Northern blotting and cDNA microarray analyses used for gene expression studies

Introduction

Gene expression studies, have been rapidly growing since the first results obtained in this field in the early 1990s (Zhao et al., 1995; DeRisi et al., 1996; Velculescu et al., 1997; Baldi and Hatfield, 2002). Protein-coding genes transfer coded information for the amino acid sequences in peptides. The protein-coding information is transmitted indirectly, through mRNA molecules, which are intermediate molecules in the central dogma process. The study of transcriptomics, aims to discover the genotype-phenotype relationship by examining the expression level of RNAs (mRNAs, non-coding RNAs and small RNAs) in a given cell population at any given moment (Horgan and Kenny, 2011).

The gene regulation in prokaryotes and eukaryotes is fundamentally different (Struhl, 1999). Prokaryotes have simple, short-term regulatory pathways controlling gene expression by the promoter region and operons, while eukaryotes have complex, both short and long-term regulatory systems at enhancer and promoter regions and requiring spliceosomes to remove introns from the coding region from the much larger genome before translation (Figure 1).

Various methods exists for gene expression profiling (Table 1) (Bowtell, 1999). Sensitivity and specificity requirements will depend on the study organism, whether the gene is known or not, whether only a single gene is the target or multiple ones are.

Gene expression studies can be applied to study human diseases (Golub et al., 1999; Iorio et al., 2005), model organisms (Greenberg et al., 1995; Su et al., 2002; Leain et al., 2007), plant growth (Shinozaki et al., 2003), food production (Görlach et al., 1996; Rabbani et al., 2003; Loor et al., 2006; Wotton et al., 2014; Hu et al., 2017) and so on.

This essay briefly introduces RNA extraction, describes the procedure of Northern blot and microarray analysis methods which are widely used in gene expression studies. These methods are somewhat considered 'old school' as more quantitative, advanced analyses are available today (e.g. qRT PCR and RNA-seq) but the Northern blot and microarray methods have led to major breakthroughs in gene expression research.

Table 1. Gene expression can be studied with various methods such as Northern blotting, real-time, quantitative reverse-transcription (qRT) PCR, Expressed Sequence Tag (EST), Serial Analysis of Gene Expression (SAGE), DNA microarrays and whole transcriptome sequencing.

Method	Description	Reference
qRT PCR	cDNA PCR dying with SYRB green, gene by gene, small sample sizes, detects and quantifies mRNA	Schmittgen et al., 2000
EST	Short cDNA sequences cloned, can be used for physical mapping	Nagaraj et al., 2006
SAGE	Library of short tags and assessing how many each tag is detected and quantified	Yamamoto et al., 2001
RNA-seq	Shotgun sequencing of cDNA, amount of each RNA molecule can be determined	Ozsolak and Milos, 2010

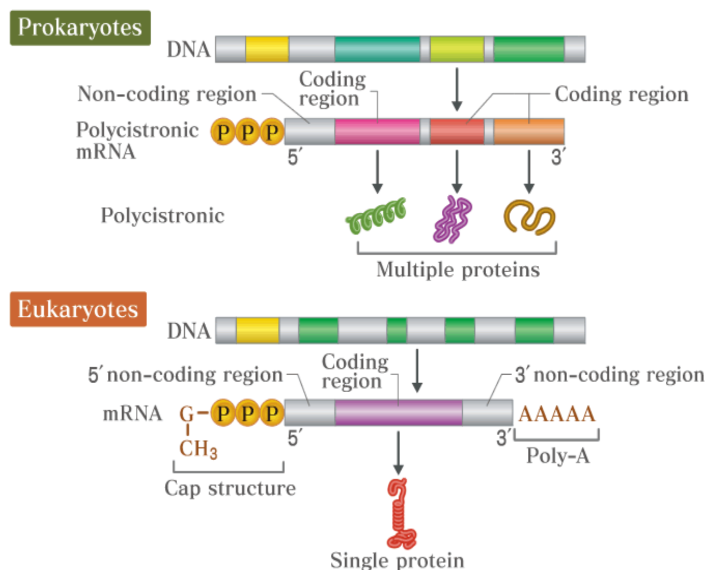

Figure 1. The DNA of prokaryotes is found in the cytoplasm, their chromosomes are circular and rarely contain non-coding regions. For eukaryotes, most of the DNA is located in their nucleus, wound on histone proteins, their chromosomes are linear and there are both introns (non-coding regions) and exons (coding regions), in their nuclear DNA. Mitochondria and chloroplasts also contain DNA. Picture by University of Tokyo.

RNA extraction

The first step for every analysis is the extraction of the total RNA content of cells which is usually carried out in a sterile environment to avoid contamination. The procedure is similar to DNA extraction, phenol extraction followed by ethanol precipitation. Commercially available QIAGEN RNeasy Mini Kit (QIAGEN, Germantown, MD, USA) is optimised for RNA extraction. The mRNAs which have 3' poly(A) tails, are isolated by either affinity chromatography or by oligo(dT) chromatography (Pemberton et al., 1975) or using Oligotex mRNA spin-column protocol (Qiagen, Valencia, CA, USA).

The first step is to size-fractionate the extracted RNA on a gel of 0.8-1.4% agarose concentration and 4-20% acrylamide. As mRNA requires stabilisation in denaturing condition, so either formaldehyde or glyoxal gel is used for mRNA (Lehrach et al., 1977). This prevents the RNA to form secondary structures which would mask strand breaks in further analysis. Up to 20 µg precisely pipetted RNA mixed with RNA loading buffer, stained usually with ethidium bromide, and is loaded onto the gel in buffer solution. Ethidium bromide cannot be used on glyoxal gels as they interact so alternative staining method is required. Denaturation can be carried out before loading the gel as well (Thurston et al., 1988).

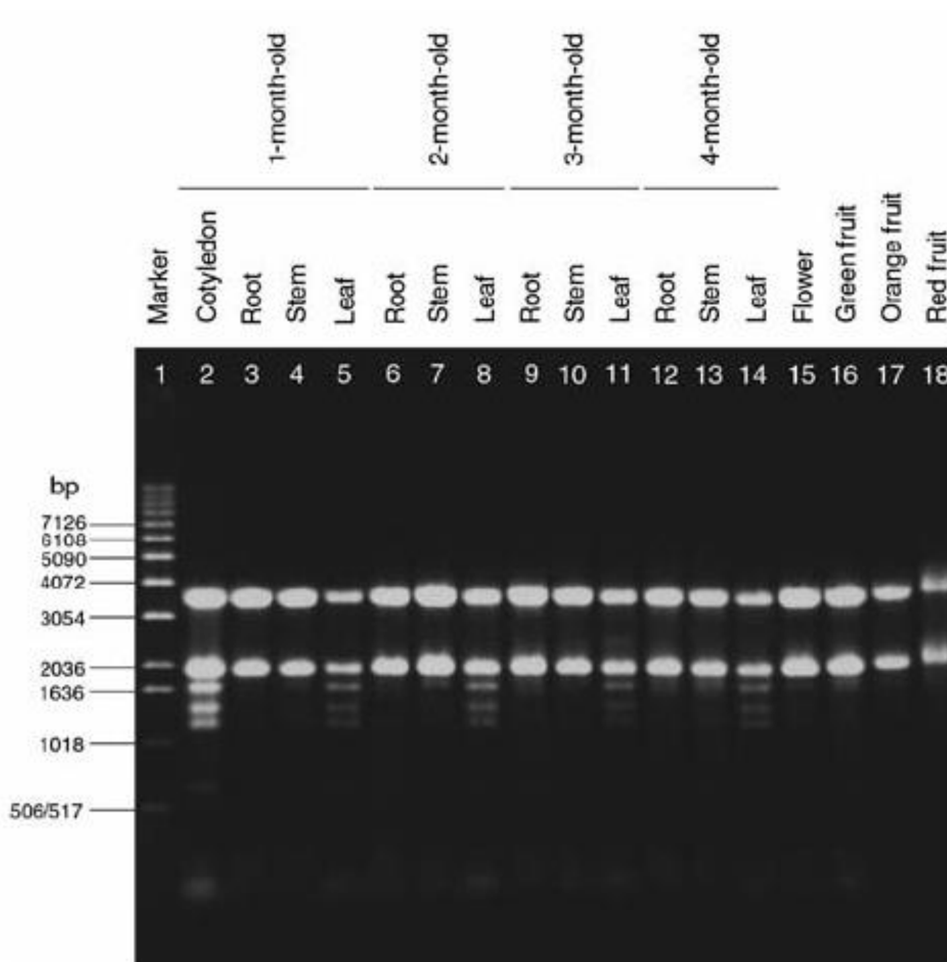

Figure 2. Study by Wang et al. (2009) examined tomato gene expression throughout development, in different plant organs. The RNA gel electrophoresis analysis of total RNA on 1.2% agarose gel shows a successful RNA extraction. In the electrical field, the mRNA molecules travel through the gel, separating them by their molecular weight, forming bands. Gels are inspected under UV light to ensure the quality of the mRNA.

Method 1: Northern Blotting analysis

This method was a development after Southern's (1997) Southern blotting procedure for DNA analysis. It provides information on the presence, size and abundance of RNA from a mixture. It is widely used for pair-wise comparisons to compare the levels of gene expression in two samples from different environments or different species. The method requires *a priori* knowledge of the gene sequences studied, which has always been a limiting factor for its use. The results can be reflected upon the organism's use of promoters, terminators or different splicing.

Following the mRNA gel electrophoresis, the gel is moved to blotting set up. During the most common capillary blotting set up the transfer buffer (saline-sodium citrate, SSC, containing NaCl) moves from reservoir below the gel into a dry stack of paper towels, above the membrane by capillary action (Figure 3). The buffer passes through the gel and carries the nucleic acid with it, until it hits the surface of the membrane above, then it stops. The filter membrane can vary according to the size of the RNA and the manufacturer's recommendation. One of the most successful membranes is positively charged nylon membrane and nitrocellulose membranes (Khandjian, 1987). The membrane has high binding capacity (4–500 μg/cm^2), with different pore sizes (0.22 to 0.45 μm). After removing the filter membrane with the mRNA bound to it, it is washed in diluted transfer buffer and dried. Then the RNA needs to be immobilised onto the membrane.

Using UV light, which activates pyrimidine bases (T and U), in alkaline conditions, the mRNA becomes covalently bound to the filter (Khandjian, 1986). Another way of immobilizing mRNA is to incubate or bake the filter at 65°C for 1–2 hours (Gillespie and Spiegelman, 1965). Therefore, the filter membrane both increases the affinity for nucleic acid and also immobilizes the mRNA on the membrane which is crucial to avoid any loss during the next step, the hybridization.

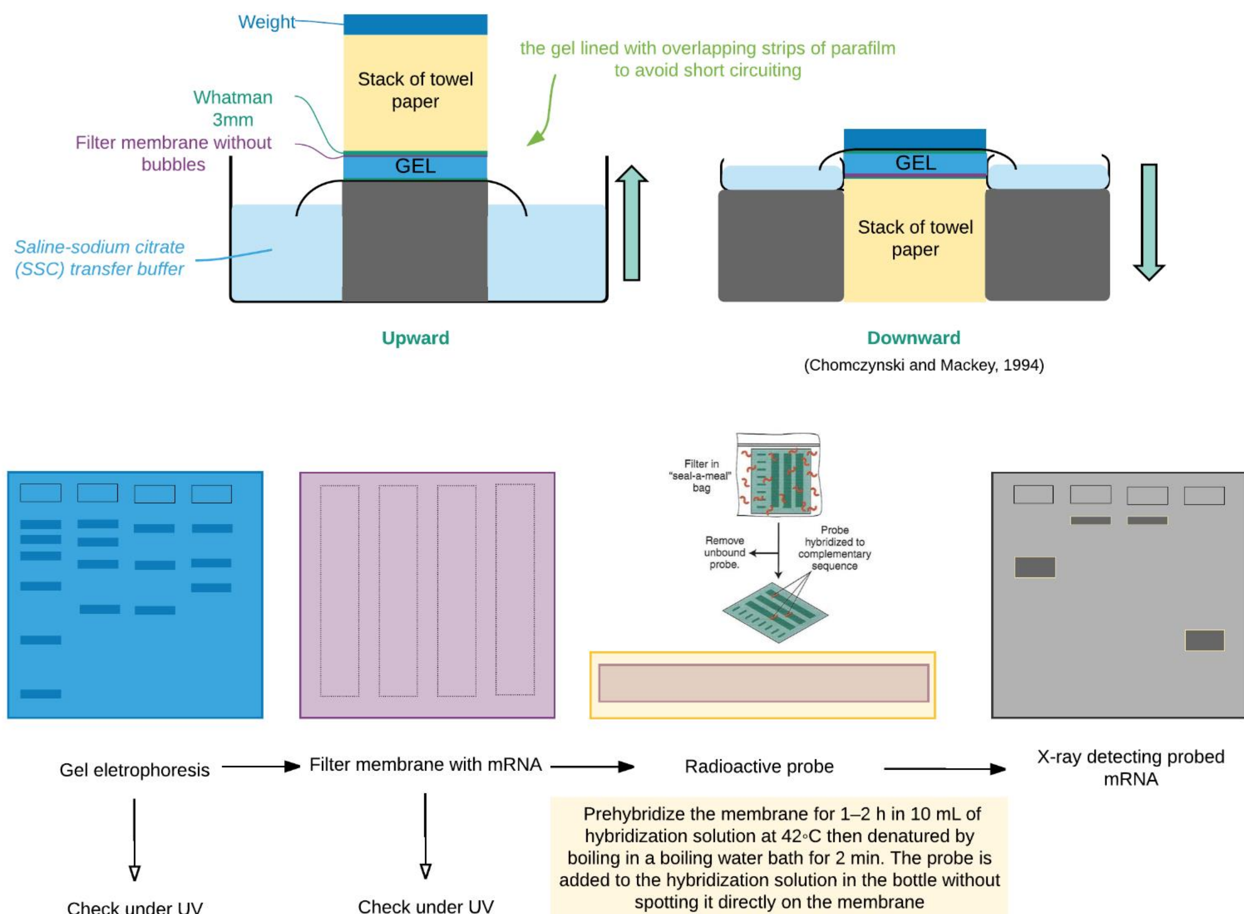

Figure 3. Northern blot upward and downward set ups showing which way the transfer buffer moves through and transfers the RNA onto the filter membrane. The RNA gel electrophoresis and filter membrane are checked under UV light. After hybridization, the band presence and width is the final output of the northern blot analysis. There are a number of commercially available blotting set ups available as well (e.g. electroblotter, semidry electroblotter, vacuum blotter, pressure blotter).

This starts with pre-hybridization, to protect the membrane from being coated by the probe, where the filter is in pre-hybridization buffer, an incubated at 42°C for 2-4 hours. Next step consists of probing the bound RNA. The most common probe is using radioactive labelling (^{32}P) which can be visualised on X-ray films. Oligonucleotides, small fragments (5-50bp) of ssDNA, can be used which are less hazardous but also less sensitive, and using digoxogenin-labelled oligonucleotide probes have been successful in various studies (Huang et al., 2014; Kim et al., 2010).

The hybridization takes place in a hybridization oven, the membrane usually on the side of a tube or in a sealed bag. It is incubated overnight at 42-50°C with constant movement (Kirakosyan et al., 2016). The labelled probe strand is hybridized in solution to the sample RNA (Figure 4). Once the solution is removed, it is washed with solution buffer (SSC and SDS). Incubation at 52°C for 30 minutes follows and the membrane is washed again. This procedure is repeated three times. This reduces the nonspecific interactions after the probes annealed to their targets. The membrane should be placed within cellophane and should be in contact with the X-ray film overnight at -70°C. Then the final output of the northern blot analysis is visualised on X-ray films.

The output of the Northern blot analysis is the image of the X-ray film, which then can be quantified with image analysis, measuring the band width the intensity. ImageQuant is a widely used software for Northern blot analysis (Zong et al., 2002; Fleming et al., 1999). It is important to use housekeeping genes, found in all cells carrying out 'housekeeping' functions, as controls and produce numerous replicates when using Northern blotting.

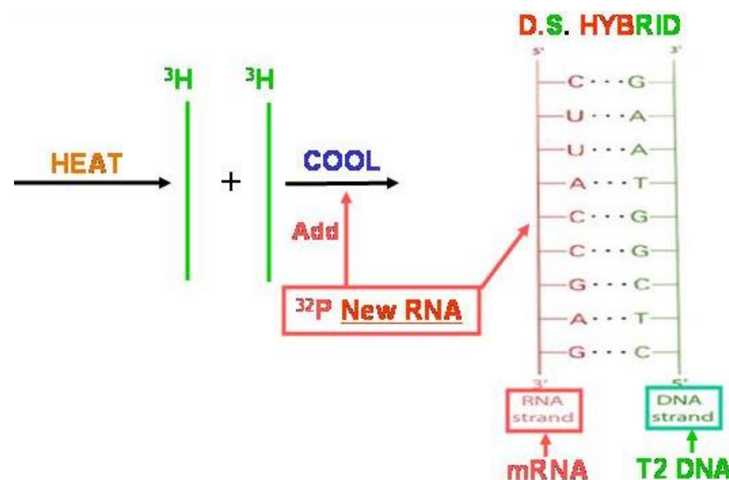

Figure 4. Hybridization of mRNA is the same process as DNA hybridization where two, complimentary ssDNA have the tendency to bind together to form a single dsDNA molecule following the Watson-Crick rule (A-T, C-G). In mRNA, T is replaced by U but binds to the complimentary probe. Illustration by cbc.arizona.edu.

Method 2: Microarrays

RNA microarrays monitor the transcription patterns of numerous genes during one procedure, contrary to the Northern blot's gene by gene approach. The DNA microarray was first introduced by Brown et al. (1999) using the complementary DNA (cDNA) of *Arabidopsis thaliana*.

As mRNA is more unstable than DNA because of the nucleophilic 2´ hydroxyl groups (Zhao and Hamilton, 2007), so after RNA extraction, reverse transcription is used to synthesize cDNA of the mRNA.This can be achieved through reverse transcription with reverse transcriptase enzyme which creates complimentary base of the mRNA. The mRNA is incubated for 1-2 hours with reverse transcriptase. Then the original RNA template can be degraded with RNase H, leaving only the single stranded (ss) cDNA. The ss cDNA is then converted into a double stranded (ds) cDNA with Polymerase

I and DNA Ligase. The complete set of cDNA copies of mRNAs is called a cDNA library and it should be large enough to contain all sequences of interested at both 5' and 3' ends.

The two cDNA libraries are labelled with substances that fluoresce when excited by light. The most commonly used dyes are Cyanine 3 (Cy3, green) and Cyanine 5 (Cy5, red) or Alexa 488 and Alexa 594 (Staal et al., 2005). These fluorescent dyes have conjugated bonds which contain electrons that absorb and emit light (Do and Choi, 2007). These can be directly labelled during the cDNA synthesise and the mRNAs are degraded by sodium hydroxide. Indirect labelling consists of incorporating an aminoallyl-modified nucleotide during reverse transcription reaction. The incorporation itself if easier than the direct labelling but it requires an additional step which prolongs the procedure.

The two labelled cDNA samples are mixed in one microfuge tube and pipetted onto the microarray (Figure 5). There is a great variety of these 'chips' available (Table 2). The pipetted cDNA hybridizes with the probes over night or ideally 17 hours (Mantione et al., 2014) at 37°C.

Table 2. There is a great variety of microarrays optimised for different applications.

Name	Brief description
cDNA microarrays	most commonly used, complementary sequences used as probes
oligonucleotide microarrays	alternative to cDNA microarray, but higher specificity and reproducibility (more expensive)
exon array	each exon of a gene is measured individually
SNP arrays	single nucleotide polymorphism measured between individuals of a population
Chip-on-Chip Array	protein binding to different states of DNA with chromatin-immunoprecipitation

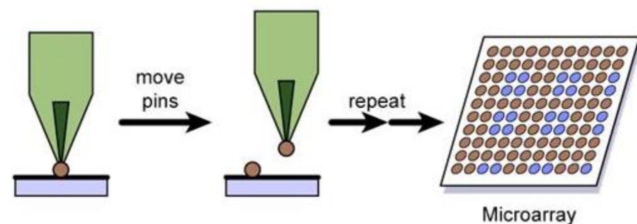

Figure 5. The mixed cDNA is pipetted with the arrayer, onto a glass slide (or silicon chip or nylon membrane) which is coated with DNA-binding chemicals (e.g. amino silane). This glass slide has a collection of ssDNA (probes) on the surface which represent different binding sites (complementary sequences) for genes. Illustration from SlidePlayer.

The microarray is washed three times (with SSC, SDS) to remove non-specific binding as in Northern blotting. The slide is then placed into a scanner. The first laser scans through the slide which excites Cy3 (532 nm) and the scanner, through a microscope takes a 16-bit grey-scale image. Then the second laser for the Cy5 dye (635 nm) produces another image. An image analysis software will replace the grey-scale pictures with green and red lights. The yellow colouring implies that both of the cDNAs bound to that location (Figure 4) and therefore that gene was expressed in both samples. The output of the microarray analysis is processed with image analysis (e.g. MAGIC tool software by Heyer et al., 2005) quantifying the expression of each gene in different samples.

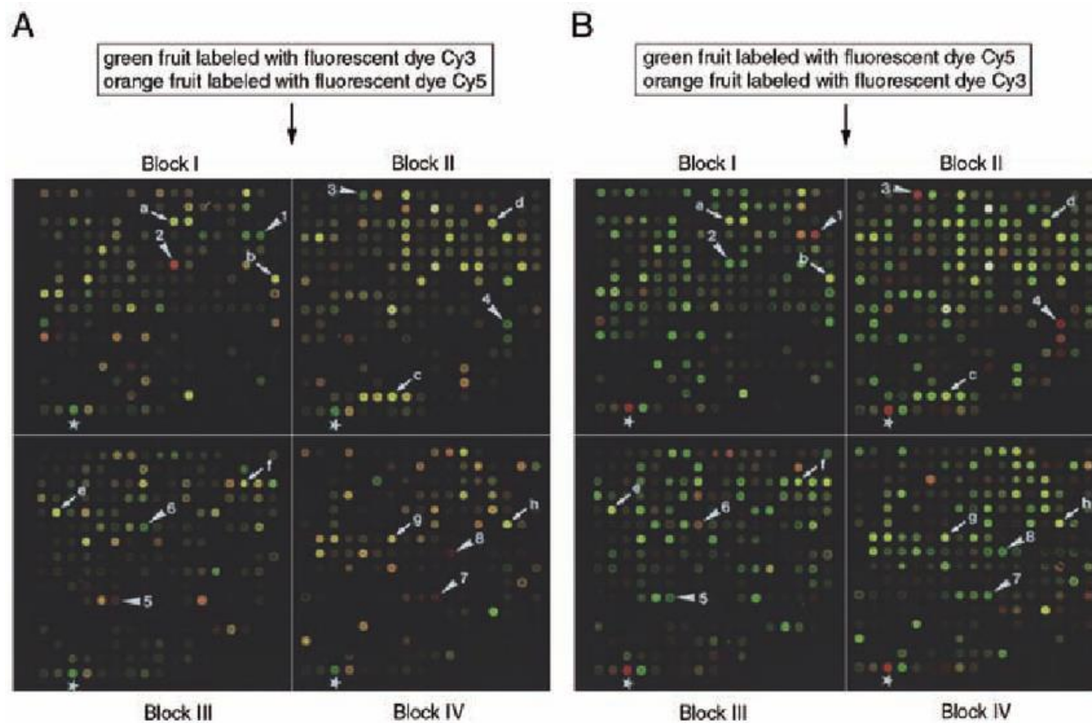

Figure 4. The final output by Wang et al. (1999) of developmental profiling by microarray analysis. The green fruit (green) and orange fruit (red) analysed in GeneSpring software. The intensity of the colours represents how strong the binding was to the probes. These two images are then combined. If a 'dot' on the microarray is green, then the first sample was bound to that specific probe (the gene was expressed), if it is red, then it is from the second sample. Overall, 98 up-regulated genes and 37 down-regulated genes showed five-fold difference in the orange coloured fruits, compared to green fruits and there was a different pattern throughout development.

Conclusion

This essay gave a brief overview of the first step of all gene expression studies, the RNA extraction, then discussed the Northern blot analysis and cDNA microarray analysis. Microarray analysis is estimated to cost $300 per sample (Mantione et al., 2014) while Northern blotting is estimated to cost 2.01€ per sample and about $700 for 34 samples and 10 genes and which takes around 6 days (Rumlow et al., 2016). When designing a gene expression study, it is important to decide whether more genes or more samples are worth researching. These methods can be used to complement each other – starting with low number of samples, exploring many genes and their expression and once genes of particular interest are identified, methods that enable fewer genes but more samples can validate the their function and mechanism (Figure 5).

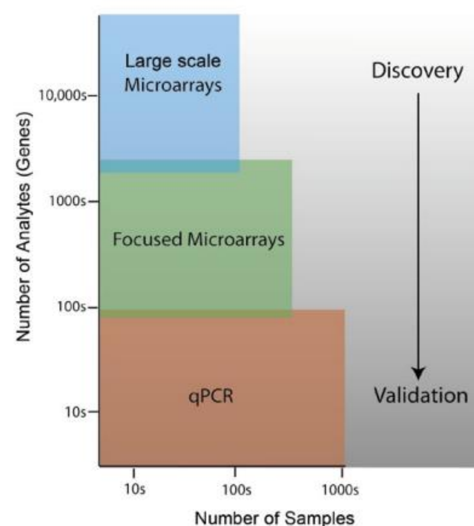

Figure 5. Different methods of gene expression studies enable researchers to move from discovery of genes to validating their functions and importance (VanGuilder et al., 2008).

Final word count: 1612

References

Baldi, P. and Hatfield, G.W., 2002. *DNA microarrays and gene expression: from experiments to data analysis and modelling*. Cambridge University Press

Bowtell, D.D., 1999. Options available—from start to finish—for obtaining expression data by microarray. *Nature Genetics*, *21*, pp.25-32

Chomczynski, P. and Mackey, K., 1994. One-hour downward capillary blotting of RNA at neutral pH. *Analytical Biochemistry*, *221*(2), pp.303-305

DeRisi, J., Penland, L., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A. and Trent, J.M., 1996. Use of a cDNA microarray to analyse gene expression. *Nature Genetics*, *14*, pp.457-460.

Do, J.H. and Choi, D.K., 2007. cDNA labeling strategies for microarrays using fluorescent dyes. *Engineering in Life Sciences*, *7*(1), pp.26-34

Fleming, J.C., Tartaglini, E., Steinkamp, M.P., Schorderet, D.F., Cohen, N. and Neufeld, E.J., 1999. The gene mutated in thiamine-responsive anaemia with diabetes and deafness (TRMA) encodes a functional thiamine transporter. *Nature Genetics*, *22*(3), pp.305-308

Gillespie, D. and Spiegelman, S., 1965. A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. *Journal of Molecular Biology*, *12*(3), pp.829-842

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A. and Bloomfield, C.D., 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. S*cience*, *286*(5439), pp.531-537

Görlach, J., Volrath, S., Knauf-Beiter, G., Hengy, G., Beckhove, U., Kogel, K.H., Oostendorp, M., Staub, T., Ward, E., Kessmann, H. and Ryals, J., 1996. Benzothiadiazole, a novel class of inducers of systemic acquired resistance, activates gene expression and disease resistance in wheat. *The Plant Cell*, *8*(4), pp.629-643

Greenberg, N.M., DeMayo, F., Finegold, M.J., Medina, D., Tilley, W.D., Aspinall, J.O., Cunha, G.R., Donjacour, A.A., Matusik, R.J. and Rosen, J.M., 1995. Prostate cancer in a transgenic mouse. *Proceedings of the National Academy of Sciences*, *92*(8), pp.3439-3443

Heyer, L.J., Moskowitz, D.Z., Abele, J.A., Karnik, P., Choi, D., Campbell, A.M., Oldham, E.E. and Akin, B.K., 2005. MAGIC Tool: integrated microarray data analysis. *Bioinformatics*, *21*(9), pp.2114-

Horgan, R.P. and Kenny, L.C., 2011. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician and Gynaecologist*, *13*(3), pp.189-195

Hu, C.J., Jiang, Q.Y., Zhang, T., Yin, Y.L., Li, F.N., Su, J.Y., Wu, G.Y. and Kong, X.F., 2017. Dietary supplementation with arginine and glutamic acid enhances key lipogenic gene expression in growing pigs. *Journal of Animal Science, 95(6), pp.2680-2689*

Huang, Q., Mao, Z., Li, S., Hu, J. and Zhu, Y., 2014. A non-radioactive method for small RNA detection by northern blotting. *Rice*, *7*(1), p.26

Iorio, M.V., Ferracin, M., Liu, C.G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M. and Ménard, S., 2005. MicroRNA gene expression deregulation in human breast cancer. *Cancer Research*, *65*(16), pp.7065-7070

Khandjian, E.W., 1986. UV crosslinking of RNA to nylon membrane enhances hybridization signals. *Molecular Biology Reports*, *11*(2), pp.107-115

Khandjian, E.W., 1987. Optimized hybridization of DNA blotted and fixed to nitrocellulose and nylon membranes. *Nature Biotechnology*, *5*(2), pp.165-167

Kim, S.W., Li, Z., Moore, P.S., Monaghan, A.P., Chang, Y., Nichols, M. and John, B., 2010. A sensitive non-radioactive northern blot method to detect small RNAs. *Nucleic Acids Research*, *38*(7), pp.e98-e98

Kirakosyan, A., Kaufman, P.B. and Tsai, C.J., 2016. 1 1 Northern Blot Hybridization. *Handbook of Molecular and Cellular Methods in Biology and Medicine*, p.211

Loor, J.J., Dann, H.M., Guretzky, N.A.J., Everts, R.E., Oliveira, R., Green, C.A., Litherland, N.B., Rodriguez-Zas, S.L., Lewin, H.A. and Drackley, J.K., 2006. Plane of nutrition prepartum alters hepatic gene expression and function in dairy cows as assessed by longitudinal transcript and metabolic profiling. *Physiological Genomics*, *27*(1), pp.29-41

Mantione, K.J., Kream, R.M., Kuzelova, H., Ptacek, R., Raboch, J., Samuel, J.M. and Stefano, G.B., 2014. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Medical science monitor basic research*, *20*, p.138

Nagaraj, S.H., Gasser, R.B. and Ranganathan, S., 2006. A hitchhiker's guide to expressed sequence tag (EST) analysis. *Briefings in Bioinformatics*, *8*(1), pp.6-21

Ozsolak, F. and Milos, P.M., 2010. RNA sequencing: advances, challenges and opportunities. *Nature reviews genetics*, *12*(2), p.nrg293

Pemberton, R.E., Liberti, P. and Baglioni, C., 1975. Isolation of messenger RNA from polysomes by chromatography on oligo (dT)-cellulose. *Analytical Biochemistry*, *66*(1), pp.18-28

Rabbani, M.A., Maruyama, K., Abe, H., Khan, M.A., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K., 2003. Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiology*, *133*(4), pp.1755-1767

Rumlow, A., Keunen, E., Klein, J., Pallmann, P., Riemenschneider, A., Cuypers, A. and Papenbrock, J., 2016. Quantitative expression analysis in Brassica napus by Northern blot analysis and reverse transcription-quantitative PCR in a complex experimental setting. *PloS One*, *11*(9), p.e0163679

Schmittgen, T.D., Zakrajsek, B.A., Mills, A.G., Gorn, V., Singer, M.J. and Reed, M.W., 2000. Quantitative reverse transcription–polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. *Analytical Biochemistry*, *285*(2), pp.194-204

Shinozaki, K., Yamaguchi-Shinozaki, K. and Seki, M., 2003. Regulatory network of gene expression in the drought and cold stress responses. *Current Opinion in Plant Biology*, *6*(5), pp.410-417

Staal, Y.C., van Herwijnen, M.H., van Schooten, F.J. and van Delft, J.H., 2005. Application of four dyes in gene expression analyses by microarrays. *BMC Genomics*, *6*(1), p.101

Struhl, K., 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell*, *98*(1), pp.1-4

Su, A.I., Cooke, M.P., Ching, K.A., Hakak, Y., Walker, J.R., Wiltshire, T., Orth, A.P., Vega, R.G., Sapinoso, L.M., Moqrich, A. and Patapoutian, A., 2002. Large-scale analysis of the human and mouse transcriptomes. *Proceedings of the National Academy of Sciences*, *99*(7), pp.4465-4470

Thurston, C.F., Perry, C.R. and Pollard, J.W., 1988. Electrophoresis of RNA denatured with glyoxal or formaldehyde. *New Nucleic Acid Techniques*, pp.1-11

VanGuilder, H.D., Vrana, K.E. and Freeman, W.M., 2008. Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques*, *44*(5), p.619

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B. and Kinzler, K.W., 1997. Characterization of the yeast transcriptome. *Cell*, *88*(2), pp.243-251

Wang, H.M., Yin, W.C., Wang, C.K. and To, K.Y., 2009. Isolation of functional RNA from different tissues of tomato suitable for developmental profiling by microarray analysis. *Botanical Studies*, *50*, pp.115-125

Wotton, H.J., Purcell, S.L., Covello, J.M., Koop, B.F. and Fast, M.D., 2014. Comparison of host selection and gene expression of adult Lepeophtheirus salmonis and Salmo salar during a cohabitation of initially infected and uninfected fish, *Journal of Aquaculture Research and Development, 5(3),* pp.1-8

Yamamoto, M., Wakatsuki, T., Hada, A. and Ryo, A., 2001. Use of serial analysis of gene expression (SAGE) technology. *Journal of Immunological Methods*, *250*(1), pp.45-66

Zhao, C. and Hamilton, T., 2007. Introns regulate the rate of unstable mRNA decay. *Journal of Biological Chemistry*, *282*(28), pp.20230-20237

Zhao, N., Hashida, H., Takahashi, N., Misumi, Y. and Sakaki, Y., 1995. High-density cDNA filter analysis: a novel approach for large-scale, quantitative analysis of gene expression. *Gene*, *156*(2), pp.207-213

Zong, H., Ren, J.M., Young, L.H., Pypaert, M., Mu, J., Birnbaum, M.J. and Shulman, G.I., 2002. AMP kinase is required for mitochondrial biogenesis in skeletal muscle in response to chronic energy deprivation. *Proceedings of the National Academy of Sciences*, *99*(25), pp.15983-15987