Pharmacogenomic Discovery Delineating the Genetic Basis of Drug Response

Wei Zhang · Yinan Zheng · Lifang Hou

Published online: 6 July 2013
© Springer Science+Business Media New York 2013

Abstract Personalized medicine has the promise to tailor medical care based on the patient's genetic make-up and clinical variables such as gender, race, and exposure to environmental stimuli. Recent progress in pharmacogenetic and pharmacogenomic studies has suggested that drug response to therapeutic treatments is likely a complex trait influenced by a variety of genetic and non-genetic factors. Identifying molecular targets (e.g., genetic variants) delineating the genetic basis of drug response could help understand the complex nature of drug response. The last decade has witnessed significant advances in genome-wide profiling technologies for genetic/epigenetic variations and gene expression. As an unbiased, cell-based model for pharmacogenomic discovery, a tremendous resource of whole-genome molecular targets has been accumulated for the HapMap lymphoblastoid cell lines (LCLs). The current progress, particularly in cancer pharmacogenomics, using the LCL model is reviewed to illustrate the potential impact of systems biology approaches on pharmacogenomic discovery.

Keywords Pharmacogenomics · Drug response · Gene expression · Genetic variation · Lymphoblastoid cell lines

Introduction

Personalized medicine provides the promise of health care tailored to an individual patient's genetic make-up and/or other clinical information such as gender, race, age, and exposure to environmental stimuli. The inter-individual variation in drug response has been an important clinical observation that may influence the efficacy of a therapeutic intervention in patients. Individual response to a medication could range from beneficial outcomes to adverse side effects and even fatal complications. Of particular concern are drugs that have a narrow therapeutic index, i.e., a small range of doses at which a medication may provide therapeutic benefits without causing severe adverse effects or complications. Warfarin, an anticoagulant normally prescribed for preventing thrombosis and thromboembolism, presents an example of how the patients' variation in drug response may require personalized decisions for clinical management [1]. In order to optimize its therapeutic efficacy without risking severe complications including bleeding, personalized dosing and close monitoring are required to achieve ideal clinical outcome of warfarin treatment. Recent pharmacogenetic and pharmacogenomic studies have indicated the associations of single nucleotide polymorphisms (SNPs) in *CYP2C9* (encoding cytochrome P450, family 2, subfamily C, polypeptide 9) and *VKORC1* (encoding vitamin K epoxide reductase complex, subunit 1) with warfarin dosing requirements [2]. The warfarin example demonstrated the importance of understanding the genetic basis of drug response, and the potential

W. Zhang (✉)
Department of Pediatrics and Institute of Human Genetics,
University of Illinois, 840 S. Wood St., CSB 1200 (M/C 856),
Chicago, IL 60612, USA
e-mail: weizhan1@uic.edu

Y. Zheng
Driskill Graduate Program in Life Sciences, Feinberg School
of Medicine, Northwestern University, Chicago, IL, USA

L. Hou
Department of Preventive Medicine and The Robert H. Lurie
Comprehensive Cancer Center, Feinberg School of Medicine,
Northwestern University, Chicago, IL, USA

applications of pharmacogenomic approaches in achieving optimal outcomes from therapeutic treatments in patients.

Given that the inter-individual response to a medication could be influenced by genes involved in both pharmacokinetics and pharmacodynamics, as well as by environmental factors such as food intake (e.g., irinotecan's interaction with St. John's wort) [3], drug response is likely to be a complex trait or phenotype affected by multiple genetic, epigenetic, and non-genetic factors. Elucidating the molecular targets (e.g., gene expression, SNPs, microRNAs, CpG methylation) associated with the variability of drug response could therefore enhance our understanding of the genetic/epigenetic basis of drug response, as well as provide potential intervention targets or diagnostic/prognostic biomarkers for achieving optimal therapeutic outcomes. In terms of experimental approaches, the last decade has witnessed significant advances in genome-wide profiling technologies (e.g., microarray and the next-generation sequencing), thus opening up the possibilities for high-throughput profiling and measuring of various molecular targets including gene expression, genetic variations, and, more recently, epigenetic variations. In the meantime, the launch of the Human Genome Project and other parallel large-scale projects such as the International HapMap Project [4, 5] and The 1000 Genomes Project [6, 7] has promoted an unprecedented

improvement of our understanding of the human genetic variation from diverse global populations. Pharmacogenetic and pharmacogenomic studies have benefited from these technical advancements and scientific breakthroughs during the past decade. In this review, the progress in identifying genome-wide molecular targets for the aims of delineating the genetic basis of drug response will be summarized using the HapMap [4, 5] lymphoblastoid cell lines (LCLs), a whole-genome, cell-based model, as an example, particularly with an emphasis from a systems biology perspective.

Cell-Based Pharmacogenomics Model

The HapMap LCLs were derived from apparently healthy individuals of major continental populations [4, 5]. The latest phase 3 HapMap samples are comprised of >1,100 samples from 11 global populations including major Asian, European, and African ethnic groups [8]. A comprehensive resource of gene expression and genetic variation data has been accumulated on the HapMap samples, particularly the CEU (Caucasian residents from Utah, USA) and YRI (Yoruba people from Ibadan, Nigeria) samples (Table 1). The availability of cell cultures for these LCLs from the Coriell Institute for Medical Research (Camden, NJ, USA) has allowed phenotyping in these samples, including drug

Table 1 Available resources for the HapMap CEU and YRI samples

Data type	Molecular target/ database	Genomic coverage	GEO accession/web resource
Genetic variation	SNP genotypes	~4 million SNPs	http://www.hapmap.org/
	SNP genotypes	~17 million SNPs	http://www.1000genomes.org/
	CNVs	>5,000	http://www.sanger.ac.uk/humgen/cnv/
	STRPs	721 STRPs	http://payseur.genetics.wisc.edu/
Transcriptome	mRNA	~22,000 transcript clusters; ~1.4 million exons	GSE7851 (gene-level); GSE9703 (exon-level)
	mRNA	8,793 transcripts	GSE10824
	mRNA	47,293 transcripts	GSE6536
	mRNA	RNA-Seq	GSE19480 ^a
Epigenome	microRNA	866 microRNAs	GSE34406
	DNA methylation	~480,000 CpGs	GSE39672
	DNA methylation	27,578 CpGs	GSE26133 ^a
	DNA methylation	27,578 CpGs	GSE27146
eQTLs	eqtl.uchicago.edu	–	http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/
	SCANdb	–	http://www.scandb.org/
Pharmacogenomic loci	PACdb	–	http://www.pacdb.org/

SNP single nucleotide polymorphism; CNV copy number variant; STRP short tandem repeat polymorphism; eQTL expression quantitative trait locus; SCANdb SNP and CNV annotation database; PACdb pharmacogenetics and cell-line database; CEU caucasians from Utah, USA; YRI Yoruba people from Ibadan, Nigeria

^a YRI only

response and transcriptomic profiling, facilitating pharmacogenomic discovery based on this model. In particular, the HapMap LCL samples have been used for modeling cytotoxicities induced by chemotherapeutic agents and response to radiation [9, 10], for which normal controls often cannot be found, given the toxicities of these treatments.

This cell-based model benefits from the availability of comprehensive genetic variation data including genotypes of >4 million SNPs for phase 1/2 [11] and >1.6 million SNPs for phase 3 samples [8] from the HapMap Project [4, 5]. The 1000 Genomes Project [6, 7] also covers a subset of the HapMap samples, thus providing even more dense coverage of genetic variants (>17 million SNPs), potentially useful for genotype imputation [12] to expand the original HapMap data. Given the fundamental role of gene expression in determining cellular phenotypes, microarray platforms including the Illumina BeadChip [13] (~45,000 transcripts), the Affymetrix Human Focused Array [14] (~9,000 genes), and the Affymetrix Human Exon Array [15, 16] (~22,000 transcript clusters and 1.4 million exons), and transcriptome sequencing (RNA-Seq) [17] were used to profile whole-genome gene expression in the HapMap samples to investigate the genetic architecture of gene expression variation both within and between human populations. Notably, eQTLs (expression quantitative trait loci), which are local (*cis*-acting) or distant (*trans*-acting) SNPs associated with quantitative gene expression phenotypes, derived from the HapMap samples [18] were demonstrated to be an important feature enriched among GWAS (genome-wide association study) loci associated with complex traits [19, 20•] and chemotherapy-induced cytotoxicities [21••]. More recently, microRNAs [22] and DNA methylation [23–25], two important epigenetic mechanisms, have also begun to be profiled using high-throughput technologies in the HapMap CEU and YRI samples, thus allowing further integrative analyses of both gene expression regulation and drug response phenotypes. In addition, internet-based, publicly accessible bioinformatic databases have been developed to disseminate primary results from studies using these samples, including, for example, the SCANDb [SNP and CNV (Copy Number Variant) Annotation database] [26•] for eQTLs and mQTLs (methylation-associated QTLs or genetic variants), and PACdb (Pharmacogenetics and Cell-line Database) [27•] for pharmacogenomic findings including SNPs and gene expression phenotypes associated with cytotoxicities. Several chemotherapeutic agents are used as examples in the following sections to illustrate how the pharmacogenomic loci or genetic/epigenetic molecular targets were identified using this model.

Genetic, Epigenetic Loci Associated with Cytotoxicities to Anticancer Agents

Genetic Mapping of Drug Response: eQTLs

Quantitative gene expression has been demonstrated to be a complex phenotype that is controlled by multiple genetic, epigenetic, and non-genetic factors (e.g., environmental factors). Common genetic variants in the form of SNPs have been detected to be associated with gene expression variation both within and between human populations [13–15, 28]. Taking advantage of the genotypic data on the HapMap LCL samples, GWAS can be performed to identify genetic variants associated with drug response phenotypes. In addition, eQTLs have been comprehensively mapped using the HapMap LCL samples and made publicly available [26•]. Notably, eQTLs have been found to be significantly enriched among GWAS loci for various complex traits/phenotypes [20•] and chemotherapy-induced cytotoxicities [21••]. Further integration of gene expression and eQTL information in the pharmacogenomic GWAS can, therefore, help prioritize and provide functional annotations (e.g., roles in gene regulation) for the drug response-associated genetic variants.

In particular, etoposide (VP-16) represents the first chemotherapeutic agent studied for pharmacogenomic loci using the HapMap LCL model [29]. Etoposide is an anticancer drug, which belongs to the class of topoisomerase II inhibitors [30]. Common adverse effects include, for example, myelosuppression and gastrointestinal toxicity, as well as risk of cardiac toxicity and secondary leukaemia [31]. Identifying genetic variants and gene expression phenotypes associated with etoposide-induced cytotoxicity can therefore enhance our understanding of the genetic basis of this cellular phenotype.

In this pioneer study, a genetics-driven ‘triangle’ approach (Fig. 1) was utilized to identify eQTLs that may contribute to etoposide-induced cytotoxicity. The study design aimed to integrate three genomic association tests, i.e., SNPs versus cytotoxicity, SNPs versus gene expression, and gene expression versus cytotoxicity. Specifically, common SNPs from the HapMap Project were evaluated for associations with etoposide-induced cytotoxicity (represented by IC_{50} , concentration at which 50 % of cell growth is inhibited), which was measured in 176 HapMap CEU and YRI samples. These cytotoxicity-associated SNPs ($P < 1E-4$) were then evaluated for associations with gene expression phenotypes, which were profiled using the Affymetrix Human Exon 1.0ST array (covering ~22,000 gene-level transcript clusters) (Table 1). The significant gene expression phenotypes were finally rechecked for associations with etoposide-induced cytotoxicity. With this unbiased, step-wise, genetics-driven strategy, pharmacogenomic loci for etoposide-induced cytotoxicity were prioritized based on their potential regulatory functions for those genes

Fig. 1 Integration of genome-wide molecular profiles in pharmacogenomic discovery. A genetics-driven ‘triangle’ approach aims to prioritize genetic variants associated with drug response by integrating information from gene expression and genetic variation. Expanding pharmacogenomic discovery to incorporate other gene regulatory mechanisms, e.g., epigenetic mechanisms such as microRNAs and DNA methylation is critical for elucidating the complex nature of drug response. Genetic variants associated with drug response can belong to various functional classes. *SNP* single nucleotide polymorphism, *CNV* copy number variant, *GWAS* genome-wide association study, *eQTL* expression quantitative trait locus, *miRSNP* SNP associated with microRNA expression, *M-eQTL* eQTL through regulating DNA methylation

associated with drug response. The final findings implicated several cancer-related genes including *AGPAT2* (encoding 1-acylglycerol-3-phosphate O-acyltransferase 2), *IL1B* (encoding interleukin 1, beta), and *WNT5B* (encoding wingless-type MMTV integration site family, member 5B), as well as some genes not yet known to be associated with sensitivity to etoposide [29] (i.e., novel candidate genes).

Genetic Mapping of Drug Response: CNVs

Common CNVs have been demonstrated to partially contribute to natural gene expression variation using the HapMap LCL model [13]. For example, SNPs and CNVs were shown to capture 83.6 and 17.7 % of the total detected genetic variation in gene expression profiled using the Illumina platform, respectively; however, the signals from the two types of variation had little overlap [13]. Previous studies also implicated CNVs in a variety of complex disorders such as rheumatoid arthritis, Crohn’s disease, type 1 diabetes, and type 2 diabetes [32]. In addition, a comprehensive scan of the NHGRI GWAS Catalog [19] showed an overrepresentation of SNPs tagging CNVs relative to frequency-matched SNPs [33], suggesting the potential roles of CNVs in determining complex traits and phenotypes. Expanding pharmacogenomic discovery to include CNVs could therefore provide novel insights into the genetic basis for drug response, beyond SNPs.

Recently, a comprehensive genome-wide study of >5,000 common CNVs [34] larger than 1 kb in size, assayed in the HapMap CEU samples, was performed to analyze their effects on cellular sensitivity to platinating agents (i.e., cisplatin and carboplatin) and topoisomerase II inhibitors (i.e., etoposide and daunorubicin) [35••], chemotherapeutic agents widely prescribed for a variety of cancers. Specifically, at the nominally significant threshold of $P < 0.05$, Gamazon et al. [35••] identified 67 CNVs associated with carboplatin IC_{50} , 70 CNVs with cisplatin IC_{50} , 73 CNVs with daunorubicin IC_{50} , and 113 CNVs with etoposide IC_{50} . Therefore, extensive CNV regions associated with cellular sensitivities were identified for functionally diverse chemotherapeutics, supporting the hypothesis that CNVs contribute to variation in drug response. Importantly, although SNPs tag some of the CNVs associated with drug sensitivities, several of the most significant CNV-drug associations were demonstrated to be independent of SNPs. Consequently, CNVs represent an independent source of genetic variations for pharmacologic phenotypes, in addition to SNPs.

Epigenetics of Drug Response: MicroRNAs

MicroRNAs (miRNAs) are short, non-coding RNA molecules that represent a class of gene regulators that generally downregulate gene expression at the post-transcriptional level, i.e., through translational repression and mRNA degradation by binding to the 3′ untranslated regions (3′UTRs) of the target mRNA [36]. Experimental evidence has been accumulated for the roles of miRNAs in regulating pharmacology-related genes and drug responses. For example, miRNAs have been shown to regulate or correlate with several VIPs (very important pharmacogenes) [e.g., miR-146a and *BRCA1* (encoding breast cancer 1, early onset), miR-126* and *CYP2A6* (encoding cytochrome P450, family 2, subfamily A, polypeptide 6)] [37, 38] maintained by the PharmGKB database [39]. Given the critical roles of miRNAs for various diseases (e.g., cancer, cardiovascular diseases) [40–43], incorporation of miRNAs into pharmacogenomic discovery could provide improved insights into drug response phenotypes [44]. In particular, whole-genome miRNA expression levels (covering ~1,000 human miRNAs using the Exiqon miRCURY LNA platform) are now available for the HapMap CEU and YRI samples [22, 45], allowing further integrative analyses of miRNAs in pharmacogenomic discovery.

For example, taking advantage of the miRNA expression data in the 60 unrelated CEU samples (Table 1), miR-193b* and miR-320 were shown to inhibit the expression of five previously identified platinating agent-associated genes: *CRIMI* (encoding cysteine rich transmembrane BMP regulator 1), *IFIT2* (encoding interferon-induced

protein with tetratricopeptide repeats 2), *OAS1* (encoding 2'-5'-oligoadenylate synthetase 1, 40/46 kDa), *KCNMA1* (encoding potassium large conductance calcium-activated channel, subfamily M, alpha member 1), and *GRAMD1B* (encoding GRAM domain containing 1B) [46••]. Furthermore, by integrating miRNAs with genetic variation and gene expression using the HapMap model, this pharmacogenomic study highlighted a potential mechanism of action for a previously observed genotype–survival outcome association in a cohort of ovarian cancer patients [46••], indicating the relevance of pharmacogenomic findings based on this model to clinical observations.

Epigenetics of Drug Response: DNA Methylation

DNA methylation, currently the most studied epigenetic mechanism, has been demonstrated to be an important class of gene regulators that putatively reflect the influence from environmental factors. Given the critical role played by DNA methylation in gene regulation, incorporation of epigenomic profiling would substantially add to the knowledge of how genetics, epigenetics, and environmental factors may contribute to gene expression and pharmacogenomics at a new level of complexity [47]. A high-throughput microarray platform (Illumina 27 K array, covering ~27,000 promoter CpGs) has been used to investigate DNA methylation variation in promoter regions in the CEU [24, 25] and YRI [24] samples. More recently, a much more comprehensive dataset (Table 1), using the Illumina 450 K array and covering more than 450,000 CpGs across the human genome [48], has been generated for the HapMap CEU and YRI samples [23, 49], thus allowing epigenomic mapping in pharmacogenomic discovery to incorporate this important gene regulatory mechanism.

The potential impact of integrating epigenomic mapping of CpG methylation in pharmacogenomic discovery can be demonstrated by a recent study on clofarabine [50••], a purine nucleoside analog, which is used in the treatment of refractory or relapsed acute leukemia and as induction therapy for stem cell transplantation [51–53]. Particularly, a GWAS using the HapMap genotypic data was performed to identify 182 SNPs associated with clofarabine cytotoxicity in the HapMap CEU and YRI samples [50••]. By integrating SNPs, gene expression, and DNA methylation data on these samples (Table 1), a set of 3 SNPs and 3 CpG sites targeting genes, including *SETBP1* (encoding SET binding protein 1), *BAG3* (encoding BCL2-associated athanogene 3), and *KLHL6* (encoding kelch-like family member 6), were found to explain 43.1 % of the observed variation in clofarabine-induced cytotoxicities, therefore indicating that incorporating these epigenetic profiles could help account for more of the observed variation in this drug response phenotype, compared to the SNPs alone (43.1 vs. 29.8 %) [50••].

Conclusion

During the past decade, technical advances in high-throughput profiling platforms, particularly microarrays, have allowed genome-wide profiling of various molecular targets including gene expression, SNP genotypes, copy number polymorphisms, microRNAs, and DNA methylation, thus facilitating pharmacogenomic studies that aim to identify genetic and epigenetic determinants associated with drug response. The HapMap LCL samples represent an example of how investigators have taken advantage of the technical advances in order to incorporate a variety of molecular targets into pharmacogenomic discovery. Though with some limitations (e.g., LCLs only represent one tissue type) [54], by using this unbiased, cell-based model, significant progress has been made to elucidate genetic and epigenetic contributions to cytotoxicities induced by anticancer chemotherapeutic agents of different mechanisms. With the next-generation sequencing technologies expected to become more accessible in the near future, a much more comprehensive genome-wide coverage of various molecular targets, including, for example, other novel gene regulatory mechanisms such as histone modifications, will allow the application of systems biology approaches to further expand the current pharmacogenomic discoveries to enhance our understanding of the complex genetic basis of drug response, ultimately facilitating the realization of personalized medical care.

Acknowledgment This work was supported, in part, by a Grant from the NHGRI/NIH (R21HG006367).

Disclosure W Zhang, Y Zheng, L Hou declares no conflicts of interest.

Compliance with Ethical Requirements This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Eriksson N, Wadelius M. Prediction of warfarin dose: why, when and how? *Pharmacogenomics*. 2012;13(4):429–40.
2. Cavallari LH, Shin J, Perera MA. Role of pharmacogenomics in the management of traditional and novel oral anticoagulants. *Pharmacotherapy*. 2011;31(12):1192–207.
3. Mathijssen RH, Verweij J, de Bruijn P, Loos WJ, Sparreboom A. Effects of St. John's wort on irinotecan metabolism. *J Natl Cancer Inst*. 2002;94(16):1247–9.

4. HapMap. The international HapMap project. *Nature*. 2003; 426(6968):789–96.
5. HapMap. A haplotype map of the human genome. *Nature*. 2005;437(7063):1299–320.
6. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467(7319):1061–73.
7. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56–65.
8. Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010;467(7311):52–8.
9. Niu N, Qin Y, Fridley BL, Hou J, Kalari KR, Zhu M, et al. Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines. *Genome Res*. 2010;20(11):1482–92.
10. Wheeler HE, Dolan ME. Lymphoblastoid cell lines in pharmacogenomic discovery and clinical translation. *Pharmacogenomics*. 2012;13(1):55–70.
11. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449(7164):851–61.
12. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)*. 2011;1(6):457–70.
13. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*. 2007; 315(5813):848–53.
14. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG. Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet*. 2007;39(2):226–31.
15. Zhang W, Duan S, Kistner EO, Bleibel WK, Huang RS, Clark TA, et al. Evaluation of genetic variation contributing to differences in gene expression between populations. *Am J Hum Genet*. 2008;82(3):631–40.
16. Zhang W, Duan S, Bleibel WK, Wisel SA, Huang RS, Wu X, et al. Identification of common genetic variants that account for transcript isoform variation between human populations. *Hum Genet*. 2009;125(1):81–93.
17. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature*. 2010;464(7289):768–72.
18. Duan S, Huang RS, Zhang W, Bleibel WK, Roe CA, Clark TA, et al. Genetic architecture of transcript-level variation in humans. *Am J Hum Genet*. 2008;82(5):1101–13.
19. Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*. 2009;106(23):9362–7.
20. • Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010;6(4):e1000888. *This paper shows that eQTLs derived from the HapMap samples are enriched in genetic loci associated with complex traits and diseases.*
21. •• Gamazon ER, Huang RS, Cox NJ, Dolan ME. Chemotherapeutic drug susceptibility associated SNPs are enriched in expression quantitative trait loci. *Proc Natl Acad Sci USA*. 2010;107(20):9287–92. *This paper shows that eQTLs derived from the HapMap samples are enriched in genetic loci associated with drug response phenotypes.*
22. Huang RS, Gamazon ER, Ziliak D, Wen Y, Im HK, Zhang W, et al. Population differences in microRNA expression and biological implications. *RNA Biol*. 2011;8(4):692–701.
23. Moen EL, Zhang X, Mu W, Delaney SM, Wing C, McQuade J et al. Genome-wide variation of cytosine modifications between European and African populations and the implications for complex traits. *Genetics*. 2013. doi:10.1534/genetics.113.151381.
24. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol*. 2011;12(1):R10.
25. Fraser HB, Lam LL, Neumann SM, Kobor MS. Population-specificity of human DNA methylation. *Genome Biol*. 2012;13(2):R8.
26. • Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL et al. SCAN: SNP and copy number annotation. *Bioinformatics*. 2010;26(2):259–62. *This paper introduces a resource of eQTLs derived from the HapMap samples.*
27. • Gamazon ER, Duan S, Zhang W, Huang RS, Kistner EO, Dolan ME et al. PACdb: a database for cell-based pharmacogenomics. *Pharmacogenet Genomics*. 2010;20(4):269–73. *This paper introduces a resource of pharmacogenomic loci derived from the HapMap samples.*
28. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, et al. Genetic analysis of genome-wide variation in human gene expression. *Nature*. 2004;430(7001):743–7.
29. Huang RS, Duan S, Bleibel WK, Kistner EO, Zhang W, Clark TA, et al. A genome-wide approach to identify genetic variants that contribute to etoposide-induced cytotoxicity. *Proc Natl Acad Sci USA*. 2007;104(23):9758–63.
30. Hande KR. Etoposide: four decades of development of a topoisomerase II inhibitor. *Eur J Cancer*. 1998;34(10):1514–21.
31. Seiter K. Toxicity of the topoisomerase II inhibitors. *Expert Opin Drug Saf*. 2005;4(2):219–34.
32. Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature*. 2010;464(7289):713–20.
33. Gamazon ER, Nicolae DL, Cox NJ. A study of CNVs as trait-associated polymorphisms and as expression quantitative trait loci. *PLoS Genet*. 2011;7(2):e1001292.
34. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, et al. Origins and functional impact of copy number variation in the human genome. *Nature*. 2010;464(7289):704–12.
35. •• Gamazon ER, Huang RS, Dolan ME, Cox NJ. Copy number polymorphisms and anticancer pharmacogenomics. *Genome Biol*. 2011;12(5):R46. *This study shows the importance of integrating CNVs in pharmacogenomic discovery.*
36. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet*. 2004;5(7):522–31.
37. Kalscheuer S, Zhang X, Zeng Y, Upadhyaya P. Differential expression of microRNAs in early-stage neoplastic transformation in the lungs of F344 rats chronically treated with the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol. *Carcinogenesis*. 2008;29(12):2394–9.
38. Shen J, Ambrosone CB, DiCioccio RA, Odunsi K, Lele SB, Zhao H. A functional polymorphism in the miR-146a gene and age of familial breast/ovarian cancer diagnosis. *Carcinogenesis*. 2008;29(10):1963–6.
39. 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(3): 167–70.
40. Gandellini P, Folini M, Zaffaroni N. Emerging role of microRNAs in prostate cancer: implications for personalized medicine. *Discov Med*. 2010;9(46):212–8.
41. Du L, Pertsemidis A. MicroRNAs and lung cancer: tumors and 22-mers. *Cancer Metastasis Rev*. 2010;29(1):109–22.

42. Yu Z, Baserga R, Chen L, Wang C, Lisanti MP, Pestell RG. MicroRNA, cell cycle, and human breast cancer. *Am J Pathol*. 2010;176(3):1058–64.
43. Lu Y, Zhang Y, Shan H, Pan Z, Li X, Li B, et al. MicroRNA-1 downregulation by propranolol in a rat model of myocardial infarction: a new mechanism for ischaemic cardioprotection. *Cardiovasc Res*. 2009;84(3):434–41.
44. Zhang W, Dolan ME. The emerging role of microRNAs in drug responses. *Curr Opin Mol Ther*. 2010;12(6):695–702.
45. Gamazon ER, Ziliak D, Im HK, LaCroix B, Park DS, Cox NJ, et al. Genetic architecture of microRNA expression: implications for the transcriptome and complex traits. *Am J Hum Genet*. 2012;90(6):1046–63.
46. •• Ziliak D, Gamazon ER, Lacroix B, Kyung Im H, Wen Y, Huang RS. Genetic variation that predicts platinum sensitivity reveals the role of miR-193b* in chemotherapeutic susceptibility. *Mol Cancer Ther*. 2012;11(9):2054–61. *This study shows the importance of microRNAs in pharmacogenomic discovery.*
47. Zhang W, Huang RS, Dolan ME. Integrating epigenomics into pharmacogenomic studies. *Pharmgenomics Pers Med*. 2008;2008(1):7–14.
48. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98(4):288–95.
49. Moen E, Mu W, Delaney S, Wing C, McQuade J, Godley L et al. Differences in DNA methylation between the African and European HapMap populations. AACR Annual Meeting, Chicago, IL; 2012.
50. •• Eadon M, Wheeler H, Stark A, Zhang X, Moen E, Delaney S et al. Genetic and epigenetic variants contributing to clofarabine cytotoxicity. *Hum Mol Genet*. 2013. doi:[10.1093/hmg/ddt240](https://doi.org/10.1093/hmg/ddt240). *This study demonstrates the impact of incorporating epigenomic data in pharmacogenomic discovery.*
51. Kantarjian H, Gandhi V, Cortes J, Verstovsek S, Du M, Garcia-Manero G, et al. Phase 2 clinical and pharmacologic study of clofarabine in patients with refractory or relapsed acute leukemia. *Blood*. 2003;102(7):2379–86.
52. Advani AS, Gundacker HM, Sala-Torra O, Radich JP, Lai R, Slovak ML, et al. Southwest oncology group study S0530: a phase 2 trial of clofarabine and cytarabine for relapsed or refractory acute lymphocytic leukaemia. *Br J Haematol*. 2010;151(5):430–4.
53. Kebriaei P, Basset R, Ledesma C, Ciurea S, Parmar S, Shpall EJ, et al. Clofarabine combined with busulfan provides excellent disease control in adult patients with acute lymphoblastic leukemia undergoing allogeneic hematopoietic stem cell transplantation. *Biol Blood Marrow Transpl*. 2012;18(12):1819–26.
54. Zhang W, Dolan ME. On the challenges of the HapMap resource. *Bioinformatics*. 2008;2(6):238–9.