UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

INSTITUTO DE BIOLOGÍA

SISTEMÁTICA

EVALUACIÓN DE LA POSICIÓN FILOGENÉTICA DE FLORES FÓSILES, Y SU RELEVANCIA EN LA INVESTIGACIÓN DE LA EVOLUCIÓN FLORAL

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

PRESENTA:

ANDREA MONSERRATH LÓPEZ MARTÍNEZ

TUTORA PRINCIPAL DE TESIS: DRA. SUSANA AURORA MAGALLÓN PUEBLA
INSTITUTO DE BIOLOGÍA, UNAM
COMITÉ TUTOR: DR. ALFONSO OCTAVIO DELGADO SALINAS
INSTITUTO DE BIOLOGÍA, UNAM
COMITÉ TUTOR: DR. LUIS ENRIQUE EGUIARTE FRUNS
INSTITUTO DE ECOLOGÍA, UNAM

CIUDAD UNIVERSITARIA, CD. MX., JUNIO, 2024.

PROTESTA UNIVERSITARIA DE INTEGRIDAD Y HONESTIDAD ACADÉMICA Y PROFESIONAL

(Graduación con trabajo escrito)

De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo a lo anterior, manifiesto que el trabajo escrito titulado:

Evaluación de la posición filogenética de flores fósiles,
y su relevancia en la investigación de la evolución floral

Que presenté para obtener el grado de DOCTOR(A) EN CIENCIAS, es original, de mí autoría y lo realicé con rigor metodológico exigido por el Programa de Posgrado en Ciencias Biológicas, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referenciadas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de obtención de mi grado académico.

Atentamente

Andrea López

LOPEZ MARTINEZ ANDREA MONSERRATH
No. de cuenta UNAM: 516011701

(Nombre, firma y número de cuenta del estudiante)

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

INSTITUTO DE BIOLOGÍA

SISTEMÁTICA

EVALUACIÓN DE LA POSICIÓN FILOGENÉTICA DE FLORES FÓSILES, Y SU RELEVANCIA EN LA INVESTIGACIÓN DE LA EVOLUCIÓN FLORAL

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

PRESENTA:

ANDREA MONSERRATH LÓPEZ MARTÍNEZ

TUTORA PRINCIPAL DE TESIS: DRA. SUSANA AURORA MAGALLÓN PUEBLA
INSTITUTO DE BIOLOGÍA, UNAM
COMITÉ TUTOR: DR. ALFONSO OCTAVIO DELGADO SALINAS
INSTITUTO DE BIOLOGÍA, UNAM
COMITÉ TUTOR: DR. LUIS ENRIQUE EGUIARTE FRUNS
INSTITUTO DE ECOLOGÍA, UNAM

CIUDAD UNIVERSITARIA, CD. MX., JUNIO, 2024.

COORDINACIÓN GENERAL DE ESTUDIOS DE POSGRADO
COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE BIOLOGÍA
OFICIO: CGEP/CPCB/IB/0432/2024
ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **11 de marzo de 2024** se aprobó el siguiente jurado para el examen de grado de **DOCTORA EN CIENCIAS** de la estudiante **LÓPEZ MARTÍNEZ ANDREA MONSERRATH** con número de cuenta **516011701** con la tesis titulada "**EVALUACIÓN DE LA POSICIÓN FILOGENÉTICA DE FLORES FÓSILES, Y SU RELEVANCIA EN LA INVESTIGACIÓN DE LA EVOLUCIÓN FLORAL",** realizada bajo la dirección de la **DRA. SUSANA AURORA MAGALLÓN PUEBLA,** quedando integrado de la siguiente manera:

Presidente:	DR. ALEJANDRO EMMANUEL GONZÁLEZ VOYER
Vocal:	DRA. CLAUDIA PATRICIA ORNELAS GARCÍA
Vocal:	DRA. TANIA HERNÁNDEZ HERNÁNDEZ
Vocal:	DR. SANTIAGO ALEJANDRO RAMÍREZ BARAHONA
Secretario:	DR. LUIS ENRIQUE EGUIARTE FRUNS

Sin otro particular, me es grato enviarle un cordial saludo.

A T E N T A M E N T E
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 27 de mayo de 2024

COORDINADOR DEL PROGRAMA

POSGRADO EN CIENCIAS BIOLÓGICAS COORDINACIÓN

DR. ARTURO CARLOS II BECERRA BRACHO

c. c. p. Expediente del alumno

ACBB/RGA/EARR/rga

AGRADECIMIENTOS INSTITUCIONALES

Al Posgrado en Ciencias Biológicas, UNAM, por las atenciones brindadas durante mi formación como estudiante de doctorado.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT), por la beca 708068 otorgada para realizar mis estudios de posgrado.

A mi tutora principal: Dra. Susana Aurora Magallón Puebla

A los miembros de mi comité tutoral: Dr. Alfonso Octavio Delgado Salinas y Dr. Luis Enrique Eguiarte Fruns

AGRADECIMIENTOS PERSONALES

A mi tutora principal, la Dra. Susana Magallón, por aceptarme como su estudiante de posgrado y por brindarme su apoyo durante estos años. Agradezco todas las enseñanzas y la motivación para dar lo mejor de mí y alcanzar mis metas profesionales.

Dr. Hervé Sauquet (Royal Botanic Gardens and Domain Trust) quien en conjunto con Dra. Susana Magallón, Dra. Maria von Balthazar y Dr. Jürg schönenberger (University of Vienna) conceptualizaron este proyecto hace varios años. Les agradezco profundamente el haberme permitido formar parte del proyecto eFLOWER. Para mí ha sido un honor y enorme satisfacción colaborar con ustedes. Son una gran fuente de inspiración y les quiero agradecer todo el apoyo, la motivación y sus enseñanzas para desarrollar los diferentes proyectos de la tesis.

Dra. Marion Chartier, University of Vienna, por aceptar colaborar en el proyecto del morfoespacio floral, por todas sus enseñanzas en los métodos de ordenación del mismo y para cuantificar disparidad morfológica, y por su cariñoso recibimiento en Viena.

Dr. Luis Eguiarte Fruns y Dr. Alfonso Delgado Salinas, por hacer comentarios constructivos a lo largo del desarrollo de los estudios que conforman esta Tesis. Les agradezco enormemente que hayan leído y realizado sugerencias en borradores previos de los artículos publicados.

Al instituto de Biología, UNAM por proporcionarme todos los recursos necesarios para mi formación académica.

A la American Society of Plant Taxonomist (ASPT) y a la International Association of Plant Taxonomy (IAPT), por otorgarme las becas de investigación (research grants) en el 2019, que me permitieron solventar los gastos de mi estancia de investigación en Viena.

Dr. Alejandro González Voyer, Dr. Santiago Ramírez Barahona, Dra. Patricia Ornelas García, Dra. Tania Hernández Hernández, por su tiempo, críticas y sugerencias que mejoraron el contenido de la tesis.

Dra. Laura Calvillo Canadell, técnico académico del Herbario Nacional (MEXU), por leer y proporcionarme comentarios valiosos al manuscrito sobre la posición filogenética de flores fósiles.

Dr. Cristian Cervantes Salcedo, técnico académico de la Unidad de Síntesis en Sistemática y Evolución (UNISSE), Instituto de Biología, UNAM, por brindarme el apoyo para correr los análisis del fechamiento en el servidor BEAGLE.

Rocío González Acosta de la Oficina del Posgrado en el Instituto de Biología, por siempre ser amable conmigo y ayudarme a esclarecer dudas durante distintas etapas de mi transcurso como estudiante en la UNAM.

A los ex y actuales integrantes del laboratorio B-123:, Adriana Benítez, Rebeca Hernández, Santiago Ramírez, Luna Sánchez, César González y Karina Yañez, por todas las discusiones sobre macroevolución, las buenas charlas y su amistad y compañía en el instituto.

A los integrantes del Sauquet Lab: Yasmin Asar, Ruby Stephens, Thomas Mesaglio, Jakub Baczynski, Julian Herting, por las grandiosas discusiones de artículos durante la pandemia, que sin duda me inspiraron y me ayudaron a entender mejor distintos aspectos metodológicos y de la evolución de angiospermas.

A Iván, mi familia, y mis amigos, por su amor, por escucharme, brindarme su apoyo incondicional, y también por motivarme a ser una mejor persona.

DEDICATORIA

A mi esposo, Iván Absalón,
que me ha dado su amor, apoyo y la mejor compañía.

Índice

Resumen	**1**
Abstract	**2**
Introducción General	**3**
Objetivos	**8**
Capítulo I. Integrando flores fósiles en la filogenia de angiospermas utilizando evidencia molecular y morfológica	
Resumen	**9**
Manuscrito publicado en *Systematic Biology*	**10**
Material Suplementario	**64**
Capítulo II. Las flores de angiospermas alcanzaron su mayor diversidad morfológica temprano en su historia evolutiva	
Resumen	**65**
Artículo publicado en *New Phytologist*	**66**
Material Suplementario	**79**
Capítulo III. Explorando el fechamiento de evidencia total bajo el modelo de nacimiento y muerte fosilizado aplicado a angiospermas	
Resumen	**80**
Manuscrito	**81**
Material Suplementario	**107**
Discusión General y Conclusiones	**139**
Referencias Bibliográficas	**147**

RESUMEN

Las angiospermas – o plantas con flor – han evolucionado hacia una gran diversidad de estructuras reproductivas desde su origen, hace más de 140 millones de años. Existe un avance significativo en diversos campos del conocimiento que nos han permitido mejorar nuestro entendimiento sobre la evolución de las flores y los procesos intrínsecos y extrínsecos que han moldeado su diversidad. La obtención de una filogenia dónde las relaciones entre los linajes principales de las angiospermas son estables en general, aunado a información derivada de múltiples estudios enfocados en la morfología comparada de las flores de especies vivientes, nos han ayudado a entender la homología de estructuras, o qué caracteres muestran mayor o menor flexibilidad. Asimismo, el estudio y la descripción detallada de numerosas flores fósiles también nos ha proporcionado una fuente de información directa sobre los patrones de diversificación floral a través del tiempo. En esta tesis integro información molecular y morfológica de 1201 especies vivientes, representantes de todas las familias de angiospermas, e información morfológica de 121 flores fósiles, para así contribuir al conocimiento sobre la evolución de las angiospermas. Primero presento un estudio evaluando la posición filogenética de flores fósiles usando distintos experimentos analíticos, y destacando la importancia de considerar la incertidumbre filogenética asociada. Proponemos estrategias para la integración de fósiles en marcos de referencia filogenéticos y perspectivas en el estudio de tiempos de divergencia y evolución de rasgos morfológicos. En segundo lugar, utilicé un enfoque del morfoespacio para determinar la ocupación de especies vivientes, fósiles y ancestros, para buscar combinaciones morfológicas exitosas y posibles límites en la evolución floral. Paralelamente, cuantifiqué la variación de la diversidad (disparidad) floral entre los principales linajes de angiospermas y entre periodos de tiempo geológico, encontrando que las angiospermas alcanzaron sus niveles más altos de disparidad morfológica temprano en su historia evolutiva. Por último, exploré – mediante la implementación del método de fechamiento de puntas (*tip dating*) bajo el modelo de nacimiento y muerte fosilizado (FBD, por sus siglas en inglés *fossilized birth-death model*) – la estimación del tiempo de origen de las familias de angiospermas y las tasas de cambio morfológico floral. Muestro que los tiempos de origen de las familias son considerablemente más jóvenes que las edades obtenidas con enfoques convencionales, y encontré una mayor heterogeneidad de cambio morfológico al inicio de la diversificación de los linajes principales de angiospermas.

ABSTRACT

Angiosperms or flowering plants have evolved into an immense diversity of reproductive structures since their origin more than 140 million years ago. There is significant progress in various fields of knowledge that has allowed us to improve our understanding of the evolution of flowers and the intrinsic and extrinsic processes that have shaped their diversity. Obtaining a phylogeny where the relationships between the major lineages of angiosperms are stable in general, coupled with the information derived from multiple studies focused on the comparative morphology of the flowers of living species, have helped us clarify the homology of structures and the characters that might show greater or lesser flexibility. In addition, the study and detailed description of numerous fossil flowers have also provided us with a direct source of information on patterns of floral diversification in deep time. In this thesis, I integrate molecular and morphological information from 1201 living species sampled to represent all angiosperm families, and morphological information from 121 fossil flowers to contribute to knowledge about some aspects of the evolution of angiosperms. First, I present a study focused on the explicit evaluation of the position of fossil flowers through different analytical experiments, and the importance of considering the associated phylogenetic uncertainty is highlighted. Based on this study, we also propose useful strategies for the integration of fossils into phylogenetic frameworks and perspectives in the study of divergence times and evolution of morphological traits. Second, I describe a morphospace approach that was used to determine the occupancy of ancestors and living and fossil species to determine successful morphological combinations and possible limits on floral evolution. In addition, I quantified variation in floral diversity (disparity) among major angiosperm lineages and between geological time periods, and found that angiosperms reached the highest levels of morphological disparity early in their evolutionary history. Lastly, an exploratory study was conducted – implementing the tip dating approach under the fossilized birth-death model (FBD) – aimed at estimating the time of origin of angiosperm families and the rates of morphological (floral) change. Based on this study, I show that the ages of origin of the families are considerably younger than those obtained with conventional dating approaches and that the greater heterogeneity of morphological change was found at the beginning of the diversification of the major angiosperm lineages.

Introducción General

Las angiospermas son el grupo de plantas dominante en los ecosistemas terrestres actuales (Crepet y Niklas, 2009). Conocer su origen y evolución es de gran importancia, por ejemplo, para entender los patrones de diversificación de otros linajes (como son los insectos, aves, mamíferos) (Herendeen et al., 2017). Las angiospermas se caracterizan por el desarrollo de múltiples novedades que les han permitido diversificarse tanto en términos de número de especies como en la diversidad de ambientes que colonizan. Una de sus mayores innovaciones es el marcado incremento en su complejidad reproductiva, mediante la construcción de flores con óvulos encerrados (protegidos) en un carpelo, de estructuras dedicadas a la producción de polen (los estambres), y de estructuras especializadas para la polinización y dispersión (Leslie et al., 2021).

La extensa evidencia paleobotánica acumulada durante las últimas décadas ha documentado el origen y radiación de diversos linajes de angiospermas desde hace ~140 millones de años (Hickey y Doyle, 1977; Lidgard y Crane, 1990). Asimismo, múltiples estudios filogenéticos han mejorado nuestro conocimiento sobre las relaciones entre linajes a niveles profundos del árbol de las angiospermas (APG IV, 2016), lo que es indispensable para la estimación de patrones macroevolutivos. A pesar de que existe un progreso significativo en diversas áreas del conocimiento, aún existe mucha incertidumbre sobre cuando se originaron las angiospermas y cuándo empezaron a diversificarse (ver Sauquet et al., 2022); cómo es la evolución de atributos florales a través del tiempo y entre linajes, y la naturaleza del cambio morfológico (ver Sauquet y Magallón, 2018). La motivación principal de esta tesis es aportar al conocimiento de algunas de estas incógnitas desde una perspectiva macroevolutiva, que involucra la aplicación de distintos enfoques metodológicos y el análisis de información conjunta de especies actuales y fósiles.

El descubrimiento y descripción de flores fósiles con atributos morfológicos excelentemente bien preservados nos han proporcionado evidencia directa sobre la diversidad y evolución de angiospermas a través del tiempo (Friis et al., 2011, Doyle y Endress, 2010; Schönenberger, 2005). Aunque los fósiles de flores son preservados por diversos procesos, como permineralización, en ámbar, por impresión/compresión, las flores carbonizadas han sido las más estudiadas. Los especímenes de flores carbonizadas se caracterizan por su excelente grado de preservación en tercera dimensión, dónde en muchos casos se conservan con detalle sus atributos morfológicos de manera interna y externa. Además, las nuevas técnicas de visualización de estos fósiles – como lo es la tomografía computarizada (Friis et al., 2014) – han mejorado la calidad y cantidad de datos morfológicos que podemos obtener, por ejemplo, de caracteres internos (e.g., placentación) a una resolución anatómica. Sin embargo, la asignación taxonómica de flores fósiles se ha restringido a la observación detallada de atributos morfológicos y su comparación con especies vivientes (Schönenberger et al., 2020; López-Martínez et al., 2023).

Relativamente pocos estudios han investigado la posición de flores fósiles en un contexto filogenético, y en su mayoría se limitan a la utilización de métodos de parsimonia y muestreos taxonómicos altamente constreñidos (e.g., un género, familia, orden). El objetivo del *Capítulo I* de esta Tesis es evaluar la posición filogenética de 24 flores fósiles en un árbol que incluye representantes de todas las familias de angiospermas. Para llevar a cabo este estudio, expandimos y curamos significativamente la matriz de datos morfológicos de flores más grande disponible hasta el momento (Sauquet et al., 2017; Schönenberger et al., 2020). Esta matriz fue construida a partir de datos estructurales de la flor (*Bauplan*; *sensu* Endress, 1994), incluyendo caracteres que son aplicables para todas las angiospermas. Adicionalmente, utilizamos diversas estrategias metodológicas para examinar el comportamiento de los datos y modelos en análisis que integran fósiles y especies vivientes a gran escala filogenética. Determinamos la incertidumbre filogenética asociada a la posición de los fósiles en análisis que varían con relación a la aplicación de diversas restricciones topológicas, criterios de optimización (parsimonia, máxima verosimilitud e inferencia bayesiana), modelos de estimación de longitudes de rama, y la cantidad de fósiles y tipo de datos analizados.

Los fósiles proveen información única sobre las tasas de origen (*origination rates*) y extinción de linajes (Raup y Sepkoski, 1982; Foote, 2003; Silvestro et al., 2015), sobre la naturaleza y las posibles causas de extinción (Marshall, 2017) y los cambios morfológicos a través del tiempo (Gould, 1989; Foote, 1994; Lupia, 1999). Particularmente, la cuantificación de la disparidad morfológica (diversidad morfológica) de fósiles, y su comparación entre clados, poblaciones o periodos de tiempo, nos ayudan a comprender patrones de evolución morfológica de los linajes (Foote, 1994).

Con base en estudios de disparidad morfológica en diversos clados de animales, principalmente invertebrados (Foote, 1992; Foote, 1999; Hughes et al., 2013), se han propuesto dos patrones generales de evolución morfológica a través del tiempo. El patrón de disparidad temprana (*early-disparity* o *bottom-heavy*), sugiere que la mayor variación morfológica es alcanzada al inicio de la diversificación de un clado, como resultado de una rápida colonización o explotación de nuevos ambientes (Gould, 1989; Hughes et al., 2013). En contraste, el modelo de disparidad tardía en la evolución (*late-disparity* o *top-heavy*; Prentice et al., 2011; Hughes et al., 2013) indica que se alcanzan los niveles altos de variación morfológica tarde en la historia evolutiva de un clado. Los patrones de disparidad tardía son esperados cuando la historia de un linaje se ve trucada por una extinción masiva (Hughes et al., 2013), o cuando ciertas regiones del morfoespacio quedan disponibles para la acumulación de nuevas morfológias después de procesos de extinción aleatorios (Oyston et al., 2015).

Las flores son estructuras altamente diversas y complejas, las cuales se encuentran en estrecha relación con procesos de especiación y diversificación (Endress, 2011; van der Niet y Johnson, 2012). La diversidad estructural de las flores se ha discutido ampliamente desde una perspectiva cualitativa, con base en estudios del desarrollo y morfología comparada de varios linajes vivientes (ver Endress, 2010; 2011). El *Capítulo II* de esta Tesis está enfocado en cuantificar esta variación morfológica floral (disparidad) entre linajes principales de angiospermas y a través del tiempo. Utilizamos diversas métricas de disparidad morfológica basadas en análisis multivariados y de ordenación del morfoespacio en el análisis una matriz morfológica que incluye caracteres estructurales de la flor (*Bauplan;* Endress, 1994) para examinar la ocupación de 121 fósiles y 1201 especies vivientes en el morfoespacio teórico. Con base en este estudio, determinamos las combinaciones florales más exitosas, así como aquellas que no han sido viables durante la evolución, sugiriendo posibles causas o limitaciones en la estructura floral. Adicionalmente, inferimos combinaciones morfológicas de ancestros para los clados principales linajes de angiospermas, y los posicionamos en el morfoespacio teórico para determinar las posibles trayectorias en la evolución floral. Por último, realizamos pruebas de correlación entre disparidad, tiempos de origen, y riqueza de especies.

En marcos de referencia filogenéticos, el registro fósil es usado con frecuencia para calibrar árboles y transformar longitudes de rama en tiempo geológico absoluto (Drummond et al., 2006; Marshall, 2008; Ho y Phillips, 2009). Algunos enfoques de fechamiento, como la calibración de nodos (*node dating*), requieren especificaciones *a priori* sobre la ubicación de los fósiles y el tipo de distribución estadística. Los avances metodológicos más recientes han permitido un uso más integral del registro fósil. Por ejemplo, tenemos el desarrollo de modelos como el de nacimiento y muerte fosilizado *(FBD,* por sus siglas en inglés *fossilized birth-death model*; Heath et al., 2014), donde los fósiles y especies vivientes forman parte de un único proceso de diversificación. La implementación del modelo de FBD ha sido útil para inferir tiempos de divergencia de linajes y otros procesos macroevolutivos, como tasas de cambio evolutivo, especiación, extinción (Heath et al., 2014; Zhang et al., 2016; Gavryushkina et al., 2017).

Adicionalmente, se han desarrollado diversas extensiones al modelo original de FBD, como la utilización de matrices de evidencia total para la co-estimación de topología y tiempos de divergencia en enfoques de fechamiento de puntas (*total evidence dating*; Zhang et al., 2016; Gavryushkina et al., 2017), la implementación de ajustes por sesgos en el muestreo taxonómico (*diversified sampling*; Zhang et al., 2016), y para relajar los supuestos de constancia en las tasas a lo largo del tiempo (*skyline model*; Zhang et al., 2023). Los árboles generados con el modelo FBD han sido un punto de inicio para la estimación de la dinámica de diversificación (Lloyd y Slater, 2021), procesos biogeográficos (Zhang et al., 2021; Coiro et al., 2023), y evolución de tasas fenotípicas (Simões et al., 2021).

En el *Capítulo III* de esta tesis presento un estudio exploratorio donde implementé el modelo de FBD con un enfoque de fechamiento de puntas a gran escala filogenética. Se utilizó este estudio como base para contestar preguntas relacionadas con los tiempos de origen y tasas de evolución morfológica en angiospermas. Los resultados principales de este estudio indican que las edades de diversificación (edad corona) de las familias de angiospermas son considerablemente más jóvenes que las obtenidas en estudios previos basados en el método de calibración de nodos. Adicionalmente, se detectaron altas tasas de cambio morfológico en el origen de los clados principales de angiospermas, que es congruente con la mayor expansión del morfoespacio floral.

OBJETIVOS

1. Evaluar la posición (e incertidumbre) asociada de 24 flores fósiles bien preservadas, distribuidas a lo largo del árbol filogenético de las angiospermas, mediante la implementación de múltiples experimentos analíticos que varían tanto con relación al criterio de optimización, como en la aplicación de diversas restricciones topológicas, y en el tipo de datos utilizados.
2. Examinar la variación en la diversidad morfológica (disparidad) de flores actuales y extintas, así como su ocupación en el morfoespacio teórico, para comprender patrones de evolución morfológica entre los linajes principales de angiospermas y a través del tiempo.
3. Estimar tiempos de divergencia y tasas de cambio morfológico entre linajes y a través del tiempo, mediante la implementación del enfoque de fechamiento de puntas bajo el modelo de nacimiento y muerte fosilizado.

CAPÍTULO I

Integrating fossil flowers into the angiosperm phylogeny using molecular and morphological evidence

Artículo publicado: **López-Martínez AM**, Schönenberger J, von Balthazar M, González-Martínez CA, Ramírez-Barahona S, Sauquet H, Magallón S. 2023. Integrating fossil flowers into the angiosperm phylogeny using molecular and morphological evidence. Systematic Biology 72: 837–855.

RESUMEN – Los fósiles son esenciales para inferir procesos evolutivos pasados. La asignación de fósiles a clados se ha basado tradicionalmente en la similitud morfológica y en apomorfías compartidas con los taxones vivientes. Hasta ahora, el uso de análisis filogenéticos explícitos para establecer afinidades fósiles ha sido limitado. En este estudio, construimos un marco integral para investigar la ubicación filogenética de 24 flores fósiles excepcionalmente preservadas. Para ello ensamblamos una matriz de datos a nivel de especie de 30 rasgos florales para 1201 especies de plantas vivientes que fueron muestreadas para capturar los nodos troncales y corona de todas las familias de angiospermas. Exploramos múltiples enfoques analíticos para integrar los fósiles en la filogenia, incluyendo diferentes métodos de estimación filogenética, análisis topológicos restringidos y combinación de datos moleculares y morfológicos de especies fósiles y vivientes. Nuestros resultados fueron ampliamente consistentes entre los enfoques y mostraron diferencias menores en el soporte de fósiles en diferentes posiciones filogenéticas. La ubicación de algunos fósiles es congruente con las relaciones sugeridas previamente, pero para otros se infiere una nueva ubicación. También identificamos fósiles que están bien sustentados dentro de ciertas familias de especies vivientes, mientras que otros mostraron una alta incertidumbre filogenética. Finalmente, presentamos recomendaciones para futuros análisis que combinen evidencia molecular y morfológica, con respecto a la selección de fósiles y metodologías apropiadas, y brindamos algunas perspectivas sobre cómo integrar los fósiles en la investigación de los tiempos de divergencia y la evolución temporal de los rasgos morfológicos.

Palabras clave: angiospermas; flores fósiles; incertidumbre filogenética; RoguePlots.

Integrating Fossil Flowers into the Angiosperm Phylogeny using Molecular and Morphological Evidence

AUTHORS

Andrea M. López-Martínez [1,2], Jürg Schönenberger[3], Maria von Balthazar[3], César A. González-Martínez[2], Santiago Ramírez-Barahona[2], Hervé Sauquet[4,5,*], Susana Magallón[2,*]

[1] Posgrado en Ciencias Biológicas, Instituto de Biología, Universidad Nacional Autónoma de México, 3er Circuito de Ciudad Universitaria, Del. Coyoacán, Ciudad de México 04510, Mexico

[2] Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, 3er Circuito de Ciudad Universitaria, Del. Coyoacán, Ciudad de México 04510, Mexico

[3] Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, Vienna A-1030, Austria

[4] National Herbarium of New South Wales (NSW), Royal Botanic Gardens and Domain Trust, Sydney, NSW 2000, Australia

[5] Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia

[*] *These authors contributed equally to this work*

Corresponding author: A. M. López-Martínez, andreaea.lopez@gmail.com

Running head: PHYLOGENETIC PLACEMENT OF FOSSIL FLOWERS

ABSTRACT

Fossils are essential to infer past evolutionary processes. The assignment of fossils to extant clades has traditionally relied on morphological similarity and on apomorphies shared with extant taxa. The use of explicit phylogenetic analyses to establish fossil affinities has so far remained limited. In this study, we built a comprehensive framework to investigate the phylogenetic placement of 24 exceptionally preserved fossil flowers. For this, we assembled a new species-level dataset of 30 floral traits for 1201 extant species that were sampled to capture the stem and crown nodes of all angiosperm families. We explored multiple analytical approaches to integrate the fossils into the phylogeny, including different phylogenetic estimation methods, topological-constrained analyses, and combining molecular and morphological data of extant and fossil species. Our results were widely consistent across approaches, with minor differences in the support of fossils at different phylogenetic positions. The placement of some fossils agrees with previously suggested relationships, but for others, a new placement is indicated. We also identified fossils that are well supported within particular extant families, whereas others showed high phylogenetic uncertainty. Finally, we present recommendations for future analyses combining molecular and morphological evidence, regarding the selection of fossils and appropriate methodologies, and provide some perspectives on how to integrate fossils into the investigation of divergence times and the temporal evolution of morphological traits.

Key words

Angiosperms; Fossil Flowers; Phylogenetic Uncertainty; RoguePlots.

The information provided by the fossil record is essential to infer diverse evolutionary processes in deep time (Marshall 2017; Benson et al. 2021). Fossils are needed to estimate the age of the origin of major living groups (e.g., O'Leary et al. 2013; Misof et al. 2014; Magallón et al. 2015), to understand diversification dynamics (Silvestro et al. 2015; Mitchell et al. 2019; Lloyd and Slater 2021), to reconstruct ancestral states (Slater et al. 2012; Betancur-R et al. 2015; Larson-Johnson 2016), or to infer the biogeographical history of lineages (Meseguer et al. 2015; Zhang et al. 2021). The recent development of novel methodological approaches has made possible a more integrative use of fossil information, such as by the incorporation of fossils into the same diversification process as extant species through the fossilized birth-death model (Heath et al. 2014); or the use of the fossil record to estimate changes in origination and extinction rates (Silvestro et al. 2014; Mitchell et al. 2019). However, some of these approaches share a requirement of *a priori* specifications about the taxonomic assignment of the fossils, which in turn are based on morphological information that, in the best-case scenario, are also supported by phylogenetic analyses. The tip dating approach, in which fossils appear as extinct tips within a phylogenetic tree, uses fossil ages and matrices combining molecular and morphological data to simultaneously infer the placement of fossils and calibrate the tree to estimate divergence times (Ronquist et al. 2012; Zhang et al. 2016; Gavryushkina et al. 2017). Tip dating is a promising approach, but it requires the assemblage of matrices that integrate molecular data of extant species with morphological data scored for both extant and fossil taxa. The assembly of such matrices is particularly challenging in broad-scale analyses because of the difficulty of generating comprehensive data for a large sample of species. The phylogenetic investigation of fossils using molecular and morphological data has so far remained limited in part due to the lack of morphological matrices for a broad sample of extant and fossil taxa (Sauquet and Magallón 2018), possible methodological artifacts linked with large amounts of missing data for fossils (Manos et al. 2007), and the use of simplified models of morphological character evolution that result in distorted topologies and branch lengths (Ronquist et al. 2016).

Fossil flowers are an invaluable source of information to comprehend floral morphological diversification and the temporal emergence of novel structures across angiosperms. These fossilized flowers also provide insights into the diverse evolutionary mechanisms in angiosperms and their interactions with pollinators and dispersal vectors (Friis et al. 2010, 2011). Although flowers are preserved through different processes such as permineralization, impressions/compressions, and amber, it is the charcoalified flowers that have been most extensively studied (Friis et al. 2006, 2011). Charcoalified flowers are most commonly found in Cretaceous sediments, possibly due to increases in paleo-fire regimes during that period (Bond and Scott 2010). These fossils have been described mainly from localities in the Eastern United States (Drinnan et al. 1990, 1991; Crepet and Nixon 1998; Gandolfo et al. 1998; von Balthazar et al. 2007; Crepet et al. 2018; Friis et al. 2020), Portugal (Friis et al. 2001, 2019), Sweden (Friis and Skarby 1981; Schönenberger and Friis 2001; Friis and Pedersen 2012), the Czech Republic (Heřmanová et al. 2021, 2022), and Japan (Takahashi et al. 2001, 2008). The preservation of charcoalified flowers is often exquisite, allowing for the detailed study of morphological and anatomical floral traits. Thanks to their rigidity and resistance against compression, many of these charcoalified specimens are preserved in 3D-shape, frequently preserving internal and external structures in great detail (Schönenberger 2005). Diverse visualization techniques, such as synchrotron radiation x-ray, have improved the quality and quantity of morphological information that can be obtained from the fossil record (Friis et al. 2014). Therefore, morphological details of charcoalified flowers may be obtained at anatomical or cellular resolution, equivalent to those obtained for living species.

The systematic assignment of fossil flowers has so far mainly relied on extensive morphological investigations and on the identification of apomorphies shared with living taxa (e.g., Drinnan et al. 1990; Magallón et al. 2001; von Balthazar et al. 2005). Studies that identify or corroborate systematic hypotheses of fossil flowers using phylogenetic methods are scarce (e.g., Keller et al. 1996; Gandolfo et al. 2002; Magallón 2007; Doyle and Endress 2010, 2014; Lee et al. 2013; Martínez et al. 2016; Schönenberger et al. 2020). Most of these analyses investigate the position of individual fossil flowers by focusing on a specific subclade of angiosperms (i.e., an order, a family, or a genus), which has been defined *a priori* based on comparative morphology and taxonomic expertise (e.g., Hermsen et al. 2003; Mendes et al. 2014; Martínez et al. 2016). Consequently, these analyses exclude the possibility that the fossil under study falls outside of the focal group (widely discussed by Schönenberger et al. 2020). Relatively few studies have applied a more broad-scale approach to cover multiple major angiosperm subclades, including Doyle and Endress (2010, 2014), and the broadest one, published recently by Schönenberger et al. (2020). In these studies, the fossils are integrated into a molecular backbone phylogeny of extant species based on parsimony, without exploring alternative methods or optimization criteria. In addition, the phylogenetic placement of fossils was assessed one fossil at the time, thereby removing the potential benefit of the interaction between multiple fossils and extant species.

Here, we integrate molecular and morphological data into a comprehensive framework to jointly analyze the phylogenetic placement of 24 well-preserved fossil flowers across the entire angiosperm phylogeny. For this, we assembled the largest dataset of floral traits ever published, and considerably expanding on previous datasets (Sauquet et al. 2017; Schönenberger et al. 2020). This dataset includes extant representatives of all currently recognized angiosperm families *sensu* the Angiosperm Phylogeny Group IV (APG IV, 2016) and mirrors the sampling of species in the molecular dataset of Ramírez-Barahona et al. (2020). We use multiple phylogenetic approaches to assess the placement of fossils and subsequently visualize their phylogenetic uncertainty. We implement previous approaches based on the application of topological constraints among extant species and the analysis of each fossil independently, as well as new unconstrained analyses based on joint molecular and morphological data for the phylogenetic estimation of multiple fossil and extant species simultaneously. Finally, based on our results we provide recommendations for the investigation on the phylogenetic placement of fossils based on analyses of combined molecular and morphological data.

MATERIALS AND METHODS

Extant and fossil taxon sampling

We selected a total of 24 exceptionally well-preserved fossil flowers with the aim of having a broad representation of fossils across all major lineages of angiosperms (according to the systematic positions established in earlier studies; Table 1 and Supplementary Information Data 1). Most of the fossils selected for this study are three-dimensionally preserved charcoalified flowers (e.g., Crane et al. 1989; Magallón-Puebla et al. 1997; von Balthazar et al. 2008). However, we also included permineralized fossils (e.g., Smith and Stockey 2007) and impressions/compressions (e.g., Manchester 1992). The set of fossils included the ten fossil flowers that were analyzed in the study by Schönenberger et al. (2020) with the aim of comparing their phylogenetic placements using different taxon sampling and methods. We used the molecular dataset of extant species from Ramírez-Barahona et al. (2020), which consisted of seven molecular markers: four protein-coding plastid genes (*rbcL*, *atpB*, *matK*, *ndhF*); and three nuclear loci (*18S*, *26S*, and *5.8S* nrDNA). The original study by Ramírez-Barahona et al. (2020) encompassed 1209 taxa, including a combination of explicit species and chimeric taxa (whereby different genes were sampled from multiple species within the same genus to maximize data coverage). Here we used the same approach as Sauquet et al. (2017) to match each of these taxa in the molecular dataset to an explicit species in our morphological dataset. This strategy has the potential weakness of not selecting those extant species with more morphological data available, resulting in a larger amount of missing data. However, we consider that by directly matching sequences and morphological traits in the same species, we avoid assumptions of supraspecific taxon monophyly and of trait distribution within such taxa (see Prendini, 2001). In this process, eight taxa could not be matched or became redundant, leading to a molecular dataset and tree of 1201 extant species representing the stem and crown nodes of all angiosperm families *sensu* APG IV (2016) and the Angiosperm Phylogeny Website (http://www.mobot.org/MOBOT/research/APweb/; Stevens, 2001).

Table 1. Main information about age, locality, previously suggested affinities of the 24 fossils included in this study. CG, Crown group; SG, Stem group; Comp, Compression; Charc, Charcoalified; Perm, Permineralized. Numbers in parentheses correspond to the number of characters scored.

Fossil	Age	Type locality	Previously suggested affinities	Reference	Preservational mode/Characters scored	Previous phylogenetic studies
Archaeanthus linnenbergeri	Albian to middle Cenomanian	Linnenberger's Ranch central Kansas, USA	Several families within Magnoliidae	Dilcher and Crane, 1984	Comp. (16)	CG Magnoliales; Doyle and Endress 2010
Archaestella verticillatus	Early Coniacian	Kamikitaba, Japan	Trochodendraceae (Trochodendrales)	Takahashi et al. 2017	Charc. (19)	Not tested
Bertilanthus scanicus	Early Campanian	Åsen Scania, Sweden	Paracryphiaceae (Paracryphiales)	Friis and Pedersen, 2012	Charc. (29)	Not tested
Canrightia resinifera	Late Barremian to Aptian	Catefica, Portugal	Chloranthaceae (Chloranthales), Piperales	Friis and Pedersen, 2011	Charc. (18)	CG Chloranthaceae, sister to Chloranthaceae; Friis and Pedersen 2011; Doyle and Endress 2014, 2018
Carpestella lacunata	Early to middle Albian	Puddledock, Virginia, USA	Nymphaeaceae (Nymphaeales), *Illicium* (Austrobaileyales)	von Balthazar et al. 2008	Charc. (12)	Laurales and Magnoliales; von Balthazar et al. 2008; CG Nymphaeales; Doyle and Endress 2014
Cecilanthus polymerus	Early Cenomanian	Elk Neck Peninsula, Maryland, USA	ANA grade, Magnoliidae, and Chloranthaceae	Herendeen et al. 2016	Charc. (24)	Nymphaeales, Magnoliales, sister to Laurales and Chloranthaceae; Herendeen et al. 2016
Chloranthistemon endressii	Early Campanian	Åsen Scania, Sweden	Chloranthaceae (Chloranthales)	Crane et al. 1989	Charc. (12)	CG Chloranthaceae; Eklund et al. 2004; Schönenberger et al. 2020
Dakotanthus cordiformis	Late Albian to Cenomanian	Dakota Formation	Quillajaceae (Fabales)	Manchester et al. 2018	Comp. (27)	CG Rosidae; Schönenberger et al. 2020

Divisestylus brevistamineus	Turonian	Old Crossman Clay Pit. New Jersey, USA	Iteaceae (Saxifragales)	Hermsen et al. 2003	Charc. (28)	CG Iteaceae; Hermsen et al. 2003
Florissantia quilchenensis	Eocene	Quilchena, British Columbia. Canada	Malvales	Manchester, 1992	Comp. (22)	Not tested
Kajanthus lusitanicus	Late Aptian to Early Albian	Chicalhão. Portugal	Lardizabalaceae (Ranunculales)	Mendes et al. 2014	Charc. (27)	Lardizabalaceae; Mendes et al. 2014; CG Ranunculales, several positions in Lardizabalaceae and Berberidaceae; Schönenberger et al. 2020
Mabelia connatifila	Turonian	Old Crossman Clay Pit. New Jersey, USA	Triuridaceae (Pandanales)	Gandolfo et al. 2002	Charc. (18)	Triuridaceae; Gandolfo et al. 2002
Mauldinia mirabilis	Cenomanian	Elk Neck Peninsula, Maryland, USA	Lauraceae (Laurales)	Drinnan et al. 1990	Charc. (21)	Sister to Lauraceae and Hernandiaceae; Doyle and Endress 2010; Schönenberger et al. 2020
Microvictoria svitkoana	Turonian	Old Crossman Clay Pit. New Jersey, USA	Nymphaeaceae (Nymphaeales)	Gandolfo et al., 2004	Charc. (20)	Nymphaceae; Gandolfo et al. 2004; Several positions in ANA grade and Calycanthaceae; Schönenberger et al. 2020
Paleoclusia chevalieri	Turonian	Old Crossman Clay Pit. New Jersey, USA	Clusiaceae (Malpighiales)	Crepet and Nixon 1998	Charc. (27)	Clusiaceae; Crepet and Nixon 1998; Ruhfel et al. 2013; CG in Malpighiales and Fabidae; Schönenberger et al. 2020
Paradinandra suecica	Early Campanian	Åsen, Scania, Sweden	Ericaceae (Ericales)	Schönenberger and Friis, 2001	Charc. (24)	CG Ericales and several positions in Pentapetalae; Schönenberger et al. 2020

Pentapetalum trifasciculandricus	Turonian	Old Crossman Clay Pit. New Jersey, USA	Theaceae s.l. (Ericales)	Martinez-Millan et al. 2009	Charc. (24)	Theaceae; Martinez-Millan et al. 2009
Platydiscus peltatus	Early Campanian	Åsen, Scania, Sweden	Cunoniaceae (Oxalidales)	Schönenberger et al. 2001	Charc. (29)	Not tested
Quadriplatanus georgianus	Coniacian to Santonian	Upatoi Creek, Georgia, USA.	Platanaceae (Proteales)	Magallón-Puebla et al. 1997	Charc. (23)	Not tested
Rariglanda jerseyensis	Turonian	Old Crossman Clay Pit. New Jersey, USA	Clethraceae (Ericales)	Martínez et al. 2016	Charc. (25)	Clethraceae; Martínez et al. 2016
Saururus tuckerae	Middle Eocene	Princeton Chert, British Columbia, Canada.	Saururaceae (Piperales)	Smith and Stockey 2007	Perm. (13)	Saururaceae; Smith and Stockey 2007
Spanomera mauldinensis	Early Cenomanian	Mauldin Mountain, Maryland, USA	Buxaceae (Buxales)	Drinnan et al. 1991	Charc. (24)	Sister to Buxaceae; Doyle and Endress 2010; CG and SG Buxales, CG Gunnerales Schönenberger et al. 2020
Tylerianthus crossmanensis	Turonian	Old Crossman Clay Pit. New Jersey, USA	Saxifragaceae (Saxifragales), Hydrangeaceae (Cornales)	Gandolfo et al. 1998	Charc. (27)	CG Saxifragaceae; Schönenberger et al. 2020
Virginianthus calycanthoides	Early to middle Albian	Puddledock locality, Virginia USA.	Calycanthaceae (Laurales)	Friis et al. 1994	Charc. (23)	SG Calycanthaceae and Laurales; Doyle and Endress 2010; CG and SG Calycanthaceae; Schönenberger et al. 2020

Morphological data

The morphological dataset consisted of 30 floral traits recorded for 1201 extant and 24 fossil species in the PROTEUS database (Sauquet 2019). This is the same set of floral characters defined and used by Schönenberger et al. (2020), including structural traits describing the sex of flowers (1 character), the perianth (8 characters), the androecium (11 characters), the gynoecium (8 characters), and pollen (2 characters). For details on our scoring philosophy and definition of characters see published appendices in Sauquet et al. (2017) and Schönenberger et al. (2020). The morphological traits were scored based on an extensive search of data in a wide range of published literature, including original papers, text books, and floras. In total, 567 extant and 10 fossil species are shared between this new dataset (of 1201 extant and 24 fossil species) and previously published datasets (of 792 extant and 10 fossil species; Sauquet et al. 2017; Schönenberger et al. 2020). We curated and filled-in gaps in these published datasets and hence, out of 13,812 total data records for these overlapping species, 13,730 were previously published, 82 are new, 169 were updated, and 52 were deleted. The majority of the remaining 634 extant and 14 fossil species were scored here for the first time, corresponding to a total of 12,785 new data records. The complete list of 26,597 data records, along with explicit references (in total 2020 distinct sources) used to score each species is provided as Supplementary Information Data 2. As in previous publications based on the PROTEUS database, the data were recorded as primary continuous or discrete characters, and then transformed into a matrix of secondary, discrete characters for the purpose of analysis (see Sauquet et al. 2017, Schönenberger et al. 2020). The final morphological matrix consists of 36,750 cells in total (36,030 for extant and 720 for fossil taxa), with a proportion of missing data of approximately 32% for extant and 23% for fossil species, including inapplicable and polymorphic characters. The matrix of joint molecular and morphological data is provided as Supplementary Information Data 3.

Constrained analyses

To estimate the phylogenetic placement of fossils, we conducted multiple independent phylogenetic analyses (summarized in Table 2). The first type of analyses incorporated a molecular backbone tree to fully constrain the topology of extant species. For this purpose, we selected the maximum clade credibility time tree obtained under the 'relaxed calibration strategy' by Ramírez-Barahona et al. (2020) (Supplementary Information Data 4), which the crown age of angiosperms was modeled by a Laplace distribution with a mean = 144.26 and scale = 4.36, and included 238 fossil-based minimum age constraints for internal nodes. First, we estimated the phylogenetic position of each fossil independently following the *phyloscan* (here *CMP-1-Morph*) approach as implemented by Schönenberger et al. (2020). For these analyses, a parsimony score was calculated for each fossil (independently of other fossils) for every possible placement in the tree (all terminal and internal branches, a total of 1201+1199), based on the morphological matrix scored for the fossil of interest and all extant representatives. Each of the 24 *CMP-1-Morph* analyses were conducted with the R script published by Schönenberger et al. (2020) using the package *phangorn* v.2.5.2 (Schliep 2011). We decided to estimate the position of fossils based on parsimony because the vast majority of fossils have been investigated using only this criterion (Eklund et al. 2004; Gandolfo et al. 2004; Doyle and Endress, 2014). Second, we conducted analyses for each fossil as described for *CMP-1-Morph* but with a maximum likelihood (ML) optimization criterion (here *CML-1-Morph*). All ML analyses were performed with RAxMLv.8.0 (Stamatakis 2014), implementing the constrained topology among extant species with the option -g, and the ASC_MULTICAT model with the 'lewis' ascertainment bias correction (Lewis 2001), which corrects for the presence of only variable morphological characters in the matrix and avoids overestimation of branch lengths. A total of 1000 replicates of non-parametric bootstrap were obtained. Third, similarly to the *CML-1-Morph* analyses, we ran ML analyses with only extant species (*CML-Morph-Extant*). Although we do not expect topological changes given that the tree topology is constrained, we conducted this analysis to explore branch length estimation with only morphological characters. Fourth, we ran a single ML analysis incorporating the 24 fossil flowers simultaneously, using the morphological matrix for extant and fossil species (*CML-24-Morph*). This analysis was conducted with the same parameters used for each of the *CML-1-Morph* analyses. Fifth, we used ML and Bayesian inference approaches to estimate the placement of all fossils simultaneously. The main difference between these two methods is that in Bayesian inferences the user specifies prior probability distributions of model parameters. For these analyses, we used a matrix consisting of molecular data for extant species and morphological data for extant and fossil species. We refer to this matrix as the total evidence matrix (such as in Manos et al. 2007), although we do not consider fossil ages. As the topology of the tree remained constrained, this analysis was aimed at investigating the effects of branch lengths as well as missing molecular data for fossils on their phylogenetic placement. The constrained ML phylogenetic analysis (*CML-24-TE*) was conducted with RAxML v.8. (Stamatakis 2014), using the ASC_MULTICAT model with the 'lewis' ascertainment bias correction (Lewis 2001) for morphological data, and the GTR model for the following molecular partitions: *rbcL* 1st and 2nd positions; *rbcL* 3rd position; *atpB* 1st and 2nd positions; *atpB* 3rd position; *matK; ndhF*; and *nuclear*, as specified by Ramírez-Barahona et al. (2020). We ran 1000 non-parametric bootstrap replicates. ML phylogenetic analyses were performed in the CIPRES science gateway (Miller et al. 2010).

Table 2. List of phylogenetic analyses conducted in this study. NA, not applicable. *Indicates the analysis that did not achieve convergence, **Indicates the analysis that retrieved unexpected positions for some fossils.

Constraint	Optimization Criterion/model of branch length	Taxa subjected to phylogenetic estimation	Data	Analysis code
Backbone (relationships among extant taxa derived from molecular data)	parsimony	1 fossil (separate analyses for each of the 24 fossils)	morphological	*CMP-1-Morph*
	ML	1 fossil (separate analyses for each of the 24 fossils)	morphological	*CML-1-Morph*
	ML	NA	morphological	*CML-Morph-Extant*
	ML	24 fossils simultaneously	morphological	*CML-24-Morph*
	ML	24 fossils simultaneously	total evidence (morphological & molecular)	*CML-24-TE*
	BI	24 fossils simultaneously	total evidence (morphological & molecular)	*CBI-24-TE*
	BI	24 fossils simultaneously	molecular	*CBI-24-Mol***
Unconstrained	ML	24 fossils & 1201 extant species simultaneously	total evidence (morphological & molecular)	*UML-24-TE*
	BI (linking branches across partitions)	24 fossils & 1201 extant species simultaneously	total evidence (morphological & molecular)	*UBI-24-TE-linked***
	BI (linking branches across partitions)	24 fossils & 1182 extant species (excluding parasitic taxa) simultaneously	total evidence (morphological & molecular)	*UBI-24-TE-linked-noparasitic*
	BI (unlinking branches across partitions)	24 fossils & 1201 extant species simultaneously	total evidence (morphological & molecular)	*UBI-24-TE-unlinked**

The constrained Bayesian analysis (*CBI-24-TE*) was performed using MrBayes v3.2.7 (Ronquist et al. 2012b) implementing the Mkv Markov Model (Lewis 2001) for categorical data, assuming that we only include variable characters (coding = variable), and for molecular data applying the same partitions as in the ML analysis. For molecular and morphological partitions, we used a gamma distribution for site-rate variation. The *CBI-24-TE* analysis used two replicate runs of 23 million MCMC generations sampling every 5000 generations. Finally, we evaluated the position of fossils in a Bayesian analysis of only molecular data for extant species with fossils scored entirely as missing data (herein *CBI-24-Mol*). The purpose of running this analysis was to identify potential methodological artifacts caused by the prior information, given that the use of inappropriate priors for some parameters such as branch lengths may lead to misleading results. The convergence for all Bayesian analyses was evaluated in Tracer v 1.7. 2 (Rambaut et al. 2018), checking that ESS values were above 200 for all parameters.

Unconstrained analyses

In the second set of analyses, no backbone tree was used to constrain the relationships among extant species, thus allowing direct phylogenetic estimation among extant and fossil species. We estimated phylogenetic relationships among extant and fossil species using the matrix combining molecular and morphological data described above, with ML and Bayesian Inference (BI). The unconstrained ML analysis (*UML-24-TE*) was conducted with RAxML v.8. (Stamatakis 2014) in the CIPRES science gateway (Miller et al. 2010), implementing the same set of partitions as in *CML-24-TE*. For the unconstrained BI analyses, we initially did not fix the relationships among extant taxa, but due to problems in MCMC convergence, and following Ramírez-Barahona et al. (2020), we constrained some well recognized groups (Supplementary Information Data 5). We ran three independent unconstrained Bayesian analyses: 1) an analysis than included all extant species (1201) with branch lengths linked across partitions (*UBI-24-TE-linked*); 2) an analysis in which we removed all parasitic species (except for Triuridaceae, because the fossil *Mabelia connatifila* has been previously assigned to this family) with branch lengths linked across partitions (*UBI-24-TE-linked-noparasitic*); and 3) an analysis that included all extant species but with branch lengths (brlens) estimated independently for molecular and morphological partitions by using the command *unlink brlens* (*UBI-24-TE-unlinked).* All Unconstrained Bayesian analyses were performed using MrBayes v3.2.7 (Ronquist et al. 2012b) implementing the same molecular and morphological partitions as the *CBI-24-TE* analysis. For each analysis, we conducted two replicate runs of four chains with a temperature of 0.2 for 30 million MCMC generations, sampling trees every 5000 generations. MCMC convergence was visualized in Tracer v 1.7. 2 (Rambaut et al. 2018). All Bayesian analyses were conducted using a parallel version of MrBayes installed on a HPC computer at the Instituto de Biología, UNAM. All files used to perform MrBayes analyses are provided in Supplementary Material.

Pseudo-fossilization analyses

To test the limitations of our morphological dataset in estimating the placement of fossils, we conducted independent phylogenetic analyses on extant taxa via pseudo-fossilization following the approach of Matsunaga et al. (2021). These analyses were aimed at testing whether, using only morphological data, we can retrieve phylogenetic positions of extant species similar to those obtained from molecular data. For this, we arbitrarily selected 27 extant species corresponding to major angiosperm groups and varying in the degree of missing data (from 0 to 83%). The selected species ("*pseudo-fossils*") were removed from the backbone tree, and their placement was subsequently inferred with the *CML-1-Morph* approach (described above). The complete list of "pseudo-fossils" is provided in Supplementary Information Table 1. To evaluate the accuracy of the phylogenetic position, we compared the original backbone tree and the trees with pseudo-fossils, including bootstrap replicates, by calculating their mean Robinson-Foulds (RF) distances (Robinson and Foulds 1981) using the package *phangorn* v2.5.5. in R (Schliep, 2011).

Visualization of phylogenetic uncertainty

We constructed RoguePlots (Klopfstein and Spasojevic 2019) as a graphical aid to visualize phylogenetic uncertainty in the placement of fossils in ML and BI phylogenetic analyses. This function summarizes all possible attachments of fossils to branches in the tree, each supported by its posterior distribution or bootstrap replicates. Instead of relying on a single summary position for a fossil, this method shows all sampled phylogenetic placements of a given fossil across the tree weighted by their support value. The RoguePlots graphs were constructed using the package *rogue.plot* (Klopfstein and Spasojevic 2019) in R (R Core Team 2021). This function colors the branches depending on the support value associated with that position. Hence, while all analyses presented in this paper included fossils as terminal taxa, RoguePlots do not show the fossil in question as a tip in the phylogeny.

RESULTS

Constrained analyses

The phylogenetic placements of fossils obtained with parsimony, ML, and BI are summarized in Table 3. The results are generally consistent across all methods, resulting in similar positions for the fossils but with slight variations. The *CMP-1-Morph* and *CML-1-Morph* analyses retrieved the same overall placement of fossils but differing in their optimal position. For example, the most parsimonious (MP) position of *Canrightia resinifera* corresponds to the family Stemonaceae (Pandanales), whereas positions one step longer than the most parsimonious one (MP+1) are in Chloranthaceae (Chloranthales), and in Thismiaceae (Dioscoreales) (Fig. S7a, Fig. S7b). In turn, the best supported maximum likelihood position is restricted to Chloranthaceae (70-80 BS) (Fig. S7c). For other ML approaches, the results were similar. However, *CML-24-TE* (Fig. S7e) associated fossils to a larger number of families than *CML-1-Morph* (Fig. S7c) and *CML-24-Morph* (Fig. S7d), as can be observed for fossils such as *Archaestella verticillatus*, *Bertilanthus scanicus*, *Dakotanthus cordiformis*, and *Florissantia quilchenensis*, among others (Table 3). In turn, the constrained Bayesian analysis of joint molecular and morphological data (*CBI-24-TE*) retrieved positions largely congruent with the other constrained analyses, with only one exception in the position of the fossil *Mabelia connatifila,* which was placed within Triuridaceae (Table 3). Our *CBI-24-Mol* analysis identified fossils attached to various branches in the tree with low support, with twelve fossils attaching to long terminal branches corresponding to parasitic species. We provide these results in Supplementary Figs. S1g-S24g.

Table 3. Summary of potential positions of fossils within major lineages and/or orders obtained across methods. * Indicates moderate confidence and ** Indicates high confidence

Fossil	CML-1-Morph	CML-24-Morph	CML-24-TE	CBI-24-TE	UML-24-TE	UBI-24-TE
Archaeanthus linnenbergeri	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	Magnoliales*	ANA grade, Magnoliidae	Magnoliales*
Archaestella verticillatus	Trochodendrales**	Trochodendrales*	ANA grade, Ranunculales, Trochodendrales	Magnoliidae, Trochodendrales	ANA grade, Ranunculales, Trochodendrales	Trochodendrales*
Bertilanthus scanicus	Saxifragales*	Saxifragales*	Saxifragales, Apiales, Caryophyllales	Saxifragales, Apiales	Saxifragales, Apiales, Asterales	Saxifragales*
Canrightia resinifera	Chloranthales, monocots	ANA grade, Chloranthales, Magnoliidae, monocots	ANA grade, Chloranthales, Magnoliidae, monocots	Monocots, Ranunculales	ANA grade, Chloranthales, Magnoliidae, monocots	Piperales**
Carpestella lacunata	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	Amborellales**
Cecilanthus polymerus	ANA grade, Chloranthales, Magnoliidae	ANA grade, Chloranthales, Magnoliidae	ANA grade, Chloranthales, Magnoliidae	ANA grade, Chloranthales	ANA grade, Chloranthales, Magnoliidae	ANA grade*
Chloranthistemon endressii	Chloranthales**	Chloranthales, Magnoliidae	Chloranthales*	Chloranthales*, Piperales, monocots	Chloranthales**	monocots
Dakotanthus cordiformis	Sapindales, Picramniales	Sapindales, Fabales, Crossosomatales	Sapindales, Fabales, Crossosomatales	Sapindales, Cucurbitales, Oxalidales	Sapindales, Fabales, Malpighiales, Crossosomatales	Oxalidales, Fabales, Sapindales
Divisestylus brevistamineus	Saxifragales, Cornales	Saxifragales*	Saxifragales*	Saxifragales, Vitales, Dilleniales	Saxifragales*	Saxifragales*
Florissantia quilchenensis	Malvales, Sapindales	Malvales, Malpighiales, Sapindales	Sapindales, Huerteales, Brassicales	Sapindales, Malpighiales, Ericales	Sapindales, Malpighiales, Huerteales, Malvales	Santalales, Malpighiales
Kajanthus lusitanicus	Ranunculales*	Ranunculales*	Ranunculales*	Ranunculales*	Ranunculales*	Ranunculales*
Mabelia connatifila	Malpighiales**	Malpighiales**	Malpighiales**	Pandanales, Piperales, Magnoliales	Malpighiales, monocots, Magnoliidae	Piperales*
Mauldinia mirabilis	Laurales*	Laurales*	Laurales*	Laurales*	Laurales*	Laurales*
Microvictoria svitkoana	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae	ANA grade, Magnoliidae

Paleoclusia chevalieri	Sapindales, Huerteales, Malvales	Sapindales, Malvales	Sapindales, Malvales	Sapindales, Malvales, Ericales	Sapindales, Malvales, Malpighiales	Ericales, Malpighiales, Sapindales
Paradinandra suecica	Sapindales, Malvales	Brassicales, Malvales	Brassicales, Malvales	Sapindales, Malvales, Malpighiales, Caryophyllales, Ericales	Sapindales, Malvales, Ericales	Caryophyllales, Sapindales
Pentapetalum trifasciculandricus	Dilleniales, Ericales	Dilleniales, Malpighiales, Ericales	Dilleniales, Caryophyllales, Ericales	Sapindales, Dilleniales	Dilleniales, Malpighiales, Ericales	Sapindales, Caryophyllales
Platydiscus peltatus	Myrtales, Crossosomatales	Myrtales, Sapindales, Crossosomatales	Myrtales, Saxifragales, Sapindales	Sapindales, Cucurbitales, Brassicales, Malvales	Cucurbitales, Myrtales, Sapindales, Saxifragales	Myrtales*
Quadriplatanus georgianus	Proteales*	Proteales, Ranunculales, ANA grade, Ceratophyllales	Proteales, Ranunculales, ANA grade	Malpighiales, Proteales	Proteales, Ranunculales, ANA grade, Saxifragales	Malpighiales**
Rariglanda jerseyensis	Geraniales, Myrtales, Sapindales	Geraniales, Sapindales, Myrtales, Brassicales, Malpighiales	Sapindales, Geraniales, Saxifragales, Crossosomatales	Ericales, Malpighiales, Sapindales	Sapindales, Ericales	Malpighiales**
Saururus tuckerae	Piperales, monocots	Piperales*	Piperales, Chloranthales, monocots	Piperales	Piperales, Chloranthales, monocots	Santalales, monocots
Spanomera mauldinensis	Buxales, Gunnerales, Trochodendrales, Saxifragales	Buxales, Gunnerales,	Buxales, Gunnerales, Ranunculales	Buxales*, Gunnerales, Santalales	Buxales, Gunnerales, Ranunculales, Saxifragales	Buxales*
Tylerianthus crossmanensis	Saxifragales, Trochodendrales	Saxifragales, Myrtales, Dilleniales, Asterales, Lamiales	Saxifragales, Lamiales, Solanales	Saxifragales*, Gentianales	Saxifragales, Lamiales, Asterales	Saxifragales, Escalloniales, Apiales
Virginianthus calycanthoides	Magnoliales, Laurales	ANA grade and Magnoliidae	ANA grade and Magnoliidae	Magnoliales*	ANA grade and Magnoliidae	Magnoliales*

Unconstrained analyses

Phylogenetic placements of fossils derived from unconstrained ML and BI analyses using molecular and morphological data are summarized in Table 3. For comparison, we only provide results from the unconstrained Bayesian analyses in which molecular and morphological branch lengths were linked (*UBI-24-TE-linked*); the MCMC of our analysis unlinking branch lengths across partitions (*UBI-24-TE-unlinked*) failed to converge to the stationary distribution. The unconstrained ML total evidence (*UML-24-TE*) phylogeny for extant and fossil species is shown in Fig. 1. Including fossils in the ML and BI analyses using molecular and morphological data did not change the topology among extant species, as inferred by Ramírez-Barahona et al. (2020) using only molecular data. However, the inclusion of morphological data in the analyses resulted in substantial decreases in the support values for most extant clades, as well as decreased resolution at deeper phylogenetic levels, presumably due to high uncertainty in the position of some fossils (Fig. S1 and Fig. S2). The placements of fossils were found to be weakly or moderately supported (indicated by blue and green triangles in Fig. 1), except for *Chloranthistemon endressii* (red triangle in Fig 1).

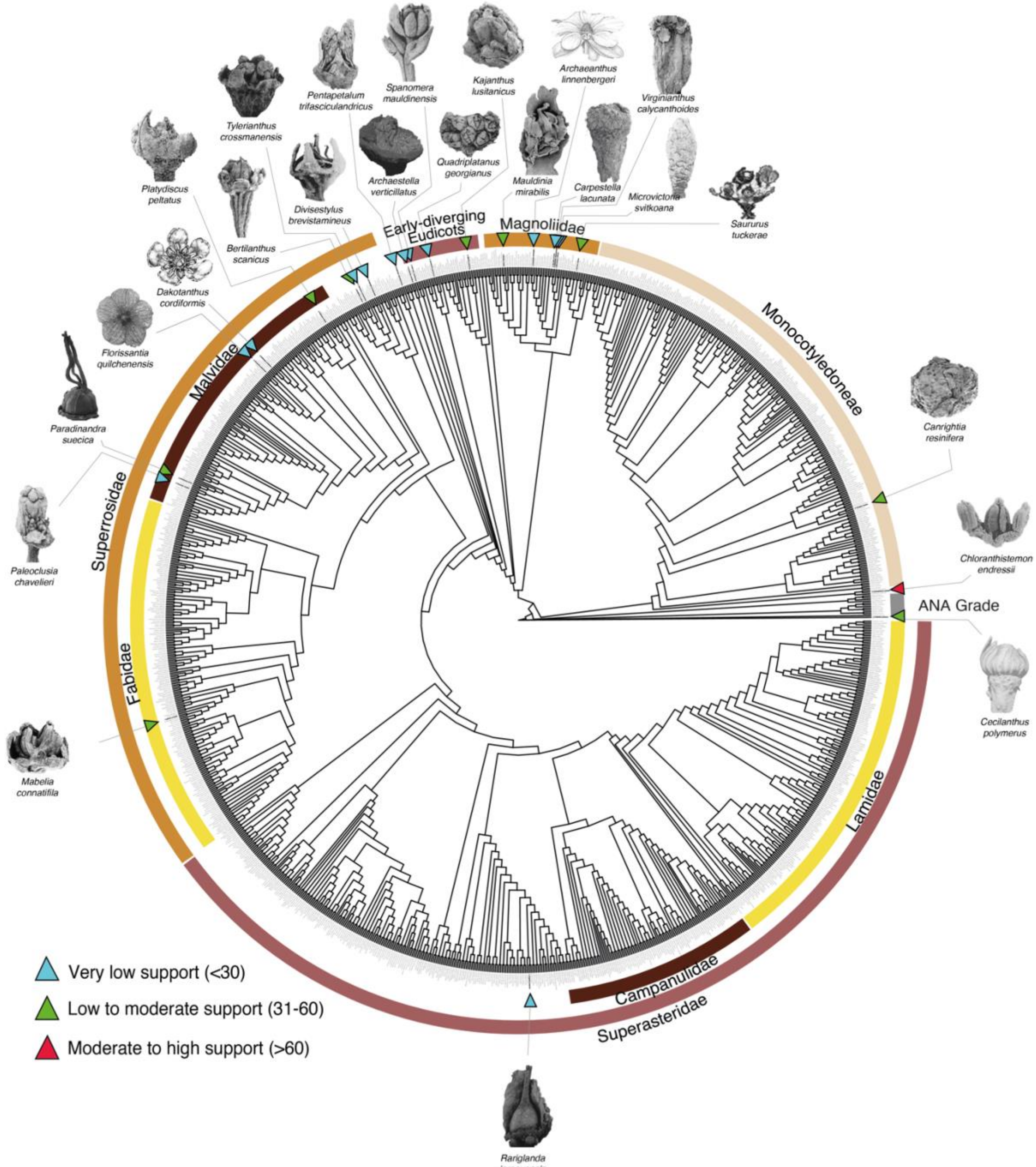

Figure 1. Maximum likelihood tree obtained with total evidence data indicating the position of 1201 extant species and 24 fossil flowers included in this study. The support values were obtained from bootstrap replicates. Blue triangles indicate very low support (<30); green triangles indicate low to moderate support (31-60); red triangles indicate moderate to high support values (>60). Although this analysis produced a phylogram (Fig S1), it is presented here as a cladogram for the sake of clarity. For detailed results of each fossil, including alternative positions, see Supplementary Information.

Major phylogenetic inconsistencies were obtained in the unconstrained BI total evidence analyses with respect to parsimony, ML analyses, and the constrained Bayesian analysis (Table 3). For instance, the results from parsimony, ML analyses, and the constrained Bayesian inference indicate *Chloranthistemon endressii* was associated with Chloranthaceae (*CML-1-Morph* (70-80 BS); *CML-24-Morph* (40-50 BS); *CML-24-TE* (60-70 BS); *CBI-24-TE* (50-60 PP); *UML-24-TE* (60-70 BS)); but in the *UBI-24-TE-linked* analysis, this fossil attached to the branch of *Pogonia japonica* (Orchidaceae) (40-50% PP). These apparent anomalous positions were also observed for other fossils such as *Florissantia quilchenensis*, *Mabelia connatifila*, *Paradinandra suecica,* and *Tylerianthus scanicus*. Additionally, we observed that some fossils were attached with high support to branches corresponding to extant parasitic lineages. For example, *Quadriplatanus georgianus,* a fossil with distinctive attributes of Platanaceae (Proteales), was placed on the branch leading to *Rafflesia keithii* (70-80% PP) (Fig. S19f), belonging to the parasitic family Rafflesiaceae in the Malpighiales; while *Canrightia resinifera* with distinctive attributes of Chloranthaceae and Piperales, was placed with high support (80-90%PP) along the branch leading to *Hydnora africana* (Aristolochiaceae, Piperales), another parasitic plant (Fig. S4f). As expected from the results of the *CBI-24-Mol* analysis, which indicated 12 out of 24 fossils attracted to long terminal branches, removing those species with elevated molecular rates (parasitic species) from the unconstrained total evidence (*UBI-24-TE-linked-noparasitic*) analysis retrieved more consistent fossil placements, which were highly congruent with the *CBI-24-TE* analysis (Supplementary Figs. S1i-S24i).

Pseudo-fossilization analyses

Results of the pseudo-fossilization analyses are available in Supplementary Information Table 1 and Supplementary Fig. S3. The mean RF distance ranged from 1.11 for *Acorus gramineus* (Acorales, monocots) to 67.97 in *Peperomia caliginigaudens* (Piperales, magnoliids). We did not identify a pattern associated with the recovery of inaccurate position of "*pseudo-fossils*". We confirmed our expectations that the number of characters scored does not appear to impact the placement of pseudo-fossils. For example, the pseudo-fossil *Trithuria cowieana* (Nymphaeales, ANA grade) has 12 characters scored, and an associated mean RF distance of 4.38 indicating highly consistent placement even when molecular data are not considered, whereas *Lathraea clandestina* (Lamiales, Lamidae) had 29 characters scored and an associated mean RF distance of 32.98. However, the pseudo-fossil *Peperomia caliginigaudens* had the highest associated RF distance and corresponds to the species with the largest amount of missing data (only five characters scored).

Uncertainty in the placement of fossils

Due to the high uncertainty associated with the placement of most fossils, we recognize the importance of considering all the relationships inferred for fossils, rather than a single most supported position. We categorized the position of fossils into distinct levels of confidence, depending on the number of estimated distinct phylogenetic positions, and the associated support. We attributed "high confidence" to the placement of fossils attached to a single branch with high support (>80%) (BS and PP) (e.g., *Chloranthistemon endressii* attached to the branch leading to *Sarcandra chloranthoides* with 90-100%, Fig. 2a). We attributed "moderate confidence" to the placement of fossils attached to a single branch in different analyses with moderate support (e.g., *Cecilanthus polymerus*, Fig. 2b), or attached to closely related branches (e.g., within the same family or order) with low support (e.g., *Kajanthus lusitanicus*, Fig. 2c; or *Bertilanthus scanicus*, Fig. 2d). Finally, we attributed "low confidence" to the placement of fossils that attached to several branches belonging to widely divergent clades with weak support (e.g., *Carpestella lacunata,* Fig. 4).

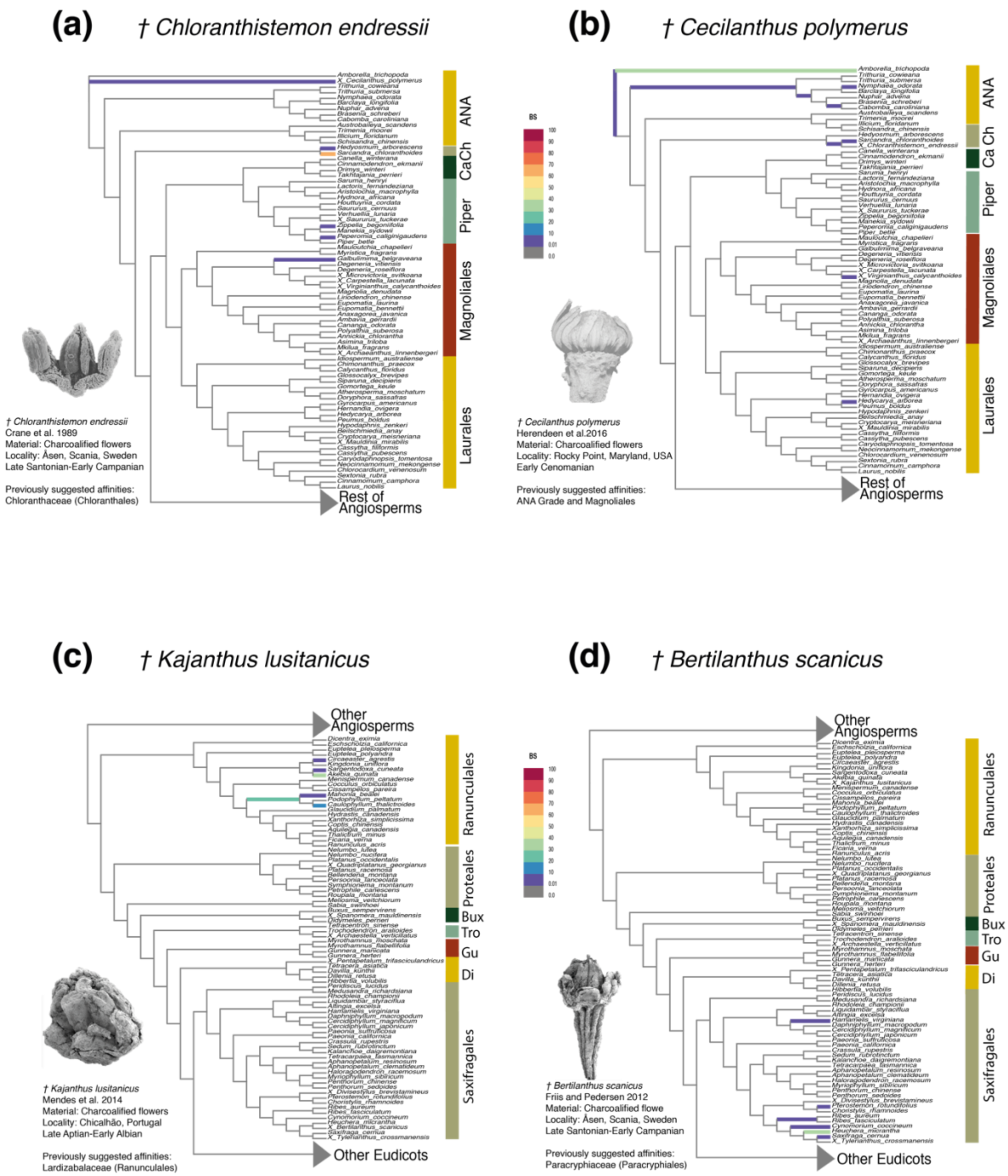

Figure 2. Partial RoguePlots derived from the unconstrained maximum likelihood analysis of molecular and morphological characters. The plots indicate all alternative phylogenetic positions of fossils and their associated uncertainty. The branches are colored according to the bootstrap values (BS values in color legend) associated with the attachment of the fossil to the branch. a) *Chloranthistemon endressii*; b) *Cecilanthus polymerus*; c) *Kajanthus lusitanicus*; d) *Bertilanthus scanicus*. Abbreviations: ANA, ANA grade (*Amborella*, Nymphaeales, Austrobaileyales); Ca, Canellales; Ch, Chloranthaceae; Piper, Piperales; Bux, Buxales; Tro, Trochodendrales, Gu, Gunnerales; Di, Dilleniales.

The phylogenetic placement of some fossils was recovered with similar levels of confidence across different phylogenetic analyses. For instance, *Archaestella verticillatus* was moderately supported as linked to the branch leading to *Trochodendron aralioides* (Trochodendraceae) in all analyses (Fig. 3). In contrast, the placement of some fossils exhibited substantial differences in support across different approaches. In general, BI analyses recovered placements that were restricted and strongly supported in particular positions. For example, *Carpestella lacunata* showed high phylogenetic uncertainty (low confidence) in ML analyses (Fig. 4), being attached with low support (<10-30 BS) to several independent branches across the ANA grade (Amborellaceae, Nymphaeaceae, and Austrobaileyaceae) and the orders Piperales, Magnoliales, and Laurales within the Magnoliidae clade. Conversely, in the total evidence BI analysis, the same fossil was attached to *Amborella trichopoda* with high support (90-100% PP). A similar pattern was observed in *Cecilanthus polymerus* (Fig. S6f).

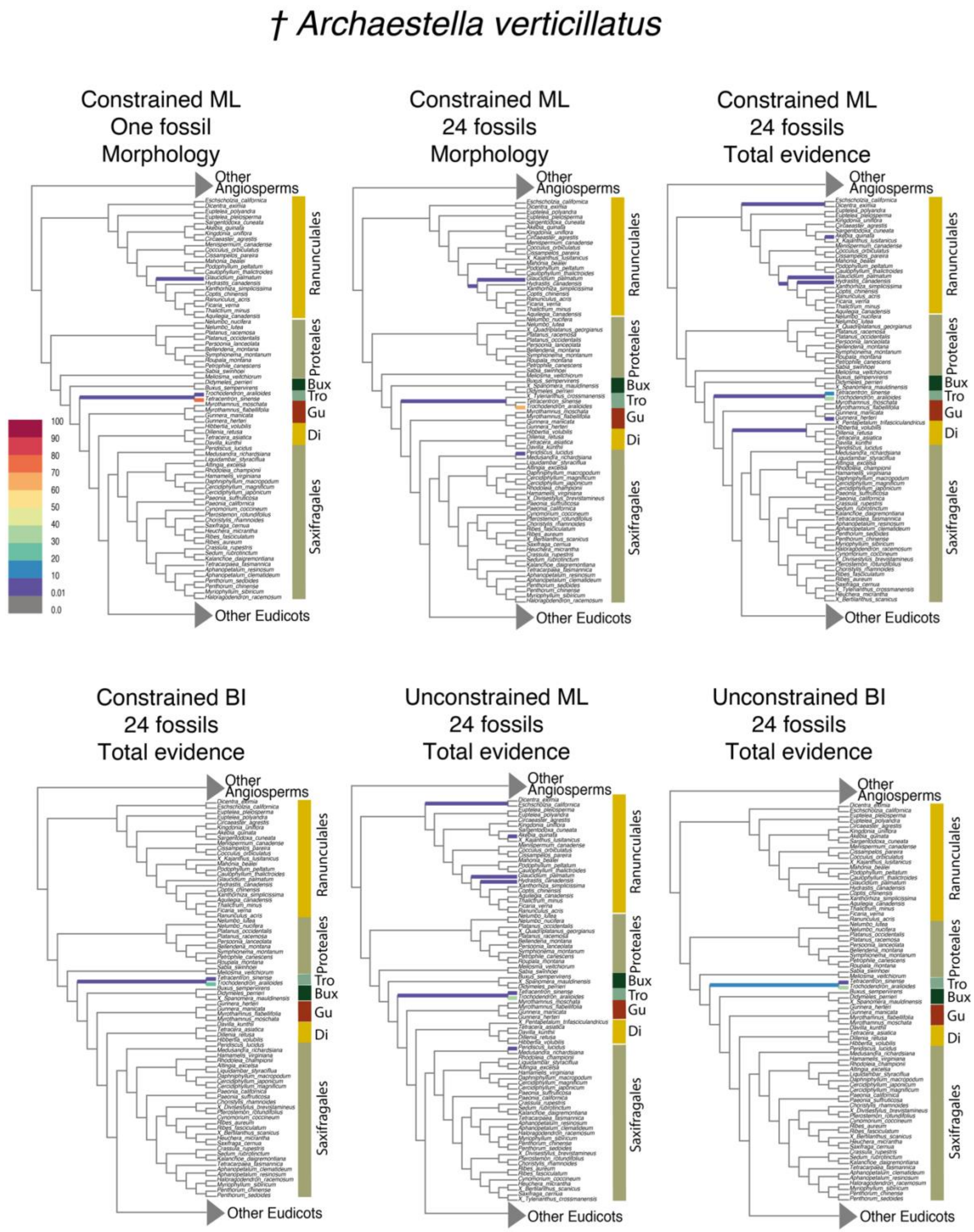

Figure 3. Partial RoguePlots indicating the phylogenetic placement of *Archaestella verticillatus* obtained from the six parametric methods implemented. The branches of the tree are colored according to the bootstrap and posterior probability (%) values associated with the positions retrieved (BS values in color legend). Abbreviations: Bux, Buxales; Tro, Trochodendrales; Gu, Gunnerales; Di, Dilleniales.

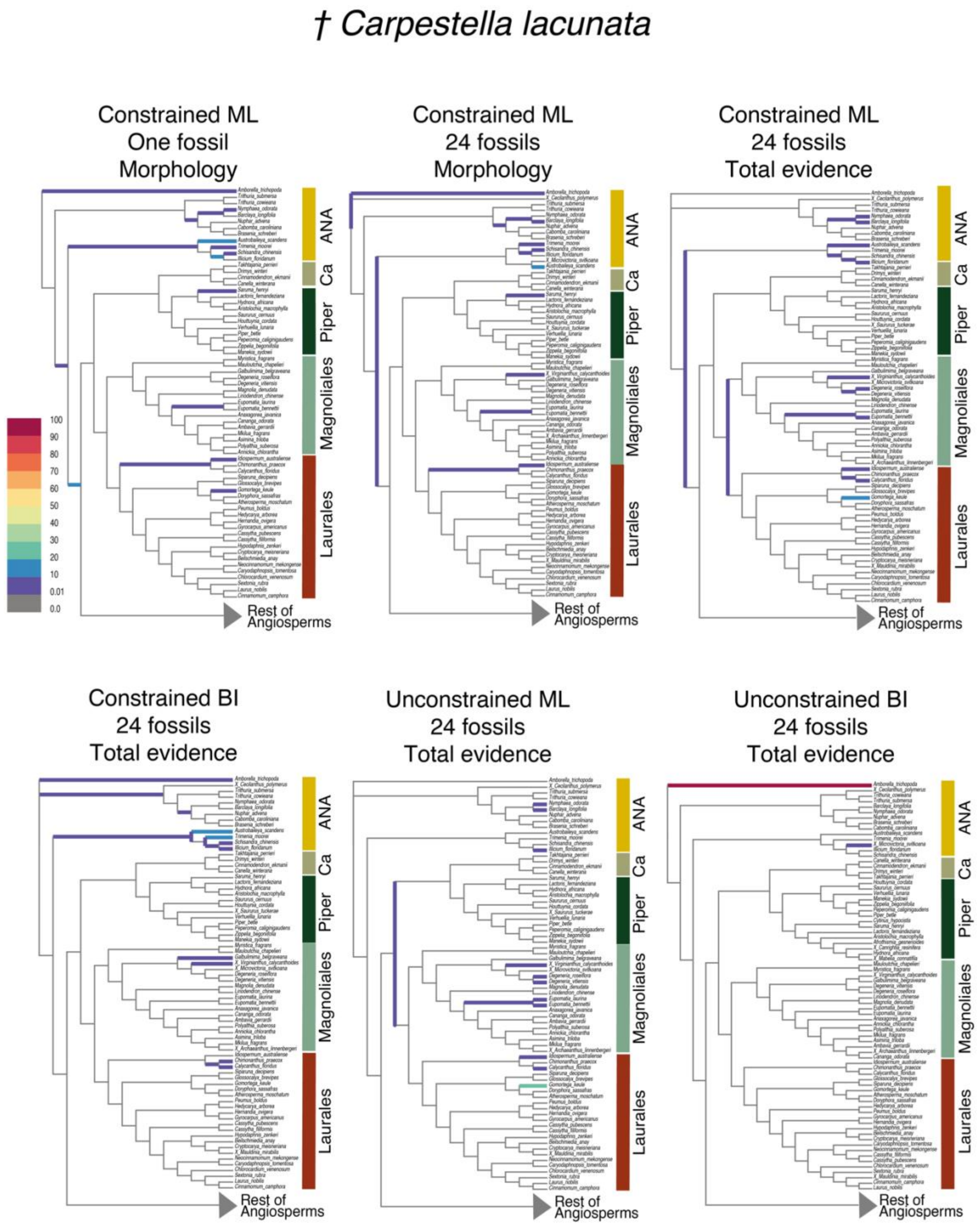

Figure 4. Partial RoguePlots indicating the placement of *Carpestella lacunata* obtained from the six parametric methods implemented. The branches of the tree are colored according to the bootstrap and probability (%) values associated with the positions retrieved (BS values in color legend). Abbreviations: ANA, ANA grade (*Amborella*, Nymphaeales, Austrobaileyales); Ca, Canellales; Piper, Piperales.

DISCUSSION

Methods to Place Fossils into Phylogenies

In this study, we estimated the placement of fossil flowers across the angiosperm phylogeny using different optimization criteria and showed that, regardless of method and datasets, the different methods produced largely similar topological placements of fossils. In the past, the phylogenetic placement of angiosperm fossils has been mostly based on parsimony approaches focusing on individual fossils by using only morphological characters (e.g., Doyle and Endress, 2014; Eklund et al. 2004). Previously, Bayesian analyses of joint molecular and morphological data have been reported to yield high accuracy in the topology in studies that incorporate extinct tips (Guillerme and Cooper 2016). Our constrained Bayesian analysis of morphological and molecular data (*CBI-24-TE*) was largely congruent with previous affinities of fossils and other approaches considered in this study (Table 3; Supplementary Figs. S1h-S24h). Compared with the *UBI-24-TE-linked* approach, the *CBI-24-TE* analysis did not identify fossils attached with strong support to parasitic lineages having long branch lengths, and in fact, this approach recovered previously suggested positions for some fossils (e.g., *Mabelia connatifila* within Triuridaceae). In contrast, the unconstrained Bayesian analysis linking branch lengths resulted in unexpected phylogenetic placements for some fossils. For example, fossils attached to long branches, which suggests this approach might not be appropriate under some conditions, such as in the incorporation of multiple extant terminals with increased substitution rates, as observed here for parasitic and mycoheterotrophic species. This result might be circumvented by unlinking branch lengths across molecular and morphological partitions, which could be a more realistic approach because it does not assume morphology and molecular share the same rates. Even though more complex models are associated with less conflict in combined morphological and molecular analyses, this might slow convergence (see Nylander et al. 2004). In complex models, the number of parameters to be estimated increases considerably, incurring greater uncertainty of topological and branch-length estimates (see Duchêne et al. 2020), and making the analysis time-consuming and nearly impossible to achieve MCMC convergence regardless of the computational resources available. Additionally, there are other potential causes that might be problematic in estimating the model posterior such as the use of uninformative and conflicting prior specifications (Van Dongen, 2006), or how weak signal interacts with the prior. Thus, based on our results, we recommend the implementation of the unconstrained maximum likelihood approach or Bayesian inferences either constraining the topology of extant species or implementing different branch-length estimation models for molecular and morphological partitions to integrate fossils into large-scale phylogenetic frameworks.

We advocate the use of broad-scale phylogenetic analyses as an important tool when estimating the placement of fossils. This approach may offer some key advantages in the assignment of fossils among living species. For example, this can reveal relationships that were not immediately obvious, and avoid the potential confinement of fossils in erroneous clades imposed *a priori*. As also stressed by Schönenberger et al. (2020), we consider that a more reliable estimation of the phylogenetic placement of any given fossil may be obtained in a two-step process. The first step relies on broad-scale analysis to identify different potential affinities of the fossil among major clades (e.g., Tank et al 2022). A second step may then focus on a particular clade and involve a more detailed morphological analysis and including a denser taxon and character sampling. Of course, it would be optimal to only conduct a single phylogenetic analysis. However, this would require the assembly of a morphological matrix that includes extensive character and taxonomic sampling to capture the full range of variation both within and among clades.

Missing Data and Morphological Signal in the Integration of Fossils into Phylogenies

Analyses of joint molecular and morphological data are often afflicted with different sources of missing data, which has been considered potentially problematic for achieving accurate phylogenetic estimations. Indeed, understanding the impact of missing data has been an important subject of active study (Kearney and Clark 2003; Wiens 2003; Magallón 2007; Manos et al. 2007; Wiens and Morrill 2011; Guillerme and Cooper 2016; Mongiardino Koch et al. 2021). Fossils represent a particularly challenging case because they lack molecular information and the number of traits that can be scored is typically limited due to the fragmentary nature of preservation, especially in plants. However, studies of empirical and simulated data have demonstrated that including fossils can have a strong positive impact on phylogenetic estimation; fossils potentially allow to break long branches and may provide character combinations not present among extant species, which may produce significant topological rearrangements (Huelsenbeck, 1991; Cobbet et al. 2007).

Our results corroborate previous findings suggesting that the number of morphological characters scored for fossils does not strongly impact their estimated phylogenetic placement (Manos et al. 2007). For example, *Chloranthistemon endressii* is a highly incomplete fossil flower, of which only the three-lobed androecium was recovered while other parts of the flower, i.e., the perianth and the gynoecium, remain unknown

(Crane et al. 1989). Given the set of floral characters we used, *C. endressii* inevitably has a high proportion of missing (or inapplicable) data; we were able to score only 12 out of 30 morphological characters. Despite this large amount of missing data, we found this fossil to be consistently placed in Chloranthaceae (with the exception of our *UBI-24-TE* analysis), in agreement with previous studies (Crane et al. 1989; Doyle and Endress, 2018; Eklund et al. 2004; Schönenberger et al. 2020). Counterintuitively, the phylogenetic placement of some fossils for which many characters were scored (e.g., *Dakotanthus cordiformis*, *Florissantia quilchenensis*, *Rariglanda jerseyensis*) remained generally equivocal, being associated with multiple distantly related families across all analyses. This suggests that obtaining reliable positions is increased for fossils displaying a particular set of highly distinctive morphological characters (as also concluded by Mongiardino Koch et al. 2021), especially if those characters are synapomorphies of particular clades (Manos et al. 2007). For example, the androecium of *C. endressii*, consisting of three stamens that are fused at the base, and with prominent connective extensions, represents a unique combination of characters shared only within modern Chloranthaceae (Crane et al. 1989).

An additional source of missing data is the often incomplete knowledge of the morphology of extant species. For instance, some studies make use of previously published morphological matrices, whose taxon sampling does not fully overlap with the molecular matrices, thereby resulting in a significant amount of missing data (e.g., Arcila et al. 2015). Based on simulations, Guillerme and Cooper (2016) proposed that the morphological data collected for extant species plays a major role in increasing the accuracy of phylogeny estimation under a total evidence approach. Our morphological matrix contained 32% of morphological characters missing, including inapplicable characters. Although we did not detect any suspicious patterns of fossil attachment to extant species with high or low proportions of missing morphological data, the results of *CML-Morph-Extant* and pseudo-fossilization analyses suggest that our morphological dataset might have some limitations in accurately placing flowers. The informativeness of morphological characters is highly heterogeneous across different clades in the angiosperm tree. This means there are some characters that might be uninformative (weak signal) to precisely determine the placement of fossils to specific family/order within the largest clade Pentapetalae, for example, for flowers exhibiting a generalized structural pattern (i.e., pentamerous flowers, bipartite perianth, whorled phyllotaxy) which are highly stable within Pentapetalae (see discussion above), but simultaneously those characters are actually informative to determine which fossils do not belong to that clade. Therefore consider that a future increase in taxon sampling and morphological characters may capture wider floral variation and consequently be helpful to further resolve the phylogenetic position of several fossil flowers retrieved here with high uncertainty.

The presence of long terminal branches might also affect the position retrieved for fossils. Our results from the *UBI-24-TE-linked* analysis reconstructed particular fossils associated with parasitic and mycoheterotrophic lineages, characterized by elevated synonymous and nonsynonymous substitution rates (Graham et al. 2017). In addition, the analysis of running the fossils under the prior (*CBI-24-Mol*) retrieved half of the fossils tested attracted to long branches, which might suggest a strong impact of the tree model. The use of combined morphological and molecular data in phylogenetic analyses, in particular, the integration of fossils (for which molecular data is missing) and extant species with increased rates of molecular evolution, might be more accurately analyzed by unlinking molecular and morphological branch-length estimation models. To analyze this, we conducted the UBI-24-TE-unlinked analysis but failed to converge, thus more work would be needed, either analyzing smaller (empirical) datasets or using simulations, to determine the effect of assuming partitioned branch-length models in the estimation of fossil positions.

Phylogenetic Placement of Fossil Flowers within a Large Angiosperm Phylogeny

Our results show that the majority of fossil flowers were placed within orders or families that had been previously suggested based on detailed morphological comparisons (e.g., *Archaestella verticillatus* in Trochodendraceae, Takahashi et al. 2017; *Quadriplatanus georgianus* in Platanaceae, Magallón-Puebla et al. 1997; *Spanomera mauldinensis* in Buxaceae, Drinnan et al. 1991). Moreover, some fossil placements coincide with results from parsimony clade-specific analyses, such as the positions of *Divisestylus brevistamineus* in Saxifragales (Hermsen et al. 2003), *Saururus tuckerae* in Piperales (Smith and Stockey 2007), and *Chloranthistemon endressii* in Chloranthales (Eklund et al. 2004). These results suggest that a priori hypotheses about the placement of certain fossils based on expert observations are highly valuable, but also highlight the ability of our approach in recovering similar placements even with a relatively low number of floral traits (compared with ~75 floral characters used in Doyle and Endress, 2014). We also obtained fossil positions similar to previous broad-scale analyses based on parsimony, comparable to the *CMP-1-Morph* approach used here, such as the placement of *Carpestella lacunata* in Nymphaeaceae and Austrobaileyaceae, two families within the ANA grade (Fig. 4; Table 3; Fig. 3 in Doyle and Endress, 2014; Fig. 4 in von Balthazar et al. 2008).

On the other hand, for some fossils we obtained new and sometimes unexpected phylogenetic placements across all phylogenetic analyses. For instance, *Mabelia connatifila* was previously suggested to belong to Triuridaceae in the monocot order Pandanales, based on morphological comparison and parsimony analyses (Gandolfo et al. 2002). Our results from *CBI-24-TE* support this finding, however the parsimony and ML approaches instead consistently placed this fossil in family Salicaceae, nested in the eudicot order Malpighiales, with high support (Table 3; Figures S12a, S12b, S12c, S12d, S12e, S12g). Another example is the consistent placement of *Bertilanthus scanicus* in Saxifragaceae (Saxifragales, Pentapetalae) across all approaches (Table 3), in contrast with suggested affinities to the distantly related Paracryphiaceae (Paracryphiales) (Friis and Pedersen, 2012).

These unexpected positions could be explained by the weak phylogenetic signal provided by some generalized morphological characters, as shown by short branch lengths given by the analysis of only morphological data (Supplementary Information Fig. S4). For this study, we attempted to include floral traits that could be comparable and scorable across all angiosperm families as well as fossil taxa. Our morphological dataset mostly comprises features of the floral organization or groundplan (sensu Endress, 1996), such as the number of organs in whorls (merosity) and their arrangement (phyllotaxy). These characters tend to be relatively stable at deeper phylogenetic levels among major angiosperms clades compared to traits of the floral style or mode, which mostly vary between closely related groups, and therefore seem to be more prone to homoplasy (Endress, 1996).

Particularly, the floral organization traits within Pentapetalae were apparently established at the onset of the diversification of eudicots and remained mostly unchanged during subsequent evolution (Endress 1994; Sauquet et al. 2017). In particular, most members of Pentapetalae display flowers characterized by a whorled phyllotaxis, a differentiated perianth (calyx and corolla), and a pentamerous (or tetramerous) perianth and androecium (Endress 2010; Sauquet et al. 2017). Fossils such as *Platydiscus peltatus, Dakotanthus cordiformis, Paradinandra suecica,* and *Paleoclusia chevalieri* all share this generalized and conserved structure and are therefore less likely to be found in restricted unique phylogenetic positions in angiosperm-wide analyses. In order to obtain stronger signals at shallower phylogenetic levels, additional characters at the level of floral construction and mode are needed (e.g., size, shape, and conformation of complex and novel structures).

Future Analyses and Perspectives

Our broad-scale phylogenetic analyses of fossil flowers, derived from a comprehensive and curated new morphological dataset for living species across all angiosperm families, provide a helpful tool for future paleobotanical investigation. We would like to reiterate that this study has corroborated the incredible taxonomic expertise that paleobotanists have made for a long time in the investigation of fossil flowers. However, we believe it is important to make use of large-scale datasets and approaches (knowing their limitations) as replicable tools to complement their work. We also document the outcome of different strategies for combining fossil and extant species in phylogenetic analyses, which resulted in varying levels of uncertainty in fossil relationships. Our approach could be used in subsequent studies, for example, to estimate the placement of a larger number of fossil flowers, including newly discovered or undescribed fossils of unknown affinities, with respect to extant families and orders as recommended by Sauquet and Magallón (2018). This approach is also useful to corroborate previous assignments of fossils based on morphological assessments and to evaluate alternative placements across the phylogeny. Our morphological matrix and analytical approaches, especially those indicating the level of phylogenetic uncertainty associated with positions, will be useful to estimate the general phylogenetic relationships of a large number of previously described fossil flowers. The morphological data set used in our analyses is the result of considerable effort in scoring and curating data covering all currently expected angiosperm families. Further improvement and extensions to this data set in the future will allow for greater certainty about the phylogenetic placement of fossil flowers. Therefore, we suggest conducting similar efforts on building and expanding morphological datasets, not only in terms of the number of living species that are scored, but most importantly in the number and type of characters (e.g., incorporating detailed floral traits or vegetative traits).

Integration of fossils into extant phylogenies has become more popular since the implementation of the Fossilized Birth-Death model (FBD) (Heath et al. 2014) and dating methods that can incorporate total evidence matrices (Ronquist et al. 2012, 2016; Zhang et al. 2016; Gavryushkina et al. 2017). The latter approach integrates molecular, morphological, and stratigraphic data of fossils into a single framework with extant species to simultaneously infer phylogenetic relationships and divergence times. Several studies have suggested that stratigraphic and morphological data are both informative in the estimation of phylogenetic relationships (e.g., King 2021), and that the use of fossil ages might be helpful to improve phylogenetic estimation, particularly when analyzing highly homoplastic morphological characters (Lee and Yates, 2018; May et al. 2021). However, simulations have shown that tip-dating under the fossilized birth-death model has difficulties in correctly placing fossils in the tree (Luo et al. 2020). Hence, we consider that undated phylogenetic analyses, as presented in this paper, and tip-dating analyses could be reciprocally informative in assessing the position of fossils, as we are starting to understand the potential limitations of both approaches. Additionally, we would also point out the importance of taking into account the phylogenetic uncertainty of fossils to carefully select those that can be incorporated into subsequent macroevolutionary inferences. For example, we can reduce the possibility of calibrating an erroneous internal node or create adequate taxonomic constraints for FBD analyses.

Phylogenetic approaches that integrate fossils and extant species are essential to provide a more comprehensive picture of deep time macroevolutionary processes. Examples of such studies include the investigation of fruit evolution through time in Fagales (Larson-Johnson 2016), understanding the timing and diversification dynamics of Cetacea (Lloyd and Slater 2021), or inferring the biogeographical history of Juglandaceae (Zhang et al. 2021) and *Hypericum* (Meseguer et al. 2015). The fossil record of angiosperms is diverse and relatively abundant since their initial appearance in the Early Cretaceous (Friis et al. 2010). As in any other group of organisms, the angiosperm fossil record is also biased in different ways (i.e., temporally, spatially, and phylogenetically, as discussed by Xing et al. 2016). However, we expect that the continued compilation of large-scale morphological data sets and the subsequent integration of fossil flowers (and other types of fossils) into the angiosperm phylogeny will lead to the elucidation of deep time patterns of floral evolution.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: doi:10.5061/dryad.cfxpnvx73

ACKNOWLEDGEMENTS

We are grateful to the University of Vienna for funding the eFLOWER server hosting the PROTEUS database. We thank the numerous colleagues who have contributed to building this new eFLOWER dataset over the past five years, in particular Elisabeth Reyes, Luka Kovac, and all the participants from the Oak Spring eFLOWER Summer School. Peter Crane, the Oak Spring Garden Foundation, and the Society of Systematic Biologists are gratefully acknowledged for their support of this Summer School. A.L.M thanks the Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, and the Consejo Nacional de Ciencia y Tecnología (CONACyT) for granting scholarship 708068. We thank the Instituto de Biología, Proyecto Fronteras de la Ciencia 2016-01-1867-CONACyT, Unidad TIC, M. I. David Velázquez and Biól. Alfredo Wong, the access to the HPC infrastructure at the Instituto de Biología, UNAM. We are grateful for the comments provided by April Wright (Associate Editor), and three anonymous reviewers which greatly improved the paper. We thank Natalia Ivalú Cacho for comments on earlier drafts; Alfonso Delgado, Luis E. Eguiarte, Rebeca Hernández, and Adriana Benítez for feedback and helpful discussions during different stages of this project.

AUTHOR CONTRIBUTIONS

A.L.M., H.S., J.S., M.v.B., and S.M. designed the study. A.L.M., J.S., M.v.B., C.G.M., and H.S. scored and curated the morphological data of extant and fossil taxa. S.R.B. provided the backbone phylogeny and the molecular alignment. A.L.M conducted the analyses and wrote this paper, with subsequent input from all co-authors.

LITERATURE CITED

Arcila D., Alexander Pyron R., Tyler J.C., Ortí G., Betancur R. 2015. An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). Mol. Phylogenet. Evol. 82:131–145.

Benson R.B.J., Butler R., Close R.A., Saupe E., Rabosky D.L. 2021. Biodiversity across space and time in the fossil record. Curr. Biol. 31:R1225–R1236.

Betancur-R R., Ortí G., Pyron R.A. 2015. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. Ecol. Lett. 18:441–450.

Bond W.J., Scott A.C. 2010. Fire and the spread of flowering plants in the Cretaceous. New Phytol. 188:1137–1150.

Chase M.W., Christenhusz M.J.M., Fay M.F., Byng J.W., Judd W.S., Soltis D.E., Mabberley D.J., Sennikov A.N., Soltis P.S., Stevens P.F. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181:1–20.

Cobbett, A., Wilkinson, M. and M. A. Wills. 2007. Fossils impact as hard as living taxa in parsimony analyses of morphology. Syst. Biol. 56 (5):753-766

Crane P.R., Friis E.M., Pedersen K.R. 1989. Reproductive structure and function in Cretaceous Chloranthaceae. Plant Syst. Evol. 165:211–226.

Crepet W.L., Nixon K.C. 1998. Fossil Clusiaceae from the Late Cretaceous (Turonian) of New Jersey and implications regarding the history of bee pollination. Am. J. Bot. 85:1122–1133.

Crepet W.L., Nixon K.C., Weeks A. 2018. Mid-Cretaceous angiosperm radiation and an asterid origin of bilaterality: diverse and extinct "Ericales" from New Jersey. Am. J. Bot. 105:1412–1423.

Dilcher D.L., Crane P.R. 1984. *Archaeanthus*: An early angiosperm from the Cenomanian of the Western Interior of North America. Ann. Missouri Bot. Gard. 71:351.

Doyle J.A., Endress P.K. 2010. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. J. Syst. Evol. 48:1–35.

Doyle J.A., Endress P.K. 2014. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: ANITA lines and relatives of Chloranthaceae. Int. J. Plant Sci. 175:555–600.

Doyle J.A., Endress P.K. 2018. Phylogenetic analyses of Cretaceous fossils related to Chloranthaceae and their evolutionary implications. Bot. Rev. 84:156–202.

Drinnan A.N., Crane P.R., Friis E.M., Pedersen K.R. 1990. Lauraceous flowers from the Potomac Group (mid-Cretaceous) of eastern North America. Bot. Gaz. 151:370–384.

Drinnan A.N., Crane P.R., Friis E.M., Pedersen K.R., Drinnan A.N., Peter R., Friis E.M., Pedersen K.R. 1991. Angiosperm flowers and tricolpate pollen of Buxaceous affinity from the Potomac Group (Mid-Cretaceous) of Eastern North America. Am. J. Bot. 78:153–176.

Duchêne DA, Tong KJ, Foster CSP, Duchêne S, Lanfear R, Ho SYW. 2020. Linking branch lengths across sets of loci provides the highest statistical support for phylogenetic inference. Mol Biol Evol. 37:1202-1210.

Eklund H., Doyle J.A., Herendeen P.S. 2004. Morphological phylogenetic analysis of living and fossil Chloranthaceae. Int. J. Plant Sci. 165:107–151.

Endress P.K. 1994. Diversity and evolutionary biology of tropical flowers. Cambridge University Press.

Endress P.K. 2010. Flower structure and trends of evolution in eudicots and their major subclades. Ann. Missouri Bot. Gard. 97:541–583.

Friis E., Pedersen K., Crane P. 2001. Fossil Evidence of water lilies (Nymphaeales) in the Early Cretaceous. Nature. 410:357–60.

Friis E.M., Crane P.R., Pedersen K.R. 2011. Early flowers and angiosperm evolution. Cambridge: Cambridge University Press.

Friis E.M., Crane P.R., Pedersen K.R. 2019. The Early Cretaceous Mesofossil Flora of Torres Vedras (Ne of Forte da Forca), Portugal: A Palaeofloristic Analysis of an Early Angiosperm Community. Foss. Impr. 75:153–257.

Friis E.M., Crane P.R., Pedersen K.R. 2020. *Melloniflora*, a new extinct multiparted flower from the early cretaceous of virginia, usa. Int. J. Plant Sci. 181:887–897.

Friis E.M., Eklund H., Pedersen K.R., Crane P.R. 1994. *Virginianthus calycanthoides*.- A calycanthaceous flower from the Potomac Group (Early Cretaceous) of Eastern North America. Int. J. Plant Sci. 155:772–785.

Friis E.M., Marone F., Pedersen K.R., Crane P.R., Stampanoni M. 2014. Three-dimensional visualization of fossil flowers, fruits, seeds, and other plant remains using synchrotron radiation X-ray tomographic microscopy (SRXTM): new insights into Cretaceous plant diversity. J. Paleontol. 88:684–701.

Friis E.M., Pedersen K.R. 2011. *Canrightia resinifera* gen. et sp. nov., a new extinct angiosperm with Retimonocolpites-type pollen from the Early Cretaceous of Portugal: Missing link in the eumagnoliid tree? Grana. 50:3–29.

Friis E.M., Pedersen K.R. 2012. *Bertilanthus scanicus*, a new asterid flower from the Late Cretaceous (Late Santonian-Early Campanian) of Scania, Sweden. Int. J. Plant Sci. 173:318–330.

Friis E.M., Pedersen K.R., Crane P.R. 2006. Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. Palaeogeogr. Palaeoclimatol. Palaeoecol. 232:251–293.

Friis E.M., Pedersen K.R., Crane P.R. 2010. Diversity in obscurity: fossil flowers and the early history of angiosperms. Philos. Trans. R. Soc. B Biol. Sci. 365:369–382.

Friis E.M., Skarby A. 1981. Structurally preserved angiosperm flowers from the Upper Cretaceous of southern Sweden. Nature. 291:484–486.

Gandolfo M.A., Nixon K.C., Crepet W.L. 1998. *Tylerianthus crossmanensis* gen. et sp. nov. (aff. Hydrangeaceae) from the Upper Cretaceous of New Jersey. Am. J. Bot. 85:376–386.

Gandolfo M.A., Nixon K.C., Crepet W.L. 2002. Triuridaceae fossil flowers from the Upper Cretaceous of New Jersey. Am. J. Bot. 89:1940–1957.

Gandolfo M.A., Nixon K.C., Crepet W.L. 2004. Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. Proc. Natl. Acad. Sci. U. S. A. 101:8056–8060.

Gavryushkina A., Heath T.A., Ksepka D.T., Stadler T., Welch D., Drummond A.J. 2017. Bayesian total-evidence dating reveals the recent crown radiation of penguins. Syst. Biol. 66:57–73.

Graham, S. W., Lam V.K.Y. Merckx, V.S.F.T. 2017. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. New Phytol.214: 48-55

Guillerme T., Cooper N. 2016. Effects of missing data on topological inference using a Total Evidence approach. Mol. Phylogenet. Evol. 94:146–158.

Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. Proc. Natl. Acad. Sci. U. S. A. 111.

Herendeen P.S., Doyle J.A., Endress P.K., Takahashi M. 2016. *Cecilanthus polymerus*, a novel multiparted flower from the mid-Cretaceous rocky point locality, Maryland. Botany. 94:787–803.

Heřmanová Z., Čepičková J., Kvaček J., von Balthazar M., Schönenberger J. 2022. A multicarpellate fruit from Late Cretaceous sediments of South Bohemia, Czech Republic. Paleontol. Electron. 25:5a.

Heřmanová Z., Kvaček J., Friis E.M. 2021. Plant mesofossils from the Late Cretaceous Klikov Formation, the Czech Republic. Foss. Impr. 77:256–270.

Hermsen E.J., Gandolfo M.A., Nixon K.C., Crepet W.L. 2003. *Divisestylus* gen. nov. (Aff. Iteaceae), a fossil saxifrage from the Late Cretaceous of New Jersey, USA. Am. J. Bot. 90:1373–1388.

Huelsenbeck, J. P. 1991. When are fossils better than extant taxa in phylogenetic analyses?. Syst. Zool. 40(4):458-469

Kearney M., Clark J.M. 2003. Problems due to missing data in phylogenetic analyses including fossils: A critical review. J. Vertebr. Paleontol. 23:263–274.

Keller J.A., Herendeen P.S., Crane P.R. 1996. Fossil flowers and fruits of the Actinidiaceae from the Campanian (Late Cretaceous) of Georgia. Am. J. Bot. 83:528–541.

King B. 2021. Bayesian tip-dated phylogenetics in paleontology: topological effects and stratigraphic fit. Syst. Biol. 70:283–294.

Klopfstein S., Spasojevic T. 2019. Illustrating phylogenetic placement of fossils using RoguePlots: An example from ichneumonid parasitoid wasps (Hymenoptera, Ichneumonidae) and an extensive morphological matrix. bioRxiv.:1–27.

Larson-Johnson K. 2016. Phylogenetic investigation of the complex evolutionary history of dispersal mode and diversification rates across living and fossil Fagales. New Phytol.

209:418–435.

Lee D.E., Conran J.G., Bannister J.M., Kaulfuss U., Mildenhall D.C. 2013. A fossil *fuchsia* (Onagraceae) flower and an anther mass with in situ pollen from the early miocene of New Zealand. Am. J. Bot. 100:2052–2065.

Lee M. S. Y. Yates A. M. 2018. Tip-dating and homoplasy: reconciling the shallow molecular divergences of modern gharials with their long fossil record. Proc. R. Soc. B. 285: 2018107120181071

Lewis P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913–925.

Lloyd G.T., Slater G.J. 2021. A total-group phylogenetic metatree for Cetacea and the importance of fossil data in diversification analyses. Syst. Biol. 70:922–939.

Luo A., Duchêne D.A., Zhang C., Zhu C., Ho S. 2020. A simulation-based evaluation of tip-dating under the fossilized birth-death process. Syst. Biol. 69:325-344.

Magallón-Puebla S., Herendeen P.S., Crane P.R., Sciences P. 1997. *Quadriplatanus georgianus* gen. et sp. nov.: staminate and pistillate platanaceous flowers. Int. J. Plant Sci. 158:373–394.

Magallón S. 2007. From fossils to molecules: Phylogeny and the core eudicot floral groundplan in Hamamelidoideae (Hamamelidaceae, Saxifragales). Syst. Bot. 32:317–347.

Magallón S., Gómez-Acevedo S., Sánchez-Reyes L.L., Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol. 207:437–453.

Magallón S., Herendeen P.S., Crane P.R. 2001. *Androdecidua endressii* gen. et sp. nov., from the Late Cretaceous of Georgia (United States): Further floral diversity in Hamamelidoideae (Hamamelidaceae). Int. J. Plant Sci. 162:963–983.

Manchester S.R. 1992. Flowers, fruits, and pollen of *Florissantia*, an extinct malvalean genus from the Eocene and Oligocene of western North America. Am. J. Bot. 79:996–1008.

Manchester S.R., Dilcher D.L., Judd W.S., Corder B., Basinger J.F. 2018. Early Eudicot flower and fruit: *Dakotanthus* gen. nov. from the Cretaceous Dakota Formation of Kansas and Nebraska, USA. Acta Palaeobot. 58:27–40.

Manos P.S., Soltis P.S., Soltis D.E., Manchester S.R., Oh S.-H., Bell C.D., Dilcher D.L., Stone D.E. 2007. Phylogeny of extant and fossil Juglandaceae inferred from the integration of molecular and morphological data sets. Syst. Biol. 56:412–430.

Marshall C.R. 2017. Five palaeobiological laws needed to understand the evolution of the living biota. Nat. Ecol. Evol. 1:1–6.

Martinez-Millan M., Crepet W.L., Nixon K.C. 2009. *Pentapetalum trifasciculandricus* gen. et sp. nov., a thealean fossil flower from the Raritan formation, New Jersey, USA (Turonian, Late Cretaceous). Am. J. Bot. 96:933–949.

Martínez C., Choo T.Y.S., Allevato D., Nixon K.C., Crepet W.L., Harbert R.S., Daghlian C.P. 2016. *Rariglanda jerseyensis*, a new ericalean fossil flower from the Late Cretaceous of New Jersey. Botany. 94:747–758.

Matsunaga, K. K.S. Smith, S. Y. 2021. Fossil palm reading: using fruits to reveal the deep roots of palm diversity. Am. J. Bot. 108:472:494.

May, M. R., Contreras D. L., Sundue, M.A., Nagalingum, N.S., Looy, C. V., Rothfels, C. J. 2021. Inferring the total-evidence timescale of marattialean fern evolution in the face of model sensitivity. Sys. Biol. 70: 1231-1255.

Mendes M.M., Grimm G.W., Pais J., Friis E.M. 2014. Fossil *Kajanthus lusitanicus* gen. et sp. nov. from Portugal: Floral evidence for Early Cretaceous Lardizabalaceae (Ranunculales, basal eudicot). Grana. 53:283–301.

Meseguer A.S., Lobo J.M., Ree R., Beerling D.J., Sanmartín I. 2015. Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: The case of *Hypericum* (Hypericaceae). Syst. Biol. 64:215–232.

Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gatew. Comput. Environ. Work. GCE 2010.

Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B., Ware J., Flouri T., Beutel R.G., Niehuis O., Petersen M., Izquierdo-Carrasco F., Wappler T., Rust J., Aberer A.J., Aspöck U., Aspöck H., Bartel D., Blanke A., Berger S., Böhm A., Buckley T.R., Calcott B., Chen J., Friedrich F., Fukui M., Fujita M., Greve C., Grobe P., Gu S., Huang Y., Jermiin L.S., Kawahara A.Y., Krogmann L., Kubiak M., Lanfear R., Letsch H., Li Y., Li Z., Li J., Lu H., Machida R., Mashimo Y., Kapli P., McKenna D.D., Meng G., Nakagaki Y., Navarrete-Heredia J.L., Ott M., Ou Y., Pass G., Podsiadlowski L., Pohl H., Von Reumont B.M., Schütte K., Sekiya K., Shimizu S., Slipinski A., Stamatakis A., Song W., Su X., Szucsich N.U., Tan M., Tan X., Tang M., Tang J., Timelthaler G., Tomizuka S., Trautwein M., Tong X., Uchifune T., Walzl M.G., Wiegmann B.M., Wilbrandt J., Wipfler B., Wong T.K.F., Wu Q., Wu G., Xie Y., Yang S., Yang Q., Yeates D.K., Yoshizawa K., Zhang Q., Zhang R., Zhang W., Zhang Y., Zhao J., Zhou C., Zhou L., Ziesmann T., Zou S., Li Y., Xu X., Zhang Y., Yang H., Wang J., Wang J., Kjer K.M., Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science (80-.). 346:763–767.

Mitchell J.S., Etienne R.S., Rabosky D.L. 2019. Inferring diversification rate variation from phylogenies with fossils. Syst. Biol. 68:1–18.

Mongiardino Koch N., Garwood R.J., Parry L.A. 2021. Fossils improve phylogenetic analyses of morphological characters. Proc. R. Soc. B Biol. Sci. 288.

Nylander, J.A., Ronquist F., Huelsenbeck J.P., Nieves-Aldrey J. 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53:47-67.

O'Leary M.A., Bloch J.I., Flynn J.J., Gaudin T.J., Giallombardo A., Giannini N.P., Goldberg S.L., Kraatz B.P., Luo Z.X., Meng J., Ni X., Novacek M.J., Perini F.A., Randall Z.S., Rougier G.W., Sargis E.J., Silcox M.T., Simmons N.B., Spaulding M., Velazco P.M., Weksler M., Wible J.R., Cirranello A.L. 2013. The placental mammal ancestor and the Post-K-Pg radiation of placentals. Science. 339:662–667.

Prendini L. 2001. Species or supraspecific taxa as terminals in cladistic analysis? groundplans versus exemplars revisited. Syst. Biol. 50: 290-300.

R Core Team. 2021. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.r-project.org/.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67:901–904.

Ramírez-Barahona S., Sauquet H., Magallón S. 2020. The delayed and geographically heterogeneous diversification of flowering plant families. Nat. Ecol. Evol. 4:1232–1238.

Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Syst. Biol. 61:973–999.

Ronquist F., Lartillot N., Phillips M.J. 2016. Closing the gap between rocks and clocks using total-evidence dating. Philos. Trans. R. Soc. B Biol. Sci. 371.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–42.

Ruhfel B.R., Stevens P.F., Davis C.C. 2013. Combined morphological and molecular phylogeny of the clusioid clade (Malpighiales) and the placement of the ancient rosid macrofossil *Paleoclusia*. Int. J. Plant Sci. 174:910–936.

Sauquet H. 2019. PROTEUS: A database for recording morphological data and calibrations. Version 1.27. .

Sauquet H., Von Balthazar M., Magallón S., Doyle J.A., Endress P.K., Bailes E.J., Barroso De Morais E., Bull-Hereñu K., Carrive L., Chartier M., Chomicki G., Coiro M., Cornette R., El Ottra J.H.L., Epicoco C., Foster C.S.P., Jabbour F., Haevermans A., Haevermans T., Hernández R., Little S.A., Löfstrand S., Luna J.A., Massoni J., Nadot S., Pamperl S., Prieu C., Reyes E., Dos Santos P., Schoonderwoerd K.M., Sontag S., Soulebeau A., Staedler Y., Tschan G.F., Wing-Sze Leung A., Schönenberger J. 2017. The ancestral flower of angiosperms and its early diversification. Nat. Commun. 8.

Sauquet H., Magallón S. 2018. Key questions and challenges in angiosperm macroevolution. New Phytol. 219:1170–1187.

Schliep K.P. 2011. Phangorn: Phylogenetic analysis in R. Bioinformatics. 27:592–593.

Schönenberger J. 2005. Rise from the ashes - The reconstruction of charcoal fossil flowers. Trends Plant Sci. 10:436–443.

Schönenberger J., von Balthazar M., López Martínez A., Albert B., Prieu C., Magallón S., Sauquet H. 2020. Phylogenetic analysis of fossil flowers using an angiosperm-wide data set: proof-of-concept and challenges ahead. Am. J. Bot. 107:1433–1448.

Schönenberger J., Friis E.M. 2001. Fossil flowers of ericalean affinity from the Late Cretaceous of Southern Sweden. Am. J. Bot. 88:467–480.

Schönenberger J., Friis E.M., Matthews M.L., Endress P.K. 2001. Cunoniaceae in the Cretaceous of Europe: Evidence from fossil flowers. Ann. Bot. 88:423–437.

Silvestro D., Cascales-Miñana B., Bacon C.D., Antonelli A. 2015. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. New Phytol. 207:425–436.

Silvestro D., Salamin N., Schnitzler J. 2014. PyRate: A new program to estimate speciation and extinction rates from incomplete fossil data. Methods Ecol. Evol. 5:1126–1131.

Slater G.J., Harmon L.J., Alfaro M.E. 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. Evolution (N. Y). 66:3931–3944.

Smith S.Y., Stockey R.A. 2007. Establishing a fossil record for the perianthless Piperales: *Saururus tuckerae* sp. nov. (Saururaceae) from the Middle Eocene Princeton Chert. Am. J. Bot. 94:1642–1657.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.

Stevens P.F. 2001. Angiosperm Phylogeny Website. Available from http://www.mobot.org/MOBOT/research/APweb/.

Takahashi M., Friis E.M., Herendeen P.S., Crane P.R. 2008. Fossil flowers of Fagales from the Kamikitaba locality (early Coniacian; Late Cretaceous) of northeastern Japan. Int. J. Plant Sci. 169:899–907.

Takahashi M., Herendeen P.S., Crane P.R. 2001. Lauraceous fossil flower from the Kamikitaba locality (Lower Coniacian; Upper Cretaceous) in northeastern Japan. J.

Plant Res. 114:429–434.

Takahashi M., Herendeen P.S., Xiao X. 2017. Two early eudicot fossil flowers from the Kamikitaba assemblage (Coniacian, Late Cretaceous) in northeastern Japan. J. Plant Res. 130:809–826.

Tank K.K., Smith S., Atkinson B.A. 2022. Extending beyong Gondwana: Cretaceous Cunoniaceae from western North America. New Phytol. 243:704-718.

Van Dongen S. 2006. Prior specification in Bayesian statistics: three cautionary tales. J. Theor. Biol. 242:90-100.

von Balthazar M., Pedersen K.R., Crane P.R., Friis E.M. 2008. *Carpestella lacunata* gen. et sp. nov., a new basal angiosperm flower from the Early Cretaceous (Early to Middle Albian) of eastern North America. Int. J. Plant Sci. 169:890–898.

von Balthazar M., Pedersen K.R., Crane P.R., Stampanoni M., Friis E.M. 2007. *Potomacanthus lobatus* gen. et sp. nov., a new flower of probable Lauraceae from the Early Cretaceous (Early to Middle Albian) of eastern North America. Am. J. Bot. 94:2041–2053.

von Balthazar M., Pedersen K.R., Friis E.M. 2005. *Teixeiria lusitanica*, a new fossil flower from the Early Cretaceous of Portugal with affinities to Ranunculales. Plant Syst. Evol. 255:55–75

Wiens J.J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52:528–538.

Wiens J.J., Morrill M.C. 2011. Missing data in phylogenetic analysis: Reconciling results from simulations and empirical data. Syst. Biol. 60:719–731.

Xing Y., Gandolfo M.A., Onstein R.E., Cantrill D.J., Jacobs B.F., Jordan G.J., Lee D.E., Popova S., Srivastava R., Su T., Vikulin S. V., Yabe A., Linder H.P. 2016. Testing the biases in the rich Cenozoic angiosperm macrofossil record. Int. J. Plant Sci. 177:371–388.

Zhang C., Stadler T., Klopfstein S., Heath T.A., Ronquist F. 2016. Total-evidence dating under the Fossilized Birth-Death Process. Syst. Biol. 65:228–249.

Zhang Q., Ree R.H., Salamin N., Xing Y., Silvestro D. 2021. Fossil-informed models reveal a boreotropical origin and divergent evolutionary trajectories in the walnut family (Juglandaceae). Syst. Biol. 0:1–17.

Material Suplementario

El material suplementario como parte de los contenidos de la publicación del *Capítulo I* se encuentra disponible de manera online en el siguiente link: https://doi.org/10.5061/dryad.cfxpnvx73.

CAPÍTULO II

Angiosperm flowers reached their highest morphological diversity early in their evolutionary history

Artículo publicado: López-Martínez AM, Magallón S, von Balthazar, M Schönenberger J, Sauquet H, Chartier M. 2023. Angiosperm flowers reached their highest morphological diversity early in their evolutionary history. New Phytologist 241: 1348–1360.

RESUMEN – Las flores son las complejas y muy diversas estructuras reproductivas de las angiospermas. Debido a su papel en la reproducción sexual, la evolución de las flores está estrechamente ligada con la especiación y diversificación. En consecuencia, la cuantificación de la diversidad (disparidad) morfológica floral entre los subgrupos de angiospermas y a lo largo del tiempo puede brindar información importante sobre la historia evolutiva de las angiospermas en su conjunto. Basándonos en una extensa base de datos que se centra en 30 caracteres que describen la estructura floral de las angiospermas, utilizamos 1201 flores existentes y 121 fósiles para medir la disparidad floral y para explorar patrones de evolución floral a través del tiempo y entre linajes. Descubrimos que las angiospermas alcanzaron su mayor disparidad floral en el Cretácico temprano. Sin embargo, la disminución de la disparidad hacia el presente probablemente no ha impedido la innovación de otros rasgos complejos a otros niveles estructurales de la flor, que probablemente desempeñaron un papel clave en la excepcional riqueza de especies de angiospermas. Las angiospermas ocupan regiones específicas del morfoespacio teórico, lo que indica que sólo una parte de las posibles combinaciones de rasgos florales se observa en la naturaleza. El grado ANA, las magnoliids y el grado de eudicotiledóneas tempranamente divergentes, ocupan grandes áreas del morfoespacio (mayor disparidad), mientras que los grupos específicos dentro de grandes clados ocupan regiones más estrechas (menor disparidad).

Palabras clave: Cretácico, disparidad a través del tiempo, evolución floral, morfoespacio categórico, Paleógeno, reconstrucción de estados ancestrales.

New Phytologist

Research

Angiosperm flowers reached their highest morphological diversity early in their evolutionary history

Andrea M. López-Martínez[1,2], Susana Magallón[2], Maria von Balthazar[3], Jürg Schönenberger[3], Hervé Sauquet[4,5] and Marion Chartier[3]

[1]Posgrado en Ciencias Biológicas, Instituto de Biología, Universidad Nacional Autónoma de México, 3er Circuito de Ciudad Universitaria, Coyoacán, Ciudad de México, 04510, Mexico; [2]Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, 3er Circuito de Ciudad Universitaria, Coyoacán, Ciudad de México, 04510, Mexico; [3]Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, Vienna, A-1030, Austria; [4]National Herbarium of New South Wales (NSW), Royal Botanic Gardens and Domain Trust, Sydney, NSW, 2000, Australia; [5]Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Biological Sciences North (D26), Sydney, NSW, 2052, Australia

Authors for correspondence:
Andrea M. López-Martínez
Email: andreaea.lopez@gmail.com

Marion Chartier
Email: marion.chartier@univie.ac.at

Received: 28 August 2023
Accepted: 20 October 2023

New Phytologist (2023)
doi: 10.1111/nph.19389

Key words: ancestral state reconstruction (ASR), categorical morphospace, Cretaceous, disparity through time, floral evolution, Paleogene.

Summary

- Flowers are the complex and highly diverse reproductive structures of angiosperms. Because of their role in sexual reproduction, the evolution of flowers is tightly linked to angiosperm speciation and diversification. Accordingly, the quantification of floral morphological diversity (disparity) among angiosperm subgroups and through time may give important insights into the evolutionary history of angiosperms as a whole.
- Based on a comprehensive dataset focusing on 30 characters describing floral structure across angiosperms, we used 1201 extant and 121 fossil flowers to measure floral disparity and explore patterns of floral evolution through time and across lineages.
- We found that angiosperms reached their highest floral disparity in the Early Cretaceous. However, decreasing disparity toward the present likely has not precluded the innovation of other complex traits at other morphological levels, which likely played a key role in the outstanding angiosperm species richness.
- Angiosperms occupy specific regions of the theoretical morphospace, indicating that only a portion of the possible floral trait combinations is observed in nature. The ANA grade, the magnoliids, and the early-eudicot grade occupy large areas of the morphospace (higher disparity), whereas nested groups occupy narrower regions (lower disparity).

Introduction

Disparity, or morphological diversity, is a quantification of the morphological variation displayed by a group of organisms (Foote, 1992a, 1993a, 1995, 1997; Wills et al., 1994; Ciampaglio et al., 2001; Erwin, 2007; Guillerme et al., 2020; Hopkins & Gerber, 2021). Measures of disparity are based on the multivariate description of morphological traits and are often coupled with the ordination of morphospaces that describe and relate the phenotypic configurations of organisms (Foote, 1994; Mitteroecker & Huttegger, 2009; Smith & Donoghue, 2022). The variation of disparity and morphospace occupation within and among clades, populations, fossil assemblages, or time periods, may provide useful information on a lineage's morphological evolution (Foote, 1994; Benton, 2015), on the evolutionary constraints shaping phenotypes (Ciampaglio, 2002; Allen et al., 2008), on the ecology and natural selection exerted on populations (Benitez-Vieyra et al., 2010), and on the effect of mass extinction on lineages (Brusatte et al., 2008a; Friedman, 2010; Korn et al., 2013; Puttick et al., 2020). In particular, studying disparity through time (DTT) using fossils gives insight into the evolutionary dynamics of lineages and the general patterns of morphological evolution on Earth (McGhee, 1991; Foote, 1993b; Hughes et al., 2013; Oyston et al., 2016).

Disparity has generally evolved nonuniformly and independently from species richness throughout the history of lineages (Foote, 1994, 1995; Slater et al., 2010; Benton et al., 2014; Moon & Stubbs, 2020; Coombs et al., 2022). Two main trends in the variation of DTT have been observed with respect to when maximal disparity was reached. Most lineages reached maximal disparity early in their evolution ('early-disparity', 'bottom-heavy' distribution: Gould et al., 1987; Gould, 1989; Hughes et al., 2013). These early-disparity trends result from abrupt and wide morphospace exploration possibly due to the rapid colonization and exploitation of new environments, or of an ecological niche that was left vacant after an extinction event (Foote, 1994; Benton et al., 2014; Benton, 2015; Leslie et al., 2021). Alternatively, new morphologies might accumulate slowly as clades diversify so that disparity might reach its maximal levels late in a clade's evolution ('late-disparity', or 'top-heavy' distribution;

Gould *et al.*, 1987; Prentice *et al.*, 2011; Hughes *et al.*, 2013; Puttick *et al.*, 2020). Although disparity variation and morphospace occupation through time have been described for various animal clades (e.g. Foote, 1991 for blastoids; Foote, 1994, 1999 for crinoids; Fortey *et al.*, 1996 for several animal groups; Brusatte *et al.*, 2008b for dinosaurs; Bapst *et al.*, 2012 for graptoloids; Hill *et al.*, 2018 for fishes), only few studies have quantified DTT in angiosperms (e.g. Lupia, 1999; Jardine *et al.*, 2022 for pollen; Oyston *et al.*, 2016; Clark *et al.*, 2023 for general morphology; Martínez-Cabrera *et al.*, 2017 for wood properties).

Flowers are a relatively recent evolutionary innovation and yet they have evolved remarkable morphological and functional diversity (Endress, 2011; Sauquet *et al.*, 2017, 2022). Because floral morphology is directly linked to angiosperm reproductive success, flowers are under strong selective pressures and studying their morphological diversity is crucial to understand trends of lineage divergence and evolutionary success (Harder & Barrett, 2006; Endress, 2011; van der Niet & Johnson, 2012; Chartier *et al.*, 2014; Sauquet & Magallón, 2018). The diversity of flowers has been extensively discussed qualitatively, based on the development and comparative morphology of various extant groups (Endress & Igersheim, 1997, 1999, 2000; Matthews & Endress, 2002, 2005, 2006; Schönenberger & von Balthazar, 2006; Endress, 2010, 2011; Schönenberger *et al.*, 2010). Based on these studies, it is for example widely recognized that floral organization (Bauplan) is relatively labile among the ANA grade lineages, magnoliids, and early-diverging eudicots, all of which exhibit, for instance, high variability in merism (organ number) and phyllotaxis (the arrangement of floral organs). By contrast, floral organization is more stable in large clades such as the monocots with almost exclusively trimerous, whorled flowers, or in Pentapetalae with mostly pentamerous, whorled flowers (Endress, 2010, 2011; Sauquet *et al.*, 2017).

The first study using a morphospace approach to analyze floral morphology across angiosperms is the theoretical morphospace created by Stebbins (1951). A theoretical morphospace allows describing the entire spectrum of theoretically possible morphologies for a group of organisms and, therefore, determining limits in the evolution of biological forms and identifying areas of possible high vs low fitness (Raup & Michelson, 1965; Raup, 1967; McGhee, 1991, 2015; Avena-Koenigsberger *et al.*, 2015). Using this approach, Stebbins determined a set of floral trait combinations that were unlikely to occur in angiosperms and others that were particularly successful (common; Stebbins, 1951). A later quantitative re-analysis of Stebbins' dataset confirmed the qualitative observations of plant morphologists that the highest floral disparity is present in the ANA grade, magnoliids, and early-diverging eudicots, whereas the lowest disparity was found in the nested clades such as lamiids, campanulids, and malvids (Chartier *et al.*, 2014). The shortcoming of these two studies is that Stebbins' original dataset only included eight binary floral characters and was coded at the family level based on a now largely outdated classification of angiosperms. The study of floral evolution and diversity during the last decades and the reconstruction of ancestral floral morphologies (e.g. Sauquet *et al.*, 2017) allow us to study disparity and morphological trajectories using unprecedently large datasets through time at the level of angiosperms as a whole. In addition, the recent description of numerous fossil flowers provides information about past floral diversification (e.g. Crepet *et al.*, 1991; Endress, 2006; Endress & Doyle, 2009; Friis *et al.*, 2010, 2011; Doyle & Endress, 2014) and makes it possible to study DTT in angiosperms. Most fossil flowers stem from Cretaceous sediments and include predominantly three-dimensionally, well-preserved, charcoalified mesofossils (e.g. Schönenberger & Friis, 2001; Gandolfo *et al.*, 2004; Friis & Pedersen, 2011, 2012) while a few are preserved as permineralizations (e.g. Atkinson *et al.*, 2015), impressions/compressions (e.g. Dilcher & Crane, 1984; Mohr & Eklund, 2004), and amber inclusions (e.g. Gandolfo *et al.*, 2018).

In this study, we use an extensive morphological dataset describing characters at the level of floral organization (Bauplan; *sensu* Endress, 1994). We include most of the best-preserved fossil flowers known and a broad representation of living species to investigate the large-scale temporal and phylogenetic patterns in the evolution of angiosperm flowers. First, we quantify and describe the changes in disparity and morphospace occupation through geologic time. Second, we explore the position of living species (split into 11 taxonomic groups) in a theoretical floral morphospace, and test for the correlation between disparity, clade age, and clade species richness in these groups.

Materials and Methods

All analyses were performed in R v.4.1.2 (R Core Team, 2022; Supporting Information Dataset S1). In the following, function names and packages are referred to as *function name* {PACKAGE NAME}.

Analyses overview

We performed three different analyses. For each analysis, a distance matrix (see Distance matrices in the Materials and Methods section) containing the taxa of interest (see Morphological dataset in the Materials and Methods section) was computed and then used to calculate disparity (see Estimating disparity in the Materials and Methods section), then an ordination was generated to visualize the morphospace (see Ordination in the Materials and Methods section).

Disparity through time We compared disparity and morphospace occupation among four stratigraphic time bins: the Early Cretaceous (145–100.5 Ma), the Late Cretaceous (100.5–66 Ma), the Paleogene (66–23.03 Ma; Cohen *et al.*, 2013), and living species (present time). Detailed morphological descriptions were only available for two species from the Neogene (23.03–5.333 Ma; Castañeda-Posadas & Cevallos-Ferriz, 2007; Hernández-Damián *et al.*, 2018; Cohen *et al.*, 2013); Neogene fossils were thus excluded from the statistical analyses.

Angiosperm trajectory in the morphospace We placed ancestral flowers reconstructed for the crown nodes of 15 key angiosperm clades (see Ancestral state reconstructions in the Materials and

Methods section) in the floral morphospace and compared their position with that of living and fossil angiosperms.

Disparity among living angiosperms We compared disparity and morphospace occupation among living angiosperms split into 11 groups (including clades and grades) and tested for the correlation among disparity, age, and species richness for these groups. We split species into the following six clades: magnoliids, commelinids, fabids, malvids, campanulids, and lamiids (*sensu* APG IV, 2016), and five grades: ANA (Amborellales, Nymphaeales, and Austrobaileyales), other monocots (all monocot lineages except commelinids), other eudicots (Ranunculales, Proteales, Trochodendrales, and Buxales), other superrosids (Saxifragales and Vitales), and other superasterids (Santalales, Berberopsidales, Caryophyllales, Cornales, and Ericales; Dataset S2). The orders Chloranthales, Ceratophyllales, Gunnerales, and Dilleniales were excluded from the statistical analyses as they contained only two to four species in our sampling and because they cannot be assigned to any previously specified group.

Morphological dataset

The dataset includes 30 categorical floral characters scored for 1201 living angiosperm species (36 030 matrix cells) sampled from all currently accepted families from APG IV (2016), as well as 121 fossil flowers (3630 matrix cells), which represented most of the best-preserved fossil flowers described to date. The final data matrix used in analyses and the complete list of fossils and information on their stratigraphic age, mode of preservation, locality, and corresponding references are available in Dataset S2.

Morphological data and references were recorded in the PROTEUS database (Sauquet, 2019). One character describes the whole flower, eight describe the perianth, 11 the androecium, eight the gynoecium, and two the pollen (Sauquet *et al.*, 2017; Schönenberger *et al.*, 2020). Data for the 1201 living species were previously scored by López-Martínez *et al.* (2023). Data for 24 of the 121 fossil species were obtained from Schönenberger *et al.* (2020) and López-Martínez *et al.* (2023), and 97 fossils were additionally scored for this study. A complete extraction of PROTEUS data for the fossil dataset (2442 data records linked to 125 explicit references) is provided as Dataset S3 and the dataset used here is provided as Dataset S2.

The proportion of missing data, including inapplicable characters, is 32% (11 573 matrix cells) for the living species (Fig. S1) and 36% (1305 matrix cells) for the fossils (Fig. S2). Additionally, there were 3% (1310 matrix cells) polymorphic entries for the living species and 2% (74 matrix cells) for the fossils. As most of our analyses do not support polymorphic data, we randomly sampled one of the polymorphic states (with equal probabilities) for each polymorphic entry before each analysis following Chartier *et al.* (2017). Our results did not change among these analyses.

Theoretical combinations

To visualize the distribution of achieved morphologies among possible ones in the ordinations, we added a background of theoretical combinations to the combinations displayed by the sampled angiosperm species (empirical data). Such an approach is helpful for categorical data like ours, where the nature of the space is discrete rather than continuous (Gerber, 2019). The theoretical morphospace contains all combinations of traits that can be obtained from the study character set. In our data, the total number of possible combinations is the product of the number of states for each character: $2^{17} \times 3^8 \times 4^3 \times 5^2 = 1.38 \times 10^{12}$. To display a subset of the theoretical morphospace in the ordinations, we randomly sampled 2000 of these theoretical combinations without replacement each time we created an ordination. We did not sample theoretical combinations containing inapplicable or impossible combinations of character states (e.g. perianth merism in flowers without perianth); in these cases, the inapplicable character state was treated as missing data. Inapplicable and impossible combinations are listed in Methods S1.

Distance matrices

For each analysis, a distance matrix was computed by calculating the *mean character difference* (here noted *D*) for each pair of taxa following Sneath & Sokal (1973 p. 135) and Foote (1999). The mean character difference is a version of the Gower index, suited for datasets like ours containing categorical ordered and categorical unordered characters. It ranges from zero (no difference) to 1 (largest difference in the dataset). Its calculation is detailed in Methods S2.

Estimating disparity

Disparity was estimated for each group (clade, grade, or pool of species belonging to the same time bin) based on four metrics capturing different aspects of morphospace occupation:

The mean pairwise difference (here *meanD*) reflects the density of taxa within a group; it is the *mean character difference* (*D*, stored in the distance matrix) averaged among each pair of taxa (Foote, 1999). This metric is moderately sensitive to sample size (Ciampaglio *et al.*, 2001), and since the data present large differences in sample size in some instances (e.g. 22 species in the Paleogene vs 1201 extant angiosperms), we additionally rarefied it. For the rarefaction analysis, *n* species were randomly sampled without replacement in each group (with *n* the size of the smallest group minus one); then, the distance matrix was computed and *meanD* calculated. This was repeated 1000 times.

The range (here noted *R*) gives information about the size of a group in the morphospace (Foote, 1992b); we calculated it as the maximum pairwise distance (*maxD*) in a group. The range is sensitive to sample size (Ciampaglio *et al.*, 2001), and we thus rarefied it as described above (we only compared rarefied values).

MeanD, rarefied *meanD*, and rarefied *R* were compared among groups with Kruskal–Wallis tests for nonparametric data (Kruskal & Wallis, 1952) using the functions *kruskal* {AGRICOLAE} (de Mendiburu & Yaseen, 2020) and *kruskalmc* {PGIRMESS} (Giraudoux *et al.*, 2018).

We also computed a measure of the contribution of a group to total disparity (here noted *Ddelta*; Foote, 1993a). We measured

Ddelta for group *i* ($Ddelta_i$), as the difference between *meanD* for the dataset without group *i*, and *meanD* for the total dataset (*Dtot*). When using *meanD* as an estimate of disparity, a positive value of $Ddelta_i$ indicates that adding group *i* contributed to an increase in *Dtot*, while a negative value of $Ddelta_i$ indicates that adding group *i* contributed to a decrease in *Dtot*. The magnitude of this contribution is proportional to the absolute value of $Ddelta_i$.

Finally, to detect the living species with the most distinct floral combinations, we estimated the divergence of each species from the average morphology (we called it eccentricity following Oyston *et al.*, 2015) by averaging *D* for each living species in the dataset. This was done by averaging each line of the distance matrix. Species with the highest eccentricity are placed at the edge of the occupied morphospace ordination.

Ordination

We visualized morphospaces with nonmetric multidimensional scaling (nMDS) using the function *metaMDS* {VEGAN} (Oksanen *et al.*, 2020), setting the number of dimensions (*k*) to two and the maximum number of random starts (trymax) to 20 (default value). For each group (time bin or taxonomic group), the position of the centroid was added to the graphs by averaging the coordinates of each point for each axis. Floral morphologies were compared among groups with a nonparametric multivariate analysis of variance (npMANOVA) with the function *adonis2* {VEGAN} (Oksanen *et al.*, 2020). Post hoc tests consisted of pairwise npMANOVAs with a Bonferroni correction.

Correlations

For living species, we investigated the correlations between disparity (*meanD*) and species richness of clades/grades (extracted from the World Flora Online; WFO, 2023), between *meanD* and clade age (crown age for six clades from Ramírez-Barahona *et al.*, 2020), and between *meanD* and *Ddelta*. Finally, we verified that sample size was positively correlated with species richness. We used Pearson correlation tests using the function *cor.test* {VEGAN} (Oksanen *et al.*, 2020).

Ancestral state reconstructions

We employed maximum likelihood (ML) and stochastic character mapping (SM) methods to reconstruct ancestral states for each floral trait at 15 angiosperm key nodes corresponding to broadly recognized major angiosperm clades (*sensu* Cantino *et al.*, 2007; APG IV, 2016). All ancestral state reconstructions (ASR) were conducted using the maximum clade credibility dated tree from the relaxed calibration complete (RC-complete) strategy from Ramírez-Barahona *et al.* (2020). Maximum likelihood reconstructions were performed implementing the equal rates (ER) and all rates different (ARD) models of discrete morphological evolution with the function *rayDISC* {CORHMM} (Beaulieu *et al.*, 2013; Dataset S4). We subsequently compared model fit based on the Akaike information criterion. SM reconstructions were then conducted using the best-fitted model obtained with ML. For SM analyses, we first dropped the tips with missing data, nonapplicable, and polymorphic data from the tree using the function *drop.tip* {PHYTOOLS} (Revell, 2012). We then inferred ancestral states using 500 simulations along the tree with the function *make.simmap* {PHYTOOLS} (Revell, 2012) and summarized with the function *describe.simmap* {PHYTOOLS} (Revell, 2012). Dropped tips were removed from the trees showing ASR with SM (Dataset S5). To display ASR in the morphospace, we included the most probable combinations of ancestral states for each node in the dataset before computing the distance matrix, according to Gerber (2019).

Results

Morphospace and disparity through time

Fossil floral disparity significantly differed among the four time-bins (Kruskal–Wallis test: $\chi^2 = 207.29$, df = 3, $P = < 2.2e^{-16}$; Fig. 1b). Angiosperm flowers showed the highest disparity in the Early Cretaceous (*meanD* = 0.372 ± SD 0.14) followed by a decrease toward the Late Cretaceous (*meanD* = 0.316 ± 0.143) and the Paleogene (*meanD* = 0.241 ± 0.13). Disparity for living species reached values intermediate between Late Cretaceous and Paleogene (*meanD* = 0.29 ± 0.137; Fig. 1b). This pattern was confirmed by the rarefaction analyses of *meanD* (Fig. S3a) and of *R*, with the exception of the Paleogene group showing a range as high as in the Early Cretaceous (Fig. S3b). Note that this pattern applies to angiosperms as a whole and that disparity within subgroups might vary differently (Fig. S3c–e).

The morphospace area occupied by fossil floral morphologies significantly varied in time (npMANOVA: $F = 21.583$ $P < 1e^{-4}$; Fig. 1a), with the position of the centroid showing a displacement from the middle bottom part to the right middle side of the morphospace (Figs 1a, 2). Although all reconstructed ancestors analyzed here dated from the Jurassic to the Early Cretaceous (Table S1), they followed the same trajectory as the position of the centroid for fossils through time and were mostly restricted to the area densely occupied by living species (Fig. 2).

Morphospace and disparity for living angiosperms

Ninety-five percent of living species occupied a restricted area of the morphospace ordination (as a rough indication, the corresponding convex hull covered 36% of the total morphospace; Fig. S6) and the remaining 5% of the living species with the highest eccentricity were distributed in otherwise largely empty areas of the theoretical morphospace (the corresponding convex hull covered 60% of the total morphospace; Fig. S6).

In our dataset, 45% of living species had bisexual flowers with one whorl of sepals and one whorl of petals, a total of 6–10 perianth parts, together with a differentiated style (Fig. S7). By contrast, the floral traits characterizing the empty area of the theoretical morphospace (traits more rarely observed in living species) included, in combination or not: dimerous perianths and androecia with more than two whorls each, diaperturate and

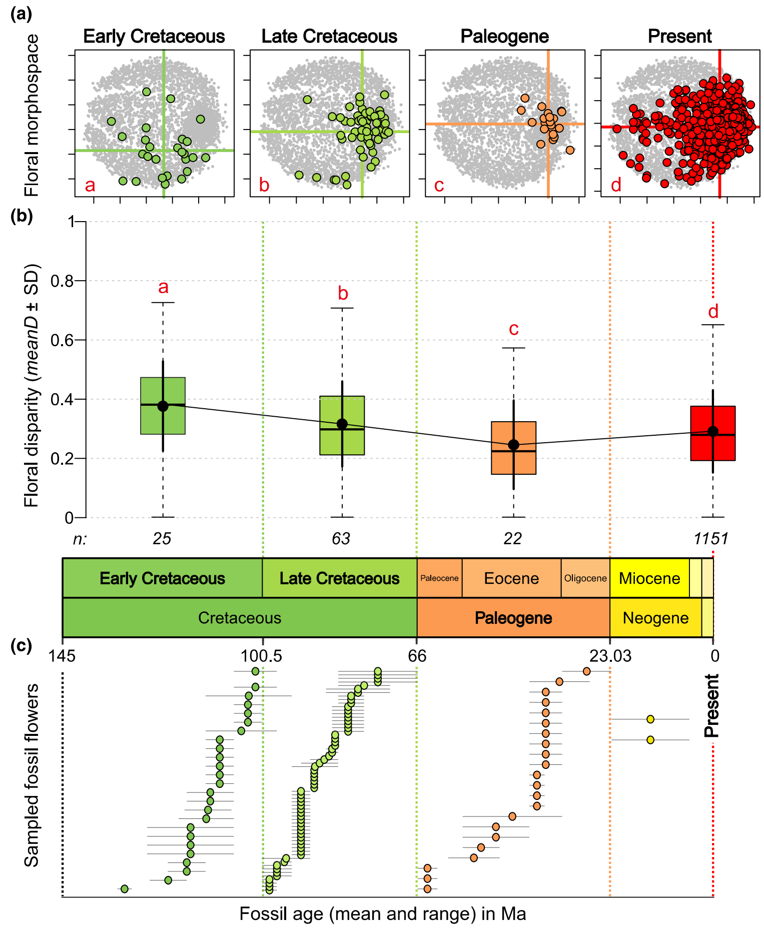

Fig. 1 Evolution of the floral morphospace and floral disparity through time for angiosperms. (a) Evolution of morphospace occupation through time (nMDS stress value = 0.34; Shepard plot: nonmetric fit $r^2 = 0.88$, linear fit $r^2 = 0.39$). Colored lines indicate the position of the centroid for each group. In (b), black dots, disparity (*meanD*) for each time bin. SD, standard deviation (black error bars). Boxplots, distribution of *D* for each group. *n*, sample sizes after computing the distance matrix. In (a) and (b), red letters indicate *post hoc* test results, groups with a different letter significantly differ from each other. (c) Stratigraphic age ranges for each fossil flower. The position of Neogene fossils in the morphospace is presented in Supporting Information Fig. S4.

inaperturate pollen grains, the absence of a perianth, spiral perianth and androecium phyllotaxis, flap-valvate and H-valvate anther dehiscence, fused ovaries, undifferentiated styles, as well as free-central and laminar placentation (Fig. S7). The five species with the highest eccentricity were as follows: *Eupomatia bennetii* F.Muell. and *E. laurina* R.Br. (magnoliids), *Cyclanthus bipartitus* Poit. ex A.Rich. (other monocots grade), *Galbulimima belgraveana* (F.Muell.) Sprague (magnoliids), and *Sarcandra chloranthoides* Gardner (Chloranthales; Fig. S6).

Disparity differed significantly among angiosperm groups (Kruskal–Wallis test: $\chi^2 = 18\,392$, df = 10, $P = < 2.2e^{-16}$; Fig. 3b; Table S2). Floral disparity was highest within magnoliids (*meanD* = 0.378 ± SD 0.16), the ANA grade (*meanD* = 0.370 ± 0.14), and the other eudicots grade (*meanD* = 0.334 ± 0.16). These three groups also contributed the most to total disparity (Fig. 4c). The lowest disparity was found in nested clades within the Pentapetalae clade, such as campanulids (*meanD* = 0.158 ± 0.09) and lamiids (*meanD* = 0.12 ± 0.08; Fig. 3b). These two groups contributed the least to total disparity (Fig. 4c). Rarefied *meanD* and rarefied *R* showed similar results (Fig. S8).

The position of living clades and grades in the space varied significantly (npMANOVA: $F = 41.78$, $r^2 = 0.23$, $P < 1e^{-04}$; Fig. 3a), meaning that each group presented specific combinations of floral traits. Exceptions (nonsignificantly different groups) were as follows: (1) the ANA grade, magnoliids, and the other eudicots grade, (2) commelinids and the other monocots grade, (3) fabids, the other superasterids grade, and the other superrosids grade, (4) the other monocots grade and the other superrosids grade (Fig. 3a).

Correlations

Disparity significantly decreased with species richness (Pearson: $t = -3.01$, df = 9, $P = 0.015$; Fig. 4a), with the exception of fabids, which showed medium disparity (*meanD* = 0.29) and a large number of species (85621). Disparity significantly increased with crown node age ($t = 5.80$, df = 4, $P = 0.004$; Fig. 4b) and

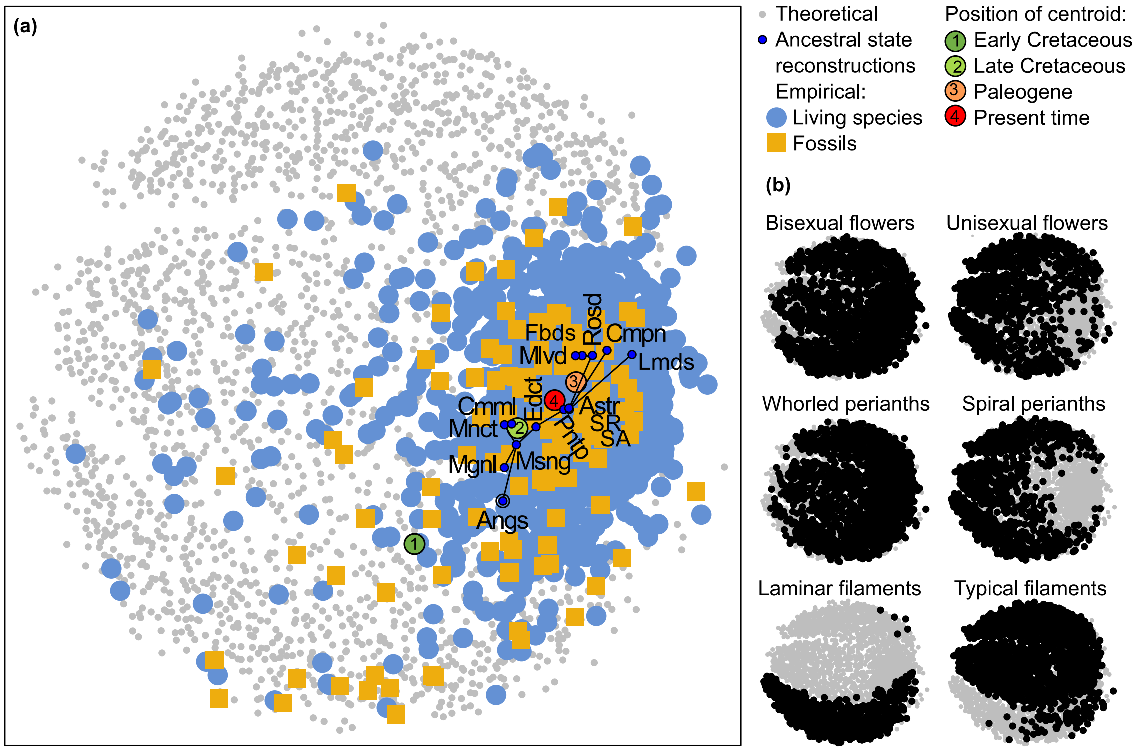

Fig. 2 Total floral morphospace of angiosperms (nMDS stress value = 0.34; Shepard plot: nonmetric fit $r^2 = 0.88$, linear fit $r^2 = 0.39$), including 1151 living species (light blue dots, *meanD* = 0.29, *R* = 1), 113 fossils (yellow squares, *meanD* = 0.34, *R* = 1), 15 ancestral state reconstructions (ASR; stochastic character mapping (SM) analyses, navy blue dots, *meanD* = 0.16, *R* = 0.35) and 2000 theoretical combinations (gray dots), displayed in (a). Angs, angiosperms; Astr, asterids; Cmml, commelinids; Cmpn, campanulids; Edct, eudicots; Fabd, fabids; Lamd, lamiids; Mgnl, magnoliids; Mlvd, malvids; Mnct, monocots; Msng, mesangiosperms; Pntp, Pentapetalae; Rosd, rosids; SA, superasterids; SR, superrosids. Links among ancestors illustrate phylogenetic relationships. In (b), distribution of some selected character states (black dots) in the same space (gray dots; see also Supporting Information Fig. S5). The values for *meanD* and *R* given here are not rarefied. Note the round shape of the theoretical space ordination, similar to the ordinations of other theoretical spaces of categorical nature (e.g. Gerber, 2019).

with the contribution to total disparity ($t = 4.84$, df = 9, $P < 1e^{-3}$; Fig. 4c). Sample size was positively correlated with species richness ($t = 4.69$, df = 9, $P = 0.001$; Fig. S10).

Discussion

Floral disparity through time

Early high levels of disparity are generally characterized by high rates of character change during the early phases of a lineage's history, resulting in maximum phenotypic variation followed by developmental canalization (Foote, 1997, 1999; Wagner, 2018; Smith & Donoghue, 2022). As we show here, angiosperm flowers also followed this pattern, reaching their maximal disparity in the Early Cretaceous, meaning that major structural variations (Bauplan, e.g. fusion, number, and arrangement of organs) were explored early in the history of the group. Interestingly, this occurred when angiosperm species richness was still low compared with other land plants (Lupia *et al.*, 1999) and the group was still mainly restricted to the paleotropics (Willis & McElwain, 2002). It remains uncertain whether the initial floral morphological diversification was associated with the exploitation of ecological opportunities, for example by the interaction with dispersers and pollinators. Insect pollination, for instance, existed before the origin of angiosperms (Peñalver *et al.*, 2015; Asar *et al.*, 2022; Peña-Kairath *et al.*, 2023) and is ancestral for the group as a whole (Friis *et al.*, 1999; Stephens *et al.*, 2023).

Disparity decreased in the Late Cretaceous as the average floral morphology shifted to become more similar to that of extant species (Fig. S11). For example, the higher frequency of undifferentiated free perianths, spiral androecia, basifixed anthers, and valvate anther dehiscence among Early Cretaceous fossils was, from the Late Cretaceous on, replaced by a majority of differentiated and fused perianths, whorled androecia, dorsifixed anthers, and dehiscence by longitudinal slits. A number of these Late Cretaceous flowers also display new traits considered as a step toward a further adaptation to animal pollination, like perianth differentiation, a lower number of floral organs, nectaries, and fusion of

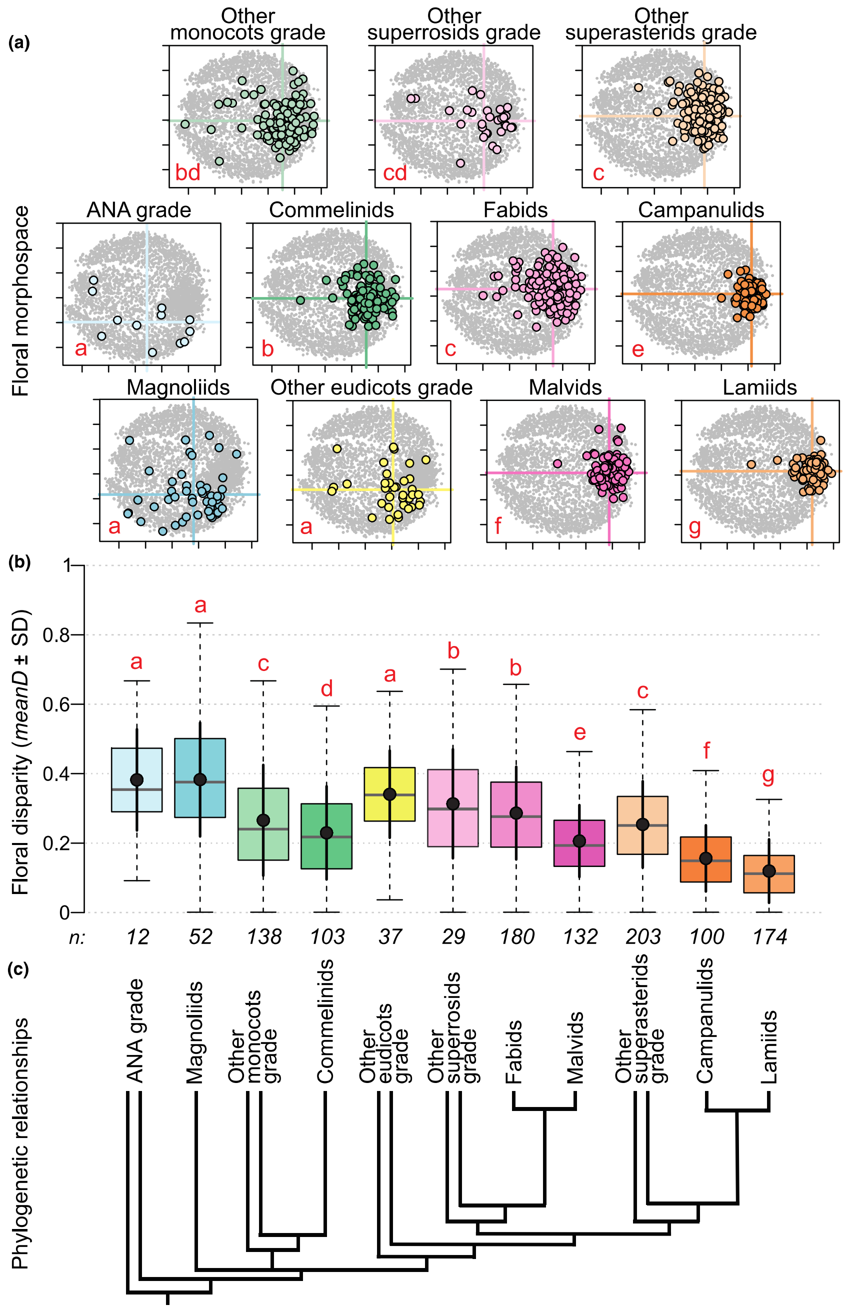

Fig. 3 Floral morphospace and disparity for living angiosperms. (a) Morphospace occupation for 11 angiosperm groups (nMDS stress value = 0.34; Shepard plot: nonmetric fit $r^2 = 0.88$, linear fit $r^2 = 0.38$). In each plot, colored dots = species belonging to the group of interest, gray dots = 2000 theoretical combinations and the remaining angiosperm species; the intersection of the colored lines indicates the position of the centroid of each group. (b) Disparity (*meanD*, black dots) ± standard deviation (black vertical plain lines). Boxplots show the distribution of pairwise distances (*D*) for each group, numbers below each boxplot correspond to sample size after computing the distance matrix. In (a) and (b), red letters indicate *post hoc* test results, groups with a different letter significantly differ from each other. (c) Phylogenetic relationships among major angiosperm lineages following Ramírez-Barahona *et al*. (2020). The position of Chloranthales, Ceratophyllales, Gunnerales, and Dilleniales in this morphospace is presented in Supporting Information Fig. S9.

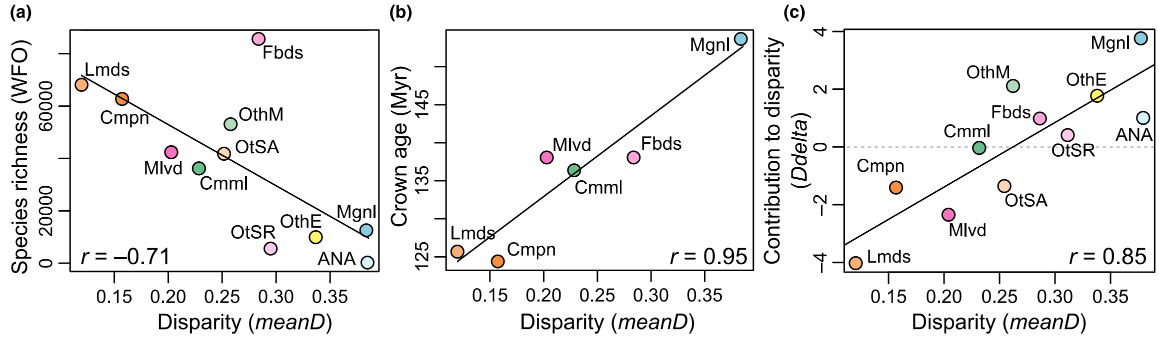

Fig. 4 Correlations between disparity (*meanD*) and (a) species richness, (b) crown node age for six clades, (c) contribution of a group to total disparity (*Ddelta*). Black lines: linear regressions. *r* = Pearson's product-moment correlation coefficient. ANA, ANA grade; Cmml, commelinids; Cmpn, campanulids; Fbds, fabids; Lmds, lamiids; Mgnl, magnoliids; Mlvd, malvids; Myr, million years ago; OthE, the other eudicots grade; OthM, the other monocots grade; OtSA, the other superasterids grade; OtSR, the other superrosids grade; WFO, World Flora Online (WFO, 2023).

carpels (syncarpy; Friis *et al.*, 2006). This timing is consistent with the diversification of various pollinating insect clades (Asar *et al.*, 2022; Benton *et al.*, 2022; Kawahara *et al.*, 2023; Peña-Kairath *et al.*, 2023), the emergence (stem age) of most extant angiosperm families (Ramírez-Barahona *et al.*, 2020), and the progressive dominance of angiosperms in terrestrial ecosystems (Lupia *et al.*, 1999; Willis & McElwain, 2002; Augusto *et al.*, 2014; Condamine *et al.*, 2020).

We also observe a decrease in floral disparity following the Cretaceous/Paleogene (K/Pg) extinction event. Although fossil-based quantitative analyses have suggested that the K/Pg extinction was highly selective to certain clades (e.g. Laurales and Cyclanthaceae; McElwain & Punyasena, 2007) and regions (e.g. tropical South America; Carvalho *et al.*, 2021), there is no consensus about the impact of the K/Pg extinction on angiosperms as a whole (Willis & McElwain, 2002; Cascales-Miñana & Cleal, 2014; Thompson & Ramírez-Barahona., 2023; Wilf *et al.*, 2023), and it remains uncertain whether it had an effect on floral disparity as Paleogene floral fossils are unevenly distributed with regard to phylogeny, geography, and preservation types (Friis *et al.*, 2011; Xing *et al.*, 2016). In our data, the morphospace area covered by these fossils is restricted, and the area emptied during the Paleogene gets occupied again in the Present, confirming a possible sample bias in the Paleogene record. The Paleogene was characterized by profound global landscape changes with the emergence of dense subtropical forests with closed canopies (Carvalho *et al.*, 2021) and the diversification (crown age) of most extant angiosperm families (Ramírez-Barahona *et al.*, 2020). Despite their lower disparity, Paleogene flowers show a higher frequency of bilateral symmetry (zygomorphy) and perianth fusion, which might indicate higher rates of specialization to different functional groups of pollinators (e.g. Stewart *et al.*, 2022).

Interestingly, angiosperm pollen disparity also rapidly increased from the Early Cretaceous (Lupia, 1999) but the highest pollen disparity in the Late Cretaceous was followed by a stabilization of the morphospace due to the accumulation of different morphologies. The same pattern was estimated for Asterales pollen (Jardine *et al.*, 2022). Conversely, Oyston *et al.* (2016) found that the overall disparity of vegetative and reproductive characters remained constant throughout the angiosperm's history. By contrast, a study on wood traits (Martínez-Cabrera *et al.*, 2017) found low initial disparity followed by an increase from the Late Cretaceous on, concomitant with the increased domination of land ecosystems by angiosperms, most probably due to vegetative key innovations leading to more adaptive growth strategies (Feild *et al.*, 2011; Condamine *et al.*, 2020). The different evolutionary trends of disparity variation between flowers, pollen, and vegetative traits probably reflect differences in function, selective value, and evolutionary potential of these traits. Knowing more about these different components would allow getting a more comprehensive understanding of the incredible success of angiosperms throughout their evolutionary history.

Limits in the evolution of floral form

Most extant flowers are restricted to a small area of the morphospace (Fig. S6): 91% of the living species have flowers with a perianth, 85% are bisexual, 83% are whorled, and 72% have a differentiated perianth (Fig. S7). In terms of trait combinations, the most common flower is bisexual, with a differentiated perianth arranged in two whorls (45% of living species in our dataset). These traits characterize *Pentapetalae*, the angiosperm clade containing > 70% of extant species (Magallon *et al.*, 1999; Christenhusz & Byng, 2016). For the flower, the higher occupation of some areas of the morphospace can be attributed to the economy of construction and/or to functional advantages (Stebbins, 1951; Endress, 1982). For example, bisexuality (85% of the living species in our dataset) increases pollination probability (function) since pollinators can deposit and take up pollen in a single flower visit; a differentiated perianth (72% of the living species in our dataset) allows sepals to protect the developing flower while petals attract pollinators at anthesis (function); a perianth with two whorls of organs (71% of the living species in our dataset) that generally alternate allows for optimal organization (economy of construction); finally, the presence of a low number of united carpels (55% of the living species in our dataset) requires a smaller amount of wall tissue during growth (economy of construction) and enables more regular pollen tube distribution among carpels (function; Endress, 1982), increasing the probability for fertilization (Armbruster *et al.*, 2002). Some trait combinations, like the presence of a corolla with zygomorphy and a small number of stamens (20% of the living species in our dataset), increase speciation rates (O'Meara *et al.*, 2016) as they give a selective advantage for pollination (function) by ensuring a more efficient placement of pollen on pollinator's bodies (Walker-Larsen & Harder, 2000; Sargent, 2004).

The distribution of living species in the floral morphospace and the quantification of disparity follow a phylogenetic pattern and support conclusions from qualitative studies (e.g. Endress & Igersheim, 1997, 1999, 2000; Matthews & Endress, 2002, 2005, 2006; Schönenberger & von Balthazar, 2006; Endress, 2010, 2011). These results are consistent with the re-analysis of Stebbin's dataset by Chartier *et al.* (2014): we found the highest floral disparity in early-diverging and species-poor grades and clades, and the lowest disparity in highly nested and species-rich clades (Fig. 3).

The ANA grade, magnoliids, and the other eudicots grade display a broad range of floral trait combinations appearing widely spread in the morphospace. In addition, most species with the highest eccentricity belong to these groups (Fig. S11). It is expected for grades to show more variability than clades per definition. Old clades are also known to present more variable floral traits. For instance, floral phyllotaxis is particularly labile in members of the ANA grade and magnoliids (Endress & Doyle, 2009), where whorled and spiral phyllotaxis often coexist at shallow phylogenetic levels, and merism is highly variable within early-diverging angiosperms and members of the other eudicots grade (Endress, 2011). The most eccentric species often show a marked reduction of organs, such as perianth loss, allowing the arrangement and number of stamens and carpels to become highly labile (e.g. *Eupomatia* and *Galbulimina* in the magnoliids, *Euptelea*, and *Trochodendron* in the other eudicots grade, and *Cyclanthus* in the other monocots grade; Endress, 2011).

Another trend is the stabilization of floral organization within highly nested groups such as malvids, campanulids, and lamiids, which occupy relatively narrow regions in the space and show low levels of disparity coupled with high species richness (Fig. 3). Some floral characters became canalized or fixed within these major clades. For example, floral phyllotaxis is generally whorled and merism became stabilized to pentamery (or tetramery) in Pentapetalae and trimery in commelinids (Endress, 2011). This tendency in angiosperm floral evolution toward fixation of the number and position of organs, and toward repeated evolution of some traits such as zygomorphy (Reyes *et al.*, 2016), is often associated with increased synorganization, that is when organs of the same module (e.g. perianth) or different modules (e.g. androecium and gynoecium) become integrated into architecturally and functionally complex structures (Endress, 2002, 2006, 2016). The lower disparity in floral organization found in nested clades does not preclude, and possibly even has enabled variation in other aspects of floral structure, for instance, traits associated with mechanical properties (e.g. size and proportions) or traits directly related to interactions with pollinators (e.g. organ shape, floral color, and rewards). Such floral traits, which have not been considered in this study, tend to be very labile, even among closely related species (Endress, 1994, 2011).

Conclusion

Our quantification of disparity through time revealed a pattern of highest disparity early during angiosperm floral evolution. Disparity in the Early Cretaceous might have been even higher than our present estimation since numerous plant fossils with reproductive organs possibly belonging to the angiosperms have not been formally described yet (Sauquet & Magallón, 2018). Studies on disparity would greatly benefit from the description of such fossils, even if they cannot be assigned to extant lineages. It seems very surprising that there are hardly any described extinct angiosperm families and orders, which one would expect when considering the long evolutionary history of the group (but see Sun *et al.*, 2002; Pessoa *et al.*, 2023). Other fossils, such as *Bevhalstia pebja* Hill, 1996 (Friis *et al.*, 2011), do not present enough characters to be added to disparity studies. It also has to be mentioned that the fossil record of angiosperms is geographically biased since most of the currently available specimens stem from the mid-northern paleolatitudes (Friis *et al.*, 2011; Xing *et al.*, 2016). Studies in extant lineages have shown that floral disparity may vary with latitude and with regional factors (Chartier *et al.*, 2021), and it would be interesting to test whether such patterns are also present in the floral fossil record. However, it is currently not possible to address this question due to the paucity of the floral fossil record from the southern hemisphere. Furthermore, our results illustrate that ASRs are, by nature, conservative. First, while the ages of the crown nodes of the evaluated groups date to the early Cretaceous (or earlier; Ramírez-Barahona *et al.*, 2020; Table S1), the corresponding ASRs fall within the small morphospace areas occupied by most extant species included in the dataset (Fig. 2). This is discrepant with the documented positive correlation between node age and disparity (Fig. 4), and hence, with our expectation that oldest nodes would occupy morphospace areas outside those occupied by the majority of extant species. Second, while ASRs may document character state combinations that are not present among the original species sample, they are methodologically precluded from estimating character states that are not present in this sample. Third, our results empirically show that fossils document greater floral disparity than ASRs (Fig. 2). Nonetheless, further ASR-based analyses would complement our approach as they would likely add crucial qualitative and quantitative information on the evolution of the group, through the exploration of morphological rates of evolution through time and through ASR incorporating fossil data. Finally, the analyses of morphospaces for vegetative and for other reproductive characters (e.g., fruits) would achieve an integrative understanding of angiosperm evolution.

Acknowledgements

We thank the University of Vienna for funding the eFLOWER server hosting the PROTEUS database. We thank S. Ramírez-Barahona for providing the dated trees used for ASR, and the Instituto de Biología, Proyecto Fronteras de la Ciencia (2016-01-1867- CONAHCYT to SM), USTIC, Instituto de Biología, D. Velázquez, and A. Wong for the access to the HPC infrastructure where the ASR were conducted. We thank A. Delgado-Salinas, L. E. Eguiarte, S. Gerber, and three anonymous reviewers for their helpful feedback on earlier drafts. This work was funded by the Consejo Nacional de Humanidades Ciencias y Tecnologías (CONAHCYT) graduate scholarship 708068 to AMLM, by the International Association for Plant Taxonomists (IAPT), by the American Society of Plant Taxonomists (ASPT), and by the Austrian Science Fund (FWF P 250077-B16 and P 31101-B29).

Competing interests

None declared.

Author contributions

MC, AML-M, HS, JS, MvB and SM designed the study. MvB, JS, AML-M and HS scored and curated the data for fossils. HS provided the scripts for ancestral state reconstructions. MC and AML-M conducted the analyses. AML-M and MC wrote the paper. All co-authors interpreted the results and edited the manuscript.

ORCID

Maria von Balthazar https://orcid.org/0000-0001-7436-3582
Marion Chartier https://orcid.org/0000-0001-6757-4760
Andrea M. López-Martínez https://orcid.org/0000-0002-2027-9753
Susana Magallón https://orcid.org/0000-0002-6838-7497
Hervé Sauquet https://orcid.org/0000-0001-8305-3236
Jürg Schönenberger https://orcid.org/0000-0001-6791-2731

Data availability

Data are available in the Supporting Information.

References

Allen CE, Beldade P, Zwaan BJ, Brakefield PM. 2008. Differences in the selection response of serially repeated color pattern characters: standing variation, development, and evolution. *BMC Evolutionary Biology* **8**: 1–13.

APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181**: 1–20.

Armbruster WS, Mulder CPH, Baldwin BG, Kalisz S, Wessa B, Nute H. 2002. Comparative analysis of late floral development and mating system evolution in Tribe Collinsieae (Scrophulariaceae SL). *American Journal of Botany* **89**: 37–49.

Asar Y, Ho SYW, Sauquet H. 2022. Early diversifications of angiosperms and their insect pollinators: were they unlinked? *Trends in Plant Science* **27**: 858–869.

Atkinson BA, Stockey RA, Rothwell GW, Mindell RA, Bolton MJ. 2015. Lauraceous flowers from the Eocene of Vancouver Island: *Tinaflora beardiae* gen. et sp. nov. (Lauraceae). *International Journal of Plant Sciences* **176**: 567–585.

Augusto L, Davies TJ, Delzon S, De Schrijver A. 2014. The enigma of the rise of angiosperms: can we untie the knot? *Ecology Letters* **17**: 1326–1338.

Avena-Koenigsberger A, Goñi J, Solé R, Sporns O. 2015. Network morphospace. *Journal of the Royal Society Interface* **12**: 20140881.

Bapst DW, Bullock P, Melchin MJ, Sheets HD, Mitchell CE. 2012. Graptoloid diversity and disparity became decoupled during the Ordovician mass extinction. *Proceedings of the National Academy of Sciences, USA* **109**: 3428–3433.

Beaulieu JM, O'Meara BC, Donoghue MJ. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* **62**: 725–737.

Benitez-Vieyra S, Ordano M, Fornoni J, Boege K, Domınguez CA. 2010. Selection on signal–reward correlation: limits and opportunities to the evolution of deceit in *Turnera ulmifolia* L. *Journal of Evolutionary Biology* **23**: 2760–2767.

Benton MJ. 2015. Exploring macroevolution using modern and fossil data. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20150569.

Benton MJ, Forth J, Langer MC. 2014. Models for the rise of the dinosaurs. *Current Biology* **24**: 87–95.

Benton MJ, Wilf P, Sauquet H. 2022. The angiosperm terrestrial revolution and the origins of modern biodiversity. *New Phytologist* **233**: 2017–2035.

Brusatte SL, Benton MJ, Ruta M, Lloyd GT. 2008a. Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. *Science* **321**: 1485–1488.

Brusatte SL, Benton MJ, Ruta M, Lloyd GT. 2008b. The first 50 Myr of dinosaur evolution: macroevolutionary pattern and morphological disparity. *Biology Letters* **4**: 733–736.

Cantino PD, Doyle JA, Graham SW, Judd WS, Olmstead RG, Soltis DE, Soltis PS, Donoghue MJ. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* **56**: E1–E44.

Carvalho M, Jaramillo C, Parra F, Caballero-Rodríguez D, Herrera F, Wing S, Turner B, D'Apolito C, Romero-Báez M, Narváez P *et al.* 2021. Extinction at the end-Cretaceous and the origin of modern Neotropical rainforests. *Science* **372**: 63–68.

Cascales-Miñana B, Cleal CJ. 2014. The plant fossil record reflects just two great extinction events. *Terra Nova* **26**: 195–200.

Castañeda-Posadas C, Cevallos-Ferriz SRS. 2007. *Swietenia* (Meliaceae) flower in late oligocene-early miocene amber from Simojovel de Allende, Chiapas, Mexico. *American Journal of Botany* **94**: 1821–1827.

Chartier M, von Balthazar M, Sontag S, Löfstrand S, Palme T, Jabbour F, Sauquet H, Schönenberger J. 2021. Global patterns and a latitudinal gradient of flower disparity: perspectives from the angiosperm order Ericales. *New Phytologist* **230**: 821–831.

Chartier M, Jabbour F, Gerber S, Mitteroecker P, Sauquet H, von Balthazar M, Staedler Y, Crane PR, Schönenberger J. 2014. The floral morphospace – a modern comparative approach to study angiosperm evolution. *New Phytologist* **204**: 841–853.

Chartier M, Löfstrand S, von Balthazar M, Gerber S, Jabbour F, Sauquet H, Schönenberger J. 2017. How (much) do flowers vary? Unbalanced disparity among flower functional modules and a mosaic pattern of morphospace occupation in the order Ericales. *Proceedings of the Royal Society B: Biological Sciences* **284**: e20170066.

Christenhusz MJ, Byng JW. 2016. The number of known plants species in the world and its annual increase. *Phytotaxa* **261**: 201–217.

Ciampaglio CN. 2002. Determining the role that ecological and developmental constraints play in controlling disparity: examples from the crinoid and blastozoan fossil record. *Evolution and Development* **4**: 170–188.

Ciampaglio CN, Kemp M, McShea DW. 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology* **27**: 695–715.

Clark JW, Hetherington AJ, Morris JL, Pressel S, Duckett JG, Puttick MN, Schneider H, Kenrick P, Wellman CH, Donoghue PC. 2023. Evolution of phenotypic disparity in the plant kingdom. *Nature Plants* **9**: 1–9.

Cohen KM, Finney SC, Gibbard PL, Fan JX. 2013; Updated. *The ICS International Chronostratigraphic Chart. Episodes 36: 199-204.* [WWW document] URL https://stratigraphy.org/chart [accessed 21 April 2022].

Condamine FL, Silvestro D, Koppelbus EB, Antonelli A. 2020. The rise of angiosperms pushed conifers to decline during global cooling. *Proceedings of the National Academy of Sciences, USA* 117: 28867–28875.

Coombs EJ, Felice RN, Clavel J, Park T, Bennion RF, Churchill M, Geisler JH, Beatty B, Goswami A. 2022. The tempo of cetacean cranial evolution. *Current Biology* **32**: 2233–2247.

Crepet WL, Friis EM, Nixon KC, Lack AJ, Jarzembowski EA. 1991. Fossil evidence for the evolution of biotic pollination. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **333**: 187–195.

Dilcher DL, Crane PR. 1984. *Archaeanthus*: an early angiosperm from the Cenomanian of the Western Interior of North America. *Annals of the Missouri Botanical Garden* 71: 351–383.

Doyle J, Endress PK. 2014. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: ANITA lines and relatives of Chloranthaceae. *International Journal of Plant Sciences* **175**: 555–600.

Endress PK. 1982. Syncarpy and alternative modes of escaping disadvantages of apocarpy in primitive angiosperms. *Taxon* **31**: 48–52.

Endress PK. 1994. *Diversity and evolutionary biology of tropical flowers.* Cambridge, UK: Cambridge University Press.

Endress PK. 2002. Origins of flower morphology. *Journal of Experimental Zoology* **291**: 105–115.

Endress PK. 2006. Angiosperm floral evolution: morphological developmental framework. *Advances in Botanical Research* **44**: 1–67.

Endress PK. 2010. Flower structure and trends of evolution in eudicots and their major subclades. *Annals of the Missouri Botanical Garden* **97**: 541–583.

Endress PK. 2011. Evolutionary diversification of the flowers in angiosperms. *American Journal of Botany* **98**: 370–396.

Endress PK. 2016. Development and evolution of extreme synorganization in angiosperm flowers and diversity: a comparison of Apocynaceae and Orchidaceae. *Annals of Botany* **117**: 749–767.

Endress PK, Doyle J. 2009. Reconstructing the ancestral flower and its initial specializations. *American Journal of Botany* **96**: 22–66.

Endress PK, Igersheim A. 1997. Gynoecium diversity and systematics of the Laurales. *Botanical Journal of the Linnean Society* **125**: 93–168.

Endress PK, Igersheim A. 1999. Gynoecium diversity and systematics of the basal eudicots. *Botanical Journal of the Linnean Society* **130**: 305–393.

Endress PK, Igersheim A. 2000. Gynoecium structure and evolution in basal angiosperms. *International Journal of Plant Sciences* **161**: S211–S223.

Erwin DH. 2007. Disparity: morphological pattern and developmental context. *Palaeontology* **50**: 57–73.

Feild TS, Brodribb TJ, Iglesias A, Chatelet DS, Baresch A, Upchurch GR Jr, Gomez B, Mohr BAR, Coiffard C, Kvacek J *et al.* 2011. Fossil evidence for Cretaceous escalation in angiosperm leaf vein evolution. *Proceedings of the National Academy of Sciences, USA* **108**: 8363–8366.

Foote M. 1991. Morphological and taxonomic diversity in clade's history: the blastoid record and stochastic simulations. *Contributions from the Museum of Paleontology* **28**: 101–140.

Foote M. 1992a. Paleozoic record of morphological diversity in blastozoan echinoderms. *Proceedings of the National Academy of Sciences, USA* **89**: 7325–7329.

Foote M. 1992b. Rarefaction analysis of morphological and taxonomic diversity. *Paleobiology* **18**: 1–16.

Foote M. 1993a. Contributions of individual taxa to overall morphological disparity. *Paleobiology* **19**: 403–419.

Foote M. 1993b. Discordance and concordance between morphological and taxonomic diversity. *Paleobiology* **19**: 185–204.

Foote M. 1994. Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. *Paleobiology* **20**: 320–344.

Foote M. 1995. Morphological diversification of Paleozoic Crinoids. *Paleobiology* **21**: 273–299.

Foote M. 1997. The evolution of morphological diversity. *Annual Review of Ecology, Evolution, and Systematics* **28**: 129–152.

Foote M. 1999. Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. *Paleobiology* **25**: 1–115.

Fortey RA, Briggs DEG, Wills MA. 1996. The Cambrian evolutionary 'explosion': decoupling cladogenesis from morphological disparity. *Biological Journal of the Linnean Society* 57: 13–33.

Friedman M. 2010. Explosive morphological diversification of spiny-finned teleost fishes in the aftermath of the end-Cretaceous extinction. *Proceedings of the Royal Society B: Biological Sciences* 277: 1675–1683.

Friis EM, Crane PR, Pedersen KR. 2011. *Early flowers and angiosperm evolution.* Cambridge, UK: Cambridge University Press.

Friis EM, Pedersen KR. 2011. *Canrightia resinifera* gen. et sp. nov., a new extinct angiosperm with *Retimonocolpites*-type pollen from the Early Cretaceous of Portugal: missing link in the eumagnoliid tree? *Grana* **50**: 3–29.

Friis EM, Pedersen KR. 2012. *Bertilanthus scanicus*, a new asterid flower from the Late Cretaceous (Late Santonian-Early Campanian) of Scania, Sweden. *International Journal of Plant Sciences* **173**: 318–330.

Friis EM, Pedersen KR, Crane PR. 1999. Early angiosperm diversification: the diversity of pollen associated with angiosperm reproductive structures in Early Cretaceous floras from Portugal. *Annals of the Missouri Botanical Garden* **86**: 259–296.

Friis EM, Pedersen KR, Crane PR. 2006. Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. *Palaeogeography, Palaeoclimatology, Palaeoecology* **232**: 251–293.

Friis EM, Pedersen KR, Crane PR. 2010. Diversity in obscurity: fossil flowers and the early history of angiosperms. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **365**: 369–382.

Gandolfo MA, Nixon KC, Crepet WL. 2004. Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. *Proceedings of the National Academy of Sciences, USA* **101**: 8056–8060.

Gandolfo MA, Nixon KC, Crepet WL, Grimaldi DA. 2018. A Late Cretaceous fagalean inflorescence preserved in amber from New Jersey. *American Journal of Botany* **105**: 1424–1435.

Gerber S. 2019. Use and misuse of discrete character data for morphospace and disparity analyses. *Palaeontology* **62**: 305–319.

Giraudoux P, Antonietti J-P, Beale C, Pleydell D, Treglia M. 2018. *PGIRMESS: spatial analysis and data mining for field ecologists. R Package v.19.* [WWW document] URL https://github.com/pgiraudoux/pgirmess [accessed 28 March 2023].

Gould SJ. 1989. *Wonderful life.* New York, NY, USA: W. W. Norton.

Gould SJ, Gilinsky NL, German RZ. 1987. Asymmetry of lineages and the direction of evolutionary time. *Science* **236**: 1437–1441.

Guillerme T, Cooper N, Brusatte SL, Davis KE, Jackson AL, Gerber S, Goswami A, Healy K, Hopkins MJ, Jones MEH *et al.* 2020. Disparities in the analysis of morphological disparity. *Biology Letters* **16**: 20200199.

Harder LD, Barrett SC, eds. 2006. *Ecology and evolution of flowers.* New York, NY, USA: Oxford University Press.

Hernández-Damián AL, Gómez-Acevedo SL, Cevallos-Ferriz SRS. 2018. Fossil flower of *Salacia lombardii* sp. nov. (Salacioideae-Celastraceae) preserved in amber from Simojovel de Allende, Mexico. *Review of Palaeobotany and Palynology* **252**: 1–9.

Hill JJ, Puttick MN, Stubbs TL, Rayfield EJ, Donoghue PC. 2018. Evolution of jaw disparity in fishes. *Palaeontology* **61**: 847–854.

Hopkins MJ, Gerber S. 2021. Morphological disparity. In: Nuno de la Rosa L, Müller G, eds. *Evolutionary developmental biology. A reference guide.* Cham, Switzerland: Springer, 965–976.

Hughes M, Gerber S, Wills MA. 2013. Clades reach highest morphological disparity early in their evolution. *Proceedings of the National Academy of Sciences, USA* **110**: 13875–13879.

Jardine PE, Palazzesi L, Tellería MC, Barreda VD. 2022. Why does pollen morphology vary? Evolutionary dynamics and morphospace occupation in the largest angiosperm order (Asterales). *New Phytologist* **234**: 1075–1087.

Kawahara AY, Storer C, Carvalho APS, Plotkin DM, Condamine FL, Braga MP, Ellis EA, St Laurent RA, Li X, Barve V *et al.* 2023. A global phylogeny of butterflies reveals their evolutionary history, ancestral hosts and biogeographic origins. *Nature Ecology & Evolution* 7: 903–913.

Korn D, Hopkins MJ, Walton SA. 2013. Extinction space – a method for the quantification and classification of changes in morphospace across extinction boundaries. *Evolution* **67**: 2795–2810.

Kruskal WH, Wallis WA. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* **47**: 583–621.

Leslie A, Simpson C, Mander L. 2021. Reproductive innovations and pulsed rise in plant complexity. *Science* **373**: 1368–1372.

López-Martínez AM, Schönenberger J, von Balthazar M, González-Martínez CA, Ramírez-Barahona S, Sauquet H, Magallón S. 2023. Integrating fossil flowers into the angiosperm phylogeny using molecular and morphological evidence. *Systematic Biology* **72**: 837–855.

Lupia R. 1999. Discordant morphological disparity and taxonomic diversity during the Cretaceous angiosperm radiation: North American pollen record. *Paleobiology* **25**: 1–28.

Lupia R, Lidgard S, Crane PR. 1999. Comparing palynological abundance and diversity: implications for biotic replacement during the Cretaceous angiosperm radiation. *Paleobiology* **25**: 305–340.

Magallon S, Crane PR, Herendeen PS. 1999. Phylogenetic pattern, diversity, and diversification of eudicots. *Annals of the Missouri Botanical Garden* **86**: 297–372.

Martínez-Cabrera HI, Zheng J, Estrada-Ruiz E. 2017. Wood functional disparity lags behind taxonomic diversification in angiosperms. *Review of Palaeobotany and Palynology* **246**: 251–257.

Matthews ML, Endress PK. 2002. Comparative floral structure and systematics in Oxalidales (Oxalidaceae, Connaraceae, Cephalotaceae, Brunelliaceae, Cunoniaceae, Elaeocarpaceae, Tremandraceae). *Botanical Journal of the Linnean Society* **140**: 321–381.

Matthews ML, Endress PK. 2005. Comparative floral structure and systematics in Celastrales (Celastraceae, Parnassiaceae, Lepidobotryaceae). *Botanical Journal of the Linnean Society* **149**: 129–194.

Matthews ML, Endress PK. 2006. Floral structure and systematics in four orders of rosids, including a broad survey of floral mucilage cells. *Plant Systematics and Evolution* **260**: 223–251.

McElwain JC, Punyasena SW. 2007. Mass extinction events and the plant fossil record. *Trends in Ecology & Evolution* **22**: 548–557.

McGhee GR. 1991. Theoretical morphology: the concept and its applications. *Short Courses in Paleontology* **4**: 87–102.

McGhee GR. 2015. Limits in the evolution of biological form a theoretical morphologic perspective. *Interface Focus* **5**: 20150034.

de Mendiburu F, Yaseen M. 2020. *AGRICOLAE: statistical procedures for agricultural research. R Package v.10.* [WWW document] URL https://cran.r-project.org/web/packages/agricolae/index.html [accessed 28 March 2023].

Mitteroecker P, Huttegger SM. 2009. The concept of morphospaces in evolutionary and developmental biology: mathematics and metaphors. *Biological Theory* **4**: 54–67.

Mohr B, Eklund H. 2004. *Araripia florifera*, a magnoliids angiosperm from the Lower Cretaceous Crato Formation (Brazil). *Review of Palaeobotany and Palynology* **126**: 279–292.

Moon BC, Stubbs TL. 2020. Early high rates and disparity in the evolution of ichthyosaurs. *Communications Biology* **3**: 1–8.

van der Niet T, Johnson SD. 2012. Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends in Ecology & Evolution* **27**: 353–361.

O'Meara BC, Smith SD, Armbruster SW, Harder LD, Hardy CR, Hileman LC, Hufford L, Litt A, Magallón S, Smith SA *et al.* 2016. Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proceedings of the Royal Society B: Biological Sciences* **283**: 20152304.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P *et al.* 2020. *VEGAN: community ecology package. R package v 2.5-7.* [WWW document] URL https://github.com/vegandevs/vegan [accessed 28 March 2023].

Oyston JW, Hughes M, Gerber S, Wills MA. 2016. Why should we investigate the morphological disparity of plant clades? *Annals of Botany* **117**: 859–879.

Oyston JW, Hughes M, Wagner PJ, Gerber S, Wills MA. 2015. What limits the morphological disparity of clades? *Interface Focus* **5**: 20150042.

Peña-Kairath C, Delclòs X, Álvarez-Parra S, Peñalver E, Engel MS, Ollerton J, Peris D. 2023. Insect pollination in deep time. *Trends in Ecology & Evolution* **38**: 749–759.

Peñalver E, Arillo A, Perez-de la Fuente R, Riccio ML, Delclos X, Barron E, Grimaldi DA. 2015. Long-proboscid flies as pollinators of Cretaceous gymnosperms. *Current Biology* **25**: 1917–1923.

Pessoa EM, Ribeiro AC, Christenhusz MJ. 2023. New evidence on the previously unknown gynoecium of *Araripia florifera* (Araripiaceae, fam. nov.), a magnoliid angiosperm from the Lower Cretaceous (Aptian) of the Crato Konservat-Lagerstätte (Araripe Basin), northeastern Brazil. *Cretaceous Research* **153**: 105715.

Prentice KC, Ruta M, Benton MJ. 2011. Evolution of morphological disparity in pterosaurs. *Journal of Systematic Palaeontology* **9**: 337–353.

Puttick MN, Guillerme T, Wills MA. 2020. The complex effects of mass extinctions on morphological disparity. *Evolution* **74**: 2207–2220.

R Core Team. 2022. *R: a language and environment for statistical computing, v4.1.3.* Vienna, Austria: R foundation for Statistical Computing. [WWW document] URL http://www.r-project.org [accessed 28 March 2023].

Ramírez-Barahona S, Sauquet H, Magallón S. 2020. The delayed and geographically heterogeneous diversification of flowering plant families. *Nature Ecology & Evolution* **4**: 1232–1238.

Raup DM. 1967. Analysis of shell coiling: coiling in ammonoids. *Journal of Paleontology* **41**: 43–65.

Raup DM, Michelson A. 1965. Theoretical morphology of the coiled shell. *Science* **147**: 1294–1295.

Revell LJ. 2012. PHYTOOLS: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.

Reyes E, Sauquet H, Nadot S. 2016. Perianth symmetry changed at least 199 times in angiosperm evolution. *Taxon* **65**: 945–964.

Sargent RD. 2004. Floral symmetry affects speciation rates in angiosperms. *Proceedings of the Royal Society B: Biological Sciences* **271**: 603–608.

Sauquet H. 2019. *PROTEUS: a database for recording morphological data and calibrations. v 1.27.* [WWW document] URL https://eflower.myspecies.info/proteus [accessed 20 April 2022].

Sauquet H, Magallón S. 2018. Key questions and challenges in angiosperm macroevolution. *New Phytologist* **219**: 1170–1187.

Sauquet H, Ramírez-Barahona S, Magallón S. 2022. What is the age of flowering plants? *Journal of Experimental Botany* **73**: 3840–3853.

Sauquet H, von Balthazar M, Magallón S, Doyle JA, Endress PK, Bailes EJ, Barroso de Morais E, Bull-Hereñu K, Carrive L, Chartier M *et al.* 2017. The ancestral flower of angiosperms and its early diversification. *Nature Communications* **8**: 16047.

Schönenberger J, von Balthazar M, López Martínez A, Albert B, Prieu C, Magallón S, Sauquet H. 2020. Phylogenetic analysis of fossil flowers using an angiosperm-wide data set: proof-of-concept and challenges ahead. *American Journal of Botany* **107**: 1433–1448.

Schönenberger J, von Balthazar M, Sytsma KJ. 2010. Diversity and evolution of floral structure among early diverging lineages in the Ericales. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **365**: 437–448.

Schönenberger J, von Balthazar MV. 2006. Reproductive structures and phylogenetic framework of the rosids-progress and prospects. *Plant Systematics and Evolution* **260**: 87–106.

Schönenberger J, Friis EM. 2001. Fossil flowers of Ericalean affinity from the Late Cretaceous of Southern Sweden. *American Journal of Botany* **88**: 467–480.

Slater GJ, Price SA, Santini F, Alfaro MF. 2010. Diversity versus disparity in the radiation of modern cetaceans. *Proceedings of the Royal Society B: Biological Sciences* **277**: 3097–3104.

Smith TJ, Donoghue PC. 2022. Evolution of fungal phenotypic disparity. *Nature Ecology & Evolution* **6**: 1489–1500.

Sneath PH, Sokal RR. 1973. *Numerical taxonomy. The principles and practice of numerical classification.* San Francisco, CA, USA: W. H. Freeman.

Stebbins GL. 1951. Natural selection and the differentiation of angiosperm families. *Evolution* **5**: 299–324.

Stephens RE, Gallagher RV, Dun L, Cornwell W, Sauquet H. 2023. Insect pollination for most of angiosperm evolutionary history. *New Phytologist* **240**: 880–891.

Stewart AB, Diller C, Dudash MR, Fenster CB. 2022. Pollination-precision hypothesis: support from native honey bees and nectar bats. *New Phytologist* **235**: 1629–1640.

Sun G, Ji Q, Dilcher DL, Zheng S, Nixon KC, Wang X. 2002. Archaefructaceae, a new basal angiosperm family. *Science* **296**: 899–904.

Thompson J, Ramírez-Barahona S. 2023. No phylogenetic evidence for angiosperm mass extinction at the Cretaceous – Paleogene (K-Pg) boundary. *Biology Letters* **19**: 20230314.

Wagner PJ. 2018. Early burst of disparity and the reorganization of character integration. *Proceedings of the Royal Society B: Biological Sciences* **285**: 20181604.

Walker-Larsen J, Harder LD. 2000. The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention. *American Journal of Botany* **87**: 1367–1384.

WFO. 2023. *World flora online*. [WWW document] URL http://www.worldfloraonline.org/ [accessed 11 April 2023].

Wilf P, Carvalho M, Stiles E. 2023. The end-Cretaceous plant extinction: heterogeneity, ecosystem transformation, and insights for the future. *Cambridge Prisms: Extinction* **1**: 1–10.

Willis KJ, McElwain JC. 2002. *The evolution of plants*. Oxford, UK: Oxford University Press.

Wills M, Briggs D, Fortey R. 1994. Disparity as an evolutionary index: a comparison of Cambrian and recent arthropods. *Paleobiology* **20**: 93–130.

Xing Y, Gandolfo MA, Onstein RE, Cantrill DJ, Jacobs BF, Jordan GJ, Lee DF, Popova S, Srivastava R, Su T *et al.* 2016. Testing the biases in the rich Cenozoic angiosperm macrofossil record. *International Journal of Plant Sciences* **177**: 371–388.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 Script used to generate the distance matrix, the ordination of the space, and calculate *D*.

Dataset S2 Morphological dataset, information about scored fossils, and about study groups.

Dataset S3 Proteus extraction used to generate the dataset, containing all references.

Dataset S4 Ancestral state reconstructions for each character obtained with a maximum likelihood approach.

Dataset S5 Ancestral state reconstructions for each character obtained with a stochastic mapping approach.

Fig. S1 Distribution of missing and nonapplicable data for living species.

Fig. S2 Distribution of missing and nonapplicable data for fossil species.

Fig. S3 Rarefied floral disparity (*meanD* and *R*) through time.

Fig. S4 Position of Neogene fossils in the floral morphospace (Fig. 1).

Fig. S5 Distribution of character states in the morphospace (Fig. 2).

Fig. S6 Position of eccentric species in the morphospace ordination.

Fig. S7 Distribution of character states in living Angiosperms.

Fig. S8 Rarefied floral disparity (*meanD* and *R*) for living angiosperms.

Fig. S9 Position of additional living angiosperm clades in the floral morphospace (Fig. 3).

Fig. S10 Correlation between sample size and group sizes.

Fig. S11 Phylogenetic distribution of species in the four time bins studied and for eccentric species.

Methods S1 List of nonapplicable and impossible combinations.

Methods S2 Calculation of the mean character difference.

Table S1 Divergence time estimates for the 15 ASR.

Table S2 Disparity (*meanD*) for the study angiosperm group.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Material Suplementario

El material suplementario como parte de los contenidos de la publicación del *Capítulo II* se encuentra disponible de manera online en el siguiente link:
https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.19389

CAPÍTULO III

Exploring total evidence dating under the skyline fossilized birth-death model applied to angiosperms*

*Se presentan resultados preliminares. Esta investigación continua en proceso y está siendo realizada en colaboración con **Dra. Susana Magallón** (Instituto de Biología, Universidad Nacional autónoma de México), **Dr. Hervé Sauquet** (National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, Australia) y **Dra. Maria von Balthazar** y **Dr. Jürg Schönenberger** (Department of Botany and Biodiversity Research, University of Vienna, Viena, Austria).

RESUMEN – El modelo de nacimiento y muerte fosilizado (FBD, por sus siglas en inglés) es un marco de referencia bayesiano para inferir tiempos de divergencia y procesos macroevolutivos, como las tasas de evolución y diversificación. Aunque el modelo FBD se ha probado en una amplia gama de escenarios de simulación, su aplicación a estudios empíricos que comprenden grupos ricos en especies sigue siendo poco explorada. En este estudio realizamos múltiples experimentos analíticos en un conjunto de datos que comprende de 62 fósiles bien conservados y 792 flores vivientes para explorar el comportamiento del modelo de nacimiento y muerte fosilizado por horizontes de tiempo (SFBD) aplicado a análisis de datos a gran escala filogenética. Aplicamos este enfoque para investigar la influencia del modelo generador del árbol, junto con la incorporación de datos morfológicos en la estimación de los tiempos de divergencia de las familias de angiospermas y la variación en las tasas de recuperación de fósiles, diversificación y tasas morfológicas. Nuestros análisis indican edades de los grupos troncales similares, pero edades de los grupos corona considerablemente más jóvenes para la mayoría de las familias de angiospermas que las reportadas en estudios previos. Detectamos las tasas de cambio morfológico más rápidas en las ramas más profundas del árbol, que corresponden al origen de los principales clados de angiospermas y a la expansión más amplia del morfoespacio floral. Finalmente, discutimos los desafíos y proponemos mejores prácticas en la aplicación del enfoque de evidencia total bajo el modelo de nacimiento-muerte fosilizado para la estimación simultánea de procesos macroevolutivos a filogenias a gran escala.

Palabras clave: angiospermas, evidencia total, modelo nacimiento-muerte fosilizado, reloj morfológico, selección, tasas morfológicas.

1 ABSTRACT

The fossilized birth-death model (FBD) is a powerful Bayesian framework for inferring divergence times and macroevolutionary processes such as evolutionary and diversification rates. Although the FBD model has been tested under a wide range of simulation scenarios, its application to empirical studies comprising species-rich groups has remained underexplored. In this study, we conducted multiple analytical experiments on a comprehensive dataset of 62 well-preserved fossils and 792 extant flowers to explore the behavior of the skyline fossilized birth-death model (SFBD) applied to large-scale datasets. We apply this approach to investigate the influence of the tree prior model along with the incorporation of morphological data in the estimation of divergence times of angiosperm families and the variation in fossil recovery, diversification, and morphological rates. Our analyses indicate similar stem ages but considerably younger crown ages of most angiosperm families than those reported in previous studies. Additionally, we detected the fastest rates of morphological change at the deepest branches of the tree, which correspond to the origin of major angiosperm clades and the broadest expansion of the floral morphospace. Finally, we discuss challenges and propose best practices in the application of the total-evidence approach under the skyline fossilized birth-death model in large-scale phylogenies.

Keywords: angiosperms, fossilized birth-death model, morphological clock, phenotypic rates, selection, total evidence.

2 INTRODUCTION

Fossils provide a unique source of information to infer past evolutionary dynamics, including the time and magnitude of mass extinction events (Raup and Sepkoski, 1982; Jablonski and Raup, 1995), the nature and potential causes of extinction (Marshall, 2017), or tracing morphological variation over time (Gould, 1989; Foote, 1994, Lupia, 1999). Comparative methods based only on extant taxa are also helpful tools for estimating deep-time processes, however, they have been widely questioned due to their lack of reliability when compared to fossil-based estimates (Quental and Marshall, 2010). Therefore, significant methodological progress has been made over the last decade to combine paleontological and neontological data into phylogenetic frameworks. One of the major advances is the development of the fossilized birth-death (FBD) model, which unifies fossils and extant species into a single macroevolutionary framework to infer divergence times and macroevolutionary parameters such as speciation, extinction, fossilization, and sampling rates (Heath et al., 2014; Zhang et al., 2016; Gavryushkina et al., 2017). Compared to previous approaches in the estimation of divergence times, the FBD model has offered several advantages, including the use of all available fossils for the group of interest, avoiding arbitrarily chosen statistical distribution for fossil calibrations, and accounting for the possibility of sample fossils as ancestors (Gavryushkina et al., 2014).

In recent years, several extensions have been developed to account for variations of the original FBD model (Zhang et al., 2016, Gavryushkina et al., 2017; Zhang et al., 2023). For instance, the model allows the use of molecular data sampled from extant species and morphological data scored from extant and fossil species to jointly estimate topology, divergence times, and evolutionary rates, an approach referred to as "tip dating" (Zhang et al., 2016, Gavryushkina, et al., 2017). This approach has proven to be an extraordinary tool for estimating divergence times due to its better account for uncertainty in the position of fossils and high accuracy (Gavryushkina et al., 2017). However, several aspects of this approach remain to be improved, including the use of the simple Mk model (with some variations) to describe the process of morphological character evolution (Ronquist et al., 2016); and potential bias associated with the analysis of large amounts of missing data, which is an inherent problem in morphological datasets containing fossils. Additional extensions of the FBD include the accommodation of taxon-sampling schemes for the nonuniform sampling of extant taxa. Zhang et al. (2016) introduced diversified sampling – extant species in the tree are sampled to maximize diversity, which is a common practice in studies of species-rich groups. Furthermore, the FBD model may also account for variation of diversification and fossil sampling rates between time intervals (the skyline-fossilized birth-death process (SFBD)), the implementation of which is more robust to deep-root attraction and sampling biases (Zhang et al., 2023).

Several studies have used the total evidence dating approach under the FBD as a starting point to infer multiple macroevolutionary processes, including speciation and extinction rate variation in whales (Mitchell et al., 2019), the biogeographical history of various plant clades (ferns: May et al., 2021; the angiosperm order Fagales: Zhang et al., 2022; cycads: Coiro et al., 2023), or to describe patterns of phenotypic evolution (amniotes: Borcklehurst and Benson, 2021; tetrapods: Simões and Pierce, 2021). However, most of these empirical studies have been performed using a relatively low number of extant species or restricted taxonomic scope. This is expected given the difficulties in the assembly of morphological matrices for a large sample of species along with the consequent increase in computational time to conducting these analyses on a large sample of species. However, evaluating the feasibility of FBD to species-rich groups is useful as we are starting to obtain denser (morphological and molecular) datasets at large phylogenetic scales.

Here we used the angiosperms, the most diverse group of plants, as a case study to evaluate the performance of the Bayesian total evidence approach under the skyline FBD model in the analysis of large-scale datasets. Compared to other land plant groups, angiosperms are characterized by a marked increase in reproductive organs complexity through the development of structures with ovules enclosed within the carpel, dedicated pollen-producing structures (stamens), and specialized organs linked to animal pollination and dispersal (Herendeen et al., 2017; Leslie et al., 2021). Angiosperms are relatively well represented in the fossil record. In particular, numerous exceptionally well-preserved fossil flowers from diverse Cretaceous localities have been described in detail (e.g., Magallón et al., 2001; Schönenberger and Friis, 2001; Gandolfo et al., 2004; von Balthazar et al., 2008; Friis and Pedersen, 2011, 2012). Such fossil flowers have provided invaluable information about the pattern of diversification of angiosperm flowers and the timing of acquisition of new structures (Friis et al., 2011). In addition, considerable progress in the integration of fossil flowers into extensive morphological matrices with extant species has improved our understanding of the phylogenetic position of fossil flowers in the angiosperm tree of life (Doyle and Endress, 2010; 2014; Schönenberger et al., 2020; López-Martínez et al., 2023a), as well as the evolutionary pattern of morphological diversity (disparity) through time (López-Martínez et al., 2023b).

Few studies have focused on estimating rates of morphological evolution for angiosperms as a whole (Sims et al., 2013; O'Meara et al., 2016; Reyes et al., 2018; Asar et al., 2023; Herting et al., 2023). For instance, some studies have revealed heterogeneous rates of morphological evolution among lineages, characters (e.g., rates differ among five perianth characters; Reyes et al., 2018), but also among floral modules (androecium, perianth, gynoecium; Herting et al., 2023). Herting et al. (2023) indicated that, although there are slow and fast-evolving rates in each floral module, the overall evolutionary rate of the androecium module is higher than the rate of the perianth and gynoecium. Additionally, O'Meara et al. (2016) examined the role of morphological transition rates and differential diversification, and showed that lineages exhibiting a trait combination characterized by the presence of corolla, bilateral symmetry, and reduced stamen number experience high diversification rates. However, trait combinations leading to the rapid diversification of a clade may not be as common as we expected.

In this study, we use the skyline fossilized birth-death model to infer the time of origin of angiosperm families and morphological (floral) rate variation through the analysis of a comprehensive dataset for angiosperm flowers, comprising 792 extant, representing most angiosperm families, and 62 well-preserved fossil flowers. We examine how the tree prior and variation in character sampling influence age estimates as well as diversification, fossilization, and morphological rates. Furthermore, we inferred significantly accelerated and decelerated rates of morphological change in the angiosperm floral bauplan across angiosperm lineages and through time.

3 METHODS

3.1 Data

We used the taxon sampling from López-Martínez et al. 2023b, which includes 1201 extant species representing all crown nodes of currently recognized angiosperm families (*sensu* APG IV, 2016), and 121 best-preserved fossil flowers described to date (for additional information about fossils, see **Supplementary Data 1**). The molecular dataset was obtained from Ramírez-Barahona et al. (2020), consisting of four plastid (*rbcL, atpB, matK, ndhF*) and three nuclear loci (*18S, 26S,* and *5.8 nrDNA*). The final concatenated matrix contained a total of 11,395 nucleotides. For morphological data, we used the matrix of López-Martínez et al. (2023b) which contains 30 discrete floral characters describing the perianth, androecium, and gynoecium. Character scoring was based on a detailed revision of the literature (e.g., floras, books, and articles) following the strict exemplar approach (*sensu* Sauquet et al., 2017). All morphological data are available in López-Martínez et al. (2023a) and López-Martínez et al. (2023b) and are stored in the PROTEUS database (Sauquet, 2019) from the project eFLOWER (https://eflower.myspecies.info/). Specifically, fossils were scored based on the original description and photographs of specimens, which in a few cases involved a reinterpretation of their morphology. Further information about the description of characters and scoring philosophy is available from Schönenberger et al. (2020).

Our previous phylogenetic analyses using a similar molecular and morphological dataset (López-Martínez et al., 2023a) resulted in difficulties in achieving convergence in Bayesian inferences. Therefore, we decided to reduce the sampling of extant species by retaining only two extant terminals per family (preserving crown nodes) and removing all parasitic or heterotrophic species that showed high molecular rates (i.e., showing extremely long terminal branches in the phylogram). Additionally, we removed fossil species that contained high levels of missing data (i.e., those species with less than 10 characters scored). A total of 409 extant species and 13 fossils were removed from the original molecular and morphological datasets after these procedures. Our final dataset included a total of 792 extant species representing 425 angiosperm families and 108 fossils.

3.2 Phylogenetic placement of fossils and visualization of uncertainty

To evaluate the phylogenetic position of fossil flowers, we conducted a maximum likelihood (ML) analysis in RAxML v.8.0 (Stamatakis, 2014). For this analysis, we followed the same *"CML-24-TE"* approach described in López-Martínez et al. (2023a). This approach consists of constraining the tree topology among extant species and focusing on estimating the placement of fossils. We applied the GTR model with a gamma-shaped rate variation across sites for six molecular partitions: *rbcL* first and second positions; *rbcL* 3rd position; *atpB* first and second positions; *atpB* third position; *matK; ndhF*; and *nuclear*. The morphological partition was analyzed using the Mkv model with "lewis" ascertainment bias correction (Lewis, 2001) for the use of only variable characters. We ran 1000 replicates of non-parametric bootstrap.

We visualized phylogenetic uncertainty in the position of fossil flowers with RoguePlots (Klopfstein and Spasojevic, 2019). These graphs summarize all sampled fossil positions across the tree according to the support value obtained from bootstrap replicates. The RoguePlots for each fossil were generated with the package rogue.plot (Klopfstein and Spasojevic, 2019) in R (R Core Team, 2022).

3.3 Divergence time estimation and evolutionary rates

3.3.1 Unconstrained analysis under the skyline fossilized birth-death model

We estimated divergence times and evolutionary rates using a Bayesian total evidence dating approach under the skyline fossilized birth-death process (SFBD) implemented in MrBayes v.3.2.7 (Ronquist et al., 2012). We first conducted a total evidence SFBD analysis, which included the full sampling of extant (792) and fossil species (108) without applying any topological constraint for fossils (we refer to this analysis as '*unconstrained SFBD*'), to compare the position of fossils with the undated analysis. For this analysis, molecular and morphological character evolution was modeled as in the maximum likelihood analysis.

We assumed a variable evolutionary rate across tree branches by implementing an independent gamma rate (IGR) relaxed clock model (*clockvarpr = igr*). The IGR clock model was linked among molecular partitions but unlinked from the morphological partition, in this way, we obtained an independent overall relative rate of molecular and morphological evolution. We applied an exponential prior (*igrvarpr = exp(10)*) for each clock model. The root of the tree was assigned a uniform prior (maximum age = 151.8 Myr, minimum age = 133 Myr) based on the estimated age of origin for angiosperms from Silvestro et al. (2015). The uncertainty of fossil ages was incorporated by using a uniform prior on the stratigraphic interval of each fossil, with minimum and maximum ages based on the current chronostratigraphic chart (Cohen et al., 2022). SFBD parameters (i.e., diversification and fossil-sampling rates) varied at 100.5, 66, and 23.03 mya (*samplestrat=divers 3: 100.5 66 23.03*). We specified an exponential distribution (*speciationpr= exp(10)*) for the speciation rate prior and a beta distribution ($\alpha = 1$, $\beta = 1$) for extinction and fossilization priors. The sampling fraction parameter (*rho*) was fixed to 0.0026 based on the ratio of the number of extant species (tips) sampled (792 species) and the total angiosperm species richness (295,283 *sensu* Christenhusz and Byng, 2016). The topology of extant species was fixed to avoid issues related to MCMC convergence.

We executed two independent runs for 40 million MCMC generations sampling every 5000 iterations. SFBD analyses were conducted in the CIPRES gateway (Miller et al., 2010) and the HPC infrastructure at the Instituto de Biología from the Universidad Nacional Autónoma de México (UNAM). Convergence was evaluated in Tracer v1.7.2 (Rambaut et al., 2018) checking that effective sample size (ESS) exceeded 200 for all parameters. Posterior estimates of FBD parameters were summarized with the *FBD_summary* function and visualized with *FBD_dens_plot* from the EVOPHYLO package (Simões et al., 2023) in R (R Core Team, 2022).

3.3.2 Constrained analysis under the skyline fossilized birth-death model (SFBD)

Our previous study focused on the phylogenetic position of a limited number of fossils (24 species) using the same morphological dataset (López-Martínez et al., 2023a), indicated some potential limitations to accurately placing fossils in the angiosperm phylogeny. This was particularly the case for fossils that do not display a distinctive or unique combination of structural traits to be placed consistently within a specific angiosperm family or order (i.e., derived clades within monocots and Pentapetalae). Therefore, we conducted another set of total evidence SFBD analyses including 792 extant species and a smaller sample of fossil species and fixing the full tree topology (see below). We used the RoguePlots resulting from the maximum likelihood analysis to select fossil taxa that met the following criteria: 1) showed low to moderate phylogenetic uncertainty (i.e., fossils attached to a single branch or several closely related branches); 2) their retrieved position (although highly uncertain) was consistent with their previously proposed systematic affinities. A total of 62 fossil flowers were selected after this procedure. The complete list of selected fossils and generated RoguePlots are available in **Supplementary Data 1 and 2**. Additionally, because our age estimations derived from the unconstrained SFBD analysis resulted in substantially young ages for families with a lack of representation of fossil flowers, we decided to incorporate 96 internal calibrations to specific nodes. The internal calibrations were based on Ramírez-Barahona et al. (2020) (**Supplementary Data 3**). We conducted two SFBD analyses varying in character sampling but fixing the topology of extant and subsampled fossils (we referred to this analysis as '*constrained SFBD*'). One analysis was performed using only morphological data ('*SFBD morphology-only*'), whereas the other included the combined matrix of molecular and morphological data ('*SFBD total evidence'*). For both analyses, we applied the same SFBD model settings as described above for the '*unconstrained SFBD*'.

3.3.3 Morphological rates of evolution

To detect significantly accelerated and decelerated rates of morphological evolution, we followed the same approach as Simões and Pierce (2021). This approach consists of comparing the variation of evolutionary (morphological) rates relative to the clock rate (background rate) calculated for every branch in the tree. Generally, values >1 indicate accelerated rates whereas <1 indicate decreasing rates of evolution. When a mean rate of evolution on a branch is greater than the background rates plus a specified threshold (e.g., one standard deviation) it is interpreted as an accelerated rate, whereas if the rate of evolution is less than the background rate minus the threshold it indicates decreasing rates. Evolutionary rate parameters were extracted from the relaxed clock analysis (consensus tip-dated tree) and visualized with the package EVOPHYLO (Simões et al., 2023) in R (R Core Team, 2022).

4 RESULTS AND DISCUSSION

4.1 Unconstrained total evidence dating analysis

Here, we first conducted a total-evidence dating analysis using the whole set of fossil flowers (108 species) without fixing their topology to any specific clade (‘*unconstrained SFBD*’). The results of this analysis show that the position of fossils differs drastically from the undated (maximum likelihood) analysis. The measured Robinson-Foulds (RF) distances (i.e., the number of bipartitions not shared by both trees) are on average high (RF mean=21.5; **see Figure S1**), indicating that almost all fossils are placed in entirely different regions in the tree. An example is the position of the fossil *Chloranthistemon endressii*, which has been consistently placed in the family Chlorantaceae based on comparative morphology (e.g., three stamens fused at the base, and each stamen bears two pairs of pollen sacs and valvate dehiscence, which are similar to extant *Chloranthus*; Crane et al., 1989) and numerous phylogenetic analyses (including those conducted in this study; **Fig. S2**) (Eklund et al., 2004; Doyle and Endress, 2018; Schönenberger et al., 2020; López-Martínez et al., 2023a). Yet surprisingly, in the unconstrained total evidence dated tree, it appeared with low uncertainty (100% of probability) as stem lineage of the order Zingiberales (RF = 22; **Supplementary Table 1**), which we consider an inaccurate result.

We expect a topological impact of the stratigraphic data associated with fossils, as the total evidence dating approach tends to improve inferences by better identifying homoplasy and obtaining fossil positions consistent with their age (Lee and Yates, 2018; King, 2021). However, we suspect that the inaccurate fossil positions obtained in this dating analysis may be due to the low number of morphological characters analyzed. A simulation study by Barido-Sottani et al. (2020) examined the impact of fossil and character sampling on the topology obtained under the FBD model and showed that when analyzing datasets with few morphological characters (e.g., 30 characters), the inferred trees tend to be far from the true tree. Similarly, using simulations, Luo et al. (2020) demonstrated that the total-evidence FBD approach has difficulties placing fossils on the tree, and increasing the numbers of morphological characters and fossils leads to better accuracy and precision of topological and divergence time estimates. However, increasing the number of characters (e.g., 300 characters as suggested by simulations) for large taxon samples (such as that used in this study), not only requires a great effort in defining and scoring applicable characters throughout the whole group, but also the analysis of such an extensive dataset would be computationally infeasible. In addition, our *unconstrained SFBD* analysis also suggests an overestimation of divergence times for clades containing fossils in unexpected positions (**Fig. S3**). This finding differs from the recent simulation study by Mongiardino Koch et al. (2023), which indicates that even when fossil taxa are incorrectly placed in the tree the divergence age estimates are accurate, however, their simulations are based on morphological datasets of 300 characters, which is unrealistic for morphological matrices like ours. Although one of the benefits of using total-evidence dating is the explicit estimation of extant and fossil positions, we strongly recommend constraining the position of fossils, for example, informed by the *RoguePlots*, when applying the SFBD model to datasets comprising high taxonomic sampling but low character sampling.

4.2 Constrained dating analyses: morphology vs total evidence

We evaluated the influence of character sampling (morphology-only vs total evidence) in the estimation of divergence times and macroevolutionary parameters. While both analyses retrieved consistent mean stem ages, they differ in the mean estimates of crown ages (**Fig. 1**). The '*SFBD morphology-only*' analysis yields mostly young mean crown ages, implying that 79% of angiosperm families diversified in the Neogene, compared to the 54% from the SFBD total evidence analysis. However, it is important to note that given that the '*SFBD morphology-only*' analysis shows wide 95% highest posterior density (HPC) intervals, the ages might completely overlap between the two analyses. The factors that may contribute to this higher uncertainty of divergence time estimates observed in the '*SFBD morphology-only*' analysis are associated with the limited number of morphological characters (as observed for molecular data; Zhu et al., 2015; Warnock et al., 2017) along with a possible high variation of morphological rates among branches and characters (Zhu et al., 2015).

We did not find differences in the posterior estimates of speciation, relative extinction, and relative fossilization rates between the two constrained FBD analyses (**Fig. S4**). Both resulted in a first burst in net diversification rate during the first time-bin (from their origin ~150 mya to 100.5 mya), but these estimates were highly variable possibly due to the small number of branches close to the root of the tree. In addition, we detect that speciation and extinction rates were the lowest for the youngest time bin (23.03 mya – present) (**Fig. S4)**. However, these results should be interpreted with caution as according to Zhang et al., 2023, based on a simulation study, the estimated diversification rates in recent intervals derived from the SFBD model under diversified sampling are sensitive to the sampling proportion of extant species (*rho*) and the taxon sampling scheme (random vs diversified).

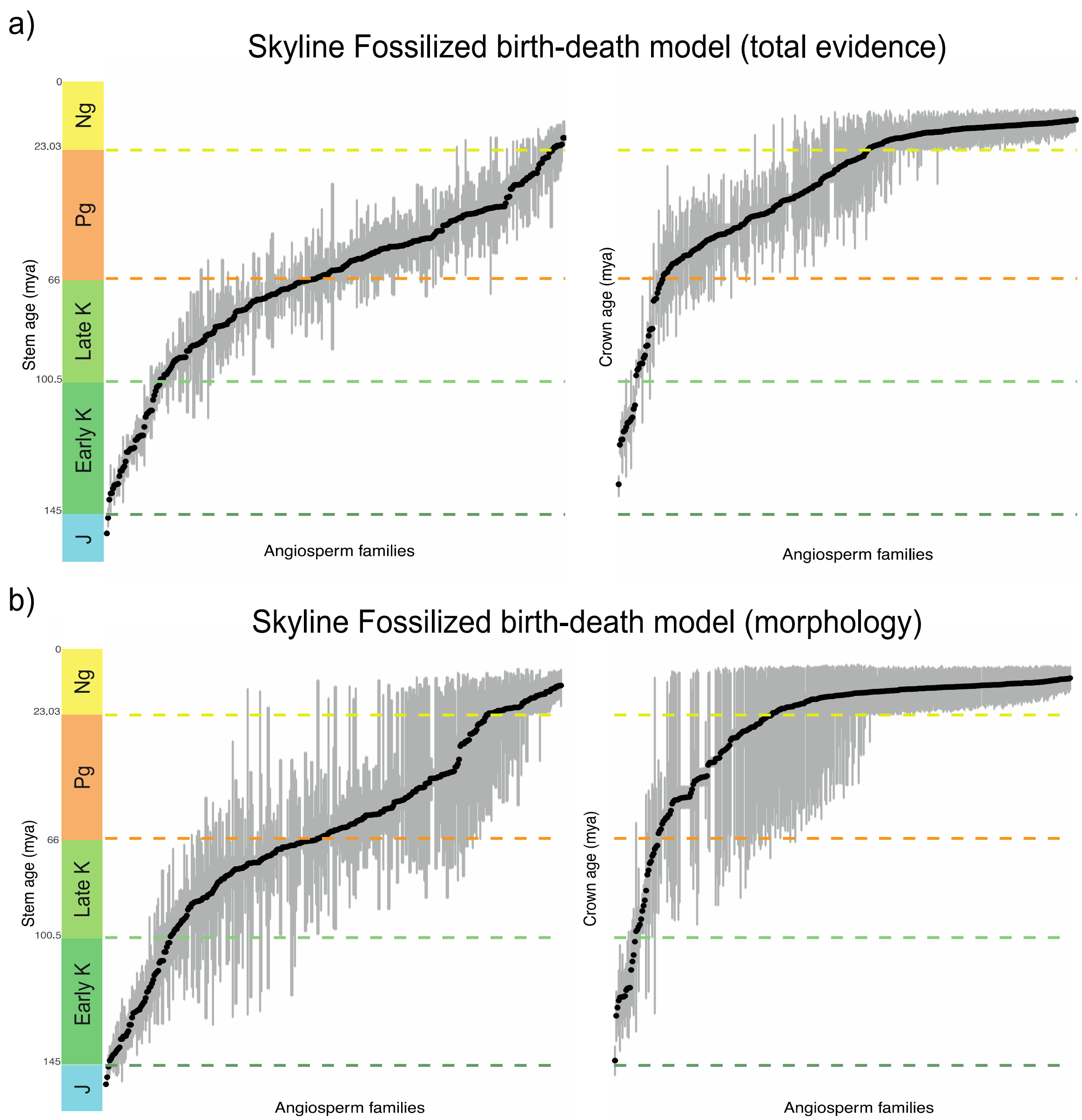

Figure 1. **Estimates of stem and crown ages of angiosperm families obtained from two analyses under the skyline fossilized birth-death model. a)** Estimated ages using the combined molecular and morphological dataset (total evidence) and **b)** estimated ages derived from the analysis of the morphological dataset. Black points indicate the mean age and gray bars correspond to 95% posterior credibility intervals. J, Jurassic; Early K, Early Cretaceous; Late K, Late Cretaceous; Pg, Paleogene; Ng, Neogene; mya, million years ago.

4.2 Total evidence dating of angiosperms

Given the high uncertainty associated with the age estimates inferred from the '*SFBD morphology-only*' analysis, we here focus on the estimates resulting from the '*SFBD total evidence*' analysis. The total-evidence dating analysis showed a stem age (i.e., divergence time from their sister lineage) of angiosperm families distributed as follows: 43 families (12%) dated from the Late Jurassic to Early Cretaceous; 113 (31%) from the Late Cretaceous, and 211 (57%) from the Cenozoic (**Fig. 1a**). In contrast, the estimated mean crown ages of most angiosperm families (331 families corresponding to 90%) were estimated in the Cenozoic (**Fig. 1a**). Detailed age estimates including the 95% highest posterior density interval of 367 families are provided in **Supplementary Table 2**.

The temporal pattern of origin and diversification of angiosperm families observed here differs slightly from those reported in recent studies using node-dating approaches (Ramírez-Barahona et al., 2020; Dimitrov et al., 2023). The study of Ramírez-Barahona et al. (2020) based on the analysis of the same molecular dataset as this study and applying 238 internal calibrations indicated a substantial time lag ("phylogenetic fuse") between the origin and diversification of angiosperm families. They suggested that although most angiosperm families originated (i.e., stem age) in the Cretaceous, their diversification (i.e., crown age) only started in the Paleocene (Ramírez-Barahona et al., 2020). This finding was also corroborated by Dimitrov et al. (2023) in their analyses of a larger dataset sampling most genera, but calibrated with the same fossils. In contrast, we obtained a higher proportion of families that originated in the Cenozoic, therefore reducing the inferred time lag. However, this result is in some way expected, due to the lack of fossil tips (or even internal node calibrations) to constrain the crown age of many angiosperm nodes. When this occurs, the resulting ages are often young under the FBD model (O'Reilly and Donoghue, 2020). Another observation is that, in a few cases, the total-evidence dating analysis retrieved older ages than node-dating for families with fossils as tips (**Fig. S5**). For instance, the mean crown age of the Lauraceae family estimated here is 134.90 mya (139.09-132.06 95% HPC), while in the node dating analysis under the relaxed calibration strategy it was 77.45 mya (52.34-106.16 95% HPC; Ramírez-Barahona et al., 2020). This could partly be due to the fossils as tips being mostly placed in crown nodes derived from our undated analysis, instead of considering the possibility they were placed as stem lineages.

Furthermore, our age estimates are quite different from those obtained by Silvestro et al. (2021), who implemented a Bayesian method (Bayesian Brownian bridge (BBB)) to infer divergence times among angiosperm families on the basis of a large dataset of meso and macrofossils occurrences and their present species richness. The results presented by Silvestro et al. (2021) suggested that the diversity of angiosperm families accumulated rapidly throughout the Cretaceous, reaching a maximum peak in the Campanian (83.6-72.1 mya), while our results indicate a much slower accumulation of families (Fig 1a). This discrepancy between age estimates could be due to the following reasons: 1) the fossils included in our SFBD analyses might not be the oldest of the clade/group; some fossil flowers provide age constraints to families that might have older records (e.g., pollen, leaves); 2) the BBB model mostly incorporated fossil occurrences to families based on original descriptions and generally without a previous explicit assessment of their systematic position. Some fossils included in our SFBD analyses are placed within different families from those originally described (e.g., the fossil *Microvictoria svitkoana* was originally described in Nymphaeaceae (Nymphaeales; Gandolfo, 2004), and here is placed as stem Degeneriaceae; Magnoliales). This is particularly problematic because the BBB method is sensitive to the earliest fossil record of each lineage and the inaccurate position of fossils in families with poor representation of fossils (Silvestro et al., 2021).

Nevertheless, we recognize that there are additional factors that would influence our divergence times estimates. Luo et al. (2023) demonstrated that the divergence times and evolutionary rates under the FBD model were strongly influenced by the root age prior. In our case, choosing an adequate prior for root age is not straightforward since the range of estimates for the crown group of angiosperms varies significantly (see Sauquet et al., 2022). Thus, the use of alternative root age priors would lead to variations in age estimates. Another factor that would lead to bias is the fossil sampling. Under the diversified sampling scheme there is a cutoff time – time at which the model assumes no fossil sampling and no rate shifting between the cutoff time and the present (Zhang et al., 2016; 2023). Previous empirical analyses have shown that the cutoff time was adjusted to be right after the youngest fossil (Zhang et al., 2016), therefore including younger fossils in the analysis might change the divergence times from the cutoff time. Moreover, further work is needed to determine the influence of the discretization of time intervals in the estimation of divergence times under the skyline FBD model.

4.3 Morphological rates of angiosperms

Our '*SFBD total evidence*' analysis indicates that the morphological rate of floral evolution is broadly uniform (stable) across the angiosperm phylogeny (**Fig. 2**). The low proportion of significant shifts in morphological rates (either low or high) is consistent with the patterns of phenotypic evolution as revealed by many animal fossils (e.g., foraminifera, mollusks, trilobites, teleost fish; Hunt, 2007). Here, we detected the highest rates of morphological evolution in the Early Cretaceous. These high rates are congruent with the high morphological diversity (disparity) estimated with fossil flowers for this period (López-Martínez et al., 2023b). Although there are also exceptions where lineages show low morphological rates and high disparity, for example by having more species or being older (see Morinaga et al., 2023), the correlation between high disparity and high morphological rate is also common for various animal groups (Morinaga et al., 2023; Yu et al., 2023). Additionally, high morphological change during a brief period of time is usually coupled with high diversification rates (Simões and Pierce, 2021). Our analysis indicated a high diversification rate for the Early Cretaceous (**Fig. S4**), however, we could not confirm whether such high morphological rates produced more species because our estimated diversification rates under the SFBD might be highly sensitive to bias as discussed above.

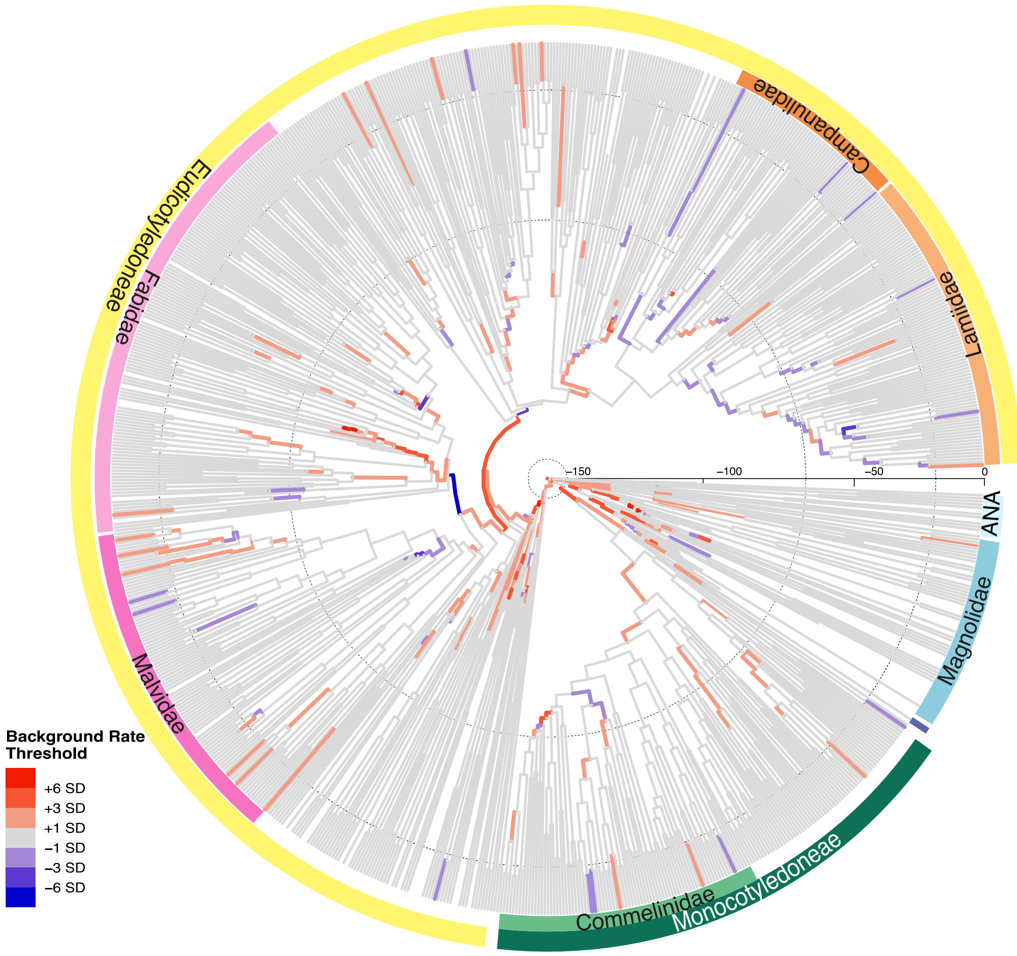

Figure 2. **Relative rates of morphological evolution through time and across major angiosperm clades.** Relative rates estimated under the skyline fossilized birth-death model with total-evidence dataset. Blue (cold) colors correspond to slow rates and red (warm) colors denote fast rates of morphological evolution.

Across the angiosperm tree, we detected the highest relative morphological rates along the deeper branches leading up to major angiosperm clades (e.g., Pentapetalae, Magnolidae, Monocotyledoneae, Rosidae, and Asteridae) (**Fig. 2**). These substantially accelerated rates are observed early in the initial acquisition and differentiation of major floral Bauplan among angiosperms. For example, the increasing rate in the branch subtending the Pentapetalae clade, which is characterized by a predominance of pentamerous flowers with differentiation of perianth into calyx and corolla, or in several branches within the Magnolidae clade that shows high variability in floral phyllotaxis and number of parts and less differentiation of perianth organs (Endress, 2011). Additionally, we identified increased morphological rates in several branches within clades characterized by an extensive occurrence of fossil flowers, for instance, magnoliids (e.g., Magnoliales and Laurales), the "early-diverging eudicots" grade (e.g., Ranunculales and Proteales), and the order Fagales. This is expected given the observed high morphological variation in Cretaceous fossils, even between sister lineages (Friis et al., 2011), and the relatively short time lag leading to such variation. However, future work is needed to examine whether the sparse sampling of fossil flowers in the Cenozoic might affect the lack of rate heterogeneity observed in that period.

Interestingly, we detected relatively few shifts to decreases in morphological rates across the tree (**Fig. 2**). Those slow rates mainly occurred in the Late Cretaceous along the branches within the clade Asteridae (Lamiidae and Campanulidae), whose members tend to have low variation in floral structure and are characterized by a stabilization of floral organization (e.g., floral characters became canalized or fixed; Endress, 2011; López-Martínez et al., 2023b). These findings are also congruent with the hypothesis that morphological rate slowdowns can occur if the morphospace becomes saturated over time (Rabosky et al., 2013), as suggested by our previous morphospace analysis showing that Lamiidae and Campanulidae occupy a restricted and overlapping area of the space (López-Martínez et al., 2023b).

Heritable shifts are characterized by either a decrease or increase in morphological rate relative to close relatives (Lloyd et al. 2012), in other words, a rate shift is considered heritable if the rate is retained in all descended taxa (i.e., all members of a specific clade/group). Here, we did not observe extremely low or high morphological rates when comparing among clades or grades (**Fig. S6**). This result is in some way expected as we are comparing an overall morphological rate encompassing characters of the entire floral bauplan, and the rates are heterogeneous among floral modules (Herting et al., 2023).

Some clades might display higher rates for the characters comprising the androecium (e.g., the order Ericales; Herting et al., 2023), while others show higher rates associated with the gynoecium (Magnoliidae; Herting et al., 2023). However, future work implementing partitioned FBD analyses would be helpful to examine the rate of morphological evolution for distinct floral modules (e.g., perianth, androecium, gynoecium) among clades and time periods.

5 CONCLUSIONS

Our age estimates using the total-evidence approach under the skyline fossilized birth-death model for a large sample of extant and fossil taxa are considerably younger than previous estimates using node-dating approaches. However, our age estimates were highly sensitive to variations in character sampling (total-evidence vs morphology-only) and fossil sampling. Therefore, our divergence time estimation would benefit from increased sampling in two directions: taxa and characters. More taxa (fossils and extant species), that is, more tips in the phylogenetic tree can help us to increase the precision of branch-length estimates and parameters of the FBD (e.g., speciation and extinction rates; Zhang et al., 2016). While more characters might improve the accuracy in the estimation of fossil positions, for that reason, the collection and accessibility of new morphological data either at low or high phylogenetic levels should be a continued task. Moreover, reliable estimates based on the total evidence approach would require further improvements in modeling morphological character evolution. Variations of the simplistic Mk model (Lewis, 2001) that would adequately capture the complexity of morphological data (see Mulvey et al., 2024) should be considered and implemented in FBD analyses.

Although previous simulation studies have improved our understanding of the potential bias and limitations of the FBD model, the simulation scenarios are far from realistic and applicable to empirical datasets such as ours. Morphological datasets available for a large sample of fossil and extant species are subject to large amounts of missing data, particularly here the fossil record of angiosperm flowers is temporally biased since most of the fossils were sampled from the Cretaceous. Future simulation studies should investigate the influence of those factors more extensively, as well as diversification and fossilization rates informed by real data (Luo et al., 2023).

Our study reveals that significantly high rates of morphological evolution across the angiosperm tree occurred in the Early Cretaceous and subtended the origin of major angiosperm clades. This pattern is consistent with the high variation in morphological disparity quantified by the fossil record of angiosperm flowers alone. A distinct trajectory corresponds to shifts towards a stabilizing (low) morphological rate in the origin of nested clades (Lamiidae and Campanulidae) that eventually occupied more restricted areas of the morphospace (López-Martínez et al., 2023b). The distribution of evolutionary change and their potential correlation with diversification rate shifts, as well as the degree to which the morphological rates are affected by the root age, or taxon-sampling schemes, need further investigation.

6 ACKNOWLEDGMENTS

We are grateful to the University of Vienna for funding the eFLOWER server hosting the PROTEUS database; the Instituto de Biología and Dr. Cristian Cervantes for the access to the HPC infrastructure where the unconstrained total-evidence analysis was conducted. We thank Dr. Santiago Ramírez-Barahona, Dr. Alfonso Delgado-Salinas, Dr. Luis Eguiarte, and M.Sc. Adriana Benítez-Villaseñor for providing helpful comments and feedback on this project. We thank the Posgrado en Ciencias Biológicas (UNAM) and the Consejo Nacional de Humanidades Ciencias y Tecnologías (CONAHCYT) for the doctoral scholarship 708068.

7 REFERENCES

1. APG IV. 2016 An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181, 1–20.
2. Asar Y, Sauquet H, Ho SYW. 2023 Evaluating the accuracy of methods for detecting correlated rates of molecular and morphological evolution. Syst. Biol. 72, 1337–1356.

3. Barido-Sottani J, van Tiel NMA, Hopkins MJ, Wright DF, Stadler T, Warnock RCM. 2020 Ignoring Fossil Age Uncertainty Leads to Inaccurate Topology and Divergence Time Estimates in Time Calibrated Tree Inference. Front. Ecol. Evol. 8, 183.
4. Borcklehurst N, Benson RJ. 2021. Multiple paths to morphological diversification during the origin of amniotes. Nat. Ecol. Evol. 5, 1243-1249.
5. Christenhusz MJ, Byng JW. 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261, 201–217.
6. Cohen KM, Finney SC, Gibbard PL, Fan JX. 2013; Updated the ICS International Chronostratigraphic Chart. Episodes 36, 199-204. URL https://stratigraphy.org/chart.
7. Crane PR, Friis EM, Pedersen KR. 1989 Reproductive structure and function in Cretaceous Chloranthaceae. Plant Syst. Evol. 165, 211–226.
8. Coiro M, Allio R, Mazet N, Seyfullah LJ, Condamine FL. 2023 Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity. New Phytol. 240, 1616–1635.
9. Dimitrov D, Xu X, Su X, Shrestha N, Liu Y, Kennedy JD, Lyu L, Nogués-Bravo D, Rosindell J, Yang Y, Fjeldså J, Liu J, Schmid B, Fang J, Rahbek C, Wang Z. 2023 Diversification of flowering plants in space and time. Nat Commun. 14, 7609.
10. Doyle J.A., Endress P.K. 2010. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. J. Syst. Evol. 48, 1–35.
11. Doyle J.A., Endress P.K. 2014. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: ANITA lines and relatives of Chloranthaceae. Int. J. Plant Sci. 175, 555–600.
12. Doyle J.A., Endress P.K. 2018 Phylogenetic analyses of Cretaceous fossils related to Chloranthaceae and their evolutionary implications. Bot. Rev. 84, 156–202.
13. Endress P. 2011 Evolutionary diversification of the flowers in angiosperms. Am. J. Bot. 98, 370–396.
14. Foote M. 1994 Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. Paleobiology 20, 320–344.

15. Friis EM, Crane PR, Pedersen KR. 2011 Early flowers and angiosperm evolution. Cambridge: Cambridge University Press.
16. Friis EM, Pedersen KR. 2011 *Canrightia resinifera* gen. et sp. nov., a new extinct angiosperm with Retimonocolpites-type pollen from the Early Cretaceous of Portugal: missing link in the eumagnoliid tree? Grana 50, 3–29.
17. Friis EM, Pedersen KR. 2012 *Bertilanthus scanicus*, a new asterid flower from the Late Cretaceous (Late Santonian–Early Campanian) of Scania, Sweden. Int. J. Plant Sci. 173, 318–330.
18. Gandolfo MA, Nixon KC, Crepet WL. 2004 Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. Proc. Natl. Acad. Sci. U.S.A. 101, 8056–8060.
19. Gavryushkina A, Welch D, Stadler T, Drummond A. 2014 Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. PLoS Comput. Biol. 10, e1003919.
20. Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D, Drummond AJ. 2017 Bayesian total-evidence dating reveals the recent crown radiation of penguins. Syst. Biol. 66, 57–73.
21. Gould SJ. 1989 Wonderful life. New York, NY, USA: W. W. Norton.
22. Heath TA, Huelsenbeck JP, Stadler T. 2014 The fossilized birth–death process for coherent calibration of divergence-time estimates. Proc. Natl Acad. Sci. USA. 111, E2957–E2966.
23. Herendeen PS, Friis EM, Pedersen KR, Crane PR. 2017 Palaeobotanical redux: revisiting the age of the angiosperms. Nat. Plants 3, 17015.
24. Herting J, Schönenberger J, Sauquet H. 2023 Profile of a flower: How rates of morphological evolution drive floral diversification in Ericales and angiosperms. Am. J. Bot. 110, e16213.
25. Hunt, G. 2007 The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. Proc. Natl Acad. Sci. USA. 104, 18404–18408.

26. King B. 2021 Bayesian tip-dated phylogenetics in paleontology: topological effects and stratigraphic fit. Syst. Biol. 70, 283–294.
27. Klopfstein S, Spasojevic T. 2019 Illustrating phylogenetic placement of fossils using RoguePlots: an example from ichneumonid parasitoid wasps (Hymenoptera, Ichneumonidae) and an extensive morphological matrix. PLoS One 14, e0212942.
28. Jablonski D, Raup DM. 1995 Selectivity of end-Cretaceous marine bivalve extinctions. Science 268, 389–391.
29. Lee MSY, Yates AM. 2018. Tip-dating and homoplasy: reconciling the shallow molecular divergences of modern gharials with their long fossil record. Proc. R. Soc. B. 285, 20181071.
30. Leslie AB, Simpson C, Mander L. 2021 Reproductive innovations and pulsed rise in plant complexity. Science 373, 1368–1372.
31. Lewis PO. 2001 A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.
32. Lloyd GT, Wang SC, Brusatte SL. 2012 Identifying heterogeneity in rates of morphological evolution: discrete character change in the evolution of lungfish (Sarcopterygii; Dipnoi). Evolution 66, 330–348.
33. López-Martinez AM, Schönenberger J, von Balthazar M, Gonzalez-Martinez CA, Ramirez-Barahona S, Sauquet H, Magallon S. 2023a. Integrating fossil flowers into the angiosperm phylogeny using molecular and morphological evidence. Syst Biol. 72, 837–855.
34. López-Martínez AM, Magallón S, von Balthazar, M Schönenberger J, Sauquet H, Chartier M. 2023b Angiosperm flowers reached their highest morphological diversity early in their evolutionary history. New Phytol. 241, 1348–1360.
35. Luo A, Duchêne DA, Zhang C, Zhu CD, Ho S. 2020 A simulation-based evaluation of tip-dating under the fossilized birth-death process. Syst Biol. 69, 325–344.
36. Luo A, Zhang C, Zhou QS, Ho SYW, Zhu CD. 2023 Impacts of taxon-sampling schemes on Bayesian tip dating under the fossilized birth–death process. Syst. Biol. 71, 781–801.

37. Lupia R. 1999 Discordant morphological disparity and taxonomic diversity during the Cretaceous angiosperm radiation: North American pollen record. Paleobiology 25, 1–28.
38. Magallón S, Herendeen PS, Crane PR. 2001 *Androdecidua endressii* gen. et sp. nov., from the Late Cretaceous of Georgia (United States): further floral diversity in Hamamelidoideae (Hamamelidaceae). Int. J. Plant Sci. 162, 963–983.
39. Mulvey LPA, May M, Brown JM, Höhna S, Wright AM, Warnock RCM. 2024 Assessing the adequacy of morphological models used in Palaeobiology. bioRxiv 2024.01.25.577179.
40. Marshall CR. 2017 Five palaeobiological laws needed to understand the evolution of the living biota. Nat. Ecol. Evol. 1, 1–6.
41. May MR, Contreras DL, Sundue MA, Nagalingum NS, Looy CV, Rothfels CJ. 2021 Inferring the total-evidence timescale of marattialean fern evolution in the face of model sensitivity. Syst. Biol. 70, 1231–1255.
42. Miller MA, Pfeiffer W, Schwartz T. 2010 Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gatew. Comput. Environ. Work. GCE. 14, 1–8.
43. Mitchell JS, Etienne RS, Rabosky DL. 2019 Inferring diversification rate variation from phylogenies with fossils. Syst. Biol. 68, 1–18.
44. Mongiardino Koch N, Garwood RJ, Parry LA. 2023 Inaccurate fossil placement does not compromise tip-dated divergence times. Palaeontology 66, e12680.
45. Morinaga G, Wiens JJ, Moen DS. 2023 The radiation continuum and the evolution of frog diversity. Nat. Commun. 14, 7100.
46. O'Meara BC, Smith SD, Armbruster WS, Harder LD, Hardy CR, Hileman LC, Hufford L, Litt A, Magallón S, Smith SA, Stevens PF, Fenster CB, Diggle PK. 2016 Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. Proc. R. Soc. B 283: 20152304.
47. O'Reilly JE, Donoghue PCJ. 2020 The effect of fossil sampling on the estimation of divergence times with the fossilized birth–death process. Syst. Biol. 69, 124–138.

48. Quental TB, Marshall CR. 2010 Diversity dynamics: molecular phylogenies need the fossil record. Trends Ecol. Evol. 25, 434–441.
49. R Core Team. 2022 R: a language and environment for statistical computing, v4.1.3. Vienna, Austria: R foundation for Statistical Computing. URL http://www.r-project.org.
50. Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, Alfaro ME. 2013 Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. 4, 1958.
51. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018 Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904.
52. Ramírez-Barahona S, Sauquet H, Magallon S. 2020 The delayed and geographically heterogeneous diversification of flowering plant families. Nat. Ecol. Evol. 4, 1232–1238.
53. Raup DM, Sepkoski JJ Jr. 1982. Mass extinctions in the marine fossil record. Science 215, 1501–3.
54. Reyes E, Nadot S, von Balthazar M, Schönenberger J, Sauquet H. 2018 Testing the impact of morphological rate heterogeneity on ancestral state reconstruction of five floral traits in angiosperms. Sci. Rep. 8, 9473.
55. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP 2012. Mrbayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
56. Ronquist F, Lartillot N, Phillips MJ. 2016 Closing the gap between rocks and clocks using total-evidence dating. Philos. Trans. R. Soc. B Biol. Sci. 371, 20150136.
57. Sauquet H, Von Balthazar M, Magallón S, Doyle JA, Endress PK, Bailes EJ, Barroso De Morais E, Bull-Hereñu K, Carrive L, Chartier M, Chomicki G, Coiro M, Cornette R, El Ottra JHL, Epicoco C, Foster CSP, Jabbour F, Haevermans A, Haevermans T, Hernández R, Little SA, Löfstrand S, Luna JA, Massoni J, Nadot S, Pamperl S, Prieu C, Reyes E, Dos Santos P, Schoonderwoerd KM, Sontag S, Soulebeau A, Staedler Y, Tschan GF, Wing-Sze Leung A, Schönenberger J. 2017 The ancestral flower of angiosperms and its early diversification. Nat. Commun. 8, 16047.

58. Sauquet H. 2019 PROTEUS: a database for recording morphological data and calibrations. v 1.27. URL https://eflower.myspecies.info/proteus
59. Sauquet H, Ramírez-Barahona S, Magallón S. 2022 What is the age of flowering plants? J. Exp. Bot. 73, 3840–3853.
60. Schönenberger J, Friis EM. 2001. Fossil flowers of ericalean affinity from the Late Cretaceous of Southern Sweden. Am. J. Bot. 88, 467–480.
61. Schönenberger J, von Balthazar M, Lopez Martinez A, Albert B, Prieu C, Magallon S, Sauquet H. 2020 Phylogenetic analysis of fossil flowers using an angiosperm-wide data set: proof-of-concept and challenges ahead. Am. J. Bot. 107, 1433–1448.
62. Silvestro D, Cascales-Miñana B, Bacon CD, Antonelli A. 2015 Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. New Phytol. 207, 425–436.
63. Silvestro D, Bacon CD, Ding W, Zhang Q, Donoghue PCJ, Antonelli A, Xing Y. 2021 Fossil data support a pre-Cretaceous origin of flowering plants. Nat. Ecol. Evol. 5, 449–457.
64. Simões TR, Pierce SE. 2021 Sustained high rates of morphological evolution during the rise of Tetrapods. Nat. Ecol. Evol. 5, 1403–1414.
65. Simões TR, Greifer N, Barido-Sottani J, Pierce, SE. 2023. EvoPhylo: An R package for pre- and postprocessing of morphological data from relaxed clock Bayesian phylogenetics. Methods Ecol. Evol. 14, 1981–1993.
66. Sims HJ. 2013 Morphological rates of angiosperm seed size evolution. Evolution 57, 1338–1346.
67. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
68. von Balthazar M, Pedersen KR, Crane PR, Friis EM. 2008 *Carpestella lacunata* gen. et sp. nov., a new basal angiosperm flower from the Early Cretaceous (Early to Middle Albian) of eastern North America. Int. J. Plant Sci. 169, 890–898.
69. Warnock RCM, Yang Z, Donoghue P. 2017 Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. Proc. R. Soc. B. 284, 20170227.

70. Yu Y, Zhang C, Xu X. 2023. Complex macroevolution of pterosaurs. Curr. Biol. 23, 770–779.
71. Zhang C, Stadler T, Klopfstein S, Heath TA, Ronquist F. 2016. Total- evidence dating under the fossilized birth–death process. Syst. Biol. 65, 228–249.
72. Zhang C, Ronquist F, Stadler T. 2023 Skyline Fossilized Birth–Death Model is Robust to Violations of Sampling Assumptions in Total-Evidence Dating. Syst Biol., syad054.
73. Zhang Q, Ree RH, Salamin N, Xing Y, Silvestro D. 2022 Fossil-informed models reveal a Boreotropical origin and divergent evolutionary trajectories in the walnut family (Juglandaceae). Syst Biol. 71, 242–258.
74. Zhu T, dos Reis M, Yang Z. 2015. Characterization of the uncertainty of divergence time estimation under relaxed molecular clock models using multiple loci. Syst. Biol. 64, 267–280.

SUPPLEMENTARY MATERIAL

*Exploring the total evidence dating under the skyline fossilized birth-death model applied to angiosperms**

*Se presentan resultados preliminares. Esta investigación continua en proceso y está siendo realizada en colaboración con **Dr. Hervé Sauquet** (National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, Australia) y **Dra. Maria von Balthazar** y **Dr. Jürg Schönenberger** (Department of Botany and Biodiversity Research, University of Vienna, Viena, Austria).

	Page
Supplementary Data 1. Detailed information about fossils	2
Supplementary Data 2. RoguePlots derived from the undated analysis	*
Supplementary Data 3. Additional calibrations for internal nodes (Node-Dating)	9
Supplementary Table 1. Detailed age estimates under the constrained skyline fossilized birth-death (SFBD) total evidence approach	13
Figure S1. Robinson-Foulds distances	28
Figure S2. Maximum likelihood tree	29
Figure S3. Comparison between unconstrained vs constrained age estimates	30
Figure S4. FBD parameters for two constrained analyses	31
Figure S5. Comparison of the total evidence under the constrained SFBD model vs node dating approaches	32
Figure S6. Morphological rates among major groups	33

* Supplementary data 2 can be found in:

https://drive.google.com/drive/folders/1J-tbGw1n-PALyJSwEm-mZz7RaF9cp8VL

Supplementary Data 1. Detailed information about fossils used in this study, including stratigraphic age, locality, and references associated. * Fossils included in the constrained analyses (rows highlighted in gray).

Fossil taxa	**Stratigraphic Age**	**Mode of Preservation**	**Mean Age**	**Min Age**	**Max Age**	**Locality**	**Reference**
Acacia eocaribbeanensis	Late Eocene	Amber	37.55	33.9	41.2	Amber deposit from the Palo Alto mine, Dominican Republic	Dilcher and Herendeen, 1992
Actinocalyx bohrii	Late Santonian to Early Campanian	Charcoalified	81.4	77.85	84.95	Höganäs AB quarry at Åsen, Scania, Sweden	Friis, 1985
*Allonia decandra**	Late Santonian to Early Campanian	Charcoalified	81.4	77.85	84.95	Roberta, Georgia, USA	Magallón-Puebla, 1996
*Androdecidua endressii**	Late Santonian	Charcoalified	84.275	83.6	84.95	Allon locality, Georgia, USA	Magallón et al. 2001
Antiquacupula sulcata	Late Santonian	Charcoalified	84.275	83.6	84.95	Gaillard locality, Georgia, USA	Sims et al., 1998
*Antiquocarya verruculosa**	Late Santonian to Early Campanian	Charcoalified	81.4	77.85	84.95	Åsen, Scania, Sweden	Friis et al. 2006
Araripia florifera	Aptian	Impression / Compression	117.2	113	121.4	Nova Olinda, Araripe Basin, Brazil	Mohr and Eklund, 2003
*Archaeanthus linnenbergeri**	Albian to Middle Cenomanian	Impression / Compression	105.1	97.2	113	Linnenberger's Ranch near Bunker Hill, Russel County, Kansas, USA	Dilcher and Crane, 1984
*Archaefagacea futabensis**	Early Coniacian	Charcoalified	88.925	88.05	89.8	Kitaba River, Kamikitaba, Japan	Takahashi et al. 2008
Archaefructus eoflora	Hautevarian	Impression / Compression	131	129.4	132.6	Beipiao, Western Liaoning, China	Qiang et al. 2004
*Archaestella verticillatus**	Early Coniacian	Charcoalified	88.925	88.05	89.8	Kitaba River, Kamikitaba, Japan	Takahashi et al. 2017
Archamamelis bivalvis	Late Santonian to Early Campanian	Charcoalified	81.4	77.85	84.95	Höganäs AB quarry at Åsen, Scania, Sweden	Endress and Friis, 1991
*Austrodiospyros cryptostoma**	Late Eocene	Amber	37.55	33.9	41.2	Anglesea, Victoria, Australia	Basinger, 1985
*Barnebyanthus buchananensis**	Bartonian	Impression / Compression	39.5	37.8	41.2	Buchanan Clay Pit, Tennessee, USA	Crepet and Herendeen, 1992

*Bedellia pusilla**	Late Santonian	Charcoalified	84.275	83.6	84.95	Gaillard, Crawford County, Georgia, USA.	Sims et al. 1999
*Bertilanthus scanicus**	Early Campanian	Charcoalified	80.725	77.85	83.6	Höganäs AB quarry at Åsen, Scania, Sweden	Friis and Pedersen, 2012
Budvaricarpus serialis	Late Turonian to Santonian	Charcoalified	87.725	83.6	91.85	Branisov, Czech Republic	Heřmanová et al. 2011
Caliciflora mauldiniensis	Early Cenomanian	Charcoalified	98.85	97.2	100.5	Mauldin Mountain, Elk Neck Peninsula, Maryland, USA	Friis et al. 2016
*Callicrypta chlamydea**	Cenomanian	Impression / Compression	97.2	93.9	100.5	Tyung River, eastern Siberia	Krassilov and Golovneva, 2004
*Canrightia resinifera**	Late Barremian to Aptian	Charcoalified	116.375	106.75	126	Catefica, Portugal	Friis and Pedersen, 2011
*Carpestella lacunata**	Early to Middle Albian	Charcoalified	109.875	106.75	113	Puddledock locality, Virginia, USA	von Balthazar et al. 2008
Caryanthus knoblochii	Late Santonian	Charcoalified	84.275	83.6	84.95	Gaillard, Crawford County, Georgia, USA.	Sims et al. 1999
Castaneoidea puryearensis	Bartonian	Impression / Compression	39.5	37.8	41.2	Puryear, Tennessee, USA	Crepet and Daghlian, 1980
Castanopsis kaulli	Late Eocene	Amber	37.55	33.9	41.2	Samland, Kaliningrad	Sadowski et al. 2018
*Catanthus dolichostemon**	Late Barremian to Albian	Charcoalified	116.375	106.75	126	Catefica, Portugal	Friis et al. 2021
*Cecilanthus polymerus**	Early Cenomanian	Charcoalified	98.85	97.2	100.5	Elk Neck Peninsula, Maryland, USA	Herendeen et al. 2016
Cedrelospermum nervosum	Middle Eocene	Impression / Compression	39.5	37.8	41.2	Colorado and Utah, USA	Manchester, 1989
Chloranthistemon alatus	Late Santonian to Early Campanian	Charcoalified	81.4	77.85	84.95	Höganäs AB quarry at Åsen, Scania, Sweden	Eklund et al. 1997
Chloranthistemon crossmanensis	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Herendeen et al. 1993
*Chloranthistemon endressii**	Early Campanian	Charcoalified, Lignitized	81.4	77.85	84.95	Höganäs AB quarry at Åsen, Scania, Sweden	Crane et al. 1989
Coahuilanthus belindae	Campanian	Impression / Compression	77.85	72.1	83.6	El Almacigo, Coahuila, México	Calvillo-Canadell and Cevallos-Ferriz, 2007

*Cohongarootonia hispida**	Early to Middle Albian	Charcoalified	109.875	106.7 5	113	Puddledock locality, Virginia, USA	von Balthazar et al. 2011
Cronquistiflora sayrevillensis	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet and Nixon, 1998
Cunoniantha bicarpellata	Danian	Impression / Compression	63.8	61.6	66	Palacio de los Loros-2, Argentina	Jud and Gandolfo, 2021
*Dakotanthus cordiformis**	Late Albian to Cenomanian	Impression / Compression	103.45	93.9	113	Ottawa County, Kansas, USA	Manchester et al. 2018
*Detrusandra mystagoga**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet and Nixon, 1998
*Divisestylus brevistamineus**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Hermsen et al. 2003
Dressiantha bicarpellata	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Gandolfo et al. 1998
Eckhartianthus lusitanicus	Late Barremian to Early Aptian	Charcoalified	121.3	117.2	125.4	Torres Vedras, Portugal	Friis et al. 2019
*Endressianthus miraensis**	Campanian to Maastrichtian	Charcoalified	74.8	66	83.6	Mira locality, Portugal	Friis et al. 2003
*Endressinia brasiliana**	Aptian	Impression / Compression	117.2	113	121.4	Nova Olinda, Araripe Basin, Brazil	Mohr and Bernardes-de-oliveira, 2004
Eoglandulosa warmanensis	Middle to Late Eocene	Impression /Compression	37.55	33.9	41.20	Northwestern Tennesse, USA	Taylor and Crepet, 1987
Eomimosoidea plumosa	Middle Eocene	Impression / Compression	37.55	33.9	41.2	Western, Tennesse, USA	Daghlian et al. 1980
Esgueiria adenocarpa	Campanian to Maastrichtian	Impression / Compression	74.8	66	83.6	Esguira, Portugal	Friis et al. 1992
Esgueiria futabensis	Early Coniacian	Charcoalified	88.925	88.05	89.8	Kitaba River, Kamikitaba, Japan	Takahashi et al. 1999
Florissantia ashwillii	Late Eocene to Early Oligocene	Impression / Compression	34.51	27.82	41.2	Sheep Rock Creek, Oregon, USA	Manchester, 1992
Florissantia quilchenensis	Late Eocene	Impression / Compression	39.5	37.8	41.2	Quilchena, southern of British Columbia, Canada	Manchester, 1992
Futabanthus asamigawaensis	Early Coniacian	Charcoalified	88.925	88.05	89.8	Kitaba River, Kamikitaba, Japan	Takahashi et al. 2008

*Glandulocalyx upatoiensis**	Early to Middle Santonian	Charcoalified	85.625	84.95	86.3	Upatoi Creek, Georgia, USA	Schonenberger et al. 2012
*Hamatia elkneckensis**	Late Albian	Charcoalified	103.625	100.5	106.75	Bull Mountain, Elk Neck Peninsula, Maryland, USA	Pedersen et al. 1994
Hymenaea protera	Late Eocene	Amber	37.55	33.9	41.2	La Toca mine, Dominican Republic	Poinar, 1991
Icacinanthium tainiaphorum	Ypresian	Amber	51.9	47.8	56	Le Quesnoy, France	Del Rio et al. 2017
*Jamesrosea burmensis**	Late Albian to Early Cenomanian	Amber	101.975	97.2	106.75	Myanmar	Crepet et al. 2016
*Jerseyanthus calycanthoides**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet et al. 2005
*Kajanthus lusitanicus**	Late Aptian to Early Albian	Charcoalified	112.875	106.7 5	119	Chicalhão, Portugal	Mendes et al. 2014
Kenilanthus marylandensis	Early to Middle Albian	Charcoalified	109.875	106.7 5	113	Kenialworth, Washington, USA	Friis et al. 2017
*Lacinipetalum spectabilum**	Danian	Impression / Compression	63.8	61.6	66	Palacio de los Loros-2, Argentina	Jud et al. 2018
*Lauranthus futabensis**	Early Coniacian	Charcoalified	88.925	88.05	89.8	Kitaba River, Kamikitaba, Japan	Takahashi et al. 2001
Lovellea wintonensis	Late Albian	Impression / Compression	103.625	100.5	106.75	Lovelle Downs, Queensland, Australia	Dettmann et al., 2009
Mabelia archaia	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Gandolfo et al. 2002
Mabelia connatifila	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Gandolfo et al. 2002
Manningia crassa	Early Campanian	Charcoalified	81.4	77.85	84.95	Höganäs AB quarry at Åsen, Scania, Sweden	Friis, 1983
*Mauldinia angustiloba**	Late Albian to Cenomanian	Charcoalified	101.975	97.2	106.75	Rhenish Massif, Germany	Viehofen et al., 2008
*Mauldinia bohemica**	Cenomanian	Charcoalified	97.2	93.9	100.5	Pecinov, Bohemia, Czech Republic	Eklund and Kvacek, 1998
*Mauldinia hirsuta**	Cenomanian to Early Turonian	Charcoalified	95.15	89.8	100.5	Sarbay Quarry, Kazakhstan	Frumin et al., 2004

*Mauldinia mirabilis**	Early Cenomanian	Charcoalified	98.85	97.2	100.5	Mauldin Mountain, Elk Neck Peninsula, Maryland, USA	Drinnan et al. 1990
*Melloniflora virginiensis**	Early to Middle Albian	Charcoalified	109.875	106.75	113	Puddledock locality, Virginia, USA	Friis et al. 2020
Microaltingia apocarpela	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Zhou, 2001
*Microlaurus perigynus**	Early Coniacian	Charcoalified	88.925	88.05	89.8	Kitaba River, Kamikitaba, Japan	Takahashi et al. 2014
*Microvictoria svitkoana**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Gandolfo et al. 2004
*Miranthus elegans**	Campanian to Maastrichtian	Charcoalified	74.8	66	83.6	Mira locality, Portugal	Friis et al. 2021
*Monetianthus mirus**	Late Aptian to Early Albian	Charcoalified	111.975	106.75	117.2	Vale de Agua, Portugal	Friis et al. 2009
Nahinda axamilpensis	Oligocene	Impression / Compression	28.465	23.03	33.9	Los Ahuehuetes, Puebla, México	Calvillo-Canadell and Cevallos-Ferriz, 2007
*Neusenia tetrasporangiata**	Santonian to Campanian	Charcoalified	79.2	72.1	86.3	Neuse River, North Carolina, USA	Eklund, 2000
Noferinia fusicarpa	Late Santonian	Charcoalified	84.275	83.6	84.95	Gaillard, Crawford County, Georgia, USA.	Lupia et al. 2002
Normanthus miraensis	Campanian to Maastrichtian	Charcoalified	74.8	66	83.6	Mira locality, Portugal	Schönenberger et al. 2001
Notiantha grandensis	Paleocene	Impression / Compression	63.8	61.6	66	Rancho Grande, Chubut, Argentina	Jud et al., 2017
Nuhliantha nyanzaiana	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Gandolfo et al. 2002
Paisia pantoporata	Late Barremian to Albian	Charcoalified	116.375	106.75	126	Catefica, Portugal	Friis et al. 2017
*Paleoclusia chevalieri**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet and Nixon, 1998
*Paleoenkianthus sayrevillensis**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet et al. 2018
Paleorosa similkameenensis	Middle Eocene	Permineralized	37.55	33.9	41.2	Princeton chert of British Columbia, Canada	Cevalloz-Ferriz et al. 1993

Paradinandra suecica	Early Campanian	Charcoalified	81.4	77.85	84.95	Höganäs AB quarry at Åsen, Scania, Sweden	Schönenberger and Friis, 2001
Parasaurauia allonensis	Early Campanian	Charcoalified	80.725	77.85	83.6	Allon locality, Georgia, USA	Keller et al. 1996
*Pecinovia annonoides**	Cenomanian	Charcoalified	97.2	93.9	100.5	Pecinov, Bohemia, Czech Republic	Kvacev and Eklund, 2003
Pennistemon portugallicus	Late Aptian to Early Albian	Charcoalified	111.975	106.7 5	117.2	Vale de Agua, Portugal	Friis et al. 2000
Pentapetalum trifasciculandricus	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Martínez-Millan et al. 2009
*Perseanthus crossmanensis**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Herendeen et al. 1994
*Platananthus synandrus**	Early to Middle Albian	Charcoalified	109.875	106.7 5	113		Pedersen et al. 1994
*Platanocarpus brookensis**	Middle Albian	Charcoalified	103.625	100.5	106.75	Bank near Brooke, Maryland, USA	Crane et al. 1993
Platanus quedlinburgensis	Santonian	Charcoalified	84.95	83.6	86.3	Blankenburg and Westerhausen, Germany	Tschan et al. 2008
Platydiscus peltatus	Early Campanian	Charcoalified	80.725	77.85	83.6	Höganäs AB quarry at Åsen, Scania, Sweden	Schönenberger et al. 2001
*Potomacanthus lobatus**	Early to Middle Albian	Charcoalified	109.875	106.7 5	113	Puddledock locality, Virginia, USA	von Balthazar et al. 2007
*Pragocladus lauroides**	Cenomanian	Charcoalified	97.2	93.9	100.5	Prague, Bohemia, CzechRepublic.	Kvacev and Eklund, 2003
Princetonia allenbyensis	Middle Eocene	Permineralized	37.55	33.9	41.2	Princeton chert of British Columbia, Canada	Stockey and Pigg, 1991
Protofagacea allonensis	Late Santonian	Charcoalified	84.95	83.6	86.3	Gaillard, Crawford County, Georgia, USA.	Herendeen et al. 1995
Protomimosoidea buchananensis	Late Paleocene to Early Eocene	Impression / Compression	53.5	47.8	59.2	Buchanan, Tennessee, USA	Crepet and Taylor, 1986
*Prunus cathybrownae**	Early Eocene	Impression / Compression	48.6	41.2	56	Boot Hill, Washington, USA	Benedict et al. 2011
*Quadriplatanus georgianus**	Coniacian to Santonian	Charcoalified	86.7	83.6	89.8	Upatoi Creek, Georgia, USA	Magallón-Puebla et al. 1997

*Rariglanda jerseyensis**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Martínez et al. 2016
Raritaniflora glandulosa	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet et al. 2013
Raritaniflora sphaerica	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet et al. 2013
Raritaniflora tomentosa	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet et al. 2013
Salacia lombardii	Miocene	Amber	14.18	5.33	23.03	Simojovel de Allende, Chiapas, México	Hernández-Damián et al. 2018
*Saportanthus brachystemon**	Late Aptian to Early Albian	Charcoalified	111.975	106.75	117.2	Vale de Agua, Portugal	Friis et al. 2017
*Saportanthus dolichostemon**	Late Aptian to Early Albian	Charcoalified	111.975	106.75	117.2	Vale de Agua, Portugal	Friis et al. 2017
*Saportanthus parvus**	Late Barremian to Albian	Charcoalified	116.375	106.75	126	Catefica, Portugal	Friis et al. 2017
*Saururus tuckerae**	Middle Eocene	Permineralized	39.5	37.8	41.2	Princeton Chert, British Columbia, Canada	Smith and Stockey, 2007
*Schenkeriphyllum glanduliferum**	Aptian	Impression / Compression	117.2	113	121.4	Nova Olinda, Araripe Basin, Brazil	Mohr et al. 2013
*Silvianthemum suecicum**	Santonian to Campanian	Charcoalified	79.2	72.1	86.3	Höganäs AB quarry at Åsen, Scania, Sweden	Friis, 1990; Friis et al. 2013
Solanites pusilus	Early Eocene	Impression / Compression	48.6	41.2	56	Claiborne, Tennesse, USA	Berry, 1930
*Spanomera mauldinensis**	Early Cenomanian	Charcoalified, Lignitized	98.85	97.2	100.5	Mauldin Mountain, Elk Neck Peninsula, Maryland, USA	Drinnan et al. 1991
*Switenia miocenica**	Miocene	Amber	14.18	5.33	23.03	Simojovel de Allende, Chiapas, México	Castañeda-Posadas and Cevallos-Ferriz, 2007
*Teixeiraea lusitanica**	Early Albian	Charcoalified	109.875	106.75	113	Vale de Agua, Portugal	von Balthazar et al. 2005
Teuschestanthes squamata	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Crepet et al. 2018
*Tinaflora beardiae**	Eocene	Permineralized	44.95	33.9	56	Appian Way, Vancouver Island, Canada	Atkinson et al. 2015

Tristaniandra alleyi	Middle Eocene	Impression / Compression	37.55	33.9	41.2	Golden Grove locality, Australia	Basinger et al. 2007
*Tylerianthus crossmanensis**	Turonian	Charcoalified	91.85	89.8	93.9	Old Crossman Clay Pit, Sayreville, New Jersey, USA	Gandolfo et al. 1998
Verneda hermaphroditica	Middle Cenomanian	Charcoalified	95.55	93.9	97.2	Saint Laurent la Vernede, France	Moreau et al., 2016
*Virginianthus calycanthoides**	Early to Middle Albian	Charcoalified	103.625	100.5	106.75	Puddledock locality, Virginia, USA	Friis et al. 1994
Zlivifructus vachae	Turonian to Santonian	Charcoalified	88.75	83.6	93.9	Třebeč, Czech Republic	Hermanova et al. 2017

Supplementary Data 3. Additional calibration points used in the skyline fossilized birth-death model using total evidence and morphological data. For further information on point calibrations see Ramírez-Barahona et al. (2020).

Order	Family	Node calibrated	Fossil taxa	Stratigraphic Age	MaxAge	MinAge
Canellales	Winteraceae	Stem Winteraceae/Crown Canellales	†*Walkeripollis gabonensis*	Barremian	129.4	121.4
Piperales	Aristolochiaceae	crown Aristolochiaceae	†*Aristolochiacidites viluiensis*	Campanian	83.6	72.1
Alismatales	Araceae	Crown Araceae	†*Mayoa portugallica*	Early Aptian	121.4	117.2
Alismatales	Alistamaceae	stem Alismataceae	†*Cardstonia toldmanii*	Campanian	83.6	72.1
Alismatales	Alistamaceae	crown Alismataceae	†*Caldesia brandoniana*	Oligocene	33.9	23.03
Alismatales	Hydrocharitaceae	crown Hydrocharitaceae	†*Stratiotes sp.*	Ypresian assumed	56	47.8
Alismatales	Aponogetonaceae	Stem Aponogetonaceae	†*Aponogeton harryi*	Campanian	83.6	72.1
Alismatales	Posidoniaceae	Stem Posidoniaceae	†*Thalassotaenia debeyi*	Maastrichtian	72.1	66
Pandanales	Pandanaceae	crown Pandanaceae	†*Pandanus sp.? cf. P. shiabensis*	Paleocene	66	56
Pandanales	Cyclanthaceae	Crown Cyclanthaceae	†*Cyclanthus messelensis*	Eocene	56	33.9
Dioscoreales	Dioscoreae	Stem Dioscoreaceae	†*Dioscorea wilkinii*	Oligocene	33.9	23.03
Liliales	Ripogonaceae	Stem Ripogonaceae	†*Ripogonum tasmanicum*	mid-Ypresian	56	47.8
Liliales	Alstroemeriaceae	crown Alstroemeriaceae	†*Luzuriaga peterbannisteri*	Chattian	27.82	23.03

Liliales	Petermanniaceae	stem Petermanniaceae	†*Petermanniopsis angleseaensis*	Late Eocene	47.8	41.2
Asparagales	Orchidaceae	Crown Orchidaceae	†*Dendrobium winikaphyllum*	Late Oligocene - EarlyMiocene	27.82	20.44
Asparagales	Asteliaceae	Crown Asteliaceae	†*Astelia antiquua*	Late Oligocene – Early Miocene	27.82	20.44
Asparagales	Asphodelaceae	Crown Asphodelaceae	†*Dianellophyllum eocenicum*	Middle Eocene (Lutetian assumed)	47.8	41.2
Asparagales	Asparagaceae	Crown Asparagaceae	†*Paracordyline kerguelensis*	Early Miocene (Aquitanian)	23.03	20.44
Arecales		Stem Arecaceae	†*Sabalites carolinensis*	Santonian	86.3	83.6
Arecales	Arecaceae	Crown Arecaeae	†*Mauritiidites crassibaculatus*	Maastrichtian	72.1	66
Commelinales	Commelinaceae	Crown Commeliaceae	†*Pollia tugenensis*	Serravallian assumed	13.82	11.63
Zingiberales		Crown Zingiberales	†*Spirematospermum chandlerae*	Campanian	83.6	72.1
Zingiberales		Stem Zingiberaceae	†*Zingiberopsis magnifolia*	Maastrichtian	72.1	66
Poales	Typhaceae	Crown Typhaceae	†*Typha sp.*	Ypresian assumed	56	47.8
Poales	Bromeliaceae	Crown Bromeliaceae	†*Tillandsia-type pollen*	Middle Eocene (Lutetian assumed)	47.8	41.2
Poales	Cyperaceae	Crown Cyperaceae	†*Volkeria messelensis*	Early Eocene assumed	56	41.2
Poales	Poaceae	Crown Poaceae	†*Changii indicum*	Maastrichtian	72.1	66
Proteales	Sabiaceae	Crown Sabiaceae	†*Sabia menispermoides*	Turonian	93.9	89.8
Dilleniales	Dilleniaceae	Crown Dilleniaceae	†*Tetracera eocenica*	Ypresian	56	47.8
Vitales	Vitaceae	Crown Vitaceae	†*Indovitis chitaleyae*	Danian	66	61.6
Geraniales	Francoaceae	Crown Francoaceae	†*Rhynchotheca type*	Burdigalian	20.44	15.97
Myrtales	Lythraceae	Crown Lythraceae	†*Decodon tiffneyi*	Campanian	83.6	72.1
Myrtales	Myrtaceae	Crown Myrtaceae	†*Paleomyrtinaea sp.*	Late Paleocene	59.2	56
Myrtales		Stem Melastomataceae	†*Melastomites montanensis*	Paleocene	66	56
Crossosomatales	Staphylaceae	Crown Staphylaceae	†*Staphylea minutidens*	Paleocene	66	56
Sapindales	Burseraceae	Crown Burseraceae	†*Bursericarpum aldwickense*	Ypresian	56	47.8
Sapindales	Sapindaceae	Crown Sapindaceae	†*Aesculus hickeyi*	Paleocene	66	56

Sapindales		Stem Sapindaceae	†*Sapindospermum nitidum*	Turonian	93.9	89.8
Sapindales	Simaroubaceae	Crown Simaroubaceae	†*Ailanthus confucii*	Middle Eocene (Lutetian assumed)	47.8	41.2
Sapindales	Rutaceae	Crown Rutaceae	†*Rutaspermum biornatum*	Maastrichtian	72.1	66
Huerteales	Tapisciaceae	Crown Tapisciaceae	†*Tapiscia occidentalis*	Middle Eocene (Lutetian assumed)	47.8	41.2
Brassicales	Akaniaceae	Crown Akaniaceae	†*Akania sp.*	Danian	66	61.6
Malvales	Bixaceae	Crown Bixaceae	†*Cochlospermum previtifolium*	Burdigalian	20.44	15.97
Malvales	Dipterocarpaceae	Crown Dipterocarpaceae	†*Foveotricolpites alveolatus*	Paleocene	66	56
Malvales		Stem Malvales	†*Javelinoxylon weberi*	Campanian	83.6	72.1
Malvales	Malvaceae	Crown Malvaceae	†*Malvaciphyllum macondicus*	Selandian	61.6	59.2
Fabales	Surianaceae	Crown Surianaceae	†*Suriana inordinata*	Ypresian (assumed)	56	47.8
Fabales		Stem Fabaceae	†*Leguminocarpon gardneri*	Paleocene	66	56
Fabales	Fabaceae	Crown Fabaceae	†*Diplotropis-like leaves and fruits*	Paleocene	66	56
Rosales	Cannabaceae	Crown Cannabaceae	†*Aphananthe cretacea*	Maastrichtian	72.1	66
Rosales	Moraceae	Crown Moraceae	†*Morus poolensis*	Ypresian (assumed)	56	47.8
Cucurbitales	Anysophyllaceae	Crown Anisophyllaceae	†*Combretocarpus rotundatus*	Middle Miocene (Serravallian assumed)	13.82	11.63
Cucurbitales	Corynocarpaceae	crown Corynocarpaceae (Cucurbitales)	†*Corynocarpus sp.*	Early Miocene	23.03	15.97
Cucurbitales		stem Cucurbitaceae (Cucurbitales)	†*Cucurbitospermum sheppeyense*	Ypresian	56	47.8
Celastrales		stem Celastrales	†*Lophopetalumoxylon indicum*	Lutetian (assumed)	47.8	41.2
Celastrales	Celastraceae	crown Celastraceae	†*Celastrus comparabilis*	Middle Eocene (Bartonian assumed)	41.2	37.71
Malpighiales	Chrysobalanaceae	crown Chrysobalanaceae (Malpighiales)	†*Chrysobalanus lacustris*	Ypresian (assumed)	56	47.8
Malpighiales	Malpighiaceae	crown Malpighiaceae	†*Tetrapterys harpyiarum*	Early Oligocene (Rupelian assumed)	33.9	27.82
Malpighiales	Phyllanthaceae	crown Phyllathaceae	†*Phyllanthus sp.*	Late Eocene	47.8	41.2

Malpighiales	Salicaceae	crown Salicaceae	†*Populus tidwelli*	Ypresian (assumed)	56	47.8
Malpighiales	Rhizophoraceae	crown Rhizophoraceae	†*Zonocostites ramonae*	Late Eocene	47.8	41.2
Malpighiales	Humariaceae	crown Humariaceae	†*Humiriaceoxylon ocuensis*	Late Eocene	47.8	41.2
Malpighiales		stem Ctenolophonaceae (Malpighiales)	†*Ctenolophonidites costatus*	Maastrichtian	72.1	66
Malpighiales	Euphorbiaceae	crown Euphorbiaceae	†*Hippomaneoidea warmanensis*	Bartonian	41.2	37.71
Santalales		stem Olaceae	†*Anacolosidites meyeorum*	Maastrichtian	72.1	66
Santalales	Schoepfiaceae	crown Schoepfiaceae	†*Schoepfia republicensis*	Ypresian (assumed)	56	47.8
Caryophyllales		stem Polygalaceae	†*Polygonocarpum johnsonii*	Maastrichtian	72.1	66
Caryophyllales	Caryophyllaceae	crown Caryophyllaceae	†*Caryophylloflora paleogenica*	Middle Eocene (Lutetian assumed)	47.8	41.2
Caryophyllales	Phytolaccaceae	crown Phytolaccaceae	†*Coahuilacarpon phytolaccoides*	Campanian	83.6	72.1
Cornales	Cornaceae	crown Cornaceae	†*Cornus hyperborea*	Paleocene	66	56
Cornales		crown Cornales	†*Hironoia fusiformis*	Coniacian	89.8	86.3
Cornales		stem Nyssaceae	†*Beckettia samuelis*	Maastrichtian	72.1	66
Cornales	Hydrangeaceae	crown Hydrangeaceae	†*Hydrangea antica*	Paleocene	66	56
Aquifoliales	Aquifoliaceae	stem Aquifoliaceae	†*Ilex hercynica*	Early Paleocene (Danian assumed)	66	61.6
Dipsacales	Caprifoliaceae	crown Caprifoliaceae	†*Diplodipelta reniptera*	Priabonian (assumed)	37.71	33.9
Apiales	Torricelliaceae	crown Torricelliaceae	†*Toricellia bonensii*	Late Paleocene	61.6	56
Apiales	Araliaceae	crown Araliaceae	†*Paleopanax oregonensis*	Middle Eocene (Lutetian assumed)	47.8	41.2
Apiales		stem Apiaceae	†*Carpites ulmiformis*	Maastrichtian	72.1	66
Asterales	Campanulaceae	crown Campanulaceae	†*Campanula palaeopyramidalis*	Miocene	23.03	5.33
Asterales	Menyanthaceae	crown Menyanthaceae	†*Menyanthes cf. trifoliata*	Miocene	23.03	5.33
Asterales	Goodeniaceae	crown Googeniaceae	†*Poluspissusites ramus*	Oligocene	33.9	23.03
Asterales		stem Asterales	†*Tubulifloridites lilliei type A*	Campanian	83.6	72.1
Asterales	Asteraceae	crown Asteraceae	†*Raiguenrayun cura*	Ypresian (assumed)	56	47.8

Garryales		stem Eucommiaceae	†*Eucommia montana*	Early Eocene (Ypresian assumed)	56	47.8
Gentianales	Rubiaceae	crown Rubiaceae	†*Emmenopterys dilcheri*	Middle Eocene (Lutetian assumed)	47.8	41.2
Gentianales	Gentianaceae	crown Gentinaceae	†*Voyrioseminites magnus*	Eocene	56	33.9
Gentianales	Apocynaceae	crown Apocynaceae	†*Apocynospermum coloradensis*	Ypresian (assumed)	56	47.8
Solanales	Convolvulaceae	crown Convolvulaceae	†*Calystegiapollis microechinatus*	Early Eocene (Ypresian assumed)	56	47.8
Solanales	Solanaceae	crown Solanaceae	†*Physalis infinemundi*	Ypresian (assumed)	56	47.8
Boraginales	Ehretiaceae	crown Ehretiaceae	†*Ehretia clausentia*	Ypresian	56	47.8
Boraginales	Cordiaceae	crown Cordiaceae	†*Cordia platanifolia*	Paleogene	66	23.03
Lamiales	Oleaceae	crown Oleaceae	†*Fraxinus rupinarum*	Rupelian	33.9	27.82
Lamiales	Byblidaceae	crown Byblidaceae	†*Golden Grove Byblidaceae parataxon*	Middle Eocene (Lutetian assumed)	47.8	41.2
Lamiales	Acanthaceae	crown Acanthaceae	†*unnamed cf. Avicennia*	Bartonian	41.2	37.71
Lamiales	Bignoniaceae	crown Bignoniaceae	†*Catalpa sp.*	Bartonian assumed	41.2	37.71
Lamiales	Lamiaceae	crown Lamiaceae	†*Melissa parva*	Early Oligocene (Rupelian assumed)	33.9	27.82

Supplementary Table 1. Detailed age estimates under the constrained skyline fossilized birth-death (SFBD) total evidence approach. L95HPD, Lower value of the 95% Highest Posterior Density; U95HPD upper value of the 95% Highest Posterior Density.

Family	CrownAge_mean	CrownAge_L95HPD	CrownAge_U95HPD	StemAge_mean	StemAge_L95HPD	StemAge_U95HPD
Amborellaceae	115.794	99.655	132.534	151.349	150.325	151.799
Ceratophyllaceae	16.158	10.628	21.677	146.140	142.888	149.173
Lauraceae	134.908	132.063	139.031	140.033	137.808	143.357
Chloranthaceae	119.874	110.677	131.157	137.984	128.310	144.356
Sabiaceae	67.415	54.421	82.321	137.885	135.522	140.266
Buxaceae	109.675	101.922	117.352	136.305	133.956	138.338

Papaveraceae	83.244	65.300	106.424	135.231	129.651	141.849
Trochodendraceae	94.460	89.783	99.070	134.987	133.128	136.898
Calycanthaceae	98.446	93.310	105.033	134.908	132.063	139.031
Myristicaceae	113.021	96.866	129.700	134.740	129.304	138.524
Nelumbonaceae	14.887	10.447	20.605	132.145	125.017	138.807
Dilleniaceae	45.127	17.864	68.201	130.459	128.356	132.971
Eupteleaceae	14.731	9.870	21.207	130.391	125.801	135.053
Hydatellaceae	43.430	29.468	53.136	129.479	118.868	141.042
Austrobaileyaceae	116.879	108.489	125.456	128.872	123.972	134.070
Siparunaceae	108.457	103.533	115.258	127.261	123.992	131.745
Proteaceae	52.416	37.745	73.830	124.083	118.284	132.458
Platanaceae	120.010	113.328	125.864	124.083	118.284	132.458
Vitaceae	60.784	45.198	85.485	123.505	120.581	127.394
Degeneriaceae	121.657	118.817	124.684	123.013	120.400	126.218
Atherospermataceae	54.540	43.494	66.614	122.961	118.505	128.153
Gomortegaceae	117.271	112.702	122.219	122.961	118.505	128.153
Schisandraceae	114.187	109.204	120.497	122.647	116.641	128.170
Menispermaceae	112.504	108.935	115.212	120.089	113.853	124.499
Acoraceae	16.085	9.855	21.681	119.899	111.665	130.369
Lardizabalaceae	114.356	110.201	119.787	119.035	112.912	124.211
Circaeasteraceae	69.337	43.070	88.017	119.035	112.912	124.211
Peridiscaceae	64.865	47.983	86.939	118.579	113.466	125.396
Nymphaeaceae	113.740	108.607	121.150	118.542	113.041	125.233
Cabombaceae	56.346	41.352	72.788	118.542	113.041	125.233
Aristolochiaceae	94.447	71.859	114.418	115.690	101.590	128.790
Rhabdodendraceae	16.554	11.137	22.460	112.373	106.484	119.461
Paeoniaceae	14.521	9.407	20.584	111.317	105.565	118.084

Primulaceae	82.824	71.692	94.200	111.019	105.788	117.383
Picramniaceae	16.900	10.957	22.892	110.549	107.513	114.409
Berberidaceae	51.216	33.001	68.379	110.181	95.336	119.563
Ranunculaceae	88.390	80.493	97.388	110.181	95.336	119.563
Altingiaceae	16.076	11.907	22.631	106.656	98.430	112.522
Tofieldiaceae	46.792	33.406	67.011	105.118	98.048	113.463
Huaceae	15.052	10.300	22.907	103.651	96.510	110.241
Hamamelidaceae	96.040	89.605	102.602	102.375	95.508	110.252
Sapotaceae	46.135	31.354	64.159	102.063	96.963	107.952
Rosaceae	66.351	54.980	78.041	101.526	94.180	107.736
Centroplacaceae	73.498	54.275	95.608	99.745	96.882	102.905
Iteaceae	94.620	90.594	99.166	99.688	94.204	106.037
Kirkiaceae	13.746	9.410	19.949	99.569	97.204	102.731
Burseraceae	56.825	38.933	72.731	98.239	95.174	102.350
Anacardiaceae	96.609	94.491	100.111	98.239	95.174	102.350
Petrosaviaceae	55.829	40.666	74.080	98.011	87.793	108.535
Araceae	53.935	40.607	69.576	97.219	86.237	111.150
Nothofagaceae	15.724	9.726	22.134	96.650	90.386	101.046
Sarraceniaceae	47.386	29.657	63.344	96.062	90.938	102.338
Clethraceae	92.597	90.411	95.560	95.321	91.775	99.603
Strombosiaceae	22.129	11.874	42.016	94.316	83.984	106.714
Saxifragaceae	85.183	81.043	89.477	93.551	88.852	99.928
Grossulariaceae	15.328	9.366	21.449	93.551	88.852	99.928
Clusiaceae	91.773	90.241	93.784	92.947	90.793	95.524
Fagaceae	55.916	39.403	72.831	92.867	88.084	96.602
Lecythidaceae	61.782	50.698	77.496	92.532	79.366	107.205
Ebenaceae	41.078	27.042	64.050	92.532	79.366	107.205

Stegnospermataceae	15.208	10.520	21.772	92.477	82.792	101.237
Droseraceae	49.212	38.419	58.981	92.416	80.702	104.265
Actinidiaceae	88.682	85.838	93.081	92.357	87.523	97.055
Roridulaceae	15.594	10.308	20.536	92.357	87.523	97.055
Hernandiaceae	58.357	43.673	69.984	90.305	77.267	110.965
Monimiaceae	56.581	35.277	76.504	90.305	77.267	110.965
Bruniaceae	58.261	42.654	76.179	90.102	85.702	94.710
Sapindaceae	44.542	30.205	61.140	89.078	70.366	100.087
Vahliaceae	15.664	9.385	21.799	88.667	86.162	91.762
Canellaceae	19.044	13.937	22.826	88.433	71.127	112.663
Winteraceae	16.397	10.086	21.682	88.433	71.127	112.663
Crassulaceae	60.990	54.012	71.734	88.376	75.766	100.845
Geraniaceae	61.114	43.524	73.174	88.045	70.300	104.782
Francoaceae	57.286	43.053	71.360	88.045	70.300	104.782
Erythropalaceae	63.896	28.285	85.248	87.774	79.721	97.346
Gunneraceae	17.357	10.297	22.738	87.393	62.817	107.132
Myrothamnaceae	14.284	9.193	19.782	87.393	62.817	107.132
Bonnetiaceae	49.143	35.256	63.381	86.693	78.502	92.810
Paracryphiaceae	64.104	49.971	77.481	86.615	82.469	90.834
Juglandaceae	42.257	14.540	57.326	85.243	82.078	88.999
Myricaceae	82.891	79.844	85.903	85.243	82.078	88.999
Magnoliaceae	19.403	10.827	37.754	84.737	59.745	112.924
Macarthuriaceae	14.108	10.311	19.755	84.680	70.103	101.713
Microteaceae	38.522	11.252	65.925	84.680	70.103	101.713
Lophiocarpaceae	53.063	31.173	73.628	84.366	79.573	89.702
Theaceae	35.765	17.079	58.122	84.299	69.347	99.339
Escalloniaceae	72.734	63.242	81.459	83.488	80.095	86.557

Kewaceae	14.735	9.593	20.470	82.169	77.595	87.328
Saururaceae	49.926	33.106	72.609	82.137	71.024	94.800
Piperaceae	62.426	41.547	74.674	82.137	71.024	94.800
Limeaceae	31.308	12.220	50.741	81.931	73.343	92.479
Burmanniaceae	38.323	16.249	57.455	81.917	70.598	97.746
Polemoniaceae	15.279	10.216	20.560	81.768	59.638	95.249
Fouquieriaceae	59.670	38.631	78.085	81.768	59.638	95.249
Balanopaceae	19.953	9.782	40.815	81.541	71.494	89.962
Malpighiaceae	30.533	27.821	33.488	81.365	68.692	92.752
Elatinaceae	57.972	48.662	67.977	81.365	68.692	92.752
Zygophyllaceae	62.854	51.013	73.799	80.583	63.727	94.728
Orchidaceae	46.098	36.689	58.669	80.091	71.028	89.014
Icacinaceae	55.085	39.892	71.537	79.959	68.613	91.691
Ximeniaceae	36.818	13.491	61.995	79.720	71.768	86.868
Cornaceae	60.843	56.442	65.866	79.531	71.926	87.611
Aizoaceae	64.360	50.591	76.755	78.683	75.347	82.461
Pennantiaceae	14.316	9.546	19.480	77.328	74.634	79.029
Nartheciaceae	34.149	22.213	46.285	76.976	65.233	92.663
Aponogetonaceae	33.107	11.601	52.050	76.861	65.637	86.783
Coulaceae	18.381	9.577	28.614	76.800	70.135	84.207
Menyanthaceae	19.008	14.292	22.950	76.774	74.477	79.122
Calophyllaceae	59.393	43.911	71.782	76.693	67.007	87.105
Velloziaceae	48.991	33.704	65.587	76.503	61.131	91.841
Gisekiaceae	14.332	9.837	20.346	76.472	73.317	81.040
Ulmaceae	26.950	10.719	45.541	76.383	72.569	80.517
Cunoniaceae	70.886	64.604	77.274	75.861	68.707	81.663
Achariaceae	52.530	39.412	72.908	75.434	67.165	84.824

Griseliniaceae	17.583	10.761	28.283	75.270	72.586	77.260
Phytolaccaceae	72.998	72.101	74.690	74.537	72.482	77.233
Goodeniaceae	31.231	27.296	33.891	74.164	72.933	75.535
Biebersteiniaceae	15.490	11.268	22.063	73.648	57.446	98.322
Nitrariaceae	19.592	9.783	34.367	73.648	57.446	98.322
Molluginaceae	59.264	43.919	71.269	73.513	63.874	86.088
Cannabaceae	68.083	66.003	70.675	73.060	68.597	76.868
Rousseaceae	58.250	40.009	69.002	73.060	59.098	82.665
Campanulaceae	16.658	11.465	22.817	73.060	59.098	82.665
Torricelliaceae	58.196	56.003	61.063	73.004	70.575	75.501
Humiriaceae	44.034	41.220	47.159	72.900	67.358	78.325
Asteraceae	51.188	47.814	55.255	72.731	72.105	73.589
Calyceraceae	14.468	11.019	18.915	72.731	72.105	73.589
Tetrachondraceae	34.461	19.575	51.822	72.560	67.299	78.406
Alismataceae	39.655	30.710	50.136	72.122	57.408	85.632
Caryocaraceae	19.683	13.285	22.923	71.865	58.320	86.277
Ixonanthaceae	15.744	12.044	21.723	71.865	58.320	86.277
Neuradaceae	23.057	10.772	44.229	71.492	52.943	83.491
Ericaceae	54.185	37.245	70.777	71.292	50.827	89.210
Cyrillaceae	15.641	11.290	21.371	71.292	50.827	89.210
Convolvulaceae	50.467	47.809	54.687	71.236	64.697	78.608
Solanaceae	51.123	47.866	55.686	71.236	64.697	78.608
Pittosporaceae	17.117	10.534	22.754	70.860	68.475	73.808
Anisophylleaceae	12.908	11.770	13.814	70.502	62.966	79.218
Nyssaceae	51.134	31.029	66.752	70.379	68.047	72.092
Montiniaceae	34.926	20.372	45.175	69.903	62.629	78.176
Elaeagnaceae	18.782	13.541	24.679	69.696	60.605	79.428

Aptandraceae	60.144	51.716	67.320	69.668	66.499	72.099
Stylidaceae	55.148	40.360	66.877	69.556	58.596	78.555
Rapateaceae	38.191	18.550	56.193	69.519	65.639	73.242
Polygonaceae	38.396	18.943	52.663	69.513	66.478	72.086
Plumbaginaceae	45.263	33.057	54.187	69.513	66.478	72.086
Columelliaceae	41.401	28.895	53.592	69.058	62.719	75.880
Araliaceae	44.484	41.423	47.678	68.882	66.574	71.573
Caryophyllaceae	45.598	42.200	47.791	68.828	53.868	84.563
Ctenolophonaceae	14.439	10.133	19.918	68.571	66.001	71.630
Chrysobalanaceae	51.556	47.871	55.293	68.479	61.013	76.963
Symplocaceae	14.521	9.855	20.485	67.788	58.527	81.926
Loranthaceae	17.742	10.008	33.147	67.613	61.978	72.567
Hypericaceae	39.902	26.997	56.326	67.427	60.803	73.507
Podostemaceae	44.858	31.413	57.219	67.427	60.803	73.507
Melanthiaceae	53.399	47.625	60.723	67.040	57.149	81.537
Passifloraceae	44.547	36.365	54.449	66.995	58.343	77.389
Myodocarpaceae	15.036	11.002	20.769	66.932	66.002	68.282
Apiaceae	51.790	38.508	63.365	66.932	66.002	68.282
Pentaphylacaceae	28.592	15.268	45.535	66.818	52.324	85.485
Sladeniaceae	56.088	39.210	69.545	66.818	52.324	85.485
Dasypogonaceae	17.234	11.440	22.852	66.476	55.225	77.458
Arecaceae	22.193	9.959	43.646	66.476	55.225	77.458
Hydrostachyaceae	14.998	9.988	18.245	66.410	62.974	69.707
Combretaceae	34.646	18.888	48.867	66.352	53.504	81.132
Thymelaeaceae	35.048	15.998	61.909	66.247	49.319	80.557
Rubiaceae	44.894	41.621	47.785	66.220	59.953	70.466
Petiveriaceae	40.107	12.373	59.225	66.147	58.001	73.464

Plantaginaceae	60.122	49.915	71.373	65.885	58.504	73.732
Cercidiphyllaceae	13.740	9.905	18.710	65.629	49.455	81.111
Rhamnaceae	43.267	27.212	63.932	65.414	51.964	80.223
Elaeocarpaceae	39.184	20.107	51.462	65.412	52.458	77.510
Stemonaceae	31.759	16.034	53.982	65.106	42.607	86.425
Erythroxylaceae	39.793	28.057	47.885	64.885	59.438	70.751
Rhizophoraceae	52.917	45.801	62.391	64.885	59.438	70.751
Bromeliaceae	17.108	12.049	22.934	64.734	56.020	72.706
Typhaceae	26.736	12.880	43.299	64.734	56.020	72.706
Violaceae	47.842	36.063	60.107	64.709	55.889	74.167
Aquifoliaceae	15.231	11.100	22.587	63.963	61.839	65.952
Meliaceae	23.743	13.976	53.039	63.714	53.827	73.971
Annonaceae	21.003	11.891	38.446	63.324	34.230	89.861
Eupomatiaceae	14.869	9.645	21.298	63.324	34.230	89.861
Scrophulariaceae	51.323	41.085	61.652	63.307	56.277	69.792
Loasaceae	44.681	35.565	54.652	63.220	59.503	66.821
Hydrangeaceae	58.516	56.002	62.252	63.220	59.503	66.821
Hydrophyllaceae	32.281	14.766	47.170	63.148	54.079	72.303
Irvingiaceae	17.263	9.931	22.971	63.028	54.394	73.113
Pandaceae	26.326	16.036	56.088	63.028	54.394	73.113
Eriocaulaceae	45.133	36.849	54.618	62.903	57.320	66.741
Urticaceae	48.039	33.272	60.263	62.135	53.975	69.822
Moraceae	50.761	47.801	54.522	62.135	53.975	69.822
Misodendraceae	25.595	12.408	41.035	61.579	56.177	67.370
Schoepfiaceae	51.956	48.099	55.673	61.579	56.177	67.370
Linderniaceae	44.050	32.255	55.881	60.951	54.965	67.768
Xyridaceae	48.728	41.818	55.112	60.656	55.471	64.472

Santalaceae	52.524	40.675	61.380	60.127	49.755	67.845
Opiliaceae	25.089	14.202	43.588	60.127	49.755	67.845
Juncaginaceae	53.122	44.177	61.734	60.015	54.036	66.645
Dioscoreaceae	51.412	41.495	59.464	59.801	52.948	66.090
Taccaceae	15.802	11.573	22.982	59.801	52.948	66.090
Casuarinaceae	15.846	11.283	22.264	59.519	39.635	75.845
Adoxaceae	43.385	31.690	53.406	59.409	52.740	65.329
Caprifoliaceae	35.919	34.134	37.709	59.409	52.740	65.329
Ochnaceae	18.982	13.217	22.928	59.401	45.895	74.157
Linaceae	32.528	14.143	48.021	59.401	45.895	74.157
Nyctaginaceae	49.763	39.414	60.560	58.954	47.915	66.277
Hydrocharitaceae	37.593	28.382	48.526	58.936	47.788	72.261
Namaceae	41.339	30.127	52.967	58.678	52.781	64.778
Limnanthaceae	17.666	10.150	28.851	58.464	43.550	75.140
Putranjivaceae	16.174	10.809	22.942	58.399	41.446	72.524
Apocynaceae	50.380	47.801	54.529	57.973	50.437	64.015
Gelsemiaceae	34.707	19.193	51.842	57.973	50.437	64.015
Picrodendraceae	16.538	10.930	22.352	57.493	48.866	67.004
Phyllanthaceae	45.374	41.882	47.793	57.493	48.866	67.004
Boryaceae	26.947	12.921	45.268	57.455	47.193	72.955
Campynemataceae	29.547	12.514	52.893	56.997	40.975	70.571
Cardiopteridaceae	36.709	16.057	55.065	56.991	47.011	71.854
Stemonuraceae	14.805	10.050	20.542	56.991	47.011	71.854
Gesneriaceae	49.065	40.915	60.296	56.602	48.045	66.602
Simaroubaceae	31.859	14.405	55.274	56.181	48.073	66.004
Rutaceae	44.987	35.427	51.516	56.181	48.073	66.004
Codonaceae	14.113	9.613	18.698	56.035	46.855	66.600

Restionaceae	47.968	39.927	60.538	55.935	46.177	62.610
Diapensiaceae	24.138	17.810	38.434	55.919	45.758	64.327
Styracaceae	37.508	19.112	56.372	55.919	45.758	64.327
Philydraceae	18.294	13.675	22.033	55.848	45.827	63.770
Euphorbiaceae	39.539	37.912	41.158	55.385	47.580	65.385
Peraceae	40.354	30.807	49.915	55.385	47.580	65.385
Doryanthaceae	16.168	10.670	22.231	55.378	44.734	71.192
Frankeniaceae	14.409	9.604	20.665	55.362	38.893	71.608
Tamaricaceae	26.430	11.286	50.835	55.362	38.893	71.608
Salicaceae	50.391	47.807	54.314	55.323	49.794	64.161
Lacistemataceae	14.764	10.092	20.567	55.323	49.794	64.161
Grubbiaceae	14.283	10.451	17.864	55.258	43.332	65.467
Helwingiaceae	13.447	9.535	18.241	55.098	44.819	61.163
Phyllonomaceae	14.255	9.421	21.021	55.098	44.819	61.163
Fabaceae	37.214	27.876	48.559	55.031	46.414	62.475
Surianaceae	21.592	12.556	36.832	55.031	46.414	62.475
Hydroleaceae	13.942	9.445	17.950	54.750	44.585	64.039
Dichapetalaceae	17.150	11.995	22.905	54.603	44.380	63.658
Trigoniaceae	37.894	20.954	49.607	54.603	44.380	63.658
Heliotropiaceae	15.817	10.768	22.954	54.499	49.909	58.660
Loganiaceae	41.619	29.898	54.142	54.136	43.006	63.332
Gentianaceae	39.229	33.963	45.858	54.136	43.006	63.332
Aphanopetalaceae	15.343	10.561	22.004	54.015	45.091	64.341
Stilbaceae	31.542	13.934	47.946	53.482	47.518	59.531
Byblidaceae	43.598	41.210	46.754	53.482	47.518	59.531
Acanthaceae	39.501	37.834	41.120	53.450	46.834	59.835
Onagraceae	44.360	30.254	53.828	53.353	34.465	63.460

Lythraceae	26.603	14.236	40.885	53.353	34.465	63.460
Cucurbitaceae	21.286	15.779	27.630	53.302	49.294	55.996
Oxalidaceae	36.188	21.322	49.329	52.891	35.217	67.318
Connaraceae	15.318	10.501	21.983	52.891	35.217	67.318
Achatocarpaceae	14.808	10.400	20.217	52.744	41.200	64.310
Amaranthaceae	27.379	17.247	42.974	52.744	41.200	64.310
Tecophilaeaceae	25.379	12.475	37.809	52.592	39.605	72.137
Ixioliriaceae	16.274	10.755	21.289	52.592	39.605	72.137
Commelinaceae	36.550	29.160	45.010	52.545	43.165	62.782
Hanguanaceae	17.056	9.944	22.788	52.545	43.165	62.782
Thurniaceae	18.929	11.681	22.981	52.515	48.887	56.268
Garryaceae	16.906	11.761	22.714	52.247	48.293	55.836
Strasburgeriaceae	17.318	12.141	22.730	51.835	36.069	69.105
Ehretiaceae	48.949	47.802	50.963	51.304	49.011	54.720
Montiaceae	22.649	10.134	42.825	50.805	35.431	68.058
Martyniaceae	15.486	9.791	22.223	50.189	44.702	56.146
Haemodoraceae	42.138	35.606	51.195	50.104	43.006	58.515
Pontederiaceae	19.945	11.825	33.049	50.104	43.006	58.515
Myrtaceae	21.450	11.220	35.224	50.025	36.990	65.435
Vochysiaceae	18.097	11.359	23.822	50.025	36.990	65.435
Blandfordiaceae	13.989	10.140	18.563	48.291	39.212	61.447
Flagellariaceae	15.685	10.530	21.989	48.228	43.182	53.983
Ancistrocladaceae	16.228	9.998	22.321	48.131	36.766	60.139
Dioncophyllaceae	15.050	9.748	19.580	48.131	36.766	60.139
Iridaceae	20.888	13.426	30.047	47.767	39.995	60.743
Calceolariaceae	15.653	10.326	21.597	47.476	34.927	56.548
Coriariaceae	14.268	10.375	20.256	46.815	37.852	57.488

Corynocarpaceae	18.473	15.987	22.032	46.815	37.852	57.488
Datiscaceae	14.626	10.339	22.358	46.607	40.877	51.809
Lamiaceae	30.924	28.108	33.841	46.565	39.655	53.680
Melastomataceae	21.898	13.867	29.648	46.538	35.769	58.673
Pandanaceae	16.503	10.073	22.998	46.142	21.790	64.295
Alseuosmiaceae	15.137	10.638	20.526	45.946	28.376	57.956
Celastraceae	39.782	37.992	41.193	45.683	42.669	47.752
Lepidobotryaceae	14.681	9.443	20.316	45.683	42.669	47.752
Musaceae	14.102	10.769	20.363	45.682	37.951	53.240
Lentibulariaceae	28.495	16.665	40.974	44.894	34.824	54.538
Pedaliaceae	21.684	10.436	43.490	44.894	34.824	54.538
Staphyleaceae	14.197	10.093	20.635	44.700	19.563	65.099
Dipentodontaceae	27.032	12.046	43.711	44.456	30.281	58.966
Tapisciaceae	25.632	11.993	39.594	44.456	30.281	58.966
Carlemanniaceae	15.705	10.712	21.890	44.425	33.644	55.399
Oleaceae	30.962	28.173	33.893	44.425	33.644	55.399
Marcgraviaceae	16.596	10.683	22.792	44.373	33.519	58.652
Haloragaceae	15.991	11.337	21.637	44.252	34.569	52.256
Penthoraceae	13.414	10.215	17.517	44.252	34.569	52.256
Bignoniaceae	39.253	37.715	40.953	44.041	40.164	48.352
Verbenaceae	29.002	17.029	44.554	44.041	40.164	48.352
Cordiaceae	35.792	25.940	45.492	43.703	35.838	51.247
Caricaceae	15.700	10.912	22.926	43.631	30.733	58.912
Moringaceae	15.383	10.074	20.889	43.631	30.733	58.912
Alstroemeriaceae	25.083	16.203	34.592	43.362	28.789	54.621
Colchicaceae	19.050	13.974	22.946	43.362	28.789	54.621
Betulaceae	19.129	12.131	27.126	42.991	28.602	58.951

Cyperaceae	27.915	16.825	42.488	42.779	33.414	46.919
Juncaceae	29.610	16.578	42.805	42.779	33.414	46.919
Bixaceae	28.280	15.447	44.357	42.552	29.872	55.043
Sphaerosepalaceae	14.662	10.801	20.869	42.552	29.872	55.043
Berberidopsidaceae	27.934	11.672	54.992	42.445	18.559	71.320
Heliconiaceae	15.069	10.628	21.318	42.442	34.420	51.997
Muntingiaceae	20.184	10.825	40.502	42.243	26.598	57.592
Malvaceae	13.452	9.812	19.336	42.243	26.598	57.592
Triuridaceae	29.693	10.174	45.512	42.060	18.793	57.766
Cyclanthaceae	25.915	11.244	45.618	42.060	18.793	57.766
Boraginaceae	18.643	14.026	22.926	42.052	30.442	52.591
Wellstediaceae	15.493	11.132	22.874	42.052	30.442	52.591
Begoniaceae	18.605	12.939	22.769	41.958	34.412	48.043
Tetramelaceae	18.996	11.182	34.572	41.958	34.412	48.043
Thomandersiaceae	14.824	10.608	20.578	41.931	33.390	50.547
Schlegeliaceae	15.090	9.392	21.392	41.931	33.390	50.547
Mazaceae	18.017	11.121	22.960	40.828	32.633	48.464
Smilacaceae	14.896	9.686	22.790	38.933	22.552	54.302
Posidoniaceae	15.206	10.894	21.528	38.869	31.139	44.901
Asteliaceae	16.985	11.657	22.752	38.582	32.169	45.726
Zosteraceae	17.827	11.915	22.933	36.040	28.576	41.641
Potamogetonaceae	17.160	9.983	22.617	36.040	28.576	41.641
Talinaceae	15.969	10.714	22.937	35.238	19.995	49.587
Cymodoceaceae	21.244	11.141	33.177	34.690	28.375	41.040
Ruppiaceae	12.957	10.001	16.224	34.690	28.375	41.040
Basellaceae	14.157	10.028	19.480	34.657	19.861	53.207
Joinvilleaceae	13.244	9.573	17.847	34.650	30.226	39.428

Phrymaceae	21.030	11.893	34.305	34.582	25.684	42.744
Argophyllaceae	15.082	9.370	20.695	34.400	21.127	45.793
Phellinaceae	13.410	9.797	18.797	34.400	21.127	45.793
Liliaceae	26.160	15.634	39.532	33.485	19.271	48.887
Resedaceae	17.162	12.340	22.926	33.134	22.275	44.712
Gyrostemonaceae	13.610	10.047	18.948	33.134	22.275	44.712
Tropaeolaceae	13.712	9.402	19.034	32.301	19.094	48.696
Akaniaceae	15.504	10.876	22.315	32.301	19.094	48.696
Tetrameristaceae	17.659	11.902	22.679	32.098	19.327	47.911
Costaceae	14.140	10.030	18.295	31.448	27.864	35.732
Zingiberaceae	15.564	11.813	21.662	31.448	27.864	35.732
Portulacaceae	14.799	10.434	21.337	30.895	19.029	43.338
Lowiaceae	13.840	9.756	17.763	30.772	21.962	45.127
Strelitiziaceae	18.079	10.801	25.593	30.772	21.962	45.127
Asphodelaceae	19.314	15.291	22.980	30.752	27.750	33.931
Didieraceae	16.973	10.057	34.545	29.259	15.771	45.769
Dipterocarpaceae	14.240	9.918	20.767	27.832	18.258	38.083
Hypoxidaceae	14.351	11.621	19.193	27.436	19.343	36.263
Crypteroniaceae	16.944	12.012	21.454	27.330	20.307	38.490
Amaryllidaceae	16.871	13.249	21.951	25.960	19.830	30.982
Asparagaceae	20.367	13.697	26.675	25.960	19.830	30.982
Marantaceae	18.166	14.792	22.333	25.489	18.473	37.660
Cannaceae	13.295	9.724	18.120	25.489	18.473	37.660
Poaceae	14.633	11.672	19.778	25.389	16.236	34.990
Orobanchaceae	19.386	14.090	28.048	23.934	16.983	35.328
Crossosomataceae	13.950	10.045	21.119	23.380	15.205	48.954
Stachyuraceae	13.186	9.860	17.110	23.380	15.205	48.954

Capparaceae	13.947	9.274	17.435	22.702	17.485	29.081
Philesiaceae	12.938	9.732	16.834	22.057	15.688	33.493
Rhipogonaceae	13.642	9.880	19.017	22.057	15.688	33.493
Anacampserotaceae	14.207	9.637	21.064	21.568	14.537	32.973
Cactaceae	13.521	10.500	19.306	21.568	14.537	32.973
Cistaceae	15.512	10.959	19.700	21.459	14.443	30.095
Sarcolaenaceae	13.205	9.590	16.974	21.459	14.443	30.095
Penaeaceae	17.006	12.202	21.561	21.111	17.217	27.139
Brassicaceae	13.191	9.750	17.096	18.942	13.323	21.690
Cleomaceae	13.912	10.639	18.704	18.942	13.323	21.690

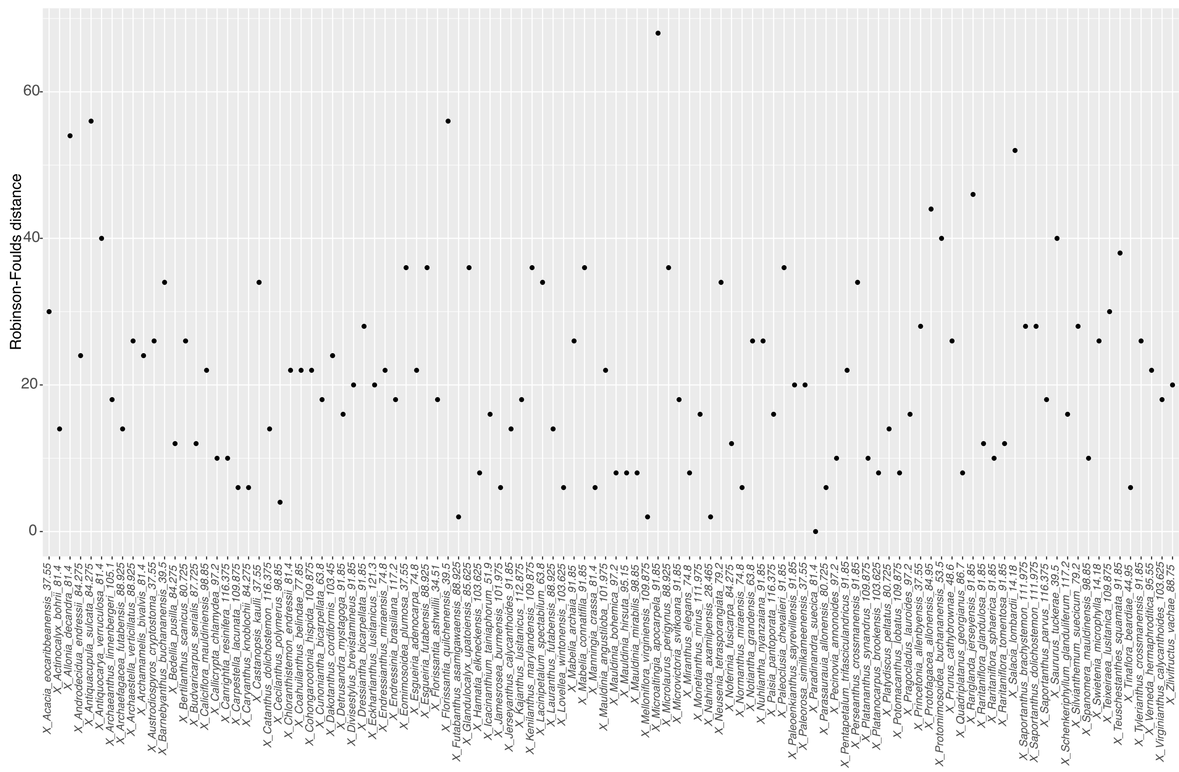

Figure S1. Robinson-Foulds distance between the undated tree (ML tree) and the maximum clade credibility tree under the unconstrained skyline fossilized birth-death model approach.

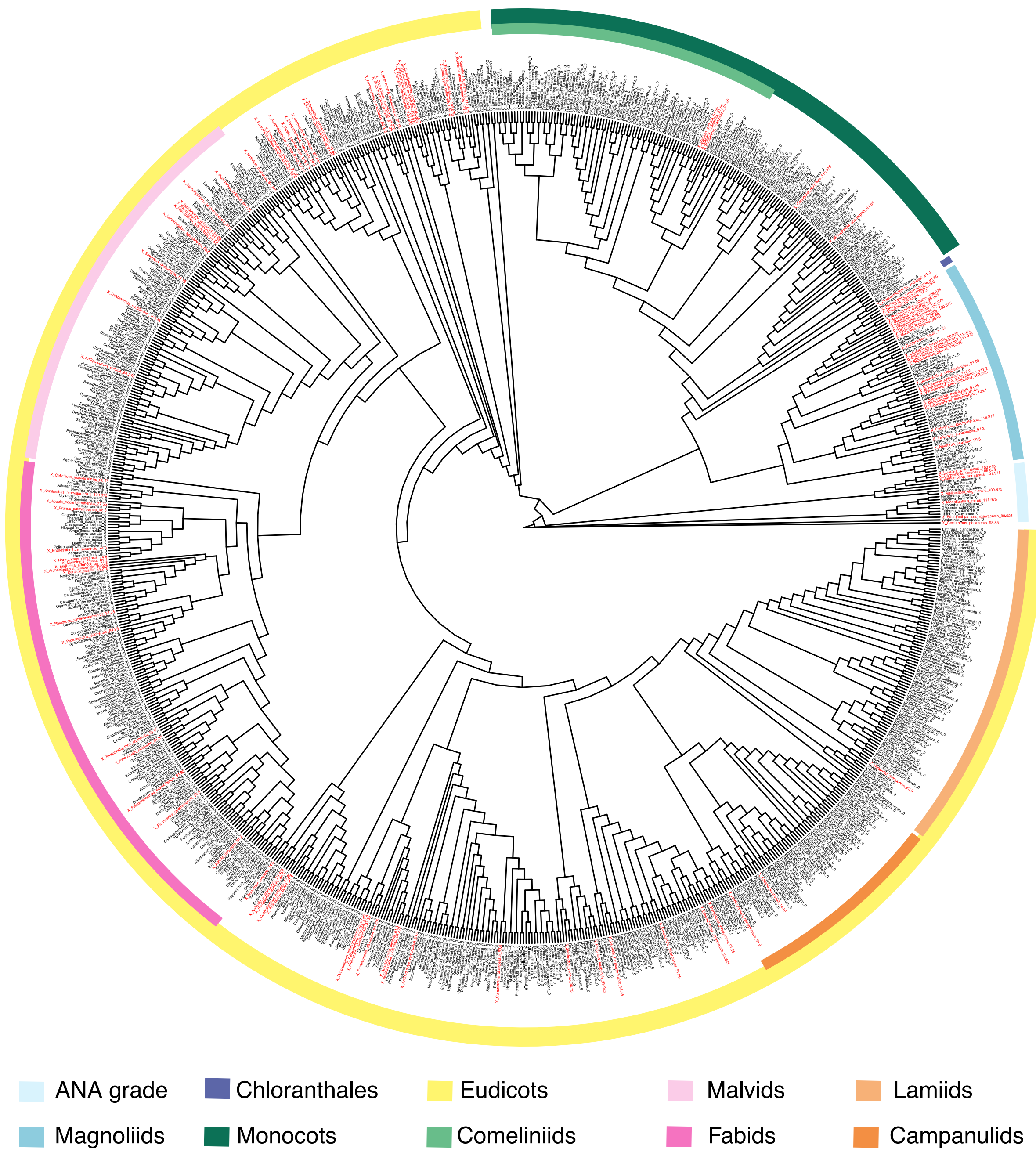

Figure S2. Maximum likelihood tree indicating the position of 792 extant species (labels in black) and 108 fossil flowers (labels in red).

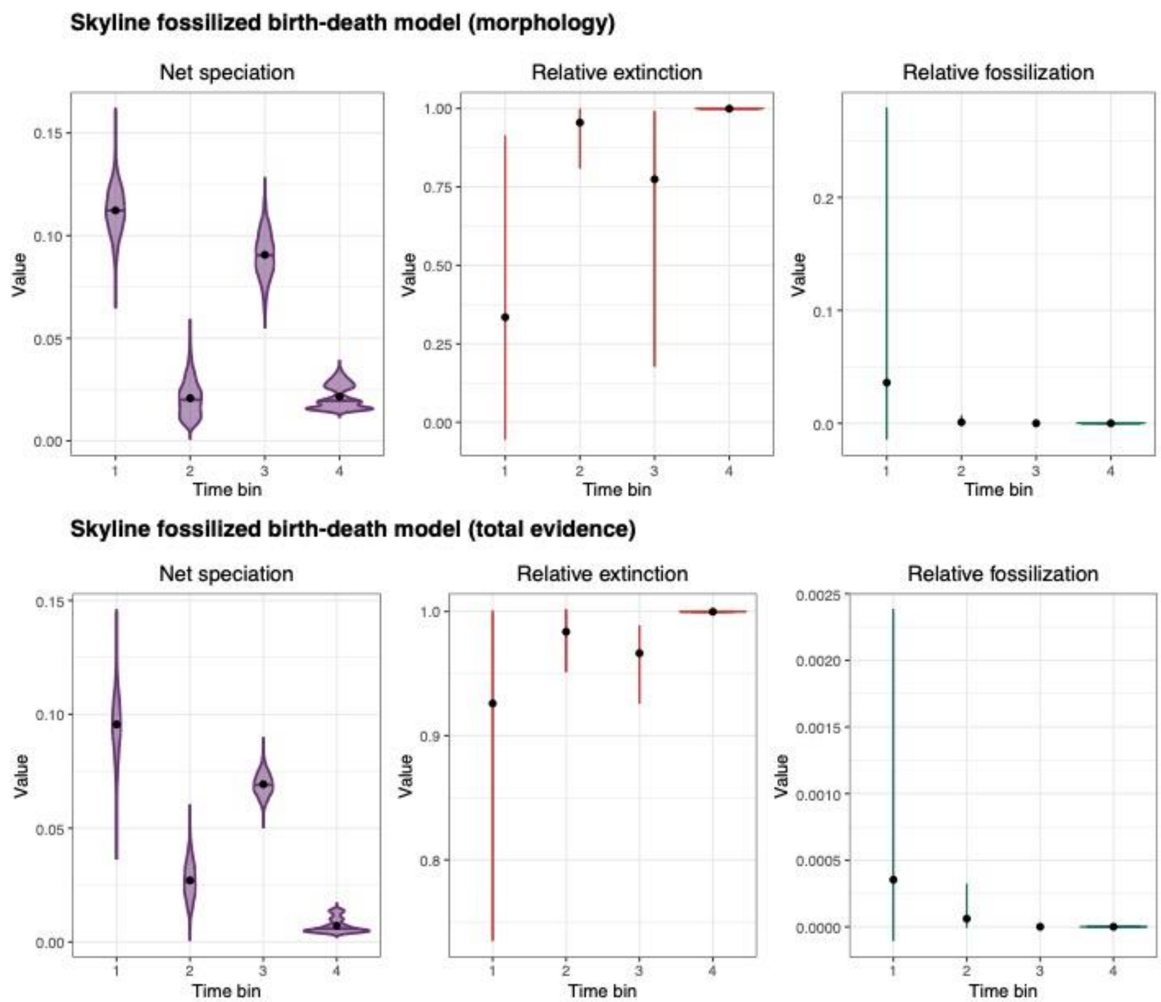

Figure S3. Visualization of posterior estimates for diversification parameters for every time bin under the constrained skyline fossilized birth death model with morphology-only and total evidence approaches.

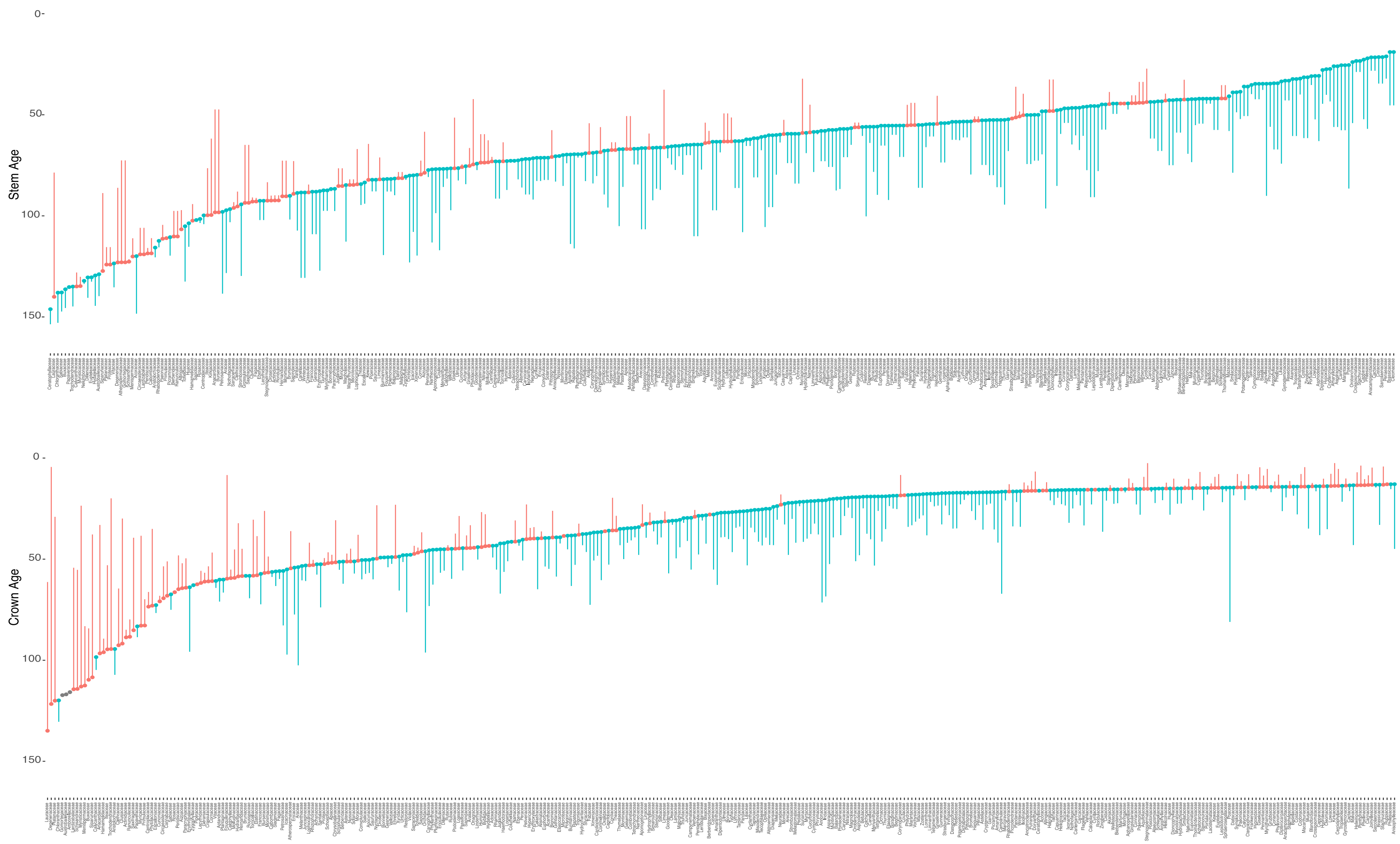

Figure S5. Comparison of the mean stem and crown ages between the SFBD (total-evidence) and node-dating from Ramírez-Barahona et al. (2020). Points correspond to mean ages obtained from Ramírez-Barahona et al. (2020), end line is the mean age obtained in this study. Red lines older estimates; blue lines younger estimates.

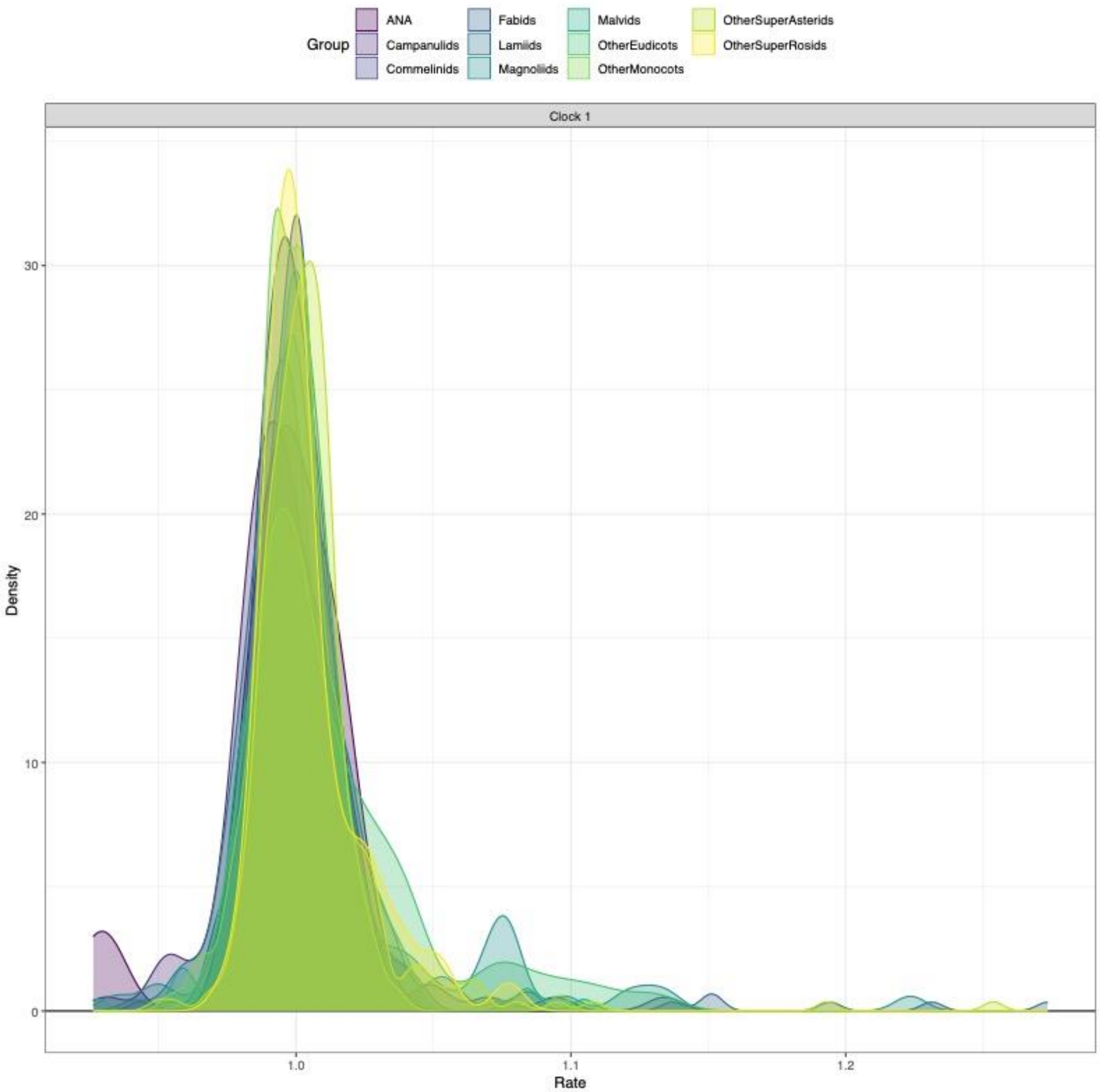

Figure S6. Densities of morphological rates across major angiosperm groups.

DISCUSIÓN GENERAL Y CONCLUSIONES

Los datos morfológicos y/o moleculares de especies vivientes y los avances recientes en el campo de la sistemática filogenética han permitido explorar cómo ha cambiado la dinámica de diversificación en angiospermas entre linajes (Smith et al., 2011; Magallón et al., 2019), reconstruir la flor ancestral y la variación en las tasas de cambio de caracteres (O´Meara et al., 2016; Sauquet et al., 2017; Herting et al., 2023), e inferir patrones espaciales de diversidad (Dimitrov et al., 2023). Sin embargo, los estudios basados en simulaciones han indicado que la falta de integración de datos del registro fósil ha limitado nuestras inferencias sobre procesos macroevolutivos a través del tiempo (Quental y Marshall, 2010). Los fósiles proveen una fuente de información directa sobre los procesos y patrones evolutivos de la vida en la Tierra, como la aparición de nuevas morfologías, su distribución espacial, y la variación en la dinámica de origen y extinción de especies (Benton et al., 2000; Foote, 2003). Sin embargo, los paleobotánicos se han enfocado a la rigurosa descripción de nuevas especies fósiles y la determinación de su afinidad taxonómica. Mientras que, en estudios filogenéticos, el uso del registro fósil se ha limitado a ser una fuente de información temporal para fechar árboles.

En la presente tesis se investigaron varios aspectos de la evolución de las angiospermas mediante la integración excepcional de flores fósiles e información masiva de especies vivientes. Un primer paso para la integración de fósiles en estudios macroevolutivos fue inferir su posición filogenética (*Capítulo I*). De manera importante, la serie de experimentos analíticos presentados en los *Capítulos I* y *III* permitieron conocer de manera más detallada las limitaciones y sesgos de los datos, modelos y métodos en los estimados obtenidos, lo cual representa una herramienta importante para futuras investigaciones en paleontología y sistemática filogenética. Adicionalmente, aporto información relevante sobre la estimación de los tiempos de origen de las familias de angiospermas mediante la integración directa del registro fósil y de especies vivientes en único proceso macroevolutivo (modelo de nacimiento y muerte fosilizado). Finalmente, se cuantificó, por primera vez, la diversidad morfológica (disparidad) y su variación entre linajes y a través del tiempo. Se destaca la importancia del registro fósil en la estimación de patrones evolutivos a gran escala que no pueden ser recuperados al analizar unicamente datos de especies vivientes. Como, por ejemplo, la presencia de combinaciones morfológicas que ya no son observadas en especies vivientes, y también permitió inferir un patrón de disminución en la disparidad a través del tiempo. Esta contribución aunada a diversos estudios (principalmente de invertebrados) abre un camino hacia una perspectiva macroevolutiva en la evolución de los organismos vivos.

Los análisis a gran escala filogenética como herramienta para el estudio de flores fósiles

En el primer *Capítulo*, ubiqué fósiles de flores en la filogenia de angiospermas, implementando un enfoque de análisis a gran escala, que considera información de representantes de todas las familias vivientes. Tradicionalmente, la asignación taxonómica de flores fósiles se ha limitado al estudio detallado de sus atributos morfológicos y a la identificación de apomorfías con linajes vivientes, pero aún son escasos los estudios que han evaluado explícitamente su posición filogenética. Los estudios que han analizado la posición filogenética de flores fósiles, en su mayoría se han basado en la utilización de matrices morfológicas que, por lo general, solo consideran representantes de una familia y/o un orden específico. Este enfoque es muy limitado, ya que no permite que el fósil pueda tener posiciones alternativas o afinidad a otros linajes no considerados *a priori* por el experto paleontólogo. Considero que gran parte de esta problemática se debe a la falta de datos morfológicos para un muestreo amplio de especies vivientes.

Nuestros resultados usando un análisis a gran escala fueron robustos y de manera general, altamente congruentes con asignaciones taxonómicas previas, lo que valida el magnífico trabajo taxonómico llevado a cabo por expertos paleobotánicos en el estudio de la diversidad flores fósiles desde hace más de 40 años. Sin embargo, es importante mencionar que la mayoría de los fósiles fueron ubicados con alta incertidumbre, y además la informatividad filogenética de los caracteres es limitada, específicamente para fósiles que muestran combinaciones de caracteres estructurales generalizados de grandes clados (e.g., familias dentro de clado Pentapetalae). Algunas limitaciones en la ubicación de los fósiles dentro de grandes clados se debe los caracteres morfológicos analizados, ya que corresponden a atributos de las estructuras que, por lo general, se encuentran altamente conservados (estables) a niveles filogenéticos profundos (e.g., dentro de familias). Aunque considero que el mejoramiento de la matriz morfológica --en términos del aumento de especies y caracteres-- puede ayudar a ubicar a los fósiles de manera más precisa, es importante destacar que la postulación de homología y la definición de caracteres aplicables para todas las especies de angiospermas debe ser muy cuidadosa, ya que la información derivada de estudios del desarrollo floral que nos pueden informar sobre la homología y los mecanismos que intervienen en la forma floral en antesis sólo está disponible para un número muy reducido de especies de angiospermas.

Por otra parte, los diferentes experimentos analíticos que hice revelaron que la incertidumbre con la que un fósil es asociado a una rama particular varía según el tipo de datos analizados. En otras palabras, el nivel de incertidumbre varía si se usa evidencia total o sólo los datos morfológicos, según el criterio de optimización, y el número de fósiles muestreados.

Los análisis basados en la matriz combinada de datos moleculares y morfológicos muestran que puede ser problemático el utilizar modelos en los que las longitudes de rama están ligadas entre particiones morfológicas y moleculares. Esta observación se debe a que algunos fósiles fueron asociados a ramas largas en el árbol, que correspondían a especies mico-heterótrofas y parásitas caracterizadas por tener altas tasas de substitución (cambio) molecular. Considero que el sesgo metodológico que genera la heterogeneidad en las tasas de substitución debe ser estudiado exhaustivamente, utilizando simulaciones y/o bases de datos con un muestreo de especies reducido, ya que podría ser aplicable a cualquier otro grupo de fósiles, no solo angiospermas.

Cambio morfológico a través del tiempo y entre linajes

Los patrones de disparidad morfológica son cruciales para documentar la historia evolutiva de los clados e inferir restricciones o selección en determinadas morfologías (Foote, 1994). Inicialmente, los estudios de disparidad se llevaron a cabo para caracterizar el origen evolutivo de los planos estructurales en animales fósiles, principalmente invertebrados (Guillerme et al., 2020). Así, hay relativamente pocos estudios de disparidad en plantas (e.g., Lupia, 1999, Chartier et al., 2014; Oyston et al., 2016).

Mediante el análisis de un muestreo amplio de especies vivientes y fósiles, se cuantificó la disparidad floral y la ocupación en el morfoespacio teórico a través del tiempo y entre los linajes principales de angiospermas. Los principales resultados de esta investigación indican que la mayor diversidad en la estructura floral fue alcanzada al inicio de la diversificación de angiospermas (*Capítulo II*; Fig. 1). Paralelamente, nuestro estudio independiente enfocado en tasas de evolución morfológica es congruente con esta hipótesis, ya que se detectaron las tasas más altas de cambio morfológico en el Cretácico temprano (*Capítulo III*; Fig. 2). Un aspecto interesante es que esta mayor exploración del morfoespacio ocurrió en un momento donde la riqueza de especies de angiospermas era aún baja comparada con otros linajes de plantas vivientes en ese periodo (Lupia, 1999). Adicionalmente, el decremento en la disparidad (estabilidad morfológica) en el Cretácico tardío (*Capítulo II*; Fig. 1; *Capítulo III*; Fig. 2) podría estar asociada a la interacción con animales polinizadores, ya que los fósiles muestran adaptaciones morfológicas que han sido asociadas a la polinización por animales, como es la diferenciación del perianto, y fusión de órganos florales.

Dado que encontramos una correlación negativa entre la disparidad morfológica y la riqueza de especies (*Capítulo II*; Fig. 4), propongo que la canalización (i.e., baja diversidad) en la estructura floral pudo haber influido en la variación en otros niveles estructurales de la flor (e.g., arquitectura, estilo “style”; *sensu* Endress, 1994), y por consiguiente en la diversificación en términos de número de especies. Considero que nuestros estimados de disparidad entre linajes vivientes no cambiarán drásticamente si incrementamos el número de especies, pero es posible que existan variaciones en los patrones de disparidad morfológica si analizamos caracteres de otros niveles estructurales de la flor.

El modelo de nacimiento y muerte fosilizado aplicado a bases de datos extensas

En esta tesis presento un estudio exploratorio del proceso de nacimiento y muerte fosilizado (FBD) aplicado a un amplio muestreo de especies de angiospermas. El enfoque de evidencia total considera tanto datos moleculares, solo para las especies vivientes, como datos morfológicos, estos si tanto para las especies vivientes como para las fósiles. Estudios previos basados en simulaciones sugieren que los estimados de edad obtenidos bajo el enfoque de evidencia total son más precisos y exactos (Gavryuskina et al., 2014; Barido-Sottani et al., 2023). Sin embargo, la aplicación de datos reales y modelos complejos, que incluyen la estimación de parámetros de tasas de diversificación, fosilización, y probabilidad de muestreo en el análisis de matrices de datos grandes, enfrenta muchos desafíos.

Los resultados indican que los estimados de edad son sensibles al tipo de datos utilizados y a la selección de fósiles. Aunque por lo general, el aumento de taxa fósiles incrementa la exactitud de los estimados, es necesario el establecimiento de un criterio adecuado para la inclusión de los fósiles. Nuestros resultados sugieren que el modelo puede presentar problemas para ubicar la posición filogenética de los fósiles, por lo tanto, es necesario mejorar la matriz de datos morfológica o constreñir los fósiles *a priori*. Una herramienta útil para constreñir la posición de los fósiles es a través de la visualización de la incertidumbre topológica mediante los *RoguePlots*. Adicionalmente, detectamos que los clados que no cuentan con representación fósil en la matriz de evidencia total resultaron en edades significativamente más jóvenes, por lo que es necesaria la aplicación de puntos de calibración internos.

El proceso de fechamiento y muerte fosilizado es un modelo complejo pero altamente prometedor para inferir filogenias y parámetros macroevolutivos de especies vivientes y fósiles. Sin embargo, aún empezamos a entender las limitaciones y sesgos del modelo, por lo que muchos estimados derivados de estudios empíricos deben ser considerados con cautela. Es curioso como varios estudios enfocados en entender las limitaciones del modelo basados en simulaciones, no consideran problemas inherentes del uso de datos empíricos. Considero que es necesario llevar a cabo simulaciones más realistas y útiles para los que hacemos estudios empíricos, que consideren muestreos taxonómicos y matrices morfológicas que pueden llevar a ser altamente incompletos --

más específicamente para fósiles-- así como los sesgos en la preservación de los fósiles entre linajes y a través del tiempo.

Perspectivas

Futuras investigaciones desde una perspectiva macroevolutiva, como son sobre la inferencia de patrones de evolución morfológica y la dinámica de diversificación e historia biogeográfica, deberán basarse en el análisis conjunto de especies fósiles y vivientes.

Considero que el mejor enfoque para la integración de fósiles, al menos un marco de referencia filogenético, es a través de la construcción y exploración de matrices morfológicas con amplia representación de especies vivientes. La determinación de caracteres morfológicos florales aplicables para todas las especies de angiospermas, incluyendo representantes fósiles, así como la codificación y curación de caracteres utilizados en esta tesis, ha sido el resultado de un esfuerzo de colaboración internacional conceptualizado hace más de 10 años como parte del proyecto eFLOWER (https://eflower.myspecies.info/).

Creo que esta base de datos morfológica constituye un elemento fundamental con un gran potencial para futuras investigaciones que involucren la evaluación de la posición filogenética de flores fósiles y su integración en estudios macroevolutivos. Sin embargo, es necesario incrementar el número de taxa vivientes y de caracteres morfológicos que proporcionen informatividad filogenética a niveles más someros de la filogenia.

Una de las limitaciones de estos estudios es la disponibilidad de información morfológica para especies vivientes y fósiles. Me impresionó cómo para muchas de las especies con las que contábamos con información de datos moleculares, no se encontró información detallada de sus atributos florales, específicamente de su morfología interna, por ejemplo para la placentación.

Por otra parte, el registro fósil actualmente disponible de las angiospermas se encuentra geográfica y temporalmente sesgado. Por ejemplo, la mayoría de las fósiles de flores se han encontrado en localidades del Cretácico del este de Estados Unidos, Portugal, Suecia y Japón. Adicionalmente, existen varios fósiles de flores que no han sido publicados formalmente, debido a la falta de caracteres distintivos para ser asignados a un grupo taxonómico en particular. A pesar de que existen varias descripciones de fósiles correspondientes al Cenozoico, el bajo nivel de preservación y/o la falta de fotografías claras o descripciones detalladas de este periodo, así que no nos permitió incluirlos en los análisis. La descripción de más fósiles de afinidades taxonómicas desconocidas, así como los nuevos descubrimientos, especialmente de los tiempos cercanos al origen del grupo de las angiospermas, serán importantes para determinar patrones macroevolutivos en la evolución temprana de las angiospermas.

Conclusiones

- El enfoque usado en esta tesis para evaluar la posición de flores fósiles a gran escala filogenética derivado de una matriz morfológica que incluye representantes vivientes de todas las familias de angiospermas, resultó en una herramienta útil para futuros estudios paleobotánicos que involucren la descripción de nuevas especies fósiles y para corroborar asignaciones taxonómicas previas.
- Distintas estrategias analíticas pueden resultar en diferencias en los niveles de incertidumbre en la ubicación de la posición filogenética de los fósiles. Se recomienda enfáticamente incorporar esta incertidumbre al momento de utilizar los fósiles en análisis subsecuentes de fechamiento con reloj molecular.
- El nivel de incertidumbre está asociada a la señal filogenética de los caracteres morfológicos y no a la cantidad de datos faltantes (o no aplicables). La expansión de la matriz morfológica --no sólo en términos de número de especies muestreadas, sino al número y tipo de caracteres morfológicos-- podría mejorar la precisión de la estimación topológica.
- Se documentaron los patrones de evolución en la disparidad floral de angiospermas a través del tiempo y entre linajes. Se propone que las angiospermas alcanzaron su mayor diversidad morfológica temprano en su evolución. La comparación de la variación en la disparidad entre linajes principales de angiospermas vivientes indicó que los miembros del grado ANA, el clado

Magnolidae, y el grado de eudicotiledóneas tempranamente divergentes, mostraron mayor diversidad de la estructura floral.

- El morfoespacio teórico permite describir el espectro completo de las posibles combinaciones morfológicas para un grupo de organismos. Las flores de angiospermas observadas en la naturaleza (tanto fósiles como vivientes) ocupan un área restringida del total del morfoespacio teóricamente posible, lo que nos podría indicar algunas limitaciones (restricciones) en la estructura floral y áreas de baja o alta adecuación.
- El enfoque de fechamiento de puntas bajo el modelo de nacimiento y muerte fosilizado nos permitió estimar los tiempos de divergencia de los nodos troncales y corona de las familias de angiospermas y las tasas de cambio morfológico. Nuestros resultados preliminares sugieren un origen más joven para los nodos corona de la mayoría de las familias de angiospermas comparado con estudios preliminares utilizando un enfoque de fechamiento de nodos.
- Encontramos una mayor heterogeneidad en las tasas de cambio morfológico en las ramas más profundas de la filogenia de angiospermas, que corresponden al origen de los clados principales de angiospermas.

REFERENCIAS BIBLIOGRÁFICAS

APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society 181: 1–20.

Barido-Sottani J, Pohle A, de Baets K, Murdock D, Warnoch R. 2023. Putting the F into FBD analysis: tree constraints or morphological data? Paleontology: e12679.

Benton MJ, Wills MA, Hitchin R. 2000. Quality of the fossil record through time. Nature 403: 534–537.

Chartier M, Jabbour F, Gerber S, Mitteroecker P, Sauquet H, von Balthazar M, Staedler Y, Crane PR, Schönenberger J. 2014. The floral morphospace – a modern comparative approach to study angiosperm evolution. New Phytologist 204: 841–853.

Coiro M, Allio R, Mazet N, Seyfullah LJ, Condamine FL. 2023. Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity. New Phytologist 240: 1616-1635.

Crepet WL, Niklas KJ. 2009. Darwin's second "abominable mystery": Why are there so many angiosperm species? American Journal of Botany 96: 366–381.

Doyle JA, Endress PK. 2010. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. Journal of Systematic and Evolution 48: 1–35.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4: e88.

Endress PK. 1994. Diversity and evolutionary biology of tropical flowers. Cambridge, UK: Cambridge University Press.

Endress PK. 2010. Flower structure and trends of evolution in eudicots and their major subclades. Annals of the Missouri Botanical Garden 97: 541–583.

Endress PK. 2011. Evolutionary diversification of the flowers in angiosperms. American Journal of Botany 98: 370–396.

Foote M. 1992. Paleozoic record of morphological diversity in blastozoan echinoderms. Proceedings of the National Academy of Sciences 89: 7325–7329.

Foote M. 1994. Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. Paleobiology 20: 320–344.

Foote M. 1999. Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. Paleobiology 25: 1-115.

Foote M. 2003. Origination and extinction through the phanerozoic: a new approach. The Journal of Geology 111: 125-148.

Friis EM, Crane PR, Pedersen KR. 2011. Early flowers and angiosperm evolution. Cambridge: Cambridge University Press.

Friis EM, Marone F, Pedersen KR, Crane PR, Stampanoni M. 2014. Three-dimensional visualization of fossil flowers, fruits, seeds, and other plant remains using synchrotron radiation X-ray tomographic microscopy (SRXTM): new insights into Cretaceous plant diversity. Journal of Paleontology 88: 684–701.

Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D, Drummond AJ. 2017. Bayesian total-evidence dating reveals the recent crown radiation of penguins. Systematic Biology 66: 57–73.

Gould SJ. 1989. Wonderful Life. New York: W. W. Norton.

Guillerme T, Cooper N, Brusatte SL, Davis KE, Jackson AL, Gerber S, Goswami A, Healy K, Hopkins MJ, Jones MEH et al. 2020. Disparities in the analysis of morphological disparity. Biology Letters 16: 20200199.

Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. Proceedings of the National Academy of Sciences 111: E2957–E2966.

Herendeen PS, Friis EM, Pedersen KR, Crane PR. 2017. Palaeobotanical redux: Revisiting the age of the angiosperms. Nature Plants 3: 1–8.

Herting J, Schönenberger J, Sauquet H. 2023. Profile of a flower: How rates of morphological evolution drive floral diversification in Ericales and angiosperms. American Journal of Botany 110: e16213.

Hickey LJ, Doyle JA. 1977. Early Cretaceous fossil evidence for angiosperm evolution. The Botanical Review 43: 3–104.

Ho SYW, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58: 367–380.

Hughes M, Gerber S, Wills MA. 2013. Clades reach highest morphological disparity early in their evolution. Proceedings of the National Academy of Sciences 110: 13875–13879.

Leslie AB, Simpson C, Mander L. 2021 Reproductive innovations and pulsed rise in plant complexity. Science 373: 1368–1372.

Lidgard S, Crane PR. 1990. Angiosperm diversification and Cretaceous floristic trends: a comparison of palynofloras and leaf macrofloras. Paleobiology 16: 77–93.

Lloyd G.T., Slater G.J. 2021. A total-group phylogenetic metatree for Cetacea and the importance of fossil data in diversification analyses. Systematic Biology 70: 922–939.

López-Martínez AM, Schönenberger J, von Balthazar M, Gónzalez-Martínez CA, Ramírez-Barahona S, Sauquet H, Magallón S. 2023. Integrating fossil flowers into the angiosperm phylogeny using molecular and morphological evidence. Systematic Biology 72: 837-855.

Lupia R. 1999. Discordant morphological disparity and taxonomic diversity during the Cretaceous angiosperm radiation: North American pollen record. Paleobiology 25: 1–28.

Magallón S, Sánchez-Reyes LL, Gómez-Acevedo SL. 2019. Thirty clues to the exceptional diversification of flowering plants. Annals of Botany 123: 491–503.

Marshall CR. 2008. A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. The American Naturalist 171: 726–742.

Marshall CR. 2017. Five palaeobiological laws needed to understand the evolution of the living biota. Nature Ecology & Evolution 1: 1–6.

O'Meara BC, Smith SD, Armbruster WS, Harder LD, Hardy CR, Hileman LC, Hufford L, Litt A, Magallón S, Smith SA, Stevens PF, Fenster CB, Diggle PK. 2016 Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. Proceedings of the Royal Society B 283: 20152304.

Oyston JW, Hughes M, Wagner PJ, Gerber S, Wills MA. 2015. What limits the morphological disparity of clades? Interface Focus 5: 20150042.

Oyston JW, Hughes M, Gerber S, Wills MA. 2016. Why should we investigate the morphological disparity of plant clades? Annals of Botany 117: 859–879.

Prentice KC, Ruta M, Benton MJ. 2011. Evolution of morphological disparity in pterosaurs. Journal of Systematic Palaeontology 9: 337–353.

Raup DM, Sepkoski J. Jr. 1982. Mass extinctions in the marine fossil record. Science 215: 1501–1503.

Sauquet H, Von Balthazar M, Magallón S, Doyle JA, Endress PK, Bailes EJ, Barroso De Morais E, Bull-Hereñu K, Carrive L, Chartier M, Chomicki G, Coiro M, Cornette R, El Ottra JHL, Epicoco C, Foster CSP, Jabbour F, Haevermans A, Haevermans T, Hernández R, Little SA, Löfstrand S, Luna JA, Massoni J, Nadot S, Pamperl S, Prieu C, Reyes E, Dos Santos P, Schoonderwoerd KM, Sontag S, Soulebeau A, Staedler Y, Tschan GF, Wing-Sze Leung A, Schönenberger J. 2017. The ancestral flower of angiosperms and its early diversification. Nature Communications 8: 16047.

Sauquet H, Magallón S. 2018: Key questions and challenges in angiosperm macroevolution. New Phytologist 219: 1170–1187.

Sauquet H, Ramírez-Barahona S, Magallón S. 2022. What is the age of flowering plants? Journal of Experimental Botany 73: 3840–3853.

Schönenberger J. 2005. Rise from the ashes - the reconstruction of charcoal fossil flowers. Trends in Plant Science. 10: 436–443.

Schönenberger J, von Balthazar M, López Martínez A, Albert B, Prieu C, Magallón S, Sauquet H. 2020. Phylogenetic analysis of fossil flowers using an angiosperm-wide data set: proof-of-concept and challenges ahead. American Journal of Botany 107: 1433–1448.

Silvestro D, Salamin N, Schnitzler J. 2014. PyRate: A new program to estimate speciation and extinction rates from incomplete fossil data. Methods in Ecology and Evolution 5: 1126–1131.

Simões TR, Pierce SE. 2021. Sustained high rates of morphological evolution during the rise of tetrapods. Nature Ecology & Evolution 5: 1403-1414.

Stephen SA, Beaulieu J, Stamatakis A, Donoghue MJ. 2011. Understanding angiosperm diversification using small and large phylogenetic trees. American Journal of Botany 98: 404–414.

van der Niet T, Johnson SD. 2012. Phylogenetic evidence for pollinator-driven diversification of angiosperms. Trends in Ecology & Evolution 27: 353–361.

Zhang C, Stadler T, Klopfstein S, Heath TA, Ronquist F. 2016. Total- evidence dating under the fossilized birth-death process. Systematic Biology 65: 228–249.

Zhang C, Ronquist F, Stadler T. 2023. Skyline fossilized birth-death model is robust to violations of sampling assumptions in total-evidence dating. Systematic Biology: syad054.

Zhang Q, Ree RH, Salamin N, Xing Y, Silvestro D. 2021. Fossil-informed models reveal a boreotropical origin and divergent evolutionary trajectories in the walnut family (Juglandaceae). Systematic Biology 71: 242–258.