2074
Stipagrostis hirtigluma Id 2069
Aristida rufescens Id 2087
Aristida purpurea Id 459
Poales

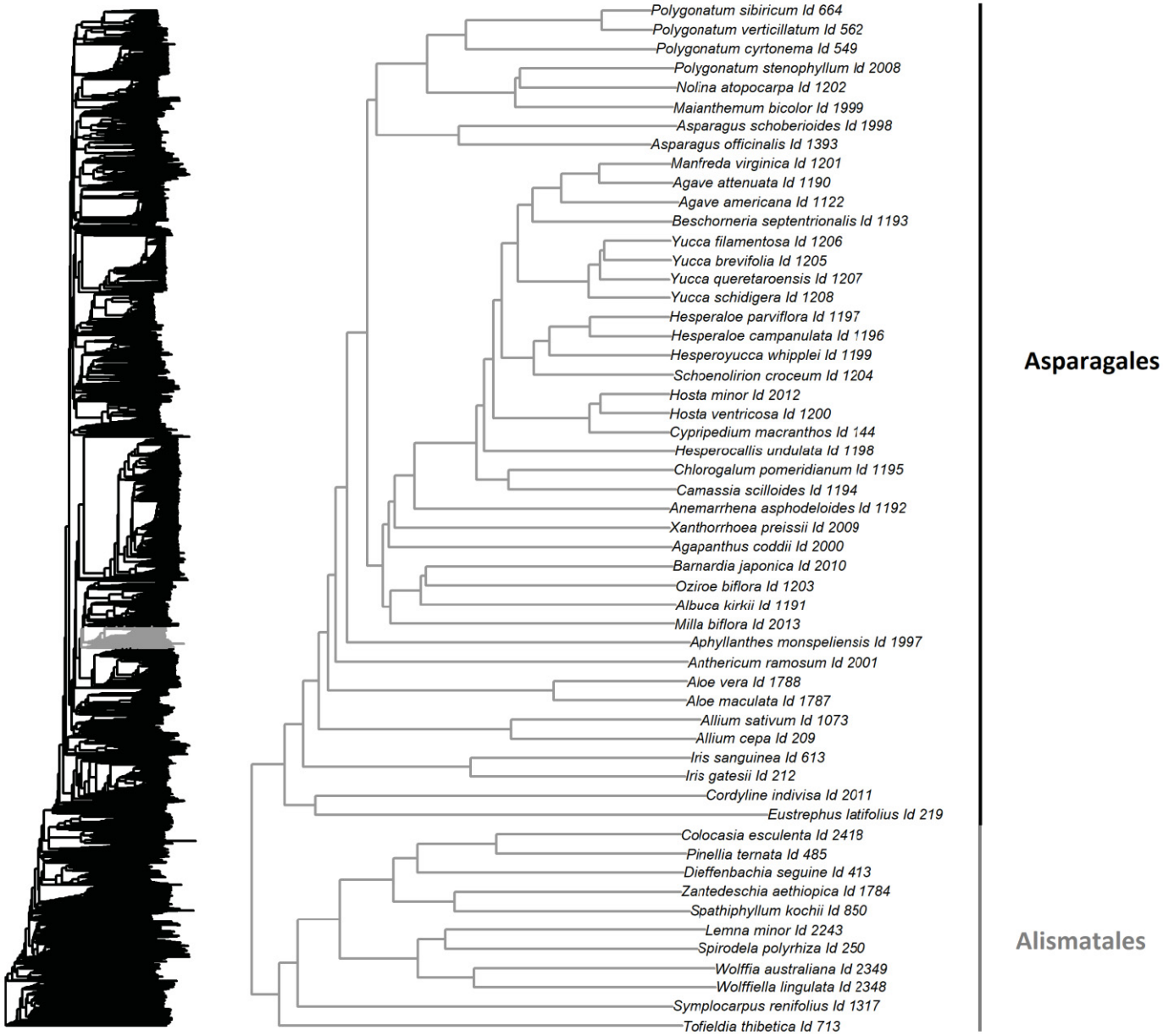

Chrysopogon gryllus Id 1602
Chrysopogon zizanioides Id 1601
Mnesithea helferi Id 1603
Saccharum arundinaceum Id 962
Rottboellia cochinchinensis Id 925
Ischaemum afrum Id 906
Sorghum bicolor Id 2204
Hemisorghum mekongense Id 1592
Saccharum hildebrandtii Id 2075
Saccharum hybrid cultivar Id 608
Eulalia aurea Id 918
Germainia capitata Id 1612
Pogonatherum paniceum Id 738
Imperata cylindrica Id 905
Sorghastrum nutans Id 914
Themeda villosa Id 1591
Themeda arundinacea Id 1563
Heteropogon contortus Id 1565
Heteropogon triticeus Id 1595
Cymbopogon flexuosus Id 1607
Themeda arguens Id 1589
Themeda triandra Id 1587
Themeda quadrivalvis Id 1783
Iseilema macratherum Id 921
Capillipedium venustum Id 932
Bothriochloa alta Id 931
Dichanthium sericeum Id 1564
Hyparrhenia diplandra Id 1604
Hyparrhenia rufa Id 1562
Hyparrhenia subplumosa Id 935
Andropogon distachyos Id 1608
Andropogon abyssinicus Id 1567
Schizachyrium spicatum Id 1606
Schizachyrium scoparium Id 1599
Schizachyrium exile Id 1566
Schizachyrium sanguineum Id 1586
Schizachyrium brevifolium Id 1585
Andropogon chinensis Id 1584
Schizachyrium imberbe Id 1611
Schizachyrium tenerum Id 1610
Andropogon fastigiatus Id 1583
Eulaliopsis binata Id 1614
Dimeria ornithopoda Id 1590
Sorghum timorense Id 93
Eulalia siamensis Id 1598
Pseudosorghum fasciculare Id 1594
Chrysopogon orientalis Id 1613
Chrysopogon serrulatus Id 741
Eriochrysis laxa Id 740
Eriochrysis cf. cayennensis 365-2 Id 739
Andropogon burmanicus Id 1605
Eulalia contorta Id 1596
Parahyparrhenia siamensis Id 1600
Kerriochloa siamensis Id 1582
Eremochloa ciliaris Id 1597
Eremochloa eriopoda Id 1593
Saccharum spontaneum Id 1400
Zea luxurians Id 836
Zea perennis Id 835
Zea diploperennis Id 851
Arthraxon lanceolatus Id 1588
Arthraxon hispidus Id 1569
Arthraxon prionodes Id 923
Arthraxon microphyllus Id 1570
Polytoca digitata Id 1568
Arundinella deppeana Id 930
Garnotia tenella Id 1615
Garnotia thailandica Id 1609
Poales

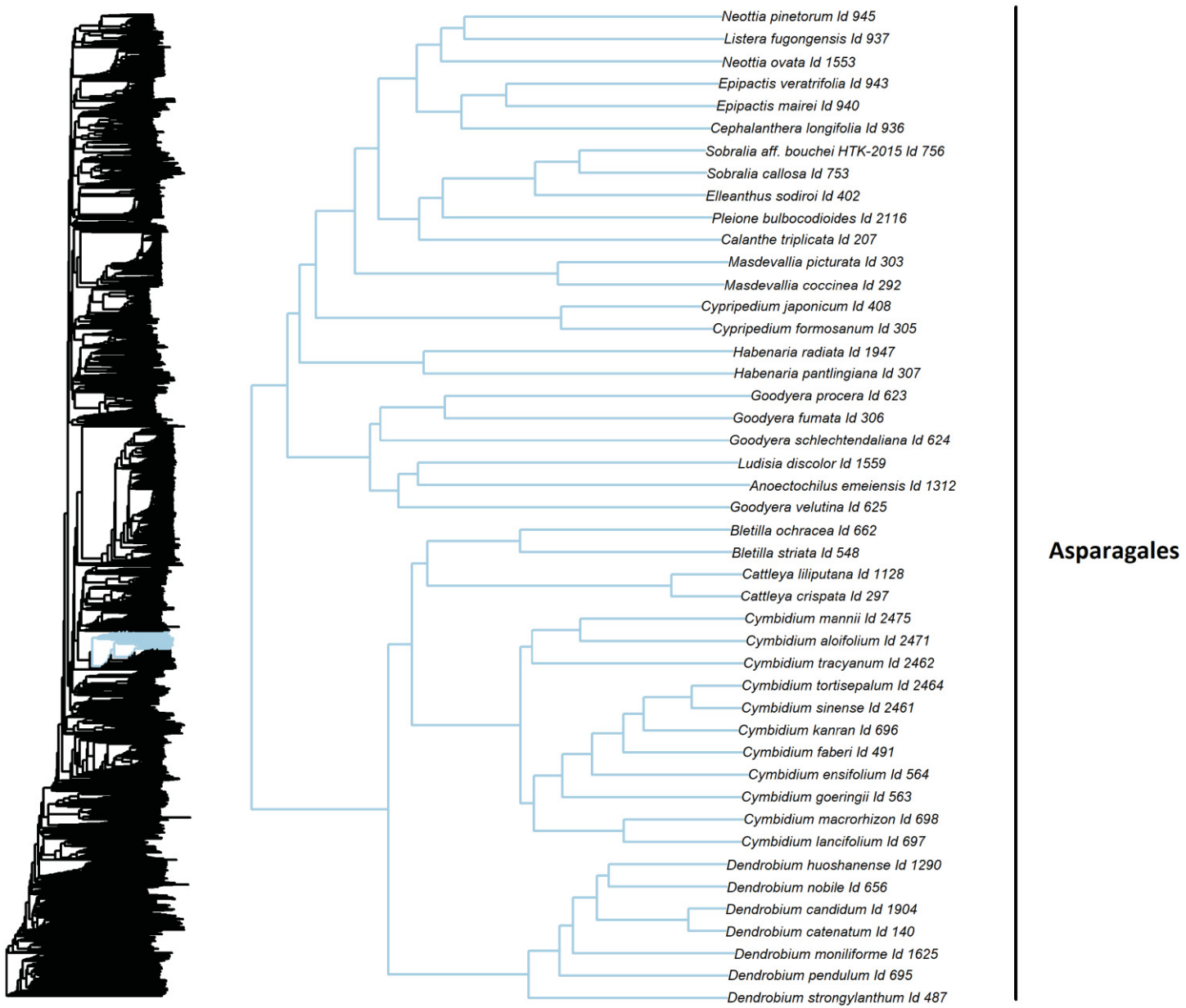

Axonopus ramosus Id 2086
Axonopus fissifolius Id 916
Streptostachys asperifolia Id 2073
Hildaea pallens Id 2081
Steinchisma laxum Id 915
Otachyrium versicolor Id 909
Plagiantha tenella Id 913
Coleataenia prionitis Id 462
Paspalum paniculatum Id 2071
Paspalum dilatatum Id 924
Paspalum fimbriatum Id 911
Paspalum glaziovii Id 912
Oncorachis ramosa Id 907
Homolepis aturensis Id 2082
Chasmanthium laxum Id 1775
Thysanolaena latifolia Id 469
Lecomtella madagascariensis Id 155
Lasiacis nigra Id 2080
Pseudolasiacis leptolomoides Id 2076
Chasechloa egregia Id 1010
Oplismenus hirtellus Id 908
Panicum lycopodioides Id 2078
Panicum pygmaeum Id 2077
Amphicarpum muhlenbergianum Id 929
Thyridolepis xerophila Id 926
Brachiaria fragrans Id 1306
Chasechloa madagascariensis Id 1307
Dichanthelium acuminatum Id 933
Entolasia imbricata Id 2083
Setaria viridis Id 532
Setaria italica Id 133
Paspalidium geminatum Id 910
Panicum capillare Id 1554
Whiteochloa capillipes Id 928
Panicum sumatrense Id 1140
Panicum miliaceum Id 699
Panicum virgatum Id 2352
Urochloa brizantha Id 805
Urochloa decumbens Id 804
Urochloa ruziziensis Id 806
Urochloa humidicola Id 807
Eriochloa meyeriana Id 934
Urochloa reptans Id 927
Megathyrsus maximus Id 1558
Digitaria exilis Id 389
Cenchrus purpureus Id 2137
Cenchrus americanus Id 388
Alloteropsis paniculata Id 1127
Alloteropsis cimicina Id 510
Alloteropsis semialata Id 495
Alloteropsis angusta Id 509
Echinochloa stagnina Id 2085
Echinochloa esculenta Id 1132
Echinochloa ugandensis Id 2084
Echinochloa colona Id 1141
Echinochloa frumentacea Id 725
Echinochloa oryzicola Id 390
Echinochloa crus-galli Id 591
Poales

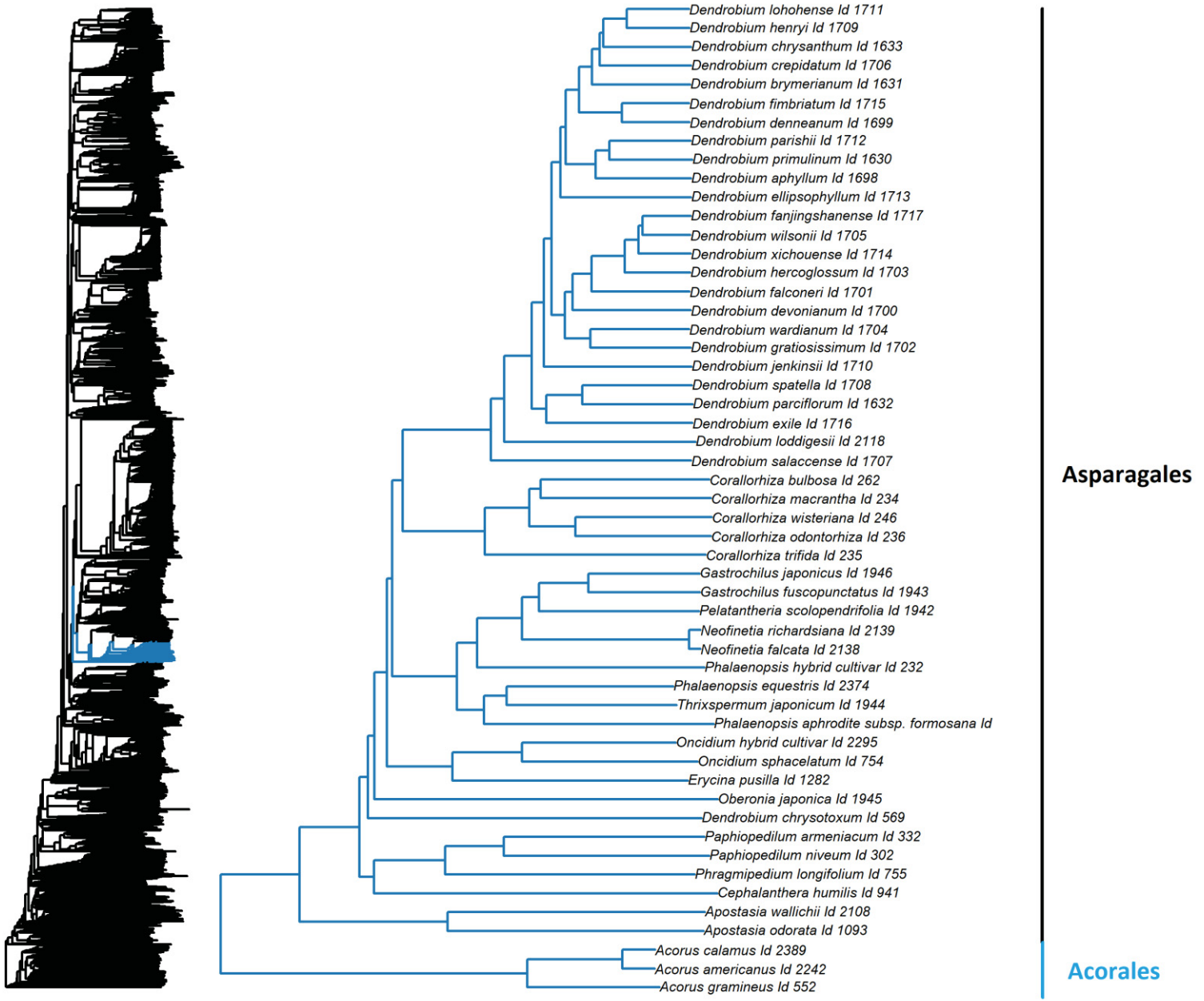

Tenaxia guillarmodiae Id 750
Danthonia californica Id 463
Chionochloa macra Id 461
Phaenanthoecium koestlinii Id 1805
Isachne albens Id 1804
Isachne distichophylla Id 467
Limnopoa meeboldii Id 1777
Coelachne africana Id 1900
Pheidochloa gracilis Id 1806
Eriachne mucronata Id 1803
Eriachne armittii Id 1396
Eriachne stipacea Id 465
Miscanthus sinensis Id 593
Miscanthus sacchariflorus Id 592
Saccharum hybrid cultivar NCo 310 Id 2381
Miscanthus transmorrisonensis Id 1907
Miscanthus floridulus Id 1905
Miscanthus junceus Id 1906
Miscanthus x giganteus Id 1908
Saccharum officinarum Id 1627
Zea mays Id 2206
Coix lacryma-jobi Id 2282
Saccharum hybrid cultivar SP80-3280 Id 2378
Oryza ridleyi Id 1390
Oryza longiglumis Id 1389
Oryza meyeriana Id 1411
Oryza brachyantha Id 1769
Leersia tisserantii Id 2359
Leersia japonica Id 1391
Rhynchoryza subulata Id 2361
Chikusichloa aquatica Id 373
Zizania latifolia Id 637
Zizania aquatica Id 343
Oryza grandiglumis Id 1410
Oryza alta Id 1409
Oryza latifolia Id 1388
Oryza eichingeri Id 1408
Oryza rhizomatis Id 1407
Oryza officinalis Id 424
Oryza longistaminata Id 423
Oryza glumipatula Id 422
Oryza barthii Id 421
Oryza glaberrima Id 1286
Oryza australiensis Id 157
Oryza sativa Indica Group Id 1284
Oryza sativa Id 1279
Oryza punctata Id 483
Oryza minuta Id 833
Oryza rufipogon Id 2375
Oryza meridionalis Id 2368
Oryza nivara Id 2379
Oryza sativa Japonica Group Id 2184
Oryza sativa Indica Group Id 2173
Festuca arundinacea Id 2279
Pharus latifolius Id 2458
Pharus lappulaceus Id 195
Leptaspis zeylanica Id 1301
Leptaspis banksii Id 1300
Streptochaeta spicata Id 1319
Anomochloa marantoidea Id 2292
Joinvillea ascendens Id 1042
Epipremnum aureum Id 542
Zostera marina Id 2019
Potamogeton perfoliatus Id 714
Elodea canadensis Id 2473
Sagittaria lichuanensis Id 715
Poales

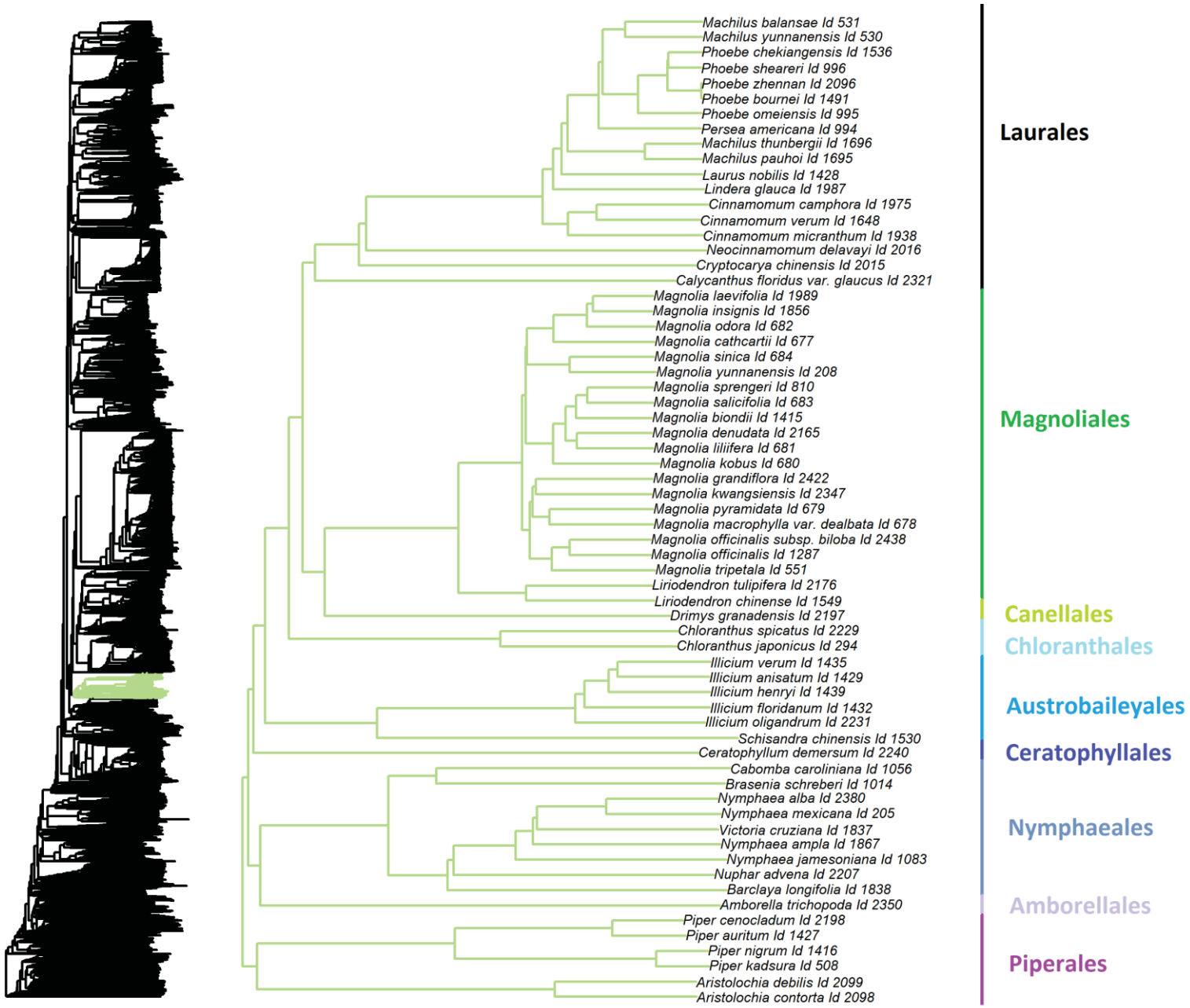

Heloniopsis tubiflora Id 384
Xerophyllum tenax Id 383
Paris vietnamensis Id 1268
Daiswa fargesii Id 1263
Paris cronquistii Id 1261
Paris polyphylla var. chinensis Id 1266
Daiswa dunniana Id 1262
Paris mairei Id 1254
Daiswa yunnanensis Id 1267
Paris marmorata Id 1265
Paris luquanensis Id 1264
Paris quadrifolia Id 1269
Trillium tschonoskii Id 793
Paris verticillata Id 156
Trillium maculatum Id 792
Trillium cuneatum Id 374
Trillium decumbens Id 414
Veratrum patulum Id 83
Campynema lineare Id 472
Luzuriaga radicans Id 222
Bomarea edulis Id 220
Gloriosa superba Id 803
Colchicum autumnale Id 802
Cardiocrinum cordatum Id 1325
Cardiocrinum cathayanum Id 1324
Cardiocrinum giganteum Id 1320
Fritillaria hupehensis Id 170
Lilium bakerianum Id 1831
Lilium taliense Id 1363
Lilium duchartrei Id 1830
Lilium leucanthum Id 1829
Lilium henryi Id 1822
Lilium fargesii Id 1314
Lilium distichum Id 758
Lilium philadelphicum Id 2007
Lilium superbum Id 474
Lilium lancifolium Id 1828
Lilium cernuum Id 1494
Lilium tsingtauense Id 482
Lilium sp. KHK-2014 Id 484
Lilium hansonii Id 481
Lilium callosum Id 2006
Lilium amabile Id 2005
Lilium brownii Id 1827
Fritillaria cirrhosa Id 183
Fritillaria taipaiensis Id 71
Fritillaria thunbergii Id 1361
Fritillaria ussuriensis Id 1362
Amana kuocangshanica Id 1422
Amana erythronioides Id 1418
Amana anhuiensis Id 1420
Amana wanzhensis Id 1419
Amana edulis Id 1421
Erythronium sibiricum Id 1868
Liliales

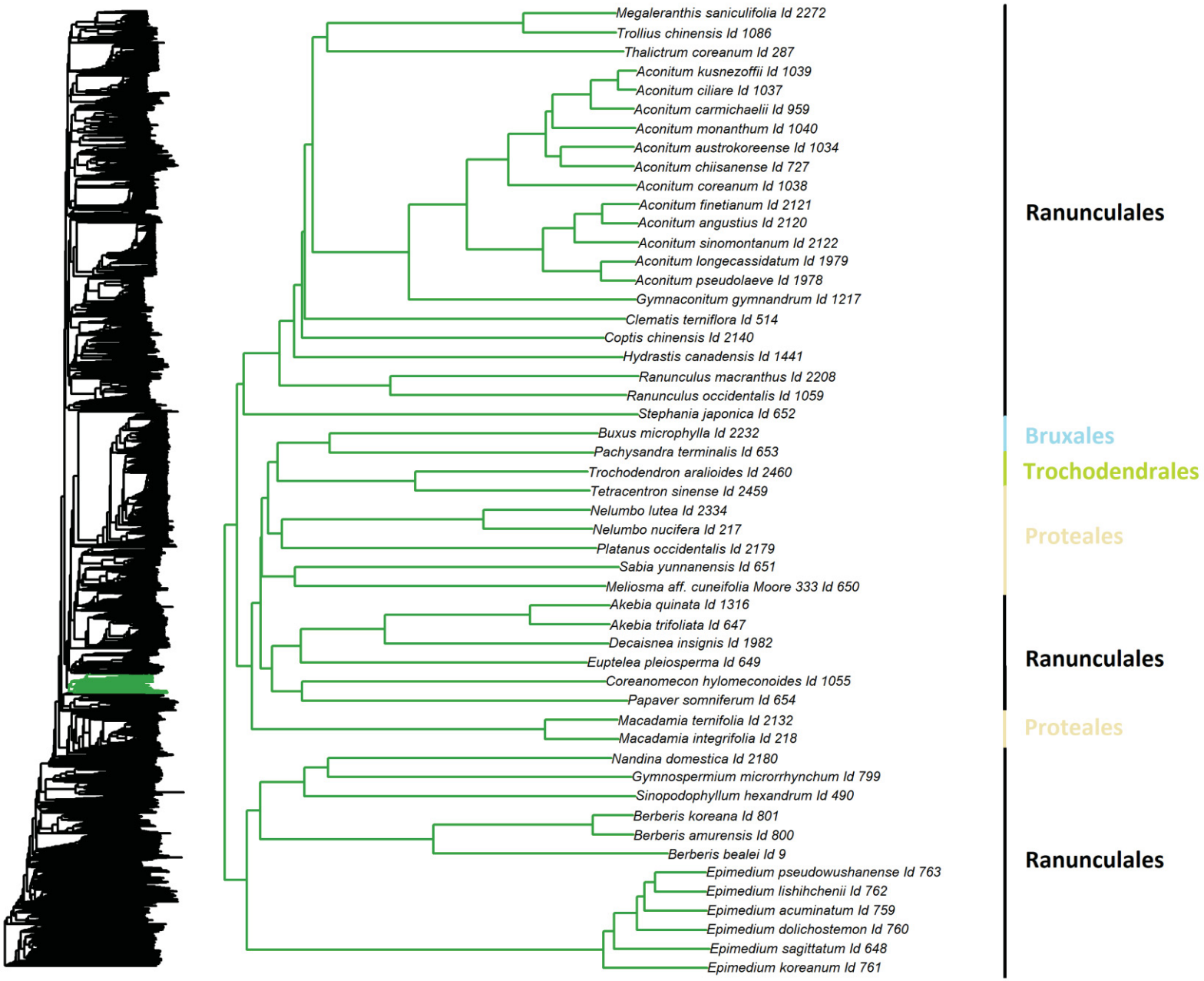

Japonolirion osense Id 2100
Carludovica palmata Id 473
Typha latifolia Id 2289
Ananas comosus Id 271
Dasypogon bromeliifolius Id 2432
Baxteria australis Id 787
Eremospatha macrocarpa Id 781
Mauritia flexuosa Id 764
Calamus caryotoides Id 2434
Pigafetta elata Id 773
Metroxylon warburgii Id 776
Salacca ramosiana Id 771
Eugeissona tristis Id 780
Pseudophoenix vinifera Id 2430
Phytelephas aequatorialis Id 774
Syagrus coronata Id 612
Cocos nucifera Id 4
Elaeis guineensis Id 2373
Areca vestiaria Id 789
Veitchia arecina Id 767
Brahea brandegeei Id 785
Chamaerops humilis Id 784
Acoelorraphe wrightii Id 790
Serenoa repens Id 770
Washingtonia robusta Id 791
Colpothrinax cookii Id 521
Pritchardia thurstonii Id 772
Leucothrinax morrisii Id 778
Sabal domingensis Id 324
Trithrinax brasiliensis Id 768
Phoenix dactylifera Id 2297
Borassodendron machadonis Id 786
Lodoicea maldivica Id 777
Bismarckia nobilis Id 2431
Corypha lecomtei Id 782
Chuniophoenix nana Id 783
Tahina spectabilis Id 769
Arenga caudata Id 788
Wallichia densiflora Id 766
Caryota mitis Id 765
Nypa fruticans Id 775
Podococcus barteri Id 403
Aletris fauriei Id 1253
Aletris spicata Id 1252
Metanarthecium luteoviride Id 607
Dioscorea elephantipes Id 2230
Dioscorea rotundata Id 397
Dioscorea villosa Id 1433
Dioscorea zingiberensis Id 365
Musa balbisiana Id 556
Musa textilis Id 107
Musa itinerans Id 1896
Musella lasiocarpa Id 1839
Heliconia collinsiana Id 2428
Ravenala madagascariensis Id 108
Curcuma flaviflora Id 597
Curcuma roscoeana Id 109
Zingiber spectabile Id 2429
Alpinia oxyphylla Id 1980
Hanguana malayana Id 779
Petrosaviales
Pandanales
Poales
Arecales
Dioscoreales
Zingiberales
Commelinales

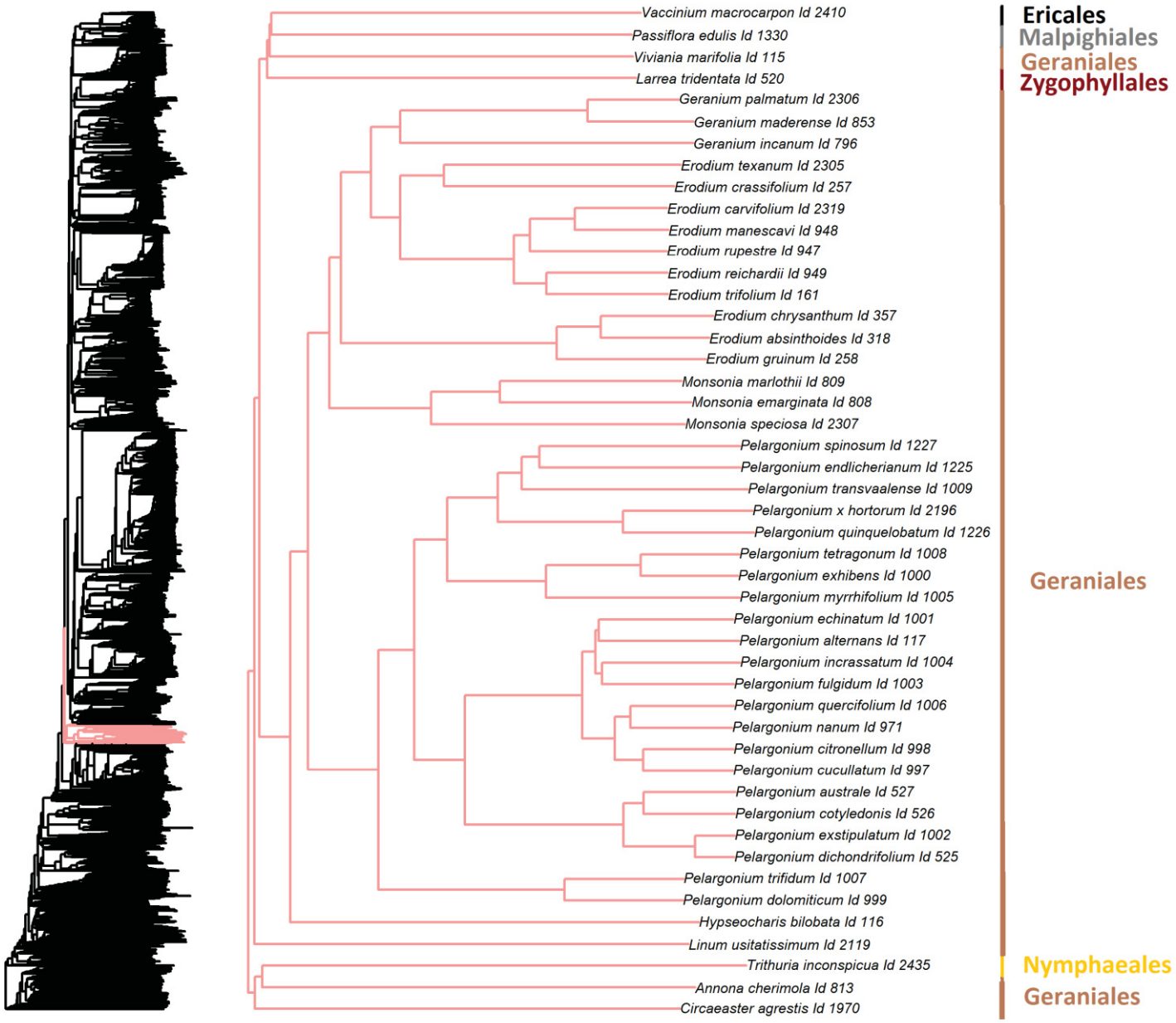

Polygonatum sibiricum Id 664
Polygonatum verticillatum Id 562
Polygonatum cyrtonema Id 549
Polygonatum stenophyllum Id 2008
Nolina atopocarpa Id 1202
Maianthemum bicolor Id 1999
Asparagus schoberioides Id 1998
Asparagus officinalis Id 1393
Manfreda virginica Id 1201
Agave attenuata Id 1190
Agave americana Id 1122
Beschorneria septentrionalis Id 1193
Yucca filamentosa Id 1206
Yucca brevifolia Id 1205
Yucca queretaroensis Id 1207
Yucca schidigera Id 1208
Hesperaloe parviflora Id 1197
Hesperaloe campanulata Id 1196
Hesperoyucca whipplei Id 1199
Schoenolirion croceum Id 1204
Hosta minor Id 2012
Hosta ventricosa Id 1200
Cypripedium macranthos Id 144
Hesperocallis undulata Id 1198
Chlorogalum pomeridianum Id 1195
Camassia scilloides Id 1194
Anemarrhena asphodeloides Id 1192
Xanthorrhoea preissii Id 2009
Agapanthus coddii Id 2000
Barnardia japonica Id 2010
Oziroe biflora Id 1203
Albuca kirkii Id 1191
Milla biflora Id 2013
Aphyllanthes monspeliensis Id 1997
Anthericum ramosum Id 2001
Aloe vera Id 1788
Aloe maculata Id 1787
Allium sativum Id 1073
Allium cepa Id 209
Iris sanguinea Id 613
Iris gatesii Id 212
Cordyline indivisa Id 2011
Eustrephus latifolius Id 219
Colocasia esculenta Id 2418
Pinellia ternata Id 485
Dieffenbachia seguine Id 413
Zantedeschia aethiopica Id 1784
Spathiphyllum kochii Id 850
Lemna minor Id 2243
Spirodela polyrhiza Id 250
Wolffia australiana Id 2349
Wolffiella lingulata Id 2348
Symplocarpus renifolius Id 1317
Tofieldia thibetica Id 713
Asparagales
Alismatales

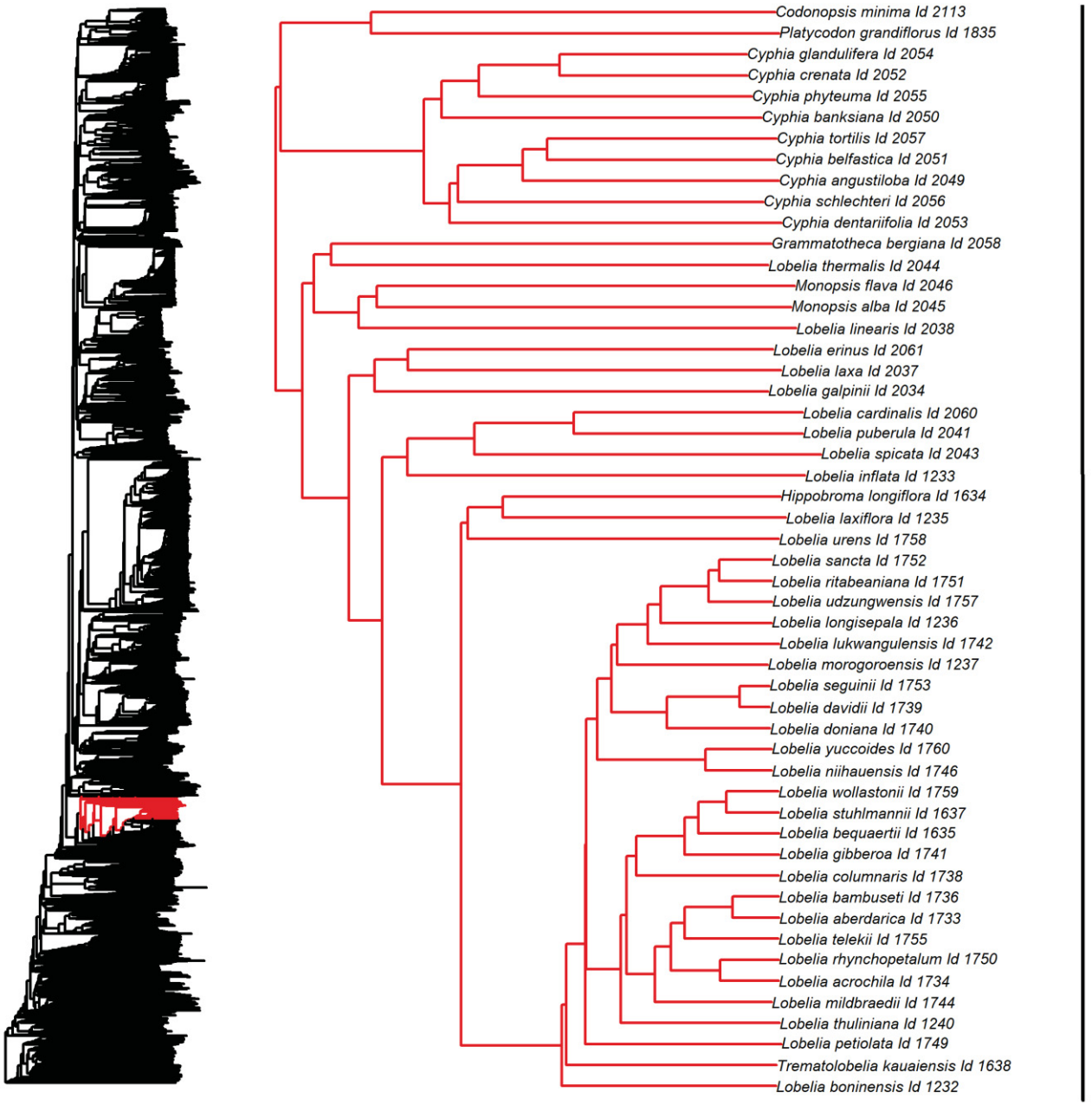

Neottia pinetorum Id 945
Listera fugongensis Id 937
Neottia ovata Id 1553
Epipactis veratrifolia Id 943
Epipactis mairei Id 940
Cephalanthera longifolia Id 936
Sobralia aff. bouchei HTK-2015 Id 756
Sobralia callosa Id 753
Elleanthus sodiroi Id 402
Pleione bulbocodioides Id 2116
Calanthe triplicata Id 207
Masdevallia picturata Id 303
Masdevallia coccinea Id 292
Cypripedium japonicum Id 408
Cypripedium formosanum Id 305
Habenaria radiata Id 1947
Habenaria pantlingiana Id 307
Goodyera procera Id 623
Goodyera fumata Id 306
Goodyera schlechtendaliana Id 624
Ludisia discolor Id 1559
Anoectochilus emeiensis Id 1312
Goodyera velutina Id 625
Bletilla ochracea Id 662
Bletilla striata Id 548
Cattleya liliputana Id 1128
Cattleya crispata Id 297
Cymbidium mannii Id 2475
Cymbidium aloifolium Id 2471
Cymbidium tracyanum Id 2462
Cymbidium tortisepalum Id 2464
Cymbidium sinense Id 2461
Cymbidium kanran Id 696
Cymbidium faberi Id 491
Cymbidium ensifolium Id 564
Cymbidium goeringii Id 563
Cymbidium macrorhizon Id 698
Cymbidium lancifolium Id 697
Dendrobium huoshanense Id 1290
Dendrobium nobile Id 656
Dendrobium candidum Id 1904
Dendrobium catenatum Id 140
Dendrobium moniliforme Id 1625
Dendrobium pendulum Id 695
Dendrobium strongylanthum Id 487
Asparagales

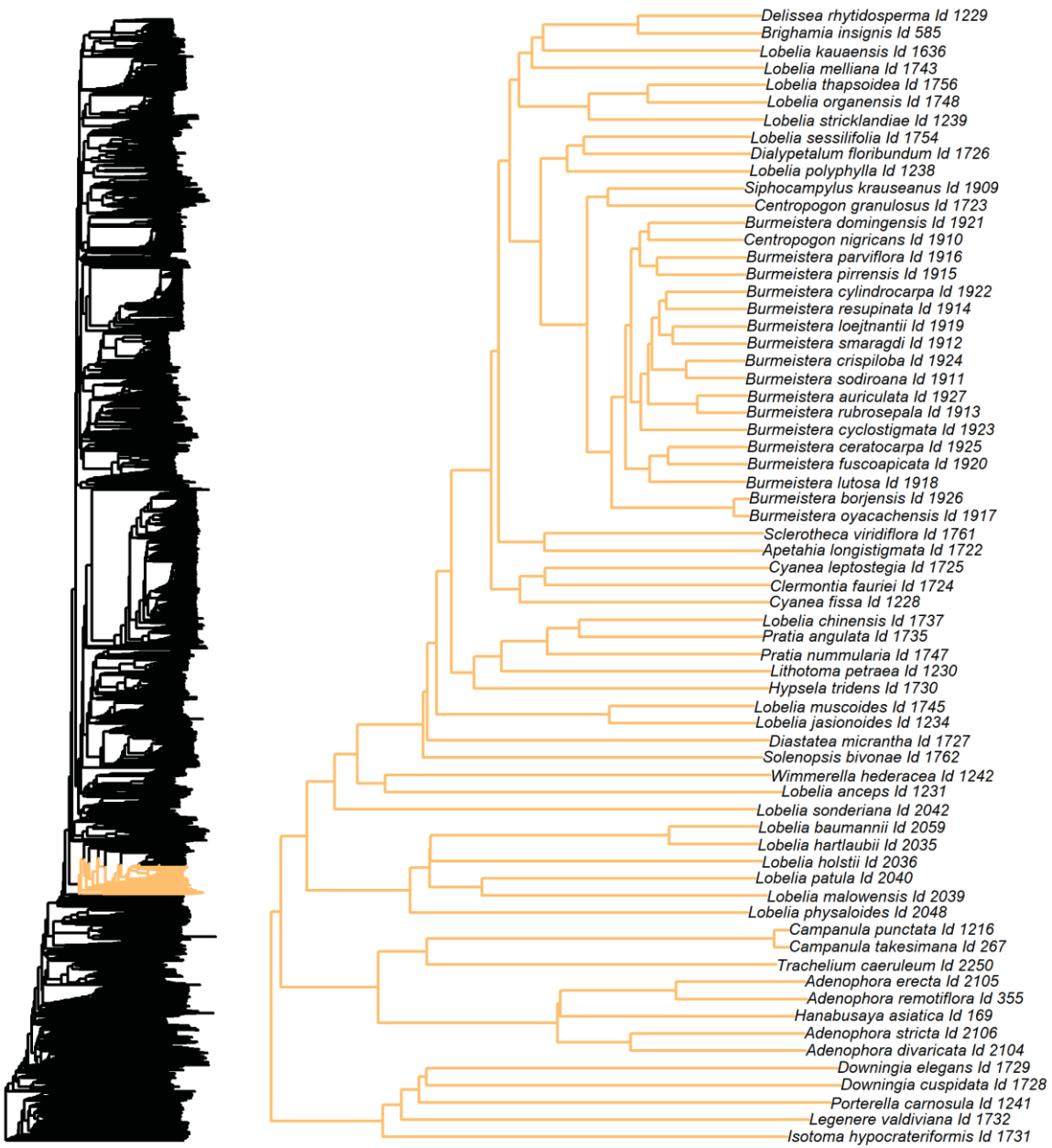

Dendrobium lohohense Id 1711
Dendrobium henryi Id 1709
Dendrobium chrysanthum Id 1633
Dendrobium crepidatum Id 1706
Dendrobium brymerianum Id 1631
Dendrobium fimbriatum Id 1715
Dendrobium denneanum Id 1699
Dendrobium parishii Id 1712
Dendrobium primulinum Id 1630
Dendrobium aphyllum Id 1698
Dendrobium ellipsophyllum Id 1713
Dendrobium fanjingshanense Id 1717
Dendrobium wilsonii Id 1705
Dendrobium xichouense Id 1714
Dendrobium hercoglossum Id 1703
Dendrobium falconeri Id 1701
Dendrobium devonianum Id 1700
Dendrobium wardianum Id 1704
Dendrobium gratiosissimum Id 1702
Dendrobium jenkinsii Id 1710
Dendrobium spatella Id 1708
Dendrobium parciflorum Id 1632
Dendrobium exile Id 1716
Dendrobium loddigesii Id 2118
Dendrobium salaccense Id 1707
Corallorhiza bulbosa Id 262
Corallorhiza macrantha Id 234
Corallorhiza wisteriana Id 246
Corallorhiza odontorhiza Id 236
Corallorhiza trifida Id 235
Gastrochilus japonicus Id 1946
Gastrochilus fuscopunctatus Id 1943
Pelatantheria scolopendrifolia Id 1942
Neofinetia richardsiana Id 2139
Neofinetia falcata Id 2138
Phalaenopsis hybrid cultivar Id 232
Phalaenopsis equestris Id 2374
Thrixspermum japonicum Id 1944
Phalaenopsis aphrodite subsp. formosana Id
Oncidium hybrid cultivar Id 2295
Oncidium sphacelatum Id 754
Erycina pusilla Id 1282
Oberonia japonica Id 1945
Dendrobium chrysotoxum Id 569
Paphiopedilum armeniacum Id 332
Paphiopedilum niveum Id 302
Phragmipedium longifolium Id 755
Cephalanthera humilis Id 941
Apostasia wallichii Id 2108
Apostasia odorata Id 1093
Acorus calamus Id 2389
Acorus americanus Id 2242
Acorus gramineus Id 552
Asparagales
Acorales

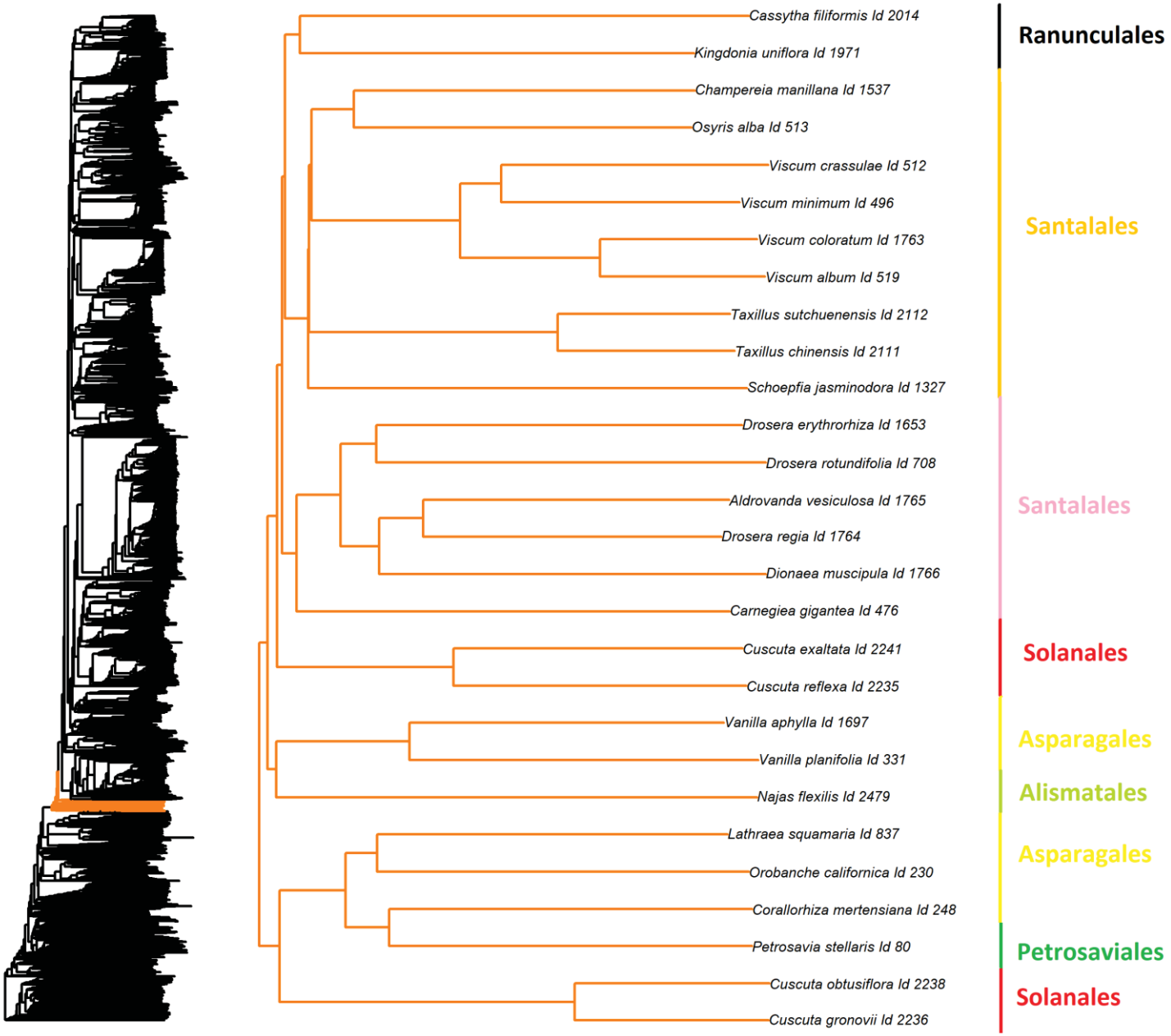

Machilus balansae Id 531
Machilus yunnanensis Id 530
Phoebe chekiangensis Id 1536
Phoebe sheareri Id 996
Phoebe zhennan Id 2096
Phoebe bournei Id 1491
Phoebe omeiensis Id 995
Persea americana Id 994
Machilus thunbergii Id 1696
Machilus pauhoi Id 1695
Laurus nobilis Id 1428
Lindera glauca Id 1987
Cinnamomum camphora Id 1975
Cinnamomum verum Id 1648
Cinnamomum micranthum Id 1938
Neocinnamomum delavayi Id 2016
Cryptocarya chinensis Id 2015
Calycanthus floridus var. glaucus Id 2321
Magnolia laevifolia Id 1989
Magnolia insignis Id 1856
Magnolia odora Id 682
Magnolia cathcartii Id 677
Magnolia sinica Id 684
Magnolia yunnanensis Id 208
Magnolia sprengeri Id 810
Magnolia salicifolia Id 683
Magnolia biondii Id 1415
Magnolia denudata Id 2165
Magnolia liliifera Id 681
Magnolia kobus Id 680
Magnolia grandiflora Id 2422
Magnolia kwangsiensis Id 2347
Magnolia pyramidata Id 679
Magnolia macrophylla var. dealbata Id 678
Magnolia officinalis subsp. biloba Id 2438
Magnolia officinalis Id 1287
Magnolia tripetala Id 551
Liriodendron tulipifera Id 2176
Liriodendron chinense Id 1549
Drimys granadensis Id 2197
Chloranthus spicatus Id 2229
Chloranthus japonicus Id 294
Illicium verum Id 1435
Illicium anisatum Id 1429
Illicium henryi Id 1439
Illicium floridanum Id 1432
Illicium oligandrum Id 2231
Schisandra chinensis Id 1530
Ceratophyllum demersum Id 2240
Cabomba caroliniana Id 1056
Brasenia schreberi Id 1014
Nymphaea alba Id 2380
Nymphaea mexicana Id 205
Victoria cruziana Id 1837
Nymphaea ampla Id 1867
Nymphaea jamesoniana Id 1083
Nuphar advena Id 2207
Barclaya longifolia Id 1838
Amborella trichopoda Id 2350
Piper cenocladum Id 2198
Piper auritum Id 1427
Piper nigrum Id 1416
Piper kadsura Id 508
Aristolochia debilis Id 2099
Aristolochia contorta Id 2098
Laurales
Magnoliales
Canellales
Chloranthales
Austrobaileyales
Ceratophyllales
Nymphaeales
Amborellales
Piperales

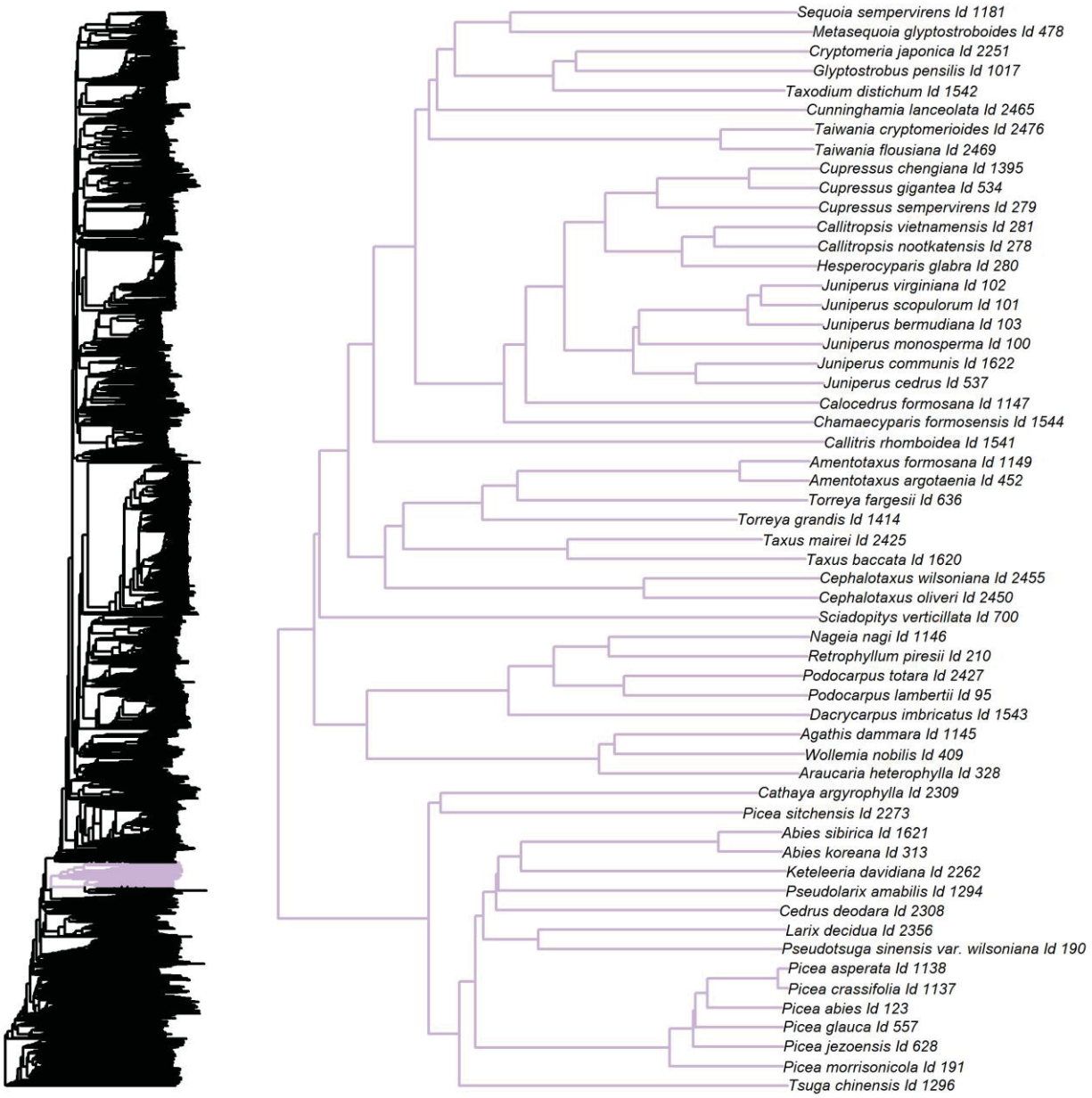

Megaleranthis saniculifolia Id 2272
Trollius chinensis Id 1086
Thalictrum coreanum Id 287
Aconitum kusnezoffii Id 1039
Aconitum ciliare Id 1037
Aconitum carmichaelii Id 959
Aconitum monanthum Id 1040
Aconitum austrokoreense Id 1034
Aconitum chiisanense Id 727
Aconitum coreanum Id 1038
Aconitum finetianum Id 2121
Aconitum angustius Id 2120
Aconitum sinomontanum Id 2122
Aconitum longecassidatum Id 1979
Aconitum pseudolaeve Id 1978
Gymnaconitum gymnandrum Id 1217
Clematis terniflora Id 514
Coptis chinensis Id 2140
Hydrastis canadensis Id 1441
Ranunculus macranthus Id 2208
Ranunculus occidentalis Id 1059
Stephania japonica Id 652
Buxus microphylla Id 2232
Pachysandra terminalis Id 653
Trochodendron aralioides Id 2460
Tetracentron sinense Id 2459
Nelumbo lutea Id 2334
Nelumbo nucifera Id 217
Platanus occidentalis Id 2179
Sabia yunnanensis Id 651
Meliosma aff. cuneifolia Moore 333 Id 650
Akebia quinata Id 1316
Akebia trifoliata Id 647
Decaisnea insignis Id 1982
Euptelea pleiosperma Id 649
Coreanomecon hylomeconoides Id 1055
Papaver somniferum Id 654
Macadamia ternifolia Id 2132
Macadamia integrifolia Id 218
Nandina domestica Id 2180
Gymnospermium microrrhynchum Id 799
Sinopodophyllum hexandrum Id 490
Berberis koreana Id 801
Berberis amurensis Id 800
Berberis bealei Id 9
Epimedium pseudowushanense Id 763
Epimedium lishihchenii Id 762
Epimedium acuminatum Id 759
Epimedium dolichostemon Id 760
Epimedium sagittatum Id 648
Epimedium koreanum Id 761
Ranunculales
Bruxales
Trochodendrales
Proteales
Ranunculales
Proteales
Ranunculales

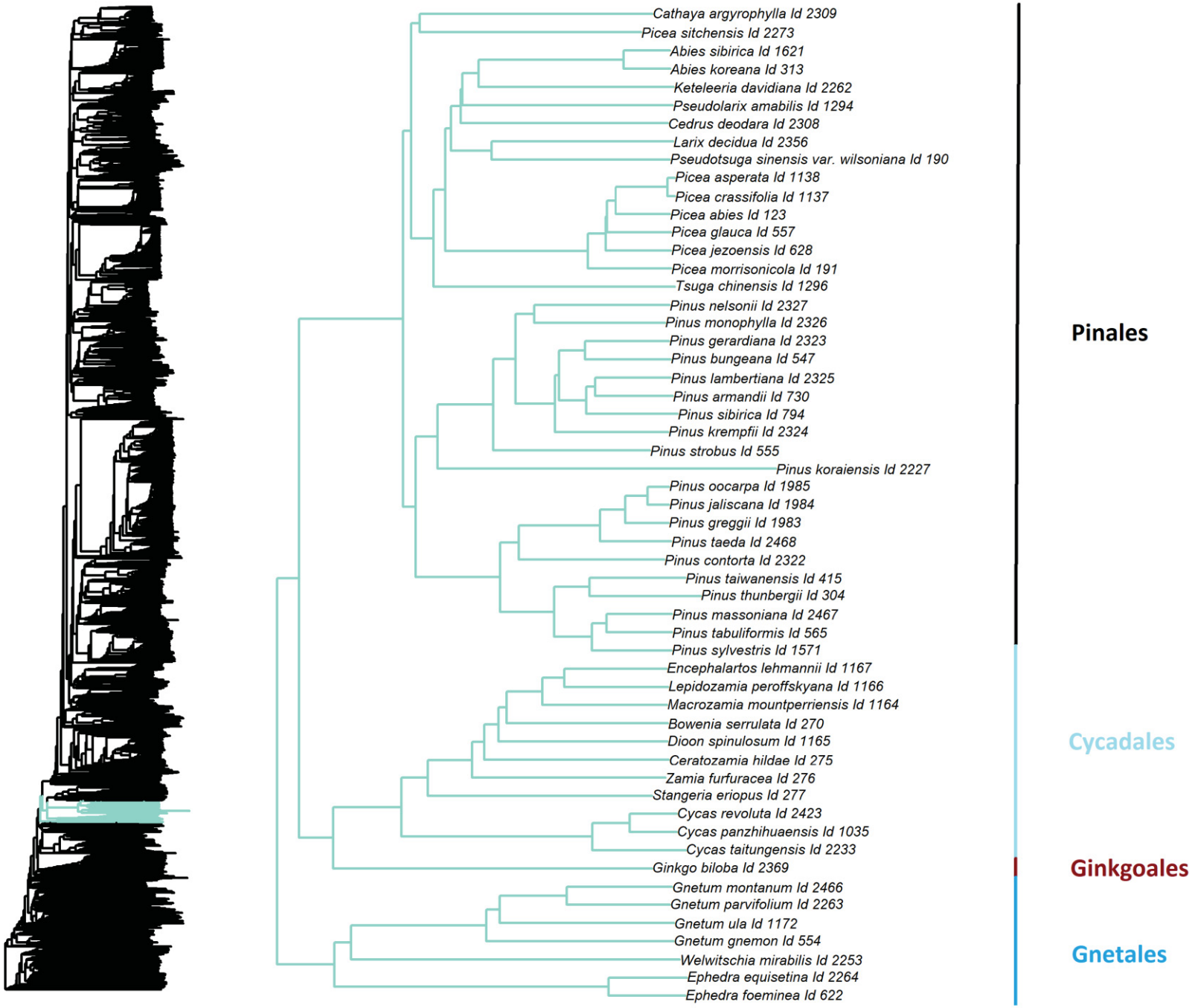

Vaccinium macrocarpon Id 2410
Passiflora edulis Id 1330
Viviania marifolia Id 115
Larrea tridentata Id 520
Geranium palmatum Id 2306
Geranium maderense Id 853
Geranium incanum Id 796
Erodium texanum Id 2305
Erodium crassifolium Id 257
Erodium carvifolium Id 2319
Erodium manescavi Id 948
Erodium rupestre Id 947
Erodium reichardii Id 949
Erodium trifolium Id 161
Erodium chrysanthum Id 357
Erodium absinthoides Id 318
Erodium gruinum Id 258
Monsonia marlothii Id 809
Monsonia emarginata Id 808
Monsonia speciosa Id 2307
Pelargonium spinosum Id 1227
Pelargonium endlicherianum Id 1225
Pelargonium transvaalense Id 1009
Pelargonium x hortorum Id 2196
Pelargonium quinquelobatum Id 1226
Pelargonium tetragonum Id 1008
Pelargonium exhibens Id 1000
Pelargonium myrrhifolium Id 1005
Pelargonium echinatum Id 1001
Pelargonium alternans Id 117
Pelargonium incrassatum Id 1004
Pelargonium fulgidum Id 1003
Pelargonium quercifolium Id 1006
Pelargonium nanum Id 971
Pelargonium citronellum Id 998
Pelargonium cucullatum Id 997
Pelargonium australe Id 527
Pelargonium cotyledonis Id 526
Pelargonium exstipulatum Id 1002
Pelargonium dichondrifolium Id 525
Pelargonium trifidum Id 1007
Pelargonium dolomiticum Id 999
Hypseocharis bilobata Id 116
Linum usitatissimum Id 2119
Trithuria inconspicua Id 2435
Annona cherimola Id 813
Circaeaster agrestis Id 1970
Ericales
Malpighiales
Geraniales
Zygophyllales
Geraniales
Nymphaeales
Geraniales

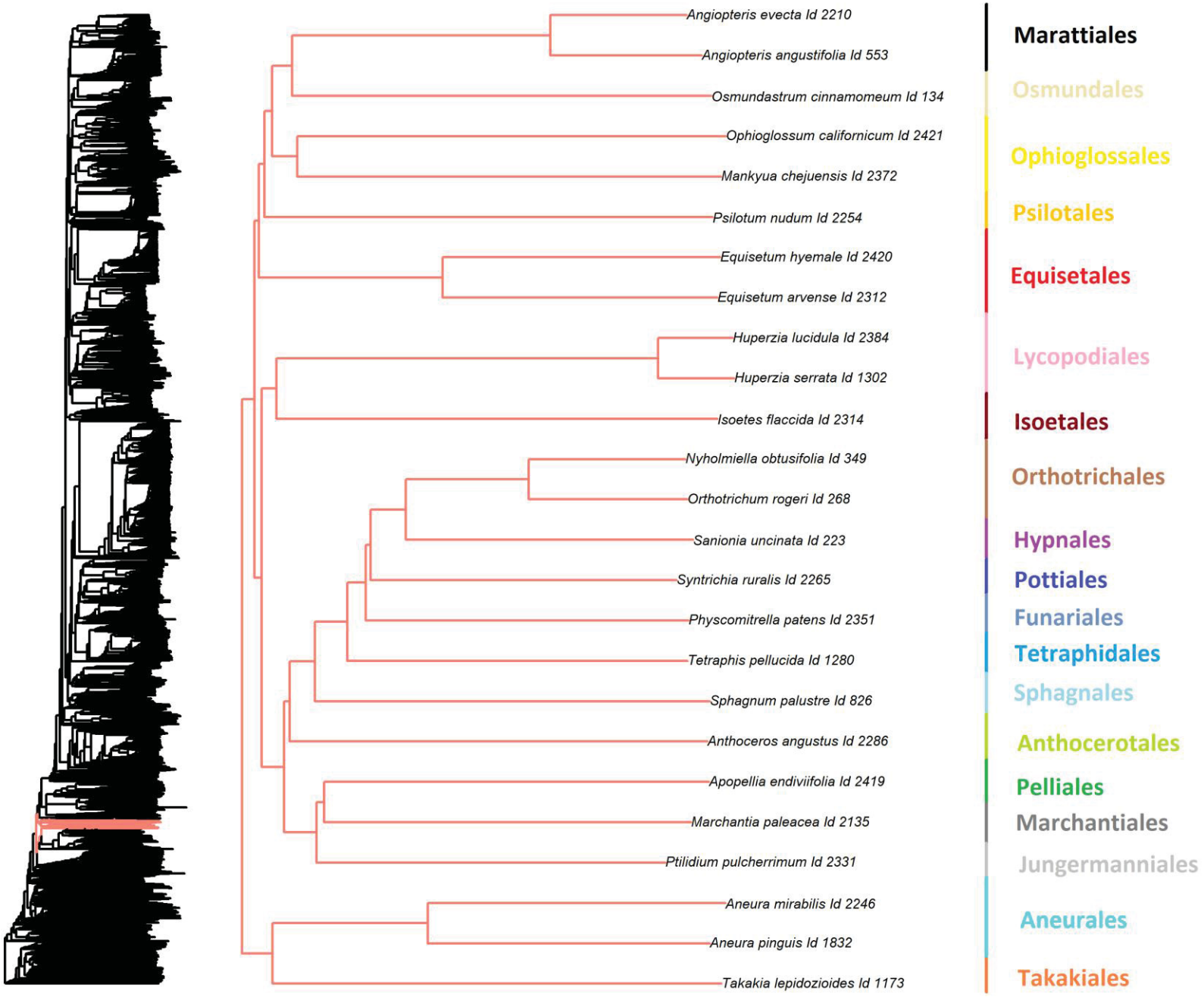

Codonopsis minima Id 2113
Platycodon grandiflorus Id 1835
Cyphia glandulifera Id 2054
Cyphia crenata Id 2052
Cyphia phyteuma Id 2055
Cyphia banksiana Id 2050
Cyphia tortilis Id 2057
Cyphia belfastica Id 2051
Cyphia angustiloba Id 2049
Cyphia schlechteri Id 2056
Cyphia dentariifolia Id 2053
Grammatotheca bergiana Id 2058
Lobelia thermalis Id 2044
Monopsis flava Id 2046
Monopsis alba Id 2045
Lobelia linearis Id 2038
Lobelia erinus Id 2061
Lobelia laxa Id 2037
Lobelia galpinii Id 2034
Lobelia cardinalis Id 2060
Lobelia puberula Id 2041
Lobelia spicata Id 2043
Lobelia inflata Id 1233
Hippobroma longiflora Id 1634
Lobelia laxiflora Id 1235
Lobelia urens Id 1758
Lobelia sancta Id 1752
Lobelia ritabeaniana Id 1751
Lobelia udzungwensis Id 1757
Lobelia longisepala Id 1236
Lobelia lukwangulensis Id 1742
Lobelia morogoroensis Id 1237
Lobelia seguinii Id 1753
Lobelia davidii Id 1739
Lobelia doniana Id 1740
Lobelia yuccoides Id 1760
Lobelia niihauensis Id 1746
Lobelia wollastonii Id 1759
Lobelia stuhlmannii Id 1637
Lobelia bequaertii Id 1635
Lobelia gibberoa Id 1741
Lobelia columnaris Id 1738
Lobelia bambuseti Id 1736
Lobelia aberdarica Id 1733
Lobelia telekii Id 1755
Lobelia rhynchopetalum Id 1750
Lobelia acrochila Id 1734
Lobelia mildbraedii Id 1744
Lobelia thuliniana Id 1240
Lobelia petiolata Id 1749
Trematolobelia kauaiensis Id 1638
Lobelia boninensis Id 1232
Asterales

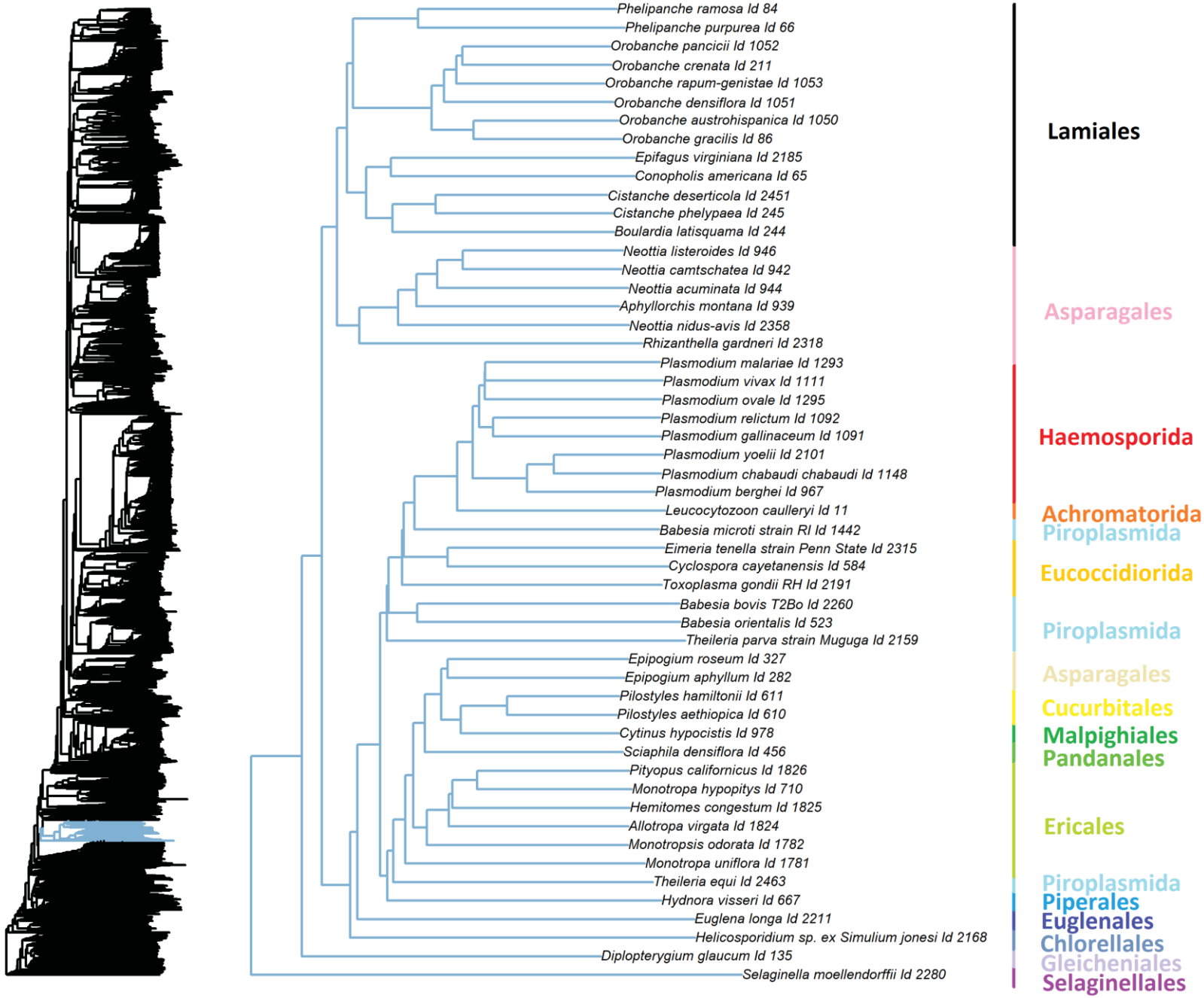

Delissea rhytidosperma Id 1229
Brighamia insignis Id 585
Lobelia kauaensis Id 1636
Lobelia melliana Id 1743
Lobelia thapsoidea Id 1756
Lobelia organensis Id 1748
Lobelia stricklandiae Id 1239
Lobelia sessilifolia Id 1754
Dialypetalum floribundum Id 1726
Lobelia polyphylla Id 1238
Siphocampylus krauseanus Id 1909
Centropogon granulosus Id 1723
Burmeistera domingensis Id 1921
Centropogon nigricans Id 1910
Burmeistera parviflora Id 1916
Burmeistera pirrensis Id 1915
Burmeistera cylindrocarpa Id 1922
Burmeistera resupinata Id 1914
Burmeistera loejtnantii Id 1919
Burmeistera smaragdi Id 1912
Burmeistera crispiloba Id 1924
Burmeistera sodiroana Id 1911
Burmeistera auriculata Id 1927
Burmeistera rubrosepala Id 1913
Burmeistera cyclostigmata Id 1923
Burmeistera ceratocarpa Id 1925
Burmeistera fuscoapicata Id 1920
Burmeistera lutosa Id 1918
Burmeistera borjensis Id 1926
Burmeistera oyacachensis Id 1917
Sclerotheca viridiflora Id 1761
Apetahia longistigmata Id 1722
Cyanea leptostegia Id 1725
Clermontia fauriei Id 1724
Cyanea fissa Id 1228
Lobelia chinensis Id 1737
Pratia angulata Id 1735
Pratia nummularia Id 1747
Lithotoma petraea Id 1230
Hypsela tridens Id 1730
Lobelia muscoides Id 1745
Lobelia jasionoides Id 1234
Diastatea micrantha Id 1727
Solenopsis bivonae Id 1762
Wimmerella hederacea Id 1242
Lobelia anceps Id 1231
Lobelia sonderiana Id 2042
Lobelia baumannii Id 2059
Lobelia hartlaubii Id 2035
Lobelia holstii Id 2036
Lobelia patula Id 2040
Lobelia malowensis Id 2039
Lobelia physaloides Id 2048
Campanula punctata Id 1216
Campanula takesimana Id 267
Trachelium caeruleum Id 2250
Adenophora erecta Id 2105
Adenophora remotiflora Id 355
Hanabusaya asiatica Id 169
Adenophora stricta Id 2106
Adenophora divaricata Id 2104
Downingia elegans Id 1729
Downingia cuspidata Id 1728
Porterella carnosula Id 1241
Legenere valdiviana Id 1732
Isotoma hypocrateriformis Id 1731
Asterales

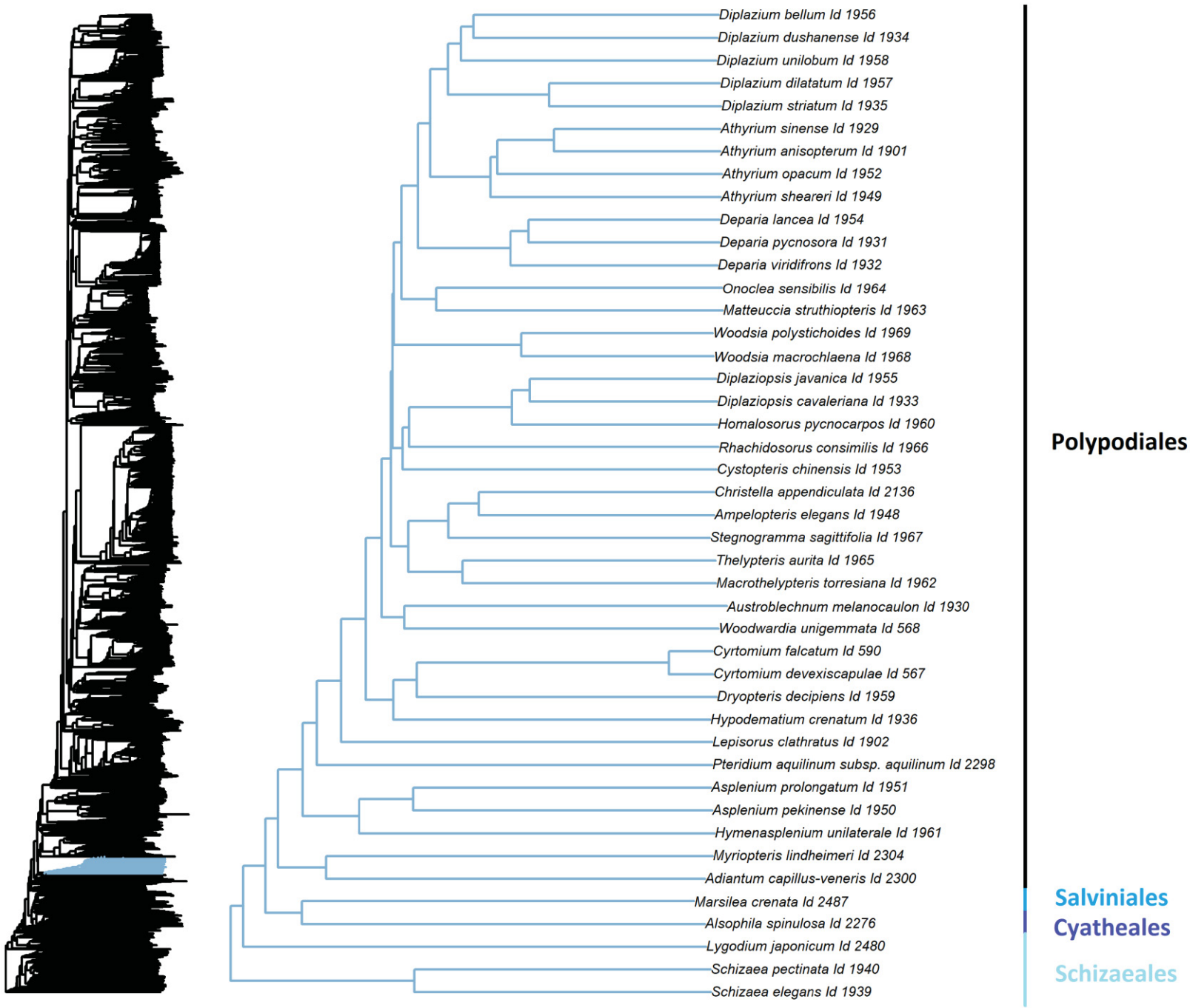

Cassytha filiformis Id 2014
Kingdonia uniflora Id 1971
Champereia manillana Id 1537
Osyris alba Id 513
Viscum crassulae Id 512
Viscum minimum Id 496
Viscum coloratum Id 1763
Viscum album Id 519
Taxillus sutchuenensis Id 2112
Taxillus chinensis Id 2111
Schoepfia jasminodora Id 1327
Drosera erythrorhiza Id 1653
Drosera rotundifolia Id 708
Aldrovanda vesiculosa Id 1765
Drosera regia Id 1764
Dionaea muscipula Id 1766
Carnegiea gigantea Id 476
Cuscuta exaltata Id 2241
Cuscuta reflexa Id 2235
Vanilla aphylla Id 1697
Vanilla planifolia Id 331
Najas flexilis Id 2479
Lathraea squamaria Id 837
Orobanche californica Id 230
Corallorhiza mertensiana Id 248
Petrosavia stellaris Id 80
Cuscuta obtusiflora Id 2238
Cuscuta gronovii Id 2236
Ranunculales
Santalales
Santalales
Solanales
Asparagales
Alismatales
Asparagales
Petrosaviales
Solanales


Sequoia sempervirens Id 1181
Metasequoia glyptostroboides Id 478
Cryptomeria japonica Id 2251
Glyptostrobus pensilis Id 1017
Taxodium distichum Id 1542
Cunninghamia lanceolata Id 2465
Taiwania cryptomerioides Id 2476
Taiwania flousiana Id 2469
Cupressus chengiana Id 1395
Cupressus gigantea Id 534
Cupressus sempervirens Id 279
Callitropsis vietnamensis Id 281
Callitropsis nootkatensis Id 278
Hesperocyparis glabra Id 280
Juniperus virginiana Id 102
Juniperus scopulorum Id 101
Juniperus bermudiana Id 103
Juniperus monosperma Id 100
Juniperus communis Id 1622
Juniperus cedrus Id 537
Calocedrus formosana Id 1147
Chamaecyparis formosensis Id 1544
Callitris rhomboidea Id 1541
Amentotaxus formosana Id 1149
Amentotaxus argotaenia Id 452
Torreya fargesii Id 636
Torreya grandis Id 1414
Taxus mairei Id 2425
Taxus baccata Id 1620
Cephalotaxus wilsoniana Id 2455
Cephalotaxus oliveri Id 2450
Sciadopitys verticillata Id 700
Nageia nagi Id 1146
Retrophyllum piresii Id 210
Podocarpus totara Id 2427
Podocarpus lambertii Id 95
Dacrycarpus imbricatus Id 1543
Agathis dammara Id 1145
Wollemia nobilis Id 409
Araucaria heterophylla Id 328
Cathaya argyrophylla Id 2309
Picea sitchensis Id 2273
Abies sibirica Id 1621
Abies koreana Id 313
Keteleeria davidiana Id 2262
Pseudolarix amabilis Id 1294
Cedrus deodara Id 2308
Larix decidua Id 2356
Pseudotsuga sinensis var. wilsoniana Id 190
Picea asperata Id 1138
Picea crassifolia Id 1137
Picea abies Id 123
Picea glauca Id 557
Picea jezoensis Id 628
Picea morrisonicola Id 191
Tsuga chinensis Id 1296
Pinales

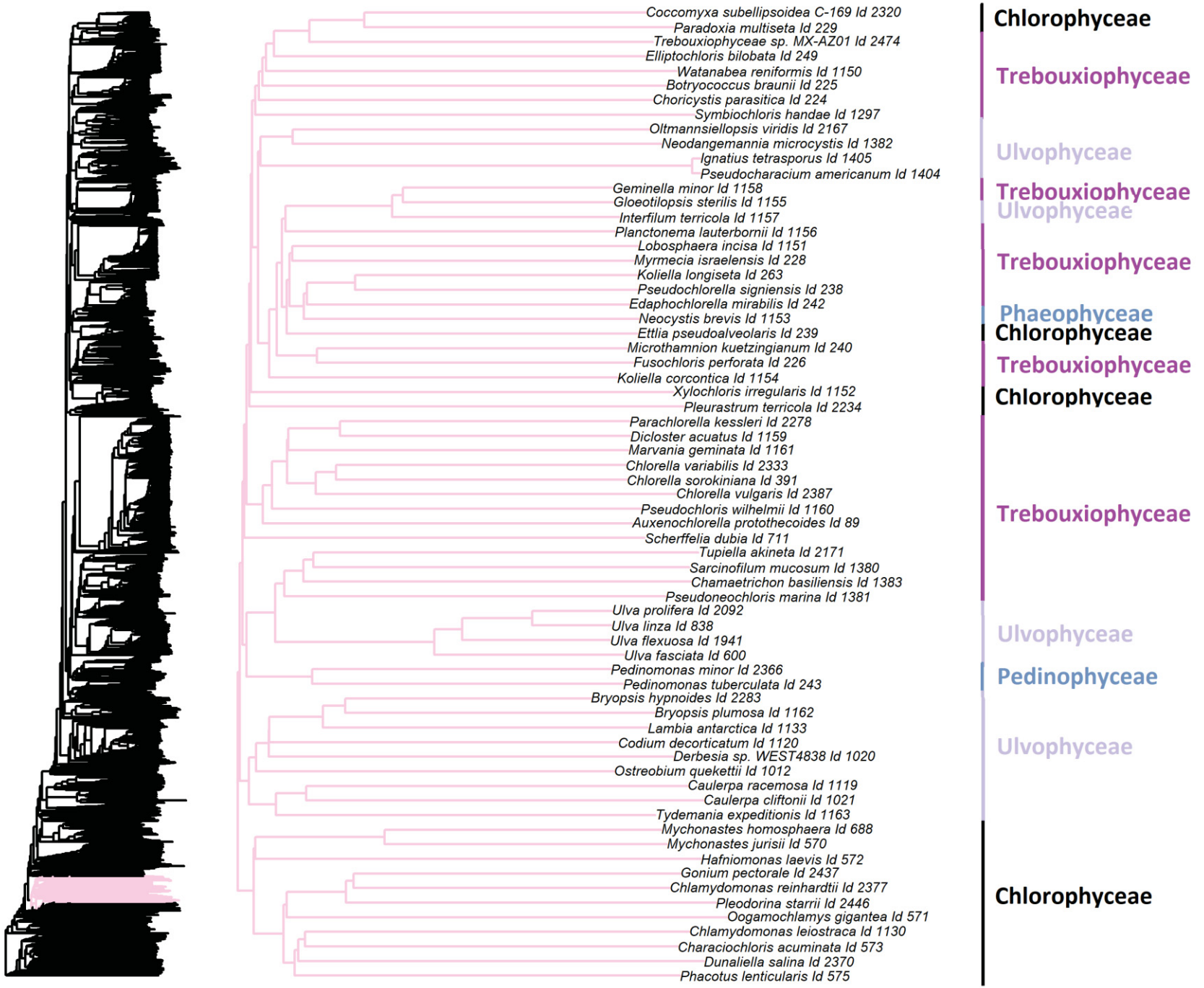

Cathaya argyrophylla Id 2309
Picea sitchensis Id 2273
Abies sibirica Id 1621
Abies koreana Id 313
Keteleeria davidiana Id 2262
Pseudolarix amabilis Id 1294
Cedrus deodara Id 2308
Larix decidua Id 2356
Pseudotsuga sinensis var. wilsoniana Id 190
Picea asperata Id 1138
Picea crassifolia Id 1137
Picea abies Id 123
Picea glauca Id 557
Picea jezoensis Id 628
Picea morrisonicola Id 191
Tsuga chinensis Id 1296
Pinus nelsonii Id 2327
Pinus monophylla Id 2326
Pinus gerardiana Id 2323
Pinus bungeana Id 547
Pinus lambertiana Id 2325
Pinus armandii Id 730
Pinus sibirica Id 794
Pinus krempfii Id 2324
Pinus strobus Id 555
Pinus koraiensis Id 2227
Pinus oocarpa Id 1985
Pinus jaliscana Id 1984
Pinus greggii Id 1983
Pinus taeda Id 2468
Pinus contorta Id 2322
Pinus taiwanensis Id 415
Pinus thunbergii Id 304
Pinus massoniana Id 2467
Pinus tabuliformis Id 565
Pinus sylvestris Id 1571
Encephalartos lehmannii Id 1167
Lepidozamia peroffskyana Id 1166
Macrozamia mountperriensis Id 1164
Bowenia serrulata Id 270
Dioon spinulosum Id 1165
Ceratozamia hildae Id 275
Zamia furfuracea Id 276
Stangeria eriopus Id 277
Cycas revoluta Id 2423
Cycas panzhihuaensis Id 1035
Cycas taitungensis Id 2233
Ginkgo biloba Id 2369
Gnetum montanum Id 2466
Gnetum parvifolium Id 2263
Gnetum ula Id 1172
Gnetum gnemon Id 554
Welwitschia mirabilis Id 2253
Ephedra equisetina Id 2264
Ephedra foeminea Id 622
Pinales
Cycadales
Ginkgoales
Gnetales


Angiopteris evecta Id 2210
Angiopteris angustifolia Id 553
Osmundastrum cinnamomeum Id 134
Ophioglossum californicum Id 2421
Mankyua chejuensis Id 2372
Psilotum nudum Id 2254
Equisetum hyemale Id 2420
Equisetum arvense Id 2312
Huperzia lucidula Id 2384
Huperzia serrata Id 1302
Isoetes flaccida Id 2314
Nyholmiella obtusifolia Id 349
Orthotrichum rogeri Id 268
Sanionia uncinata Id 223
Syntrichia ruralis Id 2265
Physcomitrella patens Id 2351
Tetraphis pellucida Id 1280
Sphagnum palustre Id 826
Anthoceros angustus Id 2286
Apopellia endiviifolia Id 2419
Marchantia paleacea Id 2135
Ptilidium pulcherrimum Id 2331
Aneura mirabilis Id 2246
Aneura pinguis Id 1832
Takakia lepidozioides Id 1173
Marattiales
Osmundales
Ophioglossales
Psilotales
Equisetales
Lycopodiales
Isoetales
Orthotrichales
Hypnales
Pottiales
Funariales
Tetraphidales
Sphagnales
Anthocerotales
Pelliales
Marchantiales
Jungermanniales
Aneurales
Takakiales

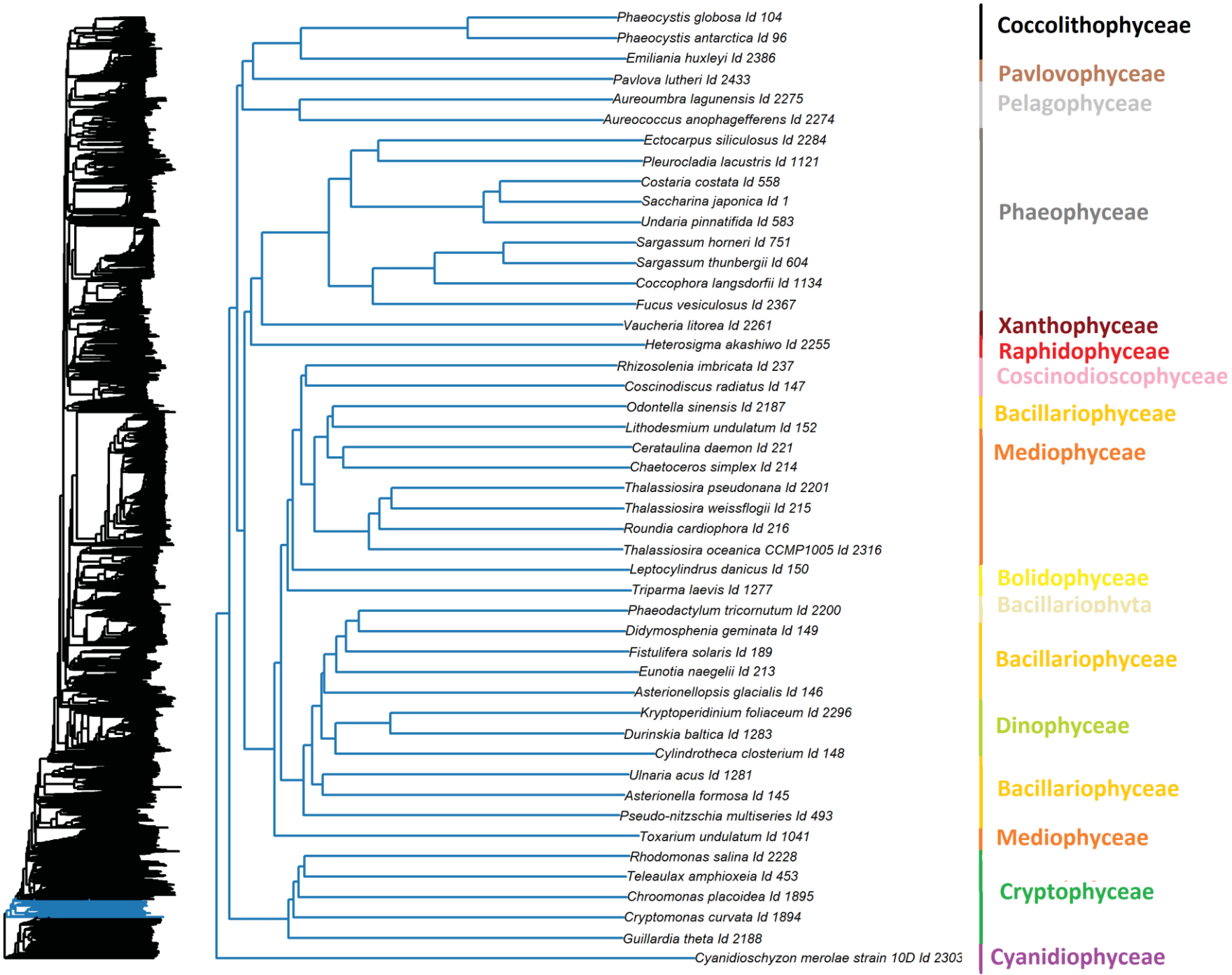

Phelipanche ramosa Id 84
Phelipanche purpurea Id 66
Orobanche pancicii Id 1052
Orobanche crenata Id 211
Orobanche rapum-genistae Id 1053
Orobanche densiflora Id 1051
Orobanche austrohispanica Id 1050
Orobanche gracilis Id 86
Epifagus virginiana Id 2185
Conopholis americana Id 65
Cistanche deserticola Id 2451
Cistanche phelypaea Id 245
Boulardia latisquama Id 244
Neottia listeroides Id 946
Neottia camtschatea Id 942
Neottia acuminata Id 944
Aphyllorchis montana Id 939
Neottia nidus-avis Id 2358
Rhizanthella gardneri Id 2318
Plasmodium malariae Id 1293
Plasmodium vivax Id 1111
Plasmodium ovale Id 1295
Plasmodium relictum Id 1092
Plasmodium gallinaceum Id 1091
Plasmodium yoelii Id 2101
Plasmodium chabaudi chabaudi Id 1148
Plasmodium berghei Id 967
Leucocytozoon caulleryi Id 11
Babesia microti strain RI Id 1442
Eimeria tenella strain Penn State Id 2315
Cyclospora cayetanensis Id 584
Toxoplasma gondii RH Id 2191
Babesia bovis T2Bo Id 2260
Babesia orientalis Id 523
Theileria parva strain Muguga Id 2159
Epipogium roseum Id 327
Epipogium aphyllum Id 282
Pilostyles hamiltonii Id 611
Pilostyles aethiopica Id 610
Cytinus hypocistis Id 978
Sciaphila densiflora Id 456
Pityopus californicus Id 1826
Monotropa hypopitys Id 710
Hemitomes congestum Id 1825
Allotropa virgata Id 1824
Monotropsis odorata Id 1782
Monotropa uniflora Id 1781
Theileria equi Id 2463
Hydnora visseri Id 667
Euglena longa Id 2211
Helicosporidium sp. ex Simulium jonesi Id 2168
Diplopterygium glaucum Id 135
Selaginella moellendorffii Id 2280
Lamiales
Asparagales
Haemosporida
Achromatorida
Piroplasmida
Eucoccidiorida
Piroplasmida
Asparagales
Cucurbitales
Malpighiales
Pandanales
Ericales
Piroplasmida
Piperales
Euglenales
Chlorellales
Gleicheniales
Selaginellales

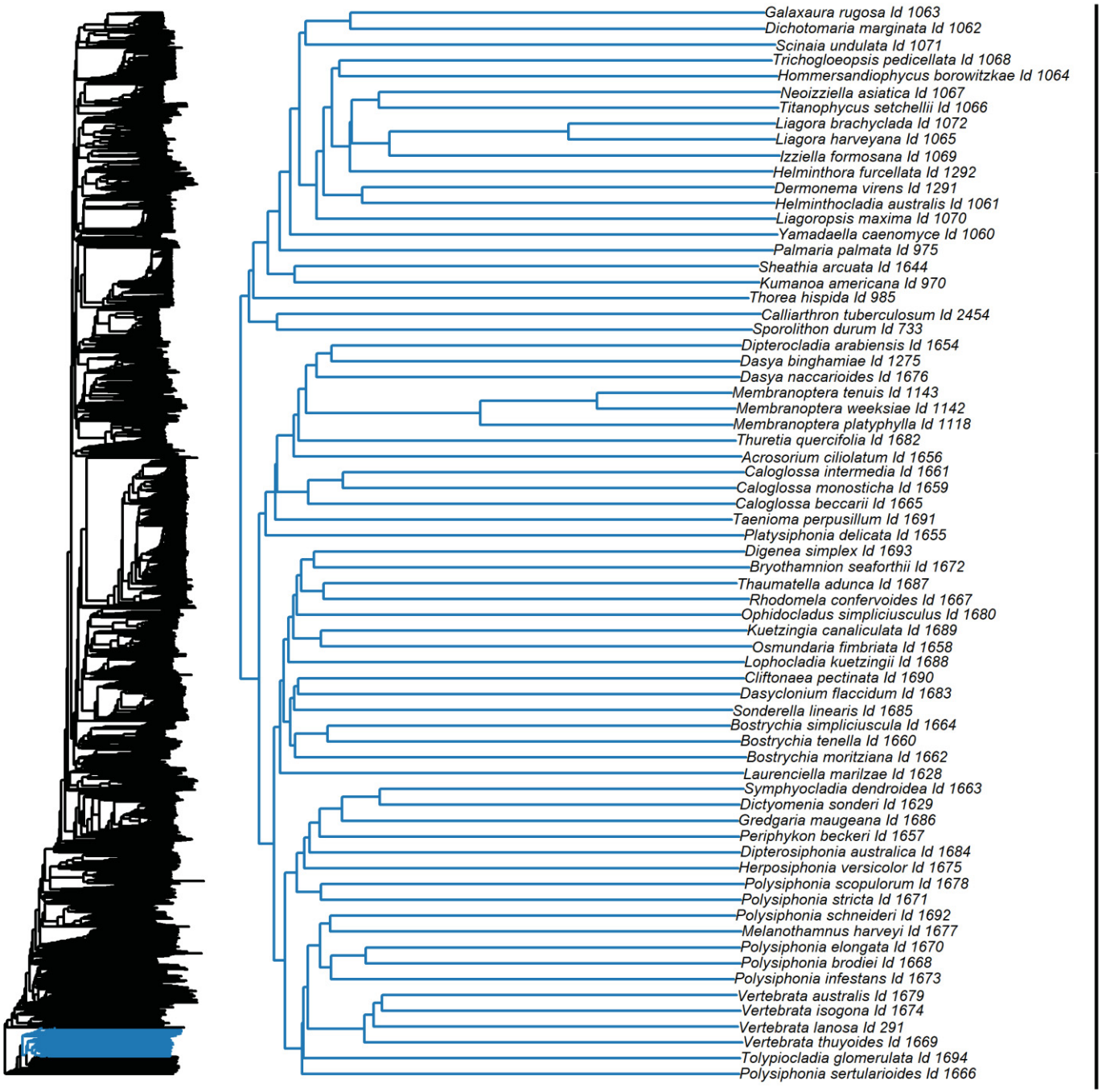

Diplazium bellum Id 1956
Diplazium dushanense Id 1934
Diplazium unilobum Id 1958
Diplazium dilatatum Id 1957
Diplazium striatum Id 1935
Athyrium sinense Id 1929
Athyrium anisopterum Id 1901
Athyrium opacum Id 1952
Athyrium sheareri Id 1949
Deparia lancea Id 1954
Deparia pycnosora Id 1931
Deparia viridifrons Id 1932
Onoclea sensibilis Id 1964
Matteuccia struthiopteris Id 1963
Woodsia polystichoides Id 1969
Woodsia macrochlaena Id 1968
Diplaziopsis javanica Id 1955
Diplaziopsis cavaleriana Id 1933
Homalosorus pycnocarpos Id 1960
Rhachidosorus consimilis Id 1966
Cystopteris chinensis Id 1953
Christella appendiculata Id 2136
Ampelopteris elegans Id 1948
Stegnogramma sagittifolia Id 1967
Thelypteris aurita Id 1965
Macrothelypteris torresiana Id 1962
Austroblechnum melanocaulon Id 1930
Woodwardia unigemmata Id 568
Cyrtomium falcatum Id 590
Cyrtomium devexiscapulae Id 567
Dryopteris decipiens Id 1959
Hypodematium crenatum Id 1936
Lepisorus clathratus Id 1902
Pteridium aquilinum subsp. aquilinum Id 2298
Asplenium prolongatum Id 1951
Asplenium pekinense Id 1950
Hymenasplenium unilaterale Id 1961
Myriopteris lindheimeri Id 2304
Adiantum capillus-veneris Id 2300
Marsilea crenata Id 2487
Alsophila spinulosa Id 2276
Lygodium japonicum Id 2480
Schizaea pectinata Id 1940
Schizaea elegans Id 1939
Polypodiales
Salviniales
Cyatheales
Schizaeales

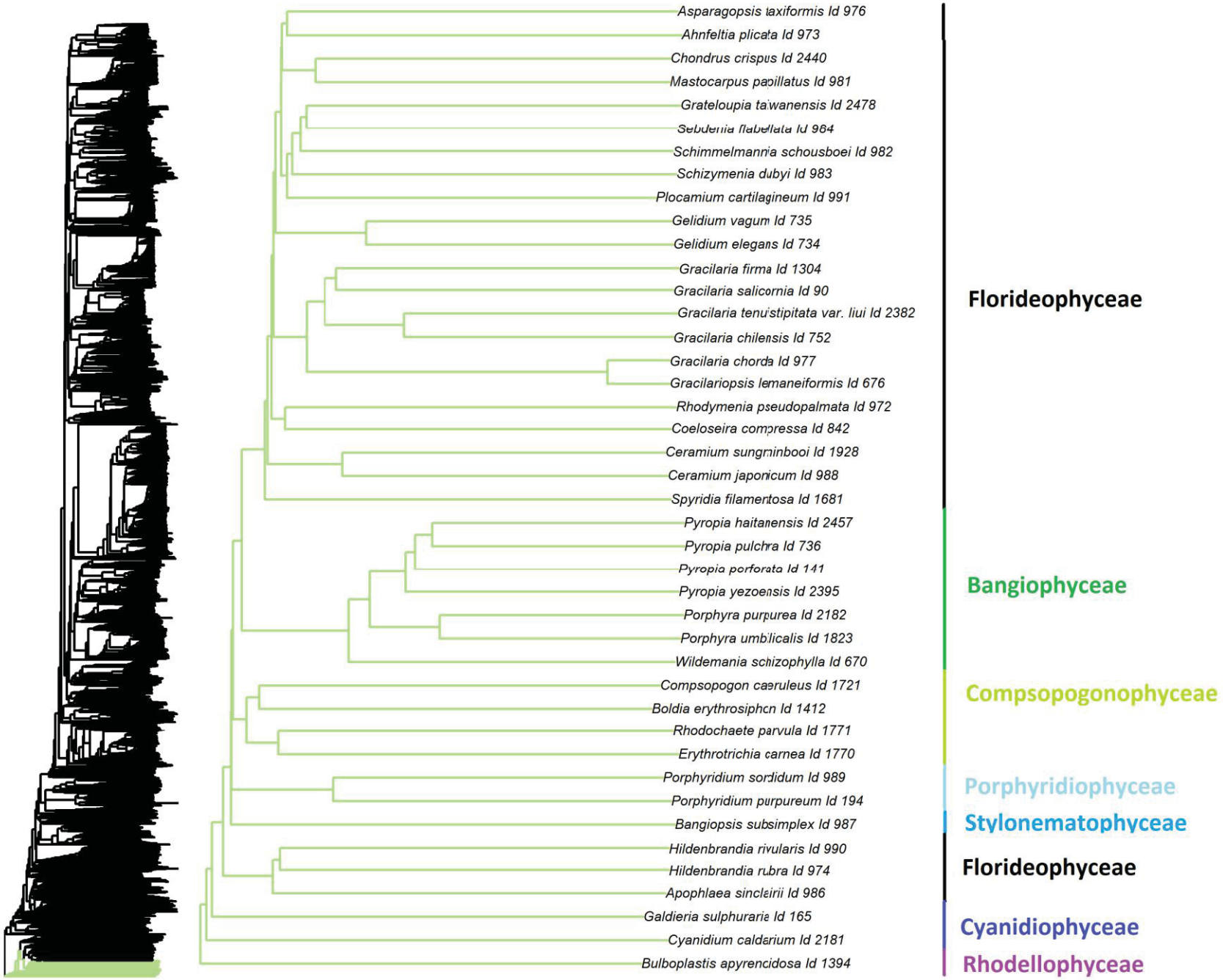

Chaetosphaeridium globosum Id 2271
Coleochaete scutata Id 1179
Chara vulgaris Id 2166
Zygnema circumcarinatum Id 2169
Cylindrocystis brebissonii Id 1180
Staurastrum punctulatum Id 2170
Cosmarium botrytis Id 1178
Closterium baillyanum Id 852
Roya obtusa Id 1175
Roya anglica Id 136
Netrium digitus Id 1177
Spirogyra maxima Id 1176
Mesotaenium endlicherianum Id 137
Entransia fimbriata Id 1174
Klebsormidium flaccidum Id 1013
Pyramimonas parkeae Id 2266
Cymbomonas tetramitiformis Id 816
Picocystis salinarum Id 394
Nephroselmis olivacea Id 2189
Nephroselmis astigmatica Id 395
Chlorokybus atmophyticus Id 2209
Mesostigma viride Id 2183
Palmophyllum crassum Id 1243
Verdigellas peltata Id 829
Pycnococcus provasolii Id 2267
Prasinoderma coloniale Id 393
Monomorphina aenigmatica Id 2413
Monomorphina parapyrum Id 405
Euglena archaeoplastidiata Id 1626
Euglenaria anabaena Id 412
Euglena viridis Id 356
Cryptoglena skujai Id 404
Trachelomonas volvocina Id 406
Euglenaformis proxima Id 371
Euglena gracilis Id 2301
Eutreptiella gymnastica Id 2409
Micromonas commoda Id 2270
Ostreococcus tauri Id 2175
Bathycoccus prasinos Id 586
Monomastix sp. OKE-1 Id 2268
Bigelowiella natans Id 2195
Partenskyella glossopodia Id 702
Gymnochlora stellata Id 709
Lotharella vacuolata Id 703
Lepidodinium chlorophorum Id 366
Coleochaetophyceae
Charophyceae
Zygnematophyceae
Klebsormidiophyceae
Pyramimonadophyceae
Picocystophyceae
Nefrophyceae
Chlorokybophyceae
Mesostigmatophyceae
Palmophyllophyceae
Pyramimonadophyceae
Palmophyllophyceae
Euglenophyceae
Mamiellophyceae
Chlorarachniophyceae
Dinophyceae


Coccomyxa subellipsoidea C-169 Id 2320
Paradoxia multiseta Id 229
Trebouxiophyceae sp. MX-AZ01 Id 2474
Elliptochloris bilobata Id 249
Watanabea reniformis Id 1150
Botryococcus braunii Id 225
Choricystis parasitica Id 224
Symbiochloris handae Id 1297
Oltmannsiellopsis viridis Id 2167
Neodangemannia microcystis Id 1382
Ignatius tetrasporus Id 1405
Pseudocharacium americanum Id 1404
Geminella minor Id 1158
Gloeotilopsis sterilis Id 1155
Interfilum terricola Id 1157
Planctonema lauterbornii Id 1156
Lobosphaera incisa Id 1151
Myrmecia israelensis Id 228
Koliella longiseta Id 263
Pseudochlorella signiensis Id 238
Edaphochlorella mirabilis Id 242
Neocystis brevis Id 1153
Ettlia pseudoalveolaris Id 239
Microthamnion kuetzingianum Id 240
Fusochloris perforata Id 226
Koliella corcontica Id 1154
Xylochloris irregularis Id 1152
Pleurastrum terricola Id 2234
Parachlorella kessleri Id 2278
Dicloster acuatus Id 1159
Marvania geminata Id 1161
Chlorella variabilis Id 2333
Chlorella sorokiniana Id 391
Chlorella vulgaris Id 2387
Pseudochloris wilhelmii Id 1160
Auxenochlorella protothecoides Id 89
Scherffelia dubia Id 711
Tupiella akineta Id 2171
Sarcinofilum mucosum Id 1380
Chamaetrichon basiliensis Id 1383
Pseudoneochloris marina Id 1381
Ulva prolifera Id 2092
Ulva linza Id 838
Ulva flexuosa Id 1941
Ulva fasciata Id 600
Pedinomonas minor Id 2366
Pedinomonas tuberculata Id 243
Bryopsis hypnoides Id 2283
Bryopsis plumosa Id 1162
Lambia antarctica Id 1133
Codium decorticatum Id 1120
Derbesia sp. WEST4838 Id 1020
Ostreobium quekettii Id 1012
Caulerpa racemosa Id 1119
Caulerpa cliftonii Id 1021
Tydemania expeditionis Id 1163
Mychonastes homosphaera Id 688
Mychonastes jurisii Id 570
Hafniomonas laevis Id 572
Gonium pectorale Id 2437
Chlamydomonas reinhardtii Id 2377
Pleodorina starrii Id 2446
Oogamochlamys gigantea Id 571
Chlamydomonas leiostraca Id 1130
Characiochloris acuminata Id 573
Dunaliella salina Id 2370
Phacotus lenticularis Id 575
Chlorophyceae
Trebouxiophyceae
Ulvophyceae
Trebouxiophyceae
Ulvophyceae
Trebouxiophyceae
Phaeophyceae
Chlorophyceae
Trebouxiophyceae
Chlorophyceae
Trebouxiophyceae
Ulvophyceae
Pedinophyceae
Ulvophyceae
Chlorophyceae

5.


Phaeocystis globosa Id 104
Phaeocystis antarctica Id 96
Emiliania huxleyi Id 2386
Pavlova lutheri Id 2433
Aureoumbra lagunensis Id 2275
Aureococcus anophagefferens Id 2274
Ectocarpus siliculosus Id 2284
Pleurocladia lacustris Id 1121
Costaria costata Id 558
Saccharina japonica Id 1
Undaria pinnatifida Id 583
Sargassum horneri Id 751
Sargassum thunbergii Id 604
Coccophora langsdorfii Id 1134
Fucus vesiculosus Id 2367
Vaucheria litorea Id 2261
Heterosigma akashiwo Id 2255
Rhizosolenia imbricata Id 237
Coscinodiscus radiatus Id 147
Odontella sinensis Id 2187
Lithodesmium undulatum Id 152
Cerataulina daemon Id 221
Chaetoceros simplex Id 214
Thalassiosira pseudonana Id 2201
Thalassiosira weissflogii Id 215
Roundia cardiophora Id 216
Thalassiosira oceanica CCMP1005 Id 2316
Leptocylindrus danicus Id 150
Triparma laevis Id 1277
Phaeodactylum tricornutum Id 2200
Didymosphenia geminata Id 149
Fistulifera solaris Id 189
Eunotia naegelii Id 213
Asterionellopsis glacialis Id 146
Kryptoperidinium foliaceum Id 2296
Durinskia baltica Id 1283
Cylindrotheca closterium Id 148
Ulnaria acus Id 1281
Asterionella formosa Id 145
Pseudo-nitzschia multiseries Id 493
Toxarium undulatum Id 1041
Rhodomonas salina Id 2228
Teleaulax amphioxeia Id 453
Chroomonas placoidea Id 1895
Cryptomonas curvata Id 1894
Guillardia theta Id 2188
Cyanidioschyzon merolae strain 10D Id 2303
Coccolithophyceae
Pavlovophyceae
Pelagophyceae
Phaeophyceae
Xanthophyceae
Raphidophyceae
Coscinodioscophyceae
Bacillariophyceae
Mediophyceae
Bolidophyceae
Bacillariophyta
Bacillariophyceae
Dinophyceae
Bacillariophyceae
Mediophyceae
Cryptophyceae
Cyanidiophyceae


Galaxaura rugosa Id 1063
Dichotomaria marginata Id 1062
Scinaia undulata Id 1071
Trichogloeopsis pedicellata Id 1068
Hommersandiophycus borowitzkae Id 1064
Neoizziella asiatica Id 1067
Titanophycus setchellii Id 1066
Liagora brachyclada Id 1072
Liagora harveyana Id 1065
Izziella formosana Id 1069
Helminthora furcellata Id 1292
Dermonema virens Id 1291
Helminthocladia australis Id 1061
Liagoropsis maxima Id 1070
Yamadaella caenomyce Id 1060
Palmaria palmata Id 975
Sheathia arcuata Id 1644
Kumanoa americana Id 970
Thorea hispida Id 985
Calliarthron tuberculosum Id 2454
Sporolithon durum Id 733
Dipterocladia arabiensis Id 1654
Dasya binghamiae Id 1275
Dasya naccarioides Id 1676
Membranoptera tenuis Id 1143
Membranoptera weeksiae Id 1142
Membranoptera platyphylla Id 1118
Thuretia quercifolia Id 1682
Acrosorium ciliolatum Id 1656
Caloglossa intermedia Id 1661
Caloglossa monosticha Id 1659
Caloglossa beccarii Id 1665
Taenioma perpusillum Id 1691
Platysiphonia delicata Id 1655
Digenea simplex Id 1693
Bryothamnion seaforthii Id 1672
Thaumatella adunca Id 1687
Rhodomela confervoides Id 1667
Ophidocladus simpliciusculus Id 1680
Kuetzingia canaliculata Id 1689
Osmundaria fimbriata Id 1658
Lophocladia kuetzingii Id 1688
Cliftonaea pectinata Id 1690
Dasyclonium flaccidum Id 1683
Sonderella linearis Id 1685
Bostrychia simpliciuscula Id 1664
Bostrychia tenella Id 1660
Bostrychia moritziana Id 1662
Laurenciella marilzae Id 1628
Symphyocladia dendroidea Id 1663
Dictyomenia sonderi Id 1629
Gredgaria maugeana Id 1686
Periphykon beckeri Id 1657
Dipterosiphonia australica Id 1684
Herposiphonia versicolor Id 1675
Polysiphonia scopulorum Id 1678
Polysiphonia stricta Id 1671
Polysiphonia schneideri Id 1692
Melanothamnus harveyi Id 1677
Polysiphonia elongata Id 1670
Polysiphonia brodiei Id 1668
Polysiphonia infestans Id 1673
Vertebrata australis Id 1679
Vertebrata isogona Id 1674
Vertebrata lanosa Id 291
Vertebrata thuyoides Id 1669
Tolypiocladia glomerulata Id 1694
Polysiphonia sertularioides Id 1666
Florideophyceae


Asparagopsis taxiformis Id 976
Ahnfeltia plicata Id 973
Chondrus crispus Id 2440
Mastocarpus papillatus Id 981
Grateloupia taiwanensis Id 2478
Sebdenia flabellata Id 964
Schimmelmannia schousboei Id 982
Schizymenia dubyi Id 983
Plocamium cartilagineum Id 991
Gelidium vagum Id 735
Gelidium elegans Id 734
Gracilaria firma Id 1304
Gracilaria salicornia Id 90
Gracilaria tenuistipitata var. liui Id 2382
Gracilaria chilensis Id 752
Gracilaria chorda Id 977
Gracilariopsis lemaneiformis Id 676
Rhodymenia pseudopalmata Id 972
Coeloseira compressa Id 842
Ceramium sungminbooi Id 1928
Ceramium japonicum Id 988
Spyridia filamentosa Id 1681
Pyropia haitanensis Id 2457
Pyropia pulchra Id 736
Pyropia perforata Id 141
Pyropia yezoensis Id 2395
Porphyra purpurea Id 2182
Porphyra umbilicalis Id 1823
Wildemania schizophylla Id 670
Compsopogon caeruleus Id 1721
Boldia erythrosiphon Id 1412
Rhodochaete parvula Id 1771
Erythrotrichia carnea Id 1770
Porphyridium sordidum Id 989
Porphyridium purpureum Id 194
Bangiopsis subsimplex Id 987
Hildenbrandia rivularis Id 990
Hildenbrandia rubra Id 974
Apophlaea sinclairii Id 986
Galdieria sulphuraria Id 165
Cyanidium caldarium Id 2181
Bulboplastis apyrencidosa Id 1394
Florideophyceae
Bangiophyceae
Compsopogonophyceae
Porphyridiophyceae
Stylonematophyceae
Florideophyceae
Cyanidiophyceae
Rhodellophyceae

6. ANEXO I: CERTIFICADO DE REGISTRO DE SOFTWARE SVECT

REPÚBLICA FEDERATIVA DO BRASIL
Ministério Da Indústria, Comércio Exterior e Serviços
Instituto Nacional da Propriedade Industrial
Diretoria de Patentes, Programas de Computador e Topografias de Circuitos Integrados

Certificado de Registro de Programas de Computador

Processo nº: BR 51 2018 000244-7

O Instituto Nacional da Propriedade Industrial expede o presente certificado de Registro de Programas de Computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de Criação: 23 de novembro de 2016 , em conformidade com o parágrafo 2º , artigo 2º da Lei Nº 9.609, de 19 de Fevereiro de 1998.

Título: **SVect - Slide Vector**

Data de Criação: 23 de novembro de 2016

Titular(es): UNIVERSIDADE FEDERAL DO PARANA

Autor(es): AMANDA WILCZEK
/ ARYEL MARLUS REPULA DE OLIVEIRA
/ BRUNO THIAGO DE LIMA NICHIO
/ CAMILLA REGINATO DE PIERRI
/ JERONIZA NUNES MARCHAUKOSKI
/ JOSUÉ OLIVEIRA CAMARGO
/ LETICIA GRAZIELA COSTA SANTOS
/ MARIANE GONCALVES KULIK
/ RICARDO VOYCEIK
/ ROBERTO TADEU RAITTZ

Linguagem: MATLAB

Campo de Aplicação: BL-01, BL-07, IF-10, MT-01

Tipo Programa: IA-01, TC-01, UT-01

Algoritmo Hash: SHA-512

Resumo Digital:
D54AAE745A1B064B83EE09D2431569DE96B996CE3AB32EC50B5CC221E513E99B6AF5127D8F7A43
710DFD14B51648DBE6C29321EDFD4F167DB6040C8A38AD77E6

Expedido em: 06 de março de 2018

Aprovado por Julio Cesar Castelo Branco Reis Moreira