QKI is a critical pre-mRNA alternative splicing regulator of cardiac myofibrillogenesis and contractile function

Xinyun Chen1,2,3, Ying Liu3,8, Chen Xu1,3,8, Lina Ba3, Zhuo Liu3, Xiuya Li1, Jie Huang1, Ed Simpson4, Hongyu Gao4, Dayan Cao5, Wei Sheng2,3, Hanping Qi3, Hongrui Ji3, Maria Sanderson3, Chen-Leng Cai3, Xiaohui Li5, Lei Yang3, Jie Na6, Kenichi Yamamura7, Yunlong Liu4, Guoying Huang2✉, Weinian Shou3✉, Ning Sun1,2✉

The RNA-binding protein QKI belongs to the hnRNP K-homology domain protein family, a well-known regulator of pre-mRNA alternative splicing and is associated with several neurodevelopmental disorders. Qki is found highly expressed in developing and adult hearts. By employing the human embryonic stem cell (hESC) to cardiomyocyte differentiation system and generating QKI-deficient hESCs (hESCs-QKIdel) using CRISPR/Cas9 gene editing technology, we analyze the physiological role of QKI in cardiomyocyte differentiation, maturation, and contractile function. hESCs-QKIdel largely maintain normal pluripotency and normal differentiation potential for the generation of early cardiogenic progenitors, but they fail to transition into functional cardiomyocytes. In this work, by using a series of transcriptomic, cell and biochemical analyses, and the Qki-deficient mouse model, we demonstrate that QKI is indispensable to cardiac sarcomerogenesis and cardiac function through its regulation of alternative splicing in genes involved in Z-disc formation and contractile physiology, suggesting that QKI is associated with the pathogenesis of certain forms of cardiomyopathies.
The transcriptional and posttranscriptional modifications are critical to mammalian gene expression\(^1\). RNA splicing is a posttranscriptional process by which introns are removed from the newly transcribed sequences of immature pre-mRNAs and this process is required for generating mature protein-coding mRNAs\(^2\). In contrast to constitutive splicing, alternative splicing is a dynamic process that is highly regulated upon cellular differentiation or in response to distinct physiological states, resulting in specific exons being either included or excluded in unique combinations to generate multiple mRNA isoforms from a single gene\(^3\). Genome-wide studies estimated that up to 95% of human genes undergo some level of alternative splicing\(^4\), and these splicing events are associated with almost all normal cellular and physiological functions as well as pathological conditions.

Alternative splicing is regulated by various cis-regulatory sequences and trans-acting factors\(^5-8\). RNA-binding proteins (RBPs) can recognize and bind to specific RNA sequences to form ribonucleoprotein complexes and act as regulators of diverse biological functions, such as modifying RNA post-transcriptionally and transporting RNA\(^9\). Many RBPs play particularly important roles in pre-mRNA splicing\(^10,11\). In general, upon receiving upstream biological signals, RBPs coordinate alternative splicing events by recognizing specific cis-elements (e.g., enhancers or silencers) to either further promote or inhibit specific splicing events. The most well-known RBP families involved in RNA splicing are the serine/arginine-rich (SR) protein family, with members such as SF2/ASF, SRp20, SRp40, SRp55, and SRp75, and the heterogeneous nuclear ribonucleoprotein (hnRNP) protein family, with members such as PTB, hnRNP A1, hnRNP C, hnRNP D, hnRNP E, hnRNP F/H, hnRNP G, and hnRNP H\(^12-14\). The heart is the first functional organ to be formed during embryonic development. Interestingly, among ~1148 RBPs in the heart, there are ~390 cardiac-specific RBPs and many of them are present in the developing hearts\(^15\), suggesting that alternative splicing is one of the major regulatory events for cardiogenesis and cardiac physiology. Indeed, it has been shown that mutations of several RBPs and cis-regulatory sequences are associated with cardiomyopathies, muscular atrophy, hypercholesterolemia, and other cardiac disorders\(^10,16-22\).

Previously, several studies have suggested that alternative splicing involved in sarcomereogenesis impacts heart development and function\(^10\). These prior investigations strongly indicated a critical role for alternative splicing and the associated RBPs in cardiac development and cardiac physiology, as well as the pathogenesis of heart diseases. RBM24 and RBM20 are two of the most well-known RBPs that have been implicated in the pathogenesis of cardiomyopathies\(^23-25\). The loss of RBM20 function resulted in aberrant splicing events that led to abnormal sarcomereogenesis in embryonic development\(^24\). Genetic ablation of SRp38 in mice resulted in mis-splicing of triadin, a cardiac protein that regulates calcium release from the sarcoplasmic reticulum during excitation–contraction coupling (E–C)\(^26\). SF2/ASF cardiomyocyte-specific ablation results in the development of dilated cardiomyopathy and rapid progression of heart failure\(^27\). RBFox1 acts as a vital regulator for the conserved splicing process of transcription factor MEF2 family members and is involved in the heart failure progression\(^28\).

One particular member of the STAR (signal transduction and activation of RNA) gene family, known as Quaking (QKI), belongs to the hnRNP K-homology (KH)-domain family of proteins, and it is a sequence-specific RBP that is enriched in the heart and central nervous system\(^29-32\). QKI contains an RNA-binding motif (KH domain), which is flanked by QUA1 and QUA2 domains. The QUA domain is involved in forming homo- or heterodimers and is required for RNA binding\(^33-36\). In addition to these functional domains, there is a tyrosin cluster located within the proline-rich PXXP motif that can be phosphorylated by Src kinases\(^37\), suggesting that QKI can be regulated by intracellular signaling. At least three major alternatively spliced mRNA isoforms are generated from the QKI gene, QKI–5, QKI-6, and QKI-7, in which exons 1–6 encode identical structures in these isoforms but differ in their C-terminal end encoded by exons 7 and 8\(^38\). QKI-5 has a nuclear localization signal (NLS)\(^39,40\) and has been shown to play a major function in pre-mRNA splicing regulators\(^41-46\). QKI-6 and QKI-7 lack NLS and apparently have different biological functions\(^38,40,47,48\). They seem to play more important roles in regulating mRNA stability and other post-transcriptional mRNA processing and intracellular transportation\(^49-52\). These functional differences reflect the complexity of QKI biological functions. Based on the early embryonic lethal phenotype of Qki-deficient mice, QKI is considered essential for early embryonic development\(^53\). Interestingly, a spontaneous mutant mouse line, known as the Qk\(^{alb}\) mouse line, in which a 1 Mb promoter/enhancer deletion upstream of the Qki transcription start site resulted in deficient myelination in the central nervous system, has been carefully studied\(^53-55\). Over the past decades, the use of these Qki mutant animal models has suggested important functions in neural progenitors, myelin formation, smooth muscle differentiation, and monocyte to macrophage differentiation\(^45,55,56\). Despite the evidence of Qki expression in developing hearts, the role of QKI in regulating normal cardiogenesis and cardiac physiology has not been carefully studied.

In this work, we analyze the biological function of QKI-mediated post-transcriptional regulation in cardiogenesis and cardiac physiology. By taking advantage of a defined in vitro system for differentiating human embryonic stem cells (hESCs) into cardiomyocytes and the Qk\(^{alb}\) mouse model, we are able to reveal that QKI is a critical pre-mRNA alternative splicing regulator in cardiomyocytes and is essential for myofibrillogenesis and contractile physiology during cardiomyocyte differentiation and maturation. Our finding shows that QKI is indispensable to normal cardiogenesis and cardiac function.

**Results**

The expression of QKI in cardiomyocytes and the generation of hESC-QKI\(^{del}\). QKI expression was previously found in the hearts\(^29\). To determine the temporal expression pattern of QKI during hESC-to-cardiomyocyte differentiation (Supplementary Fig. 1A), we analyzed QKI-5, -6, and -7 mRNA expression levels in undifferentiated hESCs and differentiating cells at Day-1, -3, -6, -8, and -10 by using quantitative reverse transcription PCR (qRT-PCR) (Fig. 1A). In this in vitro differentiation system, cells at Day-1 were considered to be early nascent mesodermal cells with high levels of expression of T (Brachyury); cells at Day-3 were considered to be cardiogenic mesodermal cells with a specific upregulation of MESP1; cells at Day-6 were considered to be cardiac progenitor cells with upregulation of ISL1; and cells at Day-8 to -10 were considered to be well-committed differentiating cardiomyocytes positive for a series of cardiomyocyte markers, such as NKX2-5 and TNNT2. Using these markers as internal references (Fig. 1B), we were able to demonstrate that QKI-5 transcripts were present in undifferentiated hESCs and there was a quick downregulation at the early differentiation stage. It maintained at a steady lower expression level until differentiation Day-6 to Day-10 (Fig. 1A), at which point cells reached the transition stage from cardiogenic progenitors to early differentiated cardiomyocytes. Comparing to QKI-5 expression, QKI-6 and QKI-7 expression levels were at a much lower level throughout, while their expression was elevated in the later stage of differentiation (Fig. 1A). Western blot analysis further
supported this observation that QKI-5 is the main isoform in hESCs and during hESC-to-cardiomyocyte differentiation (Fig. 1C and Supplementary Fig. 1B–E). This finding indicated that QKI was likely to have a specific function in the early cardiogenic process of transitioning cardiac progenitors to early cardiomyocytes. To confirm this observation in vivo, we analyzed the Qki spatiotemporal expression pattern in developing mouse embryos from E7.5 to E9.5 using the previously generated QkiβGeo mouse model (Fig. 1D, E); in this mouse line, the βGeo-reporter expression cassette is under the control of Qki promoter, thus representing Qki expression. Based on the X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) staining, we were able to visualize reporter expression in the developing hearts from the E7.5 cardiac crescent to the four-chamber heart and the expression was persistent in both the atrial and ventricular myocardia (Fig. 1D, E), confirming that QKI was likely critical for early cardiogenic events. Cardiac Qki expression persisted in the postnatal hearts (Fig. 1E), further indicating a potentially important role for QKI in maintaining normal adult cardiac function.

QKI-deficient mutant hESC lines (hESC-QKI<sup>del</sup>) were generated by using CRISPR/Cas9 gene-editing technology. The guide
Altered contractile function of cardiomyocyte sheets derived from hESCs-\(QKI^{\text{del}}\). As \(QKI\) is expressed in undifferentiated hESCs, we first analyzed the pluripotency and self-renewal capacity of hESCs-\(QKI^{\text{del}}\) by examining the expression of alkaline phosphatase and the following pluripotent markers: Oct4, Sox2, Nanog, c-Myc, Klf4, SSEA-4, and Tra-1-60. As shown in Supplementary Fig. 3A–C, hESCs-\(QKI^{\text{del}}\) had normal expression of these pluripotent markers, indicating that \(QKI\) was not required for maintaining the pluripotency of hESCs. To further confirm this conclusion, we injected \(QKI^{\text{del}}\)-hESCs into nonobese diabetic/severe combined immunodeficiency mice to test whether they could form teratomas (containing all three germ layers). Our data confirmed the formation of well-formed teratomas containing the three embryonic germ layers derived from
hESCs-QKI<sub>del</sub> (Supplementary Fig. 3D, E). In conclusion, hESCs-QKI<sub>del</sub> were normal in the maintenance of pluripotency and the differentiation potentials to the three germ layers.

To test whether QKI was specifically involved in cardiomyocyte differentiation and function, we analyzed the cardiomyocyte differentiation efficiency of hESCs-QKI<sub>del</sub> (e.g., H1-7 and H1-8) in comparison to congenic control hESCs (e.g., H1 cells). Consistent with the different hESC-QKI<sub>del</sub> clones tested, H1-7 and H1-8 were able to differentiate into monolayer cardiomyocyte sheets but with a dramatically reduced contractility and increased E<sub>0</sub> compared to normal control hESCs from early cardioprogenitors to differentiated cardiomyocytes using cluster 12 as the root cells. Under our hESC-to-cardiomyocyte differentiation protocol, early cardioprogenitors branched out two major populations of cardiomyocytes, higher proliferative cardiomyocytes, which had enriched expression of genes marking cells in the active proliferative state (Supplementary Fig. 5A), and differentiated cardiomyocytes with lower proliferative activity. Apparently, the differentiated cardiomyocytes were derived from two routes; the first route was from cardioprogenitors (cluster 12) to higher proliferating cardiomyocytes (clusters 0 and 4) (Fig. 3E), which occurred mainly in early differentiation stage (Day-6, Fig. 3F); the second route was from cardioprogenitors (cluster 12) to early differentiating cardiomyocytes (cluster 2) and differentiated cardiomyocytes (clusters 0 and 4) (Fig. 3E, G). Importantly, all cells towards cardiomyocytes differentiated from hESCs-QKI<sub>del</sub> and normal control hESCs were largely superposable in Monocle trajectory analysis, confirming overall differentiation trajectory were consistent between WT control hESCs and hESCs-QKI<sub>del</sub> under our differentiation protocol. The number of cardioprogenitors and early differentiating cardiomyocytes were similar in H1-hESC differentiation group and hESCs-QKI<sub>del</sub> group (10.9% in H1 vs. 9.71% in H1-8), suggesting that the cardiomyocyte lineage specification, determination, and early differentiation were largely preserved in hESCs-QKI<sub>del</sub>. However, it was important to note that the differentiated cardiomyocytes (clusters 0 and 4) accounted for 53.1% of total cells in the normal control hESC differentiation group and 43.3% in hESCs-QKI<sub>del</sub> group at Day-15, a roughly 18.5% reduction. This finding was similar to the significant reduction of TNNT2-positive cardiomyocytes in hESCs-QKI<sub>del</sub> measured by fluorescence-activated cell sorting analysis (Supplementary Fig. 4). Interestingly, differentiating cardiomyocytes (cluster 2) at Day-15 accounted for 2.14% of total cells in the normal control group; however, it accounted for 4.97% of total cells in hESCs-QKI<sub>del</sub> group, a dramatic 132.0% increase. This correlation further suggested that the major defect in hESCs-QKI<sub>del</sub> was at the transition stage from early differentiating cardiomyocytes to functionally more mature cardiomyocytes.

Based on the signature gene expression profiles, in all cardiogenic and early cardiomyocyte clusters at D-6, we did not find a major alteration in gene expression between the hCMs-QKI<sub>del</sub> and the control cells (Supplementary Fig. 5B). To confirm this observation, we used qRT-PCR to analyze the temporal expression patterns of several key mesodermal and cardiogenic lineage markers from Day-0 to Day-10: MESP1, Brachyury (T), ISL1, NKX2-5, GATA4, MYOCD, TNNT2, and ACTN2. As shown in Supplementary Fig. 6A, except for the downregulation of ACTN2 at Day-10, the overall expression pattern of these major lineage markers was largely unchanged during the early differentiation process. These data confirmed that the majority of hESCs-QKI<sub>del</sub> had normal early cardiomyocyte specification and determination. In the late cardiomyocyte clusters of Day-15 (clusters 0 and 4), we found that the altered expression of genes was most significant to contractile function (Fig. 3H and Supplementary Fig. 6B). Among them, ACTN2 (α-actinin),
TNNI3 (troponin I3), MYL2 (myosin light chain 2), MYH7 (β-myosin heavy chain), ACTA1 (actin α-cardiac muscle 1), ANKRD1 (ankyrin repeat domain 1), and TAGLN (transgelin) were significantly downregulated in the hCMs-QKIdel, suggesting that the diminished contractile function in cardiomyocyte sheets derived from hESCs-QKIdel was largely due to the altered regulation of sarcomere genes. Interestingly, the most significantly altered genes in the early differentiating cardiomyocytes (cluster 2) included genes involved in post-translational protein phosphorylation and cell cycle regulation (Fig. 3H and Supplementary Fig. 6B), suggesting a defect in the intracellular signaling. The downregulation of ACTN2 was confirmed in Day-15 hCMs-QKIdel (Supplementary Fig. 6C). In addition, the ratio of MYH7/MYH6 in hCMs-QKIdel at Day-15 was significantly lower compared to control hCM (Supplementary Fig. 6C), indicating a defect in the process of functional maturation of cardiomyocyte.
These data further confirmed that the role of QKI in cardiomyocyte differentiation was at the transition stage bridging early cardiomyocytes to functional cardiomyocytes, which was consistent with the overall Qki spatiotemporal expression pattern in mouse early embryos as well as in the hESC-to-cardiomyocyte differentiation process (Fig. 1).

**Bulk transcriptomic analysis of hCMs-QKIdel**. To further confirm this observation, a bulk RNA-seq analysis was performed, which provided a more in-depth transcriptome comparison between Day-15 differentiated cells of hESCs-QKIdel (H1-8) and control hESCs (H1). As demonstrated by the volcano plot in Fig. 4A, based on the false discovery rate (FDR) value of <0.05 and the fold change of 1.5, 556 genes were upregulated, and 661 genes were downregulated. Based on Gene Ontology (GO) analysis (Fig. 4B), muscle contraction, muscle tissue development, and animal organ morphogenesis were among the most affected biological processes (BP). Interestingly, despite the RNA-seq being performed using the differentiated cardiomyocytes, other top affect BP included the regulation of trans-synaptic signaling and the regulation of blood vessel diameter; the former was likely consistent with the notion that QKI was highly expressed in the developing and adult nervous system (Fig. 1)\(^6\), while the latter was likely consistent with the vascular developmental defect in Qki-deficient embryos\(^3\). The GO analysis also indicated the most significantly affected molecular functions (MFs) and cellular components (CCs), which were consistent with the finding in altered BP. Furthermore, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis indicated that pathways involved in cardiomyopathies, regulation of actin cytoskeleton, cardiac muscle contraction and adhesion, mitogen-activated protein kinase signaling, Rap1 signaling, and adrenergic signaling were the most significantly altered biological pathways (Fig. 4C). Figure 4D represents the heatmap of differentially expressed genes within these affected BPs and pathways. This finding was confirmed by additional investigation using Ingenuity Pathway Analysis (Fig. 4E, F). Collectively, our data strongly suggested that QKI was most likely a key regulator of cardiomyocyte myo-fibrillogenesis and contractile function.

**QKI-mediated alternative splicing in cardiac myogenesis**. The major MF of QKI is its regulation of alternative splicing. The data from transcriptome profiling as discussed above mostly reflected the altered cellular function and the pathogenetic pathways associated with altered function in hCMs-QKIdel. Based on the bulk RNA-seq data, we analyzed differential alternative splicing events in hCMs-QKIdel using rMATs software\(^8\). Based on the FDR value (<0.05) and the ΔPSI value (±0.3), we found that all 5 alternative splicing types were present in hCMs-QKIdel, including skipped exon (SE), intron retention, mutually exclusive exons (MXEs), alternative 5′-donor sites, and alternative 3′-acceptor sites (Fig. 5A). Among them, SE and MXE were two major alternative splicing types in hCM-QKIdel, which accounted for 93% of total AS events (among 461 total alternative splicing events, there were 363 SE events and 65 MXE events) from a total of 299 genes (Fig. 5A). Based on the RBP motif analysis, the majority of these alternatively spliced events contained the QKI RNA-binding consensus sites (Supplementary Fig. 7)\(^3\), confirming that the majority of differential alternative splicing events in hCM-QKIdel were relevant to QKI-mediated processes. Interestingly, GO analyses of these altered alternative splicing events suggested that the most significantly affected BPs in hCM-QKIdel were the actomyosin structure organization and the regulation of actin filament organization (Fig. 5B), in genes involved in cardiac myofibrillogenesis and cardiomyocyte contractile function, such as ACTN2, ABLIM1, TTN, NEBL, RYR2, and CACNA1G, were affected, as confirmed by RT-PCR (Fig. 6A), suggesting that the primary defect in hCM-QKIdel was the dysregulated formation of contractile apparatus and its associated physiological function, which was further confirmed by the most significantly affected events in MFs and CCs, as well as the molecular pathways demonstrated by KEGG pathway analysis, not only indicating that myofibrillogenesis was the major affected process but also its potential link to cardiomyopathies (Fig. 5B, C). Figure 6B shows the representative schematic diagrams of altered SE events in ACTN2, ABLIM1, NEBL, RYR2, CACNA1C, and PDLIM5, which were derived from rMATs analysis and sashimi plots shown in Supplementary Fig. 8. As QKI is known for its sequence-specific RNA binding, we used RNA immunoprecipitation (RIP)-quantitative PCR (qPCR) to confirm that several major SE events were QKI-mediated. As shown in Fig. 6C, we were able to validate that QKI was able to bind to ACTN2, CAMK2D, PDLIM5, NEBL, ABLIM1, RYR2, and CACNA1C. Collectively, these data confirmed the role of QKI in regulating alternative splicing in genes critically involved in sarcomere formation and contractile function.

**QKI-mediated alternative splicing in the ACTN2 gene and altered myofibrillogenesis in hCMs-QKIdel**. When we compared the differentially expressed genes with genes that underwent differential alternative splicing in hCM-QKIdel, we found that eight genes were associated with both groups (Supplementary Fig. 9A). Among them, ACTN2 was at the top in both the downregulated gene list (reduced over 29-fold) and the differential alternative splicing list. In hCMs-QKIdel, ACTN2-SE-8 was generated, which had an abnormally skipped exon 8a or exon 8b (FDR < 0.0000001). QKI-binding consensus sequences were mapped in ACTN2 exon 7, and in introns flanking exons 8a and 8b in ACTN2 isoform 1 and isoform 2 (Fig. 6D). Interestingly, the
analysis of the mRNA sequence of ACTN2<sup>SE-8</sup> revealed that a premature stop codon appeared in a position located in exon 9 (Fig. 6D and Supplementary Fig. 9B). To confirm whether this premature stop codon caused an activation of nonsense-mediated decay (NMD), which subsequently led to the dramatic down-regulation of ACTN2 expression, we applied NMDI14 (Millipore sigma; SML1538) to hCM-QKIdel<sup>del</sup>, a specific inhibitor disrupting the interaction of UPF1 and SMG7 in the NMD pathway<sup>64</sup>. We found that ACTN2<sup>SE-8</sup> mRNA levels were significantly enriched in NMDI14-treated cells (Fig. 6D, right panel), indicating that this rapid degradation of ACTN2<sup>SE-8</sup> mRNA in hCM-QKIdel<sup>del</sup> was in fact due to the activation of the NMD pathway. Collectively,
Fig. 4 Bulk transcriptome analysis of hESCs and hESC-QKIdel differentiated cardiomyocytes at Day-15. A A representative volcano plot demonstrates differentially expressed genes between H1 differentiated cardiomyocytes and mutant H1-8 differentiated cardiomyocytes at Day-15. Based on the FDR value (<0.05) and fold change (FC), over 1556 genes were upregulated, and 661 genes were downregulated in H1-8 differentiated cardiomyocytes. B, C Representative enrichment analysis of differentially expressed genes using Gene Ontology (GO) (B) and KEGG pathways (C) databases. The top 6 ranked GO terms altered in the categories of Biological Process, Cellular Component, and Molecular Function, and the top ten ranked KEGG terms are represented in the bar graph. GO enrichment and KEGG pathway enrichment analysis of differentially expressed genes were performed respectively using R programming based on the hypergeometric distribution. D Representative heatmap shows differentially expressed genes relevant to cardiac muscle contraction and the calcium signaling pathway. E, F Representative IPA core analysis of key terms associated with cardiovascular diseases and the interaction network. The upregulated and downregulated genes are marked in red and green, respectively.

Fig. 5 Analysis of differential alternative splicing events in hESC and hESC-QKIdel differentiated cardiomyocytes at Day-15. A Schematic diagram shows the overall distribution of alternative splicing events in hESC-QKIdel differentiated cardiomyocytes. B Enrichment analysis of differentially spliced genes using Gene Ontology (GO) and KEGG pathway databases. GO enrichment and KEGG pathway enrichment analysis of differentially expressed genes were performed, respectively, using R programming based on the hypergeometric distribution.

these findings not only confirmed that ACTN2 was a primary target of QKI but also explained the cause of the dramatic downregulation of ACTN2 mRNA expression level in hCM-QKIdel.

To further investigate myofibrillogenesis and the cause of diminished contractile function in hCMs-QKIdel, we analyzed myofibril structures in hCMs-QKIdel and compared them to those in normal control hCMs by performing dual-immunofluorescence staining for TNNT2 and ACTN2. As shown in Fig. 7A, well-organized striated myofibril structures were found in control hCMs, which was in great contrast to the largely disorganized myofibrils lacking striated structure in hCMs-QKIdel. The immunoreactivity to anti-ACTN2 was severely reduced in hCMs-QKIdel, consistent with the downregulation of ACTN2 mRNA level. The disruption of the myofibril structure in hCMs-QKIdel was further confirmed by other sets of dual-immunofluorescence staining experiments, using antibodies against NEBL/TNNT2 (Fig. 7B) and TTN/TNNT2 (Supplementary Fig. 10). To further confirm this result, transmission electron microscopy (TEM) analysis was performed. As expected, thinned and disorganized sarcomeres lacking Z-disc structures were found in hCMs-QKIdel, which was contrasted with the well-formed sarcomeres seen in control hCMs (Fig. 7C). Furthermore, in addition to ACTN2, Western blot analysis demonstrated significantly reduced protein levels of multiple sarcomeric proteins, including the myosin heavy chains (i.e., positive for MF20 antibody), TNNT2, MYOZ2, and TNNI3 (Fig. 7D), further confirming a defect in myofibrillogenesis in the hCMs-QKIdel.

QKI-5-mediated RNA-splicing activity is key to regulate cardiac myofibrillogenesis. QKIs have three major splicing isoforms, known as QKI-5 (QKI-203), QKI-6 (QKI-207), and QKI-7 (QKI-201), which differ in the last two coding exons 7 and 8. All QKI isoforms were present in heart tissue, whereas QKI5 was the highest-expressed isoform. QKI-5 contains an NLS and was exclusively localized in cardiomyocyte nuclei, which was
consistent with its implicated role in RNA splicing. QKI-6 and QKI-7, which lack an NLS, were localized in cardiomyocyte cytoplasm and implicated in maintaining RNA stability and posttranscriptional modification. To further confirm whether hCM-QKI<sup>del</sup> phenotype was due to QKI-5-mediated RNA-splicing activity or QKI-6/7-mediated regulation in RNA stability, or posttranscriptional modification, we used a PiggyBac (PB) transposon/doxycycline-inducible system to reactivate the expression of QKI-5, QKI-6, or QKI-7 in hESCs-QKI<sup>del</sup>, respectively (Supplementary Fig. 11A). A total of three hESCs-QKI<sup>5ind</sup> clones, four hESCs-QKI<sup>del;QKI-6ind</sup> clones, and three hESCs-QKI<sup>del;QKI-7ind</sup> clones were successfully established with confirmed doxycycline-induction-mediated QKI expression. Representative western blot screening is presented in Supplementary Fig. 11B.

We first tested hESCs-QKI<sup>del;QKI-5ind</sup> clones. Doxycycline induction was carried out at differentiation Day-0, -3, -6, and -8, and within 24 h, QKI-5 was successfully induced, as confirmed by western blottings. These hESCs-QKI<sup>del;QKI-5ind</sup> were differentiated into cardiomyocytes (hCMs-QKI<sup>del;QKI-5ind</sup>) and analyzed at D15 Fig. 6 QKI-mediated alternative splicing events in genes critical to myofibrillogenesis. A Representative RT-PCR confirmation of abnormal splicing events. All experiments are independently repeated at least three times with multiple independent sets of cell samples to confirm the reproducibility of the findings. B Schematic diagram of Skipped Exon events based on rMATs analysis. Solid lines indicated as normal splicing events, dash lines indicated as abnormal events in hCM-QKI<sup>del</sup>. Relevant sashimi plots are in Supplementary Fig. 8. C RIP-qPCR analysis to verify the direct targets of QKI. Data represent the mean ± SEM from at least five independent experiments, statistical significance was determined by a two-tailed Student’s t-test. D Schematic diagram to demonstrate a premature STOP codon (i.e., TAA) in ACTN2<sup>SE-8</sup> mRNA in H1-8 mutant cells.

To confirm that the NMD pathway is responsible for the downregulation of ACTN2<sup>SE-8</sup> in H1-8 mutant cells, H1-8 cardiomyocytes were treated overnight with NMDI14 (25 μM), followed by qRT-PCR analysis on the level of ACTN2<sup>SE-8</sup> in H1-8 cardiomyocytes. Data are shown as the mean ± SEM, statistical significance was determined by a two-tailed Student’s t-test.
and D30. When QKI-5 was induced at days 0, 3, and 6 (but not at day-8), hCMs-\(\text{QKIdel}\text{-QKI5ind}\) were able to be fully rescued to form well-organized beating cardiomyocyte sheets, similar to normal control cardiomyocyte sheets differentiated from hESCs (Supplementary Movies 7 and 8). Dual-immunofluorescence staining demonstrated a full recovery of striated myofibril structure (Fig. 8) and the recovery of normal splicing events for all key genes examined in Day-15 differentiated cells (Supplementary Fig. 11C, D). This finding also suggested that QKI-5 was critical for the key transition stage when cardiac progenitors became early cardiomyocytes.
cardiomyocytes, which was consistent with the temporal expression pattern of QKI in the hESC-CM differentiation system and in mouse cardiovascular development. However, similar reactivation of QKI-6 and QKI-7 expression in hESCs-QKI\textsuperscript{del} failed to rescue the abnormal contractile phenotypes at both the functional and cellular levels (Fig. 8 and Supplementary Movies 7–10). Myofibrillar structures are recovered in doxycycline-induced H1-8-QKI\textsuperscript{ind} cardiomyocytes (Day-15) and ACTN2 expression is also recovered. However, doxycycline-induced H1-8-QKI\textsuperscript{ind} and H1-8-QKI\textsuperscript{del} cardiomyocytes fail to recover ACTN2 expression and generate normal myofibrillar structures (Day-15). Scale bar: 25 \( \mu \)m. All experiments are independently repeated three times with three different sets of cell samples to confirm the reproducibility of the findings.

**Defect in myofibrillogenesis in Qki-deficient mice.** To confirm the above observations in vivo, we analyzed Qki\textsuperscript{Geo}/\textsuperscript{Geo} homozygous mutant embryos, which died in utero at E10.5 (Fig. 9A).\textsuperscript{33} Histological analysis of E9.0–9.5 embryonic hearts demonstrated a thin ventricular wall with a significantly reduced level of trabeculation in Qki\textsuperscript{Geo}/\textsuperscript{Geo} homozygous mutants relative to the level seen in the hearts of wild-type littermate or Qki\textsuperscript{Geo}/\textsuperscript{+} heterozygous embryos (Fig. 9B). Qki\textsuperscript{Geo}/\textsuperscript{Geo} mutant hearts had similar downregulation of sarcomere genes (Fig. 9C and Supplementary Fig. 12A), which was consistent with the findings demonstrated by hCM-QKI\textsuperscript{del} scRNA-seq and bulk RNA-seq data. Dual-immunofluorescence staining using antibodies against cardiac α-actinin (Actn2) and troponin I (Tnni3) demonstrated dramatically reduced expression levels of Actn2 and a significant defect in myofibrillogenesis in Qki\textsuperscript{Geo}/\textsuperscript{Geo} hearts (Fig. 9D), which closely resembled the defect seen in hCMs-QKI\textsuperscript{del}. As we tested so far, all altered alternative splicing events in Qki\textsuperscript{Geo}/\textsuperscript{Geo} mutants were identical to that found in...
hCMs-QKIdel (Fig. 9E), including the abnormal Actn2 mRNA splicing variant lacking exon 8, containing a premature stop codon in exon 9, which was confirmed by sequencing analysis (Supplementary Fig. 12B). Collectively, these findings validated the critical role of Qki in cardiac myofibrillogenesis and contractile physiology in vivo.

**Discussion**

Pre-mRNA alternative splicing is one of the most important biological features in eukaryotes. This process provides an additional mechanism for posttranscriptional regulation of gene functions at the pre-mRNA level. As we show here, QKI is an important RBP that is highly expressed in developing hearts and adult hearts. By using both the hESC-cardiomyocyte differentiation system and mouse genetic model, we are able to demonstrate four major findings in our current study: (1) despite its expression in pluripotent embryonic stem cells, QKI does not appear to have important functions in the maintenance of pluripotency, the ability to self-renew and early differentiation towards early cardiogenic progenitors; (2) QKI is essential for differentiating cardiomyocytes during the transition from cardiogenic progenitors to early cardiomyocytes when important myofibrillogenesis and contractile protein expression is compromised.
contractile physiology are initiated in the transition to a fully functional state; (3) QKI safeguards proper alternative splicing events critical to genes contributing to myofibrillogenesis in developing cardiomyocytes; and (4) QKI-5 is the major QKI isoform in mediating pre-mRNA alternative splicing in cardiomyocytes.

QKI has drawn great attention for its potentially important disease-causing contribution to various human diseases, including schizophrenia, 6q terminal deletion syndrome, myelin disorders, and cancer.\(^{56-60}\) Interestingly, a recent two-stage genome-wide association study suggested that QKI was associated with myocardial infarction and coronary heart disease.\(^ {61,62}\) In addition, it was shown that QKI-5 and QKI-6 expression was associated with hypoxia/reoxygenation and ischemia/reperfusion (I/R) injuries in cardiomyocytes and heart.\(^ {63,64}\) The overexpression of QKI-5 or QKI-6 in neonatal cardiomyocyte could prevent I/R-induced cardiomyocyte death.\(^ {59}\) Although the underlying mechanism of this association was largely elusive, the potential role of QKI in regulating vascular smooth muscle dedifferentiation in response to vessel injuries was demonstrated in a mouse model, in which QKI functioned as a splicing regulator of a key smooth muscle transcription factor, myocardin.\(^ {35}\)

As myocardin was also a key transcription factor in early cardiomyocyte differentiation,\(^ {74,75}\) we have particularly focused on potentially altered myocardin splicing in our differentiating hESCs-QKIdel. However, we did not find significant alterations in myocardin alternative splicing during cardiomyocyte differentiation, suggesting a context-specific function of QKI. This notion also identifies a critical question as to what the role of QKI plays in the differentiation of various cellular lineages. Although our current study is not able to address this question, the hESC-QKIdel mutant lines will serve as an important resource for addressing these questions in the future.

One of the major findings in our study is that QKI plays a critical role in regulating alternative splicing for genes involved in Z-disc formation, which is a critical step in sarcomereogenesis. The Z-disc, also called Z-line, provides an anchor for the sarcomere structure. ACTN2, NEBL, ABLM1, and PDLIM5 are critical components of the Z-disc, and TITIN connects the Z-disc to the M-line in the sarcomere. Importantly, hCMs-QKIdel demonstrate significantly altered alternative splicing events in these Z-disc genes. As confirmed by the RIP-qPCR assay, these Z-disc genes are direct targets of QKI. Interestingly, the skipping of exon 8 in ACTN2 also leads to the generation of a premature stop codon, which subsequently leads to the dramatic downregulation of ACTN2 via the NMD pathway. Collectively, our data demonstrate the critical safeguard function of QKI in ensuring normal processing of pre-mRNA splicing for generating a functional myofibril structure suitable for contractile physiology in cardiomyocytes. In addition to Z-disc genes, it is worth noting that many other altered alternative splicing events are found in hCMs-QKIdel. For example, there was noticeable alternative splicing of genes involved in regulating calcium dynamics and E-C coupling, which is among the most significantly altered events in hCMs-QKIdel, including the voltage-gated calcium channel (e.g., CACNA1C, CACNA1G), calcium release channel (e.g., RYR2), and CAMK2D. This alteration likely contributes to the significant delay of E-C coupling in hCMs-QKIdel (Fig. 2C, D). This notion suggests that QKI is also involved in multiple critical components essential for cardiac physiology and pathophysiology. Additional systematic analysis in the future should be given to these unique events, which will help broaden our knowledge of the role that QKI plays in cardiovascular development and function and will help identify its potential involvement in the pathogenesis of cardiomyopathies and cardiovascular diseases.

Interestingly, although our analysis was solely based on cardiomyocytes, many altered alternative splicing events were found in genes highly relevant to neuronal physiology, as demonstrated by the transcriptome analysis (Figs. 4 and 5). There appear many shared QKI-mediated alternative splicing events in the nerve system and cardiomyocytes. Interestingly, a large-scale genome-wide analysis of patients with congenital heart defects demonstrated that a similar candidate gene profile contributes to both congenital heart defects and neurodevelopmental and functional defects.\(^ {76,77}\) It is of great interest to test whether these altered splicing events share common pathways in contributing to both defects in cardiovascular development and neurodevelopment in the future.

In summary, our work has revealed the role of QKI in cardiogeneses and heart function. Our data demonstrate that QKI is a critical alternative splicing regulator in both human and mouse cardiomyocytes, which is not only indispensable for cardiogenesis but also essential to heart function. Our work finds that QKI is potentially involved in the pathogenesis of certain forms of cardiovascular diseases.

**Methods**

**Generation of QKI-null hESC lines.** hESC lines H1, H7, and H9 were purchased from WiCell. To generate the QKI-null hESC line, we utilized the CRISPR/Cas9 genome-editing tool, using the D10A-mutant nickase version of Cas9 (Cas9n) with a pair of offset sgRNAs complementary to opposite strands of the target site. We designed sgRNAa and sgRNB targeting exon 3 of the human QKI gene (http://crispr.mit.edu/8079/). pX462-sgRNAa and pX462-sgRNB were transfected into hESCs by electroporation (Neon Transfection System, ThermoFisher Scientific).

The QKI-null hESC clones were screened by western blot analysis after 1 week of puromycin selection and the positive clones were further validated by genomic sequencing.

**Generation of doxycycline-inducible QKI expression hESC-QKIdel lines.** We used a PB transposon system to generate doxycycline-inducible expression of QKI5, QKI6 and QKI7 in hESC-QKIdel lines (Supplementary Fig. 9A). The QKI5, QKI6, and QKI7 cDNA fragments were inserted into a PB-TRE3G-EGFP-PGK-Neo vector between the BamHI and BsrGI sites. These vectors were cotransfected with a Hyper-PB vector and a PB-CAG-rtTA-IRE-RES-puro vector into hESC-QKIdel lines by electroporation. Electroporated cells underwent one week of puromycin selection and neomycin selection for another week. The positive clones that expressed QKI5, QKI6 or QKI7 under the induction of doxycycline were screened by western blot analysis.

**hESC culture and cardiomyocyte differentiation.** hESCs were cultured in mTeSR medium (STEMCELL #85852). hESCs were differentiated into cardiomyocytes following previously published protocols.\(^ {53}\) In brief, the hESCs were dissociated with Accutase (Milipore Sigma, A18964) and were cultured on Matrigel-coated plates in mTeSR medium. Cardiomyocyte differentiation was initiated when cells were ~95% confluent by culturing cells in RPMI1640/B-27 minus insulin medium RPMI1640 (GIBCO) with B-27 supplement minus insulin (ThermoFisher Scientific, A1895601) supplemented with 6 μM CHIR-99021 (Selleckchem, S1263) for differentiation on days 0 and 1. For days 3 and 4, IWP-2 (3 μM, Tocris, 3533) was added to the fresh medium. From day 5 onward, cells were cultured with fresh RPMI1640/B-27 medium every 2 days. A beating cluster of cells appeared ~7–8 days post differentiation and robust spontaneous-beating sheets occurred by day 10. Cultures were maintained in Dulbecco’s modified Eagle’s medium with 10% fetal bovine serum in a 37 °C and 5% CO2 air environment.

**Immunofluorescent analyses of cultured cells and embryonic tissues.** Individual hESC clones and single cells generated from beating sheets were plated on 15 mm Matrigel-coated circular coverslips. Cells were fixed and permeabilized with cold methanol, which was followed by incubation with 5% BSA for blocking before...
the addition of antibodies for staining. To assess pluripotency, hESC colonies were stained with antibodies against OCT3/4 (Stemgent, 09-0023), SSEA-4 (ThermoFisher Scientific, 41-8000), SOX2 (Stemgent, 099-0024), and TRA-1-60 (Stemgent, 09-0010). To assess differentiated cardiomyocytes, cells were stained with antibodies against TNN2 (Abcam, ab54932), ACTN2 (Sigma, A8711), NEBL (ThermoFisher Scientific, PA5-31506), and TTN (DSHB, 9D10). Anti-β3-integrin antibody (Bethyl, A300-183A) was used in the immunofluorescence staining to confirm the localization of β3-integrin in the membrane. First antibody pairs of cells loaded at 1:100 or at a concentration specified by manufacturers. Corresponding secondary antibodies labeled with either Alexa Fluor 594 or 488 were applied at a dilution of 1:1000 (Donkey anti-Mouse IgG-594: ThermoFisher Scientific, A21203; Goat anti-Rabbit IgG: ThermoFisher Scientific, A11012; Donkey anti-Goat-IgG: ThermoFisher Scientific, A1055; Goat anti-Mouse-IgG: ThermoFisher Scientific, A11001; Goat anti-Rabbit-IgG: ThermoFisher Scientific, A11008) to visualize the immunoreactivity. Then, cells were subsequently counterstained with 4′,6-diamidino-2-phenylindole (DAPI) and sealed with a prolonged antidepressant reagent (ThermoFisher Scientific, P36931).

To determine the histological structure of developing hearts, mouse embryos were fixed, embedded in OCT (Sakura Finetek, Japan), frozen, and cut to produce 10 μm-thick sections, which was followed by the standard procedure of antibody staining, including the use of a brief fixation with 2% paraformaldehyde (PFA) for 5 min, washing with PBST (PBS + 0.1% Triton X-100), treating with M.O.M. mouse IgG blocking reagent (Vector Laboratories), incubation with primary antibody (diluted in M.O.M. reagent) and incubation with secondary antibodies labeled with horseradish peroxidase (Goat anti-Mouse IgG antibody, QKI-5 (Bethyl, A300-183A), QKI6 (Millipore, Cat#AB9906), and QKI7 (Millipore, Cat#A113752) according to the manufacturer's instructions. The antibody signals were detected using a gel documentation system (GelDoc and Image Lab, BIO-RAD). Total protein was extracted from cultured cells or isolated embryonic hearts using RIPA lysis buffer supplemented with a protease inhibitor cocktail (Sigma). Equal amounts (30 μg) of protein were boiled in 1× SDS loading buffer for 10 min and separated on a 4-20% Mini-PROTEAN TGX Stain-Free Gels (Bio-Rad, 4561093), which was followed by transfer to nitrocellulose membranes (Bio-Rad, 1704158) using the Trans-Blot Turbo Transfer System. The membranes were blocked with 3% nonfat dry milk (Bio-Rad, 1704043) in Tris-buffered saline containing Tween-20. The membranes were incubated with primary antibodies against TNN2 (DSHB, Cat#AB582495), ACTN2 (Sigma, Cat#A7811), MF20 (DSHB, Cat#AB217781), MYOZ2 (ThermoFisher, Cat#PAS-76946), TNNJ1 (Santa Cruz, Cat#SC15368), Tubulin (Sigma, Cat#T6191), Pan-QKII(Abcam, Cat#ab126742), QKI-5 (Bethyl, A300-183A), QK6 (Millipore, Cat#AB9890), and QKI7 (Millipore, Cat#A113752) at 1:1000 dilution. Membranes were washed corresponding antibodies labeled with horseradish peroxidase (Goat anti-Mouse IgG antibody, ThermoFisher Scientific, G21040; Mouse anti-Rabbit IgG antibody; Santa Cruz, sc-2357) at 1:4000 dilution. Immunoreactive bands were visualized by ECL (Thermo Scientific, 32106), which was developed and quantified using a gel documentation imaging system (GelDoc and Image Lab, BIO-RAD).

Transmission electron microscopy. Cultured differentiated cardiomyocytes were fixed with 2.5% (v/vol) glutaraldehyde for 15 min at room temperature and post-fixed with 1% osmium tetroxide for 2–3 h. The samples were dehydrated by incubation in graded ethanol (50%, 70%, 90%, and 100%) and in propylene oxide for 10 min. Samples were embedded, sectioned at 70-nm thickness, and stained with lead citrate. Images were captured using a PHILIPS CM-120 TENV (PHILIPS, Holland).

CellRanger 2.1.0 (https://support.10xgenomics.com/) was utilized to process the raw sequence data generated. Briefly, CellRanger used bcftools (https://support.10xgenomics.com/) and the UMI trim function of CellRanger to demultiplex raw but sequence-cells generated from the sequencer into sample specific FASTQ files. The FASTQ files were then aligned to the human reference genome GRCh38 with RNA-seq aligner STAR. The aligned reads were traced back to individual cells and the gene expression level of individual genes were quantified based on the number of UMIs detected in each cell. The filtered gene–cell barcode matrices generated with CellRanger were used for further analysis with the 8 package Seurat (version 2.3.1) with Retuio version 1.1.453 and R version 3.5.1. QC of the data was implemented as the first step in our analysis. We first filtered out genes that detected in less than five cells and cells with less than 200 genes. To further exclude low-quality cells in downstream analysis we used the function Outlier from R package. The scatter plot is the result of visualizing the distributions of relevant UMIs and mitochondrial gene content. Cells with extremely high or low number of detected UMIs were excluded. In addition, cells with high percentage of mitochondrial reads were also filtered out. After removing likely multiplets and low-quality cells, the gene expression levels for each cell were normalized with the `NormalizedData.frame` function in Seurat package. To reduce the false positive rates, the mitochondrial gene expression, we used the `ScaleData` function to linearly reorganize these variations. Highly variable genes were subsequently identified.

To integrate the single-cell data of the control and QKI-deicient samples, functions FindIntegrationAnchors and IntegrateData from the `10X` package v3.1.0 were implemented. We then used the single cell gene expression data from Seurat functions FindNeighbors and FindClusters were applied for shared nearest-neighbor graph-based clustering. The FindConservedMarkers function was subsequently used to identify conserved cell-type marker genes. Cell cluster identities were manually defined with the cluster-specific marker genes or known marker genes. The cell clusters were visualized using the UMAP plots. To compare average gene expression within the same cluster between different conditions, we applied the function AverageExpression. R packages ggplot2 (http://ggplot2.org) and ggrepel (https://github.com/slowkow/ggrepel) were used to plot the average gene expression. Violin plots were used to visualize specific gene expressions across clusters and different sample conditions. To infer the developmental trajectories from the data, Monocle 3 was utilized. The Seurat object from the integrative analysis was converted to SingleCellExperiment object using the function asSingleCellExperiment in Seurat. The SingleCellExperiment object was then used to generate Monocle 3 CDS object. The single-cell count matrix, dimensional reductions and clusters determined from the Seurat analysis were directly imported into the Monocle CDS object. Trajectory graph was learned from the CDS object and the cells were ordered in pseudotime along the learned trajectory. Cells of early progenitors were used as root cells in the order cell process. All single-cell RNA sequence data that support the findings of this study have been deposited in Geo Database (GSE144009).

**Bulk RNA-seq and alternative splicing analysis.** Total RNA samples from cultured cells were prepared using a RNeasy Plus Micro Kit (Qiagen) according to the manufacturer’s instructions. The RNA-seq library was prepared according to the manufacturer’s instructions (KAPA RNA Hyperprep Kit (KK8581)). The cDNA library was sequenced on an Illumina sequencing platform (Illumina HiSeq™ 4000, San Francisco, USA) using paired-end technology. The sequence data were mapped to “UCSC reference genome hg38” for expression analysis and “Ensembl GRCh38v95” for Spliced Transcripts Alignment reference [81]. To evaluate the quality of the RNA-seq data, the number of reads that fell into different annotated regions (exonic, intronic, splicing junction, intergenic, promoter, UTR, etc.) of the reference gene were determined with bamUtils [82]. Low-quality reads (including reads mapped to multiple positions) were excluded and feature Counts was used to quantify the gene level expression [83]. Differential gene expression analysis was performed with edgeR [84]. In this workflow, the statistical methodology applied used negative binomial generalized linear models with likelihood ratio tests. The RNA-seq data have been deposited in Gene Expression Omnibus (GEO).

Splicing analysis was conducted using rMATS (version 3.2.5), in which the percent-spliced-in (PSI) for each splicing event were calculated based on the number of sequencing reads supporting the inclusion and exclusion isoforms, respectively [85]. A FDR and the changes on the exon inclusion levels (deltaPSI) were calculated for each event based on four biological replicates in each condition. The cell clusters were visualized using the UMAP plots. Sequences for the primers used to confirm alternative splicing are provided in Supplementary Table 1.

Whole-mount X-Gal staining and histological analysis of mouse embryos. Mouse embryos were fixed for 15 minutes at room temperature and washed in
washed buffer three times for 15 min. Embryos were placed in X-Gal staining buffer from 4 h to overnight at 37 °C, depending on the level of LacZ activity. X-Gal-stained embryos were fixed with 4% PFA, paraffin-embedded, cut into 10 μm-thick sections, and subjected to nuclear fast red staining.

**RNA immunoprecipitation.** RIP experiments were conducted using an EZ-Magna RIP RNA-Binding Protein Immunoprecipitation Kit (Millipore). Human QK15 antibody was used to pull down RNA. Normal mouse IgG (provided by kit) was used as a negative control. Total RNA was purified using the RNeasy Plus Micro Kit (Qiagen) according to the manufacturer’s instructions.

**Statistical analysis.** Experimental data were reported as the mean ± SEM, unless otherwise indicated. Two-tailed Student’s t-tests were used when comparing the differences of two groups, and one-way analysis of variance test was used to compare among three groups. P-values < 0.05 were considered significantly different. To ensure the reproducibility of our data, all experiments were repeated at least three times with multiple sets of independent samples. GO enrichment and KEGG pathway enrichment analysis of differentially expressed genes were performed respectively using R programming based on the hypergeometric distribution.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

If needed, contact N.S. for original data described in Fig. 7C and Supplementary Fig. 3D, E, and contact W. Shou for all other original data described in this paper. Contact K.Y. for requesting QkiGFP mouse strain, N.S. for requesting hESC-Quaking (H7) line and hESC-Quaking (H7) line and J.N. for requesting the plasmids of inducible PiggyBac transposon overexpression system, and W.S. for requesting all other reagents described in this article. The bulk RNA sequence data have been deposited in GEO Database (GSE144008) and single-cell RNA sequencing data that support the findings of this study have been deposited in GEO Database (GSE144009). In addition, the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) database used in the study are available at [http://geneontology.org/](http://geneontology.org/) and [https://www.genome.jp/kegg/](https://www.genome.jp/kegg/). Source data are provided with this paper.

Received: 22 December 2019; Accepted: 27 November 2020; Published online: 04 January 2021

**References**

1. Day, D. A. & Tuite, M. F. Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. J. Endocrinol. 157, 361–371 (1998).

2. Lee, Y. & Rio, D. C. Mechanisms and regulation of alternative pre-mRNA splicing. Annu. Rev. Biochem. 84, 291–323 (2015).

3. Zhu, C., Chen, Z. & Guo, W. Pre-mRNA mis-splicing of sarcomeric genes in the developing heart. Cardiovasc. Toxicol. 11, 235–243 (2011).

4. Lotson, C. L., Hahnen, E., Androphy, E. J. & Wirth, B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc. Natl Acad. Sci. USA 96, 6307–6311 (1999).

5. Ho, C. K., Musa, F. R., Bell, C. & Walker, S. W. LDLR gene synonymous mutation c.1813C>T results in mRNA splicing variation in a kindred with familial hypercholesterolaemia. Ann. Clin. Biochem. 52, 680–684 (2015).

6. Deduc, Y., Bourque, L., Poirier, J. & Dufour, R. Role of rs8466662 and HNF1α alternative splicing in statin efficacy and baseline lipid levels in familial hypercholesterolemia. Pharmacogenet. Genomics 26, 1–11 (2016).

7. Refaat, M. M. et al. Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. Heart Rhythm 9, 390–396 (2012).

8. Wang, J. et al. RBM24 is a major regulator of muscle-specific alternative splicing. Dev. Cell 31, 87–99 (2014).

9. Poon, K. L. et al. RNA-binding protein RBM24 is required for sarcomere assembly and heart contractility. Cardiovasc. Res. 94, 418–427 (2012).

10. Feng, Y. et al. SRp38 regulates alternative splicing and is required for Cal2+ handling in the embryonic heart. Dev. Cell 16, 528–538 (2009).

11. Xu, Y. et al. ASF/SF2-regulated Elav1 alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Cell 120, 59–72 (2005).

12. Gao, C. et al. ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Circ. Res. 118, 1891–1902 (2016).

13. Ichida, M. et al. Differential regulation of exonic regulatory elements for muscle-specific alternative splicing in the developing heart. Cardiovasc. Toxicol. 14, 275–286 (2014).

14. Kalsotra, A. et al. A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. Proc. Natl Acad. Sci. USA 105, 20333–20338 (2008).

15. Lichtner, P. et al. Expression and mutation analysis of BRUNOL3, a candidate gene for heart and thymus developmental defects associated with partial heart/hypothymus agenesis. J. Mol. Med. 80, 431–442 (2002).

16. Fan, C., Chen, Q. & Wang, Q. K. Functional role of transcriptional factor TBX5 in pre-mRNA splicing and Holt-Oram syndrome via association with SC35. J. Biol. Chem. 284, 25653–25663 (2009).

17. Ichida, M. et al. Differential regulation of exonic regulatory elements for muscle-specific alternative splicing during myocardogenesis and cardiogenesis. J. Biol. Chem. 275, 15992–16001 (2000).

18. Denz, C. R. et al. Absence of mutation at the 5′-upstream promoter region of the TPM4 gene from cardiac mutant axolotl (Ambystoma mexicanum). Cardiovasc. Toxicol. 11, 235–243 (2011).

19. Denz, C. et al. Absence of mutation at the 5′-upstream promoter region of the TPM4 gene from cardiac mutant axolotl (Ambystoma mexicanum). Cardiovasc. Toxicol. 11, 235–243 (2011).

20. Leduc, Y., Bourque, L., Poirier, J. & Dufour, R. Role of rs8466662 and HNF1α alternative splicing in statin efficacy and baseline lipid levels in familial hypercholesterolemia. Pharmacogenet. Genomics 26, 1–11 (2016).

21. Refaat, M. M. et al. Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. Heart Rhythm 9, 390–396 (2012).

22. Wang, J. et al. RBM24 is a major regulator of muscle-specific alternative splicing. Dev. Cell 31, 87–99 (2014).

23. Poon, K. L. et al. RNA-binding protein RBM24 is required for sarcomere assembly and heart contractility. Cardiovasc. Res. 94, 418–427 (2012).

24. Feng, Y. et al. SRp38 regulates alternative splicing and is required for Ca2+ handling in the embryonic heart. Dev. Cell 16, 528–538 (2009).

25. Xu, Y. et al. ASF/SF2-regulated Elav1 alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Cell 120, 59–72 (2005).

26. Gao, C. et al. ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Circ. Res. 118, 1891–1902 (2016).

27. Kondo, T. et al. Genomic organization and expression analysis of the mouse CELF family. Genes Dev. 22, 2467–2477 (2008).

28. Denz, C. et al. Absence of mutation at the 5′-upstream promoter region of the TPM4 gene from cardiac mutant axolotl (Ambystoma mexicanum). Cardiovasc. Toxicol. 11, 235–243 (2011).

29. Leduc, Y., Bourque, L., Poirier, J. & Dufour, R. Role of rs8466662 and HNF1α alternative splicing in familial hypercholesterolemia. Pharmacogenet. Genomics 26, 1–11 (2016).

30. Refaat, M. M. et al. Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. Heart Rhythm 9, 390–396 (2012).

31. Wang, J. et al. RBM24 is a major regulator of muscle-specific alternative splicing. Dev. Cell 31, 87–99 (2014).

32. Poon, K. L. et al. RNA-binding protein RBM24 is required for sarcomere assembly and heart contractility. Cardiovasc. Res. 94, 418–427 (2012).

33. Ichida, M. et al. Differential regulation of exonic regulatory elements for muscle-specific alternative splicing in the developing heart. Cardiovasc. Toxicol. 11, 235–243 (2011).

34. Denz, C. et al. Absence of mutation at the 5′-upstream promoter region of the TPM4 gene from cardiac mutant axolotl (Ambystoma mexicanum). Cardiovasc. Toxicol. 11, 235–243 (2011).
Hojo, H. et al. The RNA-binding protein QKI-7 recruits the poly(A) polymerase GLD-2 for 3‘ adenylation and selective stabilization of microRNA-122. J. Biol. Chem. 295, 390–402 (2020).

Larocque, D. et al. An RNA-binding protein, Qki5, regulates embryonic organogenesis. J. Cell Mol. Med. 23, 6578–6594 (2019).

Doukhane, E., Gavino, C., Haines, J. D., Almazan, G. & Richard, S. The QKI-6 RNA binding protein regulates actin-interacting protein-1 mRNA stability during oligodendrogliocyte differentiation. Mol. Biol. Cell 21, 3029–3040 (2010).

Galarneau, A. & Richard, S. The STAR RNA binding proteins GLD-1, QKI, SAM68 and SLM-2 bind bipartite RNA motifs. Nucleic Acids Res. 44, 2475–2490 (2016).

Li, Z. et al. Defective smooth muscle development in qk-deficient mice. Dev. Growth Differ. 45, 449–462 (2003).

Yu, H.-Y., Dawson, M. R., Reynolds, R. & Hardy, J. I. Expression of QKI proteins and MAP1B identifies actively myelinating oligodendrocytes in adult rat brain. Mol. Cell Neurosci. 17, 292–302 (2001).

Wang, W. et al. A single-cell transcriptional roadmap for cardiopharyngeal fate diversification. Nat. Cell Biol. 21, 674–686 (2019).

Stuart, T. et al. Comprehensive integration of single-cell data. Cell 177, 1888–1902 e1821 (2019).

Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502 (2019).

Chen, Y., Tian, D., Ku, L., Osterhout, D. J. & Feng, Y. The selective RNA-binding protein quaking (QKI) is necessary and sufficient for promoting oligodendrogliocyte differentiation. J. Biol. Chem. 282, 23553–23560 (2007).

Hayakawa-Yano, Y. et al. An RNA-binding protein, Qki5, regulates embryonic neural stem cells through pre-mRNA processing in cell adhesion signaling. Genes Dev. 31, 1910–1925 (2017).

Shen, S. et al. rMATs: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl Acad. Sci. USA 111, E5593–E5601 (2014).

Galarneau, A. & Richard, S. The STAR RNA binding proteins GLD-1, QKI, SAM68 and SL2-2 bind bipartite RNA motifs. BMC Mol. Biol. 10, 47 (2009).

Martin, L. et al. Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer Res. 74, 3104–3113 (2014).

Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457–463 (2010).

Kelemen, O. et al. Function of alternative splicing. Gene 514, 1–30 (2013).

Aberg, K., Saetre, P., Jareborg, N. & Jazin, E. Human QKI1, a potential regulator of mRNA expression of human oligodendroglial-related genes involved in schizophrenia. Proc. Natl Acad. Sci. USA 103, 7482–7487 (2006).

Backs, L. et al. Haploinsufficiency of the gene Quaking (QKI) is associated with the 6q terminal deletion syndrome. Am. J. Med. Genet. A 152A, 319–326 (2010).

Bockbrader, K. & Feng, Y. Essential function, sophisticated regulation and pathological impact of the selective RNA-binding protein QKI in CNS myelin development. Brain Res. Mol. Brain Res. 655, 668–669 (2019).

Li, F. et al. QKI5-mediated alternative splicing of the histone variant macroH2A1 regulates gastric carcinogenesis. Oncotarget 7, 32821–32834 (2016).

de Miguel, F. J. et al. A large-scale analysis of alternative splicing reveals a key role of QKI in lung cancer. Mol. Oncol. 10, 1457–1449 (2016).

Dubre, M.-P. et al. Genome-wide association study for incident myocardial infarction and coronary heart disease in prospective cohort studies: the CHARGE Consortium. PLoS ONE 11, https://doi.org/10.1371/journal. pone.0144997 (2016).

Wang, F., Yuan, Y., Yang, P. & Li, X. Extracellular vesicles-mediated transfer of miR208b/a exaggerate hypertrophy/eosinophy injury in cardiomyocytes by reducing QKI expression. Mol. Cell Biochem. 431, 187–195 (2017).

Ueyama, T., Kasahara, H., Ishiwata, T., Nie, Q. & Izumo, S. Myocardin expression is regulated by Nkx2.5, and its function is required for cardiomyogenesis. Mol. Cell Biol. 23, 9222–9232 (2003).

Hoffnagle, M. H. et al. Myocardin is differentially required for the development of smooth muscle cells and cardiomyocytes. Am. J. Physiol. Heart Circ. Physiol. 300, H1707–H1721 (2011).

Jin, S. C. et al. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat. Genet. 49, 1593–1601 (2017).

Zaidi, S. & Brueckner, M. Genetics and genomics of congenital heart disease. Circ. Res. 120, 923–940 (2017).

Lian, X. et al. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. Proc. Natl Acad. Sci. USA 109, E1848–E1857 (2012).

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420 (2018).

McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: preprocessing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33, 1179–1186 (2017).

Dobin, A. et al. STAR: ultralast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2012).

Breeze, M. R. & Liu, Y. NGStools: a suite software for analyzing and manipulating next-generation sequencing datasets. Bioinformatics 29, 494–496 (2013).

Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2013).

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2009).

Shen, S. et al. rMATs: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl Acad. Sci. USA 111, E5593–E5601 (2014).
