Chem. Senses 27: 289–291, 2002

# High-throughput Expression Profiling Techniques

**Timothy S. McClintock, Organizer**

Department of Physiology, University of Kentucky, Lexington, KY 40536-0298, USA

*Correspondence to be sent to: Timothy S. McClintock, Department of Physiology, University of Kentucky, 800 Rose St, Lexington, KY 40536-0298, USA. e-mail: mcclint@uky.edu*

Great optimism exists that correlating genomics with function will lead to a better understanding of the detailed workings of the nervous system. The nearly complete sequencing of the genomes of several eukaryotic species and the invention of high-throughput expression profiling techniques now provide the means to rapidly investigate the molecular underpinnings of phenotypic change with reasonable accuracy. This approach is beginning to have an impact in neurobiology and we can expect it to be similarly useful for investigating problems in the chemical senses. Careful thought is necessary, however, in selecting and matching specific techniques and tissues, in confirming differences in mRNA abundance and in interpreting the results. A symposium on Functional Genomics in Neural Systems was held during the AChemS XXIII meeting to illustrate both the potential and the limitations of these techniques for investigating questions of importance to neurobiology in general and to the chemical senses in particular. A brief technical introduction, summarized herein, was followed by two keynote talks. Dr Daniel Geschwind (UCLA) shared his work identifying genes whose expression is correlated with differentiating neurons in developing neural tissues. Moving then to the other end of a nerve cell's life, Dr Tomas Prolla (University of Wisconsin) described different sets of genes that are affected in the aged brain and how many of the age-related changes can be prevented by restricting the intake of calories. Both speakers incorporated their experiences and perspectives on the technical challenges that are a critical element in the successful use of these new and still improving techniques.

Expression profiling techniques allow the simultaneous analysis of the abundance of many thousands of transcripts. By doing so, they provide three advantages. (i) They accelerate the discovery of transcripts whose abundance is correlated with any particular phenotype. (ii) Even more exciting than discoveries about individual transcripts, however, is the ability to observe patterns that emerge from comparing the known functions of all the affected transcripts. (iii) Finally, these techniques can provide hypothesis testing about gene expression patterns underlying phenotypic change on a broad scale. Clearly, expression profiling techniques provide powerful tools.

The tools available for expression profiling can be grouped into three types (Table 1). (i) Clone and count methods generate thousands of 9–20 bp sequences representing specific sites near the 3′ ends of poly(A+) RNAs (Velculescu *et al*., 1995; Brenner *et al*., 2000). Sequences of this length are sufficient to uniquely identify most transcripts. The abundance of a transcript is measured simply as the number of times its sequence tag is encountered. Overall, this is a highly effective method, but laborious and expensive enough that it is less common than the other two types. (ii) Microarray methods depend upon hybridization of probes derived from RNA samples against DNA or RNA spots bound to a solid substrate (Schena *et al*., 1995). The commercial production of microarrays, the familiarity of investigators with nucleic acid hybridization methods and the rapidity of data collection have made this the most popular type of expression profiling method. Its limitations include the obvious fact that it is restricted to analyzing only those sequences present on the array and that it is slightly less sensitive than the other methods. The paper by Prolla (Prolla, 2002) that follows describes the details of expression profiling with Affymetrix GeneChip oligonucleotide arrays. Numerous reviews describing the details of, and differences among, microarray methods are available (Duggan *et al*., 1999; Lipshutz *et al*., 1999; Hegde *et al*., 2000; Luo and Geschwind, 2001). (iii) Differential subtraction methods use the power of PCR, adapting its conditions to favor amplification of differentially abundant cDNAs. The first invented was differential display (Liang and Pardee, 1992). More recently devised techniques, such as representational difference analysis (RDA) of cDNA (Hubank and Schatz, 1994) and suppressive subtractive hybridization (Diatchenko *et al*., 1996), share its advantages and appear to be more robust and accurate. For example, RDA of cDNA rarely yields false positives, is sequence-independent, interrogates the majority of transcripts simultaneously and can be used with very small amounts of tissue. Not that it lacks limitations, however. Analysis of the products is labor-intensive and its dynamic range can be limited when many differences are present (though this can be at least partially overcome by repeating the procedure). For all three types of expression profiling methods there exists one major drawback—achieving the number of repetitions required for standard statistical analyses is often impossible or im-

**Table 1** The three types of expression profiling techniques

| | Capacity | Sensitivity | Drawbacks |
|---|---|---|---|
| Clone/count | | | |
| MPSS | approx. all mRNAs | ~1 copy per cell | not widely accessible |
| SAGE | approx. all mRNAs | ~1 copy per cell | high cost, labor intensive |
| Microarrays | | | |
| GeneChip | 1–60 000 | >5 copies per cell | high cost, access to equipment, capacity |
| Glass slides | 1–15 000 | >10 copies per cell (fluorescence, better with radioactivity) | capacity, sensitivity |
| Membranes | 1–15 000 | <5 copies per cell | capacity, sensitivity |
| Differential subtraction | | | |
| RDA | approx. all mRNAs | ~1 copy per cell | labor intensive, slow, limited range in cases of large numbers of differences (can be partially overcome) |

Some examples of each type are listed. Three categories list properties that are useful in deciding which techniques best apply to an experimental situation. Capacity refers to the number of transcripts that are interrogated with each technique. Sensitivity estimates are based on reported or expected results using the standard protocols of each technique. Drawbacks listed are relative to other expression profiling techniques. MPSS, massively parallel signature sequencing (Brenner *et al*., 2000); SAGE, serial analysis of gene expression (Velculescu *et al*., 1995); RDA, representational difference analysis of cDNA (Hubank and Schatz, 1994).

practical. Empirically defined criterion levels often correctly identify differentially abundant transcripts, but are increasingly regarded as unsatisfactory for the obvious reason that the absence of variance estimation leaves in doubt the risk of type I and type II errors. The paper by Prolla (Prolla, 2002) discusses several important statistical issues in the use of microarrays. Validation of expression profiling data by independent techniques such as Northern blot, RNase protection assay, RNA dot blot, etc. is an increasingly common solution to the statistical problem. For example, the paper that follows by Dougherty and Geschwind (Dougherty and Geschwind, 2002) describes the use of custom cDNA microarrays to screen the results of an RDA of cDNA experiment.

In some situations, choosing an expression profiling technique is easy since any would suffice. However, in most cases it is appropriate to compare carefully the advantages and disadvantages of each technique with the properties of the tissue source and the goal of the experiment. RDA of cDNA is a good choice in situations requiring a sequence-independent approach with high accuracy for detecting differential expression. Microarrays are the best choice when multiple comparisons are desired, when rapid analysis is critical, or when analyzing a subset of transcripts is sufficient. Some investigators have used both types of techniques, either in parallel or in series (Geschwind *et al*., 2001; Reick *et al*., 2001). The paper by Dougherty and Geschwind (Dougherty and Geschwind, 2002) nicely describes the benefits deriving from the complementary advantages of the two techniques. One general caveat is that all of the expression profiling techniques use genes as their basic unit. Little has yet been done to adapt them to analyze the phenotypic diversity caused by alternative splicing of mRNAs, although differential subtraction methods already have the capacity to detect some splice variants. Quantifying splice variants is important because they may represent a majority of the mRNA diversity in mammalian cells (Lander *et al*., 2001). Exons, rather than genes, might therefore be the most appropriate fundamental units on microarrays. In addition, a particular concern with expression profiling experiments in neural tissues is cellular heterogeneity. Many significant neurobiological questions involve phenotypic responses in a minority of the cells in the affected tissue. Given that many of the interesting molecular differences often belong to relatively rare transcripts, expression profiling techniques are often insufficiently sensitive to fully investigate these questions. To overcome this problem, investigators have, or will, combine expression profiling techniques with surgical microdissection, antisense RNA amplification (Eberwine *et al*., 1992), laser capture microdissection (Ohyama *et al*., 2000), production of transgenic animals expressing phenotype-specific markers and isolation of specific cell types by flow cytometry.

## References

**Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S.R., Moon, K., Burcham, T., Pallas, M., DuBridge, R.B., Kirchner, J., Fearon, K., Mao, J.** and **Corcoran, K.** (2000) *Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays.* Nat. Biotechnol., 18, 630–634.

**Diatchenko, L., Lau, Y.-F., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurshaya, N., Sverdlov, E.D.** and **Siebert, P.D.** (1996) *Suppressive subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries.* Proc. Natl Acad. Sci. USA, 93, 6025–6030.

**Dougherty, J.P.** and **Geschwind, D.H.** (2002) *Subtraction-coupled custom microarray analysis for gene discovery and gene expression studies in the CNS*. Chem. Senses, 27, 293–298.

**Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P.** and **Trent, J.M.** (1999) *Expression profiling using cDNA microarrays.* Nat. Genet., 21, 10–14.

**Eberwine J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M.** and **Coleman, P.** (1992) *Analysis of gene expression in single live neurons.* Proc. Natl Acad. Sci. USA, 89, 3010–3014.

**Geschwind, D.H., Ou, J., Easterday, M.C., Dougherty, J.D., Jackson, R.L., Chen, Z., Antoine, H., Terskikh, A., Weissman, I.L., Nelson, S.F.** and **Kornblum, H.I.** (2001) *A genetic analysis of neural progenitor differentiation.* Neuron, 29, 325–339.

**Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., Hughes, J.E., Snesrud, E., Lee, N.** and **Quackenbush, J.** (2000) *A concise guide to cDNA microarray analysis.* Biotechniques, 29, 548–562.

**Hubank, M.** and **Schatz, D.G.** (1994) *Identifying differences in mRNA expression by representational difference analysis of cDNA.* Nucleic Acids Res., 22, 5640–5648.

**Lander, E.S., Linton, L.M., Birren, B., *et al*.** (2001) *Initial sequencing and analysis of the human genome.* Nature, 409, 860–921.

**Liang, P.** and **Pardee, A.B.** (1992) *Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction.* Science, 257, 967–971.

**Lipshutz, R.J., Fodor, S.P., Gingeras, T.R.** and **Lockhart, D.J.** (1999) *High density synthetic oligonucleotide arrays.* Nat. Genet., 21, 20–24.

**Luo, Z.** and **Geschwind, D.H.** (2001) *Microarray applications in neuroscience.* Neurobiol. Dis., 8, 183–193.

**Ohyama, H., Zhang, X., Kohno, Y., Alevizos, I., Posner, M., Wong, D.T.** and **Todd, R.** (2000) *Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization.* Biotechniques, 29, 530–536.

**Prolla, T.A.** (2002) *DNA microarray analysis of the aging brain*. Chem. Senses, 27, 299–306.

**Reick, M., Garcia, J.A., Dudley, C.** and **McKnight, S.L.** (2001) *NPAS2: an analog of clock operative in the mammalian forebrain.* Science, 293, 506–509.

**Schena, M., Shalon, D., Davis, R.W.** and **Brown, P.O.** (1995) *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 270, 467–470.

**Velculescu, V.E., Zhang, L., Vogelstein, B.** and **Kinzler, K.W.** (1995) *Serial analysis of gene expression.* Science, 270, 484–487.

*Accepted December 7, 2001*