Single-Cell Transcriptomics in the Brain

Ditte Lovatt, PhD, Tae Kyung Kim, PhD, Peter Buckley, PhD, Jennifer M. Singh, PhD, and James Eberwine, PhD

Department of Pharmacology
University of Pennsylvania
Philadelphia, Pennsylvania

NOTES

Introduction to Transcriptomics Analysis

Transcriptomics analysis provides valuable qualitative and quantitative information about the global set of messenger RNAs (mRNAs) in a given sample. From such studies, tens of thousands of transcripts can be investigated simultaneously, from which information can be inferred about the sample's biochemical and functional properties. The most frequently used methods today to sample the transcriptome are cDNA microarrays and next-generation RNA sequencing (RNA-Seq) (Lockhart and Winzeler, 2000; Mortazavi et al., 2008).

Microarrays

Microarrays have been used extensively over the past decade in order to investigate the relative expression of specific mRNAs among different cell samples (Table 1). The major drawback of microarrays is that detection is based on the hybridization signal between an oligonucleotide anchored onto the chip and the fluorescently tagged nucleotide sample. This detection principle requires prior knowledge about the nucleotide sequences to be investigated and cannot lead to discoveries about, for instance, novel transcripts, splice variants, and retained introns. Also, the inherently high background noise on most commercial microarrays makes distinguishing between low-abundant RNAs and false-positives difficult, so such information must be validated using other methods, e.g., *in situ* hybridization. However, microarrays do provide a robust method for investigating sequence-specific mRNA abundances and thus, they remain a powerful quantitative method for most transcripts.

RNA sequencing

In contrast, the recent development of RNA-Seq has made unbiased mRNA sequence examination possible and eliminated concern about low-abundant transcripts, false-positives, and prior knowledge about sequence information. As in the examples given below, RNA-Seq makes unbiased sequence discoveries possible and has been applied to solving a variety of problems and discoveries, e.g., retained-introns, alternative splicing, and microRNAs (miRNAs). Although the algorithms for comparative quantification of specific transcripts are still being developed for RNA-Seq, this method is far superior to microarrays and provides a vast amount of detailed sequence information (Wang et al., 2009).

Transcriptome Data: One of a Kind or Just Average?

While transcriptome-generating methods can produce a vast amount of expression data, the interpretation of such data depends entirely on the type of sample. Studying the transcriptome of pools of cells provides a unique window into their biochemistry and function; however, information about cell-to-cell variability is lost. This becomes especially significant if the pool of cells is very heterogeneous, such as in intact brain tissue. The advantage of performing single-cell transcriptomics can easily be appreciated when considering the effect of averaging over the entire pools of cells. Furthermore, several single-cell transcriptome studies have concluded that single cells, even of the same type, are unique, and their subtleties of expression differences can have important biological functions. Limited information exists about the transcriptional differences among single neurons *in vivo*. Even so, one can easily speculate about how single neurons provide an especially unique system with inherent single-cell variability that may account for the differences in functional properties of those neurons and permit plasticity-associated changes.

The transcriptome of mRNA extracted from bulk tissue will give insight into the types of mRNAs species in the tissue. However, information about

Table 1. Comparison of transcriptome analysis methods[a]

Property	Microarrays	RNA-Seq
Quantitative	**	***
Qualitative	**	***
Low-abundant mRNA detection	*	***
Generation of false-positives	***	*
Cost[b]	*	***

[a] Increasing numbers of asterisks signify increased ability to generate the itemized data, e.g., RNA-Seq.
[b] More asterisks signify increased cost.
***, generates more quantitative data than microarrays, **.

NOTES

individual cells where these mRNAs originated from will be lost (Fig. 1). The result of this "averaging effect" will mask information about mRNA species that are present only in a subset of cells in the tissue, as their impact will be diluted. If microarray profiling is used, such subset-specific transcripts may be diluted out even beyond the detection limit; consequently, they will not be detected at all. In addition, the use of transcriptome profiling on mRNA samples from bulk tissue or pools of cells prevents us from distinguishing whether two mRNA species, *X* and Y, that function in the same signaling pathway are coexpressed in the same cell or expressed by different cells. Inferring information about the regulation of existing pathways is therefore also compromised. Although network analysis has been applied to transcriptome data from bulk tissue as a pathway-mapping tool in individual cell types, for the reasons stated above, this method can mistakenly conclude the existence of pathways. That is, *X* may be expressed exclusively by cell *A*, and *Y* may be expressed exclusively by cell *B*, so although they both are detected in the bulk mRNA, they actually cannot interact (Fig. 1).

The averaging effect will also mask information about cell-to-cell variability in the expression level of mRNAs that are expressed by the majority of cells in the tissue (Fig. 1). For instance, the majority of cells may express transcript *Z*, but at the single-cell level, *Z* can be expressed at either a high level, a low level, or not at all. However, such information is masked by the averaging effect, another argument for why single-cell transcriptomics is crucial to employ for questions related to how single cells function and interact with one another.

Single-Cell mRNA Isolation Methods

Clearly, the averaging effect has an important impact on the interpretation of transcriptome data if the tissue contains several unique cells or cell types. This is particularly true for brain tissue, which contains neurons, glia, and vascular cell types. In order to investigate how these cell types differ and how cell-to-cell variability characterizes single cells, one must apply transcriptomics to single cells instead of bulk tissue. The use of dispersed cell cultures could be an option to easily isolate single cells and perform transcriptomics on such samples. Nevertheless, the ultimate capture of a transcriptome is to sample cells from live intact tissue, in which all the synaptic architecture and cell-to-cell interactions are still in place.

Live intact tissue sampling can be accomplished by using acutely cut live brain slices or sampling cells in live animals through cranial windows. However, several technical obstacles prevent single-cell mRNA isolation in intact tissue. Brain tissue is very heterogeneous, and most cell types within the brain

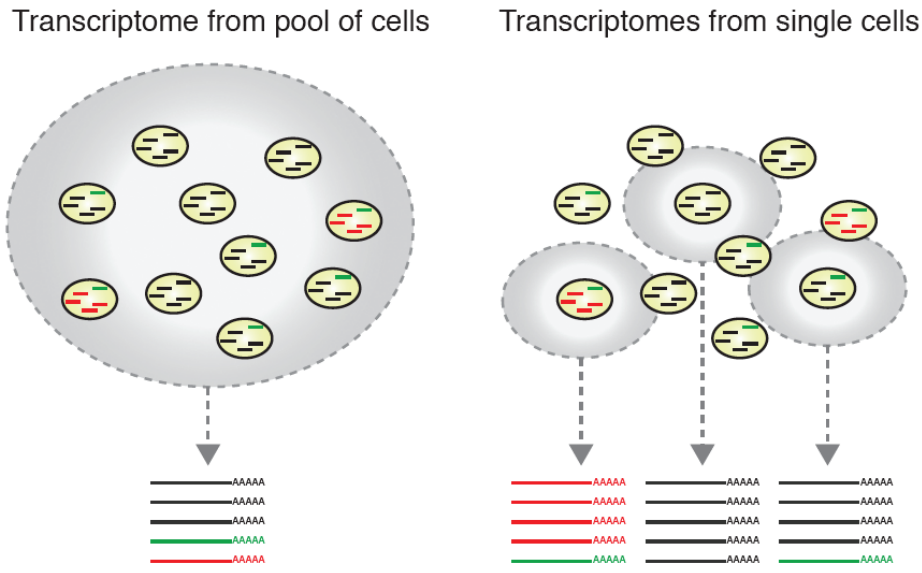

Figure 1. The "averaging effect." Transcriptomes of mRNA from bulk tissue (left) that comprises many cell types is subjected to an averaging effect in which mRNA data from each is averaged. This effect results in the dilution of mRNA species that are only present in a subset of cell transcripts (green and red). In contrast, transcriptomes from single cells (right) precisely report the abundances of each mRNA specie relative to other mRNA species in that particular cell.

NOTES

have polarized and highly branched morphologies that intermingle. This anatomical feature compromises our ability to isolate single cells because the degree of contamination from neighboring cells is significant. Although laser capture microdissection (LCM) is capable of isolating single cells from frozen and fixed tissue (Espina et al., 2006; Tang et al., 2009), these procedures adversely affect RNA quality. Also, LCM adds a significant degree of RNA contamination from neighboring structures to the dissected sample compared with other intact-tissue RNA isolation methods, such as fluorescence-activated cell sorting (FACS), immunopanning, and manual sorting (Okaty et al., 2011).

Perhaps the most successful single-cell RNA isolation method used in intact tissue to date is the micropipette approach, which isolates cytosolic mRNA from whole-cell patched cells by aspirating the cytosol (Surmeier et al., 1996; Martina et al., 1998). This method has been used in a variety of cell types including neurons of the preoptic area of the mouse hypothalamus, pyramidal neurons of the hippocampus, and serotonergnic neurons of the raphe. Indeed, it was using this approach that researchers first demonstrated that hundreds of G-protein coupled receptor (GPCR) genes can be expressed in a single cell. Data such as these offer a rationale for choosing receptor agonists and antagonists, to be used alone or in combination with other drugs (e.g., 5-HT_1 agonists), for physiological testing in selected cells or to study specific behavioral responses.

However, it should be noted that the use of a patch pipette to harvest cells from a live slice will cause mechanical damage to the slice. Thus, the development of an RNA isolation method that could isolate mRNA from single cells (or even subcellular structures, like dendrites) without contamination or induction of injury-related pathways would provide a valuable tool for studying single-cell transcriptomics.

Once mRNA has been isolated from a single cell, it has to be processed to prepare it for the downstream transcriptome method. First the mRNA has to be amplified, since the amount of mRNA from a single cell falls in the hundreds of femptograms-to-picograms range—far below the detection limit of most transcriptome methods, including RNA-Seq. In order to perform quantitative transcriptome analysis, it is crucial to use linear amplification (as opposed to PCR amplification) of the mRNA to maintain the stoichiometry among the different mRNA species (Morris et al., 2011). Linear amplification techniques are well developed. Following amplification of the mRNA to micrograms of amplified RNA (aRNA), either the aRNA need to be processed for microarray or RNA-Seq libraries need to be constructed. Altogether, these processing procedures take approximately one week before the prepped sample can be submitted to a microarray or RNA-Seq facility.

Single-Cell Transcriptomics to Distinguish TIPeR Transcriptome Transfer

Transcriptome profiling of single cells can be used to address a variety of scientific problems. In our lab, we previously used single-cell transcriptomics to validate transcriptome-induced phenotype remodeling (TIPeR)–mediated cell-to-cell transcriptome transfers (Sul et al., 2009). TIPeR is the process by which RNA populations are transferred into single cells to alter or remodel their phenotype. A successfully remodeled TIPeR cell will gradually change its transcriptome through activation and suppression of host-cell transcriptional pathways from the host cell toward that of the desired cell type. This process eventually gives rise to new cellular phenotypes in the TIPeR cells and, potentially, may be used in cell-replacement therapies.

To validate the transfer of the TIPeR cells, poly-A+ tailed mRNA from single TIPeR cells is linearly amplified (Morris et al., 2011) and processed for microarray or RNA-Seq analysis. Microarray data from TIPeR cells can be analyzed using conventional analysis software with modified algorithms that account for the 3' end amplification bias. To deal with the bias, these algorithms extract the second highest intensity values from each probe set and use them for quantitative expression analysis. Once the program obtains expression values, it selects probe sets based on their ability to distinguish the donor from the recipient TIPeR cells. Most often, such probe sets are cell-type–specific transcripts. The analysis results are presented in the form of clustering, differential gene expression profiles, and gene ontology tables to show the difference between TIPeR cells and non-TIPeR cells.

In a previous study, we used the TIPeR approach to transfer the transcriptome of cardiomyocytes into mouse fibroblasts, which converted the phenotype of the fibroblasts into cardiomyocyte-like cells (tCardiomyocyte) (Kim et al., 2011). Besides examining phenotypic signs of successful conversion, the TIPeR process was validated using single-cell transcriptomics, as described above. To this end, we isolated poly-A^+ RNAs from single adult cardiomyocytes, tCardiomyocytes, control cells, and

NOTES

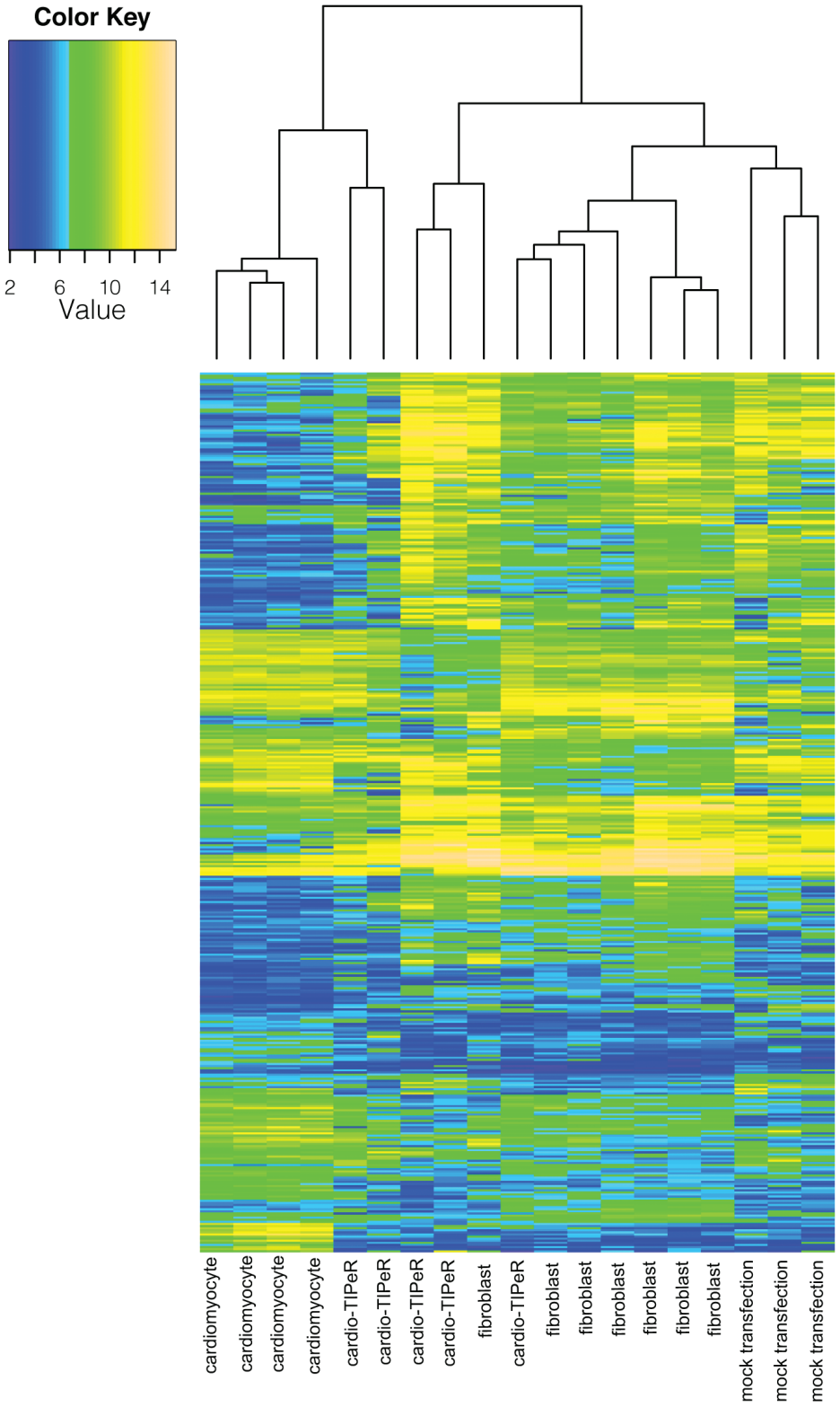

Figure 2. Transcriptomics used to validate TIPeR-mediated phenotype conversion. Global gene expression of tCardiomyocytes is reprogrammed toward adult cardiomyocytes. Dendrogram and heatmap show hierarchical clustering (Euclidean distance, complete linkage) of single cardiomyocytes, fibroblasts, cardio-TIPeR, and mock transfection using the expression values of 418 informative genes.

NOTES

fibroblasts 3–4 weeks after transfection and then amplified and processed the mRNA for microarray and transcriptome analysis. A comparison of the transcriptomes showed that tCardiomyocytes clustered closely with adult cardiomyocytes and far from fibroblasts, as expected for successful TIPeRing. However, not all TIPeR cells clustered with fibroblasts, suggesting that some TIPeR cells are not remodeled completely (Fig. 2). Global gene expression profiles also show that the expression pattern of differentially regulated genes was similar between tCardiomyocytes and adult cardiomyocytes but differed from TIPeR control cells or fibroblasts (Fig. 2). In conclusion, single-cell transcriptomics is a powerful method to validate TIPeR cells.

Single-Cell Transcriptomics for Discovering Novel Transcripts

RNA-Seq has allowed for the unbiased discovery of functionally important, low abundance transcript variants that would have been missed using conventional approaches. For example, a broad class of cytoplasmic intron-retaining transcripts (CIRTs) has been described in the dendrites of primary rat neurons (Buckley et al., 2011). Sequencing libraries constructed from the mRNA of mechanically isolated dendrites revealed not only coding region sequences for dendritically localized transcripts but also a subset of intronic regions located across the genomic organization of their respective genes. Although feasible, using microarray or PCR techniques to screen for these retained introns represents a significant challenge because it requires a priori knowledge of sequences that may be retained and decisions regarding which intronic sequences to target.

Further, these retained intron sequences have been demonstrated as functionally relevant for normal cellular function in neurons. Introducing small interfering RNA (siRNA) that targets a retained intron in *KCNMA1* leads to alterations in the protein distribution of the channel as well as changes to the intrinsic excitability of cells (Bell et al., 2008). Additionally, intron definition (ID) element sequences harbored within retained introns of the *CAMK2B* and *FMR1* transcripts are capable of competing for endogenous targeting machinery for those transcripts and impacting both RNA and protein distribution throughout the cell (Buckley et al., 2011). These results have linked relatively rare transcripts directly to observable endogenous functions. The identification of these sequences would not have been possible without current single-cell techniques like RNA-Seq.

Conclusions

Single-cell biology has undergone dramatic developments over the past decade. Performing single-cell transcriptomics from live cells in complex tissues is still difficult. Nevertheless, the development of novel methods that can isolate RNA from single cells with little resultant tissue damage promises to yield new insights into gene regulation of individual cells and how single cells in multicellular organisms work in concert. As more quantitative single-cell methods are being developed for sampling other "omes" (e.g., the proteome or metabolome), such large-scale data can be correlated to elucidate the link between gene expression and a cell's functional properties. It is through such correlations at the level of the single cell that the complexities of gene-product interactions will be identified. The goal of such research is to rationally modify these biological processes to produce predicted outcomes, including disease therapies.

References

Bell TJ, Miyashiro KY, Sul JY, McCullough R, Buckley PT, Jochems J, Meaney DF, Haydon P, Cantor C, Parsons TD, Eberwine J (2008) Cytoplasmic BK(Ca) channel intron-containing mRNAs contribute to the intrinsic excitability of hippocampal neurons. Proc Natl Acad Sci USA 105:1901–1906.

Buckley PT, Lee MT, Sul JY, Miyashiro KY, Bell TJ, Fisher SA, Kim J, Eberwine J (2011) Cytoplasmic intron sequence-retaining transcripts can be dendritically targeted via ID element retrotransposons. Neuron 69:877–884.

Espina V, Wulfkuhle JD, Calvert VS, VanMeter A, Zhou W, Coukos G, Geho DH, Petricoin EF 3rd, Liotta LA (2006) Laser-capture microdissection. Nat Protoc 1:586–603.

Kim TK, Sul JY, Peternko NB, Lee JH, Lee M, Patel VV, Kim J, Eberwine JH (2011) Transcriptome transfer provides a model for understanding the phenotype of cardiomyocytes. Proc Natl Acad Sci USA 108:11918–11923.

Lockhart DJ, Winzeler EA (2000) Genomics, gene expression and DNA arrays. Nature 405:827–836.

Martina M, Schultz JH, Ehmke H, Monyer H, Jonas P (1998) Functional and molecular differences between voltage-gated K^+ channels of fast-spiking interneurons and pyramidal neurons of rat hippocampus. J Neurosci 18:8111–8125.

Morris J, Singh JM, Eberwine JH (2011) Transcriptome analysis of single cells. J Vis Exp 50:2634.

NOTES

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold *B* (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628.

Okaty BW, Sugino K, Nelson SB (2011) A quantitative comparison of cell-type-specific microarray gene expression profiling methods in the mouse brain. PLoS One 6:e16493.

Surmeier DJ, Song WJ, Yan Z (1996) Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. J Neurosci 16:6579–6591.

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, Lao K, Surani MA (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6:377–382.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63.