



## Article

## MR-seq: measuring a single cell's transcriptome repeatedly by RNA-seq

Lu Yang<sup>a,1</sup>, Zhaochun Ma<sup>b,1</sup>, Chen Cao<sup>c,1</sup>, Yuhao Zhang<sup>a,1</sup>, Xinglong Wu<sup>c,1</sup>, Raymond Lee<sup>b</sup>, Boqiang Hu<sup>a</sup>, Lu Wen<sup>a</sup>, Hao Ge<sup>a,c,d</sup>, Yanyi Huang<sup>a,c,e,g,\*</sup>, Kaiqin Lao<sup>b,\*</sup>, Fuchou Tang<sup>a,c,f,g,\*</sup>

<sup>a</sup> Biodynamic Optical Imaging Center, School of Life Sciences, Peking University, Beijing 100871, China

<sup>b</sup> Thermo Fisher Scientific, South San Francisco, CA 94080, USA

<sup>c</sup> Peking-Tsinghua Center for Life Sciences, Beijing 100084, China

<sup>d</sup> Beijing International Center for Mathematical Research (BICMR), Peking University, Beijing 100871, China

<sup>e</sup> College of Engineering, Peking University, Beijing 100871, China

<sup>f</sup> China Center for Molecular and Translational Medicine (CMTM), Beijing 100101, China

<sup>g</sup> Beijing Advanced Innovation Center for Genomics (ICG), Peking University, Beijing 100871, China

## ARTICLE INFO

## Article history:

Received 4 January 2017

Received in revised form 20 January 2017

Accepted 20 January 2017

Available online 24 February 2017

## Keywords:

Single-cell RNA-seq

Differentially expressed genes

Intra-embryonic heterogeneity

Mouse early preimplantation embryos

## ABSTRACT

We described a novel single-cell RNA-seq technique called MR-seq (measure a single-cell transcriptome repeatedly), which permits statistically assessing the technical variation and identifying the differentially expressed genes between just two single cells by measuring each single cell twice. We demonstrated that MR-seq gave sensitivity and reproducibility similar to the standard single-cell RNA-seq and increased the positive predicate value. Application of MR-seq to early mouse embryos identified hundreds of candidate intra-embryonic heterogeneous genes among mouse 2-, 4- and 8-cell stage embryos. MR-seq should be useful for detecting differentially expressed genes among a small number of cells.

© 2017 Science China Press. Published by Elsevier B.V. and Science China Press. All rights reserved. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The single cell RNA-seq techniques developed rapidly in recent years [1–5]. They have been applied for analyzing precious and rare cell types [6–12] such as human early embryos, and for dissection of gene expression heterogeneity within a population of cells [13–20]. The technical noise of the single-cell RNA-seq techniques is much larger than the bulk RNA-seq, and separating the technical variation from the biological variation has been a research focus [21–24]. Particularly, when there are only several cells different from each other, e.g. blastomeres in early mammalian embryos, it is not possible to statistically identify the differentially expressed genes (DEGs) due to no replicates and high technical noise in the single cell RNA-seq dataset.

Here, we presented a novel method named MR-seq (measure a single-cell transcriptome repeatedly). This technique, which splits the lysate of a single cell into 2 aliquots and perform single-cell RNA-seq analysis separately, allows for measuring each single

cell's transcriptome twice, then statistically assessing the technical variation and identifying the DEGs between just two single cells.

## 2. Materials and methods

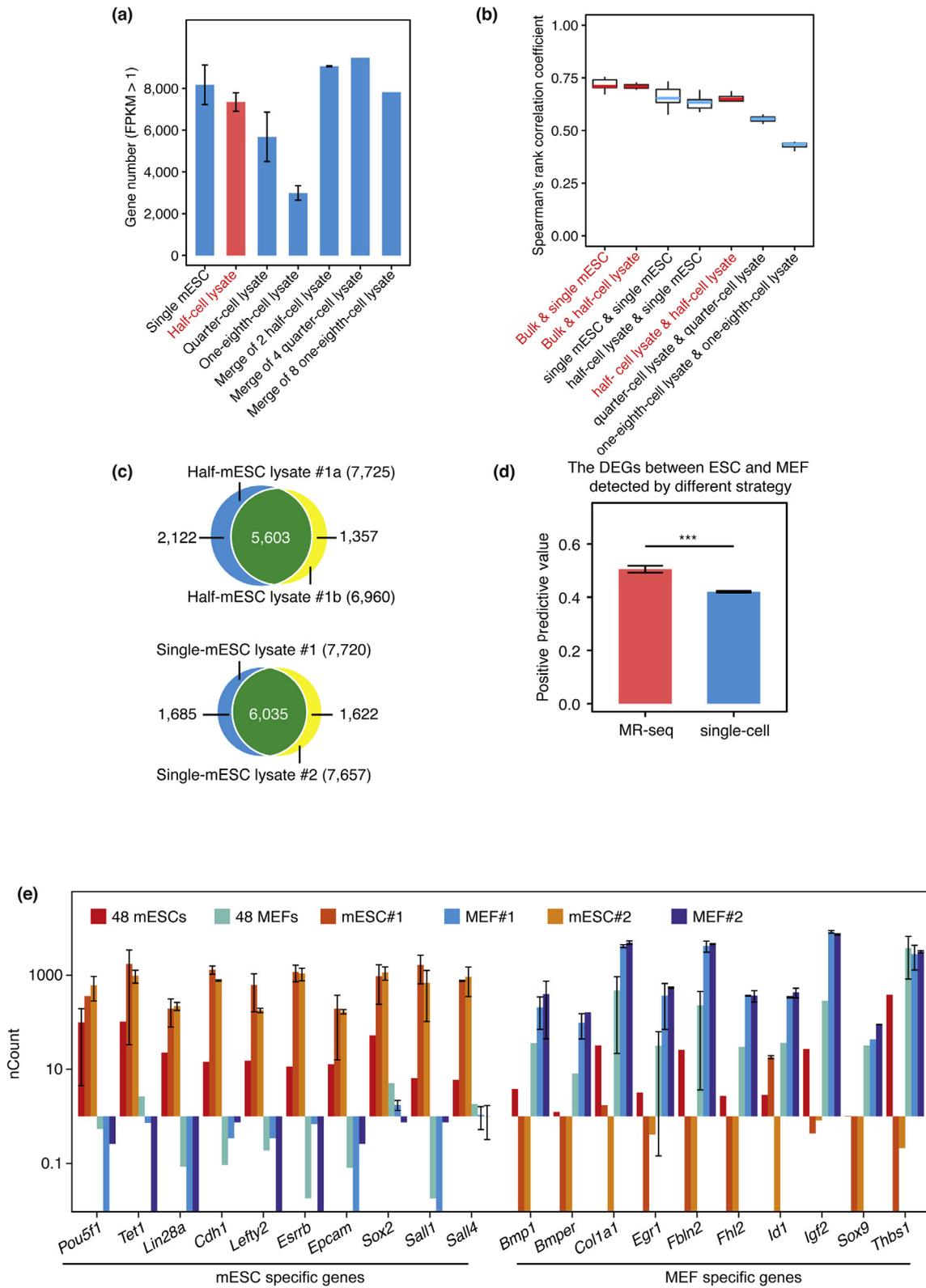
### 2.1. Isolation of zygotes; 2-, 4- and 8-cell embryos; inner cell mass (ICM) and trophectoderm (TE)

Six- to eight-week old ICR female mice were injected with 7.5 IU human chorionic gonadotropin (hCG) 46–48 h after injection with 7.5 IU pregnant mare serum gonadotrophin (PMSG). The 2-, 4- and 8-cell stage embryos were collected from the oviduct of mated mice at 48, 56 and 70 h post-hCG injection, respectively. The embryos were treated with Acid Tyrode's solution to remove the zona pellucida, transferred to trypsin-EDTA, and incubated at 37 °C for 5 min. Then, the embryos were gently washed with PBS-BSA buffer to dissociate the embryos into individual blastomeres. The ICM and TE were isolated from mouse E3.5 blastocysts [8]. The trophectoderm was isolated from the blastocysts with zona pellucida removed under the microscope by mechanically cutting using a 30 G needle. The exposed ICM was gently scraped from the trophectoderm. These two cell types were then treated with accutase or 0.005% trypsin at 37 °C to dissociate single cells.

\* Corresponding authors.

E-mail addresses: [yanyi@pku.edu.cn](mailto:yanyi@pku.edu.cn) (Y. Huang), [kai.lao@thermofisher.com](mailto:kai.lao@thermofisher.com) (K. Lao), [tangfuchou@pku.edu.cn](mailto:tangfuchou@pku.edu.cn) (F. Tang).

<sup>1</sup> These authors contributed equally to this work.



**Fig. 1.** MR-seq performs well by measure a single cell twice. **a** The average gene number (FPKM > 1) of single, half, one-quarter, and one-eighth mESC; merge of 2 half-mESC lysates, merge of 4 one-quarter-mESC lysate, and merge of 8 one-eighth-mESC lysates. The error bars showed the SD of gene number. The sequencing depths of different libraries were adjusted to  $7 \times 10^6$  reads. **b** Boxplot shows the correlations (spearman's rank correlation coefficient,  $r_s$ ) among gene expression (FPKM > 1) of bulk, single, half, quarter and eighth mESC. The blue and red horizontal lines indicate the median. **c** Venn diagrams show the gene number (FPKM > 1) of half-mESC lysates and single mESCs. **d** The positive predictive value of detected DEGs between mESCs and MEFs using two and one measurements (MR-seq and the single-cell) in each group. **e** The mESC and MEF specific genes' expression. The 48 mESCs and 48 MEFs data come from Islam et al.

2.2. Cell lines

MEFs were derived from ICR strain E13.5 embryos. Mouse embryonic stem cells (mESCs) were cultured on 0.1% gelatinized tissue culture plates and passaged every two days.

2.3. Preparation of single-cell cDNAs for MR-seq

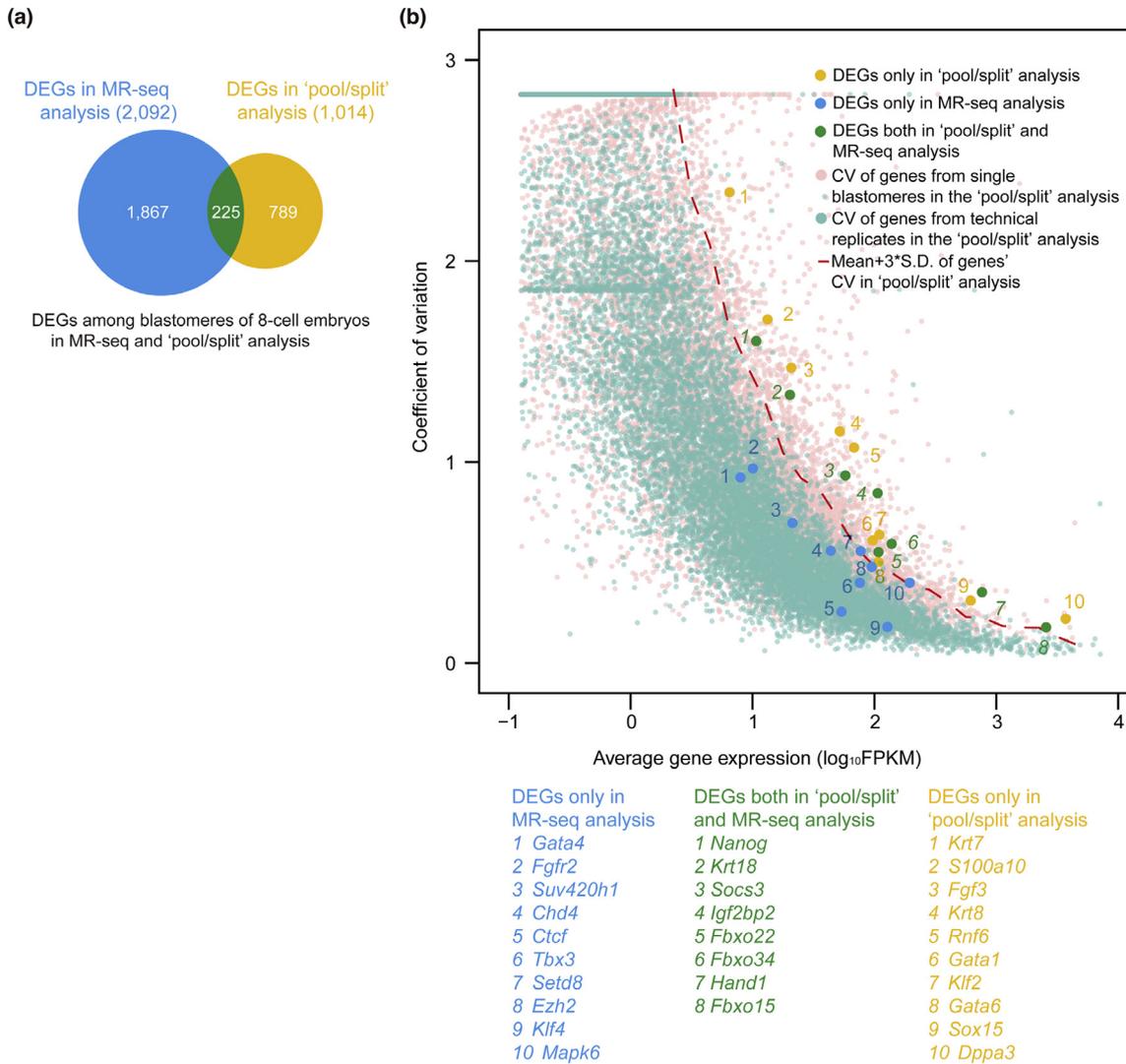
In this work, we measured 132 samples' transcriptome (Table S1, online), and sequenced the libraries on Illumina and Ion platforms. The MR-seq protocol was developed based on the single-cell RNA-seq technique we previously developed [2]. Briefly, single cells were lysed with a doubled volume (9.1  $\mu$ L) of lysis buffer, and the sample was incubated at 70 °C for 90 s. Then, a mixture of the reverse transcription enzyme solution and cell lysate was divided into two aliquots. After the reverse transcription, primer remove, A-tailing, and second strand synthesis, one round of no more than 24 cycles of PCR amplification was used to amplify the cDNAs for sequencing library preparation.

2.4. Construction of the cDNA libraries for Illumina sequencing

The 5–100 ng amplified cDNAs were frgmented to approximately 200 bp fragments using Covaris S2 and purified with DNA Clean & Concentrator™-5. Then, NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (NEB, Catalog #E7370L) was used to construct the cDNA library with half reagents. The cDNA libraries were sequenced on Illumina HiSeq 2000 and 2500 instruments with 100 bp reads and paired-end parameters.

2.5. Construction of the cDNA libraries for Ion Proton sequencing

First, 5–100 ng amplified cDNA were fragmented to approximately 250 bp fragments using Covaris S2 and purified with DNA Clean & Concentrator™-5. The Ion Xpress™ Plus Fragment Library Kit (ThermoFisher, cat. no. 4471269) and Ion Xpress™ Barcode Adapters 1-16 Kit (Thermo Fisher, cat. No. 4471250) were used to construct the library with half reagents.



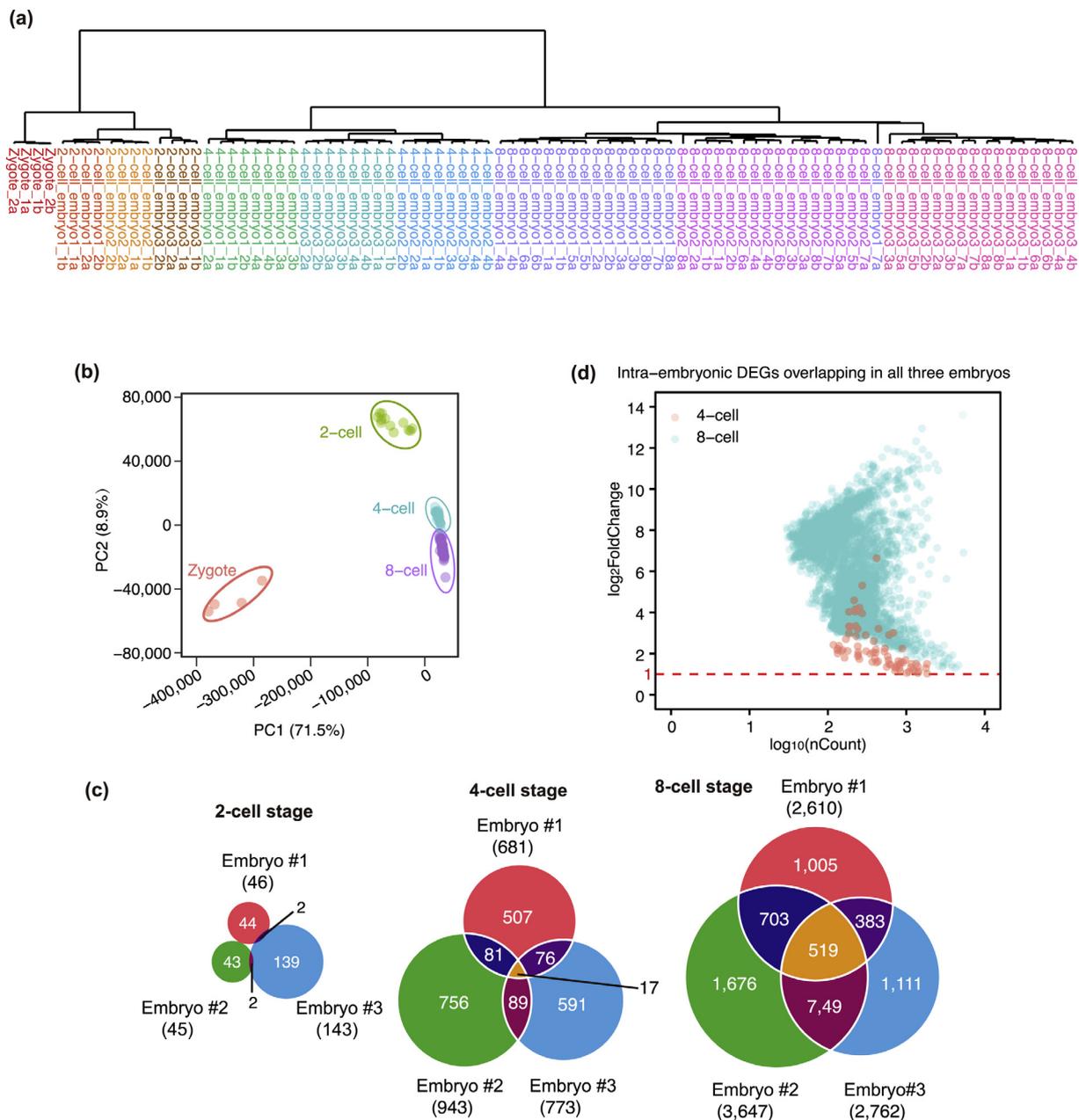
**Fig. 2.** Comparing MR-seq with the 'pool/split' approach. **a** Venn diagram of DEGs identified in MR-seq and the 'pool/split' analysis. **b** Scatter plot showing the DEGs identified in MR-seq and the 'pool/split' analysis. The CV is plotted against the  $\log_{10}$ -transformed geometric mean of expression for all genes detected in 8-cell single blastomeres (merged from MR-seq analysis dataset) and 8-cell technical replicates. The pink points showed the CV of genes from single blastomeres in the "pool/split" analysis, and the light green points showed the CV of genes from technical replicates in the "pool/split" analysis.

## 2.6. Ion PI template preparation, enrichment and sequencing on the Thermo Fisher Proton

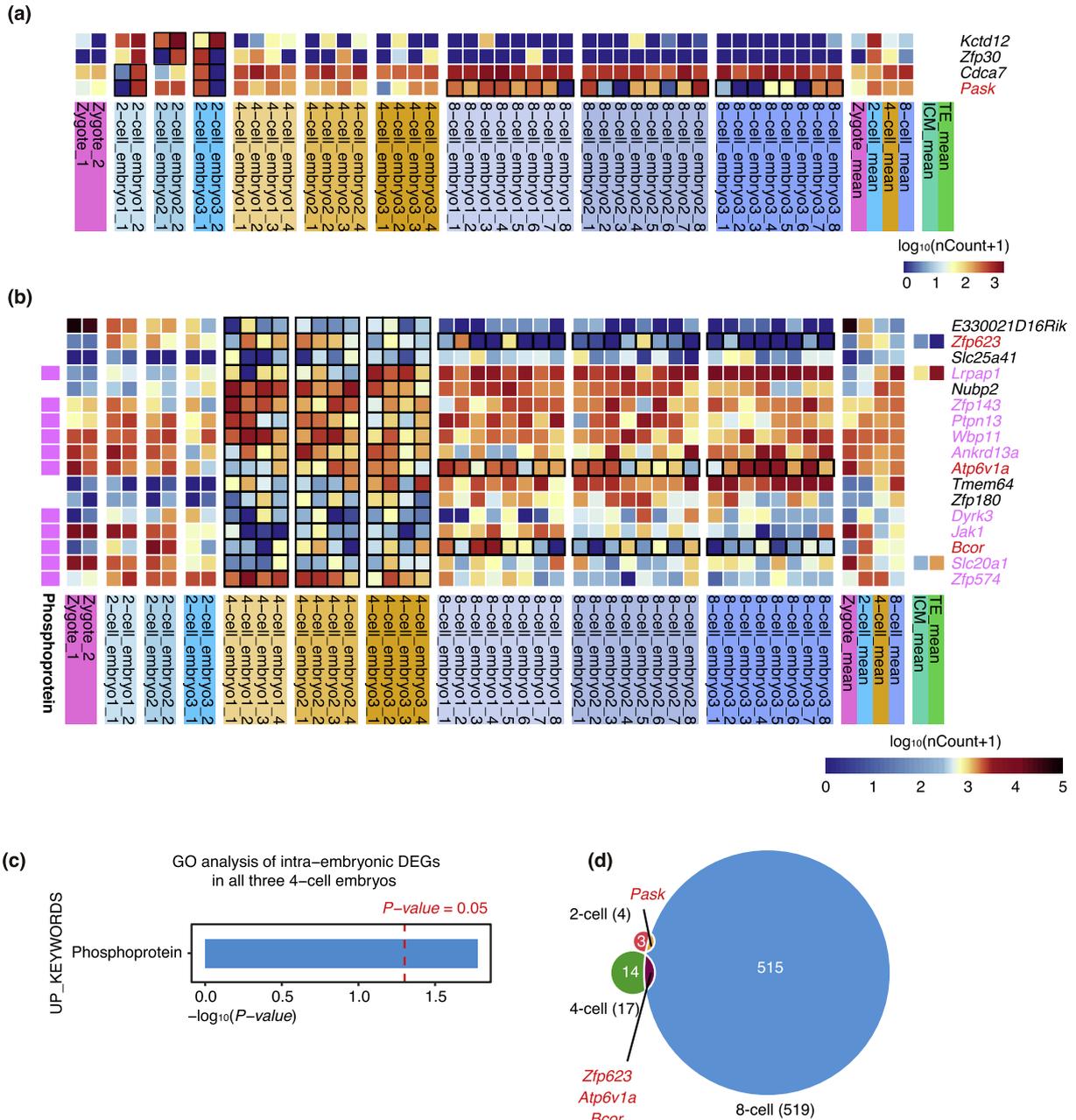
Emulsion PCR was performed using the One Touch 2 (OT2) system following the Ion PI Template OT2 200 V3 protocol (Life Tech MAN0009133 RevB.0) with 480 million DNA templates. Enriched Ion spheres were quantified, and approximately 400–600 million Ion spheres were recovered. PI chips were loaded according to the spinning protocol, and sequencing was performed using the Ion PI Sequencing 200 Kit V3 (Life Tech MAN0009136 Rev B.0) on an Ion PI Proton system. Base calls were collected using the Ion Torrent Suite v4.3 software.

## 2.7. MR-seq data alignment, analysis and graph drawings

The Illumina HiSeq 2500 and 2000 (Illumina, CA, USA) and Ion Proton (Life Technologies, CA, U.S.) sequencing systems were used for sequencing. Adaptor contamination and low-quality reads were discarded from the raw data. For the Illumina paired-end reads, TopHat (version 2.0.10) was used for sequence alignment, and FPKM values were generated by Cufflinks (version 2.1.1). For the Ion Proton RNA-seq reads, a 2-step alignment method that combined STAR (version 2.3.0) and Bowtie2 (version 2.1.0) was applied. Unmapped reads produced by STAR were re-mapped with Bowtie2 in the “very-sensitive-local” mode, and the merged bam files were used by Cufflinks as the input. Reference sequence and transcript annotation



**Fig. 3.** MR-seq detected intra-embryonic heterogeneity in mouse early embryos. **a** Unsupervised hierarchical clustering of the half blastomeres from zygote to 8-cell stages. **b** Two-dimensional principal component analysis (2D PCA) of the half blastomeres from zygote to 8-cell stages: zygote ( $n = 4$ ), 2-cell ( $n = 12$ ), 4-cell ( $n = 24$ ) and 8-cell ( $n = 48$ ) stages. PC1 and PC2 represented the top two dimensions of the genes showing differential expression among the embryos, which accounted for 71.5% and 8.9% of the expressed genes, respectively. **c** Number of intra-embryonic heterogeneous genes from 2- to 8-cell stages. The intra-embryonic heterogeneous genes were identified according following cutoff:  $|FC| > 2$ , *adjusted P* < 0.05. FC, fold change. **d** Scatter plot showed how the log<sub>2</sub>FoldChange of intra-embryonic heterogeneous blastomeres' genes change with the mean expression (nCount).



**Fig. 4.** Intra-embryonic heterogeneity in 2- and 4-cell stages. The heat map of the intra-embryonic heterogeneous genes in 2-cell **a** and 4-cell **b** stages. The embryos showing heterogeneity were highlight with black outlines. **c** The up keywords GO analysis of the 17 intra-embryonic heterogeneous genes in 4-cell stage. **d** The venn plot showed the intra-embryonic heterogeneous genes overlapped in two stages.

files were from the UCSC genome assembly mm10. DEGs were found using the R package “Deseq2” (version 1.6.1) [25], and the counts required for in Deseq2 were achieved by HTSeq (version 0.6.0). The entire correlation or distance matrixes were scaled and the unsupervised hierarchical clustering was performed using the ‘heatmap.2’ function from the “gplots” package in R. For the CV scatterplot, average CV and standard deviation of FPKM values from the technical replicates and 8-cell blastomeres were calculated within each bin (0.15) and plotted in log 10 scale. The 2D PCA plot was generated by R with the “ggbiplot” package. The Venn graphs are pictured on the <http://www.cmbi.ru.nl/cdd/biovenn/index.php> [26].

2.8. Ethical approval

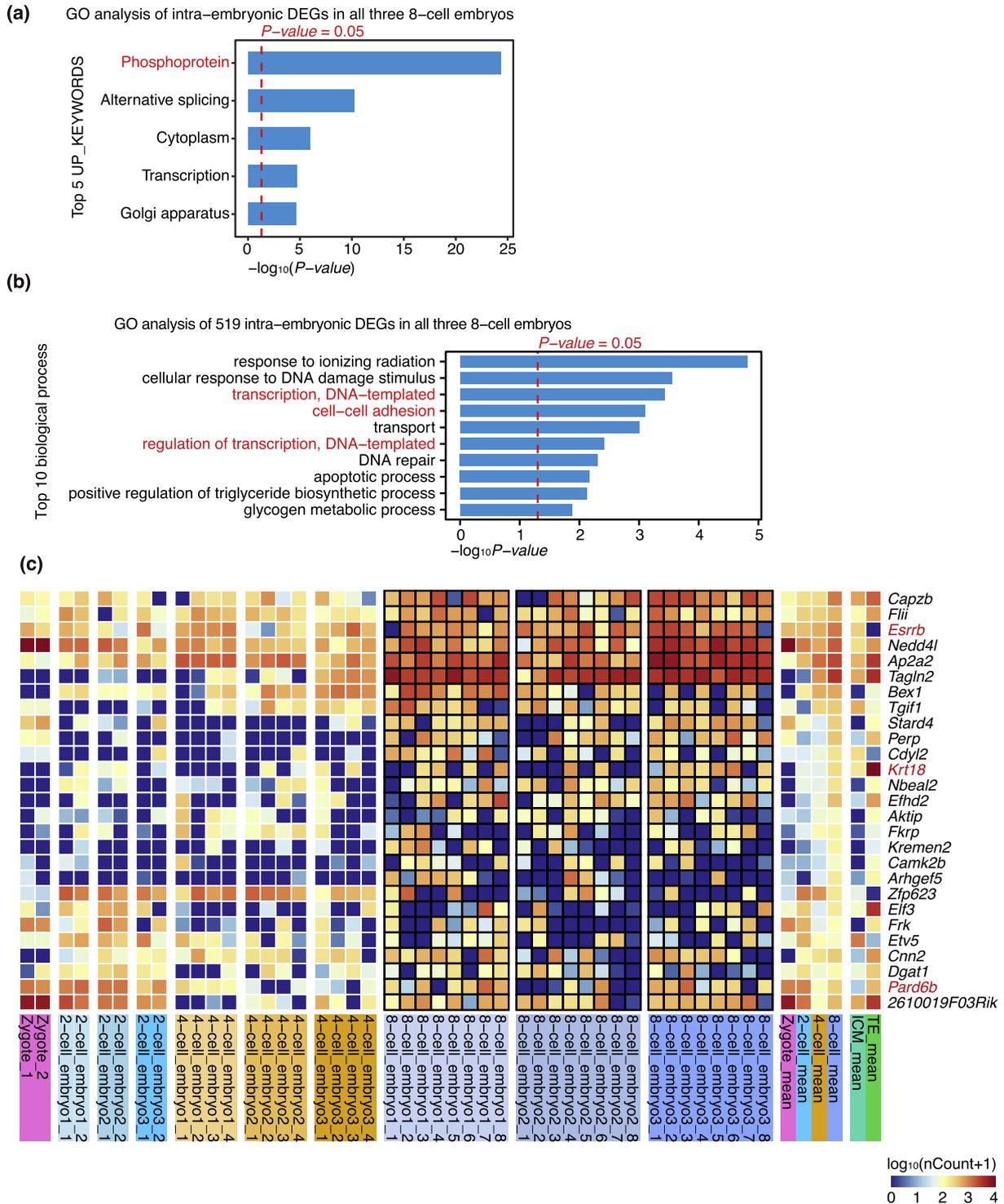
All procedures performed in studies involving human participants were in accordance with the ethical standards of the institu-

tional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. For studies with animals, all applicable institutional and/or national guidelines for the care and use of animals were followed.

3. Results

3.1. Performance of MR-seq

Firstly, we determined the sensitivity of MR-seq in comparison with the standard single-cell RNA-seq method. We split the lysate of single mouse embryonic stem cells (mESCs) into 2, 4 or 8 aliquots and constructed their libraries separately. The results showed that the number of genes (FPKM > 1) detected in half-cell lysate was comparable to that in a whole single cell (average



**Fig. 5.** Intra-embryonic heterogeneity in 8-cell stage. The top 5 up keywords **a** and top 10 biological process **b** GO analysis of the 519 intra-embryonic heterogeneous genes in 8-cell stage. **c** The heat map of the 27 intra-embryonic heterogeneous genes in 8-cell stage. The 27 genes are also the ICM and TE specific genes.

7,345 vs 8,172 for half-cell lysate vs single-cell, Fig. 1a). The gene expression correlation between half-cell lysate and the bulk mESCs was also comparable to that between the whole single cell and the bulk (the Spearman's rank correlation coefficient, median 0.71 vs 0.72 for bulk & half-cell vs bulk & single-cell, Figs. 1b and S1). The fraction of overlapping genes of two half-cell lysates was also similar to that of two single cells (Jaccard index, 62% vs 65% for half-cell lysates vs single cells, Fig. 1c). On the contrast, the detection sensitivity and reproducibility greatly decreased when divid-

ing the lysate into 4 or 8 equal aliquots (Fig. 1a and b). These results indicated that MR-seq by 2-split, but not more splits, gives a technical sensitivity comparable to the standard single-cell RNA-seq.

To assess the ability of MR-seq to identify DEGs between only two single cells, we examined the mESCs and the mouse embryonic fibroblasts (MEFs). In condition of only one sample in each group, no statistical analysis can be performed, while measuring each sample twice allows for statistical analysis by adding

technical replicates and is expected to reduce the technique variation. The results showed that two measurements conducted by MR-seq indeed significantly ( $P < 0.001$ ) increased the positive predicate value in comparison with only one measurement (average 50% vs 42% for MR-seq vs single-cell, Fig. 1d). Also, the MR-seq successfully identified known DEGs between mESCs and the MEFs [27] (Fig. 1e).

Further, we compared MR-seq with the “pool/split” experiment for identification of highly variable genes. We analyzed each blastomere of an 8-cell embryo by MR-seq for identification of the DEGs. For the “pool/split” experiment, the dataset of half-blastomere lysate from the same blastomere was merged to identify the highly variable genes by comparing with the control of the technical variation obtained from another 8-cell embryo. The MR-seq method identified 2,092 DEGs and the “pool/split” approaches identified 1,014 highly variable genes, with 225 genes overlapped (Fig. 2a). The 225 overlapping genes included embryonic development-related genes such as *Nanog* and *Krt18* (Figs. 2b and S2b). Interestingly, of the 1,867 MR-seq only genes, gene ontology (GO) terms strongly enriched in functions of cell cycle, chromatin modification, RNA processing, in utero embryonic development (Fig. S2a), as well as many transcriptional factors (Fig. 2b). In contrast, of the 789 highly variable genes identified by the “pool/split” analysis slightly enriched in functions of translation and protein localization (Fig. S2c). These results implied that MR-seq method identified many potential intra-embryonic heterogeneous genes that could be missed by the ‘pool/split’ approach (Figs. 2 and S2).

### 3.2. Analysis of early mouse preimplantation embryos by MR-seq

We next applied MR-seq to detect the intra-embryonic heterogeneous genes in early mouse preimplantation embryos from zygote to 8-cell stages. We analyzed a total of 11 embryos of 2 zygotes and 3 embryos each for 2-, 4-, and 8-cell stages. Unsupervised hierarchical clustering and PCA analysis clearly distinguished embryos of different stages, confirming the technical sensitivity of MR-seq (Fig. 3a and b). We identified 4 genes which showed potential intra-embryonic heterogeneity overlapping between two embryos in the 2-cell stage, yet no genes overlapped among all three embryos could be identified (Fig. 3c). In the 4-cell stage, we identified 17 genes showing intra-embryonic heterogeneity overlapped among three embryos. The gene number increased to 519 in the 8-cell stage (Fig. 3c, Tables S2 and S3). These results indicated that the intra-embryonic heterogeneity prominently increased from 2- to 8-cell stage. The gene expression levels of all these identified intra-embryonic heterogeneous genes were relatively high in the blastomere of the identified heterogeneous stages, suggesting they were not caused by technical variation (Fig. 3d).

Among all these potential intra-embryonic heterogeneous genes, 4 genes (*Pask*, *Zfp623*, *Atp6v1a*, *Bcor*) showed differentially expression at two developmental stages (Fig. 4a, b and d). *Pask*, which is a PAS domain containing serine/threonine kinase gene, showed intra-embryonic heterogeneity in both 2- and 8-cell stages (Fig. 4a). It seems that it was also differentially expressed in one of three 4-cell stage embryos. Three genes, *Zfp623*, *Atp6v1a*, *Bcor*, displayed differentially expression at both 4- and 8-cell stage (Fig. 4b). Interestingly, three of these 4 gene, *Pask*, *Atp6v1a*, and *Bcor*, were noted as phosphoproteins in GO, which indicated that their functions are posttranslational regulated by phosphorylation modification. We found that more than half of 4-cell stage (11 of 17 genes, 65%) and 8-cell stage (274 of 519 genes, 53%) heterogeneous genes were phosphoproteins (Figs. 4b and 5a, and Table S4). In addition, both 4- and 8-cell stage heterogeneous genes were significantly enriched in phosphoproteins ( $P = 0.02$  for 4-cell stage, and  $P = 4.8 \times 10^{25}$  for 8-cell stage, Figs. 4c and 5a).

The earliest major lineage separation of the mammalian embryo is the separation of the inner cell mass (ICM) and the trophectoderm (TE). We found that many ICM or TE specific genes, including *Zfp623*, *Lrpap1*, *Slc20a1*, *Esrrb*, *Krt18* and *Pard6b*, already showed intra-embryonic heterogeneity at 4- or 8-cell stages (Figs. 4b and 5c). The GO analysis also showed that 519 heterogeneous genes in the 8-cell stage enriched DNA-templated transcription ( $P$ -value =  $4.1e-04$ , e.g. *Yap1*, *Hsf1*, *Mbd1*, *Pou2f1*, *Esrrb* and *Dnmt1*) and cell-to-cell adhesion ( $P = 0.002$ ) (Figs. 5b, S3 and Table S5).

## 4. Discussion and conclusion

We present here a novel single-cell RNA-seq method allowing for identification of DEGs between just two single cells. We applied this method for identifying the intra-embryonic heterogeneous genes among mouse 2-, 4- and 8-cell stage embryos.

We have determined that the sensitivity of MR-seq is comparable with the standard single-cell RNA-Seq method. This indicates that although dividing the cell lysate into 2-splits reduces the amount of starting material, the loss of transcripts during the dividing step does not significantly reduce the performance of the MR-seq method. Our results suggest that MR-seq can be used for cells as small as mESCs, MEF and mouse blastomeres. Since the technical noise should increase as the size of the cell decreases, whether MR-seq also works for smaller cells needs further study.

We applied MR-seq for early mouse preimplantation embryos and identified hundreds of intra-embryonic heterogeneous genes among mouse 2-, 4- and 8-cell stage embryos. We found that more than half of these heterogeneous genes were phosphoproteins. These suggest that protein phosphorylation plays an important role in early cell fate decision, i.e. the separation of the inner cell mass (ICM) and the trophectoderm (TE), and further separation of ICM to epiblast (EPI) and primitive endoderm (PE), of mammalian blastomeres. For example, the phosphorylation of Yes-associated protein (Yap), which is encoded by a heterogeneous gene *Yap1* identified in 8-cell stage, keeps the Yap protein in the plasma and thus prevents its transcription activation of *Cdx2*, a master gene for the TE lineage [28]. Our intra-embryonic heterogeneous gene list also includes some known genes that play essential roles in early mammalian embryo development. We found that *Nanog*, which is a key transcription factor for segregation of PE vs EPI in the ICM, already showed intra-embryonic heterogeneity in the 8-cell stage [29,30]. Whether this intra-embryonic heterogeneity is related to the later cell fate decision requires further investigation. Also, *Pard6a* and *Pard6b*, which are cell polarity regulators, have been associated with the first cleavage of the mouse embryo [31]. Our finding of their intra-embryonic heterogeneous expression in the 8-cell stage may point to their roles in later cleavage stage.

### Conflict of interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

The authors would like to acknowledge the staffs of BIOPIC sequencing facility at Peking University for assistance. This work was supported by grants from the Beijing Municipal Science and Technology Commission (D15110700240000).

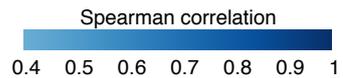
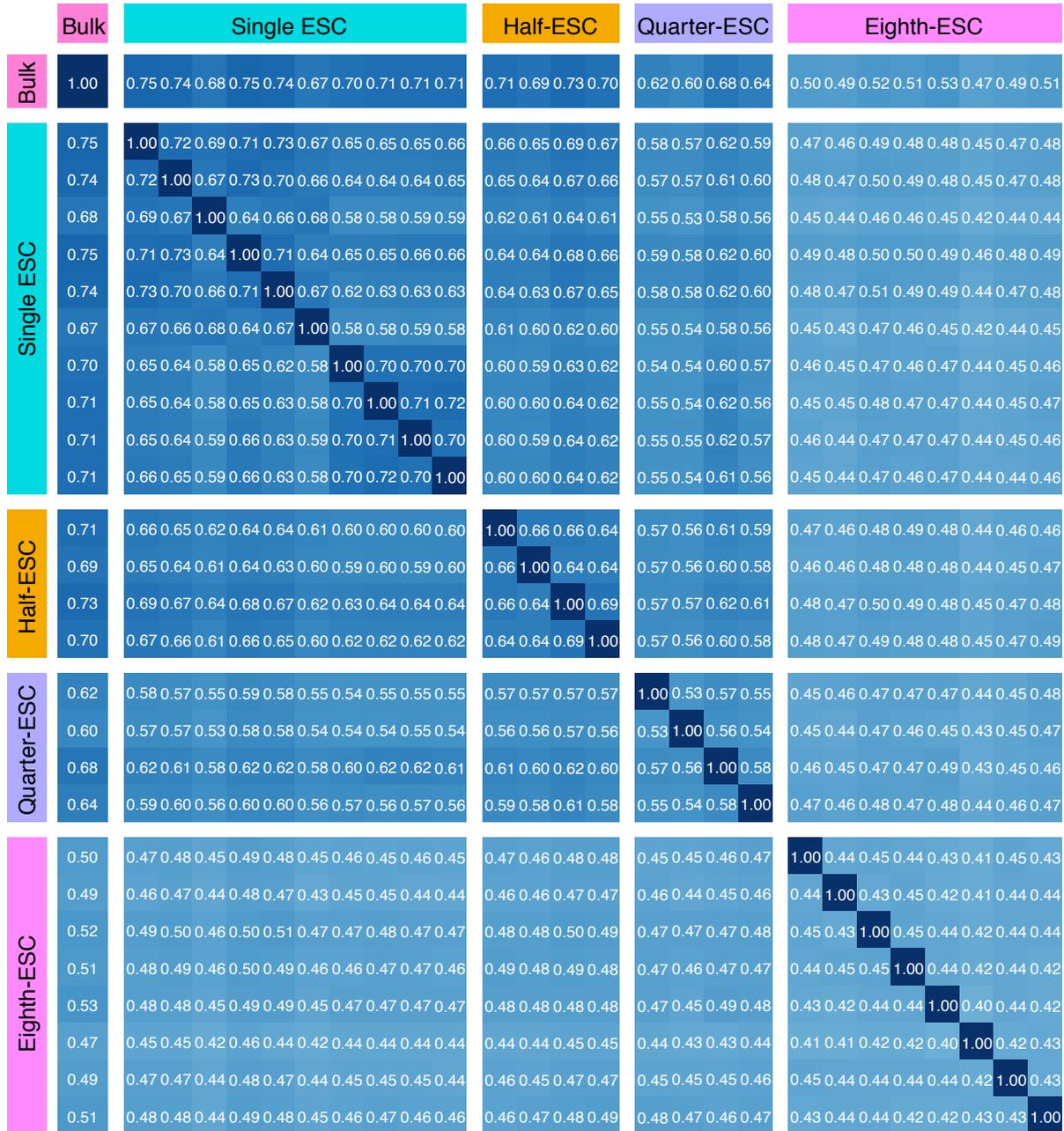
## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.scib.2017.01.029>.

## References

- [1] Tang F et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;6(5):377–82.
- [2] Tang F et al. RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc* 2010;5(3):516–35.
- [3] Streets AM et al. Microfluidic single-cell whole-transcriptome sequencing. *Proc Natl Acad Sci USA* 2014;111(19):7048–53.
- [4] Wen L, Tang F. Single-cell sequencing in stem cell biology. *Genome Biol* 2016;17(1):71.
- [5] Picelli S et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* 2014;9(1):171–81.
- [6] Tang F et al. Deterministic and stochastic allele specific gene expression in single mouse blastomeres. *PLoS One* 2011;6(6):e21208.
- [7] Xue Z et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 2013;500(7464):593–7.
- [8] Yan L et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 2013;20(9):1131–9.
- [9] Deng Q et al. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* 2014;343(6167):193–6.
- [10] Petropoulos S et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* 2016.
- [11] Goolam M et al. Heterogeneity in Oct4 and Sox2 targets biases cell fate in 4-cell mouse embryos. *Cell* 2016;165(1):61–74.
- [12] Zhou F et al. Tracing haematopoietic stem cell formation at single-cell resolution. *Nature* 2016;533(7604):487–92.
- [13] Ramskold D et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 2012;30(8):777–82.
- [14] Shalek AK et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 2014;510(7505):363–9.
- [15] Usoskin D et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat Neurosci* 2014.
- [16] Treutlein B et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 2014;509(7500):371–5.
- [17] Patel AP et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014;344(6190):1396–401.
- [18] Fan HC, Fu GK, Fodor SP. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science* 2015;347(6222):1258367.
- [19] Klein AM et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 2015;161(5):1187–201.
- [20] Macosko EZ et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015;161(5):1202–14.
- [21] Brennecke P et al. Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 2013;10(11):1093–5.
- [22] Wu AR et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods* 2014;11(1):41–6.
- [23] Marinov GK et al. From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome Res* 2014;24(3):496–510.
- [24] Grun D, Kester L, van Oudenaarden A. Validation of noise models for single-cell transcriptomics. *Nat Methods* 2014;11(6):637–40.
- [25] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15(12):550.
- [26] Hulsen T, de Vlieg J, Alkema W. BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 2008;9:488.
- [27] Islam S et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 2011;21(7):1160–7.
- [28] Nishioka N et al. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell* 2009;16(3):398–410.
- [29] Mitsui K et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 2003;113(5):631–42.
- [30] Silva J et al. Nanog is the gateway to the pluripotent ground state. *Cell* 2009;138(4):722–37.
- [31] Gray D et al. First cleavage of the mouse embryo responds to change in egg shape at fertilization. *Curr Biol* 2004;14(5):397–405.

Fig S1



Yang *et al.*, Fig S1

**Fig S2**

