Published in final edited form as:

Biol Psychiatry. 2010 January 15; 67(2): 99–100. doi:10.1016/j.biopsych.2009.11.010.

Pioneering First Steps and Cautious Conclusions

Francis J. McMahon

Genetic Basis of Mood & Anxiety Disorders, National Institute of Mental Health, National Institutes of Health, US Department of Health and Human Services, Bethesda, Maryland.

More than half of all depressed patients fail to respond to the first prescribed antidepressant (1), and physicians have no good way to identify at the start of treatment who is likely to do well and who is not. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study was undertaken in large part to address this problem. In this issue, Garriock *et al.* (2) report the results of the first genome-wide association study (GWAS) in the STAR*D sample aimed at identifying common alleles that influence response to antidepressant treatment.

Genome-wide association studies have already taught us much, including some hard facts we might not have wished to hear. Genome-wide association studies interrogate a substantial and largely unbiased proportion of the common, single base-pair (single nucleotide polymorphism [SNP]) variation in the human genome, covering almost all genes and gene regulatory regions in a single experiment. Genome-wide association studies have offered a fresh perspective on the genomic landscape of common traits and diseases, a perspective that would not have followed from studies of candidate genes. For example, novel genes contributing to type II diabetes, heart disease, and several cancers have been robustly identified (reviewed in [3]). Robust findings have also begun to emerge for psychiatric disorders (reviewed in [4]). One surprising result, observed over a broad range of disease traits, has been the rather small effect sizes of the identified markers, with odds ratios on the order of 1.1 to 1.4. This suggests that the genetic component of most common diseases comprises many genes, each of small effect, that combine to influence individual risk.

While GWAS can be very valuable, the method has important limitations. Owing to the large number of independent tests performed, very large sample sizes are needed to provide sufficient statistical power. This fact is driven home by a simple example: in a GWAS that uses the typical 500,000 markers, chance alone predicts that about 500 will be associated with the trait at a $p < 10^{-3}$ level, and 5 markers will be significantly associated at a $p < 10^{-5}$ level—all by chance alone! This is why most authorities recommend setting the p value threshold for genome-wide significance at about 7×10^{-8}, close to a $p = .05$ after Bonferroni correction. Such a stringent level of statistical significance can only be achieved when either sample sizes or effect sizes are quite large. Since large effect sizes have only rarely been observed for common traits, very large sample sizes, on the order of 2000 to 50,000 cases and control subjects, are usually needed to assure good statistical power in a GWAS.

Address correspondence to Francis J. McMahon, M.D., 35 Convent Drive, Room 1A202, Bethesda, MD 20892; mcmahonf@mail.nih.gov.

Candidate gene studies do not need to correct for such a large number of independent tests, so they can have much greater power, although candidate gene studies are limited by the choice of candidates and difficulty correcting biases that are more easily handled in genome-wide data.

Numerous confounds may contaminate even genome-wide significant findings in a single study. For this reason, findings in one sample still need to be replicated in a second or third sample before they can be considered reliable. This has been achieved for many markers and many traits already. But one consequence of the power problem is that the reproducible signals are often not within the top 100 or even 1000 hits in the original study, especially when that study is based on a limited sample. This means that a premature focus on the top hits in a single sample may actually distract attention from the true hits.

Pharmacogenetic questions are a promising target for the GWAS method, because response to many medications appears to be mediated, at least in part, by genetic factors. Previous candidate gene studies have already identified promising markers of antidepressant outcome (5–9). Effect sizes often seem to be larger than those detected in disease studies, especially for adverse events (10–12). Still, the majority of signals will probably be of modest effect size, so pharmacogenetic GWAS still need quite large samples both for discovery and replication, often far beyond what is routinely available, especially in psychiatry.

The study by Garriock *et al.*, in the current issue, must be viewed in light of this reality, one that could not have been fully anticipated years ago when the authors began this important work. This is a well-done study on the largest sample of its kind in the field, comprising close to 2000 cases. All were outpatients with at least moderate major depressive disorder; all were treated with maximally tolerated doses of citalopram for at least 6 weeks; and all were evaluated prospectively using reliable instruments. Garriock *et al.* have employed reasonable definitions of both response and remission and have applied careful quality control to their genotype data. Despite all this, the results are disappointing: no markers achieve genome-wide significance.

Why? This study does have its flaws. Some are inherent in the STAR*D sample, which was not designed as a pharmacogenetic study. These include lack of placebo control, reliance on self-report of treatment adherence, incomplete information on concomitant medications, and DNA collection in a nonrandom subset of the total sample (reviewed in [13]). Other flaws are attributable to the early genotyping platform used by Garriock *et al.*, with incomplete genomic coverage and quality control problems. Indeed, the authors estimate that less than half of the common variants now known in the genome have been sampled in their study. These flaws increase the chances of a false negative finding. Of course, it is possible that common alleles do not really have much of an effect on citalopram response, but this seems unlikely, especially in light of the previous candidate-gene association findings (reviewed in [14]), some of which are supported in the Garriock *et al.* study.

Limited statistical power for a GWAS is by far the most likely reason that this study failed to produce significant findings. The authors estimate they had power to detect a common allele that increased the odds of response or remission by a factor of >1.75, but this is probably a

substantial overestimate of the true power, which will be cut in half by the incomplete coverage of common variants noted above. A GWAS in an even smaller sample, recently published in another journal (15), also failed to find anything genome-wide significant but suggested that many genes of small effect may indeed be involved.

Where do we go next? Despite its limitations, the STAR*D sample is probably as good as any the field is likely to have for the foreseeable future. Such a valuable sample deserves a second, much more complete genotyping effort that uses a state-of-the-art genotyping platform. Nearly complete data on up to ~2.5 million common SNPs can now be derived from the over 1 million SNPs that are directly genotyped with high accuracy by most of the currently available arrays. And the new arrays do not just cover common variation. Most provide important information on rarer SNPs and on common copy number variants—a form of variation in chromosomal structure, leading to the deletion or duplication of whole segments of DNA, which has proven to be quite important for a variety of central nervous system diseases.

We should also think about what it would take to generate a second sample like STAR*D. Such a sample would not only provide an essential replication set for the next GWAS but could also be designed to address some of the key limitations of the STAR*D sample noted above, especially the lack of placebo control. A new sample might also attempt to randomly assign patients to start with one of two different antidepressants, allowing an approach to the question of whether cases who fail to respond to one agent might respond to the other and whether this differential response is associated with particular genetic profiles. This may prove to be a key issue in developing genetic markers that can play a useful role in clinical decisions (16). The ongoing Genome Based Therapeutic Drugs for Depression study has some of these design features (17) but with a sample size under 1000 will not provide much power for GWAS approaches.

Collecting the necessary samples will not be easy and will not be cheap. But without them it is difficult to see how we are to fulfill the promise of personalized medicine within psychiatry. This is not a promise that can be addressed with genetic markers alone. We also need ways to incorporate clinical predictors and biomarkers into models of outcome prediction. This strategy is effective, for example, at predicting who will develop bleeding complications on warfarin (18). Some recent antidepressant studies suggest promising clinical predictors (19) and biomarkers (20) but much more work is needed in this area as well.

In light of all these caveats, what can we conclude with confidence after a GWAS in a single, relatively small sample, without genome-wide significant findings and without replication in an independent sample? Probably very little. Still, Garriock *et al.* report a pioneering project that will help blaze the trail for larger, more powerful studies to come. The authors should be congratulated for their bold first steps and commended for their cautious conclusions.

Acknowledgments

This work was supported by the National Institute of Mental Health Intramural Research Program. The opinions expressed herein do not necessarily reflect the policy of the National Institute of Mental Health, National Institutes of Health (NIH), or the federal government.

Dr. McMahon is a co-inventor on a patent application filed by the NIH entitled “Methods to Predict the Outcome of Treatment with Antidepressant Medication.” Under federal law, the NIH is required to pay the co-inventors a portion of any royalties the NIH receives under a future patent license, if granted.

References

1. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. Am J Psychiatry. 2006; 163:28–40. [PubMed: 16390886]
2. Garriok HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, et al. A genome-wide association study of citalopram response in major depressive disorder. Biol Psychiatry. 2010; 67:133–138. [PubMed: 19846067]
3. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. J Clin Invest. 2008; 118:1590–1605. [PubMed: 18451988]
4. Maher BS, Riley BP, Kendler KS. Psychiatric genetics gets a boost. Nat Genet. 2008; 40:1042–1044. [PubMed: 19165917]
5. Binder EB, Salyakina D, Lichtner P, Wochnik GM, Ising M, Pütz B, et al. Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. Nat Genet. 2004; 36:1319–1325. [PubMed: 15565110]
6. McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF, et al. Variation in the gene encoding the serotonin 2A receptor is associated with outcome of antidepressant treatment. Am J Hum Genet. 2006; 78:804–814. [PubMed: 16642436]
7. Paddock S, Laje G, Charney D, Rush AJ, Wilson AF, Sorant AJ, et al. Association of GRIK4 with outcome of antidepressant treatment in the STAR*D cohort. Am J Psychiatry. 2007; 164:1181–1188. [PubMed: 17671280]
8. Perlis RH, Moorjani P, Fagerness J, Purcell S, Trivedi MH, Fava M, et al. Pharmacogenetic analysis of genes implicated in rodent models of antidepressant response: Association of TREK1 and treatment resistance in the STAR*)D study. Neuropsychopharmacology. 2008; 33:2810–2819. [PubMed: 18288090]
9. Uhr M, Tontsch A, Namendorf C, Ripke S, Lucae S, Ising M, et al. Polymorphisms in the drug transporter gene ABCB1 predict antidepressant treatment response in depression. Neuron. 2008; 57:203–209. [PubMed: 18215618]
10. Laje G, Paddock S, Manji H, Rush AJ, Wilson AF, Charney D, McMahon FJ. Genetic markers of suicidal ideation emerging during citalopram treatment of major depression. Am J Psychiatry. 2007; 164:1530–1538. [PubMed: 17898344]
11. 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. [PubMed: 19684573]
12. Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe’er I, Floratos A, et al. HLA-B* 5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet. 2009; 41:816–819. [PubMed: 19483685]
13. Laje G, Perlis RH, Rush AJ, McMahon FJ. Pharmacogenetics studies in STAR*D: Strengths, limitations, and results. Psychiatr Serv. 2009; 60:1446–1457. [PubMed: 19880459]
14. Kato M, Serretti A. Review and meta-analysis of antidepressant pharmacogenetic findings in major depressive disorder [published online ahead of print Nov 4]. Mol Psychiatry. 2008
15. 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. [PubMed: 19736353]

16. Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: Testing for cytochrome P450 polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. Genet Med. 2007; 9:819–825. [PubMed: 18091431]

17. Uher R, Huezo-Diaz P, Perroud N, Smith R, Rietschel M, Mors O, et al. Genetic predictors of response to antidepressants in the GENDEP project. Pharmacogenomics J. 2009; 9:225–233. [PubMed: 19365399]

18. Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, et al. International Warfarin Pharmacogenetics Consortium. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med. 2009; 360:753–764. [PubMed: 19228618]

19. Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: A STAR*D report. Am J Psychiatry. 2008; 165:342–351. [PubMed: 18172020]

20. Leuchter AF, Cook IA, Gilmer WS, Marangell LB, Burgoyne KS, Howland RH, et al. Effectiveness of a quantitative electroencephalographic biomarker for predicting differential response or remission with escitalopram and bupropion in major depressive disorder. Psychiatry Res. 2009; 169:132–138. [PubMed: 19709754]