Title

Genome-wide association studies in pharmacogenomics

Journal

Pharmacogenetics and Genomics, 23(8)

ISSN

1744-6872

Authors

Motsinger-Reif, Alison A

Jorgenson, Eric

Relling, Mary V

et al.

Publication Date

2013-08-01

DOI

10.1097/fpc.0b013e32833d7b45

Peer reviewed

Genome-wide association studies in pharmacogenomics: successes and lessons

Alison A. Motsinger-Reif[a], Eric Jorgenson[b,c], Mary V. Relling[d], Deanna L. Kroetz[c], Richard Weinshilboum[f], Nancy J. Cox[g] and Dan M. Roden[e]

Objective **As genotyping technology has progressed, genome-wide association studies (GWAS) have matured into efficient and effective tools for mapping genes underlying human phenotypes.**

Methods **Recent studies have shown the utility of the GWAS approach for examining pharmacogenomic traits, including drug metabolism, efficacy, and toxicity.**

Results **Application of GWAS to pharmacogenomic outcomes presents unique challenges and opportunities.**

Conclusion **In the current review, we discuss the potential promises and potential caveats of this approach specifically as it relates to pharmacogenomic studies. Concerns with study design, power and sample size, and analysis are reviewed. We further examine the features of successful pharmacogenomic GWAS, and describe consortia efforts that are likely to expand the reach of pharmacogenomic GWAS in the future.**

Pharmacogenetics and Genomics **23:383–394 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

Pharmacogenetics and Genomics 2013, **23**:383–394

Keywords: drug metabolism, drug response, genome-wide association, genome-wide association studies, pharmacogenetic, pharmacogenomic, toxicity

[a]Department of Statistics, Bioinformatics Research Center, North Carolina State University, Raleigh, North Carolina, [b]Department of Neurology, Ernest Gallo Clinic and Research Center, [c]Department of Bioengineering and Therapeutic Sciences, School of Pharmacy and Medicine, University of California, San Francisco, California, [d]Department of Pharmaceutical, St Jude Children's Research Hospital, Memphis, [e]Office of Personalized Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, [f]Department of Molecular Pharmacology and Experimental Therapeutics, Division of Clinical Pharmacology, Mayo Clinic, Rochester, Minnesota and [g]Department of Medicine, Department of Human Genetics, The University of Chicago, Chicago, Illinois, USA

Correspondence to Alison A. Motsinger-Reif, PhD, Department of Statistics, Bioinformatics Research Center, 1 Lampe Drive, CB 7566, North Carolina State University, Raleigh, NC 27695-7566, USA
Tel: +1 919 515 3574; fax: +1 919 515 7315;
e-mail: motsinger@stat.ncsu.edu

Received 5 March 2010 **Accepted** 23 June 2010

Introduction

Since 2005, genome-wide association studies (GWAS) have matured into a powerful tool to identify single nucleotide polymorphisms (SNPs) that can be reproducibly associated with a variety of human phenotypes. Currently, over 300 papers have reported significant associations of common variants with a range of phenotypes and diseases [1]. These successes have provided numerous insights into the relationship among genetic variants, biological pathways, and human traits, and shown how proper study design and analysis can lead to the success of GWAS. A key lesson from this first generation of GWAS is that no single approach will be appropriate for all phenotypes [2].

The genetics of drug–response outcomes, broadly referred to here as pharmacogenetic/pharmacogenomic outcomes, are a particular category of phenotypes that present unique challenges and opportunities in gene discovery [3]. In this study, we discuss the advantages and limitations of GWAS as applied to pharmacogenomic outcomes. Some of these challenges are variations on general concerns for disease gene identification, whereas others are unique to pharmacogenomic outcomes.

Like studies of disease phenotypes, the success of any pharmacogenomic GWAS will depend on the effect size and allele frequency of genetic variants that influence the trait, the sample size available to detect those variants, the population under study (treatment protocol, dosage, patient features including self-reported race/ethnicity, etc.), and study design (observational study or randomized controlled trial). Unlike most disease phenotypes, pharmacogenomic outcomes often have clear, clinically defined phenotypes and well-understood mechanisms that may underlie variation in drug response, including known systems of transport and metabolism, and sites of drug action. In addition, larger genetic effects may exist for pharmacogenomic traits than for disease phenotypes, providing greater statistical power for genetic association studies.

An important potential limitation for pharmacogenomic GWAS is the sample size. GWAS for traits like height or QT or complex diseases like diabetes need and benefit from large numbers, and currently mega–meta-analyses are identifying and validating associated loci. Such large sample sets are generally not possible for pharmacogenomic outcomes as they usually include by definition both a disease (often with low prevalence) and a well-curated drug response phenotype (which further reduces the available study population).

In this study, we discuss key issues for GWAS, including the strengths and limitations of this approach. We then elucidate issues of heightened importance in GWAS of pharmacogenomic traits. We discuss appropriate study designs and analysis strategies, and describe lessons from

1744-6872 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/FPC.0b013e32833d7b45

successful pharmacogenomic GWAS. We end with a discussion of ongoing efforts to develop consortia for the purpose of obtaining large sample sizes for drug response outcomes.

Promises

There are clear, well-understood advantages to a genome-wide association approach to phenotype association discovery. GWAS are conventionally intended as an unbiased scan of the genome, interrogating the majority of common genetic variation for disease association. In contrast to a candidate gene approach, whether narrow or broad in scope, GWAS allow the identification of totally novel susceptibility factors that promise to provide us with better biological understanding of phenotypes [4]. There are many candidate mechanisms that drive variability in drug responses: metabolism, transport, targets, target partners, immunologic pathways (e.g. for allergic reactions), etc. that have directed many successful candidate gene studies [5]. However, they cannot identify genes outside the current knowledge of those mechanisms. GWAS allow such novel discovery.

GWAS have distinct advantages as compared with more traditional linkage-based approaches [6]. There are three key general advantages of GWAS approaches for gene identification, each of which are exaggerated for pharmacogenomic outcomes:

(1) Case–control cohorts are generally less expensive and easier to collect than extended pedigrees or nuclear families. This is especially true in drug–response studies where it is rare for multiple family members to have well-characterized responses to drug challenges; that is, formal linkage analysis has not been feasible for drug response phenotypes. GWAS do not require the ascertainment of pharmacogenomic interventions in related individuals.
(2) Association studies have higher statistical power to detect small-to-modest genetic effects as compared with linkage studies [6]. For pharmacogenomic studies, especially for rare toxicities where sample sizes are limited, this advantage in power may be the difference between success and failure in gene mapping.
(3) As linkage disequilibrium (LD) typically stretches over tens of kilobases as opposed to several megabases [6], association signals are more finely localized than linkage signals, which should lead to more rapid identification of causal variants by rapidly narrowing down the regions for follow-up in functional studies – critical for novel mechanistic insights – and, thus, to more rapid translation of findings.

There are additional advantages to GWAS that are more specific to pharmacogenomic outcomes. First, GWAS provide context for understanding the relative importance of genetic contributors to pharmacogenomic traits that may otherwise be unavailable. The genetic component of human phenotypes can be assessed by estimating heritability (the proportion of variation in a trait because of genetic factors) through methods such as variance components analysis, segregation analysis, etc. Each of these methods requires family data, which, as noted above, is usually difficult to collect for pharmacogenomic outcomes [7].

Another specific application of GWAS in pharmacogenomics is the ability to rule out – with prespecified confidence intervals – contributions by unidentified genes to a drug response phenotype. As pharmacogenomic GWAS can directly investigate the role of genetic variation on clinical outcomes, the findings from pharmacogenomic GWAS can be more rapidly translated to clinical practice. As translation to the bedside is one of the goals of pharmacogenomic gene mapping [8], it is important to ensure that any unanticipated important genetic contribution to variability in a drug response is not missed [9]. Of equal importance is the identification of novel mechanisms, both for drug response and/or adverse drug reactions. So, having identified variants in gene X or Y as contributors to a variable drug response, it is key to ensure that there is no other important genetic contributor before mounting a trial. Understanding the influence of genetic variants in drug response can limit unanticipated variability in a drug treatment [9]. The role of GWAS in this process is evident in the evaluation of the genetic component of warfarin dosing [9]. The strong association of variants in *VKORC1* and *CYP2C9* for stable warfarin dosing was well established [10–12], but before the National Heart, Lung and Blood Institute in the US would mount a large clinical trial it was important to determine whether there were other genetic variants that also had large effects on stable warfarin dosing. GWAS [13,14] have now ruled out large contributions by other loci, thereby allowing clinical trials to proceed [15]. Similarly, a GWAS for clopidogrel effect on ADP-induced platelet aggregation identified only one associated locus, at *CYP2C9/19*, laying the groundwork for design of clinical trials [16]. As genotyping platforms with increased SNP density become available, the coverage of genetic variation in the human genome will become more complete, providing greater confidence that clinically important genetic effects on pharmacogenomic traits will not be missed. Thus, while many variants in drug metabolism genes have been shown to confer large clinical effects that have often been identified without GWAS (e.g. by well informed candidate gene studies), even GWAS with 'negative' results add this crucial additional information [17].

Considerations

Common disease common variant hypothesis

Despite the advantage of GWAS studies discussed above, there are important caveats that must be remembered in their design and application. Although many of these caveats are true of GWAS in general, the impact of these concerns may be different in pharmacogenomic studies than in general trait mapping.

A key assumption in GWAS is what is known as the common disease/common variant hypothesis [18]. The common disease/common variant hypothesis proposes that most of the genetic risk for common, complex diseases is attributable to relatively common [minor allele frequency (MAF) > 0.05] polymorphisms [18]. The alternative to the common disease/common variant hypothesis is that multiple rare variants cause disease at high prevalence in the population through a variety of mechanisms. Such variants can represent genetic heterogeneity of variants in a single gene, or multiple rare variants within genes in the same pathway that have cumulative effects. These two hypotheses have important implications – common variants are thought to impart subtle effects on gene function, often through changes to gene regulation. Rare variants may have larger effects on gene function, such as nonsynonymous variants that alter the amino acid sequence of the resulting protein, and as a result lead to large changes in disease risk or trait values. As a result, it is likely that both common and rare variants will contribute to common phenotypes, but the relative proportions will influence the appropriate methods for discovering associated variants. The GWAS approach is well powered to detect common variants with modest effects. GWAS is less effective in testing rare variation, a problem that is confounded by the DNA microarrays used in these studies, which have been designed to capture common variation. Even 'next generation' GWAS that will reliably interrogate (directly or indirectly) all variation with MAF greater than 0.005 may be insufficient to identify enough of the contributory variation to allow us to understand biology whether most of that variation has MAF less than 0.005, as the sample sizes required to achieve sufficient statistical power for such effects may be prohibitive. As 'next generation' sequencing becomes more accessible, and whole-genome sequencing becomes more affordable, more rare variant analysis will be possible in pharmacogenomics.

Sources of bias

An important concern in GWAS studies for pharmacogenomics is of the potential for bias in the selection of genetic variants [2]. Although large number of variants with low MAF are included in the densest GWAS platforms, GWAS have little power, given sample sizes available, to detect significant associations with low MAF SNPs. In addition, it is widely recognized that genotype quality is not as high for rare variants as it is for more common variants. Consequently, a common approach is to not assess the significance of associations with rare variants (MAF < 0.01). This further compounds the limited statistical power to detect associations with less common genetic variants. Moreover, SNPs included on high-throughput platforms must pass stringent tests for ease of genotyping, which leads regions with gene duplications (and pseudogenes) to be poorly represented on high-throughput genotyping products, and many of these – such as CYPs or the HLA locus – are precisely the genes of greatest interest for pharmacogenomic studies. The human cytochrome P-450 family of genes that encode enzymes active in xenobiotic metabolism have been associated with a large number of pharmacogenomic outcomes [19]. They are known to be highly polymorphic, with a wide range of allele frequencies across populations, and contain complex structural variation with unique haplotypic structure and copy number variations [20]. The coverage of these types of variation is limited on current GWAS genotyping platforms [21].

Study design

Experimental design is a crucial component of any successful GWAS, and pharmacogenomic studies have different advantages and limitations than traditional disease studies. The importance of proper definition and collection of phenotype data has become increasingly appreciated in the context of GWAS [17]. An important advantage in pharmacogenomic studies is that multiple response phenotypes are often collected within the same study, such as efficacy and adverse events, allowing a broader dissection of trait genetics in a single study.

However, because all pharmacogenomic outcomes are responses to the environmental exposure of the drug and because these drugs are given in response to a disease condition, there may be complex interactions between disease and drug response relevant in phenotype definition. Precise definitions are essential for both the disease and drug response phenotypes, which are often discrete diagnoses from these complex relationships. For example, in some, but not all cases, rare adverse drug reactions may represent a 'tail' of response distributions and where to define that cut-off within the distribution can be a challenge. The SEARCH Collaborative Group showed a successful approach to address this issue by combining patients with both definite and incipient statin-induced myopathy into a single case definition [22]. In other cases, a rare adverse reaction is an unexpected outcome often unrelated to the desired mechanism of action [17].

One efficient use of resources to collect pharmacogenomic phenotypes is to collect samples within the context of clinical trials, which streamlines the collection procedures. The use of clinical trial data for GWAS is not only an efficient use of resources, but has the advantage that similarly treated 'controls' for the phenotype of interest are built into the trials. However, because some trials are not designed for GWAS mapping, the study designs used for collection may not be ideal for pharmacogenomic analysis (e.g. multiple drugs used in treatment arms, etc.) [23]. Obviously, this 'challenge' is inherent to the treatment of diseases like cancer or end-stage congestive heart failure in which it would be unethical to fail to treat patients with the current standard of care for this illness. If pharmacogenomic efforts are substudies of clinical trials, sample sizes may decrease, which reduces the power

of the pharmacogenomic component. As meeting recruitment targets is a primary goal in most clinical trials, genomic and pharmacogenomic efforts are often included only as substudies to which patients may or may not consent; as a result, the power and generalizability of genomic studies is compromised. Genetic studies added as an afterthought may be viewed as creating a barrier to recruitment and thus may not be a priority for sponsors. Collecting drug response phenotypes in health care systems with electronic medical records is another method of accruing patients that is now being explored.

Sample size limitations are a challenge in any GWAS, but are amplified in many pharmacogenomic studies. Particularly when studying rare drug reactions or adverse events, it is by definition not feasible to recruit thousands of patients with rare outcomes. This is a particular limitation in pharmacogenomic GWAS, as the replication of association results in independent populations has become the 'gold standard' for validation of results [24]. If the collection of a reasonable sample size for a discovery cohort is at the edge of practicality, this makes the collection of a well-powered replication cohort often impossible. Consortia efforts (discussed below) have been motivated by this limitation, to combine samples from across the world to increase power and potentially identify replication cohorts to maximize power and provide validation to associated signals. However, even the establishment of networks of investigators cannot necessarily overcome these limitations, and the field must look for creative approaches of validation/replication possibly involving functional studies or examination of related intermediate phenotypes.

There are unique 'challenges' associated with validation/replication for pharmacogenomics. Clinical trials are expensive, and every study is unique as they are designed to represent an advance over earlier studies to answer novel therapeutic questions. Therefore, in pharmacogenomics greater emphasis may have to be placed on functional validation of GWAS 'signals' and on biological plausibility. In addition, one must recognize that the larger the sample size, the more likely that features, which confound the genotype/phenotype relationship will be undocumented or uncontrolled; thus diluting the 'purity' of the phenotype and potentially reducing the power [25].

Besides sample size, there are other practical limitations in study design for pharmacogenomic studies. As mentioned earlier, family-based designs are generally impractical with drug response outcomes, which mean the field relies heavily on cohort or case–control studies for GWAS [5]. Although the number of cases may be limited by event frequency as discussed above, finding and selecting appropriate controls presents additional challenges. Although GWAS of common diseases have taken advantage of the use of shared controls across studies, this is not often possible in pharmacogenomic studies, as typically controls must also be exposed to the drug of interest (though this may not be necessary in all cases). Other matching criteria must also be considered, such as disease interactions, population admixture, and additional environmental and clinical exposures.

Analysis

As GWAS have become more prevalent, methodologies for the analysis and interpretation of results have coevolved. Many tools have been developed and evaluated in the context of GWAS, and have resulted in the many successes seen to date. However, there are still many challenges in the analysis strategies used for GWAS in general, and particular challenges for pharmacogenomics, as discussed below.

Standard analytical approaches

The majority of earlier GWAS have relied on the use of traditional statistical methodologies for analysis, and several tools have become widely used in the field. Software packages such as PLINK [26] have become very popular in implementing logistic regression (for case–control or cohort studies), linear regression (for quantitative traits), and family-based association tests for GWAS studies.

After various types of corrections for multiple testing (Bonferroni, permutation approaches, etc), results of these analyses are typically prioritized with replication strategies. For single samples, the union of significant results from several analytical approaches (committee-based approaches) or measures of reliability from internal model validation is often used to prioritize robust signals. When more than one sample is available, multistage replication strategies are often employed to discover, prioritize, and validate signals. Finally, when multiple samples are available, meta-analysis is often used to obtain more comprehensive measure of association signals [27]. Challenges in sample collection (discussed above) can limit the use of such multistage replication and meta-analysis strategies in pharmacogenomics. One alternative approach for replication, or at least prioritization, of association signals in pharmacogenetic studies is to use nonclinical GWAS of large collections of human tissue, cell lines, and genetic model organisms [28].

Detecting complex predictive models

Such traditional approaches have been very powerful for identifying strong single-locus associations (low-hanging fruit) for a wide range of phenotypes in both common diseases and pharmacogenomic outcomes (reviewed below), and are typically applied in a way that fits within the 'unbiased' intentions of GWAS. Despite the successes of these approaches, their limitations for detecting and prioritizing more complex models have become a hot topic in the literature [29].

As many successful GWAS have been published, the sum of the genetic contributions of associated variants in many common traits is far below the estimated heritability of the

traits. These gaps in explained heritability are potentially clarified by several potential etiologies. Low power to detect low-effect sizes, the presence of rare variants contributing to phenotypes, unmeasured nucleotide or structural variation, complex methylation/epigenetic mechanisms, and gene–gene/gene–environment interactions are all hypothesized to contribute to the unexplained trait variation [29]. In response to these limitations, new analytical approaches are evolving to detect complex genetic risk models discussed below. These limitations are leading to refinement of methods for GWAS analysis, and these may be especially appropriate for pharmacogenomic studies.

Expert knowledge driven analysis

Although this 'unbiased' intent of GWAS is to detect potential new genetic associations that might not have been considered as candidate genes, there has been a recent appreciation for the fact that these simple analytical approaches ignore the large amount of expert knowledge available for a particular outcome. In response, there has recently been rapid development in the use of network and pathway analysis for analysis of GWAS data [30–33]. Literature searches (automated or hand curated), databases of earlier results, etc. are being exploited to improve the power of GWAS. As it is much known about the mechanism and metabolism of many of the drugs evaluated in pharmacogenomic studies, there is very well-directed guidance for such knowledge-driven analysis. The Pharmacogenomics Knowledge Base [34] is an important resource and data repository that summarizes and curates drug response/gene relationships through gene variant annotation, hand-curated literature review, and important pharmacogenomic genes and pathways. An example of the potential of pathway-based analysis is discussed below.

Successes in pharmacogenomics

Arguably the most important demonstrations of the utility and challenges of GWAS in pharmacogenomics are the empirical results of successful studies. A brief description of the outcomes evaluated in pharmacogenomic GWAS and the strongest signals identified is listed in Table 1. Details of each study can be found in the references provided.

The potential and drawbacks of an agnostic, unbiased approach for genetic association studies in pharmacogenetics are illustrated by a GWAS of the activity of a well-known polymorphic drug metabolizing enzyme, thiopurine methyltransferase (TPMT) in lymphoblastoid cell lines from the HapMap project [63]. The goal of the experiment was to assess whether the TPMT polymorphism could be 'rediscovered' in this fashion [62]. Although common polymorphisms in TPMT were well tagged, and TPMT polymorphisms were associated with TPMT activity, the GWAS indicated that 96 genes were ranked higher than was TPMT itself. The extent to which these higher ranked genes are false versus true positives is not yet clear, but indicate the difficulty of using GWAS approaches even for putatively monogenic traits.

An example of a GWAS for drug pharmacokinetics is provided by an analysis of methotrexate clearance determined in over 3000 courses of the drug given to 434 children with leukemia [36]. Many candidate gene studies have earlier been conducted to identify genetic variation associated with methotrexate pharmacokinetic variability with limited success. Using GWAS, the *SLCO1B1* gene was represented by multiple polymorphisms in several LD blocks, a finding that was replicated in an independent cohort of patients, suggesting that there are multiple mechanisms by which alteration of OATP1B1 (encoded by *SLCO1B1*) could affect methotrexate pharmacokinetics. Although methotrexate had been shown to be an OATP1B1 substrate, it was a rather weak one [64,65], and so the gene had not risen to the top of candidate gene lists. This finding has implications for both toxicity to methotrexate and to possible drug interactions with widely used OATP1B1 substrates, such as statins.

The utility of pathway-based analysis is shown by Hartford *et al.* [60], who performed a GWAS examining etoposide-induced leukemia with myeloid/lymphoid or mixed-lineage leukemia. They prioritized variant associations based on expression results, to identify alterations in three biological pathways: adhesion, Wnt signaling, and regulation of actin. Results in an independent validation cohort confirmed the alterations in the adhesion pathway. None of the alterations identified were significantly based on traditional association analysis, showing the potential of more complex modeling to identify pathway-level associations.

Although most of the published studies identified variants at a genome-wide significance level, many of them found strong potential signals that did not stand up to traditional analyses [45,46,66]. These negative results may represent true negative results, but it is highly likely that many of these studies were limited by many of the challenges discussed above (power, coverage, etc).

Network efforts

To address many of the limitations discussed above, particularly in regards to limited sample sizes and lack of traditional replication cohorts, researchers are successfully combining resources and establishing worldwide collaborations to support large-scale GWAS. Given the complexities of drug response phenotypes, this approach seems especially appealing in the application of GWAS to pharmacogenomics. By combining cohorts from around the globe, pharmacogenomic studies will have higher power to detect and validate response-determining variants.

The SEARCH Collaborative Group [22] shows the success of such a collaboration. The SEARCH Collaborative Group examined a rare outcome of statin therapy – myopathy,

Table 1 Details of pharmacogenomic GWAS

Study	PMID	Trait	Initial sample size	Replication sample size	Region	Reported gene(s)	Strongest SNP-risk allele	*P* value	Platform (SNPs passing QC)
Thompson *et al.* [35]	20031582	Response to statin therapy	1984 individuals	5009 individuals	NS	NS	NS	NS	Perlegen (291 998)
Trevino *et al.* [36]	19901119	Methotrexate plasma pharmacokinetics and toxicity	434 children	206 children	12p12.2	*SLCO1B1*	rs11045879-?	1.7×10^{-10}	? (500 568)
Aberg *et al.* [37]	19875103	Response to antipsychotic therapy (extrapyramidal side effects)	738 schizophrenic individuals	NR	2p12	Intergenic	rs17022444-?	1×10^{-10} (SAS)	Affymetrix (492 000)
					4q24	Intergenic	rs7669317-?	8×10^{-8} (AIMS)	
					11q24.1	*ZNF202*	rs2126709-?	4×10^{-7} (SAS)	
					14q11.2	Intergenic	rs12147450-?	6×10^{-7} (BARS)	
					9p21.3	Intergenic	rs10811771-?	8×10^{-7} (SAS)	
					8p23.1	Intergenic	rs2251301-?	1×10^{-6} (BARS)	
					14q32.2	Intergenic	rs1459148-?	2×10^{-6} (SAS)	
					16p13.2	*A2BP1*	rs9302841-?	2×10^{-6} (AIMS)	
					2q37.3	Intergenic	rs6743931-?	2×10^{-6} (AIMS)	
					9q33.1	Intergenic	rs876347-?	2×10^{-6} (SAS)	
					1q41	Intergenic	rs337161-?	3×10^{-6} (SAS)	
					1q41	*MOSC2*	rs1494373-?	6×10^{-6} (SAS)	
					4q22.1	*KIAA0914*	rs16996151-?	6×10^{-6} (SAS)	
					9q33.2	Intergenic	rs4837752-?	6×10^{-6} (SAS)	
					1p32.1	*FGGY*	rs17119280-?	7×10^{-6} (SAS)	
					20q13.32	*ZNF831*	rs12625057-?	7×10^{-6} (SAS)	
					11p13	*TRIM44*	rs7928794-?	8×10^{-6} (SAS)	
Garriock *et al.* [38]	19846067	Response to citalopram treatment	Up to 883 responders, 608 nonresponders	NR	7q36.3	*UBE3C*	rs6966038-?	4×10^{-7} (remission)	Affymetrix (430 198)
					7q36.3	*UBE3C*	rs6966038-?	5×10^{-7} (response)	
					20q13.31	*BMP7*	rs6127921-?	1×10^{-6} (remission)	
					20q13.31	*BMP7*	rs6127921-?	3×10^{-6} (response)	
					21q21.3	*EIF4A1P*	rs2830840-?	5×10^{-6} (remission)	
					15q22.2	*RORA*	rs809736-?	8×10^{-6} (response)	
					18q12.1	*NOL4*	rs7239368-?	9×10^{-6} (remission)	
Suppiah *et al.* [39]	19749758	Response to hepatitis C treatment	131 European ancestry responders, 162 European ancestry nonresponders	261 European responders, 294 European non-responders	19q13.2	*IL28A*	rs8099917-G	9×10^{-9}	Illumina (311 159)
					19q13.2	*IL28B*			
Tanaka *et al.* [40]	19749757	Response to hepatitis C treatment	72 Japanese responders, 82 Japanese nonresponders	122 Japanese responders, 50 Japanese non-responders	9q13.2	*IL28B*	rs8099917-G	3×10^{-32}	Affymetrix (621 220)
Ising *et al.* [41]	19736353	Response to antidepressant treatment	339 German individuals	1193 German individuals	NS	NS	NS	NS	Illumina (389 251) (pooled)
Laje *et al.* [42]	19724244	Response to antidepressant treatment	90 white cases, 90 white controls	30 white cases, 1652 white controls	NS	NS	NS	NS	Illumina (100 864)

McClay *et al.* [43]	19721433	Response to antipsychotic treatment	738 cases	NR	4p15.1	Intergenic	rs17390445-?	1×10^{-7} (ziprasidone)	Affymetrix & Perlegen (492 900)
					9q33.3	Intergenic	rs888219-?	2×10^{-7} (risperidone)	
					12q23.1	*ANKS1B*	rs7968606-?	3×10^{-7} (olanzapine)	
					2q14.3	*CNTNAP5*	rs17727261-?	5×10^{-7} (risperidone)	
					1q21.3	Intergenic	rs10888501-?	1×10^{-6} (olanzapine)	
					6p24.1	Intergenic	rs1040994-?	2×10^{-6} (olanzapine)	
					15q13.3	TRPM1	rs17815774-?	3×10^{-6} (risperidone)	
					3q28	Intergenic	rs7635839-?	3×10^{-6} (olanzapine)	
					6p21.33	Intergenic	rs12526186-?	3×10^{-6} (risperidone)	
Shuldiner *et al.* [16]	19706858	Response to clopidogrel therapy	429 Amish individuals	140 white, 83 African American, and 4 unspecified individuals	10q24	CYP2C18-CYP2C19-CYP2C9-CYP2C8 cluster	rs12777823-?	1.5×10^{-13}	Affymetrix (400 230)
Ge *et al.* [44]	19684573	Response to hepatitis C treatment	871 Caucasian, 191 African American, and 75 Hispanic participants	NR	19q13.2	*IL28B*	rs12979860-C	1×10^{-28} (combined)	Illumina (565 759)
					6q21	*AKD2*	rs9400317-?	7×10^{-6} (combined)	
					4q34.3	Intergenic	rs17067123-?	8×10^{-6} (combined)	
Alkelai *et al.* [45]	19680635	Response to antipsychotic treatment	199 cases, 198 controls	NR	2q24.3	*FIGN*	rs12476047-C	3×10^{-6}	Affymetrix & Perlegen (495 172)
Comabella *et al.* [46]	19667218	Response to interferon beta therapy	53 responders, 53 nonresponders	49 responders, 45 non-responders	NS	NS	NS	NS	Affymetrix (428 867) (pooled)
Teichert *et al.* [47]	19578179	Acenocoumarol maintenance dosage	1451 Caucasian patients	287 patients	16	*STX4A MYST1 BCKDK RNF40 BCL7C CTF1 VKORC1 KIA0339 NM175901 IGAM ITGAL ITGAX GZNF689 PYCARD FUS FBXC19 BCKDK FLJ23426 FLJ23436 FTGAX RNF40 RNF40 SCRAP*	rs10871454-?	2×10^{-123}	Illumina (550 000)
					10	*CYP2C9 GZNF689 PYCARD FUS CYP2C18 CYP2C19 CYP2C8*	rs4086116-?	3.3×10^{-24}	
Daly *et al.* [48]	19483685	Drug-induced liver injury (flucloxacillin)	8 cases, 282 controls	NR	6p21.33	*HCP5, HLA-B*	rs2395029-?	9×10^{-33}	Illumina (866 399)
					3q27.3	*ST6GAL1*	rs10937275-?	1×10^{-8} (B*5701 positive)	
					3q11.2	*OR5H2*	=rs1497546-?	2×10^{-7}	
					12q12	*ALG10B*	rs6582630-?	1×10^{-6}	
					9p21.2	*C9orf82*	rs10812428-?	1×10^{-6}	
					15q26.2	*MCTP2*	rs4984390-?	4×10^{-6}	

Table 1 (*continued*)

Study	PMID	Trait	Initial sample size	Replication sample size	Region	Reported gene(s)	Strongest SNP-risk allele	*P* value	Platform (SNPs passing QC)
Perlis *et al.* [49]	19448189	Response to lithium treatment in bipolar disorder	458 lithium-treated patients, 719 nonlithium treated patients	359 patients	NS	NS	NS	NS	Affymetrix (~1.4 million) (imputed)
Takeucki *et al.* [14]	19300499	Warfarin maintenance dose	1053 individuals	588 individuals	16p11.2	*VKORC1*	rs9923231-T	3×10^{-181}	Illumina (325 997)
					10q23.33	*CYP2C9*	rs1057910-?	3×10^{-79}	
					10q23.33	*CYP2C9*	rs1799853-?	1×10^{-31}	
					19p13.12	*CYP4F2*	rs2108622-?	3×10^{-10}	
Yang *et al.* [50]	19176441	Response to treatment for acute lymphoblastic leukemia	487 children	NR	10p12.33	*ST8SIA6*	rs359312-T	9×10^{-8}	Affymetrix (476 796)
					2q33.1	*C2orf47*	rs1569175-T	9×10^{-7}	
					4q31.21	*IL15*	rs17007695-C	9×10^{-7}	
					20q13.12	*NCOA3*	rs6125048-T	2×10^{-6}	
					7p14.2	*ELMO1*	rs4723619-C	3×10^{-6}	
					7p21.2	*DGKB*	rs6971925-T	3×10^{-6}	
					10q26.12	intergenic	rs2901286-A	4×10^{-6}	
					11p15.1	intergenic	rs7128311-C	5×10^{-6}	
					6q25.3	intergenic	rs35229355-T	5×10^{-6}	
					5p13.2	*LMBRD2*	rs267759-A	7×10^{-6}	
					10p14	Intergenic	rs10508343-A	8×10^{-6}	
					11q21	*MAML2*	rs7115578-A	8×10	
French *et al.* [51]	19066393	Methotrexate polyglutamate intracellular accumulation	248 patients, 176 HapMap cell lines	NR	NA	NA	NA	NA	Affymetrix (447 287)
Mick *et al.* [52]	18821564	Methylphenidate efficacy in pediatric attention deficit hyperactivity disorder treatment	187 children with attention-deficit/ hyperactivity disorder	NR	NS	NS	NS	NS	Affymetrix (319,722)
Link *et al.* [22]	18650507	Statin-related muscle toxicity	85 cases, 90 controls	19 856 individuals	12p12.1	*SLCO1B1*	rs4149056-C	2×10^{-9}	Illumina (316 184)
Liu *et al.* [53]	18615156	Response to TNF antagonist treatment	89 cases	NR	20q12	*MAFB*	rs6028945-T	2×10^{-7}	Illumina (283 348)
					6q26	*QKI*	rs10945919-G	3×10^{-7}	
					9p21.2	*IFNK*	rs7046653-A	5×10^{-7}	
					7q21.3	*PON1*	rs854555-A	2×10^{-6}	
					20p11.21	*CST5*	rs6138150-T	3×10^{-6}	
					2q24.3	*LASS6*	rs13393173-A	4×10^{-6}	
					4p15.1	*CENTD1*	rs437943-G	4×10^{-6}	
					1p22.3	*LMO4*	rs983332-A	5×10^{-6}	
Sarasquete *et al.* [54]	18594024	Development of jaw osteonecrosis after bisphosphonate in myeloma	22 cases and 65 matched controls	NR	10q23.33	*CYP2C8*	rs1934951-T	4.231×10^{-6}	Affymetrix (500 568)
Turner *et al.* [55]	18591461	Response to diuretic therapy	194 Blacks, 195 Whites	NR	12q15	*LYZ, YEATS4, FRS2*	3-SNP haplotype	6×10^{-6}	Affymetrix (up to 102 334)
Cooper *et al.* [13]	18535201	Warfarin maintenance dose	181 individuals	374 individuals	16p11.2	*VKORC1*	rs10871454-?	5×10^{-34}	Illumina (538 629)
					10q23.33	*CYP2C9*	rs4086116-?	$6x\ 10^{-12}$	
					12p13.33	*CACNA1C*	rs216013-?	9×10^{-7}	

Volpi *et al.* [56]	18521091	Response to iloperidone treatment (QT prolongation)	183 individuals	NR	10q23.1	*NRG3*	rs4933824-T	2×10^{-6}	Affymetrix (339 272)
					14q12	*NUBPL*	rs7142881-A	2×10^{-6}	
					15q26.1	*SLCO3A1*	rs3924426-T	2×10^{-6}	
					18q12.2	*BRUNOL4*	rs4799915-T	3×10^{-6}	
					2q31.3	*CERKL*	rs993648-T	3×10^{-6}	
					4q32.3	*PALLD*	rs17054392-C	3×10^{-6}	
Lavedan *et al.* [57]	18521090	Response to iloperidone treatment (PANSS-T score)	106 individuals	104 individuals	NS	NS	NS	NS	Affymetrix (334 563)
Inada *et al.* [58]	18334916	Neuroleptic-induced treatment-resistant tardive dyskinesia	50 Japanese schizophrenia patients with treatment-resistant TD and 50 Japanese schizophrenia patients without TD	36 patients with TD and 136 patients without TD	3p25.3	*SLC6A11*	?	0.0004	? (40 573)
					5q34	*GABRB2*	?	0.00007	
					15q12	*GABRG3*	?	0.0006	
Byun *et al.* [59]	18195134	Response to interferon beta therapy	206 multiple sclerosis cases	NR	NS	NS	NS	NS	Affymetrix (~100 000) (pooled)
Hartford *et al.* [60]	17673902	Etoposide-induced secondary leukemia	3 secondary leukemia/myelodysplasia cases and germline DNA from 13 matched and 156 unmatched controls	NR	NR	Genes in adhesion, Wnt signaling and actin regulation [37] pathways	NR	NR	? (116 204)
Kindmark *et al.* [61]	17505501	Ximelagatran-related liver toxicity	74 cases, 130 controls	10 cases, 16 controls	6p21.3	*HLA-DRB1*	DRB1*07	9×10^{-6}	Perlegen (~266 722)
Jones *et al.* [62]	17329987	TPMT activity	87 HapMap cell lines	NR	NA	NA	NA	NA	NA

Columns list the first author of the study, the PubMedID (PMID), the sample size of the initial sample, the replication sample (if applicable), details of the regions and variants that were statistically significant (if applicable), and the genotyping platform (and number of markers evaluated) used. Studies are arranged in reverse chronological order, according to PubMed ID assignment.
GWAS, genome-wide association study; NA, not applicable; NR, not replicated; NS, not significant; PANSS-T, positive and negative syndrome scale total; QC, quality control; SNP, single nucleotide polymorphism; TD, tardive dyskinesia.

defined as markedly elevated creatinine kinase. In its most extreme form, this can result in the potentially fatal adverse effect of rhabdomyolysis, but these cases are exceedingly rare. The SEARCH Collaborative Group also found that myopathy was rare (approximately 0.1%) with low-dose simvastatin, so they focused their efforts on 98 cases identified in 6031 patients receiving high doses (80 mg/day) of the drug. A GWAS that studied 85 of these cases and 90 controls identified rs4363657, in perfect LD with a known functional nonsynonymous SNP in *SLCO1B1* at genome-wide significance. The 5-year incidence of myopathy was 18% in individuals homozygous for the risk allele (2.1% of the study group), 3% in heterozygotes, and 0.6% in those with no risk allele. The result was replicated in a separate cohort of patients receiving a lower dose of 40 mg/day (relative risk 2.6 per C allele).

The success of this study illustrates several important points in the study design of pharmacogenomic GWAS. First, large collaborative samples can provide a valuable resource for collecting a critical mass of patients with a rare phenotype. Second, rare phenotypes are sampled from the extreme tail of drug response distributions. As a result, genetic variants that influence these traits may have larger genetic effect sizes, and therefore be detectable with small sample sizes, than more common outcomes. Third, similar outcomes can sometimes be combined into a single case group. Here, in the initial association phase, definite and incipient myopathy patients were considered together. Fourth, replication of an association should take place in a similar population. In this study, the replication cohort was treated with a lower dose, 40 mg of simvastatin daily as compared with 80 mg in the initial group. We note that selecting cases from lower dose regimen for a follow-up study may be preferable to the converse (i.e. higher doses in the follow-up cohort), as those cases have a more extreme phenotype (by developing toxicity at a lower dose). This can limit the dilution of the association signal in the confirmatory study.

Several additional pharmacogenomics consortia have been established to evaluate a number of drug response outcomes, including the International Severe Irinotecan Neutropenia Consortium (*http://www.pharmgkb.org/views/project.jsp?pId=69*), and the International Tamoxifen Pharmacogenomics Consortium (*http://www.pharmgkb.org/views/project.jsp?pId=63*). These groups have pooled data from around the world to investigate genetic predictors of drug response with high power and comparison across global populations. Although the initial study of these consortia has typically been focused on candidate/known genetic effects, they are moving towards GWAS. For example, the International Warfarin Pharmacogenetics Consortia (*http://www.pharmgkb.org/views/project.jsp?pId=56*) originally combined data for over 4000 individuals from 24 international sites, to develop and test warfarin dosing algorithms [67], and are currently using the cohort data for a GWAS (through the International Warfarin Pharmacogenetics Consortia–GWAS consortium) to identify and confirm earlier findings, and potentially discover novel variants that explain potential trait variation across multiple populations. Such collaborations are extremely important for rare events, such as adverse events. The International Serious Adverse Events Consortium (*www.saeconsortium.org*) represents one important effort in pharmacogenomics for adverse events, pulling together commercial, academic, and industry partners to collect data for well-powered GWAS.

These combined datasets represent exciting resources for pharmacogenomics GWAS, but are not without challenges. Concerns with consistent data collection, storage, data-ownership issues, etc. can be concerns in these collaborative efforts.

Conclusion

GWAS have proven to be an exciting tool for gene mapping in common human traits, and are showing their potential in pharmacogenomic outcomes as well. As pharmacogenomic GWAS mature, there is an increased appreciation for issues that are specifically related to these unique phenotypes. Practical considerations, related to study design and available sample sizes highlight the need for creative methods of replication, beyond the traditional replication cohorts that are used for common disease genetics, and the necessity of combining samples across consortia. The complex physiology of drug response outcomes highlights the need for analytical methods that incorporate this complexity, using the wealth of information available about drug mechanisms and pathways.

Acknowledgements

The authors are grateful to the Pharmacogenetics Research Network (PGRN) Publications Committee for helpful advice. Some of the study described in this review was funded by the NIH/NIGMS Pharmacogenetics Research Network grants U01 GM61393, U01 HL65962, U01 HG04603, U01 GM61388, U01 GM63340, UO1 GM61390, and Database (U01GM61374 *http://www.pharmgkb.org*). This study was additionally supported by CA21765, the American Lebanese Syrian Associated Charities (ALSAC), and by CureSearch. Dr Relling receives a portion of the income St. Jude receives from licensing patent rights related to TPMT polymorphisms. This publication (or project) was partially supported by NIH/NCRR/OD UCSF-CTSI Grant Number KL2 RR024130. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Conflicts of interest

There are no conflicts of interest.

References

1 Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 2009; **106**:9362–9367.

2 McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**:356–369.
3 Weinshilboum R, Wang L. Pharmacogenomics: bench to bedside. *Nat Rev Drug Discov* 2004; **3**:739–748.
4 Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005; **6**:95–108.
5 Roden DM, Altman RB, Benowitz NL, Flockhart DA, Giacomini KM, Johnson JA, *et al.* Pharmacogenomics: challenges and opportunities. *Ann Intern Med* 2006; **145**:749–757.
6 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; **273**:1516–1517.
7 Weiss ST, Silverman EK, Palmer LJ. Case-control association studies in pharmacogenetics. *Pharmacogenomics J* 2001; **1**:157–158.
8 Roses AD. Pharmacogenetics in drug discovery and development: a translational perspective. *Nat Rev Drug Discov* 2008; **7**:807–817.
9 Limdi NA, Veenstra DL. Expectations, validity, and reality in pharmacogenetics. *J Clin Epidemiol* 2009. [Epub ahead of print]
10 Carlquist JF, Horne BD, Muhlestein JB, Lappe DL, Whiting BM, Kolek MJ, *et al.* Genotypes of the cytochrome p450 isoform, CYP2C9, and the vitamin K epoxide reductase complex subunit 1 conjointly determine stable warfarin dose: a prospective study. *J Thromb Thrombolysis* 2006; **22**:191–197.
11 Aquilante CL, Langaee TY, Lopez LM, Yarandi HN, Tromberg JS, Mohuczy D, *et al.* Influence of coagulation factor, vitamin K epoxide reductase complex subunit 1, and cytochrome P450 2C9 gene polymorphisms on warfarin dose requirements. *Clin Pharmacol Ther* 2006; **79**:291–302.
12 Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, *et al.* Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *N Engl J Med* 2005; **352**:2285–2293.
13 Cooper GM, Johnson JA, Langaee TY, Feng H, Stanaway IB, Schwarz UI, *et al.* A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood* 2008; **112**:1022–1027.
14 Takeuchi F, McGinnis R, Bourgeois S, Barnes C, Eriksson N, Soranzo N, *et al.* A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. *PLoS Genet* 2009; **5**:e1000433.
15 Kangelaris KN, Bent S, Nussbaum RL, Garcia DA, Tice JA. Genetic testing before anticoagulation? A systematic review of pharmacogenetic dosing of warfarin. *J Gen Intern Med* 2009; **24**:656–664.
16 Shuldiner AR, O'Connell JR, Bliden KP, Gandhi A, Ryan K, Horenstein RB, *et al.* Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA* 2009; **302**:849–857.
17 Crowley JJ, Sullivan PF, McLeod HL. Pharmacogenomic genome-wide association studies: lessons learned thus far. *Pharmacogenomics* 2009; **10**:161–163.
18 Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science* 2008; **322**:881–888.
19 Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics J* 2005; **5**:6–13.
20 Daly AK, Brockmoller J, Broly F, Eichelbaum M, Evans WE, Gonzalez FJ, *et al.* Nomenclature for human CYP2D6 alleles. *Pharmacogenetics* 1996; **6**:193–201.
21 Peters EJ, McLeod HL. Ability of whole-genome SNP arrays to capture 'must have' pharmacogenomic variants. *Pharmacogenomics* 2008; **9**:1573–1577.
22 Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, *et al.* SLCO1B1 variants and statin-induced myopathy – a genome-wide study. *N Engl J Med* 2008; **359**:789–799.
23 Guessous I, Gwinn M, Khoury MJ. Genome-wide association studies in pharmacogenomics: untapped potential for translation. *Genome Med* 2009; **1**:46.
24 Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001; **29**:306–309.
25 Thornton-Wells TA, Moore JH, Haines JL. Dissecting trait heterogeneity: a comparison of three clustering methods applied to genotypic data. *BMC Bioinformatics* 2006; **7**:204.
26 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**:559–575.
27 Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am J Hum Genet* 2009; **86**:6–22.
28 Wang L, Weinshilboum RM. Pharmacogenomics: candidate gene identification, functional validation and mechanisms. *Hum Mol Genet* 2008; **17**:R174–R179.
29 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, *et al.* Finding the missing heritability of complex diseases. *Nature* 2009; **461**:747–753.
30 Elbers CC, van Eijk KR, Franke L, Mulder F, van der Schouw YT, Wijmenga C, Onland-Moret NC. Using genome-wide pathway analysis to unravel the etiology of complex diseases. *Genet Epidemiol* 2009; **33**:419–431.
31 Eleftherohorinou H, Wright V, Hoggart C, Hartikainen AL, Jarvelin MR, Balding D, *et al.* Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. *PLoS One* 2009; **4**:e8068.
32 Bush WS, Dudek SM, Ritchie MD. Biofilter: a knowledge-integration system for the multi-locus analysis of genome-wide association studies. *Pac Symp Biocomput* 2009; 368–379.
33 Torkamani A, Topol EJ, Schork NJ. Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics* 2008; **92**:265–272.
34 Klein TE, Chang JT, Cho MK, Easton KL, Fergerson R, Hewett M, *et al.* Integrating genotype and phenotype information: an overview of the PharmGKB project. Pharmacogenetics Research Network and Knowledge Base. *Pharmacogenomics J* 2001; **1**:167–170.
35 Thompson JF, Hyde CL, Wood LS, Paciga SA, Hinds DA, Cox DR, *et al.* Comprehensive whole-genome and candidate gene analysis for response to statin therapy in the Treating to New Targets (TNT) cohort. *Circ Cardiovasc Genet* 2009;**2**:173–181.
36 Trevino LR, Shimasaki N, Yang W, Panetta JC, Cheng C, Pei D, *et al.* Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. *J Clin Oncol* 2009; **27**:5972–5978.
37 Aberg K, Adkins DE, Bukszar J, Webb BT, Caroff SN, Miller DD, *et al.* Genome-wide association study of movement-related adverse antipsychotic effects. *Biol Psychiatry* 2009; **67**:279–282.
38 Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, *et al.* A genomewide association study of citalopram response in major depressive disorder. *Biol Psychiatry* 2009; **67**:133–138.
39 Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, *et al.* IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009; **41**:1100–1104.
40 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**:1105–1109.
41 Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, *et al.* A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Arch Gen Psychiatry* 2009; **66**:966–975.
42 Laje G, Allen AS, Akula N, Manji H, John Rush A, McMahon FJ. Genome-wide association study of suicidal ideation emerging during citalopram treatment of depressed outpatients. *Pharmacogenet Genomics* 2009; **19**:666–674.
43 McClay JL, Adkins DE, Aberg K, Stroup S, Perkins DO, Vladimirov VI, *et al.* Genome-wide pharmacogenomic analysis of response to treatment with antipsychotics. *Mol Psychiatry* 2009. [Epub ahead of print]
44 Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, *et al.* Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009; **461**:399–401.
45 Alkelai A, Greenbaum L, Rigbi A, Kanyas K, Lerer B. Genome-wide association study of antipsychotic-induced parkinsonism severity among schizophrenia patients. *Psychopharmacology (Berl)* 2009; **206**:491–499.
46 Comabella M, Craig DW, Morcillo-Suarez C, Rio J, Navarro A, Fernandez M, *et al.* Genome-wide scan of 500 000 single-nucleotide polymorphisms among responders and nonresponders to interferon beta therapy in multiple sclerosis. *Arch Neurol* 2009; **66**:972–978.
47 Teichert M, Eijgelsheim M, Rivadeneira F, Uitterlinden AG, van Schaik RH, Hofman A, *et al.* A genome-wide association study of acenocoumarol maintenance dosage. *Hum Mol Genet* 2009; **18**:3758–3768.
48 Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, Floratos A, *et al.*; DILIGEN Study; International SAE Consortium. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat Genet* 2009; **41**:816–819.
49 Perlis RH, Smoller JW, Ferreira MA, McQuillin A, Bass N, Lawrence J, *et al.* A genomewide association study of response to lithium for prevention of recurrence in bipolar disorder. *Am J Psychiatry* 2009; **166**:718–725.

50 Yang JJ, Cheng C, Yang W, Pei D, Cao X, Fan Y, *et al.* Genome-wide interrogation of germline genetic variation associated with treatment response in childhood acute lymphoblastic leukemia. *JAMA* 2009; **301**:393–403.

51 French D, Yang W, Cheng C, Raimondi SC, Mullighan CG, Downing JR, *et al.* Acquired variation outweighs inherited variation in whole genome analysis of methotrexate polyglutamate accumulation in leukemia. *Blood* 2009; **113**:4512–4520.

52 Mick E, Neale B, Middleton FA, McGough JJ, Faraone SV. Genome-wide association study of response to methylphenidate in 187 children with attention-deficit/hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**:1412–1418.

53 Liu C, Batliwalla F, Li W, Lee A, Roubenoff R, Beckman E, *et al.* Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis. *Mol Med* 2008; **14**:575–581.

54 Sarasquete ME, Garcia-Sanz R, Marin L, Alcoceba M, Chillon MC, Balanzategui A, *et al.* Bisphosphonate-related osteonecrosis of the jaw is associated with polymorphisms of the cytochrome P450 CYP2C8 in multiple myeloma: a genome-wide single nucleotide polymorphism analysis. *Blood* 2008; **112**:2709–2712.

55 Turner ST, Bailey KR, Fridley BL, Chapman AB, Schwartz GL, Chai HS, *et al.* Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic. *Hypertension* 2008; **52**:359–365.

56 Volpi S, Heaton C, Mack K, Hamilton JB, Lannan R, Wolfgang CD, *et al.* Whole genome association study identifies polymorphisms associated with QT prolongation during iloperidone treatment of schizophrenia. *Mol Psychiatry* 2009; **14**:1024–1031.

57 Lavedan C, Volpi S, Polymeropoulos MH, Wolfgang CD. Effect of a ciliary neurotrophic factor polymorphism on schizophrenia symptom improvement in an iloperidone clinical trial. *Pharmacogenomics* 2008; **9**:289–301.

58 Inada T, Koga M, Ishiguro H, Horiuchi Y, Syu A, Yoshio T, *et al.* Pathway-based association analysis of genome-wide screening data suggest that genes associated with the gamma-aminobutyric acid receptor signaling pathway are involved in neuroleptic-induced, treatment-resistant tardive dyskinesia. *Pharmacogenet Genomics* 2008; **18**:317–323.

59 Byun E, Caillier SJ, Montalban X, Villoslada P, Fernandez O, Brassat D, *et al.* Genome-wide pharmacogenomic analysis of the response to interferon beta therapy in multiple sclerosis. *Arch Neurol* 2008; **65**:337–344.

60 Hartford C, Yang W, Cheng C, Fan Y, Liu W, Trevino L, *et al.* Genome scan implicates adhesion biological pathways in secondary leukemia. *Leukemia* 2007; **21**:2128–2136.

61 Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, *et al.* Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. *Pharmacogenomics J* 2008; **8**:186–195.

62 Jones TS, Yang W, Evans WE, Relling MV. Using HapMap tools in pharmacogenomic discovery: the thiopurine methyltransferase polymorphism. *Clin Pharmacol Ther* 2007; **81**:729–734.

63 The International HapMap Consortium. The International HapMap Project. *Nature* 2003; **426**:789–796.

64 Abe T, Unno M, Onogawa T, Tokui T, Kondo TN, Nakagomi R, *et al.* LST-2, a human liver-specific organic anion transporter, determines methotrexate sensitivity in gastrointestinal cancers. *Gastroenterology* 2001; **120**:1689–1699.

65 Tirona RG, Leake BF, Merino G, Kim RB. Polymorphisms in OATP-C: identification of multiple allelic variants associated with altered transport activity among European- and African-Americans. *J Biol Chem* 2001; **276**:35669–35675.

66 Lavedan C, Licamele L, Volpi S, Hamilton J, Heaton C, Mack K, *et al.* Association of the NPAS3 gene and five other loci with response to the antipsychotic iloperidone identified in a whole genome association study. *Mol Psychiatry* 2009; **14**:804–819.

67 Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, *et al.* Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med* 2009; **360**:753–764.