Single Cell Transcriptome Analysis on the SOLiD™ System

Kai Q. Lao[1], Fuchou Tang[2], Catalin Barbacioru[1], Yangzhou Wang[1], Ellen Nordman[1], Clarence Lee[1], Nanlan Xu[1], Xiaohui Wang[1], John Bodeau[1], M. Azim Surani[2],
[1]Applied Biosystems, Genetic Analysis Business Unit; [2]University of Cambridge

ABSTRACT:

Here we developed a sequencing-based gene expression profiling assay at single cell resolution by combining a modified single cell whole transcriptome amplification method with the next generation sequencing technique, the SOLiD™ System. Using this assay, we showed that blastomeres in a four-cell stage embryo have similar gene expression, which is compatible with the fact that they have similar developmental potential. We proved that, compared to cDNA microarray technique, our single cell cDNA SOLiD sequencing assay can detect expression of thousands of more genes. Moreover, for the genes detected by both microarray and SOLiD sequencing, our assay detected new transcript variants for a large proportion of them, which unambiguously confirms at single cell resolution that the transcriptome complexity is higher than traditionally expected. Finally, by using our assay to Dicer knockout and Ago2 knockout oocytes, we showed that a significant amount of transposons were abnormally upregulated in Dicer/Ago2 knockout mature oocytes compared with wildtype controls.

INTRODUCTION:

By analyzing the transcriptome at spectacular and unprecedented depth and accuracy, thousands of new transcript variants/isoforms were unambiguously found expressed in mammalian tissues or organs. These advances greatly accelerate our understanding of the complexity of gene expression regulation and networks for mammalian cells. The technique usually needs μg amounts of total RNA for analysis, which corresponds to hundreds of thousands of cells. However, under certain conditions, it is practically not 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 germ line, primordial germ cells (PGCs) just specified and emerged, there are only around 30 PGC cells in an embryo. On the other hand, even for in vitro cultured stem cell, 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 cells during past 27 years, were found to contain multiple subpopulations with strong differences of both gene expression and physiological function. All together, a more sensitive next-generation sequencing assay, ideally an assay working to single cell resolution is needed for these crucial developmental processes and stem cell biology.

Here we modified a single cell whole transcriptome amplification method to make it permissive to amplify 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.

We prove that it is feasible to get gene expression profiles at single cell resolution, which enables us to ask fundamental biological questions previously not possible, especially in the field of early embryonic development, and to understand biology at the single cell resolution, which is the uniform functional unit of any organism.

MATERIALS AND METHODS:

Figure 1. Reaction Schemes

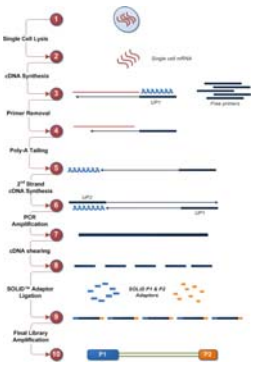


Figure 2. cDNA products

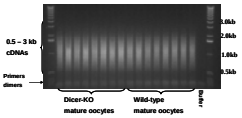
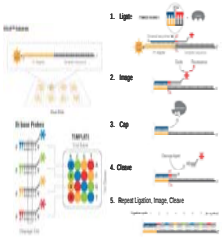


Figure 3. SOLiD™ Technology



RESULTS:

Figure 4. SOLiD™ Matching summary for two blastomeres in 4-cell embyros

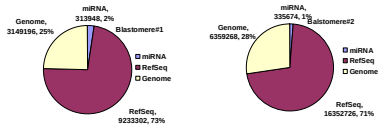


Figure 5. SOLiD results for two blastomeres in 4-cell embyros

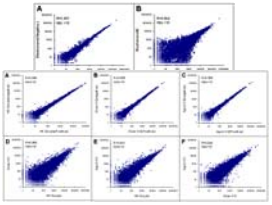


Figure 6. SOLiD™ results for Wt, Dicer-KO, and Ago2-KO oocytes

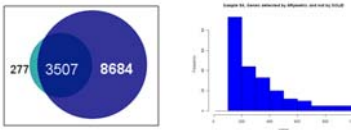


Figure 6: We compared our SOLiD sequencing data with those from microarrays of about 80 pooled four-cell stage embryos (320 blastomeres) and found that 93% (3,507 genes) of the genes detected by Affymetrix GeneChip Mouse Genome 430 2.0 Array were also detected by SOLiD sequencing. The 7% transcripts (277 genes) detected by microarray but missed by SOLiD were relatively low expressed genes, and could result from cross-hybridization. SOLID sequencing detects 8,684 more expressed genes compared with microarray.

Figure 7. Coverage of exons for Pou5f1 (Oct4)

Figure 8. Mouse Dicer locus on Chromosome 12

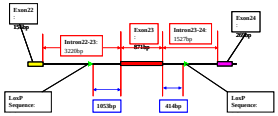


Figure 9. Coverage of Dicer Exon 23

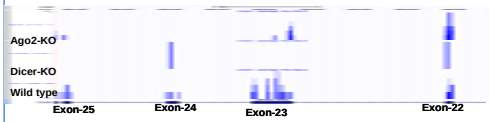


Figure 10. TaqMan® Validation Experiments

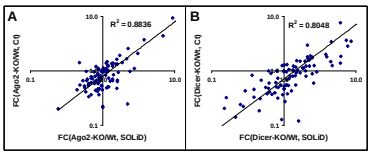


Figure 10: the correlation plots of the fold changes that are determined by SOLiD reads and real-time PCR, (A) Dicer-KO/Wt-oocyte and (B) Ago2-KO/Wt-oocyte.

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 a same cell. We also proved that in Dicer knockout mature oocytes, the transcripts of a lot of transposons and repeat elements are abnormally upregulated. This single cell sequencing assay will greatly facilitate understanding the transcriptome complexity during mammalian development, especially in the fields of stem cells and early embryonic development.

Acknowledgement:

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

Trademarks/Licensing: