



Published in final edited form as:

*Pharmacogenomics*. 2009 February ; 10(2): 161–163. doi:10.2217/14622416.10.2.161.

## Pharmacogenomic GWAS: Lessons learned thus far

James J Crowley<sup>1</sup>, Patrick F Sullivan<sup>2</sup>, and Howard L McLeod<sup>1</sup>

<sup>1</sup>University of North Carolina Institute for Pharmacogenomics and Individualized Therapy, University of North Carolina, Chapel Hill, North Carolina, USA

<sup>2</sup>Department of Genetics and Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, North Carolina, USA

Over the past four years, the once hotly debated question of whether genomewide genetic mapping of common SNPs would shed light on common diseases has been answered with a resounding “yes”. More than 100 publications have now reported the localization of common SNPs associated with a wide range of common diseases (e.g., age-related macular degeneration, type 2 diabetes, Crohn’s disease, obesity) as well as various individual traits (height, hair color, eye color, freckling). As these publications accrued, a number of lessons regarding genetic mapping by genomewide association studies (GWAS) began to emerge (for a review see Altshuler et al, 2008 [1]). These lessons include: 1) effect sizes for common variants are typically modest; 2) with currently typical sample sizes (e.g., 2000 cases and 2000 controls), the power to detect associations has been low; 3) a single genomic region can harbor both common variants of weak effect and rare variants of large effect; 4) most confirmed associations do not involve candidate genes suspected on the basis of prior theory; 5) some associations implicate non–protein-coding regions; and 6) correlations between genetic variants and phenotypes have been limited by the accuracy and validity of the phenotypic measurement.

As the number of published pharmacogenomic GWAS begins to accumulate, it is prudent to search for lessons from these early studies. A search of PubMed up to January 15<sup>th</sup>, 2009 yielded 11 articles that examined the association between a drug-induced phenotype and at least 100,000 genome-wide SNP markers. Table 1 provides an overview of these studies and their most significant findings (as always, the reader is encouraged to examine the primary literature for a more thorough description of these studies). The most prominent characteristic that distinguishes these GWAS from their disease-oriented counterparts is sample size. While disease GWAS now routinely exceed 2,000 cases and 2,000 controls, all of these early pharmacogenomic GWAS had sample sizes under 400 drug-treated individuals. This is not surprising considering the scarcity of available DNA samples from clinical studies, particularly for rare side effects, and the considerable expense and effort required to phenotype individuals for a drug response. While the cost per subject in a disease case-control study is generally \$500–1000, the typical cost per subject in a clinical trial ranges from \$5,000–30,000. While these facts have been used to argue against the use of GWAS for pharmacogenomics, we believe that these early publications have helped identify a set of important lessons:

---

Correspondence to: Jim Crowley, PhD, University of North Carolina-Chapel Hill, 4113 Neurosciences Research Building, 103 Mason Farm Road, Chapel Hill, NC 27599, lab phone: (919) 966-9576, cell phone: (215) 913-8776, fax: (919) 966-3630, crowley@unc.edu.

### Financial & competing interests disclosure

This work was supported by the NIH Pharmacogenetics Research Network (U01 GM63340). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

1. *GWAS of rare adverse drug reactions (ADRs) may be more likely to yield highly penetrant variants.* ADRs have a major impact on patients, physicians, health care providers, regulatory agencies and pharmaceutical companies. Identifying the genetic contributions to ADR risk may lead to a better understanding of the underlying mechanisms, identification of patients at risk, and clinical testing could lead to a decrease in ADR incidence. The SEARCH Collaborative group [2] carried out a GWAS in 85 subjects who developed a rare simvastatin-induced myopathy and 90 simvastatin-exposed controls who did not develop this serious side effect. They identified common variants in *SLCO1B1*, a gene involved in statin hepatic transport, with odds ratios of 4.5 for each copy of the risk allele. Sarasquete et al [3] searched a large group of multiple myeloma patients for individuals who developed bisphosphonate-induced jaw osteonecrosis, a side effect that occurs in ~4% of patients. After genotyping 22 cases and 65 matched controls, they identified SNPs in the *CYP2C8* gene that may increase risk for osteonecrosis. Kindmark et al [4] conducted a retrospective case-control pharmacogenetic study of elevated serum alanine aminotransferase (ALAT) during long-term treatment with the oral direct thrombin inhibitor ximelagatran. This relatively small study (74 cases and 130 treated controls) yielded a strong, replicated genetic association between elevated ALAT and genetic variation in the major histocompatibility complex, suggesting a possible autoimmune pathogenesis. Cooper et al [5] performed a GWAS for the daily maintenance dose of the anticoagulant warfarin. This study replicated previous candidate gene studies that associated common SNPs in *VKORC1* and *CYP2C9* with large effects on warfarin dose.
2. *Use quantitative measures if possible.* For the same sample size, quantitative measures will generally be more powerful than discrete measures. The ADR studies by Kindmark and Cooper are consistent with this idea. Turner et al [6] conducted a GWAS to identify novel genes influencing diastolic blood pressure response to hydrochlorothiazide, a commonly prescribed diuretic. They discovered a novel gene cluster (encompassing *LYZ* and *YEATS4*) that is highly associated with blood pressure response in individuals of both African and European origin.
3. *Common events (e.g., treatment outcome or non-response) may be more multi-determined and intrinsically less tractable for GWAS.* Two studies in Table 1 that examined much more common side effects (Inada et al [7] and Volpi et al [8]) failed to identify significant associations, supporting the idea that rare side effects may be more likely to resemble ‘monogenetic traits’ and yield highly penetrant variants whereas common events may be akin to common diseases like type 2 diabetes mellitus. The latter requires very large studies given greater heterogeneity and far smaller genetic effects.
4. *Define responders and non-responders as the extremes of a larger distribution.* Therapeutic response to a pharmaceutical agent is generally a complex trait that is influenced by numerous genetic and environmental factors. Therefore, assuming adherence, the magnitude of the intended drug response generally shows a continuous phenotypic distribution in outbred populations. Most patients experience a partial therapeutic response, while patients at the tails of the distribution receive either no benefit or a full response. Selective genotyping of individuals at the extremes of the distribution often provides nearly equivalent power to complete genotyping. In addition, the accuracy to which patients are called responders or non-responders is likely to be highest for patients at the extremes of the distribution. In the antihypertensive GWAS study mentioned above, Turner et al [6] screened 600 individuals for response to hydrochlorothiazide and selected the 200 “best” and 200 “poorest” responders for

genotyping. They discovered a novel gene cluster associated with blood pressure response in individuals of both African and European origin. Two studies in Table 1 that defined responders and non-responders within the entirety of a small population (Mick et al [9] and Byun et al [10]) failed to identify significant associations, supporting the idea that focusing on extreme responders may be the most fruitful approach.

These are, of course, early days in the GWAS era for the field of pharmacogenomics and these lessons outlined above may not stand the test of time. However, the take-home message seems to be that the current question of whether or not GWAS will shed light on differential drug response is beginning to look like a “yes” – provided a GWAS is done with care, thoughtfulness, and an awareness of the intricacies of the phenotype.

## Bibliography

Papers of special note have been highlighted as:

\* of interest

\*\* of considerable interest

- 1\*. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008; 322:881–8. excellent primer on the current state-of-the art in complex trait mapping. [PubMed: 18988837]
- 2\*\*. Link E, Parish S, Armitage J, et al. *SLCO1B1* variants and statin-induced myopathy--a genomewide study. *N Engl J Med*. 2008; 359:789–99. demonstrates the ability of GWAS to map a rare side effect susceptibility variant in a small population. [PubMed: 18650507]
3. Sarasquete ME, Garcia-Sanz R, Marin L, 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; 112:2709–12. [PubMed: 18594024]
4. Kindmark A, Jawaid A, Harbron CG, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. *Pharmacogenomics J*. 2008; 8:186–95. [PubMed: 17505501]
- 5\*\*. Cooper GM, Johnson JA, Langae TY, et al. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood*. 2008; 112:1022–7. through GWAS demonstrated that common SNPs with large effects on warfarin dose are unlikely to be discovered outside of the *CYP2C9* and *VKORC1* genes. [PubMed: 18535201]
6. Turner ST, Bailey KR, Fridley BL, et al. Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic. *Hypertension*. 2008; 52:359–65. [PubMed: 18591461]
7. Inada T, Koga M, Ishiguro H, 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; 18:317–23. [PubMed: 18334916]
8. Volpi S, Heaton C, Mack K, et al. Whole genome association study identifies polymorphisms associated with QT prolongation during iloperidone treatment of schizophrenia. *Mol Psychiatry*. 2008
9. 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:1412–8. [PubMed: 18821564]
10. Byun E, Caillier SJ, Montalban X, et al. Genome-wide pharmacogenomic analysis of the response to interferon beta therapy in multiple sclerosis. *Arch Neurol*. 2008; 65:337–44. [PubMed: 18195134]
11. Liu C, Batliwalla F, Li W, et al. Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis. *Mol Med*. 2008; 14:575–81. [PubMed: 18615156]

12. Hartford C, Yang W, Cheng C, et al. Genome scan implicates adhesion biological pathways in secondary leukemia. *Leukemia*. 2007; 21:2128–36. [PubMed: 17673902]

Table 1

Summary of published pharmacogenomic GWAS studies

Ref	Pharmacological effect	GWAS sample (N)	Replication sample (N)	Genotyping platform	Genome-wide significance?	Top hit (SNP)	P-value (combined)	Odds ratio (95% CI)
5	Warfarin maintenance dose	181	374	Illumina 550K	yes	<i>VKORC1</i> (rs10871454)	$4.7 \times 10^{-34}$	-
2	Statin-induced myopathy	175	~20,000	Illumina 300K	yes	<i>SLCO1B1</i> (rs4363657)	$4.1 \times 10^{-9}$	16.9 (4.7–61.1)
4	Elevation of serum ALAT during treatment with ximelagatran	204	26	Perlegen (~350K)	no	<i>MHC-DQA1</i> (rs17426385)	$7.3 \times 10^{-8}$	4.6 (2.2–9.9)
11	Efficacy of anti-TNF treatment in rheumatoid arthritis	89	0	Illumina 300K	no	<i>MAFB</i> (rs6028945)	$2.0 \times 10^{-7}$	11.2 (2.3–108.1)
6	Antihypertensive Response to Thiazide Diuretic	389	0	Affy 100K	no	Chr 12q15 (rs7297610)	$2.4 \times 10^{-7}$	-
3	Development of jaw osteonecrosis after bisphosphonate in myeloma	87	0	Affy 500K	no	<i>CYP2C8</i> (rs1934951)	$1.1 \times 10^{-6}$	12.8 (3.7–43.5)
8	QT prolongation during loperidone treatment of schizophrenia	183	0	Affy 500K	no	<i>NUBPL</i> (rs7142881)	$1.6 \times 10^{-6}$	-
9	Efficacy of methylphenidate in children with ADHD	187	0	Affy 6.0	no	Chr 22q13 (rs9627183)	$3.0 \times 10^{-6}$	-
7	Neuroleptic-induced, treatment-resistant tardive dyskinesia	100 (50 TD+)	172 (36 TD+)	Illumina 100K	no	<i>SMYD3</i> (rs6426327)	$1.0 \times 10^{-5}$	-
10	Efficacy of interferon-beta therapy in multiple sclerosis	206	81	Affy 100K (pooled)	no	<i>HAPLN1</i> (rs4466137)	$4.0 \times 10^{-3}$	-
12	Etoposide-induced secondary leukemia	182 (13 leukemic)	0	Affy 100K	no	Multi-analytic study implicates adhesion pathways in secondary leukemias		