Human genetic analyses of organelles highlight the nucleus in age-related trait heritability

Rahul Gupta[1,2,3], Konrad J. Karczewski[2,3], Daniel Howrigan[2,3], Benjamin M. Neale[2,3,*], Vamsi K. Mootha[1,2,*]

[1] Howard Hughes Medical Institute and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114
[2] Broad Institute of MIT and Harvard, Cambridge, MA 02142
[3] Analytic and Translational Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

[*]Corresponding authors: bneale@broadinstitute.org (B.M.N.), vamsi@hms.harvard.edu (V.K.M.)

Keywords: mitochondria, nucleus, aging, oxidative phosphorylation, OXPHOS, transcription factor, zinc finger, KRAB domain, gene regulation, UK Biobank, constraint, dominance, haplosufficiency, haplosufficient, PPARGC1A, ESRRA, TFAM

Abstract

Aging is associated with defects in many organelles, but an open question is whether the inherited risk for age-related disease is enriched within loci relevant to each organelle. Here, we begin with a focus on mitochondria, as mitochondrial dysfunction is a hallmark of age-related disease. We report a striking lack of enrichment of mitochondria-relevant loci across GWAS for 24 age-related traits. Analyses of nine additional organelles reveal enrichment only for the nucleus, particularly nuclear transcription factors. Consistent with these results, natural selection appears to exert stronger purifying selection against protein-truncating variants for transcription factors compared to mitochondrial pathways, underscoring the importance of inherited variation in gene-regulation in age-related traits.

Introduction

The global burden of age-related diseases such as type 2 diabetes (T2D), Parkinson's disease (PD), and cardiovascular disease (CVD) has been steadily rising due in part to a progressively aging population. These diseases are often highly heritable[1]. Genome-wide association studies (GWAS) have led to the discovery of thousands of robust associations with common genetic variants[2], implicating a complex genetic architecture as underlying much of the heritable risk. These loci hold the potential to reveal underlying mechanisms of disease and spotlight targetable pathways.

Aging has been associated with dysfunction in many cellular organelles. This is perhaps most notable for the mitochondria, dysfunction in which is frequently invoked as a key "hallmark" of aging[3] and has been nominated as a candidate driver of virtually all common age-associated diseases. In particular, deficits in mitochondrial oxidative phosphorylation (OXPHOS), the chief pathway for ATP production, is often observed in these diseases as evidenced by a decline in *in vivo* ^{31}P-NMR measures of OXPHOS activity[4,5], a decline in electron transport chain enzymatic activity[6–12] in biopsy/autopsy material, an accumulation of somatic mtDNA mutations[13–15], and a decline in mtDNA copy number (mtCN)[16]. Protein products from both the nuclear DNA (nucDNA), which encodes ~1100 proteins that primarily localize to the mitochondria[17], and the mitochondrial DNA (mtDNA), which encodes 13 proteins, are critical for proper OXPHOS homeostasis[18]. Rare, inherited mutations in any of ~300 nucDNA or mtDNA genes can cause severe "primary" mitochondrial diseases that impact multiple organ systems[19]. Dysfunction in other organelles[3], such as the nucleus (e.g., increased gene regulatory noise from epigenetic alterations[3], telomere shortening[20,21]), the lysosome[22], and the endoplasmic reticulum[23], has also been associated with aging and/or age-related disease.

A key question is whether inherited variation influencing organelle function is enriched for age-related disease risk. In the present study, we use a human genetics approach to assess common variation in loci relevant to the function of 10 cellular organelles. We begin with a deliberate focus on mitochondria given the depth of literature linking it to age-related disease, testing nucDNA and mtDNA loci relevant for mitochondrial function in 24 different age-related diseases and traits. We hypothesized that heritability for common, age-related traits would be overrepresented among mitochondria-relevant loci, namely variants near genes encoding the organelle's proteome or loci associated with quantitative readouts of mitochondrial function.

To our surprise, we find no evidence of enrichment for genome-wide association signal in mitochondria-relevant loci across any of our analyses. Of ten tested organelles, only the nucleus shows enrichment among many age-associated traits, with the signal emanating from the transcription factors. Further analysis shows that genes encoding the mitochondrial proteome tend to be tolerant to heterozygous predicted loss-of-function (pLoF) variation and thus are surprisingly "haplosufficient," whereas nuclear transcription factors are especially sensitive to gene dosage and are often "haploinsufficient". Thus, we highlight variation influencing gene-regulatory pathways, rather than organelle physiology, in the inherited risk of common age-associated diseases.

Results

Age-related diseases and traits show diverse genetic architectures

To systematically define age-related diseases, we turned to recently published epidemiological data from the United Kingdom (U.K.)[24] in order to match the U.K. Biobank (UKB)[25] cohort. We prioritized traits whose prevalence increased as a function of age (**Methods**) and were represented in UKB (https://github.com/Nealelab/UK_Biobank_GWAS) and/or had available published GWAS meta-analyses[26–35] (**Figure 1A**, **Supplementary note).** We used SNP heritability estimates from stratified linkage disequilibrium score regression (S-LDSC, https://github.com/bulik/ldsc)[36] to ensure that our selected traits were sufficiently heritable (**Methods**, **Supplementary note**). We then computed pairwise genetic

correlations between the age-associated traits to compare their respective genetic architectures (**Figure 1B, Table S2, Methods**). As expected we find a highly correlated module of primarily cardiometabolic traits with high density lipoprotein (HDL) showing anti-correlation[37]. Interestingly, several other traits (gastroesophageal reflux disease (GERD), osteoarthritis) showed moderate correlation to the cardiometabolic trait cluster while atrial fibrillation, for which T2D and CVD are risk factors[38], did not. Our final set of prioritized, age-associated traits included 24 genetically diverse, heritable phenotypes (**Table S1**). Of these, 11 traits were sufficiently heritable only in UKB, 3 were sufficiently heritable only among non-UKB meta-analyses, and 10 were well-powered in both UKB and an independent cohort.

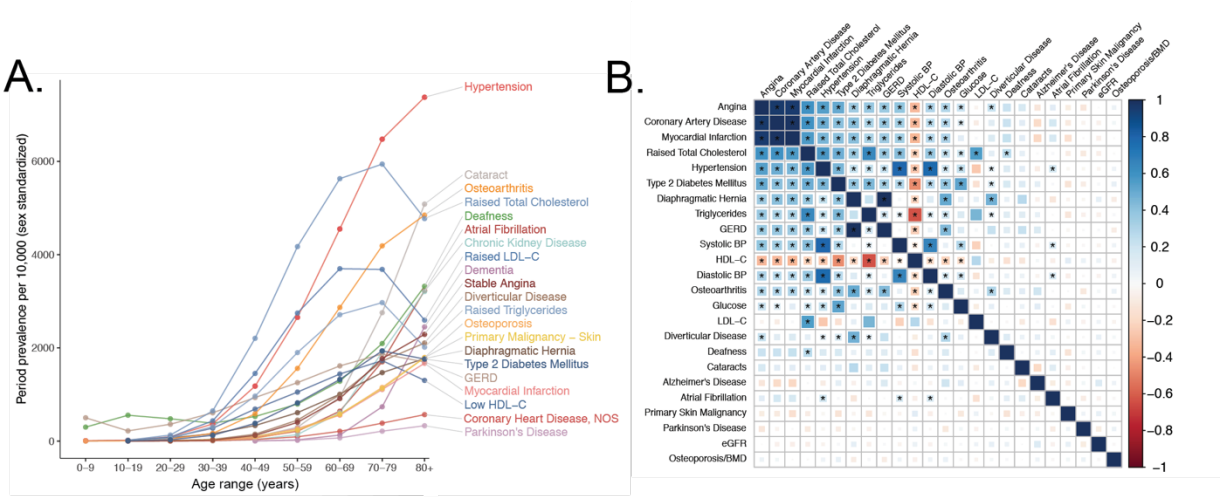

Figure 1. Selection of genetically diverse age-related diseases and traits using epidemiological data. **A.** Period prevalence of age-associated diseases systematically selected for this study **(Methods).** Epidemiological data obtained from Kuan et al. 2019. **B**: Genetic correlation between the selected age-related traits. All correlations were assessed between UK Biobank phenotypes with the exception of eGFR, Alzheimer's Disease, and Parkinson's Disease, for which the respective meta-analyses were used (**Methods**). Point estimates and standard errors reported in **Table S2**. * represent genetic correlations that are significantly different from 0 at a Bonferroni-corrected threshold for p = 0.05 x 24 traits.

No evidence for enrichment of age-related trait heritability in mitochondria-relevant loci

To test if age-related trait heritability was enriched among mitochondria-relevant loci, we began by simply asking if ~1100 nucDNA genes encoding the mitochondrial proteome from the MitoCarta2.0 inventory[17] were found near lead SNPs for our selected traits represented in the NHGRI-EBI GWAS Catalog (https://www.ebi.ac.uk/gwas/)[39] more frequently than expectation (**Methods, Supplementary note**). To our surprise, no traits showed a statistically significant enrichment of mitochondrial genes (**Figure S1A**); in fact, six traits showed a statistically significant depletion. Even more strikingly, MitoCarta genes tended to be nominally enriched in fewer traits than the average randomly selected sample of protein-coding genes (**Figure S1B**, empirical p = 0.014). This lack of enrichment was observed more broadly across virtually all traits represented in the GWAS Catalog (**Figure S1C**). We also tested several transcriptional regulators of mitochondrial biogenesis and function – *TFAM, GABPA, GABPB1, ESRRA, YY1, NRF1, PPARGC1A, PPARGC1B*. We found little evidence supporting a role for these genes in modifying risk for the age-related GWAS Catalog phenotypes, observing only a single trait (heel bone mineral density) for which a mitochondrial transcriptional regulator (*TFAM*) was nearest an associated genome-wide significant variant (**Supplementary note**).

To investigate further, we turned to U.K. Biobank (UKB). We compiled and tested three classes of "mitochondria-relevant loci" (**Figure 2A**) with which we interrogated the association between common mitochondrial variation and common disease. First, we curated literature-reported nucDNA quantitative trait loci (QTLs) associated with measures of mitochondrial function (**Table S3**): mtCN[40,41], mtRNA

abundance and modification[42,43], and plasma levels of OXPHOS dysfunction biomarkers including GDF15 protein[44,45], lactate, pyruvate, and lactate/pyruvate ratio[46–48]. Second, we considered all common variants in or near nucDNA MitoCarta genes, as well as two subsets of MitoCarta: mitochondrial Mendelian disease genes[19] and nucDNA-encoded OXPHOS genes. Third, we obtained mtDNA genotypes at up to 213 loci after quality control (**Methods**) from 360,662 individuals.

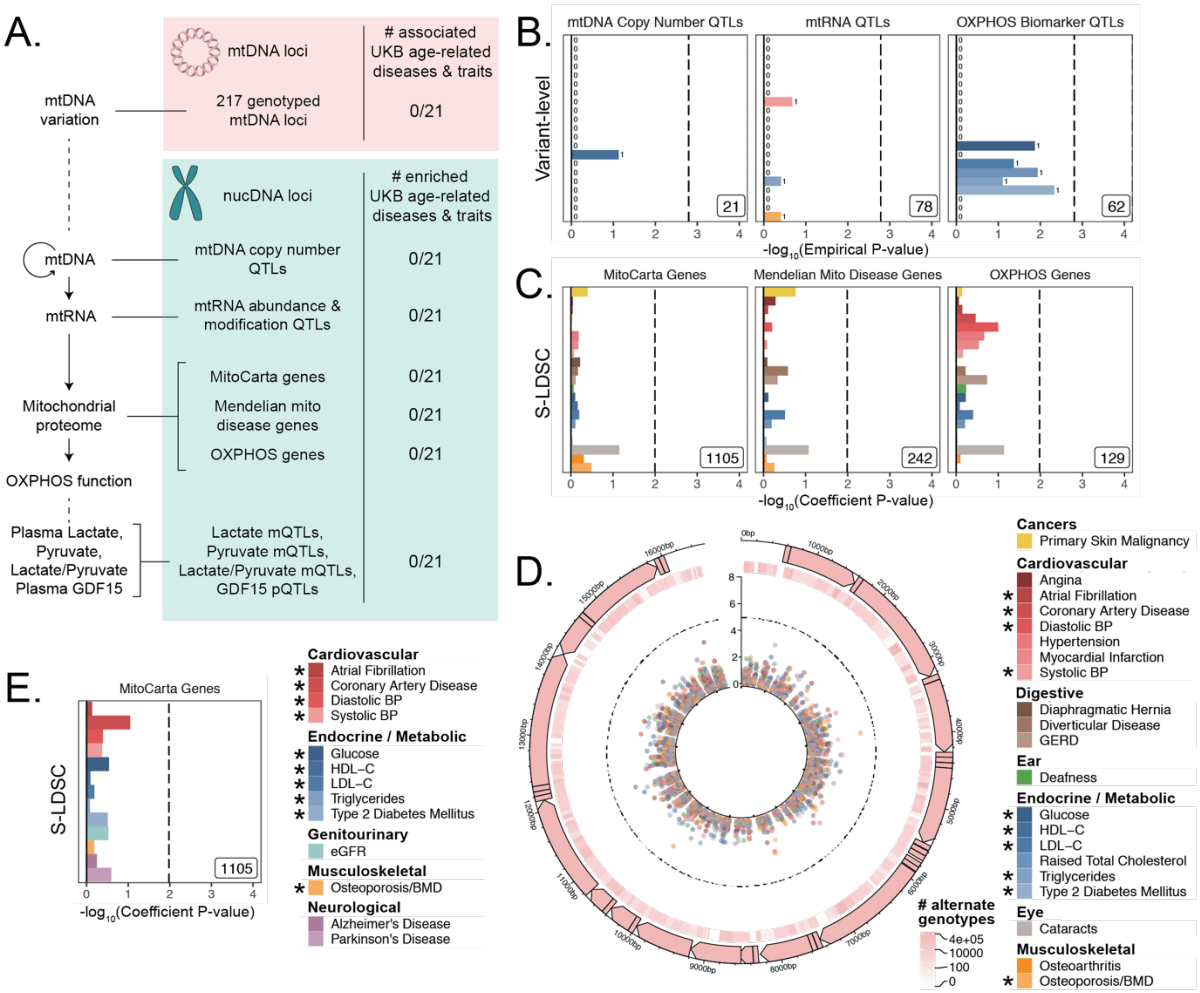

Figure 2. Assessment of the association of nucDNA and mtDNA mitochondria-relevant loci to age-related traits. **A.** Scheme outlining the aspects of mitochondrial function assessed in this study. nucDNA loci relevant to mitochondrial function are shown in teal, while mtDNA loci are shown in pink. **B.** Enrichment results for the overlap between loci associated with mtDNA copy number, mtRNA abundance/modification, and OXPHOS biomarkers and loci significantly associated with age-related disease in UKB. Inset number represents the number of tested SNPs, numbers adjacent to bars represent the absolute number of mitochondria-relevant loci overlapping the respective age-related disease. Dotted line represents Bonferroni cutoff for p = 0.1; BH FDR 0.1 threshold cannot be visualized as no tests pass the cutoff (**Supplementary note**). **C.** S-LDSC enrichment p-values on top of the baseline model in UKB. Inset labels represent gene-set size; dotted line represents BH FDR 0.1 threshold. **D.** Visualization of mtDNA variants and associations with age-related diseases. The outer-most track represents the genetic architecture of the circular mtDNA. The heatmap track represents the number of individuals with alternate genotype on log scale. The inner track represents mitochondrial genome-wide association p-values, with radial angle corresponding to position on the mtDNA and magnitude representing $-\log_{10}$ P-value. Dotted line represents Bonferroni cutoff for all tested trait-variant pairs. **E.** Replication of S-LDSC enrichment results in meta-analyses. Dotted line represents BH FDR 0.1 threshold. * represent traits for which sufficiently well powered cohorts from both UKB and meta-analyses were available. The trait color legend to the right of panel **D** applies to panels **B, C,** and **D**, representing UKB traits.

First, we tested if published QTLs for mtCN, mtRNA abundance, and OXPHOS biomarkers (**Table S3, S4**) were enriched for an overlap with genome-wide significant loci for each of our age-related traits in UKB (**Methods, Figure S2**). We observed no evidence of enrichment among QTLs available in the literature (**Figure 2B, Supplementary note;** all $q > 0.1$).

Second, we used S-LDSC[36,49] and MAGMA (https://ctg.cncr.nl/software/magma)[50], two robust methods that can be used to assess gene-based heritability enrichment accounting for LD and several confounders, to test if there was any evidence of heritability enrichment among MitoCarta genes (**Methods**). We found no evidence of enrichment near nucDNA MitoCarta genes for any trait tested in UKB using S-LDSC (**Figure 2C, S8A**), consistent with our results from the GWAS Catalog. We replicated this lack of enrichment using MAGMA at two different window sizes (**Figure S8C, S8E;** all $q > 0.1$).

Given the lack of enrichment among the MitoCarta genes, we wanted to (1) verify that our selected methods could detect previously reported enrichments and (2) confirm that common variation in or near MitoCarta genes can lead to expression-level perturbations. We first successfully replicated previously reported enrichment among tissue-specific genes for key traits using both S-LDSC (**Figure S3, S4**) and MAGMA (**Figure S5, S6, Supplementary note, Methods**). We next confirmed that we had sufficient power using both S-LDSC and MAGMA to detect physiologically relevant enrichment effect sizes among MitoCarta genes (**Figure S7, Methods, Supplementary note**). We finally examined the landscape of cis-expression QTLs (eQTLs) for these genes and found that almost all MitoCarta genes have cis-eQTLs in at least one tissue and often have cis-eQTLs in more tissues than most protein-coding genes (**Figure S9, Methods, Supplementary note**). Hence, our selected methods could detect physiologically relevant heritability enrichments among our selected traits at gene-set sizes comparable to that of MitoCarta, and common variants in or near MitoCarta genes exerted *cis*-control on gene expression.

Third, we considered mtDNA loci genotyped in UKB, obtaining calls for up to 213 common variants passing quality control across 360,662 individuals (**Methods, Supplementary note**). We found no significant associations on the mtDNA for any of the 21 age-related traits available in UKB using linear or logistic regression (**Methods, Figure 2E, S9**).

As a control and to validate our approach, we also performed mtDNA-GWAS for specific traits with previously reported associations. A recent analysis of ~147,437 individuals in BioBank Japan revealed four distinct traits with significant mtDNA associations[51]. Of these, creatinine and aspartate aminotransferase (AST) had sufficiently large sample sizes in UKB. We observed a large number of associations throughout the mtDNA for both traits ($p < 1.15 * 10^{-5}$, **Figure S9E**). Thus, our mtDNA association method was able to replicate robust mtDNA associations among well-powered traits.

Finally, we sought to replicate our negative results in an independent cohort. We turned to published GWAS meta-analyses[26–35] (**Table S1**) and successfully replicated the lack of enrichment for MitoCarta genes across all 10 traits with an available independent cohort GWAS using S-LDSC (**Figure 2E, S8B**) and MAGMA at two different window sizes (**Figure S8D, Supplementary note;** all $q > 0.1$). Importantly, while we were unable to pursue analyses for PD and Alzheimer's disease in UKB due to limited case counts, we tested MitoCarta genes among well-powered meta-analyses for these disorders (**Supplementary note**) and observed no enrichment (**Figure 2E;** all $q > 0.1$).

In summary, we tested (1) QTLs for mitochondrial physiology in UKB, (2) nucDNA loci near genes that encode the mitochondrial proteome in the GWAS Catalog, UKB, and GWAS meta-analyses, (3) mtDNA variants in UKB, and (4) known transcriptional regulators of mitochondrial biogenesis and function in the GWAS Catalog. We found no convincing evidence of heritability enrichment for common age-associated diseases among these mitochondria-relevant loci (**Table S8**).

<u>Enrichment of age-related trait heritability near genes encoding nuclear transcription factors</u>

We next asked whether heritability for age-related diseases and traits clusters among loci associated with any cellular organelle. We used the COMPARTMENTS database (https://compartments.jensenlab.org) to define gene-sets corresponding to the proteomes of nine additional organelles[52] besides mitochondria (**Methods**). We used S-LDSC to produce heritability estimates for these categories in the UKB age-related disease traits, finding evidence of heritability enrichment in many traits for genes comprising the nuclear proteome (**Figure 3A, Methods**). No other tested organelles showed evidence of heritability enrichment. Variation in or near genes comprising the nuclear proteome explained over 50% of disease heritability on average despite representing only ~35% of tested SNPs (**Figure S10, Supplementary note**). We successfully replicated this pattern of heritability enrichment among organelles using MAGMA in UKB at two window sizes (**Figure S13A, S13B**), again finding only enrichment among genes related to the nucleus.

With over 6,000 genes comprising the nuclear proteome, we considered largely disjoint subsets of the organelle's proteome to trace the source of the enrichment signal[53–55] (**Figure 3B, Methods, Supplementary note**). We found significant heritability enrichment within the set of 1,804 genes whose protein products are annotated to localize to the chromosome itself (*q < 0.1* for 9 traits, **Figure 3C, S12**). Further partitioning revealed that much of this signal is attributable to the subset classified as transcription factors[55] (1,523 genes, *q < 0.1* for 10 traits, **Figure 3D, S12**). We replicated these results using MAGMA in UKB at two window sizes (**Figure S13**), and also replicated enrichments among TFs in several (but not all) corresponding meta-analyses (**Figure S14**) despite reduced power (**Figure S7H**). We generated functional subdivisions of the TFs (**Methods, Supplementary note**), finding that the non-zinc finger TFs

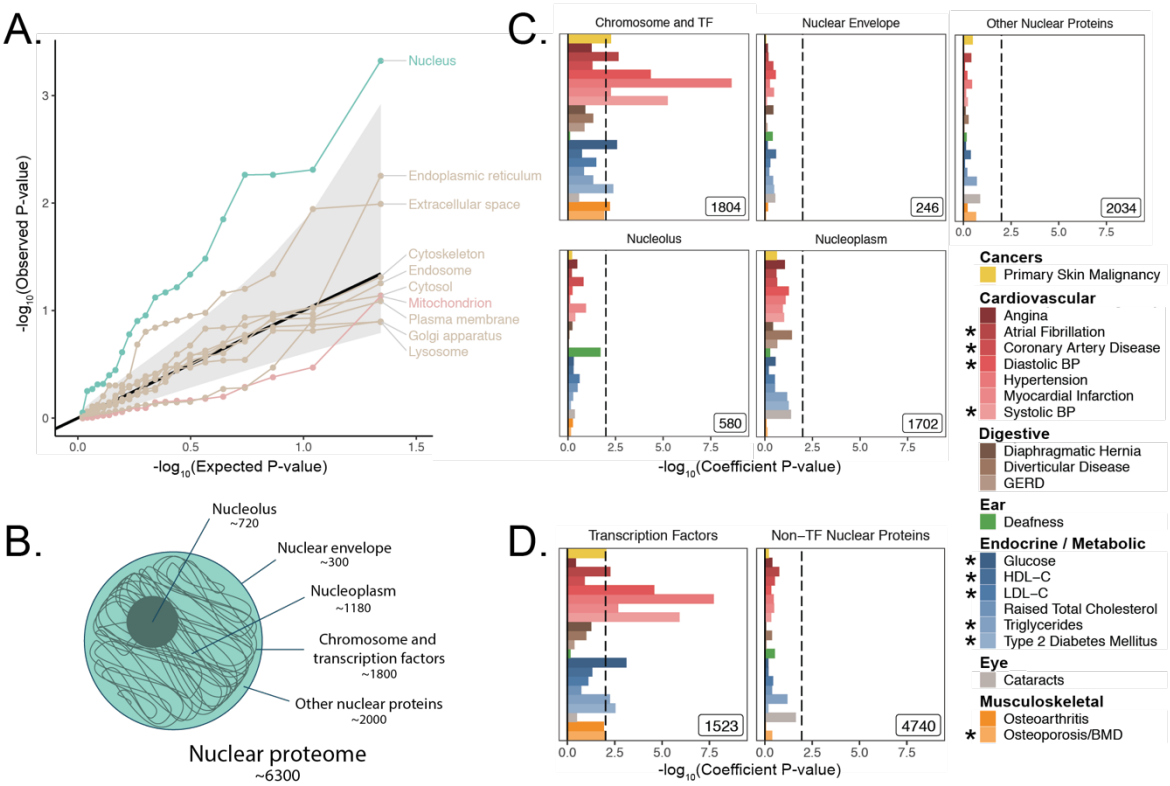

Figure 3. Heritability enrichment of organellar proteomes across age-related disease in UK Biobank. **A.** Quantile-quantile plot of heritability enrichment p-values atop the baseline model for gene-sets representing organellar proteomes, with black line representing expected null p-values following the uniform distribution and shaded ribbon representing 95% CI. **B.** Scheme of spatially distinct disjoint subsets of the nuclear proteome as a strategy to characterize observed enrichment of the nuclear proteome. Numbers represent gene-set size. **C.** S-LDSC enrichment p-values for spatial subsets of the nuclear proteome computed atop the baseline model. **D.** S-LDSC enrichment p-values for TFs and all other nucleus-localizing proteins. Inset numbers represent gene-set sizes, black lines represent cutoff at BH FDR < 10%. * represent traits for which sufficiently well powered cohorts from both UKB and meta-analyses were available.

showed enrichment for a highly similar set of traits to those enriched for the whole set of TFs (**Figure S15D, S16B, S17B, S18B**). Interestingly, the KRAB domain-containing zinc fingers (KRAB ZFs)[56], which are recently evolved (**Figure S15H**), were largely devoid of enrichment even compared to non-KRAB ZFs (**Figure S15E, S16C, S17C, S18C**). Thus, we find that variation within or near non-KRAB domain-containing transcription factor genes has an outsize influence on age-associated disease heritability (**Table S8**).

Mitochondrial genes tend to be more "haplosufficient" than genes encoding other organelles

In light of observing heritability enrichment only among nuclear transcription factors, we wanted to determine if the fitness cost of pLoF variation in genes across cellular organelles mirrored our results. Mitochondria-localizing genes and TFs play a central role in numerous Mendelian diseases[19,57–59], so we initially hypothesized that genes belonging to either category would be under significant purifying selection (i.e., constraint). We obtained constraint metrics from gnomAD

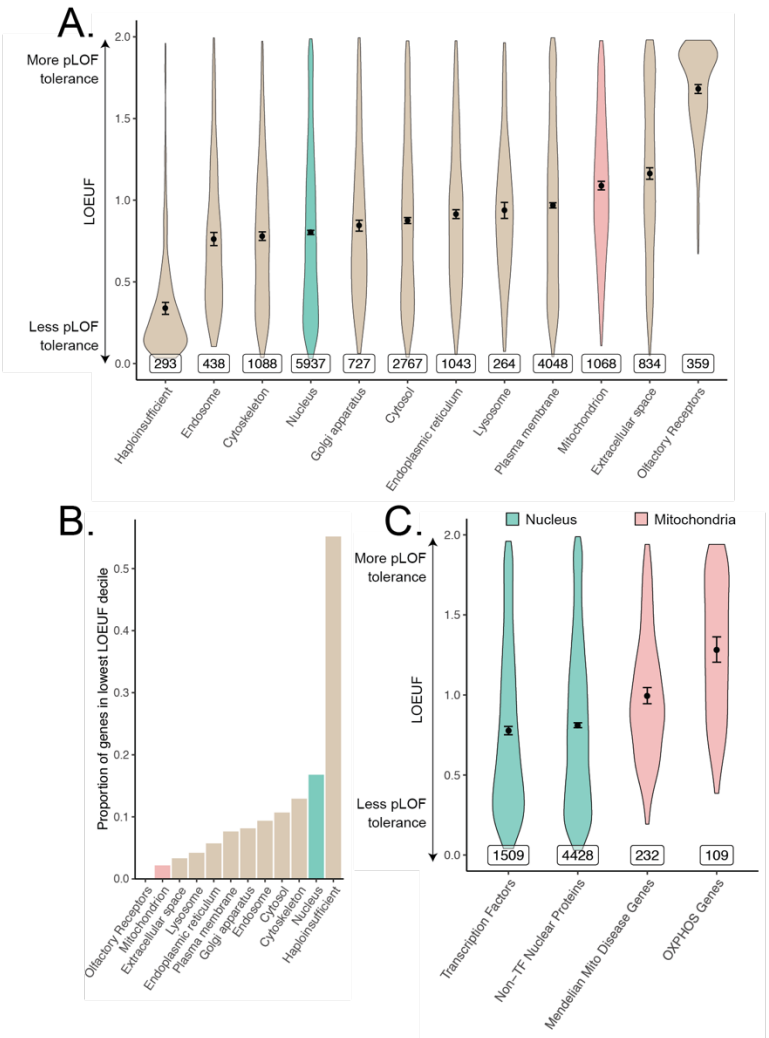

Figure 4. Differences in constraint distribution across organelles. **A.** Constraint as measured by LOEUF from gnomAD v2.1.1 for genes comprising organellar proteomes, book-ended by distributions for known haploinsufficient genes as well as olfactory receptors. Lower values indicate genes exacting a greater organismal fitness cost from a heterozygous LoF variant (greater constraint). **B.** Proportion of each gene-set found in the lowest LOEUF decile. Higher values indicate gene-sets containing more highly constrained genes. **C.** Constraint distributions for subsets of the nuclear-encoded mitochondrial proteome (red) and subsets of the nucleus (teal). Black points represent the mean with 95% CI. Inset numbers represent gene-set size.

199 (https://gnomad.broadinstitute.org)[60] as the LoF observed/expected fraction (LOEUF). In agreement with
200 our GWAS enrichment results, we observed that the mitochondrion on average is one of the least
201 constrained organelles we tested, in stark contrast to the nucleus (**Figure 4A**). In fact, the nucleus was
202 second only to the set of "haploinsufficient" genes (defined based on curated human clinical genetic
203 data[60], **Methods**) in the proportion of its genes in the most constrained decile, while the mitochondrion
204 lay on the opposite end of the spectrum (**Figure 4B**). Interestingly, even the Mendelian mitochondrial
205 disease genes had a high tolerance to pLoF variation on average in comparison to TFs (**Figure 4C, S19A**).
206 Even across different categories of TFs, we observed that highly constrained TF subsets tend to show
207 GWAS enrichment (**Figure S19B**, **S15E**) relative to unconstrained subsets for our tested traits. Indeed,
208 explicit inclusion of LOEUF as a covariate in the enrichment analysis model (**Methods**) reduced the
209 significance of (but did not eliminate) the enrichment seen for the TFs (**Figure S20B, S21B, S20E, S20F**).
210 Thus, while disruption in both mitochondrial genes and TFs can produce rare disease, the fitness cost of
211 heterozygous variation in mitochondrial genes appears to be far lower than that among the TFs. This
212 dichotomy reflects the contrasting enrichment results between the mitochondrial genes and the TFs and
213 supports the importance of gene regulation as it relates to evolutionary conservation.
214
215 **Discussion**
216 Organelle pathology has been documented in many age-related diseases[3,23,61–64]. Perhaps most notably, a
217 decline in mitochondrial abundance and activity has long been reported as one of the most consistent
218 biomarkers of aging[4,9,14,15] and of heritable age-associated diseases such as T2D[5,6,10–12], PD[8,13], and CVD[7,16].
219 This suggested that variation in loci relevant to mitochondrial function may explain a disproportionate
220 amount of heritability for these age-related traits. We tested mitochondria-relevant common variants on
221 the nucDNA and mtDNA in the GWAS Catalog, UKB, and meta-analyses using Fisher's exact tests, S-LDSC,
222 and MAGMA. We found no convincing evidence of heritability enrichment among mitochondria-relevant
223 loci in any tested trait, cohort, or method. Of 10 tested organelles, only the nucleus showed enrichment,
224 and further analysis revealed that much of this signal originated from transcription factors. These
225 contrasting results were corroborated by constraint analysis, which showed a substantial fitness cost to
226 heterozygous loss-of-function mutation in genes encoding the nuclear proteome, whereas genes
227 encoding the mitochondrial proteome were "haplosufficient."

228 For highly polygenic and well-powered traits, any large fraction of the genome may explain a statistically
229 significant amount of disease heritability[65,66]. Indeed, individual associations between mitochondria-
230 relevant loci and certain common diseases have been identified previously[67,68]. As associations have also
231 been identified among loci relevant for other organelles, enrichment analyses can place these complex
232 genetic architectures in a broader biological context and prioritize pathways for follow-up. Importantly,
233 both MAGMA and S-LDSC are capable of detecting an enrichment even in a highly polygenic background.
234 Both methods have been used in the past to identify biologically plausible disease-relevant tissues[36,49] and
235 pathway enrichments[69,70] in traits across the spectrum of polygenicity, and we identify enrichments
236 among disease-relevant tissues using both methods in several highly polygenic traits.

237 Our comprehensive analysis of common inherited mitochondrial variation is in agreement with the
238 literature and represents, to our knowledge, the most comprehensive assessment of such variation in
239 age-related diseases to date. Our negative results are consistent with a small number of isolated reports
240 interrogating either mitochondria-relevant nucDNA[71] or mtDNA[51,72–74] loci in cardiometabolic traits,
241 cataracts, osteoporosis, or osteoarthritis. We replicated mtDNA associations with creatinine and AST
242 observed previously in BioBank Japan[51], further supporting our approach. While individual mtDNA
243 variants have been previously associated with certain traits[75–77], these associations appear to be
244 conflicting in the literature, perhaps because of limited power and/or uncontrolled confounding biases
245 such as population stratification[78,79].

246 Though several organelles have been associated with aging and age-associated pathology[3,23,61–64], of the
247 ten organelles we considered, we observed age-related trait heritability enrichment for only one: the
248 nucleus. While previous work has shown that common disease GWAS can be enriched for expression in
249 specific disease-relevant organs[49,80], this framework does not appear to extend to organelles outside the
250 nucleus. This finding contrasts with our classical nosology of (typically recessive) inborn errors of
251 metabolism that tend to be mapped to "causal" organelles, e.g., lysosomal storage diseases, disorders of
252 peroxisomal biogenesis, and mitochondrial OXPHOS disorders. The enrichment of the nucleus appears to
253 derive almost entirely from transcription factors, indicating that common variation influencing genome
254 regulation impacts the inherited risk of common diseases more than variation influencing individual
255 organelles.

256 Our observation that MitoCarta genes are pLoF tolerant (**Figure 4A**) is consistent with the rare disease
257 literature. Namely, while mutations in over 300 nucDNA genes can produce severe Mendelian
258 mitochondrial disease[19], the vast majority tend to be autosomal recessive with healthy parents. The few
259 disorders that show dominant inheritance are nearly always due to dominant negativity rather than
260 haploinsufficiency. In combination with these observations, the results from the current study indicate
261 that mitochondrial pathways are generally haplosufficient.

262 The observed pLoF-intolerance among TFs (**Figure 4A**) provide a stark contrast to the results from the
263 mitochondria that is borne out in their associated Mendelian disease syndromes: TFs are known to be
264 haploinsufficient[81] and even regulatory variants modulating their expression can produce severe
265 Mendelian disease[82]. We observe heritability enrichment among TFs for 10 different diseases, consistent
266 with observed elevated purifying selection against pLoF variants in these genes. Our enrichment results
267 combined with pLoF intolerance suggest that variation among TFs may produce disease-associated
268 variants with larger effect sizes than expectation. Overall, gene-regulatory pathways at the level of the
269 TFs appear to be less resilient to common variation and more susceptible to conferring disease risk upon
270 genetic perturbation. These results highlight the importance of gene regulators and suggest that they may
271 serve as important genetic "levers" for the modulation of pathways relevant to common disease
272 heritability.

273 Why are mitochondria so robust to variation in gene dosage and hence haplosufficient? We propose three
274 possibilities. First, pathway redundancy may underlie robustness to mitochondrial insults, consistent with
275 the longstanding observation that in cell culture, defective OXPHOS can be supported thanks to the action
276 of non-mitochondrial pathways such as cytosolic glycolysis and nucleotide salvage as long as key
277 environmental nutrients are provided[83]. Second, mitochondrial pathways tend to be highly
278 interconnected, and it was already proposed by Wright[84] and later by Kacser and Burns[85] that
279 haplosufficiency arises as a consequence of physiology, i.e., network organization of metabolic reactions.
280 Kacser and Burns in fact explicitly mention that noncatalytic gene products fall outside their framework,
281 and we believe that our finding that nucleus-localizing and cytoskeletal genes are the two most pLoF-
282 intolerant compartments is consistent with their assessment. Third, mitochondria were formerly
283 autonomous microbes and hence may have retained vestigial layers of "intra-organelle buffering" against
284 genetic variation. Numerous feedback control mechanisms, including respiratory control[86], help to ensure
285 organelle robustness across physiological extremes[87] (e.g. resting versus exercising[88], shifting nutrient
286 availability). In agreement, a recent CRISPR screen showed that of the genes for which knock-out modified
287 survival under a mitochondrial poison, there is a striking over-representation of genes that themselves
288 encode mitochondrial proteins[89].

289 Throughout this study, we have tested inherited common variant associations via an additive genetic
290 model. We acknowledge the limitations of focusing on a specific genetic model and variant frequency
291 regime, though note that common variation is the largest documented source of narrow-sense

292 heritability, which typically accounts for a majority of disease heritability[90,91]. First, we consider only
293 common variants. While rare variants may prove to be instructive, it is notable that a previous rare variant
294 analysis in T2D[92] failed to show enrichment among OXPHOS genes. Second, we consider only additive
295 genetic models. A recessive model may be particularly fruitful for mitochondria-relevant loci given their
296 tolerance to pLoF variation, however these models are frequently power-limited and may not explain
297 much more phenotypic variance than additive models[93,94]. Third, we have not considered epistasis. The
298 effects of mtDNA-nucDNA interactions[95] in common diseases have yet to be assessed. While there is
299 debate about whether biologically-relevant epistasis can be simply captured by main effects[91,93,96,97] at
300 individual loci, it is possible that modeling mtDNA-nucDNA interactions will reveal new contributions.
301 Finally, it is crucial not to confuse our results with previously reported associations between somatic
302 mtDNA mutations and age-associated disease[13–15] – the present work is focused on germline variation.

303 Our study does not formally address the causality of mitochondrial dysfunction in common age-related
304 disease. Rather, we have tested if common variants in mitochondrial pathways tend to explain a
305 disproportionate amount of age-related disease heritability. The observed lack of heritability enrichment
306 in mitochondrial pathways does not preclude the possibility of a therapeutic benefit in targeting the
307 mitochondrion for age-related disease. For example, mitochondrial dysfunction is documented in brain or
308 heart infarcts following blood vessel occlusion in laboratory-based models[98,99]. Though mitochondrial
309 variants do not influence infarct risk in this laboratory model, pharmacological blockade of the
310 mitochondrial permeability transition pore can mitigate reperfusion injury and infarct size[100]. Future
311 studies will be required to determine if and how the mitochondrial dysfunction associated with common
312 age-associated diseases can be targeted for therapeutic benefit.

313 Our finding that the nucleus is the only organelle that shows enrichment for common age-associated trait
314 heritability builds on prior work implicating nuclear processes in aging. Most human progeroid syndromes
315 result from monogenic defects in nuclear components[101] (e.g., *LMNA* in Hutchinson-Gilford progeria
316 syndrome, *TERC* in dyskeratosis congenita), and telomere length has long been observed as a marker of
317 aging[102]. Heritability enrichment of age-related traits among gene regulators is consistent with the
318 epigenetic dysregulation[103] and elevated transcriptional noise[3,104] observed in aging (e.g., *SIRT6*
319 modulation influences mouse longevity[105] and metabolic syndrome[62]). An important role for gene
320 regulation in common age-related disease is in agreement with both the observation that a very large
321 fraction of common disease-associated loci corresponds to the non-coding genome and the enrichment
322 of disease heritability in histone marks and transcription factor binding sites[36,106].

323

324 **Acknowledgements**
325 We thank D. Altshuler, S.E. Calvo, H. Finucane, E.S. Lander, M.E. MacDonald, D. Palmer, E.B. Robinson,
326 A.V. Segrè, M.E. Talkowski, R.K. Walters, C.C. Winter, and members of the Mootha and Neale labs for
327 critical feedback and discussions. This research has been conducted using the UK Biobank Resource under
328 Application Number 31063. This project was supported in part by grants (NIH R35GM122455 to V.K.M.
329 and NIH T32 AG000222 to R.G.) from the National Institutes of Health. VKM is an Investigator of the
330 Howard Hughes Medical Institute.

331

332 **Data Availability**
333 Genetic correlation point estimates and standard errors plotted in Figure 1B is available in Table S2.
334 Summary statistics from mtDNA-GWAS available in Table S6. All gene-based enrichment analysis p-values
335 and point estimates are available in Table S8. Literature-reported loci associated with biomarkers of
336 mitochondrial function after clumping and QC are available in Table S4. Period prevalence data for
337 diseases in the UK can be obtained from Kuan et al. 2019. Gene-sets can be found using COMPARTMENTS

(https://compartments.jensenlab.org), MitoCarta 2.0 (https://www.broadinstitute.org/files/shared/metabolism/mitocarta/human.mitocarta2.0.html), Lambert et al. 2018 (DOI: 10.1016/j.cell.2018.01.029), Frazier et al. 2019 (DOI: 10.1074/jbc.R117.809194), Finucane et al. 2018 (https://alkesgroup.broadinstitute.org/LDSCORE/), Kapopoulou et al. 2015 (DOI: 10.1111/evo.12819), and the Macarthur laboratory (https://github.com/macarthur-lab/gene_lists). Gene age estimates were obtained from Litman, Stein 2019 (DOI: 10.1053/j.seminoncol.2018.11.002). GWAS catalog annotations can be obtained from: https://www.ebi.ac.uk/gwas. UKB heritability estimates can be obtained at: https://nealelab.github.io/UKBB_ldsc/. UKB summary statistics can be obtained from Neale lab GWAS round 2: https://github.com/Nealelab/UK_Biobank_GWAS. Annotations for the Baseline v1.1 and BaselineLD v2.2 models as well as other relevant reference data, including the 1000G EUR reference panel, can be obtained from https://alkesgroup.broadinstitute.org/LDSCORE/. eQTL and expression data in human tissues can be obtained from GTEx (https://www.gtexportal.org). Constraint estimates can be found via gnomAD: https://gnomad.broadinstitute.org. See citations for publicly available GWAS meta-analysis summary statistics[26–35].

Code Availability

Our analysis leverages publicly available tools including LDSC for heritability enrichment and genetic correlation (https://github.com/bulik/ldsc), MAGMA v1.07b for gene-set enrichment analysis (https://ctg.cncr.nl/software/magma), PLINK v1.07 for linkage disequilibrium clumping (https://zzz.bwh.harvard.edu/plink/), and Hail v0.2.51 for distributed computing and mtDNA GWAS (https://hail.is).

Competing Interests

VKM is an advisor to and receives compensation or equity from Janssen Pharmaceuticals, 5am Ventures, and Raze Therapeutics. BMN is a member of the scientific advisory board at Deep Genomics and RBNC Therapeutics. BMN is a consultant for Camp4 Therapeutics, Takeda Pharmaceutical and Biogen.

Author Contributions

R.G., B.M.N., and V.K.M. conceived of the project; R.G., K.J.K., D.H. designed analyses; R.G. performed analyses; B.M.N., V.K.M. supervised project; R.G. and V.K.M. wrote the manuscript with input from other authors.

Online Methods

Trait selection:

Sex-standardized period prevalence of over 300 diseases was obtained from an extensive survey of the National Health Service in the UK as reported previously[24]. To select high prevalence late-onset diseases, we ranked diseases with a median onset over 50 years of age by the sum of the period prevalence of all age categories above 50. We selected the top 30 diseases using this metric and manually mapped these traits to similar or equivalent phenotypes with publicly available summary statistics from UKB and/or well-powered meta-analyses (e.g., Parkinson's Disease and Alzheimer's Disease for dementia) resulting in 24 traits with data available in UKB, meta-analyses, or both (**Table S1**).

Criteria for inclusion of summary statistics:

We manually mapped selected age-related diseases and traits to corresponding phenotypes in UKB. In parallel, we searched the literature to identify well-powered EUR-predominant GWAS (referred to as meta-analyses) that (1) used primarily non-targeted arrays, (2) had publicly available full summary statistics, and (3) did not enroll individuals from UKB to serve as independent replication (**Supplementary note**). For UKB, we obtained heritability estimates (https://github.com/Nealelab/UKBB_ldsc) previously computed using stratified linkage-disequilibrium score regression (S-LDSC, https://github.com/bulik/ldsc)[36] atop the BaselineLD v1.1 model using reference LD scores computed from 1000G EUR. For meta-analyses, we computed heritability estimates with S-LDSC atop the updated BaselineLD v2.2 model using reference LD scores computed from 1000G EUR (https://alkesgroup.broadinstitute.org/LDSCORE/). We computed the heritability Z-score, a statistic that captures sample size, polygenicity, and heritability[36], and included only traits with heritability Z-score > 4 (**Supplementary note**) for further analysis.

Genetic correlations in UKB:

Pairwise genetic correlations, r_g, were computed using linkage-disequilibrium score correlation[37] on all selected age-related traits with heritability Z-score > 4. We used UKB summary statistics (https://github.com/Nealelab/UK_Biobank_GWAS) for all sufficiently powered traits; summary statistics from meta-analyses were used for eGFR[33], Alzheimer's Disease[35], and Parkinson's Disease[34] as these traits showed heritability Z-score > 4 within meta-analyses but not in UKB **(Table S1)**. P-values for genetic correlation represented deviation from the null hypothesis $r_g = 0$. Traits were ordered by their contribution to the first eigenvector of the absolute value of the correlation matrix, with point estimates and standard errors available in **Table S2**. Bonferroni correction was applied producing a p-value cutoff of $0.05/\binom{24}{2} = 1.81 * 10^{-4}$.

Assessment of mitochondria-localizing genes in the GWAS Catalog:

We mapped variants in the GWAS Catalog (obtained on September 5th, 2019, https://www.ebi.ac.uk/gwas/) meeting genome-wide significance (p < 5e-8) to genes using provided annotations, producing a set of trait-associated genes for each trait. We manually selected phenotypes represented in the GWAS Catalog matching our set of age-associated traits with over annotated 30 trait-associated genes. For each trait, we computed the proportion of trait associated genes that were mitochondria-localizing (defined via MitoCarta2.0[17]) and tested for enrichment or depletion relative to overall genome background using two-sided Fisher's exact tests correcting for multiple hypothesis tests with the Benjamini-Hochberg (BH) procedure at FDR q-value < 0.1.

We also computed the test statistic N_g^{enrich}, defined as the number of age-associated traits showing a nominal (not necessarily statistically significant) enrichment for a given gene-set g, for the MitoCarta genes. We then generated an empirical null distribution for N_g^{enrich}. We drew 1,000 random samples of

protein-coding genes, where each sample contained the same number of genes as the set of mitochondria-localizing genes and computed N_g^{enrich} for each of these gene-sets **(Figure S1B)**. The one-sided p-value, defined as $\Pr\left(N_g^{enrich} \leq x\right)$ under the null, was subsequently obtained.

We expanded our enrichment/depletion analysis to all 332 traits in the GWAS Catalog with over 30 trait-associated genes; for enrichment or depletion testing, we used two-sided Fisher's exact tests and corrected for multiple hypothesis testing with the BH procedure at FDR q-value < 0.1.

Enrichment analysis of literature-curated mitochondria-associated phenotypes:

We reviewed the literature for quantitative trait loci (QTLs) for mtDNA copy number (mtCN)[40,41], mtRNA abundance/modification[42,43], and biomarkers of OXPHOS dysfunction (namely lactate, pyruvate, lactate/pyruvate ratio[46–48], and GDF15 abundance[44,45]) (**Supplementary note**). We subsequently used PLINK v1.07 (https://zzz.bwh.harvard.edu/plink/)[107] to identify independent variants for each phenotype based on the 1000G EUR reference panel (**Supplementary note**). To test for overlap with UKB age-associated disease traits, we divided curated variants into three classes: mtCN-related (21 variants), mtRNA-related (78 variants), and OXPHOS biomarkers (62 variants). For each of the 21 UKB age-related disease traits, we computed the number of genome-wide significant (p < 5e-8) variants that overlapped the curated variants for each class, termed $N_c^{overlap}$ where c is the class. We only considered variants with INFO > 0.8 and MAF > 0.001 or expected case MAC > 25. For significance testing, we generated an empirical null distribution around $N_c^{overlap}$ including only variants with INFO > 0.8. For each class, we drew variants at random 2500 times matching on LD score, in-sample MAF, and distance to transcription start site (where the distance metric was set to 0 if the variant was located within a gene boundary). LD scores per variant were generated per-chromosome with a 1 cm window using the 1000G EUR reference panel. The $N_c^{overlap}$ was then computed for each category for each set of randomly selected variants, generating a category specific empirical null distribution for the statistic (**Figure S2**). The one-sided p-value, defined as $\Pr\left(N_c^{overlap} \geq x\right)$ under the null, was subsequently obtained. To correct for multiple hypothesis testing, we applied the BH procedure with FDR < 0.1 and also applied a Bonferroni threshold of $P = \frac{0.1}{21*3} \approx 0.0016$.

Harmonization and filtering of summary statistics for LDSC and MAGMA:

UKB summary statistics previously formatted for use with LDSC and filtered to HapMap3 (HM3) SNPs (https://github.com/Nealelab/UKBB_ldsc) were used for analysis with S-LDSC. For analysis with MAGMA v1.07b[50], we included variants from the full Neale Lab UKB Round 2 GWAS summary statistics (https://github.com/Nealelab/UK_Biobank_GWAS) with INFO > 0.8 and MAF > 0.01, and excluded any variants flagged as low confidence (a heuristic defined by MAF < 0.001 or expected case MAC < 25).

Summary statistics obtained from publicly available GWAS meta-analyses[26–35] were reported in varied formats. We manually verified the genome build upon which each meta-analysis reported results and ensured that all sets of summary statistics contained columns listing P-value, variant rsID, genome-build specific coordinates, and if available, variant-specific sample size (**Table S1**). If variant coordinates or rsID were not provided, the relevant columns were obtained from dbSNP database version 130 (for hg18) or 146 (for hg19). We used the summary statistic munging script provided with S-LDSC (https://github.com/bulik/ldsc) to generate summary statistics compatible with S-LDSC, restricting to HM3 SNPs as these tend to be best behaved for analysis with LDSC. For use of meta-analyses with MAGMA[50], we restricted analysis to variants with INFO > 0.8 and MAF > 0.01 if such information was provided.

Multiple testing correction for gene-set enrichment analysis:

To account for the multiple hypothesis tests performed throughout this study, we obtained p-value thresholds via the BH procedure at FDR < 0.1 for all gene-sets assessed for a given method and cohort type (where the two cohort types were UKB and meta-analysis).

Gene-set based enrichment analysis:

We extensively use S-LDSC and MAGMA to perform gene-set enrichment analyses among GWAS summary statistics. To test enrichment with S-LDSC, SNPs were mapped to each gene with a 100kb symmetric window as recommended[49] and LD scores were computed using the 1000G EUR reference panel (https://alkesgroup.broadinstitute.org/LDSCORE/) and subsequently restricted to the HM3 SNPs. We used S-LDSC to test for heritability enrichment controlling for 53 annotations including coding regions, enhancer regions, 5' and 3' UTRs, and others as previously described[36] (baseline v1.1, referred to as baseline model hereafter). We also used MAGMA with both 5kb up, 1.5kb down and 100kb symmetric windows to test for enrichment. MAGMA gene-level analysis was performed with the 1000G EUR LD reference panel to account for LD structure, and gene-set analysis was performed including covariates for gene length, variant density, inverse minor allele count (MAC), as well as log-transformed versions of these covariates. Statistical tests for both S-LDSC and MAGMA were one-sided, considering enrichment only. For both methods, we included the relevant superset of genes as a control to ensure that our analysis was competitive (**Supplementary note**). We refer to this approach as the 'usual approach'. All enrichment effect size estimates and p-values are available in **Table S8**.

Enrichment analysis of genes comprising the mitochondrial proteome:

We obtained the set of nuclear-encoded mitochondria-localizing genes using MitoCarta2.0[17] and used the literature to obtain the subset of MitoCarta genes involved in inherited mitochondrial disease[19] as well as those producing components of oxidative phosphorylation (OXPHOS) complexes. We used both S-LDSC and MAGMA to test for enrichment in the usual way (**Methods**) controlling for the set of protein-coding genes to ensure a competitive analysis (**Supplementary note**). We also tested mitochondria-localizing genes for enrichment in meta-analyses using S-LDSC and MAGMA with the same parameters as for UKB traits (**Supplementary note**).

Tissue-expressed gene-set enrichment analysis:

To obtain the set of genes most expressed in a given tissue versus others, we obtained t-statistics computed from GTEx v6 gene-level transcript-per-million (TPM) data corrected for age and sex as published previously[49]. For each tissue, we selected the top 2485 genes (10%) with the highest t-statistics for tissue-specific expression, producing tissue-expressed gene-sets. We selected nine tissues based on expectation of enrichment for our tested traits in UKB (e.g., liver for LDL levels, esophageal mucosa for GERD). We used both S-LDSC and MAGMA to test for enrichment in the usual way (**Methods**) controlling for the set of tissue-expressed genes to ensure a competitive analysis (**Supplementary note**). Tissue-expressed gene-set analyses were performed on meta-analyses with S-LDSC and MAGMA on the same tissues using the same parameters as used in UKB.

Power analysis:

To test for the effects of gene-set size on power, we selected ten positive control tissue-trait pairs based on (1) the presence of tissue enrichment in UKB with S-LDSC and MAGMA and (2) if the observed enrichment was biologically plausible. The pairs tested were liver-HDL, liver-LDL, liver-TG, liver-cholesterol, pancreas-glucose, pancreas-type 2 diabetes, atrial appendage-atrial fibrillation, sigmoid colon-diverticular disease, coronary artery-myocardial infarction, and visceral adipose-HDL. We then, in brief, used an empirical sampling-based approach, generating random subsamples of a selected set of tissue-expressed gene-sets at four different gene-set sizes (1523, 1105, 800, and 350 genes), defining

power as the proportion of trials showing a significant enrichment (**Supplementary note**). We used the same sub-sampled gene-sets for enrichment analysis using both S-LDSC and MAGMA in the usual way (**Methods**) controlling for the set of tissue-expressed genes to ensure a competitive analysis (**Supplementary note**). We used the same gene-sets among the subset of the positive control traits that showed enrichment in the corresponding meta-analysis to verify power for the meta-analyses (**Supplementary note**).

Cross-tissue eQTL analysis

We obtained the set of eGenes from GTEx v8 across 49 tissues (https://www.gtexportal.org), filtering to only include cis-eQTLs with q-value < 0.05. To determine how the landscape of cis-eQTLs for MitoCarta genes compared to other protein-coding genes, we regressed the number of tissues with a detected cis-eQTL for a given gene x, N_x^{eQTL}, onto an indicator for membership in a given organellar proteome ($I_x^{organelle}$), controlling for gene length, log gene length, breadth of expression (τ_x), and the number of tissues with detected expression > 5 TPM ($N_x^{express}$**, Supplementary note**). To quantify breadth of expression, we obtained median-per-tissue GTEx v8 TPM expression values and computed τ[108] after removing lowly-expressed genes with maximal cross-tissue TPM < 1, defined as:

$$\tau_x = \frac{\sum_{i=1}^{n}(1-\hat{x}_i)}{n-1} \; where \; \hat{x}_i = \frac{x_i}{\max\limits_{1\le i\le n} x_i}$$

where x_i is the expression of gene x in tissue i with n tissues. τ ranges from 0 to 1, with lower τ indicating broadly expressed gene and higher τ indicating more tissue specific expression patterns. Because GTEx sampled multiple tissue subtypes (e.g., brain sub-regions) that show correlated expression profiles[109] which bias τ_x, N_x^{eQTL}, and $N_x^{express}$ upward, for each broader tissue class (brain, heart, artery, esophagus, skin, cervix, colon, adipose) we selected a single representative tissue when computing these quantities (**Figure S14B, Supplementary note**). We used LD scores computed from the 1000G EUR reference panel. The model, fit via OLS for each tested organelle, was:

$$N_x^{eQTL} \sim I_x^{organelle} + N_x^{express} + \tau_x + \log(gene\ length) + gene\ length$$

mtDNA-wide association study:

We obtained mtDNA genotype data on 265 variants as obtained on the UK Biobank Axiom array and the UK BiLEVE array from the full UKB release[25]. To perform variant QC, we used evoker-lite[110] to generate fluorescence cluster plots per-variant and per-batch and manually inspected the results, removing 19 variants due to cluster plot abnormalities (**Table S5**, **Supplementary note**). We additionally removed any variants with heterozygous calls, within-array-type call rate < 0.95, and with less than 20 individuals with an alternate genotype. For case-control traits, we removed any phenotype-variant pair with an expected case count of alternate genotype individuals of less than 20, resulting in a maximum of 213 variants tested per trait (**Supplementary note**). To perform sample QC, we restricted samples to the same samples from which UKB summary statistics were generated (https://github.com/Nealelab/UK_Biobank_GWAS), namely unrelated individuals 7 standard deviations away from the first 6 European sample selection PCs with self-reported white-British, Irish, or White ethnicity and no evidence of sex chromosome aneuploidy. We additionally removed any samples with within-array-type mitochondrial variant call rate < 0.95, resulting in 360,662 unrelated samples of EUR ancestry. We generated the LD matrix for mitochondrial DNA variants using Hail v0.2.51 (https://hail.is) pairwise for all 213 variants tested across all post-QC samples.

We ran mtDNA-GWAS for all 21 UKB age-related phenotypes as well as creatinine and AST using Hail v0.2.51 via linear regression controlling for the first 20 PCs of the nuclear genotype matrix, sex, age, age^2, sex*age, and sex*age^2 as performed for the UKB GWAS (https://github.com/Nealelab/UK_Biobank_GWAS). We also used Hail to run Firth logistic regression with the same covariates for case/control traits (**Table S1**). As we observed that some mitochondrial DNA variants were specific to array type, we also ran linear regression including array type as a covariate; we did not perform logistic regression with array type as a covariate due to convergence issues secondary to complete separation of variants assessed only on only array type. We defined mtDNA-wide significance using a Bonferroni correction by $p = \frac{0.05}{4337} \approx 1.15e - 5$.

Enrichment analysis of components of organellar proteomes:

COMPARTMENTS (https://compartments.jensenlab.org)[52] is a resource integrating several lines of evidence for protein localization predictions including annotations, text-mining, sequence predictions, and experimental data from the Human Protein Atlas. We used this resource to obtain the degree of evidence (a number ranging from 0 to 5) linking each gene to localization to one of 12 organelles: nucleus, cytosol, cytoskeleton, peroxisome, lysosome, endoplasmic reticulum, Golgi apparatus, plasma membrane, endosome, extracellular space, mitochondrion, and proteasome. To avoid noisy localization assignments due to weak text mining and prediction evidence, we only considered localization assignments with a score > 2 as described previously[52]. We subsequently assigned compartment(s) to each gene by selecting the compartment(s) with the maximal score within each gene. We only included compartments containing over 240 genes due to limited power at these smaller gene-set sizes and used MitoCarta2.0[17] to obtain a higher confidence set of genes localizing to the mitochondrion, resulting in gene-sets representing the proteomes of 10 organelles. S-LDSC and MAGMA were used to test for enrichment across the UKB age-related traits for these gene-sets in the usual way, controlling for the set of protein-coding genes. S-LDSC was also used to obtain estimates of the percentage of heritability explained by each organelle gene-set.

Enrichment analysis of spatial components of the nucleus:

To produce interpretable sub-divisions of the nucleus, we used Gene Ontology (GO)[53,54] to identify terms listed as children of the nucleus cellular component (GO:0005634). We used Ensembl version 99[111] to obtain a first pass set of genes annotated to each sub-compartment of the nucleus (or its children). After manual review of sub-compartments with > 90 genes, we selected nucleoplasm (GO:0005654), nuclear chromosome (GO:0000228), nucleolus (GO:0005730), nuclear envelope (GO:0005635), splicosomal complex (GO:0005681), nuclear DNA-directed RNA polymerase complex (GO:0055029), and nuclear pore (GO:0005643). We excluded terms listed as 'part' due to poor interpretability and manually excluded similar terms (e.g., nuclear lumen vs nucleoplasm). To generate a high confidence set of genes localizing to each of these selected sub-compartments, we then turned to the COMPARTMENTS resource which assigns localization confidence scores for each protein to GO cellular component terms. We assigned members of the nuclear proteome to these selected nuclear sub-compartments using same the approach outlined for the organelle analysis (**Methods**). After filtering our selected sub-compartments to those containing > 240 genes, we obtained four categories: nucleoplasm, nuclear chromosome, nucleolus, and nuclear envelope. The nuclear chromosome annotation was largely overlapping with a manually curated high-quality list of transcription factors[55] however was not exhaustive; as such, we merged these lists to generate the chromosome and TF category. To improve interpretability, we removed genes from nucleoplasm that were also assigned to another nuclear sub-compartment, constructed a list of other nucleus-localizing proteins not captured in these four sub-compartments, and included only genes annotated as localizing to the nucleus (**Methods**). S-LDSC and MAGMA were used to test for enrichment

across the UKB age-related traits for these gene-sets in the usual way while controlling for the set of protein-coding genes (**Methods**).

Enrichment analysis of functionally distinct TF subsets:
We used a published curated high-quality list of TFs[55] to partition the Chromosome and TF category into transcription factors and other chromosomal proteins. To determine which TFs are broadly expressed versus tissue specific, we computed τ per TF across all selected tissues after removing lowly-expressed genes with maximal cross-tissue TPM < 1 (**Methods, Supplementary note**). The threshold for tissue-specific genes was set at $\tau \geq 0.76$ based on the location of the central nadir of the resultant bimodal distribution (**Figure S14A**). To identify terciles of TFs by age, we obtained relative gene age assignments for each gene previously generated by obtaining the modal earliest ortholog level across several databases mapped to 19 ordered phylostrata[112]. DNA binding domain (DBD) annotations for the TFs were obtained from previous manual curation efforts[55]. S-LDSC and MAGMA were used to test for enrichment across the UKB age-related traits for these gene-sets in the usual way while controlling for the set of protein-coding genes (**Methods**). We also tested TFs for enrichment in meta-analyses using S-LDSC and MAGMA with the same parameters as for UKB traits (**Supplementary note**).

Analysis of constraint across organelles and sub-organellar gene-sets:
We obtained gene-level gnomAD v2.1.1 constraint tables (https://gnomad.broadinstitute.org), haploinsufficient genes, and olfactory receptors[60] (https://github.com/macarthur-lab/gene_lists). Constraint values as loss-of-function observed/expected fraction (LOEUF) were mapped to genes within organelle, sub-mitochondrial, sub-nuclear, and TF binding domain gene-sets.

Enrichment analysis across age-related disease holding constraint as a covariate:
To test for enrichment with constraint as a covariate, we used MAGMA with UKB age-related traits. We mapped variants to genes and performed the gene-level analysis as done previously for the mitochondria-localizing gene and organelle analysis. We included LOEUF and log LOEUF as covariates for the gene-set analysis in addition to the default covariates (gene length, SNP density, inverse MAC, as well as the respective log-transformed versions) via the `--condition-residualize` flag.

631 1. Wang K, Gaitsch H, Poon H, Cox NJ, Rzhetsky A. Classification of common human diseases derived
632 from shared genetic and environmental determinants. *Nat Genet*. 2017;49(9):1319-1325.
633 doi:10.1038/ng.3931
634 2. Claussnitzer M, Cho JH, Collins R, et al. A brief history of human disease genetics. *Nature*.
635 2020;577(7789):179-189. doi:10.1038/s41586-019-1879-7
636 3. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*.
637 2013;153(6):1194. doi:10.1016/j.cell.2013.05.039
638 4. Fleischman A, Makimura H, Stanley TL, et al. Skeletal muscle phosphocreatine recovery after
639 submaximal exercise in children and young and middle-aged adults. *J Clin Endocrinol Metab*.
640 2010;95(9):69-74. doi:10.1210/jc.2010-0527
641 5. Petersen KF, Dufour S, Befroy D, Garcia R, Shulman GI. Impaired Mitochondrial Activity in the
642 Insulin-Resistant Offspring of Patients with Type 2 Diabetes. *N Engl J Med*. 2004;350(7):664-671.
643 doi:10.1056/nejmoa031314
644 6. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1α-responsive genes involved in oxidative
645 phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267-
646 273. doi:10.1038/ng1180
647 7. Fannin SW, Lesnefsky EJ, Slabe TJ, Hassan MO, Hoppel CL. Aging selectively decreases oxidative
648 capacity in rat heart interfibrillar mitochondria. *Arch Biochem Biophys*. 1999;372(2):399-407.
649 doi:10.1006/abbi.1999.1508
650 8. Schapira AH V, Cooper JM, Dexter D, Clark JB, Jenner P, Marsden CD. Mitochondrial Complex I
651 Deficiency in Parkinson's Disease. *J Neurochem*. 1990;54(3):823-827. doi:10.1111/j.1471-
652 4159.1990.tb02325.x
653 9. Trounce I, Byrne E, Marzuki S. Decline in Skeletal Muscle Mitochondrial Respiratory Chain
654 Function: Possible Factor in Ageing. *Lancet*. 1989;333(8639):637-639. doi:10.1016/S0140-
655 6736(89)92143-0
656 10. Kelley DE, He J, Menshikova E V., Ritov VB. Dysfunction of mitochondria in human skeletal muscle
657 in type 2 diabetes. *Diabetes*. 2002;51(10):2944-2950. doi:10.2337/diabetes.51.10.2944
658 11. Patti ME, Butte AJ, Crunkhorn S, et al. Coordinated reduction of genes of oxidative metabolism in
659 humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. *Proc Natl Acad Sci*
660 *U S A*. 2003;100(14):8466-8471. doi:10.1073/pnas.1032913100
661 12. Stump CS, Short KR, Bigelow ML, Schimke JM, Nair KS. Effect of insulin on human skeletal muscle
662 mitochondrial ATP production, protein synthesis, and mRNA transcripts. *Proc Natl Acad Sci U S A*.
663 2003;100(13):7996-8001. doi:10.1073/pnas.1332551100
664 13. Bender A, Krishnan KJ, Morris CM, et al. High levels of mitochondrial DNA deletions in substantia
665 nigra neurons in aging and Parkinson disease. *Nat Genet*. 2006;38(5):515-517.
666 doi:10.1038/ng1769
667 14. Taylor RW, Barron MJ, Borthwick GM, et al. Mitochondrial DNA mutations in human colonic crypt
668 stem cells Find the latest version : Mitochondrial DNA mutations in human colonic crypt stem
669 cells. *J Clin Invest*. 2003;112(9):1351-1360. doi:10.1172/JCI200319435.Introduction
670 15. Wanagat J, Cao Z, Pathare P, Aiken JM. Mitochondrial DNA deletion mutations colocalize with
671 segmental electron transport system abnormalities, muscle fiber atrophy, fiber splitting, and
672 oxidative damage in sarcopenia. *FASEB J*. 2001;15(2):322-332. doi:10.1096/fj.00-0320com
673 16. Ashar FN, Zhang Y, Longchamps RJ, et al. Association of mitochondrial DNA copy number with
674 cardiovascular disease. *JAMA Cardiol*. 2017;2(11):1247-1255. doi:10.1001/jamacardio.2017.3683
675 17. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: An updated inventory of mammalian
676 mitochondrial proteins. *Nucleic Acids Res*. 2016;44(D1):D1251-D1257. doi:10.1093/nar/gkv1003
677 18. Schon EA, Dimauro S, Hirano M. Human mitochondrial DNA: Roles of inherited and somatic
678 mutations. *Nat Rev Genet*. 2012;13(12):878-890. doi:10.1038/nrg3275

679 19. Frazier AE, Thorburn DR, Compton AG. Mitochondrial energy generation disorders: Genes,
680 mechanisms, and clues to pathology. *J Biol Chem*. 2019;294(14):5386-5395.
681 doi:10.1074/jbc.R117.809194
682 20. Shammas MA. Telomeres, lifestyle, cancer, and aging. *Curr Opin Clin Nutr Metab Care*.
683 2011;14(1):28-34. doi:10.1097/MCO.0b013e32834121b1
684 21. Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA. Association between telomere
685 length in blood and mortality in people aged 60 years or older. *Lancet*. 2003;361(9355):393-395.
686 doi:10.1016/S0140-6736(03)12384-7
687 22. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-
688 digestion. *Nature*. 2008;451(7182):1069-1075. doi:10.1038/nature06639
689 23. Özcan U, Cao Q, Yilmaz E, et al. Endoplasmic Reticulum Stress Links Obesity, Insulin Action, and
690 Type 2 Diabetes. *Science (80-)*. 2004;306(5695):457 LP - 461.
691 http://science.sciencemag.org/content/306/5695/457.abstract.
692 24. Kuan V, Denaxas S, Gonzalez-Izquierdo A, et al. A chronological map of 308 physical and mental
693 health conditions from 4 million individuals in the English National Health Service. *Lancet Digit*
694 *Heal*. 2019;1(2):e63-e77. doi:10.1016/s2589-7500(19)30012-3
695 25. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the
696 Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med*. 2015;12(3):1-10.
697 doi:10.1371/journal.pmed.1001779
698 26. Teslovich TM, Musunuru K, Smith A V, et al. Biological, clinical and population relevance of 95 loci
699 for blood lipids. *Nature*. 2010;466(7307):707-713. http://dx.doi.org/10.1038/nature09270.
700 27. Ehret GB, Munroe PB, Rice KM, et al. Genetic variants in novel pathways influence blood pressure
701 and cardiovascular disease risk. *Nature*. 2011;478(7367):103-109. doi:10.1038/nature10405
702 28. Manning AK, Hivert M-F, Scott RA, et al. A genome-wide approach accounting for body mass
703 index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat*
704 *Genet*. 2012;44(6):659-669. http://dx.doi.org/10.1038/ng.2274.
705 29. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into
706 the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*. 2012;44(9):981-990.
707 doi:10.1038/ng.2383
708 30. Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new
709 susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43(4):333-340.
710 doi:10.1038/ng.784
711 31. Estrada K, Styrkarsdottir U, Evangelou E, et al. Genome-wide meta-analysis identifies 56 bone
712 mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet*.
713 2012;44(5):491-501. doi:10.1038/ng.2249
714 32. Christophersen IE, Rienstra M, Roselli C, et al. Large-scale analyses of common and rare variants
715 identify 12 new loci associated with atrial fibrillation. *Nat Genet*. 2017;49(6):946-952.
716 doi:10.1038/ng.3843
717 33. Pattaro C, Teumer A, Gorski M, et al. Genetic associations at 53 loci highlight cell types and
718 biological pathways relevant for kidney function. *Nat Commun*. 2016;7:1-19.
719 doi:10.1038/ncomms10023
720 34. Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and
721 heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet*
722 *Neurol*. 2019;18(12):1091-1102. doi:10.1016/S1474-4422(19)30320-5
723 35. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11
724 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45(12):1452-1458.
725 doi:10.1038/ng.2802
726 36. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional annotation

727 using genome-wide association summary statistics. *Nat Genet*. 2015;47(11):1228-1235.
728 doi:10.1038/ng.3404
729 37. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human
730 diseases and traits. *Nat Genet*. 2015;47(11):1236-1241. doi:10.1038/ng.3406
731 38. Wasmer K, Eckardt L, Breithardt G. Predisposing factors for atrial fibrillation in the elderly. *J*
732 *Geriatr Cardiol*. 2017;14(3):179-184. doi:10.11909/j.issn.1671-5411.2017.03.010
733 39. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide
734 association studies (GWAS Catalog). *Nucleic Acids Res*. 2016;45(November 2016):gkw1133.
735 doi:10.1093/nar/gkw1133
736 40. Cai N, Li Y, Chang S, et al. Genetic Control over mtDNA and Its Relationship to Major Depressive
737 Disorder. *Curr Biol*. 2015;25(24):3170-3177. doi:10.1016/j.cub.2015.10.065
738 41. Guyatt AL, Brennan RR, Burrows K, et al. A genome-wide association study of mitochondrial DNA
739 copy number in two population-based cohorts. *Hum Genomics*. 2019;13(1):6.
740 doi:10.1186/s40246-018-0190-2
741 42. Ali AT, Boehme L, Carbajosa G, Seitan VC, Small KS, Hodgkinson A. Nuclear genetic regulation of
742 the human mitochondrial transcriptome. *Elife*. 2019;8:1-23. doi:10.7554/eLife.41927
743 43. Ali AT, Idaghdour Y, Hodgkinson A. Analysis of mitochondrial m1A/G RNA modification reveals
744 links to nuclear genetic variants and associated disease processes. *Commun Biol*. 2020;3(1):1-11.
745 doi:10.1038/s42003-020-0879-3
746 44. Folkersen L, Fauman E, Sabater-Lleal M, et al. Mapping of 79 loci for 83 plasma protein
747 biomarkers in cardiovascular disease. *PLoS Genet*. 2017;13(4):1-21.
748 doi:10.1371/journal.pgen.1006706
749 45. Jiang J, Thalamuthu A, Ho JE, et al. A meta-analysis of genome-wide association studies of growth
750 differentiation factor-15 concentration in blood. *Front Genet*. 2018;9(MAR):1-13.
751 doi:10.3389/fgene.2018.00097
752 46. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood
753 metabolites. *Nat Genet*. 2014;46(6):543-550. doi:10.1038/ng.2982
754 47. Suhre K, Shin SY, Petersen AK, et al. Human metabolic individuality in biomedical and
755 pharmaceutical research. *Nature*. 2011;477(7362):54-62. doi:10.1038/nature10354
756 48. Raffler J, Friedrich N, Arnold M, et al. Genome-Wide Association Study with Targeted and Non-
757 targeted NMR Metabolomics Identifies 15 Novel Loci of Urinary Human Metabolic Individuality.
758 *PLoS Genet*. 2015;11(9):1-28. doi:10.1371/journal.pgen.1005487
759 49. Finucane HK, Reshef YA, Anttila V, et al. Heritability enrichment of specifically expressed genes
760 identifies disease-relevant tissues and cell types. *Nat Genet*. 2018;50(4):621-629.
761 doi:10.1038/s41588-018-0081-4
762 50. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS
763 Data. *PLoS Comput Biol*. 2015;11(4):1-19. doi:10.1371/journal.pcbi.1004219
764 51. Yamamoto K, Sakaue S, Matsuda K, et al. Genetic and phenotypic landscape of the mitochondrial
765 genome in the Japanese population. *Commun Biol*. 2020;3(1):104. doi:10.1038/s42003-020-0812-
766 9
767 52. Binder JX, Pletscher-Frankild S, Tsafou K, et al. COMPARTMENTS: Unification and visualization of
768 protein subcellular localization evidence. *Database*. 2014;2014:1-9.
769 doi:10.1093/database/bau012
770 53. Carbon S, Douglass E, Dunn N, et al. The Gene Ontology Resource: 20 years and still GOing
771 strong. *Nucleic Acids Res*. 2019;47(D1):D330-D338. doi:10.1093/nar/gky1055
772 54. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: Tool for The Unification of Biology. *Nat*
773 *Genet*. 2000;25(1):25-29. doi:10.1038/75556
774 55. Lambert SA, Jolma A, Campitelli LF, et al. The Human Transcription Factors. *Cell*. 2018;172(4):650-

775 665. doi:10.1016/j.cell.2018.01.029
776 56. Kapopoulou A, Mathew L, Wong A, Trono D, Jensen JD. The evolution of gene expression and
777 binding specificity of the largest transcription factor family in primates. *Evolution (N Y)*.
778 2016;70(1):167-180. doi:10.1111/evo.12819
779 57. Jimenez-Sanchez G, Childs B, Valle D. Human Disease Genes. *Nature*. 2001;409:853-855.
780 58. Worman HJ, Courvalin JC. The nuclear lamina and inherited disease. *Trends Cell Biol*.
781 2002;12(12):591-598. doi:10.1016/S0962-8924(02)02401-7
782 59. Cleaver JE. It was a very good year for DNA repair. *Cell*. 1994;76(1):1-4. doi:10.1016/0092-
783 8674(94)90165-1
784 60. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from
785 variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7
786 61. Colacurcio DJ, Nixon RA. Disorders of lysosomal acidification—The emerging role of v-ATPase in
787 aging and neurodegenerative disease. *Ageing Res Rev*. 2016;32:75-88.
788 doi:10.1016/j.arr.2016.05.004
789 62. Kanfi Y, Peshti V, Gil R, et al. SIRT6 protects against pathological damage caused by diet-induced
790 obesity. *Aging Cell*. 2010;9(2):162-173. doi:10.1111/j.1474-9726.2009.00544.x
791 63. Blasco MA. Telomere length, stem cells and aging. *Nat Chem Biol*. 2007;3(10):640-649.
792 doi:10.1038/nchembio.2007.38
793 64. Bhattarai KR, Chaudhary M, Kim H-R, Chae H-J. Endoplasmic Reticulum (ER) Stress Response
794 Failure in Diseases. *Trends Cell Biol*. 2020;xx(xx):5-7. doi:10.1016/j.tcb.2020.05.004
795 65. De Leeuw CA, Neale BM, Heskes T, Posthuma D. The statistical properties of gene-set analysis.
796 *Nat Rev Genet*. 2016;17(6):353-364. doi:10.1038/nrg.2016.29
797 66. Loh PR, Bhatia G, Gusev A, et al. Contrasting genetic architectures of schizophrenia and other
798 complex diseases using fast variance-components analysis. *Nat Genet*. 2015;47(12):1385-1392.
799 doi:10.1038/ng.3431
800 67. Billingsley KJ, Barbosa IA, Bandrés-Ciga S, et al. Mitochondria function associated genes
801 contribute to Parkinson's Disease risk and later age at onset. *npj Park Dis*. 2019;5(1).
802 doi:10.1038/s41531-019-0080-x
803 68. Kraja AT, Liu C, Fetterman JL, et al. Associations of Mitochondrial and Nuclear Mitochondrial
804 Variants and Genes with Seven Metabolic Traits. *Am J Hum Genet*. 2019;104(1):112-138.
805 doi:10.1016/j.ajhg.2018.12.001
806 69. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and
807 functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 2019;51(3):404-413.
808 doi:10.1038/s41588-018-0311-9
809 70. Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in
810 mutation-intolerant genes and in regions under strong background selection. *Nat Genet*.
811 2018;50(3):381-389. doi:10.1038/s41588-018-0059-2
812 71. Segrè A V, Consortium D, investigators M, et al. Common Inherited Variation in Mitochondrial
813 Genes Is Not Enriched for Associations with Type 2 Diabetes or Related Glycemic Traits. *PLOS*
814 *Genet*. 2010;6(8):e1001058. https://doi.org/10.1371/journal.pgen.1001058.
815 72. Saxena R, De Bakker PIW, Singer K, et al. Comprehensive association testing of common
816 mitochondrial DNA variation in metabolic disease. *Am J Hum Genet*. 2006;79(1):54-61.
817 doi:10.1086/504926
818 73. Hudson G, Gomez-Duran A, Wilson IJ, Chinnery PF. Recent Mitochondrial DNA Mutations
819 Increase the Risk of Developing Common Late-Onset Human Diseases. *PLoS Genet*. 2014;10(5).
820 doi:10.1371/journal.pgen.1004369
821 74. Hudson G, Panoutsopoulou K, Wilson I, et al. No evidence of an association between
822 mitochondrial DNA variants and osteoarthritis in 7393 cases and 5122 controls. *Ann Rheum Dis*.

823 2013;72(1):136-139. doi:10.1136/annrheumdis-2012-201932
824 75. Raule N, Sevini F, Santoro A, Altilia S, Franceschi C. Association studies on human mitochondrial
825 DNA: Methodological aspects and results in the most common age-related diseases.
826 *Mitochondrion*. 2007;7(1-2):29-38. doi:10.1016/j.mito.2006.11.013
827 76. Yu X, Koczan D, Sulonen AM, et al. mtDNA nt13708A variant increases the risk of multiple
828 sclerosis. *PLoS One*. 2008;3(2):1-7. doi:10.1371/journal.pone.0001530
829 77. Hudson G, Nalls M, Evans JR, et al. Two-stage association study and meta-analysis of
830 mitochondrial DNA variants in Parkinson disease. *Neurology*. 2013;80(22):2042-2048.
831 doi:10.1212/WNL.0b013e318294b434
832 78. Samuels DC, Carothers AD, Horton R, Chinnery PF. The power to detect disease associations with
833 mitochondrial DNA haplogroups. *Am J Hum Genet*. 2006;78(4):713-720. doi:10.1086/502682
834 79. Biffi A, Anderson CD, Nalls MA, et al. Principal-Component Analysis for Assessment of Population
835 Stratification in Mitochondrial Medical Genetics. *Am J Hum Genet*. 2010;86(6):904-917.
836 doi:10.1016/j.ajhg.2010.05.005
837 80. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated
838 variation in regulatory DNA. *Science (80-)*. 2012;337(6099):1190-1195.
839 doi:10.1126/science.1222794
840 81. Seidman JG, Seidman C. Transcription factor haploinsufficiency: When half a loaf is not enough. *J*
841 *Clin Invest*. 2002;109(4):451-455. doi:10.1172/JCI0215043
842 82. Lee R Van Der, Correard S, Wasserman WW. Deregulated Regulators : Disease-Causing cis
843 Variants in Transcription Factor Genes. *Trends Genet*. 2020;xx(xx). doi:10.1016/j.tig.2020.04.006
844 83. Arroyo JD, Jourdain AA, Calvo SE, et al. A Genome-wide CRISPR Death Screen Identifies Genes
845 Essential for Oxidative Phosphorylation. *Cell Metab*. 2016;24(6):875-885.
846 doi:10.1016/j.cmet.2016.08.017
847 84. Wright S. Physiological and Evolutionary Theories of Dominance. *Am Nat*. 1934;68(714):24.
848 85. Kacser H, Burns JA. The molecular basis of dominance. *Genetics*. 1981;97(3-4):639-666.
849 86. Chance B, Williams GR. RESPIRATORY ENZYMES IN OXIDATIVE PHOSPHORYLATION: III. THE
850 STEADY STATE. *J Biol Chem*. 1955;217(1):409-428. http://www.jbc.org/content/217/1/409.short.
851 87. Vafai SB, Mootha VK. Mitochondrial disorders as windows into an ancient organelle. *Nature*.
852 2012;491(7424):374-383. doi:10.1038/nature11707
853 88. Balaban RS, Kantor HL, Katz LA, Briggs RW. Relation between work and phosphate metabolite in
854 the in vivo paced mammalian heart. *Science (80-)*. 1986;232(4754):1121-1123.
855 doi:10.1126/science.3704638
856 89. To TL, Cuadros AM, Shah H, et al. A Compendium of Genetic Modifiers of Mitochondrial
857 Dysfunction Reveals Intra-organelle Buffering. *Cell*. 2019;179(5):1222-1238.e17.
858 doi:10.1016/j.cell.2019.10.032
859 90. Golan D, Lander ES, Rosset S. Measuring missing heritability: Inferring the contribution of
860 common variants. *Proc Natl Acad Sci U S A*. 2014;111(49):E5272-E5281.
861 doi:10.1073/pnas.1419064111
862 91. Polderman TJC, Benyamin B, De Leeuw CA, et al. Meta-analysis of the heritability of human traits
863 based on fifty years of twin studies. *Nat Genet*. 2015;47(7):702-709. doi:10.1038/ng.3285
864 92. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes.
865 *Nature*. 2016;536(7614):41-47. doi:10.1038/nature18642
866 93. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for
867 complex traits. *PLoS Genet*. 2008;4(2). doi:10.1371/journal.pgen.1000008
868 94. Zhu Z, Bakshi A, Vinkhuyzen AAE, et al. Dominance genetic variation contributes little to the
869 missing heritability for human complex traits. *Am J Hum Genet*. 2015;96(3):377-385.
870 doi:10.1016/j.ajhg.2015.01.001

871 95. Rand DM, Mossman JA. Mitonuclear conflict and cooperation govern the integration of
872 genotypes, phenotypes and environments. *Philos Trans R Soc B Biol Sci*. 2020;375(1790).
873 doi:10.1098/rstb.2019.0188
874 96. Sackton TB, Hartl DL. Genotypic Context and Epistasis in Individuals and Populations. *Cell*.
875 2016;166(2):279-287. doi:10.1016/j.cell.2016.06.047
876 97. Hemani G, Shakhbazov K, Westra HJ, et al. Detection and replication of epistasis influencing
877 transcription in humans. *Nature*. 2014;508(7495):249-253. doi:10.1038/nature13005
878 98. Solenski NJ, DiPierro CG, Trimmer PA, Kwan AL, Helms GA. Ultrastructural changes of neuronal
879 mitochondria after transient and permanent cerebral ischemia. *Stroke*. 2002;33(3):816-824.
880 doi:10.1161/hs0302.104541
881 99. Flameng W, Andres J, Ferdinande P, Mattheussen M, Van Belle H. Mitochondrial function in
882 myocardial stunning. *J Mol Cell Cardiol*. 1991;23(1):1-11. doi:10.1016/0022-2828(91)90034-J
883 100. Weinbrenner C, Liu GS, Downey JM, Cohen M V. Cyclosporine a limits myocardial infarct size
884 even when administered after onset of ischemia. *Cardiovasc Res*. 1998;38(3):676-684.
885 doi:10.1016/S0008-6363(98)00064-9
886 101. Kubben N, Misteli T. Shared molecular and cellular mechanisms of premature ageing and ageing-
887 associated diseases. *Nat Rev Mol Cell Biol*. 2017;18(10):595-609. doi:10.1038/nrm.2017.68
888 102. Garcia CK, Wright WE, Shay JW. Human diseases of telomerase dysfunction: Insights into tissue
889 aging. *Nucleic Acids Res*. 2007;35(22):7406-7416. doi:10.1093/nar/gkm644
890 103. Han S, Brunet A. Histone methylation makes its mark on longevity. *Trends Cell Biol*.
891 2012;22(1):42-49. doi:10.1016/j.tcb.2011.11.001
892 104. Bahar R, Hartmann CH, Rodriguez KA, et al. Increased cell-to-cell variation in gene expression in
893 ageing mouse heart. *Nature*. 2006;441(7096):1011-1014. doi:10.1038/nature04844
894 105. Kanfi Y, Naiman S, Amir G, et al. The sirtuin SIRT6 regulates lifespan in male mice. *Nature*.
895 2012;483(7388):218-221. doi:10.1038/nature10815
896 106. Karczewski KJ, Dudley JT, Kukurba KR, et al. Systematic functional regulatory assessment of
897 disease-associated variants. *Proc Natl Acad Sci U S A*. 2013;110(23):9607-9612.
898 doi:10.1073/pnas.1219099110
899 107. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and
900 population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795
901 108. Yanai I, Benjamin H, Shmoish M, et al. Genome-wide midrange transcription profiles reveal
902 expression level relationships in human tissue specification. *Bioinformatics*. 2005;21(5):650-659.
903 doi:10.1093/bioinformatics/bti042
904 109. Melé M, Ferreira PG, Reverter F, et al. The human transcriptome across tissues and inidividuals.
905 *Science (80-)*. 2015;348(6235):660-665. http://dx.doi.org/10.1126/science.aaa0355.
906 110. Morris JA, Randall JC, Maller JB, Barrett JC. Evoker: A visualization tool for genotype intensity
907 data. *Bioinformatics*. 2010;26(14):1786-1787. doi:10.1093/bioinformatics/btq280
908 111. Yates AD, Achuthan P, Akanni W, et al. Ensembl 2020. *Nucleic Acids Res*. 2020;48(D1):D682-D688.
909 doi:10.1093/nar/gkz966
910 112. Litman T, Stein WD. Obtaining estimates for the ages of all the protein-coding genes and most of
911 the ontology-identified noncoding genes of the human genome, assigned to 19 phylostrata.
912 *Semin Oncol*. 2019;46(1):3-9. doi:10.1053/j.seminoncol.2018.11.002
913