Definition of germ cell lineage alternative splicing programs reveals a critical role for Quaking in specifying cardiac cell fate

W. Samuel Fagg[1,2*], Naiyou Liu[1], Ulrich Braunschweig[3], Xiaoting Chen[4], Steven G. Widen[2], John Paul Donohue[5], Jeffrey H. Fair[1], Matthew T. Weirauch[4,6,7], Benjamin J. Blencowe[3,8], and Mariano A. Garcia-Blanco[2,9]

[1]Transplant Division, Department of Surgery, University of Texas Medical Branch, Galveston, TX, USA.

[2]Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, Texas, USA.

[3]Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada.

[4]Center for Autoimmune Genomics and Etiology (CAGE), Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA.

[5]Sinsheimer Labs, RNA Center for Molecular Biology, Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, Santa Cruz, CA, USA.

[6]Divisions of Biomedical Informatics and Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA.

[7]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA

[8]Department of Molecular Genetics, University of Toronto, ON M5S 1A8, Canada

[9]Department of Internal Medicine, University of Texas Medical Branch, Galveston, Texas, USA.

*Correspondence to W. Samuel Fagg (wsfagg@utmb.edu)

Summary

Alternative splicing is critical for animal ontogeny; however, its role in the earliest developmental decision, the specification of the three embryonic germ layers, is poorly understood. By performing RNA-Seq on human embryonic stem cells (hESCs) and derived definitive endoderm, cardiac mesoderm, and ectoderm cell lineages, we detect distinct alternative splicing programs associated with each lineage, with the largest splicing differences observed between definitive endoderm and cardiac mesoderm. Integrative multiomics analyses predict lineage-specific RNA binding protein regulators, including a prominent role for Quaking (QKI) in the specification of cardiac mesoderm. Remarkably, knockout of QKI in hESCs disrupts the cardiac mesoderm-associated alternative splicing program and formation of myocytes, likely in part through reduced expression of BIN1 splice variants linked to cardiac development. Our results thus uncover alternative splicing programs associated with the three germ lineages and highlight an important role for QKI and its target transcripts in the formation of cardiac mesoderm.

Keywords: alternative splicing; splicing; stem cell differentiation; Quaking; QKI; cell fate; RNA processing

Introduction

Pluripotent stem cells (PSCs), such as embryonic stem cells (ESCs) and induced pluripotent stem cells, can differentiate into any adult somatic cell (Thomson et al., 1998). They are also valuable models for the study of development (O'Shea, 1999), regenerative medicine applications (Mahla, 2016; Murry and Keller, 2008), and drug discovery/toxicity testing (Baxter et al., 2010). Recent advances in human PSC differentiation have enabled the production of large numbers of nearly pure germ layer committed progenitors including definitive endoderm, mesoderm, ectoderm, and their progeny (Ang et al., 2018; Loh et al., 2014; Loh et al., 2016; Smukler et al., 2006; Tchieu et al., 2017). This has enabled the detailed characterization of the transcriptional and epigenetic/chromatin states of these lineages in comparison with hPSCs, underscoring the diversity of their transcriptional outputs (Dixon et al., 2015; Gifford et al., 2013; Loh et al., 2014; Loh et al., 2016; Tsankov et al., 2015; Xie et al., 2013). However, the use of these approaches to investigate post-transcriptional regulatory programs underlying germ layer specification remains largely unexplored.

Alternative splicing enables the widespread quantitative and qualitative regulation of gene outputs (Hamid and Makeyev, 2014; Ule and Blencowe, 2019). Accordingly, different cell and tissue types have unique splicing patterns that contribute to their identity, such as those observed in muscle and brain (Baralle and Giudice, 2017; Vuong et al., 2016). Studies have also identified splicing patterns that are unique to stem and precursor cells, including embryonic stem cells (ESCs) and comparison of these to cardiac precursor cells (Salomonis et al., 2009; Zhang et al., 2016), other late mesoderm cell types (Venables et al., 2013), and neural stem cells (Yamazaki et al., 2018; Yeo et al., 2007). Alternative splicing and indeed most RNA processing are controlled in large part by RNA binding proteins (RBPs), which are the major regulators of tissue- and developmental-specific splicing (Fu and Ares, 2014; Jangi and Sharp, 2014).

Quaking (QKI) is one such family of RBPs. From the single *Quaking* locus in human and mouse at least three major RNA and protein isoforms are produced (Ebersole et al., 1996; Kondo et al., 1999). Like other RBPs, QKI is multifunctional, regulating both nuclear and cytoplasmic RNA processing (Hall

et al., 2013; Larocque et al., 2002; Li et al., 2000; Saccomanno et al., 1999; Zhao et al., 2010), with the QKI5 isoform being necessary and sufficient for splicing function (Fagg et al., 2017). Although there is a clear role for *QKI* in the transition from myoblasts to myotubes (Hall et al., 2013), in smooth muscle (Noveroske et al., 2002; van der Veer et al., 2013), and cardiac muscle (Gupta et al., 2018), it is unknown if QKI regulates earlier cell fate decisions. Other RBPs, including RBFOX2, hnRNPF/H, and MBNL, are known to regulate splicing that influences self-renewal or differentiation (Han et al., 2013; Yamazaki et al., 2018; Yeo et al., 2009). However, the extent to which splicing programs and regulatory RBPs diverge during early cell fate lineage commitment, and their influence on downstream differentiation programs is unknown.

Here we used RNA-sequencing (RNA-seq) to address these questions by quantitatively profiling both developmentally-regulated and lineage-specific alternative splicing, as well as RBP abundance dynamics, in H9 hESCs differentiated in parallel to definitive endoderm (DE), cardiac mesoderm (CM), or neuroectoderm (ECT) cells. These analyses provide evidence for distinct lineage-specific splicing programs, with a particularly marked difference between DE and CM cells. In fact, many alternative exons are spliced in an opposite manner in DE and CM cells. We further show that many of these differentially spliced exons contain – or are flanked by – motifs that are bound and regulated by QKI. Consistent with these observations, QKI levels increase in CM and decrease in DE, compared to undifferentiated hESCs, and disruption of QKI expression reveals that it is required for a substantial fraction of CM-specific splicing changes. One such target is exon 7 in Bridging Integrator 1/Amphiphysin 2 (BIN1) whereby through inhibiting exon inclusion QKI consequently promotes intron removal, transcript maturation, and subsequent protein accumulation. Moreover, knockout of QKI prevents the transition from CM to cardiomyocytes, which may be mediated in part by a requirement of QKI for expression of BIN1. These findings thus uncover a key role for alternative splicing in defining very early lineage-specific gene expression patterns, and implicate QKI as a cardiac mesoderm-enriched RBP that is required for the control of lineage-specific splicing and differentiation of cardiac myocytes.

Results

Identification of developmentally-regulated and lineage-specific alternative splicing

To identify alternatively spliced events in human pluripotent stem cells and their germ layer-committed progenitors, we differentiated H9 hESCs in parallel to DE (Fig S1A; (Loh et al., 2014)), CM; (Loh et al., 2016), and ECT (Fig 1A; (Smukler et al., 2006)), and performed bulk RNA-sequencing (RNA-seq) to analyze their transcriptomes. Analysis of known pluripotency- or lineage-specific biomarkers at the level of RNA (Fig 1B) and protein (Fig S1B) indicate efficient differentiation, consistent with previous reports (Loh et al., 2014; Loh et al., 2016). Additionally, principal component analysis of transcript level changes shows a close clustering of the three sample replicates from each lineage, indicating a high level of reproducibility (Fig S1C). We also found that most transcript abundance changes (relative to undifferentiated cells (UD)) from protein coding genes overlapped between the three germ layer-committed progenitor cells (Fig S1D), similar to previous findings (Gifford et al., 2013).

We used VAST-tools (Irimia et al., 2014; Tapial et al., 2017) to measure splicing changes in each lineage relative to UD hESCs and found 588 statistically significant changed events in ECT, 977 in DE, and 1,144 in CM ($\geq$95% likelihood with mean delta-Percent Spliced In (dPSI) $\geq$ 10 (see methods); Fig S1E and Table S1). The majority of these are alternative cassette exons or microexons (CE/MIC), or retained introns (Fig S1E). We observe a similar frequency of increased inclusion vs skipping among differentially spliced CE/MIC events in DE, CM, and ECT, compared to UD cells, that are not significantly different (Fig 1C). The majority of alternatively spliced exons detected in DE or CM cells are specific to these lineages, but more than half of the events detected in ECT cells overlap with the other lineages (Fig 1C) irrespective of the cutoffs used, i.e. lower or higher dPSI values (Table S1). Moreover, specific CE/MIC events (see below) detected by RNA-Seq were validated at a high rate by RT-PCR assays (dPSI correlation: DE R^2 = 0.87, CM R^2 = 0.94, ECT R^2 = 0.74), supporting the robustness of the measurement (Fig S1F).

These included exons strongly affected in either CM cells (RAI14 exon 11, SLK exon 13, and MARK3 exons 17 and 18) or DE cells (PICALM exon 17 and CASK exon 19) by RT-PCR (Fig 1D). Interestingly, alternative exons in FOXP1 (exons 18A or 18B; (Gabut et al., 2011)) and DNMT3B (exons 21 and 22; (Duymich et al., 2016; Gordon et al., 2013; Van Emburgh and Robertson, 2011)), previously described as functionally promoting a differentiated state, are specifically altered with greatest magnitude in CM cells, to a lesser extent in ECT cells, and are little changed in DE cells (Fig 1D). Finally, an additional class of exons shows opposite splicing patterns between DE and CM cells, e.g. BIN1 exon 7 is more included in DE cells, more skipped in CM cells, and mostly unaffected in ECT cells, and vice-versa in the cases of CLSTN1 exon 11 and NF1 exon 23 (Fig 1D). Identification of these events suggested DE and CM splicing patterns might be more divergent than compared to UD. Accordingly, we directly compared them and found 1,382 total alternative splicing events (Fig S1G), including 144 additional unique alternative exons not revealed when compared to UD hESCs (Fig S1H, compare to Fig 1C). Many of these appear to be due to a modest alteration in inclusion in DE relative to UD, and an additional but opposite modest change in inclusion in CM relative to UD (Table S1). Consistent with this, hierarchical clustering of CE/MIC events revealed that CM and ECT splicing patterns are more similar to each other than to DE (Fig 1E). We also observed significant enrichment of functional terms associated with genes containing these differentially spliced exons that are consistent with the biology of each lineage: i.e. contractile fiber part and related terms are associated with genes displaying increased exon inclusion in CM cells, various cell adhesion/junction-based terms are enriched in genes with more skipping in DE cells, and the term synapse part is enriched in genes with exons that are more included in ECT cells (Fig 1F). These observations reveal that most alternatively spliced exons detected in DE and CM cells are spliced in a lineage-specific manner.

Inspection of differentially retained introns relative to UD hESCs shows similar characteristics as differentially regulated CE/MIC events, but with unique and noteworthy features. Overall, there are considerably more differentially retained introns specific to each lineage than observed for cassette exons, with the largest and most distinct programs detected in DE and CM cells, and most intron

retention (IR) events detected in ECT cells are unique (Fig S1E and S1I). Other features include a significantly higher frequency of reduced retention levels in DE and ECT cells, but a significantly higher frequency of increased retention levels in CM cells ($P < 0.0001$ by Mann-Whitney U, Fig S1I). Moreover, CM and ECT progenitors have more similar IR profiles, whereas DE cells are more different (Fig S1J). Additionally, 682 unique IR events were observed (Fig S1K) when comparing CM to DE, than either to UD hESCs (Fig 1D). This confirms a more dissimilar pattern than suggested by the analysis that compared each to UD hESCs (Fig 1E).We also found significant enrichment of functional terms amongst genes containing IR events that are unique to CM cells, including “RNA modification” (increased IR) and “Fibrillar collagen trimer” (decreased IR; Fig S1L). Collectively, these data suggest that alternative splicing represents a unique and orthogonal molecular regulatory program (compared to transcriptional regulation) that is executed between germ layer-committed progenitor cells. layers.

QKI5 increases in CM cells and decreases in DE cells

We hypothesized that lineage-specific alternative splicing is regulated by distinct RBPs. We measured RBPs (Cook et al., 2011) that significantly changed ($P \leq 0.01$ by Benjamini-Hochberg) (Anders and Huber, 2010) in expression across our RNA-seq data (Table S2) and found that each germ layer representative cell type has a unique network of RBPs, with DE and ECT RBP profiles are slightly more similar, and CM less so (Fig 2A). This allowed for rapid identification of particular RBPs with differential transcript levels between DE and CM. We focused on those with a known role in developmentally-regulated splicing, validated their abundance change with RT-qPCR (Fig 2B), and found: both MBNL1 and MBNL2 transcripts are significantly higher in CM compared to DE; RBFOX2 does not significantly change at the RNA level; Quaking (QKI) is significantly higher in CM, lower in DE and unchanged in ECT relative to UD hESCs; ESRP1 and ESRP2 transcripts are significantly lower in CM relative to the other lineages; CELF2 RNA increases in DE and is unchanged in other lineages (Fig 2B).

Since MBNL and QKI RNA levels change most between DE and CM, we also measured their protein abundance. Western blot analysis of pluripotency (OCT4), DE (SOX17), or CM (NKX2.5)

differentiation markers indicate specific and efficient differentiation of each cell type (Fig 2C). MBNL proteins are undetectable in UD, DE, or CM lysates, although a positive control with C2C12 myoblast lysate shows readily detectable MBNL protein (Fig S2A). In contrast, QKI showed significantly decreased levels in DE and increased levels in CM, relative to UD hESCs ($P < 0.05$ by student's t-test; Fig 2D). We also measured the abundances of QKI5, QKI6, and QKI7 isoforms relative to total QKI protein as previously described (Fagg et al., 2017), and found that the most abundant (>95%) is the nuclear isoform QKI5 (Fig S2B). Additionally, we verified that QKI protein is predominantly nuclear, and co-expressed with lineage-specific biomarkers (SSEA3 for UD, SOX17 for DE, and NKX2.5 for CM; Fig S2C).

Similar to other splicing regulatory RBPs that act by binding to specific *cis*-regulatory motifs (Barash et al., 2010; Ule et al., 2006), QKI typically promotes either exon skipping by binding the consensus motif ACUAAY (Galarneau and Richard, 2005; Hafner et al., 2010; Ryder and Williamson, 2004) located within, or upstream of exons, or exon inclusion by binding the same motif in introns downstream of regulated exons (de Bruin et al., 2016; Hall et al., 2013). We used the analysis program rMATS (Dixon et al., 2015) to ask if lineage-specific alternatively spliced exons, or their flanking sequences, are enriched for the QKI motif. As expected, we observed little to no enrichment of this motif in alternatively spliced exons from ECT cells vs. UD (Fig S2D; Table S3). In contrast, QKI motifs are significantly enriched in intron sequences that are within 250 nt downstream of included exons, and in exons and upstream introns that are more skipped in CM vs. UD cells (Fig S2E; Table S3), and vice-versa in DE vs. UD cassette exons (Fig S2F; Table S3). The enrichment of these QKI motifs is consistent with the observed significant increase in QKI abundance in CM cells and decrease in DE cells (Fig 2D; represents an 81% change; $^{**}P < 0.01$ by student's t-test), and, similarly, there is a greater enrichment of the QKI motif around CM cell alternative exons when compared to DE cells (Fig 2E, and compare to Figs S2E and S2F; Table S3). To conclude, the above data are consistent with a lineage-specific splicing regulatory role for QKI5, and suggest that QKI activity is reduced in DE and increased the in CM lineage.

QKI protein directly binds a significant number of CM-specific alternatively spliced pre-mRNAs

The above data strongly suggest QKI5 is a lineage-specific RBP that promotes CM splicing. To investigate whether QKI5 binds directly to the pre-mRNA targets inferred from the analysis described above, , we re-purposed a transcription factor/chromatin data analysis method called Regulatory Element Locus Intersection or RELI (Harley et al., 2018) to analyze RBPs activity. In brief, RELI uses a simulation-based procedure to systematically gauge the significance of the intersection between a set of input regions (e.g., and in this case, genomic coordinates of regions surrounding exons alternatively-spliced between CM and DE cells) and each member of a large library of functional genomics experiments, in this case, over 450 RBP crosslinking immunoprecipitation (CLIP)-seq experiments. This analysis revealed multiple results supporting the importance of QKI in CM-specific alternative splicing. In particular, CLIP-seq binding peaks for QKI were among the most highly enriched within intronic regions downstream of exons more included (17% of regions, $P < 10^{-16}$, 2.2-fold enriched, rank #6), within exons more skipped (5% of regions, $P < 10^{-12}$, 4.2-fold enriched, rank #65), and in introns upstream of exons more skipped (6% of regions, $P < 10^{-6}$, 2.8-fold enriched, rank #18) in CM compared to DE (Fig 2F and Table S4). In total, we found that 39.2% (+/- 2.6%) of QKI eCLIP peaks in HepG2 cells overlap with CM-specific alternative exons or are in introns 250 bp up- or downstream of these exons (Table S4). Particularly interesting examples of evidence for direct binding of QKI are found within or proximal to lineage-specific alternative exons in *BIN1*, *NF1*, *CLSTN1*, *RAI14, PICALM*, *SLK*, and *MARK3* (Fig 1F) (Figs 2G, S2G, and (Fagg et al., 2017)). The splicing outcomes observed in these transcripts are fully consistent with QKI abundance and binding location relative to the lineage-specific exon. Collectively, these data support a prominent role and direct role for QKI in the regulation of alternative splicing program specifically linked to formation of the CM lineage.

Loss of QKI results in dysregulated splicing of cardiac mesoderm-specific exons

To determine whether QKI promotes CM-specific alternative splicing, we used CRISPR/Cas9 editing to generate hESC QKI knockout (QKI-KO) cells expressing EGFP under the control of the endogenous NKX2.5 promoter, which is specifically activated in cardiac mesoderm/myocytes (Elliott et al., 2011), and then measured the inclusion levels of several CM-specific alternatively spliced exons in this line. Disruption of the QKI ORF was confirmed by analysis of PCR products derived from genomic DNA (Fig S3A; (Brinkman et al., 2014)). We also selected and expanded single cells from which three clonal NKX2.5→GFP knockout hESCs were assessed. Clones 1-12 and 4-2 have the same two lesions, while clone 4-5 shares one of the same lesions and an additional unique lesion (Fig S3B). QKI protein is undetectable in all three and the mixed clonal population of cells (Fig 3A). Correspondingly, the inclusion of predicted QKI target exons changes. RAI14 exon 11 and BIN1 exon 7 cassette exons are skipped more in wild-type (WT) CM cells consistent with the prediction that they are QKI-repressed (Fig 1D), are more included in the QKI knock-out cells (Fig 3B). Conversely, NF1 exon 23 and CLSTN1 exon 11, which are more included in CM cells (Fig 1D), consistent with the prediction that they are QKI-activated, are more skipped in the QKI knock-out lines (Fig 3B). Also, the PICALM alternative exon, which is repressed in DE cells, likely via a reduction in QKI5 protein levels, is less included in the absence of QKI (Fig S3C). The CM cell-specific alternative exon observed in DNMT3B that was not predicted to be regulated by QKI is unaffected by QKI loss (Fig S3D). These findings were independently validated following knockdown with two independent shRNAs targeting the QKI5 isoform (Figs 3C, 3D, 3E, S3E, and S3F). We also tested if RAI14 exon 11, NF1 exon 23, and BIN1 exon 7 inclusion levels were altered at various time points in these cells subjected to cardiac lineage differentiation. QKI is absolutely required for skipping of RAI14 exon 11; NF1 exon 23 inclusion increases during differentiation, but much less so in the absence of QKI; BIN1 exon 7 skipping increases during differentiation, but is little changed in the absence of QKI (Figs 3F and S3G). Therefore, in the absence of QKI, these CM cell-specific exons with evidence of QKI direct binding are mis-regulated.

QKI is required for BIN1 transcript processing and protein accumulation

Interestingly, *Bin1* is required for cardiac development (Muller et al., 2003): specifically its absence or disruption of cardiac-specific isoform expression results in either improper calcium signaling or defective myofibril formation, respectively (Hong et al., 2014; Hong et al., 2010). Given its importance in cardiac development and homeostasis, we explored the consequences of QKI loss on its overall expression (Fig 4A). Accordingly, we cloned and sequenced full-length BIN1 cDNAs from both UD and day 8 cardiomyocyte-differentiated WT and QKI-KO cells, and observed the expected increase in exon 7 inclusion in the absence of QKI, but also more inclusion of exon 13, the clathrin and AP (CLAP) binding domain-coding exons, and exon 17 (Figs S4A and 4A). Total BIN1 mRNA is lower in the absence of QKI in UD cells, but slightly increases in day 8 QKI-KO cardiomyocyte cells (Fig S4B). Its protein levels, on the other hand, are strikingly different: a predominantly lower molecular weight version is observed in QKI-KO UD; significantly lower abundance (*$P < 0.05$ by student's t-test) and altered isoform output is observed in day 4 CM QKI-KO cells; strikingly, in day 8 cardiomyocyte differentiation, the abundance is even more significantly reduced (***$P < 0.001$ by student's t-test), and only the two higher molecular weight isoforms are detectable in the absence of QKI (Fig 4B). Total BIN1 protein increases about 3-fold during WT cell cardiac differentiation, but is unchanged compared to UD in the absence of QKI (Fig S4C). High content imaging analyses also revealed reduced BIN1 protein at both the population (Fig S4D) and per cell (Fig S4E) levels in QKI-KO day 8 cardiomyocyte cells. Therefore, QKI is required for proper BIN1 protein accumulation and isoform output during cardiac differentiation.

Quaking binds the BIN1 transcript (Fig S4A) and promotes skipping of BIN1 exon 7 (Figs 3B, 3E, and 4G; (Gazzara et al., 2014; Hall et al., 2013)), which we also confirmed in QKI-KO clone 4-2 subjected to cardiac differentiation (Fig 4C). Additionally, we observed more inclusion of CLAP domain-containing exons and exons 13 + 17 in QKI-KO d4 CM and d8 cardiomyocytes, with correspondingly less of the "skipped all" isoform (Fig 4D). Together, these data indicate a loss of "muscle-specific"

(Prokic et al., 2014) and a gain of “neuron-specific” splicing patterns in *BIN1* (Butler et al., 1997; Ellis et al., 2012; Ramjaun et al., 1997; Wechsler-Reya et al., 1997), in the absence of QKI.

To explore the potential molecular mechanisms through which QKI regulates BIN1 exon 7 splicing, we cloned this exon, along with 126 bp and 184 bp of upstream and downstream intron, respectively, into the pDUP51 beta-globin splicing reporter (pDUP-BIN1 (Dominski and Kole, 1991)). Transfection of this reporter into HuH7 WT or QKI-KO cells and analysis by RT-PCR showed a modest increase in exon 7 inclusion in the absence of QKI (Fig 4E), consistent with QKI promoting skipping of this exon. Strikingly, the largest magnitude change observed was an increase in unspliced reporter RNA (Fig 4E). Closer inspection of BIN1 exon 7 and surrounding sequences revealed a potential exonic QKI motif CACUAAC and putative “half-site” in the downstream intron CUAAC (Fig 4F). To further explore the impact of these potential QKI motifs on splicing, we generated exonic or intronic deletion or substitution mutants, and double mutants to ablate the respective motifs in the pDUP-BIN1 reporter then measured their splicing in WT HuH7 cells (Fig 4F). Overall, deletion and substitution of the exonic motif resulted in more inclusion of exon 7, unspliced RNA, and mis-spliced RNA (Fig 4G). Importantly, the observed increases in exon and intron 7 inclusion from the reporter were recapitulated in the endogenous BIN1 transcript from QKI-KO cells subjected to day 8 cardiomyocyte differentiation (Fig 4H). Taken together, we conclude that QKI is required for cardiac muscle-specific BIN1 transcript processing and protein accumulation.

QKI is required for efficient cardiomyocyte cell differentiation

Our data support that QKI5 directly promotes approximately 40% of CM lineage-specific CE/MIC splicing patterns (Fig 2F and Table S4), but it is not clear whether or not QKI5 is required for cardiac differentiation. We used the QKI null hESCs to test the efficiency at which these and WT control cells form CM and cardiomyocytes (Loh et al., 2016). QKI protein increases when WT hESCs differentiate to CM (Fig 2D) and cardiomyocytes (~6-fold increase), but as expected is undetectable in UD and day 8 polyclonal QKI-KO cells subjected to the cardiomyocyte differentiation protocol (Fig 5A). The level of

QKI5 RNA increases in WT cells under this differentiation time course, and the QKI-KO cells do also, but at a larger magnitude compared to WT (Fig S5A), consistent with known autogenous regulatory features (Fagg et al., 2017). Interestingly, QKI KO does not affect NKX2.5→EGFP levels during CM differentiation (Fig 5B) but the expression of this marker is significantly reduced upon further maturation to cardiomyocytes (Fig 5B and 5C; *$P < 0.05$, **$P < 0.01$ and ****$P < 0.001$ respectively, by student's t-test), suggesting that QKI may be required for cardiomyocyte differentiation.

To further address this, we measured endogenous biomarkers for CM and cardiomyocyte cell fate in QKI-KO and WT NKX2.5→EGFP hESCs subjected to these differentiation protocols. Consistent with reporter EGFP levels, endogenous NKX2.5 mRNA is unchanged in day 4 CM, but is significantly lower in 4-2 KO day 8 cardiomyocyte cells ($P < 0.05$ by student's t-test; Fig 5D); cardiac Troponin T (cTNT; *TNNT2* gene), however, is lower in both CM and cardiomyocyte QKI KO cells (Fig 5D). Additional measurement of mRNA biomarkers for CM progenitor cells showed mixed results: ISL1 and GATA4 abundance is lower in QKI KO cells, but HAND1 and MEF2C do not significantly change (Fig S5B). Thus, QKI loss may moderately disrupt CM cell differentiation or certain sub-populations of cardiac progenitor cells. We also used high content imaging to analyze these developmental transitions in WT and QKI-KO NKX2.5→GFP hESCs, and observed no gross morphological perturbations, nor significant changes in OCT4 protein levels in the QKI-KO cells (Fig S5C). However, upon 4 day differentiation to CM, endogenous NKX2.5 protein is significantly lower on a per-cell basis in the absence of QKI ($P < 0.05$ by Mann Whitney U; Fig S5D). By day 8, both NKX2.5 and cTNT protein levels are much lower in QKI-KO cells by western blot analysis (Fig 5E). High content imaging analysis (on both a per cell (Fig 5F) and population basis (Fig S5E)) also indicate significantly reduced cTNT protein levels in day 8 QKI-KO cells (****$P < 0.0001$ by Mann-Whitney U and ***$P < 0.001$ by student's t-test, respectively). Likewise, the morphology of QKI-KO cells indicates a reduction of branched cardiac muscle-like structures at day 8 (Fig 5F and S5F). Taking these observations together, loss of QKI appears to moderately perturb CM differentiation, but significantly impairs cardiomyocyte differentiation. We conclude that QKI is required for proper cardiac differentiation.

Discussion

The Quaking ortholog *alternative splicing defective 2* (*asd2;* comparable to the QKI5 isoform) in *C. elegans* is required for proper muscle body wall development (Ohno et al., 2008); *Drosophila held out wings* (*how*) mutants show defective mesoderm formation (Baehrecke, 1997; Nabel-Rosen et al., 1999; Zaffran et al., 1997). Similarly, *quaking-a* (*qk-a*) is required for proper muscle development in zebrafish (Lobbardi et al., 2011), and *qk* null mutant mice have defects in smooth muscle development (Li et al., 2003). There is a clear role for mouse *Qk* in development (Darbelli et al., 2016; Ebersole et al., 1996; Li et al., 2003). Although the aformentioned studies indicated that Quaking is required for mesoderm/muscle development in a variety of model organisms, our study uncovered that this RNA binding protein is required at the earliest lineage commitment in human cells: either mesoderm or endoderm.

We find here that QKI regulates multiple exons in genes linked to cardiac tissue development (Brannan et al., 1994; Muller et al., 2003), and among these exons we focused on the QKI-dependent repression of exon 7 of BIN1. Inclusion of BIN1 exon 7 in skeletal muscle is associated with centronuclear myopathy in myotonic dystrophy patients (Fugier et al., 2011), and is promoted by nSR100 in neurons (Ellis et al., 2012), which do not express QKI (Hardy et al., 1996). Other BIN1 exons, 13, 13-16 (CLAP domain), and exon 17, included in neurons but skipped in cardiomyocytes, are also repressed by QKI (Fig 4), suggesting that cardiac and neuronal BIN1 splicing is determined by QKI. Dysregulation of total BIN1 level and splicing of these exons in the brain is correlated with Alzheimer's disease (Chapuis et al., 2013; De Rossi et al., 2016) . Therefore, the tight regulation of BIN1 levels and isoform output is required for proper tissue homeostasis and differentiation in cardiac muscle, skeletal muscle, and the brain, and here we establish that QKI directly regulates these via RNA splicing.

The gene expression process is highly regulated, and subject to dynamic on/off states as well as finely tuned rheostat-like adjustments to properly respond to various inputs (Lee and Young, 2013;

Rizzino, 2009; Rossi et al., 2000). Recently, extensive mapping of chromatin/histone and promoter states, and transcription factor and RNA polymerase II occupancy during commitment of hESCs to each of the three germ layers has identified both unique and shared states between these different progenitor cell types (Gifford et al., 2013; Tsankov et al., 2015). Consistent with previous findings (Gifford et al., 2013), we observe here that quantitative changes in gene expression are shared between the three germ layer derivatives (Fig S1D). This implies that most transcriptional on/off states are developmentally regulated, and are not lineage-specific. In turn, these are potential targets of alternative splicing, which can have profound effects on the ultimate gene product repertoire. Interestingly, the splicing patterns of DE and CM cells are uniquely different from hESCs, and even more different from one another (Figs 1 and 2). Of particular interest here is the finding that several exons in genes coding for transcription/chromatin modifying factors that promote differentiation are markedly different between DE and CM cells. These are examples of splicing level regulation, the products of which feedback to alter the transcriptional response, either directly (i.e., FOXP1) or indirectly by altering chromatin structure (i.e., DNMT3B). Therefore, rheostat-like modulation of gene expression via upstream alternative splicing can feedback and influence downstream transcriptional regulatory events in a lineage-specific manner. While certainly not all of these splicing changes can be attributed to QKI, we do find here that many of them in the cardiac lineage are, implicating QKI as a master regulator splicing factor (Jangi and Sharp, 2014). Orchestration by QKI of these complex and oftentimes combinatorial molecular-level transitions is critical for the exit from pluripotency and to lineage-specific gene expression.

Acknowledgments

We thank Rhonda Perriman for pre-submission review and valuable comments. We thank Seigo Hatada, David Elliot, Andrew Elefanty, and Ed Stanley for hESCs. We thank our funding sources: UTMB start-up funds to MAGB and JHF, and Sealy-Smith Endowment and John L. Hearn Endowment funds to JHF and WSF, NIH R01 CA204806 (MGB), grants from the Canadian Institutes of Health

Research and Canada First Research Excellence Fund Medicine by Design Program (to BJB), NIH R01 HG010730, R01 NS099068, R01 AR073228, R01 GM055479, U01 AI130830, and Cincinnati Children's Hospital "Center for Pediatric Genomics" and "CCRF Endowed Scholar" awards to MTW, BJB holds the Banbury Chair in Medical Research at the University of Toronto, and MAGB holds the Mildred H Vacek and John R Vacek Distinguished Chair in Honor of President Truman G. Blocker, Jr.

Author's Contributions

Conceptualization, WSF.; Methodology, WSF; Formal analysis, WSF, UB, XC, JPD, SGW MTW; Investigation, WSF, NL, UB; Resources, WSF, JHF, MTW, BJB, MAGB; Writing – Original Draft, WSF, Writing – Review and Editing, WSF, UB, SGW, MTW, BJB, MAGB; Visualization, WSF, UB; Supervision WSF; Project Administration WSF; Funding Acquisition JHF, MTW, BJB, MAGB.

Methods

Cell culture

H9-hrGN hESCs were a generous gift from Seigo Hatada, and NKX2.5→GFP hESCs were a kind gift from David Elliot, Andrew Elefanty, and Ed Stanley, and were acclimated to feeder-free conditions and propagated on plates coated with Matrigel (hESC Qualified, Corning) in mTeSR1 media (Stemcell Technologies) for general maintenance of pluripotent cultures. Cells were passed with ReLeSR or Gentle Cell Dissociation Reagent (Stemcell Technologies) as needed and described by the manufacturer. For differentiation, cells were passaged and plated at approximately 1:5 or 1:10 ratios and allowed to equilibrate in propagation media for 1-2 days. Differentiation was then initiated following protocols for DE (Loh et al., 2014) and for RNA-seq cell sorting was performed by positive selection via CXCR4 using FACS (see below, and this was due to empirical observations of variable efficiencies of differentiation and cell death due to cell density), CM and cardiomyocytes as described (Loh et al., 2016), and for ECT as described (Smukler et al., 2006) but modified by using CDM2 as described (Loh

et al., 2014) for 3 days, then with the addition of knockout serum replacement (KOSR) to CDM2 for an additional 4 days.

Plasmids, transfections, shRNAs, and antibodies

H9 or NKX2.5→GFP hESCs were transfected using Lipofectamine Stem Transfection Reagent (ThermoFisher), according to the manufacturer's instructions. HuH7 cells were transfected using Lipofectamine2000 (ThermoFisher), according to the manufacturer's instructions. NKX2.5→GFP hESCs were co-transfected with pX458-sgQKI as described (Liao et al., 2020) or control empty vector along with pCEP-puro which contains a puromycin-resistance gene. One day later puromycin was added at 1 μg/ml and cells were selected for 3 days, and then expanded in the absence of puromycin. Upon verification of reduced QKI protein in expanded cells, then directly sequencing PCR products from total population-level gDNA and analyzing with TIDE (Brinkman et al., 2014), limiting dilutions of 0.5 cells per well were performed and delivered to 96-well plates. These were monitored visually to confirm only single cells were present in each well, then expanded. Potential clonal KO lines were initially screened by western blot, then genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen) then amplified by PCR with PrimeSTAR pre-mix HS (Takara), and cloned into pCR4Blunt-TOPO (ThermoFisher). Plasmid DNA from six different colonies derived from independent hESC clones were sequenced to confirm two copy disruption of the reading frame.

Custom shRNA sequences were generated and added to pLKO.1-hPGK-Neo (MilliporeSigma) plasmids which were transfected into HEK293T cells along with packaging plasmid (pCMVR8.74) and VSV-G envelope plasmid (pMD2.G) to generate viral supernatant as described (addgene.org/protocols/lentivirus-production/). Undifferentiated H9 hESCs were infected with equal volumes of lentivirus, then selected with 200 μg/ml G418 beginning 24h later and continuing for 48-72h, at which time cells were harvested. Oligonucleotide sequences are listed in Table S5.

The mouse monoclonal antibodies used for Quaking are Anti-PanQKI (N147/6; AB_2173149)), Anti-QKI-5 (N195A/16; AB_10674117), Anti-QKI-6 (N182/17; AB_10673512), and Anti-QKI-7 (N183/15;

AB_2173148), all from UC Davis/NIH NeuroMab Facility or Millipore/Sigma. Rabbit anti-QKI was from Bethyl (cat# A300-183A). Mouse anti-Gapdh (clone GAPDH-71.1, cat# G8795) and mouse anti-MBNL1 (clone 3A4) were from Sigma. Rabbit anti-Tubulin (cat# 2148), rabbit anti-NKX2.5 (clone E17AH, cat# 8792S), and mouse anti-SOX2 (clone L1D682, cat# 4900S) were from Cell Signaling Technologies. Mouse anti-OCT4 (cat# ab184665), rabbit anti-cTNT (cat# ab91605), and rabbit anti-BIN1 (cat# ab182562) were from Abcam. Goat anti-SOX17 (cat# AF1924) and rat anti-SSEA3 (clone MC-38, cat# MAB1434) were from R&D Systems. Mouse anti-CD184 (CXCR4), APC conjugate (clone 12G5) and isotype control (clone MOPC-173) were from BioLegend. Mouse anti-MBNL1/2/3 (cat# sc-58790) was from Santa Cruz Biotech.

Western blotting

Proteins were extracted using RSB100 buffer with SDS (100 mM Tris-HCl pH 7.4, 0.5% NP-40, 0.5% Triton X-100, 0.1% SDS, 100 mM NaCl, and 2.5 mM $MgCl_2$) plus protease inhibitor cocktail and PMSF, by direct addition of this buffer to wells/plates containing cells and scraped with cell scrapers and collect by pipet. Lysates were then vortexed intermittently while kept on ice for approximately 30 minutes, then centrifuged at 14,000 RPM at 4C for 15 minutes. Clarified lysate was transferred to a clean tube and either used immediately or store at -80C. Lysates were separated on 10 or 12% SDS-PAGE then transferred to 0.4 µm nitrocellulose. Blocking was performed with 5% non-fat milk (NFM) in tris-buffered saline (TBS) for 1 hour at room temperature (RT). Primary antibody incubations were performed overnight at 4C in NFM in TBS containing 0.01% tween-20 (TBST), then membranes were washed at least 3 times with TBST at RT. Secondary infrared-conjugated antibodies were added to NFM in TBST at a 1:20,000 dilution for 680:LT conjugates or 1:15,000 for 800:CW conjugates, and incubated for 1 hour at RT. Membranes were then washed 3 times for 5 minutes and then 1-2 times for 15 minutes each before visualization on Odyssey CLx imager (Li-Cor). For quantitation, the region of interest for each band was selected and the value recorded. In certain cases, signal was normalized to a loading control (GAPDH or Tubulin) and the mean was calculated, with mean +/- standard deviation reported, and significance calculated using the student's t-test.

Immunofluorescence

Cells were fixed with 4% paraformalydehyde (Sigma Aldrich) diluted in phosphate-buffered saline (PBS) for 15-30 minutes then washed with PBS 2-3 times and stored in PBS until stained. Cells were blocked/permeabilized with 1% BSA and 0.3% Triton X100 in PBS for 1 hour at RT. Cells were then incubated overnight at 4C in primary antibody diluted in the above mentioned buffer. The following day, they were washed 3 times with PBS, then incubated in secondary antibody (AlexaFluor 568 or AlexaFluor 647 (Thermo Scientific) at 1:1000 dilution in the above blocking buffer with the addition of 1 µg/ml DAPI) for 1 hour at RT, and then washed 3 times with PBS. Cells were then imaged immediately or stored for later analysis. Imaging was performed with Opera Phenix High Content Screening microscope (Perkin Elmer). Fluorescence intensity values were calculated for each channel on a per-cell and per-well basis, and recorded then subsequently analyzed in GraphPad Prism. No correction was performed for low/no signal as some signal could be detected in any given channel due to background, and as such, this was factored into reported values. Low magnification live cell imaging for NKX2.5→EGFP show in Fig S5F were performed on an Olympus IX71 inverted fluorescence microscope using the FITC filter.

RNA extraction, RT-qPCR, and RT-PCR

RNA was extracted from cells by directly adding TRIzol (Thermo Fisher) reagent to wells containing cells, or using the ReliaPrep RNA MiniPrep system (Promega), each according to the manufacturer's instructions. In the case of TRIzol-based extractions, nucleic acid was treated with Turbo DNase (Thermo Fisher) then RNA re-extracted. Reverse transcription was performed with Super Script III (Thermo Fisher) using a combination of anchored oligo dT and random hexamers. Depending on the input material amount, cDNA was then diluted within a range of 1:10 to 1:50 then used for qPCR analysis with Applied Biosystems SYBR Green PCR Master Mix or PowerUp SYBR Green Master Mix (Thermo Fisher). Data were analyzed using the delta-delta CT method and reported as relative to *HMBS* "housekeeping gene"; we also confirmed similar results using *EEF1A1* and *GAPDH* as "housekeeping genes". For standard PCR, cDNA was added and amplified using Taq 2X Master Mix

(NEB) with cycle number determined empirically to avoid overamplification. PCR products were analyzed by either agarose gel or DNA1000 BioAnalyzer chip (Agilent). BioAnalyzer quantitation was determined by calculating the ratio of PCR products' concentration (nM) and reporting percent included of the band(s) of interest.

RNA sequencing and analysis

RNA was extracted with TRIzol (see above) and rRNA was depleted using the RiboCop rRNA Depletion Kit (Lexogen). This rRNA-depleted RNA was used as input to generate paired-end libraries using the TruSeq Stranded Library Kit (Illumina) which were then pooled and sequenced on two High Output NextSeq 550 flow cells with a 75 base paired-end protocol. Reads were mapped using STAR version 2.6.1c onto the human GRCh38 reference build, with the parameters recommended by the ENCODE consortium. DESeq2, version 1.20.0, was used to determine changes in gene expression following the authors' vignette. For splicing analyses, we used VAST-tools (version 2.0.2, database has.16.02.18) and rMATS (version 4.0.2, GENCODE v30 annotation) using the default parameters for each. Raw output was filtered based on quality scores, requiring at least 10 total reads per event in at least one replicate for a condition. Differential splicing was assessed with the VAST-tools diff module. Events were deemed significantly differential if they had a 95% likelihood of being differential and the mean PSI/PIR difference was greater than 10. The RNAseq data was deposited at Gene Expression Omnibus (GSE162649).

Computational analysis of CLIP and RNA-seq datasets

We collected functional genomic datasets from a variety of sources assaying RNA binding proteins and microRNAs, including ENCODE (2012), CLIPdb (Yang et al., 2015), and starBase (Yang et al., 2011). All datasets were indexed by their genomic coordinates, which were used to compare to the genomic coordinates of alternative exon sites and upstream or downstream introns using the RELI tool (Harley et al., 2018). In brief, RELI estimates the significance of the observed number of intersections between an input set of regions (e.g., alternative exon sites) and each member of a library of functional genomics datasets (e.g., CLIP-seq peaks). Significance is estimated by comparing to a null model formed by randomly sampling non-significant sites in each corresponding group used in this study (e.g.,

non-alternative exon sites). The distribution of the expected intersection values from the null model resembles, and therefore is modeled as, a normal distribution, with the model parameters estimated from a series of 2,000 sampling procedures. The significance of the observed number of intersections, e.g., a Z-score and the corresponding P-value, along with an enrichment score (observed intersection count divided by expected intersection count) are calculated by RELI, and are provided in Table S4. For splicing changes, GO analysis was done separately for CE/MIC and IR event types with a background comprising all genes that contain a splicing event of the type being tested that survived filtering, requiring a minimum three-fold enrichment. If categories mutually overlapped by more than 70%, only the more significantly enriched category was kept.

Flow cytometry and fluorescence-activated cell sorting (FACS)

Cells were prepared for flow cytometry or FACS by removal from plates/wells with Accumax (Innovative Cell Technologies Inc.) according to the manufacturer's instructions. Staining for CXCR4 surface marker was performed using APC anti-human CD184 (CXCR4; BioLegend) or with APC Mouse IgG_{2a}, k isotype control (BioLegend) following the instructions from the manufacturer. The cells were finally passed through a 40 μm strainer then kept on ice before sorting with BD FACS Aria IIU (Beckton Dickinson) with assistance from core facility operator. For analysis of EGFP in NKX2.5→EGFP cells, cells were detached as above or with 1x Trypsin/EDTA (0.25%; Thermo Fisher) according to the manufacturer's instructions and either measured immediately or fixed with 2% paraformaldehyde (Sigma Aldrich) dissolved in PBS (Thermo Fisher) to be measured later on the Guava EasyCyte HT (Millipore) flow cytometer. For all flow/FACS applications, gating was performed to restrict any events not identified as single cells. FlowJo software was used for data analysis.

FIGURE LEGENDS

FIGURE 1: Identification of lineage-specific alternative splicing. A. Experimental overview/model of cell lineages analyzed by RNA-seq: H9 undifferentiated (UD) hESCs can self-renew or were differentiated to definitive endoderm (DE), cardiac mesoderm (CM), or neuroectoderm (ECT) cells for further analysis. B. UCSC Genome Browser screen shot showing coverage tracks for reads from RNA-seq representing the cell lineages described above, measuring RNA abundances for POU5F1/OCT4, SOX17, NKX2.5, and PAX6. C. Change in percent spliced in (dPSI) of alternative cassette and micro exons for DE, CM, and ECT, significantly changing relative to UD shown in violin plot, and overlap of these shown in Venn diagram inset, calculated from RNA-seq data (ns indicates no significant change by Mann-Whitney U). D. Validation by RT-PCR and BioAnalyzer quantitation of alternative exons (VAST-tools alternative exon designation shown below transcript name) using RNA extracted from UD, DE, CM, and ECT (mean percent included +/- standard deviation from the mean is shown below). E. Heatmap showing hierarchical clustering of significantly changing alternative exons from DE, CM, and ECT, calculated by dPSI relative to UD. F. Gene ontology (GO) enrichment of genes with significantly changing alternative cassette/micro exons in CM, DE, or ECT relative to UD.

SUPPLEMENTAL FIGURE 1: A. Fluorescence-activated cell sorting (FACS) strategy showing density plots for day 3 differentiated DE cells (from right to left): unstained control, antibody isotype control, and three independent replicates for CXCR4:APC stained cells that were positively selected and from which RNA was extracted for RNA-seq. B. Indirect immunofluorescence of H9 hESCs with a nuclear localized humanized renilla GFP (hrGFP-NLS), showing hrGFP (green), SOX2 (magenta), SSEA3 (yellow), and DAPI (blue; SOX2, SSEA3, and DAPI in overlay) for UD; hrGFP (green), SOX17 (magenta), and DAPI (blue; SOX17 and DAPI overlay) for DE, hrGFP (green); NKX2.5 (magenta), and DAPI (blue; NKX2.5 and DAPI overlay) for CM; and hrGFP (green), SOX2 (magenta), and DAPI (blue; SOX2 and DAPI overlay) for ECT (scale bar = 100 μm). C. Principal component analysis plot of top 500 genes with the most variance between replicates of UD, DE, CM, and ECT, clustered by similarity (n = 3 each). D. Overlap of protein-coding transcript abundances significantly changing in CM, DE, and ECT relative to

UD from RNA-seq data. E. Bar charts showing total number of alternatively spliced events significantly changing in DE, CM, and ECT relative to UD hESCs by VAST-tools analysis of RNA-seq data. F. Comparison of dPSI's calculated by RT-PCR and BioAnalyzer quantitation (y-axis; see Fig 1D) versus VAST-tools analysis of RNA-seq (x-axis) data for each alternatively spliced exon in Fig 1F; R^2 values are shown for DE, CM, and ECT in the inset. G. Bar charts showing total number of alternatively spliced events significantly changing in DE relative to CM cells by VAST-tools analysis of RNA-seq data. H. Change in percent spliced in (dPSI) of alternative cassette and micro exons significantly changing in DE relative to CM cells shown in violin plot, and overlap of these compared to CM relative to UD and DE relative to UD shown in Venn diagram inset, calculated from RNA-seq data. I. dPSI values for retained introns (positive values indicate an increase in intron retention, lower values indicate a decrease) for DE, CM, and ECT, significantly changing relative to UD in violin plot, and overlap of these shown in Venn diagram inset as above (ns = not significant, and ****$P \leq 0.001$ by Mann-Whitney test). J. Heatmap showing hierarchical clustering of significantly changing retained introns from DE, CM, and ECT, calculated by column Z-score relative to UD. K. Change in percent spliced in (dPSI) of significantly changing retained introns for DE relative to CM cells shown in violin plot, and overlap of these compared to CM relative to UD and DE relative to UD shown in Venn diagram inset, calculated from RNA-seq data. L. Gene ontology (GO) enrichment of genes with significantly changing intron retention in CM relative to UD.

FIGURE 2: QKI5 increases in abundance and directly binds a significant subset of CM-specific alternatively spliced exons. A. Heatmap showing hierarchical clustering of RNA abundance measurements of annotated RBPs in DE, CM, and ECT, calculated by row Z-score relative to UD. B. Comparison and validation of RNA-seq data with RT-qPCR analysis of RNA extracted from DE, CM, and ECT cells displayed as $\log_2$ fold change relative to UD (ns = not significant, **$P \leq 0.001$, ***$P \leq 0.001$, ****$P \leq 0.0001$ by student's t-test comparing DE to CM). C. Western blot of protein extracted from triplicate cultures of UD, DE, and CM probed with antibodies for OCT4 (magenta), SOX17 (green),

NKX2.5 (magenta), and tubulin (gray). D. Western blot of protein extracted from UD, DE, and CM as above and probed with antibodies for tubulin (magenta) and panQKI (green), showing mean infrared signal +/- standard deviation for panQKI relative to tubulin (rel IR signal +/- SD) below (*$P \leq 0.05$, **$P \leq 0.01$ by student's t-test). E. RNA motif map (rMAPS) for QKI derived from rMATS analysis of alternatively spliced exons significantly changing when comparing CM relative to DE. F. Enrichment ($-\log_{10}$ p-value) of RBP eCLIP peaks mapped onto alternative exons either more included (above zero on y-axis in yellow) or skipped (below zero on y-axis in blue) in CM vs. DE and mapped positionally by upstream intron (left) within alternative exon (middle) or in the downstream intron (right), calculated by RELI analysis. G. UCSC Genome Browser screen shots showing QKI eCLIP peaks, Gencode Gene annotation for *NF1* (left) and *RAI14* (right), with RNA-seq coverage tracks for UD (black), DE (pink), CM (cyan), and ECT (yellow), and conservation of 100 vertebrates track.

SUPPLEMENTAL FIGURE 2: A. Western blot of protein extracted from UD, DE, CM (100 µg each) and decreasing volumes of C2C12 myoblasts probed with antibodies for Tublin (magenta) and MBNL1 and then subsequently panMBNL antibodies (green). B. Western blot of protein extracted from UD, DE, and CM probed with antibodies for QKI5 (magenta, top panel) and panQKI (green), QKI6 (magenta, middle panel) and panQKI (green), or QKI7 (magenta, bottom panel) and panQKI (green); quantitation of each isoform relative to total panQKI is shown below by bar graph. C. Indirect immunofluorescence of H9 hrGFP-NLS showing panQKI (magenta), SSEA3 (yellow), hrGFP (green), and each overlaid for UD; panQKI (magenta), SOX17 (yellow), and hrGFP (green), with each overlaid for DE; panQKI (magenta), NKX2.5 (yellow), and hrGFP (green), with each overlaid for CM (scalebar denotes 100µm). D. RNA motif map (rMAPS) for QKI derived from rMATS analysis of alternatively spliced exons significantly changing in ECT relative to UD. E. RNA motif map (rMAPS) for QKI derived from rMATS analysis of alternatively spliced exons significantly changing in CM relative to UD. F. RNA motif map (rMAPS) for QKI derived from rMATS analysis of alternatively spliced exons significantly changing in DE relative to UD. G. UCSC Genome Browser screen shots showing QKI eCLIP peaks, Gencode Gene annotation for *CLSTN1,*

PICALM, SLK, and *MARK3*, with RNA-seq coverage tracks for UD (black), DE (pink), CM (cyan), and ECT (yellow), and conservation of 100 vertebrates track.

FIGURE 3: Loss of QKI results in changes in splicing of CM-activated exons. A. Western blot of protein extracted from UD NKX2.5→EGFP cells WT, polyclonal QKI-KO (poly), 1-12 QKI-KO clone (1-12), 4-2 QKI-KO clone (4-2), and 4-5 QKI-KO clone (4-5) probed with antibodies for Tubulin (magenta) and panQKI antibodies (green; result shown is representative of 3 biological replicates). B. RT-PCR and BioAnalyzer gel-like image of RNA extracted from UD NKX2.5→EGFP cells described above to measure alternative exon inclusion for RAI14, BIN1, CLSTN1, and NF1 (result shown is representative of 3 biological replicates with mean percent included +/- standard deviation shown below). C. RT-qPCR of RNA extracted from UD H9 hESCs and measuring mean QKI5 RNA relative to HMBS mRNA, normalized to control shRNA (shNT), under shRNA knockdown for QKI with shQKI1 and shQKI2, showing 3 independent biological replicates as individual points (ns = not significant for shQKI1 (P = 0.0614), $^{**}P \leq 0.01$ for shQKI2 by student's t-test; n = 3 independent replicates; error bars denote standard deviation from the mean). D. Western blot of 3 independent biological replicates of protein extracted from UD H9 hESCs, transduced with shNT, shQKI1, and shQKI2 probed with antibodies for panQKI (green) and GAPDH (magenta), and showing panQKI protein abundance relative to GAPDH in bar graph below, normalized to GAPDH ($^{**}P \leq 0.01$ and $^{***}P \leq 0.001$ by student's t-test; n = 3 independent replicates). E. RT-PCR and BioAnalyzer gel-like image of RNA extracted from UD H9 hESCs transduced with shRNAs described above (in biological triplicate) to measure alternative exon inclusion of RAI14, BIN1, CLSTN1, and NF1 (the mean values +/- standard deviation are reported below). F. RT-PCR and BioAnalyzer gel-like image showing quantitation of RAI14, NF1, and BIN1 alternative exons using RNA extracted from WT or polyKO NKX2.5→EGFP cells either UD, or subjected to the following differentiation protocols: day 2 lateral mesoderm (d2 lat meso), day 4 CM, day 6 cardiomyocyte (d6 c.myo), and day 8 cardiomyocyte (d8 c.myo) (biological replicate n=3; mean percent

included +/- is shown below by heatmap; see Fig S3G for mean numerical values and standard deviation from the mean).

SUPPLEMENTAL FIGURE 3: A. Sequence decomposition analysis of PCR products generated by amplification of the 5' end of *QKI* genomic DNA by TIDE. Top panel shows the Indel Spectrum detected in polyclonal QKI KO NKX2.5→GFP cells: the total efficiency of mutation is 81.6% (sum of y-axis values) with rate of nucleotide number deleted (to the left) or inserted (to the right) on the x-axis; red bars denote $P < 0.001$, black $P \geq 0.001$. Lower panel shows percent aberrant sequence signal (y-axis) detected in WT (black) compared to QKI-KO, with expected bp cut site shown by nucleotide number on x-axis. Right panel shows percentages and p-values for a given number of nucleotides deleted (negative numbers) or inserted (positive numbers). B. Chromatograms showing reads from genomic DNA amplified by PCR then cloned and sequenced from NKX2.5→GFP WT and QKI KO cells: top panel shows WT sequence with guide RNA target noted in box. Next two panels show two independent sequences derived from polyclonal QKI KO cells (vertical box indicates nucleotide insertion); the following two panels show two independent sequences from single cell clone number 1-12 (deletion noted and vertical box indicates nucleotide insertion); the next two panels show two independent sequences from clone number 4-2 which has the same mutations as described in clone 1-12; the final two panels show two independent sequences from clone number 4-5 which has a single nucleotide insertion (vertical box) and 7 nucleotide deletion. C. RT-PCR and BioAnalyzer gel-like image of RNA extracted from UD NKX2.5→EGFP cells described (see Figs 3A and 3B) above to measure alternative exon inclusion for PICALM (result shown is representative of 3 biological replicates with mean percent included +/- standard deviation shown below). D. RT-PCR and agarose gel analysis of RNA extracted from UD NKX2.5→EGFP cells as described above to measure alternative exon inclusion for DNMT3B (result shown is representative of 3 biological replicates). E. RT-PCR and BioAnalyzer gel-like image of RNA extracted from UD H9 hESCs transduced in triplicate with shRNAs described above (see Fig 3E) to measure alternative exon inclusion of PICALM (the mean values +/- standard deviation are reported below). F. RT-PCR and agarose gel analysis of RNA extracted from UD H9 hESCs transduced in triplicate with shRNAs described above to measure

alternative exon inclusion of DNMT3B. G. Numerical mean values of percent included +/- standard deviation from the mean for data shown in Fig 3F.

FIGURE 4: QKI promotes skipping of BIN1 exon 7, exons 13+17, and exons 13-17 and total BIN1 transcript maturation through splicing. A. BIN1 transcript overview graphic: location of PCR and qPCR primers are noted, as are alternatively spliced exons and the tissue type in which they are correlated. B. Western blot of protein extracted from WT, polyKO, and 4-2KO NKX2.5→EGFP UD, d4 CM, and d8 c.myo cells probed with a BIN1 antibody that recognizes all isoforms (green) and GAPDH antibody (magenta); three independent biological replicates are shown, with mean protein levels relative to WT and normalized to GAPDH shown below +/- standard deviation from the mean (ns denotes not significant, *$P < 0.05$, and ***$P < 0.001$ by student's t-test). C. RT-PCR and BioAnalyzer gel-like image showing quantitation of BIN1 alternative exon 7 using RNA extracted from WT, polyKO, or 4-2KO NKX2.5→EGFP cells either UD, d4 CM, or d8 c.myo (biological replicate n=3; mean percent included +/- standard deviation from the mean is shown below). D. RT-PCR and BioAnalyzer gel-like image showing quantitation of BIN1 products amplified by primers annealing to exons 12 (forward) and 18 (reverse) using RNA extracted from WT, polyKO, or 4-2KO NKX2.5→EGFP cells either UD, d4 CM, or d8 c.myo (biological replicate n=3; mean percent product represented +/- standard deviation from the mean is shown below). E. Western blot of protein extracted from four independent replicates of HuH7 WT or HuH7 QKI-KO clone #3 cells transfected with pDUP-BIN1 exon 7 splicing reporter probed with panQKI antibody (top) or Tubulin antibody (bottom), with RT-PCR and BioAnalyzer gel-like image showing quantitation of splicing reporter products derived from RNA extracted from WT and KO#3 cells as above; below gel-like image are the mean values of each representative product +/- standard deviation from the mean. F. Overview graphic of pDUP-BIN1 exon 7 splicing reporter with WT and mutant sequences shown in inset below. G. RT-PCR and BioAnalyzer gel-like image showing quantitation of WT or mutant splicing reporter products derived from RNA extracts from WT HuH7 cells transfected in triplicate; as above, the mean value percentage of each representative product +/- standard deviation from the mean

is shown below the gel-like image. H. RT-qPCR of RNA extracted from d8 c.myo cells from NKX2.5→EGFP WT, polyKO, or 4-2KO measuring (from left to right) total exon 6, intron 6 included, exon 6-7 spliced, exon 6-8 spliced, exon 7-8 spliced, intron 7 included, and total exon 8 levels from the endogenous BIN1 transcript; each value is shown normalized to HMBS mRNA and relative to WT; graphic below bars indicate approximate location of forward and reverse primer pairs; n = 3 biological replicates, ns = not significant, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, relative to WT by student's t-test).

SUPPLEMENTAL FIGURE 4: A. UCSC Genome Browser screen shot showing QKI eCLIP from K562 cells, Qk iCLIP from C2C12 myoblasts, Gencode Genes with annotated BIN1 transcripts, unique full-length BIN1 cDNA clones sequenced (Sanger sequencing) from UD and d8 c.myo WT and QKI-KO NKX2.5→EGFP cells (KO11 represents one full-length clone from two different forward and reverse sequencing reactions), coverage tracks for UD, DE, CM, and ECT RNA-seq datasets, and conservation tracks. B. RT-qPCR of RNA extracted from NKX2.5→EGFP WT, polyKO, or 4-2KO cells from UD, d4CM, or d8 c.myo, and measuring either the 5' end of BIN1 (primer pair spanning exons 4 and 5) or the 3' end of BIN1 (primer pair spanning exons 19 and 20); value shown is ddCt relative to HMBS; n = 2 or 3 independent replicates, ns = not significant, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ by student's t-test compared to WT. C. BIN1 infrared protein signal relative to GAPDH infrared protein signal from western blot shown in Fig 4B (n = 3 independent replicates, ns = not significant, *$P < 0.05$, ***$P < 0.001$ by student's t-test compared to WT). D. BIN1 fluorescence intensity relative to GAPDH per well (n = 2), measured by high content imaging in UD, d4 CM, and d8 c.myo NKX2.5→EGFP WT, polyKO, or 4-2KO cells (ns = not significant, *$P < 0.05$, **$P < 0.01$ by student's t-test compared to WT). E. BIN1 fluorescence intensity relative to GAPDH per cell (n = ~50,000 per cell type) in d8 c.myo NKX2.5→EGFP WT, polyKO, or 4-2KO cells (****$P < 0.0001$ by Mann Whitney U).

FIGURE 5: QKI is required for efficient cardiomyocyte cell differentiation. A. Western blot of protein extracted from NKX2.5→GFP UD and day 8 cardiomyocyte (d8 c.myo) cells either wild-type (WT) or lacking QKI (polyclonal KO "pKO") and probed with antibodies for tubulin (magenta) and panQKI (green).

B. Population-level EGFP abundance (relative to GAPDH) calculated per well, measured by high content imaging in WT, polyclonal QKI-KO (polyKO), and clone 4-2 QKI-KO (4-2KO) in UD, day 4 CM (d4 CM), and day 8 cardiomyocytes (d8 c.myo) NKX2.5→EGFP cells (n = 4 independent replicates; ns = not significant, *$P < 0.05$, and **$P < 0.01$ relative to WT, determined by student's t-test). C. Flow cytometry analysis of NKX2.5→GFP UD (gray), WT, or KO day 8 c.myo (light or dark red, respectively); mean percent GFP positive is shown above each histogram (+/- standard deviation; biological replicate n = 3) and ****$P < 0.0001$ measured by student's t-test; y-axis denotes number of events measured. D. RT-qPCR measuring abundance of RNA extracted from UD, d4 CM, and d8 c.myo for endogenous NKX2.5, and TNNT2 (cTNT) mRNAs relative to HMBS mRNA in WT, polyKO, or 4-2KO cells (biological replicate n = 2 or 3 each; ns = not significant, *$P < 0.05$, **$P < 0.01$, measured by student's t-test). E. Western blot of protein extracted from NKX2.5→GFP UD and d8 c.myo cells either WT or pKO and probed with antibodies for tubulin (green, top) cTNT (magenta, middle (TNNT2)), and endogenous NKX2.5 (green, bottom). F. Indirect immunofluorescence with high content imaging analysis of NKX2.5→GFP WT (left) or QKI polyclonal KO (right) for d8 cardiomyocyte cells showing QKI5 (magenta), cTNT (yellow), EGFP (green), brightfield (gray), and DAPI (blue) with each fluorescent channel overlaid (left). On the right, the fluorescence intensity values per cell (relative to DAPI) is shown for QKI5 and cTNT with *P* values calculated by Mann-Whitney U.

SUPPLEMENTAL FIGURE 5: A. RT-qPCR measuring abundance of RNA extracted from UD, d2 lat meso, d4 CM, d6 c.myo, and d8 c.myo for QKI5 RNA relative to HMBS mRNA in WT or QKI-KO cells (biological replicate n = 3 each; ns = not significant, *$P \leq 0.05$, **$P \leq 0.01$ measured by student's t-test). B. RT-qPCR measuring abundance of RNA extracted from NKX2.5→EGFP UD, d4 CM, and d8 c.myo cells for ISL1, GATA4, HAND1, and MEF2C mRNAs relative to HMBS mRNA in WT, polyKO, or 4-2KO cells (biological replicate n = 2 or 3 each; ns = not significant, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$ measured by student's t-test). C. Indirect immunofluorescence of UD NKX2.5→GFP WT (left) or polyclonal QKI-KO (right) UD hESCs analyzed by high content imaging, showing QKI5 (magenta), OCT4 (yellow), EGFP (green), brightfield (gray), and DAPI (blue) with each fluorescent channel overlaid (left

side of figure). On the right, the fluorescence intensity per cell (relative to DAPI) is shown for QKI5 and OCT4 per cell with *P* values calculated by Mann-Whitney U. D. Indirect immunofluorescence of NKX2.5→GFP WT (left) or polyclonal QKI-KO (right) d4 CM cells analyzed by high content imaging, showing panQKI (magenta), endogenous NKX2.5 (yellow), EGFP (green), brightfield (gray), and DAPI (blue) with each fluorescent channel overlaid (left). On the right, the fluorescent intensity per cell (relative to DAPI) is shown for panQKI and NKX2.5 with *P* values calculated by Mann-Whitney U. E. Per well population values of fluorescent intensities measured by indirect immunofluorescence as described in Fig 4F for QKI5 and cTNT (TNNT2) each normalized to DAPI in WT and polyclonal QKI-KO day 8 c.myo NKX2.5→GFP cells (n = 3 independent biological replicates; ***$P \leq 0.001$ by student's t-test). F. Low magnification epifluorescence imaging with FITC filter of live d8 c.myo NKX2.5→EGFP cells WT, polyKO, and 4-2KO (scale bar denotes 200 µm).

References

(2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.
Ang, L.T., Tan, A.K.Y., Autio, M.I., Goh, S.H., Choo, S.H., Lee, K.L., Tan, J., Pan, B., Lee, J.J.H., Lum, J.J.*, et al.* (2018). A Roadmap for Human Liver Differentiation from Pluripotent Stem Cells. Cell Rep *22*, 2190-2205.
Baehrecke, E.H. (1997). who encodes a KH RNA binding protein that functions in muscle development. Development (Cambridge, England) *124*, 1323-1332.
Baralle, F.E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. Nat Rev Mol Cell Biol *18*, 437-451.
Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. Nature *465*, 53-59.
Baxter, M.A., Rowe, C., Alder, J., Harrison, S., Hanley, K.P., Park, B.K., Kitteringham, N.R., Goldring, C.E., and Hanley, N.A. (2010). Generating hepatic cell lineages from pluripotent stem cells for drug toxicity screening. Stem Cell Res *5*, 4-22.
Brannan, C.I., Perkins, A.S., Vogel, K.S., Ratner, N., Nordlund, M.L., Reid, S.W., Buchberg, A.M., Jenkins, N.A., Parada, L.F., and Copeland, N.G. (1994). Targeted disruption of the neurofibromatosis type-1 gene leads to developmental abnormalities in heart and various neural crest-derived tissues. Genes & development *8*, 1019-1029.
Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res *42*, e168.
Butler, M.H., David, C., Ochoa, G.C., Freyberg, Z., Daniell, L., Grabs, D., Cremona, O., and De Camilli, P. (1997). Amphiphysin II (SH3P9; BIN1), a member of the amphiphysin/Rvs family, is concentrated in the cortical cytomatrix of axon initial segments and nodes of ranvier in brain and around T tubules in skeletal muscle. J Cell Biol *137*, 1355-1367.
Chapuis, J., Hansmannel, F., Gistelinck, M., Mounier, A., Van Cauwenberghe, C., Kolen, K.V., Geller, F., Sottejeau, Y., Harold, D., Dourlen, P.*, et al.* (2013). Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. Mol Psychiatry *18*, 1225-1234.
Cook, K.B., Kazan, H., Zuberi, K., Morris, Q., and Hughes, T.R. (2011). RBPDB: a database of RNA-binding specificities. Nucleic Acids Res *39*, D301-308.
Darbelli, L., Vogel, G., Almazan, G., and Richard, S. (2016). Quaking Regulates Neurofascin 155 Expression for Myelin and Axoglial Junction Maintenance. J Neurosci *36*, 4106-4120.
de Bruin, R.G., Shiue, L., Prins, J., de Boer, H.C., Singh, A., Fagg, W.S., van Gils, J.M., Duijs, J.M., Katzman, S., Kraaijeveld, A.O.*, et al.* (2016). Quaking promotes monocyte differentiation into pro-atherogenic macrophages by controlling pre-mRNA splicing and gene expression. Nature communications *7*, 10846.
De Rossi, P., Buggia-Prévot, V., Clayton, B.L., Vasquez, J.B., van Sanford, C., Andrew, R.J., Lesnick, R., Botté, A., Deyts, C., Salem, S.*, et al.* (2016). Predominant expression of Alzheimer's disease-associated BIN1 in mature oligodendrocytes and localization to white matter tracts. Mol Neurodegener *11*, 59.
Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.
Dominski, Z., and Kole, R. (1991). Selection of splice sites in pre-mRNAs with short internal exons. Mol Cell Biol *11*, 6075-6083.
Duymich, C.E., Charlet, J., Yang, X., Jones, P.A., and Liang, G. (2016). DNMT3B isoforms without catalytic activity stimulate gene body methylation as accessory proteins in somatic cells. Nature communications *7*, 11453.

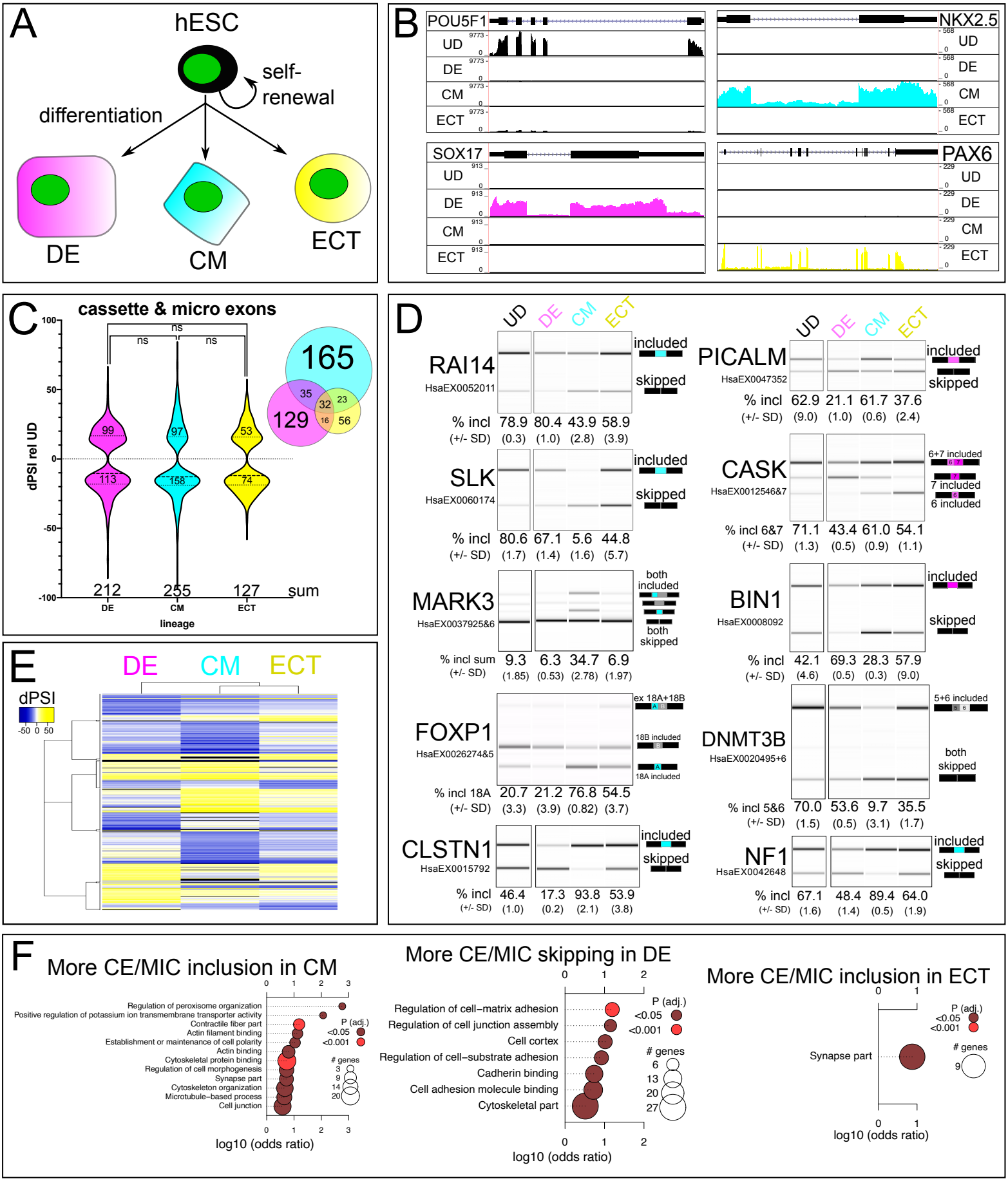
Ebersole, T.A., Chen, Q., Justice, M.J., and Artzt, K. (1996). The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins. Nat Genet *12*, 260-265.
Elliott, D.A., Braam, S.R., Koutsis, K., Ng, E.S., Jenny, R., Lagerqvist, E.L., Biben, C., Hatzistavrou, T., Hirst, C.E., Yu, Q.C.*, et al.* (2011). NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. Nat Methods *8*, 1037-1040.
Ellis, J.D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J.A., Wang, X., Pan, Q., O'Hanlon, D., Kim, P.M.*, et al.* (2012). Tissue-specific alternative splicing remodels protein-protein interaction networks. Molecular cell *46*, 884-892.
Fagg, W.S., Liu, N., Fair, J.H., Shiue, L., Katzman, S., Donohue, J.P., and Ares, M., Jr. (2017). Autogenous cross-regulation of Quaking mRNA processing and translation balances Quaking functions in splicing and translation. Genes & development *31*, 1894-1909.
Fu, X.D., and Ares, M., Jr. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. Nature reviews Genetics *15*, 689-701.
Fugier, C., Klein, A.F., Hammer, C., Vassilopoulos, S., Ivarsson, Y., Toussaint, A., Tosch, V., Vignaud, A., Ferry, A., Messaddeq, N.*, et al.* (2011). Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. Nat Med *17*, 720-725.
Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.K., Alvarez, M., Talukder, S., Pan, Q., Mazzoni, E.O.*, et al.* (2011). An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. Cell *147*, 132-146.
Galarneau, A., and Richard, S. (2005). Target RNA motif and target mRNAs of the Quaking STAR protein. Nat Struct Mol Biol *12*, 691-698.
Gazzara, M.R., Vaquero-Garcia, J., Lynch, K.W., and Barash, Y. (2014). In silico to in vivo splicing analysis using splicing code models. Methods *67*, 3-12.
Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R.*, et al.* (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell *153*, 1149-1163.
Gordon, C.A., Hartono, S.R., and Chédin, F. (2013). Inactive DNMT3B splice variants modulate de novo DNA methylation. PloS one *8*, e69486.
Gupta, S.K., Garg, A., Bär, C., Chatterjee, S., Foinquinos, A., Milting, H., Streckfuß-Bömeke, K., Fiedler, J., and Thum, T. (2018). Quaking Inhibits Doxorubicin-Mediated Cardiotoxicity Through Regulation of Cardiac Circular RNA Expression. Circ Res *122*, 246-254.
Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M.*, et al.* (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell *141*, 129-141.
Hall, M.P., Nagel, R.J., Fagg, W.S., Shiue, L., Cline, M.S., Perriman, R.J., Donohue, J.P., and Ares, M., Jr. (2013). Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation. RNA (New York, NY) *19*, 627-638.
Hamid, F.M., and Makeyev, E.V. (2014). Emerging functions of alternative splicing coupled with nonsense-mediated decay. Biochem Soc Trans *42*, 1168-1173.
Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N.*, et al.* (2013). MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature *498*, 241-245.
Hardy, R.J., Loushin, C.L., Friedrich, V.L., Jr., Chen, Q., Ebersole, T.A., Lazzarini, R.A., and Artzt, K. (1996). Neural cell type-specific expression of QKI proteins is altered in quakingviable mutant mice. J Neurosci *16*, 7941-7949.
Harley, J.B., Chen, X., Pujato, M., Miller, D., Maddox, A., Forney, C., Magnusen, A.F., Lynch, A., Chetal, K., Yukawa, M.*, et al.* (2018). Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. Nat Genet *50*, 699-707.

Hong, T., Yang, H., Zhang, S.S., Cho, H.C., Kalashnikova, M., Sun, B., Zhang, H., Bhargava, A., Grabe, M., Olgin, J.*, et al.* (2014). Cardiac BIN1 folds T-tubule membrane, controlling ion flux and limiting arrhythmia. Nat Med *20*, 624-632.
Hong, T.T., Smyth, J.W., Gao, D., Chu, K.Y., Vogan, J.M., Fong, T.S., Jensen, B.C., Colecraft, H.M., and Shaw, R.M. (2010). BIN1 localizes the L-type calcium channel to cardiac T-tubules. PLoS Biol *8*, e1000312.
Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallières, M., Tapial, J., Raj, B., O'Hanlon, D.*, et al.* (2014). A highly conserved program of neuronal microexons is misregulated in autistic brains. Cell *159*, 1511-1523.
Jangi, M., and Sharp, P.A. (2014). Building robust transcriptomes with master splicing factors. Cell *159*, 487-498.
Kondo, T., Furuta, T., Mitsunaga, K., Ebersole, T.A., Shichiri, M., Wu, J., Artzt, K., Yamamura, K., and Abe, K. (1999). Genomic organization and expression analysis of the mouse qkI locus. Mamm Genome *10*, 662-669.
Larocque, D., Pilotte, J., Chen, T., Cloutier, F., Massie, B., Pedraza, L., Couture, R., Lasko, P., Almazan, G., and Richard, S. (2002). Nuclear retention of MBP mRNAs in the quaking viable mice. Neuron *36*, 815-829.
Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in disease. Cell *152*, 1237-1251.
Li, Z., Takakura, N., Oike, Y., Imanaka, T., Araki, K., Suda, T., Kaname, T., Kondo, T., Abe, K., and Yamamura, K. (2003). Defective smooth muscle development in qkI-deficient mice. Dev Growth Differ *45*, 449-462.
Li, Z., Zhang, Y., Li, D., and Feng, Y. (2000). Destabilization and mislocalization of myelin basic protein mRNAs in quaking dysmyelination lacking the QKI RNA-binding proteins. J Neurosci *20*, 4944-4953.
Liao, K.C., Chuo, V., Fagg, W.S., Bradrick, S.S., Pompon, J., and Garcia-Blanco, M.A. (2020). The RNA binding protein Quaking represses host interferon response by downregulating MAVS. RNA Biol *17*, 366-380.
Lobbardi, R., Lambert, G., Zhao, J., Geisler, R., Kim, H.R., and Rosa, F.M. (2011). Fine-tuning of Hh signaling by the RNA-binding protein Quaking to control muscle development. Development (Cambridge, England) *138*, 1783-1794.
Loh, K.M., Ang, L.T., Zhang, J., Kumar, V., Ang, J., Auyeong, J.Q., Lee, K.L., Choo, S.H., Lim, C.Y., Nichane, M.*, et al.* (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. Cell stem cell *14*, 237-252.
Loh, K.M., Chen, A., Koh, P.W., Deng, T.Z., Sinha, R., Tsai, J.M., Barkal, A.A., Shen, K.Y., Jain, R., Morganti, R.M.*, et al.* (2016). Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell *166*, 451-467.
Mahla, R.S. (2016). Stem Cells Applications in Regenerative Medicine and Disease Therapeutics. Int J Cell Biol *2016*, 6940283.
Muller, A.J., Baker, J.F., DuHadaway, J.B., Ge, K., Farmer, G., Donover, P.S., Meade, R., Reid, C., Grzanna, R., Roach, A.H.*, et al.* (2003). Targeted disruption of the murine Bin1/Amphiphysin II gene does not disable endocytosis but results in embryonic cardiomyopathy with aberrant myofibril formation. Mol Cell Biol *23*, 4295-4306.
Murry, C.E., and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell *132*, 661-680.
Nabel-Rosen, H., Dorevitch, N., Reuveny, A., and Volk, T. (1999). The balance between two isoforms of the Drosophila RNA-binding protein how controls tendon cell differentiation. Molecular cell *4*, 573-584.
Noveroske, J.K., Lai, L., Gaussin, V., Northrop, J.L., Nakamura, H., Hirschi, K.K., and Justice, M.J. (2002). Quaking is essential for blood vessel development. Genesis *32*, 218-230.
O'Shea, K.S. (1999). Embryonic stem cell models of development. Anat Rec *257*, 32-41.

Ohno, G., Hagiwara, M., and Kuroyanagi, H. (2008). STAR family RNA-binding protein ASD-2 regulates developmental switching of mutually exclusive alternative splicing in vivo. Genes & development *22*, 360-374.
Prokic, I., Cowling, B.S., and Laporte, J. (2014). Amphiphysin 2 (BIN1) in physiology and diseases. J Mol Med (Berl) *92*, 453-463.
Ramjaun, A.R., Micheva, K.D., Bouchelet, I., and McPherson, P.S. (1997). Identification and characterization of a nerve terminal-enriched amphiphysin isoform. The Journal of biological chemistry *272*, 16700-16706.
Rizzino, A. (2009). Sox2 and Oct-3/4: a versatile pair of master regulators that orchestrate the self-renewal and pluripotency of embryonic stem cells. Wiley Interdiscip Rev Syst Biol Med *1*, 228-236.
Rossi, F.M., Kringstein, A.M., Spicher, A., Guicherit, O.M., and Blau, H.M. (2000). Transcriptional control: rheostat converted to on/off switch. Molecular cell *6*, 723-728.
Ryder, S.P., and Williamson, J.R. (2004). Specificity of the STAR/GSG domain protein Qk1: implications for the regulation of myelination. RNA (New York, NY) *10*, 1449-1458.
Saccomanno, L., Loushin, C., Jan, E., Punkay, E., Artzt, K., and Goodwin, E.B. (1999). The STAR protein QKI-6 is a translational repressor. Proceedings of the National Academy of Sciences of the United States of America *96*, 12605-12610.
Salomonis, N., Nelson, B., Vranizan, K., Pico, A.R., Hanspers, K., Kuchinsky, A., Ta, L., Mercola, M., and Conklin, B.R. (2009). Alternative splicing in the differentiation of human embryonic stem cells into cardiac precursors. PLoS Comput Biol *5*, e1000553.
Smukler, S.R., Runciman, S.B., Xu, S., and van der Kooy, D. (2006). Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. J Cell Biol *172*, 79-90.
Tapial, J., Ha, K.C.H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallières, M., Permanyer, J., Sodaei, R., Marquez, Y.*, et al.* (2017). An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. Genome Res *27*, 1759-1768.
Tchieu, J., Zimmer, B., Fattahi, F., Amin, S., Zeltner, N., Chen, S., and Studer, L. (2017). A Modular Platform for Differentiation of Human PSCs into All Major Ectodermal Lineages. Cell stem cell *21*, 399-410.e397.
Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.
Tsankov, A.M., Gu, H., Akopian, V., Ziller, M.J., Donaghey, J., Amit, I., Gnirke, A., and Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. Nature *518*, 344-349.
Ule, J., and Blencowe, B.J. (2019). Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. Molecular cell *76*, 329-345.
Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature *444*, 580-586.
van der Veer, E.P., de Bruin, R.G., Kraaijeveld, A.O., de Vries, M.R., Bot, I., Pera, T., Segers, F.M., Trompet, S., van Gils, J.M., Roeten, M.K.*, et al.* (2013). Quaking, an RNA-binding protein, is a critical regulator of vascular smooth muscle cell phenotype. Circ Res *113*, 1065-1075.
Van Emburgh, B.O., and Robertson, K.D. (2011). Modulation of Dnmt3b function in vitro by interactions with Dnmt3L, Dnmt3a and Dnmt3b splice variants. Nucleic Acids Res *39*, 4984-5002.
Venables, J.P., Lapasset, L., Gadea, G., Fort, P., Klinck, R., Irimia, M., Vignal, E., Thibault, P., Prinos, P., Chabot, B.*, et al.* (2013). MBNL1 and RBFOX2 cooperate to establish a splicing programme involved in pluripotent stem cell differentiation. Nature communications *4*, 2480.
Vuong, C.K., Black, D.L., and Zheng, S. (2016). The neurogenetics of alternative splicing. Nat Rev Neurosci *17*, 265-281.
Wechsler-Reya, R., Sakamuro, D., Zhang, J., Duhadaway, J., and Prendergast, G.C. (1997). Structural analysis of the human BIN1 gene. Evidence for tissue-specific transcriptional regulation and alternate RNA splicing. The Journal of biological chemistry *272*, 31453-31458.

Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D.*, et al.* (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell *153*, 1134-1148.
Yamazaki, T., Liu, L., Lazarev, D., Al-Zain, A., Fomin, V., Yeung, P.L., Chambers, S.M., Lu, C.W., Studer, L., and Manley, J.L. (2018). TCF3 alternative splicing controlled by hnRNP H/F regulates E-cadherin expression and hESC pluripotency. Genes & development *32*, 1161-1174.
Yang, J.H., Li, J.H., Shao, P., Zhou, H., Chen, Y.Q., and Qu, L.H. (2011). starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res *39*, D202-209.
Yang, Y.C., Di, C., Hu, B., Zhou, M., Liu, Y., Song, N., Li, Y., Umetsu, J., and Lu, Z.J. (2015). CLIPdb: a CLIP-seq database for protein-RNA interactions. BMC Genomics *16*, 51.
Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.D., and Gage, F.H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol *16*, 130-137.
Yeo, G.W., Xu, X., Liang, T.Y., Muotri, A.R., Carson, C.T., Coufal, N.G., and Gage, F.H. (2007). Alternative splicing events identified in human embryonic stem cells and neural progenitors. PLoS Comput Biol *3*, 1951-1967.
Zaffran, S., Astier, M., Gratecos, D., and Sémériva, M. (1997). The held out wings (how) Drosophila gene encodes a putative RNA-binding protein involved in the control of muscular and cardiac activity. Development (Cambridge, England) *124*, 2087-2098.
Zhang, T., Lin, Y., Liu, J., Zhang, Z.G., Fu, W., Guo, L.Y., Pan, L., Kong, X., Zhang, M.K., Lu, Y.H.*, et al.* (2016). Rbm24 Regulates Alternative Splicing Switch in Embryonic Stem Cell Cardiac Lineage Differentiation. Stem cells (Dayton, Ohio) *34*, 1776-1789.
Zhao, L., Mandler, M.D., Yi, H., and Feng, Y. (2010). Quaking I controls a unique cytoplasmic pathway that regulates alternative splicing of myelin-associated glycoprotein. Proceedings of the National Academy of Sciences of the United States of America *107*, 19061-19066.

FIGURE 1


A
hESC
self-renewal
differentiation
DE
CM
ECT
B
POU5F1
NKX2.5
SOX17
PAX6
UD
DE
CM
ECT
C
cassette & micro exons
dPSI rel UD
lineage
212
255
127
sum
165
129
D
RAI14
HsaEX0052011
% incl 78.9 80.4 43.9 58.9
(+/- SD) (0.3) (1.0) (2.8) (3.9)
SLK
HsaEX0060174
% incl 80.6 67.1 5.6 44.8
(+/- SD) (1.7) (1.4) (1.6) (5.7)
MARK3
HsaEX0037925&6
% incl sum 9.3 6.3 34.7 6.9
(+/- SD) (1.85) (0.53) (2.78) (1.97)
FOXP1
HsaEX0026274&5
% incl 18A 20.7 21.2 76.8 54.5
(+/- SD) (3.3) (3.9) (0.82) (3.7)
CLSTN1
HsaEX0015792
% incl 46.4 17.3 93.8 53.9
(+/- SD) (1.0) (0.2) (2.1) (3.8)
PICALM
HsaEX0047352
% incl 62.9 21.1 61.7 37.6
(+/- SD) (9.0) (1.0) (0.6) (2.4)
CASK
HsaEX0012546&7
% incl 6&7 71.1 43.4 61.0 54.1
(+/- SD) (1.3) (0.5) (0.9) (1.1)
BIN1
HsaEX0008092
% incl 42.1 69.3 28.3 57.9
(+/- SD) (4.6) (0.5) (0.3) (9.0)
DNMT3B
HsaEX0020495+6
% incl 5&6 70.0 53.6 9.7 35.5
(+/- SD) (1.5) (0.5) (3.1) (1.7)
NF1
HsaEX0042648
% incl 67.1 48.4 89.4 64.0
(+/- SD) (1.6) (1.4) (0.5) (1.9)
E
DE CM ECT
dPSI
F
More CE/MIC inclusion in CM
Regulation of peroxisome organization
Positive regulation of potassium ion transmembrane transporter activity
Contractile fiber part
Actin filament binding
Establishment or maintenance of cell polarity
Actin binding
Cytoskeletal protein binding
Regulation of cell morphogenesis
Synapse part
Cytoskeleton organization
Microtubule-based process
Cell junction
log10 (odds ratio)
More CE/MIC skipping in DE
Regulation of cell-matrix adhesion
Regulation of cell junction assembly
Cell cortex
Regulation of cell-substrate adhesion
Cadherin binding
Cell adhesion molecule binding
Cytoskeletal part
log10 (odds ratio)
More CE/MIC inclusion in ECT
Synapse part
log10 (odds ratio)

SUPPLEMENTAL FIGURE 1

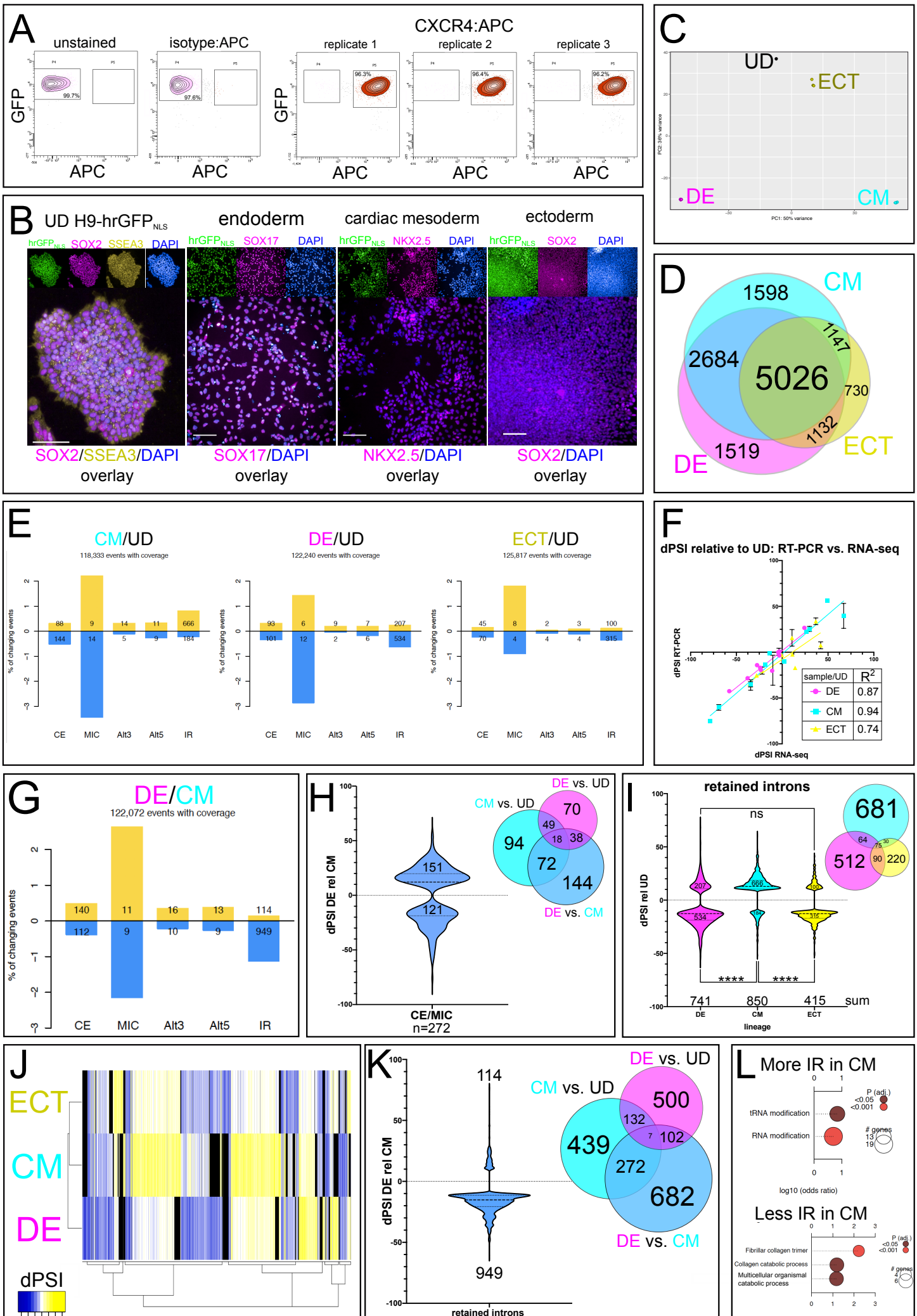
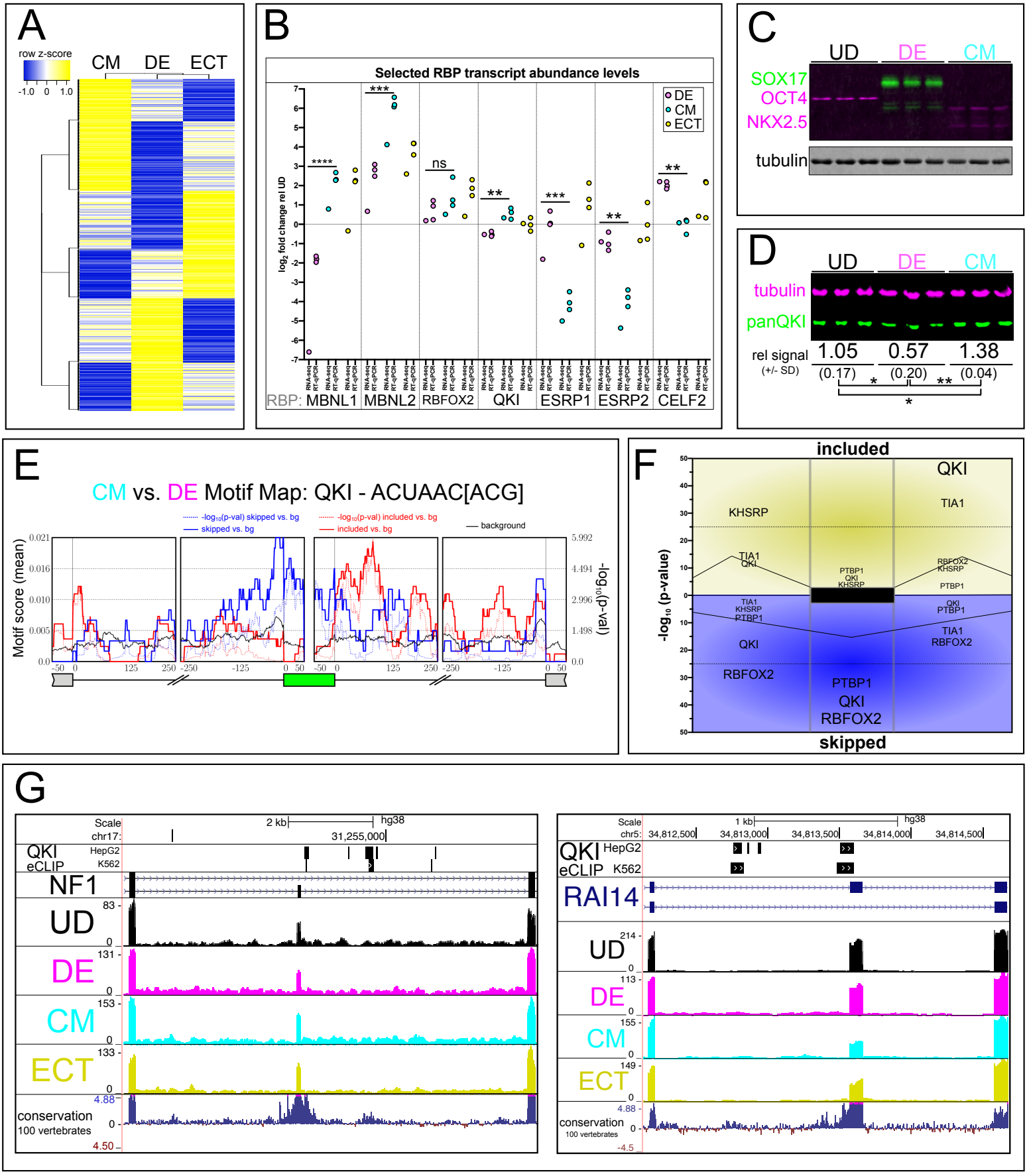


FIGURE 2


A
row z-score
-1.0 0 1.0
CM DE ECT
B
Selected RBP transcript abundance levels
log2 fold change rel UD
DE
CM
ECT
RBP: MBNL1 MBNL2 RBFOX2 QKI ESRP1 ESRP2 CELF2
C
UD DE CM
SOX17
OCT4
NKX2.5
tubulin
D
UD DE CM
tubulin
panQKI
rel signal (+/- SD)
1.05 (0.17)
0.57 (0.20)
1.38 (0.04)
E
CM vs. DE Motif Map: QKI - ACUAAC[ACG]
-log10(p-val) skipped vs. bg
skipped vs. bg
-log10(p-val) included vs. bg
included vs. bg
background
Motif score (mean)
-log10(p-val)
F
included
skipped
-log10 (p-value)
QKI
TIA1
KHSRP
RBFOX2
PTBP1
G
Scale
chr17:
2 kb
hg38
31,255,000
QKI eCLIP
HepG2
K562
NF1
UD
DE
CM
ECT
conservation 100 vertebrates
chr5:
1 kb
34,812,500
34,813,000
34,813,500
34,814,000
34,814,500
RAI14

SUPPLEMENTAL FIGURE 2

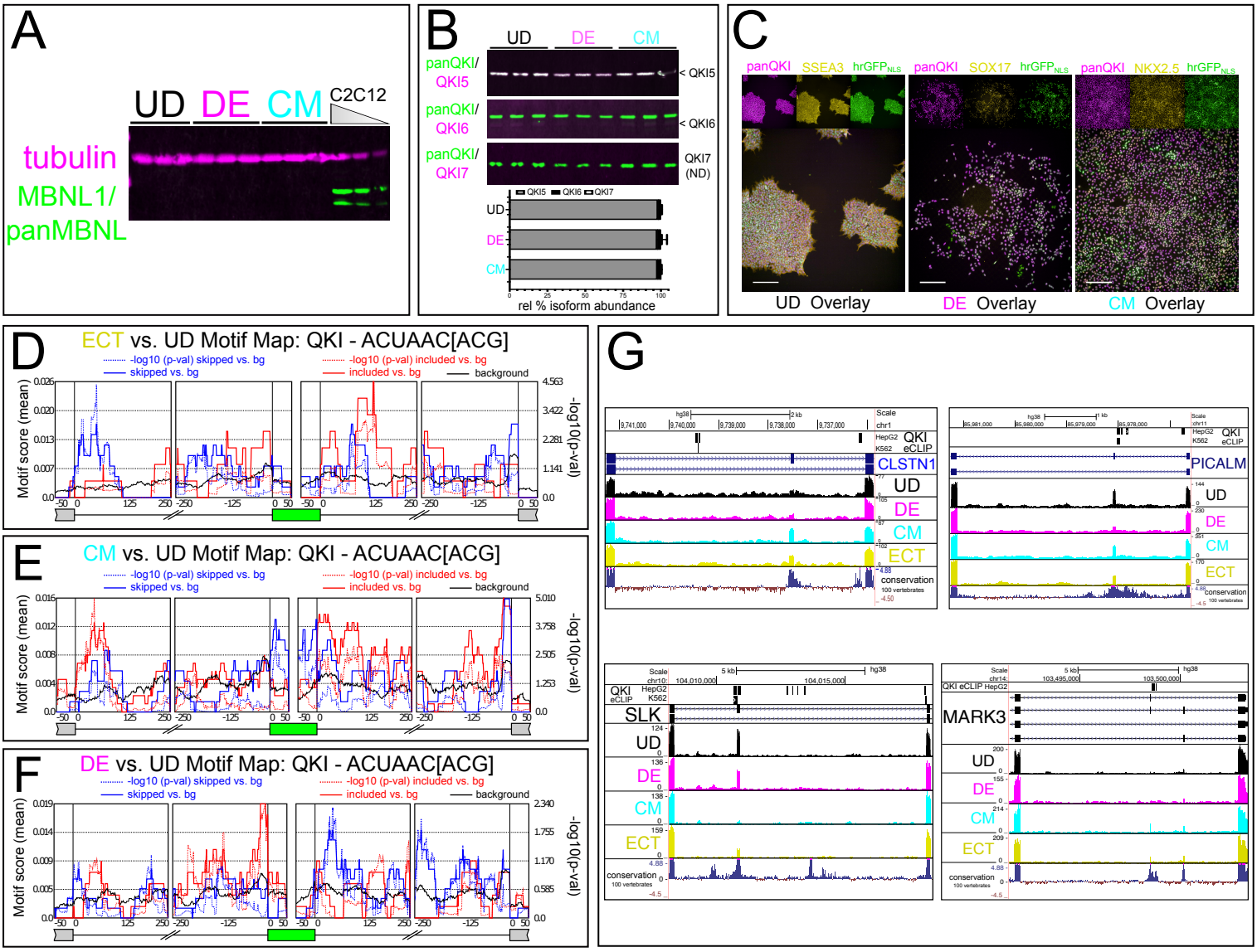
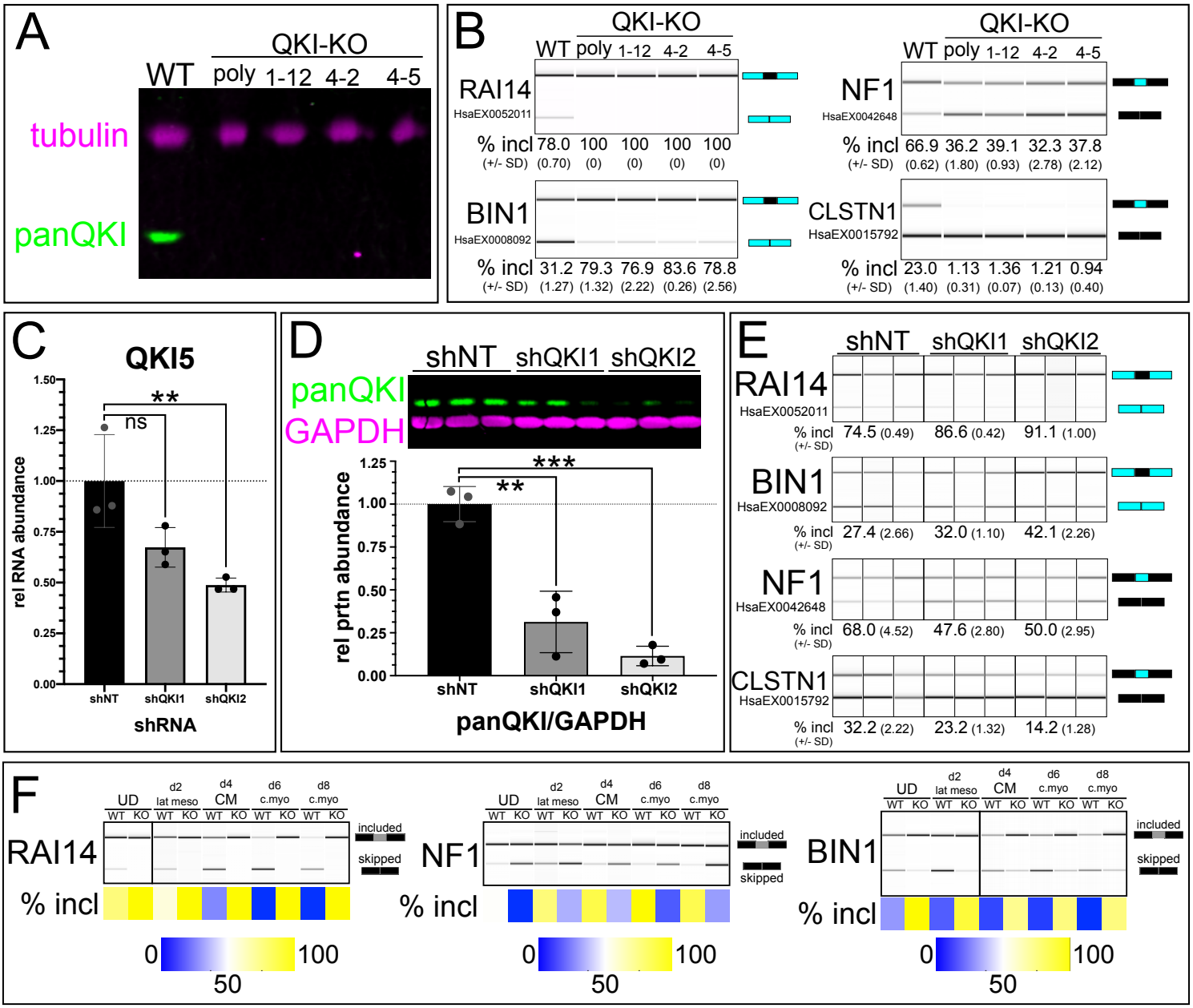
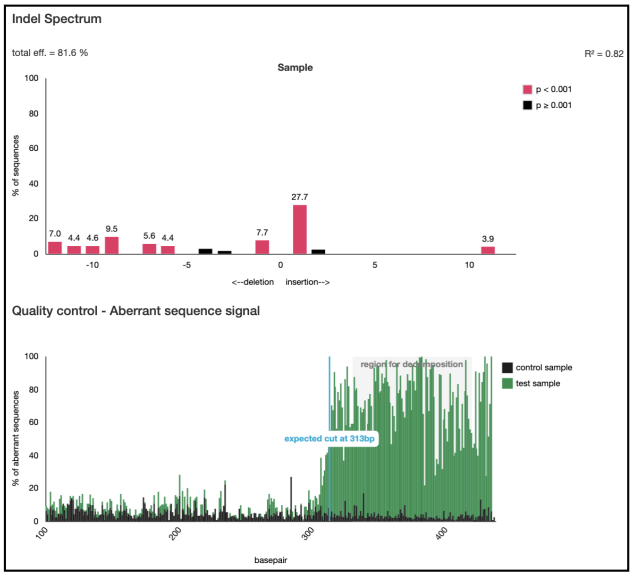


FIGURE 3


A
QKI-KO
WT poly 1-12 4-2 4-5
tubulin
panQKI
B
QKI-KO
WT poly 1-12 4-2 4-5
RAI14
HsaEX0052011
% incl 78.0 100 100 100 100
(+/- SD) (0.70) (0) (0) (0) (0)
BIN1
HsaEX0008092
% incl 31.2 79.3 76.9 83.6 78.8
(+/- SD) (1.27) (1.32) (2.22) (0.26) (2.56)
NF1
HsaEX0042648
% incl 66.9 36.2 39.1 32.3 37.8
(+/- SD) (0.62) (1.80) (0.93) (2.78) (2.12)
CLSTN1
HsaEX0015792
% incl 23.0 1.13 1.36 1.21 0.94
(+/- SD) (1.40) (0.31) (0.07) (0.13) (0.40)
C
QKI5
rel RNA abundance
1.50
1.25
1.00
0.75
0.50
0.25
0.00
**
ns
shNT shQKI1 shQKI2
shRNA
D
shNT shQKI1 shQKI2
panQKI
GAPDH
rel prtn abundance
1.25
1.00
0.75
0.50
0.25
0.00
***
**
shNT shQKI1 shQKI2
panQKI/GAPDH
E
shNT shQKI1 shQKI2
RAI14
HsaEX0052011
% incl 74.5 (0.49) 86.6 (0.42) 91.1 (1.00)
(+/- SD)
BIN1
HsaEX0008092
% incl 27.4 (2.66) 32.0 (1.10) 42.1 (2.26)
(+/- SD)
NF1
HsaEX0042648
% incl 68.0 (4.52) 47.6 (2.80) 50.0 (2.95)
(+/- SD)
CLSTN1
HsaEX0015792
% incl 32.2 (2.22) 23.2 (1.32) 14.2 (1.28)
(+/- SD)
F
UD d2 lat meso d4 CM d6 c.myo d8 c.myo
WT KO WT KO WT KO WT KO WT KO
RAI14
included
skipped
% incl
0 100
50
NF1
included
skipped
% incl
0 100
50
BIN1
included
skipped
% incl
0 100
50

SUPPLEMENTAL FIGURE 3

A

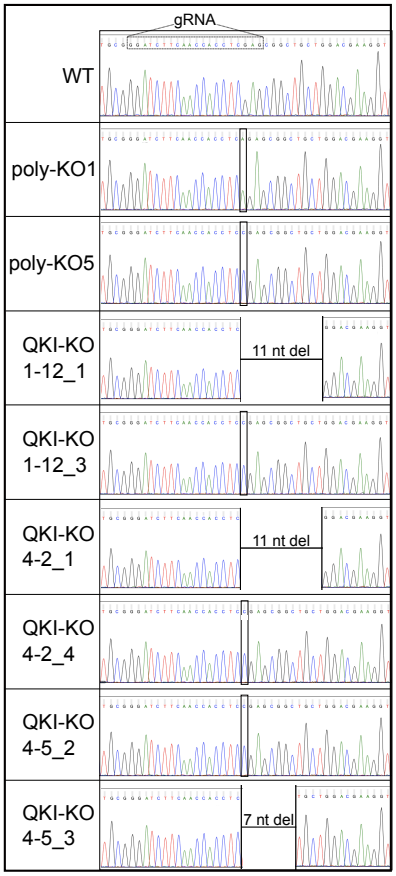


Quantification Indel Frequencies

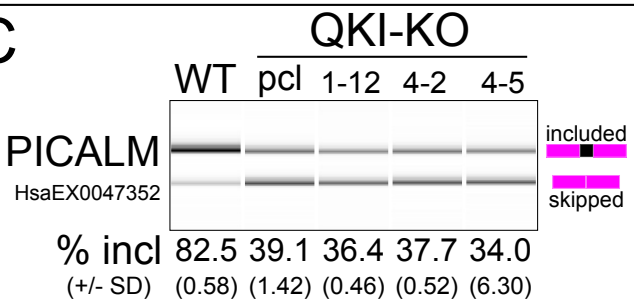
overall efficiency = 81.6 %

	percentage	pvalue
-12	7.0	1.1e-14
-11	4.4	1e-06
-10	4.6	6.2e-07
-9	9.5	1.1e-24
-8	0.0	1
-7	5.6	1.6e-09
-6	4.4	1.8e-06
-5	0.0	1
-4	2.9	0.0019
-3	1.5	0.11
-2	0.0	1
-1	7.7	3.3e-16
0	0.0	1
+1	27.7	1.3e-198
+2	2.3	0.012
+3	0.0	1
+4	0.2	0.81
+5	0.1	0.92
+6	0.0	1
+7	0.0	1
+8	0.0	1
+9	0.0	1
+10	0.0	1
+11	3.9	7.4e-06
+12	0.0	1

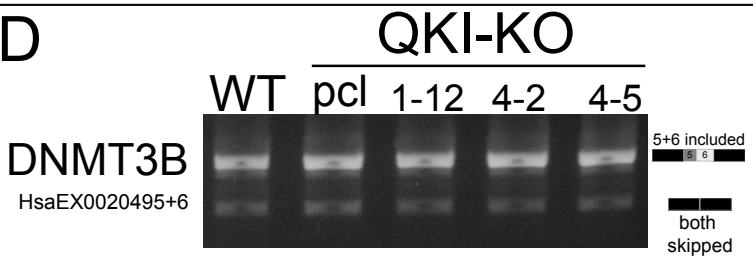
B



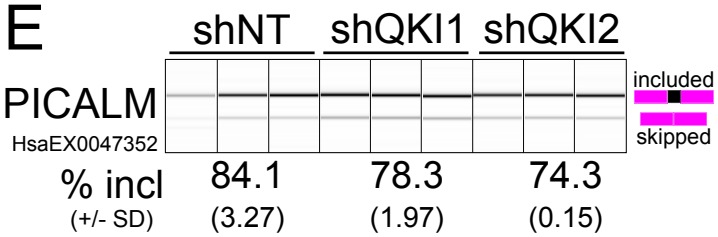
C



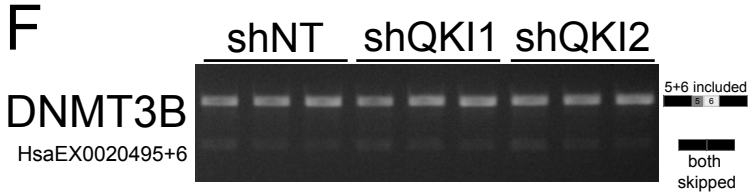
D



E



F

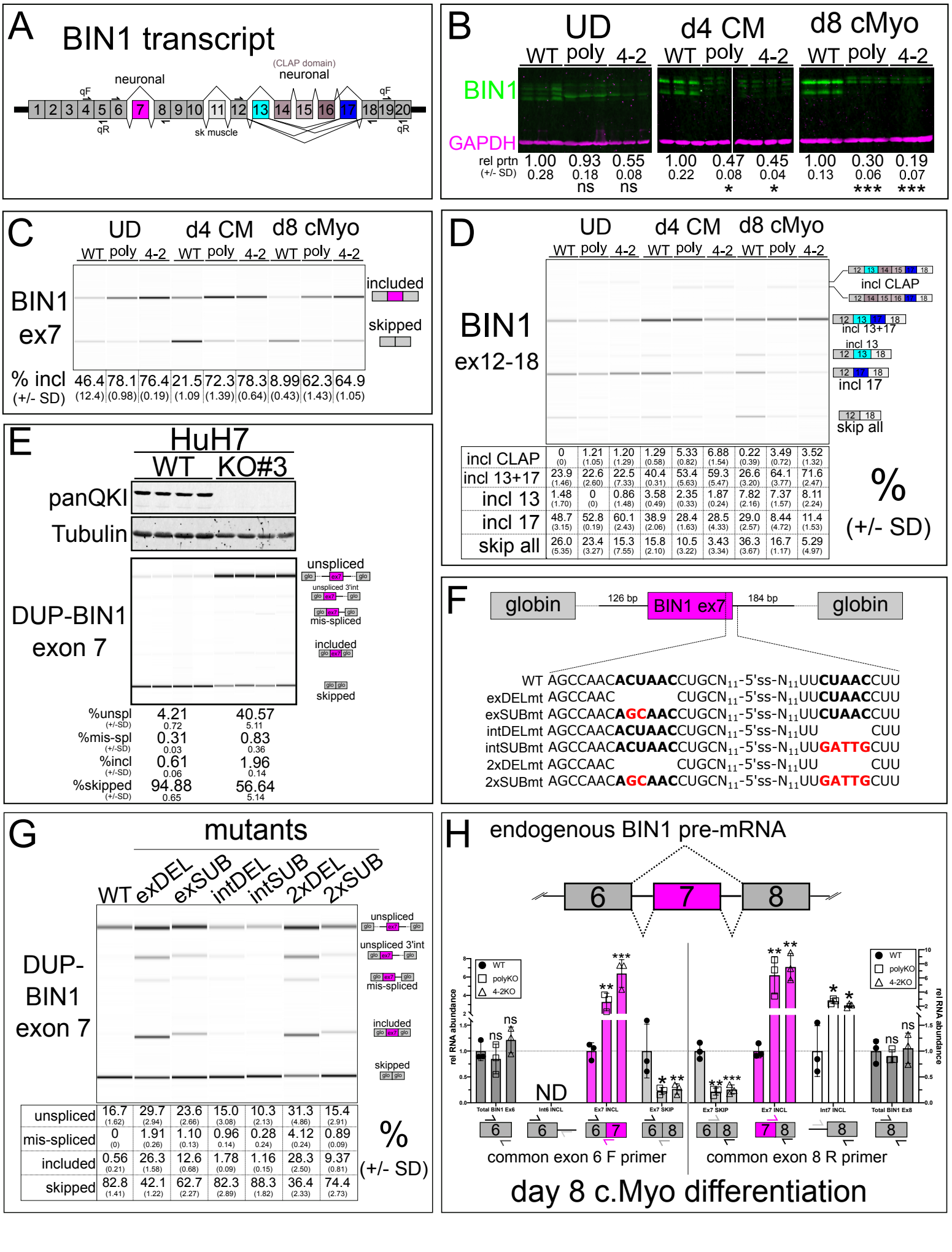


G

	UD		d2 lat meso		d4 CM		d6 c.myo		d8 c.myo	
	WT	KO	WT	KO	WT	KO	WT	KO	WT	KO
RAI14	82.5 (0.49)	100 (0)	62.8 (2.59)	100 (0)	29.5 (1.75)	99.6 (0.64)	10.1 (3.06)	100 (0)	10.5 (0.61)	100 (0)
NF1	67.0 (0.61)	40.0 (1.77)	83.2 (0.62)	56.4 (1.78)	88.0 (1.46)	57.8 (1.03)	93.2 (6.03)	46.3 (2.17)	87.3 (0.76)	54.0 (2.22)
BIN1	41.2 (5.13)	83.1 (1.78)	33.9 (1.41)	78.9 (3.37)	31.0 (1.95)	69.8 (3.25)	29.8 (15.67)	71.7 (2.58)	27.0 (1.40)	71.6 (0.49)

% incl (+/- SD)

FIGURE 4


A
BIN1 transcript
(CLAP domain)
neuronal
neuronal
qF
qF
qR
qR
sk muscle
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
B
UD
d4 CM
d8 cMyo
WT poly 4-2
WT poly 4-2
WT poly 4-2
BIN1
GAPDH
rel prtn (+/- SD)
1.00 0.93 0.55
0.28 0.18 0.08
ns ns
1.00 0.47 0.45
0.22 0.08 0.04
* *
1.00 0.30 0.19
0.13 0.06 0.07
*** ***
C
UD
d4 CM
d8 cMyo
WT poly 4-2
WT poly 4-2
WT poly 4-2
BIN1
ex7
included
skipped
% incl (+/- SD)
46.4 78.1 76.4 21.5 72.3 78.3 8.99 62.3 64.9
(12.4) (0.98) (0.19) (1.09 (1.39) (0.64) (0.43) (1.43) (1.05)
D
UD
d4 CM
d8 cMyo
WT poly 4-2
WT poly 4-2
WT poly 4-2
BIN1
ex12-18
incl CLAP
incl 13+17
incl 13
incl 17
skip all
incl CLAP 0 (0) 1.21 (1.05) 1.20 (1.29) 1.29 (0.58) 5.33 (0.82) 6.88 (1.54) 0.22 (0.39) 3.49 (0.72) 3.52 (1.32)
incl 13+17 23.9 (1.46) 22.6 (2.60) 22.5 (7.33) 40.4 (0.31) 53.4 (5.63) 59.3 (5.47) 26.6 (3.20) 64.1 (3.77) 71.6 (2.47)
incl 13 1.48 (1.70) 0 (0) 0.86 (1.48) 3.58 (0.49) 2.35 (0.33) 1.87 (0.24) 7.82 (2.16) 7.37 (1.57) 8.11 (2.24)
incl 17 48.7 (3.15) 52.8 (0.19) 60.1 (2.43) 38.9 (2.06) 28.4 (1.63) 28.5 (4.33) 29.0 (2.57) 8.44 (4.72) 11.4 (1.53)
skip all 26.0 (5.35) 23.4 (3.27) 15.3 (7.55) 15.8 (2.10) 10.5 (3.22) 3.43 (3.34) 36.3 (3.67) 16.7 (1.17) 5.29 (4.97)
% (+/- SD)
E
HuH7
WT KO#3
panQKI
Tubulin
DUP-BIN1
exon 7
unspliced
unspliced 3'int
mis-spliced
included
skipped
%unspl (+/-SD) 4.21 0.72 40.57 5.11
%mis-spl (+/-SD) 0.31 0.03 0.83 0.36
%incl (+/-SD) 0.61 0.06 1.96 0.14
%skipped (+/-SD) 94.88 0.65 56.64 5.14
F
globin
126 bp
BIN1 ex7
184 bp
globin
WT AGCCAACACUAACCUGCN11-5'ss-N11UUCUAACCUU
exDELmt AGCCAAC CUGCN11-5'ss-N11UUCUAACCUU
exSUBmt AGCCAACAGCAACCUGCN11-5'ss-N11UUCUAACCUU
intDELmt AGCCAACACUAACCUGCN11-5'ss-N11UU CUU
intSUBmt AGCCAACACUAACCUGCN11-5'ss-N11UUGATTGCUU
2xDELmt AGCCAAC CUGCN11-5'ss-N11UU CUU
2xSUBmt AGCCAACAGCAACCUGCN11-5'ss-N11UUGATTGCUU
G
mutants
WT exDEL exSUB intDEL intSUB 2xDEL 2xSUB
DUP-BIN1
exon 7
unspliced
unspliced 3'int
mis-spliced
included
skipped
unspliced 16.7 (1.62) 29.7 (2.94) 23.6 (2.66) 15.0 (3.08) 10.3 (2.13) 31.3 (4.86) 15.4 (2.91)
mis-spliced 0 (0) 1.91 (0.26) 1.10 (0.13) 0.96 0.14) 0.28 0.24) 4.12 (0.24) 0.89 (0.09)
included 0.56 (0.21) 26.3 (1.58) 12.6 (0.68) 1.78 (0.09) 1.16 (0.15) 28.3 (2.50) 9.37 (0.81)
skipped 82.8 (1.41) 42.1 (1.22) 62.7 (2.27) 82.3 (2.89) 88.3 (1.82) 36.4 (2.33) 74.4 (2.73)
% (+/- SD)
H
endogenous BIN1 pre-mRNA
6 7 8
WT
polyKO
4-2KO
rel RNA abundance
ns
ND
**
***
*
Total BIN1 Ex6
Int6 INCL
Ex7 INCL
Ex7 SKIP
Ex7 SKIP
Ex7 INCL
Int7 INCL
Total BIN1 Ex8
common exon 6 F primer
common exon 8 R primer
day 8 c.Myo differentiation

SUPPLEMENTAL FIGURE 4

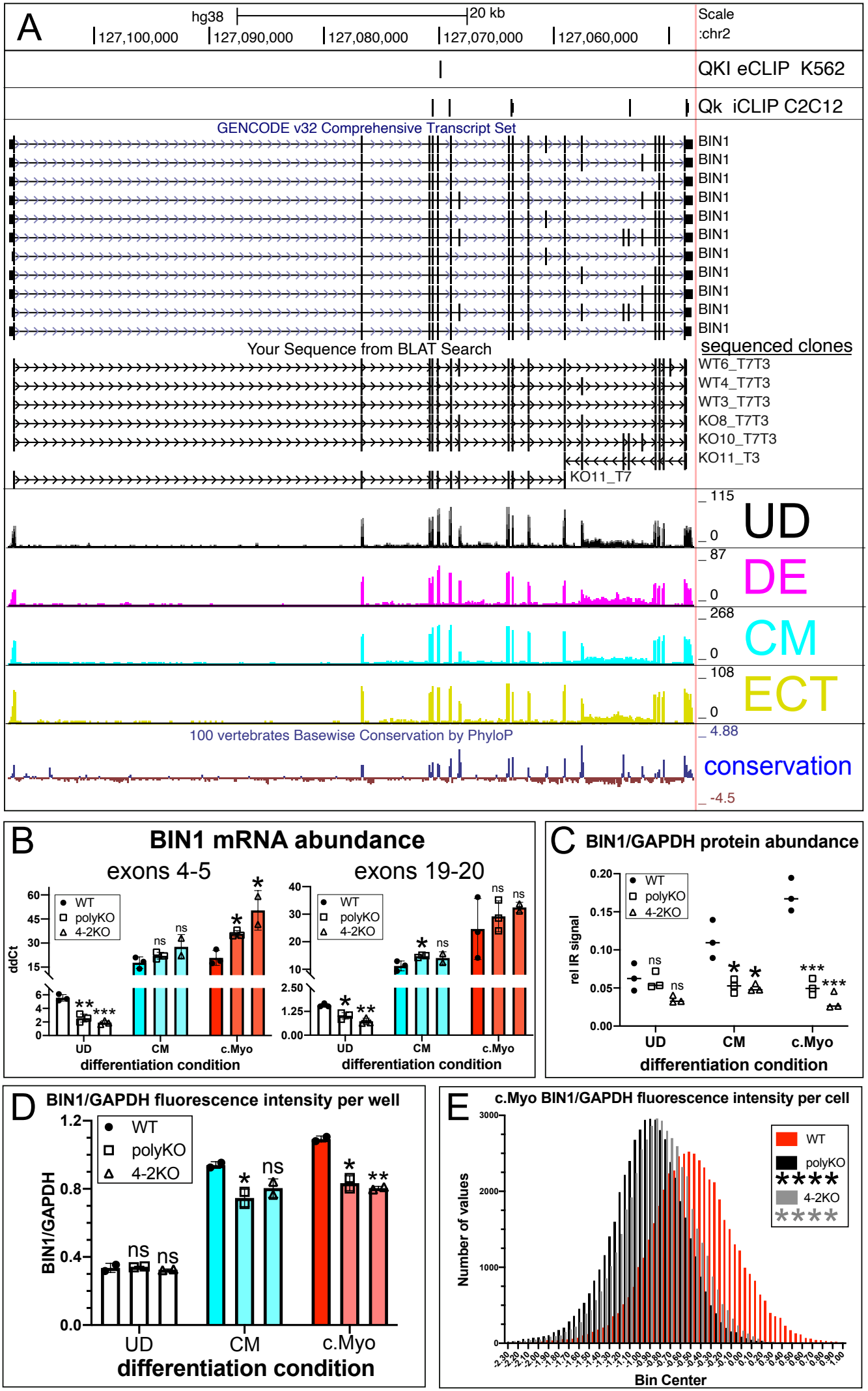
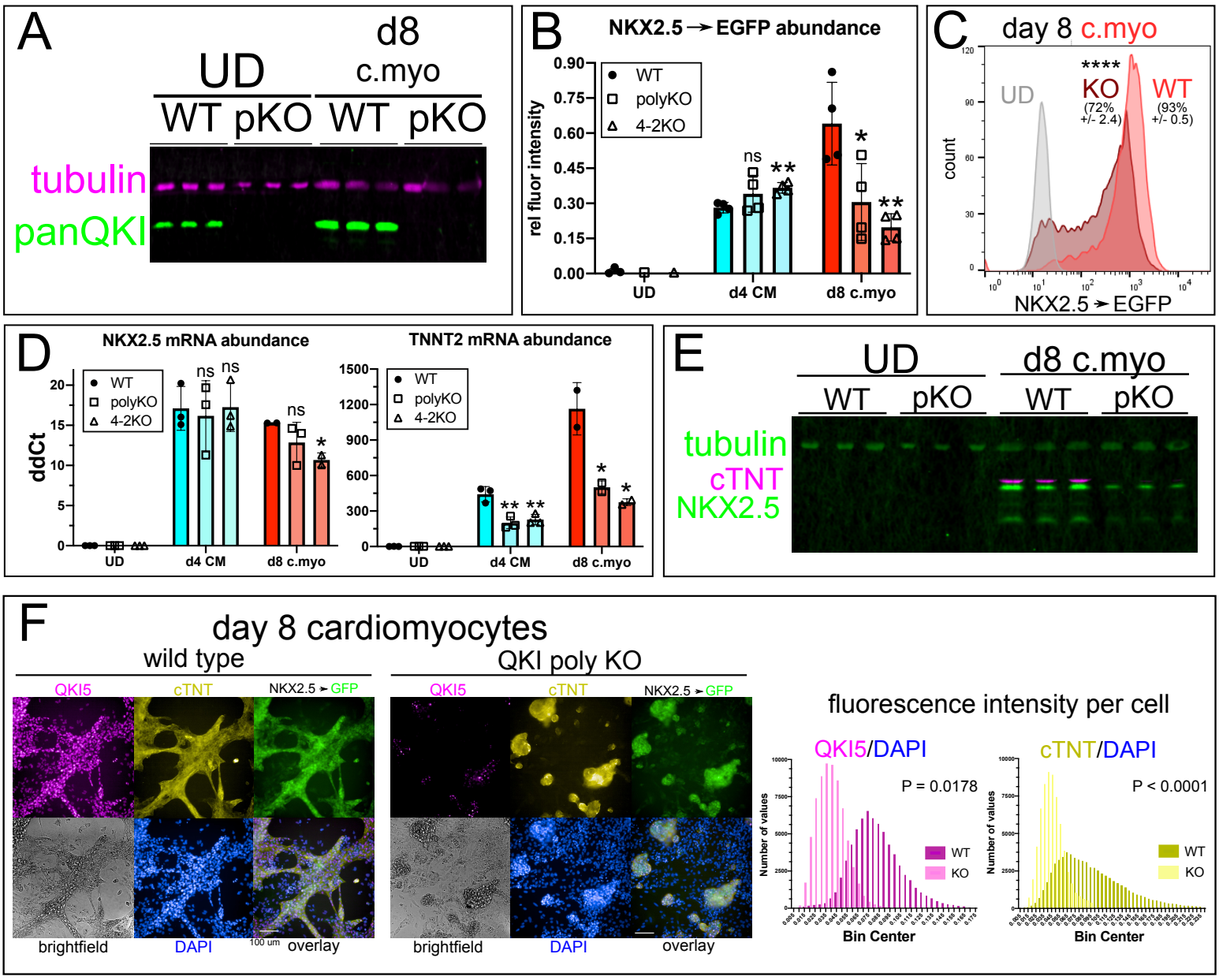
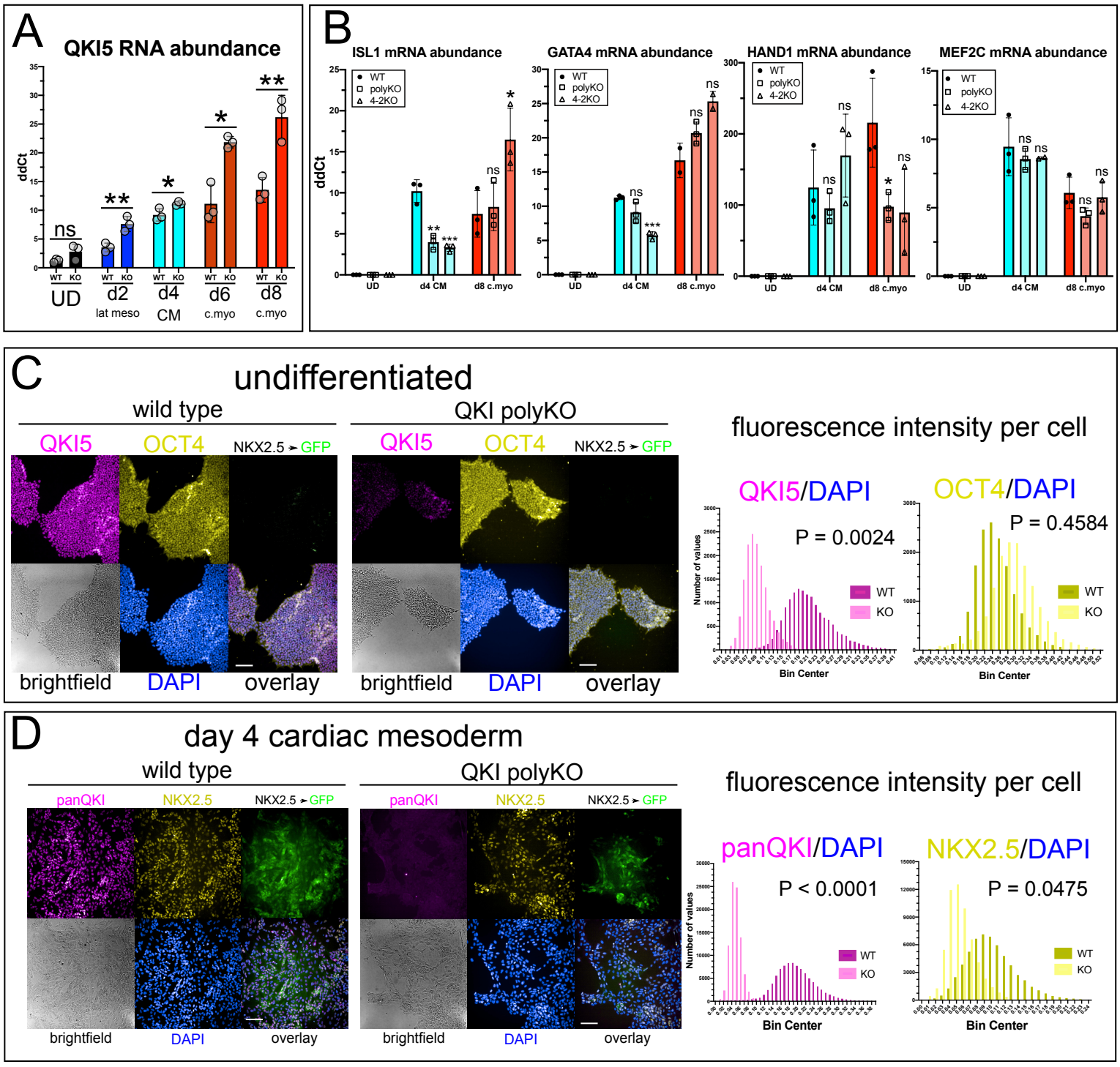


FIGURE 5


A
UD
d8 c.myo
WT pKO WT pKO
tubulin
panQKI
B
NKX2.5 → EGFP abundance
rel fluor intensity
WT
polyKO
4-2KO
ns
UD
d4 CM
d8 c.myo
C
day 8 c.myo
UD
KO
(72% +/- 2.4)
WT
(93% +/- 0.5)
count
NKX2.5 → EGFP
D
NKX2.5 mRNA abundance
TNNT2 mRNA abundance
ddCt
E
UD
d8 c.myo
WT pKO WT pKO
tubulin
cTNT
NKX2.5
F
day 8 cardiomyocytes
wild type
QKI poly KO
QKI5
cTNT
NKX2.5 → GFP
brightfield
DAPI
overlay
100 um
fluorescence intensity per cell
QKI5/DAPI
P = 0.0178
cTNT/DAPI
P < 0.0001
Number of values
Bin Center
WT
KO

SUPPLEMENTAL FIGURE 5