The completed human genome: implications for chemical biology

Christopher P Austin

The recently completed human genome sequence represents an enormous opportunity to understand biology and accelerate the development of new therapeutics. However, it also presents equally large logistical, scientific and paradigmatic challenges to efficiently translate the enormous cache of sequence data into functional information that will be the precursor of new drug development. Small-molecule chemical biology applied on a genomic scale promises to speed this translation to novel therapeutics.

Addresses
National Human Genome Research Institute, National Institutes of Health Building 31, Room 4B09, 31 Center Drive, Bethesda, MD 20892, USA
e-mail: austinc@mail.nih.gov

Current Opinion in Chemical Biology 2003, **7**:511–515

This review comes from a themed issue on Next-generation therapeutics
Edited by John Kozarich and Hugh Rosen

1367-5931/$ – see front matter
© 2003 Elsevier Ltd. All rights reserved.

DOI 10.1016/S1367-5931(03)00083-8

Abbreviations

CFTR	cystic fibrosis transmembrane conductance regulator
FXR	farnesoid X receptor
GPCR	G-protein-coupled receptor
HDAC	histone deacetylase
HGP	Human Genome Project
siRNA	small inhibitory RNA

Introduction

The official completion of the Human Genome Project (HGP) in April 2003 was a landmark event in the history of biology [1,2]. However, with celebrations of this accomplishment came realization that the research community is now faced with the exciting but daunting task of identifying, and assigning function and therapeutic potential to, all ~30 000 genes in the human genome. A plan for beginning to accomplish this is part of a recently published vision statement for the future of genome research [3]. This review discusses a potentially important role for chemical biology in accomplishing these goals.

The genome era

The National Research Council report that first proposed the HGP in 1988 clearly envisioned benefits to human health as an expected outcome [4], and subsequent research plans have reaffirmed this view [5]. However, the perceived value of the human genome sequence has followed an undulating course [6–8]. The problem of translating gene sequences into tangible improvements in human health was initially thought to be relatively straightforward. In the commercial sector, optimism derived from the early success of recombinant protein therapeutics such as human insulin and erythropoietin, and led to ephemeral market success for companies that aimed to capitalize on the genome sequence. As it became apparent that new drugs would not be quickly forthcoming from the HGP, interest shifted to companies offering 'functional genomics' technologies that promised to identify therapeutically important genes for drug development. Many of these techniques found utility in limited classes of gene products [9,10], but none has proven to be the generally useful divination tool that was hoped for. This is perhaps not surprising, since definition of a gene's function and therapeutic potential rarely results from a single approach. The variety of tools and time required to translate a novel gene discovery into a drug has led to suggestions that the HGP will actually slow therapeutic development in the short term [11]. Thus, despite much prognostication that drug development and medicine will be changed dramatically by the HGP, this has for the most part yet to occur [12].

Although access to whole genomes has not yet translated into major therapeutic advances, it has made possible fundamentally new approaches to basic biological research. Where data were traditionally gathered in a hypothesis-driven fashion, much data are now collected systematically for later mining; where experiments previously focused on single genes, many now examine entire networks; and where previously research focused on the identification of novel genes, much work now focuses on determining the function of genes, now that the bounded set of gene products in genomes is becoming defined. Currently, relatively little biological information is known about most genes identified by human genome sequencing [13], and very few genes that confer susceptibility to common diseases (e.g. cardiovascular disease, diabetes, Alzheimer's disease) are known. Bioinformatic querying and comparison of genomes and has identified important novel human genes [14,15], as well as conserved regions that are candidates for yet-undefined protein coding and regulatory sequences [15]. Experimental whole-genome querying has become increasingly common with tools such as RNA interference and targeted mutagenesis, particularly in organisms with smaller genomes such as yeast [16], *Caenorhabditis elegans* [17], and *Drosophila melanogaster* [18•]. In human genetics, the impact of the genomic view is evident in genetic linkage studies, in which a comprehensive list of genes in a

linkage region can now be quickly defined, candidates prioritized, and etiologic gene identified [19], a process that previously required months to years [20]. We have truly entered the Genome Era, and this can and should change how biological research is done.

The genome translation problem

Though the number of human genes is estimated at ~30 000, it is likely that the number of gene products encoded by the human genome is greater than this by at least two to threefold, because of alternative splicing, intergenic recombination, RNA editing and posttranslational modifications [13]. The number of human gene products that will represent therapeutic targets has been the subject of much debate in the literature, and has ranged from <1000 [21] to >10 000 [22], depending on what criteria are used. Ideally, experimental determination of the function of as many gene products as possible will result in the most complete set of potential targets for therapeutic development. However, the number of gene products to be evaluated is so large, and the kinds of data required to validate a target so diverse, that an agenda for functionation of the genome analogous to the HGP itself is needed to accomplish the genome-wide prioritization of gene products appropriate for intervention.

The success of the HGP in scientific and organizational terms provides a template for how this might be achieved. Two strengths of the HGP were its comprehensiveness, and its commitment to open access to the sequence as a research tool without encumbrance [6]. Similar principles, applied to the development and distribution of research tools for determining the function and therapeutic potential of genes, would empower the biomedical research community as did the sequence itself. Tools for this 'genome functionation toolbox' would include a subset of those routinely used to investigate the function of novel genes. Applicable tools would be well validated in individual gene studies, flexible and applicable to a broad range of gene classes, producible by high-throughput methods, and outside the reach of individual investigators due to lack of expertise or resources. Nucleic-acid-based tools that meet these criteria might include whole-genome transcriptome [23,24] and proteome [25] characterization, and initiatives to develop small inhibitory RNAs (siRNAs) [26] and knockout mice [27] for every gene.

Small organic molecules have several potential advantages as tools for genome functionation. Unlike nucleic-acid-based methods, small molecules most often affect the function of a protein directly, rather than acting indirectly via DNA or RNA. Small molecules can augment or diminish gene product activity, whereas genetic methods are used most often to increase or decrease the quantity of gene product present. Small molecules can be routinely made to be cell penetrant, whereas analogous nucleic acid tools (e.g. antisense or siRNA) are often limited by lack of cell permeability. In addition, small molecules can selectively affect one of several functions of a gene product [28••], and can be used in experiments on extracts, cells or whole organisms, at any stage of an organism's lifespan; the versatility of genetic methods is limited in these regards. Perhaps most importantly for the study of dynamic systems, the onset and duration of gene function perturbation can be chosen precisely and repeated if desired; this is difficult or impossible with genetic methods. Finally, the diversity of small-molecule chemical space is commensurate to the scope of biological protein-folding space encoded by the human genome. Nucleic-acid-based tools are generally limited to a single dimension of diversity defined by their primary sequence; biologically active small molecules, by contrast, have a much broader range of chemical composition and characteristics. Estimates of 'chemical space' have ranged above 10^{60} molecules [29].

Use of small molecules as genome functionation tools has several potential drawbacks, however. It has been suggested that only certain classes of gene products are capable of binding drug-like molecules to produce important biological or therapeutic effects; collectively these genes have been referred to as the 'druggable genome' [21]. If true, this would suggest that a small-molecule approach would not be equally effective for all classes of gene products, something that is essential to whole-genome analysis. The division of the genome into 'druggable' and 'non-druggable' is, however, based principally on extrapolation from currently developed drugs and compound libraries rather than on empirical evidence about the genome. Recent success in modulating distinctly non-traditional drug targets using small molecules suggests that the true druggable genome may be larger than is commonly believed [30]. For small molecules to be used as broadly as nucleic acid methods in genome functionation, they must be as broadly available; however, access to small-molecule libraries and screening is currently very limited in most academic institutions. Furthermore, many biologists and geneticists in the public sector are less familiar with the concept and use of small molecules for gene function studies than are their colleagues in the pharmaceutical sector, for whom these methods are routine [31]. However, in the past several years, several collaborative biology–chemistry facilities have been initiated in academic institutions, suggesting that there is increasing appreciation of the utility of small molecules in these settings ([32], http://iccb.med.harvard.edu/screening/faq_hts_facility.htm). It is nonetheless true that for small molecules to be used routinely in the functionation of genomes, new infrastructure for screening would be needed, and project-driven collaborations between academic biologists and chemists would need to become routine, as they are now in the pharmaceutical sector.

Using small molecules to functionate the genome

Several groups over the past two years have demonstrated the promise of small molecules to define function of novel genes and elucidate biological pathways. For example, the nuclear farnesoid X receptor (FXR) is in a gene family amenable to small-molecule manipulation, but its signaling has remained difficult to define because its native bile acid ligands have multiple signaling effects. Downes *et al.* [33••] identified a novel high-affinity agonist of FXR that has distinct transcriptional effects from its natural bile acid ligands, thus allowing chemical dissection of the FXR signaling pathway. This study is typical of recent 'chemical genetics' approaches to function, in that it utilized combinatorial chemistry (in this case, around a natural product scaffold), microarray profiling of compound effects, and structural biology information to produce novel insights of potential therapeutic importance that had not been possible with the use of genetic methods or natural ligands.

Other studies have identified small-molecule modulators of less conventional targets. The cystic fibrosis transmembrane conductance regulator (CFTR) is a cAMP-activated Cl^- channel, mutations in which cause cystic fibrosis by altering channel synthesis, maturation or intrinsic function. Such defects would not be thought of as typically 'druggable', but several classes of compounds that appear to activate CFTR by direct interaction have been found, including some that have activity on mutant channels [30]. Inhibitors of the channel with potential utility for secretory diarrhea have also been identified [34]. Chen *et al.* [35•] identified small-molecule agonists and antagonists of Smoothened (Smo), a protein that mediates Hedgehog signaling during development and tumorigenesis. Smo has been intensively studied via molecular and genetic methods in several species [36], but elucidation of its mechanisms of activation and coupling to downstream signaling proteins has been difficult and has been greatly facilitated by these compounds.

Other recent studies have emphasized the capacity of small molecules to influence only a subset of functions of their targets. Haggarty *et al.* [37•] identified a novel compound, tubacin, that inhibits only one of the two actions of its target, HDAC6, using cellular and transcriptional readouts. Inhibitors of histone deacetylases (HDACs) have been suggested as clinical treatments for cancer and neurodegenerative disorders [38], but broad-spectrum HDAC inhibitors such as trichostatin also inhibit tubulin acetylation, making interpretation of the cellular effects of non-specific HDAC inhibitors complex; development of agents such as tubacin will allow determination of the functional effects of selective inhibition of the deacetylation actions of HDAC6. A similar theme was developed in earlier work from the Schreiber group, in which a compound was discovered via diversity-oriented synthesis and small-molecule microarray binding that modulated only a subset of functions of the yeast protein Ure2p [28••]. The compound, uretupamine, regulated only the glucose-response arm of signaling through Ure2p without affecting nitrogen source response, making its effects more specific than those obtained upon deletion or inhibition of the *URE2* gene.

These studies illustrate an important utility of small molecules — modulating specific functions of multifunctional proteins — that makes them particularly suited to the task of assigning function to novel genes. One of the more surprising findings of the HGP was that humans have only ~30 000 genes, the same number as the mouse and only two to three times the number of the fruit fly *D. melanogaster* or roundworm *C. elegans*. It appears that many human genes may produce more than one protein, and that many human proteins may subserve more than one function, allowing a relatively constant number of genes to support increasing organismal complexity. The implication for efforts to determine gene function is that manipulations at the genetic level may produce a multiplicity of effects, whereas small molecules may affect only one of several cellular phenotypes for which a gene is responsible [39].

Several investigators have expanded the utility of small molecules for defining novel biology by altering the cell line or target being screened. In an example of the former, 'synthetic lethal' screens have identified compounds with activity against cell lines altered in their expression of particular cancer-related genes, but not against their normal isogenic cell counterparts, leading to the expectation that these compounds will have increased specificity for neoplastic cells [40,41]. In a related approach, sequence changes have been made in drug targets rendering them unable to bind their normal ligands, and small molecules identified that bind only the mutated target. This process, which has been referred to as 'orthogonal chemical genetics' [42] has produced important insights into physiological functions of G-protein-coupled receptors (GPCRs) [43••], nuclear receptors [44] and kinases [45••], and extends the concept of screening ligand-binding domain mutated targets to identify allosteric modulators of receptors [46]. In another example of target modification, Mallet *et al.* [47•] expressed an engineered version of caspase-3 having FKBP-binding domains in transgenic mice, which allowed induction of caspase-3 activation and *in vivo* hepatocyte apoptosis upon addition of a small-molecule derivative of the semisynthetic FK506 dimer FK1012.

Finally, small molecules offer promise in defining biological pathways by their combination with genomic readouts such as DNA microarrays. Identification of pathways operative in the mechanism of action of pharmaceutical compounds by microarray analysis is now routine [48].

Extension of this approach to compounds in the public sector, and a mechanism to compare results of small-molecule perturbations in a common database, will be critical to realize the power of genomic profiling of small-molecule compounds. Taken one step further, genome-wide transcriptional readouts of series of related compounds could begin to allow inference of compound action and novel transcript function on the basis of patterns of gene regulation elicited by related small molecules with known, or even unknown, targets ([49]; SL Schreiber, personal communication). By comparing the patterns of gene and protein alteration produced by related molecules, pathways targeted by the compounds and the pharmacophores within them could begin to be distinguished. Applied on a large scale, this approach could become a powerful tool for defining signaling networks within living cells.

Conclusion

Chemistry and molecular genetics were closely allied disciplines at the time of the discovery of the structure of the DNA double helix, 50 years ago this year [50]. In fact, many of those who were instrumental in creating the new field of molecular biology were trained as chemists, among them Sydney Brenner, Linus Pauling, Rosalind Franklin and Max Perutz. Though close collaboration among chemists and biologists has long been the rule in the pharmaceutical industry, academic departments of genetics, biology and chemistry have been less closely allied in recent years. With the completion of the sequencing of the human genome, these fields are again finding fertile ground for collaboration as well as common themes of study in areas of structural/sequence diversity and informatics. As powerful modulators of gene function, small molecules will have an outsized role to play in realization of the promise of the Human Genome Project for biology and medicine.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Pennisi E: **Human genome: reaching their goal early, sequencing labs celebrate**. *Science* 2003, **300**:409.
2. Roses AD: **The genome era begins**. *Nat Genet* 2003, **33**:S217.
3. Collins FS, Green ED, Guttmacher AE, Guyer MS: **A vision for the future of genomics research**. *Nature* 2003, **422**:835-847.
4. National Research Council: *Mapping and Sequencing the Human Genome*. Washington DC: National Academy Press; 1988.
5. Collins F, Galas D: **A new five-year plan for the U.S. Human Genome Project**. *Science* 1993, **262**:43-46.
6. Collins FS, Morgan M, Patrinos A: **The Human Genome Project: lessons from large-scale biology**. *Science* 2003, **300**:286-290.
7. Tilghman SM: **Lessons learned, promises kept: a biologist's eye view of the Genome Project**. *Genome Res* 1996, **6**:773-780.
8. Holtzman NA, Marteau TM: **Will genetics revolutionize medicine?** *N Engl J Med* 2000, **343**:141-144.
9. Jamieson AC, Miller JC, Pabo CO: **Drug discovery with engineered zinc-finger proteins**. *Nat Rev Drug Discov* 2003, **2**:361-368.
10. Godfray J, Estibeiro P: **The potential of antisense as a CNS therapeutic**. *Expert Opin Ther Targets* 2003, **7**:363-376.
11. Anon: *The Fruits of Genomics: Drug Pipelines Face Indigestion Until the New Biology Ripens*. New York: Lehman Brothers, McKinsey & Co.; 2001.
12. Bell JI: **The double helix in clinical practice**. *Nature* 2003, **421**:414-416.
13. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W *et al.*: **Initial sequencing and analysis of the human genome**. *Nature* 2001, **409**:860-921.
14. Pennacchio LA, Olivier M, Hubacek JA, Cohen JC, Cox DR, Fruchart JC, Krauss RM, Rubin EM: **An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing**. *Science* 2001, **294**:169-173.
15. Thomas JW, Touchman JW, Blakesley RW: **Comparative analyses of multi-species sequences from targeted genomic regions**. *Nature* 2003, in press.
16. Winzeler EA, Castillo-Davis CI, Oshiro G, Liang D, Richards DR, Zhou Y, Hartl DL: **Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays**. *Genetics* 2003, **163**:79-89.
17. Grant BD, Wilkinson HA: **Functional genomic maps in *Caenorhabditis elegans***. *Curr Opin Cell Biol* 2003, **15**:206-212.
18. • Lum L, Yao S, Mozer B, Rovescalli A, Von Kessler D, Nirenberg M, Beachy PA: **Identification of Hedgehog pathway components by RNAi in *Drosophila* cultured cells**. *Science* 2003, **299**:2039-2045.

This paper demonstrates the use of RNA interference to knock down a total of 43% of all predicted *Drosophila* genes in cultured cells, as a screen for genes active in a critical developmental pathway.

19. Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, Erdos MR, Robbins CM, Moses TY, Berglund P *et al.*: **Recurrent *de novo* point mutations in lamin A cause Hutchinson-Gilford progeria syndrome**. *Nature* 2003, **423**:293-298.
20. Huntington's Disease Collaborative Research Group: **A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes**. *Cell* 1993, **72**:971-983.
21. Hopkins AL, Groom CR: **The druggable genome**. *Nat Rev Drug Discov* 2002, **1**:727-730.
22. Drews J: **Drug discovery: a historical perspective**. *Science* 2000, **287**:1960-1964.
23. Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, Dillon W, Lee KF, Clark KE, Haverty P *et al.*: **A compendium of gene expression in normal human tissues**. *Physiol Genomics* 2001, **7**:97-104.
24. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M *et al.*: **Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays**. *Nat Biotechnol* 2000, **18**:630-634.
25. Chanda SK, Caldwell JS: **Fulfilling the promise: drug discovery in the post-genomic era**. *Drug Discov Today* 2003, **8**:168-174.
26. Frankish H: **Consortium uses RNAi to uncover genes' function**. *Lancet* 2003, **361**:584.
27. Zambrowicz BP, Sands AT: **Knockouts model the 100 best-selling drugs–will they model the next 100?** *Nat Rev Drug Discov* 2003, **2**:38-51.
28. •• Kuruvilla FG, Shamji AF, Sternson SM, Hergenrother PJ, Schreiber SL: **Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays**. *Nature* 2002, **416**:653-657.

This paper utilizes two novel methods, diversity oriented synthesis and screening of small molecules on microarrays, to identify the first known modulator of a yeast protein, Ure2p, known to function in nitrogen metabolism. Using genome-wide transcriptional readouts, the authors then show that Ure2p also has a role in glucose sensing, a function

previously unappreciated from genetic studies of the target. This paper demonstrates the power of novel synthesis, screening and compound evaluation methods to discover new biology.

29. Smit WA, Bochkov F, Caple RE: *Organic Synthesis: The Science Behind the Art*. Cambridge, UK: Royal Society of Chemistry; 1998.

30. Ma T, Vetrivel L, Yang H, Pedemonte N, Zegarra-Moran O, Galietta LJ, Verkman AS: **High-affinity activators of cystic fibrosis transmembrane conductance regulator (CFTR) chloride conductance identified by high-throughput screening**. *J Biol Chem* 2002, **277**:37235-37241.

31. Rohrer SP, Birzin ET, Mosley RT, Berk SC, Hutchins SM, Shen DM, Xiong Y, Hayes EC, Parmar RM, Foor F *et al.*: **Rapid identification of subtype-selective agonists of the somatostatin receptor through combinatorial chemistry**. *Science* 1998, **282**:737-740.

32. Alper J: **Drug development. Biotech thinking comes to academic medical centers**. *Science* 2003, **299**:1303-1305.

33. Downes M, Verdecia MA, Roecker AJ, Hughes R, Hogenesch JB,
•• Kast-Woelbern HR, Bowman ME, Ferrer JL, Anisfeld AM, Edwards PA *et al.*: **A chemical, genetic, and structural analysis of the nuclear bile acid receptor FXR**. *Mol Cell* 2003, **11**:1079-1092.
The authors demonstrate a multifaceted approach to the dissection of a well known but poorly understood signaling pathway, using combinatorial and natural products chemistry, transcriptional profiling and crystal structure data.

34. Ma T, Thiagarajah JR, Yang H, Sonawane ND, Folli C, Galietta LJ, Verkman AS: **Thiazolidinone CFTR inhibitor identified by high-throughput screening blocks cholera toxin-induced intestinal fluid secretion**. *J Clin Invest* 2002, **110**:1651-1658.

35. Chen JK, Taipale J, Young KE, Maiti T, Beachy PA: **Small molecule**
• **modulation of Smoothened activity**. *Proc Natl Acad Sci USA* 2002, **99**:14071-14076.
This paper identifies small-molecule agonists and antagonists of a critical protein in embryonic development that had been resistant to genetic dissection of its function.

36. Zhang XM, Ramalho-Santos M, McMahon AP: **Smoothened mutants reveal redundant roles for Shh and Ihh signaling including regulation of L/R asymmetry by the mouse node**. *Cell* 2001, **105**:781-792.

37. Haggarty SJ, Koeller KM, Wong JC, Grozinger CM, Schreiber SL:
• **Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation**. *Proc Natl Acad Sci USA* 2003, **100**:4389-4394.
The authors identify a small-molecule inhibitor of the tubulin-deacetylation activity of a protein better known as a histone deacetylase.

38. Steffan JS, Bodai L, Pallos J, Poelman M, McCampbell A, Apostol BL, Kazantsev A, Schmidt E, Zhu YZ, Greenwald M *et al.*: **Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila**. *Nature* 2001, **413**:739-743.

39. Strausberg RL, Schreiber SL: **From knowing to controlling: a path from genomics to drugs using small molecule probes**. *Science* 2003, **300**:294-295.

40. Dolma S, Lessnick SL, Hahn WC, Stockwell BR: **Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells**. *Cancer Cell* 2003, **3**:285-296.

41. Fantin VR, Berardi MJ, Scorrano L, Korsmeyer SJ, Leder P: **A novel mitochondriotoxic small molecule that selectively inhibits tumor cell growth**. *Cancer Cell* 2002, **2**:29-42.

42. Shokat K, Velleca M: **Novel chemical genetic approaches to the discovery of signal transduction inhibitors**. *Drug Discov Today* 2002, **7**:872-879.

43. Scearce-Levie K, Coward P, Redfern CH, Conklin BR: **Tools for**
•• **dissecting signaling pathways *in vivo*: receptors activated solely by synthetic ligands**. *Methods Enzymol* 2002, **343**:232-248.
The authors are the creators of 'RASSLs', or receptors activated solely by synthetic ligands, which are engineered GPCRs that respond only to exogenously applied small-molecule compounds but not to their endogenous ligands. These constructs have proven valuable in elucidating multiple aspects of GPCR signaling and physiology.

44. Tedesco R, Thomas JA, Katzenellenbogen BS, Katzenellenbogen JA: **The estrogen receptor: a structure-based approach to the design of new specific hormone-receptor combinations**. *Chem Biol* 2001, **8**:277-287.

45. Eblen ST, Kumar NV, Shah K, Henderson MJ, Watts CK,
•• Shokat KM, Weber MJ: **Identification of novel ERK2 substrates through use of an engineered kinase and ATP analogs**. *J Biol Chem* 2003, **278**:14926-14935.
These authors molecularly engineered an ERK2 kinase mutant that could utilize an analog of ATP (cyclopentyl ATP) that cannot be used by wild-type ERK2 or other cellular kinases, thus allowing them to identify two novel ERK2 substrates.

46. Spalding TA, Trotter C, Skjaerbaek N, Messier TL, Currier EA, Burstein ES, Li D, Hacksell U, Brann MR: **Discovery of an ectopic activation site on the M(1) muscarinic receptor**. *Mol Pharmacol* 2002, **61**:1297-1302.

47. Mallet VO, Mitchell C, Guidotti JE, Jaffray P, Fabre M, Spencer D,
• Arnoult D, Kahn A, Gilgenkrantz H: **Conditional cell ablation by tight control of caspase-3 dimerization in transgenic mice**. *Nat Biotechnol* 2002, **20**:1234-1239.
This paper demonstrates the use of a small molecule to study apoptosis *in vivo*, by creating a genetically engineered mouse that expresses a form of the apoptosis-inducing protein caspase-3 that can be activated by the small molecule.

48. Gerhold DL, Jensen RV, Gullans SR: **Better therapeutics through microarrays**. *Nat Genet* 2002, **32(Suppl)**:547-551.

49. Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai H, He YD *et al.*: **Functional discovery via a compendium of expression profiles**. *Cell* 2000, **102**:109-126.

50. Watson JD, Crick FHC: **A structure for deoxyribose nucleic acid**. *Nature* 1953, **171**:737-738.