A transcription network underlies the 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 biogenesis requires the expression of genes encoded by both the nuclear and mitochondrial genomes. However, aside from a handful transcriptional factors regulating specific subsets of mitochondrial genes, the overall architecture of the transcriptional control of mitochondrial biogenesis remains to be elucidated. The mechanisms coordinating these two genomes are largely unknown. We performed a targeted RNAi screen in developing eyes with reduced mitochondrial DNA content, anticipating a synergistic disruption of tissue development due to impaired mitochondrial biogenesis and mtDNA deficiency. Among 638 transcription factors annotated in Drosophila genome, 77 were identified as potential regulators of mitochondrial biogenesis. Utilizing published ChIP-seq data of positive hits, we constructed a regulatory network revealing the logic of the transcription regulation of mitochondrial biogenesis. Multiple transcription factors in core layers had extensive connections, collectively governing the expression of nearly all mitochondrial genes, whereas factors sitting on the top layer may respond to cellular cues to modulate mitochondrial biogenesis through the underlying network. CG1603, a core component of the network, was found to be indispensable for the expression of most nuclear mitochondrial genes, including those required for mtDNA maintenance and gene expression, thus coordinating nuclear genome and mtDNA activities in mitochondrial biogenies. Additional genetics analyses validated YL-1, a transcription factor upstream of CG1603 in the network, as a regulator controlling CG1603 expression and 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

59 regulating ETC biogenesis. The nuclear respiration factors, NRF1 and NRF2 activate the
60 expression of many nuclear-encoded ETC subunits and genes essential for mtDNA replication
61 and transcription (Scarpulla, 2008). This regulation allows NRFs to indirectly control the
62 expression of mtDNA-encoded genes, and hence coordinate the activities of both genomes in
63 ETC biogenesis. The peroxisome proliferator-activated receptors (PPARs), upon activation by
64 diverse lipid ligands, induce the expression of nuclear genes in fatty acids oxidation pathway
65 (Berger and Moller, 2002). Another family of nuclear receptors, the estrogen-related receptors
66 (ERRs) regulate nuclear genes involved in oxidative phosphorylation, including ETCs and the
67 citric acid (Deblois and Giguere, 2011). Notably, all forementioned transcription factors share a
68 common co-activator, PPARγ coactivator-1α (PGC-1α) that directly stimulates the basal
69 transcriptional machinery (Finck and Kelly, 2007; Hock and Kralli, 2009; Scarpulla et al., 2012).
70 PGC-1α and related coactivators are dynamically regulated in responses to various physiological
71 or environmental cues, to adjust metabolic program and energy metabolism accordingly (Hock
72 and Kralli, 2009; Scarpulla *et al.*, 2012). Members of PPARs or ERRs families often show tissue-
73 specific expression and regulate subsets of mitochondrial genes (Deblois and Giguere, 2011;
74 Hock and Kralli, 2009). However, given the large number and diverse evolution origins of
75 mitochondrial genes (Kurland and Andersson, 2000; Rath et al., 2021), the current understanding
76 on transcriptional regulation of mitochondrial biogenies appears incomplete. Additionally,
77 mechanisms coordinating nuclear genome and mtDNA activities in ETC biogenesis remain
78 unclear.
79
80 Recently, the modERN (model organism Encyclopedia of Regulatory Networks) project generated
81 genome-wide binding profiles of a large set of transcription factors in *C. elegans* and *Drosophila*
82 *melanogaster* (Kudron et al., 2018). The global mapping of transcription factor (TF)-DNA
83 interactions could potentially be applied to identify the transcriptional network governing
84 mitochondrial biogenesis. However, the DNA binding profiles of TFs have their limitations in
85 understanding the true biological functions of TFs in gene regulation (Jiang and Mortazavi, 2018).
86 Gene expression is, in most cases, subject to the combined influences of multiple TFs.
87 Additionally, an individual TF may have either activating or repressive roles based on the local
88 chromatin environment (Jiang and Mortazavi, 2018). Furthermore, despite the substantial
89 progress in bioinformatics analyses, the interpretation of genome-wide omics data still has its

limitations 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 RNAi targeting a mitotic cyclin, *CycB* (*Figure 1A*), would disrupt eye development. Therefore, assaying eye morphology alone is not sufficient to enrich candidates regulating ETC biogenesis.

The mitochondrial genome of wt *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 wild type 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 wild type 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

143 vial, thereby minimizing any potential discrepancy caused by environmental factors. Most RNAi
144 flies survived to adult stage but had reduced eye size. A few RNAi flies were lethal at pupae stage,
145 due to a lack of head capsule that is derived from the eye antenna disc.

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

163 To understand transcriptional regulations of ETC biogenesis, we screened 1264 RNAi lines that
164 cover 638 genes annotated as transcriptional regulators in *Drosophila* genome. Total 77
165 enhancers were identified (*Figure 1F, G* and *Supplementary file 1*), including all known factors
166 involved in ETC biogenesis such as Myc, TFAM (Scarpulla, 2008; Wang et al., 2019). We also
167 recovered 20 suppressors, of which, eyes of RNAi+MitoXhoI flies were larger than the
168 corresponding RNAi-only flies.
169

170 **Regulatory network of mitochondrial biogenesis**

171 Among 77 TFs identified in the initial modifier screen, 49 TFs have ChIP-seq data available in
172 modERN (Kudron *et al.*, 2018). To further understand the transcriptional regulation of ETC

173 biogenesis, we performed the network analysis on these 49 TFs using the "*vertex sort"* algorithm
174 (Jothi et al., 2009), and constructed a regulatory network (*Figure 2A* and *Supplementary file 2*).
175 All 49 TFs had binding sites on the promoter region of at least one nuclear mitochondrial gene
176 (*Figure 2B,* Figure *2—figure supplement* and *Supplementary file 3*). Respectively, 89% nuclear-
177 encoded mitochondrial genes (851), including nearly all Mito-EBR genes (*Supplementary file 3),*
178 were bound by at least one TF. Given that 28 hits were not included in the analyses due to a lack
179 of ChIP-seq data, the actual coverage of total 77 TFs on nuclear mitochondrial genes would be
180 more comprehensive. Six TFs bound to more than half of mitochondrial genes (*Figure 2B,* Figure
181 *2—figure supplement* and *Supplementary file 3*). However, not a single TF covered all
182 mitochondrial genes, or all genes in a specific mitochondrial process, which is consistent with the
183 diverse evolution origin of mitochondrial genes (Kurland and Andersson, 2000). It also indicates
184 that there is no such a "master" regulator controlling all aspects of mitochondrial genesis. Forty-
185 seven TFs were identified as strongly connected component due to their extensive connections
186 and were classified in the core or bottom layer of the hierarchical structure, suggesting complexed
187 co-regulations and potential redundancy among these TFs in controlling mitochondrial biogenesis
188 (*Figure 2A* and *Supplementary file 2*). Through this interconnected network, every single node
189 can link to all 851 mitochondrial genes in the coverage. Two transcription factors, Crg-1 and
190 CG15011 were identified as the top-layer TFs with no upstream regulators in the network (*Figure*
191 *2A* and *Supplementary file 2*). Crg-1 is a circadian regulated gene (Rouyer et al., 1997). CG15011
192 is an X-box binding protein and had a binding profile to another X box binding protein, Xbp1
193 (*Supplementary file 3)*, a stress respondent and regulator (Acosta-Alvear et al., 2007). These top-
194 layer TFs may sense physiological oscillations and stresses to modulate mitochondrial biogenesis
195 through the underlying network. Additionally, YL-1 and E(bx), two components in the middle layer,
196 are involved in chromatin-remodeling (Liang et al., 2016; Wysocka et al., 2006), suggesting a
197 potential regulation of mitochondrial biogenesis at a chromatin level.
198
199 **CG1603 regulates ETC gene expression and mitochondrial biogenesis.**
200 To validate the efficacy of this integrated genetic and bioinformatic approach, we next followed up
201 on *CG1603*, one of the strongest hits from the primary screen (*Figure 1A, G*) and exhibited binding
202 to a diverse array of genes associated with ETC biogenesis (*Figure 2B,* Figure *2—figure*
203 *supplement* and *Supplementary file 3*). *CG1603* RNAi slightly reduced eye size. However, the

204 combination of *CG1603* RNAi with MitoXhoI expression in the heteroplasmic background resulted
205 in markedly smaller eyes, indicating a clear synergy between the inhibition of *CG1603* and the
206 mtDNA deficiency. We next asked whether CG1603 was involved in mtDNA maintenance. The
207 *Drosophila* midgut is essentially a monolayer epithelium and composed of intestine stem cells,
208 enteroblasts, enteroendocrine cells and enterocytes (EC). The large, flattened enterocytes allow
209 high-resolution imaging of mitochondria and mitochondrial nucleoids. Additionally, the simple
210 organization and distinct cell types, containing both proliferative and terminally differentiated cells,
211 render the midgut an ideal model to evaluate the impact of mitochondrial disruptions on cell
212 proliferation and differentiation (Zhang et al., 2020). We used a "flip-out" method to activate
213 *CG1603* RNAi in a subset of cells (Prober and Edgar, 2000; Zhang *et al.*, 2020), and imaged
214 TFAM-GFP (Zhang et al., 2016), a marker for mitochondrial nucleoids in midgut clones. Both the
215 total TFAM-GFP level and the number of mtDNA nucleoids (TFAM puncta), were markedly
216 reduced in *CG1603* RNAi clones (*Figure 3A-C*), suggesting that CG1603 is necessary for
217 maintaining the steady-state level of mtDNA. We constructed a *SDHA-mNG* reporter line by
218 inserting the *mNeonGreen (mNG)* cDNA in-frame, downstream of the endogenous locus of
219 *SDHA*, a subunit of ETC Complex II that are entirely encoded by the nuclear genome. SDHA-
220 mNG level was notably reduced in *CG1603* RNAi clones (*Figure 3D, E*), suggesting that CG1603
221 is also required for the expression of nuclear-encoded ETC subunits. Different from TFAM-GFP
222 that marks mitochondrial nucleoids and appears as puncta in mitochondria (Chen et al., 2020),
223 SDHA-mNG uniformly distributed in the mitochondrial matrix (*Figure 3D*). By quantifying the total
224 volume of SDHA-mNG positive voxels in the 3-D rendering, we found that the total mitochondrial
225 volume was also reduced in *CG1603* RNAi clones (*Figure 3F*). Collectively, these results
226 demonstrate that CG1603 regulates the expression of genes essential for both ETC function and
227 mitochondrial biogenesis. *CG1603* RNAi produced very few EC clones (*Figure 3A, D* and *Figure*
228 *3—figure supplement*), consistent with the notion that mitochondrial respiration is necessary for
229 ISCs differentiation (Zhang *et al.*, 2020).
230
231 **CG1603 regulates cell growth and differentiation.**
232 *CG1603* encodes a C2H2 zinc finger protein. It has one C2H2 Zinc finger (C2H2-ZF) at its N-
233 terminus, followed by two MADF (myb/SANT-like domain in Adf-1) domains, and six additional
234 zinc fingers at the C-terminus (*Figure 4A, B*). A PiggyBac transgene,

235 *PBac[SAstopDSRed]LL06826* is inserted between the exons 2 and 3 of *CG1603* locus. This
236 modified Piggybac mutator transgene contains splicing donors and stop codons in all 3 reading
237 frames (Schuldiner et al., 2008), and thereby would disrupt the translation of the full length
238 CG1603 protein. Homozygous *PBac[SAstopDSRed]LL06826* was lethal, arrested at the 2nd instar
239 larval stage and eventually died after 10 days (*Figure 4C)*. Both the steady state level of mtDNA,
240 and total mitochondrial mass assessed by the levels of several mitochondrial proteins were
241 reduced in these larvae (*Figure 4D, E*), as well as the integrities and activities of ETC complexes
242 (*Figure 4—figure supplement A*). The lethality of this PiggyBac transgene was mapped to a
243 genomic region spanning the *CG1603* locus (*Figure 4—figure supplement B, C*). Importantly, a
244 *P[CG1603^{gDNA}]* transgene that covers the genomic region of *CG1603* fully rescued its viability
245 (*Figure 4A, F*). These results demonstrate that the lethality of *PBac[SAstopDSRed]LL06826* was
246 caused by the loss of function of CG1603, and we hence named it *CG1603^{PBac}* thereafter. Using
247 FLP/FRT-mediated recombination, we generated homozygous *CG1603^{PBac}* mutant clones in both
248 germline and follicle cells in adult ovaries. Consistent with the results of "flip-out" RNAi
249 experiments in the midgut, both the total TFAM level and the number of mtDNA nucleoids,
250 visualized by an endogenously expressed TFAM-mNG reporter, were significantly reduced in
251 *CG1603^{PBac}* clones (*Figure 5A-D* and *Figure 5—figure supplement A, B*). In most *CG1603^{PBac}*
252 clones, TFAM-mNG puncta were hardly observed, demonstrating an essential role of CG1603 in
253 mtDNA maintenance. Compared to twin clones, *CG1603^{PBac}* follicle cell clones contained
254 significantly fewer cells, and these cells were smaller, indicating that CG1603 promotes both cell
255 growth and cell proliferation (*Figure 5A, E)*. *CG1603^{PBac}* egg chambers were also notably small,
256 even smaller than the adjacent anterior egg chambers that are at earlier developmental stages in
257 the same ovariole (*Figure 5A)*. We assessed $\Delta\psi_m$ using the ratiometric imaging of TMRM and
258 MitoTracker Green (Zhang et al., 2019). $\Delta\psi_m$ was nearly abolished in *CG1603^{PBac}* clones with
259 reduced MitoTracker Green staining (*Figure 5F and Figure 5- figure supplement A*). All together,
260 these observations demonstrate that CG1603 promotes mitochondrial biogenesis and is essential
261 for ETC biogenesis.
262
263 **CG1603 is a transcription factor regulating nuclear mitochondrial gene expression.**
264 CG1603 protein exclusively localized to nucleus when expressed in cultured cells (*Figure 6A*).
265 We generated a transgene expressing CG1603-mNG fusion protein by inserting *mNeonGreen*

266 cDNA into the endogenous locus of *CG1603*. CG1603-mNG localized to nuclei in ovaries (*Figure*
267 *6B*) and directly bound to polytene chromosomes in salivary gland (*Figure 6C*). Notably, CG1603-
268 mNG was highly enriched on less condensed chromatin regions that had weak Hoechst staining
269 (*Figure 6C*). We performed RNA sequencing (RNA-seq) in larvae to uncover potential targets of
270 CG1603. Between wt and *CG1603*^{PBac} larvae, total 7635 genes were differentially expressed,
271 including 86% nuclear-encoded mitochondrial genes (*Figure 6D* and *Supplementary file 4, 5*).
272 Nearly half of nuclear-encoded mitochondrial genes were among 1698 genes that were reduced
273 by more than 2-folds in *CG1603* mutant (*Figure 6E* and *Supplementary file 5a*), including many
274 structural subunit genes of all five ETC complexes (*Supplementary file 5b*), some of which were
275 further confirmed by quantitative real-time PCR (*Figure 6F)*. Gene Ontology (GO) enrichment
276 analyses on these 1698 genes also revealed that all top ten significantly enriched biological
277 processes were related to mitochondria, including "mitochondrial translation", "mitochondrial gene
278 expression", "electron transport chain", "aerobic respiration", "cellular respiration" and "ATP
279 metabolic process" (*Figure 6G*).

280 CG1603 had 8963 binding sites (peaks) distributing over all four chromosomes (*Figure 7A* and
281 *Supplementary file 6*). A subset of peaks showed high intensity evaluated by signalValue (*Figure*
282 *7A* and *Supplementary file 6*), which may correspond to these high intensity CG1603-mNG bands
283 on the polytene chromosomes of salivary gland (*Figure 6C*). Most CG1603 binding sites (6799)
284 were found at promoter regions, close to the transcription start site (*Figure 7B, C* and
285 *Supplementary file 6*), which is a key feature of a typical TF. Using the RSAT "peak-motifs" tool
286 (Thomas-Chollier et al., 2012), an 8-bp palindromic sequence, "TATCGATA" emerged as the most
287 prevalent CG1603 binding motif (*Figure 7D* and *Supplementary file 7*). CG1603 bound to the
288 genomic regions of 50% nuclear-encoded mitochondrial genes, and among these genes, 79.5 %
289 were down-regulated in the *CG1603*^{PBac} mutant (*Figure 7E* and *Supplementary file 6),* indicating
290 a great accordance between ChIP data and RNAseq results. Most nuclear-encoded mitochondrial
291 genes that were both bound by CG1603 and down-regulated in *CG1603* mutant were ETC genes
292 or related to ETC biogenesis (*Figure 7F* and *Supplementary file 6)*. Collectively, CG1603 appears
293 to be essential for mitochondrial biogenesis and coordinates the expression of both nuclear and
294 mtDNA genes in ETC biogenesis.
295

296 **The integrated approach identifies YL-1 as an upstream regulator of CG1603.**

297 In the network analyses, CG1603 was positioned in the middle layer, linked to 7 TFs above and
298 6 TFs below by integrating the RNA-seq result with ChIP-seq data (*Figure 8A).* Through these
299 TFs below, CG1603 may indirectly control the expression of 2230 genes, including 291
300 mitochondrial genes downregulated in *CG1603*^{PBac} but not bound by CG1603 (*Figure 7E*). Using
301 the "flip-out" RNAi system in the midgut, we found that among 7 TFs upstream of CG1603 in the
302 network, E(bx), YL-1, trem, STAT92E and Myb were also required for maintaining TFAM levels
303 (*Figure 8—figure supplement A, B*). To further verify their potential roles in regulating CG1603,
304 we performed RNAi against these genes in midgut clones carrying CG1603-mNG reporter. Only
305 *YL-1* RNAi clones displayed a markedly reduction of CG1603 protein compared with neighboring
306 cells (*Figure 8B, C*). Furthermore, overexpression of CG1603 restored eye size, TFAM-GFP,
307 SDHA-mNG and mtDNA levels cause by *YL-1* RNAi (*Figure 8D-J and Figure 8—figure*
308 *supplement C*). These results indicate that YL-1 is indeed an upstream regulator of CG1603, and
309 through which to regulate ETC biogenesis.
310

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 mtDNA-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 core regulator in a transcription network regulating 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

342 size (*Figure 1G* and *Supplementary file 1*). Noteworthy, 5 of them were lethal due to the lack of
343 head capsule that is developed from the eye antenna disc, but the viability of these RNAi flies was
344 restored by MitoXhoI expression. Given that MitoXhoI expression also disrupts eye development,
345 it is perplexing that the combination of RNAi and MitoXhoI expression, two genetic conditions
346 causing the same phenotype, led to a milder phenotype. Perhaps, mitochondrial DNA deficiency
347 caused by MitoXhoI expression triggers a retrograde signal, which boosts cellular stress
348 responses and thereby mitigates the cell growth defects in these RNAi backgrounds.
349
350 The CG1603 belong to a large family of C2H2 Zinc finger (C2H2-ZF) transcription factors that
351 contains 272 genes in *Drosophila* genome (https://flybase.org/reports/FBgg0000732.html). It has
352 one N-terminus C2H2-ZF, followed by two MADFs and a cluster of six C2H2-ZFs at the C-
353 terminus (*Figure 4A*). In addition to the C2H2-ZF cluster, which predominantly mediates
354 sequence-specific DNA binding (Persikov et al., 2015; Wolfe et al., 2000), C2H2-ZF transcription
355 factors often possess additional N-terminal protein-protein interaction domains, such as KRAB,
356 SCAN and BTB/POZ domains in vertebrates, ZAD and BTB/POZ in *Drosophila*, for binding to
357 transcription co-regulators (Fedotova et al., 2017; Perez-Torrado et al., 2006; Sobocinska et al.,
358 2021). These interactions allow them to either activate or repress gene expression. CG1603 binds
359 to the genomic regions of 4687 genes in the *Drosophila* genome. Among these genes, 602 and
360 562 genes were, respectively, decreased or increased more than two folds in *CG1603* mutant
361 (*Supplementary file 6)*. Thus, CG1603 is likely a dual-function transcription factor, capable of both
362 activating and repressing transcription depends on the chromosomal environment of its targets.
363 Ying Yang 1, a well characterized dual-function transcription factor in mammals, contains both a
364 transcription activation domain and a repression domain, in addition to four C2H2-ZFs at its C-
365 terminus (Verheul et al., 2020). The N-terminus C2H2-ZF and the first MADF domain of CG1603
366 are negatively charged, thus have a low likelihood of binding to DNA that is also negatively
367 charged (*Figure 4B*). In the predicted 3-D structure of CG1603, the positively charged C-terminal
368 zinc fingers and MADF-2 domain cluster in the center, while the negatively charged N-terminal
369 C2H2-ZF and the MADF-1 extend outward to the opposite directions (*Figure 4B*), resembling the
370 domain arrangement of Ying Yang 1 (Verheul *et al.*, 2020). MADF domains share significant
371 similarity with Myb/SANT domains that may bind to either DNA or proteins (Maheshwari et al.,
372 2008). Some MADF domains, due to their negative charge, have been proposed to interact with

373 positively charged histone tails, similar to the Myb/SANT domain in a well-known chromatin
374 remodeler ISWI, suggesting a potential role in chromatin remodeling (Maheshwari et al., 2008).
375 Notably, some chromatin remodelers possess tandem Myb/SANT domains that can directly
376 interact with histones and DNA or histone remodeling enzymes like ISWI and SMRT (Boyer et al.,
377 2004). It is plausible that CG1603 may play a role in chromatin remodeling directly or by recruiting
378 nucleosome remodeling factors to its binding sites, thereby modulating gene transcription in those
379 regions.
380
381 Notably, CG1603 had no impact on the expression of one third of its binding genes
382 (*Supplementary file 6)*, highlighting that DNA binding profiling along is not sufficient to predict the
383 function of a transcription factor. Nonetheless, the network analyses on TFs' ChIP-seq data
384 allowed us to construct a potential regulatory network among these transcription factors, which
385 subsequently served as a blueprint for genetic analyses to verify potential regulations. Total 7
386 transcription factors were upstream of CG1603 in the network and emerged as positive hits in the
387 initial screen in the eye. RNAi against 5 of them led to reduced TFAM level in the midgut, while
388 the other two had no noticeable phenotype, suggesting that these two TFs may regulate
389 mitochondrial biogenesis in a tissue specific manner. Only YL-1 was confirmed to act upstream
390 of CG1603 based on the genetic epistasis analysis, further indicating the necessity of combining
391 genomic, bioinformatic, and genetics analyses to gain more reliable and comprehensive
392 understanding on transcriptional regulations. YL-1 is one of DNA binding subunits of the SRCAP
393 complex, which is essential for histone H2A.Z incorporation and replacement (Liang *et al.*, 2016).
394 Recently, it has been shown that both SRCAP complex and H2A.Z are necessary for the
395 transcription of nuclear-encoded mitochondrial genes (Lowden et al., 2021; Xu et al., 2021). Our
396 work offers a mechanistic insight into how CG1603 and its upstream regulator, YL-1 may regulate
397 mitochondrial biogenesis at nucleosome and chromatin level. Currently, most known transcription
398 paradigms controlling mitochondrial biogenesis are centered on transcription factors and co-
399 activators that stabilize or directly stimulate the core transcription machinery. The YL-1 to CG1603
400 cascade may represent a previously underappreciated layer of transcriptional regulation on ETC
401 biogenesis and could act in concert with NRF1 and other transcription factors to coordinate both
402 the nuclear and mitochondrial genome in ETC biogenesis.
403

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 wt 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).

482 ChIP-sequencing reads were aligned to the Drosophila Melanogaster reference genome (Dmel6) using
483 BWA (v0.7.17) (Li and Durbin, 2009). SAM files were sorted and compressed into BAM format with
484 Samtools (v1.16.1) (Li et al., 2009). Replicates were merged by Picard tools (v2.27.3,
485 https://broadinstitute.github.io/picard) using lenient criteria, and all alignments with a MAPQ value less than
486 20 were removed. Lags prediction and peak-calling were done with MACS2 (v2.2.7.1) (Zhang et al., 2008)
487 following the ENCODE TF ChIP pipeline with IDR analysis performed for consistency analysis
488 (https://github.com/mforde84/ENCODE_TF_ChIP_pipeline). Peak annotation and analysis of profile of
489 ChIP peaks binding to TSS regions were performed with ChIPseeker R package (Yu et al., 2015).
490 Transcription network was analyzed and visualized with VertexSort (Jothi *et al.*, 2009) and igraph
491 (https://igraph.org) R packages, respectively, and ChIP peaks of each TF identified in the gene promoter
492 regions (<2kb) were used for analyses. CG1603 binding motif discovery was done by online integrated
493 pipeline 'peak-motifs' of RSAT tools (https://rsat.france-bioinformatique.fr/rsat/peak-motifs_form.cgi)
494 (Thomas-Chollier *et al.*, 2012).

495 **Imaging analyses**

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

518 Mitochondrial membrane potential was detected using a protocol adopted from a previously study (Zhang
519 et al., 2020; Zhang et al., 2019). Briefly, after dissection, adult ovaries were incubated in the Schneider's
520 medium containing TMRM (200nM, Thermo Fisher Scientific) and MitoTracker Green (200nM, Thermo
521 Fisher Scientific) for 20 min at room temperature, rinsed with PBS for 3 times, and then imaged live within
522 1 h. TMRM and MitoTracker signal intensities were quantified and ratiometric image was generated using

523 Fiji / Image J software (NIH). Mitochondrial membrane potential was computed as the ratio of the mean
524 intensity of TMRM to MitoTracker fluorescence with background correction.

525 **Western blot**

526 Protein extracts from wt and *CG1603* mutant larvae tissues (48h after egg laying) were prepared using the
527 RIPA buffer (MilliporeSigma) with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific), 5 mM NaF
528 (MilliporeSigma) and 1 mM Na3VO4 (MilliporeSigma). Western blot was performed using a XCell
529 SureLock™ Mini-Cell and XCell II™ Blot Module (Thermo Fisher Scientific). Samples were
530 electrophoresed under a reducing condition on NuPAGE™ 4 to 12% Bis-Tris Mini Protein Gels (Thermo
531 Fisher Scientific). Proteins on the gel were transferred to a Polyvinylidene Difluoride (PVDF) membrane
532 (Thermo Fisher Scientific). The membranes were blocked with 5% BSA or non-fat milk (MilliporeSigma) in
533 TBST (Tris buffered saline with 0.1% Tween-20). After a serial of washes and incubations with primary
534 antibodies, TBST and secondary antibodies, the immunoreactivity was visualized using SuperSignal West
535 Dura Chemiluminescent Substrate (Thermo Fisher Scientific) and Amersham ImageQuant 800 system
536 (Cytiva). The antibodies used were: Mouse anti-Actin antibody (C4, MilliporeSigma), Mouse anti-ATP5A
537 antibody (15H4C4, abcam), Mouse anti-ND30 antibody (17D95, abcam), rabbit anti-TFAM antibody (Liu
538 *et al.*, 2022), rabbit anti-HSP60 antibody (#4870, Cell Signaling), Anti-rabbit IgG, HRP-linked Antibody
539 (#7074, Cell Signaling) and Anti-mouse IgG, HRP-linked Antibody (#7076, Cell Signaling).

540 **Blue native PAGE and in-gel activity assays of ETC complexes**

541 Mitochondria from fly larvae were isolated by homogenization and differential centrifugation following
542 previous protocol (Chen et al., 2015). Solubilized protein samples from isolated mitochondria were
543 prepared with NativePAGE™ Sample Prep Kit (Thermo Fisher Scientific) and the concentrations were
544 determined by Pierce BCA protein assay (Thermo Fisher Scientific). Blue native PAGE was performed
545 using NativePAGE™ 4–16% Bis-Tris gels and NativePAGE™ Running Buffer Kit (Thermo Fisher
546 Scientific) according to the manufacturer's protocol. 60 µg proteins each sample were used. For in-gel
547 activity assays, gels were incubated with one of the following solutions: Complex I buffer (5mM Tris-HCl
548 pH 7.4; 0.1mg/ml NADH; 2.5mg/ml Nitro Blue Tetrazolium), Complex II buffer (5mM Tris-HCl pH 7.4; 20mM
549 sodium succinate; 0.2mM phenazine methasulfate; 2.5mg/ml Nitro Blue Tetrazolium) or Complex IV buffer
550 (50mM sodium phosphate pH 7.2; 0.05% 3,3'-diaminobenzidine tetrahydrochloride, 50µM horse heart
551 cytochrome c) at room temperature for hours, stopped by fixation with 50% methanol and 10% acetic acid
552 for 30 min and washed with 10% acetic acid. All chemicals from MilliporeSigma.

553 **Quantitative real-time PCR**

554 Total genomic DNAs or RNAs were isolated using the DNeasy Blood & Tissue Kit (Qiagen) and RNeasy
555 Mini Kit (Qiagen), respectively, following the manufacturers' instructions. cDNAs were synthesized by the
556 SuperScript VILO cDNA Synthesis Kit (Thermo Fisher Scientific). Real-time PCRs were performed in
557 triplicate using the PowerTrack SYBR Green Master Mix (Thermo Fisher Scientific), MicroAmp™Optical
558 96-Well Reaction Plate with Barcode (Thermo Fisher Scientific), and QuantStudio™ 3 Real-Time PCR
559 System (Thermo Fisher Scientific). Primers for amplifying mtDNA and nuclear DNA (nuDNA), as well as
560 for measuring gene expression levels are listed in *Supplementary file 8*. The relative mtDNA levels of fly
561 larvae or eye discs were measured in three biological replicates for each group using total DNAs extracted

562 from twenty larvae or eye discs. The relative mRNA levels of ETC genes were measured in three biological
563 replicates for each group using total RNAs extracted from twenty larvae.

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

568 **Statistical analysis**
569 Two-tailed Student's t-test was used for statistical analysis. Difference was considered statistically
570 significant when $P<0.05$. Results are represented as mean ± SD of the number of determinations.
571

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

References

Acosta-Alvear, D., Zhou, Y., Blais, A., Tsikitis, M., Lents, N.H., Arias, C., Lennon, C.J., Kluger, Y., and Dynlacht, B.D. (2007). XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks. Mol Cell *27*, 53-66. 10.1016/j.molcel.2007.06.011.

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.

Deblois, G., and Giguere, V. (2011). Functional and physiological genomics of estrogen-related receptors (ERRs) in health and disease. Biochim Biophys Acta *1812*, 1032-1040. 10.1016/j.bbadis.2010.12.009.

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.

632 Kurland, C.G., and Andersson, S.G. (2000). Origin and evolution of the mitochondrial proteome. Microbiol Mol Biol
633 Rev *64*, 786-820. 10.1128/MMBR.64.4.786-820.2000.
634 Laws, K.M., and Drummond-Barbosa, D. (2015). Genetic Mosaic Analysis of Stem Cell Lineages in the Drosophila
635 Ovary. Methods Mol Biol *1328*, 57-72. 10.1007/978-1-4939-2851-4_4.
636 Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics
637 *25*, 1754-1760. 10.1093/bioinformatics/btp324.
638 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome
639 Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-
640 2079. 10.1093/bioinformatics/btp352.
641 Liang, X., Shan, S., Pan, L., Zhao, J., Ranjan, A., Wang, F., Zhang, Z., Huang, Y., Feng, H., Wei, D., et al. (2016).
642 Structural basis of H2A.Z recognition by SRCAP chromatin-remodeling subunit YL1. Nat Struct Mol Biol *23*, 317-323.
643 10.1038/nsmb.3190.
644 Liao, T.S., Call, G.B., Guptan, P., Cespedes, A., Marshall, J., Yackle, K., Owusu-Ansah, E., Mandal, S., Fang, Q.A.,
645 Goodstein, G.L., et al. (2006). An efficient genetic screen in Drosophila to identify nuclear-encoded genes with
646 mitochondrial function. Genetics *174*, 525-533. 10.1534/genetics.106.061705.
647 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).
648 The PPR domain of mitochondrial RNA polymerase is an exoribonuclease required for mtDNA replication in
649 Drosophila melanogaster. Nat Cell Biol *24*, 757-765. 10.1038/s41556-022-00887-y.
650 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data
651 with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.
652 Lowden, C., Boulet, A., Boehler, N.A., Seecharran, S., Rios Garcia, J., Lowe, N.J., Liu, J., Ong, J.L.K., Wang, W.,
653 Ma, L., et al. (2021). Homeostatic control of nuclear-encoded mitochondrial gene expression by the histone variant
654 H2A.Z is essential for neuronal survival. Cell Rep *36*, 109704. 10.1016/j.celrep.2021.109704.
655 Maheshwari, S., Wang, J., and Barbash, D.A. (2008). Recurrent positive selection of the Drosophila hybrid
656 incompatibility gene Hmr. Mol Biol Evol *25*, 2421-2430. 10.1093/molbev/msn190.
657 Mandal, S., Guptan, P., Owusu-Ansah, E., and Banerjee, U. (2005). Mitochondrial regulation of cell cycle progression
658 during development as revealed by the tenured mutation in Drosophila. Dev Cell *9*, 843-854.
659 10.1016/j.devcel.2005.11.006.
660 Owusu-Ansah, E., Yavari, A., Mandal, S., and Banerjee, U. (2008). Distinct mitochondrial retrograde signals control
661 the G1-S cell cycle checkpoint. Nature genetics *40*, 356-361. 10.1038/ng.2007.50.
662 Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet *10*, 669-680.
663 10.1038/nrg2641.
664 Perez-Torrado, R., Yamada, D., and Defossez, P.A. (2006). Born to bind: the BTB protein-protein interaction domain.
665 Bioessays *28*, 1194-1202. 10.1002/bies.20500.
666 Persikov, A.V., Wetzel, J.L., Rowland, E.F., Oakes, B.L., Xu, D.J., Singh, M., and Noyes, M.B. (2015). A systematic
667 survey of the Cys2His2 zinc finger DNA-binding landscape. Nucleic Acids Res *43*, 1965-1984. 10.1093/nar/gku1395.
668 Prober, D.A., and Edgar, B.A. (2000). Ras1 promotes cellular growth in the Drosophila wing. Cell *100*, 435-446.
669 10.1016/s0092-8674(00)80679-0.
670 Putri, G.H., Anders, S., Pyl, P.T., Pimanda, J.E., and Zanini, F. (2022). Analysing high-throughput sequencing data
671 in Python with HTSeq 2.0. Bioinformatics. 10.1093/bioinformatics/btac166.
672 Rath, S., Sharma, R., Gupta, R., Ast, T., Chan, C., Durham, T.J., Goodman, R.P., Grabarek, Z., Haas, M.E., Hung,
673 W.H.W., et al. (2021). MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and
674 pathway annotations. Nucleic Acids Res *49*, D1541-D1547. 10.1093/nar/gkaa1011.
675 Rouyer, F., Rachidi, M., Pikielny, C., and Rosbash, M. (1997). A new gene encoding a putative transcription factor
676 regulated by the Drosophila circadian clock. EMBO J *16*, 3944-3954. 10.1093/emboj/16.13.3944.
677 Scarpulla, R.C. (2008). Transcriptional paradigms in mammalian mitochondrial biogenesis and function. Physiol Rev
678 *88*, 611-638. 10.1152/physrev.00025.2007.
679 Scarpulla, R.C., Vega, R.B., and Kelly, D.P. (2012). Transcriptional integration of mitochondrial biogenesis. Trends
680 Endocrinol Metab *23*, 459-466. 10.1016/j.tem.2012.06.006.
681 Schuldiner, O., Berdnik, D., Levy, J.M., Wu, J.S., Luginbuhl, D., Gontang, A.C., and Luo, L. (2008). piggyBac-based
682 mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. Dev Cell *14*,
683 227-238. 10.1016/j.devcel.2007.11.001.
684 Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., and Bork, P. (2000). SMART: a web-based tool for the study of
685 genetically mobile domains. Nucleic Acids Res *28*, 231-234. 10.1093/nar/28.1.231.

686 Shen, K., Pender, C.L., Bar-Ziv, R., Zhang, H., Wickham, K., Willey, E., Durieux, J., Ahmad, Q., and Dillin, A. (2022).
687 Mitochondria as Cellular and Organismal Signaling Hubs. Annu Rev Cell Dev Biol *38*, 179-218. 10.1146/annurev-
688 cellbio-120420-015303.
689 Sobocinska, J., Molenda, S., Machnik, M., and Oleksiewicz, U. (2021). KRAB-ZFP Transcriptional Regulators Acting
690 as Oncogenes and Tumor Suppressors: An Overview. Int J Mol Sci *22*. 10.3390/ijms22042212.
691 Spinelli, J.B., and Haigis, M.C. (2018). The multifaceted contributions of mitochondria to cellular metabolism. Nat Cell
692 Biol *20*, 745-754. 10.1038/s41556-018-0124-1.
693 Taylor, R.W., and Turnbull, D.M. (2005). Mitochondrial DNA mutations in human disease. Nat Rev Genet *6*, 389-402.
694 nrg1606 [pii]

695 10.1038/nrg1606.
696 Thomas-Chollier, M., Darbo, E., Herrmann, C., Defrance, M., Thieffry, D., and van Helden, J. (2012). A complete
697 workflow for the analysis of full-size ChIP-seq (and similar) data sets using peak-motifs. Nat Protoc *7*, 1551-1568.
698 10.1038/nprot.2012.088.
699 Verheul, T.C.J., van Hijfte, L., Perenthaler, E., and Barakat, T.S. (2020). The Why of YY1: Mechanisms of
700 Transcriptional Regulation by Yin Yang 1. Front Cell Dev Biol *8*, 592164. 10.3389/fcell.2020.592164.
701 Wang, Z.H., Liu, Y., Chaitankar, V., Pirooznia, M., and Xu, H. (2019). Electron transport chain biogenesis activated
702 by a JNK-insulin-Myc relay primes mitochondrial inheritance in Drosophila. Elife *8*. 10.7554/eLife.49309.
703 Wang, Z.H., Zhao, W., Combs, C.A., Zhang, F., Knutson, J.R., Lilly, M.A., and Xu, H. (2023). Mechanical stimulation
704 from the surrounding tissue activates mitochondrial energy metabolism in Drosophila differentiating germ cells. Dev
705 Cell. 10.1016/j.devcel.2023.08.007.
706 Wiedemann, N., and Pfanner, N. (2017). Mitochondrial Machineries for Protein Import and Assembly. Annu Rev
707 Biochem *86*, 685-714. 10.1146/annurev-biochem-060815-014352.
708 Wolfe, S.A., Nekludova, L., and Pabo, C.O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev
709 Biophys Biomol Struct *29*, 183-212. 10.1146/annurev.biophys.29.1.183.
710 Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T.,
711 Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin
712 remodelling. Nature *442*, 86-90. 10.1038/nature04815.
713 Xu, H., DeLuca, S.Z., and O'Farrell, P.H. (2008). Manipulating the metazoan mitochondrial genome with targeted
714 restriction enzymes. Science *321*, 575-577. 321/5888/575 [pii]

715 10.1126/science.1160226.
716 Xu, M., Yao, J., Shi, Y., Yi, H., Zhao, W., Lin, X., and Yang, Z. (2021). The SRCAP chromatin remodeling complex
717 promotes oxidative metabolism during prenatal heart development. Development *148*. 10.1242/dev.199026.
718 Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes
719 among gene clusters. OMICS *16*, 284-287. 10.1089/omi.2011.0118.
720 Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation,
721 comparison and visualization. Bioinformatics *31*, 2382-2383. 10.1093/bioinformatics/btv145.
722 Zhang, F., Pirooznia, M., and Xu, H. (2020). Mitochondria regulate intestinal stem cell proliferation and epithelial
723 homeostasis through FOXO. Mol Biol Cell *31*, 1538-1549. 10.1091/mbc.E19-10-0560.
724 Zhang, F., Qi, Y., Zhou, K., Zhang, G., Linask, K., and Xu, H. (2015). The cAMP phosphodiesterase Prune localizes
725 to the mitochondrial matrix and promotes mtDNA replication by stabilizing TFAM. EMBO Rep *16*, 520-527.
726 10.15252/embr.201439636.
727 Zhang, Y., Chen, Y., Gucek, M., and Xu, H. (2016). The mitochondrial outer membrane protein MDI promotes local
728 protein synthesis and mtDNA replication. EMBO J *35*, 1045-1057. 10.15252/embj.201592994.
729 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown,
730 M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137. 10.1186/gb-
731 2008-9-9-r137.
732 Zhang, Y., Wang, Z.H., Liu, Y., Chen, Y., Sun, N., Gucek, M., Zhang, F., and Xu, H. (2019). PINK1 Inhibits Local
733 Protein Synthesis to Limit Transmission of Deleterious Mitochondrial DNA Mutations. Mol Cell *73*, 1127-1137 e1125.
734 10.1016/j.molcel.2019.01.013.
735 Zhou, W., Sherwood, B., and Ji, H. (2016). Computational Prediction of the Global Functional Genomic Landscape:
736 Applications, Methods, and Challenges. Hum Hered *81*, 88-105. 10.1159/000450827.

737

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 (Mito-EBR). 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.

769

770 **Figure 2. Regulatory network of mitochondrial biogenesis.**

771 (A) The potential transcriptional regulatory network of nuclear-encoded mitochondrial genes. (B) Bar
772 graphs illustrating the promoter binding profiles of the 49 synergistic enhancer TFs within different groups
773 of genes (nuclear-encoded mitochondrial genes, Mito-EBR genes, 49 synergistic enhancer TFs and ETC
774 genes). Total gene number of the group (orange) and the number bound by CG1603 (cyan) were highlighted.

775

776 **The following figure supplement is available for figure 2:**

777 Figure 2—figure supplement. Bar graphs summarizing the promoter binding profiles of the 49 synergistic
778 enhancer TFs within subgroups of genes involved in different processes of ETC biogenesis. Total gene
779 number of the subgroup (orange) and the number bound by CG1603 (cyan) were highlighted.

780 *Supplementary file 2. Vertices, edges and vertex.sort analysis information of the potential transcriptional*
781 *regulatory network of nuclear-encoded mitochondrial genes.*

782 *Supplementary file 3. Gene binding profiles of 49 synergistic enhancer TFs, including TF-target matrix,*
783 *TF binding profile summary, and list of nuclear-encoded mitochondrial genes with subgroup information*
784 *as well as the counts of 49 TFs binding to each promoter.*

785

786 **Figure 3. CG1603 promotes ETC gene expression and mitochondrial biogenesis.**

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

794

795 **The following figure supplement is available for figure 3:**

796 Figure 3—figure supplement. Number of EC clone cells in control RNAi (Ctrl) and *CG1603* RNAi
797 (*CG1603*-i) midguts. n= 8, error bar: SD. ***: $p<0.001$.

798

799

800 **Figure 4. *CG1603* gene model, product, mutant and the genomic DNA transgene.**

801 (A) Schematic representation of *CG1603* genomic locus, showing the *CG1603* transcript (5' and 3'UTR in
802 black bar and four exons in white), its protein product (586 amino acids in length, and characterized by
803 seven C2H2-ZF and two MADF domains), the *CG1603*^{PBac} mutant allele (with a piggyBac insertion in the
804 second intron, which is marked by fluorescent DsRed driven by an eye-specific *3xP3* promoter and flanked
805 by stop codons in all three reading frames terminating translation through downstream), and the genomic
806 region (in square brackets, from 955 bp upstream of the *CG1603* 5'UTR to 656 bp downstream of *CG1603*
807 3'UTR) used for *P[CG1603*^{gDNA}*]* transgene. (B) Predicted 3D structure of the CG1603 protein by
808 AlphaFold. Labels indicate the N-and C-terminus, as well as the specific protein domains along with their
809 predicted isoelectric point (pI). (C) Images of *CG1603*^{PBac} / *CyO, Act-GFP* and homozygous *CG1603*^{PBac}
810 larvae cultured together at 25°C, day 4 after egg laying. Green: GFP; Red: DsRed. Scale bars: 1 mm. (D)
811 Relative mtDNA levels in *CG1603*^{PBac} mutant larvae to wild type (wt) control. n= 3, error bar: SD. ***:
812 p<0.001. (E) Western blots of mitochondrial proteins in *CG1603*^{PBac} mutant larvae to wild type (wt) control.
813 (F) *P[CG1603*^{gDNA}*]* restored viability of *CG1603*^{PBac}flies. The number of progenies for each genotype is
814 listed.
815
816 **The following figure supplement is available for figure 4:**
817 Figure 4—figure supplement A. Blue native PAGE and in-gel activity analyses of ETC Complex I, II & IV
818 isolated from wild type (wt) and *CG1603*^{PBac} mutant. An equal amount of protein was used for each sample.
819 Figure 4—figure supplement B. Summary of adult viability phenotypes of combinations of *CG1603*^{PBac}
820 mutant, *P[CG1603*^{gDNA}*]* transgene and deficiency chromosomes.
821 Figure 4—figure supplement C. Schematic map of deficiency chromosomes spanning *CG1603* genomic
822 region.
823
824 **Figure 5. Clonal analyses confirmed CG1603's role in mitochondrial biogenesis and activity.**
825 (A) Representative images of *CG1603*^{PBac} mutant germline (top and middle panel) and follicle (bottom
826 panel) clones in late-stage egg chambers of adult ovaries with endogenously expressed TFAM-mNG
827 visualized in green. Homozygous mutant clones were marked by the absence of RFP and compared with
828 either flanking RFP-positive cysts (germline) or homozygous wt twin (follicle). White dash lines aided in
829 illustrating clones. The wt (RFP+/RFP+) follicle clone showed markedly higher RFP intensity than the
830 heterozygous (RFP+/RFP-) cells, also see *Figure 5—figure supplement B*. Red: nls-RFP; Blue: DAPI. Scale
831 bars: 10 μm. (B) Quantification of the relative TFAM-mNG level in the homozygous *FRT42D* control and

832 *CG1603*^{PBac} mutant germline clones in early-stage egg chamber to their anterior flanking RFP-positive cysts
833 within the same ovariole. also see *Figure 5—figure supplement A.* n=10 for each group, error bar: SD. ***:
834 p<0.001. (C-E) Quantification of the relative TFAM-mNG level (C), the relative mtDNA nucleoids number
835 (D) and the relative clone cell number (E) in the homozygous *FRT42D* control and *CG1603*^{PBac} mutant
836 follicle clones to their wt twins. n=10 for each group, error bar: SD. ***: p<0.001. (F) TMRM / MitoTracker
837 Green (MT) ratiometric live imaging and quantification of ovarioles containing homozygous *CG1603*^{PBac}
838 mutant germline clones (highlighted by white dash lines). Notably, in contrast to flanking control cysts
839 (highlighted by white lines), $\Delta\psi_m$ was almost absent in mutant clones. Please note that compared to TMRM,
840 nls-RFP signal was too low to be detected in ratiometric imaging. Nonetheless, the nls-RFP was readily
841 detected in control cysts, but not in homozygous *CG1603*^{PBac} clones, *via* visual observation, as depicted in
842 *Figure 5A* and *Figure 5—figure supplement*. A twin pair of follicle clones in the same egg chamber were
843 also highlighted (cyan line for control and cyan dash line for homozygous *CG1603*^{PBac} mutant). The MT
844 intensity was reduced in both the germline and follicle *CG1603*^{PBac} clones, compared to germ cells in
845 adjacent egg chambers and follicle cells in the same egg chamber, respectively. Quantification with
846 background correction for MT intensity in germline clones is shown in *Figure 5—figure supplement C.*
847 Blue: Hoechst. Scale bars: 10 μm. n= 8, error bar: SD. ***: p<0.001.
848
849 **The following figure supplement is available for figure 5:**
850 Figure 5—figure supplement A. Representative images of homozygous *FRT42D* control and *CG1603*^{PBac}
851 mutant germline clones in early-stage egg chambers of adult ovaries with endogenously expressed TFAM-
852 mNG visualized in green. Homozygous mutant clones were marked by the absence of RFP and compared
853 with flanking RFP-positive cysts. Red: RFP; Blue: DAPI. Scale bars: 10 μm.
854 Figure 5—figure supplement B. Representative images of homozygous *FRT42D* control follicle cell clone
855 (RFP-/RFP-) and its wt twin (RFP+/RFP+) with endogenously expressed TFAM-mNG visualized in green.
856 Red: nls-RFP; Blue: DAPI. Scale bars: 10 μm.
857 Figure 5—figure supplement C. The relative intensity of MitoTracker (MT) Green in *CG1603*^{PBac} mutant
858 germline clones to control (Ctrl). n= 8, error bar: SD. **: p<0.01. Related to *Figure 5F.*
859
860
861 **Figure 6. CG1603 localizes in the nucleus and is essential for regulating nuclear mitochondrial gene**
862 **expression.**

863 (A-B) Representative images showing nuclear localization of CG1603 protein in cultured S2 cell (A) and
864 adult ovary (B). Green: MitoTracker Green in S2 cell, and CG1603-mNG in tissues; Red: CG1603-mCH;
865 Blue & Magenta: Hoechst. Scale bars: 10 µm. (C) Representative images showing bindings of endogenously
866 expressed CG1603 proteins to less condensed euchromatin regions in the polytene chromosomes of salivary
867 gland. High intensity CG1603-mNG bands were visualized in green in upper panel as indicated by arrow,
868 and low intensity bands were visualized in white in lower panel as indicated by arrow heads with image
869 B&C adjusted. Magenta: Hoechst. Scale bars: 10 µm. (D) Density plot illustrating the distribution of
870 expression changes of the nuclear-encoded mitochondrial and non-mitochondrial genes in *CG1603*^{PBac}
871 mutant. (E) Graph illustrating the overlap between nuclear-encoded mitochondrial genes and differentially
872 expressed genes (DEGs) that down-regulated > 2-fold, as well as the distribution of the overlapped genes
873 in different mitochondrial function categories. (F) Relative mRNA levels of several ETC biogenesis-related
874 genes in *CG1603*^{PBac} mutant larvae to control, measured by real-time PCR. n= 3, error bar: SD. (G) Gene
875 Ontology (GO) enrichment analyses of DEGs that down-regulated > 2-fold. The top 10 enriched biological
876 processes are shown.
877
878
879 **The following figure supplement is available for figure 6:**
880 *Supplementary file 4. List of nuclear-encoded mitochondrial genes with symbols, IDs, subgroup*
881 *information, and RNA-seq status.*
882 *Supplementary file 5. a. List of differentially expressed nuclear-encoded genes in CG1603* ^{PBac} *mutant*
883 *flies compared to controls. b. Summary of nuclear-encoded ETC genes that down-regulated more than*
884 *two-fold in CG1603*^{PBac} *mutant.*
885
886 **Figure 7. ChIP analysis identified nuclear mitochondrial genes that may be directly regulated by**
887 **CG1603.**
888 (A) CG1603 ChIP peaks over all chromosomes. (B) Genomic distribution of CG1603 peaks. (C) Average
889 profile of CG1603 peaks binding to transcription start site (TSS) regions. (D) Representative binding motif
890 discovered with CG1603 ChIP peaks. (E) Summary of the number of nuclear-encoded mitochondrial and
891 non-mitochondrial mRNA coding genes bound by CG1603, and the overlapping down-regulated
892 differentially expressed genes (DEGs) in each group. (F) Scatterplot illustrating the signalValue of CG1603
893 ChIP peaks (y-axis) and log2 fold change in expression of DEGs between *CG1603*^{PBac} mutant and control

894 (x-axis). MitoES: the genes belong to the categories that are clearly essential to mitochondrial ETC
895 biogenesis and maintenance, including ETC subunits and assembly factors, mtDNA replication and
896 transcription, mitochondrial RNA metabolism and translation, as well as mitochondrial protein import and
897 membrane insertion machinery.
898
899 The following figure supplement is available for figure 7:
900 *Supplementary file 6. List of CG1603 peaks from modERN ChIP-seq data, including gene annotation,*
901 *group information, combined with RNAseq analysis result of CG1603* ^{PBac} *mutant flies.*
902 *Supplementary file 7. CG1603 binding motifs discovered by RSAT peak-motifs.*
903
904 **Figure 8. YL-1 is an upstream regulator of CG1603.**
905 (A) Schematic graph illustrating the CG1603 upstream and downstream (co-)TFs involved in regulating
906 mitochondrial ETC biogenesis, inferred from ChIP-seq, RNAseq and genetics data. (B) Representative
907 images of control RNAi (Ctrl), *CG1603* RNAi (*CG1603*-i) and *YL-1* RNAi (*YL-1*-i) midgut EC clones with
908 endogenously expressed CG1603-mNG visualized in green or white. Clones were labeled by mCherry red
909 and compared with wt neighbors. White dash lines aided in illustrating clones. Blue: Hoechst. Scale bars:
910 10 μm. (C) Quantification of the relative CG1603-mNG level in the EC clones to their wt neighbors. n=16
911 from 8 midguts for each group, error bar: SD. ***: p<0.001. (D) Representative eye image and Index-R
912 ratio (RNAi + mitoXhoI / RNAi-only) of adult flies with indicated genotypes. Three biological repeats were
913 performed for each group, error bar: SD. ***: p<0.001. (E-F) Representative images of *YL-1* RNAi (*YL-1*-
914 i) and *YL-1* RNAi + *CG1603* overexpression (*YL-1*-i + *CG1603*-OE) midgut EC clones with endogenously
915 expressed TFAM-GFP (E) or SDHA-mNG (F) visualized in green. Clones were labeled by mCherry red
916 and compared with wt neighbors. Blue: Hoechst. Scale bars: 10 μm. (G-J) Quantification of the relative
917 TFAM-GFP level (G), the relative mtDNA nucleoids number (H), the relative SDHA-mNG level (I) and
918 the relative mitochondrial volume (J) in the EC clones to their wt neighbors. n=16 from 8 midguts for each
919 group, error bar: SD. ***: p<0.001.
920
921 **The following figure supplement is available for figure 8:**
922 Figure 8—figure supplement A. Representative images of control RNAi (Ctrl), *CG1603* RNAi (*CG1603*-
923 i), *STAT92E* RNAi (*STAT92E*-i), *YL-1* RNAi (*YL-1*-i), *Myb* RNAi (*Myb* -i), *trem* RNAi (*trem*-i) and *E(bx)*

924 RNAi (*E(bx)*-i) midgut ISC/EB clones labeled by mCherry red, with endogenously expressed TFAM-GFP
925 visualized in green. White dash lines aided in illustrating clones. Scale bars: 10 μm.
926 Figure 8—figure supplement B. Quantification of the relative TFAM-GFP level in midgut ISC/EB clones
927 to their wt neighbors for different groups. n=10, error bar: SD. ***: $p<0.001$.
928 Figure 8—figure supplement C. Relative mtDNA levels in eye discs of different RNAi groups. n= 3, error
929 bar: SD. *: $P<0.05$, ***: $p<0.001$.

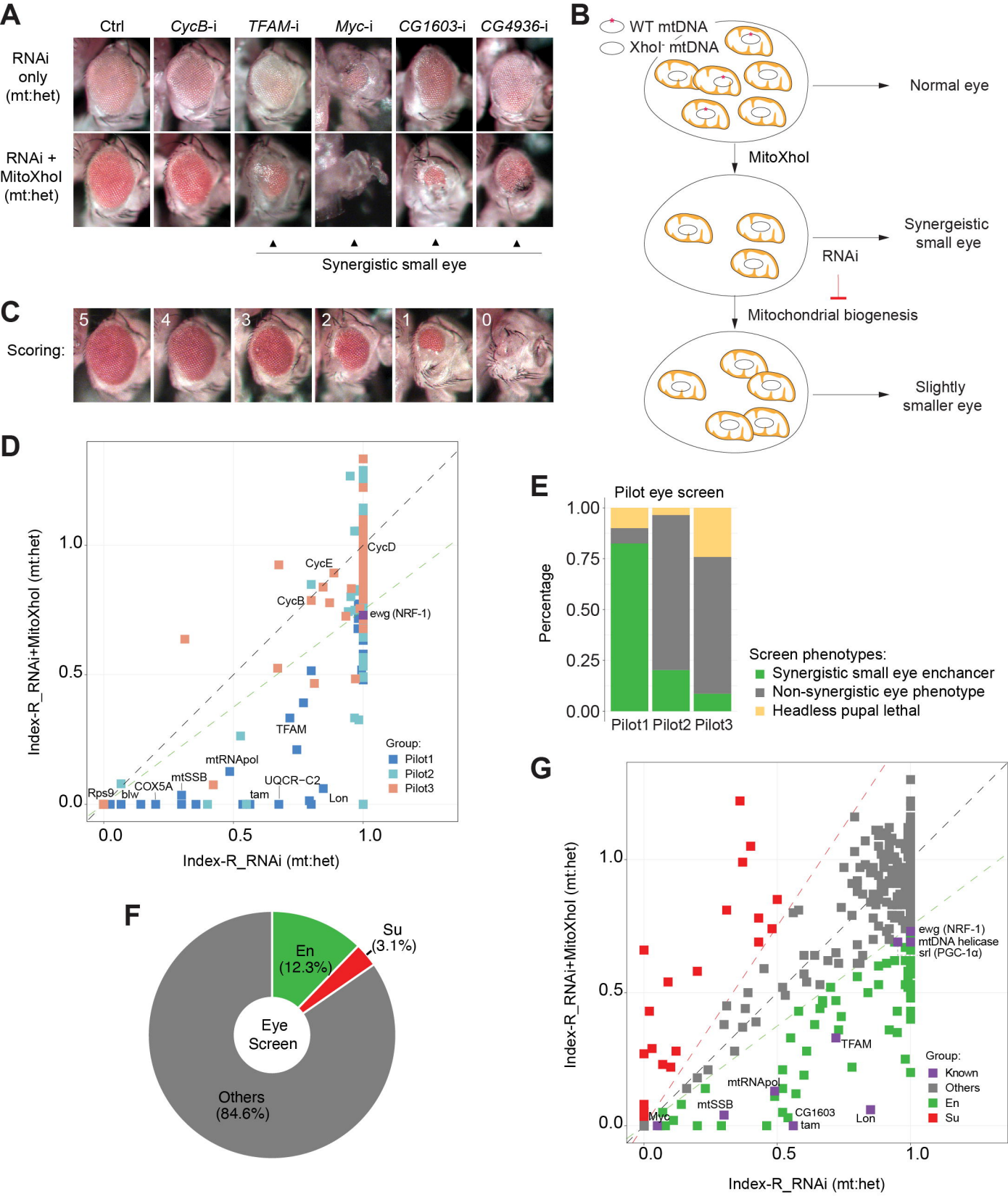

A
Ctrl
CycB-i
TFAM-i
Myc-i
CG1603-i
CG4936-i
RNAi only (mt:het)
RNAi + MitoXhoI (mt:het)
Synergistic small eye
B
WT mtDNA
XhoI⁻ mtDNA
Normal eye
MitoXhoI
Synergeistic small eye
RNAi
Mitochondrial biogenesis
Slightly smaller eye
C
Scoring:
5
4
3
2
1
0
D
Index-R_RNAi+MitoXhoI (mt:het)
Index-R_RNAi (mt:het)
CycD
CycE
CycB
ewg (NRF-1)
TFAM
mtRNApol
mtSSB
UQCR-C2
COX5A
Rps9
blw
tam
Lon
Group:
Pilot1
Pilot2
Pilot3
E
Pilot eye screen
Percentage
Pilot1 Pilot2 Pilot3
Screen phenotypes:
Synergistic small eye enchancer
Non-synergistic eye phenotype
Headless pupal lethal
F
En (12.3%)
Su (3.1%)
Eye Screen
Others (84.6%)
G
Index-R_RNAi+MitoXhoI (mt:het)
Index-R_RNAi (mt:het)
ewg (NRF-1)
mtDNA helicase
srl (PGC-1α)
TFAM
mtRNApol
mtSSB
Myc
CG1603
tam
Lon
Group:
Known
Others
En
Su

Figure 1

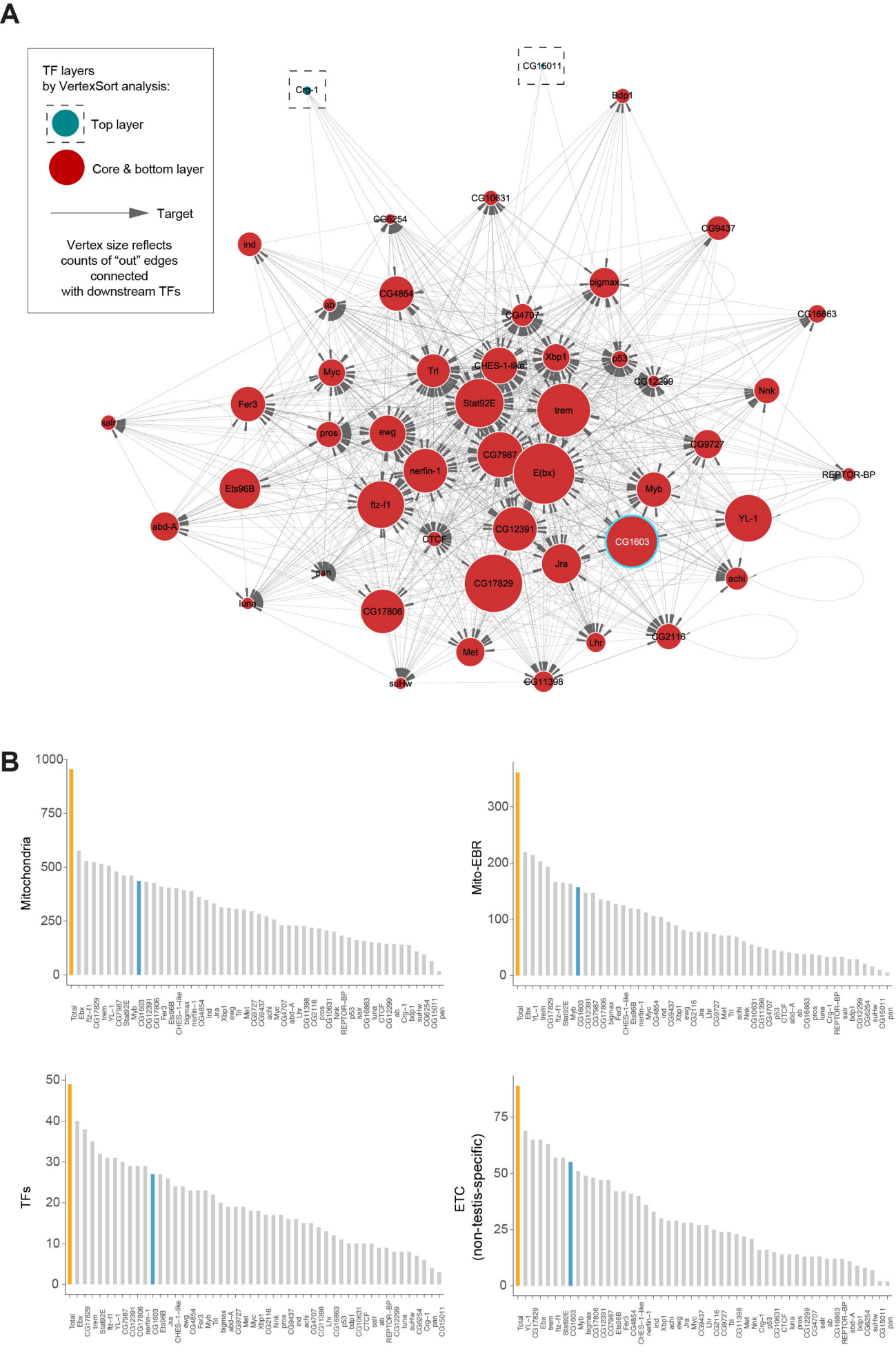
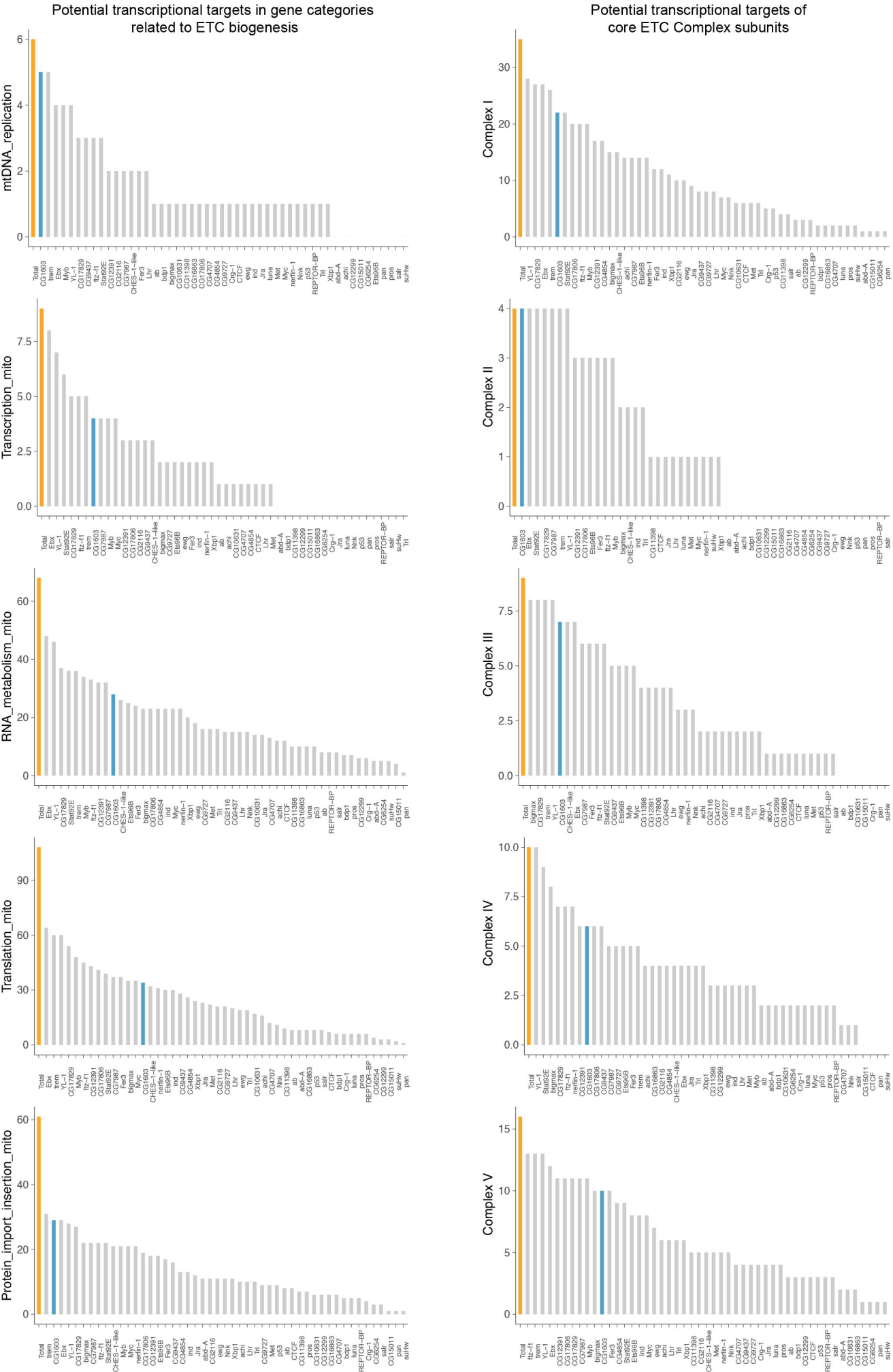

A
TF layers
by VertexSort analysis:
Top layer
Core & bottom layer
Target
Vertex size reflects
counts of "out" edges
connected
with downstream TFs
Crc-1
CG16011
Bdp1
CG10631
CG6254
CG9437
ind
bigmax
CG4854
ab
CG4707
CG16863
Xbp1
p53
Myc
Trl
CHES-1-like
CG12299
Nnk
Fer3
Stat92E
trem
salr
pros
ewg
CG9727
CG7987
nerfin-1
E(bx)
REPTOR-BP
Ets96B
Myb
ftz-f1
YL-1
abd-A
CG12391
CTCF
CG1603
pan
Jra
achi
CG17829
luna
CG17806
Lhr
CG2116
Met
suHw
CG11398
B
Mitochondria
Mito-EBR
TFs
ETC
(non-testis-specific)

Figure 2

Figure 2—figure supplement



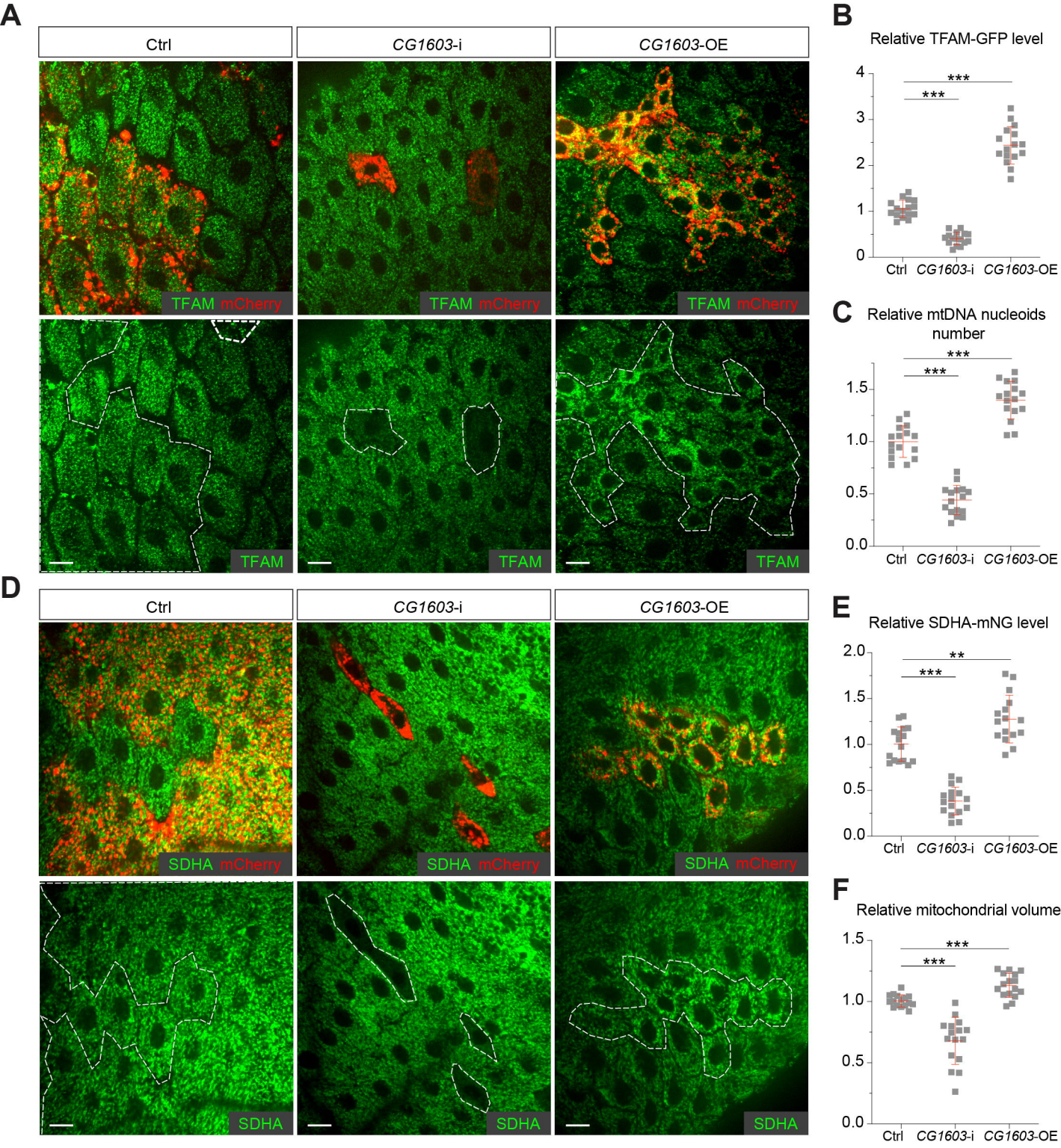
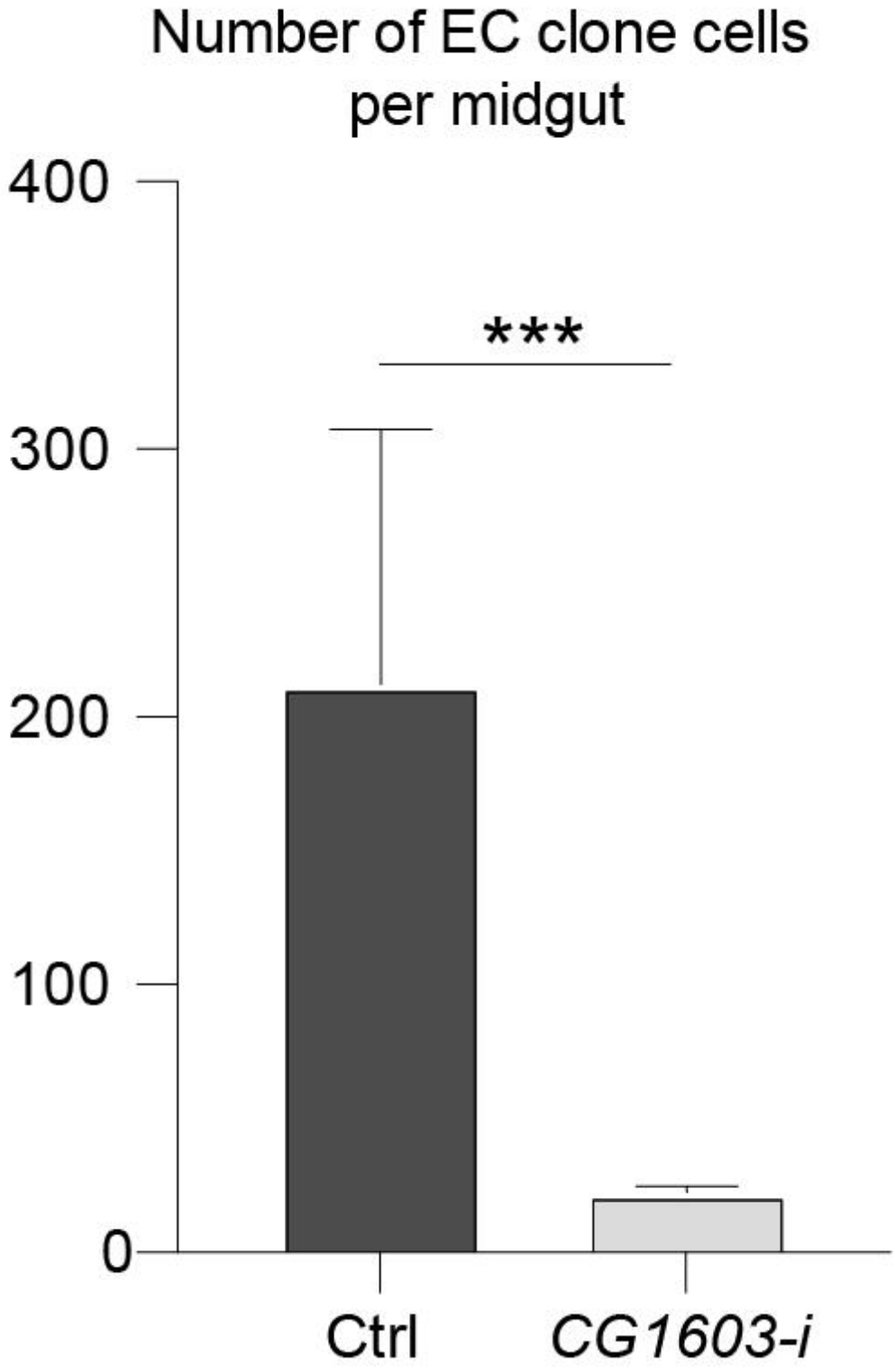

A
Ctrl
CG1603-i
CG1603-OE
TFAM mCherry
TFAM
B
Relative TFAM-GFP level
***
***
Ctrl
CG1603-i
CG1603-OE
C
Relative mtDNA nucleoids number
D
SDHA mCherry
SDHA
E
Relative SDHA-mNG level
**
F
Relative mitochondrial volume

Figure 3

Figure 3—figure supplement



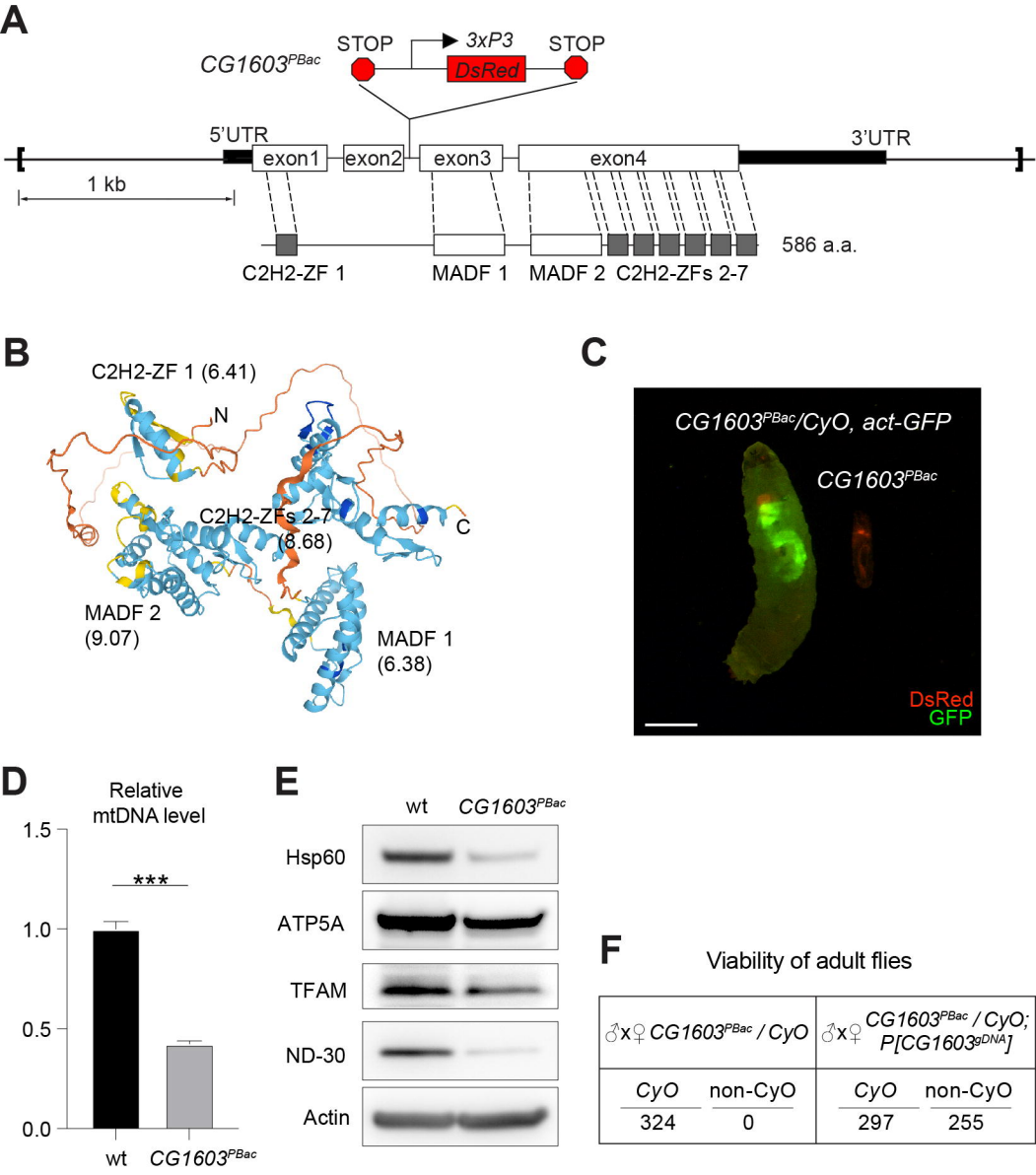
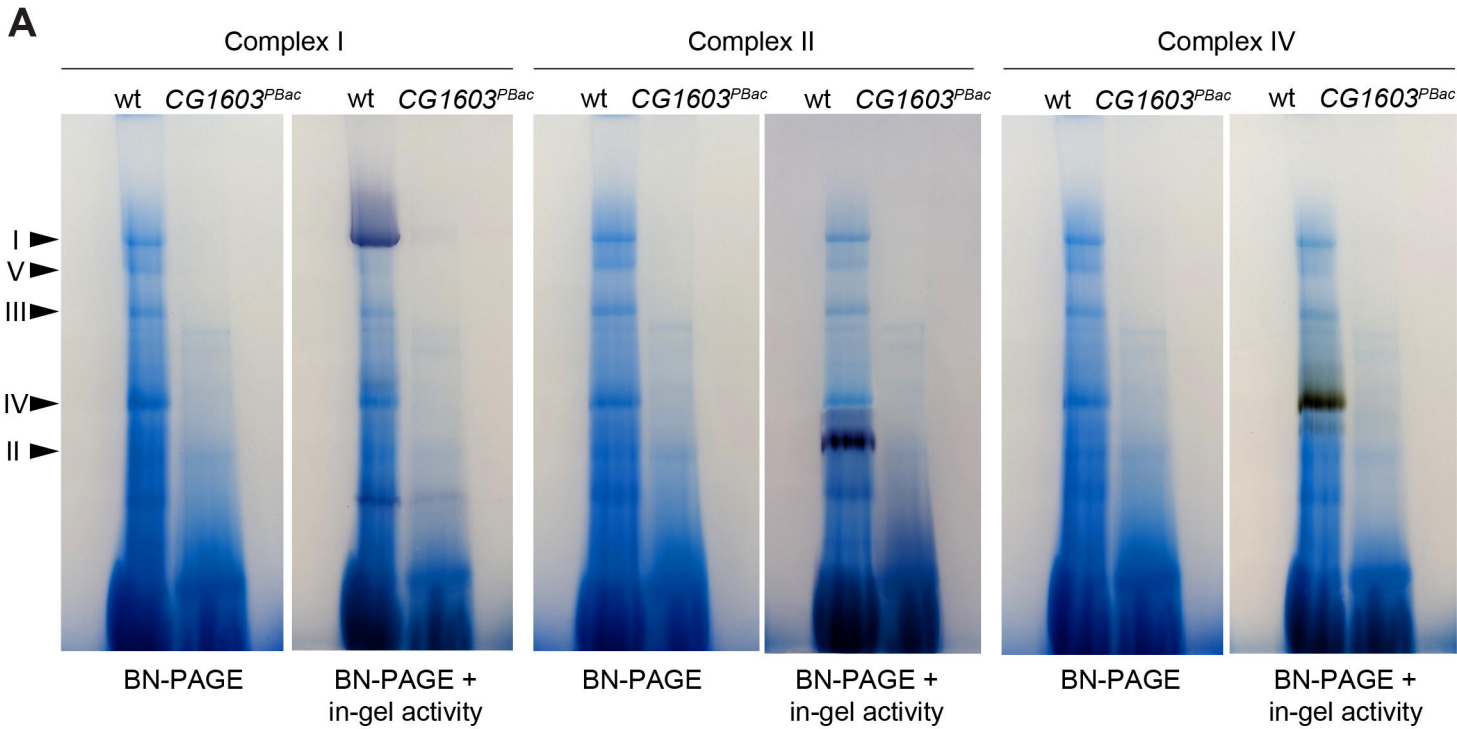

A
CG1603PBac
STOP
3xP3
DsRed
STOP
5'UTR
exon1
exon2
exon3
exon4
3'UTR
1 kb
586 a.a.
C2H2-ZF 1
MADF 1
MADF 2
C2H2-ZFs 2-7
B
C2H2-ZF 1 (6.41)
N
C2H2-ZFs 2-7
(8.68)
C
MADF 2
(9.07)
MADF 1
(6.38)
C
CG1603PBac/CyO, act-GFP
CG1603PBac
DsRed
GFP
D
Relative
mtDNA level
1.5
1.0
0.5
0.0
***
wt
CG1603PBac
E
wt
CG1603PBac
Hsp60
ATP5A
TFAM
ND-30
Actin
F
Viability of adult flies
♂x♀ CG1603PBac / CyO
♂x♀ CG1603PBac / CyO; P[CG1603gDNA]
CyO
non-CyO
324
0
CyO
non-CyO
297
255

Figure 4

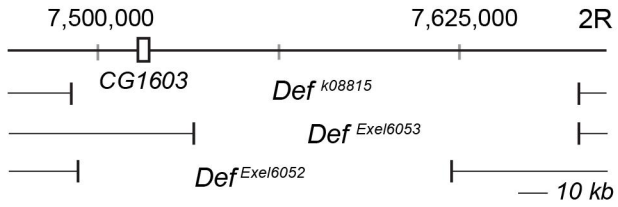
Figure 4—figure supplement



B

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	

C



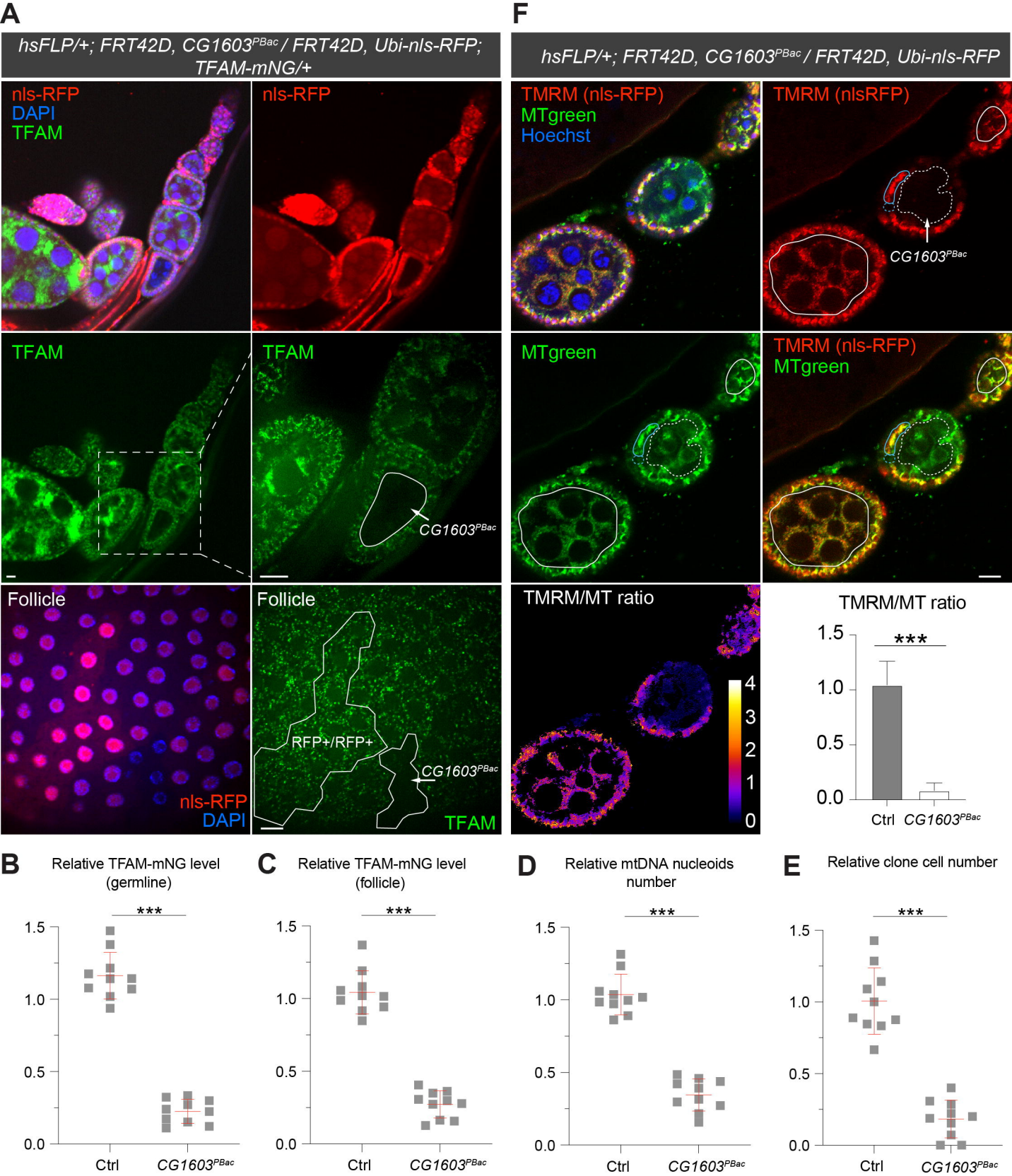
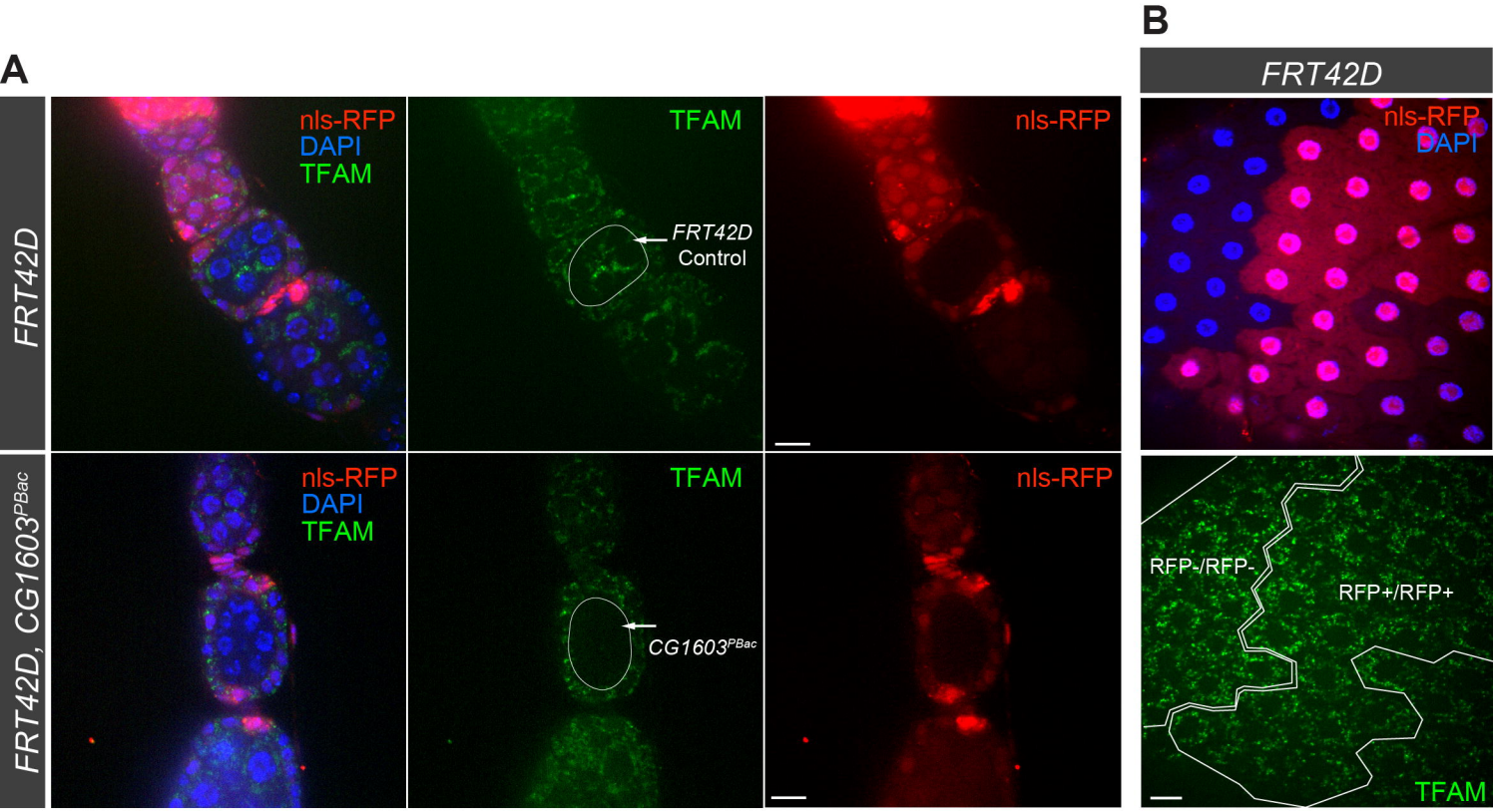
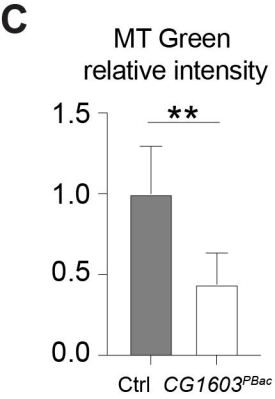

A
hsFLP/+; FRT42D, CG1603PBac / FRT42D, Ubi-nls-RFP; TFAM-mNG/+
nls-RFP
DAPI
TFAM
nls-RFP
TFAM
TFAM
CG1603PBac
Follicle
Follicle
nls-RFP
DAPI
RFP+/RFP+
CG1603PBac
TFAM
F
hsFLP/+; FRT42D, CG1603PBac / FRT42D, Ubi-nls-RFP
TMRM (nls-RFP)
MTgreen
Hoechst
TMRM (nlsRFP)
CG1603PBac
MTgreen
TMRM (nls-RFP)
MTgreen
TMRM/MT ratio
4
3
2
1
0
TMRM/MT ratio
1.5
1.0
0.5
0.0
***
Ctrl
CG1603PBac
B
Relative TFAM-mNG level (germline)
C
Relative TFAM-mNG level (follicle)
D
Relative mtDNA nucleoids number
E
Relative clone cell number

Figure 5

Figure 5—figure supplement





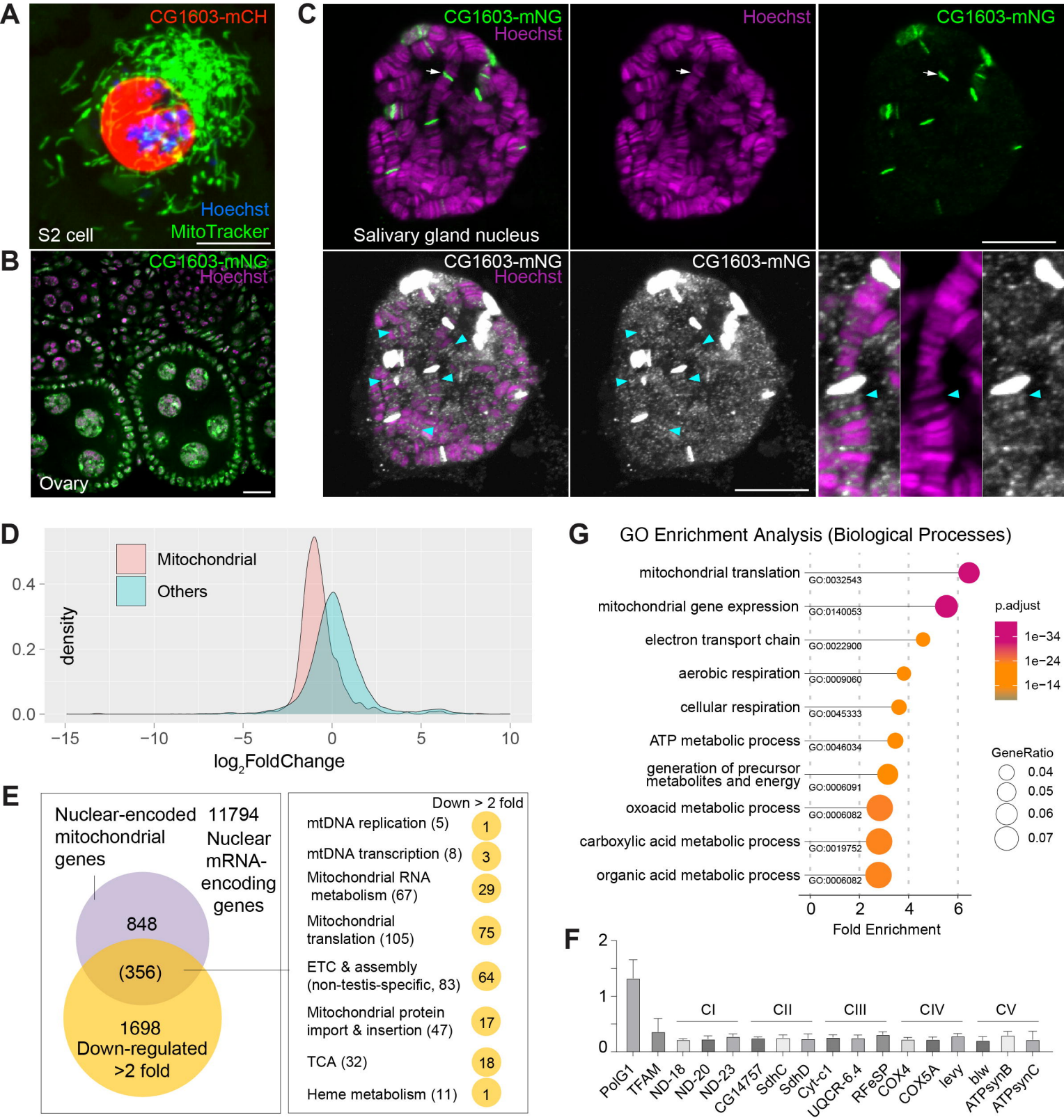
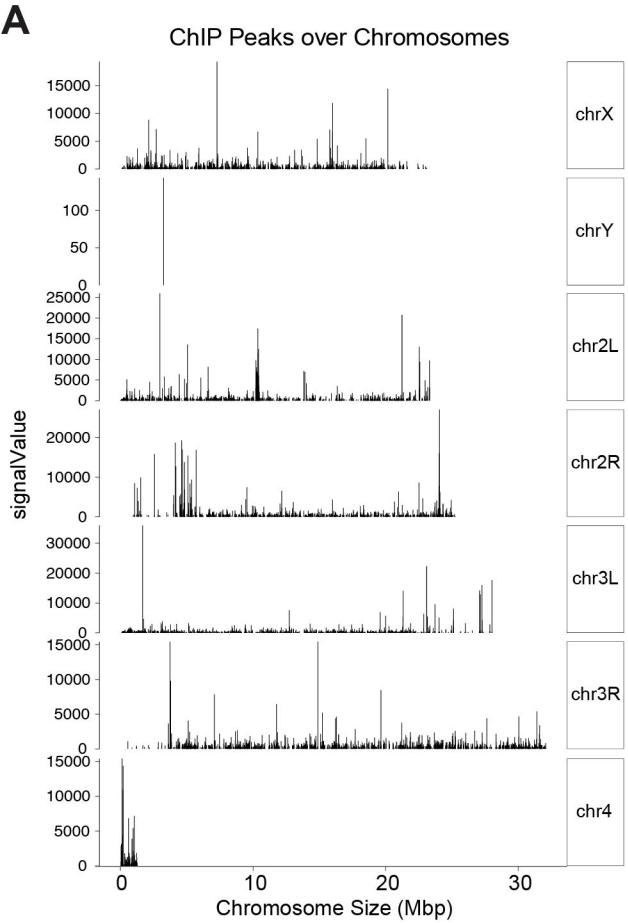
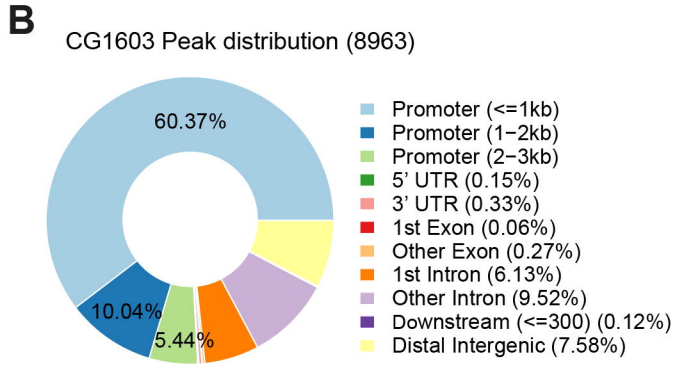
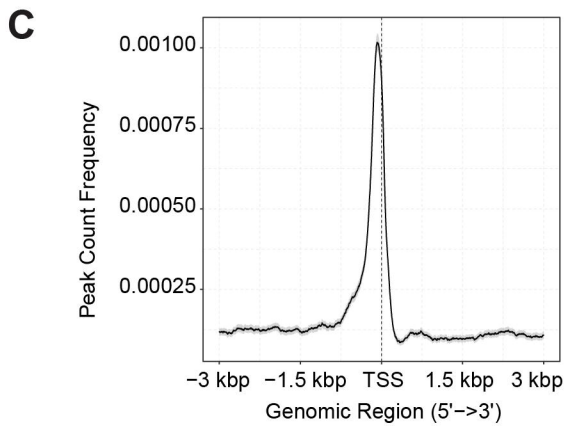
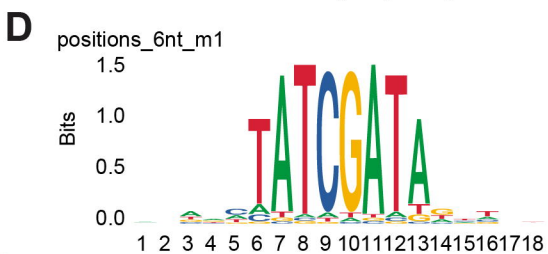

A
CG1603-mCH
Hoechst
MitoTracker
S2 cell
B
CG1603-mNG
Hoechst
Ovary
C
CG1603-mNG
Hoechst
Salivary gland nucleus
Hoechst
CG1603-mNG
CG1603-mNG
Hoechst
CG1603-mNG
D
Mitochondrial
Others
density
0.4
0.2
0.0
−15
−10
−5
0
5
10
log2FoldChange
E
Nuclear-encoded mitochondrial genes
11794 Nuclear mRNA-encoding genes
848
(356)
1698 Down-regulated >2 fold
Down > 2 fold
mtDNA replication (5) 1
mtDNA transcription (8) 3
Mitochondrial RNA metabolism (67) 29
Mitochondrial translation (105) 75
ETC & assembly (non-testis-specific, 83) 64
Mitochondrial protein import & insertion (47) 17
TCA (32) 18
Heme metabolism (11) 1
G
GO Enrichment Analysis (Biological Processes)
mitochondrial translation GO:0032543
mitochondrial gene expression GO:0140053
electron transport chain GO:0022900
aerobic respiration GO:0009060
cellular respiration GO:0045333
ATP metabolic process GO:0046034
generation of precursor metabolites and energy GO:0006091
oxoacid metabolic process GO:0006082
carboxylic acid metabolic process GO:0019752
organic acid metabolic process GO:0006082
p.adjust
1e−34
1e−24
1e−14
GeneRatio
0.04
0.05
0.06
0.07
0
2
4
6
Fold Enrichment
F
2
1
0
CI
CII
CIII
CIV
CV
PolG1
TFAM
ND-18
ND-20
ND-23
CG14757
SdhC
SdhD
Cyt-c1
UQCR-6.4
RFeSP
COX4
COX5A
levy
blw
ATPsynB
ATPsynC

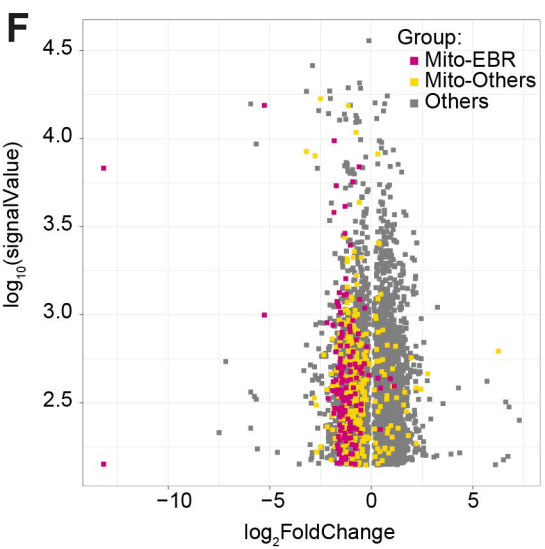
Figure 6

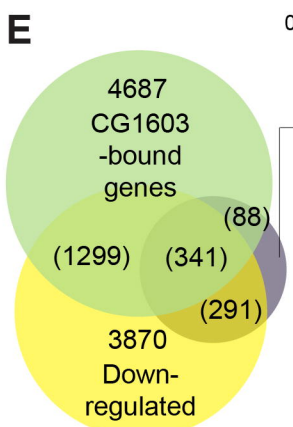












848 Nuclear-encoded mitochondrial genes

Nuclear-encoded	Mitochondrial (848)	Others (10946)
CG1603-bound	429 (50.6%)	4258 (38.9%)
Down	341 (79.5%)	1299 (30.5%)
Down > 2 fold	182 (42.4%)	420 (9.9%)

Figure 7

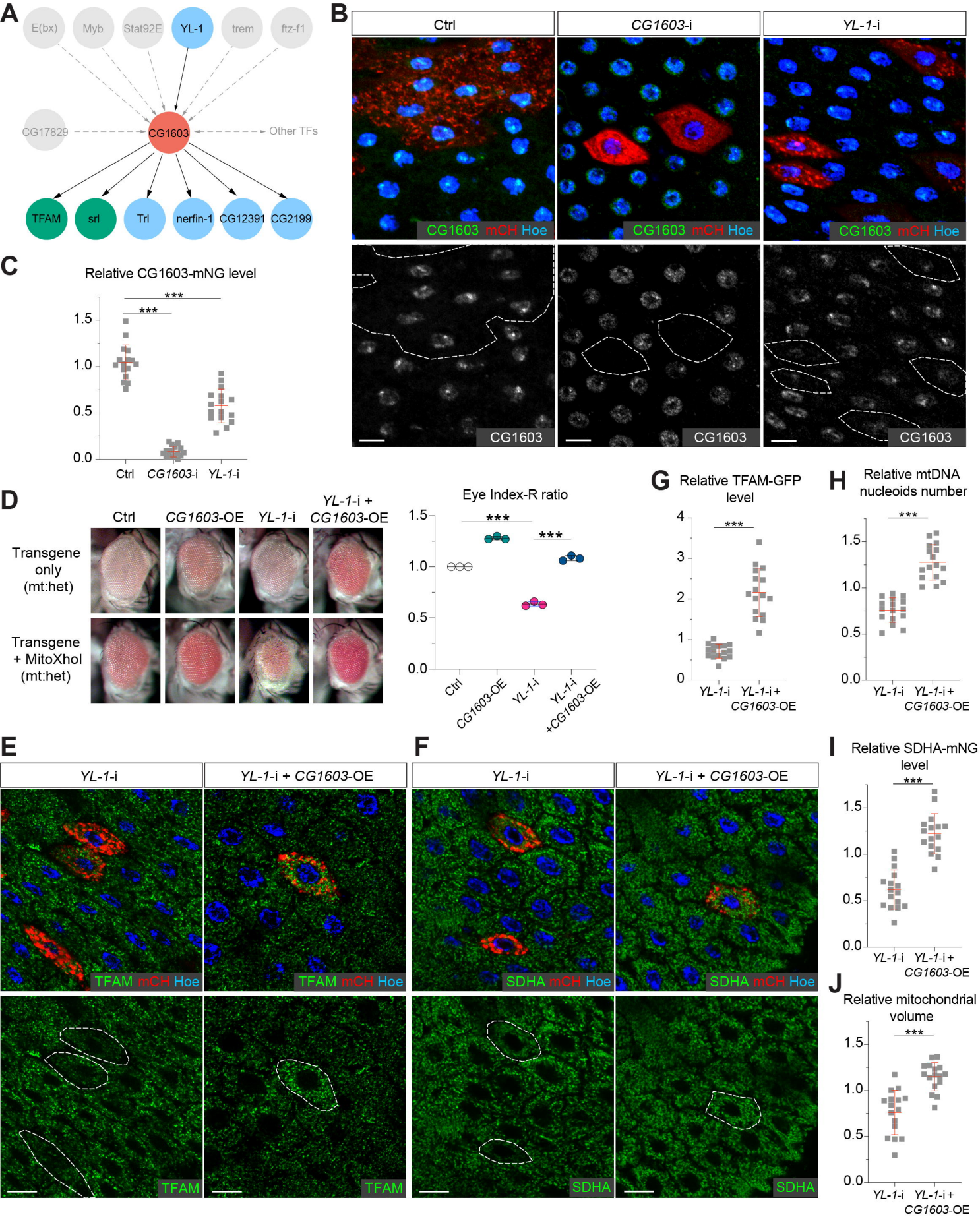

A
E(bx)
Myb
Stat92E
YL-1
trem
ftz-f1
CG17829
CG1603
Other TFs
TFAM
srl
Trl
nerfin-1
CG12391
CG2199
B
Ctrl
CG1603-i
YL-1-i
CG1603 mCH Hoe
CG1603
C
Relative CG1603-mNG level
***
***
1.5
1.0
0.5
0.0
Ctrl
CG1603-i
YL-1-i
D
Ctrl
CG1603-OE
YL-1-i
YL-1-i + CG1603-OE
Transgene only (mt:het)
Transgene + MitoXhoI (mt:het)
Eye Index-R ratio
***
***
1.5
1.0
0.5
0.0
Ctrl
CG1603-OE
YL-1-i
YL-1-i +CG1603-OE
G
Relative TFAM-GFP level
***
4
3
2
1
0
YL-1-i
YL-1-i + CG1603-OE
H
Relative mtDNA nucleoids number
***
1.5
1.0
0.5
0.0
YL-1-i
YL-1-i + CG1603-OE
E
YL-1-i
YL-1-i + CG1603-OE
TFAM mCH Hoe
TFAM
F
YL-1-i
YL-1-i + CG1603-OE
SDHA mCH Hoe
SDHA
I
Relative SDHA-mNG level
***
1.5
1.0
0.5
0.0
YL-1-i
YL-1-i + CG1603-OE
J
Relative mitochondrial volume
***
1.5
1.0
0.5
0.0
YL-1-i
YL-1-i + CG1603-OE

Figure 8

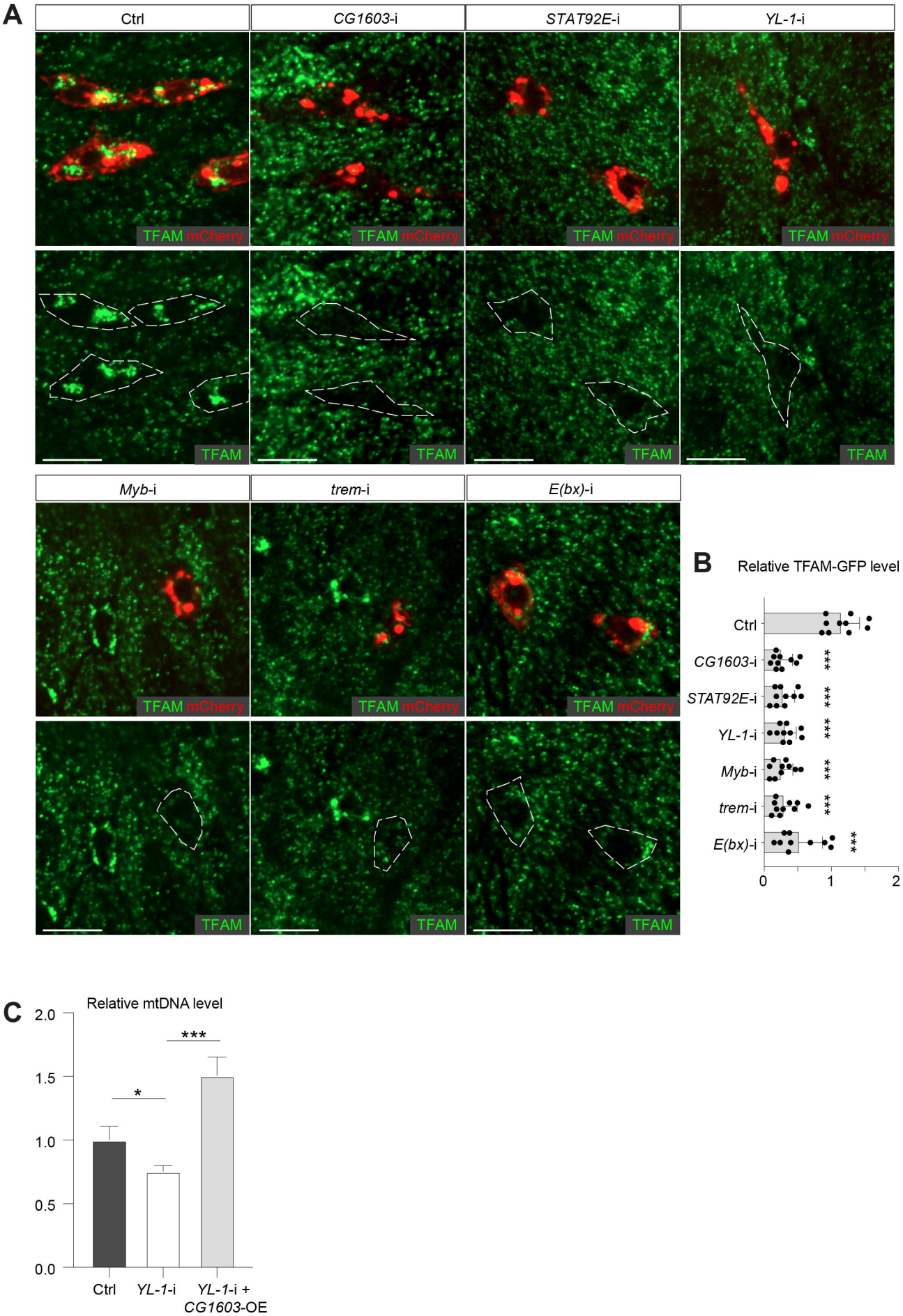



Figure 8—figure supplement