



## OPEN

## SUBJECT AREAS:

BIOMEDICAL  
ENGINEERING

GENE EXPRESSION PROFILING

LAB-ON-A-CHIP

# Co-detection and sequencing of genes and transcripts from the same single cells facilitated by a microfluidics platform

Lin Han<sup>1\*</sup>, Xiaoyuan Zi<sup>1,2,3\*</sup>, Lana X. Garmire<sup>4</sup>, Yu Wu<sup>1</sup>, Sherman M. Weissman<sup>2,5</sup>, Xinghua Pan<sup>2</sup> & Rong Fan<sup>1,5</sup>Received  
25 November 2013Accepted  
5 September 2014Published  
26 September 2014

Correspondence and requests for materials should be addressed to S.M.W. (Sherman.weissman@yale.edu); X.P. (xinghua.pan@yale.edu) or R.F. (rong.fan@yale.edu)

\* These authors contributed equally to this work.

<sup>1</sup>Department of Biomedical Engineering, Yale University, New Haven, CT 06520, USA, <sup>2</sup>Department of Genetics, Yale School of Medicine, New Haven, CT 06520, USA, <sup>3</sup>Department of Cell Biology, Second Military Medical University, Shanghai 200433, China, <sup>4</sup>Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813 USA, <sup>5</sup>Yale Comprehensive Cancer Center, New Haven, CT 06520, USA.

Despite the recent advance of single-cell gene expression analyses, co-measurement of both genomic and transcriptional signatures at the single-cell level has not been realized. However such analysis is necessary in order to accurately delineate how genetic information is transcribed, expressed, and regulated to give rise to an enormously diverse range of cell phenotypes. Here we report on a microfluidics-facilitated approach that allows for controlled separation of cytoplasmic and nuclear contents of a single cell followed by on-chip amplification of genomic DNA and cytoplasmic mRNA. When coupled with off-chip polymerase chain reaction, gel electrophoresis and Sanger sequencing, a panel of genes and transcripts from the same single cell can be co-detected and sequenced. This platform is potentially an enabling tool to permit multiple genomic measurements performed on the same single cells and opens new opportunities to tackle a range of fundamental biology questions including non-genetic cell-to-cell variability, epigenetic regulation, and stem cell fate control. It also helps address clinical challenges such as diagnosing intra-tumor heterogeneity and dissecting complex cellular immune responses.

The Central Dogma, also called a DNA-RNA-protein axis, describes how genetic information is transcribed to messenger RNAs (mRNAs) and expressed to produce proteins that form the building blocks of a living cell and fulfill all biological functions<sup>1,2</sup>. However, it turns out the correlation between genomic DNA variation, mRNA copy numbers, and the cognate protein levels is very poor, which is quite puzzling<sup>3-5</sup> and represents a major challenge to accurate prediction of cell fate and function from genetic information – one of the main goals of future genomic medicine. This poor correlation is due in part to the following reasons. First, the regulatory mechanism of gene expression is much more complex than initially expected<sup>6-8</sup>. The genes are interacting with each other and regulated by a range of epigenetic alterations<sup>9-14</sup>, suggesting the need to examine a panel of genes at the same time. Second, there is a significant degree of non-genetic cell-to-cell variability<sup>15,16</sup> and stochastic fluctuation of RNAs/proteins<sup>5,17-23</sup>, which was previously underestimated, requiring the use of single-cell resolution analysis. Despite recent advances in genomic technologies and next generation sequencing<sup>24-29</sup>, it is still challenging to investigate the genetic information flow through multiple levels of the Central Dogma (e.g., from DNA to RNA) at a single-cell level.

Microfluidic technologies emerged as a new approach to prepare single cell RNAs for gene expression analysis<sup>30,31</sup> and to quantify molecular targets in single cell<sup>32</sup>. It was reported that this approach significantly increased the mRNA-to-cDNA conversion efficiency by ~5 fold to reach 54% as compared to 12% for bulk-scale qPCR detection<sup>33</sup>. Microfluidics offers fundamental new capabilities for the manipulation of fluids, molecules, and cells that are very pertinent for the development of high-throughput, high-precision single-cell analysis methods<sup>34-37</sup>. Mathies *et al.* developed an agarose-droplet-based microfluidic platform that leverages emulsion-generator-array technology for high-throughput single-cell genetic analysis<sup>38</sup>. Quake *et al.* applied microfluidics to single-cell whole-genome amplification, which enabled increased parallelization and improved amplification performance<sup>39,40</sup>. To our best knowledge, there are no reports to date on processing single cells in microfluidics for simultaneous analyses of transcriptional and genomic signatures.

Here we report on an integrated microchip platform that can capture single cells, extract and process genomic DNA (gDNA) and messenger RNA (mRNA), respectively, from single cells, followed by whole pool amplification

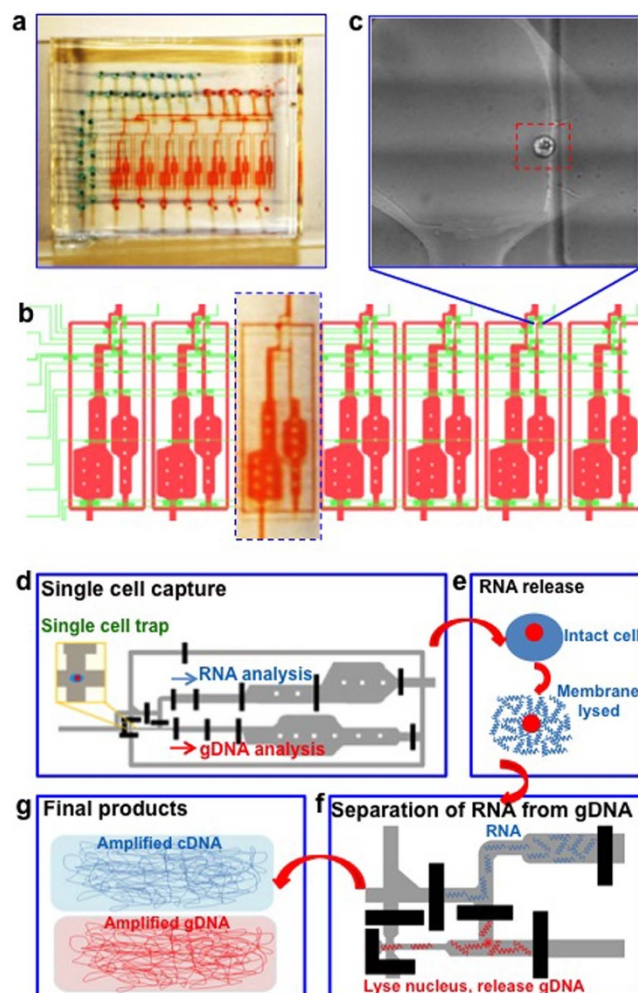


on chip. In conjunction with off-chip polymerase chain reaction (PCR), gel electrophoresis and Sanger sequencing, it enables co-detection of multiple transcripts and their cognate genes in the same single cell. This platform opens new opportunities to address the unexpectedly poor correlation between genomic and transcriptional signatures. It can help better delineate how gene expression is regulated at the single cell level, which is central to a range of fundamental biology questions, for example, stem cell fate control and cancer initiation.

## Results

**A microfluidic processor for separation of gDNA and mRNA from the same single cells.** The microfluidic chip was fabricated via soft lithography<sup>41,42</sup>. It contains flow channels (Fig 1a,b red) and control channels (Fig 1a,b green). The flow channels are used to load and process cells and conduct gDNA/mRNA analysis. The control channels permit programmable switching of the membrane valves to introduce reagents into or remove them from individual microchannels. The integrated microchip allows for conducting single-cell capture, selective lysis and separation of cytoplasmic and nuclear contents followed by on-chip reverse transcription of mRNA to generate complementary DNA (cDNA) and whole pool amplification of both cDNA and gDNA from the same single cell. This microchip has seven identical units, and each unit has two branches; the first is used to conduct mRNA analysis and the second is to amplify the whole pool of gDNA. The procedure to use this microchip is the following (Supplementary Figure S1). First, single cells were captured in the cell-loading microchamber. To successfully separate cytoplasm from the nucleus, a highly selective lysis buffer was developed. This buffer can break down the cell membrane to release cytoplasm (to be discussed in more details in the next section) while keeping the nucleus completely intact. We found that the nucleus can remain intact when submerged in this lysis buffer for one hour, which is more than enough to allow the following step - physical separation of cytoplasm from nucleus - to be completed. Second, the cytoplasmic content was slowly introduced to the RNA branch carried by the 1<sup>st</sup> strand synthesis reverse transcription reagent while the valve between the cell capture chamber and the RNA analysis branch was partly opened such that the nucleus was retained in the cell capture chamber for further analysis. Third, when the cytoplasm separation was completed, the reverse transcription of mRNA was conducted right away. Fourth, the nucleus was then lysed by a high-pH buffer to release gDNA, which was introduced to the next chamber of the gDNA analysis branch where gDNA was denatured and re-neutralized. Fifth, back to the RNA analysis branch, reagents were introduced to synthesize the second strand cDNAs, which were circularized by ligation. Finally, the circularized cDNA and single strand gDNA were amplified at the same time at 30°C overnight using a whole pool amplification procedure (WPA)<sup>43</sup>.

**Selective lysis of cell membrane in a microfluidic chip.** We developed a highly selective cell lysis protocol to facilitate the separation of cytoplasm from nucleus. A well-characterized chronic myelogenous leukemia cell line (K562) was used in this study. A series of lysis buffers prepared at varying concentrations of detergent was prepared and tested in the microfluidic chip to examine the lysis selectivity and performance (Figure 2a). The lysis buffer A containing 2% of ionic detergent (e.g., sodium dodecyl sulphate, SDS) quickly lysed both cell membrane and nucleus. The lysis buffer B spiked with high concentration of non-ionic detergent Triton (~5%) selectively lysed the membrane but was found to inhibit the reverse transcriptase activity in the next step. A low concentration non-ionic detergent Triton (~1%) buffer D failed to release the cytoplasmic RNAs completely, which may cause the loss of mRNA for transcriptional profiling. An optimized lysis buffer C, which contained 2% of non-ionic detergent Triton, was able to



**Figure 1 | Microfluidic device designed for single-cell mRNA/gDNA separation and whole pool amplification.** (a). Photograph showing a PDMS microfluidic device bound on a glass slide. The microfluidics device is filled with food dye for illustration. All flow channels are filled with red dye, and the control channel are filled with green dye. (b) Layout of microfluidic device with 7 identical units. Optical photograph of a single unit is overlaid on top. (c). A single K562 cell highlighted by the red square captured in the micro-chamber. (d–f). Schematic depiction of the process flow. (d) A single cell was captured in the microchamber of each unit. (e) The cell membrane was lysed by the membrane-selective lysis buffer. (f) RNA was separated from the intact nucleus. (g) Final products of cDNA and gDNA in separated microchambers are subjected to whole pool amplification.

selectively lyse the cell membrane in 10 min and had little effect on the activity of reverse transcriptase in the next steps.

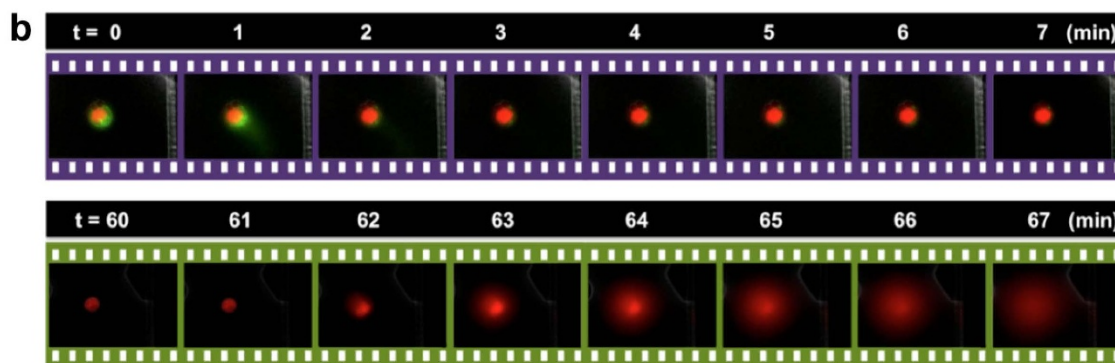
In order to visualize selective lysis and separation of cytoplasmic content from the nucleus, K562 cells were pre-stained with live cell cytosolic dye (green) and nuclear DNA dye (red). A pre-stained cell was captured in the microchamber and then the membrane-selective lysis buffer C was introduced (time=0 min). In 1 min, an observable amount of cytoplasmic content was already released as indicated by the green comet tail and the cell membrane was disrupted although the cell boundary was still visible. In 2 min, the majority of the cytoplasmic content was released and carried away by the lysis buffer. In 7 min, there remained only trace amount of cytoplasmic content (green) presumably the endoplasmic reticulum (ER) adjacent to nuclear membrane (Supplementary Video 1). Afterward, the naked nucleus was left in the lysis buffer C on chip for up to 1 hour and found



### a Optimization of the cell membrane-selected lysis buffers

Lysis buffers <sup>[a]</sup>	Time needed to completely lyse cell membrane	Intactness of nuclei	Impact on the following reverse transcription step
A	<1min	No	Inhibit RT
B	<1min	Yes	Not tested
C	<10min	Yes	No effect on RT
D	>60min, incomplete	Yes	No effect on RT

[a] A, 2% ionic detergent, SDS; B, 5% non-ionic detergent, Triton; C, 2% Non-ionic detergent, Triton; D, 1% non-ionic detergent, Triton.



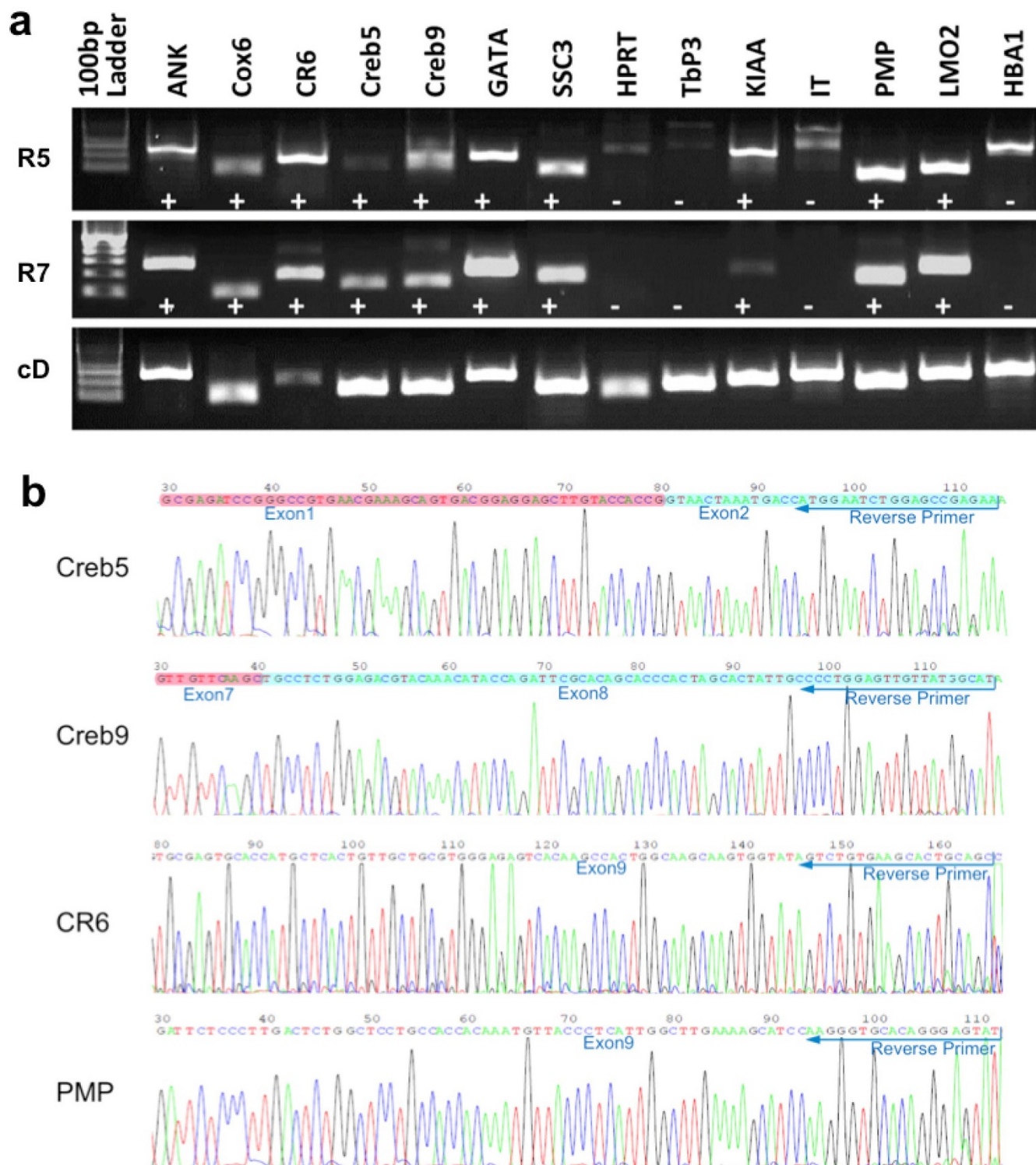
**Figure 2 | Selective lysis of cell membrane and extraction of cytoplasmic and nuclear contents respectively.** (a). Summary of 4 representative lysis buffers in terms of their performance for selective cell membrane lysis and subsequent mRNA analysis. (b). Live cell imaging of selective cell lysis in a microfluidic chip. The upper panel shows rapid lysis of cell membrane and release of cytoplasm (pre-stained in green) with no effect on nucleus. The lower panel shows the nucleus remains intact in the lysis buffer for 1 hr and then lysed by a high-pH buffer to release gDNA (pre-stained in red).

no observable signs of nuclear lysis. The cytoplasmic content was separated and carried to the RNA analysis microchannel directly by the 1<sup>st</sup> strand synthesis reagent at a very low flow rate <1 mm/min driven by an external pressure of 0.5–1 psi such that the shear force is insufficient to move the nucleus out of the cell capture chamber. In each of the next steps of mRNA analysis starting from reverse transcription to whole mRNA pool amplification, the corresponding reaction mix was pushed into the next empty chamber and the volume of different reagents introduced in each step was determined by the volume of microchambers. Afterwards, to extract gDNA from the nucleus for genomic analysis, a high pH solution (buffer A in Figure 2a: 400 mM KOH, 10 mM EDTA)<sup>43</sup> was introduced and the nucleus was quickly “dissolved” in 5–10 minutes as shown in Figure 2b lower panels (Supplementary Video 2). Upon optimization of the lysis reagents and the biomedical procedure implemented in microfluidics, we were able to successfully separate cytoplasmic content from nuclear content in a reliable and controllable manner. The separation purity can be examined in the next steps when the amplicons were subjected to PCR and sequencing analysis.

**Synthesis and amplification of cDNA on-chip from cytoplasmic mRNAs of single cells.** Here we describe how mRNAs in the separated cytoplasm were processed and amplified to give single-cell cDNA amplicons for PCR detection and sequencing of transcripts. Our approach for on-chip whole mRNA pool amplification was modified from a procedure we have developed, which is called Phi29 DNA polymerase-based full-length mRNA transcriptome amplification (PMA), which allows for improved recovery of 5'-end sequences according to our previous study<sup>44</sup>. To implement PMA in a microfluidics platform, we modified the protocol by eliminating a complicated cDNA purification step after the reverse transcription and prior to the on-chip whole pool amplification. Through systematic optimization of the reagents and substantial increase of each next step reaction mixture, a new system was built up upon each previous reaction.

We were able to perform reverse transcription and PMA to generate amplicons step-by-step on the microchip by pushing the reaction mix into the next chamber along the RNA analysis branch with no need of product purification between steps. First, we synthesized the first strand cDNA from the total RNA isolated on-chip using the enzyme SSRT III at 50°C for efficient reverse transcription that produces single-stranded cDNA (sscDNA), which was then converted to the double-stranded cDNA (dscDNA) followed by a circularization step using DNA ligase, which improves uniformity of full-length mRNA amplification. Second, the circularized cDNA was subjected to phi-29 polymerase-based isothermal amplification on chip at 30°C over night to give the initial cDNA amplicon. Afterwards, the amplicon was collected at the outlet by adding diluted BSA and blown out of the main amplification chamber using high-purity nitrogen gas. Because the total volume of the microchamber was limited and the quantity of cDNA product is typically <1 ng, the collected amplicon was then re-amplified off-chip in a conventional PCR tube using the same whole DNA-pool amplification procedure<sup>43</sup>. Finally, we were able to obtain >2 µg quantities of amplified cDNA. The product was examined by gel electrophoresis (Supplementary Figure S2) and quantified by UV absorption analysis.

**Amplification of genomic DNA on-chip from the nucleus of the same single cell.** gDNA isolated from the same single cell was introduced to the gDNA analysis branch and subjected to a simple denaturation and neutralization step before it can be amplified on chip. The final amplification procedure is nearly identical to the cDNA amplification and actually conducted at the same time because they both were carried out at the same temperature and duration. Similarly the gDNA amplicon was collected from the outlet and re-amplified in a PCR tube. Finally, sufficient quantities of gDNA (>2 µg) were obtained from single cells (Supplementary Figure S2) as demonstrated by the gel electrophoresis, in which 2 µL of cDNA and gDNA amplicons from the same single cells were analyzed. Quantification by UV absorption (Nanodrop) showed

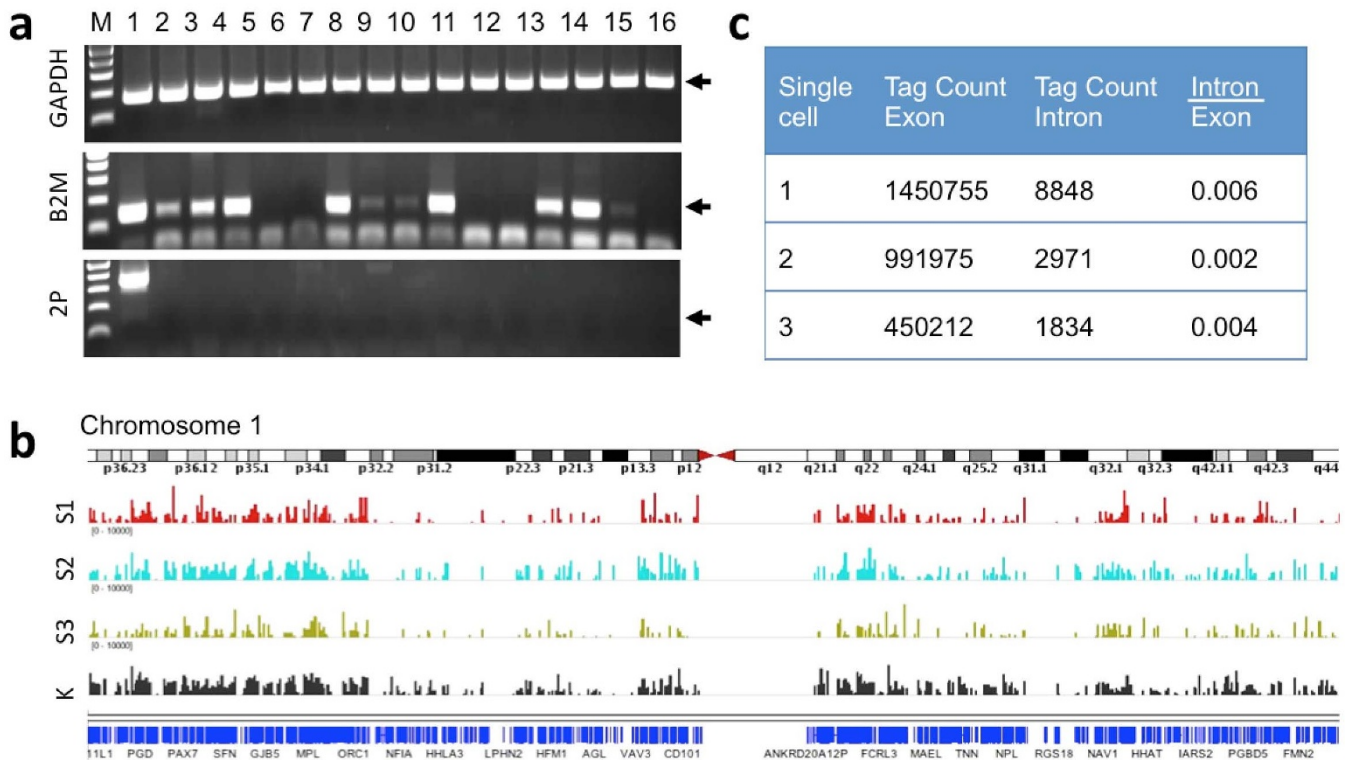


**Figure 3 | PCR detection and sequencing of multiple transcripts in single cells.** (a). PCR detection of transcripts by electrophoretic analysis of PCR products from single K562 cell cDNA amplicons and population controls. 14 pairs of primers were used to detect the amplification efficiency of single K562 cell by PCR. R5 and R7 are cDNA amplicons from 2 single K562 cells; cD is cDNA from a population of K562 cells; gD is genomic DNA from a population K562 cells. (b). Sanger sequencing results with the PCR products from the bands corresponding to Creb5, Creb9, CR6, and PMP.

that a few micrograms of DNA products (both cDNA and gDNA) were routinely obtained from single cells (Table S1), and this amount is sufficient for PCR detection and preparation of sequencing libraries.

**PCR detection and targeted sequencing of transcripts from single cells.** A set of primers was designed to detect gene expression by PCR

(Figure 3a and Supplementary Figure S3). This included primers for both house-keeping genes and the genes that are known to be expressed in K562 cells (Supplementary Table S2). These primers were designed using the NCBI primer design tool, and the forward/reverse primer pairs were chosen to cross exons when the sequences allowed. Two single-cell mRNA analysis samples (R5 and R7) were measured for 14 transcripts; 10 were successfully detected. A



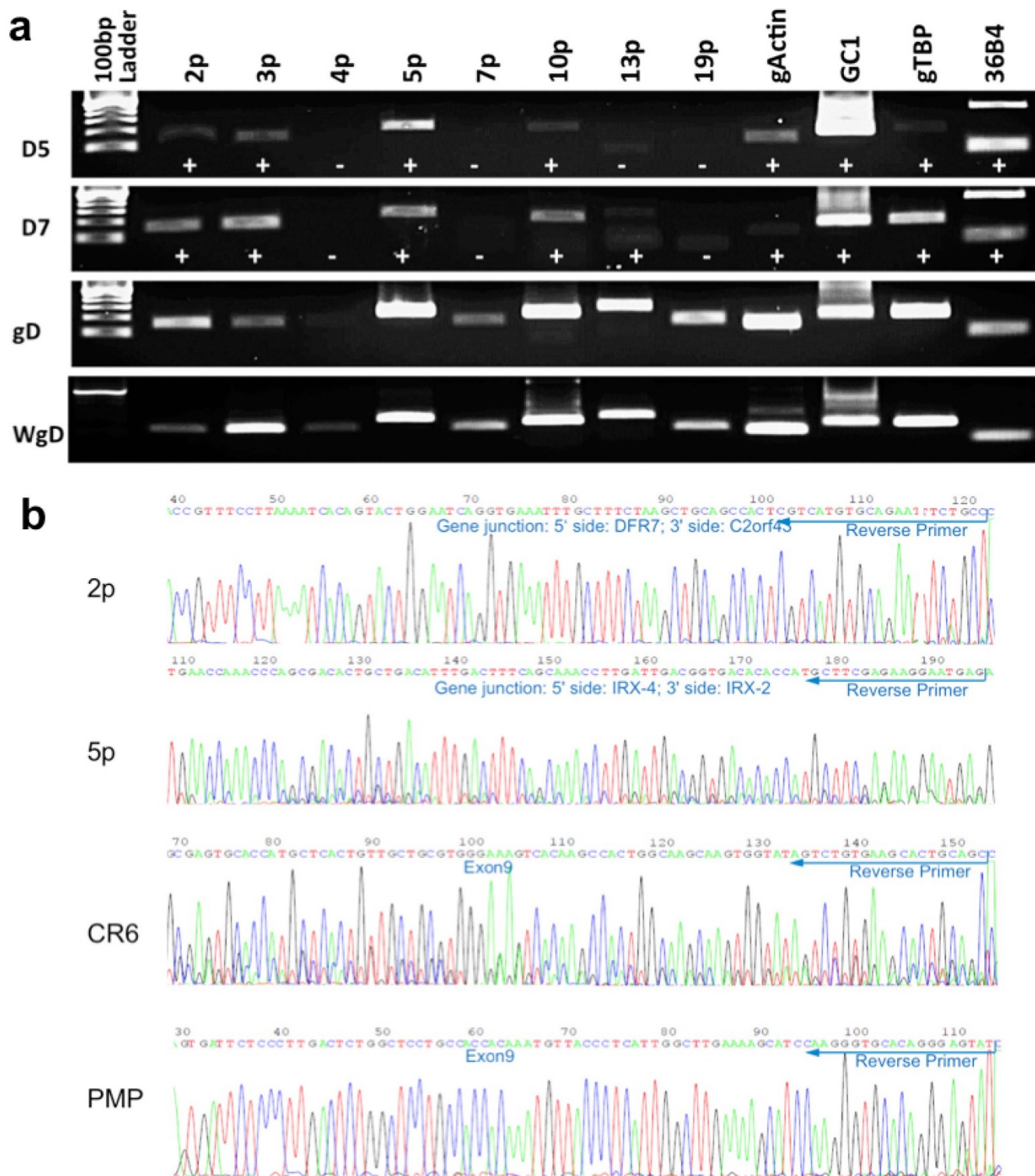
**Figure 4 | Single-cell mRNA amplicons: quality and purity.** (a). PCR detection of transcripts GAPDH and B2M in 16 single-cell mRNA amplicons. The same amplicons were also checked for the presence of genomic DNA using a primer pair 2p, which detects a sequence in the intron region of chromosome 2. M: molecular weight marker. 1–16: single-cell cDNA amplicons. (b). Next-generation sequencing of three single-cell mRNA amplicons (S1, S2, and S3) and a 1000-cell amplicon (K) showing excellent coverage over the gene sequences in chromosome 1. Blue bars indicate the human genome reference sequence. Non-coding regions are not present in any of these samples. (c). Summary of tag counts for exons and introns. The results are from three single-cell amplicons. The intron tag counts in all these samples are negligible ( $\sim 0.01\%$ ).

population control experiment was performed by analyzing the cDNA sample (cD) generated from a large number of K562 cells ( $\sim 2$  million cells) using the conventional method, in which the RNA was extracted using an RNeasy mini kit (Qiagen) and reverse transcribed into single-strand cDNA. 20 ng of cDNA was loaded to each gel electrophoresis lane. All 14 genes were detected in the population sample (cD). The results demonstrate the detection of multiple transcripts from single cells using cytoplasmic mRNA amplicons generated using the microchip described in Figure 1. To further confirm the sequences detected by PCR, the gel bands of the right molecular weight were cut out, re-amplified, and then analyzed by Sanger sequencing. All the transcripts of interest (Fig 3a) were analyzed (see the results in Supplementary Table S2) and the sequences identified by Sanger sequencing did match the NCBI database. Four representative sequencing traces are shown in Figure 3b. Two sequences (CR6 and PMP) flanked by the primer pairs are located within single exons and the other two are crossing exon-exon junctions: the exon1-exon2 junction at Creb5 and the exon7-exon8 junction at Creb9, which confirms that these cDNA products are indeed from transcripts not genomic DNA contaminants.

**Evaluation of single-cell cDNA amplicons: quality and purity.** Although we have confirmed the amplicons from single-cell cytoplasmic extracts are indeed from mRNAs, the possibility of gDNA contamination needs to be further excluded. While the nucleus appeared to be intact as seen by naked eyes under a microscope (Fig. 2b), it is a qualitative assessment. It requires further evaluation to judge whether or not gDNAs contaminate the cDNA amplicons. In addition, we prepared cDNA amplicons from 16 single cells using the same cytoplasm separation protocol and performed PCR analysis with three sets of primers. GAPDH and

B2M are housekeeping genes for cDNA and 2p is designed to detect a specific non-coding region between genes in chromosome 2. While GAPDH was detected in all 16 samples, B2M detected from 11 samples, the gDNA specific sequence (2p) is not detectable in any of these samples, indicating that single-cell cDNA amplicons generated using our method have no or minimal gDNA contamination. To further examine the coverage and purity, three samples were subjected to next-generation sequencing (Hi-Seq, Illumina). Figure 4b shows the readout at the chromosome 1 from three single-cell cDNA amplicons and a population (1,000 cells) amplicon. Single-cell data, although containing less reads compared to the population data, show excellent coverage and, importantly, little gDNA contamination when aligned to the human genome reference sequence (blue). We also conducted genome-wide tag counting of exons and introns from three representative single-cell cDNA amplicons (Figure 4c). The result shows more than 99.4% of tag reads are from exons. These results confirm that single-cell cDNA amplicons generated using our method is of high quality and has minimal gDNA contamination.

**PCR detection and targeted sequencing of genes in genomic DNA from the same single cells.** We designed 12 pairs of primers to detect genes from genomic DNA (Supplementary Table S3). These include 8 genomic DNA-specific primers and the regions to be detected using these primers are located on different chromosomes such that we can assess qualitatively the amplification uniformity. It also includes 4 pairs of primers for genes containing single exons (3 are of house-keeping genes). gDNA amplicons (D5 and D7) from the same single cells as mentioned in single-cell mRNA analysis (cell #5 and #7) were subjected to PCR analysis using the whole panel of primers (Figure 4a and Supplementary Figure S4). The control gDNA (gD)



**Figure 5 | PCR detection and sequencing of multiple genes from the gDNA of the same single cells. (a).** PCR detection of genes by electrophoretic analysis of PCR products from single K562 cell gDNA amplicons and population controls. 12 pair primers were used to measure the amplification efficiency of single K562 cell gDNA by PCR. D5 and D7 are gDNA amplicons from the same single K562 cells as of R5 and R7 (Fig. 3); gD is genomic DNA of a population of K562 cells. WgD is the gDNA product amplified from a small amount of sample gD used access amplification bias. (b). Sanger sequencing of the PCR products from 2p, 5p, CR6, and PMP bands.

extracted from a large amount of cells (>2million cells) using DNeasy blood and tissue kit (Qiagen) was also analyzed. 20 ng of gDNA amplicon was used for PCR and the PCR product was examined by gel electrophoresis as shown in Fig. 5a. All the genes were successfully detected from the population gDNA control

although 4p is relatively weak. 8 (D5) or 9 (D7) out of the 12 genes were detected in single-cell gDNA amplicons with the corresponding bands at the right sizes and different intensities. In order to assess the amplification bias, 100 ng of the population gDNA control was amplified using the same WPA method as for generating single



cell amplicons. We found that the amplification was generally uniform except 2p (under-amplified) and 3p (over-amplified). Finally, the PCR bands were all cut out from the gel for Sanger sequencing and the results are summarized in Supplementary Table S3. The sequencing results did match the expected gDNA sequence. The selected examples are shown in Figure 5b. The regions flanked by the primers for 2P and 5P are located in the space between genes and thus are only detectable in gDNA. Sanger sequencing result shows exactly the expected sequence in these regions. We also performed Sanger sequencing of genes located within single exons, e.g., CR6 and PMP, which indeed can be detected from both gDNA and mRNA. The results confirm the correct sequence within the exons. Analysis of these genes (e.g., CR6 and PMP) is the first demonstration of co-measurement and sequencing of the same transcripts and cognate genes from the same single cells.

## Discussions

We developed a microfluidics-facilitated platform that can reliably separate and extract mRNA and gDNA from the same single cells followed by on-chip processing and amplification of transcripts and genes. Using off-chip amplification and PCR targeting, we were able to detect most transcripts in our panel in single cell samples. The transcripts that were not always detected in our single-cell amplicons but supposedly expressed at significant levels according to the population data are likely associated with stochastic cell-to-cell variability. But the exact mechanisms require future studies using genome-wide single-cell RNA sequencing. The microfluidic platform provides nanoliter-level reaction space, and the small volume enables rapid mixing and increases the efficiency of reverse transcription reactions. The sealed reaction space can effectively prevent contamination of RNA or DNA from the outside environment. A single cell has only ~10–40 pg of total RNA<sup>45,46</sup> and ~6 pg of DNA<sup>47</sup>. The small amount of labile mRNA often degrades before it can be reversely transcribed into a stable cDNA copy during traditional processing<sup>48</sup>. For example, using standard laboratory reagents and hardware, only a small number of genes can be quantitatively assessed per cell<sup>49</sup>. Rapid mixing of microliter quantities of reagents is able to increase the efficiency of reverse transcriptase so that mRNA can be effectively converted to cDNA before the degradation occurs. However, this is not trivial because at the microliter scale liquid flow is laminar, and current mixing methods (i.e. shaking, vortexing and trituration) cannot produce sufficient chaotic flow to effectively mix reagents. In our microfluidic device, the reagents were flowed through all the previous reaction microchambers, which helps enhance efficiency and speed of mixing in nanoliter-level volumes. However, an active mixing mechanism, if successfully implemented in the microfluidic device, will further improve the device performance. Our microfluidic device is made of a soft rubber material – PDMS, which is sticky and can absorb DNA or RNA in a non-specific manner. Due to the low copy number of RNA or DNA molecules from a single cell, the loss of RNA or DNA due to non-specific binding to channel walls may not be negligible. In order to reduce the loss of DNA or RNA and increase the coverage over the genome scale, new methods are needed to block the surface of the microchannels prior to cell lysis. Nevertheless, this work, for the first time, demonstrates co-measurement of multiple genes and transcripts from the same single cells. The correlation between gene expression patterns and genomic alterations at the level of single cells can potentially address a widely range of basic biology questions such as stem cell fate control, cellular immunity and cancer development.

## Methods

**Device design and operation.** The microfluidics chip has seven identical processing units. The chip contains 27 control lines operating 105 valves. The function of valves are the following: 7 valves to control seven unit inputs separately, 7 valves to control single cell capture, 13x7 valves for mRNA reverse transcription and genomic DNA

amplification operation. Figure 1 shows the overall design of the microfluidics system. The microfluidic valves in the device are controlled by individual pressure regulators and are interfaced via 23 gauge stainless steel pins and tygon tubing.

**Device fabrication.** The chip was fabricated by a multilayer soft lithography technique using polydimethylsiloxane (PDMS; RTV 615A and B). Two photomasks were first generated with microscale patterns designed by computer-aided design software L-Edit and printed on high-resolution transparencies. The control layer mold was made of negative photoresist SU-8 2010 spun on a 4-inch silicon wafer. After development, the control mold was hard-baked at 150 °C for 30 minutes to remove the moisture and harden the photoresist. The region where a control valve and a flow channel intersect is at least 120 μm x 120 μm. The control lines connecting the valve regions are 20 μm in width and 20 μm in height. The flow layer mold was made of positive photoresist SPR-220 spun on a 4-inch silicon wafer. After development, the flow mold was baked at 140 °C for 40 minutes to allow the SPR-220 to reflow, thus yielding a template for molding channels with a round cross section. The flow channels has varying channel width in the range of 30–800 μm, and the channel height of ~20 μm. Afterwards, the PDMS pre-polymer at a mass ratio of RTV615A : RTV615B = 5 : 1 was poured onto the control mold in a Petri dish to fabricate a ~5 mm thick control layer, and the PDMS pre-polymer at a mass ratio of RTV615A : RTV615B = 20 : 1 was spun onto the flow modes at 3000 rpm for 1 minute resulting in a thin flow layer (~20 μm). Both layers of PDMS were cured at 80 °C for 40 minutes. Then access holes were punched in the control layer chip to produce inlets for pressure control. The control layer chip was then aligned and bonded to the flow layer under an optical microscope. The bonded two-layer structure was baked at 80 °C for 1 hour, cut out from the mold and then the hole were punched to produce inlets and outlets to access the flow layer microchannels. Once the two-layer PDMS and a pre-cleaned glass slide were treated with oxygen plasma, they were immediately brought into contact to complete the device assembly. Finally the assembled chip was baked at 80 °C for 1 hour to strengthen the bonding between PDMS and the glass slide.

**Single-cell processing on chip.** Nucleic acid processing was performed in one of two ways: single reactions in a serial manner or seven reactions in parallel with the 7plex device. The device contains a module for mRNA extraction and reverse transcription and a module for gDNA extraction. Single-cell mRNA processing includes (1) cell capture (Fig. 1Sb), (2) cytoplasm lysis (Fig. 1Sc) (3) cytoplasm separation and 1<sup>st</sup> strand generation (Fig. 1Sd), (4) 2<sup>nd</sup> strand generation (Fig. 1Se), (5) cDNA ligation (Fig. 1Sf), and (g) whole pool amplification (Fig. 1Sg). Single-cell gDNA processing includes (1) nuclei lysis and denaturation (Fig. 1Sh), (2) denaturation buffer neutralization (Fig. 1Si), and DNA whole pool amplification (Fig. 1Sj). Single cells were captured one by one in the seven identical units. Valve 0 is switched off, when a single cell travels into the small chamber identified by observing under a microscope, we switch on v0 to trap one cell in the chamber. The mRNA reverse transcription is conducted first. The lysis buffer is loaded into the microchamber. Valve 1 is used to introduce cell lysis and valve 3 is switched off to direct the increased volume of cell lysate to move into the empty processing chamber. Then valve 4 is switched off, and valve 1 is opened to introduce 1<sup>st</sup> strand reagents to separate cytoplasm from the nuclei. After the space between v2, v4 and v5 is filled, v1 and v4 are switched on and the separated mRNA is ready to be used for generating 1<sup>st</sup> strand cDNA. Valve 2 will become the inlet for the following mRNA reverse transcription. Multiple valves downstream of valve 4 (v7, v9, v11 in the mRNA reverse transcription reaction chain) will allow for multi-step processing for 2<sup>nd</sup> strand generation, cDNA ligation, and whole pool amplification. After the mRNA reverse transcription, the nucleus is lysed and denatured by a basic buffer introduced by switching on v1 till the region flanked by v5 and v8 and v5 is switched off. After turning off v8, the neutralizing buffer is introduced through v1. Finally, the whole pool amplification of gDNA is conducted by switching off v1 and v10 to flow in the amplification reagents. The same procedure is conducted in parallel for the 7 identical units at the same time. The amplified cDNA and gDNA in the microfluidics devices are collected respectively, and are amplified in thin-well PCR tubes (200 μl) for the second time amplification in order to obtain enough DNA for PCR detection and sequencing.

**K562 cell culture.** The K562 cell line was maintained in RPMI 1640 media supplemented with 10% FBS at a cell density between  $2 \times 10^5$  and  $2 \times 10^6$  and incubated at 37 °C in a 5% CO<sub>2</sub> atmosphere. Cells were passaged every other day.

**Selective lysis of cell membrane.** Single cells were captured in the microchamber of each unit on the microfluidic chip. The cytoplasm was lysed by loading the lysis buffer which contains 1 or 2% Triton, 20 mM NaCl, and 20 mM tris(hydroxymethyl)aminomethane(Tris). The buffer is highly selective. Even after 1–2 hours, the nucleus remains intact in the membrane lysis buffer. After the membrane was completely lysed, the inlet valve was opened, and the first-strand reaction solution was introduced at a low flow rate and carried the lysate into a different microchamber while the nucleus in the previous microchamber was kept. The valve between the two microchambers was partially-opened in order to retain the nucleus in the previous chamber. The valve was initially closed and then slowly released by decreasing the control line pressure until the fluid can pass through but the nucleus remains in the original chamber. In a typical experiment, the pressure to completely close the control valve is 12 Psi, and the pressure of 4–6 Psi will not be able to fully close V4 so that the liquid can keep going through. The membrane and nucleus lysis procedure is shown in the Supplementary videos. The optimized buffer (Buffer C) was



also tested for its applicability to other cell types including human embryonic stem cells and brain tumour cells (U87 cell line) to selectively lyse the cell and extract messenger RNAs (Supplementary Figure S5).

**Reverse transcription and PMA.** Assuming that the RNA solution in the chip is 1 unit volume, the RT was carried out in a 2-unit volume with introducing the first-strand synthesis buffer (final 1×: 50 mM TrisHCl, pH 8.3, 75 mM KCl, and 6 mM MgCl<sub>2</sub>; Invitrogen), MgCl<sub>2</sub> (6 mM), DTT (2 mM; Invitrogen), RNaseOut (0.8 U/μL; Invitrogen), and SuperScript Reverse Transcriptase III (10 U/μL, SSRTIII; Invitrogen). The thermal cycling program was set as 55°C × 5 min, 37°C × 3 min, 42°C × 3 min, 45°C × 3 min, 50°C × 90 min, and 70°C × 10 min, and then the sample was cooled to 4°C. Because double-strand cDNA (ds-cDNA) was required for robust ligation, the second strand was generated in a 4-unit reaction volume using the product described above by introducing the following reagents: second Strand Buffer (0.8× concentration: 16 mM Tris-HCl, 9.6 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 8 mM MgCl<sub>2</sub>, 0.128 mM β-NAD; NEB), RNaseH (0.1 U/μL; Epicentre), Escherichia coli DNA ligase (0.125 U/μL; Epicentre), E. coli DNA polymerase (0.15 U/μL), and dNTPs (0.125 mM). The reaction was conducted at 16°C × 90 min, followed by 70°C × 10 min for inactivation of the enzymes. For phi29 DNA polymerase, the End-It DNA End-Repair Kit (Epicentre) plus T4 DNA ligase (Epicentre) were then combined for DNA end blunting, 5'-end phosphorylation, and ligation. This included End-it buffer 1×, 1 mM dNTPs, 1 mM ATP, 4% total enzyme mixture, and T4 DNA ligase (0.4 U/μL). The reaction volume was 5 units and was incubated at room temperature for 120 min. Without deactivation of enzyme, and without denaturation of the DNA template, 7.5 units of 2× whole DNA pool amplification procedure (WPA) premixture [containing buffer, trehalose, dNTPs, random oligonucleotide N9, based on the WPA procedure (ref. 1)] plus RepliPhi phi29 DNA polymerase (1,000 U/μL × 0.4 μL/60 μL; Epicentre) and fresh DTT (1 mM; Invitrogen) were added to the DNA template prepared above, with a final reaction volume of 12.5 units. The reaction was carried out at 30°C for 16 h.

**Whole gDNA pool amplification.** First, the nucleus was lysed and gDNA was denatured by treating with the denaturing buffer for 5 minutes at room temperature. Then an equal volume of neutralization buffer was flowed in to neutralize the reaction solution. The ssDNA was amplified by applying 20-fold whole pool amplification master mixture<sup>13</sup>. For high-efficiency amplification, the microfluidic chip was incubated at 30°C for 16 hours.

**Amplicon collection.** After the whole pool amplification of cDNA and gDNA, the amplicons were collected by flowing water or diluted BSA solution from the inlet of each unit to push the amplicon out of the outlet. The amplicons were collected into the Tygon tubing at the outlets. In order to ensure the amplicon was fully recovered, ~6 μL of water was used to wash it out of the tubing and collected in a PCR tube.

**Purification of cDNA and gDNA.** Genomic DNA and polyadenylated RNA-derived cDNA amplicons were purified by Genomic DNA Clean and Concentrator kit (Zymo) column. After purification, 2–4 μg gDNA and cDNA amplicon was obtained, which was then evaluated by PCR and Sanger sequencing. Gel electrophoresis results of mRNA reverse transcriptional cDNA and amplified gDNA from the same single cells as corresponding cDNA. Ch1 - Ch7 are the seven identical units on one microfluidic device.

**Sanger sequencing.** Selected PCR products were cut from the gel and purified by Zymoclean Gel DNA recovery Kit. 20 ng of purified PCR products were submitted for Sanger sequencing.

**Next generation sequencing.** The cDNA amplicons (3 to 5 μg) from single cells were fragmented to an approximately 200–500 bp size range by a Bioruptor Sonicator (Diagenode). After purification with DNA Clean & Concentrator kit (Zymo), end-repairing, 3'-A tailing and ligation with adaptor was performed. Then, a DNA band between 250–300 bp was selected by gel electrophoresis (E-gel EX 2%, invitrogen) and barcode added by PCR using Phusion High-Fidelity DNA polymerase (NEB) for 8 cycles. The product was size selected again and the DNA concentration was quantitated by a Bioanalyzer (Agilent). Eight samples were pooled and loaded to the Hi-Seq2000 for sequencing performed with 50 bp single end reads.

- Crick, F. Central Dogma of Molecular Biology. *Nature* **227**, 561–& (1970).
- Sustar, P. Crick's notion of genetic information and the 'central dogma' of molecular biology. *Brit J Philos Sci* **58**, 13–24 (2007).
- Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol* **4** (2003).
- Yeung, E. S. Genome-wide Correlation between mRNA and Protein in a Single Cell. *Angew Chem Int Edit* **50**, 583–585 (2011).
- Cai, L., Friedman, N. & Xie, X. S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358–362 (2006).
- Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
- Peter, I. S. & Davidson, E. H. A gene regulatory network controlling the embryonic specification of endoderm. *Nature* **474**, 635–U110 (2011).
- Smith, J., Theodoris, C. & Davidson, E. H. A gene regulatory network subcircuit drives a dynamic pattern of gene expression. *Science* **318**, 794–797 (2007).
- Dawson, M. A. & Kouzarides, T. Cancer Epigenetics: From Mechanism to Therapy. *Cell* **150**, 12–27 (2012).
- Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R. & Misteli, T. Epigenetics in Alternative Pre-mRNA Splicing. *Cell* **144**, 16–26 (2011).
- Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A landscape takes shape. *Cell* **128**, 635–638 (2007).
- Dove, A. Epigenetics: The Final Frontier? *Science* **326**, 303–305 (2009).
- Berger, S. L. Chromatin and gene regulation - Molecular mechanisms in epigenetics. *Science* **300**, 252–254 (2003).
- Jones, P. A. & Takai, D. The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068–1070 (2001).
- Altschuler, S. J. & Wu, L. F. Cellular Heterogeneity: Do Differences Make a Difference? *Cell* **141**, 559–563 (2010).
- Niepel, M., Spencer, S. L. & Sorger, P. K. Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr Opin Chem Biol* **13**, 556–561 (2009).
- Blake, W. J., Kaern, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
- Bahar, R. *et al.* Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**, 1011–1014 (2006).
- Sigal, A. *et al.* Variability and memory of protein levels in human cells. *Nature* **444**, 643–646 (2006).
- Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. & Huang, S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544–U510 (2008).
- Eldar, A. & Elowitz, M. B. Functional roles for noise in genetic circuits. *Nature* **467**, 167–173 (2010).
- Li, G. W. & Xie, X. S. Central dogma at the single-molecule level in living cells. *Nature* **475**, 308–315 (2011).
- Gupta, P. B. *et al.* Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. *Cell* **146**, 633–644 (2011).
- Perkel, J. M. Sangerwho? Sequencing the Next Generation. *Science* **324**, 275–279 (2009).
- Foo, J. N., Liu, J. J. & Tan, E. K. Next-generation sequencing diagnostics for neurological diseases/disorders: from a clinical perspective. *Hum Genet* **132**, 721–734 (2013).
- Mardis, E. R. Next-generation DNA sequencing methods. *Annu Rev Genom Hum Genet* **9**, 387–402 (2008).
- Lizardi, P. M. Next-generation sequencing-by-hybridization. *Nat Biotechnol* **26**, 649–650 (2008).
- Shendure, J. & Ji, H. L. Next-generation DNA sequencing. *Nat Biotechnol* **26**, 1135–1145 (2008).
- Smith, D. R. *et al.* Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res* **18**, 1638–1642 (2008).
- White, A. K. *et al.* High-throughput microfluidic single-cell RT-qPCR. *Proc Natl Acad Sci U S A* **108**, 13999–14004 (2011).
- 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).
- Huang, B. *et al.* Counting low-copy number proteins in a single cell. *Science* **315**, 81–84 (2007).
- Zhong, J. F. *et al.* A microfluidic processor for gene expression profiling of single human embryonic stem cells. *Lab chip* **8**, 68–74 (2008).
- Zeng, Y., Novak, R., Shuga, J., Smith, M. T. & Mathies, R. A. High-Performance Single Cell Genetic Analysis Using Microfluidic Emulsion Generator Arrays. *Anal Chem* **82**, 3183–3190 (2010).
- Andersson, H. & van den Berg, A. Microtechnologies and nanotechnologies for single-cell analysis. *Curr Opin Biotech* **15**, 44–49 (2004).
- El-Ali, J., Sorger, P. K. & Jensen, K. F. Cells on chips. *Nature* **442**, 403–411 (2006).
- Toriello, N. M. *et al.* Integrated microfluidic bioprocessor for single-cell gene expression analysis. *Proc Natl Acad Sci U S A* **105**, 20173–20178 (2008).
- Novak, R. *et al.* Single-Cell Multiplex Gene Detection and Sequencing with Microfluidically Generated Agarose Emulsions. *Angew Chem Int Edit* **50**, 390–395 (2011).
- Marcy, Y. *et al.* Nanoliter reactors improve multiple displacement amplification of genomes from single cells. *PLoS Genet* **3**, 1702–1708 (2007).
- Wang, J. B., Fan, H. C., Behr, B. & Quake, S. R. Genome-wide Single-Cell Analysis of Recombination Activity and De Novo Mutation Rates in Human Sperm. *Cell* **150**, 402–412 (2012).
- Whitesides, G. M., Ostuni, E., Takayama, S., Jiang, X. Y. & Ingber, D. E. Soft lithography in biology and biochemistry. *Annu Rev Biomed Eng* **3**, 335–373 (2001).
- Unger, M. A., Chou, H. P., Thorsen, T., Scherer, A. & Quake, S. R. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science* **288**, 113–116 (2000).
- Pan, X. H. *et al.* A procedure for highly specific, sensitive, and unbiased whole-genome amplification. *Proc Natl Acad Sci U S A* **105**, 15499–15504 (2008).
- Pan, X. H. *et al.* Two methods for full-length RNA sequencing for low quantities of cells and single cells. *Proc Natl Acad Sci U S A* **110**, 594–599 (2013).
- Roosmond, R. C. Ultramicrochemical Determination of Nucleic-Acids in Individual Cells Using Zeiss Umsp-1 Microspectrophotometer - Application to





Isolated Rat Hepatocytes of Different Ploidy Classes. *Histochem J* **8**, 625–638 (1976).

46. Uemura, E. Age-Related-Changes in Neuronal Rna-Content in Rhesus-Monkeys (Macaca-Mulatta). *Brain Res Bull* **5**, 117–119 (1980).
47. Kumar, G., Garnova, E., Reagin, M. & Vidali, A. Improved multiple displacement amplification with phi 29 DNA polymerase for genotyping of single human cells. *Biotechniques* **44**, 879–+ (2008).
48. Boon, W. C. *et al.* Increasing cDNA yields from single-cell quantities of mRNA in standard laboratory reverse transcriptase reactions using acoustic microstreaming. *J Vis Exp* e3144 (2011).
49. Aumann, T. D. *et al.* SK channel function regulates the dopamine phenotype of neurons in the substantia nigra pars compacta. *Exp Neurol* **213**, 419–430 (2008).

## Acknowledgments

This study was supported by the State of Connecticut Stem Cell Research Grant (Award No. 11SCA15, PIs: R.F. and S.M.W.), the Packard Fellowship for Science and Engineering (Award No. 201238215, PI: R.F.), the Single Cell Profiling Core supported by the U.S. National Cancer Institute Physical Sciences - Oncology Program Grant (U54 CA143798, subaward PI: R.F.), the National Science Foundation (NSF) Faculty Early Career Development Award (CBET-1351443, PI: R.F.), the University of Hawaii Cancer Center startup funding (PI: L.X.G.), and the U.S. National Institutes of Health Grant (1P01GM099130-01, Co-PI: SMW). We also acknowledge the Yale Institute for Nanoscience and Quantum Engineering (YINQE) and the Yale Nanofabrication Center to allow us to use their facilities.

## Author contributions

R.F., X.P. S.M.W. conceived the research project. L.H., X.Z. and X.P. performed the wet bench experiments, data acquisition and quantification. L.X.G. participated in sequencing analysis. Y.W. participated in microchip fabrication. All the authors contributed to data analysis and the writing of the manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Han, L. *et al.* Co-detection and sequencing of genes and transcripts from the same single cells facilitated by a microfluidics platform. *Sci. Rep.* **4**, 6485; DOI:10.1038/srep06485 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>