Plant Biotechnology Journal

WILEY

RESEARCH ARTICLE OPEN ACCESS

Genome and Single-Cell Transcriptome Reveal the Evolution of Holoparasitic Plants: A Case Study of *Cistanche deserticola*

Xinke Zhang[1] | Yujing Miao[1,2] | Yuanyuan Xing[1] | Guoshuai Zhang[1] | Meiqiao Zhang[3] | Chris J. Thorogood[4] | Shilin Chen[5] | Linfang Huang[1]

[1]Key Laboratory of Chinese Medicine Resources Conservation, State Administration of Traditional Chinese Medicine of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China | [2]National key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China | [3]College of Horticulture, China Agricultural University, Beijing, China | [4]Department of Plant Sciences, University of Oxford, Oxford, UK | [5]Chengdu University of Traditional Chinese Medicine, Chengdu, China

Correspondence: Shilin Chen (slchen@icmm.ac.cn) | Linfang Huang (lfhuang@implad.ac.cn)

Received: 11 July 2025 | **Revised:** 14 October 2025 | **Accepted:** 31 October 2025

Funding: This work was supported by the National Natural Science Foundation of China (82211540726, 82073960 and 82274045); the Open Fund of the State Key Laboratory of Southwestern Chinese Medicine Resources (SCMR2022015); the Beijing Municipal Natural Science Foundation (24G40321); and the China Postdoctoral Science Foundation (2024M763263).

Keywords: *Cistanche deserticola* | gene loss | genome | holoparasitic plants | parasitism-related cells | single-cell transcriptome

ABSTRACT

The Orobanchaceae family, the largest group of parasitic plants, spans a complete spectrum from autotrophic to holoparasitic species. As a typical endangered holoparasitic species within this family, *Cistanche deserticola* is a parasitic plant that is widely harvested for traditional medicine in desertic regions, and of growing importance as a cash crop. However, the evolution of *C. deserticola* at the molecular and cellular level is poorly understood. Here, we constructed the first chromosome-level genome map of *C. deserticola*. Comparative genomic analyses demonstrated that the *C. deserticola* genome exhibited a substantial loss of genes related to photosynthesis and immunity (21.58% of the total genes) and contained 115 horizontally transferred genes. This suggested that the genomic evolution of holoparasitic plants was driven by the interplay between the acquisition of functional genes and the loss of genes specific to plant tissues or functions. Additionally, parasitism-related cells were identified using a high-resolution single-cell transcriptomic atlas, revealing stage-specific differentiation during the parasitic process. Early cells (cluster 11) highly expressed dopamine/tyrosine metabolism pathways genes (e.g., polyphenol oxidase), driving phenylethanoid glycoside biosynthesis. By contrast, mature cells (cluster 10) show high levels of gene expression relating to carbohydrate metabolism in association with nutrient acquisition. Connecting these insights, we developed a comprehensive *C. deserticola* database to integrate multi-omics and ecological data (http://60.30.67.246:7006/Home). This builds a robust molecular foundation for exploring pathways to parasitism in plants more broadly.

Xinke Zhang and Yujing Miao contributed equally to this work.

1 | Introduction

Parasitic plants account for approximately 1% of all angiosperms and evolved independently at least 12 times (Nickrent 2020). These plants extract water and nutrients from host plants. Hemiparasitic plants have retained chlorophyll and the ability to photosynthesise; holoparasites, on the other hand, have completely lost the ability to photosynthesise and lack functional leaves and roots (Westwood et al. 2010). As the largest family of parasitic plants, the Orobanchaceae is unusual in encompassing a rich spectrum spanning fully autotrophic plants, hemiparasites, and holoparasites (Masumoto et al. 2021)—a diversity linked closely to regional and climatic adaptations. While parasitism evolved once in the family, the multiple independent evolutions of holoparasitism in Orobanchaceae further complicate its phylogeny and evolutionary history (Fu et al. 2017), making this an ideal family with which to explore the evolution of parasitism.

The evolution of parasitic plants has been associated with gene loss, mirroring their shift to heterotrophy. During this process, numerous genes associated with photosynthesis, nutrient synthesis or autotrophic functions are either lost or rendered nonfunctional (Gruzdev et al. 2019). Additionally, parasitic plants can acquire new genes through horizontal gene transfer (HGT). For example, root-parasitic *Orobancha aegyptiaca* and stem-parasitic *Cuscuta australis* obtained Brassicaceae-specific strictosidine synthase-like genes by HGT (Zhang et al. 2014). Indeed, the genetic functional integration of parasitic plants with their hosts relies on HGT (Yang et al. 2015). Taken together, both gene loss and HGT are essential drivers in parasitic plant evolution. They provide a molecular foundation for the evolution of parasitism, enabling organisms to acquire novel functions and survive across a range of diverse ecological niches linked to those of their hosts. However, the exact extent and patterns of gene loss within the genomes of holoparasitic plants remain unclear. Significant differences exist among various holoparasitic plants regarding the quantity, function, and underlying mechanisms of gene loss, which likely reflect variations in their parasitic strategies, host preferences, and evolutionary histories. Although HGT is widely recognised as a critical adaptive mechanism for parasitic plants, systematic studies on the spatial distribution of horizontally transferred genes (e.g., the localisation within specific chromosomes or cell types) and their functional integration into the holoparasitic plants genome are still lacking. Moreover, at the cellular level, little is known about how holoparasitic plants efficiently acquire and regulate horizontally transferred genes. In particular, the functional specialisation of different cell types in mediating resource acquisition and adaptive regulation during parasitism remains poorly understood and requires further investigation.

Recent advances in genomics have provided tools for uncovering adaptive mechanisms and molecular regulatory networks in parasitic plants. High-quality genome sequencing enables researchers to comprehensively analyse gene loss and HGT in parasitic plants with varying degrees of parasitism (Chen et al. 2023). Meanwhile, the emergence of single-cell transcriptomics offers a new platform for deciphering the functional heterogeneity of cells and tissues (Jean-Baptiste et al. 2019). Nevertheless, cellular-level regulatory mechanisms in parasitic plants are still poorly understood, and systematic studies using single-cell sequencing are absent. The complex tissue structures of parasitic plants and the intricate interactions with hosts pose substantial challenges to deciphering the differentiation and functions of parasitism-related cells at the single-cell level.

Cistanche deserticola, commonly known as Desert Ginseng, is a desert plant in northwest China that has a long history of use in Traditional Chinese Medicine (Li et al. 2015). *C. deserticola* parasitises desert shrubs (e.g., *Haloxylon ammodendron*) and is becoming an increasingly important local cash crop, providing both food and medicine (Shen et al. 2019). Phenylethanoid glycoside isolated from the stem of *C. deserticola* has various pharmacological properties, including neuroprotection, immune system modulation, anti-ageing, anti-inflammatory, anti-osteoporosis, liver protection, antioxidant and antibacterial properties. As a holoparasite with important economic potential, it is an ideal candidate for genomic and single-cell work. Previous studies have shown that the plastid genome of *C. deserticola* has undergone extensive gene loss (Miao et al. 2022), especially from the plastid genome (Fan et al. 2023). However, due to the lack of a high-quality reference genome, the patterns of gene loss in the nuclear genome of *C. deserticola* remain poorly characterised, and the precise localisation and mechanisms of horizontally transferred genes have yet to be thoroughly explored. Furthermore, studies at the cellular level are also notably limited. Owing to the absence of high-resolution single-cell sequencing datasets, our understanding of the specific roles of different cell populations during the parasitic process remains minimal. These knowledge gaps severely constrain a comprehensive understanding of the molecular and cellular mechanisms underlying parasitism in *C. deserticola*.

Here, we present a high-quality reference genome of *C. deserticola* together with single-cell atlases from the stem to elucidate the genetic control and parasitic adaptation mechanisms in *C. deserticola*. Our objectives are threefold: (1) to assemble a chromosome-scale *C. deserticola* genome; (2) to construct a cell atlas using single-cell transcriptomics; and (3) to develop a *C. deserticola* database to integrate the multi-omics data (http://60.30.67.246:7006/Home). This research enhances the understanding of the parasitic adaptation mechanisms and offers important scientific groundwork for developing and conserving parasitic medicinal plants.

2 | Results

2.1 | Genome Assembly and Annotation of *C. deserticola*

The genome size of *C. deserticola* was estimated to be around 5.27 Gb through k-mer analysis (Figure S1 and Table S1). Based on 332.36 Gb of Illumina, 181.26 Gb of PacBio, and 795.95 Gb of Hi-C reads (Table S2), we assembled a draft genome for *C. deserticola* (5.43 Gb in size, contig N50 of 89.36 Mb) (Table S1). Through Hi-C-assisted assembly, 5.33 Gb of the genome sequence was mapped to 21 chromosomes, accounting for 98.19% of the total sequence. The sequence lengths of the assembled 21 chromosomes were highly correlated with the physical chromosomal sizes measured by karyotyping ($R^2 = 0.97$; Figure S1),

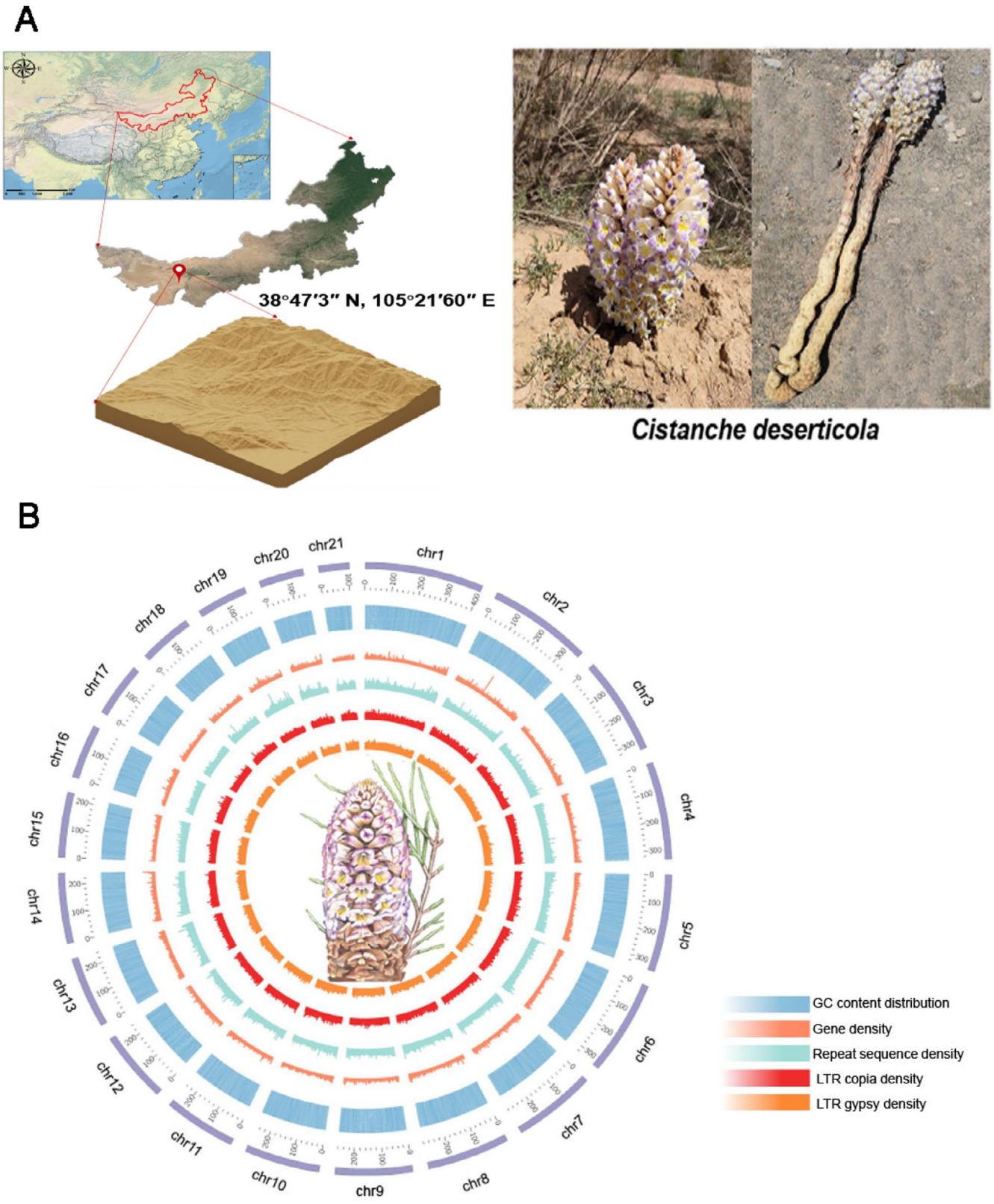

FIGURE 1 | Habitat and genomic landscape of *Cistanche deserticola*. (A) Habitat and plant morphology of *C. deserticola*. (B) Genomic landscape of *C. deserticola*. Inside to outside: Circle 1: LTR gypsy density; Circle 2: LTR copia density; Circle 3: Repeat sequence density; Circle 4: Gene density; Circle 5: GC content distribution.

spanning from 404.89 Mb to 103.78 Mb with a scaffold N50 of 263.24 Mb (Figure 1B and Table S3). The assembled genomes showed good quality BUSCO completeness (96.5%), LAI index (17.29%), QV values (43.23%) and read mapping rates from both second-generation and third-generation sequencing data (99.9%/100%) (Tables S4–S6). In total, we annotated 42 738

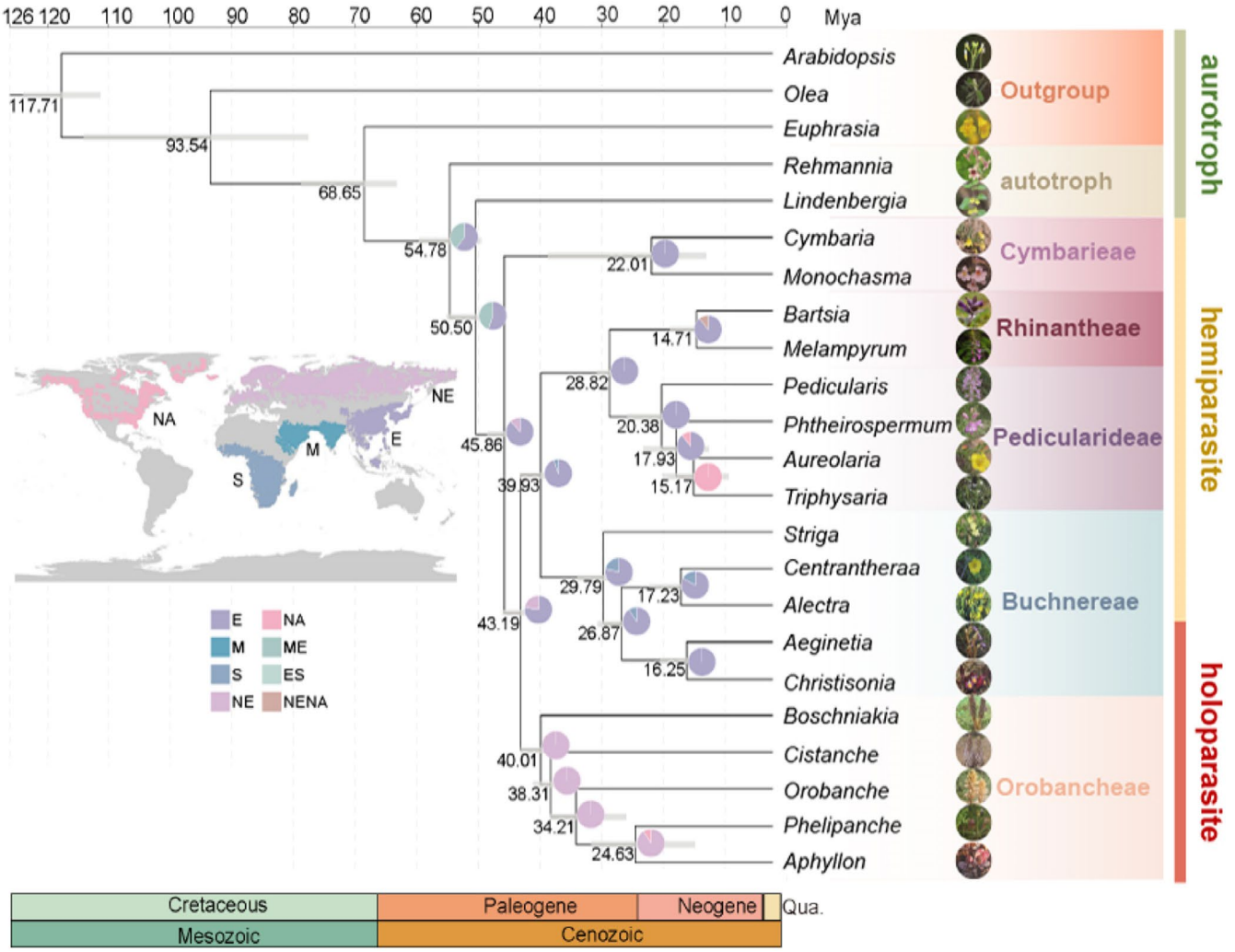

FIGURE 2 | Phylogenetic relationships and historical biogeography of Orobanchaceae. Low-copy orthogroups from 20 Orobanchaceae species and three outgroup species were used to build the phylogenetic tree. The numbers on the internal nodes show divergence times, and the grey bars indicate error ranges. Each node contains a pie chart that suggests potential ancestral distributions. E., East Asia and Southeast Asia; M., Middle East, South Asia; S., South Central Africa; NE., Northern Eurasia; NA., North America.

protein-coding genes (Table S7). Of these, a significant 87.12% (37 234 genes) have been annotated in at least one of the databases: KEGG, InterPro, NR, GO, SwissUniprot and Pfam (Table S8). Additionally, we have identified a total of 190 miRNAs, 2305 tRNAs, 6353 rRNAs and 6873 snRNAs. These noncoding RNAs span a range of lengths and collectively account for a small but significant portion of the genome (Table S9).

2.2 | Phylogenetic Relationships and Ancestral Reconstruction of the Orobanchaceae

Autotrophic *Rehmannia* was the most basal branch of Orobanchaceae. Another autotrophic (*Lindenbergia*) diverged from other parasitic plants 50.50 million years ago (Mya) (Figure 2). All parasitic plants can be divided into five groups. Within this classification, the holoparasitic lineage Orobanchaceae separated from other parasitic lineages 40.01 Mya ago. Buchnereae was the closest relative to all other hemiparasitic lineages, including species with varying degrees of parasitism. The holoparasitic species (*Christisonia*, *Aeginetia* and *Alectra*) diverged from other hemiparasitic species about 16.25 Mya, which signified another independent origin of holoparasitism in the Orobanchaceae family. In addition, ancestral character state reconstruction for Orobanchaceae suggested that autotrophy was the ancestral state (Figure S2). Both hemiparasitism and holoparasitism were derived from autotrophy (Figure S3).

The BioGeoBEARS results for the best model (DIVALIKE+J) indicated that the Orobanchaceae family likely originated from the Middle East, South Asia, East Asia and Southeast Asia. The ancestors of parasitic species were mainly distributed in East Asia and Southeast Asia (Figure 2). Subsequently, the holoparasitic lineage of Orobanchaceae migrated northward to the Eurasian continent and North America and spread westward into North America approximately 24.63 Mya. Another branch of holoparasitic species, Buchnereae, originated in East Asia and Southeast Asia, with some descendants migrating southward to South Central Africa (Figure S4). According to the biogeographic stochastic mapping (BSM) analysis, diversification within the family was primarily driven by within-area

speciation, which accounted for 54.43% of the total speciation events, followed by dispersal events (37.17%). This indicates that the diversity observed in the Orobanchaceae was attributable not only to local adaptation but also to geographic migration. Only a few events involved a subset (4.33%) and vicariance (4.06%) (Table S10).

2.3 | Evolution of Gene Families and Genome Size of Parasitic Plants

Parasitic plants show extreme morphological divergence, alongside which gene families may have undergone unique patterns of expansion and contraction (Sun et al. 2018). To reveal the changes in the gene family sizes, eight species with reference genomes (holoparasitic: *C. deserticola, Phelipanche aegyptiaca* and *Orobanche cumana*; hemiparasitic: *Pedicularis kansuensis, Phtheirospermum japonicum* and *Striga asiatica*; autotrophy: *Lindenbergia luchunensis* and *Rehmannia glutinosa*) were chosen for analysis of gene family expansion and contraction. According to our results, most gene families underwent substantial contraction during the initial phase of parasitism and holoparasitism, while only a small portion exhibited expansion (Figure S5). The results were similar to those of previous research (Xu et al. 2022). By calculating the F-index, we found that about 68%–77% of conserved gene families in hemiparasitic species had gene counts below the average, while in holoparasitic species, approximately 60%–89% of the conserved gene families had fewer genes than the average (Figure S5). This suggests that as the degree of parasitism increased, gene families shrank in size and number. In addition, GO enrichment analysis was carried out for the gene families that showed expansion and contraction (Figure S5). When parasitism first appeared, the primary enrichment of contracted gene families occurred in biological rhythm regulation and intercellular communication signal transduction. The expanded gene families primarily participated in energy metabolism (such as ATP synthesis and energy gradient establishment), gene regulation and transmembrane transport of substances. As parasitism becomes more derived, the contracted gene families are further enriched in photosynthesis, transcriptional regulation, immune response and secondary metabolite synthesis pathways. The expanded gene families, initially enriched in pathways such as the mitochondrial protein-containing complex, also become increasingly enriched in new pathways such as those involved in serine-tRNA ligase and aspartic-type endopeptidase activity.

The Orobanchaceae shows considerable variation in genome size, a trait potentially linked to a diversity in ecology and life history. An increase in the degree of heterotrophy among species within the Orobanchaceae may correlate with a trend towards larger genome size (Lyko and Wicke 2021). *C. deserticola* has the largest genome size of 5.4 Gb among the eight species examined, while *L. luchunensis* has the smallest at 0.2 Gb, showing a 27-fold difference. This significant difference in genome size is likely attributable to an increase in repetitive elements (REs) and, in certain instances, whole-genome duplication (WGD). To understand what influences genome expansion in these parasitic plants, we compared the REs among the eight species. An increase in the number of REs was the main reason for the large genome. The largest genome, *C. deserticola*, consisted of 93.4% REs (Figure S6). The long terminal repeat (LTR) family constituted the main portion of these REs at 84.22%, with Gypsy and Copia contributing 43.42% and 23.28%, respectively (Table S11). The analysis of the insertion timing of the LTR indicated a markedly accelerated accumulation within the *C. deserticola* genome over the past million years. Copia insertions peaked around 0.54 Mya, while Gypsy insertions expanded more recently, about 0.42 Mya (Figure S6). Considering the crucial role of WGD in increasing genome size, we investigated whether *C. deserticola* had undergone any WGD events. Based on previous studies, *Erythranthe guttatus* shared a WGD event with Orobanchaceae. Therefore, we selected *E. guttatus* as an outgroup to compare with *C. deserticola*. We first identified WGD events in *C. deserticola* and *E. guttatus* through synteny analysis, revealing that both species underwent one WGD event (Figure S7). Furthermore, the Ks peak corresponding to species divergence was smaller than the Ks peaks of duplication events, indicating that the shared WGD event occurred prior to species divergence (Figure S7). The findings suggested that throughout evolutionary history, there was one shared WGT event that transpired in *C. deserticola*.

2.4 | Gene Loss Patterns in Parasitic Plants

Extensive gene loss is a major trait of parasitic plants (Yoshida and Kee 2021). We first identified 11 922 conserved orthogroups present in autotrophic plants to facilitate the analysis of gene loss in parasitic plants. According to our research, there was a positive association between the extent of gene loss and the degree of parasitism. In hemiparasitic plants, the rate of conserved gene loss ranged from 3.30% to 5.45%, whereas in holoparasitic species, it was significantly higher, ranging from 10.83% to 21.58% (Figure 3A). Among holoparasitic species, *C. deserticola* exhibited the highest number of orthogroup losses, with its loss of unique orthogroups significantly exceeding that of other species. The holoparasitic species possessed a total of 479 conserved orthogroups, whereas hemiparasitic species shared 10 conserved orthogroups. Notably, only 5 conserved orthogroups were common to both parasitic plant types, suggesting that the evolution from hemiparasitism to holoparasitism was associated with the loss of specific gene families (Figure 3A).

We conducted a functional enrichment analysis of the lost genes to delve deeper into the patterns of gene loss in parasitic plants. We found that photosynthesis, flowering regulation, response to nutrient levels, immune response, cellular development, external stimuli response and reproductive processes were most significant. Specifically, hemiparasitic plants showed a high rate of gene loss in the cellular development pathway. In holoparasitic plants, gene loss predominantly occurs in fundamental pathways essential for autonomous survival, particularly in photosynthesis, flowering regulation and immune response pathways (Figure 3B). For other pathways, the differences between the two groups were not significant. Compared to hemiparasitic plants, holoparasitic plants have lost photosynthetic capacity because they are fully dependent on their hosts, leading to a substantial reduction in plastid genomes, characterised by extensive gene

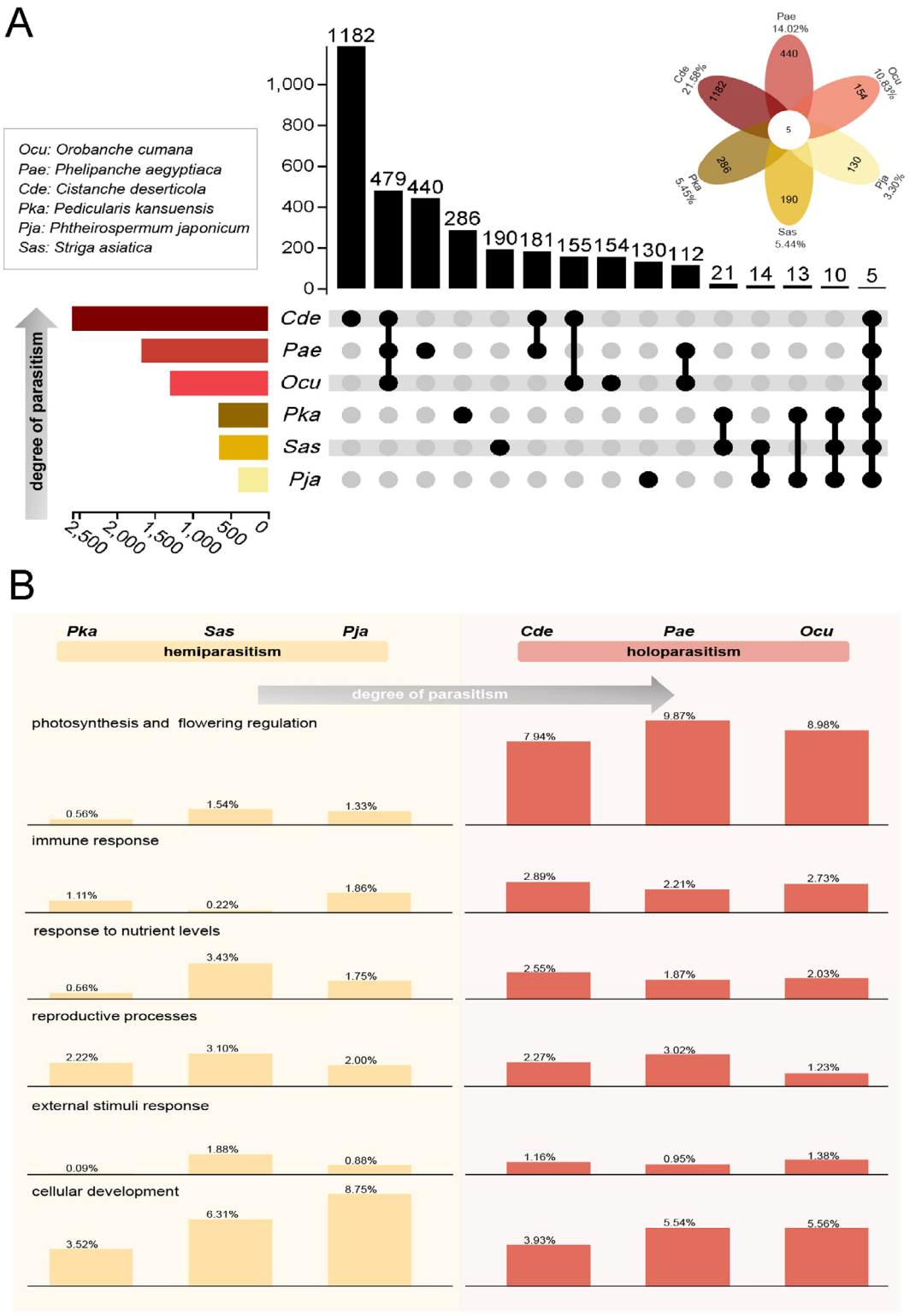

FIGURE 3 | Characterising gene losses in hemiparasitic and holoparasitic plants. (A) The number of commonly and specifically lost orthogroups in hemiparasitic and holoparasitic plants. Under the species name, the number shows the proportion of all lost orthogroups in that species. (B) Convergent and functional-biassed gene loss in parasitic plants. Above each bar, the percentage reflects the fraction of lost genes in that functional category relative to all lost genes.

loss (Chen et al. 2020). Therefore, we examined plastid gene loss in three holoparasitic plants. The findings confirmed that the plastids of the three holoparasitic species had undergone significant gene loss, with *P. aegyptiaca* losing the most and *C. deserticola* the least (Figure S8). Most of these lost genes are related to photosynthesis and self-replication (Figure S8). *C. deserticola* exhibited the loss of two plastid-specific genes, *atpH* and *atpI*, which are responsible for encoding subunits of ATP synthase, hinting at a distinct evolutionary pathway for the plastid genome in *C. deserticola*.

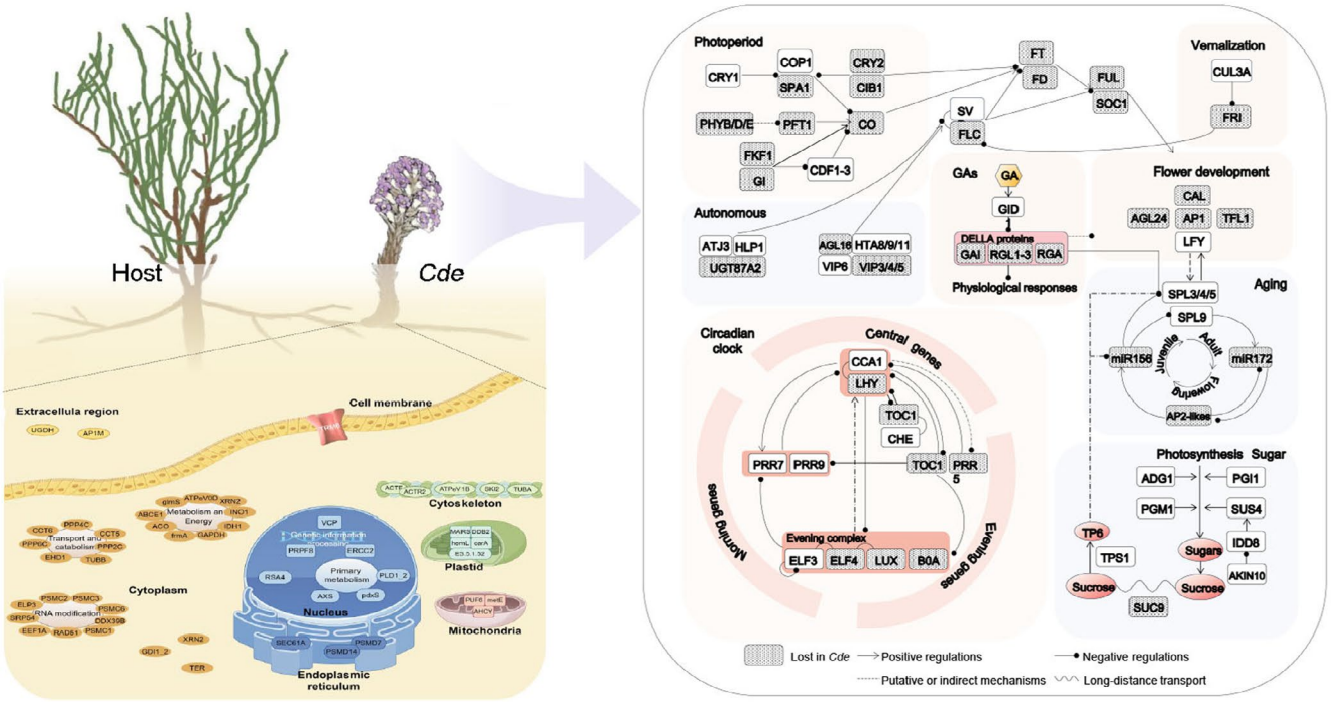

FIGURE 4 | Illustration of horizontal gene transfer and loss of genes associated with the flowering time regulation in *C. deserticola*. The left panel illustrates the parasitic interaction between the host and *C. deserticola*. Distinctly coloured regions indicate the localisation of proteins encoded by horizontally transferred genes within the cellular structure (by Figdraw, ID: TRWRO15a33). The right panel displays the gene loss of a simplified gene network controlling flowering development and regulation in *C. deserticola*. Genes depicted within grey boxes represent lost genes. *Cde: C. deserticola*.

2.5 | Specific Gene Loss and Acquisition in *C. deserticola*

In contrast to autotrophic plants, holoparasitic plants are devoid of leaves and are unable to directly regulate flowering through conventional pathways (Song et al. 2015). Previous research has demonstrated that holoparasitic plants frequently exhibit a loss of genes associated with flowering (Yoshida et al. 2019). Nevertheless, holoparasitic plants flower to reproduce, indicating the evolution of a distinct regulatory mechanism for this process. Using the FLOR-ID database, we examined the genes that control flowering in *C. deserticola* (Bouché et al. 2016). Interestingly, our findings indicate a near-complete loss of genes associated with the photoperiod, vernalisation, flower development, gibberellin receptors and circadian clock pathways in *C. deserticola* (Figure 4). Notably, key flowering genes such as *FLC*, *FT* and *SOC1*, which are integral components of the flowering regulatory network (Li et al. 2022), were absent in *C. deserticola*. Conversely, genes associated with the sugar, aging, and autonomous pathways were relatively less affected, with the sugar pathway genes being the most conserved, as only *SUC9* was lost. This preservation suggests that the sugar pathway may still contribute to the growth and development of *C. deserticola*, particularly in the regulation of flowering.

Through intimate interactions with host plants, parasitic plants can effectively acquire host genes via HGT, as an adaptive trait (Kim et al. 2014). We identified a total of 115 horizontally transferred genes from the host in the *C. deserticola* genome, of which 91.3% were located on 20 assembled chromosomes, while 8.7% were mapped to 5 unmapped scaffolds (Table S12). These genes were predominantly distributed on chromosomes Chr13, Chr1, Chr4 and Chr10, with most located in the distal regions of the chromosomes (Figure S9). The majority of proteins encoded by these genes were localised in the cytoplasm and were involved in processes such as energy metabolism, RNA modification, transport and catabolism (Figure 4). A smaller subset of these proteins was localised in the nucleus and organelles, where they participated in primary metabolism and the metabolism of genetic information. The statistical analysis of GO enrichment results revealed that horizontally transferred genes in *C. deserticola* display an array of functions, predominantly associated with the regulation of metabolism, energy production, protein degradation, cell division and ion regulation (Figure S10). This gene distribution and functional enrichment might represent a critical driving force by which *C. deserticola* transitions to holoparasitism, paving the way to a new biochemical basis for host interaction and adaptation.

2.6 | Discovery of Parasitism-Related Cells Through Single-Cell RNA-Seq

To perform scRNA-seq on *C. deserticola* stem cells, we collected *C. deserticola* stems and used a prior technique to obtain the cell suspension (Figure S11) (Farmer et al. 2021). A dissection was used to examine the various cell types present in the *C. deserticola* stem (Figure S12). A total of 406 219 346

reads were acquired using the 10× Genomics scRNA-seq platform. By mapping reads to the reference genome, the expression levels of genes in single cells were quantified, with 75% of the reads aligning, which validated the transcriptome sequence's accuracy and quality. We identified 32 074 cells, with each cell exhibiting a median expression of 1114 genes among different cell types (Table S13).

In a plant single-cell atlas, cell types are generally annotated with the help of one or multiple marker genes. This approach depends on homologous marker genes (Sun et al. 2023), but some of the markers were absent in *C. deserticola*. Cell types can be identified using comprehensive gene databases from model plants (Wu et al. 2024). Here, we downloaded 23 154 *Arabidopsis thaliana* marker genes from PlantscRNAdb. Then, we identified the orthologs between *A. thaliana* and *C. deserticola* by sequence alignment. The 12 distinct cell clusters were annotated (Figure 5A, Figure S13 and Table S14), with the cells from clusters 0, 1, and 2 identified as metaphloem sieve element cells, the cells from clusters 3, 4 and 8 as vasculature cells, cells from cluster 5 as phloem pole pericycle cells, and cells from cluster 7 as epidermis cells. Clusters 6 and 9 did not have enough markers to support cell annotation. In particular, most of the horizontally transferred genes were highly expressed in clusters 10 and 11, and these two clusters have no corresponding genes in *A. thaliana*. Therefore, we speculate that clusters 10 and 11 were related to the expression of parasitism-related genes and define them as parasitism-related cells (Figure 5B and Figure S14). In situ hybridisation was conducted to localise the expression of the marker genes of the metaphloem sieve element and the vasculature cell. The results demonstrated the spatial distribution of marker genes in the vascular tissue (CdesChr4G086830, CdesChr1G000260, CdesChr1G007780 and CdesChr9G200950) (Figure S15).

To assess the variations in expression profiles across different stages, pseudotime analysis was carried out on cells associated with parasitism (clusters 10 and 11). Cells associated with parasitism were mapped to the end of the pseudotime path, encompassing three trajectory states, and primarily clustered at one of the larger ends (Figure 6A–C). Monocle 2 was used to calculate the gene expression patterns of cells in various states in pseudotime order. Genes with differential expression were categorised into three states, indicating shifts in gene expression differences from the start to the end of pseudotime (Figure 6D). GO terms related to nucleic acid-binding function, ATP binding, and metal ion binding were enriched at the beginning of the pseudotime. Genes related to carbohydrate metabolic processes were enriched in the end stage of pseudotime (Figure 6E).

2.7 | Cell Type-Specific Distribution of Phenylethanoid Glycoside Biosynthesis

To decipher the spatial distribution of the phenylethanoid glycoside biosynthesis pathway in the *C. deserticola* stem, we analysed the expression patterns of the associated genes at single-cell resolution. In higher plants, phenylethanoid glycoside metabolism is divided into the phenylalanine metabolic pathway, the dopamine/tyrosine pathway and the downstream acylation and glycosylation catalytic pathways (Figure S16) (Delazar et al. 2011). Our single-cell transcriptomic analysis revealed distinct spatial expression patterns: genes associated with the phenylalanine metabolic pathway were predominantly expressed in vasculature cells (Table S15), suggesting their primary role in early biosynthetic steps. Dopamine/tyrosine pathway genes showed cell-type-specific enrichment: aldehyde dehydrogenases (ALDH), copper amine oxidases (CuAO), and UDP-glycosyltransferases (UGT) in the dopamine/tyrosine pathway were preferentially expressed in vasculature cells (Figure S17). The enzymes located in the central section of the dopamine/tyrosine pathway from tyrosine to tyramine, as well as L-DOPA to dopamine, were highly enriched in parasitism-related cells (clusters 11) (Table S15). And polyphenol oxidase (PPO) was also enriched in parasitism-related cells. These results indicate that vasculature cells and parasitism-related cells might be important for maintaining phenylethanoid glycoside diversity in *C. deserticola* (Table S16).

2.8 | A Multiomics Database for *C. deserticola*

The absence of a database has hindered the comprehensive utilisation of the valuable multi-omics data derived from *C. deserticola*. To overcome this limitation, we created a dedicated *C. deserticola* database that combines genomics and transcriptomics data in a multi-omics format. This comprehensive database comprises two principal modules: (1) the Genome module, which offers genome sequences, gene structures, and functional annotations, all accessible through the JBrowse browser for enhanced visualisation; and (2) the Single-cell transcriptome module, which provides a single-cell atlas, cell-type annotation and enrichment analysis (Figure 7). Furthermore, the database offers various analytical tools, including BLAST, Heatmap, MSA, Venn, Sequence Fetch and Collinearity Analysis, empowering users to conduct studies in conjunction with the *C. deserticola* dedicated database to achieve desired outcomes. This database provides comprehensive and user-friendly visualisation tools for comparative genomic sequences, gene structures and gene annotations of *C. deserticola*, thereby promoting the preservation of genetic data, molecular breeding and innovation in germplasm resources. Users can access the database interface by visiting the following URL: http://60.30.67.246:7006/Home.

3 | Discussion

3.1 | Evolution of Plant Parasitism Through Space and Time

We present the first chromosome-level genome for *C. deserticola* and contextualise this by examining the evolutionary history of parasitism in the Orobanchaceae more broadly. By integrating published transcriptome and genome data from other plants in the Orobanchaceae, we were able to explore aspects of temporal and geographic divergence in the family. Our findings suggest that the Orobanchaceae likely originated in the low-latitude regions of the Middle East and Asia near the Cretaceous-Palaeocene (K–Pg) boundary. During this time, the swift development of angiosperms not only provided the necessary ecological and climatic conditions for the emergence of Orobanchaceae but also promoted a diversification of potential hosts, thereby opening up new opportunities

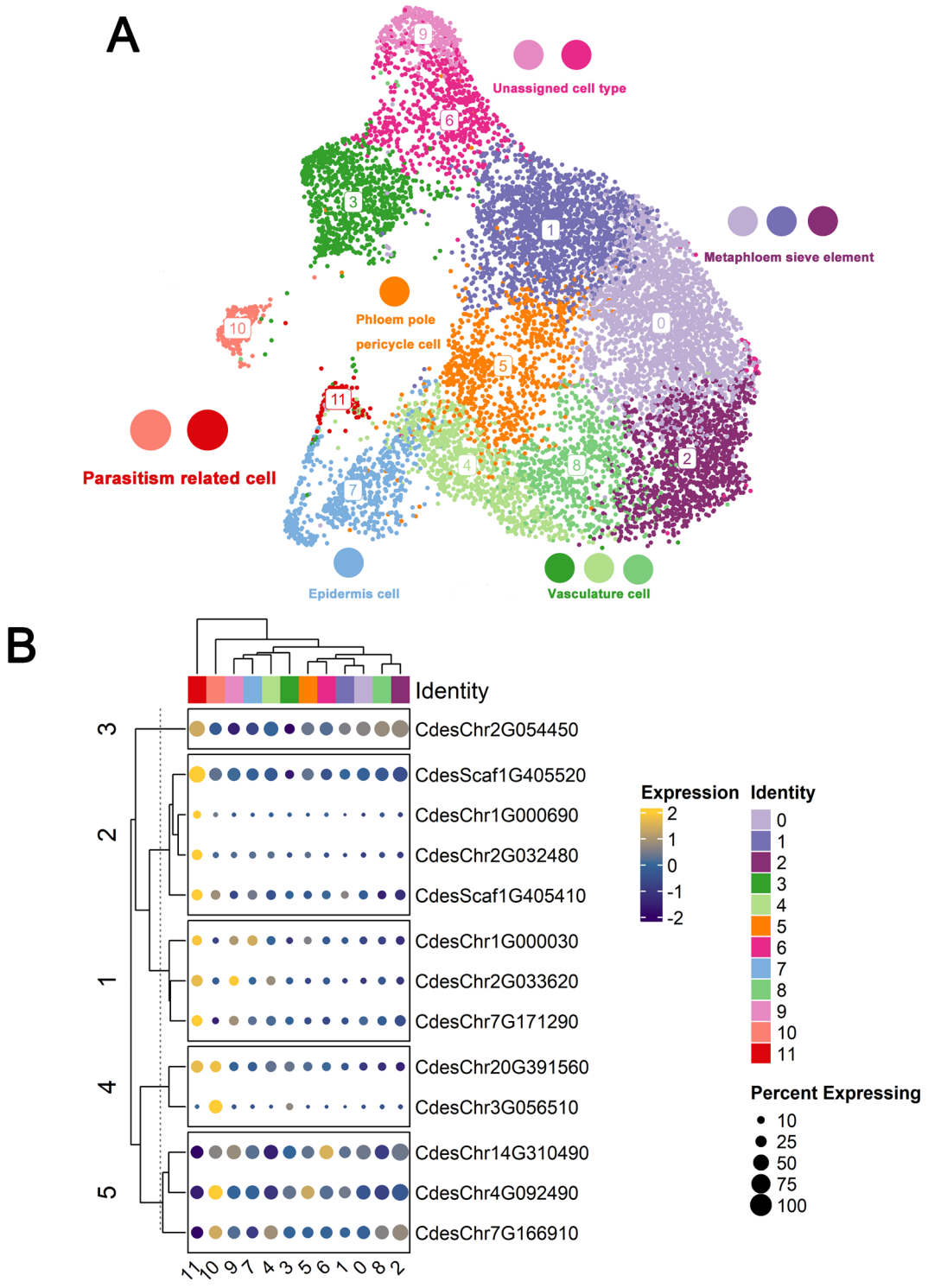

FIGURE 5 | A single-cell map of the *C. deserticola* stem and identification of parasitism-related cells. (A) A UMAP visualisation with reduced dimensions shows individual stem cells, each colour indicating a unique cluster. A total of 12 distinct cell clusters were identified. The three clusters of metaphloem sieve element cells (0, 1, and 2) are visualised in shades of purple; the three clusters of vasculature cells (3, 4, and 8) are visualised in shades of green; cells of cluster 5, visualised in orange, are assigned as phloem pole pericycle cells; cells of cluster 7, visualised in blue, are assigned as epidermis cells. Cells of clusters 10 and 11, visualised in red, are assigned as parasitism-related cells. Each dot symbolises a single cell. (B) The expression profiles of the typical cell horizontally transferred genes. Dot size indicates the proportion of cells expressing a particular gene in each cluster, and the colours represent the scaled average expression.

for parasitism (Frailey et al. 2018). The Cymbarieae, predominantly consisting of facultative hemiparasitic plants, represent the earliest diverging parasitic lineage, originating in East Asia. This suggests that East Asia has been a key centre for the emergence and early evolution of parasitic plants in the Orobanchaceae family. By the Eocene (56.0–33.9 Mya), the holoparasitic lineage of

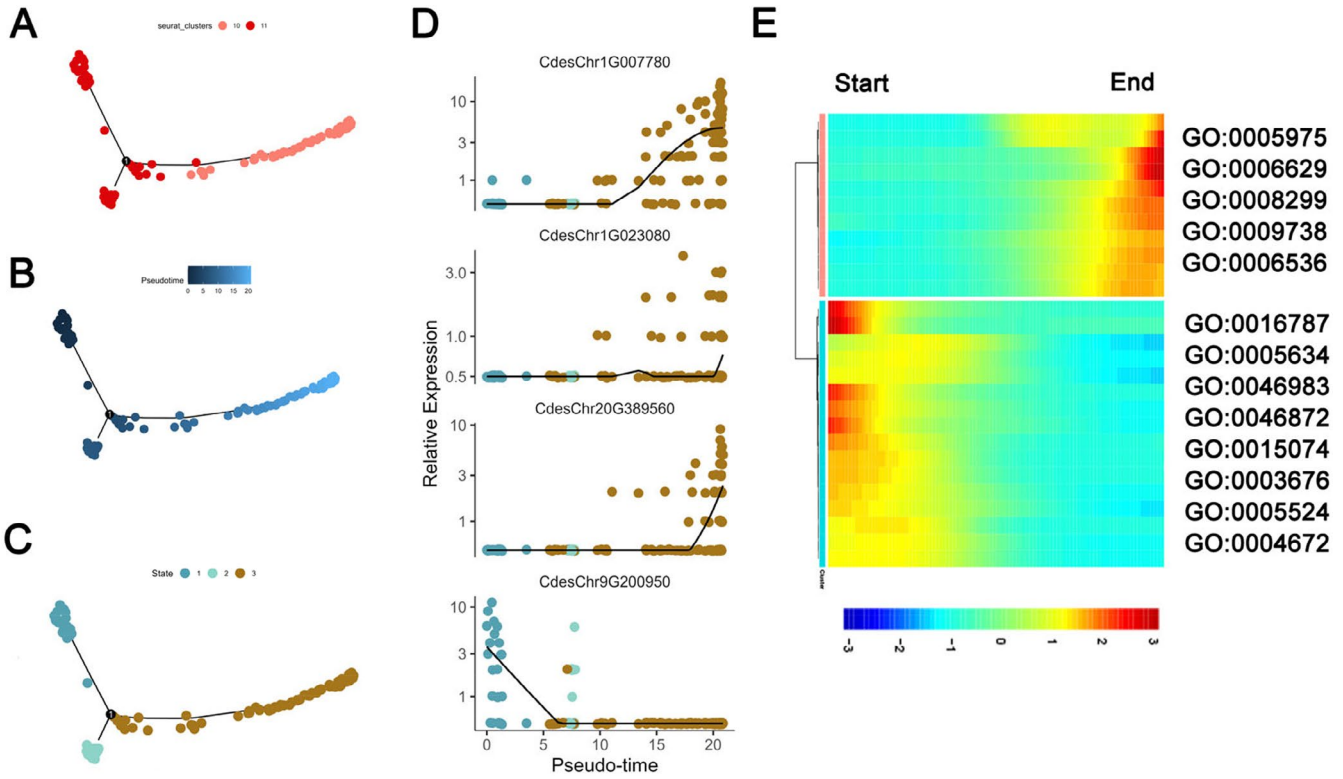

FIGURE 6 | Reconstruction of the developmental trajectories of parasitism-related cells. (A) A successive differentiation trajectory of parasitism-related cells. The colour of the dots indicates the cluster. (B) Distribution of cells in the parasitism related cell subclusters on the pseudotime trajectory. The colour of the dots indicates the pseudotime score. (C) Distribution of cells in the parasitism related cell subclusters on the pseudotime trajectory. The colour of the dots indicates the state. (D) Gene expression kinetics of representative genes along a pseudotime progression. (E) Heatmap showing the expression of pseudotime-dependent genes over the pseudotime trajectory of parasitism related cell subclusters. Representative GO terms of modules are labelled on the right.

the Orobanchaceae had migrated to high-latitude regions of the Eurasian continent and subsequently dispersed to North America. The Eocene was marked by global warming, facilitating the expansion of many angiosperms into mid- and high-latitude regions (Frailey et al. 2018). Then, as the climate cooled, high-latitude regions became inhospitable for parasitic plants (Lu et al. 2022), resulting in a decline in the diversity of Orobanchaceae in these areas. Deciphering this spatio-temporal evolutionary history is fundamental for understanding host–parasite interactions at the molecular and cellular level, which we consider below.

3.2 | Genomic Evolution in Parasitic Plants

Our results reveal significant gene family contraction in parasitic plants in the Orobanchaceae. During the initial phases of evolution, hemiparasitic plants shifted to heterotrophy through functional optimisation and specialisation of particular gene families. This evolutionary process was accompanied by the simplification of intercellular communication and signal transduction pathways, a reduction in the regulation of circadian rhythms and an enhancement of energy metabolism, gene regulation, and transmembrane transport functions. These genomic adjustments laid the foundation for a transition from hemiparasitism to holoparasitism. The metabolic and physiological functions of parasitic plants become increasingly simplified during this shift. Accordingly, resources became allocated to functions that directly support this derived life history, such as haustorium development and host resource utilisation; meanwhile, investment in non-essential functions, such as photosynthesis and secondary metabolism, was reduced. These dynamic changes in genomic function not only illustrate the adaptive capabilities of parasitic plants to their environments and host plants but also highlight how genome evolution facilitates transitions in plant life histories through functional specialisation and simplification. Moreover, our observations indicate that, in contrast to plastid genome reduction observed in parasitic plants (Al-Juhani et al. 2022), nuclear genomes of parasitic species within Orobanchaceae are generally larger than those of autotrophic plants. Notably, among these species, *C. deserticola* possesses the largest genome. Although the genomes of *O. cumana* and *P. aegyptiaca* have undergone WGD, their genome sizes are less than half that of *C. deserticola* (Figure S5). This finding suggests that following LTR insertion events and WGD, the *C. deserticola* genome has been more effectively preserved, reflecting a distinct evolutionary trajectory compared to *O. cumana* and *P. aegyptiaca* (Xu et al. 2024).

3.3 | Pathway to Holoparasitism

The transition to holoparasitism represents an obligate shift to heterotrophy and was accompanied by significant gene loss (Frailey et al. 2018), the extent and mechanisms of which are variable (Xu et al. 2016). Hemiparasitic plants, while having lost

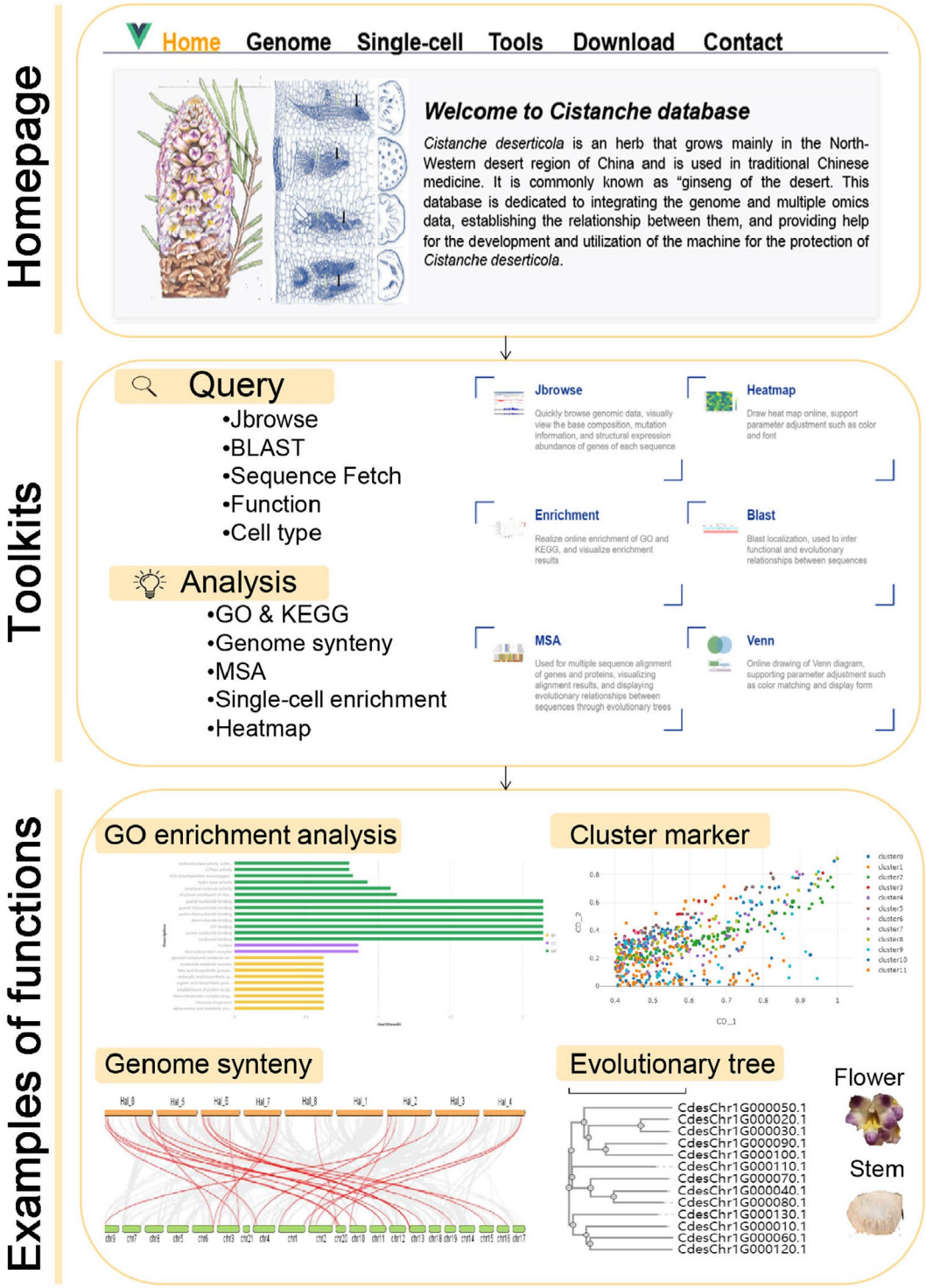

FIGURE 7 | Summary of the data modules and toolkits, along with examples of the functions of the *C. deserticola* dedicated database.

some functions, still retain a certain degree of photosynthetic ability (Casadesús and Munné-Bosch 2021), associated with a relatively low rate of gene loss. This gene loss is associated predominantly with non-essential physiological functions. During the evolutionary path to holoparasitism, losing genes associated with immunity could represent a co-evolutionary adaptation of holoparasitic plants to the defence strategies of their hosts (Xu et al. 2021). Perhaps by lacking these genes, holoparasitic plants sidestepped the immune responses of their hosts, thereby reducing the host resistance to invasion. This process is presumably essential for holoparasitic plants to successfully attach to and penetrate their hosts' tissues. The reallocation of resources resulting from the loss of immunity-related genes in parasitic plants facilitates a shift in energy focus towards parasitic functions, while the significance of photosynthesis and autonomous reproduction diminishes. Therefore, we propose that the gene loss model for holoparasitic plants is characterised by a loss of immunity-related genes in association with host invasion, followed by a gradual loss of genes associated with photosynthesis and flowering, thereafter.

In autotrophic plants, flowering genes are essential for reproduction, as they regulate the formation and development of floral organs necessary for seed production (Thomson and Wellmer 2019). However, the reproductive strategies of parasitic plants have undergone significant changes during their adaptation to novel host-based environments, leading to a gradual loss of flowering genes. Indeed, the loss of flowering genes has occurred independently across various evolutionary lineages of parasitic plants, suggesting a shared adaptive trend to parasitism (Xu et al. 2022). *C. deserticola* is entirely dependent on its host for nutrients, so its reproductive processes are presumably strongly influenced by environmental signals. Compared to other flowering pathways, sugar pathway genes in *C. deserticola* remain relatively conserved, with only *SUC9* being lost (Figure 4). Sugar functions as a signalling molecule in plants and is essential for controlling growth and developmental activities (Peng et al. 2018). *C. deserticola* may utilise the sugar pathway to 'sense' carbon sources supplied by the host as an alternative signal to regulate flowering time. Meanwhile, genes associated with autonomous pathways are less impacted, indicating that *C. deserticola* predominantly relies on internal signals to govern the regulation of flowering, rather than external environmental cues such as photoperiod or temperature. This unique regulatory mechanism reflects the extensive adaptation of *C. deserticola* to holoparasitism. It illustrates that, despite extensive gene loss, *C. deserticola* has not entirely forsaken flowering functionality; rather there has been a functional replacement, and integration, through alternative pathways.

HGT is ubiquitous in parasitic plants and has enabled them to acquire novel metabolic pathways and improve stress resistance (Yang et al. 2016), increasing their competitiveness and fitness (Saroha et al. 2022). In *C. deserticola*, HGT plays a critical role in energy metabolism and transport, enabling the efficient absorption and utilisation of host resources. Some horizontally transferred genes are localised within the nucleus and organelles, indicating their integration into the core metabolic network and direct involvement in the fundamental metabolic functions of *C. deserticola*. The functional integration of HGT highlights the specialised evolutionary trajectory of the *C. deserticola* genome in the shift to heterotrophy. The enhancement of genes associated with primary and energy metabolism allows *C. deserticola* to become entirely dependent on its host, thereby relinquishing certain functions essential for independent survival. Taken together, it is clear that extensive gene loss, coupled with the functional integration of horizontally transferred genes, was a driver in the evolution of *C. deserticola*—a specialist desert holoparasite.

3.4 | Host–Parasite Coordination at the Cellular Level

Our single-cell transcriptome atlas analysis of *C. deserticola* stems identified parasitism-related cells (clusters 10 and 11) with high expression of horizontally transferred genes. Compared to other cell clusters, the parasitism-related cells exhibited distinct transcriptomic features, indicating their highly specialised functions. Pseudotime analysis provided a temporal dimension of the cellular developmental trajectory, revealing the dynamic changes of parasitism-related cells from the initial to mature stage (Figure 6A–C). The results showed that cluster 11 was located at the beginning of development, marking the starting point of the parasitic process. During the early parasitic phase, parasitism-related cells establish an initial connection with the host by expressing nucleic acid-binding genes and metabolism-regulating genes, completing the development of the initial parasitic structure (Alexander et al. 2016). This process lays the foundation for the subsequent absorption of resources from the host. At the end of the pseudo-time trajectory, cluster 10 signified the complete maturation of cells associated with parasitism. During the mature phase, the primary function was nutrient acquisition—a critical process for a parasitic plant. High levels of carbohydrate metabolism gene expression suggest that mature parasitism-related cells (cluster 10) obtain energy by absorbing sugars from the host. With the development of the parasitic structure completed, the main function of mature parasitism-related cells shifts towards the efficient absorption and metabolism of host-derived nutrients. This transition reflects a stage-specific division of labour among parasitism-related cells in *C. deserticola*, revealing a highly coordinated developmental mechanism that has evolved in close association with its host.

Phenylethanoid glycosides are key compounds in *C. deserticola* that have an array of pharmacological effects (Leberecht et al. 2022). By analysing the spatial distribution of pathway gene transcripts, we found that the biosynthesis of phenylethanoid glycosides in *C. deserticola* exhibited multicellular compartmentation. The phenylalanine metabolism pathway was localised in vascular cells. The dopamine/tyrosine pathway was initiated in early parasitism-related cells (cluster 11), and intermediate steps predominantly occurred in vascular cells. This multicellular compartmentalisation of secondary metabolism can effectively reduce substrate inhibition of enzymes, regulate the size and direction of metabolic flux, lessen the harmful effects of intermediate products, and distribute the metabolic burden across cells, which is vital for their survival and adaptation to environmental changes (Sun et al. 2023). The involvement of early parasitism-related cells in the biosynthesis of phenylethanoid glycosides demonstrates their pivotal role in the metabolic network, reflecting the parasite's ability to efficiently convert host-derived resources into bioactive secondary metabolites while simultaneously supporting primary metabolism. Once again, this metabolic pathway highlights the extraordinarily tightly coupled relationship that has co-evolved between *C. deserticola* and its host, whereby HGTs enable the former to synthesise phenylethanoid glycosides, enabling adaptability to environmental challenges.

Understanding the mechanisms of interaction between parasitic plants, their hosts, and the environment offers crucial insights into plant–plant dynamics, with practical implications for food security, environmental management, and our broader understanding of plant ecology and evolution in a changing world. Genome and single-cell transcriptome sequencing of *C. deserticola* shows that the molecular and cellular mechanisms governing gene simplification, functional integration and cell specialisation underlie an intricate, multi-faceted evolutionary pathway to parasitism. Our analysis of the spatial distribution of the *C. deserticola* phenylethanoid glycoside biosynthesis pathway reveals a role in early parasitism-related cells through the overexpression of genes in the dopamine/tyrosine metabolism

pathway (particularly PPO). This finding provides the first single-cell-level evidence of the crucial role played by early parasitism-related cells in the core steps of phenylethanoid glycoside synthesis. We take a holistic approach to exploring host–parasite evolution. Through the development of a dedicated *C. deserticola*, we create a robust foundation of data and tools to support future research endeavours. We hope that this will pave the way for future insights into *C. deserticola*, enabling improvements in quality and yield. Furthermore, it will provide practical measures for the conservation of this endangered species in its natural habitat.

4 | Experimental Procedures

4.1 | Sampling and Estimation of Genome Size

Field experiments were performed in Alxa, Inner Mongolia, with geographical coordinates of 38°47′3″ N and 105°21′60″ E, situated in northwest China. The young stems of *C. deserticola* were harvested (Figure 1A), and chopped up the sample with a sterile blade for 2 min (repeat 1–3 times), until no large organisation remained. To ascertain the genome size of *C. deserticola*, the k-mer counting method was utilised. This process entailed examining the frequency of k-mers among the cleaned Illumina short reads, facilitated by the Jellyfish and GenomeScope software (Marçais and Kingsford 2011). To estimate genome size, use the formula: Genome size equals K-mer Number over K-mer Depth.

4.2 | Genome Sequencing and Assembly

The complete genome sequencing, assembly, and annotation pipeline was summarised in Figure S18. To construct chromosome-level assemblies for *C. deserticola*, data from three technologies were integrated: SMRT sequencing with the PacBio Sequel, Hi-C for chromosome conformation capture and polished short reads (Ashburner et al. 2000). DNA samples underwent processing for both next-generation sequencing (NGS) and SMRT sequencing. For NGS, libraries were created using the Nextera Flex DNA library kit (Illumina), fragmented using a Covaris ultrasonicator and sequenced on either the Illumina NovaSeq 6000 or BGI DNBSEQ-T7 platforms. Quality control was conducted utilising Qubit 2.0 and Agilent 2100, followed by quantitative PCR (qPCR) for precise quantification. SMRT sequencing was conducted using a PCR-free SMRTbell library preparation and the PacBio Sequel II platform for sequencing. Data filtering was performed using FastQC for NGS and SMRTLink 8.0 for SMRT, applying standards like removing low-quality reads, adapter sequences for NGS, and length and quality thresholds for SMRT. The effective data from both second and third-generation sequencing were statistically summarised. The genome assembly was carried out using hifiasm (v 0.16.1), which utilises an Overlap Layout Consensus (OLC) algorithm.

The Hi-C library was constructed using the Dovetail Hi-C Library kit, with young *C. deserticola* stems processed as per a previously established protocol (Belton et al. 2012). Libraries for paired-end sequencing were constructed from chimeric fragments and sequenced via Illumina technology. Using the 3D-DNA pipeline, mis-joined contigs were corrected after aligning Hi-C reads to the contig-level assembly (v 201008), which is specifically designed to detect abrupt changes in long-range contact patterns indicative of incorrect contig joins (Dudchenko et al. 2017). Interaction intensity and positional relationships of contigs were visualised using HiCExplorer (v 3.6) (Wolff et al. 2020). Using the contig N50 size metric, the assembly was assessed, and BWA (v 0.7.17) was employed to align next-generation clean reads to the reference genome for evaluating mapping rates and coverage. BUSCO (v 4) (Benchmarking Universal Single-Copy Orthologs) was used to assess the completeness of the assembly (v 5.3.0; parameter: *e*-value 1e-05), using a eukaryotic reference data set with 255 conserved genes (eukaryota_odb10.2019-11-20) (Simão et al. 2015). Genome continuity was evaluated using LAI in the LTR_retriver package (Ou et al. 2018). Mequery (v 1.3) was used to assess the consensus quality (QV) value and completeness of the genome assembly (Rhie et al. 2020). RNA-seq data were mapped to our genome assembly using HISAT2 (v 2.2.0) (HiSat2, RRID:SCR_015530) (Kim et al. 2015).

4.3 | Nuclear Suspension Isolation and Single-Cell RNA-Seq

The chopped sample was gently moved to a centrifuge tube with NIB solution, centrifuged at a speed of 300 g for 1 min; then, the supernatant was transferred into a new centrifuge tube. The supernatant was filtered and centrifuged; then, the wash buffer was used for resuspension. We adjusted the nuclear suspension concentration to 1000–2000 nuclei/μL with the wash buffer and stored it on ice for final FANS sorting. The construction of the library utilised the Chromium Single Cell 3′ Library and Single Cell 3′ v3 Gel Beads kit. The automated Chromium Controller instrument was used to load a nuclear suspension for generating single-cell GEMs. Each GEM underwent reverse transcription to ensure its integrity, generating cDNA with a unique barcode for each cell. After synthesising cDNA, GEMs were disrupted to free the cDNA from each cell. The cDNA that had been released was pooled and underwent PCR amplification. Once amplified, the cDNA was divided into fragments, and sequencing adapters were added to assemble libraries. The Illumina HiSeq X Ten platform was used to generate paired-end short reads of 150 base pairs.

See Appendix S1 (Supplemental Methods) for genome annotation, phylogenetic analysis, ancestral area reconstruction, comparative genomics analysis, scRNA-seq data analysis, HGT identification and database implementation.

Author Contributions

All authors contributed to the manuscript revision, read and approved the submitted version. X.Z. and Y.M. contributed equally to this work. X.Z. and Y.M. contributed to the conception and design of the study; X.Z. and Y.X. performed the analysis; Y.M. and M.Z. wrote the first draft of the manuscript; L.H., S.C., G.Z. and Chris wrote sections of the manuscript.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (82211540726, 82073960 and 82274045); the Open Fund of the State Key Laboratory of Southwestern Chinese Medicine Resources (SCMR2022015); the Beijing Municipal Natural Science Foundation (24G40321); and the China Postdoctoral Science Foundation (2024M763263).

Conflicts of Interest

The authors declared no conflicts of interest.

Data Availability Statement

All of the datasets and a useful toolkit can be accessed from the *C. deserticola* database website (http://60.30.67.246:7006/Home). Sequencing data have been submitted to the NCBI Sequence Read Archive (NCBI BioProject PRNA1344256). All data are also available from the corresponding author on request.

References

Alexander, W. G., J. H. Wisecaver, A. Rokas, and C. T. Hittinger. 2016. "Horizontally Acquired Genes in Early-Diverging Pathogenic Fungi Enable the Use of Host Nucleosides and Nucleotides." *Proceedings of the National Academy of Sciences* 113: 4116–4121.

Al-Juhani, W., N. T. Al Thagafi, and R. N. Al-Qthanin. 2022. "Gene Losses and Plastome Degradation in the Hemiparasitic Species Plicosepalus Acaciae and Plicosepalus Curviflorus: Comparative Analyses and Phylogenetic Relationships Among Santalales Members." *Plants* 11: 1869.

Ashburner, M., C. A. Ball, J. A. Blake, et al. 2000. "Gene Ontology: Tool for the Unification of Biology." *Nature Genetics* 25: 25–29.

Belton, J.-M., R. P. McCord, J. H. Gibcus, N. Naumova, Y. Zhan, and J. Dekker. 2012. "Hi–C: A Comprehensive Technique to Capture the Conformation of Genomes." *Methods* 58: 268–276.

Bouché, F., G. Lobet, P. Tocquin, and C. Périlleux. 2016. "FLOR-ID: An Interactive Database of Flowering-Time Gene Networks in *Arabidopsis thaliana*." *Nucleic Acids Research* 44: D1167–D1171.

Casadesús, A., and S. Munné-Bosch. 2021. "Holoparasitic Plant–Host Interactions and Their Impact on Mediterranean Ecosystems." *Plant Physiology* 185: 1325–1338.

Chen, J., R. Yu, J. Dai, Y. Liu, and R. Zhou. 2020. "The Loss of Photosynthesis Pathway and Genomic Locations of the Lost Plastid Genes in a Holoparasitic Plant *Aeginetia indica*." *BMC Plant Biology* 20: 1–10.

Chen, X., D. Fang, Y. Xu, et al. 2023. "Balanophora Genomes Display Massively Convergent Evolution With Other Extreme Holoparasites and Provide Novel Insights Into Parasite–Host Interactions." *Nature Plants* 9: 1627–1642.

Delazar, A., M. R. Delnavazi, L. Nahar, et al. 2011. "Lavandulifolioside B: A New Phenylethanoid Glycoside From the Aerial Parts of Stachys Lavandulifolia Vahl." *Natural Product Research* 25: 8–16.

Dudchenko, O., S. S. Batra, A. D. Omer, et al. 2017. "De Novo Assembly of the *Aedes aegypti* Genome Using Hi-C Yields Chromosome-Length Scaffolds." *Science* 356: 92–95.

Fan, Y., Q. Zhao, H. Duan, et al. 2023. "Large-Scale mRNA Transfer Between Haloxylon Ammodendron (Chenopodiaceae) and Herbaceous Root Holoparasite *Cistanche deserticola* (Orobanchaceae)." *IScience* 26: 105880.

Farmer, A., S. Thibivilliers, K. H. Ryu, J. Schiefelbein, and M. Libault. 2021. "Single-Nucleus RNA and ATAC Sequencing Reveals the Impact of Chromatin Accessibility on Gene Expression in Arabidopsis Roots at the Single-Cell Level." *Molecular Plant* 14: 372–383.

Frailey, D. C., S. R. Chaluvadi, J. N. Vaughn, C. G. Coatney, and J. L. Bennetzen. 2018. "Gene Loss and Genome Rearrangement in the Plastids of Five Hemiparasites in the Family Orobanchaceae." *BMC Plant Biology* 18: 1–12.

Fu, W., X. Liu, N. Zhang, et al. 2017. "Testing the Hypothesis of Multiple Origins of Holoparasitism in Orobanchaceae: Phylogenetic Evidence From the Last Two Unplaced Holoparasitic Genera, Gleadovia and Phacellanthus." *Frontiers in Plant Science* 8: 1380.

Gruzdev, E. V., V. V. Kadnikov, A. V. Beletsky, et al. 2019. "Plastid Genomes of Carnivorous Plants Drosera Rotundifolia and Nepenthes × Ventrata Reveal Evolutionary Patterns Resembling Those Observed in Parasitic Plants." *International Journal of Molecular Sciences* 20: 4107.

Jean-Baptiste, K., J. L. McFaline-Figueroa, C. M. Alexandre, et al. 2019. "Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*." *Plant Cell* 31: 993–1011.

Kim, D., B. Langmead, and S. L. Salzberg. 2015. "HISAT: A Fast Spliced Aligner With Low Memory Requirements." *Nature Methods* 12: 357–360.

Kim, G., M. L. LeBlanc, E. K. Wafula, C. W. DePamphilis, and J. H. Westwood. 2014. "Genomic-Scale Exchange of mRNA Between a Parasitic Plant and Its Hosts." *Science* 345: 808–811.

Leberecht, C., M. Schroeder, and D. Labudde. 2022. "A Multiscale Model of the Regulation of Aquaporin 2 Recycling." *npj Systems Biology and Applications* 8: 16.

Li, C., J. Lei, Y. Zhao, X. Xu, and S. Li. 2015. "Effect of Saline Water Irrigation on Soil Development and Plant Growth in the Taklimakan Desert Highway Shelterbelt." *Soil and Tillage Research* 146: 99–107.

Li, X., L. Chen, L. Yao, J. Zou, J. Hao, and W. Wu. 2022. "Calcium-Dependent Protein Kinase CPK32 Mediates Calcium Signaling in Regulating Arabidopsis Flowering Time." *National Science Review* 9: nwab180.

Lu, X., R. Jiang, and G. Zhang. 2022. "Predicting the Potential Distribution of Four Endangered Holoparasites and Their Primary Hosts in China Under Climate Change." *Frontiers in Plant Science* 13: 942448.

Lyko, P., and S. Wicke. 2021. "Genomic Reconfiguration in Parasitic Plants Involves Considerable Gene Losses Alongside Global Genome Size Inflation and Gene Births." *Plant Physiology* 186: 1412–1423.

Marçais, G., and C. Kingsford. 2011. "A Fast, Lock-Free Approach for Efficient Parallel Counting of Occurrences of k-Mers." *Bioinformatics* 27: 764–770.

Masumoto, N., Y. Suzuki, S. Cui, et al. 2021. "Three-Dimensional Reconstructions of Haustoria in Two Parasitic Plant Species in the Orobanchaceae." *Plant Physiology* 185: 1429–1442.

Miao, Y., H. Chen, W. Xu, Q. Yang, C. Liu, and L. Huang. 2022. "Structural Mutations of Small Single Copy (SSC) Region in the Plastid Genomes of Five Cistanche Species and Inter-Species Identification." *BMC Plant Biology* 22: 412.

Nickrent, D. L. 2020. "Parasitic Angiosperms: How Often and How Many?" *Taxon* 69: 5–27.

Ou, S., J. Chen, and N. Jiang. 2018. "Assessing Genome Assembly Quality Using the LTR Assembly Index (LAI)." *Nucleic Acids Research* 46: e126.

Peng, Y., L. Chen, S. Li, et al. 2018. "BRI1 and BAK1 Interact With G Proteins and Regulate Sugar-Responsive Growth and Development in Arabidopsis." *Nature Communications* 9: 1522.

Rhie, A., B. P. Walenz, S. Koren, and A. M. Phillippy. 2020. "Merqury: Reference-Free Quality, Completeness, and Phasing Assessment for Genome Assemblies." *Genome Biology* 21: 245.

Saroha, T., V. Chaudhry, and P. B. Patil. 2022. "Novel Insights Into the Role of the Mobilome in Ecological Diversification and Success of

Staphylococcus haemolyticus as an Opportunistic Pathogen." *Microbial Genomics* 8: e000755.

Shen, L., R. Xu, S. Liu, et al. 2019. "Parasitic Relationship of Cistanche Deserticola and Host-Plant Haloxylon Ammodendron Based on Genetic Variation of Host." *Chinese Herbal Medicines* 11: 267–274.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. "BUSCO: Assessing Genome Assembly and Annotation Completeness With Single-Copy Orthologs." *Bioinformatics* 31: 3210–3212.

Song, Y. H., J. S. Shim, H. A. Kinmonth-Schultz, and T. Imaizumi. 2015. "Photoperiodic Flowering: Time Measurement Mechanisms in Leaves." *Annual Review of Plant Biology* 66: 441–464.

Sun, G., Y. Xu, H. Liu, et al. 2018. "Large-Scale Gene Losses Underlie the Genome Evolution of Parasitic Plant *Cuscuta australis*." *Nature Communications* 9: 2683.

Sun, S., X. Shen, Y. Li, et al. 2023. "Single-Cell RNA Sequencing Provides a High-Resolution Roadmap for Understanding the Multicellular Compartmentation of Specialized Metabolism." *Nature Plants* 9: 179–190.

Thomson, B., and F. Wellmer. 2019. "Molecular Regulation of Flower Development." *Current Topics in Developmental Biology* 131: 185–210.

Westwood, J. H., J. I. Yoder, M. P. Timko, and C. W. Depamphilis. 2010. "The Evolution of Parasitism in Plants." *Trends in Plant Science* 15: 227–235.

Wolff, J., L. Rabbani, R. Gilsbach, et al. 2020. "Galaxy HiCExplorer 3: A Web Server for Reproducible Hi-C, Capture Hi-C and Single-Cell Hi-C Data Analysis, Quality Control and Visualization." *Nucleic Acids Research* 48: W177–W184.

Wu, S., A. L. M. Morotti, J. Yang, E. Wang, and E. C. Tatsis. 2024. "Single-Cell RNA Sequencing Facilitates the Elucidation of the Complete Biosynthesis of the Antidepressant Hyperforin in St. John's Wort." *Molecular Plant* 17: 1439–1457.

Xu, F., J. Jerlström-Hultqvist, M. Kolisko, et al. 2016. "On the Reversibility of Parasitism: Adaptation to a Free-Living Lifestyle via Gene Acquisitions in the Diplomonad Trepomonas sp. PC1." *BMC Biology* 14: 1–15.

Xu, S., R. Chen, X. Zhang, et al. 2024. "The Evolutionary Tale of Lilies: Giant Genomes Derived From Transposon Insertions and Polyploidization." *Innovation* 5: 100726.

Xu, Y., Y. Lei, Z. Su, et al. 2021. "A Chromosome-Scale Gastrodia Elata Genome and Large-Scale Comparative Genomic Analysis Indicate Convergent Evolution by Gene Loss in Mycoheterotrophic and Parasitic Plants." *Plant Journal* 108: 1609–1623.

Xu, Y., J. Zhang, C. Ma, et al. 2022. "Comparative Genomics of Orobanchaceous Species With Different Parasitic Lifestyles Reveals the Origin and Stepwise Evolution of Plant Parasitism." *Molecular Plant* 15: 1384–1399.

Yang, Z., Y. Zhang, E. K. Wafula, et al. 2016. "Horizontal Gene Transfer Is More Frequent With Increased Heterotrophy and Contributes to Parasite Adaptation." *Proceedings of the National Academy of Sciences of the United States of America* 113: E7010–E7019.

Yang, Z., Y. Zhou, J. Huang, et al. 2015. "Ancient Horizontal Transfer of Transaldolase-Like Protein Gene and Its Role in Plant Vascular Development." *New Phytologist* 206: 807–816.

Yoshida, S., and Y. J. Kee. 2021. "Large-Scale Sequencing Paves the Way for Genomic and Genetic Analyses in Parasitic Plants." *Current Opinion in Biotechnology* 70: 248–254.

Yoshida, S., S. Kim, E. K. Wafula, et al. 2019. "Genome Sequence of *Striga asiatica* Provides Insight Into the Evolution of Plant Parasitism." *Current Biology* 29: 3041–3052.

Zhang, D., J. Qi, J. Yue, et al. 2014. "Root Parasitic Plant *Orobanche aegyptiaca* and Shoot Parasitic Plant *Cuscuta australis* Obtained Brassicaceae-Specific Strictosidine Synthase-Like Genes by Horizontal Gene Transfer." *BMC Plant Biology* 14: 1–14.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Tables S1–S17:** pbi70464-sup-0001-TableS1-S17.docx. **Appendix S1:** pbi70464-sup-0002-AppendixS1.docx. **Figures S1–S18:** pbi70464-sup-0003-FigureS1-S18.docx.