Karyotype reconstruction provides new insights into polyploid evolution in grasses (Poaceae)

Longjiang Fan

fanlj@zju.edu.cn

Zhejiang University https://orcid.org/0000-0003-4846-0500

Haijun Cheng

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Xuefang Huang

Institute of Biotechnology, Zhejiang University

Yujie Huang

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Linfeng Jin

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Qinjie Chu

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Kun Yang

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Shiyu Zhang

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Shenjun Zhu

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Ruiyao Zhou

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Zhefu Li

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Yinuo Zhao

Institute of Crop Sciences & Institute of Bioinformatics, Zhejiang University

Dongya Wu

Center for Evolutionary & Organismal Biology, Zhejiang University

Pengchuan Sun

College of Life Sciences, Sichuan University

Research Article

Keywords: Ancestral karyotype, Chromosome fusion, Whole-genome duplication, Poaceae

Posted Date: March 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8993757/v1

Additional Declarations: The authors declare no competing interests.

Abstract

Chromosomal rearrangements following whole-genome duplication (WGD) shape the diversity of genome structures in Poaceae. An accurate karyotype construction based on high-quality genomes of extant species will facilitate a deeper understanding of the origin and evolution of Poaceae. Here, we determine ancestral karyotypes for the BOP and PACMAD clades (ABPK, $n = 12$) and further infer ancestral karyotypes for subfamilies and tribes with distinct chromosome structures. Evolutionary trajectories from ancestral karyotypes to modern genomes are reconstructed for 29 species across 7 subfamilies through the characterization of multiple cascading fusion events. Our results reveal that subfamily-specific fusions contribute to the divergence and establishment of distinct subfamilies, while shared fusions among closely related species within each subfamily reflect the stable inheritance of karyotypic architecture. Based on the comparative karyotype analysis, we elucidate that *Cenchrus purpureus* and *C. fungigraminus* share two diploid progenitors contributing to their respective subgenomes A and B, and that Paleotetraploid *S. maritimus* and *S. alterniflorus* share the first allotetraploid WGD (WGD1) but experience independent second WGD (WGD2) events. Our results highlight the importance of chromosome fusion in descending dysploidy after recurrent polyploidization and shed light on the polyploidization and karyotype evolution across the Poaceae family.

Introduction

Poaceae, one of the most ecologically and economically significant plant lineages, ranks as the fifth-largest angiosperm family, encompassing approximately 12,000 species distributed among 12 subfamilies[1–4]. The family's diversity is structured around two major clades: BOP (Bambusoideae, Oryzoideae, and Pooideae) and PACMAD (Panicoideae, Arundinoideae, Chloridoideae, Micraioideae, Aristidoideae, and Danthonioideae). Additionally, three early-diverging subfamilies (Anomochlooideae, Pharoideae, and Puelioideae) occupy successive sister positions to the main BOP-PACMAD lineages[4–6].

In general, the Poaceae family has undergone three rounds of WGD events (σ, τ, and ρ)[7–11]. While the first two events are shared with other monocots, the third ρ-WGD event is specific to Poaceae and is thought to have driven the family's extensive adaptive radiation. Following the ρ-WGD event, numerous additional species-specific polyploidization events have occurred in the Poaceae family, such as tetraploidization and hexaploidization in *Echinochloa* species[12]. Polyploidization plays an essential role in plant evolution by providing raw genetic material for functional innovation and adaptive diversification[13]. This process has been recognized as a major driver of speciation, contributing to reproductive isolation and facilitating rapid adaptation to environmental changes[14,15]. Autopolyploids are formed through genome duplication, while allopolyploids arise from hybridization between different species with independent evolutionary histories, both of which are accompanied by chromosome doubling and gamete fusion[16]. Among polyploid species, allopolyploids are predominant, accounting for nearly 90% of cases, whereas autopolyploids represent only a small proportion[16,17]. The initial evolutionary stages of recently formed allotetraploids can be traced, as exemplified by *Tragopogon*

species[18]. However, long-term evolution and divergence of allopolyploid lineages, such as *Cenchrus* and *Sporobolus* species, are more difficult to infer due to mutations, genomic rearrangements, and differential retention or loss of genes across subgenomes[19–21]. Disentangling their evolutionary histories is possible through karyotype reconstruction using chromosome-level genomes, which provide phylogenetic markers based on conserved collinear gene blocks and shared chromosome fusions.

Post-polyploidization genome instability triggers chromosome breakages, abnormal recombination, and extensive gene losses, providing evolutionary opportunities for plant diversification[22,23]. Chromosome number reduction typically occurs through chromosome fusions accompanied by structural rearrangements[24]. By performing comparative analyses of extant genomes, researchers have attempted to reconstruct the ancestral karyotype and elucidate evolutionary trajectories of chromosomes[25–30]. Traditional methods identify contiguous ancestral regions (CARs) by detecting collinear gene blocks. However, chromosomal rearrangements and short syntenic blocks create undefined gap regions, resulting in parameter-sensitive outputs that limit downstream analysis of chromosomal evolution[31]. To address these limitations, pairwise genomic comparisons were employed to identify shared chromosome-like syntenic blocks (CLSBs) as protochromosomes[32]. Evolutionary trajectories from ancient origins to extant species can be reconstructed by tracing three principal types of chromosomal rearrangements: nested chromosome fusion (NCF), end-to-end joining (EEJ), and reciprocal chromosome translocation (RCT)[33]. NCF occurs when one chromosome with free ends invades and integrates within the pericentromeric region of another chromosome[28]. EEJ involves a crossover event near the termini of two chromosomes, resulting in the formation of one large chromosome and a dispensable mini-chromosome[33]. Both NCF and EEJ reduce chromosome number by one. However, RCT exchanges arms between two chromosomes through interstitial crossover, producing two recombinant chromosomes without altering chromosome numbers[24]. Additionally, different types of chromosomal rearrangements generate distinct numbers of fusion breakpoints: EEJ produces one breakpoint, while RCT and NCF each produce two breakpoints[34].

Karyotype evolution research in Poaceae has historically focused on only a few crops like rice and sorghum, without a systematic cross-lineage perspective[35,36]. With the accumulation of chromosome-level genomic resources in Poaceae and advances in karyotyping methodologies, more comprehensive and accurate karyotype reconstruction will become increasingly achievable, further advancing our knowledge of grass genome evolution. In this work, we reconstruct multiple ancestral karyotypes characterized by distinct chromosome structures within the Poaceae family. We use available chromosome-level genome assemblies from extant Poaceae species and analyze their karyotype evolution by explicitly considering three types of chromosomal rearrangements (RCTs, EEJs, and NCFs), rather than merely counting the number of fission and fusion events. Additionally, we investigate the origin and evolution of allopolyploid species in two genera based on comparative karyotype analysis.

Results

Ancestral karyotype of BOP and PACMAD

We established a comprehensive genomic resource encompassing 29 species spanning 7 Poaceae subfamilies for karyotype construction (Fig. 1; **Supplementary Fig. 1**). These species exhibited diverse chromosome numbers and ploidy levels, representing the karyotypic diversity of the Poaceae family. All genome assemblies achieved high completeness (BUSCO > 95%) and continuity (LAI > 10), providing a reliable foundation for accurate karyotype construction. The only exception was *S. maritimus*, which displayed a lower LAI value (7.4) due to its scaffold-level assembly. Nevertheless, its high BUSCO score (98.6%) ensured the reliability of gene-based karyotype analysis.

We reconstructed the ancestral karyotype of BOP and PACMAD following the workflow of WGDI tool (details in Methods)[32]. Synteny comparisons between *O. sativa* and species from different subfamilies of the BOP clade revealed that each *O. sativa* chromosome exhibited CLSBs with corresponding chromosomes in BOP species (**Supplementary Fig. 2a**), supporting *O. sativa* chromosomes as protochromosomes of the BOP clade. Using the same approach, chromosomes from *P. australis* subgenome A were identified as protochromosomes of the PACMAD clade (**Supplementary Fig. 2b**). Notably, the karyotypes of *O. sativa* and *P. australis* subgenome A were highly similar. Therefore, we designated the 12 chromosomes of the *O. sativa* genome as protochromosomes of the ancestral karyotype of BOP and PACMAD, named ABPK1 to ABPK12 (**Supplementary Fig. 2c**).

Karyotype evolution in BOP clade

We systematically examined chromosome number variation and fusion events across the Poaceae family (Fig. 1). Within the BOP clade, the paleodiploid *Zizania latifolia* (*n* = 17, Oryzoideae) provides a clear instance of chromosome number reduction from a post-WGD state (*n* = 24) via seven EEJs and one RCT (**Supplementary Fig. 3**). *Raddia guianensis* and *Olyra latifolia* (both Bambusoideae) shared a nested fusion (abbreviated as ABPK_10 + 12_NCF), indicating that their divergence postdated this fusion. Notably, *R. guianensis* represented the only documented case of chromosome fission in Poaceae (purple-dashed ellipse in **Supplementary Fig. 4**).

As the most species-rich subfamily of Poaceae, Pooideae exhibits remarkable karyotypic variation. The ancestral karyotype of the Brachypodieae tribe (ABK, *n* = 10) was derived from ABPK through two NCFs (ABPK_8 + 10_NCF and ABPK_9 + 12_NCF), followed by one ABPK_(9 + 12_NCF)+11_RCT event (**Supplementary Fig. 5**). The *Brachypodium distachyon* genome (*n* = 5) was derived from the ABK genome through four NCFs (ABK_(2 + 6_NCF)+7_NCF, ABK_1 + 8_NCF, and ABK_3 + 4_NCF) and one EEJ event (ABK_5 + 10_EEJ), with the EEJ shared with *Brachypodium sylvaticum* (*n* = 9) (**Supplementary Fig. 6**). In the tetraploid *Brachypodium hybridum* genome (*n* = 15), all five *B. distachyon*-specific chromosome fusions were identified in the 5-chromosome subgenome, whereas the 10-chromosome subgenome had the *Brachypodium stacei* genome structure without any rearrangement (**Supplementary Fig. 7**). Based on the above results combined with the subgenome-level phylogenetic tree (Fig. 3), we

concluded that *B. distachyon* and *B. stacei* were most likely the donors of the two *B. hybridum* subgenomes, respectively. This was consistent with previous conclusions[37]. The ancestral karyotype of the Poeae and Triticeae tribes (APTK, *n* = 7) was derived from ABPK through descending dysploidy involving four NCFs (ABPK_3 + 11_NCF, ABPK_4 + 7_NCF, ABPK_6 + 8_NCF, and ABPK_5 + 10_NCF) and one EEJ (ABPK_9 + 12_EEJ) (**Supplementary Fig. 8a, c**). Within the Poeae tribe, *Alopecurus myosuroides* evolved from APTK through one APTK_1 + 4_RCT (**Supplementary Fig. 9**). Subsequently, the APTK genome was altered by a reciprocal translocation (APTK_1 + 6_RCT) towards the ancestral karyotype of Triticeae (ATK, *n* = 7) preceding the diversification of this tribe (**Supplementary Fig. 8b, c**). In relation to the Triticeae genome (**Supplementary Fig. 10a**), *Aegilops umbellulata* shared the ATK_4 + 5_RCT event with *Triticum monococcum* and had a reciprocal translocation (ATK_4 + 6_RCT) that arose independently from the similar translocation in *Elymus sibiricus*, as indicated by their different fusion breakpoint positions (**Supplementary Fig. 10b**).

Karyotype evolution in PACMAD clade

Within the PACMAD clade, although five subfamilies are recognized, chromosome-level genome assemblies for karyotype reconstruction are currently limited to three subfamilies: Panicoideae, Chloridoideae, and Arundinoideae (Fig. 2). Two NCF events (ABPK_3 + 10_NCF and ABPK_7 + 9_NCF) shaped the ancestral karyotype of Panicoideae subfamily (APEK, *n* = 10) from ABPK, and a subsequent NCF event (APEK_8 + 9_NCF) generated the ancestral karyotype of Paniceae tribe (APAK, *n* = 9) (**Supplementary Fig. 11**). *Miscanthus sinensis* (Andropogoneae) underwent a nested fusion between chromosomes 4 and 7 (APEK_4 + 7_NCF), reducing its chromosome number to *n* = 19 (**Supplementary Fig. 12**). In the Panicoideae tribe, *Setaria italica* and *S. viridis* retained the ancestral chromosome number (*n* = 9) altered by one APAK_3 + 7_RCT (**Supplementary Fig. 13**). Both subgenomes of the allotetraploid *C. fungigraminus* retained 7 chromosomes but differed in their evolutionary trajectories (**Supplementary Fig. 14a**). A shared APAK_3 + 6_NCF occurred prior to subgenome divergence (**Supplementary Fig. 14b**); subsequent APAK_4 + 9_RCT breakpoints differed between subgenomes, indicating independent fusion origins (**Supplementary Fig. 14c**).

The ancestral karyotype of the Chloridoideae subfamily (ACK, *n* = 10) was shaped by two NCF events (ABPK_2 + 10_NCF and ABPK_6 + 9_NCF) (**Supplementary Fig. 15**). Within the Cynodonteae tribe, the allotetraploid *Cleistogenes songorica* displayed asymmetric karyotype evolution: subgenome A retained the ancestral ACK structure without rearrangement, whereas subgenome B underwent two independent reciprocal translocations (ACK_1 + 5_RCT and ACK_3 + 7_RCT) (**Supplementary Fig. 16a**). The shared presence of ACK_3 + 10_NCF in both subgenomes of *Cynodon dactylon* and ACK_9 + 10_NCF in both subgenomes of *Eleusine coracana* indicated these fusion events occurred in diploid ancestors prior to the respective WGD events (**Supplementary Fig. 16b, c**).

Within the Zoysieae tribe, pairwise karyotype comparisons of fused chromosomes between subgenomes (A and B) of *S. maritimus* and subgenomes (C and D) of *S. alterniflorus* revealed patterns

of chromosomal rearrangement (**Supplementary Fig. 17**). All four subgenomes shared four chromosome fusion events (ACK_4 + 9_NCF, ACK_6 + 8_NCF, ACK_5 + 10_NCF, and ACK_2 + 3_RCT), indicating these rearrangements predated the divergence of these subgenomes. In addition, three subgenomes (A-C) each underwent two further chromosomal rearrangements (ACK_4 + 6_NCF and ACK_1 + 10_NCF), reducing the chromosome number to $n = 15$. By contrast, subgenome D experienced a unique ACK_4 + 10_NCF event, bringing chromosome number down to $n = 16$. This distinct karyotypic pattern in subgenome D suggested that it diverged earlier from the other subgenomes, consistent with the *K*s results (**see the last section**). Moreover, the currently available tetraploid and hexaploid Arundinoideae species retained an ABPK-like karyotype (Fig. 2).

Phylogenetic congruence of chromosome fusion distribution and uncertainty in the ancestral Poaceae karyotype

To comprehensively illustrate the karyotype evolution across the Poaceae family, we systematically identified 57 distinct chromosome fusion events (**Supplementary Table 2**) in representative species (excluding *Pharus latifolius*), and constructed a subgenome-level phylogeny, with *P. latifolius* as the outgroup (Fig. 3a). The resulting phylogeny resolved the evolutionary relationships of subgenomes across different subfamilies, which is consistent with previous studies[38]. The characterization of multiple cascading fusion events allowed us to construct the ABPK and other ancestral karyotypes, revealing different evolutionary trajectories from a common ancestor (Fig. 1). As shown in the heatmap, (i) chromosome fusions were not shared between subfamilies, providing evidence for independent karyotype evolution in each lineage, (ii) within each subfamily, species belonging to the same clade shared chromosome fusions, and the distribution of these shared fusions (marked by black rectangles) was highly congruent with the phylogenetic relationships inferred from single-copy orthologous genes (Fig. 3a). Notably, several species retained the same karyotype as ABPK without undergoing any chromosome fusion events.

Within Poaceae, Pharoideae serves as the sister lineage to the BOP and PACMAD clades. Karyotype comparisons between the *P. latifolius* genome, regarded as the ancestral karyotype of Pharoideae (APHK), and the ancestral ABPK genome revealed a reciprocal translocation (ABPK_1 + 7_RCT or APHK_1 + 7_RCT), but the evolutionary direction of this RCT event could not be determined (Fig. 3b). Earlier studies suggested that the rice genome (referred to as ABPK) was the ancestral karyotype of Poaceae[36]. However, chromosome fusion analysis revealed two possible scenarios: either the ancestral karyotype resembled APHK, with APHK_1 + 7_RCT generating the ABPK genome configuration, or ABPK represented the ancestral karyotype state, with ABPK_1 + 7_RCT producing the APHK genome structure.

To determine the directionality of the RCT event and thus infer the ancestral karyotype state, we extracted 100 genes flanking each side of the fusion breakpoints from RCT (ABPK_1 + 7_RCT or APHK_1 + 7_RCT) as individual segments and performed synteny analysis with closely related non-Poaceae outgroup species (*Ananas comosus* and *Juncus effusus*). In principle, when synteny blocks spanning

the 200-gene segment are detected in outgroup genomes, the karyotype (ABPK or APHK) possessing the 200-gene segment is more closely related to the outgroup species and is considered to represent the ancestral karyotype of Poaceae[34]. However, we could not identify any syntenic blocks near the breakpoints (**Supplementary Fig. 18**) owing to extensive chromosomal rearrangements during the Poaceae-specific ρ-WGD ($n = 7$ to $n = 12$)[39]. Hence, the true ancestral Poaceae karyotype remains uncertain.

The phylogenomic history of Cenchrus species

The phylogeny of *C. purpureus*, *C. fungigraminus*, and *C. americanus* is currently undetermined. To investigate their evolutionary relationships, we conducted pairwise karyotype comparisons of (sub)genomes of three species. Although subgenomes A and B within each allotetraploid species showed significant karyotypic divergence (**Supplementary Fig. 19**), subgenome A of both *C. purpureus* and *C. fungigraminus* exhibited highly conserved karyotypes largely consistent with genome of *C. americanus* (Fig. 4a). Similarly, subgenome B of both *C. purpureus* and *C. fungigraminus* also exhibited high karyotypic concordance.

We further performed whole-genome mapping analyses by aligning 150-mer reads derived from the diploid *C. americanus* genome against both *C. purpureus* and *C. fungigraminus* genomes. Subgenome A of *C. purpureus* and *C. fungigraminus* showed moderate coverage depths of approximately 45% and 50%, respectively (**Supplementary Fig. 20**), suggesting that *C. americanus* was likely a close relative of the diploid progenitor that contributed subgenome A to both *C. purpureus* and *C. fungigraminus*. Following the same analytical approach, we aligned 150-mer reads from subgenomes A and B of *C. purpureus* against the *C. fungigraminus* genome, achieving coverage depths of approximately 80% and 85% for subgenomes A and B of *C. fungigraminus*, respectively (Fig. 4d, e). The high level of sequence similarity between corresponding subgenomes strongly supported the hypothesis that *C. purpureus* and *C. fungigraminus* share two diploid progenitors for their respective A and B subgenomes. This interpretation is further validated by the subgenome tree (Fig. 3a).

We calculated the synonymous substitution rate (*K*s) values for homeologous gene pairs between *Cenchrus* (sub)genomes. According to the *K*s peak, we determined that the two unknown diploid progenitors (diploids A and B) of the *Cenchrus* species diverged approximately 9.5 million years ago (mya). Subsequently, diploid A diverged from *C. americanus* around 4.5 mya (Fig. 4b, c). The "bubble" peak in the transposable element (TE) divergence curve indicated that the allotetraploid speciation event occurred approximately 2.9 mya (**Supplementary Fig. 21**). Furthermore, the divergence time between *C. purpureus* and *C. fungigraminus* was estimated at 0.6 mya.

Allopolyploid evolution of Sporobolus species

By using *Eragrostis tef* subgenome A as a reference and Chr3 as an example, we were able to readily identify four homeologous blocks in *S. maritimus* and *S. alterniflorus*, respectively. Based on the

complementarity and completeness of these synteny blocks, we phased them into distinct sub-subgenomes (details in Methods; Fig. 5a). To further test relationships among sub-subgenomes, we then extracted collinear genes from these synteny blocks and constructed gene trees, rooted on *E. tef* Chr3. Phylogenies constructed from 1243 collinear genes of sub-subgenomes with at least six collinear genes resulted in 1243 topologies, from which we obtained a consensus tree (Fig. 5b). The coexisting sub-subgenomes within the genus *Sporobolus* formed two clusters (labeled S1 and S2) that fell into different clades, suggesting that the extant subgenomes (A-D) of *Sporobolus* species originated from an allotetraploid event. We further tested this hypothesis by phasing the *Sporobolus* subgenomes and constructing 10 phylogenetic trees using individual chromosomes from *E. tef* subgenome A as an outgroup, all of which consistently supported the allopolyploid origin (**Supplementary Figs. 22 and 23**).

We reconstructed karyotype schematics to trace the evolutionary history of two *Sporobolus* species (Fig. 5d). Based on synchronous *Ks* peaks at 0.22 (Fig. 5c), the diploid ACK genome is estimated to have undergone WGD1 approximately 15 mya, yielding a shared tetraploid intermediate (n = 2x = 20). Subsequent descending dysploidy via chromosome fusions restructured the intermediate into two diploidized subgenomes with $n = x = 16$ and $n = x = 15$ chromosomes, respectively. These subgenomes then underwent lineage-specific WGD2, generating *S. maritimus* ($n = 2x = 30$) at approximately 0.5 mya and *S. alterniflorus* ($n = 2x = 31$) at approximately 0.9 mya (**Supplementary Fig. 24a, b**). Unlike the ancient shared WGD1, these recent WGD2 events occurred independently of each other.

Chromosome-scale synteny analysis of selected homologous chromosomes and an ancient chromosome (marked by red triangles in Fig. 5d) revealed distinct chromosomal syntenic scenarios (Fig. 5e). Chromosomes with two distinct color blocks reflected their derivation from ancestral chromosome fusions. Notably, Chr31 of *S. alterniflorus* (Salt_31) displayed one-to-one syntenic correspondence with Chr6 of *E. tef* (Etef_6), demonstrating that Salt_31 represented an intact karyotype structure inherited from the ancestral chromosome, rather than a product of non-reciprocal translocation as previously proposed[21].

Discussion

We have reconstructed ABPK ($n = 12$) as the ancestral karyotype of the analyzed genomes of clades BOP and PACMAD. Our results show that modern genomes can be traced back to the ancestral karyotype through multiple cascading chromosomal rearrangements. Based on the order in which these chromosome fusions occur, we can determine the evolutionary trajectories of extant species. The Paniceae tribe provides a clear example: the ancestral karyotype ABPK underwent sequential chromosomal changes, first forming APEK, which then evolved into APAK, and subsequently gave rise to *S. italica* (Fig. 2**)**. This stepwise progression of fusion events reveals the evolutionary pathway more precisely than previous strategies[40,41].

By determining if genes on either side of a fusion breakpoint lie within a single synteny block, we could identify shared chromosomal rearrangements between species with significantly higher accuracy.

Sometimes, seemingly identical fusion events from the same ancestral chromosome pairs show distinct breakpoint positions, such as APAK_4 + 9_RCT in the two subgenomes of *C. fungigraminus* and ATK_4 + 6_RCT in *A. umbellulata* versus *E. sibiricus* (**Supplementary Figs. 10 and 14)**. Corresponding ancestral chromosomes as rearrangement hotspots experience multiple fusions, suggesting that certain chromosomal regions may be particularly prone to structural instability during genome evolution, similar fusions patterns have also been reported in the karyotype evolution of other plant species[34].

In this study, we identified only one chromosome fission event, consistent with a growing body of research suggesting that fusions, rather than fissions, are the primary drivers of karyotype changes[42,43]. The distribution of shared versus lineage-specific chromosome fusions provides a clear signature of karyotype evolution in Poaceae. The subfamily-specific fusions acted as key drivers of reproductive isolation, contributing to the divergence and establishment of distinct subfamilies, while shared fusions among closely related species reflect stable inheritance of karyotypic configurations (Fig. 3a**)**. Fusion events that predated divergence will be shared across descendant species, while those occurring post-divergence will be species-specific. For instance, in *O. latifolia*, ABPK_10 + 12_NCF occurred before divergence from *Raddia guianensis*, whereas ABPK_2 + 5_EEJ arose after divergence (**Supplementary Fig. 4)**.

Our karyotype reconstruction further proves its utility by providing crucial structural insights into evolutionary relationships of polyploid lineages. For Cenchrus species, pairwise karyotype comparisons between (sub)genomes, integrated with *K*s analysis and whole-genome sequence mapping, provide a high-resolution lens on their evolutionary history (Fig. 4). Our results clarify that *C. americanus* is a sister lineage to the diploid progenitor of subgenome A in the allopolyploids *C. purpureus* and *C. fungigraminus*, rather than the direct donor, consistent with previous conclusions[44]. We confirm that *C. purpureus* and *C. fungigraminus* share common diploid progenitors for their respective subgenomes A and B, resolving evolutionary relationships that remained ambiguous in prior studies[20]. We further reconstructed the karyotype of *C. fungigraminus* and characterized specific chromosome fusion events derived from the ancestral APAK genome (**Supplementary Fig. 14)**, which significantly improved the accuracy of evolutionary trajectory inference relative to previous studies that only estimated the number of likely chromosomal rearrangements occurring during karyotype evolution[20].

Similarly, for *Sporobolus* species, through manual scaffold reordering (**Supplementary Fig. 25)**, we generated a near-chromosome-level *S. maritimus* assembly and successfully partitioned the genomes of *S. maritimus* and *S. alterniflorus* into four subgenomes (A-D), each comprising two sub-subgenomes, confirming that both species are paleotetraploids rather than hexaploids as previously proposed[45](**Supplementary Fig. 22)**. The polyploid origin of *Sporobolus* species has remained contentious, recent work showed that distinguishing between autopolyploidy and allopolyploidy was not feasible due to the absence of extant diploid progenitors and tetraploid intermediates[21]. To address this challenge, we employed the *E. tef* subgenome A as a reference for partitioning homeologous chromosome blocks into their corresponding sub-subgenomes and constructed 10 phylogenetic trees

using individual chromosomes of *E. tef* subgenome A as an outgroup, thereby establishing the initial allopolyploid origin of both Sporobolus species[45] (**Supplementary Fig. 23)**. Karyotype evolution in *Sporobolus* species involves complex chromosomal rearrangements following polyploidization. Previous studies only characterized post-rediploidization karyotype changes and hypothesized the number of chromosome fusions in Sporobolus species, without considering the detailed karyotype changes from the ACK to *Sporobolus* species[21]. Here, we present a comprehensive karyotype reconstruction to trace the evolutionary trajectory of *Sporobolus* species. Contrary to the two kinds of karyotype evolution trajectories hypotheses proposed by Salmon et al.[21], our findings revealed that the ACK genome underwent a shared WGD1 event followed by lineage-specific chromosomal rearrangements that generated distinct subgenomes (A-D). Subgenomes A and B subsequently underwent WGD2 to form *S. maritimus*, with karyotype evolution involving 12 chromosome fusions (including two RCTs) that reduced the chromosome complement from theoretical $n = 40$ (octoploid) to $n = 30$ (paleotetraploid) (**Supplementary Fig. 17**). Similarly, WGD2 of subgenomes C and D produced S. alterniflorus, with 11 chromosome fusions (including two RCTs) decreasing the chromosome number from theoretical octoploid number ($n = 40$) to the current paleotetraploid complement ($n = 31$). The exact match between chromosome number reduction and chromosome fusion events robustly validates the accuracy of our karyotype constructions. Unlike previous reports[21], our results indicate that WGD2 is not shared between *S. maritimus* and *S. alterniflorus*. By using the ACK genome as a reference, combined with chromosome karyotype reconstruction, we successfully disentangled the complex evolutionary relationships between two *Sporobolus* species.

Identification of shared chromosome fusions between species provides an independent line of evidence that can corroborate and refine topology inferences, particularly when sequence-based phylogenetic methods are challenged by introgression, homoplasy, and incomplete lineage sorting over deep evolutionary timescales[46]. Nevertheless, the approach has inherent limitations that should be carefully considered. It is largely ineffective for phylogenetic inference when comparing genomes with conserved karyotypes, such as genomes from subfamily Arundinoideae[47]. Widespread gene losses in genomes can lead to the disappearance of gene collinearity, which may render the inference of chromosomal karyotypes extremely difficult. Additionally, karyotype reconstruction requires high-quality chromosome-level genome assemblies. Poor assembly quality can lead to the false-positive identification of chromosomal rearrangements and imprecise location of fusion breakpoints, resulting in misleading interpretations of evolutionary relationships between species.

To date, the fate of diploid progenitors of subgenomes A and B in *Cenchrus* remains unclear—they may have gone extinct or may persist but lack chromosome-level genomic resources. Expanded taxonomic sampling across *Cenchrus* species will be necessary to resolve the outstanding question. Additionally, the structure of the most ancestral Poaceae karyotype remains unresolved. It may eventually be resolved as more chromosome-level genomes become available, especially those from early-diverging subfamilies.

Methods

Data collection

We collected all available chromosome-level genomes in Poaceae, together with several phylogenetically related outgroup genomes from public databases[48]. Given the comprehensive characterization of Bambusoideae karyotypes in recent studies[49], only two representative species from this subfamily were selected here. Following stringent criteria for assembly contiguity and completeness, we selected 29 representative genomes spanning the major Poaceae lineages (a list of genomes see Supplementary Table 1) for karyotype reconstruction and included additional species for comparative karyotype analysis.

Phasing subgenomes with subphaser

Subphaser is an automated pipeline based on subgenome-specific *k*-mers[50]. The assembled genome and homeologous chromosome relationships were input into SubPhaser (v 1.2.6)[50] to obtain subgenome phasing results. Potential inter-subgenomic translocations and homeologous exchanges between subgenomes were also detected. For *S. maritimus*, initial subgenome phasing revealed ambiguous assignments and large-scale apparent homeologous exchanges in several regions (marked by red dashed ellipses in **Supplementary Fig. 25a, c**), likely reflecting scaffold ordering errors caused by high sequence similarity between homeologous chromosomes rather than genuine biological recombination events. Following manual scaffold repositioning (**Supplementary Table 3)**, the optimized assembly achieved correct subgenome partitioning, with chromosomes 1–15 assigned to subgenome A and chromosomes 16–30 assigned to subgenome B (**Supplementary Fig. 25b, d**). The two subgenomes of *S. alterniflorus* were successfully phased in a default subphaser run, with chromosomes 1–15 assigned to subgenome C and chromosomes 16–31 assigned to subgenome D (**Supplementary Fig. 26a, b**).

Ancestral karyotype reconstruction

To reconstruct ancestral karyotypes of BOP and PACMAD clades, we performed pairwise genome comparisons using WGDI[32]. Within the BOP clade, we selected *O. sativa* (*n* = 12) as the anchor genome because it has the highest haploid chromosome number within the clade and thus likely retains more ancestral karyotypic features than other BOP members. Three species from distinct subfamilies (Pooideae, Bambusoideae, and Oryzoideae) were then independently aligned against *O. sativa* using WGDI (with the '-d' parameter) to generate homologous dotplots. Similarly, for the PACMAD clade, we used *P. australis* subgenome A (*n* = 12) as the anchor genome and performed alignments with three species representing Chloridoideae, Panicoideae, and Arundinoideae subfamilies, respectively. Given that fusion or fission of ancestral chromosomes occur gradually and randomly in the offspring lineages, intact protochromosomes are expected to be preserved in some descendants. Based on this reasoning, we identified shared CLSBs and extracted intact chromosomes as putative protochromosomes.

Karyotypic evolution in Poaceae

To reconstruct karyotype evolution in Poaceae, we employed a comparative genomics approach that inferred chromosomal changes from ancestral karyotypes to extant genomes. The WGDI workflow began with the identification of collinear gene pairs between species using the improved collinearity module ('-icl') followed by synteny block integration ('-bi'). Stringent filtering was then applied through the correspondence module ('-c') to eliminate syntenic regions potentially derived from earlier polyploidization events. Protochromosomes were mapped onto the chromosomes of the sample genomes via the karyotype mapping function ('-km'). Furthermore, the parameter '-sf' was employed to rapidly identify chromosome fusions, with shared fusions and their corresponding fusion breakpoints subsequently recorded. Fusion breakpoint coordinates were extracted using the '-fpd' parameter, which generated a comprehensive dataset recording fusion positions. Finally, we applied the '-fd' module to determine the occurrence of these breakpoints in other genomes.

Phylogenomic analysis

We reconstructed phylogenetic relationships using 43 subgenomes from 29 Poaceae species. Orthogroups were identified using OrthoFinder (v 2.5.4)[51] with subgenome protein sequences. We retained high-confidence orthogroups present in at least 35 subgenomes (> 80% taxonomic sampling), with a single gene copy per subgenome. Protein sequences within each orthogroup were aligned using MUSCLE (v 5.1)[52], and the resulting multiple sequence alignments were concatenated into a supermatrix. Maximum-likelihood phylogenetic inference was performed using IQ-TREE (v 2.1.3)[53] under the best-fit substitution model selected by ModelFinder, with branch support assessed by 1,000 ultrafast bootstrap replicates.

Divergence and merger time estimation

We estimated the divergence time between subgenomes and subsequent merger time for allopolyploid *Sporobolus* species (*S. alterniflorus* and *S. maritimus*) and *Cenchrus* species (*C. americanus*, *C. fungigraminus*, and *C. purpureus*). We first aligned protein sequences among subgenomes using BLASTP (v 2.14.0)[54] with an E-value cutoff of $1e^{-10}$. DAGchainer (r02-06-2008)[55] was used to identify syntenic blocks with at least five homologous gene pairs within a region of 10 gene models based on best-hit blastp results. The *K*s values for syntenic homeologous gene pairs were calculated by *Ka*/*K*s_Calculator in NG mode. The *K*s distribution peaks were used to calculate the divergence time using the formula $T = Ks/2r$, where r is the nucleotide substitution rate ($r = 6.5 \times 10^{-9}$ mutations × bp^{-1} × $generation^{-1}$). TE divergence was assessed using PercDivs (percentage of substitutions in the matching region compared with the consensus), calculated by RepeatMasker (v 4.2.0) (https://github.com/rmhubley/repeatmasker). The TE sequence divergence between two subgenomes displaying a high degree of overlap suggested consistency in the TE evolutionary rate in the two subgenomes. The non-overlapping segregation region indicated the timeframe from diploid progenitor divergence to tetraploid genome formation.

K -mer-based whole-genome mapping analysis

To investigate evolutionary relationships between *Cenchrus* species, *k*-mer-based whole-genome mapping analysis was performed as previously described[12]. Query (sub)genome sequences were fragmented into 150-mers with a 5 bp step size, and these short reads were subsequently mapped to the reference genome using Bowtie2 (v 2.5.1)[56]. Coverage depth was calculated using SAMtools (v 1.23)[57] to quantify sequence similarity.

Phasing sub-subgenomes with WGDI

Subphaser toolkit[50] independently partitioned *S. alterniflorus* and *S. maritimus* into two subgenomes, each comprising two chromosome sets (termed sub-subgenomes). To resolve the evolutionary relationships among sub-subgenomes within the *Sporobolus* species, we used *E. tef* subgenome A as the reference for phasing of *S. alterniflorus* and *S. maritimus*. First, protein sequences from *Sporobolus* species were compared against those of the reference genome using DIAMOND (v 2.0.15)[58]. Collinear regions were identified and integrated using WGDI with the '-icl' and '-bi' parameters. Orthologous syntenic blocks derived from recent polyploidy events were retained using WGDI with the '-c' filtering option. The reference genome karyotype was then mapped onto *Sporobolus* chromosomes based on syntenic relationships using WGDI with the '-km' parameter. Sub-subgenomes were manually assigned based on complementarity and completeness of synteny blocks. Hierarchical gene lists (one sub-subgenome per column) were generated using WGDI with the '-pc' and '-a' parameters. For each chromosome of the reference genome, maximum likelihood gene trees were inferred from the merged hierarchical gene lists using built-in IQ-TREE with automatic selection of the best-fit substitution model (-m MFP) through the WGDI '-at' function. These gene trees were input to ASTRAL-Pro (v 1.19.3.6)[59] to construct the sub-subgenome phylogeny. The resulting phylogeny was visualized and evaluated using Newick utilities (v 1.6)[60] and Phytop (v 0.3)[61]. To ensure robust sub-subgenome classification, the analysis pipeline ('-pc', '-a', and '-at' options in WGDI, followed by ASTRAL-Pro) was iterated until topological consistency converged.

Chromosome synteny visualization

Chromosome-level synteny was analyzed across ten chromosomes from *E. tef* (Chr 6), *S. alterniflorus* (Chr 31, 3, 18, 2, and 17), and *S. maritimus* (Chr 3, 18, 2, and 17). Pairwise syntenic blocks between chromosomes were detected using JCVI (v 1.5.7)[62], requiring ≥ 30 gene pairs per block with ≤ 20 intervening genes. Coordinates of anchor genes were extracted from BED files and converted to NGenomeSyn link format with strand directionality preserved. The resulting syntenic links were visualized using NGenomeSyn (v 1.43)[63] with customized color schemes for different collinearity relationships.

Declarations

Author countribution

L.F., P.S., and D.W. led and coordinated the project. H.C. and L.J. collected the assemblies. X.H. and K.Y. performed quality control. H.C., S.Z., S.Z., R.Z., and Z.L. analyzed Poaceae genomes. H.C., X.H., and Y.Z. finished karyotype and phylogenetic analyses. H.C., P.S., D.W., Q.C., and Y.H. provided gramine quantification. L.F., P.S., D.W., H.C., and Y.H. discussed the results. L.F. and P.S. supervised all analyses. H.C. wrote the initial manuscript with input from all co-authors. All authors discussed the results and commented on the manuscript.

Competing interests

All authors declare no competing interests.

Acknowledgements

This work was supported by National Key Research and Development Program (2023YFD1400502).

References

1. Kellogg EA (2000) The grasses: A case study in macroevolution. Annu Rev Ecol Evol Syst 31:217–238
2. Kellogg EA (2001) Evolutionary history of the Grasses1. Plant Physiol 125:1198–1205
3. Soreng RJ et al (2015) A worldwide phylogenetic classification of the poaceae (gramineae). J Syst Evol 53:117–137
4. Soreng RJ et al (2017) A worldwide phylogenetic classification of the poaceae (gramineae) II: An update and a comparison of two 2015 classifications. J Syst Evol 55:259–290
5. Clark LG, Judziewicz EJ (1996) The grass subfamilies anomochlooideae and pharoideae (poaceae). Taxon 45:641–645
6. Clark LG, Kobayashi M, Mathews S, Spangler RE, Kellogg EA (2000) The puelioideae, a new subfamily of poaceae. Syst Bot 25:181–187
7. Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci.* 101, 9903–9908
8. Tang H, Bowers JE, Wang X, Paterson AH (2010) Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proc. Natl. Acad. Sci.* 107, 472–477
9. Ming R et al (2015) The pineapple genome and the evolution of CAM photosynthesis. Nat Genet 47:1435–1442
10. McKain MR et al (2016) A phylogenomic assessment of ancient polyploidy and genome evolution across the poales. Genome Biol Evol 8:1150–1164
11. Qiao X, Zhang S, Paterson AH (2022) Pervasive genome duplications across the plant tree of life and their links to major evolutionary innovations and transitions. Comput Struct Biotechnol J

20:3248–3256

12. Ye C-Y et al (2020) The Genomes of the Allohexaploid *Echinochloa crus-galli* and Its Progenitors Provide Insights into Polyploidization-Driven Adaptation. Mol Plant 13:1298–1310
13. Van de Peer Y, Maere S, Meyer A (2009) The evolutionary significance of ancient genome duplications. Nat Rev Genet 10:725–732
14. Madlung A (2013) Polyploidy and its effect on evolutionary success: Old questions revisited with new tools. Heredity 110:99–104
15. Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA (2016) On the relative abundance of autopolyploids and allopolyploids. New Phytol 210:391–398
16. Chen ZJ (2010) Molecular mechanisms of polyploidy and hybrid vigor. Trends Plant Sci 15:57–71
17. Wang J et al (2019) Polyploidy index and its implications for the evolution of polyploids. Front Genet 10
18. Soltis DE et al (2004) Recent and recurrent polyploidy in tragopogon (asteraceae): Cytogenetic, genomic and genetic comparisons. Biol J Linn Soc 82:485–501
19. Wolfe KH (2001) Yesterday's polyploids and the mystery of diploidization. Nat Rev Genet 2:333–341
20. Zheng H et al (2023) A near-complete genome assembly of the allotetrapolyploid *Cenchrus fungigraminus* (JUJUNCAO) provides insights into its evolution and C4 photosynthesis. Plant Commun 4:100633
21. Salmon A et al (1997) On the way to diploidization and unexpected ploidy in the grass Sporobolus section Spartina mesopolyploids. *Nat. Commun.* 16, (2025)
22. Leitch AR, Leitch IJ (2008) Genomic plasticity and the diversity of polyploid plants. Science 320:481–483
23. Wendel JF (2015) The wondrous cycles of polyploidy in plants. Am J Bot 102:1753–1756
24. Schubert I, Lysak MA (2011) Interpretation of karyotype evolution should consider chromosome structural constraints. Trends Genet 27:207–216
25. Salse J et al (2009) Reconstruction of monocotelydoneous proto-chromosomes reveals faster evolution in plants than in animals. *Proc. Natl. Acad. Sci.* 106, 14908–14913
26. Salse J, Abrouk M, Murat F, Quraishi UM, Feuillet C (2009) Improved criteria and comparative genomics tool provide new insights into grass paleogenomics. Briefings Bioinf 10:619–630
27. Murat F, de Peer YV, Salse J (2012) Decoding plant and animal genome plasticity from differential paleo-evolutionary patterns and processes. Genome Biol Evol 4:917–928
28. Wang X et al (2015) Telomere-centric genome repatterning determines recurring chromosome number reductions during the evolution of eukaryotes. New Phytol 205:378–389
29. Wang X et al (2015) Genome alignment spanning major poaceae lineages reveals heterogeneous evolutionary rates and alters inferred dates for key evolutionary events. Mol Plant 8:885–898
30. Wang M, Liu C, Xing T, Wang Y, Xia G (2015) Asymmetric somatic hybridization induces point mutations and indels in wheat. BMC Genomics 16:807

31. Pont C et al (2019) Paleogenomics: Reconstruction of plant evolutionary trajectories from modern and ancient DNA. Genome Biol 20:29
32. Sun P et al (2022) A user-friendly toolkit for evolutionary analyses of whole-genome duplications and ancestral karyotypes. Mol Plant 15:1841–1851
33. Lysak MA (2022) Celebrating mendel, McClintock, and darlington: on end-to-end chromosome fusions and nested chromosome fusions. Plant Cell 34:2475–2491
34. Jiang X et al (2025) Chromosome fusions shaped karyotype evolution and evolutionary relationships in the model family brassicaceae. Nat Commun 16:4631
35. Murat F et al (2010) Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. Genome Res 20:1545–1557
36. Murat F, Armero A, Pont C, Klopp C, Salse J (2017) Reconstructing the genome of the most recent common ancestor of flowering plants. Nat Genet 49:490–496
37. Gordon SP et al (2020) Gradual polyploid genome evolution revealed by pan-genomic analysis of *Brachypodium hybridum* and its diploid progenitors. Nat Commun 11:3670
38. Huang W et al (2022) A well-supported nuclear phylogeny of poaceae and implications for the evolution of C_4 photosynthesis. Mol Plant 15:755–777
39. Bellec A et al (2023) Tracing 100 million years of grass genome evolutionary plasticity. Plant J 114:1243–1266
40. Mandáková T, Lysak MA (2008) Chromosomal phylogeny and karyotype evolution in x = 7 crucifer species (brassicaceae). Plant Cell 20:2559–2570
41. Lysak MA, Mandáková T, Schranz ME (2016) Comparative paleogenomics of crucifers: Ancestral genomic blocks revisited. Curr Opin Plant Biol 30:108–115
42. Yin Y et al (2021) Molecular mechanisms and topological consequences of drastic chromosomal rearrangements of muntjac deer. Nat Commun 12:6858
43. Shao Y et al (2023) Phylogenomic analyses provide insights into primate evolution. Science 380:913–924
44. Zhang S et al (2022) Chromosome-scale genome assembly provides insights into speciation of allotetraploid and massive biomass accumulation of elephant grass (*pennisetum purpureum* schum). Mol Ecol Resour 22:2363–2378
45. Ainouche ML et al (2009) Hybridization, polyploidy and invasion: Lessons from spartina (poaceae). Biol Invasions 11:1159–1173
46. Whitfield JB, Lockhart PJ (2007) Deciphering ancient rapid radiations. Trends Ecol Evol 22:258–265
47. Jeunen G-J et al (2023) crabs—a software program to generate curated reference databases for metabarcoding sequencing data. Mol Ecol Resour 23:725–738
48. Xie L et al (2024) Technology-enabled great leap in deciphering plant genomes. Nat Plants 10:551–566

49. Ma P-F et al (2024) Genome assemblies of 11 bamboo species highlight diversification induced by dynamic subgenome dominance. Nat Genet 56:710–720
50. Jia K-H et al (2022) SubPhaser: A robust allopolyploid subgenome phasing method based on subgenome-specific k-mers. New Phytol 235:801–809
51. Emms DM, Kelly S, OrthoFinder (2019) Phylogenetic orthology inference for comparative genomics. Genome Biol 20:238
52. Edgar RC (2022) Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun 13:6968
53. Minh BQ et al (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534
54. Camacho C et al (2009) BLAST+: Architecture and applications. BMC Bioinf. 10
55. Bhasi K, Forrest A, Ramanathan MSPLINDID (2005) A semi-parametric, model-based method for obtaining transcription rates and gene regulation parameters from genomic and proteomic expression profiles. Bioinformatics 21:3873–3879
56. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. Nat Methods 9:357–359
57. Li H et al (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079
58. Buchfink B, Reuter K, Drost H-G (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18:366–368
59. Zhang C, Mirarab S (2022) ASTRAL-pro 2: Ultrafast species tree reconstruction from multi-copy gene family trees. Bioinform (Oxf Engl) 38:4949–4950
60. Junier T, Zdobnov EM (2010) The newick utilities: High-throughput phylogenetic tree processing in the unix shell. Bioinformatics 26:1669–1670
61. Shang H-Y et al (2025) Phytop: A tool for visualizing and recognizing signals of incomplete lineage sorting and hybridization using species trees output from ASTRAL. Hortic Res 12:uhae330
62. Tang H et al (2011) Screening synteny blocks in pairwise genome comparisons through integer programming. BMC Bioinf. 12
63. Gentili M, Martini L, Sponziello M, Becchetti L (2022) Biological random walks: Multi-omics integration for disease gene prioritization. Bioinformatics 38:4145–4152

Figures

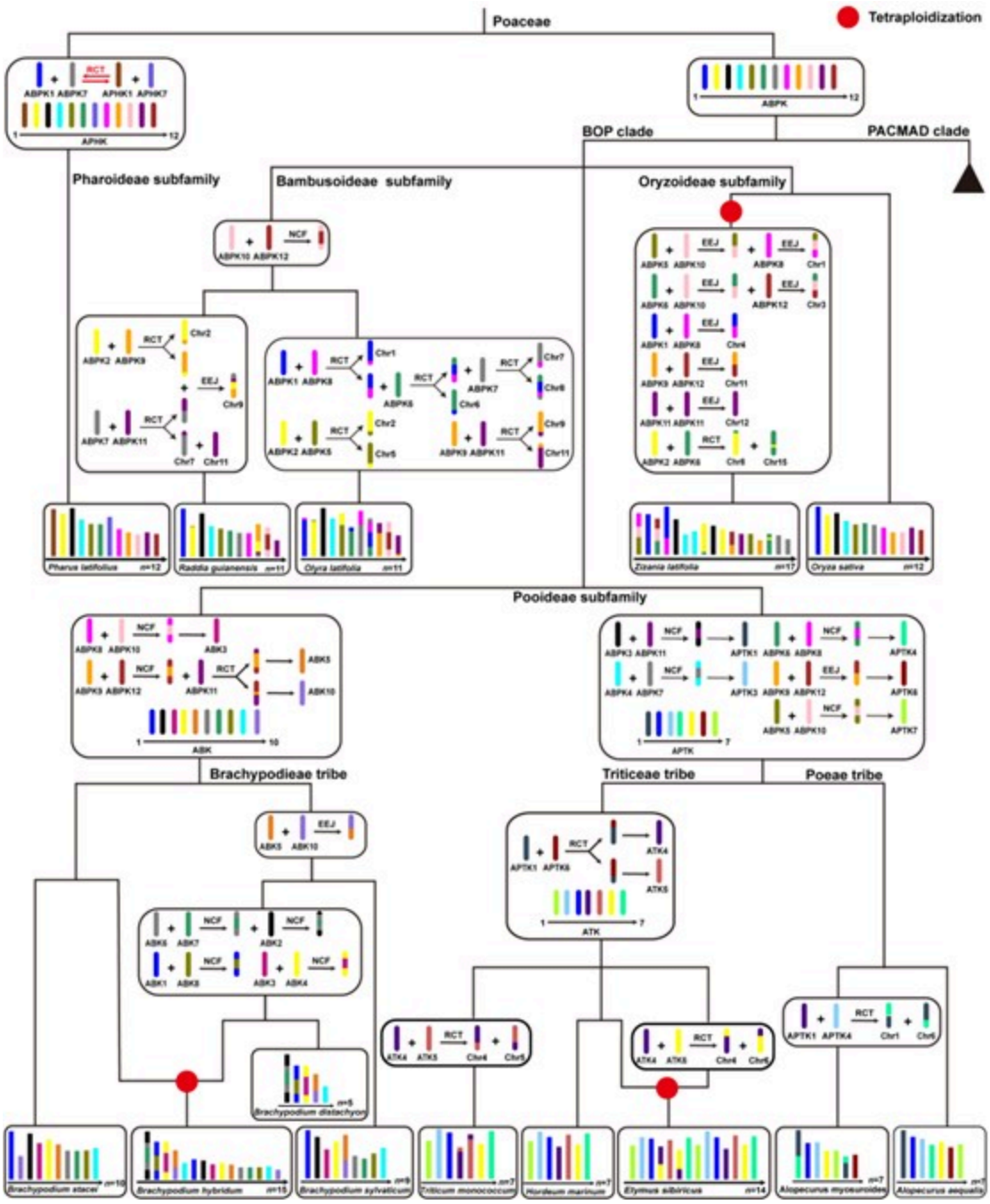

Figure 1

Karyotype reconstruction and evolution for a portion of the Poaceae family (including subfamily Pooideae, Oryzoideae, Bambusoideae from BOP clade and subfamily Pharoideae).

Ideograms are based on the published chromosome-level genome assemblies (**Supplementary Table 1**) and their karyotype reconstruction using the WGDI. ABPK: Ancestral karyotype of BOP and PACMAD clades, APHK: Ancestral Karyotype of Pharoideae subfamily, ABK: Ancestral Karyotype of Brachypodieae tribe, APTK: Ancestral Karyotype of Poeae and Triticeae tribes, ATK: Ancestral Karyotype of Triticeae tribe, RCT: reciprocal chromosome translocation, EEJ: end-to-end joining, NCF: Nested chromosome fusion. Arrows illustrate the evolutionary directions of the fusion events, while bidirectional arrows

indicate uncertainty about the direction of evolution. Tetraploidization ismarked by a red circle. Triangle represents the PACMAD clade, and the reconstructed karyotype evolution is shown in Fig. 2.

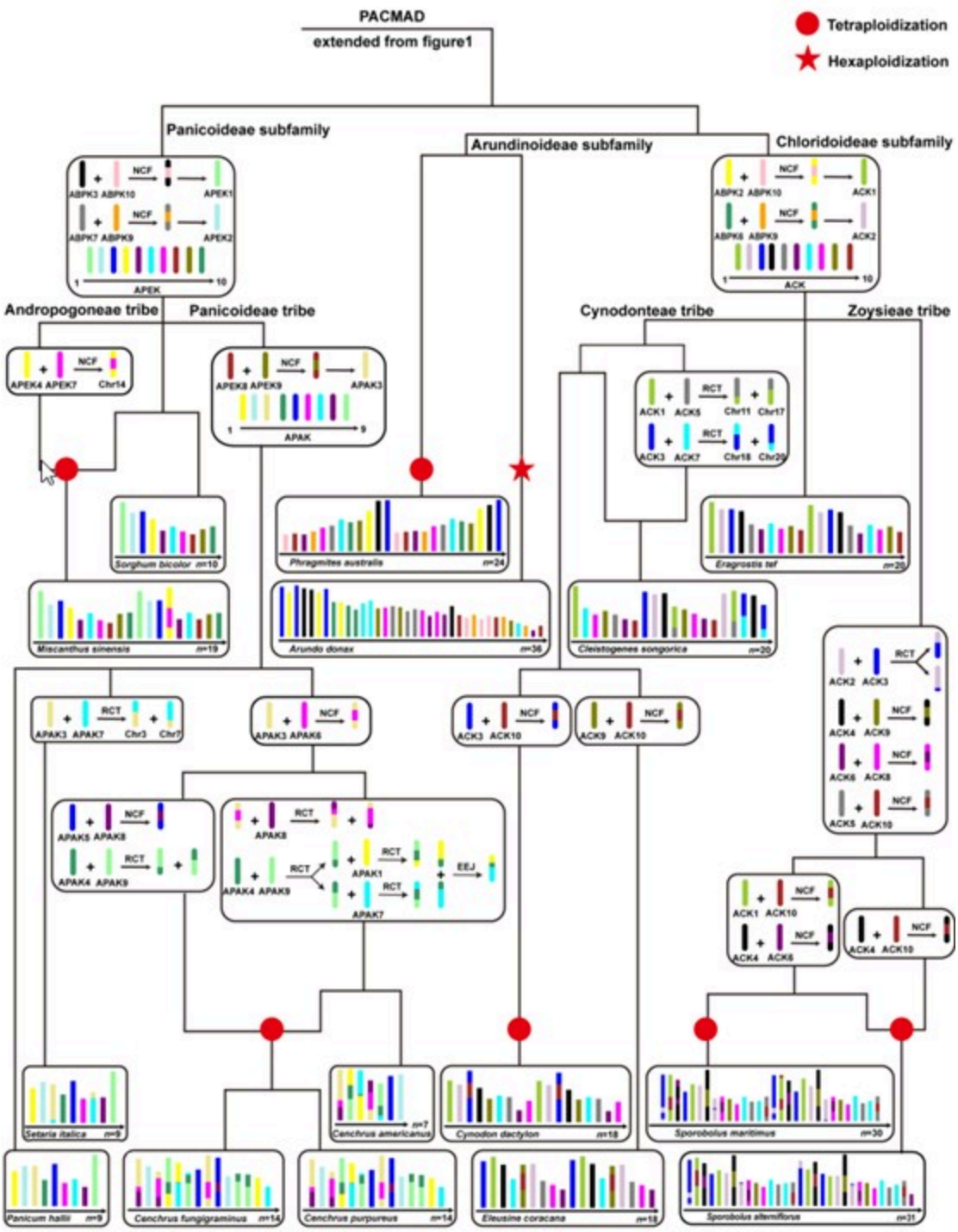

Figure 2

Reconstructed karyotype evolution for a portion of the Poaceae family (including subfamily Arundinoideae, Chloridoideae and Panicoideae from PACMAD clade).

APEK: Ancestral Karyotype of Panicoideae subfamily, ACK: Ancestral Karyotype of Chloridoideae subfamily, APAK: Ancestral Karyotype of Panicoideae tribe. RCT: reciprocal chromosome translocation,

EEJ: end-to-end joining, NCF: Nested chromosome fusion. Tetraploidization ismarked by a red circle; hexaploidization is marked by a red pentagram. See also legend for Fig. 1.

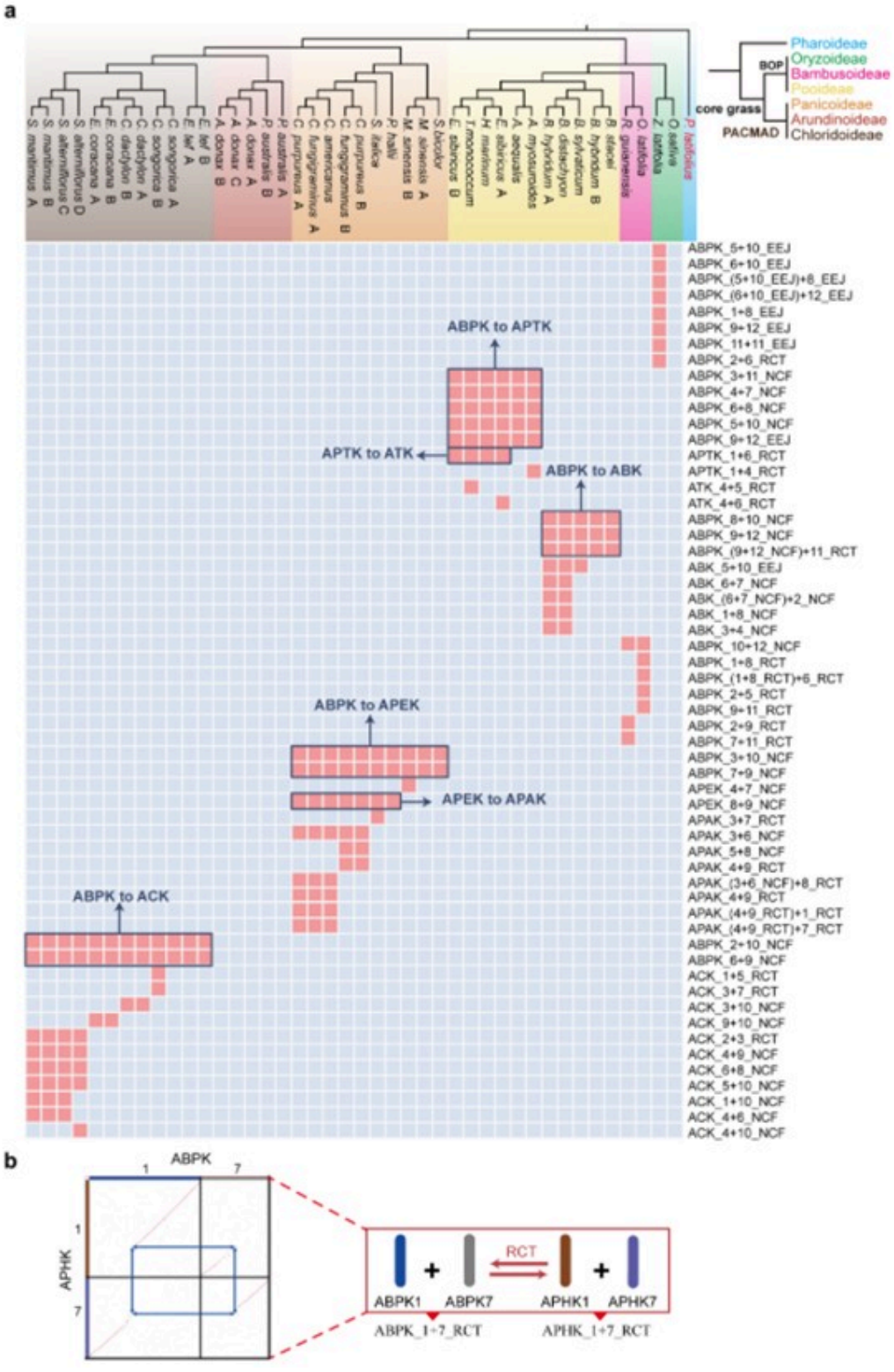

Figure 3

Chromosome fusions in Poaceae during their karyotype evolution.

(a) Subgenome-level phylogenetic tree with *P. latifolius* as an outgroup, accompanied by a heatmap illustrating the distribution of chromosome fusions across subgenomes from BOP and PACMAD clades. (b) The ancestral ABPK (identical to *O. sativa*) genome and APHK (identical to *P. latifolius*) genome are distinguished by a reciprocal translocation (ABPK_1+7_RCT or APHK_1+7_RCT). Bidirectional arrows indicate uncertain fusion direction. ABPK: Ancestral karyotype of BOP and PACMAD clades, APHK: Ancestral Karyotype of Pharoideae subfamily, ABK: Ancestral Karyotype of Brachypodieae tribe, APTK: Ancestral Karyotype of Poeae and Triticeae tribes, ATK: Ancestral Karyotype of Triticeae tribe, APEK: Ancestral Karyotype of Panicoideae subfamily, ACK: Ancestral Karyotype of Chloridoideae subfamily, APAK: Ancestral Karyotype of Paniceae tribe. RCT: reciprocal chromosome translocation, EEJ: end-to-end joining, NCF: Nested chromosome fusion.

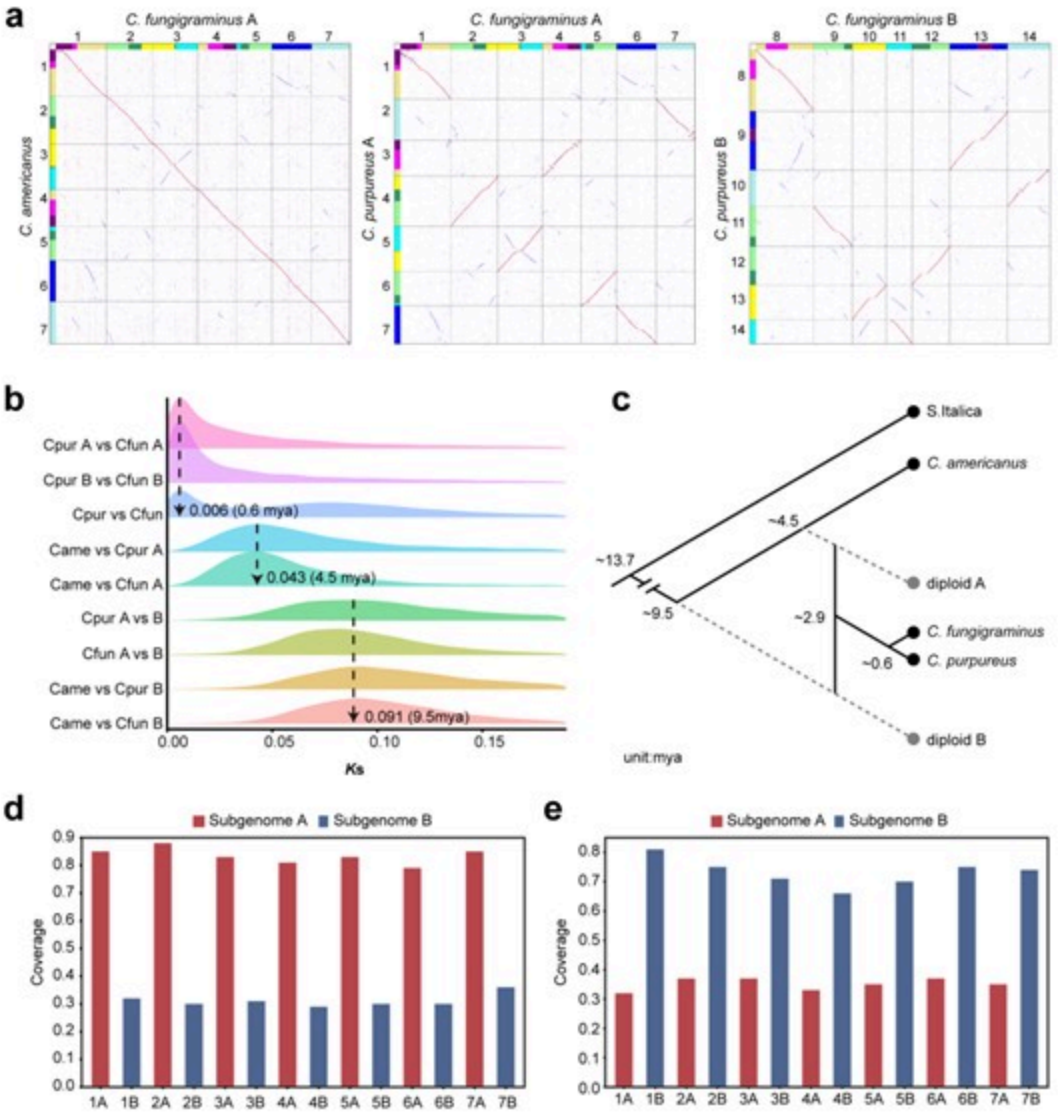

Figure 4

Evolutionary processes of the genomes in three *Cenchrus* species.

(a) Pairwise karyotype comparisons among the (sub)genomes of *Cenchrus* species. (b) The synonymous substitution rate (*K*s) distribution of pairs of homeologous genes among the (sub)genomes from *C. purpureus* (Cpur), *C. fungigraminus* (Cfun), and *C. americanus* (Came). (c) Brief diagram showing the evolution of *Cenchrus* species and estimation of polyploidization times. The grey dashed line represents two unknown ancestral diploid progenitors. Numbers beside each branch point show divergence or hybridization time. mya: million years ago. Coverage depth of two subgenomes in *C. fungigraminus* by mapping 150-mer sequence from *C. purpureus*subgenome A (d) or subgenome B (e).

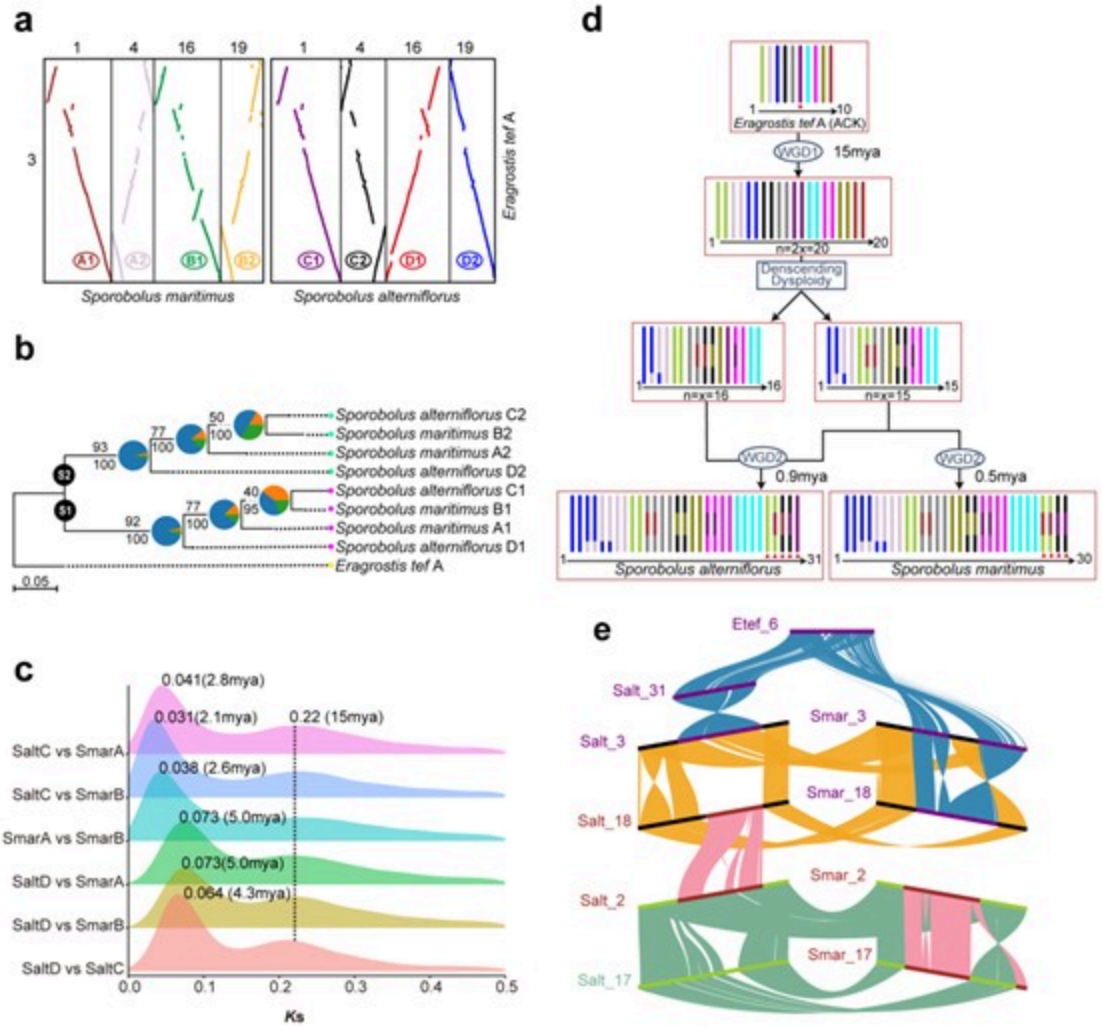

Figure 5

Evolutionary relationships of the genomes between two *Sporobolus* species.

(a) Identification of chromosome-like synteny blocks (CLSBs) using *Eragrostis tef* Chr3 as the reference against *S. maritimus* and *S. alterniflorus*. Letters within each oval indicate the sub-subgenome origin of the corresponding chromosome. (b) A phylogeny of the coexisting sub-subgenomes based on 1243 collinear genes from homologous blocks to *E. tef* Chr3. Numbers above branches indicate concordance percentages between gene-based and chromosome-scale phylogenies; numbers below branches represent local posterior probabilities calculated by ASTRAL. Pie charts at nodes display the frequencies of the three primary gene tree topologies (q1, q2, and q3) as quantified by ASTRAL, reflecting

discordance among gene trees. Branch lengths are proportional to substitutions per site (scale bar = 0.05). Species labels are color-coded to distinguish different accessions: *S. maritimus* (A1, A2, B1, B2) and*S. alterniflorus* (C1, C2, D1, D2), with *E. tef* subgenome A as the outgroup. (c) The synonymous substitution rate (*K*s) distribution of pairs of homeologous genes between subgenomes of *S. maritimus* and *S. alterniflorus*. (d) Karyotype reconstruction depicting the evolutionary trajectory from the Chloridoideae ancestor to extant *Sporobolus* species. Colored bars represent individual chromosomes, with each color denoting distinct ancestral chromosome segments. The red triangle highlights the chromosomes selected for downstream synteny analysis. (e) Synteny visualization of 10 selected chromosomes (red triangle in d) from *E. tef* (Etef), *S. maritimus* (Smar), and *S. alterniflorus* (Salt).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.xlsx
- SupplementaryFigures.docx