

Mendelian randomization study of spermine oxidase and cancer risk.

João Fadista^{1,2,3*}; Victor Yakimov^{4,7}; Urmo Võsa⁵; Christine S. Hansen^{4,7}; Silva Kasela⁵; Line Skotte¹; Frank Geller¹; Julie Courraud^{6,7}; Tõnu Esko^{5,11}; Viktorija Kukuškina⁵; Alfonso Buil^{7,8}; Mads Melbye^{9,10}; Thomas M. Werge^{7,8}; David M. Hougaard^{4,7}; Lili Milani⁵; Jonas Bybjerg-Grauholm^{4,7}; Arieh S. Cohen^{6,7}; Bjarke Feenstra¹

1: Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark

2: Department of Clinical Sciences, Lund University Diabetes Centre, Malmö, Sweden

3: Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

4: Section of Neonatal Genetics, Danish Centre for Neonatal Screening, Department of Congenital Diseases, Statens Serum Institut, Copenhagen, Denmark

5: Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia

6: Section of Clinical Mass Spectrometry, Danish Center for Neonatal Screening, Statens Serum Institut, Copenhagen, Denmark

7: iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Aarhus, Denmark

8: Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Roskilde, Denmark

9: Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

10: Department of Medicine, Stanford University School of Medicine, Stanford, California, USA

11: Program in Medical and Population Genetics, Broad Institute, Cambridge MA, USA

*Address correspondence to: João Fadista, PhD (jpsf@ssi.dk). Department of Epidemiology Research, Statens Serum Institut, Artillerivej 5, 2300 Copenhagen, Denmark. Phone: +45 32688153.

Word count of the manuscript text: 3033

ABSTRACT

Importance: Spermine oxidase (SMOX) catalyzes the oxidation of spermine to spermidine. Observational studies have reported SMOX as a source of induced reactive oxygen species associated with cancer, implying that inhibition of SMOX could be a target for chemoprevention.

Objective: To test the causality of SMOX levels with cancer risk using a Mendelian randomization analysis.

Design, Setting, and Participants: We performed a GWAS of spermidine/spermine ratio, from blood of 534 infants, to identify genetic variants associated with regulation of SMOX activity. In two additional data sets of 262 newborns and 508 adults, we quantified *SMOX* gene expression using RNA-sequencing and performed expression and methylation QTL lookups. We then did a Mendelian randomization analysis by testing the association between the SMOX genetic instrument and various cancer types using GWAS summary statistics.

Main Outcomes and Measures: Neuroblastoma, gastric, lung, breast, prostate, and colorectal cancers.

Results: The GWAS of spermidine/spermine ratio identified a genome-wide significant locus ($P=1.34 \times 10^{-49}$) explaining 32% of the variance. The lead SNP rs1741315 was also associated with *SMOX* gene expression in newborns ($P=8.48 \times 10^{-28}$) and adults ($P=2.748 \times 10^{-8}$) explaining 37% and 6% of the variance, respectively. rs1741315 was not associated with neuroblastoma (OR=0.95; 95% CI:0.88, 1.03; $P=0.18$), gastric (OR=0.99; 95% CI:0.95, 1.03; $P=0.54$), lung (OR=0.97; 95% CI:0.94, 1.00; $P=0.08$), breast (OR=0.99; 95% CI:0.96, 1.02; $P=0.47$), prostate (OR=0.98; 95% CI:0.96, 1.00; $P=0.05$) nor colorectal cancer (OR=1.03; 95% CI:0.99, 1.07; $P=0.10$). A PheWAS of rs1741315 did not reveal any associations with risk factors of the cancers tested.

Conclusions and Relevance: Genetic variation in the *SMOX* gene was strongly associated with SMOX activity in newborns, and less strongly in adults. Genetic down-regulation of SMOX was not significantly associated with lower odds of neuroblastoma, gastric, lung, breast, prostate and colorectal cancer. Further studies are needed to understand the effect of SMOX inhibition in relation to cancer risk.

ARTICLE SUMMARY

Question: Is SMOX causally associated with risk of cancer?

Findings: In this Mendelian randomization study, genetically lower levels of SMOX were not associated with decreased risk of neuroblastoma, gastric, lung, breast, prostate and colorectal cancer.

Meaning: These findings do not support a causal association between SMOX activity and risk of cancer, suggesting that ongoing efforts to identify SMOX inhibitors for chemoprevention may not be successful.

STRENGTHS AND LIMITATIONS

- Previous studies which examined SMOX activity and cancer risk were susceptible to recall bias, confounding and reverse causation, none of which are concerns of this Mendelian randomization study.
- Our genetic instrument explained a sizeable proportion of the variance of SMOX activity
- We used summary statistics from the largest meta-analyses of primary cancer GWAS to date.
- Elevated SMOX levels in cancer could also be due to environmental factors not captured by genetics.
- Our genetic instrument was developed based on normal range SMOX activity data, thus additional genetic variants might play a role in aberrant expression of this enzyme.

INTRODUCTION

Although polyamines are essential for normal cell growth and development^{1,2}, dysregulation of polyamine metabolism is involved in tumorigenesis³, and hence recognized as a potential target for chemotherapy and chemoprevention⁴. Already in the 1960s, ornithine decarboxylase (ODC), the first rate-limiting enzyme in polyamine biosynthesis, was demonstrated to be at high levels in human cancer specimens⁵. Moreover, a genetic variant in ODC, regulating its enzymatic activity, confirmed the role of this enzyme in human colon cancer risk^{6,40}, and ODC levels have been shown to be elevated in human skin, breast and prostate cancer specimens⁷⁻⁹. Other polyamine metabolic enzymes have not yet been reported to be associated with tumorigenesis in humans.

Spermine oxidase (SMOX), a member of the mammalian polyamine catabolic pathway, encodes an enzyme with cytoplasmic and nuclear expression in most tissues⁴², that catalyzes the oxidation of spermine to spermidine with the production of hydrogen peroxide (H₂O₂) and 3-aminopropanal (3-AP)¹⁰⁻¹¹. SMOX has been reported as a source of induced reactive oxygen species (ROS) associated with neuroblastoma, gastric, lung, breast, prostate and colon cancers^{12-16,36,37}, implying that inhibition of SMOX could be a target for chemoprevention³. However, so far, studies of SMOX inhibition have been observational in their nature with no direct inference on causation^{12-16,36,37}. Genome-wide association studies (GWAS) of SMOX activity are also lacking. A GWAS could potentially detect genetic determinants of SMOX activity that might serve as instruments to assess the causality of SMOX on cancer using a Mendelian randomization framework¹⁷.

To our knowledge, there has not been any Mendelian randomization study assessing the potential causal relationship between SMOX activity and risk of any type of cancer. In this study, we present the first GWAS of spermidine/spermine ratio as a proxy for SMOX activity. We then use the genetic determinants of SMOX activity to perform Mendelian randomization analysis to assess causal relationships between SMOX activity and risk of lung, breast, prostate and colon cancer. In addition, we perform a phenome-wide association study (PheWAS) to query possible deleterious effect of altered SMOX activity through all disease categories and to test possible pleiotropic effects.

METHODS

Study population

The GWAS of spermidine/spermine ratio reflecting SMOX activity, was based on dried blood spot samples taken during routine newborn screening from 534 individuals of Danish ancestry, recruited in a matched case-control study of infantile hypertrophic pyloric stenosis (IHPS), as previously described^{22,23} (SSI-IHPS cohort). In addition, dried blood spot samples of 262 newborns of Danish ancestry were included for expression quantitative trait locus (eQTL) analysis (PSYCH-twin cohort). These 262 newborns were taken at random from each of 262 monozygotic twin pairs discordant for later diagnosis of psychiatric disorders as previously described³⁴. Of these 262 newborns, 235 were included for methylation quantitative trait loci (mQTL) analysis. Furthermore, 671 whole blood samples taken from adults of the Estonian Genome Center University of Tartu (EGCUT) cohort⁴⁶ were also used for eQTL (N=508)^{35,47} and mQTL (N=305; overlap with gene expression dataset 141 samples) lookups. The Danish Scientific Ethics Committee, the Danish Data Protection Agency and the Danish Neonatal Screening Biobank Steering Committee approved the PSYCH-twins study as well as the GWAS of metabolites in newborns. Usage of EGCUT RNA-seq dataset was approved by Estonian Committee on Bioethics and Human Research, protocol nr. 1.1-12/624, data extraction nr. N26.

The study also included GWAS summary statistics from four cancer-specific consortia cohorts: The ELLIPSE consortium (prostate cancer)¹⁸ with 79,194 cases and 61,112 controls; the BCAC consortium (breast cancer)¹⁹ with 62,533 cases and 60,976 controls; the TRICL consortium (lung cancer)²⁰ with 29,266 cases and 56,450 controls, and the North American-based Children's Oncology Group (neuroblastoma)³⁸ with 2,101 cases and 4,202 controls. GWAS summary statistics for gastric and colorectal cancers relied on population-based cohorts: UK Biobank with 5,693 colorectal cancer cases and 386,740 controls²¹, and the BioBank Japan Project with 6,563 gastric cancer cases and 195,745 controls³³. Analysis were restricted to individuals of European ancestry, except for gastric cancer due to the lack of publicly available European gastric cancer GWAS data.

Patients or the public WERE NOT involved in the design, or conduct, or reporting, or dissemination plans of our research.

Metabolite measurements and GWAS

The metabolites spermidine and spermine were quantified in whole blood of dried blood spots using the AbsoluteIDQ® p400 Kit (Biocrates Life Sciences AG, Innsbruck, Austria). The AbsoluteIDQ® p400 Kit allows quantification of 408 metabolites which were described in our previous study²³. In our previous study²³, we quantified 148 of the 408 metabolites, from which spermidine and spermine measurements were also used for this study. In this study we performed genome-wide association scans of all these 148 metabolites and their biologically relevant ratios (e.g. direct substrate and product of an enzyme), including spermidine to spermine ratio, which had the most significant genome-wide hit of all the metabolites and ratios tested. For accuracy and computing time issues, the GWAS analysis was first implemented in PLINK⁴¹ and then repeated in R²⁴ only for the genome-wide significant variants. The PLINK⁴¹ and R²⁴ analyses were done using the following linear model: spermidine/spermine inverse normally transformed concentration ratios ~ SNP + sex + YOB (as factor) + GA + Parity + IHPS. YOB pertains to year of birth, GA as gestational age in weeks, and IHPS as case/control status. As described before²², the SSI-IHPS cohort was array-genotyped with the Illumina Multi-Ethnic Global_v2_A2 array, and after genotyping QC, unobserved genotypes were imputed from the Haplotype Reference Consortium panel⁴³. Altogether, 6,846,507 SNPs with minor allele frequency>=1% and imputation info score>=0.8 were used in the analysis.

eQTL and mQTL lookups

To replicate the genetic association with *SMOX* activity at the gene expression level, eQTL analysis was performed in whole blood from two additional data sets of 262 newborns (a subset of the PSYCH-twin cohort) and 508 adults (a subset of the EGCUT cohort) that had previously been RNA-sequenced and GWAS-genotyped^{34,35,47}. Preprocessing of the EGCUT expression data is described elsewhere³⁵. Briefly, in EGCUT, gene expression matrix was normalized by trimmed mean of M values (TMM)⁴⁸, log₂ transformed, gene expression values were centered and scaled, and genotype multidimensional scaling (MDS) was regressed out. In addition, in order to remove non-genetic variance from the data, up to 20 gene expression-based PCs which were not associated with genetics were regressed out from the residuals of the expression data. Finally, inverse normal transformation was applied on the residuals from the previous step. Moreover, 235 out of 262 individuals from the newborn PSYCH-twin cohort³⁴ also had genome-wide methylation array data from the Infinium HumanMethylation450 BeadChip (450k) array, as previously described³⁴. For this study, all samples were filtered by Call Rate > 0.99 (for detection P < 0.01 and bead count > 2) and median methylation and unmethylation intensity signal > 2000 to ensure proper sample measurement quality. All probes (CpGs) were filtered by Call Frequency > 0.99 (for detection P < 0.01 and bead count > 2) to ensure sufficient probe representation across most samples. The data was normalized with the noob background dye correction method after sample filtering³⁹, implemented in the R²⁴ package minfi. For all the 480 CpG sites within 1 Mb of the *SMOX* gene, we tested their methylation fraction against the lead (most significant) *SMOX* SNP. The analysis was implemented in R²⁴ as the following linear model: CpG methylation fraction ~ lead *SMOX* SNP + sex + YOB + GA + BW (birth weight). An association was deemed significant if P-value < 0.05/480 = 1.04 × 10⁻⁴ to correct for the number of CpGs tested. Furthermore, 305 individuals from the adult EGCUT cohort also had genome-wide methylation array data from the Infinium HumanMethylation450 BeadChip (450k) array. Briefly, methylation data was normalized according to the GoDMC (<http://www.godmc.org.uk/>) pipeline⁴⁴ using functional normalization method as implemented in the R package meffil⁴⁵. In step 1, we adjusted normalized methylation proportion for age, sex, predicted cell counts, predicted smoking and genetic PCs. In step 2, PCA was applied on the residuals from the previous step, and top non-genetic PCs were regressed out. To be able to estimate the SNP × Age effect on methylation, inverse normally transformed methylation proportions were regressed against covariates from step 1 and methylation PCs from step 2.

Variance explained, Mendelian randomization, and PheWAS analyses

The lead SNP located on chromosome 20q12 (in cis with *SMOX*) being genome-wide significantly ($P < 5 \times 10^{-8}$) associated with spermidine/spermine ratio was used as the genetic instrument in the Mendelian randomization analyses. Variance explained by the genetic determinant of *SMOX* activity was calculated as the adjusted R² of lm function in R²⁴ implemented as spermidine/spermine inverse normally transformed concentrations ratio ~ SNP, or normalized and inverse normally transformed gene expression value of *SMOX* ~ SNP. The inverse-variance weighted method²⁵, implemented in the R²⁴ package MendelianRandomisation (v0.3.0) was used to test whether the association between *SMOX* genotype and incident cancers could be caused by differences in *SMOX* activity due to genotype. The primary analysis measured the association between genetically determined *SMOX* activity and the risk of neuroblastoma, gastric, lung, breast, prostate and colorectal cancers. The PhenoScanner bioinformatic tool²⁷, a curated database of publicly available results from large-scale genetic association studies, was used for the PheWAS analysis and detect possible concurrent risk factors of the investigated cancer types (pleiotropy). The PhenoScanner database was queried from the MendelianRandomisation R²³ package. In our main analyses we tested six key outcomes, namely neuroblastoma, gastric, lung, breast, prostate and colorectal cancers. A P-value of less than 0.05/6 (number of cancers tested) = 0.008 was considered statistically significant.

RESULTS

Participant characteristics

A total of 534 newborn participants (SSI-IHPS cohort) were included in the GWAS of spermidine/spermine ratio to find the genetic instrument of SMOX activity (Methods). An additional 262 newborns (PSYCH-twin cohort) and 508 adults (EGCUT cohort) were included to confirm the genetic instrument of SMOX activity at the gene expression level (Table 1 and Methods). A total of 950,575 participants, including 79,194 participants who had pancreatic cancer, 62,533 with breast cancer, 29,266 with lung cancer, 5,693 with colorectal cancer, 6,563 with gastric cancer and 2,101 with neuroblastoma were included as the cancer outcome cohorts (Table 2 and Methods). Demographic characteristics of each cohort are described in the original studies^{18-23,33-35}.

Genetic associations reflecting SMOX activity

We performed a GWAS of spermidine/spermine ratio, with two loci reaching genome-wide significance ($P < 5 \times 10^{-8}$) (Figure 1, Methods). One locus was located on chromosome 11p13, having four intergenic SNPs in perfect LD ($r^2=1$) with each other, namely rs3891324, rs7124637, rs2067666 and rs11031804 ($P=1.50 \times 10^{-8}$). These SNPs explained 5% of the variance in the spermidine/spermine ratio. The other locus was on chromosome 20q12, with the most significant SNP rs1741315 ($P=1.34 \times 10^{-49}$) being an intronic variant in the *SMOX* gene (Figure 2). This lead SNP explained 32% of the variance in the spermidine/spermine ratio (Figure 3, Table 1). Despite that the locus on chromosome 11p13 reaches genome-wide significance, it is located in trans regarding *SMOX*. Since trans-acting SNPs can often be pleiotropic (violating one of the principles of Mendelian randomization)²⁶, and since their variance explained is relatively small (5%) compared to the top cis-acting variant rs1741315 (32%), we chose to use only the latter as a genetic instrument of SMOX for the Mendelian randomization analysis. To assess effects of rs1741315 on SMOX activity at the gene expression level, we performed eQTL analyses in two additional cohorts. In a newborn cohort of 262 individuals (PSYCH-twin)³⁴ rs1741315 was associated with *SMOX* gene expression ($P=2.47 \times 10^{-30}$) explaining 37% of the variance (Table 1). In a cohort of 508 adults (EGCUT)³⁵, rs1741315 was also associated with *SMOX* gene expression (eQTL $P=2.75 \times 10^{-8}$) explaining 6% of the variance (Table 1, eFigure 1). Of note, rs1741315 was array-genotyped in all cohorts, in contrast to being just imputed. We stratified the adult cohort (EGCUT) into different age groups (10-years intervals) in order to detect a possible age-dependent eQTL effect in early vs. late adulthood (eFigure 2), but found no significant interaction between donor age at blood draw and rs1741315 genotype ($P=0.50$) on *SMOX* expression (eFigure 3). Due to some missing values in the data, this age-stratified analysis was performed on 494 out of the total 508 adults. Moreover, 235 out of 262 individuals from the newborn PSYCH-twin cohort³⁴ had genome-wide methylation array data (Methods). Zooming in on CpGs at the *SMOX* locus, we found rs1741315 to be associated with the methylation proportion of the CpG cg07472708 (mQTL $P=4.63 \times 10^{-5}$) explaining 0.5% of the variance in methylation proportion for this CpG (Table 1). Contrary to the newborn data, we did not find rs1741315 to be associated with the methylation proportion of the CpG cg07472708 (mQTL $P=0.371$) in 305 individuals from the adult EGCUT cohort who also had genome-wide methylation array data (eFigure 4, Table 1 and Methods). Similar to the eQTL data, we also stratified the adult cohort (EGCUT) into different age groups (10-years intervals) in order to detect a possible age-dependent mQTL effect (eFigure 5), but found no significant interaction between donor age at blood draw and rs1741315 genotype ($P=0.48$) on cg07472708 methylation levels (eFigure 6). The CpG cg07472708 is located in the same *SMOX* intron as rs1741315 and is under a region of transcription factor binding site for the transcription factor GATA-1 (Methods). It should be noted that the 50-mer probe containing cg07472708 overlaps with rs1741317 (37 bp apart from the C nucleotide of cg07472708), a high LD ($r^2=0.996$) SNP with our lead rs1741315. This could lead to a technical artifact from poorer measurement conditions in those individuals carrying the minor allele, so the results from mQTL analysis should be interpreted with caution.

PheWAS association of SMOX genetic instrument

One of the assumptions of Mendelian randomization is that genetic variants should only affect the outcome through their effect on the risk factors, i.e. no pleiotropic effects²⁹⁻³⁰. To test if this assumption holds, we checked the association of the SMOX genetic instrument rs1741315 against the PhenoScanner database of genotype-phenotype associations²⁷ at $P < 1 \times 10^{-5}$ (to correct for the 2998 traits/diseases tested in the database). With rs1741315 'A' allele being the effect allele, we detected reticulocyte count ($\beta = -0.150$; 95% CI: -0.157, -0.143; $P = 4.12 \times 10^{-285}$), lymphocyte count ($\beta = 0.016$; 95% CI: 0.009, 0.023; $P = 6.44 \times 10^{-6}$) and hemoglobin concentration ($\beta = 0.016$; 95% CI: 0.009, 0.023; $P = 8.05 \times 10^{-6}$) as traits associated with this variant³¹. One potential way to test and adjust for pleiotropic effects through associations with these blood traits would be by multivariable Mendelian randomization analysis²⁶. However, such analysis is only possible for multi-SNP genetic instruments, which we did not have.

Association of SMOX genetic instrument with cancer risk

We found little evidence that the genetic instrument of SMOX activity, rs1741315, was associated with any of the five non-pediatric cancers evaluated (Table 2). Although genetically lower levels of SMOX were associated with slightly lower risk of prostate cancer at $P = 0.047$, this finding should be interpreted with caution as it becomes non-significant when correcting for the six different cancers tested.

Since the SMOX genetic instrument, rs1741315, explained 5 to 6 times more variance of SMOX activity in newborns than in adults, we also tested the association between rs1741315 and neuroblastoma, a pediatric cancer in which SMOX has been reported to play a role^{36,37}. Genetically lower levels of SMOX did not associate with lower risk of developing neuroblastoma³⁸ (OR=0.95; 95% CI: 0.88, 1.03; $P = 0.182$) (Table 2).

DISCUSSION

We conducted the first GWAS of spermidine/spermine ratio to identify genetic variants regulating spermine oxidase (SMOX) activity. Variants in the *SMOX* gene explained a large proportion of the variance in this ratio in newborns and were eQTLs for *SMOX* expression in newborns as well as adults. We did not find genetically determined SMOX activity to be associated with risk of either pediatric (neuroblastoma) or adult cancers (gastric, lung, breast, prostate, and colorectal cancer).

Observational studies have reported elevated SMOX levels in gastric, lung, breast, prostate, and colorectal cancers¹²⁻¹⁶, and it has been hypothesized that SMOX inhibition could be an effective target for chemoprevention³. Our results do not indicate that genetically driven variation in SMOX activity plays a major role in cancer risk. Several factors might explain this discrepancy between the observational epidemiological evidence¹²⁻¹⁶ and our Mendelian randomization results. Our results indicate that the genetic instrumental variable more accurately reflects SMOX activity in infancy, compared with later life, where most cancers occur. The positive association between observed SMOX levels and cancer in observational studies might be driven by reverse causation, in which cancer in an individual could cause SMOX to be elevated due to inflammation. In addition, elevated SMOX levels in cancer could also be due to environmental factors not captured by genetics, e.g. short-term pharmacological changes, induced by SMOX inhibitors. Our genetic instrument was developed based on normal range SMOX activity data, thus additional genetic variants might play a role in aberrant expression of this enzyme. Furthermore, despite we have corrected for disease status, half the samples from the discovery SSI-IHPS cohort were selected to be IHPS cases, and are therefore not representative of the general population. However, the association of rs1741315 with spermidine/spermine ratio did not differ between IHPS cases and controls (eFigure 7). The instrumental variable SNP is also associated with hematological traits³¹ that could potentially affect cancer risk. To assess whether pleiotropic effects drove our results, we performed a PheWAS analysis, with no diseases or non-hematological traits detected. We detected reticulocyte count, lymphocyte count and hemoglobin concentration associated with rs1741315. However, it is unclear if and how the variability in those blood cell phenotypes could affect the cancer predisposition. Therefore there was no reason to conclude that pleiotropy would affect our Mendelian Randomization analyses.

Our study had important strengths. Several previous studies which examined SMOX activity and cancer risk were susceptible to recall bias, confounding and reverse causation¹¹⁻¹⁷ none of which are concerns of Mendelian randomization studies¹⁷. In addition, the fact that our genetic instrument explained a sizeable proportion of the variance of SMOX activity together with the fact that we used summary statistics from the largest meta-analyses of primary GWAS of these cancer types to date^{18-21,33,38} are factors enhancing the statistical power to detect causal effects³².

In conclusion, lifelong genetic exposure to low levels of SMOX was not associated with lower cancer risk. However, this finding cannot be assumed to represent a short-term strong drug inhibition of SMOX, thus establishing causality would require a randomized clinical trial. To better interpret the complexity of the relationship between SMOX activity and cancer, future studies should effectively distinguish newborn versus adulthood SMOX activity and short versus long exposures. This study is the first to use Mendelian randomization to assess the possible benefits of SMOX inhibitors on cancer risk, paving the way for other polyamine catabolic pathway enzymes to be tested with the same methodology.

ACKNOWLEDGMENTS

This study received partial support from the Danish Medical Research council (DFF 4004-00512), the Novo Nordisk Foundation (NNF18OC0053228) and the Oak Foundation (OCA-18-598). It was conducted using the Danish National Biobank resource, supported by the Novo Nordisk Foundation, grant number 2010-11-12 and 2009-07-28. We would also like to acknowledge iPSYCH and The Lundbeck Foundation for their QTL validation datasets. The iPSYCH (The Lundbeck Foundation Initiative for Integrative Psychiatric Research) team acknowledges funding from The Lundbeck Foundation (grant numbers R102-A9118 and R155-2014-1724). This research has also been conducted using the UK Biobank Resource (project ID 31063) available from the Pan-UKB team. This work has also been supported by the European Regional Development Fund and the programme Mobilitas Plus (MOBTP108), SP1G18045T, No. 2014-2020.4.01.15-0012 GENTRANSMED and 2014-2020.4.01.16-0125. This study was also funded by EU H2020 grant 692145, Estonian Research Council Grant PUT1660 and PRG687. Data analyzes with Estonian datasets were carried out in part in the High-Performance Computing Center of University of Tartu. Furthermore, we acknowledge the Genetic Associations and Mechanisms in Oncology (GAME-ON)/Elucidating Loci Involved in Prostate Cancer Susceptibility (ELLIPSE) Consortium, the Transdisciplinary Research Into Cancer of the Lung (TRICL) consortium, the Breast Cancer Association Consortium (BCAC), and the North American-based Children's Oncology Group. We further thank all the participants, staff and researchers of the Danish National Biobank, Estonian, Japan and UK biobanks, along with the cancer case-control cohorts above described for their contribution to this research.

REFERENCES

1. Pegg AE. Polyamine metabolism and its importance in neoplastic growth and a target for chemotherapy. *Cancer research*. 1988; 48(4):759–74.
2. Gerner EW, Meyskens FL Jr. Polyamines and cancer: old molecules, new understanding. *Nature reviews. Cancer*. 2004; 4(10):781–92
3. Casero RA Jr, Murray Stewart T, Pegg AE. Polyamine metabolism and cancer: treatments, challenges and opportunities. *Nat Rev Cancer*. 2018 Nov;18(11):681–695.
4. Casero RA Jr, Marton LJ. Targeting polyamine metabolism and function in cancer and other hyperproliferative diseases. *Nature reviews. Drug discovery*. 2007; 6(5):373–90.
5. Russell D, Snyder SH. Amine synthesis in rapidly growing tissues: ornithine decarboxylase activity in regenerating rat liver, chick embryo, and various tumors. *Proceedings of the National Academy of Sciences of the United States of America*. 1968; 60(4):1420–7.
6. Martinez ME, et al. Pronounced reduction in adenoma recurrence associated with aspirin use and a polymorphism in the ornithine decarboxylase gene. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100(13):7859–64.
7. Gupta S, et al. Chemoprevention of prostate carcinogenesis by alpha- difluoromethylornithine in TRAMP mice. *Cancer research*. 2000; 60(18):5125–33.
8. Elmetts CA, Athar M. Targeting Ornithine Decarboxylase for the Prevention of Nonmelanoma Skin Cancer in Humans. *Cancer Prev Res (Phila Pa)*. 2010; 3(1):8–11.
9. Manni A, et al. Prognostic influence on survival of increased ornithine decarboxylase activity in human breast cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 1996; 2(11):1901–6.
10. Wang Y, Devereux W, Woster PM, Stewart TM, Hacker A, Casero RA Jr. Cloning and characterization of a human polyamine oxidase that is inducible by polyamine analogue exposure. *Cancer Res*. 2001; 61(14):5370–5373.
11. Vujcic S, Diegelman P, Bacchi CJ, Kramer DL, Porter CW. Identification and characterization of a novel flavin-containing spermine oxidase of mammalian cell origin. *Biochem J*. 2002; 367(Pt 3): 665–675.
12. Babbar N & Casero RA Jr. Tumor necrosis factor- α increases reactive oxygen species by inducing spermine oxidase in human lung epithelial cells: a potential mechanism for inflammation-induced carcinogenesis. 2006; 66(23):11125–11130.
13. Goodwin AC et al. Increased spermine oxidase expression in human prostate cancer and prostatic intraepithelial neoplasia tissues. *Prostate* 68, 766–772 (2008).
14. Chaturvedi R et al. Spermine oxidase mediates the gastric cancer risk associated with *Helicobacter pylori* CagA. *Gastroenterology* 141, 1696–1708 (2011).
15. Goodwin AC et al. Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl Acad. Sci. USA* 108, 15354–15359 (2011).
16. Pledge A et al. Spermine oxidase SMO(PAOh1), not N1-acetylpolyamine oxidase PAO, is the primary source of cytotoxic H₂O₂ in polyamine analogue-treated human breast cancer cell lines. *J. Biol. Chem* 280, 39843–39851 (2005).
17. Ebrahim, S. & Davey Smith, G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum. Genet.* 123, 15–33 (2008).
18. Schumacher FR et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet.* 2018 Jul;50(7):928–936. doi: 10.1038/s41588-018-0142-8.
19. Michailidou K et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet.* 2015 Apr;47(4):373–80. doi: 10.1038/ng.3242.
20. McKay JD et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet.* 2017 Jul;49(7):1126–1132. doi: 10.1038/ng.3892.
21. Pan-UKB team. <https://pan.ukbb.broadinstitute.org>. 2020.
22. Fadista J et al. Genome-wide meta-analysis identifies BARX1 and EML4-MTA3 as new loci associated with infantile hypertrophic pyloric stenosis. *Hum Mol Genet.* 2019 Jan 15;28(2):332–340. doi: 10.1093/hmg/ddy347.
23. Fadista J et al. Integrating genetics with newborn metabolomics in infantile hypertrophic pyloric stenosis. medRxiv. 2020.02.22.20026724.
24. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> (2020).
25. Stephen Burgess, Adam S Butterworth, Simon G Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic Epidemiology* 2013; 37:658–665. doi: 10.1002/gepi.21758.
26. Stephen Burgess, Simon G Thompson. Multivariable Mendelian Randomization: the use of pleiotropic genetic variants to estimate causal effects. *American Journal of Epidemiology* 2015; 181(4):251–260. doi: 10.1093/aje/kwu283.
27. Kamat MA et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics*. 2019 Nov 1;35(22):4851–4853. doi: 10.1093/bioinformatics/btz469.
28. Vösa U et al. Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. bioRxiv. October 2018.
29. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. *Epidemiology* 2014;25:427–35.
30. Thomas DC, Conti DV. Commentary: The concept of ‘Mendelian randomization’. *Int J Epidemiol* 2004;33:21–25.
31. Astle WJ et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*. 2016 Nov 17;167(5):1415–1429.e19. doi: 10.1016/j.cell.2016.10.042.

32. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol* 2014;43:922–29.
33. Ishigaki K, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nature Genetics*. 2020; 52(7):669-679.
34. Hansen CS, et al. Elevated DNA Methylation Gestational Age is associated with the Risk of Later Bipolar Disorder and Anorexia Nervosa in Twins. *medRxiv* 2020.07.16.20155010
35. Vösa U, et al. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. 2018. *bioRxiv* 447367.
36. Amendola R, et al. Direct oxidative DNA damage, apoptosis and radio sensitivity by spermine oxidase activities in mouse neuroblastoma cells. *Biochimica et Biophysica Acta*. 2005;1755(1):15-24.
37. Fratini E, et al. Link between spermine oxidase and apoptosis antagonizing transcription factor: A new pathway in neuroblastoma. *Int J Oncol*. 2019;55(5):1149-1156.
38. McDaniel LD, et al. Common variants upstream of MLF1 at 3q25 and within CPZ at 4p16 associated with neuroblastoma. *PLoS Genet*. 2017;13(5):e1006787.
39. Triche TJ Jr, et al. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res*. 2013;41(7):e90.
40. Hubner RA, et al. Ornithine decarboxylase G316A genotype is prognostic for colorectal adenoma recurrence and predicts efficacy of aspirin chemoprevention. *Clin Cancer Res*. 2008 Apr 15; 14(8):2303-9.
41. Chang CC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015 Feb 25;4:7.
42. Uhlén M, et al. Tissue-based map of the human proteome. *Science*. 2015; 347(6220):1260419.
43. McCarthy S, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48,1279–1283.
44. Min JL, et al. Genomic and phenomic insights from an atlas of genetic effects on DNA methylation. *medRxiv* 2020.09.01.20180406.
45. Min JL et al. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics*. 2018. Dec 1;34(23):3983-3989.
46. Leitsalu L et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int J Epidemiol*. 2015 Aug;44(4):1137-47.
47. Lepik K, et al. C-reactive protein upregulates the whole blood expression of CD59 - an integrative analysis. *PLoS Comput Biol*. 2017 Sep 18;13(9):e1005766.
48. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25.

Table 1. SNP rs1741315 associations with SMOX activity, *SMOX* gene expression and cg07472708 methylation fraction. SMOX activity and *SMOX* gene expression estimates were based on inverse normally transformed values (Methods). SNP rs1741315 'A' allele is the effect allele. Pos, genomic position of rs1741315 (hg19); rsID, variant identifier; EAF, effect allele frequency; N, sample size; Beta, effect estimate; SE, standard error; R², variance explained of rs1741315 on SMOX activity, *SMOX* gene expression and cg07472708 methylation fraction.

Pos	rsID	Cohort	Analysis	N	EAF	Beta	SE	P Value	R ²
20:4155948	rs1741315	Newborns (SSI-IHPS)	SMOX activity	534	0.458	-0.757	0.046	1.34x10 ⁻⁴⁹	0.318
		Newborns (PSYCH-twins)	<i>SMOX</i> expr.	262	0.471	-0.830	0.062	2.47x10 ⁻³⁰	0.366
			cg07472708	235		-0.012	0.003	4.63x10 ⁻⁰⁵	0.005
		Adults (EGCUT)	<i>SMOX</i> expr.	508	0.452	-0.366	0.065	2.75x10 ⁻⁰⁸	0.057
cg07472708	305		-0.062	0.069		0.371	0.002		

Table 2. Associations between genetically predicted SMOX activity, based on SNP rs1741315, and six site-specific cancers (prostate, breast, lung, colorectal and neuroblastoma). All estimations were based on the inverse-variance weighted method with fixed-effects, with A being the effect allele. EAF, effect allele frequency.

Outcome	Cases, No.	Controls, No.	EAF	Odds Ratio (95% CI)	P Value
Prostate cancer	79,194	61,112	0.430	0.98 (0.96-1.00)	0.046
Breast cancer	62,533	60,976	0.434	0.99 (0.96-1.02)	0.465
Lung cancer	29,266	56,450	0.442	0.97 (0.94-1.00)	0.084
Colorectal cancer	5,693	386,740	0.448	1.03 (0.99-1.07)	0.096
Gastric cancer	6,563	195,745	0.361	0.99 (0.95-1.03)	0.543
Neuroblastoma	2,101	4,202	0.412	0.95 (0.88-1.03)	0.182

Figure 1. Genetic determinants of SMOX activity. Manhattan plot of the genome-wide association study (GWAS) of spermidine/spermine ratio (534 samples; SSI-IHPS cohort).

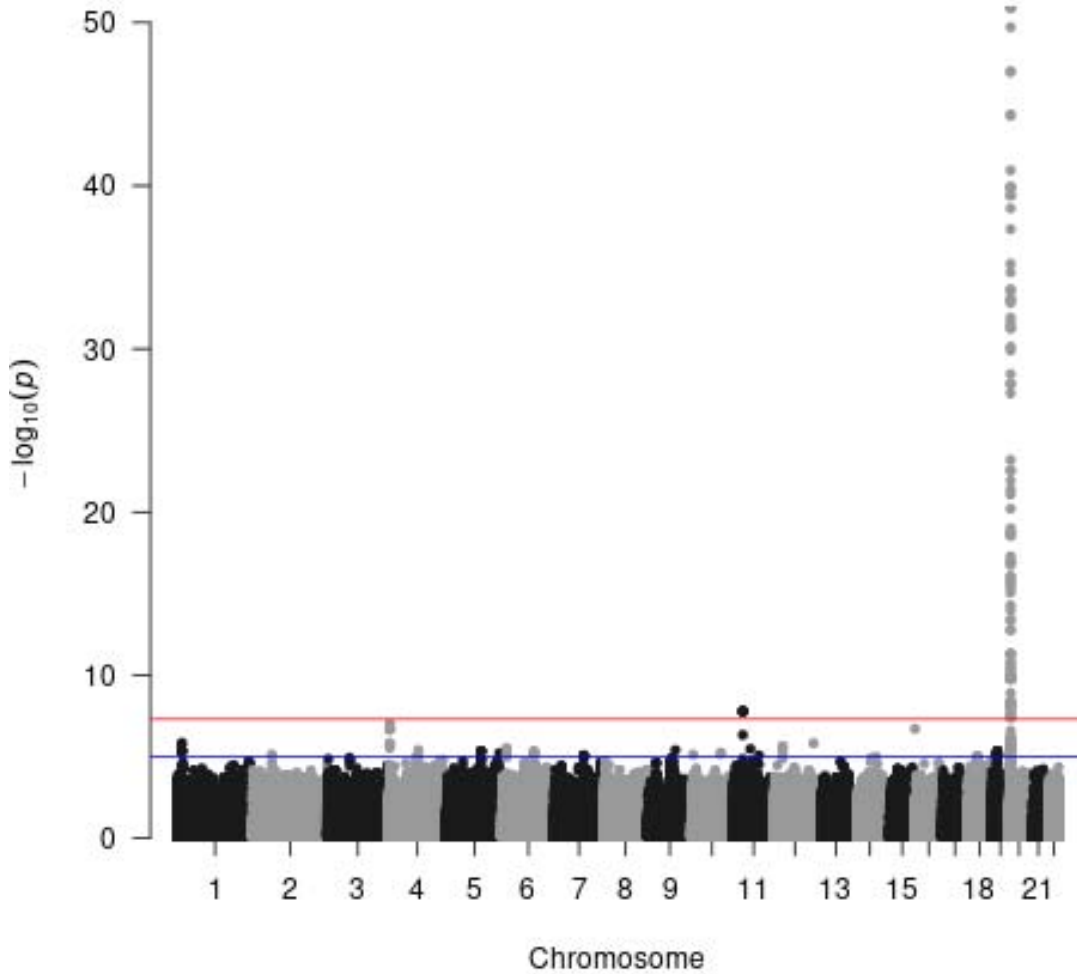


Figure 2. Regional association plot of the new genome-wide significant loci for spermidine/spermine ratio at the SMOX locus. Color-coded linkage disequilibrium (LD) is shown for the top SNP rs1741315 (LD determined with 1000Genomes EUR population). The x-axis represents the genomic region (hg19 assembly) surrounding 200 kb of SMOX gene, while the y-axis represents the strength of the association in $-\log_{10}(P\text{-value})$.

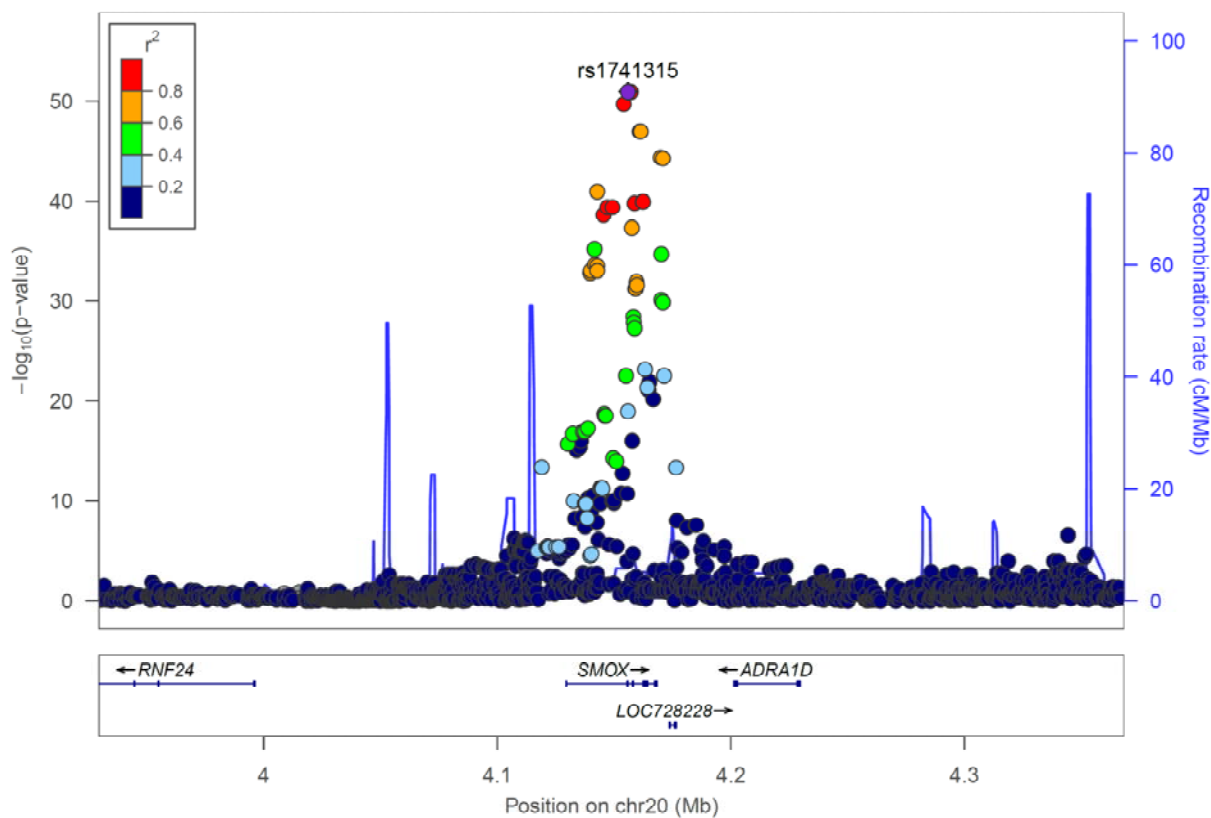


Figure 3. Boxplot of the rs1741315 alleles versus spermidine/spermine ratio (534 samples; SSI-IHPS cohort).

