Expert Opinion

Ashley Publications
www.ashley-pub.com

High-throughput analysis of mRNA expression: microarrays are not the whole story

Stefan Lorkowski† & Paul Cullen
†*Institute of Arteriosclerosis Research, University of Münster, Domagkstraße 3, 48149 Münster, Germany*

Analysis of gene expression has become one of the most widely used techniques in biomedical research. The publication of the human genome, together with the technical ability to automatically detect and process very large numbers of samples, has forced many research groups in academia and industry to shift to the complex and demanding study of genome-wide gene expression. DNA microarrays or 'gene chips', have become one of the most popular techniques for this purpose but their use is limited by technical deficiencies and the fact that prior knowledge on the genome of interest is a prerequisite. During the last few years, several approaches have been developed that overcome many of the limitations of DNA microarrays, while yielding greater sensitivity and a greater dynamic range in the absence of prior genomic knowledge. This review is an overview of selected high-throughput methods that provide a viable alternative to the use of DNA microarrays for analysing gene expression. It is a companion paper to one specifically on DNA microarrays, recently published in this journal by the same authors.

Keywords: ADGE, AFLP, DEPD, differential display, GeneCalling, genome-wide expression profiling, microarrays, MPSS, OFP, real-time RT-PCR, RNA fingerprinting, SAGE, TALEST, TOGA

Expert Opin. Ther. Patents (2004) 14(3):377-403

1. Introduction

During the past 30 years, since the development of approaches such as northern blotting, nuclear run-on and western blotting, techniques for the analysis of gene expression have become one of the most widely used set of methods in biomedical research. During this time, the approach to analysing gene expression has changed dramatically. Where once the analysis of differential gene expression was viewed in terms of a single gene, it is now possible to assess the expression of nearly all genes of a genome simultaneously.

The genome-wide landscape of expressed genes is generally termed the 'expression profile'. This profile is a complex, context-dependent pattern of expressed genes that is as typical for a certain cell or tissue in a defined biological state as a fingerprint of an individual finger tip. Therefore, these profiles are often referred to as RNA or protein fingerprints.

Expression analysis at the level of a single gene at a single time point is a relatively simple task. The message of a gene in the experimental situation is either upregulated, downregulated or unchanged compared to its expression in the control situation. By contrast, the analysis of the expression of a multitude of genes or the totality of genes in response to a stimulus is an immensely complex and difficult endeavour, in particular in the case of time- and/or dose-dependent experiments (**Figure 1**) [1].

Due to the increasing interest in profiling genome-wide expression in numerous samples, it is not surprising that high-throughput techniques have burgeoned during the past years [2]. However, the impressive range in complexity of methods that have been developed for the analysis of mRNA expression points to the importance of such analyses and demonstrates how technically demanding the analysis of mRNA expression profiling is. This review aims to provide an

Copyright © Ashley Publications Ltd to Printing and distribution. Strictly prohibited

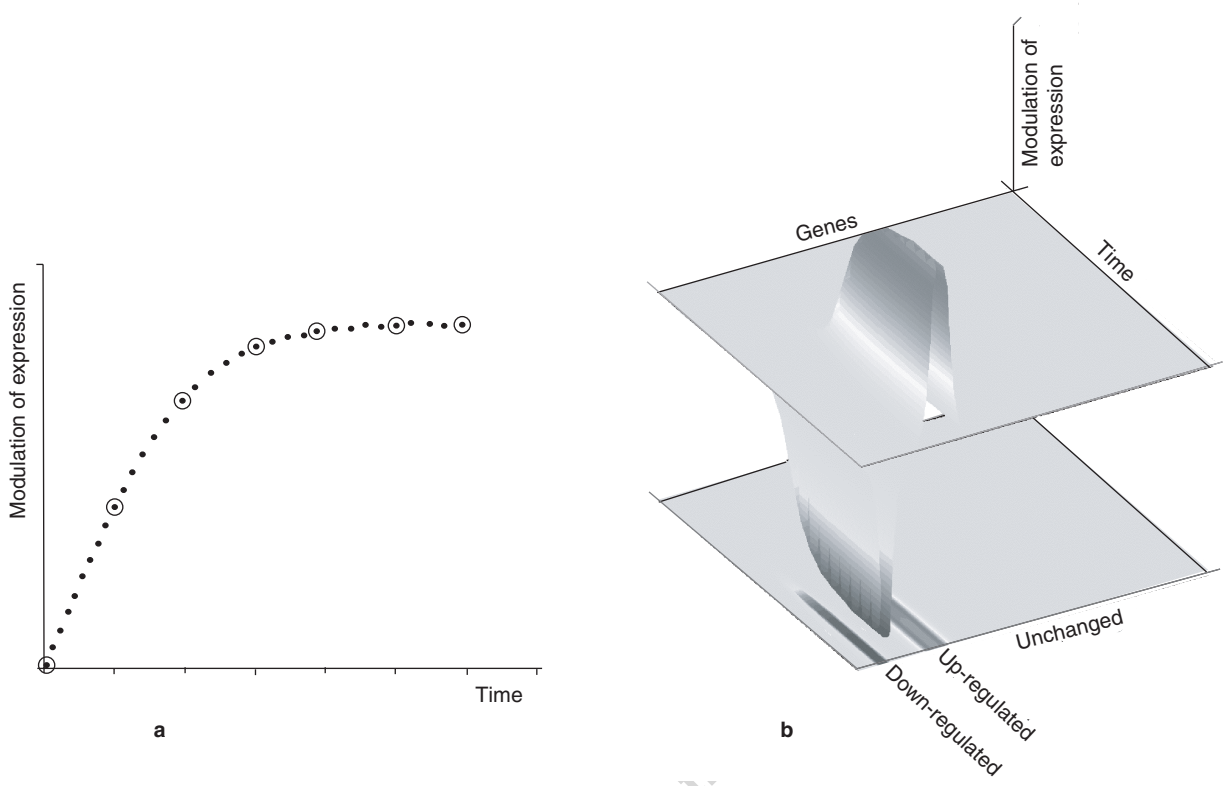

Figure 1. Graphical representation of examples of the relative complexity of single gene and genome-wide assessment of expression analysis. The increase in relative scale and complexity of data management becomes obvious if one compares the hypothetical schemes of the data generated by **a**) northern blot analysis or quantitative reverse transcriptase polymerase chain reaction, with a handful of data points, to the data output from **b**) a genome-wide expression analysis, for example generated by SAGE™ or DNA microarrays, with several thousand data points. Assuming that there are 30,000 – 40,000 transcripts present in a cell, as estimated for a typical human cell, a time course with five time points, a single experiment without replication, may be expected to yield at least 150,000 data points. The problems of dealing with this amount of data are obvious. In this figure, the grey-scaled surface represents the relative change in the expression of a certain gene at a given time point compared to the initial state (time point zero). The expression of genes that are upregulated or downregulated (up to a few hundred) is coloured in dark grey. The plane coloured in light grey shows cohorts of genes that are not modulated in their expression. These genes represent the majority of all measured genes (about several hundreds to several thousands). The figure is adapted from Green *et al.* and Lorkowski & Cullen [1,2].
SAGE: Serial analysis of gene expression.

introduction to a selected list of methods that have been established as high-throughput procedures allowing genome-wide profiling or simultaneous screening of huge amounts of samples for a selected number of genes. This review specifically explores alternatives to the use of DNA microarrays and is a companion paper to one specifically focusing on microarrays that the authors have recently published in this journal [3]. Readers seeking more detail on the techniques described here and on gene expression in general are also referred to a comprehensive recent work on this topic edited by the authors [2]. In addition, **Table 1** gives an overview of methods for profiling mRNA expression that are not described here.

The development of expression profiling has been paralleled by the generation of bioinformatics tools and algorithms to manage and analyse the huge amounts of data retrieved from these very complex experiments. Indeed it is only by use of these tools that extraction of useful physiological and biological information from gene expression studies is possible. For reasons of brevity, however, this review will not include detailed information on bioinformatics but it is of note that the methods presented below have different degrees of dependence on bioinformatics. For example, microplate-based assays require only statistical input because they have been developed with the aim to quantify the expression of known mRNAs. By contrast, methods such as GeneCalling® (CuraGen Corporation, New Haven, CT, USA) or massively parallel signature sequencing (MPSS™, Lynx Therapeutics, Hayward, CA, USA), which have been developed for genome-wide expression profiling display their full capabilities only if comprehensive and well-annotated information on the transcriptome or the genome of the organism of interest is available. At first glance, open system methods such as differential display or cDNA-amplified fragment length polymorphism (AFLP) require only limited

Copyright © Ashley Publications Ltd. All rights reserved. Unauthorized reproduction prohibited

Table 1. Overview of methods for mRNA expression profiling that are not described in this review.

Methods	Abbreviations	References
Hybridisation-based methods		
Barcode-based single cell analysis	–	[85]
Northern blot and related techniques	–	[86]
Nuclear run-on assay	–	[87,88]
Subtractive hybridisation		
Differentiation induction subtraction hybridisation	DISH	[89,90,222]
Differential subtraction chain	DSC	[91,92]
Differential subtraction display	DSD	[93,223,224]
Enzymatic degrading subtraction	EDS	[94,95]
Gene expression screen	GES	[96,97]
Linker capture subtraction	LCS	[98,99,225]
Rapid subtraction hybridisation	RaSH	[100,101,226]
Representational difference analysis	RDA	[102,103]
Reciprocal subtraction differential RNA display	RSDD	[104,105]
Selective amplification via biotin and restriction-mediated enrichment	SABRE	[106,107,158]
Selectively primed adaptive driver RDA	SPAD-RDA	[108,227]
Subtractive hybridisation of transcriptionally active genomic DNA fragments	–	[109,110]
Subtracted RNA fingerprinting	SuRF	[111,112]
Suppressive subtractive hybridisation	SSH	[113,114,228-230]
Multiplexed RNA analysis on fluorescent microbeads	–	[115,231,232]
Ribonuclease protection assay	RPA	[116,117]
Virtual northern blot	–	[118,119]
Polymerase chain reaction-based methods		
Amplification of double-stranded cDNA end restriction fragments	ADDER	[6,7]
Adapter-tagged competitive PCR	ATAC-PCR	[120,121,233-235]
Codon optimised differential display PCR	CODD-PCR	[122,236-238]
Gene expression fingerprinting	GEF	[123-125,239-241]
High coverage expression profiling	HiCEP	[126,242]
Introduced amplified fragment length polymorphism	iAFLP	[127,128]
Indexing-based differential display	IBDD	[129,130]
Molecular indexing	–	[131-133,243,244]
Multiplex PCR using colour-tagged module-shuffling primers	MSP-PCR	[134,135,245,246]
Nuclear expressed sequence tag analysis	NEST analysis	[136]
Ordered differential display	ODD	[4,5]
Preferential amplification of coding sequences	PACS	[137,138,247]
Quantitative (competitive) RT-PCR	–	[139-141]
RNA arbitrarily primed PCR fingerprinting	RAP-PCR	[142,143,248,249]
Restriction display polymerase chain reaction	RD-PCR	[144,250]
Restriction fragment differential display	RFDD	[145]
Restriction fragment length polymorphism-coupled domain-directed differential display	RC4D	[146,147]
Reverse transcription and subsequent amplification by PCR using sequence-independent primers	SIP RT-PCR	[148,149]

PCR: Polymerase chain reaction; RDA: Representational difference analysis; RT-PCR: Reverse transcriptase-polymerase chain reaction.

Copyright

Table 1. Overview of methods for mRNA expression profiling that are not described in this review *(continued)*.

Methods	Abbreviations	References
Polymerase chain reaction-based methods **(continued)**		
Targeted display	–	[150,151]
Sequencing-based techniques		
Comparative expressed sequence tags analysis/electronic subtraction	–	[15,152]
Open reading frame expressed sequence tags	ORESTES	[153,154,251]
Methods based on other principles		
Digital analysis of cDNA abundance	DADA	[155,252]
Restriction landmark cDNA scanning	RLCS	[156,253]
RNA patterns method	–	[157,254]

prior knowledge on expressed sequences of the genome of interest. But if only limited information is available these methods can be used only for identifying a limited number of differentially expressed genes with the aim to further characterise their physiological role in the context of the primary scientific query. However, at present, high-throughput expression profiling is aimed at investigating complete expression patterns of cells or tissues under different conditions to answer more global problems: instead of identifying single candidate genes, the identification of groups (clusters) of differentially expressed genes is in the spotlight. Gene expression profiling with the aim to investigate global changes in mRNA expression patterns, independent from the method being used, is only reasonable if a well-annotated expressed sequence database or at least genome databases and powerful bioinformatics tools such as normalisation procedures, clustering algorithms and data management tools can be used. Otherwise, these analyses are limited to the identification of new targets for further characterisation and do not answer the purpose of state-of-the-art genome-wide mRNA expression profiling.

2. Classification of methods for analysing mRNA expression

State-of-the-art methods for analysing gene expression can be divided into two systems [1], those with a closed and those with an open architecture. While closed systems only allow analysis of predefined known genes, open systems allow analysis of the expression of genes that were previously unknown. The quality of closed systems therefore depends on the quality of the prior knowledge. A complete analysis of the transcriptome of a genome using closed systems is strictly dependent upon the completeness of knowledge of that transcriptome, thus severely limiting comprehensive applicability in the near future, in all but the most well-characterised and commonly used species.

Although closed systems work well for certain types of analyses, as has been shown in several studies using conventional techniques such as northern blotting, reverse transcriptase-polymerase chain reaction (RT-PCR) or DNA microarrays, their usefulness is intrinsically limited by the fact that the conformation of expressed genes is not completely known. Even in species such as man where the whole genome is known, the transcriptome is thought to be far more complex than the number of genes would at first suggest because of the use of alternative promoters, alternative splicing, RNA editing, etc. Therefore, one must keep in mind that even closed systems with huge screening capacity are limited by the current knowledge of the transcriptome.

In contrast to closed architecture systems, open systems require no *a priori* knowledge of the transcriptome. Therefore, open systems can identify novel transcripts in the absence of information on the modulation of expression of such a transcript. At least in theory, it is possible to identify all alternatively spliced or otherwise modified RNAs, so that with an open architecture system the complexity of a given transcriptome does not affect the quality of expression analysis.

To unfold their power, open systems require the support of available data. For example, in the simplest case, it is only by comparing a transcriptome investigated with an open architecture approach with the pre-existing closed-architecture transcriptome that novel and modified transcripts are identified. Comparison of the expression modulation of a novel transcript with data of known genes may also provide a hint on the function of a novel transcript and thus help in the generation of testable hypotheses.

For technical reasons, the capacity of open systems is limited, so that in situations where comprehensive knowledge of a transcriptome exists, they cannot really compete with closed architecture systems in terms of time and labour intensity. Even an open system that theoretically permits analysis of a whole genome is limited by the sensitivity of the method in detecting rare transcripts and splice variants.

3. High-throughput methods

The latest procedures that are suitable for automated high-throughput expression profiling or that may, at least in principle, be adapted to this use are now described. The reader should

Copyright of Ashley Publications Ltd Printing and distribution st... prohibited

note that these methods are in no way technologically inferior to several that are not described, such as ordered differential display (ODD) [4,5] or amplification of double-stranded cDNA ends restriction fragments (ADDER) [6,7,158,159].

The following content has been divided into three parts. The first deals with techniques that exploit the ability of DNA to bind in a specific manner to itself (hybridisation); the second describes techniques based on PCRs using short or arbitrary oligonucleotides to simultaneously amplify large numbers of cDNA fragments; while the third summarises techniques based on DNA sequencing which generally operate by the incorporation of short sequence tags into the gene whose expression is to be quantified.

3.1 Methods relying on hybridisation

3.1.1 *Microarrays*

DNA microarrays exploit a variety of techniques in order to apply DNA molecules in perpendicular fashion to a solid surface. In technical terms, DNA arrays are a logical extension of the method, first described > 30 years ago by Dr Edward Southern at Cambridge University, of applying DNA to a treated cellulose surface [8,160-163]. In recent years, a number of methods have been developed that allow very large numbers of DNA oligonucleotides or cDNA fragments to be applied to surfaces in ordered two-dimensional arrays, thus allowing the massively parallel analysis of hybridisation events [2,3]. The starting gun for this field was sounded in 1995 when a group at Stanford University published the first paper describing the use of DNA microarrays for gene expression profiling [9,164,165]. Since then, the number of published scientific studies employing microarray technology has increased sharply. Coupled with the explosion of information arising from the sequencing of the human and other genomes, such arrays hold promise in a wide range of research, diagnostic and therapeutic contexts.

The principles of using and producing DNA microarrays have been described in much detail in the literature (e.g., [2,3]). However, it should be pointed out that DNA-microarray-based analysis is an emerging technique that still has many technical problems [10]. Chief among these are problems for reproducibility, probe sensitivity, non-linearity in signal detection [11], probe cross-hybridisation due to homologous cDNA sequences [12], technical difficulties of spotting/synthesising probes, closed system architecture, as well as data analysis and management [13]. Although some of these problems have been (partly) solved in the case of oligonucleotide microarrays, it is certainly a mistake to believe that DNA microarray data are as accurate and sensitive as that derived from northern blotting. Nevertheless, DNA microarrays are an important tool for expression profiling in a wide variety of research applications.

3.1.2 *Beads array for the detection of gene expression (BADGE)*

Only a few of the high-throughput methods for analysing gene expression are suitable for detection of moderate numbers of genes in thousands of samples with high speed and low cost although there is great demand for such a method for use in diagnostics and drug target screening. To address this need, Yang *et al.* developed the beads array for the detection of gene expression (BADGE) assay for gene expression analysis using microspheres and flow cytometry analysis [14]. In a typical BADGE assay, up to 100 genes can be monitored within 1 h in a single reaction.

The underlying principle of BADGE is similar to that of microarrays. In microarray-based analyses, oligonucleotide probes are mounted on a solid surface, whereas in BADGE, the oligonucleotide capture probes are coupled to colour-coded microspheres and the RNA samples are processed in the same way as is for microarray-based analyses: total RNA is extracted and the poly(A)$^+$ mRNA fraction is converted into double-stranded cDNA using an oligo(dT) primer plus a 5′-end T7 RNA polymerase promoter sequence for first-strand synthesis. Purified double-stranded cDNA is converted into labelled cRNA by *in vitro* transcription and biotinylated UTP. The purified cRNA is randomly fragmented by heat incubation and hybridised to capture–probe-coupled microspheres. Phycoerythrin-conjugated streptavidin is used to detect bound biotinylated targets. The signal of each target hybridised to its specific capture probe can be determined by the fluorescence intensity of phycoerythrin. By counting single microspheres, the corresponding number of targets bound to the microspheres can be determined.

An interesting variant of this method has been patented by Shi *et al.* [166]. Using this enzymatic-ligation-based screening method, oligonucleotide pairs, each consisting of specific sensor probes having a part that is complementary to the targeted polynucleotide and a part comprising a common adapter primer binding site, are used for each target polynucleotide molecule. The complementary part of the sensor probes is immediately adjacent on the targeted polynucleotide. After hybridisation of the probes to the target molecule, the sensor probes are ligated, amplified and biotinylated using adapter primers and hybridised to complementary detector oligonucleotides attached to microspheres. As described above, single beads are counted by determining the fluorescence intensity of streptavidin-coupled phycoerythrin bound to the biotinylated nucleic acid. This method can be used to identify single nucleotide polymorphisms as well as for expression profiling as mentioned above.

BADGE has the advantages of affordable cost, rapid speed and high flexibility with the potential for high-throughput analysis. It is suitable for applications such as diagnostic detection of disease genes from a large number of clinical samples and screening of characteristic marker genes from many biological systems. However, as a closed architecture system, it is limited to the analysis of known transcripts. In addition, the method is not suitable at present for genome-wide expression profiling.

3.1.3 *Oligonucleotide fingerprinting (OFP)*

Sequencing of expressed sequence tags (ESTs) or cDNA libraries [15] has been widely used for expression profiling due to the

Copyright © Ashley Publications Ltd. Printing and distribution strictly prohibited

representation of all expressed genes of a given source. These libraries provide direct access to splice variants and give valuable insight into the abundance levels of different mRNAs. Despite these advantages, it is necessary to screen libraries consisting of up to several million independent cDNA clones to ensure that very rare transcripts are represented [16].

One way of analysing such libraries is the classical comparative EST sequencing approach in which, in principle, cDNA clones from these libraries are sequenced using classical sequencing techniques and are counted to reveal an expression profile. However, the oligonucleotide fingerprinting (OFP) approach, which was conceptually developed more than a decade ago [16,17], is an elegant alternative means of analysing these libraries. In principle, the OFP method is based on the concept of sequencing by hybridisation where a complete target sequence is deduced from independent hybridisations with a set of all possible oligonucleotides of a given length. OFP differs from this concept in that only partial sequence information (a fingerprint) is generated. Through a series of up to 250 successive hybridisations of short, labelled oligonucleotides of known sequence, vectors of hybridisation signal intensities (oligonucleotide fingerprints) are generated for each clone of an arrayed cDNA library. According to their individual fingerprints, clones are grouped into clusters ideally representing one gene and indicating its relative level of expression. A representative clone of each cluster is selected for subsequent experiments, such as the generation of gene expression arrays. Thus, a highly normalised almost non-redundant clone set is created representing the expressed gene repertoire of the analysed biological source.

As a result of its multiple robotic steps and automated bioinformatics elements, the OFP technology can be employed in numerous large scale applications. Nevertheless, the current technology shows some inherent limitations which need to be overcome in order to further increase the degree of automation and, hence, the throughput [16]. In particular, bottlenecks, which are the high number of oligonucleotide probes needed, the confined array density and the restriction to serial hybridisations.

3.1.4 *Microplate-based mRNA quantification assays*

Microplate-based mRNA quantification assays such as the Quantikine™ mRNA assay from R&D systems (Minneapolis, MN, USA), the QuantiGene™ expression assay and the VERSANT™ RNA assay from Bayer Diagnostics (Fernwald, Germany) [18], the Invader® RNA assay developed and patented by Third Wave Technologies (Madison, WI, USA) [19-21,167,168], the Xpress-Screen® mRNA detection assay from Applied Biosystems (Weiterstadt, Germany) [169] and the Array-Plate™ multiplexed molecular profiling system developed and patented by High Throughput Genomics (Tucson, AZ, USA) [170,171] are hybridisation-based systems that offer notable advantages, including ease of use and sample preparation, as well as high-throughput and fast processing capabilities. Microplate-based assays can be completed in a few hours, allow the direct use of cell lysate, total RNA or poly(A)$^+$ mRNA samples and are performed in a 96-well microplate format. Furthermore, these assays do not make use of radioactivity or target amplification and are read using conventional detection systems.

In the following, the Quantikine mRNA assay is exemplarily described. The Quantikine mRNA assay first involves hybridising of samples with gene-specific biotin-labelled capture probes and digoxigenin-labelled detection probes in a microplate well. The hybridisation solution is then transferred to a streptavidin-coated microplate where the biotinylated mRNA/probe hybrid is captured by binding the streptavidin. The captured mRNA/probe hybrid is incubated with an antidigoxigenin, alkaline phosphatase-conjugated antibody and, subsequently, a substrate solution is added, followed by an amplifier solution, to allow colour development. The resulting intensity of the colour is proportional to the amount of gene-specific mRNA in the sample. Colour development is stopped and the intensity of the colour is measured spectrophotometrically.

One advantage of the microplate-based mRNA quantification technologies is the utilisation of signal amplification instead of target amplification. The Quantikine mRNA assay uses a chemical reaction that amplifies both the signal generated from the alkaline phosphatase-conjugated detection antibody and the colorimetric product. The Xpress-Screen kit also makes use of an alkaline phosphatase-conjugated antibody but the antibody recognises RNA/DNA hybrids and allows luminescent signal detection [169]. Bayer Diagnostics' kits utilise branched DNA, in conjunction with alkaline phosphatase [18] and a luminescence substrate for signal amplification [22]. The Invader RNA assay is based on the cleavage of an oligonucleotide fluorescence resonance energy transfer probe [20,167]. Once the probe is cleaved, fluorophore and quencher are separated and a fluorescent signal is generated. The Array-Plate technology is based on ribonuclease protection assay principles and uses peroxide-conjugated probes and luminescence detection [23,170]. One drawback of many of these assays is the requirement for specialised fluorescent or luminescent detection equipment, whereas the Quantikine mRNA assay can be read on a conventional microplate reader.

Another feature of the microplate-based mRNA assay platform is its suitability to high-throughput screening and the ability to directly use cell lysates. The Array-Plate assay stands out among the rest of the mRNA detection assays by being able to measure up to 16 different mRNA species in a single well [23]. Invader RNA assays can measure two targets per well by using two different fluorescence probes. In contrast, the Quantikine mRNA assay, the branched DNA assays and the Xpress-Screen kit can measure multiple targets on the same plate but not within the same well.

The sensitivity of microplate-based mRNA quantification assays is sufficient for most purposes. For example, the Quantikine mRNA kit has a sensitivity of 0.1 amol, which is equivalent to 60,000 copies of mRNA. Similarly, the Array-Plate

Copyright © Ashley Publications Ltd Printing and distribution strictly prohibited

multiplexed molecular profiling assays detect 0.25 amol or 150,000 copies of mRNA. The assays using branched DNA are able to detect 0.005 amol or 3,000 copies of mRNA [24], whereas the Invader RNA assay and the Xpress-Screen assay detect 0.01 amol or 6000 copies of mRNA [20].

When compared to other methods, microplate-based mRNA quantification assays offer some clear advantages by obviating the use of radioactivity, polyacrylamide gels, electrophoresis and RT-PCR. However, from the perspective of ease of use and low entry cost, microplate-based mRNA quantification assays are an interesting alternative for mRNA quantification in a high-throughput fashion.

3.2 Procedures based on polymerase chain reaction

3.2.1 Automated fluorescent mRNA differential display (FDD)

RT-PCR-based differential display was developed by Liang & Pardee in 1992 [25,172] and is one of the major tools for analysing gene expression. Differential display integrates three of the most commonly used molecular biological methods, namely RT-PCR, DNA sequencing gel electrophoresis and cDNA cloning. In contrast to most of the other PCR-based methods currently available, no second strand cDNA synthesis, purification of cDNA, restriction enzyme digestion, adapter primer ligation, probe labelling and normalisation, hybridisation or washing steps are required. This increases the power of the method, as each of these steps has the potential to introduce errors.

In differential display [25,26], mRNAs are converted to single-stranded cDNAs using a set of three individual anchored oligo(dT)V primers or a set of nine individual anchored oligo(dT)VN primers. The use of different anchored primers enables the division of a complex cDNA pool representing the whole set of mRNAs of a cell into three or nine, less complex samples. The resulting subpopulations are further amplified and labelled with either isotopes or fluorescent dyes by PCR in the presence of a second set of upstream primers that are short and arbitrary in sequence and the corresponding downstream primer which was used for cDNA synthesis. The length of an arbitrary primer is designed so that by probability each primer will recognise 50 – 100 mRNAs under a given PCR condition [27]. As a result, 3′-termini of almost all mRNAs can be amplified and displayed by denaturing gel electrophoresis. Side-by-side comparisons of such cDNA patterns between, or among, relevant RNA samples reveal differences in gene expression. Differentially expressed cDNA bands are excised, cloned and sequenced for further molecular characterisation.

A genome-wide comprehensive differential display screening is possible but requires hundreds of PCR reactions for each RNA sample being compared. This makes a high-throughput and high-precision differential display platform, which incorporates robotic liquid handling alongside digital data acquisition and analysis, highly desirable. To achieve this, fluorescent mRNA differential display (FDD) was developed [28] and combined with robotics and digital data analysis [29]. Data analysis tools such as spectra overlay, which allow digital data presentation and quantification, have been developed by Hitachi Genetics Systems (San Francisco, CA, USA), for use with the FMBIO® (MaraiBio) series of fluorescent scanners (Hitachi Genetics Systems). Further streamlining the FDD process, any differentially expressed cDNAs of interest could be reamplified and sequenced directly with a given arbitrary primer without subcloning [30]. Furthermore, computer programs were developed to automatically allow positive band identification from an FDD image [31-33]. The most sophisticated attempt in FDD downstream automation could have been the development of a prototype computer-controlled system for positive band identification and retrieval [34]. This approach employed capillary array gel electrophoresis coupled with fraction collection using sheath flow technology. FDD is currently offered by GenHunter (Nashville, TN, USA) as a commercial service. Several other companies offer differential display kits for laboratory use.

Differential display has many advantages over other methods. It is an open system that requires no prior knowledge of the genes to be analysed. As such, differential display can discover novel genes and is readily applicable to any biological system. Due to PCR amplification of cDNAs, differential display achieves high sensitivity. Furthermore, differential display can compare more than two RNA samples side-by-side without the need for data normalisation. Differential display has been used in several thousand applications for expression profiling and through further refinement and automation, will undoubtedly continue to play a key role in gene discovery research in the future.

3.2.2 Amplified differential gene expression (ADGE)

Amplified differential gene expression (ADGE) was developed to magnify the magnitude of the expression difference of a gene between two samples before displaying them and is, in principle, a PCR-enhanced subtractive hybridisation method. The method was described for the first time in 2001 [35], the patent was published in 2002 [173]. ADGE seeks to improve the accuracy and sensitivity of gene expression profiling.

In ADGE, two different RNA samples are selected, one as control, the other as tester. After cDNA synthesis, both cDNAs are cut with a restriction enzyme. This generates three types of fragments depending on the number of restriction sites on any particular gene: i) a type A fragment with a restriction site at one end; ii) a type B fragment with restriction sites on both ends and iii) a type C fragment without a restriction site at either end. The restricted control and tester DNAs are ligated to a CT adapter oligonucleotide and a TT adapter oligonucleotide, respectively. The adapter-linked control and tester DNA samples are mixed in equal amounts, as in the usual method for subtractive hybridisation-based methods, denatured and annealed. After filling in the ends, the reassociated DNA is amplified using CT primers complementary to the CT adapter and TT primers complementary

Copyright © Ashley Publications Ltd Printing and distribution strictly prohibited

to the TT adapter. The PCR products are separated on gels or hybridised onto DNA microarrays. The bands on gels displaying differences between the two samples are isolated and the corresponding genes are identified by sequencing.

The integration of DNA reassociation and PCR amplification make it possible to obtain quadratic magnification of the ratio of an mRNA between two samples. After DNA reassociation, three different duplexes of type B fragments are formed: i) control DNA with the CT adapters on both ends; ii) tester DNA with the TT adapters on both ends and iii) hybrid DNA with the CT adapter on one end and the TT adapter on the other end. Although the ratio of control DNA and tester DNA has been magnified quadratically after DNA reassociation, control and tester DNA have not yet been separated from each other or from hybrid DNA. This is achieved by PCR. The CT primers amplify control DNA exponentially and hybrid DNA linearly, as control DNA has the CT adapter on both ends while hybrid DNA contains the CT adapter on one end only. Equally, the TT primers amplify tester DNA exponentially and hybrid DNA linearly. After PCR, the exponentially amplified control or tester DNA is a million times more abundant than the linearly amplified hybrid DNA. Thus, the PCR step both represses the hybrid DNA and separates control DNA and tester DNA from each other. The ratio of type A control DNA to type A tester DNA does not change after DNA reassociation. Type A DNA is also linearly amplified and thus repressed by PCR.

The major advantage of ADGE is that it quadratically magnifies the ratio of a gene between two samples before displaying them. Thus, ADGE raises the magnitude of the expression ratios above the threshold of detection error before displaying them and identifies smaller differences in gene expression that might elude detection.

3.2.3 cDNA-based amplified-fragment length polymorphism fingerprinting (cDNA–AFLP)

Amplified fragment length polymorphism (AFLP) was developed as a technique for the visualisation of genomic DNA polymorphisms [36]. It was initially used to construct genetic linkage maps in segregating populations and to identify individuals with unknown genotypes. In AFLP, restriction enzyme digestion using two enzymes differing in the frequency of their recognition sites within a given genome is used to fragment the DNA sample, after which adapters are ligated to the resulting sticky ends. Primers corresponding to the adapters with one or more additional nucleotides extending beyond the restriction recognition site and into the target fragments are used to amplify the DNA and to create a sample-specific fingerprint. In 1996, the application of AFLP to cDNA with the aim of visualising mRNA expression patterns was published [37-39]. This application has been widely used for mRNA expression profiling and is suitable for use in a standard laboratory.

In cDNA–AFLP, mRNA is isolated using biotin-labelled oligo(dT)V oligonucleotides attached to streptavidin beads from DNase-I-treated total RNA [40]. Complementary DNA is synthesised using the oligo(dT)V primer attached to the microbead. After synthesis of double-stranded cDNA, the sample is processed in the same manner as DNA is processed in genomic AFLP to prepare a template that is in turn amplified to generate an RNA fingerprint. After synthesis, the cDNA attached to the bead is digested enzymatically and anchor adapters are ligated. The 3′-ends of the cDNA bound to the streptavidin beads are removed and the 5′-ends are discarded for amplification and separation by polyacrylamide gel electrophoresis. The intensity of individual signals in the fingerprint is taken as a measure of the strength of expression of the mRNA corresponding to each band. Similar to differential display RT-PCR, differentially expressed genes can be identified by comparing band patterns of different samples, lane by lane. Interestingly, transcript-derived fragments can be isolated from gels after auto-radiographic detection or silver staining by excising the relevant section from the polyacrylamide gel and by reamplifying the extracted DNA using the same primers. After reamplification, DNA fragments can be sequenced directly or cloned into a vector before sequencing.

Due to the complexity of eukaryotic genomes, it is impossible in practice to visualise all transcript-derived fragments on a single gel. To achieve a selective reduction, two strategies are implemented in cDNA–AFLP. First, the restriction enzyme used to digest the cDNA limits the number of transcripts that are visualised. Second, by using different lengths of so-called 'selective bases' on the primer termini, tuning of the number of targeted fragments per amplification is possible.

In the first digestion step, the cDNA is digested with a rare cutting enzyme which is usually an enzyme with a six-nucleotide recognition sequence. The second enzyme is used to generate fragments of the desired size usually using a restriction enzyme recognising four nucleotides. Due to the differential frequency of restriction sites of the two enzymes in cDNA, most of the fragments carry sticky ends for the frequent cutting enzyme only. These fragments are not visualised, as only the primer binding to the adaptor ligated to the rare cutting enzyme site is labelled in the final PCR, thereby revealing only fragments carrying either both sites or only sites of the rare cutter.

The selection of enzymes is a crucial point in cDNA–AFLP as it will affect the number of different sequences represented in the fingerprint. Ideally, the rare cutter should cut every cDNA only once but, for example, in potato, the highest theoretically achievable digestion frequency is ~ 50% [38]. For a variety of reasons, the digestion frequency that can be achieved experimentally is < 50%. Several strategies have been considered to increase the success rate of the rare cutter enzyme without generating a large number of fragments per cDNA. One is to use multiple rare cutting enzymes generating sticky ends compatible with the same anchor or to use enzymes with a recognition sequence including one or more wobble bases and to then use degenerate primers in the amplification. In addition, large differences in the restriction frequency of different transcripts encoded in the genomes of different organisms occur.

Copyright © Ashley Publications Ltd Printing and distribution strictly prohibited

The effectiveness of the enzymes therefore needs to be determined in order to optimise the results.

By incorporating a restriction enzyme site in the primer used for cDNA synthesis and initial digestion with the frequent cutting enzyme, all 5′-fragments of the cDNAs can be discarded leaving a single fragment for every cDNA attached to the beads. This fragment can be released from the beads by digestion with a second (rare cutting) enzyme. It should be noted, however, that a 3′-bias is established in this variation of the method. Due to the low level of conservation in such sequences between homologues, using this procedure in a species with little sequence data available in the databases makes it difficult to identify transcripts directly by sequence similarity.

The cDNA–AFLP method is a robust and reliable method for monitoring the differential expression of genes in a wide range of biological systems as it is able to detect gene expression of very rare messages. Along with other RNA-fingerprinting methods, it has the great advantage of not requiring prior sequence information and allows the screening of a large number of controls simultaneously with the samples of interest. Due to the highly stringent conditions used during PCR, the rate of false-positives produced by mis-priming is low. However, as with all analysis involving DNA retrieval from denaturing polyacrylamide gels, false-positives are generated by co-amplifying unrelated sequences that may either underlie the targeted band or co-migrate very closely with the chosen fragment.

3.2.4 *Digital expression pattern display (DEPD®)*

Researchers at Biofrontera Pharmaceuticals developed and patented an automated PCR-based procedure called digital expression pattern display (DEPD)® [41,42,174,175]. In principle, DEPD is a modification of the differential display RT-PCR method. In contrast to differential display, DEPD makes use of more stringent PCR conditions by introducing adapter molecules to enzymatically restricted cDNA samples.

In DEPD, double-stranded cDNA is synthesised following extraction of total RNA. The cDNA is subdivided into three groups to achieve high resolution. Each cDNA pool is digested with a different kind of class IIs restriction enzyme. This step produces short DNA fragments containing single-stranded overhangs in all possible sequence combinations. Sixteen different specific DNA adapter oligonucleotides are then ligated to the restricted cDNAs to allow specific PCR-based amplification of the fragments. Subsequently, 1,024 PCR reactions are performed for each of the three cDNA pools using 1,024 different 5′-primers. Improved results are obtained if the 3′-poly(A) sequences derived from the mRNAs are removed during the process and are replaced with a mixed-sequence adapter. Because PCR primers are complementary to the adapters, it is possible to perform the PCR reactions at high temperatures, resulting in high reproducibility and reliability. This depicts a crucial difference in comparison to the original differential display RT-PCR method invented by Liang & Pardee in 1992, in which the use of randomised PCR primers requires lower temperatures, probably leading to a high error rate. Amplified DNA fragments of each PCR are separated by electrophoresis using 96 capillary electrophoresis sequencers, allowing the separation of up to 300,000 distinct DNA fragments. Each of the distinct fragments that are produced by DEPD represents an EST, which can be identified by sequencing and reference to Biofrontera's proprietary database, Brain-Sort®, containing the sequence information of all transcripts already analysed in Biofrontera's prior experiments.

Isolation and sequence analysis of all fragments obtained in DEPD experiments showed that DEPD produces a redundancy of ~ 2.5 [41]. Therefore, 100,000 identified fragments correspond to 40,000 different transcripts in the cell. This approximates to the probable number of genes in the human genome.

DEPD is a massively high-throughput method for the analysis of gene expression that has been transformed into a fully automated, industrial scale process by Biofrontera Pharmaceuticals. As an open architecture system, DEPD allows both detection of known and unknown transcripts.

3.2.5 *GeneCalling®*

GeneCalling permits high-throughput reproducible detection of mRNAs with a sensitivity of > 1 in 100,000 transcripts. It allows the discovery of known and novel differentially expressed genes from any species [43,44,176,177].

The GeneCalling procedure is comprised of three basic steps: i) restriction endonuclease digestion of cDNA; ii) adapter ligation and iii) PCR amplification. Following double-stranded cDNA synthesis from poly(A^+) mRNA, cDNA pools are digested using different pairs of restriction enzymes with 6 bp recognition sites. Complementary adapters are ligated to the digested cDNA and adapter-specific primers are used for PCR amplification of fragments containing sites for the pairs of restriction enzymes used. One adapter-specific primer is biotin-labelled while the other is fluorescently labelled. Following amplification, the biotin-labelled DNA is purified on immobilised streptavidin. Denatured single-stranded DNA fragments are resolved by electrophoresis on ultra-thin polyacrylamide gels or capillaries and labelled fragments are detected upon laser excitation. Since the biotin label is necessary for purification and the fluorescence label is necessary for detection, all detected fragments result from restriction digestion with both enzymes. cDNA fragments representing differentially expressed genes can be identified by database searching with the 6 bp restriction enzyme recognition sequences at the fragment ends and the exact length of each fragment.

A method for rapid confirmation of the identity of DNA fragments determined by database searching was also developed as part of the GeneCalling procedure. The reaction revealing the fragment of interest is performed a second time using the same end primers but in the absence and presence of an excess of an unlabelled oligonucleotide whose sequence is derived from the predicted gene fragment. If the identity of the fragment was predicted correctly, the unlabelled oligonucleotide will outcompete the universal oligonucleotide for priming that fragment and appear in the chromatogram to ablate

Copyright © Ashley Publications Ltd Printing and distribution strictly prohibited

that peak specifically without affecting the amplification of the other fragments [43]. This method also allows more precise estimation of the magnitude of gene expression differences.

GeneCalling has been applied to a wide variety of both animal and plant systems for trait identification, marker and pathway identification and has utility both in the drug discovery arena as well as in the study of drug response. The advantages of GeneCalling include the flexibility of an open architecture system, the capturing of the transcript's centre, which provides protein-coding information, and the ability to i) apply this technology to any organism; ii) sensitively distinguish rare and abundant transcripts; iii) independently measure transcript abundance multiple times in a single experiment and iv) comprehensively measure the majority of transcripts in a cell. These characteristics make GeneCalling an attractive system for drug discovery and other applications.

3.2.6 *Ligation specificity-based expression analysis display (LEAD)*

The open architecture mRNA profiling technology ligation specificity-based expression analysis display (LEAD) was developed by Li *et al.* in 2002 [45] to improve specificity of the cDNA fractionation step of expression profiling methods using enzymatic digestion of double-stranded cDNA such as restriction enzyme analysis of differential sequences (READS™, Gene Logic, La Jolla, CA, USA, see below), ADGE or total gene expression analysis (TOGA™, Digital Gene Technologies, see below). Several of these methods use sequence variation flanking restriction enzyme sites in different cDNA fragments to distinguish them. By contrast, the LEAD method utilises restriction enzymes with N(m) degeneracy in their recognition/cleavage sequences to fractionate cDNA populations (N represents any of the four bases A, T, G or C; while m represents the number of degenerate bases).

In LEAD, total RNA is used to synthesise double-stranded cDNA using oligo(dT) primers for reverse transcription. The double-stranded cDNA is digested by *Bsa*JI (recognition site: C/CNNGG) in combination with ten different 6-base cutter enzymes. Restricted cDNAs are selectively ligated to specific double-stranded adapters, amplified and fluorescently labelled by PCR. Fingerprints of the resulting cDNA fragments are obtained by gel electrophoresis-based separation and quantified electronically by using a specially designed computer program.

Transcripts of low abundance (1 of 100,000 copies) are detectable using LEAD, allowing the status of nearly all mRNAs to be monitored. Because of its sequence independence, LEAD can be used to monitor gene expression in fully sequenced organisms and in systems lacking whole genome information.

3.2.7 *Restriction enzyme analysis of differential sequences (READS™)*

READS is a PCR-based method that was first described by Prashar & Weissman in 1996 [46-48,178,179]. Similar to the differential display RT-PCR approach, this method uses PCR amplification but this is performed under more stringent conditions using more specific adapter primers [47,48].

In the first step of READS, DNA-free total RNA is isolated from samples of interest. Different subsets of the RNA are independently reverse-transcribed into cDNA using different oligo(dT)VN primers with a 20 bp heel (for details, see [48]). Following second strand synthesis, double-stranded cDNA is digested using restriction enzymes to form 3′-cDNA fragments with a single-stranded GATC 5′-overhang. For the selective subsequent PCR amplification of the 3′-cDNA fragments, a Y-shaped double-stranded adapter, called a fly adapter, with a single-stranded GATC 5′-overhang at one site is used (for details, see [48]). After digestion of double-stranded cDNA, the adapter is annealed and ligated to the digested cDNA fragments to form two different kinds of cDNA fragments: 5 ′-fragments with the Y-shaped fly adapter at each of the two ends and 3′-fragments with the Y-shaped fly adapter at the 5′-end and the modified oligo(dT)VN primer sequence at the 3′-end. In the next step, the 3′-cDNA fragments are selectively amplified by PCR using 3′-primers that are identical either to the primers used for cDNA synthesis or the heel of these primers; while an oligonucleotide that is a part of the non-complementary stretch of the Y-shaped fly adapter serves as the 5′-primer. During this PCR, only those 3′-end cDNA fragments with the Y-shaped adapter ligated to their 5′-ends and the heel sequence on the 3′-ends are amplified. The adapter primers may be radioactively labelled, allowing detection of PCR products by autoradiography. The PCR fragments derived from different samples are separated on polyacrylamide gels and cDNA bands showing different intensities in two samples are extracted from the gel, reamplified, directly sequenced, or subcloned and sequenced as performed in the classic differential display approach [25,27,48,49].

Using different kinds of restriction endonucleases (12 – 24 6-base cutting restriction enzymes) each being used with several different heeled oligo(dT)VN primers, it is possible to systematically analyse the complete mRNA pool of a eukaryotic cells as a pattern of bands derived from gel electrophoresis [47,48]. As with differential display RT-PCR, no prior knowledge of the transcript sequences is needed.

READS is in principle a modification of the differential display RT-PCR and the refinements make it a sensitive tool leading to reproducible results. It commonly allows detection of relative changes in the level of mRNA expression of the order 1.5 to 2-fold at the lower limit. At the upper end, relative differences in abundance of three to four orders of magnitude can be measured [47,48]. The most important advantage of the READS procedure is that it is an open system, thus allowing the identification of known and unknown genes.

In summary, READS is an elegant strategy to identify differentially expressed mRNAs. It bears the same risks that have been proposed for differential display RT-PCR and related methods such as DEPD. READS technology is provided as a commercial high-throughput service for the analysis of gene expression.

Copyright © Ashley Publications Ltd Printing and distribution strictly prohibited

3.2.8 *Total gene expression analysis (TOGA™)*

The TOGA method was first developed as a manual technology. Digital Gene Technologies subsequently automated the process and built the informatics, which accomplish the integration of TOGA as a high-throughput method and offers this patented technology as a commercial service [50-52,180,181]. The main advantages of TOGA are very low amounts of starting material and its immense high-throughput capacities.

TOGA begins with a few nanograms of total RNA. Synthesis of cDNA is performed using a pool of primers that anneal to the very beginning of the poly(A) tail on the mRNAs. After cDNA synthesis, the cDNAs are cleaved with a restriction endonuclease that uses a 4 bp recognition sequence. The resulting 3′-fragments are isolated and ligated to a short primer-binding sequence at the 5′-end, thus permitting use of this PCR-product in subsequent PCR steps. The important step is the last one, in which the binding site primer is extended by four nucleotides into the cDNA inserts. Since each position may contain A, C, G or T, 256 (4^4) primers are required to cover all possible sequence combinations These 256 reactions are performed with fluorescently-labelled 3′ in order to generate fluorescent PCR products.

To identify expressed sequences, a digital address is created for each transcript identified. The TOGA addressing begins with the four nucleotides that are recognised by a restriction endonuclease which cleaves the RNAs near the 3′ end. The cleaved molecules are sorted into 256 pools by virtue of the 4 nucleotides adjacent to the cleavage site. That provides eight nucleotides of sequence. Since the cDNA was originally cleaved at a discrete site upstream of the poly(A) tail, when the length is included, an address for each RNA is obtained. The identity or address, of the eight nucleotides plus the length of the fragment is sufficient information to recognise an RNA if its sequence has already been deposited in a database. These sequence tags are called digital sequence tags.

The whole TOGA process is repeated with a total of four different restriction enzymes, so that > 98% of all transcripts are sampled with at least one of the enzymes, most of them with three or four. The products are separated by capillary electrophoresis and each product fragment is detected by fluorescence; its length and intensity are recorded. The TOGA analysis is accomplished through the TOGA Portal, a web-based application that integrates gene expression data with information from genomics databases and provides a series of bioinformatics tools within a single graphical user interface.

The data generated by TOGA are highly reproducible and are stored in a permanent record. Using the database, RNA concentrations in different samples that might have been generated in completely unrelated experiments can be compared. The detection level of capillary electrophoresis systems is below one part in a million, which is between one-tenth and three-tenths of a copy of RNA per cell. TOGA requires only ~ 20 ng of total RNA. About 6,000 samples can be processed per robot per year with very short turnaround time. In addition, the method recognises both known and novel RNAs and can be used to examine RNAs in any species [50]. Thus, TOGA is an elegant and interesting high-throughput method that fits several research applications.

3.3 Sequencing-based methods

3.3.1 *Massively parallel signature sequencing (MPSS™)*

MPSS is an open-ended, open architecture platform that analyses the level of expression of virtually all transcripts in a sample by simultaneously counting a large number of individual mRNA molecules in a sample. The steady improvement of the MPSS procedure is reflected by the numerous patents describing enhancements and modifications of the methods comprising the MPSS technology [53,54,182-205]. Using MPSS, individual mRNAs are identified through the generation of a 17 base signature sequence at a unique site on the molecule.

In the first step of MPSS, the Megaclone™ technology developed at Lynx Therapeutics, is used to clone millions of individual double-stranded cDNA molecules onto 5 µm microbeads [55,189-191,195,196,201]. For this purpose, cDNA is prepared from poly(A)$^+$ mRNA using a biotin-labelled oligo(dT) primer. Double-stranded cDNAs are digested and the purified 3′-most poly(A) cDNA fragments are tagged with 1 of 16.8 million different 32-base synthetic oligonucleotide tags. After amplification of the tagged cDNA molecules, the 32-base tag at the 3′-end of each molecule is made single-stranded with T4 DNA polymerase and the tags on the cDNA fragments are hybridised to a set of 16.8 million different microbeads, each containing a covalently attached 32-base oligonucleotide (anti-tag) that is complementary to one of the 32-base oligonucleotides used to tag the molecules. On average, 100,000 amplified copies of each cDNA will hybridise to a single bead. The nick between the hybridised tag and the complementary tag on the bead is repaired to create a covalent linkage between the cDNA molecules and the bead. The result of the Megaclone process yields a library of microbeads where each molecule of mRNA in the starting sample is represented by one microbead.

The microbeads loaded with cDNA fragments are used directly for MPSS sequencing [53]. Since each cDNA molecule is attached at its poly(A) end, the 17-base signature sequencing reactions occur at the restricted end of each molecule. Approximately one million beads are packed into a specially designed flow-cell in a way that allows them to form a tight monolayer by stacking along channels in the flow-cell. Each microbead is stationary in the flow-cell and is monitored with a camera that is capable of detecting a fluorescent signal on each of the microbeads. Fluorescent signals arising from each microbead are monitored during the sequencing reactions in order to determine the sequence from the end of each cDNA molecule attached to the beads.

To initiate sequencing, a 4-base single-stranded overhang is produced at the end of each cDNA molecule. This is done by first digesting the DNA on the microbeads with *Dpn*II and then ligating an initiating adapter to the resulting sticky end. The initiating adapter contains a type IIs restriction

Copyright © Ashley Publications Ltd. Printing and distribution strictly prohibited

recognition site. Digestion of the sample with a type IIs restriction enzyme produces a 4-base single-stranded overhang at a position 9 – 13 nucleotides away from the recognition sequence. This generates initiated molecules with a 4-base single-stranded overhang immediately adjacent to the *Dpn*II recognition site. Two different initiating adapters, called 2-stepper and 4-stepper, are used for initiating the sequencing reaction in order to stagger the first four bases in the single-stranded overhang. This is done to address the issue of palindromes that arise in the 4-base overhangs during the sequencing reactions.

Once the molecules are initiated, the sequencing reactions are performed directly in the flow-cell. The system generates a sequence in four nucleotide increments through the hybridisation of a synthetic adapter to the 4-base single-stranded overhang on each initiated cDNA. Each adapter is part of a set of molecules that each contain a defined 4-base single-stranded overhang with 1 of all 256 different possible combinations of 4 nucleotides. Each molecule is also encoded in a manner that allows the sequence of its 4-base overhang to be decoded through a series of hybridisation reactions in the flow-cell. Therefore, the sequence of the first four nucleotides on the cDNA is generated by simply deciphering which member of the adapter set has hybridised to the cDNA on each bead. Each synthetic adapter also contains a type IIs restriction recognition site, so additional sequence beyond the first four nucleotides is produced by repeating the process after removing the first synthetic adapter with a type IIs restriction enzyme. The recognition site for the enzyme is positioned in the adapter in a way that produces a new single-stranded overhang that is immediately adjacent to that from the previous round of analysis.

The first step of the process involves the addition of a set of encoded adapters to the flow-cell. The members of the encoded adapter set each contain two important features: the recognition sequence, which is a defined 4-base single-stranded sequence at one end of each molecule and the encoder sequence, which is a single-stranded sequence at the other end of the molecule. The 4-base recognition sequence on an encoded adapter will hybridise to a complementary sequence on the single-stranded end of the initiated cDNA on the beads and the encoded adapter is ligated to the ends of the molecules. The next step involves determining which encoded adapter has become attached to each cDNA molecule. This is accomplished through a series of hybridisation reactions with 16 different fluorescent-labelled decoder probes, where each probe hybridises only to the encoder sequence on specific members of the encoded adapter set. The recognition sequence for the encoded adapters on each microbead is then deciphered using the encoded sequence as a guide. To collect additional sequence, the encoded adapter from the first round is removed by digestion with a type IIs restriction enzyme to produce a 4-base single-stranded overhang that is immediately adjacent to first four nucleotides from the first round of sequencing. This is possible because the encoded adapters contain a synthetic type IIs restriction recognition site at a defined position within the molecule. The process is then repeated several times in order to generate a 17-base signature sequence for the cDNA on the bead in the flow-cell.

The encoded adapter set of 1,024 different molecules is comprised of 4 groups, each with 256 different members [53]. Each group contains all possible combinations of nucleotides in the recognition sequence, so there will be an encoded adapter for any possible 4-base overhang on the initiated cDNA molecules on the beads. Sixteen different encoder sequences are used to build the encoded adapter set and each is used to identify a specific nucleotide at one of the positions on the 4-base recognition sequence.

This approach makes it possible to decipher 256 different sequence combinations on the 4-base overhangs on the cDNA molecules using only 16 different fluorescent-labelled decoder probe hybridisation reactions. It is important to note that four different encoded adapters hybridise with the cDNA on each microbead. These different members of the encoded adapter set contain the same recognition sequence but each has a different decoder sequence that allows the nucleotide at each position to be deciphered.

Using multiple flow-cells loaded with beads derived from the same mRNA sample, a typical MPSS experiment produces data sets that contain 17-base signature sequences from more than one million counted molecules per sample. This provides greater than a threefold oversampling per cell, thus giving potentially the capability to detect a single mRNA per cell. Additional sensitivity can be achieved by sequencing more beads of the same sample.

Each signature sequence in an MPSS data set is identified, compared to other signature sequences and counted. The data for each gene is normalised and expressed as transcripts per million. Analysis of a complete MPSS data set makes it possible to calculate the numbers of genes that are expressed at various levels within the sample. Signature sequences can be connected to known genes by comparison with databases for the organism under study.

Compared to several existing technologies, MPSS provides in-depth quantification of virtually all transcripts in a sample. Due to the open architecture design of MPSS, it is possible to generate gene expression data sets from any organism. Additionally, MPSS has the sensitivity to quantify genes that are expressed at very low levels. The digital nature of MPSS data sets makes it easy to be shared and integrated into databases. These features make MPSS an ideal technology for generating data content for a wide variety of applications.

3.3.2 *Serial analysis of gene expression (SAGE)*™

Serial analysis of gene expression (SAGE)™ (Invitrogen, Karlsruhe, Germany) is a high-throughput technique based on cloning and sequencing of short tags that was designed by Velculescu *et al.* in 1995 [56] to gain a detailed quantification of gene expression in eukaryotic cells in the absence of

Copyright © Ashley Publications Ltd Printing and distribution strictly prohibited

sequence information [57,58,206-213]. SAGE is based on the isolation of unique DNA sequence tags derived from a defined position in individual mRNAs, which are concatenated serially into long DNA molecules for sequencing. SAGE is useful for the study of rare and abundant transcripts and for the discovery of novel genes.

In the original SAGE approach, mRNA is isolated and double-stranded cDNA is synthesised using biotinylated oligo(dT) primers. Digestion of the cDNA using the so-called anchoring restriction enzyme results in cDNA fragments, 3′-biotinylated ends of which can be isolated using streptavidin beads. The 3′-end cDNA fragments are divided into two portions, which are ligated to two different linkers (A and B) containing the recognition site for the so-called tagging restriction enzyme. Short fragments (tags) are enzymatically released by digestion with the tagging restriction enzyme, which cleaves on average 13 – 14 bp downstream of its recognition site. Blunt ends are created and the released linker tags from pools A and B are ligated tail-to-tail to form a ditag. The linker ditags are amplified by PCR and again cleaved by using the anchoring restriction enzyme to release the ditag. Ditags are separated by polyacrylamide gel electrophoresis and isolated. Overhanging CATG sequences at both sides of the ditags allow ligation of ditags to form concatemers. Longer concatemers are isolated and cloned into a plasmid. By comparing the concatemer sequences with SAGE databases, qualitative and quantitative information about the obtained SAGE tags can be obtained. Careful statistical analysis is required to establish significant differences between expression levels of a certain gene, present in different libraries.

As explained above, the SAGE technique is based on the isolation of unique sequence tags, derived from a defined position in individual mRNAs. The position is defined by the anchoring restriction enzyme used to digest the cDNA. The short SAGE tags are able to discriminate between several thousand transcripts (in the case of a 9-base tag, 262,144 transcripts = 4^9). The human genome contains between 30,000 and 40,000 genes, so that theoretically each SAGE tag is unique to a single gene transcript.

The limitations of SAGE are as follows: i) large amounts of mRNA are required; ii) the average size of cloned concatemers is small; iii) unidentified tags that represent uncharacterised transcripts require further analysis; iv) sequencing errors may lead to an invalid tag-to-gene assignment; v) the short SAGE tags are not always unique; vi) polymorphisms may result in wrong assignment of tags to genes and vii) bias in GC content may occur due to spontaneous denaturation of AT-rich ditags at higher temperatures.

Nevertheless, SAGE is a powerful technique providing massively high-throughput capacities and allowing nearly complete automation of the process. The importance of SAGE in the field of analysing gene expression is demonstrated by the large number of studies performed in particular in the field of cancer research. The data from these experiments have been collated in several cancer databases such as the Cancer Genome Anatomy Project (CGAP) at the National Cancer Institute [301]. Many other SAGE databases exist and a good starting-point is the SAGE database at National Center for Biotechnology Information [302].

3.3.2.1 Rapid analysis of gene expression (RAGE)

Rapid analysis of gene expression (RAGE; Capital Genomix, Gaithersburg, MD, USA) was conceived as a low cost SAGE alternative appropriate for a standard molecular biology laboratory interested in analysing the expression of intermediate numbers of genes. The original design parameters included the potential for analysing any chosen gene without extensive development of new primers or alteration of reaction conditions, high sensitivity, direct assessment of the statistical significance of measurements and the potential for gene discovery. The description was published in 1999 [59,60] and patents covering the technique were issued in 2001 [214,215]. The technology is currently available in kit form.

The basic strategy employed in RAGE is to construct a library of RAGE-tags, short but defined fragments of each mRNA present in a population, that will be used as a template in subsequent PCR reactions. Bioinformatics software supports primer selection with primers selected combinatorially from a small, predesigned set that specifically amplifies the RAGE-tag from a chosen gene based on the sequence of the RAGE-tag. The relative concentration of the corresponding mRNA in the original population is inferred from the level of product formation after PCR amplification. In contrast to SAGE, RAGE gives only relative amounts of any given mRNA between different samples.

For RAGE, cDNA is prepared from an mRNA sample using a biotinylated oligo(dT) primer and the cDNA is immobilised on a streptavidin magnetic bead. To produce sticky ends, the immobilised cDNAs are cleaved with an enzyme 'A', leaving only the 3′-most 'A' fragment attached to the beads and the cleaved 5′-fragments are washed off and discarded. The RAGE-tag is then cleaved from the beads with an enzyme 'B', along with other 3′-fragments and collected. At this point, only the RAGE-tags contain sticky ends derived from enzyme A; the other fragments of cDNA that contaminate the preparation have 'B' sites at both ends. Taking advantage of the unique sticky ends left by the 'A' and 'B' enzymes, the RAGE-tags are then ligated to two unique linkers that distinguish the 'A' and 'B' ends; these provide common 'A' and 'B' primer binding sites for subsequent PCR analysis. The 'A'-end linker is biotinylated so that the RAGE-tags can be purified in the next step by binding to streptavidin beads, eliminating the unwanted cDNA fragments that contain only 'B'-ends. This library of immobilised RAGE-tags can be used as a template in PCRs to amplify particular gene products.

Specificity in the PCR amplification step is provided by using primers that extend past the 'A' or 'B' restriction sites into the unique portion of the RAGE-tag. Good specificity is obtained with a set of 'A'-end primers that contain the 'A'-end linker sequence but extend four nucleotides into the specific

Copyright © Ashley Publications Ltd. Printing and distribution strictly prohibited

portion of the RAGE-tag and a set of 'B'-end primers that contain the 'B'-end linker sequence and extend three nucleotides into the RAGE-tag from the opposite direction. Thus, the total set of primers needed for RAGE analysis is 256 'A'-end primers and 64 'B'-end primers. For any known gene, the sequence of the corresponding RAGE-tag can be determined from GenBank entries and the sequence of the specific 'A'-end and 'B'-end primers that will amplify this RAGE-tag can be inferred. Thus, PCR amplification of the RAGE-tag library with the specific 'A' and 'B'-end primers should give rise to a product of known size. This size is defined by the distance between the 3′-most 'A' restriction site in the gene's cDNA and the closest 'B' restriction site in the 3′-direction and can of course be predicted from the mRNA sequence.

About 5 – 10% of known genes were found that lack either one or both kinds of restriction sites or have closely spaced or overlapping sites that do not give a specific PCR product. However, about half of all mRNAs do not have a 'A' site to the left of the 3′-most 'B' site and therefore the RAGE-tag library will not represent this half of the transcriptome [59]. To assay these genes, a second RAGE-tag library is made by reversing the order in which the restriction cuts are made. Use of both libraries therefore allows to assay ~ 90 – 95% of the transcriptome [59].

To completely assay the entire transcriptome, each library would have to be amplified with each possible combination of RAGE primers (32,768 PCRs). However, in most cases the size of the PCR products produced from two different genes will be distinguishable using standard electrophoretic separation techniques [59]. Since the average RAGE-tag is ~ 130 bp in length, sequencing unknown PCR products after gel purification usually gives enough information to uniquely identify corresponding ESTs.

Overall, RAGE is a flexible open architecture system that can be used for a wide variety of applications. Although it has the potential to be used as a high-throughput system, it has been currently used only for the analysis of medium numbers of targets within a sample.

3.3.2.2 MicroSAGE

Datson *et al.* developed microSAGE [61,62], a modification of the SAGE procedure that allows the use of 500 – 5,000-fold less starting material than in the conventional SAGE approach. The microSAGE method was recently enhanced by improved *Nla*III digestion of the PAGE-purified ditags by addition of a single purification step [63].

Leaving the basic principle of SAGE unaltered, microSAGE is a 'single-tube' procedure in which all steps from RNA isolation to tag release are performed in a single tube. This is achieved by directly immobilising the poly(A)$^+$ mRNA fraction of a total RNA sample to the streptavidin-coated wall of the reaction tube using a biotinylated oligo(dT) primer without mRNA extraction. The biotinylated oligo(dT) primer also serves as a primer in the subsequent cDNA synthesis step and the cDNA remains immobilised to the tube. Overall, this avoids extraction and precipitation between each subsequent reaction step of the SAGE procedure. The enzymes of the previous reactions are simply removed by heat inactivation and disposal of the solution. After washing and change of buffer, the next reaction can be performed in the same tube. Thus, the number of manipulations and the accompanying loss of material is reduced. Another change from the original protocol is that a limited number of cycles of PCR are performed to reamplify the gel-purified ditag bands to generate sufficient ditags to be cloned and sequenced.

Using microSAGE, an expression profile can be obtained from as little as 1 – 5 ng of mRNA, allowing expression profiling in small tissue specimens, and microdissections of complex heterogeneous tissues, biopsies, scarce biological material etc. Thus, microSAGE is an interesting alternative to the original SAGE approach, in particular if only limited amount of RNA is available. However, Blackshaw *et al.* showed very recently that microSAGE is highly representative and reproducible but reveals major differences in gene expression among samples obtained from similar tissues so that care should be taken in generalising results obtained from libraries constructed from tissue obtained from different individuals and/or processed or stored differently [64].

3.3.2.3 MiniSAGE

To allow the use of minute amounts of starting material, usually an additional amplification step is performed either at the cDNA level before SAGE ditags have been established or after formation of the ditags. The MiniSAGE method described by Yea *et al.* dispenses with an additional PCR amplification step [65] but nevertheless requires only 1 µg of total RNA, several hundred times less than that required by the original SAGE protocol. This is advantageous because PCR amplification potentially introduces bias and may comprise the quantitative aspects of the SAGE method.

The MiniSAGE technique is based on three key modifications of the original SAGE procedure: i) use of a phase-lock gel system to increase recovery and purity of DNA material after each extraction step; ii) reduction of the amount of linker oligonucleotides in the ligation, thereby minimising their interference with SAGE ditag amplification and increasing the yield of SAGE ditags and iii) integration of an mRNA capture kit to allow the first five steps of the SAGE procedure [mRNA isolation; reverse transcription using a biotin-labelled oligo(dT) primer; enzymatic digestion of cDNA; binding of digested biotin-labelled 3′-terminal cDNA fragments to streptavidin-coupled magnetic beads, ligation of linker oligonucleotides containing recognition sites for a tagging enzyme to the bound cDNA fragments and release of cDNA tags] to be performed within one tube, thus preventing loss of material between successive steps.

The MiniSAGE protocol allows expression profiling from only 1 µg of total RNA without an additional PCR amplification, as in those performed in other SAGE modifications,

Copyright of Ashley Publications Ltd. Printing and distribution strictly prohibited

such as microSAGE, PCR-SAGE and SAGE-Lite [61,62,66-68]. This is an incontestable advantage. As yet, however, the only experience gathered with this procedure is that generated by the researchers who described the method.

3.3.2.4 PCR-SAGE and SAGE-Lite

PCR-SAGE and SAGE-Lite were developed to overcome the limiting large quantities of RNA required for SAGE [66-68]. Both methods utilise PCR to amplify transcripts as a first step, lifting the amount of starting material for SAGE from picogram to the requisite microgram levels.

For PCR-SAGE [66,68], total RNA is extracted and SMART™ II cDNA Synthesis (BD Biosciences, Franklin Lakes, NJ, USA) system is used with an adapter-modified oligo(dT) primer to initiate first-strand cDNA synthesis. The modified primer ensures the integration of a PCR priming site at the 3′-end of each cDNA. The SMART II system completes cDNA synthesis at the 5′-end of the mRNA by adding a short tract of cytosines utilising the ability of the Superscript II reverse transcriptase to add cytosines to a cDNA strand [69,70,216,217]. This tract hybridises with an oligonucleotide containing a short tract of guanines. Following hybridisation, the reverse transcriptase switches strands and extends the cDNA using the hybridised oligonucleotide sequence as a template. This process incorporates a second priming sequence at the 5′-end of each cDNA molecule. The technology is called strand-switching technology and is used widely to capture small amounts of mRNA. To ensure that the final SAGE template is only biotinylated at the poly(A) end, cDNA synthesis at the 5′-end of the mRNA is completed by the incorporation of a modified SMART II oligonucleotide containing a restriction site for a rarely cutting enzyme not present in the 3′-cDNA synthesis primer. Design of the PCR-SAGE primers is that both cDNA strands have identical 3′-termini, allowing amplification with a single primer. PCR amplification maintains the restriction site at the 5′-end of the cDNA but does not introduce this restriction enzyme site at the 3′-end corresponding to the poly(A) tail. Using a 5′-biotinylated primer, PCR generates cDNA with biotin tags at both ends. Digestion of PCR-amplified cDNA removes the biotin tag and a small amount of attached sequence only from the 5′-end of the mRNA. The cleaved 5′-ends are small (31 bp in length) and removed by size fractionation to yield cDNAs biotinylated solely at the end corresponding to the poly(A)$^+$ tail of the original mRNA.

SAGE-Lite is based on the same principle as PCR-SAGE but utilises slightly different primer sequences for cDNA synthesis [67,68]. Biotinylation of cDNA at the end corresponding to the poly(A) tail is achieved by performing PCR with the same primers used to synthesise cDNA. In this case, only the oligo(dT) primer is biotinylated. The result is an amplified SAGE template that is biotinylated at one end only. The cDNA amplified by SAGE-Lite produces 2 – 5 µg of material, enough for the construction of a SAGE catalogue [67]. PCR-SAGE and SAGE-Lite both use conventional SAGE protocols for further processing of the cDNAs.

Both procedures rely on PCR to amplify cDNA synthesised from < 50 ng of total RNA and both utilise strand switching by reverse transcriptase to introduce universal priming sites at each end of the cDNA. SAGE-Lite and PCR-SAGE differ only in the mechanisms used to ensure unilateral biotinylation of cDNA. With the advent of new technologies, such as laser capture microdissection, for harvesting individual cells from tissue, the demand for transcriptome catalogues from small quantities of RNA will increase. SAGE-Lite and PCR-SAGE offer alternative solutions to this problem.

3.3.2.5 SAGE adaption for downsized extracts (SADE)

The principle of SAGE adaption for downsized extracts (SADE) procedure is similar to the original SAGE protocol but is essentially a single-tube assay from the obtained sample lysate to cDNA tag recovery [71-73,218-220]. In addition, SADE includes the following modifications compared to the original SAGE protocol: i) a single-step method for mRNA extraction from tissue lysate; ii) use of a reverse transcriptase lacking ribonuclease H activity; iii) use of a different anchoring enzyme; iv) modification of procedure for blunt-ending cDNA tags and v) design of modified primers for PCR amplification.

In the first step of SADE, poly(A)$^+$ mRNAs are isolated directly from the tissue lysate through binding to oligo(dT) primers covalently bound to magnetic beads and the cDNA is synthesised immediately. Following second strand synthesis, the double-stranded cDNA fraction still present on beads is digested with the anchoring enzyme. The remaining 3′-ends of bound cDNA are divided into two fractions. Each fraction is ligated to either linker primer A or linker primer B and then digested enzymatically. The released cDNA tags are blunt-ended and the two fractions are then ligated to each other. Following PCR amplification of the resulting ditag, the sample is digested with the anchoring enzyme and ditags are purified and concatenated as described in the original SAGE method. Concatemers of > 350 bp are recovered, cloned into plasmids and counted by sequencing.

The modifications implemented in SADE are not of equal importance. The studies performed by Virlon *et al.* [73] showed that SADE increased the efficiency 4-fold at the stage of synthesis of double-stranded cDNA and ~ 400-fold at the stage of digesting cDNA by the anchoring enzyme. However, the modifications integrated in SADE allow the construction of tag libraries from a limited amount of starting material of 15,000 – 50,000 cells but SADE is also flexible enough to allow the analysis of macro-amounts of cells.

3.3.2.6 Small amplified RNA-SAGE (SAR-SAGE)

Unless PCR steps are added to the SAGE protocol, SAGE is still limited by the need for ~ 1 µg of total RNA. As extra PCR amplification might introduce representation biases, current SAGE protocols are not fully suitable for the study of small, microdissected tissue samples. Small amplified RNA-SAGE (SAR-SAGE) is an alternative method involving the linear amplification of small mRNA fragments containing the SAGE tags [74].

Copyright of Ashley Publications Ltd Printing and distribution strictly prohibited

According to classical SAGE protocols, 50 ng of total RNA are transformed into double-stranded cDNA. After cleavage with the anchoring enzyme, a T7-SAGE adaptor containing the promoter of the T7 RNA polymerase is ligated to the cohesive end of the cDNAs; this allows for the synthesis of small cRNA molecules harbouring the SAGE tags. These cRNAs are then used in the classical microSAGE procedure described above.

SAR-SAGE offers an interesting alternative to prePCR amplification in the preparation of SAGE libraries from small amounts of microdissected tissue samples. The procedure allows preparation of libraries of > 100,000 tags from as few as 2,500 cells and is thus applicable for all current applications.

3.3.2.7 Tandem arrayed ligation of expressed sequence tags (TALEST)™

A drawback of SAGE is its reliance on PCR amplification to generate ditags. In order to circumvent the PCR biases, a cDNA tag-based technique called tandem arrayed ligation of expressed sequence tags (TALEST™, Chugai Biopharmaceuticals, San Diego, CA, USA) was developed that does not rely on PCR amplification to generate tag arrays [75,76,221]. The TALEST technique retains the high-throughput advantages of SAGE but avoids quantitative biases associated with PCR by using traditional bacterial amplification of clones rather than PCR amplification.

In TALEST, double-stranded cDNA is prepared using a biotinylated oligo(dT) primer. The cDNA is treated with *Eco*RI methylase to protect internal *Eco*RI sites and digested with a 4-base cutting restriction enzyme (the punctuating enzyme). The 3′-most fragment is isolated using streptavidin-coated magnetic particles. A double-stranded adapter molecule with a 5′-overhang compatible with the punctuating enzyme is ligated to the cDNA. The adapter introduces a type IIs restriction site immediately 5′ to the ligated cDNA and contains an *Eco*RI site at its 5′-terminus to facilitate later cloning. Digestion of the adapter-ligated cDNA with a type IIs restriction enzyme releases a linear DNA fragment consisting of the adapter and additional 10 – 12 nucleotides of unknown cDNA sequence separated from the adapter by the punctuation sequence. This fragment is ligated to a second biotinylated adapter molecule containing a *Not*I site at its 5′-end and a 16-fold degenerate 3′-overhang which renders it compatible with all possible cDNA sequences released by the punctuating enzyme. This adapter introduces a *Msp*I site to the 3′ end of the original DNA fragment such that all the molecules contain a 12-base cDNA-derived tag sequence flanked at both ends by *Msp*I sites. The resulting molecule is double-digested with *Eco*RI/*Not*I and the biotinylated fragment is removed by using streptavidin-coated magnetic particles. The remaining *Eco*RI/*Not*I fragment is isolated by polyacrylamide gel electrophoresis to resolve it from adapter dimers. The result is an *Eco*RI/*Not*I-tailed DNA fragment containing a 12 bp cDNA tag flanked at both ends by the *Msp*I sequence which is eluted and cloned. The recombinant plasmids are transformed into *E. coli* in order to generate a tag library. Plasmid DNA is isolated and digested with *Msp*I to release the tags. The tag fragments are purified by polyacrylamide gel electrophoresis and ligated to form concatemers. Concatemers of sufficient minimal length are isolated by agarose gel electrophoresis and cloned into standard plasmids for sequencing. Each array consists of 20 – 60 12-base tag sequences separated from each other and from the plasmid backbone by the defined 4-base punctuating enzyme sequence. Sequencing and counting of tags reveal levels of expression of transcripts by allocating tags to transcripts.

TALEST is a reproducible high-throughput alternative to SAGE that can be applied to several biological questions and offers high sensitivity avoiding biases usually introduced by using PCR amplification.

3.4 Real-time RT-PCR

The fluorescence-based real-time RT-PCR technique integrates the amplification and analysis steps of the PCR reaction, thereby eliminating the need for post-PCR processing. It does this by monitoring the amount of DNA produced during each PCR cycle and its sensitivity, specificity and wide dynamic range have revolutionised the approach to PCR-based quantification of RNA, making it the method of choice for quantifying steady-state mRNA levels [77-79]. Three key innovations make this technology possible. First, a process was developed that permitted the simultaneous amplification and detection of specific DNA sequences. The original system involved the addition of an intercalator such as ethidium bromide to the PCR reaction and the use of a thermal cycler with an ultraviolet light source and a computer-controlled charge-coupled device (CCD) camera [80,81]. A plot of increased fluorescence, caused by ethidium bromide binding to newly synthesised DNA, against cycle allowed the calculation of the quantity of starting material. The drawback of this approach was that this measurement could not easily distinguish between specific and nonspecific products. The second innovation provided specificity by adding a separate oligonucleotide probe, labelled with ^{32}P at its 5′-end and designed to hybridise within the target amplicon during the amplification. This generates a substrate suitable for the double-strand-specific 5′–3′ exonuclease activity of *Taq* DNA polymerase [82] and cleavage of the probe can occur only if the target sequence is being amplified. However, measurement of the ^{32}P-labelled cleavage products was not straightforward and required their separation by thin layer chromatography. The introduction of fluorogenic probes [83] made it possible to amalgamate the two methodologies in the 5′-nuclease assay, the first practical real-time fluorescence-based quantitative PCR method [84].

The principles of this technique and the different approaches (such as the DNA-binding dye SYBR® Green [Molecular Probes, Leiden, The Netherlands] and different kinds of fluorescence probes: TaqMan® probes [Molecular Systems, Pleasanton, CA, USA], hybridisation probes, Molecular Beacons, Scorpions® [DxS, Manchester, UK]) that have been developed to enhance this method have been extensively discussed in the literature and the reader is therefore referred to previously

Copyright © Ashley Publications Ltd. Printed and distribution strictly prohibited

published reviews [77-79]. However, real-time RT-PCR is a powerful closed architecture technique allowing rapid and sensitive screening of the expression of large numbers of transcripts simultaneously in numerous samples. This strength of real-time RT-PCR is further enhanced by the new microfluidic card system provided by Applied Biosystems (Weiterstadt, Germany) that allows fast sample preparation within a few minutes in low reaction volumes of 2 µl in a 384-well format.

4. Limitations

Nearly all of the methods presented here begin with the isolation of total RNA or mRNA. In order to ensure accurate analyses, it is important that the RNA analysed truly reflects the *in vivo* expression pattern of the sample. Changes in the expression profile incursed by downregulation of mRNA expression and enzymatic degradation of RNA can occur very rapidly during sample collection and handling and during RNA isolation. Flawless handling and RNA isolation are therefore prerequisite for accurate expression analysis with all the methods described. The underlying molecular biological principles also cause inherent limitations and introduce problems to each method. In general, important problems facing PCR-based methods is the amplification bias in favour of smaller cDNA molecules causing larger transcripts to be under-represented, the co-amplification of cDNA fragments due to unspecific binding of short arbitrary primers and the large number of PCR cycles introducing false-positive and false-negative results. In addition, integration of second strand DNA synthesis, purification of cDNA, restriction enzyme digestion or adapter primer ligation, probe labelling, hybridisation or additional washing steps can introduce and amplify errors or can lead to the loss of particularly low abundant mRNAs. Problems of reproducibility, probe sensitivity, non-linearity in signal detection, probe cross-hybridisation due to homologous cDNA sequences and faulty data management have been reported mainly for hybridisation-based methods such as microarrays [26]. Sequencing-based methods, in general, require larger amounts of starting material (the introduction of amplification procedures reduces the amount of starting material but increases the risk of introducing some of the problems just mentioned), huge numbers of sequence tags to be counted, under-representation of low abundant mRNAs and biases in GC content. For reasons of brevity, the advantages and disadvantages of each method is not discussed in detail but an overview is presented in **Table 2**.

5. Future trends

Each of the methods described here has its particular niche and each has its specific drawbacks. Although analysis of gene expression using DNA microarrays is undoubtedly a powerful and appealing technique, it suffers from the inherent drawback that only known sequences can be presented on the microarray. In addition, this technique is expensive and has been beset by problems related to quality control and lack of sensitivity. Differential display RT-PCR, for example, is at first sight an inexpensive method but is technically unforgiving and labour-intensive. However, differential display RT-PCR has the great advantage of revealing changes in the level of expression of unknown genes. The methods based on gene sequencing such as MPSS [53,54] and SAGE [56,57] also reveal the regulation of unknown genes but are expensive and require a level of infrastructure that effectively limits their use to the more non-academic setting. The introduction of lower-cost techniques such as RAGE and microSAGE may expand the use of sequence-based methods to the academic research laboratory in coming years, although at present these laboratories appear to favour the use of DNA microarrays for gene expression analysis.

The analysis of gene expression is a dynamic and rapidly moving field. so that despite the dizzying array of available technologies, new techniques may be expected to emerge in the coming years. However, these new techniques will need to demonstrate clear advantages in terms of cost, robustness or throughput if they are to gain acceptance in this very competitive sector. One development that may be expected is the attempt to combine methods in order to exploit their specific advantages. For example, the convenience of DNA microarrays might be combined with the sensitivity of PCR-based methods, while avoiding the bias typically associated with a PCR-based approach. Further developments are also likely to focus on widening the dynamic range of methods, increasing the degree of automation and improving sample preparation, processing and analysis.

Clearly, laboratory procedures for performing massively parallel large-scale analyses of mRNA expression of multiple samples at high sensitivity are material-, cost- and labour-intensive. They also require equipment of high technical standard and increased input of bioinformatics. Thus, genome-wide expression profiling using the methods described here may be out of reach for most modest university laboratories. Not surprisingly, many such methods have been developed for industrial use and are offered as a commercial service. Nevertheless, some methods like differential display RT-PCR (Genhunter), SAGE and RAGE are available as kits and are suitable for well-equipped university laboratories. Several of the high-throughput methods presented here can likewise be adapted to the needs of smaller laboratories if only partial screening of the expression patterns of a few samples is of interest. For example, cDNA-AFLP and differential display RT-PCR have been widely and successfully used to identify genes that are differentially expressed in certain diseases. In most such studies, expression profiling was used only initially to identify a few genes which were subsequently studied in more detail to characterise their physiological role in the context of the primary scientific query. In contrast, methods such as GeneCalling and MPSS are tailored for genome-wide expression profiling. Their high technical demands make them inaccessible for small-scale research programmes. It is of note that methods such as SAGE, OFP and MPSS enfold their full potential only when they are used on a large scale.

Copyright of Ashley Publications Ltd Printing and distribution strictly prohibited

Table 2. Most prominent advantages and disadvantages of the methods reviewed.

Method	Advantages	Disadvantages
ADGE	Open architecture system Low cost method Highly sensitive Applicable for university laboratory	Multiple-step sample processing Low sample throughput Use of multiple restriction enzymes is required for genome-wide profiling Identification of DNA fragments is difficult
AFLP	Open architecture system Low cost method Highly sensitive Applicable for university laboratory	Multiple-step sample processing Low sample throughput Use of multiple restriction enzymes is required for genome-wide profiling Identification of DNA fragments is difficult PCR end-point analysis only
BADGE	High sample throughput Fast sample processing Standardised protocol	Closed architecture system Multiple-step sample processing Parallel analysis of limited number of targets Cross-hybridisation is possible
DEPD®	Open architecture system Highly sensitive Applicable for university laboratory Commercial service available Automated procedure available	Multiple-step sample processing Low sample throughput Use of multiple restriction enzymes is required for genome-wide profiling Identification of DNA fragments is difficult PCR end-point analysis only
FDD	Open architecture system Two-step sample processing Applicable for university laboratory Protocol easy to follow Low cost method Highly sensitive Commercial kits available Commercial service available Automated procedure available Widely used method	Low sample throughput Identification of DNA fragments is difficult PCR end-point analysis only Low stringency conditions during PCR due to the use of short arbitrary primers
GeneCalling®	No target amplification required Multiple-step but fast sample processing Automated procedure available	Closed architecture system Use of multiple restriction enzymes is required for genome-wide profiling High dependence on annotation quality of databases Use restricted to well-sequenced organisms
LEAD	Open architecture system Applicable for university laboratory Low cost method Highly sensitive	Multiple-step sample processing Identification of DNA fragments is difficult PCR end-point analysis only
Microarrays	Protocol easy to follow and standardised Large numbers of genes measured in parallel Fast sample processing Commercial service available Commercial systems and kits available	Closed architecture system Low dynamic range of quantification Relatively low sensitivity Probe inhomogeneity and array-to-array heterogeneity Cross-hybridisation possible High dependence on annotation quality of databases High cost of both arrays and equipment
Microplate-based assays	Protocol easy to follow and standardised Fast sample processing High sample throughput Commercial kits available Applicable for university laboratory Highly sensitive and reproducible (No target amplification required) (Direct use of cell lysates)	Closed architecture system Multiple-step sample procedure Parallel analysis of limited number of targets

ADGE: Amplified differential gene expression; AFLP: Amplified fragment length polymorphism; BADGE: Beads array for the detection of gene expression; DEPD: Digital expression pattern display; FDD: Fluorescent mRNA differential display; LEAD: Ligation specificity-based expression analysis display; PCR: Polymerase chain reaction.

Table 2. Most prominent advantages and disadvantages of the methods reviewed *(continued)*.

Method	Advantages	Disadvantages
MicroSAGE	Open architecture system Absolute quantification Lower amounts of starting material (compared to SAGE™) Single-tube procedure Faster sample processing (compared to SAGE)	Multiple-step sample processing Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags
MiniSAGE	Open architecture system Absolute quantification Lower amounts of starting material (compared to SAGE) Faster sample processing (compared to SAGE)	Multiple-step sample processing Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags
MPSS™	Open architecture system Absolute quantification Commercial service available Highly sensitive and reproducible Fully automated process	Multiple-step sample processing Low sample throughput High costs and technically demanding Not applicable for university laboratory Closed architecture system
OFP	Open architecture system Highly sensitive and reproducible Fully automated process	Multiple-step sample processing Low sample throughput High costs and technically demanding Not applicable for university laboratory
PCR-SAGE	Open architecture system Absolute quantification Low amounts of starting material Highly sensitive	Multiple-step sample processing Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags
RAGE	Open architecture system Low amounts of starting material Highly sensitive	Multiple-step sample processing Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags Only relative quantification possible
READS™	Open architecture system Highly sensitive Applicable for university laboratory	Multiple-step sample processing Low sample throughput Use of multiple restriction enzymes is required for genome-wide profiling Identification of DNA fragments is difficult PCR end-point analysis only
Real-time RT-PCR	Absolute quantification possible Protocol easy to follow and standardised Fast sample processing No PCR end-point analysis Highly sensitive Highly reproducible High linear dynamic range of measurement Applicable for university laboratory	Closed architecture system Relatively high cost of equipment Labour-intensive sample preparation (if no robot is available)
SADE	Open architecture system Absolute quantification Single-tube procedure Lower amounts of starting material (compared to SAGE) Faster sample processing (compared to SAGE)	Multiple-step sample processing Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags

MPSS: Massively parallel signature sequencing; OFP: Oligonucleotide fingerprinting; PCR: Polymerase chain reaction; RAGE: Rapid analysis of gene expression; READS: Restriction enzyme analysis of differential sequences; RT-PCR: Reverse transcriptase-polymerase chain reaction; SADE: SAGE adaption for downsized extracts; SAGE: Serial analysis of gene expression.

Copy

Table 2. Most prominent advantages and disadvantages of the methods reviewed *(continued)*.

Method	Advantages	Disadvantages
SAGE	Open architecture system Absolute quantification Commercial kits available	Multiple-step sample processing Time-consuming Low sample throughput Requires relatively large amounts of starting material Sensitivity and reproducibility depends on the number of sequenced tags
SAGE-Lite	See PCR-SAGE	See PCR-SAGE
SAR-SAGE	Open architecture system Absolute quantification Avoids PCR amplification Highly sensitive	Multiple-step sample processing Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags
TALEST™	Open architecture system Absolute quantification Avoids PCR amplification of target	Time-consuming Low sample throughput Sensitivity and reproducibility depends on the number of sequenced tags
TOGA™	Open architecture system Highly sensitive Commercial service available Automated procedure available	Closed architecture system Use of multiple restriction enzymes is required for genome-wide profiling High dependence on annotation quality of databases Use restricted to well-sequenced organisms

SAGE: Serial analysis of gene expression; SAR-SAGE: Small amplified RNA-SAGE; TALEST: Tandem arrayed ligation of expressed sequence tags; TOGA: Total gene expression analysis.

Although the subject has not been dealt with here, development is also likely to concentrate more on the 'software', the bioinformatics of gene expression analysis, rather than on the technical platforms or 'hardware'. Here, parallels with developments within the computer industry may be expected. Such tools will include improved comparison methods and methods for the classification of genes into metabolic or functional pathways. In addition, the process of the annotation of genes (i.e., the putative assignment of function based on sequence information) is likely to undergo rapid development.

6. Expert opinion

Despite their undoubted advantage in terms of ease of use, microarrays continue to be beset by problems of small dynamic range, lack of sensitivity, quality control and their closed architecture. It is therefore important to remember that many other good methods exist for the analysis of gene expression. The principal advantage of many of these methods is that they require no prior sequence information. This feature is likely to increase in importance in the postgenomic era, in which there is a shift from sequence analysis to analysis of function and physiological relevance. In practical terms, a major drawback of both DNA microarrays and many of the techniques described in the present report is their high cost. Substantial cost reductions will be required if these techniques are to become universally applicable, particularly in the academic research setting. In addition, much effort will be needed to be devoted to the standardisation of gene expression data and to techniques for the storage of this data in universally accessible databases.

Copyright of Ashley Publications Ltd Printing a

Bibliography

1. GREEN CD, SIMONS JF, TAILLON BE, LEWIN DA: Open systems: panoramic views of gene expression. *J. Immunol. Methods* (2001) **250**:67-79.
2. LORKOWKSI S, CULLEN P *et al.*: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002).
3. CULLEN P, LORKOWSKI S: DNA microarrays in medicine. *Expert Opin. Ther. Patents* (2002) **12**:1783-1794.
4. MATZ M, USMAN N, SHAGIN D, BOGDANOVA E, LUKYANOV S: Ordered differential display: a simple method for systematic comparison of gene expression profiles. *Nucleic Acids Res.* (1997) **25**:2541-2542.
5. MATZ M: ODD – Ordered differential display. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):326-332.
6. KORNMANN B, PREITNER N, RIFAT D, FLEURY-OLELA F, SCHIBLER U: Analysis of circadian liver gene expression by ADDER, a highly sensitive method for the display of differentially expressed mRNAs. *Nucleic Acids Res.* (2001) **29**:e51.
7. KORNMANN B, PREITNER N, RIFAT D, SCHIBLER U: ADDER – Amplification of double-stranded cDNA end restriction fragments. In: *Analysing Gene Expression. A Handbook of Methods*: *Possibilities and Pitfalls.* Lorkowski S, Cullen P, eds. Wiley-VCH, Weinheim, Germany (2002):267-277.
8. SOUTHERN EM, MITCHELL AR: Chromatography of nucleic acid digests on thin layers of cellulose impregnated with polyethyleneimine. *Biochem. J.* (1971) **123**:613-617.
9. SCHENA M, SHALON D, DAVIS RW, BROWN PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* (1995) **270**:467-470.
10. KING HC, SINHA AA: Gene expression profile analysis by DNA microarrays: promise and pitfalls. *JAMA* (2001) **286**:2280-2288.
11. RAMDAS L, COOMBES KR, BAGGERLY K *et al.*: Sources of non-linearity in cDNA microarray expression measurements. *Genome Biol.* (2001) **2**:RESEARCH0047.
12. RICHMOND CS, GLASNER JD, MAU R, JIN H, BLATTNER FR: Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* (1999) **27**:3821-3835.
13. GIBBS WW: Shrinking to enormity: DNA microarrays are reshaping basic biology – but scientists fear they may soon drown in the data. *Sci. Am.* (2001) **284**:33-34.
14. YANG L, TRAN DK, WANG X: BADGE, beads array for the detection of gene expression, a high-throughput diagnostic bioassay. *Genome Res.* (2001) **11**:1888-1898.
15. ADAMS MD, KELLEY JM, GOCAYNE JD *et al.*: Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* (1991) **252**:1651-1656.
16. BAUER O, JANITZ M, GUERASIMOVA A, HERWIG R, LEHRACH H, RADELOF U: OFP – Oligonucleotide fingerprinting. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowksi S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):463-471.
17. POUSTKA A, POHL T, BARLOW DP *et al.*: Molecular approaches to mammalian genetics. *Cold Spring Harb. Symp. Quant. Biol.* (1986) **51**:131-139.
18. WARRIOR U, FAN Y, DAVID CA *et al.*: Application of QuantiGene nucleic acid quantification technology for high throughput screening. *J. Biomol. Screen*. (2000) **5**:343-52.
19. DE ARRUDAM, LYAMICHEV VI, EIS PS *et al.*: Invader technology for DNA and RNA analysis: principles and applications. *Expert Rev. Mol. Diagn.* (2002) **2**:487-496.
20. KWIATKOWSKI RW, LYAMICHEV V, DE ARRUDA M, NERI B: Clinical, genetic, and pharmacogenetic applications of the Invader assay. *Mol. Diagn.* (1999) **4**:353-364.
21. LYAMICHEV VI, KAISER MW, LYAMICHEVA NE *et al.*: Experimental and theoretical analysis of the invasive signal amplification reaction. *Biochemistry* (2000) **39**:9523-9532.
22. URDEA MS: Branched DNA signal amplification. *Biotechnology (NY)* (1994) **12**:926-928.
23. SELIGMANN B: Improved genomics and proteomics array technology. *Am. Biotechnol. Lab* (2001) **19**:34.
24. SHEN L, SHERIDAN P, CAO WW *et al.*: Quantification of cytokine mRNA in peripheral blood mononuclear cells using branched DNA (bDNA) technology. *J. Immunol. Methods* (1998) **215**:123-134.
25. LIANG P, PARDEE AB: Differential display of eukaryotic mRNA by means of the polymerase chain reaction. *Science* (1992) **257**:967-971.
26. FISHER JS, MEADE J, WALDEN J, GUO Z, LIANG P: FDD – Automated fluorescent mRNA differential display. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowksi S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):494-501.
27. LIANG P, PARDEE AB: Recent advances in differential display. *Curr. Opin. Immunol.* (1995) 7:274-280.
28. ITO T, KITO K, ADATI N, MITSUI Y, HAGIWARA H, SAKAKI Y: Fluorescent differential display: arbitrarily primed RT-PCR fingerprinting on an automated DNA sequencer. *Fed. Eur. Biochem. Soc. Lett.* (1994) **351**:231-236.
29. CHO Y, MEADE J, WALDEN J, GUO Z, LIANG P: Multicolor fluorescent differential display. *Biotechniques* (2001) **30**:562-572.
30. BUESS M, MORONI C, HIRSCH HH: Direct identification of differentially expressed genes by cycle sequencing and cycle labelling using the differential display PCR primers. *Nucleic Acids Res*. (1997) **25**:2233-2235.
31. AITTOKALLIO T, OJALA P, NEVALAINEN TJ, NEVALAINEN O: Analysis of similarity of electrophoretic patterns in mRNA differential display. *Electrophoresis* (2000) **21**:2947-2956.
32. AITTOKALLIO T, OJALA P, NEVALAINEN TJ, NEVALAINEN O: Automated detection of differentially expressed fragments in mRNA differential display. *Electrophoresis* (2001) **22**:1935-1945.
33. QIN L, PRINS P, JONES JT *et al.*: GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. *Nucleic Acids Res.* (2001) **29**:1616-1622.
34. IRIE T, OSHIDA T, HASEGAWA H *et al.*: Automated DNA fragment collection by capillary array gel electrophoresis in search of differentially expressed genes. *Electrophoresis* (2000) **21**:367-374.

Copyright © Ashley Publications. Unauthorized reproduction and distribution strictly prohibited

35. CHEN ZJ, SHEN H, TEW KD: Gene expression profiling using a novel method: amplified differential gene expression (ADGE). *Nucleic Acids Res.* (2001) **29**:e46.

36. VOS P, HOGERS R, BLEEKER M *et al.*: AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* (1995) **23**:4407-4414.

37. BACHEM CW, VAN DER HOEVEN RS, DE BRUIJN SM, VREUGDENHIL D, ZABEAU M, VISSER RG: Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant J.* (1996) **9**:745-753.

38. BACHEM CWB, DEL CARMEN AF, VISSER RGF: cDNA–AFLP – cDNA-based amplified-fragment length polymorphism fingerprinting. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowksi S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):285-296.

39. MONEY T, READER S, QU LJ, DUNFORD RP, MOORE G: AFLP-based mRNA fingerprinting. *Nucleic Acids Res.* (1996) **24**:2616-2617.

40. BACHEM CWB, OOMEN RFJF, VISSER RGF: Transcript imaging with cDNA–AFLP: a step-by-step protocol. *Plant Mol. Biol. Reporter* (1998) **16**:157-173.

41. LÜBBERT H: DEPD® – Digital expression pattern display. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):484-494.

42. MAELICKE A, LÜBBERT H: DEPD, a high resolution gene expression profiling technique capable of identifying new drug targets in the central nervous system. *J. Recept. Signal. Transduct. Res.* (2002) **22**:283-295.

43. SHIMKETS RA, LOWE DG, TAI JT *et al.*: Gene expression analysis by transcript profiling coupled to a gene database query. *Nat. Biotechnol.* (1999) **17**:798-803.

44. SHIMKETS RA: GeneCalling™. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P, (Eds), Wiley-VCH, Weinheim, Germany (2002):501-505.

45. LI BY, KWAN WK, TURGEON G *et al.*: Analysis of differential gene expression by ligation specificity-based transcript profiling. *OMICS* (2002) **6**:175-185.

46. LORKOWSKI S., CULLEN P: READS™ – Restriction enzyme analysis of differential sequence. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds) Wiley-VCH, Weinheim, Germany (2002)b:518-529.

47. PRASHAR Y, WEISSMAN SM: Analysis of differential gene expression by display of 3′ end restriction fragments of cDNAs. *Proc. Natl. Acad. Sci. USA* (1996) **93**:659-663.

48. PRASHAR Y, WEISSMAN SM: READS: a method for display of 3′ end fragments of restriction enzyme-digested cDNAs for analysis of differential gene expression. *Methods Enzymol.* (1999) **303**:258-272.

49. WERLE E, SCHNEIDER C, RENNER M, VOLKER M, FIEHN W: Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res.* (1994) **22**:4354-4355.

50. HILBUSH BS, LO D, CHEN L, SUTCLIFFE JG: TOGA™ – Total gene expression analysis. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):529-537.

51. LO D, HILBUSH B, SUTCLIFFE JG: TOGA analysis of gene expression to accelerate target development. *Eur. J. Pharm. Sci.* (2001) **14**:191-196.

52. SUTCLIFFE JG, FOYE PE, ERLANDER MG *et al.*: TOGA: an automated parsing technology for analyzing expression of nearly all genes. *Proc. Natl. Acad. Sci. USA* (2000) **97**:1976-1981.

53. BRENNER S, JOHNSON M, BRIDGHAM J *et al.*: Gene expression analysis by massively parallel signature sequencing (MPSS™) on microbead arrays. *Nat. Biotechnol.* (2000) **18**:630-634.

54. WOYCHIK R, BRENNER S, BURCHAM T, CORCORAN K, ALBRECHT G, RUSSELL N: MPSS™ – Massively parallel signature sequencing. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):538-550.

55. BRENNER S, WILLIAMS S, VERMASS EH *et al.*: *In vitro* cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA* (2000) **97**:1665-1670.

56. VELCULESCU VE, ZHANG L, VOGELSTEIN B, KINZLER KW: Serial analysis of gene expression. *Science* (1995) **270**:484-487.

57. DE WAARD V, BEAUCHAMP NJ, VAN ZONNEVELD A-J, PANNEKOEK H: SAGE™ – Serial analysis of gene expression. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowksi S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):551-557.

58. VELCULESCU VE, VOGELSTEIN B, KINZLER KW: Analysing uncharted transcriptomes with SAGE. *Trends Genet.* (2000) **16**:423-425.

59. RAY JM, MACLEOD MC: RAGE – Rapid analysis of gene expression. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds.) Wiley-VCH, Weinheim, Germany (2002):511-518.

60. WANG A, PIERCE A, JUDSON-KREMER K *et al.*: Rapid analysis of gene expression (RAGE) facilitates universal expression profiling. *Nucleic Acids Res*. (1999) **27**:4609-4618.

61. DATSON NA, VAN DER PERK-DE JONG J, VAN DEN BERG MP, DE KLOET ER, VREUGDENHIL E: MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucleic Acids Res.* (1999) **27**:1300-1307.

62. DATSON NA, VREUGDENHIL E: MicroSAGE. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowksi S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):557-568.

63. ANGELASTRO JM, KLIMASCHEWSKI LP, VITOLO OV: Improved NlaIII digestion of PAGE-purified 102 bp ditags by addition of a single purification step in both the SAGE and microSAGE protocols. *Nucleic Acids Res.* (2000) **28**:e62.

64. BLACKSHAW S, KUO WP, PARK PJ *et al.*: MicroSAGE is highly representative and reproducible but reveals major differences in gene expression among samples obtained from similar tissues. *Genome Biol.* (2003) **4**:R17.

Copyright © Ashley Publications Ltd Reprinting and distribution strictly prohibited

65. YE SQ, ZHANG LQ, ZHENG F, VIRGIL D, KWITEROVICH PO: MiniSAGE: gene expression profiling using serial analysis of gene expression from 1 μg total RNA. *Anal. Biochem.* (2000) **287**:144-152.

66. NEILSON L, ANDALIBI A, KANG D *et al.*: Molecular phenotype of the human oocyte by PCR-SAGE. *Genomics* (2000) **63**:13-24.

67. PETERS DG, KASSAM AB, YONAS H, O'HARE EH, FERRELL RE, BRUFSKY AM: Comprehensive transcript analysis in small quantities of mRNA by SAGE-lite. *Nucleic Acids Res.*(1999) **27**:e39.

68. STANTON J-AL, GREEN DPL: PCR-SAGE and SAGE-Lite. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):577-581.

69. ZHUMABAYEVA B, CHANG C, MCKINLEY J, DIATCHENKO L, SIEBERT PD: Generation of full-length cDNA libraries enriched for differentially expressed genes for functional genomics. *Biotechniques* (2001) **30**:512-520.

70. ZHUMABAYEVA B, DIATCHENKO L, CHENCHIK A, SIEBERT PD: Use of SMART-generated cDNA for gene expression studies in multiple human tumors. *Biotechniques* (2001) **30**:158-163.

71. CHEVAL L, VIRLON B, BILLON E, AUDE JC, ELALOUF JM, DOUCET A: Large-scale analysis of gene expression: methods and application to the kidney. *J. Nephrol.* (2002) **15**:S170-S183.

72. LORKOWSKI S, CULLEN P, ELALOUF J-M: SADE – SAGE adaption for downsized extracts. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P, (Eds), Wiley-VCH, Weinheim, Germany (2002):581-585.

73. VIRLON B, CHEVAL L, BUHLER JM, BILLON E, DOUCET A, ELALOUF JM: Serial microanalysis of renal transcriptomes. *Proc. Natl. Acad. Sci. USA* (1999) **96**:15286-15291.

74. VILAIN C, LIBERT F, VENET D, COSTAGLIOLA S, VASSART G: Small amplified RNA-SAGE: an alternative approach to study transcriptome from limiting amount of mRNA. *Nucleic Acids Res.* (2003) **31**:e24.

75. SPINELLA DG, BERNARDINO AK, REDDING AC *et al.*: Tandem arrayed ligation of expressed sequence tags (TALEST): a new method for generating global gene expression profiles. *Nucleic Acids Res.* (1999) **27**:e22.

76. SPINELLA D: TALEST – Tandem arrayed ligation of expressed sequence tags. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds) Wiley-VCH, Weinheim, Germany (2002):585-592.

77. BUSTIN SA: Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* (2000) **25**:169-193.

78. BUSTIN SA: Real-time reverse transcription polymerase chain reaction. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowksi S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):345-358.

79. BUSTIN SA: Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J. Mol. Endocrinol.* (2002) **29**:23-39.

80. HIGUCHI R, DOLLINGER G, WALSH PS, GRIFFITH R: Simultaneous amplification and detection of specific DNA sequences. *Biotechnology (NY)* (1992) **10**:413-417.

81. HIGUCHI R, FOCKLER C, DOLLINGER G, WATSON R: Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (NY)* (1993) **11**:1026-1030.

82. HOLLAND PM, ABRAMSON RD, WATSON R, GELFAND DH: Detection of specific polymerase chain reaction product by utilizing the 5′–3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. USA* (1991) **88**:7276-7280.

83. LEE CK, KLOPP RG, WEINDRUCH R, PROLLA TA: Gene expression profile of aging and its retardation by caloric restriction. *Science* (1999) **285**:1390-1393.

84. HEID CA, STEVENS J, LIVAK KJ, WILLIAMS PM: Real time quantitative PCR. *Genome Res.* (1996) **6**:986-994.

85. LEVSKY JM, SHENOY SM, PEZO RC, SINGER RH: Single-cell gene expression profiling. *Science* (2002) **297**:836-840.

86. HESS B: Northern blot and related techniques. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):167-173.

87. PROFOUS-JUCHELKA HR, REUBEN RC, MARKS PA, RIFKIND RA: Transcriptional and post-transcriptional regulation of globin gene accumulation in murine erythroleukemia cells. *Mol. Cell. Biol.* (1983) **3**:229-232.

88. LANGER C, SEEDORF U: Nuclear run-on assay. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):173-177.

89. GOPALKRISHNAN RV, SU ZZ, KANG DC *et al.*: Differentiation induction subtraction hybridization. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):178-185.

90. HUANG F, ADELMAN J, JIANG H, GOLDSTEIN NI, FISHER PB: Differentiation induction subtraction hybridization (DISH): a strategy for cloning genes displaying differential expression during growth arrest and terminal differentiation. *Gene* (1999) **236**:125-131.

91. LUO JH, PUC JA, SLOSBERG ED *et al.*: Differential subtraction chain, a method for identifying differences in genomic DNA and mRNA. *Nucleic Acids Res.* (1999) **27**:e24.

92. LUO JH, YU YP: DSC – Differential subtraction chain. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):185-191.

93. PARDINAS JR, COMBATES NJ, PROUTY SM, STENN KS, PARIMOO S: Differential subtraction display: a unified approach for isolation of cDNAs from differentially expressed genes. *Anal. Biochem.* (1998) **257**:161-168.

94. ZENG J, GORSKI RA, HAMER D: Differential cDNA cloning by enzymatic degrading subtraction (EDS). *Nucleic Acids Res.* (1994) **22**:4381-4385.

95. STOLLE K, CULLEN P, LORKOWSKI S: EDS – Enzymatic degrading subtraction. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):192-196.

96. WANG Z, BROWN DD: A gene expression screen. *Proc. Natl. Acad. Sci. USA* (1991) **88**:11505-11509.

97. WANG Z, BROWN DD: GES – Gene expression screen. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):196-201.

98. YANG M, SYTKOWSKI AJ: Cloning differentially expressed genes by linker capture subtraction. *Anal Biochem.* (1996) **237**:109-114.

Copyright of Ashley Publications Ltd. Personal and distribution strictly prohibited

99. WANG Y, SYTKOWSKI AJ: LCS – Linker capture subtraction. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds) Wiley-VCH, Weinheim, Germany (2002):201-205.

100. JIANG H, KANG DC, ALEXANDRE D, FISHER PB: RaSH, a rapid subtraction hybridization approach for identifying and cloning differentially expressed genes. *Proc. Natl. Acad. Sci.* USA (2000) **97**:12684-12689.

101. KANG DC, JIANG H, SU ZZ, VOLSKY DJ, FISHER PB: RaSH – Rapid subtraction hybridization. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P, (Eds) Wiley-VCH, Weinheim, Germany (2002):206-214.

102. HUBANK M, SCHATZ DG: Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic Acids Res*. (1994) **22**:5640-5648.

103. BOWLER LD: cDNA-RDA – Representational difference analysis of cDNA. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):214-219.

104. KANG DC, LAFRANCE R, SU ZZ, FISHER PB: Reciprocal subtraction differential RNA display: an efficient and rapid procedure for isolating differentially expressed gene sequences. *Proc. Natl. Acad. Sci. USA* (1998) **95**:13788-13793.

105. KANG DC, SU ZZ, FISHER PB: RSDD – Reciprocal subtraction differential RNA display. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):219-227.

106. LAVERY DJ, LOPEZ-MOLINA L, FLEURY-OLELA F, SCHIBLER U: Selective amplification via biotin- and restriction-mediated enrichment (SABRE), a novel selective amplification procedure for detection of differentially expressed mRNAs. *Proc. Natl. Acad. Sci. USA* (1997) **94**:6831-6836.

107. BOUCHARDON A, DEHAY C, LAVERY DJ: SABRE – Selective amplification via biotin and restriction-mediated enrichment. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):227-232.

108. BIRKENMEYER LG, LEARY TP, MUERHOFF AS, DAWSON GJ, MUSHAHWAR IK, DESAI SM: Selectively primed adaptive driver RDA (SPAD-RDA): an improved method for subtractive hybridization. *J. Med. Virol.* (2003) **71**:150-159.

109. CUI K, FELDMAN L, SYTKOWSKI AJ: Isolation of differentially expressed genes by cloning transcriptionally active DNA fragments. *Methods* (1999) **17**:265-2671.

110. GERDEMANN C: Subtractive hybridisation of transcriptionally active genomic DNA fragments. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):233-237.

111. FUCHS B, ZHANG K, BOLANDER ME, SARKAR G: Identification of differentially expressed genes by mutually subtracted RNA fingerprinting. *Anal. Biochem.* (2000) **286**:91-98.

112. SHERWOOD A, ZHANG K, BOLANDER ME, SARKAR G: SuRF – Subtracted RNA fingerprinting. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):237-244.

113. DIATCHENKO L, LAU YF, CAMPBELL AP *et al.*: Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. USA* (1996) **93**:6025-6030.

114. HESS B: SSH – Suppressive subtractive hybridization. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):244-249.

115. LACKNER KJ: Multiplexed DNA and RNA analysis on fluorescent microbeads. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds) Wiley-VCH, Weinheim, Germany (2002):249-255.

116. ZINN K, DIMAIO D, MANIATIS T: Identification of two distinct regulatory regions adjacent to the human beta-interferon gene. *Cell* (1983) **34**:865-879.

117. SEEDORF U: RPA – Ribonuclease protection assay. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):255-260.

118. FRANZ O, BRUCHHAUS I I, ROEDER T: Verification of differential gene transcription using virtual northern blotting. *Nucleic Acids Res.* (1999) **27**:e3.

119. BRUCHHAUS I I, ROEDER T: Virtual northern blot. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds) Wiley-VCH, Weinheim (2002):260-266.

120. KATO K: ATAC – Adapter-tagged competitive PCR. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds) Wiley-VCH, Weinheim, Germany (2002): 278-285.

121. SAITO S, MATOBA R, KATO K: Adapter-tagged competitive PCR (ATAC-PCR) – a high-throughput quantitative PCR method for microarray validation. *Methods* (2003) **31**:326-331.

122. LOPEZ-NIETO CE, NIGAM SK: Selective amplification of protein-coding regions of large sets of genes using statistically designed primer sets. *Nat. Biotechnol*. (1996) **14**:857-861.

123. IVANOVA NB, BELYAVSKY AV: Identification of differentially expressed genes by restriction endonuclease-based gene expression fingerprinting. *Nucleic Acids Res.* (1995) **23**:2954-2958.

124. SHMELKOV SV, VISSER JW, BELYAVSKY AV: Two-dimensional gene expression fingerprinting. *Anal. Biochem.* (2001) **290**:26-35.

125. SHMELKOV SV, BELYAVSKY AV: GEF – Gene expression fingerprinting. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):302-308.

126. FUKUMURA R, TAKAHASHI H, SAITO T *et al.*: A sensitive transcriptome analysis method that can detect unknown transcripts. *Nucleic Acids Res*. (2003) **31**:e94.

127. KAWAMOTO S, OHNISHI T, KITA H, CHISAKA O, OKUBO K: Expression profiling by iAFLP: A PCR-based method for genome-wide gene expression profiling. *Genome Res.* (1999) **9**:1305-1312.

128. OKUBO K, KAWAMOTO S: iAFLP – Introduced amplified fragment length polymorphism. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):506-511.

Copyright © Ashley Publications Ltd Printed and distributed. Strictly prohibited

129. RYAN M, STARKEY M, FAULL R, EMSON P, BAHN S: Indexing-based differential display-studies on post-mortem Alzheimer's brains. *Brain Res. Mol. Brain. Res.* (2001) **88**:199-202.

130. STARKEY MP, BAHN S, MAHADEVA H: IBDD – Indexing-based differential display. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):308-312.

131. KATO K: Description of the entire mRNA population by a 3′ end cDNA fragment generated by class IIS restriction enzymes. *Nucleic Acids Res*. (1995) **23**:3685-3690.

132. UNRAU P, DEUGAU KV: Non-cloning amplification of specific DNA fragments from whole genomic DNA digests using DNA 'indexers'. *Gene* (1994) **145**:163-169.

133. KATO K: Molecular indexing of 3′-cDNA ends. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P, (Eds), Wiley-VCH, Weinheim, Germany (2002):312-321.

134. UEMATSU C, NISHIDA J, OKANO K, MIURA F, ITO T, SAKAKI Y, KAMBARA H: Multiplex polymerase chain reaction (PCR) with color-tagged module-shuffling primers for comparing gene expression levels in various cells. *Nucleic Acids Res*. (2001) **29**:E84.

135. UEMATSU C, NISHIDA J, OKANO K *et al.*: MSP-PCR – Multiplex PCR using colour-tagged module-shuffling primers. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):321-326.

136. MACAS J, LAMBERT GM, DOLEZEL D, GALBRAITH DW: Nuclear expressed sequence Tag (NEST) analysis: a novel means to study transcription through amplification of nuclear RNA. *Cytometry* (1998) **33**:460-468.

137. FUCHS B, ZHANG K, BOLANDER ME, SARKAR G: Differential mRNA fingerprinting by preferential amplification of coding sequences. *Gene* (2000) **258**:155-163.

138. SHERWOOD A, ZHANG K, BOLANDER ME, SARKAR G: PACS – Preferential amplification of coding sequences. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*. Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):332-339.

139. BECKER-ANDRÉ M, HAHLBROCK K: Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcript titration assay (PATTY). *Nucleic Acids Res.* (1989) **17**:9437-9446.

140. HINOJOS CA, DORIS PA: Competitive RT-PCR. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):296-302.

141. VOLKENANDT M, DICKER AP, BANERJEE D *et al.*: Quantitation of gene copy number and mRNA using the polymerase chain reaction. *P. Soc. Exp. Biol. Med.* (1992) **200**:1-6.

142. HORSTMANN S: RAP-PCR – RNA arbitrarily primed PCR fingerprinting. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):340-345.

143. WELSH J, CHADA K, DALAL SS, CHENG R, RALPH D, MCCLELLAND M: Arbitrarily primed PCR fingerprinting of RNA. *Nucleic Acids Res.* (1992) **20**:4965-4970.

144. GERDEMANN C: RD-PCR – Restriction display polymerase chain reaction. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):358-363.

145. GRAVESEN A, WARTHOE P, KNOCHEL S, THIRSTRUP K: Restriction fragment differential display of pediocin-resistant Listeria monocytogenes 412 mutants shows consistent overexpression of a putative β-glucoside-specific PTS system. *Microbiology* (2000) **146**:1381-1389.

146. FISCHER A, SAEDLER H, THEISSEN G: Restriction fragment length polymorphism-coupled domain-directed differential display: a highly efficient technique for expression analysis of multigene families. *Proc. Natl. Acad. Sci. USA* (1995) **92**:5331-5335.

147. THEISSEN G, FISCHER A: RC4D-restriction fragment length polymorphism-coupled domain-directed differential display. *Methods Mol. Biol*. (1997) **85**:123-133.

148. THEILGAARD-MÖNCH K, COWLAND J, BORREGAARD N: Profiling of gene expression in individual hematopoietic cells by global mRNA amplification and slot blot analysis. *J. Immunol. Methods* (2001) **252**:175-189.

149. THEILGAARD-MÖNCH K, RASMUSSEN T: SIP RT-PCR. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):363-370.

150. BROWN AJ, HUTCHINGS C, BURKE JF, MAYNE LV: Application of a rapid method (targeted display) for the identification of differentially expressed mRNAs following NGF-induced neuronal differentiation in PC12 cells. *Mol. Cell. Neurosci.* (1999) **13**:119-130.

151. SCHNOOR M, CULLEN P, LORKOWSKI S. Targeted display. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim, Germany (2002):363-370.

152. HENNIG S, PANOPOULOU G, LEHRACH H, POUSTKA A: Comparative expressed sequence tag analysis. In: *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls.* Lorkowski S, Cullen P (Eds), Wiley-VCH, Weinheim (2002):770-778.

153. DE SOUZA SJ, CAMARGO AA, BRIONES *et al.*: Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags. *Proc. Natl. Acad. Sci. USA* (2000) **97**:12690-12693.

154. CAMARGO AA, SAMAIA HP, DIAS-NETO SANTOS ML *et al.*: The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome. *Proc. Natl. Acad. Sci. USA* (2001) **98**:12103-12108.

155. HOF P, ORTMEIER C, PAPE K *et al.*: Digital analysis of cDNA abundance; expression profiling by means of restriction fragment fingerprinting. *BMC Genomics* (2002) **3**:7.

156. SUZUKI H, YAOI T, KAWAI J, HARA A, KUWAJIMA G, WATANABE S: Restriction landmark cDNA scanning (RLCS): a novel cDNA display system using two-dimensional gel electrophoresis. *Nucleic Acids Res*. (1996) **24**:289-294.

157. CHENCHIK AA, DIACHENKO LB, BEABEALASHVILLI RSH: Application of poly(A)$^+$ RNA patterns method for searching of differentially expressed genes. *FEBS Lett.* (1993) **324**:136-139.

Copyright of Ashley Publications Ltd (Printed) and distribution of this is prohibited

Patents

201. ROCHE DIAGNOSTICS GMBH: US6090548 (2000).
202. ROCHE DIAGNOSTICS GMBH: WO9743443 (1997).
203. MED. RES. COUNCIL (GB): WO8702133 (1987).
204. UNIV. OXFORD: US4886589 (1989).
205. BECKMAN COULTER, INC.: US5429807 (1995).
206. BECKMAN COULTER, INC.: WO9511748 (1995).
207. THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIV.: US5807522 (1998).
208. LELAND STANFORD JUNIOR UNIV.: WO9535505 (1995).
209. SYNGENTA PARTICIPATIONS AG: WO03076566 (2003).
210. THIRD WAVE TECH., INC.: WO0190337 (2001).
211. PORTELA & CA SA: WO02090572 (2002).
212. TROPIX, INC.: WO0060116 (2000).
213. HIGH THROUGHPUT GENOMICS, INC.: US6238869 (2001).
214. HIGH THROUGHPUT GENOMICS, INC.: WO03002750 (2003).
215. DANA-FARBER CANCER INST., INC.; THE PRESIDENT AND FELLOWS OF HARVARD COLLEGE: US5599672 (1997).
216. FOX CHASE CANCER CENTER: WO02090564 (2002).
217. NOVARTIS AG: DE19806431 (1999).
218. HOFFMANN R, ZWILLING S, LUEBBERT H *ET AL.*: WO9942610 (1999).
219. ROTHBERG JM, MCKENNA M, PREDKI P, WINDMUTH A, SHIMKETS RA: US2002090612 (2002).
220. CURAGEN CORP.: WO0040757 (2002).
221. YALE UNIV.: US5712126 (1998).
222. YALE UNIV: WO9705286 (1997).
223. DIGITAL GENE TECH., INC.: US2002012922 (2002).
224. DIGITAL GENE TECH., INC.: WO02061045 (2002).
225. LYNX THERAPEUTICS, INC.: WO9746704 (1997).
226. LYNX THERAPEUTICS, INC.: US6013445 (2000).
227. LYNX THERAPEUTICS, INC.: US6265163 (2001).
228. LYNX THERAPEUTICS, INC.: WO9935293 (2001).
229. LYNX THERAPEUTICS, INC.: US6511802 (2003).
230. SPECTRAGEN, INC.: US5552278 (1996).
231. SPECTRAGEN, INC.: US5599675 (1997).
232. SPECTRAGEN, INC.: US5654413 (1997).
233. SPECTRAGEN, INC.: US5635400 (1997).
234. SPECTRAGEN, INC.: US5604097 (1997).
235. LYNX THERAPEUTICS, INC.: US5714330 (1998).
236. LYNX THERAPEUTICS, INC.: US5888737 (1999).
237. SPECTRAGEN, INC.: WO9641011 (1996).
238. LYNX THERAPEUTICS, INC.: WO0020639 (2000).
239. LYNX THERAPEUTICS, INC.: US5846719 (1998).
240. LYNX THERAPEUTICS, INC.: US6150516 (2000).
241. LYNX THERAPEUTICS, INC.: US6235475 (2001).
242. LYNX THERAPEUTICS, INC.: WO9846621 (1998).
243. LYNX THERAPEUTICS, INC.: US5888737 (1999).
244. LYNX THERAPEUTICS, INC.: WO0026411 (2000).
245. LYNX THERAPEUTICS, INC.: US6175002 (2001).
246. LYNX THERAPEUTICS, INC.: WO0227029 (2002).
247. LYNX THERAPEUTICS, INC.: US2003027157 (2003).
248. LYNX THERAPEUTICS, INC.: WO9853300 (1998).
249. KINZLER KW, VOGELSTEIN B, VELCULESCU VE, ZHANG L: US5695937 (1997).
250. JOHNS HOPKINS UNIV. SCHOOL OF MEDICINE: WO9710363 (1997).
251. JOHNS HOPKINS UNIV. SCHOOL OF MEDICINE: US5866330 (1999).
252. UNIV. JOHNS HOPKINS: US6383743 (2002).
253. UNIV. JOHNS HOPKINS: US2003049653 (2003).
254. UNIV. JOHNS HOPKINS: US6498013 (2002).
255. UNIV. JOHNS HOPKINS: WO0210438 (2002).
256. UNIV. JOHN HOPKINS: US2003008290 (2003).
257. UNIV. TEXAS: US6221600 (2001).
258. UNIV. TEXAS SYSTEMS: WO0127329 (2001).
259. CLONTECH LABS, INC.: US5962271 (1999).
260. CLONTECH LABS, INC.: US5962272 (1999).
261. CHEVAL L, ELALOUF J-M, VIRLON B: US6506561 (2000).
262. COMMISSARIAT ENERGIE ATOMIQUE; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE: WO0044936 (2000).
263. COMMISSARIAT ENERGIE ATOMIQUE: US2003165910 (2003).
264. CHUGAI PHARM. CO. LTD: WO0053806 (2000).
265. TRUSTEES OF COLUMBIA UNIV.: WO9829570 (1998).
266. JOHNSON & JOHNSON CONSUMER PRODUCTS, INC.: WO9737045 (1997).
267. JOHNSON & JOHNSON CONSUMER PRODUCTS, INC.: US6045998 (2000).
268. BETH ISRAEL DEACONESS MEDICAL CENTER, INC.: US5804382 (1998).
269. COLUMBIA UNIV.: WO0231205 (2002).
270. ABBOTT LABS: WO0210458 (2002).
271. CLONTECH LABS, INC.: US5759822 (1996).
272. CLONTECH LABS, INC.: WO9623079 (1996).
273. CLONTECH LABS, INC.: US5759822 (1998).
274. LUMINEX CORP.: WO0114589 (2001).
275. LUMINEX CORP.: US6649414 (2002).
276. KATO K; UNITECH CO. LTD: US6248534 (2001)
277. JAPAN SCI. AND TECH. CORP.: EP-0870842 (1998).
278. JAPAN SCI. AND TECH. CORP.: US6090556 (2000).

Copyright of Ashley Publications Ltd Printing and distribution strictly prohibited

279. BRIGHAM AND WOMEN'S HOSPITAL; LOPEZ-NIETO CE: WO9531574 (1995).

280. BRIGHAM AND WOMEN'S HOSPITAL: EP-0760009 (1995).

281. BRIGHAM AND WOMEN'S HOSPITAL: US6110667 (2000).

282. NEW YORK BLOOD CENTER, INC.: US5814445 (1998).

283. NEW YORK BLOOD CENTER, INC.: US6120996 (2000).

284. NEW YORK BLOOD CENTER, INC.: US2003017490 (2003).

285. MAZE, INC.: WO0248352 (2002).

286. RES. DEV. CORP. OF JAPAN: US5707807 (1998).

287. RES. DEV. CORP. OF JAPAN: EP-0735144 (2003).

288. HITACHI LTD: US6225064 (2001).

289. UEMATSU C, KAMBARA H, OKANO K: US2002098490 (2002).

290. HITACHI LTD: US6562572 (2003).

291. STRATAGENE: WO9207095 (1992).

292. STRATAGENE; CALIFORNIA INST. OF BIOLOGICAL RES.: US5861245 (1999).

293. US HEALTH (US); WEINSTEIN J; BOULAMWINI J: WO9729211 (1997).

294. LUDWIG INST. FOR CANCER RES.: WO02074994 (2002).

295. SWITCH BIOTECH. AG: US6303308 (2001).

296. SUZUKI H, YAOI T, WATANABE S: WO9717466 (1997).

297. SHIONOGI & CO.: US5665544 (1993).

Websites

301. http://cgap.nci.nih.gov/ Cancer Genome Anatomy Project (CGAP) at the National Cancer Institute

302. http://www.ncbi.nlm.nih.gov/SAGE SAGE database at National Center for Biotechnology Information

Affiliation

Stefan Lorkowski[†1,2] and Paul Cullen[1,3]
[†]Author for correspondence
[1]Institute of Arteriosclerosis Research, University of Münster, Domagkstraße 3, 48149 Münster, Germany
Tel: +49 251 835 2072;
Fax: +49 251 835 2062;
E-mail: stefan.lorkowski@uni-muenster.de
[2]Institute of Biochemistry, Wilhelm-Klemm-Straße 2, 48149 Münster, Germany
[3]Ogham GmbH, Mendelstraße 11, 48149 Münster, Germany

Copyright of Ashley Publications Ltd Printing and distribution strictly prohibited