RNA-Seq analysis to capture the transcriptome landscape of a single cell

Fuchou Tang[1], Catalin Barbacioru[2], Ellen Nordman[2], Bin Li[2], Nanlan Xu[2], Vladimir I Bashkirov[2], Kaiqin Lao[2] & M Azim Surani[1]

[1]Wellcome Trust/Cancer Research UK Gurdon Institute of Cancer and Developmental Biology, University of Cambridge, Cambridge, UK. [2]Genetic Systems, Applied Biosystems, part of Life Technologies, Foster City, California, USA. Correspondence should be addressed to M.A.S. (a.surani@gurdon.cam.ac.uk) or K.L. (kai.lao@lifetech.com).

Published online 25 February 2010; doi:10.1038/nprot.2009.236

We describe here a protocol for digital transcriptome analysis in a single mouse oocyte and blastomere using a deep-sequencing approach. In this method, individual cells are isolated and transferred into lysate buffer by mouth pipette, followed by reverse transcription carried out directly on the whole cell lysate. Free primers are removed by exonuclease I and a poly(A) tail is added to the 3′ end of the first-strand cDNAs by terminal deoxynucleotidyl transferase. Single-cell cDNAs are then amplified by 20 + 9 cycles of PCR. The resulting 100–200 ng of amplified cDNAs are used to construct a sequencing library, which can be used for deep sequencing using the SOLiD system. Compared with cDNA microarray techniques, our assay can capture up to 75% more genes expressed in early embryos. This protocol can generate deep-sequencing libraries for 16 single-cell samples within 6 d.

INTRODUCTION

Developing tools for transcriptome analysis

The identity and function of a cell is determined by its entire RNA component, which is called the transcriptome[1,2]. The transcriptome is the functional readout of the genome and epigenome. In an organism, essentially every cell has the same genome, although every cell type and potentially each individual cell has a unique transcriptome. Ideally, transcriptome analysis should capture the exact quantity of all full length RNAs of all classes at single-base resolution in the smallest functional unit of an organism, i.e., an individual cell[2]. Eventually, the transcriptome analysis may even become non-invasive, allowing us to read the sequences of every RNA molecule of a living cell without destroying the cell. During the past decade, the most successful and widely used transcriptome analysis method has been the cDNA microarray[3–8], and this technique is now applicable to virtually any model organisms with a known genome. These arrays represent a powerful way to capture the expression pattern of tens of thousands of known genes through hybridization. However, this technology has significant drawbacks and limitations[1,2], including (1) cross-hybridization between genes of similar sequence; (2) detection of expression levels limited to the 1,000-fold range or three orders of magnitude, compared with actual cellular dynamic range of gene expression that spans six orders of magnitude; (3) lack of information about the exact length and sequence of the mRNAs being analyzed; and (4) inability to detect novel transcripts.

Use of a tiling array can resolve some of these problems[9,10], but recently developed methods for deep-sequencing-based transcriptome analysis or RNA-Seq, can potentially overcome all of these problems[1,2,11–14]. RNA-Seq can achieve single-base resolution, and the dynamic range of gene expression levels that it can capture is theoretically unlimited, depending only on the depth of sequencing. More importantly, with the help of complete genome information, the exact length and sequence of all RNAs analyzed can be captured accurately. The past 2 years have seen astonishingly fast development of the RNA-Seq technique and deepening of our understanding of the complexity of the eukaryotic transcriptome, from yeast to humans and from adult tissues to embryonic development[15–23]. However, the sensitivity of the method generally requires microgram amounts of total RNA as the starting material, and most RNA-Seq studies have used tissues containing a mixture of different types of cells or cell lines, which are at least a mixture of cells at different stages of the cell cycle. Recent investigations into the stochastic nature of transcription and gene expression have shown that even in the same cell type at the same cell cycle stage, the mRNA copy number from an expressed gene can be affected by both the microenvironment and the intrinsic noise of the transcription process[24–29]. Ideally, the RNA-seq transcriptome analysis should be done using individual cells or even the sub-compartment of a cell, such as the cytoplasm or nucleus. For the study of early embryonic development or stem cells *in vivo*, it is extremely difficult or even practically impossible to isolate millions of cells of a single type. RNA-Seq at single-cell resolution will greatly promote the development of these fields by permitting comprehensive capture of the expression dynamics of all genes at all developmental stages.

During the past few years, people have developed cDNA microarray techniques for working with a small amount of starting material, or even a single cell[30,31], through either *in vitro* transcription (IVT)-based linear amplification[32–35], PCR-based exponential amplification[36–39] or a combination of the two methods[40,41]. However, these approaches have inherited the limitations of the microarray platform. In recent work[42], our group described how we improved a widely used single-cell cDNA amplification strategy and combined it with the SOLiD deep-sequencing system to develop a digital transcriptome analysis method, single-cell RNA-Seq.

Strengths and limitations of single-cell RNA-Seq

We have demonstrated that single-cell RNA-Seq has greater accuracy than single-cell cDNA microarrays[40,42]. There are two reasons for this. First, owing to the greater sensitivity of the deep sequencing compared with the cDNA microarray, IVT is unnecessary, thus eliminating the amplification bias introduced by this step. Second, the greater dynamic range of the deep-sequencing method means that RNA-Seq more accurately captures levels of gene expression. Using this method, we captured 5,270 (75%) more genes than

the cDNA microarray (i.e., the commercial array (Affymetrix Mouse Genome 430 2.0 array))[43], expanding the transcriptome of a blastomere of a four-cell stage mouse embryo from 7,050 expressed genes to 11,920 expressed genes[42] (i.e., we detected 5,270 more genes than microarray, but we also missed 400 genes from the 7,050 genes detected by microarray). We observed that around 10% of genes with multiple known transcript isoforms express more than two different transcript isoforms in the same blastomere at the same time point, which has never been detectable by single-cell cDNA microarray[42]. We also identified thousands of previously unknown splice junctions from already known genes, indicating the extent to which our previous appreciation of the complexity of the eukaryotic transcriptome was limited based on the cDNA microarray[42].

This strategy has many possible future applications. It creates the possibility of isolating one or several cells from patient organs or tumors and analyzing their transcriptome by RNA-seq, reducing the invasiveness of biopsies and clarifying the diagnosis of diseases. In addition, the subcellular distribution of mRNAs could potentially be explored by comparing, e.g., the axon and cell body of a neuron.

Current limitations of single-cell RNA-Seq include[42]:

(1) The assay can only capture mRNAs with a poly(A) tail. Those mRNAs without poly(A) tails or other classes of RNAs, such as small non-coding RNAs, will not be detected.
(2) The assay cannot discriminate sense from antisense transcripts.
(3) The assay can only capture cDNA fragments extending upto 3 kb from the 3′ end of a mRNA. For genes with mRNAs longer than 3 kb, which constitute about 36% of all known genes, the 5′ end part of the mRNA can not be detected. We are at present improving the method to overcome these limitations.

RNA isolation and preparation

The use of this technique requires efficient and reliable methods for dissociating the embryos or tissues into a single-cell suspension. Any suitable strategies can be used to disaggregate the tissues into single-cell suspension while maintaining the integrity of living single cells under physiological conditions. In general, quicker and milder dissociation processes should be adopted. These will depend on the type of cells and which stage the embryos are at.

For picking up and transferring individual cells, we usually use a mouth tube to control an attached micropipette under a dissection microscope, which permits swift and efficient control of individual cell collection and release[41,44]. In this method, four-cell stage mouse embryos are isolated from the oviduct[45]. Acidic Tyrode's solution is used to dissolve and remove the zona pellucida and the embryos are then transferred to a calcium-free medium to dissociate the zona-free embryos into individual blastomeres. PBS–BSA is used to wash and hold individual blastomeres; the use of PBS alone will cause the blastomeres to attach tightly to the bottom of the dish or inside wall of the micropipette, leading to rupture or loss of these cells. The individual blastomeres should be placed in a small drop of PBS–BSA, usually only 50–150 μl on a 6-cm cell culture dish; this will allow accurate and reliable manipulation of single cells. The portion of the micropipette outside the mouth tube should be long enough (>4 cm) to permit the tip of it to dip into the lysis buffer at the bottom of the tube without the mouth tube touching the lid or rim of the tube, thus avoiding any potential contamination. Gloves should be worn during the manipulation of single cells to prevent potential contamination.

It is crucial to make sure that the individual cells to be analyzed have actually been transferred into the lysis buffer from the drop of PBS–BSA via the micropipette. Initially, the operator can practice by transferring individual cells into a 4–5 μl drop of PBS–BSA on a 6-cm cell culture dish, and then recovering it to ascertain whether they can accurately transfer an individual cell.

We use a relatively mild lysis buffer[41,42], which enables efficient cell lysis without severely interfering with reverse transcription efficiency. If more than one cell is used for each lysate reaction, dissociated individual cells must still be used rather than a clump of cells to make sure that all the cells are thoroughly and completely lysed under these conditions. As the method is extremely sensitive, it is essential to include a negative control sample, which involves repeating all of the steps described above, but without any cells. This negative control will ensure the absence of contamination at all steps from the start to the end of the whole procedure. The UP1 primer used here for first-strand cDNA synthesis contains a 24-nt poly(dT) tail at the 3′ end, which permits priming of the reverse transcription of all mRNAs with poly(A) tails and a 24-nt anchor sequence to add a universal tail to the cDNAs[41].

The protocol can also be carried out with purified total RNA instead of single cells. In such cases, simply use total RNA to replace the single cell and keep the final volume of the reverse transcription reaction at 5 μl (although the volume of water in the lysis buffer should be adjusted accordingly). For each sample, 20 pg–1 ng of total RNA should work.

Before cDNA synthesis, prepare the master mix based on the number of samples to be analyzed. In order to minimize pipetting error, prepare an excess stock solution of master mix, at least 12× of reagents for ten samples for all steps except poly-A tailing and PCR step, for which an 11× master mix is sufficient. For the reverse transcription step, it is essential to add the 0.45 μl of RT mix accurately into each reaction tube. In the free primer removal step and poly(A) tailing step, it is important to prepare the master mixes immediately before adding them into the reaction tube. For the PCR step, it is crucial to use hot-start Ex Taq to minimize background amplification of single-cell cDNAs during pipetting and mixing. And for all the steps, it is very important to mix the reaction solutions mildly but thoroughly to ensure that all the reactions are efficient and that the amplification of cDNAs is accurate and quantitative. The UP2 primer used here has a 24-nt poly(dT) at the 3′ end, which permits priming during the second-strand cDNA synthesis and the 24-nt anchor sequence adds a universal tail to the second-strand cDNAs[41]. During PCR, the extension time is increased by 6 s after each cycle to compensate for the reduced efficiency of PCR amplification because of the accumulation of byproducts and consumption of the active reagents.

After the first 20 cycles of PCR, the PCR product is purified and part of it is used as the template for an additional 9 cycles of PCR. There are two purposes for this step: (1) to increase both the total amount and concentration of the amplified cDNAs for construction of the deep-sequencing library and other relevant analysis and (2) to shift the primers from unmodified UP1 and UP2 to 5′-amine (NH2)-blocked AUP1 and AUP2 to suppress the contamination of primer dimers in the deep-sequencing library. The AUP1 and AUP2 primers have an NH2-modification at their 5′ ends with a C_6 linker that prevents ligation of the 5′-end fragments of the double-stranded cDNA (after shearing) to the SOLiD library P1/P2 adaptors. Because of this, any undesired small amplification products that may have been formed will not be carried over into the library.

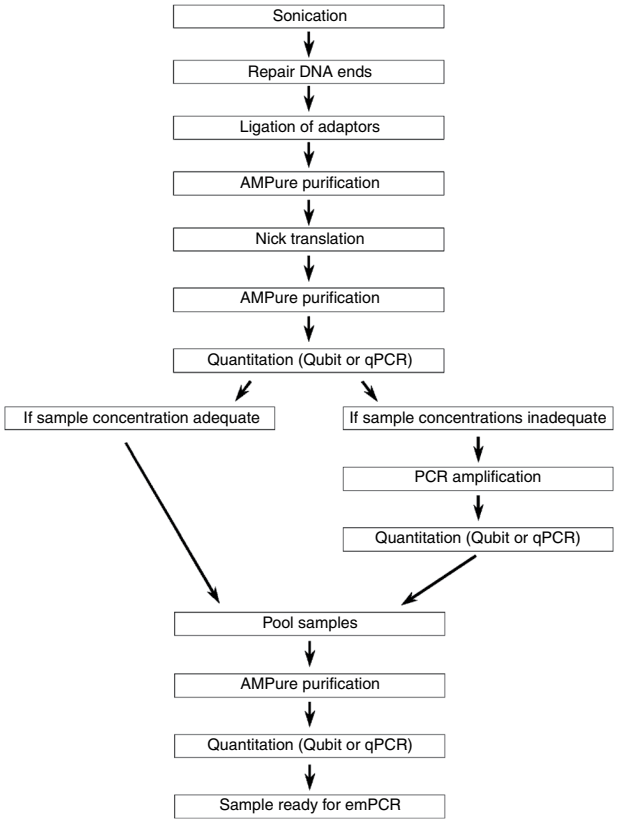

Figure 1 | Workflow for SOLiD System Express library preparation for fragment libraries and multiplexed fragment libraries. Following generation of cDNA from a single cell, typically 100–200 ng is used for standard fragment-library preparation. Using the Covaris S2 system, cDNA (0.5–3 kb) is sheared into short 80–130 bp fragments. The ends of the fragments are repaired and subsequently ligated to SOLiD P1 and P2 adaptors. The fractionated, adaptor-ligated DNA is then subjected to 8–10 cycles of PCR amplification. Ultimately, emulsion PCR reactions are carried out by mixing 1-2 ng of single-cell libraries with 1.6 billion 1-µm diameter beads with P1 primers covalently attached to their surfaces. The 50-base sequences were obtained using Applied Biosystems SOLiD sequencing.

The number of PCR amplification cycles can be adjusted according to the size or mRNA content of the cell under investigation; for a blastomere from a four-cell stage embryo we use 20 cycles of PCR, whereas for a mature oocyte we use 18 cycles of PCR. The template amount used for the second round of PCR can also be flexible. Out of 50 µl of the purified first round PCR product, 1–4 µl can be used for each 90-µl PCR reaction. This round of PCR amplification can also be adjusted according to the final amount of cDNA needed; the number of cycles can be between 6 and 12, but should be minimized to reduce bias. The combined number of cycles in the first and second rounds of PCR should be <30. The final PCR product should be gel purified to remove primer dimers and any other potential non-specific amplification byproducts. In our experience, from one blastomere of a four-cell stage embryo (or 20 pg of input total RNA), 20 + 9 cycles of PCR will yield at least 200 ng of gel-purified cDNA.

Library preparation

For a workflow of this part of the process, consult **Figure 1**. The primary steps are as follows:

- *The prepared cDNA samples are sheared using sonication, with the majority of fragments between* 80–130 *bp* (see **Fig. 2a**). Fine-tuning the shearing protocol may be necessary for certain DNA samples. We have tested the conditions described below in Steps 43 and 44 for shearing 50 ng–10 µg of DNA in a total volume of 100 µl in a Covaris S2 System. For this system, it is important to use the appropriate tube holder, program the instrument correspondingly and set the chiller temperature at 2–5 °C. This ensures that the water bath stays around 5 °C, preventing excessive heating of the sample during sonication.
- *Fragment ends are repaired to enable ligation of double-stranded adaptors.* The quantity of double-stranded adaptors must be calculated based on the quantity of cDNA sample that was sonicated. For a fragment library, P1 and P2 adaptors are ligated to the fragment ends. All oligonucleotides needed for fragment library construction are included in the SOLiD Fragment Library Oligo Kit. Alternatively, multiplexed P1 and P2 adaptors can be used for large sample sets, allowing the user to prepare a multiplexed library for emulsion PCR and running on the SOLiD instrument. A multiplexed library reduces workflow and allows the user to run (in this protocol) up to 16 samples in parallel on SOLiD, which reduces sequencing time and reagents. All multiplexed oligonucleotides must be ordered (see **Table 1**) and pair-wise annealed (see Procedure, **Box 1**) to create double-stranded adaptors.

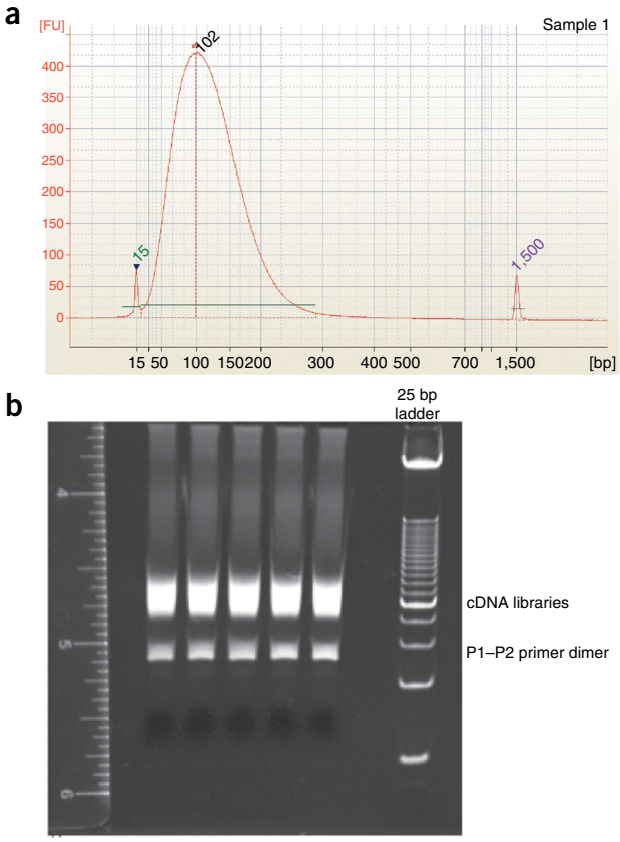

Figure 2 | cDNA library fragment QC. (**a**) Electropherogram of DNA sheared in C4011-10 tubes generated using Agilent's Bioanalyzer. (**b**) Gel picture of library cDNAs. The 68 bp bands are P1–P2 primer-dimer products, whereas the cDNA libraries are bands between 150 bp and 200 bp. (The size of the sheared cDNA is 80–130 bp. After ligation of the two adaptors, totaling 68 bp, the size of the DNA is 150-200 bp).

TABLE 1 | All oligos needed for the single cell RNA-Seq protocol.

Oligo name	Oligo sequence (all oligos sequence read 5′→3′)
UP1	ATATGGATCCGGCGCGCCGTCGACTTTTTTTTTTTTTTTTTTTTTTT
UP2	ATATCTCGAGGGCGCGCCGGATCCTTTTTTTTTTTTTTTTTTTTTTT
AUP1[a]	(NH2)ATATGGATCCGGCGCGCCGTCGACTTTTTTTTTTTTTTTTTTTTTTT
AUP2[a]	(NH2)ATATCTCGAGGGCGCGCCGGATCCTTTTTTTTTTTTTTTTTTTTTTT
Fragment Library P1 PCR Primer	CCACTACGCCTCCGCTTTCCTCTCTATG
Fragment Library P2 PCR Primer	CTGCCCCGGGTTCCTCATTCT
Fragment Library P1 Adaptor -5′end	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT
Fragment Library P1 Adaptor 3′ end	ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGGTT
Fragment Library P2 Adaptor 5′ end	AGAGAATGAGGAACCCGGGGCAGTT
Fragment Library P2 Adaptor 3′ end	CTGCCCCGGGTTCCTCATTCTCT
P1-Adaptor-5-end	ATCACCGACTGCCCATAGAGAGGTT
P1-Adaptor-3-end	*CCTCTCTATGGGCAGTCGGTGAT*
P2-Bar-1-5-end	CGCCTTGGCCGTACAGCAG**GGGCTT**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-1-3-end	CTGCCCCGGGTTCCTCATTCTCT**AAGCCC**CTGCTGTACGGCCAAGGCG
P2-Bar-2-5-end	CGCCTTGGCCGTACAGCAG**GGTGTG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-2-3-end	CTGCCCCGGGTTCCTCATTCTCT**CACACC**CTGCTGTACGGCCAAGGCG
P2-Bar-3-5-end	CGCCTTGGCCGTACAGCAG**AAGGGG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-3-3-end	CTGCCCCGGGTTCCTCATTCTCT**CCCCTT**CTGCTGTACGGCCAAGGCG
P2-Bar-4-5-end	CGCCTTGGCCGTACAGCAG**CCGATG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-4-3-end	CTGCCCCGGGTTCCTCATTCTCT**CATCGG**CTGCTGTACGGCCAAGGCG
P2-Bar-5-5-end	CGCCTTGGCCGTACAGCAG**CAACGA**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-5-3-end	CTGCCCCGGGTTCCTCATTCTCT**TCGTTG**CTGCTGTACGGCCAAGGCG
P2-Bar-6-5-end	CGCCTTGGCCGTACAGCAG**GTGCCC**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-6-3-end	CTGCCCCGGGTTCCTCATTCTCT**GGGCAC**CTGCTGTACGGCCAAGGCG
P2-Bar-7-5-end	CGCCTTGGCCGTACAGCAG**GTCTGG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-7-3-end	CTGCCCCGGGTTCCTCATTCTCT**CCAGAC**CTGCTGTACGGCCAAGGCG
P2-Bar-8-5-end	CGCCTTGGCCGTACAGCAG**ACGGAG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-8-3-end	CTGCCCCGGGTTCCTCATTCTCT**CTCCGT**CTGCTGTACGGCCAAGGCG
P2-Bar-9-5-end	CGCCTTGGCCGTACAGCAG**GAAGGG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-9-3-end	CTGCCCCGGGTTCCTCATTCTCT**CCCTTC**CTGCTGTACGGCCAAGGCG
P2-Bar-10-5-end	CGCCTTGGCCGTACAGCAG**GACCGC**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-10-3-end	CTGCCCCGGGTTCCTCATTCTCT**GCGGTC**CTGCTGTACGGCCAAGGCG
P2-Bar-11-5-end	CGCCTTGGCCGTACAGCAG**CTCAGG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-11-3-end	CTGCCCCGGGTTCCTCATTCTCT**CCTGAG**CTGCTGTACGGCCAAGGCG
P2-Bar-12-5-end	CGCCTTGGCCGTACAGCAG**AGCGTT**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-12-3-end	CTGCCCCGGGTTCCTCATTCTCT**AACGCT**CTGCTGTACGGCCAAGGCG
P2-Bar-13-5-end	CGCCTTGGCCGTACAGCAG**CGGGTC**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-13-3-end	CTGCCCCGGGTTCCTCATTCTCT**GACCCG**CTGCTGTACGGCCAAGGCG
P2-Bar-14-5-end	CGCCTTGGCCGTACAGCAG**CGTCTG**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-14-3-end	CTGCCCCGGGTTCCTCATTCTCT**CAGACG**CTGCTGTACGGCCAAGGCG
P2-Bar-15-5-end	CGCCTTGGCCGTACAGCAG**TAGCGT**AGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-15-3-end	CTGCCCCGGGTTCCTCATTCTCT**ACGCTA**CTGCTGTACGGCCAAGGCG
P2-Bar-16-5-end	CGCCTTGGCCGTACAGCAG**GCGTT**TAGAGAATGAGGAACCCGGGGCAGTT
P2-Bar-16-3-end	CTGCCCCGGGTTCCTCATTCTCT**AAACGC**CTGCTGTACGGCCAAGGCG

(continued)

TABLE 1 | Continued.

Oligo name	Oligo sequence (all oligos sequence read 5′→3′)
16Barcode Library PCR Primer-1	CCACTACGCCTCCGCTTT*CCTCTCTATGGGCAGTCGGTGAT*
16Barcode Library PCR Primer-2	CTGCCCCGGGTTCCTCATTCT

Bold–indicates barcode sequence. Underline–16Barcode Library PCR primer-2 sequence for amplification during emulsion PCR. Italics–segment of 16Barcode Library PCR Primer-1 that matches P1 adaptor.
[a]The amine group is covalently linked to the 5′ end bases with a C6 linker.

- *The ligated samples are purified using the Agencourt AMPure kit.* This system utilizes reversible immobilization of DNA on paramagnetic beads, with reaction conditions optimized to selectively bind DNA fragments 100 bp and larger. Excess oligonucleotides, nucleotides, salts and enzymes are removed by a simple washing procedure. Purification will eliminate 68-bp primer-dimer contamination and reduce carryover of this product into the fragment libraries (**Fig. 2b**).
- *Purified samples are subjected to nick translation.* If the operator has multiplexed libraries and wishes to conserve reagents, they may quantitate their samples and pool the multiplexed libraries so that they are of equal concentration, then continue on to nick translation. Alternatively, if the operator has fragment libraries or does not wish to pool at this stage, they may directly proceed with nick translation.
- *Multiplexed libraries must be quantitated.* This can be done using either Qubit (according to the manufacturer's protocol) or using quantitative real-time PCR. For multiplexed samples analyzed by real-time quantitative PCR, normalize the Ct values and pool accordingly. For determining relative concentrations, two titration points are adequate per multiplexed sample. When quantitating samples for pooling by real-time qPCR, make sure that all SYBR Green PCR assay reactions are assembled and run on the same plate to avoid variation that can occur between plates and instruments. If the library concentration is determined to be too low for emulsion PCR (<1 ng μl^{-1}), it can be subjected to additional PCR amplification. For the final pooled multiplexed library, Qubit may also be used to determine the final concentration of the pooled samples, following the manufacturer's protocol. For a fragment library sample, either Qubit or qPCR may be used to determine the final concentration. For qPCR of fragment libraries, a Standard Library is used to generate a standard curve. Throughout this protocol, for all steps requiring the use of a 1.5 ml tube, the operator should use 1.5 ml Lo-Bind tubes (Eppendorf, Hamburg, Germany). For all centrifugation steps, adjust according to the g-forces specified in the protocol.

Sample analysis

Sequence data collected from the SOLiD instrument (Applied Biosystems, Foster City, CA, USA) are subjected to a whole transcriptome analysis pipeline that we have developed, which is an off-instrument software package for the analysis of experiment runs. The software and a more detailed description can be found at http://solidsoftwaretools.com/gf/project/transcriptome/docman/?subdir=43.

The pipeline maps 50 base reads from the transcript sample to the mouse genome. It starts from the assumption that transcriptome reads may cross an exon–exon boundary. To overcome this issue, we use an 'anchor-extend' method, in which subsequences from the beginning and end of the read are selected as 'anchors'. These 'anchors' are then mapped to the reference genome sequence, finding global alignments with an upper limit for the maximum number of mismatches. The next step is to extend the global alignment of the anchors as a local alignment of the entire read with the reference. A simple ungapped alignment algorithm is applied to find the shortest alignment with the maximum score. Reads that have a unique alignment or have multiple alignments, where the second-best alignment contains at least two more mismatches than the best alignment, are considered to be unambiguously aligned and used in downstream analysis. These reads are sorted according to their genomic locations and stored in a separate file.

BOX 1 | MULTIPLEX SAMPLE PREPARATION

1. Hybridization of 16 Barcode Library adaptor oligonucleotides for multiplexing (● **TIMING** 30 min). Prepare 1 mM stock of multiplexed oligonucleotides (see **Table 1** for oligo sequences of the 16 Barcode Library adaptors).
2. Mix equal volumes of 1 mM 5′ end and 3′ end oligonucleotides with enough 5× ligase buffer for a final concentration of 1× ligase buffer.
3. Hybridize the oligonucleotides by the following thermocycler program:

Temperature	Time
95 °C	5 min
72 °C	5 min
60 °C	5 min
50 °C	3 min
40 °C	3 min
30 °C	3 min
20 °C	3 min
10 °C	3 min
4 °C	—

4. Store hybridized multiplexed oligonucleotides at −20 °C until ready for use.

These data are used to generate feature (exons/transcripts/genes) counts, which can be normalized using the r.p.m. (read per million counts) method[20]. The same reads are used to generate base coverage files (in a wiggle format) that can be viewed directly in the UCSC genome browser.

Unambiguously mapped reads can also be used to detect new exon–exon junctions by aligning them to a reference containing all exon–exon junctions not reported in RefSeq and Ensembl as a component part of any transcript. Each exon–exon junction present in this reference is represented by 84 nucleotides, 42 on each side of the junction. Junctions that are covered by at least five reads are considered to represent new exon–exon junctions. Typically, 20–40 million reads are needed to cover the entire transcriptome for complex genomes (such as mouse or human). However, 100 million reads would be ideal for alternative splicing and new junction studies.

MATERIALS

REAGENTS

- Mouse blastomeres ▲ CRITICAL All experiments that use human and animals should comply with the institutional and national guidelines.
- Primers (sequences of UP1, UP2, AUP1, AUP2 and all multiplexed adaptors and primers are provided in **Table 1**; these can be synthesized at Sigma or IDT at 0.2 μM scale and HPLC purified)
- Ac-BSA (Sigma, cat. no. B8894)
- Acidic Tyrode's solution (Sigma, cat. no. T1788)
- 1× PBS (pH 7.2) (Gibco, cat. no. 14249-95)
- 10× PCR buffer II and 25 mM $MgCl_2$ (Applied Biosystems, cat. no. 4379878)
- Nonidet P-40 SP (Roche, cat. no. 11332473001) ! CAUTION Harmful; irritating to the skin, the eyes and the respiratory tract. Handle using appropriate safety equipment.
- SuperScript III reverse transcriptase with 0.1 M DTT (Invitrogen, cat. no. 18080-044 or 18080-085)
- RNase inhibitor (cloned) (40 U μl^{-1}) (Applied Biosystems, cat. no. AM2682)
- SUPERase-In (20 U μl^{-1}) (Applied Biosystems, cat. no. AM2694)
- T4 gene 32 protein (Roche, cat. no. 972983)
- Exonuclease I (New England Biolabs, cat. no. M0293S)
- Terminal deoxynucleotidyl transferase (TdT) (Invitrogen, cat. no. 10533-065 or 10533-073)
- 100 mM dATP (Promega, cat. no. U1201)
- RNase H (Invitrogen, cat. no. 18021-014 or 18021-071)
- TaKaRa ExTaq HS (with 10× ExTaq buffer with Mg^{2+} and 2.5 mM each dNTP mix) (Takara Bio, cat. no. RR006A or RR006B) ▲ CRITICAL Hot Start ExTaq is necessary to minimize background amplification of single-cell cDNAs during pipetting and mixing. Takara labels their dNTP mix as '2.5 mM each dNTP', meaning it contains 2.5 mM dATP, 2.5 mM dTTP, 2.5 mM dCTP and 2.5 mM dGTP
- QIAquick PCR purification kit (Qiagen, cat. no. 28106)
- QIAquick gel extraction kit (Qiagen, cat. no. 28706)
- GeneAmp dNTP blend (100 mM) (Applied Biosystems, cat. no. N8080261)
- 3 M sodium acetate (pH 5.5) (Ambion, cat. no. AM9740) ! CAUTION Harmful if swallowed or inhaled, causes irritation to the skin, the eyes and the respiratory tract. Handle using appropriate safety equipment. ▲ CRITICAL This reagent is required for carrying purification using the QIAGEN kit.
- 1 M Tris, pH 8.0 (100 ml) (Ambion, cat. no. AM9855G)
- Nuclease-free water (1 liter) (Ambion, cat. no. AM9932)
- 10× NEB buffer 2 (New England Biolabs, cat. no. B7002S)
- Ethanol (Sigma-Aldrich, cat. no. E7023) ! CAUTION Flammable. Handle using appropriate safety equipment.
- Ethylene glycol (American Bioanalytical, cat. no. AB00455-01000)
- End-It DNA end-repair kit (Epicentre, cat. no. ER0720)
- 0.5 M EDTA, pH 8.0 (Ambion, cat. no. AM9260G)
- DNA polymerase I (*E. coli*), (10 U μl^{-1}) (New England Biolabs, cat. no. M0209L)
- Quick ligation kit (New England Biolabs, cat. no. M2200L)
- SYBR GreenER qPCR SuperMix Universal (Invitrogen, cat. no. 11762-100) or SYBR Green PCR master mix (Applied Biosystems, cat. no. 4309155) ! CAUTION Harmful and combustable. It is readily absorbed through skin and may cause eye, skin and respiratory irritation. Handle using appropriate safety equipment and keep away from heat sources as well as open flames.
- Agilent DNA 1,000 kit (Agilent Technologies, cat. no. 5067-1504)
- Agencourt AMPure 60 ml kit (Agencourt, cat. no. 000130)
- AmpliTaq DNA polymerase, LD with buffer I (Applied Biosystems, cat. no. N8080157)
- Cloned Pfu polymerase (2.5 U μl^{-1}) (Stratagene, cat. no. 600153)
- Invitrogen Platinum PCR SuperMix (Invitrogen, cat. no. 11306-016)
- Quant-iT dsDNA HS assay kit (Invitrogen, cat. no. Q32854)
- SOLiD fragment library oligos kit (includes P1 and P2 adaptors, P1 and P2 library PCR primers) (Applied Biosystems, cat. no. 4401151)
- 4% NuSieve 3:1 plus agarose TBE gel (Lonza, cat. no. 54927)
- 2% E-Gel (Invitrogen, cat. no. G4010-02)
- 2.2% FlashGel DNA cassette (Lonza, cat. no. 57031)

EQUIPMENT

- PCR tubes, 0.5 ml thin-wall (Applied Biosystems, cat. no. N8010180)
- DNA LoBind Tubes, 1.5 ml (Eppendorf, cat. no. 022431021)
- Brown-flaming micropipette puller (Sutter Instrument, P-80)
- Covaris S2 System (Covaris, cat. no. 4387833) (110 V)
- Microcentrifuge (5417R, refrigerated, without rotor 120 V/60 Hz; Eppendorf, cat. no. 022621807)
- PCR system (96-Well GeneAmp 9700; Applied Biosystems, cat. no. N8050200)
- Real-time PCR system with 96-well block (ABI PRISM 7500, 7900HT real-time PCR systems; Applied Biosystems, cat. nos. 4366605 and 4329001)
- Spectrophotometer (computer required) (NanoDrop ND1000; Thermo Scientific, cat. no. ND-1000)
- Rotating tube mixer (Barnstead/Thermolyne Labquake Rotisserie rotator; VWR, cat. no. 56264–312)
- Qubit fluorometer (Invitrogen, cat. no. Q32857)
- 6-tube Magnetic stand (Applied Biosystems, cat. no. AM10055)
- Agilent 2100 bioanalyzer (Agilent Technologies, cat. no. G2938C)
- Covaris microTUBE with AFA fiber and Snap-Cap with pre-slit Teflon/silicone/Teflon septa (Covaris, cat. no. 520045)
- Heating block (VWR, cat no. 13259)
- FlashGel Dock system (Lonza, cat. no. 57025)
- E-Gel iBase power system (Invitrogen, cat. no. G6400)

REAGENT SETUP

PBS–BSA Dissolve Ac-BSA (20 mg ml^{-1}) in PBS at 1 mg ml^{-1} and aliquot into 1.5 ml Eppendorf tubes. The solution can be stored at −20 °C for at least 6 months.

PROCEDURE

▲ CRITICAL STEP Use a thermal cycler with heated lid for all incubations throughout this protocol.

Single-cell lysis ● TIMING 3h

1| Dilute the UP1 Primer to 0.5 μM by adding 1 μl of 100 μM UP1 Primer and 199 μl of nuclease-free water to a tube and mix well. (All primer sequences are listed in **Table 1**).

2| Prepare cell lysis buffer (4.45 μl per sample) in a 0.5 ml thin-wall PCR tube by combining and mixing the following components:

Component	Final concentration	Volume (μl) (1×)
10× PCR buffer II (without $MgCl_2$)	0.9×	0.45
25 mM $MgCl_2$	1.35 mM	0.27
10% NP40	0.45%	0.225
0.1 M DTT	4.5 mM	0.225
SUPERase-In (20 U μl^{-1})	0.18 U μl^{-1}	0.045
RNase inhibitor (40 U μl^{-1})	0.36 U μl^{-1}	0.045
0.5 μM UP1 Primer	12.5 nM	0.125
dNTP mix (2.5 mM each)	0.045 mM (each)	0.09
Nuclease-free water	—	2.975
Total volume	—	4.45

! CAUTION Nonidet P-40 is harmful if swallowed or inhaled and is irritating to the skin, the eyes and the respiratory tract. Handle using appropriate safety equipment.

3| Isolate four-cell stage embryos from the oviduct (see ref. 45 for details) and transfer by mouth pipette to a drop of acidic Tyrode's solution to remove the zona pellucida.
▲ CRITICAL STEP All experiments that use human and animals should comply with the institutional and national guidelines.
▲ CRITICAL STEP In all transfer steps, ensure that no bubbles are formed in the drops by the micropipette before transferring a single cell into the lysate buffer.

4| Transfer the embryos into calcium-free medium and gently pipette until the individual blastomeres separate.

5| Sequentially transfer the blastomeres into three drops (typical volume per drop = 50–150 μl) of PBS–BSA to wash them. Then, transfer blastomeres into a final drop of PBS–BSA for picking.

6| Seed each single cell (with PBS–BSA carryover) into a 0.5 ml thin-wall PCR tube containing 4.45 μl of freshly prepared cell lysis buffer. Suck a small volume of PBS–BSA using the mouth pipette, then gently suck the single cell into the micropipette so that the cell is inside the micropipette but still near the tip. When pushing the cell out of the micropipette into the lysis buffer, eject all carryover PBS–BSA as well.
▲ CRITICAL STEP Each micropipette can be used only once for transferring each individual cell into lysis buffer. Never use the same micropipette repeatedly for transferring multiple cells into lysis buffer.
▲ CRITICAL STEP It is essential to include a negative control sample, in which no individual cell is picked; instead, transfer PBS–BSA solution equivalent to the carryover volume collected during single-cell transfer into lysis buffer. This negative control will enable detection of contamination in subsequent steps of the procedure.
? TROUBLESHOOTING

7| Centrifuge samples for 30 s at 7,500*g* at 4 °C and put on ice immediately.

8| Incubate at 70 °C for 90 s, then immediately transfer to ice.

9| Centrifuge tubes for 30 s at 7,500*g* at 4 °C, then immediately transfer to ice for 1 min. After this step, all mRNAs will be released from the single cells.

Reverse transcription ● TIMING 2 h
10| Prepare and add 0.45 μl RT mix (recipe below) to each tube, so that total volume of each RT reaction is 5 μl per tube: 4.45 μl lysis buffer, ~0.1 μl PBS–BSA carryover and 0.45 μl RT mix.

Component	Final concentration (U μl^{-1})	1× Volume (μl)
SuperScript III reverse transcriptase (200 U μl^{-1})	13.2	0.33
RNase inhibitor (40 U μl^{-1})	0.4	0.05
T4 gene 32 protein (1–10 U μl^{-1})	0.07 U	0.07
Total volume	—	0.45

11| Incubate at 50 °C for 30 min.

12| Inactivate the reverse transcriptase at 70 °C for 15 min.

13| Centrifuge tubes for 30 s at 7,500*g* at 4 °C and immediately place on ice for 1 min. After this step, first strand cDNAs for all mRNAs are synthesized.

Free primer removal ● TIMING 2 h

14| Prepare the exonuclease I mix by combining and mixing the following components, and add 1.0 μl of mix to each reaction:

Component	Final concentration	1× Volume (μl)
10× Exonuclease I buffer	1×	0.1
Nuclease-free water	—	0.8
Exonuclease I (5 U μl^{-1})	0.5 U μl^{-1}	0.1
Total volume	—	1

15| Incubate at 37 °C for 30 min.

16| Inactivate the exonuclease I at 80 °C for 25 min.

17| Centrifuge the tubes for 30 sec at 7,500*g* at 4 °C and immediately place on ice for 1 min. After this step, all free UP1 primers are destroyed, leaving 5′ cDNA ends (UP1 sequence) intact.

3′ Poly(A) tailing ● TIMING 1h

18| Prepare TdT reaction mix by combining and mixing the following components, and add 6.0 μl of mixture to each reaction:

Component	Final concentration	1× Volume (μl)
10× PCR buffer II (without $MgCl_2$)	1×	0.6
25 mM $MgCl_2$	1.5 mM	0.36
100 mM dATP	3 mM	0.18
Nuclease-free water	—	4.26
Terminal transferase (15 U μl^{-1})	0.75 U μl^{-1}	0.3
RNase H (2 U μl^{-1})	0.1 U μl^{-1}	0.3
Total volume	—	6

19| Incubate at 37 °C for 15 min.

20| Inactivate the TdT at 70 °C for 10 min.

21| Centrifuge tubes for 30 s at 7,500*g* at 4 °C and immediately place on ice for 1 min. After this step, the 3′end of the first-stranded cDNAs has a poly(dA) tail.

Second strand synthesis ● TIMING 1 h

22| Prepare 76 μl PCR mix 1 for each reaction by combining and mixing the following components:

Component	Final concentration	1× Volume (μl)
10× Ex Taq buffer (with $MgCl_2$)	1×	7.6
dNTP mix (2.5 mM each)	0.25 mM	7.6
UP2 primer (100 μM)	1 μM	0.76
Nuclease-free water	—	59.28
TaKaRa Ex Taq HS (5 U μl^{-1})	0.05 U μl^{-1}	0.76
Total volume	—	76

23| Divide the poly(dA) tailed RT product into four empty 0.2 ml thin-wall PCR tubes (3 μl for each tube).

24| Add 19 µl PCR mix 1 to each tube.

25| Carry out the following PCR program in the thermal cycler: one cycle of 95 °C for 3 min, 50 °C for 2 min and 72 °C for 10 min.

26| Put tubes on ice for 1 min.

27| Centrifuge tubes for 30 s at 7,500*g* at 4 °C and immediately place on ice. After this step, the second-strand cDNAs are 5′-UP2-$(T)_n$-cDNA-$(A)_n$-UP1-3′.

PCR amplification ● **TIMING 5 h**

28| Prepare 76 µl PCR Mix 2 for each reaction by combining and mixing the following components:

Component	Final concentration	1× Volume (µl)
10× Ex Taq buffer (with $MgCl_2$)	1×	7.6
dNTP mix (2.5 mM each)	0.25 mM (each)	7.6
100 µM UP1 primer	1 µM	0.76
Nuclease-free water	—	59.28
TaKaRa Ex Taq HS (5 U μl^{-1})	0.05 U μl^{-1}	0.76
Total volume	—	76

29| Add 19 µl PCR mix 2 to each tube.

30| Carry out the following PCR program: 95 °C for 3 min, then 20 cycles of 95 °C for 30 sec, 67 °C for 1 min and 72 °C for 6 min (+ 6 s each cycle), with a 4 °C hold. After this step, all cDNAs have been amplified. Save the PCR product at –80 °C or transfer to ice if continuing to Step 31 immediately.

■ **PAUSE POINT** The single-cell cDNA PCR product can be saved at –80 °C for 6 months.

▲ **CRITICAL STEP** 1st cycle = 6 min, 2nd cycle = 6 min and 6 s, 3rd cycle = 6 minutes and 12 seconds and so on.

DNA purification ● **TIMING 3–7 h**

31| Combine the four PCR reactions for each individual sample.

32| Take 10 µl from each combined single-cell PCR product and combine with 90 µl nuclease-free water to dilute ten-fold. Use 1 µl or 2 µl of this diluted PCR product as a template for a 20 µl SYBR Green real-time PCR reaction to check the expression of housekeeping genes (such as GAPDH, HPRT and ACTB).

! CAUTION SYBR Green master mix is harmful and combustible. It is readily absorbed through the skin and may cause eye, skin and respiratory irritation. Handle using appropriate safety equipment and keep away from heat sources and open flames.

? TROUBLESHOOTING

33| Purify the remaining undiluted PCR product by QIAquick PCR purification kit and elute with 50 µl EB buffer.

34| Store at −80 °C.

■ **PAUSE POINT** The purified PCR product can be saved at −80 °C for 6 months.

Second round of PCR

35| For each sample, setup four 90 µl PCR reactions as follows:

Component	Final concentration	1× Volume (µl)
Purified first-round PCR products	—	1.2
10× Ex Taq buffer	1×	9
dNTP mix (2.5 mM each)	0.25 mM (each)	9
Amine-blocked UP1 (AUP1) (100 µM)	1 µM	0.9
Amine-blocked UP2 (AUP2) (100 µM)	1 µM	0.9
Nuclease-free water	—	68.1
TaKaRa Ex Taq HS (5 U μl^{-1})	0.05 U μl^{-1}	0.9
Total volume	—	90

36| Run the following PCR program: 95 °C for 3 min, then 9 cycles of 95 °C for 30 s, 67 °C for 1 min and 72 °C for 6 min (+ 6 s each cycle) with a 4 °C hold. Combine the four 90 µl PCR reactions from each sample and save at −80 °C.
■ **PAUSE POINT** The second round PCR product can be saved at −80 °C for 6 months.

Quality check ● TIMING 1 h
37| *Optional*: Perform QC on PCR products by running 20 µl of each sample on a 2% E-Gel.

Purification of PCR production ● TIMING 2 h
38| Purify amplified cDNAs using QIAquick PCR purification kit. Elute in 30 µl EB buffer and quantify by Nanodrop. The total amount of cDNAs should be more than 200 ng.

39| Store at −80 °C.
■ **PAUSE POINT** The purified PCR product can be stored at −80 °C for 6 months.

Gel purification ● TIMING 6–7 h
40| Load cDNA on a 1% agarose gel alongside an appropriate size standard and run the gel until the bromophenol blue dye marker is 2 cm from the well. Excise cDNA from the size range 0.5–3 kb and purify by QIAquick gel extraction kit. Elute the cDNA in 30 µl EB buffer and quantify using Quant-it HS dsDNA kit.

41| Store at −80 °C.
■ **PAUSE POINT** The gel purified PCR product can be saved at −80 °C for 6 months.
▲ **CRITICAL STEP** The cDNA may be weakly stained if the primer dimer band saturates the ethidium bromide. If this happens, simply cut the gel between 0.5 kb and 3 kb and proceed to gel extraction; alternately, the operator can stain the gel with ethidium bromide to see the cDNA smear.
▲ **CRITICAL STEP** Do not over-stain or over-irradiate DNA; long exposure of ethidium bromide stained DNA to UV light causes DNA damage, including strand breaks and base modifications, which affect the efficiency of polymerase-mediated reactions. SYBR safe gel staining and blue light transillumination can be used instead of ethidium bromide and UV light.
? TROUBLESHOOTING

Quality check ● TIMING 1.5 h
42| Check quality of the gel-purified cDNA by running 50–100 ng of the purified cDNA on a 2% agarose gel (see **Fig. 3**).

Library fragment preparation ● TIMING 30–60 min
43| In borosilicate glass microTUBE, dilute about 200 ng of each DNA sample into a final volume of 100 µl of 10 mM Tris, pH 7.5.

44| Shear the DNA using the Covaris S2 shearing program described below:

No. of cycles: 15
Bath temperature: 5 °C
Bath temperature limit: 30 °C
Mode: frequency sweeping
Water quality testing function: off

Treatment 1:
Duty cycle: 20%
Intensity: 5
Cycles/burst: 200
Time: 60 s

45| Transfer the sheared DNA into a clean 1.5 ml LoBind tube.

46| For QC, 2 µl of the reaction should be saved and used to confirm that shearing has created the desired range of fragment sizes using the Agilent 2100 Bioanalyzer.

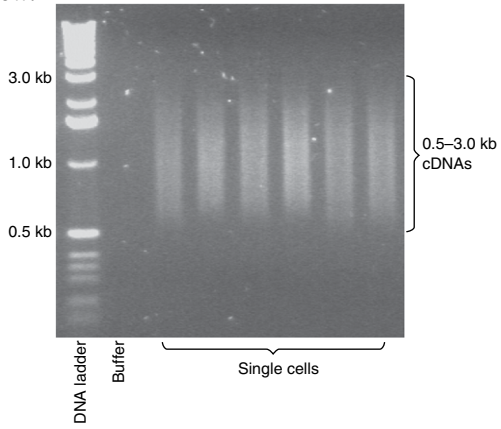

Figure 3 | PCR amplified single-cell cDNA libraries. Typical agarose gel electrophoresis for single-cell cDNAs after 20 + 9 cycles of PCR post-gel purification. From left to right, lane #1 is the 1-kb DNA ladder (the size of three representative bands is labeled at the left side of the figure); lane #2 is a negative control (picking buffer only) showing no contamination; lane #3–#8 are individual cell-derived libraries. The cDNA products are evenly distributed as a smear between 0.5 kb and 3.0 kb.

Repair DNA ends ● TIMING 1 h

47| Add the following to the sheared DNA:

Component	Volume (μl)
Sheared DNA	98
10× End-it buffer	14
End-it ATP (10 mM)	7
End-it dNTPs (2.5 mM)	14
End-it enzyme mix	2
Nuclease-free water	5
Total volume	140

48| Incubate at room temperature (20–25 °C) for 30 min.

49| Inactivate End-it Enzymes by transferring the tube to heating block and incubating for 20 min at 65 °C. Centrifuge the tube for 5 sec at 7,500*g* at 4 °C.

Ligate library adaptors to DNA ● TIMING 20 min

50| (Note: for multiplex adaptor preparation, see **Box 1**) Thaw P1 and P2 adaptors on ice or at room temperature.

51| Use the information below to calculate the pmoles of adaptor needed for the reaction. Average insert fragment size is the size that the sample was sonicated to; if the provided sonication parameters were used, the average should be 105 bp.

1 μg DNA × 10^6 pg $μg^{-1}$ × 1 pmol 660pg^{-1} × 1/(average insert fragment size) = X pmoles
(# μg DNA) × (X pmoles DNA) = (# pmoles DNA for adaptor ligation)
(# pmoles DNA for adaptor ligation) × (30) = (# pmoles adaptors needed)
(# pmoles adaptors needed) / (# pmoles $μl^{-1}$ stock) = (# μl adaptor needed)

52| In a tube containing the end-repaired DNA from the previous step, combine and mix the following components. For DNA inputs higher than 2 μg, scale up the total volume of the reaction and amount of components (buffer, adaptors, ATP and ligase) proportionally. Add 1 μl of Quick Ligase per 20 μl of reaction volume. Use one adaptor set per sample.

Component	Volume (μl)	i.e., Volume for 2 μg input DNA (μl)
P1 (ds) Adaptor (500 pmol $μl^{-1}$)	As determined above	1.8
P2 (ds) Adaptor (500 pmol $μl^{-1}$)	As determined above	1.8
10× End-it buffer	2	2
End-it ATP (10 mM)	2	2
Quick ligase enzyme	8	8
End-repaired DNA	140	140
Nuclease-free water	Variable	4.4
Total volume	160	160

53| Incubate at room temperature for 10–15 min.

Purification of ligated DNA ● TIMING 1h

54| Add 1.8× volumes of Agencourt AMPure beads to the sample (e.g., 288 μl of beads for 160 μl) and incubate for 5 min at room temperature on a rotator.

55| Prepare 70% ethanol (~650 μl per sample).

! CAUTION Ethanol is flammable. Handle using appropriate safety equipment.

▲ CRITICAL STEP Using precisely 70% freshly prepared ethanol is critical, as a higher percentage will result in inefficient washing of smaller-sized molecules, whereas <70% ethanol could cause loss of sample.

56| Place the tube of beads in the magnetic rack to separate beads from solution. Wait for the solution to clear before proceeding to the next step.

57| Remove the supernatant and discard.

58| Dispense 200 µl of freshly prepared 70% ethanol, vortex the tube thoroughly and incubate for 30 s at room temperature.

59| Place the tube of beads in the magnetic rack to separate beads from solution. Wait for the solution to clear before proceeding to the next step.

60| Aspirate out the ethanol and discard.

61| Repeat Steps 58–60 two more times.

62| Place the tube of beads in the magnetic rack and remove the supernatant and pulse-spin to remove the residual ethanol.

63| Repeat Step 62 for two to three more times to remove the residual ethanol.

64| Dry the beads at room temperature for ~5 min.

65| Elute the DNA by adding 36 µl 10 mM Tris pH 8, vortexing for 10 s and ensuring homogeneity by pipetting the solution up and down for several times.

66| Place the tube of beads in the magnetic rack to separate beads from solution. Wait for the solution to clear before proceeding to the next step.

67| Transfer the eluant to a 1.5 ml LoBind tube.

68| Place the eluted sample in the magnetic rack again to separate any remaining beads from solution. Wait for the solution to clear before proceeding to the next step.

69| Transfer the eluted sample in a new 1.5 ml LoBind tube.

70| Repeat Steps 68–69 once.

71| Store DNA in a 1.5 ml LoBind tube.

72| Save 2 µl of the AMPure-purified ligated library DNA. The aliquot can be used for running a library QC for troubleshooting purposes.

■ **PAUSE POINT** The sample(s) may be stored at −20 °C overnight or upto 6 months.

Quantification of multiplexed samples ● TIMING 1 h

73| *Optional*: **Quantitation of multiplexed samples.** Using Qubit (following the manufacturer's protocol) or qPCR (see Steps 79–84) determine the concentration of each of the multiplexed samples.

Pooling multiplexed samples ● TIMING 1 h

74| *Optional*: **Pooling multiplexed samples.** Pool equal concentrations of the samples and purify with AMPure beads as above, eluting with 34 µl of 10 mM Tris, pH 8 solution and continue with nick translation.

▲ **CRITICAL STEP** For fragment libraries or multiplexed libraries that are not yet pooled, the operator may skip quantitation and go directly on to nick translation; pooling before nick translation is primarily to conserve reagents.

Nick-translation of DNA ● TIMING 1.5 h

75| Combine and mix the following components in a LoBind tube:

Component	Volume (µl)
P1 (ds) and P2 (ds) adapter ligated fragment DNA	34
10× NEBuffer 2	4
GeneAmp dNTP Blend (100 mM)	0.8
DNA Polymerase I (10 U μl^{-1})	1.0
Nuclease-free water	Variable
Total volume	40

76| Incubate at 16 °C for 30 min.

77| Stop the reaction by adding 0.5 µl 0.5 M EDTA pH 8.0.

78| Purify the nick-translated DNA with the Agencourt AMPure Kit as before (see Steps 54–72), eluting with 40 µl of 10 mM Tris, pH 8.0 solution.
▲ **CRITICAL STEP** Save 2 µl of the AMPure-purified ligated library DNA. The aliquot can be used for running a library QC for troubleshooting purposes.
■ **PAUSE POINT** The sample(s) may be stored at −20 °C overnight or upto 6 months.

Quantification of library by qPCR ● TIMING 2 h
79| Dilute EB-eluted library 1:10 in H_2O (final volume 30 µl) and use this solution to prepare further serial ten-fold dilutions ranging from 10^{-2} to 10^{-6} (final volume 50 µl each).

80| Prepare samples containing 390 fg, 39 fg, 3.9 fg, 0.39 fg and 0.039 fg of Standard Library (*E. coli* DH10B fragment library, 150 bp mean fragment size) in 5 µl H_2O. These values should be used for absolute quantification when programming the 7900HT fast real-time PCR system to generate a standard curve. For each reaction, prepare Pre-Mix:

	Final concentration	1× Volume (µl)
2× SYBR Green master mix (ABI)	1×	15
Library PCR primer 1 (50 µM)	0.9 µM	0.54
Library PCR primer 2 (50 µM)	0.9 µM	0.54
H_2O		8.92
Total volume		25

▲ **CRITICAL STEP** If the sample is multiplexed, use 16Barcode Library PCR primers 1 and 2 for qPCR. If the sample is a fragment library, use Fragment Library PCR primers 1 and 2.

81| Distribute 25 µl Pre-Mix in corresponding wells. Then, add 5 µl of the serially diluted template DNA (or H_2O for negative control). Two serial dilutions per sample are adequate for quantitation. Seal the plate with adhesive optical cover.

82| When carrying out absolute quantification with 7900HT fast real-time PCR system, set up the following thermal profile:

95 °C	2 min	
95 °C	15 s	} 40 cycles
62 °C	15 s	
70 °C	1 min	

Also set:
Thermal cycler protocol mode: standard
Data Collection: at 70 °C (extension step)

83| Run qPCR and quantify the amount of library DNA/number of P1-insert-P2 molecules. Typical standard curve and amplification plots for standards and Library samples are shown in **Figure 4**.

84| If the sample concentration is adequate (above 1 ng μl^{-1}), the sample is then ready for emulsion PCR. Otherwise continue with PCR amplification, Steps 85–89.

(*Optional*) PCR amplification ● TIMING 4 h
85| Amplify 25 µl of sample using reaction conditions described below. For a fragment library, use the Fragment Library PCR primers 1 and 2 provided in the Fragment Library oligo kit; for a multiplexed library, use the 16 Barcode Library PCR primers 1 and 2.

Component	Volume (μl)
Sample	25
Invitrogen SuperMix	75
Library PCR Primer 1	1.5
Library PCR Primer 2	1.5
Pfu polymerase	0.25
AmpliTaq	0.2
Total volume	103.45

Temperature (°C)	Time	Condition/note
95	5 min	Hold
95	15 s	8–16 cycles
62	15 s	
70	1 min	
70	5 min	Hold
4	—	Hold

▲ **CRITICAL STEP** The number of cycles should be decided based on the amount of starting material used for shearing. Minimal cycling is desirable for avoiding over-amplification. Use the table below as a guideline to determine the number of PCR cycles based on the amount of input DNA:

Starting amount of DNA	Number of PCR cycles
1–2 μg	8–10
100 ng–1 μg	10–12
20–100 ng	13–16

86| To determine if amplification was adequate, load and run 4 μl of sample on a 2.2% Lonza FlashGel cassette for 6 min at 275 V to ensure amplification after a minimal number of cycles as described above.
? TROUBLESHOOTING

87| If fairly robust amplification products are visible, proceed with Step 88. If little or no amplification products are observed at this point, return the tubes to the thermal cycler and run the PCR cycling program below:

Temperature (°C)	Time	Condition
95	5 min	Hold
95	15 s	2–3 cycles
62	15 s	
70	1 min	
70	5 min	Hold
4	—	Hold

88| Repeat AMPure purification as in Steps 54–72.

89| Quantify sample(s) with either Qubit quantitation or qPCR.

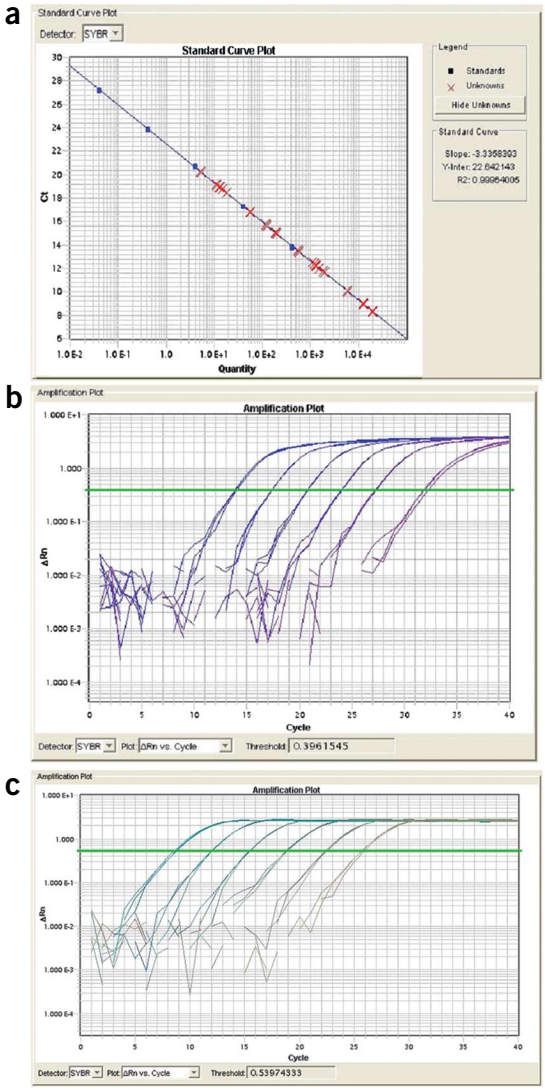

Figure 4 | Standard curve of cDNAs following qPCR quantification. (**a**) Standard curve. (**b**) Standard plot. (**c**) Library plot. Quantitative PCR was carried out using real-time SYBR Green assay, with the Standard Library used to generate a standard curve. In all, 5 μl of each library sample and standard were used per PCR reaction; the no-template control uses H_2O instead of template. All samples were prepared in triplicate as follows. (1) Prepare serial tenfold dilutions of the library eluted in EB at the previous step. (2) Prepare standard samples containing 390, 39, 3.9, 0.39 and 0.039 fg of Standard Library (*E. coli* DH10B fragment library, 150 bp mean fragment size) in 5 μl H_2O. These values should be used for absolute quantification when programming PCR instruments to generate a standard curve. (3) Assay is carried out in 96-well plates in 30 μl final volume with an AB7500 real-time PCR instrument.

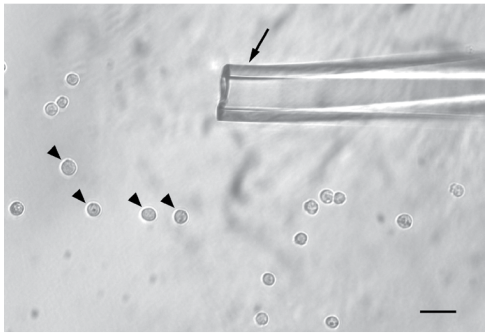

Figure 5 | Single mouse embryonic stem (ES) cells in a PBS–BSA drop. Black arrowheads indicate individual ES cells. Black arrow indicates the tip of micropipette for picking single cells. The inner diameter of the micropipette is about two- to threefold greater than the diameter of an individual ES cell. Scale bar: 30 μm.

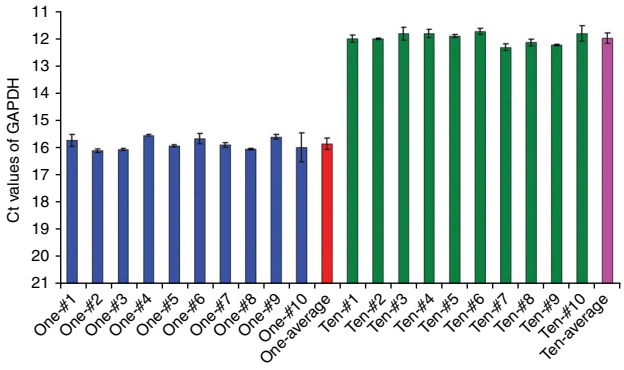

Figure 6 | GAPDH expression in mouse embryonic stem (ES) cells measured by real-time PCR. One or ten ES cells were picked into each tube to amplify cDNAs by 20 cycles of PCR. The PCR products were then diluted ten-fold (take 20 μl out of 164 μl PCR product to dilute to 200 μl), with 2 μl of the diluted PCR product used as template for a 20 μl real-time PCR reaction. The error bar represents standard deviation of technical duplicates of real-time PCR.

● TIMING

Day 1
Steps 1–9, Single-cell lysis: 3 h
Steps 10–13, Reverse transcription: 2 h
Steps 14–17, Free primer removal: 2 h
Steps 18–21, 3′ Poly(A) tailing: 1 h
Steps 22–27, Second-strand cDNA synthesis: 1 h
Steps 28–30, PCR amplification: 5 h

Day 2
Steps 31–34, DNA purification: 3–7 h

Day 3
Steps 35 and 36, Second round of PCR: 3 h
Step 37, (*optional*) Quality check: 1 h
Steps 38 and 39, Purification of PCR product: 2 h

Day 4
Steps 40 and 41, Gel purification: 5 h
Step 42, Quality check: 1.5 h

Day 5
Steps 43–46, Library fragment preparation: 30–60 min, if bioanalyzer is used
Steps 47–49, Repair DNA ends: 1 h
Box 1 (*optional*), Hybridization of oligonucleotides for multiplexing: 30 min
Steps 50–53, Ligate library adaptors to DNA: 20 min
Steps 54–72, Purification of ligated DNA: 1 h (time depends on sample numbers; for 16 samples, ~1 h)
Step 73, (*optional*) Quantification of multiplexed samples: 1 h
Step 74, (*optional*) Pooling multiplexed samples: 1 h
Steps 75–77, Nick-translation of DNA: 1.5 h
Step 78, Purify nick-translated DNA: 1 h (time depends on sample numbers; for 16 samples, ~1 h)
Steps 79–84, Quantification of library by qPCR: 2 h

Day 6
Steps 85–89, (*optional*) PCR amplification: 4 h

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2**.

TABLE 2 | Troubleshooting table.

Step	Problem	Possible reason	Solution
6	Cell not transferred into lysis tube	Cell stuck to the inside wall of the glass capillary (micropipette)	First, suck a small volume of PBS–BSA using the mouth pipette, then suck the single cell into the glass capillary gently to keep the cell already inside the capillary but still near the tip of the capillary. When pushing the cell out of the capillary into the lysis, push all the carryover PBS–BSA until a bubble is visible from the capillary into the lysis buffer
32	Picking buffer, only control, shows positive signal when using real-time PCR to check expression of housekeeping genes	PBS–BSA drop is contaminated by lysed cells	Make sure the PBS–BSA drop for holding single cells is not contaminated by lysed cells Wash the single cells through several PBS–BSA drops before picking them. Aliquot all reagents for lysis, RT, cutting, tailing and PCR steps in small batches Ensure that each aliquot is only used once and the remainder is discarded. Only load the exact number of single cells that will be picked into the final PBS–BSA drop. After picking each of them, check the number of remaining single cells to make sure every cell is correctly picked and no cell is by chance lysed or incorrectly picked
	The expression level of housekeeping genes is much lower than expected in some single-cell cDNA samples	Some dead or partially damaged single cells were picked	Before picking, try to check the quality of the single cells by either morphology or trypan blue staining. Only pick healthy single cells
		RNase contamination during the lysate and reverse transcription step	Keep the bench and surrounding area very clean Wear a dust mask to avoid breathing contaminants into reaction reagents Change gloves regularly during these steps Try to do all manipulations in a clean hood
		Loss of activity of some enzymes and reagents	Some reagents are not very stable, such as dATP, dNTPs or primers at low concentrations—aliquot them into small batches. Avoid repeated freeze–thaws. Make sure reagents are not expired
41	Low recovery of gel purification	Gel fragment recovered more than the kit requested	Try to cut the gel fragment containing the cDNA smear as accurately as possible. If too much gel is already cut, divide it between two gel purification columns and combine the purified cDNAs. Each QIAGEN column can hold upto 400 mg of gel
86	Primer dimers present on Lonza gel	Too many cycles of PCR were done or the amount of adaptors added was too large relative to sample size	Gel purification must be done to remove primer dimers. SOLiD Fragment Library protocol recommends the use of 3% or 4% agarose gel run in 1× TAE, using a 25 bp Track-It ladder as reference. Cut the gel between 125 bp and 200 bp and purify using QIAGEN MinElute gel extraction kit following the manufacturer's instructions
	Size distribution of fragment after shearing is not in the desired range	Sonication protocol was not calibrated correctly for the sample	Readjust shearing protocol and redo shearing. Make sure the Covaris S2 System is properly setup
	Low yield after AMPure purification	Ethanol was not freshly prepared and/or was not at 70%	If there is adequate sample remaining, PCR can be done to increase the quantity, see Steps 85–87. If not, protocol must be redone, using freshly prepared 70% ethanol
	Overall yield is low	Sample quality was poor	Confirm quality of sample before beginning Express protocol using a housekeeping gene assay or use Qubit to confirm that the sample's concentration is adequate for input
		Sample loss occurred during AMPure purification	See above. Check QC aliquots using Qubit or NanoDrop to determine where has loss occurred
			Loss of sample can occur during steps when sample is transferred or purified. Take care to transfer the entire volume during purification steps and minimize transferring of samples. Sample input may need to be increased if initial input was low and loss occurs

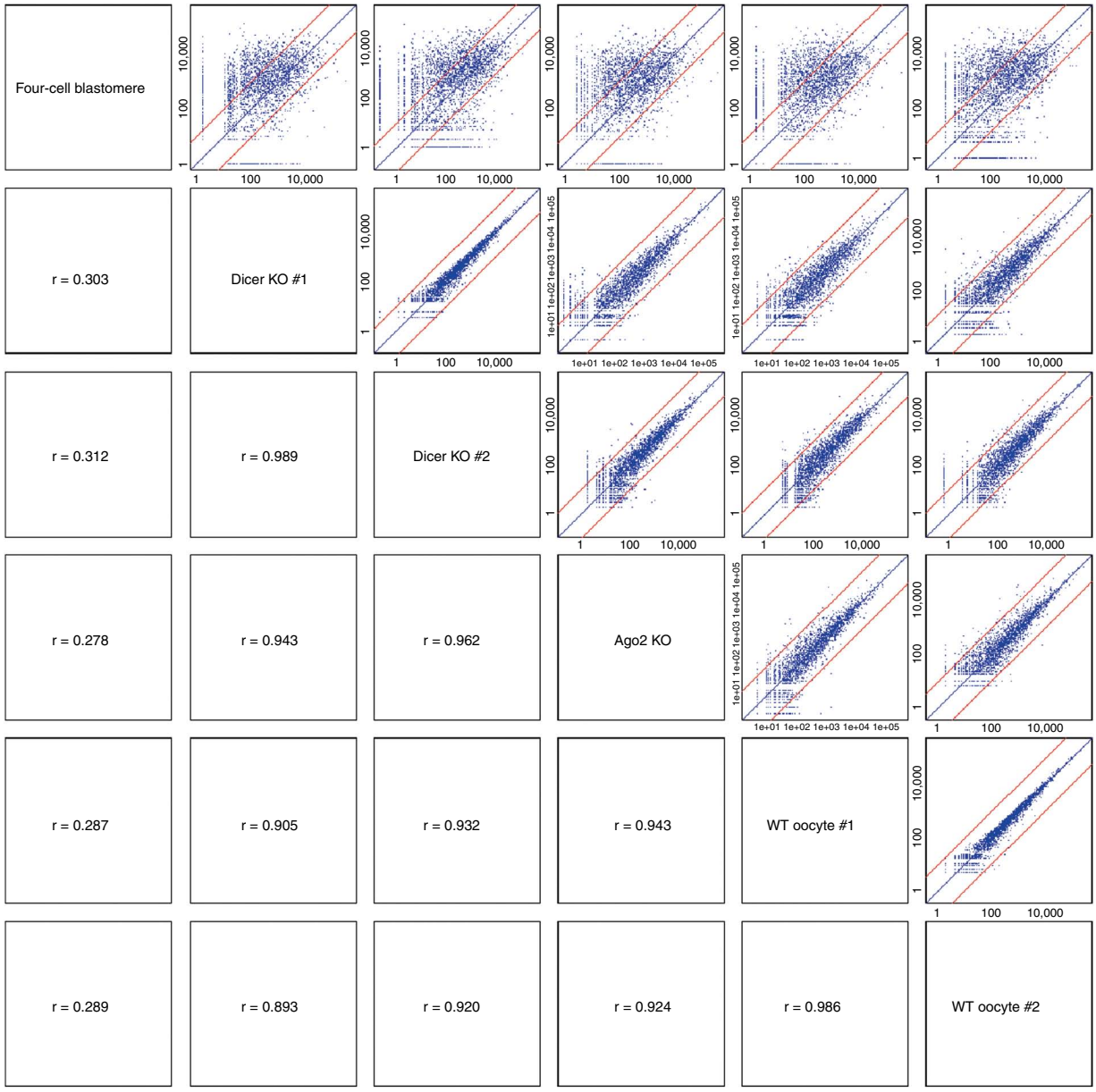

Figure 7 | Assessing accuracy and reproducibility of the single-cell RNA-Seq method. Pearson's coefficient plots for single-cell RNA-Seq of one blastomere of a four-cell stage embryo, two wild-type mature oocytes, two Dicer-knockout mature oocytes and one Ago2-knockout mature oocyte. The reads mapped to RefSeq were used for the scattering plots. The values of correlation coefficients, *r*, are listed in the corresponding boxes and the red lines indicate >2- or <0.5-fold changes in expression. The labels of *x* axis and *y* axis are the log2 of the reads.

ANTICIPATED RESULTS

We first used relatively small cells—mouse embryonic stem (ES) cells, which only contain 10–20 pg of total RNA in each individual cell—to show that the method is sensitive and reproducible. We picked into each tube either one individual ES cell or ten ES cells together (**Fig. 5**) and then ran the single-cell cDNA amplification procedure and real-time PCR to analyze the expression of the housekeeping gene GAPDH (**Fig. 6**). The results show that variation in the measurement of the expression of GAPDH is reasonably minimal and the expression differences between one cell and ten cells are as expected.

We then confirmed the accuracy of the method by analyzing the correlation efficiency between two separately processed biological replicates. Based on our previous single-cell cDNA microarray results (unpublished) and results from other laboratories, undifferentiated ES cells have strong sub-populations with dramatic differences in gene expression[46–48], and so we analyzed biological replicates of mature oocytes, as we know that these represent a much more homogeneous population[42]. We can achieve correlation coefficients as high as 0.986 (**Fig. 7**), which is comparable to that achieved using

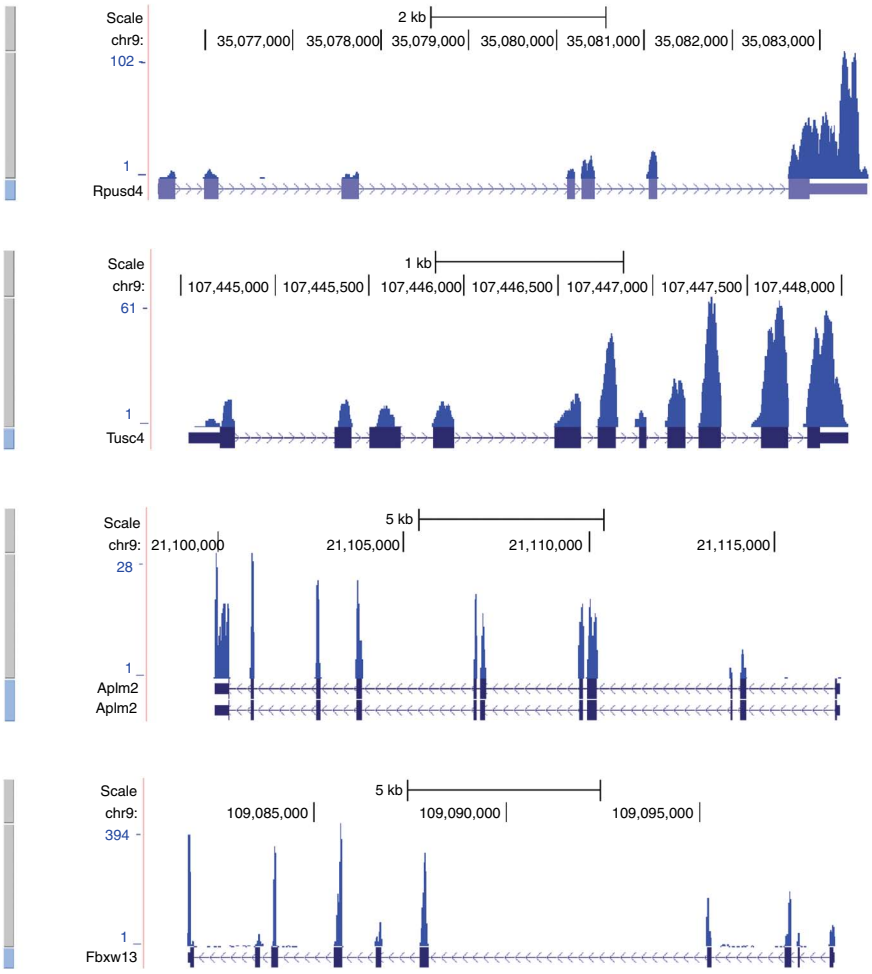

Figure 8 | Coverage plots of RNA-Seq reads from a single wild-type mature oocyte. Analysis was performed using the UCSC genome browser. Depicted here are the base coverage files for Rpusd4, Tusc4, Ap1m2 and Fbxw13 on chromosome 9.

bulk amount of cells, confirming the accuracy and reproducibility of our single-cell RNA-Seq analysis. The coverage plots[42] of several individual genes are shown in **Figure 8**, clearly showing the single-exon resolution of the RNA-Seq reads and that the reads can cover up to 3 kb of the 3'end of the cDNAs. The top 20 most abundant genes[42] in the mature oocyte and the blastomere of a four-cell stage embryo respectively are shown in **Table 3**. The majority of the most abundantly expressed genes in oocytes and blastomeres are not housekeeping genes, but rather developmental stage-specific genes, which show upto several-hundred-fold variability between these two types of cells.

TABLE 3 | Top 20 most abundant genes in the mature oocyte and the blastomere of four-cell stage embryos, respectively.

Gene name	RefSeq ID	Counts in mature oocyte	Counts in four-cell blastomere
H1foo	NM_138311	209,872	2,117
Bcl2l10	NM_013479	205,060	1,027
Spin1	NM_011462	200,333	23,757
Gdf9	NM_008110	165,482	14,100
Obox1	NM_027802	149,944	2,247
Oog1	NM_178657	136,086	100,086
Tcl1b2	NM_013775	131,326	76,552
Tcl1	NM_009337	129,237	22,122
Omt2b	NM_205822	122,678	258,088

(continued)

TABLE 3 | Continued.

Gene name	RefSeq ID	Counts in mature oocyte	Counts in four-cell blastomere
Tcl1b1	NM_013773	118,580	74,377
Omt2a	NM_001111286	113,360	243,955
Obox5	NM_145709	110,924	2,669
E330034G19Rik	NM_001033214	106,956	1,162
Zbed3	NM_028106	106,540	3,005
Slc45a3	NM_145977	98,658	14,813
Tcl1b5	NM_013776	95,977	68,387
Khdc1b	NM_001113187	95,938	23,715
Oosp1	NM_133353	93,718	14,502
Bpgm	NM_007563	86,249	4,611
EG194588	NM_001038676	84,998	2,979
Map1lc3b	NM_026160	984	279,836
Omt2b	NM_205822	12,2678	258,088
Omt2a	NM_001111286	113,360	243,955
Oaz1	NM_008753	30,967	165,822
Rplp0	NM_007475	1,844	151,774
Pdxk	NM_172134	0	126,122
Akp5	NM_007433	17	121,700
Ubb	NM_011664	21,871	118,549
Sp110	NM_175397	1,238	115,350
Klf17	NM_029416	24,531	112,781
H3f3b	NM_008211	47,128	112,188
Dppa3	NM_139218	29,844	106,509
Rps5	NM_009095	129	105,279
Tubb2c	NM_146116	3,013	101,600
Oog1	NM_178657	136,086	100,086
EG547109	NM_001034906	0	98,682
Serf2	NM_011354	17,828	95,696
Impdh2	NM_011830	27	90,365
Prps1	NM_021463	20	88,663
Ccnb1	NM_172301	35,328	87,846

ACKNOWLEDGMENTS We thank John Bodeau, Clarence Lee, Yangzhou Wang, Umberto Ulmanella, Karen Li, Cinna Monighetti and Swati Rande for their generous help.

COMPETING INTERESTS STATEMENT The authors declare competing financial interests (see the HTML version of this article for details). C.B., E.N., B.L., N.X., V.I.B. and K.L. are currently employees of Applied Biosystems (part of Life Technologies).

AUTHOR CONTRIBUTIONS K.L. designed the project; F.T. and E.N. carried out the experiments; C.B. analyzed the data; B.L., V.I.B. and N.X. improved the protocols; F.T., E.N., K.L. and M.A.S. wrote the manuscript with contributions from all the authors.

Published online at http://www.natureprotocols.com/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
2. Cloonan, N. & Grimmond, S.M. Transcriptome content and dynamics at single-nucleotide resolution. *Genome Biol.* **9**, 234 (2008).
3. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).
4. DeRisi, J. *et al.* Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **14**, 457–460 (1996).
5. Lockhart, D.J. *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**, 1675–1680 (1996).
6. Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J.M. Expression profiling using cDNA microarrays. *Nat. Genet.* **21**, 10–14 (1999).
7. Lipshutz, R.J., Fodor, S.P., Gingeras, T.R. & Lockhart, D.J. High density synthetic oligonucleotide arrays. *Nat. Genet.* **21**, 20–24 (1999).

8. Blackshaw, S. & Livesey, R. Applying genomics technologies to neural development. *Curr. Opin. Neurobiol.* **12**, 110–114 (2002).
9. Royce, T.E. *et al.* Issues in the analysis of oligonucleotide tiling microarrays for transcript mapping. *Trends Genet.* **21**, 466–475 (2005).
10. Preker, P. *et al.* RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851–1854 (2008).
11. Mardis, E.R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**, 133–141 (2008).
12. Wold, B. & Myers, R.M. Sequence census methods for functional genomics. *Nat. Methods* **5**, 19–21 (2008).
13. Schuster, S.C. Next-generation sequencing transforms today's biology. *Nat. Methods* **5**, 16–18 (2008).
14. Shendure, J. The beginning of the end for microarrays? *Nat. Methods* **5**, 585–587 (2008).
15. Wilhelm, B.T. *et al.* Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239–1243 (2008).
16. Nagalakshmi, U. *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349 (2008).
17. Sultan, M. *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956–960 (2008).
18. Wang, E.T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
19. Cloonan, N. *et al.* Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat. Methods* **5**, 613–619 (2008).
20. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
21. Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M. & Gilad, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**, 1509–1517 (2008).
22. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. & Blencowe, B.J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
23. Li, H. *et al.* Determination of tag density required for digital transcriptome analysis: application to an androgen-sensitive prostate cancer model. *Proc. Natl. Acad. Sci. USA.* **105**, 20179–20184 (2008).
24. Blake, W.J., Kærn, M., Cantor, C.R. & Collins, J.J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
25. Raser, J.M. & O'Shea, E.K. Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010–2013 (2005).
26. Arias, A.M. & Hayward, P. Filtering transcriptional noise during development: concepts and mechanisms. *Nat. Rev. Genet.* **7**, 34–44 (2006).
27. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).
28. Losick, R. & Desplan, C. Stochasticity and cell fate. *Science* **320**, 65–68 (2008).
29. Shahrezaei, V. & Swain, P.S. The stochastic nature of biochemical networks. *Curr. Opin. Biotechnol.* **19**, 369–374 (2008).
30. Kawasaki, E.S. Microarrays and the gene expression profile of a single cell. *Ann. N Y Acad. Sci.* **1020**, 92–100 (2004).
31. Livesey, F.J. Strategies for microarray analysis of limiting amounts of RNA. *Brief. Funct. Genomic Proteomic.* **2**, 31–36 (2003).
32. Hamatani, T., Carter, M.G., Sharov, A.A. & Ko, M.S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev. Cell* **6**, 117–131 (2004).
33. Wang, Q.T. *et al.* A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev. Cell* **6**, 133–144 (2004).
34. Zeng, F., Baldwin, D.A. & Schultz, R.M. Transcript profiling during preimplantation mouse development. *Dev. Biol.* **272**, 483–96 (2004).
35. Bontoux, N. *et al.* Integrating whole transcriptome assays on a lab-on-a-chip for single cell gene profiling. *Lab Chip* **8**, 443–450 (2008).
36. Iscove, N.N. *et al.* Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. *Nat. Biotechnol.* **20**, 940–943 (2002).
37. Klein, C.A. *et al.* Combined transcriptome and genome analysis of single micrometastatic cells. *Nat. Biotechnol.* **20**, 387–392 (2002).
38. Hartmann, C.H. & Klein, C.A. Gene expression profiling of single cells on large-scale oligonucleotide arrays. *Nucleic Acids Res.* **34**, e143 (2006).
39. Jensen, K.B. & Watt, F.M. Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence. *Proc. Natl. Acad. Sci. USA* **103**, 11958–11963 (2006).
40. Kurimoto, K. *et al.* An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res.* **34**, e42 (2006).
41. Kurimoto, K., Yabuta, Y., Ohinata, Y. & Saitou, M. Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis. *Nat Protoc.* **2**, 739–752 (2007).
42. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
43. Maekawa, M., Yamamoto, T., Kohno, M., Takeichi, M. & Nishida, E. Requirement for ERK MAP kinase in mouse preimplantation development. *Development* **134**, 2751–2759 (2007).
44. Tang, F. *et al.* 220-plex microRNA expression profile of a single cell. *Nat. Protoc.* **1**, 1154–1159 (2006).
45. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. Manipulating the mouse embryo. 3rd ed., 194–200 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 2003).
46. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
47. Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K. & Niwa, H. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**, 909–918 (2008).
48. Hayashi, K., Lopes, S.M., Tang, F. & Surani, M.A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401 (2008).