SUPPLEMENTAL INFORMATION

PGC-1α Senses the CBC of Pre-mRNA to Dictate the Fate of Promoter-Proximally Paused RNAPII

Xavier Rambout, Hana Cho, Roméo Blanc, Qing Lyu, Joseph M. Miano, Joe V. Chakkalakal, Geoffrey M. Nelson, Hari K. Yalamanchili, Karen Adelman, and Lynne E. Maquat

SUPPLEMENTAL FIGURES

Rambout et al., Figure S1

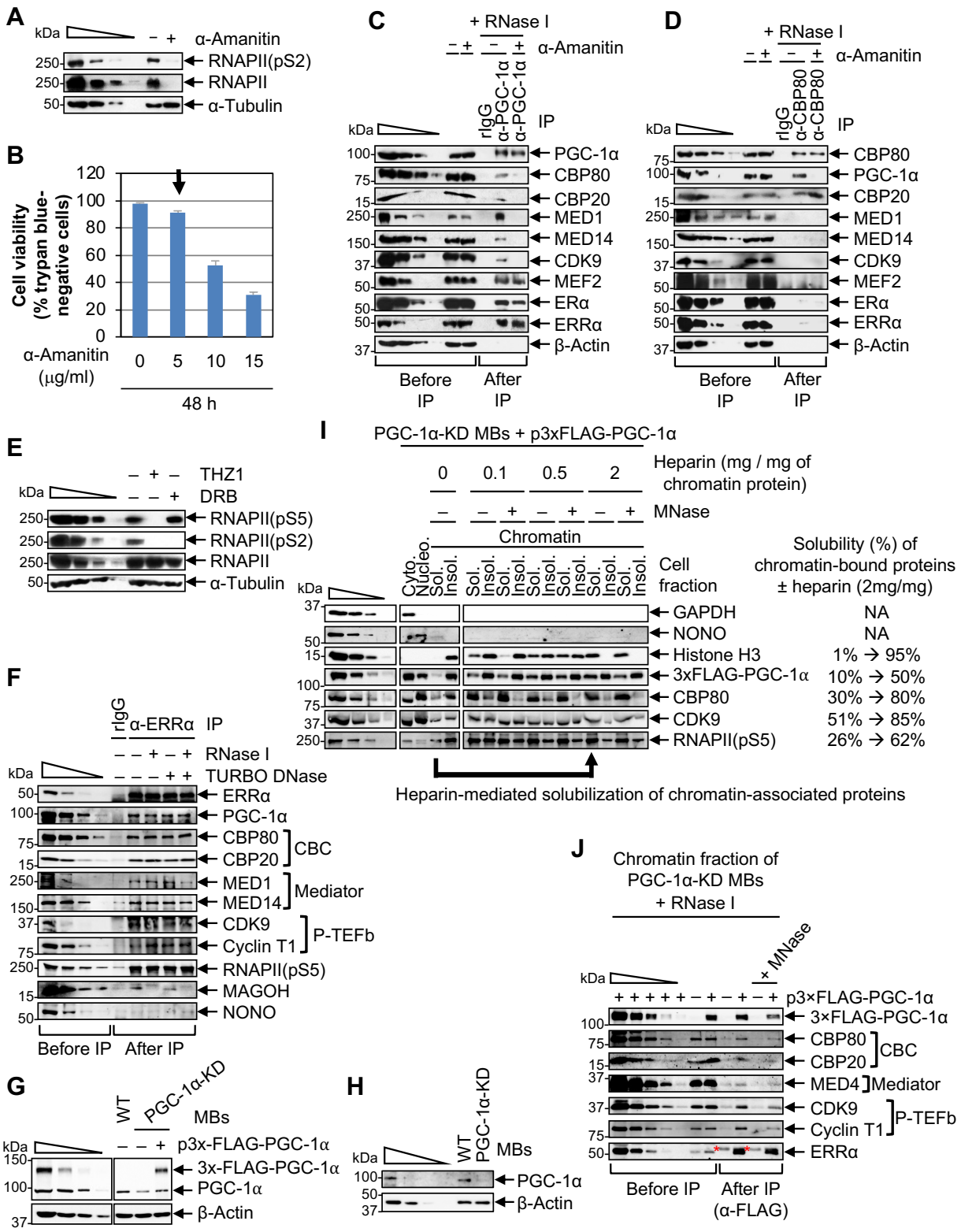

Figure S1. Characterizations of C2C12 MBs cultured in the presence or absence of inhibitors, α-ERRα IPs, and cell fractionations. Related to **Figure 1**.

(**A**) Western blots of lysates of C2C12 MBs cultured with 5 μg/ml of α-amanitin for 48 hours to confirm that gene transcription was inhibited and RNAPII was degraded. In this Western blot and others, the level of α-tubulin serves to control for variations in protein loading and, when relevant, IP specificity, and the upper left wedge denotes 3-fold serial dilutions of samples to provide semi-quantitative comparisons, and n = 2-3 biological replicates.

(**B**) Histogram representations of C2C12-MB viability, measured as the percentage of trypan blue-exclusion, i.e. trypan blue-negative cells, demonstrating that >90% of cells were viable after culturing with 5 μg/ml of α-amanitin for 48 hours (condition indicated with an arrow). Results are means ± S.D (n = 2 biological replicates).

(**C**) Western blots of lysates of C2C12 MBs that were (+) or were not (–) treated with α-amanitin as described in **A**, before or after IP in the presence of RNase I using anti(α)-PGC-1α or, as a control, rabbit (r)IgG. Here and when utilized elsewhere, the level of β-Actin serves to control for variations in protein loading and IP specificity.

(**D**) As in **C**, but using anti-CBP80 in place of anti-PGC-1α.

(**E**) Western blots of lysates of C2C12 MBs cultured with THZ1 or DRB, demonstrating conditions under which THZ1 inhibits transcription initiation, i.e. inhibits RNAPII(pS5) and RNAPII(pS2) formation, and DRB inhibits promoter-proximal pausing, i.e., inhibits RNAPII(pS2) formation, each without affecting the level of total-cell RNAPII.

(**F**) Western blots of lysates of C2C12 MBs before or after IP in the presence (+) or absence (–) of RNase I and/or TURBO DNase using anti-ERRα or, as a control, rIgG. The percentage (%) of input refers to the amount of lysate loaded in the left-most lane of the titration wedge. The paraspeckle constituent NONO, also called is $p54^{nrb}$, serves to control for IP specificity. The exon-junction complex constituent MAGOH serves to control for an RNase I-sensitive interaction.

(**G**) Western blots of lysates of wild-type (WT) C2C12 MBs or PGC-1α-KD C2C12 MBs, i.e. a pool of C2C12 MBs stably expressing *PGC-1α* shRNA so that the level of PGC-1α is ~40% of normal, transiently transfected with a plasmid (p) producing FLAG-PGC-1α(WT) (+) or FLAG alone (–).

(**H**) Western blots of lysates of WT or PGC-1α-KD C2C12 MBs using a PGC-1α antibody specific to the first 300-amino acids of PGC-1α1/-a isoform (sc-518025). All other PGC-1α western blots utilized an antibody raised against a peptide encoded by exon 8 of human PGC-1α mRNA (Novus, NBP1-04676), which is present in all known long PGC-1α isoforms (see also **Limitations of the study**).[S1]

(**I**) Western blots demonstrating effective fractionation of PGC-1α-KD MBs transiently expressing 3×FLAG-PGC-1α(WT) and effective solubilization of chromatin-associated proteins using heparin and MNase. Cyto., cytoplasmic fraction; Nucleo., nucleoplasmic fraction; Sol., soluble; Insol., insoluble. The condition used to solubilize chromatin in **Figure 1E,F** is indicated with an arrow.

(**J**) Western blots of the solubilized chromatin fraction of PGC-1α-KD MBs transiently expressing 3×FLAG-PGC-1α(WT) or FLAG alone (–) before or after anti-FLAG IP in the presence of RNase I and, where specified, in the presence (+) of micrococcal nuclease (MNase). Red asterisks denote a non-specific band in the western blot using anti-ERRα that does not interfere with analyses.

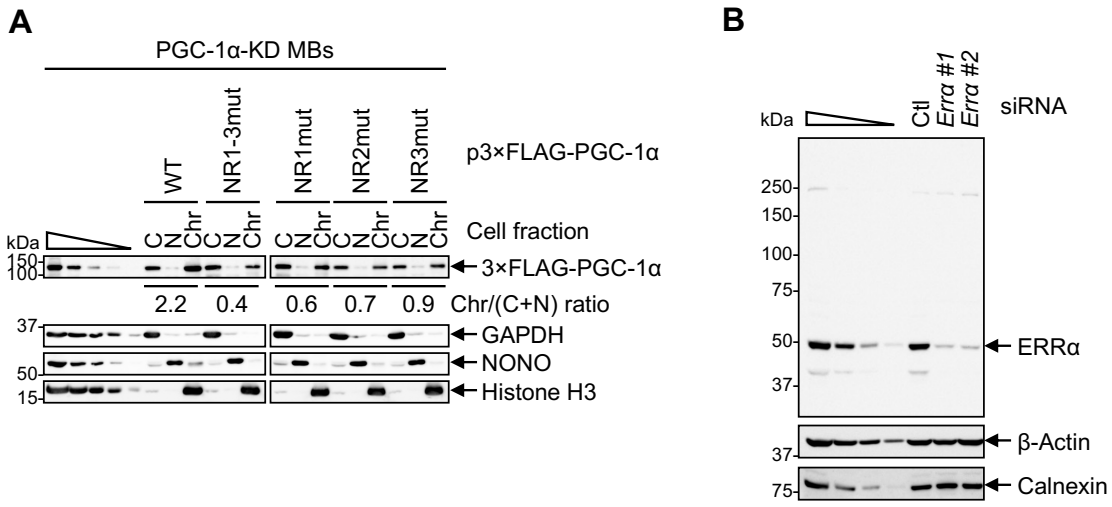

Figure S2. Demonstration of effective fractionation of PGC-1α-KD C2C12 MBs expressing a FLAG-PGC-1α variant and of effective ERRα downregulation using each of two siRNAs. Related to **Figure 2**.

(**A**) As **Figure S1I**, but showing the fractionation of PGC-1α-KD MBs treated as in **Figure 2C**.

(**B**) Western blots demonstrating that in WT MBs transiently transfected with each of two *Errα* siRNA, the level of ERRα is downregulated to ~10% of its normal level, i.e. its level in WT MBs transiently transfected with control (Ctl) siRNA. β-Actin and Calnexin serve to control for variations in protein loading.

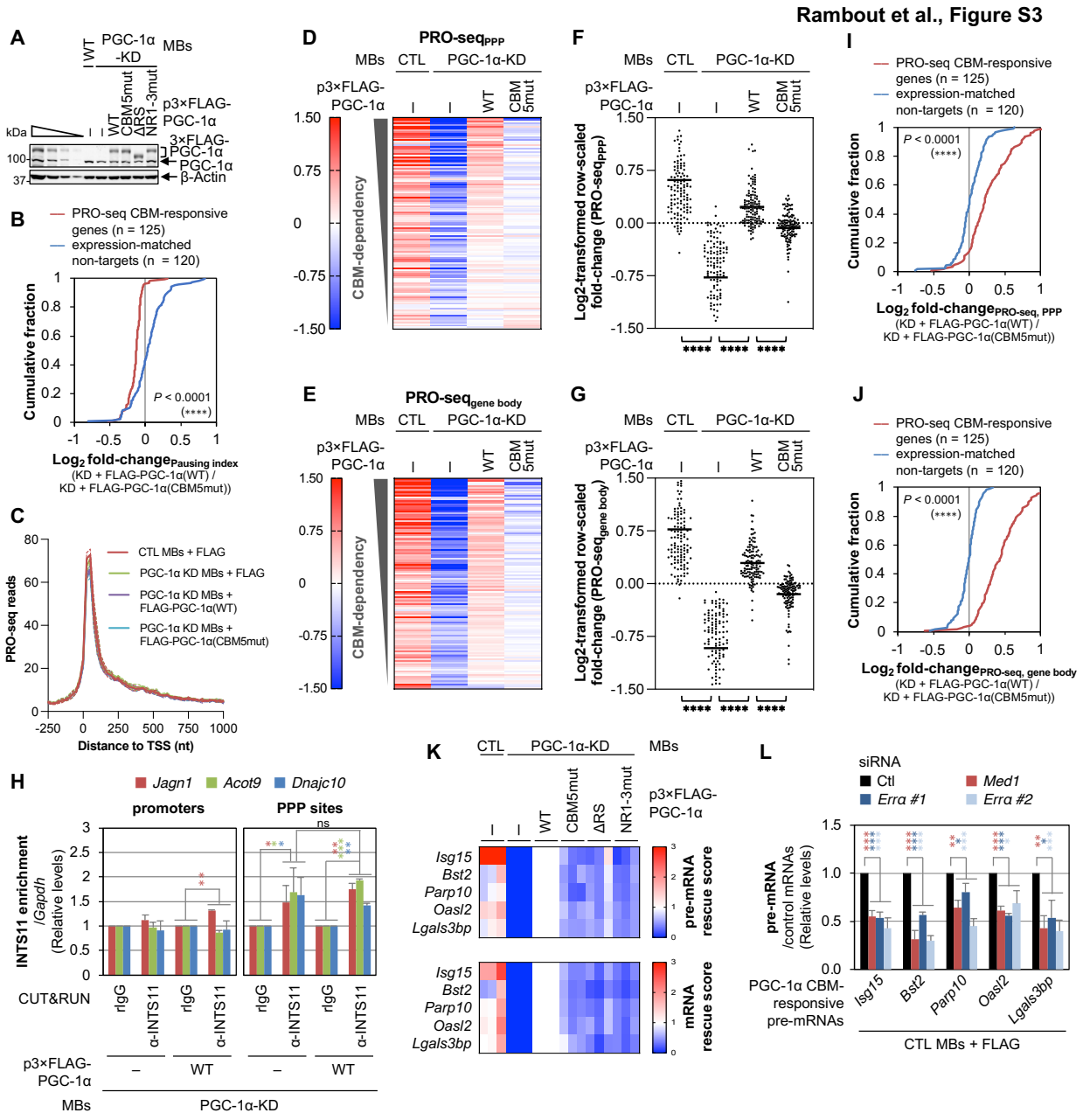

Figure S3. Supporting evidence that the PGC-1α(CBM) activates the transcription of PGC-1α-responsive genes by favoring release of PPP over early transcription termination. Related to **Figure 4**.

(**A**) Western blots showing downregulation of PGC-1α in PGC-1α-KD MBs relative to CTL MBs, and expression of the specified FLAG-PGC-1α variants in PGC-1α-KD MBs used in **Figure 4**. Results are representative of the three biological replicates.

(**B**) As in **Figure 4C**, but for the distribution of average changes in pausing indices for PRO-seq PGC-1α-responsive genes (red) or expression-matched non-target genes (blue) (n = 2 biological replicates).

(**C**) Average distribution of PRO-seq signal (n = 2 biological replicates) in the specified MBs at and around the transcription start site (TSS) of 120 PGC-1α CBM-independent genes whose expression matches the expression of the 125 PGC-1α-responsive genes identified in **Figure 4A**. Read counts are summed in non-overlapping 25-nucleotide bins.
(**D**) Heatmap representation of raw-scaled PRO-seq $\log_2$-fold changes at PPP sites, in the specified MBs, for genes shown in **Figure 4A**. Genes are ranked by CBM-dependency. Results are means (n = 2 biological replicates).
(**E**) As in **D**, but representing raw-scaled PRO-seq $\log_2$-fold changes in gene bodies.
(**F**) Scatter plot representation of raw-scaled PRO-seq $\log_2$-fold changes at PPP sites for the specified MBs. Horizontal bars indicate mean values for each cell condition. *P*-values compare the indicated conditions using a nonparametric Wilcoxon matched-pairs signed rank test. ∗∗∗∗, $P < 0.0001$.
(**G**) As in **F**, but representing PRO-seq $\log_2$-fold changes in gene bodies.
(**H**) Histogram representations of CUT&RUN qPCR quantitations of INTS11 binding to the promoter (left) or PPP site (right) of three control genes in the specified MBs, where values were normalized to binding values obtained for the corresponding locations in the *Gapdh* gene. Results are means ± S.D. (n = 4 biological replicates). *P*-values compare the specified MBs using a two-tailed unpaired Student's *t*-test. ∗, $P < 0.05$; no asterisks or ns, $P \geq 0.05$. Asterisk color corresponds to gene color.
(**I**) Differential cumulative distribution of the average (n = 2 biological replicates) changes in the PRO-seq signal at PPP sites, in the specified MBs. ∗∗∗∗, $P < 0.0001$ comparing PRO-seq PGC-1α-responsive genes (red line) to non-target genes (blue line).
(**J**) As in **I**, but representing changes in the PRO-seq signal in gene bodies.
(**K**) Heatmap representations of "CBM rescue scores" (n = 3 biological replicates) in the specified MBs for five PGC-1α-activated genes identified using PRO-seq, calculated using RT-qPCR. The CBM rescue scores are defined as pre-mRNA (top) or mRNA (bottom) levels in a given MB, normalized to the difference between pre-mRNA or mRNA levels in PGC-1α-KD MBs transiently expressing FLAG-PGC-1α(WT) or pre-mRNA or mRNA levels in PGC-1α-KD MBs transiently expressing FLAG alone. Pre-mRNA and mRNA levels were first normalized to the geometric mean of *Hprt* and *Gapdh* mRNA levels.
(**L**) Histogram representations of RT-qPCR quantitations of pre-mRNA from five PGC-1α-responsive genes normalized to the geometric mean of *Hprt* and *Gapdh* mRNA levels using lysates of CTL MBs transiently transfected with a plasmid encoding FLAG alone and the specified siRNA. Results are means ± S.D. (n = 3 biological replicates), *P*-values compare the indicated conditions using a two-tailed unpaired Student's *t*-test. ∗, $P < 0.05$; ∗∗, $P < 0.01$; ∗∗∗, $P < 0.001$; the color of the stars referring to the color of the gene tested.

Rambout et al., Figure S4

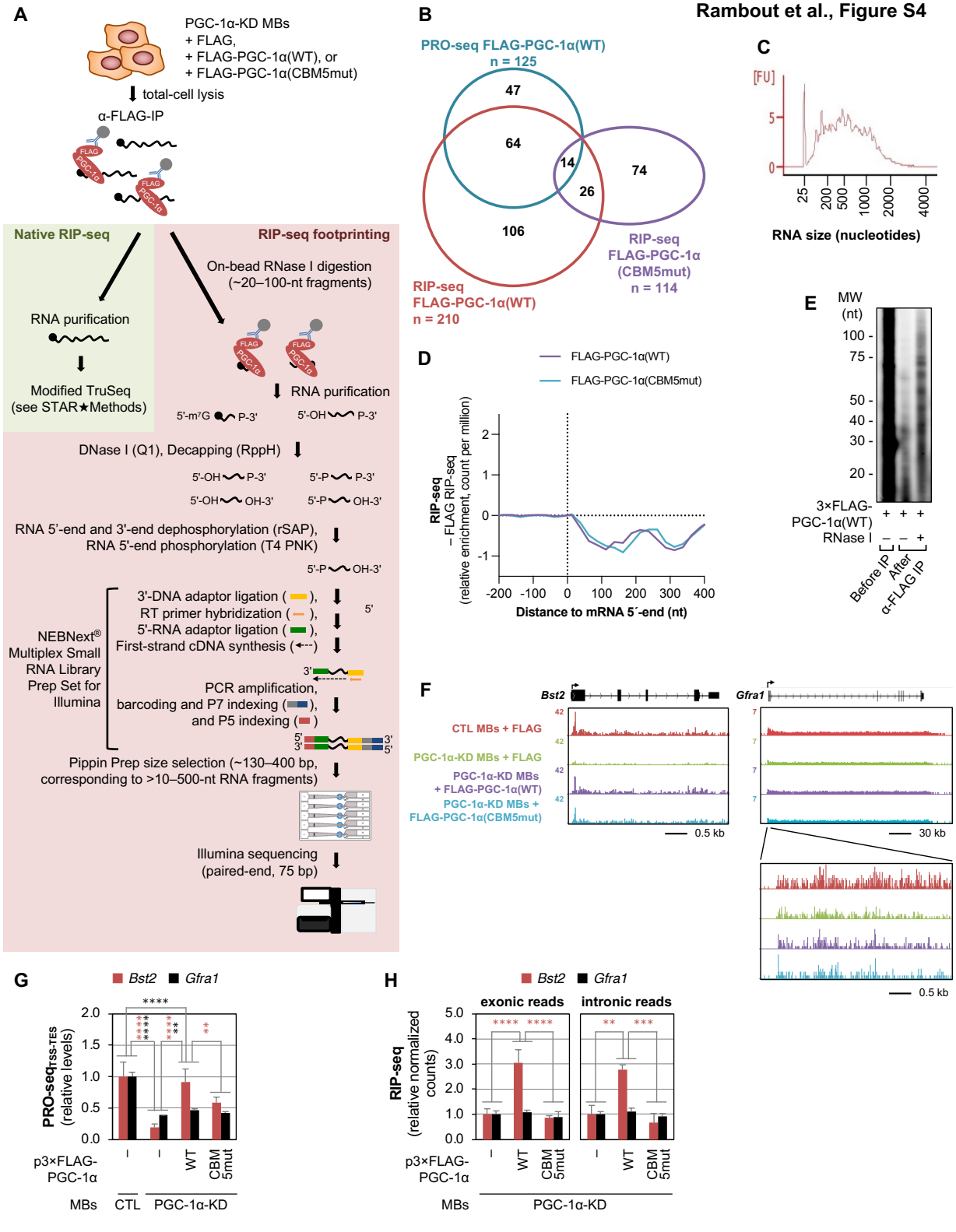

Figure S4. Protocol for native RIP-seq and RIP-seq footprinting, and supporting evidence that the PGC-1α CBM does not bind to non-target RNAs. Related to **Figure 5**.
(**A**) Schematic representations of the native RIP-seq and RIP-seq footprinting methods used to identify mRNAs and/or pre-mRNAs (native RIP-seq) and ~20–100-nt RNA fragments (RIP-seq footprinting) bound by 3×FLAG-PGC-1α(WT), but not 3×FLAG-PGC-1α(CBM5mut), transiently expressed in PGC-1α-KD C2C12 MBs.
(**B**) Venn Diagram representation of the overlay between the 125 transcriptional target genes of 3×FLAG-PGC-1α(WT) (defined by PRO-seq), and the 210 or 144 genes whose transcripts co-immunoprecipitate with 3×FLAG-PGC-1α(WT) or 3×FLAG-PGC-1α(CBM5mut), respectively (defined by RIP-seq, using total-gene reads). More than 60% of PRO-seq target genes are 3×FLAG-PGC-1α(WT) RIP-seq targets ((64+14)/125 = 0.624). More than 80% of these transcripts are not 3×FLAG-PGC-1α(CBM5mut) RIP-seq targets (64/(64+14) = 0.821), i.e. the binding of more than 80% of these transcripts is strictly CBM-dependent.
(**C**) Representative Bioanalyzer trace image (n = 3 biological replicates) of RNA purified from anti-FLAG-PGC-1α(WT) IPs performed in the absence of RNase inhibitors. FU, arbitrary fluorescence units. The first peak in the trace corresponds to the 25-nucleotide marker.
(**D**) Average distribution (n = 3 biological replicates) of the relative enrichment of Illumina reads in anti-FLAG IPs using lysates of PGC-1α-KD MBs expressing FLAG-PGC-1α(WT) or FLAG-PGC-1α(CBM5mut), relative to PGC-1α-KD MBs expressing FLAG alone, for the 5′-region of transcripts from non-target genes. Relative read counts are summed in non-overlapping 25-nucleotide bins.
(**E**) SYBR Gold staining, after electrophoresis in 15% polyacrylamide−TBE-urea, of RNA (1 μg) isolated from PGC-1α-KD C2C12 MBs transiently expressing 3×FLAG-PGC-1α(WT), before or after IP in the presence (+) or absence (–) of RNase I using anti-FLAG. Representative of three biological replicates.
(**F**) PRO-seq signal densities (n = 2 biological replicates) throughout *Bst2* (a PGC-1α CBM-responsive gene in the IFN pathway) or *Gfra1* (a PGC-1α-responsive gene that regulates glucose metabolism), in the specified MBs. Black boxes, exons; black lines, introns; arrows, transcription direction.
(**G**) Histogram representations of full-gene PRO-seq densities for *Bst2* and *Gfra1* in the specified MBs. Results are means ± S.D. (n = 2 biological replicates), *P*-values compare the indicated conditions using DESeq2.[S2] **, $P < 0.01$; ****, $P < 0.0001$; the color of the stars referring to the color of the gene tested.
(**H**) Histogram representations of relative normalized RIP-seq reads mapped to exons or introns of *Bst2* or *Gfra1* genes in the specified MBs. Results are means ± S.D. (n = 3 biological replicates), *P*-values compare the indicated conditions using DESeq2 for exonic reads and a two-tailed unpaired Student's *t*-test for intronic reads (intronic read count is too low for DESeq2 analysis). **, $P < 0.01$; ***, $P < 0.01$; ****, $P < 0.0001$; the color of the stars referring to the color of the gene tested.

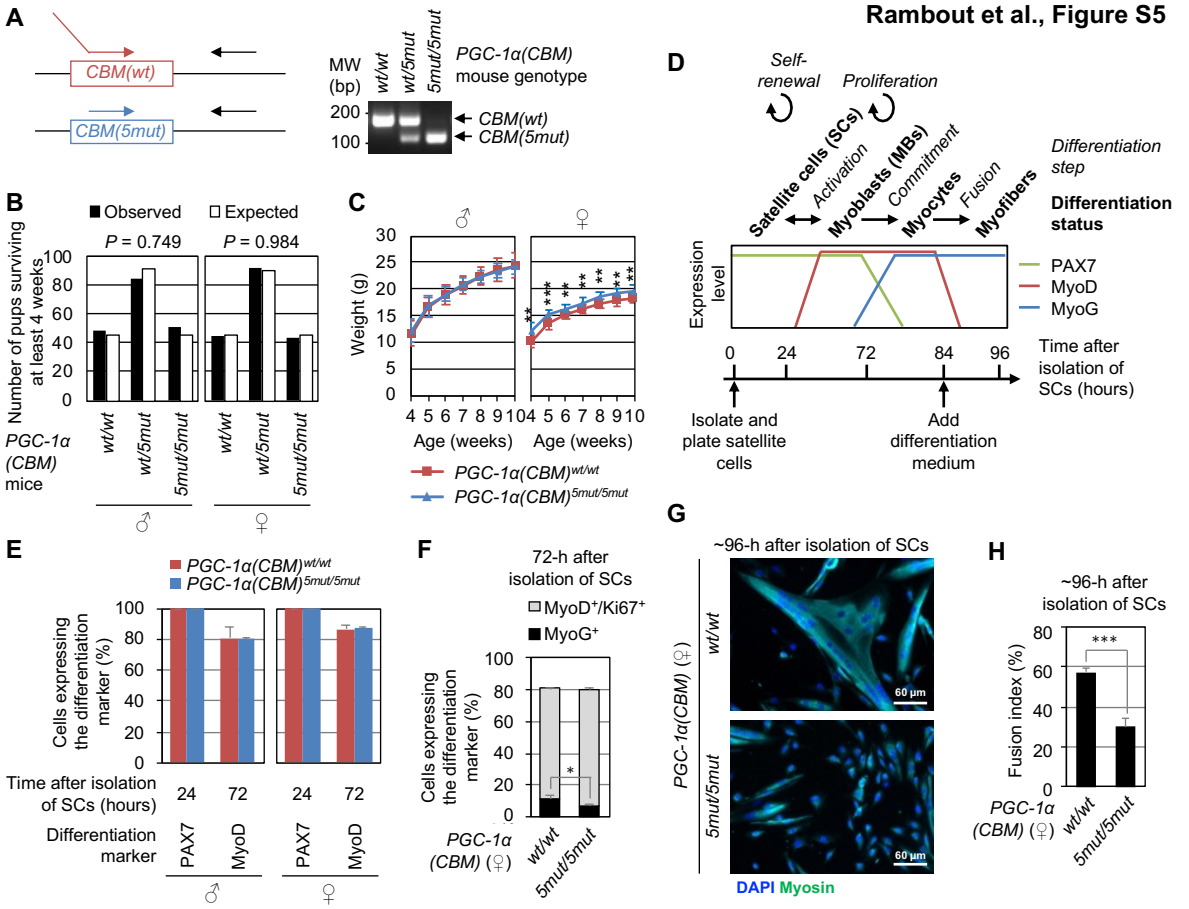
A
CBM(wt)
CBM(5mut)
MW (bp)
wt/wt
wt/5mut
5mut/5mut
PGC-1α(CBM) mouse genotype
200
100
CBM(wt)
CBM(5mut)
B
Observed
Expected
P = 0.749
P = 0.984
Number of pups surviving at least 4 weeks
PGC-1α (CBM) mice
C
Weight (g)
Age (weeks)
PGC-1α(CBM)wt/wt
PGC-1α(CBM)5mut/5mut
D
Self-renewal
Proliferation
Satellite cells (SCs)
Activation
Myoblasts (MBs)
Commitment
Myocytes
Fusion
Myofibers
Differentiation step
Differentiation status
Expression level
PAX7
MyoD
MyoG
Time after isolation of SCs (hours)
Isolate and plate satellite cells
Add differentiation medium
E
Cells expressing the differentiation marker (%)
Time after isolation of SCs (hours)
Differentiation marker
F
72-h after isolation of SCs
MyoD+/Ki67+
MyoG+
G
~96-h after isolation of SCs
PGC-1α(CBM) (♀)
60 µm
DAPI Myosin
H
~96-h after isolation of SCs
Fusion index (%)
I
Bst2
Parp10
Isg15
pre-mRNA /control mRNAs (Induced levels)
mRNA /control mRNAs (Induced levels)
IFN-γ [pg/ml] (24 h)
J
IFN-γ [h] (500 pg/ml)
K
PGC-1α(CBM)wt/wt MBs (♂)
+ DRB
IFN-γ (500 pg/ml, 24h)
IP + RNase I
rIgG
α-CBP80
kDa
CBP80
PGC-1α
MED1
MED4
CDK9
RNAPII(pS5)
ERRα
β-Actin
Before IP
After IP

Figure S5. Characterization of the *PGC-1α(CBM)$^{5mut/5mut}$* mice, supporting evidence that the PGC-1α(CBM) contributes to myogenesis and to IFN-ɣ-dependent activation of gene expression *ex vivo*. Related to **Figure 6**.

(**A**) (**Left**) PCR strategy used to define the *PGC-1α CBM* genotype of mice. →, sense PCR primers used to differentiate PCR products deriving from wild-type (*wt*) alleles (longer products) from mutated (*5mut*) alleles (shorter products); ←, common antisense PCR primer. (**Right**) Ethidium bromide staining of a 2.5% agarose gel showing the products of PCR-amplification using ear-punch lysates deriving from *PGC-1α(CBM)$^{wt/wt}$*, *PGC-1α(CBM)$^{wt/5mut}$*, and *PGC-1α(CBM)$^{5mut/5mut}$* mice.

(**B**) Histogram representations of the number of male (♂) or female (♀) live-born pups, by genotype, that derived from heterozygous intercrosses and reached weaning age. *P*-value is calculated for observed numbers (black bars) compared to expected Mendelian numbers (white bars), using a χ^2 test.

(**C**) Broken line representations of the weight of male and female *PGC-1α(CBM)$^{wt/wt}$* mice (red rectangles and lines) or *PGC-1α(CBM)$^{5mut/5mut}$* mice (blue triangles and lines) as a function of age. Results are mean ± S.D. (n ≥ 8 biological replicates). ∗∗, $P < 0.01$; ∗∗∗, $P < 0.001$; no asterisks, $P \geq 0.05$.

(**D**) Schematic representations of the differentiation of primary mouse SCs into myofibers (top), indicating the timely expression of key protein markers used to define the differentiation status in cultures over time (middle). The protocol used to differentiate freshly isolated SCs into myofibers is illustrated (bottom).

(**E**) Histogram representations of the percentage of SCs isolated from male or female *PGC-1α(CBM)$^{wt/wt}$* (red bars) or *PGC-1α(CBM)$^{5mut/5mut}$* (blue bars) adult mice, expressing the SC marker PAX7 or the MB marker MyoD 24-hours or 72-hours, respectively, post-isolation. Results are mean ± S.D. (n = 4 biological replicates).

(**F**) As in **Figure 6D**, but using cells derived from female mice.

(**G**) As in **Figure 6E**, but using cells derived from female mice.

(**H**) As in **Figure 6F**, but using cells derived from female mice.

(**I**) Broken line representations of RT-qPCR quantitations of the induction of pre-mRNA (top) and mRNA (bottom) from three PGC-1α(CBM)-responsive genes in primary MBs derived from *PGC-1α(CBM)$^{wt/wt}$* (red lines) or *PGC-1α(CBM)$^{5mut/5mut}$* (blue lines) adult mice following treatment with the specified concentrations of IFN-ɣ for 24 hours. Results are means ± S.E.M. (n = 3–6 biological replicates), where values for untreated cells were set to 1. ∗, $P < 0.05$; ∗∗, $P < 0.01$; ∗∗∗, $P < 0.001$; no asterisks, $P \geq 0.05$.

(**J**) As in **I**, but after different times of treatment with IFN-ɣ at 500 pg/ml (n = 4).

(**K**) WBs of lysates of primary MBs derived from male (♂) *PGC-1α(CBM)$^{wt/wt}$* mice, that were treated with DRB and were (+) or were not (–) treated with IFN-ɣ (500 pg/ml, 24 hours), before or after IP in the presence of RNase I using α-CBP80 or, as a control, rIgG.

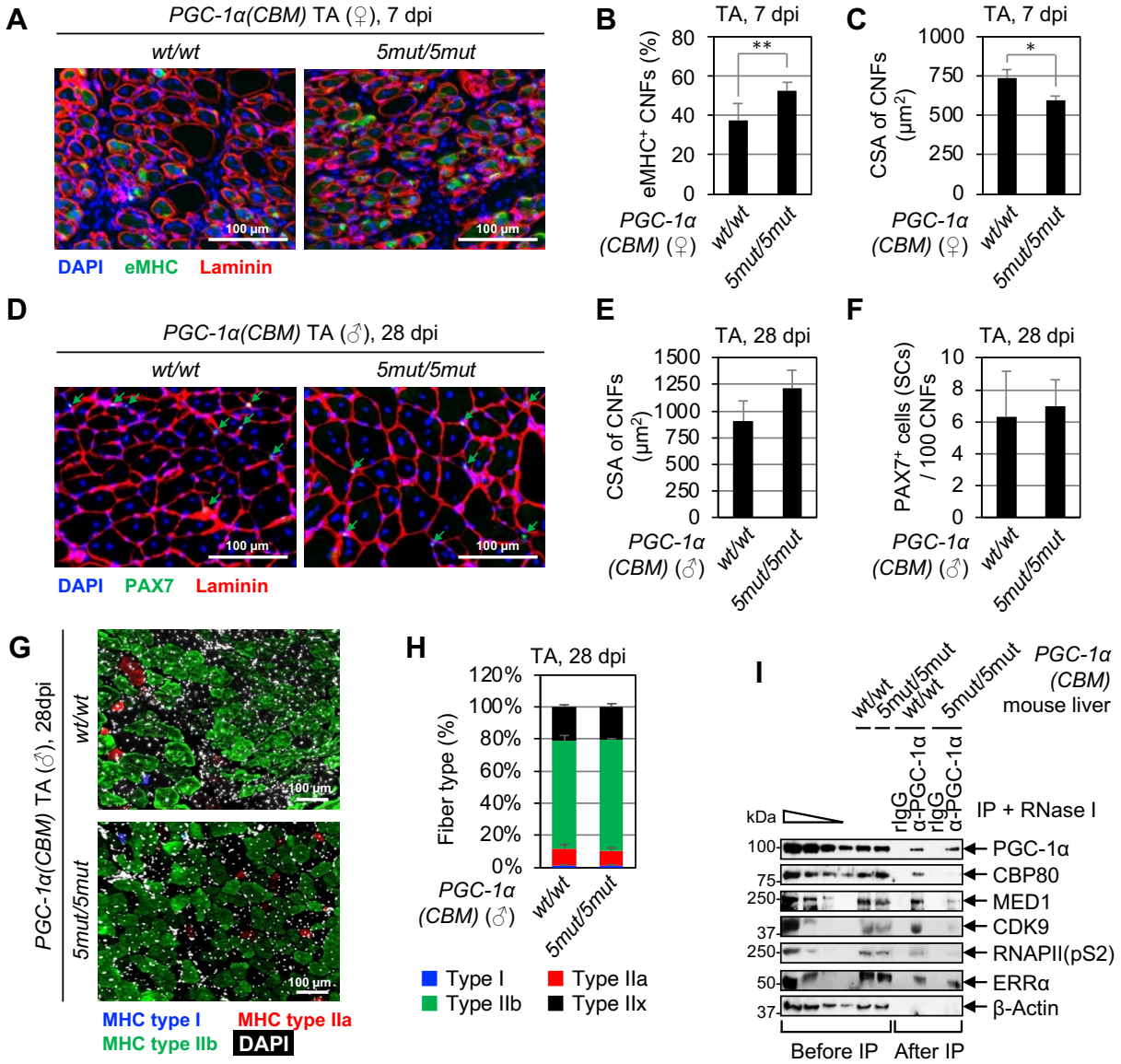

Figure S6. Supporting evidence that the PGC-1α(CBM) contributes to skeletal muscle regeneration after injury, and functions in liver. Related to **Figure 6**.

(**A**) As in **Figure 6G**, but using tissues derived from female mice.

(**B**) As in **Figure 6H**, but using tissues derived from female mice.

(**C**) As in **Figure 6I**, but using tissues derived from female mice.

(**D**) Representative images of immunofluorescence staining of PAX7 (green), Laminin (red), and nuclei (DAPI, blue) in transverse sections of injured tibialis anterior (TA) muscles from male *PGC-1α(CBM)*$^{5mut/5mut}$ or *PGC-1α(CBM)*$^{wt/wt}$ mice, isolated 28-days post $BaCl_2$-induced injury (7 dpi) (n = 4 mice).

(**E**) As in **Figure 6I**, but 28 dpi (n = 4 mice).

(**F**) Histogram representations of the number of PAX7-positive cells, i.e. SCs, per 100 centrally nucleated fibers (CNFs), i.e. newly generated fibers, in transverse sections of regenerated TA muscles from male *PGC-1α(CBM)*$^{5mut/5mut}$ or *PGC-1α(CBM)*$^{wt/wt}$ mice (28 dpi) (n = 4 mice).
(**G**) Representative images of immunofluorescence staining of myosin heavy chain (MHC) type I (blue), MHC type IIa (red), type IIb (green), and nuclei (DAPI, white) in transverse sections of regenerated TA muscles from male *PGC-1α(CBM)*$^{wt/wt}$ mice (28 dpi).
(**H**) Histogram representations of the percentage of type I, type IIa, type IIb and type IIx muscle fibers in transverse sections of regenerated TA muscles from male *PGC-1α(CBM)*$^{wt/wt}$ mice (28 dpi) (n = 4 mice).
(**I**) WBs of lysates of whole liver from male *PGC-1α(CBM)*$^{wt/wt}$ or *PGC-1α(CBM)*$^{5mut/5mut}$ mice, before or after IP in the presence of RNase I using α-PGC-1α or, as a control, rabbit (r)IgG.

SUPPLEMENTAL REFERENCES

S1. Martínez-Redondo, V., Pettersson, A.T., and Ruas, J.L. (2015). The hitchhiker's guide to PGC-1alpha isoform structure and biological functions. Diabetologia *58*, 1969-1977. 10.1007/s00125-015-3671-z.

S2. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. 10.1186/s13059-014-0550-8.