

ORIGINAL ARTICLE

The ancestral karyotype of the Heliantheae Alliance, herbicide resistance, and human allergens: Insights from the genomes of common and giant ragweed

Martin Laforest¹ | Sara L. Martin² | Katherine Bisailon¹ | Brahim Soufiane¹ |
Sydney Meloche³ | François J. Tardif⁴ | Eric Page³ 

¹ Saint-Jean-sur-Richelieu Research and Development Centre, Agriculture and Agri-Food Canada, Saint-Jean-sur-Richelieu, Quebec, Canada

² Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada

³ Harrow Research and Development Centre, Agriculture and Agri-Food Canada, Harrow, Ontario, Canada

⁴ Department of Plant Agriculture, University of Guelph, Guelph, Ontario, Canada

Correspondence

Eric Page, Harrow Research and Development Centre, Agriculture and Agri-Food Canada, 2585 County Rd 20, Harrow, ON N0R 1G0, Canada.
Email: eric.page@agr.gc.ca

Assigned to Associate Editor Katrien Devos.

Funding information

Agriculture and Agri-Food Canada, Grant/Award Numbers: J-001751, J-002275

Abstract

Ambrosia artemisiifolia and *Ambrosia trifida* (Asteraceae) are important pest species and the two greatest sources of aeroallergens globally. Here, we took advantage of a hybrid to simplify genome assembly and present chromosome-level assemblies for both species. These assemblies show high levels of completeness with Benchmarking Universal Single-Copy Ortholog (BUSCO) scores of 94.5% for *A. artemisiifolia* and 96.1% for *A. trifida* and long terminal repeat (LTR) Assembly Index values of 26.6 and 23.6, respectively. The genomes were annotated using RNA data identifying 41,642 genes in *A. artemisiifolia* and 50,203 in *A. trifida*. More than half of the genome is composed of repetitive elements, with 62% in *A. artemisiifolia* and 69% in *A. trifida*. Single copies of herbicide resistance-associated genes *PPX2L*, *HPPD*, and *ALS* were found, while two copies of the *EPSPS* gene were identified; this latter observation may reveal a possible mechanism of resistance to the herbicide glyphosate. Ten of the 12 main allergenicity genes were also localized, some forming clusters with several copies, especially in *A. artemisiifolia*. The evolution of genome structure has differed among these two species. The genome of *A. trifida* has undergone greater rearrangement, possibly the result of chromoplexy. In contrast, the genome of *A. artemisiifolia* retains a structure that makes the allotetraploidization of the most recent common ancestor of the Heliantheae Alliance the clearest feature of its genome. When compared to other Heliantheae Alliance species, this allowed us to reconstruct the common ancestor's karyotype—a key step for furthering of our understanding of the evolution and diversification of this economically and allergenically important group.

Abbreviations: ALS, acetolactate synthase; BUSCO, Benchmarking Universal Single-Copy Ortholog; EPSPS, enolpyruvylshikimate-3-phosphate synthase; HPPD, 4-hydroxyphenylpyruvate dioxygenase; Hi-C, chromatin conformation capture; LAI, LTR Assembly Index; LTR, long terminal repeat; MRCA, most recent common ancestor; MYA, million years ago; NCF, nested chromosome fusion; P106S, proline-to-serine mutation at position 106 of EPSPS; PPX, protoporphyrinogen oxidase; TE, transposable elements.

Martin Laforest and Sara L. Martin contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 His Majesty the King in Right of Canada. *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America. Reproduced with the permission of the Minister of Agriculture and Agri-Food Canada.

Plain Language Summary

Common and giant ragweed are important weeds with agricultural and health impacts worldwide. Here, we used a hybrid of common and giant ragweed to simultaneously produce chromosome-level assemblies of both species. These assemblies uncovered new molecular mechanisms of herbicide resistance and indicated that many of the genes believed to be responsible for the allergenicity of ragweed pollen occur in clusters of several copies. Finally, we used the genome assemblies of these and related species to reconstruct the genome of the most recent common allopolyploid ancestor.

1 | INTRODUCTION

Ambrosia artemisiifolia L. (common ragweed) and *Ambrosia trifida* L. (giant ragweed) are two of the most widely distributed and economically damaging pest species in the world (Cowbrough et al., 2003; Harrison et al., 2001; Montagnani et al., 2017; Smith et al., 2013). The detrimental effects of the species are in two main areas: agriculture and human health. Across their native North and Central American range and their introduced range in Europe, Asia, Africa, and Australia (Chauvel et al., 2021; Chen, 2008; Iamónico, 2016a, 2016b; Kil et al., 2004; Smith et al., 2013), these species are important weeds of agriculture that cause significant yield losses when left uncontrolled. Pollen production from *Ambrosia* spp. is directly linked to seasonal allergies that affect millions of people each year (D'Amato et al., 2007). For example, if *A. artemisiifolia* was to be controlled throughout its invasive European range, Schaffner et al. (2020) estimate that approximately 2.3 million fewer patients would seek assistance from public health services and that this in turn would save close to 1.1 billion Euros annually (~1.45 billion USD for the same period). Climate change is expected to expand the species range and lengthen this pollen production period, with recent evidence suggesting that, in North America, the *Ambrosia* spp. pollen season north of the 44th parallel has already expanded by as much as 13–27 days since 1995 (Case & Stinson, 2018; Wayne et al., 2002; Ziska & Beggs, 2012; Ziska et al., 2011). The combined effect of additional costs for agriculture and human health is estimated at 1 billion yuan (155 M USD) and 1.45 billion yuan (~224 M USD), respectively, for China in 2011 (Xian et al., 2023). Herbicides have been extensively used to control *Ambrosia* spp. for both agricultural and public health reasons and have generally been effective. However, widespread and frequent herbicide use coupled with the inherent high genetic variability within these two species has resulted in the evolution of resistance to many commonly used herbicides, including inhibitors of acetolactate synthase (ALS) (Taylor et al., 2002; Van Wely et al., 2015) and enolpyruvylshikimate-3-phosphate synthase (EPSPS), the molecular target site for glyphosate. While first

reported in the US Midwest and southern United States in the early 2000s, glyphosate resistant biotypes of *A. artemisiifolia* and *A. trifida* have subsequently been reported in Canada (Van Wely et al., 2015; Vink et al., 2012), not long after the introductions of glyphosate-tolerant corn (*Zea mays* L.) and soybean [*Glycine max* (L.) Merr.] cropping systems (Duke, 2018b). This development, along with historic changes in cropping practices (Page & Nurse, 2015; Regnier et al., 2016), has exacerbated the difficulties in managing these problematic weeds. The availability of genomic resources is key to understanding the genetic basis of herbicide resistance and allergenicity in these species and enable the development of management solutions to mitigate both agricultural and health impacts.

Ambrosia artemisiifolia and *A. trifida* are two of approximately 40 species in the genus *Ambrosia*, which has a center of diversity in North America and includes wind-pollinated, annual or perennial weeds and woody desert shrubs (Chen et al., 2018; Martin et al., 2018). The species are estimated as having diverged from each other approximately 5.3 million years ago (MYA) (Kumar et al., 2022). *Ambrosia* is part of the Ambrosiinae subtribe, and the Heliantheae tribe within Asteraceae. As one might intuit, this tribe also includes the more charismatic species within *Helianthus* (sunflowers) and with related tribes, such as the Eupatorieae, forms part of the Heliantheae Alliance. Members of this alliance share a whole-genome duplication event that occurred approximately 23–28 MYA at its crown node (Barker et al., 2008; Fan et al., 2022; Mandel et al., 2019) with a transition from $2n = 18$ to $2n = 36$ (Semple & Watanabe, 2009). The mechanism of this transition, whether it was the result of autopolyploidization (the duplication of genomes from within a lineage) or allopolyploidization (the result of hybridization and duplication), has not been determined. *Ambrosia artemisiifolia* ($2n = 36$) and *A. trifida* ($2n = 24$) (Mulligan, 1957; Payne et al., 1964) share this polyploidization event with each other and with the approximately 5500 species in the Heliantheae Alliance (Baldwin, 2009; Dvorak et al., 2018; Mandel et al., 2019; Semple & Watanabe, 2009) (Figure 1). There are currently three other genomes sequenced to

chromosome level with available annotations from the Heliantheae Alliance: *Helianthus annuus* L. (sunflower; $2n = 34$), *Mikania micrantha* Kunth (mile-a-minute weed; $2n = 38$) (Liu et al., 2020), and *Smallanthus sonchifolius* (Poepp.) H. Rob. (yacon; $2n = 58$) (Fan et al., 2022). These genomes, along with a representative of the lineage prior to the polyploidization event, *Conyza canadensis* (L.) Cronquist, ($2n = 18$) can provide insight into the karyotype of the ancestor, allowing inference of its structure. In turn, the ancestral karyotype can clarify how these genomes have evolved and has been a powerful tool for understanding genome evolution in the Brassicaceae (Lysak et al., 2016; Schranz et al., 2006), Poaceae (Murat et al., 2010), *Cucumis* (Zhao et al., 2021), and, more broadly, flowering plants (Murat et al., 2017). The nuclear content of *A. artemisiifolia* and *A. trifida* has been evaluated as approximately 1.2 and 1.9 pg (1C-value), respectively. These numbers correspond to a haploid genome size of 1.1 gigabase pairs (Gbp) and 1.8 Gbp for *A. artemisiifolia* and *A. trifida* (Bai et al., 2012; Kubesova et al., 2010).

At present, there is no publicly available genome assembly for *A. trifida*, while a short-read assembly for *A. artemisiifolia* was recently published (Bieker et al., 2022). Although an important step forward, this assembly remains highly fragmented, consisting of 12,228 ordered contigs with an N50 of 270.6 kbp. More recently, Battlay et al. (2023) reported a chromosomal-level assembly with an N50 of 58Mb; however, at the time of writing, the data had not been made publicly available for comparison. Producing genome assemblies for the members of the Asteraceae, such as *H. annuus* (sunflower), has proven to be a challenging endeavor as they often consist of long and highly similar repeats (Badouin et al., 2017; Cerca et al., 2022; Nakano et al., 2021). Additional challenges are also expected when assembling the genomes of outcrossing species with high levels of heterozygosity as this variation complicates haplotype phasing and the collapse of homologous regions of the genome. To address these issues, genome sequencing projects have employed various strategies to avoid heterozygosity including sequencing of inbred lines, haploids, and megagametophytes (Lin et al., 2021; Zimin et al., 2014). Another approach is to take advantage of interspecies hybrids to resolve haplotypes. The software CANU can sort sequencing reads of an F1 progeny provided that data from both parents are provided separately (Koren et al., 2018). In this method, sequences from the parents are used to sort reads from the F1 hybrid and resolve haplotypes, as exemplified in a recent paper describing haplotype sequencing of the wood tiger moth (*Arctia plantaginis* L.) genome (Yen et al., 2020).

Although uncommon, *A. artemisiifolia* and *A. trifida* have been observed to hybridize and the hybrid, *Ambrosia* × *helenae* Rouleau., is noted in several floras (Reznicek et al., 2011; Strother, 1993; Wylie, 1915). The first report of a hybrid between these species (hereafter referred to as hybrid

Core Ideas

- Production of high-quality chromosome-level assemblies for *Ambrosia artemisiifolia* and *Ambrosia trifida* is reported.
- Trio binning of parental haploid genomes from within hybrid ragweed (*Ambrosia* × *helenae* Rouleau) was used for the assembly of both parental haplotypes.
- Target site mutation conferring resistance to glyphosate in *A. artemisiifolia* and *A. trifida* were identified.
- The karyotype of most recent common ancestor of the Heliantheae Alliance—an allopolyploid—was reconstructed.
- The genome evolution in the Heliantheae Alliance has involved nested chromosome fusions and possibly chromoplexy.

ragweed) was made by Wylie (1915). The author reports observing an unusual specimen of *Ambrosia*, with long marginless petioles, leaves palmately divided in three, and segments that were pinnatifid (Figure 2). They note that “from another point of view the leaf might be looked upon as a pinnate leaf with greatly enlarged and opposite basilar segments, remotely suggestive of *Ambrosia artemisiifolia*.” Hybrid ragweed has subsequently been reported by Jones (1943), Wagner (1958), Vincent and Cappadocia (1987), and Vollenberg et al. (2005). Wagner (1958) provides perhaps the most in-depth description of hybrid ragweed, including a summary of its observation from Michigan and Iowa to Quebec and France, as well as a plate of leaf silhouettes from cotyledon to highest bract, which are extremely instructive for its identification in the field. Both Jones (1943) and Vincent and Cappadocia (1987) indicate that the number of chromosomes in hybrid ragweed is 30, 18 from *A. artemisiifolia* and 12 from *A. trifida*.

The current paper describes the approach undertaken to sequence haploid genomes of both species where parents as well as the hybrid were sequenced using Pacific Bioscience long-read technology (GenBank PRJNA967341). The genomes of the parental species were assembled individually and these results served as reference to sort the sequence reads of the hybrids into two groups, referring to the origin of the chromosomes and therefore permitting the resolution of haplotypes for both species. We then annotated these genomes, investigated the location and copy number of genes associated with herbicide resistance and allergens, and reconstructed the ancestral karyotype of the Heliantheae Alliance’s common ancestor. This work provides the first chromosome-scale, publicly available annotated genome sequences for *A.*

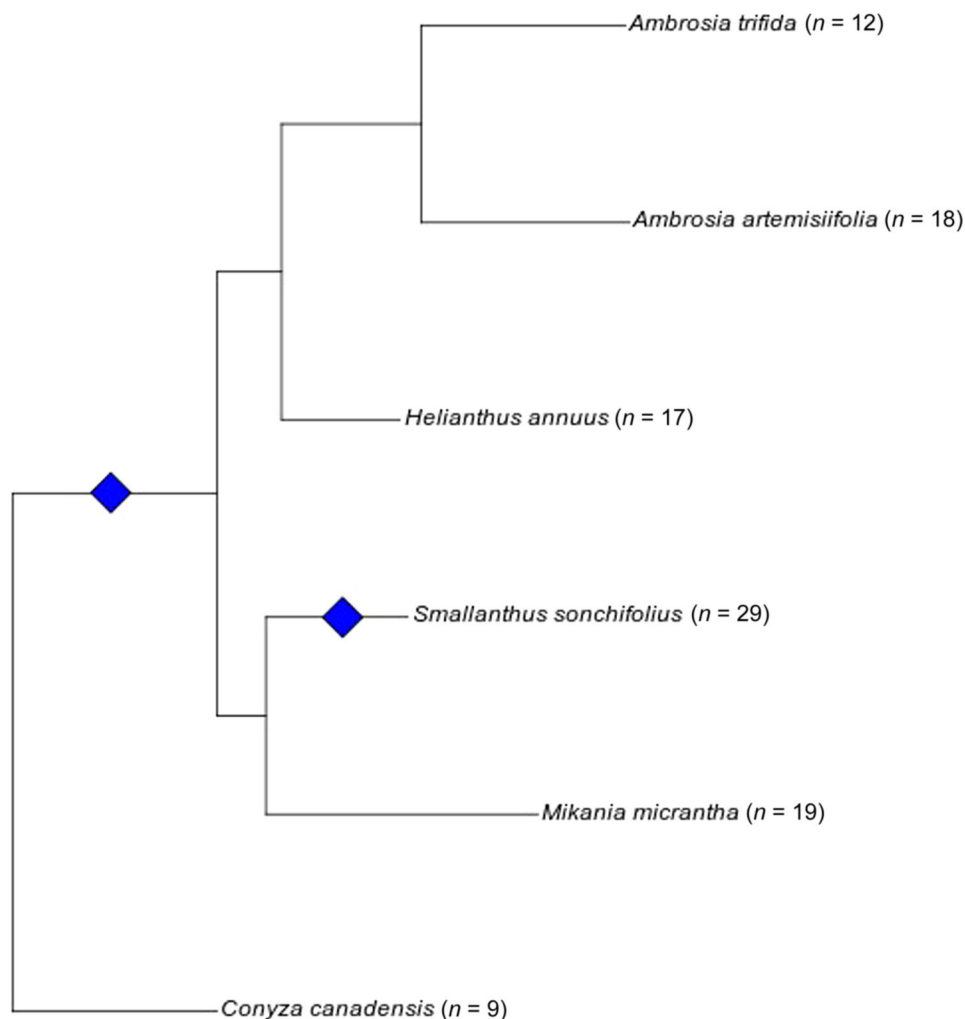


FIGURE 1 Evolutionary relationships among genomes examined in this paper with polyploidization events marked with blue diamonds including the polyploidization event shared by the *Helianthus* Alliance and a second event experienced by *Smallanthus sonchifolius*.

artemisiifolia and *A. trifida*, species of worldwide importance from both agronomic and human health perspectives.

2 | MATERIALS AND METHODS

2.1 | Plant material

In 2020, two interspecific hybrids were observed among the progeny of a spontaneous *A. artemisiifolia* ♂ × *A. trifida* ♀ cross. A single *A. trifida* individual arising from three generations of self-fertilization was co-located in a greenhouse compartment with six glyphosate-resistant *A. artemisiifolia* plants. The original sources for the parental *A. trifida* biotype were from a collection made on the grounds of the Harrow Research and Development Centre (42.035°N, 82.902°W) in 2015. For *A. artemisiifolia*, the parental material for the hybrid arose from a cross between a susceptible biotype collected from a soybean field in Ridgetown, Ontario (42.27°N,

81.52°W) and a glyphosate resistant biotype from a farm in Essex County, Ontario (42.28°N, 82.84°W). The two interspecific hybrids were observed in successive propagations of progeny from the same maternal *A. trifida* individual but separated by a period of several months. The first of the two interspecific hybrids was used to verify DNA content relative to the parental species via flow cytometry, as described below. Tissue from the second hybrid individual was used for subsequent genome sequencing, as well as the *A. artemisiifolia* and *A. trifida* parental biotypes mentioned above. Material from this hybrid has been deposited in the Department of Agriculture, Ottawa (DAO) herbarium with accession number DAO-01-01709745.

2.2 | Flow cytometry

Tissue for flow cytometry analysis was collected from fresh, young leaves of *A. artemisiifolia*, *A. trifida*, and hybrid



FIGURE 2 Representative leaves from *Ambrosia artemisiifolia* (top right), *Ambrosia trifida* (bottom right), and the hybrid *Ambrosia* × *helenae* (middle left).

ragweed and placed on ice. In addition to this sample, leaves from *Camelina sativa* (L.) Crantz. or *Raphanus sativa* L. were also collected as an internal standard with a known 2C DNA contents of 1.59 pg (Martin et al., 2017) and 1.11 pg (Doležel & Bartos, 2005), respectively. Leaf samples from each of *A. artemisiifolia*, *A. trifida*, and their hybrid were tested with this internal standard on three separate days as recommended for precise DNA estimates. Sample and standard leaf tissue were chopped in Galbraith buffer (0.7 mL) with a razor blade and kept in the dark while they stained with propidium iodide for 30–40 min (Doležel et al., 1992). Samples were run at low speed on a Gallios flow cytometer (Beckman Coulter). The relative DNA content was determined using fluorescence area (585/42 nm detector), and all characteristics of the peaks including mean, coefficients of variation, and nuclei numbers were determined using the R package flowPloidy (Smith et al., 2018). DNA content in picograms was converted to base pairs using the relationship that 1 pg equals 978 Mbp (Doležel et al., 2003).

2.3 | DNA extraction

Genomic DNA was extracted using QIAGEN Genomic-tip 20/G and Genomic DNA Buffer Set (Qiagen Inc.) according to a user-developed protocol. Isolation of genomic DNA from plants and filamentous fungi was performed using the QIAGEN Genomic-tip – (EN) (<https://www.qiagen.com/ca/resources/resourcedetail?id=cb2ac658-8d66-43f0-968e-7bb0ea2c402a&lang=en>). The enzyme for digesting cell wall material (step 2 of Qiagen protocol) selected was Hemicellulase from *Aspergillus niger* (Millipore Sigma) (for 2 h at 37°C). DNA final resuspension was in 75 µL of transposable elements (TE).

2.4 | Genome sequencing, assembly, and annotation

The DNA library was prepared following the Pacific Biosciences *Preparing gDNA Libraries Using the SMRTbell Express Template Preparation kit 2.0 protocol* (PN# 100-693-800). A total of 15.0 µg of high-molecular-weight genomic DNA (final volume of 100 µL) was sheared using the Covaris g-TUBES (Covaris Inc.) at 4000 RPM for 120 s on each side (total six passes), on an Eppendorf centrifuge 5424 (Eppendorf). The DNA damage repair, end repair, SMRT bell ligation, and nuclease treatment steps were performed as described in the template preparation protocol with the SMRTbell Template Prep Kit 2.0 reagents (Pacific Biosciences) and a DNA input of 10 µg. The DNA library was size selected on a BluePippin system (Sage Science Inc.) using a cutoff range of 15–50 kbp. The sequencing primer was annealed with sequencing primer v2 at a final concentration of 1 nM, and the Sequel II 2.0 polymerase was bound at 0.5 nM. The library went through an AMPure bead cleanup (following the SMRTlink v8 calculator procedure) before being sequenced on a PacBio Sequel II instrument at a loading concentration (on-plate) of 120 pM using the diffusion loading protocol, Sequel II Sequencing Kit 2.0, SMRT Cell 8 M, and 30-h movies with 4-h pre-extension time. The data processing was done using the SMRT Link Analysis v8.0.0.80501 pipeline at the Genome Quebec Innovation Center (Montreal, Canada), producing 866 Gbp of data. For *A. artemisiifolia*, a further 41 Gbp of Oxford Nanopore sequencing data were produced. For the latter dataset, libraries were prepared with the Rapid Sequencing kit (SKU: SQK-RAD004; Oxford Nanopore Technologies) and sequenced on MK1 R9 flow cells (Oxford Nanopore Technologies) following manufacturer guidelines.

For *A. trifida*, PacBio sequence assembly was performed using CANU (Koren et al., 2018) with the following parameters: `genomesize=1800m correctedErrorRate=0.085 corMhapSensitivity=normal minReadLength=15000 corMhapFilterThreshold=0.000000002 corMhapOptions="--threshold 0.80 --num-hashes 256 --num-min-matches 3 --ordered-sketch-size 1000 --ordered-kmer-size 14 --min-olap-length 2000 --repeat-idf-scale 50"`. For *A. artemisiifolia*, the CANU parameters were: `genomesize=1100m correctedErrorRate=0.085 corMhapSensitivity=normal minReadLength=7000 corMhapFilterThreshold=0.000000002 corMhapOptions="--threshold 0.80 --num-hashes 256 --num-min-matches 3 --ordered-sketch-size 1000 --ordered-kmer-size 14 --min-olap-length 2000 --repeat-idf-scale 50"`. The trimmed reads of these two assemblies were used to split the hybrid reads according the trio binning procedure running CANU with the following parameters: `genomesize=1800m correctedErrorRate=0.085 corMhapSensitivity=normal minReadLength=12500 corMhapFilterThreshold=0.000000002 corMhapOptions="--threshold 0.80 --num-hashes 256 --num-min-matches 3 --ordered-sketch-size 1000 --ordered-kmer-size 14 --min-olap-length 2000 --repeat-idf-scale 50"`.

Chromatin conformation capture (Hi-C) data were generated for the hybrid using a Phase Genomics Proximo Hi-C 2.0 Kit, which is a commercially available version of the Hi-C protocol (Lieberman-Aiden et al., 2009). Following the manufacturer's instructions for the kit, intact cells from the hybrid plant were crosslinked using a formaldehyde solution, digested using the restriction enzymes *DpnII*, *DdeI*, *HinF*, and *MseI* with cut sites at GATC, CTNAG, GANTC, and TTAA, respectively, and proximity ligated with biotinylated nucleotides to create chimeric molecules composed of fragments from different regions of the genome that were physically near in vivo, but not necessarily genomically contiguous. Continuing with the manufacturer's protocol, molecules were pulled down with streptavidin beads and processed into an Illumina-compatible sequencing library. Sequencing was performed on an Illumina HiSeq 4000, generating a total of 31,561,366 paired-end (PE) 150-bp reads. Reads were aligned to the PacBio/CANU assembly also following the manufacturer's recommendations (<https://phasegenomics.github.io/2019/09/19/hic-alignment-and-qc.html>). Briefly, reads were aligned using BWA-MEM (Li & Durbin, 2010) with the `-5SP` and `-t 8` options specified, and all other options default. SAMBLASTER (Faust & Hall, 2014) was used to flag PCR duplicates, which were later excluded from analysis. Alignments were then filtered with samtools (Li et al., 2009) using the `-F 2304` filtering flag to remove nonprimary and secondary alignments. Phase Genomics' Proximo Hi-C genome scaffolding platform was used to create chromosome-scale scaffolds from the draft assembly as described in Bickhart et al. (2017). As in the LACHESIS method (Burton et al.,

2013), this process computes a contact frequency matrix from the aligned Hi-C read pairs, normalized by the number of *Sau3AI* restriction sites (GATC) on each contig, and constructs scaffolds in such a way as to optimize expected contact frequency and other statistical patterns in Hi-C data. Three separate Proximo runs were performed to optimize the number of scaffolds and scaffold construction in order to make the scaffolds as concordant with the observed Hi-C data as possible. Juicebox (Durand et al., 2016; Rao et al., 2014) was used to correct scaffolding errors, resulting in a final set of 18 and 12 chromosome-scale scaffolds, for common ragweed and giant ragweed, respectively.

The completeness, quality, and level of repetitiveness of the sequence assembly were evaluated with BUSCOs (Benchmarking Universal Single-Copy Orthologs [Waterhouse et al., 2017]) (version 5.0.0) against the database eudicots_odb10. Repeated sequences were identified with RepeatModeler (Smit & Hubley, 2008–2015) and the genome was masked using RepeatMasker (Smit et al., 2013–2015) using default values. The completeness of the genomes based on repetitive elements was examined using the long terminal repeat (LTR) Assembly Index (LAI) (Ou et al., 2018), which was generated using LTRHarvest (Ellinghaus et al., 2008), LTR_Finder (Xu & Wang, 2007), and LTR_retriever (Ou & Jiang, 2018). The Extensive de novo TE annotator (EDTA) was used to form a list of intact elements (Ou et al., 2019). AUGUSTUS (Keller et al., 2011) was used for ab initio gene prediction before and after repeat masking (RepeatMasker version 4.1.2-p1, default mode, run with rmblastn version 2.11.0+). AUGUSTUS hints for in silico gene discovery were deduced from RNAseq datasets (GenBank PRJNA1020868). Potential proteins identified by AUGUSTUS were annotated using InterProScan (Jones et al., 2014) version 5.57-90.0 with the `-goterms` argument.

2.5 | Genes of interest

The genes identified as potential allergens (Bordas-Le Floch, Groeme, et al., 2015; Ghosh et al., 1993; Mole et al., 1975) were mapped onto the genome sequence using BLAST and BLASTp (Chen et al., 2018). The identified proteins have different functions and are of different molecular weight ranging from 5 to 48 kDa (Table S1). GenBank accessions MH220501 (Amb a 1), MK415805 (Amb a 3), FN687761 (Amb a 4), M84987 (Amb a 5), U89793 (Amb a 6), MH220504 (Amb a 8), AY894657 (Amb a 9), AY894659 (Amb a 10), MH220502 (Amb a 11), and KU593514 (Amb a 12) were compared to *A. artemisiifolia* and *A. trifida* genome sequences, and coordinates from BLASTn hits with an expected value less than 1×10^{-50} were used to place genes on the genome assemblies. BLASTp hits were also screened based on at least 50% match identity to the reference protein. GenBank

accessions KX870184.1 (*A. artemisiifolia* ALS), MT537192 (*A. artemisiifolia* EPSPS), JN966736 (*A. artemisiifolia* protoporphyrinogen oxidase [PPX2]), and MK096764 (*A. artemisiifolia* 4-hydroxyphenylpyruvate dioxygenase [HPPD]) were similarly used to map candidate herbicide-resistant genes.

2.6 | Comparative genomics

The genomes and annotation files for *H. annuus* ($2n = 34$) (sunflower) (Badouin et al., 2017) (HanXRQr2.0-SUNRISE), *M. micrantha* ($2n = 38$) (mile-a-minute weed) (ASM936387v1) (Liu et al., 2020), and *S. sonchifolius* (yacon) ($2n = 58$) (ASM2352597v1) (Fan et al., 2022) were downloaded from the National Library of Medicines NCBI database (<https://www.ncbi.nlm.nih.gov/data-hub/genome/>). These genomes and that of *C. canadensis* (Laforest et al., 2020) were compared with the *Ambrosia* genomes using tools from Mummer v4.0 (Marçais et al., 2018) and the Whole-Genome Duplication Integrated (WGDI) analysis tool kit (Sun et al., 2021), which relies on synteny determined by MCScanX (Wang et al., 2012). In order to isolate synteny resulting from the most recent duplication for the reconstruction of an ancestral karyotype analog, the block files produced by WGDI were filtered based on K_s scores, with a p -value of less than 0.075 and block length of at least 20; otherwise, default settings were used. Reconstruction then followed the steps outlined in the WGDI manual, though selecting the chromosomes most likely to be representative of the ancestral karyotype and then inverting the gene order of sections based on inversions shared by the majority of species. The positions of centromeres were tentatively estimated based on gene density, with the lowest point of gene density being chosen as a likely position. This point was determined by using R's density function to calculate the density of genes in the gff file, excluding the first and last 4% of the chromosome, and determining the position of the minimum. Biased retention of genes from the two ancestral genomes was investigated by first dividing species genomic information into subgenomes based on the ancestral karyotype and running the WGDI pipeline determine retention compared with *C. canadensis* with the same filtering parameters as above. We then examined the subset of genes where only one homolog to *C. canadensis*'s genes was found in the species' genome and determined if it was more likely that the retained gene was from one ancestral genome or the other with a proportion test. In all cases, R (R Core Team, 2021) v4.1.1 "Kick Things" was used to convert data into appropriate formats, visualize results, and perform analyses. The following R packages were used in addition to the base packages: ape (Paradis & Schliep, 2019), apex (Schliep et al., 2020), Biostrings (Pagès et al., 2021), ggplot2 (Wickham, 2016), Iragenes (Lawrence

et al., 2013), msa (Bodenhofer et al., 2015), phangorn v2.10.0 (Schliep, 2011; Schliep et al., 2017), plyr (Wickham, 2011), Rsamtools (Morgan et al., 2021), seqinr (Charif & Lobry, 2007), dplyr (Wickham et al., 2022), berryFunctions (Boessenkool, 2023), diagram (Soetaert, 2020), pBrackets (Schulz, 2021), and scales (Wickham & Seidel, 2022). Plotting of synteny between genomes was completed with the R package GENESPACE (Lovell, 2023; Lovell et al., 2022), which relies on Orthofinder (Emms & Kelly, 2019). The phylogenetic tree produced by Orthofinder was converted to an R-friendly format using figtree (v1.4.4) (Rambaut, 2006–2018). Orthofinder (v2.55) was also used to evaluate the relationship among genes of interest across different taxa.

3 | RESULTS

3.1 | Genome size estimation

The average total nuclear DNA contents (2C) for *A. artemisiifolia* and *A. trifida* were 2.17 (± 0.36) pg and 3.44 (± 0.08) pg, respectively, giving estimated haploid genome sizes of 1.06 and 1.68 Gbp (Table 1). The hybrid between the species had a total nuclear DNA content of 2.71 (± 0.04) pg and estimated haploid genome size of 1.32 Gbp.

3.2 | Assembly and annotation of the genomes

Sequence data for the parental biotypes of *A. artemisiifolia* and *A. trifida* were first assembled individually yielding 18,851 and 38,288 contigs, respectively (Table 2). The trimmed reads produced during the assembly process were used to split the data resulting from sequencing the hybrid. The individual assemblies of the data for each parent species of the hybrid, representing a haploid genome, yielded more than four times fewer contigs (4656 and 9408 for *A. artemisiifolia* and *A. trifida*, respectively, data not shown). These assemblies were used for the Hi-C scaffolding protocol that further reduced the number of contigs to 3521 and 7530. More importantly, however, were the gains made with Hi-C in terms of the length of the contigs and the number needed to cover 75% of the genome sequence. For *A. artemisiifolia*, the N75 value increased from 122 kilobases (kbp) to more than 41 megabase (Mbp), and the L75 value dropped from 8297 to 16, which is smaller than the number of chromosomes in this species. In the case of giant ragweed, the N75 value increased from 51 kbp to more than 122 Mbp, and the L75 value dropped from 20,805 to 12, the haploid chromosome number in this species. Final assembly size was 1.130 Gbp for *A. artemisiifolia* and 2.028 Gbp for *A. trifida*, while guanine-cytosine content was similar for both species at approximately 37%.

TABLE 1 Flow cytometry estimates of DNA content for the lines sequenced here, common ragweed (*Ambrosia artemisiifolia*), giant ragweed (*Ambrosia trifida*), and their hybrid.

Taxon	Number of individuals	2C DNA estimate (pg)	2C DNA stdev (pg)	Haploid genome size (Kbp)
<i>Ambrosia trifida</i>	3	3.44	0.08	1682
<i>Ambrosia artemisiifolia</i>	3	2.17	0.26	1061
<i>Ambrosia</i> × <i>helenae</i>	1	2.71	0.04	1325

Note: Samples from individuals were run at least three times across three different days with either *Raphanus sativa* or *Camelina sativa* as an internal standard.

TABLE 2 Quantity of data produced in this project and assembly statistics.

Species	<i>Ambrosia artemisiifolia</i>	<i>Ambrosia trifida</i>	Hybrid	<i>Ambrosia artemisiifolia</i> (Bieker et al., 2022)	
PacBio	194,273,509,376	291,015,016,448	380,813,787,136		
Oxford Nanopore	41,162,363,750	–	–		
Total	235,435,873,126	291,015,016,448	380,813,787,136		
Expected genome size	1134.5–1163 Gbp	1837–1865.5 Gbp	~2935 Gbp		
Coverage	202.4–207.5x	156.2–158.4x	129.7x		
			<i>A. artemisiifolia</i> from hybrid	<i>A. trifida</i> from hybrid	
Assembly size	1,678,677,417	1,734,447,773	1,129,637,797	2,027,844,680	1,258,364,641
Number of contigs (>1000 bp)	18,851	38,288	3521	7530	12,224
N50	122,960	51,135	51,200,794	122,405,267	270,627
L50	3400	10,225	10	8	1355
N75	62,417	33,383	41,317,028	107,573,200	135,390
L75	8297	20,805	16	12	2971
% guanine-cytosine	37.08%	37.24%	37.12%	37.04%	36.25%
Number of N per 100 kbp	–	–	10.48	9.25	29,122

Note: Numbers indicated base pairs, unless mentioned otherwise. Gbp = gigabase pairs, x = fold coverage. L50 and L75 values expressed in number of contigs.

The post-scaffolding Hi-C heat map indicates that the highest density of matching reads that can be found was in contigs in close proximity, within the same scaffold (Figure S1).

To assess the completeness of the assembled draft sequence, a BUSCO analysis was performed, revealing that 94.5% of the genes surveyed were complete, with 74.2% as single copy, 20.3% duplicated, 1.5% fragmented, and 4.1% missing (Table 3) for *A. artemisiifolia*. Similarly, 96.1% of the *A. trifida* genes surveyed were complete, with 50.6% as single copy, 45.5% duplicated, 0.9% fragmented, and 3% missing. The LAI, calculated as the ratio between the length of intact LTR retrotransposons and the total length of retrotransposons in the assembly, indicated that the repeatome of these genomes is highly complete with indexes of 26.6 and 23.6, respectively, and that both genomes would be considered of gold quality (≥ 20) (Ou et al., 2018). RepeatModeler identified a high prevalence of repetitive elements in both genomes. Nearly 62% (699,505,010 bp) and 69% (1,397,540,865 bp)

of *A. artemisiifolia* and *A. trifida* genomes, respectively, are made of repetitive elements (Table 4). The repeated elements with the highest copy number were LTRs of the *Copia*-type at 16.76% and 15.05%, followed by *Gypsy*-like transposons at 7.95% and 15.50%, for *A. artemisiifolia* and *A. trifida*, respectively. L1-like LINE (a transposable element with homology to a type of class 1 long interspersed retrotransposable element) elements were another repeated sequence occurring at a frequency higher than 1%, found at 2.02% and 2.08%, respectively, for *A. artemisiifolia* and *A. trifida*. DNA transposons also represented a high proportion of the repeats at 3.05% and 3.54% of the assemblies. A large proportion (30.64% and 31.69% of the respective genomes of the species sequenced) of the identified repeated elements were not classified and were categorized as unknown. Helitron density is particularly high in both *A. artemisiifolia* and *A. trifida*, with values of 24.75 and 20.52 helitrons per megabase of sequence. For comparisons, helitron densities are 9.77, 6.48,

TABLE 3 Benchmarking Universal Single-Copy Ortholog (BUSCO) analysis for the de novo *Ambrosia artemisiifolia* and *Ambrosia trifida* genome assemblies.

	<i>Ambrosia artemisiifolia</i>		<i>Ambrosia artemisiifolia</i> (Bieker et al. 2022)		<i>Ambrosia trifida</i>	
Complete	2197	94.5%	2073	89.1%	2236	96.1%
Complete and single-copy	1725	74.2%	1360	58.5%	1177	50.6%
Complete and duplicated	472	20.3%	713	30.7%	1059	45.5%
Fragmented	34	1.5%	98	4.2%	20	0.9%
Missing	95	4.1%	155	6.7%	70	3.0%
Total groups searched	2326		2326		2326	

Note: The genome sequence for *A. artemisiifolia* of Bieker et al. (2022) is presented for comparison. The number of gene in each category and their relative importance are given for each of the three assemblies.

2.38, 2.86, 3.27, and 10.46 in *H. annuus*, *M. micrantha*, *Chrysanthemum lavandulifolium*, *Lactuca saligna*, *Lactuca sativa*, and *C. canadensis*, respectively. Helitron density is similarly variable—from 0.2368 to 26.0412 in Brassicaceae (Hu et al., 2019). In silico gene discovery, assisted with hints computed from RNAseq data, has identified 41,642 genes in *A. artemisiifolia* (Table S1) and 50,203 genes in *A. trifida* (Table S2). Gene distribution, expressed in number of genes per 100 kbp, reveals that gene content is generally higher near telomeres in both species (Figures 3, 4, and S2).

3.3 | Genes of interest for herbicide resistance and allergenicity

Ten of the 11 classes of genes (Table S3) coding for potential allergens were mapped onto the genome sequences of *A. artemisiifolia* and *A. trifida* using BLASTn and BLASTp. Using the DNA sequences for the allergens, in *A. artemisiifolia*, we see that most allergens map at least twice to the genome. For example, the Amb a 1 gene maps to both chromosome 2 and 13. Amb a 3 and Amb a 6 form clusters of five repeats of the same gene on chromosomes 8 and 14, respectively (Figure 3). Amb a 3 also forms a small cluster of two genes in *A. trifida* on chromosome 12, while three copies of Amb a 1 can be found in close vicinity on chromosome 7 (Figure 4). When we examine amino acid sequences predicted for the genomes, there are hits with at least 50% match identity with the reference protein for Amb a 1, 8, 9, 11, and 12 for all taxa (Tables 5 and S4). Homologs for Amb a 6 were found in *A. artemisiifolia*, *H. annuus*, *M. micrantha*, and *S. sonchifolius*. No homologs were detected for Amb a 4, 5, or 7 for any taxa. The proteins identified were all within orthogroups, groups of genes descended from a common gene, determined by Orthofinder with a total of 33 orthogroups represented among the top BLASTp results (Table S5). In general, the number of proteins is fewer than expectations based on genome duplication and gene count in *C. canadensis*. Two exceptions to this apparent pattern of gene

loss are (i) an increase in the number of Amb a 1 proteins in the *Ambrosia* lineage, where 10 homologs were detected in each *Ambrosia* species but only three in *C. canadensis*, and (ii) the number of Amb a 11 proteins in *H. annuus*, where we found 32 homologous proteins but only six in *C. canadensis*.

Four genes typically associated with herbicide resistance were mapped to the assembly using DNA sequence similarity. Single copies of *PPX2L*, *HPPD*, and *ALS* were found in the genomes of the two species (Figures 3 and 4), while two complete copies of the gene coding for EPSP synthase were located in each *Ambrosia* species.

3.4 | The legacy of allotetraploidy and the ancestral karyotype

The most obvious feature of *A. artemisiifolia*'s ($n = 18$) genome is that it is the result of a genome duplication event, as expected based on its evolutionary history. This is particularly clear when the genome is compared to *C. canadensis* ($n = 9$) (Figures 5A and 6) as *C. canadensis* did not share the Heliantheae Alliance's genome duplication event. Nine of *A. artemisiifolia*'s 18 chromosomes show strong synteny with *Conyza*'s chromosomes along their whole length indicating that these chromosomes have changed relatively little since the divergence of these taxa (Glover et al., 2016) (Figure 5A). For example, aa10/16 and aa05/11 are syntenic with cc05 and cc06, respectively. Additionally, as chromosome number has been retained following duplication in *A. artemisiifolia*, there are still pairs of homoeologous chromosomes (i.e., chromosomes that originated by speciation of a common ancestor, but were brought back together in the same genome) that can be picked out as having a similar structure with the same rearrangements compared with *C. canadensis*. As an example, aa08 and aa02 show the same pattern of synteny with cc01 and cc04 (Figure 5A). Divergence between the proteins coded by these homoeologs in *A. artemisiifolia* and other species of the Heliantheae Alliance—particularly *A. trifida* (Figures 6 and 7)—indicates a pattern with homoeologs showing two

TABLE 4 Repeat content analysis for the de novo assemblies of *Ambrosia artemisiifolia* and *Ambrosia trifida*.

	<i>Ambrosia artemisiifolia</i>			<i>Ambrosia trifida</i>		
	Number of elements	Length occupied	Percentage of sequence	Number of elements	Length occupied	Percentage of sequence
Retroelements	215,502	318,965,598 bp	28.24%	430,093	683,113,674 bp	33.69%
	SINES:	28,779 bp	0.00%	0	0 bp	0.00%
	Penelope	723,714 bp	0.06%	2460	309,045 bp	0.02%
	LINES:	33,601,390 bp	2.97%	91,665	58,722,694 bp	2.90%
	CRE/SLACS	6,960,297 bp	0.62%	27,121	12,519,742 bp	0.62%
	L2/CR1/Rex	52,579 bp	0.00%	0	0 bp	0.00%
	R1/LOA/Jockey	0 bp	0.00%	0	0 bp	0.00%
	R2/R4/NeSL	0 bp	0.00%	0	0 bp	0.00%
	RTE/Bov-B	2,329,849 bp	0.21%	10,421	3,644,693 bp	0.18%
	L1/CIN4	22,828,360 bp	2.02%	51,316	42,121,796 bp	2.08%
DNA	LTR elements:	285,335,429 bp	2.26%	338,428	624,390,980 bp	30.79%
	BEL/Pao	460,007 bp	0.04%	0	0 bp	0.00%
	Ty1/Copia	189,332,760 bp	16.76%	132,138	305,241,525 bp	15.05%
	Gypsy/DIRS1	89,838,540 bp	7.95%	191,343	314,214,750 bp	15.50%
	Retroviral	3,840,506 bp	0.34%	11,417	3,723,783 bp	0.18%
	Transposons	34,410,325 bp	3.05%	191,246	71,753,423 bp	3.54%
	hobo-Activator	5,413,402 bp	0.48%	47,350	16,323,791 bp	0.80%
	Tc1-IS630-Pogo	351,094 bp	0.03%	1961	428,829 bp	0.02%
	En-Spm	0 bp	0.00%	0	0 bp	0.00%
	MuDR-IS905	0 bp	0.00%	0	0 bp	0.00%
Other	PiggyBac	0 bp	0.00%	0	0 bp	0.00%
	Tourist/Harbinger	22,309	0.74%	41,079	15,014,316 bp	0.74%
	Other (Mirage, P-element, Transib)	0	0.00%	0	0 bp	0.00%
	Rolling circles	13,638	0.87%	21,118	11,886,472 bp	0.59%
	Unclassified	1,219,296	30.64%	2,025,972	642,673,768 bp	31.69%
	Total interspersed repeats	699,505,010 bp	61.92%	6502	1,397,540,865 bp	68.92%
	Small RNA	530,124 bp	0.05%	14,916	2,677,556 bp	0.13%
	Satellites	1,860,260 bp	0.16%	487,838	2,846,261 bp	0.14%
	Simple repeats	15,898,703 bp	1.41%	54,967	24,379,926 bp	1.20%
	Low complexity	1,877,703 bp	0.17%	54,967	2,776,498 bp	0.14%

Abbreviation: LINES, long interspersed retrotransposable elements; LTR, long terminal repeat; SINES, short interspersed retrotransposable elements.

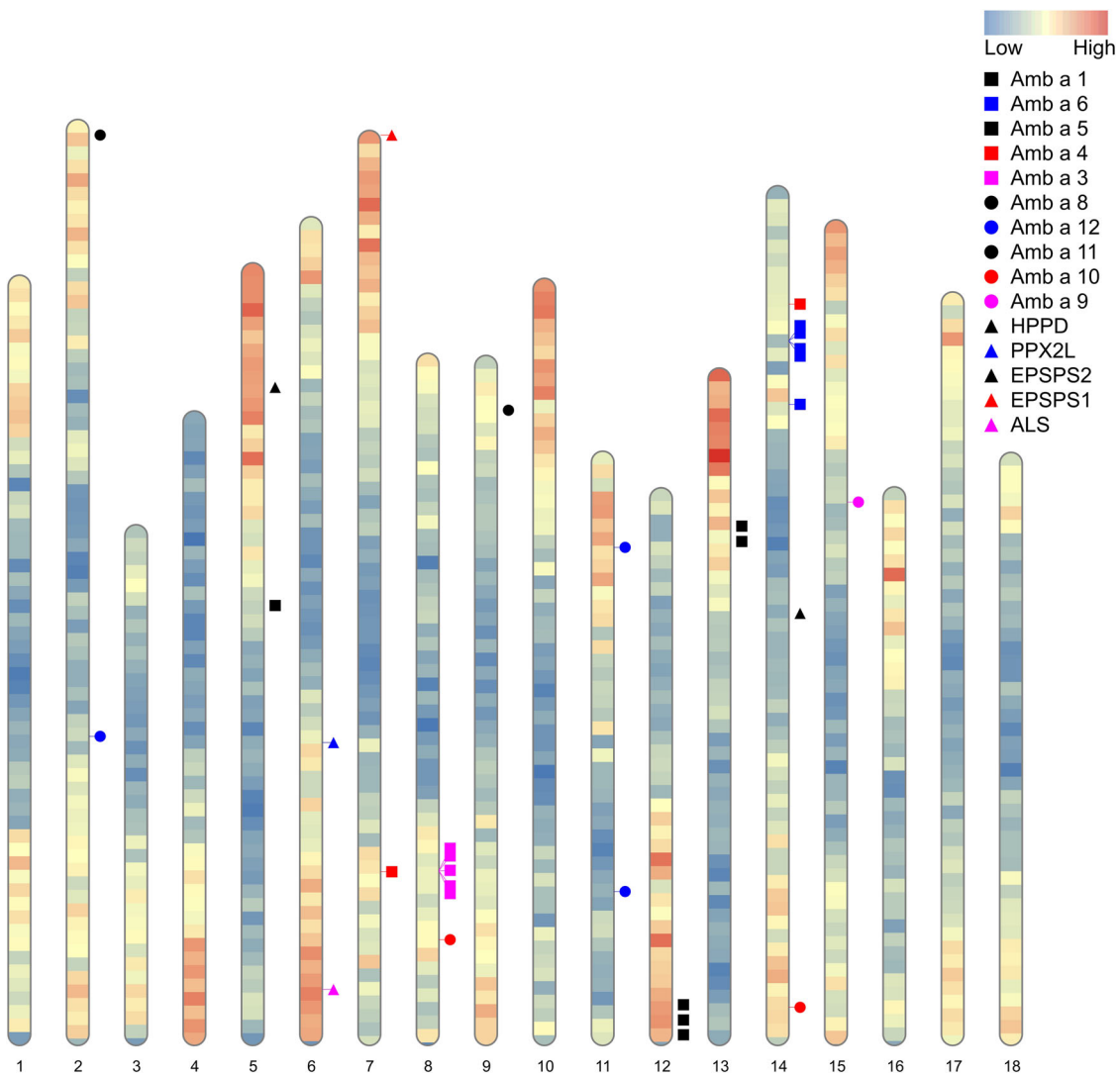


FIGURE 3 *Ambrosia artemisiifolia* ideograms with gene density, per 100 kb, shown as a color gradient. Ten allergenicity-related genes and five genes related to herbicide resistance have been localized on the physical map by sequence comparison.

levels of divergence. For example, aa17 and aa18 represent homoeologs in *A. artemisiifolia*'s genome syntenic with the lower half of cc01 (as shown in Figure 5A) with at05 and at06 homoeologs of the same region in *A. trifida* (Figure 5B). Considered as a group of four, there is striking similarity in the pattern of synteny and when colored by K_s score they contrast, indicating that aa17 and at06 are less diverged and aa17 and at05 are more diverged and reciprocally, aa18 and at05 are less diverged and aa18 and at06 more (Figure 7). This pattern of divergence is also evident in the peak in the K_s score distribution (Figure 6). This indicates that the genomes involved in the polyploidization event that shaped the most recent common ancestor (MRCA) of the Heliantheae Alliance involved diverged genomes; it was an allopolyploidization event.

Four chromosomes aa09, aa13, aa14, and aa15 from *A. artemisiifolia* are not easily grouped with homoeologs

and could represent differences either in the two ancestral genomes or represent post-allopolyploidization changes. Chromosomes aa09 and aa13 are similar to chromosomes from *M. micrantha*, mm05 and mm15, respectively, indicating these are likely differences in the ancestral genomes (Figure S3A). Chromosome aa15, which formed through a nested chromosome fusion (NCF) (Luo et al., 2009; Lysak, 2022), appears to be a post-allopolyploidization change instead. In this nesting, a chromosome syntenic with the majority of cc08 is nested inside a chromosome syntenic with cc07 (Figure 5). Chromosome aa15 is syntenic with *A. trifida*'s at02 (Figure 7) and with *H. annuus*'s ha04 and ha15, but not with chromosomes from *M. micrantha* or *S. sonchifolius*. This suggests that this chromosome likely represents post-allopolyploidization changes that may be specific to the *Ambrosia*–*Helianthus* lineage following the divergence from the *Mikania*–*Smallanthus* lineage. Chromosome

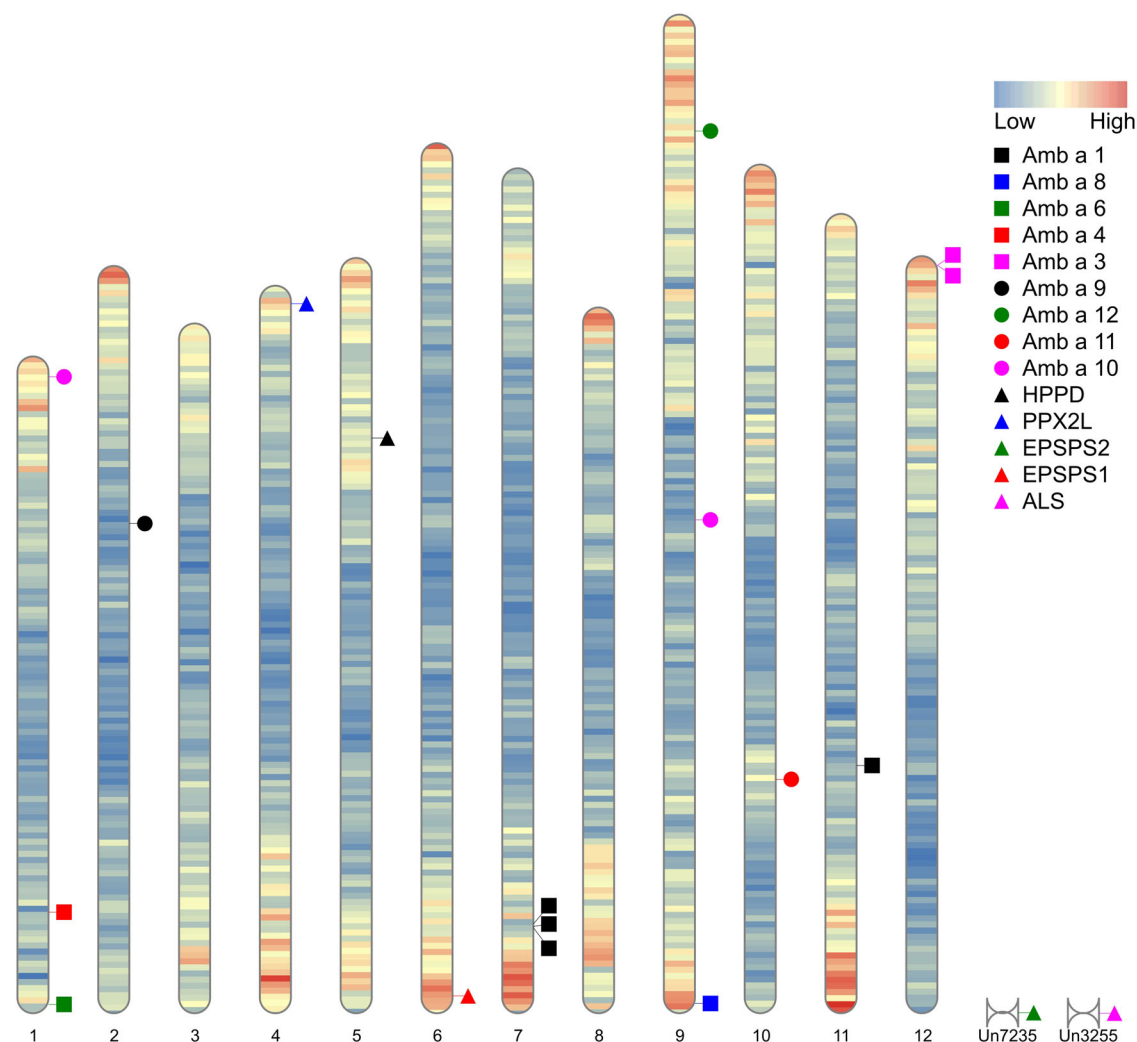


FIGURE 4 *Ambrosia trifida* ideograms with gene density, per 100 kb, shown as a color gradient. Ten allergenicity-related genes and five genes related to herbicide resistance have been localized on the physical map by sequence comparison. ALS and EPSPS2 are located on contigs that were not scaffolded.

aa14 also appears to be a post-allopolyploidization change. Chromosome aa14 is syntenic with the central section of chromosome cc04, but no other chromosomes within *A. artemisiifolia* are syntenic with the same region. *Ambrosia artemisiifolia*'s chromosome aa14 is syntenic with *A. trifida*'s at01 (Figure 7), but not with chromosomes from any of the other species examined here. Comparison with the chromosomes of the other species and examination of the K_s scores of the blocks using WGDI indicate that chromosomes aa14 and at01 are the result of a fusion of two regions, one from each ancestral genome, both syntenic with half of chromosome cc04 ("top," red chromosome Figure 5A). As a result, compared to our reconstructed analog of the ancestral karyotype, only two major differences are seen in the *A. artemisiifolia* genome—a fusion of homoeologous portions of ak07 and ak16 to create aa14, and the fusion of the remaining portion of ak16 and a portion of ak17 to form aa02 (with the remaining portion of ak17 forming aa18) (Figure 8; Table S6).

The maintenance of chromosome structure from the ancestral karyotype seen in *A. artemisiifolia*'s genome is not unique among the species examined here. The genome of *M. micrantha* has 10 of 19 chromosomes that are largely syntenic with the genome of *C. canadensis* and with the reconstructed ancestral karyotype (Figure S4). Seven major changes separate the inferred ancestral structure from that of *M. micrantha*, involving the fission of one ancestral chromosome into two, the movement of five large portions of chromosome arms, and a large insertion. Chromosome number increase in the species resulted from the division of ak17 into two chromosomes—mm08 and mm18 (Figure 8).

The genome of *S. sonchifolius* ($n = 29$) has undergone an additional round of polyploidization since the Heliantheae Alliance's MRCA (Fan et al., 2022). Fan et al. (2022), however, indicated that it was unclear whether this was an allopolyploidization or autopolyploidization event. Through comparison with the reconstructed ancestor, it is clear that

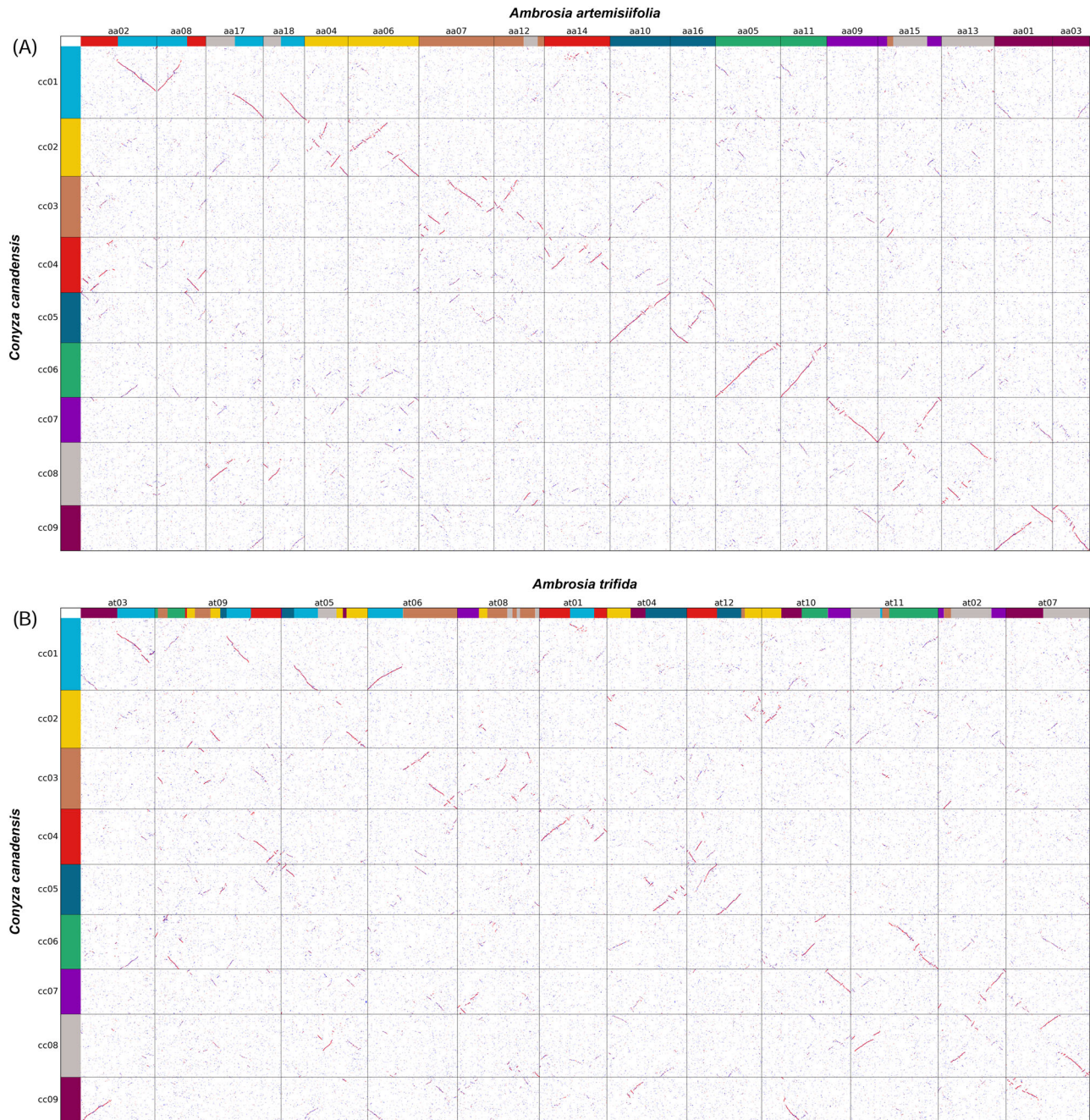


FIGURE 5 Dot plots showing synteny between *Conyza canadensis* and (A) *Ambrosia artemisiifolia* and (B) *Ambrosia trifida*. Points have been colored red for best two matches between protein and gene models and blue for any additional matches. *Conyza canadensis*' chromosomes have been assigned colors arbitrarily and synteny with these chromosomes has been used to color the genomes of the two *Ambrosia* species (WGDI).

this most recent event was an autopolyploidization. *Smallanthus sonchifolius* has six of 29 chromosomes that show synteny along their whole length compared to *C. canadensis* and 10 when compared to the reconstructed ancestor (Figure S5). The decrease from the expected chromosome count of 36, however, is the result of more extensive changes than seen in the genomes of *A. artemisiifolia* and *M. micrantha*. These reductions are primarily the result of NCF. For example, two chromosomes, ss01 and ss17, are the result of

NCFs after autopolyploidization with the nesting of chromosomes like ak10 within ak03 and ak13 within ak14, respectively, reducing the chromosome count by 2 (Figure 8). Three other NCFs have occurred, one before genome duplication and two after duplication, though these are less obvious from the current chromosome structure. Before duplication, a chromosome similar to ak12 was nested within one similar to ak05. These chromosomes then split into ss07 and ss20 and into ss29 and ss10 (Figure S6). The portion forming part of

TABLE 5 Count of homologous proteins found for each allergen across species with an e-value greater than 1×10^{-50} and greater than 50% match identity, the number of orthogroups (OG) to which they belong, information on which ancestral genome the gene is within, and the average match identity (ID%) for all homologs found.

	Conyza canadensis					Ambrosia artemisiifolia					Ambrosia trifida					Helianthus annuus					Mikania micrantha					Smallanthus sonchifolius				
	No.	OG	ID%	No.	OG	ID%	G1	G2	No.	OG	ID%	G1	G2	No.	OG	ID%	G1	G2	No.	OG	ID%	G1	G2	No.	OG	ID%	G1	G2		
Amb a 1	3	2	61	10	5	65	4	6	10	3	65	4	6	5	2	63	2	3	3	5	3	60	2	3	9	58	5	4		
Amb a 3				6	1	91		6	1	1	84	1												1	1	71	1			
Amb a 6				2	1		2							1	1	66		1						1	1	65	1			
Amb a 8	4	1	75	5	1	77		5	6	1	77	5	1	8	1	78	4	4	7	75	5	2	11	82	5	6				
Amb a 9	3	1	60	2	2	78	1	1	2	1	90		2	3	1	67	1	2	1	1	57	1		5	67	5				
Amb a 10				1	1	96	1		1	1	96	1																		
Amb a 11	6	5	53	5	3	59	1	4	5	4	62		5	32	3	57	28	4	5	57	2	3	10	62	1	9				
Amb a 12	4	1	87	5	1	88	2	3	6	3	79	3	3	8	3	83	6	2	6	75	3	3	21	79	10	11				

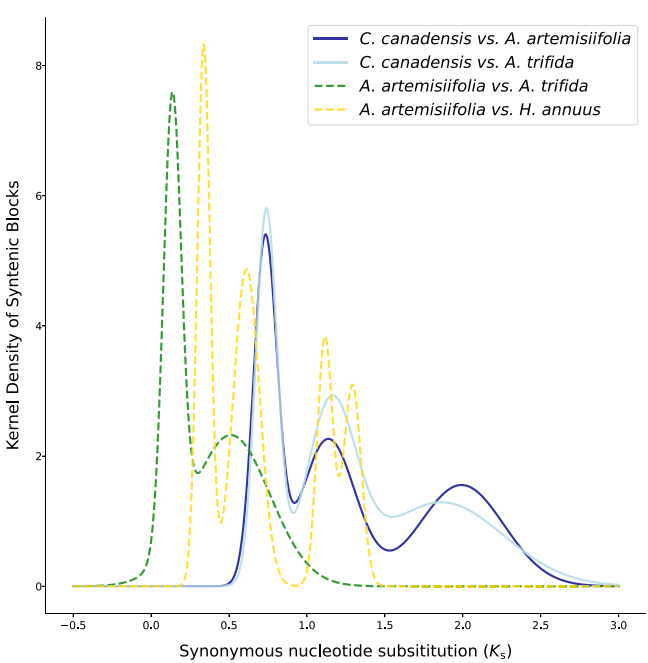


FIGURE 6 K_s score distributions for *Conyza canadensis* versus *Ambrosia artemisiifolia* (solid dark blue) and versus *Ambrosia trifida* (solid light blue); *A. artemisiifolia* versus *A. trifida* (dashed green); and *A. artemisiifolia* versus *Helianthus annuus* (dashed yellow) showing the signatures of genome duplication—multiple peaks in the kernel density of K_s scores showing different portions of the genome have different levels of divergence. For example, for *A. artemisiifolia* versus *A. trifida*, the first of the two peaks indicates low divergence when blocks from one ancestral subgenome are compared with blocks from the same ancestral subgenome, while the second peak indicates greater divergence when these blocks are compared to blocks from the other ancestral subgenome. As blocks from one of the ancestral subgenome diverge, additional peaks appear as in the *A. artemisiifolia* versus *H. annuus* comparison.

ss10 fused with one of the chromosomes similar to ss16 further reducing the chromosome count by 1. After duplication, a double chromosome fusion occurred in chromosomes homologous to ak10, ak11, and ak04; these chromosomes then broke into ss22 and ss25 (Figure S7). Interestingly, the other chromosome in the ak11 pair also underwent a NCF with a chromosome similar to ak13. This chromosome then split into two—one forming ss15 and the other joining one of the chromosomes similar to ak14. The other change resulting in chromosome reduction occurred before autopolyploidization and was the result of incorporation of the ak08-like chromosome into three other chromosomes, ones similar to ss04, ss02, and ss16. The majority of the chromosome was incorporated within one of the arms of ak17 before polyploidization. This suggests that just prior to autopolyploidization, the lineage had $n = 17$ chromosomes. If we take into consideration the changes that occurred after *S. sonchifolius*'s most recent polyploidization, the species' genome would have shown strong synteny with the ancestral genome with 10 of 17 of

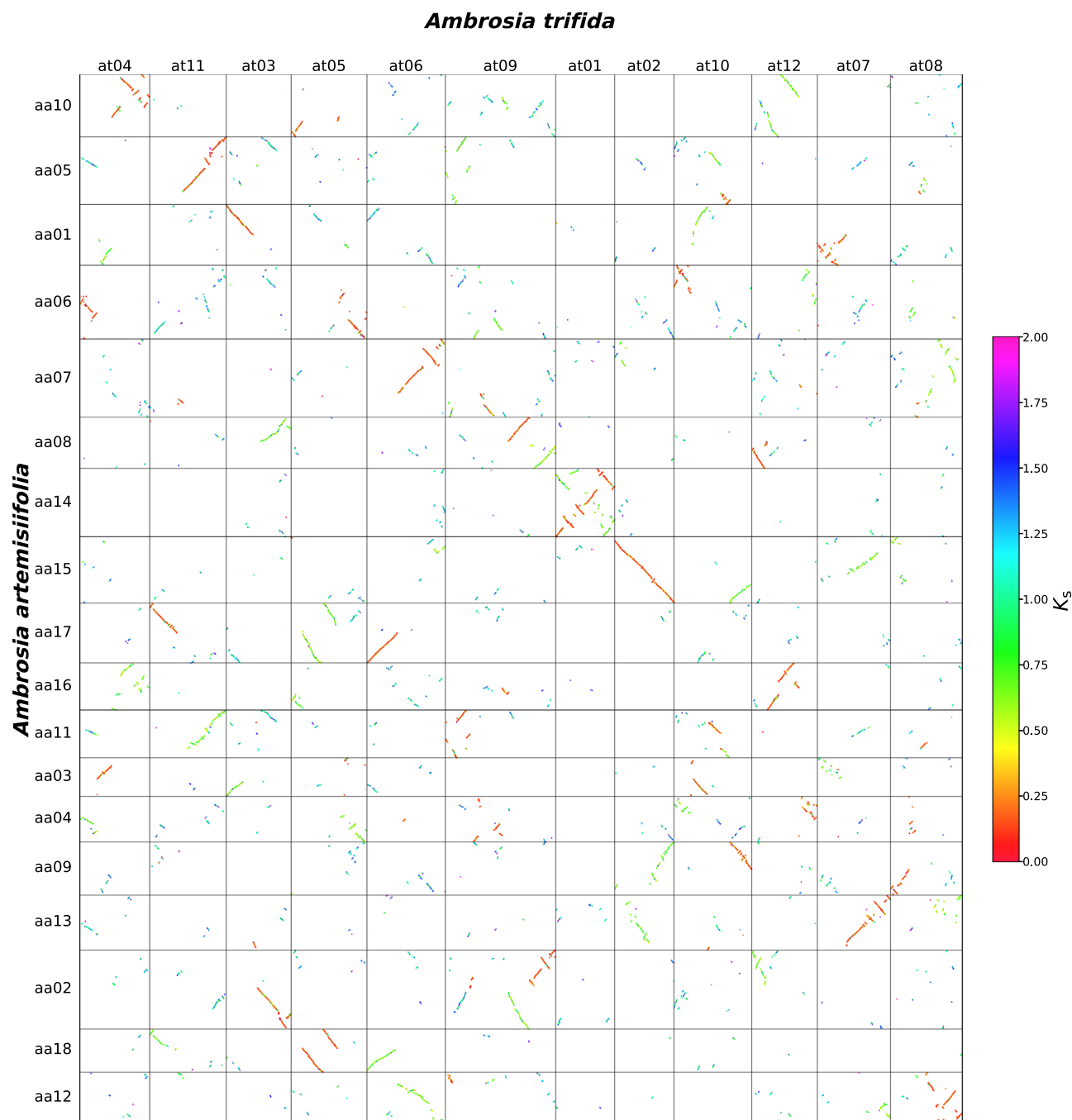


FIGURE 7 K_s scores for syntenic blocks proteins between *Ambrosia artemisiifolia* and *Ambrosia trifida* indicating divergence of the blocks. Blocks with colors toward the red end of the spectrum indicate lower divergence and originate within the same ancestral genome, while colors toward the magenta end indicate higher divergence.

the chromosomes showing similar structure. This could be confirmed with chromosome-level sequencing and annotation of a close relative of *S. sonchifolius* that did not share the autoploidization event, such as *Smallanthus riparius* (Kunth) H. Rob., *Smallanthus macvaughii* (J.R. Wells) H. Rob., or another member of the Millerieae.

In contrast to the other three genomes studied here, it is harder to pick out pairs of homoeologous chromosomes with similar structure for *H. annuus* ($n = 17$) or *A. trifida* ($n = 12$)

using comparison to *C. canadensis*. *Helianthus annuus* has three chromosomes that show full synteny with chromosomes from *C. canadensis* and four when compared to the reconstructed ancestor (Figure S8), while *A. trifida* has only one (at02–ak08) (Figures 5B and 8). In *H. annuus*, chromosome reduction appears to be the result of ak07 fusing with the majority of ak15 and the remaining portion of ak15 fusing with ak06, though several other ancestral chromosomes, such as ak13 and ak16, have been distributed among three

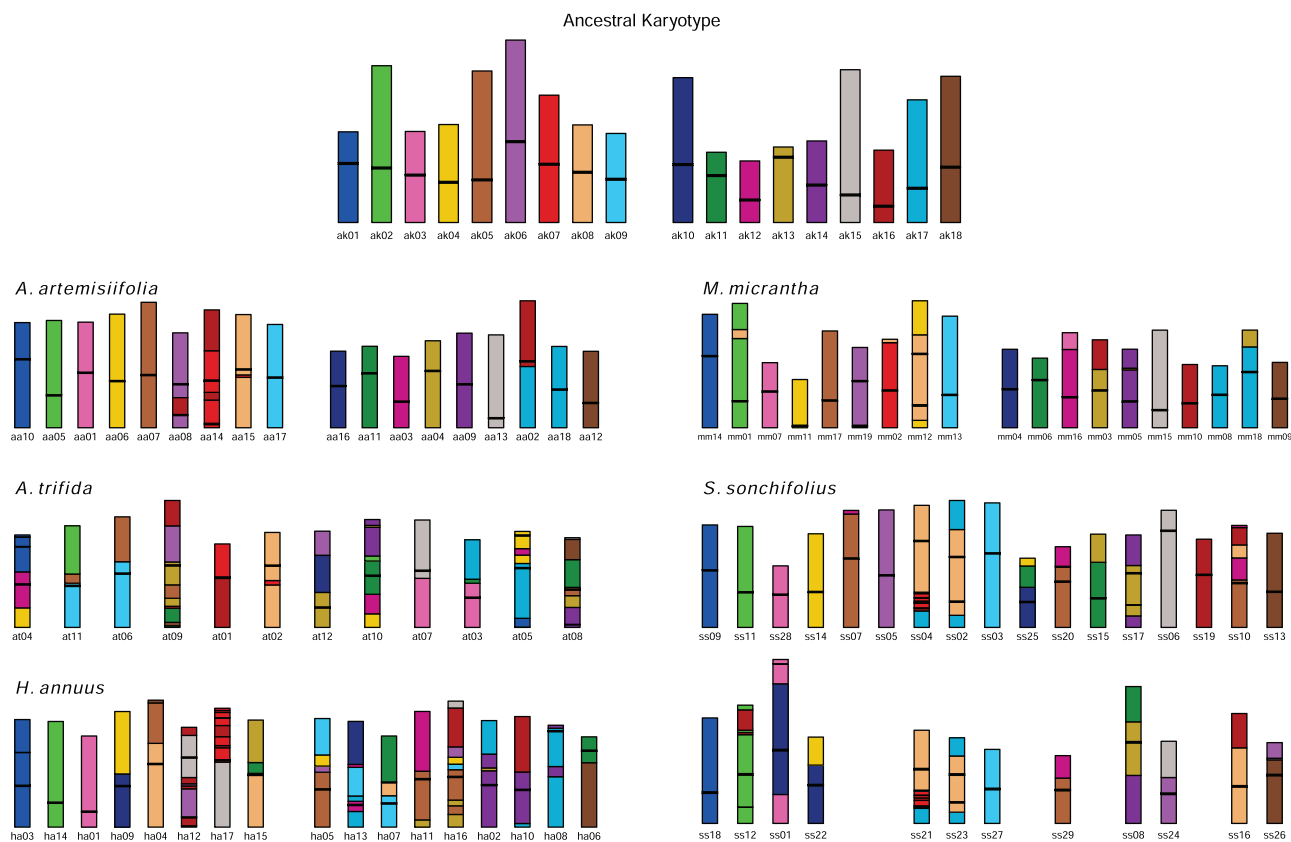


FIGURE 8 Inferred karyotype of the ancestral genomes with chromosomes with a common ancestor, but brought back together in the same genome by allopolyploidization colored in similar colors (e.g., ak01 and ak10). The other genomes studied here are colored based on synteny with these ancestral chromosomes.

chromosomes. In *A. trifida*, there are no clear chromosome number decreases through NCFs or end-to-end joining with extensive mixing of the ancestral chromosomes. For example, genes syntenic with ak04 are scattered through *A. trifida*'s genome on at04, at05, and at10, even though these genes remain largely together in *A. artemisiifolia*. Indeed, six chromosomes have genes from three or more inferred ancestral chromosomes. As a result, the number of changes needed to transition to the ancestral chromosome structure would be much greater for *H. annuus* and *A. trifida* than the other species studied. For example, ha04 and ha15 have different arms of ak08, but this structure cannot be the result of a single reciprocal translocation as they are united with arms of ak05 and ak13, respectively, rather than with the arms of a single other chromosome. This extensive remixing of material from multiple chromosomes may suggest that chromosome restructuring or chromoplexy has occurred in both genomes (Guo et al., 2023). In this case, many of the changes in structure observed here would be the result of a single large event within each genome rather than a stepwise accumulation of changes. Despite mixing and fragmentation, most duplicated regions have been retained, as indicated by the similarity in retention of genes in *A. artemisiifolia* and *A. trifida* compared to *C. canadensis* (Figure 9A,B; Table 6) (Martin et al., 2018;

Rice et al., 2015; Strother, 1993; Tomasello et al., 2019).

For all of the species included in this analysis, there has been a slight overall bias in the retention of genes from ancestral genome 1 (Figure 9A,B; Table 7). If we examine retention of genes broken down by the position of homologous genes on the chromosomes of *C. canadensis*, there is variation in the bias, with 18 of the 45 groups retaining more genes from ancestral genome 2 and 3 showing no bias. The ancestral genome retained is consistent across genomes with the exception that genes homologous to *C. canadensis*'s chromosome 6 in *A. trifida* show higher retention from ancestral genome 2 rather than 1 (Table S7; Figures S9–S11).

4 | DISCUSSION

4.1 | Genome assembly

The ragweeds, *A. artemisiifolia* and *A. trifida*, are invasive, weedy species from the Asteraceae known around the world due to their impact as agricultural weeds and as producers of highly allergenic pollen. The genomes of these species, like other members of the Asteraceae, are expected to have high levels of repetitive DNA, and as wind-pollinated outcrossing

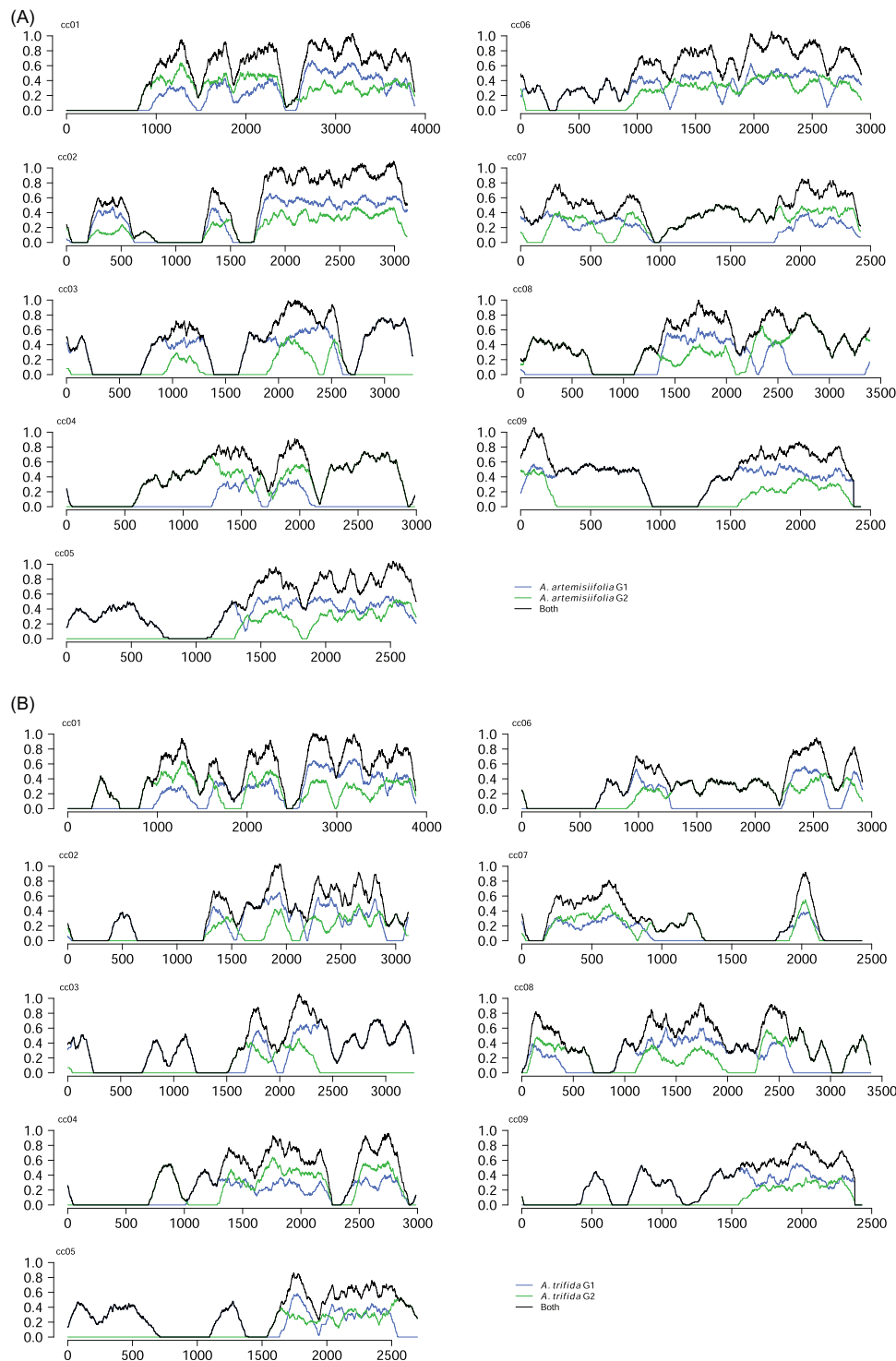


FIGURE 9 Retention of *Conyza canadensis* genes in (A) *Ambrosia artemisiifolia* and (B) *Ambrosia trifida* by *C. canadensis* chromosome. Genes found to be within syntenic blocks between the species were given a weight of 1, and a moving average with a window size of 100 was used. To calculate retention across both ancestral genomes, if a syntenic gene was found in both genomes, the weight of the gene was increased to 2.

species, they are expected to have high levels of heterozygosity. In addition to these complications for genome assembly, the species shared a polyploidization event with other members of the Heliantheae Alliance approximately 25 MYA (Fan et al., 2022; Mandel et al., 2019). Here, we took advantage of

hybrids generated between *A. artemisiifolia* and *A. trifida* to reduce the complexity of genome assembly using the Canu trio binning procedure (Koren et al., 2018) and Hi-C. This approach was beneficial as it reduced the number of contigs in the initial assembly created by CANU from 18,851 to

TABLE 6 Retention of genes in syntenic blocks homologous with *Conyza canadensis* within subgenomes of Helianthus Alliance species.

Species	<i>Conyza canadensis</i> genes with one homolog in species	Genes within AG1 subgenome		<i>p</i> -value	χ^2
		Count	(%)		
<i>Ambrosia artemisiifolia</i>	9407	5101	54	<0.001	67.08
<i>Ambrosia trifida</i>	7603	4319	57	<0.001	140.62
<i>Helianthus annuus</i>	11,875	6434	54	<0.001	82.869
<i>Mikania micrantha</i>	9154	4928	54	<0.001	53.682
<i>Smallanthus sonchifolius</i>	11,717	6594	56	<0.001	184.42

TABLE 7 Bias in retention of genes in syntenic blocks homologous with *Conyza canadensis* within subgenomes of Helianthus Alliance species broken down by *C. canadensis* chromosome.

<i>Conyza canadensis</i> chromosome	<i>Ambrosia artemisiifolia</i>	<i>Ambrosia trifida</i>	<i>Helianthus annuus</i>	<i>Mikania micrantha</i>	<i>Smallanthus sonchifolius</i>
cc01	G2		G2	G2	
cc02	G1	G1	G1	G1	G1
cc03	G1	G1	G1	G1	G1
cc04	G2	G2	G2	G2	G2
cc05	G1	G1	G1	G1	G1
cc06	G1	G2	G1	G1	G1
cc07	G2	G2	G2	G2	G2
cc08	G2		G2	G2	G2
cc09	G1	G1	G1	G1	G1

Note: Blank spaces represent no statistically significant bias at $p < 0.01$.

4656 and 38,288 to 9408 for *A. artemisiifolia* and *A. trifida*, respectively. More importantly, the gain in length of the contigs obtained through this analysis was sufficient to build a high confidence chromosome-scale assembly following Hi-C scaffolding.

4.2 | Genes of interest

From an applied perspective, interest in the ragweeds stems from a desire to control weed populations reducing their extensive health and agricultural impact. However, populations of both *A. artemisiifolia* and *A. trifida* have developed resistance (Heap, 2023) to glyphosate (*N*-(phosphonomethyl) glycine), the most widely used herbicide in the world (Duke, 2018a). As reviewed by Sammons and Gaines (2014), the molecular mechanisms conferring resistance to glyphosate can be broken down into two broad categories: (i) alterations to the target site gene, EPSPS, and (ii) those involving non-target site mechanisms, which can include vacuolar sequestration, reduced translocation, copy number variation, among others. Here, we report that the genomes of *A. artemisiifolia* and *A. trifida* each contain two full length copies of the EPSPS gene, annotated as EPSPS1 and EPSPS2 based on homology with *C. canadensis* (Figure S12), and provide the genomic

resources needed to further investigate the genetic basis of this resistance.

Previous studies of glyphosate resistance in *A. artemisiifolia* and *A. trifida* have uniformly reported that resistance was conferred through nontarget site mechanisms after sequencing EPSPS and not observing proline-to-serine mutation at position 106 (P106S) or other target site mutations (Byker et al., 2022; Ganie et al., 2017; Moretti et al., 2018; Nandula et al., 2017, 2015; Van Horn et al., 2018). Through the sequencing of the genomes, however, it has become clear that EPSPS2 in both ragweed species contains a P106S mutation. This mutation has been widely reported to confer resistance to glyphosate in other weed species, including *C. canadensis* (Page et al., 2017) and *Eleusine indica* (L.) Gaertn. (Baerson et al., 2002), among others. When the primers utilized in the previous studies of glyphosate resistance in the ragweeds (Nandula et al., 2017, 2015) were aligned to the genomes, it was evident that they were amplifying the copy of the gene without the target site mutation (i.e., EPSPS1). Given this, it is highly likely that P106S was present in previously described glyphosate-resistant biotypes of *A. artemisiifolia* and *A. trifida*. This conclusion is supported by several lines of evidence from the current work. For example, the P106S mutation was also observed in the RNAseq data used to annotate the *A. trifida* genome. These data were produced from

a glyphosate-resistant biotype described by Van Horn et al. (2018) as displaying a rapid cell death response and were initially reported as not having a resistance conferring EPSPS mutation (F. Tardif, personal communication, 2023). Second, the paternal *A. artemisiifolia* parent of the hybrid used to produce the genome assemblies was descendant from the same source population as the first case of glyphosate-resistant *A. artemisiifolia* in Canada (Byker et al., 2022; Van Wely et al., 2015). The observation of P106S in the genomic assembly suggests this parental population also contained the P106S target site mutation. Subsequent sequencing with primers specific to EPSPS2 confirmed the near ubiquity of P106S in this glyphosate-resistant *A. artemisiifolia* biotype (Figure S13).

Studies to identify new allergens from nonmodel plants are often hindered by a lack of information about species' genome and proteome (Bordas-Le Floch, Le Mignon, et al., 2015). Currently, 11 allergenic proteins (with various isoforms) have been identified in *A. artemisiifolia* and recognized by the World Health Organization/International Union of Immunological Societies (Mari, 2023; Mari et al., 2009). Here, the predicted proteins contained homologs for eight of these proteins in *A. artemisiifolia* and seven in *A. trifida* with a further two with homology based on DNA sequences, allowing mapping of 10 and nine of the 11 allergenic proteins to the *A. artemisiifolia* and *A. trifida* genomes, respectively (Figures 3 and 4; Tables S4 and S5). These genes, such as Amb a 3 (chromosome 8) and Amb a 6 (chromosome 8), form clusters suggesting tandem amplification has increased their copy number. The two major allergens from *A. artemisiifolia* are Amb a 1, a pectate lyase, and Amb a 11, a cysteine protease. Homology has been noted between Amb a 1 and allergens in *Artemisia vulgaris* L. (mugwort) and the Cupressaceae as well as with a pectate lyase from *H. annuus* (Chen et al., 2018). Here, we see that the Amb a 1 protein maps with 90% identity to a protein with a sequence located on chromosome 05 of *A. artemisiifolia*. This gene forms an orthogroup (genes identified as descendants of the same gene) with six other proteins in *A. artemisiifolia*, two in *C. canadensis*, eight in *A. trifida*, four in *H. annuus*, three in *M. micrantha* and six in *S. sonchifolius*. It is possible that these genes underly allergic responses to this protein in these other species like the pectate lyase in the heavier pollen of *H. annuus* noted previously (Chen et al., 2018). If, like *C. canadensis*, the ancestors of the Heliantheae Alliance (see below) also had two copies of the gene, the expectation would be that all these species would have four copies of the gene (two from each ancestral genome), except *S. sonchifolius* where the expectation would be eight (four from each ancestral genome) (Table 5). This suggests that these genes have increased in copy number in *Ambrosia*, while some copies have been lost in *M. micrantha* and *S. sonchifolius*. Homology between Amb a 11 and proteases in *Actinidia* Lindl. (kiwi), *Ananas comosus* (L.) Merr. (pineapple), and species of Pyroglyphidae (dust mites) has

been noted (Chen et al., 2018). Here, the Amb a 11 protein maps best, with 100% identity, to a protein with a sequence on chromosome 9 of *A. trifida*. This gene is part of a smaller orthogroup with members in *H. annuus* (2), *S. sonchifolius* (4), and *M. micrantha* (1), but not in *A. artemisiifolia* (Tables S5 and S6). Additional annotations using RNA from pollen would improve the probability of determining the location of these genes, their homologs, and orthologs across species. Further studies would be required to link the sequences identified here with the production of allergenic proteins, determine where the genes are expressed, and determine the level of homology that produces protein that elicits a hypersensitivity response. However, given the full genome sequence of these two ragweed, it is clear that there could be many loci underlying the production of these proteins and that similar proteins are produced in the majority of species examined here. The public availability of the *Ambrosia* genomes sequenced with annotations here will support further analysis of the allergome of these species.

4.3 | Allopolyploidy and the ancestral karyotype

As we examine the evolutionary secrets hidden in the genomes of plant species, their complex history of hybridization, whole-genome duplication, and genome reorganization is being revealed. Species of *Ambrosia* are no exception with genomes that share successive rounds of polyploidization: a duplication shared by the seed plants (WGD- ζ), a duplication shared by angiosperms (WGD- ϵ), the triplication shared by dicots (WGT- γ), a triplication shared by the asterids II lineage (WGT-1), and the most recent duplication shared with other closely related tribes in the Asteraceae including Heliantheae and Eupatorieae (WGD-2) (Jiao et al., 2011; Liu et al., 2020; Mandel et al., 2019). Species within this group of tribes, called the Helianthus Alliance, include a stunning array of morphological diversity that includes oil seed crops such as *H. annuus* and root crops such as *S. sonchifolius* and *Helianthus tuberosus* L. (sunchoke). It also includes weed species including *M. micrantha*, which was ranked among the top 100 invasive alien species by the World Conservation Union (Lowe et al., 2000), and the two *Ambrosia* species studied here. Here, we present evidence that the two genomes involved in this genome duplication event had diverged, indicating it was an allopolyploidization event between two distinct species (Figures 6 and 7). As a result, the genomes of members of Heliantheae Alliance are each an example of the potential outcome of allopolyploidization given approximately 25 million years of evolution.

The *Ambrosia* genomes sequenced and assembled here represent two different trajectories for that evolution. Chromosome number and genome size appear to have remained

relatively stable in *A. artemisiifolia* since the MRCA of the Heliantheae Alliance. As a result, the whole-genome duplication is the most obvious feature of *A. artemisiifolia* with many duplicated regions still clear and the similar structure of homoeologous chromosomes still evident when compared to the genome of *C. canadensis*, which did not undergo WGD-2 (Figure 5A). The genomes of *S. sonchifolius* and *M. micrantha* show similar relative stability and allowed us to reconstruct the karyotype of their MRCA (Figure 8). Refinements of this ancestral karyotype will no doubt be made as more representatives of the approximately 5500 species in the group (Baldwin, 2009) are sequenced. Additionally, a species more closely related to the Heliantheae Alliance, but which did not share the genome duplication at the base of the group, could provide more information than that of *C. canadensis* alone. In contrast to these three genomes, the genome of *A. trifida* has undergone more extensive rearrangement and a reduction from $n = 18$ to $n = 12$. This has resulted in greater fragmentation of regions syntenic with *C. canadensis*, or with the reconstructed ancestral karyotype, obscuring the homoeologous origins of the chromosomes. This is also the case for the genome of *H. annuus* ($n = 17$) where numerous changes have altered the structure compared to the ancestral karyotype. *Helianthus annuus* has been shown to have an exceptionally high rate, one of the highest among plants and animals, of chromosomal rearrangement with an estimated 5.5–7.9 chromosomal rearrangements per million years (Burke et al., 2004; Ostevik et al., 2020). This raises the possibility that *A. trifida* has a similarly high rate of chromosomal rearrangement contributing to the divergence seen between *A. trifida* and *A. artemisiifolia*.

Given the number of rearrangements needed to form *A. trifida*'s chromosomes and the relative stability observed in *A. artemisiifolia* from which its lineage diverged approximately 5 MYA (Kumar et al., 2022), it is possible that these chromosomal rearrangements did not occur in a stepwise process, but rather primarily as the result of a single cataclysmic chromosome restructuring event, chromoplexy. During this event, a group of chromosomes would have had double-stranded breaks resulting in fragments that were reassembled like a pique assiette mosaic via nonhomologous end joining or microhomology-mediated end joining (Burssed et al., 2022; Guo et al., 2023). The process is similar to chromothripsis, chromosome shattering, but is distinct as it involves multiple chromosomes rather than just one (Guo et al., 2023). Restructuring of chromosomes can have large consequences for gene regulation leading to numerous changes to the development, morphology, and reproductive compatibility. As a result, this process is likely frequently deleterious, as in prostate cancer where it was discovered (Baca et al., 2013), but occasionally innovative, supplying novel, likely disjunctive, variation on which selection can act, contributing to diversification and speciation (Pellestor & Gatinois, 2019). Chromoplexy has

been suggested to have shaped the genomes (Guo et al., 2023) of *Camelina* spp. (Mandakova et al., 2019) and *Cucumis* spp. (Zhao et al., 2021) and contributed to variation in *Vitis* spp. (Dvorak et al., 2018), but has not yet been documented in the Asteraceae. It is possible that chromoplexy has contributed to the high rate of rearrangement noted in *H. annuus* and other plant taxa such as grasses (Dvorak et al., 2018). Whether it is primarily extrinsic factors, such as exposure to radiation or incomplete apoptosis, or if there are characteristics that predispose genomes to undergo chromoplexy is unclear (Burssed et al., 2022; Guo et al., 2023; Pellestor et al., 2022). The presence of extra chromosomes, aneuploidy, in trisomic and tetrasomic human cell lines was associated with increased chromosomal rearrangement and has been suggested as a contributing factor to chromoanagenesis (Pellestor et al., 2022). In the case of these Helianthus Alliance species, however, the genomes have a similar history of chromosome duplication except *S. sonchifolius*, which has undergone an additional round of autopolyploidization but has one of the more stable genomes. TE have also been cited promoters of rapid chromosome restructuring as they can induce secondary structures in DNA that disrupt replication and repair mechanisms (Burssed et al., 2022) (Pellestor & Gatinois, 2019). The species studied here have a similar or higher percent of repetitive elements in their genomes compared to what we have in ours (*Homo sapiens*—45%) or is found in other plant species suspected to have undergone the process: *Camelina* species (35%–50%) (Martin et al., 2022), sequenced *Cucumis* species (24%–50%) (Huang et al., 2009; Ling et al., 2021), or *Vitis* (41%) (Jailon et al., 2007). TE were determined to make up over 60% *A. artemisiifolia* and *A. trifida*'s genomes sequenced here, while they make up 70% of *H. annuus* (Badouin et al., 2017; Liu et al., 2020) and *M. micrantha* (Liu et al., 2020) genomes and over 80% of the genome of relatively stable *S. sonchifolius* (Fan et al., 2022; Pellestor et al., 2022). Further investigation and more examples of genomes where chromoplexy has been verified as having occurred are needed before this type of chromoanagenesis in plants is understood. The Heliantheae Alliance is an example of an allopolyploid lineage that has flourished and diversified. Assessing how often this diversification has been a consequence of dramatic or high rates of change in chromosomal organization as in *A. trifida* or *H. annuus* compared to a more conservative process as in *A. artemisiifolia*, *M. micrantha*, and *S. sonchifolius* requires a broader understanding of the genomes of species within the Heliantheae Alliance. Specifically, analysis of additional genomes within *Helianthus* and *Ambrosia* may reveal intermediate stages and elucidate stepwise changes that resulted in the current structure of their genomes. Twenty-nine of the 47 *Ambrosia* species have available chromosome counts (Rice et al., 2015; Strother, 1993). The majority (18/29) of these taxa have chromosome counts of $n = 18$, and species that deviate usually have higher counts with a base of 18 suggesting recent

polyploidization events (e.g., *Ambrosia psilostachya*: $n = 18$, 36, 54, and 72). Other than *A. trifida* ($n = 12$), the only exception to this is *A. trifida*'s sister species *A. bidentata* (Martin et al., 2018; Tomasello et al., 2019), which shows the loss of one chromosome ($n = 17$). As a result, the genome of *A. bidentata* would provide additional context for understanding how the changes in *A. trifida*'s genome arose.

Another noteworthy characteristic of the genomes studied here is the prevalence of NCF in the reduction of genome size of *S. sonchifolius*. Five NCFs occurred within the genome, one before and four after the latest round of the genome duplication in the lineage—an autopolyploidization event of a taxon with $n = 17$. NCFs, where one chromosome is inserted between the arms of another, have been described as the dominant mechanism for chromosome number reduction in grass species (Murat et al., 2010). Since the first description of this type of fusion within the Triticeae tribe in Poaceae (Luo et al., 2009), one or two examples of NCF have been noted within the genomes of species from the Betulaceae, Brassicaceae, Cucurbitaceae, Fabaceae, and Rosaceae (Lysak, 2022). Comparison with these cases indicates that genome of *S. sonchifolius* has a high number of NCFs, similar to what has been seen in some grass species, and is, to our knowledge, the first reported case observed within the Asteraceae. We also see evidence of an older NCF shared by the *Ambrosia* species and *Helianthus*, when compared to the *C. canadensis* genome. This supports the suggestion of Lysak (2022) that NCFs may have played a more important role in eukaryotic karyotype evolution than previously appreciated. Species in the Heliantheae Alliance may provide interesting insights into the evolutionary significance and frequency of these types of fusions.

A feature of some, but not all, allopolyploid genomes is fractionation bias where the genes from one ancestral genome are preferentially retained compared to the genes from the other ancestral genome (Alger & Edger, 2020; Bird et al., 2018; Thomas et al., 2006; Wendel et al., 2018; Woodhouse et al., 2014). For example, *Arabidopsis thaliana* (Thomas et al., 2006), *Zea mays* L. (corn) (Schnable et al., 2011), *Gossypium* (cotton) (Flagel et al., 2008), *Triticum aestivum* L. (wheat) (Eckardt, 2014), and *Brassica rapa* L. (Wang et al., 2011) genomes show evidence of fractionation bias or subgenomes dominance. However, fractionation has not been observed in other allopolyploid genomes such as *Camelina sativa* (L.) Crantz (Kagale et al., 2014), *Capsella bursa-pastoris* (L.) Medik. (Douglas et al., 2015), *Cucurbita* (Sun et al., 2017), *Mangifera indica* L. (mango) (Wang et al., 2020), higher ploidy *Brassica*'s from the triangle of U (Chalhoub et al., 2014; Song et al., 2021; Yang et al., 2016), or *Poa annua* L. (annual bluegrass) (Benson et al., 2023). Here, we see evidence of a slight overall bias toward the retention of genes from ancestral genome 1 (Table 6). This varies when genes are divided by their location on *C. canadensis*'s chromosomes, with 40% of these groups showing bias toward preferential retention of genes from ancestral genome 2 (Table 7). These biases are consistent across the five species with only one exception, suggesting that these biases were already established in the MRCA of the Helianthus Alliances after allopolyploidization rather than following similar tracks in each species. This is expected from studies indicating that rapid changes in DNA methylation, TE activity, and gene expression determine subgenome dominance immediately upon formation (Alger & Edger, 2020; Edger et al., 2017; Zhang et al., 2015). Further work is needed to understand whether the expression patterns of these genes reflect a bias toward ancestral genome 1 and to investigate the role of genes retained from both ancestral genomes (Birchler & Yang, 2022).

Here, we produced two high-quality chromosome-level assemblies for *A. artemisiifolia* and *A. trifida* with high levels of completeness as measured by BUSCOs and the LAI. Sequencing a hybrid between the two species was central in achieving high contiguity in both individual assemblies. Like other members of the Heliantheae Alliance, the majority of these genomes are composed of repetitive elements and dominated by LTRs. Sequencing these genomes allowed us to develop a fuller understanding of the genes that impact human activities. For example, it revealed a second EPSPS gene is present within the genome of the ragweeds and that this gene carried a previously unreported target site mutation associated with resistance to the herbicide glyphosate. It also provides the necessary genomic foundation for the analysis of allergens from *A. artemisiifolia* and *A. trifida* that cause seasonal rhinitis in humans. Access to the complete genome sequences provides a full picture of an organism's gene set and associated regulatory regions, which is essential for the analysis of traits such as herbicide resistance and allergenicity. Comparison of the available annotated genomes from Heliantheae Alliance with each other and *C. canadensis* has allowed us to determine an allopolyploid origin for the MRCA of the Heliantheae Alliance and reconstruct its karyotype. We see that the chromosome organization in *A. artemisiifolia*, *M. micrantha*, and *S. sonchifolius* has been relatively conserved, retaining a structure that more clearly reflects the MRCA of the Heliantheae Alliance. In contrast, the genome structures of *A. trifida* and *H. annuus* show a greater number of rearrangements and may have been shaped by high rates of chromosome rearrangements or by chromoplexy. Finally, in *S. sonchifolius*, which has undergone an additional autopolyploidization, NCFs have contributed to chromosome number reduction. Finally, all the Heliantheae Alliance species examined here show a slight fractionation bias toward ancestral

5 | CONCLUSION

Here, we produced two high-quality chromosome-level assemblies for *A. artemisiifolia* and *A. trifida* with high levels of completeness as measured by BUSCOs and the LAI. Sequencing a hybrid between the two species was central in achieving high contiguity in both individual assemblies. Like other members of the Heliantheae Alliance, the majority of these genomes are composed of repetitive elements and dominated by LTRs. Sequencing these genomes allowed us to develop a fuller understanding of the genes that impact human activities. For example, it revealed a second EPSPS gene is present within the genome of the ragweeds and that this gene carried a previously unreported target site mutation associated with resistance to the herbicide glyphosate. It also provides the necessary genomic foundation for the analysis of allergens from *A. artemisiifolia* and *A. trifida* that cause seasonal rhinitis in humans. Access to the complete genome sequences provides a full picture of an organism's gene set and associated regulatory regions, which is essential for the analysis of traits such as herbicide resistance and allergenicity. Comparison of the available annotated genomes from Heliantheae Alliance with each other and *C. canadensis* has allowed us to determine an allopolyploid origin for the MRCA of the Heliantheae Alliance and reconstruct its karyotype. We see that the chromosome organization in *A. artemisiifolia*, *M. micrantha*, and *S. sonchifolius* has been relatively conserved, retaining a structure that more clearly reflects the MRCA of the Heliantheae Alliance. In contrast, the genome structures of *A. trifida* and *H. annuus* show a greater number of rearrangements and may have been shaped by high rates of chromosome rearrangements or by chromoplexy. Finally, in *S. sonchifolius*, which has undergone an additional autopolyploidization, NCFs have contributed to chromosome number reduction. Finally, all the Heliantheae Alliance species examined here show a slight fractionation bias toward ancestral

genome 1. With 5500 morphologically and karyotypically diverse species that share this common allopolyploidization event, further study of members of the group will provide a greater understanding of the potential trajectories for genome evolution following allopolyploidization.

AUTHOR CONTRIBUTIONS

Martin Laforest: Conceptualization; data curation; formal analysis; funding acquisition; investigation; project administration; supervision; validation; visualization; writing—original draft. **Sara L. Martin:** Conceptualization; data curation; formal analysis; investigation; validation; visualization; writing—original draft. **Katherine Bisailon:** Investigation. **Brahim Soufiane:** Investigation. **Sydney Meloche:** Investigation. **François J. Tardif:** Investigation; visualization; writing—review and editing. **Eric Page:** Conceptualization; funding acquisition; project administration; supervision; writing—original draft.

ACKNOWLEDGMENTS

The funding for this project was provided by Agriculture and Agri-Food Canada (Project J-001751 and Project J-002275).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The sequence assemblies for common ragweed (*Ambrosia artemisiifolia*) and giant ragweed (*Ambrosia trifida*) have been made available in the genome section of GenBank under project PRJNA967341.

ORCID

Eric Page  <https://orcid.org/0000-0003-3361-5151>

REFERENCES

- Alger, E. I., & Edger, P. P. (2020). One subgenome to rule them all: Underlying mechanisms of subgenome dominance. *Current Opinion in Plant Biology*, 54, 108–113. <https://doi.org/10.1016/j.cpb.2020.03.004>
- Baca, S. C., Prandi, D., Lawrence, M. S., Mosquera, J. M., Romanel, A., Drier, Y., Park, K., Kitabayashi, N., MacDonald, T. Y., Ghandi, M., Van Allen, E., Kryukov, G. V., Sboner, A., Theurillat, J. P., Soong, T. D., Nickerson, E., Auclair, D., Tewari, A., Beltran, H., ... Garraway, L. A. (2013). Punctuated evolution of prostate cancer genomes. *Cell*, 153(3), 666–677. <https://doi.org/10.1016/j.cell.2013.03.021>
- Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., Lelandais-Brière, C., Owens, G. L., Carrère, S., Mayjonade, B., Legrand, L., Gill, N., Kane, N. C., Bowers, J. E., Hubner, S., Bellec, A., Bérard, A., Bergès, H., Blanchet, N., ... Langlade, N. B. (2017). The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature*, 546(7656), 148–152. <https://doi.org/10.1038/nature22380>
- Baerson, S. R., Rodriguez, D. J., Tran, M., Feng, Y., Biest, N. A., & Dill, G. M. (2002). Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Plant Physiology*, 129(3), 1265–1275. <https://doi.org/10.1104/pp.001560>
- Bai, C., Alverson, W. S., Follansbee, A., & Waller, D. M. (2012). New reports of nuclear DNA content for 407 vascular plant taxa from the United States. *Annals of Botany*, 110(8), 1623–1629. <https://doi.org/10.1093/aob/mcs222>
- Baldwin, B. G. (2009). Heliantheae alliance. In V. A. Funk, A. Susanna, T. F. Stuessy, & R. J. Bayer (Eds.), *Systematics, evolution, and biogeography of Compositae* (pp. 689–711). International Association for Plant Taxonomy.
- Barker, M. S., Kane, N. C., Matvienko, M., Kozik, A., Micheltore, R. W., Knapp, S. J., & Rieseberg, L. H. (2008). Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Molecular Biology and Evolution*, 25(11), 2445–2455. <https://doi.org/10.1093/molbev/msn187>
- Battlay, P., Wilson, J., Bieker, V. C., Lee, C., Prapas, D., Petersen, B., Craig, S., van Boheemen, L., Scalone, R., de Silva, N. P., Sharma, A., Konstantinovic, B., Nurkowski, K. A., Rieseberg, L. H., Connallon, T., Martin, M. D., & Hodgins, K. A. (2023). Large haploblocks underlie rapid adaptation in the invasive weed *Ambrosia artemisiifolia*. *Nature Communications*, 14(1), 1717. <https://doi.org/10.1038/s41467-023-37303-4>
- Benson, C. W., Sheltra, M. R., Maughan, P. J., Jellen, E. N., Robbins, M. D., Bushman, B. S., Patterson, E. L., Hall, N. D., & Huff, D. R. (2023). Homoeologous evolution of the allotetraploid genome of *Poa annua* L. *BMC Genomics*, 24(1), 350. <https://doi.org/10.1186/s12864-023-09456-5>
- Bickhart, D. M., Rosen, B. D., Koren, S., Sayre, B. L., Hastie, A. R., Chan, S., Lee, J., Lam, E. T., Liachko, I., Sullivan, S. T., Burton, J. N., Huson, H. J., Nystrom, J. C., Kelley, C. M., Hutchison, J. L., Zhou, Y., Sun, J., Crisa, A., Ponce de Leon, F. A., ... Smith, T. P. (2017). Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nature Genetics*, 49(4), 643–650. <https://doi.org/10.1038/ng.3802>
- Bieker, V. C., Battlay, P., Petersen, B., Sun, X., Wilson, J., Brealey, J. C., Bretagnolle, F., Nurkowski, K., Lee, C., Barreiro, F. S., Owens, G. L., Lee, J. Y., Kellner, F. L., van Boheeman, L., Gopalakrishnan, S., Gaudeul, M., Mueller-Schaerer, H., Lommen, S., Karrer, G., ... Martin, M. D. (2022). Uncovering the genomic basis of an extraordinary plant invasion. *Science Advances*, 8(34), eabo5115. <https://doi.org/10.1126/sciadv.abo5115>
- Birchler, J. A., & Yang, H. (2022). The multiple fates of gene duplications: Deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. *Plant Cell*, 34(7), 2466–2474. <https://doi.org/10.1093/plcell/koac076>
- Bird, K. A., VanBuren, R., Puzey, J. R., & Edger, P. P. (2018). The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytologist*, 220(1), 87–93. <https://doi.org/10.1111/nph.15256>
- Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C., & Hochreiter, S. (2015). msa: An R package for multiple sequence alignment. *Bioinformatics*, 31(24), 3997–3999. <https://doi.org/10.1093/bioinformatics/btv494>
- Boessenkool, B. (2023). *berryFunctions: Function collection related to plotting and hydrology*. R package version 1.22.0. <https://CRAN.R-project.org/package=berryFunctions>

- Bordas-Le Floch, V., Groeme, R., Chabre, H., Baron-Bodo, V., Nony, E., Mascarell, L., & Moingeon, P. (2015). New insights into ragweed pollen allergens. *Current Allergy and Asthma Reports*, 15(11), 63. <https://doi.org/10.1007/s11882-015-0565-6>
- Bordas-Le Floch, V., Le Mignon, M., Bouley, J., Groeme, R., Jain, K., Baron-Bodo, V., Nony, E., Mascarell, L., & Moingeon, P. (2015). Identification of novel short ragweed pollen allergens using combined transcriptomic and immunoproteomic approaches. *PLoS ONE*, 10(8), e0136258. <https://doi.org/10.1371/journal.pone.0136258>
- Burke, J. M., Lai, Z., Salmaso, M., Nakazato, T., Tang, S., Heesacker, A., Knapp, S. J., & Rieseberg, L. H. (2004). Comparative mapping and rapid karyotypic evolution in the genus *Helianthus*. *Genetics*, 167, 449–457.
- Burssed, B., Zamariolli, M., Bellucco, F. T., & Melaragno, M. I. (2022). Mechanisms of structural chromosomal rearrangement formation. *Molecular Cytogenetics*, 15(1), 23. <https://doi.org/10.1186/s13039-022-00600-6>
- Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature Biotechnology*, 31(12), 1119–1125. <https://doi.org/10.1038/nbt.2727>
- Byker, H. P., Soltani, N., Nissen, S. J., Gaines, T. A., Westra, P. E., Martin, S. L., Tardif, F. J., Robinson, D. E., Lawton, M. B., & Sikkema, P. H. (2022). Mechanisms of glyphosate resistance in common ragweed (*Ambrosia artemisiifolia*): Patterns of absorption, translocation, and metabolism. *Weed Science*, 70(2), 151–159.
- Case, M. J., & Stinson, K. A. (2018). Climate change impacts on the distribution of the allergenic plant, common ragweed (*Ambrosia artemisiifolia*) in the eastern United States. *PLoS ONE*, 13(10), e0205677. <https://doi.org/10.1371/journal.pone.0205677>
- Cerca, J., Petersen, B., Lazaro-Guevara, J. M., Rivera-Colón, A., Birkeland, S., Vizueta, J., Li, S., Li, Q., Loureiro, J., Kosawang, C., Díaz, P. J., Rivas-Torres, G., Fernández-Mazuecos, M., Vargas, P., McCauley, R. A., Petersen, G., Santos-Bay, L., Wales, N., Catchen, J. M., ... Martin, M. D. (2022). The genomic basis of the plant island syndrome in Darwin's giant daisies. *Nature Communications*, 13(1), 3729. <https://doi.org/10.1038/s41467-022-31280-w>
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corréa, M., Da Silva, C., Just, J., Falentin, C., Koh, C. S., Le Clainche, I., Bernard, M., Bento, P., Noel, B., ... Wincker, P. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, 345, 950–953.
- Charif, D., & Lobry, J. R. (2007). Seqin{R} 1.0-2: A contributed package to the {R} project for statistical computing devoted to biological sequences retrieval and analysis. In U. Bastolla, M. Porto, H. E. Roman, & M. Vendruscolo (Eds.), *Structural approaches to sequence evolution: Molecules, networks, populations* (pp. 207–232). Springer Verlag.
- Chauvel, B., Fried, G., Follak, S., Chapman, D., Kulakova, Y., Le Bourgeois, T., Marisavljevic, D., Monty, A., Rossi, J.-P., & Starfinger, U. (2021). Monographs on invasive plants in Europe N° 5: *Ambrosia trifida* L. *Botany Letters*, 168(2), 167–190.
- Chen, K. W., Marusciac, L., Tamas, P. T., Valenta, R., & Panaitescu, C. (2018). Ragweed pollen allergy: Burden, characteristics, and management of an imported allergen source in Europe. *International Archives of Allergy and Immunology*, 176(3–4), 163–180. <https://doi.org/10.1159/000487997>
- Chen, Y. (2008). *Ambrosia artemisiifolia* Linnaeus, *Sp. Pl.* 2: 988. 1753. http://www.efloras.org/florataxon.aspx?flora_id=2&taxon_id=200023072
- Cowbrough, M. J., Brown, R. B., & Tardif, F. J. (2003). Impact of common ragweed (*Ambrosia artemisiifolia*) aggregation on economic thresholds in soybean. *Weed Science*, 51, 947–954. <https://doi.org/10.1614/02-036>
- D'Amato, G., Cecchi, L., Bonini, S., Nunes, C., Annesi-Maesano, I., Behrendt, H., Liccardi, G., Popov, T., & Van Cauwenberge, P. (2007). Allergenic pollen and pollen allergy in Europe. *Allergy*, 62(9), 976–990.
- Doležel, J., & Bartos, J. (2005). Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany*, 95(1), 99–110. <https://doi.org/10.1093/aob/mci005>
- Dolezel, J., Bartos, J., Voglmayr, H., & Greilhuber, J. (2003). Nuclear DNA content and genome size of trout and human. *Cytometry Part A: The journal of the International Society for Analytical Cytology*, 51(2), 127–129. <https://doi.org/10.1002/cyto.a.10013>
- Dolezel, J., Sgorbati, S., & Lucretti, S. (1992). Comparison of three nuclear DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum*, 85, 625–631.
- Douglas, G. M., Gos, G., Steige, K. A., Salcedo, A., Holm, K., Josephs, E. B., Arunkumar, R., Ågren, J. A., Hazzouri, K. M., Wang, W., Platts, A. E., Williamson, R. J., Neuffer, B., Lascoux, M., Slotte, T., & Wright, S. I. (2015). Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid *Capsella bursa-pastoris*. *Proceedings of the National Academy of Sciences of the United States of America*, 112(9), 2806–2811. <https://doi.org/10.1073/pnas.1412277112>
- Duke, S. O. (2018a). Glyphosate: The world's most successful herbicide under intense scientific scrutiny. *Pest Management Science*, 74(5), 1025–1026. <https://doi.org/10.1002/ps.4902>
- Duke, S. O. (2018b). The history and current status of glyphosate. *Pest Management Science*, 74(5), 1027–1034. <https://doi.org/10.1002/ps.4652>
- Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., & Aiden, E. L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Systems*, 3(1), 99–101. <https://doi.org/10.1016/j.cels.2015.07.012>
- Dvorak, J., Wang, L., Zhu, T., Jorgensen, C. M., Deal, K. R., Dai, X., Dawson, M. W., Muller, H. G., Luo, M. C., Ramasamy, R. K., Dehghani, H., Gu, Y. Q., Gill, B. S., Distelfeld, A., Devos, K. M., Qi, P., You, F. M., Gulick, P. J., & McGuire, P. E. (2018). Structural variation and rates of genome evolution in the grass family seen through comparison of sequences of genomes greatly differing in size. *Plant Journal*, 95(3), 487–503. <https://doi.org/10.1111/tpj.13964>
- Eckardt, N. A. (2014). Genome dominance and interaction at the gene expression level in allohexaploid wheat. *Plant Cell*, 26(5), 1834. <https://doi.org/10.1105/tpc.114.127183>
- Edger, P. P., Smith, R., McKain, M. R., Cooley, A. M., Vallejo-Marin, M., Yuan, Y., Bewick, A. J., Ji, L., Platts, A. E., Bowman, M. J., Childs, K. L., Washburn, J. D., Schmitz, R. J., Smith, G. D., Pires, J. C., & Puzey, J. R. (2017). Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower. *Plant Cell*, 29(9), 2150–2167. <https://doi.org/10.1105/tpc.17.00010>
- Ellinghaus, D., Kurtz, S., & Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR

- retrotransposons. *BMC Bioinformatics*, 9, 18. <https://doi.org/10.1186/1471-2105-9-18>
- Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20(1), 238. <https://doi.org/10.1186/s13059-019-1832-y>
- Fan, W., Wang, S., Wang, H., Wang, A., Jiang, F., Liu, H., Zhao, H., Xu, D., & Zhang, Y. (2022). The genomes of chicory, endive, great burdock and yacon provide insights into Asteraceae palaeopolyploidization history and plant inulin production. *Molecular Ecology Resources*, 22(8), 3124–3140. <https://doi.org/10.1111/1755-0998.13675>
- Faust, G. G., & Hall, I. M. (2014). SAMBLASTER: Fast duplicate marking and structural variant read extraction. *Bioinformatics*, 30(17), 2503–2505. <https://doi.org/10.1093/bioinformatics/btu314>
- Flagel, L., Udall, J., Nettleton, D., & Wendel, J. (2008). Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. *BMC Biology*, 6, 16. <https://doi.org/10.1186/1741-7007-6-16>
- Ganie, Z., Jugulam, M., Varanasi, V., & Jhala, A. J. (2017). Investigating mechanism of glyphosate resistance in a common ragweed (*Ambrosia artemisiifolia* L.) biotype from Nebraska. *Canadian Journal of Plant Science*, 97(6). <https://doi.org/10.1139/cjps-2017-0036>
- Ghosh, B., Perry, M. P., Rafnar, T., & Marsh, D. C. (1993). Cloning and expression of immunologically active recombinant Amb A V allergen of short ragweed (*Ambrosia artemisiifolia*) pollen. *The Journal of Immunology*, 150(12), 5391–5399. <https://doi.org/10.4049/jimmunol.150.12.5391>
- Glover, N. M., Redestig, H., & Dessimoz, C. (2016). Homoeologs: What are they and how do we infer them? *Trends in Plant Science*, 21(7), 609–621. <https://doi.org/10.1016/j.tplants.2016.02.005>
- Guo, W., Comai, L., & Henry, I. M. (2023). Chromoanagenesis in plants: Triggers, mechanisms, and potential impact. *Trends in Genetics*, 39(1), 34–45. <https://doi.org/10.1016/j.tig.2022.08.003>
- Harrison, S. K., Regnier, E. E., Schmoll, J. T., & Webb, J. E. (2001). Competition and fecundity of giant ragweed in corn. *Weed Science*, 49(2), 224–229. [https://doi.org/10.1614/0043-1745\(2001\)049\[0224:cafogr\]2.0.co;2](https://doi.org/10.1614/0043-1745(2001)049[0224:cafogr]2.0.co;2)
- Heap, I. (2023). *International survey of herbicide resistant weeds*. www.weedscience.org
- Hu, K., Xu, K., Wen, J., Yi, B., Shen, J., Ma, C., Fu, T., Ouyang, Y., & Tu, J. (2019). *Helitron* distribution in Brassicaceae and whole Genome *Helitron* density as a character for distinguishing plant species. *BMC Bioinformatics*, 20(1), 354. <https://doi.org/10.1186/s12859-019-2945-8>
- Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W. J., Wang, X., Xie, B., Ni, P., Ren, Y., Zhu, H., Li, J., Lin, K., Jin, W., Fei, Z., Li, G., Staub, J., Kilian, A., ... Li, S. (2009). The genome of the cucumber, *Cucumis sativus* L. *Nature Genetics*, 41(12), 1275–1281. <https://doi.org/10.1038/ng.475>
- Iamónico, D. (2016a). *Ambrosia artemisiifolia* (common ragweed). CABI Digital Library. <https://www.cabidigitallibrary.org/doi/10.1079/cabicompendium.4691#core-ref-61>
- Iamónico, D. (2016b). *Ambrosia trifida* (giant ragweed). CABI Digital Library. <https://www.cabidigitallibrary.org/doi/10.1079/cabicompendium.4693>
- Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choise, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Huguency, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyere, C., ... The French-Italian Public Consortium for Grapevine Genome Characterization. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449(7161), 463–467. <https://doi.org/10.1038/nature06148>
- Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., Tomsho, L. P., Hu, Y., Liang, H., Soltis, P. S., Soltis, D. E., Clifton, S. W., Schlarbaum, S. E., Schuster, S. C., Ma, H., Leebens-Mack, J., & dePamphilis, C. W. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, 473(7345), 97–100. <https://doi.org/10.1038/nature09916>
- Jones, K. L. (1943). Studies on *Ambrosia*: III. Pistillate *Ambrosia elatior* X *A. trifida* and its bearing on matroclinic sex inheritance. *Botanical Gazette*, 105(2), 226–232.
- Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30(9), 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
- Kagale, S., Koh, C., Nixon, J., Bollina, V., Clarke, W. E., Tuteja, R., Spillane, C., Robinson, S. J., Links, M. G., Clarke, C., Higgins, E. E., Huebert, T., Sharpe, A. G., & Parkin, I. A. (2014). The emerging bio-fuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nature Communications*, 5, 3706. <https://doi.org/10.1038/ncomms4706>
- Keller, O., Kollmar, M., Stanke, M., & Waack, S. (2011). A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics*, 27(6), 757–763. <https://doi.org/10.1093/bioinformatics/btr010>
- Kil, J., Shim, K., Park, S., Koh, K., Suh, M., Ku, Y., Suh, S., Oh, H., & Kong, H. (2004). Distributions of naturalized alien plants in South Korea. *Weed Technology*, 18(sp1), 1493–1495.
- Koren, S., Rhie, A., Walenz, B. P., Dilthey, A. T., Bickhart, D. M., Kingan, S. B., Hiendleder, S., Williams, J. L., Smith, T. P., & Phillippy, A. M. (2018). De novo assembly of haplotype-resolved genomes with trio binning. *Nature Biotechnology*, 36(12), 1174–1182.
- Kubesova, M., Moravcova, L., Suda, J., Jarosik, V., & Pysek, P. (2010). Naturalized plants have smaller genomes than their non-invading relatives: A flow cytometric analysis of the Czech alien flora. *Preslia*, 82, 91–96.
- Kumar, S., Suleski, M., Craig, J. E., Kaspruwicz, A. E., Sanderford, M., Li, M., Stecher, G., & Hedges, S. B. (2022). TimeTree 5: An expanded resource for species divergence times. *Molecular Biology and Evolution*, 39, msac174. <https://doi.org/10.1093/molbev/msac174>
- Laforest, M., Martin, S. L., Bisaillon, K., Soufiane, B., Meloche, S., & Page, E. (2020). A chromosome-scale draft sequence of the Canada fleabane genome. *Pest Management Science*, 76(6), 2158–2169. <https://doi.org/10.1002/ps.5753>
- Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M., & Carey, V. (2013). Software for computing and annotating genomic ranges. *PLoS Computational Biology*, 9(8), e1003118. <https://doi.org/10.1371/journal.pcbi.1003118>
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format

- and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Iander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326, 289–293.
- Lin, G., He, C., Zheng, J., Koo, D. H., Le, H., Zheng, H., Tamang, T. M., Lin, J., Liu, Y., Zhao, M., Hao, Y., McFrand, F., Wang, B., Qin, Y., Tang, H., McCarty, D. R., Wei, H., Cho, M. J., Park, S., ... Liu, S. (2021). Chromosome-level genome assembly of a regenerable maize inbred line A188. *Genome Biology*, 22(1), 175. <https://doi.org/10.1186/s13059-021-02396-x>
- Ling, J., Xie, X., Gu, X., Zhao, J., Ping, X., Li, Y., Yang, Y., Mao, Z., & Xie, B. (2021). High-quality chromosome-level genomes of *Cucumis metuliferus* and *Cucumis melo* provide insight into *Cucumis* genome evolution. *Plant Journal*, 107(1), 136–148. <https://doi.org/10.1111/tbj.15279>
- Liu, B., Yan, J., Li, W., Yin, L., Li, P., Yu, H., Xing, L., Cai, M., Wang, H., Zhao, M., Zheng, J., Sun, F., Wang, Z., Jiang, Z., Ou, Q., Li, S., Qu, L., Zhang, Q., Zheng, Y., ... Wan, F. (2020). *Mikania micrantha* genome provides insights into the molecular mechanism of rapid growth. *Nature Communications*, 11(1), 340. <https://doi.org/10.1038/s41467-019-13926-4>
- Lovell, J. T. (2023). *GENESPACE: Synteny- and orthology-constrained comparative genomics*. R package version 1.0.13.
- Lovell, J. T., Sreedasyam, A., Schranz, M. E., Wilson, M., Carlson, J., Harkess, A., Emms, D., Goodstein, D. M., & Schmutz, J. (2022). GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *eLife*, 11, e78526. <https://doi.org/10.7554/eLife.78526>
- Lowe, S., Browne, M., Boudjelas, S., & De Poorter, M. (2000). *100 of the world's worst invasive alien species: A selection from the Global Invasive Species Database*. Invasive Species Specialist Group (ISSG). <http://www.iucngisd.org/gisd/pdf/100English.pdf>
- Luo, M. C., Deal, K. R., Akhunov, E. D., Akhunova, A. R., Anderson, O. D., Anderson, J. A., Blake, N., Clegg, M. T., Coleman-Derr, D., Conley, E. J., & Crossman, C. C. (2009). Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), 15780–15785.
- Lysak, M. A. (2022). Celebrating mendel, mcclintock, and darlington: On end-to-end chromosome fusions and nested chromosome fusions. *Plant Cell*, 34(7), 2475–2491. <https://doi.org/10.1093/plcell/koac116>
- Lysak, M. A., Mandakova, T., & Schranz, M. E. (2016). Comparative paleogenomics of crucifers: Ancestral genomic blocks revisited. *Current Opinion in Plant Biology*, 30, 108–115. <https://doi.org/10.1016/j.pbi.2016.02.001>
- Mandakova, T., Pouch, M., Brock, J. R., Al-Shehbaz, I. A., & Lysak, M. A. (2019). Origin and evolution of diploid and allopolyploid camelina genomes were accompanied by chromosome shattering. *Plant Cell*, 31(11), 2596–2612. <https://doi.org/10.1105/tpc.19.00366>
- Mandel, J. R., Dikow, R. B., Siniscalchi, C. M., Thapa, R., Watson, L. E., & Funk, V. A. (2019). A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. *Proceedings of the National Academy of Sciences of the United States of America*, 116(28), 14083–14088. <https://doi.org/10.1073/pnas.1903871116>
- Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., & Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PLoS Computational Biology*, 14(1), e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>
- Mari, A. (2023). *Allergome—A database of allergenic molecules*. www.allergome.org
- Mari, A., Rasi, C., Palazzo, P., & Scala, E. (2009). Allergen databases: Current status and perspectives. *Current Allergy and Asthma Reports*, 9(5), 376–383. <https://doi.org/10.1007/s11882-009-0055-9>
- Martin, M. D., Quiroz-Claros, E., Brush, G. S., & Zimmer, E. A. (2018). Herbarium collection-based phylogenetics of the ragweeds (*Ambrosia*, Asteraceae). *Molecular Phylogenetics and Evolution*, 120, 335–341. <https://doi.org/10.1016/j.ympev.2017.12.023>
- Martin, S. L., Lujan Toro, B., James, T., Sauder, C. A., & Laforest, M. (2022). Insights from the genomes of 4 diploid *Camelina* spp. G3: *Genes, Genomes, Genetics*, 12(12), jkac182. <https://doi.org/10.1093/g3journal/jkac182>
- Martin, S. L., Smith, T. W., James, T., Shalabi, F., Kron, P., & Sauder, C. A. (2017). An update to the Canadian range, abundance, and ploidy of *Camelina* spp. (Brassicaceae) east of the Rocky Mountains. *Botany*, 95(4), 405–417. <https://doi.org/10.1139/cjb-2016-0070>
- Mole, L. E., Goodfriend, L., Lapkoff, C. B., Kehoe, J. M., & Capra, J. D. (1975). The amino acid sequence of ragweed pollen allergen Ra5. *Biochemistry*, 14(6), 1216–1220. <https://doi.org/10.1021/bi00677a019>
- Montagnani, C., Gentili, R., Smith, M., Guarino, M. F., & Citterio, S. (2017). The worldwide spread, success, and impact of ragweed (*Ambrosia* spp.). *Critical Reviews in Plant Sciences*, 36(3), 139–178. <https://doi.org/10.1080/07352689.2017.1360112>
- Moretti, M. L., Van Horn, C. R., Robertson, R., Segobye, K., Weller, S. C., Young, B. G., Johnson, W. G., Douglas Sammons, R., Wang, D., Ge, X., d'Avignon, A., Gaines, T. A., Westra, P., Green, A. C., Jeffery, T., Lesperance, M. A., Tardif, F. J., Sikkema, P. H., Christopher Hall, J., ... Schulz, B. (2018). Glyphosate resistance in *Ambrosia trifida*: Part 2. Rapid response physiology and non-target-site resistance. *Pest Management Science*, 74(5), 1079–1088. <https://doi.org/10.1002/ps.4569>
- Morgan, M., Pagès, H., Obenchain, V., & Hayden, N. (2021). *Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import* (Version 2.10.0). <https://bioconductor.org/packages/Rsamtools>
- Mulligan, G. A. (1957). Chromosome numbers of Canadian weeds. I. *Canadian Journal of Botany*, 35(5), 779–789.
- Murat, F., Armero, A., Pont, C., Klopp, C., & Salse, J. (2017). Reconstructing the genome of the most recent common ancestor of flowering plants. *Nature Genetics*, 49(4), 490–496. <https://doi.org/10.1038/ng.3813>
- Murat, F., Xu, J. H., Tannier, E., Abrouk, M., Guilhot, N., Pont, C., Messing, J., & Salse, J. (2010). Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Research*, 20(11), 1545–1557. <https://doi.org/10.1101/gr.109744.110>
- Nakano, M., Hirakawa, H., Fukai, E., Toyoda, A., Kajitani, R., Minakuchi, Y., Itoh, T., Higuchi, Y., Kozuka, T., & Bono, H. (2021). A chromosome-level genome sequence of *Chrysanthemum seticuspe*, a model species for hexaploid cultivated chrysanthemum. *Communications Biology*, 4, 1167.

- Nandula, V. K., Tehranchian, P., Bond, J. A., Norsworthy, J. K., & Eubank, T. W. (2017). Glyphosate resistance in common ragweed (*Ambrosia artemisiifolia* L.) from Mississippi, USA. *Weed Biology and Management*, 17(1), 45–53.
- Nandula, V. K., Wright, A. A., R Van Horn, C., T Molin, W., Westra, P., & N Reddy, K. (2015). Glyphosate resistance in giant ragweed (*Ambrosia trifida* L.) from Mississippi is partly due to reduced translocation. *American Journal of Plant Sciences*, 06(13), 2104–2113. <https://doi.org/10.4236/ajps.2015.613211>
- Ostevik, K. L., Samuk, K., & Rieseberg, L. H. (2020). Ancestral reconstruction of karyotypes reveals an exceptional rate of nonrandom chromosomal evolution in sunflower. *Genetics*, 214(4), 1031–1045. <https://doi.org/10.1534/genetics.120.303026>
- Ou, S., Chen, J., & Jiang, N. (2018). Assessing genome assembly quality using the LTR Assembly Index (LAI). *Nucleic Acids Research*, 46(21), e126. <https://doi.org/10.1093/nar/gky730>
- Ou, S., & Jiang, N. (2018). LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiology*, 176(2), 1410–1422. <https://doi.org/10.1104/pp.17.01310>
- Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J. R. A., Hellings, A. J., Lugo, C. S. B., Elliott, T. A., Ware, D., Peterson, T., Jiang, N., Hirsch, C. N., & Hufford, M. B. (2019). Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology*, 20(1), 275. <https://doi.org/10.1186/s13059-019-1905-y>
- Page, E. R., Grainger, C. M., Laforest, M., Nurse, R. E., Rajcan, I., Bae, J., & Tardif, F. J. (2017). Target and non-target site mechanisms confer resistance to glyphosate in Canadian accessions of *Conyza canadensis*. *Weed Science*, 66, 234–245. <https://doi.org/10.1017/wsc.2017.69>
- Page, E. R., & Nurse, R. E. (2015). Cropping systems and the prevalence of giant ragweed (*Ambrosia trifida*): From the 1950's to present. *Field Crops Research*, 184, 104–111. <https://doi.org/10.1016/j.fcr.2015.09.013>
- Pagès, H., Aboyoun, P., Gentleman, R., & DebRoy, S. (2021). *Biostrings: Efficient manipulation of biological strings*. R package version 2.62.0. <https://bioconductor.org/packages/Biostrings>
- Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in {R}. *Bioinformatics*, 35, 526–528. <https://doi.org/10.1093/bioinformatics/bty633>
- Payne, W. W., Raven, P. H., & Kyhos, D. W. (1964). Chromosome numbers in Compositae. IV. Ambrosieae. *American Journal of Botany*, 51(4), 419–424.
- Pellestor, F., Gaillard, J. B., Schneider, A., Puechberty, J., & Gatinois, V. (2022). Chromoanagenesis, the mechanisms of a genomic chaos. *Seminars in Cell & Developmental Biology*, 123, 90–99. <https://doi.org/10.1016/j.semcdb.2021.01.004>
- Pellestor, F., & Gatinois, V. (2019). Chromothripsis, a credible chromosomal mechanism in evolutionary process. *Chromosoma*, 128(1), 1–6. <https://doi.org/10.1007/s00412-018-0679-4>
- R Core Team. (2021). *R: A language and environment for statistical computing* (Version 4.1.1 “Kick Things”). R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rambaut, A. (2006–2018). *FigTree* (Version 1.4.4). <http://github.com/rambaut/figtree/>
- Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., & Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7), 1665–1680. <https://doi.org/10.1016/j.cell.2014.11.021>
- Regnier, E. E., Harrison, S. K., Loux, M. M., Holloman, C., Venkatesh, R., Diekmann, F., Taylor, R., Ford, R. A., Stoltenberg, D. E., Hartzler, R. G., Davis, A. S., Schutte, B. J., Cardina, J., Mahoney, K. J., & Johnson, W. G. (2016). Certified crop advisors' perceptions of giant ragweed (*Ambrosia trifida*) distribution, herbicide resistance, and management in the corn belt. *Weed Science*, 64(2), 361–377. <https://doi.org/10.1614/ws-d-15-00116.1>
- Reznicek, A. A., Voss, E. G., & Walters, B. S. (2011). *Ambrosia artemisiifolia* L. Herbarium of The University of Michigan. <https://lsa-miflora-p.lsa.umich.edu/#/record/213>
- Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., & Mayrose, I. (2015). The Chromosome Counts Database (CCDB) – A community resource of plant chromosome numbers. *New Phytologist*, 206(1), 19–26. <https://doi.org/10.1111/nph.13191>
- Sammons, R. D., & Gaines, T. A. (2014). Glyphosate resistance: State of knowledge. *Pest Management Science*, 70(9), 1367–1377. <https://doi.org/10.1002/ps.3743>
- Schaffner, U., Steinbach, S., Sun, Y., Skjøth, C. A., de Weger, L. A., Lommen, S. T., Augustinus, B. A., Bonini, M., Karrer, G., & Šikoparija, B. (2020). Biological weed control to relieve millions from *Ambrosia* allergies in Europe. *Nature Communications*, 11, 1745.
- Schliep, K., Jombart, T., Kamvar, Z. N., Archer, E., & Harris, R. (2020). *apex: Phylogenetic methods for multiple gene data*. R package version 1.0.4. <https://CRAN.R-project.org/package=apex>
- Schliep, K., Potts, A. J., Morrison, D. A., & Grimm, G. W. (2017). Inter-twining phylogenetic trees and networks. *Methods in Ecology and Evolution*, 8, 1212–1220. <https://doi.org/10.1111/2041-210X.12760>
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, 27(4), 592–593.
- Schnable, J. C., Springer, N. M., & Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10), 4069–4074. <https://doi.org/10.1073/pnas.1101368108>
- Schranz, M. E., Lysak, M. A., & Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the Brassicaceae: Building blocks of crucifer genomes. *Trends in Plant Science*, 11(11), 535–542. <https://doi.org/10.1016/j.tplants.2006.09.002>
- Schulz, A. (2021). *pBrackets: Plot brackets*. R package version 1.0.1. <https://CRAN.R-project.org/package=pBrackets>
- Semple, J. C., & Watanabe, K. (2009). A review of chromosome numbers in Asteraceae with hypotheses on chromosomal base number evolution. In V. A. Funk, A. Susanna, T. F. Stuessy, & R. J. Bayer (Eds.), *Systematics, evolution, and biogeography of Compositae* (pp. 61–72). International Association for Plant Taxonomy.
- Smit, A., & Hubley, R. (2008–2015). *RepeatModeler Open-1.0*. <http://www.repeatmasker.org>
- Smit, A., Hubley, R., & Green, P. (2013–2015). *RepeatMasker Open-4.0*. <http://www.repeatmasker.org>
- Smith, M., Cecchi, L., Skjøth, C., Karrer, G., & Šikoparija, B. (2013). Common ragweed: A threat to environmental health in Europe. *Environment International*, 61, 115–126.
- Smith, T. W., Kron, P., & Martin, S. L. (2018). flowPloidy: An R package for genome size and ploidy assessment of flow cytometry data. *Applications in Plant Sciences*, 6(7), e01164. <https://doi.org/10.1002/aps3.1164>

- Soetaert, K. (2020). *diagram: Functions for visualising simple graphs (networks), plotting flow diagrams*. R package version 1.6.5. <https://CRAN.R-project.org/package=diagram>
- Song, X., Wei, Y., Xiao, D., Gong, K., Sun, P., Ren, Y., Yuan, J., Wu, T., Yang, Q., Li, X., Nie, F., Li, N., Feng, S., Pei, Q., Yu, T., Zhang, C., Liu, T., Wang, X., & Yang, J. (2021). *Brassica carinata* genome characterization clarifies U's triangle model of evolution and polyploidy in Brassica. *Plant Physiology*, 186(1), 388–406. <https://doi.org/10.1093/plphys/kiab048>
- Strother, J. L. (1993). *Ambrosia*. http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=101325
- Sun, H., Wu, S., Zhang, G., Jiao, C., Guo, S., Ren, Y., Zhang, J., Zhang, H., Gong, G., Jia, Z., Zhang, F., Tian, J., Lucas, W. J., Doyle, J. J., Li, H., Fei, Z., & Xu, Y. (2017). Karyotype stability and unbiased fractionation in the paleo-allotetraploid *Cucurbita* genomes. *Molecular Plant*, 10(10), 1293–1306. <https://doi.org/10.1016/j.molp.2017.09.003>
- Sun, P., Jiao, B., Yang, Y., Shan, L., Li, T., Li, X., Xi, Z., Wang, X., & Liu, J. (2021). *WGDI: A user-friendly toolkit for evolutionary analyses of wholegenome duplications and ancestral karyotypes*. bioRxiv. <https://doi.org/10.1101/2021.04.29.441969>
- Taylor, J. B., Loux, M. M., Harrison, S. K., & Regnier, E. (2002). Response of ALS-resistant common ragweed (*Ambrosia artemisiifolia*) and giant ragweed (*Ambrosia trifida*) to ALS-inhibiting and alternative herbicides. *Weed Technology*, 16(4), 815–825. [https://doi.org/10.1614/0890-037X\(2002\)016\[0815:ROARCR\]2.0.CO;2](https://doi.org/10.1614/0890-037X(2002)016[0815:ROARCR]2.0.CO;2)
- Thomas, B. C., Pedersen, B., & Freeling, M. (2006). Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Genome Research*, 16(7), 934–946. <https://doi.org/10.1101/gr.4708406>
- Tomasello, S., Stuessy, T. F., Oberprieler, C., & Heubl, G. (2019). Ragweeds and relatives: Molecular phylogenetics of Ambrosiinae (Asteraceae). *Molecular Phylogenetics and Evolution*, 130, 104–114. <https://doi.org/10.1016/j.ympev.2018.10.005>
- Van Horn, C. R., Moretti, M. L., Robertson, R. R., Segobye, K., Weller, S. C., Young, B. G., Johnson, W. G., Schulz, B., Green, A. C., Jeffery, T., Lesperance, M. A., Tardif, F. J., Sikkema, P. H., Hall, J. C., McLean, M. D., Lawton, M. B., Sammons, R. D., Wang, D., Westra, P., & Gaines, T. A. (2018). Glyphosate resistance in *Ambrosia trifida*: Part I. Novel rapid cell death response to glyphosate. *Pest Management Science*, 74(5), 1071–1078. <https://doi.org/10.1002/ps.4567>
- Van Wely, A. C., Soltani, N., Robinson, D. E., Hooker, D. C., Lawton, M. B., & Sikkema, P. H. (2015). Glyphosate and acetolactate synthase inhibitor resistant common ragweed (*Ambrosia artemisiifolia* L.) in southwestern Ontario. *Canadian Journal of Plant Science*, 95(2), 335–338. <https://doi.org/10.4141/cjps-2014-290>
- Vincent, G., & Cappadocia, M. (1987). Interspecific hybridization between common ragweed (*Ambrosia artemisiifolia*) and giant ragweed (*A. trifida*). *Weed Science*, 35(5), 633–636.
- Vink, J. P., Soltani, N., Robinson, D. E., Tardif, F. J., Lawton, M. B., & Sikkema, P. H. (2012). Occurrence and distribution of glyphosate-resistant giant ragweed (*Ambrosia trifida* L.) in southwestern Ontario. *Canadian Journal of Plant Science*, 92(3), 533–539. <https://doi.org/10.4141/cjps2011-249>
- Volenberg, D. S., Rayburn, A. L., Zheng, D., & Tranel, P. (2005). *Natural hybridization between giant and common ragweed*. North Central Weed Science Society.
- Wagner, W. H. (1958). The hybrid ragweed, *Ambrosia artemisiifolia* × *trifida*. *Rhodora*, 60(720), 309–316.
- Wang, P., Luo, Y., Huang, J., Gao, S., Zhu, G., Dang, Z., Gai, J., Yang, M., Zhu, M., Zhang, H., Ye, X., Gao, A., Tan, X., Wang, S., Wu, S., Cahoon, E. B., Bai, B., Zhao, Z., Li, Q., ... Chen, Y. (2020). The genome evolution and domestication of tropical fruit mango. *Genome Biology*, 21(1), 60. <https://doi.org/10.1186/s13059-020-01959-8>
- Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J. H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J. C., Paterson, A. H., Chalhoub, B., ... The Brassica rapa Genome Sequencing Project Consortium. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, 43(10), 1035–1039. <https://doi.org/10.1038/ng.919>
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., Lee, T. H., Jin, H., Marler, B., Guo, H., Kissinger, J. C., & Paterson, A. H. (2012). MCSanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, 40(7), e49. <https://doi.org/10.1093/nar/gkr1293>
- Waterhouse, R. M., Seppey, M., Simao, F. A., Manni, M., Ioannidis, P., Kliuchnikov, G., Kriventseva, E. V., & Zdobnov, E. M. (2017). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, 35(3), 543–548. <https://doi.org/10.1093/molbev/msx319>
- Wayne, P., Foster, S., Connolly, J., Bazzaz, F., & Epstein, P. (2002). Production of allergenic pollen by ragweed (*Ambrosia artemisiifolia* L.) is increased in CO₂-enriched atmospheres. *Annals of Allergy, Asthma & Immunology*, 88(3), 279–282. [https://doi.org/10.1016/S1081-1206\(10\)62009-1](https://doi.org/10.1016/S1081-1206(10)62009-1)
- Wendel, J. F., Lisch, D., Hu, G., & Mason, A. S. (2018). The long and short of doubling down: Polyploidy, epigenetics, and the temporal dynamics of genome fractionation. *Current Opinion in Genetics & Development*, 49, 1–7. <https://doi.org/10.1016/j.gde.2018.01.004>
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1), 1–29. <https://www.jstatsoft.org/v40/i01/>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Wickham, H., François, R., & Müller, H. L. (2022). *dplyr: A grammar of data manipulation*. R package version 1.0.10. <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., & Seidel, D. (2022). *scales: Scale function for visualization*. R package version 1.2.1. <https://CRAN.R-project.org/package=scales>
- Woodhouse, M. R., Cheng, F., Pires, J. C., Lisch, D., Freeling, M., & Wang, X. (2014). Origin, inheritance, and gene regulatory consequences of genome dominance in polyploids. *Proceedings of the National Academy of Sciences of the United States of America*, 111(14), 5283–5288. <https://doi.org/10.1073/pnas.1402475111>
- Wylie, R. B. (1915). A hybrid ragweed. *Proceedings of the Iowa Academy of Science*, 22(1), 127–128.
- Xian, X., Zhao, H., Wang, R., Huang, H., Chen, B., Zhang, G., Liu, W., & Wan, F. (2023). Climate change has increased the global threats posed by three ragweeds (*Ambrosia* L.) in the Anthropocene. *Science of the Total Environment*, 859(Pt 2), 160252. <https://doi.org/10.1016/j.scitotenv.2022.160252>
- Xu, Z., & Wang, H. (2007). LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Research*, 35(Suppl_2), W265–W268. <https://doi.org/10.1093/nar/gkm286>

- Yang, J., Liu, D., Wang, X., Ji, C., Cheng, F., Liu, B., Hu, Z., Chen, S., Pental, D., Ju, Y., Yao, P., Li, X., Xie, K., Zhang, J., Wang, J., Liu, F., Ma, W., Shopan, J., Zheng, H., . . . Zhang, M. (2016). The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. *Nature Genetics*, 48(10), 1225–1232. <https://doi.org/10.1038/ng.3657>
- Yen, E. C., McCarthy, S. A., Galarza, J. A., Generalovic, T. N., Pelan, S., Nguyen, P., Meier, J. I., Warren, I. A., Mappes, J., Durbin, R., & Jiggins, C. D. (2020). A haplotype-resolved, de novo genome assembly for the wood tiger moth (*Arctia plantaginis*) through trio binning. *Gigascience*, 9(8), gaaa088. <https://doi.org/10.1093/gigascience/gaaa088>
- Zhang, D., Pan, Q., Cui, C., Tan, C., Ge, X., Shao, Y., & Li, Z. (2015). Genome-specific differential gene expressions in resynthesized *Brassica* allotetraploids from pair-wise crosses of three cultivated diploids revealed by RNA-seq. *Frontiers in Plant Science*, 6, 957. <https://doi.org/10.3389/fpls.2015.00957>
- Zhao, Q., Meng, Y., Wang, P., Qin, X., Cheng, C., Zhou, J., Yu, X., Li, J., Lou, Q., Jahn, M., & Chen, J. (2021). Reconstruction of ancestral karyotype illuminates chromosome evolution in the genus *Cucumis*. *Plant Journal*, 107(4), 1243–1259. <https://doi.org/10.1111/tpj.15381>
- Zimin, A., Stevens, K. A., Crepeau, M. W., Holtz-Morris, A., Koriabine, M., Marçais, G., Puiu, D., Roberts, M., Wegrzyn, J. L., de Jong, P. J., Neale, D. B., Salzberg, S. L., Yorke, J. A., & Langley, C. H. (2014). Sequencing and assembly of the 22-gb loblolly pine genome. *Genetics*, 196(3), 875–890. <https://doi.org/10.1534/genetics.113.159715>
- Ziska, L., Knowlton, K., Rogers, C., Dalan, D., Tierney, N., Elder, M. A., Filley, W., Shropshire, J., Ford, L. B., Hedberg, C., Fleetwood, P., Hovanky, K. T., Kavanaugh, T., Fulford, G., Vrtis, R. F., Patz, J. A., Portnoy, J., Coates, F., Bielory, L., & Frenz, D. (2011). Recent warming by latitude associated with increased length of ragweed pollen season in central North America. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10), 4248–4251. <https://doi.org/10.1073/pnas.1014107108>
- Ziska, L. H., & Beggs, P. J. (2012). Anthropogenic climate change and allergen exposure: The role of plant biology. *Journal of Allergy and Clinical Immunology*, 129(1), 27–32. <https://doi.org/10.1016/j.jaci.2011.10.032>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Laforest, M., Martin, S. L., Bisailon, K., Soufiane, B., Meloche, S., Tardif, F. J., & Page, E. (2024). The ancestral karyotype of the Heliantheae Alliance, herbicide resistance, and human allergens: Insights from the genomes of common and giant ragweed. *The Plant Genome*, e20442. <https://doi.org/10.1002/tpg2.20442>