Chapter 3

Introduction to the Gene Expression Analysis

Ignacio San Segundo-Val and Catalina S. Sanz-Lozano

Abstract

In 1941, Beadle and Tatum published experiments that would explain the basis of the central dogma of molecular biology, whereby the DNA through an intermediate molecule, called RNA, results proteins that perform the functions in cells. Currently, biomedical research attempts to explain the mechanisms by which develops a particular disease, for this reason, gene expression studies have proven to be a great resource. Strictly, the term "gene expression" comprises from the gene activation until the mature protein is located in its corresponding compartment to perform its function and contribute to the expression of the phenotype of cell.

The expression studies are directed to detect and quantify messenger RNA (mRNA) levels of a specific gene. The development of the RNA-based gene expression studies began with the Northern Blot by Alwine et al. in 1977. In 1969, Gall and Pardue and John et al. independently developed the in situ hybridization, but this technique was not employed to detect mRNA until 1986 by Coghlan. Today, many of the techniques for quantification of RNA are deprecated because other new techniques provide more information. Currently the most widely used techniques are qPCR, expression microarrays, and RNAseq for the transcriptome analysis. In this chapter, these techniques will be reviewed.

Key words Gene expression, qPCR, Microarrays, RNA sequencing, Transcriptome

1 Quantitative PCR

In 1984, Kary Mullis developed the Polymerase Chain Reaction (PCR). This technique amplifies a specific segment of DNA to obtain hundreds of millions of copies in few hours [1]. It was used for qualitative studies and it was not until 1992 when Higuchi et al. developed the quantitative Polymerase Chain Reaction (qPCR), which employs the PCR technique to gene expression studies.

To do this, they used the same material as for conventional PCR, a pair of specific oligonucleotides as primers; deoxynucleotide triphosphates (dNTPs); a reaction buffer; a thermo stable DNA polymerase; and they added a fluorochrome that fluoresces when excited [2]. Additionally, they designed a system capable of detecting in real-time PCR products accumulated. The system uses

María Isidoro-García (ed.), *Molecular Genetics of Asthma*, Methods in Molecular Biology, vol. 1434,
DOI 10.1007/978-1-4939-3652-6_3,

a camera that detects the increase in fluorescence that occurs when the ethidium bromide is intercalated in new strands of DNA formed in each cycle [3, 4]. Therefore, in the real-time PCR or qPCR, the processes of amplification and detection occur simultaneously in the same vial.

The PCR reaction consists in a series of cyclical temperature changes. Each cycle is divided into three stages [5]:

- Denaturation: separation of double-stranded DNA when subjected to 95 °C.
- Hybridization of primers: alignment of the primers to the DNA template at temperature around 50–60 °C.
- Elongation or Polymerization: binding of the corresponding dNTPs, to the DNA elongation chain at temperature around 68–72 °C.

Theoretically, if the reaction efficiency is 100%, the number of DNA molecules will double with each cycle. The reality is that efficiency in optimal conditions will be slightly lesser than 100% [6].

1.1 Methods of Detection

The qPCR technique allows the quantification of starting material (DNA, cDNA, or RNA), by using fluorophores. The fluorescence is measured in each cycle and is proportional to the amount of the PCR product.

On the one hand, fluorophores can be fluorescent dyes that nonspecifically bind to double-stranded DNA and produce a fluorescent amount that correlates with the DNA copy number. In this group we find dyes such as SYBR Green [7]. On the other hand, can be used fluorochromes attached to probes, which specifically hybridize to the amplified DNA strands. Thus, the reaction is more specific and the signal only is generated when the probe hybridizes with its complementary region. In this group, we find hydrolysis probes "TaqMan," FRET probes, Beacon probes, and Scorpions probes [8].

1.1.1 Nonspecific Fluorescent Dyes

SYBR Green I is an organic compound with the chemical formula $C_{32}H_{37}N_4S$ [9]. It is associated with the DNA molecule by interacting with the minor groove. It is used in DNA staining for electrophoresis analysis of PCR products, or as a means of direct visualization of the PCR products in real time [9].

The melting curve analysis is necessary to detect problems of nonspecific binding of SYBR Green to DNA. Comparison of the melting curves allows an interpretation of the amplified products in PCR. Different products have different temperatures because the melting temperature depends on the size of the amplicon, the GC content, and the secondary and tertiary structures. At the end of the PCR reaction, a gradient from 50 to 95 °C is made to denature the double-stranded DNA. When the double-stranded DNA

becomes single-stranded DNA, the decrease in fluorescence can be observed as peaks, represented by performing the second derivative of fluorescence. If the reaction is not specific, the PCR products show different melting curve peaks [10]. Since more than a decade there are other fluorescent dyes such as LC Green, ResoLight, EvaGreen, Chromofy SYTO, and BEBO [11].

1.1.2 Sequence Specific Probes

- TaqMan® probe: belongs to the group of hydrolysis probes and is based on the 5′-3′ exonucleasa activity of the Taq polymerase. The Taqman probe is a sequence complementary to a PCR product that is not part of the primers. The probe is labeled with a fluorophore that covalently bonds to its 5′ end called "donor" and with a quencher covalently bonded to its 3′ end also called "acceptor" whose function is to quench the fluorescence emitted by the fluorophore when is excited by the light. The 5′ exonuclease activity degrades the probe and releases the fluorophore that will emit fluorescence when not being close to the quencher. The liberated fluorophore is proportional to the amount of DNA amplified in the PCR. The probe should be close to a primer and the amplicon size should not be greater than 200 pair bases. Another drawback is that these probes do not allow melting curve analysis because the hydrolysis of the probe prevents its reutilization [12, 13].
- Molecular Beacons: they have a fluorophore covalently linked to 5′ called "donor" and a quencher covalently linked to 3′ called "acceptor"; they also have a hairpin-like secondary structure which brings the fluorophore to the quencher. When the probe specifically hybridizes to DNA, the distance between the fluorophore to the quencher is opened, allowing registering the signal of the annealing phase. The difficulty is in the probe designing because the secondary hairpin structure of the probe–amplicon hybrid should be more stable than that formed by the molecular beacon [12–14].
- FRET probe: two probes system in which one oligonucleotide contains a donor fluorophore and the acceptor fluorophore is in another oligonucleotide. When both probes bind, the fluorophores are close and the power is transmitted from the donor to acceptor emitting fluorescence. A good design of the two primers and the two probes is critical to obtain good results [12–14].
- Scorpion probe: it acts both as a probe and as a PCR primer and has hairpin structure. The molecule has a fluorophore and a quencher similar to molecular beacon probes that bind to amplicon by the same principle. The reaction leading to the fluorescent signal is immediate. This is because it is attached to primer and not collides with targeted region; therefore, signals are stronger [14].

- MGB probe (Minor Groove Binder): it is considered a variant of Taqman probe but it is not hydrolyzed and has the bases modified. The probe is labeled with a fluorophore in 3′ and in 5′ with a quencher. The Minor Groove Binder is a part of the probe that binds to minor groove of DNA. Minor Groove Binder allows protecting the probe from the 5′-3′ exonucleasa activity of the polymerase. When the probe hybridizes the fluorophore is separated of quencher and a fluorescent signal is generated. The fluorescent signal is directly proportional to DNA quantity. An advantage of these probes is that they are shorter and more stable. They include superbases which are bases modified to increase the temperature of melting curve [12, 13].

1.2 Quantitative PCR Characteristics

1.2.1 Advantages of the Technique

The main advantages of this technique are as follows [15]:

- Speed: the testing time is approximately 1 h.
- Simplicity: the assay requires a pair of primers, an enzyme, dNTPs, and optionally a probe.
- Convenience: does not require postamplification processing.
- Sensitivity: samples with very few copies of messenger RNA (mRNA) can be quantified.
- Specificity: a well-designed assay is specific for a single target.
- Robustness: a well-designed trial will give results in a wide range of reaction conditions.
- High performance: thousands of reactions can be carried out in a single experiment.
- Familiarity: the PCR is well known; their advantages and disadvantages are well understood.
- Cost: the price of reagents is affordable and with small reaction volumes, the costs decrease.

1.2.2 Basic Terms of qPCR

All systems similarly quantify and record fluorescent signal each cycle of the PCR. The reaction kinetics shows a representation forming a sigmoid shape. The graphical representation is defined by the background baseline, threshold, and threshold cycle [13] (*see* Fig. 1).

- Background: Fluorescence not specific of the reaction. The mathematical algorithm removes it.
- Baseline: noise level in early cycles of PCR where it is not detected an increase in fluorescence of PCR product. Determines basal fluorescence.
- Threshold: threshold value is set just above the baseline level where the exponential amplification begins. Threshold level can be determined automatically or manually.

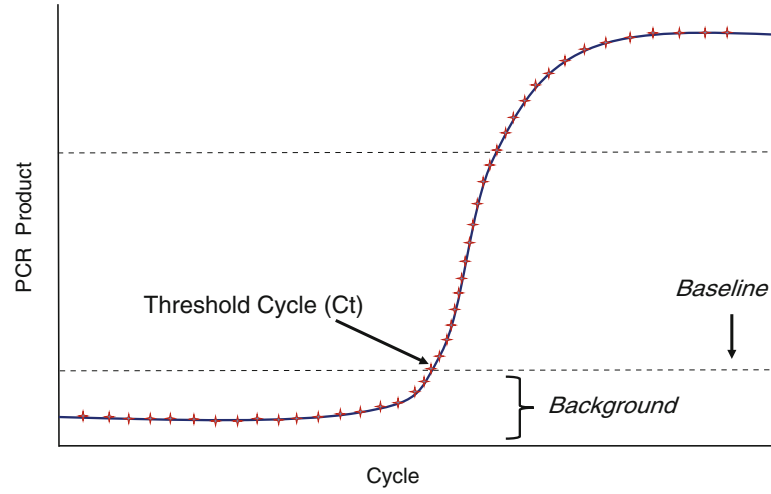

Fig. 1 Basic terms of qPCR

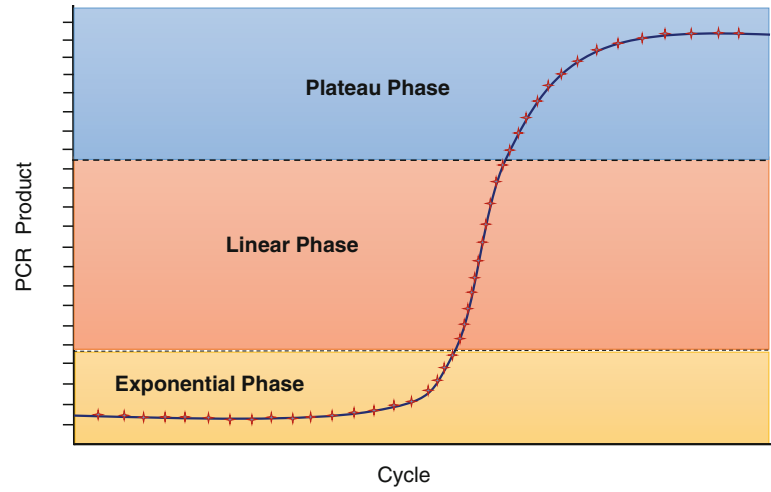

Fig. 2 Phases of qPCR

- Threshold cycle: the cycle at which the fluorescence exceeds the threshold. This cycle is used for relative quantification of gene expression.

1.2.3 *Phases of qPCR*

The qPCR reaction involves three phases that may be represented as a sigmoidal curve [16]. The three phases are as follows [12] (Fig. 2):

- Exponential phase: the amount of product is small, the PCR product is exponentially generated because of that the enzyme and reagents are not limited, so that reaction can reach the maximum efficiency. This exponential growth is difficult to detect because the quantity of product is insufficient. The amount of

product at this stage is proportional to the amount of initial sample.

- Linear phase: the amount of product increases linearly because the quantity of enzyme and reagents begin to be limited so that reaction efficiency decreases.
- Plateau Phase: the reaction slows until the dNTPs and primers needed for new synthesis are depleted.

The cycle at which the fluorescence begins to exceed the background level is called the threshold cycle (Ct) and is the beginning of the logarithmic phase. Therefore, the Ct is inversely correlated with the amount of sample: the lower Ct, the greater cDNA amount.

1.2.4 Quantification

To quantify the expression of the gene under study there are two methods, called absolute quantification and relative quantification. In the first case, the absolute expression amount is expressed as number of copies obtained. In the second, the quantification is based on a calibrator, whose value is benchmark for all others, assigning it a value of 1 [17].

To perform absolute quantification is necessary to know the number of copies of target gene in a standard sample. Generally, this standard sample is a plasmid DNA or complementary DNA (cDNA) whose concentrations are measured by a spectrophotometer. In the assay, serial dilutions of the standard sample in which the target gene will be amplified should be prepared.

The computer registers the threshold cycles (Ct) for each sample. The dilutions of a standard sample show a standard curve that reflects the number of copies by interpolation of the Ct and allows quantifying the samples [12, 17].

There are three methods for relative quantification; the most widely used is the method 2-ΔΔCT [12]. This method assumes that the procedure doubles the DNA content in each cycle, that reaction efficiency is 100%, and that a reference gene is expressed at a constant level between all samples [18]. This reference gene is a constitutive gene that is used as an endogenous control to correct intra and inter-assay variability.

The expression of the reference gene is changeless in the different samples of the assay as it is a gen which function is related to the cell maintenance, therefore, it is also called constitutive gene. The following formula expresses the ratio obtained from the relationship between the Ct values of the sample and the Ct values of the calibrator:

$$\text{Ratio} = 2^{-[\Delta Ct\ \text{sample} - \Delta Ct\ \text{calibrator}]}$$

$$\text{Ratio} = 2^{-\Delta\Delta Ct}$$

ΔCt sample is the difference between the Ct of the gene under study in sample minus the Ct of the reference gene in sample; ΔCt calibrator is difference between the Ct of the gene under study in the calibrator minus the Ct of the reference gene in the calibrator.

Another model for relative quantification was proposed by Pfaffl [19, 20]. In this model, different PCR efficiencies of the genes under study are taken into account, as shown in the following equation:

$$\text{Ratio}(\text{fold}) = \frac{(E \text{ target gene})\,\Delta\text{Ct target}\,(\text{calibrator} - \text{sample})}{(E \text{ reference gene})\,\Delta\text{Ct reference}\,(\text{calibrator} - \text{sample})}$$

In this model, it is necessary to know the efficiency of each pair of primers for each gene. The efficiencies are obtained from the slopes of standard curves obtained from serial dilutions according to the formula [16]:

$E = 10[-1/\text{slope}] - 1$

The third model is called standard curve method or method-E that analyzes the efficiency of the gene under study and reference gene using standards, for which serial dilutions of a single sample are made. The standard curve method calculates the efficiency for each pair of primers of each gene [19, 21].

1.2.5 Selection of Reference Genes

Selection of the proper reference gene is a critical step to assess correctly data obtained by quantitative PCR. Most authors agree that the use of reference genes is the most effective and simple method for correcting bugs and glitches such as [8, 22]:

- Problems in the process of extraction, purification, or storage of RNA
- Bad performance of reverse transcription to cDNA synthesis
- Errors in pipetting or on transferring of material
- Polymerase inhibitors
- Poorly designed primers
- Inappropriate statistical analysis

Many qPCR experiments have been wrongly designed and are difficult to reproduce due to poor quality of data [8]. For this reason, in recent years the process of normalization of reference genes has become a recurring problem addressed by scientists. This has led to development of a variety of protocols and methodologies. Moreover, related publications showed big differences regarding the published information on development of their researches. This is being solved by guidelines for all publications associated with this methodology that are published in the MIQE guide (Minimum Information for Publication of Quantitative Digital PCR Experiments) [23].

Most publications agree that constitutive genes should show minimal variability among the tissues, cells, physiological conditions, or treatments under investigation [8, 24]. Therefore, it is necessary validate the reference genes for each tissue and treatment analyzed. Furthermore, it is desirable that the constitutive gene

presents a threshold cycle (Ct) as close as possible to the problem gene, but it is not always possible, so that general recommendations should be followed. These recommendations do not advise choosing a constitutive gene with low expression (Ct > 30) or with high expression (Ct < 15) [8].

It is always recommended to use at least two reference genes, since the use of one can only result in relatively large errors [8, 23]. Although the use of a single reference gene is acceptable if it was previously tested in an experiment with similar conditions and was properly validated, it should be avoided to minimize the possible bias [8].

Several programs and models to select reference genes have been proposed with different statistical approaches and algorithms. Among the best-known applications are statistics to normalize, Normfinder, geNorm, and BestKeeper. Although there are approaches such as the ΔCt method or the classic ANOVA model for comparing the stability of the reference gene expression in the different study conditions by comparing the average Cts [25].

It also should be mentioned that many algorithms exist to study the variability of the constitutive gene expression, therefore the classification of the candidate reference gene might vary depending on the software or statistical technique used [26–29]. Finally, in selecting reference genes is advisable to select genes with different functions that do not have a common regulation that may affect its expression [8].

2 Microarrays

The Microarrays technique is based on complementarity between nucleic acid strands allowing detecting specific sequences by what is called hybridization [30]. The in situ hybridization was independently used for the first time by Gall and Pardue, and by John et al. in the same year. It was not used until 1986 to detect mRNA, by Coghlan [31]. In the late 1980s, Ekins et al. in University College London developed the first array for immunoassay studies, this array began to be manufactured and marketed in 1991 [32].

In the field of gene expression, it could study between 500 and 18,000 cDNAs. These human cDNAs were obtained from bacterial libraries and generally were radiolabeled with 33P-dNTPs. The most important limitations of macroarrays were the low density of probes onto the nylon support, large volumes of sample necessary for hybridization and that bacterial libraries could not be composed of pure colonies [33].

Technological efforts were focused on the miniaturization of arrays to overcome the limitations of the existing technology. In 1991, a group formed by Stephen Fodor, Leighton Read, Michael Pirrung, and Luberc Stryer fabricated a microarray. In 1993,

Stephen Fodor creates a spin-off that was dedicated to the microarrays development and in 1994 began the manufacture and sale of microarrays [34, 35].

Later, Patrick Brown used a glass support allowing a higher density of probes. Miniaturization decreases the amount of sample required for the study of gene expression, on the other hand, the use of radioactively labeled nucleotides was replaced by fluorophores [36]. In 1996, discloses the design, tools, and knowledge to let other research groups can make their microarrays in their laboratories. This information will boost the use of DNA microarrays [33].

The microarray, also known as DNA chip or biochip, is a solid support of glass, plastic, or nylon which is joined to oligonucleotides which sequences corresponds to all regions of the genome. The manufactures used the same technology used in the semiconductor chips but vertically placing million DNA strands on the support. Each oligonucleotide is placed in a specific area called "probe cell," where billions of copies of the oligonucleotide are found. These oligonucleotides are synthesized prior to bonding to the support or on the support. In the second case, there are many different methods: photolithography, phosphoramidite injection, or activation of precursors by an electric field [37, 38].

2.1 Types of Microarrays

Two types of microarrays are distinguished: expression microarrays, where specific RNA sequences are detected and genotyping microarrays for detecting specific DNA sequences [38]. By comparing the results of both arrays it could be established the relationship of polymorphisms with gene expression, which, among others, it could explain the different response to the treatment observed among patients with different genetic background.

2.1.1 Oligonucleotide Microarrays

Fragments of 25 specific base pairs are immobilized while chemically synthesized. By homogeneity, reproducibility, robustness, and high density are the most used since they can study up to 20,000 genes at once [38]. The main limitation is the cost of specific oligonucleotides selection and synthesis; it is more than three times more expensive than a cDNA microarray. The most important advantage is that photolithography can be used to direct synthesis of oligonucleotides which allows a high density of probes. Another advantage is that targeting sequence synthesis prevents cross-hybridization between sequences related genes. Furthermore, all the oligonucleotides of the microarray are of the same size, the same temperature melting, and the same concentration, therefore, the experimental variation is decreased, and the statistical power is increased [39].

2.1.2 cDNA Microarrays

The cDNA probes of 600–2000 base pairs are immobilized on a glass, nylon, or nitrocellulose base [38]. The main advantage is that these microarrays are cheaper than oligonucleotide

microarrays. Another advantage is that it can be created from cDNA libraries, many of them of public domain.

The most important limitation is the data treatment due to the possibility of cross-hybridization that occurs between related genes. Another important limitation is that the variations of probe sizes and melting temperatures can diminish the statistical power [39].

On the other hand, these microarray experiments can be used in co-hybridization of two colors. These experiments allow direct comparison of mRNA abundance in two populations, although by this approach it is obtained comparative ratios rather than absolute levels of expression. On the other hand, the use of ratios reduces the inter-assay variation [39].

2.2 Methodology in Microarrays

The methodology of this technique includes the following steps [38, 40]:

- Extraction and preparation of RNA: the RNA is extracted from specific tissues, trying to obtain RNA with the highest purity and quality possible and thus avoiding a major source of variability.
- RNA amplification: the RNA is amplified by PCR to facilitate hybridization because the amount of mRNA in cells may not be sufficient.
- Reverse transcription: to convert mRNA into cDNA.
- Labeling the probes: the cDNA is fragmented and labeled with biotin. After, the fluorescent molecule that binds to biotin is added.
- Hybridization of the probes: the time required for complete hybridization is directly proportional to sample concentration. At the end of this process, the hybridized microarray is rinsed to remove unbound chains.
- Scanning of microarray: the fluorescent light detection indicates that hybridization has occurred at a specific point to a specific sequence. Reading is performed by a laser and the fluorescence is recorded by scanning. The fluorescence intensity is proportional to the amount of probe bound to each sample.

2.3 Limitations of Microarrays

The expression microarrays have produced a lot of information, but have limitations. Some of the main limitations of using microarrays are as follows [41, 42]:

- Limited dynamic range of detection: the detection of expression levels is limited to two or three orders of magnitude due to the background and the signal saturation. Thus, they are not suitable for studies of genes with low or very high expression.
- Reliance on existing knowledge: only can study known genome sequences, an errors in the database can worsen outcomes.

- Difficult comparison between experiments: they require complicated normalization methods for comparing expression levels from different experiments.
- Cross Hybridization: between genes of similar sequence; generates background and reduces the dynamic range.
- New transcripts can not be detected: new transcripts produced by alternative splicing can not be detected if the sequence is unknown.
- Necessity of validation: the microarrays need to be validated by qPCR for obtaining reliable microarray expression data. To do this, you must select and validate by using appropriate reference genes [43].

3 RNA Sequencing (RNAseq)

In 1964, Holley made the first complete sequencing of a gene. In the 1970s, Maxam and Gilbert developed a DNA sequencing technology based on chemical modification of DNA and subsequent cleavage at specific bases, while Sanger developed a DNA sequencing method based on the chain termination method. The Sanger sequencing was imposed by high efficiency and low radioactivity as the first-generation sequencing [44, 45].

The first automatic sequencer that used the Sanger method appeared in 1987 adopted the capillary electrophoresis as a more accurate and faster sequencing method. These sequencers have evolved from the 500 kilobases sequenced per day with the first model that appeared on the market until the 2.88 Megabases of the current model. These sequencers allowed to complete the human genome project in 2001 [45].

In 2005, the great revolution occurred in the field of DNA sequencing when the first sequencing of high performance appeared in the market. Several companies are now responsible for the development of high-throughput sequencers [45].

3.1 Transcriptome and RNA Sequencing

The high-throughput sequencers allow investigating the transcriptome. The transcriptome is the set of ribonucleic acids in the cell, including messenger ribonucleic acid (mRNA), transfer ribonucleic acid (tRNA), ribosomal ribonucleic acid (rRNA), small nuclear ribonucleicacid (snRNA), noncoding ribonucleic acids (ncRNA), and others. These RNAs are differently expressed according to the tissue, the physiological condition, or the stage of development [46].

The interpretation of the transcriptome complexity is a crucial objective to understand the functional elements of the genome and thus the functioning of the disease and its progression. In this sense, it has recently been shown that the amount of noncoding

DNA increases with the complexity of the organism, 0.25% in the prokaryotic genome and 98.8% in the human genome.

The existing level of complexity attached to the discoveries of endogenous small interfering RNA (siRNA), long interspersed noncoding RNA (lincRNA), transcription initiation RNA (tiRNA), microRNAs (miRNAs), transcription start site-associated RNA (TSSa-RNA) among others, represent parts of the transcription puzzle that we must decipher to understand how the genome works [45].

For this, the new RNA sequencing technology must start with cataloging all RNAs from mRNA; through the noncoding RNA to reach small RNAs determine transcription start sites and quantify changes in the expression levels of genes during development and in different conditions [41].

RNA sequencing is a technique in which sequenced fragments of 25–400 base pairs are used depending on the technology. To this end, a population of RNAs is fractionated and is transformed into a population of cDNA. These cDNAs are joined to adapters by one or both ends. Each cDNA molecule is sequenced to obtain short readings of one or both ends [41].

3.2 RNAseq Methodology

The design of a transcriptome study follows these steps [46]:

1. Selection of the tissue of interest in which the RNAs are to be studied.
2. Building the cDNA libraries.
3. Using a massive sequencing system.
4. Analyze the data using bioinformatics tools. Million short readings are obtained, which are mapped to a genome or transcriptome. The reads must be aligned to summarize all data. Finally, the data are normalized and the statistical tests are applied for studying differential gene expression, resulting in a classification of genes with their expression levels, the *p*-values, and the fold-changes [47].

There are several methods and programs to treat RNA sequencing data which aim not only estimating the differences in expression levels between samples, but also, detecting alternative splicing, analyzing the RNA editing, calculating the abundance of a transcript, etc. [47].

3.3 Advantages and Limitations of RNAseq

The RNAseq technology presents several advantages over other methodologies that aim to study gene expression:

- Resolution of a base: it can also detect changes in a single nucleotide (SNP) in the transcripts, microsatellites, isoforms, and allelic variants.
- Wide Dynamic Range: no upper limit for quantification and correlates with the number of sequences obtained, consequently

has a wide dynamic range of expression levels, estimated four to five orders of magnitude. So, it can analyze genes expressed at very low or very high levels that are not detected by other techniques. The low background-signal helps to improve the dynamic range.

- Short Readings: readings from 30 bp to allow accurate information about how to connect two exons.
- High accurate: it has been shown by validation studies with quantitative PCR.
- High levels of reproducibility
- Lower RNA sample: because there are no cloning steps.
- There is no need of reference genome: a reading can be performed without reference genome and the transcriptome can be de novo assembled. This is an advantage for species whose genome has not been sequenced yet [41, 45, 46].

However, this technique also presents some limitations as the high cost. It requires lots of resources compared to other techniques. The Data set is large and complex, the large amount of generated data makes the interpretation difficult, and usually a bioinformatic adviser is needed [46].

Acknowledgments

This work was supported by grants of the Junta de Castilla y León ref. GRS1047/A/14, GRS1189/A/15, and BIO/SA73/15; and by the project "Efecto del Ácido Retinóico en la enfermedad alérgica. Estudio transcripcional y su traslación a la clínica," PI13/00564, integrated into the "Plan Estatal de I+D+I 2013–2016" and cofunded by the "ISCIII-Subdirección General de Evaluación y Fomento de la investigación" and the European Regional Development Fund (FEDER).

References

1. Saiki RK, Scharf S, Faloona F et al (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350–1354
2. Mallona I (2008) Selección de genes de normalización para RT-PCR cuantitativa en Petunia hybrida. (Normalization gene selection for quantitative RT-PCR in Petunia hybrida). Available via http://repositorio.bib.upct.es/dspace/handle/10317/723. Accessed 25 Nov 2014
3. Higuchi R, Dollinger G, Walsh PS et al (1992) Simultaneous amplification and detection of specific DNA sequences. Biotechnology 10:413–417
4. Higuchi R, Fockler C, Dollinger G et al (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology 11:1026–1030
5. Clewley JP (1994) The polymerase chain reaction (PCR) for human viral diagnosis. CRC Press, Boca Ratón
6. Taylor S, Wakem M, Dijkman G et al (2010) A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. Methods 50:S1–S5

7. Huggett J, Bustin S (2011) Standardization and reporting for nucleic acid quantification. Accred Qual Assur 16:399–405
8. Kozera B, Rapacz M (2013) Reference genes in real-time PCR. J Appl Genet 54:391–406
9. Zipper H, Brunner H, Bernhagen J et al (2004) Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. Nucleic Acids Res 32(12):e103
10. Sigma-Aldrich (2008) qPCR Technical Guide. Available via http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/General_Information/qpcr_technical_guide.pdf. Accessed 22 Nov 2014
11. http://www.gene-quantification.de/hrm-dyes.html
12. VanGuilder HD, Vrana KE, Freeman WM (2008) Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques 44:619–626
13. Qiagen (2006) Critical factors for successful real time PCR. Integrated solutions-real time PCR applications. Available via http://jornades.uab.cat/workshopmrama/sites/jornades.uab.cat.workshopmrama/files/Critical_factors_successful_real_time_PCR.pdf. Accessed 28 Nov 2014
14. Leonard DGB (2007) Molecular pathology in clinical practice. Springer Science & Business Media, Berlin
15. Bustin SA, Kessler HH (2010) Amplification and detection methods. In: Kessler HH (ed) Molecular diagnostics of infectious diseases. De Gruyter, Berlín
16. Louw TM, Booth CS, Pienaar E et al (2011) Experimental validation of a fundamental model for PCR efficiency. Chem Eng Sci 66:1783–1789
17. Diez GO (2006) Técnicas de Genética Molecular II (Molecular Genetic Techniques II). In: Lasa A (ed) PCR cuantitativa (quantitative PCR). SEQC, Barcelona
18. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25:402–408
19. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:2002–2007
20. Pfaffl MW (2004) Quantification strategies in real-time PCR. In: Bustin SA (ed) A-Z of quantitative PCR. International University Line, La Jolla
21. Bohla L, Dusanic D, Narat M et al (2012) Comparison of methods for relative quantification of gene expression using real-time PCR. Acta Agric Slov 100:97–106
22. Mallona I, Lischewski S, Weiss J et al (2010) Validation of reference genes for quantitative real-time PCR during leaf and flower development in Petunia hybrida. BMC Plant Biol 10:4
23. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622
24. Valente V, Teixeira SA, Neder L et al (2009) Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR. BMC Mol Biol 10:17
25. Podevin N, Krauss A, Henry I et al (2012) Selection and validation of reference genes for quantitative RT-PCR expression studies of the non-model crop Musa. Mol Breed 30:1237–1252
26. Fu W, Xie W, Zhang Z et al (2013) Exploring valid reference genes for quantitative real-time PCR analysis in *Plutella xylostella*. Int J Biol Sci 9:792–802
27. Gantasala NP, Papolu PK, Thakur PK et al (2013) Selection and validation of reference genes for quantitative gene expression studies by real-time PCR in eggplant (Solanum melongena L). BMC Res Notes 6:312
28. Paim RM, Pereira MH, Di Ponzio R et al (2012) Validation of reference genes for expression analysis in the salivary gland and the intestine of Rhodniusprolixus (Hemiptera, Reduviidae) under different experimental conditions by quantitative real-time PCR. BMC Res Notes 5:128
29. Tunbridge EM, Eastwood SL, Harrison PJ (2011) Changed relative to what? Housekeeping genes and normalization strategies in human brain gene expression studies. Biol Psychiatry 69:173–179
30. Southern E, Mir K, Schepinov M (1999) Molecular interactions on microarrays. Nat Genet 21:5–9
31. Eberwine JH, Valentino KL, Barchas JD (1994) In situ hybridization in neurobiology: advances in methodology. Oxford University Press, Oxford
32. McLachlan G, Do K, Ambroise C (2005) Analyzing microarray gene expression data. Wiley, Hoboken
33. Faiz A, Burgess JK (2012) How can microarrays unlock asthma? J Allergy 2012:241314
34. Affymetrix (2002) Affymetrix, Stanford University and incyte resolve patent oppositions and interferences. Available via http://investor.affymetrix.com/phoenix.

zhtml?c=116408&p=irol-newsArticle_pf&ID=362094. Accessed 21 Dec 2014
35. Times Higher Education (2006) Background memo on the winners of the European inventor of the year 2006 awards. Available via http://www.timeshighereducation.co.uk/news/background-memo-on-the-winners-of-the-european-inventor-of-the-year-2006-awards/203002.article. Accessed 22 Dec 2014
36. Shalon D, Smith SJ, Brown PO (1996) A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res 6:639–645
37. Lopez M, Mallorquín P, Vega M (2002) Microarrays y biochips de DNA, Informe de vigilancia tecnológica (DNA microarrays and biochips, technological surveillance report). Genoma España/CIBT-FGUAM
38. Daudén E (2007) Farmacogenética II. Métodos moleculares de estudio, bioinformática y aspectos éticos (Molecular study methods, bioinformatics and ethical aspects). Actas Dermosifiliogr 98:3–13
39. Alba R, Fei Z, Payton P et al (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. Plant J 39:697–714
40. Lin SM, Johnson KF (2002) Methods of microarray data analysis II. Springer Science & Business Media, Berlin
41. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63
42. Malone JH, Oliver B (2011) Microarray, deep sequencing and the true measure of the transcriptome. BMC Biol 9:34
43. Fernández AI, Óvilo C, Fernández A et al (2008) Luces y sombras del análisis de expresión génica utilizando microarrays. Un ejemplo en cerdo ibérico (Lights and shadows of gene expression analysis using microarrays. An example Iberian pig.) ITEA 104:99–105
44. Liu L, Li Y, Li S et al (2012) Comparison of next-generation sequencing systems. J Biomed Biotechnol 2012:251364
45. Costa V, Angelini C, De Feis I et al (2010) Uncovering the complexity of transcriptomes with RNA-Seq. J Biomed Biotechnol 2010: 853916
46. Santos CA, Blanck DV, de Freitas PD (2014) RNA-seq as a powerful tool for penaeid shrimp genetic progress. Front Genet 5:298
47. Oshlack A, Robinson MD, Young MD (2010) From RNA-seq reads to differential expression results. Genome Biol 11:220