Additional information for January Journal Club

1- Overview of mechanism of transcription

Transcription is the first step in gene expression in which a gene's DNA sequence is copied (transcribed) to make an RNA molecule that will be used to translate the genetic information into proteins. This essential and fundamental cellular process has conserved throughout all three kingdoms of life. It is a highly regulated and its control is essential to maintain homeostasis.

Transcription begins when RNA polymerase binds to a **promoter** sequence near the beginning of a gene (directly or through helper proteins). RNA polymerase uses one of the DNA strands (the **template strand**) as a template to make a new, complementary RNA molecule. Transcription ends in a process called **termination**. Termination depends on sequences in the RNA, which signal that the transcript is finished.

2- Study of the transcriptome

The analysis of the transcriptome, which is the complete set of transcripts in a cell, provides insightful information on a cell state or its characteristics. Elucidating what genes are expressed and their relative time course and/or localization allows to determine their biological role in a specific cell or pathway or disease. It also provides clues on characteristics of subpopulation of cells and its potential relevance for disease progression for example.

Over the years, several methods have been developed to study the transcriptome. By in large the principle of these different methods is based on the fact that to detect the abundance of mRNA, the only element needed is the cDNA of the mRNAs of interest.

How to make cDNA?
The synthesis of DNA from an RNA template, via reverse transcription, results in complementary DNA (cDNA) (Figure 1). cDNA can then serve as template in a variety of downstream applications. By the action of reverse transcriptase and DNA polymerase, the formed cDNA can be used as a probe to hybridize with a specific gene sequence (microarray and *in situ* hybridization) or amplified by PCR and sequenced (RNA seq).

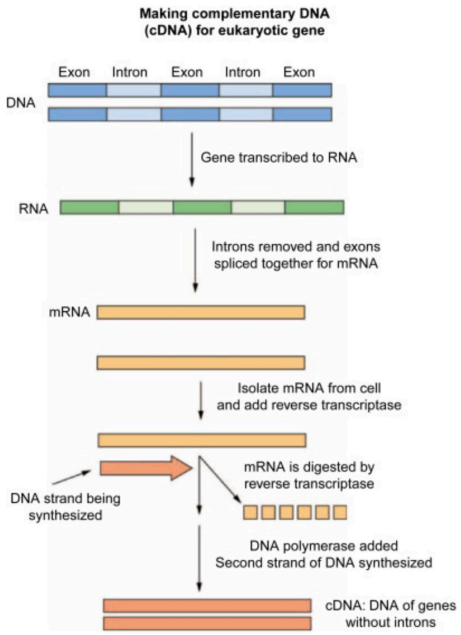


Figure 1: Generation of a cDNA from an isolated eukaryotic gene. cDNA, Complementary DNA. Figure published in Nucleic Acids and Molecular Genetics,
Gerald Litwack PhD, in Human Biochemistry (Second Edition), 2022

1.1 Microarray

A microarray is a collection of microscopic DNA spots attached to a solid surface (membrane or glass). Microarrays can be manufactured in different ways, depending on various criteria such as the number of probes under examination, costs, and goals of experiments. Commercially available arrays may have as few as 10 probes or as many as 5 million or more micrometre-scale probes.

Each DNA spot contains picomoles (10^{-12} moles) of a specific DNA sequence, known as probes. The target is defined as the mRNA of interest that will be reverse transcribed. Probe-target hybridization is usually detected and quantified by detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine relative abundance of nucleic acid sequences in the target (Figure 2).

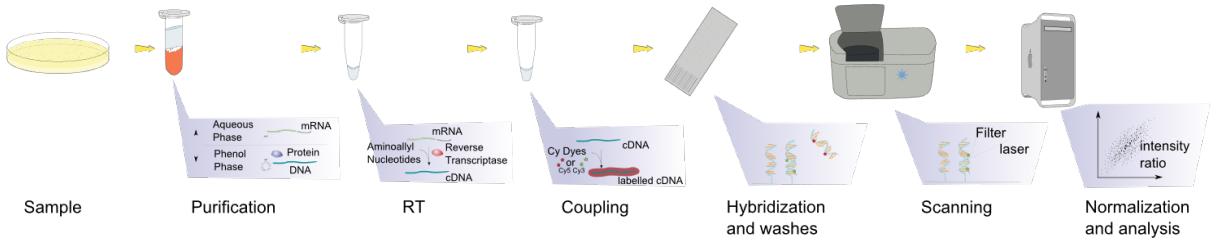


Figure 2: The steps required in a microarray experiment.
https://commons.wikimedia.org/w/index.php?curid=39423104

Following preparation of cDNAs and hybridization methods, the microarrays are scanned by a machine that uses a laser to excite the dye and measures the emission levels with a detector. The image is gridded with a template and the intensities of each feature (composed of several pixels) is quantified.

The raw data is normalized; the simplest normalization method is to subtract background intensity and scale so that the total intensities of the features of the two channels are equal, or to use the intensity of a reference gene to calculate the t-value for all of the intensities. The analysis required 2 samples to be compared.

The obtention of a signal is relatively straight forward however the analysis of the microarrays could be challenging as well as the validity of the results obtained.

Things to consider when using microarray technology:

- Replication in experimental design
- Statistical treatment of the data
- Mapping each probe to the mRNA transcript that it measures
- Volume of data
- Ability to share and compare data

Three main limitations to array technology: 1) a reference genome and transcriptome need to be available before the microarray itself can be designed; 2) can only detect expression of preselected probes; 2) Sensitivity as gene expression measurement is limited by background at the low end and signal saturation at the high end.

Most likely to study a transcriptome of interest related to your project you will not use the microarray technology but rather RNA sequencing or Assay for Transposase-Accessible Chromatin (ATAC) sequencing or *in situ* hybridization (ISH) assay. Below is a short summary of these different technologies.

1.2 RNA sequencing

RNA sequencing (RNA-Seq) technology enables rapid profiling and deep investigation of the whole transcriptome, for any species. This quantitative technology offers a number of advantages compared to microarray analysis, as it can determine RNA expression levels more accurately than microarrays. RNA-Seq is a sequencing technique which uses next-generation sequencing (NGS) to reveal the presence and quantity of RNA in a biological sample at a given moment. RNA-Seq can be used for various research objectives such as alternative gene spliced transcripts, post-transcriptional modifications, gene fusion, mutations/SNPs and changes in gene expression over time, or differences in gene expression in different groups or treatments. Furthermore, RNAseq doesn’t only use mRNA but could also provide tools to study total RNA, small RNA, such as miRNA, tRNA, and ribosomal profiling.

RNA seq can be used to study the transcriptome of multiple cells (Bulk RNA seq) or of single cells (single cells RNA seq or scRNA seq). The workflow is described below

(Figure 3)

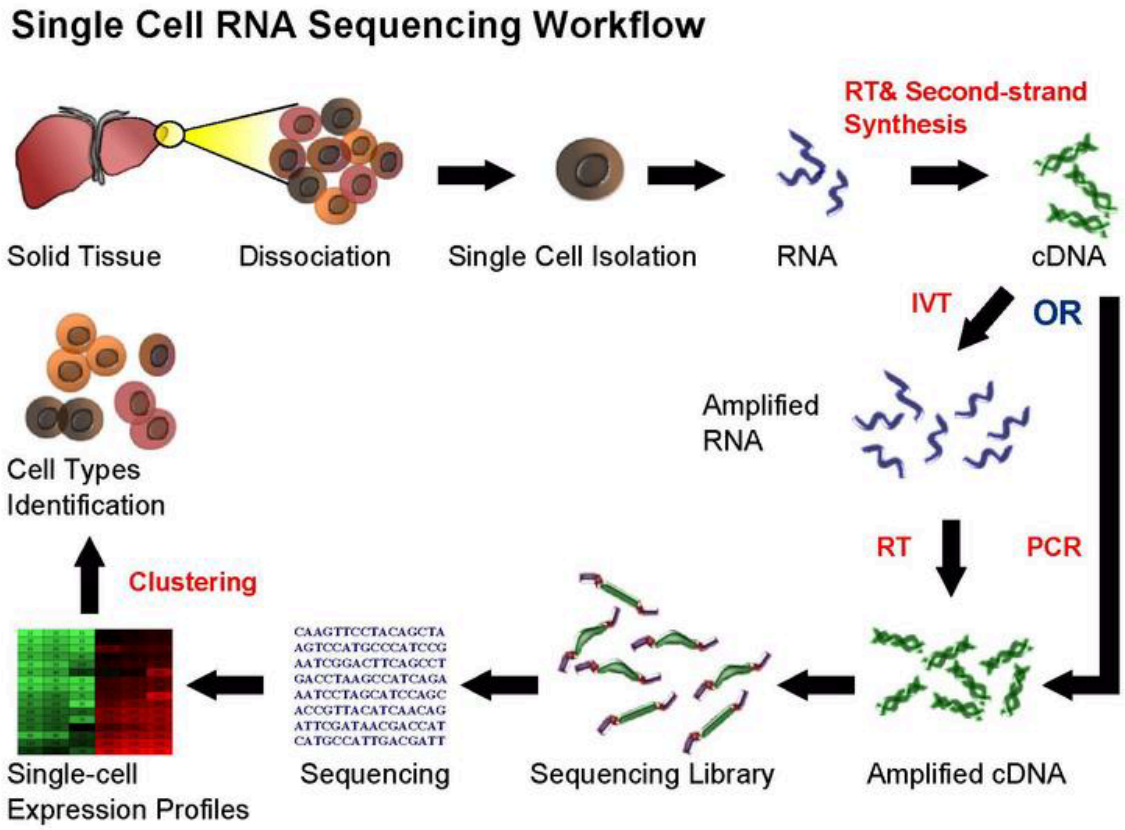
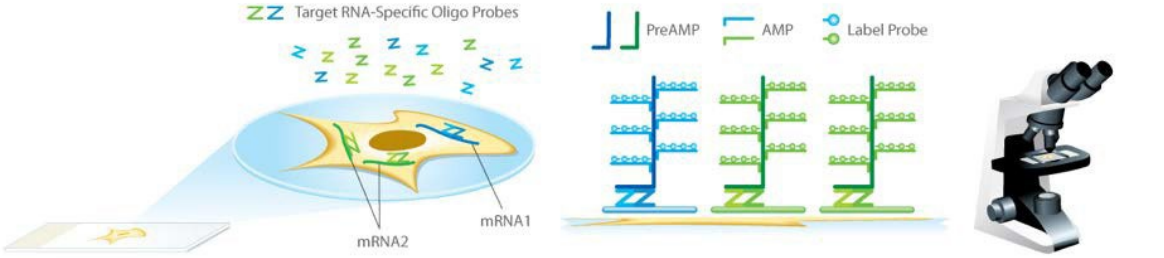


Figure 3: Single cell RNA sequencing workflow. https://en.wikipedia.org/wiki/File:RNA-Seq_workflow-5.pdf

Similarly, to microarray technology, sample collection up to sequencing is relatively straightforward but the analysis component of the technology (bioinformatics analysis) is challenging. However, scRNA-Seq is becoming widely used across biological disciplines as this technology can provide considerable insights. It's worth noting that RNA-Seq data provide a unique snapshot of the transcriptomic status of the disease and look at an unbiased population of transcripts that allows the identification of novel transcripts, fusion transcripts and non-coding RNAs that could be undetected with different technologies. Furthermore, it allows the discovery of clusters of cells with common and specific transcriptome allowing to develop new hypotheses.

1.3 *in situ* hybridization (ISH) assay

In situ hybridization (ISH) is a type of hybridization that uses a labeled complementary DNA, RNA or modified nucleic acids strand (i.e., probe) to localize a specific DNA or RNA sequence in a portion or section of tissue, a crucial step for understanding the organization, regulation, and function of genes. It's complementary to RNA-seq as it allows to spatially detect the expression of genes. For example, localized a cluster of cells that has been identified by RNA-seq. In addition, this technology allows the detection of mRNA for target with no or poor antibodies.



1: Tissue section	2: Hybridize to target RNA	3: Amplify signal	4: Image
Start with properly prepared tissue sections and pretreat to allow access to target RNA.	Hybridize multiple sets of gene-specific probe pairs to target mRNAs.	Use up to four signal amplification systems to detect multiple target RNAs. Probes are hybridized to a cascade of signal amplification molecules, culminating in binding of dye-labeled probes visible in different fluorescent chanels.	Visualize target RNA using a standard fluorescent microscope.

Figure 4: Procedure overview (from RNAscope Protocol Manuel)

3- Diauxic shift

The preferred source of carbon and energy for yeast cells is glucose. When yeast cells are grown in liquid cultures, they metabolize glucose predominantly by glycolysis, releasing ethanol in the medium. When glucose becomes limiting, the cells enter diauxic shift by switching metabolism from glycolysis to aerobic utilization of ethanol. This switch is associated with decreased growth rate. An extensive reprogramming of the transcription machinery is required. However, the events in transcriptional regulation during diauxic shift and quiescence were incompletely understood. Understanding when and where genes are expressed can provide clues on the state of the cells. Studying expression pattern of genes at different times during the cell culture can provide insight on regulatory mechanisms. The proposed article used the microarray technology to identify what genes are expressed at different time points. In addition, microarray technology allowed the investigators to uncover sets of genes with similar expression suggesting common transcriptional regulatory mechanisms. These data led to additional hypotheses that can be tested for example by specifically delete a gene, encoding a transcriptional factor, identified by the microarray analysis.