New Phytologist

Research

Polyploidy promotes species diversification of *Allium* through ecological shifts

Ting-Shen Han[1,2,3], Quan-Jing Zheng[1,4], Renske E. Onstein[5], Blanca M. Rojas-Andrés[6], Frank Hauenschild[6], Alexandra N. Muellner-Riehl[5,6] and Yao-Wu Xing[1,2]

[1]CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan 666303, China; [2]Center of Plant Ecology, Core Botanical Gardens, Chinese Academy of Sciences, Mengla, Yunnan 666303, China; [3]Department of Biology, Duke University, Box 90338, Durham, NC 27708, USA; [4]University of Chinese Academy of Sciences, Beijing 100049, China; [5]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, Leipzig D-04103, Germany; [6]Department of Molecular Evolution and Plant Systematics & Herbarium (LZ), Leipzig University, Johannisallee 21-23, Leipzig D-04103, Germany

Author for correspondence:
Yao-Wu Xing
Tel: +86 691 8713882
Email: ywxing@xtbg.org.cn

Received: *15 April 2019*
Accepted: *1 August 2019*

New Phytologist (2020) **225:** 571–583
doi: 10.1111/nph.16098

Key words: *Allium*, diversification, ecological shifts, evolution, polyploidy, speciation rate.

Summary

- Despite the role of polyploidy in multiple evolutionary processes, its impact on plant diversification remains controversial. An increased polyploid frequency may facilitate speciation through shifts in ecology, morphology or both. Here we used *Allium* to evaluate: (1) the relationship between intraspecific polyploid frequency and species diversification rate; and (2) whether this process is associated with habitat and/or trait shifts.
- Using eight plastid and nuclear ribosomal markers, we built a phylogeny of 448 *Allium* species, representing 46% of the total. We quantified intraspecific ploidy diversity, heterogeneity in diversification rates and their relationship along the phylogeny using trait-dependent diversification models. Finally, we evaluated the association between polyploidisation and habitat or trait shifts.
- We detected high ploidy diversity in *Allium* and a polyploidy-related diversification rate shift with a probability of 95% in East Asia. *Allium* lineages with high polyploid frequencies had higher species diversification rates than those of diploids or lineages with lower polyploid frequencies. Shifts in speciation rates were strongly correlated with habitat shifts linked to particular soil conditions; 81.7% of edaphic variation could be explained by polyploidisation.
- Our study emphasises the role of intraspecific polyploid frequency combined with ecological drivers on *Allium* diversification, which may explain plant radiations more generally.

Introduction

Polyploidy plays an important role in plant evolutionary processes, such as speciation (Wood *et al.*, 2009), novel trait generation (Soltis & Soltis, 2016), crop domestication (Salman-Minkov *et al.*, 2016), invasion (Suda *et al.*, 2015) and community assembly (Segraves, 2017; Rice *et al.*, 2019). Although nearly 15% of speciation events in angiosperms could have resulted from polyploidy (Wood *et al.*, 2009), it has long been debated whether polyploidy promotes diversification or not. Polyploids were once considered as 'evolutionary dead-ends' with reduced selection efficacy and low rates of adaptation (Stebbins, 1950). Phylogeny-based comparative work has suggested that polyploids diversify more slowly, with higher extinction rates and lower speciation rates than those of their diploid relatives (Mayrose *et al.*, 2011). Nevertheless, recent studies have highlighted a positive correlation of polyploidy with species diversification (Levin & Soltis, 2018; Ren *et al.*, 2018). This is attributed to the advantages of polyploids with respect to genomic plasticity (Leitch & Leitch, 2008), ecological transformation (Ramsey & Ramsey, 2014), or long-term adaptation (Van de Peer *et al.*, 2017). For example, polyploidisation can be temporally associated with speciation events (Smith *et al.*, 2018; Cai *et al.*, 2019), yet with lags of up to millions of years (Schranz *et al.*, 2012). Polyploidisation may also be an important mechanism underlying the formation of biodiversity in geographic areas, such as the Pan-Himalayan region (Wen *et al.*, 2014), the Andes (Luebert & Weigend, 2014) or the Mediterranean Basin (Marques *et al.*, 2018). Therefore, a thorough reevaluation of the evolutionary significance of polyploidy and its impact on diversification is needed (Kellogg, 2016).

Previous studies investigating the effect of ploidy on diversification mostly focused on simplified binary traits (e.g. diploid or polyploid), which underestimates the complexity of polyploids (e.g. mixed ploidy) (Soltis *et al.*, 2014; Kolář *et al.*, 2017). As the origin and establishment of polyploidy are both under frequency-dependent selection to avoid minority cytotype exclusion (MCE; in which rare cytotypes tend to become extinct due to rapid reshuffling by mating with the dominant cytotype in a mixed-ploidy population) (Levin, 1975), the usage of intraspecific

frequency of polyploidy or multistate traits based on it, may be more appropriate to evaluate the association between ploidy and species diversification. These traits can not only be used to quantify the mixed-ploidy variation in angiosperms (e.g. the median count used in Rice *et al.*, 2019), but also make several quantitative or hidden state-dependent models operable in the issue of polyploidy-related diversification (such as FitzJohn, 2010; Beaulieu & O'Meara, 2016). A high intraspecific frequency of polyploidy can be achieved by various mechanisms, including recurrent origins (Soltis & Soltis, 1999), triploid mediation (Köhler *et al.*, 2010), or mating-system transition to asexuality (Freeling, 2017). Interactions among these factors can overcome MCE-related disadvantages and promote polyploid speciation (Ramsey & Schemske, 1998; Husband, 2000; Kreiner *et al.*, 2017). Additionally, a high intraspecific polyploid frequency will increase the population size as well as contacts within or among populations (Fowler & Levin, 2016; Levin, 2019). Gene flow under these frequent intra- or interspecific contacts can result in adaptive introgression (Petit *et al.*, 1999; Chapman & Abbott, 2010; Han *et al.*, 2015), which may further stabilise polyploid populations. It is unclear whether plant lineages with high polyploid frequency (at intraspecific level hereafter) exhibit higher speciation rates than those of diploids and lineages with lower polyploid frequencies.

Furthermore, the underlying drivers of polyploid speciation are unclear (Rothfels & Otto, 2016). Habitat shifts are thought to be an important factor for driving polyploid evolution. This idea is founded on studies using simulated data (Rodriguez, 1996; Marchant *et al.*, 2016) and on empirical observations (Godfree *et al.*, 2017), both for deep time (Estep *et al.*, 2014; Cai *et al.*, 2019) and more recent time (Abbott & Lowe, 2004; Chao *et al.*, 2013). Modelling studies have shown that niche separation or intermediate states between polyploidy and diploidy can counteract MCE-related interference (Levin, 1975; Parisod & Broennimann, 2016) and promote polyploid expansion (Rodriguez, 1996; Marchant *et al.*, 2016). Moreover, robust transformative effects of polyploids make them more responsive to stress (Chao *et al.*, 2013; Godfree *et al.*, 2017; Čertner *et al.*, 2019), which consequently improves their survival and long-term persistence. Therefore, habitat shifts may contribute to polyploid speciation through ecological radiations. However, it is not clear whether habitat shifts could affect speciation rates of polyploid lineages. Additionally, ecologically driven diversification events seem to be associated with key innovations of functional traits (Čertner *et al.*, 2019; Wei *et al.*, 2019). Accordingly, we also hypothesise that trait shifts, including shifts in mating system or growth form, could facilitate polyploid speciation (Freeling, 2017; Van Drunen & Husband, 2019).

In this study, we examined the relationship between intraspecific polyploid frequency and diversification in *Allium* L. and its association with shifts in habitats or traits. *Allium* is one of the largest monocotyledonous genera with 971 species in the family Amaryllidaceae, including important vegetable crops, such as onion and garlic (Hauenschild *et al.*, 2017). *Allium* species are widely distributed in the Northern Hemisphere, occupying habitats from the dry subtropics to boreal zones (Fritsch & Fritsch, 2002). Previous studies have suggested that *Allium* exhibited several diversification rate shifts (Hauenschild *et al.*, 2017; Xing & Ree, 2017), implying that accelerated species diversification may have occurred. Moreover, the genus is rich in mixed-ploidy species (Peruzzi *et al.*, 2017), making it an ideal system to study the impact of polyploidy on diversification. Here, our main objectives were to test: (1) whether *Allium* lineages with high polyploid frequencies have higher species diversification rates than lineages with low polyploid frequencies or diploids, thereby contributing to the species diversification of *Allium*; and (2) whether habitat or trait shifts are positively related to species diversification and polyploid frequency, explaining the adaptive spread of *Allium* across its current distribution. Our study is therefore designed to provide new insights into the effect of polyploidy on diversification, considering both the effects of intraspecific polyploid frequency and habitat or trait shifts.

Materials and Methods

Sampling

To build a dated phylogeny of the genus *Allium*, DNA sequences for eight plastid loci (*aptB-rbcL*, *matK*, *psbA-trnH*, *rbcL*, *rps16*, *trnL-rpl32*, *trnL-trnF,* and *trnL-trnP*) and one nuclear locus (ITS) were downloaded from GenBank (accessed 27 September 2017; Supporting Information Table S1). The sequences were aligned using MAFFT v.7.222 (Katoh & Standley, 2013) and adjusted in GENEIOUS v.11.0.4 (Kearse *et al.*, 2012). The final dataset included 448 species, 30 subspecies and 30 varieties of *Allium* covering all 15 subgenera. Sixty-five genera including taxa from three subfamilies and 15 tribes of Amaryllidaceae were sampled as outgroups.

Phylogenetic reconstruction

Maximum likelihood (ML) trees were built to test the congruence of topologies among datasets: for example, the plastid (470 species), nuclear (ITS, 539 species), and concatenated datasets (573 species) (Fig. S1). The inferred phylogenies based on either plastid or nuclear datasets did not yield any strongly supported incongruences as assessed by CONCATERPILLAR v.1.8a (Leigh *et al.*, 2008) (Fig. S1; *P*-values = 1.00 for tests of both topology and branch-length). Therefore, the concatenated dataset was used for all subsequent analyses.

The tree topology and divergence times were then estimated using BEAST v.1.8.4 (Drummond *et al.*, 2012). The general time-reversible (GTR) model was used as the nucleotide substitution rate model, an uncorrelated relaxed clock with a lognormal distribution as a clock model, and uniform prior distributions for the calibration points. As available reliable fossils are limited for *Allium* and Amaryllidaceae, secondary calibrations were used for constraining the crown nodes of Allioideae and Amaryllidaceae, with the 95% highest posterior density (HPD) defined as 27.8–44.5 million years (Myr) and 42.0–61.7 Myr, respectively (Chen *et al.*, 2013). This is in temporal agreement with the age of the first described *Allium*-like fossil (*Paleoallium billgenseli*) from the

early Eocene, dating back to 49.42 ± 0.54 million years ago (Ma) (Pigg *et al.*, 2018). We recalibrated the divergence time by defining the 95% HPD for the crown age of Amaryllidaceae as 49.4–76.7 Myr according to the age of the *P. billgenseli* fossil (Pigg *et al.*, 2018) and the estimated 95% HPD maximum age of Amaryllidaceae (Magallón *et al.*, 2015), with or without constraining the crown age of Allioideae as 27.8–44.5 Myr (Chen *et al.*, 2013), respectively. Four independent runs with 200 million Markov chain Monte Carlo (MCMC) generations were performed for each analysis. Convergence was checked using TRACER v.1.7 (Rambaut *et al.*, 2018) to make sure the effective sample size of parameters > 200. The tree files for the four runs were combined using LOGCOMBINER v.1.8.4 (BEAST Developers), with 10% of runs as burn-in and resampling at a frequency of 1/72 000. The final maximum clade credibility (MCC) tree was annotated in TREEANNOTATOR v.1.8.4 (BEAST Developers) under mean node heights.

Chromosome counts and ploidy diversity

Raw chromosome data for 568 *Allium* species were extracted from the Chromosome Counts Database (CCDB, http://ccdb.tau.ac.il/; accessed by 30 September 2017) using the R package CHROMER v.0.1 (Rice *et al.*, 2015), supplemented with chromosome data from published literature for 286 species (as of 24 October 2017; Table S2). To ensure their credibility, the corresponding species vouchers and chromosome slides were checked when available. Records under re-citation were excluded accordingly. In total, 5513 records for 401 species with phylogenetic data were available, among which 5204 records were from CCDB (312 species) and 309 records were taken from the literature (89 species) (Tables S2, S3). Among these, nearly 77% species had more than one chromosome record (with 66% species ≥ 3 records) and nearly all the remaining species with only one record were found as being diploids.

First, the basic chromosome number (x) was determined for each species, following the criterion that its frequency is over 50% out of the total number of records (in case of species with more than one x type). Second, the ploidy was defined for each taxon using the ploidy index (pi, the multiplication factor of the haploid chromosome number relative to x) (Rice *et al.*, 2015). Species with $pi > 1.4$ were defined as polyploids, as this threshold can be used to distinguish polyploids (e.g. triploids with $pi = 1.5$) from diploids ($pi = 1.0$) (Rice *et al.*, 2015). Third, the frequency of each cytotype per species was calculated for those that had more than two chromosome records, from which the frequency of the dominant polyploid type (if applicable) was used to define the intraspecific polyploid frequency applied in the following analyses. To fully determine the polyploid frequency, it would be necessary to do a substantial screening of ploidy level for each species and have a representative sampling across its native ranges. However, this was unfeasible for the present study due to the large extent of the genus' distribution area. To reduce the dimensionality of the ploidy data, a principal component analysis (PCA) was performed based on eight variables for the numbers of: (1) basic chromosomes, (2) ploidy types, and (3) dominant ploidy level; and the frequencies of: (4) dominant ploidy level, (5) dominant polyploids, (6) total polyploids, (7) polyploids with abnormal chromosome numbers (e.g. aneuploidy-like), and (8) polyploids with odd chromosome numbers (e.g. triploids) (Table S3). Individual principal components (PCs) for the first two PCA axes (ploidy PC1 and PC2) were extracted to quantify ploidy diversity across *Allium* species.

Climate, soil and trait divergence

In total, 18 728 GPS coordinates for *Allium* were retrieved from GBIF (https://www.gbif.org/; accessed 28 September 2017), eFloras (http://www.efloras.org/), eMonocot (http://www.e-monocot.org/), CVH (http://www.cvh.ac.cn/), JSTOR (https://plants.jstor.org/), and published articles (newly added 1660 GPS data for 360 species independent of GBIF). As results may be biased by the GPS distribution for *Allium* species grown as vegetable crops, only points in native ranges were included to obtain a final set of 17 318 georeferenced occurrences for 674 species (> 59% of species had ≥ 3 records).

To quantify the climatic extent of each taxon, the median values of 19 bioclimatic variables and altitude were obtained from WorldClim (http://worldclim.org) using RASTER v.2.6-6 in R (R Core Team, 2018), together with five variables, including the aridity index (AI), the annual actual or potential evapotranspiration (AAE or APE), relative humidity (ARH) and ultraviolet B (UV-B) light (AUVB) from GeoNetwork (http://www.fao.org/geonetwork/srv/en/main.home), SAGE (https://nelson.wisc.edu/sage/) and gIUV (http://www.ufz.de/gluv/). A final set of 13 low correlated variables was used for subsequent analyses (absolute Pearson's correlation coefficient $r < 0.70$, mean ± SD = 0.29 ± 0.20; Table S4), including mean diurnal temperature range, isothermality, mean temperature of wettest, warmest or coldest quarters, mean precipitation of driest, warmest or coldest quarters, altitude, AAE, AI, ARH and AUVB. The median values for 10 soil variables, such as soil pH, cation capacity, moisture (storage capacity and maximum availability), effective depth, drainage class, dominant taxonomy, production index, texture or carbon content, were retrieved from GeoNetwork and SAGE (Table S4). As described previously (Aryakia *et al.*, 2016), eight trait variables, like bulb form and size (bulb or rhizome length and width), leaf form and size (leaf length and width), scape length or perianth colour, were selected based on descriptions in eFloras, eMonocot, and publications or were measured directly from herbarium specimens from CVH and JSTOR (including 1077 specimens for 260 species, for quantitative variables only) (Table S4). Median values for each trait variable were used in the following analysis. To reduce the dimensionality of the climatic niche, soil and trait datasets, PCA was performed separately in R (R Core Team, 2018). The species scores from the first two principle components (PC1 and PC2) were extracted to capture the main variation in each dataset.

Character-independent diversification rate analysis

Bayesian Analysis of Macroevolutionary Mixtures (BAMM v.2.5.0) was used to evaluate character-independent heterogeneity in evolutionary rates on the *Allium* MCC tree after excluding outgroups (Rabosky, 2014). The global sampling fraction was set to

0.46 (448 out of the total 971). MCMC runs with 50 000 000 generations were performed, with the first 10% as burn-in. Rate shift inferences and visualisation were conducted using the R package 'BAMMTOOLS v.2.1.6' (Rabosky *et al.*, 2014).

Character-dependent diversification rate estimate

The effect of ploidy variation on diversification rates was estimated using both the hidden state speciation and extinction (HiSSE) model and the multistate HiSSE (MuHiSSE) model (Beaulieu & O'Meara, 2016). The shape or extent of correlation between diversification rates and intraspecific polyploid frequency or other continuous variables (e.g. climate, soil or trait) was estimated using the quantitative state speciation and extinction (QuaSSE) model (FitzJohn, 2010) and multiple tip rate correlations-based (TRCs-based) tests (Harvey & Rabosky, 2018). Analyses were performed in the R packages DIVERSITREE v.0.9-10 (FitzJohn, 2012), HISSE v.1.8.9 (Beaulieu & O'Meara, 2016) or using R scripts provided previously (Harvey & Rabosky, 2018). Custom codes can be found on GitHub (https://github.com/Ting-Shen/AlliumPolyploidy).

General settings Firstly, to assess the effects of polyploidy on diversification rates, species were divided into diploids and polyploids at four levels according to the quantile of dominant intraspecific polyploid frequency ($q0.00$, $q0.05$, $q0.25$, $q0.50$; frequency = 0.00, 0.13, 0.29, 0.50). A cytotype can be considered dominant when its frequency is > 20% (Kolář *et al.*, 2017); accordingly, $q0.05$ was used as a lower boundary for the delimitation of polyploid dominance in subsequent analyses. Secondly, to obtain an estimated proportion of extant species for each state, the states for samples originally missed in our dataset were evaluated using PHYLOPARS (Bruggeman *et al.*, 2009). Thirdly, a set of biologically meaningful models were assembled, among which transition rates between diploidy and polyploidy were specifically incorporated in modelling works. Intuitively, there is an irreversible relationship between diploidy and polyploidy; diploids can transition to polyploids, whereas polyploidy is unlikely to transition back ($q_{DP} = 0$) (Mayrose *et al.*, 2011). However, most *Allium* polyploids are complex with mixed-ploidy levels, including diploids (e.g. 90% of mixed-ploidy species contain diploids). As polyploidy may also originate from ancestral polyploidy (Soltis *et al.*, 2014) and may revert to diploidy through a triploid bridge, diploid−tetraploid−dihaploid cycle or postpolyploid diploidisation (Savidan & Pernès, 1982; Kolář *et al.*, 2017; Mandakova & Lysak, 2018), bidirectional transitions between 'diploid' and 'polyploid' states defined here (dominant cytotype) may occur. Therefore, we incorporated $q_{PD} = 0$ into a set of constraints for HiSSE or MuHiSSE model selection.

HiSSE and MuHiSSE To detect whether hidden factors could impact the diversification in *Allium*, 30 HiSSE models were fitted to the datasets across quantiles (Fig. S2). Twelve of these models were binary state speciation and extinction (BiSSE)-like models that excluded hidden states or constrained specific parameters of turnover, extinction or transition rates, 12 corresponded to hidden state-dependent models with various constraints on the above parameters among states, and six were null HiSSE models with various character-independent diversification (CID) forms. The performance of these models was compared using the Akaike information criteria (AIC) corrected for sample size (AICc) and Akaike weights (w_i). The estimates for each parameter of the best-fitted model were obtained through adaptive sampling of the likelihood surface for 10 000 points.

To quantify the relative importance of mixed ploidy or basic chromosome number for diversification in *Allium*, binary variation in both ploidy (diploidy vs polyploidy) and additional data about mixed ploidy (pure vs mixed) or basic chromosome number ($x = 7$ vs 8) were used in MuHiSSE models, respectively (Fig. S3). As a result, the four newly combined states for the mixed ploidy or not were diploid lineages with pure (DP) or mixed cytotypes (DM, diploids-dominated complexes), and polyploid lineages with pure (PP) or mixed cytotypes (PM, polyploids-dominated complexes); for different x were diploid lineages with $x = 7$ (D7) or $x = 8$ (D8), and polyploid lineages with $x = 7$ (P7) or $x = 8$ (P8). In total, 20 MuHiSSE models were assembled, including five multistate speciation and extinction (MuSSE)-like models and five MuHiSSE models without or with hidden states and differing in turnover or transition rates, and 10 multistate CID (MuCID) models with increasing numbers of hidden states from 2 to 8. The performance of these models was also compared using AICc and w_i. The estimates for each parameter of the best-fitted MuHiSSE model were obtained through adaptive sampling as described above. To test the effect of quantiles on diversification analysis, the best-fitted MuHiSSE model was reanalysed on 18 datasets generated under polyploid frequency from 0.05 to 0.90, with 0.05 as the interval.

TRC tests To test whether there are linear associations between continuous variables (e.g. polyploid frequency, PCs of climate, soil or trait) and tip speciation rates in *Allium*, eight TRC tests were performed, including three phylogenetic generalised least squares (PGLS) tests and five simulation tests (Harvey & Rabosky, 2018). Methods of node density (ND), terminal branch lengths (TB) and equal splits (ES) were used to estimate speciation rate for each tip, respectively. Specifically, TRCs of the inverse equal splits with simulation (*ES-sim*) were performed by three kinds of correlation tests, including Pearson's parametric test, Spearman's and Kendal's nonparametric tests. In total, 1000 simulations were run as null models for each test.

QuaSSE To detect whether ploidy, habitat or trait divergence could (additionally) drive species diversification in *Allium*, multiple QuaSSE tests were performed under different models. Five models with increasing complexity were constructed to fit the changes in speciation rates driven by quantitative variables characterised by polyploid frequency or ploidy PCs (at $q0.05$), climate PCs, soil PCs and trait PCs. Polyploid frequency was included as a continuous character in these QuaSSE analyses. The performance of these state speciation and extinction (SSE) models was compared using chi-squared tests and AIC. We only fitted speciation functions to simplify the analysis (FitzJohn, 2010).

Correlation between ploidy diversity and habitat or trait shifts

Multiple PGLS regressions were implemented to evaluate the effect of ploidy variation (explanatory variable, PC1 or PC2) on habitat or trait variation (response variable, PC1 or PC2), while correcting for phylogenetic dependence in the data. This was done using the R package CAPER v.1.0.1 (Orme *et al.*, 2013). The best transformation structure of the covariance matrix (λ, κ and δ) was selected for each PGLS pair by fitting ML models to the data. Two of the transformative parameters were fixed to determine the optimal value for the remaining parameters. The final PGLS analysis was performed using the best selected transformation structure.

Results

Phylogeny

The inferred phylogenetic tree of *Allium* based on the concatenated dataset was well resolved (Figs 1a, S1a). The phylogeny was congruent with previous studies, hypothesising that there are three main evolutionary lines (ELs) in *Allium* (Friesen *et al.*, 2006; Hauenschild *et al.*, 2017), with EL1 dominated by species from North America (NAM, 58% of EL1 species) and EUR (Europe, 14%), EL2 mainly including species from central and western Asia (CWA, 71% of EL2 species), and EL3 from a combination of CWA (35% of EL3), EA (temperate and boreal East Asia, 27%), and EUR (18%) (Fig. S1b).

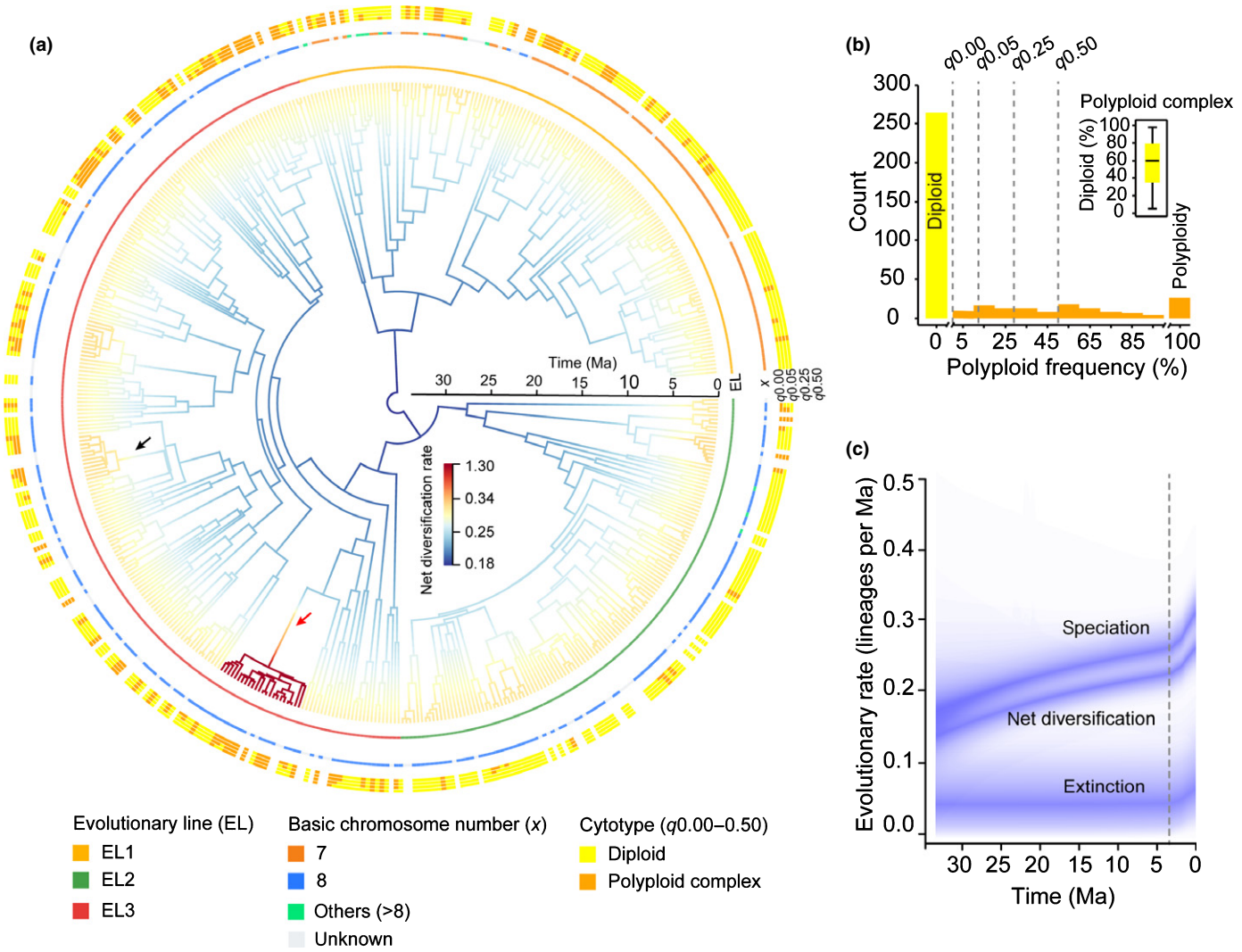

Fig. 1 Phylogeny of *Allium* with diversification rate shifts and the distribution of polyploid frequency. (a) Dated Bayesian tree of *Allium* indicating heterogeneity in the net diversification rate, the three evolutionary lines (EL1–EL3), the distribution of basic chromosome numbers (*x*), and the four quantiles of polyploid frequency (*q*0.00–0.50). The phylorate plot was obtained based on Jenks natural breaks method, in which rate variances are minimised within bins but maximised between bins. Arrows show the position of branches in the 95% credible interval of distinct shift configurations, with the most substantial shift indicated by a red arrow. Ma, million years ago. (b) Polyploid frequencies of 401 *Allium* species used for the classification of diploids (yellow) and polyploids (orange) at four quantiles; inner boxplot shows the frequency of diploids within polyploid complexes. (c) Rate-through-time density plot for *Allium*, with the dashed line showing an increase in diversification rate at 3 Ma.

The mean ages of the main branches were younger when we only used markers with less missing data or used younger ages for calibration, but the general time intervals overlapped significantly among datasets and methods with different constraints (Table S5). Amaryllidaceae originated *c.* 44.7 Ma (95% HPD: 42.0–51.0 Ma), Allioideae originated *c.* 41.4 (35.8–44.5) Ma, and *Allium* originated *c.* 33.5 (27.3–39.9) Ma. All three ELs diversified simultaneously *c.* 26.6–27.0 Ma.

Ploidy diversity

The basic chromosome number (x) in *Allium* varied from 7 to 11 (Fig. 1a; Table S6). Species in EL1 were dominated by $x = 7$ (88% of the total species in EL), while most species had $x = 8$ in EL2 (97%) and EL3 (99%) (Fig. 1a). EL1 included all five x values. Species with mixed-x accounted for 7.2% of species (17/401 species, 7 + 8; 12/401 species, 7 + 8 + others) (Table S6).

According to the ploidy index (*pi*), 65.6% of species can be defined as pure diploids (Fig. 1b), 3.7% of species can be defined as pure polyploids, and the remaining 30.6% including more than one cytotype were polyploid complexes or mixed-ploidy species (122/401 species). Polyploid complexes comprised 2.5 ± 1.1 cytotypes (mean ± SD, range = 2–9), and $2x + 4x$ was the dominant combination (63/122 species) (Table S7). The most frequent cytotype in polyploid complexes was diploidy (15/25 cytotype combinations; accounting for 60% for average) (Fig. 1b; Table S7).

For each species, we extracted scores from the first two principle components of the PCA based on eight ploidy characters (explaining 62% of the total ploidy variation) (Fig. S4). Ploidy PC1 reflected changes in intraspecific polyploid frequency and was positively linked to ploidy level and frequency. Ploidy PC2 was associated with rare cytotypes and increased with the number of basic chromosomes and the occurrence of polyploids with abnormal chromosome number.

Diversification and polyploidy

We detected two shifts in net diversification rates with a cumulative probability of 0.96 according to BAMM (Fig. 1a). The most substantial shift in the net diversification rate was observed in EL3 (probability, 95%; Fig. 1a), which includes a group of species from East Asia with higher polyploid frequencies (mean, 0.42) than random (0.19; Wilcoxon's signed-rank test: $P < 2.2e\text{-}16$; Fig. S5). Furthermore, we detected a significant increase in the net diversification rate for the whole genus at *c.* 3.0 Ma (Fig. 1c).

Based on AICc and w_i, the models with hidden states best fitted the datasets across quantiles (Table S8). At the quantile level of $q0.05$, the optimal model was HiSSE with three transition rates which included bidirectional transitions between diploid and polyploid lineages within hidden scenario, and same transition rate between diploid lineages across hidden scenarios. At the quantile levels of both $q0.25$ and $q0.50$, the best-fitted HiSSE model was one with free transition across cytotypes and hidden scenarios but under equal extinction fraction (extinction/speciation). Under these models, the net turnover/diversification rates or speciation rates for polyploids were higher than those for diploids (Figs 2, S6), and their differences increased across quantiles (within at least one hidden scenario), except the nearly equal ones under the second hidden scenario at $q0.25$. The observation remains, irrespective of hidden states or samples in the strongest rate shift detected by BAMM (Fig. S7).

The optimal MuHiSSE models fitted the datasets without interploidal transitions between pure and mixed cytotypes (e.g. no transition between DP and PM, or DM and PP), or between $x = 7$ and $x = 8$ (no transition between D7 and P8, or D8 and P7), under either equal ($q0.25$) or unequal ($q0.05$ and $q0.50$) extinction fractions (Fig. S3; Tables S9, S10). These models were selected against either nonhidden MuSSE-like models or nontrivial MuCID models (Tables S9, S10). Under them, the net turnover/diversification rates or speciation rates for diploids were always lower than those for polyploids (DP/DM < PP/PM;

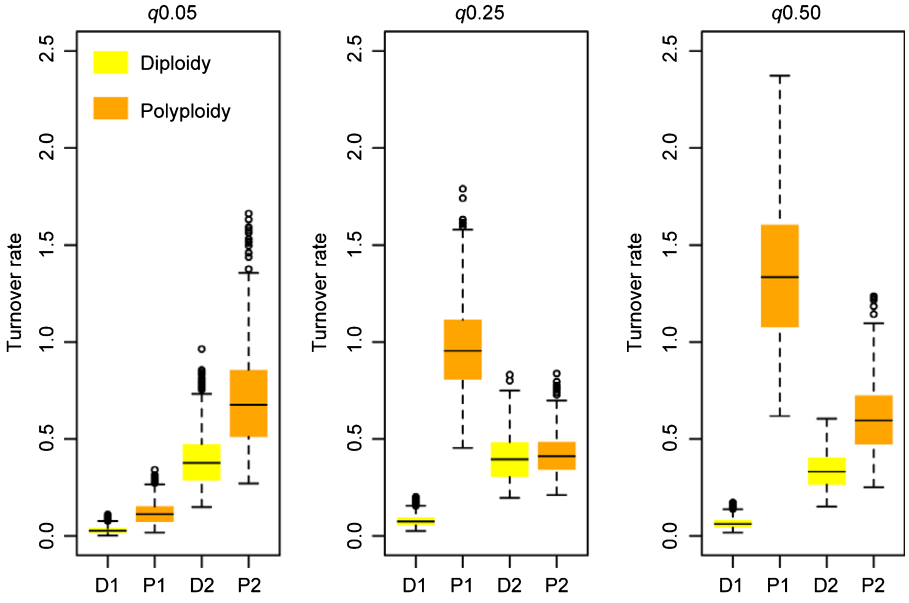

Fig. 2 Turnover rate estimates for *Allium* lineages with diploid (D, yellow) and polyploid (P, orange) cytotypes by HiSSE at $q0.05$–0.50, under the two hidden states (1 and 2). Box-and-whisker plots indicate the median (horizontal line), 25th and 75th percentiles (bottom and top of the box), and limits of the 95% confidence intervals (lower and upper whiskers) of the net turnover rates estimated through adaptive sampling of the likelihood surface for 10 000 points. Outliers beyond the 95% confidence intervals are shown as dots.

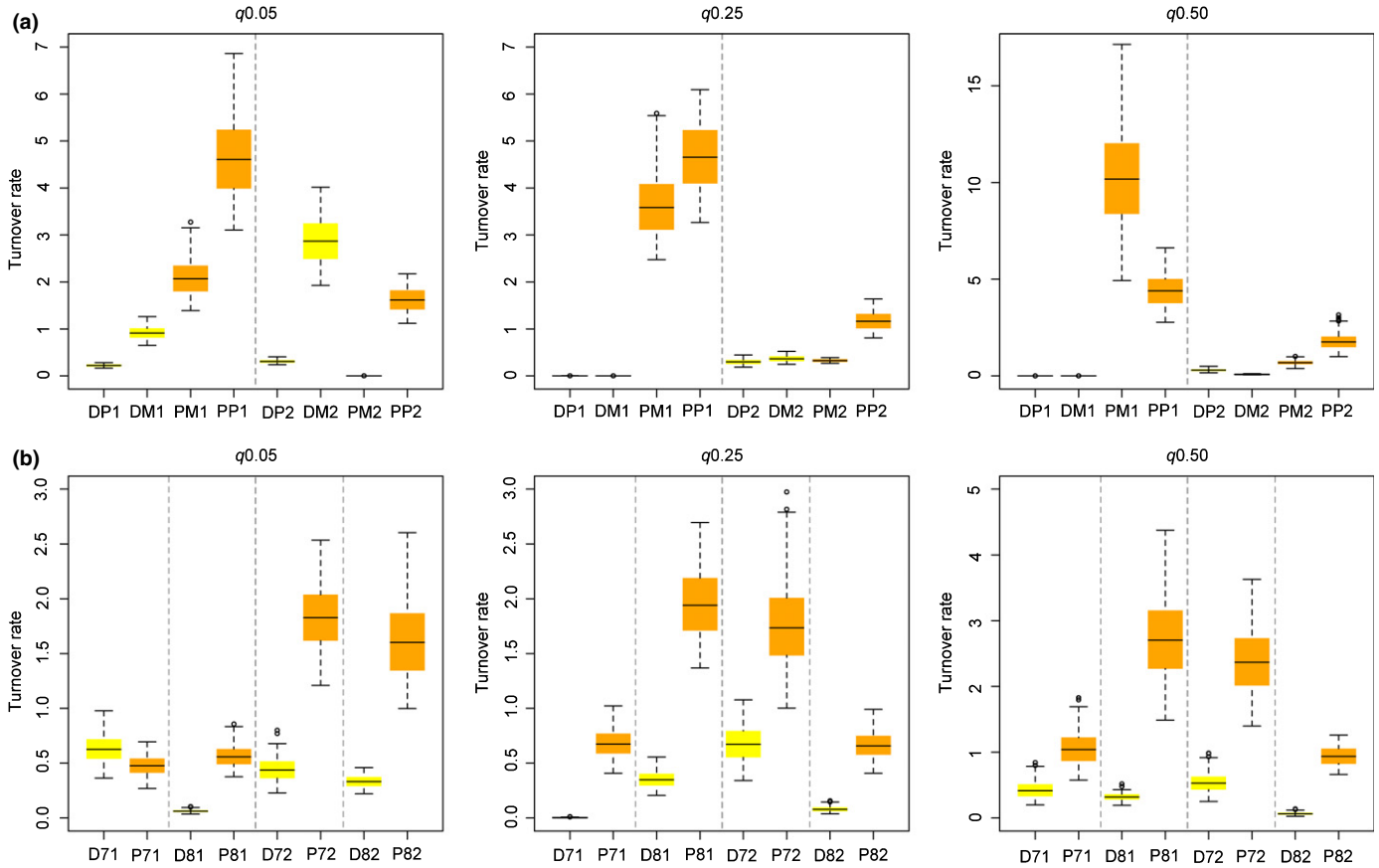

Fig. 3 Turnover rates of the best-fit MuHiSSE models at $q0.05$–0.50, under the two hidden states (1 and 2) and the additional multistate variation of mixed ploidy (a) and basic chromosome number (b) in *Allium*. The multistates of mixed-ploidy variation included diploids (yellow) with pure (DP) or mixed cytotypes (DM, diploids-dominated complexes), and polyploids (orange) with pure (PP) or mixed cytotypes (PM, polyploids-dominated complexes). The multistates of basic chromosome number variation included diploidy with $x = 7$ or 8 (D7 or D8) and polyploidy with $x = 7$ or 8 (P7 or P8). The comparison of turnover rates was divided by grey dashed lines. Box-and-whisker plots indicate the median (horizontal line), 25th and 75th percentiles (bottom and top of the box), and limits of the 95% confidence intervals (lower and upper whiskers) of the net turnover rates estimated through adaptive sampling of the likelihood surface for 10 000 points. Outliers beyond the 95% confidence intervals are shown as dots.

Figs 3, S8), no matter whether including the hidden states or the additional variation of mixed ploidy or basic chromosome number, except in one hidden scenario at $q0.05$. We found a similar pattern when using datasets without samples shown in the strongest rate shift (Fig. S9) or across quantiles of polyploid frequency (Fig. S10). Our results also indicated that in 31 out of 36 hidden scenarios (Fig. S10), the turnover rates increased from pure diploids to diploids-dominated complexes and to mixed ploidy or pure polyploids (shown as DP < DM < PM/PP), which to some extent mimicked the changes of intraspecific polyploid frequency from low to high.

Under the QuaSSE model, a positive sigmoidal (with diffusion) relationship between polyploid frequency ($q0.05$) and speciation rate best fitted the data (Fig. 4; Table S11). Under this model, the speciation rate of mixed-ploidy *Allium* lineages would not change the stably low state until their polyploid frequency reaching at 61.9% (midpoint; Fig. 4). We obtained similar results when we corrected for the influence of phylogenetic pseudoreplication (Fig. S11). Similarly, we detected a positive relationship between speciation rates and ploidy diversity represented by ploidy PC2 following a sigmoidal curve at the $q0.05$ level (Fig. S12). The TRCs results revealed that most variables cannot be linearly correlated with tip-specific speciation rates in any TRCs (mean $R^2 = 0.008$; Table S12). However, stronger speciation-related TRCs can be found in intraspecific polyploid frequency (at $q0.05$, $R^2 = 0.029$, TB-sim), although with a nonsignificant P-value (0.346).

Niche and trait divergence

We found that species niches were largely differentiated with respect to climate and soil variables (Fig. S4b,c). Diversity along the climate PC1 axis was mostly correlated with changes in temperature variability (from stable to harsh; for example BIO2, increased mean diurnal range of temperature) and diversity along climate PC2 was correlated with both temperature and precipitation (from dry to wet); together, these axes explained 49.7% of the total climatic variation in *Allium* (Fig. S4b). Soil PC1 and PC2 explained 46.6% of the total variation of soil fertility in *Allium* (Fig. S4c). The PCA based on the trait dataset indicated that trait PC1 explains 30.8% of variation in organ size from small to large, such as leaf length and width, bulb width, and

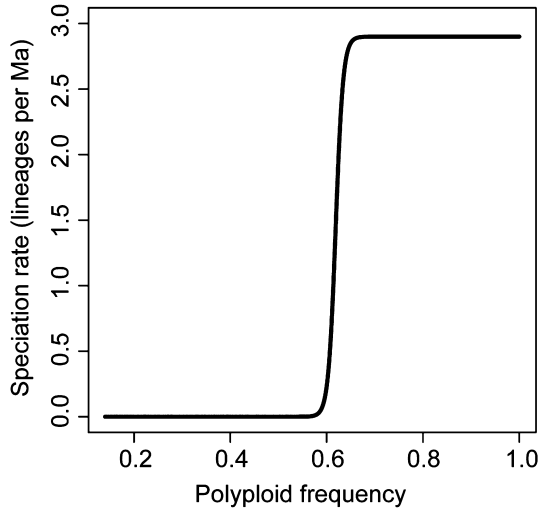

Fig. 4 Speciation rates for *Allium* lineages across polyploid frequencies (quantile, *q*0.05) based on the QuaSSE best-fitted ML model. The best-fitted ML model of the speciation rate (*y*-axis) modelled against the intraspecific polyploid frequency at quantile of *q*0.05 (*x*-axis), shown as a positive sigmoidal curve with the midpoint of 0.619 on the *x*-axis. Ma, million years ago.

scape length and trait PC2 explains 15.6% of variation in growth forms of leaves (from fistulose to solid), bulbs (from solitary to cluster), and perianth colour (from other colours to pinkish) (Fig. S4d).

The QuaSSE results indicate that speciation rates increase with temperature variability (climate PC1) and soil fertility (soil PC1 and PC2) following sigmoidal curves (Table S11; Fig. S12). In addition, speciation rates decrease with air humidity (climate PC2, linear) and organ size (trait PC1, sigmoidal). These results were partially in coincidence with TRC tests (Table S12), in which significant associations existed between tip-specific speciation rates and climate PC2 ($R^2 = 0.035$, $P = 0.046$, *ES-sim* Spearman's nonparametric test) or soil PC1 ($R^2 = 0.037$, $P = 0.036$, Pearson's parametric test).

Polyploid frequency and habitat or trait shift

Correlation between ploidy variation and habitat or trait divergence were detected in PGLS analyses (Table S13). Corrected for phylogeny, intraspecific polyploid frequency (represented by ploidy PC1) explained 81.7% and 60.2% of the variation in niches related to soil fertility and moisture (soil PC1) at the level of both *q*0.00 and *q*0.05 (Table 1; Fig. 5). Simultaneously, ploidy complexity (ploidy PC2) reflected divergence in temperature stability (climate PC1; 2.5% and 12.6%), soil character (soil PC2; 2.3% and 63%), and growth form (trait PC2; 1.9% and 11.5%) at both level of *q*0.00 and *q*0.05 (Table 1; Fig. 5).

Discussion

Accounting for polyploid frequency and intraspecific variation in ploidy, we quantified the effects of polyploidy on species diversification rates in *Allium*. We show that *Allium* lineages with high polyploid frequencies have increased diversification or speciation rates in conjunction with habitat shifts under particular soil conditions. Furthermore, we detected a significant polyploidy-related diversification rate shift in East Asia. These results suggest that the intraspecific frequency of polyploidy is a determinant of the rate of polyploid speciation and ecological radiation in *Allium*.

High ploidy diversity in *Allium*

Allium is a well known model system for cytological studies (Levan, 1931), and extensive datasets on its chromosomal variation are available (Rice *et al.*, 2015; Peruzzi *et al.*, 2017). We observed substantial differences in basic chromosome number and ploidy level among *Allium* species (Fig. 1a; Tables S6, S7). The large number of mixed-ploidy species implies that new cytotypes are generated recurrently before extinction. In *Allium*, nearly 30.6% of species show intraspecific variation in ploidy levels (Table S7), higher than the mean in angiosperms, estimated at 12–13% (Wood *et al.*, 2009) or 16.2% (Rice *et al.*, 2015). This variation may lead to reproductive isolation among offspring and facilitate speciation (Husband, 2004; Chester *et al.*, 2012; but see Zhang *et al.*, 2013). Furthermore, rare or intermediate cytotypes can influence the spatiotemporal dynamics of ploidy diversity by reproductive interactions or gene flow across ploidy levels, which may therefore facilitate polyploid diversification (Köhler *et al.*, 2010; Kolář *et al.*, 2017; Mandakova & Lysak, 2018). We observed a large number of polyploids with abnormal

Table 1 Estimates of the effects of polyploid diversity on climate, soil, and trait variation in *Allium* using phylogenetic generalised least squares (PGLS) regression.

	Ploidy PC1 (*q*0.00)		Ploidy PC2 (*q*0.00)		Ploidy PC1 (*q*0.05)		Ploidy PC2 (*q*0.05)	
Response variable	Estimated coefficient	Adjusted R^2	Estimated coefficient	Adjusted R^2	Estimated coefficient	Adjusted R^2	Estimated coefficient	Adjusted R^2
Climate PC1	0.096*	0.008	0.257***	0.025	0.113	0.000	1.008****	0.126
Climate PC2	−0.040	0.002	0.218****	0.041	−0.002	0.000	0.076	0.000
Soil PC1	0.583****	0.817	−0.070	0.002	0.806****	0.602	0.201	0.009
Soil PC2	−0.092****	0.031	0.179**	0.023	−0.111	0.022	0.701****	0.630
Trait PC1	−0.046	0.000	−0.112	0.001	−0.052	0.000	−0.281**	0.029
Trait PC2	−0.028	0.000	0.144*	0.019	−0.018	0.000	0.433**	0.115

We fixed δ = 3 across all analyses according to the estimate obtained using the maximum likelihood method.
Significance levels: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

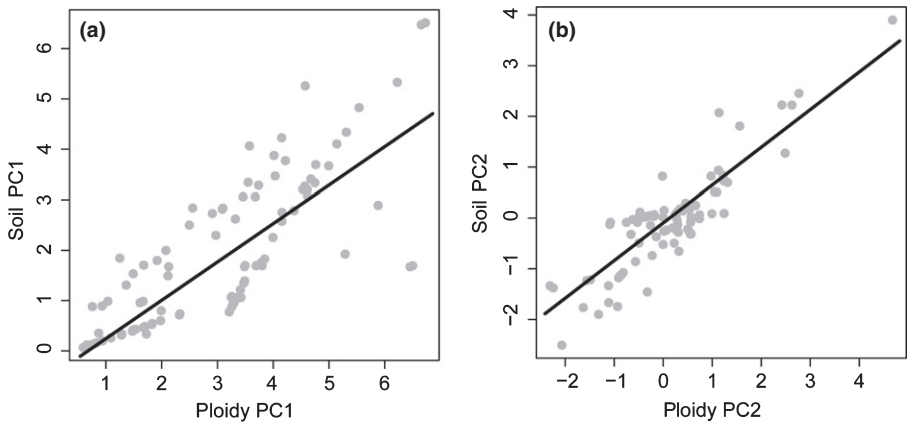

Fig. 5 Linear relationships between ploidy PC1 and soil PC1 (a) and between ploidy PC2 and soil PC2 (b) for *Allium* species, shown as scatter plots with the best fitting linear regression (solid black line).

(15.9%) or odd (28.4%) chromosome number in mixed-ploidy *Allium* species, nearly five and two times higher than those in recently surveyed plants, respectively (Kolář *et al.*, 2017). The roles of aneuploidy or odd polyploids (like triploids) in species diversification have been demonstrated in plant systems, such as *Chamerion augustifolium* (Husband, 2004) or *Tragopogon miscellus* (Chester *et al.*, 2012). This result suggests that they can facilitate polyploid speciation through promoting the polyploidy–diploidy coexistence or contributing to genetic variation needed by neopolyploids. Therefore, these chromosomal variations can provide substantial opportunity for divergence among *Allium* plants (Leitch & Leitch, 2008; Soltis *et al.*, 2014).

Frequency-based polyploid diversification

Polyploidisation is a major mechanism of speciation (Rothfels & Otto, 2016). A high polyploid frequency is predicted to accelerate speciation, whereas a low polyploid frequency will induce extinction at an early age as a result of MCE (Levin, 1975; Baduel *et al.*, 2018). Our results support this hypothesis by demonstrating a high species diversification rate for *Allium* lineages with high polyploid frequencies (Fig. 2–4), particularly when polyploids are dominant in mixed-ploidy complexes (e.g. at the level of $q0.50$). We found that diversification shifts can be predicted by the observed characters of polyploidisation, which have been detected across both hidden states (Fig. 2) and basic chromosome numbers (Fig. 3b). Our implemented MuHiSSE and QuaSSE models further proved that intraspecific polyploid frequency in state of mixed ploidy can influence the likelihood of species diversification in *Allium* (Figs 3a, 4). Importantly, BAMM analysis detected a significant polyploidy-related diversification rate shift along the *Allium* phylogeny (red arrow in Fig. 1a). These observations indicate several aspects about the origin and establishment of polyploid species. First, the origin of polyploidy largely depends on a high frequency of unreduced gametes and individuals with the same ploidy level (Ramsey & Schemske, 1998; Kreiner *et al.*, 2017). And then, polyploid populations can be established and stabilised when MCE challenges are overcome at an early stage (Husband, 2000). Second, elevated polyploid frequencies can contribute to avoiding the risk of a small population size and enhance reproductive connectivity, thereby promoting demographic stability or local adaptation within population (Ramsey & Schemske, 2002; Fowler & Levin, 2016; Levin, 2019). Third, adaptive introgression may take place across ploidy levels during secondary contact (Petit *et al.*, 1999; Chapman & Abbott, 2010). Therefore, the frequent occurrence of polyploids in *Allium* can contribute to rapid speciation and impact diversification.

Polyploid frequency-based diversification can be achieved through recurrent origins (Soltis & Soltis, 1999), triploid bridges (Köhler *et al.*, 2010), asexual reproduction (Freeling, 2017), or perennial life forms (Stebbins, 1950). Recurrent origins of polyploidy are common in *Allium*, given the high rate of unreduced pollen grains (Levan, 1931). For example, the tetraploid *A. przewalskianum* Regel arose through autopolyploidisation at least eight times independently, enabling its colonisation and survival in the arid habitats of the Qinghai−Tibetan Plateau (Wu *et al.*, 2010). In addition, potential triploid bridges may provide an important genetic connection between distinct cytotypes given the high occurrence of odd polyploids (28.4%) in *Allium*, which can contribute to the stabilisation of newly generated rare cytotypes. Furthermore, asexual reproduction through bulbs or rhizomes and perennial life forms are common characters well established by the latest early Eocene in *Allium* (Pigg *et al.*, 2018), which may promote polyploid survival at early stages or during the establishment of populations (Van Drunen & Husband, 2019).

A higher proportion of polyploidy (> 40%) than diploidy can be achieved in a mixed-ploidy population with nonlinear effects of fertility or unreduced gametes on their coexistence (Suda & Herben, 2013). Higher fitness (than that of diploids) can be obtained when polyploids are in the majority (e.g. 67% in *C. angustifolium*; Husband, 2000). Our results indicate a positive sigmoidal relationship between polyploid frequency and speciation rate, switching at the midpoint of 61.9% (Fig. 4). This suggests that the speciation rate will increase quickly in response to increases in polyploid frequency at this point, possibly through the acquisition of adaptive traits, such as flower or ovule numbers, as suggested previously for other species (Husband, 2000).

Role of polyploids in ecological radiations and diversification in *Allium*

Polyploids are proposed to inhabit new ecological niches given their genetic plasticity, suggesting a causal relationship between polyploidy and ecological radiations in plants (Leitch & Leitch, 2008; Van de Peer *et al.*, 2017). Except for a few case studies (Ramsey, 2011; Godfree *et al.*, 2017), we know little information about the functional interaction between these two processes. In *Allium*, previous work has reported climatically driven diversification (Cui *et al.*, 2008; Wu *et al.*, 2010; Huang *et al.*, 2014; Wang *et al.*, 2015; Herden *et al.*, 2016; M-J. Li *et al.*, 2016; Q-Q. Li *et al.*, 2016; Zhang *et al.*, 2017; Xie *et al.*, 2019). However, few studies have ever quantified these processes to identify the underlying mechanisms (Hauenschild *et al.*, 2017).

Our results indicate that *Allium* speciation rates can increase with habitat shifts into either dry climates or fertile soils as well as with trait shifts to small organs (Fig. S12). Species in dry habitats usually occur in arid or semiarid areas, whereas species in wet habitats often occur near the moist alpine or subalpine grasslands in high mountains (Fritsch & Fritsch, 2002). During the late Miocene to Pliocene (*c.* 5–3 Ma), global climate fluctuations led to the expansion of cooler and drier habitats, allowing species, such as those of *Allium*, to colonise these habitats and diversify (Fig. 1c) (Herden *et al.*, 2016; M-J. Li *et al.*, 2016; Q-Q. Li *et al.*, 2016; Xie *et al.*, 2019). *Allium* diversification was also accompanied by the divergence of traits related to drought tolerance, such as reduced bulb or leaf sizes, suggesting that ecologically driven diversification is associated with functional traits, as it has been shown in other plant groups (Onstein *et al.*, 2016; Ebersbach *et al.*, 2017). Interestingly, we also found a positive relationship between *Allium* speciation rates and soil fertility (Table S12; Fig. S12). This implies that belowground environments can also influence diversification rates of bulbiferous species, such as *Allium* (Veselý *et al.*, 2011).

Based on the evolutionary advantages of polyploids over diploids and the availability of newly opened habitats (Wu *et al.*, 2010; M-J. Li *et al.*, 2016), polyploidy may have played a crucial role during the ecological radiation of *Allium*. The colonisation of new habitats allows polyploids to avoid competition with their diploid ancestors (Parisod & Broennimann, 2016) and contributes to the establishment of an ecological edge to reduce the effects of MCE (Levin, 1975). We showed that polyploid diversification can be driven by polyploid frequency-related shifts in habitats or traits, particularly concerning edaphic conditions (Fig. 5). They would act as the hidden states into the process of species diversification driven by polyploid frequency. In *Allium*, at least 60.2% of variation in soil fertility and moisture can be explained by polyploid frequency (at q0.05; 81.7% at q0.00) after correcting for the confounding effect of phylogeny (Table 1; Fig. 5). Previous studies have indicated that polyploidisation can directly interact with either abiotic (e.g. salinity stress) or biotic (e.g. microbes) factors in soil through polyploidisation-induced changes in salinity tolerance or species interaction (Chao *et al.*, 2013; Segraves & Anneberg, 2016). These interactions are important for understanding the elevated polyploid speciation rate in habitats under fertile conditions (Walczyk & Hersch-Green, 2019). Additionally, our results indicate that rare cytotypes, such as aneuploids/dysploids or polyploids with uncommon basic chromosome numbers (e.g. ploidy PC2), are more likely to accumulate with climatic shifts from stably wet to extremely dry environments (e.g. climate PC1) (Table 1). These kinds of genomic plasticity imply that newly generated or hybridised polyploids are prone to occur under stress (Ramsey, 2011), which may contribute to polyploid diversification in new ecological niches (Chester *et al.*, 2012; Zhang *et al.*, 2013). However, more work should be done to fully understand the driving forces behind the diversification of *Allium*, especially concerning the accurate quantification of the contribution of ecology or polyploidy to it at population level (Ebersbach *et al.*, 2018).

Conclusion

Despite extensive studies, the association between polyploidy and species diversification remains enigmatic (Kellogg, 2016). We quantified ploidy variation in the genus *Allium* and demonstrated that diversification in the genus can be driven by intraspecific polyploid frequency-based forces through extrinsic habitat shifts (climatic or edaphic). Given the frequent occurrence of mixed-ploidy species in *Allium*, these results are likely to reflect an ongoing process of polyploid evolution (Kolář *et al.*, 2017). We recognise that our datasets might suffer from sampling biases, which may partially affect our inference. However, given that the number of diploid records was much higher than those of polyploids (Fig. 1b), the phenomenon of polyploid frequency-based diversification would be underestimated accordingly. Intriguingly, we still detected the interaction between polyploid frequency and species diversification, which suggests the limited influence of sampling bias on the issue of robustness of our observation. This study therefore provides new insights into the role of polyploidy in diversification.

Acknowledgements

We are grateful to all the collectors of *Allium* chromosome data. We thank the Public Technology Service Centre of XTBG. This work was supported by the National Natural Science Foundation of China (U1802242, 31800177); and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-FZT 118 to REO and ANM-R. T-SH is also supported by China Scholarship Council and Y-WX is supported by National Key R&D Program of China (2017YFC0505200) and CAS 135 programme (2017XTBG-F01). We thank Yanis Bouchenak-Khelladi and four anonymous reviewers for discussions and constructive comments on the manuscript, which highly improved this work.

Author contributions

T-SH and Y-WX designed the research. T-SH collected the data and performed analyses. Q-JZ collected the data. REO provided analysis tools. T-SH and Y-WX wrote the first draft of the

manuscript, and REO, BMR-A, FH and ANM-R contributed to writing. All authors reviewed the final manuscript.

ORCID

Ting-Shen Han https://orcid.org/0000-0002-8612-6581
Frank Hauenschild https://orcid.org/0000-0002-4870-0108
Alexandra N. Muellner-Riehl https://orcid.org/0000-0002-2710-469X
Renske E. Onstein https://orcid.org/0000-0002-2295-3510
Blanca M. Rojas-Andrés https://orcid.org/0000-0001-7164-1313
Yao-Wu Xing https://orcid.org/0000-0001-6709-4492

References

Abbott RJ, Lowe AJ. 2004. Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biological Journal of the Linnean Society* **82**: 467–474.

Aryakia E, Karimi HR, Naghavi MR, Fazeli SAS. 2016. Morphological characterization of intra- and interspecific diversity in some Iranian wild *Allium* species. *Euphytica* **211**: 185–200.

Baduel P, Bray S, Vallejo-Marin M, Kolář F, Yant L. 2018. The "Polyploid Hop": shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Frontiers in Ecology and Evolution* **6**: 1–19.

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

Bruggeman J, Heringa J, Brandt BW. 2009. PhyloPars: estimation of missing parameter values using phylogeny. *Nucleic Acids Research* **37**(suppl_2): W179–W184.

Cai L, Xi Z, Amorim AM, Sugumaran M, Rest JS, Liu L, Davis CC. 2019. Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytologist* **221**: 565–576.

Čertner M, Sudová R, Weiser M, Suda J, Kolář F. 2019. Ploidy-altered phenotype interacts with local environment and may enhance polyploid establishment in *Knautia serpentinicola* (Caprifoliaceae). *New Phytologist* **221**: 1117–1127.

Chao D-Y, Dilkes B, Luo H, Douglas A, Yakubova E, Lahner B, Salt DE. 2013. Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis*. *Science* **341**: 658–659.

Chapman MA, Abbott RJ. 2010. Introgression of fitness genes across a ploidy barrier. *New Phytologist* **186**: 63–71.

Chen S, Kim D-K, Chase MW, Kim J-H. 2013. Networks in a large-scale phylogenetic analysis: reconstructing evolutionary history of Asparagales (Lilianae) based on four plastid genes. *PLoS ONE* **8**: e59472.

Chester M, Gallagher JP, Symonds VV, da Silva AVC, Mavrodiev EV, Leitch AR, Soltis PS, Soltis DE. 2012. Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae). *Proceedings of the National Academy of Sciences, USA* **109**: 1176–1181.

Cui X-K, Ao C-Q, Zhang Q, Chen L-T, Liu J-Q. 2008. Diploid and tetraploid distribution of *Allium przewalskianum* Regel. (Liliaceae) in the Qinghai-Tibetan Plateau and adjacent regions. *Caryologia* **61**: 192–200.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.

Ebersbach J, Muellner-Riehl A, Favre A, Paule J, Winterfeld G, Schnitzler J. 2018. Driving forces behind evolutionary radiations: Saxifraga section Ciliatae (Saxifragaceae) in the region of the Qinghai–Tibet Plateau. *Botanical Journal of the Linnean Society* **186**: 304–320.

Ebersbach J, Schnitzler J, Favre A, Muellner-Riehl A. 2017. Evolutionary radiations in the species-rich mountain genus *Saxifraga* L. *BMC Evolutionary Biology* **17**: 119.

Estep MC, McKain MR, Diaz DV, Zhong J, Hodge JG, Hodkinson TR, Layton DJ, Malcomber ST, Pasquet R, Kellogg EA. 2014. Allopolyploidy, diversification, and the Miocene grassland expansion. *Proceedings of the National Academy of Sciences, USA* **111**: 15149–15154.

FitzJohn RG. 2010. Quantitative traits and diversification. *Systematic Biology* **59**: 619–633.

FitzJohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3**: 1084–1092.

Fowler NL, Levin DA. 2016. Critical factors in the establishment of allopolyploids. *American Journal of Botany* **103**: 1236–1251.

Freeling M. 2017. Picking up the ball at the K/Pg boundary: the distribution of ancient polyploidies in the plant phylogenetic tree as a spandrel of asexuality with occasional sex. *Plant Cell* **29**: 202–206.

Friesen N, Fritsch RM, Blattner FR. 2006. Phylogeny and new intrageneric classification of *Allium* (Alliaceae) based on nuclear ribosomal DNA ITS sequences. *Aliso* **22**: 372–395.

Fritsch RM, Fritsch N. 2002. Evolution, domestication and taxonomy. In: Rabinowitch HD, Currah L, eds. *Allium crop science: recent advances*. London, UK: CABI Publishing, 5–30.

Godfree RC, Marshall DJ, Young AG, Miller CH, Mathews S. 2017. Empirical evidence of fixed and homeostatic patterns of polyploid advantage in a keystone grass exposed to drought and heat stress. *Royal Society Open Science* **4**: 170934.

Han T-S, Wu Q, Hou X-H, Li Z-W, Zou Y-P, Ge S, Guo Y-L. 2015. Frequent introgressions from diploid species contribute to the adaptation of the tetraploid Shepherd's purse (*Capsella bursa-pastoris*). *Molecular Plant* **8**: 427–438.

Harvey MG, Rabosky DL. 2018. Continuous traits and speciation rates: alternatives to state-dependent diversification models. *Methods in Ecology and Evolution* **9**: 984–993.

Hauenschild F, Favre A, Schnitzler J, Michalak I, Freiberg M, Muellner-Riehl AN. 2017. Spatio-temporal evolution of *Allium* L. in the Qinghai–Tibet-Plateau region: immigration and *in situ* radiation. *Plant Diversity* **39**: 167–179.

Herden T, Hanelt P, Friesen N. 2016. Phylogeny of *Allium* L. subgenus *Anguinum* (G. Don. ex WDJ Koch) N. Friesen (Amaryllidaceae). *Molecular Phylogenetics and Evolution* **95**: 79–93.

Huang D-Q, Li Q-Q, Zhou C-J, Zhou SD, He X-J. 2014. Intraspecific differentiation of *Allium wallichii* (Amaryllidaceae) inferred from chloroplast DNA and internal transcribed spacer fragments. *Journal of Systematics and Evolution* **52**: 341–354.

Husband BC. 2000. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**: 217–223.

Husband BC. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biological Journal of the Linnean Society* **82**: 537–546.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.

Kellogg EA. 2016. Has the connection between polyploidy and diversification actually been tested? *Current Opinion in Plant Biology* **30**: 25–32.

Köhler C, Scheid OM, Erilova A. 2010. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics* **26**: 142–148.

Kolář F, Čertner M, Suda J, Schönswetter P, Husband BC. 2017. Mixed-ploidy species: progress and opportunities in polyploid research. *Trends in Plant Science* **22**: 1041–1055.

Kreiner JM, Kron P, Husband BC. 2017. Frequency and maintenance of unreduced gametes in natural plant populations: associations with reproductive mode, life history and genome size. *New Phytologist* **214**: 879–889.

Leigh JW, Susko E, Baumgartner M, Roger AJ. 2008. Testing congruence in phylogenomic analysis. *Systematic Biology* **57**: 104–115.

Leitch A, Leitch I. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* **320**: 481–483.

Levan A. 1931. Cytological studies in *Allium*: a preliminary note. *Hereditas* **15**: 347–356.

Levin DA. 1975. Minority cytotype exclusion in local plant populations. *Taxon* **24**: 35–43.

Levin DA. 2019. Why polyploid exceptionalism is not accompanied by reduced extinction rates. *Plant Systematics and Evolution* **305**: 1–11.

Levin DA, Soltis DE. 2018. Factors promoting polyploid persistence and diversification and limiting diploid speciation during the K-Pg interlude. *Current Opinion in Plant Biology* **42**: 1–7.

Li M-J, Tan J-B, Xie D-F, Huang D-Q, Gao Y-D, He X-J. 2016. Revisiting the evolutionary events in *Allium* subgenus *Cyathophora* (Amaryllidaceae): insights into the effect of the Hengduan Mountains Region (HMR) uplift and Quaternary climatic fluctuations to the environmental changes in the Qinghai–Tibet Plateau. *Molecular Phylogenetics and Evolution* **94**: 802–813.

Li Q-Q, Zhou S-D, Huang D-Q, He X-J, Wei X-Q. 2016. Molecular phylogeny, divergence time estimates and historical biogeography within one of the world's largest monocot genera. *Annals of Botany* **8**: 1–17.

Luebert F, Weigend M. 2014. Phylogenetic insights into Andean plant diversification. *Frontiers in Ecology and Evolution* **2**: 1–17.

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

Mandakova T, Lysak MA. 2018. Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology* **42**: 55–65.

Marchant DB, Soltis DE, Soltis PS. 2016. Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytologist* **212**: 708–718.

Marques I, Loureiro J, Draper D, Castro M, Castro S. 2018. How much do we know about the frequency of hybridisation and polyploidy in the Mediterranean region? *Plant Biology* **20**: 21–37.

Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, Otto SP. 2011. Recently formed polyploid plants diversify at lower rates. *Science* **333**: 1257.

Onstein RE, Jordan GJ, Sauquet H, Weston PH, Bouchenak-Khelladi Y, Carpenter RJ, Linder HP. 2016. Evolutionary radiations of Proteaceae are triggered by the interaction between traits and climates in open habitats. *Global Ecology and Biogeography* **25**: 1239–1251.

Orme D, Freckleton R, Thomas G, Petzoldt T. 2013. The caper package: comparative analysis of phylogenetics and evolution in R. *R Package v.5* [WWW document] URL https://CRAN.R-project.org/package=caper.

Parisod C, Broennimann O. 2016. Towards unified hypotheses of the impact of polyploidy on ecological niches. *New Phytologist* **212**: 540–542.

Peruzzi L, Carta A, Altinordu F. 2017. Chromosome diversity and evolution in *Allium* (Allioideae, Amaryllidaceae). *Plant Biosystems* **151**: 212–220.

Petit C, Bretagnolle F, Felber F. 1999. Evolutionary consequences of diploid–polyploid hybrid zones in wild species. *Trends in Ecology & Evolution* **14**: 306–311.

Pigg KB, Bryan FA, DeVore ML. 2018. *Paleoallium billgenseli* gen. et sp. nov.: Fossil Monocot remains from the Latest Early Eocene Republic Flora, Northeastern Washington State, USA. *International Journal of Plant Sciences* **179**: 477–486.

R Core Team. 2018. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* **9**: 1–15.

Rabosky DL, Grundler M, Anderson C, Title P, Shi JJ, Brown JW, Huang H, Larson JG. 2014. BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* **5**: 701–707.

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

Ramsey J. 2011. Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences, USA* **108**: 7096–7101.

Ramsey J, Ramsey TS. 2014. Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**: 1–20.

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29**: 467–501.

Ramsey J, Schemske DW. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* **33**: 589–639.

Ren R, Wang H, Guo C, Zhang N, Zeng L, Chen Y, Ma H, Qi J. 2018. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Molecular Plant* **11**: 414–428.

Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, Mayzel J, Chay O, Mayrose I. 2015. The Chromosome Counts Database (CCDB)–a community resource of plant chromosome numbers. *New Phytologist* **206**: 19–26.

Rice A, Šmarda P, Novosolov M, Drori M, Glick L, Sabath N, Meiri S, Belmaker J, Mayrose I. 2019. The global biogeography of polyploid plants. *Nature Ecology & Evolution* **3**: 265–273.

Rodriguez DJ. 1996. A model for the establishment of polyploidy in plants. *The American Naturalist* **147**: 33–46.

Rothfels CJ, Otto SP. 2016. Polyploid speciation. In: Kliman RM, ed. *Encyclopedia of evolutionary biology*. Oxford, UK: Academic Press, 317–326.

Salman-Minkov A, Sabath N, Mayrose I. 2016. Whole-genome duplication as a key factor in crop domestication. *Nature Plants* **2**: 1–4.

Savidan Y, Pernès J. 1982. Diploid-tetraploid-dihaploid cycles and the evolution of *Panicum maximum* Jacq. *Evolution* **36**: 596–600.

Schranz ME, Mohammadin S, Edger PP. 2012. Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. *Current Opinion in Plant Biology* **15**: 147–153.

Segraves KA. 2017. The effects of genome duplications in a community context. *New Phytologist* **215**: 57–69.

Segraves KA, Anneberg TJ. 2016. Species interactions and plant polyploidy. *American Journal of Botany* **103**: 1326–1335.

Smith SA, Brown JW, Yang Y, Bruenn R, Drummond CP, Brockington SF, Walker JF, Last N, Douglas NA, Moore MJ. 2018. Disparity, diversity, and duplications in the Caryophyllales. *New Phytologist* **217**: 836–854.

Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, Majure L, Miles NM, Mavrodiev EV, Mei W, Cortez MB, Soltis PS, Gitzendanner MA. 2014. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose *et al.* (2011). *New Phytologist* **202**: 1105–1117.

Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution* **14**: 348–352.

Soltis PS, Soltis DE. 2016. Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology* **30**: 159–165.

Stebbins GL. 1950. *Variation and evolution in plants.* New York, NY, USA: Columbia University Press.

Suda J, Herben T. 2013. Ploidy frequencies in plants with ploidy heterogeneity: fitting a general gametic model to empirical population data. *Proceedings of the Royal Society B: Biological Sciences* **280**: 1–10.

Suda J, Meyerson LA, Leitch IJ, Pyšek P. 2015. The hidden side of plant invasions: the role of genome size. *New Phytologist* **205**: 994–1007.

Van de Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics* **18**: 411–424.

Van Drunen WE, Husband BC. 2019. Evolutionary associations between polyploidy, clonal reproduction, and perenniality in the angiosperms. *New Phytologist* **224**: 1266–1277.

Veselý P, Bureš P, Šmarda P, Pavlíček T. 2011. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.

Walczyk AM, Hersch-Green EI. 2019. Impacts of soil nitrogen and phosphorus levels on cytotype performance of the circumboreal herb *Chamerion angustifolium*: implications for polyploid establishment. *American Journal of Botany* **106**: 1–16.

Wang X, Li Y, Liang Q, Zhang L, Wang Q, Hu H, Sun Y. 2015. Contrasting responses to Pleistocene climate changes: a case study of two sister species *Allium cyathophorum* and *A. spicata* (Amaryllidaceae) distributed in the eastern and western Qinghai–Tibet Plateau. *Ecology and Evolution* **5**: 1513–1524.

Wei N, Cronn R, Liston A, Ashman TL. 2019. Functional trait divergence and trait plasticity confer polyploid advantage in heterogeneous environments. *New Phytologist* **221**: 2286–2297.

Wen J, Zhang J, Nie Z-L, Zhong Y, Sun H. 2014. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Frontiers in Genetics* **5**: 1–16.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences, USA* **106**: 13875–13879.

Wu L-L, Cui X-K, Milne RI, Sun Y-S, Liu J-Q. 2010. Multiple autopolyploidizations and range expansion of *Allium przewalskianum* Regel. (Alliaceae) in the Qinghai-Tibetan Plateau. *Molecular Ecology* **19**: 1691–1704.

Xie C, Xie D-f, Zhong Y, Guo X-L, Liu Q, Zhou S-D, He X-J. 2019. The effect of Hengduan Mountains Region (HMR) uplift to environmental changes in the HMR and its eastern adjacent area: tracing the evolutionary history of *Allium* section *Sikkimensia* (Amaryllidaceae). *Molecular Phylogenetics and Evolution* **130**: 380–396.

Xing Y, Ree RH. 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proceedings of the National Academy of Sciences, USA* **114**: E3444–E3451.

Zhang H, Bian Y, Gou X, Zhu B, Xu C, Qi B, Li N, Rustgi S, Zhou H, Han F. 2013. Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat. *Proceedings of the National Academy of Sciences, USA* **110**: 3447–3452.

Zhang Y, Yu Q, Zhang Q, Hu X, Hu J, Fan B. 2017. Regional-scale differentiation and phylogeography of a desert plant *Allium mongolicum* (Liliaceae) inferred from chloroplast DNA sequence variation. *Plant Systematics and Evolution* **303**: 451–466.

Supporting Information

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

Fig. S1 Phylogenetic reconstruction of *Allium*.

Fig. S2 HiSSE models for testing the effect of a hidden binary trait (diploidy vs polyploidy) on *Allium* diversification.

Fig. S3 MuHiSSE models for testing the combined effect of binary trait (diploidy vs polyploidy) and mixed-ploidy (or basic chromosome number) variation on *Allium* diversification.

Fig. S4 PCA biplots of ploidy (a), climatic (b), soil (c) and trait (d) variation, with the annotation of their abbreviations (e).

Fig. S5 Test for the accumulation of polyploids on branches with the strongest signal of a rate shift detected by BAMM.

Fig. S6 Diversification rates across *Allium* using HiSSE.

Fig. S7 Diversification rates of HiSSE across *Allium* without samples in the strongest rate shift detected by BAMM.

Fig. S8 Diversification rates across *Allium* using MuHiSSE.

Fig. S9 Diversification rates of MuHiSSE across *Allium* without samples in the strongest rate shift detected by BAMM.

Fig. S10 Turnover rates across *Allium*.

Fig. S11 Speciation rates across polyploid frequencies (quantile, $q0.05$) based on QuaSSE best-fit maximum likelihood model.

Fig. S12 Relationships between speciation rate and ploidy principal components (PCs), climate PCs, soil PCs and trait PCs in *Allium* based on QuaSSE analyses.

Table S1 GenBank accession numbers for DNA sequences of *Allium* and outgroups used for the phylogenetic reconstruction.

Table S2 Chromosome data from published literature.

Table S3 Ploidy variables for 401 *Allium* species used in this study.

Table S4 Climate, soil and trait datasets used in this study.

Table S5 Divergence times based on the maximum clade credibility (MCC) tree obtained using BEAST.

Table S6 Count and frequency of each basic chromosome number (x).

Table S7 Count and frequency of different cytotype combinations.

Table S8 HiSSE models and model selection for the relationship between polyploid frequency and diversification rates in *Allium*.

Table S9 MuHiSSE models and model selection for the relationship between polyploid frequency plus with mixed-ploidy variation and diversification rates in *Allium*.

Table S10 MuHiSSE models and model selection for the relationship between polyploid frequency plus with basic chromosome number variation and diversification rates in *Allium*.

Table S11 Models and model selection for the relationship between ploidy, climate, soil and trait characters and diversification rates at the whole genus level.

Table S12 Summary estimates of tip rate correlation tests.

Table S13 Best transformation structure of the covariance matrix (λ, κ and δ) for each phylogenetic generalised least squares pair by the maximum likelihood method.

Please note: Wiley Blackwell are 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.