Exon Proteogenomic Abundance Predicts Advanced Heart Failure in Titin Cardiomyopathy

Peter Deckerman[1], Jihyeon Lee[1], Aleksandra Binek[1], Sarah Parker[1], Daniel J. Luthringer[1], Michelle M. Kittleson[1], Donald Richards[1], Ian Williamson[1], Daria Ma[1], Lawrence S.C. Czer[1], Nan Jiang[1], Jason Cowan[2], Carl A. Starkey[2], Ray E. Hershberger[2], Jennifer E. Van Eyk[1], Evan P. Kransdorf[1]

1-Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

2-Davis Heart and Lung Research Institute, The Ohio State University, Columbus, OH 43210, USA

Correspondence:

Evan Kransdorf, MD, PhD

8670 Wilshire Blvd., Suite 200

Beverly Hills, CA 90211

Evan.kransdorf@cshs.org

(310) 248-300

Short Title: Proteogenomic Quantification in TTN Cardiomyopathy

Word Count: 5,485

Abstract

Background: Truncating variants in the titin (*TTN*) gene (TTN-TV) are the most common genetic cause of dilated cardiomyopathy (DCM) and confer a significant risk of progression to advanced heart failure (AHF). Disease penetrance of TTN-TV has been linked to the level of expression of the exon containing the TTN-TV, quantified using the percent spliced in (PSI). We hypothesized that recalculating PSI using long-read RNA sequencing and including all 15 *TTN* isoforms would provide more accurate predictions of cardiomyopathy (penetrance) and advanced heart failure [AHF] (expressivity) in patients with TTN cardiomyopathy. Additionally, transcript and protein abundance can be discordant due to post-translational regulation in myocardium which motivated us to compare PSI values to exon-specific TTN peptide abundance.

Methods: We performed long-read RNA sequencing on cardiac tissue from 10 unused organ donors and 15 DCM patients and identified all *TTN* isoforms. We also performed mass spectrometry-based peptide mapping specific for each *TTN* isoform. Exon abundance was quantified using: 1) PSI-DCM-LR-15, a novel PSI metric calculated from long-read RNA sequencing which includes all 15 TTN isoforms and 2) quantile peptide intensity (QPI), a novel quantitative metric reflecting exon-specific peptide abundance. We then assessed the ability of PSI-DCM-LR-15 and QPI to predict AHF in two cohorts of patients with cardiomyopathy due to TTN-TV.

Results: Multiple *TTN* transcript isoforms are expressed in myocardium. PSI-DCM-LR-15 values differed from the original PSI values, especially for the I-band. Proteomic profiling revealed discordance between mRNA and protein-level exon abundance across multiple domains, also highest for the I-band. A hybrid metric, PSI-QPI, combining transcriptional and proteomic exon abundance improved prediction of TTN-TV disease penetrance and expressivity.

Conclusions: A novel hybrid proteogenomic metric, PSI-QPI, that incorporates both transcript and protein abundance more accurately predicts cardiomyopathy (penetrance) and AHF (expressivity) in patients with TTN-TV. This updated tool has direct clinical implications for patient management and suggests that combined proteogenomic strategies may enhance risk stratification for other genetic cardiomyopathies.

Clinical Perspective

What Is New?

- The percent spliced in (PSI) is used to quantify exon-specific mRNA abundance. This study introduces a novel PSI metric calculated using long-read RNA sequencing, PSI-DCM-LR-15, that includes all 15 titin (*TTN*) isoforms. PSI-DCM-LR-15 revealed differences from the original PSI calculations, particularly in the I-band of TTN.

- Proteomic analysis demonstrated significant discordance between mRNA and protein-level exon abundance. To better quantify exon abundance at the protein level, a novel metric termed quantile peptide intensity (QPI) was developed.

- The authors combined PSI-DCM-LR-15 and QPI into a novel hybrid proteogenomic metric, PSI-QPI, that combined transcriptional and proteomic exon abundance. This metric improves the prediction of TTN truncating variant (TTN-TV) penetrance and expressivity.

What are the clinical implications?

- Compared to the original PSI values, the PSI-QPI metric more accurately predicts which persons with TTN-TVs will develop cardiomyopathy (penetrance) and which patients with cardiomyopathy will progress to advanced heart failure (expressivity).

- PSI-QPI can be combined with other clinical risk factors to assist clinicians in the management of patients with TTN-TVs.

- The successful application of a combined proteogenomic strategy in this study suggests that similar approaches could be used to enhance risk stratification for genes with a high frequency of truncating variants, such as desmoplakin (*DSP*) and filamin C (*FLNC*).

Introduction

Truncating variants in the titin (*TTN*) gene, comprising stop gain, frameshift, or splice donor/acceptor mutations, are the most common genetic cause of dilated cardiomyopathy (DCM), being identified in 10%-20% of patients [1]. Cardiomyopathy due to *TTN* truncating variants (CM-TTN-TV) is among the most responsive to heart failure medical therapy, with left ventricular (LV) reverse remodeling observed in 50-60% of cases [2-4]. Paradoxically, however, TTN-TVs are also the most frequently identified genetic cause of end-stage heart failure requiring heart transplantation (HT) [5].

The metatranscript of the *TTN* gene has 363 coding exons which in turn encode the 35,991 amino acid protein that connects the M-line to Z-line of the sarcomere [6] and contributes to the passive tension of the cardiomyocyte [7]. *TTN* is transcribed into 15 isoforms (as per Gencode), with the predominantly expressed isoforms in adult cardiomyocytes being N2BA and N2B [7,8]. However, the shorter isoforms, Novex-3 [9] and Cronos [10], also have important functions in both fetal and adult cardiomyocytes. The function of the other 11 transcriptional isoforms remains unclear.

TTN-TVs can be identified in up to 1-3% of the population [8,11,12]. Prior studies have shown that patients with DCM carry a preponderance of TTN-TVs located in the A-band [8,13]. The mRNA expression levels of transcribed *TTN* exons within 3 main isoforms (*TTN* N2BA, N2B, and Novex-3) have been used to calculate the "percent spliced in" (PSI), which quantifies the fraction of mRNA transcripts that include a particular exon. Patients with TTN-TVs affecting exons with a PSI > 0.9 (included in >90% of transcripts) are more likely to develop DCM [8]. As a result, the American College of Medical Genetics and Genomics (ACMG) criteria for variant

interpretation in DCM originally required pathogenic TTN-TV to originate in the A-band [14]. These criteria for pathogenicity were revised to include variants affecting exons outside the A-band with high PSI (PSI > 0.9) [1] because there is an association between DCM and TTN-TVs in high PSI exons within the Z-disk, I-band, and M-band. Interestingly, DCM is also associated with TTN-TV in a limited number of low PSI (0.15 to 0.48) exons within the I-band [15]. Furthermore, there are numerous A-band exons with PSI > 0.9 that have not exhibited significant enrichment in DCM [15]. These findings suggest that A-band restrictive views of TTN-related DCM may be both scientifically and clinically limiting.

The development of DCM in patients carrying TTN-TV, *i.e.* penetrance, differs from the severity of DCM that develops, *i.e.* expressivity [16,17]. The ability to predict which patients with CM due to TTN-TV (penetrance) will have high expressivity and develop advanced heart failure (AHF) requiring specialized therapies including durable left ventricular assist device (LVAD) implantation or HT represents a critical but unmet clinical need. We hypothesized that PSI would predict the risk of AHF in patients with CM-TTN-TV. However, given that PSI values were derived using short-read RNA sequencing and including only 3 of the 15 isoforms, we further hypothesized that recalculating PSI using long-read RNA sequencing and including all 15 *TTN* isoforms would provide more accurate predictions of penetrance and expressivity. Additionally, transcript and protein abundance can be discordant in myocardium [18], so we were motivated to develop a method to compare PSI values to exon-specific TTN protein abundance using quantitative peptide mass spectrometry, a value we termed “quantile peptide intensity” (QPI). Finally, we proposed that a hybrid metric, PSI-QPI, combining transcriptional and proteomic

exon abundance would improve prediction of the development of TTN cardiomyopathy (penetrance) and its severity (expressivity).

Methods

Cardiac Tissue for RNA, Spatial Transcriptomics, and Proteomic Studies

This study was performed under protocols approved by the Cedars-Sinai Institutional Review Board (Pro00010979, STUDY00001617, STUDY00004096) and OneLegacy. All tissue donors provided consent to participate. We utilized a total of 43 samples from 35 heart explants: 25 samples for RNA sequencing, 4 for spatial transcriptomics, and 14 for proteomics. Samples for RNA sequencing were derived from donors from 2 groups: 10 organ donors not utilized for HT (group 1) and 15 patients undergoing HT for end-stage DCM (group 2). For spatial transcriptomics, samples were derived from 1 organ donor not utilized for HT and 3 patients undergoing HT for end-stage DCM (group 3). For proteomics, samples were derived from 14 patients undergoing HT for end-stage DCM (group 4). Characteristics and genetic testing results for all tissue donors is provided in Supplemental Tables 1 and 2, respectively.

Genetic Testing of Tissue Donors and Cedars-Sinai Titin Cardiomyopathy Cohort

Genetic testing was performed using commercial vendors or via whole exome sequencing as part of a biobank study as previously described [19]. All genes classified by ClinGen as having a definitive, strong, or moderate association for DCM were tested (Supplemental Table 4) [20].

Variant consequence and genomic position were determined using Ensembl Variant Effect Predictor [21] based on the *TTN* metatranscript (ENST00000589042.5).

Cedars-Sinai Titin Cardiomyopathy Cohort

We identified 115 patients from 104 families that presented to the Advanced Heart Disease or Cardiogenetics clinics at Cedars-Sinai Medical Center between 2015 and 2025 for cardiomyopathy due to a TTN-TV (CM-TTN-TV). In addition, we identified patients through search of the Cedars-Sinai electronic medical record via Deep6 AI [22]. We defined CM-TTN-TV as: 1) left ventricular ejection fraction (LVEF) < 50% by echocardiography, with or without LV dilation (as defined below), and 2) the presence of a likely pathogenic (LP) or pathogenic (P) variant in the *TTN* gene. Genetic testing results for the Cedars-Sinai CM-TTN-TV cohort are provided in Supplemental Table 3. AHF was defined as the need for LVAD implantation or HT. In patients who did not have AHF and who presented with an initial LVEF < 40%, LV reverse remodeling was defined as improvement of EF from first to last available echocardiogram of ≥ 10% or achievement of an ejection fraction of ≥ 40% [23]. We excluded 3 of 115 patients due to the presence of an additional cause of cardiomyopathy: 1 with a P variant in *RBM20*, 1 with critical coronary artery disease, and 1 with cardiac amyloidosis.

Dilated Cardiomyopathy Precision Medicine Cohort

The design [24], patient characteristics and genetics [1] of the Dilated Cardiomyopathy Precision Medicine (DCM-PM) study have been previously reported. The Institutional Review Boards at the Ohio State University and all clinical sites approved the initial study, followed by single

institutional review board oversight at the University of Pennsylvania. All participants gave written informed consent. Due to differences in variant interpretation between the Cedars-Sinai Titin Cardiomyopathy cohort and DCM-PM cohort (specifically the requirement for TV-containing exons to have a PSI > 0.9 to meet criteria for LP/P in DCM-PM) we defined CM-TTN-TV for this cohort as: 1) left ventricular ejection fraction (LVEF) < 50% by echocardiography, 2) left ventricular dilation as ≥ 95th percentile for sex/height, and 3) the presence of a variant of undetermined significance (VUS), LP, or P variant in the *TTN* gene. AHF was again defined as the need for LVAD implantation or HT. We excluded 2 patients due to the presence of an additional cause of cardiomyopathy: 1 with a LP variant in *FLNC* and 1 with a LP variant in *TCAP*.

Long-read RNA Sequencing and Bioinformatics

RNA was isolated from explanted cardiac tissue using Zymo RNA extraction kit (Zymo Research). RNA purification was performed using columns (Qiagen). Library preparation was performed using the Kinnex full-length RNA kit (PacBio). Parameters of RNA libraries are detailed in Supplementary Table 5. Sequencing was performed using a SMRT cell on a Revio system (Novogene). The resulting BAM files were mapped to the transcriptome (Gencode version 48) [25], modified to include the *TTN* isoform Cronos which originates at a transcriptional start site in the intron between exon 239 and 240 of the *TTN* metatranscript [10]. Transcripts for all *TTN* isoforms were quantified using Oarfish [26,27]. We assessed for novel *TTN* isoforms using "Bambu" [28]. Counts were compared between samples using DESeq2 [29].

Short-read RNA Sequencing

RNA was isolated and purified as above. Library preparation was performed using xGen Broad-Range RNA Library Prep kit (IDT). Parameters of RNA libraries are detailed in Supplementary Table 5. RNA sequencing was performed using NovaSeq 6000 (Illumina).

Spatial Transcriptomics

Formalin-fixed paraffin-embedded slides from 4 heart explants were obtained: 1 from an organ donor not utilized for HT, 1 from a patient undergoing HT for end-stage DCM due to a *DSP* variant, and 2 patients undergoing HT for end-stage DCM due to a TTN-TV. The Xenium Human Immuno-Oncology Profiling Panel was utilized with the addition of custom probes for the *TTN* isoforms N2BA, Novex-3, and metatranscript (designed to span a unique exon-exon junction), *DSP*, and the housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). Custom probe sequences are provided in Supplemental Table 6. Probe hybridization, ligation, and rolling circle amplification were performed, following the manufacturer's protocol. Imaging and signal decoding were done using the Xenium Analyzer instrument (10x Genomics). Each slide's Xenium output was exported as a "spatialexperiment" object and normalized using SpaNorm [30]. Normalized counts for each slide were rescaled to between 0 and 1000 and plotted using the R package Voyager [31].

Calculation of PSI

We tabulated exon-exon junctions for all *TTN* isoforms in Gencode version 48 with the addition of Cronos (Supplemental Table 7). Junction sequences 20 base pairs in length (10 bp from each

exon) were counted in a sample-wise fashion using Bioconductor [32] Biostrings [33]. PSI was calculated for the *TTN* metatranscript per the original definition [34]:

$$PSI_{exonN} = \frac{\sum IR_{exonN}}{\sum(IR_{exonN} + ER_{exonN})}$$

where N=exon number, IR=inclusion reads, ER=exclusion reads. The median PSI value for each exon among all samples was utilized.

Mass Spectrometry-based Proteomics and Development of QPI Metric

14 samples underwent proteomic mass spectrometry (MS) sample processing as previously reported [35]. Briefly, protein was isolated from explanted cardiac tissue by the "in sequence" method (add reference) *i.e.*: each tissue was fractionated into cytosolic-, myofilament-, and insoluble-enriched protein fractions, and trypsin digested on S-Trap columns (ProtiFi). Peptides from each patient sample were analyzed in triplicate by liquid chromatography (LC)-MS/MS in data independent acquisition (DIA) mode on an Ultimate 3000 NanoLC connected to an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific) equipped with an EasySpray ion source. Data acquisition was completed blinded to experimental groups. LC-MS/MS method was adapted and modified from previously published workflow [36]. Peptides were separated on a C18 column (15 cm with 300 µm ID, 3 µm Omega Polar C18 beads, and 100 Å pore size, Phenomenex) over the course of total 90 minutes at a flow rate of 8µL/min. Mobile phase A was 0.1% v/v formic acid in water and mobile phase B was 90% ACN, 0.1% formic acid in water. Initial loading condition was 1% B for 3 min and then 3% for another 2 min at a flow rate of 8 µL/min to allow for equilibration and for samples to reach the column, followed by a linear gradient of 3–30% B over 70 min, 30–40% B over 10 min and 40–45% B over 5 min. Each

MS1 scan was followed by 150 DIA MS2 scans [37] (4Da size precursor window). Peptide identification and quantification was performed using DIA-NN [38].

DIA-NN and Quantile Peptide Intensity Analyses

The whole proteome sequence, excluding the canonical *TTN* sequence and supplemented with human *TTN* isoform FASTA entries, was used to generate the spectral library. *In silico* digestion was performed using cleavage at K and R. A deep-learning model was applied to generate an in silico spectral library from the resulting peptide list. The fragment m/z range was set to 200–1800. Peptide length was restricted to 7–30 amino acids. The precursor m/z range was set to 300–1800, and precursor charge states were limited to 1–4. Cysteine carbamidomethylation was included as a fixed modification, and N-terminal methionine was excluded. A maximum of one missed cleavage was allowed. Quantification used a fixed-width center of each elution peak. DIA-NN automatically optimized mass accuracy using the first run of the experiment. RT-dependent cross-run normalization was applied, and results were filtered at a 0.01 FDR threshold. For each sample peptides were mapped to the corresponding exon of the *TTN* metatranscript. Only proteotypic peptides, peptides, that are unique to a single exon, were used and peptides mapping to more than one exon were discarded. Peptide intensities were summed per exon, the median exon-specific value among all 14 samples was then ranked as a quantile within the context of the full protein. We termed this exon-specific metric "quantile peptide intensity (QPI)."

Codon Usage Bias Analysis

During translation, specific aminoacyl-tRNAs are selected via a degenerate, triplet genetic code. The preferential selection of specific synonymous codons that encode the same amino acid is termed codon usage bias. We utilized the codon adaptation index (CAI) [39] to assess codon usage bias across the *TTN* gene. CAI is a numerical score between 0.0 and 1.0 that quantifies how closely a gene's sequence matches the preferred codons across the human proteome. The average CAI value across the human proteome is 0.8 [40], and also varies among protein domains [41]. CAI was calculated using "cubar" [42].

Statistical Analysis

All analysis was performed using R [43]. PSI values were compared using symmetric mean absolute percentage error (sMAPE) [44]; correlation coefficient cannot be used since all PSI values of the A-band are equal to 1. sMAPE is given by the formula:

$$sMAPE = \frac{100\%}{n} \sum_{e=1}^{n} \frac{|PSI_x - PSI_y|}{|PSI_x| + |PSI_y|}$$

Where PSI_x and PSI_y represent the values of two different PSI metrics for each exon and n represents the total number of exons. sMAPE values range between 0 and 100%. Univariate logistic modeling with AHF as the response variable was performed using a Gaussian generalized linear model. Due to a scarcity of endpoints for women and patients of black race, these observations were weighted w=2, whereas all other endpoints were weighted w=1. Multivariable logistic modeling with AHF as the response variable was performed using "glmnet" [45] with 10-fold cross-validation and α=0.5 (elastic-net). Coefficients were determined at the minimum λ. Prediction performance was assessed using leave-one-out cross validation.

Prediction performance was assessed using "pROC" [46] and "PRROC" [47]. Figures were made using "ggplot2" [48].

Results

TTN Transcript Isoform Quantification

We quantified *TTN* mRNA transcript isoforms using heart tissue collected from 25 donors: 10 organ donors not utilized for HT, 8 patients with end-stage DCM without a TTN-TV (non-TTN-DCM) undergoing HT, and 7 patients with end-stage DCM due to a TTN-TV (TTN-DCM) undergoing HT. Among the 7 patients with non-TTN-DCM, 5 had disease-causing variants in other genes associated with DCM, specifically *DMD*, *DSP*, *FLNC*, *LMNA*, and *RBM20* (Supplemental Table 2). In the 7 patients with TTN-DCM, all *TTN* variants were located in exons with PSI of 1.0 (Supplemental Table 2).

To assess *TTN* transcript isoform expression levels, we quantified long-read sequencing-derived RNA expression levels of all 15 *TTN* transcript isoforms curated in Gencode, as well as Cronos, which was manually curated, with long-read RNA sequencing. We did not identify any novel *TTN* transcript isoforms. The top 5 most highly expressed isoforms were N2BA, N2B, TTN-208, TTN-215, and Novex-3 (Table 1). The *TTN* metatranscript, which is the default *TTN* transcript as defined by the Matched Annotation from NCBI and EMBL-EBI (MANE) project, was in the lower half of *TTN* transcript isoforms by expression levels. Given that this transcript is used for variant interpretation, it is notable that it is transcribed at low levels. Cronos was the transcript with the lowest expression.

We compared *TTN* mRNA transcript isoform expression levels between unused donor hearts (controls) and heart explants with dilated cardiomyopathy, both non-TTN-DCM and TTN-DCM. We found that the expression of 4 *TTN* isoforms differed significantly between the groups: TTN-206, metatranscript, and TTN-209 were higher in DCM, whereas TTN-211 was lower in controls (Table 1). There was also a trend for higher expression of N2B in DCM (p-value = 0.054). We also compared *TTN* mRNA transcript isoform expression levels between non-TTN-DCM and TTN-DCM heart explants and found that TTN-205 was significantly lower in TTN-DCM than in non-TTN-DCM (Supplemental Table 8). Taken together, these results show: (1) multiple *TTN* transcript isoforms are expressed in humans, (2) mRNA isoform expression is mostly similar in DCM as compared to normal tissue with exception of 4 isoforms, and (3) mRNA isoform expression is generally similar in non-TTN-DCM as compared to TTN-DCM.

Spatial Transcriptomics

We used spatial transcriptomics to validate the *TTN* isoform expression levels we measured using long-read RNA sequencing. We performed spatial transcriptomics using the Xenium platform on heart tissue collected from 4 donors: 1 organ donor, 1 patient that underwent HT for DCM due to a *DSP* disease-causing variant, and 2 patients that underwent HT for DCM due to TTN-CM. We developed isoform-specific probes for highly expressed *TTN* isoforms that contain unique splice junctions: N2BA, Novex-3, and metatranscript. For all 4 donors, we found that the count ratio of *TTN* Novex-3 to N2BA and meta-transcript to N2BA were similar between those derived from long-read RNA sequencing and spatial transcriptomics

(Figure 1 and Supplemental Table 9). These results validate the accuracy of our *TTN* isoform quantification obtained by long-read RNA sequencing.

Calculation of PSI Metrics

We quantified exon-exon junction expression for all junctions present in *TTN* transcript isoforms in Gencode, calculated exon-specific PSI (based on exons present in the *TTN* metatranscript) for each sample, and found the median value across samples. We hypothesized that a PSI metric calculated using long-read RNA sequencing and including all 15 *TTN* isoforms, which we termed PSI-DCM-LR-15, would provide more accurate predictions of penetrance and expressivity. We calculated 3 additional versions of PSI to determine potential sources of variation: 1) PSI-DCM-SR-3, calculated from DCM samples, quantified using short-read RNA sequencing, and considering 3 *TTN* isoforms (*TTN* N2BA, N2B, Novex-3), 2) PSI-DCM-LR-3, calculated from DCM samples, quantified using long-read RNA sequencing, and considering the same 3 *TTN* isoforms, and 3) PSI-CNTL-LR-15, calculated from control samples, quantified using long-read RNA sequencing, and considering 15 *TTN* isoforms. We also compared these PSI values to those calculated by Roberts et al. (PSI-Original) [8].

We compared these 5 PSI metrics for each exon within the 4 domains of the TTN protein: Z-disk, I-band, A-band, and M-band. For the Z-disk, A-band, and M-band, the 5 PSI metrics were similar as assessed using sMAPE (Figure 2B and Supplemental Table 10). In contrast, for the I-band, the 5 PSI metrics differed substantially. The highest similarity was between PSI-DCM-LR-3/PSI-DCM-LR-15 (sMAPE 15.5%), PSI-DCM-LR-15/PSI-DCM-CNTL-15 (sMAPE 22.9%), and PSI-DCM-SR-3/PSI-DCM-LR-3 (sMAPE 23.8%), suggesting that neither the

number of *TTN* isoforms, use of DCM versus control samples, or short-read versus long-read RNA sequencing substantially changed PSI values. In contrast, PSI-Original exhibited the highest dissimilarity from other PSI values, including PSI-DCM-SR-3/PSI-Original (sMAPE 44.7%) which should exhibit high similarity since both metrics used DCM samples, short-read RNA sequencing and considered the 3 most prevalent *TTN* isoforms. These results suggest that the discrepancies between our PSI metrics and PSI-Original are due to intrinsic differences in the method of calculation rather than differences in sample origin, number of isoforms considered, or mRNA sequencing technology.

Mass Spectrometry-based TTN Peptide Mapping

To analyze *TTN* isoform expression at the protein level, we performed mass spectrometry-based proteomics on explanted cardiac tissue from 14 patients undergoing HT: 10 for DCM not due to a TTN-TV and 4 for DCM due to a TTN-TV. We obtained excellent proteome coverage and TTN and observed 2,490 unique TTN proteotypic peptides, represented by 3,580 total precursors comprised of redundant TTN peptides with various charge states and/or post-translational modifications (Supplemental Data Set 1). First, we sought to identify which *TTN* isoforms were represented by the 2,490 unique proteotypic TTN peptides. Of the 16 *TTN* transcript isoforms, 15 yield unique amino acid sequences (Supplemental Table 7), as TTN-211 does not contain an open reading frame and thus does not generate a unique isoform. Of these 15 isoforms, 9 isoforms generate peptides that are unique to one specific isoform. The majority (96%, 2,397/2,490 peptides) of the *TTN* peptides were not unique to a specific isoform and mapped to two or more of 6 identified isoforms. Thus, overall 4% (93/2,490 peptides) were

specific to a single isoform: 92 mapped uniquely to Novex-3, and 1 peptide mapped uniquely to TTN-215.

We assessed the abundance of peptides translated from *TTN* metatranscript exons 2 to 363 based on 3,424 peptides that mapped to a unique site within the *TTN* metatranscript protein product. There were 1/27 exons in the Z-disk (4%), 105/223 exons in the I-band (47%), 1/106 exons in the A-band (1%), and 0/5 exons in the M-band for which we did not identify any peptides. We quantified peptide abundance by summing the peptide intensities for all peptides derived from an exon, finding the median exon-specific peptide intensity sum among the 14 samples, and then ranked each exon's sum intensity as a quantile (considering all exons in the *TTN* metatranscript). We termed this exon-specific metric "quantile peptide intensity (QPI)." We found that QPI varied substantially between the domains of *TTN*: the median QPI was 0.55, 0.03, 0.73, and 0.39 for the Z-disk, I-band, A-band, and M-band, respectively (Figure 3A). We then compared QPI to PSI-DCM-LR-15 for each domain of TTN (Figure 3B and Supplemental Table 11) and found that the 2 metrics were most similar for the A-band (sMAPE 19.1%) and were less similar for the Z-disk (sMAPE 30.5%), I-band (sMAPE 66.1%), and M-band (sMAPE 40.1%).

Taken together, these protein-level data suggest several findings: (1) the non-canonical *TTN* isoforms Novex-3 and TTN-215 are present *in vivo* in patients with DCM with AHF and (2) exon abundance was similar at the mRNA and protein levels for the A-band but discordant for the other 3 domains of TTN.

Codon Usage Bias Across the TTN Gene

We postulated that codon usage bias could lead to the observed discordance between exon abundance at the mRNA and protein levels across the *TTN* gene. We used the codon adaptation index (CAI) to quantify codon usage bias. We calculated CAI across the *TTN* gene and found that CAI was low at the 5' end and increased towards the 3' end of the mRNA transcript (Figure 4A). The minimum CAI value of 0.77 was found in exon 2 of the Z-disc and the maximum CAI value of 0.83 was found in exon 326 of the A-band. The mean CAI values per domain were 0.77, 0.79, 0.81 and 0.79 for the Z-disk, I-band, A-band, and M-band, respectively, which were significantly different (Figure 4B).

Predictors of Penetrance in TTN Cardiomyopathy

We sought to determine the potential of exon abundance as measured using PSI and QPI to predict penetrance of TTN-TV. We utilized data from Vatta et al. [15] who identified in 2,446 TTN-TV cardiomyopathy/arrhythmia cases referred for genetic testing and 922 TTN-TV in gnomAD v3.1. We performed case-control analysis using a logistic model. We found that PSI-DCM-LR-15 (odds ratio [OR] 1.23 per 0.1 increase, 95% confidence intervals [95%CI] 1.20-1.26), PSI-Original (OR 1.24, 95%CI 1.21-1.27), and QPI (OR 1.20, 95%CI 1.18-1.23) were all associated with an increased odds of an exon containing cases (Supplemental Table 12). Given the discordance between RNA and protein levels for the Z-disk, I-band, and M-band, we hypothesized that a hybrid metric comprised of QPI for these 3 domains plus PSI-DCM-LR-15 for the A-band would be a better predictor of penetrance. This hybrid metric, which we termed PSI-QPI, was associated with an increased odds of an exon containing cases (OR 1.21, 95%CI 1.18-1.23) and provided the best performance in terms of area under the receiver operating

characteristic curve (AUROC) and area under the precision recall curve (AUPRC). The values of PSI-QPI are provided in Supplemental Table 13.

Predictors of Advanced Heart Failure in TTN Cardiomyopathy

We next determined the clinical relevance of the PSI-QPI metric for 112 patients in the Cedars-Sinai TTN Cardiomyopathy Cohort. All patients had been treated with guideline-directed medical therapy. During the study period 39 (35%) patients developed AHF, as defined by the need for implantation of LVAD or HT. Characteristics of CM-TTN-TV patients without and with AHF are provided in Table 2. Significant differences between the non-AHF and AHF groups included higher systolic blood pressure (118 vs. 106 mmHg, $p<0.001$) , lower frequency of ventricular tachycardia (28% vs. 68%, $p<0.001$), higher ejection fraction (EF; 28% vs. 20%, $p=0.015$) and smaller left ventricular internal diastolic dimension (LVIDD) at presentation (58 vs 65 mm, $p=0.007$). At their last follow up echocardiogram at a median of 2.8 years later (interquartile range 1.0 - 6.4 years) the EF was higher (46% vs. 16%, $p<0.001$) and LVIDD was lower (54 vs. 67 mm, $p<0.001$) in the non-AHF group. There were significantly more patients in the non-AHF group that experienced left ventricular reverse remodeling (52% vs. 3%, $p<0.001$). Genetic characteristics of the TTN-TV also differed between the groups with non-AHF patients possessing I-band and Z-disk variants more frequently (25% vs 8% for I-band, 7% vs. 0% for Z-disk) and A-band variants less frequently (69% vs. 92%) than AHF patients ($p=0.015$).

We performed logistic modeling with AHF as the endpoint and 11 potential predictor variables (Supplemental Table 14), including 4 PSI/QPI metrics. We found that increasing age of CM onset, female sex, and a non-A-band location for the TTN-TV were associated with a

decreased risk of AHF, while black race, diabetes mellitus, hypertension, and frameshift variant consequence were associated with an increased risk of AHF. With regards to the PSI/QPI metrics, we found that PSI-DCM-LR-15, QPI, and PSI-QPI were associated with a significantly increased risk of AHF while PSI-Original was not associated with a significantly increased risk of AHF (Table 3 and Supplemental Table 14). Model coefficients for PSI-Original were uninterpretable due to multicollinearity, as all patients with AHF had PSI-Original values of 1.

Table 3 lists the performance of these 4 PSI/QPI metrics for predicting AHF (Table 3). We found that PSI-QPI had the best performance (highest AUROC and AUPRC) for prediction of AHF as compared to PSI-DCM-LR-15, PSI-Original, and QPI. The AUROC of PSI-QPI was significantly higher than that of PSI-Original (p=0.0479), but not PSI-DCM-LR-15 (p=0.623) or QPI (p=0.219). PSI-QPI also provided significant reclassification, with a net reclassification index (continuous) of 0.211 (95% CI 0.072-0.349, p=0.003). We also performed multivariable logistic modeling with AHF as the endpoint and 7 potential predictor variables with 1 model containing PSI-QPI and 1 model PSI-Original. PSI-QPI was retained in the first model (Figure 5A) with a OR of 1.22 per 0.1 increase in PSI-QPI and an overall AUROC in leave-one-out cross-validation of 0.667. PSI-Original was retained in the second model (Figure 5B) with a OR of 1.33 per 0.1 increase in PSI-Original and an overall AUROC in leave-one-out cross-validation of 0.664.

We then validated the clinical relevance of the PSI-QPI metric in an independent cohort of 180 patients with CM-TTN-TV from the DCM-PM Study. We compared the characteristics of the DCM-PM cohort to those of the Cedars-Sinai Titin Cardiomyopathy cohort (Supplementary Table 15). The two groups were similar for frequency of AHF (35% vs. 39%, p=0.504), women (29% vs. 36%, p=0.268), hypertension (39% vs. 41%, p=0.852), diabetes mellitus (20% vs. 22%,

p=0.791), and in the domain of the TTN protein containing the TTN-TV (p=0.180). Significant differences between the cohorts included an older median age of CM onset in the CS group (48 vs. 42 years, p=0.010), a more diverse racial composition (p<0.001), and a differing profile of TTN-TV consequences (p=0.001), specifically characterized by higher rates of frameshift (47% vs. 41%) and stop-gain (44% vs. 34%) mutations alongside lower rates of canonical splice variants (9% vs. 16%).

With regards to the PSI/QPI metrics, we found that only PSI-QPI was associated with a significantly increased risk of AHF, while PSI-DCM-LR-15, QPI, and PSI-Original were not associated with a significantly increased risk of AHF (Table 3). We found that QPI and PSI-QPI had the best performance (highest AUROC and AUPRC) for prediction of AHF as compared to PSI-DCM-LR-15 and PSI-Original (Table 3).

Discussion

In this study, we developed PSI-QPI, a novel quantitative metric for penetrance and expressivity of CM-TTN-TV. We first investigated *TTN* mRNA transcript isoform levels using long-read RNA sequencing in explanted cardiac tissue. We found that multiple *TTN* transcript isoforms are expressed in humans, isoform expression is mostly similar in DCM as compared to normal tissue with exception of 4 isoforms, and isoform expression is generally similar in non-CM-TTN-TV as compared to CM-TTN-TV. We developed 4 versions of PSI and found that substantial differences exist between our novel PSI metrics and the original PSI [8]. We performed mass spectrometry-based proteomics on explanted cardiac tissue and found that exon abundance was similar at the mRNA and protein levels for the A-band but discordant for

the other 3 domains of *TTN*. Given this discordance, we developed a new quantitative metric for exon abundance at the protein level we termed QPI. A hybrid metric comprised of PSI and QPI provided the best performance for prediction of both penetrance and AHF and led to significant reclassification of risk as compared to the original PSI.

Our results have immediate clinical implications. Given that 1-3% of the general population carry TTN-TV [8,11,12] and the 10-20% prevalence of TTN-TV in patients with DCM [1], genetic testing laboratories frequently incorporate PSI values into the reporting of genetic results for *TTN* to provide the ordering clinician with context for interpretation of the results. The performance of PSI-QPI as compared to PSI-Original for both penetrance (Supplemental Table 12) and expressivity (Table 3) indicates that PSI-QPI may be a superior metric for predicting the penetrance and expressivity of TTN-TV.

While PSI-QPI predicted AHF in both the Cedars-Sinai TTN Cardiomyopathy and DCM-PM cohorts, its discriminative performance was better in the Cedars-Sinai cohort (Table 3). Based on a higher frequency of AHF, albeit not statistically significant, and a younger age of onset (Supplementary Table 15), we postulate that the DCM-PM cohort had a higher propensity to develop AHF, thereby making PSI-QPI less predictive in that cohort. Nevertheless, PSI-Original was not significantly associated with AHF in the DCM-PM cohort, whereas PSI-QPI was significantly associated with AHF (Table 3). Our results show that exon abundance, as quantified by PSI-QPI is one factor among several that predict AHF (Supplementary Table 14 and Figure 5). We envision the incorporation of PSI-QPI in conjunction with other clinical and genetic factors into a score to assist clinicians in the prediction of CM in persons with TTN-TV (penetrance) and AHF in patients with CM-TTN-TV (expressivity).

Roberts et al. [8] established that exon utilization, as quantified using PSI, serves as a predictor of the penetrance and expressivity of TTN-TV. Here we have extended these findings from the transcriptional level to the protein level, as we find that our metric of exon-specific peptide abundance, QPI, also predicts the penetrance and expressivity of TTN-TV. Consideration of isoform mRNA expression has previously been shown to improve prediction of variant expressivity [49]. We postulate that combined proteogenomic metrics, as developed in this study, could help improve penetrance/expressivity for other genes such as desmoplakin (*DSP*) and filamin C (*FLNC*) that are also associated with cardiomyopathy with a high frequency of truncating variants.

We observed discordance between exon abundance at the mRNA and protein levels within the Z-disk, I-band, and M-band, whereas exon abundance was concordant within the A-band. The preferential selection of specific synonymous codons that encode the same amino acid is termed codon usage bias. We observed lower usage of optimal codons (as quantified by CAI values) for exons in the Z-disk, I-band, and M-band but higher usage of optimal codons in the A-band. Lower usage of optimal codons is associated with slower translation and increased nonsense mediated decay (NMD) [50]. Our findings provide a potential explanation for the observation of truncated TTN proteins within the sarcomere [51-53], *i.e.* faster translation within the A-band, mediated by higher usage of optimal codons, could decrease NMD of mutated TTN proteins thereby permitting their incorporation into the sarcomere. Intriguingly, we found the maximum CAI value to be located in exon 326 of the TTN mRNA transcript and TTN-TV in exon 327 have been associated with decreased NMD [54]. Further investigation utilizing ribosome profiling will be needed to confirm this observation.

This study has strengths and weaknesses. Many studies of *TTN* mRNA expression have relied on animal models. By having examined explanted heart tissue from human organ donors and patients with CM-TTN-TV, a strength of this work is that these observations offer direct clinical translation. Nevertheless, as only a small number of samples were analyzed, the PSI values need to be replicated, even though very similar PSI values between DCM and control samples were observed, suggesting that larger number of samples may not be necessary. The two cohorts of patients with CM-TTN-TV were derived from AHF programs and represented a more severely affected group of patients that may be most relevant for enhanced risk assessment. Nevertheless, these results need to be validated in patients with a greater range of CM-TTN-TV risk.

In conclusion, a novel hybrid proteogenomic metric, PSI-QPI, that incorporated both transcript and protein measures of exon abundance more accurately predicted cardiomyopathy (penetrance) and AHF (expressivity) in patients harboring TTN-TV. This updated tool has direct clinical implications for patient management and suggests that combined proteogenomic strategies may enhance risk stratification for other genetic cardiomyopathies.

Acknowledgements/Funding

This work was funded by seed money support from the Smidt Discovery Fund of the Smidt Heart Institute (E.K.), the American Heart Association (AHA) Postdoctoral Fellowship, AHA Award Number: 829444 (A.B.) and National Institutes of Health (NIH) 1R01HL155346-01A1 (J.V.E.) and 1R01HL144509-01 (J.V.E.). J.V.E. is the Erika Glazer Endowed Chair in Women's Heart Health. We would also like to acknowledge access to mass spectrometers via the Cedars-

Sinai Medical Center Proteomics and Metabolomics Core. Research reported in this publication was also supported by a parent award from the National Heart, Lung, and Blood Institute of the NIH under Aware Number R01HL128857 to R.E.H., which included a supplement from the National Human Genome Research Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Disclosures

The authors report no relevant financial relationships or conflicts of interest to disclose

References

1. Jordan E, Kinnamon DD, Haas GJ, Hofmeyer M, Kransdorf E, Ewald GA, Morris AA, Owens A, Lowes B, Stoller D, et al. Genetic Architecture of Dilated Cardiomyopathy in Individuals of African and European Ancestry. *JAMA*. 2023;330:432–441. doi: 10.1001/jama.2023.11970
2. Jansweijer JA, Nieuwhof K, Russo F, Hoorntje ET, Jongbloed JD, Lekanne Deprez RH, Postma AV, Bronk M, van Rijsingen IA, de Haij S, et al. Truncating titin mutations are associated with a mild and treatable form of dilated cardiomyopathy. *Eur J Heart Fail*. 2017;19:512–521. doi: 10.1002/ejhf.673
3. Verdonschot JAJ, Hazebroek MR, Wang P, Sanders-van Wijk S, Merken JJ, Adriaansen YA, van den Wijngaard A, Krapels IPC, Brunner-La Rocca HP, Brunner HG, et al. Clinical Phenotype and Genotype Associations With Improvement in Left Ventricular Function in Dilated Cardiomyopathy. *Circ Heart Fail*. 2018;11:e005220. doi: 10.1161/CIRCHEARTFAILURE.118.005220
4. Escobar-Lopez L, Ochoa JP, Mirelis JG, Espinosa MA, Navarro M, Gallego-Delgado M, Barriales-Villa R, Robles-Mezcua A, Basurte-Elorz MT, Gutierrez Garcia-Moreno L, et al. Association of Genetic Variants With Outcomes in Patients With Nonischemic Dilated Cardiomyopathy. *J Am Coll Cardiol*. 2021;78:1682–1699. doi: 10.1016/j.jacc.2021.08.039
5. Kim Y, Gunnarsdottir OB, Viveiros A, Reichart D, Quiat D, Willcox JAL, Zhang H, Chen H, Curran JJ, Kim DH, et al. Genetic Contribution to End-Stage Cardiomyopathy Requiring Heart Transplantation. *Circ Genom Precis Med*. 2023;16:452–461. doi: 10.1161/CIRCGEN.123.004062
6. Furst DO, Osborn M, Nave R, Weber K. The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of

ten nonrepetitive epitopes starting at the Z line extends close to the M line. *J Cell Biol*. 1988;106:1563–1572. doi: 10.1083/jcb.106.5.1563

7. van der Pijl R, Nusayr E, Strom J, Slater R, Gohlke J, Hourani Z, Saripalli C, Kolb J, Hermanson K, Brynnel O, et al. Importance of N2BA Titin in Maintaining Cardiac Homeostasis and Its Role in Dilated Cardiomyopathy. *Circ Heart Fail*. 2025;18:e012083. doi: 10.1161/CIRCHEARTFAILURE.124.012083
8. Roberts AM, Ware JS, Herman DS, Schafer S, Baksi J, Bick AG, Buchan RJ, Walsh R, John S, Wilkinson S, et al. Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci Transl Med*. 2015;7:270ra276. doi: 10.1126/scitranslmed.3010134
9. Hashimoto K, Ohira M, Kodama A, Kimoto M, Inoue M, Tone S, Usui Y, Hanashima A, Goto T, Ogura Y, et al. Loss of connectin novex-3 leads to heart dysfunction associated with impaired cardiomyocyte proliferation and abnormal nuclear mechanics. *Sci Rep*. 2024;14:13727. doi: 10.1038/s41598-024-64608-1
10. Zou J, Tran D, Baalbaki M, Tang LF, Poon A, Pelonero A, Titus EW, Yuan C, Shi C, Patchava S, et al. An internal promoter underlies the difference in disease severity between N- and C-terminal truncation mutations of Titin in zebrafish. *Elife*. 2015;4:e09406. doi: 10.7554/eLife.09406
11. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. *Circ Cardiovasc Genet*. 2012;5:391–399. doi: 10.1161/CIRCGENETICS.112.962928
12. Akinrinade O, Koskenvuo JW, Alastalo TP. Prevalence of Titin Truncating Variants in General Population. *PLoS One*. 2015;10:e0145284. doi: 10.1371/journal.pone.0145284
13. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, Conner L, DePalma SR, McDonough B, Sparks E, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med*. 2012;366:619–628. doi: 10.1056/NEJMoa1110186
14. Morales A, Kinnamon DD, Jordan E, Platt J, Vatta M, Dorschner MO, Starkey CA, Mead JO, Ai T, Burke W, et al. Variant Interpretation for Dilated Cardiomyopathy: Refinement of the American College of Medical Genetics and Genomics/ClinGen Guidelines for the DCM Precision Medicine Study. *Circ Genom Precis Med*. 2020;13:e002480. doi: 10.1161/CIRCGEN.119.002480
15. Vatta M, Regalado E, Parfenov M, Swartzlander D, Nagl A, Mannello M, Lewis R, Clemens D, Garcia J, Ellsworth RE, et al. Analysis of TTN Truncating Variants in >74 000 Cases Reveals New Clinically Relevant Gene Regions. *Circ Genom Precis Med*. 2025;18:e004982. doi: 10.1161/CIRCGEN.124.004982
16. Shah RA, Asatryan B, Sharaf Dabbagh G, Aung N, Khanji MY, Lopes LR, van Duijvenboden S, Holmes A, Muser D, Landstrom AP, et al. Frequency, Penetrance, and Variable Expressivity of Dilated Cardiomyopathy-Associated Putative Pathogenic Gene Variants in UK Biobank Participants. *Circulation*. 2022;146:110–124. doi: 10.1161/CIRCULATIONAHA.121.058143
17. Massier M, de Groote P, Donal E, Magnin-Poull I, Coubes C, Le Guillou Horn X, Rooryck C, Reant P, Troadec Y, Brehin AC, et al. Exploring the Familial Phenotypic Variability Associated With TTN Truncating Variants in Cardiomyopathies: Variant Spectrum,

Genotype-Phenotype Correlation and Consequences in Genetic Counseling. *Clin Genet*. 2025;107:425–433. doi: 10.1111/cge.14679
18. Jani VP, Yoo EJ, Binek A, Guo A, Kim JS, Aguilan J, Keykhaei M, Jenkin SR, Sidoli S, Sharma K, et al. Myocardial Proteome in Human Heart Failure With Preserved Ejection Fraction. *J Am Heart Assoc*. 2025;14:e038945. doi: 10.1161/JAHA.124.038945
19. Cao L, Rushakoff J, Williamson I, Karlstaedt A, Kittleson M, Czer L, Kransdorf EP. Similar burden of rare genetic variants in ischemic and non-ischemic dilated cardiomyopathy. *Front Cardiovasc Med*. 2025;12:1542653. doi: 10.3389/fcvm.2025.1542653
20. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, Celeghin R, Edwards M, Fan J, Ingles J, et al. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. *Circulation*. 2021;144:7–19. doi: 10.1161/CIRCULATIONAHA.120.053033
21. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17:122. doi: 10.1186/s13059-016-0974-4
22. Deep 6 AI. https://deep6.ai/.
23. Wilcox JE, Fang JC, Margulies KB, Mann DL. Heart Failure With Recovered Left Ventricular Ejection Fraction: JACC Scientific Expert Panel. *J Am Coll Cardiol*. 2020;76:719–734. doi: 10.1016/j.jacc.2020.05.075
24. Kinnamon DD, Morales A, Bowen DJ, Burke W, Hershberger RE, Consortium* DCM. Toward Genetics-Driven Early Intervention in Dilated Cardiomyopathy: Design and Implementation of the DCM Precision Medicine Study. *Circ Cardiovasc Genet*. 2017;10. doi: 10.1161/CIRCGENETICS.117.001826
25. Mudge JM, Carbonell-Sala S, Diekhans M, Martinez JG, Hunt T, Jungreis I, Loveland JE, Arnan C, Barnes I, Bennett R, et al. GENCODE 2025: reference gene annotation for human and mouse. *Nucleic Acids Res*. 2025;53:D966–D975. doi: 10.1093/nar/gkae1078
26. Wissel D, Mehlferber MM, Nguyen KM, Pavelko V, Tseng E, Robinson MD, Sheynkman GM. A Systematic Benchmark of High-Accuracy PacBio Long-Read RNA Sequencing for Transcript-Level Quantification. *bioRxiv*. 2025. doi: 10.1101/2025.05.30.656561
27. Zare Jousheghani Z, Singh NP, Patro R. Oarfish: enhanced probabilistic modeling leads to improved accuracy in long read transcriptome quantification. *Bioinformatics*. 2025;41:i304–i313. doi: 10.1093/bioinformatics/btaf240
28. Chen Y, Sim A, Wan YK, Yeo K, Lee JJX, Ling MH, Love MI, Goke J. Context-aware transcript quantification from long-read RNA-seq data with Bambu. *Nat Methods*. 2023;20:1187–1195. doi: 10.1038/s41592-023-01908-w
29. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. doi: 10.1186/s13059-014-0550-8
30. Salim A, Bhuva DD, Chen C, Tan CW, Yang P, Davis MJ, Yang JYH. SpaNorm: spatially-aware normalization for spatial transcriptomics data. *Genome Biol*. 2025;26:109. doi: 10.1186/s13059-025-03565-y
31. Moses L, Einarsson PH, Jackson K, Luebbert L, Booeshaghi AS, Antonsson S, Bray N, Melsted P, Pachter L. Voyager: exploratory single-cell genomics data analysis with geospatial statistics. *bioRxiv*. 2023. doi: 10.1101/2023.07.20.549945

32. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12:115–121. doi: 10.1038/nmeth.3252
33. H P. Biostrings.
34. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008;456:470–476. doi: 10.1038/nature07509
35. Meddeb M, Koleini N, Binek A, Keykhaei M, Darehgazani R, Kwon S, Aboaf C, Margulies KB, Bedi KC, Jr., Lehar M, et al. Myocardial ultrastructure of human heart failure with preserved ejection fraction. *Nat Cardiovasc Res*. 2024;3:907–914. doi: 10.1038/s44161-024-00516-x
36. Mc Ardle A, Binek A, Moradian A, Chazarin Orgel B, Rivas A, Washington KE, Phebus C, Manalo DM, Go J, Venkatraman V, et al. Standardized Workflow for Precise Mid- and High-Throughput Proteomics of Blood Biofluids. *Clin Chem*. 2022;68:450–460. doi: 10.1093/clinchem/hvab202
37. Robinson AE, Binek A, Venkatraman V, Searle BC, Holewinski RJ, Rosenberger G, Parker SJ, Basisty N, Xie X, Lund PJ, et al. Lysine and Arginine Protein Post-translational Modifications by Enhanced DIA Libraries: Quantification in Murine Liver Disease. *J Proteome Res*. 2020;19:4163–4178. doi: 10.1021/acs.jproteome.0c00685
38. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17:41–44. doi: 10.1038/s41592-019-0638-x
39. Sharp PM, Li WH. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res*. 1987;15:1281–1295. doi: 10.1093/nar/15.3.1281
40. Arshad M, Uchmanowicz M, Rana V, Trost B, Scherer SW, Rafiq MA. Mapping the inter- and intra-genic codon-usage landscape in Homo sapiens. *NAR Genom Bioinform*. 2026;8:lqag024. doi: 10.1093/nargab/lqag024
41. Homma K, Noguchi T, Fukuchi S. Codon usage is less optimized in eukaryotic gene segments encoding intrinsically disordered regions than in those encoding structural domains. *Nucleic Acids Res*. 2016;44:10051–10061. doi: 10.1093/nar/gkw899
42. Liu M, Zi B, Zhang H, Zhang H. Cubar: a versatile package for codon usage bias analysis in R. *Genetics*. 2026;232. doi: 10.1093/genetics/iyaf191
43. Computing RFfS.
44. Shahid A, Kim J, Byon SS, Hong S, Lee I, Lee BD. An end-to-end pipeline for automated scoliosis diagnosis with standardized clinical reporting using SNOMED CT. *Sci Rep*. 2025;15:17274. doi: 10.1038/s41598-025-01952-w
45. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33:1–22.
46. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77. doi: 10.1186/1471-2105-12-77

47. Grau J, Grosse I, Keilwagen J. PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. *Bioinformatics*. 2015;31:2595–2597. doi: 10.1093/bioinformatics/btv153
48. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer International Publishing; 2016.
49. Cummings BB, Karczewski KJ, Kosmicki JA, Seaby EG, Watts NA, Singer-Berk M, Mudge JM, Karjalainen J, Satterstrom FK, O'Donnell-Luria AH, et al. Transcript expression-aware annotation improves rare variant interpretation. *Nature*. 2020;581:452–458. doi: 10.1038/s41586-020-2329-2
50. Barrington CL, Galindo G, Koch AL, Horton ER, Morrison EJ, Tisa S, Stasevich TJ, Rissland OS. Synonymous codon usage regulates translation initiation. *Cell Rep*. 2023;42:113413. doi: 10.1016/j.celrep.2023.113413
51. McAfee Q, Chen CY, Yang Y, Caporizzo MA, Morley M, Babu A, Jeong S, Brandimarto J, Bedi KC, Jr., Flam E, et al. Truncated titin proteins in dilated cardiomyopathy. *Sci Transl Med*. 2021;13:eabd7287. doi: 10.1126/scitranslmed.abd7287
52. Fomin A, Gartner A, Cyganek L, Tiburcy M, Tuleta I, Wellers L, Folsche L, Hobbach AJ, von Frieling-Salewsky M, Unger A, et al. Truncated titin proteins and titin haploinsufficiency are targets for functional recovery in human cardiomyopathy due to TTN mutations. *Sci Transl Med*. 2021;13:eabd3079. doi: 10.1126/scitranslmed.abd3079
53. McAfee Q, Caporizzo MA, Uchida K, Bedi KC, Jr., Margulies KB, Arany Z, Prosser BL. Truncated titin protein in dilated cardiomyopathy incorporates into the sarcomere and transmits force. *J Clin Invest*. 2024;134. doi: 10.1172/JCI170196
54. Kim YG, Ha C, Shin S, Park JH, Jang JH, Kim JW. Enrichment of titin-truncating variants in exon 327 in dilated cardiomyopathy and its relevance to reduced nonsense-mediated mRNA decay efficiency. *Front Genet*. 2022;13:1087359. doi: 10.3389/fgene.2022.1087359

Figure Legends

Figure 1: Spatial transcriptomics using explanted heart tissue obtained from an unused organ donor, patient with DCM due to a *DSP* truncating variant, and two patients with DCM due to TTN-TV. Using probes specific for the N2BA, Novex-3, and metatranscript *TTN* isoforms, we found that the ratio of Novex-3:N2BA and metatranscript:N2BA was similar between all four donors and yielded similar values to those measured using long-read RNA sequencing.

Figure 2: (A) Plot of PSI-DCM-LR-15 values against exon number in the TTN metatranscript. (B) Heatmap of symmetric mean absolute percent error (sMAPE) values between percent spliced in (PSI) metrics for each domain of the titin protein.

Figure 3: (A) Plot of quantitative peptide intensity (QPI) values against exon number in the TTN metatranscript. (B) Plot of PSI-DCM-LR-15 values (grey), QPI values (blue), and their percent difference (orange) for each exon in the *TTN* metatranscript. The gray, blue, and orange lines represent the fitted trend estimated by a generalized additive model for PSI-DCM-LR-15, QPI, and their percent difference, respectively. *TTN* domains are indicated on the bottom of the plot: blue Z-line, purple I-band, green A-band, red M-band.

Figure 4: (A) Plot of the codon adaptation index every 90 base pairs along the *TTN* metatranscript mRNA (grey dots). The blue line represents the fitted codon adaptation index trend estimated by a generalized additive model. The red line represents the average codon adaptation index across the human genome (0.8). *TTN* domains are indicated on the bottom of the plot: blue Z-line, purple I-band, green A-band, red M-band. (B) Violin plot of the fitted codon adaptation index estimated by a generalized additive model for each exon of the Z-disk

(blue), I-band (purple), A-band (green), and M-band (red) domains. The means for each domain were significantly different (p-value by ANOVA < 0.001).

Figure 5: (A) Forest plot of odds ratios with their 95% confidence intervals from a multivariable model for advanced heart failure including PSI-QPI. (B) Forest plot of odds ratios with their 95% confidence intervals from a multivariable model for advanced heart failure including PSI-Original.

Tables

Table 1: Comparison of *TTN* transcript isoform expression between unused donor hearts (control) and dilated cardiomyopathy.

Transcript	Control n=10	DCM N=15	Log2 Fold Change	P-value
TTN-212/Metatranscript	365	1938	2.40	0.008
TTN-209	102	479	2.39	0.008
TTN-206	1354	3416	1.69	0.01
TTN-207	549	1011	0.99	0.062
TTN-205	276	397	0.73	0.194
Cronos	8	31	0.68	0.389
TTN-210/N2B	6752	11234	0.67	0.054
TTN-204/Novex-3	3210	3408	0.16	0.458
TTN-202/N2A	1617	1546	0.05	0.977
TTN-213/N2BA	67712	59179	-0.02	0.977
TTN-201/Novex-2	84	73	-0.02	0.977
TTN-208	5646	4520	-0.36	0.458
TTN-215	3291	2691	-0.41	0.207
TTN-211	193	26	-1.78	0.008

Notes: 1: Expression levels are provided in median normalized counts. 2: Counts compared using the negative binomial Wald test. 3. Transcript isoforms are ordered by log2 fold change.

Abbreviations: DCM = dilated cardiomyopathy.

Table 2: Characteristics, echocardiographic and genetic parameters of patients in the Cedars-Sinai TTN Cardiomyopathy cohort based on whether they developed advanced heart failure.

Characteristic	CM-TTN-TV Patients Without AHF n=73	CM-TTN-TV Patients With AHF n=39	P-value
Age of CM onset	50.7 [37.7;59.8]	43.8 [36.6;51.5]	0.048
Sex: female	25 (34.2%)	7 (17.9%)	0.110
Race			0.077
Asian	8 (11.0%)	3 (7.69%)	
Black	5 (6.85%)	9 (23.1%)	
Other	5 (6.85%)	4 (10.3%)	
White	55 (75.3%)	23 (59.0%)	
HTN	24 (32.9%)	20 (51.3%)	0.090
DM	11 (15.1%)	11 (28.2%)	0.156
SBP at presentation	118 [108;129]	106 [96.0;116]	0.001
Family history of CM	30 (52.6%)	23 (67.6%)	0.236
Atrial fibrillation	30 (43.5%)	22 (57.9%)	0.220
Ventricular tachycardia	19 (27.5%)	26 (68.4%)	<0.001
Echocardiographic Parameters			
Echo at presentation			
EF	27.5 [18.0;41.0]	20.0 [15.0;28.8]	0.015
LVIDD	57.9 (8.01)	64.9 (11.9)	0.007
Echo at last follow up			
EF	46.0 [40.0;54.0]	15.5 [14.2;20.0]	<0.001
LVIDD	54.0 [49.0;58.0]	67.0 [60.0;71.5]	<0.001
Change in EF	10.5 [5.00;25.0]	-3.00 [-9.75;0.75]	<0.001
LVRR	38 (52.1%)	1 (2.63%)	<0.001
Genetic parameters			
Location of TTN-TV			0.015
A-band	50 (68.5%)	36 (92.3%)	
I-band	18 (24.7%)	3 (7.69%)	
Z-disk	5 (6.85%)	0 (0.00%)	
TTN-TV consequence			
frameshift	31 (42.5%)	22 (56.4%)	0.147
stop-gain	33 (45.2%)	16 (41.0%)	
canonical splice	9 (12.3%)	1 (2.56%)	

Abbreviations: AHF: advanced heart failure; BP: blood pressure; CM: cardiomyopathy; DM: diabetes mellitus; EF: ejection fraction; HTN: hypertension; LVEDD: left ventricular end-diastolic diameter; LVRR: left ventricular reverse remodeling; SBP: systolic blood pressure; TTN-TV: titin truncating variant.

Table 3: Performance of PSI and QPI metrics for prediction of advanced heart failure in the Cedars-Sinai Titin Cardiomyopathy cohort and the Dilated Cardiomyopathy-Precision Medicine cohort.

Metric	Odds Ratio (95% CI)	p-value	AUROC	AUPRC
Cedars-Sinai Titin Cardiomyopathy				
PSI-DCM-LR-15	1.54 (1.11-2.53)	0.049	0.584	0.391
PSI-Original	46.02	0.984	0.575	0.386
QPI	1.17 (1.02-1.36)	0.034	0.586	0.390
PSI-QPI	1.41 (1.14-1.92)	0.007	0.623	0.417
Dilated Cardiomyopathy-Precision Medicine				
PSI-DCM-LR-15	1.12 (0.96-1.37)	0.189	0.529	0.409
PSI-Original	1.11 (0.93-1.39)	0.285	0.518	0.403
QPI	1.05 (0.96-1.14)	0.281	0.559	0.444
PSI-QPI	1.15 (1.03-1.30)	0.019	0.557	0.426

Abbreviations: AUROC: area under the receiver operating characteristic curve; AUPRC: area under precision-recall curve; PSI: percent spliced in; QPI: quantile peptide intensity.

Notes: Odds ratios for PSI and QPI metrics are given per 0.1 increase.

Figures

Figure 1

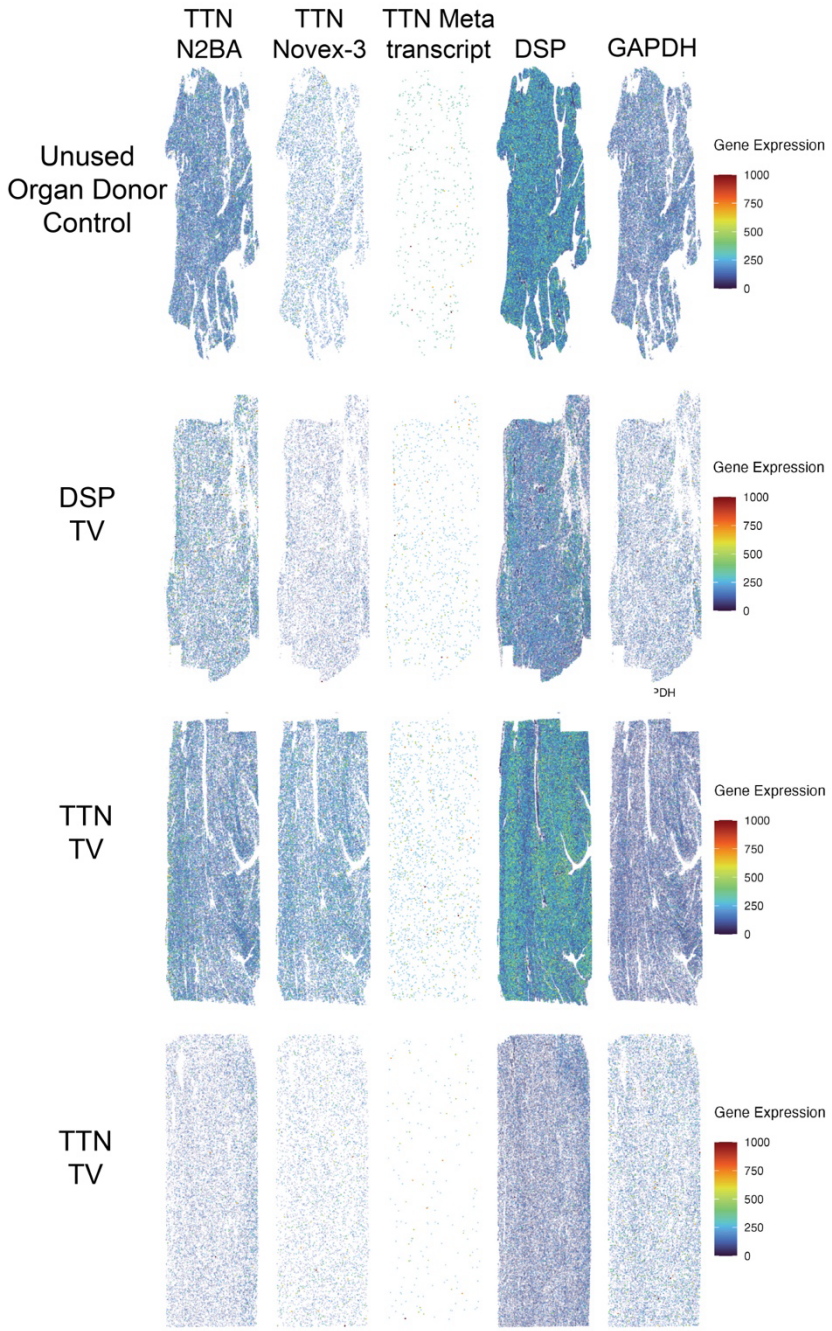
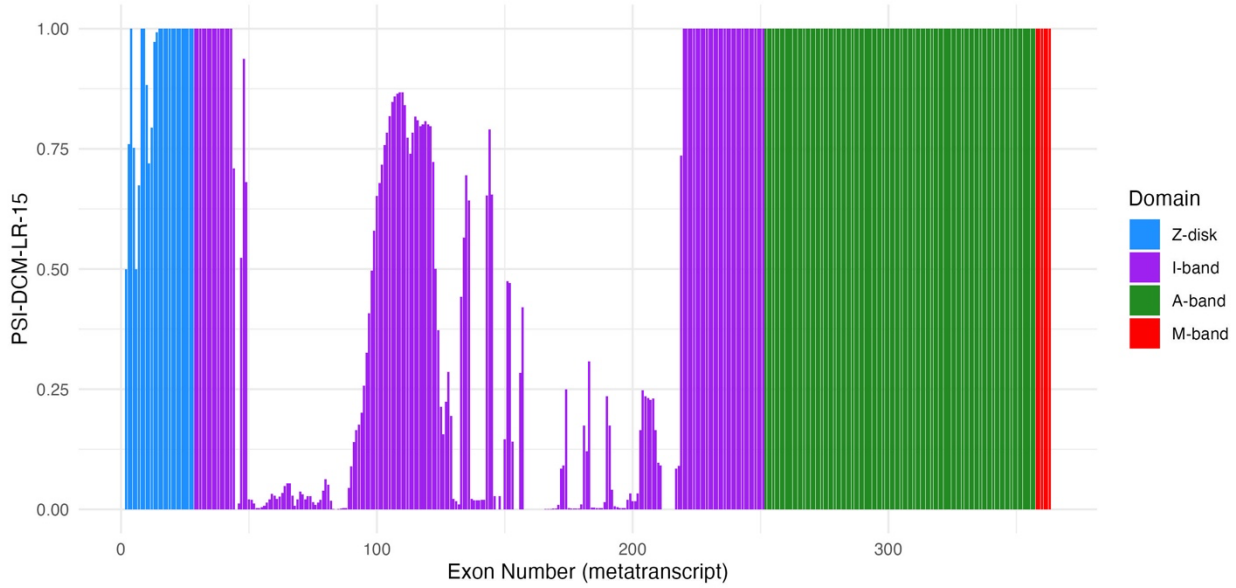


Figure 2

A


1.00
0.75
0.50
0.25
0.00
PSI-DCM-LR-15
0
100
200
300
Exon Number (metatranscript)
Domain
Z-disk
I-band
A-band
M-band

B

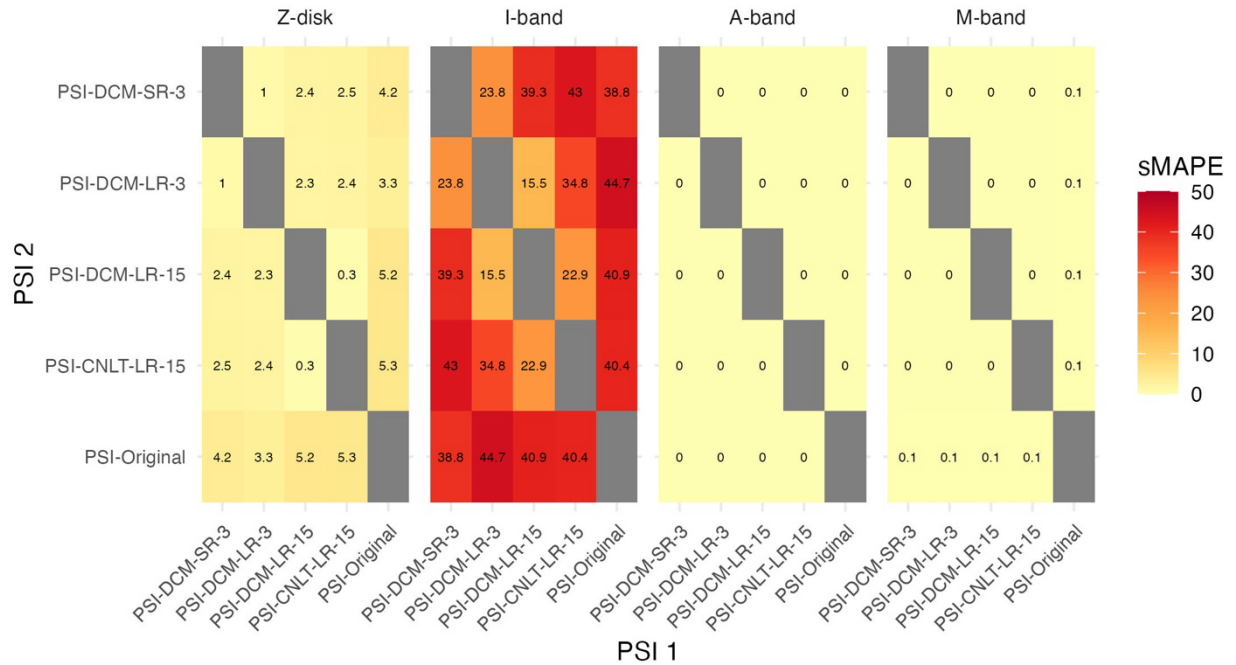
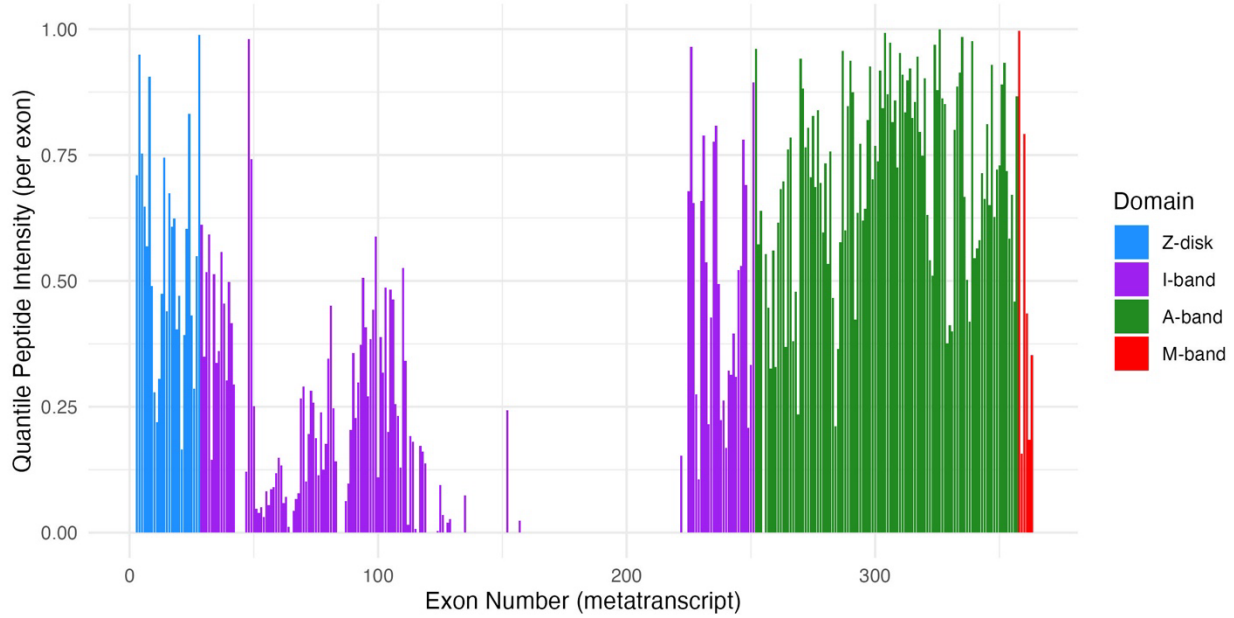

Z-disk
I-band
A-band
M-band
PSI-DCM-SR-3
PSI-DCM-LR-3
PSI-DCM-LR-15
PSI-CNLT-LR-15
PSI-Original
PSI 2
PSI 1
sMAPE
50
40
30
20
10
0

Figure 3

A


1.00
0.75
0.50
0.25
0.00
Quantile Peptide Intensity (per exon)
0
100
200
300
Exon Number (metatranscript)
Domain
Z-disk
I-band
A-band
M-band

B

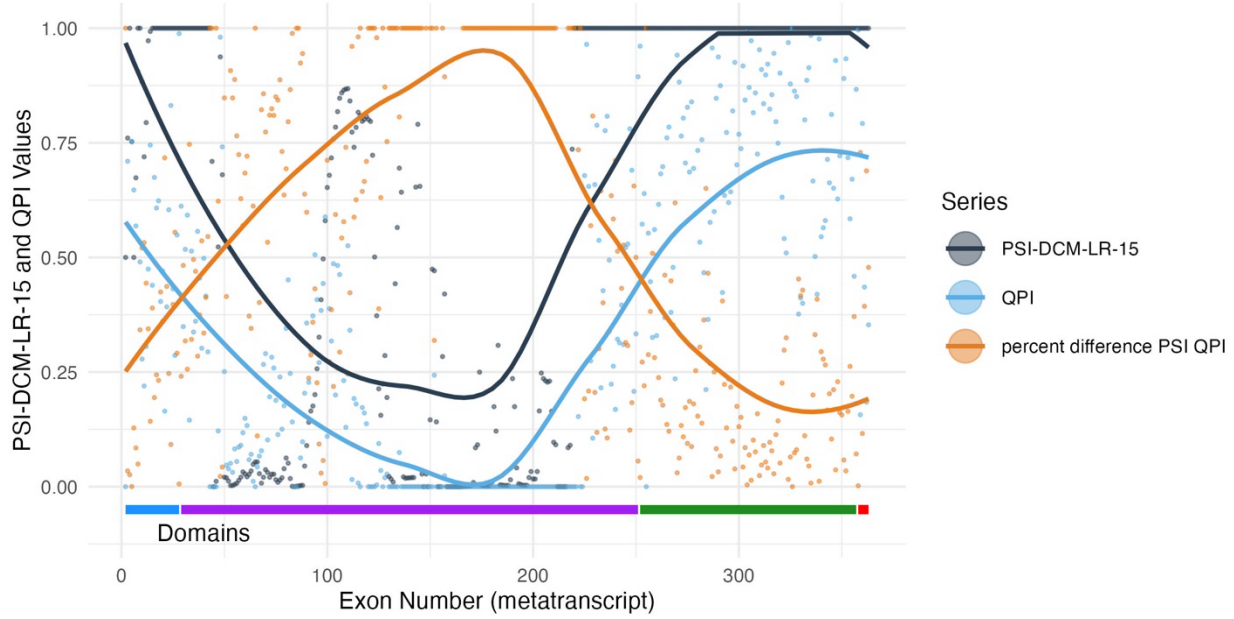
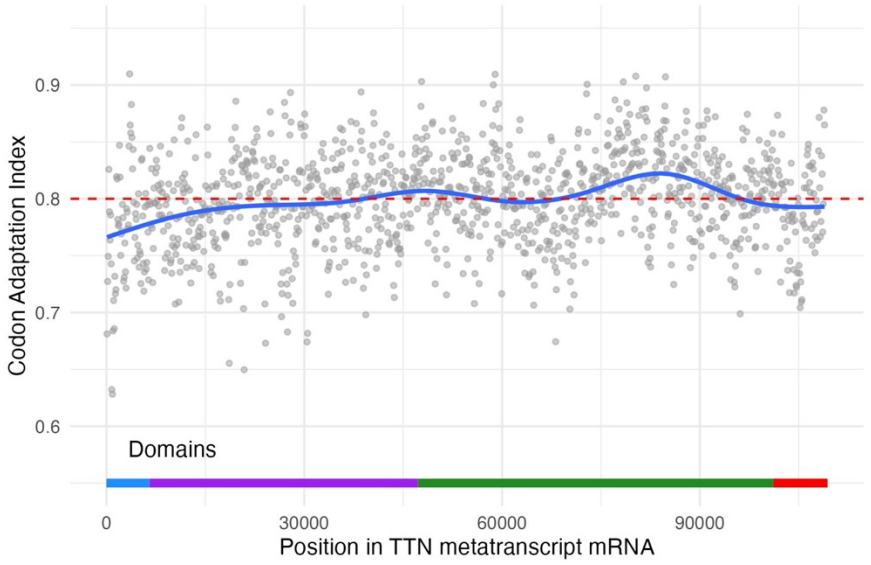

1.00
0.75
0.50
0.25
0.00
PSI-DCM-LR-15 and QPI Values
Domains
0
100
200
300
Exon Number (metatranscript)
Series
PSI-DCM-LR-15
QPI
percent difference PSI QPI

Figure 4

A


Codon Adaptation Index
0.9
0.8
0.7
0.6
Domains
0
30000
60000
90000
Position in TTN metatranscript mRNA

B

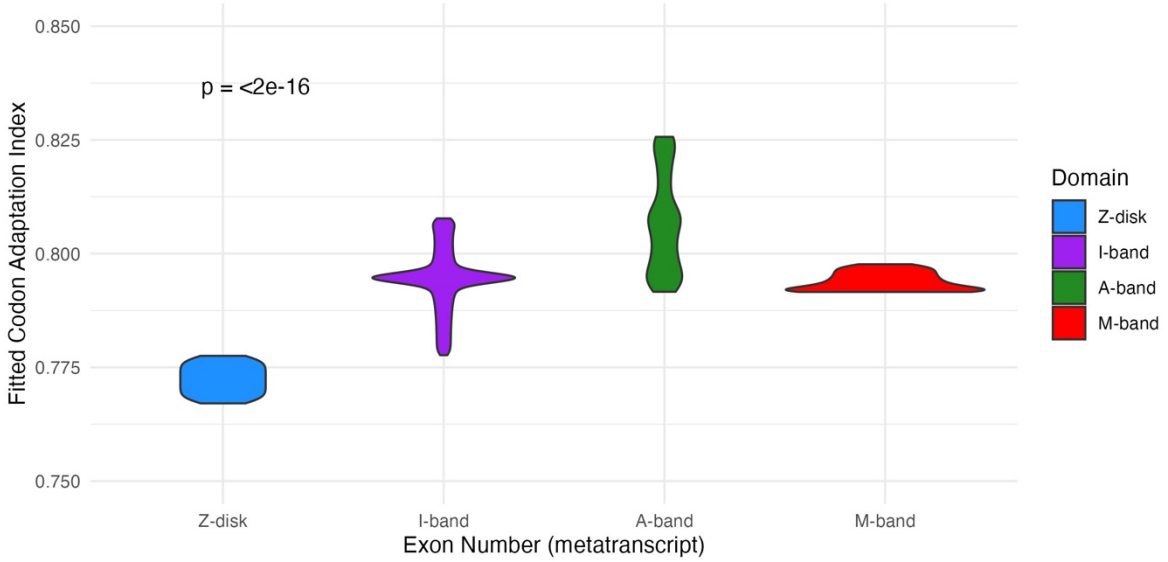
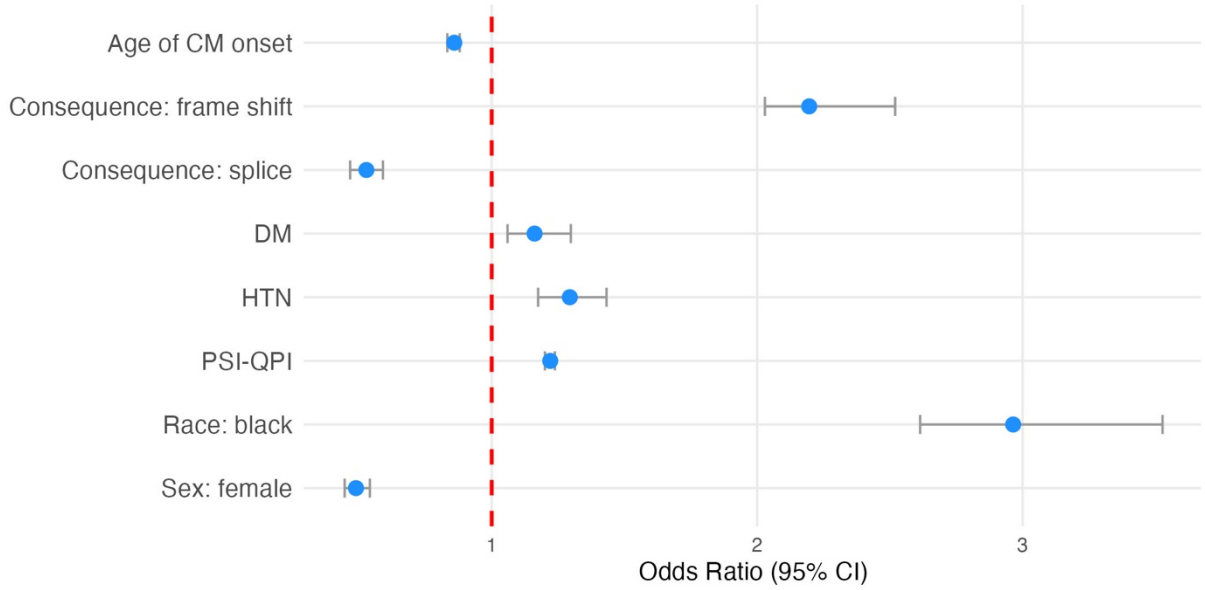
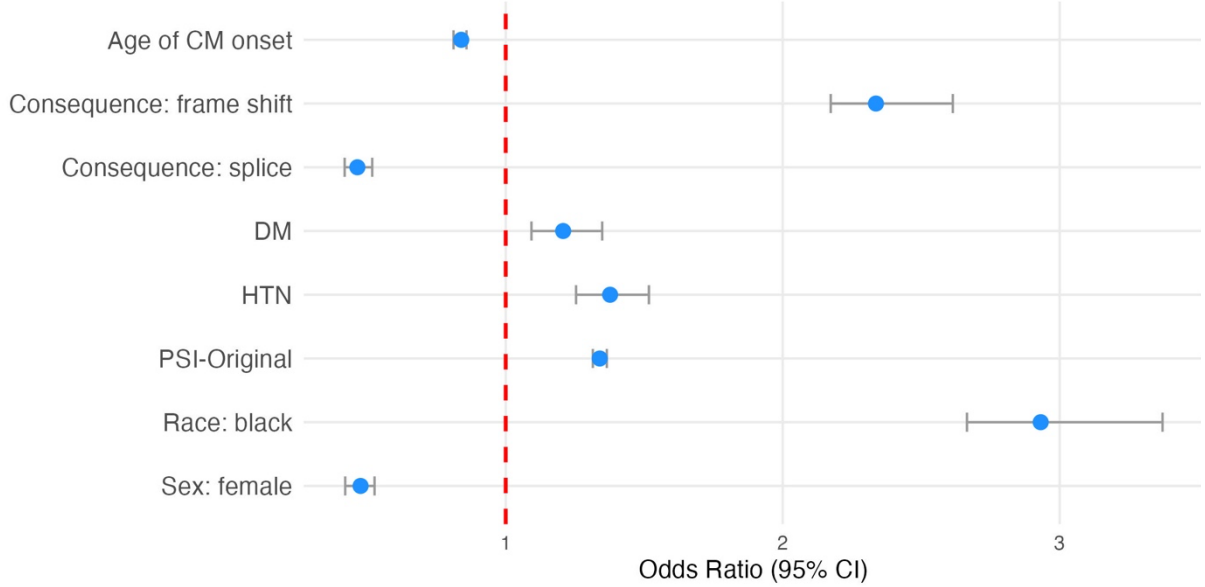

Fitted Codon Adaptation Index
0.850
0.825
0.800
0.775
0.750
p = <2e-16
Z-disk
I-band
A-band
M-band
Exon Number (metatranscript)
Domain
Z-disk
I-band
A-band
M-band

Figure 5

A


Age of CM onset
Consequence: frame shift
Consequence: splice
DM
HTN
PSI-QPI
Race: black
Sex: female
1
2
3
Odds Ratio (95% CI)

B


Age of CM onset
Consequence: frame shift
Consequence: splice
DM
HTN
PSI-Original
Race: black
Sex: female
1
2
3
Odds Ratio (95% CI)