

Chapter 2

Decoding the Transcriptome of Neuronal Circuits

Paul J. Bonthuis and Christopher Gregg

Abstract Genomics is fostering broad discoveries across biological disciplines, including the neurosciences. However, the analysis of gene expression and gene regulation in the brain is complicated by the extraordinary cellular heterogeneity, complex connectivity, and dynamic physiology of the tissue. Indeed, one of the great challenges of modern neuroscience involves the functional and molecular classification of cells in the brain within the context of network connectivity. In parallel, a major area of focus in the field of genomics involves the development of technologies that can profile the transcriptome of single or small numbers of cells [38]. Thus, major objectives in these two fields are well aligned. Here, we review modern approaches for the analysis of gene expression at the cellular level in the brain. As detailed below, these new technologies involve both *ex vivo* genomics approaches and new and emerging technologies for *in situ* and *in vivo* imaging of molecules in the brain.

2.1 Introduction

Brain functions and behaviors emerge through the coordinated responses and activity of different neurons organized into networks. Neural networks are composed of neurons with unique molecular features such as the expression of specific neurotransmitters, neuropeptides, ion channels, receptors, and transcription factors. Neurons are typically classified based on the expression of one or a few molecular markers. However, these broad classifications fail to capture the complexity of network connectivity and functionality. For example, a heavily studied neuron population in the arcuate nucleus that plays a role in feeding behavior has been defined by its expression of agouti-related peptide (AgRP). Yet, optogenetic studies have revealed that AgRP neurons are functionally heterogeneous and only a specific subpopulation controls the drive to feed. The subset of AgRP neurons that drive hunger have been shown to interact with cells in the paraventricular nucleus that express

P. J. Bonthuis (✉) · C. Gregg
Department of Neurobiology and Anatomy, University of Utah School of Medicine,
Salt Lake City, UT, USA
e-mail: Paul.Bonthuis@neuro.utah.edu

oxytocin, thyrotropin-releasing hormone, and pituitary adenylate cyclase-activating polypeptide [3, 35]. Currently, the gene expression programs that define the unique connectivity patterns and functions of AgRP neuron subpopulations are unknown. Subpopulations of neurons can be defined based on connectivity patterns for every major, molecularly defined class of neuron in the brain. Recent collaborative efforts have begun to define wiring diagrams (connectomes) in mouse and human brains at the macro-, meso- and microscales (<http://www.humanconnectome.org>; <http://www.mouseconnectome.org>; <https://www.alleninstitute.org>). The studies reveal extraordinary complexity and cellular diversity in terms of connection patterns [54, 93]. Further, efforts to classify neurons based on morphological and physiological criteria are underway and the number of defined different cell types is constantly growing. In addition to the 85 billion neurons that are estimated to exist in the human brain, there are even more glial cells, which also perform essential supporting functions. Underlying the formation and function of all these cells is the transcriptome.

The transcriptome was once thought to be largely composed of ribosomal RNA, transfer RNA, and a small number of protein-coding messenger RNAs (~ 2 % of the total). However, it is now clear that transcription is pervasive in the genome and that ~ 75 % of the genome is transcribed [15]. In most cases, the function of these transcripts is unknown. GENCODE estimates that the human genome (version 19) contains 57,820 genes that give rise to 196,520 different transcripts (<http://www.genCODEgenes.org/stats.html>). In total, 20,345 protein-coding genes have been identified, each of which gives rise to ~ 3 different transcripts due to the effects of alternative promoters, polyadenylation sites, and exon splicing. Different transcript isoforms from a given gene can have very different functions, and their highly regulated expression can change in response to different stimuli in a developmental stage and cell-type-specific manner. The noncoding portion of the human transcriptome includes 13,870 long noncoding RNAs and 9013 small noncoding RNAs. Noncoding RNA species range from small noncoding RNAs, such as microRNAs (1973 in total), endogenous small-interfering RNAs (siRNAs), small nucleolar RNAs (snoRNAs, 1530 in total), small nuclear RNA (snRNA, 1383 in total), and piwi-interacting RNAs (piwiRNAs), to long noncoding RNAs that are between 200 bp and several kilobases in length. The functions of most features of the transcriptome are poorly defined.

It has been estimated that 86 % of all human genes are expressed in the brain, and most of these are differentially localized to different brain regions and/or different developmental stages [33, 50]. In addition, 90 % of all genes expressed in the brain exhibit differential exon usage between brain regions and/or developmental stages [33, 50]. Thus, the brain transcriptome is dynamic and extremely complex. Some features of gene expression appear to be unique to the brain. For example, 3' untranslated regions (3'UTRs) are generally longer for transcripts in the brain compared to peripheral tissues [88]. These sequences can be over 10 kb in length and are known to contain binding sites for miRNAs and regulatory proteins, which could make transcript regulation through the 3'UTR especially complex in the brain. In addition to differences between brain regions and developmental stages, expression levels and isoform usage can change in response to metabolic changes, hormonal

changes, neuronal activity, circadian rhythms, and other events. Some isoforms contain signal sequences that result in their transport specifically into dendrites, while others are localized to axons or specific organelles, or are retained in the nucleus [7, 30]. Thus, the transcriptome of the brain is not a disorganized cloud of RNA molecules but rather a highly regulated system of transcripts that arise under specific conditions and are trafficked to precise locations for local translation or other functions. Understanding these processes is central to the greater goal of understanding the mechanisms that underlie specific brain functions.

For the most part, our understanding of the brain transcriptome arises from studies of discrete regions: microdissected chunks of brain, which is composed of hundreds of different cell types. Since brain functions arise through the activities of thousands of different cell types organized into different circuits, this level of analysis generally fails to capture the complex relationship between the transcriptome and circuit physiology. Encouragingly, emerging approaches are beginning to allow accurate measurements of gene expression at the cellular level in the brain. Defining the gene expression programs that establish the morphological, anatomical, and physiological traits of specific cell types is becoming a realistic undertaking. Further, we are gaining access to cell type-specific transcriptional programs that respond to changes in neuronal activity, disease, and other events.

In this chapter, we describe established and emerging techniques for measuring the transcriptome in specific neuronal populations. We cover the following topics: (1) how different cell types are defined and identified; (2) techniques and technologies to label and isolate RNA selectively from a desired cell type; (3) methods to detect and quantify RNA transcripts from a small number of cells of a specific type or from single cells; and finally, (4) quantitative *in vivo* and *in situ* analyses to measure both the location and expression levels of genes in individual cells. Although newer techniques are constantly emerging and not all of them can be comprehensively described here, it is our aim to discuss the benefits and limitations of some of the most widely used and potentially impactful approaches available.

2.2 Cell Type Identification and Labeling

Brain cells are categorized by anatomical location and characteristic features. The broadest intrinsic characteristic is whether a cell is neuronal or glial. Central nervous system glia consist of astrocytes that perform homeostatic functions, myelin-producing oligodendrocytes, ventricular ependymal cells that secrete cerebrospinal fluid and form the blood–brain barrier, and phagocytic microglia differentiated from hematopoietic stem cells. Neurons are polarized cells that send and receive electro-chemical information through synaptic connections with other neurons. A typical neuron receives numerous excitatory and inhibitory inputs onto its dendritic processes, and if excited above a threshold electrical potential, it fires depolarizing action potentials that trigger the release of small molecule and/or peptide neurotransmitters from their axon terminals. Anatomical location, morphology (e.g., soma size and shape, dendritic arborization, axon length), electrophysiological

properties, connectivity, neurotransmitter content, and molecular marker expression are common ways to differentiate neuronal cell types [52]. For example, two major classes of neurons are pyramidal shaped, long-distance projecting, glutamatergic excitatory neurons, and γ -aminobutyric acid (GABAergic), inhibitory interneurons; however, classifying subtypes of these two classes is difficult [76]. One way to identify neural subtypes is through electrophysiological characteristics. These can include properties of intrinsic firing, responses to neurotransmitters, and electrical conductance through the composition of ion channels [19].

Transgenic technologies for labeling specific cell types based on molecular expression patterns are particularly useful in organisms where these tools are well advanced, such as mice. Typical transgenic labeling tools include CRE recombinase and fluorescent reporter proteins acquired from bioluminescent organisms, such as green fluorescent protein (GFP) derived from jellyfish. Old methods for molecular reporter labeling genetically engineered GFP fused to the amino- or carboxy-terminal end of a gene of interest, and transgenic knock-in insertion by homologous recombination replaced the endogenous gene with the protein-fusion reporter. However, recombinant fusion proteins can be problematic if the conjugated reporter perturbs the function of the endogenous protein. Fluorescent proteins can also be cloned into bacterial artificial chromosomes (BACs) such that their expression is controlled by several kilobases of the transcriptional regulatory element from an endogenous gene of interest. Ideally, when the BAC construct is inserted into the genome following pronuclear injection, the reporter is only expressed in the cell types that express the endogenous gene [25, 44, 85]. In practice, however, BAC transgenics often have off-target labeling due to positional effects related to the site of genomic integration. Further, one must often screen several founder lines in order to identify a line in which the correct cells are labeled. In model organisms where homologous recombination can be utilized for gene targeting, many labs have turned to modified, knock-in strategies. For example, a CRE recombinase allele or fluorescent reporter of interest can be engineered with an upstream internal ribosome entry sequence (IRES) and targeted to the 3'UTR region of the endogenous gene of interest [71, 82]. In this approach, a bicistronic transcript is generated under the control of the endogenous enhancer and promoter elements. Thus, the expression of the CRE allele or reporter matches the endogenous transcript.

CRE lines can be crossed to floxed reporter lines, where expression of the reporter is conditional upon a CRE recombination event in order to label specific cell types. A challenge for conditional reporter lines is often that the intensity of the fluorescent label is weak, limited to a subset of cells and/or variegated. To address these issues in mice, Liqun Lou's laboratory developed an approach to increase expression by knocking the reporter gene into the permissive *ROSA26* locus in the mouse genome under the control of a strong and ubiquitous promoter, such as CAG [51, 94]. CAG is a synthetic sequence composed of a cytomegalovirus enhancer, the chicken beta-actin promoter, and the splice acceptor site of the rabbit beta-globin gene [53]. Subsequently, the Allen Institute for Brain Science used this approach to generate robust conditional reporter lines with three spectrally distinct fluorescent proteins: EYFP, ZsGreen, and tdTomato [46]. The reporters were engineered with a woodchuck hepatitis virus post-transcriptional regulatory element that is added to

the end of the transcript to increase mRNA stability. Numerous CRE driver lines are now available from the Allen Brain Institute and other public repositories (GENSAT or Jackson Laboratory) for use with these reporter lines, and the expression patterns of different CRE drivers are being characterized and made publically available at <http://transgenic-mouse.alleninstitute.org>. These are outstanding resources to begin studying the transcriptome of specific cell types in the brain.

Ideally, one would be able to purify subtypes of cells from the brain not just on the basis of the expression of a single marker gene, but also based on connectivity patterns and physiological properties. Currently, we have limited approaches to integrate all of these features. In principle, fluorescent retrograde tracers can be used in combination with reporter mice to label specific subpopulations of molecularly defined neurons for purification. This approach would allow for the isolation of neurons based on both connectivity patterns and molecular markers. A pioneering study by Arlotta and colleagues previously employed retrograde fluorescent tracers to define and purify motor neurons in the developing cortex for transcriptome analysis [2]. Virus-based approaches to label specific circuits have also been developed [8, 43, and 57]. However, these labeling methods are not ideal for transcriptome analysis, since they influence the health of the infected cells and can change gene expression. Sugino et al. [76] were able to characterize 12 distinct populations of both GABA and glutamate neurons, as distinguished by a combination of factors including anatomical location, transgene expression, and by connectivity using a retrograde tracer.

Approaches to isolate cells based on their activity patterns are very limited. Immediate early gene expression can be used as a proxy for neuronal activity. Recent studies exploited this by expressing a destabilized fluorescent reporter under the control of enhancers from the immediate early genes *c-fos* and *arc* [10]. These reporters have been used to label cells that respond to fear conditioning [63], oriented visual stimuli [89], and motor learning [65]. Cells labeled in this way could be purified for transcriptome analysis within a functionally related group of cells. In principle, this approach could be used in combination with other reporter lines and/or neuron tracing strategies to further enhance specificity. An alternative approach, involving activity-dependent ribosomal protein phosphorylation, is detailed below.

2.3 Cell Type and RNA Isolation Strategies

The strategy chosen to capture and isolate RNA from specific cell types is intrinsically linked to the methods used to identify or label the cell types of interest. Laser-capture microdissection (LCM) isolates brain regions or specific cell types under a microscope from thin cryosections of frozen or fixed tissue based on anatomical location, morphology, and molecular marker expression [26, 45, 52, 72]. LCM can attain two types of samples: whole tissue from a well-outlined and defined brain region, or single cells of a specifically labeled type. These methods offer improved accuracy, precision, and selectivity compared to manual dissection of fresh whole tissue. One issue with standard LCM is that the wide cuts that are generated ($\sim 7.5 \mu\text{m}$) cannot dissect fine cell contours, such as neuronal processes, and

material from these compartments are lost. However, laser-directed microdissection (LDM) systems make narrow cuts ($\sim 0.5 \mu\text{M}$) that can trace the shape of the cell [56, 66]. Both LCM and LDM require cryosectioning of either fresh-frozen or fixed tissue; therefore, the RNA might be more degraded as compared to RNA extracted directly from live tissue.

For higher quality RNA and less contamination from surrounding tissue, live cells expressing a molecular reporter can be purified for RNA extraction by enzymatically dissociating tissue into single cells and then picking the labeled cells under a fluorescent microscope. In this approach, live tissue vibratome sections are prepared ($\sim 400 \mu\text{M}$) and transferred into oxygenated artificial cerebral spinal fluid (ACSF), as would be done for electrophysiological recordings [29]. The slice preparation, or a microdissected portion thereof, is then dissociated by protease digestion and gentle mechanical trituration while keeping the cells alive and intact. Unfortunately, these procedures tear off neural processes, leaving the contents of axonal and dendritic compartments behind after cell sorting. Once the tissue is dissociated, individual cells are sorted by their expression of molecular (i.e., fluorescent) markers. Manually sorting live cells is a labor-intensive, yet highly accurate method for purifying individual cell types. Fluorescently labeled neurons are manually sorted in a culture dish by scanning for labeled cells under a dissecting microscope, and healthy cells are picked using a pulled glass mouth pipette and deposited into a lysis buffer for RNA extraction [29]. Manual sorting is convenient and useful when a highly pure sample of ~ 30 – 100 cells is sufficient [52, 56, 76].

A high-throughput purification approach to isolate dissociated cells involves fluorescent-activated cell sorting (FACS). In FACS, live cells are streamed single file through a narrow nozzle, as a detector measures their fluorescence. As the single cells exit the bottom of the nozzle, different electrostatic charges are applied to them before they pass through an electric field that deflects and sorts them into separate receptacle tubes based upon fluorescence [84]. This approach has been used in numerous gene expression-profiling studies [11, 17, 44].

A final method to sort live, dissociated cells, which does not need a transgenic organism or other means of fluorescent labeling, is called immunopanning (PAN) [6]. This technique uses antibody-covered culture plates to separate different cell types based on their expression of cell-surface proteins. Cells are placed into the immunolabeled plates over a period of time to allow antibody binding, and unbound cells that do not present the conjugate surface antigen are washed away from the adherent cells. By using a series of plates with antibodies against unique antigens, cells are separated according to specific protein expression profiles. Unfortunately, empirical evidence indicates that PAN induces immediate early, stress, and apoptosis genes, likely because the process takes a relatively long period of time and exposes the cell surfaces to antibodies [55]. Once cells of a desired type are acquired by any of the methods above, RNA is extracted from the purified live-cell population.

While the methods described above are common, they have major limitations. The axonal and dendritic processes are stripped away, yet these cellular compartments contain a large portion of the transcriptome due to local translation at synapses, growth cones, and other sites [48]. In addition, the cells are extensively manipulated and separated from their *in vivo* environment, which presumably leads to

dramatic changes to physiological gene expression. Generally, these purification approaches are appropriate for discovering cell type-specific marker genes that distinguish one cell type from another, but are not ideal for studying endogenous gene expression programs. The limitations of these methods have led to the development of other strategies for RNA purification from defined cell populations.

RNA-tagging methods avoid the need for tissue dissociation and cell screening steps. In these approaches, transgenic organisms express epitope-tagged RNA-binding proteins in a cell type-specific manner. Antibodies against the epitope tags are used to immunoprecipitate mRNA-protein complexes from whole-tissue homogenates, and then the RNA is extracted from the immunoprecipitate. Approaches of this type include poly-A-binding protein (PABP) purification, translating ribosome affinity purification (TRAP), and RiboTag [55]. As the name implies, PABP binds to the poly-A tail of mRNA transcripts and stabilizes eukaryotic initiation factor (EIF) subunit binding to the 5'-cap of mRNA. Protein-protein interactions between PABP and EIF complexes promotes mRNA circularization, enhances mRNA stability, and increases protein translation presumably due in part to increased ribosome reinitiation [21]. Since PABP naturally binds to poly-A tails (i.e., mRNA), cell type-specific mRNA isolation is made possible with a recombinant FLAG-tagged PABP expressed under the control of a cell-specific promoter [36, 47, 67, 92]. RNA is fixed to RNA-binding proteins *in situ* with formaldehyde, and total cell homogenates are then immunoprecipitated with FLAG-antibody conjugated beads. After washing away the rest of the cellular content, including RNA not captured by the FLAG beads, poly-A mRNA fixed to FLAG-tagged PABP is reverse-crosslinked and eluted. Thus, PABP technology captures all polyadenylated RNA in a cell type-specific manner, including many untranslated RNAs.

Translated RNA can be purified from untranslated RNA by exploiting the fact that actively translated mRNA is loaded with ribosomes. In TRAP technology, EGFP is fused to the N-terminus of the large ribosomal-subunit L10a and is expressed in transgenic organisms under the control of a cell type-specific promoter [16, 28]. The EGFP-L10a fusion protein integrates into polysome complexes, and immunoprecipitation with an EGFP antibody enriches for actively translated RNA from the targeted cells. In contrast to PABP isolation, TRAP does not involve fixation of RNA-protein complexes prior to immunoprecipitation, but it does require rapid dissection and homogenization in lysis buffer supplemented with magnesium and cycloheximide. The lysis buffer performs several functions, including keeping ribosomes bound to polysomal RNA and solubilizing rough endoplasmic reticulum. High-affinity EGFP antibodies that can withstand high salt washes are conjugated to magnetic beads and used for the immuno-isolation [28]. The RiboTag method is similar to TRAP in that it purifies cell type-specific polyribosomal mRNA. The strategy uses Cre-lox technology to conditionally knock-in a c-terminally hemagglutinin (HA)-tagged version of ribosomal protein subunit *Rpl22* exon4, which is transgenically inserted just downstream of the endogenous exon 4 [68]. Cell type-specific expression of CRE removes the endogenous exon 4 by recombination and puts the HA-tagged exon 4 in frame to express HA-tagged RPL22 protein (RPL22^{HA}). This versatile strategy can be combined with many existing cell type-specific CRE driver lines. Finally, phosphorylated ribosome capture is a variation of TRAP that is

specifically geared toward neuroscience applications, and selectively purifies RNA from neurons based upon changes in firing activity rather than on *a priori* selected molecular markers [34]. This approach depends upon the fact that the S6 subunit of the ribosome is phosphorylated by the PI3-K/mTOR, MAPK, and PKA signaling pathways in activated neurons [34]. An antibody specific for the phosphorylated epitope of S6 subunit is used to pull down RNA undergoing translation within the activated cells. This technique provides a powerful new method to study activity-induced gene expression profiles in cells that fire under specific conditions. Overall, ribosomal-tagging methods are powerful for studying actively translated, coding RNAs, but they are not applicable for noncoding RNAs.

A cell type-specific RNA-tagging approach that does not depend upon mRNA-binding proteins, and can capture all RNA species including noncoding RNA, is the thioracil RNA-tagging (TU-tagging) method. In TU tagging, CRE-lox technology is used to conditionally express a heterologous thioracil phosphoribosyltransferase (UPRT) enzyme derived from *Taxoplasma gondii* [49] in specific cell types of interest. Next, 4-thioracil (TU), which is incorporated into actively transcribed RNA, is injected into the UPRT transgenic organism. Only cells that express the UPRT transgene will incorporate TU into the transcriptome [24]. RNA is then purified from whole-tissue dissections, and TU-incorporated RNAs within the pool of total RNA are chemically conjugated to biotin via the thiol group of TU. Streptavidin beads are then used to purify biotin-conjugated, TU-tagged RNA for downstream profiling by RNASeq. This method provides temporal information, since the TU injection allows for pulse labeling of newly transcribed RNAs. Overall, the immunoprecipitation and TU-tagging methods are powerful, but suffer from high background. The ongoing efforts are focused on improving purification chemistries to address this problem.

The choice of which method to use must be guided by the central goals of the experiment. For example, if the goal is to identify cell type-specific marker genes, then cell dissociation and purification-based strategies should work well. If the goal is to study endogenous gene expression programs in disease models or under different experimental conditions, then other purification strategies might more effectively represent the physiological state. Empirical evidence shows that only 30–60 cells are needed to get consistently reproducible results in the number of RNA transcripts detected from cell type-specific pools [56]. Unfortunately, none of these methods are universally optimal and new strategies with improved efficiencies are greatly needed.

2.4 Single Cell Transcriptomics

Analyzing the transcriptome from several cells provides an averaged readout of gene expression at the cellular level. The genomics and neuroscience communities have sought to accurately profile gene expression at the single cell level for over two decades. The pioneering study that first achieved this feat was performed by Catherine Dulac and led to the discovery of chemoreceptors in the vomeronasal organ of the mouse in 1994 [18]. However, subsequent attempts to profile the entire transcriptome of a single cell achieved limited success. Encouragingly, advances in

single cell genomics are now occurring at a rapid rate [72]. An ideal application for this technology would involve the integration of electrophysiological methods, such as patch clamp, with gene expression profiling on the same neuron to learn how physiological properties relate to gene expression [61].

To perform single cell transcriptome analysis, new technologies are being developed to isolate and process many single cells using microfluidic chambers. For example, the C1 Single-Cell Auto Prep System (Fluidigm, South San Francisco, CA) isolates 96 cells into individual chambers within a microfluidic chip and automatically performs lysis, cDNA synthesis, and amplification [80]. The single cell cDNA samples can then be used for qPCR analysis of targeted genes or made into libraries for next-generation sequencing. Since loading the chip only requires pipetting a single sample, hands-on time and chances for technical errors and contamination artifacts are greatly reduced. A direct comparison of single cell RNAseq libraries constructed at the nanoliter scale in C1 microfluidic devices, to libraries constructed at the microliter scale in tubes, found the C1-generated libraries performed better in regard to sensitivity, accuracy, and false positives [91]. Moreover, combining the data from all 96 single cells processed in the C1 device quantitatively recapitulates measurements from a bulk RNA sample RNAseq experiment, giving high confidence that the single cell measurements were accurate. These results demonstrated the ability to differentiate discrete cell identities and/or physiologies within tissues by individually assessing the transcriptomes of single cells. The nanoliter volumes used in microfluidic devices not only improve single cell sequencing quality, but also reduce the cost of consumable reagents.

2.5 Amplification of Low Input RNA

After separating a pool of cell type-specific live cells, RNA-tagged molecules, or individual single cells from the surrounding tissue, the resulting RNA isolation yields are usually low and need to be amplified before transcriptome measurement. New methods for amplifying different amounts of starting material into cDNA libraries usable for transcriptome measurement are being created at a rapid rate. Each method has inherent strengths, weaknesses, limitations, and biases [1, 72], and so, investigators must chose which amplification method will best be suited to their particular question.

In vitro transcription (IVT) has been in use for over three decades to linearly amplify an RNA sample [86], and IVT is often used to amplify and convert an RNA sample into labeled cRNA for microarray analysis. IVT starts by reverse-transcribing an RNA sample. The oligo dT primer used in this reaction contains a 5' leader overhang encoding a T7-transcriptional promoter sequence (Fig. 2.1). The oligo dT sequence nonselectively anneals to polyA-tailed mRNA, while the 5' overhang introduces a T7 RNA-polymerase loading site into each cDNA. Subsequent IVT with a T7 RNA polymerase is used to make copy RNA (cRNA). A second round of reverse transcription and IVT from the first round cRNA may be performed to amplify the library and label the cRNA for microarray hybridization. This approach

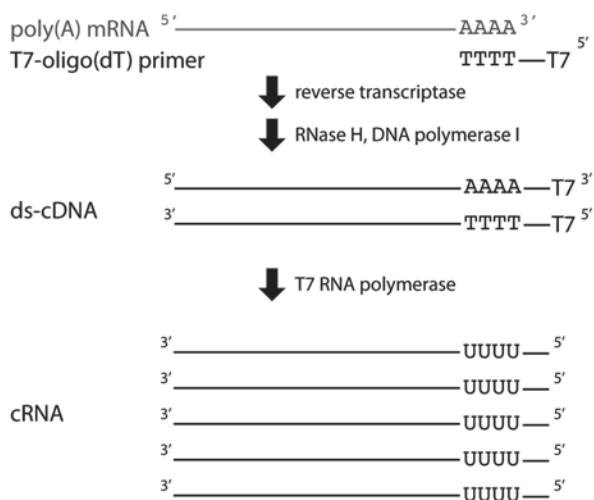


Fig. 2.1 In vitro transcription (IVT). Polyadenylated total mRNA (*polyA mRNA*) is selectively reverse-transcribed to copy-DNA (*cDNA*) with an oligo(dT) primer carrying a T7 RNA polymerase binding site on the 5'-end (**T7-oligo(dT) primer**). After reverse transcription, the cDNA:mRNA duplexes are converted to double-stranded cDNA (**ds-cDNA**) by the addition of **RNase H**, to cleave the mRNA into short sequences, and **DNA polymerase I** uses the cleaved mRNA to prime synthesis of second-strand cDNA. The ds-cDNA, with a T7 site, is then used as template for in vitro transcription with **T7 RNA polymerase** to synthesize several molecules of copy-RNA (**cRNA**). A second round of IVT, using the cRNA as starting template, may be used to further amplify and label the library for transcriptome measurement (e.g., microarray analysis). RNA sequence is depicted in *grey letters and lines*, and DNA in *black letters and lines*. (Figure was adapted from [86])

has been successfully used to amplify a range of RNA starting concentrations acquired by several isolation methods. For example, RNA isolated from 5000 to 10,000 FACS purified EGFP positive cells (yielding 3–10ngs of total RNA; [44]), RNA isolated from ~ 30 to 50 cells acquired by manual sorting (yielding 0.25–1 ng of total RNA [76]), TRAP-purified RNA from pooled tissue of several (3–7) mice [16, 28], and even RNA isolated from single cells after electrophysiological recording [17]. Since IVT is a process of linear amplification, artifacts due to exponential amplification (e.g., PCR) are avoided [19].

Like IVT, RNA single primer isothermal amplification (Ribo-SPIA) by NuGEN is a linear amplification technology [12, 37]. NuGEN (San Carlos, CA) has several products to generate libraries for different applications. The first step in Ribo-SPIA is to reverse-transcribe mRNA into cDNA using a 5'-RNA-DNA-3' hybrid reverse transcription primer (RT primer). The 3'-DNA ends of the RT primers anneal to the mRNA template and prime first-strand cDNA synthesis with reverse transcriptase. The reverse transcription reaction can either use RT primers that all have a 3'-DNA poly-(T) sequence for 3'-biased mRNA amplification, or use a mixture RT primers with both poly-(T) and random nucleotide 3'-DNA sequence for whole of transcriptome amplification. The 5' ends of the RT primers are composed of a single

New Techniques in Systems Neuroscience

Douglass, A. (Ed.)

2015, X, 301 p. 72 illus., 64 illus. in color., Hardcover

ISBN: 978-3-319-12912-9