Main Manuscript for

Temporal orchestration of transcriptional and epigenomic programming underlying maternal embryonic diapause in a cricket model

Kosuke Kataoka[1,2*†], Yuta Shimizu[3†], Ryuto Sanno[4], Yuichi Koshiishi[5], Ken Naito[6], Kei Yura[7,8], Toru Asahi[2,4,8], Shin G. Goto[3*]

1. Division of Biotechnology and Life Science, Institute of Engineering, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei-shi, Tokyo 184-8588, Japan.
2. Comprehensive Research Organization, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan.
3. Graduate School of Science, Osaka Metropolitan University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan.
4. Graduate School of Advanced Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan.
5. NODAI Genome Research Center, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan.
6. Research Center of Genetic Resources, National Agriculture and Food Research Organization, 2-1-2 Kannondai, Tsukuba-shi, Ibaraki 305-8602, Japan.
7. Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan.
8. Department of Life Science and Medical Bioscience, Waseda University, 513 Waseda Tsurumaki-cho, Shinjuku-ku, Tokyo 162-0041, Japan.

[†] These authors equally contributed to this work.

* Kosuke Kataoka: kataokak@go.tuat.ac.jp
* Shin G. Goto: shingoto@omu.ac.jp

Author Contributions: K.K., Y. S., and S.G.G. designed research; K.K., Y.S., R.S., Y.K., K.N., K.Y., and T.A. performed research; K.K. analyzed data; and K.K., Y.S., and S.G.G. wrote the paper.

Competing Interest Statement: The authors have no competing interests to disclose.

Abstract

Maternal perception of environmental conditions can direct offspring developmental trajectories, providing adaptive flexibility across taxa. In the band-legged ground cricket *Dianemobius nigrofasciatus*, maternal exposure to short days induces embryonic diapause at the cellular blastoderm stage in offspring. Here, we investigate molecular mechanisms underlying this transgenerational adaptation through genome assembly (1.45 Gbp) and time-series transcriptomic analyses of diapause and non-diapause eggs from 12 to 72 hours post-oviposition. Despite morphological similarity, diapause-destined eggs show early upregulation of ATP-dependent chromatin remodeling genes at 24 hours. ATAC-seq reveals reduced chromatin accessibility at neural and cell cycle-related genes. Time-series clustering identifies precocious shifts in RNA processing machinery (peaking at 24 versus 40 hours in non-diapause eggs), followed by metabolic regulation toward amino acid catabolism and gluconeogenesis sustaining long-term survival during developmental arrest. Our findings reveal diapause as actively coordinated molecular programming involving epigenetic, transcriptional, and metabolic remodeling, providing insights into transgenerational environmental adaptation.

Main Text

Introduction

Across the tree of life, organisms have evolved remarkable strategies to survive temporarily hostile environments through dormancy, a physiological state characterized by dramatically reduced metabolic activity that enhances survival during adverse conditions [1,2]. Among the most widespread environmental challenges is seasonality, which poses significant physiological demands on organisms inhabiting temperate regions. To survive periods unsuitable for growth and reproduction, diverse taxa have independently evolved dormancy strategies: hibernation and torpor in mammals and birds, as well as embryonic developmental arrest in vertebrates and developmental pauses in numerous invertebrates, notably insects [3].

Among these strategies, diapause emerges as a programmed form of dormancy employed by many insect species in anticipation of seasonal stressors. Diapause can occur at any developmental stage—egg, larva, pupa, or adult—depending on the species-specific life cycle strategy [4]. A number of common physiological events are associated with diapause, including metabolic suppression and developmental arrest [5], enhanced resistance to cold and desiccation [6,7], and accumulation of energy reserves [8].

In many species, individuals enter diapause upon receiving a token stimulus. However, a particularly intriguing mechanism involves maternal influence, whereby the mother receives environmental signals and transfers this information to her progeny. This transgenerational effect is referred to as the maternal effect [9]. Maternal diapause induction is particularly valuable for embryonic diapause, a situation in which an individual has limited access to or time to independently evaluate seasonal conditions. This transgenerational adaptation has been well-documented across diverse taxa, including mites [10], crickets [11–13], rice leaf bug [14], flies [15], mosquitoes [16], and parasitic wasps [17,18].

Despite the growing interest in maternal diapause induction, the underlying molecular mechanisms remain poorly understood. Recent omics studies have highlighted the importance of alterations in endocrine signaling, cell proliferation, metabolism, and energy production in

maternal diapause induction [19,20]; however, these investigations have not comprehensively addressed how maternal influences impact specific molecular pathways initiating diapause, due to limited sampling periods. A comprehensive understanding of the molecular processes through which maternal environmental signals redirect the developmental trajectory towards diapause requires investigation of multiple time points throughout early development.

Dianemobius nigrofasciatus, the band-legged ground cricket, serves as a prominent model for investigating the maternal effect on embryonic diapause (Fig. 1a, b) [13,21]. Adult females reared under long-day conditions produce eggs that develop directly into nymphs without undergoing diapause (hereafter called long-day eggs), whereas those exposed to short-day conditions lay diapause-destined eggs (hereafter called short-day eggs) (Fig. 1b) [12,13]. A previous morphological study [22] established that long- and short-day eggs remain visually indistinguishable until 56 hours post-oviposition (hpo). During the first 40 hpo, energids (early embryonic cells) proliferate and migrate to the egg surface in both egg types. Subsequently, between 40 and 56 hpo, both types of eggs undergo cellular blastoderm formation. However, by 72 hpo, a clear morphological divergence emerges: long-day eggs proceed to form a germ band at the posterior pole, while short-day eggs arrest at the cellular blastoderm stage without germ band formation. Although the involvement of suppressed cell cycle regulators, small silencing RNAs, and segment patterning genes has been suggested [23], the precise molecular mechanisms underlying maternal diapause induction remain unknown.

In this study, we assembled a high-quality genome of *D. nigrofasciatus* and conducted a comprehensive time-series transcriptomic analyses of long- and short-day eggs to elucidate the molecular events occurring before and after the onset of embryonic diapause. Based on the morphological milestones, we strategically selected timepoints that would capture molecular changes before (12 and 24 hpo), during (40 and 56 hpo), and after (72 hpo) the critical developmental decision point where diapause and non-diapause trajectories diverge. Our findings suggest that this process is orchestrated not by a straightforward developmental switch, but rather by a sophisticated temporal orchestration of molecular programs.

Results

Chromosome-scale genome assembly and gene annotation of *D. nigrofasciatus*

To investigate the molecular mechanisms of maternal diapause induction, we generated a chromosome-scale genome assembly of *D. nigrofasciatus* using short-read, long-read, and Omni-C sequencing technologies. Flow cytometry revealed genome size differences between sexes, with males (1496.9 Mbp) having smaller genomes than females (1612.6 Mbp) (Fig. 1c; Supplementary Table 1), consistent with the XO/XX sex determination system in crickets [24]. *K*-mer analysis estimated a haploid genome size of 1.24 Gbp, which is smaller than the flow cytometry estimates, likely due to underrepresentation of repetitive sequences in short-read data (Supplementary Fig. 1; Supplementary Table 2).

The hybrid assembly strategy integrated 100.84× Illumina, 19.87× Oxford Nanopore, and 78.82× Omni-C coverage (Supplementary Table 3), resulting in a 1.45 Gbp assembly across 482 scaffolds with a scaffold N50 of 154 Mbp and minimal contamination (Supplementary Fig. 2). 87.1% of sequences were anchored into 10 linkage groups (Fig. 1d; Supplementary Fig. 3; Supplementary Tables 4, 5). Coverage analysis of male genomic sequences revealed that LG8 and LG9 constitute the X chromosome (Supplementary Fig. 4; Supplementary Table 6).

Assembly quality assessment using BUSCO (Benchmarking Universal Single-Copy Orthologs) [25] demonstrated high completeness, with 90.0% complete single-copy and 6.2% duplicated orthologs for the Arthropoda dataset (n=1,013) (Fig. 1e; Supplementary Table 4). When evaluated against the more specific Insecta dataset (n=1,367), the assembly showed even higher completeness with 93.2% single-copy and 3.0% duplicated orthologs (Fig. 1e; Supplementary Table 4). Gene annotation using BRAKER3 [26–28], StringTie2 [29], and GeMoMa [30] yielded 20,812 protein-coding genes (Supplementary Table 7). Phylogenetic analysis of 334 single-copy orthologs placed *D. nigrofasciatus* within Trigonidiidae, sister to Gryllidae (Supplementary Fig. 5),consistent with previous studies [31–33].

This high-quality genome assembly and gene annotation provide crucial resources for elucidating the mechanisms of transgenerational diapause programming.

Gradual molecular divergence in maternal diapause induction

To elucidate the molecular mechanisms underlying maternal diapause induction, we performed comprehensive time-series RNA-seq analysis of both long- and short-day eggs at five developmental timepoints: 12, 24, 40, 56, and 72 hpo (Fig. 2a). Principal component analysis (PCA) of the transcriptome data revealed intriguing patterns in the temporal dynamics of developmental divergence (Fig. 2b). Given that maternal diapause induction presumably involves signaling from mother to offspring, we initially expected substantial transcriptomic differences even at the earliest timepoint. However, molecular differences between long- and short-day eggs were minimal at 12-24 hpo, despite both egg types having been exposed to different maternal environments. The transcriptomic distance between the two egg types then expanded progressively from 24 to 72 hpo, with the most dramatic separation occurring between 56 and 72 hpo, coinciding with morphological divergence.

Differential expression analysis substantiated these observations, revealing a progressive increase in the number of differentially expressed genes (DEGs) between long- and short-day eggs over time ($|\log_2 \text{fold change}| > 1$, adjusted p-value < 0.05) (Fig. 2c, left panel; Supplementary Table 8). Remarkably, at 12 hpo, merely 31 genes showed differential expression (16 upregulated, 15 downregulated in short-day eggs). This number increased modestly to 120 DEGs at 24 hpo, then expanded more substantially to 371 DEGs at 40 hpo, 486 DEGs at 56 hpo, and culminated in 1,994 DEGs at 72 hpo (1,078 upregulated, 916 downregulated in short-day eggs).

Analysis of temporal gene expression changes between consecutive timepoints revealed three distinct developmental phases (Fig. 2c, right panel; Supplementary Table 9). The first major transcriptional transition (24-40 hpo) coincided with energid proliferation and migration, showing 3,316 DEGs in long-day eggs and 3,104 DEGs in short-day eggs. Transcriptional activity then decreased during cellular blastoderm formation (40-56 hpo), with 858 and 722 DEGs in long- and short-day eggs, respectively. The final phase (56-72 hpo) showed divergent patterns, particularly in long-day eggs (2,922 DEGs) compared to short-day eggs (2,031 DEGs), reflecting the initiation of germ band formation in long-day eggs while short-day eggs undergo molecular changes associated with diapause [22].

Temporal coordination and divergence of gene expression patterns reveal diapause preparation dynamics

To comprehensively characterize these temporal dynamics and identify genes with similar expression patterns across developmental time, we performed fuzzy c-means clustering using Mfuzz [34] on the time-series transcriptome data. This soft clustering assigns genes to multiple clusters with varying membership degrees. This feature makes it particularly suitable for capturing the continuous nature of developmental processes. Following independent clustering of each dataset (long- and short-day eggs), we identified cluster pairs between the two egg types that showed significant gene overlap using upper-tail hypergeometric tests (Fig. 3a; Supplementary Fig. 6, 7). These overlapping cluster pairs were further classified based on their temporal expression patterns: "homochronic clusters" displayed synchronized expression dynamics between egg types, while "heterochronic clusters" exhibited temporal shifts in peak expression timing, revealing potential differences in developmental timing between the two trajectories (Supplementary Fig. 8; Supplementary Table 10).

To understand which developmental processes remain conserved, we first examined homochronic gene clusters that maintained synchronized expression between egg types. The majority of homochronic cluster pairs exhibited peak expression during the early developmental periods (12-24 hpo), suggesting that fundamental embryonic processes are largely conserved before the diapause trajectory diverges (Supplementary Fig. 9; Supplementary Table 11). Analysis of these early-peaking homochronic clusters revealed common activation of essential cellular machinery: genes encoding components of clathrin-coated vesicles and membrane trafficking systems showed high initial expression followed by continuous decline, likely reflecting the early embryo's use of maternally supplied membrane and vesicle pools required for cellularization, together with maternally provided trafficking proteins and maternally deposited mRNAs that are translated during early development (Fig. 3b) [35–37]. Both egg types also showed coordinated activation of autophagy-related genes, peptidases, and catabolic processes at 12-24 hpo, suggesting a conserved program for recycling maternal components during early embryogenesis (Fig. 3b) [37–40].

In contrast, heterochronic clusters unveiled critical timing differences in key molecular pathways (Supplementary Fig. 10; Supplementary Table 12). Notably, genes involved in RNA processing and spliceosome function exhibited marked temporal advancement in short-day eggs (Fig. 3c). While these genes peaked at 40 hpo in long-day eggs, corresponding to the onset of cellular blastoderm formation, they reached maximum expression already at 24 hpo in short-day eggs. This precocious activation of RNA processing machinery coincides with the upregulation of chromatin remodeling factors, as shown in the subsequent epigenetic analyses (Fig. 4a, b), suggesting a coordinated early establishment of the molecular foundation required for developmental arrest.

Early epigenetic remodeling drives diapause preparation

Differentially expressed genes unveiled the molecular mechanisms initiating diapause (Supplementary Table 13). At 12 hpo, the sole significantly enriched pathway was glycolysis/gluconeogenesis, showing downregulation in short-day eggs (Fig. 4a). However, by 24 hpo, this initial metabolic shift had expanded into comprehensive molecular programming, with multiple pathways emerging as significantly enriched between egg types. Most strikingly, ATP-dependent chromatin remodeling genes showed specific upregulation in short-day eggs at this timepoint (Fig. 4a). Detailed examination revealed increased expression of core subunits of the Brahma (SWI/SNF) complex, including *Bap111*, *pont*, and *Bap60* (Fig. 4b). The Brahma complex is a critical and conserved regulator of chromatin accessibility and gene expression [41–43], suggesting that epigenetic remodeling may serve as a molecular switch for diapause induction.

To test this hypothesis, we performed ATAC-seq analysis on 24 hpo eggs to assess genome-wide chromatin accessibility. Short-day eggs exhibited reduced chromatin accessibility specifically near genes associated with neural development and cell cycle regulation (Supplementary Fig. 11; Supplementary Table 14, 15). Notable examples include *Kr-h1*, encoding a juvenile hormone-responsive transcription factor essential for developmental progression [44], and *crol*, an ecdysone-inducible transcription factor that recruits Polycomb group proteins [45,46]. Additionally, chromatin closure was observed at *Tlk* (Tousled-like kinase), which triggers S-phase entry and histone assembly [47], and *Ssrp*, a FACT complex subunit required for transcription initiation [48]. Neural signaling genes also showed reduced accessibility, including synaptic components such as *Tollo* (Toll-6) and *nAChRα1*. This selective chromatin closure likely contributes to the developmental arrest characteristic of diapause. The timing of these chromatin changes before morphological divergence suggests that epigenetic remodeling is not a consequence but rather a driver of diapause preparation.

Metabolic regulation and developmental signaling orchestrate diapause preparation

Following the early establishment of epigenetic and transcriptional regulatory programs, a comprehensive metabolic and developmental restructuring emerged as the next major phase of diapause preparation. While initial changes were subtle, a dramatic transformation unfolded from 40 hpo onwards, intensifying through the morphological divergence point.

The metabolic shift was characterized by coordinated activation of multiple amino acid degradation pathways. By 56-72 hpo, valine, leucine, and isoleucine degradation, tryptophan metabolism, lysine degradation, and beta-alanine metabolism were significantly upregulated in short-day eggs (Fig. 4a). This was accompanied by increased expression of key gluconeogenic enzymes, with phosphoenolpyruvate carboxykinase 1 (*Pepck1*) showing marked upregulation from 40 hpo and pyruvate carboxylase (*PCB*) increasing progressively from 56 hpo onwards (Fig. 4c). Genes encoding malate-aspartate shuttle components (*Got2* and *Mdh1/2*) also showed elevated expression in short-day eggs at 56-72 hpo (Supplementary Fig. 12), suggesting enhanced mitochondrial-cytoplasmic metabolite exchange. These enzymes enable the conversion of amino acid-derived carbon skeletons into glucose [8,49], suggesting a metabolic strategy to maintain energy homeostasis during impending developmental arrest.

Complementing this metabolic regulation, protein homeostasis pathways showed coordinated adjustment. From 56 hpo onwards, coinciding with peak metabolic restructuring, proteasome, protein processing in the endoplasmic reticulum, and spliceosome pathways were significantly suppressed in short-day eggs (Fig. 4a). This comprehensive downregulation of both protein synthesis and degradation machinery effectively pauses the cellular protein economy to preserve resources for extended diapause.

Concurrent with these metabolic changes, key developmental signaling pathways diverged between diapause and non-diapause trajectories. At 40 hpo, we observed significant downregulation of TGF-beta signaling in short-day eggs (Fig. 4a). As a master regulator of cell proliferation and differentiation [50,51], suppression of TGF-beta signaling effectively applies molecular brakes on developmental progression, occurring before morphological divergence.

Discussion

Our comprehensive molecular analysis reveals that maternal diapause induction in *D. nigrofasciatus* proceeds through a precisely orchestrated temporal cascade of molecular events (Fig. 4d). Both long- and short-day eggs initially share common early developmental processes, including endocytosis and autophagy for recycling maternal components (12-24 hpo). However, short-day eggs diverge through a distinctive molecular program: early activation of chromatin remodeling machinery and precocious RNA processing (24-40 hpo), followed by progressive metabolic restructuring with upregulation of amino acid catabolism and gluconeogenesis pathways that continues through 72 hpo. In contrast, long-day eggs maintain sustained protein synthesis machinery, RNA processing at normal developmental timing, and active cell cycle progression leading to cellular differentiation. This temporal progression demonstrates that diapause preparation is not a simple developmental arrest but rather an active programming process requiring substantial molecular investment before the morphological divergence point.

An important consideration in interpreting this temporal cascade is the timing of the maternal-to-zygotic transition (MZT), during which developmental control shifts from maternally deposited transcripts to embryo genome–derived transcription [52]. Although the precise timing of MZT has not been formally characterized in *D. nigrofasciatus*, previous analyses in this species demonstrated clear temporal increases in multiple developmental regulators, including segmentation-related genes, between 12 and 40 hpo, suggesting the onset of zygotic transcription during this period [23]. In line with these observations, our RNA-seq data show minimal photoperiod-dependent differences at 12 hpo, followed by progressive divergence from 24 hpo onward, suggesting that embryo genome transcription is already active by 12–24 hpo. Under this framework, maternal diapause induction is unlikely to operate solely through differential deposition of maternal mRNAs. Instead, maternal photoperiodic experience likely establishes distinct epigenetic states in the egg, such as chromatin accessibility, that subsequently bias embryo-driven transcriptional programs toward diapause or direct development. A similar principle has been demonstrated in *Drosophila melanogaster*, in which MZT occurs well before cellularization and early chromatin and transcriptional states critically shape later developmental trajectories [52].

The temporal advancement of components of RNA processing in short-day eggs, peaking at 24 hpo versus 40 hpo in long-day eggs, represents a striking example of heterochronic gene expression. This precocious activation likely serves multiple adaptive functions. First, the early investment in RNA processing capacity could enable the preparation of stable mRNA pools that persist throughout the diapause period [53,54]. Second, this enhanced RNA processing machinery may facilitate rapid transcriptional responses upon receiving diapause-terminating signals, allowing quick resumption of development [53,54]. Third, it may establish diapause-specific alternative splicing programs that generate protein isoforms optimized for dormancy [55]. The sustained elevation of these components throughout diapause preparation suggests they play a central role in maintaining the molecular readiness for eventual developmental resumption.

The early upregulation of Brahma complex components at 24 hpo, preceding any morphological changes, suggests that chromatin remodeling serves as the primary molecular switch for diapause induction. This finding aligns with emerging evidence that epigenetic mechanisms underlie phenotypic plasticity across diverse organisms [56–59]. The selective reduction in chromatin accessibility at developmental regulators such as *Kr-h1* and *crol* effectively locks cells into a pre-differentiation state, preventing the normal progression of embryonic development [44,45]. Both genes are central components of juvenile hormone (JH) and ecdysteroid signaling cascades, which are essential endocrine pathways governing insect growth and developmental progression.

Notably, however, our RNA-seq data did not reveal clear differences in juvenile hormone or ecdysteroid signaling pathways between long- and short-day eggs at 24 hpo. Consistent with this, prior physiological assays in *D. nigrofasciatus* did not detect egg JH III or photoperiod-dependent differences in egg 20-hydroxyecdysone during early development [60]. These observations suggest that early chromatin closure at endocrine-responsive developmental regulators does not reflect active hormonal signaling at this stage, but rather a preemptive epigenetic mechanism that suppresses the embryo's capacity to respond to later endocrine cues required for developmental progression. Within this framework, rather than being directly specified by maternally deposited Brahma transcripts, we propose that maternal photoperiodic perception establishes distinct epigenetic contexts in the egg, within which the Brahma complex might function as an executor that stabilizes and amplifies these maternally primed regulatory states. This epigenetic regulation occurs during a critical window (within 24 hpo), demonstrating how maternal environmental signals can be translated into stable epigenetic modifications that fundamentally alter offspring developmental trajectory. Intriguingly, this 24-hour window coincides with the period during which temperature pulses can still reverse diapause fate decisions [61], suggesting that chromatin remodeling may define developmental fate plasticity.

In *D. nigrofasciatus*, parental RNAi is technically feasible and could, in principle, enable functional validation of the Brahma complex [62]. However, knockdown of Brahma has been reported to severely impair fertility and fecundity in other insect species, including coleopteran and hemipteran models [63], rendering such approaches impractical for maintaining sufficient egg production in this system. Importantly, diapause fate in *D. nigrofasciatus* eggs can be experimentally induced or prevented by temperature manipulation [63], and diapause termination occurs spontaneously after prolonged arrest and is accelerated by winter-like chilling [64]. Therefore, examining Brahma complex expression and chromatin states under experimentally altered diapause induction and termination trajectories provides a realistic and informative framework for assessing its functional relevance *in vivo*.

The progressive metabolic restructuring observed from 56 hpo onwards in short-day eggs reveals a sophisticated survival strategy distinct from simple developmental arrest. The coordinated upregulation of amino acid catabolism pathways and gluconeogenic enzymes indicates a fundamental shift from anabolic growth metabolism to a catabolic maintenance mode. This metabolic flexibility enables short-day eggs to utilize maternally deposited amino acids as an energy source while preserving essential proteins needed for eventual development [53,54]. The timing of these changes, beginning before morphological divergence and intensifying through 72 hpo, indicates that metabolic programming is an integral preparatory component rather than a consequence of arrest. This preemptive metabolic adjustment, coupled with the suppression of protein synthesis and proteasome pathways, creates a molecular economy optimized for long-term survival while maintaining developmental competence [13,21].

The remarkable plasticity of *D. nigrofasciatus*, which can facultatively switch between diapause and direct development based on maternal photoperiod experience, makes it a unique model for understanding how environmental information is translated into developmental outcomes [64,65]. Our findings illuminate how subtle maternal signals can be progressively amplified into dramatically different developmental outcomes through self-reinforcing molecular cascades that combine epigenetic, transcriptional, and metabolic regulation. This temporal orchestration framework likely extends beyond insects to other organisms that employ maternal programming of developmental dormancy, including delayed implantation in mammals [65,66], suggesting that coordinated molecular timing represents a fundamental mechanism for transgenerational

371 environmental adaptation across diverse organisms. This cricket model system provides a
372 powerful framework for understanding the universal principles governing how organisms across
373 the tree of life integrate environmental information to orchestrate developmental decisions
374 through conserved molecular mechanisms.

375 Methods

376 Sample collection
377
378 The *D. nigrofasciatus* individuals used in this study was collected in a grassy field on the campus
379 of Osaka Metropolitan University (34°35′N, 135°30′E) and the bank of the Yamato-gawa River
380 (34°35′N, 135°30′E), Japan, from June to November 2022 and 2023. The rearing temperature
381 (25.0 ± 1.0 °C) was chosen to approximate ambient temperatures commonly experienced during
382 the natural oviposition seasons of this species in the Osaka region. Approximately 40–60
383 individuals were reared in a plastic case (29.5 cm width, 19.0 cm depth, and 17.0 cm height)
384 under long-day conditions (16-h light:8-h dark, 25.0 ± 1.0 °C) with a piece of moist cotton as a
385 water source and oviposition site. They were fed an artificial diet (Masu Chigyo Super EPC-2,
386 Marubeni Nisshin Feed Co., Ltd., Tokyo, Japan) and fresh carrot. The collected eggs were placed
387 on a piece of moist cotton in plastic dishes (50 mm diameter, 12 mm depth) and continuously
388 maintained under long-day conditions [66]. For the experiments, first instar nymphs were kept in
389 long-day conditions or were transferred to short-day conditions (12-h light:12-h dark, 25.0 ±
390 1.0 °C). Adult females reared under long-day conditions lay eggs destined to direct development
391 (non-diapause), whereas those under short-day conditions lay eggs destined to enter diapause [67].
392 In the present study, females reared under long- and short-day conditions are referred to as long-
393 and short-day females, respectively. Eggs laid by long- and short-day females are referred to as
394 long- and short-day eggs, respectively. Eggs deposited within a 2 h period were collected and
395 used for experiments. These eggs were continuously maintained under long-day conditions until
396 use for experiments. Under these conditions, the incidence of diapause was 14.6% in eggs laid
397 by long-day females and 95.7% in eggs laid by short-day females, confirming that diapause fate
398 was determined by maternal photoperiodic history rather than by environmental conditions
399 experienced by the eggs.
400

401 Genome and RNA sequencing
402
403 Alive *D. nigrofasciatus* was used to extract its genomic DNA. Total genomic DNA was extracted
404 from 10 whole male bodies after removal of the digestive tract, as well as from the head and hind
405 legs of a single male and female *D. nigrofasciatus*, using NucleoBond® HMW DNA (Macherey-
406 Nagel, Germany) according to the manufacturer's instructions. The resulting genomic DNA was
407 size-selected using a Short Read Eliminator Kit (PacBio, CA, USA). Oxford Nanopore
408 Technologies (ONT) sequencing libraries were then constructed and sequenced on the
409 PromethION platform (Oxford Nanopore Technologies, UK) with the ligation sequencing kit SQK-
410 LSK112 and flow cell R10.4. Base-calling was performed using Guppy basecalling software
411 v6.3.8 with the model dna_r10.4_e8.1_sup. Finally, we obtained 30.40 Gbp ONT sequencing
412 data; average and N50 read lengths were 19.11 Kbp and 31.18 Kbp, respectively. Additionally,
413 using the same DNA sample, we prepared a whole genome sequencing library using the TruSeq
414 DNA PCR-Free Library Prep Kit (Illumina, CA, USA) following the manufacturer's protocol. This
415 library was sequenced on the Illumina NovaSeq 6000 platform with 150 bp paired-end reads
416 (PE150), generating 152.7 Gbp of short-read data. For chromosome-scale scaffolding by the
417 Dovetail™ Omni-C™ Kit (Dovetail Genomics, CA, USA), the head and hind legs of another single
418 male *D. nigrofasciatus* were used according to the manufacturer's instructions. The Hi-C

419 sequencing library was built on the Illumina NovaSeq 6000 platform with PE150 and generated
420 127.1 Gbp raw data. DNA purity and concentrations were measured by spectrometry using
421 NanoPhotometer NP80-TOUCH (Implen, Germany) and fluorometry using Qubit 4 (Thermo
422 Fisher Scientific, MA, USA).
423
424 Total RNA was extracted from whole eggs 12, 24, 40, 56, and 72 hours post-oviposition (hpo)
425 using TRIzol Reagent with the PureLink RNA Mini Kit and PureLink DNase Set (Thermo Fisher
426 Scientific) after homogenizing with the BioMasher II tissue grinder (Nippi, Inc., Tokyo, Japan)
427 (Supplementary Table 16). Total RNA of 12, 24, 40, and 56 hpo eggs was purified with ssDNA /
428 RNA Clean & Concentrator (ZYMO Research, CA, USA). RNA integrity was confirmed using a
429 NanoDrop spectrophotometer (Thermo Fisher Scientific). For 12, 24, 40, and 56 hpo eggs,
430 libraries were constructed using the NEBNext® Single Cell/Low Input RNA Library Prep Kit (New
431 England Biolabs, MA, USA) and sequenced on an Illumina NextSeq 1000 platform with PE150.
432 For 72 hpo eggs, libraries were prepared using the SMART-Seq® HT PLUS Kit (Takara Bio,
433 Japan) and sequenced on an Illumina NovaSeq 6000 platform with PE150. To improve gene
434 annotation and provide tissue-specific expression profiles, we also performed RNA-seq on adult
435 tissues. Ovaries and neural tissues (brains and suboesophageal ganglions) were dissected from
436 10-day-old adult females at ZT 9-10, maintained under either long- or short-day photoperiod
437 conditions. For these adult tissue samples, mRNA was enriched using the NEBNext® Poly(A)
438 mRNA Magnetic Isolation Module, and libraries were prepared using the NEBNext® Ultra™ II
439 Directional RNA Library Prep Kit (New England Biolabs). Both brain and ovary libraries were
440 sequenced on an Illumina NovaSeq 6000 platform with PE150.
441
442 **Genome size estimation**
443
444 Genome size was computationally estimated using k-mer frequency distribution analysis. Raw
445 Illumina sequencing reads were processed using KMC v3 (K-mer Counter) [68] with k-mer length of
446 21. K-mers were counted. The resulting k-mer frequency histogram was generated using
447 kmc_tools with a maximum coverage cutoff of 10,000. The k-mer frequency distribution was
448 analyzed using GenomeScope2 [69] with k=21 and ploidy set to 2 (diploid).
449
450 Genome size was also estimated using flow cytometry to provide an independent validation of the
451 computational estimates. The 1C (haploid) genome size of males and females of *D.*
452 *nigrofasciatus* was estimated with reference to the method previously described [70]. Uniform
453 suspensions of single nuclei were extracted from a single *D. nigrofasciatus* head, a single
454 *Drosophila virilis* female head (1C =328 Mbp) and a portion of brain from a *Periplaneta americana*
455 male (1C =3300 Mbp) using the BD Cycletest Plus DNA Reagent Kit (BD Biosciences, CA, USA).
456 After trypsin and RNase A treatments, the samples were incubated with propidium iodide for 10
457 min in the dark at 4 °C. Finally, the sample was filtered through a 50-µm nylon mesh. Cellular
458 DNA content was assessed on the BD Accuri C6 Flow Cytometer (BD Biosciences). The obtained
459 data were analyzed using the FCS Express software (De Novo Software, CA, USA). Haploid (1C)
460 DNA quantity was calculated as (2C sample mean fluorescence/2C standard mean fluorescence)
461 times 328 Mbp for the *D. virilis* standard and times 3300 Mbp for the *P. Americana* standard. The
462 estimates based on the two standards were averaged to produce a 1C estimate for each sample.
463
464 ***De novo* genome assembly**
465
466 Initial genome assembly was performed using MaSuRCA v4.1.0 [71] with a hybrid approach
467 combining Illumina paired-end reads and Oxford Nanopore long reads. Illumina paired-end reads
468 (153 bp mean insert size, 102 bp standard deviation) were provided as input along with the

469 parameters FLYE_ASSEMBLY=0 to use CABOG assembly algorithm. The resulting primary
470 assembly scaffolds were used for subsequent processing.
471
472 To remove redundant haplotigs from the assembly, we employed Purge Haplotigs v1.1.2 [72].
473 Illumina paired-end reads were mapped to the primary assembly using minimap2 v2.24-r1122 [73]
474 with the -ax sr option. Read-depth histogram was generated using the purge_haplotigs hist
475 command to identify haploid and diploid peaks. Based on the histogram analysis showing an
476 average coverage of 94.5×, cutoff values were set as: low=5, midpoint=70, high=190, with junk
477 coverage threshold at 101. Coverage statistics were calculated using purge_haplotigs cov,
478 followed by the actual purging step using purge_haplotigs purge.
479
480 Potential contamination in the assembly was removed using BlobTools v1.1.127 [74], which
481 analyzes unexpected coverage, GC content, or similarity to bacterial and other contaminant
482 sequences. Sequence coverage was determined by mapping Illumina reads with bwa v0.7.17-
483 r1188. Similarity analysis was performed using BLASTn v2.13.0+ [75] against NCBI NT database
484 v5 (options: -task megablast -culling_limit 10 -evalue 1e-25 -outfmt '6 qseqid staxids bitscore std
485 sscinames sskingdoms stitle'). Mitochondrial genomes were identified through gene prediction
486 using the MITOS2 webserver (accessed on November 28, 2023) and subsequently removed. As
487 a result, one contig was removed as a bacterial genome and another as a mitochondrial genome.
488
489 The purged and decontaminated assembly was further scaffolded using Hi-C data. The input *de*
490 *novo* assembly and Dovetail OmniC library reads were processed using HiRise [76], a software
491 pipeline designed specifically for using proximity ligation data to scaffold genome assemblies.
492 Dovetail OmniC library sequences were aligned to the draft input assembly using bwa [77]. The
493 separations of Dovetail OmniC read pairs mapped within draft scaffolds were analyzed by HiRise
494 to produce a likelihood model for genomic distance between read pairs. This model was used to
495 identify and break putative misjoins, score prospective joins, and make joins above a threshold.
496
497 A Circos plot illustrating the distribution of genomic elements was generated using Circos v0.69-9
498 [78]. The plot displays the major scaffolds arranged in a circular layout with tracks showing gene
499 density, GC content, repeat element distribution, and other genomic features. The completeness
500 of the genome assembly was evaluated using BUSCO (Benchmarking Universal Single-Copy
501 Orthologs) v5.1.2 [78] implemented in gVolante [79]. Assembly completeness was assessed against
502 both the arthropoda_odb10 and insecta_odb10 lineage-specific ortholog datasets to evaluate the
503 representation of conserved genes at different taxonomic levels.
504
505 **Prediction of repeat regions and annotation of protein-coding genes**
506
507 For *de novo* repeat prediction, RepeatModeler v2.0.1 [80] was first used to identify novel repetitive
508 elements in the genome. This *de novo* library was supplemented with known repeat sequences
509 from closely related species to create a comprehensive repeat database. The combined repeat
510 library was then used to identify and softmask repetitive elements in the genome using
511 RepeatMasker v4.1.2 [81].
512
513 Structural annotation of protein-coding genes was performed on the soft-masked genome using
514 three complementary approaches: *ab initio* prediction, homology-based prediction, and RNA-seq-
515 based prediction. Each prediction method incorporated RNA-seq data from eggs, ovaries, and
516 brains as supporting evidence.
517

The *ab initio* prediction was carried out using BRAKER v3.0.2 [26–29,82,83], incorporating protein data from OrthoDB 10's arthropoda dataset and the mapping data of the filtered RNA-seq reads. This BRAKER prediction served as the foundation for our gene set. To complement and improve this base set, we employed two additional approaches. GeMoMa v1.9 [29] was used for homology-based prediction with gene sets from four model species (*Apis mellifera*, *Drosophila melanogaster*, and *Tribolium castaneum*), while StringTie2 v2.1.7 [83] was used for RNA-seq-based transcript assembly and gene prediction. The predictions from GeMoMa and StringTie2 were used to identify and add genes that BRAKER3 had missed. These combined predictions were merged and duplicate genes were removed using GffCompare v0.12.6 [84] with transcript classification codes set to u, i, and y to form a final, comprehensive consensus gene set.

Gene functional annotation was conducted using eggNOG-mapper online (http://eggnog-mapper.embl.de/) and BLASTp-based methods. For the BLASTp-based annotation, we searched against databases including *Homo sapiens* (GRCh38), *Mus musculus* (GRCm39), *Caenorhabditis elegans* (WBcel235), *Drosophila melanogaster* (BDGP6.32), and UniProt Swiss-Prot to identify the best hits for annotation (E-value < 1.0×10^{-5}). The completeness of the genome assembly was also evaluated using BUSCO v5.1.2 [78] implemented in gVolante [79].

Phylogenetic analysis

To reconstruct the phylogenetic relationships among cricket species, we analyzed protein sequences from eight species: *Acheta domesticus* [85], *Apteronemobius asahinai* [86], *Dianemobius nigrofasciatus*, *Gryllus bimaculatus* [87], *Laupala kohalensis* [87], *Locusta migratoria* [88], *Schistocerca gregaria* (GCF_023897955.1), and *Teleogryllus occipitalis* [89]. Translated coding sequences (CDS) from each species were used as input for ortholog identification.

Orthologous gene groups were identified using OrthoFinder v2.5.4 [90] with default parameters. Single-copy orthologous genes (SCOs) were extracted for phylogenetic reconstruction, resulting in 334 SCOs shared across all eight species. Each SCO was aligned using MAFFT v7.505 [91] with the --auto option, which selected the L-INS-i strategy. The aligned sequences were subsequently trimmed using trimAl v1.4.rev15 [92] with the -automated1 option to remove poorly aligned regions and gaps.

The trimmed alignments were concatenated, and a partition file was generated to define gene boundaries within the concatenated alignment. Maximum likelihood phylogenetic analysis was performed using IQ-TREE v2.2.0.3 [93] with the following parameters: -bb 1000 for ultrafast bootstrap analysis with 1000 replicates, -alrt 1000 for SH-like approximate likelihood ratio test with 1000 replicates, -m MFP for automatic model selection using ModelFinder [94], and partition model optimization (-sp).

Divergence times were obtained from TimeTree (https://timetree.org/) [95] and mapped onto the resulting phylogeny to provide temporal context for the evolutionary relationships among the analyzed species.

RNA-seq analysis

To prepare the reference transcriptome for quantification, the longest isoform for each gene was extracted from the gene models. Transcript abundance estimation was performed using Salmon v1.10.1 [96]. First, a quasi-mapping index was built from the longest isoform cDNA sequences

derived from the *D. nigrofasciatus* genome annotation. RNA-seq data from eggs collected at five developmental time points (12, 24, 40, 56, and 72 hpo) under both diapause (short-day) and non-diapause (long-day) conditions were analyzed. Paired-end reads were mapped using Salmon quant with automatic library type detection (-l A). Transcript-level estimates were imported and processed using the tximport package [97] in R. Transcripts per million (TPM) values and estimated read counts were extracted for downstream analysis.

Differential expression analysis was performed using DESeq2 v1.38.0 [98]. Low-expressed genes were filtered by removing genes with total read counts less than 10 across all samples. Library size normalization and dispersion estimation were performed using default DESeq2 parameters, followed by the Wald test for differential expression. Results were ordered by adjusted *p*-value (Benjamini-Hochberg correction) and exported for further analysis.

Gene Ontology (GO) and KEGG pathway enrichment analyses were performed using g:Profiler (https://biit.cs.ut.ee/gprofiler/) [99]. Differentially expressed genes were mapped to their *Drosophila melanogaster* orthologs based on the BLASTp annotation results, and the *D. melanogaster* gene IDs from the BDGP6.32 reference genome were used as input for enrichment analysis.

Principal component analysis

To visualize the overall transcriptional differences between diapause and non-diapause conditions across developmental time points, three-dimensional principal component analysis (PCA) was performed. TPM values from all samples were used. Genes with zero variance across all samples were removed prior to analysis. PCA was performed using the prcomp function in R with scaling (scale. = TRUE).

The first three principal components were extracted and visualized using the rgl package for 3D plotting. Samples were colored according to their diapause status (diapause vs. non-diapause) to assess the separation of conditions in the reduced dimensional space. The interactive 3D visualization was generated using plot3d functions.

Time-series clustering analysis, cluster comparison and classification

To identify genes with similar temporal expression patterns during embryonic development, time-series clustering was performed using the Mfuzz package [98]. First, normalized count data from DESeq2 were obtained for each developmental time point (12h, 24h, 40h, 56h, and 72h). The mean expression values were calculated separately for diapause and non-diapause conditions, and genes with zero standard deviation were removed.

To focus on genes with substantial expression changes across time points, genes with a coefficient of variation less than 1.0 were filtered out. This filtering step ensures that the clustering analysis captures biologically meaningful expression dynamics rather than noise from low-variance genes.

The filtered expression data were converted to ExpressionSet objects and standardized using the Mfuzz standardize function. The fuzziness parameter (m) was estimated using the mestimate function. Based on the elbow method analysis and practical considerations, the number of clusters was set to 20 for both diapause and non-diapause datasets.

To identify relationships between diapause and non-diapause expression patterns, cosine similarity was calculated between all pairwise combinations of cluster centroids. Cluster pairs were classified as "homochronic" (similar temporal dynamics) or "heterochronic" (distinct temporal dynamics) based on a cosine similarity threshold of 0.8. This classification allows for the identification of genes that maintain similar expression patterns across conditions versus those that show condition-specific temporal regulation. Gene overlap significance between clusters was assessed using upper-tail hypergeometric tests. Cluster pairs with *p*-values < 10^{-10} were considered significantly overlapping, indicating non-random association of genes between diapause and non-diapause conditions.

For functional interpretation, genes from cluster pairs in the top and bottom 10th percentiles of cosine similarity were extracted. These gene sets, representing the most homochronic and heterochronic patterns respectively, were subjected to GO and KEGG enrichment analysis using g:Profiler [99] to identify biological processes associated with conserved versus divergent temporal expression dynamics during diapause induction.

ATAC-seq library preparation and sequencing

Chromatin accessibility profiling was performed using ATAC-seq on 24 hpo eggs. For either long- or short-day conditions, two biological replicates were prepared (Supplementary Table 17).

Libraries were prepared using the ATAC-seq Kit (Active Motif, CA, USA). Frozen egg samples were lysed in 100 µL ice-cold ATAC Lysis Buffer and mechanically disrupted using scissors to release nuclei. Cell debris and eggshell fragments were removed by filtration through pluriStrainer-Mini 40 µm filters (pluriSelect, Germany). Subsequent library preparation steps were performed according to the manufacturer's protocol.

Library quality was assessed using NanoDrop One (Thermo Fisher Scientific) for concentration measurement and DNA ScreenTape (Agilent Technologies, CA, USA) for fragment size distribution analysis. Qualified libraries were sequenced on an Illumina NovaSeq 6000 platform with 150PE. The sequencing generated 11.15-11.60 Gbp of data per sample.

ATAC-seq data analysis

Raw sequencing reads were quality-trimmed using Skewer v0.2.2 [100] with the following parameters: maximum allowed error rate of 0.1, maximum allowed indel error rate of 0.03, and minimum read length of 18 bp after trimming. Paired-end mode was enabled for all samples.

Trimmed reads were aligned to the *D. nigrofasciatus* reference genome using Bowtie2 v2.3.4.2 [101] with parameters -N 1 (maximum mismatch of 1) and -X 2000 (maximum fragment length of 2000 bp). Peak calling was performed using MACS2 v2.1.2 [102] with a q-value threshold of 0.05 (-q 0.05) and parameters --nomodel --shift -75 --extsize 150, which defines a 150 bp region centered on the read start position.

To create a unified peak set, detected peaks from all samples were sorted and merged using bedtools v2.27.1 [103] with default settings. The resulting consensus peak regions were formatted into SAF format for downstream analysis.

Genomic annotation of peaks was performed by categorizing each peak into the following genomic regions: 1-5 kb upstream (1to5kb), <1 kb upstream (promoter), 5'UTR, exon, intron, 3'UTR, exon/intron boundary (defined as 100 bp surrounding exon-intron junctions), and intergenic regions. When peaks overlapped multiple features, priority was given in the following order: 5'-/3'-UTR > exon/intron boundary > exon/intron. Genes located within 1 Mbp of each peak were assigned as nearby genes.

Read counts for consensus peaks were quantified using featureCounts v1.6.3 [104]. Differential accessibility analysis between diapause-inducing and non-diapause conditions was performed using edgeR v3.26.8 [105] with TMM normalization. Log2 fold changes were calculated for each peak region. Peaks with uncorrected *p*-values < 0.05 were considered differentially accessible, and their associated nearby genes from different genomic contexts were extracted for functional analysis. Genomic contexts included: 1-5 kb upstream regions (1-5 kbp upstream of transcription start sites, TSS), promoter regions (<1 kbp upstream of TSS), 5'UTR, exon, intron, 3'UTR, exon/intron boundaries (within 100 bp of exon-intron junctions), and intergenic regions (areas not classified into any of the above categories).

Data availability

Raw read data used in this article (RNA-seq, ATAC-seq, Omni-C and WGS) have been submitted at DNA Data Bank of Japan (DDBJ) repository under accession no. PRJDB18895 and are publicly available as of the date of publication. The datasets are available in following link; https://ddbj.nig.ac.jp/resource/bioproject/PRJDB18895. The assembled genome and annotation datasets are available in https://doi.org/10.6084/m9.figshare.29665742.v1. The scripts used for the analyses in this study are available on GitHub (https://github.com/Kataoka-K-Lab/Dnigrofasciatus_EggDiapause.git).

Acknowledgments

This work was supported by the Cooperative Research Grant of the Genome Research for BioResource (NODAI Genome Research Center, Tokyo University of Agriculture), Cross-ministerial Moonshot Agriculture, Forestry and Fisheries Research and Development Program, "Technologies for Smart Bio-industry and Agriculture" (Bio-oriented Technology Research Advancement Institution, BRAIN) [Grant Number: JPJ009237], JSPS KAKENHI Grant-in-Aid for Scientific Research (C) [Grant Number: 21K05614], and Grant-in-Aid for JSPS Fellows [Grant Number: 21J23478/22KJ2609]. We thank Dr. Tarô Adati (Department of International Agricultural Development, Faculty of International Agriculture and Food Studies, Tokyo University of Agriculture) for his valuable advice and support.

Figures legends

Figure 1. Life cycle and genomic characterization of the band-legged ground cricket, *Dianemobius nigrofasciatus*. (a) Adult female *D. nigrofasciatus*. **(b)** Annual life cycle diagram in Osaka, Japan, illustrating photoperiod-dependent developmental transitions and diapause regulation. The outer ring represents the developmental stages (egg, nymph, and adult) and the diapause egg state. The inner circle shows seasonal progression with photoperiod-dependent phases: the pink area indicates the period with the longer critical day length, whereas the blue area indicates the period with the shorter critical day length. The first-generation adults emerge in September-October under short-day conditions and produce predominantly diapause-destined eggs that overwinter in the cellular blastoderm stage, whereas the second-generation adults appear in June-August under long-day conditions and produce non-diapause eggs. **(c)** Flow cytometric analysis of brain cells from adult females and males showing genome size estimation. Box plots show the distribution of genome size measurements, with blue numbers indicating mean ± 95% confidence interval for each sex. **(d)** Circos plot depicting genomic features across the ten linkage groups (LG1-10). From outer to inner tracks: gene density (I), repetitive element content (II), and GC content distribution (III). **(e)** BUSCO completeness assessment of the genome assembly and gene annotation against Arthropoda and Insecta databases via the gVolante web server. Colors indicate complete single-copy (solid green), completely duplicated (blue with diagonal stripes), fragmented (yellow crosshatched pattern), and missing (red with black dot pattern) genes.

Figure 2. Transcriptome dynamics during early embryogenesis in long- and short-day eggs of *D. nigrofasciatus*. (a) Schematic representation of developmental progression from 0 to 72 hours post-oviposition (hpo) at 25°C. Adults reared under short-day conditions produce diapause-destined eggs (short-day eggs), while those under long-day conditions produce eggs that undergo direct development (long-day eggs). The morphological features of early embryonic stages were characterized by Tanigawa et al. (2009). In both egg types, nuclei migrate from the yolk mass to the egg surface by 24 hpo. Cellular blastoderm formation occurs between 40-56 hpo in both trajectories. While long-day eggs proceed to form a germ band by 72 hpo, short-day eggs arrest development at the cellular blastoderm stage, maintaining this state throughout the diapause period. **(b)** Three-dimensional principal component analysis of transcriptome profiles across developmental time points. Red and blue marks indicate long- and short-day eggs, respectively. Dashed line circles indicate clustering of samples corresponding to each developmental time point. **(c)** *Left panel*: Number of differentially expressed genes (DEGs) between long- and short-day eggs at each developmental timepoint (12, 24, 40, 56, and 72 hpo). DEGs were defined as genes with $|\log_2$ fold change (FC)$| > 1$ and adjusted p-value < 0.05. Green bars with crosshatched pattern indicate genes upregulated in short-day eggs, while yellow bars with diagonal stripes represent genes downregulated in short-day eggs. *Right panels*: Temporal dynamics of gene expression changes within each egg type, comparing later versus earlier developmental timepoints. For long- (left) and short-day eggs (right), bars show the number of genes upregulated (green with crosshatched pattern) or downregulated (yellow with diagonal stripes) at the later timepoint relative to the earlier timepoint in each comparison ($|\log_2 FC| > 1$, adjusted p-value < 0.05).

Figure 3. Comparative analysis of temporal gene expression dynamics between long- and short-day eggs. (a) Heat map showing the significance of gene overlap between temporal expression clusters in long- and short-day eggs across five developmental timepoints (12, 24, 40, 56, and 72 hpo) as derived from an upper-tail hypergeometric test. Gene expression profiles were

clustered using the Mfuzz soft clustering algorithm. Numbers on the right and bottom axes represent cluster IDs for long-day and short-day eggs, respectively. Color intensity represents -$\log_{10}$(adjusted *p*-value), with darker red indicating more significant overlap between clusters. Homochronic clusters (highlighted in blue boxes) show conserved temporal expression patterns between the two egg types. Heterochronic clusters (highlighted in green boxes) exhibit shifted temporal dynamics between long- and short-day eggs. **(b1, b2, c1, and c2)** Expression trajectories show representative homochronic **(b)** or heterochronic **(c)** cluster pairs, where color intensity reflects membership scores calculated by Mfuzz (magenta indicating high membership and green indicating low membership). Bar graphs display functional enrichment analysis results using overlapping genes, with terms categorized as GO Biological Process (GO:BP, orange), GO Cellular Component (GO:CC, green), GO Molecular Function (GO:MF, blue), or KEGG pathways (red). For each category, the top five most significantly enriched terms (ranked by adjusted *p*-value) are shown; when fewer than five terms were significant, all enriched terms are displayed. Bar length represents -$\log_{10}$(adjusted *p*-value).

Figure 4. Temporal coordination of biological processes during early embryogenesis in long- and short-day eggs. (a) KEGG pathway enrichment analysis of differentially expressed genes (DEGs) identified between long-day and short-day conditions at each time point (adjusted *p*-value < 0.05). Dot size represents the percentage of DEGs annotated to each pathway, and color intensity indicates the adjusted *p*-value. **(b and c)** Temporal expression patterns (TPM values) of genes involved in chromatin remodeling **(b)** and gluconeogenesis **(c)**. The red frames in (b) highlight the critical periods of expression divergence between long- and short-day eggs. **(d)** Schematic representation of key biological processes and their temporal dynamics from 0 to 72 hpo. Both long- and short-day eggs show similar timing of membrane trafficking processes (endocytosis and autophagy) during early embryogenesis (12-24 hpo). Chromatin remodeling occurs within a specific time window around 24 hpo. Prior to diapause entry at the cellular blastoderm stage, short-day eggs show advanced activation of RNA processing and protein synthesis compared to long-day eggs. During later development (>56 hpo), short-day eggs activate amino acid metabolism and gluconeogenesis pathways, while long-day eggs proceed with neural development and cell cycle progression. Blue-shaded processes indicate molecular programs that are activated in both short-day (diapause) and long-day (non-diapause) eggs, whereas red-shaded processes indicate pathways that are specifically activated in only one developmental trajectory.

References

1. Lennon, J. T., den Hollander, F., Wilke-Berenguer, M. & Blath, J. Principles of seed banks and the emergence of complexity from dormancy. *Nat. Commun.* **12**, 4807 (2021).
2. Özgüldez, H. Ö. & Bulut-Karslioğlu, A. Dormancy, quiescence, and diapause: Savings accounts for life. *Annu. Rev. Cell Dev. Biol.* **40**, 25–49 (2024).
3. Wilsterman, K., Ballinger, M. A. & Williams, C. M. A unifying, eco-physiological framework for animal dormancy. *Funct. Ecol.* **35**, 11−31 (2021).
4. Denlinger, D. L. *Insect Diapause*. (Cambridge University Press, Cambridge, England, 2022).
5. J., M., Tauber, C. A. & Masaki, S. *Seasonal Adaptations of Insects*. (Oxford University Press, New York, NY, 1985).
6. Rinehart, J. P. *et al.* Up-regulation of heat shock proteins is essential for cold survival during insect diapause. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11130–11137 (2007).
7. Benoit, J. B. Water management by dormant insects: comparisons between dehydration resistance during summer aestivation and winter diapause. *Prog. Mol. Subcell. Biol.* **49**, 209–229 (2010).
8. Hahn, D. A. & Denlinger, D. L. Energetics of insect diapause. *Annu. Rev. Entomol.* **56**, 103–121 (2011).
9. Mousseau, T. A. & Fox, C. W. The adaptive significance of maternal effects. *Trends Ecol. Evol.* **13**, 403–407 (1998).
10. Oku, K., Yano, S. & Takafuji, A. Different maternal effects on diapause induction of tetranychid mites, *Tetranychus urticae* and *T. kanzawai* (Acari: Tetranychidae). *Appl. Entomol. Zool. (Jpn.)* **38**, 267–270 (2003).
11. Huestis, D. L. & Marshall, J. L. Interaction between maternal effects and temperature affects diapause occurrence in the cricket *Allonemobius socius*. *Oecologia* **146**, 513–520 (2006).
12. Masaki, S. Climatic adaptation and photoperiodic response in the band-legged ground cricket. *Evolution* **26**, 587–600 (1972).
13. Shiga, S. & Numata, H. Seasonal changes in the incidence of embryonic diapause in the band-legged ground cricket, *Dianemobius nigrofasciatus*. *Zool. Sci.* **14**, 1015–1018 (1997).
14. Shintani, Y. Effect of seasonal variation in host-plant quality on the rice leaf bug, *Trigonotylus caelestialium*. *Entomol. Exp. Appl.* **133**, 128–135 (2009).
15. Rockey, S., Miller, B. B. & Denlinger, D. A diapause maternal effect in the flesh fly, *Sarcophaga bullata*: Transfer of information from mother to progeny. *Journal of Insect Physiology* **35**, 553–558 (1989).
16. Bova, J., Soghigian, J. & Paulson, S. The prediapause stage of *Aedes japonicus japonicus* and the evolution of embryonic diapause in Aedini. *Insects* **10**, 222 (2019).
17. Schneiderman, H. A. & Horwitz, J. The induction and termination of facultative diapause in the chalcid wasps *Mormoniella vitripennis* (Walker) and *Tritneptis klugii* (Ratzeburg). *J. Exp. Biol.* **35**, 520–551 (1958).
18. Saunders, D. S. Larval diapause of maternal origin: Induction of diapause in *Nasonia vitripennis* (Walk.) (Hymenoptera: Pteromalidae). *J. Exp. Biol.* **42**, 495–508 (1965).
19. Poelchau, M. F., Reynolds, J. A., Elsik, C. G., Denlinger, D. L. & Armbruster, P. A. Deep sequencing reveals complex mechanisms of diapause preparation in the invasive mosquito, *Aedes albopictus*. *Proc. Biol. Sci.* **280**, 20130143 (2013).
20. Gong, J. *et al.* Early molecular events during onset of diapause in silkworm eggs revealed by transcriptome analysis. *Int. J. Mol. Sci.* **21**, 6180 (2020).

21. Kidokoro, T. & Masaki, S. Photoperiodic response in the relation to variable voltinism in the ground cricket, *Pteronemobius fascipes* Walker (Orthoptera: Gryllidae). *Jpn. J. Ecol* **28**, 291–298 (1978).
22. Tanigawa, N., Matsumoto, K., Yasuyama, K., Numata, H. & Shiga, S. Early embryonic development and diapause stage in the band-legged ground cricket *Dianemobius nigrofasciatus*. *Dev. Genes Evol.* **219**, 589–596 (2009).
23. Shimizu, Y., Tamai, T. & Goto, S. G. Cell cycle regulator, small silencing RNA, and segmentation patterning gene expression in relation to embryonic diapause in the band-legged ground cricket. *Insect Biochem. Mol. Biol.* **102**, 75–83 (2018).
24. Blackmon, H., Ross, L. & Bachtrog, D. Sex determination, sex chromosomes, and karyotype evolution in insects. *J. Hered.* **108**, 78–93 (2017).
25. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
26. Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* **24**, 637–644 (2008).
27. Stanke, M., Schöffmann, O., Morgenstern, B. & Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).
28. Gabriel, L., Hoff, K. J., Brůna, T., Borodovsky, M. & Stanke, M. TSEBRA: transcript selector for BRAKER. *BMC Bioinformatics* **22**, 566 (2021).
29. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019).
30. Keilwagen, J., Hartung, F. & Grau, J. GeMoMa: Homology-based gene prediction utilizing intron position conservation and RNA-seq data. *Methods Mol. Biol.* **1962**, 161–177 (2019).
31. Ito, Y. *et al.* Chromosome-scale whole genome assembly and annotation of the Jamaican field cricket *Gryllus assimilis*. *Sci. Data* **12**, 826 (2025).
32. Murata, K. *et al.* Complete mitochondrial genome sequences of two ground crickets, *Dianemobius fascipes nigrofasciatus* and *Polionemobius taprobanensis* (Orthoptera: Grylloidea: Trigonidiidae). *Mitochondrial DNA Part B* **8**, 1311–1315 (2023).
33. Sanno, R. *et al.* Comparative analysis of mitochondrial genomes in Gryllidea (Insecta: Orthoptera): Implications for adaptive evolution in ant-loving crickets. *Genome Biol. Evol.* **13**, evab222 (2021).
34. Kumar, L. & E Futschik, M. Mfuzz: a software package for soft clustering of microarray data. *Bioinformation* **2**, 5–7 (2007).
35. Pelissier, A., Chauvin, J.-P. & Lecuit, T. Trafficking through Rab11 endosomes is required for cellularization during *Drosophila* embryogenesis. *Curr. Biol.* **13**, 1848–1857 (2003).
36. Lecuit, T. Junctions and vesicular trafficking during *Drosophila* cellularization. *J. Cell Sci.* **117**, 3427–3433 (2004).
37. Lefebvre, F. A. & Lécuyer, É. Flying the RNA nest: *Drosophila* reveals novel insights into the transcriptome dynamics of early development. *J. Dev. Biol.* **6**, 5 (2018).
38. Kuhn, H., Sopko, R., Coughlin, M., Perrimon, N. & Mitchison, T. The Atg1-Tor pathway regulates yolk catabolism in *Drosophila* embryos. *Development* **142**, 3869–3878 (2015).
39. Harrison, M. M. & Eisen, M. B. Transcriptional activation of the zygotic genome in *Drosophila*. *Curr. Top. Dev. Biol.* **113**, 85–112 (2015).

875 40. Laver, J. D., Marsolais, A. J., Smibert, C. A. & Lipshitz, H. D. Regulation and function of
876 maternal gene products during the maternal-to-zygotic transition in *Drosophila*. *Curr. Top.*
877 *Dev. Biol.* **113**, 43–84 (2015).
878 41. Shi, J. *et al. Drosophila* Brahma complex remodels nucleosome organizations in multiple
879 aspects. *Nucleic Acids Res.* **42**, 9730–9739 (2014).
880 42. Möller, A., Avila, F. W., Erickson, J. W. & Jäckle, H. *Drosophila* BAP60 is an essential
881 component of the Brahma complex, required for gene activation and repression. *J. Mol. Biol.*
882 **352**, 329–337 (2005).
883 43. Papoulas, O. *et al.* The HMG-domain protein BAP111 is important for the function of the
884 BRM chromatin-remodeling complex *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 5728–5733
885 (2001).
886 44. Beck, Y., Pecasse, F. & Richards, G. Krüppel-homolog is essential for the coordination of
887 regulatory gene hierarchies in early *Drosophila* development. *Dev. Biol.* **268**, 64–75 (2004).
888 45. Mitchell, N., Cranna, N., Richardson, H. & Quinn, L. The ecdysone-inducible zinc-finger
889 transcription factor Crol regulates Wg transcription and cell cycle progression in *Drosophila*.
890 *Development* **135**, 2707–2716 (2008).
891 46. Erokhin, M. *et al.* Crol contributes to PRE-mediated repression and Polycomb group
892 proteins recruitment in *Drosophila*. *Nucleic Acids Res.* **51**, 6087–6100 (2023).
893 47. Carrera, P. *et al.* Tousled-like kinase functions with the chromatin assembly pathway
894 regulating nuclear divisions. *Genes Dev.* **17**, 2578–2590 (2003).
895 48. Chen, J. *et al.* Piwi regulates the usage of alternative transcription start sites in the
896 *Drosophila* ovary. *Nucleic Acids Res.* **53**, gkae1160 (2025).
897 49. Izadi, H. Endocrine and enzymatic shifts during insect diapause: a review of regulatory
898 mechanisms. *Front. Physiol.* **16**, 1544198 (2025).
899 50. Zhang, X.-S., Wang, Z.-H., Li, W.-S. & Xu, W.-H. FoxO induces pupal diapause by
900 decreasing TGFβ signaling. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2210404119 (2022).
901 51. Li, H.-Y., Lin, X.-W., Geng, S.-L. & Xu, W.-H. TGF-β and BMP signals regulate insect
902 diapause through Smad1-POU-TFAM pathway. *Biochim. Biophys. Acta Mol. Cell Res.* **1865**,
903 1239–1249 (2018).
904 52. Hamm, D. C. & Harrison, M. M. Regulatory principles governing the maternal-to-zygotic
905 transition: insights from Drosophila melanogaster. *Open Biol.* **8**, 180183 (2018).
906 53. Stadler, M. & Fire, A. Conserved translatome remodeling in nematode species executing a
907 shared developmental transition. *PLoS Genet.* **9**, e1003739 (2013).
908 54. Clouser, P. R., Riggs, C. L., Romney, A. L. T. & Podrabsky, J. E. Diapause and anoxia-
909 induced quiescence are unique states in embryos of the annual killifish, *Austrofundulus*
910 *limnaeus*. *Biomolecules* **15**, 515 (2025).
911 55. Steward, R. A., Pruisscher, P., Roberts, K. T. & Wheat, C. W. Genetic constraints in genes
912 exhibiting splicing plasticity in facultative diapause. *Heredity (Edinb.)* **132**, 142–155 (2024).
913 56. Li, B. *et al.* DNA methylation is correlated with gene expression during diapause termination
914 of early embryonic development in the silkworm (*Bombyx mori*). *Int. J. Mol. Sci.* **21**, 671
915 (2020).
916 57. Tian, Z. *et al.* The PBAP chromatin remodeling complex mediates summer diapause via
917 H3K4me3-driven juvenile hormone regulation in *Colaphellus bowringi*. *Proc. Natl. Acad. Sci.*
918 *U. S. A.* **122**, e2422328122 (2025).
919 58. Singh, P. P. *et al.* Evolution of diapause in the African turquoise killifish by remodeling the
920 ancient gene regulatory landscape. *Cell* **187**, 3338–3356 (2024).
921 59. Wrona, M. *et al.* The BRAHMA-associated SWI/SNF chromatin remodeling complex
922 controls *Arabidopsis* seed quality and physiology. *Plant Physiol.* **197**, kiae642 (2024).

923 60. Shimizu, Y. & Goto, S. G. No evidence for the involvement of juvenile hormone III and 20-
924 hydroxyecdysone in maternal decisions for embryonic diapause and diapause entry in the
925 band-legged ground cricket *Dianemobius nigrofasciatus* (Orthoptera: Trigonidiidae). *Appl.*
926 *Entomol. Zool.* (2023) doi:10.1007/s13355-023-00853-6.
927 61. Fukumoto, E., Numata, H. & Shiga, S. Effects of temperature of adults and eggs on the
928 induction of embryonic diapause in the band-legged ground cricket, *Dianemobius*
929 *nigrofasciatus*. *Physiol. Entomol.* **31**, 211−217 (2006).
930 62. Goto, S. G. & Nagata, M. The circadian clock gene (Clock) regulates photoperiodic time
931 measurement and its downstream process determining maternal induction of embryonic
932 diapause in a cricket. *Eur. J. Entomol.* **119**, 12–22 (2022).
933 63. Fishilevich, E. *et al.* Use of chromatin remodeling ATPases as RNAi targets for parental
934 control of western corn rootworm (*Diabrotica virgifera virgifera*) and Neotropical brown stink
935 bug (*Euschistus heros*). *Insect Biochem. Mol. Biol.* **71**, 58–71 (2016).
936 64. Masaki, S. Thermal relations of diapause in the eggs of certain crickets (Orthoptera:
937 Gryllidae). *Bull. Fac. Agric. Hirosaki Univ.* **6**, 5–20 (1960).
938 65. Ptak, G. E. *et al.* Embryonic diapause is conserved across mammals. *PLoS One* **7**, e33027
939 (2012).
940 66. McGowen, M. R., Erez, O., Romero, R. & Wildman, D. E. The evolution of embryo
941 implantation. *Int. J. Dev. Biol.* **58**, 155–161 (2014).
942 67. Goto, S. G., Doi, K., Nakayama, S. & Numata, H. Maternal control of cold and desiccation
943 tolerance in eggs of the band-legged ground cricket *Dianemobius nigrofasciatus* in relation
944 to embryonic diapause. *Entomol. Res.* **38**, 17–23 (2008).
945 68. Kokot, M., Dlugosz, M. & Deorowicz, S. KMC 3: counting and manipulating k-mer statistics.
946 *Bioinformatics* **33**, 2759–2761 (2017).
947 69. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot
948 for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
949 70. Johnston, J. S., Bernardini, A. & Hjelmen, C. E. Genome size estimation and quantitative
950 cytogenetics in insects. *Methods Mol. Biol.* **1858**, 15–26 (2019).
951 71. Zimin, A. V. *et al.* The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
952 72. Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotigs: Allelic contig
953 reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* **19**, 460 (2018).
954 73. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–
955 3100 (2018).
956 74. Laetsch, D. R. & Blaxter, M. L. BlobTools: Interrogation of genome assemblies. *F1000Res.*
957 **6**, (2017).
958 75. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421
959 (2009).
960 76. Putnam, N. H. *et al.* Chromosome-scale shotgun assembly using an *in vitro* method for
961 long-range linkage. *Genome Res.* **26**, 342–350 (2016).
962 77. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
963 *arXiv [q-bio.GN]* (2013).
964 78. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome*
965 *Res.* **19**, 1639–1645 (2009).
966 79. Nishimura, O., Hara, Y. & Kuraku, S. Evaluating genome assemblies and gene models
967 using gVolante. *Methods Mol. Biol.* **1962**, 247–256 (2019).
968 80. Flynn, J. M. *et al.* RepeatModeler2 for automated genomic discovery of transposable
969 element families. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 9451–9457 (2020).

81. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **Chapter 4**, 4.10.1-4.10.14 (2009).
82. Gabriel, L. *et al.* BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res.* **34**, 769–777 (2024).
83. Brůna, T., Lomsadze, A. & Borodovsky, M. GeneMark-ETP significantly improves the accuracy of automatic annotation of large eukaryotic genomes. *Genome Res.* **34**, 757–768 (2024).
84. Pertea, G. & Pertea, M. GFF utilities: GffRead and GffCompare. *F1000Res.* **9**, 304 (2020).
85. Dossey, A. T. *et al.* Genome and genetic engineering of the house cricket (*Acheta domesticus*): A resource for sustainable agriculture. *Biomolecules* **13**, 589 (2023).
86. Satoh, A., Takasu, M., Yano, K. & Terai, Y. *De novo* assembly and annotation of the mangrove cricket genome. *BMC Res. Notes* **14**, 387 (2021).
87. Ylla, G. *et al.* Insights into the genomic evolution of insects from cricket genomes. *Commun Biol* **4**, 733 (2021).
88. Wang, X. *et al.* The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* **5**, 2957 (2014).
89. Kataoka, K. *et al.* The draft genome dataset of the Asian cricket *Teleogryllus occipitalis* for molecular research toward entomophagy. *Front. Genet.* **11**, 470 (2020).
90. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
91. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
92. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
93. Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
94. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
95. Kumar, S. *et al.* TimeTree 5: An expanded resource for species divergence times. *Mol. Biol. Evol.* **39**, msac174 (2022).
96. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
97. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
98. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
99. Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
100. Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
101. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
102. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
103. Quinlan, A. R. BEDTools: The Swiss-army tool for genome feature analysis: BEDTools: The Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinformatics* **47**, 11.12.1-34 (2014).

1019 104. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for
1020 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
1021 105. Chen, Y., Chen, L., Lun, A. T. L., Baldoni, P. L. & Smyth, G. K. edgeR v4: Powerful
1022 differential analysis of sequencing data with expanded functionality and improved support
1023 for small counts and larger datasets. *Nucleic Acids Res.* **53**, gkaf018 (2025).

Figure 1

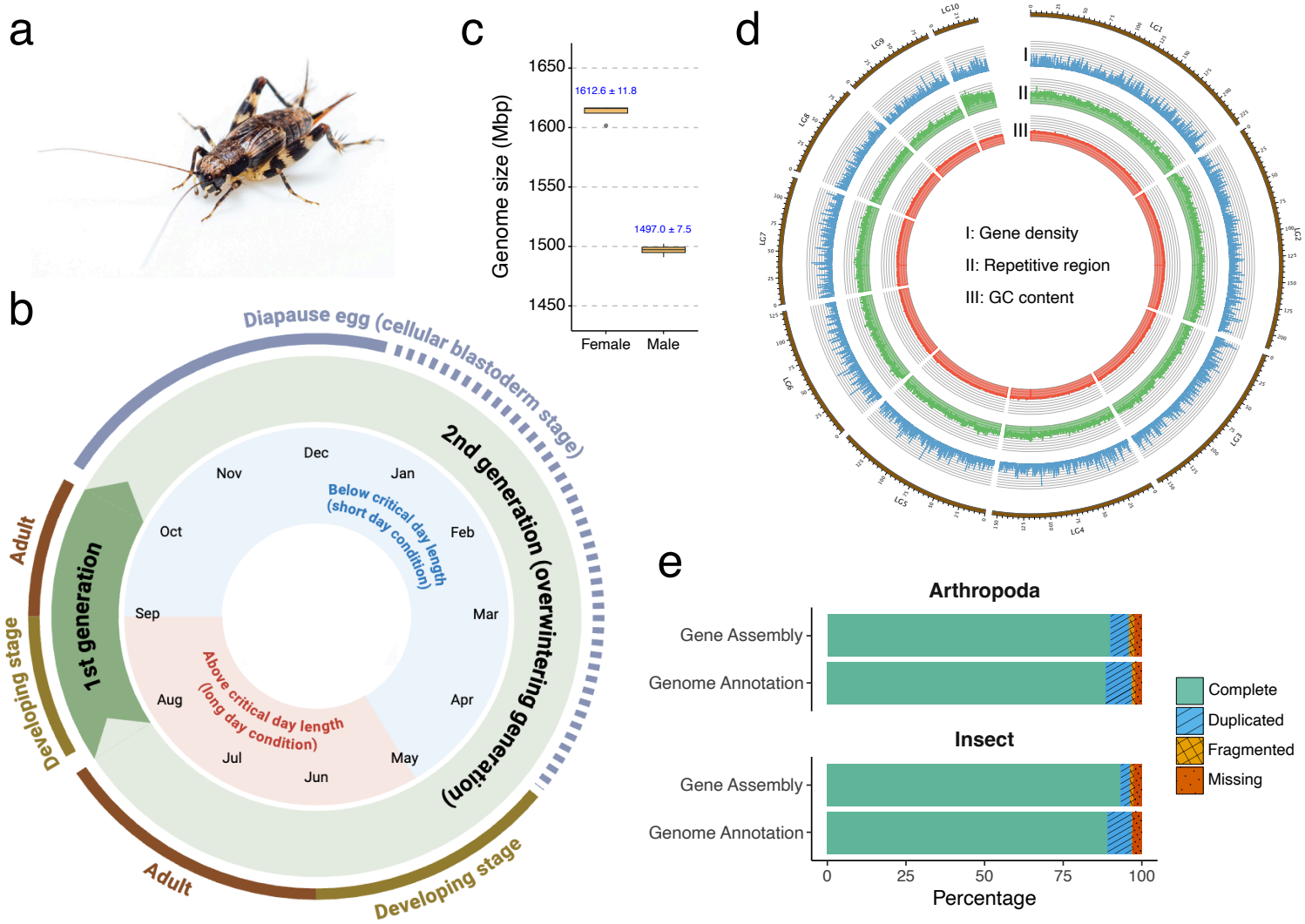

a
b
c
d
e
Genome size (Mbp)
1650
1600
1550
1500
1450
1612.6 ± 11.8
1497.0 ± 7.5
Female
Male
Diapause egg (cellular blastoderm stage)
2nd generation (overwintering generation)
1st generation
Adult
Developing stage
Adult
Developing stage
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Below critical day length
(short day condition)
Above critical day length
(long day condition)
I
II
III
I: Gene density
II: Repetitive region
III: GC content
Arthropoda
Insect
Gene Assembly
Genome Annotation
Complete
Duplicated
Fragmented
Missing
0
25
50
75
100
Percentage

Figure 2

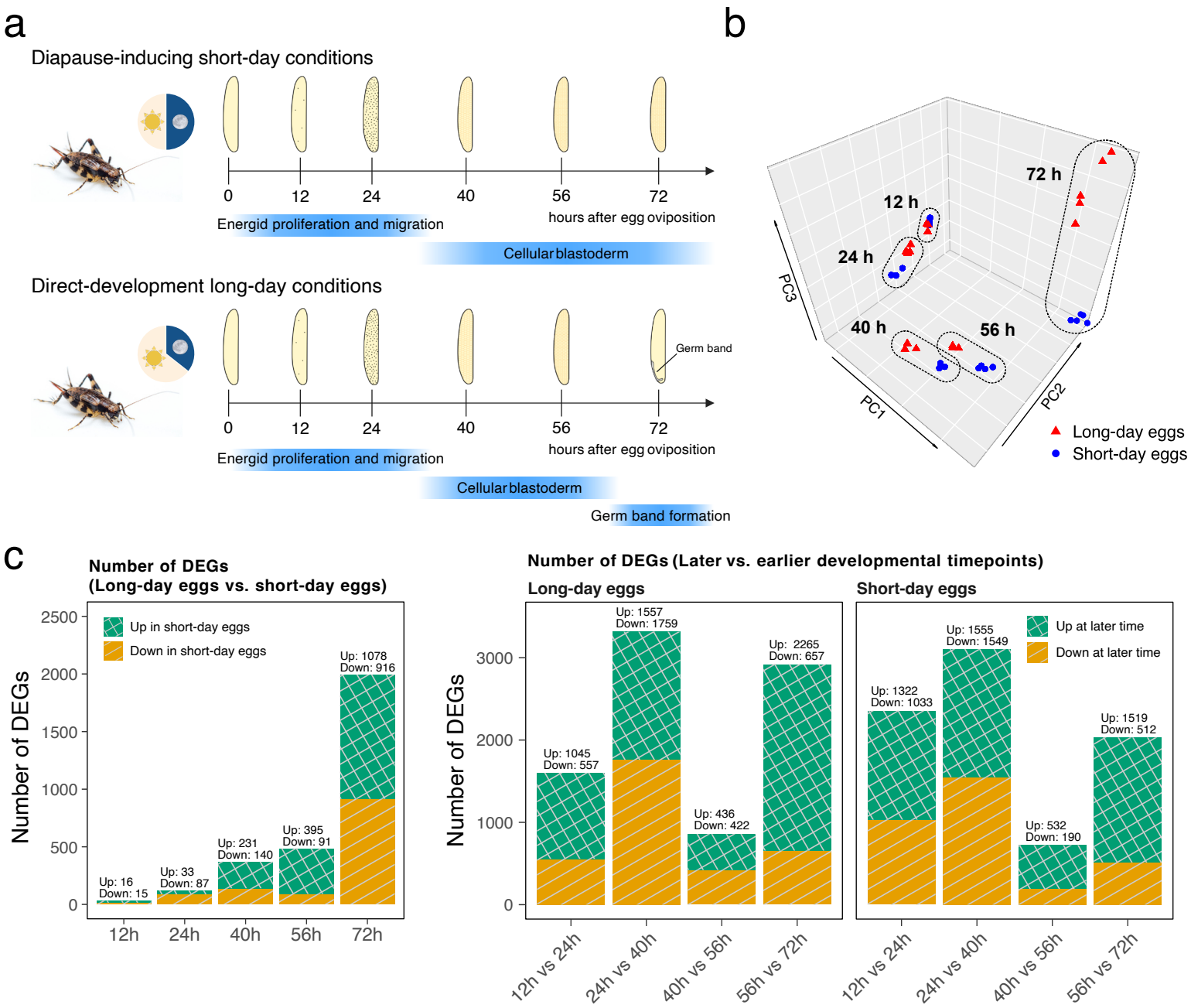

a
Diapause-inducing short-day conditions
0
12
24
40
56
72
hours after egg oviposition
Energid proliferation and migration
Cellular blastoderm
Direct-development long-day conditions
Germ band
Germ band formation
b
72 h
12 h
24 h
40 h
56 h
PC1
PC2
PC3
Long-day eggs
Short-day eggs
c
Number of DEGs
(Long-day eggs vs. short-day eggs)
Up in short-day eggs
Down in short-day eggs
Number of DEGs
Up: 16
Down: 15
Up: 33
Down: 87
Up: 231
Down: 140
Up: 395
Down: 91
Up: 1078
Down: 916
12h
24h
40h
56h
72h
Number of DEGs (Later vs. earlier developmental timepoints)
Long-day eggs
Short-day eggs
Up at later time
Down at later time
Up: 1045
Down: 557
Up: 1557
Down: 1759
Up: 436
Down: 422
Up: 2265
Down: 657
Up: 1322
Down: 1033
Up: 1555
Down: 1549
Up: 532
Down: 190
Up: 1519
Down: 512
12h vs 24h
24h vs 40h
40h vs 56h
56h vs 72h

Figure 3

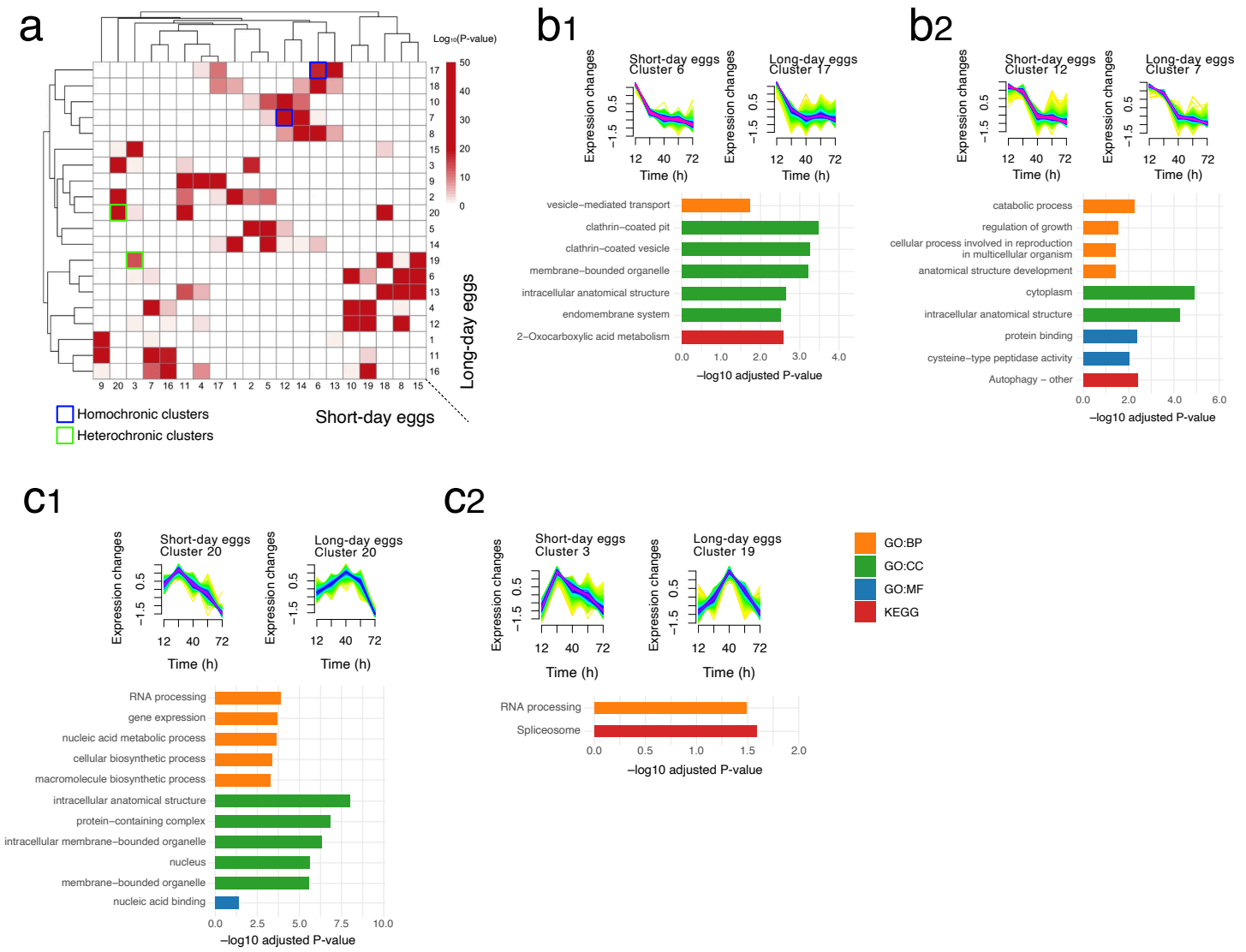
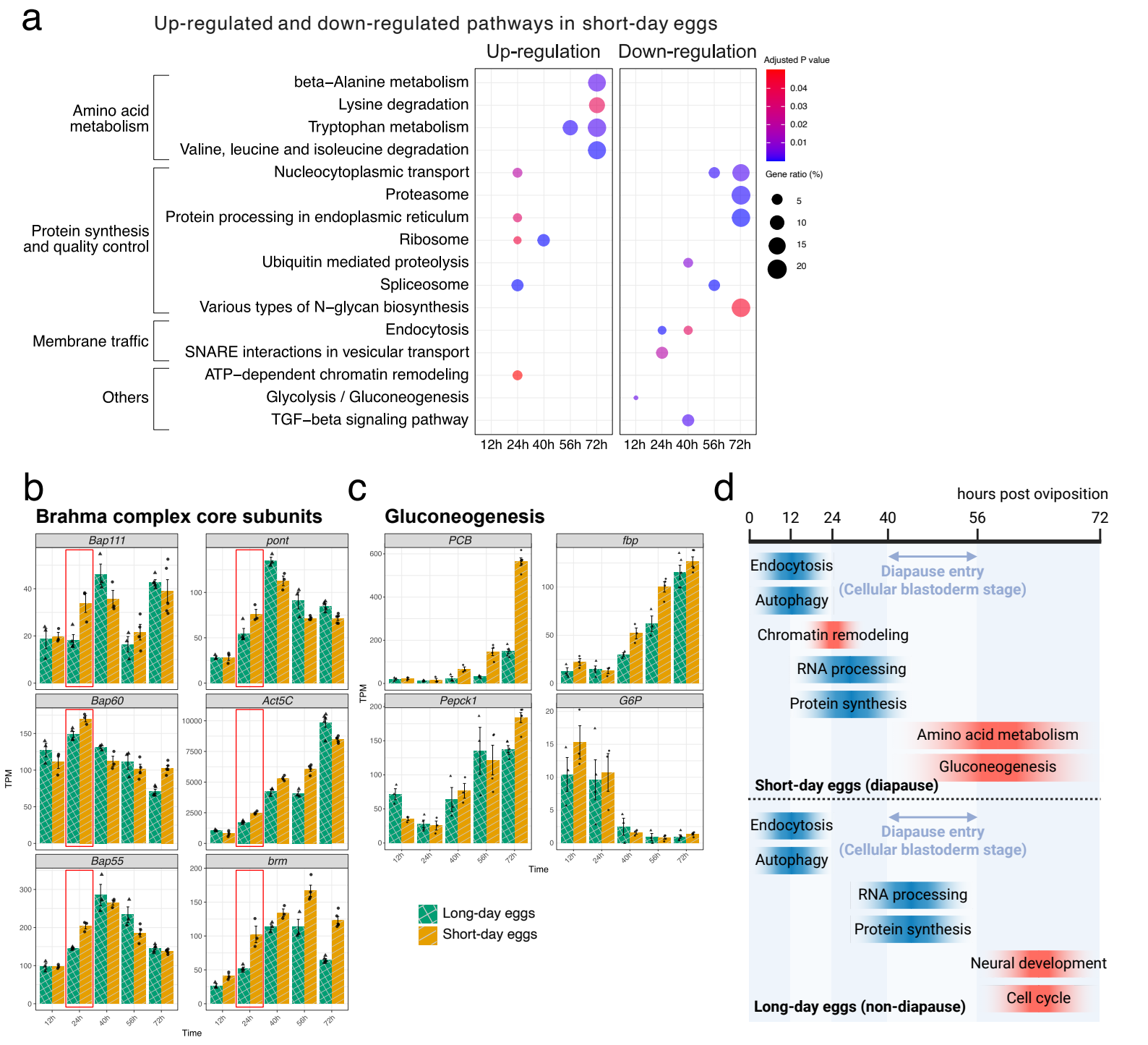

a
Log10(P-value)
50
40
30
20
10
0
17 18 10 7 8 15 3 9 2 20 5 14 19 6 13 4 12 1 11 16
9 20 3 7 16 11 4 17 1 2 5 12 14 6 13 10 19 18 8 15
Long-day eggs
Short-day eggs
Homochronic clusters
Heterochronic clusters
b1
Short-day eggs
Cluster 6
Long-day eggs
Cluster 17
Expression changes
−1.5 0.5
12 40 72
Time (h)
vesicle−mediated transport
clathrin−coated pit
clathrin−coated vesicle
membrane−bounded organelle
intracellular anatomical structure
endomembrane system
2−Oxocarboxylic acid metabolism
0.0 1.0 2.0 3.0 4.0
−log10 adjusted P-value
b2
Short-day eggs
Cluster 12
Long-day eggs
Cluster 7
catabolic process
regulation of growth
cellular process involved in reproduction in multicellular organism
anatomical structure development
cytoplasm
intracellular anatomical structure
protein binding
cysteine−type peptidase activity
Autophagy − other
0.0 2.0 4.0 6.0
−log10 adjusted P-value
C1
Short-day eggs
Cluster 20
Long-day eggs
Cluster 20
RNA processing
gene expression
nucleic acid metabolic process
cellular biosynthetic process
macromolecule biosynthetic process
intracellular anatomical structure
protein−containing complex
intracellular membrane−bounded organelle
nucleus
membrane−bounded organelle
nucleic acid binding
0.0 2.5 5.0 7.5 10.0
−log10 adjusted P-value
C2
Short-day eggs
Cluster 3
Long-day eggs
Cluster 19
RNA processing
Spliceosome
0.0 0.5 1.0 1.5 2.0
−log10 adjusted P-value
GO:BP
GO:CC
GO:MF
KEGG

Figure 4


a
Up-regulated and down-regulated pathways in short-day eggs
Up-regulation
Down-regulation
Adjusted P value
0.04
0.03
0.02
0.01
Gene ratio (%)
5
10
15
20
Amino acid metabolism
beta–Alanine metabolism
Lysine degradation
Tryptophan metabolism
Valine, leucine and isoleucine degradation
Protein synthesis and quality control
Nucleocytoplasmic transport
Proteasome
Protein processing in endoplasmic reticulum
Ribosome
Ubiquitin mediated proteolysis
Spliceosome
Various types of N–glycan biosynthesis
Membrane traffic
Endocytosis
SNARE interactions in vesicular transport
Others
ATP–dependent chromatin remodeling
Glycolysis / Gluconeogenesis
TGF–beta signaling pathway
12h 24h 40h 56h 72h
12h 24h 40h 56h 72h
b
Brahma complex core subunits
Bap111
pont
Bap60
Act5C
Bap55
brm
TPM
Time
c
Gluconeogenesis
PCB
fbp
Pepck1
G6P
TPM
Time
Long-day eggs
Short-day eggs
d
hours post oviposition
0
12
24
40
56
72
Endocytosis
Autophagy
Diapause entry
(Cellular blastoderm stage)
Chromatin remodeling
RNA processing
Protein synthesis
Amino acid metabolism
Gluconeogenesis
Short-day eggs (diapause)
Endocytosis
Autophagy
Diapause entry
(Cellular blastoderm stage)
RNA processing
Protein synthesis
Neural development
Cell cycle
Long-day eggs (non-diapause)