What will whole genome searches for susceptibility genes for common complex disease offer to clinical practice?

H. Lango & M. N. Weedon

From the Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, UK


In the developed world the majority of disease results from common, but complex disorders such as diabetes, obesity and cancer. Genetic variation explains a large proportion of an individual’s risk of developing these diseases; however, success in identifying the particular gene variants involved has been limited. Recent advances in high-throughput genotyping technology, and a better understanding of the genetic architecture of complex disease has led to the development of genome-wide association studies (GWA), which are providing novel and important insights into disease processes. The results from these studies could be of substantial clinical importance in the relatively near future. In this review, we present some recent, exciting findings from studies that have used the GWA approach, and discuss the clinical application of identifying disease susceptibility genes and variants.

Keywords: common genetic susceptibility variants, complex disease, genetics, genome-wide association, polygenic.

Introduction

Many human diseases have a genetic component. Some diseases are caused entirely by a genetic mutation, and much success has been had in identifying the genes that, when mutated, cause these monogenic disorders [1]. Over 1500 genes for monogenic diseases such as cystic fibrosis and maturity onset diabetes of the young (MODY) have been identified (http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM). Identifying these monogenic disease genes, has led to deep insights into the biology of these and closely related disorders and has led to the development of therapeutic measures [2]. Monogenic diseases, however, are relatively rare within populations, explaining only a small percentage of the overall disease burden.

In the developed world the majority of disease results from common, but complex disorders such as diabetes, obesity and cancer. Environmental and lifestyle factors such as diet, exercise and smoking are important risk factors for the development of these diseases; however, twin, family, admixture and migration studies have also demonstrated a large genetic component to an individual’s disease risk. Heritability estimates for some common diseases are given in Table 1. A heritability of 50% suggests that half of the variation in disease risk for an individual in a population can be explained by genetic variation.

Success in identifying the genes and variants which explain the genetic component of common complex disease has been slow, but recent advances in the understanding of the genetic architecture of complex diseases, together with advances in high-throughput genotyping technology, has led to a new era of genetic analyses – genome-wide association (GWA) studies – which are providing novel and important insights into common polygenic disorders. These findings could be of substantial clinical importance in the relatively near future.
In this review, we first discuss the genetic architecture of common complex disease. We then describe how we go about finding polygenic disease genes using the GWA study approach. We go on to present some recent, exciting findings in the genetics of complex disease. Finally, we discuss the clinical application of finding disease susceptibility genes and variants.

### The genetic architecture of common complex disease

In the search for genetic variants (see Box 1) that predispose to common disease, the most powerful strategies depend on the (often unknown) underlying genetic model. For monogenic disorders the genetic model can, most often, be defined as dominant (where a single mutated gene causes the disease) or recessive (where two copies of the disease genes are required). These models produce distinct inheritance patterns in families that allow, with sufficiently large families, the single highly penetrant causal mutations to be identified through classical linkage studies. This is in contrast to polygenic diseases that strongly cluster in families and are highly heritable, but do not demonstrate simple inheritance patterns. Figure 1 compares monogenic and polygenic disease inheritance. One explanation for the polygenic inheritance pattern is that many (tens to hundreds) of common genetic variants (minor allele frequency >1% in the population), each with only a modest effect on disease risk (affecting relative risk by <50%) are responsible for the heritability of polygenic disease. This is the common disease/common variant hypothesis (CDCV) [3].

A number of people have argued against the CDCV hypothesis, suggesting that rare, modest-risk alleles may explain a large proportion of the variation in susceptibility to common disease [4–6], and it is likely that both common and rare alleles are important in polygenic disease. However, given the current near impossibility of reliably detecting effects of rare alleles (owing to sample size and sequencing constraints), studies have focused on finding common disease alleles.

### Finding common susceptibility genes for common complex disease: past approaches

Until very recently, the two major strategies that have been used to identify complex disease genes were positional cloning through genome-wide linkage scans, and candidate gene association studies. The
The linkage approach has been very successful in identifying genes responsible for monogenic diseases following Mendelian pattern of inheritance, but very few linkage studies of diseases with polygenic inheritance patterns have provided reproducible evidence for linkage, and only a small number of disease genes have been identified through this approach [7].

In their 1996 paper, Risch and Merikangas [8] showed that genetic association studies are much more powerful than linkage studies in identifying common variants of modest effects. Since then genetic association studies have become the method of choice for identifying common gene variants predisposing to disease. At their core, genetic association studies of disease are straightforward. In its simplest form, when applied to unrelated individuals, the frequency of a variant (allele) of a single nucleotide polymorphism (SNP) is determined in a sample of subjects with a particular disease, and a sample of subjects without the disease. A statistically significant higher frequency of a variant of a SNP (or other genetic variant) in the cases versus the controls suggests that it is associated with a particular disease.

There are around 10 million SNPs in the human genome and, until recently, it was not possible to assay such a large number of variants. Instead, many studies...
used a candidate gene approach and analysed a small number of candidate SNPs in these genes. However, most of these studies used sample sizes that provided insufficient power to detect the association unless the allele had extremely high odds ratio [9]. When an association was detected it was usually a false-positive or, in a few cases, a true-positive with a greatly overestimated risk effect [10, 11]. In type 2 diabetes, for example, years of research had identified only two reproducibly associated susceptibility variants, in PPARG [12] and KCNJ11 [13] genes. Lessons have been learned from the poor performance of the candidate gene studies (Box 2), and the power of the association approach is now being demonstrated with the advent of the large-scale GWA study.

**The era of the genome-wide association study**

We are now in the era of the GWA study, whereby using the phenomenon of linkage disequilibrium (LD; association between alleles in a population because of their proximity on a chromosome such that recombination at meiosis has not had a chance to ‘separate’ them) and DNA chips that allow the assaying of several hundred SNPs simultaneously, we are able to evaluate a large proportion of the 10 000 000 common genetic variants across the human genome. The Affymetrix 500 K chip (Affymetrix Inc., Santa Clara, CA, USA) captures approximately 65% of common variation across the genome, and the Illumina 317 K chip (Illumina Inc., San Diego, CA, USA) captures approximately 75% [14, 15]. Coverage of the genome for these chips can be substantially increased using the statistical technique of imputation, whereby information from a group of typed SNPs and LD information from the HapMap (http://www.hapmap.org) can be used to infer genotypes at untyped SNPs [16]. We now discuss some of the recent exciting findings from GWA studies.

**Recent findings in genome-wide association studies**

The largest, most comprehensive GWA study to date was carried out by the Wellcome Trust Case-Control Consortium (WTCCC) [17]. The WTCCC study used the Affymetrix 500 K GeneChip and examined 3000 shared controls and 2000 cases, all of UK Caucasian ancestry, for seven common complex diseases: bipolar disorder, coronary artery disease, Crohn’s disease, hypertension, rheumatoid arthritis, type 1 diabetes and type 2 diabetes [17]. The study identified 25 independent association signals at a stringent level of significance ($P < 5 \times 10^{-7}$). Association signals were identified for all diseases except hypertension, where the strongest signal had $P = 7.7 \times 10^{-7}$. This initial analysis of the WTCCC study therefore doubled the number of known complex disease genes. However, the WTCCC was

**Box 2. Lessons learned from past approaches**

Many lessons have been learned from past failing of complex trait genetic studies, the most important of which are as follows.

1. A large sample size is crucial. Polygenic variants have small effects, which to detect reliably requires many thousands of subjects. Past studies used just tens to just a few hundred subjects, which will be powered to find nothing but the strongest polygenic effects.

2. Statistical significance levels need to be interpreted with much caution. A $P$-value of 0.05 in a genetic association study cannot be considered significant; as genetic association studies test hundreds of thousands of variants against many traits a much more stringent threshold is required. $P$-values of $<5 \times 10^{-7}$ and less are required for a SNP to be considered to have strong evidence for association.

3. Replication of findings in essential. The multiple hypothesis testing problem makes replication of GWA results (in suitably powered follow-up studies) essential.

4. Comprehensive coverage of the common variation across the genome is required. By focusing on biological candidates there is less chance of important novel insights into disease pathophysiology. Also, the lack of success of this approach in type 2 diabetes, for example, bears testament to the lack of underlying knowledge of common disease pathophysiology.
primarily a hypothesis generating study, with only the ‘low hanging fruit’ being convincingly identified in this ‘first pass’ analysis: many of the SNPs with \( P \)-values \( >5 \times 10^{-7} \) will also be disease-causing variants. As will be described in the sections below, follow-up studies in sufficiently powered replication cohorts, and the combination of findings from several GWA scans have confirmed many other complex disease variants.

**Type 2 diabetes**

It is perhaps the results of the first type 2 diabetes genome-wide scans [18–22] that best illustrate the power of the GWA approach for identifying novel genes that are important in the aetiology of complex disease. Following up the results from the initial WTCCC GWA scan [17], we worked closely with the DGI [20] and FUSION [19] groups, who had performed similar studies. We used the combined information from these three studies to prioritize variants for follow-up. Including replication samples, these three studies provided data from 14 586 cases and 17 968 controls. The combined analyses identified three entirely novel type 2 diabetes susceptibility genes: **CDKAL1** [cyclin-dependent kinase 5 (CDK5) regulatory subunit-associated protein 1-like 1; OR 1.12, combined \( P = 4.1 \times 10^{-11} \)], **IGF2BP2** (insulin-like growth factor 2-binding protein 2; OR 1.14, combined \( P = 8.6 \times 10^{-16} \)) and **CDKN2A/CDKN2B** (cyclin-dependent kinase inhibitor 2 A/B) gene region (OR 1.20, combined \( P = 7.8 \times 10^{-15} \)) and demonstrated that integrating the results from multiple genome scans can aid the prioritization of signals for replication, and allow confirmation of genes at appropriate levels of statistical confidence not possible with individual GWA studies.

Other type 2 diabetes GWA studies have also been published. The deCODE study [10] of several European and a Chinese population replicates the association of the **CDKAL1** variant (OR 1.20 in Europeans and 1.25 in Chinese). These four studies also confirm the association of variants near **HHEX** (homeobox, hematopoietically expressed) and **SLC30A8** [solute carrier family 30 (zinc transporter), member 8] genes, originally published by Sladek et al. [21]. Importantly, as a positive control, associations for variants in **PPARG** [12], **KCNJ11** [13] and **TCF7L2** [23], originally identified through candidate gene and positional cloning methods, were also seen in the GWA scans, with expected odds ratios.

Of the variants identified through the GWA approach, the two near the **CDK2NA/CDK2NB** gene are particularly interesting. **CDK2NA** encodes P16\(^{\text{INK4a}}\), and is a known tumour-suppressor gene [24]. Mutations of **CDK2NA** cause diverse neoplasias. **CDK2NA** is an inhibitor of cyclin-dependent kinase 4 (CDK4), which is important for beta-cell replication [25]. Overexpression of *Cdkn2a* in mice leads to decreased islet proliferation, whilst *Cdkn2a* knockout mice demonstrate enhanced islet proliferation and survival after beta-cell ablation [26]. Overexpression of *Cdkn2b* causes islet hypoplasia and diabetes in murine models [27]. Together with the **CDKAL1** association, the **CDK2NA/B** finding implicates the cyclin-dependent kinase pathway in the pathophysiology of type 2 diabetes.

Another interesting feature of the **CDK2NA/B** finding is that, as described below, variants of the **CDK2NA/B** gene have also recently been shown to predispose to myocardial infarction (MI). Determining why a gene predisposes to type 2 diabetes and heart disease may lead to an explanation for the link between these two disorders.

The **CDK2NA/B** finding also highlights the power of GWA studies to identify variants outside described genes: whilst one of the signals occurs in the **CDK2NA/B** region, the other (much stronger) association signal occurs \( >200 \) kb from these genes, in a gene desert. This association would not have been picked up by a candidate gene approach. Identifying the mechanism by which this variant (presumably) affects **CDK2NA/B** expression will provide new insights into the regulation of this important gene(s).

The other newly identified type 2 diabetes genes are generally involved in beta-cell development and function, and insulin secretion [18–20]. For example, the
**HHEX** gene is highly expressed in foetal and adult pancreas, and is implicated in pancreatic development [28, 29]. It is a target of the WNT signalling pathway, which has been shown to be critical for the development of the pancreas and islets during embryonic growth [30]. Importantly, TCF7L2 also has an important role in WNT signalling, acting as a nuclear receptor for β-catenin [31]. Together, these findings highlight the importance of the WNT signalling pathway in glucose homeostasis.

**Obesity**

In addition to the newly identified type 2 diabetes genes described above, the WTCCC study found strong association with *FTO* (fat mass and obesity associated) gene region (OR 1.27, *P* = 2.0 × 10\(^{-8}\)) [18]. This finding, which was the strongest susceptibility locus outside TCF7L2, showed strong replication in a further 3757 type 2 diabetes cases and 5346 controls from the UK (OR 1.22, *P* = 5.4 × 10\(^{-7}\)) [18]. However, the lack of such strong association in the DGI study [20], which matched cases and controls for BMI, and the FUSION study [19], where there was minimal BMI differences between cases and controls, suggested that the association with type 2 diabetes was caused by the primary effect on adiposity. Indeed, adjustment for BMI in the UK replication samples abolished the type 2 diabetes association (OR 1.03, *P* = 0.44). This exciting observation lead to the study of association of *FTO* gene variation with BMI and the risk of being overweight and obese in an additional 19,424 adults and 10,172 children, all of white European origin [32]. In the combined data set each additional copy of the rs9939609 risk allele is associated with a BMI increase of approximately 0.4 kg m\(^{-2}\) (*P* = 3 × 10\(^{-35}\)). Individuals homozygous for the A allele (16% of the population) are at a substantially increased risk of being overweight and obese (OR 1.38, *P* = 4 × 10\(^{-11}\)) and obese (OR 1.67, *P* = 1 × 10\(^{-14}\)) compared to those homozygous for the low-risk T allele (37% of the population). This association was observed in children at ages 7–11, but not at birth, and reflects a specific increase in fat mass [32].

*FTO* is a gene of unknown function in an unknown pathway. It seems to be widely expressed in both foetal and adult tissues, with highest levels in the brain [32]. One possibility therefore is that *FTO* is an important regulator of appetite. This would be consistent with the role of monogenic obesity genes, such as *leptin*, but much work is needed to determine whether this is the case. It is clear though that understanding how variants of the *FTO* gene increase fat mass will lead to the identification of a new obesity pathway, with implications for drug development and treatments.

**Age-related macular degeneration**

Age-related macular degeneration (AMD), the main cause of blindness in developed countries, is a chronic, common and complex disease characterized by progressive destruction of retina’s central region and drusen formation behind the retina (reviewed in Ref. [33]). Currently, there is no broadly effective therapy available. The major environmental risk factor for AMD is smoking (smokers have up to 2.5-fold increased risk of AMD than nonsmokers) [34, 35]. One of the first published genome-wide case–control studies was by Klein and colleagues [36]. Using the relatively sparse Affymetrix 100 K chip (Affymetrix Inc.) they identified a common variant in the complement factor H gene (*CFH*) as the SNP most strongly associated with AMD [36]. Although this was a small study (96 cases and 50 controls), it increased its power by using enriched samples (severe AMD cases and older controls to increase the probability of them not developing AMD). A more recent case–control candidate gene study replicated the association of *CFH* gene, and confirmed that individuals homozygous for the most strongly associated risk allele have over sevenfold higher risk for AMD than those homozygous for the nonrisk allele [37]. Human CFH is a regulator of the innate complement system that responds to infection by normally attacking only the diseased cells. Observations of activated complement components within drusen of AMD patients, and of strong effects of smoking and age on CFH plasma levels, suggest that AMD may result from abnormal complement activation in an anomalous inflammatory
response [36]. Although the CFH polymorphisms are noncoding, they may alter the binding of CFH to heparin and C-reactive protein [36]. Furthermore, as CFH is a member of the complement and coagulation cascade pathway, these findings highlight that several different complement and coagulation factors may be potential drug targets and justify further research.

**Crohn’s disease**

Crohn’s disease, most commonly affecting ileum and colon, is a common form of idiopathic inflammatory bowel disease (IBD) where genetic predisposition has been supported by twin studies showing concordance rate of 50% in monozygotic compared to 10% in dizygotic pairs. Previously, years of research effort involving linkage, candidate gene and targeted association studies, identified only two genuinely associated variants, in CARD15 gene and the IBD5 haplotype. A recent GWA study by Rioux and colleagues [39] identified and replicated several new susceptibility loci for ileal Crohn’s disease. The most associated SNP, independently identified by a smaller German study [38], was a nonsynonymous amino acid change in ATG16 autophagy-related 16-like 1 (ATG16L1) gene. The risk allele is a major allele (it has a frequency of about 52% and 60% in controls and cases, respectively), and individuals carrying one copy are at a 35–45% higher risk of developing the disease than those carrying no ATG16L1 risk alleles [38, 39]. This SNP is in strong LD ($r^2 = 0.97$) with the strongest signal in the WTCCC scan for Crohn’s disease. Autophagy is a constitutive biological process involved in immune pathogen recognition, and the variants in ATG16L1 gene may alter innate immune control or antigen presentation in the adaptive immune pathways [39]. The WTCCC study identified four novel association signals, all of which have since been replicated. These map to IRGM (immunity-related guanosine triphosphatase), MST1 (macrophage stimulating 1), NKX2-3 (NK2 transcription factor related, locus 3) and PTPN2 (protein tyrosine phosphatase, nonreceptor type 2) gene regions. These novel findings highlight that defects in a number of components of innate and adaptive immune pathways, such as those in autophagy and the processing of phagocytosed bacteria, are a major cause of Crohn’s disease.

**Prostate cancer**

Until recently, the only risk factors established for prostate cancer, the most prevalent noncutaneous malignancy in males in developed countries, were family history and African-American ethnic background [40]. Two genome-wide case–control association studies have now both confirmed a known associated variant and separately identified new independent risk variants in the linked 8q24 region [41, 66]. Yeager et al. [41] estimate that their new and a previously reported variant contribute to over threefold risk of disease in double homozygotes for the risk alleles compared to double homozygotes for the protective alleles, giving the combined population attributable risk (PAR) of 27%, whilst Gudmundsson et al. [66] estimate a combined PAR of 13% in European populations for their two variants. As none of the independent 8q24 signals lie in known genes, it is possible, that there are several unknown prostate cancer susceptibility genes in the region. Alternatively, the risk variants may independently affect the regulation of genes outside the linked region, such as the near-by proto-oncogene MYC, by making the whole region prone to somatic amplification, a common event in prostate tumours [42].

**Myocardial infarction**

Coronary heart disease (CHD), including MI, has reached endemic proportions worldwide and is the major cause of mortality in Western countries [43]. Two GWA studies seem to have discovered the major genetic risk factor [44, 45]. Both studies identified a strong signal on chromosome 9p21, close to CDKN2A and CDKN2B genes (MI association $P = 1.2 \times 10^{-20}$). Homozygotes for the risk allele (20–25% of Caucasians) are 64% more likely to suffer heart attack [44] and have up to 40% increase in CHD [45] than homozygotes for the nonrisk allele. In the study by Helgadottir et al. [44], each copy of the associated variant reduced the age at onset of MI by approximately 1 year ($P = 2.9 \times 10^{-7}$), and had PAR
of 21% (31% for early onset cases). Resequencing of the associated 58-kb interval identified a copy number variation in a putative noncoding RNA of unknown function, suggesting that specific variation in the transcript expression or function may predispose to heart disease [45].

Although the relative risks explain a small proportion of the familial clustering of the disease, this susceptibility locus is the same one found to associate with type 2 diabetes, as described above. It is possible that the risk allele is located within a regulatory element that controls the expression of a gene outside of the associated region, or that the functional variant itself is located outside the region that was not well covered by the genotyping platforms used in the studies. Therefore, fine mapping and further studies of this region are needed to find the causative variant(s) for both cardiovascular disease and type 2 diabetes, providing potentially the same drug target for two common diseases.

What will these findings mean for clinical practice

Identifying the genes and pathways involved in predisposing individuals to complex disease is giving us insights into the pathophysiology of these diseases, which may eventually lead to the development of novel treatments. These insights are the most valuable thing to come out of these genetic studies, and will, in time, no doubt lead to better clinical management and treatment of these diseases. But is there any immediate clinical utility of knowing an individual’s genotypes at these disease variants?

Pharmacogenetics

Pharmacogenetics is a study of how genetically determined variation affects an individual’s response to drugs. It is well known that adverse side effects and therapeutic failure of drugs may both have a strong genetic component [46]. If a patient’s genotypes in the relevant genes are known, it may be possible to truly personalize medication and optimize treatment by selecting the most effective drug and its dose. It has been estimated that adverse drug reactions account for 6.5% of hospital admissions, 4% of hospital bed capacity and 0.15% fatalities in England, at a projected annual cost to the NHS of £466 million [47]. Although a large portion of this figure can be attributed to prescription errors and accidental overdoses, pharmacogenetics could be used to identify individuals both at a highest and lowest risk of developing adverse effects to particular drugs or doses. Furthermore, clinical efficacy could be much improved if drugs are prescribed only to individuals likely to benefit from them, thus reducing the number of ineffective treatments. Pharmacogenetic testing will be particularly desirable in cases where it proves to be more effective than the current practice of just careful monitoring of a patient’s response to a drug/dose.

There are many examples of where this has already happened, such as for monogenic types of diabetes. The definition of six different genetic subtypes in MODY has led to recognition of different clinical phenotypes [48]. Molecular genetics is now commonly used as a diagnostic test, and the diagnosis has a dramatic effect on the treatment decisions in MODY. Patients with glucokinase mutations have life-long mild but stable fasting hyperglycaemia from birth, and require no treatment because they essentially have a glucose sensing defect where glucose is regulated at a higher level [48]. In contrast, patients with HNF1A mutation have a progressive hyperglycaemia but are very sensitive to hypoglycaemic effects of sulphonylureas [49]. The correct diagnosis is very important in this case because many of these patients are misdiagnosed as having type 2 diabetes, in which case the most common pharmacological treatment is metformin, but HNF1A patients have a fourfold greater response to sulphonylurea gliclazide than to metformin [50].

The ability to analyse thousands of SNPs in a large number of individuals will be essential for defining genetic heterogeneity of diabetes, which could then be translated into clinical heterogeneity between the patients. Even if a gene does not cause or predispose to a disease, it can still interfere with the drug-targeted metabolic pathway and modify the toxicity
(effectiveness and side effects) of the drug and the clinical response. For example, the putative role of TCF7L2 in beta-cell function has lead to a hypothesis that TCF7L2 variation may have an effect on glycemic response to sulphonylureas, but not to metformin [51]. A study that tested this using 747 and 864 metformin users with type 2 diabetes found that, whilst the tested variant had no effect on metformin response, in the sulphonylurea treatment group only 40% of patients homozygous for the risk allele reached the target HbA1c <7%, compared to 61% of patients with the other two genotypes (OR 0.46, \( P < 0.001 \)) [51]. This study provides a strong example of pharmacogenetics in type 2 diabetes, where a common variant predicts treatment response. If a patient’s genotypes in several of the relevant genes are known, it may be possible to personalize medication and optimize treatment.

**Prediction**

In type 2 diabetes and other complex polygenic diseases the number of confirmed common risk variants is small. This means that, for now, molecular genetics cannot be used as a diagnostic test because family history and environmental and lifestyle influences, such as BMI, smoking, physical activity and social class, have much higher predictive power. One way of assessing predictive power of polygenic variant information is to use the area under the ROC curve (\( AUC \)), which measures the discriminatory power of the test [52]. The test is 100% accurate when the \( AUC \) is 1, and is no better than chance when \( AUC \) is 0.5. It has been estimated that for an \( AUC \) of 0.8 it will be necessary to genotype \( \sim50 \) risk variants with allele frequencies of 10% and ORs of 1.5 [67]. In multifactorial diseases where preventative measures exist, coupled with the low cost of genotyping, it may be practically and economically justifiable to identify individuals with the highest risk of developing the disease if the preventative measures are still effective in such high-risk groups. This should certainly be the case in type 2 diabetes where lifestyle factors have a big impact on onset and severity of the disease.

A recent type 2 diabetes case–control study demonstrated that, although individual susceptibility variants only moderately increase the risk of type 2 diabetes and are of limited use in disease prediction, by combining the information from the three replicated risk variants (known at the time) it is possible to identify individuals at significantly greater risk of developing the disease than when a single polymorphism is used [53]. Individuals with all six risk alleles had an OR of 5.71 (95% CI: 1.15–28.3) compared to those with no risk alleles. The area under the ROC curve for the three polymorphisms was 0.58, which should be much improved as the number of reproducibly associated variant increases, and if evidence of gene–gene and gene–environment interaction is found for any of these variants.

Similarly, the predictive value of specific genotype combinations was estimated for three known AMD risk variants, in the \( CFH \), \( LOC387715 \) and \( C2-CFB \) genes: 1% of the population homozygous for high-risk alleles at all three loci has a 250-fold higher risk of AMD than 2% of the population carrying no risk alleles, translating into >50% lifetime risk of developing the disease [37]. Clearly, the advantage of identifying individuals at such high risk warrants further investigation into how these genes might interact with the known lifestyle and environmental factors, whilst the lack of interaction between the three known AMD genetic factors [37] suggests that, as expected, several different drugs may be needed to successfully treat polygenic diseases.

**Conclusions**

With the relative failing of linkage and candidate gene association studies, there had been much scepticism about the ability of geneticists to identify genes for common, complex disorders. Vast amounts of resources were poured into these studies, with very limited success. Recent technological advances, and a better understanding of the genetic architecture of complex diseases, has led to the development of the GWA study, which is a powerful new approach to identifying genes influencing predisposition to...
common, complex disease. Genome-wide studies are ‘hypothesis-free’ and allow the identification of previously unsuspected genes and pathways in disease aetiology. The initial success of many GWA studies has certainly brought much excitement to the genetics community, and we finally have the ability to systematically identify the genetic variants underlying a range of common disorders from type 2 diabetes to Crohn’s disease, but when will these findings benefit clinical medicine, and are we on the right track to using this genetic information to improve clinical practice?

It should be remembered that this type of genetic research is still in its infancy, and as we have only just acquired the ability, through GWA studies, to identify a large number of common disease genes, it is unsurprising that there has been no major impact of this research on clinical practice as yet. However, the ability of genetic research to alter treatment and diagnosis, given sufficient time, is clearly demonstrated by the success in the monogenic field, for example, in MODY. We expect there will be a great benefit to clinical medicine in the years to come, and whilst prediction of disease and pharmacogenetics may eventually prove valuable, the greatest clinical benefit of GWA studies is likely to come from etiological insights into disease processes.

Conflict of interest statement
No conflict of interest was declared.

References
13 Glyn AL, Weedon MN, Owen KR et al. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. Diabetes 2003; 52: 568–72.
17 The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661.
31 Smith U. TCF7L2 and type 2 diabetes – we WNT to know
30 Papadopoulou S, Edlund H. Attenuated Wnt signaling perturbs
29 Bogue CW, Ganea GR, Sturm E, Ianucci R, Jacobs HC. Hex
27 Moritani M, Yamasaki S, Kagami M et al.
26 Krishnamurthy J, Ramsey MR, Ligon KL et al.
24 Wolfel T, Hauer M, Schneider J
23 Grant SFA, Thorleifsson G, Reynisdottir I et al. Variant of tran-
21 Krishnamurthy J, Ramsey MR, Ligon KL et al.
8 Seddon JM, Cote J, Page WF, Aggen SH, Neale MC. The US Twin Study of Age-Related Macular Degeneration: Relative Roles of Genetic and Environmental Influences. Arch Ophthal-


Correspondence: Michael Weedon, Peninsula Medical School, St Lukes Campus, Magdalen Road, Exeter EX1 2LU, UK.

(fax: +44 (0) 1392 262926; e-mail: michael.weedon@pms.ac.uk).