# The importance of using human-based models in gene and drug discovery

Completion of the human genome sequence has opened an unprecedented opportunity to discover the genetic contribution to many of the complex diseases that afflict humans and to use this information for new approaches to drug discovery. Mice and humans have more than 95% of their genes in common yet mice are not men (or women). Although cell-based systems and animal models of disease have been the cornerstone of drug discovery it is becoming increasingly apparent that they are of limited predictive value for complex disorders. This is particularly evident for disorders of the brain. Recent advances in technologies for studying the function of genes in human tissues have benefited from a massively parallel analysis of gene and protein expression and made possible the efficient use of difficult to obtain human tissue. However, there are still many technical and ethical hurdles to be addressed before the mainstream use of human-based discovery models dependant upon access to human tissues can play a substantial role early in the process of gene and drug discovery and development. This review highlights both the challenges and the unprecedented opportunity to use human cells and tissues to improve the drug discovery process.

**By Dr Michael G. Palfreyman, Dr Vinod Charles and Jeffrey Blander**

Drug discovery is for the most part directed at human diseases. Why then do we wait until the clinical phase to determine if our potential human therapeutic is effective in humans? Historically, this has been due to several constraints on the use of human tissue: historically access to high quality materials has been severely limited, the tissue commonly available for research is frequently too poorly defined for the sort of precise measurements that are required by the modern drug discoverer and there have been a number of ethical and legal concerns regarding the patient donor. Fortunately, things are beginning to change. In response to incorporation of human genetic and genomic information into the discovery and development of new drugs, more focus is being placed on the use of human cells and tissues early in the discovery process. The cornerstones of the drug discovery process, target identification and validation and predictive high throughput screening are facing unprecedented challenges. In spite of the massive investment made by the industry, the number of new chemical entities reaching the marketplace is insufficient to satisfy unmet medical needs and the appropriate financial goals of the pharmaceutical industry. Both time and cost of reaching the marketplace have grown to astronomical proportions. A significant contribution to this time/cost equation is the massive attrition that occurs throughout the drug discovery and development process. Having better pre-clinical

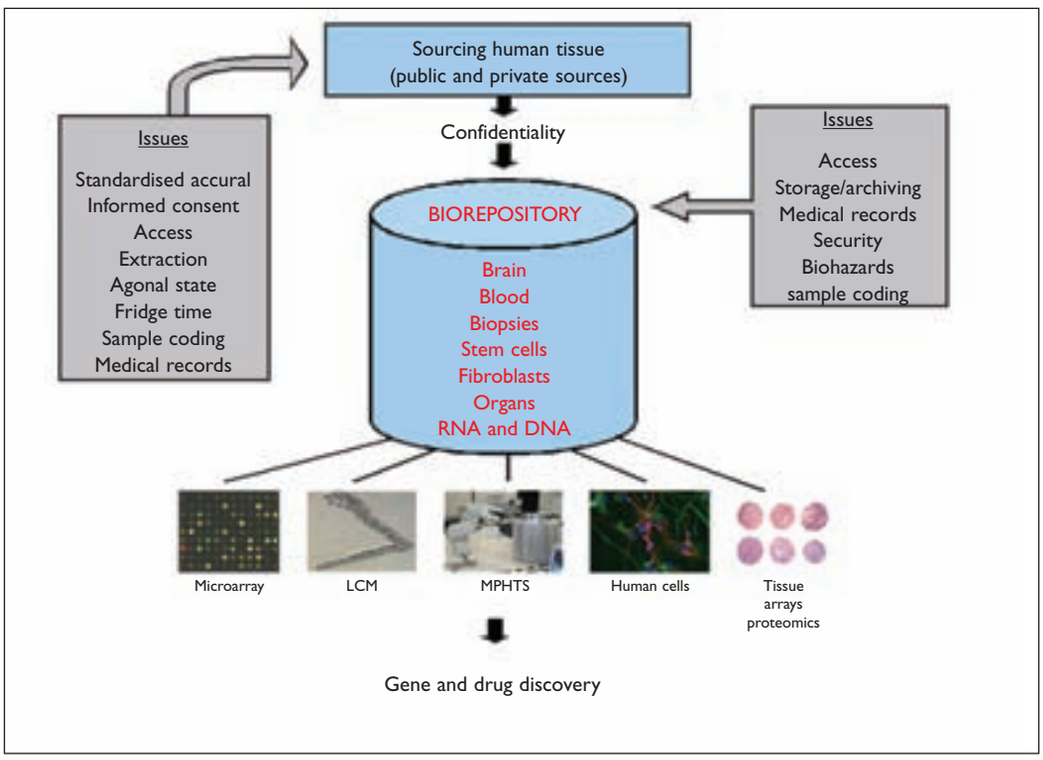


**Figure 1**
High quality biorepositories require careful management to address a number of issues and attention to detail regarding medical information while respecting all aspects of patient confidentiality. Tissues or cells must be collected and stored under controlled conditions to allow access to DNA, RNA, proteins, cyto architecture for laser capture micro-dissection and for creating tissue arrays for *in situ* hybridisation

predictors of clinical efficacy and side effects and identifying failures very early on would significantly address these issues. Obviously testing of novel therapeutics cannot take place in humans until all the potential therapeutic and toxicity issues have been addressed. A significantly better chance of success in developing both therapeutics and diagnostics is to work with the human targets and preferentially in the naturalistic human tissue.

One of the major challenges facing the drug discovery community is the limitation and poor predictability of animal-based strategies. Over the last decade drug discovery has largely been based on finding targets in animal models and then identifying the human homologue. Frequently, the human target is expressed in an artificial environment to allow high throughput drug screening to be undertaken. Many drugs have failed in later stages of development because the animal data were poor predictors of efficacy in the human subject. One of the overriding interests of the pharmaceutical and biotechnologies industry is to reduce the cost of discovering and developing new therapeutics and diagnostics by creating alternative development strategies that are less reliant on poor animal predictor models of human disease.

## Table 1

| | |
|---|---|
| 1 | Need for high quality and quantity of representative disease diversity in commercially or publicly available repositories |
| 2 | Consistent protocols for the collection and preparation of human materials across medical institutions |
| 3 | Standardised medical vocabularies to collect comprehensive donor medical histories for post-hoc statistical and bioinformatics analysis |
| 4 | Maintenance of the integrity of tissue to maximise quality of data for genomics and proteomics and *in situ* applications (maintenance of RNA, DNA, protein and cytoarchitecture) |
| 5 | Ensure bioethical issues surrounding accrual (confidentiality, informed consent, commercialisation of data, patentability and intellectual property of findings) |

## The promise of human-based discovery models

Sequencing of the human genome[1,2] offers an expanded opportunity to develop novel diagnostics and therapeutics based on understanding the underlying pathophysiology and pathogenesis of many complex medical conditions that are so prevalent. The case for using human-based models early in the discovery process is being elegantly made by scientists from major pharmaceutical and biotechnology companies. This is further reinforced by a recent study[3] that compared the expression of genes between chimpanzees and humans. Although the species share more than 98.9% gene identity, the expression of genes in the brain was more than five-fold greater in humans

than in the chimpanzees. Differences in gene expression level in the liver and blood were much smaller suggesting that metabolism and distribution could be modelled in such a species, although disorders of higher order mental functioning such as schizophrenia would be poorly represented. Differences from mice were even greater. These differences reinforce the importance of using human disease models in drug discovery as a real predictor of human efficacy.

Genomics-based research needs access to high quality human tissue and cell lines to achieve this objective. It has been estimated that there are more than 300 million human tissue samples in various repositories throughout the world. Access to different types of human tissue from living patients occurs on a routine basis. Blood is drawn from nearly every human being; skin, muscle, fat, bone marrow and liver biopsies are commonplace. Cancer tissues are regularly collected and processed and many tissue banks are replete with tumour tissues. More recently, stem cells have become a potential resource that can be expanded and differentiated into a whole range of different tissues. More ways to obtain tissue from human subjects are now being identified. Unfortunately, these riches have been of little use for gene discovery, target identification and drug discovery for reasons pertaining to inconsistent collection protocols that lead to low quality sample preservation as well as the inability to properly associate relevant clinical information because the initial informed consent process was ill-defined or non-existent. In **Table 1** we outline several issues and challenges facing the application and use of human tissues in drug and gene discovery.

Access to DNA from blood samples has now been very well established and many population wide studies are being conducted, eg the Icelandic population studies being conducted by DeCode (www.decode.com) and the Ashkenazi Jewish studies undertaken by IDGene (www.idgene.com). To date however, there is no comprehensive source of human tissues. A number of tissue banks exist in academic laboratories, eg the Harvard Brain Bank (www.brainbank.mclean.org) and University of Maryland Brain and Tissue Bank (medschool.umaryland.edu/BTBank), but the majority of tissues come from diverse hospital, surgical and pathological services, from doctor's offices, blood banks and government-funded disease-specific biorepositories such as those supported by the National Institutes of Health. More recently, in response to the recognition of the role that genomics will play in the healthcare of nations, a number of countries have committed significant resources to creating national biorepositories. For example, Norway recently announced a significant initiative. Unfortunately, virtually none of these tissue banks apply uniform standards of collection, storage, data retrieval, medical information or analysis of the samples and very few sources are suitable for analysis of RNA or protein. Procuring and handling human tissue is an expensive and laborious process that requires considerable technical skills to meet the exacting requirements of modern genomic research. To fill this need for both quality control and well documented medical information a number of biotechnology companies, such as American Biogenetic Sciences (www.mabxa.com), Ardais Corporation (www.ardais.com), Asterand (www.asterand.com), Clinomics Biosciences (www.clinomicslabs.com), Genomics Collaborative (www.getdna.com) and LifeSpan BioSciences (www.lsbio.com), have been established. A recent report[4] highlights many of these new companies. **Figure 1** details the issues and challenges for the accrual and facilitation of use of human tissue in gene and drug discovery.

As discussed, although there are many millions of human samples in biorepositories throughout the world very few of them are of a suitable quality for extraction of RNA or proteins. Genomics- and proteomics-based drug discovery research predicates, the maintenance and integrity of quality RNA and protein. Moreover, in applying stringent standards for controlling; freezing/fixation

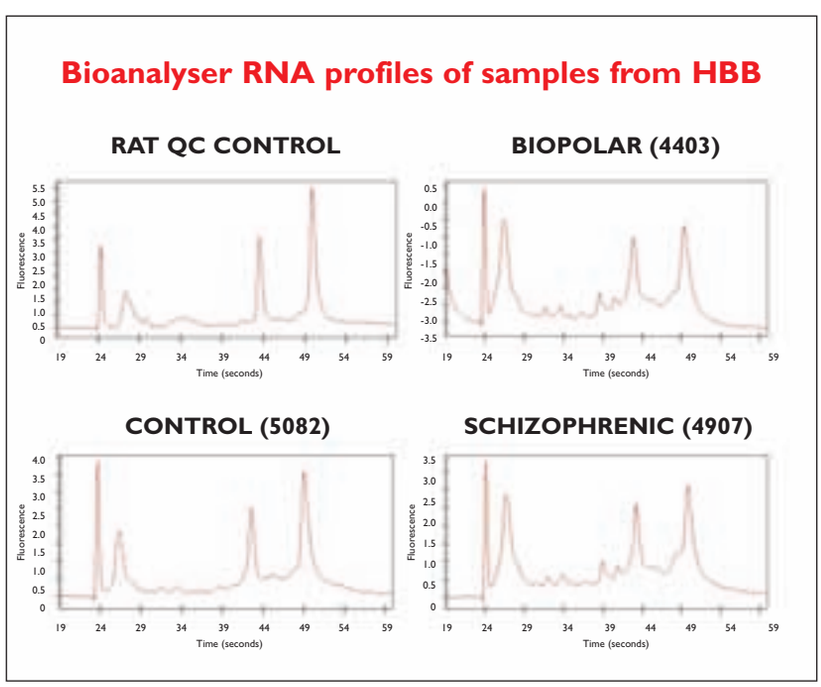


**Figure 2**
Profile of RNA samples obtained post-mortem from rat brain pre-frontal cortex and from samples of pre-frontal cortex from the brain of a normal subject and subjects diagnosed with schizophrenia and bipolar disorder. Human brain samples were provided by the Harvard Brain Bank

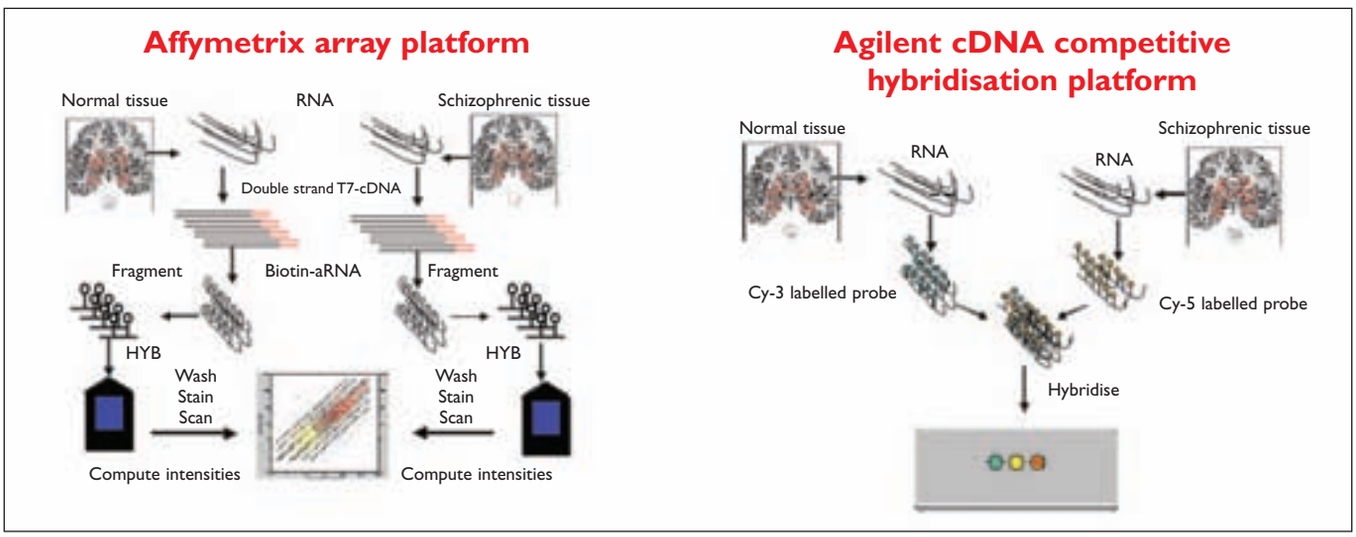


**Figure 3**
Analysis of the human transcriptome using Affymetrix and Agilent micro-arrays. The examples shown are for processing human brain samples from patients diagnosed with schizophrenia and were obtained from the Stanley Research Institute

time, Post-mortem Interval and thereby cytoarchitectural integrity will not only minimise RNA degradation but also preserve cellular membranes and proteins. Quality control analysis of the samples becomes an essential priority. For example, DNA, RNA and protein quality can be assessed with the Agilent 2100 BioAnalyser based on a Calliper designed 'lab-on-a-chip' electrophoresis methods (www.agilent.com). Results from such an analysis of brain RNA are shown in **Figure 2**.

## Identifying targets from human cells and tissues

The drug discovery community is in the midst of a major retoolling of its approach to target discovery, target validation and high throughput screening. Historically, drug discovery has focused on finding and validating targets, many of which came from animal models, and then setting up a high throughput screen. Screening is often undertaken with the same animal protein expressed in a rodent cell system or as the isolated target in an appropriate HTS format, such as an enzyme assay. This approach focused mostly on finding therapeutics with activity on a single target with the emphasis on potency and specificity. We are now moving from the era of ultra-specific drugs with activity on just one target to a realisation that complex disorders require a systems biology approach to therapeutic management. The reductionist approaches we have used previously where the 'validated' target is expressed in a non-physiologically relevant environment was largely done to address the constraints of high throughput screening. Cell-based screening has become more effective and with better access to stem cells, transformed cells and engineered human cell lines a move away from animal tissues and cells has begun.

We are becoming increasingly aware of the interplay of systems needed to maintain homeostasis in the organism and how disease arise when these homeostatic mechanisms come out of balance. Moreover, such systems have considerable in-built redundancy and drug effects on one system are often compensated by adaptive changes in alternative systems. Facilitating this systems biology approach to target identification is the use of the massively parallel processing power of micro-arrays. Examples of such systems are the Affymetrix and Agilent DNA micro-arrays (www.affymetrix.com; www.agilent.com). Nowhere is this approach to understanding the complexity of systems and their interaction more apparent than in diseases of the central nervous system (**Figure 3**).

## Target validation and drug discovery

The pharmaceutical industry is now realising the potential of the multiple points at which micro-array technology and human tissue could be integrated to aid in target identification, target validation, safety testing, compound selection and critical decision-making on parameters to advance products to clinical testing. Much of the technology from the biotech industry offers novel and potentially very powerful ways to integrate the target discovery phase with the drug discovery phase based on the creative use of human tissues, stem cells and micro-array capability.

Applying the latest technologies to target identification and validation directly in human tissues is the first and critical step in the holy grail of drug

discovery. It has been elegantly illustrated how the use of micro-array technology and/or proteomics could allow a massively parallel approach to target identification. These technologies are gradually becoming mainstream. Although they still have a number of technical challenges to address, such as sensitivity and reproducibility, they are proving their value in multiple ways. Excellent examples of using micro-array technologies to obtain unique insights into the role of myelination in the pre-frontal cortex of schizophrenics have been conducted by Kenneth L. Davis (Mt. Sinai School of Medicine); several new targets identified in post-mortem tissue from schizophrenic and bipolar patients have been identified by scientists at Psychiatric Genomics[5] (www.psygenomics.com), many novel cancer targets have been identified by scientists from Avalon Pharmaceuticals (www.avalonrx.com); by Emmanual Petricoin's group at CBER/FDA and Krishnarao Appasani's group at Perkin Elmer™ Life Sciences (www.perkinelmer.com); substantial databases of gene-expression from normal and diseased human tissue are available from Gene-Logic (www.genelogic.com). It is clear that the micro-array technologies can also be used to validate targets, examine toxicity profiles and eventually be used for drug discovery (see below). Of particular note is the way that several of the premier pharmaceutical companies have integrated these new technologies into their target identification and validation strategies. In all cases the use of human tissue was central to maximising the value of these powerful technologies. Micro-array technologies are particularly suited for use with human tissues because the widest availability of genes on the chips is human. Rat and mouse micro-arrays are now becoming available so it is a valuable coincidence that this powerful technology has further complemented the important and growing place that human tissues have established in the drug discoverer's armamentarium.

Drug targets are for the most part proteins. To date, the high throughput proteomics technologies are not so advanced as the micro-array technologies in terms of numbers of items per array, nor in sensitivity. However, many new technologies are being

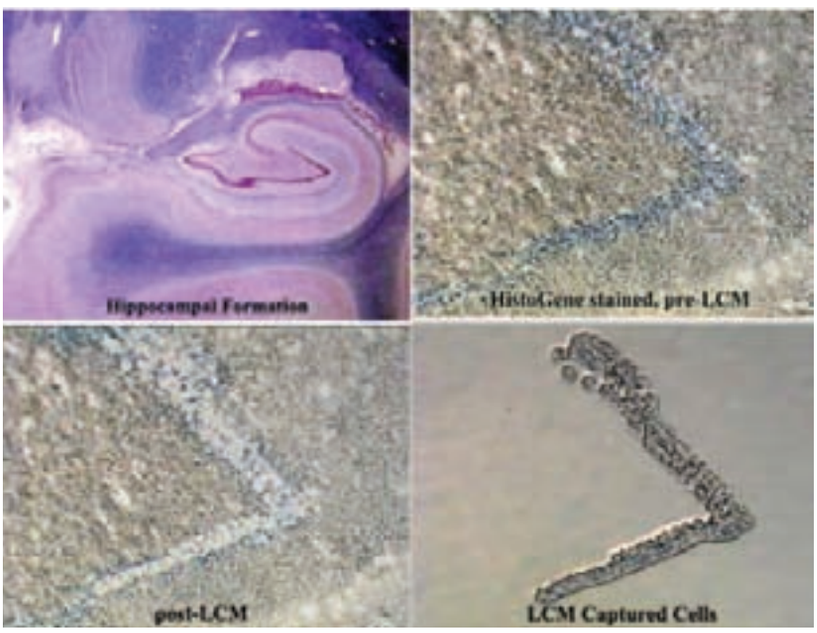


**Figure 4**
Laser-Capture microdissection of human hippocampus, illustrating the specific staining, identification and removal of dentate granule cells. The last panel (bottom right) shows the individual captured cells for RNA extraction and amplification

applied and valuable data generated. Through the use of high-resolution 2DE (NEPHGE technology, Wita Proteomics, AG, www.wita-proteomics.com) comparisons between diseased and healthy tissue, and drug treated and untreated human cell lines, protein targets can be validated, drug efficacy determined and drug toxicity profiles examined. Other examples include technologies developed by GeneProt (www.geneprot.com), ActivX (www.activx.com), Aclara Biosciences (www.aclara.com), and Large Scale Proteomics (www.lsbc.com).

Although neither proteomics nor gene expression analysis are ideal technologies they both have the powerful potential to accelerate the discovery of new targets, particularly when used in combination with other tools such as RT-PCR, *in situ* hybridisation, antibody localisation and functional studies. These technologies, when applied to human tissue produce the sort of breakthrough synergies that significantly accelerate and improve the efficiency of the drug discovery process.

## Drug discovery applications using human tissues

### Microarray

The validity and utility of gene-based drug discovery would not be possible without the advent of microarray technology. Until recently, comparing expression levels across different tissues or cells was limited to tracking changes in a few genes at a time. Utilising these 'limited sets' of gene expression changes made it historically difficult to investigate broader effects, mechanisms or pathways for a given drug or disease. Microarrays, however, typically consist of thousands of immobilised DNA sequences spotted on a glass slide or gene-chip, allowing for the simultaneous scanning of effects across thousands of genes. The marriage of accessible human tissue and microarray technology now enable a multi-faceted approach towards the ultimate goal of drug and diagnostic development. Using human brain, for example, microarray analyses can be conducted to compare gene expression profiles between normal and diseased cases – a 'disease or diagnostic signature'. Assuming proper collection of patient's records, the tissue can also be used to tease out important variables such as gene expression effects of medications, drug activity and toxicity – a 'pharmacogenomic signature'.

The composition of DNA on microarrays is of two general types; Oligonucleotide-based and cDNA-based. Oligonucleotide-based arrays (Affymetrix Inc; www.affymetrix.com), allow the user to do both genotyping-based and gene expression comparisons of multiple samples. cDNA-based arrays (Agilent Technologies, Inc; www.agilent.com) are generally used for the comparison of relative gene expression in multiple samples only. Affymetrix platform uses a series of 11 oligonucleotides, each 25 bases in length that are a perfect sequence match to the target gene. Paired with each perfect match oligo is an oligo that contains a single nucleotide mismatch to control for cross hybridisation. The Affymetrix platform requires conversion of the RNA sample to a double stranded cDNA pool containing a T7 polymerase recognition site and amplification with T7 polymerase in the presence of biotin-UTP. The probe is fragmented and hybridised to an Affymetrix GeneChip. Differential expression is determined by comparing the results of one chip to another. The Agilent platform uses cDNA clones with an average length of one Kb for each gene. To compare biological samples, RNA from one sample is converted to cDNA labelled with Cy-3 and the other RNA is converted to cDNA labelled with Cy-5. The two labelled cDNAs are mixed together and simultaneously hybridised to the array. The intensity of each dye hybridised to each feature is determined by scanning the array with two lasers specific for each dye. Differential expression is determined by the ratio of signal intensities. Both array platforms have been used with varying results by numerous laboratories to measure gene expression of human postmortem brain tissues. DNA array analysis provides an excellent tool to enlarge our knowledge of genome function for subsequent use in pharmacogenomics research.

**Laser-capture microdissection**

It was recently shown that gene expression profiles can be obtained from differing neuronal subtypes in post-mortem tissue using laser capture microdissection technology and T7-based RNA amplification with cDNA microarrays. In 1997 LCM was commercially developed through a Collaborative Research and Development Agreement (CRADA) partnership with the National Institutes for Health and Arcturus Systems, Inc (www.arctur.com). LCM was developed to automate and standardise microdissection and has greatly increased reproducibly and accuracy of selecting specific cells from a complex tissue for subsequent molecular analysis. Using laser-capture microdissection, experiments can now be performed to study changes in gene expression from individual cell types within specific regions of the brain in normal and disease states. In disease states, the diseased cells of interest are surrounded by heterogeneous tissue elements. Cell types undergoing similar molecular changes, such as those thought to be most definitive of the disease progression, may constitute less than 5% of the volume of the tissue biopsy sample (NIH, LCM core facility, unpublished data). Using heterogenous tissue samples, the presence of gene expression as well as differential gene expression between diseased and control samples are often undetectable due to dilution from surrounding, non-relevant cells. LCM thus, uniquely enables the evaluation and detection of genes, and gene expression changes that are only expressed in small subpopulations of cells. Laser-capture microdissection is essential to apply molecular analysis methods to study evolving disease lesions in actual tissue. These types of analyses can greatly increase our understanding of; interactions and dynamics within neuronal circuits, changes in gene expression of homogeneous neuronal populations, and discreet mechanisms of drug action (**Figure 4**).

**Human cells**

Blood cells have been routinely collected from human patients and used in target and drug discovery, particularly in the areas of immunology and oncology. More controversial is the use of stem cells. Presidential restraints on funding embryonic stem cell research have had the unexpected effect of highlighting their potential. Stem cells have the unique property of being able to differentiate under the appropriate conditions into any cell in the body, including brain cells. A number of companies are considering using stem cells for cell therapy or transplantation. More recently, it has become evident that human stem cells could also be used for target identification and drug discovery. A number of academic labs and a few companies are now working with stem cells for target identification. For example, ReNeuron (www.reneuron.com) and Neural Stem (www.neuralstem.com) have developed a number of cell lines from different regions of the brain and can efficiently differentiate these pluripotent cells into neurons and glia with phenotypes that differ and reflect the brain region from which they were collected. Rhinoneuroepithelial cells obtained from the nasal passage are neuronal progenitor cells that readily differentiate into neurons and have the potential to be obtained by biopsy from living patients. Several other sources of adult stem cells from the skin, placenta, fat and bone marrow are also being evaluated. Considerable debate has occurred regarding the potential of adult stem cells to differentiate into a variety of organs suitable for study. For example, a recent study from Irving Weissman's group at Stanford (www.med.stanford.edu) has shown that embryonic stem cells have a significantly greater potential than hemopoetic stem cells to differentiate into various cell types. Neuronal precursor cells, such as rhinoneuroepithelial cells, can be differentiated into neurons and used as *in vitro* models CNS functions and used to study gene function and drug actions. Psychiatric Genomics is using differentiated neuronal stem cells in this way but has expanded their use to drug discovery. Although the adult precursor and stem cells can be obtained from patients with various diseases, it is not always possible to obtain human cells with the appropriate pathology. Once we know the underlying pathophysiology of a disorder it may be possible to manipulate 'normal' human cells to exhibit the appropriate disease-specific abnormalities. One

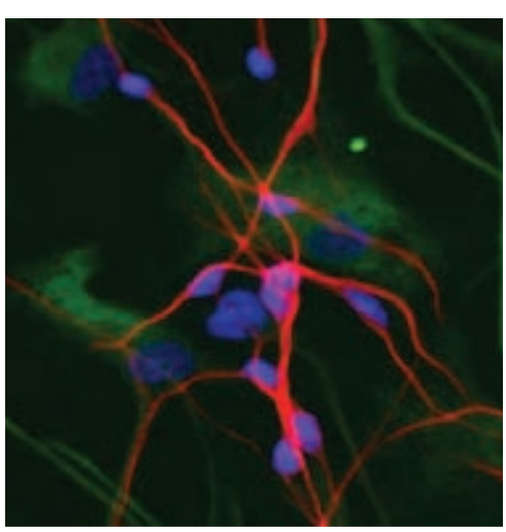

**Figure 5**
ReNeuron cortical stem cell treated by LIF (leukemia inhibitory factor) HSC were expanded as neurospheres with bFGF and EGF.
Cells were fixed with 4% paraformaldehyde and double immunostained for Tubulin Beta III (red, neuron), GFAP (green, astrocyte) and Hoesht (blue, nucleus).
High magnification (x40)

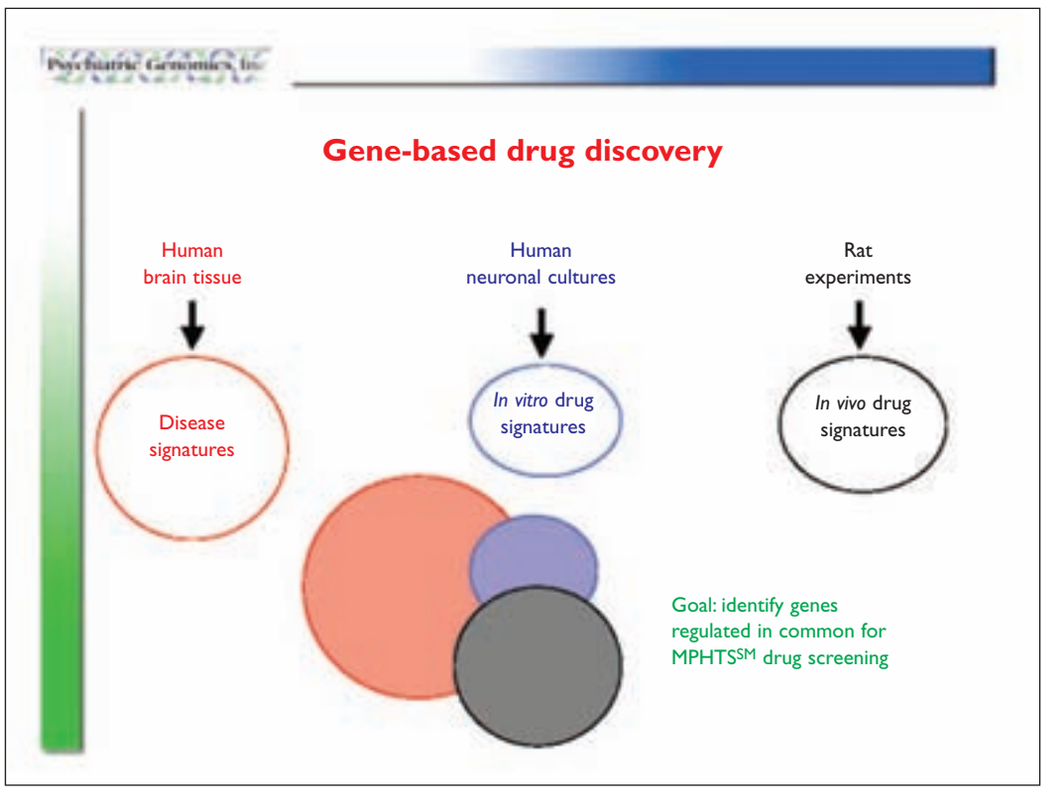


**Figure 6**
The changes in gene expression associated with a disease: the 'Disease Signature' and the changes associated with drug action: the 'Drug Signature' can be used as the basis of a high throughput drug discovery process to find novel agents that will regulate these genes in a common fashion

such potential technique involving regulated gene expression (RHeoPlex™) has been developed by RheoGene (www.rheogene.com). It is a ligand-inducible gene regulation technology that enables the precise regulation of one or more target genes in the same cell allowing the study of gene expression, cell signaling and protein interactions *in situ*. Most gene manipulation has been conducted in cell-based systems with antisense knock-downs and companies such as Sequitur (www.sequiturinc.com). Other examples are 'gene activation' technologies developed by Transkaryotic Therapies (www.tktx.com) that turn on selected genes and can be used to silence gene expression. These manipulations of human cells offer numerous possibilities for advancing the drug discovery and development process.

There are many sources of human cells that are used in target identification and drug discovery and although several pharmaceutical and biotechnology companies are developing a battery of different human cell lines there is clearly no consensus on the predictive value of such human-based systems. However, they are clearly superior to those obtained from animals. For example, human adipocytes are more readily available, and these cells developed by Zen-Bio (www.zen-bio.com) are the subject of considerable interest. Major antiobesity research in the area of adipocyte metabolism has been conducted on rodent adipocyte-like cell lines. There are key differences in the genetic and metabolic make up of humans and rodents, which can only be evaluated after careful investigation using human adipocytes. These can form the basis of a useful target and drug discovery tool (**Figure 5**).

## Drug discovery

Although there has been significant progress in identification of new targets, many companies feel they are awash with targets. Predictions of tens of thousands of new targets versus the 500 or so we currently have will require new technologies for validation and certainly new tools for drug discovery. Using human tissues to prioritise relevant targets is a first step, as is validation using diseased tissue and human cells. However, there are currently very limited technologies that take the gene expression information into a high throughput drug discovery screen. Typically, the gene expression information has been used to pick a favourite gene for expression in a suitable cell system and subsequent high throughput screening using the familiar substrate or ligand displacement assays. A newer and potentially more powerful approach is to use Multi-Parameter High Throughput Screening℠ (MPHTS℠). This new technology from Psychiatric

Genomics (www.psygenomics.com) uses human cell-based screening and quantitative changes in gene expression as the read out[6]. This technology has the major advantage that multiple genes can be read in parallel and the screening is very information rich. Typical high throughput screening measures a single target and is binary in its output (ie, it shows if the unknown compound produces an effect or not). MPHTS$^{SM}$ screens compound against multiple genes in parallel and even if the effect on the 'gene signature' is not the desired one it could be of interest for another objective. Thus such screening is very information rich. Most importantly, MPHTS$^{SM}$ does not require that the function of the gene product be understood in great detail, only that the change in gene expression is a reliable indicator of disease or drug action. This allows assays to be set up very quickly as it is not necessary, as in typical screens, to express and purify the protein or create a recombinant cell line expressing the receptor, provided of course the cells used in MPHTS$^{SM}$ express the desired genes. One significant advantage of using differentiated human stem cells is the similarity of gene expression to the native tissue.

The technology has been developed to allow up to 16 genes to be measured at any one time but it is certainly likely that mini-arrays of 100 or more genes could be devised. This approach, although universal in its application to any therapeutic area has particular attraction for complex polygenic disorders such as cancer and psychiatric disorders. Considerable progress has been made in the latter application. Micro-array analysis of tissue from patients with these complex disorders facilitates an analysis of the interaction between genes in a pathway and identifies multiple genes that contribute to the 'disease signature'. This 'disease signature' can then form the basis of a drug screening approach to return such aberrant gene expression patterns back to normal using small molecular drugs. This is illustrated in **Figure 6**.

Complementary to this approach is analysis of the gene expression patterns that result from the action of therapeutic agents. These 'drug signatures' can also be used to find new drug classes with improved properties. A critical part of this 'gene signature' discovery paradigm is access to human nerve cells that can be grown in culture under controlled conditions. One such system uses human neuronal stem cells that can be differentiated into neuronal systems that really reflect the complexity of the CNS and include the multiple interactions between neurons, astrocytes and oligodendrocytes that most closely approximate the function of both the normal and diseased brain. Finally, discovery of drugs that act on the human central nervous system are best studied in human

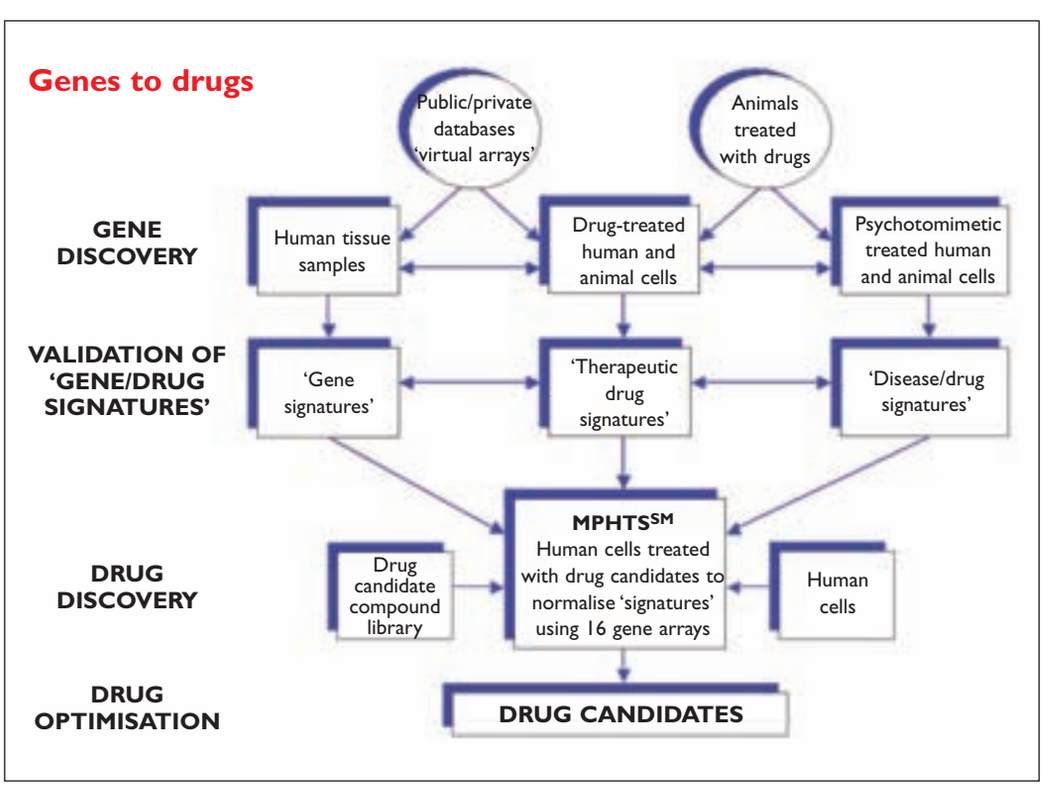


**Figure 7**
Integration of the application of massively parallel genomic and proteomic technologies to the target discovery, target validation, drug discovery and drug optimisation process using human cells and tissues will significantly reduce attrition and accelerate the drug discovery and development process

cell-based systems and differentiated neural stem cells present such an opportunity for MPHTS[SM]. MPHTS[SM] using human neuronal cells provides drug candidates for these families of new targets that will enable discovery of class-defining therapeutics for these complex psychiatric diseases.

## Pharmacokinetics, pharmacogenetics, pharmacogenomics and diagnostics

Study of pharmacokinetic parameters plays a critical role in helping to assess the necessary drug exposure for effective but non-toxic drug effect. Multiple factors play a role in determining *in vivo* drug exposure and although animal studies are essential prerequisites of human dosing, the more one can determine the relationship between the animal and human data the better the predictive power. Examples of human systems suitable for *in vitro* studies of ADME (Absorption, Distribution, Metabolism and Excretion) properties include CaCo-2 cell lines for determining the relative permeability of different analogues of a given chemotype, human hepatic microsomal or S9 fractions of human or animal hepatocytes for determining metabolic stability and human blood cells for protein binding and metabolism studies. A comprehensive utilisation of these human systems can aid in the study of new drug candidates. Pharmacogenetics/genomics and diagnostics require blood samples for DNA analysis but increasingly assessment of gene expression levels in human cells is being considered as a potential pharmacogenomic or diagnostic predictor of drug response and as such will increasingly become a valuable tool.

## Bioinformatics and databases

Tools are being developed to allow the accurate tracking of donors, collected tissues and the associated medical information in such a way that all aspects of quality control and all ethical standards such as patient privacy and informed consent are tracked and monitored. They support the enrollment and informed consent of donors, the collection and storage of human specimens and the accumulation of associated clinical information in a manner consistent with the highest principles of patient confidentiality. All biorepository companies are attempting to create such databases and are illustrated by the systems such as those developed by Ardais Corporation and Genomics Collaborative. A number of companies are developing databases of gene expression or proteomics information on very large collections of normal and diseased human tissue under carefully controlled conditions. These databases can become a valuable resource for comparative information, identification of novel targets, target validation and as a platform for examining the effects of drugs on the tissues. One such database from Pharmagene (www.pharmagene.com) contains gene expression data from more than 70 human tissues. LifeSpan Biosciences is creating localisation databases of major gene families in normal and diseased tissues. Another example is GeneLogic's suite of gene expression databases with comparative data between humans and animal models as well as a comprehensive analysis of gene expression following toxic drugs (ToxExpess[TM]).

The challenges of implementing a suite of informatics tools to manage a repository of human tissue products and associated clinical information are multi-fold. The prototypical supply chain process for preparing human materials into research quality resources requires information systems to support sourcing of unfinished goods inventory at the medical institution (both human materials and associated clinical information), the transfer, storage and processing of unfinished goods inventory into finished goods products (micro array technologies, RNA, DNA, etc) and a mechanism to provide unencumbered access to the research community to perform detailed search and product requests based upon specific genomic study criteria.

Sourcing of human tissue and clinical information requires the use of technology solutions that provide a seamless interface for the operations personnel involved in the banking and shipping of human materials. The major tenet for deploying such systems is that they must not interfere with the routine care of the patient or normal workflow of the clinical staff working within the hospital setting. Web-enabled technologies and wireless networks have provided marked enhancements to the software deployment process and data entry experience to user communities as compared to older client server-based applications. The incorporation of structured clinical data otologies within an informatics tool set for human tissue collection and clinical data capture is critical to facilitating the standardisation of tissue processing protocols and quality control parameters. For example within each medical sourcing institution there is wide variation in clinical practice, standards of care, and logistical operating procedures. Data capture tools that use controlled vocabularies implemented through searchable pick list's for diagnostic procedures, clinical diagnoses, treatments and medication histories are a pre-requisite

to ensure generalisability across disease states for researchers to accurately assess detailed clinical histories and patient outcomes information. Comparable collection protocols that track and notify operating personnel about outliers in surgical resection times, size of extracted tissue blocks, and time in and out of formalin fixative or cryopreservation storage units provide a quantifiable approach to analysing quality standards for histology review and derivative product preparation. Additionally improvements in electronic patient donor screening and consenting techniques coupled with the automation of study inclusion criteria within the sourcing institutions electronic medical record system will continue to systematically help to improve the accrual rates of patient donor's. Ongoing enhancements to the donor enrollment and data collection processes will greatly facilitate the collection of large numbers of human tissue samples across a growing diversity of neoplastic and non-neoplastic disease areas. Systems and processes that improve the availability of high quality human materials have helped to increase the number of studies being conducted and fuelled the demand for high quality samples needed to adequately power clinical genomics based research.

## Integration of gene/target discovery, target validation, drug discovery and development

The common theme of using human tissues from gene discovery through to drug development and integrating the power of gene expression analysis into the process can lead to both a more rapid and higher probability of success approach to identifying NCE to fill the rapidly drying pipeline of the major pharmaceutical companies. The integration of this process is illustrated in **Figure 7**.

## Bioethics and legal issues

No discussion on the use of human tissues in drug discovery would be possible without consideration of all the bioethical issues. Moreover, the use of

## References

1 International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409, 860-921 (2001).
2 Venter, J et al. The sequence of the human genome. Science 291, 1304-1351 (2001).
3 Editorial. Biorepositories. Start-up, July/August, 19-27 (2001).
4 Enard, W et al. Intra- and interspecific variation in primate gene expression patterns. Science 296, 340-343 (2002).
5 Palfreyman, MG et al. Novel directions in antipsychotic target identification using gene arrays. Current Drug Targets-CNS & Neurological Disorders, 1, 227-238 (2002).
6 Palfreyman, MG et al. Drug discovery of the mind: Therapeutic approaches to polygenic psychiatric disorders. PharmaGenomics 2, 26-32 (2002).


human stem cells is further influenced by debate on the sanctity of human life. All scientists who work with human tissues have a legal as well as moral obligation to adhere to all protocol guidelines, and take significant steps to respect and adhere to all bioethical guidelines on the use of human tissues. The importance of this aspect of research with human tissues is central to the use of human tissue in research and must be further emphasised in all projects that use human tissues. In particular, we must reach a clear consensus about the meaning of informed consent being 'informed', the appropriate restrictions to be placed on the use of information, the right of refusal and the potential impact on discrimination of employment and insurance when using genetic information or extracting biological information from human materials that could in any way negatively influence the individual or group of individuals who generously provided the tissue samples. Many of the issues such as confidentiality of information are similar to those that are faced with all types of medical information, but there is a strong perception that genetic information is more like 'life's future diary'. Such a 'diary' of an individual reveals their potential for illness and could easily, if not properly protected, be the basis for discrimination. Privacy issues, state laws vs federal laws, legal issues, intellectual property rights, pass through of benefit both medical and financial to the donor of tissue, ownership of material, and most importantly, the confidentiality issues are all subject to intense debate and need to be resolved satisfactorily before tissue samples can be made readily available.

Important to the process of maintaining patient donor privacy is to ensure the de-identification of personalised clinical information as it is captured from the source documentation at the medical institution prior to research use. This is an area in which informatics tools have a high value added impact in helping to facilitate the scrubbing of raw and unstructured electronic source documentation to ensure a high confidence in the collection and storage for de-identified clinical information. HIPPA compliance standards and 21 CFR 11 rules as they apply to drug discovery and clinical genomics research are continuing to evolve. The more flexible a system's architecture can be in supporting audit reports, source documentation archiving, automated scheduling of de-identification checks and electronic submission of study results, the greater likelihood that it will meet the ever increasing regulatory challenges facing the collection of human materials as they are used in drug discovery, applied clinical trials and post marketing surveillance study research.


## Acknowledgements

We would like to thank Pascal Laeng, PhD for use of the figures on human differentiated neuronal stem cells and Tracy Young for help in preparation of this article. DDW



***Dr Michael Palfreyman*** *has been a seasoned leader in the biotechnology and pharmaceutical industries for more than 25 years. He is Founder, President and Chief Scientific Officer at Psychiatric Genomics, Inc. From 1994-2000 he was Senior Vice President of Research & Development at the biotechnology company Scriptgen Pharmaceuticals. Before joining Scriptgen in 1994, he held a series of executive positions at Marion Merrell Dow Research Institute (now Aventis), including Head of CNS Research, Director of Pharmacological Sciences, Vice-President of Global Biological and Scientific Affairs and Vice-President of Research. Michael's PhD and DSc are from the University of Nottingham in neuroscience and rational drug design, respectively. He is the author of 33 patents and more than 150 scientific articles and book chapters.*

***Dr Vinod Charles*** *is a Research Scientist at Psychiatric Genomics, Inc managing the Biorepository and* in vivo *Target Validation groups. Prior to joining Psychiatric Genomics, he served four years as a PRAT Staff Fellow in Neurogenetics at the National Human Genome Research Institute at the National Institutes of Health. Dr Charles earned his PhD from Rush Medical College in Chicago in the field of Systems Neuroscience studying neuronal rescue strategies for Huntington's disease. His work has primarily centred on developing novel pharmacotherapies, neurotransplantation models and mechanisms of neurodegenerative diseases.*

***Jeffrey Blander*** *has specialised in the design and implementation of automated clinical outcomes and applied clinical research systems for more than 10 years. He has worked extensively in interfacing between academic investigators, commercial researchers and IT product development specialists to support the drug discovery and regulatory approval processes within the US, Europe and SE Asia. Jeffrey lectures on technology and business at Harvard and MIT and is the recent co-author of the book* Business Expectations: Are You Using Your Technology to Its Fullest. *JWiley: ISBN 0-471-20834-5.*