2 Genomic Resources and Technical Toolkit

2.1 Reference genomes, assemblies, and databases

The initial monarch genome assembly (~273 Mb, ~16 866 predicted protein-coding genes; Table 1) provided a foundational resource for trait mapping and comparative genomics in Lepidoptera (Zhan et al., 2011). Subsequent chromosome-level assemblies, improved annotations, and curated community databases have substantially expanded the utility of monarch genomics for population, functional, and evolutionary analyses (MonarchBase team, 2012; Mongue et al., 2017). These resources have enabled identification of candidate loci underlying migration, circadian regulation, detoxification, pigmentation, and host-plant interactions, as well as comparative analyses of sex chromosome evolution within *Danaus*.

Recent high-quality assemblies have further resolved structural variants, repetitive regions, and the neo-Z chromosome, providing insight into genomic features that may contribute to adaptation and phenotypic divergence among migratory and non-migratory populations (Table 1). Collectively, these monarch-specific genomic resources now support both hypothesis-driven functional studies and broad-scale evolutionary inference.

Table 1 Genomic and functional resources for *D. plexippus* research

Resource type	Description	Year(s)	Applications	Key references
Genome assemblies (draft → chromosome-scale)	Initial draft genome (~273 Mb) followed by improved, chromosome-scale assemblies using long-read sequencing and Hi-C scaffolding; includes annotation of coding genes, repeats, structural variants, and neo-Z chromosome	2011-2020	Trait mapping, comparative genomics, population genomics, sex-chromosome evolution, structural variant discovery	Zhan et al., 2011; Mongue et al., 2017; Zhan et al., 2020
MonarchBase and genomic databases	Community-curated genome browser and annotation resource integrating gene models, transcriptomes, and functional annotations; linked to NCBI and other repositories	2012-present	Gene discovery, annotation refinement, comparative analyses, education and outreach	MonarchBase Team, 2012; Zhan et al., 2011
RNA-seq atlases (tissues, developmental stages)	Bulk RNA-seq from antennae, brain, fat body, flight muscle, wing discs, larvae, pupae, and adults across migratory and reproductive states	2009-present	Gene expression profiling, circadian biology, diapause regulation, migration physiology, developmental genetics	Merlin et al., 2009; Zhan et al., 2011; de Roode et al., 2011
CRISPR/Cas9 and TALEN applications	Targeted genome editing to disrupt or modify candidate genes; functional validation of regulatory and coding loci in monarchs and related Lepidoptera	2016-present	Causal tests of gene function (navigation, pigmentation, circadian clocks), regulatory element validation	Markert et al., 2016; Zhang and Reed, 2016
Metabolomic datasets (milkweeds and monarchs)	LC-MS/MS and untargeted metabolomics of milkweed secondary metabolites and monarch tissues; quantification of cardenolide diversity, sequestration, and biotransformation	2013-present	Chemical ecology, host-plant adaptation, parasite resistance, eco-genomic integration	Petschenka et al., 2013; Dreisbach et al., 2023; Agrawal et al., 2025
OE parasite genomic resources	Genomic and transcriptomic resources for *Ophryocystis elektroscirrha*, a specialist protozoan parasite of monarchs	2015-present	Host-parasite coevolution, disease ecology, immunity and chemical defense interactions	de Roode et al., 2008; Satterfield et al., 2015

2.2 Functional genomics, genome editing, and multi-omic tools

Functional genomic studies in monarchs have leveraged transcriptomic, proteomic, and metabolomic data to link genetic variation with key migratory, physiological, and defensive traits (Table 1). For example, tissue-specific transcriptomic analyses of antennae revealed circadian clock gene expression patterns essential for time-compensated sun-compass navigation (Merlin et al., 2009), while expression profiling of fat body and flight muscle tissues clarified metabolic shifts associated with long-distance migration and lipid utilization (de Roode et al., 2011).