Dysregulation of alternative splicing patterns in the ovaries of reproductively aged mice

Nehemiah S. Alvarez † [1,2], Pavla Brachova † [1]

[1] Department of Biomedical and Translational Sciences, Division of Woman and Child Health, Eastern Virginia Medical School, Macon and Joan Brock Virginia Health Sciences, Old Dominion University, Norfolk, VA, USA

[2] Advanced Sequencing Program, Eastern Virginia Medical School, Macon and Joan Brock Virginia Health Sciences, Old Dominion University, Norfolk, VA, USA

† These authors contributed equally to this work

Nehemiah Seth Alvarez, Ph.D.

Assistant Professor

Director, Advanced Sequencing Program

Department of Biomedical and Translational Sciences

Division of Woman and Child Health

Eastern Virginia Medical School

Macon and Joan Brock Virginia Health Sciences

Old Dominion University

Lewis Hall

700 W Olney Road

Norfolk VA 23507

Pavla Brachova, Ph.D.

Assistant Professor

Department of Biomedical and Translational Sciences

Division of Woman and Child Health

Eastern Virginia Medical School

Macon and Joan Brock Virginia Health Sciences

Old Dominion University

Lewis Hall

700 W Olney Road

Norfolk VA 23507

Author e-mail addresses: NSA: AlvareNS@odu.edu; PB: BrachoP@odu.edu

Running title: Splicing disruption in the aging mouse ovary

Keywords: splicing; direct RNA-sequencing; post-transcriptional regulation; ovary

Word count of the full article, excluding references and figure legends: 12,791

In brief

Reproductive aging is associated with changes in alternative splicing patterns in the mouse ovary. Our study identifies novel transcripts resulting from alterations in exon usage that may have altered protein function and important roles in ovarian physiology as well as in reproductive aging.

Abstract

Female reproductive aging is marked by a decline in oocyte quality, yet the underlying molecular mechanisms remain incompletely understood. Here, we used long-read direct RNA-sequencing to map transcript isoform changes in mouse ovaries across reproductive age. Comparing young and aged mice under controlled gonadotropin stimulation, we identified widespread alternative splicing changes, including shifts in exon usage, splice site selection, and transcript boundaries. Aged ovaries exhibited increased isoform diversity, favoring distal start and end sites, and a significant rise in exon skipping and intron retention events. Many of these age-biased splicing events altered open reading frames, introduced premature stop codons, or disrupted conserved protein domains. Notably, mitochondrial genes were disproportionately affected. We highlight *Ndufs4*, a Complex I subunit, as a case in which aging promotes a truncated isoform lacking the canonical Pfam domain. Structural modeling suggests this splice variant could impair Complex I assembly, providing a mechanistic link between splicing and mitochondrial dysfunction in the aging ovary. These findings support the existence of a splicing-energy-aging axis in ovarian physiology, wherein declining mitochondrial function and adaptive or maladaptive splicing changes are intertwined. Our study reveals that alternative splicing is not merely a byproduct of aging but a dynamic, transcriptome-wide regulatory layer that may influence oocyte competence and ovarian longevity. These insights open new avenues for investigating post-transcriptional mechanisms in reproductive aging and underscore the need to consider isoform-level regulation in models of ovarian decline.

Introduction

Metabolic Challenges in the Aging Ovary

Female reproductive aging is one of the earliest manifestations of organismal aging, marked by a steady decline in fertility beginning in the mid-30s and culminating in menopause by the early 50s. This decline is driven not only by the depletion of ovarian follicles but also by a pronounced reduction in oocyte quality with advanced maternal age (Bao *et al.*, 2024). The molecular mechanisms underlying age-related oocyte dysfunction remain incompletely understood, but accumulating evidence implicates mitochondrial dysfunction and energetic insufficiency in both oocytes and their surrounding somatic cells as key contributors. Mitochondria are critical for oocyte development and survival, providing the ATP needed for processes like meiotic spindle assembly and chromosome segregation. Consistent with this, aged oocytes often exhibit impaired ATP production and elevated oxidative stress, leading to spindle instability and increased chromosomal segregation errors during meiosis (Ju *et al.*, 2024). Mitochondrial dysfunction has indeed been identified as a hallmark of ovarian aging: a decline in mitochondrial DNA copy number and respiratory capacity has been correlated with oocyte quality loss and accelerated ovarian cellular senescence (Ju *et al.*, 2024). In the follicular niche, the aging microenvironment further exacerbates these issues, senescent granulosa cells in aged ovaries show limited mitochondrial energy output and undergo increased apoptosis, creating an inflammatory, suboptimal milieu for the oocyte (Ju *et al.*, 2024). Notably, a recent study of human oocytes from advanced-age women confirmed significant mitochondrial defects and an apparent shift toward alternative energy substrates in these cells (Smits *et al.*, 2023). Together, these findings highlight the heavy energetic demands on the ovary and suggest that reproductive aging imposes metabolic stress on ovarian cells. This raises the question of how ovarian cells might adapt at the molecular level to maintain energy homeostasis as mitochondrial function declines with age.

The Splicing–Energy–Aging Axis

Emerging evidence from other systems suggests that one adaptive strategy cells use to cope with metabolic stress is the dynamic regulation of mRNA splicing (Ferrucci *et al.*, 2022). In fact, it has been hypothesized that there exists an “energy–splicing–resilience axis” of aging, wherein changes in mitochondrial energy status actively drive changes in pre-mRNA splicing, and this splicing plasticity in turn helps cells preserve energetic

homeostasis during aging (Ferrucci *et al.*, 2022). Ferrucci and colleagues articulated this concept, proposing that when cellular energy supplies are ample, the splicing program remains in a basal state, but under conditions of mitochondrial impairment and ATP shortfall, cells broaden their repertoire of transcript isoforms as a resilience mechanism (Ferrucci *et al.*, 2022). In this model, alternative splicing is upregulated to produce isoforms that can increase energy production or reduce energy expenditure for nonessential processes, partly through signaling cascades that engage energy sensors like AMP-activated protein kinase (AMPK) (Ferrucci *et al.*, 2022).

Empirical studies across tissues lend support to this splicing–energy–aging axis. In neuronal cells, for example, inducing mitochondrial damage (and consequent ATP depletion) triggers widespread shifts in splicing patterns, universally biasing transcripts toward shorter isoforms, while constitutively spliced exons remain largely unchanged (Maracchioni *et al.*, 2007). Importantly, this splicing modulation in neurons has been shown to correlate specifically with ATP loss rather than with oxidative damage per se (Maracchioni *et al.*, 2007), underscoring that the splicing machinery is directly sensitive to cellular energy levels. In skeletal muscle, age-related splicing changes have been tied to the energetic capacity of individuals: older adults with low fitness (indicative of reduced muscle oxidative capacity) exhibit an upregulation of splicing activity and greater alternative isoform diversity compared to high-fitness older adults (Donega *et al.*, 2025). The splicing variants enriched in low-fitness muscle were associated with pathways like RNA–protein granule assembly, hinting at a compensatory remodeling of gene expression under energy-poor conditions (Donega *et al.*, 2025). The immune system shows a similar linkage between metabolism and splicing regulation. During T-lymphocyte activation, when cellular energy demand sharply rises, there is increased expression and stability of certain splicing regulators (e.g. CELF2) that in turn drive broad alternative splicing changes supporting T cell survival and function (Choi *et al.*, 2022). At a gene-specific level, a classic illustration of metabolic-splicing crosstalk is provided by the PKM gene: alternative splicing of PKM yields the PKM1 and PKM2 isoforms, which dictate whether glucose is primarily metabolized through oxidative phosphorylation or diverted to aerobic glycolysis, respectively (Biamonti *et al.*, 2018). This well-known switch in pyruvate kinase isoforms demonstrates how splicing decisions can rewire core metabolic pathways to meet cellular energy needs. Furthermore, direct manipulation of energy-sensing pathways can influence splicing outcomes. Activation of AMPK, central energy gauge of a cell, has been shown to modulate splicing factor activity; for instance, in a progeroid aging model, pharmacological AMPK activation (using metformin) inhibited the splicing factor SRSF1 and thereby reduced the production of a toxic alternative splicing product of LMNA (the progerin isoform), alleviating some aging-associated defects (Finley, 2018). Collectively, these examples illustrate that alternative splicing networks respond to energetic and mitochondrial stress across diverse cell types. Rather than being a passive consequence of aging, splicing changes often appear to act as a compensatory mechanism; a means to adjust metabolic gene expression, optimize energy utilization, and maintain homeostasis when cells are under energetic duress (Ferrucci *et al.*, 2022).

Given this backdrop, it is plausible that the ovary might also engage a splicing-mediated adaptive program as it ages, in order to cope with mounting mitochondrial dysfunction. However, the contribution of alternative splicing to the ovarian aging process remains unexplored. Recent transcriptome studies of ovarian aging (including single-cell RNA-sequencing efforts) have catalogued changes in gene expression with age, but these analyses have been limited to gene-level measurements and lack information on full-length isoforms due to the short-read sequencing technologies used. As a result, potential age-related shifts in specific transcript isoforms, such as those involving subtle exon usage changes, alternative splicing junctions, or 5′/3′ UTR length remodeling, could not be resolved. To overcome this limitation and directly address the role of alternative splicing in reproductive aging, we employed a direct RNA-sequencing approach using the Oxford Nanopore platform. This long-read, single-molecule sequencing technology allows us to read intact native mRNA molecules without reverse transcription or fragmentation, thereby capturing complete transcript isoforms with base-level resolution. In the present study, we applied direct RNA-sequencing to ovaries from young and aged mice to comprehensively characterize age-dependent changes in the ovarian transcriptome at the isoform level. This approach enabled us to identify differential exon inclusion events, shifts in splice site usage, and alternative first/last exon selections (“transcript boundary” remodeling) that emerge with advanced reproductive age. We hypothesized that reproductive aging would be accompanied by specific alternative splicing alterations, potentially as an adaptive response to declining mitochondrial function, and our objective was to map these isoform changes to gain insight into how the ovarian gene expression program is modulated during

aging. By delineating the splicing landscape of young vs. aged ovaries in unprecedented detail, our study sets the stage for understanding whether an energy-linked splicing axis is active in the ovary and how such post-transcriptional changes might influence ovarian function and the decline of fertility with age.

Materials and Methods

Ovary collection and RNA isolation

Young (6-8 week-old) C57BL/6J female mice were purchased from Jackson Laboratory, while reproductively aged (14 month-old) C57BL/6 mice were from National Institutes of Aging. 14 month-old mice were selected as the reproductively aged cohort because significant reduction in egg quality and changes in the ovarian microenvironment occur at that age (Chiang *et al.*, 2010; Briley *et al.*, 2016). Groups of young (3 mice) and aged mice (3 mice) were given an I.P. injection of 5 international units pregnant mare's serum gonadotropin to synchronously stimulate follicle growth (PMSG, Catalog # 493-10, Lee Biosolutions) 46 hours prior to harvest of whole ovaries. Groups of young (3 mice) and aged mice (3 mice) were given an I.P. injection 5 international units PMSG and then with 5 international units human chorionic gonadotropin (hCG) 46 hours later; ovaries were collected 16 hours later as previously described (Brachova *et al.*, 2019). Ovaries were cleaned from bursal sacs and placed in TRIzol (Invitrogen) for RNA isolation according to the manufacturer's protocol. Mice were maintained in environmentally controlled facilities at the Eastern Virginia Medical School (EVMS) animal facility in a room with a 12-hour light, 12-hour dark cycle (7 am to 7 pm) with ad libitum access to food and water. All animal procedures were performed according to an approved Eastern Virginia Medical School's Institutional Animal Care and Use Committee protocol.

Direct RNA-sequencing of young and aged ovaries

Direct RNA-sequencing libraries were prepared for one ovary from each mouse for each age and condition using a Oxford Nanopore Direct RNA-sequencing kit (SQK-RNA002, ONT) the manufacturer's protocol. Briefly, polyadenylated (poly(A)) RNA was purified from total RNA using a Magnetic mRNA Isolation Kit (NEB, S1550S). 100 ng of poly(A) RNA was annealed and ligated to an oligo-d(T) sequencing adaptor that anchors to the 3′ end of RNA T4 DNA Ligase (NEB, M0202). Adaptor ligated libraries were reverse transcribed using Superscript IV (ThermoFisher, 18090010) and purified with RNAClean XP beads (Beckman, A63987). Twelve libraries were loaded onto individual flow cells (FLO-MIN106D, ONT) and sequenced on a MinION Mk1B (ONT).

Alignment and isoform quantification

Raw direct-RNA FASTQ files were concatenated to form a single input dataset for uniform processing. The *Mus musculus* reference genome (GRCm38, build 68) and its accompanying Ensembl gene annotation (release 102) were obtained to provide a consistent coordinate framework for read mapping and splice-site correction (Yates *et al.*, 2020). Long reads were processed with FLAIR (Full-Length Alternative Isoform Analysis of RNA) (Tang *et al.*, 2020). FLAIR's workflow, comprising *align*, *correct*, *collapse*, and *quantify* stages, first aligns nanopore reads to the genome, refines splice junctions against the annotation, collapses redundant transcripts into non-redundant isoforms, and finally derives transcript-per-million (TPM)–normalized abundance estimates. Stringent splice-checking and annotation-reliant settings were used to maximize isoform confidence, and all operations were parallelized on 120 CPU threads to accommodate the large read volume. The resulting isoform GTF, FASTA, BED, and read-to-isoform mapping files constituted a sample-wide reference transcriptome suitable for downstream analysis.

In-silico Translation of Isoforms

The collapsed cDNA isoforms were translated *in silico* with bioseq2seq, a transformer-based sequence-to-sequence model trained on mammalian coding sequences (Valencia and Hendrix, 2023). Translations were generated with beam-size 1 to select the single most likely reading frame per transcript; maximum token and decode-length limits prevented over-extension beyond biologically plausible CDS lengths. Predicted peptide sequences were parsed into FASTA format for compatibility with standard annotation tools.

Protein Domain Annotation

Translated isoforms were scanned against the Protein family (Pfam)-A profile HMM collection (release 35.0) using *hmmsearch* to assign conserved domain annotations and to facilitate functional interpretation of novel splice variants (Mistry *et al.*, 2021).

Preprocessing and Isoform Proportions

To identify genome-wide splicing changes between young and aged ovaries, we implemented a Bayesian hierarchical model to compare isoform usage proportions across conditions. We first grouped raw transcript counts by their promoter of origin and filtered out any promoter group lacking expression in one of the conditions (each group was required to have at least one non-zero count in both young and aged samples). Within each remaining promoter group, transcript counts for each sample were converted to proportions by dividing by the group total count, yielding the fraction of transcripts originating from each isoform. These isoform proportions (analogous to percent-spliced-in values) were bounded away from 0 and 1 (rescaled to lie between 0.001 and 0.999) to avoid numerical issues. This preprocessing ensures that differential splicing is assessed on a relative usage scale, effectively controlling for overall gene expression differences. Notably, using relative proportions addresses the compositional nature of isoform data—changes in an isoform fraction reflect true shifts in splicing, independent of gene-level expression changes.

Classical RNA-seq analysis methods (e.g., edgeR or DESeq2) model absolute read counts with negative binomial distributions and are designed for detecting differential total expression per gene or transcript (Robinson *et al.*, 2010; Love *et al.*, 2014). These count-based tools are not optimized for compositional data like isoform ratios, as they assume independent features and do not account for the constraint that isoform proportions sum to one within a gene. In fact, applying gene-level methods directly to isoform counts can be problematic: it fails to capture the varying estimation uncertainty across isoforms and may lead to reduced power or inflated false positives (Leng *et al.*, 2013). Leng et al. (2013) specifically cautioned that using standard count-based techniques on isoform-level estimates is *"not recommended"* due to these issues, advocating for hierarchical modeling to properly share information between isoforms. Therefore, a Bayesian approach was chosen to naturally accommodate the relative nature of isoform usage data and to better model the inherent uncertainties in isoform proportion estimates.

Bayesian Model for Differential Isoform Usage

We applied a custom Bayesian Beta-binomial model to assess shifts in isoform ratios between conditions. In our hierarchical model, the proportion values for each isoform (across all replicates and conditions) are treated as random draws from an underlying Beta distribution with isoform-specific shape parameters (α and β). This choice is well-suited for data bounded between 0 and 1, and it flexibly captures different usage biases: for example, an isoform with $\alpha >> \beta$ has a propensity for high usage (fraction near 1), whereas $\beta >> \alpha$ would imply the isoform is usually lowly used (fraction near 0). We placed weakly informative Gamma priors on each isoform α and β (shape=5, scale=1), reflecting a mild prior belief that extreme proportions are unlikely without substantial evidence. These priors act as regularization, shrinking estimates toward modest usage levels in the absence of strong data support, which is particularly advantageous given our limited sample size. Such prior-informed smoothing is a key benefit of the Bayesian framework – it helps stabilize inference with small *n* by "borrowing strength" across isoforms and samples (Leng *et al.*, 2013). Indeed, empirical Bayes and fully Bayesian methods have demonstrated improved performance in RNA-seq analyses by sharing information in this way (Leng *et al.*, 2013; Tiberi and Robinson, 2020).

Posterior sampling was performed using Markov Chain Monte Carlo and yields a posterior distribution for each isoform usage parameters and, by extension, for each isoform difference in usage between conditions. From the posterior samples, we directly computed the $\log_2$ fold-change in isoform proportion (aged vs. young) along with its 95% highest posterior density (HPD) interval. Importantly, the Bayesian approach provides a probabilistic assessment of differential splicing: for each isoform, we calculated the posterior probability that its usage is higher in aged ovaries (and conversely in young ovaries). This probability metric offers an intuitive measure of confidence in the direction of splicing change, as opposed to relying only on p-values or FDR thresholds. We designated an isoform as significantly differentially spliced if two criteria were met: (1) the 95% HPD interval of its $\log_2$ fold-change in proportion did not include zero (i.e. no overlap with "no change"), and (2) the posterior probability of a usage increase in one condition exceeded 0.95. In other words, we require high-

confidence that the isoform usage shifted substantially between young and aged samples. This Bayesian decision criterion is analogous to using a stringent significance cutoff, but it has the advantage of incorporating the full posterior uncertainty of the estimate. Methods like MISO and MATS have similarly leveraged Bayesian posterior distributions to call differential isoform usage with high confidence (Katz *et al.*, 2010; Shen *et al.*, 2012), and more recent Bayesian frameworks have shown that such probability-based thresholds control false discoveries well while improving detection power in practice (Tiberi and Robinson, 2020).

Advantages of a Bayesian modeling approach: In summary, adopting a Bayesian hierarchical model for isoform usage offers several key advantages over classical approaches. First, it models proportions directly (using Beta/Beta-binomial distributions) and thus properly accounts for the compositional structure of transcript data, unlike methods built for independent counts. Second, the Bayesian framework accounts for uncertainty in a principled way: rather than outputting only point estimates and p-values, it yields posterior distributions and credible intervals for isoform usage changes, enabling statements like “there is a 98% posterior probability that isoform X usage is higher in aged ovaries.” This quantification of uncertainty is crucial when read mapping ambiguity or limited read depth can obscure isoform quantification (Katz *et al.*, 2010; Tiberi and Robinson, 2020). Third, Bayesian modeling (especially with informative priors and hierarchical sharing) often improves performance with small sample sizes. By incorporating prior knowledge and pooling information across genes, it mitigates the instability that plagues frequentist estimates in low-replicate scenarios (Shen *et al.*, 2012). Studies have shown that Bayesian or empirical Bayes methods can achieve greater sensitivity and specificity in detecting differential isoform expression, even with few replicates, compared to traditional count-based tests (Leng *et al.*, 2013; Tiberi and Robinson, 2020). These advantages motivated our choice of a Bayesian model for differential splicing analysis, as it provides a robust, uncertainty-aware inference of isoform regulation changes between young and aged ovarian samples.

Alternative Splicing Event Annotation Pipeline

Input Transcript Models: The starting point was a collapsed transcriptome GTF produced by FLAIR, containing multi-exon and single-exon transcript models for each gene. All transcripts were grouped by gene, and a custom Python module was used to compare transcript structures and integrate coverage data to identify various classes of alternative splicing events (implemented as custom code). The analysis proceeded gene-by-gene in a strand-aware manner, leveraging both direct GTF structural comparisons and an inferred expression-support signal derived from normalized exon coverage. The following event categories were annotated: alternative promoters, alternative 5′ splice sites, alternative 3′ splice sites, alternative exons (including cassette and mutually exclusive exons), retained introns, and alternative transcription end sites.

Alternative Promoters (TSS Clusters): To identify *alternative promoter* events, transcripts of each gene were first clustered by their transcription start site (TSS). Transcripts with TSS positions within a small genomic window (±1 bp) were considered to share the same promoter cluster. If a transcript TSS did not overlap any existing cluster, it initiated a new promoter group. This ensured that transcripts initiating from distinct genomic positions were assigned to separate promoter groups, representing distinct promoter usage. Each unique TSS cluster was recorded as an alternative promoter for the gene. For genes with multiple promoter clusters, a fractional distance metric was computed to indicate the relative position of each promoter along the gene span (0.0 for the most upstream TSS, up to 1.0 for the most downstream). All transcripts were annotated with their promoter group membership, and those originating from downstream clusters (when multiple clusters exist) were thus classified under alternative promoters.

Normalized Exon Coverage Profile: For each gene (and for each promoter-based subgroup of its transcripts), we generated a strand-specific normalized coverage profile across the gene exonic span to help delineate exon and intron features. This profile was derived by superimposing all exons from the transcripts in the group onto a single coordinate axis from the first exon start to the last exon end. A coverage value was calculated at each nucleotide position as the sum of 1/(number of transcripts spanning that exon) for each overlapping exon. In effect, bases that are exonic in all transcripts of the group received a normalized coverage of ~1.0, whereas bases present in only a subset of transcripts had fractional coverage (<1.0), and intronic regions (absent from all transcript exons) had coverage ~0. This normalized coverage approach provides an approximate “exon inclusion” frequency across the transcripts. We then applied a peak-finding algorithm to this profile to locate local peaks (high points corresponding to putative exon centers) and valleys (low points corresponding to

intronic regions). Peaks represented candidate exonic segments, with their height reflecting the fraction of transcripts including that segment, while valleys indicated boundaries between exons (intron centers with minimal coverage). Each valley position was expanded leftward and rightward to determine the span of the intron (valley region) between flanking exons. This provided approximate intron coordinates for further analysis. The normalized coverage profile, along with the identified peak and valley positions, was then used to support the detection of alternative splicing events as described below.

Alternative 5′ Splice Sites (Donor Site Variation): We defined *alternative 5′ splice site* events (alternate donor sites) as cases where an intron 5′ end occurs at multiple positions across transcripts, resulting in alternative starts of the downstream exon. Using the normalized coverage profile and identified intronic regions, we scanned for abrupt drops in coverage at exon boundaries that would indicate such donor site variation. On the positive strand, the algorithm iterated through exon regions in the 5′→3′ direction and flagged a position i as a candidate alternative 5′ site if the coverage at i was lower than at the preceding base $i-1$, signaling a step down in coverage within what should be an exon. Importantly, any candidate position overlapping an intronic region or lying beyond the last exon end was ignored. This logic detects a point where one transcript exon ends (entering the intron), while another transcript exon continues further, a hallmark of an alternative donor splice site. For negative-strand genes, the scan was inverted (iterating 3′→5′) to detect drops in coverage when moving in the opposite orientation (since a 5′ splice site is at the genomic downstream end of an intron on the minus strand). By comparing coverage differences across exon–intron junctions in a strand-aware fashion, we pinpointed bases that serve as alternate donor splice junctions supported by the transcript models and coverage dips. We additionally ensured that such events were not misattributed to transcript termination differences by excluding any donor-site candidate that coincided with an alternative transcript end position (see below). Each confirmed position was annotated as an alternative 5′ splice site, and any exon ending at that genomic coordinate was labeled accordingly in the output GTF (feature type exon_alt5).

Alternative 3′ Splice Sites (Acceptor Site Variation): Similarly, *alternative 3′ splice site* events (alternate acceptor sites) were identified as positions where an intron 3′ end (the acceptor junction for the upstream exon) varies between transcripts. For plus-strand transcripts, we examined each coverage peak (exon) and scanned backwards (toward the 5′ direction) from the exon center toward its start, looking for a position where coverage increases from one base to the next (when moving right-to-left). A position j was flagged as an alternative acceptor if coverage at j was lower than at j+1, indicating the point at which an incoming intron coverage rises into the exon – i.e. a transcript that has a longer intron (skipping some exon portion) versus another that starts the exon at a downstream acceptor j+1. The algorithm skipped over any positions lying inside intronic regions (ensuring the candidate is at an exon boundary) and stopped once it reached another peak or the transcript start. For minus-strand genes, the procedure was symmetric but in the forward direction (scanning from exon center toward the 3′ end) due to reversed transcription orientation. This effectively finds alternate acceptor junctions where one transcript includes extra sequence at an exon 5′ end that another transcript splices out. To avoid conflating alternative 3′ splice sites with alternative TSS events, any candidate acceptor site that coincided with a known alternative promoter (TSS cluster) position was not counted as an alternative 3′ splice site. Verified alternative acceptor positions were annotated in the GTF, and exons beginning at those coordinates were labeled exon_alt3 in the feature field.

Alternative Exons (Cassette and Mutually Exclusive Exons): We annotated *alternative exons* as internal exonic segments that are included in some transcript isoforms but skipped (omitted) in others. This category covers simple cassette exons (one exon optional vs. skipped) as well as mutually exclusive exons (where two alternative exons occupy the same relative position, each present in complementary subsets of transcripts). Detection of alternative exons relied on the normalized coverage profile peak analysis. Specifically, any internal exon that did not appear in all transcripts of the group exhibited a reduced normalized coverage peak (<1.0). We extracted all coverage peak positions and identified those with height significantly below 1.0 (indicating partial exon inclusion). To focus on well-defined exon-skipping events, we further required that such a sub-maximum peak be flanked by pronounced valleys on both sides (i.e. it is an isolated internal exon bounded by introns). In practice, for each peak with fractional coverage, we found the nearest coverage minima (valleys) before and after the peak and retrieved the exact exon coordinates falling between those valley positions. We excluded the first and last exons from this analysis (to distinguish alternative internal exons from alternative promoters or ends). Any exon that met these criteria – an intermediate exon used by only a subset of transcripts – was labeled as an alternative exon. In the updated GTF, such exons were

marked with feature type exon_ce (denoting cassette/alternative exon). Notably, mutually exclusive exon pairs would both appear as alternative exons by this method: each alternative exon of the pair shows incomplete inclusion across transcripts (each is missing in the transcripts that use the alternate exon), yielding sub-maximal coverage peaks for both. Thus, the pipeline can capture mutually exclusive exons by labeling each involved exon as exon_ce (if they satisfy the coverage and structural criteria). Each transcript that contains an alternative exon was flagged in its attributes (see Output, below), and transcripts skipping the exon simply lack that exon_ce segment.

Retained Introns: Retained intron events were detected by identifying intronic regions that show evidence of being transcribed (not spliced out) in at least one isoform. Two complementary approaches were used. Structural comparison: for each intron (interval between exons) defined by one transcript, the pipeline checked if any other transcript in the gene has a continuous exon that spans across that intron coordinates. If so, this suggests that the region is not spliced out in that transcript, i.e. an intron retention event. Coverage analysis: we leveraged the normalized coverage valleys corresponding to introns. After defining all intron intervals from the coverage minima (valley spans) as described above, we calculated the average normalized coverage within each intron interval. Introns that had non-zero (and especially significant fractional) coverage indicated that some transcripts contribute exonic reads in those normally intronic positions. In other words, if the coverage profile does not drop to zero in an intron, it implies one or more transcripts include that sequence. Such intronic regions were flagged as retained intron candidates. In the output GTF, any exon overlapping a flagged intron region was relabeled with exon_ir to denote that it represents a retained intron sequence. This labeling was achieved by marking all base positions within each identified intron interval and intersecting these with exon coordinates. As a result, an exon that fully or partially covers a normally intronic span is annotated as exon_ir. Each transcript containing one or more retained introns was also annotated in its attributes (incrementing an intron retention count; see below). Conventional introns (with zero coverage and not spanned by any other transcript exon) were not labeled as retained.

Alternative Transcription End Sites (Alternative TES): We identified *alternative transcription end site* events by clustering transcript 3′ end positions (transcription end sites, TES) on a per-gene basis (accounting for strand orientation). For each transcript, the TES was taken as the coordinate of the transcript last exon end if on the + strand, or the first exon start if on the – strand. Transcripts with end sites within a small window (±1 nt) were grouped, meaning their 3′ ends were considered the same cluster. Distinct clusters thus represent different polyadenylation sites or termination positions used by different isoforms. If a transcript end did not fall in the vicinity of any other, it initiated a new cluster. This method effectively merges transcripts ending at essentially the same nucleotide into one group, separating them from transcripts that end elsewhere. Each unique end cluster was recorded as an alternative end (alternative poly(A) site) for the gene. To quantify relative positions of multiple end clusters, a fractional distance along the gene was computed: for plus-strand genes, end sites closer to the gene 3′ end were assigned higher fractional values (approaching 1.0), whereas for minus-strand genes the calculation was inverted (since a smaller genomic coordinate actually lies downstream in transcription). This provided a normalized measure of how far downstream each alternative TES was. Alternative end events were annotated by marking the corresponding transcript end exons in the GTF as exon_alt_end. In practice, the last exon of transcripts using a downstream poly(A) site was labeled exon_alt_end (and similarly for the first exon of a minus-strand transcript using an alternative downstream end). Transcripts were also annotated with their respective end cluster (and fractional position) in the output attributes.

Output GTF Annotation and Transcript Attributes: The pipeline outputs an updated GTF file in which exons are relabeled according to the alternative splicing features identified. Each exon feature field (column 3 of the GTF) is replaced with a composite label indicating any alternative splicing event that overlaps that exon coordinates. For example, an exon that was determined to use an alternative 5′ donor site would be labeled exon_alt5, and one with both an alternative acceptor and being an underutilized cassette exon would be labeled exon_alt3_ce (the labels are concatenated with an underscore, all prefixed by “exon”). The possible exon labels include: exon_alt_pro (exon at an alternative promoter TSS, i.e. first exon of a downstream promoter group), exon_alt5 (exon with an alternate 5′ splice site), exon_alt3 (exon with an alternate 3′ splice site), exon_ce (cassette or alt exon), exon_ir (exon representing a retained intron sequence), and exon_alt_end (exon at an alternative transcript end site). These labels were assigned based on the positional overlap of the exon with the features

detected (the set of labels for each base position of the exon was aggregated). In addition to relabeled exons, the method appends informative attributes to each transcript entry in the GTF. Each transcript is annotated with a promoter group identifier (indicating which alternative promoter cluster it belongs to) and with numeric attributes reflecting that transcript involvement in various splicing events. For example, the number of alternative donor sites (a5) utilized by the transcript, the number of alternative acceptors (a3), the count of alternative exons it includes (ce), and any retained introns (i), along with the total possible events of each type identified for the gene or promoter group (a5t, a3t, cet, it, etc.), were recorded in the attributes. Transcripts arising from multiple promoter groups have a fractional promoter usage value (ap) and those with multiple end site options have a fractional TES value (ae), corresponding to the relative positions described above. These transcript-level attributes provide a summary of each isoform splicing pattern and its grouping by promoter. The custom code implementing this pipeline writes out the annotated GTF with the modified exon feature labels and transcript attributes (promoter group and splicing event metrics) as the final output (Custom Python Module). All analyses were performed with this in-house Python script, and the methodology described here reflects the exact logic in that code. (No transcript abundance quantification or TPM filtering steps were applied in this annotation pipeline, focusing purely on structural and coverage-based splicing evidence.)

Alternative splicing event calling (transcript-centric model): After isoform collapsing, we treated each full-length transcript, rather than individual splice junctions, as the primary unit for splicing analysis. The workflow proceeds gene-by-gene and relies only on the GTF structure plus an internally derived, strand-aware exon–coverage profile; no junction counts or short-read PSI estimates are used. For every transcript we stored a vector of binary flags indicating which events it includes; the complementary transcripts implicitly exclude those same events.

Percent spliced (PS) assignment: Within each promoter group, transcript TPMs were converted to isoform-usage proportions (TPM of a transcript divided by the sum of TPMs of all transcripts that share its promoter). The $\log_2$ fold change of this proportion between the two biological conditions was taken as that transcript percent spliced (PS) value. Because every alternative splicing event is physically embedded in the full-length isoform, the PS value of an event is inherited directly from the PS of the transcript(s) that include it (and conversely from the excluding transcript for the “skipped” form). Thus each event receives a single, internally consistent magnitude that exactly matches the change in usage of the complete transcript that embodies it.

Why a transcript-centric model is preferable for Nanopore direct-RNA data: *Full combinatorial context*. Nanopore reads span the entire mRNA, so alternative promoters, splice sites, alternative exons, intron retention, and poly(A) choices are observed simultaneously in each molecule. Scoring events through whole-isoform proportions preserves these naturally co-occurring combinations and avoids double counting overlapping junctions. *Coherence between splicing and expression*. The PS metric is exactly the differential usage of a biologically intact transcript; downstream functional consequences (protein isoforms, UTR changes) are therefore interpreted on the same quantitative scale. *Elimination of incompatible read allocations*. Short-read PSI methods must partition ambiguous junction reads among multiple events, often producing inconsistent or inflated estimates. Long-read isoform proportions bypass this ambiguity entirely. *Promoter-aware normalization*. By computing proportions within promoter groups, variation caused by alternative TSSs is separated from downstream splicing changes, preventing dilution or exaggeration of splice-site effects. *Direct compatibility with coverage-based QC*. The same normalized coverage profile used to discover events also confirms their quantitative support, ensuring that low-depth artefacts are not inflated into false positives. Together, these features make the transcript-centric, proportion-based PS framework the most logical and internally consistent way to quantify alternative splicing with Nanopore direct RNA-sequencing. All event calls and PS assignments were produced by the custom Python module available in our public repository.

Statistical Tests

Figure 2B-E, utilized two-sided exact binomial test for imbalance between up- and down-regulated counts. For Figure 2F–G, distal- versus proximal-boundary counts in up- and down-regulated (and unaffected) isoforms were compared with two-sided Fisher’s exact tests (2 × 2 tables, Bonferroni-corrected for multiple comparisons). Figure 3 analyses used Pearson’s chi-square tests to compare the proportion of single- versus multi-isoform genes and two-sided Fisher’s exact tests to assess enrichment of alternative splicing event categories. For Figure 5A and B, chi-squared tests were used to evaluate differences in the distribution of

splicing event types contributing to Pfam domain gains or losses between young and aged ovaries under PMSG and PMSG+hCG stimulation.

Data and Code Availability

All command scripts, exact software versions, and detailed runtime parameters are available in the public repository: https://github.com/DrNAlvarez/Dysregulation-of-alternative-splicing-patterns-in-the-ovaries-of-reproductively-aged-mice-. Raw direct-RNA FASTQ files were deposited in the Sequence Read Archive (Leinonen *et al.*, 2011), under SRA project ID: PRJNA1263509.

Results

Direct RNA-sequencing reveals novel transcripts in the mouse ovary

Direct RNA-sequencing was performed on ovaries from mice at two ages (6 weeks-old and 14 months-old), following gonadotropin stimulation with PMSG alone or PMSG followed by hCG (see Methods for details). Across twelve sequencing runs, each sample yielded a high number of RNA molecules: on average ~3.63×10^5 RNA molecules were sequenced per run (SEM $\pm 1.05 \times 10^5$; Figure 1A). Analysis of read lengths showed that the transcripts spanned a wide size range (from a few hundred to over 10^4 nucleotides), with a mean RNA length of approximately 964 nt (SEM ±0.36 nt; Figure 1B). Notably, the length distribution was highly similar across all experimental conditions (young vs. aged ovaries, with or without hormonal stimulation), indicating a consistent transcript size profile in the ovarian transcriptome.

To comprehensively characterize the ovarian transcriptome, we pooled all direct RNA-seq data from the 12 runs and annotated the unique transcripts by Ensembl transcript biotype (Figure 1C, left). In total, 56,862 unique transcripts were identified. A striking majority of these transcripts (72.7%) were classified as *unannotated*, indicating they do not correspond to any known Ensembl transcript. Among the annotated fraction, *protein-coding* transcripts accounted for 19.4%, and *retained intron* transcripts made up 2.8%. Smaller proportions of known transcript types were detected as well, including long noncoding RNA (lncRNA, 1.3%) and protein-coding transcripts with undefined coding sequence (CDS not defined, 1.3%). All remaining minor categories (each contributing <1% of transcripts) were combined into a single “Other” bin, together representing 2.5% of the transcriptome. Because such a large fraction of ovarian transcripts lacked annotation, we next evaluated their protein-coding potential using the Bioseq2seq deep-learning model for mRNA classification (Valencia and Hendrix, 2023). This analysis revealed that 51.6% of the detected transcripts exhibit protein-coding potential, whereas 48.4% are predicted to be noncoding (Figure 1C, right). Notably, the proportion of transcripts with coding potential far exceeds the fraction annotated as known protein-coding, suggesting that many of the novel, unannotated transcripts may indeed encode proteins.

Transcriptome wide detection of alternative splicing in the reproductive-aged ovary

An advantage of using direct RNA-sequencing is the ability to assess splicing events at single molecule resolution. We interrogated alternative splicing patterns in the reproductively-aged mouse ovary by identifying and quantifying splicing events, followed by statistical analysis to determine the significance of observed changes. An example is provided for the gene ENSMUSG00000029810 (*Tmem176b*) (Figure 2A).

We analyzed alternative splicing within the context of an individual promoter within a gene (see Methods). Our analysis assumes a model of post-transcriptional gene regulation where pre-mRNA generated from a promoter is alternatively spliced into mRNA independently and uniquely, compared to other pre-mRNA generated within the same gene. In Figure 2A, alternative 5' splice sites are marked with orange dashed lines, alternative 3' splice sites with red dotted lines, and alternative ends with black dash-dot lines. Alternative exons are represented by red shaded rectangles with low opacity, introns are highlighted in green, and retained introns are shown in a darker green, indicating retention. These visual markers help to identify and differentiate between various splicing events and their locations within the gene. Additionally, we grouped transcripts into promoter groups within a gene, Figure 2A is an example of this grouping.

To quantify how reproductive aging affects specific alternative splicing events, we measured isoform usage changes in 6-week versus 14-month ovaries under two hormone conditions: PMSG alone or PMSG followed by hCG. For each transcript, we computed the $\log_2$ fold-change in percent spliced (*PS*), the statistical

significance, and propagated that value to each alternative splicing event embedded within the transcript (see Methods). This allowed us to examine event-level changes in inclusion (the transcript contains the event) and exclusion (the transcript skips the event). Panels B–E of Figure 2 summarize these *PS* shifts for alternative exons, alternative 5′ and 3′ splice sites, and intron retention.

Under PMSG stimulation (Figure 2B), alternative exon inclusion events showed a strong downward bias in aged ovaries: 283 events were down-regulated compared to only 51 up-regulated (binomial $p < 10^{-39}$), with 514 unaffected. This trend intensified after hCG (Figure 2C), with 762 CE inclusion events down-regulated and 535 up-regulated ($p = 3 \times 10^{-10}$). Alternative 5′ splice site usage were similarly down-biased under both conditions (Figure 2B: 46 down / 13 up, $p = 2 \times 10^{-5}$; 2C: 103 down / 70 up, $p = 0.011$). Alternative 3′ splice sites inclusion events showed no significant imbalance under PMSG (Figure 2B, $p = 0.30$) but became slightly up-regulated after hCG (Figure 2C: 99 up / 72 down, $p = 0.046$). In contrast, intron retention inclusion events displayed a significant upward shift after hCG (80 up / 26 down, $p = 1 \times 10^{-7}$) but not under PMSG alone (Figure 2B, $p = 0.18$).

For transcripts that skip the annotated events, we observed different trends. Intron retention exclusion events were significantly down-regulated in aged ovaries under both conditions (Figure 2D: 194 down / 78 up, $p = 1 \times 10^{-12}$; 2E: 420 down / 302 up, $p = 1 \times 10^{-5}$). Alternative exon exclusion events showed no directional bias after PMSG alone (Figure 2D: 60 down / 58 up, $p = 0.93$), but a significant upward shift after hCG (Figure 2E: 475 up / 188 down, $p = 1 \times 10^{-29}$), consistent with increased exon skipping. For alternative 5′ and 3′ splice site exclusion events, aged ovaries showed strong upregulation (Figure 2E: alternative 5′ 288 up / 214 down, $p = 0.0014$; alternative 3′ 1,871 up / 1,301 down, $p = 3.8 \times 10^{-24}$), indicating increased production of transcripts that skip alternative splice sites.

In aged ovaries we detected marked shifts in transcript boundary choice (Figure 2F, G). Transcripts whose *PS* value increased with age were far more likely to originate from distal start and end sites, whereas transcripts whose PS decreased favored proximal boundaries. Under PMSG stimulation (Figure 2F), 44 of 100 PS-increased promoter isoforms (44 %) used a distal TSS, compared with 320 of 1 392 PS-decreased isoforms (23 %; odds ratio ≈ 2.6, Fisher's exact $p = 9.6 \times 10^{-6}$). For transcript ends, 29 of 100 PS-increased events (29 %) employed a distal TES versus 121 of 1 392 PS-decreased events (8.7 %; OR ≈ 4.3, $p = 2.5 \times 10^{-8}$). The same pattern held after the LH-like surge (PMSG + hCG; Figure 2G): 347 of 760 PS-increased promoter isoforms (45.7 %) used distal starts compared with 1248 of 4 751 PS-decreased isoforms (26.3 %; OR ≈ 2.4, $p = 5.1 \times 10^{-26}$), and 203 of 760 PS-increased transcripts (26.7 %) ended distally versus 346 of 4 751 PS-decreased transcripts (7.3 %; OR ≈ 4.6, $p = 3.3 \times 10^{-48}$). No significant interaction with hormone treatment was detected, indicating that the age-associated preference for distal transcript boundaries in PS-increased isoforms, and for proximal boundaries in PS-decreased isoforms, remains consistent across both PMSG and PMSG + hCG conditions.

Reproductive aging alters the alternative splicing landscape

To investigate how reproductive aging affects the splicing landscape, we first assessed gene-level transcript complexity by comparing the prevalence of single- versus multi-isoform genes. Under gonadotropin stimulation with PMSG, the fraction of genes expressing only a single transcript dropped from 59.1% in young ovaries to 40.9% in aged ovaries, with a corresponding rise in multi-transcript genes from 40.9% to 59.1% (Figure 3A). A similar age-related shift was observed following the PMSG+hCG treatment, as single-transcript genes declined from 55.2% in young to 44.8% in aged ovaries (Figure 3B). These changes in transcript complexity were highly significant (chi-square test, $p < 0.0001$), indicating that reproductive aging broadly increases isoform diversity per gene.

Aging also broadly reconfigured the spectrum of alternative splicing event types across transcripts (Figure 3C). In aged ovaries, transcripts were far more likely to utilize alternative exons at their ends: the incidence of alternative last exon usage approximately doubled (~7.5% in young vs ~15.6% in aged), and the use of alternative 3′ terminal splice sites/polyadenylation increased over three-fold (~4.8% vs ~17.7%). Alternative first exon usage also rose with age (from ~31% to ~38%), consistent with shifts in promoter usage. Conversely, intron retention events became comparatively rare in aged transcripts, dropping from ~7.5% of splicing events in young to under 1% in aged. Collectively, these landscape-level changes illustrate that aging preferentially

remodels splicing at gene boundaries, promoting greater use of alternative transcript start and end sites, while reducing the occurrence of retained introns. This pattern suggests that reproductive aging involves altered selection of promoters and polyadenylation sites, ultimately favoring the production of longer transcript isoforms in the aged ovary.

Reproductive aging influences stop codon and ORF distribution through alternative splicing

Alternative splicing not only alters exon composition but also impacts coding potential by introducing premature stop codons or modifying open reading frames (ORFs). To assess how reproductive aging affects these outcomes, we focused on splicing events that were significantly different between young (6-week-old) and aged (14-month-old) ovaries under gonadotropin stimulation. In total, hundreds of splicing changes with age were predicted to alter coding sequences. Under PMSG treatment, 757 splicing events showed significant age-related differences that affected ORFs, and 247 such events introduced stop codons. With the addition of hCG (PMSG+hCG), the number of age-affected events increased to 2,829 for ORF-altering events and 983 for stop codon–introducing events (pie chart totals in Figure 4B,C).

We next examined the types of splicing events contributing to these coding changes. Figure 4B summarizes the distribution of splicing event types among the age-regulated events that alter ORFs. In both hormone conditions, *alternative exon* usage was the predominant driver of ORF changes, accounting for ~42% of events in both PMSG (left pie) and PMSG+hCG (right pie). *Alternative last exon* usage was the second most common category (33.4% under PMSG vs. 30.9% under PMSG+hCG), indicating that many age-associated splicing changes involved the usage of different terminal exons. *Alternative 3′ end* and *alternative 5′ end* splicing contributed more modestly to ORF changes (PMSG: 13.7% and 10.4%, respectively; PMSG+hCG: 15.2% and 9.8%). *Intron retention* events were relatively rare among ORF-altering changes (0.5% under PMSG, increasing to 2.5% under PMSG+hCG). Notably, the higher proportion of retained-intron events in aged mice after ovulatory hCG stimulation suggests that intron retention became more prevalent with age in that context. A chi-square test comparing the two stimulation conditions confirmed that intron retention events were significantly enriched among the age-affected ORF changes in the PMSG+hCG condition ($\chi^2 = 10.50$, $p = 0.0012$). Fisher's exact test likewise indicated a disproportionate impact of intron retention on ORF alteration with age ($p = 0.027$). By contrast, differences in the contributions of other splicing types (alternative exons or 5′/3′ ends) to ORF changes were not statistically significant (all $p > 0.05$).

Alternative splicing not only affects exon composition but also influences coding potential by introducing novel stop codons and modifying open reading frames. To assess the impact of reproductive aging on these processes, we examined the distribution of stop codons introduced by age-affected splicing events in young (6-week) and aged (14-month) mouse ovaries following PMSG or PMSG+hCG stimulation (Figure 4C). We further compared the positions of these splicing-associated stop codons relative to the final exon in transcripts from young vs. aged ovaries under each hormonal condition (Figures 4D, 4E).

We found that the majority of age-sensitive alternative splicing changes leading to a stop codon were due to alternative last exon usage in both hormone conditions. Following PMSG stimulation (Figure 4C, left), alternative last exons accounted for 80.2% of such events, with smaller contributions from alternative 3′ end splicing (9.7%), alternative exon inclusion (6.1%), alternative 5′ end splicing (2.8%), and intron retention (1.2%). After the ovulatory PMSG+hCG stimulus (Figure 4C, right), alternative last exons remained the dominant category (75.3%), while the relative contribution of other event types increased (for example, alternative exon events 8.3%, intron retention 4.9%, and alternative 5′ end events 5.2%). A Chi-square analysis of event-type frequencies indicated that intron retention events were significantly enriched among those splicing changes favoring the aged group ($\chi^2 = 5.79$, $p = 0.016$), whereas no significant age bias was observed for alternative exon ($\chi^2 = 1.10$, $p = 0.29$), alternative 5′ end ($\chi^2 = 1.94$, $p = 0.16$), or alternative 3′ end events ($\chi^2 = 3.02$, $p = 0.08$).

We next examined whether aging influences where these splicing-induced stop codons occur within transcripts. For each condition, events were classified by whether the introduced stop codon occurred upstream of the final exon or within the final exon of the transcript. In the PMSG-only condition (Figure 4D), young and aged ovaries showed a very similar distribution of stop codon positions, with no significant difference between age groups. Under PMSG+hCG stimulation (Figure 4E), aged ovaries tended toward a

higher proportion of splicing events that generate stop codons before the last exon compared to young. This shift was most pronounced for intron retention events: in aged ovaries, intron retention was significantly more likely to produce a stop codon prior to the final exon than in young ovaries (Fisher's exact test, $p = 0.027$). Other splicing event types did not show a significant age-related shift in stop codon location (for example, alternative exon events showed a similar trend toward more upstream stops in aged ovaries, but this did not reach significance, $p = 0.057$). Thus, although reproductive aging increases the incidence of splicing alterations that introduce stop codons, the overall positional distribution of these stop codons remains largely comparable between young and aged ovaries, with only specific event types (notably intron retentions under an ovulatory stimulus) showing a marked difference in stop codon placement.

Reproductive aging alters protein domain architecture through splice type–specific remodeling

To investigate how reproductive aging reshapes protein architecture, we compared Pfam domain usage between young and aged ovaries after gonadotropin stimulation. Figure 5 summarizes these results, with panels A and B depicting the distribution of splicing event types that alter Pfam domains, and panel C highlighting specific domain enrichments by splice type. In PMSG-stimulated ovaries, a total of 498 Pfam domain-altering splicing events were observed. These events were dominated by Alternative Exon splicing (44.0% of all domain events), with substantial contributions from Alternative Last Exon events (34.1%). Alternative 3′ End and Alternative 5′ End splice events were less common (12.0% and 9.4%, respectively), and Intron Retention was very rare (0.4%). The distribution of splice event types differed by age (overall χ^2 tests, $p<0.01$). Notably, aged ovaries exhibited a higher proportion of Alternative Last Exon-mediated domain changes compared to young under PMSG stimulation, whereas intron retention remained ≤1% of events in both age groups (showing no significant age difference). Pfam domain modifications at transcript start sites (Figure 5A, middle; 108 total Pfam start events) similarly were largely driven by Alternative Exon usage (50.0%), followed by Alternative 3′ End events (24.1%) and Alternative 5′ End events (18.5%), with smaller contributions from Alternative Last Exon events (6.5%) and only ~1% from intron retention. Likewise, Pfam domain stop alterations (Figure 5A, right; 197 total Pfam stop events) were overwhelmingly mediated by Alternative Last Exon splicing (72.1%), with fewer contributions from Alternative Exon (14.2%), Alternative 3′ End (7.6%), and Alternative 5′ End (5.6%) events; only a single intron retention event was observed (0.5%). No significant age-dependent shifts were detected in the Pfam start or stop categories under PMSG alone (intron retention remained minimal in both young and aged).

Following the LH-like surge (PMSG+hCG), the overall prevalence of domain-altering splicing events increased (1627 total events), but the distribution by splice type remained similar. Alternative Exon events again constituted the largest category (45.3%), with Alternative Last Exon events next (31.3%), followed by Alternative 3′ End (12.0%) and Alternative 5′ End events (9.0%). Intron Retention events, while still infrequent (2.3%), were more prevalent under PMSG+hCG than with PMSG alone. Importantly, intron retention increased from <1% of domain-altering events in young mice to ~3% in aged mice, reflecting a significantly higher fraction of Pfam-affecting events in aged ovaries after the hCG surge ($\chi^2=6.71$, $p = 9.59\times10^{-3}$ for the overall distribution). Concomitantly, the contribution of Alternative Last Exon events to domain changes was relatively lower in aged hCG-treated ovaries. Splicing events altering Pfam start sites under PMSG+hCG (Figure 5B middle; 324 events) showed a comparable pattern to PMSG: Alternative Exon events comprised 51.9%, with Alternative 3′ End (26.2%) and Alternative 5′ End (14.5%) usage next, a small share of Alternative Last Exon events (4.3%), and few intron retentions (3.1%). In Pfam stop site alterations (Figure 5B right; 663 events), Alternative Last Exon splicing remained dominant (66.4% of stop-related events), and Alternative Exon usage was similar to PMSG (15.2%). However, intron retention events affecting Pfam stops, while still uncommon (4.2%), were markedly higher with PMSG+hCG than under PMSG. This age-associated rise in intron retention at protein 3′ ends was statistically significant ($\chi^2=5.35$, $p = 2.08\times10^{-2}$), with aged mice exhibiting a noticeable intron retention contribution to domain truncation (~4% in aged vs ~0.5% in young). Aside from intron retention, the proportions of Alternative 3′ End (6.3%) and Alternative 5′ End (7.8%) events in Pfam stop panels under PMSG+hCG were roughly comparable to those seen with PMSG.

Figure 5C presents heatmaps of Pfam domain enrichment for transcripts differentially spliced between 6-week-old and 14-month-old ovaries under two hormonal conditions (PMSG+hCG and PMSG alone). Across both treatments, transcripts with alternative exon usage and altered splice sites (3′ or 5′) exhibited the most

extensive patterns of Pfam domain enrichment. Many of the enriched Pfam domains correspond to mitochondrial proteins, including subunits of respiratory Complexes I and IV, and these were more prevalent in transcripts from aged ovaries. These mitochondrial-related domains were distributed across multiple splicing event categories, particularly among transcripts with alternative exons and splice-site changes. For example, *Ndufs4* (a Complex I subunit gene) was among the transcripts significantly differentially spliced with age.

Ndufs4 is differentially spliced following PMSG + hCG, with aged-associated increases in two alternative transcript isoforms (Figure 5D, ~29-fold and ~4-fold). In contrast, following PMSG alone, *Ndufs4* splicing is not significantly altered by age. *Figure 5D* illustrates the two *Ndufs4* transcript isoforms: the longer isoform contains a larger open reading frame that encodes a conserved Pfam domain (blue), whereas the shorter isoform lacks this domain. Correspondingly, Figure 5E shows AlphaFold 3 (Abramson *et al.*, 2024) models of the mitochondrial Complex I N-module assembled with either the short (left image) or long (right image) NDUFS4 isoform. In the model containing the short isoform (Figure 5E, left), the NDUFS4 subunit (dark orange) appears truncated, while the arrangement of the other subunits remains similar to the model with the long isoform.

Discussion

Post-transcriptional remodeling in aged ovaries

Our direct RNA-sequencing analysis reveals that ovarian aging is accompanied by extensive post-transcriptional changes. In comparing young and aged mouse ovaries, we observed multiple alterations in mRNA splicing and isoform expression. Key findings include: 1) increased isoform diversity: aged ovaries express a broader repertoire of transcript isoforms per gene, indicating expanded transcriptomic heterogeneity with age. This suggests that aged ovaries produce more alternative mRNA variants than young ovaries, potentially as a result of widespread splicing changes. 2) Age-biased exon usage: we detected numerous instances of differential exon inclusion or skipping between young and aged ovaries. Certain exons are used preferentially in one age group over the other, pointing to systematic shifts in splicing patterns as a function of age. This age-biased exon usage implies that pre-mRNA processing is dynamically regulated during aging, leading to different protein-coding sequences in the transcripts of aged cells. 3) Shifts in transcript boundaries: aging also influenced transcript start and end sites. We observed changes in the usage of alternative transcription start sites and polyadenylation sites, resulting in mRNA with altered 5′ or 3′ untranslated regions in aged ovaries. These boundary shifts can affect mRNA stability, localization, or translation, underscoring another layer of post-transcriptional regulation that diverges with age. 4) Altered splicing of mitochondrial gene transcripts: notably, genes related to mitochondrial function exhibited pronounced splicing alterations in aged ovaries. Several nuclear-encoded mitochondrial components, particularly subunits of the electron transport chain, showed age-dependent isoform expression. This includes the emergence of truncated or otherwise modified transcripts for critical mitochondrial proteins in aged ovaries. Such changes hint that the gene programs sustaining cellular energetics are especially vulnerable to splicing modulation during aging.

These results paint a picture of an aging ovary that is not only characterized by changes in gene expression levels, but also by qualitative changes in mRNA isoforms. In other words, it is not just *how much* of a gene is expressed that changes with age, but *which versions* of the gene transcript are produced. This expanded isoform diversity and biased exon usage in aging aligns with observations in other systems that aging is associated with pervasive alterations in RNA splicing (Angarola and Anczuków, 2021). Age-related splicing alterations have been documented in multiple tissues, often affecting genes tied to fundamental aging processes (Angarola and Anczuków, 2021). Our findings extend this knowledge to the ovarian context, revealing that the ovarian transcriptome undergoes a similar diversification with age.

Isoform diversity and the splicing–energy–aging axis

An intriguing interpretation of the increased isoform diversity and exon skipping in aged ovaries is through the lens of the recently proposed “splicing–energy–aging axis”. This hypothesis (Ferrucci *et al.*, 2022) posits that cellular energy status and RNA splicing are interconnected, with changes in mitochondrial function influencing splicing as a potential adaptive response to stress (Ferrucci *et al.*, 2022). Our results provide an opportunity to critically examine this idea in the context of reproductive aging.

Consistent with the energy–splicing axis hypothesis, many splicing changes we observed in aged ovaries involve genes related to metabolism and stress responses. Large-scale analyses of aging tissues have noted that alternative splicing events with age often occur in genes linked to mitochondrial function, metabolic processes, and apoptosis (Angarola and Anczuków, 2021). For instance, an aging study in multiple mouse tissues found that skipped exon events were frequent in aged animals, particularly in genes involved in mitochondrial processes and inflammation (Angarola and Anczuków, 2021). Our ovarian data echo this pattern: we see preferential alternative splicing in genes that are integral to cellular energetics (e.g., components of the electron transport chain and metabolic enzymes) and stress pathways. This convergence suggests that as ovarian cells confront age-related mitochondrial decline, they may modulate mRNA splicing of relevant genes, perhaps as a compensatory mechanism.

Experimental evidence from other systems supports a direct link between energetic stress and splicing modulation. In cell culture, disrupting mitochondrial function with agents like paraquat (which induces an energy deficit and oxidative stress) triggers widespread changes in alternative splicing patterns (Angarola and Anczuków, 2021). Notably, these splicing alterations correlate with ATP depletion, and oxidative stress alone (without energy loss) is insufficient to elicit the same splicing changes (Angarola and Anczuków, 2021). This finding implies that a drop in cellular energy availability can act as a signal to reshape the splicing program. In light of this, the aged ovary – known to suffer mitochondrial dysfunction and lower bioenergetic efficiency – might undergo splicing reprogramming as a resilience strategy. The increased production of alternative isoforms could be an attempt by cells to preserve homeostasis under energetic duress (Angarola and Anczuków, 2021). For example, skipping certain exons might produce shorter isoforms that are quicker or less energetically costly to translate, or might eliminate segments of proteins that are dispensable under stress conditions. Through such adjustments, cells could potentially conserve resources or redirect their functional capacity in the face of diminishing ATP levels.

However, our data also urge a critical nuance to the splicing–energy–aging axis hypothesis: not all splicing changes in aging appear beneficial. While some alternative splicing may be adaptive, the overall increase in splicing variability might equally reflect a breakdown in regulatory precision. The machinery responsible for splicing (the spliceosome and its regulatory factors) itself is energy-intensive and subject to age-related deregulation (Angarola and Anczuków, 2021). In aged ovaries, the observed splicing dysregulation – exemplified by aberrant exon skipping or the use of cryptic splice sites – could indicate a loss of splicing fidelity as cells age. This loss of fidelity might be a direct consequence of energetic stress (since correct splicing requires sufficient ATP and functional spliceosomal components) or other age-associated changes (such as altered expression of splicing factors). In other words, what we detect as “increased isoform diversity” could be a double-edged sword: part of a regulated response to stress, but also part of the stochastic noise introduced by an aging, less efficient cellular machinery.

Thus, our findings support the idea of a splicing–energy–aging axis, while also refining it. The aged ovary does show splicing changes consistent with an energy-stress response (supporting the notion of splicing as a resilience mechanism (Ferrucci *et al.*, 2022)), but simultaneously, some of these changes may undermine cellular function. This complex picture aligns with a broader view that RNA splicing in aging can be bi-directional in its effects (Angarola and Anczuków, 2021). It may help cells cope up to a point, yet beyond that, splicing errors or maladaptive isoforms can accumulate and contribute to aging phenotypes. In the context of ovarian aging, this means alternative splicing might help adjust to metabolic stress in early or mid-reproductive life, but in later life it could become a factor that exacerbates ovarian functional decline.

Mitochondrial Complex I and NDUSF4: a case of age-dependent mis-splicing

One striking example from our dataset underscores how age-related splicing changes can have concrete functional repercussions: the *Ndufs4* gene, which encodes an accessory subunit of mitochondrial Complex I (NADH:ubiquinone oxidoreductase). In aged ovaries, we found that *Ndufs4* pre-mRNA is differentially spliced to produce a truncated isoform that was undetected in young ovaries. This truncated transcript likely results from skipping of a critical exon (or inclusion of a cryptic exon/intron) that introduces a premature stop codon, yielding an mRNA that encodes a shortened NDUFS4 protein. The prevalence of this *Ndufs4* isoform in aged ovaries suggests an age-biased splicing switch that favors a non-canonical, likely non-functional variant of a key mitochondrial gene.

The functional importance of NDUFS4 in mitochondria is well-established. NDUFS4 is essential for the stability and assembly of Complex I, although it is an “accessory” subunit (not part of the catalytic core). Loss-of-function mutations in *Ndufs4* are known to destabilize Complex I and reduce its activity. In mice, complete knockout of *Ndufs4* leads to a severe phenotype with neurodegeneration and metabolic deficits, due to an inability to form fully functional Complex I. Even at the cellular level, NDUFS4 deficiency has measurable consequences: *Ndufs4*-null mouse cells show an accumulation of inactive Complex I subcomplexes and exhibit elevated levels of reactive oxygen species (ROS) (Valsecchi *et al.*, 2013). In other words, without functional NDUFS4, Complex I cannot effectively transfer electrons, leading to electron leak and increased ROS generation. Given this background, the age-dependent splicing change in *Ndufs4* is highly pertinent. By favoring a truncated *Ndufs4* isoform, aged ovaries may be functionally reducing NDUFS4 protein levels or producing a defective version of the protein. Either outcome would impair Complex I assembly or activity. We can infer that aged ovarian cells, especially oocytes or granulosa cells which rely on mitochondrial respiration, could suffer a partial Complex I deficiency as a result of this splicing change. Even a moderate reduction in Complex I efficiency can perturb cellular metabolism; one likely consequence is increased production of ROS by mitochondria that are running below optimal capacity.

This mechanism dovetails with a known hallmark of ovarian aging: elevated oxidative stress in the oocyte and its microenvironment. Prior studies have documented that oocytes from aged females accumulate mitochondrial damage and ROS, which contribute to their reduced developmental competence (Sasaki *et al.*, 2019). Our findings provide a novel molecular explanation for this phenomenon. Mis-splicing of *Ndufs4* (and potentially other electron transport chain genes) in aged ovaries could be one cause of mitochondrial dysfunction, leading to excess ROS generation. Excess ROS, in turn, can damage cellular components in oocytes and somatic follicle cells, hastening their attrition. Thus, the aberrant splicing of *Ndufs4* links post-transcriptional regulation to a bioenergetic and quality control failure in aging oocytes: as *Ndufs4* transcripts skew toward a nonfunctional form, Complex I activity drops, ROS rise, and the viability of the oocyte is compromised.

It is worth highlighting that alternative splicing of *Ndufs4* is not just an oddity observed in our aging study, but part of a broader regulatory theme. There is evidence that cells can modulate NDUFS4 levels via splicing under certain conditions. For example, recent work showed that the protein IQGAP1 can influence *Ndufs4* pre-mRNA splicing, thereby affecting Complex I activity as a means of communication between the nucleus and mitochondria (Ferrucci *et al.*, 2022). This indicates that the splicing of *Ndufs4* is a lever that cells may pull to adjust mitochondrial function. In the context of aged ovaries, one might speculate that the increased truncated *Ndufs4* isoform could initially arise from a regulatory attempt to fine-tune metabolism under stress (consistent with the “resilience” idea). However, if that regulation overshoots, or if the truncated isoform accumulates due to a loss of splicing fidelity, it could become detrimental. The aged ovary may essentially be caught in a vicious cycle: energetic stress leads to splicing changes (like those in *Ndufs4*), which then further impair mitochondrial function and amplify stress (through ROS production), ultimately accelerating ovarian aging. This splicing-mediated mitochondrial feedback loop represents a new angle on how cellular aging processes can reinforce themselves.

Importantly, our identification of the *Ndufs4* truncated isoform in aged ovaries provides a concrete example where an age-associated splice variant is likely to have a direct pathogenic impact. While many of the isoform changes we cataloged will require functional validation, the *Ndufs4* case stands out because of the well-known consequences of its loss. Complex I deficiency from NDUFS4 loss mimics certain mitochondrial diseases and is known to trigger redox imbalances (Valsecchi *et al.*, 2013). Therefore, we propose that the age-biased splicing of *Ndufs4* contributes to the mitochondrial ROS burden in the aged ovary, linking a specific molecular change to a higher-order phenotype (increased oxidative damage and oocyte loss). This mechanistic link exemplifies how alternative splicing changes are not merely correlative markers of aging but can be causal players in age-related functional decline.

Alternative splicing in reproductive aging: adaptive vs. detrimental roles

Our findings contribute to current models of reproductive aging by adding a nuanced perspective: alternative splicing reprogramming is both a response to and a potential driver of ovarian aging. Classic models of ovarian aging have emphasized factors such as mitochondrial dysfunction, accumulated DNA damage, telomere

shortening, and changes in intercellular signaling within the follicle niche. We now see that post-transcriptional regulation – especially splicing – is an important piece of this puzzle. How might these splicing changes fit into the larger framework of ovarian aging, and do they help or hurt the aging ovary? The answer appears to be a bit of both.

On one hand, we can interpret some splicing changes as adaptive or compensatory. The aged ovary is an environment of chronic stress: follicles are fewer and often exposed to elevated oxidative stress and metabolic strain (Sasaki *et al.*, 2019). In such conditions, the ability to produce alternative protein isoforms could allow cells to adjust their physiology. For instance, an alternative isoform might lack a regulatory domain that is normally active under rich nutrient conditions, thereby dialing down a pathway to match the low-energy state. Alternatively, splicing could produce a protein variant that is more stable or effective in an oxidized environment. These kinds of adjustments would represent an effort by aging cells to preserve function despite deteriorating conditions. The concept of splicing as an adaptive mechanism is directly in line with the splicing–energy–aging axis hypothesis (Ferrucci *et al.*, 2022), which envisions splicing changes as a form of cellular resilience. In the ovary, where energy supply to oocytes and granulosa cells is critical for follicle survival, such splicing adaptations might help prolong the functionality of these cells under stress. In summary, from this angle, the complex splicing landscape in aged ovaries could be seen as a *catalog of survival tricks* – the ovary exploring different transcript isoforms to cope with mitochondrial insufficiency and other age-related challenges.

On the other hand, our data also reveal splicing changes that seem clearly detrimental. The production of nonfunctional or deleterious isoforms (like the truncated NDUFS4 discussed above) can actively undermine cellular health. These mis-spliced transcripts may lead to dominant-negative protein products or trigger nonsense-mediated decay that depletes crucial mRNA. In aged ovaries, such detrimental splicing events could contribute to the decline in ovarian reserve and oocyte quality. For example, if key DNA repair genes or cell cycle regulators are mis-spliced, oocytes might accumulate damage or enter arrest, hastening their loss. It is known from human and animal studies that the rate of oocyte attrition accelerates with advancing age, and while apoptosis and follicular atresia have been blamed on hormonal or external factors, our findings suggest an intrinsic factor – splicing errors – could be at play as well. In this view, alternative splicing acts as a double-edged sword during aging: the same process that provides flexibility and adaptation can also generate faulty components that exacerbate aging. This dualism has been noted in other aging contexts; for instance, studies in muscle and brain have found that greater alternative splicing activity accompanies both stress resistance and degenerative changes (Angarola and Anczuków, 2021). Our ovarian data resonate with those observations, indicating a conserved theme across tissues.

In integrating these perspectives, we propose that alternative splicing in the ovary serves as both a barometer and an effector of aging. It is a barometer in the sense that the splicing profile reflects the internal state of a cell (energy levels, stress, etc.) – essentially taking the pulse of the health of a cell. At the same time, it is an effector because changes in splicing can actively modify that state – either bolstering the cellular defenses or, if aberrant, contributing to dysfunction. This insight extends current reproductive aging models by linking metabolic and proteostatic theories of aging with gene regulatory changes at the mRNA level. It urges researchers to consider not just *how many* oocytes remain or *how intact* their DNA is, but also *which isoforms* those oocytes are expressing as they age. The path to ovarian senescence may well be paved by a series of splice site choices.

Unresolved questions and future directions

While this study significantly advances our understanding of post-transcriptional regulation in ovarian aging, it also raises important questions that remain to be answered: 1) *Cell-type specificity of splicing changes*: The ovary is composed of multiple cell types (oocytes, granulosa cells, theca cells, stromal and immune cells), and our bulk tissue RNA-seq cannot distinguish which cell type is driving each splicing change. Do the splicing alterations predominantly occur in oocytes themselves, or are they a feature of somatic support cells in the ovarian niche? For instance, the *Ndufs4* truncated isoform might be originating in energy-intensive oocytes, or perhaps in granulosa cells which also have high metabolic activity supporting the oocyte. It is also possible that changes in cell population proportions with age (e.g. fewer oocytes and growing proportion of stromal tissue) contribute to the observed isoform shifts. Clarifying the cellular origin of these splicing differences is crucial.

Future studies employing single-cell RNA-seq or spatial transcriptomics could resolve isoform expression at cell-specific resolution, identifying exactly where (and in which cell types) these age-related splicing events occur within the ovary. 2) *Functional impact of novel isoforms*: Although we have cataloged many age-associated isoform changes, the functional consequences of most of these changes are unknown. Do the alternative isoforms encode proteins with different or dominant-negative functions? Do they alter the efficiency of translation or mRNA stability? For example, we hypothesize that the truncated NDUFS4 isoform impairs Complex I, but direct evidence (such as measuring Complex I activity in young vs. aged ovarian cells, or rescuing the phenotype by forced expression of the canonical isoform) is needed. Similarly, other splicing changes might affect pathways like DNA repair, apoptosis, or follicular development, but these effects must be demonstrated. We need targeted functional experiments to test the impact of specific isoform switches. Such experiments could include overexpressing or knocking down particular isoforms in ovarian cell culture or mouse models, CRISPR/Cas9 editing to prevent or mimic certain splicing events, and examining the resulting effects on cellular metabolism, viability, and oocyte maturation. Determining which isoform changes are benign bystanders and which are key drivers of aging will be a major step forward. 3) *Feedback loops between splicing and mitochondrial function*: Our findings hint at a feedback loop where mitochondrial dysfunction (a hallmark of aging) and splicing dysregulation may reinforce each other. However, the causality and mechanisms of this interplay are not fully understood. Does declining ATP availability in aging cells directly cause splicing errors by impairing spliceosome function, or are there active signaling pathways (e.g. involving AMPK or other energy sensors) that reprogram splicing in response to low energy? Conversely, to what extent do splicing changes feed back to worsen mitochondrial performance (as in the case of *Ndufs4*)? Answering these questions will likely require integrative approaches, such as manipulating cellular energy status and observing splicing outcomes, or vice versa. For instance, one could experimentally boost mitochondrial function in an aged ovary model (perhaps via supplementing key metabolites) to see if the splicing profile becomes “younger” or more normal. Alternatively, one could disrupt specific splicing regulators in young ovaries and assess if mitochondrial gene expression and function begin to resemble an aged state. Investigating these feedback relationships will shed light on whether breaking the splicing–mitochondria vicious cycle can slow down ovarian aging. 4) *Adaptive vs. maladaptive splicing – what tips the balance?*: We have discussed that alternative splicing changes can be both adaptive and maladaptive. A pressing question is: what determines whether a given splicing change is beneficial or harmful for the cell? Is it the magnitude of change, the context, or perhaps the regulatory control behind it? It will be important to discern whether the aging-associated splicing changes are actively orchestrated by the cell (via signaling pathways and splicing factor modulation) or if they result from a loss of splicing fidelity. Approaches like profiling the expression or phosphorylation of splicing factors in young vs. aged ovaries could be informative, as would studies of interventions known to affect lifespan (e.g. dietary restriction, which affects cellular energy metabolism) on the splicing landscape. Notably, caloric restriction and mTOR inhibition can extend reproductive lifespan in model organisms, and it would be fascinating to see if they also curb the splicing drift we observed with age. Such investigations could reveal whether improving cellular energetics or bolstering splicing accuracy can maintain a more youthful transcriptome and delay functional decline.

Conclusion

Moving forward, future research should integrate genomic, bioenergetic, and reproductive biology approaches to build a more complete picture of ovarian aging. Our study underscores that the transcriptomic landscape (particularly splicing) is an important frontier in aging research. By pursuing the questions above, researchers can deepen our understanding of how alternative splicing interfaces with known aging mechanisms in the ovary. Ultimately, this knowledge will help to refine models of reproductive aging, for example, by identifying molecular markers of oocyte health related to splicing, or by revealing points of intervention where age-related decline might be slowed in model systems. We emphasize that alternative splicing is both a consequence and a contributor of the aging process in ovaries. This insight expands the paradigm of ovarian aging beyond static DNA damage or hormonal changes, framing it as a dynamic interplay between gene regulation and metabolism. By clearly delineating how and why splicing shifts occur with age, and what their downstream effects are, we edge closer to a holistic understanding of why ovarian function fails over time. Such understanding, grounded in our data, avoids speculative overreach but opens the door to more targeted hypotheses. In conclusion, our findings reveal that the regulation of mRNA splicing is a key facet of ovarian aging, one that likely acts in concert with energy stress and ROS to influence the lifespan of the female

reproductive system. This work lays a foundation for further studies to unravel the complex splicing–energy–aging axis in female germline longevity, and it prompts a re-evaluation of how post-transcriptional regulation shapes the fate of oocytes as organisms age.

Figure Legends

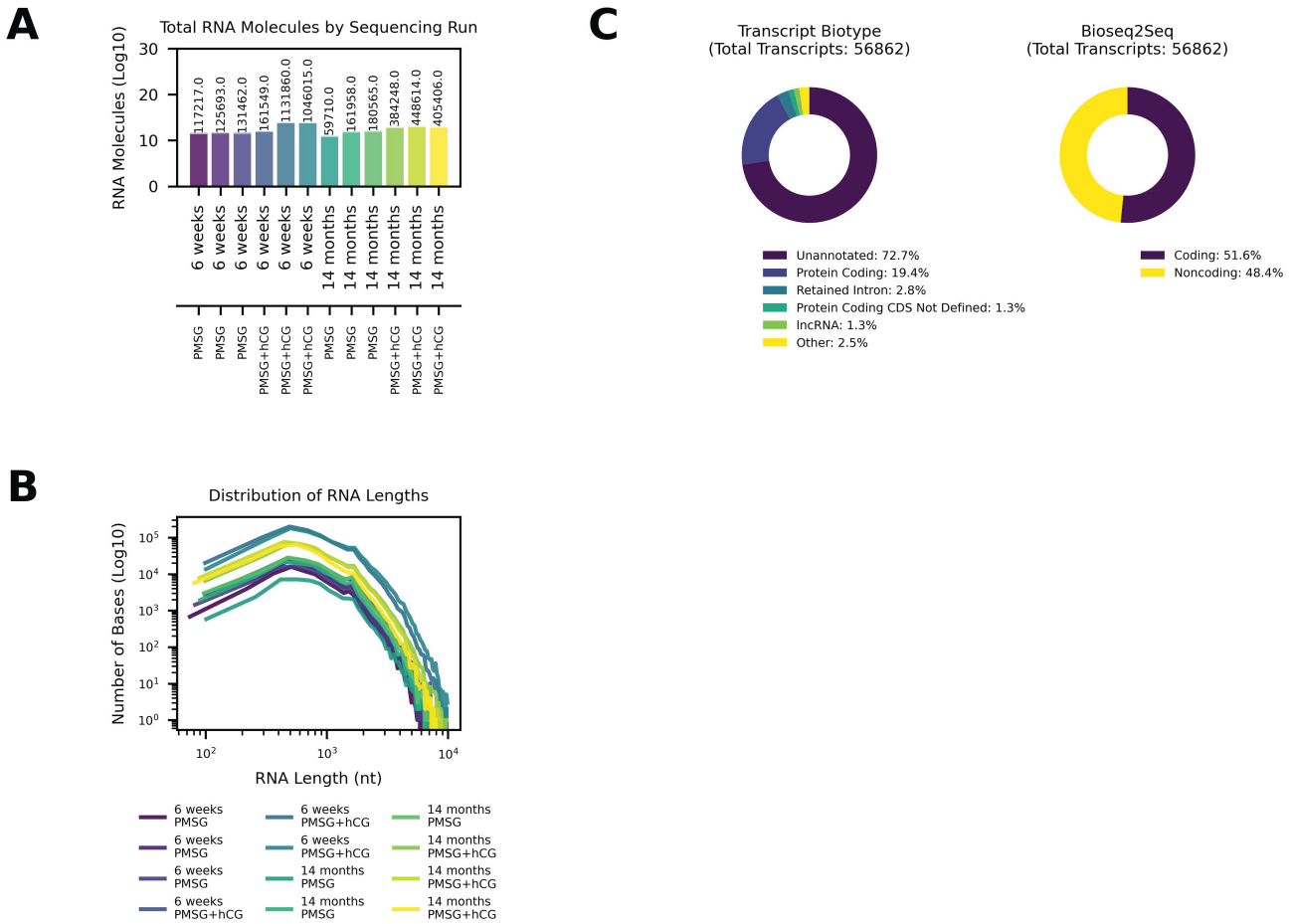

Figure 1. Direct RNA-sequencing yields a high transcript count and reveals novel transcripts in the mouse ovary. (A) Bar plot showing the total number of RNA molecules sequenced in each of 12 runs (y-axis on a $\log_{10}$ scale). Bars are colored for ease of viewing, age and gonadotropin stimulation marked on x-axis, and each bar represents an individual sequencing run. (B) Length distribution of sequenced RNA molecules. Each curve corresponds to one experimental condition (6-week or 14-month, +PMSG/+PMSG+hCG, color-coded as in the figure legend) and plots the total number of bases sequenced (y-axis, $\log_{10}$ scale) across transcript lengths (x-axis, in nucleotides). (C) Ovarian transcript composition and coding potential. Left: Donut chart illustrating the proportion of transcripts by Ensembl biotype among all unique transcripts identified. Right: Donut chart classifying these transcripts by predicted coding potential (Bioseq2seq model).

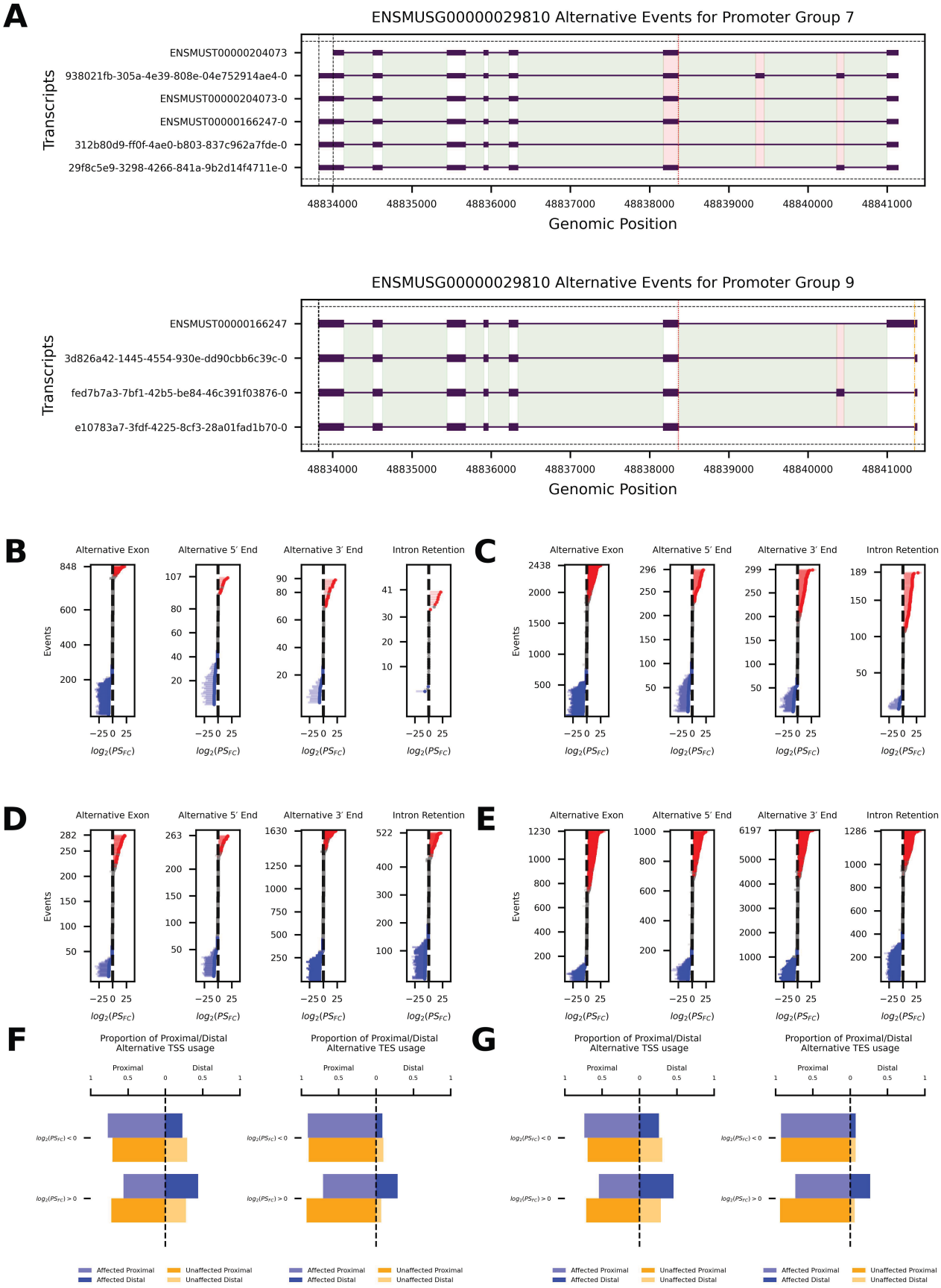

Figure 2. Reproductive aging skews alternative splicing outcomes and transcript boundary selection under gonadotropin stimulation. (A) Example gene (Tmem176b; ENSMUSG00000025955) with transcripts

from two different promoter groups. Alternative 5′ splice sites are indicated by orange dashed lines, alternative 3′ splice sites by red dotted lines, and alternative transcript end sites by black dash-dot lines. Alternative exons are shown as semi-transparent red rectangles, introns are highlighted in green, and retained introns are marked in dark green. (B–E) Summary of age-associated splicing events (alternative exons, alternative 5′/3′ splice sites, intron retention) for young vs. aged. Panels B and D show inclusion and exclusion events for 6-week vs. 14-month PMSG. Panels C and D show inclusion and exclusion events for 6-week vs. 14-month PMSG+hCG. Y-axis are individual alternative splicing events, x-axis is $\log_2$ fold change of percent spliced (*PS*). Horizontal red and blue lines are significant events, horizontal grey lines are not significant. (F–G) Proportional usage of proximal vs. distal transcript boundaries significantly affected (blue) or not affected (orange) isoforms in 6 weeks vs. 14 months under PMSG (F) or PMSG+hCG (G). F and G left sub-panel compares transcription start sites (TSS) and right sub-panel compares transcription end sites (TES).

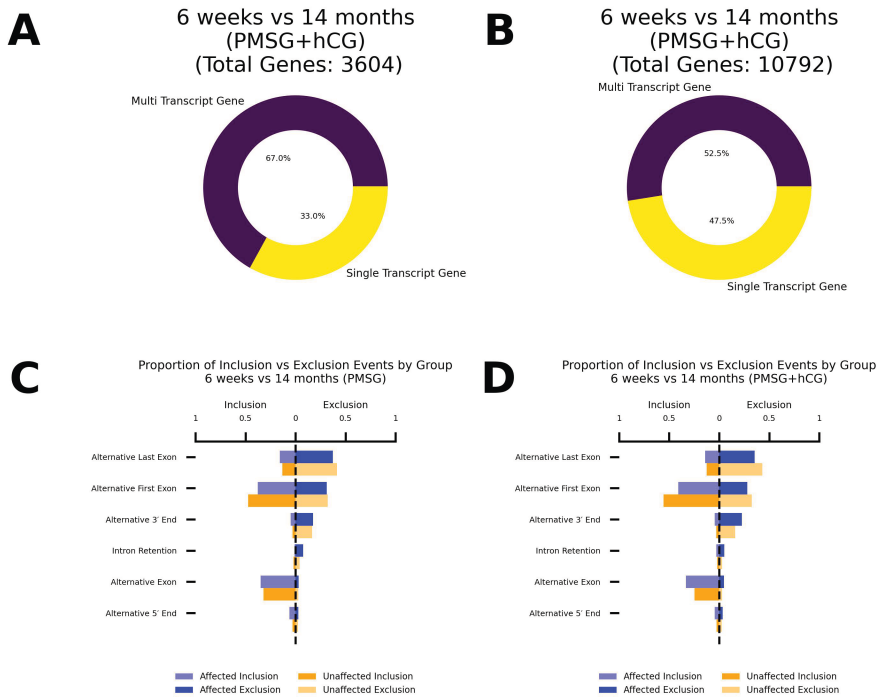

Figure 3. Reproductive aging increases transcript complexity and differentially affects splicing event types. Gene transcript complexity in 6-week vs. 14-month ovaries after PMSG (A) and PMSG+hCG (B). Donut charts show the percentage of single-transcript genes (yellow) versus multi-transcript genes (purple). (C, D) Distribution of age-affected vs. unaffected splicing events leading to exon inclusion or exclusion. Bar graphs illustrate the proportion of alternative splicing events (categorized by type) that are affected by aging (significantly different between 6-week vs. 14-month, blue bars) versus unaffected (no significant change, orange bars) in the PMSG (C) and PMSG+hCG (D) conditions. For each panel, the left side represents events resulting in inclusion, and the right side represents events leading to exclusion.

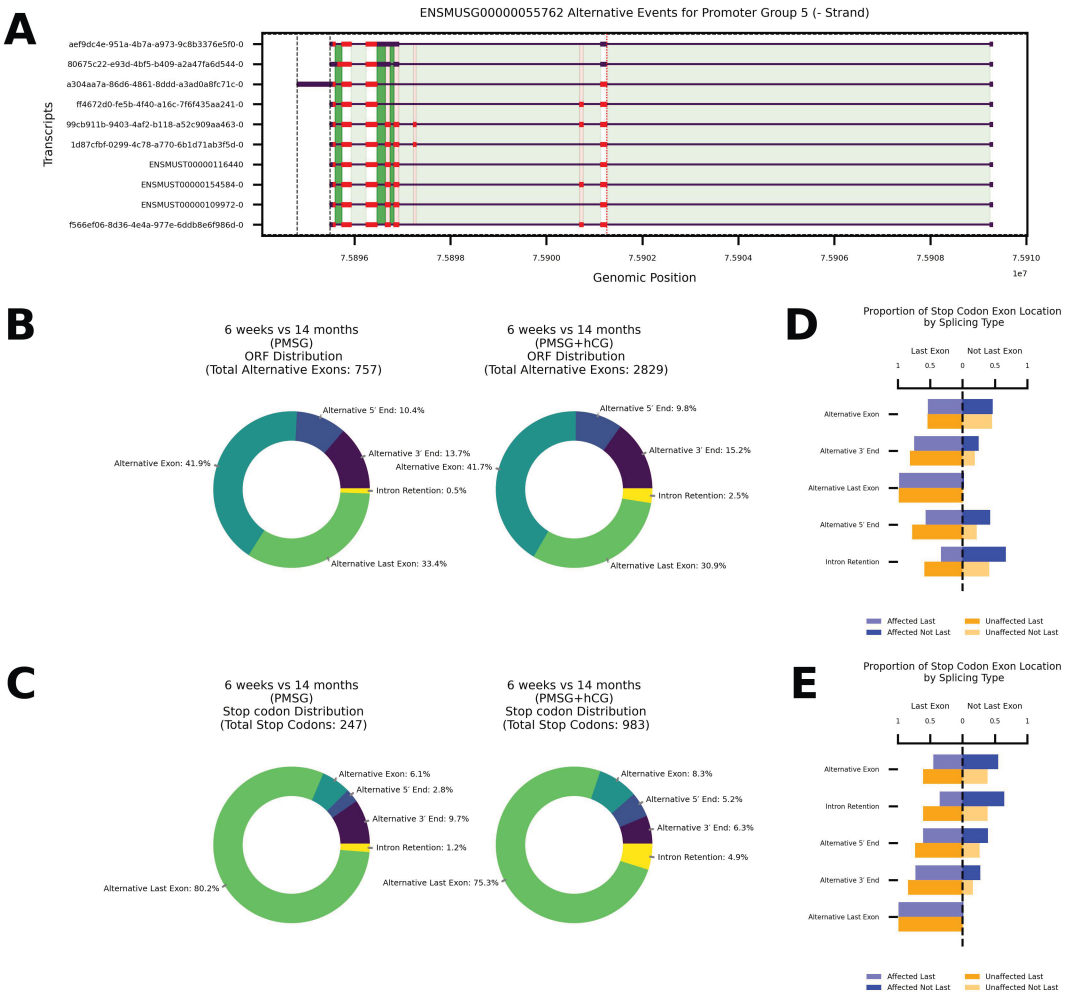

Figure 4. Age-associated changes in alternative splicing alter the protein-coding capacity of the ovary. (A) Example gene (Hscb; ENSMUSG00000043510) that illustrates how alternative splicing affects coding sequence length in 6-week vs. 14-month ovaries. Exons are depicted as black boxes and the translated open reading frame is shown in red. Alternative 5′ splice sites are indicated by orange dashed lines, alternative 3′ splice sites by red dotted lines, and alternative transcript end sites by black dash-dot lines. Alternative exons are shown as semi-transparent red rectangles, introns are highlighted in green, and retained introns are marked in dark green. (B) Pie charts showing the distribution of significantly affected alternative splicing events that alter open reading frame (ORF) length, categorized by event type, 6-week vs. 14-month ovaries after PMSG(left) and PMSG+hCG (right). (D, E) Horizontal bar plots comparing the location of splicing-induced stop codons in transcripts from affected (blue) and unaffected (orange) ovaries under PMSG (D) or PMSG+hCG (E) treatment. Bars represent the percentage of events resulting in a stop codon in the last exon (left side of each panel) versus within exons that are not last (right side).

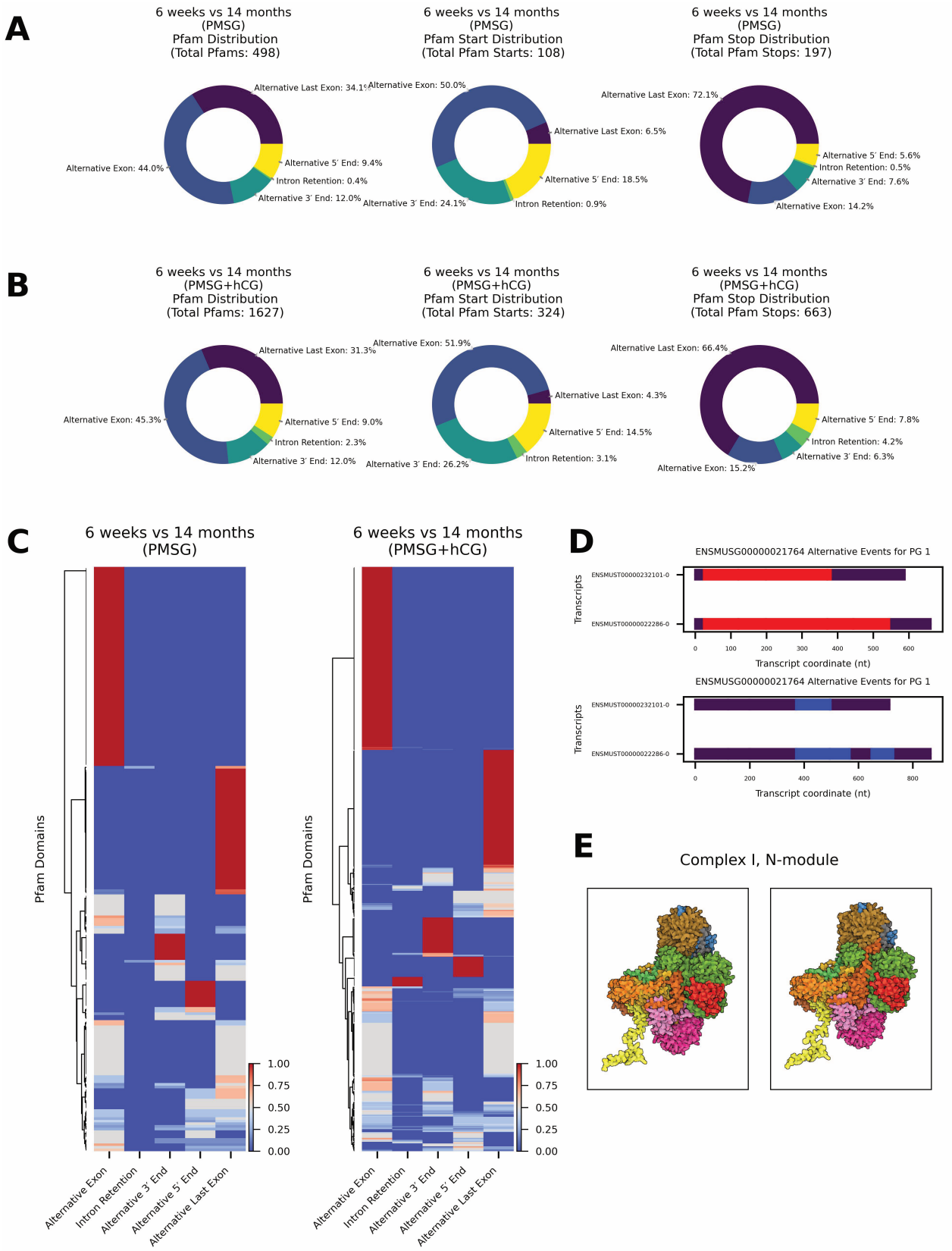

Figure 5. Reproductive aging alters protein domain architecture through splice type–specific remodeling. (A–B) Donut charts showing the distribution of splicing event types associated with Pfam domain gains or losses in significantly affected transcripts of in 6-week vs. 14-month ovaries after PMSG(A) and PMSG+hCG (B). Each panel depicts the proportion of events contributing to overall domain distribution (left),

domain start sites (middle), or domain stop sites (right), categorized by alternative exon, intron retention, alternative 5′ or 3′ splice sites, or alternative last exons. (C) Heatmaps showing the relative enrichment of Pfam domains across splicing event categories in significantly affected transcripts from 6-week vs. 14-month ovaries. Left: PMSG condition. Right: PMSG+hCG condition. Rows represent Pfam domains; columns represent splice event categories. Color scale indicates the normalized fraction of events in which each domain was detected. (D) Transcript architecture of *Ndufs4* gene identified as significantly differentially spliced under PMSG+hCG treatment. Red bars denote the predicted open reading frames. The Pfam domain (blue) is present in the long isoform but absent in the short isoform due to exon skipping. (E) AlphaFold-predicted structures of the mitochondrial Complex I N-module incorporating either the long or short NDUFS4 isoform. Subunits are color-coded: NDUFS4 is shown in dark orange. The short isoform model lacks the C-terminal portion of NDUFS4, while other subunits retain similar spatial arrangement across both models.

Declaration of interest

The authors declare no conflict of interest.

Funding

This work was supported by the National Institutes of Health (R00HD099269 to P.B., R21HD112742 to P.B. and N.S.A., and 1L50HD112918-01 to N.S.A.).

Author contribution statement

Conceptualization, P.B. and N.S.A.; methodology, P.B. and N.S.A.; validation, P.B. and N.S.A.; computational analysis, N.S.A.; investigation, P.B. and N.S.A.; writing—original draft preparation, P.B. and N.S.A.; writing—review and editing, P.B. and N.S.A. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

Not applicable.

References

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J *et al.* (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630** 493–500.

Angarola BL and Anczuków O (2021) Splicing alterations in healthy aging and disease. *Wiley Interdisciplinary Reviews. RNA* **12** e1643.

Bao S, Yin T and Liu S (2024) Ovarian aging: energy metabolism of oocytes. *Journal of Ovarian Research* **17** 118.

Biamonti G, Maita L and Montecucco A (2018) The Krebs cycle connection: Reciprocal influence between alternative splicing programs and cell metabolism. *Frontiers in Oncology* **8** 408.

Brachova P, Alvarez NS, Hong X, Gunewardena S, Vincent KA, Latham KE and Christenson LK (2019) Inosine RNA modifications are enriched at the codon wobble position in mouse oocytes and eggs. *Biology of Reproduction*.

Briley SM, Jasti S, McCracken JM, Hornick JE, Fegley B, Pritchard MT and Duncan FE (2016) Reproductive age-associated fibrosis in the stroma of the mammalian ovary. *Reproduction* **152** 245–260.

Chiang T, Duncan FE, Schindler K, Schultz RM and Lampson MA (2010) Evidence that weakened centromere cohesion is a leading cause of age-related aneuploidy in oocytes. *Current Biology: CB* **20** 1522–1528.

Choi JO, Ham JH and Hwang SS (2022) RNA metabolism in T lymphocytes. *Immune Network* **22** e39.

Donega S, Banskota N, Gupta E, Gonzalez-Freire M, Moore AZ, Ubaida-Mohien C, Munk R, Zukley L, Piao Y, Bergeron C *et al.* (2025) Skeletal muscle mRNA splicing variants association with four different fitness and energetic measures in the GESTALT study. *Journal of Cachexia, Sarcopenia and Muscle* **16** e13603.

Ferrucci L, Wilson DM 3rd, Donegà S and Gorospe M (2022) The energy-splicing resilience axis hypothesis of aging. *Nature Aging* **2** 182–185.

Finley J (2018) Cellular stress and AMPK activation as a common mechanism of action linking the effects of metformin and diverse compounds that alleviate accelerated aging defects in Hutchinson-Gilford progeria syndrome. *Medical Hypotheses* **118** 151–162.

Ju W, Zhao Y, Yu Y, Zhao S, Xiang S and Lian F (2024) Mechanisms of mitochondrial dysfunction in ovarian aging and potential interventions. *Frontiers in Endocrinology* **15** 1361289.

Katz Y, Wang ET, Airoldi EM and Burge CB (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nature Methods* **7** 1009–1015.

Leinonen R, Sugawara H, Shumway M and International Nucleotide Sequence Database Collaboration (2011) The sequence read archive. *Nucleic Acids Research* **39** D19–D21.

Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM and Kendziorski C (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics (Oxford, England)* **29** 2073–2073.

Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. In *bioRxiv*. bioRxiv.

Maracchioni A, Totaro A, Angelini DF, Di Penta A, Bernardi G, Carrì MT and Achsel T (2007) Mitochondrial damage modulates alternative splicing in neuronal cells: implications for neurodegeneration. *Journal of Neurochemistry* **100** 142–153.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ *et al.* (2021) Pfam: The protein families database in 2021. *Nucleic Acids Research* **49** D412–D419.

Robinson MD, McCarthy DJ and Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26** 139–140.

Sasaki H, Hamatani T, Kamijo S, Iwai M, Kobanawa M, Ogawa S, Miyado K and Tanaka M (2019) Impact of oxidative stress on age-associated decline in oocyte developmental competence. *Frontiers in Endocrinology* **10** 811.

Shen S, Park JW, Huang J, Dittmar KA, Lu Z-X, Zhou Q, Carstens RP and Xing Y (2012) MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Research* **40** e61.

Smits MAJ, Schomakers BV, van Weeghel M, Wever EJM, Wüst RCI, Dijk F, Janssens GE, Goddijn M, Mastenbroek S, Houtkooper RH *et al.* (2023) Human ovarian aging is characterized by oxidative damage and mitochondrial dysfunction. *Human Reproduction (Oxford, England)* **38** 2208–2220.

Tang AD, Soulette CM, van Baren MJ, Hart K, Hrabeta-Robinson E, Wu CJ and Brooks AN (2020) Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nature Communications* **11** 1438.

Tiberi S and Robinson MD (2020) BANDITS: Bayesian differential splicing accounting for sample-to-sample

variability and mapping uncertainty. *Genome Biology* **21** 69.

Valencia JD and Hendrix DA (2023) Improving deep models of protein-coding potential with a Fourier-transform architecture and machine translation task. *PLoS Computational Biology* **19** e1011526.

Valsecchi F, Grefte S, Roestenberg P, Joosten-Wagenaars J, Smeitink JAM, Willems PHGM and Koopman WJH (2013) Primary fibroblasts of NDUFS4(-/-) mice display increased ROS levels and aberrant mitochondrial morphology. *Mitochondrion* **13** 436–443.

Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R *et al.* (2020) Ensembl 2020. *Nucleic Acids Research* **48** D682–D688.