Recent developments in personalised medicine

THE ROYAL SOCIETY

Summary

This short report summarises progress that has been made in some aspects of the field of personalised medicine since the publication in 2005 of the Royal Society's report *Personalised Medicine: Hopes and Realities*.

It focuses on scientific progress, rather than policy developments, in this important field. There now seems little doubt that genomics and its related fields will play an increasing role in medical research and practice in the future. Given the many complexities of disease it is not surprising that progress in some areas of this field has been slow. More rapid progress will be made towards the assessment of this approach to the common diseases of Western society by developing tighter partnerships between clinicians who have the day to day care of patients with these diseases and those who are studying them at the molecular level; lessons from several fields have made it absolutely clear that the phenotype must not be neglected as part of the basic investigations of the role of the genotype and related fields in the evolution of a more personalised approach to medical care.

The Report was compiled by the following group:

Professor Sir David Weatherall FRS (Chairman)
Weatherall Institute of Molecular Medicine, University of Oxford.

Professor Sir Walter Bodmer FRS
Weatherall Institute of Molecular Medicine, University of Oxford.

Professor Sir Alasdair Breckenridge FRSE
University of Liverpool.

Professor Dame Kay Davies FRS
Department of Physiology Anatomy and Genetics, University of Oxford.

Professor Munir Pirmohamed FMedSci
The Wolfson Centre for Personalised Medicine, Institute of Translational Medicine, University of Liverpool.

Professor Andrew Wilkie FRS
Weatherall Institute of Molecular Medicine, University of Oxford.

Introduction

Although the personalised approach to dealing with the protean manifestations of the same or similar diseases has always been the critical part of good clinical practice, it was the results of the Human Genome Project that offered to put it on a much stronger scientific basis.

In 2005 the Royal Society produced a report entitled *Personalised Medicines: Hopes and Realities*. The Report was confined largely to the field of pharmacogenetics, that is the application of the better understanding of the human genome to its more logical application for drug development, and, in particular, their safer usage. As well as discussing the technology involved it made a number of recommendations for the broader aspects of the development of this field, including how it might be integrated into our health services and many of the educational and ethical issues that might be involved in the future.

Although by 2005 pharmacogenetics had made relatively little impact on clinical practice, there were some new products resulting from its technology on the market, predominantly in the field of cancer, and some examples of the value of genetic testing for the safety of drugs. In short there were already indications of the likely benefit of the application of pharmacogenetics in the future.

Over the last nine years there have been some remarkable developments in the technical aspects of genomics that are undoubtedly relevant to the further development of personalised medicine, or stratified medicine as it is also known. Sequencing the human genome was completed in 2001; it took 10 years to complete and cost $3 bn. Since then sequencing has reduced in price over 50% every year and the day of the $1,000 genome is now with us. This has led to major commercial pressure and there are now at least 10 biotechnology companies competing to develop a whole-genome sequencing platform that is commercially robust for both research and clinical purposes. Already, multiplexing technology makes DNA sequencing of a significant number of genes for relatively large numbers of people quite affordable for routine clinical-genetics services. Remarkable advances have also been made in a variety of fields relating to genomics and considerable progress has been made towards an understanding of gene regulation, particularly at the RNA level. These new disciplines include transcriptomics, epigenomics, proteomics, metabolomics and others. All these new areas of research offer the possibility of producing valuable biomarkers for different aspects of clinical practice.

Despite this remarkable technical progress doubts are still being cast about the value of the relationship between genomics and personalised medicine and its likely role in the future. In this brief update of the further development of this field since 2005 we will focus first on particular areas that have been foremost in its clinical application and then try briefly to summarise the outlook for its future development in relation to its application for day-to-day clinical practice.

Readers who are not familiar with the basic issues of inherited disease, the mechanisms of gene action, and recent advances in the study of the human genome are referred to the Glossary and several key references which are highlighted at the end of this paper.

Cancer

In the application of personalised medicine to cancer it is very important to distinguish somatic from germ-line variation.

Each cancer is the product of an independent somatic evolutionary process within a human body, in which changes in the expression of key genes, especially due to genetic mutations, are successively selected for their effect on the ability of the cells of the developing cancer to evade their normal growth controls and eventually expand and spread without limit. It is these genetic changes that define the biology of a cancer and which may be specifically targeted by appropriate drugs, and which largely determine the response of a cancer to treatment. These genetic changes are not inherited in the conventional sense of being passed on in the germ line (sperm and eggs) from parent to offspring. Pharmacogenetics, which was the main topic of the 2005 Royal Society Report, is concerned with germ-line inherited genetic differences in the response of an individual to specific drugs. These are properties with which an individual is born and which may also occur in that individual's relatives. Inherited susceptibilities to cancers are similarly the products of inherited germ line differences. **There are, thus, three quite different aspects to personalised medicine and cancer:**

a) The treatment of a cancer as a function of its genetic constitution,

b) The influence of an individual's genetic make up on the effectiveness of a particular cancer treatment, and

c) The influence of an individual's genetic make up on the chance of them developing a particular type of cancer.

The impact of the developments in molecular genetics for cancer personalised medicine since the publication of the 2005 report will be considered separately for each of these three aspects.

The treatment of a cancer as a function of its genetic constitution

Over the last 5 years there has been a dramatic increase in the number of cancers that have been fully sequenced and a corresponding increase in the number of selectively advantageous mutations identified. This, together with an increase in the amount of information about changes in gene expression in cancers, using both micro-array and cDNA sequencing techniques, has considerably increased the amount of information available on cancer properties that can be associated with response to therapy. At the same time, analysis of data from cancer cell-line studies and from clinical trials has revealed important properties that determine response to current and newly developed therapies. Thus, for example, it is now clear that the direct effects of therapy for colon cancer with anti-EGFR monoclonal antibodies are not associated with EGFR levels but depend on the absence of mutations in the *KRAS* and other oncogenes. It is now an essential requirement that only patients with *KRAS* non-mutant cancers are treated with anti-EGFR therapy, and this will soon be extended to requiring that the *BRAF*, *NRAS* and *PIK3CA* oncogenes are also non-mutant.

The techniques of gene knockdown in cell lines were used to show that a specific class of inhibitors, called PARP inhibitors, was active on *BRCA1* and *BRCA2* mutation carriers and then on a subset of sporadic breast cancers called 'triple negatives'. It is becoming clear that resistance to some of the new targeted small molecular weight compounds directed against tyrosine kinase activity arises through mutations in the target genes in the cancer as well as in genes regulating the intracellular concentrations of the drugs through active influx and efflux, for example in transporters belonging to the multidrug resistance protein (MDR) family. These and other similar discoveries will require that many if not most cancers and even subclones of a single cancer will need to

be well characterised at the genomic level in order to optimise their treatment. This will become increasingly important as more drug combinations are used for cancer therapy in order to overcome the inevitable development of resistance to single or dual therapies. Such approaches to treatment, personalised to the cancer, are already being accepted by clinicians and undoubtedly will rapidly become more widely applied. The limiting factor to their application will therefore become the availability of the facilities and resources within the NHS for the necessary characterisation of cancers as a part of routine clinical management.

Cancer Pharmacogenetics

There has not been the extent of new developments in the area of individual genetic differences in responses to cancer chemotherapy that was predicted in the 2005 Royal Society report. This may, in part, reflect the types of new drugs that have recently been introduced into the clinic for cancer treatment which include the low molecular weight receptor tyrosine kinase inhibitors, already mentioned, that may not be so susceptible to modification by the P450 cytochrome enzymes as some of the more commonly used chemotherapeutic drugs. Especially prominent amongst the newer drugs is a range of monoclonal antibodies, starting with the well known herceptin antibody against *HER2/NEU* in breast cancers and the anti-EGFR antibody, cetuximab. While the prevailing view so far has been that these act by directly blocking the activity of the receptors, there is increasing evidence for immune-mediated effects of therapeutic antibodies. The immune effects work by killing cells by antibody-directed cellular cytotoxicity (ADCC), and similar mechanisms, based simply on antibodies binding to their targets on the cancer cells. The efficiency of this binding is influenced by polymorphic genetic variation in the antibody-binding receptor on the immune cells that mediate ADCC. This is a novel form of pharmacogenetic variation that may be important to take into account when using antibody-based therapies. It also turns out that, for example, for the immune based effect of cetuximab against EGFR in the treatment of colorectal cancer, it is the level of expression of EGFR on the cancer cells that matters rather than being non-mutant for *KRAS* and the other oncogenes. As newer recombinant DNA techniques are being used to create more sophisticated antibody-based therapies, the requirements for the characterisation of both the patient's germ-line response and the cancers' properties will no doubt change.

Though it seems that so far it has been hard to persuade both the clinicians and the pharmaceutical companies of the importance of pharmacogenetic evaluation of cancer patients before treatment, this may well change as it becomes apparent that, with newer combined therapies, such evaluation may lead to substantially improved outcomes.

Inherited Cancer Susceptibility

The last 7 years has seen a huge increase in the number and scale of genome wide association studies (GWAS) on cancers. Most recently, more than 200,000 selected single nucleotide polymorphisms (SNPs) have been tested on tens of thousands of breast, ovarian and prostate cancer patients in large internationally coordinated studies. These studies have identified many variants with statistically significant, but very small, effects. Few of the variants are even associated with as much a 10% increase or decrease in relative incidence. It seems doubtful that any, or perhaps at most a very few, of such variants will be useful either for identifying therapeutically relevant functions or, even in combination, for defining at risk subgroups appropriate for some form of increased prophylactic screening type intervention. However, they may lead to the identification of novel "drugable" targets.

The most interesting development from the application of next generation high throughput DNA sequencing has been the ability to pin down quite rare but clearly inherited susceptibilities to cancers that have been found in only a few, even sometimes just one, families. This parallels, of course, similar studies in many other rare inherited syndromes. A good example of this for cancer is the recent description of clearly inherited susceptibility to colorectal cancer caused by very rare variants in two different DNA polymerase sub-units. Even though such cases are rare, they can provide valuable functional insights to disease mechanisms and potential therapeutic developments.

The DNA sequencing advances over the last 5 or so years have already made possible the provision of sequencing services for the more common, though still rare, clearly inherited cancer susceptibilities such as familial adenomatous polyposis (*APC*), hereditary non-polyposis colon cancer (*HNPCC*) and breast cancer (*BRCA1* and *BRCA2*). This is now being routinely done in many clinical genetics units. As further rare inherited susceptibilities are discovered, the demand for such services will inevitably increase. The advantage of finding cancer susceptibilities, whatever the mechanism, is that early discovery of a cancer nearly always leads to improved treatment, or sometimes even prevention when, as in the case of colorectal cancer, precursor lesions can be identified and removed by colonoscopy.

It is very important that such work on rare cancer susceptibilities, a fair amount of which is done within the NHS, gets adequate support, both at the research and service levels.

Common non-communicable diseases

Of course the genetic basis for the cause or susceptibility to a few relatively common non-communicable diseases was known before the era of the complete human genome sequence.

For example, type I diabetes was clearly related to certain HLA-DR types and several genes had been identified as the cause of the rarer forms of juvenile diabetes; the latter had a dramatic effect on the management of these conditions. Similarly, several types of hypercholesterolaemia had been related to single gene mutations, again leading to their better management. But for the bulk of the common killers of Western society, although there was a hint of a genetic element in some of them, very little was known about their cause.

Large numbers of GWAS have now been carried out in populations with common diseases such as cardiac or renal disease, major psychiatric disorders, gastrointestinal disease, and many more. With a few exceptions, where genes have been identified that might increase or decrease susceptibility to these particular diseases, the overall effect has been extremely small. An additional difficulty is that a high proportion of these GWAS associations locate at a distance from clearly defined genes so that it is not always clear which gene is affected by the association, let alone what the mechanism of the effect might be. This is particularly the case where these susceptibility loci have been determined by association with SNPs.

Although these studies have produced a number of possible approaches to drug development for particular diseases, the extremely low additive genetic component in these associations seems to be unlikely to be of value for predictive testing at the personalised medicine level. Perhaps this is not surprising since in many of the studies the clinical phenotypes have not been analysed in great detail and it is possible that they have been examining more than one disease entity. Furthermore, many of the common non-communicable diseases have an extremely large environmental factor together with other important issues including lifestyle, diet, stress, not to mention the ill-understood effects of the biology of ageing.

Another factor that is complicating this particular aspect of personalised medicine is the remarkable difference in susceptibility to many of these common diseases in different ethnic groups. There is no doubt that there is a global epidemic of type 2 diabetes and related disorders. Although this has increased over recent years it can be traced back to what was described as the New World Syndrome which became apparent after World War II and is characterised by very high prevalence of obesity at an early adult age, type 2 diabetes, and gallstones and gall bladder cancer, particularly in women. Particularly in the case of type 2 diabetes there are some extreme examples of these ethnic differences. For example, the population of Pima has a 19 – fold greater incidence of type 2 diabetes than Caucasian populations in other parts of the Americas. A similar picture has been observed in some of the island populations of Micronesia and Polynesia. For example, the prevalence of type 2 diabetes on Nauru Island reaches 60% of the adult population. Some years ago Neel, Weiss and others

suggested that these observations reflected a ‘thrifty’ genotype which had been selected to take advantage of sporadic food and long periods of privation in primitive societies. There is no doubt that westernisation of many diets in low-income countries has caused a rapid increase in the frequency of diabetes and related diseases. Given the increasingly large immigrant populations from many of these high-frequency regions to the USA and Europe it is likely that the genetic component of susceptibility to these conditions may reflect the action of different genes in their widely varied ethnic groups. While a number of international partnerships have been developed in the field of trans-ethnic disease it may be a long time before their studies provide information that is appropriate to the question of personalised medicine.

In short, some progress is being made towards an understanding of the varying genetic components that underlie susceptibility to common diseases in Western society. However, because many of these components have only a small effect on susceptibility, and may well differ between different ethnic groups, it will be a long time before it is clear whether individual genome sequencing will be of genuine value in the control of these diseases. However, there is already promise that this approach may offer the pharmaceutical industry ways of developing new therapeutic agents provided that the functions of such biomarkers that are discovered are studied in detail.

Communicable disease

There is no doubt that technologies derived from various aspects of genomics have had a major effect on the diagnostic aspects of communicable disease.

This includes the early identification of new infective organisms, together with the increasing ease of diagnosis of infections due to known organisms, particularly those that are difficult to grow in culture. And whole-genome sequencing of both vectors and infectious organisms is beginning to disclose vaccine candidates, approaches to vector control, and prospects for new therapeutic agents.

There are also hints that a more personalised approach to the control of some common infectious diseases may become possible. Malaria is a good example. Over many years the exploration of candidate genes, and more recently GWAS, have disclosed a wide range of mutations which result in relative resistance to malaria; the classical example is the sickle cell gene but many others have now been identified. Given the level of protection conferred by some of these genetic variants against different aspects of malarial infection it seems likely that in the future they will have to become part of clinical trials for testing potential vaccines or therapeutic agents. And the recent discovery of some of the remarkable epistatic interactions between some of these protective mutations promises to provide more information about the precise mechanism whereby they reduce the severity of different aspects of malarial infection and hence offer new approaches to drug design. A particularly good example of this type of interaction is evidenced by the observation that individuals who are heterozygous for the sickle cell gene and those who are heterozygous or homozygous for a mild form of α thalassaemia both have strong protection against P. falciparum malaria whereas if they inherit both these mutations the protective effect is completely abolished. Although less is known about protective polymorphisms in relationship to other infectious diseases there are increasing numbers of examples and hence the same principles will apply to them.

Single gene (Mendelian) disorders

Further developments in genome technology are proving of particular value for the study of single-gene disorders. The vast majority of these conditions are extremely rare although there are a few that occur at a high frequency and are posing major health problems in different countries.

There has been major progress towards identifying genes for rare disorders and there are now several large Rare Disease Consortia where DNA from patients in families with single gene defects are being sequenced. Indeed, in almost half of the cases collected by some Consortia, those focusing on disorders associated with cardiac defects for example, the underlying disease mutation has been identified in almost 10% of cases. A good example of the way that such information can be translated into novel therapy is the discovery of the mutations responsible for a rare form of hypercholesterolaemia. A form of this disorder was found to result from mutations in a gene which encodes an enzyme called pro-protein convertase subtilisn kexin type 9 (*PCSK9*). It was found that both normal and mutant forms of *PCSK9* promote degradation of hepatic LDL receptors, which are involved in the removal of circulating LDL, suggesting that those that cause hypercholesterolaemia confer a gain of function. Later, loss of function mutations in *PCSK9* were found to increase the number of receptors and decrease circulating LDL levels; they are associated with a reduction in LDL concentrations and a reduction in coronary artery disease. These findings suggested that the inhibition of *PCSK9* function might be an approach to the prevention of coronary artery disease. Although no small molecule inhibitors of *PCSK9* have been identified this objective has been achieved by developing monoclonal antibodies against *PCSK9* which have been extremely successful in reducing plasma concentrations of LDL. Phase 3 clinical trials showing the effect of these antibodies as add-on therapy on LDL-C lowering have already been published, but there is a need for longer-term trial evaluating their effect on clinical outcomes.

Clearly the analysis of rare single gene disorders has important implications for the future of personalised medicine. First, there is increased ability to integrate gene mutation with sophisticated studies of disease pathogenesis in order to design targeted therapeutic strategies; and, second, the definition of these conditions at the molecular level, and at least in some cases, the similarity to some of the common diseases of western society, offer possible approaches both to determining the underlying basis of the latter disorders together with the development of agents for their better management.

There are some single gene disorders that occur at relatively high frequencies in different populations. The prime example are the inherited disorders of haemoglobin, notably the sickle cell disorders and the thalassaemias. It is estimated that between 300,000 and 400,000 babies are born each year with these diseases, over 90% of the births occurring in low- or middle-income countries. One of the major features of these conditions is the wide phenotypic diversity associated with the same or similar mutations at the loci which control the synthesis of the α or β chains of adult haemoglobin. Originally by the candidate gene approach and more recently by GWAS it has been possible to define some of the reasons for the remarkable variation in clinical phenotypes associated with the same genetic defects. In the case of the thalassaemias for example, it has been possible to classify these genetic modifiers into three classes; primary, which describe the action of different mutations of the α or β globin genes reflected by variation in the resulting reduction in globin chain production; secondary, which involves genetic modifiers which have a direct effect on ameliorating the degree of imbalanced globin chain

production, and tertiary, which represent a wide range of genetic modifiers of the many complications of these disorders which are not related directly to haemoglobin production. There has also been considerable progress towards an understanding of why there are differences in the way in which patients are able to adapt to the low haemoglobin levels which occur in these conditions and at least a start has been made at studying the neglected topic of the effect of the environment on these monogenic disorders.

At least in some forms of thalassaemia it has been possible to detect the reason for approximately 50% of the heterogeneity of disorders due to the same genetic defect and a start has been made to use these modifiers as an approach to defining better forms of management for individual patients. For example, some of these modifiers result in unusually high levels of fetal haemoglobin which compensate for the defect in adult haemoglobin caused by the thalassaemia mutations; some progress is already being made towards attempting to raise levels of fetal haemoglobin in patients who carry these modifier genes. Similar progress is being made towards a better understanding of the remarkable clinical diversity of sickle cell disease. There is no doubt that much of this progress has followed careful repeated observation of the phenotypes of the different forms of thalassaemia and sickle cell disease. The importance of long term and detailed clinical follow up of patients with these conditions has been accentuated by the finding of relative instability of the phenotype in younger patients. The lessons learnt from this detailed clinical phenotyping are probably applicable to many other studies, including searches for the genetic component of common disease.

Remarkable progress has been also made in the development of population screening for carriers of the severe haemoglobin disorders in high-frequency populations followed by prenatal diagnosis and the termination of affected pregnancies. This approach was first carried out by developing methods to study haemoglobin synthesis in fetal blood samples and by 1990 over 14,000 successful prenatal diagnoses had been carried out in different high-frequency countries. Later it became possible to carry out prenatal diagnosis earlier in pregnancy using DNA obtained by chorion-villus sampling and this is now applied widely, not just in the richer countries but in increasing numbers of the developing countries where these diseases occur at such a high frequency.

Methods to enhance protein expression

Particularly (but not exclusively) in recessive disorders, various ingenious methods are being developed to target the defect in protein synthesis that causes particular single-gene disorders. In many cases this is dependent on knowledge of the specific causative mutation and the intricate details of the cellular process that give rise to the disease phenotype.

A good example is provided by new therapies for Duchenne muscular dystrophy. This is caused by very many different mutations of the *DMD* gene that lead to loss of function of the protein. However, overall 83% of mutations are potentially amenable to correction by driving the skipping of specific exons and hence restoring the reading frame of the messenger RNA that encodes the protein. For specific mutations, proof-of-principle has been achieved using antisense oligonucleotides (AONs) in Phase II clinical trials, but there are still major challenges for efficient delivery to heart and skeletal muscle. Moreover, success with this approach raises challenging regulatory issues because the therapy is highly personalised - requiring the development of multiple AONs, with direct cost and regulatory implications if each individual AON was classed as a novel drug. Similar exon skipping approaches are achieving promising results in other diseases including spinal muscular atrophy and cancer.

It should also be noted that this is only one of several strategies to tackle protein deficiency – others include viral gene therapy (replacement of the missing gene using specifically engineered viruses), read-through of premature termination codons, and pharmacological methods to increase the expression of a compensatory protein. It is very unclear at this stage which will be the “winning” technology; only considerable investment in all potential therapeutic avenues will enable the multiple, but different challenges of each approach to be identified and, hopefully, overcome.

Pharmacogenetics and its relationship to the increasingly broader technology of personalised medicine

The report of the Royal Society in 2005 on the future of personalised medicine focused almost entirely on the personalisation of the efficacy and safety of drugs based on genetic studies.

Since then this field has rapidly extended beyond this approach to improving the dosing and choice of drugs or other interventions for patient care. For example, molecular imaging modalities are now able to identify metastases which may help to inform clinical decision-making about the level of intensity of particular therapy.

There is also increasing emphasis on the value of protein biomarkers. For example, periostin appears to have the potential to be used in the stratification of different asthma phenotypes and, in particular, to assess response to steroids and to agents such as anti-IL13.

We have discussed earlier some of the new approaches and usage of biomarkers for the discovery of new therapeutic agents. The field of drug safety has undoubtedly improved quite dramatically since our last report. Apart from the agents already mentioned in the cancer field there have been a number of other areas of advance, particularly in evaluating the major histocompatibility locus (*HLA*) for this purpose. For example, it has been found that *HLA-A*31:01* is an important pre-disposing locus for carbamazepine hypersensitivity. *HLA-B*57:01* is now routinely used to prevent hypersensitivity to abacavir; the need for this particular genetic test is included in the drug label and recommended in all clinical guidelines. This has already led to a fall in the incidence of abacavir hypersensitivity from 5 – 7% to less than 1%. *HLA-B*1502* genotyping has also been shown to prevent carbamazepine induced side effects in the Chinese population and, again, is also now recommended in the drug label. Indeed, there are now 24 strong HLA associations known to be related to different forms of idiosyncratic drug toxicity. There are many more examples of knowledge of genes which predispose towards drug toxicity. For example, response to interferon-alpha in patients with hepatitis C is dependent on genetic polymorphisms in *IFNL3* (IL-28B). This particular relationship has now been confirmed in many different studies and in many different populations. The field of hepatitis C therapeutics has seen major advances in recent years with the introduction of a number of novel compounds; all trials currently being conducted by pharmaceutical companies which use interferon-alpha therapy as a comparator or as part of combined therapy now require IL-28B testing. Although this will become less common as the new interferon-free therapies become more commonly used, it nevertheless provides an excellent example of how genomic investigations can provide novel insights into the mechanisms of action of therapies.

Another interesting development relates to the relationship between pharmacogenetics and single gene disorders. The relatively common condition cystic fibrosis has been found to result from several different mutations in the *CFTR* gene. One particular mutation, G551D, which affects only a low proportion of patients, has been found to be responsive to the agent ivacaftor, a small molecule that increases the time that activated CFTR channels remain open. The response to this agent has been found to be significantly greater than from attempts to deal with this condition by the recently developed methods of somatic gene therapy. The FDA has recently

expanded the indication for ivacaftor to encompass eight other *CFTR* mutations (G178R, S549N, S549R, G551S, G1244E, S1251N, S1255P, and G1349D). Further advances in this area with drugs such as lumecaftor targeting the most common CF mutation (Phe508del) are expected.

Industry

The number of therapeutic agents that are now being developed by Pharma in which the development of the drug is accompanied by the development of a diagnostic test has increased over the last few years; reported figures from the large companies range from 60% to 75%. At least 50% of these new drug-diagnostic combinations are for cancer. A recent report from one of the large companies suggests that the attrition rate has fallen from over 50% to less than 13% since the introduction of personalised healthcare approaches. Furthermore, their latest anti-cancer drug was licensed by the FDA based on only about 200 patients in a definitive phase III trial, while approval for vemurafenib was amongst the fastest in regulatory history, taking only four months. This tendency has been reported in several other cases and suggests that the use of this approach will considerably reduce the cost of clinical trials for new therapeutic agents.

Recently the Association of the British Pharmaceutical Industry (ABPI) issued a report entitled *The Stratification of Disease for Personalised Medicines*. The report concluded that the ability to classify individuals into sub-populations that differ in their susceptibility to a particular disease or their response to a specific treatment means that interventions can be concentrated on those who will benefit, sparing expense and side effects for those who will not. There certainly appears to be a move away from the block-buster model of drug development to a niche-buster model.

Regulation

The regulatory agencies have started issuing guidances on the use of pharmacogenetics in drug development. For example, the European Medicines Agency (EMA) has issued guidance on the need for dosing advice based on genetic polymorphisms in drug metabolising enzyme genes. And in the UK the National Institute for Care and Health Excellence (NICE) has formed a Diagnostics Panel Committee that will assess diagnostics for stratified medicine approaches. They are considering an Oncotype diagnostic test for women with breast cancer. This is a test that evaluates the expression of 21 genes in a tumour sample and quantifies the likelihood of disease recurrence in women with early-stage estrogen-receptor (ER)-positive-only breast cancer (prognostic significance) and assesses the likely benefits from certain types of chemotherapy (predictive significance).

A number of professional clinical bodies are also beginning to take a major interest in the applications of personalised medicine including some of their more practical ethical issues. For example, the American College of Medical Genetics and Genomics (ACMG) have recently published an important paper regarding the procedures to be followed with incidental findings that are discovered after human genome sequencing. Their advice is to feed this information back to patients and their families. In particular, genetic information on children must also be fed back to families. Some aspects of this guidance remain controversial in a European setting however.

Funding

This field has been highlighted as an area for further research which has led to some specific funding calls, including the MRC Stratified Medicine Initiative, TSB Stratified Medicine Innovation Platform, NIHR Stratified Medicine Competition and the alliance between AR-UK and TSB. In December 2012, the UK government announced that the genomes of up to 100,000 patients or infections in patients will be sequenced over the next five years. This will improve understanding, leading to better and earlier diagnosis and personalised care. Based on expert scientific advice, the Department of Health (DH) has initially prioritised sequencing of lung and paediatric cancer, rare diseases and infectious diseases. The project will be run by Genomics England, an organisation entirely owned by DH. Genomics England which will manage contracts for specialist UK-based companies, universities and hospitals to supply services on sequencing, data linkage and analysis. It will set standards for obtaining patients' consent and also strictly manage storage of personal data in accordance with existing NHS rules designed to securely protect patient information. It will have the independence and influence to drive innovation across systems and healthcare economies.

Up to £100 million of funding pledged by the government will be made available not only to sequence the 100,000 genomes but also to train a new generation of British genetic scientists to develop life-saving new drugs, treatments and to train the wider healthcare community to use the technology. The aim is to build secure NHS data linkage to ensure that this new technology leads to better care for patients.

Broader Issues

Many of the issues that have been considered briefly in this section are covered in more detail in a recent report from the UK Academy of Medical Sciences.

Conclusions

Overall, there is no doubt that considerable progress has been made in the field of personalised medicine. At least in part this has resulted from the extraordinary developments in the speed and accuracy of genome sequencing combined with equally remarkable progress in the disciplines related to genomics, notably transcriptomics and proteomics.

These, and related fields, are producing large numbers of novel biomarkers for the application of research directed at the cause or prevention of disease.

Largely because it has been the main practical success story we have focused this update on the cancer field and emphasised the different roles of biomarkers relating either to the prevention or treatment of malignant tumours. As well as some of the successes in other diseases we have also emphasised the value of the new sequencing technology with regard to both rare and common monogenic diseases which, as well as being of extreme value to individual families, are starting to produce lessons for the study of the common killers of western society.

There is abundant evidence that the pharmaceutical industry and the medical research community have taken on these new developments and that the field is in a very much more active state than it was ten years ago.

Not surprisingly, there have been disappointments and criticisms of the reductionist approach that is being applied to personalised or stratified medicine. Overall, the results of extremely large GWAS involving thousands of patients have, with a few exceptions, not disclosed anything more than multiple, individually minor genetic components to these conditions. These findings have raised questions about the potential clinical value of genome sequencing, however cheap it becomes, for the better control of many of our major non-communicable diseases. Recently published reviews on this topic arrive at widely different conclusions.

Some suggest that, within the foreseeable future, sequencing a patient's genome will become a routine part of day to day clinical practice, while others, because they estimate that the risk of developing these common diseases will be little different to that of the general population, suggest that genome sequencing will not be a major determinant for better patient care in the future. But it is early days, and there is still a definite possibility that some of the genetic markers for these conditions, even though they occur at a low frequency, will provide biomarkers in the future for a better understanding of the mechanism of these conditions and hence how they may be attacked with various forms of therapy more effectively.

It is not surprising that many of the major disorders of Western society have proved disappointing so far as regards defining a genetic basis for them. A key issue in this respect is the quite remarkable phenotypic variation of what are thought to be the same basic disease. It is inevitable that, by the very size of many of the GWAS it has not been possible to describe in detail the phenotypes of all the patients involved. These vary very much from case to case and many environmental factors are involved together with stress, individual reaction to disease, the social environment of the patient and, in particular, the very limited knowledge of the biology of ageing. Most of the common disorders of the advanced countries occur at an increased frequency as the population ages and we still only have the most limited knowledge of the effect of ageing on key issues like drug metabolism and the effect of age on the efficacy and toxicity of different forms of treatment.

There is an even more fundamental reason why patient phenotypes have been neglected to some degree in the early stages of gene hunting by GWAS. In many advanced societies the medical profession has become less concerned with the basic details of clinical history taking and detailed and regular patient examination. The reasons for this are diverse and relate to changes in medical and nursing education together with reorganisations of health services which do not allow doctors to spend enough time with their patients or, and this is of particular importance, to follow up their course sufficiently. For this reason it seems vital that this aspect of medical practice can be analysed in more detail and put right so that it will be possible in the future for a much closer collaboration between clinicians and research workers, with greater input of phenotypic data into studies related to a more personal approach to medical practice.

However, the biggest challenge of whole genome sequencing still remains the analysis and interpretation of the data. It is also acknowledged that for many diseases it will be necessary to share data on both phenotype and genotype in order to elucidate novel disease mechanisms. In recognition of this, the Global Alliance for Genomics and Health (GA4GH), has been formed which is an alliance of 148 of the world's leading biomedical research institutions, healthcare providers, information technology and life science companies, funders of research, and disease and patient advocacy organizations. This is a world-wide effort to responsibly aggregate and analyse large amounts of genomic and clinical information – seeking best practices where they exist, and developing new approaches where needed – in order to advance the understanding, diagnosis, and treatment for cancer, inherited diseases, infectious diseases, and drug responses.

The challenges of delivery of genomic information into the clinic remain numerous. In the UK, Genomics England Ltd is a company owned by the Department of Health (DH) which has been established to promote good practice and public engagement in this field. The aim is to generate of 100,000 genome sequences focussing initially on rare disease, cancer and infectious disease.

Acknowledgement

We thank Liz Rose for her help in preparing this report.

Some Sources and References for Further Reading

ACMG policy statement: updated recommendations regarding analysis and reporting os secondary findings in clinical genome-scale sequencing. (2015) *Genetics in Medicine* **17**, 68 – 69

Bodmer, W., and Bonilla, C. (2008). Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* **40**, 695 – 70

Bodmer, W. & Tomlinson, I. (2010) Rare genetic variants and the risk of cancer. *Current Opinion in Genetics Development*, **20**, 1 – 6.

Chen, R., Mias, G.I., Li-Pook-Than, J., et al. (2012) Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell*, **148**, 1293 – 1307.

Cohen, J.C. & Hobbs, H.H. (2013) Genetics. Simple genetics for a complex disease. *Science*, **340**, 689 – 690.

DH Commissioned Working Group: Recommendations on Strategic Priorities. https://www.gov.uk/government/publications/mapping-100000-genomes-strategic-priorities-data-and-ethics

European Medicines Agency Guidelines on the Use of Pharmacogenetics in Evaluating Pharmacokinetics of Medicines. http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2012/02/news_detail_001434.jsp&mid=WC0b01ac058004d5c1

Evangelou, E. & Ioannidis, J.P. (2013) Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews Genetics*, **14**, 379 – 389.

Fairclough, R.J., Wood, M.J. & Davies, K.E. (2013) Therapy for Duchenne muscular dystrophy: renewed optimism from genetic approaches. *Nature Reviews Genetics*, **14**, 373 – 378.

Fugger, L., McVean, G. & Bell, J.I. (2012) Genomewide association studies and common disease--realizing clinical utility. *The New England Journal of Medicine*, **367**, 2370 – 2371.

Ge, D., Fellay, J., Thompson, A.J., Simon, J.S., Shianna, K.V., Urban, T.J., Heinzen, E.L., Qiu, P., Bertelsen, A.H., Muir, A.J., Sulkowski, M., McHutchison, J.G. & Goldstein, D.B. (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* **461**, 399 – 401.

Goldstein, D.B., Allen, A., Keebler, J., Margulies, E.H., Petrou, S., Petrovski, S. & Sunyaev, S. (2013) Sequencing studies in human genetics: design and interpretation. *Nature Reviews Genetics*, **14**, 460 – 470.

Hershfield, M.S., Callaghan, J.T., Tassaneeyakul, W., Mushiroda, T., Thorn, C.F., Klein, T.E. & Lee, M.T. (2013) Clinical Pharmacogenetics Implementation Consortium guidelines for human leukocyte antigen-B genotype and allopurinol dosing. *Clinical pharmacology and therapeutics*, **93**, 153 – 158.

Horwitz, R.I., Cullen, M.R., Abell, J. & Christian, J.B. (2013) Medicine. (De)personalized medicine. *Science*, **339**, 1155 – 1156.

Houle, D., Govindaraju, D.R. & Omholt, S. (2010) Phenomics: the next challenge. *Nature reviews. Genetics*, **11**, 855 – 866.

Katsanis, S.H. & Katsanis, N. (2013) Molecular genetic testing and the future of clinical genomics. *Nature Reviews Genetics*, **14**, 415 – 426.

The Malaria Genomic Epidemiology Network (2008) A global betwork for investigating the genomic epidemiology of malaria. *Nature Reviews Genetics*, **456**, 732 – 737.

Ng, P.C., Levy, S., Huang, J., Stockwell, T.B., Walenz, B.P., Li, K., Axelrod, N., Busam, D.A., Strausberg, R.L. & Venter, J.C. (2008) Genetic variation in an individual human exome. *PLoS Genetics*, **4**, e1000160.

Pang, T. & Weatherall, D. (2012) Genomics and world health: a decade on. *Lancet*, **379**, 1853 – 1854.

Personalised Medicines: Hopes and Realities. The Royal Society (2005), http://www.royalsoc.ac.uk,

Pirmohamed, M. (2012) Genetics and the potential for predictive tests in adverse drug reactions. *Chemical immunology and allergy* **97**, 18 – 31.

Pirmohamed, M. & Hughes, D.A. (2013) Pharmacogenetic tests: the need for a level playing field. *Nat Rev Drug Discov* **12**, 3 – 4.

Realising the Potential of Stratified Medicine. Academy of Medical Sciences (2013). http://www.acmedsci.ac.uk/p99puid280.html

Roberts, N.J., Vogelstein, J.T., Parmigiani, G., Kinzler, K.W., Vogelstein, B. & Velculescu, V.E. (2012) The predictive capacity of personal genome sequencing. *Science Translational Medicine*, **4**, 133ra158.

Shazad Q. Ashraf, Angela M. Nicholls, Jennifer L. Wilding, Triantafyllia G. Ntouroupia, Neil J. Mortensen and Walter F. Bodmer (2012). Direct and immune mediated antibody targeting of ERBB receptors in a colorectal cancer cell-line panel. *Proc Natl Acad Sci USA* **109**, 21046 – 21051

Stratton, M.R. (2011) Exploring the genomes of cancer cells: progress and promise. *Science*, **331**, 1553 – 1558.

Symposium: Personalised Medicine (2015). *J Intern Med* **277**, 152 – 259

Van der Harst, P., Zhang, W., Mateo Leach, I. et al. (2012) Seventy-five genetic loci influencing the human red blood cell. *Nature*, **492**, 369 – 375.

Weatherall, D.J. (2013) The role of the inherited disorders of hemoglobin, the first 'Molecular Diseases', in the future of human genetics. *Annual Review of Genomics and Human Genetics*, **14**, 1 – 24.

Wilding, J.L. & Bodmer, W.F. (2014) Cancer cell lines for drug discovery and development. *Cancer Research*, **74**, 2377 – 2384.

Williams, T.N., Mwangi, T.W., Wambua, S., Peto, T.E., Weatherall, D.J., Gupta, S., Recker, M., Penman, B.S., Uyoga, S., Macharia, A., Mwacharo, J.K., Snow, R.W. & Marsh, K. (2005) Negative epistasis between the malaria-protective effects of alpha+-thalassemia and the sickle cell trait. *Nat Genet*, **37**, 1253 – 1257.

Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M.R., Leduc, M.S., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., Plon, S.E., Lupski, J.R., Beaudet, A.L., Gibbs, R.A. & Eng, C.M. (2013) Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *The New England journal of medicine*, **369**, 1502 – 1511.

Zaidi, S., Choi, M., Wakimoto, H., et al. (2013) De novo mutations in histone-modifying genes in congenital heart disease. *Nature*, **498**, 220 – 224.

Glossary

Abacavir	Anti-retroviral drug used to treat HIV and AIDS
Allele	One of a number of alternative forms of the same gene or genetic locus
Amino acids	Structural units which make up proteins. Amino acids are composed of amine and carboxylic acid functional groups together with a side-chain specific to each amino acid
Antibodies	Large proteins produced by plasma cells and used by the immune system to identify and neutralize foreign objects (such as bacteria and viruses). Antibodies recognise a unique part of the target called an antigen
Antibody-dependent cell-mediated cytotoxicity (ADCC)	Mechanism of cell-mediated immune defence whereby immune cells actively break down a target cell whose membrane-surface antigens have been bound by specific antibodies
Antigen	Substance/molecule which causes the immune system to produce antibodies against it
Anti-IL13	Monoclonal antibody currently under investigation to determine efficacy as an immunosuppressive drug in the treatment of asthma
Antisense oligonucleotides	Single strands of DNA or RNA derivatives which are complementary (in terms of the bases they contain) and hence can bind mRNA produced by a chosen gene and altering activity. Used both as research tools and in the clinic, for example, in treatment of certain genetic disorders
APC	See 'Familial adenomatous polyposis'
Bioinformatics	An interdisciplinary field for development of methods to study and process biological data. Bioinformatics has become an important part of many areas of biology, particularly in analysis of large quantities of biological data. Bioinfomatics has played an integral role within developments in genetics and genomics, for example in aiding genome sequencing
Biomarkers	Measurable indicators of biological state, often evaluated to examine normal biological or pathogenic processes and responses to drugs
BRAF	The gene *BRAF* encodes for protein BRAF which is involved in cell signalling and directing cell growth. BRAF has been shown to be mutated in some human cancers
BRCA1/BRCA2	Human tumour suppressor genes which encode for proteins BRCA1 and BRCA2 respectively. *BRCA2* and *BRCA1* are normally expressed in the cells of breast and other tissue, where they either help to repair damaged DNA or destroy cells where DNA cannot be repaired. If *BRCA1* or *BRCA2* are themselves damaged by mutation, damaged DNA (eg in breast cells) cannot be repaired properly and hence the risk for breast cancer is increased

Carbamazepine	Carbamazepine is an anticonvulsant which can cause various forms of hypersensitivity reactions, ranging from headache and dizziness to severe blistering reactions
Cell-line studies	Experiments with populations of cells (for example, certain types of cancer cells) grown outside of the body (*in vitro*)
Cetuximab	A monoclonal antibody which binds and inhibits epidermal growth factor receptors (EGFR) and is used in the treatment of bowel cancer
CFTR	The cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, which encodes for CTFR, a transporter protein/ion channel responsible for regulating salt and water movement between cells. Mutation(s) of the *CFTR* gene lead to dysregulation of epithelial fluid transport in the lung, pancreas and other organs and result in cystic fibrosis
Chorionic-villus sampling	Form of prenatal diagnosis used to determine fetal chromosomal or genetic disorders. Involves taking samples of the chorionic villus (placental tissue) derived from the fetus which are then tested for relevant disorders
Clinical trials	Clinical trials investigate whether a medical strategy, treatment, or device is effective and safe for use in humans. There are a number of phases within a full clinical trial, including Phase 1 in which a small group of people (20 – 80) is tested to evaluate safety and dosage range of a drug, Phase 2 where larger groups of people (100 – 300) are tested to determine efficacy and safety of a drug or device, and Phase 3 in which increased numbers of people (1000 – 3000) are tested for final confirmation of safety and efficacy
Communicable Disease(s)	Diseases which spread between people, animals or people to animals (and vice versa) Disease spread via airborne viruses or bacteria, or blood and other bodily fluids
Complementary DNA (cDNA)	DNA synthesized from a messenger RNA (mRNA) template in a reaction catalysed by the enzyme reverse transcriptase
Cytochrome P450	Family of enzymes with important roles in the degradation and elimination of drugs from the human body.
DNA (deoxyribonucleic acid)	Molecule which carries the genetic code for all living organisms
DNA polymerases	Enzymes which build DNA molecules through assembling nucleotides
DNA Microarray	Used in analysis of gene expression and in gene discovery. Uses base pairing rules for identification or matching of DNA samples
Duchenne muscular dystrophy	A disease, primarily displayed in males, characterised by muscle degeneration and caused by a mutation of the dystrophin gene

Efficacy	Capacity to produce effect. For example, in pharmacology the term efficacy is often used to describe the efficiency and effectiveness of a drug in achieving the desired response
EGFR	The epidermal growth factor receptor (EGFR) is involved in the proliferation and survival of cancer cells. EGFR was the first molecular target against which monoclonal antibodies (mAb) have been developed for cancer therapy (anti-EGFR therapy). For example, see Cetuximab
Epigenomics	Study of the complete set of epigenetic features. These are cellular and physiological traits not caused by changes in DNA sequence
Epistatic interactions	Interactions in which one gene affects the expression of another
Exons	Sections of DNA or mature RNA which contain coding information
Familial adenomatous polyposis (APC)	Inherited condition in which numerous polyps form in the large intestine. Colon cancer can occur when left untreated
Gene	Specific sequence/region of a chromosome responsible for encoding specific trait(s)
Gene knockdown/out	Experimental techniques by which expression of one or more of an organism's genes are reduced
Gene regulation	Mechanisms used by cells to increase or decrease production of gene products (RNA or protein)
Genetic locus	Specific location of a gene/DNA sequence
Genome	The complete set of genes in a cell or organism
Genome wide association studies (GWAS)	Examination of genetic variants, for example single-nucleotide polymorphisms (SNPs) in different individuals to determine whether a variant is associated with a trait
Genomics	The study of the genome. Genomic studies will often utilise and apply recombinant DNA, DNA sequencing methods and bioinformatics to sequence, assemble and analyse genome structure and function
Genotype	An organism's full hereditary information
Germ line/cells	Germ cells give rise to gametes. A gamete is a cell (eg ovum, sperm) which fuses with another during conception in sexual reproduction. The term germ line refers to the cell population which passes on genetic information to the progeny (offspring)
Globins	Family of haem-containing globular proteins involved in binding and/or transporting oxygen

Haemoglobin	Iron-containing oxygen transport protein found in red blood cells. Haemoglobin is made up from 4 polypeptide chains: 2 α- (alpha) and 2 β-(beta) chains
Hepatitis C	A virus that can infect and damage the liver
Hepatic LDL receptors	Low-density lipoprotein (LDL) receptors, of which there are many in the liver (hepatic LDL receptors), play a critical role in regulating the amount of cholesterol in the blood
Herceptin antibody	A monoclonal antibody which interferes with HER2, a receptor frequently found to be over expressed and responsible for causing cancer cells to reproduce uncontrollably in certain breast cancers
Hereditary nonpolyposis colorectal cancer (HNPCC)	Genetic condition with a high risk of colon and other cancers
Heterozygous/ homozygous	Humans contain two copies (alleles) of each gene, one from the father and one from the mother. Where both copies are identical the individual is considered homozygous whilst those with a mutation in one copy of the gene would be deemed heterozygous for that trait
HLA-DR	Cell surface receptor involved in numerous autoimmune conditions, disease susceptibility and resistance
Human leukocyte antigen (HLA) LA	Group of genes encoding for proteins on the surface of cells that are in turn responsible for regulation of the immune system
Hypercholesterolaemia	Presence of high levels of cholesterol in the blood, which although asymptomatic in itself can lead to atherosclerosis in cases of longstanding elevation of serum levels
Idiosyncratic drug toxicity	Idiosyncratic drug reactions are drug reactions that occur rarely and unpredictably amongst the population
Introns	Segments of DNA within genes that are translated into RNA but then removed (spliced out) so that they do not contribute to the mature RNA
Interferon-alpha	A protein which occurs naturally in the body in very small amounts. It can also be made outside the body and used as a drug, for example, when given to stimulate the body's immune system to fight some types of cancer
Ion channels	Ion channels are pore-forming membrane proteins which control the movement of ions (for example, sodium ions) across cell membranes
Ivacaftor	A drug treatment for cystic fibrosis
KRAS	Protein encoded by the *KRAS* gene. *KRAS* performs an essential function in normal tissue signalling. Mutation of the *KRAS* gene is an essential step in the development of many cancers

Low-density lipoprotein (LDL)	Lipoproteins are proteins which transport fat molecules (lipids) around the body. LDL cholesterol is often referred to as bad cholesterol because too much is unhealthy whereas HDL (high density lipoprotein) is often referred to as "good cholesterol" because it is protective against heart disease
Malaria	A serious tropical disease spread by mosquitoes
Major histocompatibility locus (HLA)	Human leukocyte antigen (*HLA*) genes encode for proteins on the surface of cells responsible for regulation of the human immune system
Metabolomics	Study of chemical processes involving metabolites. Metabolites have various functions including signalling and stimulatory and inhibitory effects on enzymes
Metastases	Tumours formed by cells that have spread from original site
Monoclonal antibodies	Identical antibodies produced by a single clone of cells or a cell line. Monoclonal antibodies are used to bind specific targets and have become an important tool in biochemistry, molecular biology and medicine. For example, in cancer research, development is ongoing to find monoclonal antibodies which bind specific antigens and induce an immunological response against target cancer cells
Monogenic disease	Disease resulting from modification/mutation within a single gene
Multidrug resistant protein (MDR) family	A family of several membrane proteins which are able to transport a wide range of anticancer drugs out of cells and hence are considered as possible suspects in unexplained cases of drug resistance
Multiplexing technology	Multiplex assays are tools for measuring multiple variables (for example, multiple proteins or mRNAs) in a single sample
Mutation	A permanent change in the genome of an organism. Mutations result from unrepaired damage to DNA or RNA (for example, as caused by radiation), mistakes in the process of replication, or from insertion or deletion of segments of DNA by mobile genetic elements such as transposons. Mutations are important in both normal and abnormal biological processes including evolution and cancer
Messenger RNA (mRNA)	Copied from genes by the process of transcription conveys genetic information from DNA to the ribosome for the synthesis of proteins
NRAS	*NRAS* encodes for the protein *NRAS* involved in regulating cell division
Oncogene	Gene with ability to cause cancer, often where mutated or expressed at increased levels
Pathogenesis	The origin and mechanism of development of a disease
PARP inhibitors	Compounds which inhibit the activity of the enzyme poly ADP ribose polymerase (PARP). Since several forms of cancer are more dependent on PARP than regular cells, PARP is an attractive target for cancer therapy

Plasmodium falciparum (P falciparum)	A parasite transported by mosquitos which causes malaria in humans. Malaria caused by this species is the most dangerous form of the disease
Periostin	Protein which supports adhesion and migration of epithelial cells. Known to be associated in asthma and several types of cancer
Phenotype	An organism's actual observed properties, such as morphology, development, or behaviour
PIK3CA	A gene responsible for encoding factors fundamental to cell proliferation and survival. *PIK3CA* mutations occur in approximately 1/3rd of breast cancers
Proteins	Large biological molecules consisting of long chains of amino acids. Proteins perform a vast range of functions including catalysing metabolic reactions and transporting molecules from one location to another. Proteins differ from one another in their sequence of amino acids
Proprotein convertase subtilisin kexin type 9 (PCSK9)	Enzyme encoded by the *PCSK9* gene. *PCSK9* acts in cholesterol homeostasis. Drugs that block *PCSK9* can lower low-density lipoprotein cholesterol and are currently being assessed to determine their safety and efficacy in humans in improving outcome in heart disease
Proteome	Complete protein complement of a cell or organism
Proteomics	The study of proteins, particularly their structures and functions
Polymorphism (genetic)	Differences in DNA sequence which make each human genome unique. Genetic polymorphism promotes diversity within a population, for example in relation to the different allelic forms that give rise to different blood types in humans
Recombinant DNA	DNA which has been synthesised artificially through combining constituents from different sequences
Recessive disorders	Disorders which occur only when an individual carries two mutant copies of the relevant gene
Ribonucleic acid (RNA)	Multifunctional, but its primary responsibility is to make proteins, according to the instructions encoded within a cell's DNA
Sequencing (DNA)	The process of determining the nucleotide order of a given DNA fragment
Sickle cell disease	An inherited blood disorder featuring abnormal structure and function of red blood cells
Sickle cell gene	A mutation in the β globin (*HBB*) gene which cause sickle cell disease
Single-nucleotide polymorphism (SNP)	Variation in DNA sequence of a single nucleotide (adenine, thymine, cytosine or guanine). Commonly found in the DNA between genes, SNPs can act as biological markers and help locate genes associated with disease

Somatic cell	Any cell within an organism not involved in the germline. For example, somatic cells make up internal organs, skin, bones, blood and connective tissue in humans
Subunit (protein)	A single protein molecule. Assembles with others to form a protein complex
Termination codon	A termination codon (or 'stop' codon) is a triplet of nucleotides within mRNA that signals a termination of translation, so stopping the further addition of amino acids to a protein
Thalassaemias	A group of inherited blood disorders characterised by severe anaemia due to defective synthesis of the α or β chains of adult haemoglobin
Transcriptome	The set of all RNA molecules transcribed in a cell (or population of cells)
Transcriptomics	Study of the transcriptome. Transcriptomics is an emerging and continually growing field with use in a number of areas, for example, biomarker discovery
Translation	Synthesis of proteins by the ribosome based on the sequence of nucleotides in the messenger RNA
Transporter	A protein responsible for movement of molecules across a biological membrane
Transposon	A DNA sequence able to change its position within the genome
Triple negative breast cancers	Breast cancers which do not express receptors for oestrogen, progesterone or HER2 (a protein)
Tyrosine kinases	Enzymes which function in a variety of processes, and are responsible for key events in the body
Toxicity	The degree to which a substance can cause damage for example to an organ or entire organism
Vector	In biology, the term vector usually refers to either an organism that transmits a pathogen to a host or a vehicle used to transfer genetic material to a target cell (for example, in molecular biology where small rings of DNA called plasmids are used to incorporate genes of interest into bacteria for study)

The Royal Society is a self-governing Fellowship of many of the world's most distinguished scientists drawn from all areas of science, engineering, and medicine. The Society's fundamental purpose, as it has been since its foundation in 1660, is to recognise, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity.

The Society's strategic priorities emphasise its commitment to the highest quality science, to curiosity-driven research, and to the development and use of science for the benefit of society. These priorities are:

- Promoting science and its benefits
- Recognising excellence in science
- Supporting outstanding science
- Providing scientific advice for policy
- Fostering international and global cooperation
- Education and public engagement

For further information
The Royal Society
6 – 9 Carlton House Terrace
London SW1Y 5AG

T +44 20 7451 2500
W royalsociety.org

Registered Charity No 207043
Issued: December 2015 DES3599