

Article

De Novo Assembly of Eight Commercial Crossbred Pig Genomes Provides Insights into the Potential Functional Impact of Structural Variation Hotspots

Jiaolong Wen ^{1,2,†}, Haiqi Qiu ^{1,†}, Shaoxiong Deng ¹, Shiyuan Wang ¹, Yiyi Liu ¹, Meng Lin ^{1,3}, Jie Yang ^{1,3} , Zhenfang Wu ^{1,2,3}, Langqing Liu ^{1,3,4,*} and Yibin Qiu ^{1,2,4,*} 

¹ National Engineering Research Center for Breeding Swine Industry, South China Agricultural University, Guangzhou 510642, China; dragon123@stu.scau.edu.cn (J.W.); haiqi_qiu@stu.scau.edu.cn (H.Q.); dengsx1998@stu.scau.edu.cn (S.D.); wsy20221024010@stu.scau.edu.cn (S.W.); yiyiliu@stu.scau.edu.cn (Y.L.); linmeng@scau.edu.cn (M.L.); jieyang@scau.edu.cn (J.Y.); wzf@scau.edu.cn (Z.W.)

² Yunfu Subcenter of Guangdong Laboratory for Lingnan Modern Agriculture, Yunfu 527400, China

³ National and Regional Livestock Genebank, Guangdong Gene Bank of Livestock and Poultry, South China Agricultural University, Guangzhou 510642, China

⁴ Guangdong Provincial Key Laboratory of Agro-Animal Genomics and Molecular Breeding, South China Agricultural University, Guangzhou 510642, China

* Correspondence: langqing.liu@scau.edu.cn (L.L.); 13422157044qyb@gmail.com (Y.Q.)

† These authors contributed equally to this work.

Abstract

The Duroc × (Landrace × Yorkshire) (DLY) pig is a cornerstone of three-way crossbreeding system. Nevertheless, advances in commercial crossbred performance have been constrained by the dearth of high-resolution genomic resources for this key population. Here, we report the sequencing and assembly of 16 haplotype-resolved, chromosome-level genome assemblies derived from eight DLY pigs. These assemblies exhibited high continuity (contig N50: 18.17–29.54 Mb) and completeness (BUSCO: 99.3–99.4%), with sequences successfully localized to the 19 chromosomes. Genome annotation revealed an average of 21,922 protein-coding genes and 44.66% repetitive sequences per assembly. Comparative genomic analysis against the current reference genome Sscrofa11.1 enabled the construction of a non-redundant SV catalog comprising 130,416 variants, nearly half of which (48.99%) were novel relative to existing pig pan-genome SV panel. These SVs clustered non-randomly into 231 “SV hotspots” that were significantly enriched in protein-coding genes and putative regulatory elements. Functional analyses further linked these SV hotspots to quantitative trait loci (QTLs) associated with economically important traits. A focused analysis of a 3.43 Mb hotspot on chromosome 1, overlapping a known QTL for average daily gain, revealed eight high-frequency SVs in open chromatin regions near candidate genes (*NCS1*, *HMCN2*, *FUBP3*, *ABL1*, and *FIBCD1*), suggesting a cis-regulatory mechanism that may influence gene expression. Collectively, this work provides the first haplotype-resolved genomic resource for commercial crossbred pigs, and establishes a foundational framework for deciphering the genomic architecture of hybrid vigor and advancing precision breeding in swine.

Keywords: haplotype-resolved; genome assembly; Duroc × (Landrace × Yorkshire) pig; comparative analysis



Academic Editor: Laszlo Patthy

Received: 7 January 2026

Revised: 27 January 2026

Accepted: 27 January 2026

Published: 31 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The pig is a critically important livestock species for meat production worldwide [1]. Modern commercial production primarily relies on a three-breed terminal crossing systems to optimize productivity, wherein F1 sows (Landrace $\times$ Large White) are bred with purebred Duroc boars selected for superior production traits such as growth rate, leanness, and feed efficiency [2,3]. The resulting hybrid offspring Duroc $\times$ (Landrace $\times$ Yorkshire), commonly termed DLY pigs, constitute a substantial portion of the meat supply, meeting growing consumer demand for high-quality protein [4].

Beyond their production value, DLY pigs offer a distinct advantage for the genetic dissection of economically important traits. Compared with purebred populations, their hybrid genomes exhibit shorter linkage disequilibrium (LD) due to the recombination of parental haplotypes [5]. This accelerated LD decay enables more precise mapping of quantitative trait loci (QTL), allowing finer resolution of genomic regions associated with key performance traits [6]. Consequently, DLY populations have been widely adopted in pig genetic studies [6–9].

Over the past decade, genomic and genetic research in pigs has largely relied on a single Duroc-origin reference genome [1,10]. While invaluable, a single reference poses limitations for comprehensively understanding genomic architecture, haplotype diversity, and the genetic basis of heterosis in hybrid breeding systems [11]. The absence of high-quality, haplotype-resolved reference genomes for widely used commercial hybrids like DLY constrains the full potential of genomic selection [12]. Thus, developing such resources is essential to advance the accuracy and efficiency of modern pig breeding programs.

Structural variants (SVs) are now recognized as key determinants of phenotypic diversity [13–15]. Recent research has highlighted the importance of population-level SV catalogs as critical resources for understanding genomic diversity and its functional implications [16–18]. However, the contribution of large-size genomic variants (≥ 50 bp), e.g., genome assembly dependent SVs, in shaping the genomic architecture of pigs remains poorly characterized. A major limiting factor is the scarcity of precise genetic variation information in hybrid lines. Of the 45 pig genomes currently available in NCBI database <https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=9823> (accessed on 9 December 2025), only two (USMARCv1.0 and NCMD) represents crossbred animals [10,19].

To bridge these gaps, we sampled eight DLY pigs and successfully assembled high-quality, haplotype-resolved genomes by integrating Oxford Nanopore long-read and short-read sequencing data. We further identified genomic variants, with a focus on SVs, to construct a comprehensive variant catalog for this key commercial population and evaluated their potential functional contributions. This comprehensive dataset provides a valuable genetic resource that will enhance our understanding of the biological mechanisms underlying economically important traits and disease resilience in pigs.

2. Materials and Methods

2.1. Sample Collection

Ear tissue samples were collected from eight Duroc $\times$ (Landrace $\times$ Yorkshire) three-way crossbred (DLY) pigs (three males and five females) at 180–200 days of age. To avoid full- and half-sibling relationships, all individuals were selected based on a three-generation pedigree. The pigs were provided by the Wens Foodstuff Group Co., Ltd. (Yunfu, China). Samples were immediately snap-frozen in liquid nitrogen and stored at -80 °C until DNA extraction. All experimental protocols were approved by the Animal Care and Use Committee of the South China Agricultural University (approval number: SYXK 2019-0136, Guangzhou, China). No anesthesia or euthanasia was performed on the animals throughout this study.

2.2. Data Generation

Sequencing services were provided by Novogene Biotech Co., Ltd. (Beijing, China). Briefly, high-quality genomic DNA was isolated from the ear tissues (see Supplementary Method). For short-read sequencing, libraries were sequenced on a DNBSEQ-T7 platform, generating 150 bp paired-end reads. In total, 628.81 Gb of short-read data were produced, achieving a coverage ranging from $26.44\times$ to $57.42\times$ per individual (Table S1). To obtain long-read data, DNA libraries were prepared and sequenced on a Nanopore PromethION platform following standard Oxford Nanopore Technologies (ONT) protocols. Base-calling was performed with dorado (v1.3; <https://github.com/nanoporetech/dorado> (accessed on 9 December 2025)) using default parameters, retaining reads with an average quality score above 7. This yielded 565.44 Gb of ONT data with a mean read N50 of 19.28 kb and an average quality score of 12.89, providing $20.61\times$ to $38.09\times$ coverage across the eight samples (Table S1).

2.3. De Novo Genome Assembly

For each individual, sequence data were pre-processed before assembly. Short reads were trimmed and quality-controlled using Fastp [20] (v0.23.4). Corresponding ONT reads were error-corrected with Ratatosk [21] (v0.9.0) using the trimmed short reads as reference (Table S1). De novo genome assembly was performed per sample using the corrected ONT reads with Flye [22] (v2.9.6). The resulting draft assemblies were then polished into diploid contig-level sequences using Hypo-hybrid [23] (v1.0.3), which integrates both short and long-read data. Finally, chromosome-level scaffolding was carried out for each diploid assembly using the reference-guided tool RagTag [24] (v2.1.0), yielding two haplotype-resolved genomes per individual. Heterozygous regions between the two haplotypes were visualized with Bandage [25] (v0.9.0) by examining bubbles across the assembly graph of chromosomes (<https://github.com/T2T-CN1/CN1/tree/main/heterozygosity> (accessed on 9 December 2025)).

Assembly base quality (QV) was estimated using merqury [26] (v1.3), with k-mer databases constructed from short reads via meryl [26] (v1.4.1). Scaffold continuity was assessed with QUAST (v5.3.0; <https://github.com/ablab/quast> (accessed on 9 December 2025)). Assembly completeness was evaluated with Benchmarking Universal Single-Copy Orthologs (BUSCO) [27] (v6.1.0) against the mammalian single-copy ortholog set (mammalia_odb10) using the “--mode genome” option. Synteny between the newly assembled genomes and the reference genome (Sscrofa11.1) was analyzed using minimap2 [28] (v2.30) with “-asm 5” parameter and visualized with the pafr R package (v0.0.2; <https://github.com/dwinter/pafr> (accessed on 9 December 2025)).

2.4. Genome Annotation

Protein-coding genes were annotated by mapping the Sscrofa11.1 annotation file (Sus_scrofa.Sscrofa11.1.115.gff3) onto the assembled genomes using LiftOn [29] (v1.7.0). Transcript and protein sequences were extracted using gffread [30] (v0.12.7). Gene pairs located within collinearity blocks—identified from coding sequence alignments between the assembled genomes and the pig reference—were visualized in karyotype plots using JCVI [31] (v1.5.9). The completeness of the annotated transcriptomes and proteomes for each assembled genome was evaluated separately with BUSCO [27] (v6.1.0) using “--mode transcriptome” and “--mode proteins” options, respectively.

Repeat sequences were identified with a homology-based approach using RepeatMasker (v4.1.1; <https://www.repeatmasker.org> (accessed on 9 December 2025)). The RMBlast (v2.9.0; <http://www.repeatmasker.org/rmbblast/> (accessed on 9 December 2025)) search engine was employed with the transposable element (TE) databases from Dfam (v3.2; <https://dfam.org> (accessed on 9 December 2025)) and Repbase [32] (v20181026). The

repeat landscape was visualized using the RepeatMasker utility scripts `calcDivergenceFromAlign.pl` and `createRepeatLandscape.pl`.

2.5. Structural Variant Calling

SVs were identified using two complementary approaches. First, ONT reads were aligned to the reference genome (Sscrofa11.1) using minimap2 [28] (v2.30) and SVs were called with Sniffles2 (v2.7.1; <https://github.com/fritzsedlazeck/Sniffles> (accessed on 9 December 2025)) under default parameters. Second, to detect SVs from genome assemblies, pairwise whole-genome alignments were performed with minimap2 [28] (v2.30) and SVs were called using syri [33] (v1.7.1) with default parameters. From both call sets, only variants labeled as “PASS” and located on autosomes were retained. Furthermore, deletions (DELs), duplications (DUPs), inversions (INVs), and insertions (INs) larger than 50 bp were kept for downstream analysis. The SVs derived from the two methods were then merged using Truvari [34] (v5.4.0) with the “collapse” option to generate a non-redundant SV set. A Pig pan-genome SV panel [35] was downloaded (<http://animal.omics.pro/code/index.php/panPig> (accessed on 9 December 2025)) and compared against the non-redundant SVs set using Truvari [34] (v5.4.0) with the “bench” option.

2.6. Functional Annotation of SVs

Functional annotation of the identified SVs was performed using ANNOVAR [36]. SVs were categorized into seven groups based on genomic context: exonic and splicing (coding sequence variant), downstream (downstream gene variant), upstream (upstream gene variant), intronic (intron variant), intergenic (intergenic variant), UTR (UTR3 and UTR5), and others (ncRNA exonic, ncRNA intronic, and ncRNA splicing).

2.7. SVs Hotspot Identification

SV hotspots were identified using the “hotspotter” function from the *primat* R package (<https://github.com/daewoooo/primatR> (accessed on 9 December 2025)) with the parameters “bw = 200,000, pval = 1×10^{-8} , num.trial = 2000”. To assess whether these SV hotspots were enriched in protein-coding genes and functional genomic regions. We extracted 22,018 unique protein-coding genes from Sscrofa11.1 annotation file (Sus_scrofa.Sscrofa11.1.115.gff3) and downloaded annotation of potential regulatory elements from Pan et al., 2021 [37]. Permutation tests for each feature set were carried out with the *regioner* package (<https://github.com/bernatgel/regioner> (accessed on 9 December 2025)) over 1000 iterations to evaluate statistical significance. For comparison, the same permutation tests were performed using all SVs as the background.

2.8. Functional Enrichment Analysis

Genes and quantitative trait loci (QTLs) that overlapped with SV hotspots were selected for functional enrichment analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses were performed using KOBAS [38] (v3.0). For QTL enrichment analysis, QTL data were downloaded from PigBiobank (<https://pigbiobank.piggtex.bio/download> (accessed on 9 December 2025)), and enrichment was assessed using the GALLO R package [39]. Statistical significance was defined as an adjusted *p*-value < 0.05 based on the Benjamini–Hochberg value [40].

3. Results

3.1. High-Quality De Novo Assemblies for the DLY Pigs

Using the combined long- and short-read sequencing data, we assembled draft genomes of the eight DLY pigs with Flye [22]. These assemblies had an average total length of 2.49 Gb,

comprising 1787 contigs with a contig N50 of 24.36 Mb (Table S2). Subsequent phasing and chromosomal scaffolding yielded haplotype-resolved genome assemblies for each individual (Figure 1B). Relative to the pig reference genome (Sscrofa11.1), the haplotype-resolved assemblies ranged in size from 2.43 to 2.44 Gb, consisting of 687–1359 contigs with a contig N50 of 18.17–29.54 Mb and BUSCO completeness of 99.3–99.4% (Table 1; Figure 1C). Assembly base quality (QV) scores, estimated per individual with merquary using short-read data, averaged 43.02, exceeding that of the pig reference genome (QV = 36.48). Cumulative scaffold lengths of the 16 haplotype-resolved genomes demonstrated high assembly continuity (Table 1; Figure 1D). Moreover, the assembled sequences, ordered according to the reference genome, showed strong synteny with the reference (Figure S1).

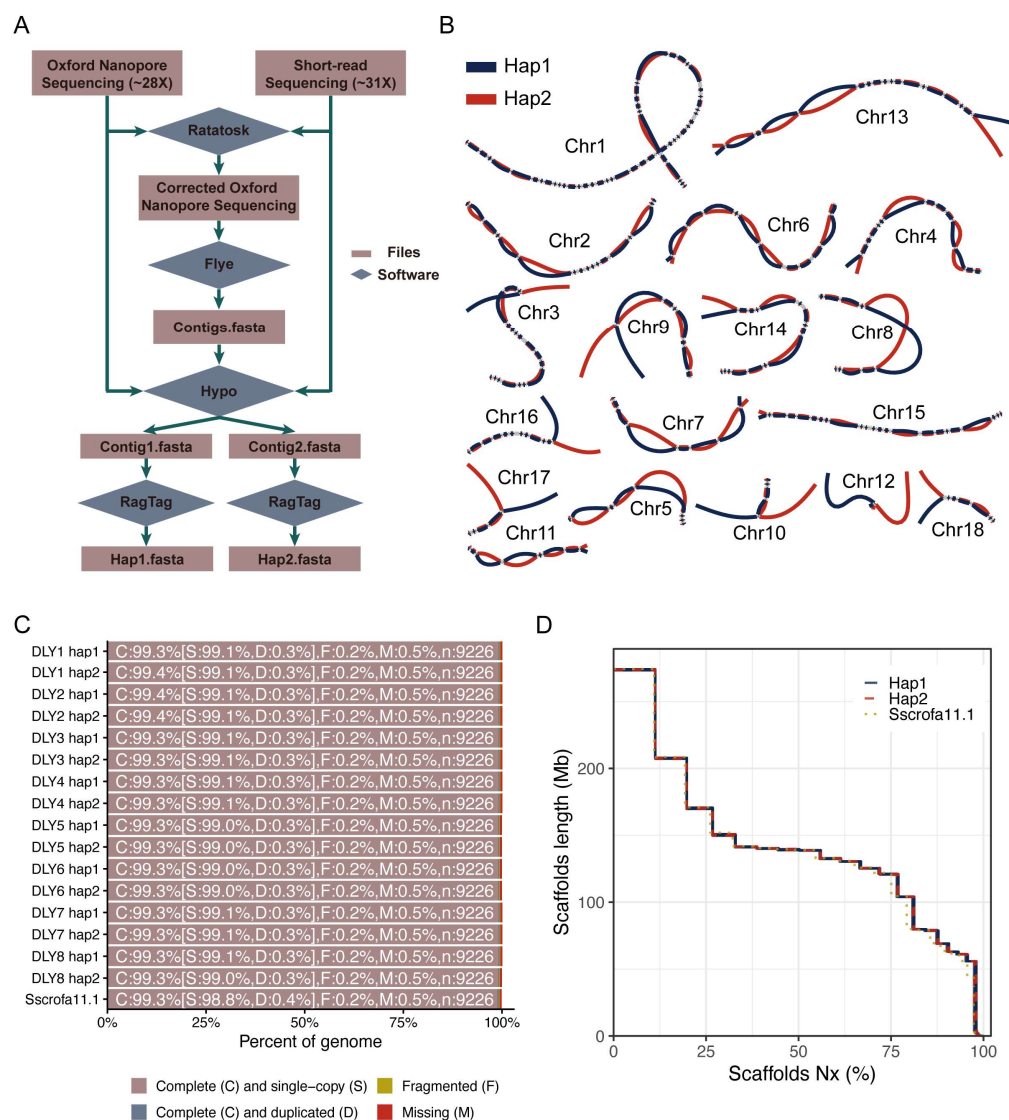


Figure 1. The genome assembly of the eight DLY pigs. **(A)** Workflow for the genome assembly of the eight DLY pigs. **(B)** Haplotype-resolved assemblies in the DLY pigs (using DLY2 as an example). Heterozygous regions between haplotype 1 (hap1) and haplotype 2 (hap2) in DLY2 are shown using bubbles across the assembly graphs of chromosomes (Chr). **(C)** BUSCO analysis of the DLY haplotype-resolved assemblies and the reference genome (Sscrofa11.1). **(D)** Cumulative lengths of the scaffolds in DLY haplotype-resolved assemblies and the reference genome (Sscrofa11.1). The x-axis indicates a scaffold Nx and the y-axis indicates the length of a scaffold Nx.

Table 1. Genome statistics of the DLY haplotype-resolved assemblies and the reference genome (Sscrofa11.1).

ID	Assembly	Assembly Length (Gb)	Contig Number	Contig N50 (Mb)	Scaffold N50 (Mb)	Largest Scaffold Length (Mb)	Quality Values (QV)
DLY1	Hap1/Hap2	2.44/2.44	858/873	21.95/21.96	139/138.78	274.13/274.18	43.08/40.76
DLY2	Hap1/Hap2	2.43/2.43	796/791	29.54/29.54	138.76/138.79	274.14/274.10	45.87/42.89
DLY3	Hap1/Hap2	2.44/2.44	1359/1355	18.18/18.17	138.79/138.81	274.02/273.93	43.76/42.47
DLY4	Hap1/Hap2	2.44/2.44	810/805	28.14/28.15	138.89/138.92	273.19/273.21	46.09/43.03
DLY5	Hap1/Hap2	2.43/2.43	802/816	24.28/24.29	139.68/139.89	273.33/273.36	42.53/40.79
DLY6	Hap1/Hap2	2.43/2.43	855/857	25.12/25.12	138.85/138.87	273.26/273.35	46.56/43.12
DLY7	Hap1/Hap2	2.43/2.43	754/757	24.22/24.23	139.26/139.15	273.84/273.84	43.32/41.79
DLY8	Hap1/Hap2	2.43/2.43	690/687	22.67/22.68	139.14/139.16	273.85/273.89	42.27/40.04
Sscrofa11.1	Primary	2.5	1157	41.89	138.97	274.33	36.48

3.2. Genome Annotation of the DLY Pigs

Repetitive elements were annotated across the 16 haplotype-resolved assemblies (Table S3; Figure S2). On average, 44.66% of each assembly was identified as repetitive sequence. Consistent with previous reports in pigs [19,41,42], LINEs constituted the most abundant repeat class (20.98% of each assembly), followed by SINEs (14.41%), LTRs (4.74%), and DNA transposons (2.45%).

Protein-coding gene annotation was performed by lifting over the reference annotation to the assemblies (Table 2). Between 21,846 and 22,001 genes (99.22–99.92%) were successfully transferred. The resulting annotations showed an average of 2.08 transcripts per gene, an average mRNA length of 59,939.94 bp, 11.7 exons per mRNA, and an average exon length of 269.31 bp. Completeness of the annotated transcriptomes and proteomes was assessed with BUSCO using the “mammalia_odb10” dataset (Table S4). For transcriptomes, an average of 96.32% of BUSCOs were complete, 1.42% fragmented, and 2.26% missing. For proteomes, 95.38% were complete, 1.96% fragmented, and 2.66% missing.

Table 2. Gene annotation statistics of the DLY haplotype-resolved assemblies and the reference genome (Sscrofa11.1).

ID	Assembly	Number of Putative Coding Genes	Number of mRNA	Average mRNA Length (bp)	Average CDS Length (bp)	Average Exons per mRNA	Average Exon Length (bp)
DLY1	Hap1/Hap2	21,930/21,910	45,688/45,652	59,961/59,833	1712/1695	11.7/11.7	269/269
DLY2	Hap1/Hap2	21,976/21,909	45,732/45,664	59,829/59,966	1705/1687	11.7/11.7	270/269
DLY3	Hap1/Hap2	21,923/21,867	45,663/45,607	59,832/60,018	1714/1711	11.7/11.7	269/269
DLY4	Hap1/Hap2	21,948/21,973	45,718/45,738	59,903/59,876	1715/1704	11.7/11.7	270/269
DLY5	Hap1/Hap2	22,001/21,970	45,740/45,709	59,872/59,925	1711/1699	11.7/11.7	270/270
DLY6	Hap1/Hap2	21,898/21,846	45,636/45,578	60,012/60,021	1713/1702	11.7/11.7	269/269
DLY7	Hap1/Hap2	21,934/21,887	45,692/45,641	59,940/60,036	1711/1700	11.7/11.7	270/269
DLY8	Hap1/Hap2	21,924/21,862	45,634/45,564	59,963/60,052	1692/1654	11.7/11.7	269/269
Sscrofa11.1	Primary	22,018	45,958	60,112	1732	11.8	268

Furthermore, gene pairs located within collinearity blocks between each haplotype-resolved assembly and Sscrofa11.1 were compared, demonstrating a high degree of coding sequence conservation (Figure S3). Together, these annotation results confirm the high quality and functional completeness of the DLY pig genome assemblies.

3.3. Generating and Characterizing a Catalog of SVs in DLY Pigs

SVs represent a major class of genetic variation [43], yet their detection in individual samples is often obscured by large-scale genomic synteny and collinearity (Figures S1 and S3). To comprehensively characterize SVs in the DLY pigs, we employed a combined alignment-

and assembly-based detection strategy. On average, we detected 51,851 SVs (ranging from 51,094 to 52,409) in each genome, covering 34.06 Mb (ranging from 32.56 Mb to 35.83 Mb) (Figure 2A). Consequently, we merged the high-confidence SVs detected from all the samples, and constructed a set of 130,416 non-redundant SV catalog (length ≥ 50 bp), comprising 55,379 deletions, 73,500 insertions, 928 inversions, and 609 duplications (Table S5).

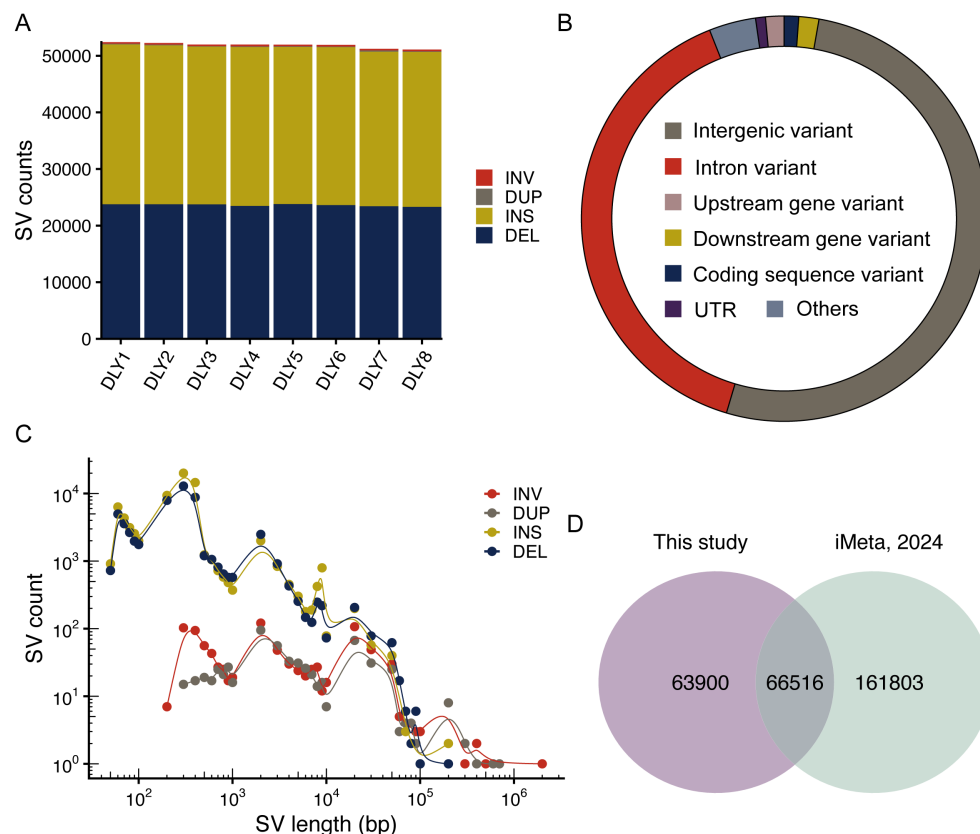


Figure 2. Characterization of SVs in DLY pigs. (A) The number of SV for each type per sample. (B) Functional annotation of SVs. (C) Length distribution of the SVs of each type. (D) Comparison analysis of the number of SVs.

In-depth annotation revealed that the majority of SVs were located in intergenic (51.83%) or intronic regions (39.42%), while only a small fraction (1.15%) of SVs was found overlapped with protein-coding sequences (Figure 2B; Table S6). The size distribution of SVs showed distinct peaks corresponding to known transposable elements (Figure 2C). For example, two peak at lengths of ~ 55 bp and ~ 276 bp were mainly annotated as SINEs, a peak at ~ 1407 bp corresponds to LTR, and a peak at ~ 7964 bp corresponds to LINEs. This pattern is consistent with previous reports [15,44] and underscores the role of transposable elements as a major source of SVs in the pig genome.

We further compared our SV call set with a previously published pig pan-genome SV panel (Figure 2D) [35]. Notably, 48.99% of the SVs identified here were novel, highlighting the enhanced detection sensitivity afforded by long-read sequencing relative to assembly-based methods. This expanded SV catalog provides a valuable resource for future pan-genome and structural variation studies in pig.

3.4. Identification and Enrichment Analysis of SV Hotspots

The genomic SVs altered gene expression, and the resulting transcriptional changes effectively explained variation in heterosis, supporting the dominance model and highlighting a prevalent role of SVs in its genetic basis [11]. Although three-way crossbreeding systems

are designed to utilize heterotic effects, the genome-wide landscape of structural variation in DLY pigs has not yet been systematically characterized. Here, we identified genomic regions enriched for SVs, hereafter referred to as “SV hotspots”, across the DLY pig genomes. Analysis revealed that SVs are non-randomly distributed [43,45], with 231 SV hotspots identified spanning approximately 203.69 Mb of the genome (Figure 3A; Table S7). To assess their functional relevance, we examined the overlap between these hotspots and annotated genomic features (Figure S4). SV hotspots showed significant enrichment in protein-coding genes (permutation test: $p = 0.001$, Z-score = 5.003). In contrast, when all SVs (non-redundant SVs) were tested against the same gene set, a significant depletion was observed (permutation test: $p = 0.001$, Z-score = -6.082). Similarly, comparison with putative regulatory elements [37] revealed significant overrepresentation of SV hotspots in these regions (permutation test: $p = 0.002$, Z-score = 3.145), whereas using all SVs as background showed significant depletion in regulatory elements (permutation test: $p = 0.001$, Z-score = -21.294).

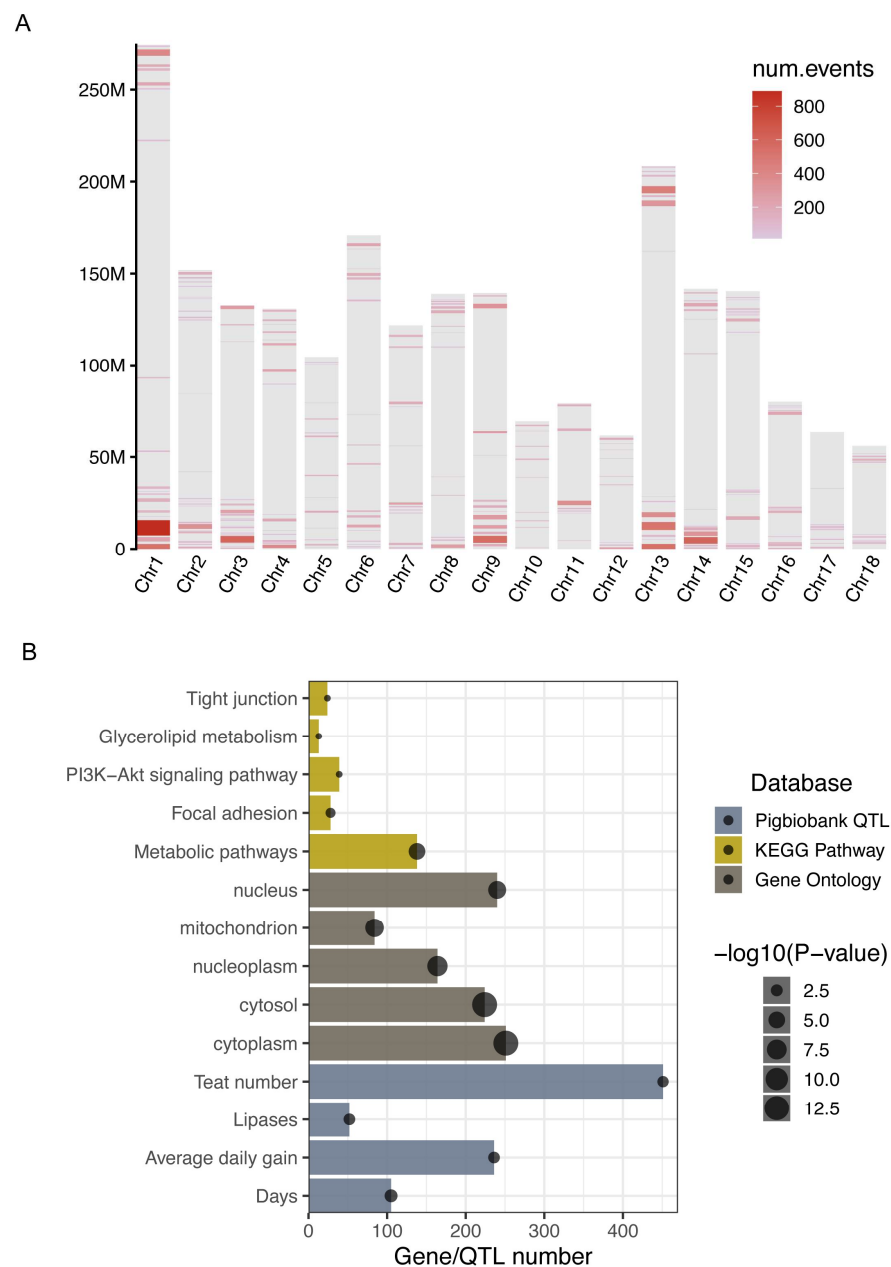


Figure 3. Enrichment analysis of SV hotspots. (A) Genome-wide distribution of SV hotspots. (B) KEGG, GO, and QTL enrichment analysis of SV hotspots. The detailed definitions for the enriched QTL terms are provided in Table S9.

A total of 2705 protein-coding genes and 4510 pig QTLs overlapped these SV hotspots. Functional enrichment analysis of the overlapping genes revealed several significantly enriched GO terms and KEGG pathways, including ATP binding, oxidation–reduction process, metabolic pathways, and fatty acid degradation (Figure 3B; Table S8). QTL enrichment analysis further indicated that these hotspots are implicated in growth-related traits such as days and average daily gain (Figure 3B; Table S9). As a representative example, a notable 3.43 Mb SV hotspot (Chr1:268,359,526–271,794,677) overlapped with a previously reported QTL (Chr1:270,153,237–271,111,196) associated with average daily gain (Table S10). Within this region, five candidate genes (*NCS1*, *HMCN2*, *FUBP3*, *ABL1*, and *FIBCD1*) were prioritized based on gene function and literature support. Further annotation identified eight high-frequency candidate SVs (present in $\geq 87.5\%$ of DLY pigs) within this hotspot-QTL overlap that were enriched in open chromatin regions (Table S11), highlighting its possible role in modulating gene expression by altering the cis-regulatory elements that may contribute to heterosis in DLY pigs.

4. Discussion

Advances in long-read sequencing technologies and assembly algorithms have revolutionized genome assembly. Nevertheless, only 24 chromosome-level pig genomes are currently available in the NCBI database. In this study, we combined Oxford Nanopore long reads with short reads to generate 16 haplotype-resolved, chromosome-level genome assemblies that accurately capture the haplotype diversity of DLY commercial pigs. Our assemblies exhibit high accuracy, continuity, and completeness, and are expected to serve as an indispensable genomic resource for future studies. They will facilitate detailed haplotype comparisons, enhance the identification of heterozygous variants, and enable the assessment of genetic diversity at the individual level.

Although strong synteny and collinearity were observed between the newly assembled genomes and the pig reference genome (*Sscrofa11.1*), numerous SVs and sequence differences were also detected. The SVs identified in this study spanned approximately 98.94 Mb, representing about 3.95% of the pig reference genome. This SV profile differs from previously reported patterns of genomic variation [14,15], which can largely be attributed to the unique genetic background of DLY pigs. As a crossbred population derived from crossing F1 sows (Landrace $\times$ Large White) with purebred Duroc boars, a substantial proportion of the genomic variation in DLY pigs may originate from the Landrace and Large White lineages, as well as from recombination-derived SVs generated during hybridization. Additionally, the eight DLY pigs studied included three males and five females. Notably, we observed no substantial differences in SV profiles between sexes (Figure 2A), which is likely due to our analysis being confined to autosomes.

Consistent with other studies [14,15], we observed that SVs were enriched near chromosomal ends, likely because telomeres and subtelomeric regions are particularly prone to mutation [46]. Taking the SV hotspot region on chromosome 1 (Chr1:268.35–271.79 Mb) as an example, previous studies have reported a highly significant peak in this interval (Table S10). Alleles in several genes located here are linked to growth performance. For instance, *FUBP3* has been implicated in skeletal development and loin eye area [47,48], while *ABL1* is a reported candidate gene for backfat thickness [49] and meat-to-fat ratio in pigs [50]. Other genes in this region, such as *NCS1*, *HMCN2*, and *FIBCD1*, although not yet studied in pigs, have been associated with bone mineral density and body mass index in humans [51,52], suggesting potential conserved roles in growth regulation. Further investigation revealed eight high-frequency SVs located in open chromatin regions within this hotspot (Table S11). Given that presence/absence variants represent a major class of SVs and have been shown to play important roles in gene-expression heterosis [53],

we speculate that these SVs may influence the expression of nearby genes by altering cis-regulatory elements, thereby contributing to phenotypic heterosis in pigs.

Lastly, it should be noted that scaffolding against Sscrofa11.1 in this study may have limited the detection of large-scale structural rearrangements. Furthermore, while chromosome-level continuity was achieved, gaps and potential misassemblies remain in highly complex genomic regions. Future efforts toward population-scale telomere-to-telomere assemblies will therefore be an important direction for refining structural variant discovery and genome completeness. We also observed variation in the number of putative coding genes annotated across haplotypes and individuals (Table 2). This variation likely reflects both biological differences, such as presence/absence variants that alter gene content (Table S12), and technical aspects of annotation transfer. Specifically, the accuracy of LiftOn is critically dependent on the accuracy of the source annotation, which may be less reliable for uncharacterized protein-coding genes. Consequently, the approach may fail to map such genes located in structurally complex or poorly aligned regions between the reference and the newly assembled genomes [29]. Indeed, many genes that were not successfully transferred are uncharacterized protein-coding genes, and a subset of these overlap with SV hotspots, assembly gaps, or unplaced contigs of the reference in our study (Table S12). These observations highlight that differences in gene counts across haplotypes and individuals result from a combination of genetic variation and limitations inherent to current annotation pipelines. Moving forward, integrating homology-based, RNA-seq-assisted, and ab initio gene predictions will be essential to achieve a more complete and accurate annotation of protein-coding genes in the pig genome. In addition, while we identified several candidate genes overlapping SVs, gene expression is often cell-, tissue-, stage-, or environment-specific [54,55]. The absence of parental genotype data and detailed growth phenotypes for the DLY pigs also limits our ability to directly correlate specific SVs with hybrid performance. Therefore, future studies incorporating multi-omics data, including transcriptomic, epigenomic, and phenotypic information, will be essential to functionally characterize these SVs and elucidate their roles in pigs.

5. Conclusions

Collectively, this study provides high-quality, haplotype-resolved, chromosome-level genome assemblies and constructs a comprehensive catalog of SVs for eight DLY commercial pigs. We demonstrate that SV hotspots are significantly enriched in protein-coding genes and potential regulatory elements, and highlight that high-frequency SVs within these regions likely contribute to economically important traits in commercial crossbred pigs. Our work establishes a foundational genomic resource that will support the fine-mapping of complex traits, facilitate haplotype-based selection, and advance the understanding of the genetic architecture underlying heterosis in pigs.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/biom16020214/s1>, Supplementary Method: Genomic DNA Extraction, Short-read Sequencing (DNBSEQ-T7), and Long-read Sequencing (Nanopore); Figure S1: Synteny analysis between the 16 DLY haplotype-resolved genomes and Sscrofa11.1; Figure S2: Sequence divergence of repetitive elements in the 16 DLY haplotype-resolved genomes; Figure S3: The alignment of coding sequences between 16 DLY haplotype-resolved genomes and Sscrofa11.1; Figure S4: The permutation result of annotated genomic features intersected with 231 SV hotspots; Table S1: The statistics of short- and long-reads sequencing sample; Table S2: The genome assembly statistics using Flye; Table S3: Repetitive elements across the 16 haplotype-resolved assemblies of the DLY pigs.; Table S4: Completeness of the annotated transcriptomes and proteomes; Table S5: Non-redundant SV catalog in DLY pigs; Table S6: Functional annotation of non-redundant SV catalog; Table S7: SV hotspots identified in this study; Table S8: GO and KEGG analysis using genes that

overlapped with SV hotspots.; Table S9: QTL enrichment analysis using QTLs that overlapped with SV hotspots; Table S10: A 3.4 Mb SV hotspot overlapped with QTLs that associated with average daily gain and days; Table S11: The chromatin state overlapped with 14 candidate SVs; Table S12: The genes that failed to transfer from the reference genome annotation.

Author Contributions: Z.W., J.Y., L.L. and Y.Q. conceived and designed the experiment. H.Q., Y.Q., S.D., S.W., Y.L. and M.L. collected the samples, and performed the experiments. J.W., H.Q. and Y.Q. analyzed the data. J.W., H.Q., Y.Q. and L.L. wrote the manuscript. Z.W., L.L. and J.Y. improved the manuscript. Z.W. contributed to the materials. All authors have read and agreed to the published version of the manuscript.

Funding: This project was supported by the National Major Agricultural Science and Technology Project (Grant Number: NK20221101), National Key Research and Development Program of China (Grant Number: 2023YFD1300200), Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 32502859 to L.L.), Key Technologies R&D Program of Guangdong Province project (Grant Number: 2022B0202090002 to Z.W.), Local Innovative and Research Teams Project of Guangdong Province (Grant Number: 2019BTO2N630 to J.Y.), South China Agricultural University discipline construction and development project (Grant Number: 2023B10564001 to L.L.), China Overseas Postdoctoral Recruitment Program (to L.L.).

Institutional Review Board Statement: All experimental protocols were approved by the Animal Care and Use Committee of the South China Agricultural University (Guangzhou, China) (approval No. SYXK 2019-0136; Approval date: 8 March 2019). No anesthesia or euthanasia was performed on the animals throughout this study.

Informed Consent Statement: Not applicable.

Data Availability Statement: The assemblies of the 16 DLY haplotype-resolved, chromosome-level genomes can be obtained from <https://doi.org/10.6084/m9.figshare.30951155> (accessed on 26 December 2025). Individual sequenced animals were proprietary properties of Guangdong Gene Bank of Livestock and Poultry. They may be requested by wzf@scau.edu.cn, respectively. Pig reference genome (Sscrofa11.1) and annotations (v115) can be obtained from ENSEMBL (<https://ftp.ensembl.org/pub/release-115> (accessed on 26 December 2025)). QTL data were downloaded from <https://pigbiobank.piggtex.bio/download> (accessed on 26 December 2025) (PigBiobank_release1. The list of lead SNP in 300 studies.csv.gz). The workflow for the DLY pig genome assembly is available at https://github.com/YibinQiu/ONT_assembly_workflow (accessed on 26 December 2025).

Acknowledgments: We are grateful to all individuals who contributed to this study but are not listed as authors. We specifically thank Zhibin Cao, Linhao Huang, Ying Sun, Donglin Ruan, Danyang Lin, Zekai Yao, Shanpeng Wang, and Fuchen Zhou for their technical assistance in DNA extraction from pig ear tissue samples. We also thank Gengyuan Cai, Jiajin Wu, Enqin Zheng, Sixiu Huang, and Zebin Zhang for their coordination and liaison work with the pig core breeding farms of Wens Foodstuff Group Co., Ltd. (Guangdong, China) facilitating the tissue sampling process. Finally, we thank all staff at the farms for their help with sample collection.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DLY	Duroc × (Landrace × Yorkshire)
SV	Structural variant
LD	Linkage disequilibrium
QTL	Quantitative trait locus
ONT	Oxford Nanopore Technologies
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology

References

1. Groenen, M.A.M.; Archibald, A.L.; Uenishi, H.; Tuggle, C.K.; Takeuchi, Y.; Rothschild, M.F.; Rogel-Gaillard, C.; Park, C.; Milan, D.; Megens, H.-J.; et al. Analyses of Pig Genomes Provide Insight into Porcine Demography and Evolution. *Nature* **2012**, *491*, 393–398. [\[CrossRef\]](#)
2. Kuhlers, D.L.; Jungst, S.B.; Little, J.A. An Experimental Comparison of Equivalent Terminal and Rotational Crossbreeding Systems in Swine: Pig Performance. *J. Anim. Sci.* **1994**, *72*, 2578–2584. [\[CrossRef\]](#)
3. Christensen, O.F.; Legarra, A.; Lund, M.S.; Su, G. Genetic Evaluation for Three-Way Crossbreeding. *Genet. Sel. Evol.* **2015**, *47*, 98. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Kim, S.W.; Less, J.F.; Wang, L.; Yan, T.; Kiron, V.; Kaushik, S.J.; Lei, X.G. Meeting Global Feed Protein Demand: Challenge, Opportunity, and Strategy. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 221–243. [\[CrossRef\]](#)
5. Veroneze, R.; Bastiaansen, J.W.; Knol, E.F.; Guimarães, S.E.; Silva, F.F.; Harlizius, B.; Lopes, M.S.; Lopes, P.S. Linkage Disequilibrium Patterns and Persistence of Phase in Purebred and Crossbred Pig (*Sus scrofa*) Populations. *BMC Genet.* **2014**, *15*, 126. [\[CrossRef\]](#)
6. Li, J.; Peng, S.; Zhong, L.; Zhou, L.; Yan, G.; Xiao, S.; Ma, J.; Huang, L. Identification and Validation of a Regulatory Mutation Upstream of the BMP2 Gene Associated with Carcass Length in Pigs. *Genet. Sel. Evol.* **2021**, *53*, 94. [\[CrossRef\]](#)
7. Zhang, C.; Wang, Z.; Bruce, H.; Kemp, R.A.; Charagu, P.; Miar, Y.; Yang, T.; Plastow, G. Genome-Wide Association Studies (GWAS) Identify a QTL Close to PRKAG3 Affecting Meat pH and Colour in Crossbred Commercial Pigs. *BMC Genet.* **2015**, *16*, 33. [\[CrossRef\]](#)
8. Zhuang, Z.; Wu, J.; Qiu, Y.; Ruan, D.; Ding, R.; Xu, C.; Zhou, S.; Zhang, Y.; Liu, Y.; Ma, F.; et al. Improving the Accuracy of Genomic Prediction for Meat Quality Traits Using Whole Genome Sequence Data in Pigs. *J. Anim. Sci. Biotechnol.* **2023**, *14*, 67. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Qiu, Y.; Zhuang, Z.; Meng, F.; Ruan, D.; Xu, C.; Ma, F.; Peng, L.; Ding, R.; Cai, G.; Yang, M.; et al. Identification of Candidate Genes Associated with Carcass Component Weights in Commercial Crossbred Pigs through a Combined GWAS Approach. *J. Anim. Sci.* **2023**, *101*, skad121. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Warr, A.; Affara, N.; Aken, B.; Beiki, H.; Bickhart, D.M.; Billis, K.; Chow, W.; Eory, L.; Finlayson, H.A.; Flicek, P.; et al. An Improved Pig Reference Genome Sequence to Enable Pig Genetics and Genomics Research. *GigaScience* **2020**, *9*, gaa051. [\[CrossRef\]](#)
11. Dan, Z.; Chen, Y.; Huang, W. Structural Variations Contribute to Subspeciation and Yield Heterosis in Rice. *Plant Biotechnol. J.* **2025**. [\[CrossRef\]](#)
12. Liu, S.; Yao, T.; Chen, D.; Xiao, S.; Chen, L.; Zhang, Z. Genomic Prediction in Pigs Using Data from a Commercial Crossbred Population: Insights from the Duroc × (Landrace × Yorkshire) Three-Way Crossbreeding System. *Genet. Sel. Evol.* **2023**, *55*, 21. [\[CrossRef\]](#)
13. Qiu, Y.; Liu, L.; Huang, M.; Ruan, D.; Ding, R.; Zhang, Z.; Zheng, E.; Wang, S.; Deng, S.; Meng, X.; et al. Origins, Dispersal, and Impact: Bidirectional Introgression between Chinese and European Pig Populations. *Adv. Sci.* **2025**, *12*, e2416573. [\[CrossRef\]](#)
14. Du, H.; Zhuo, Y.; Lu, S.; Li, W.; Zhou, L.; Sun, F.; Liu, G.; Liu, J.-F. Pangenome Reveals Gene Content Variations and Structural Variants Contributing to Pig Characteristics. *Genom. Proteom. Bioinform.* **2024**, *22*, qzae081. [\[CrossRef\]](#)
15. Jiang, Y.-F.; Wang, S.; Wang, C.-L.; Xu, R.-H.; Wang, W.-W.; Jiang, Y.; Wang, M.-S.; Jiang, L.; Dai, L.-H.; Wang, J.-R.; et al. Pangenome Obtained by Long-Read Sequencing of 11 Genomes Reveal Hidden Functional Structural Variants in Pigs. *iScience* **2023**, *26*, 106119. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Li, Z.; Liu, X.; Wang, C.; Li, Z.; Jiang, B.; Zhang, R.; Tong, L.; Qu, Y.; He, S.; Chen, H.; et al. The Pig Pangenome Provides Insights into the Roles of Coding Structural Variations in Genetic Diversity and Adaptation. *Genome Res.* **2023**, *33*, 1833–1847. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Liu, L.; Yi, G.; Yao, Y.; Liu, Y.; Li, J.; Yang, Y.; Liu, M.; Fang, L.; Mo, D.; Zhang, L.; et al. Multiomics Analysis Reveals Signatures of Selection and Loci Associated with Complex Traits in Pigs. *iMeta* **2024**, *3*, e250. [\[CrossRef\]](#)
18. Miao, J.; Wei, X.; Cao, C.; Sun, J.; Xu, Y.; Zhang, Z.; Wang, Q.; Pan, Y.; Wang, Z. Pig Pangenome Graph Reveals Functional Features of Non-Reference Sequences. *J. Anim. Sci. Biotechnol.* **2024**, *15*, 32. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Kwon, D.; Park, N.; Wy, S.; Lee, D.; Chai, H.-H.; Cho, I.-C.; Lee, J.; Kwon, K.; Kim, H.; Moon, Y.; et al. A Chromosome-Level Genome Assembly of the Korean Crossbred Pig Nanchukmacdon (*Sus scrofa*). *Sci. Data* **2023**, *10*, 761. [\[CrossRef\]](#)
20. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [\[CrossRef\]](#)
21. Holley, G.; Beyter, D.; Ingimundardottir, H.; Møller, P.L.; Kristmundsdottir, S.; Eggertsson, H.P.; Halldorsson, B.V. Ratatosk: Hybrid Error Correction of Long Reads Enables Accurate Variant Calling and Assembly. *Genome Biol.* **2021**, *22*, 28. [\[CrossRef\]](#)
22. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of Long, Error-Prone Reads Using Repeat Graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Darian, J.C.; Kundu, R.; Rajaby, R.; Sung, W.-K. Constructing Telomere-to-Telomere Diploid Genome by Polishing Haploid Nanopore-Based Assembly. *Nat. Methods* **2024**, *21*, 574–583. [\[CrossRef\]](#) [\[PubMed\]](#)

24. Alonge, M.; Lebeigle, L.; Kirsche, M.; Jenike, K.; Ou, S.; Aganezov, S.; Wang, X.; Lippman, Z.B.; Schatz, M.C.; Soyk, S. Automated Assembly Scaffolding Using RagTag Elevates a New Tomato System for High-Throughput Genome Editing. *Genome Biol.* **2022**, *23*, 258. [\[CrossRef\]](#)
25. Wick, R.R.; Schultz, M.B.; Zobel, J.; Holt, K.E. Bandage: Interactive Visualization of de Novo Genome Assemblies. *Bioinformatics* **2015**, *31*, 3350–3352. [\[CrossRef\]](#)
26. Rhie, A.; Walenz, B.P.; Koren, S.; Phillippy, A.M. Merquy: Reference-Free Quality, Completeness, and Phasing Assessment for Genome Assemblies. *Genome Biol.* **2020**, *21*, 245. [\[CrossRef\]](#)
27. Tegenfeldt, F.; Kuznetsov, D.; Manni, M.; Berkeley, M.; Zdobnov, E.M.; Kriventseva, E.V. OrthoDB and BUSCO Update: Annotation of Orthologs with Wider Sampling of Genomes. *Nucleic Acids Res.* **2025**, *53*, D516–D522. [\[CrossRef\]](#)
28. Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Chao, K.-H.; Heinz, J.M.; Hoh, C.; Mao, A.; Shumate, A.; Pertea, M.; Salzberg, S.L. Combining DNA and Protein Alignments to Improve Genome Annotation with LiftOn. *Genome Res.* **2025**, *35*, 311–325. [\[CrossRef\]](#)
30. Pertea, G.; Pertea, M. GFF Utilities: GffRead and GffCompare. *F1000Research* **2020**, *9*, 304. [\[CrossRef\]](#)
31. Tang, H.; Krishnakumar, V.; Zeng, X.; Xu, Z.; Taranto, A.; Lomas, J.S.; Zhang, Y.; Huang, Y.; Wang, Y.; Yim, W.C.; et al. JCVI: A Versatile Toolkit for Comparative Genomics Analysis. *iMeta* **2024**, *3*, e211. [\[CrossRef\]](#)
32. Bao, W.; Kojima, K.K.; Kohany, O. Repbase Update, a Database of Repetitive Elements in Eukaryotic Genomes. *Mob. DNA* **2015**, *6*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Goel, M.; Sun, H.; Jiao, W.-B.; Schneeberger, K. SyRI: Finding Genomic Rearrangements and Local Sequence Differences from Whole-Genome Assemblies. *Genome Biol.* **2019**, *20*, 277. [\[CrossRef\]](#) [\[PubMed\]](#)
34. English, A.C.; Menon, V.K.; Gibbs, R.A.; Metcalf, G.A.; Sedlazeck, F.J. Truvari: Refined Structural Variant Comparison Preserves Allelic Diversity. *Genome Biol.* **2022**, *23*, 271. [\[CrossRef\]](#)
35. Li, D.; Wang, Y.; Yuan, T.; Cao, M.; He, Y.; Zhang, L.; Li, X.; Jiang, Y.; Li, K.; Sun, J.; et al. Pangenome and Genome Variation Analyses of Pigs Unveil Genomic Facets for Their Adaptation and Agronomic Characteristics. *iMeta* **2024**, *3*, e257. [\[CrossRef\]](#)
36. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. *Nucleic Acids Res.* **2010**, *38*, e164. [\[CrossRef\]](#)
37. Pan, Z.; Yao, Y.; Yin, H.; Cai, Z.; Wang, Y.; Bai, L.; Kern, C.; Halstead, M.; Chanthavixay, G.; Trakooljul, N.; et al. Pig Genome Functional Annotation Enhances the Biological Interpretation of Complex Traits and Human Disease. *Nat. Commun.* **2021**, *12*, 5848. [\[CrossRef\]](#)
38. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J.; et al. KOBAS-i: Intelligent Prioritization and Exploratory Visualization of Biological Functions for Gene Enrichment Analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [\[CrossRef\]](#)
39. Fonseca, P.A.S.; Suárez-Vega, A.; Marras, G.; Cánovas, Á. GALLO: An R Package for Genomic Annotation and Integration of Multiple Data Sources in Livestock for Positional Candidate Loci. *GigaScience* **2020**, *9*, gaa149. [\[CrossRef\]](#)
40. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300. [\[CrossRef\]](#)
41. Du, H.; Hu, J.; Zhang, Z.; Wu, Z. Chromosome-Level Genome Assembly of the Meishan Pig and Insights into Its Domestication Mechanisms. *Animals* **2025**, *15*, 603. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Ma, H.; Jiang, J.; He, J.; Liu, H.; Han, L.; Gong, Y.; Li, B.; Yu, Z.; Tang, S.; Zhang, Y.; et al. Long-Read Assembly of the Chinese Indigenous Ningxiang Pig Genome and Identification of Genetic Variations in Fat Metabolism among Different Breeds. *Mol. Ecol. Resour.* **2022**, *22*, 1508–1520. [\[CrossRef\]](#)
43. Kosugi, S.; Momozawa, Y.; Liu, X.; Terao, C.; Kubo, M.; Kamatani, Y. Comprehensive Evaluation of Structural Variation Detection Algorithms for Whole Genome Sequencing. *Genome Biol.* **2019**, *20*, 117. [\[CrossRef\]](#)
44. Dai, X.; Bian, P.; Hu, D.; Luo, F.; Huang, Y.; Jiao, S.; Wang, X.; Gong, M.; Li, R.; Cai, Y.; et al. A Chinese Indicine Pangenome Reveals a Wealth of Novel Structural Variants Introgressed from Other *Bos* Species. *Genome Res.* **2023**, *33*, 1284–1298. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Ebert, P.; Audano, P.A.; Zhu, Q.; Rodriguez-Martin, B.; Porubsky, D.; Bonder, M.J.; Sulovari, A.; Ebler, J.; Zhou, W.; Serra Mari, R.; et al. Haplotype-Resolved Diverse Human Genomes and Integrated Analysis of Structural Variation. *Science* **2021**, *372*, eabf7117. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Alonge, M.; Wang, X.; Benoit, M.; Soyk, S.; Pereira, L.; Zhang, L.; Suresh, H.; Ramakrishnan, S.; Maumus, F.; Ciren, D.; et al. Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. *Cell* **2020**, *182*, 145–161.e23. [\[CrossRef\]](#)
47. Sato, S.; Uemoto, Y.; Kikuchi, T.; Egawa, S.; Kohira, K.; Saito, T.; Sakuma, H.; Miyashita, S.; Arata, S.; Kojima, T.; et al. SNP- and Haplotype-Based Genome-Wide Association Studies for Growth, Carcass, and Meat Quality Traits in a Duroc Multigenerational Population. *BMC Genet.* **2016**, *17*, 60. [\[CrossRef\]](#)

48. He, Y.; Ma, J.; Zhang, F.; Hou, L.; Chen, H.; Guo, Y.; Zhang, Z. Multi-Breed Genome-Wide Association Study Reveals Heterogeneous Loci Associated with Loin Eye Area in Pigs. *J. Appl. Genet.* **2016**, *57*, 511–518. [[CrossRef](#)]
49. Ma, G.; Tan, X.; Yan, Y.; Zhang, T.; Wang, J.; Chen, X.; Xu, J. A Genome-Wide Association Study Identified Candidate Regions and Genes for Commercial Traits in a Landrace Population. *Front. Genet.* **2024**, *15*, 1505197. [[CrossRef](#)]
50. Falker-Gieske, C.; Blaj, I.; Preuß, S.; Bennewitz, J.; Thaller, G.; Tetens, J. GWAS for Meat and Carcass Traits Using Imputed Sequence Level Genotypes in Pooled F2-Designs in Pigs. *G3* **2019**, *9*, 2823–2834. [[CrossRef](#)]
51. He, D.; Liu, H.; Wei, W.; Zhao, Y.; Cai, Q.; Shi, S.; Chu, X.; Qin, X.; Zhang, N.; Xu, P.; et al. A Longitudinal Genome-Wide Association Study of Bone Mineral Density Mean and Variability in the UK Biobank. *Osteoporos. Int.* **2023**, *34*, 1907–1916. [[CrossRef](#)] [[PubMed](#)]
52. Pulit, S.L.; Stoneman, C.; Morris, A.P.; Wood, A.R.; Glastonbury, C.A.; Tyrrell, J.; Yengo, L.; Ferreira, T.; Marouli, E.; Ji, Y.; et al. Meta-Analysis of Genome-Wide Association Studies for Body Fat Distribution in 694 649 Individuals of European Ancestry. *Hum. Mol. Genet.* **2019**, *28*, 166–174. [[CrossRef](#)] [[PubMed](#)]
53. Zhang, Y.; Fu, J.; Wang, K.; Han, X.; Yan, T.; Su, Y.; Li, Y.; Lin, Z.; Qin, P.; Fu, C.; et al. The Telomere-to-Telomere Gap-Free Genome of Four Rice Parents Reveals SV and PAV Patterns in Hybrid Rice Breeding. *Plant Biotechnol. J.* **2022**, *20*, 1642–1644. [[CrossRef](#)] [[PubMed](#)]
54. Chen, L.; Li, H.; Teng, J.; Wang, Z.; Qu, X.; Chen, Z.; Cai, X.; Zeng, H.; Bai, Z.; Li, J.; et al. Construction of a Multitissue Cell Atlas Reveals Cell-Type-Specific Regulation of Molecular and Complex Phenotypes in Pigs. *Adv. Sci.* **2025**, e04961. [[CrossRef](#)]
55. Han, B.; Li, H.; Zheng, W.; Zhang, Q.; Chen, A.; Zhu, S.; Shi, T.; Wang, F.; Zou, D.; Song, Y.; et al. A Multi-Tissue Single-Cell Expression Atlas in Cattle. *Nat. Genet.* **2025**, *57*, 2546–2561. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.