Human primitive brain displays negative mitochondrial-nuclear expression correlation of respiratory genes

Gilad Barshad, Amit Blumberg, Tal Cohen, and Dan Mishmar

Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel

Oxidative phosphorylation (OXPHOS), a fundamental energy source in all human tissues, requires interactions between mitochondrial (mtDNA)- and nuclear (nDNA)-encoded protein subunits. Although such interactions are fundamental to OXPHOS, bi-genomic coregulation is poorly understood. To address this question, we analyzed ~8500 RNA-seq experiments from 48 human body sites. Despite well-known variation in mitochondrial activity, quantity, and morphology, we found overall positive mtDNA-nDNA OXPHOS genes' co-expression across human tissues. Nevertheless, negative mtDNA-nDNA gene expression correlation was identified in the hypothalamus, basal ganglia, and amygdala (subcortical brain regions, collectively termed the "primitive" brain). Single-cell RNA-seq analysis of mouse and human brains revealed that this phenomenon is evolutionarily conserved, and both are influenced by brain cell types (involving excitatory/inhibitory neurons and nonneuronal cells) and by their spatial brain location. As the "primitive" brain is highly oxidative, we hypothesized that such negative mtDNA-nDNA co-expression likely controls for the high mtDNA transcript levels, which enforce tight OXPHOS regulation, rather than rewiring toward glycolysis. Accordingly, we found "primitive" brain-specific up-regulation of lactate dehydrogenase B (*LDHB*), which associates with high OXPHOS activity, at the expense of *LDHA*, which promotes glycolysis. Analyses of co-expression, DNase-seq, and ChIP-seq experiments revealed candidate RNA-binding proteins and CEBPB as the best regulatory candidates to explain these phenomena. Finally, cross-tissue expression analysis unearthed tissue-dependent splice variants and OXPHOS subunit paralogs and allowed revising the list of canonical OXPHOS transcripts. Taken together, our analysis provides a comprehensive view of mito-nuclear gene co-expression across human tissues and provides overall insights into the bi-genomic regulation of mitochondrial activities.

[Supplemental material is available for this article.]

Mitochondrial oxidative phosphorylation (OXPHOS) is the major energy source in most human tissues. Thus, it is not surprising that many congenital mitochondrial disorders display systemic phenotypes (Chinnery and Hudson 2013; Craven et al. 2017) and that altered mitochondrial DNA (mtDNA) gene expression was observed in complex phenotypes such as Alzheimer's disease (Coskun et al. 2004; Lunnon et al. 2017), ALS (Ladd et al. 2017), diabetes (Brun et al. 2015), and various types of cancer (Reznik et al. 2017). Despite the global importance of mitochondrial activity for cellular function, mitochondrial biogenesis, morphology, and number vary among tissues (Fernandez-Vizarra et al. 2011). Since OXPHOS involves interaction of mtDNA- and nuclear DNA (nDNA)-encoded factors, it is unclear whether the tissue variability in mitochondrial activity affects the required coordination between mtDNA- and nDNA-encoded OXPHOS genes and their regulation.

OXPHOS employs five multi-subunit protein complexes harboring ~80 nDNA-encoded proteins, and 13 mtDNA-encoded subunits (Calvo and Mootha 2010). As during the course of evolution, many OXPHOS genes were transferred from the ancestral mtDNA to the host nucleus (Gabaldón and Huynen 2007), these genes are currently located throughout the human karyotype and are individually regulated. In contrast, human mtDNA-encoded genes have retained their prokaryotic ancestral cotranscription strategy in two heavy- and one light-strand polycistrons (Aloni and Attardi 1971; Gustafsson et al. 2016). These strand-specific polycis-

trons dramatically differ in their coding content, with the first harboring 12 of the 13 OXPHOS subunits, and the remaining protein, the MT-ND6 subunit of OXPHOS complex I (i.e., NADH:ubiquinone oxidoreductase), being encoded by the light strand. Such prokaryotic-like expression is thought to be regulated mainly by a dedicated mitochondrial-targeted RNA polymerase (POLRMT) and two transcription factors (TFs), TFAM and TFB2M (Gustafsson et al. 2016). Although several known regulators of nuclear gene transcription were recently shown to be imported into the mitochondria and regulate mtDNA transcription (Chen et al. 2004; Reyes et al. 2011; She et al. 2011; Blumberg et al. 2014), the dispersal of OXPHOS genes throughout the human genome, in addition to their division between the nuclear and cytoplasmic genomes, apparently interferes with transcriptional coregulation.

It was previously argued that genes encoding factors belonging to the same pathway tend to be cotranscriptionally regulated (Eisen et al. 1998; Shyamsundar et al. 2005). This hypothesis has been tested for nDNA-encoded OXPHOS genes at the level of gene expression using microarrays (van Waveren and Moraes 2008). van Waveren and Moraes argued for the tendency of transcripts encoding OXPHOS subunits to co-express and that such co-expression tended to be subdivided according to the five OXPHOS protein complexes, i.e., complex I, succinate

dehydrogenase (complex II), cytochrome bc1 (complex III), cytochrome oxidase (complex IV), and ATP synthase (complex V). Nevertheless, although the gene expression microarray profiles considered multiple human tissues ($N = 79$), they were derived from only a few individuals, thus interfering with assessment of OXPHOS genes co-expression per tissue. Additionally, these experiments did not employ mtDNA probes and did not consider tissue-specific splice variants. Another analysis that included mtDNA genes was limited to only a few nDNA genes (Garbian et al. 2010). Recent analysis of RNA-seq data from The Cancer Genome Atlas (TCGA) sample collection enabled assessing the expression patterns of mtDNA-encoded OXPHOS subunits and corresponding nDNA-encoded OXPHOS genes in a variety of cancer types relative to normal samples (Reznik et al. 2017). This study revealed down-regulation of OXPHOS transcripts encoded by the two genomes in most tested cancer types, with some exceptions (i.e., lung adenocarcinoma and thyroid cancer). However, as OXPHOS gene expression was only measured in tumors relative to normal adjacent tissues, one cannot easily extend these observations to understand the pattern of OXPHOS gene expression across and within normal, healthy tissues.

Differences in mitochondrial quantity, morphology, network interactions (Liesa and Shirihai 2013), and OXPHOS function were observed in several model organisms (Fernandez-Vizarra et al. 2011). Such variation interferes with any generalization concerning OXPHOS genes coregulation from one tissue to others. Furthermore, the identification of tissue-specific subunit paralogs in OXPHOS complex IV (Schlerf et al. 1988; Lightowlers et al. 1990; Segade et al. 1996), as well as tissue-specific splice variants in the complex I subunit NDUFV3 (Bridges et al. 2017; Dibley et al. 2017; Guerrero-Castillo et al. 2017), provided first clues of possible tissue variation in OXPHOS subunit composition. Given the above, although OXPHOS function is required for most tissues, it is not trivial to assume similarity in OXPHOS genes' coregulation across tissues. Therefore, assessing OXPHOS genes co-expression in a large variety of tissues, including both mtDNA and nDNA genes, is essential.

Here, we utilized the human gene-tissue expression (GTEx) database (The GTEx Consortium 2015), containing experimental information (~8500 RNAs-seq experiments) collected from a variety of body sites ($N = 48$) from at least 50 individuals per tissue. Our analysis enabled assessment of gene expression correlation between the mtDNA and nDNA across multiple tissues, identification of tissue-dependent expression of OXPHOS genes and splice-variants, and allowed providing first mechanistic/regulatory explanation for the identified landscape of OXPHOS gene co-expression in humans.

Results

nDNA-encoded OXPHOS structural subunits are co-expressed across human body sites

Since proteins that participate in the same biochemical activity tend to co-express at the RNA level (Eisen et al. 1998; Shyamsundar et al. 2005), we hypothesized that mtDNA- and nDNA-encoded OXPHOS genes (Supplemental Tables S1, S2) would likely be co-expressed. To test this hypothesis, we assessed correlation between OXPHOS gene expression across multiple human body sites using the publicly available GTEx RNA-seq expression portal. For the sake of consistency, and to avoid sample size-related effects, we only considered body sites for which expression data from at least

50 samples from unrelated donors was available ($N = 48$ tissues), and extracted expression data for OXPHOS genes (Supplemental Table S1), which were then normalized for age, gender, and cause of death (Supplemental Fig. S1). Calculation of expression pattern correlations (Spearman's rank correlation coefficient, ρ_s; see details in Methods) revealed clusters of co-expressed OXPHOS genes (Supplemental Table S3; Fig. 1). Specifically, such analysis revealed high similarity in the expression pattern of all nDNA-encoded OXPHOS genes, consistent with previous analysis (van Waveren and Moraes 2008). Expression level analysis suggested that although certain OXPHOS subunits are present in more than one copy at the protein level in the assembled complexes (i.e., two copies of NDUFAB1 in complex I, three copies of ATP-synthase F1 subunits α and β, and the eight copies of subunit c in ATP-synthase F0), their transcripts did not overexpress as compared to other subunits from the same complex. This suggests that OXPHOS protein stoichiometry is not regulated at the level of transcripts but rather either at the protein level or at the copy number of genes encoding each subunit (such as the three genes that encode the c-ring of complex V).

Nevertheless, although overall clustering of gene expression was observed between nDNA-encoded OXPHOS genes, we noticed that nine genes (NDUFS4, NDUFC2, NDUFA11, SDHC, SDHD, COX7A2, ATP5MC2, ATP5IF1, and ATP5F1E) showed relatively lower similarity to the rest of the nDNA-encoded OXPHOS genes (hereafter called the "lower similarity cluster") (Fig. 1). This suggests that a subset of the OXPHOS structural subunits are differentially expressed in certain tissues. Notably, two of these genes (i.e., COX7A2 and ATP5MC2) have known paralogs (see further expansion on this point below).

Gene expression patterns of nDNA- and mtDNA-encoded OXPHOS subunits are correlated

As four of the five OXPHOS protein complexes involve interactions of mtDNA- and nDNA-encoded subunits, we asked whether their gene expression patterns correlate. To avoid sequence contamination by mtDNA fragments that migrated to the nucleus and integrated in the nDNA during the course of evolution, i.e., nuclear mitochondrial pseudogenes (NUMTs) (Hazkani-Covo et al. 2003; Mishmar et al. 2004), we removed NUMT reads that were identified by the GTEx Consortium from further analyses (see Supplemental Methods). Our analysis (after NUMT exclusion) revealed high correlation between the expression patterns of all mtDNA-encoded OXPHOS genes, with lower expression similarity observed for the MT-ND6 gene (Fig. 1). Such a pattern is consistent with most (12) protein-coding genes being transcribed from the heavy mtDNA strand, in contrast to MT-ND6, which is the only protein-coding gene encoded by the light mtDNA strand.

Although forming a distinct expression cluster, the expression of all mtDNA-encoded genes was positively correlated with nDNA-encoded OXPHOS genes, apart from five redundant OXPHOS genes' paralogs (ATP5MG2, COX6B2, COX7B2, COX8C, and COX7A2L—further discussed below) and six OXPHOS genes belonging to the "lower similarity cluster" (SDHC, SDHD, NDUFS4, ATP5MC2, NDUFA11, and ATP5F1E) (Fig. 1). Taken together, our observations support a general cross-tissue mtDNA-nDNA co-expression.

In agreement with a previous report (van Waveren and Moraes 2008), we found significantly higher positive correlation between subunits of the same OXPHOS complex compared to OXPHOS subunits from different complexes (median within-
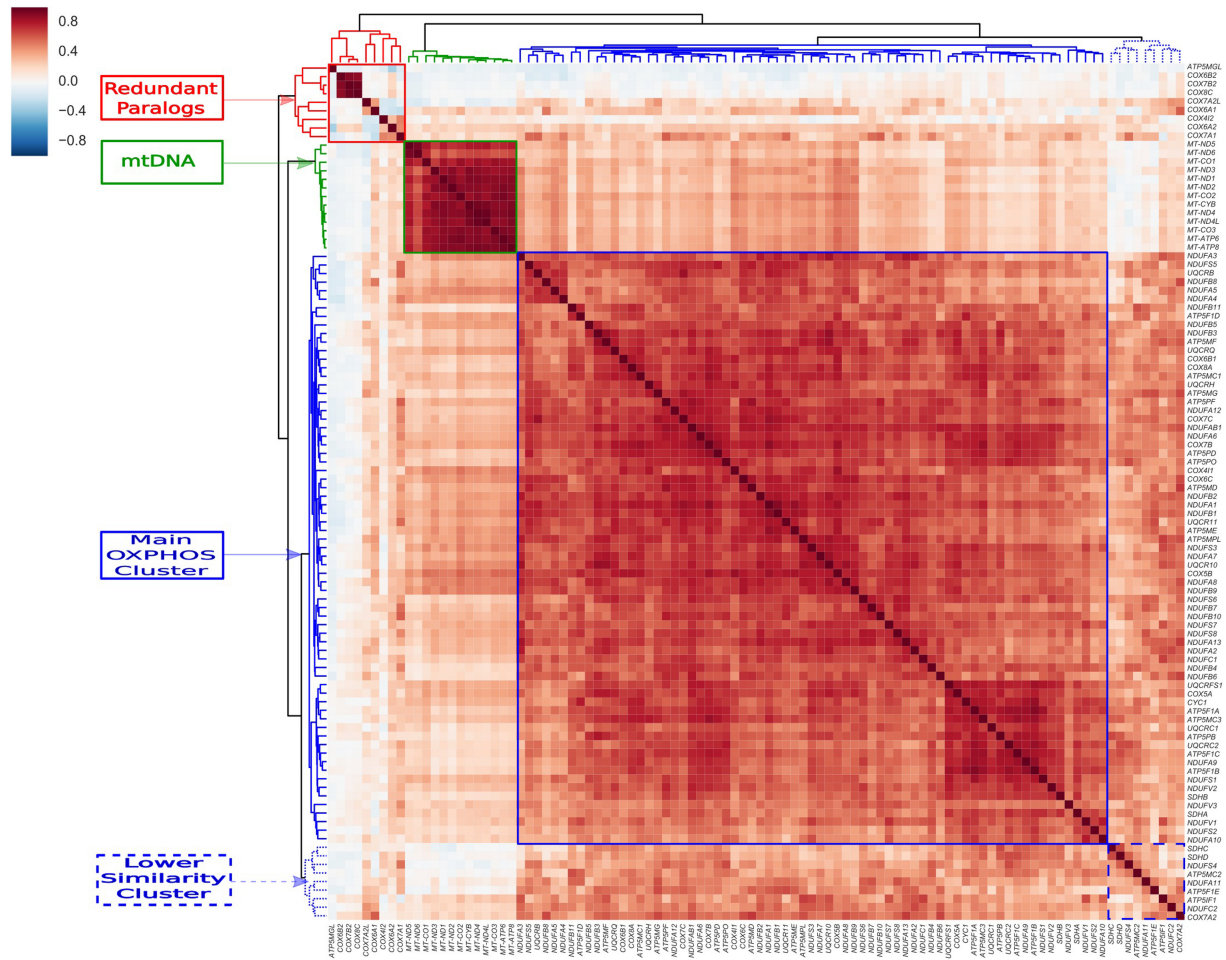
Figure 1. OXPHOS gene expression patterns across all tissues support the general coregulation of mito-nuclear gene expression. A heat map of correlation values for the expression of all OXPHOS genes. The four clusters (redundant paralogs, mtDNA, main OXPHOS cluster, and lower similarity cluster) are framed in squares (red, green, full blue, and dashed blue line, respectively). A key for the color code of the Spearman's correlation coefficients is shown in *top left* corner.

complex correlation = 0.634, median between-complex correlation = 0.619, Mann-Whitney U [MWU] = 3.54×10^6, $P = 5.34 \times 10^{-5}$). Specifically, significant expression correlation within each of the OXPHOS complexes was evident for OXPHOS complexes II, IV, and V, with complex I showing marginal significance and complex III presenting the same trend although statistically insignificant (Supplemental Table S4). Although the electron transport chain (ETC) complexes (complexes I–IV) and ATP synthase (complex V) are considered two separate metabolic modules, no significant differential expression correlation was observed within ETC as compared to ETC-complex V (median within ETC = 0.59, median ETC-complex V = 0.60, MWU = 2.5×10^6, $P = 0.10$). Finally, subunits belonging to complex II, which participates both in OXPHOS and in the tricarboxylic acid (TCA) cycle, showed slightly higher (yet significant) expression correlation with the OXPHOS compared to members of the TCA cycle (median with OXPHOS = 0.45, median with TCA = 0.41, MWU = 1.6×10^4, $P = 0.04$).

We next compared the correlation between the expression patterns of OXPHOS gene pairs to the correlation of OXPHOS gene expression with randomly chosen protein-coding genes (measurements were recalculated in 1000 replicates). The results showed a significantly higher correlation value between OXPHOS gene pairs than between OXPHOS and randomly selected genes (median = 0.54 and median = −0.017, respectively, MWU = 0.0, $P < 1 \times 10^{-100}$) (Supplemental Fig. S2; Fig. 2A). The expression correlation of mtDNA- and nDNA-encoded structural OXPHOS genes remained highly significant even when tested against (1) the entire known mitochondrial proteome (MitoCarta2.0) (Calvo et al. 2016) (median correlation = −0.021, MWU = 1.06×10^{10}, $P < 1 \times 10^{-100}$), or (2) while testing against genes encoding non-OXPHOS proteins that (similar to the OXPHOS complexes) are embedded within the mitochondrial inner membrane (Da Cruz et al. 2003) (median correlation = −0.015, MWU = 6.18×10^8, $P < 1 \times 10^{-100}$) (Fig. 2B). Thus, the correlation of mtDNA- and nDNA-encoded OXPHOS genes is highly specific and is stronger than their correlation with non-OXPHOS protein-coding genes with mitochondrial function.

In addition to the structural and catalytic subunits that comprise the OXPHOS protein complexes, the construction of these complexes requires assembly factors. Therefore, we assessed whether the strong correlation between gene expression patterns of nDNA- and mtDNA-encoded OXPHOS protein
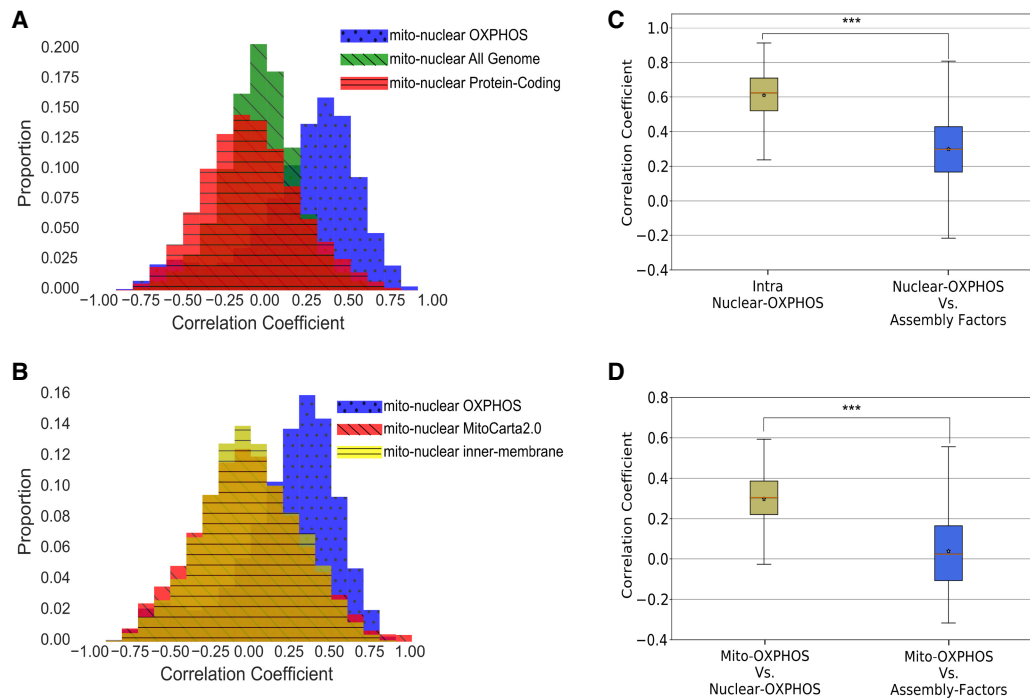
Figure 2. The overall positive co-expression of mitochondrial and nuclear genes is limited to OXPHOS structural subunits, excluding assembly factors. (*A*) A histogram panel demonstrating expression correlation distributions between mtDNA-encoded OXPHOS genes and either genome-wide non-OXPHOS genes (green), non-OXPHOS protein coding genes (red), or nDNA-encoded OXPHOS genes (blue). (*B*) A histogram panel demonstrating expression correlation distributions between mtDNA-encoded OXPHOS gene and either nDNA-encoded non-OXPHOS mitochondrial proteins listed in MitoCarta2.0 (Calvo et al. 2016) (red), a MitoCarta subset of mitochondrial inner-membrane proteins (yellow), and nDNA-encoded OXPHOS genes as a reference (blue). (*C*) Box plots representing the correlations (Spearman's correlation coefficients) between nDNA-encoded OXPHOS genes (intra-OXPHOS) versus their correlations with OXPHOS complexes assembly factors genes. Star: average value. (*D*) Box plots representing the correlations (Spearman's correlation coefficients) of mtDNA OXPHOS genes with nDNA OXPHOS genes versus their correlations with OXPHOS assembly factors. Red lines: median correlation coefficient; stars: average correlation coefficient. (***) $P < 1 \times 10^{-100}$.

subunits is shared by known OXPHOS complexes' assembly factors (Supplemental Table S1). First, we noticed that the correlation of tissue expression patterns among nDNA-encoded OXPHOS genes was significantly stronger than their expression correlation with the assembly factors (median intra-OXPHOS correlation = 0.623, median OXPHOS-assembly factor correlation = 0.298, MWU = 1.35×10^6, $P < 1 \times 10^{-100}$) (Fig. 2C). Similarly, the correlation of mtDNA-encoded genes expression with nDNA-encoded OXPHOS genes (median = 0.280) was significantly higher than their correlation with the OXPHOS assembly factors (median = −0.018, MWU = 4.9×10^6, $P < 1 \times 10^{-100}$) (Fig. 2D). Therefore, mito-nuclear gene expression correlation across human body sites is highly specific to structural components of the OXPHOS system.

Certain OXPHOS gene paralogs co-express with the main OXPHOS expression cluster in a tissue-dependent manner

Seven OXPHOS subunits (COX4I, COX6A, COX6B, COX7A, COX7B, COX8, and ATP5MG) have gene paralogs, of which some were previously described as tissue-specific transcripts, although in a small number of tissues (Pierron et al. 2012). To assess tissue-dependent expression of OXPHOS subunit paralogs, we first tested which of the paralogs is most correlated with OXPHOS genes expression, which tends to be tissue-dependent or tissue-specific, i.e., correlate with OXPHOS genes expression only in 1–2 tissues (Fig. 3). Notably, ATP5IF1 was excluded from further analysis as it codes for a nonstructural regulatory protein of com-

plex V (Campanella et al. 2008), in contrast to its HUGO definition as a complex V structural subunit. As a reference for each of the above-mentioned seven subunits, we assessed the tissue expression pattern of the grouped paralogs per subunit (here, termed "collapsed expression vector"). Five of the seven analyzed subunits (i.e., COX4I, COX6B, COX7B, COX8, and ATP5MG) had a single gene paralog with a significantly higher expression correlation with OXPHOS genes expression as compared to the collapsed expression vector (i.e., COX4I1, COX6B1, COX7B, COX8A, and ATP5MG). Hence, these genes should be considered part of the "main" OXPHOS gene collection. Notably, the other paralogs of these five OXPHOS genes (i.e., COX4I2, COX6B2, COX7B2, COX8C, and ATP5MG2) had only marginal expression correlation with OXPHOS genes expression in all tested tissues, thus suggesting lower importance for OXPHOS. Gene paralogs of the complex IV subunits COX6A and COX7A (COX6A1, COX6A2, COX7A1, COX7A2, and COX7A2L) had significantly lower correlation with OXPHOS genes expression as compared to their corresponding collapsed expression vectors (Fig. 3A), likely reflecting a tissue-dependent role in OXPHOS. To test for such possible tissue-dependent expression correlation, we calculated the correlation values of these paralogs with OXPHOS genes separately in each of the available 48 body sites. Firstly, we found that COX7A2L co-expresses with the rest of the OXPHOS genes in 11 out of 13 tested brain regions. Secondly, the expression of COX6A2 clustered with that of COX7A1, but not with COX6A1 and COX7A2, especially in the heart, lung, and liver, suggesting preferential coregulation of
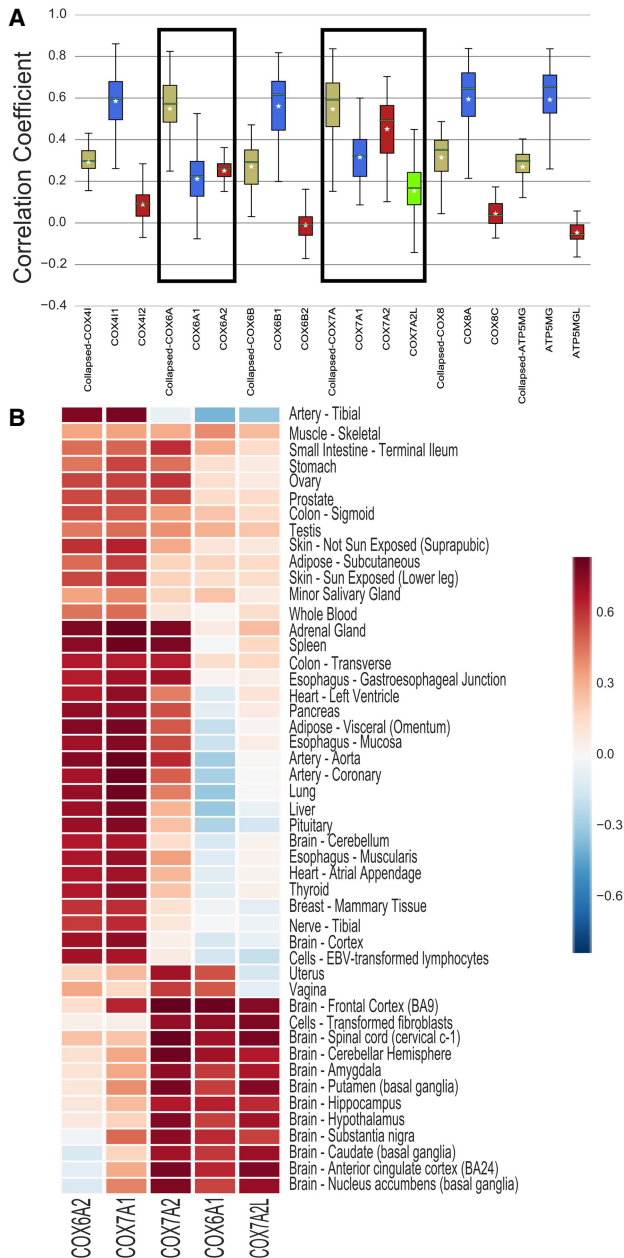
Figure 3. The various gene paralogs of *COX6A* and *COX7A* are co-expressed with OXPHOS genes in various tissues. (*A*) Box plot demonstrating the distributions of expression correlations of OXPHOS genes with the grouped (collapsed) paralogs per each of the following seven OXPHOS subunits: *COX4I*, *COX6A2*, *COX6B*, *COX7A2*, *COX7B*, *COX8*, and *ATP5MG*) (greenish gray) and the expression vectors of the indicated gene paralogs (blue, red, and green). (*B*) A heat map demonstrating OXPHOS genes expression correlation values with *COX6A* and *COX7A* paralogs, considering the 48 different tissues tested in the current study.

specific *COX6A* and *COX7A* gene paralog combinations (Fig. 3B). Accordingly, we observed high similarity of the correlation patterns of *COX7A2* and *COX6A1*, along with that of *COX7A2L*, with OXPHOS genes in most brain regions and in female reproductive organs (uterus and vagina). This suggests that the mentioned subunit paralog pairs have a mutually exclusive role in OXPHOS function across different tissues.

mtDNA and nDNA gene expression positively correlate in all body sites apart from the "primitive" brain

Although we observed positive correlation between the expression patterns of mtDNA- and nDNA-encoded OXPHOS genes, such correlation considered all tested body sites, thus masking possible inter-tissue variability. To address this caveat, we analyzed each body site separately. Such analysis revealed positive mito-nuclear correlation for all nonbrain body sites, as well as in the neocortex and the cerebellum (cerebellum and cerebellar hemisphere). These regions were expanded after the radiation of mammals and, for the sake of simplicity, are collectively termed here as "mammalian" brain (McGeer 2013; Barton and Venditti 2014). However, in the hypothalamus, basal ganglia (nucleus accumbens, putamen, and caudate), amygdala, and the upper spinal cord (cervical c-1), i.e., regions that retained a relatively similar size during vertebrate evolution (MacLean 1990; Emery and Clayton 2005; Mashour and Alkire 2013) (hereby termed "primitive" brain), we found striking negative correlation between nDNA- and mtDNA-encoded OXPHOS genes expression. Such anti-correlation between mtDNA- and nDNA-encoded OXPHOS genes expression within "primitive" brain regions (median = −0.493) was significantly more negative than the correlation between mtDNA genes expression and other randomly chosen genes in these tissues (median = −0.098, MWU = 5.47×10^7, *P*-value = 2.13×10^{-79}). In contrast to the "primitive" brain regions, the so-called "mammalian" brain regions displayed higher correlation values of mtDNA- and nDNA-encoded OXPHOS genes (median = 0.304) compared to the expression correlation of mtDNA genes with other randomly chosen genes (median = −0.001, MWU = 6.35×10^7, *P*-value < 1×10^{-100}) (Fig. 4B). Hence, the "primitive" brain displays a unique mito-nuclear co-expression pattern.

There were several exceptions to this rule. First, in the substantia nigra and anterior cingulate cortex (Brodmann area 24), some nDNA OXPHOS genes positively correlated with the mtDNA genes, whereas certain nDNA genes exhibited anti-correlation with mtDNA genes (Supplemental Data Sets S1, S2). Secondly, the hippocampus, a subcortical region which is considered part of the "primitive" brain (see definition above), revealed a generally positive mito-nuclear expression correlation, apart from four nDNA-encoded OXPHOS genes (*SDHA*, *NDUFV1*, *NDUFS2*, and *NDUFA10*), which showed somewhat negative correlation with mtDNA-encoded genes. Taken together, despite these exceptions, there is an overall strong tendency toward distinct mito-nuclear coregulation of OXPHOS gene expression in "primitive" brain regions.

OXPHOS genes from the mtDNA and nDNA are overexpressed in "primitive" brain tissues

It is possible that the negative mito-nuclear gene expression correlation is accompanied by altered absolute expression levels of both mtDNA- and nDNA-encoded OXPHOS genes in the different brain regions. To address this possibility, we first normalized the medial tissue expression values of all genes to their medial expression across all body sites. This was performed to avoid possible skew due to genes that have a tissue-independent, very high, or very low expression relative to other genes. We found significantly higher normalized expression values of mtDNA-encoded genes in the "primitive" brain (median mtDNA-gene relative expression = 2.412) compared to the rest of the body (median mtDNA-gene relative expression = 0.970, MWU = 47, *P* = 1.79×10^{-46}) and to other brain regions (median mtDNA-gene relative expression = 0.827,
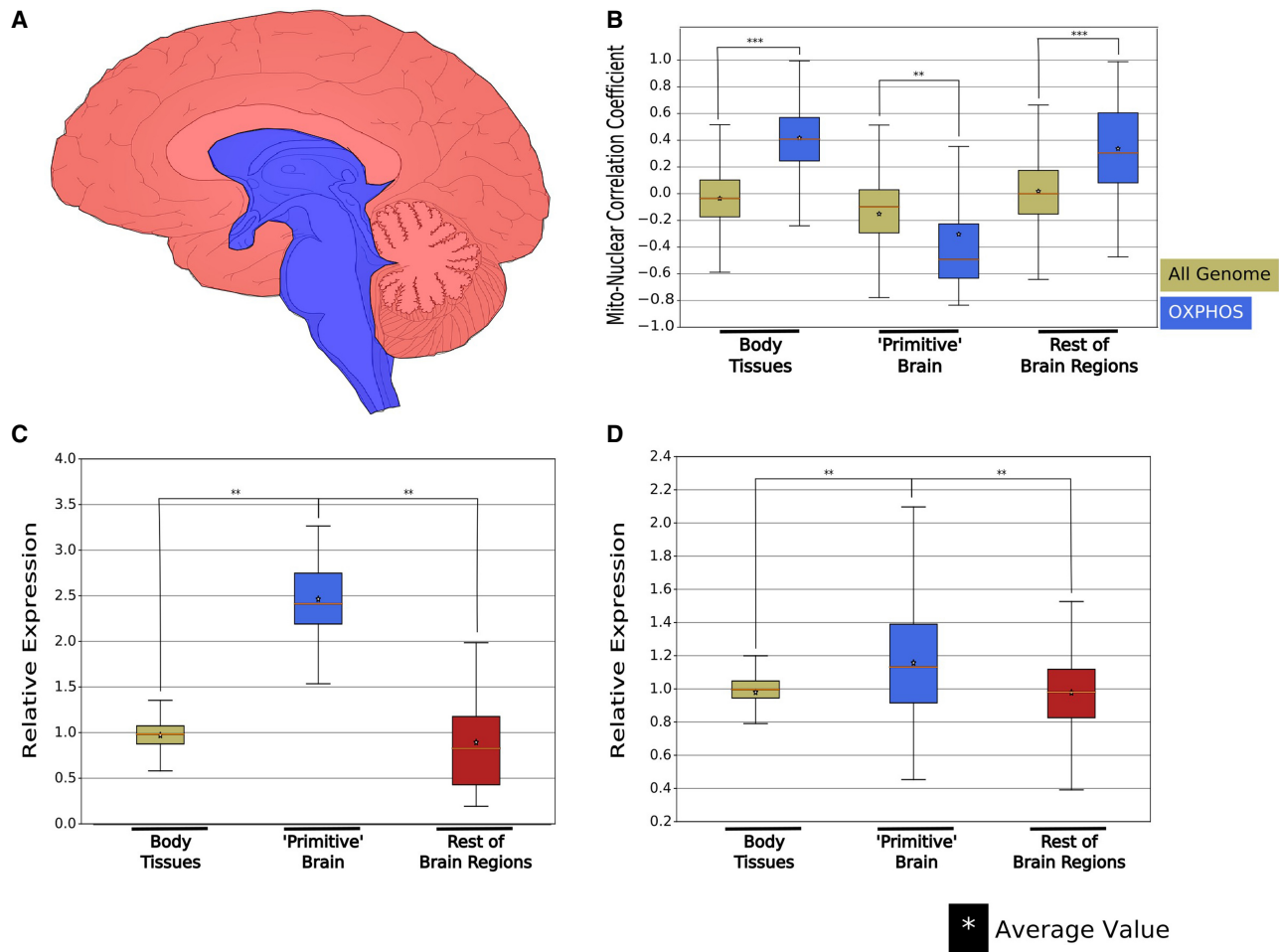
Figure 4. "Primitive" brain reveals a negative mito-nuclear expression pattern and an overall increase in OXPHOS gene expression levels. (*A*) A schematic illustration of human brain cross-section, roughly dividing the brain into associated with the "primitive" (blue) or other brain regions (red). (*B*) A box plot representing Spearman's correlation coefficients between either the entire genome (greenish gray) or nDNA-encoded OXPHOS genes with mtDNA genes (blue) in all nonbrain body sites, in an aggregate of all "primitive" brain regions (that each showed negative mito-nuclear OXPHOS genes correlations, separately) and in all other brain regions as well (see Supplemental Data Sets S1, S2). Red lines: median correlation coefficient; stars: average correlation coefficient. (***) $P < 1 \times 10^{-100}$. (*C,D*) Box plots representing the expression of mtDNA-encoded (*C*) and nDNA-encoded (*D*) OXPHOS genes relative to the median expression across all 48 tested tissues, in all nonbrain body sites (greenish-gray), in an aggregate of all "primitive" brain regions that showed (separately) negative mito-nuclear OXPHOS genes correlations (blue), and in all other brain tissues (red) (see Supplemental Data Sets S1, S2). Red lines: median correlation coefficient; stars: average correlation coefficient. (**) $P < 1 \times 10^{-10}$.

$MWU = 41$, $P = 9.69 \times 10^{-29}$) (Fig. 4A,C). A similar pattern was obtained while examining nDNA-encoded OXPHOS genes expression in the "primitive" brain (median nDNA-gene relative expression = 1.132) compared to all other brain and body regions (median nDNA-gene relative expression = 0.995, $MWU = 6.5 \times 10^5$, $P = 4.89 \times 10^{-26}$) and, specifically, to other brain regions (median nDNA-gene relative expression = 0.980, $MWU = 1.2 \times 10^5$, $P = 6.78 \times 10^{-16}$) (Fig. 4D).

These findings, along with the connection of brain disease phenotypes to OXPHOS dysfunction in "primitive" brain regions (Ikebe et al. 1990; Bossi et al. 1993; Lim et al. 2008; Lindfors et al. 2011; Pickrell et al. 2011a, b), led us to hypothesize that our observed negative mito-nuclear gene expression correlation reflects a distinct pattern of OXPHOS regulation rather than departure from the dependence on OXPHOS. This could be tested by assessing the expression of lactate dehydrogenase (LDH), a key enzyme in determining whether pyruvate enters the TCA cycle, thus pro-

moting OXPHOS, or, alternatively, whether pyruvate is fermented into lactate and sustains glycolysis. LDH forms a tetramer in two major forms—LDH-1 and LDH-5: Whereas LDH-5 contains only *LDHA*-encoded subunits and is more abundant in glycolytic tissues, LDH-1 harbors only *LDHB*-encoded subunits, which is more abundant in high OXPHOS tissues (Read et al. 2001; Porporato et al. 2011). Our analysis indicates that, while in all nonbrain and "mammalian" brain tissues, *LDHA* and *LDHB* expression levels are comparable (Supplemental Fig. S3), the "primitive" brain had a decreased *LDHA* and an increased *LDHB* expression level (~1.5- to 2.0-fold compared to the tissue expression average). This was also observed in the substantia nigra, the anterior cingulate cortex, and the hippocampus (Supplemental Fig. S3), which also showed anti-correlation between mtDNA- and nDNA-encoded OXPHOS genes expression. These findings suggest that the negative mito-nuclear gene expression correlation in the "primitive" brain reflects tight OXPHOS regulation.

Brain cell types and their spatial location modulate OXPHOS mito-nuclear coregulation

We hypothesized that the unique negative mito-nuclear co-expression pattern in subcortical brain regions (the so-called "primitive" brain), partially stems from differential cell types distribution across brain areas. This hypothesis predicts that the expression correlation between mtDNA- and nDNA-encoded OXPHOS genes differs among brain cell types. To test for this hypothesis, we used the publicly available comprehensive collection of single cell RNA-seq data from both cortical and subcortical brain regions in mouse (Allen Brain Atlas; mouse.brain-map.org). Specifically, we utilized single-cell RNA-seq data from two cortical regions (primary visual cortex [PVC] and anterior lateral motor area [ALM]) and one subcortical brain region (lateral geniculate complex [LGd]).

While considering the data from mouse, we analyzed RNA-seq experiments from 8329 PVC cells, 5992 cells from the ALM, and 1772 cells from the LGd, comprising a total of 16,093 cells. To avoid data sparsity due to known high proportions of zero read counts in single-cell RNA-seq data (Vallejos et al. 2017), we included in our analysis only OXPHOS genes whose sequencing read count was higher than zero in a minimum of 15,000 cells (Supplemental Fig. S4). This quality control step resulted in a set of 13 mtDNA- and 76 nDNA-encoded OXPHOS genes (a total of 89 OXPHOS genes). This set of genes had no zero read counts in 14,433 cells, which were used for further analyses.

To enable assessment of gene expression patterns according to brain cell types, we classified the cells into excitatory-neurons (glutamatergic), inhibitory-neurons (GABAergic), and nonneuronal cells, based on marker genes' expression patterns as previously described (help.brain-map.org/display/celltypes/Documentation?preview/8323525/10813526/CellTypes_Transcriptomics_Overview.pdf). Such classification resulted in 7800 excitatory-neurons (including 2477 from the ALM, 1097 in the LGd, and 4244 in the PVC), 6553 inhibitory-neurons (ALM: 2931, LGd: 537, and PVC: 3085), and 80 nonneuronal cells (ALM: 27, LGd: 15, and PVC: 38). The small number of nonneuronal cells, mostly comprising glia (>97%) (Lauwers et al. 2008), may either reflect their known low representation (~26%) in the mouse brain (Herculano-Houzel 2014) and\or sampling bias. Hence, to avoid sample size issues, we grouped the nonneuronal cells from all brain regions for further analyses.

As the single-cell RNA-seq data originated from mouse, we asked whether the human mito-nuclear correlation landscape in the brain is similar to that of the mouse. We therefore first compared the OXPHOS mito-nuclear gene expression correlation in the three available brain regions (ALM, PVC, and LGd), while grouping the cells (bulk analysis) according to these regions. Our results indicate a significantly lower medial OXPHOS mito-nuclear expression correlation value in the subcortical (LGd) relative to both cortical regions (LGd-ALM: median LGd correlation = 0.102, median ALM correlation = 0.274, MWU = 2.2×10^5, $P < 1 \times 10^{-100}$; LGd-PVC: median PVC correlation = 0.239, MWU = 2.5×10^5, $P < 2.19 \times 10^{-80}$) (Fig. 5A). Thus, similar to humans, subcortical mouse brain tends toward negative mito-nuclear OXPHOS gene expression correlation.

Next, we analyzed the OXPHOS mito-nuclear expression correlation values of the different brain cell types. We found an overall negative and significantly lower mito-nuclear expression correlation values in the nonneuronal cell population (medi-
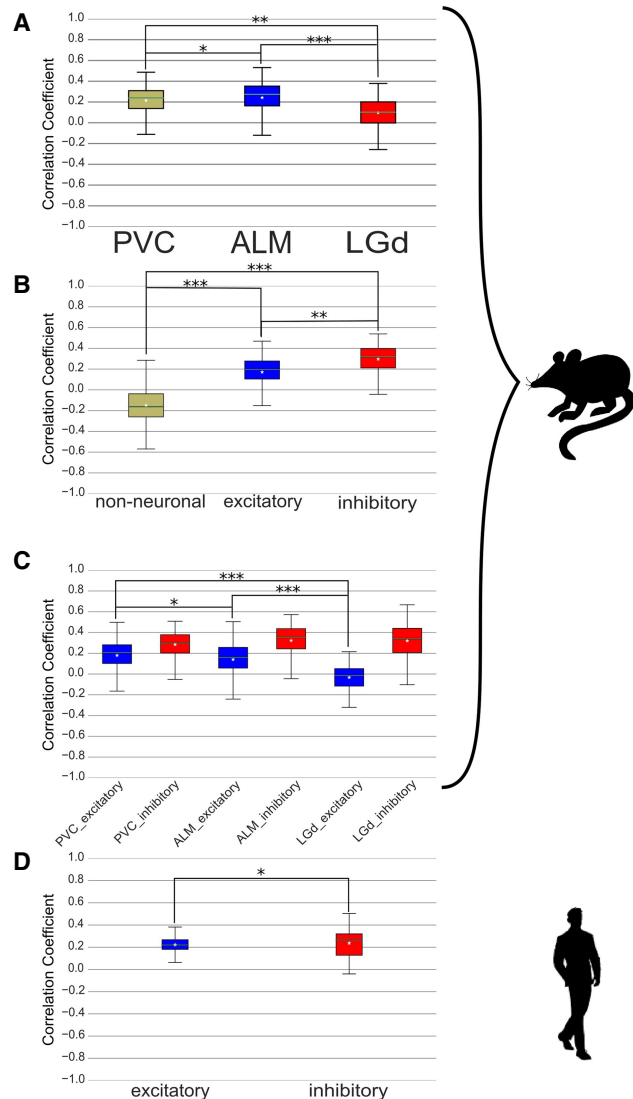

Figure 5. Single-cell RNA-seq analysis indicate spatial and cell-type–dependent effect on OXPHOS mito-nuclear genes co-expression. (*A*) Box plots showing Spearman's rank OXPHOS mito-nuclear expression correlation values in a bulk of mouse brain cells, grouped according to their brain region collection site. Star: average value (in all subsequent panels). (*B*) Box plots showing Spearman's rank OXPHOS mito-nuclear expression correlation values in mouse brain cells, bulked according to their assigned cell type. (*C*) Box plots showing Spearman's rank OXPHOS mito-nuclear expression correlation values in mouse brain neurons, bulked according to brain regions and neuronal type. (*D*) Box plots showing Spearman's rank OXPHOS mito-nuclear expression correlation values in human brain cells (nuclei), bulked according to their assigned neuronal cell type.

an correlation = −0.161) relative to both the excitatory-neurons (median correlation = 0.195, MWU = 7.3×10^4, $P < 1 \times 10^{-100}$) and the inhibitory-neurons (median correlation = 0.320, MWU = 1.8×10^4, $P < 1 \times 10^{-100}$). Secondly, we found significantly higher mito-nuclear expression correlation values in the inhibitory-neurons relative to the excitatory-neurons (MWU = 2.5×10^5, $P = 4.33 \times 10^{-78}$) (Fig. 5B).

Despite the relative paucity of nonneuronal cells from the LGd in the current study (~1%), we observed a significantly reduced mito-nuclear OXPHOS expression correlation in this region

(Fig. 5A), suggesting that certain neuronal cell types contributed to this phenomenon. If this is true, then not only cell-type composition, but also the brain regional cells' location, affect the mito-nuclear OXPHOS gene expression. To test for that, we analyzed OXPHOS mito-nuclear gene expression correlations of the two neuronal cell types (excitatory and inhibitory neurons), divided into the three available mouse brain regions. Notably, as mentioned above, nonneuronal cells were not subjected to such analysis due to small overall sample size ($N = 80$). We found a significant reduction of mito-nuclear OXPHOS gene expression correlation values of excitatory-neurons in the subcortical region (LGd) to a negative median correlation of −0.015. This suggests that more than 50% of the mito-nuclear gene expression correlations of excitatory-neurons in the LGd were negative. The inhibitory-neuronal populations, however, showed positive medial correlation values (between 0.29 and 0.35) across all three brain regions (Fig. 5C). Taken together, the negative OXPHOS mito-nuclear gene expression correlation in the mouse brain subcortical regions likely results from both cell-type composition variation across brain regions and from spatial effects. As our bulk RNA-seq analysis indicated relatively negative mito-nuclear gene expression correlation in subcortical mouse brain regions, we are tempted to speculate that the cell type and spatial impact may also explain our observed negative mito-nuclear gene expression correlation in human subcortical ("primitive") brain regions.

Human single-cell RNA-seq data became recently available from cell nuclei isolated from cortical brain regions. Firstly, we used human single-nuclei RNA-seq data from the middle temporal gyrus cortical area (MTG) (Allen Brain Atlas). To this end, and similar to the mouse analysis, we divided the cells into cell types (inhibitory neurons, excitatory neurons, and nonneuronal cells), using human orthologs of the mouse cell-type marker genes. Such division revealed that <6% of the tested nuclei (913 out of 15,928) were considered nonneuronal cells. Hence, similar to the mouse, there is a dramatic underrepresentation of this cell type compared to their expected ~50% representation in the human brain (Azevedo et al. 2009). Notably, the mtDNA mapped read count was low, probably due to the analysis of human cell nuclei rather than whole cells. Only four mtDNA-encoded OXPHOS genes had no-zero read counts in at least 10,700 isolated nuclei. We found that 28 nDNA-encoded OXPHOS genes fulfilled the same criterion. Nevertheless, while sifting for cell nuclei that had no-zero read counts in both the four mtDNA- and the 28 nDNA-encoded OXPHOS genes (a total of 32 OXPHOS genes), only 710 nuclei remained. After dividing these remaining nuclei into cell types (i.e., excitatory and inhibitory neurons, as well as nonneuronal cells), only five nuclei could be assigned to nonneuronal cells, thus excluding them from further analysis. Similar to the mouse, the majority of the human nuclei (627 out of 710) were defined as excitatory-neurons, with the rest (78 nuclei) being inhibitory-neurons. Our analysis indicated that despite the relatively smaller sample size and small number of calculated gene expression correlations (112 mito-nuclear correlations in human nuclei as compared with 988 correlations in the mouse), the human inhibitory-neuron nuclei exhibited a significantly higher median OXPHOS mito-nuclear expression correlation value (median correlation = 0.265) compared to the nuclei from excitatory-neurons (median correlation = 0.220, MWU = 5.4×10^3, $P = 0.043$) (Fig. 5D). Thus, although this analysis is limited to the human cortex, it further suggests that differential cell-type distributions across brain regions partially contributes to

our observed variation in mito-nuclear OXPHOS co-expression in the brain.

Candidate regulatory factors that best explain the coregulation of nDNA-encoded OXPHOS genes

As the first step to identifying candidates contributing to the common mechanism underlying the co-expression of OXPHOS genes, we aimed at identifying *cis*-regulatory elements upstream of all nDNA-encoded OXPHOS genes. To this end, we analyzed DNase-seq experimental results from 125 human cell lines generated by the Encyclopedia of DNA Elements Consortium (ENCODE) and identified DNase genomic footprints (DGFs) in a window of 1000 base pairs (bp) upstream of the transcription start site (TSS) of all nDNA-encoded OXPHOS genes (Supplemental Fig. S5). Then, to correlate in vivo TF binding sites to the identified DGF sites, we analyzed available chromatin immunoprecipitation (ChIP)-seq results for 161 TFs from ENCODE and assessed the overlap between the above-mentioned DGF sites and ChIP-seq peaks (Supplemental Tables S5–S7; Supplemental Fig. S5). Although our results supported some previous in silico predictions (van Waveren and Moraes 2008), there were quite a few disparities. For example, van Waveren and Moraes (2008) suggested that only a low percentage (23%) of OXPHOS gene promoters harbored TATA-box binding protein (TBP) motifs and further suggested that there should be a requirement for binding of additional factors to promote transcription in TATA-less promoters. Although DGF sites upstream of many OXPHOS genes were occupied by Sp1 transcription factor (SP1) and YY1 transcription factor (YY1) (68% and 86%, respectively), which associate with transcription of TATA-less promoters, the vast majority of OXPHOS genes' upstream DGF sites harbored strong binding signals for TBP and for the TATA-box binding protein associated factor 1 (TAF1) (91% and 89%, respectively). Hence, in contrast to previous in silico analysis, our in vivo analysis reveals that TATA-binding proteins likely participate in regulating OXPHOS genes transcription.

To test for the impact of TF binding on the correlation coefficient between pairs of in vivo-bound nDNA OXPHOS genes (ChIP-seq), we used a least-squares linear regression model. Such analysis enabled the identification of 15 TFs, which bound candidate *cis*-regulatory elements of OXPHOS genes, and revealed a significant positive impact on the Spearman's rank correlation coefficient between the expression patterns of OXPHOS gene pairs (Supplemental Table S7). Firstly, this list included POLR2A, which in vivo bound 93% of the OXPHOS gene promoters and had a very high impact on the OXPHOS expression correlation coefficient (ρ_s; 0.57 Spearman correlation units [SCU]). Secondly, we found in vivo binding signals for the histone lysine demethylase, PHD finger protein 8 (PHF8), and MYC associated zinc finger protein (MAZ) in 86% and 81% of the OXPHOS gene promoters, respectively. Such abundant binding signals were accompanied by relatively high impact on the ρ_s of OXPHOS genes, which was positively affected (0.17 and 0.056 SCU, respectively). Thirdly, three AP-1 transcription factors had a significantly positive impact on OXPHOS gene expression correlations. Specifically, JUND, FOS, and FOSL2 bound 73%, 36%, and 25% of the candidate OXPHOS gene promoters, respectively. These TFs also significantly affected OXPHOS gene expression pattern correlations (0.042, 0.036, and 0.048 SCU, respectively) (Fig. 6A). Taken together, our analysis identified in vivo *cis*-regulatory elements in all nDNA-encoded OXPHOS genes and underlined the best candidates for mediating their co-expression.
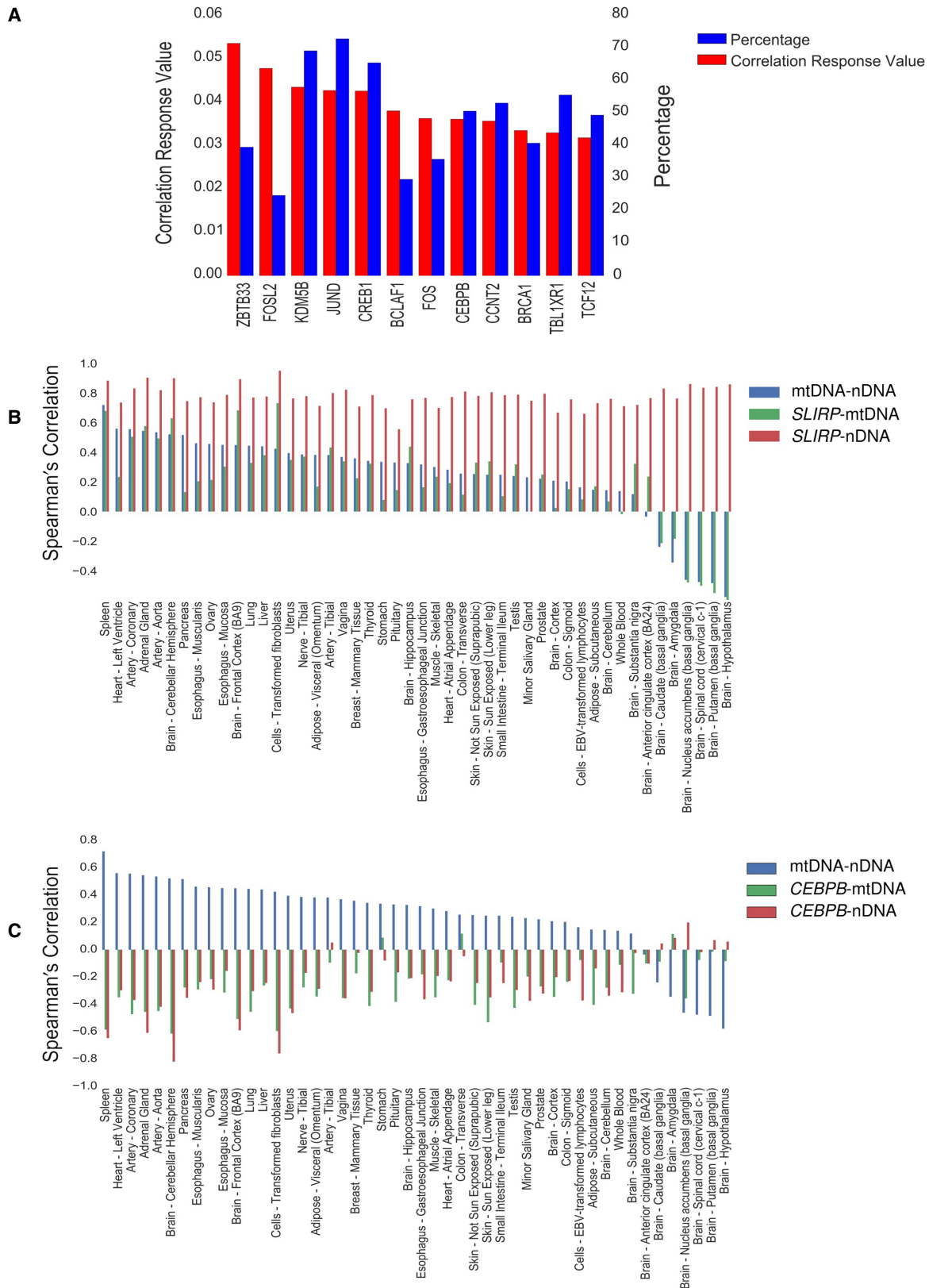
Figure 6. Candidate factors that best explain mito-nuclear coregulation. (*A*) Bar plot showing 15 transcription factors that bind *cis*-regulatory elements of nDNA-encoded OXPHOS genes in vivo and correlate with their co-expression. (*B*,*C*) Bar plots of the expression correlations between either the genes encoding mtRNA-binding protein SLIRP (*B*) or the mtDNA-binding transcription factor CEBPB (*C*) with the median expression values of mtDNA- (green bars) and nDNA-encoded (red bars) OXPHOS genes. Blue bars represent the median value of correlation between mtDNA- and nDNA-encoded OXPHOS genes.

Twelve RNA-binding proteins and CEBPB are the best candidates to explain the co-expression pattern of mtDNA- and nDNA-encoded OXPHOS genes across human body sites

We next sought the best candidate factors to explain the co-expression of mtDNA and nDNA OXPHOS genes. To this end, we focused our analysis on a set of genes encoding both TFs and predicted RNA-binding proteins with known mitochondrial localization in human cells. Notably, the listed TFs included were shown to bind the mtDNA in vivo (Chen et al. 2004; Reyes et al. 2011; She et al. 2011; Blumberg et al. 2014; Wolf and Mootha 2014). To identify the factors that best explain the increase in mito-nuclear gene expression correlation, we sought factors whose expression positively correlated with OXPHOS genes encoded by both the mtDNA and the nDNA in most tissues, except for "primitive" brain regions (in which mito-nuclear expression correlation was negative). Twelve genes which encode predicted RNA-binding proteins (*ACO2*, *SLIRP*, *TIMM8B*, *MRPL57*, *c6orf203*, *ENDOG*, *MRPS17*, *MRPS7*, *COQ3*, *GAT3A*, *DLD*, and *TACO1*) showed expression correlation with both nDNA- and mtDNA-encoded OXPHOS genes. Although these genes positively correlated with OXPHOS gene expression from both genomes in most body sites, in the "primitive" brain, their expression negatively correlated with the mtDNA genes while maintaining positive correlation with nDNA-encoded genes (Supplemental Fig. S6). Notably, one of these genes (*SLIRP*) (Fig. 6B) was previously shown to promote the stability of mRNA transcripts of both nDNA- and mtDNA-encoded OXPHOS genes (Baughman et al. 2009). We therefore suggest that products of these genes serve as the best candidates to participate in post-transcriptional regulation of mtDNA OXPHOS gene expression and their putative participation in mito-nuclear coregulation.

The expression pattern of a single gene in our list of 117 genes, *CEBPB*, was negatively correlated with the expression of both mtDNA and nDNA OXPHOS genes in most human tissues. Such negative expression correlation with genes of both genomes tended to be stronger in body sites with highly positive mito-nuclear expression correlation, such as spleen, adrenal gland, and the cerebellar hemisphere, but was lost specifically in "primitive" brain regions (Fig. 6C). *CEBPB* encodes for the CCAAT enhancer binding protein beta (CEBPB), which was shown to modulate gene expression regulation of nDNA-encoded OXPHOS genes (Brady et al. 1992) and bind the mtDNA in vivo (Blumberg et al. 2014). Our DNase-seq and ChIP-seq analysis of 1000-bp regions upstream of all nDNA-encoded OXPHOS genes' TSSs (Supplemental Fig. S5), revealed in vivo CEBPB binding to 51% of the candidate promoter regions upstream of nDNA OXPHOS genes. It also had a significantly positive effect on mito-nuclear gene expression correlation across human tissues (0.036 SCU). We, therefore, argue that CEBPB is likely a candidate repressor of OXPHOS gene expression that concomitantly acts on both genomes to coordinate OXPHOS gene expression.

The landscape of OXPHOS subunit splice variants and their tissue expression patterns

So far we identified co-expression of mtDNA- and nDNA-encoded OXPHOS genes across human body sites. We also identified subunit gene-paralogs that either co-express or have tissue-dependent correlation with the major OXPHOS gene expression cluster. A yet understudied source of variation in OXPHOS subunit composition, which might be tissue-dependent, is alternatively spliced variants. Recent identification of a tissue-dependent inclusion of *NDUFV3* splice isoforms in OXPHOS complex I (Bridges et al.

2017; Dibley et al. 2017; Guerrero-Castillo et al. 2017) urged us to assess the extent of such phenomena in the entire OXPHOS system. To this end, we analyzed the expression correlation patterns of in vivo-identified splice variants of all nDNA-encoded OXPHOS genes across and within the 48 body sites analyzed in the current study, while employing the same analysis steps used for the above-described gene expression patterns. Such analysis underlined the main expression cluster of OXPHOS gene transcripts (Supplemental Fig. S7). We found that this cluster harbors at least one splice isoform representing the known canonical proteins (UniProt) of 43 out of the 87 nDNA-encoded OXPHOS genes (here, termed "canonical transcripts"). This was defined as the "main OXPHOS transcript cluster" and was used as a reference for further identification of both cross-tissue and tissue-dependent expression patterns of OXPHOS subunit splice isoforms (Supplemental Data Sets S1, S2; Supplemental Fig. S7). Firstly, we found 15 genes for which a particular, noncanonical transcript (splice variant) co-expresses with the "main OXPHOS transcript cluster" (i.e., *ATP5F1A*, *ATPF51D*, *ATP5MC1*, *ATP5PF*, *ATP5MF*, *ATP5PO*, *COX6B1*, *NDUFA2*, *NDUFA3*, *NDUFA5*, *NDUFA6*, *NDUFB11*, *NDUFB6*, *NDUFV1*, and *UQCRB*). After verifying the presence of a mitochondrial localization signal peptide in these splice isoforms (TargetP 1.1; www.cbs.dtu.dk/services/TargetP/), we identified five OXPHOS subunits (*ATP5MF*, *NDUFA2*, *NDUFB11*, *NDUFB6*, and *NDUFA3*) having alternatively spliced variants that co-expressed with the "main OXPHOS transcript cluster" instead of the known canonical transcripts (Table 1). This offers alternative canonical transcripts for these five subunits. To assess the impact of the translation products of these transcripts on the structure of their corresponding OXPHOS complexes (complexes I and V), we compared their secondary structures with the canonical proteins or (when available) their 3D structure predictions while incorporating either the canonical or the alternatively suggested splice isoforms in the relevant OXPHOS complexes (Supplemental Fig. S8). Our analysis indicated that the translation product of the alternatively spliced variant of *NDUFA2* is predicted to lack a beta-strand and an alpha-helix in its C-terminus, yet the role of these elements in NDUFA2 interactions with its neighboring subunit (NDUFS1) is yet unclear. Secondly, the translation product of the alternatively spliced variant of *NDUFB11* displayed no predicted changes of the two alpha-helixes that participate in its interactions with either ND4, NDUFB10, or NDUFB8. Similarly, the translation product of the alternatively spliced variant of *NDUFB6* still contains all three helixes that participate in its interaction with ND5. Thirdly, although the alternatively spliced variant of *NDUFA3* underwent major sequence changes compared to its known canonical transcript, its translation product still retained the alpha-helixes that physically interact with ND1. Finally, although the product of the alternatively spliced variant of *ATP5MF* lacks five residues compared to its known canonical isoform (positions 5–10), this change did not significantly alter the protein's secondary structure. Taken together, our results support incorporating the noncanonical splice isoforms of *ATP5MF*, *NDUFA2*, *NDUFB11*, *NDUFB6*, and *NDUFA3* as revised canonical transcripts, replacing the current canonical ones (Table 1).

We next aimed at assessing tissue-dependent co-expression of the different splice-variants with the "main OXPHOS transcript cluster." To this end, we screened for splice variants whose median correlation vectors with the "main OXPHOS transcript cluster" were higher as compared to other splice isoforms of the same subunit in at least one tissue (Fig. 7; Supplemental Data Set S3).

Table 1. The revisited list of canonical OXPHOS splice variants

Gene	Revised canonical transcript (Ensembl ID)	Revised inferred protein (UniProt ID)	Previous canonical transcript (Ensembl ID)	Previous inferred protein (UniProt ID)
ATP5MF	ENST00000394186.3	P56134-2	ENST00000292475.7	P56134-1
NDUFA2	ENST00000512088.1	O43678-2	ENST00000252102.8	O43678-1
NDUFA3	ENST00000391763.3	A8MUI2-1	ENST00000485876.5	O95167-1
NDUFB11	ENST00000276062.8	Q9NX14-2	ENST00000377811.3	Q9NX14-1
NDUFB6	ENST00000350021.2	O95139-2	ENST00000379847.7	O95139-1

See also Supplemental Figure S8.

Notably, below-zero median correlations were considered zero for the sake of this analysis. Our results identified tissue-dependent expression of several splice isoforms of OXPHOS subunits, which are best exemplified by NDUFV3 and ATP5F1C. First, NDUFV3 encodes for two splice isoforms differing in ~300 amino acids (aa) (ESNT00000340344.4, encoding NDUFV3-S, and ESNT00000354250.6, encoding NDUFV3-L), which were recently thoroughly studied (Bridges et al. 2017; Dibley et al. 2017; Guerrero-Castillo et al. 2017). Our screen revealed that, in most examined human tissues (especially in the coronary artery and the aorta), NDUFV3-S displayed a relatively higher correlation with the "main OXPHOS transcript cluster" as compared to NDUFV3-L (Fig. 7A). Secondly, ATP5F1C displayed two splice variants (which differ in a single C-terminal aspartate residue), of which one was previously described as liver-specific (ENST00000356708.11, encoding ATP5F1C-L) and the other was considered heart-specific (ENST00000335698.4, encoding ATP5F1C-H). Our analysis revealed that the expression of ATP5C1-H correlates with the "main OXPHOS transcript cluster" in most tissues, while in many brain regions, skeletal muscle, testis, uterus, and vagina, the correlation of ATPF15C-L with the main OXPHOS cluster was more prominent (Fig. 7B).

Discussion

In the current study, we quantitatively analyzed the RNA expression of OXPHOS genes encoded by the mtDNA and nDNA in 48 human body sites. First, positive correlation of gene expression was evident when considering all tested body sites grouped together. Tissue-by-tissue analysis revealed that all nonbrain body sites, separately, retained this positive mito-nuclear OXPHOS gene expression correlation despite known variation in various other parameters of mitochondrial function, such as mtDNA copy number (Fernandez-Vizarra et al. 2011).

Nevertheless, striking negative correlation of mtDNA- and nDNA-encoded OXPHOS genes was found in subcortical brain regions (the so-called "primitive" brain). This finding drew our attention, since the brain regions that were especially enlarged during primate radiation, i.e., the cortex, cerebellar hemispheres, and cerebellum (the so-called "mammalian" brain), displayed positive mtDNA-nDNA OXPHOS expression correlation. This suggests that adjustment of mito-nuclear coregulation occurred in the "mammalian" brain during evolution, possibly in adaptive response to its different bioenergetics requirements. Such findings may shed new light on the molecular basis of neurodegenerative diseases, such as Alzheimer's disease (Liang et al. 2008), especially since mitochondrial gene expression alterations have been reported in various cortical tissues of Alzheimer's disease patients.

What could underlie the negative mito-nuclear co-expression pattern in "primitive" brain regions? Our single-cell RNA-seq data analysis in the mouse brain revealed that the mito-nuclear co-expression pattern both differs among cell types and was influenced by the spatial location of cell types in the brain. It is thus conceivable that both cell-type distribution and cell-cell communication play a role in modulating patterns of mito-nuclear gene co-expression across the brain. Since human subcortical brain regions comprise mostly nonneuronal (glia) cells (90%) (Azevedo et al. 2009), we speculate that overrepresentation of glia cells in such brain regions likely contributes to the negative OXPHOS mito-nuclear expression correlation in man and mouse. This interpretation awaits future assessment of mito-nuclear co-expression patterns in human nonneuronal cells.

What could be the outcome of the negative mtDNA-nDNA gene expression correlation in the "primitive" brain? Analysis of mitochondrial OXPHOS activity in mouse tissues revealed a significantly higher ratio between the activity of complex I and the activity of complex II in the cortex than in the hypothalamus (Burtscher et al. 2015). Since, unlike complex II, complex I harbors subunits encoded by both the mitochondrial and nuclear genomes, the negative mito-nuclear coregulation in the "primitive" brain regions may be accompanied by changes in mitochondrial activity. Secondly, mtDNA depletion in the rat striatum (a "primitive" brain region) led to a significantly worse mitochondrial phenotype compared to mtDNA depletion in the rat neocortex (Pickrell et al. 2011a). These pieces of evidence suggest that "primitive" brain regions are more sensitive to functional changes in assembled OXPHOS complexes I, III, IV, and V and supercomplexes, which are likely limited by the negative mito-nuclear OXPHOS correlation in the "primitive" brain. This interpretation is consistent with the higher vulnerability of these brain regions in the context of mitochondrial defects, such as the striatum of Huntington's disease patients (Coyle and Schwarcz 1976). Nevertheless, a smaller possible range of active OXPHOS complexes in the "primitive" brain cannot be intuitively accepted in light of the extremely oxidative nature of this brain region. This is partially reconciled by our observed higher expression levels of both mtDNA- and nDNA-encoded OXPHOS genes in "primitive" brain regions. Hence, the negative mito-nuclear expression correlation may likely serve to tightly regulate and maintain high OXPHOS levels, rather than reflecting glycolysis preference. Accordingly, we found that the ratio between the expression levels of LDH subunits (LDHA versus LDHB) was dramatically lower in "primitive" brain tissues compared to all other body sites. Since the protein product of LDHA is more efficient than that of LDHB in converting pyruvate-to-lactate at glycolytic conditions (Read et al. 2001; Porporato et al. 2011), we suggest that "primitive" brain tissues have a lower tendency to convert pyruvate into lactate, suggesting higher dependence on oxidative metabolism in the "primitive" brain. Such interpretation is supported by the observed lower tendency to accumulate lactate in "primitive" brain sites (brainstem
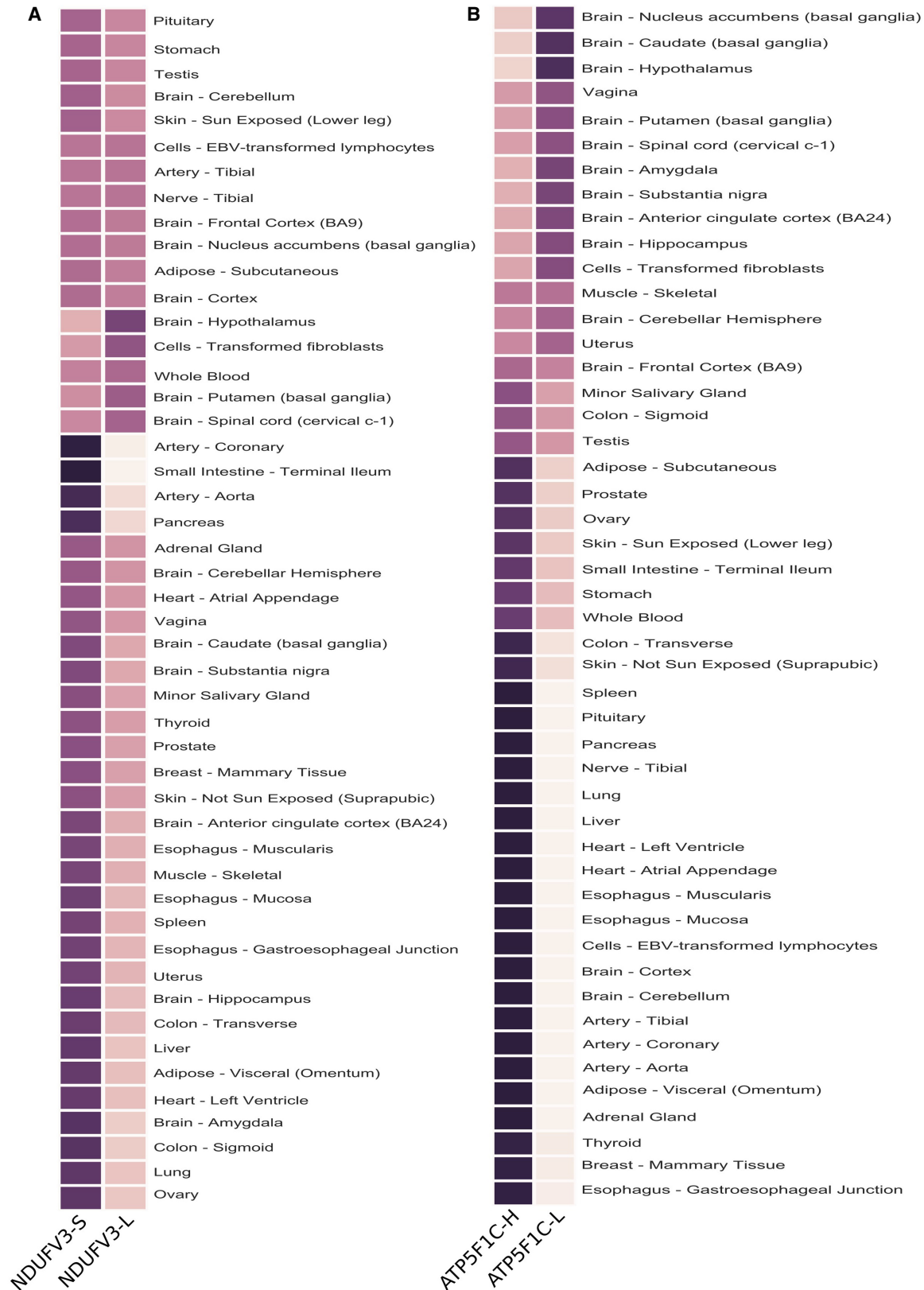
A

		Tissue
		Pituitary
		Stomach
		Testis
		Brain – Cerebellum
		Skin – Sun Exposed (Lower leg)
		Cells – EBV-transformed lymphocytes
		Artery – Tibial
		Nerve – Tibial
		Brain – Frontal Cortex (BA9)
		Brain – Nucleus accumbens (basal ganglia)
		Adipose – Subcutaneous
		Brain – Cortex
		Brain – Hypothalamus
		Cells – Transformed fibroblasts
		Whole Blood
		Brain – Putamen (basal ganglia)
		Brain – Spinal cord (cervical c-1)
		Artery – Coronary
		Small Intestine – Terminal Ileum
		Artery – Aorta
		Pancreas
		Adrenal Gland
		Brain – Cerebellar Hemisphere
		Heart – Atrial Appendage
		Vagina
		Brain – Caudate (basal ganglia)
		Brain – Substantia nigra
		Minor Salivary Gland
		Thyroid
		Prostate
		Breast – Mammary Tissue
		Skin – Not Sun Exposed (Suprapubic)
		Brain – Anterior cingulate cortex (BA24)
		Esophagus – Muscularis
		Muscle – Skeletal
		Esophagus – Mucosa
		Spleen
		Esophagus – Gastroesophageal Junction
		Uterus
		Brain – Hippocampus
		Colon – Transverse
		Liver
		Adipose – Visceral (Omentum)
		Heart – Left Ventricle
		Brain – Amygdala
		Colon – Sigmoid
		Lung
		Ovary

NDUFV3-S NDUFV3-L

B

		Tissue
		Brain – Nucleus accumbens (basal ganglia)
		Brain – Caudate (basal ganglia)
		Brain – Hypothalamus
		Vagina
		Brain – Putamen (basal ganglia)
		Brain – Spinal cord (cervical c-1)
		Brain – Amygdala
		Brain – Substantia nigra
		Brain – Anterior cingulate cortex (BA24)
		Brain – Hippocampus
		Cells – Transformed fibroblasts
		Muscle – Skeletal
		Brain – Cerebellar Hemisphere
		Uterus
		Brain – Frontal Cortex (BA9)
		Minor Salivary Gland
		Colon – Sigmoid
		Testis
		Adipose – Subcutaneous
		Prostate
		Ovary
		Skin – Sun Exposed (Lower leg)
		Small Intestine – Terminal Ileum
		Stomach
		Whole Blood
		Colon – Transverse
		Skin – Not Sun Exposed (Suprapubic)
		Spleen
		Pituitary
		Pancreas
		Nerve – Tibial
		Lung
		Liver
		Heart – Left Ventricle
		Heart – Atrial Appendage
		Esophagus – Muscularis
		Esophagus – Mucosa
		Cells – EBV-transformed lymphocytes
		Brain – Cortex
		Brain – Cerebellum
		Artery – Tibial
		Artery – Coronary
		Artery – Aorta
		Adipose – Visceral (Omentum)
		Adrenal Gland
		Thyroid
		Breast – Mammary Tissue
		Esophagus – Gastroesophageal Junction

ATP5F1C-H ATP5F1C-L

Figure 7. Tissue-dependent expression correlation of *NDUFV3* and *ATP5F1C* splice-isoforms with the main OXPHOS cluster. (*A,B*) Heat maps representing relative expression correlations between *NDUFV3* (*A*) and *ATP5F1C* (*B*) splice-variants and the "main OXPHOS transcript cluster."

and hypothalamus), relative to the cortex in rats after exercise (Takimoto and Hamada 2014). Nevertheless, this interpretation requires further investigation.

Mito-nuclear expression coregulation across human body sites suggests an active regulatory phenomenon. Our expression pattern analysis of all known transcriptional and post-transcriptional regulators of mitochondrial gene expression, as well as the preferential in vivo binding of TFs at upstream *cis*-regulatory elements of OXPHOS genes, revealed candidate factors that best explain our observed cross-tissue OXPHOS genes co-expression pattern. Hence, regulatory cross-talk between the mtDNA and nDNA is imperative, which may likely be partially mediated via anterograde and/or retrograde signaling (Chae et al. 2013). Such signals may participate in the recently discovered translocation of nuclear transcription regulators into mammalian mitochondria, their direct mtDNA binding, and their regulation of mtDNA transcription (She et al. 2011; Blumberg et al. 2014). The latter serve as the best candidates to regulate mito-nuclear gene expression, in general, and particularly transcription of OXPHOS genes in the two genomes. Still, further work is required to identify the mechanism that governs this phenomenon.

Despite the overall co-expression of OXPHOS genes, distinct clustering of mtDNA gene expression was identified in all tested tissues. This is not surprising, as known core-regulators of mtDNA gene expression are dedicated to mitochondrial function and have no clear impact on nuclear gene regulation (Quirós et al. 2016). While focusing solely on mtDNA genes, *MT-ND6* had lower correlation scores with the rest of the genes encoded by this genome, consistent with *MT-ND6* transcript being the only mRNA encoded by the light mtDNA strand, which in humans has a separate promoter from the heavy mtDNA strand, that harbors most mtDNA genes. Accordingly, recent analysis of nascent RNA transcripts revealed quantitative differences between the transcription levels of the light and heavy mtDNA strands among human cell lines (Blumberg et al. 2017). In an evolutionary perspective, it would be of great interest to assess correlation within mtDNA gene expression in birds and amphibians, where a bidirectional mtDNA promoter governs the transcription of both mtDNA strands (Bogenhagen and Yoza 1986; L'Abbé et al. 1991). As our expression correlation was observed at the population level, this analysis awaits the availability of RNA-seq from multiple unrelated individuals from these taxa.

It is well known that certain OXPHOS subunits have gene paralogs. Our analysis indicated that, while some OXPHOS subunit paralogs co-express with the OXPHOS genes in many tissues, some are tissue-dependent and others had only residual expression correlation with OXPHOS genes. This suggests that, during the course of evolution, the emergence of such gene duplicates was followed by selective pressure to either retain their OXPHOS activity, at least in certain tissues, or to adopt a different function. For example, it has been suggested that one such OXPHOS subunit paralog, COX7A2L, is pivotal to the assembly of complex III-complex IV super-OXPHOS complex in mouse fibroblasts (Lapuente-Brun et al. 2013), although others questioned this conclusion (Mourier et al. 2014). Nevertheless, recent findings support the importance of COX7A2L for respirasome formation in mouse kidney, liver, lung, and brain, but not in the heart and skeletal muscle (Cogliati et al. 2016), suggesting that some of the above disagreement stems from a tissue-dependent OXPHOS function of this subunit. These findings are consistent with our observed tissue-dependent expression correlation of *COX7A2L* with the OXPHOS genes. Additionally, our results demonstrate that specific gene-

paralog combinations of COX7A and COX6A correlated with the expression of OXPHOS genes in a tissue-dependent manner. As specific paralog combinations of these OXPHOS subunits, namely COX6A2 and COX7A1, are more abundant in complex IV dimers (Cogliati et al. 2016), we are tempted to suggest that such protein composition is already regulated at the transcript levels. Taken together, *COX6A* and *COX7A* gene paralogs, and specifically *COX7A2L*, exemplify how certain gene paralogs retain OXPHOS-related function in a tissue-dependent fashion.

RNA-seq analysis enabled the identification of both canonical and noncanonical splice variants of all nDNA-encoded OXPHOS genes and assessment of their expression patterns in human body sites. Our analysis showed that, while most OXPHOS genes co-expressed in most tissues, such co-expression clusters were not always represented by a combination of the known canonical gene transcripts. Close inspection revealed that five genes encoding OXPHOS subunits, including *ATP5MF*, *NDUFA2*, *NDUFB11*, *NDUFB6*, and *NDUFA3*, were mostly represented by a different, noncanonical splice variant, resulting in a noncanonical protein. Therefore, based on clustering and abundance in the different tissues, we offer a revised list of OXPHOS canonical transcripts, which will be useful for future analyses of OXPHOS genes (Table 1). We also found that, in some genes, such as *NDUFV3* and *ATP5F1C*, different splice-variants correlate with the expression of OXPHOS transcripts in a tissue-dependent manner. This suggests that the subunit composition of several OXPHOS complexes varies among tissues. Such hypothesis is supported by the existence of tissue-specific paralogs of cytochrome C-oxidase (complex IV) (Hüttemann et al. 2003; Kadenbach and Huettemann 2015) and by the recent identification of a tissue-specific splice isoform of the complex I subunit, NDUFV3, also at the protein level (Bridges et al. 2017; Dibley et al. 2017; Guerrero-Castillo et al. 2017). Notably, such interpretation awaits proteomics analysis of multiple human tissues that potentially will differentiate between paralogs and splice isoforms. Regardless, our analyses provide a first glance into the landscape of the tissue expression pattern of OXPHOS subunits, their splice variants and that of subunits paralogs, thus providing first assessment of this phenomenon across the human body.

In summary, to the best of our knowledge, we have presented here the most comprehensive analysis to date of the OXPHOS transcription regulatory landscape across multiple human body sites, involving both cross-tissue and single-cell RNA-seq analyses. Our results strongly support mito-nuclear gene expression regulation across tissues and revealed the best transcriptional and post-transcriptional candidates to underlie such bigenomic regulation. The unexpected altered pattern of mito-nuclear coregulation in the "primitive" brain lends new insight into variation in mito-nuclear coregulation among cell types, in different brain regions. Our revised list of canonical OXPHOS transcripts and paralogs provided a comprehensive view of tissue-dependent alternatively spliced transcripts and gene paralogs, thus suggesting tissue variability not only in the regulatory pattern but also in the subunit composition of OXPHOS complexes.

Methods

Source of RNA-seq data in RPKM (reads per kilobase per million) mapped reads

Gene expression data and individual information, including gender and cause of death, were obtained from the publicly available

data at the GTEx portal (www.gtexportal.org). Individual ages (rather than age groups) were obtained from the GTEx-authorized data table for donor information (pht002742.v6.p1.c1), "age" column (phv00169063.v6.p1.c1). Detailed information about the tissue sample collection is available at: www.gtexportal.org/home/documentationPage#staticTextSampleCollection. Notably, we used the HUGO gene nomenclature committee (HGNC)-approved gene symbols. To bridge over nomenclature inconsistency between GTEx and HGNC, we provide a list of gene symbols used in this work that differ between the two databases (Supplemental Table S2).

Extraction of alternatively spliced variants

The expression patterns of splice variants of the tested genes were extracted from the publicly available data at the GTEx portal. Only protein-coding variants, determined according to the Ensembl genome browser (www.ensembl.org), were considered in our analysis of co-expression.

Control for age, gender, and cause of death

To control for the possible contribution of age, gender, and cause of death to gene expression patterns, we normalized for these factors using linear regression (see below), such that the calculated residual expression per gene reflected age-/gender-/cause-of-death-independent gene expression. To this end we first divided the RPKM values of each sample by the sum of RPKM values associated with a given gene across all samples. In this way, we defined the expression value of each gene in a given sample as a fraction of its expression out of the sum of its expression values in all available samples. Hence, when calculating correlation of gene expression, we asked whether changes in the residual relative expression of genes are correlated.

For gene "i" in sample "S", the relative gene expression is calculated as follows:

$$\text{Relative Expression} = \frac{\text{RPKM}(\text{Gene}_i)_{(s)}}{\Sigma\ \text{RPKM}(\text{Gene}_i)}.$$

Identification of DNase genomic footprinting sites upstream of OXPHOS genes and their occupation by transcription factors

DNase-seq experiments performed by the ENCODE Consortium were analyzed in terms of a 1000-bp region upstream of the transcription start site of all genes encoding OXPHOS subunits and assembly factors in the nuclear genome. The DNase-seq footprinting data was taken from the ENCODE DNase Hypersensitive Site Master List (http://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg19&g=wgEncodeAwgDnaseMasterSites). Transcription start sites were annotated according to the GRCh37/hg19 RefSeq genes track and were taken from the UCSC Table Browser (https://genome.ucsc.edu/cgi-bin/hgTables) (see Supplemental Methods for a note about reference assembly). Transcription factor-binding sites were obtained from the ENCODE clustered TFs list (http://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg19&g=wgEncodeRegTfbsClusteredV3). For each OXPHOS gene, a positive TF-binding site was listed if its binding peak was identified within the 1000-bp window upstream of the TSS and overlapped with an identified DNase-seq footprinting site/s.

Analysis of co-expression

To assess co-expression between genes across the available multiple tissues in the GTEx database, we focused our analysis on 48 tissues, for which a minimum of 50 samples had available RNA-seq data. Then, we sampled 50 samples from each tissue and calculated Spearman's rank correlation coefficients (ρ_s) between pairs of genes/transcripts according to their 2400 (50×48) residual relative expression value vectors. We repeated this calculation 1000 times and considered the correlation to be significant only if the P-value for the Spearman's rank correlation was smaller than 0.05 in a minimum of 950 iterations (95% of the iterations). Any correlation that failed to match such cut-off value was considered as zero. Otherwise, the median correlation coefficient of each pair of genes/transcripts was obtained after the 1000 random iterations and considered as the median Spearman's rank correlation coefficient for this pair (Supplemental Fig. S1).

Assessment of tissue-dependent co-expression in genes, gene paralogs, and alternatively spliced transcripts

In the analysis of gene paralogs, we calculated the collapsed (grouped/added) expression vectors for all paralogs of the same gene as compared to that of the rest of the OXPHOS genes. Additionally, we calculated the correlation of expression vectors of each of the paralogs and compared to other OXPHOS genes, separately. High tissue dependence of gene expression was indicated by significantly higher correlation values of the collapsed expression vectors relative to each of the different paralogs of the same gene. In the case of alternatively spliced transcripts, we first identified the transcripts that showed a significantly higher correlation values with the "main OXPHOS transcript cluster" (Supplemental Fig. S7), as compared to other transcripts of the same gene (grouped). If there was more than one such transcript in the same tissue for a given gene, we defined the transcript with the highest correlation values to the "main OXPHOS transcript cluster" as the representative of that gene in that tissue. Secondly, we identified genes which are represented by more than one transcript across the different tissues (hereafter termed "candidate transcripts"). Notably, transcripts were considered different only if they differed in their protein translation sequence.

Single-cell RNA-seq data analysis

Normalized single-cell data for mouse and human brains (either FPKM or CPM) was obtained from the Allen Brain Atlas (http://celltypes.brain-map.org/download). To avoid zero-values inflation, a known characteristic of single-cell RNA-seq data, we included in our analysis OXPHOS genes that had a minimum of one count in a minimum of 15,000 cells (mouse database) and then sifted for the cells that share the same set of OXPHOS genes with no zero sequencing read counts. In the analysis of RNA-seq data from single human brain cell nuclei, we noticed low read counts for all mtDNA-encoded genes, as explained in the Results section. Thus, to allow further analysis in the human cell nuclei, we set a value of no zero count in any of four mtDNA-encoded OXPHOS genes in at least 10,700. This approach enabled identifying 28 nDNA-encoded OXPHOS genes with no zero count, yet reduced significantly the number of cells sharing no zero counts in both mtDNA- and nDNA-encoded OXPHOS genes (see Results section). Cells were considered excitatory-neurons, inhibitory-neurons, or nonneuronal according to the expression pattern of known gene markers—the same orthologous markers in human and mouse samples (see Results section). Such cell-type assignment allowed comparison of mito-nuclear OXPHOS gene expression correlations across brain regions and cell types, while maintaining a minimum of 50 cells per group per analysis (Supplemental Fig. S4).

Statistical tools and data visualization

All statistical analyses and data visualizations were performed using Python 2.7 scipy.stats, pandas, matplotlib, and seaborn packages (Supplemental Data Set S4). Validations to the performance of the Python scripts were performed using STATISTICA 12.

References

Aloni Y, Attardi G. 1971. Symmetrical *in vivo* transcription of mitochondrial DNA in HeLa cells. *Proc Natl Acad Sci* **68:** 1757–1761.

Azevedo FA, Carvalho LR, Grinberg LT, Farfel JM, Ferretti RE, Leite RE, Lent R, Herculano-Houzel S. 2009. Equal numbers of neuronal and non-neuronal cells make the human brain an isometrically scaled-up primate brain. *J Comp Neurol* **513:** 532–541.

Barton RA, Venditti C. 2014. Rapid evolution of the cerebellum in humans and other great apes. *Curr Biol* **24:** 2440–2444.

Baughman JM, Nilsson R, Gohil VM, Arlow DH, Gauhar Z, Mootha VK. 2009. A computational screen for regulators of oxidative phosphorylation implicates SLIRP in mitochondrial RNA homeostasis. *PLoS Genet* **5:** e1000590.

Blumberg A, Sailaja BS, Kundaje A, Levin L, Dadon S, Shmorak S, Shaulian E, Meshorer E, Mishmar D. 2014. Transcription factors bind negatively selected sites within human mtDNA genes. *Genome Biol Evol* **6:** 2634–2646.

Blumberg A, Rice EJ, Kundaje A, Danko CG, Mishmar D. 2017. Initiation of mtDNA transcription is followed by pausing, and diverges across human cell types and during evolution. *Genome Res* **27:** 362–373.

Bogenhagen DF, Yoza BK. 1986. Accurate in vitro transcription of *Xenopus laevis* mitochondrial DNA from two bidirectional promoters. *Mol Cell Biol* **6:** 2543–2550.

Bossi S, Simpson J, Isacson O. 1993. Age dependence of striatal neuronal death caused by mitochondrial dysfunction. *Neuroreport* **4:** 73–76.

Brady P, Park E, Liu J, Hanson R, Brady L. 1992. Isolation and characterization of the promoter for the gene coding for the 68 kDa carnitine palmitoyltransferase from the rat. *Biochem J* **286:** 779–783.

Bridges HR, Mohammed K, Harbour ME, Hirst J. 2017. Subunit NDUFV3 is present in two distinct isoforms in mammalian complex I. *Biochim Biophys Acta* **1858:** 197–207.

Brun T, Li N, Jourdain AA, Gaudet P, Duhamel D, Meyer J, Bosco D, Maechler P. 2015. Diabetogenic milieus induce specific changes in mitochondrial transcriptome and differentiation of human pancreatic islets. *Hum Mol Genet* **24:** 5270–5284.

Burtscher J, Zangrandi L, Schwarzer C, Gnaiger E. 2015. Differences in mitochondrial function in homogenated samples from healthy and epileptic specific brain tissues revealed by high-resolution respirometry. *Mitochondrion* **25:** 104–112.

Calvo SE, Mootha VK. 2010. The mitochondrial proteome and human disease. *Annu Rev Genomics Hum Genet* **11:** 25–44.

Calvo SE, Clauser KR, Mootha VK. 2016. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res* **44**(D1): D1251–D1257.

Campanella M, Casswell E, Chong S, Farah Z, Wieckowski MR, Abramov AY, Tinker A, Duchen MR. 2008. Regulation of mitochondrial structure and function by the F_1F_0-ATPase inhibitor protein, IF_1. *Cell Metab* **8:** 13–25.

Chae S, Ahn BY, Byun K, Cho YM, Yu M-H, Lee B, Hwang D, Park KS. 2013. A systems approach for decoding mitochondrial retrograde signaling pathways. *Sci Signal* **6:** rs4.

Chen JQ, Delannoy M, Cooke C, Yager JD. 2004. Mitochondrial localization of ERα and ERβ in human MCF7 cells. *Am J Physiol Endocrinol Metab* **286:** E1011–E1022.

Chinnery PF, Hudson G. 2013. Mitochondrial genetics. *Br Med Bull* **106:** 135–159.

Cogliati S, Calvo E, Loureiro M, Guaras AM, Nieto-Arellano R, Garcia-Poyatos C, Ezkurdia I, Mercader N, Vázquez J, Enriquez JA. 2016. Mechanism of super-assembly of respiratory complexes III and IV. *Nature* **539:** 579–582.

Coskun PE, Beal MF, Wallace DC. 2004. Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication. *Proc Natl Acad Sci* **101:** 10726–10731.

Coyle JT, Schwarcz R. 1976. Lesion of striatal neurones with kainic acid provides a model for Huntington's chorea. *Nature* **263:** 244–246.

Craven L, Alston CL, Taylor RW, Turnbull DM. 2017. Recent advances in mitochondrial disease. *Annu Rev Genomics Hum Genet* **18:** 257–275.

Da Cruz S, Xenarios I, Langridge J, Vilbois F, Parone PA, Martinou JC. 2003. Proteomic analysis of the mouse liver mitochondrial inner membrane. *J Biol Chem* **278:** 41566–41571.

Dibley MG, Ryan MT, Stroud DA. 2017. A novel isoform of the human mitochondrial complex I subunit NDUFV3. *FEBS Lett* **591:** 109–117.

Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* **95:** 14863–14868.

Emery NJ, Clayton NS. 2005. Evolution of the avian brain and intelligence. *Curr Biol* **15:** R946–R950.

Fernandez-Vizarra E, Enriquez JA, Perez-Martos A, Montoya J, Fernandez-Silva P. 2011. Tissue-specific differences in mitochondrial activity and biogenesis. *Mitochondrion* **11:** 207–213.

Gabaldón T, Huynen MA. 2007. From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. *PLoS Comput Biol* **3:** e219.

Garbian Y, Ovadia O, Dadon S, Mishmar D. 2010. Gene expression patterns of oxidative phosphorylation complex I subunits are organized in clusters. *PLoS One* **5:** e9985.

The GTEx Consortium. 2015. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348:** 648–660.

Guerrero-Castillo S, Cabrera-Orefice A, Huynen MA, Arnold S. 2017. Identification and evolutionary analysis of tissue-specific isoforms of mitochondrial complex I subunit NDUFV3. *Biochim Biophys Acta* **1858:** 208–217.

Gustafsson CM, Falkenberg M, Larsson NG. 2016. Maintenance and expression of mammalian mitochondrial DNA. *Annu Rev Biochem* **85:** 133–160.

Hazkani-Covo E, Sorek R, Graur D. 2003. Evolutionary dynamics of large *numts* in the human genome: rarity of independent insertions and abundance of post-insertion duplications. *J Mol Evol* **56:** 169–174.

Herculano-Houzel S. 2014. The glia/neuron ratio: how it varies uniformly across brain structures and species and what that means for brain physiology and evolution. *Glia* **62:** 1377–1391.

Hüttemann M, Jaradat S, Grossman LI. 2003. Cytochrome *c* oxidase of mammals contains a testes-specific isoform of subunit VIb—the counterpart to testes-specific cytochrome *c*? *Mol Reprod Dev* **66:** 8–16.

Ikebe S-i, Tanaka M, Ohno K, Sato W, Hattori K, Kondo T, Mizuno Y, Ozawa T. 1990. Increase of deleted mitochondrial DNA in the striatum in Parkinson's disease and senescence. *Biochem Biophys Res Commun* **170:** 1044–1048.

Kadenbach B, Huettemann M. 2015. The subunit composition and function of mammalian cytochrome *c* oxidase. *Mitochondrion* **24:** 64–76.

L'Abbé D, Duhaime J-F, Lang BF, Morais R. 1991. The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. *J Biol Chem* **266:** 10844–10850.

Ladd AC, Brohawn DG, Thomas RR, Keeney PM, Berr SS, Khan SM, Portell FR, Shakenov MZ, Antkowiak PF, Kundu B, et al. 2017. RNA-seq analyses reveal that cervical spinal cords and anterior motor neurons from amyotrophic lateral sclerosis subjects show reduced expression of mitochondrial DNA-encoded respiratory genes, and rhTFAM may correct this respiratory deficiency. *Brain Res* **1667:** 74–83.

Lapuente-Brun E, Moreno-Loshuertos R, Acín-Pérez R, Latorre-Pellicer A, Colás C, Balsa E, Perales-Clemente E, Quirós PM, Calvo E, Rodríguez-Hernández M. 2013. Supercomplex assembly determines electron flux in the mitochondrial electron transport chain. *Science* **340:** 1567–1570.

Lauwers F, Cassot F, Lauwers-Cances V, Puwanarajah P, Duvernoy H. 2008. Morphometry of the human cerebral cortex microcirculation: general characteristics and space-related profiles. *Neuroimage* **39:** 936–948.

Liang WS, Reiman EM, Valla J, Dunckley T, Beach TG, Grover A, Niedzielko TL, Schneider LE, Mastroeni D, Caselli R. 2008. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *Proc Natl Acad Sci* **105:** 4441–4446.

Liesa M, Shirihai OS. 2013. Mitochondrial dynamics in the regulation of nutrient utilization and energy expenditure. *Cell Metab* **17:** 491–506.

Lightowlers R, Ewart G, Aggeler R, Zhang Y-Z, Calavetta L, Capaldi RA. 1990. Isolation and characterization of the cDNAs encoding two isoforms of subunit C_{IX} of bovine cytochrome c oxidase. *J Biol Chem* **265:** 2677–2681.

Lim D, Fedrizzi L, Tartari M, Zuccato C, Cattaneo E, Brini M, Carafoli E. 2008. Calcium homeostasis and mitochondrial dysfunction in striatal neurons of Huntington disease. *J Biol Chem* **283:** 5780–5789.

Lindfors C, Nilsson IA, Garcia-Roves PM, Zuberi AR, Karimi M, Donahue LR, Roopenian DC, Mulder J, Uhlén M, Ekström TJ. 2011. Hypothalamic mitochondrial dysfunction associated with anorexia in the *anx/anx* mouse. *Proc Natl Acad Sci* **108:** 18108–18113.

Lunnon K, Keohane A, Pidsley R, Newhouse S, Riddoch-Contreras J, Thubron EB, Devall M, Soininen H, Kloszewska I, Mecocci P, et al. 2017. Mitochondrial genes are altered in blood early in Alzheimer's disease. *Neurobiol Aging* **53:** 36–47.

MacLean PD. 1990. *The triune brain in evolution: role in paleocerebral functions*. Springer Science & Business Media, New York.

Mashour GA, Alkire MT. 2013. Evolution of consciousness: phylogeny, ontogeny, and emergence from general anesthesia. *Proc Natl Acad Sci* **110:** 10357–10364.

McGeer P. 2013. *Molecular neurobiology of the mammalian brain*. Springer Science & Business Media, New York.

Mishmar D, Ruiz-Pesini E, Brandon M, Wallace DC. 2004. Mitochondrial DNA-like sequences in the nucleus (NUMTs): insights into our African origins and the mechanism of foreign DNA integration. *Hum Mutat* **23:** 125–133.

Mourier A, Matic S, Ruzzenente B, Larsson N-G, Milenkovic D. 2014. The respiratory chain supercomplex organization is independent of *COX7a2l* isoforms. *Cell Metab* **20:** 1069–1075.

Pickrell AM, Fukui H, Wang X, Pinto M, Moraes CT. 2011a. The striatum is highly susceptible to mitochondrial oxidative phosphorylation dysfunctions. *J Neurosci* **31:** 9895–9904.

Pickrell AM, Pinto M, Hida A, Moraes CT. 2011b. Striatal dysfunctions associated with mitochondrial DNA damage in dopaminergic neurons in a mouse model of Parkinson's disease. *J Neurosci* **31:** 17649–17658.

Pierron D, Wildman DE, Huttemann M, Markondapatnaikuni GC, Aras S, Grossman LI. 2012. Cytochrome c oxidase: evolution of control via nuclear subunit addition. *Biochim Biophys Acta* **1817:** 590–597.

Porporato PE, Dhup S, Dadhich RK, Copetti T, Sonveaux P. 2011. Anticancer targets in the glycolytic metabolism of tumors: a comprehensive review. *Front Pharmacol* **2:** 49.

Quirós PM, Mottis A, Auwerx J. 2016. Mitonuclear communication in homeostasis and stress. *Nat Rev Mol Cell Biol* **17:** 213–226.

Read JA, Winter VJ, Eszes CM, Sessions RB, Brady RL. 2001. Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins* **43:** 175–185.

Reyes A, He J, Mao C, Bailey L, Di Re M, Sembongi H, Kazak L, Dzionek K, Holmes J, Cluett T. 2011. Actin and myosin contribute to mammalian mitochondrial DNA maintenance. *Nucleic Acids Res* **39:** 5098–5108.

Reznik E, Wang Q, La K, Schultz N, Sander C. 2017. Mitochondrial respiratory gene expression is suppressed in many cancers. *eLife* **6:** e21592.

Schlerf A, Droste M, Winter M, Kadenbach B. 1988. Characterization of two different genes (cDNA) for cytochrome c oxidase subunit VIa from heart and liver of the rat. *EMBO J* **7:** 2387.

Segade F, Hurlé B, Claudio E, Ramos S, Lazo PS. 1996. Identification of an additional member of the cytochrome c oxidase subunit VIIa family of proteins. *J Biol Chem* **271:** 12343–12349.

She H, Yang Q, Shepherd K, Smith Y, Miller G, Testa C, Mao Z. 2011. Direct regulation of complex I by mitochondrial MEF2D is disrupted in a mouse model of Parkinson disease and in human patients. *J Clin Invest* **121:** 930–940.

Shyamsundar R, Kim YH, Higgins JP, Montgomery K, Jorden M, Sethuraman A, van de Rijn M, Botstein D, Brown PO, Pollack JR. 2005. A DNA microarray survey of gene expression in normal human tissues. *Genome Biol* **6:** R22.

Takimoto M, Hamada T. 2014. Acute exercise increases brain region-specific expression of MCT1, MCT2, MCT4, GLUT1, and COX IV proteins. *J Appl Physiol* **116:** 1238–1250.

Vallejos CA, Risso D, Scialdone A, Dudoit S, Marioni JC. 2017. Normalizing single-cell RNA sequencing data: challenges and opportunities. *Nat Methods* **14:** 565–571.

van Waveren C, Moraes CT. 2008. Transcriptional co-expression and co-regulation of genes coding for components of the oxidative phosphorylation system. *BMC Genomics* **9:** 18.

Wolf AR, Mootha VK. 2014. Functional genomic analysis of human mitochondrial RNA processing. *Cell Rep* **7:** 918–931.