Genetic and bioinformatic analyses reveal transcriptional networks underlying dual genomic coordination of mitochondrial biogenesis

Fan Zhang, Annie Lee, Anna Freitas, Jake Herb, Zongheng Wang, Snigdha Gupta, Zhe Chen, Hong Xu [1]

National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA

[1] Correspondence: Hong.Xu@nih.gov

Abstract

Mitochondrial genome encodes handful genes of respiratory chain complexes, whereas all the remaining mitochondrial proteins are encoded on the nuclear genome. However, the mechanisms coordinating these two genomes to control mitochondrial biogenesis remain largely unknown. To identify transcription circuits involved in these processes, we performed a candidate RNAi screen in developing eyes that had reduced mitochondrial DNA contents. We reasoned that impaired mitochondrial biogenesis would synergistically interact with mtDNA deficiency in disrupting tissue development. Over 638 transcription factors annotated in the fly genome, we identified 77 transcription factors that may be involved in mitochondrial genome maintenance and gene expression. Additional genetic and genomic analyses revealed that a novel transcription factor, CG1603, and its upstream factor YL-1 are essential for mitochondrial biogenesis. We constructed a regulator network among positive hits using the published CHIP-seq data. The network analysis revealed extensive connections, and complex hierarchical organization underlying the transcription regulation of mitochondrial biogenesis.

Introduction

Mitochondria respiration, carried out by the electron transport chain complexes (ETC), converts the energy in chemical fuels to the electrochemical potential across the mitochondrial inner membrane ($\Delta\psi_m$) that drives the synthesis of ATP. Deficient ETC not only impairs energy metabolism, but may also disrupt cellular redox balance and various biosynthetic pathways (Shen et al., 2022; Spinelli and Haigis, 2018), and is associated with various human diseases (Gorman et al., 2016). Mitochondria are under dual genetic control. Their own genome, mitochondrial DNA (mtDNA) encodes 13 core subunits of ETC, alongside 2 rRNAs and 22 tRNAs that are required for the translation of these protein coding genes inside the mitochondrial matrix (Chen et al., 2019)

The majority of more than 1000 mitochondrial proteins including the remaining ETC subunits, factors for mtDNA replication and transcription, and mitochondrial ribosomal proteins are all encoded on the nuclear genome (Hock and Kralli, 2009; Taylor and Turnbull, 2005). The mitochondrial transcription factor A (TFAM) compacts mtDNA into nucleoids and is a key regulator of mtDNA copy number (Alam et al., 2003; Scarpulla, 2008). TFAM, together with other auxiliary factors including mtTFB1 and mtTFB2, promotes the transcription of mtDNA by mitochondrial RNA polymerase (POLRMT) into long polycistronic precursor RNAs, which are further processed into individual RNAs (Chen *et al.*, 2019; Falkenberg et al., 2002). POLRMT can also generate short RNA oligos, owning to its exoribonuclease activity, to prime mtDNA replication by polymerase γ (Liu et al., 2022). Nuclear encoded mitochondrial proteins are synthesized in cytoplasm and imported into mitochondria (Wiedemann and Pfanner, 2017). Hence, mitochondrial biogenesis is influenced by the abundance and activities of mitochondrial translocases as well. The intricate interplay between mitochondrial and nuclear-encoded components demands coordinate activities of these two genomes, to maintain the efficiency and integrity of oxidative phosphorylation system and other critical mitochondrial processes (Hock and Kralli, 2009; Scarpulla, 2008).

Mitochondrial respiration, particularly, the contents and the activity of ETC, is fine tuned to cope with the developmental and tissue-specific metabolic demands (Fernandez-Vizarra et al., 2011). Various transcriptional cascades have emerged as effective and adaptable mechanisms

regulating ETC biogenesis. The nuclear respiration factors, NRF1 and NRF2 activate the expression of many nuclear encoded ETC subunits and genes essential for mtDNA replication and transcription (Scarpulla, 2008). This regulation allows NRFs to indirectly control the expression of mtDNA encoded genes, and hence coordinate the activities of both genomes in ETC biogenesis. The peroxisome proliferator-activated receptors (PPARs), upon activation by diverse lipid ligands, induce the expression of nuclear genes in fatty acids oxidation pathway (Berger and Moller, 2002; Hock and Kralli, 2009). Another family of nuclear receptors, the estrogen-related receptors (ERRs) regulate nuclear genes involved in oxidative phosphorylation, including ETCs and the citric acid cycle (Hock and Kralli, 2009; Scarpulla et al., 2012). Different members of PPARs or ERRs families often show tissue specific expression, therefore control mitochondrial biogenesis in tissue and context specific manners by docking different transcription co-factors (Hock and Kralli, 2009; Scarpulla *et al.*, 2012). Notably, all forementioned transcription factors share a common co-activator, PPARγ coactivator-1α (PGC-1α) that directly stimulates the basal transcriptional machinery (Finck and Kelly, 2007; Hock and Kralli, 2009; Scarpulla *et al.*, 2012). The array of transcription factors that interact with PGC-1α enables the coordinated expression of genes responsible for various aspects of mitochondrial functions and biogenesis. Additionally, PGC-1α and related coactivators are dynamically regulated in responses to various physiological or environmental cues, to adjust metabolic program and energy metabolism accordingly (Hock and Kralli, 2009; Scarpulla *et al.*, 2012). Given the large number of mitochondrial genes and their diverse evolution origins (Kurland and Andersson, 2000; Rath et al., 2021), additional transcription factors involved in these processes likely remain to be identified.

Recently, the ModERN (model organism Encyclopedia of Regulatory Networks) project generated genome-wide binding profiles of a large set of transcription factors in *C. elegans* and *Drosophila melanogaster* (Kudron et al., 2018). The global mapping of transcription factor (TF)-DNA interactions could potentially be applied to identify the transcriptional network governing mitochondrial biogenesis. However, the DNA binding profiles of TFs have their limitations in understanding the true biological functions of TFs in gene regulation (Jiang and Mortazavi, 2018). Gene expression is, in most cases, subject to the combined influences of multiple TFs. Additionally, an individual TF may have either activating or repressive roles based on the local

chromatin environment (Jiang and Mortazavi, 2018). Furthermore, despite the substantial progress in bioinformatics analyses, the interpretation of genome-wide omics data still has its limitation due to a lack of robust statistical algorithms, variations in biological contexts, and intrinsic experimental variations (Angelini and Costa, 2014; Zhou et al., 2016). The integration of the DNA binding profiles with functional genetic and genomic studies is ideal to study gene expression regulations (Jiang and Mortazavi, 2018; Park, 2009).

The *Drosophila* eye is an excellent model for genetic analyses due to its ease of assessment and minimal impact on other physiological processes in the presence of developmental abnormalities. The cell proliferation and differentiation during eye development require robust mitochondrial respiration, and adult eyes are severely disrupted by mutations affecting nuclear ETC subunits or mitochondrial translation apparatus (Liao et al., 2006; Owusu-Ansah et al., 2008). We previously developed a genetic scheme to generate mtDNA deficiency by expressing a mitochondrially targeted restriction enzyme (Chen et al., 2015; Xu et al., 2008). In this study, we performed a RNAi screen, targeting 638 transcription factors annotated in the *Drosophila* genome, in the presence of mtDNA deficiency in developing eyes. We recovered 77 TFs, RNAi against which had synergistic effects with the mtDNA deficiency in causing the small-eye phenotype. We further followed up on CG1603, one of the strongest hits from the initial modifier screen and revealed that it was essential for coordinating nuclear genome and mtDNA for ETC biogenesis. Additional network analyses on the recovered hits using published DNA binding profiles illustrated potential regulatory connections and a complex hierarchy of the transcription regulations on mitochondrial biogenesis. The combination of genetic and bioinformatic analyses also facilitated the identification of YL-1 as an upstream regulator of CG1603.

Results

The design of a genetic modifier screen for genes regulating mtDNA maintenance and expression

Mutations on nuclear encoded ETC genes block cell cycle and disrupt the differentiation and morphogenesis of developing eyes (Mandal et al., 2005; Owusu-Ansah *et al.*, 2008). RNAi against *tfam* or *myc*, which promotes the expression of ETC genes encoded on mtDNA and nuclear genome, respectively, also reduced adult eye size (*Figure 1A*). These observations prompted us to use fly eye as a model to identify transcription factors regulating ETC biogenesis. However, inhibitions of any genes essential for cell viability, proliferation, or differentiation, exemplified by the RNAi targeting a mitotic cyclin, *CycB* (*Figure 1A*), would disrupt eye development. Therefore, assaying the eye morphology alone is not sufficient to enrich candidates regulating ETC biogenesis.

The mitochondrial genome of wild type *Drosophila melanogaster* contains a single XhoI site. The expression of a mitochondrially targeted restriction enzyme, XhoI (MitoXhoI) in *Drosophila* ovary effectively selects for escaper progeny carrying mtDNA mutations that abolish the XhoI site (Xu *et al.*, 2008). In a heteroplasmic background containing both wild type and XhoI resistant genome (*XhoI*$^-$), the expression of MitoXhoI can effectively remove the wildtype genome and hence generate mtDNA deficiency (Chen *et al.*, 2015). As a result, the adult eyes were slightly smaller than the control (*Figure 1A*). Considering that mtDNA encodes core components of ETC, we reasoned that inhibiting a gene related ETC biogenesis would have synergistic effect with the mtDNA deficiency on eye development, and the combination of these two genetic manipulations should lead to a stronger disruption of eye development than either of these conditions individually *(Figure 1A, B)*. On this basis, we devised a scheme of modifier screen in eye for genes involved in ETC biogenesis *(Figure 1B).*

The RNAi modifier screen identifying transcription factors regulating ETC biogenesis

To assess the efficacy of this scheme, we carried out a pilot RNAi screen, covering 124 nuclear-encoded mitochondrial genes and 58 non-mitochondrial genes annotated in various cellular processes (*Figure 1C-E*, and *Supplementary file 1*). In practice, male flies carrying a *UAS-IR* transgene were crossed with *Sco/CyO, mitoXhoI; eyeless-GAL4* heteroplasmic female flies (carrying both wildtype and *XhoI*$^-$ mtDNA). This cross generated two groups of offspring, RNAi

only and RNAi together with MitoXhoI expression (RNAi+MitoXhoI) that were cultured in a same vial, thereby minimizing any potential discrepancy caused by environmental factors. Most RNAi flies survived to adult stage but had reduced eye size. A few RNAi were lethal at pupae stage, due to a lack of head capsule that is derived from the eye antenna disc.

For most genes tested in the pilot screen, eyes of RNAi+MitoXhoI flies were smaller than the corresponding RNAi-only flies. To rule out a simple additive effect between RNAi and MitoXhoI expression, we carried out additional analyses to semi-quantify a potential synergy between RNAi intervention and mtDNA deficiency caused by MitoXhoI expression. The eye size of progeny was arbitrarily scored on a scale from 0 to 5 (*Figure 1C*). The indexes of eye size reduction (Index-R) of RNAi and RNAi+MitoXhoI flies were calculated by normalizing the mean eye size scores of each genotype to the corresponding values of control RNAi or control RNAi+MitoXhoI, respectively, and were subsequently plotted against each other on a linear graph (*Figure 1D*). If a RNAi intervention had synergistic effect with mtDNA deficiency, it would lie below the diagonal line. We also included *ewg*, the fly homolog of *NRF-1*, in the pilot screen to set the threshold for calling out positive hits. Of total 40 genes that are either nuclear encoded ETC subunits or related to mtDNA maintenance and gene expression, 82.5% (33 genes) emerged as enhancers (*Figure 1E*). The proportions of synergistic enhancers were much lower in the group of genes involved in other mitochondrial processes (20.2%) or groups of other essential genes not related to mitochondria (8.6%), indicating the efficacy of this modifier screen in enriching genes related to ETC biogenesis (*Figure 1E*).

To understand transcriptional regulations of ETC biogenesis, we screened 1264 RNAi lines that cover 638 genes annotated as transcriptional regulators in *Drosophila* genome. Total 77 enhancers were identified (*Figure 1F, G* and *Supplementary file 1*), including all known factors involved in ETC biogenesis such as Myc, TFAM (Scarpulla, 2008; Wang et al., 2019). We also recovered 20 suppressors, of which, eyes of RNAi+MitoXhoI flies were larger than the corresponding RNAi-only flies.

CG1603 promotes ETC gene expression and mitochondrial biogenesis.

CG1603 emerged as one of the strongest hits from the primary screen (*Figure 1A, G*). *CG1603* RNAi slightly reduced eye size. However, the combination of *CG1603* RNAi with MitoXhoI

expression in the heteroplasmic background resulted in markedly smaller eyes, indicating a clear synergy between the inhibition of *CG1603* and the mtDNA deficiency. We next asked whether CG1603 was involved in mtDNA maintenance. The *Drosophila* midgut is essentially a monolayer epithelium and composed of intestine stem cells, enteroblasts, enteroendocrine cells and enterocytes (EC). The large, flattened enterocytes allow high-resolution imaging of mitochondria and mitochondrial nucleoids. Additionally, the simple organization and distinct cell types, containing both proliferative and terminally differentiated cells, render the midgut an ideal model to evaluate the impact of mitochondrial disruptions on cell proliferation and differentiation (Zhang et al., 2020). We used a "flip-out" method to activate *CG1603* RNAi in a subset of cells (Prober and Edgar, 2000; Zhang *et al.*, 2020), and imaged TFAM-GFP (Zhang et al., 2016), a marker for mitochondrial nucleoids in midgut clones. Both the total TFAM-GFP level and the number of mtDNA nucleoids (TFAM puncta), were markedly reduced in *CG1603* RNAi clones (*Figure 2A-C*), suggesting that CG1603 is necessary for maintaining the steady-state level of mtDNA. We constructed a *SDHA-mNG* reporter line by inserting the *mNeonGreen (mNG)* cDNA in-frame, downstream of the endogenous locus of *SDHA*, a subunit of ETC Complex II that are entirely encoded by the nuclear genome. SDHA-mNG level was notably reduced in *CG1603* RNAi clones (*Figure 2D, E*), suggesting that CG1603 is also required for the expression of nuclear encoded ETC subunits. Different from TFAM-GFP that marks mitochondrial nucleoids and appears as puncta in mitochondria (Chen et al., 2020), SDHA-mNG uniformly distributed in the mitochondrial matrix (*Figure 2D*). By quantifying the total volume of SDHA-mNG positive voxels in the 3-D rendering, we found that the total mitochondrial volume was also reduced in *CG1603* RNAi clones (*Figure 2F*). Collectively, these results demonstrate that CG1603 promotes the expression of both nuclear and mtDNA encoded ETC genes and boosts mitochondrial biogenesis in general. *CG1603* RNAi produced very few EC clones, consistent with the notion that mitochondrial respiration is necessary for ISCs differentiation (Zhang *et al.*, 2020).

CG1603 regulates cell growth and differentiation.

CG1603 encodes a C2H2 zinc finger protein. It has one C2H2 Zinc finger (C2H2-ZF) at its N-terminus, followed by two MADF (myb/SANT-like domain in Adf-1) domains, and six additional zinc fingers at the C-terminus (*Figure 3A, B*). A PiggyBac transgene,

PBac[SAstopDSRed]LL06826 is inserted between the exons 2 and 3 of *CG1603* locus. This modified Piggybac mutator transgene contains splicing donors and stop codons in all 3 reading frames (Schuldiner et al., 2008), and thereby would disrupt the translation of the full length CG1603 protein. Homozygous *PBac[SAstopDSRed]LL06826* was lethal, arrested at the 2nd instar larval stage and eventually died after 10 days (*Figure 3C)*. Both the steady state level of mtDNA and total mitochondrial mass assessed by the levels of several mitochondrial proteins were reduced in these larvae (*Figures 3D, E*). The lethality of this PiggyBac transgene was mapped to a genomic region spanning the *CG1603* locus (*Figure 3—figure supplement A, B*). Importantly, a *P[CG1603^{gDNA}]* transgene that covers the genomic region of *CG1603* fully rescued its viability (*Figure 3A, F*). These results demonstrate that the lethality of *PBac[SAstopDSRed]LL06826* was caused by the loss of function of CG1603, and we hence named it *CG1603^{PBac}* thereafter. Using FLP/FRT-mediated recombination, we generated homozygous *CG1603^{PBac}* mutant clones in both germline and follicle cells in adult ovaries. Consistent with the results of "flip-out" RNAi experiments in the midgut, both the total TFAM level and the number of mtDNA nucleoids, visualized by an endogenously expressed TFAM-mNG reporter, were significantly reduced in *CG1603^{PBac}* clones (*Figure 4A-D* and *Figure 4—figure supplement A, B*). In most *CG1603^{PBac}* clones, TFAM-mNG puncta were hardly observed, demonstrating an essential role of CG1603 in mtDNA maintenance. Compared to twin clones, *CG1603^{PBac}* follicle cell clones contained significantly fewer cells, and these cells were smaller, indicating that CG1603 promotes both cell growth and cell proliferation (*Figure 4A, E)*. *CG1603^{PBac}* egg chambers were also notably small, even smaller than the adjacent anterior egg chambers that are at earlier developmental stages in the same ovariole (*Figure 4A)*. We assessed $\Delta\psi_m$ using the ratiometric imaging of TMRM and MitoTracker Green (Zhang et al., 2019). $\Delta\psi_m$ was nearly abolished in *CG1603^{PBac}* clones (*Figure 4F, Figure 4—figure supplement C*). All together, these observations demonstrate that CG1603 promotes mitochondrial biogenesis and is essential for ETC biogenesis.

CG1603 is a transcription factor regulating nuclear mitochondrial gene expression.

CG1603 protein exclusive localized to nucleus when expressed in cultured cells (*Figure 5A*). We generated a transgene expressing CG1603-mNG fusion protein by inserting *mNeonGreen* cDNA into the endogenous locus of *CG1603*. CG1603-mNG localized to nuclei in ovaries

(*Figure 5B*) and directly bound to polytene chromosomes in salivary gland (*Figure 5C*). Notably, CG1603-mNG was highly enriched on less condensed chromatin regions that had weak Hoechst staining (*Figure 5C*). We performed RNA sequencing (RNA-seq) in larvae to uncover potential targets of CG1603. Between wild type and *CG1603*^{PBac} larvae, total 7635 genes were differentially expressed, including 86% nuclear-encoded mitochondrial genes (*Figure 5D* and *Supplementary file 2, 3*). Nearly half of nuclear-encoded mitochondrial genes were among 1698 genes that were reduced by more than 2-folds in *CG1603* mutant (*Figure 5E* and *Supplementary file 3*). Gene Ontology (GO) enrichment analyses on these 1698 genes also revealed that all top ten significantly enriched biological processes were related to mitochondria, including “mitochondrial translation”, “mitochondrial gene expression”, “electron transport chain”, “aerobic respiration”, “cellular respiration” and “ATP metabolic process” (*Figure 5F*).

To identify genes that may be directly regulated by CG1603, we analyzed the chromatin immunoprecipitation sequencing (ChIP-seq) data that was deposited in modERN (Kudron *et al.*, 2018). Total 8963 peaks were recovered and distributed over all four chromosomes (*Figure 6A* and *Supplementary file 4*). A subset of peaks showed high intensity evaluated by signalValue (*Figure 6A* and *Supplementary file 4*), which may correspond to these high intensity CG1603-mNG bands on the polytene chromosomes of salivary gland (*Figure 5C*). Most CG1603 binding sites (6799) were found at promoter regions, close to the transcription start site (*Figures 6B, C* and *Supplementary file 4*), which is a key feature of a typical TF. Using the RSAT “peak-motifs” tool (Thomas-Chollier et al., 2012), an 8-bp palindromic sequence, “TATCGATA” emerged as the most prevalent CG1603 binding motif (*Figure 6D* and *Supplementary file 5*). CG1603 bound to the genomic regions of 50% nuclear-encoded mitochondrial genes, and among these genes, 79.5 % were down-regulated in the *CG1603*^{PBac} mutant (*Figure 6E* and *Supplementary file 4),* indicating a great accordance between CHIP data and RNAseq results. Most nuclear-encoded mitochondrial genes that were both bound by CG1603 and down-regulated in *CG1603* mutant were ETC genes or related to ETC biogenesis (*Figure 6F* and *Supplementary file 4)*. These are mitochondrial protein import and membrane insertion machinery, ETC assembly factors, and proteins related to the expression of mtDNA encoded ETC subunits including mtDNA replication and transcription, mitochondrial RNA metabolism and translation. Collectively, CG1603 appears

to be a master regulator of mitochondrial biogenesis and coordinates the expression of both nuclear and mtDNA genes in ETC biogenesis.

Network analyses of CG1603 and other transcription factors

Among 77 TFs identified in the initial modifier screen, 49 TFs have ChIP-seq data available in modREN (Kudron *et al.*, 2018). To gain a compressive understanding on the transcriptional regulation of ETC biogenesis, we performed the network analysis on these 49 TFs using the "*vertex sort"* algorithm (Jothi et al., 2009), and constructed a regulatory network (*Figure 7—figure supplement A* and *Supplementary file 6, 7*). Most TFs were identified as strongly connected component due to their extensive connections and were classified in the core or bottom layer of the hierarchical structure, suggesting complexed co-regulations among these TFs in controlling ETC biogenesis. Two transcription factors, Crg-1 and CG15011 were identified as the top-layer TFs with no upstream regulators in the network (*Figure 7—figure supplement A)*. CG1603 is positioned in the middle layer, linked to 7 TFs above and 6 TFs below by integrating the RNA-seq result with CHIP-seq data (*Figure 7A).* Using the "flip-out" RNAi system in the midgut, we found that among 7 TFs upstream of CG1603 in the network, E(bx), YL-1, trem, STAT92E and Myb were also required for maintaining TFAM levels (*Figure 7—figure supplement B*). To further verify their potential roles in regulating CG1603, we performed RNAi against these genes in midgut clones carrying CG1603-mNG reporter. Only *YL-1* RNAi clones displayed a markedly reduction of CG1603 protein compared with neighboring cells (*Figure 7B, C*). Furthermore, overexpression of CG1603 restored eye size, TFAM-GFP and SDHA-mNG levels cause by *YL-1* RNAi (*Figures 7D-J*). These results indicate that YL-1 is indeed an upstream regulator of CG1603, and through which to regulate ETC biogenesis.

Discussion

The dual genetic control of mitochondria presents a fundamental challenge: how are the nuclear genome and mitochondrial DNA coordinated to ensure the efficiency and integrity of oxidative phosphorylation system and other critical mitochondrial processes? In *Drosophila* ovary, the mitochondrial A-kinase-anchor-protein, MDI promotes the translation of a subset of nuclear mitochondrial proteins by cytosolic ribosomes on the mitochondrial outer membrane (Zhang *et al.*, 2016). MDI's targets are predominantly ETC subunits and proteins essential for mitochondrial genome maintenance and gene expression (Zhang *et al.*, 2019). This mechanism coordinates the nuclear and mitochondrial genomes to augment the ETC biogenesis that takes place in differentiating germ cells (Wang *et al.*, 2019; Wang et al., 2023). Cytosolic and mitochondrial translation are up-regulated in concert to boost ETC biogenesis in budding yeast undergoing a metabolic shift from glycolysis to oxidative phosphorylation (Couvillion et al., 2016), further supporting the synchronized expression of ETC components from dual genetic origins at the translational level. Nevertheless, nuclear-encoded mitochondrial ETC subunits often exhibited a concordant expression pattern at the RNA level (Eisen et al., 1998), and mitochondrial-encoded ETC RNAs consistently exhibited similar trends, albeit with a more gradual increase compared to their nuclear-encoded counterparts accompanying the metabolic shift (Couvillion *et al.*, 2016). These observations suggest a potential coordination at the transcriptional level as well. We uncovered a zinc-finger protein encoded by the *CG1603* locus as a master regulator of mitochondrial biogenesis. CG1603 promoted the expression of more than half of nuclear encoded mitochondrial proteins and the inhibition of CG1603 severely reduced mitochondrial mass and mtDNA contents. CG1603 targets were highly enriched in nuclear-encoded ETC subunits and essential factors required for mitochondrial DNA genome maintenance and gene expression. Thus, CG1603 not only promotes mitochondrial biogenesis in general, but also affords a transcriptional coordination of the nuclear and mitochondrial genomes in ETC biogenies.

The modifier screen in the developing eyes took advantage of the mtDNA deficiency resulted from the expression of MitoXhoI in a heteroplasmic background. Besides 77 enhancers, we also recovered 20 suppressors, of which "RNAi + MitoXhoI" flies had larger eyes than "RNAi-only" (*Figure 1G* and *Supplementary file 1*). Knockdown of these genes alone severely reduced eye

size (*Figure 1G* and *Supplementary file 1*). Noteworthy, 5 of them were lethal due to the lack of head capsule that is developed from the eye antenna disc, but the viability of these RNAi flies was restored by MitoXhoI expression. Given that MitoXhoI expression also disrupts eye development, it is perplexing that the combination of RNAi and MitoXhoI expression, two genetic conditions causing the same phenotype, led to a milder phenotype. Perhaps, mitochondrial DNA deficiency caused by MitoXhoI expression triggers a retrograde signal, which boosts cellular stress responses and thereby mitigates the cell growth defects in these RNAi backgrounds.

The CG1603 belong to a large family of C2H2 Zinc finger (C2H2-ZF) transcription factors that contains 272 genes in *Drosophila* genome (https://flybase.org/reports/FBgg0000732.html). It has one N-terminus C2H2-ZF, followed by two MADFs and a cluster of six C2H2-ZFs at the C-terminus (*Figure 3A*). In addition to the C2H2-ZF cluster, which predominantly mediates sequence-specific DNA binding (Persikov et al., 2015; Wolfe et al., 2000), C2H2-ZF transcription factors often possess additional N-terminal protein-protein interaction domains, such as KRAB, SCAN and BTB/POZ domains in vertebrates, ZAD and BTB/POZ in *Drosophila*, for binding to transcription co-regulators (Fedotova et al., 2017; Perez-Torrado et al., 2006; Sobocinska et al., 2021). These interactions allow them to either activate or repress gene expression. CG1603 binds to the genomic regions of 4687 genes in the *Drosophila* genome. Among these genes, 602 and 562 genes were, respectively, decreased or increased more than two folds in *CG1603* mutant (*Supplementary file 4)*. Thus, CG1603 is likely a dual-function transcription factor, capable of both activating and repressing transcription depends on the chromosomal environment of its targets. Ying Yang 1, a well characterized dual-function transcription factor in mammals, contains both a transcription activation domain and a repression domain, in addition to four C2H2-ZFs at its C-terminus (Verheul et al., 2020). The N-terminus C2H2-ZF and the first MADF domain of CG1603 are negatively charged, thus have a low likelihood of binding to DNA that is also negatively charged (*Figure 3A*). In the predicted 3-D structure of CG1603, the positively charged C-terminal zinc fingers and MADF-2 domain cluster in the center, while the negatively charged N-terminal C2H2-ZF and the MADF-1 extend outward to the opposite directions (*Figure 3B*), resembling the domain arrangement of Ying Yang 1 (Verheul *et al.*, 2020). MADF domains share significant similarity with Myb/SANT domains that may bind to either DNA or proteins (Maheshwari et al., 2008). Some MADF

domains, due to their negative charge, have been proposed to interact with positively charged histone tails, similar to the Myb/SANT domain in a well-known chromatin remodeler ISWI, suggesting a potential role in chromatin remodeling (Maheshwari et al., 2008). Notably, some chromatin remodelers possess tandem Myb/SANT domains that can directly interact with histones and DNA or histone remodeling enzymes like ISWI and SMRT (Boyer et al., 2004). It is plausible that CG1603 may play a role in chromatin remodeling directly or by recruiting nucleosome remodeling factors to its binding sites, thereby modulating gene transcription in those regions.

Notably, CG1603 had no impact on the expression of one third of its binding genes (*Supplementary file 4)*, highlighting that DNA binding profiling along is not sufficient to predict the function of a transcription factor. Nonetheless, the network analyses on TFs' ChIP-seq data allowed us to construct a potential regulatory network among these transcription factors, which subsequently served as a blueprint for genetic analyses to verify potential regulations. Total 7 transcription factors were upstream of CG1603 in the network and emerged as positive hits in the initial screen in the eye. RNAi against 5 of them led to reduced TFAM level in the midgut, while the other two had no noticeable phenotype, suggesting that these two TFs may regulate mitochondrial biogenesis in a tissue specific manner. Only YL-1 was confirmed to act upstream of CG1603 based on the genetic epistasis analysis, further indicating the necessity of combining genomic, bioinformatic, and genetics analyses to gain more reliable and comprehensive understanding on transcriptional regulations. YL-1 is one of DNA binding subunits of the SRCAP complex, which is essential for histone H2A.Z incorporation and replacement (Liang et al., 2016). Recently, it has been shown that both SRCAP complex and H2A.Z are necessary for the transcription of nuclear-encoded mitochondrial genes (Lowden et al., 2021; Xu et al., 2021). Our work offers a mechanistic insight into how CG1603 and its upstream regulator, YL-1 may regulate mitochondrial biogenesis at nucleosome and chromatin level. Currently, most known transcription paradigms controlling mitochondrial biogenesis are centered on transcription factors and co-activators that stabilize or directly stimulate the core transcription machinery. The YL-1 to CG1603 cascade may represent a previously underappreciated layer of transcriptional regulation on ETC biogenesis and could act in concert with NRF1 and other transcription factors to coordinate both the nuclear and mitochondrial genome in ETC biogenesis.

382

Materials and methods

Fly genetics

Flies were maintained on standard cornmeal medium at 25°C, unless otherwise stated. Heteroplasmic lines that contain ~50% XhoI-resistant *mt:CoI*^{T300I} genome (Hill et al., 2014) were maintained at 18 °C. The heteroplasmic *w*^{1118}; *Sco / CyO*, *UAS-mitoXhoI; eyeless-GAL4* females were crossed with different RNAi lines to generate male offspring for assessing adult eye morphology. RNAi lines used in the screen were obtained from the Bloomington *Drosophila* Stock Center (BDSC), or Vienna *Drosophila* Resource Center, and listed in *Supplementary file 1*. *UAS-Luciferase* (BDSC#35788) was used as the transgene control. *TFAM-GFP* reporter line was described previously (Zhang *et al.*, 2016). *Act>CD2>GAL4, UAS-mCD8::mCherry and hsFLP* (BDSC#7) were used to generated "flip-out" clones in midguts. We found that the leakage expression of flippase at 22°C was sufficient to induce "flip-out" clones. *PBac[SAstopDsRed]LL06826* (Kyoto#141919) was obtained from Kyoto Drosophila Stock Center, and backcrossed to *w*^{1118} for six generations before phenotypic analyses. A fluorescent "*CyO, act-GFP*" (BDSC#4533) was used for selecting homozygous mutant larvae. *PBac[SAstopDsRed]LL06826* was recombined with *FRT42D* (BDSC #1802) to generate *FRT42D, CG1603*^{PBac}, which was crossed with *hs-flp; FRT42D, Ubi-nls-RFP* (derived from BDSC#35496) for generating mitotic clones in ovaries (Laws and Drummond-Barbosa, 2015). Briefly, 0–2 days old females were transfer along with sibling males to the Kimwipe-semi-covered vials, then passed to 37°C water bath, heat shock for one hour, twice daily, for three consecutive days. The clones were assessed 7-10 days after the final heat shock. *Def*^{k08815}(BDSC#10818), *Def*^{Exel6052}(BDSC#7534) and *Def*^{Exel6053}(BDSC#7535) were obtained from BDSC.

Cell culture and Gene expression

S2 cells from Drosophila Genomics Resource Center (DGRC) were cultured as previously described (Zhang et al., 2015) following the online instruction (DRSC, https://fgr.hms.harvard.edu/fly-cell-culture). Briefly, cells were maintained in Schneider's medium (Thermo Fisher Scientific) with 10% heat inactivated Fetal Bovine Serum (FBS, Thermo Fisher Scientific) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific) at 27 ºC. Effectene Transfection Reagent (Qiagen) was used for plasmids transfection following manufacturer's instructions. For expression in S2 cells, the coding sequence of *CG1603* was cloned into a pIB vector (Thermo Fisher Scientific), with an *mCherry* coding sequence fused at the 3' end.

Transgenic flies

UASz-CG1603 plasmid was generated by inserting *CG1603* coding sequence between the Acc65I and XbaI sites of pUASz1.0 (https://dgrc.bio.indiana.edu//stock/1431; RRID: DGRC_1431). *UASz-CG1603* was inserted into either attP2 or attP40 (Bestgene Inc.) using PhiC31 integrase-mediated site-specific transformation, to generate transgenes on 3rd and 2nd chromosome, respectively.

The DNA fragment spanning *CG1603* genomic region was amplified by PCR and subcloned into a pattB vector (https://dgrc.bio.indiana.edu//stock/1420; RRID: DGRC_1420). The resulted plasmid was inserted into attP2 site (Bestgene Inc.) using PhiC31 integrase-mediated -specific transformation to generate the transgene *P[CG1603*^{gDNA}*]*.

SDHA-mNeonGreen reporter line was generated using a previous published method (Wang *et al.*, 2019). The targeting cassette comprising of 1 kb genomic DNA fragment upstream of *SDHA* stop codon, *mNeonGreen* coding sequence, a fragment containing *GMR-Hid* flanked by two FRT sites, and 1 kb genomic DNA fragment downstream of *SDHA* stop codon was inserted into a pENTR vector to make the homology donor construct. This donor construct and a *SDHA* chiRNA construct (*gRNA-SDHA* recognizes GTAGACATCCGTACGAGTGA[TGG]) were injected into the *Vasa-Cas9* expressing embryos (BDSC#51323). G0 adults were crossed with w1118 files, and progeny with small eye phenotype were selected as candidates due to the expression of *GMR-Hid*. To remove the *GMR-Hid* cassette, the *SDHA-mNeon-GMR-Hid* flies were crossed with *nos-Gal4; UASp-FLP*. The F1 progeny with the genotype of *nos-Gal4 / SDHA-mNeon-GMR-Hid; UASp-FLP /* + were selected and crossed with *Sco / CyO.* The F2 flies of *SDHA-mNeon / CyO* with normal white eyes were selected and maintained.

For *TFAM-mNeonGreen* and *CG1603-Halo-mNeonGreen* knock-In lines, the targeting cassette comprising of 1 kb genomic DNA fragment upstream of the stop codon, either *mNeonGreen* or *Halo-mNeonGreen* coding sequence, and 1 kb genomic DNA fragment downstream of the stop codon was inserted into pOT2 vector to generate the donor constructs. Each donor construct and the corresponding chiRNA construct (*gRNA for TFAM*: ATGATTTGTGAATTATGTGATGG; *gRNA for CG1603*: GGAATGAACTCTCGCCTTGAGGG) were injected into Vasa-Cas9 expressing embryos (BDSC#51323 or BDSC#51324). G0 adults were crossed with w1118 files, and the progeny carrying the *mNeonGreen* insertions were screened by PCR. Primers for TFAM-mNeonGreen are GCTCGCTGATCAACAAAGTC & GGTGGACTTCAGGTTTAACTCC. Primers for CG1603-mNeonGreen are AGTGCGAGTTCCTCAGT-GTG & CGCCCAGGACTTCCACATAA.

RNA sequencing analysis

For bulk RNA sequencing analysis, total RNA was extracted from wild type and CG1603 mutant larvae (48h after egg laying) by Trizol (Thermo Fisher Scientific) following standard protocol. Three samples were used for each genotype. Poly (A) capture libraries were generated at the DNA Sequencing and Genomics Core, NHLBI, NIH. RNA sequencing was performed with using an Novaseq 6000 (Illumina) and 100-bp pair-end reads were generated at the DNA Sequencing and Genomics Core, NHLBI, NIH. Sequencing data were analyzed following the Bioinformatics Pipeline of mRNA Analysis, NCI, NIH. After quality assessment of FASTQ files using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc), paired-end reads were aligned against Drosophila Melanogaster reference genome (Dmel6) using a two-pass method with STAR (v2.7.9a) (Dobin et al., 2013). Gene level read counts were produced by HTseq (v0.11.4) (Putri et al., 2022). Differential expression analysis at the gene-level was carried out using DESeq2 open-source R package (Love et al., 2014) with an FDR cut-off of 0.05. Gene Ontology (GO) enrichment analysis was performed by clusterProfiler R package (Yu et al., 2012) with the log2 fold change cut off >1 and < -1 for upregulated and down regulated genes, respectively. Density plot was generated by ggplot2 R package (https://ggplot2.tidyverse.org). Drosophila mitochondrial genes and subgroups were referenced against a modified MitoCarta 3.0 inventory (Rath *et al.*, 2021; Wang *et al.*, 2019).

ChIP-seq computational analysis

ChIP-sequencing reads in FASTQ format and the narrowPeak output files for each candidate TF were downloaded from ENCODE Project open resource (https://www.encodeproject.org) (Kudron *et al.*, 2018). ChIP-sequencing reads were aligned to the Drosophila Melanogaster reference genome (Dmel6) using BWA (v0.7.17) (Li and Durbin, 2009). SAM files were sorted and compressed into BAM format with Samtools (v1.16.1) (Li et al., 2009). Replicates were merged by Picard tools (v2.27.3, https://broadinstitute.github.io/picard) using lenient criteria, and all alignments with a MAPQ value less than 20 were removed. Lags prediction and peak-calling were done with MACS2 (v2.2.7.1) (Zhang et al., 2008) following the ENCODE TF ChIP pipeline with IDR analysis performed for consistency analysis (https://github.com/mforde84/ENCODE_TF_ChIP_pipeline). Peak annotation and analysis of profile of ChIP peaks binding to TSS regions were performed with ChIPseeker R package (Yu et al., 2015). Transcription network was analyzed and visualized with VertexSort (Jothi *et al.*, 2009) and igraph (https://igraph.org) R packages, respectively, and ChIP peaks of each TF identified in the gene promoter regions (<2kb) were used for analyses. CG1603 binding motif discovery was done by online integrated pipeline 'peak-motifs' of RSAT tools (https://rsat.france-bioinformatique.fr/rsat/peak-motifs_form.cgi) (Thomas-Chollier *et al.*, 2012).

Imaging analyses

Imaging analyses were performed as previously described (Zhang *et al.*, 2020) using the Zeiss Axio Observer equipped with a Perkin Elmer spinning disk confocal system or a Zeiss LSM880 confocal system. Tissues were dissected out and rinsed in room temperature Schneider's medium (Thermo Fisher Scientific) supplied with 10% heat inactivated fetal bovine serum (FBS; Thermo Fisher Scientific), and then used for either direct imaging or further staining and fixation. For live imaging, a Zeiss incubation system was used to maintain proper temperature, humidity. Live tissues were mounted with medium on the coverslip in a custom-made metal frame and then covered with a small piece of Saran wrap before imaging. For tissue fixation, PBS containing 4% PFA were used, followed by three times PBS washing. Hoechst 33342 and DAPI (5 µg/mL, Thermo Fisher Scientific) incubation in PBS for 5 min was used for nuclear staining of live tissues and fixed tissues, respectively. The image processing and quantification were performed by Volocity (Perkin Elmer, for image acquisition), Zen (Zeiss, for image acquisition), Imaris (Oxford Instruments, https://imaris.oxinst.com/, for 3D surface, voxels and intensity statistics) and Fiji / Image J software (NIH, https://fiji.sc/, for image processing and statistics) based on the previous published methods (Liu *et al.*, 2022; Wang *et al.*, 2023). The relative TFAM-GFP or TFAM -mNeonGreen level, or relative SDHA-mNeonGreen level was calculated as the ratio of the mean fluorescence intensity in the RNAi or mutant clone to that of its neighboring control, with background correction. The relative CG1603-mNG level was calculated as the ratio of the mean nuclear fluorescence intensity in the RNAi clone to that of its neighboring control, with background correction. The relative mtDNA nucleoids number was determined by calculating the ratio of the TFAM-GFP or TFAM -mNeonGreen puncta number in the RNAi or mutant clone, standardized by clone volume, to that of its neighboring control. The relative mitochondrial volume was calculated as the ratio of the total SDHA-GFP positive voxels with local contrast in the RNAi clone, standardized by clone volume, to that of its neighboring control.

Mitochondrial membrane potential was detected using a protocol adopted from a previously study (Zhang et al., 2020; Zhang et al., 2019). Briefly, after dissection, adult ovaries were incubated in the Schneider's medium containing TMRM (200nM, Thermo Fisher Scientific) and MitoTracker Green (200nM, Thermo

Fisher Scientific) for 20 min at room temperature, rinsed with PBS for 3 times, and then imaged live within 1 h. TMRM and MitoTracker signal intensities were quantified and ratiometric image was generated using Fiji / Image J software (NIH). Mitochondrial membrane potential was computed as the ratio of the mean intensity of TMRM to MitoTracker fluorescence with background correction.

Western blot

Protein extracts from wt and *CG1603* mutant larvae tissues (48h after egg laying) were prepared using the RIPA buffer (MilliporeSigma) with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific), 5 mM NaF (MilliporeSigma) and 1 mM Na3VO4 (MilliporeSigma). Western blot was performed using a XCell SureLock™ Mini-Cell and XCell II™ Blot Module (Thermo Fisher Scientific). Samples were electrophoresed under a reducing condition on NuPAGE™ 4 to 12% Bis-Tris Mini Protein Gels (Thermo Fisher Scientific). Proteins on the gel were transferred to a Polyvinylidene Difluoride (PVDF) membrane (Thermo Fisher Scientific). The membranes were blocked with 5% BSA or non-fat milk (MilliporeSigma) in TBST (Tris buffered saline with 0.1% Tween-20). After a serial of washes and incubations with primary antibodies, TBST and secondary antibodies, the immunoreactivity was visualized using SuperSignal West Dura Chemiluminescent Substrate (Thermo Fisher Scientific) and Amersham ImageQuant 800 system (Cytiva). The antibodies used were: Mouse anti-Actin antibody (C4, MilliporeSigma), Mouse anti-ATP5A antibody (15H4C4, abcam), Mouse anti-ND30 antibody (17D95, abcam), rabbit anti-TFAM antibody (Liu *et al.*, 2022), rabbit anti-HSP60 antibody (#4870, Cell Signaling), Anti-rabbit IgG, HRP-linked Antibody (#7074, Cell Signaling) and Anti-mouse IgG, HRP-linked Antibody (#7076, Cell Signaling).

Quantitative real-time PCR and measurement of mtDNA levels

Total genomic DNAs were isolated using the DNeasy Blood & Tissue Kit (Qiagen), following the manufacturers' instructions. Real-time PCRs were performed in triplicate using the PowerTrack SYBR Green Master Mix (Thermo Fisher Scientific), MicroAmp™Optical 96-Well Reaction Plate with Barcode (Thermo Fisher Scientific), and QuantStudio™ 3 Real-Time PCR System (Thermo Fisher Scientific). Primers for amplifying mtDNA were AGCTCATCATATATTTACCGTTGGA & AGCTGGAGAATAAGAAA-GTTGAGT. A nuclear gene, *his4* was used for the nuclear DNA reference. Primers for *his4* were TCCAAGGTATCACGAAGCC & AACCTTCAGAACGCCAC. The relative mtDNA levels of fly larvae were measured in three biological replicates for each group using total DNA extracted from twenty larvae.

Prediction of protein domains, isoelectric point (pI), net charge and structure

Protein domains were predicted via SMART (Schultz et al., 2000). Protein domain pI and net charge were predicted using bioinformatic toolbox, Prot pi (https://www.protpi.ch/Calculator/ProteinTool). Protein 3D structure was predicted by AlphaFold (Jumper et al., 2021).

Statistical analysis

Two-tailed Student's t-test was used for statistical analysis. Difference was considered statistically significant when $P<0.05$. Results are represented as mean ± SD of the number of determinations.

Acknowledgements

We thank Bloomington Drosophila Stock Center, Vienna Drosophila Resource Center and Kyoto Drosophila Genomics and Genetics Resources for various fly lines; NHLBI Light Microscope Core and NHLBI DNA Sequencing and Genomics Core for technical assistance. This work is supported by NHLBI Intramural Research Program.

References

Alam, T.I., Kanki, T., Muta, T., Ukaji, K., Abe, Y., Nakayama, H., Takio, K., Hamasaki, N., and Kang, D. (2003). Human mitochondrial DNA is packaged with TFAM. Nucleic Acids Res *31*, 1640-1645. 10.1093/nar/gkg251.

Angelini, C., and Costa, V. (2014). Understanding gene regulatory mechanisms by integrating ChIP-seq and RNA-seq data: statistical solutions to biological problems. Front Cell Dev Biol *2*, 51. 10.3389/fcell.2014.00051.

Berger, J., and Moller, D.E. (2002). The mechanisms of action of PPARs. Annu Rev Med *53*, 409-435. 10.1146/annurev.med.53.082901.104018.

Boyer, L.A., Latek, R.R., and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? Nat Rev Mol Cell Biol *5*, 158-163. 10.1038/nrm1314.

Chen, Z., Qi, Y., French, S., Zhang, G., Covian Garcia, R., Balaban, R., and Xu, H. (2015). Genetic mosaic analysis of a deleterious mitochondrial DNA mutation in Drosophila reveals novel aspects of mitochondrial regulation and function. Mol Biol Cell *26*, 674-684. 10.1091/mbc.E14-11-1513.

Chen, Z., Wang, Z.H., Zhang, G., Bleck, C.K.E., Chung, D.J., Madison, G.P., Lindberg, E., Combs, C., Balaban, R.S., and Xu, H. (2020). Mitochondrial DNA segregation and replication restrict the transmission of detrimental mutation. J Cell Biol *219*. 10.1083/jcb.201905160.

Chen, Z., Zhang, F., and Xu, H. (2019). Human mitochondrial DNA diseases and Drosophila models. J Genet Genomics *46*, 201-212. 10.1016/j.jgg.2019.03.009.

Couvillion, M.T., Soto, I.C., Shipkovenska, G., and Churchman, L.S. (2016). Synchronized mitochondrial and cytosolic translation programs. Nature *533*, 499-503. 10.1038/nature18015.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868. 10.1073/pnas.95.25.14863.

Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.G., and Gustafsson, C.M. (2002). Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nature genetics *31*, 289-294. 10.1038/ng909.

Fedotova, A.A., Bonchuk, A.N., Mogila, V.A., and Georgiev, P.G. (2017). C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors. Acta Naturae *9*, 47-58.

Fernandez-Vizarra, E., Enriquez, J.A., Perez-Martos, A., Montoya, J., and Fernandez-Silva, P. (2011). Tissue-specific differences in mitochondrial activity and biogenesis. Mitochondrion *11*, 207-213. 10.1016/j.mito.2010.09.011.

Finck, B.N., and Kelly, D.P. (2007). Peroxisome proliferator-activated receptor gamma coactivator-1 (PGC-1) regulatory cascade in cardiac physiology and disease. Circulation *115*, 2540-2548. 10.1161/CIRCULATIONAHA.107.670588.

Gorman, G.S., Chinnery, P.F., DiMauro, S., Hirano, M., Koga, Y., McFarland, R., Suomalainen, A., Thorburn, D.R., Zeviani, M., and Turnbull, D.M. (2016). Mitochondrial diseases. Nat Rev Dis Primers *2*, 16080. 10.1038/nrdp.2016.80.

Hill, J.H., Chen, Z., and Xu, H. (2014). Selective propagation of functional mitochondrial DNA during oogenesis restricts the transmission of a deleterious mitochondrial variant. Nature genetics *46*, 389-392. 10.1038/ng.2920.

Hock, M.B., and Kralli, A. (2009). Transcriptional control of mitochondrial biogenesis and function. Annu Rev Physiol *71*, 177-203. 10.1146/annurev.physiol.010908.163119.

Jiang, S., and Mortazavi, A. (2018). Integrating ChIP-seq with other functional genomics data. Brief Funct Genomics *17*, 104-115. 10.1093/bfgp/ely002.

Jothi, R., Balaji, S., Wuster, A., Grochow, J.A., Gsponer, J., Przytycka, T.M., Aravind, L., and Babu, M.M. (2009). Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. Mol Syst Biol *5*, 294. 10.1038/msb.2009.52.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-03819-2.

Kudron, M.M., Victorsen, A., Gevirtzman, L., Hillier, L.W., Fisher, W.W., Vafeados, D., Kirkey, M., Hammonds, A.S., Gersch, J., Ammouri, H., et al. (2018). The ModERN Resource: Genome-Wide Binding Profiles for Hundreds of Drosophila and Caenorhabditis elegans Transcription Factors. Genetics *208*, 937-949. 10.1534/genetics.117.300657.

Kurland, C.G., and Andersson, S.G. (2000). Origin and evolution of the mitochondrial proteome. Microbiol Mol Biol Rev *64*, 786-820. 10.1128/MMBR.64.4.786-820.2000.

Laws, K.M., and Drummond-Barbosa, D. (2015). Genetic Mosaic Analysis of Stem Cell Lineages in the Drosophila Ovary. Methods Mol Biol *1328*, 57-72. 10.1007/978-1-4939-2851-4_4.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760. 10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.

Liang, X., Shan, S., Pan, L., Zhao, J., Ranjan, A., Wang, F., Zhang, Z., Huang, Y., Feng, H., Wei, D., et al. (2016). Structural basis of H2A.Z recognition by SRCAP chromatin-remodeling subunit YL1. Nat Struct Mol Biol *23*, 317-323. 10.1038/nsmb.3190.

Liao, T.S., Call, G.B., Guptan, P., Cespedes, A., Marshall, J., Yackle, K., Owusu-Ansah, E., Mandal, S., Fang, Q.A., Goodstein, G.L., et al. (2006). An efficient genetic screen in Drosophila to identify nuclear-encoded genes with mitochondrial function. Genetics *174*, 525-533. 10.1534/genetics.106.061705.

Liu, Y., Chen, Z., Wang, Z.H., Delaney, K.M., Tang, J., Pirooznia, M., Lee, D.Y., Tunc, I., Li, Y., and Xu, H. (2022). The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in Drosophila melanogaster. Nat Cell Biol *24*, 757-765. 10.1038/s41556-022-00887-y.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.

Lowden, C., Boulet, A., Boehler, N.A., Seecharran, S., Rios Garcia, J., Lowe, N.J., Liu, J., Ong, J.L.K., Wang, W., Ma, L., et al. (2021). Homeostatic control of nuclear-encoded mitochondrial gene expression by the histone variant H2A.Z is essential for neuronal survival. Cell Rep *36*, 109704. 10.1016/j.celrep.2021.109704.

Maheshwari, S., Wang, J., and Barbash, D.A. (2008). Recurrent positive selection of the Drosophila hybrid incompatibility gene Hmr. Mol Biol Evol *25*, 2421-2430. 10.1093/molbev/msn190.

Mandal, S., Guptan, P., Owusu-Ansah, E., and Banerjee, U. (2005). Mitochondrial regulation of cell cycle progression during development as revealed by the tenured mutation in Drosophila. Dev Cell *9*, 843-854. 10.1016/j.devcel.2005.11.006.

Owusu-Ansah, E., Yavari, A., Mandal, S., and Banerjee, U. (2008). Distinct mitochondrial retrograde signals control the G1-S cell cycle checkpoint. Nature genetics *40*, 356-361. 10.1038/ng.2007.50.

Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet *10*, 669-680. 10.1038/nrg2641.

Perez-Torrado, R., Yamada, D., and Defossez, P.A. (2006). Born to bind: the BTB protein-protein interaction domain. Bioessays *28*, 1194-1202. 10.1002/bies.20500.

Persikov, A.V., Wetzel, J.L., Rowland, E.F., Oakes, B.L., Xu, D.J., Singh, M., and Noyes, M.B. (2015). A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. Nucleic Acids Res *43*, 1965-1984. 10.1093/nar/gku1395.

Prober, D.A., and Edgar, B.A. (2000). Ras1 promotes cellular growth in the Drosophila wing. Cell *100*, 435-446. 10.1016/s0092-8674(00)80679-0.

Putri, G.H., Anders, S., Pyl, P.T., Pimanda, J.E., and Zanini, F. (2022). Analysing high-throughput sequencing data in Python with HTSeq 2.0. Bioinformatics. 10.1093/bioinformatics/btac166.

Rath, S., Sharma, R., Gupta, R., Ast, T., Chan, C., Durham, T.J., Goodman, R.P., Grabarek, Z., Haas, M.E., Hung, W.H.W., et al. (2021). MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. Nucleic Acids Res *49*, D1541-D1547. 10.1093/nar/gkaa1011.

Scarpulla, R.C. (2008). Transcriptional paradigms in mammalian mitochondrial biogenesis and function. Physiol Rev *88*, 611-638. 10.1152/physrev.00025.2007.

Scarpulla, R.C., Vega, R.B., and Kelly, D.P. (2012). Transcriptional integration of mitochondrial biogenesis. Trends Endocrinol Metab *23*, 459-466. 10.1016/j.tem.2012.06.006.

Schuldiner, O., Berdnik, D., Levy, J.M., Wu, J.S., Luginbuhl, D., Gontang, A.C., and Luo, L. (2008). piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. Dev Cell *14*, 227-238. 10.1016/j.devcel.2007.11.001.

Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., and Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res *28*, 231-234. 10.1093/nar/28.1.231.

Shen, K., Pender, C.L., Bar-Ziv, R., Zhang, H., Wickham, K., Willey, E., Durieux, J., Ahmad, Q., and Dillin, A. (2022). Mitochondria as Cellular and Organismal Signaling Hubs. Annu Rev Cell Dev Biol *38*, 179-218. 10.1146/annurev-cellbio-120420-015303.

Sobocinska, J., Molenda, S., Machnik, M., and Oleksiewicz, U. (2021). KRAB-ZFP Transcriptional Regulators Acting as Oncogenes and Tumor Suppressors: An Overview. Int J Mol Sci *22*. 10.3390/ijms22042212.

Spinelli, J.B., and Haigis, M.C. (2018). The multifaceted contributions of mitochondria to cellular metabolism. Nat Cell Biol *20*, 745-754. 10.1038/s41556-018-0124-1.
Taylor, R.W., and Turnbull, D.M. (2005). Mitochondrial DNA mutations in human disease. Nat Rev Genet *6*, 389-402. nrg1606 [pii] 10.1038/nrg1606.
Thomas-Chollier, M., Darbo, E., Herrmann, C., Defrance, M., Thieffry, D., and van Helden, J. (2012). A complete workflow for the analysis of full-size ChIP-seq (and similar) data sets using peak-motifs. Nat Protoc *7*, 1551-1568. 10.1038/nprot.2012.088.
Verheul, T.C.J., van Hijfte, L., Perenthaler, E., and Barakat, T.S. (2020). The Why of YY1: Mechanisms of Transcriptional Regulation by Yin Yang 1. Front Cell Dev Biol *8*, 592164. 10.3389/fcell.2020.592164.
Wang, Z.H., Liu, Y., Chaitankar, V., Pirooznia, M., and Xu, H. (2019). Electron transport chain biogenesis activated by a JNK-insulin-Myc relay primes mitochondrial inheritance in Drosophila. Elife *8*. 10.7554/eLife.49309.
Wang, Z.H., Zhao, W., Combs, C.A., Zhang, F., Knutson, J.R., Lilly, M.A., and Xu, H. (2023). Mechanical stimulation from the surrounding tissue activates mitochondrial energy metabolism in Drosophila differentiating germ cells. Dev Cell. 10.1016/j.devcel.2023.08.007.
Wiedemann, N., and Pfanner, N. (2017). Mitochondrial Machineries for Protein Import and Assembly. Annu Rev Biochem *86*, 685-714. 10.1146/annurev-biochem-060815-014352.
Wolfe, S.A., Nekludova, L., and Pabo, C.O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct *29*, 183-212. 10.1146/annurev.biophys.29.1.183.
Xu, H., DeLuca, S.Z., and O'Farrell, P.H. (2008). Manipulating the metazoan mitochondrial genome with targeted restriction enzymes. Science *321*, 575-577. 321/5888/575 [pii] 10.1126/science.1160226.
Xu, M., Yao, J., Shi, Y., Yi, H., Zhao, W., Lin, X., and Yang, Z. (2021). The SRCAP chromatin remodeling complex promotes oxidative metabolism during prenatal heart development. Development *148*. 10.1242/dev.199026.
Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284-287. 10.1089/omi.2011.0118.
Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382-2383. 10.1093/bioinformatics/btv145.
Zhang, F., Pirooznia, M., and Xu, H. (2020). Mitochondria regulate intestinal stem cell proliferation and epithelial homeostasis through FOXO. Mol Biol Cell *31*, 1538-1549. 10.1091/mbc.E19-10-0560.
Zhang, F., Qi, Y., Zhou, K., Zhang, G., Linask, K., and Xu, H. (2015). The cAMP phosphodiesterase Prune localizes to the mitochondrial matrix and promotes mtDNA replication by stabilizing TFAM. EMBO Rep *16*, 520-527. 10.15252/embr.201439636.
Zhang, Y., Chen, Y., Gucek, M., and Xu, H. (2016). The mitochondrial outer membrane protein MDI promotes local protein synthesis and mtDNA replication. EMBO J *35*, 1045-1057. 10.15252/embj.201592994.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137. 10.1186/gb-2008-9-9-r137.
Zhang, Y., Wang, Z.H., Liu, Y., Chen, Y., Sun, N., Gucek, M., Zhang, F., and Xu, H. (2019). PINK1 Inhibits Local Protein Synthesis to Limit Transmission of Deleterious Mitochondrial DNA Mutations. Mol Cell *73*, 1127-1137 e1125. 10.1016/j.molcel.2019.01.013.
Zhou, W., Sherwood, B., and Ji, H. (2016). Computational Prediction of the Global Functional Genomic Landscape: Applications, Methods, and Challenges. Hum Hered *81*, 88-105. 10.1159/000450827.

Figures and Figure Legends

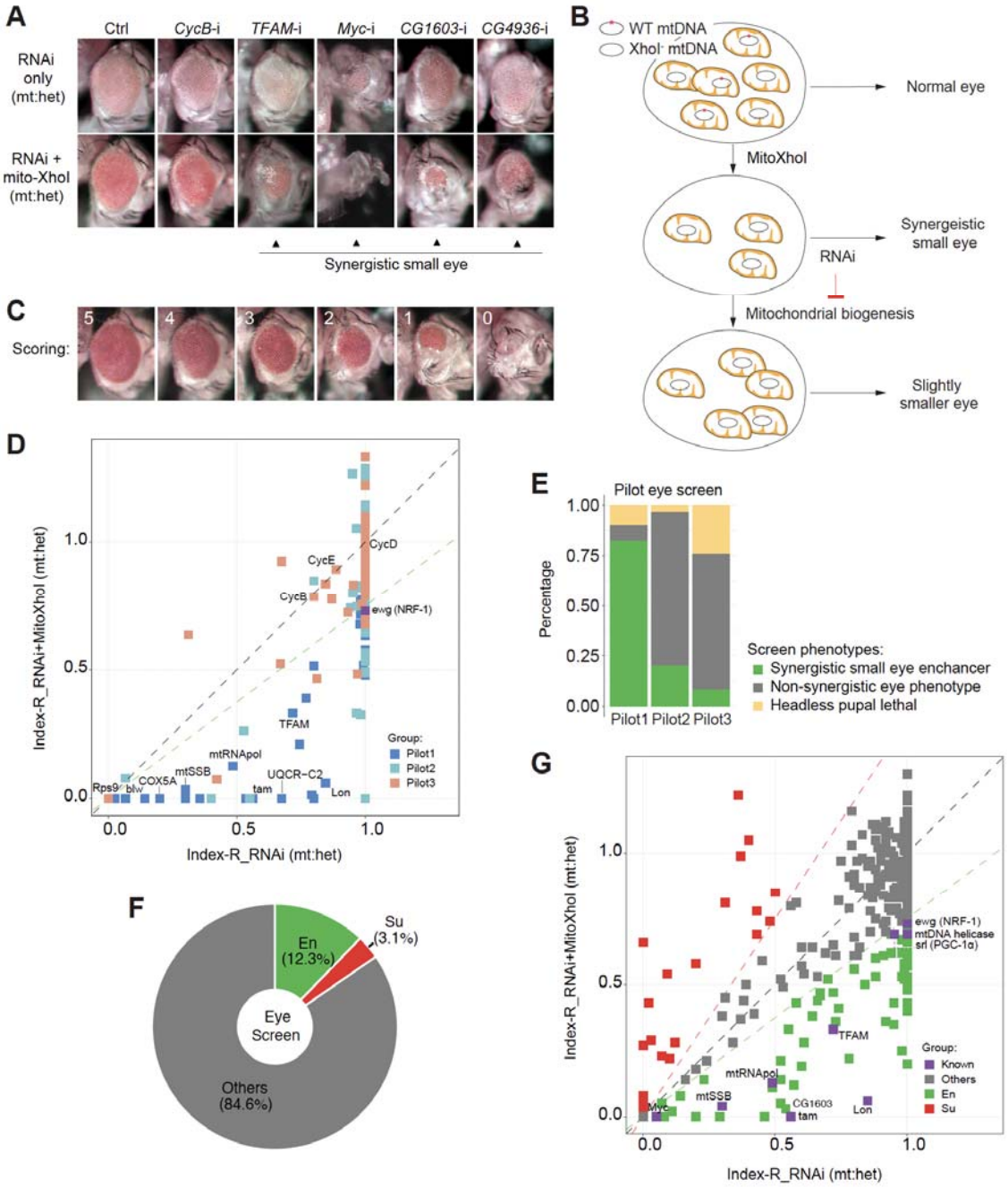


Figure 1. A genetic modifier screen identifying transcription factors regulating ETC biogenesis.

(A) Representative images of adult eye for the control RNAi (Ctrl) and RNAi of selected genes tested in the eye screen, including *CycB* RNAi (*CycB*-i), *TFAM* RNAi (*TFAM*-i), *Myc* RNAi (*Myc*-i), *CG1603* RNAi (*CG1603*-i) and *CG4936* RNAi (*CG4936*-i). The upper panel shows eyes from RNAi only offspring and lower panel displays eyes from RNAi+MitoXhoI offspring cultured at the same condition.

Arrowheads indicate the synergistic small eye phenotype resulting from the combination of gene knockdown and the mtDNA deficiency caused by mitoXhoI in the background of heteroplasmic mtDNAs. (B) Schematic of the genetic modifier screen methodology (see text for details). (C) Representative images illustrating the scoring of eye size. (D) A plot illustrating the calling of positive hits in the pilot screen. Each datapoint represents the Index-R of RNAi (X values) or RNAi+MitoXhoI flies (Y values) for each gene belonging to the different groups (see *Figure 1E* and *Supplementary file 1* for details). Genes with datapoints lie below the grey diagonal dash line exhibited a synergistic effect when combining their RNAi with mtDNA deficiency suggesting a potential role in regulating ETC biogenesis. The datapoint for *ewg*, the fly homolog of *NRF-1*, is labeled in purple. The green dashed line of slope 0.75 outlines the threshold for calling out positive hits based on *ewg*'s performance in the screen. (E) Graph summarizing the pilot screen of nuclear-encoded genes, demonstrating the efficacy of this screen in identifying genes involved in mitochondrial ETC biogenesis. Pilot group 1 (Pilot1) has 40 genes that are either nuclear encoded ETC subunits or related to mtDNA maintenance and gene expression. Pilot2 has 84 genes involved in other mitochondrial processes. Pilot3 has 58 essential genes from other cellular components. (F) Graph summarizing the percentages of synergistic enhancers (En) and suppressors (Su) identified in the screen (see *Figure 1G* and *Supplementary file 1* for details). (G) A plot illustrating the calling of positive hits in the screen of TF genes. Factors that are known to be involved in mitochondrial or ETC biogenesis are marked in purple (Known). The green dashed line outlines the threshold for calling out synergistic enhancers (En, green square). The red dashed line of slope 1.5 outlines the threshold for calling out suppressors (Su, red square).

The following figure supplements are available for figure 1:

Supplementary file 1. List of all genes assessed in the eye screen, including gene IDs, symbols, group information, representative RNAi lines, and the Index-R of the "RNAi only" and "RNAi+ mitoXhoI" flies under the same heteroplasmic-mtDNAs background.

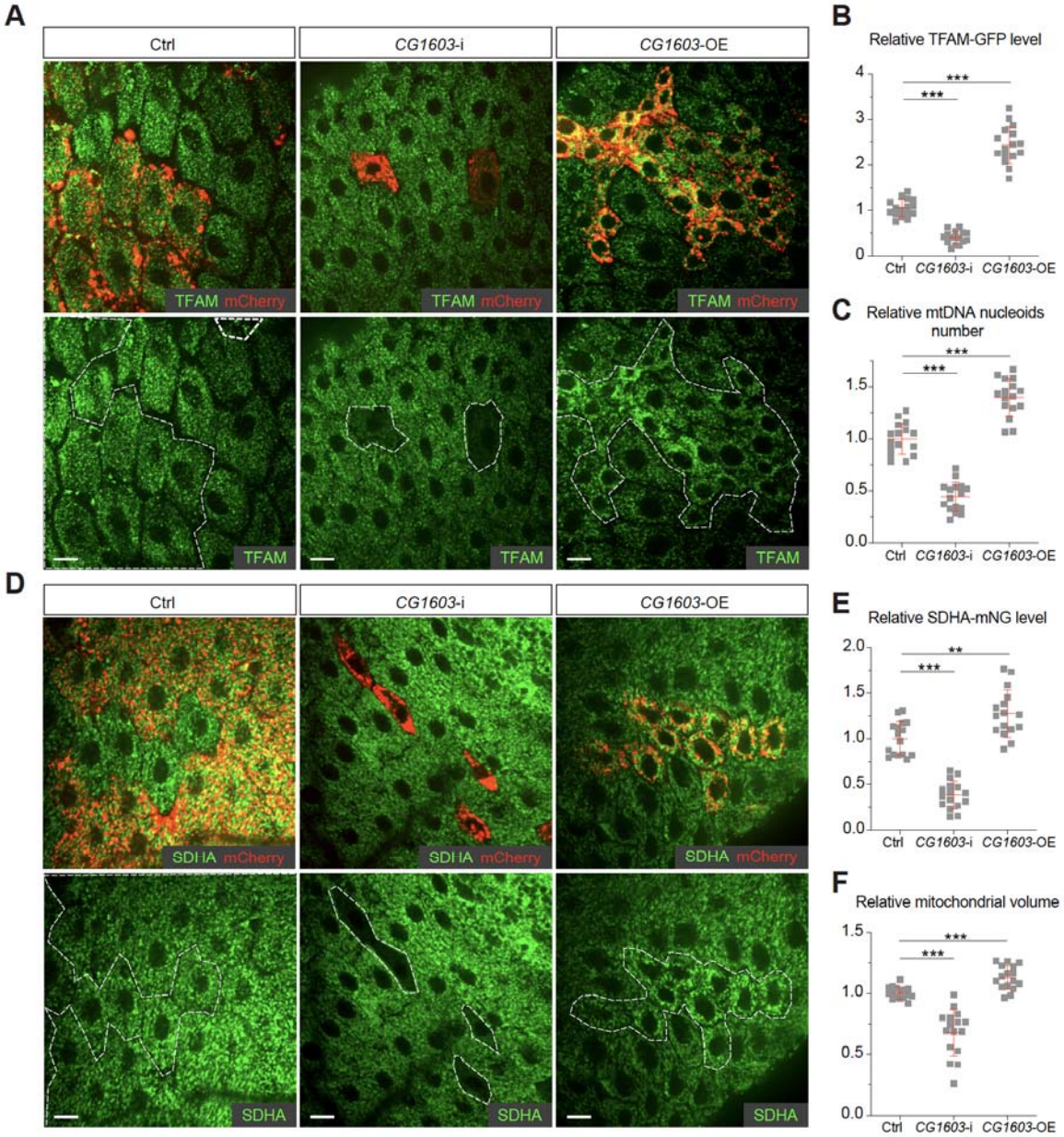


Figure 2. CG1603 promotes ETC gene expression and mitochondrial biogenesis.

(A, D) Representative images of control RNAi (Ctrl), *CG1603* RNAi (*CG1603*-i*) and CG1603* overexpression (*CG1603*-OE*)* midgut EC clones with endogenously expressed TFAM-GFP (A) or SDHA-mNG (D) visualized in green. Clones were labeled by mCherry red and compared with wild-type neighbors. White dash lines aided in illustrating clones. Scale bars: 10 µm. (B, C, E, F) Quantification of the relative TFAM-GFP level (B), the relative mtDNA nucleoids number (C), the relative SDHA-mNG level (E), and the relative mitochondrial volume (F) in the EC clones to their wild-type neighbors. n=16 from 8 midguts for each group, error bar: SD. **: $p<0.01$, ***: $p<0.001$.

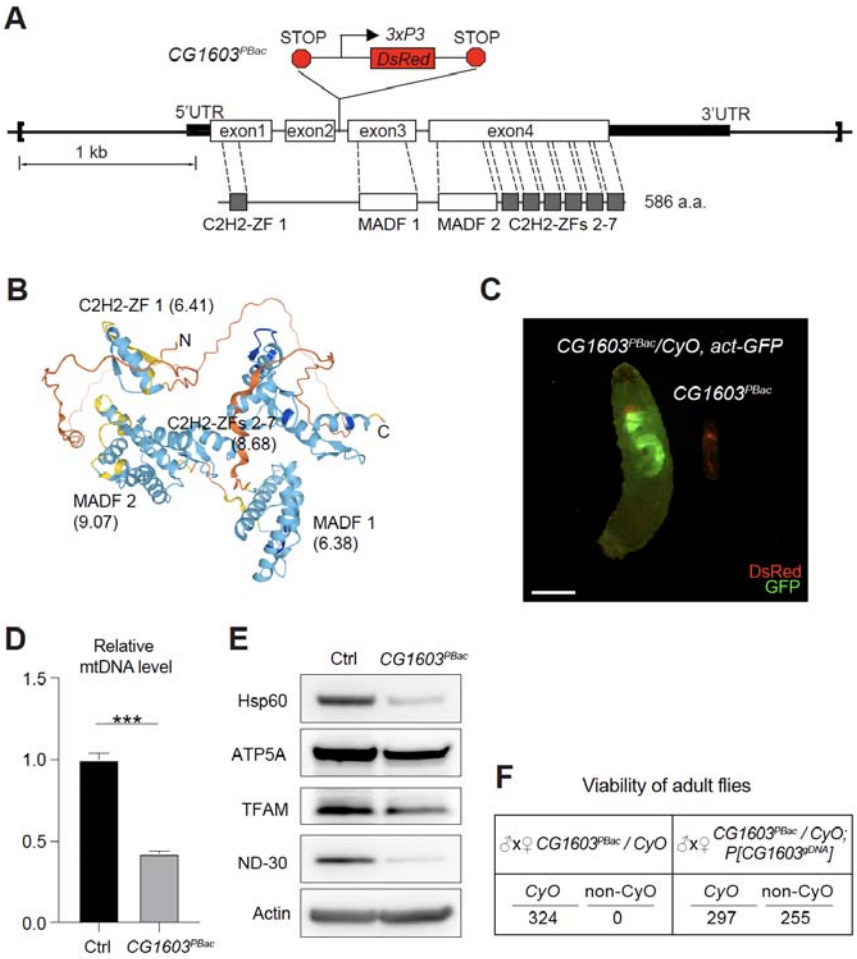


Figure 3. *CG1603* gene model, product, mutant and the genomic DNA transgene.

(A) Schematic representation of *CG1603* genomic locus, showing the *CG1603* transcript (5' and 3'UTR in black bar and four exons in white), its protein product (586 amino acids in length, and characterized by seven C2H2-ZF and two MADF domains), the $CG1603^{PBac}$ mutant allele (with a piggyBac insertion in the second intron, which is marked by fluorescent DsRed driven by an eye-specific *3xP3* promoter and flanked by stop codons in all three reading frames terminating translation through downstream), and the genomic region (in square brackets, from 955 bp upstream of the *CG1603* 5'UTR to 656 bp downstream of *CG1603* 3'UTR) used for $P[CG1603^{gDNA}]$ transgene. (B) Predicted 3D structure of the CG1603 protein by AlphaFold. Labels indicate the N-and C-terminus, as well as the specific protein domains along with their predicted isoelectric point (pI). (C) Images of $CG1603^{PBac}$ / *CyO, Act-GFP* and homozygous $CG1603^{PBac}$ larvae cultured together at 25°C, day 4 after egg laying. Green: GFP; Red: DsRed. Scale bars: 1 mm. (D) Relative mtDNA levels in $CG1603^{PBac}$ mutant larvae to control. n= 3, error bar: SD. ***:

p<0.001. (E) Western blots of mitochondrial proteins in *CG1603*^{PBac} mutant larvae to control. (F) *P[CG1603*^{gDNA}*]* restored viability of *CG1603*^{PBac}flies. The number of progeny for each genotype is listed.

The following figure supplement is available for figure 3:

Figure 3—figure supplement

A

Genotype	Phenotype	with *P[CG1603*^{gDNA}*]* transgene
CG1603^{PBac} / *CG1603*^{PBac}	Lethal	Viable
CG1603^{PBac} / *Def*^{k08815}	Lethal	Viable
CG1603^{PBac} / *Def*^{Exel6052}	Lethal	Viable
CG1603^{PBac} / *Def*^{Exel6053}	Viable	

B

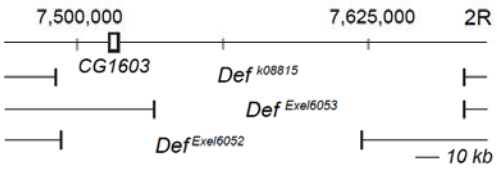


Figure 3—figure supplement A Summary of adult viability phenotypes of combinations of *CG1603*^{PBac} mutant, *P[CG1603*^{gDNA}*]* transgene and deficiency chromosomes.

Figure 3—figure supplement B Schematic map of deficiency chromosomes spanning *CG1603* genomic region.

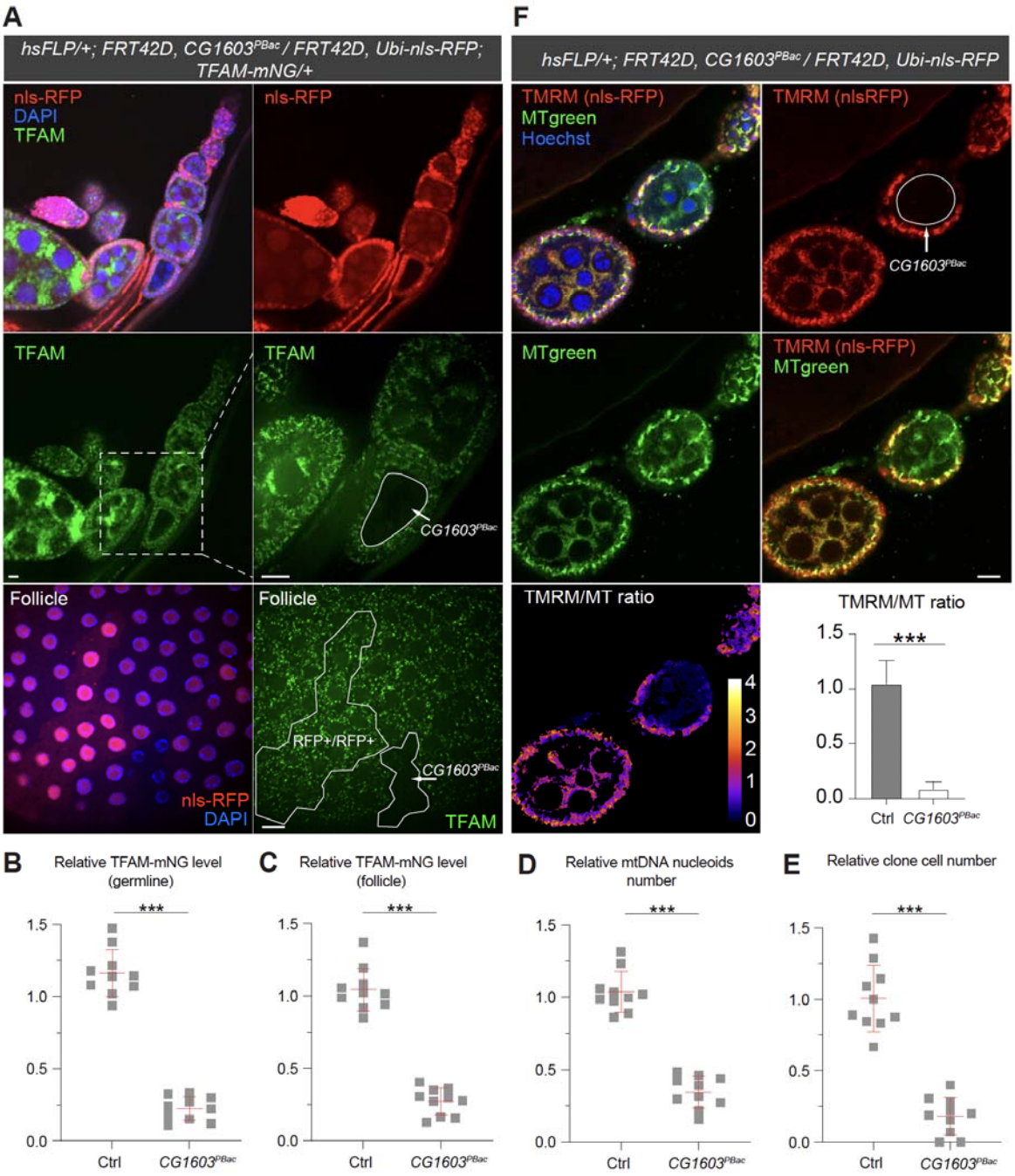


Figure 4. Clonal analyses confirmed CG1603's role in mitochondrial biogenesis and activity.

(A) Representative images of *CG1603*^{PBac} mutant germline (top and middle panel) and follicle (bottom panel) clones in late-stage egg chambers of adult ovaries with endogenously expressed TFAM-mNG visualized in green. Homozygous mutant clones were marked by the absence of RFP and compared with either flanking RFP-positive cysts (germline) or homozygous wild-type twin (follicle). White dash lines aided in illustrating clones. The wild-type (RFP+/RFP+) follicle clone showed markedly higher RFP

intensity than the heterozygous (RFP+/RFP-) cells, also see *Figure 4—figure supplement B*. Red: nls-RFP; Blue: DAPI. Scale bars: 10 μm. (B) Quantification of the relative TFAM-mNG level in the homozygous *FRT42D* control and *CG1603*^{PBac} mutant germline clones in early-stage egg chamber to their anterior flanking RFP-positive cysts within the same ovariole. also see *Figure 4—figure supplement A.* n=10 for each group, error bar: SD. ***: p<0.001. (C-E) Quantification of the relative TFAM-mNG level (C), the relative mtDNA nucleoids number (D) and the relative clone cell number (E) in the homozygous *FRT42D* control and *CG1603*^{PBac} mutant follicle clones to their wild-type twins. n=10 for each group, error bar: SD. ***: p<0.001. (F) TMRM / MitoTracker Green (MT) ratiometric live imaging and quantification of ovarioles containing homozygous *CG1603*^{PBac} mutant germline clones. Notably, in contrast to flanking control cysts, $\Delta\psi_m$ was almost absent in mutant clones. Please note that compared to TMRM, nls-RFP signal was too low to be detected in ratiometric imaging. Nonetheless, the nls-RFP was readily detected in control cysts, but not in homozygous *CG1603*^{PBac} clones, *via* visual observation, as depicted in *Figure 4A* and *Figure 4—figure supplement*. Blue: Hoechst. Scale bars: 10 μm. n= 8, error bar: SD. ***: p<0.001.

The following figure supplement is available for figure 4:

Figure 4—figure supplement

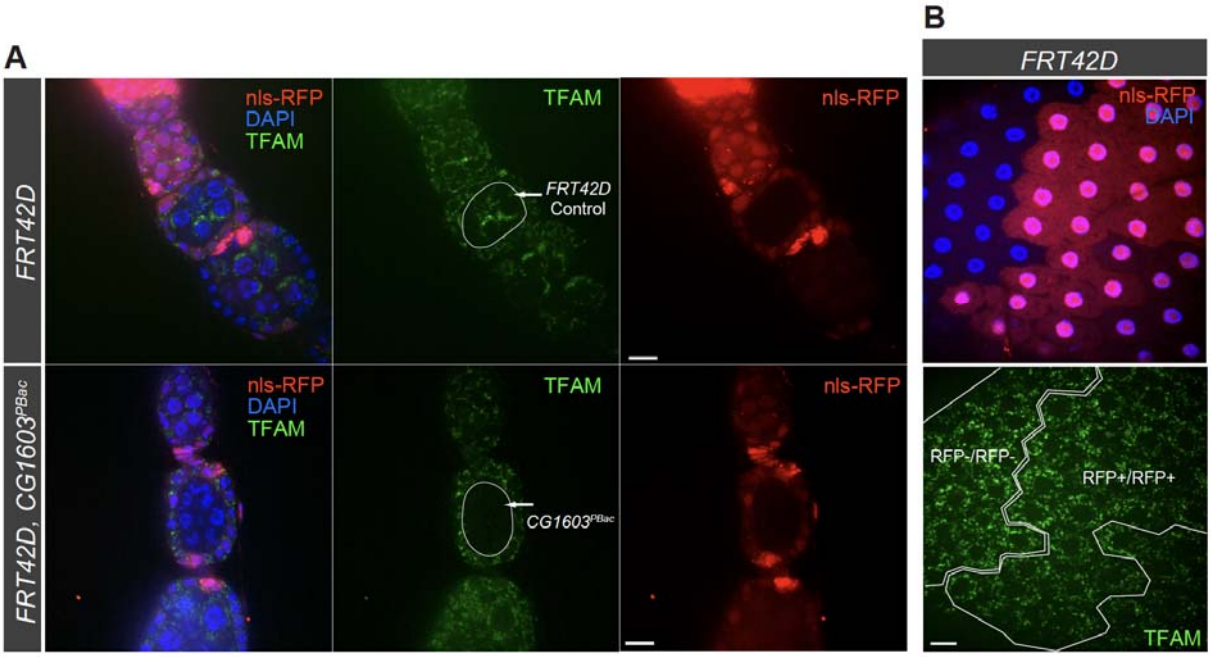


Figure 4—figure supplement A. Representative images of homozygous *FRT42D* control and *CG1603*^{PBac} mutant germline clones in early-stage egg chambers of adult ovaries with endogenously expressed TFAM-mNG visualized in green. Homozygous mutant clones were marked by the absence of RFP and compared with flanking RFP-positive cysts. Red: RFP; Blue: DAPI. Scale bars: 10 μm.

Figure 4—figure supplement B. Representative images of homozygous *FRT42D* control follicle cell clone (RFP-/RFP-) and its wild-type twin (RFP+/RFP+) with endogenously expressed TFAM-mNG visualized in green. Red: nls-RFP; Blue: DAPI. Scale bars: 10 μm.

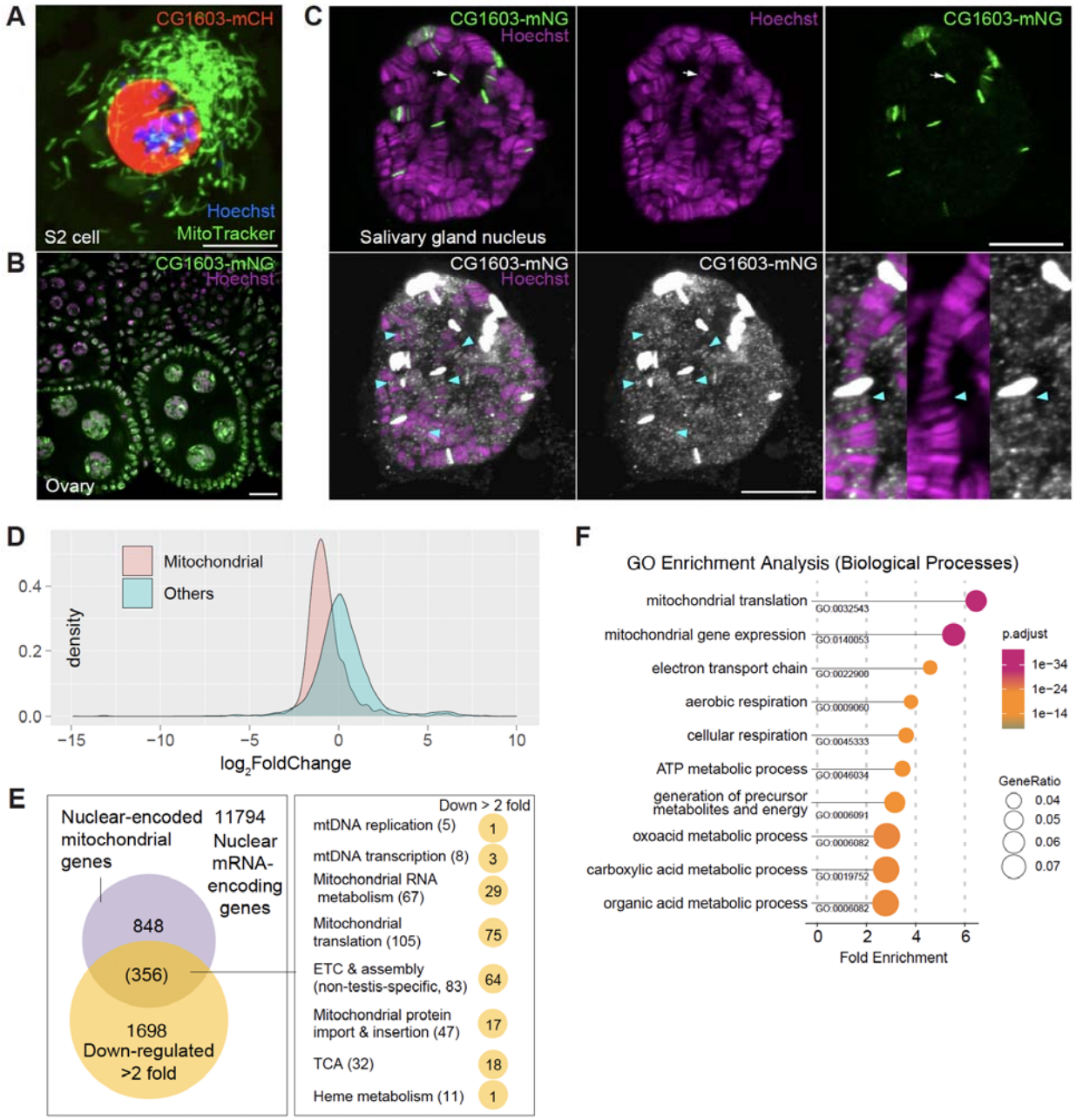


Figure 5. CG1603 localizes in the nucleus and is essential for regulating nuclear mitochondrial gene expression.

(A-B) Representative images showing nuclear localization of CG1603 protein in cultured S2 cell (A) and adult ovary (B). Green: MitoTracker Green in S2 cell, and CG1603-mNG in tissues; Red: CG1603-mCH; Blue & Magenta: Hoechst. Scale bars: 10 μm. (C) Representative images showing bindings of endogenously expressed CG1603 proteins to less condensed euchromatin regions in the polytene chromosomes of salivary gland. High intensity CG1603-mNG bands were visualized in green in upper

panel as indicated by arrow, and low intensity bands were visualized in white in lower panel as indicated by arrow heads with image B&C adjusted. Magenta: Hoechst. Scale bars: 10 μm. (D) Density plot illustrating the distribution of expression changes of the nuclear-encoded mitochondrial and non-mitochondrial genes in *CG1603*^{PBac} mutant. (E) Graph illustrating the overlap between nuclear-encoded mitochondrial genes and differentially expressed genes (DEGs) that down-regulated > 2-fold, as well as the distribution of the overlapped genes in different mitochondrial function categories. (F) Gene Ontology (GO) enrichment analyses of DEGs that down-regulated > 2-fold. The top 10 enriched biological processes are shown.

The following figure supplement is available for figure 5:

Supplementary file 2. List of nuclear-encoded mitochondrial genes, including gene symbols, IDs and subgroup information.

Supplementary file 3. List of differentially expressed nuclear-encoded genes in CG1603^{PBac} *mutant flies compared to controls.*

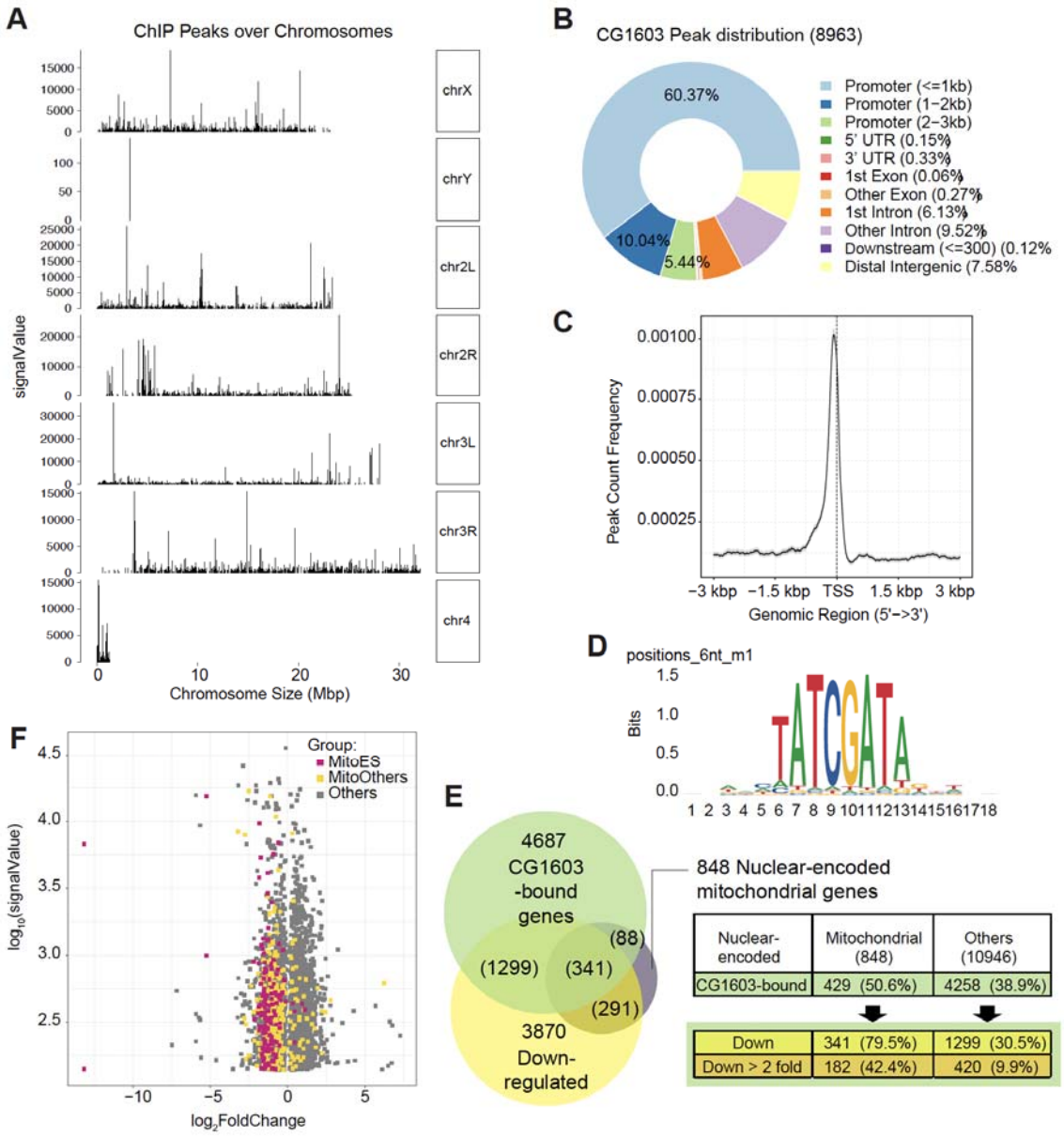


Figure 6. ChIP analysis identified nuclear mitochondrial genes that may be directly regulated by CG1603.

(A) CG1603 ChIP peaks over all chromosomes. (B) Genomic distribution of CG1603 peaks. (C) Average profile of CG1603 peaks binding to transcription start site (TSS) regions. (D) Representative binding motif discovered with CG1603 ChIP peaks. (E) Summary of the number of nuclear-encoded mitochondrial and non-mitochondrial mRNA coding genes bound by CG1603, and the overlapping down-regulated differentially expressed genes (DEGs) in each group. (F) Scatterplot illustrating the signalValue of CG1603 ChIP peaks (y-axis) and log2 fold change in expression of DEGs between *CG1603*^{PBac} mutant and control (x-axis). MitoES: the genes belong to the categories that are clearly essential to mitochondrial

ETC biogenesis and maintenance, including ETC subunits and assembly factors, mtDNA replication and transcription, mitochondrial RNA metabolism and translation, as well as mitochondrial protein import and membrane insertion machinery.

The following figure supplement is available for figure 6:

Supplementary file 4. List of CG1603 peaks from modERN ChIP-seq data, including gene annotation, group information, combined with RNAseq analysis result of CG1603 ^{PBac} *mutant flies.*

Supplementary file 5. CG1603 binding motifs discovered by RSAT peak-motifs.

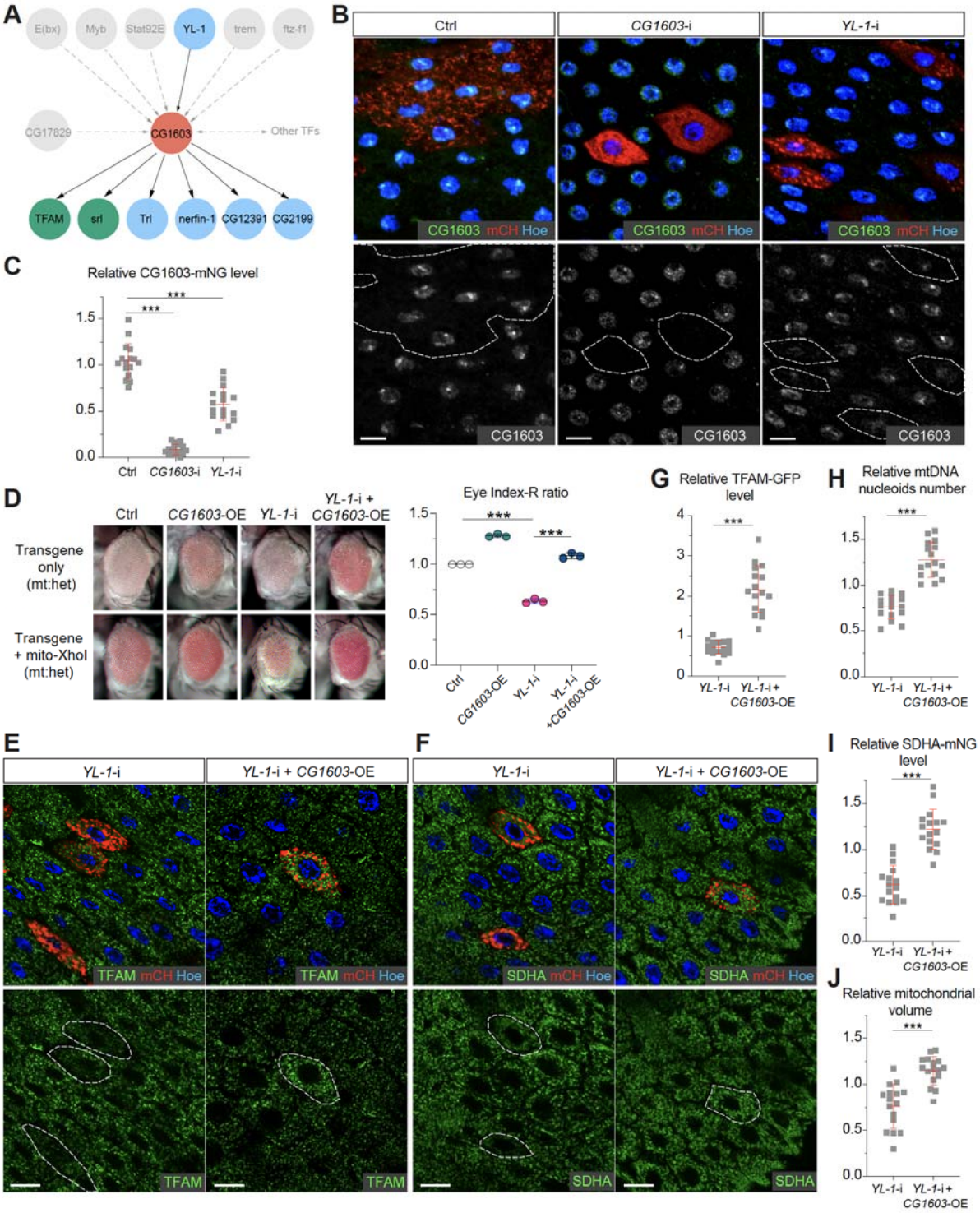


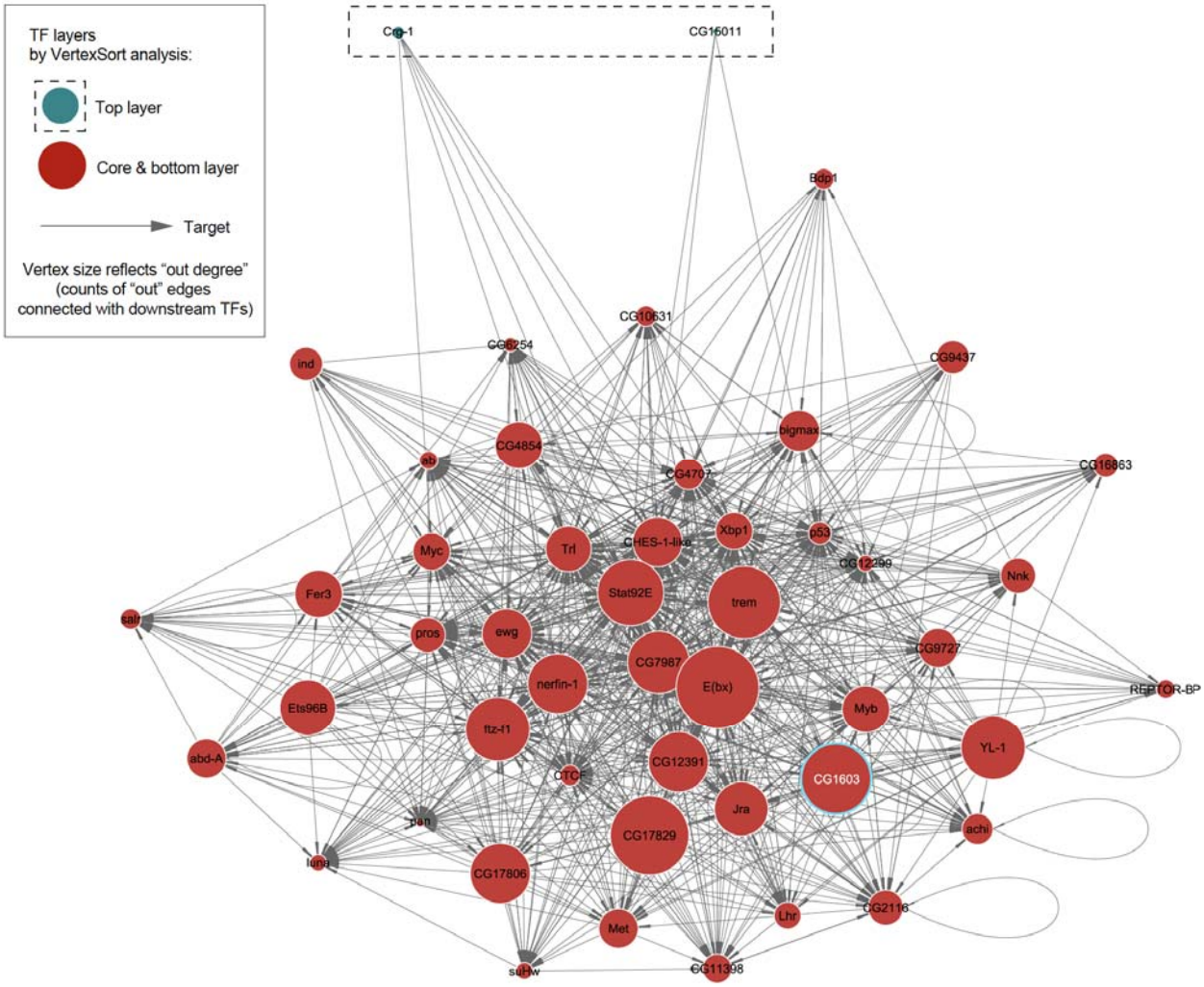
Figure 7. Network analyses of CG1603 and other transcription factors

(A) Schematic graph illustrating the CG1603 upstream and downstream (co-)TFs involved in regulating mitochondrial ETC biogenesis, inferred from ChIP-seq, RNAseq and genetics data. (B) Representative images of control RNAi (Ctrl), *CG1603* RNAi (*CG1603*-i) and *YL-1* RNAi (*YL-1*-i) midgut EC clones

with endogenously expressed CG1603-mNG visualized in green or white. Clones were labeled by mCherry red and compared with wild-type neighbors. White dash lines aided in illustrating clones. Blue: Hoechst. Scale bars: 10 μm. (C) Quantification of the relative CG1603-mNG level in the EC clones to their wild-type neighbors. n=16 from 8 midguts for each group, error bar: SD. ***: $p<0.001$. (D) Representative eye image and Index-R ratio (RNAi + mitoXhoI / RNAi only) of adult flies with indicated genotypes. Three biological repeats were performed for each group, error bar: SD. ***: $p<0.001$. (E-F) Representative images of *YL-1* RNAi (*YL-1*-i) and *YL-1* RNAi + *CG1603* overexpression (*YL-1*-i + *CG1603*-OE) midgut EC clones with endogenously expressed TFAM-GFP (E) or SDHA-mNG (F) visualized in green. Clones were labeled by mCherry red and compared with wild-type neighbors. Blue: Hoechst. Scale bars: 10 μm. (G-J) Quantification of the relative TFAM-GFP level (G), the relative mtDNA nucleoids number (H), the relative SDHA-mNG level (I) and the relative mitochondrial volume (J) in the EC clones to their wild-type neighbors. n=16 from 8 midguts for each group, error bar: SD. ***: $p<0.001$.

855 **The following figure supplement is available for figure 7:**

Figure 7—figure supplement A



856

857 Figure 7—figure supplement A. The potential transcriptional regulatory network of nuclear-encoded

858 mitochondrial genes.

Figure 7—figure supplement B

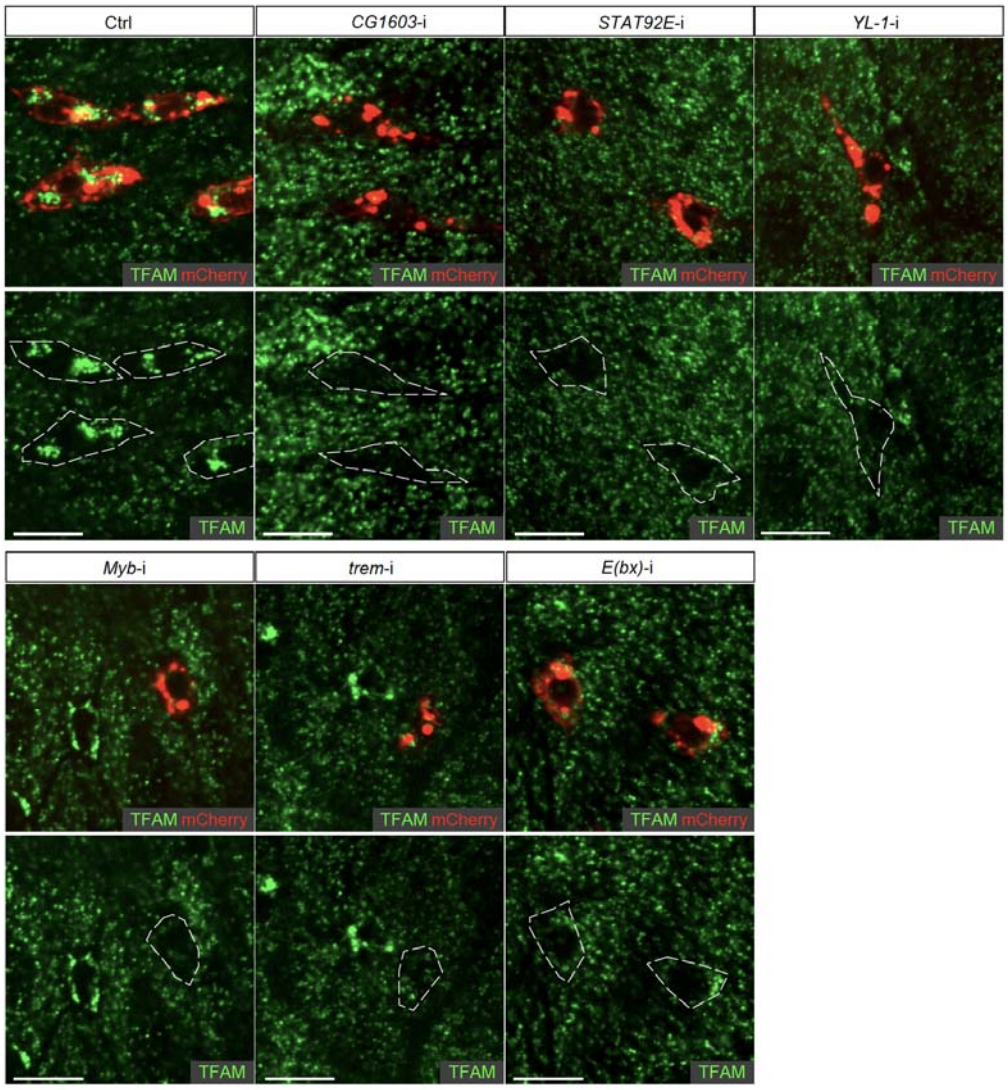


Figure 7—figure supplement B. Representative images of control RNAi (Ctrl), *CG1603* RNAi (*CG1603*-i), *STAT92E* RNAi (*STAT92E*-i), *YL-1* RNAi (*YL-1*-i), *Myb* RNAi (*Myb* -i), *trem* RNAi (*trem*-i) and *E(bx)* RNAi (*E(bx)*-i) midgut ISC/EB clones labeled by mCherry red, with endogenously expressed TFAM-GFP visualized in green. White dash lines aided in illustrating clones. Scale bars: 10 µm.

Supplementary file 6. Vertices, edges and vertex.sort analysis information of the potential transcriptional regulatory network of nuclear-encoded mitochondrial genes.

Supplementary file 7. Gene binding profile of 49 synergistic enhancer TFs.