Science@Life

mRNA-Seq Whole Transcriptome Analysis of a Single Cell

Kai Q. Lao[1], Fuchou Tang[2], Catalin Barbacioru[1], Yangzhou Wang[1], Ellen Nordman[1], Clarence Lee[1], Nanlan Xu[1], David Wang[1], Brain Tuch[1], John Bodeau[1], Asim Siddiqui[1], Azim Surani[2], 1Genetic Systems, Applied Biosystems, Foster City, CA 94404, USA, [2]Wellcome Trust/Cancer Research UK Gurdon Institute of Cancer and Developmental Biology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QN, UK

ABSTRACT

Next generation sequencing technology is a new powerful tool for transcriptome analysis. However, under certain conditions only a small amount of material is available, which requires more sensitive techniques that can preferably work down to the single cell level. Here we describe a single cell digital gene expression-profiling assay. Using only a single mouse blastomere, our mRNA-Seq assay detected the expression of 75% (5,270) more genes than microarray techniques, and identified 1,753 previously unknown splice junctions called by at least 5 reads. Moreover, 8 - 19% of the genes with multiple known transcript isoforms express at least two isoforms in the same blastomere or oocyte, which unambiguously demonstrates the complexity of the transcript variants at the whole genome scale within individual cells. Finally, for DICER$^{-/-}$ and AGO2$^{-/-}$ oocytes, we showed that 1,696 and 1,553 genes respectively were abnormally upregulated compared to wild-type controls, with 619 of them in common.

INTRODUCTION

By analyzing the transcriptome at a spectacular and unprecedented depth and accuracy, thousands of new transcript variants/isoforms were unambiguously found to be expressed in mammalian tissues or organs. These advances have greatly accelerated our understanding of the complexity of gene expression regulation and networks for mammalian cells. The technique usually requires mg amounts of total RNA for analysis, which corresponds to hundreds of thousands of cells. However, under certain conditions, it is not practically possible to get such amounts of materials for analysis, e.g. for early embryonic development. In fact, during mouse early development, when the founder population of a germ line, primordial germ cells (PGCs), are specified and have just emerged, there are only around 30 PGC cells in an embryo.

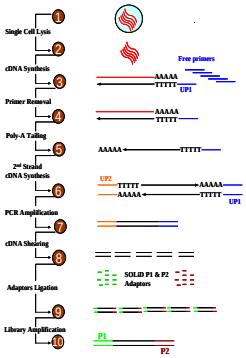
On the other hand, even for in vitro cultured stem cells, for which the cell amount available for analysis is unlimited, there are limitations. For example, mouse embryonic stem (ES) cells, probably the most thoroughly analyzed type of stem cell, were found to contain multiple subpopulations with strong differences of both gene expression and physiological function. A more sensitive next-generation sequencing assay is needed to illuminate these crucial developmental processes and stem cell biology—ideally, an assay capable of single cell resolution.

Here we modified a single cell whole transcriptome amplification method to permit amplification of cDNAs as long as 3kb in an efficient and unbiased manner (8, 9). We combined this modified single cell cDNA amplification method with Applied Biosystems' next generation sequencing (NGS) technology, the SOLiD™ System, to set up a single cell whole transcriptome assay.

Detecting gene expression at the resolution of a single cell enables us to ask fundamental biological questions that were previously not possible, especially in the field of early embryonic development, and allows us to understand biology at the smallest functional unit of any organism—a single cell.

MATERIALS AND METHODS:

Figure 1. Reaction Schemes



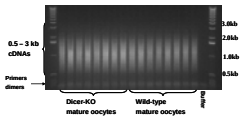
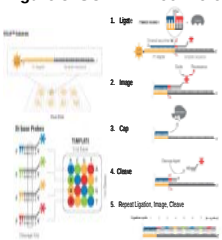
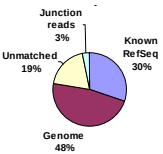


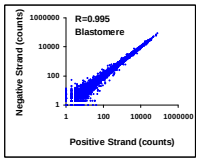
Figure 3. SOLiD™ Technology

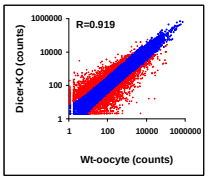


RESULTS

Figure 4. SOLiD™ matching summary for mouse blastomeres in 4-cell embyros







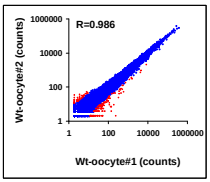
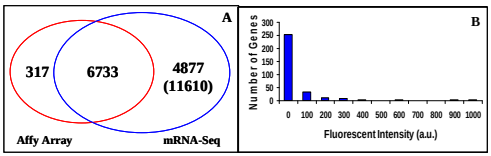


Figure 6: We compared our SOLiD™ quenching data with those from microarrays of about 80 pooled four-cell stage embryos (320 blastomeres) and found that 6,733 genes detected by Affymetrix GeneChip Mouse Genome 430 2.0 Array were also detected by SOLiD™ sequencing. SOLiD was unable to detect only 317 genes that the were detected by the microarray. However, these genes had relatively low expression, and they were likely detected due to cross-hybridization. SOLiD™ sequencing detected 4,877 more genes than the microarray.



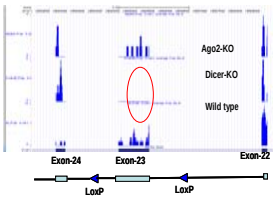
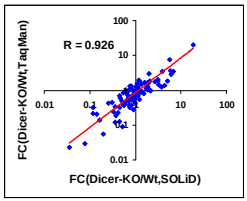


Figure 9: Correlation plots of the fold changes that are determined by SOLiD™ reads and real-time PCR



Conclusions:

In summary, we have established a SOLiD™ sequencing-based gene expression profiling assay at single cell resolution. We proved that thousands of genes express two or more transcript variants in the same cell. We also proved that in Dicer-knockout mature oocytes, the transcripts of many of transposons and repeat elements are abnormally upregulated. This single cell sequencing assay will greatly facilitate our understanding of the transcriptome's complexity during mammalian development, especially in the fields of stem cell and early embryonic development.

Acknowledgement:

We would like to thank Caroline Lee, Melissa A Barker, Roland Wicki, Cinna Monighetti, Francisco M. De La Vega, and Neil A. Straus for excellent technical help.

Trademarks: