Long-term coexistence of mtDNA variations and nuclear responses of host

Article *in* Entomologia Generalis · September 2024

DOI: 10.1127/entomologia/2024/2530

8 authors, including:

Tao Wang
Nankai University
21 PUBLICATIONS **92** CITATIONS
SEE PROFILE

Guan-Hong Wang
Chinese Academy of Sciences
36 PUBLICATIONS **630** CITATIONS
SEE PROFILE

Da-Wei Huang
Institute of Zoology, Chinese Academy of Sciences
132 PUBLICATIONS **923** CITATIONS
SEE PROFILE

Jin-Hua Xiao
Nankai University
145 PUBLICATIONS **1,286** CITATIONS
SEE PROFILE

Entomologia Generalis
Published online September 2024

Open Access
Early Access Article

Long-term coexistence of mtDNA variations and nuclear responses of host

Tao Wang[1], Xin-Rui Ren[1], Zhi-Kun Guo[1], Yu-Xuan Zhao[1], Jin Geng[1], Guan-Hong Wang[2], Da-Wei Huang[1,*], Jin-Hua Xiao[1,*]

[1] College of Life Sciences, Nankai University, 300071 Tianjin, China
[2] Institute of Zoology, Chinese Academy of Sciences, 100101 Beijing, China
* Corresponding authors: huangdw@ioz.ac.cn; xiaojh@nankai.edu.cn

With 4 figures

Abstract: Numerous studies have accumulatively discovered mitochondrial genome (mtDNA) diversity in the natural populations of the same species, and some of the mtDNA variations may be selected by the host's environment. However, it remains unclear about the molecular mechanisms by which this long-term coexistence of mtDNA variations in the same species affects the metabolism and evolution of the host. By comparing two mitochondrial genomes of cultured population of *Hermetia illucens*, our study reveals that the mtDNAs of both strains (Ref-strain and Sub-strain) have great structural divergences, and mitochondria of the Sub-strain may be functionally defective, which is consistent with the observed lower body weights and higher oxidative stress levels in the midgut of Sub-strain. Moreover, the differentially expressed genes and differential metabolites between the midguts of both strains were related to the mitochondrial functions including oxidative stress, antioxidant and electron transport chain. Interesting, the midgut microbial compositions differed significantly in both strains. Additionally, 25 of 310 the potentially positively selected genes were related to mitochondrial function. Combination of these multidimensional investigations of both strains helped to reveal how the host insects adapt to mtDNA variations through cyto-nuclear interactions. This study can provide new evidence for understanding the nuclear response to the mitochondria dysfunction in insects, and its role in differentiation of the natural populations and even in the process of speciation.

Keywords: black soldier fly; mitochondrial dysfunction; nuclear responses; metabolism; intestinal flora

1 Introduction

Mitochondria are essential organelles in eukaryotic cells, primarily producing ATP for energy supply through oxidative phosphorylation, and participating in various cellular functions. The function of eukaryotic cells requires the intimate interaction between the mitochondrial genome (mtDNA) and the nuclear genome. However, animal mtDNA is circular and contains only 37 genes (including 13 protein-coding genes), so the functional proteins in mitochondria are co-encoded by both the mitochondrial and nuclear genomes. Additionally, the production of ATP can generate reactive oxygen species (ROS) (Brand 2010), which can act as signaling molecules in regulating intracellular biological processes (Santabárbara-Ruiz et al. 2019) and can even play an important role in immunity of the insect intestine (Bai et al. 2021).

Since all mtDNA-encoded protein genes are the subunits of the electron transport chain, the accumulation of deleterious mutations in mtDNA will affect mitochondrial oxidative phosphorylation function (Wallace 2010). However, nuclear genome can evolve to compensate for the effects of deleterious mutations in mtDNA, so the nuclear genome and mtDNA can co-evolve (Gershoni et al. 2014). Recently, some studies constructed conplastic strains by introducing different mtDNA types into hosts of the same nuclear genetic background to explore the impacts of mtDNA variations on hosts' gene expression and metabolism (Latorre-Pellicer et al. 2016; Healy & Burton 2020), which are essentially studies of host metabolic responses under nuclear-mitochondrial incompatibility. However, it still remains unclear how long-term coexisting mitochondrial variations in the same species, under the influence of natural selection, affect metabolism and evolution of the host. In fact, many studies have recently identified mtDNA diversity in the natural populations of the same species (Li et al. 2022). Among these mtDNA variations, some functional mutations can facilitate hosts' adaptations to new environments (Wallace 2016). For example in

DOI: 10.1127/entomologia/2024/2530
E. Schweizerbart'sche Verlagsbuchhandlung, 70176 Stuttgart, Germany, www.schweizerbart.de

natural populations of *Drosophila melanogaster*, the mtDNA variations may reflect adaptations to climatic environments (Lajbner et al. 2018), and in small brown planthoppers, *Laodelphax striatellus*, the emergence of two mitochondrial haplotypes is correlated with the different temperatures (Sun et al. 2019).

The black soldier fly (BSF), *Hermetia illucens* L. (Diptera: Stratiomyidae), is a globally distributed and economically important insect (Kaya et al. 2021). The bioconversion of organic waste by BSF larvae (BSFL) is considered one of the most promising technologies, and the economic value of BSFL is largely depend on their digestion and absorption of food in which the intestinal microbiota plays an important role, and related applied research has rapidly expanded in recent years (IJdema et al. 2022). Moreover, recent phylogenetic studies have revealed that global mtDNAs of BSF can be divided into two haplogroups, which diverged over two million years ago (Guilliet et al. 2022). Hence, the BSF is a good model for studying the effects of mtDNA variations on metabolism and evolution of host.

Here, from the same cultured population, we isolated two strains of BSF with the mtDNAs belonging to each of the two phylogenetic haplogroups. First, we comparably studied the evolution of both mitochondrial genomes. Then, we compared the physiological characterizations, the metabolism and microbiota in the midguts of both strains of BSFL. Finally, based on analyses of the gene expression and molecular evolution between both strains, we explored how the host adapted to mtDNA variations in nuclear genome. Understanding these mechanisms can provide new evidence for the molecular mechanism of host adaptation, which in turn will help understand the role of mtDNA variations in the differentiation and even speciation of natural populations of insects.

2 Material and methods

2.1 Breeding of black soldier flies

The BSF population was initially collected from the wild in Guangxi, China, and then reared on a large scale (millions in population) at Jiangsu Difei Biotechnology Co., Ltd. Taken from this population, about 50 g eggs were randomly selected and reared (fed with wheat diet; 27 ± 1 °C; relative humidity of 60 ± 1%; 12:12 L:D photoperiod) in the laboratory for four generations. Two strains, Ref-strain and Sub-strain, were then isolated using two mitochondrial molecular markers, *cox1* and *cob* (primers see Table S1), and the founder flies were setted as F0 generation. Both strains (each containing five pairs of founder flies) were reared separately for nine generations. A schematic of the two BSFL strain isolated processes is shown in Fig. S1A.

When the eggs of the 10th generation (F10 generation) hatched and reared on the wheat diet for five days (more than 2000 larvae), 60 larvae were randomly picked and placed in a plastic box with 100 g wheat diet and 200 mL distilled water (for each strain, 3 replicates from 3 different larval boxes were set). The same food (20 g of wheat diet and 20 ml of distilled water) was added to each box at 17:00 every day for the next nine days. On the 15th day, the washed and surface-sterilized larvae were dissected in sterile phosphate buffered saline (PBS) to isolate the midguts, which were immediately frozen in liquid nitrogen and then stored at −80 °C. A schematic processing of the two BSFL strain processes is shown in Fig. S1B.

The BSF strain with the mtDNA marker sequences of *cox1* and *cob* identical to the published mtDNA sequences of BSF (accession ID: NC_035232.1) was named as Ref-strain (Ref means reference) and the other strain was named as Sub-strain (Sub represents subpar, which implied that the body weights of this strain were subpar compared to the Ref-strain).

2.2 Sequencing and analysis of mitochondrial genomes

For each BSF strain, one male adult was selected to extract total DNA with EasyPure® Genomic DNA Kit (TransGen, Beijing, China), and build sequencing library with NEB Next® Ultra™ DNA Library Prep Kit (NEB, USA). 250 bp paired-end sequencing was carried out on a NovaSeq 6000. After quality control and filtering by FastQC and fastp, clean data was used to assemble the mitochondrial genome using Getorganelle. Assembled mtDNA was annotated using MITOS and manually corrected for protein-coding genes (PCGs) and tRNAs. Meanwhile, the published 57 BSF mtDNAs (Guilliet et al. 2022) were downloaded. Subsequently, 12 PCGs (except *ND5* which was provided partial sequence in the 57 BSF mtDNAs) and both rRNAs were extracted from the mtDNA and later aligned with MAFFT (v. 7.313, data type selected "codon", align strategy selected "auto" and codon table selected "5"). These alignments were concatenated using PhyloSuite and ModelFinder was used to find the best model. Then Bayesian inference and maximum-likelihood trees were constructed. Due to one bp of insertion in *nad1* gene, the protein tertiary structures for ND1 of both BSF strains were predicted using RoseTTAFold (Baek et al. 2021), followed by alignment of the tertiary structures using Pymol 2.5.2.

2.3 Physiological characterization of both BSF strains

Thoracic tissues of adult BSF of both strains were taken and homogenized, and Complex I activity was measured using Mitochondrial Respiratory Chain Complex I Activity Assay Kit (Abbkine, Wuhan, China), according to the manufacturer's instructions (4 replicates for each strain). The concentration of ROS in the intestine was measured as described previously (Wang et al. 2022), and the malondialdehyde (MDA) concentration in the intestine was measured using a Malondialdehyde assay kit (TBA method) (Nanjing Jiancheng

Bioengineering Institute, China), with more detailed information available in the supplementary methods. The weight of each BSFL individual was measured before midgut dissection (a total of 45 individuals were averagely and randomly selected from the 3 boxes of larvae per strain). The larvae were selected from the food, washed with water, dried with paper tissue, and accurately measured using an analytical balance (CP224C; OHAUS, USA), retaining four decimal places. Statistical tests for all the above measurement data were performed in Minitab 19. The data was tested whether conformed to a normal distribution, and the independent samples t-test was used for normally distributed data.

2.4 Measurements and analyses of metabolites in the BSFL midguts

The metabolites of BSFL midguts were extracted with an extraction buffer (6 individuals per strain). We used the LC/MS system for metabolomics analysis. The metabolite identification was performed based on the Progenesis QI software online METLIN database and a self-built library, and at the same time, theoretical fragment identification and mass deviation were within 100 ppm. The identified metabolites were searched in the KEGG, HMDB, and lipidmaps databases for classification and pathway information. Based on the grouping information, the fold-change and p-value (T-test) were calculated for each metabolite. The R language package Ropls was used to perform orthogonal partial least-squares discrimination analysis (OPLS-DA) modeling. The metabolites with VIP > 1, p-value < 0.05 and |fold change| (FC) ≥ 1.5 or FC ≤ 0.67 were identified as differential metabolites.

2.5 Microbiota sequencing and analyses of the BSFL midguts

Total genomic DNA was extracted from the midgut of BSFL per individual using the TGuide S96 Magnetic Soil /Stool DNA Kit (DP812, TIANGEN®) (a total of 21 individuals were randomly selected from 3 boxes of larvae per strain). After quality control, the clean data was merged and denoised in QIIME2 v2020.6. Then, according to the SILVA release 138 (Quast et al. 2012) classifier, feature abundance was normalized using the relative abundance of each sample. α- and β-diversity was calculated.

2.6 RNA-seq and bioinformatics analyses of BSFL midguts

Total RNA was isolated from the midgut of BSFL individuals according to the instructions of TansZol Up Plus RNA Kit (TransGen, Beijing, China). 12 samples (6 sample replicates per strain) were used to extracted the total RNA. For each sample, the sequencing library was constructed (Berry Genomics, Beijing, China) and 250 bp paired-end sequencing was carried out on a NovaSeq 6000.

The gene expression levels were estimated by calculating the value of transcripts per kilobase of exon model per million mapped reads (TPM). The differentially expressed genes were selected with a |fold change| > 2 and p-value < 0.05. KEGG enrichment was performed in R with clusterProfiler package, and the pathways with p-value < 0.05 were selected as significantly enriched pathways. Before GO enrichment, the eggNOG v5.0 was used to assigned GO terms to each DEGs. GO enrichment was performed in TBtools, and we used corrected p-value < 0.05 to filter out the significantly enriched terms.

2.7 Quantitative real-time PCR (qPCR)

To validate whether the regulation of genes were the same as that detected by RNA-seq, and whether the differences in the relative expression of genes were significant, the genes related to fatty acid β-oxidation and mitochondrial encoded genes were quantified with qPCR. By using total RNA from the same samples of RNA-seq, and after total RNA reverse transcribed into cDNAs, three-step qPCR was performed on a CFX96 system. *Actin-5C* (LOC119646467) was used as a reference gene for normalization (Gao et al. 2019) and the Ref-strain was set as the control group. The fold changes were calculated as 2^(-ΔΔCq). The primers for qPCR are provided in Table S2.

2.8 Whole genome resequencing and bioinformatics analysis of F0 generation

The thoracic tissues of F0 generation adults were collected, and total DNA was extracted using DNeasy Blood & Tissue Kits (Qiagen, Germany). Six biological replicates were set for each strain, with each replicate containing one individual. The paired-end sequencing libraries were sequenced on the DNBSEQ platform (BGI, Shenzhen, China), with a data volume of 30 Gb for each sample. The quality control of raw data was performed by fastp to obtain the clean data, and the clean data was mapped to the BSF reference genome (GCF_905115235.1) with BWA-MEM.

Potential PCR duplicates were marked and SNP calling were performed following the Genome Analysis Toolkit standard process to filter out SNPs, then SNPs of more than three within 10 bp or near Indel within 5 bp were removed to avoid false positives. To obtain high-quality SNPs, additional filtering was performed using vcftools V0.1.16 (--min-meanDP 12 --minDP 12 --max-meanDP 30 --maxDP 3 --max-missing 0.85) and PLINK (--maf 0.05 --geno 0.1 --mind 0.1). After obtaining the high-quality SNPs, we used gcta to analyze PAC and PLINK to generate distance, and then Neighbor Joining (NJ) tree based on the p-distance was constructed in MEGA X. Population structure analysis was performed using Admixture. Nucleotide diversity ($\theta\pi$) and fixation index (Fst) was calculated using vcftools within 100kb windows with a step size of 5 kb across the genome. The region with top 5% Fst & top 5% Pi ratio ($\theta\pi_{Ref\text{-}strain}/\theta\pi_{Sub\text{-}strain}$) was considered to be the putatively selected region of Sub-strain, and all genes within the selected region were considered as the potentially positively selected genes (PPSGs). The eggNOG v5.0 was used to assigned GO terms

to each PPSGs, then GO enrichment of PPSGs was performed in TBtools, and KEGG enrichment was done in R using the clusterProfiler package, with organism= 'his' being set. For GO and KEGG enrichment, corrected p-value < 0.05 was considered as the significantly enriched term or pathway. The mitochondrial targeting sequence (MTS) of PPSGs was predicted by MitoFates.

3 Results

3.1 mtDNA divergences analyses of both BSF strains

In laboratory-fed population, two mtDNAs (Ref-mtDNA and Sub-mtDNA) of BSF were identified with an average nucleotide distance of 3.62%. There were two main differences in the structure of both mitochondrial genomes, with Sub-mtDNA having a 15 bp deletion between *tRNA-Gln* and *tRNA-Met*, and an insertion of one bp at mt.11670. The one bp insertion falls in *nad1* gene, leading to a frame-shift that results in the extension of 29 bp which shaped an overlap with the *cob* gene (Fig. S2A). In the phylogenetic tree with other published BSF mtDNAs, Ref-mtDNA and Sub-mtDNA belonged to each of the two main clades of BSF (Fig. S2B). Further alignment of the tertiary structures of the ND1 protein revealed a longer α-helix at the C-terminal (Fig. S2C), which might affect the assembly of mitochondrial complex I, and thereby affecting primary function of electron transport to produce ATP.

3.2 Differentiations of oxidative stress levels and metabolites between the midguts of both BSFL strains

To further explore whether the mtDNA variations between both strains lead to divergences in oxidative stress levels and mitochondrial functions, the activity of mitochondrial complex I and the oxidative stress levels were measured in the larval midguts of both strains that had been reared separately for ten generations. The activity of mitochondrial complex I was significantly lower in Sub-strain (p-value = 0.0195; Fig. 1A), which was consistent with the results in previous section that the assembly of complex I may be affected in Sub-strain. Additionally, the levels of both ROS concentrations (p-value = 0.011; Fig. 1B) and MDA (p-value < 0.01; Fig. 1C) were significantly higher in the Sub-strain, with MDA reflecting the degree of damage to the intestine by lipid oxidation. Moreover, at the equal feeding periods, the Sub-strain individuals were significantly lighter than those of the Ref-strain (p-value < 0.001; Fig. 1D).

Of the detected 5817 metabolites, there were remarkable differences between both BSFL strains (Fig. 1E). The results of OPLS-DA and permutation tests showed that the model was stable and reliable (R2Y = 0.992, Q2Y = 0.904; Fig. S3A), and not over-fitted (Fig. 1F). We identified 1462 differentially expressed metabolites, with 947 up-regulated and 515 down-regulated in Sub-strain (Fig. 1G). Among the differentially expressed metabolites, the most relatively abundant metabolites were carboxylic acids and derivatives, followed by fatty acyls and prenol lipids (Fig. 1H). KEGG enrichment analyses of the differential metabolites detected that lipid metabolism (sphingolipid, α-linolenic acid, and arachidonic acid metabolism) and retinol metabolism were enriched in the Sub-strain, whereas amino acid metabolism (histidine and β-alanine), pyrimidine metabolism, propanoate metabolism, and nicotinate and nicotinamide metabolism were enriched in the Ref-strain (Fig. 1I–J).

By classifying the differential metabolites, we found that oxidative stress-related (arachidonate, 12(S)-HETE, 13-HODE, and 13-oxo-ODE) and antioxidant-related (Vitamin A, delta-tocotrienol, glutathione peroxide, α-linolenic acid and ergothioneine) metabolites significantly increased in the Sub-strain (Fig. 2); NAD metabolism-related metabolites (NADPH, deamino-NAD+, and nicotinamide ribotide) (Fig. S3B–D), and the respiratory substrates of mitochondrial complex I (NADH) and complex II (succinic acid), together with NAD+, were also significantly reduced in the Sub-strain; while butanoic acid produced by bacteria was notably increased in the Sub-strain (Fig. 2). In addition, a substantial of acylcarnitines (Fig. S3E) and CoQ10 (Fig. 2) also differed significantly between both BSFL strains, further indicating differences in their fatty acid β-oxidation pathway. Interestingly, the metaboltics of short-fatty acid and antibiotic were significantly increased in the Sub-strain (Fig. 2).

3.3 Composition of midgut microbiota differed between both BSFL strains

We compared the bacterial diversity in the midguts between both BSFL strains, and found that the α-diversity indices of the midgut microbiota were significantly higher in the Sub-strain (Fig. 3A–D), and the midgut microbiota of both BSFL strains distinctly separated according to Principal component analyses (PCA) analyses (Fig. 3E–F). Two strains shared a low number of amplicon sequence variants (ASVs), whereas each strain had a large number of unique ASVs, with the Sub-strain having a greater number of unique ASVs (Fig. 3G). Among the four major phyla of bacteria (Proteobacteria, Firmicutes, Bacteroidetes, and Actinobacteria) contained in both strains, compared to the Ref-strain, the Sub-strain displayed an increase in Bacteroidetes and Actinobacteria while a decrease in Proteobacteria and Firmicutes (Fig. 3H). At the genus level, the Ref-strain had higher relative abundance of *Providencia* (10.41% in Ref-strain and 1.78% in Sub-strain), *Bacillus* (1.05% in Ref-strain and 0.33% in Sub-strain) and an unidentified genus of Rhizobiaceae (22.97% in Ref-strain and 2.4% in Sub-strain), while the Sub-strain had a higher relative abundance of *Paenochrobactrum* (9.78% in Ref-strain and 17.65% in Sub-strain), *Alcaligenes* (1.16% in Ref-strain and 12.53% in Sub-strain), *Ochrobactrum* (1.98% in Ref-strain and 9.68% in Sub-strain), *Sphingobacterium* (0.73% in Ref-strain and 5.51% in Sub-strain) and *Dysgonomonas* (1.66% in Ref-strain and 5.22% in Sub-strain) (Fig. 3I).

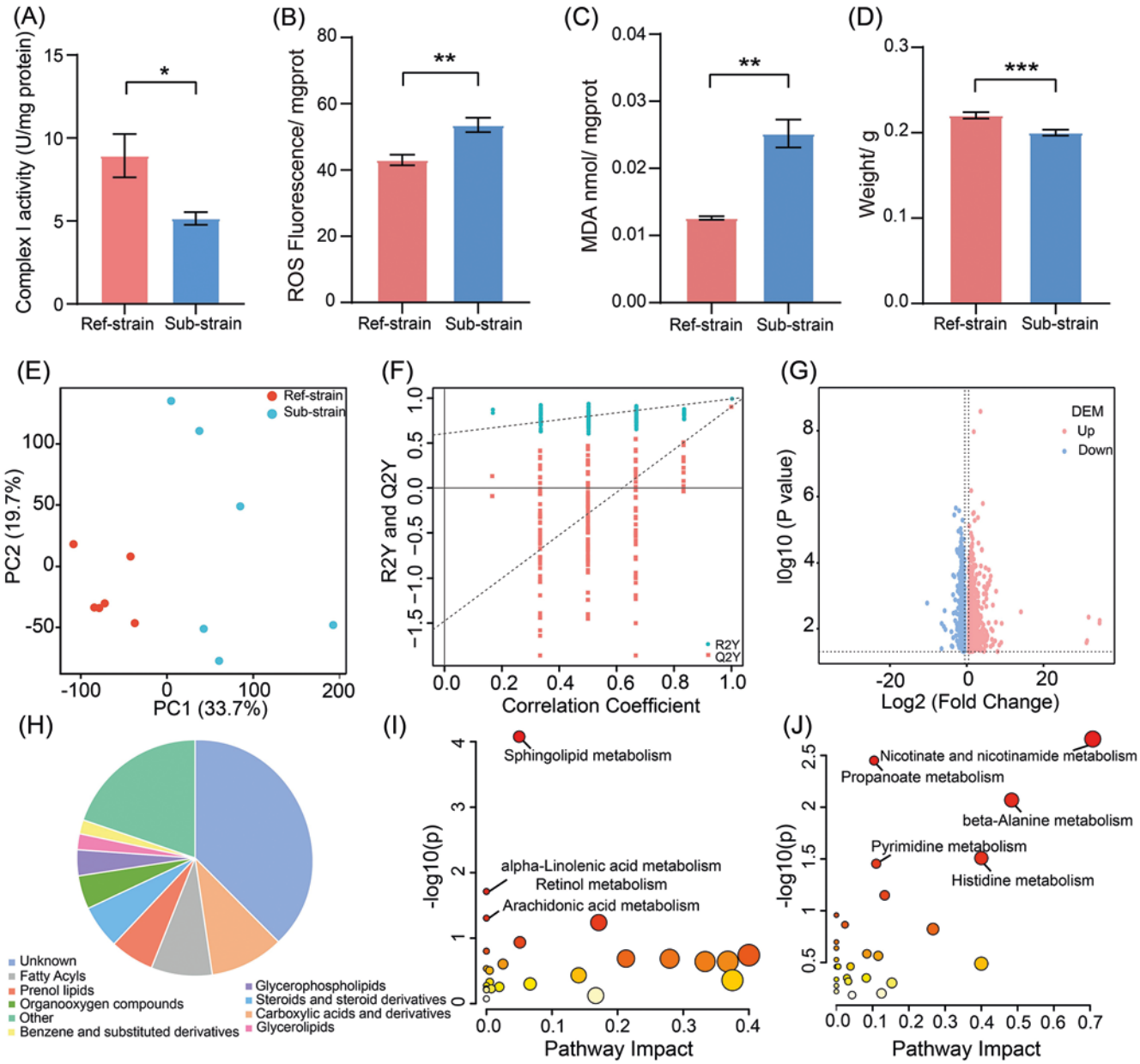

Fig. 1. Physiological characteristics and midgut metabolome of the two BSFL strains. (**A**) Complex I activity (n = 4 individuals/strain); (**B**) ROS concentration of midgut (n = 4 individuals/strain); (**C**) MDA concentration of midgut (n = 4 individuals/strain); (**D**) Weight of larvae (n = 45 individuals/strain); (**E**) PCA analysis (n = 6 sample/strain); (**F**) the permutation test for OPLS-DA; (**G**) volcano plot of differential metabolites, with "up" indicating up-regulated in Sub-strain; (**H**) classification of differential metabolites; (**I**–**J**) the KEGG functional enrichment of differential metabolites (I: Sub-strain enrichment and J: Ref-strain enrichment). *** represents $p < 0.001$, ** represents $p < 0.01$, * represents $p < 0.05$ (independent-sample t test).

3.4 Gene expression patterns indicated altered mitochondrial metabolism in Sub-strain of BSFL

Based on RNA sequencing data of larval midguts, the results of PCA revealed that both strains significantly differentiated along the PC1 (Fig. 4A), indicating that the midguts had significantly different gene expression profiles. Totally, there were 319 genes down-regulated and 545 genes up-regulated in the Sub-strain (Fig. 4B). Interestingly, 11 of the 13 protein-coding genes in mtDNA were significantly down-regulated in the Sub-strain. For KEGG enrichment, the down-regulated genes were significantly enriched in oxidative phosphorylation and mismatch repair (Fig. 4C), and up-regulated genes were significantly enriched in fatty acid degradation, glutathione metabolism and some other pathways involved in metabolism (Fig. 4D–E). For GO enrichment, the down-

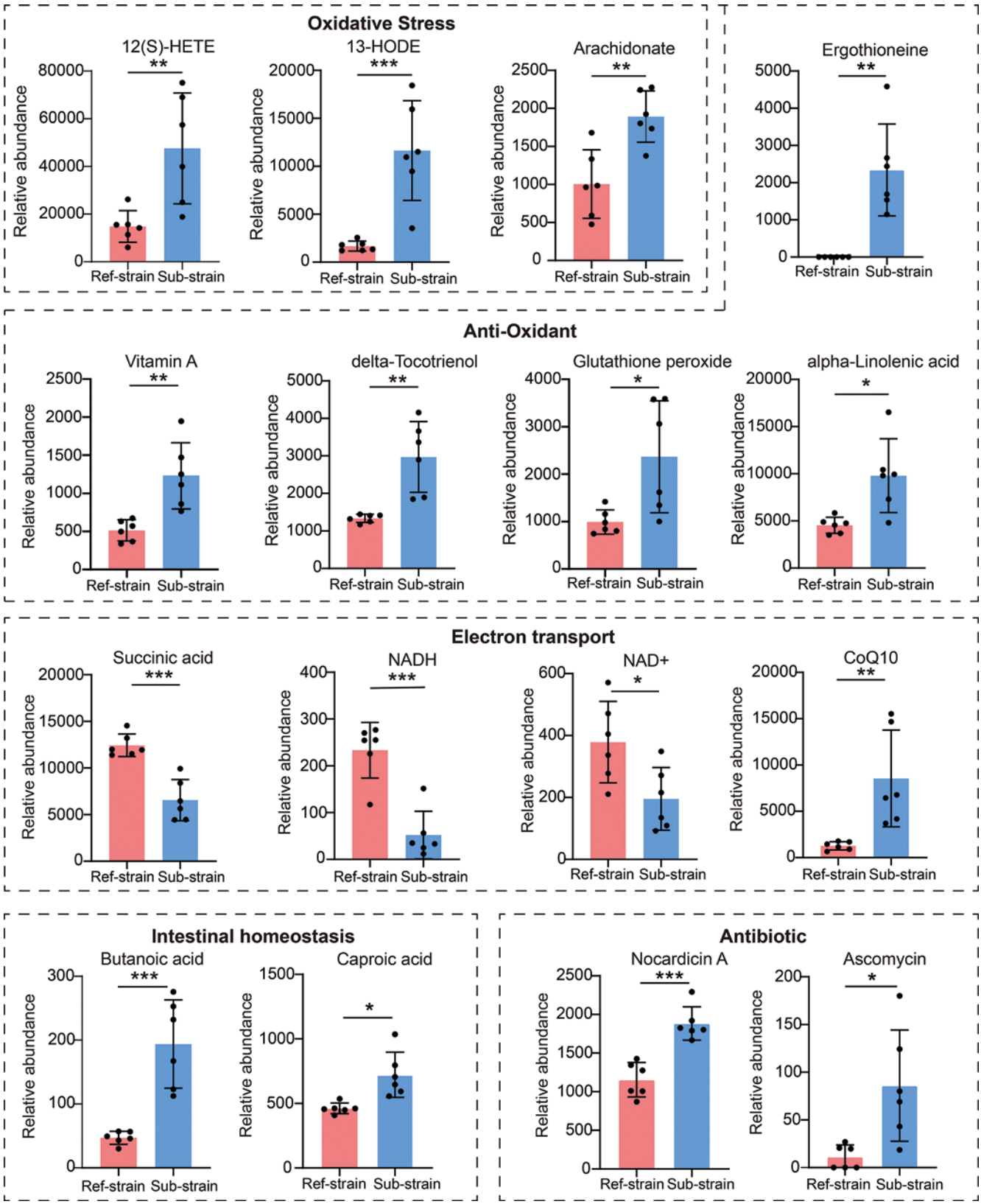

Fig. 2. Histogram of relative abundance of interested differential metabolites. (n = 6 individuals/strain); *** represents $p < 0.001$, ** represents $p < 0.01$, and * represents $p < 0.05$ (independent-sample t test / Mann-Whitney U test).

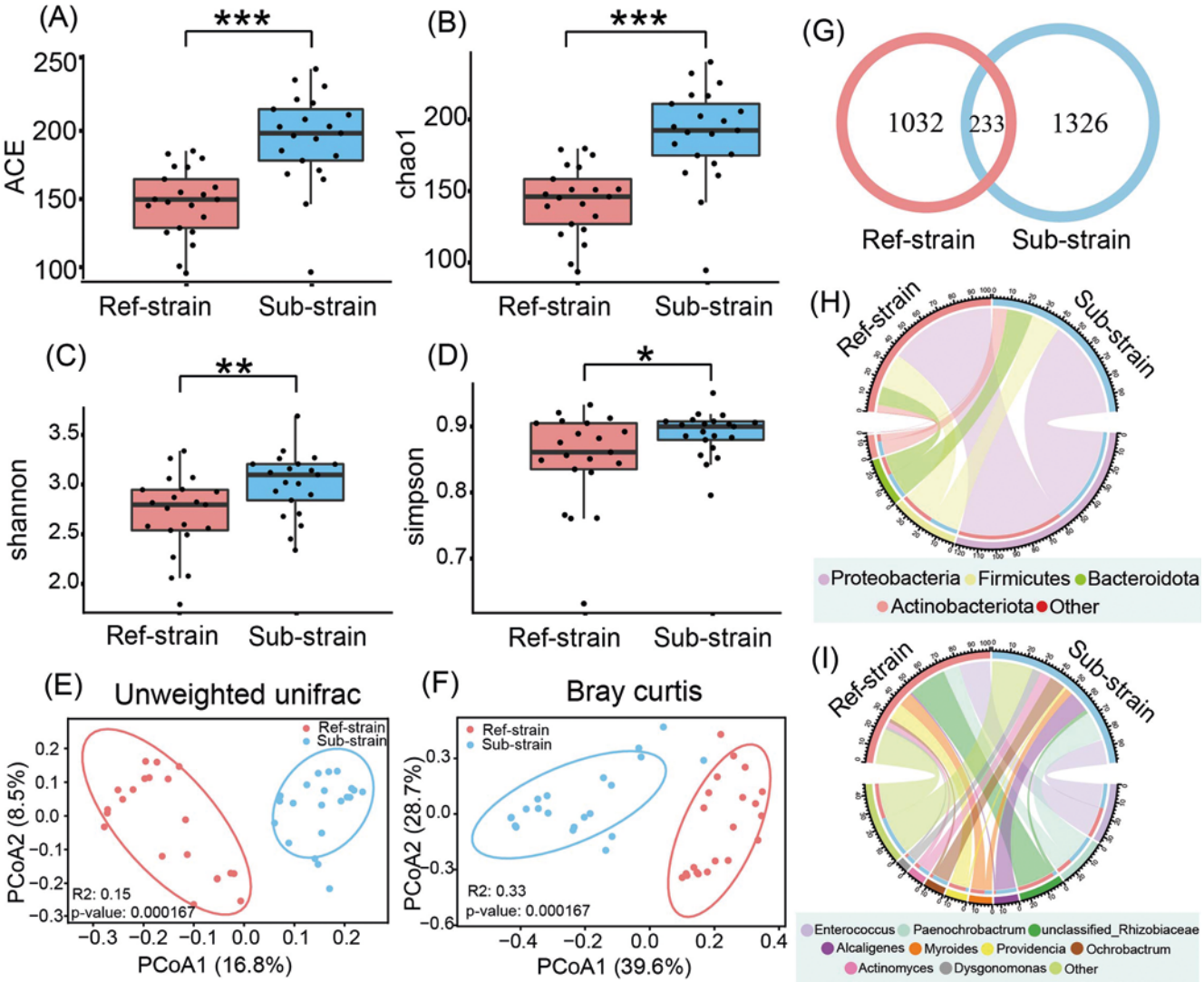

Fig. 3. Comparison of midgut microbiota between the two BSFL strains (n = 21 individuals/strain). (**A**–**D**) Four α-diversity indices (A: ACE index, B: chao1 index, C: Shannon index, D: Simpson index); (**E**–**F**) β-diversity analyses (E: unweighted Unifrace, F: Bray-Curtis distance); (**G**) Venn analysis of ASVs; (**H** & **I**) microbial composition (H: phylum level, I: genus level). *** represents $p < 0.001$, ** represents $p < 0.01$, and * represents $p < 0.05$ (independent-sample t test).

regulated genes were significantly enriched in ribosome biogenesis, complex and NADH dehydrogenase (quinone) activity (Table S3), and the up-regulated genes were significantly enriched in many pathways, including glutathione transferase activity, fatty acid and glutathione metabolic process (Table S4).

Additionally, the expression of mitochondria-encoded genes and key nuclear genes in the pathway of fatty acid β-oxidation were examined by qPCR. The results confirmed that the relative expressions of mitochondria-encoded genes in complex I were significantly down-regulated in the Sub-strain, consistent with results in RNA-seq analyses (Fig. 4F). The nuclear genes of fatty acid β-oxidation pathway were also significantly up-regulated in the Sub-strain, as detected by RNA-seq (Fig. 4F), indicating greater activity of fatty acid β-oxidation in the mitochondria of the Sub-strain, except for *CPT1*, excluded from RNA-Seq DEGs for its low log2fold change (0.889).

3.5 Selection of nuclear genome of the Sub-strain BSF in F0 generation

Considering the functional cooperation of the nuclear and mitochondrial genomes, the mtDNA variation in the Sub-strain BSF may have selective pressure on the evolution of its nuclear genome. To accurately detect the footprint of selection on the nuclear genome of the Sub-strain, we measured genome-wide variations, and obtained 1,008,371 high quality single nucleotide polymorphisms (SNPs) between

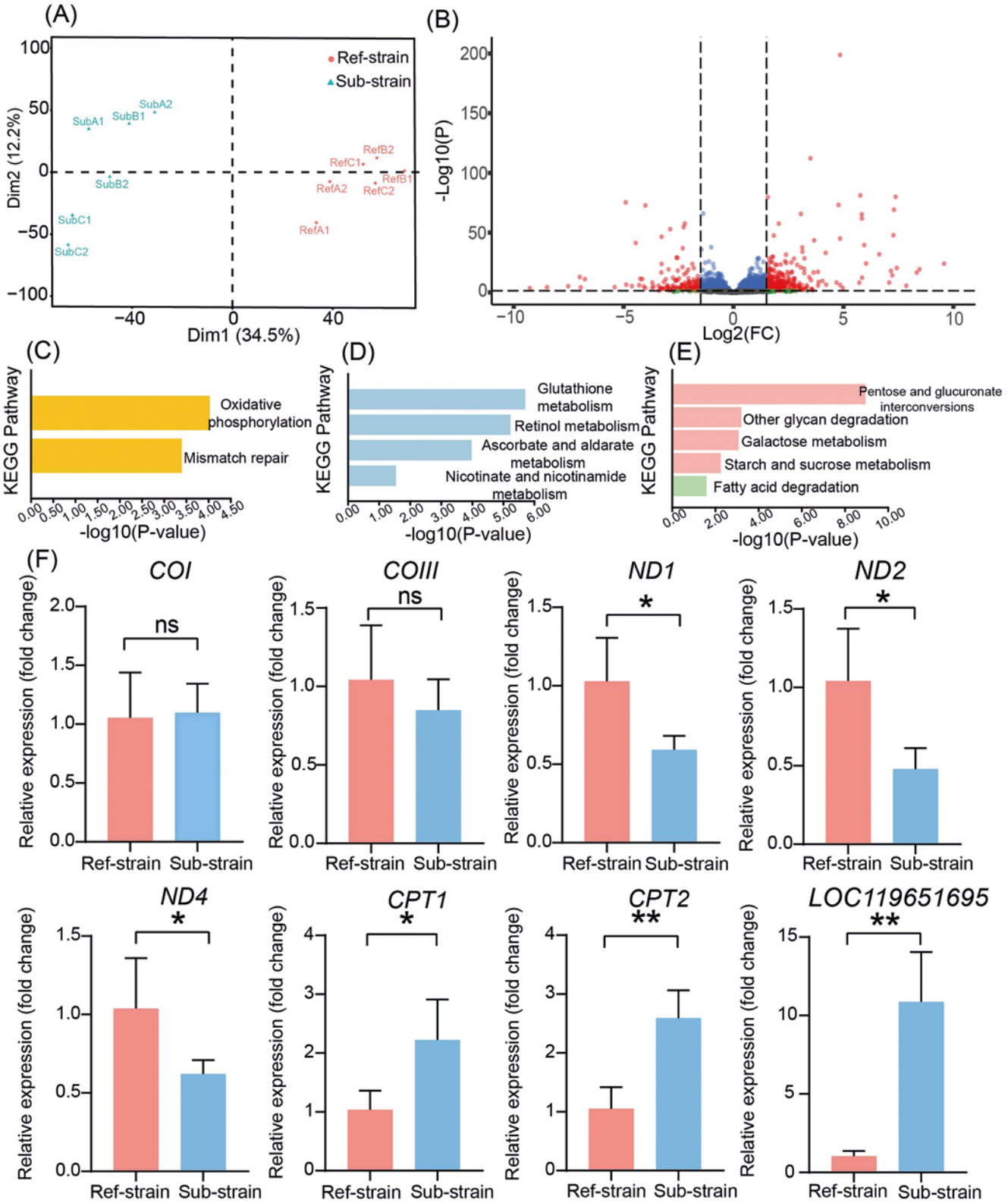

Fig. 4. RNA-seq analysis and qPCR validation of candidate genes between the midgut of two BSFL strains. (**A**) PCA analysis (n = 6 individuals/strain); (**B**) Volcano plot; (**C-E**) Significantly enriched KEGG pathways: (**C**) down-regulated genes; (**D-E**) up-regulated genes; (**F**) qPCR assays of mitochondria-encoded genes and fatty acid β-oxidation-related genes. ** represents $p < 0.01$, * represents $p < 0.05$ (independent-sample t test).

both BSF strains in F0 generation. The neighbor-joining tree based on the SNPs supported separation of both BSF strains (Fig. S4), and PCA results also showed that both BSF strains could be distinguished along PC1 (Fig. S5A). The average Fst value between genomes of both strains was 0.26624, indicating great genetic differentiations between both strains, however results of the population structure analyses showed that both BSF strains still belonged to the same population (Fig. S5B).

Compared with Ref-strain, chromosomes of 2, 4 and 5 were more differentiated on the nuclear genome of Sub-strain (Fig. S5C). By combined analyzing Fst and Pi ratio values, a total of 23.15 MB of putatively selected regions were identified in the genome of Sub-strain (Fig. S5D), including 310 potentially positively selected genes (PPSGs). KEGG analyses of these PPSGs obtained no significantly enriched KEGG pathway (under corrected p-value < 0.05, Table S5). The results of GO enrichment showed that the Molecular Function (MF) mainly included ion transporter activity and glutathione transferase activity (Fig. S5E), the Cellular Component (CC) was only enriched in neuron projection membrane (Fig. S5E), and there was no remarkable enriched item in Biological Process (BP) (Table S6). Interestingly, PPSGs were remarkably enriched in glutathione transferase activity (GO term, MF), highly consistent with GO enrichment results of the up-regulated genes detected by RNA seq.

In order to find out whether the 310 PPSGs were enriched or comprehended nuclear-encoded mitochondrial proteins even mitochondrial functions, we predicted Mitochondrial Targeting Sequences (MTSs) of the 310 PPSGs and found that 22 of them had MTSs (Table S7). Moreover, we screened the function of the 310 PPSGs based on literature and found that 25 of them were related to mitochondrial functions such as fatty acid β-oxidation, mitochondrial fission and fusion and electron transport chains (Table S8), with 14 proteins owning MTS of these 25 mito-function proteins. The above results imply that suggesting a certain extent of mitochondrial and nuclear genome co-evolution driven by the different cyto-nuclear interaction of the two BSF strain.

4 Discussion

In eukaryotes, the interaction between nuclear genome and mtDNA has attached wide interest and has been supported by evidence from multiple perspectives (Luo et al. 2013; Desai et al. 2020). Compared with the laboratory-constructed model of nuclear-mitochondrial incompatibility, the species natively harboring various mtDNA haplotypes can reveal more clearly how mtDNA variations affect metabolism and adaptive evolution of the host, since these mtDNA haplotypes may be ancient adaptive mutations (Ludwig-Słomczyńska & Rehm 2022). Here we focus on two mtDNA haplotypes that have co-existed stably for over two million years in BSF populations, to study the molecular mechanisms underlying the effects of mtDNA variations on the metabolism and evolution of the host.

First, despite the significant differentiations of mtDNA and nuclear genomes, both BSF strains still belong to the same population. However, subsequent sequences analyses predict that mitochondria in the Sub-strain may have functional defects in mitochondrial complex I, consistent with further investigation results of the physiological characterizations, metabolism, microbiota of the larval midguts, and even natural selection on the nuclear genomes. Therefore, a hypothesis about the cyto-nuclear interaction occurs in the Sub-strain of BSF, the putative functional defects in the mitochondrial Complex I can cause increased consumption of alternative energy substances by host cells to meet energy requirements through electron transport pathway, leading to an increase in ROS levels, which in turn can affect host metabolism and evolution of the nuclear genome.

To test this hypothesis, we obtain multidimensional evidence (as illustrated in Fig. S6), which helps reveal the possible mechanism of how ROS levels are elevated within Sub-strain individuals. Since mitochondrial complex I serves as the main entrance for electrons into the electron transport chain (ETC) and is the main rate-limiting enzyme in the electron transport process (Sharma et al. 2009), and owing to its possible functional defects in the Sub-strain, the hosts can alternatively enhance the pathway of fatty acid β-oxidation for energy requirements, by significantly upregulating the expression of related genes such as *CPT1, CPT2*, and *ACS*. The β-oxidation of fatty acids can increase the FADH2/NADH ratio (F/N ratio) (Speijer 2016); the electrons of the increased FADH2 can thus be directly transferred to co-enzyme Q (CoQ) via an electron transfer flavoprotein, thereby bypassing complex I but producing more ROS (Aw et al. 2018). Thus, the Sub-strain BSFL can bypass the potentially constrained mitochondrial complex I by using long-chain fatty acids for fatty acid β-oxidation to meet energy requirements of the cell, which will result in the production of more ROS. It is interesting that due to the consumption of long-chain fatty acids for fatty acid β-oxidation, the fat storage in the Sub-strain larvae will be reduced, which can explain why the BSFL individuals of Sub-strain exhibit lower weight than those of Ref-strain. As a holometabolous insect, majority energy required for the pupal development of BSF comes from the fat stored in the larval stage. This reduce of fat storage of the Sub-strain BSFL may have a negative impact on pupal development and consequently on adult eclosion. However, based on this metabolic flexibility (Aw et al. 2018), we predict that the Sub-strain of BSFL may have a higher fitness in the diet environment with higher fat content, reflecting the differentiation of feeding adaptation between the two strains. Therefore, further research on these mitochondrial metabolic pathways will help to explore how mtDNA variations affect the divergence and adaptation of natural populations in BSF.

In addition, we reveal that while the composition of gut microbiota may be influenced by mtDNA variations through elevated levels of ROS, the gut microbes can also help the host maintain gut homeostasis. We detect that compared to the Ref-strain which has a healthier midgut microbiota with a higher relative abundance of some potential probiotic bacteria of *Bacillus* and *Providenca* (Ao et al. 2021; Klüber et al. 2022), the microbiota of Sub-strain may be disordered due to the higher abundance of inflammatory bacteria (for example *Erysipelothrix* (Eriksson et al. 2010)). This midgut disturbances in Sub-strain larvae may be closely related to its elevated ROS level. However, the bacteria such as *Actinomyces* that produce antibiotics (IJdema et al. 2022) to help eliminate pathogenic bacteria has increased significantly in the Sub-strain; the content of butanoic acid produced by gut microbiota has also significantly increased, which plays an important role in maintaining gut homeostasis (Wang & Liu 2022), suggesting that these gut microbes play an important role in assisting the host to maintain intestinal homeostasis.

A currently prevailing view is that co-existence of various mtDNA haplotypes in a species is in response to different environmental conditions in different localities (Lajbner et al. 2018). However, we screen two mtDNA haplotypes in the same breeding population, which is possibly due to the fact that both mtDNA haplotypes may be reserved genetic diversity for the host to adapt to different environments. In the history of environmental changes, if the environment at a certain stage was more suitable for a certain mtDNA haplotype, the frequency of that haplotype would increase significantly in the population (Lajbner et al. 2018), to improve the population's fitness to the living environment. We find that under the same rearing conditions, the hosts containing different mitochondrial haplotypes exhibit very different metabolic profiles, gene expression patterns and midgut microbiota, suggesting that the interaction of nuclear-mitochondrial genomes can help hosts with divergent mitochondrial haplotypes adapt to different environmental conditions by mainly relying on distinct metabolic pathways and gut microbiota. These differentiated mechanisms of nuclear-mitochondrial interactions will affect the co-evolution of mitochondrial and nuclear genomes, which in turn will affect the differentiation of natural populations.

Acknowledgements: We thank the National Natural Science Foundation of China (Nos. of 31830084 & 32070466) and "the Fundamental Research Funds for the Central Universities", Nankai University (Nos. 96172158, 96173250 & 91822294) for funding. The sequencing datasets were deposited in the NCBI under the Bioprojects PRJNA917807, PRJNA917823 and PRJNA1033065. The mtDNA of Sub-strain (OQ158799) and the sanger sequences for cox1 (OQ160427 and OQ160428), and cob (OQ186632 and OQ186633) were deposited in NCBI.

References

Ao, Y., Yang, C., Wang, S., Hu, Q., Yi, L., Zhang, J., … Yu, C. (2021). Characteristics and nutrient function of intestinal bacterial communities in black soldier fly (*Hermetia illucens* L.) larvae in livestock manure conversion. *Microbial Biotechnology, 14*(3), 886–896. https://doi.org/10.1111/1751-7915.13595

Aw, W. C., Towarnicki, S. G., Melvin, R. G., Youngson, N. A., Garvin, M. R., Hu, Y., … Ballard, J. W. O. (2018). Genotype to phenotype: Diet-by-mitochondrial DNA haplotype interactions drive metabolic flexibility and organismal fitness. *PLoS Genetics, 14*(11), e1007735. https://doi.org/10.1371/journal.pgen.1007735

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., … Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science, 373*(6557), 871–876. https://doi.org/10.1126/science.abj8754

Bai, S., Yao, Z., Raza, M. F., Cai, Z., & Zhang, H. (2021). Regulatory mechanisms of microbial homeostasis in insect gut. *Insect Science, 28*(2), 286–301. https://doi.org/10.1111/1744-7917.12868

Brand, M. D. (2010). The sites and topology of mitochondrial superoxide production. *Experimental Gerontology, 45*(7–8), 466–472. https://doi.org/10.1016/j.exger.2010.01.003

Desai, R., East, D. A., Hardy, L., Faccenda, D., Rigon, M., Crosby, J., … Campanella, M. (2020). Mitochondria form contact sites with the nucleus to couple prosurvival retrograde response. *Science Advances, 6*(51), eabc9955. https://doi.org/10.1126/sciadv.abc9955

Eriksson, H., Brännström, S., Skarin, H., & Chirico, J. (2010). Characterization of *Erysipelothrix rhusiopathiae* isolates from laying hens and poultry red mites (*Dermanyssus gallinae*) from an outbreak of erysipelas. *Avian Pathology, 39*(6), 505–509. https://doi.org/10.1080/03079457.2010.518313

Gao, Z., Deng, W., & Zhu, F. (2019). Reference gene selection for quantitative gene expression analysis in black soldier fly (*Hermetia illucens*). *PLoS One, 14*(8), e0221420. https://doi.org/10.1371/journal.pone.0221420

Gershoni, M., Levin, L., Ovadia, O., Toiw, Y., Shani, N., Dadon, S., … Mishmar, D. (2014). Disrupting mitochondrial–nuclear coevolution affects OXPHOS complex I integrity and impacts human health. *Genome Biology and Evolution, 6*(10), 2665–2680. https://doi.org/10.1093/gbe/evu208

Guilliet, J., Baudouin, G., Pollet, N., & Filée, J. (2022). What complete mitochondrial genomes tell us about the evolutionary history of the black soldier fly, *Hermetia illucens. BMC Ecology and Evolution, 22*(1), 72. https://doi.org/10.1186/s12862-022-02025-6

Healy, T. M., & Burton, R. S. (2020). Strong selective effects of mitochondrial DNA on the nuclear genome. *Proceedings of the National Academy of Sciences of the United States of America, 117*(12), 6616–6621. https://doi.org/10.1073/pnas.1910141117

IJdema, F., de Smet, J., Crauwels, S., Lievens, B., & van Campenhout, L. (2022). Meta-analysis of larvae of the black soldier fly (*Hermetia illucens*) microbiota based on 16S rRNA gene amplicon sequencing. *FEMS Microbiology Ecology, 98*(9), 1–13. https://doi.org/10.1093/femsec/fiac094

Kaya, C., Generalovic, T. M., Stahls, G., Hauser, M., Samayoa, A. C., Nunes-Silva, C. G., … Sandrock, C. (2021). Global

population genetic structure and demographic trajectories of the black soldier fly, *Hermetia illucens*. *BMC Biology, 19*(1), 94. https://doi.org/10.1186/s12915-021-01029-w

Klüber, P., Müller, S., Schmidt, J., Zorn, H., & Rühl, M. (2022). Isolation of bacterial and fungal microbiota associated with *Hermetia illucens* larvae reveals novel insights into entomopathogenicity. *Microorganisms, 10*(2), 319. https://doi.org/10.3390/microorganisms10020319

Lajbner, Z., Pnini, R., Camus, M. F., Miller, J., & Dowling, D. K. (2018). Experimental evidence that thermal selection shapes mitochondrial genome evolution. *Scientific Reports, 8*(1), 9500. https://doi.org/10.1038/s41598-018-27805-3

Latorre-Pellicer, A., Moreno-Loshuertos, R., Lechuga-Vieco, A. V., Sánchez-Cabo, F., Torroja, C., Acín-Pérez, R., … Enríquez, J. A. (2016). Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. *Nature, 535*(7613), 561–565. https://doi.org/10.1038/nature18618

Li, T. C., Miao, Y. H., Wang, T., Zu, G. H., Huang, D. W., & Xiao, J. H. (2022). Impact of mitotype diversity on metabarcoding biodiversity estimations in Insecta and Arachnida using different sample preparation strategies. *Molecular Ecology Resources, 22*(8), 2967–2980. https://doi.org/10.1111/1755-0998.13678

Ludwig-Słomczyńska, A. H., & Rehm, M. (2022). Mitochondrial genome variations, mitochondrial-nuclear compatibility, and their association with metabolic diseases. *Obesity (Silver Spring, Md.), 30*(6), 1156–1169. Retrieved from https://dor.org/10.1002/oby.23424 https://doi.org/10.1002/oby.23424

Luo, D. P., Xu, H., Liu, Z. L., Guo, J. X., Li, H. Y., Chen, L. T., … Liu, Y. G. (2013). A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nature Genetics, 45*(5), 573–577. Retrieved from https://dor.org/10.1038/ng.2570 https://doi.org/10.1038/ng.2570

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research, 41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Santabárbara-Ruiz, P., Esteban-Collado, J., Pérez, L., Viola, G., Abril, J. F., Milán, M., … Serras, F. (2019). Ask1 and Akt act synergistically to promote ROS-dependent regeneration in *Drosophila*. *PLoS Genetics, 15*(1), e1007926. https://doi.org/10.1371/journal.pgen.1007926

Sharma, L., Lu, J., & Bai, Y. (2009). Mitochondrial respiratory complex I: Structure, function and implication in human diseases. *Current Medicinal Chemistry, 16*(10), 1266–1277. https://doi.org/10.2174/092986709787846578

Speijer, D. (2016). Being right on Q: Shaping eukaryotic evolution. *The Biochemical Journal, 473*(22), 4103–4127. https://doi.org/10.1042/BCJ20160647

Sun, J., Duan, X. Z., Hoffmann, A. A., Liu, Y., Garvin, M. R., Chen, L., … Hong, X. Y. (2019). Mitochondrial variation in small brown planthoppers linked to multiple traits and probably reflecting a complex evolutionary trajectory. *Molecular Ecology, 28*(14), 3306–3323. https://doi.org/10.1111/mec.15148

Wallace, D. C. (2010). Mitochondrial DNA mutations in disease and aging. *Environmental and Molecular Mutagenesis, 51*(5), 440–450. https://doi.org/10.1002/em.20586

Wallace, D. C. (2016). Mitochondrial DNA in evolution and disease. *Nature, 535*(7613), 498–500. https://doi.org/10.1038/nature18902

Wang, L., & Liu, J. (2022). Localized butyrate restores gut homeostasis. *Nature Biomedical Engineering, 7*(1), 3–5. https://doi.org/10.1038/s41551-022-00994-z

Wang, T., Li, T. C., Miao, Y. H., Wu, L. N., Chen, Y. Q., Huang, D. W., & Xiao, J. H. (2022). The gender-specific impact of starvation on mitotypes diversity in adults of *Drosophila melanogaster*. *Open Biology, 12*(9), 220108. https://doi.org/10.1098/rsob.220108

Manuscript received: February 9, 2024
Revisions requested: March 25, 2024
Revised version received: July 7, 2024
Manuscript accepted: July 10, 2024

The pdf version (Adobe JavaScript must be enabled) of this paper includes an electronic supplement:
Figure S1, Table S1–S2