SCIENTIFIC REPORTS

OPEN

Comparison of multi-lineage differentiation of hiPSCs reveals novel miRNAs that regulate lineage specification

Lu Li[1,2], Kai-Kei Miu[1], Shen Gu[1,3], Hoi-Hung Cheung[1] & Wai-Yee Chan[1]

MicroRNAs (miRNAs) are known to be crucial players in governing the differentiation of human induced pluripotent stem cells (hiPSCs). Despite their utter importance, identifying key lineage specifiers among the myriads of expressed miRNAs remains challenging. We believe that the current practice in mining miRNA specifiers via delineating dynamic fold-changes only is inadequate. Our study, therefore, provides evidence to pronounce "lineage specificity" as another important attribute to qualify for these lineage specifiers. Adopted hiPSCs were differentiated into representative lineages (hepatic, nephric and neuronal) over all three germ layers whilst the depicted miRNA expression changes compiled into an integrated atlas. We demonstrated inter-lineage analysis shall aid in the identification of key miRNAs with lineage-specificity, while these shortlisted candidates were collectively known as "lineage-specific miRNAs". Subsequently, we followed through the fold-changes along differentiation via computational analysis to identify *miR-192* and *miR-372-3p*, respectively, as representative candidate key miRNAs for the hepatic and nephric lineages. Indeed, functional characterization validated that *miR-192* and *miR-372-3p* regulate lineage differentiation via modulation of the expressions of lineage-specific genes. In summary, our presented miRNA atlas is a resourceful ore for the mining of key miRNAs responsible for lineage specification.

Extensive evidence has shown that human pluripotent stem cells (hPSCs) can give rise to all somatic cell types in an adult human[1–4]. Understanding the mechanisms by which hPSCs differentiate during early embryogenesis is crucial to understanding human development. Moreover, a comprehensive understanding of the mechanisms underlying hPSC differentiation is required for their application in disease studies, where intentionally differentiated functional cells derived from hPSCs have been widely used for disease modeling and replacement therapy[5–7].

A previous in-depth analysis of the transcriptome reported a temporal change in microRNA (miRNA) expression levels during hPSC differentiation[8], suggesting that the expression dynamics of miRNAs are associated with hPSC differentiation[9]. In addition, the identification of miRNAs regulating lineage specification such as *miR-375*, which regulates pancreatic islet formation, further confirms that miRNAs can serve as cell fate determinants[8,10–15]. However, identification of miRNAs regulating lineage specification, hereafter referred to as key miRNAs in this manuscript, remains challenging since there are massive numbers of miRNAs potentially regulating lineage formation[16]. In previous studies, key miRNAs were mainly revealed by the analysis of expression changes during differentiation of a single lineage (intra-lineage analysis)[9,17,18]. Following such analysis, only those key miRNAs with the greatest fold-changes were selected for further study[19–21], while miRNAs that changed less dramatically were avoided.

Another criterion, lineage specificity, has also been used to identify key miRNAs, which alleviates the drawbacks of solely considering fold-change as the filtration criterion. Lineage-specific miRNAs have been reported previously[22,23]; several of them have been confirmed to determine cell identity[10]. For instance, *miR-124* and *miR-9*, which are neuron-specific miRNAs, have been reported to direct neuronal differentiation and regulate brain

[1]CUHK-CAS GIBH Joint Research Laboratory on Stem Cell and Regenerative Medicine, School of Biomedical Sciences, the Chinese University of Hong Kong, Shatin, N.T., Hong Kong SAR. [2]Present address: School of Pharmacy, University of Southern California, Los Angeles, CA, USA. [3]Present address: M&H Genetics/Baylor Genetics Laboratories, Baylor College of Medicine, Houston, TX, USA. Correspondence and requests for materials should be addressed to H.-H.C. (email: cheunghh@cuhk.edu.hk) or W.-Y.C. (email: chanwy@cuhk.edu.hk)

development[24]. Many lineage-specific miRNAs have been revealed by the comparison of the miRNA transcriptome (miRNAome) between terminally differentiated tissues[23,25,26]. However, few studies have examined different lineages at the earliest stages of development. Therefore, profiling of the miRNAome of the three germ layers and further lineages induced from hPSCs would allow the identification of lineage-specific miRNAs that affect the earliest cell fate decisions.

Human embryonic stem cells (hESCs) are the most frequently used hPSCs in *in vitro* studies of human development[27], although the full-scale application of hESCs in regenerative medicine and disease modeling is hindered by ethical and technical issues[28]. Human induced PSCs (hiPSCs), which resemble hESCs, have become an alternative cell type for disease modeling and drug selection[7]. hiPSCs induced from somatic cells of individual patients enable the establishment of hiPSC banks corresponding to all HLA haplotypes. Considering the necessity of integrating information from different patients, generation of a reference dataset of miRNAome based on hiPSCs of healthy individuals that could be used for comparison with any patient-specific hiPSCs is urgently required[29].

In this study, we performed comparisons among three lineages (inter-lineage analysis) to reveal lineage-specific miRNAs at the early stages of development. From these miRNAs, we further identified novel key miRNAs and validated their key regulatory roles in early lineage specification. To perform the inter-lineage analysis, we induced hiPSCs into three representative lineages for the three germ layers and profiled the expression changes of miRNAs during differentiation. By integrating all profiling results, we built a miRNA atlas that allows comparisons of miRNA expression across three lineages. Based on this atlas, those miRNAs with lineage-specific expressions, such as *miR-192* and *miR-372-3p*, were easily identified. After selecting key miRNA candidates by inter-lineage analysis, we conducted extensive experiments to determine whether they were true key regulators. Targets of *miR-192* and *miR-372-3p* in lineage specification were predicted computationally. Their repression on predicted targets and their regulatory effects on lineage differentiation were validated experimentally. These results strongly supported the hypothesis that new key miRNAs could be precisely identified from our atlas using inter-lineage analysis.

Taken together, we have developed the inter-lineage analysis methodology to identify key miRNAs of early hPSCs differentiation by considering lineage specificity. The inter-lineage analysis is complementary to intra-lineage analysis and allows key miRNAs with less dramatic fold-changes to be identified. Our study differs from previous studies primarily in terms of: (1) determining which miRNAs affect the very earliest stages of hPSCs differentiation (key miRNAs); (2) the development of a computational method to identify key miRNAs from lineage-specific miRNAs; and (3) the generation of an integrated miRNA expression atlas for all three germ layers and their derived lineages.

Results

Step-wise *in vitro* differentiation of hiPSCs into hepatocytes, nephron progenitors, and neural progenitors.

To clarify the miRNA dynamics involved in hiPSCs differentiation, we established a multi-lineage differentiation system following published protocols that show high differentiation efficiency[30–32]. We differentiated hiPSCs into representative lineages of the three germ layers: hepatocytes for the endoderm (12 days), nephron progenitors for the mesoderm (18 days), and neural progenitors for the ectoderm (11 days), respectively (Fig. 1A and Supplementary Fig. S1A). The hiPSC line used in this study (iBC 1.2) meets all established criteria for hiPSCs, as reported in our previous studies[33,34]. The expression of representative markers for each lineage was detected by quantitative real-time PCR (qPCR) to confirm the cell identity at each time-point (Fig. 1B and Supplementary Fig. S1B–D). The expression of the pluripotent marker *OCT4* declined during differentiation of all three lineages, indicating a loss of pluripotency. Meanwhile, germ layer markers and lineage-specific markers appeared in a sequential order, as expected. Markers for hepatocytes, metanephric mesenchyme (MM), and neural progenitors were confirmed using immunofluorescent staining (IFC) analysis and flow-cytometric analysis, indicating that the purities of the three lineages were all around 50% (Supplementary Fig. S1E–K). Hepatocytes and neurons were additionally confirmed by functional assays and IFC, respectively (Fig. 1C).

Intra-lineage analysis of differentially expressed miRNAs confirms previously reported key miRNAs.

We first asked how miRNAs are differentially expressed when hiPSCs are differentiated into individual lineages (intra-lineage analysis). RNA was collected at days 0, 3, 6, and 10 (n = 2) of each lineage differentiation, as illustrated in Fig. 1D. The miRNA transcriptome (miRNAome) was examined by small RNA microarray analysis. We processed the microarray results with Partek Genomic Suite software following a standard workflow for miRNA expression. The differentially expressed small RNAs within each lineage are summarized in Supplementary Table S1. Specifically, a one-way ANOVA analysis was used to determine which small RNAs had differences in expression between time-points. Subsequent pairwise comparison between successive time-points was used to identify when small RNAs demonstrated significant differential expression (post-hoc testing, fold-change ≥2 or ≤−2, false discovery rate <0.05). We used the false discovery rate (FDR) instead of the *P* value to decrease false positives, as performed in previous studies[8,33]. These analyses identified 170, 177, and 1040 differentially expressed small RNAs in hepatocyte differentiation (HD), nephron progenitor differentiation (KD), and neural progenitor differentiation (ND), respectively (Fig. 2A–C).

Bi-directional hierarchical clustering indicated distinct expression dynamics of small RNAs when different lineages were established (Fig. 2A–C). In HD, stage 1 and 2 samples clustered together, while stage 3 and 4 samples shared a similar pattern (Fig. 2A). In contrast, in KD and ND, stage 2, 3 and 4 samples clustered together, separate from stage 1 samples (Fig. 2B,C). Within each lineage, small RNAs were separated into different groups according to their differential dynamics changes. To clearly reveal these dynamics, miRNAs with the greatest fold-changes in each group were plotted in different colors (Fig. 2D–F). To ascertain the dynamic range of the microarray platform is appropriate for reliable detection, the expression of the five most-upregulated miRNAs in each lineage were further validated by qPCR (Supplementary Fig. S2).
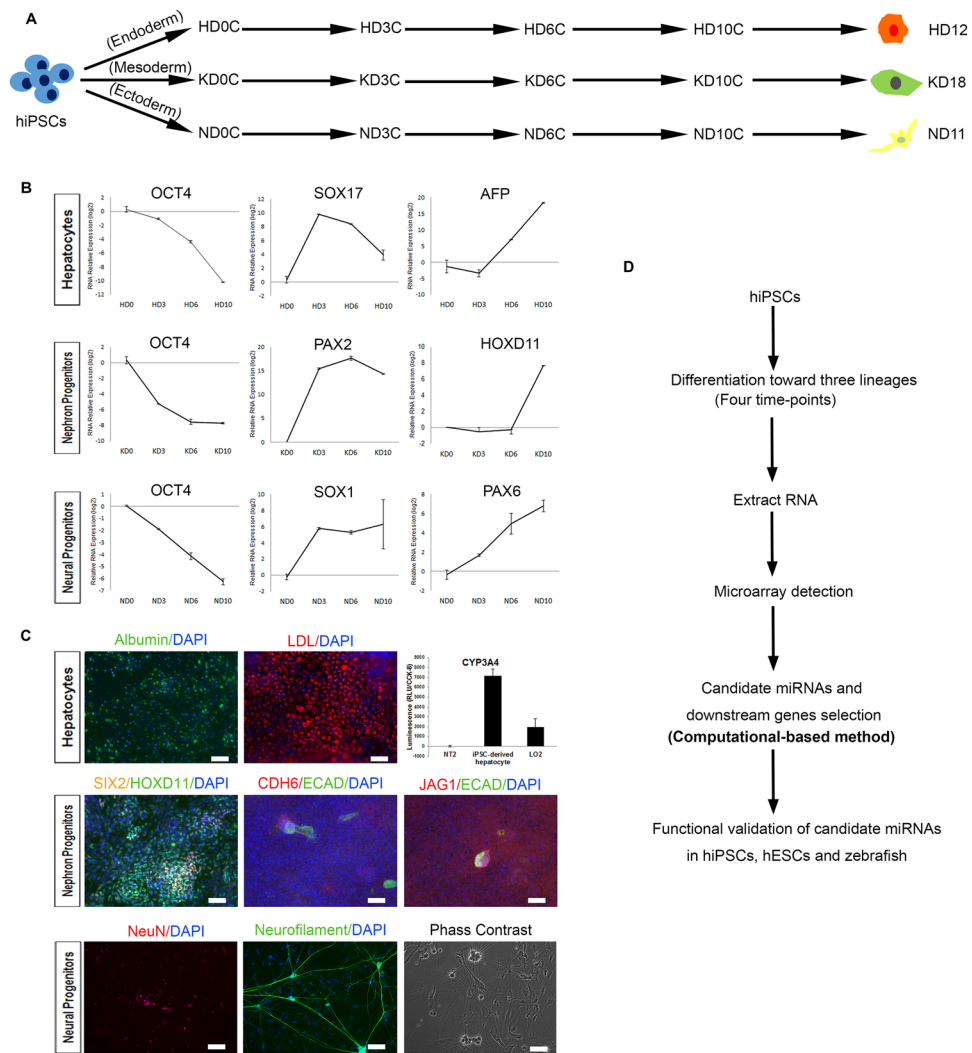
Figure 1. Outline of the experimental design and summary of the multi-lineage induction. (**A**) Schematic overview of hiPSC differentiation into hepatocytes, nephron progenitors, and neural progenitors. (**B**) qPCR results showing the expression tendencies of pluripotency marker (*OCT4*), markers for endoderm (*SOX17*), mesoderm (*PAX2*), ectoderm (*SOX1*), and representative markers for hepatocytes (*AFP*), metanephric mesenchyme (*HOXD11*), and neural progenitors (*PAX6*) at four time-points. Values represent means ± SD (n = 2 independent cultures for each time-point). (**C**) Functional characterization of terminal cells, including hepatocytes (12 days), nephron progenitors (18 days) and neurons (40 days). For hepatocyte differentiation, *Albumin* (green fluorescence) indicates hepatocytes; LDL uptake assay indicates the LDL receptor activity in hepatocytes; and CYP450 assays show the cytochrome P450 activity of hepatocytes. For kidney differentiation, early metanephric mesenchyme marked by SIX2 (yellow fluorescence) and HOXD11 (green fluorescence), and nephron vesicles marked by CDH6 (green fluorescence), ECAD (red fluorescence) and JAG1 (red fluorescence) were induced successfully. For neuron differentiation, the nucleus and the axons of neurons were marked by NeuN (red fluorescence) and Neurofilament (green fluorescence), respectively. Phase contrast imaging showing the morphology of an induced neuron. Scale bars represent 100 μm. (**D**) Schematic overview of the experimental design. HD: hepatocyte differentiation; KD: nephron progenitor differentiation; ND: neural progenitor differentiation.

Using the intra-lineage analysis, we were able to recapitulate previously reported key miRNAs. Specifically, we ranked the absolute fold-changes of miRNAs during lineage differentiation. Functions of miRNAs with the largest fold-change (top 10 ranking) were identified from previously reported studies and are summarized in Table 1. As expected, most of the miRNAs with large fold-changes (9 of 10) in both HD and KD were associated with corresponding lineage establishment. For example, *miR-122*, which changes the most from HD 0 to HD 6 (positive fold-change of 1,022), has been reported to be a liver-specific miRNA that functions in the maintenance of liver phenotype and metabolism of cholesterol[35,36].

In summary, the success in recapitulating many previously reported key miRNAs by intra-lineage analysis effectively supports the use of our profiling data as a reliable source for identifying key miRNAs.
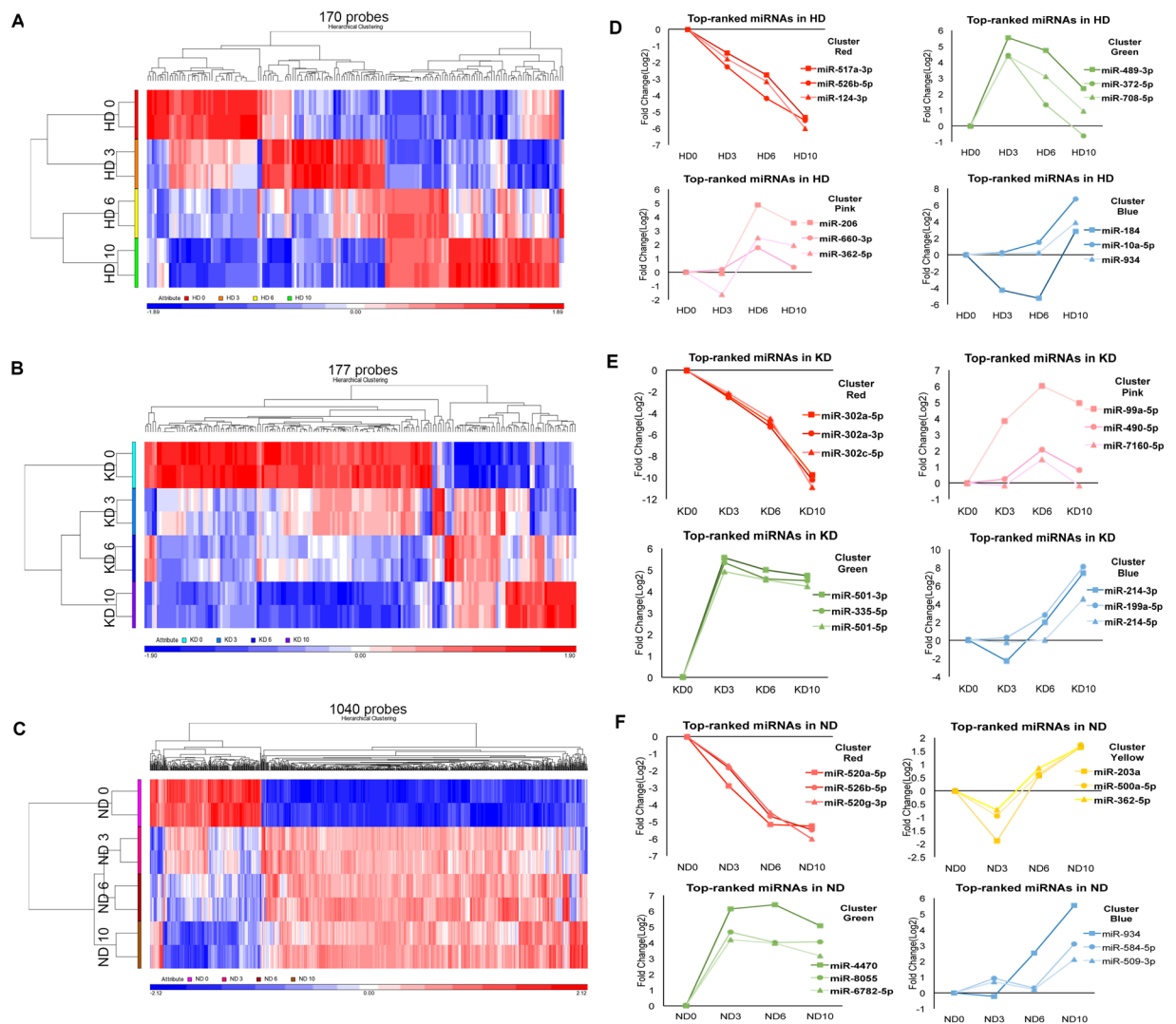
Figure 2. Intra-lineage analysis of differentially expressed miRNAs. (**A**–**C**) Hierarchical clustering of 170 differentially expressed small RNAs in the first ten days of HD, 177 differentially expressed small RNAs of KD, and 1,040 differentially expressed small RNAs of ND, respectively. The values of two biological replicates (indicated in color) for each time-point are shown. (**D**–**F**) The miRNAs with differential dynamics changes were plotted in different colors: small RNAs that decreased with differentiation (red); small RNAs enriched in stage 1 and 4 but downregulated at stage 2 (yellow); small RNAs increased at stage 2 but decreased at later stages (green); small RNAs enriched in stage 3 (pink); and small RNAs only increased at the latest stage (blue).

Inter-lineage analysis reveals lineage-specific miRNAs.   Identification of key miRNAs by intra-lineage analysis has traditionally relied on fold-change, leading to the identification of only key miRNAs with large fold-changes. Here, we applied another criterion to identify key miRNAs, one that focuses on their differential expression across lineages in addition to considering fold-change[23,25,26].

According to previous studies, miRNAs with expression restricted to one lineage contribute to cell identity[10], suggesting that the expression changes of miRNAs (upregulation or downregulation) have a significant potential to specify lineages. Therefore, we hypothesized that miRNAs that specifically change at one lineage are potential candidates for key miRNAs. For convenience, we termed these miRNAs "lineage-specific miRNAs". Here, we describe how we identify them by comparing the expression of miRNAs across lineages (inter-lineage analysis).

First, we subjected all small RNA transcriptome data (24 samples in total for all three lineages) to principal component analysis (PCA). The 24 samples automatically segregated into four groups, including undifferentiated hiPSCs with different pre-induction methods (HD 0, ND 0, KD 0, yellow sphere), differentiating hepatocytes (HD 3–10, pink sphere), differentiating nephron progenitors (KD 3–10, green sphere), and differentiating neural progenitors (ND 3–10, blue sphere) (Fig. 3A).

Next, to further show the expression patterns of small RNAs in the establishment of different lineages, we performed hierarchical clustering of all identified differentially expressed small RNAs (1,812 small RNAs) in three lineages (Fig. 3B). The bi-direction hierarchical clustering automatically resulted in several subclusters with distinct expression patterns (Fig. 3B)[37]. Within these subclusters of small RNAs, we focused primarily on miRNAs

miRNAs	Expression pattern	Reported functions in lineages
miR-10a-5p	HD 10 (up)	Liver fibrosis[62]
miR-122-5p	HD 6-10 (up)	Expressed in developing liver tissue[63]/Liver-specific miRNA and maintenance of adult-liver phenotype[35,36]/Cholesterol and lipid metabolism[36,64]
miR-124-3p	HD 3–10 (down)	Cholangiocyte proliferation[65]/Brain-specific miRNA[65,66]
miR-146b-5p	HD 10 (up)	Intestinal epithelial cell differentiation[67]/Visceral preadipocyte proliferation and differentiation[68]/Regulator of inflammation in epithelial cells, lung fibroblasts and others[69]
miR-184	HD 10 (up)	—
miR-192-5p/194-3p	HD 6–10 (up)	Liver-abundant miRNAs[70]/Intestinal epithelial cell differentiation[71]
miR-489-3p	HD 3 (enriched)	Increased in all germ layers, highest in the endoderm[11]
miR-517a-3p/-526b-5p	HD 3–10 (down)	Members of C19MC cluster, which is involved in pluripotent stem cell status[72,73]
miR-10a-5p	KD 6 (enriched)	Biomarkers of kidney injury[74]
miR-10b-5p	KD 6 (enriched)	—
miR-181a-5p	KD 10 (up)	Expressed in kidney tissue[75]/Hematopoietic lineage differentiation[75]
miR-214–3p	KD 10 (up)	Differentiation of ESCs into endothelial cells[76]/Osteogenic differentiation[77]
miR-302a-5p/3p/-302c-5p/-302d-5p	KD 10 (down)	hPSC-specific miRNAs[22]/Definitive endoderm formation[12]
miR-371a-3p	KD 10 (down)	hPSC-specific miRNAs[22]/Mesendoderm specification[78]
let-7e-5p	KD 10 (up)	Early nephrogenic differentiation[38]
miR-10a-5p	ND 10 (up)	Neuron differentiation[79]
miR-181a-2-3p	ND 3 (down)	Regulation of synaptic functions[80]
miR-1208	ND 6 (enriched)	—
miR-375	ND 6–10 (up)	Inhibition of neurite differentiation[81]/Spinal motor neuron development[82]
miR-4470	ND 3–6 (up)	—
miR-520a-5p/-520g-3p/-526b-5p	ND 3–10 (down)	Members of C19MC cluster, which is involved in pluripotent stem cell status[72,73]
miR-664b-5p	ND 3–10 (up)	—
miR-934	ND 6–10 (up)	—

Table 1. Top 10 miRNAs with previously reported development-associated functions.

and denoted them as lineage-specific miRNAs. To clearly show these lineage-specific miRNAs in the atlas, we delineated them with black boxes in the heat-map and colored them in the dendrogram (Fig. 3B). For instance, the subcluster of miRNAs colored green in the dendrogram was specifically elevated in KD 3 but not in HD or ND (Fig. 3B). All lineage-specific miRNAs with normalized expression values are summarized in Supplementary Table S2. As expected, several of them have been previously reported to specify lineage, such as *miR-124* in ND and *let-7e* in KD[38,39], supporting lineage specificity as an effective criterion to identify key miRNAs.

Identification of novel key miRNAs from lineage-specific miRNAs. Lineage-specific miRNAs form a candidate set of key miRNAs (273 lineage-specific miRNAs). We next sought to select the most likely potential candidates for validation. However, since the cardinality of the candidate set (Supplementary Table S2) was large, it was difficult to evaluate every candidate and identify key miRNAs. Therefore, we focused on the HD 6–10 subcluster (purple and blue subcluster) and the KD 10 subcluster (orange subcluster). At these time-points, early lineages have formed, considering that hepatoblasts and metanephric mesenchyme are progenitors of hepatocytes and nephrons, respectively.

One straightforward method to identify key miRNAs from lineage-specific miRNAs is based on fold-change. By ranking the fold-change of lineage-specific miRNAs, we found that both *miR-192-5p* and *miR-192-3p* ranked the highest among the HD 6–10 subcluster (Fig. 3B and Supplementary Table S2), increasing our suspicion that they were candidate key miRNAs. However, selection by fold-change can only identify key miRNAs that change substantially, while it ignores those changing less dramatically. We therefore developed a computational method to expand the search range (Supplementary Fig. S3). Briefly, given that key miRNAs should regulate critical genes in specific lineages, we first searched miRNAs targeting critical genes in KD, and then intersected these identified miRNAs with lineage-specific miRNAs to obtain final key miRNA candidates (Supplementary Fig. S3A). We focused on *PKD1* and *PKD2*, which are key regulatory genes that encode Polycystin 1 (PC1) and Polycystin 2 (PC2), respectively. They are upregulated synergistically during the early stages of embryonic renal development and are important in the developing kidney for maintaining the differentiated phenotype of tubular epithelium[40]. We then asked whether there were any miRNAs contributing to kidney development via regulation of *PKD1* or *PKD2*. TargetScan generated a list of potential miRNAs for targeting *PKD1/PKD2*. Because *PKD1/PKD2* were upregulated in KD (Supplementary Fig. S3B), we searched the specifically downregulated miRNAs in KD and compared the results with the TargetScan analysis. Using this approach, we recapitulated the *miR-106b~25* cluster and *miR-17~92* cluster (Supplementary Fig. S3C), which showed an increased expression tendency in a polycystic kidney disease (PKD) mouse model[41]. More importantly, the overexpression of the *miR-17~92* cluster was sufficient to produce cyst-like structures in the mouse model[41]. In addition to previously reported miR-NAs, we identified *miR-372-3p*, which has well-established enrichment in hESCs but has never been reported in kidney development[42] (Supplementary Fig. S3D). Our array results indicated that *miR-372-3p* was specifically
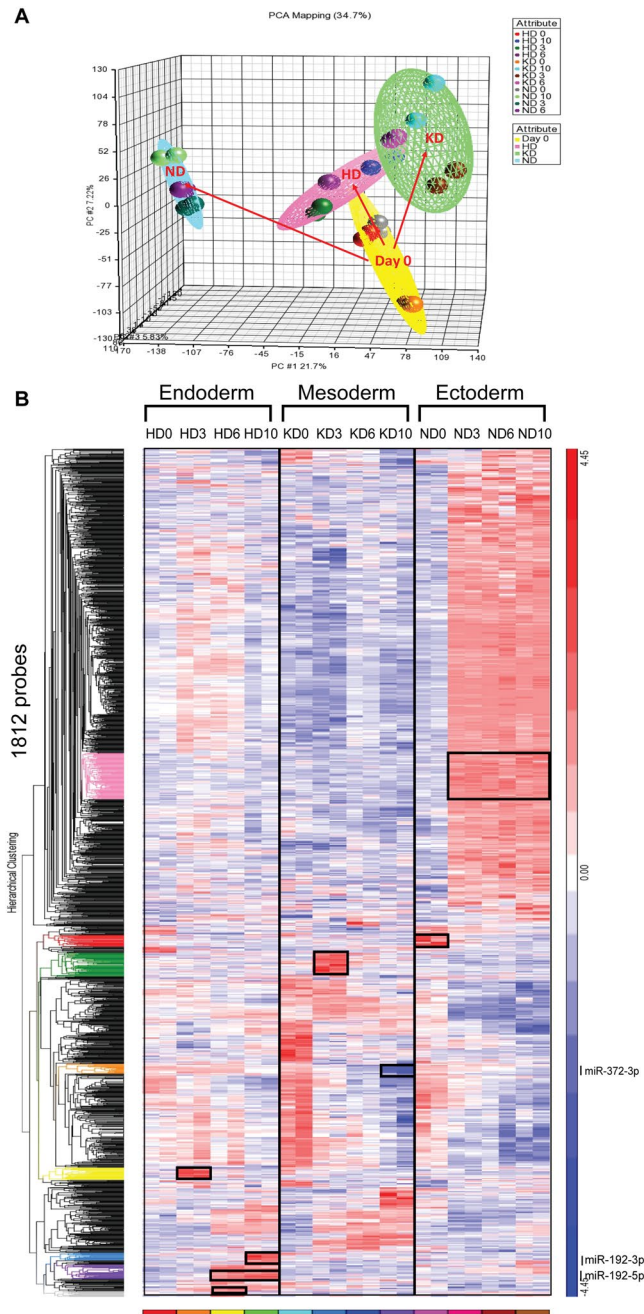
Figure 3. Inter-lineage analysis reveals lineage-specific miRNAs. (**A**) Principal component analysis (PCA) showing the unsupervised division of samples into a day 0 group (yellow sphere), an HD group (pink sphere), a KD group (green sphere) and a ND group (blue sphere). (**B**) Comprehensive heat map showing distinct dynamics of small RNA expression patterns during lineage specification. Colored dendrograms and black boxes indicate small RNAs that were only upregulated or downregulated in one lineage. Particularly, gray, purple, blue and yellow dendrograms indicate small RNAs that were only changed in HD. Orange and green dendrograms indicate small RNAs that were only changed in KD, Red and pink dendrograms indicate small RNAs that were only changed in ND. The newly identified key miRNAs (*miR-192-3p/5p* and *miR-372-3p*) are indicated in the dendrograms that they are included.

downregulated from KD 6 to 10 (Fig. 3B). TargetScan analysis showed that both *PKD1* and *PKD2* are targeted by *miR-372-3p* (Supplementary Fig. S3E).

Effects of novel key miRNAs on lineage specification. To experimentally validate that *miR-192* and *miR-372-3p* indeed regulate lineage specification, we performed functional characterization to determine their effects. The experimental validations were performed in both hESCs (H1) and hiPSCs (iBC 1.2) to rule out the possibility of hiPSC-specific "artifacts".

Effects of miR-192 on hepatocyte differentiation.   We first sought to determine what the targets of *miR-192* in HD. To predict targets of *miR-192* in a more deterministic manner, we applied bioinformatics analysis by combining published transcriptional datasets and the *in silico* target prediction (Fig. 4A). The qPCR analysis confirmed that *miR-192-5p/3p* were only upregulated during HD (Fig. 4B and Supplementary Fig. S4A). Due to inhibitory effects of miRNAs on target mRNAs, only downregulated genes (fold-change ≤ -2) from two independent pro-filing studies were identified (Supplementary Fig. S4B and Table S3)[43,44]. With further filtration by miRWalk, in which the interactions between miRNAs and mRNAs could be predicted by different online tools, common targets of *miR-192-3p* and *miR-192-5p* were identified (Supplementary Fig. S4C and Table S3). By overlapping the list of downregulated genes and targets of *miR-192*, we identified 123 potential targets of *miR-192* in HD (Supplementary Fig. S4D and Table S3).

As the fold-changes of genes detected by the two datasets were modulated, we randomly selected several genes for validation in our HD samples. The qPCR analysis showed that Calneuron 1 (*CALN1*), MGAT4 family member C (*MGAT4C*) and adenylate cyclase activating polypeptide 1 (pituitary) receptor type I (*ADCYAP1R1*) were downregulated (Fig. 4C and Supplementary Fig. S4E). A negative correlation between *miR-192-5p/3p* and *CALN1/MGAT4C/ADCYAP1R1* was observed (Fig. 4D and Supplementary Fig. S4F). Since *CALN1* is a human brain-specific gene, whereas *MGAT4C* is normally expressed in mesodermal tissues like the kidney and *ADCYAP1R1* encodes the receptor for peptide signals in response to stress in the brain[45–47], we reasoned that *miR-192* may promote HD via repressing genes normally expressed in the other two germ layer-derived lineages.

Given that both microarray and qPCR analysis detected a significant downregulation of *MGAT4C*, we further examined the molecular effects of *miR-192* on this gene. Luciferase reporter assays showed that *miR-192-5p* and *miR-192-3p* mimics could suppress the *MGAT4C* 3′UTR (Fig. 4E). In addition to this, we also examined the expression change of *MGAT4C* upon disturbance of *miR-192* expression. When H1 cells were transfected with *miR-192-5p* and *miR-192-3p* mimics from HD 0–6, *MGAT4C* transcript levels were moderately decreased (Fig. 4F). However, when *miR-192-5p* and *miR-192-3p* inhibitors were transfected from HD 6–12, we observed an upregulation of the *MGAT4C* transcript levels at HD 12 (Fig. 4G). Similar results were obtained in iBC 1.2 cells (Supplementary Fig. S4G,H).

Next, we determined the regulatory effects of *miR-192* on lineage specification. We examined the expression changes of lineage-specific genes while modulating *miR-192* expression during HD. When H1 cells were trans-fected with *miR-192-5p/3p* mimics from HD 0–6, *GATA4* and *GATA6*—which are expressed during definitive endoderm differentiation[48]—were elevated with the overexpression of *miR-192-5p/3p* (Fig. 5A). Conversely, the intermediate mesoderm marker *PAX2* and the MM marker *HOXD11* were significantly increased upon inhibition of *miR-192-5p/3p* from HD 6–12 (Fig. 5B and Supplementary Fig. S5), supporting that *miR-192-5p/3p* repress mesodermal differentiation.

To demonstrate the effects of *miR-192* on HD, we also determined the functional characteristics of HD 12 cells when modulating the expression of *miR-192*. Indocyanine green (ICG) is a non-toxic organic anion that can be eliminated exclusively by mature hepatocytes[49]. Therefore, uptake of ICG is used to measure the hepatic functions. When H1 cells were transfected with *miR-192-5p* mimics, ICG analysis showed that the percentage of ICG-positive cells at HD12 was increased substantially (Fig. 5C,D).

Taken together, we concluded that *miR-192* is likely a key miRNA that promotes endodermal hepatic differentiation while inhibiting the formation of mesoderm.

Effects of miR-372-3p on the expression of KD-specific genes.   To assess the regulatory roles of *miR-372-3p* in KD, we first performed qPCR analysis to confirm the downregulation of *miR-372-3p* (Fig. 6A and Supplementary Fig. S6A). A negative correlation between *miR-372-3p* and *PKD1/PKD2* was identified (Fig. 6B and Supplementary Fig. S6B). Next, we performed luciferase reporter assays and found that the 3′UTRs of *PKD1/PKD2* were significantly inhibited by *miR-372-3p* (Fig. 6C). Since miRNAs can lead to mRNA decay or inhibit translation, we further measured the protein levels of PC1/PC2. Knockdown of *miR-372-3p* during early nephron differentiation (KD 0–6) resulted in increased PC1 and PC2 protein levels (Fig. 6D and Supplementary Fig. S6C). Conversely, overexpression of *miR-372-3p* at KD 8–14 resulted in decreased PC1 and PC2 protein levels (Fig. 6E and Supplementary Fig. S6D).

To demonstrate the effects of *miR-372-3p* on KD, we also determined the efficiency of KD when modulating the expression of *miR-372-3p*. Briefly, we measured the expression of cellular markers that are important for kidney development including SIX2 and HOXD11. IFC analysis showed that the percentage of cells expressed the early metanephric marker HOXD11 at KD 12 was decreased substantially when H1 cells were transfected with *miR-372-3p* mimics (Fig. 6F). Similarly, the percentage of cells that eventually expressed both HOXD11 and the nephric marker SIX2 at KD 18 was decreased upon the transfection (Fig. 6F), indicating that *miR-372-3p* inhibits the formation of nephron progenitors.

Taken together, these experiments validated that *miR-372-3p* is a key miRNA. According to intra-lineage analysis (Supplementary Table S1), *miR-372-3p* expression was not changed significantly; the change from KD 0 to KD 3 was ranked 99th among 177 small RNAs. As a result, it was difficult to identify *miR-372-3p* as a key miRNA by intra-lineage analysis. In contrast, inter-lineage analysis combined with bioinformatics analysis allowed us to identify *miR-372-3p* as a key miRNA, supporting the use of inter-lineage analysis to discover novel key miRNAs.

Discussion

MiRNAs have been found to act as critical regulators in lineage specification. Identification of key miRNAs regulating hPSCs differentiation has become a leading research topic. Intra-lineage analysis, which solely considers miRNAs with large fold-changes, has frequently been used to identify key miRNAs. More recently, inter-lineage analysis has been used to identify key miRNAs with lineage specificity. However, investigation of
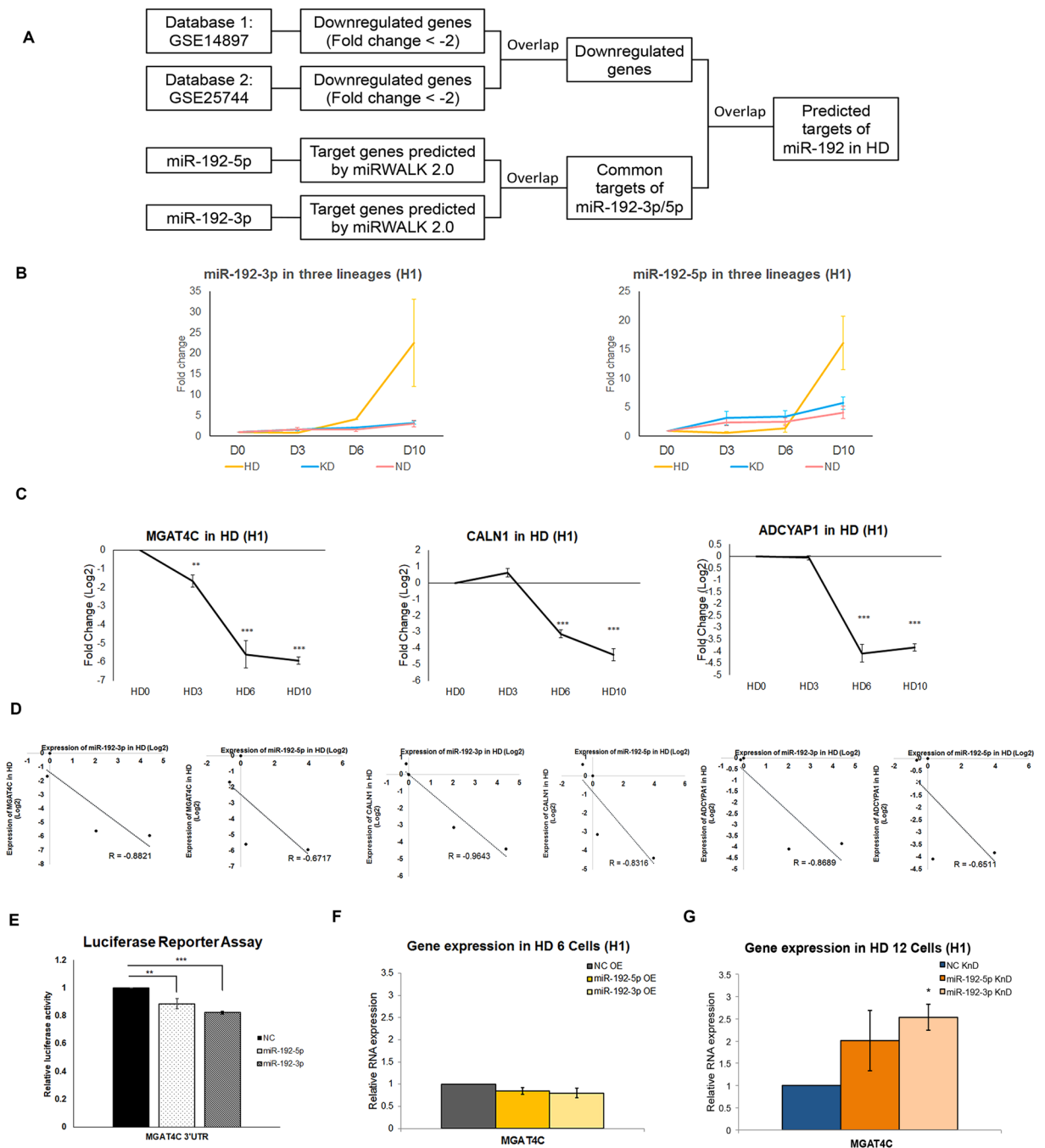
Figure 4. Identification of target genes of key miRNAs during hepatocyte differentiation. (**A**) Strategy for predicting common targets of *miR-192-3p* and *miR-192-5p* during HD. (**B**) TaqMan qPCR analysis confirming that *miR-192-3p/5p* were specifically upregulated during HD (n = 3 independent cultures for each time-point). (**C**) qPCR results showing the expression tendencies of common targets of *miR-192-3p/5p* during HD (n = 3 independent cultures for each time-point). (**D**) Correlation plot revealing reverse-correlations between *miR-192-3p/5p* and *MGAT4C/CALN1/ADCYAP1*, respectively. (**E**) Luciferase reporter assay confirming that *miR-192-3p/5p* could inhibit the 3'UTR of *MGAT4C* (n = 3 independent cultures for each group). (**F**) qPCR results showing the expression of ectodermal marker *MGAT4C* in HD 6 cells upon transfection of *miR-192-3p/5p* mimics (n = 3 independent cultures for each group). (**G**) qPCR results showing the expression of *MGATC4C* in HD 12 cells upon transfection of *miR-192-3p/5p* inhibitors (n = 3 independent cultures for each group). In (**C**), data are presented as the means ± SD. *P < 0.05, **P < 0.01, ***P < 0.001 for statistical comparisons between day 0 and other time-points (ANOVA plus Bonferroni's post hoc test). In (**E–G**), data are presented as the means ± SD. *P < 0.05, **P < 0.01, ***P < 0.001 for statistical comparisons between control groups and experimental groups (ANOVA plus Bonferroni's post hoc test). OE: overexpression; KnD: Knockdown; KD: nephron progenitor differentiation; NC: non-targeting control.
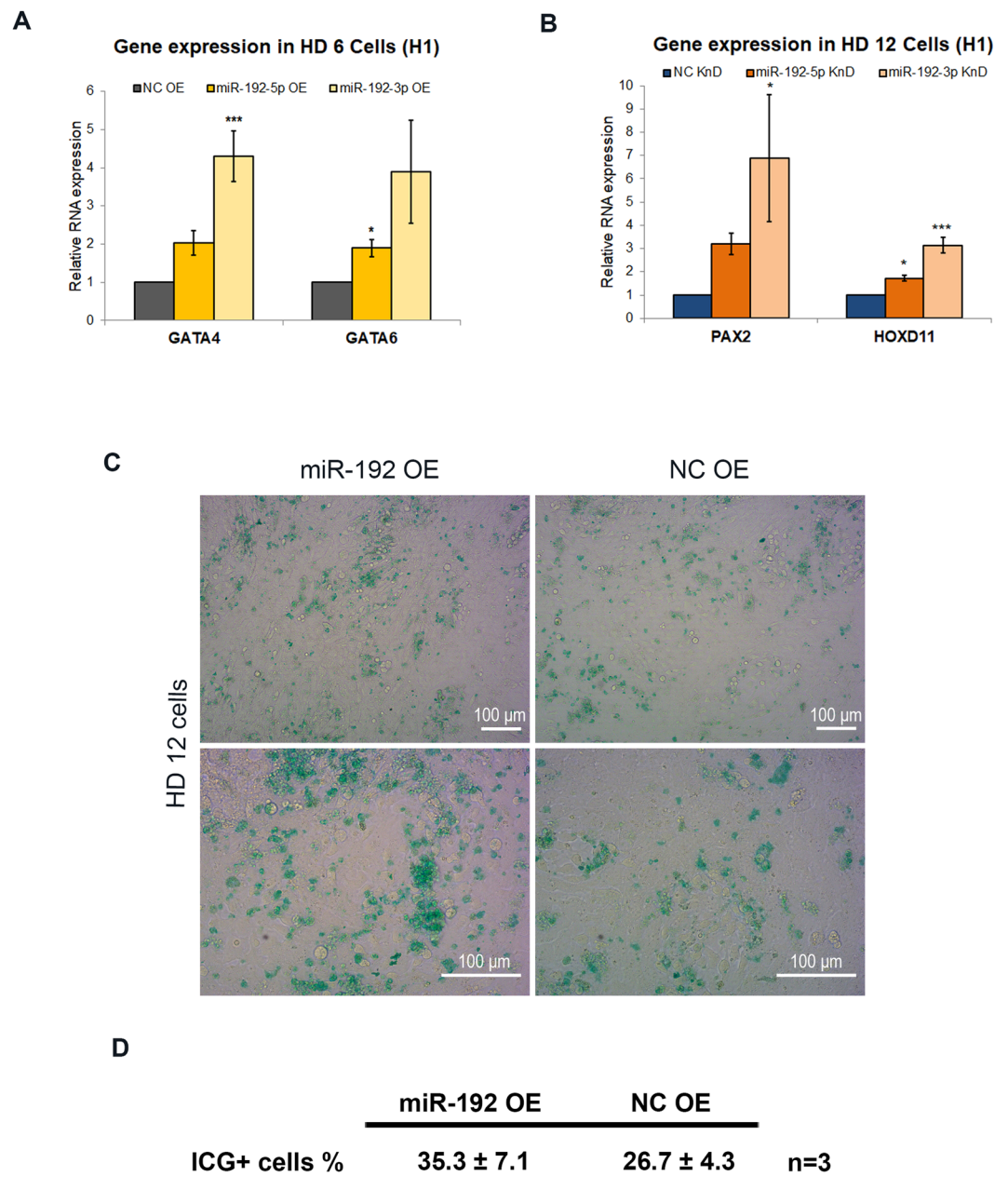
Figure 5. Regulation of lineage-specific gene expression by lineage-specific miRNAs during hepatocyte differentiation. (**A**) qPCR results showing the expression of definitive endoderm markers *GATA4* and *GATA6* in HD 6 cells with transfection of *miR-192-3p/5p* mimics (n = 3 independent cultures for each group). (**B**) qPCR results showing the expression of intermediate mesoderm marker *PAX2* and metanephric mesenchyme marker *HOXD11* in HD 12 cells upon transfection of *miR-192-3p/5p* inhibitors (n = 3 independent cultures for each group). (**C**) HD 12 cells were examined by indocyanine green staining (ICG). The cells transfected with *miR-192* mimics (left column) and scramble controls (right column) were compared (Scale bars represent 100 μm). (**D**) Table showing the percentage of ICG⁺ cells at HD 12 with the transfection of *miR-192* mimics and non-targeting controls (data are presented as the means ± SD. n = 3 independent cultures for each group). In (**A**,**B**), data are presented as the means ± SD. *P < 0.05, **P < 0.01, ***P < 0.001 for statistical comparisons between control groups and experimental groups (ANOVA plus Bonferroni's post hoc test). OE: overexpression; KnD: Knockdown; NC: non-targeting control.

human-specific key miRNAs at the earliest stages of development has still been hindered by a lack of human samples; hPSC-derived lineages have been developed to overcome this challenge. Herein, we applied inter-lineage analysis to hPSC-derived lineages to identify key miRNAs with lineage specificity.

To identify key miRNAs by inter-lineage analysis, we built a miRNA atlas for human *in vitro* early cell lineage specification. This atlas depicts the dynamic expression of miRNAs during the three germ layer formation and lineage differentiation. With this atlas, we can easily identify lineage-specific miRNA in a readily-visualized manner.
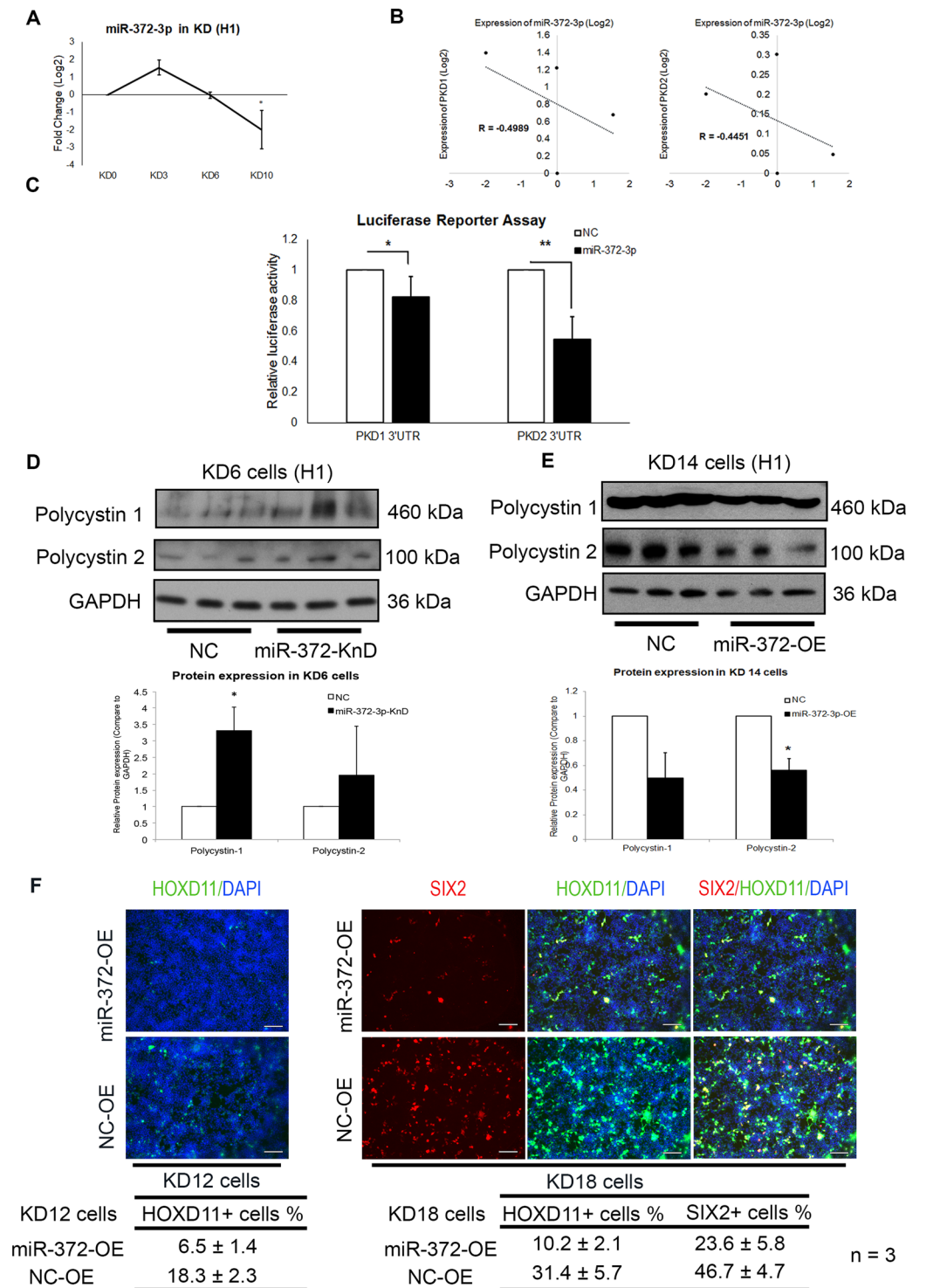
Figure 6. Regulation of *PKD1/PKD2* expression by lineage-specific miRNAs during nephron progenitor differentiation. (**A**) Taqman qPCR analysis showing the expression tendency of *miR-372-3p* during KD (n = 3 independent cultures for each time-point). (**B**) Correlation plot revealing reverse-correlations between *miR-372-3p* and *PKD1/PKD2*, respectively. (**C**) Luciferase reporter assay results confirming that *miR-372-3p* could inhibit the 3′UTR of both *PKD1* and *PKD2* (n = 4 independent cultures for each group). (**D,E**) Western blot results (upper panel) showing expression of Polycystin 1 and Polycystin 2 in KD 6 cells (**D**) with transfection of *miR-372-3p* mimics or non-targeting controls, and expression of Polycystin 1 and Polycystin 2 in KD 14 cells (**E**) upon transfection of *miR-372-3p* inhibitors or non-targeting controls. Normalized protein expressions of Polycystin 1 and Polycystin 2 are shown in lower panel (**D,E**). The expression of NC group of each experiment is set as 1. n = 3 independent cultures for each group. (**F**) KD 12 cells (left column) and KD 18 cells (right column) were examined by immunofluorescent staining (IFC) analysis. The cells transfected with *miR-372-3p* mimics (upper row) and scramble controls (lower row) were compared (Scale bars represent 200 μm). Table showing the

percentage of HOXD11[+] cells at KD 12, SIX2[+] cells and HOXD11[+] cells at KD 18, with the transfection of *miR-372-3p* mimics and non-targeting controls (data are presented as the means ± SD. n = 3 independent cultures for each group). In (**A**), data are presented as the means ± SD. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ for statistical comparisons between day 0 and other time-points (Paired two-tailed *t*-test). In (**C**–**E**), data are presented as the means ± SD. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ for statistical comparisons between control groups and experimental groups (Paired two-tailed *t*-test). OE: overexpression; KnD: Knockdown; KD: nephron progenitor differentiation, NC: non-targeting control.

In this study, we identified *miR-192* and *miR-372-3p* as key miRNAs regulating HD and KD, respectively. To further validate *miR-192* and *miR-372-3p* as key miRNAs, we performed comprehensive experiments to validate their regulatory functions in HD and KD. First, we confirmed that *miR-192* and *miR-372-3p* directly repressed their downstream targets. Next, we modulated their expression during lineage differentiation. The disruptions skewed lineage formation by affecting the expression of lineage markers. Particularly, our data showed that the overexpression of *miR-192* increased the expressions of endodermal markers *GATA4/6*, while knockdown of *miR-192* increased the expression of mesodermal marker *PAX2*. Given that endoderm and mesoderm share the same progenitors, perhaps the upregulation of *miR-192* potentially affects the segregation of endodermal and mesodermal cell fate, leading to an increase in the formation of hepatic endoderm at the expense of mesodermal formation[50]. Finally, we performed functional characterization of *miR-192* and *miR-372-3p* in lineage differentiation, establishing that *miR-192* and *miR-372-3p* affect HD and KD, respectively. These validations prove that *miR-192* and *miR-372-3p* are truly key miRNAs.

In addition to revealing lineage-specific miRNAs, our miRNA atlas can also be used as a roadmap for lineage distinction, as it reveals lineage-specific patterns of miRNA expression. In HD, for example (Fig. 2A), four clusters represented small RNAs highly expressed at HD 0, 3, 6 and 10, respectively. Another cluster remained high from HD 6 to 10. Similar to HD, KD was also characterized by stage-specific clusters (Fig. 2B). In ND, however, the pattern was remarkably different (Fig. 2C). We observed two major clusters, the expression levels of which were dramatically changed as early as ND 3. Meanwhile, this atlas clearly identified many small RNAs that changed specifically. For instance, the miRNAs that increased from ND 3 to ND 10 (purple dendrogram) had low expression in the other two lineages. These lineage-specific clusters can be used as markers to identify lineages. Notably, since the hiPSCs were induced to generate HD 0 cells, KD 0 cells and ND 0 cells with different pre-induction methods (Supplementary data), the expression patterns of miRNAs at day 0 were variable in different lineages (Fig. 3).

Table 1 shows that the top 10 miRNAs in ND were much less likely to have known functions in neural development compared to those in HD and KD. This might be due to the increasing complexity of development from small mammals to primates, especially in the neuronal system[51]. More miRNAs are involved in ND of primates during evolution, whereas the majority of developmental studies have been performed in small mammals[52], leading to many neuron-specific primate miRNAs that have remained functionally undescribed, requiring specific studies with primate samples[53].

In addition to human developmental processes, our miRNA atlas can be applied to the study of complex congenital disorders. Mutations in the *PKD1/PKD2* genes or abnormal expression of PC1/PC2 are known to cause the formation of renal cysts[54–56]. We newly identified the regulation of *miR-372-3p* on *PKD1/PKD2*, indicating the possible association of *miR-372-3p* with polycystic kidney disease for the first time. Moreover, the key miRNAs identified from our atlas are also potential markers and therapeutic targets for various diseases, as treatments in which the miRNA activity is inhibited are in various stages of preclinical or clinical trials[57,58]. There was a differential change of *miR-372-3p* expression between hiPSC- and hESC-derived KD 3 cells that might be due to known differences in microRNA expression between these two hPSC types[59]. However, after KD 3, *miR-372-3p* expression decreased in both hPSC-derived cells. Together with its inhibitory effects on PC1/PC2 at KD 6, the substantial decline in *miR-372-3p* expression levels after KD 3 was sufficient to increase the expression levels of PC1/PC2, despite the differences at earlier time-points.

Importantly, we selected hiPSCs but not hESCs for the construction of the atlas. Considering the necessity of integrating information from different patients, we attempted to produce a reference dataset based on hiPSCs of healthy individuals that could be used for comparison with any patient-specific hiPSCs[29]. Given that hESCs are a widely used model for experimental subjects[27], we used these for testing to provide fair comparisons between hiPSCs and hESCs and to confirm the reliability of our dataset. Compatible expression patterns of the most upregulated miRNAs in hiPSC- and hESC-derived cells (Supplementary Fig. S2) suggested that differential open chromatin states did not affect the expression patterns of important miRNAs. The hiPSCs were reprogrammed from male fibroblasts, which is the most common model despite the existence of male skin lineage memories. Given that the expression profile of miRNAs is different between women and men in many somatic tissues[60], our miRNA atlas is more applicable for studies of male tissues.

To construct the atlas, we followed three well-established protocols to establish the multi-lineage differentiation systems. After successful inductions of HD, KD, and ND, we measured the differentiation efficiencies using IFC and flow-cytometric analysis to, by which the purities of desired somatic cell types were indicated up to 70%. These results suggested that the data outputted from these derived cells represents a mixture of miRNAs changed from wanted cells and unwanted cells. Therefore, to construct a more reliable atlas, an analysis of pure population will be important and necessary in the future.

A statistical test with non-adjusted *P* values has an increased chance of drawing false conclusions when multiple tests are performed[61]. As a result, a non-adjusted $P < 0.05$ is more likely to appear in our case when testing many outcomes from a single intervention. Herein, we adopted a more stringent criterion for multiple testing, the

FDR-adjusted *P* value[61]. Compared to the non-adjusted *P* value, fewer but more reliable results have been generated by FDR. Using HD as an example, 826 and 170 differentially expressed miRNAs could be identified with a $P < 0.05$ and an FDR < 0.05, respectively. As a tradeoff, an FDR < 0.05 excluded several important candidates, such as *miR-375*; FDR was too conservative in this case. When we ran the analysis with a $P < 0.05$, *miR-375* was the number 1 upregulated miRNA (fold-change = 104, $P = 0.001291$) from HD 0–3, which is consistent with its implicated functions in the endodermal formation[8,11–15]. Therefore, the statistical criterion using the *P* value can be applied to our dataset when more candidate miRNAs need to be identified.

In summary, our inter-lineage analysis generated a list of lineage-specific miRNAs during hiPSC differentiation. This list can serve as a source for the identification of key miRNAs and further investigation of the molecular mechanisms underlying hPSC differentiation. Moreover, the integrated miRNA expression atlas can further serve as a reference for studying human developmental processes and human congenital disease.

Materials and Methods

hiPSC culture and differentiation.
iBC 1.2 is a hiPSC line that has been previously reported. iBC 1.2 was generated from a "normal" individual by a viral infection with genetic integration. H1 is a widely used hESC line. Both iBC 1.2 and H1 are male PSC lines. hiPSCs and hESCs were maintained with mTeSR medium (05870, StemCell Technologies, Vancouver, BC, Canada) on Matrigel matrix (354277, Corning, NY, USA)-coated plates and passaged with dispase (354235, Corning, NY, USA) every 4–6 days. They were induced according to previously established protocols for hepatocytes, nephron progenitors, and neural progenitors. The details of the three lineage induction procedures are described in the Supplemental Methods.

Immunofluorescence Microscopy.
Cells were fixed in 4% formaldehyde for 20 minutes and permeabilized with 0.1% Triton X-100 for 45 min, after which the cells were blocked with 1% BSA and 4% goat serum in PBS for 45 min. The cells were then incubated with primary antibodies at 4 °C overnight and with secondary antibodies at room temperature for 1 hour. Finally, nuclei were stained with DAPI. The details regarding the antibodies and dilution ratio are listed in the Supplemental Information. Cells were imaged with a Nikon Ti-E Live-Cell Imaging System.

RNA extraction and real-time quantitative PCR.
Total RNA, including miRNAs, were extracted from cells using the RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (AM1975, Thermo Fisher Scientific). RNA quantitation was performed using a NanoDrop 2000 Spectrophotometer (Thermo Scientific). RNA quality was examined using an Agilent 2100 Bioanalyzer. For cell samples prepared for microRNA microarray detection, an RNA integrity number larger than 9.0 was accepted. For gene expression analysis, cDNA was synthesized with PrimeScirpt™ Reverse Transcriptase (RR036A, Takara). SYBR Green master mix was used for PCR in a QuantStudio™ 7 Flex System (Life Technologies). *GAPDH* and 18 S RNA were used as internal controls. For microRNA expression analysis, RNA was reverse-transcribed with a TaqMan MicroRNA Reverse Transcription Kit, and amplification was performed using TaqMan probes with Taqman Universal Master Mix II, without UNG (Thermo Fisher Scientific). *RNU6B* was used as an endogenous control for normalization of miRNAs. The samples were plotted relative to control samples (HD0, KD0, and ND0), and the standard deviation of at least three measurements was calculated.

MicroRNA microarray analysis and data processing.
MiRNA microarray expression data were acquired on an Affymetrix miRNA 4.0 platform (2,578 mature miRNAs, 2,025 stem-loop miRNAs, and 1,996 other small RNAs), using 500 ng total RNA per sample. Microarray data were deposited in the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GEO: GSE97952. miRNA expression data were then analyzed in the Partek® Genomics Suite® following the workflow for miRNA microarray analysis to detect differentially expressed miRNAs between samples. MicroRNAs of interest were filtered using a cut-off fold-change value of ≥ 2 or ≤ -2 with an FDR < 0.05. Both hierarchical clustering results and a heat map were generated with the Partek® Genomics Suite® platform. miRNA target prediction algorithms were performed with TargetScanHuman 6.2 (http://www.targetscan.org/vert_61/) and miRWALK (http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/). The accession projects supporting the prediction of downstream targets of *miR-192* are GSE14897 and GSE25744.

Luciferase reporter assay.
50,000 HEK 293 cells were plated in wells of 24-well plates 24 h before transfection. Reporter plasmids with *MGAT4C*, *PKD1*, *PKD2* 3′UTR were purchased from GeneCopoeia. 10 pmol synthetic miRNA mimics of specific miRNAs, including *miR-192-3p*, *miR-192-5p*, and *miR-372-3p* or non-targeting control (GenePharma, Shanghai, China), and 200 ng reporter plasmids were co-transfected with 1 μl Lipofectamine 2000 according to the manufacturer's instructions. The medium was changed to DMEM supplemented with 10% FBS for 18 h post-transfection. At 48–72 h post-transfection, media were collected for analysis. Luciferase activities were determined with a Secrete-Pair™ Dual Luminescence Assay Kit (SPDA-D010, GeneCopoeia, Rockville, MD, USA).

Knockdown and overexpression of miRNAs.
Mimics of non-targeting control, *miR-192-3p*, *miR-192-5p*, and *miR-372-3p* miRNAs were purchased from GenePharma. Inhibitors of negative control, *miR-192-3p*, *miR-192-5p*, and *miR-372-3p* were purchased from Thermo Fisher Scientific. Cells cultured in 12-well plates were transfected with either mimics or inhibitors. First, 12.5 pmol mimics or inhibitors were diluted in 80 μl Opti-MEM. Second, 3.5 μl Lipofectamine RNAiMAX reagent was diluted in 80 μl Opti-MEM. Third, the Opti-MEM containing miRNA mimics and the Opti-MEM containing Lipofectamine® RNAiMAX Reagent were mixed at a ratio of 1:1. The mixture was incubated at room temperature for 20 minutes and added to the

differentiated cells. The medium was changed the next day. Transfection was performed every two days three times. Then, the transfected cells were collected for further analysis.

Western blot analysis.   Cells were lysed in RIPA buffer with 2% proteinase inhibitor. 30 μg of total proteins were resolved on 4–20% precast polyacrylamide gels (4561096, Bio-Rad, Hercules, CA, USA), and transferred to PVDF membranes (ISEQ. 00010, Merck Millipore, Darmstadt, Germany). After transfer, the membranes were blocked with 5% Blotting-grade Blocker (1706404, Bio-Rad, Hercules, CA, USA) in PBS-T for 1 hour, and incubated with specific primary antibodies in the blocking solutions overnight at 4 °C. The membranes were then washed with PBST three times and incubated with secondary antibodies for 1 hour at room temperature. Then, the antibodies were detected using an ECL HRP substrate system (K-12045-D20, Advansta, Menlo Park, CA, USA). The protein band intensities were quantified using Image Lab software. Information regarding the antibodies is shown in the Supplemental Information.

Statistical analysis.   For microarray analysis, one-way ANOVA testing was used to determine which small RNAs had significant differences in expression between time-points. Subsequent pairwise comparisons between successive time-points were used to identify when small RNAs demonstrated significant differential expression (post-hoc testing, fold-change ≥ 2 or ≤ -2, false discovery rate < 0.05). For qPCR analysis, luciferase reporter assays and Western blotting analysis, statistical significance was determined using one-way ANOVA followed by Bonferroni's post-hoc testing. A paired two-tailed Student's t-test was performed when only two groups were compared. A single asterisk indicates $P < 0.05$, a double asterisk indicates $P < 0.01$, and a triple asterisk indicates $P < 0.001$. In all Figures, the mean $\pm$ SD is plotted.

Data availability statement.   All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

References

1. Leonardo, T. R., Schultheisz, H. L., Loring, J. F. & Laurent, L. C. The functions of microRNAs in pluripotency and reprogramming. *Nature cell biology* **14**, 1114–1121, https://doi.org/10.1038/ncb2613 (2012).
2. Boland, M. J., Nazor, K. L. & Loring, J. F. Epigenetic regulation of pluripotency and differentiation. *Circulation research* **115**, 311–324, https://doi.org/10.1161/CIRCRESAHA.115.301517 (2014).
3. Thiagarajan, R. D., Morey, R. & Laurent, L. C. The epigenome in pluripotency and differentiation. *Epigenomics* **6**, 121–137, https://doi.org/10.2217/epi.13.80 (2014).
4. Miyamoto, T., Furusawa, C. & Kaneko, K. Pluripotency, Differentiation, and Reprogramming: A Gene Expression Dynamics Model with Epigenetic Feedback Regulation. *PLoS computational biology* **11**, e1004476, https://doi.org/10.1371/journal.pcbi.1004476 (2015).
5. Jung, Y., Bauer, G. & Nolta, J. A. Concise review: Induced pluripotent stem cell-derived mesenchymal stem cells: progress toward safe clinical products. *Stem Cells* **30**, 42–47, https://doi.org/10.1002/stem.727 (2012).
6. Fox, I. J. *et al.* Stem cell therapy. Use of differentiated pluripotent stem cells as replacement therapy for treating disease. *Science* **345**, 1247391, https://doi.org/10.1126/science.1247391 (2014).
7. Avior, Y., Sagi, I. & Benvenisty, N. Pluripotent stem cells in disease modelling and drug discovery. *Nature reviews. Molecular cell biology* **17**, 170–182, https://doi.org/10.1038/nrm.2015.27 (2016).
8. Liao, X. *et al.* Matched miRNA and mRNA signatures from an hESC-based *in vitro* model of pancreatic differentiation reveal novel regulatory interactions. *Journal of cell science* **126**, 3848–3861, https://doi.org/10.1242/jcs.123570 (2013).
9. Ivey, K. N. & Srivastava, D. microRNAs as Developmental Regulators. *Cold Spring Harbor perspectives in biology* **7**, a008144, https://doi.org/10.1101/cshperspect.a008144 (2015).
10. Young, R. A. Control of the embryonic stem cell state. *Cell* **144**, 940–954, https://doi.org/10.1016/j.cell.2011.01.032 (2011).
11. Hinton, A. *et al.* A distinct microRNA signature for definitive endoderm derived from human embryonic stem cells. *Stem cells and development* **19**, 797–807, https://doi.org/10.1089/scd.2009.0224 (2010).
12. Hinton, A. *et al.* sRNA-seq analysis of human embryonic stem cells and definitive endoderm reveals differentially expressed microRNAs and novel IsomiRs with distinct targets. *Stem Cells* **32**, 2360–2372, https://doi.org/10.1002/stem.1739 (2014).
13. Ishikawa, D. *et al.* miRNome Profiling of Purified Endoderm and Mesoderm Differentiated from hESCs Reveals Functions of miR-483-3p and miR-1263 for Cell-Fate Decisions. *Stem cell reports* **9**, 1588–1603, https://doi.org/10.1016/j.stemcr.2017.10.011 (2017).
14. Joglekar, M. V., Joglekar, V. M. & Hardikar, A. A. Expression of islet-specific microRNAs during human pancreatic development. *Gene expression patterns: GEP* **9**, 109–113, https://doi.org/10.1016/j.gep.2008.10.001 (2009).
15. Fogel, G. B., Tallon, T., Wong, A. S., Lopez, A. D. & King, C. C. 93–102 (Springer International Publishing).
16. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic acids research* **42**, D68–73, https://doi.org/10.1093/nar/gkt1181 (2014).
17. Letzen, B. S. *et al.* MicroRNA expression profiling of oligodendrocyte differentiation from human embryonic stem cells. *PloS one* **5**, e10480, https://doi.org/10.1371/journal.pone.0010480 (2010).
18. Petriv, O. I. *et al.* Comprehensive microRNA expression profiling of the hematopoietic hierarchy. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15443–15448, https://doi.org/10.1073/pnas.1009320107 (2010).
19. Garcia, D. M. *et al.* Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nature structural & molecular biology* **18**, 1139–1146, https://doi.org/10.1038/nsmb.2115 (2011).
20. Hsu, S. D. *et al.* miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic acids research* **39**, D163–169, https://doi.org/10.1093/nar/gkq1107 (2011).
21. Wong, N. & Wang, X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic acids research* **43**, D146–152, https://doi.org/10.1093/nar/gku1104 (2015).
22. Suh, M. R. *et al.* Human embryonic stem cells express a unique set of microRNAs. *Dev. Biol.* **270**, 488–498, https://doi.org/10.1016/j.ydbio.2004.02.019 (2004).
23. Londin, E. *et al.* Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1106–1115, https://doi.org/10.1073/pnas.1420955112 (2015).
24. Li, X. & Jin, P. Roles of small regulatory RNAs in determining neuronal identity. *Nature reviews. Neuroscience* **11**, 329–338, https://doi.org/10.1038/nrn2739 (2010).
25. McCall, M. N. *et al.* Toward the human cellular microRNAome. *Genome research* **27**, 1769–1781, https://doi.org/10.1101/gr.222067.117 (2017).

26. Juzenas, S. *et al.* A comprehensive, cell specific microRNA catalogue of human peripheral blood. *Nucleic acids research* **45**, 9290–9301, https://doi.org/10.1093/nar/gkx706 (2017).

27. Zhu, Z. & Huangfu, D. Human pluripotent stem cells: an emerging model in developmental biology. *Development* **140**, 705–717, https://doi.org/10.1242/dev.086165 (2013).

28. Medvedev, S. P., Shevchenko, A. I. & Zakian, S. M. Induced Pluripotent Stem Cells: Problems and Advantages when Applying them in Regenerative Medicine. *Acta naturae* **2**, 18–28 (2010).

29. Singh, V. K., Kalsan, M., Kumar, N., Saini, A. & Chandra, R. Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery. *Frontiers in cell and developmental biology* **3**, 2, https://doi.org/10.3389/fcell.2015.00002 (2015).

30. Chen, Y. F. *et al.* Rapid generation of mature hepatocyte-like cells from human induced pluripotent stem cells by an efficient three-step protocol. *Hepatology* **55**, 1193–1203, https://doi.org/10.1002/hep.24790 (2012).

31. Mohamad, O. *et al.* Vector-free and transgene-free human iPS cells differentiate into functional neurons and enhance functional recovery after ischemic stroke in mice. *PloS one* **8**, e64160, https://doi.org/10.1371/journal.pone.0064160 (2013).

32. Takasato, M. *et al.* Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. *Nature cell biology* **16**, 118–126, https://doi.org/10.1038/ncb2894 (2014).

33. Liu, X. *et al.* Idiopathic Autism: Cellular and Molecular Phenotypes in Pluripotent Stem Cell-Derived Neurons. *Molecular neurobiology* **54**, 4507–4523, https://doi.org/10.1007/s12035-016-9961-8 (2017).

34. Cheung, H. H. *et al.* Telomerase protects werner syndrome lineage-specific stem cells from premature aging. *Stem cell reports* **2**, 534–546, https://doi.org/10.1016/j.stemcr.2014.02.006 (2014).

35. Lagos-Quintana, M. *et al.* Identification of tissue-specific microRNAs from mouse. *Current biology: CB* **12**, 735–739 (2002).

36. Girard, M., Jacquemin, E., Munnich, A., Lyonnet, S. & Henrion-Caude, A. miR-122, a paradigm for the role of microRNAs in the liver. *Journal of hepatology* **48**, 648–656, https://doi.org/10.1016/j.jhep.2008.01.019 (2008).

37. Yin, H. *et al.* Phylogenetic tree-informed microRNAome analysis uncovers conserved and lineage-specific miRNAs in Camellia during floral organ development. *Journal of experimental botany* **67**, 2641–2653, https://doi.org/10.1093/jxb/erw095 (2016).

38. Vinas, J. L. *et al.* miRNA let-7e modulates the Wnt pathway and early nephrogenic markers in mouse embryonic stem cell differentiation. *PLoS One* **8**, e60937, https://doi.org/10.1371/journal.pone.0060937 (2013).

39. Makeyev, E. V., Zhang, J., Carrasco, M. A. & Maniatis, T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Molecular cell* **27**, 435–448, https://doi.org/10.1016/j.molcel.2007.07.015 (2007).

40. Gainullin, V. G., Hopp, K., Ward, C. J., Hommerding, C. J. & Harris, P. C. Polycystin-1 maturation requires polycystin-2 in a dose-dependent manner. *The Journal of clinical investigation* **125**, 607–620, https://doi.org/10.1172/JCI76972 (2015).

41. Patel, V. *et al.* miR-17~92 miRNA cluster promotes kidney cyst growth in polycystic kidney disease. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 10765–10770, https://doi.org/10.1073/pnas.1301693110 (2013).

42. Lakshmipathy, U. *et al.* MicroRNA expression pattern of undifferentiated and differentiated human embryonic stem cells. *Stem cells and development* **16**, 1003–1016, https://doi.org/10.1089/scd.2007.0026 (2007).

43. Sullivan, G. J. *et al.* Generation of functional human hepatic endoderm from human induced pluripotent stem cells. *Hepatology* **51**, 329–335, https://doi.org/10.1002/hep.23335 (2010).

44. Jozefczuk, J., Prigione, A., Chavez, L. & Adjaye, J. Comparative analysis of human embryonic stem cell and induced pluripotent stem cell-derived hepatocyte-like cells reveals current drawbacks and possible strategies for improved differentiation. *Stem cells and development* **20**, 1259–1275, https://doi.org/10.1089/scd.2010.0361 (2011).

45. Wu, Y. Q., Lin, X., Liu, C. M., Jamrich, M. & Shaffer, L. G. Identification of a human brain-specific gene, calneuron 1, a new member of the calmodulin superfamily. *Molecular genetics and metabolism* **72**, 343–350, https://doi.org/10.1006/mgme.2001.3160 (2001).

46. Yang, T. L. *et al.* Runs of homozygosity identify a recessive locus 12q21.31 for human adult height. *The Journal of clinical endocrinology and metabolism* **95**, 3777–3782, https://doi.org/10.1210/jc.2009-1715 (2010).

47. Ressler, K. J. *et al.* Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor. *Nature* **470**, 492–497, https://doi.org/10.1038/nature09856 (2011).

48. Tiyaboonchai, A. *et al.* GATA6 Plays an Important Role in the Induction of Human Definitive Endoderm, Development of the Pancreas, and Functionality of Pancreatic beta Cells. *Stem cell reports* **8**, 589–604, https://doi.org/10.1016/j.stemcr.2016.12.026 (2017).

49. Berk, P. D. & Stremmel, W. Hepatocellular uptake of organic anions. *Progress in liver diseases* **8**, 125–144 (1986).

50. Li, N., Wei, C., Olena, A. F. & Patton, J. G. Regulation of endoderm formation and left-right asymmetry by miR-92 during early zebrafish development. *Development* **138**, 1817–1826, https://doi.org/10.1242/dev.056697 (2011).

51. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nature reviews. Neuroscience* **10**, 724–735, https://doi.org/10.1038/nrn2719 (2009).

52. Berezikov, E. *et al.* Diversity of microRNAs in human and chimpanzee brain. *Nature genetics* **38**, 1375–1377, https://doi.org/10.1038/ng1914 (2006).

53. Lopez, J. P. *et al.* miR-1202 is a primate-specific and brain-enriched microRNA involved in major depression and antidepressant treatment. *Nature medicine* **20**, 764–768, https://doi.org/10.1038/nm.3582 (2014).

54. Harris, P. C. & Torres, V. E. Polycystic kidney disease. *Annual review of medicine* **60**, 321–337, https://doi.org/10.1146/annurev.med.60.101707.125712 (2009).

55. Torres, V. E. & Harris, P. C. Autosomal dominant polycystic kidney disease: the last 3 years. *Kidney international* **76**, 149–168, https://doi.org/10.1038/ki.2009.128 (2009).

56. Chapin, H. C. & Caplan, M. J. The cell biology of polycystic kidney disease. *The Journal of cell biology* **191**, 701–710, https://doi.org/10.1083/jcb.201006173 (2010).

57. Patel, V. & Noureddine, L. MicroRNAs and fibrosis. *Current opinion in nephrology and hypertension* **21**, 410–416, https://doi.org/10.1097/MNH.0b013e328354e559 (2012).

58. Noureddine, L., Hajarnis, S. & Patel, V. MicroRNAs and Polycystic Kidney Disease. *Drug discovery today. Disease models* **10**, e137–e1743, https://doi.org/10.1016/j.ddmod.2013.10.001 (2013).

59. Chin, M. H. *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell stem cell* **5**, 111–123, https://doi.org/10.1016/j.stem.2009.06.008 (2009).

60. Dai, R. & Ahmed, S. A. Sexual dimorphism of miRNA expression: a new perspective in understanding the sex bias of autoimmune diseases. *Therapeutics and clinical risk management* **10**, 151–163, https://doi.org/10.2147/TCRM.S33517 (2014).

61. Verhoeven, K. J. F., Simonsen, K. L. & McIntyre, L. M. Implementing false discovery rate control: increasing your power. *Oikos* **108**, 643–647, https://doi.org/10.1111/j.0030-1299.2005.13727.x (2005).

62. Van Keuren-Jensen, K. R. *et al.* microRNA changes in liver tissue associated with fibrosis progression in patients with hepatitis C. *Liver international: official journal of the International Association for the Study of the Liver* **36**, 334–343, https://doi.org/10.1111/liv.12919 (2016).

63. Chang, J. *et al.* miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1. *RNA Biol* **1**, 106–113 (2004).

64. Esau, C. *et al.* miR-122 regulation of lipid metabolism revealed by *in vivo* antisense targeting. *miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting* **3**, 87–98, https://doi.org/10.1016/j.cmet.2006.01.005 (2006).

65. Xiao, Y. *et al*. Dysregulated miR-124 and miR-200 expression contribute to cholangiocyte proliferation in the cholestatic liver by targeting IL-6/STAT3 signalling. *Journal of hepatology* **62**, 889–896, https://doi.org/10.1016/j.jhep.2014.10.033 (2015).
66. Cheng, L. C., Pastrana, E., Tavazoie, M. & Doetsch, F. miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche. *Nature neuroscience* **12**, 399–408, https://doi.org/10.1038/nn.2294 (2009).
67. Liao, Y., Zhang, M. & Lonnerdal, B. Growth factor TGF-beta induces intestinal epithelial cell (IEC-6) differentiation: miR-146b as a regulatory component in the negative feedback loop. *Growth factor TGF-beta induces intestinal epithelial cell (IEC-6) differentiation: miR-146b as a regulatory component in the negative feedback loop* **8**, 69–78, https://doi.org/10.1007/s12263-012-0297-3 (2013).
68. Chen, L. *et al*. MiR-146b is a regulator of human visceral preadipocyte proliferation and differentiation and its expression is altered in human obesity. *Mol. Cell. Endocrinol.* **393**, 65–74, https://doi.org/10.1016/j.mce.2014.05.022 (2014).
69. Comer, B. S. *et al*. MicroRNA-146a and microRNA-146b expression and anti-inflammatory function in human airway smooth muscle. *Am. J. Physiol. Lung Cell Mol. Physiol.* **307**, L727–734, https://doi.org/10.1152/ajplung.00174.2014 (2014).
70. Starkey Lewis, P. J. *et al*. Circulating microRNAs as potential markers of human drug-induced liver injury. *Hepatology* **54**, 1767–1776, https://doi.org/10.1002/hep.24538 (2011).
71. Hino, K. *et al*. Inducible expression of microRNA-194 is regulated by HNF-1alpha during intestinal epithelial cell differentiation. *RNA* **14**, 1433–1442, https://doi.org/10.1261/rna.810208 (2008).
72. Hromadnikova, I. *et al*. Expression profile of C19MC microRNAs in placental tissue in pregnancy-related complications. *DNA and cell biology* **34**, 437–457, https://doi.org/10.1089/dna.2014.2687 (2015).
73. Stadler, B. *et al*. Characterization of microRNAs involved in embryonic stem cell states. *Stem cells and development* **19**, 935–950, https://doi.org/10.1089/scd.2009.0426 (2010).
74. Wang, N. *et al*. Urinary microRNA-10a and microRNA-30d serve as novel, sensitive and specific biomarkers for kidney injury. *PloS one* **7**, e51140, https://doi.org/10.1371/journal.pone.0051140 (2012).
75. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83–86, https://doi.org/10.1126/science.1091903 (2004).
76. Yoder, M. C. Differentiation of pluripotent stem cells into endothelial cells. *Curr. Opin. Hematol.* **22**, 252–257, https://doi.org/10.1097/MOH.0000000000000140 (2015).
77. Shi, K. *et al*. MicroRNA-214 suppresses osteogenic differentiation of C2C12 myoblast cells by targeting Osterix. *Bone* **55**, 487–494, https://doi.org/10.1016/j.bone.2013.04.002 (2013).
78. Rosa, A., Spagnoli, F. M. & Brivanlou, A. H. The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. *Dev. Cell* **16**, 517–527, https://doi.org/10.1016/j.devcel.2009.02.007 (2009).
79. Beveridge, N. J., Tooney, P. A., Carroll, A. P., Tran, N. & Cairns, M. J. Down-regulation of miR-17 family expression in response to retinoic acid induced neuronal differentiation. *Cellular signalling* **21**, 1837–1845, https://doi.org/10.1016/j.cellsig.2009.07.019 (2009).
80. Saba, R. *et al*. Dopamine-regulated microRNA MiR-181a controls GluA2 surface expression in hippocampal neurons. *Molecular and cellular biology* **32**, 619–632, https://doi.org/10.1128/MCB.05896-11 (2012).
81. Abdelmohsen, K. *et al*. miR-375 inhibits differentiation of neurites by lowering HuD levels. *Molecular and cellular biology* **30**, 4197–4210, https://doi.org/10.1128/MCB.00316-10 (2010).
82. Bhinge, A. *et al*. MiR-375 is Essential for Human Spinal Motor Neuron Development and May Be Involved in Motor Neuron Degeneration. *Stem Cells* **34**, 124–134, https://doi.org/10.1002/stem.2233 (2016).

Acknowledgements

Author Contributions

L.L., K.K.M., S.G. and W.Y.C. conceived the project. L.L., K.K.M., S.G. and H.H.C. designed the experiments. L.L., K.K.M. and S.G. conducted the experiments. L.L. analyzed the data and wrote the manuscript. W.Y.C. and H.H.C. edited the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-27719-0.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.