

Single-cell cDNA microarray profiling of complex biological processes of differentiation

Kazuki Kurimoto^{1,2} and Mitinori Saitou^{1,2}

¹Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.

²JST, CREST, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.

Correspondence should be addressed to:

Kazuki Kurimoto, Ph.D.

E-mail: kurimoto@anat2.med.kyoto-u.ac.jp; Tel: +81-75-753-4337; Fax: +81-75-751-7286

or

Mitinori Saitou, M.D., Ph.D.

E-mail: saitou@anat2.med.kyoto-u.ac.jp; Tel: +81-75-753-4335; Fax: +81-75-751-7286

Summary

Gene expression profiling at the single-cell level has been used to identify genes expressed in specific cell populations, in attempts to address various fundamental questions in multicellular organisms. In this article, we review the advance of single-cell cDNA amplification techniques in the last decade, and introduce a recently developed, reliable, quantitative method that is applicable to genome-wide transcriptional analyses with high-density oligonucleotide microarray and massively parallel sequencing. This method has been applied to a variety of biological studies, including developments of blastocyst inner cell mass, neurons, and primordial germ cells, to profile the molecular properties, dynamics during differentiation, and impacts of gene alterations in the individual cells in depth. These studies uncovered complex behaviors of the cells during differentiation *in vivo*, and identified previously unknown, transient populations that emerged in specific stages of development. These achievements clearly demonstrated that it is now more feasible to analyze gene expression in any cell type of interest in a quantitative, genome-wide manner at the single-cell resolution.

Introduction

Multicellular organisms are complex arrays of numerous functionally and phenotypically distinct cell types, with essentially the same genomic information. Such variation is achieved by differential gene expression, and therefore, the quantitative measurement of expression in a small number of cells, ideally single cells, is essential for the understanding of properties or states of cells in any biological context. This review therefore first describes recent advances in the methods for single-cell expression analysis.

Modern microarray platforms [1] and massively parallel sequencing techniques [2] provide major opportunities for quantitative, genome-wide transcriptional analyses, and open the possibility of a systems level understanding of life. These technologies, however, usually require large amounts of starting materials obtained from typically more than ten thousand cells. Due to this limitation, there is a risk that the methods will fail to detect differences among individual cells in a population, even if they are marked with particular genes or surface antigens. To overcome this problem, methods for single-cell cDNA analyses have been developed and used in many biological studies.

Recently, a fully validated, quantitative cDNA amplification method applicable to commercially available oligonucleotide microarrays (e.g., Affymetrix GeneChip) and massively parallel sequencing has been developed [3, 4]. This method has been applied to early mouse embryogenesis [4-6], stem cell biology [7-11], neurogenesis [12]

and primordial germ cell (PGC) development [13-17], and some of these efforts are a focus of the latter part of this review. These studies quantitatively revealed the expression dynamics of individual cells in differentiation and the impact of specific gene alterations, and clearly demonstrated the feasibility of transcriptome analysis at the single-cell resolution.

Methods for single-cell cDNA analysis

In order to profile gene expression in single cells, there are two major strategies; one is global amplification to make cDNAs applicable to various analytical methods [e.g., quantitative PCR (Q-PCR), microarray, massively parallel sequencing, serial analysis of gene expression (SAGE)]; the other is highly sensitive measurement with an unamplified single-use cDNA library.

Recent studies using sophisticated Q-PCR methods have achieved highly quantitative performance with the latter strategy [18-21]. Moreover, high-throughput analyses employing this strategy in conjunction with microfluidics devices (e.g., Fluidigm Digital Array chips) (for review, see [22]) have been reported [23-26]. More recently, this strategy was applied to five hundred single cells from mouse early embryos at developmental time points from the one-cell zygote to the sixty-four-cell blastocyst stages, and revealed expression dynamics of selected forty eight genes [27]. However, the single-cell cDNA library without global amplification is less suitable for repetitive use, and less applicable to gene screening or genome-wide analyses. Therefore, these methods would be useful for quantifying a limited number of known genes for a large number of cells.

cDNA amplification methods, on the other hand, have been used in a wide variety of biological studies, with at least two major strategies currently available; one is exponential amplification by Brady et al., based on polymerase chain reaction (PCR) [28, 29], and the other is multiple rounds of isothermal linear amplification by Van Gelder et al., based on *in vitro* transcription (IVT) with T7 RNA polymerase [30, 31].

Linear amplification has been widely used in the standard protocol for probe preparation with biotin-labeled substrates in modern oligonucleotide microarray systems, including commercially available Affymetrix GeneChip and CodeLink microarrays. It has also been applied to small amounts of material consisting of less than one hundred cells [32-34], including single-cell analyses[30, 31, 35-37]. Despite the potential advantage of the linear amplification that, in principle, it avoids the risk of rapid increase of random fluctuation or gene bias, such as often occurs in exponential amplification, it bears several practical disadvantages, including the need for complicated procedures with repeated buffer exchange with DNA/RNA purifications, the production of single-stranded RNA as the final product, and the limited amplification efficiency (up to

one-thousand fold per round) [38, 39].

Development and advances of exponential amplification

Brady et al. developed a global single-cell cDNA amplification method with relatively simple procedures relying on PCR [28, 29]. In this method, single cells were isolated directly in tubes that contained cell lysis buffer, with no need for purification throughout the procedure (Fig. 1A). After the cell lysis, the first-strand cDNA was synthesized in a relatively short amount of time using low concentrations of dNTP and primers, so as to be restricted to a few hundred base pairs (five hundred in average) from the 3'-end of mRNA. The first-strand cDNA was tailed with poly dA with terminal deoxynucleotidyl transferase (TdT), to which the same oligo dT-tailed primer annealed. The cDNA was then subjected to a large number of PCR cycles, typically fifty [28, 29] (Fig. 1A). This method amplifies cDNA fairly robustly, especially those highly expressed, and has been widely used in various biological studies. Among the discoveries made using this method, for example, was the identification of the pheromone receptors from sensory neurons in rat vomeronasal organs [40].

More recently, a number of genes specifically expressed in the primordial germ cells (PGCs) have been identified with this method [41]. In mice, PGCs originate from the posterior epiblast and are specified in the extraembryonic mesoderm at around embryonic (E) day 7.5 as a small number of cells (typically around forty) [42, 43]. A differential cDNA screening between putative single PGCs and their closest somatic neighbors has identified a number of PGC-specific genes [41], among which *Blimp1/Prdm1* turned out to be a critical determinant of the germ cell lineage [44]. As described below, the molecular program and the role of *Blimp* in this developmental process has been fully elucidated with improved methods [4, 14, 16, 17].

This method and its modified versions have also been applied to gene profiling on microarrays [38, 45-49]. At the same time, however, it has been reported that the original method exhibited significant variation among individual amplifications [squared correlation coefficient (R^2) 0.69] and distortion from the original expression profile ($R^2=0.41$) [4, 45]. A modified method by Iscove et al. was shown to produce a presumable byproduct without the poly A signal of mRNA at a relatively high frequency (~28% of amplicon) [38]. Therefore, improvement of these methods would lead to a better outcome. Moreover, since these methods compromised sense-antisense orientation by tagging cDNAs with the same primer sequence, the amplified cDNAs are no longer suitable for the standard labeling protocols employed by the commercially available oligonucleotide microarrays, which use isothermal linear amplification as described above. Therefore, there has been a demand for an amplification strategy that preserves the orientation of cDNA for quantitative single-cell microarray analysis.

The quantitative performance of the single-cell cDNA amplifications has been greatly improved, at least in part due to advances in the enzymes, the reduction of the PCR cycle numbers (~25 cycles), the optimization of primer sequences, the use of the PCR buffer throughout the procedure [50], and the employment of spike RNAs as an external control (artificially synthesized poly A tailed RNA of known copy numbers), which together have resulted in proportional amplification from several tens copies of mRNA [15, 48, 51]. These methods have been used to characterize melanocyte stem cells [51] and to explore the dynamics of multiple genes expressed during 36 hours in the PGC specification stage [15].

Development of a quantitative single-cell microarray method

To preserve the sense-antisense orientation of the mRNA, the first-strand cDNA must be tagged with different primer sequences. To do so, several approaches have been developed [3, 4, 52, 53], among which the method by Kurimoto et al. [3, 4] has been employed in various biological studies as mentioned below. In principle, this method eliminated the residual first primer [V1 (dT)₂₄] with Exonuclease I after the first-strand synthesis (Fig. 1B). The second primer [V3 (dT)₂₄] was added to the reaction mixture after the cDNA tailing, so that the second-strand cDNA was tagged with two different sequences (Fig. 1B). To minimize the risk of random fluctuation and gene bias in amplification efficiency, the cycle number was kept small (20 cycles), and the amplification was performed in four split tubes that were combined again after PCR (Fig. 1B).

This method turned out to be robust, with a DNA amplification success rate more than 90% for purified diluted total RNA (10 pg) and 75% for real single cells [3]. The sequence analysis demonstrated that all of the amplified cDNA preserved the sense-antisense orientation with bona fide transcript ends (about seven hundred base pairs from the 3'-end in average), demonstrating faithful amplification [4]. The efficiency of amplification was high, with the amount of cDNA being nearly doubled in each cycle (Fig. 1B). The proportional amplification was ensured for genes expressed at twenty copies per cell or more, as demonstrated with the spike RNAs and more than twenty genes in the diluted total RNA [4]. Microarray analysis showed that the false negative and positive rates were sufficiently low for genes expressed at more than twenty copies per cell (6% and 3%, respectively). Two independently amplified cDNAs from the single-cell-level diluted total RNA showed high reproducibility of this method ($R^2 = 0.89$, 83% of genes detected within 3.5-fold difference). The comparison between the amplified and nonamplified cDNAs showed faithful preservation of the transcript abundance ($R^2 = 0.7$, 84% genes within 3.5-fold difference). These systematic evaluations clearly demonstrated the reliability and quantitative performance of the new amplification method, and stringently defined its dynamic range [4].

This method has been used for expression profiling in various biological studies [4-8, 10-13, 16, 17]: the next three sections will focus on some of the applications to developmental biology in particular. In addition, this method has recently been applied to an mRNA-sequencing whole-transcriptome analysis with a slight modification, and shown to be able to identify various previously unidentified transcript variants, including splice variants [5]. This approach was used to analyze the impacts of Dicer1 and Ago2 mutation in oocytes [5] and the process of embryonic stem (ES) cell deviation from the inner cell mass (ICM) of blastocyst [10, 11]. This expands the application range of this cDNA amplification method, enabling an analysis of post transcriptional regulation in single cells.

Inner cell mass differentiation of mouse blastocysts

To assess the performance of the new single-cell cDNA amplification method combined with oligonucleotide microarrays, it was first applied to the ICM of early mouse blastocysts at E3.5 [4]. The ICM cells form a morphologically homogeneous population at E3.5, and one day later (E4.5), some of them are differentiated into primitive endoderm (PE) while the others are differentiated into pluripotent primitive ectoderm (PEct). E3.5 and E4.5 mouse ICMs were dissociated into single cells, randomly isolated, and subjected to cDNA amplification with the new method. Extensive Q-PCR and genome-wide analysis with a oligonucleotide microarray, Affymetrix GeneChip Mouse Genome 430 2.0, demonstrated that the morphologically homogeneous ICM cells at E3.5 had already started a molecular segregation to PE and PEct. Notably, this early segregation has been also demonstrated with different approaches [54, 55], underscoring the reliability of this amplification method.

Neurogenesis in mouse cerebral cortex development

The mammalian cerebral cortex develops from the epithelial cells of telencephalon, the most anterior part of the neural tube. The neural progenitors are composed of at least two subtypes based on mitotic positions; the apical progenitor in the ventricular zone (VZ) and the basal progenitor in the subventricular zone (SVZ), which produce distinct sets of neurons and progenitors through symmetric or asymmetric cell divisions. The variation of molecular properties of the neural progenitors, which may underlie the cellular diversity of the cortical neurons, has long been elusive. A single-cell microarray analysis with the new method, combined with extensive *in situ* hybridization analyses, revealed that a large part of the mouse E14 cerebral cells (~20% of the cells examined) formed a previously unidentified population with no known combination of marker gene expressions [12]. With the genome-wide transcriptional profiles, however, these cells turned out to be very similar to the apical progenitors, indicating the diversity of this undifferentiated progenitor population [12].

Specification and development of mouse primordial germ cells

As mentioned above, the molecular properties of the nascent PGCs, the dynamics of their transcriptome in the specification period, and the impact of the deficiency of a key gene, *Blimp1*, have been fully elucidated at the single-cell resolution [13, 15]. Single-cell cDNAs were generated from randomly isolated PGCs during the first forty-eight hours of the specification stage at six- to twelve-hour intervals (E6.25~E8.25) (Fig. 2). A comparative genome-wide analysis identified more than eight hundred genes differentially expressed between PGCs and surrounding somatic mesoderm cells, among which *Prdm14* and *Ap2γ/Tcfap2c* have since been identified as essential factors for PGC development [15, 17, 56]. A time-course analysis of transcriptome further revealed that the nascent PGCs started to show a profile quite similar to neighboring somatic cells under a strong influence of the mesoderm formation signal [13, 15]. However, the germline cells with high *Blimp1* expression then drastically repressed the proceeding somatic developmental programs (e.g., Hox cluster activation, DNA methyltransferase machinery), reactivated the genes associated to pluripotency (e.g., *Sox2*, *Nanog*), and up-regulated PGC-associated genes (*Prdm14*, *stella/Pgc7/Dppa3*, *Nanos3*, *Kit*, *Ap2γ*). These genome-wide transition events occurred simultaneously in a short time span at around E7.0, causing the formation of at least two populations of PGCs with distinct molecular properties, which were identifiable only through the single-cell analysis.

Furthermore, single-cell microarray analyses of the *Blimp1* mutant PGC-like cells demonstrated that the mutant cells no longer resist the mesoderm formation signal, with almost all of the somatic programs activated at a level indistinguishable from the neighboring somatic cells (Fig. 2). Interestingly, even in these cells, about half of the genes acquired in the wild-type PGCs were more or less up-regulated, whereas genes highly specific to PGCs were shown to tend to depend on *Blimp1*. The degree of the impact of *Blimp1* deficiency on the PGC-specific genes was predictable from the expression patterns in the wild-type PGCs; if the expression of a gene is highly correlated with the level of *Blimp1* in the wild type, it tends to be highly impaired in *Blimp1* mutants, and vice versa [13]. This may imply a direct role of *Blimp1* in the activation of these genes. The PGC-specific genes were apparently stochastically expressed in the *Blimp1* mutants, and they were not expressed within individual cells simultaneously [13]. These lines of evidence identified *Blimp1* as the dominant repressor of the somatic developmental program, and an activator and the coordinator of the PGC-specific genetic program [13].

The genes up-regulated under the *Blimp1* mutation indicated the existence of *Blimp1*-independent genetic pathways in the germline specification [13], which are the consequence of the PGC induction signal by *Bmp4* from the extraembryonic ectoderm [44, 57, 58]. These pathways include *Prdm14*, another PR domain-containing protein

essential for the PGC specification [17]. With single-cell cDNA analyses, *Prdm14* mutant PGC-like cells were shown to normally repress somatic mesoderm genes and up-regulate multiple PGC-associated genes, with the exception of *Sox2* and a few other genes[17]. Consistent with the failure of *Sox2* reacquisition, the *Prdm14* mutant cells failed to derive embryonic germ (EG) cells[59], indicating that *Prdm14* was an essential factor for reacquisition of the potential pluripotency of PGCs [17].

After the specification stage, PGCs start to migrate, and enter developing gonads at around E10. *Nanog*, which functions as a gateway and safeguard for pluripotency [60, 61], has been shown to be essential for maintenance of PGCs [60]. In a study using highly controlled gene knockdown (KD) systems, *Nanog* was shown to be required for the survival of migrating PGCs until E10.5, but dispensable once they had arrived at the gonads. A single-cell microarray analysis revealed that, although there were many genes influenced by *Nanog* KD, the vast majority of them were not classified as genes involved in the development of known cell types, suggesting that the *Nanog* KD was unlikely to lead PGCs to transdifferentiation [16]. Moreover, many genes associated with PGC development and pluripotency were normally expressed in the *Nanog* KD PGC-like cells, except for an RNA-binding protein required for germline development, *Tiar/Tial1*[62]. Although it has not yet been evident that the decreased PGC number in the *Tiar* mutants is due to apoptotic cell death[62], down-regulation of this gene may in part cause the phenotype of *Nanog* KD in PGC development.

Collectively, these studies have clearly demonstrated that the single-cell microarray analysis is a powerful tool for identification of genes specifically expressed in particular cell populations, clarification of molecular properties of individual cells, discovery of previously unidentified, transient cell populations, and elucidation of the impact of gene alterations in depth.

Perspective

Along with the improvement of cDNA amplification methods, it has become feasible to solve biologically important questions through quantitative measurement of genome-wide expression profiles at the single-cell level. To date, this method has been applied mainly to embryonic development. A future challenge will be to apply the method to more complex, dynamic systems, including homeostasis, growth, regeneration, and diseases of adult tissues. Recent studies have revealed balanced, dynamic behaviors and interactions of mammalian tissue stem cells and their niches in physiological and pathological processes (for review, see [63]). Quantitative single-cell analyses would contribute to the characterization and better understanding of such complex systems, with a profiling of transitions among multiple stem cell states, quantification of responses to environmental signals, and detection of onsets of aberration.

Further improvements with respect to the quality and simplicity of the present amplification method will be possible by optimizing the procedure. Automation with appropriate devices will improve the throughput of this method, which is currently based on manual labor. Such improvements may enable more precise and/or high-throughput analyses of various biological questions and may open a possibility of addressing even more sophisticated issues that deal with the behavior of individual cells in a cell population.

Acknowledgements

This study was supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, by JST, CREST, and by the Takeda Science Foundation.

References and recommended reading

1. Lockhart DJ, Dong H, Byrne MC, Follett MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, et al.: **Expression monitoring by hybridization to high-density oligonucleotide arrays.** *Nat Biotechnol* 1996, **14**: 1675-80.
2. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10**: 57-63.
3. 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* 2007, **2**: 739-52.
4. Kurimoto K, Yabuta Y, Ohinata Y, Ono Y, Uno KD, Yamada RG, Ueda HR, Saitou M: **An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis.** *Nucleic Acids Res* 2006, **34**: e42.
5. ●● Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, et al.: **mRNA-Seq whole-transcriptome analysis of a single cell.** *Nat Methods* 2009, **6**: 377-82.

This study modified the new exponential amplification method so that amplified cDNA is suitable for the massively parallel sequencing using the Applied Biosystems SoLiD system. The authors applied this modified method to single *Dicer1* and *Ago2* mutant oocytes, and identified numerous aberrantly regulated genes, illustrating the global importance of small RNAs for oogenesis.

6. Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA: **Maternal microRNAs are essential for mouse zygotic development.** *Genes Dev* 2007, **21**: 644-8.
7. Hayashi K, Lopes SM, Tang F, Surani MA: **Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states.** *Cell Stem Cell* 2008, **3**: 391-401.
8. Kobayashi T, Mizuno H, Imayoshi I, Furusawa C, Shirahige K, Kageyama R: **The cyclic gene *Hes1* contributes to diverse differentiation responses of embryonic stem cells.** *Genes Dev* 2009, **23**: 1870-5.
9. Hayashi K, Surani MA: **Self-renewing epiblast stem cells exhibit continual delineation of germ cells with epigenetic reprogramming in vitro.** *Development* 2009, **136**: 3549-56.
10. Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, Surani MA: **Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis.** *Cell Stem Cell* 2010, **6**: 468-78.
11. Tang F, Barbacioru C, Nordman E, Li B, Xu N, Bashkirov VI, Lao K, Surani MA: **RNA-Seq analysis to capture the transcriptome landscape of a single cell.** *Nat Protoc*

2010, **5**: 516-35.

12. • Kawaguchi A, Ikawa T, Kasukawa T, Ueda HR, Kurimoto K, Saitou M, Matsuzaki F: **Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis.** *Development* 2008, **135**: 3113-24.

Using single-cell microarray analysis, this study identified a previously unidentified, large population of cerebral cells that were very similar to the undifferentiated progenitors in terms of genome-wide gene expression profiles.

13. •• Kurimoto K, Yabuta Y, Ohinata Y, Shigeta M, Yamanaka K, Saitou M: **Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice.** *Genes Dev* 2008, **22**: 1617-35.

This study revealed the transcriptome of mouse PGCs in wild-type and *Blimp1* mutant mice at the specification stage with the new exponential single-cell cDNA amplification method combined with the Affymetrix GeneChip oligonucleotide microarrays. This is the first report of gene expression dynamics and the impact of the loss of an essential gene in a differentiating cell lineage at the single-cell level in the genome-wide manner.

14. Kurimoto K, Yamaji M, Seki Y, Saitou M: **Specification of the germ cell lineage in mice: a process orchestrated by the PR-domain proteins, Blimp1 and Prdm14.** *Cell Cycle* 2008, **7**: 3514-8.
15. Yabuta Y, Kurimoto K, Ohinata Y, Seki Y, Saitou M: **Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling.** *Biol Reprod* 2006, **75**: 705-16.
16. Yamaguchi S, Kurimoto K, Yabuta Y, Sasaki H, Nakatsuji N, Saitou M, Tada T: **Conditional knockdown of Nanog induces apoptotic cell death in mouse migrating primordial germ cells.** *Development* 2009, **136**: 4011-20.
17. Yamaji M, Seki Y, Kurimoto K, Yabuta Y, Yuasa M, Shigeta M, Yamanaka K, Ohinata Y, Saitou M: **Critical function of Prdm14 for the establishment of the germ cell lineage in mice.** *Nat Genet* 2008, **40**: 1016-22.
18. • Taniguchi K, Kajiyama T, Kambara H: **Quantitative analysis of gene expression in a single cell by qPCR.** *Nat Methods* 2009, **6**: 503-6.

This study established a Q-PCR method featuring a reusable single-cell cDNA library immobilized on beads for measuring the expression of multiple genes in a single cell.

19. Diercks A, Kostner H, Ozinsky A: **Resolving cell population heterogeneity: real-time PCR for simultaneous multiplexed gene detection in multiple single-cell samples.** *PLoS One* 2009, **4**: e6326.
20. Gibson JD, Jakuba CM, Boucher N, Holbrook KA, Carter MG, Nelson CE: **Single-cell transcript analysis of human embryonic stem cells.** *Integrative Biology* 2009, **1**: 540-551.

21. Bengtsson M, Hemberg M, Rorsman P, Stahlberg A: **Quantification of mRNA in single cells and modelling of RT-qPCR induced noise.** *BMC Mol Biol* 2008, **9**: 63.
22. Melin J, Quake SR: **Microfluidic large-scale integration: the evolution of design rules for biological automation.** *Annu Rev Biophys Biomol Struct* 2007, **36**: 213-31.
23. Warren L, Bryder D, Weissman IL, Quake SR: **Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR.** *Proc Natl Acad Sci U S A* 2006, **103**: 17807-12.
24. Diehn M, Cho RW, Lobo NA, Kalisky T, Dorie MJ, Kulp AN, Qian D, Lam JS, Ailles LE, Wong M, et al.: **Association of reactive oxygen species levels and radioresistance in cancer stem cells.** *Nature* 2009, **458**: 780-3.
25. Toriello NM, Douglas ES, Thaitrong N, Hsiao SC, Francis MB, Bertozzi CR, Mathies RA: **Integrated microfluidic bioprocessor for single-cell gene expression analysis.** *Proc Natl Acad Sci U S A* 2008, **105**: 20173-8.
26. Ottesen EA, Hong JW, Quake SR, Leadbetter JR: **Microfluidic digital PCR enables multigene analysis of individual environmental bacteria.** *Science* 2006, **314**: 1464-7.
27. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P: **Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst.** *Dev Cell* 2010, **18**: 675-85.
28. Brady G, Iscove NN: **Construction of cDNA libraries from single cells.** *Methods Enzymol* 1993, **225**: 611-23.
29. Brady G, Barbara M, Iscove N: **Representative in vitro cDNA amplification from individual hemopoietic cells and colonies.** *Methods Molec. Cell. Biol.* 1990, **2**: 17-25.
30. Eberwine J, Yeh H, Miyashiro K, Cao Y, Nair S, Finnell R, Zettel M, Coleman P: **Analysis of gene expression in single live neurons.** *Proc Natl Acad Sci U S A* 1992, **89**: 3010-4.
31. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH: **Amplified RNA synthesized from limited quantities of heterogeneous cDNA.** *Proc Natl Acad Sci U S A* 1990, **87**: 1663-7.
32. Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP: **Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome.** *Development* 2003, **130**: 889-900.
33. Hamatani T, Carter MG, Sharov AA, Ko MS: **Dynamics of global gene expression changes during mouse preimplantation development.** *Dev Cell* 2004, **6**: 117-31.
34. Sugino K, Hempel CM, Miller MN, Hattox AM, Shapiro P, Wu C, Huang ZJ, Nelson SB: **Molecular taxonomy of major neuronal classes in the adult mouse forebrain.** *Nat Neurosci* 2006, **9**: 99-107.

35. Kamme F, Salunga R, Yu J, Tran DT, Zhu J, Luo L, Bittner A, Guo HQ, Miller N, Wan J, et al.: **Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity.** *J Neurosci* 2003, **23**: 3607-15.
36. • Sul JY, Wu CW, Zeng F, Jochems J, Lee MT, Kim TK, Peritz T, Buckley P, Cappelleri DJ, Maronski M, et al.: **Transcriptome transfer produces a predictable cellular phenotype.** *Proc Natl Acad Sci U S A* 2009, **106**: 7624-9.

This study revealed dynamics of transcriptome of rat astrocytes converted to neurons with a “transcriptome-induced phenotype remodeling protocol”, at the single-cell level over a ten-day period, using the multiple-round linear amplification method.

37. Tougan T, Okuzaki D, Nojima H: **Chum-RNA allows preparation of a high-quality cDNA library from a single-cell quantity of mRNA without PCR amplification.** *Nucleic Acids Res* 2008, **36**: e92.
38. Iscove NN, Barbara M, Gu M, Gibson M, Modi C, Winegarden N: **Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA.** *Nat Biotechnol* 2002, **20**: 940-3.
39. Subkhankulova T, Livesey FJ: **Comparative evaluation of linear and exponential amplification techniques for expression profiling at the single-cell level.** *Genome Biol* 2006, **7**: R18.
40. Dulac C, Axel R: **A novel family of genes encoding putative pheromone receptors in mammals.** *Cell* 1995, **83**: 195-206.
41. Saitou M, Barton SC, Surani MA: **A molecular programme for the specification of germ cell fate in mice.** *Nature* 2002, **418**: 293-300.
42. Ginsburg M, Snow MH, McLaren A: **Primordial germ cells in the mouse embryo during gastrulation.** *Development* 1990, **110**: 521-8.
43. Lawson KA, Meneses JJ, Pedersen RA: **Clonal analysis of epiblast fate during germ layer formation in the mouse embryo.** *Development* 1991, **113**: 891-911.
44. Ohinata Y, Payer B, O'Carroll D, Ancelin K, Ono Y, Sano M, Barton SC, Obukhanych T, Nussenzweig M, Tarakhovsky A, et al.: **Blimp1 is a critical determinant of the germ cell lineage in mice.** *Nature* 2005, **436**: 207-13.
45. Tietjen I, Rihel JM, Cao Y, Koentges G, Zakhary L, Dulac C: **Single-cell transcriptional analysis of neuronal progenitors.** *Neuron* 2003, **38**: 161-75.
46. Chiang MK, Melton DA: **Single-cell transcript analysis of pancreas development.** *Dev Cell* 2003, **4**: 383-93.
47. Hayashi T, Shibata N, Okumura R, Kudome T, Nishimura O, Tarui H, Agata K: **Single-cell gene profiling of planarian stem cells using fluorescent activated cell sorting and its "index sorting" function for stem cell research.** *Development Growth*

- & Differentiation* 2010, **52**: 131-144.
48. Jensen KB, Watt FM: **Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence.** *Proc Natl Acad Sci U S A* 2006, **103**: 11958-63.
 49. Ramos CA, Bowman TA, Boles NC, Merchant AA, Zheng Y, Parra I, Fuqua SA, Shaw CA, Goodell MA: **Evidence for diversity in transcriptional profiles of single hematopoietic stem cells.** *PLoS Genet* 2006, **2**: e159.
 50. Saito H, Kubota M, Roberts RW, Chi Q, Matsunami H: **RTP family members induce functional expression of mammalian odorant receptors.** *Cell* 2004, **119**: 679-91.
 51. Osawa M, Egawa G, Mak SS, Moriyama M, Freter R, Yonetani S, Beermann F, Nishikawa S: **Molecular characterization of melanocyte stem cells in their niche.** *Development* 2005, **132**: 5589-99.
 52. Hartmann CH, Klein CA: **Gene expression profiling of single cells on large-scale oligonucleotide arrays.** *Nucleic Acids Res* 2006, **34**: e143.
 53. Esumi S, Wu SX, Yanagawa Y, Obata K, Sugimoto Y, Tamamaki N: **Method for single-cell microarray analysis and application to gene-expression profiling of GABAergic neuron progenitors.** *Neurosci Res* 2008, **60**: 439-51.
 54. Chazaud C, Yamanaka Y, Pawson T, Rossant J: **Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway.** *Dev Cell* 2006, **10**: 615-24.
 55. Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis AK: **Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst.** *Development* 2008, **135**: 3081-91.
 56. Weber S, Eckert D, Nettersheim D, Gillis AJ, Schafer S, Kuckenberg P, Ehlermann J, Werling U, Biermann K, Looijenga LH, et al.: **Critical function of AP-2 gamma/TCFAP2C in mouse embryonic germ cell maintenance.** *Biol Reprod* 2010, **82**: 214-23.
 57. Ohinata Y, Ohta H, Shigeta M, Yamanaka K, Wakayama T, Saitou M: **A signaling principle for the specification of the germ cell lineage in mice.** *Cell* 2009, **137**: 571-84.
 58. Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, Wright CV, Korving JP, Hogan BL: **Bmp4 is required for the generation of primordial germ cells in the mouse embryo.** *Genes Dev* 1999, **13**: 424-36.
 59. Matsui Y, Zsebo K, Hogan BL: **Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture.** *Cell* 1992, **70**: 841-847.
 60. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A: **Nanog safeguards pluripotency and mediates germline**

- development.** *Nature* 2007, **450**: 1230-4.
61. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A: **Nanog is the gateway to the pluripotent ground state.** *Cell* 2009, **138**: 722-37.
62. Beck AR, Miller IJ, Anderson P, Streuli M: **RNA-binding protein TLAR is essential for primordial germ cell development.** *Proc Natl Acad Sci U S A* 1998, **95**: 2331-6.
63. Quante M, Wang TC: **Stem cells in gastroenterology and hepatology.** *Nat Rev Gastroenterol Hepatol* 2009, **6**: 724-37.

Figure legends

Figure 1

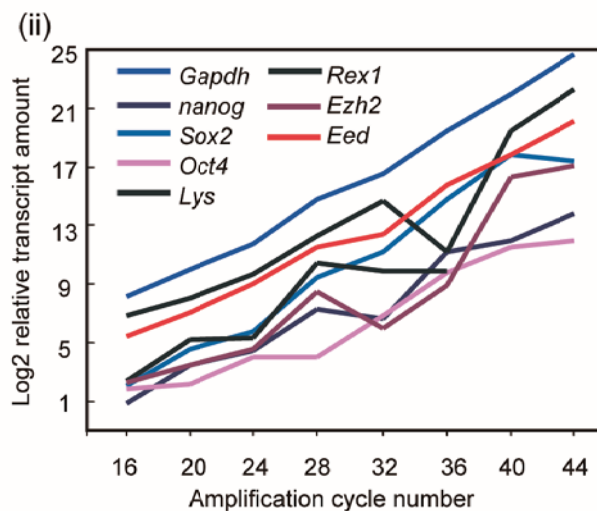
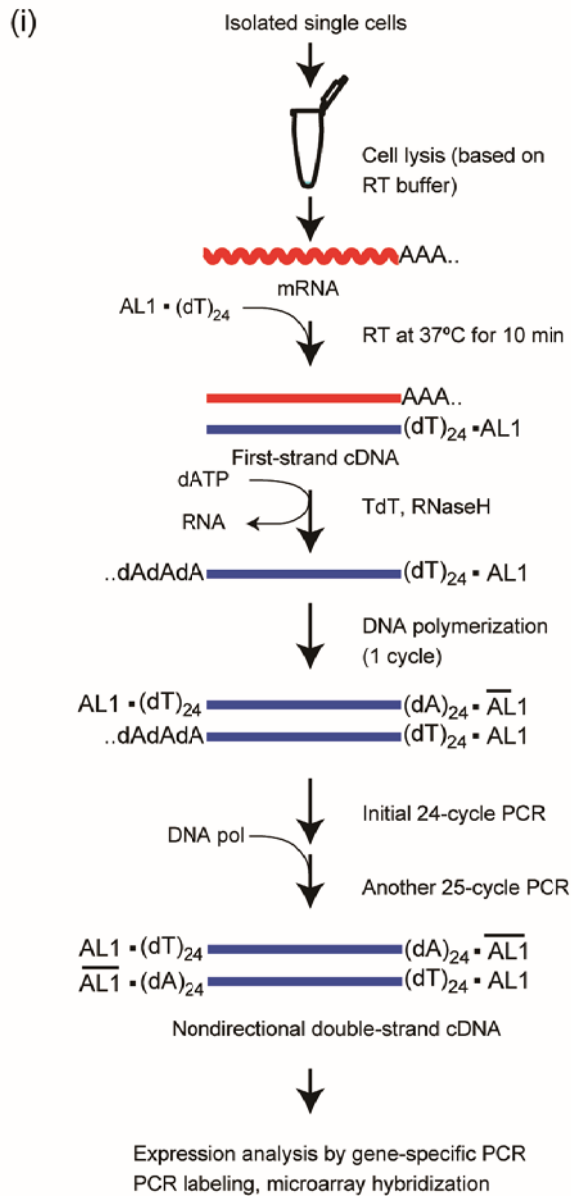
Single-cell cDNA amplification methods by Brady et al. [28, 29] (A) and by Kurimoto et al. [3, 4] (B). (i) Schematic representation of the methods. (ii) Performance of cDNA amplification. The $\log_2$ relative transcript amounts of the indicated genes in ES total RNA (10 pg) are plotted against the cycle numbers of cDNA amplification. If cDNAs are amplified exponentially, these two values exhibit a linear relationship. The efficiency of amplification is measured by the slopes of the regression lines of exponential amplification (0.49 in (A) and 0.91 in (B), in average, between 16 and 28 cycles); when the PCR doubles cDNA at each cycle, the slope is expected to be 1.0. The data indicate that the amplification in (B) reaches more than 90% efficiency of the ideal value, and is 1.86 times efficient as that in (A).

Figure 2

Single-cell microarray analysis of mouse PGC specification. (A-C) Schematic representation of single-cell isolation. (A) Mouse embryos from which single cells are isolated. PGCs and their progenitors are represented with red circles. (B) An embryonic fragment at E7.25 that contains a PGC cluster. (C) Dissociated single cells from the embryonic fragment depicted in (B). (D) Screening of amplified cDNA libraries generated from the cells isolated in (C), using gene-specific PCR of *Blimp1* and *stella*. The *Blimp1*-positive, *stella*-positive cells are PGCs at E7.25. (E) Heat map representation of genes differentially expressed between PGC and somatic neighbors at E7.25[13]. The gene expression levels are color coded from blue (estimated copy number 20) to red (estimated copy number 500).

Fig. 1

(A) Method of Brady et al.



(B) Method of Kurimoto et al.

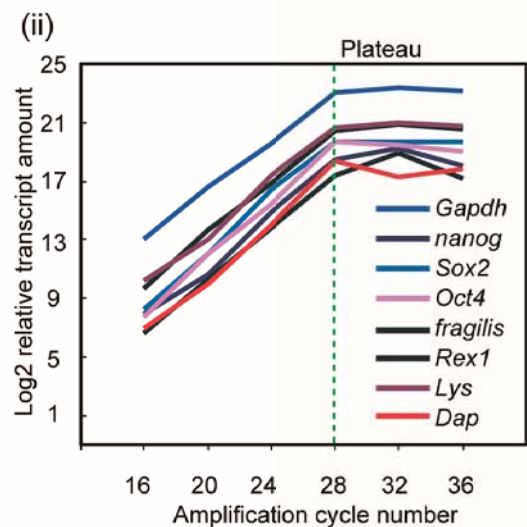
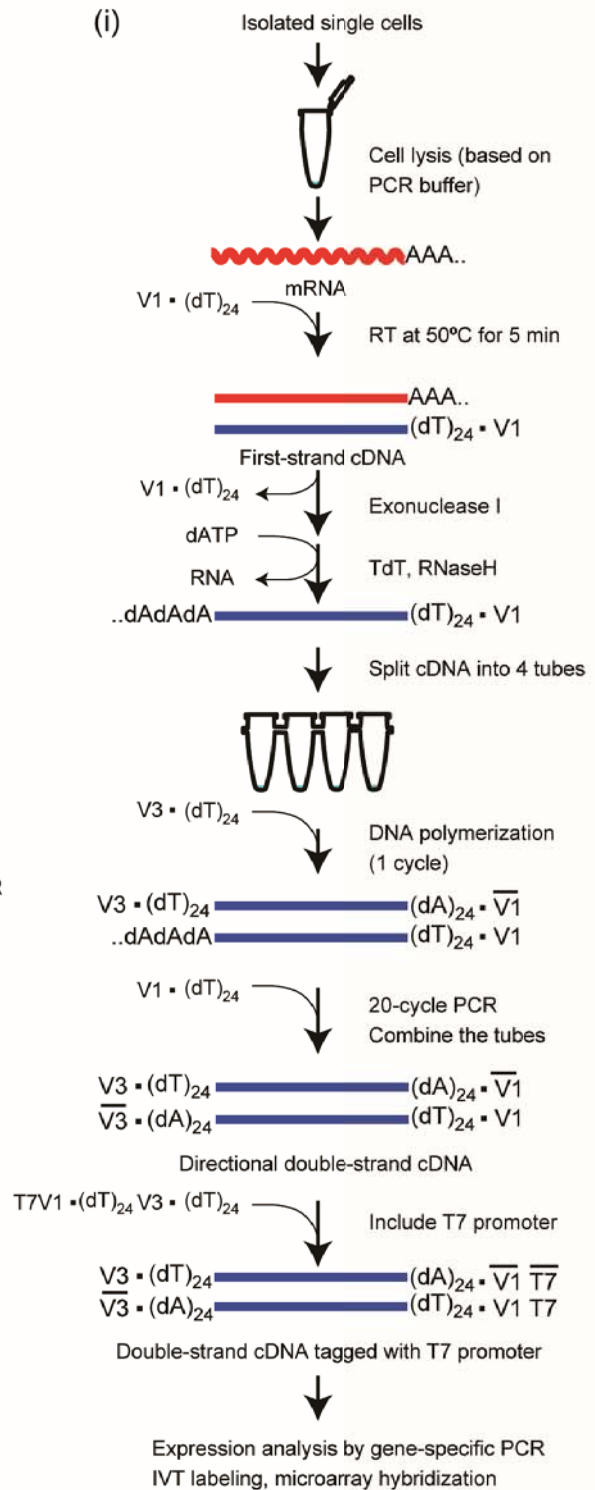


Fig 2

