HapMap, pharmacogenomics, and the goal of personalized prescribing

Kevin M. O'Shaughnessy

Clinical Pharmacology Unit, University of Cambridge, Cambridge, UK

Correspondence

Kevin M. O'Shaughnessy, Clinical Pharmacology Unit, University of Cambridge, Cambridge, UK.

E-mail: kmo22@medschl.cam.ac.uk

Keywords

HapMap, pharmacogenomics, personalized prescribing

Most of the gene variants that have been identified in the last 50 years have affected individual drug responses through effects on pharmacokinetics. Gene variants that affect a drug's pharmacodynamics, such as polymorphisms in receptors or signalling molecules, have been more difficult to identify. This often requires a more systematic and genome-wide approach, prompting the emergence of the new discipline of pharmacogenomics. The provision of the human genomic sequence and now data from the HapMap project look set to bring 'personalized prescription' a step closer.

In the postgenomic era, pharmacogenetics and pharmacogenomics are terms we encounter with increasing frequency. They are transforming the way we approach adverse drug reactions and drug responses; they may even deliver truly 'personalized' prescribing. It is important to remember, however, that pharmacogenetics is not a recent discipline; Freidrich Vogel coined the term originally in 1959, only a few years after the structure of DNA had been established [1]. However, pharmacokinetics has traditionally focused on the interplay of a drug with a single gene or a small set of genes. There is no denying that this approach has been extremely successful in explaining pharmacokinetic variability between individuals. It has led us to appreciate the polymorphic nature of many drug-metabolizing enzymes and drug transporters and this knowledge now plays an important part the preclinical assessment of any new drug.

However, it has become apparent that drug responses (after allowing for pharmacokinetic variability) are unlikely to be controlled by similar large single-gene effects. Instead, we anticipate that drug responses will pose complex genetic problems that reflect a number of

genes, possibly in single figures, possibly a few tens of genes. The gene may encode a drug target, such as a TM7 receptor, or perhaps a down-stream signalling molecule, but the individual contribution will be small to modest and affected by gene-gene and gene-environment interactions. The dissection of gene effects such as these calls for a much more comprehensive and genome-wide approach. This is the territory of pharmacogenomics, but the scale of the problem has meant that pharmacogenomics has lagged behind pharmacokinetics. It has taken the development of the necessary bioinformatics tools to manipulate the huge amounts of data involved, and of course the data themselves. Those data became a reality in 2005 with the publication of phase I of the HapMap project [2] and the promise of an even larger dataset from phase II of the project.

The HapMap project [3] evolved naturally from the sequencing of the human genome that highlighted single nucleotide polymorphisms (SNPs, pronounced 'snips') as the commonest source of variation between individuals. This is a remarkable pace of discovery, considering that the two drafts of the human sequence with some 3 billion bases and around 30 000 genes were not pub-

lished until early 2001 [4, 5]. That project had taken 11 years, an international consortium and an uneasy amalgam of public and private capital totalling well over \$3 billion. Currently over 10 million SNPs have been identified within the human genome [6], although probably 10–15% are false positives. HapMap's remit was to establish the frequency of some 1 million of these SNPs (and over 5 million in phase II), using a representative panel of 269 DNAs from four different ethnic groups: Yorubans from Nigeria, White Americans from Utah, Han Chinese from Beijing and Japanese from Tokyo. Evidence suggests that although these samples were not specifically selected to represent these populations, they do broadly represent African, European and Asian ancestry.

The importance of the HapMap project was to use these data to identify common haplotypes within the genome, because for many SNPs there is a highly predictable adjacent SNP. The correlation between them is referred to as linkage disequilibrium and it arises from the fixing of the mutation creating the SNP to the particular combination of SNP genotypes on the ancestral chromosome. Over successive generations coupling is slowly lost by recombination between the new SNP and its neighbours, so creating the blocks of SNPs in tight linkage disequilibrium that we find today; these are the haplotypes. Tight linkage disequilibrium also builds in redundancy, in the sense that the haplotypes in a genomic segment can be identified without sequencing all of its SNPs. Enough information is actually contained in a small subset of SNPs (so-called tagging SNPs) and for many genes these are in single figures. Hence, HapMap will deliver a resource of TagSNPs to define common haplotypes in the major ethnic groupings. The attraction of haplotypes is the powerful tool that they provide for genetic association studies. A gene's contribution can be assessed anonymously, without the need to identify the 'functional' SNP or SNPs that might be operational in a genetic association. The whole genome is simply interrogated systematically as linkage disequilibrium blocks defined by TagSNPs. In practice, this means genotyping each subject for perhaps about 3–500 000 SNPs, but this is entirely feasible with current chip array technology (Figure 1).

In what directions might the arrival of the HapMap resource and the technology to genotype a genome-wide panel of TagSNPs take pharmacogenomics? It will certainly enable rapid dissection of genetic contributions to drug responses, both wanted and unwanted. A good proof-of-principle of this has yet to emerge, but typical recent pharmacogenetic successes that have had a clear candidate gene could have been tackled by an anony-

Figure 1

A commercial bead-array chip (www.illumina.com) the size of a microscope slide, which allows over 300,000 HapMap SNPs to be genotyped in a single hybridization step

mous pharmacogenomic approach. The discovery that the haplotype for the vitamin K epoxide reductase (VKORC1) gene affects warfarin sensitivity (more so than CYP2C9) actually did use information provided by HapMap [7]. The role of HLA haplotypes in determining severe cutaneous reactions to drugs is also ideally suited [8, 9].

There are additional spin-offs from the pharmacogenomics approach. Identifying the genes that affect a drug response could reveal novel mechanisms of disease. These will in turn suggest new drug targets; cancer therapeutics is an area in which this approach has already had a significant impact [10, 11].

Powerful economic drivers are steering the uptake of pharmacogenomics. On the one hand, pharmaceutical companies need to get new drugs to market sooner and with reduced risks of failure. On the other hand, prescribers (in the guise of healthcare systems) need better ways of identifying those who are most likely to benefit from the new and expensive drugs that exit the pharmaceutical pipeline. Pharmaceutical companies have been slow to incorporate pharmacogenomics as a standard tool in clinical drug development, at least in terms of studies that enter the public domain. This is important, because pharmacogenetic studies to date have almost exclusively come from academic departments with rel-

atively limited access to large sample sizes. Perhaps not surprisingly, the limited statistical power of these small studies explains why genotypic prediction of drug response is often difficult to replicate. Consider, for example, the confusion that has surrounded the role of the promoter polymorphism in the 5HT transporter gene in predicting the outcome of selective serotonin reuptake inhibitor therapy in depression (the short form usually being associated with a poor outcome) [12].

Part of industry's reticence to publicize pharmacogenomic results has been the concern that agencies such as the Food and Drug Administration (FDA) would delay the registration process if these data were routinely available. However, the FDA has encouraged the voluntary deposition of supplementary pharmacogenomics data and has emphasized that this will have a neutral impact on the registration process [13]. This should pave the way for the use of pharmacogenomics in the large cohorts that are routinely studied in phase III of the drug development process and beyond. On the assumption that drug responses are controlled by multiple gene variants that are relatively frequent in the population (1–50%), phase II studies will generally be too small for the task. The SNPs or gene variants that underpin a serious adverse drug reaction will inevitably be rarer (perhaps in the 0.01–0.1% range) and large phase IV cohorts will probably be needed for effective use of pharmacogenomics as a pharmacovigilance tool. However, the rewards are potentially huge. If the SNP signature of these adverse drug reactions were diagnostic, subjects who should never receive a particular drug could be easily identified. This would safely avert many of the drug withdrawals that occur after registration, which not only are costly for pharmaceutical companies but also deny nonsusceptible patients potentially very important drugs, such as ximelagatran, which was recently withdrawn because of three deaths from severe hepatocellular failure.

A lot has been said about the promise of pharmacogenomics in delivering individually optimized drug therapy and risk prediction of adverse drug reactions. Genome-wide SNP genotyping is currently an expensive research tool, but this will inevitably change. It is certainly going to be affordable and practical to apply routinely in the clinic in the next 5–10 years. So will personalized medicine then be a reality? It is not quite that simple. Although we have known, for example, the impact of drug polymorphisms such as CYP2D6 for over 2 decades, drugs that are metabolized by CYP2D6 have never been prescribed in a genotype-specific way. This is partly because of the lack of rapid and cheap genotyping (as opposed to the phenotypic

metabolizer status that is based on formal pharmacokinetic testing) and partly because of the lack of the necessary data to make detailed dosage adjustments. Manufacturers also recommend average dosages that are usually based on Caucasian samples with a minority of poor CYP2D6 metabolizers. The frequencies of most drug-metabolizing polymorphisms are very different in nonCaucasian groups, especially in South-east Asia. In the same way, exploiting pharmacogenomic data will require radical changes to the way data are collected during the drug development process, with much greater emphasis on genotype-driven data recovery and trial design [14]. Recovering material suitable for genotyping serious adverse drug reactions is another area that needs to be addressed. There are also issues of privacy and disclosure of pharmacogenomic data that will need to be tackled, because the process is likely to identify pharmacogenomic 'untouchables' for some drugs. All of this is going to take time.

The suggestion that individuals will be genotyped at birth and their 'HapMap genotype' carried lifelong as an implantable identity may be too Orwellian for some. However, it may be close to the reality of clinical pharmacology practice in decades to come. By then we will have the data needed to extract the genetic risk from a genome and effectively model dosages and the risks of adverse reactions based on an individual's genotype. This is further than most imagined electronic prescription would go but it is an attractive prospect. Pharmacogenomics has a long way to go to achieve this, but its potential is clear enough.

References

- 1 Vogel F. Moderne problem der humangenetik. *Ergeb Inn Med U Kinnerheilk* 1959; 12: 52–125.
- 2 Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005; 437 (7063): 1299–320.
- 3 International HapMap Project. <http://www.hapmap.org> Last accessed 2 April 2006.
- 4 Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001; 291 (5507): 1304–51.
- 5 International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; 409 (6822): 860–921.
- 6 National Center for Biotechnology Information. Single Nucleotide Polymorphism. <http://www.ncbi.nlm.nih.gov/SNP> Last accessed 2 April 2006.
- 7 Geisen C, Watzka M, Sittlinger K, Steffens M, Daugela L, Seifried E, Muller CR, Wienker TF, Oldenburg J. VKORC1 haplotypes and their impact on the inter-individual and inter-

- ethnic variability of oral anticoagulation. *Thromb Haemost* 2005; 94: 773–9.
- 8 Hung SI, Chung WH, Liou LB, Chu CC, Lin M, Huang HP, Lin YL, Lan JL, Yang LC, Hong HS, Chen MJ, Lai PC, Wu MS, Chu CY, Wang KH, Chen CH, Fann CS, Wu JY, Chen YT. HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proc Natl Acad Sci USA* 2005; 102: 4134–9.
- 9 Lonjou C, Thomas L, Borot N, Ledger N, de Toma C, Lelouet H, Graf E, Schumacher M, Hovnanian A, Mockenhaupt M, Roujeau JC. A marker for Stevens – Johnson syndrome . . . : ethnicity matters. *Pharmacogenomics J* 2006; [Epub] DOI: 10.1038/sj.tpj.6500356.
- 10 Cheok MH, Lugthart S, Evans WE. Pharmacogenomics of acute leukemia. *Annu Rev Pharmacol Toxicol* 2006; 46: 317–53.
- 11 Simon R, Wang SJ. Use of genomic signatures in therapeutics development in oncology and other diseases. *Pharmacogenomics J* 2006; [Epub] DOI: 10.1038/sj.tpj.6500349.
- 12 Murphy GM Jr, Hollander SB, Rodrigues HE, Kremer C, Schatzberg AF. Effects of the serotonin transporter gene promoter polymorphism on mirtazapine and paroxetine efficacy and adverse events in geriatric major depression. *Arch Gen Psychiatry* 2004; 61: 1163–9.
- 13 Lesko LJ, Salerno RA, Spear BB, Anderson DC, Anderson T, Brazell C, Collins J, Dorner A, Essayan D, Gomez-Mancilla B, Hackett J, Huang SM, Ide S, Killinger J, Leighton J, Mansfield E, Meyer R, Ryan SG, Schmith V, Shaw P, Sistare F, Watson M, Worobec A. Pharmacogenetics and pharmacogenomics in drug development and regulatory decision making: report of the first FDA-PWG-PhRMA-DruSafe Workshop. *J Clin Pharmacol* 2003; 43: 342–58.
- 14 Kirchheiner J, Fuhr U, Brockmoller J. Pharmacogenetics-based therapeutic recommendations – ready for clinical practice? *Nat Rev Drug Discov* 2005; 4: 639–47.