New Phytologist

Research

Macroevolutionary inference of complex modes of chromosomal speciation in a cosmopolitan plant lineage

Carrie M. Tribble[1,2,3]*, José Ignacio Márquez-Corro[4,5]*, Michael R. May[6], Andrew L. Hipp[7], Marcial Escudero[8†] and Rosana Zenil-Ferguson[9†]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA; [2]Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA; [3]School of Life Sciences, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA; [4]Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK; [5]Department of Molecular Biology and Biochemistry Engineering, Universidad Pablo de Olavide, Sevilla, 41013, Spain; [6]Department of Evolution and Ecology, University of California Davis, Davis, CA, USA; [7]Herbarium and Center for Tree Science, The Morton Arboretum, Lisle, IL 60532, USA; [8]Department of Plant Biology and Ecology, Faculty of Biology, University of Sevilla, Sevilla, 41012, Spain; [9]Department of Biology, University of Kentucky, Lexington, KY 40506, USA

Author for correspondence:
Carrie M. Tribble
Email: tribblec@uw.edu

Received: *4 November 2024*
Accepted: *28 November 2024*

New Phytologist (2024)
doi: 10.1111/nph.20353

Key words: *Carex*, chromosome number, diversification, dysploidy, macroevolution, macroevolutionary process variation, RevBayes, TensorPhylo.

Summary

- The effects of single chromosome number change—dysploidy – mediating diversification remain poorly understood. Dysploidy modifies recombination rates, linkage, or reproductive isolation, especially for one-fifth of all eukaryote lineages with holocentric chromosomes. Dysploidy effects on diversification have not been estimated because modeling chromosome numbers linked to diversification with heterogeneity along phylogenies is quantitatively challenging.
- We propose a new state-dependent diversification model of chromosome evolution that links diversification rates to dysploidy rates considering heterogeneity and differentiates between anagenetic and cladogenetic changes. We apply this model to *Carex* (Cyperaceae), a cosmopolitan flowering plant clade with holocentric chromosomes.
- We recover two distinct modes of chromosomal evolution and speciation in *Carex*. In one diversification mode, dysploidy occurs frequently and drives faster diversification rates. In the other mode, dysploidy is rare, and diversification is driven by hidden, unmeasured factors. When we use a model that excludes hidden states, we mistakenly infer a strong, uniformly positive effect of dysploidy on diversification, showing that standard models may lead to confident but incorrect conclusions about diversification.
- This study demonstrates that dysploidy can have a significant role in speciation in a large plant clade despite the presence of other unmeasured factors that simultaneously affect diversification.

Introduction

Unveiling the primary drivers of diversification remains one of the most important goals in evolutionary biology (Sauquet & Magallón, 2018). Hundreds of studies have focused on estimating changes in plant diversification processes through time (Magallón & Castillo, 2009), across clades (Magallón *et al.*, 2019), or in association with trait evolution (Helmstetter *et al.*, 2023). Chromosome number changes and rearrangements are particularly likely to influence lineage diversification (Freyman & Höhna, 2018a). Yet, most plant diversification studies that address chromosome evolution focus only on the role of polyploidy, rather than on dysploidy – gains and losses of single chromosomes – which does not involve large DNA content changes (i.e. dysploidy, gains or losses of single chromosomes; Escudero *et al.*, 2014; Mandáková & Lysak, 2018). A recent review of trait-dependent diversification in angiosperms, for example, cites seven studies linking speciation and polyploidy (Helmstetter *et al.*, 2023), only one of which considered dysploidy linked to diversification (Freyman & Höhna, 2018a). While dysploidy has the potential to influence lineage diversification through effects on recombination or reproductive isolation, the macroevolutionary effects of dysploidy on speciation and extinction remain unknown, and far fewer macroevolutionary studies have focused on speciation and extinction as a result of dysploidy.

Dysploidy and holocentric chromosomes

There are two classic models of chromosomal speciation in a macroevolutionary context. In the hybrid-dysfunction model, dysploidy is linked to speciation events. Under this model,

*These are co-first authors.
†These are co-senior authors.

dysploidy causes an immediate reproductive barrier, as reproduction between individuals with different chromosome numbers would cause problems during meiosis. Thus, most – if not all – dysploidy events across a phylogeny would occur cladogenetically. Alternatively, the recombination–suppression model posits that chromosomal rearrangements may become fixed in lineages via either drift or selection (as some rearrangements may physically link adaptive loci or locally reduce recombination). Under this model, dysploidy would evolve primarily anagenetically since fissions and fusions do not necessarily restrict gene flow; rather, they evolve within populations and lineages (Baker & Bickham, 1986). The two above-described models are not mutually exclusive. In a recent review, Lucek *et al.* (2022) introduce a third option: the hybrid-dysfunction/recombination–suppression model, under which dysploidy evolves both anagenetically and cladogenetically (fig. 4 in Lucek *et al.*, 2022). While populations may be able to continue interbreeding despite some dysploidy events (which then may become fixed in the lineage), other dysploidy events may cause speciation, either because of an accumulation of differences that eventually leads to incompatibility or because of the genomic signature of the dysploidy event itself. We describe additional theory on dysploidy and macroevolution in Supporting Information Notes S1.

Holocentricity – having chromosomes without clear centromeres/primary constrictions – is distributed broadly across the Tree of Life, including 18 different lineages in animals, plants, and rhizaria (Escudero *et al.*, 2016; Márquez-Corro *et al.*, 2019), *c.* 15–20% of eukaryotic species (Márquez-Corro *et al.*, 2018). Two particularly diverse holocentric clades show extraordinary chromosome number variation: the insect order Lepidoptera (de Vos *et al.*, 2020) and the angiosperm sedge family Cyperaceae ($2n = 4$–224; Márquez-Corro *et al.*, 2019,2021). Holocentric chromosomes, instead of having kinetochore activity concentrated in a single point (i.e. at a single centromere in monocentric chromosomes), have centromeric regions distributed along the whole chromosome where the kinetochores assemble in most of holocentric organisms (Marques *et al.*, 2015; Márquez-Corro *et al.*, 2019). In monocentric chromosomes, many chromosome fissions are expected to result in a loss of genetic material during meiosis and inviable gametes, as chromosome fragments without centromeres are unable to segregate normally, and fusions are generally inherited by combining two telocentric chromosomes (Robertsonian translocations; Robertson, 1916). Holocentric chromosomes, by contrast, allow chromosome fragments to segregate normally during meiosis (Faulkner, 1972). Holocentricity may promote chromosome number variation via fission and fusion, as these changes are expected to be neutral or nearly so in holocentric organisms (Márquez-Corro *et al.*, 2019). Thus, holocentric organisms provide a unique system in which to study chromosomal speciation (Lucek *et al.*, 2022). Nonetheless, of the few applied macroevolutionary studies that address the role of dysploidy in lineage diversification, most have focused on monocentric chromosomes, which have a single centromere (e.g. Ayala & Coluzzi, 2005; Freyman & Höhna, 2018a, but see relevant work on Lepidoptera, e.g. de Vos *et al.*, 2020).

Carex – the largest genus in Cyperaceae – is particularly well suited to studying the effect of dysploidy on plant diversification because all *Carex* have holocentric chromosomes, the genus represents 40% of the third most species-rich monocot family (among the 10th in angiosperms, POWO, 2023), and well-developed phylogenetic and chromosome number datasets for the genus facilitate macroevolutionary studies (Martín-Bravo *et al.*, 2019; Márquez-Corro *et al.*, 2021). In *Carex*, karyotype evolves mainly through fusion, fission, and translocation, in contrast to other sedge lineages where karyotype evolves through both dysploidy and polyploidy (Márquez-Corro *et al.*, 2019; Elliott *et al.*, 2022; Shafir *et al.*, 2023). *Carex* also has exceptional variability in chromosome number, ranging from $2n = 10$ to $2n = 132$ (Márquez-Corro *et al.*, 2021). *Carex* has experienced several rapid radiations (Martín-Bravo *et al.*, 2019), and shifts in optimal chromosome number are thought to have played a role in some of these radiations (e.g. *Carex* sect. *Cyperoideae*; Hipp, 2007; Márquez-Corro *et al.*, 2021).

Previous studies that have tested for a correlation between dysploidy and diversification (e.g. Márquez-Corro *et al.*, 2021) have relied on models that fail to account for alternative sources of variation in diversification rates – for example, morphological traits, climatic niche, or biotic interactions – that are not the study's focal trait (in our case, chromosome number; Beaulieu & O'Meara, 2016). Models that fail to account for alternative sources of variation in diversification rates have high type-I error rates (Rabosky & Goldberg, 2015) because they misattribute underlying diversification-rate variation caused by unmeasured factors to variation due to the states included in the model. In other words, the null hypothesis of those models assumes that there is no underlying diversification-rate variation. This high error rate has motivated the development of hidden-state diversification models that account for underlying diversification-rate variation that is not caused by the focal trait (Beaulieu & O'Meara, 2016; Caetano *et al.*, 2018). These models not only more accurately test for associations between the focal trait and diversification but also provide an opportunity to model how focal evolutionary processes vary across the phylogeny. Here, we design a new model of joint chromosome evolution and lineage diversification that incorporates process variation in chromosome number evolution and disentangles the effects of dysploidy from other unobserved factors that may also affect diversification rates (described in detail later). We apply our model to the most recent *Carex* time-calibrated phylogeny with chromosome number information that contains over 700 taxa and > 50 states (Martín-Bravo *et al.*, 2019; Márquez-Corro *et al.*, 2021).

We test for a detectable effect of chromosome gains and losses on the rate of species formation in a large, diverse group of plants with high variation in chromosome number: *Carex* (Cyperaceae, Poales). We demonstrate that dysploidy is sometimes associated with faster rates of lineage formation, but diversification and dysploidy play out through two evolutionary modes across *Carex*. In one mode, chromosome rearrangements evolve rapidly and are linked to new species formation. In the other mode, rearrangements evolve less frequently and are not associated with new

species formation; instead, other factors – hidden states unmeasured in our analysis – likely drive the formation of new species. Our results demonstrate the complexity of the diversification process and the important role of genomic rearrangements in determining biodiversity patterns at geological timescales. Furthermore, we illustrate how computational and statistical advances in modeling permit increasingly nuanced models that better represent the underlying complexity of biological systems.

Materials and Methods

Modeling chromosome number evolution

We developed the Chromosome number and Hidden State-dependent Speciation and Extinction model (ChromoHiSSE) to test the interaction between single chromosome number change, speciation, extinction, and diversification heterogeneity. Our model can be considered a natural extension of the ChromoSSE model (Freyman & Höhna, 2018a). Under the ChromoSSE model, lineages evolve independently under a continuous-time Markov chain (CTMC) that describes changes in chromosome numbers, speciation events, and extinction events; each of these events occurs at a particular rate (interpreted as the expected number of events per lineage per unit time). The ChromoHiSSE model includes an additional hidden trait with $m > 1$ states. The states of this hidden trait correspond to different sets of ChromoSSE parameters, and lineages evolve among hidden states as a Markov process with an estimated rate.

Under the ChromoHiSSE model, the state of a lineage is both the chromosome number, n, and the hidden state, i. For numerical tractability (Mayrose *et al.*, 2010; Zenil-Ferguson *et al.*, 2017; Freyman & Höhna, 2018a), we place an upper bound on the possible number of chromosomes (k; transitions to $n > k$ are prohibited, that is have rate 0), and the lower bound is 0. We denote the two hidden states as i and ii. For example, the state for a lineage could be $10i$, indicating that the lineage has 10 haploid chromosomes and is in the hidden state i.

The ChromoHiSSE model is a stochastic process that begins with two lineages at the root, which evolve independently forward in time. As the process evolves, a lineage can experience anagenetic events (changes in chromosome number or hidden state that happen along the branches of the phylogeny, denoted with subscript a), cladogenetic events (speciation events that involve changes in the number of chromosomes or hidden state for one of the daughter lineages, therefore associated with speciation and denoted with subscript c), and extinction events. As in all continuous-time Markov chain models, each event happens at an instantaneous point in time, and multiple events cannot occur concurrently. The anagenetic events are as follows:

(**1**) n increases by one (increasing dysploidy) but the hidden state stays the same, which occurs at rate γ_{a_i};

(**2**) n decreases by one (decreasing dysploidy) but the hidden state stays the same, which occurs at rate δ_{a_i};

(**3**) n stays the same, but the hidden state changes (e.g. i to ii), which occurs at rate χ_a.

Cladogenetic events produce two daughter lineages that then evolve independently. The states of the daughters depend on the type of event:

(**1**) both daughters inherit the state of the ancestor, which occurs at rate φ_{c_i};

(**2**) one daughter inherits n, the other inherits $n + 1$ (increasing dysploidy), and both inherit the same hidden state i, which occurs at rate γ_{c_i};

(**3**) one daughter inherits n, the other inherits $n-1$ (decreasing dysploidy), and both inherit the same hidden state i, which occurs at rate δ_{c_i};

(**4**) both daughters inherit n from the ancestor, but one daughter changes hidden state (i to ii), which occurs at rate χ_c.

These events are depicted in Fig. 1. Additionally, all lineages go extinct at rate μ, independent of n or the hidden state. The lineages evolve forward in time until the present, at which point they are sampled independently with probability f. Extinct and unsampled lineages are pruned from the tree, and the hidden state for sampled lineages is ignored; this produces a realization comprising a reconstructed phylogeny relating the sampled lineages and a chromosome number for each sampled lineage. This stochastic process allows us to compute the probability of an observed dataset (i.e. the probability that a realization under this process corresponds to our observed data) given a set of parameter values.

Note that we assume that there are only two hidden states and that rates of change between hidden states are symmetric, that is, the anagenetic rate of change from i to ii is the same as the rate of change from ii to i, and likewise for cladogenetic changes. We do not model density-dependent dysploidy rates (rates varying based on the number of chromosomes), as previous studies across all angiosperms that have implemented density-dependent rates have found that the relationship between number of chromosomes and rates of chromosome evolution is very weak to nonexistent (Carta *et al.*, 2020). In addition, we assume that the rate of extinction is constant among all lineages, regardless of the number of chromosomes or hidden states. We assume a constant extinction rate because our model is biologically informed by our hypothesis that speciation is linked to chromosome number change, but that dysploid chromosome number changes have little effect on fitness and consequently little effect on extinction. In Notes S2, we show how to relax some of these assumptions by specifying a full ChromoHiSSE model. These specifications may guide future researchers who wish to include polyploidization (- and/or demipolyploidization); similarly, extinction could be allowed to vary in future implementations for empirical scenarios where such variation is warranted.

Testing differences between transition rates

Most hidden state SSE (HiSSE) models include two (or more) sets of diversification rates for each state hypothesized to affect diversification (e.g. Helmstetter *et al.*, 2023). For example, if a researcher wants to know if herbivorous beetles diversify faster than carnivorous ones, the hidden state approach would include

New Phytologist

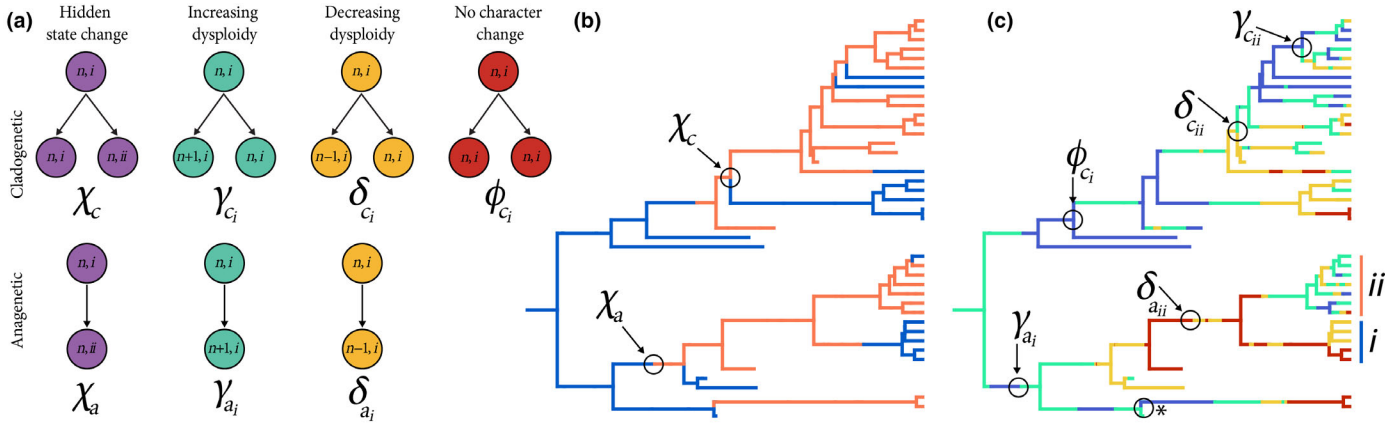

Fig. 1 The Chromosome number and Hidden State-dependent Speciation and Extinction (ChromoHiSSE) model. Panel (a) describes the event rates allowed in the model for both cladogenetic (upper) and anagenetic (lower) events. Panels (b) and (c) demonstrate those rates on a tree simulated under ChromoHiSSE. Branches that do not reach the present represent extinct, unsampled lineages. Panel (b) shows the anagenetic and cladogenetic changes in the hidden states, indicated by blue (*i*) vs orange (*ii*). Panel (c) shows the anagenetic and cladogenetic gains and losses of chromosomes, as well as speciation with no corresponding change in chromosome numbers. More red colors in (c) correspond to more chromosomes. The vertical blue and orange bars in (c) indicate clades in the blue (*i*) vs orange (*ii*) hidden states, displaying that chromosome number changes are less frequent in the blue hidden state than in the orange. The asterisk in (c) demarcates where a cladogenetic dysploidy event could appear as an anagenetic event because of unsampled lineages.

two diversification parameters for the carnivorous state and two diversification parameters for the herbivorous state. If both herbivory-specific rates are higher than the carnivory rates, then the analysis suggests that herbivores diversify faster than carnivores. However, our new model differs from previous HiSSE models, in that we parameterize diversification rates associated with *types* of state transitions rather than with the states themselves. This means that our goal is not to test whether $n = 15$ or $n = 16$ chromosomes have different modes of diversification, but rather whether changes in the karyotype (i.e. increase: $n = 15$ to $n = 16$ or decrease: $n = 15$ to $n = 14$) are associated with faster or slower diversification rates compared with heterogeneity (or noise) in the diversification process. To do this, we compare speciation rates *not* associated with dysploidy (φ_{c_i} and χ_c) to speciation rates *associated* with dysploidy (γ_{c_i} and δ_{c_i}) by defining the test statistic $T_i = \gamma_{c_i} + \delta_{c_i} - (\varphi_{c_i} + \chi_c)$ (Fig. 2d). A posterior probability of $T_i > 0$ greater than 0.95 ($P(T_i > 0) \geq 0.95$) represents high probability of speciation by single chromosome number change. By contrast, $P(T_i < 0) \geq 0.95$ suggests that speciation rates unrelated to chromosome number change are significantly faster than speciation by single chromosome number change. Test statistic T_{ii} is defined analogously for the second hidden state.

Simulation study

We performed a simulation study to validate ChromoHiSSE using parameter values that approximate our empirical parameter estimates, with the exception of extinction, which we varied across two simulation scenarios (simulating values are presented in Table S1). For each scenario (low and high extinction), we simulated 100 datasets (see Notes S3 for details). We analyzed all datasets using ChromoHiSSE to confirm that our model is able to recover the true simulating values and with ChromoSSE to test for the effect of not accounting for rate variation when it is present in the dataset.

Chromosomal and phylogenetic data for *Carex*

We implemented our proposed ChromoHiSSE model on a large dataset that includes a phylogeny of *Carex* with 755 taxa (*c.* 40% of extant diversity), representing all *Carex* subgenera and most of the sections for which chromosome counts have been reported. Our dataset also includes haploid chromosome number (n) for all tips in the tree; both the tree and chromosome number data come from Márquez-Corro *et al.* (2021). The original tree, based on a HybSeq backbone and three DNA regions (ITS, ETS, and *mat*K) from *c.* 1400 of 2000 *Carex* species, was published by Martín-Bravo *et al.* (2019). Tips without chromosome number information were pruned for this study. Chromosome number in *Carex* evolves through dysploid events, except in the small subgenus *Siderostictae* (13 species in our tree) – sister to the rest of the genus – which includes species with different reported ploidy levels (Márquez-Corro *et al.*, 2019; Márquez-Corro *et al.*, 2021). To avoid modeling rare polyploidy events that occur only in a small part of the tree, we removed *Siderostictae* from the primary analysis (though see Notes S4 for results with this clade included). For the remaining taxa, we coded chromosome numbers as the most frequent haploid number or the lowest haploid cytotype in polyploid lineages from the Márquez-Corro *et al.* (2021) dataset for those tips. The final data matrix included n-values ranging from 5 to 66 chromosomes for 742 taxa.

Computational implementation of ChromoHiSSE

We implemented ChromoHiSSE in RevBayes (Höhna *et al.*, 2016), a software for specifying Bayesian probabilistic graphical models primarily for phylogenetics and phylogenetic

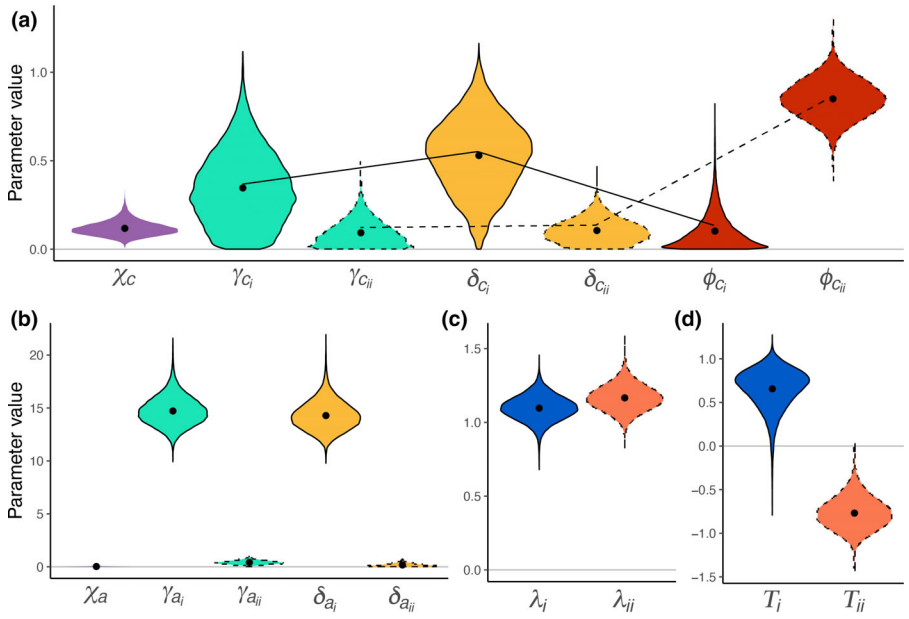

Fig. 2 Posterior distributions of rate estimates from our Chromosome number and Hidden State-dependent Speciation and Extinction (ChromoHiSSE) analysis of *Carex*. Solid lines indicate the posterior mean estimates of rates corresponding to hidden state *i* and dashed lines indicate the posterior mean estimates of hidden state *ii*. Panel (a) shows the posterior distributions of cladogenetic rates, panel (b) shows the anagenetic rates, panel (c) shows the total speciation rate per hidden state (λ_i and λ_{ii}), and panel (d) shows the test statistic (T_i and T_{ii}) of the difference in speciation rate associated with chromosome number change and speciation rate without chromosome number change. T_i is significantly greater than 0 (faster speciation when associated with chromosome number change $P(T_i > 0) = 0.986$), and T_{ii} is significantly less than 0 (slower speciation when associated with chromosome number change: $P(T_{ii} < 0) = 1.000$).

comparative methods. Due to the large state space and cladogenesis in our model, we used the newly developed TensorPhylo plugin (May & Meyer, 2022) to accelerate likelihood calculations and thus achieve convergence in a more reasonable time frame. We monitored ancestral states through stochastic character mapping as implemented in RevBayes (Freyman & Höhna, 2018b). We ran two chains of the analysis and assessed convergence in the R programming language (R Core Team, 2013). We processed the traces and removed 10% of the generations per chain as burnin using RevGadgets (Tribble *et al.*, 2022). We then calculated the effective sample size (ESS) value for each parameter in each chain using coda (Plummer *et al.*, 2006) and verified that the harmonic mean of the ESS values of each chain was greater than 200. We additionally visually inspected all model parameters across both chains in Tracer (Rambaut *et al.*, 2018). For any parameters that appeared to have strikingly non-normal posterior distributions, we also estimated a transformed ESS following Vehtari *et al.* (2021). We subsequently combined both runs for all downstream analyses.

Analysis and postprocessing

We summarized posteriors and plotted results in R (R Core Team, 2013) using the R package revgadgets (Tribble *et al.*, 2022). We additionally transformed model parameters to produce two types of useful summary statistics. First, we calculated the total speciation rate in hidden state *i* as $\lambda_i = \chi_{c_i} + \gamma_{c_i} + \delta_{c_i} + \varphi_{c_i}$ (and likewise for hidden state *ii*). Second, we wanted to know whether the rate of speciation concurrent with changes in chromosome number (i.e. $\gamma_{c_i} + \delta_{c_i}$) is greater or less than the rate of speciation with no change in chromosome number (i.e. $\chi_c + \varphi_{c_i}$). We thus estimated an additional summary statistic T (defined above); positive values of T indicate more speciation with chromosome number change and negative values indicate less speciation with chromosome number change. We estimate T for each hidden state to estimate how chromosome number changes associated with speciation vary between modes of the model.

All code for implementing and running the model and processing and plotting the results *i* available at Zenodo doi: 10.5281/zenodo.14035743.

Results

Two modes of chromosomal anagenesis

We recover two modes of anagenetic dysploidy. Branches in hidden state *i* have high rates of single chromosome number increases and decreases ($\gamma_{a_i} = 14.63$ and $\delta_{a_i} = 14.19$ chromosome number change events per lineage per million years (E/L/Myr); Table 1; Fig. 2b). For the other branches – in hidden state *ii* – chromosome number rarely changes via dysploidy (increases are $\gamma_{a_{ii}} = 0.39$ and decreases are $\delta_{a_{ii}} = 0.14$ E/L/Myr; Table 1; Fig. 2b). Rarely, lineages transition between hidden

New Phytologist

Table 1 Summary statistics of posterior distributions for parameter estimates for the Chromosome number and Hidden State-dependent Speciation and Extinction (ChromoHiSSE) model implemented in *Carex*.

Model parameter	Parameter type	Hidden state	Median	2.5% Quantile	97.5% Quantile
χ_c	Hidden state change	Cladogenetic	*i*, *ii*	0.113	0.051	0.215
γ_{c_i}	Increasing dysploidy	Cladogenetic	*i*	0.330	0.021	0.772
$\gamma_{c_{ii}}$	Increasing dysploidy	Cladogenetic	*ii*	0.077	0.003	0.261
δ_{c_i}	Decreasing dysploidy	Cladogenetic	*i*	0.539	0.120	0.916
$\delta_{c_{ii}}$	Decreasing dysploidy	Cladogenetic	*ii*	0.099	0.007	0.250
φ_{c_i}	No character change	Cladogenetic	*i*	0.072	0.003	0.363
$\varphi_{c_{ii}}$	No character change	Cladogenetic	*ii*	0.849	0.642	1.058
χ_a	Hidden state change	Anagenetic	*i*, *ii*	0.021	0.001	0.093
γ_{a_i}	Increasing dysploidy	Anagenetic	*i*	14.638	12.282	17.626
$\gamma_{a_{ii}}$	Increasing dysploidy	Anagenetic	*ii*	0.397	0.121	0.769
δ_{a_i}	Decreasing dysploidy	Anagenetic	*i*	14.191	11.903	17.246
$\delta_{a_{ii}}$	Decreasing dysploidy	Anagenetic	*ii*	0.142	0.005	0.505

Light blue rows correspond to parameter estimates for hidden state *i*, and light orange rows correspond to parameter estimates for hidden state *ii*. All rate estimates are given in units of events per lineage per million years (E/L/Myr).

states ($\chi_a = 0.02$). In Table 1, we show the credible intervals for these estimates, and in Fig. 2(b), we show the posterior distributions of all the anagenetic rate parameters.

Two modes of chromosomal evolution and speciation

We identify variation in chromosomal evolution across the tree via two distinct modes, corresponding to the two hidden states *i*, *ii*. While the overall total speciation rate (λ_i, λ_{ii}) between the two modes is very similar (Fig. 2c), rates of dysploidy and dysploidy-driven speciation vary significantly between the modes.

Hidden state *i* is characterized by a positive association between speciation and dysploidy and generally high rates of cladogenetic dysploidy (Fig. 2, solid lines; Table 1). In this mode, the difference between the rate of speciation with chromosome number change and speciation without chromosome number change is larger than zero with 98.6% probability ($P[T_{ii} > 0] = 0.986$). Thus, in hidden state *i*, there is a highly probable and positive correlation between chromosome number change and speciation rate (Fig. 2d, blue). Every estimated dysploidy-related cladogenetic rate is higher than non-dysploidy rates (e.g. speciation with no state change, Fig. 2a), which suggests that speciation happens more frequently when associated with changes to karyotype.

Cladogenetic increasing and decreasing dysploidy (γ_{c_i} and δ_{c_i}) are higher than 0 with 99% probability (Table 1) and faster than the cladogenetic rate of no character change (speciation without change in chromosome number, φ_{c_i}). While δ_{c_i} is slightly faster than γ_{c_i}, these rates overlap (Table 1).

By contrast, hidden state *ii* is characterized by a negative association between speciation and dysploidy and slower rates of anagenetic and cladogenetic dysploidy (Fig. 2, dashed lines; Table 1). The difference between speciation associated with chromosome number change and speciation without chromosome number change is less than 0 E/L/Myr ($P(T_{ii} < 0) = 1.000$), suggesting that speciation is slower when associated with a chromosome number change in hidden state *ii* (Fig. 2d, orange). Anagenetic rates ($\gamma_{a_{ii}}$ and $\delta_{a_{ii}}$) are quite low, approximating 0 (Table 1; Fig. 2b), suggesting that little to no anagenetic change in chromosome number happens in hidden state *ii*.

Cladogenetic dysploidy rates are also quite low, and while the rate of decreasing dysploidy ($\delta_{c_{ii}}$) is slightly higher than increasing dysploidy ($\gamma_{c_{ii}}$), this difference is minimal (Table 1). In summary, we identify variation in modes of chromosomal evolution across the tree and via two distinct modes, corresponding to the two hidden states *i*, *ii*. In hidden state *i*, dysploidy is linked to speciation, whereas in hidden state *ii*, with the same total speciation rate, dysploidy is not an important contributor speciation.

Reconstruction of chromosome number evolution

Just over half (51.85%) of all branches in the phylogeny showed a net change in chromosome number (Fig. 3). Overall, the dominant pattern of chromosome number in *Carex* is one of frequent but gradual change; 75.05% of per-branch net chromosome number changes are three or fewer gains or losses. Most clades vary in chromosome number despite some shallow evolutionary divergences (Fig. 3a). Yet, there are some remarkable exceptions. For example, Clade 2 – which includes sect. *Pictae*, the Hirta Clade, and sect. *Praelongae* – shows substantial variation in chromosome number, from $n = 13$ to $n = 66$. However, a clade of hooked sedges (*Carex* sect. *Uncinia*; Fig. 3, Clade 1 in gray) has a constant chromosome number ($n = 44$) across the 33 species included in this study, with one exception: *Carex perplexa* has a

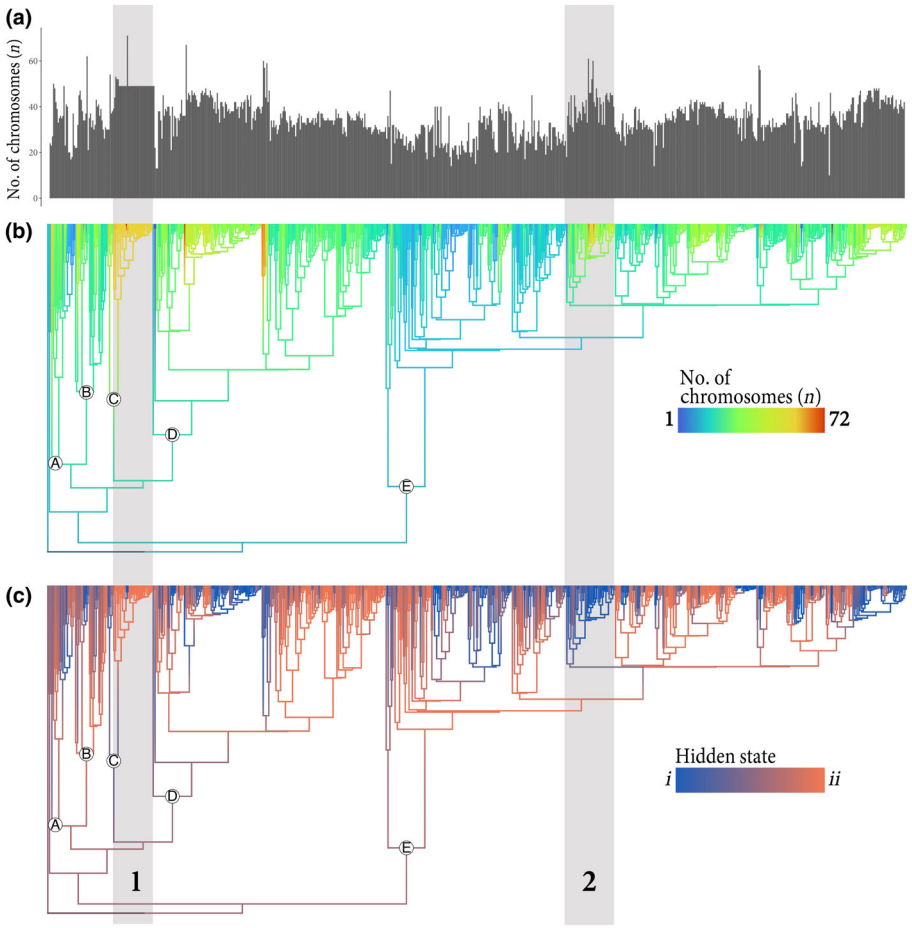

Fig. 3 Reconstruction of chromosome numbers (a and b) and hidden states (c) on the *Carex* phylogeny. Panel (a) shows the distribution of haploid chromosome numbers for all extant taxa included in the analysis. Panel (b) shows the reconstructed evolution of chromosome number along branches of the phylogeny, where redder colors indicate more chromosomes. Panel (c) shows the reconstructed evolution of the hidden state along branches of the phylogeny, where blue indicates strong statistical support for state *i*, orange indicates strong support for state *ii*, and intermediary colors indicate uncertainty in the estimates. Gray bars highlight two clades (labeled 1 and 2) that are discussed in the main text, and circled letters on nodes indicate the position of major subgenera: A – *Psyllophora*; B – *Euthyceras*; C – *Uncinia*; D – *Vignea*; E – *Carex*.

count of $n = 66$, suggesting that this species may be a demipolyploid. While both clades highlighted in Fig. 3 have similar crown ages, number of species, and average chromosome number ($n = 35.7$ vs $n = 44.6$), they differ dramatically in karyotype variability and, consequently, in the inferred hidden state (Clade 1 = *ii*, Clade 2 = *i*, Fig. 3c).

Simulation study

Our ChromoHiSSE analyses of datasets simulated under ChromoHiSSE (with low- and high-extinction scenarios) demonstrate that our model generally recovers the true, simulating parameter values, with three exceptions, discussed later. By contrast, when we analyze data simulated under ChromoHiSSE with a ChromoSSE model, parameter estimates are severely compromised.

For analyses under ChromoHiSSE, our simulations with a high extinction rate reveal a tendency to underestimate anagenetic dysploidy rates for the hidden state with the highest rates of dysploidy (γ_{a_i} and δ_{a_i}), as characterized by the error in the posterior mean point estimate (squared error of the posterior mean normalized by the true value averaged across simulations of *c.* 0.6 and *c.* 0.55 for the two parameters, respectively) as well as the frequency with which the 95% credible interval of the posterior contains the true value (the 'coverage'; *c.* 26% and *c.* 32% for the two parameters, respectively). However, the difference between these rates $(\gamma_{a_i} - \delta_{a_i})$ is less biased (normalized squared error of *c.* 28), and the coverage is nominal (*c.* 95%). The tendency to underestimate high rates could be a result of saturation (so many chromosome number changes that subsequent changes become 'invisible' in the simulated datasets) or prior sensitivity (as our prior mean on these rates is significantly lower than the true value, concentrating the prior probability on low rates; see prior specification details in Notes S3). Additionally, the posterior mean rates of change between hidden states (both anagenetic χ_a and cladogenetic χ_c) are higher than the simulating values (Figs S1, S2). However, the coverage for these parameters is high ($> 95\%$), suggesting that our ability to detect the true rate of change between hidden states is low. All of these results are similar for high-extinction and low-extinction simulation scenarios and are presented in more detail in Notes S3 (Figs S1–S4; Tables S2 and S3).

For analyses under ChromoSSE, the posterior distributions are intermediate between the true values for the hidden states, suggesting that ChromoSSE is essentially 'averaging' the rates for the two hidden states together. Additionally, the '95%' credible intervals of these distributions do not contain the true value of either hidden state. (See Figs S5, S6; Tables S4, S5 for detailed results under the ChromoSSE model.)

Discussion

Although the impact of dysploidy on *Carex* diversification has been addressed previously (e.g. Faulkner, 1972; Hipp, 2007; Márquez-Corro *et al.*, 2019; Márquez-Corro *et al.*, 2021), this is the first study to jointly model chromosome number change and diversification and to demonstrate an association between higher speciation rates and dysploidy in parts of the phylogeny despite heterogeneity in the process of diversification. While gains and losses in chromosome number spur diversification along parts of the phylogeny (hidden state *i*), they have the opposite effect elsewhere in the tree (hidden state *ii*, Fig. 3). Furthermore, while dysploidy does not lead to higher rates of speciation across the tree on average, speciation in some clades is driven strongly by dysploidy. We propose that these discrepancies between clades may be due to the nature of holocentric chromosomes, where a single dysploidy event in isolation may not be enough to trigger reproductive isolation (Whitkus, 1988; Hipp *et al.*, 2010; Escudero *et al.*, 2016; Márquez-Corro *et al.*, 2019; Lucek *et al.*, 2022), but the accumulation of sufficient chromosomal rearrangements in a lineage may form a reproductive barrier and thus trigger speciation (Baker & Bickham, 1986; Whitkus, 1988; Escudero *et al.*, 2016). This is a central hypothesis discussed in fig. 4 of Lucek *et al.* (2022) known as the recombination-suppression/hybrid-dysfunction chromosomal speciation model. We elaborate on the differences between speciation modes, discuss the evidence supporting the recombination-suppression/hybrid-dysfunction model, and provide suggestions for future research later.

Some *Carex* species are known to have striking chromosome number polymorphism, even within populations (Whitkus, 1988; Luceño & Castroviejo, 1991; Escudero *et al.*, 2013; Escudero *et al.*, 2023). For example, *Carex scoparia* varies from $n = 28$ to $n = 35$ (Escudero *et al.*, 2013), and individuals with different chromosome numbers are able to reproduce and exchange alleles, maintaining gene flow despite chromosome number differences. This polymorphism suggests that some chromosome number differences are insufficient to create reproductive barriers (Hipp *et al.*, 2009), which is supported by our results; often, inferred chromosome number changes do not result in an immediate speciation event on the phylogeny. This constant gene flow might also result in a difficulty to estimate anagenetic chromosome number change when it is rampant. As our simulations have indicated, it is possible that anagenetic dysploidy is underestimated frequently for holocentric chromosomes. The question remains, why does dysploidy result in speciation in some cases and not in others? One possibility is that an accumulation of changes may eventually lead to reproductive isolation, where one 'last straw' dysploidy event triggers speciation (referred to as the last-straw hypothesis). We discuss potential model developments that could test this theory later in the *Modeling chromosomal speciation* section. Another possibility is that rearrangements in some parts of the genome are more stable than others, and the genomic architecture – where the fragmentation or fusion occurs in the genome – determines the evolutionary effects of dysploidy.

Patterns of repeat DNA – *for example* LINEs, LTRs, and Helitrons – differ significantly between *Carex* lineages whose chromosome arrangements evolve at different rates (Cornet *et al.*, 2023). Moreover, repeat regions in investigated *Carex* genomes correlate with chromosome breakpoints across deep phylogenetic splits as well as within species (Escudero *et al.*, 2023; Höök *et al.*, 2023), pointing to a mechanism by which natural selection could operate on the genome architecture of chromosome rearrangements in the genus. In fact, synteny in holocentric sedge chromosomes is more conserved than we would expect if chromosome evolution were unconstrained, even in comparisons between species that span deep nodes in the phylogeny (Escudero *et al.*, 2023). This suggests that selection may maintain large blocks of the genome, likely comprising numerous, tightly linked genes (Escudero *et al.*, 2023). The massive diversity of sedges is thus likely an outcome of two processes: recombination suppression in rearranged regions of the genome, shaping ecological divergence in during species divergence; and the typically gradual – or, occasionally, immediate – evolution of reproductive isolation as chromosomes split and fuse. Such information could be incorporated into a macroevolutionary framework by relaxing the assumption in SSE models that speciation occurs instantly, and instead integrating microevolutionary dynamics of how genomic changes become fixed in lineages via ancestral recombination graphs (e.g. Brandt *et al.*, 2024; Deng *et al.*, 2024).

An important caveat to our study is the presence of missing data – particularly for tropical lineages – in terms of both sequenced taxa represented in the phylogeny and available chromosome counts. While the phylogeny used in this study was assembled with a HybSeq backbone and three DNA regions for *c.* 1400 out of 2000 species, phylogenetic relationships in some areas of the genus are still tenuous (Jiménez-Mejías *et al.*, 2016; Martín-Bravo *et al.*, 2019; Roalson *et al.*, 2021). Additionally, there are only chromosome counts for *c.* 700 species in the phylogeny, and data availability is more sparse in tropical lineages than temperate ones. For example, one of the most species-diverse lineages within *Carex* – the Decora Clade/*Carex* sect. *Indicae* (Roalson *et al.*, 2021) – only includes five species with reported chromosome numbers (Márquez-Corro *et al.*, 2021). This pantropical lineage alone would constitute *c.* 10% of the genus (Roalson *et al.*, 2021). Also, we include only a few New Zealand species from *Carex* sect. *Uncinia* (Clade 1 in Fig. 3), as there are few known karyotypes for the South American representatives of the lineage. Dispersal traits such as hooked utricles and epizoochory might be as or more important for diversification as karyotype in this clade, but we cannot tease these effects apart without additional chromosome count data. *Carex* biodiversity is greatest in temperate regions including the Global North, and this bias in data availability may suggest that our findings are most robust for northern temperate *Carex*. Filling these data gaps would facilitate testing whether evolutionary modes of dysploidy play out differently based on geography and/or dispersal traits.

Chromosome number evolution

The contrasting Clades 1 and 2 from Fig. 3 reflect the two modes of macroevolution described previously in 'Results: Reconstruction of chromosome number evolution' and in 'Discussion:

Multiple modes of chromosomal speciation' exemplify the importance of including hidden states in our model. *Carex* sect. *Uncinia* (hook sedges, Clade 1 in Fig. 3) is characterized by hooked utricles that allow for long-distance dispersal through epizoochory (García-Moro *et al.*, 2022). The hook sedges exemplify a *Carex* clade that is characterized by low-to-zero dysploidy but high rates of speciation (Martín-Bravo *et al.*, 2019) – though this was not formally tested in our study – and thus may be a particularly appealing candidate for future studies focusing on potential adaptive traits (currently 'hidden' and not related to chromosome number) that drive diversification. This clade is characterized by similar karyotypes of many ($n = 44$) chromosomes (Márquez-Corro *et al.*, 2021), mostly reported from New Zealand, where *c.* half of the section diversified. However, this only gives partial information on their evolutionary history, as the karyotypes of the Andean relatives have not yet been studied in detail. Most of the New Zealand chromosome counts correspond to a single, old report, and there is some variation in the few existing counts from South America, so these results may be taken with caution. If the South American species show wider karyotypic variation than currently detected in comparison with the New Zealand lineage, there may be different drivers of karyotype evolution and speciation on each side of the South Pacific.

Modeling chromosomal speciation

The methodological innovations presented here allow us to integrate process noise and variation in chromosome number evolution, despite the computational challenges associated with the huge number of states defined by chromosome numbers. Process variation is fundamental to our conclusions; only by incorporating multiple diversification modes do we discover that while dysploidy is strongly associated with speciation in some clades, in other parts of the phylogeny, dysploidy is rare and does not lead to faster speciation. Leaving the hidden states out leads to a false inference about the significance of chromosome evolution to diversification: When we implement our model without process variation, our results suggest a strong, uniform boost in speciation rate associated with dysploidy (results presented in Notes S4; Figs S7–S10). Furthermore, our simulation study demonstrates that when hidden rate variation is present, analyses by a model without hidden rate variation (e.g. ChromoSSE) consistently fail to recover the true parameter values (Figs S5, S6). Our approach permits estimation of the past via stochastic maps, allowing for the detection of lineages in which dysploidy has been linked or not to diversification. Reconstruction of the past and identification of clades and lineages where a trait makes a difference in diversification is key for future, smaller comparative studies aimed at understanding the genomic underpinnings of plant speciation.

RevBayes' graphical modeling framework permits flexibility in future modifications to our approach. Our ChromoHiSSE model does not include parameters for polyploidy, but we provide the mathematical framework for such additions in Notes S2. Our model includes two hidden states, but future implementations could increase the number of hidden states (with a significant increase in the computational effort required) and/or limit which parameters vary across hidden states. Model selection – typically quite challenging and computationally intensive to implement in Bayesian approaches – could assist researchers with smaller datasets, who lack the ability to estimate all parameters in this parameter-rich model, to decide between models that allow all or some parameters to vary between hidden states. In particular, we believe a Bayesian model averaging approach using reversible-jump MCMC may prove particularly useful (Freyman & Höhna, 2018a), as it avoids the need to compute computationally expensive marginal likelihoods and could automatically consider all ways that parameters could be shared between hidden states.

Future work that builds off our ChromoHiSSE approach will allow us to pursue promising avenues for innovative research; we highlight two examples. First, like all birth–death models, our ChromoHiSSE model operates with species as the fundamental unit of analysis (the tips in the tree) and thus does not formally model the chromosome number polymorphism that is present in some *Carex* lineages. Second, while ChromoHiSSE tests for the effect of single changes in chromosome number on diversification rates, it cannot test for the effect of an accumulation of changes (the last-straw hypothesis). However, ChromoHiSSE could be modified to include tip-state polymorphism (e.g. a dysploid series) as additional hidden states (e.g. a particular tip either has 12, 13, or 14 chromosomes). Additionally, (Goldberg & Foo, 2020) described a mechanism for modeling memory (thus the accumulation of chromosome number changes) in a macroevolutionary framework using hidden states, which could be applied to test the last-straw hypothesis.

Conclusions

Our work demonstrates the important role of dysploidy on diversification in *Carex*, a model lineage for understanding how karyotype rearrangements via dysploidy affect speciation and macroevolutionary dynamics. Our results – using the new ChromoHiSSE model – paint a complex picture of how dysploidy affects speciation in a clade characterized by high species diversity, high morphological disparity, and holocentric chromosomes (Martín-Bravo *et al.*, 2019), and our results support the recombination-suppression/hybrid-dysfunction chromosomal speciation model, in which only some karyotype rearrangements trigger reproductive isolation and thus speciation. Future work on the underlying genomic mechanisms of chromosomal speciation via comparative genomics will be particularly powerful for linking across scales, from molecules to lineages. Ultimately, our novel modeling approach also serves as a critical step toward even more complex and powerful macroevolutionary analyses that incorporate intraspecific chromosome number variation and track the accumulation of change through time.

Acknowledgements

CMT was supported by the School of Life Sciences at the University of Hawai'i at Mānoa, funded through the Office of the

Vice President of Research. This material is based upon work supported by the NSF Postdoctoral Research Fellowships in Biology Program under grant no. 2109835 to CMT. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. JIM-C was funded by the 'Next Generation EU' funding, the Recovery Plan, Transformation and Resilience, and the Ministry of Universities, under the grant 'Fundación Margarita Salas' for the requalification of the Spanish university system 2021–2023, called by the Universidad Pablo de Olavide, Seville (2021/0005425). ME was supported by the Ministerior de Ciencia e Innovación-Fondo Europeo de Desarrollo Regional (MICINN-FEDER), Project DiversiChrom (PID2021-122715NB-I00). RZF was supported by NSF-DEB 2323170. We thank three anonymous reviewers for their comments and feedback on earlier versions of the manuscript and associate editor Daniel Ortiz-Barrientos for a thorough revision.

Competing interests

None declared.

Author contributions

RZ-F and ME conceived the study. CMT, JIM-C, ME and RZ-F wrote the article. JIM-C, ME and ALH gathered all *Carex* chromosome number and phylogenetic data. CMT, MRM and RZ-F designed models and methods, and verified mathematical/statistical notation. CMT and MRM wrote computational code and created figures. All authors revised and edited the manuscript.

ORCID

Marcial Escudero https://orcid.org/0000-0002-2541-5427
Andrew L. Hipp https://orcid.org/0000-0002-1241-9904
José Ignacio Márquez-Corro https://orcid.org/0000-0003-4277-2933
Michael R. May https://orcid.org/0000-0002-5031-4820
Carrie M. Tribble https://orcid.org/0000-0001-7263-7885
Rosana Zenil-Ferguson https://orcid.org/0000-0002-9083-2972

Data availability

RevBayes model code and R code for processing and plotting the results are available at Zenodo doi: 10.5281/zenodo.14035743.

References

Ayala FJ, Coluzzi M. 2005. Chromosome speciation: humans, *Drosophila*, and mosquitoes. *Proceedings of the National Academy of Sciences, USA* **102**: 6535–6542.

Baker RJ, Bickham JW. 1986. Speciation by monobrachial centric fusions. *Proceedings of the National Academy of Sciences, USA* **83**: 8245–8248.

Beaulieu JM, O'Meara BC. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* **65**: 583–601.

Brandt DY, Huber CD, Chiang CW, Ortega-Del Vecchyo D. 2024. The promise of inferring the past using the ancestral recombination graph. *Genome Biology and Evolution* **16**: evae005.

Caetano DS, O'Meara BC, Beaulieu JM. 2018. Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution* **72**: 2308–2324.

Carta A, Bedini G, Peruzzi L. 2020. A deep dive into the ancestral chromosome number and genome size of flowering plants. *New Phytologist* **228**: 1097–1106.

Cornet C, Pablo M, Augustijnen H, Nguyen P, Escudero M, Lucek K. 2023. Holocentric repeat landscapes: from microevolutionary patterns to macroevolutionary associations with karyotype evolution. *Molecular Ecology* **33**: e17100.

Deng Y, Nielsen R, Song YS. 2024. Robust and accurate bayesian inference of genome-wide genealogies for large samples. *bioRxiv*. doi: 10.1101/2024.03.16.585351.

Elliott TL, Zedek F, Barrett RL, Bruhl JJ, Escudero M, Hroudová Z, Joly S, Larridon I, Luceño M, Márquez-Corro JI *et al.* 2022. Chromosome size matters: genome evolution in the cyperid clade. *Annals of Botany* **130**: 999–1014.

Escudero M, Arroyo JM, González-Ramírez S, Jordano P. 2023. Founder events and subsequent genetic bottlenecks underlie karyotype evolution in the ibero – North African endemic *Carex helodes*. *Annals of Botany* **133**: 871–882.

Escudero M, Hahn M, Brown BH, Lueders K, Hipp AL. 2016. Chromosomal rearrangements in holocentric organisms lead to reproductive isolation by hybrid dysfunction: the correlation between karyotype rearrangements and germination rates in sedges. *American Journal of Botany* **103**: 1529–1536.

Escudero M, Marques A, Lucek K, Hipp AL. 2023. Genomic hotspots of chromosome rearrangements explain conserved synteny despite high rates of chromosome evolution in a holocentric lineage. *Molecular Ecology* **33**: e17086.

Escudero M, Márquez-Corro JI, Hipp AL. 2016. The phylogenetic origins and evolutionary history of holocentric chromosomes. *Systematic Botany* **41**: 580–585.

Escudero M, Martín-Bravo S, Itay M, Fernández-Mazuecos M, Fiz-Palacios O, Hipp AL, Pimentel M, Valcárcel V, Vargas P, Luceño M. 2014. Karyotypic changes through dysploidy persist longer over evolutionary time than polyploid changes. *PLoS ONE* **9**: e85266.

Escudero M, Weber JA, Hipp AL. 2013. Species coherence in the face of karyotype diversification in holocentric organisms: the case of a cytogenetically variable sedge (*Carex scoparia*, Cyperaceae). *Annals of Botany* **112**: 515–526.

Faulkner J. 1972. Chromosome studies on *Carex* section *Acutae* in north-west Europe. *Botanical Journal of the Linnean Society* **65**: 271–301.

Freyman WA, Höhna S. 2018a. Cladogenetic and anagenetic models of chromosome number evolution: a Bayesian model averaging approach. *Systematic Biology* **67**: 195–215.

Freyman WA, Höhna S. 2018b. Stochastic character mapping of state-dependent diversification reveals the tempo of evolutionary decline in self-compatible Onagraceae lineages. *Systematic Biology* **68**: 505–519.

García-Moro P, Otero A, Benítez-Benítez C, Costa L, Martín-Bravo S, Naczi RF, Reznicek A, Roalson EH, Starr JR, Jiménez-Mejías P. 2022. Biogeography and systematics of *Carex* subgenus *Uncinia* (Cyperaceae): a unique radiation for the genus *Carex* in the Southern Hemisphere. *Taxon* **71**: 587–607.

Goldberg EE, Foo J. 2020. Memory in trait macroevolution. *The American Naturalist* **195**: 300–314.

Helmstetter AJ, Zenil-Ferguson R, Sauquet H, Otto SP, Méndez M, Vallejo-Marin M, Schönenberger J, Burgarella C, Anderson B, de Boer H *et al.* 2023. Trait-dependent diversification in angiosperms: patterns, models and data. *Ecology Letters* **26**: 640–657.

Hipp A, Rothrock P, Roalson E. 2009. The evolution of chromosome arrangements in *Carex* (Cyperaceae). *The Botanical Review* **75**: 96–109.

Hipp AL. 2007. Nonuniform processes of chromosome evolution in sedges (*Carex*: Cyperaceae). *Evolution* **61**: 2175–2194.

Hipp AL, Rothrock PE, Whitkus R, Weber JA. 2010. Chromosomes tell half of the story: the correlation between karyotype rearrangements and genetic diversity in sedges, a group with holocentric chromosomes. *Molecular Ecology* **19**: 3124–3138.

Höhna S, Landis MJ, Heath TA, Boussau B, Lartillot N, Moore BR, Huelsenbeck JP, Ronquist F. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology* **65**: 726–736.

Höök L, Näsvall K, Vila R, Wiklund C, Backström N. 2023. High-density linkage maps and chromosome level genome assemblies unveil direction and frequency of extensive structural rearrangements in wood white butterflies (*Leptidea* spp.). *Chromosome Research* **31**: 2.

Jiménez-Mejías P, Hahn M, Lueders K, Starr JR, Brown BH, Chouinard BN, Chung K-S, Escudero M, Ford BA, Ford KA *et al.* 2016. Megaphylogenetic specimen-level approaches to the *Carex* (Cyperaceae) phylogeny using ITS, ETS, and *mat*K sequences: implications for classification the Global *Carex* Group. *Systematic Botany* **41**: 500–518.

Lucek K, Augustijnen H, Escudero M. 2022. A holocentric twist to chromosomal speciation? *Trends in Ecology & Evolution* **37**: 655–662.

Luceño M, Castroviejo S. 1991. Agmatoploidy in *Carex laevigata* (Cyperaceae). Fusion and fission of chromosomes as the mechanism of cytogenetic evolution in Iberian populations. *Plant Systematics and Evolution* **117**: 149–159.

Magallón S, Castillo A. 2009. Angiosperm diversification through time. *American Journal of Botany* **96**: 349–365.

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

Mandáková T, Lysak MA. 2018. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Current Opinion in Plant Biology* **42**: 55–65.

Marques A, Ribeiro T, Neumann P, Macas J, Novák P, Schubert V, Pellino M, Fuchs J, Ma W, Kuhlmann M *et al.* 2015. Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. *Proceedings of the National Academy of Sciences, USA* **112**: 13633–13638.

Márquez-Corro JI, Escudero M, Luceño M. 2018. Do holocentric chromosomes represent an evolutionary advantage? a study of paired analyses of diversification rates of lineages with holocentric chromosomes and their monocentric closest relatives. *Chromosome Research* **26**: 139–152.

Márquez-Corro JI, Martín-Bravo S, Jiménez-Mejías P, Hipp AL, Spalink D, Naczi RF, Roalson EH, Luceño M, Escudero M. 2021. Macroevolutionary insights into sedges (*Carex*: Cyperaceae): the effects of rapid chromosome number evolution on lineage diversification. *Journal of Systematics and Evolution* **59**: 776–790.

Márquez-Corro JI, Martín-Bravo S, Pedrosa-Harand A, Hipp AL, Luceño M, Escudero M. 2019. *Karyotype evolution in holocentric organisms*. New York, NY, USA: John Wiley & Sons, 1–7.

Márquez-Corro JI, Martín-Bravo S, Spalink D, Luceño M, Escudero M. 2019. Inferring hypothesis-based transitions in cladespecific models of chromosome number evolution in sedges (Cyperaceae). *Molecular Phylogenetics and Evolution* **135**: 203–209.

Martín-Bravo S, Jiménez-Mejías P, Villaverde T, Escudero M, Hahn M, Spalink D, Roalson EH, Hipp AL, Group, G. C, Benítez-Benítez C *et al.* 2019. A tale of worldwide success: behind the scenes of *Carex* (Cyperaceae) biogeography and diversification. *Journal of Systematics and Evolution* **57**: 695–718.

May, M. R. & Meyer, X. (2022), *TensorPhylo RevBayes plugin*. [WWW document] URL https://bitbucket.org/mrmay/tensorphylo/src/master.

Mayrose I, Barker MS, Otto SP. 2010. Probabilistic models of chromosome number evolution and the inference of polyploidy. *Systematic Biology* **59**: 132–144.

Plummer M, Best N, Cowles K, Vines K. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* **6**: 7–11.

POWO. 2023. *Plants of the World Online, facilitated by the Royal Botanic Gardens, Kew*. [WWW document] URL http://www.plantsoftheworldonline.org/.

R Core Team. 2013. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rabosky DL, Goldberg EE. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology* **64**: 340–355.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901–904.

Roalson EH, Jiménez-Mejías P, Hipp AL, Benítez-Benítez C, Bruederle LP, Chung K-S, Escudero M, Ford BA, Ford K, Gebauer S *et al.* 2021. A framework infrageneric classification of *Carex* (Cyperaceae) and its organizing principles. *Journal of Systematics and Evolution* **59**: 726–762.

Robertson W. 1916. Chromosome studies. I. Taxonomic relationships shown in the chromosomes of Tettigidae and Acrididae: Vshaped chromosomes and their significance in Acrididae, Locustidae, and Gryllidae: chromosomes and variation. *Journal of Morphology* **27**: 179–331.

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

Shafir A, Halabi K, Escudero M, Itay M. 2023. A non-homogeneous model of chromosome-number evolution to reveal shifts in the transition patterns across the phylogeny. *New Phytologist* **238**: 1733–1744.

Tribble CM, Freyman WA, Landis MJ, Lim JY, Barido-Sottani J, Kopperud BT, Höhna S, May MR. 2022. REVGADGETS: an R package for visualizing Bayesian phylogenetic analyses from RevBayes. *Methods in Ecology and Evolution* **13**: 314–323.

Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C. 2021. Rank-normalization, folding, and localization: an improved $\widehat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis* **16**: 667–718.

de Vos JM, Augustijnen H, Bätscher L, Lucek K. 2020. Speciation through chromosomal fusion and fission in Lepidoptera. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **375**: 20190539.

Whitkus R. 1988. Experimental hybridizations among chromosome races of *Carex pachystachya* and the related species *C. macloviana* and *C. preslii* (Cyperaceae). *Systematic Botany* **13**: 146–153.

Zenil-Ferguson R, Ponciano JM, Burleigh JG. 2017. Testing the association of phenotypes with polyploidy: an example using herbaceous and woody eudicots. *Evolution* **71**: 1138–1148.

Supporting Information

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

Fig. S1 Absolute parameter estimates under the true model for high extinction.

Fig. S2 Relative parameter estimates under the true model with high extinction.

Fig. S3 Absolute parameter estimates under the true model with low extinction.

Fig. S4 Relative parameter estimates under the true model with low extinction.

Fig. S5 Absolute parameter estimates under the ChromoSSE model (i.e. no hidden states) with high extinction.

Fig. S6 Absolute parameter estimates under the ChromoSSE model (i.e. no hidden states) with low extinction.

Fig. S7 Posterior distributions of rate estimates from a ChromoSSE model that allows anagenetic polyploidy (ρ_a; estimated to be *c.* 0) on data that include Carex subg. Siderostictae.

Fig. S8 Posterior distributions of rate estimates from a ChromoSSE model that does not allow anagenetic polyploidy on data that include Carex subg. Siderostictae.

Fig. S9 Posterior distributions of rate estimates from a ChromoSSE model that allows anagenetic polyploidy (ρ_a; estimated to be *c.* 0) on data that do not include Carex subg. Siderostictae.

Fig. S10 Posterior distributions of rate estimates from a ChromoSSE model that does not allow anagenetic polyploidy on data that does not include Carex subg. Siderostictae.

Notes S1 Background on dysploidy and diversification.

Notes S2 General Chromosome number and Hidden State-dependent Speciation and Extinction model.

Notes S3 Simulation study of datasets simulated under Chromosome number and Hidden State-dependent Speciation and Extinction with low- and high-extinction scenarios.

Table S1 True parameter values used to simulated datasets for the high- and low-extinction scenarios.

Table S2 Parameter estimates, error, and coverage (how often the true value falls within the 95% confidence interval) under the Chromosome number and Hidden State-dependent Speciation and Extinction model with high extinction.

Table S3 Parameter estimates, error, and coverage (how often the true value falls within the 95% confidence interval) under the Chromosome number and Hidden State-dependent Speciation and Extinction model with low extinction.

Table S4 Parameter estimates, error, and coverage (how often the true value falls within the 95% confidence interval) under the ChromoSSE model (i.e. no hidden states) with high extinction.

Table S5 Parameter estimates, error, and coverage (how often the true value falls within the 95% confidence interval) under the ChromoSSE model (i.e. no hidden states) with low extinction.

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.