Complexity and Innovation in Carnivorous Plant Genomes

*Albert VA[1], #Avilá Robledillo L[2], #Fleck SJ[1], Guo Y[3,4], Kanamori S[3], Kirshner J[1], *Marques A[2,5], *Fukushima K[3,4]

*Correspondence: carnivorous.plant.genomes@gmail.com (V. A. Albert), amarques@mpipz.mpg.de (A. Marques), kenji.fukushima@nig.ac.jp (K. Fukushima)

#Equal contributions

Affiliations:

1. Department of Biological Sciences, University at Buffalo, Buffalo, New York, USA

2. Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Cologne, Germany

3. National Institute of Genetics, Mishima, Shizuoka, Japan

4. Graduate Institute for Advanced Studies, SOKENDAI, Mishima, Shizuoka, Japan

5. Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine University, Düsseldorf, Germany

Keywords:
carnivorous plants, genomes, polyploidy, gene content, adaptation, holocentricity

Highlights

- Complex polyploidy reveals convoluted hybrid origins: chromosome-scale genome assemblies for Caryophyllales carnivores uncover decaploidy in *Nepenthes* pitcher plants, tetraploidy in Venus flytrap, and hexaploidy/dodecaploidy in sundews.
- Genome compaction defies expectation: the bladderwort *Utricularia gibba* achieves extreme genome size reduction via massive intergenic DNA deletion despite multiple polyploidizations, retaining a typical angiosperm gene complement.
- Centromere architecture dictates genome dynamics: contrasting monocentric and holocentric organizations in *Drosera* species fundamentally shape chromatin structure, satellite repeat evolution, and karyotype plasticity.
- Multilevel molecular convergence underpins carnivory: digestive enzyme recruitment from defense pathways, lineage-specific gene family expansions (e.g., cysteine proteases), and recurrent amino acid substitutions demonstrate pervasive convergent evolution.

Abstract

Carnivorous plants are a paradigm of convergent evolution, but their genomes reveal even deeper layers of complexity. Recent work uncovers widespread polyploidy, including the decaploid East Asian pitcher plant (*Nepenthes gracilis*) genome and hybrid origins for the tetraploid Venus flytrap (*Dionaea muscipula*) and queen (hexaploid) and Cape (dodecaploid) sundews (*Drosera regia and D. capensis*, respectively). The bladderwort (*Utricularia gibba*) experienced extreme genome compaction while retaining otherwise typical gene number, challenging assumptions about genome size. Molecular convergence is conspicuous, from digestive enzyme recruitment to repeated amino acid substitutions under functional constraints. *Drosera* species further illustrate how centromere type (monocentric versus holocentric) shapes genome architecture. These discoveries position carnivorous plants as models for studying the plasticity and adaptive landscapes of plant genomes, including tradeoffs between local and global gene duplication and intergenic DNA deletion.

Genomic Perspectives on the Evolution of Plant Carnivory

Carnivorous plants are a paradox within botanical diversity in terms of their morphologies and ecophysiological adaptations for trapping and digesting prey animals. Equally extraordinarily, specialized trapping formats have evolved multiple times across the flowering plant family tree[1]. These trap types include sticky-fingered, flypaper-like leaves that have arisen through **convergent evolution** at least four times, liquid-containing pitchers that have three independent origins, tiny suction bladders, substrate-penetrating corkscrew traps, and nematode-trapping subterranean leaves[1-3]. Pitcher plants, for example, which bear tubular, pitfall-trapping leaves, have evolved in three distantly related plant families separated by tens of millions of years of evolutionary history[1-3]. In contrast, two distinct lineages that each include closely interrelated carnivores display highly divergent trap forms within them. For example, the East Asian pitcher plant (*Nepenthes* L.) is closely related to the Venus flytrap (*Dionaea* J. Ellis) and flypaper-trapping sundews (*Drosera* L.; Fig. 1). However, another sticky-fingered trapper, the terrestrial butterwort (*Pinguicula* L.), is closely related to the often-aquatic bladderwort plant (*Utricularia* L.)[2,3]. Moreover, a second snap-trapping species, aquatic *Aldrovanda vesiculosa* L., is closely related to *Nepenthes*, *Dionaea* and *Drosera* (Fig. 1), and the morphologically divergent, lobster-pot-trapping corkscrew plant *Genlisea* A.St.-Hil. is *Utricularia*'s immediate sister lineage[2,3]. Perhaps not surprisingly, the genomes and gene-content trajectories of many of these plants are unusually complex as well, showing perplexing histories of genome multiplication events, interwoven ancestral lineages, and distinct genetic adaptations for their unusual nutrient-acquiring lifestyle.

While angiosperms are well known for showing signatures of past **polyploidy** events and lineage admixtures during such events[4], few juxtapose these with the evolution of such bizarre morphological complexity along both similar and dissimilar themes. Moreover, among genome adaptive landscapes characterized across angiosperms, genetic specializations for the carnivorous lifestyle are unusually prominent[5-7]. In some cases, gene families encoding digestive enzymes have themselves been convergently recruited from ancient plant defense pathways[8,9] that are otherwise ubiquitous in plants. Further highlighting the astonishing molecular specializations of carnivores, some of these enzymes are themselves under convergent evolutionary pressures to incorporate specific amino acids at certain protein positions[5]. Altogether, the evolution of the carnivorous plant strategy can be likened to a Matryoshka doll, with convergence at the physiological level encasing morphological innovations, which further encloses genomic complexity that is itself wrapped around co-option of enzymatic diversity and protein sequence evolution within it. Here, we review recent research that has uncovered some of the more interesting genomic oddities and specializations among carnivorous plants.

Genome Structural Evolution

Polyploidy and Lineage Admixtures. **Whole genome multiplications** (WGMs; see **Glossary**), wherein entire chromosome complements are multiplied (most often duplicated), can take place within species or following hybridization events between species. The former is termed **autopolyploidy**, the latter **allopolyploidy**. Allopolyploidies, at least in a sense broad enough to constitute two different lineages fusing and needing to rebalance meiotic pairing, are likely the most common[10]. Despite the observed prevalence of polyploidy among plant lineages, the long-term survival of polyploids is probably rare, as immediate impacts of WGM will usually be deleterious[11]. Thus, while polyploidy appears to be superficially common across

angiosperm lineages, these cases likely represent the rare survivors of otherwise disadvantageous genomic upheavals.

In the case of carnivorous plants, all genomes sequenced except one, the Australian pitcher plant, *Cephalotus follicularis* Labill. (which occupies a phylogenetic outpost in the starfruit lineage, Oxalidales Bercht. & J.Presl[1,5]) have undergone additional WGMs following the ancient genome hexaploidization that occurred in the stem lineage of all core eudicots (approximately 75% of angiosperm species)[12,13]. This genomic architecture was first discovered in grapevine (*Vitis vinifera* L.)[14], and is now known to be found across a broad diversity of taxa that have not undergone any WGMs since. In contrast to *Cephalotus*, the other pitcher plants, *Nepenthes* and *Sarracenia* Tourn. *ex* L. (the American pitcher plant), have both undergone post-hexaploid WGMs, with *Nepenthes* presenting an especially confusing case[7] (see below). While no well-assembled genome sequence of *Sarracenia* has been published at the time of this writing, its phylogenetic position in the Ericales Bercht. & J.Presl (a large asterid order containing 22 plant families), some preliminary transcriptomic research[15] and genome sequence data (J. Kirshner, C. Page, N. Pratt and V. A. Albert, unpublished) paint a case for (allo)tetrapolyploidy. In fact, the entire order may have an allopolyploid hybrid origin[16].

The Confusing Genomes of Caryophyllales Carnivores. Within the tightly related carnivore clade found in the spinach order (**Caryophyllales** Juss. ex Bercht. & J.Presl; Fig. 1), the *Nepenthes gracilis* Korth. genome, an extreme allopolyploid outlier, was the first assembled to the chromosomal level[7]. Revealed to have five haploid subgenomes of eight chromosomes each (Fig. 2), *N. gracilis* is a decaploid in comparison with *Cephalotus*, which only possesses the core eudicot paleohexaploid event. Detailed **synteny** and post-polyploid gene retention analyses (see below) revealed a likely allopolyploidization scenario wherein one ancestral lineage contributed four x=8 subgenomes from its own two WGMs to one *Nepenthes* ancestor, whereafter a second, distinct lineage contributed a fifth 8-chromosome subgenome to yield a hybrid clade. Neither the 4-subgenome ancestor nor the contributor of the fifth subgenome appear to have left modern descendants.

Altogether, compared to the relict species *Amborella trichopoda* Baill., which represents one branch of the first evolutionary split among flowering plants and shows no evidence for WGMs following the flowering-plant-wide polyploid event[17], the haploid genome of *Nepenthes* represents a 15-fold multiplication: 1× (*Amborella*) times 3× (core eudicot hexaploidy) times 5 subgenomes. All other Caryophyllales carnivores share *Nepenthes*'s x=8 ancestral subgenome, or a modification of it (Fig. 2). *Dionaea* is an n=16 allotetraploid hybrid (Fig. 1) of two different 8-chromosome ancestors (Fig. 2)[18]. *Drosera regia* Stephens appears to share the two subgenomes of that species, along with a third joining in to yield another hybrid lineage, an allohexaploid one (Fig. 2) that precludes phylogenetic depiction as deriving from a single phylogenetic split (Fig. 1)[18].

Indeed, this hybrid origin of *D. regia* likely involves other *Drosera* species in addition to *Dionaea*, as is apparent from extreme conflict among independent phylogenies based on a large collection of single-copy genes (Fig. 1). Originating from three x=8 ancestral subgenomes, the immediate progenitor of *D. regia* was n=24 prior to chromosomal rearrangements that led to its modern n=17 karyotype (Fig. 2). The *Drosera capensis* L. genome also contains an underlying triplicate structure (a hexaploidy likely shared with *D. regia*), and it has undergone a recent, lineage-specific WGM to yield a dodecaploid (Fig. 2)[18].

The remarkable African carnivore *Triphyophyllum peltatum* (Hutch. & Dalziel) Airy Shaw, which bears three different leaf types - one a carnivorous flypaper-like tendril (Fig. 1), the second a fully laminate leaf borne on the same basal rosette with the tendrils, and a third with distinct hooks occupying the climbing vine phase of the plant[19] - includes, through a

lineage-specific WGM, *n*=18 chromosomes descendant from a *x*=9 progenitor derived from a chromosomal fission event within the carnivorous *Caryophyllales* *x*=8 ancestor (Fig. 2). The vine *Ancistrocladus* Wall.[19], a non-trapping revertant from the carnivorous habit[3] (Fig. 1), is closely related to *Triphyophyllum* and shares its karyotype (Fig. 2). Their close flypaper-trapping relative *Drosophyllum lusitanicum* (L.) Link (Fig. 1), which has a large genome of over 10 gigabases (Gb)[20], remains to be sequenced. Together, from *Nepenthes* through *Ancistrocladus*, this incredible lineage of trapping/non-trapping morphological and physiological plant diversity is equally astonishing for its polyploid complexity.

Chromosome Structure and Genome Size

Contrasting Centromere Architectures and Repeat Dynamics in *Drosera* Genomes. The carnivorous plant genus *Drosera*, which includes both monocentric and holocentric species[21], presents a rare natural experiment in **centromere** evolution. In monocentric chromosomes, kinetochore function is confined to a single locus, meaning that broken chromosomal arm fragments lacking a centromere are typically lost. Holocentric chromosomes, by contrast, distribute kinetochore activity along their length, allowing chromosome fragments to segregate properly during cell division. This structural difference may enhance chromosomal flexibility and accelerate karyotype evolution in holocentric lineages[22,23].

Comparative genomic analyses of *D. capensis* (monocentric) and *D. regia* (holocentric) revealed striking contrasts in centromere structure, chromatin organization, and repeat dynamics[18]. Hi-C contact maps of *D. capensis* showed well-defined centromeric versus chromosome arm compartments (Fig. 3). Polymers of a single **satellite DNA** sequence were concentrated in a narrow region flanked by transposable elements, typical of a pericentromeric structure surrounding a monocentromere. In contrast, *D. regia* lacked large-scale chromatin compartmentalization, and its satellite DNA, nonhomologous to that of *D. capensis* in sequence, was distributed in short arrays along the chromosome, which had no localized monocentromere[18], matching the chromosomal organization of other holocentric plants with satellite-based holocentromeres (Fig. 3)[24-26].

Immunostaining supported these differences. In *D. capensis*, centromere-specific proteins CENH3 and KNL1 was restricted to single foci, co-localizing with spindle attachment sites. These signals were absent in *D. regia*, although α-tubulin was found along the full length of its chromosomes[18], indicating spindle attachment across the entire chromosome, which is a hallmark of holocentricity. This pattern suggested that *D. regia* may lack canonical centromere proteins or employ nonstandard kinetochore assembly mechanisms. Fluorescence *in situ* hybridization (FISH) using a satellite-specific probe on *D. regia* pachytene chromosomes further highlighted dispersed foci consistent with holocentric centromere function, in sharp contrast to the single clustered signal found in *D. capensis* chromosomes[18].

As alluded to above, satellite DNA evolution in the two *Drosera* species also differed markedly. In *D. capensis*, its satellite repeat sequences formed chromosome-specific clades with high intra-array sequence similarity, indicative of localized amplification and concerted evolution. In *D. regia*, its satellite repeats showed more of a starburst phylogeny, with short branches and extensive intermixing across chromosomes, suggesting recent, genome-wide expansion and homogenization via gene conversion[18].

A broader survey across 12 *Drosera* species revealed rapid repeat turnover that usually mirrored phylogeny but also appeared linked to centromere type (Fig. 4)[18]. The *D. capensis* repeat sequence was shared with closely related monocentric species, while the *D. regia* satellite sequence was unique to that species. Repeat content was especially divergent in *D.*

regia and *Drosera scorpioides* Planch. (Fig. 4), the latter a species with suspected holocentric features[27]. In *D. scorpioides*, the complete absence of shared tandem repeats may reflect an especially high turnover rate, contrasting with the more conserved repeat composition observed in the *D. capensis* and *Drosera aliciae* Raym.-Hamet species pair[18]. These contrasting dynamics in repeat stability may reflect underlying differences in centromere organization.

Together, these findings suggest that centromere architecture profoundly shapes chromatin structure, repeat evolution, and genome dynamics in this lineage. Furthermore, the *Drosera* system underscores how shifts between mono- and holocentricity can reshape not only centromere function, but also the evolutionary trajectory of repetitive DNA and chromosome structure in general.

Genome Size Reduction in Carnivorous Plants: The *Utricularia gibba* Model. Carnivorous plants exhibit striking genomic diversity, with the bladderwort *Utricularia gibba* L. representing a paradigm of extreme genome minimization[6,28]. Along with its complex vegetative morphology, intricate suction traps, and a typically-sized angiosperm gene complement, *U. gibba* has one of the smallest known flowering plant genomes at about 100 megabases (Mb)[6]. This compact architecture arises from a drastic reduction in non-genic DNA, challenging the notion that genome evolution is principally unidirectional and size-expanding.

Remarkably, *U. gibba* underwent at least two and probably three sequential tetraploidizations following its divergence from tomato and grape, rendering it effectively a 16-ploid relative to the core eudicot diploid ancestor[6,28]. Synteny analyses demonstrated extreme **fractionation** following these WGMs - nearly two-thirds of duplicated genes syntenic with tomato reverted to single copy, accompanied by massive contraction of intergenic spaces. Repetitive DNA constitutes less than 5% of the genome, which is exceptionally low for angiosperms. In keeping with the overall repetitive DNA loss, retrotransposons were severely depleted. Less than 400 retrotransposons were identified, with just about 100 potentially active. This indicates that the species' WGMs did not help drive genome expansion, perhaps instead providing redundant genetic material subsequently pruned substantially by deletion mechanisms. Therefore, the *U. gibba* genome reveals a counterintuitive interplay of polyploidization and aggressive DNA deletion driving reduction instead.

Multiple processes contributed to DNA loss in *U. gibba.* Microdeletions were apparent by highly compact regulatory regions: functional (experimentally confirmed) promoters are as small as 400 bp (including bidirectional promoters), intergenic spaces are ca. 50% shorter than in *Arabidopsis thaliana* (L.) Heynh. and introns are reduced in size and average number per gene. Recombinational deletion likely played a major role, as indicated by abundant solo long terminal repeat retrotransposons (solo-LTRs) in the genome. These solo-LTRs result from ectopic recombination between identical sequences at the ends of retrotransposons, often excising DNA in between them and drastically reducing repetitive DNA load in the process. Crucially, this deletion bias targeted nuclear DNA specifically; plastid and mitochondrial genomes showed no comparable contraction of intergenic regions. Interestingly as well, as expected for a genome under deletion pressure, given *U. gibba*'s rootless body plan, the losses of genes vital for root development such as the MADS-box genes *ANR1*[29] and *XAL1*[30], and the homeobox gene *WOX5* (which is crucial for the root stem cell niche[31,32]), aligns with their ready dispensability when purifying selection for retention is reduced or absent[28].

The persistence of a nearly full gene complement amidst such pervasive DNA loss suggests a model wherein polyploidies may have buffered against deleterious effects[28]. Following each WGM, the transient redundancy afforded by gene duplicates may have permitted aggressive deletion of DNA (and dispensable genes) without immediate loss of essential functions. Thereafter, fractionation could have progressively restored single-copy status for most genes. This "deletion buffering" hypothesis posits that in lineages with an inherent molecular deletion

bias, WGMs can facilitate genome reduction by providing a genetic safety net during DNA loss. In *U. gibba*, this process appears to have occurred repeatedly, compounded by suppressed retrotransposition. Consistent with this hypothesis, *U. gibba* also exhibits a high rate of duplicate gene turnover compared to model plant systems, indicating a highly dynamic pattern of gene additions and deletions over time[31].

With a mechanistic bias toward DNA deletion potentially due to compromised double-strand break repair while effectively suppressing retrotransposon activity, *U. gibba*'s genome could have reduced in size through random genetic drift alone, which would passively amplify such biases in small or inbred populations. Alternatively, some have proposed that this reduction reflects adaptation to nutrient-poor aquatic environments[33]. In this view, efficient DNA replication and repair might offer advantages under stress, with a smaller genome reducing phosphorus and nitrogen demands. However, this hypothesis lacks support - only some *Utricularia* and *Genlisea* species exhibit such tiny genomes[34], despite other carnivorous plants occupying similarly nutrient-poor habitats.

The *U. gibba* case demonstrates that angiosperms can achieve extreme genome compaction while maintaining phenotypic complexity. This challenges assumptions about the functional necessity of vast non-coding DNA in complex organisms, showing that a minimal intergenic landscape suffices for regulating a complete plant gene repertoire. Comparative genomics across the Lentibulariaceae family, exhibiting genome sizes from 60 Mb to 1.5 Gb[34], will promise further insights into the dynamics of carnivorous plant genome evolution.

Two Vignettes That Also Highlight *Utricularia gibba* genome size. First, a preliminary long-read genome assembly of *Utricularia macrorhiza* Leconte, a close relative of *U. gibba*[34], reveals a genome approximately twice the size (at ca. 228 Mb; J. Kirshner, C. Page, and V.A. Albert, unpublished data) of *U. gibba*'s (at ~100 Mb). Notably, the *U. macrorhiza* genome also exhibits an additional WGM relative to the *U. gibba* genome, which is already a high polyploid. These findings underscore unexpectedly dynamic polyploid trajectories within the genus. As such, further investigation of the *U. macrorhiza* genome promises to clarify the relative contributions of WGMs versus DNA repeat expansions in genome size evolution among closely related carnivorous plants.

Second, and remarkably, synteny between the genomes of *U. gibba* and *Dionaea* is strongly conserved despite over 100 million years of independent evolution. However, homologous genes retained from their common ancestor are concentrated within ~350-kilobase (Kb) regions in *U. gibba*, compared to syntenic blocks spanning 7.5 Mb or more in *D. muscipula* (Fig. 5). This contrast reflects substantial genome compaction in *U. gibba*, with retained gene space distributed across up to eight fractionated subgenomes, in contrast to only two in *Dionaea*.

Evolution of the Carnivorous Plant "Gene Space"

Gene Space Remodeling in Lentibulariaceae: Divergent Outcomes of Polyploidy and Tandem Duplication. Within the carnivorous Lentibulariaceae, contrasting modes of gene duplication have shaped gene space in markedly different ways. Comparative analysis of *Utricularia* and *Pinguicula*[35] revealed how WGM and **tandem duplication** can lead to divergent evolutionary outcomes, even within closely related genera. In *U. gibba*, which has likely experienced three tetraploidy events since the core-eudicot hexaploidization, retained homeologs that are enriched in transcription factors and other regulatory genes[6]. This pattern is consistent with observations in other angiosperms[36,37], and underscores a conserved trend following polyploidy wherein dosage balance (among other possible mechanisms) favors

survival of entire gene regulatory networks as opposed to one-off duplications of network members, which could throw off stoichiometry of regulation[38,39].

In contrast, tandem duplication, by which genes are duplicated locally and iteratively, has played a central role in shaping the adaptive gene space of *Utricularia*[6]. Tandemly duplicated genes in *U. gibba* are disproportionately involved in secondary metabolism, transport, and defense and digestive processes. Similar profiles favoring retention of secondary metabolic and other defense-related genes have been observed across many angiosperm species, in keeping with the notion that pathway addition (e.g., to generate new metabolites) is simple and subject to environmental adaptation[37,39]. However, digestive protein genes are a uniquely conspicuous tandem duplicate enrichment among some carnivores. In *Utricularia*, such duplicates include a lineage of genes encoding cysteine **proteases** involved in prey digestion (Fig. 6); these show strong or even exclusive expression in bladder traps, suggesting tissue-specific functional specialization[6]. These arrays, which may have been retained under natural selection pressures due to their direct roles in prey digestion, conceivably occur as "capacitors" for rapid and high-response digestive enzyme production. They also appear to have initially expanded independently of polyploid events. Despite the genome-wide deletion pressure associated with *U. gibba*'s ongoing genome size reduction (Fig. 5), some of these arrays have expanded further after some of its lineage-specific WGMs. Moreover, phylogenetic and protein structural analyses suggested that some of these duplicates have diverged functionally, whereby certain amino acid substitutions are located in the substrate-binding cleft and were predicted to affect enzyme activity. Notably, the tandem arrays on the dominant subgenome, where they are expected to be subject to stronger purifying selection, are better preserved, supporting a model in which local gene expansion and functional refinement acted together during genome-size reduction in the species.

A parallel, yet independently-evolved expansion of genes encoding cysteine proteases has occurred in *Pinguicula gigantea* Luhrs[35] (Fig. 6). Eighteen tandemly duplicated copies of this gene family are located on its chromosome 8, with a syntenic region on chromosome 5 that lacks the array, indicating the expansion occurred after the most recent WGM in this lineage[35]. Homologous genomic regions in other *Pinguicula* species (*Pinguicula agnata* Casper and *Pinguicula moctezumae* Zamudio & R. Z. Ortega) also show this asymmetry, reinforcing the conclusion that tandem duplication occurred post-polyploidy and were confined to a specific chromosomal context.

Despite their close phylogenetic relationship, *Utricularia* and *Pinguicula* cysteine protease genes do not intermix in a phylogenetic tree (Fig. 6). Instead, gene lineages follow assignment by genus, supporting independent (convergently evolved) expansions. Both *Utricularia* and *Pinguicula* tandem arrays are embedded in regions rich in repetitive DNA, such as large retrotransposon derivatives (LARDs) and other transposable elements[6,35]. Such local genome architecture may have facilitated tandem duplications via, e.g., unequal recombination.

Together, these findings illustrate how gene space can evolve through complementary but functionally distinct duplication mechanisms. While polyploidy can permit sub-/neofunctionalization of duplicate dosage-sensitive regulatory networks, tandem duplication can enable lineage-specific expansions of gene families under direct ecological or physiological selection. In *Utricularia* and *Pinguicula*, these parallel expansions of cysteine proteases exemplify how convergent pressures can repeatedly shape gene space through similar structural mechanisms, even in closely related species.

Gene Numbers in Carnivorous Plants are Not Reduced. Angiosperm genomes typically encode tens of thousands of genes. For example, the latest account of gene model number in the Arabidopsis genome (araport11) comprises 27,655 protein-coding genes[40]; rice, 42,189 genes (v7)[41]; and poplar, 34,699 genes (v4.1)[42]. Recent WGMs can inflate gene counts further

- hexaploid bread wheat, for example, carries 99,386 genes (v2.2)[43]. Thus, a "normal" angiosperm gene count is greater than 20,000, with higher values in polyploid lineages.

Recent genomic data show that carnivorous plants generally conform to this norm. Despite their highly specialized lifestyle, these plants have retained a gene repertoire comparable to that of other angiosperms. For example, the Southeast Asian pitcher plant *N. gracilis* possesses 34,010 predicted genes in a ~0.7 Gb genome[7], and the Australian pitcher plant *C. follicularis* has 36,503 predicted genes[5], both well within the usual range for flowering plants. Even the bladderwort *U. gibba*, with a diminutive ~100 Mb genome, contains a typical number of genes for a plant (29,666 genes), afforded by drastically reducing non-coding DNA instead of gene content[6]. A potential outlier is *Genlisea aurea* A. St.-Hil. (Lentibulariaceae), for which 17,755 genes were reported within a highly fragmented 43.4 Mb assembly of a genome with estimated size of 63.6 Mb[44]. Independent research will be requisite to establish the validity of these findings. Such possible size and gene-content exceptions remain rare, at least in carnivorous plants where well-annotated species sequenced to date exhibit gene numbers comparable to those of non-carnivorous plants.

Aside from the genome of *G. aurea*, as discussed above, those of the family Droseraceae illustrate how initial impressions of gene reduction need re-evaluation. For example, the draft genome assembly of *D. muscipula*, published together with assemblies for *Aldrovanda* and *Drosera spatulata* Labill., was initially described as among the more gene-poor among land plants[45]. This fueled the idea that carnivorous plants might have shed genes associated with non-carnivorous functions during the evolution of their novel nutritional strategy. Indeed, this first *Dionaea* assembly appeared to have an especially low gene count, with several genes important for its trapping physiology missing entirely in the draft genome assembly, for example, the jasmonate receptor *CORONATINE INSENSITIVE1*[46]. However, more recent evidence demonstrates that the low gene count was an artifact of incomplete data rather than an evolutionary genome streamlining. A new chromosome-scale assembly of *Dionaea* revealed 38,887 protein-coding genes[47], roughly double the number annotated in the earlier draft[45], bringing the Venus flytrap's gene count in line with other angiosperms. This improved assembly, aided by long-read sequencing and Omni-C scaffolding, recovered ~1 Gb of genomic sequence that had been missing entirely from the earlier draft. As a result, ~17,000 gene models absent in the 2020 assembly[45] were newly predicted, overturning the notion that the Venus flytrap genome is intrinsically gene-poor.

Understanding why the earlier misinterpretation arose highlights the importance of genome assembly completeness. The 2020 assembly[45] relied on short-read data and covered only ~3/5 of the large Venus flytrap genome, leaving many genic regions unassembled. Additionally, overly conservative gene prediction pipelines may have excluded small, divergent or partial gene sequences, further underestimating total gene number. In summary, the weight of genomic evidence indicates that carnivorous plants have not broadly reduced their gene numbers. Although genes required for arbuscular mycorrhizal symbiosis have been lost in most carnivorous lineages[48], and those involved in root development are sometimes undetected in rootless aquatic carnivores[6,31], such losses are modest at the genomic scale. This refined view dispels earlier claims of gene scarcity and underscores that carnivorous plants are no less genetically endowed than other flowering plants. As the adage reminds us, "absence of evidence is not evidence of absence", particularly when interpreting incomplete genome assemblies.

Subgenomic Adaptation in Polyploids. As discussed above, following polyploidization, genomes commonly undergo extensive gene loss (fractionation) to achieve a more diploid-like genomic structure[49]. During fractionation, subgenome dominance can emerge, where dominant subgenomes retain more genes (under enhanced purifying selection) and show higher

expression levels compared to recessive subgenomes (which are exposed to relaxed negative selection pressure)[50]. As such, recessive subgenomes experience greater gene loss, and the duplicate divergence afforded by the relaxed purifying selection they experience can permit them to serve as reservoirs for genetic novelty through neofunctionalization or subfunctionalization[51].

Recent research on the *N. gracilis* genome[7] provides a concrete example of these concepts. Through a chromosome-scale assembly, as described above, the genome was revealed to possess a decaploid karyotype with clear evidence of a 1:4, dominant:recessive subgenome relationship. Several novel trait-associated genes, linked to **dioecy** and specialized pitcher leaf functions, were located within specific subgenomes. Notably, a male-specific 1 Mb region was discovered within recessive chromosome 20 that contains three male-specific transcriptional regulatory genes. Among them is *DYSFUNCTIONAL TAPETUM1* (*DYT1*), a previously reported male-specific gene in *Nepenthes*[52], which in Arabidopsis directly regulates tapetum development[52]. Also present is *MALE MEIOCYTE DEATH1* (*MMD1*), a gene involved in pollen meiosis[53]. Surprisingly, this region also harbors a duplicated copy of the conserved transcription factor *LEAFY* (*LFY*), which is a hub-like regulator of central importance to reproductive development[54]. This rare duplicate, designated *LFY-Y* and located on the Y chromosome, is a strong candidate for direct involvement in the evolution of dioecy in *Nepenthes*.

In addition, a tandem cluster of *SENESCENCE-RELATED GENE 1* (*SRG1*) paralogs[55] was detected on recessive chromosomes. These genes are likely involved in scavenging reactive oxygen species (ROS) during prey digestion and nutrient absorption. Many of them have acquired tissue-specific expression in the digestive zone of *Nepenthes* pitcher leaves.

Collectively, these findings highlight the adaptive potential embedded in recessive subgenomes, demonstrating their important role as genetic reservoirs for evolutionary innovation and trait diversification in polyploid lineages such as *Nepenthes*.

Convergent Genetic Changes Associated with Carnivory. Convergent genetic and phenotypic solutions in distantly related taxa offer insights into evolutionary constraint, adaptation, and predictability[56]. Carnivorous plants, which evolved independently several times, illustrate such multi-level convergence. Genomic surveys now reveal recurring patterns in which identical or similar genetic changes underlie the evolution of carnivory.

Even in independently evolved carnivorous lineages, genes that encode digestive enzymes are frequently orthologous, indicating repeated co-option of the same gene lineages that encode pathogenesis-related proteins[5]. In addition, genes encoding digestive enzymes exhibit repeated bursts of duplication that supply raw material for novel functions[57]. Within Lentibulariaceae, both *Utricularia* and *Pinguicula* independently expanded their cysteine protease repertoires, as described above (Fig. 6). A similar surge of cysteine protease duplicates is seen in *D. capensis*[58], which belongs to a different carnivorous lineage. Aspartic proteases display a similar pattern in other taxa, with independent bursts of gene copy-number expansion in both *Nepenthes*[7] and *Cephalotus*[57]. Known as nepenthesins in *Nepenthes*, these aspartic proteases exhibit exceptional stability across a wide range of temperature and pH, suggesting adaptation to the unique chemical conditions of the pitcher fluid[59].

Convergent gene loss also shapes carnivorous genomes. Across most carnivorous lineages, dozens of genes required for arbuscular mycorrhizal (AM) symbiosis, a pathway otherwise conserved across land plants, have been independently lost, implying functional redundancy between AM associations and carnivory as alternative nutrient acquisition strategies[48]. Plastid *ndh* genes, which encode subunits of the thylakoid NADH dehydrogenase complex, have likewise been lost in many carnivorous clades[60]. Similar patterns in other heterotrophs[48,61] imply relaxed purifying selection on these loci once nutritional modes shift.

At the protein-sequence level, digestive enzymes such as glycoside hydrolase family 19 chitinases, purple acid phosphatases, RNase T2 and S1-P1 nuclease display an excess of convergent amino-acid substitutions between *Cephalotus* and Caryophyllales carnivores, including *Nepenthes*[5,62]. Remarkably, in the *Cephalotus* genome, some purple acid phosphatase and RNase T2 genes are adjacent, and their co-expression indicates they likely share local regulatory control[5]. Convergence extends beyond digestive enzymes to polygalacturonase-inhibiting proteins (PGIPs) and HIGH AFFINITY K^+ TRANSPORTER 5 (HAK5). PGIP substitutions cluster at residues predicted to contact fungal polygalacturonase[62], hinting at a role in suppressing fungal growth in pitcher fluid, whereas HAK5 likely promotes prey-derived potassium uptake, as characterized in *D. muscipula*[63]. The biochemical consequences of many pinpointed amino acid changes remain to be tested.

Collectively, repeated gene duplication, targeted gene loss, and convergent amino-acid substitutions point to a shared genetic toolkit redeployed throughout the evolution of plant carnivory. Ongoing functional and comparative studies will clarify whether the same mechanisms also shape other facets of the carnivorous syndrome.

Concluding Remarks and Future Perspectives

Genomic investigations into carnivorous plants have revealed systems with great complexity, considerable convergence, and compelling evolutionary paradoxes. While significant progress has elucidated the contributions of polyploidy (particularly allopolyploidy among the Caryophyllales carnivores), extremes in genome architecture (e.g., holocentricity and genome size reduction), and molecular parallelism accompanying the carnivorous syndrome, crucial issues remain to be addressed that define a diversity of future research opportunities (see **Outstanding Questions**). Unresolved questions include: What core developmental genetic pathways are recruited to transform standard leaf developmental programs into the diverse, complex trap types observed within closely interrelated lineages (e.g., snap-traps versus pitchers within the Caryophyllales carnivore clade)? Correspondingly, what regulatory networks might be recurrently co-opted among the convergent trap forms of the three independently-evolved pitcher plants? How do transitions in centromere organization (monocentric to holocentric) impact genome stability and potentially influence rates of morphological radiation, as inferred from the great species diversity in the *Drosera* genus and the contrasting repeat dynamics revealed there? What specific biochemical properties, such as altered substrate affinity or catabolic activity and enhanced stability under acidic trap conditions, are conferred by convergent amino acid substitutions identified in key digestive enzymes? Is the "deletion buffering" hypothesis, which has been proposed to explain *U. gibba*'s genome reduction alongside polyploidy, a general mechanism among parallel genome size reductions in the genus as well as among *Genlisea* species? What molecular machinery implements this DNA loss, and might it differ among lineages that convergently undergo such losses? And regarding these genome size reductions, are some genes or functional genetic categories lost in parallel? Since allopolyploidy appears to be common among the Caryophyllales carnivores, how consistently (for example, in *Dionaea* or *Drosera regia*) might subgenome dominance serve as an intermediary condition for evolutionary innovation, particularly within recessive genomic compartments? Addressing these challenges will necessitate a multifaceted approach. Firstly, bridging critical sampling gaps through high-quality, chromosome-scale assemblies for key missing taxa (*Sarracenia*, *Drosophyllum*, *Genlisea*, *Triphyophyllum*, *Aldrovanda* - and others such as *Brocchinia* J.H.Schult. ex J.A.Schult. & J.H.Schult., *Heliamphora* Benth., *Darlingtonia* Torr., *Roridula* Forssk., *Byblis*

Salisb., *Philcoxia* P. Taylor & V. C. Souza, and *Triantha* (Nutt.) Baker, which were not addressed here) will be vital to complete a more complete comparative framework. Secondly, the field must advance beyond descriptive genomics towards functional dissection. Such research may utilize state-of-the-art genome editing and other molecular assays to validate the roles of candidate genes (such as the *LFY-Y* gene of *Nepenthes*), or of specific amino acid changes. Finally, integrating genomic research with ecophysiological data will be vital for elucidating the selective pressures that shaped the remarkable genome-to-phenome architectural evolution of carnivorous plants.

Outstanding Questions

- What specific molecular mechanisms drive extreme genome compaction in *Utricularia gibba*? Can the proposed "deletion buffering" model (where polyploidy facilitates DNA loss) be experimentally validated, and does it apply to other *Utricularia*/*Genlisea* species? What biases in DNA repair pathways might enable it?
- How do transitions between monocentric and holocentric chromosome organization in *Drosera* directly impact genome stability, repeat dynamics, and rates of karyotype evolution/speciation? Does holocentricity accelerate morphological diversification in this lineage?
- What core developmental genetic pathways are recurrently rewired to generate diverse trap morphologies (e.g., snap-traps versus pitchers) within closely interrelated lineages such as carnivorous Caryophyllales? Similarly, what regulatory networks might be convergently recruited to generate similar forms (e.g., pitchers) in distantly related carnivore lineages>
- Do the convergent amino acid substitutions identified in key digestive enzymes (e.g., chitinases, phosphatases, proteases) across carnivorous lineages confer specific biochemical advantages, such as enhanced stability in trap fluids or altered substrate specificity? What are their precise functional impacts?
- How pervasive might the functional potential of subgenome dominance be in the complex polyploid genomes of Caryophyllales carnivores beyond *Nepenthes*? Do recessive subgenomes consistently act as a reservoir for evolutionary innovation (e.g., novel trap features, dioecy genes)?
- Beyond digestive enzymes, are there broader, parallel patterns of copy number dynamism within gene families associated with the carnivorous habit across independent lineages?
- What role does local tandem duplication versus whole-genome multiplication play in the rapid evolution and functional refinement of key adaptive traits (e.g., prey digestion, nutrient sensing) across different carnivorous plant clades?

Glossary

Allopolyploidy: A form of whole genome multiplication in which chromosome sets from different species combine through hybridization, requiring rebalancing of meiotic pairing, and sometimes leading to novel genomic and phenotypic traits.

Autopolyploidy: Whole-genome multiplication involving chromosome set duplication within a single species, without hybridization.

Caryophyllales: An order of flowering plants including carnivorous genera such as *Nepenthes*, *Drosera*, *Dionaea*, and *Triphyophyllum*, and non-carnivores such as beets and spinach.

Centromere: A chromosomal region where the kinetochore assembles and spindle fibers attach during cell division. Centromeres are usually composed of satellite DNA repeats with foci that may be localized (monocentric) or distributed along the chromosome (holocentric).

Convergent evolution: The independent evolution of similar traits or functions in distantly related lineages, often driven by similar selective pressures, as seen in repeated origins of trap types in carnivorous plants.

Dioecy: A sexual system in which individual plants are either male or female. In *Nepenthes*, dioecy is associated with Y-linked reproductive genes.

Fractionation: The process of gene loss following polyploidy, which can result in one subgenome becoming dominant in gene retention and expression.

Polyploidy: The condition of possessing more than two complete sets of chromosomes, arising through autopolyploidy or allopolyploidy, and a major driver of plant genome evolution.

Proteases: Digestive enzymes that cleave peptide bonds. In some carnivores, they are aspartic proteases (cleaving at aspartic acid residues) while in other lineages they are cysteine proteases (which cleave at cysteine residues).

Satellite DNA: Tandemly repeated non-coding sequences that are often associated with centromeres. Their distribution patterns differ between monocentric and holocentric species.

Synteny: Conservation of gene order between chromosomes of different species or subgenomes, used to infer evolutionary relationships and genome rearrangements.

Tandem duplication: Local gene copy number expansion via unequal crossing over or replication slippage. In carnivores, tandem duplications have sometimes expanded digestive enzyme gene families.

Whole genome multiplication (WGM): Events in which an organism's entire chromosome set is duplicated or combined with another species' set, increasing ploidy level and creating opportunities for genetic innovation.

Acknowledgments

We acknowledge the following sources for funding: United States National Science Foundation grants (2030871 to V.A.A.), Human Frontier Science Program Young Investigators grant (RGY0082/2021 to K.F.), JSPS KAKENHI (23K20050 to K.F.), Max Planck Society (core funding to A.M.), European Research Council (ERC StG HoloRECOMB, grant no. 101114879 to A.M.), and German Science Foundation (DFG MA 9363/6-1 to A.M.). L.A.R. is financially supported by DFG (German Research Foundation) grant MA 9363/6-1. The DFG also funded this work under Germany's Excellence Strategy—EXC 493 2048/1–390686111 (to A.M.). We also thank Christopher Page and Nicholas Pratt for their contributions toward unpublished genome data.

Disclosure: AI tools (ChatGPT and DeepSeek) were used to ensure the text was free of errors in grammar, spelling, punctuation and tone. No generative editorial work or autonomous content creation was performed.

Author Contributions

V.A.A., K.F., A.M., L.A.R., Y.G. and S.K. wrote the first draft with input from all authors. L.A.R., V.A.A., S.K. and J.K. generated graphics, with analytical support from S.J.F.

Competing Interests

The authors declare no competing interests.

Figures

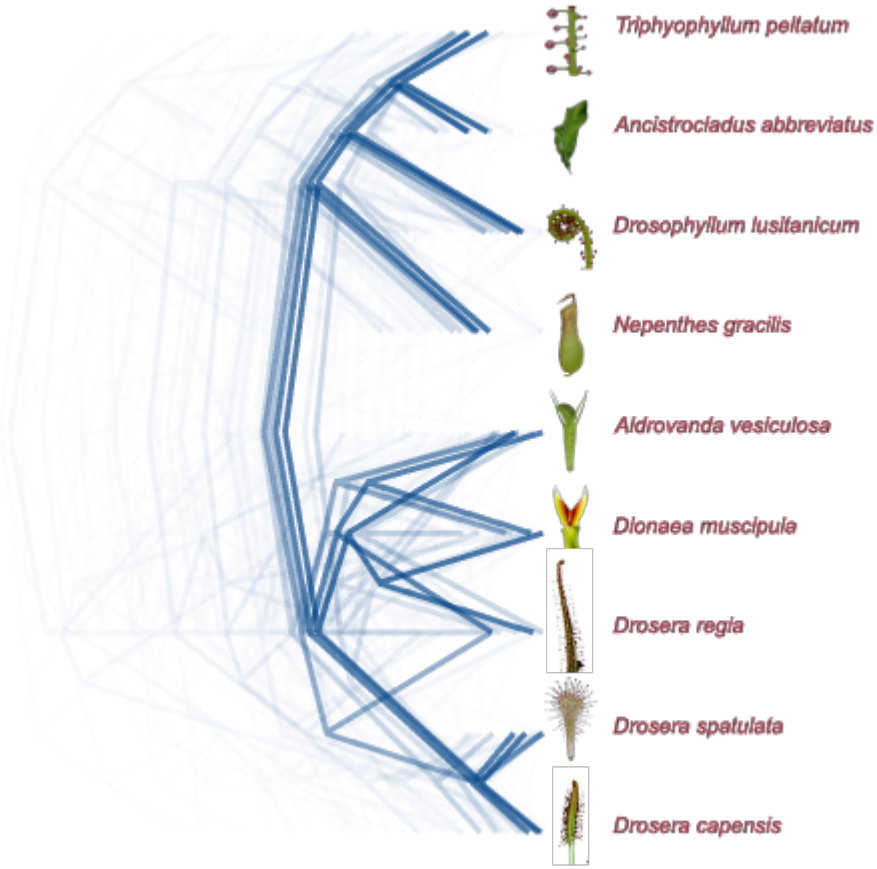

Figure 1. A DensiTree cloudogram[64] demonstrates gene tree conflict among 563 phylogenies of single-copy orthologous genes. Consensus trees representing distinct tree topologies found among the set are shown, with their intensities proportional to their frequency. Some loci favor a close relationship between *D. regia* and *Aldrovanda* plus *Dionaea*, whereas others place *D. regia* as sister to the other *Drosera* species pair. These conflicting topologies highlight the allopolyploid hybrid origins of *Dionaea* and *D. regia*. Interestingly, the 5-subgenome allopolyploid *N. gracilis* shows very little gene tree incongruence among this set of taxa and loci, possibly reflecting the absence (extinction) of lineages that could have donated its subgenomes. See also Ávila Robledillo et al., 2025[18].

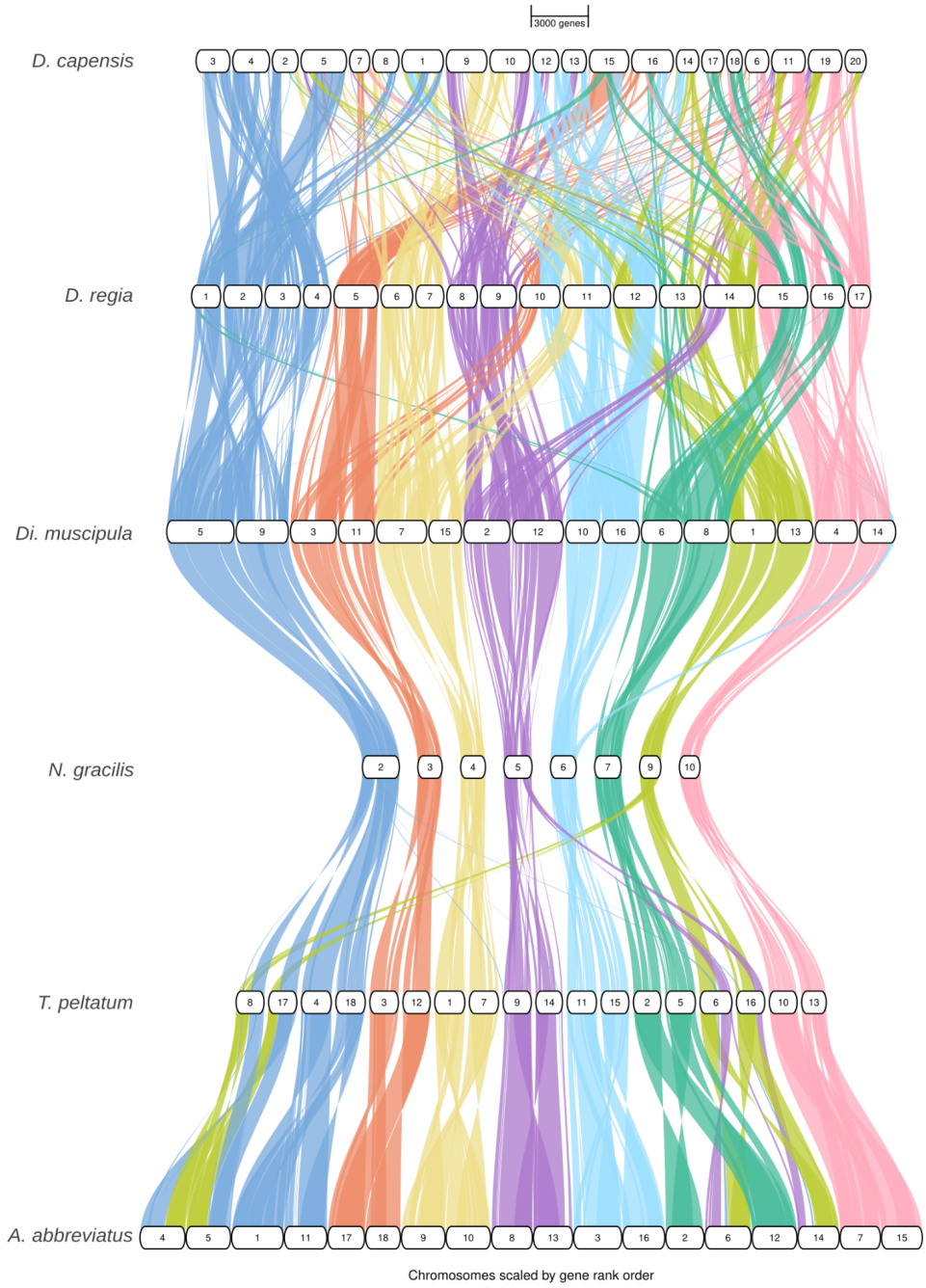

Figure 2. Syntenic interrelationships among chromosomes of species of the Caryophyllales carnivorous plant lineage. The syntenic lines based on gene orders of syntenic orthologs pass through the 8-chromosome *Nepenthes gracilis* dominant subgenome. *Dionaea* is duplicate in structure compared to the 8-chromosome *Nepenthes* subgenome. *D. regia* shows extensive chromosomal fusions within a triplicated structure based on the same ancestral *x*=8 chromosome set. The highly rearranged *D. capensis* genome, on the other hand, underwent a duplication of its triplicate architecture, which may or may not be shared with *D. regia*'s. *Triphyophyllum* and *Ancistrocladus* share a chromosomal fission and a unique translocation that generated nine chromosomes from the *x*=8 ancestor also seen in *Nepenthes*, followed by a subsequent shared WGM[18].

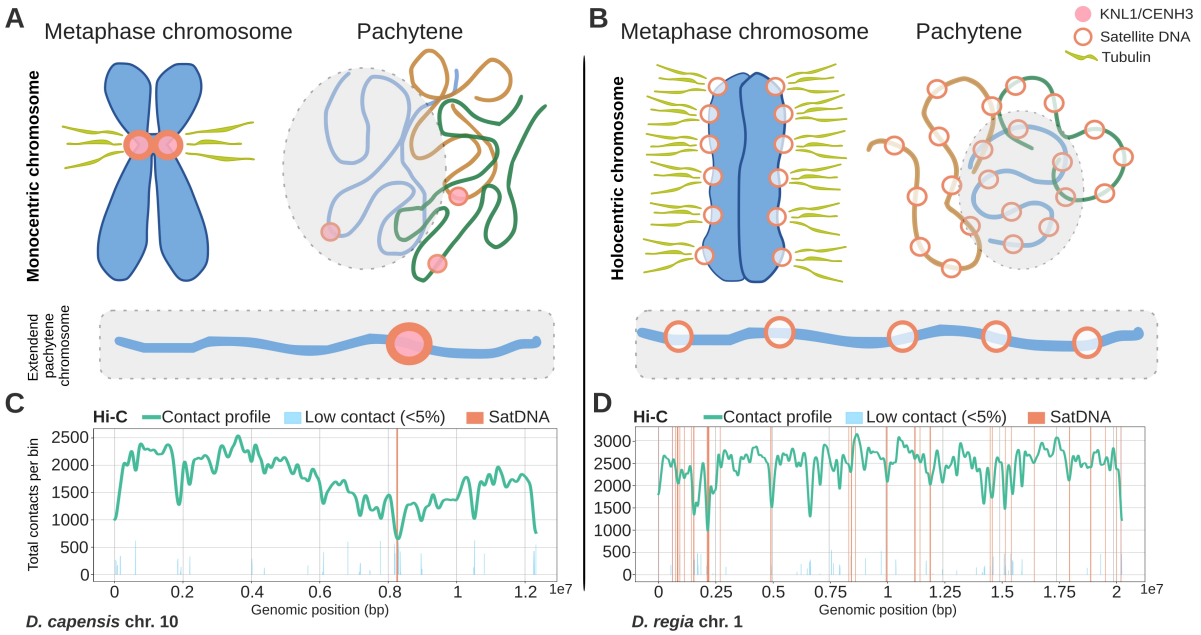

Figure 3. Comparative centromere organization and its impact on 3D genome architecture. (A-B) Schematic representations of chromosome structure in monocentric **(A)** and holocentric **(B)** species across three configurations: condensed metaphase chromosome, pachytene nucleus, and extended pachytene chromosome. In monocentric species, centromeric activity is restricted to a single chromosomal region, which is typically observed as a primary constriction at metaphase. In contrast, holocentric species exhibit diffuse centromeric activity along the entire length of the chromosome. **(C-D)** The one-dimensional Hi-C contact profiles of chromosome 10 of *D. capensis* and chromosome 1 of *D. regia* reflect the differences in chromosomal organization. The total contacts per bin are shown as a smoothed line (green), and genomic bins falling within the bottom 5th percentile of contact intensity are highlighted in light blue, indicating regions with locally reduced contact frequency. Regions containing satDNA repeats are shown as vertical orange lines. See also Ávila Robledillo et al., 2025[18].

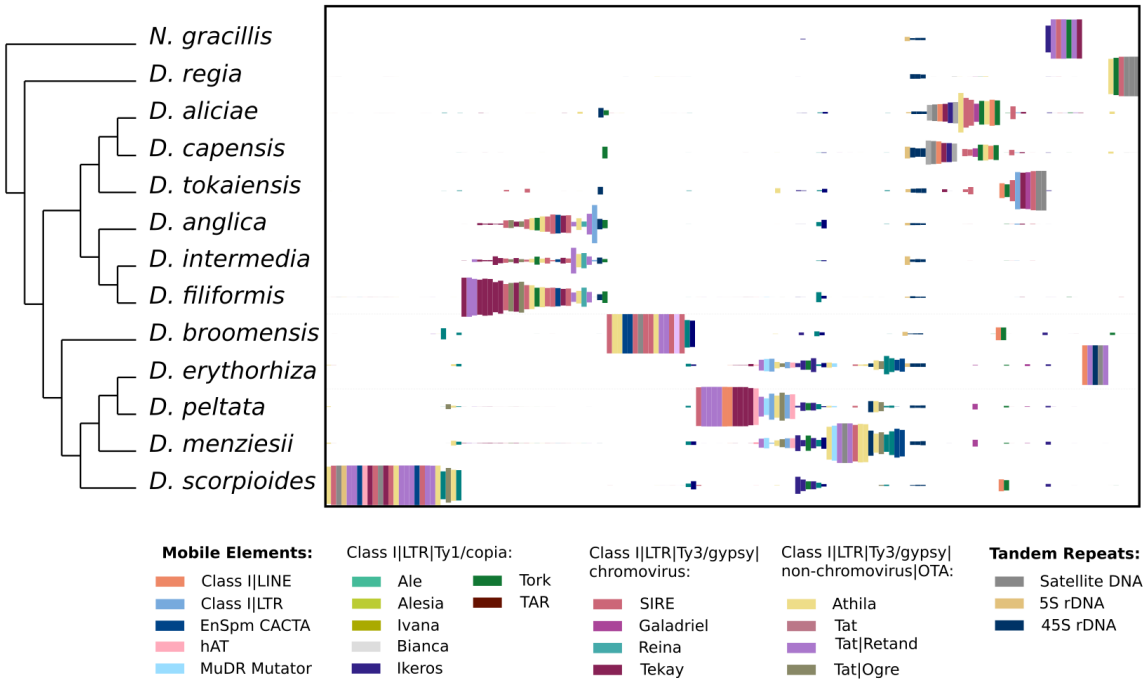

Figure 4. Comparative analysis of 12 *Drosera* species and their close relative *Nepenthes gracilis*. This graphic summarizes the results of a RepeatExplorer comparative analysis, showing the distribution and relative abundance of repetitive DNA elements across genomes, grouped by major repeat families. Each column represents a family of repeat elements, and each vertical bar within a row (species) shows the proportion of that group within the total analyzed DNA reads for each species. The figure highlights how repeat composition varies across species, without considering genome size. See also Ávila Robledillo et al., 2025[18].

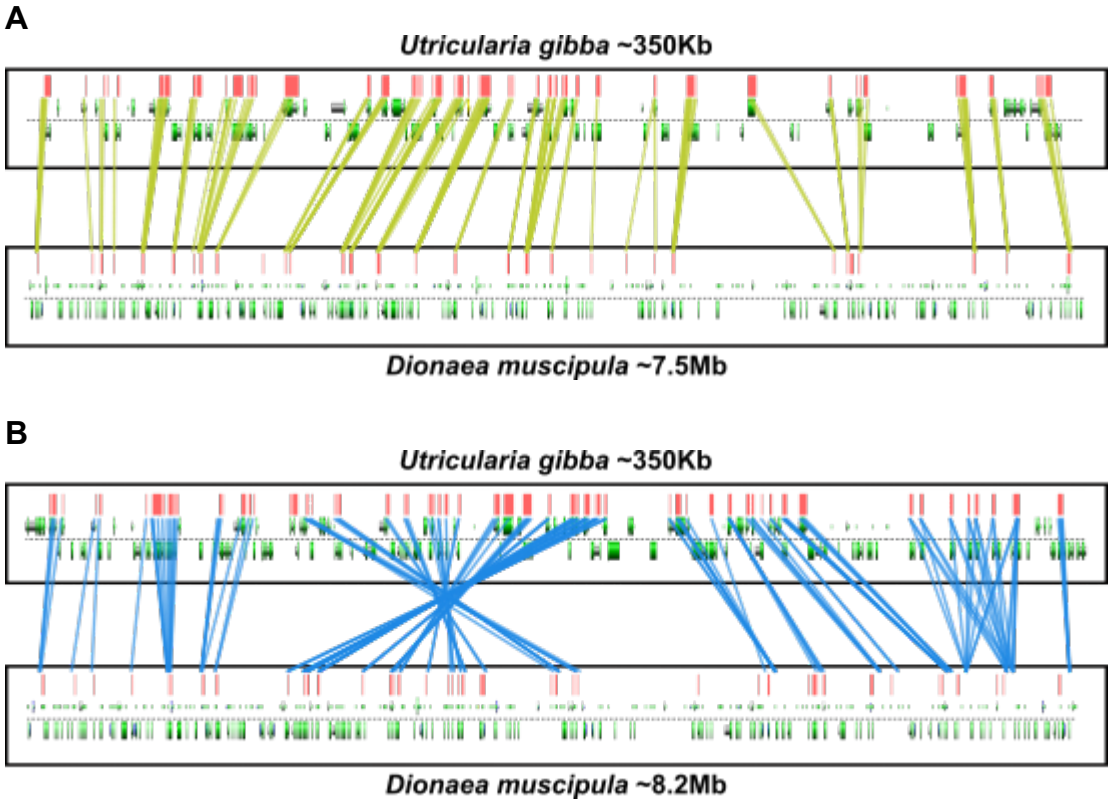

Figure 5. Syntenic views of homologous blocks from the *U. gibba* (~100 Mb) and D. *muscipula* (~2500 Mb) genomes. These examples (**A** and **B**) show massive genome size and gene density differences while linear synteny is preserved since ancient common ancestry. Only about 350 Kb encompasses the gene space in the *U. gibba* genomic blocks, while syntenic genes cover about 7.5 (**A**) and 8.2 (**B**) Mb in the *Dionaea* windows. The *Utricularia* genomic block is heavily fractionated, containing far fewer genes in between syntenic homologs than *Dionaea*. A large inversion can be seen in view **B**, but synteny between the species is strongly conserved despite this rearrangement.

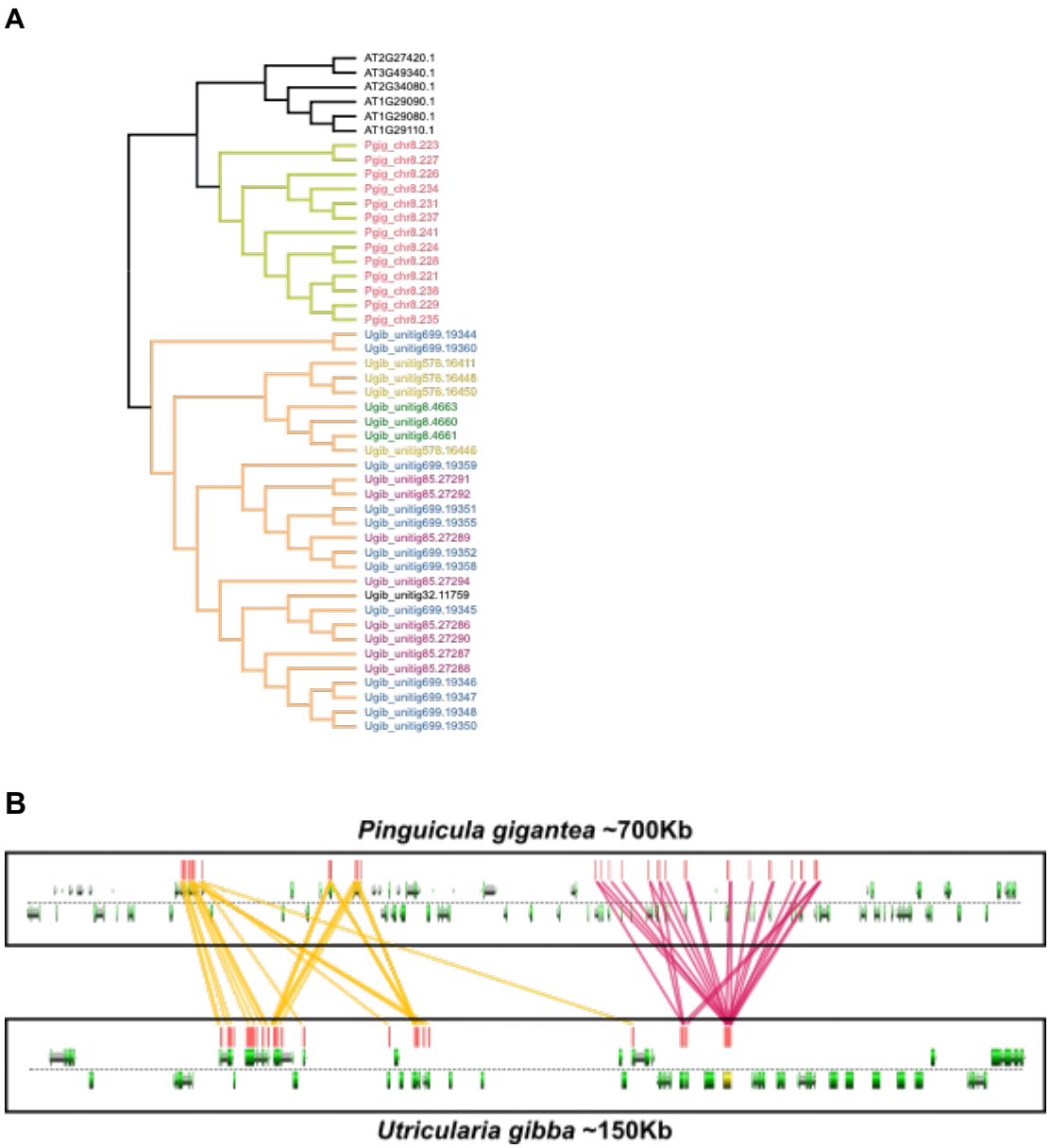

Figure 6. Cysteine protease gene phylogeny (A) shows convergent tandem gene expansions in *U. gibba* and *P. gigantea*. Since the gene copies from the two species do not form mixed clades in the tree, the tandem duplications occurred independently. In *Pinguicula*, one tandem array lies on a single chromosome, while in *Utricularia*, there are 4 tandem arrays on different chromosomes (unitigs), which are color coded. Interestingly, the tree topology within the *Utricularia* lineage does not wholly reflect position along the genome, with one possible explanation being seeding of new genomic locations by translocating duplicates with origins on other chromosomes. Also interesting is that in *Utricularia*, one small collection of tandem duplicates on unitig 8 is syntenic with the large tandem array on *Pinguicula* chromosome 8 (**B**). In (**A**), a lineage of Arabidopsis cysteine protease genes (with prefix AT) is sister to the *Pinguicula* array. Gene model IDs for *U. gibba* and *P. gigantea* have the prefixes Ugib and Pgig, respectively, and gene number (following the period sign) reflects physical position along the chromosome (chr) or unitig. Greater syntenic preservation is present but not shown in this view. Data are from Lan et al., 2017[6] and Fleck et al., 2025[35].

References

1. Albert, V.A., Williams, S.E., and Chase, M.W. (1992). Carnivorous plants: phylogeny and structural evolution. Science *257*, 1491-1495.
2. Fleck, S.J., and Jobson, R.W. (2023). Molecular phylogenomics reveals the deep evolutionary history of carnivory across land plants. Plants *12*, 3356.
3. Fleischmann, A., Schlauer, J., Smith, S.A., Givnish, T.J., Ellison, A., and Adamec, L. (2018). Evolution of carnivory in angiosperms. Carnivorous plants: physiology, ecology, and evolution *1*, 22-42.
4. Soltis, P.S., Marchant, D.B., Van de Peer, Y., and Soltis, D.E. (2015). Polyploidy and genome evolution in plants. Current opinion in genetics & development *35*, 119-125.
5. Fukushima, K., Fang, X., Alvarez-Ponce, D., Cai, H., Carretero-Paulet, L., Chen, C., Chang, T.-H., Farr, K.M., Fujita, T., and Hiwatashi, Y. (2017). Genome of the pitcher plant Cephalotus reveals genetic changes associated with carnivory. Nature Ecology & Evolution *1*, 0059.
6. Lan, T., Renner, T., Ibarra-Laclette, E., Farr, K.M., Chang, T.-H., Cervantes-Pérez, S.A., Zheng, C., Sankoff, D., Tang, H., and Purbojati, R.W. (2017). Long-read sequencing uncovers the adaptive topography of a carnivorous plant genome. Proceedings of the National Academy of Sciences *114*, E4435-E4441.
7. Saul, F., Scharmann, M., Wakatake, T., Rajaraman, S., Marques, A., Freund, M., Bringmann, G., Channon, L., Becker, D., and Carroll, E. (2023). Subgenome dominance shapes novel gene evolution in the decaploid pitcher plant Nepenthes gracilis. Nature plants *9*, 2000-2015.
8. Bemm, F., Becker, D., Larisch, C., Kreuzer, I., Escalante-Perez, M., Schulze, W.X., Ankenbrand, M., Van de Weyer, A.-L., Krol, E., and Al-Rasheid, K.A. (2016). Venus flytrap carnivorous lifestyle builds on herbivore defense strategies. Genome Research *26*, 812-825.
9. Pavlovič, A., and Mithöfer, A. (2019). Jasmonate signalling in carnivorous plants: copycat of plant defence mechanisms. Journal of experimental botany *70*, 3379-3389.
10. Barker, M.S., Arrigo, N., Baniaga, A.E., Li, Z., and Levin, D.A. (2016). On the relative abundance of autopolyploids and allopolyploids. New phytologist *210*, 391-398.
11. Arrigo, N., and Barker, M.S. (2012). Rarely successful polyploids and their legacy in plant genomes. Current opinion in plant biology *15*, 140-146.
12. Chanderbali, A.S., Berger, B.A., Howarth, D.G., Soltis, D.E., and Soltis, P.S. (2017). Evolution of floral diversity: genomics, genes and gamma. Philosophical Transactions of the Royal Society B: Biological Sciences *372*, 20150509.
13. Jiao, Y., Leebens-Mack, J., Ayyampalayam, S., Bowers, J.E., McKain, M.R., McNeal, J., Rolf, M., Ruzicka, D.R., Wafula, E., and Wickett, N.J. (2012). A genome triplication associated with early diversification of the core eudicots. Genome biology *13*, R3.
14. Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature *449*, 463-467. 10.1038/nature06148.
15. Srivastava, A., Rogers, W.L., Breton, C.M., Cai, L., and Malmberg, R.L. (2011). Transcriptome analysis of Sarracenia, an insectivorous plant. DNA research *18*, 253-261.
16. Stull, G.W., Pham, K.K., Soltis, P.S., and Soltis, D.E. (2023). Deep reticulation: the long legacy of hybridization in vascular plant evolution. The Plant Journal *114*, 743-766.
17. Project, A.G., Albert, V.A., Barbazuk, W.B., dePamphilis, C.W., Der, J.P., Leebens-Mack, J., Ma, H., Palmer, J.D., Rounsley, S., and Sankoff, D. (2013). The Amborella genome and the evolution of flowering plants. Science *342*, 1241089.
18. Ávila Robledillo, L., Fleck, S.J., Kirshner, J., Becker, D., Bhatia, A., Bringmann, G., Brock, J.R., Drautz-Moses, D., Freund, M., Hedrich, R., et al. (2025). Contrasting centromere structures in carnivorous sundews uncover a dynamic evolutionary trajectory shaped by polyploidy and chromosomal remodeling. bioRxiv, 2025.2007.2027.666937. 10.1101/2025.07.27.666937.
19. Metcalfe, C. (1951). The anatomical structure of the Dioncophyllaceae in relation to the taxonomic affinities of the family. Kew Bulletin *6*, 351-368.

20. Rothfels, K., and Heimburger, M. (1968). Chromosome size and DNA values in sundews (Droseraceae). Chromosoma *25*, 96-103.
21. Veleba, A., Šmarda, P., Zedek, F., Horová, L., Šmerda, J., and Bureš, P. (2017). Evolution of genome size and genomic GC content in carnivorous holokinetics (Droseraceae). Annals of Botany *119*, 409-416.
22. Marques, A., and Pedrosa-Harand, A. (2016). Holocentromere identity: from the typical mitotic linear structure to the great plasticity of meiotic holocentromeres. Chromosoma *125*, 669-681.
23. Steckenborn, S., and Marques, A. (2025). Centromere diversity and its evolutionary impacts on plant karyotypes and plant reproduction. New Phytologist *245*, 1879-1886.
24. Hofstatter, P.G., Thangavel, G., Lux, T., Neumann, P., Vondrak, T., Novak, P., Zhang, M., Costa, L., Castellani, M., and Scott, A. (2022). Repeat-based holocentromeres influence genome architecture and karyotype evolution. Cell *185*, 3153-3168. e3118.
25. Kuo, Y.-T., Câmara, A.S., Schubert, V., Neumann, P., Macas, J., Melzer, M., Chen, J., Fuchs, J., Abel, S., and Klocke, E. (2023). Holocentromeres can consist of merely a few megabase-sized satellite arrays. Nature Communications *14*, 3502.
26. Mata-Sucre, Y., Krátká, M., Oliveira, L., Neumann, P., Macas, J., Schubert, V., Huettel, B., Kejnovský, E., Houben, A., and Pedrosa-Harand, A. (2024). Repeat-based holocentromeres of the woodrush Luzula sylvatica reveal insights into the evolutionary transition to holocentricity. Nature Communications *15*, 9565.
27. Zedek, F., Veselý, P., Horová, L., and Bureš, P. (2016). Flow cytometry may allow microscope-independent detection of holocentric chromosomes in plants. Scientific Reports *6*, 27161.
28. Ibarra-Laclette, E., Lyons, E., Hernández-Guzmán, G., Pérez-Torres, C.A., Carretero-Paulet, L., Chang, T.-H., Lan, T., Welch, A.J., Juárez, M.J.A., and Simpson, J. (2013). Architecture and evolution of a minute plant genome. Nature *498*, 94-98.
29. Gan, Y., Bernreiter, A., Filleur, S., Abram, B., and Forde, B.G. (2012). Overexpressing the ANR1 MADS-box gene in transgenic plants provides new insights into its role in the nitrate regulation of root development. Plant and Cell Physiology *53*, 1003-1016.
30. Zluhan-Martínez, E., Castañón-Suárez, C.A., Gutiérrez-Rodríguez, M.A., Lledías, F., Zhang, T., Peng, J.T., Dickinson, J., Sánchez Rodríguez, D.B., Sánchez, M.d.l.P., and García-Ponce, B. (2025). The MADS-box gene XAANTAL1 participates in Arabidopsis thaliana primary root growth and columella stem cell patterns in response to ROS, via direct regulation of PEROXIDASE 28 and RETINOBLASTOMA-RELATED genes. Journal of Experimental Botany *76*, 411-432.
31. Carretero-Paulet, L., Librado, P., Chang, T.-H., Ibarra-Laclette, E., Herrera-Estrella, L., Rozas, J., and Albert, V.A. (2015). High gene family turnover rates and gene space adaptation in the compact genome of the carnivorous plant Utricularia gibba. Molecular Biology and Evolution *32*, 1284-1295.
32. Zhang, N., Bitterli, P., Oluoch, P., Hermann, M., Aichinger, E., Groot, E.P., and Laux, T. (2025). Deciphering the molecular logic of WOX5 function in the root stem cell organizer. The EMBO Journal *44*, 281-303.
33. Veleba, A., Zedek, F., Horová, L., Veselý, P., Srba, M., Šmarda, P., and Bureš, P. (2020). Is the evolution of carnivory connected with genome size reduction? American Journal of Botany *107*, 1253-1259.
34. Veleba, A., Bureš, P., Adamec, L., Šmarda, P., Lipnerová, I., and Horová, L. (2014). Genome size and genomic GC content evolution in the miniature genome-sized family Lentibulariaceae. New Phytologist *203*, 22-28.
35. Fleck, S.J., Kirshner, J., Rajaraman, S., Lan, T., Tomlin, C., Herrera-Estrella, L., Jobson, R.W., Ibarra-Laclette, E., Rosas, M.M., and Michael, T.P. (2025). Genomic Insights into the Evolution of Carnivory in the Giant Butterwort, Pinguicula gigantea: Chromosome-Level Assembly and Phylogenomic Analysis. bioRxiv, 2025.2004. 2005.646448.
36. Conant, G.C., Birchler, J.A., and Pires, J.C. (2014). Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. Current opinion in plant biology *19*, 91-98.
37. Salojärvi, J., Smolander, O.-P., Nieminen, K., Rajaraman, S., Safronov, O., Safdari, P., Lamminmäki, A., Immanen, J., Lan, T., and Tanskanen, J. (2017). Genome sequencing and

population genomic analyses provide insights into the adaptive landscape of silver birch. Nature genetics *49*, 904-912.
38. Birchler, J.A., and Veitia, R.A. (2012). Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. Proceedings of the National Academy of Sciences *109*, 14746-14753.
39. Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. Annual review of plant biology *60*, 433-453.
40. Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S., and Town, C.D. (2017). Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. The Plant Journal *89*, 789-804.
41. Hamilton, J.P., Li, C., and Buell, C.R. (2025). The rice genome annotation project: an updated database for mining the rice genome. Nucleic Acids Research *53*, D1614-D1622.
42. Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., and Salamov, A. (2006). The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). science *313*, 1596-1604.
43. Consortium, I.W.G.S., Mayer, K.F., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., and Lukaszewski, A.J. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science *345*, 1251788.
44. Leushkin, E.V., Sutormin, R.A., Nabieva, E.R., Penin, A.A., Kondrashov, A.S., and Logacheva, M.D. (2013). The miniature genome of a carnivorous plant Genlisea aurea contains a low number of genes and short non-coding sequences. BMC genomics *14*, 476.
45. Palfalvi, G., Hackl, T., Terhoeven, N., Shibata, T.F., Nishiyama, T., Ankenbrand, M., Becker, D., Förster, F., Freund, M., and Iosip, A. (2020). Genomes of the Venus flytrap and close relatives unveil the roots of plant carnivory. Current Biology *30*, 2312-2320. e2315.
46. Yan, J., Zhang, C., Gu, M., Bai, Z., Zhang, W., Qi, T., Cheng, Z., Peng, W., Luo, H., and Nan, F. (2009). The Arabidopsis CORONATINE INSENSITIVE1 protein is a jasmonate receptor. The Plant Cell *21*, 2220-2236.
47. Trebing, S., Freund, M., Iosip, A.L., Diblasi, C., Krennerich, V., Kirshner, J., Sato, M.P., Marques, A., Saitou, M., and Becker, D. (2025). Impaired trap closure in the counting-deficient Venus flytrap mutant DYSCALCULIA is caused by cell wall biomechanics. bioRxiv, 2025.2006. 2026.661685.
48. Montero, H., Freund, M., and Fukushima, K. (2025). Convergent losses of arbuscular mycorrhizal symbiosis in carnivorous plants. bioRxiv, 2025.2004. 2003.646726.
49. Mandakova, T., Li, Z., Barker, M.S., and Lysak, M.A. (2017). Diverse genome organization following 13 independent mesopolyploid events in Brassicaceae contrasts with convergent patterns of gene retention. The Plant Journal *91*, 3-21.
50. Cheng, F., Wu, J., Cai, X., Liang, J., Freeling, M., and Wang, X. (2018). Gene retention, fractionation and subgenome differences in polyploid plants. Nature plants *4*, 258-268.
51. Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. science *290*, 1151-1155.
52. Zhang, W., Sun, Y., Timofejeva, L., Chen, C., Grossniklaus, U., and Ma, H. (2006). Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor.
53. Yang, X., Makaroff, C.A., and Ma, H. (2003). The Arabidopsis MALE MEIOCYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis. The Plant Cell *15*, 1281-1295.
54. Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M. (1992). LEAFY controls floral meristem identity in Arabidopsis. Cell *69*, 843-859.
55. Callard, D., Axelos, M., and Mazzolini, L. (1996). Novel molecular markers for late phases of the growth cycle of Arabidopsis thaliana cell-suspension cultures are expressed during organ senescence. Plant Physiology *112*, 705-715.
56. Losos, J.B. (2011). Convergence, adaptation, and constraint. Evolution *65*, 1827-1840.
57. Nei, M., and Rooney, A.P. (2005). Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. *39*, 121-152.

58. Butts, C.T., Bierma, J.C., and Martin, R.W. (2016). Novel proteases from the genome of the carnivorous plant Drosera capensis: structural prediction and comparative analysis. Proteins: Structure, Function, and Bioinformatics *84*, 1517-1533.
59. Athauda, S.B., Matsumoto, K., Rajapakshe, S., Kuribayashi, M., Kojima, M., Kubomura-Yoshida, N., Iwamatsu, A., Shibata, C., Inoue, H., and Takahashi, K. (2004). Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. Biochemical Journal *381*, 295-306.
60. Fu, C.-N., Wicke, S., Zhu, A.-D., Li, D.-Z., and Gao, L.-M. (2023). Distinctive plastome evolution in carnivorous angiosperms. BMC Plant Biology *23*, 660.
61. Lin, C.S., Chen, J.J., Chiu, C.C., Hsiao, H.C., Yang, C.J., Jin, X.H., Leebens-Mack, J., de Pamphilis, C.W., Huang, Y.T., and Yang, L.H. (2017). Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. The Plant Journal *90*, 994-1006.
62. Wakatake, T., and Fukushima, K. (2025). Transcriptomic prey-capture responses in convergently evolved carnivorous pitcher plants. bioRxiv, 2025.2007. 2016.665073.
63. Scherzer, S., Böhm, J., Krol, E., Shabala, L., Kreuzer, I., Larisch, C., Bemm, F., Al-Rasheid, K.A., Shabala, S., and Rennenberg, H. (2015). Calcium sensor kinase activates potassium uptake systems in gland cells of Venus flytraps. Proceedings of the National Academy of Sciences *112*, 7309-7314.
64. Bouckaert, R.R. (2010). DensiTree: making sense of sets of phylogenetic trees. Bioinformatics *26*, 1372-1373. 10.1093/bioinformatics/btq110.