Drug Metabolizing Enzyme and Transporter Gene Variation, Nicotine Metabolism, Prospective Abstinence, and Cigarette Consumption

Andrew W. Bergen1,*, Martha Michel2, Denise Nishita1, Ruth Krasnow1, Harold S. Javitz1, Karen N. Conneely3, Christina N. Lessov-Schlaggar4, Hyman Hops5, Andy Z. X. Zhu6, James W. Baurley7, Jennifer B. McClure8, Sharon M. Hall9, Timothy B. Baker10, David V. Conti11, Neal L. Benowitz12, Caryn Lerman13, Rachel F. Tyndale14, Gary E. Swan5,15, Transdisciplinary Research in Cancer of the Lung Research Team16

1 Center for Health Sciences, SRI International, Menlo Park, California, United States of America, 2 Academic Research Systems, University of California San Francisco, San Francisco, California, United States of America, 3 Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, United States of America, 4 Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, United States of America, 5 Oregon Research Institute, Eugene, Oregon, United States of America, 6 Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada, 7 BioRealm, LLC, Monument, Colorado, United States of America, 8 Group Health Research Institute, Seattle, Washington, United States of America, 9 Department of Psychiatry, University of California San Francisco, San Francisco, California, United States of America, 10 Center for Tobacco Research and Intervention, Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States of America, 11 Department of Preventive Medicine, University of Southern California, Los Angeles, California, United States of America, 12 Departments of Medicine and of Bioengineering & Therapeutic Sciences, University of California San Francisco, San Francisco, California, United States of America, 13 Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, 14 Cambell Family Mental Health Research Institute, Centre for Addiction and Mental Health, and Departments of Psychiatry, and of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada, 15 Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Palo Alto, California, United States of America

†Membership of the Transdisciplinary Research In Cancer of the Lung (TRICL) Research Team can be found in the Acknowledgments.

* andrew.bergen@sri.com

Abstract

The Nicotine Metabolite Ratio (NMR, ratio of trans-3'-hydroxycotinine and cotinine), has previously been associated with CYP2A6 activity, response to smoking cessation treatments, and cigarette consumption. We searched for drug metabolizing enzyme and transporter (DMET) gene variation associated with the NMR and prospective abstinence in 2,946 participants of laboratory studies of nicotine metabolism and of clinical trials of smoking cessation therapies. Stage I was a meta-analysis of the association of 507 common single nucleotide polymorphisms (SNPs) at 173 DMET genes with the NMR in 449 participants of laboratory studies. Nominally significant associations were identified in ten genes after adjustment for intragenic SNPs: CYP2A6 and two CYP2A6 SNPs attained experiment-wide significance adjusted for correlated SNPs (CYP2A6 $P_{ACT}$=4.1E-7, rs4803381 $P_{ACT}$=4.5E-5, rs1137115, $P_{ACT}$=1.2E-3). Stage II was mega-regression analyses of 10
Pharmacogenetics of Nicotine Addiction Treatment Consortium. Therefore I will not submit individual level data from Stage II of this analysis to dbGaP as the remaining data from Stage II (10 SNPs in three RCTs) would be an incomplete dataset (representing 0% of the RCT data with baseline NMR, and 32% of the RCT data with baseline nicotine dependence and prospective abstinence measures). Readers may contact Andrew W Bergen PhD to request data from the three remaining RCTs.

**Funding:** This work was supported, in part, by: a Research Agreement between Medco Health Solutions, Inc., Affymetrix, Inc., and SRI International; grants and a contract from the National Institute of Drug Abuse, and grants from the National Cancer Institute, of the United States Department of Health and Human Services [R21 DA33813 to AWB; NIH/ NIDA Contract No. HHSN271201300004C to JWB; R01 DA03706 to HH; R01s DA16752, DA18691, DA15732 and DA9253 to SH; P50 CA84724, K05 CA139871 and P50 DA19706 to TBB; R01s CA71358 and DA11170 to GES; and PGRN U01 DA20830 to NLB (2005-2010), and to RFT and CL (2010-2015)]; a grant from the University of California Tobacco Related Diseases Research Program (TPT2004 to NLB); by the Centre for Addiction and Mental Health (RFT); and by an endowed Chair in Addictions (RFT). Clinical trials (NCT Trial IDs) 00301145, 01621009, 01621022 and 00332644 were funded by the National Cancer Institute, 00087880 and 00086385 were funded by the National Institute of Drug Abuse, and 00326781 and 00322205 were funded by the National Cancer Institute and by the National Institute of Drug Abuse. Varenicline and nominal support for recruiting clinical trial 00301145 participants was provided by Pfizer. GlaxoSmithKline provided medication for clinical trials 01621022 and 00332644. Medco Health Solutions, Inc. and Affymetrix, Inc. participated with SRI International in the selection of the two clinical laboratory metabolic studies, of the metabolic outcome, and of the DMET Plus Array for genotyping and Affymetrix performed the DMET Plus Array genotyping of the two clinical laboratory metabolic study DNA samples. No other funders or sponsors had any other role in the design of the laboratory studies, the randomized controlled trials or of this two-stage study, the collection, analysis, or interpretation of data, the writing of, or the decision to submit the ms. BioRealm, LLC provided support in the form of a salary for author JWB, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of the authors are articulated in the ‘author contributions’ section.

DMET SNPs with pretreatment NMR and prospective abstinence in up to 2,497 participants from eight trials. rs4803381 and rs1137115 SNPs were associated with pretreatment NMR at genome-wide significance. In post-hoc analyses of CYP2A6 SNPs, we observed nominally significant association with: abstinence in one pharmacotherapy arm; cigarette consumption among all trial participants; and lung cancer in four case:control studies. CYP2A6 minor alleles were associated with reduced NMR, CPD, and lung cancer risk. We confirmed the major role that CYP2A6 plays in nicotine metabolism, and made novel findings with respect to genome-wide significance and associations with CPD, abstinence and lung cancer risk. Additional multivariate analyses with patient variables and genetic modeling will improve prediction of nicotine metabolism, disease risk and smoking cessation treatment prognosis.

**Introduction**

The predominant enzyme involved in nicotine and cotinine metabolism is CYP2A6 [1, 2], where nicotine is metabolized primarily to cotinine, and cotinine primarily to trans-3′-hydroxycotinine. The Nicotine Metabolite Ratio (NMR) is the ratio of the metabolites trans-3′-hydroxycotinine and cotinine. The ratio reflects the enzymatic activity of CYP2A6 and is a biomarker of the rate of nicotine clearance [3]. The ratio is associated with smoking topography [4], responsiveness to smoking cessation treatments [5], and the number of cigarettes smoked per day (CPD) [6]. Significant correlations between the NMR and a) oral clearance of nicotine, b) oral clearance and t_{1/2} of cotinine, and c) lack of production of trans-3′-hydroxycotinine in individuals homozygous for null CYP2A6 alleles, support the validity of the NMR as a marker of CYP2A6 activity [3]. In addition to CYP2A6, genetic and biochemical studies have identified contributions of additional DMET loci to measures of nicotine metabolism in diverse samples and study designs. Gene variants associated with differences in nicotine metabolism [7] influence smoking behavior, tobacco exposures and attributable disease risks, and could serve as biomarkers for disease risk, and treatment prognosis [8].

To discover and develop novel biomarkers of nicotine metabolism and related phenotypes, we performed a planned analysis to: a) identify DMET SNPs associated with the laboratory study-based NMR [9, 10] using the DMET Plus Array [11]; b) validate nominally significant DMET SNP associations with baseline NMR in clinical trial participants from two trials [12]; and, c) further validate nominally significant DMET SNPs with end-of-treatment and six-month seven day point prevalence abstinence (abstinence) in clinical trial participants from eight trials [12–18]. We also conducted post-hoc analyses of validated DMET SNPs in clinical trial participants with abstinence by smoking cessation pharmacotherapy assignment, and with nicotine dependence measures among all participants. Finally, we interrogated a meta-GWAS lung cancer database [19] for association results of two validated DMET SNPs.

**Methods**

**Human Subjects**

Institutional Review Board approval for each study, and informed written consent from each participant, was obtained by the Principal Investigators of each study. Institutional Review Board approval for these analyses was obtained from the Committee on Human Research at the University of California San Francisco and the Human Subjects Committee at SRI.
Competing Interests: AWB reports employment at SRI International. This project was partly funded by Medco Health Solutions, Inc. (acquired by Express Scripts in 2012) and Affymetrix, Inc. JWB is an employee of Riptide Bio. Varenicline and nominal support for recruiting clinical trial 00301145 participants was provided by Pfizer. GlaxoSmithKline provided medication for clinical trials 01621022 and 00332644. GES reports serving as a Pfizer consultant for a one-day meeting in 2008. SMH has served as a consultant to Gilead, Inc., and Sleepio, Inc. TBB reports no competing interests in the last five years. He discloses that in the five years prior to the trials included in this analysis and that he conducted (NCT IDs 01621009, 01621022 and 00332644), he served as a PI or Co-PI on research sponsored by pharmaceutical companies, including GlaxoSmithKline and Pfizer. He also discloses that he is an author and collaborator on work done by Washington University School of Medicine colleagues LS Chen, AJ Bloom, & LJ Bierut on work examining associations of CYP2A6 with cessation. NLB reports having received research funding from pharmaceutical companies, including GlaxoSmithKline and Pfizer. He also discloses that he served as a consultant to Gilead, Inc., and Medco Health Solutions, Inc. (acquired by Express Scripts in 2012) and Affymetrix, Inc. JWB is an employee and a LLC member of BioRealm. International, and for the TRICL study by the institutional ethics review committees at the involved institutions.

Stage I Participants

Stage I participants were derived from two distinct studies (Table 1). Healthy twin pairs and siblings were recruited from the Northern California Twin Registry, age ≥18 and <65 years, <130% of height-adjusted ideal weight, neither pregnant nor intending to become pregnant, and after exclusion of chronic medical and psychiatric conditions (PKTWIN study) [9]. Participants underwent a 30 minute infusion of deuterium-labeled nicotine (3',3'-dideuteronicotine) and cotinine (2',4',5',6'-tetradeuterocotinine), monitoring, and blood and urine collection in a hospital setting [9]. Probands and two first-degree relatives from 158 pedigrees with ≥three ever-smoking individuals per pedigree were recruited to assess the relations between genetic factors, environmental factors and tobacco use (SMOFAM study) [10]. Individuals from 61 pedigrees completed a clinical study of nicotine metabolism; oral administration of a fixed dose of deuterium-labeled nicotine and cotinine at home, monitored by a nurse, was followed by collection, aliquoting and freezing of saliva samples at multiple time points [10]. Participants completed a detailed questionnaire and provided a blood sample for DNA extraction and analysis [10]. Levels of nicotine, cotinine trans-3'-hydroxycotinine and their glucuronides were estimated via liquid chromatography-mass spectrometry and pharmacokinetic phenotypes calculated, as described [5], from each study.

Stage I Analysis

Genomic DNA was extracted from whole blood [20], and quantified [21]. DNA samples were genotyped using the Affymetrix DMET Plus Array [11] (S1 File). Genotype and phenotype data are accessible through application to a data access committee using dbGaP Study ID phs000931.v1.p1. We performed association analyses with common biallelic SNPs and indels [≥0.05 minor allele frequency (MAF) in each dataset] with the NMR derived from the six hour biospecimen collection. In the PKTWIN dataset, the NMR was square-root transformed and adjusted for age, age-squared, BMI, sex, smoking status (current versus other), and any hormone use (menopausal status and/or reproductive hormone use versus no hormone use). In the SMOFAM dataset, the NMR was log transformed and adjusted for age, aged-squared, BMI, sex, and smoking status. We constructed principal components of population genetic variation using 655 unlinked (r²<0.5) DMET Plus Array SNPs in 323 PKTWIN individuals and 212 SMOFAM individuals, and used the first ten principal components to further adjust the NMR in each dataset. To avoid heterogeneity due to potential differential linkage disequilibrium among different continental population samples, genotype-phenotype analyses presented were performed on individuals who self-identified as White. Additive models were evaluated for each common SNP via Hierarchical Linear Modeling (HLM) [22] as our primary analysis (S1 File). We performed a sample-size weighted fixed-effect meta-analysis followed by PACT adjustment using test statistics from the HLM analyses. The meta-PACT procedure [23] assesses meta-analysis significance at two levels: adjusted for all SNPs tested within a gene, and within the entire experiment, accounting for intragenic SNP correlation.

Stage II Participants

We used DNA samples and clinical data from self-identified White participants from eight clinical trials of smoking cessation therapies conducted in six US sites [12–18] to validate nominally significant DMET SNPs from Stage I with pretreatment NMR and prospective abstinence (Table 2). Treatment-seeking smokers (≥10 CPD, ≥18 years of age, recruited using local
media and screened for eligibility), were randomized to placebo or active pharmacotherapy [12, 14–18], or were prescribed varenicline and randomized to three modes of behavioral therapy [13]; all participants were provided with multiple sessions of supportive therapy (individual, group, telephone, internet or combined modes). Trial participants provided a blood [12, 14–18] or a saliva [13] sample on study entry; blood was used for pretreatment NMR determination from two clinical trials [12] and blood or saliva for DNA extraction from eight trials [12–18].

### Stage II Analysis

For validation, we chose one or more SNPs with uncorrected meta-analysis \( p \)-values < 0.05 from each of the ten genes with meta-\( P_{ACT} \) \( p \)-values < 0.05 and evaluated availability of TaqMan SNP Genotyping Assays (Life Technologies). Of 16 selected SNPs, 12 SNPs had pre-designed assays, two were proxy SNPs (\( r^2 = 1 \), Utah residents with ancestry from northern and western Europe HapMap sample, CEU) with predesigned assays, and two had custom assays designed (S1 File and S4 Table). We performed genotyping on the TaqMan OpenArray SNP Genotyping System or ViiA 7 using laboratory study, clinical trial, and HapMap (Coriell Cell Repositories, Camden, NJ) DNA samples. Stage II SNP genotyping assays were evaluated for clustering, concordance and completion rates prior to removing DNA samples with low completion rates, then for concordance, completion and Hardy Weinberg Equilibrium (HWE). Quality control procedures and summary results are described in S1 File, and completion rate

<table>
<thead>
<tr>
<th>Dataset</th>
<th>PKTWIN</th>
<th>SMOFAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>N individuals</td>
<td>247</td>
<td>202</td>
</tr>
<tr>
<td>N families</td>
<td>120</td>
<td>59</td>
</tr>
<tr>
<td>Age, Years, mean (SD)</td>
<td>Overall</td>
<td>247 (37.5, 12.7)</td>
</tr>
<tr>
<td></td>
<td>Twins</td>
<td>235 (37.2, 12.6)</td>
</tr>
<tr>
<td></td>
<td>Siblings</td>
<td>12 (43.7, 13.2)</td>
</tr>
<tr>
<td></td>
<td>Offspring</td>
<td>113 (28.1, 4.6)</td>
</tr>
<tr>
<td>Sex (N, %)</td>
<td>Parents</td>
<td>178 (72.1%)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>44 (17.8%)</td>
</tr>
<tr>
<td>Smoking status (N, %)</td>
<td>Former</td>
<td>62 (25.1%)</td>
</tr>
<tr>
<td></td>
<td>Never</td>
<td>141 (57.1%)</td>
</tr>
<tr>
<td>NMR (mean, SD)</td>
<td>VA, PA</td>
<td>247 (0.27, 0.14)</td>
</tr>
<tr>
<td></td>
<td>OA, SA</td>
<td>202 (0.28, 0.14)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>MZ</td>
<td>247 (25.1, 4.8)</td>
</tr>
<tr>
<td>Zygosity</td>
<td>Pre-menopausal</td>
<td>130 (77.4%)</td>
</tr>
<tr>
<td></td>
<td>During menopause</td>
<td>18 (10.7%)</td>
</tr>
<tr>
<td></td>
<td>Post-menopausal</td>
<td>20 (11.9%)</td>
</tr>
<tr>
<td>Hormone use</td>
<td>Oral Contraceptive</td>
<td>40 (25.8%)</td>
</tr>
<tr>
<td></td>
<td>HRT</td>
<td>16 (10.3%)</td>
</tr>
</tbody>
</table>

\( a^p<0.001 \) for \( t \), \( \chi^2 \), and \( t \) tests, respectively.  
\( b^v\text{enous administration, plasma analysis.} \)  
\( ^c\text{oral administration, saliva analysis. } \)  
\( ^d\text{hormone replacement therapy.} \)

DOI:10.1371/journal.pone.0126113.t001
and HWE significance by assay and by clinical trial are described in S5 and S6 Tables. Six of 16 selected DMET SNPs failed genotype clustering, completion rate and HWE quality thresholds. Thus, ten SNPs from seven of 10 genes passed quality thresholds and were included in subsequent analyses. In the course of evaluating the specificity of custom assays for one CYP2A6 SNP (rs4803381), we resolved an ambiguous genomic location, i.e., evidence for rs4803381’s location proximal to CYP2A6 rather than proximal to CYP2A7 (S1 File).

We used linear regression to estimate association of ten DMET SNPs on natural log transformed NMR in treatment-seeking smokers from two trials [12], adjusted for age, age-squared, BMI, sex, the first three principal components of population genetic variation obtained from analysis of 45 ancestry informative markers [24], and site. We used logistic regression to estimate association between ten DMET SNPs on end of treatment and on six month abstinence in treatment-seeking smokers from eight trials [12–18], adjusting for age, age-squared, BMI, sex, education (college degree or less than a college degree), marital status (married or other), principal components, and 26 treatment arms. For both analyses, we imputed missing values for BMI (N = 43), education (N = 10) and marital status (N = 7) 20 times, analyses were performed on each data set, and results combined with adjustment to reflect the variance attributable to the imputations [25].

**Post-hoc analyses**

For two CYP2A6 SNPs (rs4803381 and rs1137115) we estimated the proportions of variance and significance of association with: cigarettes per day (CPD), coded as a continuous variable and as in the Fagerström Test for Nicotine Dependence (FTND) [26] or Cigarette Dependence (FTCD) [27]; time to first cigarette after waking (TTFC); and total FTND/FTCD...
score. Model covariates in nicotine dependence regressions were age, age squared, BMI, sex, education, marital status, three principal components of population genetic variation, and site. We performed regression analyses to evaluate the influence of rs4803381 on prospective abstinence outcomes stratified by pharmacotherapy. Model covariates in abstinence regressions were the same as in the nicotine dependence regressions with the addition of clinical trial arm. Power analyses of SNP associations with NMR and six month prospective abstinence were performed. The TRICL meta-GWAS database, comprised of multiple case-control studies of lung cancer [19], was interrogated for existing association results at two CYP2A6 SNPs.

Analysis software, annotation, and model parameters
We used: SAS (Cary, NC) for data curation and HWE testing; plink [28] to evaluate Mendelian transmission; GCTA [29] to estimate principal components; STATA (StataCorp, College Station, TX) to perform HLM, imputation and other regression analyses; \( P_{\text{ACT}} [30] \) and meta-\( P_{\text{ACT}} [23] \) to adjust for multiple testing; Haploview to calculate linkage disequilibrium [31]; SNAP [32] to identify proxy SNPs; Quanto [33] for power analyses; and dbSNP [34] and Affymetrix resources for SNP annotation. Chromosome coordinates are from the NCBI36/hg18 assembly [34]. SNP models were additive, tests were two-sided, and all alphas were 0.05.

Results
Stage I participants, common DMET SNPs and the NMR
PKTWIN participants are significantly (\( P < .001 \)) more likely to be female, less likely to be a current smoker, and, lower in adiposity than SMOFAM participants (Table 1). Five SMOFAM individuals were excluded from analysis, one with a NMR value more than five standard deviations above the mean, and four due to genotypes inconsistent with interview-based family relationships. We excluded 121 and 28 DMET SNPs in PKTWIN and SMOFAM, respectively, from analysis due to nominally significant deviation (\( p \)-values < 0.05) from HWE. We tested 608 DMET SNPs common in PKTWIN and 531 SNPs common in SMOFAM, and adjusted for multiple tests in genes with more than one SNP tested using the \( P_{\text{ACT}} \) procedure. We meta-analyzed 507 SNPs within 173 genes in 449 individuals from the PKTWIN and SMOFAM datasets and adjusted for multiple SNPs within a gene, and for two datasets, using the meta-\( P_{\text{ACT}} \) procedure. S1–S3 Tables contain genotype counts and HLM analyses results, meta-analysis Z scores and \( p \)-values, and meta-\( P_{\text{ACT}} \) gene-wise \( p \)-values, respectively. Ten genes exhibited gene-wise meta-\( P_{\text{ACT}} \) \( p \)-values < 0.05; CYP2A6, CYP2D6, and SPG7 were the top three ranked genes (Table 3). Forty-one SNPs had uncorrected meta-analysis \( p \)-values < 0.05 with CYP2A6 and CYP2D6 SNPs in the top three ranked SNPs. rs4803381 was the top ranked SNP with an uncorrected meta-analysis \( P = 1.33E-07 \). After multiple test correction for 507 SNPs, two CYP2A6 SNPs attained experiment-wide significance: rs4803381 (\( P_{\text{ACT}} = 4.53E-05 \)) and rs1137115 (\( P_{\text{ACT}} = .0012 \)), located in the promoter (c.-1013) and first exon (c.51) of CYP2A6. These SNPs are in strong linkage disequilibrium with each other (\( r^2 = 0.63 \) and 0.62 and \( D^* = 0.99 \) and 1.00 in PKFAM and SMOFAM), and in ex vivo hepatic tissue have been shown to be associated with significantly reduced protein and activity [35]. rs4803381 and rs1137115 are associated with many CYP2A6* star allele (*) haplotypes (’1B/’1B2, ’1B5, ’1B6, ’1B8, ’1B9, ’1B10, ’1B11, ’1D, ’1J, ’9A, ’9B, ’18C, ’24A, ’24B, ’31A, ’31B, ’35A and ’1A, ’1B14, ’1B17, ’2, ’14, ’18B, ’20, ’21, ’28A, ’28B, ’41, ’42, ’44, ’45, respectively) [36].
Stage II participants, NMR and prospective abstinence

DNA samples from 2,499 clinical trial participants (Table 2) were subject to genotyping at 16 DMET SNPs, and to quality control filtering (S1 File and S4–S6 Tables). Two CYP2A6 SNPs evaluated for association with pretreatment NMR were associated at genome-wide significance [β = -0.280, 95%CI (-0.336, -0.225), P = 1.25E-21, N = 633 individuals, and -0.240, 95%CI (-0.301, -0.178), P = 7.30E-14, N = 614 individuals]. The Stage II mean (standard deviation) completion rates for the two CYP2A6 SNPs over the eight randomized clinical trials were: rs4803381, 0.9946 (0.0045) and rs1137115, 0.9834 (0.0218). One of eight clinical trial HWE p-values for rs4803381 was <.05, while all other p-values were >.155; no clinical trial HWE p-value for rs1137115 was <.05. The two CYP2A6 SNPs are in strong linkage disequilibrium (r² = 0.58 and D' = 0.98 in the two clinical trials). No other analyzed DMET SNPs were statistically significantly associated with pretreatment NMR, including the CYP2D6 SNP rs28371725 (Table 4). In joint analysis of rs4803381 and rs1137115 with baseline NMR (N = 605 individuals), rs4803381 was experiment-wide statistically significantly associated [β = -0.209, 95%CI (-0.293, -0.125), P = 1.31E-6], while rs1137115 was not [β = -0.075, 95%CI (-0.168, 0.017), P = 0.110]. Pretreatment NMR variance accounted for by covariates alone, with rs4803381, with rs1137115, and with both SNPs were 9.2%, 21.9%, 18.5%, and 22.2%. The mean (SD) and N of the residualized transformed NMR in individuals with SNP reference, heterozygote and minor allele homozygote genotypes were: -0.947 (0.153) 263, -0.950 (0.166) 296; and -0.953 (0.161) 74 for rs4803381; and -0.955 (0.156) 351, -0.951 (0.166) 228, -0.947 (0.143) 35 for rs1137115. Coefficient signs from Stage I and II analyses matched for both CYP2A6 SNPs. No DMET SNPs were associated with prospective abstinence over all participants, at either time point (S7 Table); power to detect ORs of 1.2 to 1.4-fold was good to excellent (S8 Table).

Post-hoc analyses

rs4803381 and rs1137115 were statistically significantly associated with cigarettes per day at different levels of significance (continuous, p-values of .0001 and .0017; categorical, p-values of .0020 and .0215), but not with TTFC or with total FTND (Table 5 and S9 Table). In joint analysis of rs4803381 and rs1137115, rs4803381 remained nominally associated with CPD (continuous, P = .0279; categorical, P = .0376).

Table 3. DMET genes associated with the laboratory study-based NMR, by meta-PACT P<0.05.

<table>
<thead>
<tr>
<th>Gene</th>
<th>meta-P&lt;sub&gt;ACT&lt;/sub&gt;</th>
<th>chr:coor of transcript&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SNPs ranked by meta-P&lt;sub&gt;ACT&lt;/sub&gt; p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2A6</td>
<td>4.05E-07</td>
<td>chr19:46,041,283–46,048,192</td>
<td>rs4803381, rs1137115, rs4079369, rs8192729</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>1.03E-03</td>
<td>chr22:40,852,445–40,856,827</td>
<td>rs1080985, rs28371725, rs16947, rs1080983, rs1065852</td>
</tr>
<tr>
<td>SPG7</td>
<td>1.08E-02</td>
<td>chr16:88,102,306–88,151,165</td>
<td>rs12960, rs2292954</td>
</tr>
<tr>
<td>XDH</td>
<td>1.11E-02</td>
<td>chr2:31,410,692–31,491,115</td>
<td>rs1884725, rs2295475</td>
</tr>
<tr>
<td>CHST8</td>
<td>1.49E-02</td>
<td>chr19:38,804,701–38,956,254</td>
<td>rs1064349</td>
</tr>
<tr>
<td>CYP2A6</td>
<td>1.53E-02</td>
<td>chr2:21,175,404–21,283,997</td>
<td>rs11045819, rs2291075, rs2306283, rs4149057, rs4149056</td>
</tr>
<tr>
<td>CYP4F3</td>
<td>2.66E-02</td>
<td>chr19:15,587,029–15,601,447</td>
<td>rs1805041, rs1805042</td>
</tr>
<tr>
<td>SLC15A1</td>
<td>4.42E-02</td>
<td>chr1:39,134,057–98,202,909</td>
<td>rs2297322, rs1339067</td>
</tr>
<tr>
<td>CBR1</td>
<td>4.76E-02</td>
<td>chr2:36,364,155–36,367,332</td>
<td>rs2835265, rs1005695, rs998383, rs3787728</td>
</tr>
</tbody>
</table>

<sup>a</sup>NCBI36/hg18

doi:10.1371/journal.pone.0126113.t003
In analyses of end-of-treatment and six month abstinence among individuals randomized to one of six (end-of-treatment) and to one of nine (six months) different pharmacotherapies, we observed two results of interest (S10 Table). In the first result of interest, we observed increased abstinence in individuals randomized to nicotine replacement therapy (NRT) (S10 Table). After analyses by mode of administration, we observed increased abstinence at end-of-treatment and at six months, with a trend towards significance (p-values < .10) in individuals randomized to NRT patch [in 339 individuals, ORend-of-treatment (95%CI) \( P = 1.348 \ (0.961–1.890) \ 0.084 \) and ORsix months(95%CI) \( P = 1.386 \ (0.973–1.975) \ 0.071\]. These observations, while not statistically significant, are the most directly interpretable. We also observed a nominally statistically significant rs4803381 association with six month abstinence in one arm of 66 Table 4. DMET SNP association with NMR, Stage I (PKTWIN, SMOFAM, Meta-analysis) and Stage II (RCT), by Stage I Meta-P\( \text{ACT} \).

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNPa</th>
<th>( \beta_{PK} )</th>
<th>( SE_{PK} )</th>
<th>( P_{PK} )</th>
<th>( \beta_{SM} )</th>
<th>( SE_{SM} )</th>
<th>( P_{SM} )</th>
<th>( Z_{\text{Meta}} )</th>
<th>( P_{\text{Meta}} )</th>
<th>( \beta_{RCT} )</th>
<th>( SE_{RCT} )</th>
<th>( P_{RCT} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2A6</td>
<td>rs4803381</td>
<td>-0.054</td>
<td>0.014</td>
<td>0.000</td>
<td>-0.070</td>
<td>0.020</td>
<td>0.000</td>
<td>-5.274</td>
<td>1.3E-7</td>
<td>-0.280</td>
<td>0.028</td>
<td>1E-21</td>
</tr>
<tr>
<td>CYP2A6</td>
<td>rs1137115</td>
<td>-0.053</td>
<td>0.015</td>
<td>0.000</td>
<td>-0.068</td>
<td>0.023</td>
<td>0.003</td>
<td>-4.657</td>
<td>3.2E-6</td>
<td>-0.240</td>
<td>0.031</td>
<td>4E-14</td>
</tr>
<tr>
<td>XDH</td>
<td>rs1884725</td>
<td>0.040</td>
<td>0.014</td>
<td>0.004</td>
<td>0.014</td>
<td>0.024</td>
<td>0.569</td>
<td>2.768</td>
<td>0.0056</td>
<td>0.002</td>
<td>0.033</td>
<td>0.949</td>
</tr>
<tr>
<td>SLC1B1</td>
<td>rs11045819c</td>
<td>0.047</td>
<td>0.018</td>
<td>0.007</td>
<td>0.024</td>
<td>0.028</td>
<td>0.394</td>
<td>2.752</td>
<td>0.0059</td>
<td>-0.012</td>
<td>0.040</td>
<td>0.767</td>
</tr>
<tr>
<td>CYP4F3</td>
<td>rs1805041</td>
<td>0.007</td>
<td>0.015</td>
<td>0.644</td>
<td>0.074</td>
<td>0.021</td>
<td>0.000</td>
<td>2.456</td>
<td>0.0140</td>
<td>0.039</td>
<td>0.031</td>
<td>0.216</td>
</tr>
<tr>
<td>CYP1B1</td>
<td>rs2835265</td>
<td>0.056</td>
<td>0.022</td>
<td>0.012</td>
<td>0.022</td>
<td>0.033</td>
<td>0.497</td>
<td>2.447</td>
<td>0.0144</td>
<td>0.068</td>
<td>0.045</td>
<td>0.133</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>rs28371725</td>
<td>-0.048</td>
<td>0.019</td>
<td>0.011</td>
<td>-0.005</td>
<td>0.035</td>
<td>0.888</td>
<td>-2.299</td>
<td>0.0215</td>
<td>0.013</td>
<td>0.047</td>
<td>0.779</td>
</tr>
<tr>
<td>SLC15A1</td>
<td>rs2297322</td>
<td>-0.022</td>
<td>0.020</td>
<td>0.275</td>
<td>-0.079</td>
<td>0.031</td>
<td>0.012</td>
<td>-2.284</td>
<td>0.0224</td>
<td>0.031</td>
<td>0.044</td>
<td>0.481</td>
</tr>
<tr>
<td>CYP4F3</td>
<td>rs1805042</td>
<td>-0.015</td>
<td>0.012</td>
<td>0.229</td>
<td>-0.041</td>
<td>0.018</td>
<td>0.023</td>
<td>-2.272</td>
<td>0.0231</td>
<td>-0.020</td>
<td>0.031</td>
<td>0.520</td>
</tr>
<tr>
<td>SLC1B1</td>
<td>rs2306283</td>
<td>0.013</td>
<td>0.013</td>
<td>0.319</td>
<td>0.049</td>
<td>0.020</td>
<td>0.013</td>
<td>2.178</td>
<td>0.0294</td>
<td>0.018</td>
<td>0.030</td>
<td>0.558</td>
</tr>
</tbody>
</table>

aSee S4 Table for details on SNPs.
b\( \beta \) = coefficient and \( P \) from PKTWIN [9] (PK), SMOFAM [10] (SM), and two clinical trials [12] (RCT), respectively; \( SE \) = Standard Error; \( Z_{\text{Meta}} \) = meta-analysis Z score; and \( P_{\text{Meta}} \) = meta-analysis \( P_{\text{ACT}} \).
c,dProxy SNPs, \( r^2 = 1.0 \), CEU for laboratoryc and clinical triald analyses, respectively.

doi:10.1371/journal.pone.0126113.t004

In analyses of end-of-treatment and six month abstinence among individuals randomized to one of six (end-of-treatment) and to one of nine (six months) different pharmacotherapies, we observed two results of interest (S10 Table). In the first result of interest, we observed increased abstinence in individuals randomized to nicotine replacement therapy (NRT) (S10 Table). After analyses by mode of administration, we observed increased abstinence at end-of-treatment and at six months, with a trend towards significance (p-values < .10) in individuals randomized to NRT patch [in 339 individuals, \( OR_{\text{end-of-treatment}} \ (95\%CI) \ P = 1.348 \ (0.961–1.890) \ 0.084 \) and \( OR_{\text{six months}} \ (95\%CI) \ P = 1.386 \ (0.973–1.975) \ 0.071\]. These observations, while not statistically significant, are the most directly interpretable. We also observed a nominally statistically significant rs4803381 association with six month abstinence in one arm of 66 Table 5. Post-hoc analyses of CYP2A6 SNPs and measures of nicotine dependencea in treatment-seeking smokers.

<table>
<thead>
<tr>
<th></th>
<th>CPD (continuous)</th>
<th>CPD (FTND coding)</th>
<th>TTFC</th>
<th>FTND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>21.67 (8.95)</td>
<td>1.34 (0.77)</td>
<td>1.99 (0.90)</td>
<td>5.23 (2.18)</td>
</tr>
<tr>
<td>Min, Max</td>
<td>1, 100</td>
<td>0, 3</td>
<td>0, 3</td>
<td>0, 10</td>
</tr>
<tr>
<td>Covariatesb ( r^2 )</td>
<td>0.1064</td>
<td>0.1056</td>
<td>0.0633</td>
<td>0.0674</td>
</tr>
<tr>
<td>Covariates and rs4803381 ( r^2 )</td>
<td>0.1118</td>
<td>0.1092</td>
<td>0.0633</td>
<td>0.0678</td>
</tr>
<tr>
<td>rs4803381 ( p )-values</td>
<td>0.0001</td>
<td>0.0020</td>
<td>0.9045</td>
<td>0.2938</td>
</tr>
<tr>
<td>Covariates and rs1137115, ( r^2 )</td>
<td>0.1101</td>
<td>0.1076</td>
<td>0.0633</td>
<td>0.0678</td>
</tr>
<tr>
<td>rs1137115, ( p )-values</td>
<td>0.0017</td>
<td>0.0021</td>
<td>0.8415</td>
<td>0.2938</td>
</tr>
<tr>
<td>Covariates and both CYP2A6 SNPs, ( r^2 )</td>
<td>0.1119</td>
<td>0.1092</td>
<td>0.0633</td>
<td>0.0678</td>
</tr>
<tr>
<td>rs4803381 and rs1137115, ( p )-values</td>
<td>0.0279 0.6966</td>
<td>0.0376, 0.9522</td>
<td>0.9601, 0.8650</td>
<td>0.6966, 0.6892</td>
</tr>
</tbody>
</table>

aPairwise correlations of cigarettes per day (continuous) with categorical CPD, TTFC and FTND were 0.914, 0.400 and 0.614, of categorical CPD with TTFC and FTND were 0.379 and 0.622, and of TTFC with FTND were 0.799.
bAge, age squared, BMI, sex, education, marital status, three principal components of population genetic variation and site. Descriptive and regression analyses were performed on a sample of 2,418 individuals with called genotypes at both rs4803381 and rs1137115.

doi:10.1371/journal.pone.0126113.t005
individuals first treated with combined NRT and bupropion from baseline to 12 weeks and then randomized to chronic bupropion \( (P = 0.023) \). This result has a more speculative interpretation.

rs1137115 was nominally significantly associated with lung cancer in a meta-analysis of four studies from the TRICL database \( \text{[fixed effects odds ratio (95%CI)} P = 0.92 (0.87–0.97) \ 0.0022] \), and in two of the individual case-control studies \( \text{(S11 Table). rs4803381 does not have results in the TRICL database. The direction of effect of the rs1137115 association with lung cancer is consistent with the reduced nicotine metabolism and cigarette consumption associated with the rs1137115 minor allele we observed in our Stage I and II analyses, and with the functional effects of rs1137115 evaluated in \textit{ex vivo} hepatic tissue \text{[35, 37].}} \)

**Discussion**

**Findings**

There were five novel findings in this study, two in the \textit{a priori} phase, and three in the \textit{post-hoc} phase. \textit{CYP2A6} was associated with laboratory-based NMR at experiment-wide significance in an analysis of 173 DMET genes. rs4803381 and rs1137115 were associated with the pretreatment NMR at genome-wide significance. In \textit{post-hoc} analyses in clinical trial participants, these two \textit{CYP2A6} SNPs were shown to be associated at experiment-wide significance with baseline CPD in treatment-seeking smokers of eight clinical trials, and rs4803381 was associated with six month abstinence in individuals randomized to chronic bupropion treatment at nominal significance. In the \textit{post-hoc} meta-analysis of four case-control studies, rs1137115 was associated with lung cancer at nominal significance. The minor alleles of these \textit{CYP2A6} SNPs were associated with reduced nicotine metabolism, smoking heaviness and lung cancer risk in a mechanistically interpretable fashion (reduced nicotine metabolism rate reduces cigarette consumption reduces lung cancer risk).

**rs4803381, rs1137115, and correlates**

Our multivariate analysis of two \textit{CYP2A6} SNPs, demographics, population genetic variation and clinical trial site accounted for up to 22.2% of the variance of the pretreatment NMR with the individual SNPs accounting for approximately half of this variance. Validation of the association of \textit{CYP2A6} SNPs with the NMR confirms the role \textit{CYP2A6} plays in nicotine metabolism \text{[7, 38, 39]}. Both SNPs are common polymorphisms in continental ancestry samples, where the rs4803381 European ancestry minor allele is the major allele in Asian and African HapMap populations. The minor alleles of rs4803381 and rs1137115 are associated with reduced \textit{CYP2A6} transcript levels \text{[35, 37]}, \textit{CYP2A6} protein level \text{[35]}, and \textit{CYP2A6} activity \text{[35, 37]} and are linked with reduced function \textit{CYP2A6} alleles \text{[36]} \text{(S4 Table).}

rs4803381 and rs1137115 were significantly associated with cigarette consumption coded continuously or as in the FTND/FTCD \text{(Table 5). The influence of a single allele of rs4803381 and of rs1137115 in our multivariate regressions of treatment-seeking smokers were -0.11 and -0.098 standard deviates of continuously coded CPD, or a reduction of 0.99 and 0.88 cigarette per day (unadjusted means in S9 Table), which represent effect sizes larger than those associated with other \textit{CYP2A6} SNPs, e.g., rs1801272 (c.479T>A/p.L160H), with effect size of 0.68 CPD \( (N = 66,380, P = 1.1E-4) \) \text{[40]. rs1137115 was nominally significantly associated with lung cancer in a meta-GWAS database, adding to existing evidence that other \textit{CYP2A6} variants contribute to lung cancer risk \text{(S2 File). The functional effects of rs4803381 and rs1137115, i.e., reduced transcription, protein level and activity \text{[35, 37]} suggest that \textit{CYP2A6} variation that influences regulation may be more important than \textit{CYP2A6} non-synonymous variants that directly influence enzyme activity, e.g., rs1801272, but with lower prevalence \text{[41].}} \)
The validated CYP2A6 SNPs were not associated with the clinically relevant outcome of abstinence in all trial participants, adjusted for pharmacotherapy. However, when we performed analyses stratified by pharmacotherapy, we observed interesting effect sizes in several smaller strata, including trending associations with increased abstinence in individuals randomized to NRT patch, and nominally significant decreased abstinence in individuals randomized to chronic bupropion. The former result is consistent with reported results in individuals from two trials [18, 42] analyzed together here, with the influence of CYP2A6 activity on abstinence previously evaluated using the biochemical NMR [42], or the CYP2A6 activity model [43]. In our analysis of individuals randomized to NRT patch from two trials [18, 42], we find that a single CYP2A6 SNP is a statistically trending predictor where reduced nicotine metabolism implies increased nicotine plasma levels and increased abstinence. The latter result in individuals randomized to chronic bupropion therapy reaches our a priori nominally statistically significant threshold but is of uncertain clinical significance due to the small size of the pharmacotherapy group and multiple pharmacotherapy groups examined. A speculative interpretation might suggest that linkage disequilibrium between rs4803381 and CYP2B6 SNPs that influence expression or activity might influence six month abstinence in individuals undergoing chronic treatment with bupropion. Interaction between CYP2A6 and CYP2B6 variants in association with nicotine metabolism has been reported [44]. CYP2B6 SNPs that influence bupropion metabolism also influence abstinence in individuals at six months in individuals treated with bupropion for 10 weeks [45], and linkage disequilibrium between relevant CYP2A6 and CYP2B6 alleles is low [46]. CYP2B6 association with smoking heaviness has been reported at robust statistical significance (rs7260329, $P = 6E^{-6}$) in a genome-wide scan [40]. Mechanistic understanding of the potential association of rs4803381 and six month abstinence in individuals randomized to chronic bupropion treatment awaits replication and more data to permit testing of CYP2A6 and CYP2B6 metabolic and treatment outcome hypotheses.

Limitations

In this study, we did not identify the monooxygenase gene $FMO3$ in our Stage I meta-analysis; $FMO3$ haplotypes were previously associated with the ratio of cotinine to the sum of nicotine and cotinine in a laboratory study [47], with CPD in dependent smokers [47], and with the NMR among slow metabolizers [48]. Future work should include haplotype and diplotype analyses to evaluate the joint influence of multiple variants. Limitations of DMET Plus array gene coverage contributed to the inability to identify some DMET gene variants previously associated with nicotine metabolism. Finally, cytochrome P450 SNPs were over-represented among DMET SNPs that failed quality control thresholds in Stage II of our analysis, suggesting that methods that more effectively interrogate cytochrome P450 loci variation may enable further analyses of NMR variance and related phenotypes at these loci. Significant CYP2A6 SNP associations with CPD, with prospective abstinence by pharmacotherapy, and with lung cancer were post-hoc and were not replicated in this study. In addition, findings from analyses of treatment-seeking cigarette smokers may not generalize to all tobacco product smokers.

Conclusions

In this two-stage scan of DMET gene variation association with the NMR, we utilized two of the largest laboratory-based studies of nicotine metabolism to nominate DMET genes and SNPs, and eight randomized controlled trials of smoking cessation therapies and a lung cancer meta-GWAS database, to validate novel SNP associations with the NMR and to explore associations with related phenotypes. We ranked DMET candidate genes and SNPs by their association with the laboratory study NMR using a single model. We validated association of CYP2A6
SNPs with a similar phenotype, the clinical trial-based pretreatment NMR. This suggests that multiple CYP2A6 polymorphisms and non-genetic covariates will improve the predictive power of CYP2A6 activity models and support a continued focus on CYP2A6 activity models for use where phenotypic measures of nicotine metabolites are not available.

Enhanced knowledge of the genes that influence nicotine metabolism, smoking behaviors and clinical outcomes will help to characterize the risks from gene variants on smoking behaviors and smoking-attributable disease, identify novel biomarkers for therapeutic efficacy, and novel targets for the development of smoking cessation therapies. Based on this analysis and additional studies cited, CYP2A6 SNPs account for large fractions of the variance of the NMR, smaller fractions of the variance of cigarette consumption, and influence risk for lung cancer, but do not account for other nicotine dependence factors. This study does not provide guidance on the use of individual DMET SNPs to assign clinical treatment for tobacco dependence as the findings of interest in this study were unreplicated. Research must continue to improve our understanding of factors influencing tobacco product exposures and attributable disease, to develop improved interventions, and to protect the public health.

Supporting Information

S1 Table. PKTWIN and SMOFAM DMET Plus SNP HLM Results.

S2 Table. PKFAM DMET Plus SNP Meta-analysis Results.

S3 Table. Meta-PACT Gene-wise Results.

S4 Table. SNPs chosen for TaqMan SNP Genotyping Assay Genotyping.

S5 Table. TaqMan SNP Genotyping Assay completion rate, by RCT.

S6 Table. TaqMan SNP Genotyping Assay HWE Exact p-value, by RCT.

S7 Table. DMET SNPs and Prospective Abstinence in Eight RCTs.

S8 Table. Power to Detect SNP Odds Ratios, Six Month Abstinence.

S9 Table. rs4803381, rs1137115 and Measures of Nicotine Dependence.

S10 Table. rs4803381 and Abstinence by Pharmacotherapy Randomization.

S11 Table. TRICL rs1137115 Lung Cancer Results.

S1 File. Discovery (Stage I) and Validation (Stage II) Analyses.

S2 File. Prior CYP2A6 Associations with NMR, Abstinence, CPD and Lung Cancer Risk.
Acknowledgments

We thank: the laboratory study and clinical trial participants for their contributions to research; Eric J. Stanek, Jay Kaufman and Richard Hockett for their contributions to Stage I genotyping; Kirsten Copren and Kathryn Thompson at the University of California San Francisco Helen Diller Family Comprehensive Cancer Center Genome Core for their contributions to Stage II genotyping; Gregory Kronmal, Pius Brzoska, and Elizabeth Pitts of ThermoFisher for their custom TaqMan SNP Genotype assay designs; and Heike Bickeböller of the University of Göttingen, Richard Houlston of the Institute of Cancer Research, Maria Teresa Landi and Neil Caporaso of the National Cancer Institute, RayJean Hung of the Lunenfeld-Tanenbaum Research Institute, and William Wheeler of Information Management Services for TRICL meta-GWAS database results.

AWB, MM, RK, DN and HSJ had full access to all of the laboratory study and clinical trial data in the study and take responsibility for the integrity and the accuracy of the data analysis of the laboratory study and clinical trial data. The TRICL Research Team takes responsibility for the integrity and the accuracy of the data analysis of the lung cancer meta-GWAS. The content of this ms is solely the responsibility of the authors and does not represent the official views of Affymetrix, GlaxoSmithKline, Medco Health Solutions, the National Cancer Institute, the National Institute of Drug Abuse, Pfizer, SRI International, ThermoFisher or the University of California Tobacco Related Disease Research Program.

Author Contributions

Conceived and designed the experiments: AWB MM HSJ GES. Performed the experiments: MM RK DN HSJ. Analyzed the data: AWB MM DN RK HSJ KNC. Contributed reagents/materials/analysis tools: HSJ DN RK KNC CNL-S HH JBM SMH TBB DVC NLB CL RFT GES TRICL. Wrote the paper: AWB MM DN RK HSJ KNC CNL-S AZXJ JWB JBM SMH TBB NLB CL RFT GES.

References


