Gene-expression analysis at the single-cell level

Alistair K. Dixon, Peter J. Richardson, Robert D. Pinnock and Kevin Lee

The manner in which a cell responds to and influences its environment is ultimately determined by the genes that it expresses. To fully understand and manipulate cellular function, identification of these expressed genes is essential. Techniques such as RT-PCR enable examination of gene expression at the tissue level. However, the study of complex heterogeneous tissue, such as the CNS or immune system, requires gene analysis to be performed at much higher resolution. In this article, the various methods that have been developed to enable RT-PCR to be performed at the level of the single cell are reviewed. In addition, how, when carried out in combination with techniques such as patch-clamp recording, single-cell gene-expression studies extend our understanding of biological systems is discussed.

The human genome encodes ~100 000 different genes and it is anticipated that sequence information for the majority of these will be available early next year1–3. However, knowledge of a gene sequence is only the first step to understanding the role of any given gene and its product. It is also necessary to determine the function of the gene product, the mechanisms regulating where and when it is expressed, and any potential modifications that occur during RNA and protein processing. Of fundamental importance are the interactions between the thousands of gene products that determine cell phenotypes. Until recent times it has only been possible to study the expression of small numbers of genes in any given investigation, but significant advances in genomics and proteomics research now permit the co-expression of thousands of genes to be analysed in a given tissue4. This has provided the tools with which to link tissue function with specific patterns, and levels, of gene expression.

A variety of approaches are routinely used to assay the expression of specific genes in cells and tissues (northern blot and reverse-transcriptase polymerase chain reaction (RT-PCR) for mRNA, electrophoresis and western blot, and mass spectroscopy, among others, for proteins). All such techniques used to analyse gene expression are limited when applied to heterogeneous tissues in which it can be difficult to assign the expression of a given gene to a specific cell type. Obviously in heterogeneous tissues (e.g. CNS and immune system) this problem is compounded. In such tissues, it is important to identify and classify each cell group within a tissue before assessing gene expression in the different groups. To date, the identification of cells and their classification into specific groups has been based primarily on the measurement of single parameters such as: (1) the location of the cell (anatomy); (2) the shape of the cell (morphology); (3) the proteins that the cell expresses (immunohistochemistry); (4) the genes that the cell expresses (in situ hybridization); or (5) a functional approach such as electrophysiology is used. However, cells that are identical both anatomically and morphologically might have divergent properties and as a result be functionally distinct. Immunohistochemistry and in situ hybridization are more useful techniques because they permit the study of the expression of proteins or mRNAs in thousands of cells simultaneously, usually in tissue sections that have retained many aspects of their original morphology. However, classification of cells based on these techniques might be a result of only a single or a small number of proteins or mRNAs being routinely used as markers. Because the total population of mRNAs expressed in a cell is 6000–10 000 (Ref. 4), these techniques do not determine the potential underlying genetic diversity of cells. Thus, the variety of cellular identities is probably far greater than that suggested from results of these approaches. A typical functional approach is the use of specific receptor agonists and antagonists to determine the expression of a selected population of receptor or ion channels in conjunction with single-cell electrophysiological recording. This technique is obviously restricted by the pharmacological tools available and by the current understanding of the receptor or ion channel family being studied.

Despite the above caveats, useful classifications of cells in complex tissues have been made using both molecular and functional approaches. As suggested, these classifications might underestimate cellular diversity but their major flaw is that it is difficult to identify cells within groups whose functions have changed as a result of previous activity (e.g. long-term potentiation (LTP) and long-term depression (LTD) in the CNS). More importantly, these approaches are not suitable for detecting changes in gene expression that are fundamental to such activity-based changes in cellular activity because the repertoire of genes whose expression might change is very large – much larger than the immediate early gene family (e.g. Fos and Jun) that is often studied. In addition, genes do not act in isolation (i.e. the properties of the cells are the result of the combined interactive action of many gene products), so techniques that can assess the expression of many genes and their relationship to function in single cells is required. It will only then be possible to begin to relate cellular physiology to gene expression in a meaningful way.

Accessing genetic material

The first hurdle to analysing gene expression in single cells involves accessing the genetic material, usually the cytoplasmic mRNA, from the cells of interest before the application of the amplification protocol. Patch-clamp techniques provide the ideal method by which to couple functional properties, such as neuronal excitability, and molecular studies at the level of the single cell. Standard electrophysiological analysis could be performed, on completion of which the patch-pipette forms a gateway for the harvesting of cytoplasmic contents1–3 (Fig. 1).
The application of gentle negative pressure results in the collection of the cytoplasm (b) – February 2000 (Vol. 21)

Alternatively, single cells can be obtained after tissue dissection or from blood. The simplest way of isolating such single cells is to visualize them in micro-drops under a phase-contrast microscope and then to pick them up using a conventional pipette. Access to the cytoplasmic contents can be gained using a detergent that selectively lyses the plasma membrane, leaving the nuclear membrane intact. The nucleus, and all nuclear genetic material, can then be partitioned by centrifugation.13 A more elegant method of cell isolation employs the technology of fluorescence-activated cell sorting (FACS) with flow cytometry. Such an approach provides a fast, accurate way of isolating single or any number of cells. Linkage to function can also be made before genetic analysis. Using fluorescent probes that are selective for specific cell properties, selection can be made in relation to cell cycle, membrane potential or the expression of surface or cytoplasmic markers.

A more ‘hands-on’ approach might be necessary when dealing with certain specimens, particularly when obtaining material from stained histological slides. This could involve laser disruption of potentially contaminating tissue before the use of a micromanipulator to physically remove the target cell.14 This operation could be made more accurate and high throughput by the use of a laser capture microdissection system.15 One consideration when using routinely processed archival tissue is the degree of nucleic acid degradation that might have taken place; however, such studies have proven to be feasible.16

Fig. 1. Harvesting of cell contents from a neurone situated within a 300-μm brain slice. (a) The patchclamp electrode has been placed on the surface of the neurone and the whole-cell recording configuration is formed. (b) The application of gentle negative pressure results in the collection of the cytoplasm and importantly within spaces therein, into the patch pipette. (c) The electrode has been withdrawn from the neurone and taken out of the brain slice with subsequent formation of a nucleated patch, thus preserving tissue or contamination of the sample. (d) Withdrawal of the nucleated patch through the air–electrolyte interface results in the loss of the nucleated patch integrity. Scale bar: 10 μm.

Approaches to facilitate gene-expression analysis at the level of the single cell

The quantity of mRNA that can be harvested from a single cell is in the order of 1 pg at best.17 Therefore, to obtain meaningful gene-expression data, well optimized or specialized amplification protocols must be applied. Using conventional PCR, the theoretical limit of detection is one copy of a single-stranded DNA molecule and so with efficient harvesting of cytoplasm and a well-optimized PCR assay, single-cell PCR is feasible.18 However, this is not a trivial undertaking and rare or particularly labile transcripts would prove to be a technically demanding procedure for many. The scope of such studies is limited by the fact that only one or two genes can be analysed at any one time. Therefore, certain modifications to this approach have been applied to provide more comprehensive single-cell expression profiles.

Nested PCR

A straightforward method of expanding the results obtained from single-cell PCR involves the use of ‘nested’ PCR primers (Fig. 2a). Essentially, a primary conventional PCR between conserved regions of gene family members increases the target concentration such that a second PCR reaction can be carried out to assay for the presence or absence of genes within that group. This approach has the advantage of being relatively simple to apply – the overall protocol deviates little from conventional RT-PCR – and as such is of obvious use when applied to the analysis of sets and subsets of homologous genes. However, in practice the development of a working protocol might not be entirely straightforward. In particular, the design of multiple primer pairs that work efficiently and specifically in concert can sometimes prove difficult. In addition, when applying an approach that involves two rounds of exponential amplification great care must be taken to ensure that contamination is not carried over giving rise to false positives. However, in most instances the main drawback of the nested approach is that only an extremely small group of predeter-

Homomorphic tailing

The only way of surmounting these problems is through the use of techniques that result in the global amplification of the mRNA population possessed by a single cell. One approach to achieving this goal involves the combination of a tailing reaction and subsequent universal PCR amplification.18 However, to date, this method has not been widely used to link functional phenotype to genotype. This might well be a result of the specificity problems inherent in the use of repetitive homomeric primers for PCR. In addition, this method will tend to result in the preferential amplification of genes with conserved regions of gene family members increases the target concentration such that a second PCR reaction can be carried out to assay for the presence or absence of genes within that group. This approach has the advantage of being relatively simple to apply – the overall protocol deviates little from conventional RT-PCR – and as such is of obvious use when applied to the analysis of sets and subsets of homologous genes. However, in most instances the main drawback of the nested approach is that only an extremely small group of predeter-

Three-prime-end amplification PCR

More recently, a PCR-based method has been developed that circumvents some of these problems. Three-prime-end amplification (TPEA) PCR results in the global amplification of the most 3’ end of all mRNAs present in a single cell (Fig. 2c). Again, conventional gene-specific PCR can now be used to assay for the presence or absence of any gene, at present, the expression of up to 40 genes has been assessed.19 The primary advantage of this technique is the relative ease with which
Fig. 2. Single-cell gene-expression methodologies. (a) In nested PCR, primers that are homologous to conserved regions of closely related genes (in red) are used in a primary PCR reaction to amplify target DNA to the level at which the individual gene family members can be assayed in a secondary gene-specific PCR reaction.

(b) Global amplification by homomeric tailing. All mRNA is converted to first strand cDNA by reverse transcription with an oligo d(T) primer, thus targeting the 3′ poly(A) tract unique to expressed genes. The enzyme terminal deoxynucleotidyl transferase (TdT) is used to link a homomeric poly d(A) tail to the cDNA produced. This provides a priming site for PCR amplification, using a single oligo d(T) primer until sufficient cDNA has been generated for gene-specific PCR.

(c) In three-prime-end amplification (TPEA) PCR, a unique sequence is incorporated at the 3′ end of the first strand cDNA during reverse transcription. During second-strand synthesis, a second unique sequence is incorporated as part of the second-strand primer, which is designed to anneal once every kb. This provides a set of amplified 3′ ends, uniform in size, which represent the original population of expressed genes within the cell. Gene-specific PCR can then be performed.

(d) In RNA amplification, the internal solution of the patch-pipette contains the reagents to initiate first-strand cDNA synthesis. This reaction is primed with an oligo d(T) primer that contains the T7 RNA polymerase promoter sequence. Following production of double-stranded cDNA, complementary RNA is transcribed providing a 100–1000 fold amplification. After reverse transcription of the RNA produced, further amplification can be achieved by performing another round of RNA transcription. Gene detection can now be performed using PCR with specific primers.
RAR REVIEW

The overall protocol can be performed. Additionally, the global amplification that this technique achieves would appear to provide access to all gene-expression information within a single cell. To date, only 40 genes have been analysed, however this is a limitation of the reads rather than the pre-amplification protocol. Indeed, with the application of cDNA-array technology to assay TPEA-amplified products it is possible to envisage that the expression of many thousands of genes could be analysed from a single cell. However, because this approach is PCR based, the data obtained are not quantitative. In addition, because the amplification is biased towards the 3' end, a certain amount of 5' information is lost, which might hamper the detection of certain splice variants.

RNA amplification

A non-PCR-based approach, involving RNA amplification2, has been developed for the analysis of any expressed gene in a single cell (Fig. 3). This process could commence within the cell if a patch-pipette is used to gain access. Following production of double-stranded cDNA, complementary RNA is transcribed providing a 100–1000 fold amplification. After reverse transcription of the RNA so produced, further amplification can be achieved by performing another round of RNA transciption. Our advantage of not applying exponential PCR, amplification is such that a linear quantity of RNA can be amplified. That is, the number of copies of the RNA is doubled with each round of the reaction, has been used to demonstrate that sensitivity to neurosteroid modulation of GABA receptors22.

Applications of single-cell PCR

Receptor pharmacology

Single-cell PCR increasingly provides receptor information that is inaccessible to more traditional pharmacological approaches. For example, studies on ionotropic glutamate receptor expression have been restricted by the lack of suitable ligands, thus limiting understanding of function in this area. However, the use of multiplex PCR, in which several primer pairs that target several genes are combined in a single PCR reaction, has been used to demonstrate that sensitivity to nikinate in cultured hippocampal neurons correlates with GluK expression, whereas GluR5 and GluR7 RNAs were not detected18. Similarly, the use of ‘nested’ primers at the single-cell level has been applied to correlate AMPA-receptor subtype expression to electrophysiological characteristics in single parvalbumin cells20. This particular approach has also been applied to determine the subunit composition that gives rise to the functional diversity of glutamate receptors20 and to decoding the complexities of the mammalian olfactory system21.

The specificity afforded by single-cell RT-PCR has proven to be useful when studying complex heteropolymers. For example, it has been known for some time that neuroactive steroids activate GABA receptors directly and single-cell RT-PCR has been employed to demonstrate that the presence of the β-subunit has an inhibitory effect on neurosteroid modulation of GABA receptors21.

A direct correlation between functional and molecular observations has also been made in the study of nicotinic acetylcholine (ACh) receptors. Whole-cell recording and single-cell RT-PCR were combined to investigate the effect of nicotinic agonists on different types of neurons in the rat neocortex22. Pharmacological analysis suggested that a subpopulation of interneurons that were responsive to nicotinic receptor agonists lacked the α7-subunit. Single-cell PCR studies subsequently confirmed that the responsive interneurons possessed nicotinic receptors that comprised α4-, α5- and β2-subunits, thereby unravelling the molecular architecture of the neuronal regulation of cortical flow mediated by ACh.

More recently, the global amplification achieved by techniques such as TPEA and RNA amplification have been applied to address the nature of other types of surface receptors. The use of TPEA to analyse the expression of specific...
adrenosine and tachykinin receptors has helped to explain the complex physiology of the striatal ACh-containing interneurones (Fig. 3). This technique has also been used to determine the identity of functional bombesin receptors present in hippocampal interneurones (5). Similarly, RNA amplification has been used to re-direct neuronal complexity, demonstrating the co-expression of multiple dopamine receptor subtypes in nigrostriatal neurones and the widespread expression of glutamate decarboxylase, showing its expression in glutamate- and GABA-containing neurones (6,7).

**Ion channel pharmacology**

As in the study of cell-surface receptors, single-cell RT-PCR has proven a useful tool for the accurate identification of ion channel subunits involved in a wide range of physiological responses. Once again when dealing with multi-component molecules, and in the absence of sufficient pharmacological tools to dissect individual subunits, such approaches are invaluable. TPEA PCR has been used to analyse the heteromeric composition of ATP-sensitive K<- channels (KATP) in both the striatal ACh-containing interneurones and glucose-responsive neurons of the ventromedial hypothalamus (8,9). A single-cell multiplex PCR approach has also been used to investigate the molecular machinery responsible for the production of such channels in rat hippocampal neurones (10). These studies have been able to unequivocally determine which channel subunits (Kir6.1 and Kir6.2) and sulphonylurea receptor (SUR1 and SUR2) might or might not be present in KATP channels in phenotypically distinct neurones. Similarly, in the study of the different pharmacological properties of the delayed rectifier current in hippocampal interneurones and pyramidal cells, single-cell PCR has been used to establish that Kv3.1 and Kv3.2 give rise to this current in interneurones, whereas the current is generated by Kv4.1- and Kv4.2-containing channels (11). Similar approaches have also been applied to the molecular characterisation of Na<- channels in cerebellar purkinje cells and Ca<- channels in motoneurones of the rat facial nucleus (12). RNA amplification at the single-cell level has also proved useful and has been used to examine the effect of corticosterone exposure on the expression of genes contributing to Ca<sup>2+</sup>-induced and Ca<sup>2+</sup>-mediated synaptic transmission in CA1 pyramidal neurones (13).

**Neuropil plasticity and development**

The technique of RNA amplification at the single-cell level has provided insight into the molecular mechanisms of neuronal development and plasticity. To date, many investigations of temporal and spatial gene expression during embryogenesis have focused solely on the analysis of isolated single gene events. However, when the technique of RNA amplification was applied to neural tubes dissected from the developing mouse, the gene expression of several genes in single neuronal dendrites and dendritic growth cones has been used to track and detect the expression of several genes in single neuronal dendrites and dendritic growth cones (14,15).

**Data interpretation**

Negative results obtained in experiments designed to detect gene expression are essentially meaningless using most techniques. However, because PCR allows single copies to be detected, it will be possible, using appropriate controls, to show a true negative within a cell (i.e. if a single copy can be detected in a given gene-specific assay and negative results are still obtained, it can be stated that this gene was not expressed within the cell at the time of sampling). All techniques that are designed to detect mRNA suffer from the potential drawback that mRNA expression does not necessarily mean protein synthesis. Resolution of this debate awaits increases in the sensitivity of proteomics, although it should be noted that the single-cell approaches outlined above have linked the presence of specific mRNAs to specific cell functions. One disadvantage of the approaches that have been used thus far is that they are limited to known gene sequences. Although this will not be a problem in the near future when studying human cells, this might be a severe drawback in some experimental systems. Another potential problem is that as the sensitivity of single-cell PCR increases, it might become necessary to perform hundreds of gene-specific PCRs on each cell extract if the information gained is to be maximised. To overcome this, it might be advantageous to link to hybridisation-based technologies (e.g. slot blots or cDNA micro-arrays). For example, using slot blots single-cell RT-PCR has been used to correlate functional changes in GABA<sub>A</sub> receptor physiology to gene-expression patterns in temporal kobe epilepsy (16).

**Concluding remarks and future prospects**

In recent years, the use of single-cell RT-PCR has proved a useful tool to aid pharmacological investigation. As the technology improves and our ability to relate the expression of many genes to function in single cells increases, one can anticipate that this technology will become even more invaluable to the pharmacologist. As the sensitivity of these techniques improves it is likely that it will soon be possible to assess the expression of all known genes within a genome in samples derived from individual cells.

**Selected references**

5-HT7 receptors: current knowledge and future prospects

Peter Vanhoenecker, Guy Haegeman and Josee E. Leysen

Identification of three splice variants of the 5-HT7 receptor suggests a possible diversity in 5-HT receptor action. Indeed, 5-HT7 receptors have been implicated in the pathophysiology of several disorders; they play a role in smooth muscle relaxation within the vasculature and in the gastrointestinal tract. However, most of these assignments are derived from receptor localization studies and investigations using nonselective ligands, and are therefore mainly suggestive. The development of selective 5-HT7 receptor antagonists will be of utmost importance in determining the actual physiological and pharmacological roles of this receptor. Major challenges of 5-HT7 receptor research are determination of the transcriptional regulation of the gene encoding the 5-HT7 receptor and elucidation of the differences in regulation and signalling of its four gene products.

5-HT7 is implicated in a large variety of behavioural and physiological processes1. Molecular cloning studies revealed the existence of 14 different genes, each encoding a distinct 5-HT receptor subtype2. Although all these receptor subtypes respond differently to the tryptophan metabolite, an even higher level of diversity exists because isoforms have been identified for several of these receptor subtypes. Isoforms are either created by RNA editing, as is the case for the 5-HT2C receptor, or by alternative splicing3.

The 5-HT7 receptor is the most recently identified member of the family of G-protein-coupled 5-HT7 receptors. Three different splice variants have been described, all with a putative intracellular and human tissues. The fact that these variants differ in the length of their C-terminus and in the number of phosphorylation