The influence of Neanderthal alleles on cytotoxic response

Farida S. Akhtari[1,2], Tammy M. Havener[3], Masahide Fukudo[4], John R. Jack[2,5], Howard L. McLeod[6], Tim Wiltshire[3,7] and Alison A. Motsinger-Reif[2,5]

[1] Department of Biological Sciences, North Carolina State University, Raleigh, NC, United States of America
[2] Bioinformatics Research Center, North Carolina State University, Raleigh, NC, United States of America
[3] Pharmacotherapy and Experimental Therapeutics, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States of America
[4] Department of Pharmacy, Kyoto University, Kyoto, Japan
[5] Department of Statistics, North Carolina State University, Raleigh, NC, United States of America
[6] The DeBartolo Family Personalized Medicine Institute, Moffitt Cancer Center, Tampa, FL, United States of America
[7] Center for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States of America

ABSTRACT

Various studies have shown that people of Eurasian origin contain traces of DNA inherited from interbreeding with Neanderthals. Recent studies have demonstrated that these Neanderthal variants influence a range of clinically important traits and diseases. Thus, understanding the genetic factors responsible for the variability in individual response to drug or chemical exposure is a key goal of pharmacogenomics and toxicogenomics, as dose responses are clinically and epidemiologically important traits. It is well established that ethnic and racial differences are important in dose response traits, but to our knowledge the influence of Neanderthal ancestry on response to xenobiotics is unknown. Towards this aim, we examined if Neanderthal ancestry plays a role in cytotoxic response to anti-cancer drugs and toxic environmental chemicals. We identified common Neanderthal variants in lymphoblastoid cell lines (LCLs) derived from the globally diverse 1000 Genomes Project and Caucasian cell lines from the Children's Hospital of Oakland Research Institute. We analyzed the effects of these Neanderthal alleles on cytotoxic response to 29 anti-cancer drugs and 179 environmental chemicals at varying concentrations using genome-wide data. We identified and replicated single nucleotide polymorphisms (SNPs) from these association results, including a SNP in the *SNORD-113* cluster. Our results also show that the Neanderthal alleles cumulatively lead to increased sensitivity to both the anti-cancer drugs and the environmental chemicals. Our results demonstrate the influence of Neanderthal ancestry-informative markers on cytotoxic response. These results could be important in identifying biomarkers for personalized medicine or in dissecting the underlying etiology of dose response traits.

Corresponding author
Alison A. Motsinger-Reif,
motsinger@stat.ncsu.edu,
alison.motsinger@gmail.com

INTRODUCTION

Research has revealed that anatomically modern humans and archaic hominins, such as Neanderthals and Denisovans, coexisted and interbred (*Green et al., 2010*; *Higham et al., 2014*; *Prüfer et al., 2014*; *Sankararaman et al., 2014*). The Neanderthal introgression has led to modern Eurasian genomes containing ~1 to 4% of DNA inherited from Neanderthals (*Sankararaman et al., 2012*; *Wall et al., 2013*; *Vernot & Akey, 2014*). Recent studies have shown that Neanderthal haplotypes and alleles are under strong selective pressures, with regions undergoing both positive and negative selection (*Abi-Rached et al., 2011*; *Vernot & Akey, 2014*; *Sankararaman et al., 2014*). Importantly, these variants have been shown to be strongly associated with a number of clinically important diseases and phenotypes, such as hypercoagulation, depression, actinic keratosis, and urinary tract disorders among others (*Simonti et al., 2016*). These associations were discovered and replicated in electronic medical records data. Several of the single nucleotide polymorphisms (SNPs) in the Neanderthal PheWAS catalog are listed as being associated with drug response phenotypes such as 'Adverse drug events and allergies', 'Adverse effects of sedatives or other central nervous system depressants and anesthetics', 'Antineoplastic and immunosuppressive drugs causing adverse effects' and 'Allergic/adverse effects of penicillin'. This inspired us to investigate the effect of these Neanderthal alleles on drug or chemical response using the Lymphoblastoid Cell Lines (LCL) model.

Drug and chemical dose response are clinically important traits that are influenced by genetic factors. Variability in response to a drug or chemical is observed across individuals and populations, and it has been demonstrated that genetic variation is responsible for much of this variation. Understanding the genetic factors responsible for the variability in individual response to a drug or a chemical is what defines pharmacogenomics and toxicogenomics. The understanding of this variability in response is required to tailor medications to an individual or a population, thus leading to optimized clinical treatments (*Ritchie, 2012*). There are a number of high impact translational examples in this important area of gene mapping (*Ritchie, 2012*; *McCarthy, McLeod & Ginsburg, 2013*; *Motsinger-Reif et al., 2013*).

While many of the studies in pharmacogenomics and toxicogenomics focus on humans, a large proportion of studies rely on model systems. In particular, model systems allow for precise measurements of response, including mechanistic changes, but also very few studies in humans are designed to specifically measure xenobiotic responses that are known to induce a toxicity phenotype. Mostly, these data have been collected from retrospective and observational studies. Lymphoblastoid Cell Lines (LCLs) provide a convenient model for xenobiotic testing with their associated accumulated genomic data. However, they still have limitations due to their inability to provide response data where metabolism of compounds is an important part of the response, or where whole organism functions, such as immune response, are important.

In this study, we use LCLs as a model of cytotoxic response. The use of LCLs over clinical trials for pharmacogenomics studies has several advantages. LCLs from the CEPH pedigrees (*Dausset et al., 1990*), the International Hapmap Project (*International HapMap*

Consortium, 2003), The 1000 Genomes Project (*The 1000 Genomes Project Consortium, 2015*) and so on are publicly available resources with extensive genomic data. As opposed to traditional pharmacogenomic studies conducted via clinical trials, using the LCL model is more scalable, efficient and cost-effective. The LCL model is conducive to robotic automation for high-throughput assays. Also, there are fewer limitations with LCL models with respect to sample size and confounding effects from treatment variability (*Wheeler & Dolan, 2012*). The LCL model has been indicated to possess numerous successes (*Huang et al., 2008*; *Stark et al., 2010*; *Watson et al., 2011*; *Brown et al., 2012a*) and several of the results from these studies have been identified as clinically relevant and they include the association of SNPs in the MGMT gene with temozolomide response (*Brown et al., 2012a*), the identification of QTLs associated with dose response for 29 FDA-approved drugs (*Peters et al., 2011*) and the identification and evaluation of SNPs as predictors for platinum response in head and neck cancer patients (*Ziliak et al., 2011*).

Research has demonstrated that ethnicity and race influence the dose response phenotype, both in clinical samples and in the LCL model of drug response (*Ortega & Meyers, 2014*). Recently, we have shown strong differences in drug response in the LCL model for both pharmaceuticals and environmental chemicals. Differences in dose response phenotypes were observed between Hispanic and Caucasian populations for 28 chemotherapy drugs (*Jack et al., 2015*) and significant variation in EC_{10} was observed across the globally diverse populations from the 1000 Genomes Project (1000G) for 179 environmental chemicals (*Abdo et al., 2015*).

However, the influence of Neanderthal ancestry on dose response has not been interrogated. In this study, we examine if introgressed Neanderthal DNA influences the dose response phenotype in the LCL model. We assessed the effects of genome-wide Neanderthal alleles on cytotoxic response in 520 LCLs derived from unrelated Caucasian individuals from the Children's Hospital of Oakland Research Institute (CHORI) to 29 anti-cancer drugs, each assayed at six different concentrations (*Brown et al., 2014*) and in 1,086 LCLs from the 1000 Genomes (1000G) Project to 179 environmental chemicals, each assayed at eight different concentrations (*Abdo et al., 2015*). We first identified common Neanderthal variants that influence dose response in LCLs derived from Caucasian cell lines from CHORI. We then replicated this result in LCLs derived from the globally diverse 1000G Project. We also ascertain the cumulative effect of Neanderthal alleles on dose response using burden tests, again using the LCLs derived from Caucasian cell lines from CHORI as our discovery dataset and the LCLs from the 1000G Project as our replication dataset.

MATERIALS AND METHODS

To identify Neanderthal derived SNPs in the genome-wide dose response data, we used the Neanderthal SNPs listed in the Neanderthal PheWAS catalog (*Simonti et al., 2016*) which listed 1328 unique Neanderthal derived SNPs. These Neanderthal SNPs were identified by comparing the sequences from the 1000 Genomes Project (*The 1000 Genomes Project Consortium, 2015*) with the Altai Neanderthal genome (*Prüfer et al., 2014*;

Vernot & Akey, 2014). The Neanderthal PheWAS catalog provides the association results of 1328 Neanderthal SNPs with 1152 electronic health record (EHR) –derived phenotypes (*Simonti et al., 2016*).

We used the genome-wide association mapping results and dose response profile data from two previous studies. The first was from a study by *Brown et al. (2014)*, which we used as our discovery dataset. This study provided data from 520 Caucasian cell lines in response to 29 chemotherapy drugs, each assayed at six different concentrations, with 2.1 million SNPs used in the association analysis. Details of the drugs and concentrations used in this study are listed in Table S1. Details of the methods used to assess drug response can be found in *Brown et al. (2014)*. Potential batch effects were controlled by both the experimental design and by our quality control procedures. Briefly, cell lines were assayed in random order with each cell line being seeded on two 384 well plates with 4,000 cells per well. Each plate also included a control for background fluorescence signal, 10% dimethyl sulfoxide (DMSO) and another control for drug vehicle, water or DMSO at 0.01, 0.1, 1 and 2 percent. Each cell line was incubated with each of the 29 anti-cancer drugs at each of the six concentrations for 72 h, dyed with Alamar Blue (BioSource International) and incubated for another 24 h. At this point, LCL viability measurements were taken using an Infinite F200 microplate reader with Connect Stacker (Tecan Group Ltd, Männedorf, Switzerland) and iControl software (Version 1.6), which measures fluorescence intensity in raw fluorescence units (RFUs) at EX535nm and EM595nm. All individuals in this study were genotyped using HumanHap300 bead chip or HumanQuad610 bead chip platforms and these markers were used to impute 2.5 million SNPs from HapMap Release 22, using the Caucasian CEPH reference population. Genotypic quality control on these 2.5 million SNPs resulted in ~2.1 million SNPs per drug for association analysis. The second set of dose response data was obtained from a study by *Abdo et al. (2015)*, which we used as our replication dataset. This study provided data from 1,086 LCLs from the 1000 Genomes Project, representing nine populations from Europe, Asia, Africa and the Americas. The dose response profile data from this study consists of the cytotoxic response to 179 environmental chemicals each assayed at 8 different concentrations, with 1.3 million SNPs used in the association analysis. Chemicals and concentrations used in this study are listed in Table S2. Cell lines were divided into screening batches equally distributed by population and sex in each batch to limit the effect of any systematic bias, such as that of population substructure, on the assay results. Each cell line was plated on one or two, 1,536-well plates with 2,000 cells/5μL/well. DMSO at 0.46% vol/vol was used as the negative control and tetra-octyl-ammonium bromide was used as the positive control. There were no significant differences in the control values between ethnic groups. The CellTiter-Glo Luminescent Cell Viability (Promega, Madison, WI, USA) assay was used to assess intracellular ATP concentration, a marker for viability/cytotoxicity, 40 h after treatment. Luminescent intensity was detected using a ViewLux plate reader (PerkinElmer). From the 1,086 cell lines, a subset of 884 "unrelated" individuals was chosen. Of the 884 individuals, genotyped SNPs were obtained from the Illumina HumanOmni2.5 platform for 761 individuals and from HapMap for the remaining 123 individuals. A set of 875 samples from the 1000 Genomes set were used as reference to impute 1.3 million SNPs for

association analysis. Genotypic quality control on these SNPs resulted in 699,068 SNPs per chemical for association analysis.

The genomic/genotype data used for the chemical response is publicly available through the 1000 Genomes Project (http://www.internationalgenome.org). The genomic/genotype data used for the chemotherapy response can be found on dbGAP: Cholesterol and Pharmacogenetics (CAP) Study dbGaP Study Accession: phs000481.v2.p1.

We retrieved all the SNPs identified in each of the above genome-wide association mapping studies. We performed quality control on the data, which mainly involved removing rare variants, which was determined to be markers with <20 individuals for any genotype. We identified the Neanderthal-derived SNPs from the Neanderthal PheWAS catalog (*Simonti et al., 2016*) in each of the genome-wide data based on rs-ID. We then used these results as our initial data to identify the influence of individual Neanderthal-derived SNPs on dose response. A false discovery rate of $q < 0.25$, using the R function *p.adjust()* (package *stats v3.4.0*) (*R Development Core Team, 2016*), applied per drug or chemical was used to identify significant SNPs. For each of the significant Neanderthal SNPs that replicated in both datasets, we inspected a region of ± 50 kilobases (kb) using LocusZoom (*Pruim et al., 2010*) (version 1.3, Human Genome assembly (build hg18) *Lander et al., 2001*) to locate flanking genes. The workflow for this analysis is depicted in Fig. 1.

To identify the cumulative effect of Neanderthal alleles on dose response, we performed burden tests. We used the dbSNP database (*Sherry et al., 2001*) to identify the Neanderthal allele for each of the significant Neanderthal-derived SNPs. We used PLINK v1.07 (*Purcell*; *Purcell et al., 2007*) to aggregate information across all significant Neanderthal-derived SNPs to compute a single Neanderthal allele dosage number representing the cumulative effect of these Neanderthal-derived SNPs in each individual, for each study. We then modeled each response, at each concentration, as a function of the Neanderthal allele dosage number using a linear regression approach (R function *lm()*, package *stats v3.4.0*) (*R Development Core Team, 2016*). Significant results were obtained after correcting for multiple testing, using the Benjamini–Hochberg method with a false discovery rate of $q < 0.25$, on a per drug or chemical basis (R function *p.adjust()*, package *stats v3.4.0*) (*R Development Core Team, 2016*). The workflow for this analysis is depicted in Fig. 2. We also performed burden testing using a univariate summary measure, the half maximal inhibitory concentration (EC_{50}). We used linear regression to model EC_{50} as a function of the Neanderthal allele dosage number.

We attempted to perform phenome-wide association study (PheWAS)-based analyses to identify different phenotypes or phenotype groups associated with the Neanderthal-derived SNPs in our datasets as compared to the non-Neanderthal-derived SNPs. We isolated the phenotypes associated with the significant Neanderthal and non-Neanderthal SNPs in each of our datasets, using the PheWAS catalog (*Denny et al., 2013*). However, none or very few of the significant non-Neanderthal SNPs from the discovery and replication dataset were listed in the PheWAS catalog, thus preventing further analyses.

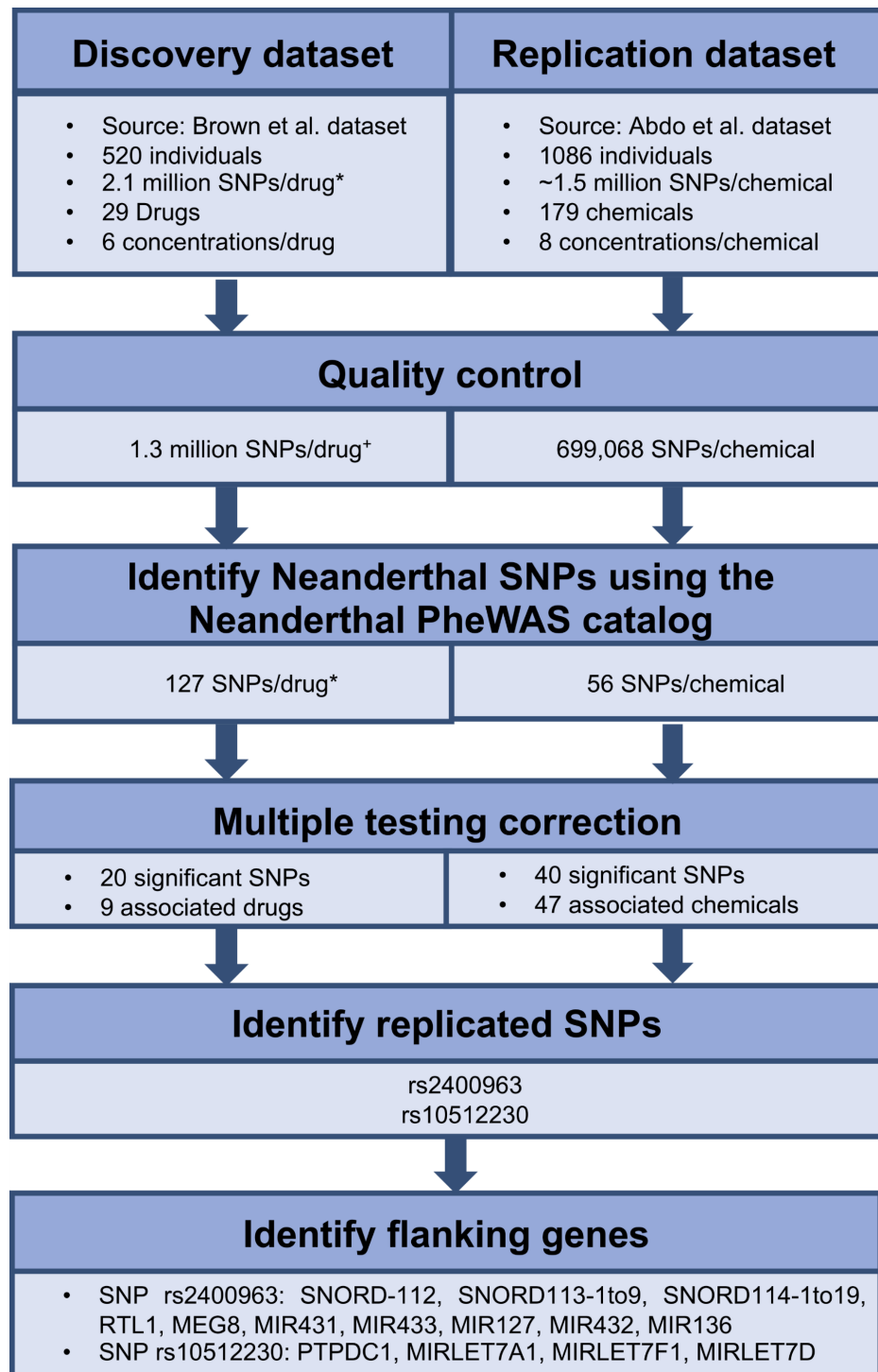
All data used in the study have been previously collected. All data are de-identified, and IRB exceptions have been filed.

+ except for Five-Fluorouracil (~1 million) and Nilotinib (~815k)
* except for Five-Fluorouracil (50) and Nilotinib (19)

Figure 1 **Workflow used to identify the effect of individual Neanderthal SNPs on dose response in the discovery and replication datasets.**

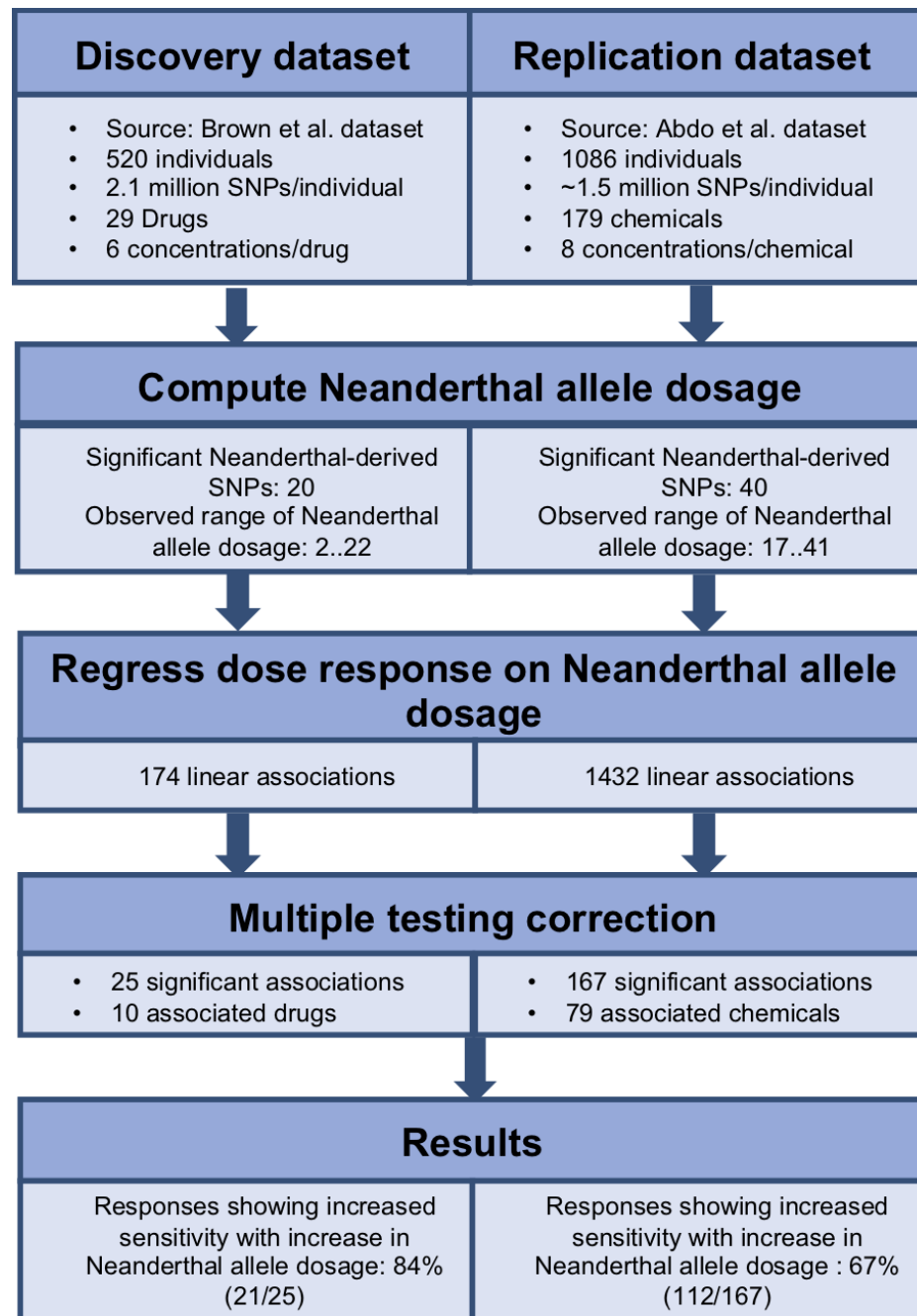



Figure 2 Workflow used to determine the cumulative effect of Neanderthal SNPs on dose response in the discovery and replication datasets.

Table 1 **Individual Neanderthal SNPs with significant and replicating dose response associations along with their associated drug or chemical and flanking genes over a ±50 kb region.** *SNORD-* small nucleolar RNA, C/D Box, *RTL1* retrotransposon Gag like 1, *MEG8-* maternally Expressed 8, *MIR* microRNA, *PTPDC1-* protein tyrosine phosphatase domain containing 1. Chromosomal position for SNPs was determined using the Human Genome assembly (build hg18). Multiple testing correction was done using the Benjamini-Hochberg method with a false discovery rate of *q* < 0.25, on a per drug or chemical basis.

Chr:position	SNP	Drug/Chemical	*p*-value	*q*-value	Flanking genes
14:100465459	rs2400963	Amiloride hydrochloride	3.97×10^{-4}	2.22×10^{-2}	*SNORD-112,*
		2,3,4,5-Tetrachloronitrobenzene	1.13×10^{-3}	9.45×10^{-2}	*SNORD113-1to9,*
		Nilotinib	1.04×10^{-2}	1.98×10^{-1}	*SNORD114-1to19,*
					RTL1, MEG8, MIR431,
					MIR433, MIR127,
					MIR432, MIR136
9:95956051	rs10512230	Nitazoxanide	2.58×10^{-3}	1.45×10^{-1}	*PTPDC1, MIRLET7A1,*
		Tetra-N-Octylammonium bromide	2.23×10^{-3}	1.25×10^{-1}	*MIRLET7F1,*
		2-Biphenylamine	8.07×10^{-3}	2.40×10^{-1}	*MIRLET7D*
		Chlordane (technical grade)	3.74×10^{-3}	2.07×10^{-1}	
		Cytarabine	1.21×10^{-2}	2.21×10^{-1}	

RESULTS

Effect of individual Neanderthal SNPs

Twenty Neanderthal-derived SNPs in the *Brown et al. (2014)* data and 40 Neanderthal-derived SNPs in the *Abdo et al. (2015)* data were deemed significant with a false discovery rate of *q* < 0.25 applied per drug or chemical. Two Neanderthal SNPs replicated in both data sets, rs2400963 and rs10512230. These significant replicated SNPs along with their flanking genes and the associated drugs or chemicals are listed in Table 1. The LocusZoom (*Pruim et al., 2010*) plots showing the regional genes for each SNP are shown in Figs. S1 and S2. All except the protein tyrosine phosphatase domain containing 1 (*PTPDC1*) gene are non-coding RNA genes. Several of the flanking genes have been previously shown to be associated with cancer. SNP rs2400963 is in the *SNORD113-3* gene. SNP rs2400963 and its flanking genes (*SNORD-112, SNORD113-1* to *-9, SNORD114-1* to *-19, RTL1, MEG8, MIR431 (miR431), MIR433 (miR433), MIR432 (miR432), MIR136 (miR136)*) are located in the *DLK1-DIO3* locus at the *14q32* region. The *DLK1-DIO3* region contains several maternally and paternally expressed imprinted genes. This region hosts several miRNA clusters, many of which are differentially expressed in various cancers (*Benetatos et al., 2013*). For example, *MIR127* and other microRNA (miRNA) genes located at *14q32* region show overexpression in acute myeloid leukemia bearing a t(15;17) translocation (*Valleron et al., 2012*). Overexpression of the *DLK1-DIO3* miRNA cluster at *14q32.2* is associated with increased expression of hepatocellular carcinoma (HCC) stem cell markers and poor survival rate in HCC patients (*Luk et al., 2011*). A specific small nucleolar RNA (snoRNA) profile has been reported in acute promyelocytic leukemia (APL) with ectopic expression of *SNORD112, SNORD113* and *SNORD114* snoRNA clusters (*Valleron et al., 2012*). *SNORD113-1* has been shown to function as a tumor suppressor in HCC (*Xu et al., 2014*). SNP rs10512230 is an expression quantitative trait loci (eQTL) of the *RP11-307E17.8* gene (*LOC100132077*) while *RP11-307E17.8* is a non-coding RNA gene. Data in the Genotype-Tissue Expression (GTEX) Project (*The GTEx Consortium, 2013*)

(accessed on 06/27/2017) shows that the Neanderthal allele at rs10512230 is significantly associated with increased expression of the *RP11-307E17.8* gene in transformed fibroblast cells and in the muscularis mucosae of the esophagus. SNP rs10512230 is flanked by the *PTPDC1*, *MIRLET7A1 (Let-7a-1)*, *MIRLET7F1 (Let-7f-1)* and *MIRLET7D (Let-7d)* genes. *PTPDC1* is a Protein Coding gene, in the p21-activated kinases (PAK) pathway. PAKs are involved in cell proliferation signaling, cell death resistance and metastasis, which are key hallmarks of cancer (*Ye & Field, 2012*). *MIRLET7A1*, *MIRLET7F1*, *MIRLET7D* are microRNA genes which are involved in cancer and DNA damage response pathways. Thus, as described above, various studies have reported links of the identified flanking genes in cancer pathways. Hence, we hypothesize that the Neanderthal-derived SNPs rs2400963 and rs10512230 may be associated with cytotoxic response.

Cumulative effect of Neanderthal alleles

We employed a linear regression model to analyze the cumulative effect of Neanderthal alleles on the dose–response phenotype, for each individual at each concentration for each drug or chemical. Each drug or chemical was considered as a separate test and a false discovery control was applied per drug or chemical. Results with false discovery rates of $q < 0.25$ were considered significant. Further, 25 Neanderthal allele dosage-cytotoxic response associations for 10 chemotherapy drugs from the *Brown et al. (2014)* data were declared significant. Also, 167 Neanderthal allele dosage-cytotoxic response associations for 79 environmental chemicals from the *Abdo et al. (2015)* data were declared significant.

For a given drug or chemical, at a fixed concentration, as the Neanderthal allele dosage increased, cell viability consistently decreased. This is shown in Fig. 3 for the chemotherapy drug Paclitaxel. In this test, 84% (21/25) of responses showed increased sensitivity to the chemotherapy drugs with increase in Neanderthal allele dosage. On the other hand, 67% (112/167) of responses showed increased susceptibility to environmental chemicals with increase in Neanderthal allele dosage. This is a remarkable trend observed in a large percentage of responses in data from both studies. Thus, we conclude that an increase in Neanderthal allele dosage leads to increased sensitivity to anti-cancer drugs and environmental chemicals.

We did not find consistent results in the linear regression model with EC_{50} as a function of the Neanderthal allele dosage number. This may be due to the fact that the univariate summary measure may not capture the true differences in the dose response profile if the differences are not entirely due to EC_{50}. Univariate summary measures have been shown to be less robust and have less power as compared to the full dose–response profile (*Brown et al., 2012b*). Since the *Abdo et al. (2015)* data consists of cell lines from geographically and ancestrally diverse populations, we repeated the linear regression model for this data with the cell lines stratified by continental origin. Results from this stratified analysis showed the same trend as above for the cell lines from European and Asian populations but not for American and African populations. These results are as expected because the Eurasian genome contains the largest percentage of DNA inherited from Neanderthals (*Sankararaman et al., 2012*; *Wall et al., 2013*; *Vernot & Akey, 2014*). While the European

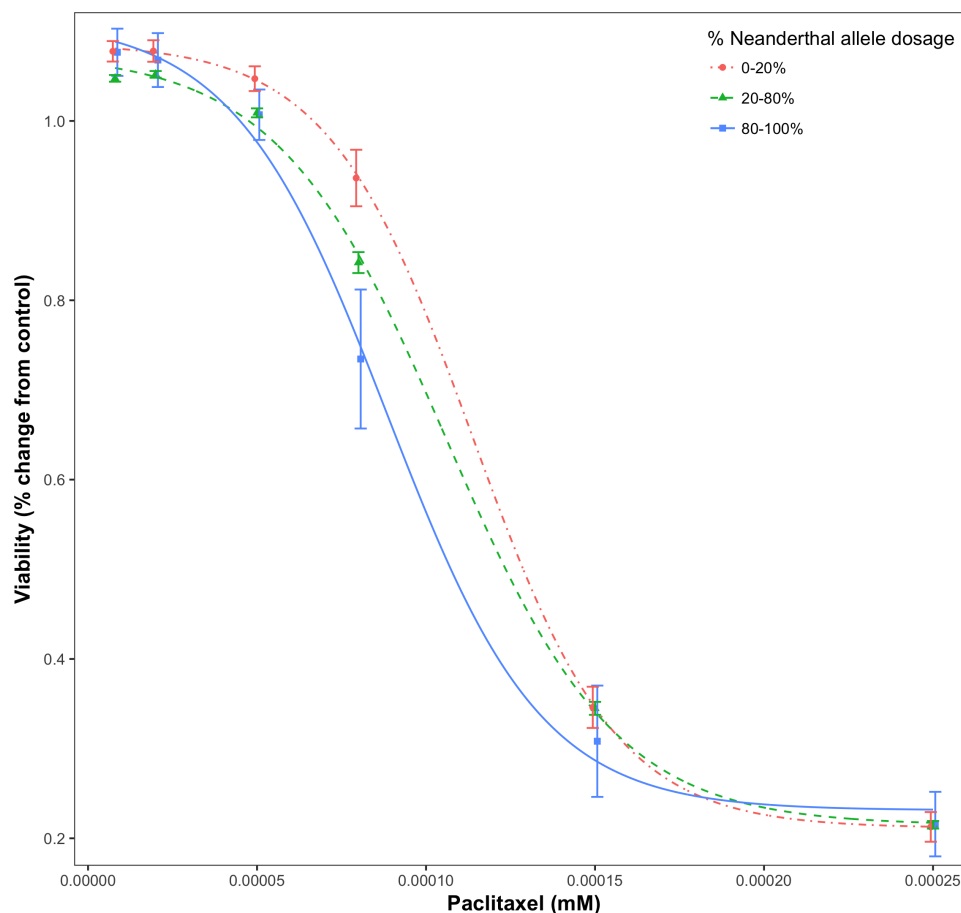



Figure 3 **Dose-response curve for Paclitaxel, a chemotherapy drug.** Dose-response is stratified by percent Neanderthal allele dosage as shown. Bars represent the standard error of the mean. Increase in Neanderthal allele dosage is accompanied by a decrease in cell viability at a fixed drug concentration.

population showed the trend of increased sensitivity with increase in Neanderthal allele dosage, we did observe fewer significant results in the European population than expected.

Detailed results

In the main text we focus on the top results, but additional details of all the results can be found in the Supplemental Material. Table S3 includes details of all SNP-drug and SNP-chemical associations for the single SNP analysis, including the EC_{50} values for each of the genotypes. Table S4 expands on the hypothesis generation aspect of our analyses to provide additional annotation and references to support further investigation of putative associations.

DISCUSSION

Various studies have shown that archaic admixture has led to genetic variability that has played a role in human adaptation and continues to influence modern human traits. Denisovan introgression has been identified as the cause of an unusual haplotype in the

EPAS1 gene region leading to favorable altitude adaptation in Tibetans (*Huerta-Sánchez et al., 2014*). Neanderthal variants have been found to be associated with various modern human phenotypes, such as hypercoagulation, depression, actinic keratosis, urinary tract disorders, etc. (*Simonti et al., 2016*). Introgressed Neanderthal and Denisovan alleles were found to be associated with increased immune system response and increased susceptibility to allergic diseases (*Dannemann, Andrés & Kelso, 2016*).

However, to the best of our knowledge, the effect of introgressed Neanderthal alleles within the cell line model of drug response has not been studied. In this study, we demonstrate the influence of individual Neanderthal SNPs and cumulative Neanderthal allele dosage on cytotoxic response to 29 chemotherapy drugs and 179 environmental chemicals. Our results show another association between Neanderthal alleles and a clinically relevant modern human phenotype, cytotoxic response. Thus, our results provide additional insight into how Neanderthal introgression influences modern human traits. The cumulative effect of Neanderthal alleles on cytotoxic response is striking. Our results show that an increase in Neanderthal allele dosage is associated with an increased sensitivity to anti-cancer drugs and environmental chemicals. Of course, the response to anti-cancer agents and toxic chemicals is the measurable phenotype that associates with the residual Neanderthal variance in anatomically modern humans, but we hypothesize that Homo sapiens may have evolved resistance to a diverse array of phytochemicals due to a more heterogeneous diet relative to the predominantly meat-based diet of Neanderthals. Previous studies have shown the influence of race and ethnicity on the dose response phenotype (*Shord et al., 2006*; *Sharma, Cogswell & Li, 2008*; *Fu et al., 2014*; *Abdo et al., 2015*; *Jack et al., 2015*). In this study, we show a strong association between Neanderthal ancestry and the dose response phenotype. In a previous study consisting of patients with similar access to health care resources and treatment with adjuvant 5-fluorouracil for colon cancer, it was shown that individuals with Caucasian ancestry experienced higher treatment-related toxicity rates than individuals with African-American ancestry (*McCollum et al., 2002*). In another study from our group analyzing associations between race and variations in drug response to Paclitaxel in breast cancer patients (manuscript in preparation), we have found that at higher drug concentrations, cell lines from white individuals show a significant increase in sensitivity compared to cell lines from black individuals. This corroborates our results of increased sensitivity to cytotoxic chemicals with increased Neanderthal allele dosage, because the Caucasian genome contains a higher proportion of Neanderthal DNA than any other population, whereas the African genome contains very little to no Neanderthal DNA (*Sankararaman et al., 2012*; *Wall et al., 2013*; *Vernot & Akey, 2014*). Thus, the results from this study are consistent with the two studies described above, with all of them having the same direction of effect. Interestingly, the results of the *Abdo et al. (2015)* study showed that overall, Caucasian populations were less sensitive to environmental chemicals compared to the other global populations (the overall distribution of EC_{10} values for Caucasian populations is higher than for other populations). This reinforces our results in that while Caucasian populations have the overall highest Neanderthal allele dosage, our results are not confounded by the overall relative sensitivity of Caucasian populations.

While the cumulative effect of Neanderthal allele dosage is remarkable, there are some limitations that must be noted. In order to translate these results into a clinical setting, additional follow-up studies are needed, both related to understanding the biology of the result and to the translatability of the finding to patients. siRNA knock-down experiments, gene expression profiling assays and/or CRISPR/Cas9 gene editing techniques could be performed to validate the functional effects of Neanderthal alleles on cytotoxic response. The dose response profile data used in this study was assayed in LCLs, which have inherent limitations that must be addressed for the results of this study to be clinically relevant. Dose response is a complex trait and is influenced by various criteria such as absorption, distribution, metabolism, excretion and toxicity (ADMET). The LCL dose response assay is a measure of cellular toxicity which is only one of the many components of dose response. Thus, the LCL dose response does not completely represent an individual's complex response to a chemical or drug. This limitation could be addressed by replicating this study in data from a clinical trial on cancer patients undergoing chemotherapy treatment to examine if Neanderthal alleles have an analogous effect on clinical outcome. In this study, we examined the influence of Neanderthal ancestry on the dose response phenotype; it would be interesting to inspect if Denisovan ancestry also has a similar influence on cytotoxic response.

CONCLUSION

This study establishes that Neanderthal ancestry-informative markers (AIMs) do have an influence on cytotoxic response, which is a clinically important trait. We have identified Neanderthal-derived SNPs that could be associated with cytotoxic response and showed that an increase in Neanderthal allele dosage leads to increased sensitivity to anti-cancer drugs and environmental chemicals. The results from this study shed light on another phenotype influenced by Neanderthal DNA. The findings in this study provide another dimension of the genetic factors that influence how an individual or a population will respond to a drug or a chemical, which is a key goal of pharmacogenomics and toxicogenomics.

ACKNOWLEDGEMENTS

We would like to thank Nour Abdo, Menghang Xia, Chad C. Brown, Oksana Kosyk, Ruili Huang, Srilatha Sakamuru, Yi-Hui Zhou, Paul Gallins, Kai Xia, Yun Li, Weihsueh A. Chiu, Christopher P. Austin, Raymond R. Tice, Ivan Rusyn, Fred A. Wright, Marisa W. Medina, Ronald M. Krauss and John A. Capra for sharing the results of their study for this work.

Funding

This work was supported by NIH NCI RO1CA161608 from the National Cancer Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
NIH NCI: RO1CA161608.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Farida S. Akhtari and John R. Jack analyzed the data, contributed reagents/materials/-analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Tammy M. Havener and performed the experiments, contributed reagents/materials/-analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Masahide Fukudo performed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper.
- Howard L. McLeod conceived and designed the experiments, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Alison A. Motsinger-Reif conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The research in this article did not generate any data or code. All data used in the study had been previously collected. All data are de-identified, and IRB exceptions have been filed.

The genomic/genotype data used for the chemical response is publicly available through the 1000 Genomes Project, and can be found here: http://www.internationalgenome.org. The genomic/genotype data used for the chemotherapy response can be found on dbGAP: Cholesterol and Pharmacogenetics (CAP) Study.

dbGaP Study Accession: phs000481.v2.p1.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5691#supplemental-information.

REFERENCES

Abdo N, Xia M, Brown CC, Kosyk O, Huang R, Sakamuru S, Zhou Y-H, Jack JR, Gallins P, Xia K, Li Y, Chiu WA, Motsinger-Reif AA, Austin CP, Tice RR, Rusyn I, Wright FA. 2015. Population-based in vitro hazard and concentration—response assessment of chemicals: the 1000 genomes high-throughput screening study. *Environmental Health Perspectives* **123**:458 DOI 10.1289/EHP.1408775.

Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, Gragert L, Babrzadeh F, Gharizadeh B, Luo M, Plummer FA, Kimani J, Carrington M, Middleton D, Rajalingam R, Beksac M, Marsh SGE, Maiers M, Guethlein LA, Tavoularis S, Little A-M, Green RE, Norman PJ, Parham P. 2011. The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science* **334**:89–94 DOI 10.1126/science.1209202.

Benetatos L, Hatzimichael E, Londin E, Vartholomatos G, Loher P, Rigoutsos I, Briasoulis E. 2013. The microRNAs within the DLK1-DIO3 genomic region: involvement in disease pathogenesis. *Cellular and Molecular Life Sciences* **70**:795–814 DOI 10.1007/s00018-012-1080-8.

Brown CC, Havener TM, Medina MW, Auman JT, Mangravite LM, Krauss RM, McLeod HL, Motsinger-Reif AA. 2012a. A genome-wide association analysis of temozolomide response using lymphoblastoid cell lines shows a clinically relevant association with MGMT. *Pharmacogenetics and Genomics* **22**:796–802 DOI 10.1097/FPC.0b013e3283589c50.

Brown CC, Havener TM, Medina MW, Jack JR, Ronald M, Mcleod HL, Motsinger-reif AA. 2014. Genome-wide association and pharmacological profiling of 29 anticancer agents using lymphoblastoid cell lines. *Pharmacogenomics* **15**:137–146 DOI 10.2217/pgs.13.213.

Brown CC, Havener TM, Medina MW, Krauss RM, McLeod HL, Motsinger-Reif AA. 2012b. Multivariate methods and software for association mapping in dose-response genome-wide association studies. *BioData Mining* **5**:Article 21 DOI 10.1186/1756-0381-5-21.

Dannemann M, Andrés AM, Kelso J. 2016. Introgression of neandertal- and denisovan-like haplotypes contributes to adaptive variation in human toll-like receptors. *American Journal of Human Genetics* **98**:22–33 DOI 10.1016/j.ajhg.2015.11.015.

Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, White R. 1990. Centre d'etude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. *Genomics* **6**:575–577 DOI 10.1016/0888-7543(90)90491-C.

Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, Field JR, Pulley JM, Ramirez AH, Bowton E, Basford MA, Carrell DS, Peissig PL, Kho AN, Pacheco JA, Rasmussen LV, Crosslin DR, Crane PK, Pathak J, Bielinski SJ, Pendergrass SA, Xu H, Hindorff LA, Li R, Manolio TA, Chute CG, Chisholm RL, Larson EB, Jarvik GP, Brilliant MH, McCarty CA, Kullo IJ, Haines JL, Crawford DC, Masys DR, Roden DM. 2013. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nature Biotechnology* **31**:1102–1111 DOI 10.1038/nbt.2749.

Fu S, Wu J, Li Y, Liu Y, Gao Y, Yao F, Qiu C, Song L, Wu Y, Liao Y, Sun D. 2014. Urinary arsenic metabolism in a Western Chinese population exposed to high-dose inorganic arsenic in drinking water: influence of ethnicity and genetic polymorphisms. *Toxicology and Applied Pharmacology* **274**:117–123 DOI 10.1016/j.taap.2013.11.004.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH-Y, Hansen NF, Durand EY, Malaspinas A-S, Jensen JD,

Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Ž, Gušic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, De la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Pääbo S. 2010. A draft sequence of the neandertal genome. *Science* **328**:710–722 DOI 10.1126/science.1188021.

Higham T, Douka K, Wood R, Ramsey CB, Brock F, Basell L, Camps M, Arrizabalaga A, Baena J, Barroso-Ruíz C, Bergman C, Boitard C, Boscato P, Caparrós M, Conard NJ, Draily C, Froment A, Galván B, Gambassini P, Garcia-Moreno A, Grimaldi S, Haesaerts P, Holt B, Iriarte-Chiapusso M-J, Jelinek A, Jordá Pardo JF, Maíllo-Fernández J-M, Marom A, Maroto J, Menéndez M, Metz L, Morin E, Moroni A, Negrino F, Panagopoulou E, Peresani M, Pirson S, De la Rasilla M, Riel-Salvatore J, Ronchitelli A, Santamaria D, Semal P, Slimak L, Soler J, Soler N, Villaluenga A, Pinhasi R, Jacobi R. 2014. The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* **512**:306–309 DOI 10.1038/nature13621.

Huang RS, Duan SW, Kistner EO, Hartford CM, Dolan ME. 2008. Genetic variants associated with carboplatin-induced cytotoxicity in cell lines derived from Africans. *Molecular Cancer Therapeutics* **7**:3038–3046 DOI 10.1158/1535-7163.Mct-08-0248.

Huerta-Sánchez E, Jin X, Asan, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, Ni P, Wang B, Ou X, Huasang, Luosang J, Cuo ZXP, Li K, Gao G, Yin Y, Wang W, Zhang X, Xu X, Yang H, Li Y, Wang J, Wang J, Nielsen R. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**:194–197 DOI 10.1038/nature13408.

International HapMap Consortium. 2003. The International HapMap Project. *Nature* **426**:789–796 DOI 10.1038/nature02168.

Jack J, Havener TM, McLeod HL, Motsinger-Reif AA, Foster M. 2015. Evaluating the role of admixture in cancer therapy via *in vitro* drug response and multivariate genome-wide associations. *Pharmacogenomics* **16**:1451–1463 DOI 10.2217/PGS.15.85.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson

DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, De la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921 DOI 10.1038/35057062.

Luk JM, Burchard J, Zhang C, Liu AM, Wong KF, Shek FH, Lee NP, Fan ST, Poon RT, Ivanovska I, Philippar U, Cleary MA, Buser CA, Shaw PM, Lee C-N, Tenen DG, Dai H, Mao M. 2011. DLK1-DIO3 genomic imprinted microRNA cluster at 14q32.2 defines a stemlike subtype of hepatocellular carcinoma associated with poor survival. *The Journal of Biological Chemistry* **286**:30706–30713 DOI 10.1074/jbc.M111.229831.

McCarthy JJ, McLeod HL, Ginsburg GS. 2013. Genomic medicine: a decade of successes, challenges, and opportunities. *Science Translational Medicine* **5**:189sr4 DOI 10.1126/scitranslmed.3005785.

McCollum AD, Catalano PJ, Haller DG, Mayer RJ, Macdonald JS, Benson AB, Fuchs CS. 2002. Outcomes and toxicity in african-american and caucasian patients in a randomized adjuvant chemotherapy trial for colon cancer. *Journal of the National Cancer Institute* **94**:1160–1167 DOI 10.1093/JNCI/94.15.1160.

Motsinger-Reif AA, Jorgenson E, Relling MV, Kroetz DL, Weinshilboum R, Cox NJ, Roden DM. 2013. Genome-wide association studies in pharmacogenomics: successes and lessons. *Pharmacogenetics and Genomics* **23**:383–394 DOI 10.1097/FPC.0b013e32833d7b45.

Ortega VE, Meyers DA. 2014. Pharmacogenetics: implications of race and ethnicity on defining genetic profiles for personalized medicine. *The Journal of Allergy and Clinical Immunology* **133**:16–26 DOI 10.1016/j.jaci.2013.10.040.

Peters EJ, Motsinger-Reif A, Havener TM, Everitt L, Hardison NE, Watson VG, Wagner M, Richards KL, Province MA, McLeod HL. 2011. Pharmacogenomic characterization of US FDA-approved cytotoxic drugs. *Pharmacogenomics* **12**:1407–1415 DOI 10.2217/pgs.11.92.

Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, De Filippo C, Li H, Mallick S, Dannemann M, Fu Q,

Kircher M, Kuhlwilm M, Lachmann M, Meyer M, Ongyerth M, Siebauer M, Theunert C, Tandon A, Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green RE, Hellmann I, Johnson PLF, Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES, Bakken TE, Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M, Reich D, Kelso J, Pääbo S. 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**:43–49 DOI 10.1038/nature12886.

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**:2336–2337 DOI 10.1093/bioinformatics/btq419.

Purcell S. PLINK (v1.07). *Available at http://pngu.mgh.harvard.edu/purcell/plink/*.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**:559–575 DOI 10.1086/519795.

R Development Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. *Available at http://www.r-project.org/*.

Ritchie MD. 2012. The success of pharmacogenomics in moving genetic association studies from bench to bedside: study design and implementation of precision medicine in the post-GWAS era. *Human Genetics* **131**:1615–1626 DOI 10.1007/s00439-012-1221-z.

Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N, Reich D. 2014. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**:354–357 DOI 10.1038/nature12961.

Sankararaman S, Patterson N, Li H, Pääbo S, Reich D. 2012. The date of interbreeding between neandertals and modern humans. *PLOS Genetics* **8**:e1002947 DOI 10.1371/journal.pgen.1002947.

Sharma AJ, Cogswell ME, Li R. 2008. Dose-response associations between maternal smoking during pregnancy and subsequent childhood obesity: effect modification by maternal race/ethnicity in a low-income US cohort. *American Journal of Epidemiology* **168**:995–1007 DOI 10.1093/aje/kwn223.

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research* **29**:308–311 DOI 10.1093/NAR/29.1.308.

Shord SS, Thompson DM, Krempl GA, Hanigan MH. 2006. Effect of concurrent medications on cisplatin-induced nephrotoxicity in patients with head and neck cancer. *Anti-Cancer Drugs* **17**:207–215 DOI 10.1097/00001813-200602000-00013.

Simonti CN, Vernot B, Bastarache L, Bottinger E, Carrell DS, Chisholm RL, Crosslin DR, Hebbring SJ, Jarvik GP, Kullo IJ, Li R, Pathak J, Prato JD, Bush WS, Akey JM, Denny JC, Capra JA. 2016. The phenotypic legacy of admixture between modern humans and Neandertals. *Science* **351**:737–741 DOI 10.1126/science.aad2149.

Stark AL, Zhang W, Mi S, Duan S, O'Donnell PH, Huang RS, Dolan ME. 2010. Heritable and non-genetic factors as variables of pharmacologic phenotypes in lymphoblastoid cell lines. *Pharmacogenomics Journal* **10**:505–512 DOI 10.1038/tpj.2010.3.

The 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature* **526**:68–74 DOI 10.1038/nature15393.

The GTEx Consortium. 2013. The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**:580–585 DOI 10.1038/ng.2653.

Valleron W, Laprevotte E, Gautier E-F, Quelen C, Demur C, Delabesse E, Agirre X, Pró Sper F, Kiss T, Brousset P. 2012. Specific small nucleolar RNA expression profiles in acute leukemia. *Leukemia* **26**:2052–2060 DOI 10.1038/leu.2012.111.

Vernot B, Akey JM. 2014. Resurrecting surviving neandertal lineages from modern human genomes. *Science* **343**:1017–1021 DOI 10.1126/science.1245938.

Wall JD, Yang MA, Jay F, Kim SK, Durand EY, Stevison LS, Gignoux C, Woerner A, Hammer MF, Slatkin M. 2013. Higher levels of Neanderthal ancestry in east Asians than in Europeans. *Genetics* **194(1)**:199–209 DOI 10.1534/genetics.112.148213.

Watson VG, Motsinger-Reif A, Hardison NE, Peters EJ, Havener TM, Everitt L, Auman JT, Comins DL, McLeod HL. 2011. Identification and replication of loci involved in camptothecin-induced cytotoxicity using CEPH pedigrees. *PLOS ONE* **6**:e17561 DOI 10.1371/journal.pone.0017561.

Wheeler HE, Dolan ME. 2012. Lymphoblastoid cell lines in pharmacogenomic discovery and clinical translation. *Pharmacogenomics* **13**:55–70 DOI 10.2217/pgs.11.121.

Xu G, Yang F, Ding C-L, Zhao L-J, Ren H, Zhao P, Wang W, Qi Z-T. 2014. Small nucleolar RNA 113—1 suppresses tumorigenesis in hepatocellular carcinoma. *Molecular Cancer* **13**:Article 216 DOI 10.1186/1476-4598-13-216.

Ye DZ, Field J. 2012. PAK signaling in cancer. *Cellular Logistics* **2**:105–116 DOI 10.4161/cl.21882.

Ziliak D, O'donnell PH, Im HK, Gamazon ER, Chen P, Delaney S, Shukla S, Das S, Cox NJ, Vokes EE, Cohen EEW, Dolan ME, Huang RS. 2011. Germline polymorphisms discovered via a cell-based, genome-wide approach predict platinum response in head and neck cancers. *Translational Research* **157**:265–272 DOI 10.1016/j.trsl.2011.01.005.