Novel genomic methods for drug discovery and mechanism based toxicological assessment

Timothy W. Gant*, Shu-Dong Zhang and Emma L. Taylor

Address: Systems Toxicology Group, Medical Research Council Toxicology Unit, University of Leicester, Lancaster Road, Leicester LE1 9HN.

*Email: twg1@le.ac.uk

Abstract

Genomics today encompasses a range of powerful technologies which find application at all levels of gene expression from transcription to mRNA translation. Collectively, these technologies have great potential for improving drug discovery, both target and molecule recognition, and development. In this article we review the current and potential future status of established and novel genomic methods within drug discovery.

Introduction

Genomics is a tool that has been available to the scientific community for over a decade and was developed from the success of DNA sequencing technology [1]. Introduction of the microarray was a direct response to the wealth of cDNA probes that were first made available from mRNA library sequencing that demanded a high throughput format for their use. Sub-sequentially as whole genome information became available these cDNA probes were replaced by oligos designed against parts of the whole genome [2]. Therefore, though the microarray in various formats has been the technology driver, it is in essence only a small platform to allow the simultaneous use of multiple probes. Prior to the introduction of mRNA libraries and genome sequencing, the number of probes available to the research community did not justify a high density format such as the microarray. A second driver for genomics was the parallel development of computing resources which gave the required power of analysis, storage and transfer of data that allowed genomic science to develop. The two technologies were partners in enabling us to reach this 21st century point where the development of large quantities of information about gene expression in a given system is now a routine event [3]. What is still not such a routine event however is the interpretation of genomic data. Despite the access to numerous bioinformatic tools interpretation is still a human activity assisted by computation.

There are three major arms to the application of genomics in new drug research and deployment; 1) discovery, 2) development and 3) personalisation. The first involves using the profiles of altered gene expression to find new targets and/or molecules, the second the assessment and evolution of these molecules into drugs, and finally matching individual to

drug [4,5]. This is often called personalising medicine and the desired benefits are two-fold, first a reduction in toxicity, particularly of adverse events, and second making sure the drug is likely to be effective [6]. For example this means ensuring that the drug target is expressed in the individual disease state. This is particularly important in diseases that can have very individual characteristics such as cancer.

From the birth of genomics there was a great deal of interest in its potential to resolve difficult problems in drug development, in particular forecast of potential toxicity and species extrapolation [7] [8]. The most optimistic predictions stated that toxicity assessment in drug development would become an activity undertaken using genomics and *in vitro* systems alone early on in the development process. The data obtained through pattern matching to a database would be so indicative of the potential and type of toxicity associated with a new molecule in man that other forms of toxicity assessment would be unnecessary. Thus, those compounds showing an 'adverse' gene expression profile would be deemed unsuitable for further development, while those that were 'clean' would advance [9]. The reality has been a little different and we are now at a point where genomics technologies are not generally viewed as a replacement science, but rather as a powerful tool in the armoury to assist in overall compound evaluation [10]. In addition, genomics enables the discovery and development of new biomarkers that give better warning of the occurrence of (adverse) reactions.

Genomics has to date focussed more on the analysis of mRNA levels (transcriptomics). However the application of genomics within biology and drug development can be extended beyond this level [11]. Transcriptomics measures mRNA levels, which can be altered either by increased transcription or decreased mRNA degradation rate. Microarrays do not differentiate between these two possibilities. However another important mechanism of the regulation of gene expression has recently been discerned and occurs at the level of mRNA translation. Interest in this area has been driven primarily from the discovery of microRNA (miRNA) species and their functions in cellular biochemistry [12]. These small RNA species, while transcribed from the genome in the same manner as protein coding genes are not themselves transcribed into protein. Instead they control the rate of translation of mRNAs [13]. Genomics can be applied to both miRNA expression and mRNA translational analysis allowing a more complete picture of gene expression and regulation to be incorporated into the processes of drug discovery and development.

This review is divided into subsections based on the major potential impact areas of genomics in drug discovery and follows the regulation of gene expression from the genome to the ribosome. As the technical processes of microarray manufacture, labelling etc are now well established these methodologies are not covered here except in brief for the less familiar analysis of miRNA species and mRNA translation. Figure 1 and table 1 present summaries of all the methods discussed here.

Application of genomic technologies in drug discovery and development.

Array comparative genome hybridisation (ArrayCGH)

Array CGH is the process by which gene copy number variations (CNVs) in DNA are identified on a whole or partial genome basis [14] [15]. Hybridised to the array is genomic DNA isolated from either fresh tissue or frequently now from archived samples. The degree of resolution for genomic change is dependent on the length of the probes on the microarray. Short probes spaced close together give the greatest resolution while longer probes such as those from bacterial artificial chromosome (BACS) give greater coverage [16]. The technique has been extensively employed in the identification of genomic change associated with tumor formation and progression, for example in breast carcinomas [17] [18]. Furthermore, the method has been utilised to map the extent of human CNVs that may contribute to adverse drug reactions.

What is the value of ArrayCGH in drug discovery? First is the applicability of the method to archived tissue. Many valuable human studies are stored as archived paraffin blocks and extracting usable mRNA from these blocks for an expression analysis is difficult. Even when mRNA extraction is possible it is never an absolute certainty that the composition of the complex mRNA has not been altered by selective degradation, which will give rise to variable and false positive results. DNA is however a stable molecule and relatively easily extracted from archived paraffin block samples [19]. This gives rise to the possibility of screening such samples for genomic changes that may indicate new drug targets. The expression of genes of interest from such an analysis can be verified later using fresh tissue using fewer samples than might otherwise have been required.

A potential application of ArrayCGH in drug development is for the recognition of aneugens and clastogens. Both types of chemical can give rise to numerical changes in gene numbers on affected chromosome(s). One example of a chemical giving rise to a change in copy number is 4NQO, which increases the copy number of the SV40 gene in immortalised CO631 cells [20]. While this study was not conducted using an array analysis it indicates the potential applicability of ArrayCGH in this aspect of drug development.

Recent studies employing the ArrayCGH method have shown that about 12% of the human genome carries CNVs [21], which could lead to differences in gene expression and potentially phenotype. Several disease states (review [22]) and acquired phenotypes such as drug resistance in cancer [23] are linked to CNV. CNVs have also been found in inbred laboratory mice [24]. The effect of copy number and subsequent gene dosage in mediating adverse drug reactions has not been investigated to any great extent though one partial deletion in Cyp2b6 has been described [25]. These data suggest that CNV alterations in genes responsible for metabolism of drugs could be responsible for altering pharmacokinetics, efficacy and may be partially or wholly responsible for some adverse drug reactions. This remains a substantially underexplored area but appears of critical importance both for drug development and usage in terms of personalized medicine.

Epigenetics

The importance of epigenetic change in controlling the expression of genes has recently moved to the forefront of genomic analysis with the development of high throughput sequencing methods directed to the recognition of epigenomic change, in particular alteration of DNA methylation patterns [26]. Application of these methods led to the first full chromosome methylation analysis of three human chromosomes and represented a milestone in the analysis of epigenetic change [27]. On an array based format the technique of methyl cytosine (MeDIP) immunopreciptation coupled with analysis using a promoter region microarray allows the identification of those regions of DNA that have undergone methylation change in response to a altered state or drug [28]. However if changes in specific genes are of interest then direct sequencing combined with bisulfite treatment is probably a better option [29]. Global patterns of DNA methylation change on genes will be reflected in their gene expression profiles where transcriptomics has a major role to play in assessing the magnitude and type of these effects. Methylation changes are associated with several disease states and are frequently altered during cancer progression [30] [31]. DNA methylases mediating DNA methylation therefore make promising drug targets and several small molecule inhibitors have been developed [32] [33].

Transcriptional genomics

The most widely applied of the genomics methods in drug development has been the analysis of gene transcription (mRNA levels) [11]. Whether using a one or two color system the approach is the same. The probe is a fluorescently labelled complex mixture of cRNA or cDNA derived from mRNA that is hybridised to an array containing the targets. Hybridisation is detected and quantitated by measuring the degree of fluorescence associated with each target [34,35]. Analysis of the resulting data has typically taken one of three protocols. First global analysis for profiles indicative of an event using multivariate mathematical techniques such as clustering or principal component analysis, or other network or self learning type approaches. Second, for single differentially expressed genes whose altered expression in a disease state can act both as biomarker of the disease and as a prognostic or pathological indicator. The protein products of these genes may offer the prospect of a new 'druggable' target. Finally pathway analysis where genes are treated in groups according to their roles in established biochemical pathways. There is overlap from the first analysis to the second, as multivariate approaches can successfully identify single genes which are discriminatory and potentially 'druggable'. This approach requires the application of rigorous statistical tests to avoid false positives[36]. All these methods have the potential to deliver new drug targets through the exploration of differences in gene expression, for example a normal cell and one that is transformed. This is the process of target identification. If a target is validated then the process of lead generation and optimization can begin. There is often debate amongst genomic practitioners about the right approach to data mining. The opinion of these authors is that there is no single 'right' method. There is only a scale of appropriate

methods, the order of which will depend on the question being asked and the amount and type of the data.

One such example of target identification is OGN (osteoglycin). Using transcriptomic analysis OGN expression was recently shown to be correlated with left ventricular cardiac mass (LVM). Increased LVM leads to cardiac dysfunction. The close association of OGM expression with LVM identifies it as a potentially druggable target for LVM therapy. The association of OGN with LVM relied on an inherited trait in a model (the rat) which could be accurately measured and then developing an expression quantitative trait locus (eQTL) whose significance of association with the phenotype could be ascertained [37].

The above example illustrates how identification of a potentially druggable target in a disease can be undertaken. However can transcriptomics also be used to actually identify molecules that may have potential as drugs? Traditional drug development against a target such as OGN takes the form of a cell line being developed with expression of high levels of the target of interest usually coupled to a reporter system or assay. Compound libraries are applied, and molecules with pharmacological activity identified by the read out from the reporter. This method however has a fundamental disadvantage in drug discovery; it is a single target screening approach. If it is subsequently decided that another target is of interest then screening has to take place again utilising the new target. A more rational approach would be a generic screen utilising a non-hypothesis driven approach with the potential of identifying different pharmacologies.

In 2006 Lamb *et al* [38] described such a method where gene expression profiles could be used to link pharmacological profiles to recognise novel pharmacology. This is known as the connectivity map and has been reviewed by Micknich [39]. Screening takes place in cell lines that have not usually been engineered to overexpress any form of drug target. Gene expression profiles in response to the chemical are derived at multiple concentrations and then ordered by up and down regulated genes. These are then matched to a database and commonality is indicated by the number of matched genes (Figure 2). Lamb *et al* [38,40] showed the applicability of this method to a number of drug molecules. The method of analysis was reviewed and modified by Zhang and Gant [41] to provide more statistical rigor. Recently the database of compounds has been increased substantially by Lamb et al (http://www.broad.mit.edu/cmap/). The net result of these efforts is the provision of a resource that allows non-hypothesis driven drug discovery and therefore has the potential to recognise novel pharmacology, or off target pharmacology in existing molecules. Zhang and Gant [41] tested the method on the estrogen receptor modulators and HDAC inhibitors using the data originally collected by Lamb and new data collected from public sources as the query signature. The new data was successfully used to identify not only estrogen receptor agonists from source data but also antagonists by virtue of negatively correlated gene expression maps. This can be argued to be a greater achievement than that with the estrogen receptor active compounds since the alteration of gene expression by HDAC

inhibitors may potentially be more generic than that with estrogen receptor modulators, and therefore difficult to identify. This testing shows the potential of the connectivity mapping for drug discovery.

Similar methods have the potential to recognise adverse toxicology in new chemical entities (NCE). For example, it is possible that genotoxins may produce a common gene expression signature that can be used to recognized this particular activity in a NCE. However, the problem of generic signatures associated with cellular toxicity may render this analysis not possible. Thus connectivity analysis may only work acceptably when operated within an environment of a compound/nuclear receptor interaction where a specific set of genes under the control of the transcription factor recognition site will be activated. These hypotheses remain to be explored, but initial data derived by Zhang (not published) shows that the method has applicability. There is also a question of whether the biological system will allow the correct matching. For example, a compound that requires metabolic activation for an effect may not map correctly when the gene expression profile is generated in a non-metabolically competent cell line. This is likely to be a problem more relevant to the recognition of toxicology using this method than recognition of pharmacology where activity is more often a property of the parent molecule. The limits of the methods will only be discovered with extensive further testing.

mRNA splicing

Alternated RNA transcript splicing is a common feature of disease states and results in altered proteins [42]. Differential splicing can, with care, be detected using microarrays utilising the same methods as for transcriptomics (Figure 3). The difference is that the microarray will contain multiple probes for each genes directed against the exons known or hypothesised to be differentially spliced [43](Figure 3). There are two major applications of such technology in drug discovery and development. If a particular splice variant of a gene is associated with a specific or particular disease then there is the potential to develop a molecule to target the variant protein. Given that virtually every gene has the potential to be alternatively spliced into multiple forms this area of disease biology greatly expands the portfolio of potential drug targets. However a great deal of further analysis utilising the methods above needs to be carried out to find and characterise those alternatively spliced mRNAs suitable as targets. This approach could also be used to screen for molecules that modulate alternative splicing, a recent example of which is digoxin [44]. Furthermore polymorphic profiles in splice variants can give rise to differential pharmacokinetics which can cause adverse drug reactions.

Antisense and RNAi

Identification of a transcript closely associated with a disease represents the possibility of a druggable target, and connectivity analysis the possibility of identifying a small molecule that may display the requisite pharmacological activity for the target [45]. Another approach

to modulation of the target is through the use of a biological therapeutic agent. One such biological drug approach is the direct antisense targeting of an mRNA strand using an oligonuleotide with the goal of causing mRNA degradation through activation of RNAses. Another approach is to use RNAi technology (RNAi = siRNA). Instead of using a single strand antisense molecule, RNAi utilizes a short (typically 23bp) double stranded RNA molecule which is activated by Dicer and incorporated into the RISC complex [46]. Upon binding the mRNA is then targeted for degradation by RNAses. Despite the appeal of antisense RNAi therapeutic theory only one such molecule has made it to the clinic, Fomivirsen (5'-GCGTTTGCTCTTCTTCTTGCG-3') which is an antisense oligo used for the treatment of cytomegalovirus retinitis [47]. Many more however are in clinical trials [48].

Transcriptomics has particular application in the development of direct mRNA targeting therapeutic molecules, not just for the target identification but also for the recognition of off-target, or downstream mRNA effects. Similarly if downstream effects occur, for example on transcription rates, then these will be also easily recognised using whole genome profiling. RNAi transcriptomic analysis will be particularly important for RNAi technology where the mechanism of action relies on amplification through the RISC complex. This mechanism can be saturated *in vivo* preventing the normal processing of miRNA (the physiological equivalent of RNAi) species with undesirable consequences [49]. Both transcriptomic profiling of mRNA species and miRNA (see below) have potential in drug development for recognising these undesirable off target RNAi effects.

miRNA and mRNA translation analysis in drug development

miRNA species are the physiological equivalent of RNAi though their mechanism of action is a little different. miRNAs do not require complete homology to the mRNA sequence, utilising instead a 7-8bp target sequence in the 3'UTR of the mRNA to which they bind through the miRNA-RISC complex [50]. This leads to suppression of mRNA translation by repression of initiation or elongation [51] [13] rather than causing the mRNA to be targeted for degradation.

These small RNA species are transcribed from the genome in the same manner as protein coding genes and are under the control of many of the same transcriptional regulatory mechanisms [52]. Once processed and transported to the cytoplasm (for a review of miRNA biochemistry see [12] [53]) they have the potential to alter the translation of mRNA species to which they have complementarity. Thus, these genes have network properties whereby a differential expression in a miRNA can affect the protein expression of many downstream genes. The effect of this translational suppression on specific protein levels has been elegantly shown in two recent proteomic studies [54] [55]. Many studies are now demonstrating the importance of miRNA expression in normal physiology and development and in the alteration of gene expression during xenobiotic exposure and disease progression [56]. Genomic methods similar to those used for transcriptomic analysis can be used to identify differentially expressed miRNA species in a cell or tissue. The only essential

differences in the technique are that the labelling is altered to a form of end labelling because the miRNA species are short and lack a poly A tail [57] [58], and probes on the microarray can be altered to increase the stability of the hybridisation kinetics, for example by including locked nucleic acids [59,60]. Standard genomics methods can therefore be used to recognise druggable miRNA species. Furthermore, these species can be utilised as targets relatively simply by using the same technology as that for antisense methods. Antisense species to miRNA species are called 'Antagomirs' [61]. Additionally the miRNA may itself be employed as a drug. For example, expression of several miRNA species are downregulated in inflammatory skin conditions such as Psoriasis [62]. This suggests that these miRNA species may be involved with the normal physiological suppression of genes involved in the inflammatory process, and that if replaced may act to reduce the inflammation in such conditions by suppressing translation of the inflammatory genes [63]. This type of approach has been used successfully in cardiac hypertrophy [64] and the whole miRNA field has huge potential for future drug development built on a miRNA genomic analysis foundation [65].

Translationalomics

Using transcriptomics to verify any off-target effects of a miRNA therapy is not as simple as that for siRNA because most miRNA species suppress translation and so do not cause degradation of the mRNA. Therefore off-target mRNAs may not be detected by a microarray measure of mRNA levels. What is required in this instance is a measure of mRNA translation. This can be achieved by using sucrose density gradients to separate the mRNA species under active translation followed by analysis on microarrays [66]. Great care has to be taken with the experimental design and in particular data normalisation to prevent the mistaken identification of transcriptional rather than translational events [67]. With this caveat however it is possible to use such methods to identify both on, and off, target mRNA transcripts undergoing translational regulation in response to the use of miRNAs or antagomirs as drugs.

It is also becoming clear, primarily driven by research into the physiological roles of miRNA, that translational control represents a fundamental process in normal development and can be disrupted in disease states. Using the above technique it has been shown that mir21 plays a role in the progression of colorectal cancer to a metastatic phenotype [68]. Furthermore Gabriely *et al* demonstrated that mir21 targets matrix metalloproteinases and by doing so promotes glioma invasion [69]. Therefore as proposed by these authors mir21 appears to be a very good target for antagomir drugs. There is highly likely to be a great deal more development in this field and novel genomic methods will have a significant role to play in both discovery and development.

Proteins, cell and plasmid 'omics.

Finally, there is another whole world of arrays that are constructed with plasmids, proteins (antigens and antibodies), glycoproteins, cells and even chemicals. All of these have

application in drug discovery and development, but are outside the scope of genomics. Many of these applications have been reviewed by Angres [70] and the interested reader is referred to this article.

Conclusion

Genomic analysis applied at the levels of the genome, mRNA and miRNA transcription and mRNA translation has application in all areas of drug discovery and development. Firstly for the recognition of 'druggable' targets, secondly for the identification of pharmacological potential, thirdly for identification of off target effects, and finally for the recognition and elucidation of toxic mechanisms. It is hard to see how a feasible drug discovery and development program can proceed without application of genomics in one or more of these areas.

Acknowledgement.

Thanks are due to the systems toxicology group whose hard work enables us to apply genomics in many areas of work and explore the boundaries of the methodology. In addition to the authors they are: Kate M. Phillips, Amy V. Pointon, JinLi Luo and Joan Riley.

Method	Technology	Use in drug discovery	Use in drug development	References
Methylation analysis	Methylation immunoprecipitation and microarrays/bisulfite sequencing	New targets – disease related gene methylation changes	Off target effects of drugs in the epigenome	[26,28,29]
ArrayCGH	Microarray analysis of genome duplications	New targets - Disease related alteration of gene copy number	Detection of genotoxins causing aneuploidy or clastogenesis	[11,14,16,18]
Transcriptomics	Microarray analysis with conserved exon sequence targets	Discovery of new drug targets	Early detection of toxicity or unexpected pharmacology	[71-73]
Splice variation	Microarray analysis of target sequences for splice variants	Discovery of new drug targets	Adverse splicing related to drug exposure	[74-76]
miRNA analysis	Microarray analysis for mature miRNAs	Disease specific miRNAs can act as targets for biomolecule drugs 'Antagomirs'	Altered miRNA changes related to drug exposure	[60,61,67,77]
mRNA translation	Density gradient and microarray based analysis of mRNA translation	Discovery of new targets; validation of mRNA translation	Altered translation resulting from cellular stress reactions/toxicity	[78-80] [81,82]

Table 1. Applications of genomic technologies in drug discovery and development.

Figure legends

Figure 1 – From genome to protein. Control stages in the regulation of gene expression where different genome technologies can be applied in either a drug discovery or development mode.

Figure 2 – Connectivity analysis. Pictorial representation of the manner in which connectivity analysis can be performed for the discovery of novel activity in small molecules using gene transcription profiles.

Figure 3 – Splice variants as drug targets. Differential gene splicing during disease states giving rise to altered proteins represents a major area for the discovery of new drug targets. Splice variation can be discovered using sequences directed against different exons of a protein coding gene and microarrays provide the ideal format for the use of these target.

References

1. Ewis AZ, Z, Bakalova R, Fukuoka S, Shinohara Y, Ishikawa M, Baba Y: **A history of microarrays in biomedicine**. *Expert reviews in molecular diagnosis* 2005, **5**:315-328.
2. Wixon j: **Genome sequencing and bioloy 2001** *Comparative and Functional Genomics* 2001, **2**:246-251.
3. Yu U, Lee S, Kim Y, Kim S: **Bioinformatics in the post-genome era**. *Journal of biochemisty and molecular biology* 2004, **37**:75-82.
4. Workman P: **The opportunities and challenges of personalized genome-based molecular therapies for cancer: targets, technologies, and molecular chaperones**. *Cancer Chemother. Pharmacol.* 2003, **52**:S45-S56.
5. Workman P, deBono J: **Targeted therapeutics for cancer treatment: major progress towards personalised molecular medicine**. *Current opinion in Pharmacology* 2008, **In Press**.
6. Collins I, Workman P: **New approaches to molecular cancer therapeutics**. *Nature Chemical biology* 2006, **2**:689-700.
7. Albertini S: **Toxicogenomics in the pharmaceutical industry: hollow promises or real benefit?** *Mutation Research* 2005, **575**:102-115.
8. Gant TW: **Application of toxicogenomics in drug development**. *Drug News & Perspectives* 2003, **16**:217-221.
9. Luhe A, Suter L, Ruepp S, Singer T, Weiser T, Albertini S: **Toxicogenomics in the pharmaceutical industry: hollow promises or real benefit?** *Mutation Research* 2005, **575**:102-115.
10. Lord P, Nie A, McMillian M: **Application of Genomics in Preclinical Drug Safety Evaluation**. *Basic and clinicl pharmacology and toxicology* 2005, **98**:537-546.
11. Gant T: **Novel and future applications of microarrays in toxicological research**. *Expert opinion in drug metabolism and toxolcology* 2007, **3**:599-608.
12. Kim V, Nam J-W: **Genomics of microRNA**. *Trends in Genetics* 2006, **22**:165-176.
13. Valencia-Sanchez M, Liu J, Parker R: **Control of translation and mRNA degradation by miRNAs and siRNAs**. *Genes and Development* 2006, **20**:515-524.
14. Barrett MT, Scheffer A, Ben-Dor A, Sampas N, Lipson D, Kincaid R, Tsang P, Curry B, Baird K, Meltzer P, et al.: **Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA**. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**:17765-17770.
15. Pinkel D, Seagraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, et al.: **High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays**. *Nature Genetics* 1998, **20**:207-211.
16. Pinkel D, Albertson D: **Comparative genomic hybridization**. *Annual reviews in genomics and human genetics* 2005, **6**:331-354.
17. Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, Ringner M, Sauter G, Monni O, Elkahloun A, et al.: **Impact of DNA amplification on gene expression patterns in breast cancer**. *Cancer Research* 2002, **62**:6240-6245.
18. Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, Brown PO: **Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors**. *Proceedings of the National Academy of Sciences of the United States of America* 2002, **99**:12963-12968.
19. Joosse S, vanBeers E, Nederlof P: **Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material**. *BMC Cancer* 2007, **7**:43-54.
20. Fahrig F, Steinkamp-Zucht A: **Induction or suppression of SV40 amplification by genotoxic carcinogens, non-genotoxic carcinogens or tumor promoters**. *Mutation Research* 1996, **356**:217-224.

21. Redon R, Ishikawa S, Fitch K, Feuk L, Perry G, Andrews T, Fiegler H, Shapero M, Carson A, Chen W, et al.: **Global variation in copy number in the human genome**. *Nature* 2006, **444**:444-454.
22. Beckmann J, Estivill X, Antonarakis SE: **Copy number variants and genetic traits:closer to the resolution of phenotypic to genotypic variability**. *Nature reviews* 2007, **8**:639-645.
23. Turton NJ, Judah DJ, Riley J, Davies R, Lipson D, Styles JA, Smith AG, Gant TW: **Gene expression and amplification in breast carcinoma cells with intrinsic and acquired doxorubicin resistance**. *Oncogene* 2001, **20**:1300-1306.
24. Egan C, Sridhar S, Wigler M, Hall I: **Recurrent DNA copy number variation in the laboratory mouse**. *Nature Genetics* 2007, **39**:1384-1389.
25. Rotger M, Saumoy M, Zhang K, Flepp M, Sahli R, Decosterd L, Telenti A, Study SHC: **Partial deletion of CYP2B6 owing to unequal crossover with CYP2B7**. *Pharmacogenetics and Genomics* 2007, **17**:885-890.
26. Beck S, Rakyan V: **The Methylome: Global approaches for DNA methylation profiling**. *Trends in Genetics* 2008, **24**:231-237.
27. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, et al.: **DNA methylation profiling of human chromosomes 6, 20 and 22**. *Nature Genetics* 2006, **38**:1378-1385.
28. Jacinto F, Ballestar E, Esteller M: **Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome**. *BioTechniques* 2008, **44**:35-39.
29. Schilling E, Rehli M: **Global, comparative analysis of tissue-specific promoter CpG methylation**. *Genomics* 2007, **90**:314-323.
30. Sawan C, Vaissiere T, Murr R, Herceg Z: **Epigenetic drivers and genetic passengers on the road to cancer**. *Mutation Research* 2008, **642**:1-13.
31. Jones P, Baylin S: **The epigenomics of cancer**. *Cell* 2007, **128**:683-392.
32. Ghoshal K, Bai S: **DNA methyltransferase as targets for cancer therapy**. *Drugs of Today* 2007, **43**:395.
33. Egger G, Liang G, Aparicio A, Jones P: **Epigenetics in human disease and prospects for epigenetic therapy**. *Nature* 2004, **429**:457-463.
34. Lockhart D, Winzeler E: **Genomics, gene expression and DNA arrays**. *Nature* 2000, **405**:827-836.
35. Schena M, Heller RA, Theriault TP, Konrad K, Lachenmeier E, Davis RW: **Microarrays: biotechnology's discovery platform for functional genomics**. *Trends in Biotechnology* 1998, **16**:301-306.
36. Zhang S-D, Gant TW: **A statistical framework for the design of microarray experiments and effective detection of differential gene expression**. *Bioinformatics* 2004, **20**:2821-2828.
37. Petretto E, Sarwar R, Grieve I, Lu H, Kumaran M, Muckett P, Mangion J, Schroen B, Benson M, Punjabi P, et al.: **Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass**. *Nature Genetics* 2008, **40**:546-552.
38. Lamb J, Crawford E, Peck D, Modell J, Blat I, Wrobel M, Lerner J, Brunet J, Subramanian A, Ross K, et al.: **The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.** *Science* 2006, **313**:1929-1935.
39. Micknich S: **The connectivity map**. *Nature Chemical biology* 2006, **2**:663-664.
40. Lamb J: **The Connectivity Map: a new tool for biomedical research**. *Nature reviews Cancer* 2007, **1**:54-60.
41. Zhang S, Gant T: **A simple and robust method for connecting small-molecule drugs using gene-expression signatures**. *Bmc Bioinformatics* 2008, **9**:258-268.
42. Wang G, Cooper T: **Splicing in disease: disruption of the splicing code and the decoding machinery**. *Nature Revews Genetics* 2007, **8**:749-761
43. Fehlbaum P, Guihal C, Bracco L, Cochet O: **A microarray configuration to quantify expression levels and relative abundance of splice variants**. *Nucleic Acids Research* 2005, **33**:e47-e58.

44. Stoilov P, Lin C-H, Damoiseaux R, Nikolic J, Black D: **A high-throughput screening strategy identifies cardiotonic steroids as alternative splicing modulators**. *Proceedings of the National Academy of Sciences of the United States of America* **32**:11218-11223.
45. Corey D: **RNA learns from antisense**. *Nature Chemical biology* 2007, **3**:8-11.
46. Achenbach T, Brunner B, Heermeier K: **Oligonucleotide-based knockdown technologies: antisense versus RNA interference**. *Chembiochem* 2003, **4**:928-935.
47. Geary R, Henry S, Grillone L: **Fomivirsen: clinical pharmacology and potential drug interactions.** *Clinical Pharmacokinetics* 2002, **41**:255-260.
48. Rayburn E, Zahng R: **Antisense, RNAi, and gene silencing strategies for therapy: Mission possible or impossible?** *Drug Discovery Today* 2008, **13**:514-521.
49. Grimm D, Streetz K, Jopling C, Storm T, Pandey K, Davis C, Marion P, Salazar F, Kay M: **Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways**. *Nature* 206, **441**:537-541.
50. Lewis B, Shih I-h, Jones-Rhoades M, Bartel D, Burge C: **Prediction of Mammalian MicroRNA Targets**. *Cell* 2003, **115**:787-798.
51. Chu C-Y, Rana T: **Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54.** *PLOS biology* 2006, **4**:1122-1136.
52. Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek S, Kim V: **MicroRNA genes are transcribed by RNA polymerase II**. *EMBO Journal* 2004, **23**:4051-4060.
53. He L, Hannon G: **microRNAs: Small RNAs with a big role in gene regulation**. *Nature Genetics* 2004, **5**:522-531.
54. Baek D, Villén J, Shin C, Camargo F, Gygi S, Bartel D: **The impact of microRNAs on protein output**. *Nature* 2008, **455**:64-71.
55. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N: **Widespread changes in protein synthesis induced by microRNAs**. *Nature* 2008, **455**:58-63.
56. Pasquinelli A, Hunter S, Bracht J: **MicroRNAs: a developing story**. *Current Opinion in Genetics and Development* 2005, **15**:200-205.
57. Thompson J, Parker J, Perou CM, Hammond SM: **A custom microarray platform for analysis of microRNA gene expression**. *Nature Methods* 2004, **1**:47-53.
58. Thompson J, Parker J, Hammond S: **Microarray analysis of miRNA gene expression**. *Methods in Enzymology* 2007, **427**:107-119.
59. Castoldi M, Schmidt S, Benes V, Noerholm M, Kulozik A, Hentze M, Muckenthaler M: **A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA)**. *RNA* 2006, **12**:913-920.
60. Castoldi M, Schmidt S, Benes V, Hentze M, Muckenthaler M: **miChip: an array-based method for microRNA expression profiling using locked nucleic acid capture probes**. *Nature Protocols* 2008, **3**:321-329.
61. Hammond S: **MicroRNA therapeutics: a new niche for antisense nucleic acids**. *Trends in Molecular Medicine* 2006, **12**:99-101.
62. Sonkoly E, Wei T, Janson P, Sääf A, Lundeberg L, Tengvall-Linder M, Norstedt G, Alenius H, Homey B, Scheynius A, et al.: **MicroRNAs: novel regulators involved in the pathogenesis of Psoriasis?** *Plos one* 2007, **11**:e610.
63. Sonkoly E, Ståhle M, Pivarcsi A: **MicroRNAs: novel regulators in skin inflammation**. *Clinical and experimental dermatology* 2008, **33**:312-315.
64. Carè A, Catalucci D, Felicetti F, Bonci D, Addario A, Gallo P, Bang M, Segnalini P, Gu Y, Dalton N, et al.: **MicroRNA-133 controls cardiac hypertrophy**. *Nature Medicine* 2007, **13**:613-618.
65. Latronico M, Catalucci D, Condorelli G: **MicroRNA and cardiac pathologies**. *Physiological Genomics* 2008, **34**:239-242.
66. Melamed D, Arava Y: **Genome-wide analysis of mRNA polysomal profiles with spotted DNA microarrays**. *Methods in Enzymology* 2007, **431**:177-201.

67. Taylor E, Gant T: **Emerging fundamental roles for non-coding RNA species in toxicology**. *Toxicology* 2008, **249**:34-39.
68. Asangani I, Rasheed S, Nikolova D, Leupold J, Colburn N, Post S, Allgayer H: **MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer**. *Oncogene* 2008, **27**:2128-2136.
69. Gabriely G, Wurdinger T, Kesari S, Esau C, Burchard J, Linsley P, Krichevsky A: **MicroRNA 21 promotes glioma invasion by targeting matrix metalloproteinase regulators**. *Molecular and Cellular Biochemistry* 2008, **28**:5369-5380.
70. Angres B: **Cell microarrays**. *Expert reviews in molecular diagnosis* 2005, **5**:766-779.
71. Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, Ahr H: **Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver**. *Mutation Research* 2005, **575**:61-84.
72. Gant T, Zhang S-D: **In pursuit of effective toxicogenomics**. *Mutation Research* 2005, **575**:4-16.
73. Sausville E, Holbeck S: **Transcription profiling of gene expression in drug discovery and development: the NCI experience**. *European journal of cancer* 2004, **40**:2544-2549.
74. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C: **Rapid, transcript-specific changes in splicing in response to environmental stress**. *Molecular Cell* 2007, **27**:928-937.
75. Shoemaker D, Schadt E, Armour C, He Y, Garrett-Engele P, McDonagh P, Loerch P, Leonardson A, Lum P, Cavet G, et al.: **Experimental annotation of the human genome using microarray technology**. *Nature* 2001, **409**:922-927.
76. Blencowe B: **Alternative splicing: new insights from global analysis**. *Cell* 2006, **126**:37-47.
77. Lee S, Dutta A: **MicroRNA in Cancer**. *Annual Reviews in Pathology* 2008, **4**:199-227.
78. Reiter A, Crozier S, Kimball S, Jefferson L: **Meal feeding alters translational control of gene expression in rat liver**. *Journal of Nutrition* 2005, **135**:367-375.
79. Shenton D, Smirnova J, Selley J, Carroll K, Hubbard S, Pavitt G, Ashe M, Grant C: **Global translational responses to oxidative stress impact upon multiple levels of protein synthesis**. *Journal of Biological Chemistry* 2006, **281**:29011-29021.
80. Iguchi N, Tobias J, Hecht N: **Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**:7712-7717.
81. Provenzani A, Fronza R, Loreni F, Pascale A, Amadio M, Quattrone A: **Global alterations in mRNA polysomal recruitment in a cell model of colorectal cancer progression to metastasis**. *Carcinogenesis* 2006, **27**:1323-1333.
82. Mazan-Mamczarz K, Kawai T, Martindale J, Gorospe M: **En masse analysis of nascent translation using microarrays**. *Biotechniques* 2005, **39**:61-67.

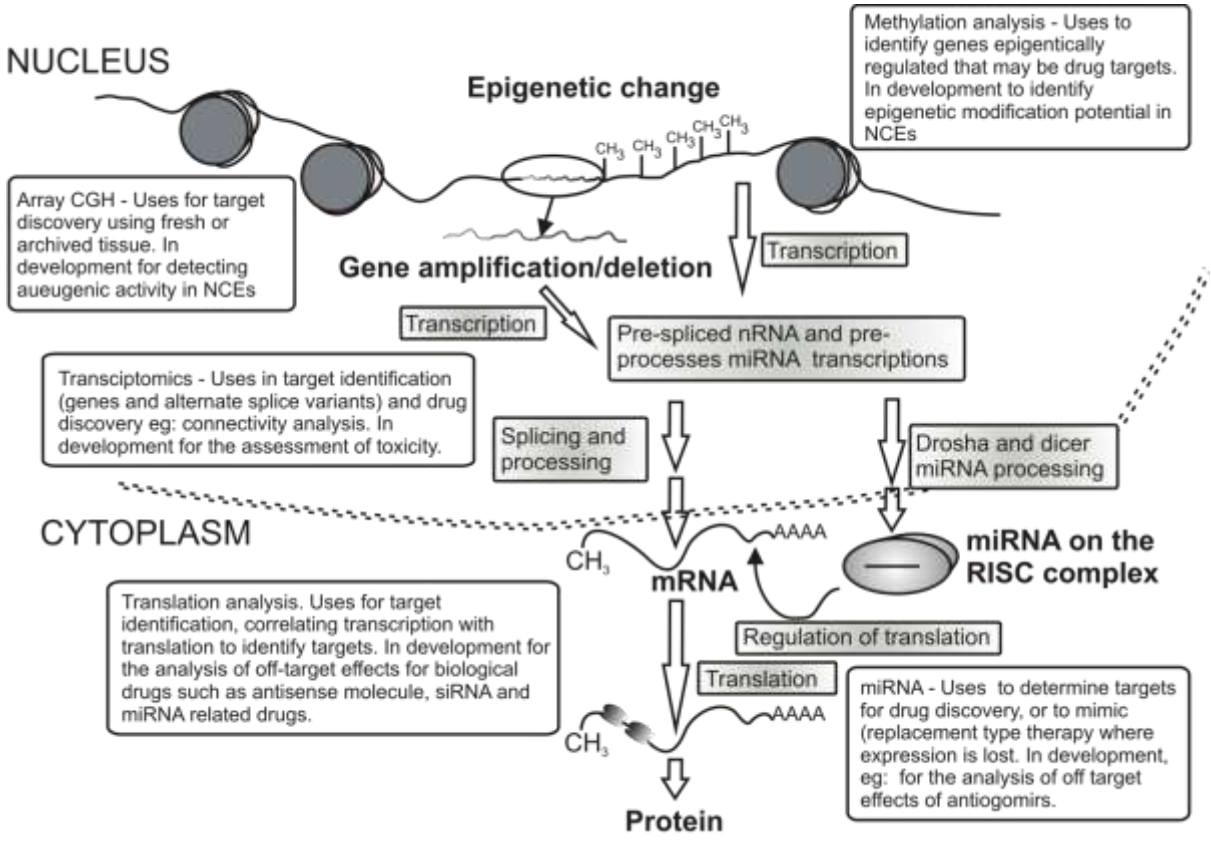

Figure 1

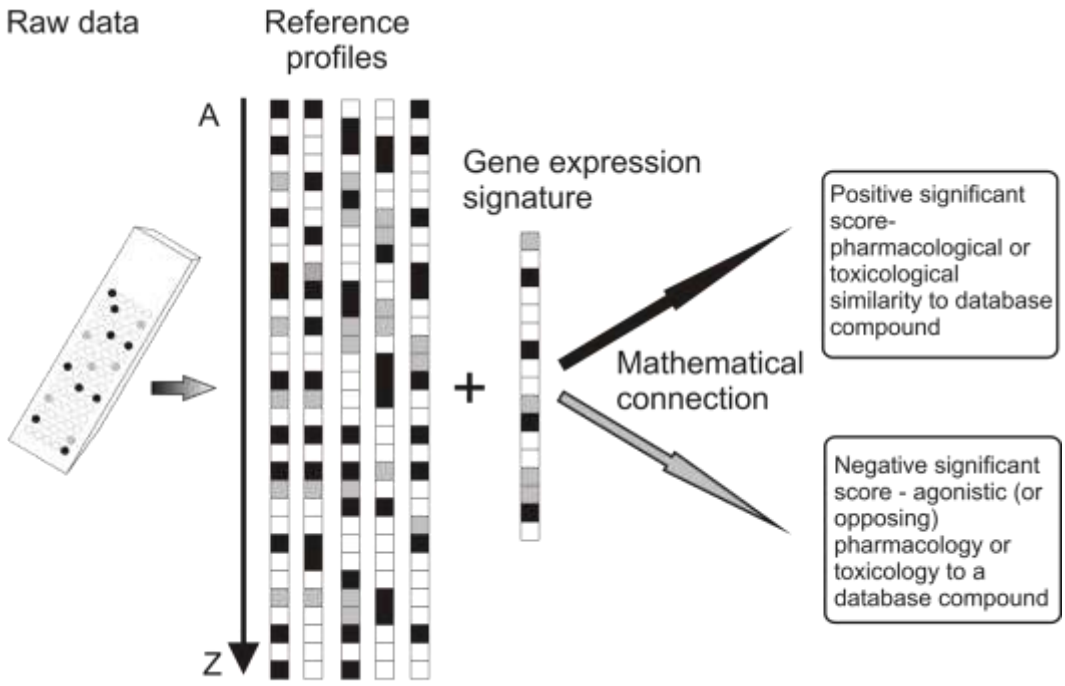

Figure 2

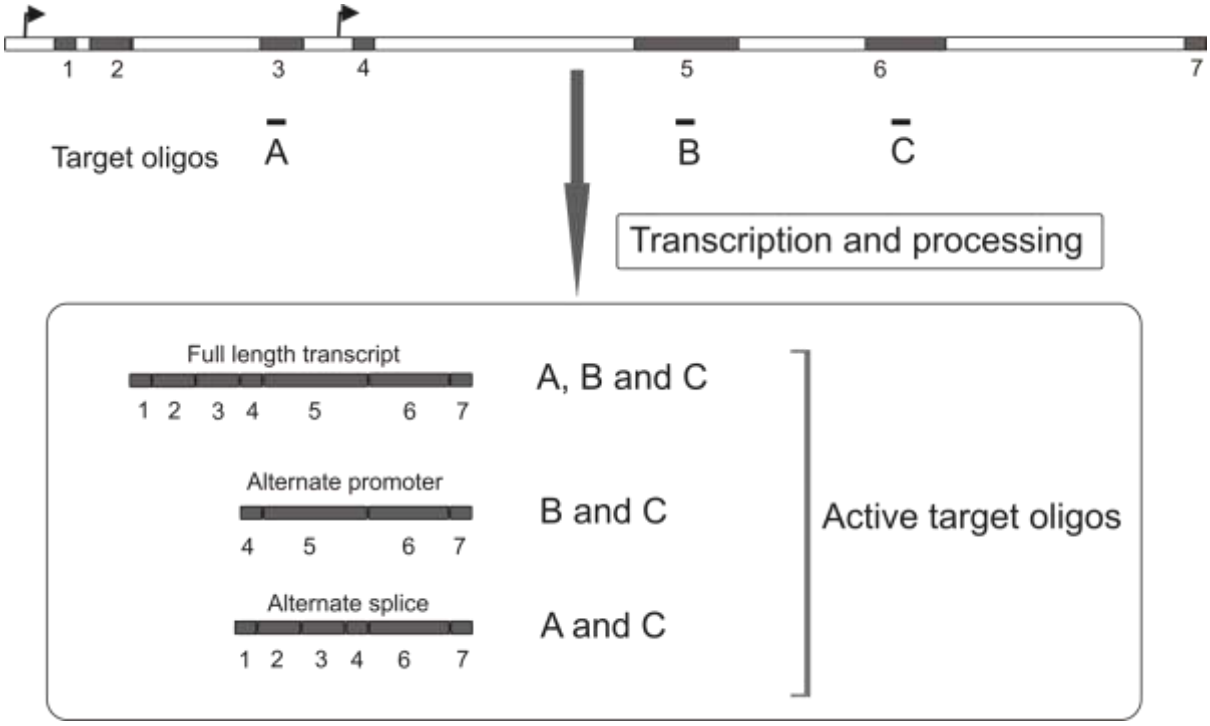
1 2 3 4 5 6 7
Target oligos
A
B
C
Transcription and processing
Full length transcript
1 2 3 4 5 6 7
A, B and C
Alternate promoter
4 5 6 7
B and C
Alternate splice
1 2 3 4 6 7
A and C
Active target oligos

Figure 3