SCIENCE, MEDICINE, AND THE FUTURE

Translating genomics into improved healthcare

Aroon D Hingorani,[1 2] Tina Shah,[2] Meena Kumari,[1] Reecha Sofat,[2] Liam Smeeth[3]

[1]Genetic Epidemiology Group, Department of Epidemiology and Public Health, University College London, UK
[2]Centre for Clinical Pharmacology, Division of Medicine, University College London
[3]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK
Correspondence to: L Smeeth liam.smeeth@lshtm.ac.uk

Cite this as: *BMJ* 2010;341:c5945
doi: 10.1136/bmj.c5945

Since the work of Mendel,[1] genetic research has been punctuated by key, shifting advances (box 1), which culminated in the first draft sequence of the human genome in 2001.[2-4] Although most of the human genome sequence is shared by everyone, a small proportion varies between individuals. This variation, acting together with environmental factors, is thought to underlie differences in physiology, susceptibility to disease, and responses to drugs. We summarise the recent discoveries and review the implications of newly acquired knowledge for medical practice and public health.

SOURCES AND SELECTION CRITERIA

Because of the wide remit of this article, we did not attempt a systematic search covering the whole of translational genetics. Instead, we used personal collections of major journal articles and reviews accumulated individually by all authors over several years of academic work in translational genetics.

Box 1 | Key early milestones in genetic research

Discovery of nucleic acid (Meisscher 1869)

Identification of its key constituents—the organic bases adenine (A), thymine (T), guanine (G), and cytosine (C) (Kossel 1879) held on a backbone of deoxyribose or ribose sugars in DNA and RNA, respectively

The Boveri-Sutton chromosome theory of inheritance (1902)

The proposal that DNA (not protein) is the heritable genetic material that distinguishes organisms of different species (Chargaff 1950)

The discovery of the double helical structure of DNA (Crick and Watson 1953) provided the necessary insight into how base sequence information can be replicated and transmitted from parent to offspring. It also pointed the way to understanding how the DNA sequence is translated into the primary amino sequence of proteins, which, in turn, determines the structure and function of an organism. The unidirectional flow of information from DNA to RNA to protein was subsequently summarised in the central dogma (Crick 1958)

SUMMARY POINTS

Until recently the genetic basis of most common diseases remained elusive

Genome wide association analysis has now uncovered thousands of regions of human DNA where sequence variation influences susceptibility to common diseases

Common single nucleotide polymorphisms associated with disease are distributed across multiple chromosomal regions, have modest affects on disease risk, act additively, and explain only part of disease heritability

Association analysis is starting to provide data on the causes of common human diseases that should accelerate the design and development of new treatments

Emerging technologies, including rapid, less costly sequencing of whole genomes, bring the prospect of better diagnostics and treatment

Clinicians will need to keep updated on genetic advances that have healthcare applications

What is known about the molecular basis of single gene disorders?

Uncommon single gene (Mendelian) disorders such as cystic fibrosis and familial forms of type 2 diabetes, colon cancer, and breast cancer are caused mainly by mutations, usually in the coding sequence of a gene, that produce a major structural or functional disruption of an encoded protein. A mutation is both necessary and sufficient for disease to develop, and the pattern of disease transmission from parent to offspring is predictable. Mendelian disorders are categorised as autosomal dominant, autosomal recessive, or X linked, depending on whether the mutation is located on an autosome or sex chromosome and on whether one or two mutant alleles are needed for the disease to manifest itself. The responsible genes were identified mainly through analysis of DNA samples from multigenerational families (pedigrees) with affected and unaffected members, a technique called linkage analysis.[5 6] Coinheritance of a genetic marker of known chromosomal location with the disease phenotype allows the position of the disease gene to be "mapped" in these families. For some conditions, mutations in different genes can produce the same disease phenotype, a phenomenon known as locus heterogeneity.[7] DNA sequencing of candidate genes in the linked region allows the disease causing mutations to be identified. For some diseases, the precise disease causing mutations, even within the same gene, can also differ from family to family, and this is referred to as allelic heterogeneity.[8] These discoveries provided great insight into the normal biological function of the affected genes and proteins and in some cases have led to the development of predictive genetic

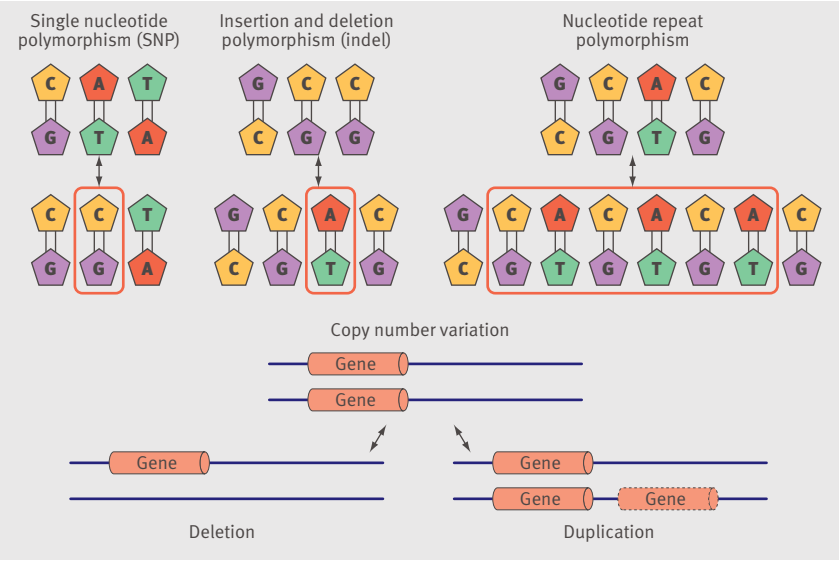

Fig 1 | The spectrum of common genetic variation includes single nucleotide polymorphism, insertion and deletion polymorphism, nucleotide repeat polymorphism, and copy number variation, all of which may affect coding or non-coding regions of DNA

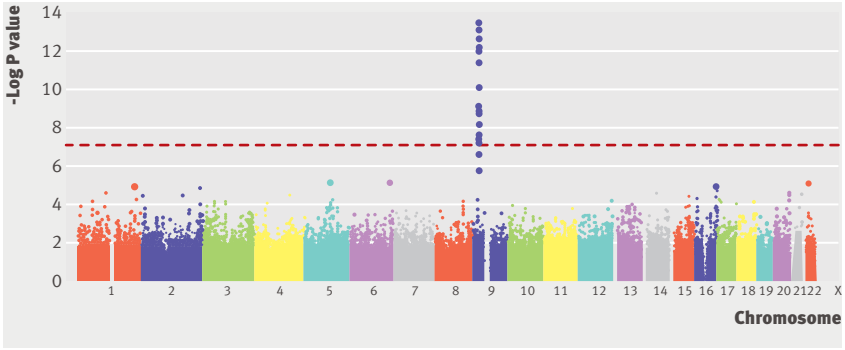

Fig 2 | Manhattan plot from a genome-wide association study of myocardial infarction (adapted, with permission, from Samani and colleagues[13]). The x axis refers to points along the genome (separated by chromosome) at which each of the several hundred thousand single nucleotide polymorphisms (SNPs; represented by a dot) evaluated are located. The y axis refers to the negative logarithm of the P value for a test of association between each SNP and the binary outcome—the presence or absence of disease. For example, a –log P value of 7 (shown by the dashed red line) equates to a P value for the association of an SNP with disease of 1×10^{-7} or 0.0000001, and a –log P value of 7.3 equates to 5×10^{-8}

tests, gene based or drug treatments. A comprehensive list of single gene disorders and their molecular basis can be found online (www.ncbi.nlm.nih.gov/omim/).

What is known about the genetic basis of complex diseases?

Diseases that commonly affect adults (such as cardiovascular disease, diabetes, Parkinson's disease, Alzheimer's disease, and common cancers) result from a more complex interplay between genetic and environmental factors. Such diseases exhibit familial clustering, but there is no clear inheritance pattern because of the polygenic aetiology and the substantial contribution from the environment. Family based linkage analysis can rarely identify genetic mutations associated with these disorders because people are typically over 50 years before clinical presentation; by this time their parents may be dead and susceptible children too young to manifest the disorder. The "common-disease common-variant" hypothesis proposes that these diseases arise from many common DNA variants, each with a modest influence on disease risk.[9 10] The presence of a particular genetic variation is neither necessary nor sufficient to cause disease but confers a modest increase in risk.

After publication of the draft human sequence in 2001, attention turned to cataloguing and studying the effects of common variations in the DNA sequence on the risk of common diseases. Single nucleotide polymorphisms (SNPs) are variations at a single base pair (fig 1) and are the most common type of human sequence variation, occurring about every 500-1000 DNA base pairs. Less common types of variation include single nucleotide insertions and deletions (indels) and deletion or duplication of longer tracts of DNA (copy number variation) (fig 1).

The SNP Consortium developed the SNP database (www.ncbi.nlm.nih.gov/SNP/) to identify the most common single nucleotide variations in the genome. Each SNP in this database is allocated a unique reference number. The Human HapMap consortium (http://hapmap.ncbi.nlm.nih.gov/index.html.en) was then established to quantify the association between SNPs in the genomes of human populations with differing ancestry (linkage disequilibrium). Coupled with new technological developments, this provided the framework to develop arrays capable of typing around 500 000 SNPs across the genome. Information on another million or more SNPs could then be inferred on the basis of the known associations between SNPs. These arrays or "SNP chips" provide a cost effective way to genotype many thousands of people.[11]

To investigate the genetic basis of common diseases, genome-wide case-control studies have compared the frequency of typed (and inferred) SNPs in large numbers of unrelated people affected by a disease and unaffected unrelated controls.[12] Instead of the single exposure evaluated in a non-genetic case-control study (such as smoking), these studies examine hundreds of thousands of genetic exposures simultaneously. Points in the genome at which the frequency of alleles differs between cases and controls harbour the genetic variants contributing to disease risk. Figure 2 shows the results of such a study of myocardial infarction.[13] The Office of Population Genetics website contains an up to date list of genome-wide association studies and publications (www.genome.gov/GWAStudies/).

Genome wide association studies involve hundreds of thousands of tests of association, so the number of false positive SNP-disease associations would be high if conventional thresholds were used to determine significant associations. Thus, stringent criteria are used before "genome-wide significance" can be declared. Typically, P value thresholds of 1×10^{-7} or 5×10^{-8} are used.[14] A P value of 5×10^{-8} can be thought of as a P value of 0.05 with a Bonferroni correction for one million statistical tests—a typical number of variants genotyped or inferred in a genome-wide association study. Even with such stringent thresholds, positive findings are routinely replicated in independent datasets to confirm or refute association. Because SNPs of smaller effect size can be identified by increasing the available sample size, research consortiums focus on a single disease so that data from several case-control collections can be pooled and summarised using meta-analysis. This approach has increased the number of genetic loci identified for many disorders.[15]

Box 2 | Insight into disease aetiology from genome-wide association studies

In Crohn's disease, many of the single nucleotide polymorphisms (SNPs) associated with disease susceptibility lie in and around genes concerned with autophagy, which was previously not considered an important disease mechanism in Crohn's disease.

In type 2 diabetes many of the loci encode proteins concerned with insulin secretion rather than insulin signalling, which had previously been the focus of research.

Several of the genetic loci associated with the risk of coronary heart disease influence low density lipoprotein-cholesterol, but some are located distant from known genes and seem not to influence any of the known risk factors for myocardial infarction. Thus, previously unknown disease mechanisms might be at work.

Data from different disorders show that some genetic regions or SNPs affect the risk of more than one disease. For example, different SNPs in the same region on chromosome 12 influence the risk of coeliac disease, type 1 diabetes, and myocardial infarction. The same SNP near the *TCP2* gene on chromosome 8 is associated with the risk of developing type 2 diabetes and prostate cancer. These, and other examples indicate that some common disorders have a partially overlapping aetiological basis, and this may lead to new disease taxonomy.

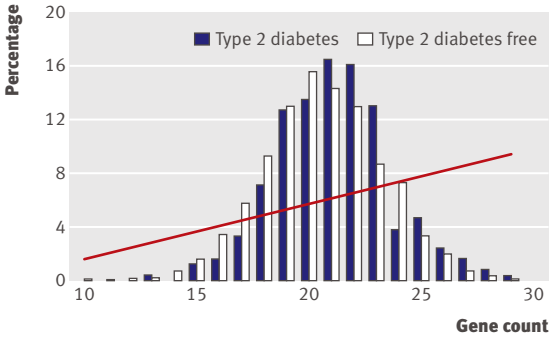

Fig 3 | Association between the population frequency distribution of type 2 diabetes risk alleles (bars) and risk of incident diabetes (red line) in the Whitehall II study (adapted from Talmud and colleagues[19])

How might findings from genome-wide association studies affect healthcare?

Accumulated evidence from such studies is now helping to determine the direction of future research and to clarify where the future healthcare applications will be. Typically, many genetic regions contribute to increased risk of complex disease (20 loci have been identified for type 2 diabetes and 40 for Crohn's disease), but the effect at each region is weak (5-10% increase in risk) and seems to be additive and independent. As a result, genome wide association studies to date have explained only a small part of the heritability of common disorders (table). The findings currently have limited value for predicting disease risk but may have other important implications for healthcare provision (box 2).

Can genetics help predict risk of common disease?

Predictive genetic tests based on findings from genome-wide association studies are being offered commercially, despite concerns about their clinical value.[16 17] With a few notable exceptions (such as age related macular degeneration[18]; table), carrying any one common risk allele increases the chance of experiencing a common disease event by only a fraction (typically 5-25%). This makes tests based on only one SNP poorly predictive. Several risk alleles, often on different genes, may contribute to increased risk of disease, so would information from a panel of common modest effect alleles be better at predicting disease than a single common risk allele? A person with 10 risk alleles, each in a different gene and each conferring a 20% increase in risk, might be expected to be at double the risk of disease compared with someone carrying none, and such a high risk person might benefit from a targeted preventive intervention. However, people with 10 risk alleles are rare in the population. For example, if the average frequency of a risk allele in a population is 30% (0.3), the probability of inheriting 10 such independent alleles is 0.3^{10} (0.0000059). Because the frequency distribution of common independently inherited risk alleles is normal (bell shaped) and the association with risk is additive, more cases of disease would be expected in the many people with an intermediate number of risk alleles than the minority with a large number of alleles. Thus, the frequency distributions of risk alleles should overlap substantially among eventual cases and controls, making it difficult to separate the two groups by the number of risk alleles carried (fig 3).

In disease prevention the aim is often to stratify risk rather than to discriminate events. This is because many preventive interventions produce the same relative risk reduction whatever the risk so that the absolute benefit is larger (and the number needed to treat smaller) in people at high risk. If the number needed to treat to prevent one event in a person with 10 risk alleles was 100, the number needed to screen to prevent one event would be 100/0.000006 (16 666 666). Although genotype based tests are cheap, screening such a large number of people just to alter risk of disease in one would be very costly.

Genetic tests that capture a wider range of variability in a given gene or region, rather than simply a few SNPs, may be better for prediction. The discovery of rare or intermediate frequency alleles that have a much larger effect size than common HapMap alleles may also open up greater opportunities.[20] Tests of rarer alleles may be useful in family based screening, such as that used for monogenic familial hypercholesterolaemia, which currently uses cholesterol measurement rather than genotyping.[21]

Recently identified genetic variants associated with disease*

Disease	Examples of research consortiums	Examples of associated genes	Insight on mechanism
Crohn's disease	WTCCC	*NOD2*, *IRGM*, *ATG16L1*, *IL23R*, *PTPN2*, *PTGER4*	Autophagy, immunity
Type 2 diabetes	DIAGRAM WTCCC DGI	*TCF7L2*, *CAPN10*, *VNTR*, *SLC2A2*, *IRS1*, *CDKAL1*, *FTO*, *CDKN2B*	Insulin secretion unknown
Myocardial infarction and coronary heart disease	CARDIOGRAM WTCCC	*APOE*, *PCSK9*, *CELSR2*/*PSRC1*, *CDKN2A*/*CDKN2B*	Low density lipoprotein-cholesterol pathway unknown
Age related macular disease		*CFH*, *C2*/*CFB*, *C3*, *CFI*, *ARMS2*, *TIMP3*, *LIPC*, *CETP*, *LPL*, *ABCA1*	Complement pathway, high density lipoprotein-cholesterol pathway?
Breast cancer	Asia Breast Cancer Consortium	*BRCA1*, *BRCA2*, *PTEN*, *TP53*, *FGFR2*, *TNRC9*	DNA and cell repair, tumour suppression pathway
Meningococcal disease	International Meningococcal Genetics Consortium	*CFH*/*CFHR3*	Complement pathway
Tuberculosis	African Tuberculosis Genetics Consortium and WTCCC	*GATA6*, *CTAGE1*, *RBBP8*, *CABLES1*	As yet unknown

*The variants vary in their strength of effect. Some variants (such as *BRCA1* in breast cancer and *CFH* in age related macular disease) are strongly associated with disease risk and thus highly predictive of disease risk, whereas others (such as those listed for myocardial infarction and coronary heart disease) are only weakly associated with disease risk and provide little predictive power.

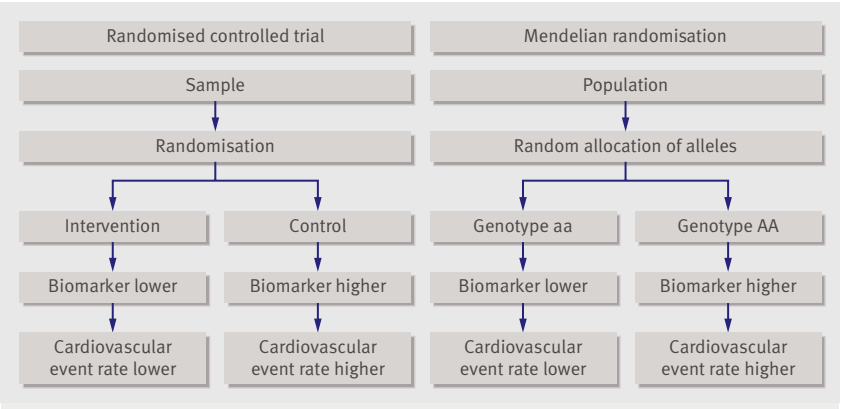

Fig 4 | Conceptual parallels between a randomised controlled trial and a Mendelian randomisation experiment to judge the causal relevance of a biomarker associated with risk of cardiovascular disease (adapted from Casas and colleagues[28])

Can genetics improve understanding of the non-genetic causes of disease?

The genotype is unique among naturally occurring differences between people because it is allocated at random,[1] fixed throughout life, and not modified by disease. Interpretation of genetic associations is therefore not limited by confounding (where exposure and disease seem to be associated because of common association with a third factor) or reverse causation (where the association between an exposure and disease is caused by the disease itself leading to an alteration in the exposure).[22-24] Thus when the function of a gene is known, its association with a disease (however weak) can provide clear insight into the causal mechanisms leading to disease (fig 4). For example, *ADH1* (alcohol dehydrogenase) gene variants are associated with differences in long term alcohol consumption, which is otherwise difficult to measure accurately. *ADH1* variants can provide a useful index of long term usual alcohol consumption. Association between these variants and oesophageal cancer provides evidence for a causal role of alcohol consumption in the disease.[25] Another example is C reactive protein (CRP) and its role in coronary heart disease. Raised concentrations of CRP are associated with an increased risk of heart disease. However, bias and confounding by reverse causality may partially or wholly explain these associations, so we do not know whether lowering CRP would reduce the risk of cardiovascular events. Genetic variants that influence CRP values are less prone to confounding, and reverse causality is not a problem. The presence or absence of an association between CRP genetic variants and disease can thus provide clear evidence on whether CRP actually plays a causal role in disease.[26 27]

Can genetics lead to improved therapeutics?

Using pharmacogenetics to develop genotype based predictive tests of drug response may help to personalise or stratify therapeutic interventions at an individual or group level. The number of pharmacogenetic tests currently approved for clinical use is limited. A recent systematic review of pharmacogenetic studies highlighted several methodological and other problems in this area,[29] notably small underpowered studies and the widespread use of surrogate outcome measures. However, genetic loci associated with the risk of statin myopathy have recently been identified,[30] as well as loci associated with hypersensitivity to the protease inhibitor abacavir,[31] response to interferon treatment in hepatitis C infection,[32] and dose requirements in people taking warfarin.[33] Pharmacogenetic testing may eventually become common in some therapeutic areas, but the clinical value and cost effectiveness of emergent pharmacogenetic tests have not yet been subject to the same level of scrutiny and careful appraisal as other diagnostic tests.

As well as offering the potential for new pharmacogenetic tests, emerging genome-wide association studies of drug response could provide new insights into the pathways by which even well established drugs are handled by the body. The concept of pharmacogenetics is expanding to incorporate not only the use of genetic data to guide treatment but also to inform drug development. This is because studies in populations of variants in genes encoding a drug target protein can be considered to be a type of natural randomised trial and could be used to help predict the on-target effect of modifying the same target pharmacologically. For example, common variants in the *CETP* gene, which encodes cholesteryl ester transfer protein, the target of the CETP inhibitor torcetrapib, were associated with the same lipid and lipoprotein changes seen with torcetrapib treatment but were not associated with high blood pressure, an off-target effect of torcetrapib.[34] The hope is that by providing randomised evidence of the effects of drugs in humans without requiring participants' exposure to a new molecule of uncertain safety and efficacy new drug targets might be validated and the risks of late stage failure in drug development reduced through the application of genetic studies during early drug development.

Future directions

Much of the heritability of common diseases cannot be explained by common SNPs, so the focus is now moving towards other types of genetic variation, such as copy number and rare single nucleotide variation. Interest in heritable changes in gene expression caused by other processes, such as DNA methylation or histone modification, is also increasing.[20 35]

Some alleles with large effects on disease risk are likely to be rare (because natural selection reduces their frequency over time) and thus are not well represented on whole genome arrays. Efforts in this area are being stimulated by another technological advance, the ability to sequence a single human genome (exome sequencing) using "next generation" sequencing technology. In contrast to the first draft human genome sequence, which took several years and millions of pounds to complete, this now takes a few days and costs about £8000 (€9000; $12 600). Systematic rare variant discovery is now being undertaken as part of the international 1000 genomes project (www.1000genomes.org/page.php) and the UK 10k Consortium. Whole exome sequencing is expected to identify rare variants that influence risk in common disorders and identify mutations underlying sporadic single gene disorders that have not been amenable to linkage analysis because of the absence of multigenerational pedigrees. The variants that are discovered may eventually prove valuable as family based genetic tests.

ADDITIONAL EDUCATIONAL RESOURCES

Resources for healthcare professionals

Online Mendelian Inheritance in Man (OMIM) (www.ncbi.nlm.nih.gov/omim)—Comprehensive database of human genes and genetic phenotypes that contains information on all known mendelian disorders and more than 12 000 genes

The SNP Consortium Database (www.ncbi.nlm.nih.gov/SNP/)—National Center for Biotechnology Information database of all known single nucleotide polymorphisms

Human HapMap Consortium (http://hapmap.ncbi.nlm.nih.gov/index.html.en)—Public resource developed by a partnership of scientists and funding agencies to help researchers find genes associated with human disease

Office of Population Genetics (www.genome.gov/GWAStudies/)—Catalogue of published genome wide association studies

1000 Genomes Project (www.1000genomes.org/page.php)—Catalogue of human genetic variation

UK Biobank Project (www.ukbiobank.ac.uk/)—Resource to support a diverse range of research intended to improve the prevention, diagnosis, and treatment of illness, and the promotion of health

Resources for patients and the public

BBC Frontiers (www.bbc.co.uk/programmes/b00vkxy0)—BBC Radio 4 documentary about PLX4032, a new cancer drug based on new genetic knowledge

Wellcome Trust Centre for Human Genetics: Understanding human genetics (www.well.ox.ac.uk/monaco/Group/Complex_Disorders.shtml)—Good introduction to genetics

NHS National Genetics Education and Development Centre (www.geneticseducation.nhs.uk/)—Includes a wide range of introductory material about many areas of genetics

NHS Choices (www.library.nhs.uk/GENETICCONDITIONS/ViewResource.aspx?resID=287600&tabID=289&catID=8700)—Site that discusses and analyses news stories on genetic breakthroughs

NHS UK Genetic Testing Network (www.ukgtn.nhs.uk/gtn/Home/Public)—Advises the NHS on all aspects of genetic testing; this section for the public has links to a range of other resources

GLOSSARY

Allele: Alternative forms of a genetic locus; a single allele for each locus is inherited separately from each parent

Autosome: A chromosome not involved in sex determination

Complex disease genetics: The study of the patterns of inheritance of common diseases resulting from the combined action of alleles of more than one gene (such as heart disease, diabetes, and some cancers)

Epigenetics: The study of heritable changes in gene expression that are not caused by changes in DNA sequence

Genotype: The specific combination of two alleles inherited for a particular gene

Human genome: The total set of chromosomes found in an individual

Linkage analysis: Analysis of DNA markers that are near or within a gene of interest among families to identify the inheritance of a disease causing mutation in a given gene

Linkage disequilibrium: Where alleles occur together more often than can be accounted for by chance alone

Locus: The physical position of a gene or marker on a chromosome

Marker: An identifiable physical location on a chromosome (for example, a single nucleotide polymorphism or a gene), the inheritance of which can be assessed. Markers can be regions of DNA that are expressed (genes) or a segment of DNA with no known coding function but whose pattern of inheritance can be determined

Mutation: Any heritable change in DNA sequence that occurs in less than 1% of the population

Phenotype: The observable characteristics of an organism produced by the genotype (or environment, or both)

Sex chromosome: The X and Y chromosomes in humans that determine sex

Single nucleotide polymorphism (SNP): Variation at a single base pair (A, T, C, or G) in the DNA sequence

The SNP Consortium: A public-private partnership that was established to identify and map common SNPs. As part of the international HapMap project, it aimed to generate a high quality, extensive, publicly available map of SNPs as markers evenly distributed throughout the human genome in different populations

Finally, analysis of the effects of newly discovered genetic loci in representative population based cohort studies (not case-control studies) are beginning to provide better information on the absolute (not relative) risk of common diseases, as well as insight into the modification of genetic effects by environmental factors. The UK Biobank project (www.ukbiobank.ac.uk/)—a prospective study that has recruited more than 500 000 volunteers, stored millions of biological samples (including DNA), and recorded information on lifestyle measures—will provide a new resource for scientists studying the environmental and genetic determinants of a wide range of common diseases in future decades.

Conclusions

As the use of whole genome sequencing becomes more widespread, an improved understanding of the causes of disease, better targeted drug treatments, and perhaps prediction of risk are realistic expectations. As in any area of medical advance, rigorous evaluation of new genetic based technologies will be needed. A major challenge for clinicians will be to keep updated on genetic advances with potential healthcare applications and to develop the ability to critically appraise research findings in this fast moving field.

Contributors: LS and AH had the original idea for the article and developed the outline. All authors contributed to drafts and approved the final draft. LS and AH are guarantors.

Funding: LS is supported by a Wellcome Trust Senior Clinical Fellowship (082178). AH is supported by a senior fellowship from the British Heart Foundation (FS05/125) and is the principal investigator on a MRC Biomarker Award with funding from Pfizer. The funders had no role in the manuscript.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work (TS, MK, RS); LS had support from a Wellcome Trust senior clinical fellowship and ADH had support from a British Heart Foundation senior fellowship for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years (LS, TS, MK, RS); ADH is on the editorial board of *Drug and Therapeutics Bulletin*, a BMJ Group publication, has provided non-remunerated advice to GlaxoSmithKline and London Genetics, and has received honorariums for speaking at educational meetings with sponsorship from the drug industry, which have been donated in whole or large part to charity. ADH has received a Medical Research Council biomarker research award on which Pfizer is a co-sponsor; no other relationships or activities that could appear to have influenced the submitted work (all authors).

Provenance and peer review: Commissioned; externally peer reviewed.

1 Mendel G. *Experiments in plant hybridization*. 1865.
2 The International Human Genome Mapping Consortium. A physical map of the human genome. *Nature* 2001;409:934-41.
3 International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860-921.
4 Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* 2001;291:1304-51.
5 White R, Lalouel JM, Leppert M, Lathrop M, Nakamura Y, O'Connell P. Linkage maps of human chromosomes. *Genome* 1989;31:1066-72.
6 Pericak-Vance MA, Bebout JL, Gaskell PC Jr, Yamaoka LH, Hung WY, Alberts MJ, et al. Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage. *Am J Hum Genet* 1991;48:1034-50.
7 Haddad L, Day IN, Hunt S, Williams RR, Humphries SE, Hopkins PN. Evidence for a third genetic locus causing familial hypercholesterolemia. A non-LDLR, non-APOB kindred. *J Lipid Res* 1999;40:1113-22.
8 Benlian P, Amselem S, Loux N, Pastier D, Giraud G, de Gennes JL, et al. ALDL receptor gene homozygous mutation: PCR amplification, direct genomic sequencing, associated haplotype, rapid screening for frequency. *Ann Genet* 1990;33:65-9.

bmj.com archive

Previous articles in this series

- Capsule endoscopy (*BMJ* 2009;339:b3420)
- Cell-free fetal DNA and RNA in maternal blood: implications for safer antenatal testing (*BMJ* 2009;339:b2451)
- Preventing transmission of maternally inherited mitochondrial DNA diseases (*BMJ* 2009; 338:b94)

9 Lander ES. The new genomics: global views of biology. *Science* 1996;274:536-9.
10 Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17:502-10.
11 De La Vega FM, Dailey D, Ziegle J, Williams J, Madden D, Gilbert DA. New generation pharmacogenomic tools: a SNP linkage disequilibrium map, validated SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies. *BioTechniques* 2002;(Jun suppl):48-50, 52, 54.
12 Wellcome Trust Case Control Consortium. Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 2007;447:661-78.
13 Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443-53.
14 Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genom* 2008;9:516.
15 Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008;40:638-45.
16 Edelman E, Eng C. A practical guide to interpretation and clinical application of personal genomic screening. *BMJ* 2009;339:b4253.
17 Annes JP, Giovanni MA, Murray MF. Risks of presymptomatic direct-to-consumer genetic testing. *N Engl J Med* 2010;363:1100-1.
18 Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;308:385-9.
19 Talmud PJ, Hingorani AD, Cooper JA, Marmot MG, Brunner EJ, Kumari M, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. *BMJ* 2010;340:b4838.
20 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747-53.
21 Heath KE, Humphries SE, Middleton-Price H, Boxer M. A molecular genetic service for diagnosing individuals with familial hypercholesterolaemia (FH) in the United Kingdom. *Eur J Hum Genet* 2001;9:244-52.
22 Davey Smith G, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1-22.
23 Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol* 2004;33:30-42.
24 Hingorani A, Humphries S. Nature's randomised trials. *Lancet* 2005;366:1906-8.
25 Lewis SJ, Davey Smith G. Alcohol, ALDH2, and esophageal cancer: a meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. *Cancer Epidemiol Biomarkers Prev* 2005;14:1967-71.
26 Casas JP, Shah T, Cooper J, Hawe E, McMahon AD, Gaffney D, et al. Insight into the nature of the CRP-coronary event association using Mendelian randomization. *Int J Epidemiol* 2006;35:922-31.
27 Elliott P, Chambers JC, Zhang W, Clarke R, Hopewell JC, Peden JF, et al. Genetic loci associated with C-reactive protein levels and risk of coronary heart disease. *JAMA* 2009;302:37-48.
28 Casas JP, Shah T, Hingorani AD, Danesh J, Pepys MB. C-reactive protein and coronary heart disease: a critical review. *J Intern Med* 2008;264:295-314.
29 Holmes MV, Shah T, Vickery C, Smeeth L, Hingorani AD, Casas JP. Fulfilling the promise of personalized medicine? Systematic review and field synopsis of pharmacogenetic studies. *PLoS One* 2009;4:e7960.
30 SEARCH Collaborative Group. SLCO1B1 variants and statin-induced myopathy—a genomewide study. *N Engl J Med* 2008;359:789-99.
31 Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, et al. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet* 2002;359:727-32.
32 Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of IL28B with response to pegylated interferon-[alpha] and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009;41:1105-9.
33 Freeman BD, Zehnbauer BA, McGrath S, Borecki I, Buchman TG. Cytochrome P450 polymorphisms are associated with reduced warfarin dose. *Surgery* 2000;128:281-5.
34 Sofat R, Hingorani AD, Smeeth L, Humphries SE, Talmud PJ, Cooper J, et al. Separating the mechanism-based and off-target actions of cholesteryl ester transfer protein inhibitors with CETP gene polymorphisms. *Circulation* 2010;121:52-62.
35 Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, et al. Parental origin of sequence variants associated with complex diseases. *Nature* 2009;462:868-74.

ANSWERS TO ENDGAMES, p 1057. For long answers go to the Education channel on bmj.com

ON EXAMINATION QUIZ

Ketoconazole treatment

Answer C is correct.

STATISTICAL QUESTION

Skewed distributions

Answers *a* and *b* are true, whereas *c* and *d* are false.

ANATOMY QUIZ

Axial computed tomography image through the face (bone windows)

A Nasal concha (turbinate)
B Lateral pterygoid plate
C Ramus of mandible
D Vomer
E Opening of eustachian tube
F Torus tubarius
G Lateral pharyngeal recess (of Rosenmueller)

CASE REPORT

Pain after amputation of the lower leg

1 Phantom limb pain.
2 Increasing evidence suggests that central and peripheral mechanisms contribute to phantom limb pain.
3 Pain before amputation and genetic predisposition may contribute to the development of neuropathic pain.
4 Preoperative epidural insertion and nerve sheath catheters have been investigated with varying success.
5 Conservative options include transcutaneous electrical nerve stimulation, hypnosis, massage, ultrasound, and mirror box therapy. Medical options include anticonvulsants, tricyclic antidepressants, opiates, NMDA antagonists, and benzodiazepines. Surgical options include neurectomy, stump revision, and dorsal column stimulation.